2014

On Initializing CGCMs for Seasonal Predictability of ENSO

J-P Michael

FLORIDA STATE UNIVERSITY

COLLEGE OF ARTS AND SCIENCES

ON INITIALIZING CGCMS FOR SEASONAL PREDICTABILITY OF ENSO

By

J-P MICHAEL

A Dissertation submitted to the
Department of Earth, Ocean and Atmospheric Science
in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Degree Awarded:
Summer Semester, 2014

J-P Michael defended this dissertation on April 1 2014.

The members of the supervisory committee were:

Vasu Misra
Professor Directing Dissertation

Mike Burmester
University Representative

Eric P. Chassignet
Committee Member

Zhaohua Wu
Committee Member

Philip Sura
Committee Member

The Graduate School has verified and approved the above-named committee members, and certifies that the dissertation has been approved in accordance with university requirements.

ACKNOWLEDGMENTS

A special thanks to Dr. Robert van Engelen whose contributions as university representative for a majority of the lifespan of this project were greatly appreciated and to Dr. Mike Burmester who graciously presided as university representative for dissertation defense proceedings.

We acknowledge the World Climate Research Programme's Working Group on Coupled Modeling, which is responsible for CMIP, and the climate modeling groups for producing and making available their model output. This work was supported by grants from NOAA (NA12OAR4310078, NA10OAR4310215, NA11OAR4310110), USGS (06HQGR0125), and USDA (027865).

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

Initializing Coupled General Circulation Models (CGCMs) for routine seasonal ENSO prediction is currently an onerous task. This is one of the main reasons on why the CGCMs participating in the Coupled Model Intercomparison Project 5 (CMIP5), which represents the state-of-the-art in climate modeling, is infrequently used for routine seasonal prediction of El Niño and the Southern Oscillation (ENSO), the largest known natural variability that affects the global climate. In this work we propose a simple ocean initialization technique that can be adopted for any CGCM for seasonal predictability studies of ENSO. The technique entails finding the best analogues from a long historical simulation of the CGCM to the targeted air-ocean initial state. Since this study is on seasonal ENSO predictability, the metrics chosen to pick the analogues were confined to a set of 4 variables in the tropical Pacific that were found sensitive to the Niño3.4 SST variations. They were Tropical Pacific SSTs, thermocline depth, time tendency of thermocline depth, and the zonal wind stress. The multiobjective optimization technique was used to optimize the overall analogue match across the four variables giving equal weighting to each. This in effect uses the minimum root mean square difference between the targeted initial state and the model states to pick the analogue from the historical simulation of the CGCM that matched the targeted initial state. The chosen analogues were then perturbed using empirical singular vectors to provide additional initial conditions to generate in total 12 ensemble members per seasonal hindcast.

The methodology for ocean initialization was first tested with the Cane-Zebiak model, a two layer reduced gravity ocean model coupled to a statistical atmosphere. We found that the methodology is sensitive to the length of the library generated from the historical simulation of the model and also on the fidelity of the model in simulating the ENSO. These toy model experiments also revealed the benefit of using a multi-variate metric to choose the analogues. Before proceeding to conduct the proposed work with a CGCM, the CMIP5 historical simulations for the 20th century were analyzed for their ENSO simulation. The mean-state and ENSO variations were analyzed in both the atmosphere and ocean. It was found that most of the CMIP5 models exhibit cold (warm) biases in the equatorial (subtropical eastern) Pacific Ocean sea surface temperature that are reminiscent of the split inter-tropical convergence zone phenomenon. There is, however, a major improvement in the representation of the power

spectrum of the Niño3.4 sea surface temperature variations which shows that, as in the observations, a majority of the models display a spectral peak in the 2-7 year range, have a near linear relationship with the displacement of the equatorial thermocline and exhibit a robust atmospheric response to ENSO variations. Several issues remain in the CMIP5 simulations such as erroneous amplitudes in the Niño3.4 sea surface temperature spectrum's peak and a width of the spectral peak that is either too broad or too narrow. It is also seen that most CMIP5 models unlike the observations extend the ENSO variations in the equatorial Pacific too far westward beyond the dateline and there is very little asymmetry in event duration between the warm and cold phases. ENSO variability forces a dominant mode of rainfall variability in the southeastern United States, especially in the boreal winter season. The CMIP5 exhibited a wide range of response in this metric with several displaying weak to non-existent, some showing relatively strong, and one indicating excessively zonally-symmetric teleconnection over the southeastern United States. Based on this study we choose to use the CCSM4, which displayed a reasonable ENSO simulation for our experimental seasonal hindcasts with the proposed ocean initialization strategy.

The seasonal hindcasts were initiated in beginning of March of each year from 1980-2012 follows from seeking a model state that minimizes the RMS difference in SST, zonal wind stress, thermocline depth and thermocline depth tendency from a 600 year continuous integration of the CCSM4 with the corresponding metric in the Global Ocean Data Assimilation (GODAS) of the National Centers for Environmental Prediction (NCEP). The four variables are jointly optimized by multi-objective optimization of the resulting root mean squared (RMS) difference curves, essentially minimizing the normalized RMS in all four parameters.
Some of the main highlights of our results from the seasonal hindcasts are:

i) The deterministic skill as measured by the anomaly correlation of the monthly ensemble mean and observed SST anomalies in the Niño3.4 region at 9-month lead is 0.71.
ii) The probabilistic prediction of the Niño3.4 SST anomalies at 9-month lead for warm and cold ENSO events as measured by the area under the Relative Operating Characteristic Curve is 0.7 and 0.8 respectively. Likewise the brier skill score for warm and cold ENSO events at 9-month lead for Niño3.4 SST anomalies is 0.11 and 0.21 respectively.

iii) The global teleconnection patterns in SST, precipitation and 500hPa geopotential heights with Niño3.4 SST variations in the seasonal hindcast in Oct-November-December season (7 month lead) is reasonable.

From these results we demonstrate that the proposed initialization strategy is viable to deploy many other existing CMIP5 models for either operational seasonal ENSO prediction or ENSO predictability studies.

CHAPTER ONE

INTRODUCTION

The El Niño–Southern Oscillation (ENSO) is one of the most well-known and largest sources of natural variability at interannual scales (Philander 1990). ENSO is a coupled ocean-atmosphere phenomenon, which manifests most apparently with prominent Sea Surface Temperature Anomalies (SSTA) appearing in the eastern equatorial Pacific Ocean and off the coast of Peru, which persists for several months. These SSTA appear irregularly with a broad periodic range of 2-7 years (Philander 1990). The warm (cold) ENSO episodes are characterized by warm (cold) SST anomalies appearing in the central and eastern equatorial Pacific Ocean. Such variations of SSTA in the equatorial Pacific have global ramifications broadly felt in the changes in seasonal mean precipitation and temperature changes across the world (Rasmusson and Carpenter 1982, Ropelewski and Halpert 1987, Trenberth et al. 1998, Goddard and Dilley 2005). This is illustrated in Figure B.1. More specifically, it is a well-known teleconnection pattern that warm (cold) episodes of ENSO are coincident with the cold and wet (warm and dry) winters in the southeastern United States. ENSO also affects the extreme weather events like the Atlantic tropical cyclone (TC) frequency (Gray 1984) and tornado outbreaks in the US (Lee et al. 2013).

ENSO is most commonly quantified by a simple one-dimensional metric such as the SST in the equatorial Pacific or the difference in sea level pressure (Southern Oscillation Index, or SOI) between Darwin, Australia, and Tahiti. Figure B.2 shows a map of the various critical Niño regions, where SST is monitored for ENSO variations. For example, Niño1+2 (0-10S; 80-90W) is one of the first areas where the SSTA associated with ENSO events begin to unfold. Similarly, Niño3 (5S-5N; 150W-90W) is a region with the largest SSTA variability associated with ENSO. Likewise the slightly overlapping region but further west of Niño3 is the Niño3.4 (5S-5N; 170W-120W) region whose variations in SSTA are considered to be critical for manifestation of global teleconnections of precipitation and temperature (Barnston and Chelliah 1997). And lastly Niño4 (160E-150W; 5S-5N) region is further west of Niño3.4 region, which characterizes the eastern most extent of the 27.5C isotherm that is considered critical for sustaining deep convection in the atmosphere associated with ENSO variability. On the other

hand the SOI is an ENSO index of the longest record that dates back to the late 19th century. However it is quite noisy as the index is measured as the mean sea level pressure difference between two points that can be easily affected by transient weather disturbances.

ENSO is defined in several different ways (Hanley et al. 2003). For example, the Climate Prediction Center at NCEP defines ENSO when the 3 month running mean of Niño3.4 SSTA for 5 overlapping season exceeds 0.5C. On the other hand the Japanese Meteorological Agency (JMA) defines ENSO in the Niño3 region when its 3 month running mean anomalies are either below the 25th percentile or above the 75th percentile. The reader is referred to Hanley et al. (2003) for a more detailed discussion on ENSO indices.

ENSO prediction now relies heavily on dynamical modeling using coupled general circulation models (CGCM) because ENSO is a coupled air-sea phenomenon. Saha et al. (2006) clearly demonstrated in a systematic manner for the first time that a dynamical model has seasonal prediction skill of ENSO at par or superior to the statistical prediction schemes. The dynamical prediction models have been used for operational ENSO prediction ever since the first success of predicting ENSO (Cane and Zebiak 1985). More recently there is an ongoing national effort, where multiple CGCM's are being run routinely for generating ENSO predictions (Kirtman et al. 2013). The multimodel ensemble forecast for ENSO is based on the premise that it will fetch a better estimate of the influence of the model uncertainty on ENSO predictions.

a) A number of studies examining the model errors in simulating ENSO (Davey et al. 2001, AchutaRao and Sperber 2006, Lin 2007, Misra et al. 2007, Guilyardi et al. 2009, Bellenger et al. 2013) have identified the following common biases:
b) A split ITCZ phenomenon where the ITCZ is zonally symmetric in either hemisphere about the equator
c) A cold bias in the eastern equatorial Pacific Ocean
d) A warm bias in the eastern subtropical Pacific Oceans
e) A tropical thermocline that is diffuse
f) Insufficient upwelling along the eastern boundaries
g) Interannual variability of SST that extends too far to the west (west of dateline) along the equatorial Pacific Ocean

In a related study, Vanniere et al. (2012) found that the main source for biases in seasonal forecasts using CGCM's occur during the transition from initialization to stabilized mean state.

Misra et al. (2008) illustrated the ill effects of poor initialization on ENSO prediction. The study showed that the strong observed correlation (=0.77) between the depth of the thermocline and the SSTA in the Niño3 region (Figure B.3a) is completely lost when the CGCM is poorly initialized (Figure B.3c). The observed correlation of the SST and thermocline depth is nearly restored in the free running CGCM, which is far removed from the initial conditions (Figure B.3b). This study clearly illustrates the effect of the initialization shock on ENSO, which results from the imbalance in the prescribed initial state and the model climate.

In this study, we will initialize a CGCM with initial states that conform to the climate of the model, thereby producing no initialization shock in the forecast. In Misra et al. (2008) and Kirtman and Min (2009), the argument is that imposing an assimilated (i.e., near perfect) initial condition on an imperfect climate model leads to an adjustment or initialization shock; the analogue method results in no adjustment as the initial states already meet the CGCMs imperfect climate. As a consequence of the success of this methodology of initialization, there is scope for wider deployment of CGCM's for routine seasonal ENSO prediction, which is often used in research or climate change type long-term integrations.

The overarching objective of this study is to develop an effective initialization technique for coupled ocean–atmosphere models that will lead to improved ENSO prediction skills. The approach is to develop an easily adaptable initialization methodology of analogue matching (Vikhliaev et al. 2007), so that many of the climate models used in a research environment can potentially be used for routine ENSO prediction. Most common initialization techniques produce imbalances in the initialized state of the coupled ocean–atmosphere models that result in significant errors in ENSO prediction (Misra et al. 2008; Kirtman and Min 2009). In analogue matching, the history of climate states is compared with the current observed state for the closest match. When this match is found, the past evolution of the climate is used to predict the current state. We also supplement this initial state by generating ensemble members from two methods: analogues and bred vectors.

The following chapter will provide an in-depth review of the global effects of ENSO, ENSO theory, and ENSO prediction. Chapter 3 will explain the objectives of this research work. Chapter 4 describes the CGCM used for this model followed by a description of the initialization technique in Chapter 5. We give a brief background into the choice of CGCM used in Chapter 6 followed by the results from the Cane-Zebiak model, a simple toy model of ENSO and the

results from CCSM4 retrospective forecasts. Finally, we close with concluding remarks and perspectives on future work in Chapter 7. For reference, a list of all acronyms used in this document with a brief description can be found in Table B.1.

CHAPTER TWO

BACKGROUND

2.1 Global Effects of ENSO

The global effects of ENSO dominate the interannual climate variability of many regions to the extent that ENSO is the primary basis for seasonal to interannual climate prediction. Ropelewski and Halpert (1987) first presented a comprehensive overview of ENSO teleconnections (seen in Figures B.4 and B.5). They showed the remote influence of ENSO variations in SST over the eastern equatorial Pacific Ocean with regional precipitation (Figure B.4) and surface temperature anomalies (Figure B.5) across the globe. ENSO is also associated the modulation of the frequency and intensity of extreme events such as floods, droughts, tropical cyclones (TCs), and monsoons (Glantz 2000). Furthermore, the strength of the ENSO event can impact these global teleconnections nonlinearly (Hoerling and Kumar 2002). In addition to the canonical ENSO that is measured by the eastern equatorial Pacific SSTs, the newly discovered Central Pacific warming (CPW), often referred to as "Modoki El Niño," has its own distinct climatological teleconnections (Ashok et al. 2007).

2.1.1 ENSO Teleconnections over North America

Interannual variability of temperature and rainfall are measurably influenced by ENSO. North American winters are heavily influenced by the warm and cold phases of ENSO. The Southeastern United States experiences a cooler and wetter winter during the warm phase and a warmer and drier winter during the cold phase (Diaz et al. 2001, Glantz 2000). Conversely, the northwestern portion of the continent has a milder winter during the warm phase and a colder winter during the cold phase. The ENSO-related shift of the subtropical jet over the Pacific Ocean and the North American continent is responsible for these changes (Glantz 2000). The changes in climate linked to ENSO can result in increased wildfires in the Florida Keys (Harley et al. 2013). During Modoki events, the seasonal variations are typically opposite of canonical ENSO (Hoerling and Kumar 2002, Larkin and Harrison 2005b). Rainfall teleconnections over the southeast and southern California remain similar for canonical and Modoki ENSO whereas much of the US reverses the teleconnection patterns (Larkin and Harrison 2005a).

Atlantic hurricane activity is also substantially influenced by ENSO (Gray 1984). During a warm phase there is less TC activity because increased vertical shear in the TC development region inhibits the strengthening of storms. A cold ENSO phase reduces the amount of shear and therefore more TCs are observed in cold years. Tang and Neelin (2004) have shown that the onset period of a warm phase can also alter TC activity. If the beginning of the non-neutral phase coincides with the start of a hurricane season, then the adjustment from neutral to warm (cold) causes the troposphere over the TC development region to be more (less) stable than it is during a neutral year, thus acting to diminish (promote) TC activity for several months. Kim et al. (2009) found that tropical storms in the North Atlantic basin make landfall more often during Modoki El Niño than during canonical El Niño. Furthermore, Hoerling and Kumar (2002) note that weak warm ENSO events results in drier than normal northwest US winters. However, this teleconnection is non-linear: for moderate and strong warm ENSO events, the Pacific storm track takes a slightly southerly route which causes more precipitation in California. In a more predictable fashion, in the southeastern US, winter climate teleconnection with ENSO is linear, with temperature and precipitation teleconnections becoming more pronounced with stronger ENSO events.

In accordance with the North American teleconnections described above, modeled ENSO conditions in the northwest are drier during a weaker warm ENSO phase (Hoerling and Kumar 2002). It might be assumed that the stronger the warm anomaly is, the drier the conditions in this region would be. However, it was found that for moderate and strong warm phases, the Pacific storm track takes a different route and causes more precipitation in California rather than less. In the model, the cool–wet teleconnection in the Southeastern United States does prove to have a more linear relationship with increasingly stronger warm events.

2.1.2 ENSO Teleconnections over Asia and the Pacific Region

For years the India summer monsoon had been observed to be strongly influenced by ENSO (Sikka 1980, Shukla and Paolino 1983) because of the changes in the zonnally averaged flow during ENSO events over the region (Nigam 1994). The predictability of the variation in rainfall over India which results in devastating floods and droughts relies largely on ENSO prediction. More recently, the ENSO teleconnection with Indian monsoons has been found to be weaker than before (Shukla 1995, Kinter et al. 2002). It has been suggested that this change was a product of climate change (Kumar et al. 1999). Despite this weakening teleconnection, recent

studies continue to suggest that ENSO has significant influence on the seasonal predictability of the Indian monsoon (Gadgil et al. 2004, Jiang et al. 2013, Kim et al. 2012, Chowdary et al. 2013).

In East Asia, ENSO also has a significant impact on rainfall patterns both by affecting Western North-Pacific monsoon and North Pacific TCs (Yonekura and Hall 2013). The prominent feature associated with the warm phase of ENSO is a precipitation anomaly wrapping around the area of the Niño3 region in the Eastern Pacific and extending to the northeast and southeast from the western equatorial Pacific like a sideways horseshoe. This anomalous "horseshoe" pattern is produced by the convective branch of the Walker cell being displaced further eastward from Indonesia to the open ocean of the central equatorial Pacific.

During warm ENSO events rainfall significantly diminishes in Northern East Asia (Chou et al. 2009). While the Indian monsoon is affected by ENSO in the summer months, the East Asian monsoon is modulated during the winter with 20% of the fall-winter rainfall variance explained by ENSO (Wu et al. 2003). In the Philippines, the rain increases during warm events as a result of a large-scale anomalous anticyclone (Wang et al. 2000). It should also be noted that the SSTs in the South China Sea are also modulated by ENSO via the Luzon Strait (Fang et al. 2003, Qu et al. 2004). During the Modoki type ENSO events that are not coincident with positive Indian Ocean Dipole (IOD) phase, the East Asian-Western North Pacific summer monsoon becomes stronger while less rainfall occurs in the Meiyu–Baiu (Feng and Chen 2013).

It should be no surprise that ENSO plays a role in the interannual variability of North Pacific TCs given the proximity of the development region of the North Pacific TCs and the anomalous equatorial Pacific SSTs of ENSO. During cold events, changes to track and genesis site result in increased landfall rate of TCs over Vietnam, China, Korea, and Japan; during warm ENSO events the changes to track and genesis location negate any changes to the frequency of land-falling TCs (Yonekura and Hall 2013).

2.1.3 ENSO Teleconnections over the Southern Hemisphere and Africa

One of the leading methods for measuring ENSO is with the Southern Oscillation Index – the difference in sea-level pressure of Tahiti and Darwin, Australia. As this suggests, the effects of ENSO on the Southern hemisphere and Australia are significant. During warm (cold) ENSO events, the eastern half of the continent (the populated half) has strong droughts (floods).

Australia is the content most affected by ENSO. The most pronounced effects of ENSO are on the interannual variability of rainfall (Nicholls 1992). The 2011 Queensland floods which resulted in billions of dollars in damages were largely attributed to the La Niña event that year (www.bom.gov.au). ENSO forecasts have been found to be very helpful in Australian for managing fresh water resources (Simpson et al. 1993) and in aiding in fire and predator management (Letnic and Dickman 2006). The strong effects of ENSO on the region make it possible to find information on ENSO before the recorded era via coral samples and tree rings (Diaz and Markgraf 1992).

In South America, the effects of ENSO vary widely across the content. Droughts in Brazil increase during the cold phase of ENSO (Ropelewski and Halpert 1987, Halpert and Ropelewski 1992). In Monagas, Venezuela, El Niño years tend to be 5 to 33% drier than average whereas La Niña years increase the rainfall during the normally dry January to April period (Lopez Perez 2012). The effects of ENSO on precipitation also allow for skillful prediction reservoir levels in many areas of South America including Columbia and Ecuador (Gutiérrez and Dracup 2001, Gelati et al. 2013).

While the developed world has spent years of research on the effects of ENSO on their regional climate, ENSO also has strong effects on less inhabited regions of the world, which should not be ignored. ENSO teleconnections in the Southern Ocean can affect sea ice concentrations (Deb et al. 2013).

The Southern Ocean plays a crucial role in the global climate. While it is one of the least inhabited places on earth, this Ocean affects the global air-ocean CO2 exchange, the formation and deterioration of Antarctic sea ice area (SIA), and the transmission of climate signals between Atlantic, Pacific, and Indian Oceans (Gille 2002). ENSO phase plays a hand in the circulation and temperature patterns of the Southern Ocean, which can result in changes in SIA. The ENSO index has a negative correlation with SIA at 6-12 month lags. These correlations suggest that warm (cold) events are associated with smaller (larger) SIA in the following winter. Much like canonical warm events, Modoki ENSO reduces SIA, however, unlike canonical ENSO, the SIA reduction continues into the following summer and winter (Deb et al. 2013). Another teleconnection in the Southern Ocean involves the southern hemisphere blocking events, an event which dramatically changes the flow of the jet stream dividing it into two branches. For this teleconnection the impacts are influences not only by ENSO but also by the southern annular

mode (SAM). During warm events when the Southern Annular Mode (SAM) is in a negative phase, there are more blocking events over the East Pacific region. When ENSO is in a cold phase and SAM is strong there are far fewer blocking events (Oliveira et al. 2013).

ENSO phase has been observed to influence rainfall over Africa's largest lake, Lake Victoria, which supplies fresh water to one of the largest populations in Africa (Gissila et al. 2004). Over-lake rainfall increases during warm ENSO events and decreases during cold ENSO events (Awange et al. 2013). Similarly, ENSO accounts for 10% of rainfall variability in Tanzania (Enos 2013). The influence of ENSO throughout Africa also aids in climate forecasting of rainfall in Ethiopia (Gissila et al. 2004). In southern Africa, warm ENSO events are correlated with warmer temperatures from October to the following June and drier conditions from November to May (Ropelewski and Halpert 1987; Halpert and Ropelewski 1992).

2.1.4 Biological Effects of ENSO

The ability to predict ENSO events can lead not only to better climate forecasts but also the ability to protect fragile biological ecosystems. Drastic changes in ecosystems also occur during ENSO events and this is most notable in the island and coastal regions near the Eastern Pacific. Changes in the eastern Pacific SST can set off a chain of biological successes and failures that range from phytoplankton blooms to near extinction of mackerel and seals (Barber and Chavez 1983; Pauly et al. 1989). During a warm ENSO event, sunlight and nutrients are less available to the phytoplankton near the coast of Peru. Nutrients are absent during this phase because the depth of the thermocline has increased to a level below the depth of entrainment. As a result the upwelling is not transporting as much nutrient-rich water to the surface. Second, less sunlight will reach the bottom of the mixed layer because the thickness of the mixed layer has also increased during the warm phase (Cane 1983) and sunlight decreases rapidly through seawater. Since the phytoplankton are roughly uniformly distributed in the mixed layer the amount of sunlight reaching any given phytoplankton is reduced and so the population is smaller. Less phytoplankton in this region leads to less biomass for higher organisms.

These changes in local ecosystems, when coupled with human interaction, can affect species survival in a particular region. For example, overfishing alone in a region has lead to species reduction, and when combined with a change in fish population related to ENSO phase, near extinction has been observed for several species (Barber and Chavez 1983). Near or total extinction of a species is a tragedy, and the economic impact can be the permanent dismantling

of a regional industry. In the Australia outback ENSO affects the populations of rodents and their predators. Letnic and Dickman (2006) claim that predators are vulnerable to extinction from the combined effect of the natural ENSO-induced climate variably and the anthropogenic changes to the landscape. Agriculture is also often a victim when an extreme unanticipated ENSO event occurs. When crops fail, great social and economic strife can follow.

Understanding how ENSO can affect the biological realm so profoundly requires looking at the basic physical process that leads to the change in ecosystem. For the coast of Peru, the resulting change in ecosystem alters the phytoplankton population, which then causes proportional changes in fish, bird, and marine mammal populations (Barber and Chavez 1983). The jack mackerel were severely affected by the 1982 ENSO event. Because of the changing environment and diminished food source, the mackerel, usually seen offshore, moved into a tight 30-km band along shore. The increased density allowed fishermen and predatory fish to nearly eradicate the mackerel (Cane 1983). The Galapagos fur seal was also gravely impacted. Because of the reduction of fish, adult seals were less successful at finding food. Observations show that adult females stayed at sea for about 5 days (3.5 days longer than normal) and returned with too little milk to feed their pups. The effects of the 1982 event were strong enough to kill not only the majority of pups, but also many adults (Pauly et al. 1989). Forecasts of ENSO can help governments regulate fishing quotas to avoid overfishing and allow for maximum sustainable yields (Ortiz and Sánchez 2011).

The effects of ENSO on the agriculture industry are significant as well. Changes in seasonal rainfall in the Southeastern United States related to ENSO are a key influence on agriculture in that region. For instance, in the Southeastern states, crops such as corn and tobacco, which are sensitive to precipitation, are significantly affected by ENSO (Hansen et al. 1998). Interestingly, the failure of the fishing season due to warm ENSO events may also impact crops planted in the Southeastern United States. In years when demand for fish is high, the price of livestock feed, which contains fish protein, increases. In response to this increase in the price of and demand for livestock feed, soybean farmers tend to plant more soybeans, which can substitute for fish protein, two years after warm events (Hansen et al. 1998).

Reliable seasonal climate predictions can inform decision makers on hedge their risk on when and where to plant crops. Being forewarned about the nature of upcoming ENSO events or even its predictability can not only protect farmers from crop failures but also can help them

choose crops that may be in higher demand. During warm ENSO event winters, the prices of bell peppers and snap beans increase. The increase in hard-freezes for the Southeastern United States reduces crop yields of winter-harvest vegetables. In Florida, diminished efficacy of fertilizers and damaged plant roots result from the excess rainfall during El Niño years. Because Florida is the primary producer of winter-harvest bell peppers and snap beans, farmers in regions outside Florida with mild winters can capitalize on this shortage (Hansen 1999).

2.2 ENSO Theory

The above sections illustrate the importance of ENSO on our planet and value of skillful ENSO forecasts. This section will lay the ground work to understand the physical mechanisms underlying ENSO variability based on current theories.

Currently, two ENSO theories are widely accepted by the scientific community. The first is the delayed oscillator theory (DOT, Suarez and Schopf 1988), which explains ENSO in terms of equatorial Kelvin and Rossby waves in the Pacific Ocean. The second, the recharge–discharge theory (Jin 1997), explains ENSO in terms of Sverdrup transport of warm water volume in the ocean. In addition to these two paradigms of ENSO, several other theories describe the air–sea interactions responsible for ENSO. For example, Wang (2001) put forth a theory that combined all the leading contemporary theories of ENSO into one unified theory. Both DOT and recharge-discharge theory can be useful in explaining the fundamental physics of the canonical ENSO. Theories on the Central Pacific warming events, Modoki ENSO, are still not widely agreed upon (Takahashi et al. 2011).

The underlying principal of DOT (Suarez and Schopf 1988) is that perturbations in zonal wind stress in the central equatorial Pacific will generate equatorial Kelvin and Rossby waves. These waves propagate zonally and displace the thermocline vertically as they travel across the Pacific basin to counteract anomalies in the eastern equatorial Pacific that grow from the positive air–sea feedback (Bjerknes 1969). To the east of the wind stress anomaly, Kelvin waves move toward the eastern boundary, the western coast of South America, and then partially reflect back to the west as Rossby waves. Likewise, to the west of the wind stress anomaly, equatorial Rossby waves travel to the eastern boundary and reflect back as Kelvin waves.

Action of these waves to displace the thermocline vertically (either upwelling or downwelling) depends on the sign of the wind stress anomaly over the western-central Pacific Ocean. If the anomaly is positive (an eastward anomaly), the eastward propagating Kelvin wave

will advect the thermocline downward (downwelling Kelvin wave), creating a positive SST anomaly in the eastern Pacific, and the initial westward propagating Rossby wave will upwell in the western Pacific. If the zonal wind stress anomaly is westward over the western-central Pacific Ocean, the initial waves will lead to a cold ENSO phase. Upon reflection, the upwelling/downwelling waves reach the opposite edge of the Pacific Ocean and reverse the SST anomaly.

The time scale for the Kelvin waves and Rossby waves to cause a change in SST in the Pacific Ocean is much less than the observed return period of ENSO. It takes about one month for a Kelvin wave generated in the central Pacific to reach the eastern boundary. Rossby waves, moving at one third the speed of Kelvin waves, require three months to reach the western boundary. Thus, it takes approximately four months delay time for the reflected waves to reach the central Pacific—much shorter than the observed ENSO period.

Since Suarez and Schopf's study (1988), there have been many modifications and suggested improvements to the DOT that help to better explain the observed return period of ENSO. Cane et al. (1990) proposed that instead of one wave propagating away from the central Pacific, multiple waves propagate and the cumulative effect alters the SSTs in the eastern Pacific and reverses the anomalies with the observed period of ENSO. This hypothesis was further refined by Schneider et al. (1995). They proposed that the much slower westward propagating off-equatorial Rossby waves interact with the equatorial Kelvin waves to produce the observed ENSO period.

As shown by Jin (1997), the ENSO oscillations can be explained by the time evolution of the equatorial SST and the thermocline depth. Figure B.6 illustrates a schematic view of the Pacific Ocean during four phases of ENSO evolution (Meinen and McPhaden 2000). To begin, the basic state of the thermocline can be expressed in terms of its depth in the eastern and the western Pacific and the zonal wind stress along the equator. At the equator, wind stress anomalies to the west will drive the thermocline to be deeper in the west and shallower in the east. Conversely, a westerly anomaly will shoal the thermocline in the west and deepen it in the east. As stated previously, a westerly wind stress anomaly will cause the thermocline in the west to shoal. As the trade winds slacken in the east, SST tendency is positively forced because there will be less Ekman pumping, bringing less cool water to the surface. Furthermore there will also be less evaporative cooling from the cooler temperatures and weak winds over the eastern

equatorial Pacific Ocean. The atmosphere responds to these changes in equatorial Pacific SSTs: warmer eastern equatorial Pacific SSTs cause anomalous westerlies over most of the equatorial Pacific (with a westerly anomaly to the west and an easterly anomaly to the east of the SST anomaly).

Starting from an initial neutral state, a small positive SST anomaly will lead to a growth of the anomaly from air-sea feedback. This will result in basin-averaged positive wind stress anomalies as a response to the change in large-scale equatorial Pacific SST gradients. In turn, this will result in a deepening thermocline in the east and a shoaling thermocline in the west equatorial Pacific Ocean. Thus, recharging the warm water volume in the equatorial Pacific Ocean. This in turn culminates at the height of the recharge phase by discharging heat to off-equatorial region via the meridional Sverdrup transport out of the equator as a result of the disequilibrium between the zonal winds and the zonal mean thermocline. This leads to the SST anomaly in the equatorial Pacific Ocean coming to a neutral state with an anomalous shallow thermocline across the equatorial Pacific Ocean. The anomalous shallow thermocline is also the phase with the lowest warm water volume in the equatorial Pacific, which marks that discharge of warm water is complete. However, SSTs in the equatorial eastern Pacific Ocean continue to cool because of the upwelling of anomalous cold water in the subsurface ocean marked by a very shallow thermocline. As the SSTs cool below climatology, anomalous easterlies develop that further force the SSTs down because of enhanced upwelling and evaporative cooling. The strengthened trades from change in SST gradients cause eventually a recharge of the warm water volume through meridional Sverdrup transport of warm water from off-equatorial regions as a response to the growing disequilibrium between the zonal winds and the zonal mean thermocline. The result is the return to neutral SST anomalies in the eastern Pacific Ocean with anomalously deep thermocline across the equatorial Pacific Ocean, which marks the recharge of the equatorial Pacific Ocean. With continued upwelling in the eastern Pacific Ocean, the climatological tilt of the equatorial Pacific thermocline is restored, which marks the true neutral state of ENSO.

A fundamental property of ENSO is its apparent phase locking to the seasonal cycle (Rasmusson and Carpenter 1982). ENSO events tend to gradually build to a peak by the end of a calendar year and then rapidly subside by mid-spring. This phenomenon is often referred to as to the spring persistence barrier or the spring predictability barrier because there is a large drop in

SST persistence in April and May, the cause of poor model prediction of ENSO during the spring.

To understand the mechanism for this phase locking, the seasonal propagation of the anomalous trade winds needs to be identified. During an ENSO event, the changes in the Walker circulation lead to anomalous convection in the western Pacific and anomalous winds over the central equatorial Pacific. Because of the change in the solar angle throughout the year, the latitude of these anomalies shifts south in January–March as the peak solar radiation moves south of the equator (Clarke 2008). Gadgil et al. (1984) showed that when SSTs reach 28°C convection can occur. So as the water south of the equator warms, the convection shifts southward, causing wind anomalies to do the same.

The propagation of the anomalous winds southward at the beginning of the calendar year removes the positive feedback of ENSO. For instance, during a warm event the anomalous westerlies allow the SSTs in the eastern Pacific to warm. As this anomaly moves south, the thermocline depth begins to shoal and SSTs cool in the eastern equatorial Pacific, thus ending the warm phase. Therefore, the demise of ENSO events is significantly phase locked to the seasonal cycle even though these events do not always occur at regular intervals.

2.3 ENSO Forecasting

ENSO forecasting can be divided into two groups: statistical forecasting and dynamical forecasting. Statistical models are usually Bayesian with the forecast model based on pre-existing relationships between ENSO predictands and its predictors. This class of forecast model cannot predict how ENSO may change in future climates since they rely solely on data from prior ENSO events. Bunge and Clarke (2014) demonstrated the potential problems such models face when these long-standing climatic relationships change. Dynamical ENSO models generate forecasts based on approximating fundamentals of ocean-atmosphere system and simulating the evolution of ENSO based on current conditions. In practice both types can produce skillful forecasts and are often combined in a multi-model ensemble forecast of ENSO.

2.3.1 Statistical and Dynamical Models

A large range of both statistical and dynamical models of ENSO are currently in use operationally forecasting ENSO. The International Research Institute for Climate and Society

(IRI) releases seasonal forecasts using an ensemble of many of the best available models; the February 2014 ENSO prediction is shown in Figure B.7.

Statistical models of ENSO work exceedingly well at forecasting ENSO metrics such as the Niño 3.4 Index. In Fact, statistical forecasting was used before dynamical models of ENSO were available. These models typically use SSTs, wind stress, and related variables as inputs to predict seasonal evolution of ENSO. One type of statistical model used is a low-order Markov model (Xue et al. 1994). The Climate Prediction Center (NCEP/CPC) Markov model is a prominent statistical model currently in use which includes SSTs, wind stress, sea-level pressure and seasonality (Xue et al. 2000). Another type of statistical model is the linear inverse model (LIM) which uses the lag covariance of predictor variables. These models excel when linear dynamics dominate, during times of rapid changing SSTs rather than persistent conditions (Penland and Magorian 1993). Still other statistical models are regression-based; Clarke and Van Gorder (2003) incorporate upper ocean heat content and Indo-Pacific winds to forecast Niño3.4 SSTs. While some statistical models produce global SST forecasts, most do not go beyond a simple prediction of ENSO metrics as discussed in Chapter 1. Figure B.8 shows an example of model performance of a statistical model, the Clarke and Van Gorder ENSO model.

The first generation of dynamical models included only a simple ocean and atmosphere coupled system and consisted of only a small domain covering the Pacific Ocean. The Cane-Zebiak (CZ) model simulated ENSO by forcing an anomaly model with an initial zonal wind stress (Zebiak and Cane 1987). Many more models improved on the basic CZ model by increasing the complexity of the model such as by allowing flow in and out of the western boundary (Battisti 1988). In turn some dynamical models are then linearized and used as statistical models such as the Linearized Ocean Atmosphere Model (LOAM) based on the Battisti (1988) model (Thompson 1998). By today's standards these models are considered "toy models" in that they are very simplistic and require little computer time to run. Nevertheless, many of these early models of ENSO are still used in research today (Mukhin et al 2012, Roberts et al. 2014).

Modern dynamical ENSO forecast models such as those shown in Figure B.7 are global CGCMs containing many more levels of complexity compared to the first generation models. As such, not all CGCMs are well suited for ENSO forecasting as many lack ENSO simulation (Guilyardi et al. 2009). Those models that are used for ENSO forecasting are not necessarily the

only model which could produce skillful ENSO forecast. Most international forecasting agencies can invest time to develop only one model as an operational forecasting tool and as such must weigh the strengths and weaknesses of the possible models. Some CGCMs used as ENSO forecast models include the NCEP Climate Forecast System model (CFS), the JMA climate model, the NASA Global Modeling and Assimilation Office (GMAO) model, and the European Centre for Medium-Range Weather Forecasting (ECMWF) model (JMA 2002, Saha et al. 2006, Troccoli et al. 2003, Van Oldenborgh et al. 2005, Wang et al. 2005)

2.3.2 Initialization Procedures

Most forecast initialization methods rely on data assimilation techniques that result in initialization shock. This shock refers to the unbalanced state of the initial conditions with respect to the model climate. The unbalanced state leads to an adjustment that produces unrealistic conditions in the coupled climate system—such as the breakdown of correlations between thermocline depth and SST mentioned in Chapter 1 (Misra et al. 2008).

The data assimilation process must be redeveloped for each new model before that model can be used to forecast. As a result, most CGCM models are not used in operational forecasting. This could be argued as a significant waste of resources as seasonal forecasts are increasingly being found to benefit from multi-model ensembles (Kirtman and Min 2009, Kirtman 2013). For example, Kirtman and Min (2009) showed that for two models that have orthogonal skill in forecasting ENSO, the combined multi-model ensemble produces a better overall skill over the whole hindcast period examined.

The analogue method of initialization is an alternative that does not require data assimilation specific to the forecast model because it draws on existing model states. Because the initial state used is taken from a state within the model climate, initialization shock errors are not introduced and logically the forecast skill is not diminished by this mechanism. In its most basic form, the analogue method of forecast initialization is nothing more than comparing two instances of a system and surmising that two relatively similar states will develop similarly in time. In the case of Earth's climate, both the complexity (degrees of freedom) and the instability (chaos) of the system make it virtually impossible to find exact analogues to make a skillful prediction. In application, the efficacy of the analogue method can be augmented with increased climate records (van den Dool 1994), incorporation of inverse analogues (van den Dool 1987), averaging an ensemble of analogues (Bergen and Harnack 1982), reducing the degrees of freedom (Gutzler and

Shukla 1984), or increasing the analogue library by using existing model simulations as will be done in this study.

The earliest attempts at forecasting with analogues produced poor results, leading to little or no confidence in the future of analogue initialization (Lorenz 1969, Radinovic 1975, Bergen and Harnack 1982, Gutzler and Shukla 1984). Lorenz (1969), who examined upper-air data for five consecutive years, was unable to find any analogues fit for a forecast initialization. Even though the likelihood of finding truly good analogues was small, it was suggested that a much larger dataset might lead to better analogues. Bergen and Harnack (1982) found that an analogue approach to forecasting could improve on persistence in some instances. In this case, 40 years of data were used as the analogue library to generate one- and two-month temperature forecasts. An ensemble of analogues was used to reduce the sampling error.

Although Bergen and Harnack (1982) improved the analogue approach by using an ensemble, other studies attempted to refine the use of analogues by other means. The analogue method's primary drawback is the difficulty of finding suitable analogues. The system has high degrees of freedom and the available data are limited. Gutzler and Shukla (1984) examined analogues for wintertime 500-millibar heights using a 15-year library. To reduce the degrees of freedom Gutzler and Shukla (1984) filtered for the long wave component and focused on RMS differences for regional forecasts. Van den Dool (1987) examined a way to increase the number of possible analogue matches by also testing for mirror analogues, dubbed "antilogues." Despite these efforts, results for analogue-based forecasts remain relatively discouraging: it would take an estimated 10^{30} years to find analogues that match over the entire Northern Hemisphere (van den Dool 1994). Given that climate records are on the order of 10^2 years, the only option is to reduce the degrees of freedom.

CHAPTER THREE

OBJECTIVES

The overarching goal of this proposal is to develop a simple initialization technique for coupled ocean–atmosphere models which will contribute to improving seasonal prediction skill of ENSO while also accelerating the effort to bring more climate models under the ambit of operational seasonal ENSO prediction. There is a growing consensus that a multi-model ensemble strategy is a pragmatic approach to resolve the forecast uncertainty (Palmer et al. 2004, Hagedorn et al. 2005, Doblas-Reyes et al. 2005, Palmer et al. 2008, and Kirtman et al. 2013). The uncertainty associated with model formulation including the discretization and the parameterization schemes for the various physical processes get further compounded in a non-linear system such as in a coupled general circulation model. Kirtman and Min (2009) clearly demonstrated the superiority of a multi-model approach to ENSO prediction compared to a single model, even when their multi-model comprised of just two climate models. More recently Kirtman et al. (2013) has shown that the global seasonal prediction skill (especially that related to ENSO) with nine different global climate models under the aegis of the North American Multi-Model Ensemble (NMME) improves upon the National Center for Environmental Prediction's (NCEP's) operational CFS version 2. Furthermore, the uncertainty demonstrated by seasonal climate forecast is considered to be very useful information, irrespective of its magnitude, when it is expressed in probabilistic terms (Mason et al. 1999, Palmer et al. 2000, Kirtman 2003, Misra 2008). The forecast uncertainty in such probabilistic measures is ideally a reflection of the reliability of the forecast. Therefore, any effort to use more climate models for routine ENSO predictions (raising the robustness of the measure of uncertainty) will translate to better information in the seasonal forecasts.

There is also a growing movement to have seamless prediction systems that address numerical weather prediction (NWP) with climate change projections with one numerical model (Randall et al. 2003, Hurrell et al. 2009, Shukla et al. 2009, Brunet et al. 2010). This stems from the recognition that interactions across all spatial/temporal scales are important for the climate system. One argument for promoting this seamless system is that the large-scale climate influences the small-scale variations in the environment (e.g. microscale ~1km or mesoscale ~10km) and these small-scale variations then feedback to the large-scale climate. Another reason

for promoting the seamless or the unified prediction tool is that the climate and the weather system evolves together with the coupled interactions of the various components of the earth system (the hydrosphere, biosphere, cryosphere, atmosphere, etc.), and is therefore erroneous to model with reductionist, singleton systems (e.g. an atmospheric general circulation model or an ocean general circulation model by itself). Therefore, there is good reason to begin bringing the class of models used in the latest IPCC assessment reports into use for routine operational predictions. Additionally, Rodwell and Palmer (2007) claim that some of the largest systematic errors in the climate projection models can be traced back to the first few time steps of the model integration of the initial time. Rodwell and Palmer (2007) indicate that fast physics (that which relates to convection in the atmosphere) contribute significantly to some of the systematic errors displayed by the climate model. Therefore by bringing the CMIP5 class of models into routine seasonal ENSO prediction, we can diagnose the model errors more efficiently, which could potentially lead to accelerated efforts for improving them. The proposed simple initialization strategy for ENSO prediction could help in improving the climate change projection models in the long run. It should be mentioned that a majority of the Coupled Model Intercomparison Project 5 (CMIP5) models (which are the primary basis for the IPCC AR5) are not deployed for either ENSO seasonal prediction or predictability studies. Even the NMME project recognized the bottleneck of initializing climate models for which none existed previously, and gave all the contributing modeling centers full freedom to develop their initialization protocol independently (Kirtman et al. 2013).

An analogue method of initialization for the ocean is proposed in this research as a viable option for a simple and effective method for seasonal ENSO prediction. The premise of the proposed initialization methodology is to reconcile the errors in the climate model to produce a balanced initial state, which could potentially lead to substantial improvement in ENSO prediction. A major shortcoming of data assimilative models is that the analysis is tightly constrained by observations, which then lead to significant initialization shocks when the model is integrated in the absence of the observational constraints. This is because the model will tend to drift to its own climatology. Another important motivation for developing a simple initialization technique is to make it easily adaptable for climate models that are traditionally run in a research environment or for long-term climate change experiments that are not sensitive to

initial conditions (e.g. CMIP5) so that they could also be leveraged for ENSO predictability studies. The success of this study could therefore potentially lead to:

a) Bringing more of the CMIP5 class of models in seasonal forecast settings and consequently provide better resolution of forecast uncertainty
b) Highlight the model errors and achievements with regard to ENSO (the biggest natural variability of our planet), which could lead to accelerated improvement of the climate models.

To fulfill the overarching goal of this study, two points must be fulfilled:

a) Reduce or eliminate the initial shock for which the initial conditions must be brought as close as possible to the model climatology, while it also relates in a optimal way to the observed state
b) The implementation must be feasible for a large array of dynamical models.

If the forecast initial states to be used originated from an existing integration of the same model, point 'a' would be satisfied because the initial state is in fact already in alignment with the model physics assuming a long running simulation is used. Using the analogue method to match the model simulation to observations would be feasible if a multi-centennial integration for the dynamical model existed.

With the analogue approach it is possible to avoid the initialization shock problem but several limitations apply.

a) First, the available model states (the library of possible analogues) must conform somewhat to observations. If there are no modeled states sufficiently close to observations, the forecast will have no bearing on reality.
b) The forecast model must be frozen, meaning that no changes to that model can be made. If the model climatology differs from the climatology of the library, then point goal 'a' above will not be met. In the current protocol of generating a substantial set of seasonal hindcasts and evaluating their skill before the models are put in operation, at least for seasonal prediction avoids periodic incremental changes to the model (Saha et al. 2006, 2010, Kirtman and Min 2009, Kirtman et al. 2013).

c) The analysis system will also be frozen. This would include freezing the data assimilation model and the observations that are assimilated so that the observed states are not contaminated with artificial discontinuities.

The basic approach of this study consists of two parts:

1. An exploratory research preformed with a toy model of ENSO and
2. A thirty-two-year (1980-2012) seasonal ensemble hindcasts analysis generated with a fully coupled CGCM initiated in the beginning of March of each year.

The toy model experiments serve to establish the proof of concept before launching the methodology in the more complex CGCM. The seasonal hindcasts with the CGCM are designed with the information garnered with the toy model. Standard benchmarks of forecast skill then provide a quantitative analysis of the analogue method of initialization. The toy model experiment provides several purposes. First, the toy model can prove if the analogue method of initialization is possible for ENSO forecasting. Second, because the toy model is computationally inexpensive compared to the CGCM, which allows us to create a much larger library to choose the analogues from; this allows for a sensitivity study on library length.

The fully coupled CGCM captures a much richer ENSO (e.g. it exhibits more diversity) than the toy model and can better represent the mean global climate and the teleconnections associated with ENSO (at least for some of the current generation of the CMIP5 models; Michael et al. 2013). In this study, the CGCM is used to predict real ENSO events in a hindcast setting in contrast to the toy model which is only tested against other model simulations (perfect model setup). Limitations of the CGCM will be the model bias, the shortened library length, and the smaller number of forecasts performed. The CGCM will be able to provide a global view of ENSO and possibly a wider range of ENSO variability as observations exhibit.

The specific objectives of this research work are:

i) Present the proof of concept of the initialization strategy with the toy model.

ii) Diagnose the ENSO features of the chosen CGCM vis-à-vis other CGCM's in the CMIP5 suite.

iii) Conduct seasonal hindcasts of 9-month duration for a 32-year period (1980-2012) using the CGCM with the adopted initialization strategy. We buttress the

analogue method further by perturbing the chosen initial condition with empirical singular vectors (Kug et al. 2011). Therefore a total of 12 ensemble members per season are generated, which is a combination of 4 initial states from the analogue method and two additional sets of perturbations superposed on every given analogue initial state.

iv) Analysis of the sensitivity of the metric chosen for picking ENSO analogues.

v) Analysis of the seasonal hindcast skill.

CHAPTER FOUR

METHODOLOGY

At the most fundamental level, the analogue method is very simple. Simply put the objective of the analogue method is to find the closest match to the target state. In our case the target is the observed state of the ocean at the initial time of a seasonal prediction. Of course, the job of matching a highly complex ocean-atmosphere state to a large number of modeled states is too tedious to conduct mechanically for any human. We therefore automate the procedure for identifying the best analogues with the same simple underlying principles of analogue matching. But the implementation is technical with many nuances.

In this chapter, the methodology used for analogue matching is described followed by the description of the empirical singular vectors to generate a larger set of ensemble members than what analogue matching can offer. As stated in the previous chapter, robust estimates of uncertainty of the seasonal forecast is viewed with equal interest as the seasonal climate forecast itself, especially when it is used in some form of decision support model. The analogue matching methodology is described both in the toy model and the CGCM. In each section of this chapter, we will provide a description of each procedure for the toy model model first followed by that with CGCM.

4.1 Model Description

For this study two different models were used: the toy model follows from Cane-Zebiak (CZ, Zebiak and Cane 1987) model and the Community Climate System Model version 4 (CCSM4). Along with CCSM4, 16 other CMIP5 models were also examined for their ENSO simulation (explained in the following chapter).

The CZ model used for this experiment follows from Zebiak and Cane (1987). By today's standards, the CZ is run at a very course resolution and large time step of 10 days. The total domain for both the ocean and the atmosphere in the CZ model extends over the Tropical Pacific, from 124°E – 80°W and 29°S – 29°N. Both ocean and atmosphere are anomaly driven with observed climatology.

The atmosphere is very similar to Gill atmosphere; they are linear shallow-water equations on an equatorial beta plane which includes Rayleigh friction and Newtonian cooling.

Departing from the standard steady-state Gill atmosphere, there is a convergence feedback from the coupled ocean such that the atmosphere can evolve in time. The feedback term is has two parts: a convergence feedback from the anomalous heating of SSTs and anomalous moisture at the surface. The moisture-driven convergence term responds only when the local atmosphere has positive convergence at the surface; for more information on the formulation of this feedback see Zebiak (1986). Below is the governing equations for the atmosphere

$$+\epsilon {u_a}^n - \beta_0 y {v_a}^n = -\left(\frac{p^n}{\rho_0}\right)_x$$

$$\epsilon {v_a}^n - \beta_0 y {u_a}^n = -\left(\frac{p^n}{\rho_0}\right)_y$$

$$\epsilon\left(\frac{p^n}{\rho_0}\right) + {c_a}^2\left[({u_a}^n)_x + ({v_a}^n)_y\right] = -Q$$

where Q is the convergence feedback and the parameter values are as follows:

$$\epsilon = (2\ \text{days})^{-1}, c_a = \frac{60m}{s}$$

The ocean is a simple 2-layer reduced gravity model, which allows for changes in the thermocline depth. A shallow 50-meter frictional layer of is also included to produce Ekman currents in the flow. The evolution of the ocean depends on feedbacks from the anomalous atmosphere. Below show the governing equations for the ocean model:

$$u_t - \beta_0 y v = -g' h_x + \frac{\tau^{(x)}}{\rho H} - ru$$

$$\beta_0 y u = -g' h_y + \frac{\tau^{(y)}}{\rho H} - rv$$

$$h_t + H(u_x + v_y) - rh$$

For further information on the model specifics, the reader is referred to Zebiak and Cane (1987).

The CZ model showed that stronger SST-heating in the equatorial oceans and stronger air-sea coupling lead to stronger amplitude ENSO. Furthermore the seasonal cycle of the ENSO variability in the CZ model showed that the mean conditions in the boreal summer season promoted the strongest growth and the interannual variability of ENSO manifested with the transition in the sub-surface ocean heat content in the equatorial Pacific (Zebiak and Cane 1987). With this configuration of the CZ model, an initial starting anomaly was needed to being ENSO

evolution. As following Zebiak and Cane (1987), this initial push used was in the form of a small westerly wind anomaly for four months. After the initial forcing, a warm ENSO event is triggered and from then on no external forcing is needed to maintain the oscillation. However the model exhibited limitations in explaining the evolution of the mean state and the trigger for the ENSO onset in the model aka the prescribed anomalous surface winds. But the model is still useful as it is able to "realistically" simulate the sub-surface equatorial ocean variations associated with ENSO with very limited computational resources.

CCSM4 model used for this study is publically available from the University Corporation for Atmospheric Research (UCAR) and can be downloaded from www.cesm.ucar.edu/models/ccsm4.0. This model is available in a large variety of configurations depending on the intended use. For this study a coupled ocean-atmosphere configuration was necessary. Because we wish to generate a multi-century library of model output, a fixed forcing was also needed. The configuration chosen to satisfy these requirements consisted of a fully coupled model free running simulation with the atmosphere and land components at 1.9 degrees latitude and 2.5 degrees longitude and with a curvilinear ocean and ice grid approximately 0.5 by 1.125 degrees (known as fv1.9x2.5 gx1v6). The north pole of the ocean/ice grid is displaced over Greenland. The only external forcing was fixed for present day (year 2000) conditions with an atmospheric CO_2 concentration of 367.0ppm (known as B_2000). CCSM4 contains five distinct model components: ocean, atmosphere, land, sea-ice, and a coupler. The coupler's main task is to translate and pass data between the differing grids of the other four models at each coupling time step. Notable improvements were made to the convection parameterizations used by this model configuration; for more detail see Gent et al. (2010).

4.2 Analogue Method

The salient variables for ENSO prediction must first be identified and then isolated for analogue matching of ocean-atmosphere states to be effective in initializing for ENSO forecasts. For example, a state that matches 500hPa over North America may have very little in common with the phase of ENSO. This type of "red herring" matching may lead to improper initial states for the ENSO forecast. From ENSO theory (Zebiak and Cane 1987; Kirtman 1997), we find that the most important fields are SST, thermocline depth, thermocline depth tendency, and zonal wind stress in the Tropical Pacific Ocean. Thus, only these four *predictor variables* are

considered when making any matching. The way in which these four variables are used will be discussed later in this chapter.

Following the procedures employed with the CZ model, 1-to 9-month forecasts will be made for ENSO. The proposed work will include an accumulation of the work previously done on CZ with the addition of several considerations necessary for working with a fully coupled GCM. First, the high degrees of freedom in CCSM4 versus CZ require that either a larger library be used or data in CCSM be filtered by some means. For this, we plan to employ ensemble empirical mode decomposition to reduce the degrees of freedom in CCSM4. Second, the wall clock time needed to execute CCSM4 code is considerably longer than that needed for CZ. Therefore, model experiments will be limited to those that will most contribute to attaining the proposed goals. Unlike the library length in CZ, which could be rapidly generated, the library length in CCSM4 will be limited by the computational resources and human resources.

4.2.1 Analogue Matching by RMS Difference

There are three common metrics that could potentially be used for analogue matching between the observed state and the corresponding model state with each having their own advantages and disadvantages based on the features of the variables being matched: 1) total root mean squared (RMS) difference 2) spatial covariance and 3) spatial correlation. Matching RMS emphasizes similar amplitudes of anomaly patterns while downplaying the phase over the domain. Covariance matching selects for the most pronounced anomalies ignoring areas that are near neutral. Correlation matching gives less weight to large anomalies and more to phase differences (Gutzler and Shukla 1984).

The analogue matching technique used here is similar to the RMS method used in Gutzler and Shukla (1984). RMS was chosen as the primary method because RMS difference is sensitive to amplitude of anomaly patterns. Furthermore, given the systematic errors over the equatorial Oceans in most CGCM's, we find that it is much harder to meet the analogue matching criterion using the correlation or the covariance metric than the RMS metric. In addition, the impact of an erroneous forecast on the strength of ENSO relative to its phase on agro-economy is found to be contrastingly different with the former having a larger impact than the latter in the southeastern United States (Solís 2013).

With the CZ model, a 5000-year long continuous simulation was generated for use as the library of potential model states that could be matched with the desired observed state. This

library was not however compared to any observations. Instead, two versions of the CZ model were used, with one (modified) version used as a proxy to observation and the other standard version used for model state. This was so done because the ENSO variability in a specific version of the CZ model is rather periodic and similar to each other, while systematically different from the observations.

Due to the simplicity and the limited use of resources used in running the CZ model, it allows us to examine the analogue matching approach in a perfect and imperfect model set up. Therefore to test the analogue matching in a perfect model scenario, an additional 100 years of integration was conducted with the standard version of the CZ model, which was used as proxy for observed ENSO states. The comparison of the results between perfect and imperfect model scenario will give us an idea of how model errors can impact the effectiveness of the analogue matching. In a realistic situation, the model states from the model library will be matched to an observed ENSO state from an Ocean Data Assimilation (ODA) that possibly uses another model than the seasonal forecast model. Therefore the imperfect model scenario of the CZ model would test this realistic situation. The perfect model scenario mimics the situation when the ODA is generated from the same model as the seasonal forecast model.

The modified version of the CZ model was generated by modifying two parameters from the standard version, which are viz., the parameter controlling the strength of the SST-related component to atmospheric heating and the parameter that relates to the decay time (diffusive timescale) of the ocean. An illustration of the spectrum of the Niño3 SST from the observations along with the standard and modified versions of the CZ is shown in Figure B.9.

In the modified model, the atmosphere feedback is made stronger and the ocean is made to adjust more rapidly than the standard version. The increased stochastic forcing of the atmosphere on the ocean in the modified CZ model results in a broader peak in the spectrum. What is important about these changes is that the modified CZ system will be harder to forecast than the original CZ model given the diversity of possible ENSO states.

For the perfect model experiments the analogues were found six different ways: Methods 1-4: by calculating the RMS differences of each the four predictor variables which were then used individually to pick the best analogue with the least RMS

Method 5: by taking the average normalized RMS of the first four and picking the one with the least RMS

Method 6: by taking the weighted average of the RMS of first four. In this case, the weighted average was based on the lag-correlation of the predictor variable with the Niño 3.4 SST.

Once moving on to the experiments with CCSM4, a 5000-year library was no longer possible as each year took approximately one day of Wall Clock Time (WCT) at the FSU High Performance Computing Center (circa 2011) to complete, which was later reduced by porting the code to a faster super computer (Stampede at Texas Advance Computing Center). In total, 600 years of library length were available for analogue matching in CCSM4 for this study. An attempt was made to use a lower resolution version of CCSM4, approximately 2 degree grid spacing. Figure B.10 shows the regression of Niño 3.4 SST on tropical Pacific SST for the low-resolution version of CCSM4 and the higher (standard) resolution. The figure shows that the lower resolution model has a pattern of SST warming similar to both observed and higher resolution versions of the model. However, the wider zonal extent of warming in the low resolution model is indicative of poorer ENSO fidelity so the use of low resolution was abandoned. Furthermore, we did not want any doubts to remain on the efficacy of the proposed methodology from using a coarser resolution of the model than routinely used in operational ENSO forecasting (Saha et al. 2006, 2010, Kirtman et al. 2013).

Using a combined metric following method 6 (see above) was found to yield a superior seasonal forecast than any of the other 5 methods of choosing the metric for analogue matching (not shown). This is because choosing the metric for matching based exclusively on each variable influences the complex evolution of ENSO in different ways. For instance, matching wind stress may prove to be most useful for prediction of onset or termination of events (Lengaigne et al. 2004, Harrison and Vecchi 1999) while SST is best for persisting anomalies. Therefore analogues based on the RMS differences of any single variable is found to find model states that are not optimal for initializing the CZ model for forecasting a given ENSO event. The method of combining the four predictor variables to obtain a single metric for analogue matching is explained in the next section.

4.2.2 Multiobjective Optimization

To improve on the simple averaging or weighting averaging of the RMS differences, Multi-Objective Optimization (MOO, Hwang and Masud 1979) was employed to match all four jointly. As highlighted by Bunge and Clarke (2014), the use one predictor variable may lead to

reduced predictability in different climate patterns. MOO is used to find the best compromise of two or more objective curves (J_i) simultaneously as shown in Figure B.11a. Each of the objective curves has a different time of optimal value that conflicts with the other curves. To select the best overall time MOO determines the optimal balance between these variables as shown by the highlighted time in Figure B.11a. On the other hand the choice of a bad optimal state is illustrated in Figure B.11b when MOO is not used. In this case we wish to find the optimal balance of the RMS difference curves for each predictor variable jointly to isolate the best analogue. Figure B.11b shows how variables might require scaling or weighting in order to have all variables considered in the joint optimization. Thus the scaling of each J_i is critical for the success of MOO.

The objective curves can be scaled and normalized in many different ways depending on the situation at hand. In some Engineering practices, each curve is normalized by the highest or lowest value. Alternatively, each objective curve can be normalized by the standard deviation of the curve. After this normalizing is performed, weighting is given to each objective, for instance, a higher weighting may be given to the variable of greater importance. For our RMS difference curves, it is important that the minimal values are all treated equally. First, each RMS difference curve (for each of the four variables) was first normalized. The histogram of the four RMS error curves is shown in Figure B.12. It is evident from this plot that for smaller changes in RMS differences there is relatively rapid change in the distribution of the zonal wind stress and thermocline depth increase than that of SST and thermocline depth tendency. Therefore performing MOO on these four objectives in a combined manner (following method 5) would give greater importance to zonal wind stress and thermocline depth implicitly. To correct this feature, a new objective distribution was produced by averaging the normalized distribution of all four variables. We then used this average distribution to remap the individual objective time curves, which is often referred as quantile mapping (Li et al. 2010, Bastola and Misra 2014, Maraun 2013). This ensured that the best fit for all four variables would be taken. Contrary to the weighting strategy used with CZ, an equal weighting was applied to CCSM4 because it was shown that weighting had a marginal effect on the outcome.

4.2.3 Sensitivity of Analogue Matching to Domain

Next, we attempt to limit the degrees of freedom as other studies have attempted. Previous studies found improvement in analogue forecasts by reducing the degrees of freedom

by filtering either spatially or temporally (Barnett and Preisendorfer 1978, Bergen and Harnack 1982, Gutzler and Shukla 1984). For this study the data has already been temporally reduced by using monthly mean values for all variables. The monthly average is also removed decreasing the spectra to timescales mostly longer than one year. In addition, we attempt to reduce the degrees of freedom by selecting a sub domain in which to find the RMS difference curves because the ENSO phenomenon is primarily occurring in the Tropical Pacific.

The method used to select the subdomain in CZ followed from the same technique used in the weighting, namely lag-correlation of the predictor variables with Niño3 SST. The region selected for the subdomain should in effect be the area most related to the evolution of ENSO. For each predictor variable, each month of the year, and each lead-lag time from 1 to 9 months a correlation table was created containing the lag correlation values for every grid point; an example for thermocline depth is shown in Figure B.13. In this way the region was limited to areas of positive correlation values. If the value was negative or small, that grid point would have no influence in the analogue matching. Where correlation values were greater, that variable at that location would have a larger influence on the matching taking place.

Being a global model, CCSM4 is not limited to a Tropical Pacific domain for analogue matching. The ability to use a larger area for analogue matching does not mean better results. For example, matching the SST in the Indian Ocean may improve ENSO forecast skill since there is a likely feedback at play. But matching variables in the Southern Ocean may in effect hamper results by arbitrarily excluding all cases out of sync with the southern annular mode, which has little bearing on the current state of ENSO.

To test the sensitivity of the domain choice for analogue matching in a CGCM, we chose three domains over which the RMS is calculated. For these three, the hindcast results from each are compared:

1. Global (60S-60N)
2. Tropical Pacific (20S-20N by 120E-80W) and
3. Subregions (individual sub regions for each variable) shown in Figure B.14.

In the final case, subregions were chosen based on the maxima of lag correlation between the predictor variables and Niño 3.4 SST. The extent of each subregion included lag correlation values greater than 1σ of the correlation values over the Tropical Pacific (20S-20N by 120E-80W). Figure B.14 illustrates the lag correlations used; the black boxes in each panel

indicate the areas of the individual subregions. Results show that limiting the domain to the Tropical Pacific produced the best analogues of the three domains tested.

4.2.4 Forecast Period and Length

As stated before the forecasts done with CZ were verified only within perfect model scenario and imperfect model scenario. For these experiments the period is chosen arbitrarily from the end of the library simulation onward for 100 years so that the library would be created without any influence from the later model data used to validate it.

For CCSM4 hindcast experiments, the observationally-based dataset chosen for validation and analogue matching is the GODAS data provided by the NOAA/OAR/ESRL PSD, Boulder, Colorado, USA, from their website at http://www.esrl.noaa.gov/psd/. This dataset contains monthly ocean analysis at 0.333x1.0 degree resolution for 1980 to present (Behringer and Xue 2004). The temporal range of GODAS is from 1980 to present day allowing for a total of 32 hindcast years on which to test the analogue method. For this study we take the surface potential temperature, zonal wind stress, and derive the thermocline depth by linear interpolating the depth of the 20°C isotherm of potential temperature. These three variables remain the same as those used with the CZ model experiments. Additionally, we derived the thermocline tendency using backwards time differencing with time interval of two months which provides a time-evolution aspect to the analogue matching. Thermocline depth and thermocline depth tendency along with surface temperature, and momentum flux were then used to match analogues with CCSM4. Figure B.15a shows the average SST for GODAS.

Next, we define the length of the forecast and the starting time of the initialization. To decide when the forecasts would end we examine GODAS for the month of peak variability. ENSO has a strong tendency to phase lock to the seasonal cycle. It was recognized that the value of the forecast will be greatly enhanced by forecasting for the month of peak variability at the longest lead possible (approximately 9 months).

In conclusion, we have selected the years 1980-2011 as the years for conducting the seasonal hindcast with CCSM4 with the proposed initialization scheme. The start time for initialization is March of each year and the resulting hindcasts will be verified after 9 months of integration in December of the same year.

4.3 Ensemble Hindcasts

Establishing the method of analogue matching as the initialization technique is only half way to generating a probabilistic forecast. In order to have a robust probabilistic measure of the forecast uncertainty, a large enough set of ensemble members per seasonal hindcast is desired (Palmer et al. 2004, Hagedorn et al. 2005, Doblas-Reyes et al. 2005, Palmer et al. 2008, Misra 2008, Kirtman et al. 2013). Traditionally, this is done by optimally perturbing the initial conditions in such a way that the most unstable modes triggered to grow within the lifespan of the seasonal forecast period. For the analogue method of initialization, a unique option exists: selecting multiple analogue matches. When dealing with analogues this method should be utilized in order to avoid sampling error. Typically the library length of the model states is limited and furthermore we have artificially reduced the degrees of freedom of a CGCM to identify suitable analogues that give rise to sampling errors.

The hindcasts generated contain a total of twelve ensemble members. These ensembles consist of four best initial states from CCSM found by analogue matching. Eight more initial states are obtained by modifying the first four with additive and subtractive singular vector perturbations for a total of twelve ensemble members per season. Therefore a total of 384 (=32 years x 12 ensemble members) seasonal hindcasts were performed with CCSM4 to examine the fidelity of the proposed initialization scheme.

4.3.1 Analogue Ensemble

Choosing multiple analogue matches to include in the ensemble reduces the sampling error and provides confidence limits to the resulting forecast (Zhu et al. 2013). The exact number of members is dependent on many factors but very rarely will the closest match be the best selection for the forecast at hand. It was found that for both CZ and CCSM4, the forecasts generated benefited from selecting multiple analogues of the initial state. In the CZ model, a detailed experiment was done to determine the optimal number of analogues to select (not shown). For this experiment the forecast results for 50 cases were compared including one to twenty analogue ensemble members. After the optimal number of analogues was found, the rest of the forecasts were generated.

For CCSM4, the first four ensemble members come from selecting the first through fourth closest analogues from the method described in the previous sub-section. These members

contain the most variation. It was found that using four initial states from the analogue library minimized the overall forecast error having less sampling error than using fewer matches and better agreement with observations than using more matching. After the four best analogues were taken, small perturbations were added to these to create a larger, more accurate ensemble than could be attained through only selecting analogues to fill the ensemble.

4.3.2 Empirical Singular Vector Ensemble

As pointed in Lorenz (1969), the fastest growing error measured by the largest singular value of a forward tangent model may vary by an order of magnitude due to variations in the initial state. Singular vectors are designed to maximize error growth over a finite time interval (Buizza and Palmer 1995). Therefore singular vectors are identified as the fastest growing perturbations in a dynamical system before non-linearity becomes important. The singular vector approach requires adjoint and linear tangent of the forecast model to determine these growing directions and the perturbations (Buizz and Palmer 1995, Molteni et al. 1996, Errico 1997). This approach although widely used in numerical weather prediction has also been used for ENSO prediction (Xue et al. 1997). However it has been done for simple models like the CZ model (Xue et al. 1997) or in intermediate coupled ocean-atmosphere models (Moore and Kleeman 1997). However to develop adjoint and linear tangent of CGCM's is not trivial and therefore this methodology is not preferred for ENSO prediction relative to the bred vector approach (Toth and Kalnay 1993, 1997, Cai et al. 2003, Vikhliaev et al. 2007).

In the context of this study it is important to examine how the initial states isolated by the analogue matching evolve over time when these singular vector perturbations are superposed on them. Do they evolve in a stable manner comparable to that in the library or do they diverge from the evolution in the library significantly? In order to circumvent the necessity to develop the linear tangent and its adjoint for CCSM4, we used the empirical method of generating singular vectors following (Kug et al. 2011). For accomplishing this task, the complex dynamical system is first broken down to a near linear approximation: the linear relationship (L) between SST (X) and thermocline (Y) at 6 months (τ) such as:

$$Y_n = X_{n+\tau} = LX_n + \varepsilon$$

Normally, this would require linearizing the model to find the relationship. The linear operator, conveniently, can be derived empirically from the same data used as the model library. First, the linear operator is estimated from the first 5 empirical orthogonal functions (EOF) of

thermocline depth and the 6-month lead of SST for the tropical Pacific (15S-15N; 120E-80W). This reduced system is then used to calculate the singular values; the largest singular value which is also the optimal perturbation mode is used to generate the perturbations. Figure B.16 shows the first and largest mode for SST and thermocline. Now the perturbations to be used have been found for SST and thermocline depth in the tropical Pacific region. To carry over these perturbations to the rest of the field variables, a simple regression is done between the EOFs of thermocline depth and all other variables. The magnitude of the perturbations is scaled to 10% of the variance of the thermocline depth anomalies in CCSM4. Finally, the four analogue matches are modified with the additive and subtractive perturbation giving a total of 12 ensemble members per hindcast.

CHAPTER FIVE

RESULTS

5.1 Why CCSM4?

Selecting the right CGCM for this work is vital to the overall success. While the methodology developed can, in theory, be applied to any model, its efficacy is however intimately dependent on the model fidelity of simulating ENSO. With typical initialization techniques, forecast error is a combination of model bias and initialization errors (and shock). With the analogue methodology for ocean initialization, there is a possibility of having no matching analogue with predetermined tolerance level because of poor ENSO simulation in the model. For this reason, it is best to choose a model that conforms to the natural ENSO variability as much as possible but at a minimum meets the following criteria:

1. Has relatively low errors in mean state over the equatorial Pacific Ocean

The model used should reasonably reproduce mean SST and thermocline in the tropical Pacific to assure that the model ENSO variations about a mean state are in some way comparable to the observations (Latif et al. 2001, Misra et al. 2008).

2. Has reasonable variability of eastern equatorial Pacific SST at ENSO time scales

Ideally the simulated ENSO should be of similar strength and have a similar return period as observations. A model with weak ENSO variability may lack analogues for strong peak events. Bias in return period might produce systematic bias in forecasts such as progressing from one event to another too quickly. Another issue is that the long model simulations should have reasonable diversity (and aperiodicity) of ENSO to be able to sample the initial state analogues for a given ENSO event.

The above two criterion seem rather abstract, when 12 different features critical to ENSO simulation were identified in Misra et al. (2008). In fact several studies (Federov and Philander 2001, Wittenberg et al. 2006) suggest that fidelity of the mean state is critical for successful ENSO simulation. However, the manifestation of these observed surface ocean features in the equatorial Pacific is also a reflection of the sub-surface ocean evolution and potential local and remote teleconnections in the atmosphere.

One of the tasks of this research work was to develop an initialization strategy that could be easily adopted by any of the CMIP5 suite of models for seasonal ENSO predictability studies. Therefore before demonstrating the proposed initialization strategy, all the available CMIP5 models in 2012 were analyzed for the fidelity of their ENSO simulation (Michael et al. 2013). A brief summary of the published work (Michael et al. 2013, see Appendix A) is presented here.

The mean observed SST field (Figure B.17a) shows the equatorial cold tongue off the coast of Peru and the tropical western Pacific warm pool. The replication of this equatorial zonal temperature gradient is important for the coupled feedbacks of the ENSO variations (Clarke 2008). The CMIP5 model SST error fields (Figure B.17b-r) exhibit a rather pathological bias in three areas: a cold bias over the cold tongue in the equatorial Pacific Ocean, a warm bias in the eastern oceans of the subtropical region in the southeastern Pacific region, and a cold bias in the western portion of the subtropical Pacific Ocean. The annual mean SST errors averaged over these three regions are indicated in Table B.4. All of these errors were quite prevalent in the CMIP3 models (AchutaRao and Sperber 2006; Capotondi et al. 2006), and, unfortunately, these errors continue to persist in the CMIP5 suite of models. The most extreme cold bias in the equatorial Pacific is seen in the CSRIO-Mk3.6 (Figure B.17f) and it is the only model to have a cold bias over nearly the entire tropical Pacific basin. However, GISS-E2-H (Figure B.17j) and GISS-E2-R (Figure B.17k) display a widespread warm bias across the eastern oceans, covering even the cold tongue region. In GFDL-CM3 (Figure B.17g), MIROC5 (Figure B.17o), and NorESM1-M (Figure B.17r) the errors at the equator and over the Peruvian coast are comparatively far less but they show significant cold biases in northwestern Pacific. Table B.4 and Figure B.17c indicate that CanESM2 has the least bias in these three regions. For the purpose of selecting analogues, errors in the cold tongue, the warm pool, and the stratocumulus regions of the eastern subtropical oceans would contribute to poor analogues and ENSO forecasts. On a more encouraging note, nine of the seventeen models listed in Table B.4 have less than 1°C errors for all three regions (CanESM2, CCSM4, CNRM-CM5, GFDL-CM3, GFDL-ESM2M, HadGEM2-ES, INM-CM4, MIROC5, and NorESM1-M). Similarly, Table B.5 shows a comparison of thermocline depths in the equatorial Pacific for each model. The maximum and minimum thermocline depths in Table B.5 correspond roughly to the western and eastern part of the equatorial Pacific Ocean. In Table B.5 we note that CCSM4, CSIRO-Mk3.6, IPSL-CM5A-LR, and NorESM1-M are the four models with the closest maximum depths to the

observations in the western Pacific while CanESM2, CCSM4, MRI-CGCM3, and NorESM1-M are the closest match in minimum depth to the observations in the east.

Figure B.18 shows the spectra of Niño 3.4 SST for 17 CMIP5 models in comparison to the observed (ERSSTv3b) spectra in black. It is clear from this figure that there are considerable differences among the most recent generation of climate models with respect to ENSO simulation. For the purposes of this study a suitable model should have a peak nearly matching the observed ENSO peak around the 4-year time scale. Additionally, models with very low peaks should be excluded as the ENSO variability would be too weak in those models. The three panels in Figure B.18 have classified the 17 CMIP5 models with the ones closest to or weaker than the observed spectrum (Figure B.18a), relatively stronger (Figure B.18b), and comparatively much stronger than the observed spectrum (Figure B.18c).

In the observational study of Zelle et al. (2004) it is shown that the depth of the thermocline anomalies are closely related to the overlying (Niño3) SST anomalies contemporaneously and at various lead time (with former leading the latter). Figure B.19 is the elliptical fit of the scatter between the anomalies of the thermocline depth with the SST in the Niño3 region. In observations (black) this scatter has a linear spread with positive (negative) SST anomalies increasing with positive (negative) thermocline depth anomalies. The observed scatter has a correlation of 0.78. Several CMIP5 models show a comparable correlation between the two variables as in the observations, but very few are able to get the observed range of Niño3 SST anomalies and the associated spread of thermocline variations for a given Niño3 SST anomalies (e.g., CanESM2, CCSM4, NorESM1-M).

CCSM4 as the CGCM for this study works well for several reasons. As shown above, CCSM4 is in the top tier of reproducing the observed mean state among CMIP5 models. Because the peak return period of ENSO is roughly the same in CCSM4 as observed, there is a better probability that matching analogues will succeed when using CCSM4 as compared to other CMIP5 models—models that may have spurious peaks at higher or lower frequencies or that may lack strong variability altogether. The availability of CCSM4 source code also makes for an ideal model candidate. Although CCSM4 is not the only suitable model, the work proposed is best limited to just one model for practical considerations. For a more in-depth examination of ENSO simulation CMIP5 models see, Kim and Yu 2012, Zhang and Jin 2012,

Michael et al. 2013; Bellenger et al. 2013. For a more detailed analysis of the ENSO simulation in CCSM4 the readers are directed to Neale et al. 2008, Deser et al. 2012, and Capotondi 2013.

5.2 Results with Cane-Zebiak Toy Model

The CZ model owing to its simplicity was found convenient to conduct preliminary experiments to demonstrate proof of concept. In this regard we used the CZ model to develop a simple analogue matching algorithm. We then tested this algorithm to its sensitivity to the library length of the model state in a perfect model approach followed by examining the efficacy of choosing more than one analogue for a given observed state as a basis for generating ensemble members. As pointed earlier, the perfect model approach although theoretical, allows us to gauge the role of model errors in our methodology when it is compared with the imperfect model experiments. The algorithm developed for this perfect model experiment is specially tailored for the CZ model. However, the steps followed in the procedure are the same for any model. In a perfect model approach, the CZ model is integrated for a multi-millennial period from which the model states are identified for analogue matching. The standard model is then integrated further for another 100 years, which is then utilized as a proxy for target observed ENSO forecasts and their corresponding initial states 9 months before. The following subsections will describe the results of the CZ which include the influence of multi-variable analogue approach, the effects of library length, and the perfect and imperfect model cases.

5.2.1 Development of an Efficient Analogue Matching Algorithm

Before analogue matching can occur, the salient variables for ENSO prediction must be isolated. From ENSO theory, these fields are SST, thermocline depth, and zonal wind stress over the equatorial Pacific Ocean. For each of the field variables, a lag-correlation with the 1-to-9-month lag Niño3 SST is calculated for each month of the year over the tropical Pacific. The result is 108 correlation maps for each variable. The importance of using separate correlation maps for each month can be seen most prominently in the correlation values for SST, which varies seasonally; likewise, the importance of repeating the correlation at each forecast range is of known importance for correlations with thermocline depth (Clarke 2008). From the correlation maps, the region of importance can be chosen on the basis of the zonal and meridional range of significant correlation values. Consequently a mean squared difference

(MSD) computed over this region between the model state and observed initial state is used as a basis for determining the best fit analogue.

Area variables identified from the correlation maps are used for analogue matching. For the CZ model, the highest lag-correlated field to the Niño3 SST is SST within (150W−90W, 5S−5N). The correlation with SST is highest for lags of 1 to 4 months; at greater lags, correlations vary significantly by season. The thermocline used for matching extends farther east (150W−80W, 5S−5N) than the SST area. The rationale behind this choice can be seen in Figure B.13, where there are heightened correlation values along the eastern boundary. The thermocline variations are most important for long-term forecasts of the ENSO SST anomalies owing to their longer inertial timescales. Zonal wind stress proved to be the variable least correlated with Niño 3 SST, but remained important in efficient analogue matching. The region chosen for zonal wind stress was (170E−150W, 20S−EQ).

We examine matching methods using the three variables described in the previous paragraph. First, the MSD of all model states in the library and target observed initial state are computed for each variable. Second, the modeled state with the lowest MSD value is used as the initial state for the forecast. This is done following eight different formulas: three based on the MSD of one variable (SST, thermocline, and zonal wind stress), three based on MSD of paired variables (SST and thermocline, etc), one based on the MSD with all three variables, and one based on the MSD with all three variables weighted proportionally to the correlation values described earlier in this section. The results for a 9-month forecast show that no single variable or variable pair exceeds the skill of using all three variables and that weighting the variables improves the skill only slightly Figure B.20. It is also seen that persisting the anomalies for 9 months yields a very poor forecast relative to the other analogue initialization experiments with the CZ model.

5.2.2 Examining Library Length and Ensembles

As mentioned before earlier studies found that the analogue approach was not viable because of the short record of model states usable for initialization. With today's computing power we were able to produce 5000 years of model data from the CZ model to use as the analogue library within a short period of time. This valuable tool (CZ model) allows us to measure the improvement of forecast skill with increased library length. One might intuitively

think that increasing the library would follow the law of diminishing returns. As shown in Figure B.21, this is in fact the case: the forecast error follows an exponential curve resulting in no significant improvement in skill beyond a library size of ~2000 years. This curve is highly dependent on the number of degrees of freedom in the model system and therefore will not be applicable to a GCM like CCSM4. Figure B.21 does however indicate that the effort of producing very long library length for CCSM4 could be less productive beyond a certain integration period. It may also be noted that many of the CMIP5 models produce ENSO, which is far more periodic than observations as noted in their sharp peak of the Niño3.4 spectrum. Therefore, the diversity of ENSO cannot be improved in such models by increasing the length of the library, which is one of the main objectives of having large enough library to pick analogues that match with the observed initial state.

Another important consideration is the number of analogues chosen for each forecast. Since we cannot choose analogues because of imperfections in the model, and analogue algorithm, it is prudent to choose more than one analogue for a given observed state. If too few analogues are chosen the forecast skill can suffer from sampling error and if too many are chosen then the likelihood of the consequent forecast skill degrading is significantly high from poorly matching analogues. To answer the question of how many analogues to include, 48 forecasts were produced for ensemble sizes ranging from 2 to 20 members. If the first analogue chosen was suboptimum as shown in Figure B.22, the forecast skill could be improved by adding additional analogues to an ensemble size of up to 10 members in many cases. However, if the first analogue was good, adding ensemble members decreased the forecast skill linearly. The average number of ensemble members with the highest skill is four (Figure B.22) for a library size of 5000 years with CZ. Overall, the number of ensembles chosen impacts the results less than either the library size or the variable combination.

5.2.3 Testing the Analogue Method with an Imperfect Model

The experiments discussed up to this point focused on forecasting a system with a perfect model. This however, is far from the truth where models are highly imperfect. For this reason, it was paramount that we attempt to use the analogue method on forecasting a system with an imperfect model.

To perform this experiment, the CZ model was used as before to select analogues, but the model used for the forecast verification was a modified version of CZ model. By modifying the two parameters of the CZ (α, the parameter controlling the strength of the SST-related component of atmospheric heating and the decay time of the ocean), we were able to produce a spectrum more closely resembling the observed Niño 3 SST variance (Figure B.9). Although it may appear that we have improved the CZ model in some way, the apparent improvement should not be judged by this figure alone. In fact, the values of the new parameters are not realistic compared to the ocean and atmosphere. In the modified CZ model, the atmosphere feedback is stronger and the ocean is adjusting more rapidly. The increased stochastic forcing of the atmosphere on the ocean results in a broader peak in the spectrum. We would expect that such a system would be harder to forecast than the original CZ model given the wider range of possible states.

The limitation of the imperfect model is dependent on the differences of the model attractor and the attractor of the forecasted system. The more the two attractors overlap, the more likely the model is to have a suitable analogue for the forecast. The distribution of the modeled states for CZ can be compared to the modified CZ model in Figures B.23a and B.23b, respectively. In Figure B.23a we can see that the density of points in the modified model is much higher for positive SST anomalies (~ 2.5°C) and negative thermocline anomalies (~ -20m). The maximum thermocline anomalies are also limited to about 20m in the standard CZ model but reach 50m in the modified CZ model. From this we expect that analogue forecasts for states not captured by CZ will contain the most error, which is the result shown in Figure B.23c.

The imperfect forecast model experiment highlights an important shortcoming of the analogue method of initialization: the initial state can only be good as the model that produced it. In addition to library size, having a good model is necessary to produce skilled seasonal ENSO forecasts.

5.3 Results from CCSM4

The CCSM4 results shown here were found using a 600-year library. As was shown in the section above, upwards of 2000 years was needed before forecast skill stabilized in the CZ model within the perfect model experiment. If the same relationship holds true, then at least as many years, if not more noting the 10^{30} years cited by Van den Dool (1994), would be needed of a CGCM library for best results. However, we are not able to obtain a record of restart data that

long for CCSM4 and can only speculate on the improvement a longer library would yield. It may also be noted that CCSM4 also suffers with a rather periodic ENSO (as displayed by the sharpness in the Niño3.4 SST spectrum, Michael et al. 2013) and therefore it is plausible that very large library lengths may be unnecessary.

The seasonal hindcasts of 9-month period were conducted with CCSM4 initialized in beginning of March from 1980 to 2012. We deliberately chose this initialization time as it is in the beginning of the month of the spring predictability barrier (Webster and Yang 1992). This barrier results usually in poor model fidelity with weakening SST gradients over the equatorial Pacific Ocean and persistence of observed SST anomalies are extremely poor alternative for seasonal ENSO forecasts. Therefore models initialized in this month are being challenged to produce seasonal forecasts in the near absence of any skill from observed persistence. Furthermore at 9 months from March, ENSO usually exhibits peak variability, which is often referred as the seasonal phase locking feature of ENSO. Therefore anticipating such large Niño3.4 SST anomalies with sufficient lead time provide ample opportunities to respond effectively and can render the seasonal forecast for very useful applications.

5.3.1 Multiobjective Optimization

Nine-month hindcasts for 32 seasons from 1980 to 2012 were generated from analogues using the four predictor variables (SST, thermocline depth, thermocline depth tendency, and zonal wind stress) to match CCSM4 to observationally based states in GODAS. Each predictor variable was used separately to match analogues and then jointly using MOO. For lead times less than 5 months, SST matching provided the best analogue based forecasts (not shown); for lead times longer than 5 months, thermocline depth provided the best matching (not shown). Overall, the matching of the combined four variables using the MOO preformed best. It should be noted that while zonal wind stress matched analogues never out preformed any other method, the combined matching method improved with the addition of zonal wind stress (not shown).

While we chose to ignore weighting, future studies may find value in weighting the predictor variables depending on the length of the forecast. For short lead times, the initial SSTs can be persistent enough to require a higher emphasis in the analogue matching algorithm. For lags greater than 5 months, the slower ocean dynamics of the thermocline are most important and should be given higher weighting.

5.3.2 Sensitivity of Domain Choice

To further improve the forecast skill using analogue matching, a simple test to reduce the degrees of freedom was performed under a perfect model approach. The choice to limit the domain size stemmed from two main issues. First, RMS errors unrelated to ENSO may influence the analogue matching in such a way as to exclude possible best matches for ENSO phase and strength. Second, the mean state of CCSM4, contained biases which may reflect in disproportionate RMS errors based on factors not relevant to ENSO.

Three different domains were used as the limits for analogue matching as described previously. Comparing the Niño3.4 SST anomaly errors from hindcasts generated with the three different analogue domains to persistence forecasts show that all three cases produced forecasts with significantly lower errors. Figure B.24 shows that for the 32 years over which the analogue hindcasts were produced, the Tropical Pacific domain worked best for most lead times. Although the global domain was able to match the tropical Pacific at longer lead times, this is possibly because the errors are reaching saturation limits. The forecast errors from analogues with the metric to choose analogues encompassing global versus Tropical Pacific domain were not significantly different at the 90% confidence limit and so the Tropical Pacific domain was used henceforth.

Domain choice for selecting analogues proved to be more complex than anticipated. During the initial domain sensitivity test, a perfect model approach was taken. In the perfect model case, the global domain performed best at shorter lead times in contrast to the hindcast experiments, which showed tropical Pacific domain as the clear leader at 5 months. This is possibly a result of the short-term SST being forced by global factors not directly related to ENSO and also because the global biases in CCSM4 caused a deterioration of the analogue quality when matched globally (see Figure B.17).

5.3.3 The Fidelity of the Mean of CCSM4 Seasonal Hindcasts

Several general error metrics were generated to examine the overall climatology of the hindcasts. These include, the seasonal cycle of standard deviation of SST, the composite SST fields for each phase of ENSO, and a time series plot of the Niño3.4 SST. It may be noted that these fields were generated by using all 12 ensemble members per seasonal hindcast (which includes ensemble members generated by the 4 analogues and 8 more perturbed by empirical

singular vectors). An analysis of the seasonal hindcast skill will follow in the subsequent sub-section.

Figure B.25 shows the seasonal cycle of standard deviation of SSTs for GODAS and the 12-member ensemble average of the hindcast results. In this figure it is apparent that the peak variability in the CCSM4 hindcast (Figure B.25b) occurs east of the dateline as observed (Figure B.25a). Furthermore, this variability propagates westward (Figure B.25) as observations indicate in Figure B.25a. However these anomalies extend further west of the dateline in the CCSM4 hindcast than the observations. But more gravely, the CCSM4 hindcasts appear to show a single peak of equatorial Pacific SST variability in late Spring season unlike observations, which shows the variations in the SST over equatorial Pacific to peak in Spring and Fall in the far eastern Pacific Ocean (Niño 1+2 longitudes).

The composite global SSTAs from GODAS for warm, neutral, and cold events are shown in Figure B.26. Here we see that the average warm and cold events are usually on the order of 1°C. Figure B.27 shows the same plot but for the composite of the CCSM4 ensemble hindcasts. Here we can see that overall, the hindcast does predict warming and cooling of SSTs in the same region as observed but, on average, has slightly reduced magnitudes (Figure B.27a, Figure B.27c). There is also a notable reduction of SSTA in the coastal region of Peru in the CCSM4 hindcasts. To gain a better perspective of these biases, Figure B.28 shows the difference of Figures B.27 and B.26. Here we can see a reduction of the SSTAs during warm ENSO events of around 0.5°C whereas the cold ENSO events warmer compared to GODAS.

We examined the systematic errors of the ENSO SST anomalies from the 600 year simulation of the CCSM4 (Figure B.29. in comparison to Figure B.28). The systematic errors of the ENSO composites in 600 year simulation are quite similar with warm events being colder and cold events being warmer than GODAS. In other words, the similarity of the errors in the hindcasts and the long-term simulation illustrate the reduction of the initialization errors. It may be noted that difference in the magnitude of the errors in Figures B.28 and B.29 are a result of the difference in the sample sizes in creating the composites.

5.3.4 Hindcast Verification

The ESV is used to increase the total set of ensemble members to obtain a robust estimate of the uncertainty of the seasonal hindcast. In Figure B.30 we show the forecast plume of the 32 seasonal hindcasts with all 12 ensemble members. The 4-member ensemble mean from the analogue initial conditions and the ensemble mean from all 12 members are highlighted by bold lines. This figure shows that the ensemble spread is uniformly large across all seasons. It should be noted that there is no evidence to be gathered from these plumes that initial conditions corresponding to analogues 3 and 4 produce consistently poorer seasonal hindcasts compared to analogues 1 and 2, which have objectively less RMS differences with the target initial state. It should be remembered that these multiple analogues were chosen to account for the sampling error from the limited library length of the historical simulation of the CCSM4 and also to account for the uncertainty in the choice of the metric. In other words, the quantitative differences in the analogues does not manifest in systematic differences in the seasonal hindcast. But the diversity or the ensemble spread of the seasonal hindcast does suggest the manifestation of the initial condition uncertainty.

In Figure B.30 it is seen that in some instances the initial condition of the Niño3.4 SSTA seems to be significantly different from the observed initial condition (e.g. 1980, 2004), This so happens because we use a multi-variate metric to choose the analogues, which is dependent on the combined RMS differences of the other three variables (thermocline depth, its tendency and zonal wind stress in the Niño4 region) and is not solely based on the Niño3.4 SSTA. Another important feature to be noted from Figure B.30 is that the ensemble-mean of the 4 analogues and 12-member mean are nearly consistent with each other for all seasons, which suggests that the analogues are isolating initial conditions that are quite robust and not sensitive to any growing unstable modes perturbed by the empirical singular vectors.

The evolution of the equatorial Pacific SSTA in the ensemble mean of the hindcast and the corresponding observations for 1997, 1998, and 2007 are shown. The 1997 El Niño event was considered to be the strongest such event of the 20th century (McPhaden 1999). It was one of the poorly forecasted events by the then operational dynamical climate models initiated in March 1997 (Barnston et al. 1999). Although majority of these operational models did pick the evolving El Niño event, they failed to capture its unprecedented magnitude at 9-month lead. Nevertheless, the same models initiated in the summer of 1997 were able to forecast the event far more

accurately (Barnston et al. 1999). In contrast, the CCSM4 hindcasts in this study did a reasonable job of picking the magnitude and phase of the El Niño event in 1997 reasonably well (Figure B.31a and Figure B.33) in comparison to the observations. Similarly the evolution of the 1998 La Niña event is shown in Figure B.31b and Figure B.34. Here again the CCSM4 hindcast does a reasonable job of capturing the event at 9 month lead. In contrast the 2007 event was poorly forecasted by the CCSM4 (Figure B.32 and Figure B.35) evolved the anomalies contrary to observations.

There are instances when some of the ensemble members diverge significantly from the rest of the members by the culmination of the seasonal hindcast. For example in Figure B.32, for 2007 seasonal hindcast some of the ensemble members evolved ENSO significantly differently compared to the rest of the members. On the other hand 1997 El Niño event and 1998 La Niña event did not exhibit that large of a divergence amongst the ensemble members of the seasonal hindcast (Figure B.31). In fact the ensemble mean forecast of 2007 was extremely poor, which ran contrary to the evolving La Niña anomalies in the Niño3.4 region. But the forecasts of 1997 and 1998 events were far more reasonable where the ensemble spread was far less. Figure B.36 shows the scatter between the standard deviation of the ensemble spread and the ensemble mean Niño3.4 SST error. Despite the anecdotal cases explained earlier, the relationship between ensemble spread and the mean Niño3.4 SST errors is rather poor to suggest that models are more uncertain when there is larger forecast error. In other words, contrary to the desired feature the CCSM4 hindcasts do not display forecast uncertainty arising from initial condition uncertainty as a good proxy for the forecast skill.

Figure B.37 shows the scatter between the multi-objective optimization criteria used to choose the analogues with the 9-month lead Niño3.4 SST errors. This figure clearly suggests that the hindcast skill is independent of the quality of the analogue. In fact the correlation of the two variables in Figure B.37 is merely 0.2. This points to the fact that the analogues have been well chosen in that there is consistency in their quality since it is based on an objective criterion. There are several possible answers to explain model hindcast errors (like those in 2007), which are as follows:

a) Events like the 2007 La Niña event represent a limit of the ENSO predictability at least in CCSM4. In other words it is possible that such ENSO events are not

initial condition dependent and could be generated from internal dynamics of the coupled ocean atmosphere system.

b) The CCSM4 does not display sufficient ENSO diversity as displayed by the sharp peak in the Niño3.4 SST spectrum. Therefore it is possible for CCSM4 not to systematically capture some ENSO events or systematically predict a periodic ENSO event. This point is further illustrated with the distribution of the Niño3.4 SSTA as a function of the thermocline depth anomalies from GODAS (Figure B.38a), and the CCSM4 library (Figure B.38b). The same scatter is shown for the 9-month hindcast in Figure B.38c. The forecast error of the Niño3.4 SSTA from the 32 seasonal hindcasts is also shown in Figure B.38d. Unlike the CZ imperfect results we see the model errors seem be larger when the anomalies are nearer to zero but still rather uniform over the whole plane which 1) exposes the poor skill at forecasting neutral events and 2) shows that the model is well suited for use in the analogue method of initialization.

c) The library length of 600 years could be short and it is possible that we have not sampled the initial states sufficiently to capture the subsequent evolution of such ENSO events as 2004, 2007.

d) CCSM4 like most other CMIP5 models displays a lack of asymmetry between the warm and cold ENSO events.

We leave further investigation of the above points for future work.

Several tests were conducted to measure the skill of the seasonal hindcasts objectively. Hindcast verification tests undertaken include RMS error of Niño3.4 at 1 to 9 months lead time, correlation of Niño 3.4 SST at 1 to 9 months lead time, and Relative Operative Characteristic (ROC) Curve analysis. The latter is a measure of the probabilistic skill as opposed to correlations and RMS error which are deterministic.

Figure B.38 shows the average RMS error of the Niño3.4 SST from the 372 hindcasts (=32 years x 12 ensemble members) initialized in March as a function of lead time for the ensemble mean and the corresponding range of the RMSE produced by the ensemble members. The corresponding RMS error of a persistent SST forecast is also shown (Figure B.38) for

comparison in red. Except at the initial few months of the hindcast when the persistence skill is very high, the seasonal hindcast from CCSM4 outperforms persistence in terms of this metric.

Considering that forecasts initialized in March provide more skill than other seasons is somewhat counter intuitive. The reason for this improvement is likely because the subsurface ocean heat content is able to predict what the surface SSTs are unable to show: the underlying phase and strength of ENSO. Figure B.40 shows the correlation of the Niño 3.4 SST at 1 to 9 month lead-time. Here we can see that persistence drops off rapidly after outperforming the seasonal hindcasts from CCSM4 in the first couple of months of the hindcast. It is important to note that the CCSM4 hindcast results maintain a correlation of above 0.5 for all 9 months starting in March and the persistence precipitously drops below zero correlation value after August.

The probabilistic skill of the ensemble hindcasts from 1980-2012 for 9-month lead was assessed by plotting the relative operating characteristic (ROC) curve following Green and Swets (1966). Figure B.41 shows the ROC curves for warm, neutral, and cold ENSO events. Both warm and cold events are predicted reasonably well with an area under the curve greater than 0.5, which suggests that it is better than climatology. The accuracy of warm event prediction is 69% with specificity of 76%; the area under the fitted curve is 0.711. For cold events, the analogue ensemble accuracy is 75% with a specificity of 90% and an area of 0.837.

Next, we examined the Brier skill scores for the 12-member ensemble hindcasts. With a Brier score of 0 being the best and a score of 1 being the worst, we can see that the hindcasts do have skill for both warm and cold events. The Brier Score 0.116 for warm events and 0.200 for cold events indicates that the warm events are being predicted more accurately than cold events, contrary to that found in the probabilistic skill of the ROC curves. This indicates that the model skill in predicting the correct phase of ENSO and the correct magnitude of an ENSO event are distinct.

5.3.5 Teleconnections

ENSO variations impact local and remote climate through atmospheric teleconnections (Ropelewski and Halpert 1987, Halpert and Ropelewski 1992). These atmospheric teleconnections provide a basis for seasonal climate prediction at regional scales. Figure B.42 shows the regression of the October-November-December (OND) averaged Niño3.4 SST on rainfall at zero lag from the observations and ensemble mean of the seasonal hindcast

respectively. It may be noted that the OND season is a 7-8-9 month lag season of the hindcasts initiated in March of the CCSM4 hindcasts. The precipitation observations used in Figure B.42a are from Climate Prediction Center Merged Precipitation Analysis (CMAP; Xie and Arkin 1997). The shift in the equatorial rainfall anomalies from the west Pacific to central Pacific is consistent with notion of the shift in the atmospheric Walker circulation anomalies associated with ENSO variability. Similarly, the regression over southern California and southeastern US Figure B.42a is consistent with the well-known teleconnection pattern of warm (cold) ENSO events associated with wet (dry) winters over these regions. However, like most other climate models, the rainfall anomalies in the CCSM4 hindcasts in Tropical Pacific display a stronger north-south gradient (Figure B.42b) than the observed east-west gradient suggesting that the Hadley cell atmospheric circulation anomalies in the hindcasts are over represented compared to the Walker circulation anomalies.

Similarly, the regression of the OND Niño3.4 SST anomalies with contemporaneous global SST anomalies from observations and the CCSM4 hindcasts are shown in Figure B.43a and Figure B.43b respectively. The well-known horse shoe pattern of ENSO variations centered over the western Pacific is reasonably well picked by the CCSM4 hindcast (Figure B.43a) in comparison to the observations (Figure B.43b). In addition the observed co-variation of the SSTA in the western Indian Ocean and Niño3.4 SSTA (Krishnamurthy and Kirtman 2003) is also captured in the hindcasts (Figure B.43b). However, the equatorial Pacific anomalies in the hindcast (Figure B.43b) unlike the observations (Figure B.43a) indicate that they extend too far to the west beyond the dateline. Furthermore the associated SSTA in the tropical Atlantic Ocean are far more uniform than what observations indicate.

The regressions of the Niño3.4 SSTA on the 500hPa geopotential heights from the hindcasts and observations are shown in Figure5.28a and Figure B.44b respectively. In a typical El Niño event there is tropical wide warming of the troposphere, and from the shift in the deep convection to central Pacific, Rossby waves get propagated along the great circle over North America (Wallace and Gutzler 1981). These features are well captured in the hindcasts (Figure B.44b) in comparison to the observations (Figure B.44a).

CHAPTER SIX

CONCLUSION

As stated in the beginning of this dissertation, the primary motivation of conducting this research was to develop a simple initialization technique that would minimize the initialization shock and can be easily adopted for any of the existing CGCMs to conduct seasonal ENSO predictability studies. ENSO is the largest natural variability that influences the global climate variations and forms the basis for seasonal climate prediction.

6.1 Summary of Work

The proposed initialization technique is based on an analogue approach, which seeks the best match from a relatively long library of historical simulations to the target initial state. The metrics for the analogue were tested from the most simple to a multi-variate (multiobjective) metric that includes SST, zonal wind stress, thermocline depth and the thermocline depth tendency over the Tropical Pacific region. It was shown that this multi approach for choosing analogues was better than the choice of any single variable for picking the best analogue for ENSO predictability.

This research work began with proof of concept of the proposed ocean initialization technique demonstrated in the Cane-Zebiak model, a two layer reduced gravity model coupled to a statistical atmosphere. The efficacy of the approach was tested both in perfect and an imperfect model set up of the Cane-Zebiak model. The results from using this toy model can be summarized as follows:

a) A sufficiently long library length is needed for the analogue method to be successful. The exact length in dependent on the degrees of freedom but also the tolerance for errors in the initial conditions. In the perfect model scenario, as the library length approaches infinity the initial errors tend toward zero.
b) In the modified model scenario, the initial error is a function of (a) library length (as in the perfect model scenario) and (b) the ratio of overlap of two models' attractors to total attractor. Thus, if the attractors for the two models are different in any way, the initial errors will not approach zero as the library length in increased; if the attractors are substantially different or do not overlap, it is

possible that increasing the library length will not lead to initial errors within the tolerance limits.

c) Increasing the library length follows the law of diminishing returns. By increasing the library by 250 years, greater improvement was seen when starting with a shorter library length than when starting with a longer library length.
d) With limited library length, sampling error can result in suboptimum choice for analogue matching. This occurred often even with a library of 5000 years in the perfect model scenario.
e) Adding additional analogue matches to an ensemble reduced sampling error but adding too many ensemble members introduced initial errors. Generating an ensemble of the top four analogue matches minimized the errors from sampling and departure from observations.
f) These toy model experiments also demonstrated the value of multivariate matching which lead to the development of the multiobjective optimization technique used in the CGCM experiments.

Before proceeding to demonstrate the proposed initialization technique on a CGCM, a comprehensive analysis of the CMIP5 models was conducted. This follows from a desired goal of this study to provide a simple enough initialization technique so that a majority of the CMIP5 models, which represent the state-of-the-art in climate modeling, can be deployed for routine ENSO seasonal predictions. Furthermore, the effectiveness of the proposed initialization procedure is critically dependent on the fidelity of the ENSO simulation in the CGCM. Therefore it is important to pick a CGCM that has a reasonable simulation of ENSO. Our ENSO diagnostics on the CMIP5 historical simulations show that some of the old biases from the CMIP3 legacy like the westward extension of the variance beyond the dateline over the equatorial Pacific, relatively little asymmetry between the warm and cold ENSO events, continue to persist. However the CCSM4 model displays a Niño3.4 SST spectrum with the peak frequency generally within the 4-7 year range of the observations as do many other CMIP5 models, which is a significant improvement from the previous generation of models (CMIP3). However the Niño3.4 SST spectrum in majority of these models displays a spectral peak, which is either too narrow or too broad with power far greater than observations suggest. Furthermore it

is shown in the CCSM4 historical simulation that the seasonal phase locking feature is also reasonably well captured as are the teleconnections of the 200hPa geopotential heights and rainfall. In addition given the user support and the free availability of the source code, CCSM4 was our choice of CGCM for this study.

An extensive set of seasonal hindcasts is conducted with the proposed initialization technique of matching analogues with the CCSM4. To provide a robust estimate of the forecast uncertainty, the 4 chosen best analogues for initial state of the ocean are perturbed with empirical singular vectors to generate a total of 12 ensemble members per season. The seasonal hindcasts were initiated in the beginning of March and extended to end of December of the year for a 9-month hindcast. The seasonal hindcasts were conducted for the period of 1982-2012.

The analysis of the seasonal hindcasts of the CCSM4 reveal that they display skill that is superior to persistence up to 9 month lead time over the Niño3.4 region. Both the deterministic and probabilistic skill amply demonstrates useful skill up to 9 months over the equatorial Pacific and the related global teleconnections are also reasonably well captured. It may be however noted that the persistence skill is superior to the ensemble mean of the seasonal hindcasts for the first 3 months of the forecast period. But after this period the persistence skill drops precipitously while the hindcast skill reduces comparatively gradually.

It is shown that the differences in the climatology of the hindcasts and the simulations are small. The relationship of the variability of the Niño3.4 SST anomalies and the thermocline depth is maintained in the hindcasts from the long-term simulations. Both these features indicate that the initialization shock has been reduced appreciably in the seasonal hindcasts. However the errors in the initialization is dependent on the model fidelity as the quality of the analogues will be determined by it.

It is therefore demonstrated that the proposed analogue approach with the superposed perturbations from empirical singular vectors is a viable initialization technique for ENSO predictability. Furthermore it is easily adaptable to any CGCM given its simplicity in implementation. It should be however recognized that the approach requires a long historical simulation with a frozen version of the CGCM. Despite this limitation, the CMIP5 community will have a lot to gain from understanding the errors in seasonal ENSO predictability just as they can contribute usefully to grand multi-model seasonal ENSO prediction to provide robust estimates of forecast uncertainty.

6.2 Future Work

The underlying simple framework for initializing CGCMs for ENSO predictability is an excellent beginning. The work can be further extended by:

a) Comparing the seasonal hindcasts of this study with those where CCSM4 is initialized with initial conditions from a data assimilation product (e.g. GODAS). This comparison will provide some perspective on the amelioration of the initialization shock and any benefits to the seasonal prediction skill from the analogue and empirical singular vector approach.
b) Understand the sensitivity of the library length on the choice of the analogues and the consequent seasonal prediction skill.
c) Adopt the proposed initialization approach on many other CMIP5 models to compare and contrast their seasonal prediction skill results. This study can reveal the sensitivity of the model fidelity on the choice of the analogues and its consequent impact on the seasonal prediction skill. Furthermore the impact of resolution of model uncertainty on seasonal forecast uncertainty can also be studied with greater depth.
d) Examine the bimodal nature of the ensemble forecasts with respect to hard-to-predict forecast years. Understanding such behavior of the model could potentially lead to better understanding of the ENSO system as a whole.

APPENDIX A

THE EL NIÑO SOUTHERN OSCILLATION IN THE HISTORICAL CENTENNIAL INTEGRATIONS OF THE NEW GENERATION OF CLIMATE MODELS

J. –P. Michael[1]
Vasubandhu Misra[2]
Eric P. Chassignet[2]

1 Dept. of Earth, Ocean and Atmospheric Science &
Center for Ocean-Atmospheric Prediction Studies,
2 Florida Climate Institute
Florida State University, Tallahassee, Florida

Appeared in special issue "Multi-disciplinary assessment of southeast US climate" in Regional Environmental Change

A.1 Abstract

In this study we compare the simulation of El Niño and the Southern Oscillation (ENSO) in the historical integrations of 17 Coupled Model Intercomparison Project 5 (CMIP5) models with corresponding observations. The mean-state and ENSO variations are analyzed in both the atmosphere and ocean and it is found that most of the CMIP5 models exhibit cold (warm) biases in the equatorial (subtropical eastern) Pacific Ocean sea surface temperature that are reminiscent of the split inter-tropical convergence zone phenomenon found in previous studies. There is, however, a major improvement in the representation of the power spectrum of the Niño3.4 sea surface temperature variations which shows that, as in the observations, a majority of the models display a spectral peak in the 2-7 year range, have a near linear relationship with the displacement of the equatorial thermocline and exhibit a robust atmospheric response to ENSO variations. Several issues remain such as erroneous amplitudes in the Niño3.4 sea surface temperature spectrum's peak and a width of the spectral peak that is either too broad or too narrow. It is also seen that most CMIP5 models unlike the observations extend the ENSO variations in the equatorial Pacific too far westward beyond the dateline and there is very little asymmetry in event duration between the warm and cold phases. ENSO variability forces a dominant mode of rainfall variability in the southeastern United States, especially in the boreal winter season. The CMIP5 exhibited a wide range of response in this metric with several displaying weak to non-existent, some showing relatively strong, and one indicating excessively zonally symmetric teleconnection over the southeastern United States.

A.2 Introduction

Validating El Niño and the Southern Oscillation (ENSO) in coupled ocean-atmosphere climate models is considered to be vital to understand and build confidence in the fidelity of the model (Guilyardi et al. 2009). This is partly because ENSO is one the best known natural climate variations (Philander 1990), which is relatively well observed (Zebiak and Cane 1987; Battisti 1988; Battisti & Hirst 1989; Hayes et al. 1991; McPhaden 1993; Jin 1997; Neelin et al. 1998) and, in comparison to other natural climate signals, is well understood theoretically (Kirtman 1997; Clarke 2008). There is also a good precedent for ENSO intercomparison studies which have been benchmarking the progress of the global coupled ocean-atmosphere model development (Mechoso et al. 1995; AchutaRao and Sperber 2006; Guilyardi et al. 2012). In the

process of this constant engagement of ENSO verification from one generation of models to the other, the community is slowly getting a better comprehension of the metrics that critically evaluate the model simulations (Guilyardi et al. 2009; NRC 2010). Most importantly, however, ENSO variations strongly affect the climate variations in the southeastern United States (Ropelewski and Halpert 1986; Kiladis and Diaz 1989; Diaz et al. 2001; Misra and DiNapoli 2012). Therefore the validation of the ENSO simulation in a climate model is relevant to evaluating its efficacy for use in understanding the southeastern US climate variations and change.

From the first climate model intercomparison study that displayed no ENSO features (Mechoso et al. 1995) there has been a steady, incremental progress in the simulation of ENSO (Guilyardi et al. 2012). One of the significant challenges in improving ENSO representation climate models is that we have not come across a panacea that works universally across all coupled models. For example, several modeling groups have successfully improved the tropical mean state and the ENSO simulation of their climate models by changing one of the following a) deep convective parameterization scheme in the atmospheric model (Zhang and Wang 2006; Neale et al. 2008), b) diffusion in the ocean models (Meehl et al. 2001), and c) resolution of the atmospheric and oceanic models (Gent et al. 2012). However, the same modification in other climate models did not lead to the same improvement and often yielded undesired results (http://www.iges.org/ctbm05/meetingreport.html). It is important to note that most of these climate models continue to systematically display a split Inter-Tropical Convergence Zone (ITCZ) (Mechoso et al. 1995), which is considered to be detrimental to the ENSO simulation. Modest improvements in the intensity of this split-ITCZ bias have been reported in (Zhang and Wang 2006; Neale et al. 2008).

Other common errors in the representation of ENSO in coupled models are an erroneous westward extension of the ENSO variability beyond the dateline, a narrower than observed meridional extent of the equatorial Pacific SST anomalies, and an ENSO which is too periodic and symmetric (AchutaRao and Sperber 2006; Joseph and Nigam 2006). Features of ENSO that are traditionally well represented are the seasonal phase locking of ENSO variability and the teleconnection of ENSO over North America (Misra et al. 2007; Joseph and Nigam 2006).

In this study, we examine the mean state as well as the variations on the ENSO time scales in the historical simulations of the Coupled Model Intercomparison Project 5 (CMIP5)

models, which will be used in preparing the International Panel for Climate Change (IPCC) Assessment Report 5 (AR5). It should be noted that the CMIP3 suite of models was used in the preparation of AR4 (Solomon et al. 2007); there was no CMIP4. A brief description of the model output and the validation datasets used in the analysis are provided in the supplementary material. The results are discussed in the following section followed by conclusions in section 3.

A.3 Results

A.3.1 Mean state

Several studies (Federov and Philander 2001; Wittenberg et al. 2006) suggest that fidelity of the mean state is critical for successful ENSO simulation. The mean observed SST field (Fig. 1a) shows the equatorial cold tongue off the coast of Peru and the tropical western Pacific warm pool. The replication of this equatorial zonal temperature gradient is important for the coupled feedbacks of the ENSO variations (Clarke 2008). The CMIP5 model SST error fields (Figs. 1b-r) exhibit biases in three areas: a cold bias over the cold tongue in the equatorial Pacific Ocean, a warm bias in the eastern oceans of the subtropical region, and a cold bias in the western portion of the subtropical Pacific Ocean. The annual mean SST errors averaged over these three regions are indicated in Table 2.1. All of these errors were quite prevalent in the CMIP3 models (AchutaRao and Sperber 2006; Capotondi et al. 2006), and, unfortunately, these errors are still present in the CMIP5 suite of models. The most extreme cold bias in the equatorial Pacific is seen in the CSRIO-Mk3.6 (Fig. 1f) and it is the only model to have a cold bias over nearly the entire tropical Pacific basin. However, GISS-E2-H (Fig. 1j) and GISS-E2-R (Fig. 1k) display a widespread warm bias across the eastern oceans, covering even the cold tongue region. In GFDL-CM3 (Fig. 1g), MIROC5 (Fig. 1o), and NorESM1-M (Fig. 1r) the errors at the equator and over the Peruvian coast are comparatively far less but they show significant cold biases in northwestern Pacific. Table 2.1 and Fig. 1c indicate that CanESM2 has the least bias in these three regions.

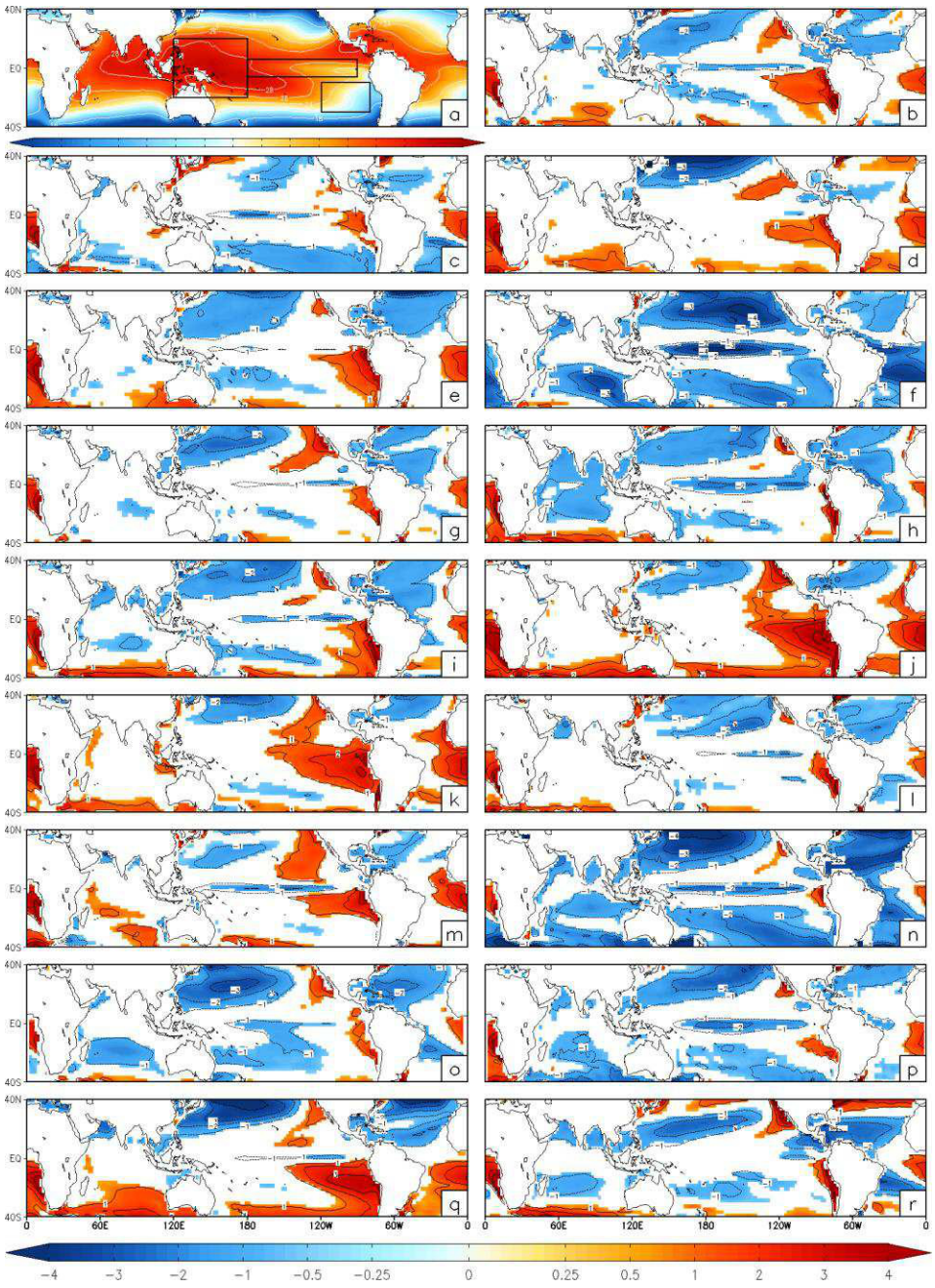

Figure A1: a) Observed annual mean SST (ERSSTv3b) and annual mean SST errors from, b) BCC-CSM1-1, c) CanESM2, d) CCSM4, e) CNRM-CM5, f) CSIRO-Mk3-6, g) GFDL-CM3, h) GFDL-ESM2G, i) GFDL-ESM2M, j) GISS-E2-H, k) GISS-E2-R, l) HadGEM2-ES, m) INM-CM4, n) IPSL-CM5A-LR, o) MIROC5, p) MPI-ESM-LR, q) MRI-CGCM3, and r) NorESM1-M The units are in °C. Values significant at 95% confidence limit are shaded.

In Fig. 2, we show the Pacific equatorial cross-section of the mean thermocline depth (defined as the depth of the 20°C isotherm) averaged between 5°S and 5°N for each of the models overlaid with the depth from the observations-based reanalysis GODAS. It is generally seen that the slope of the thermocline depth is reasonably well captured by the models with deeper (shallower) depths in the western (eastern) equatorial Pacific Ocean. However, the model biases tend to cluster with a majority of the climate models having steeper gradient across equatorial Pacific Ocean. In GISS-E2-R the slope of the thermocline is relatively weaker than most other models and observations while CCSM4 seem to nearly replicate the zonal gradient of the equatorial Pacific thermocline depth in GODAS. Further information on the climatological seasonal cycle of the equatorial Pacific Ocean in the CMIP5 models is provided in the supplementary material.

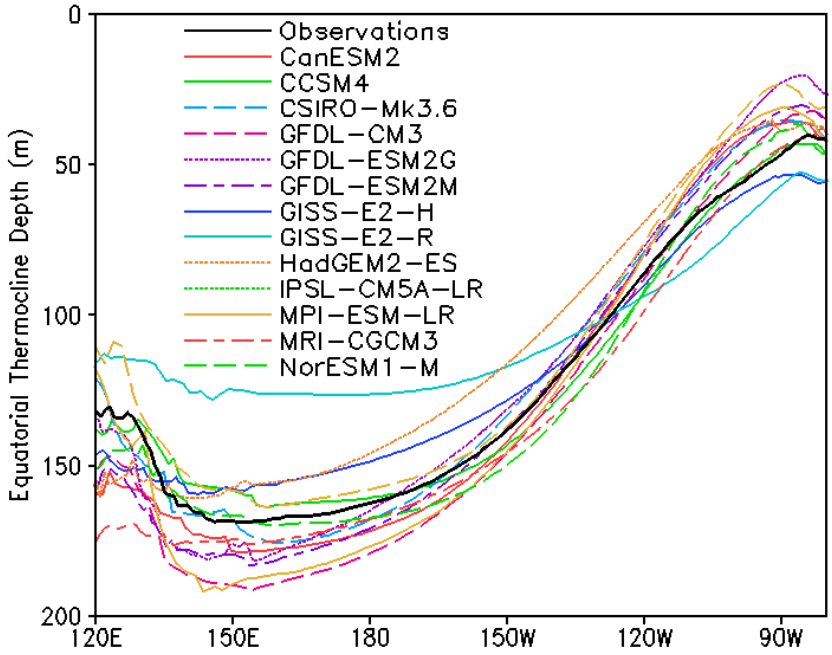

Figure A2: The annual mean thermocline depth (20°C isotherm) in meters averaged between 5S-5N.

A.3.2 Spectral Analysis

Using the maximum entropy method of order 40 (Ghil et al. 2002), the spectra based on the Niño3.4 SST is shown in Fig. 3; the models have been subdivided into three classes by the strength of the ENSO signal. The Niño 3.4 SST spectral peak of CSIRO-Mk3.6, GFDL-ESM2G, GISS-E2-H, GISS-E2-R, INM-CM4, and IPSL-CM5-LR (Fig. 3a) is broadest of all the other CMIP5 models. Furthermore in this group of models it may be noted that the amplitude of the spectrum at peak and at the biennial time scale is comparable to the observations. However, four of the models (CSIRO-Mk3.6, GISS-E2-H, GISS-E2-R, INM-CM4, and IPSL-CM5A-LR) in Fig. 3a also exhibit the largest bias in SST (Table 2.1 and Fig. 1). The spectra of BCC-CSM1-1, CanESM2, CNRM-CM5, GFDL-CM3, HadGEM-ES, and MPI-ESM-LR (Fig. 3b) display a very strong ENSO signal (in terms of the amplitude) compared to the observations. Furthermore, with the exception of MPI-ESM-LR, all have a significant secondary peak at around the biennial time scale. Lastly, in Fig. 3c we show the spectrum of the Niño3.4 SST anomalies from CCSM4, GFDL-ESM2M, MIROC5, and NorESM1-M, which have peak variability closer to the observed frequency than other CMIP5 models. However, the power of the ENSO spectrum in this class of models is relatively much higher than any other group of models including that in Fig. 3b. Further, the slender peaks of the spectrum are also indicative of the ENSO events being too periodic in this class of models. The supplementary material further discusses the evaluation of the duration of the ENSO events and the seasonal phase locking feature of ENSO variability. A key feature of the ENSO variations is that it is phase locked with the seasonal cycle of the eastern equatorial Pacific SST (Chang et al. 1995). It is seen that ENSO variability usually peaks near the end of the year, when the SST are coolest in the eastern equatorial Pacific. This feature of ENSO is further analyzed in the supplementary material.

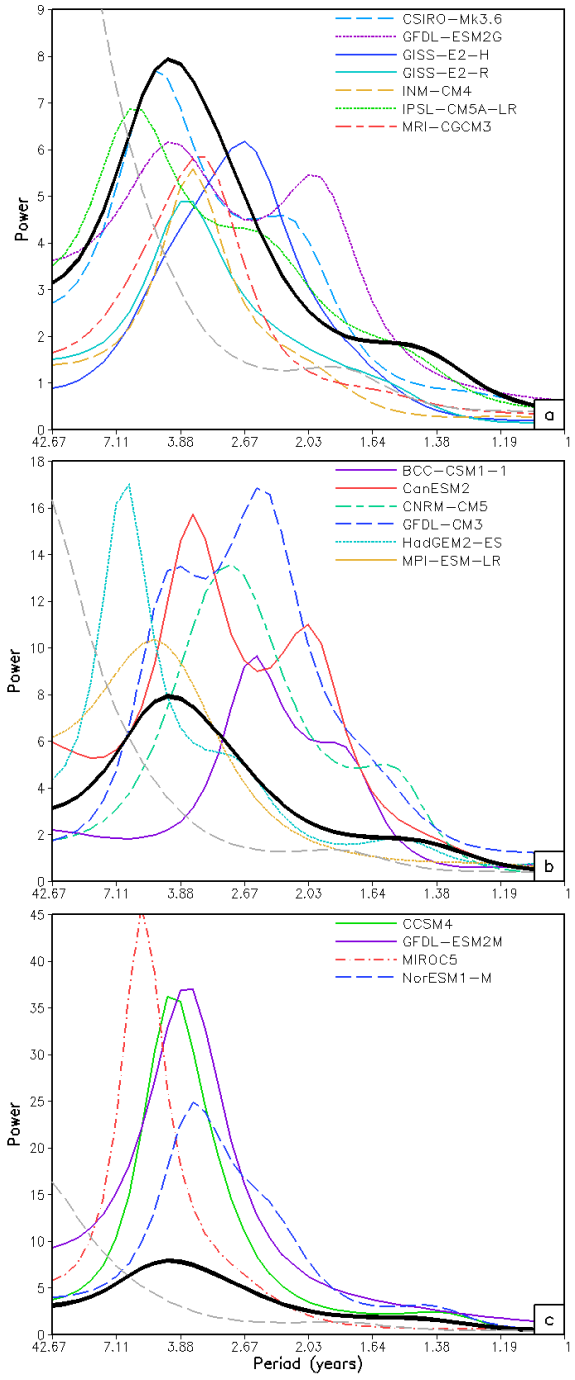

Figure A3: Sample spectrum of the Niño 3.4 SST using maximum entropy method (Ghil et al. 2002) with observations (ERSSTv3b) in black and AR1 model of observations in gray.

A.3.3 Lag/lead Relationship with Equatorial Pacific

The lag/lead relationship of the Niño3.4 SST anomalies with the SST anomalies over the equatorial Pacific Ocean illustrates the asymmetry between the phases of ENSO, the zonal extent of the ENSO anomalies in the equatorial Pacific and the period of the ENSO oscillation. Fig. 4 shows the normalized lag/lead regressions for the equatorial Pacific SST (averaged between 5°N and 5°S) with Niño3.4 SST anomalies leading (lagging) equatorial Pacific SST anomalies for negative (positive) lags. All the models in Fig. 4, including the observations, show an asymmetry between the ENSO phases, with the amplitude of one phase being relatively stronger than the other. However, a majority of these models fail to show that the duration of one phase is longer than the other with possible exception of GISS-E2-R (Fig. 4k) and MRI-CGCM3 (Fig. 4p). The westward propagation of anomalies is not well simulated by a number of models; CNRM-CM5 (Fig. 4e), GISS-E2-R (Fig. 4k), HadGEM2-ES (Fig. 4l), and MIROC (Fig. 4o) all lack a distinct propagation and GFDL-ESM2M (Fig. 4i) appears to propagate eastward with time. Another major issue with the CMIP5 simulation is the erroneous westward extension of the SST anomalies beyond the dateline as displayed most acutely by CanESM2 (Fig. 4c), CSIRO-Mk3.6 (Fig. 4f), GFDL-CM3 (Fig. 4g), GFDL-ESM2G (Fig. 4h), GFDL-ESM2M (Fig. 4i), GISS-E2-H (Fig. 4j), INM-CM4 (Fig. 4m), IPSL-CM5A-LR (Fig. 4m), MIROC5 (Fig. 4o), and MPI-ESM-LR (Fig. 4p). In contrast, some models are unable to get the anomalies far enough to the east near the Peruvian coast (e.g. GISS-E2-H [Fig. 4j], GISS-E2-R [Fig. 4k], INM-CM4 [Fig. 4m], IPSL-CM5A-LR [Fig. 4n], and MRI-CGCM3 [Fig. 4q]). The variations of the eastern equatorial Pacific SST are closely related to the associated variations in the depth of the underlying thermocline (Zelle et al. 2004), which is further discussed in the supplementary material.

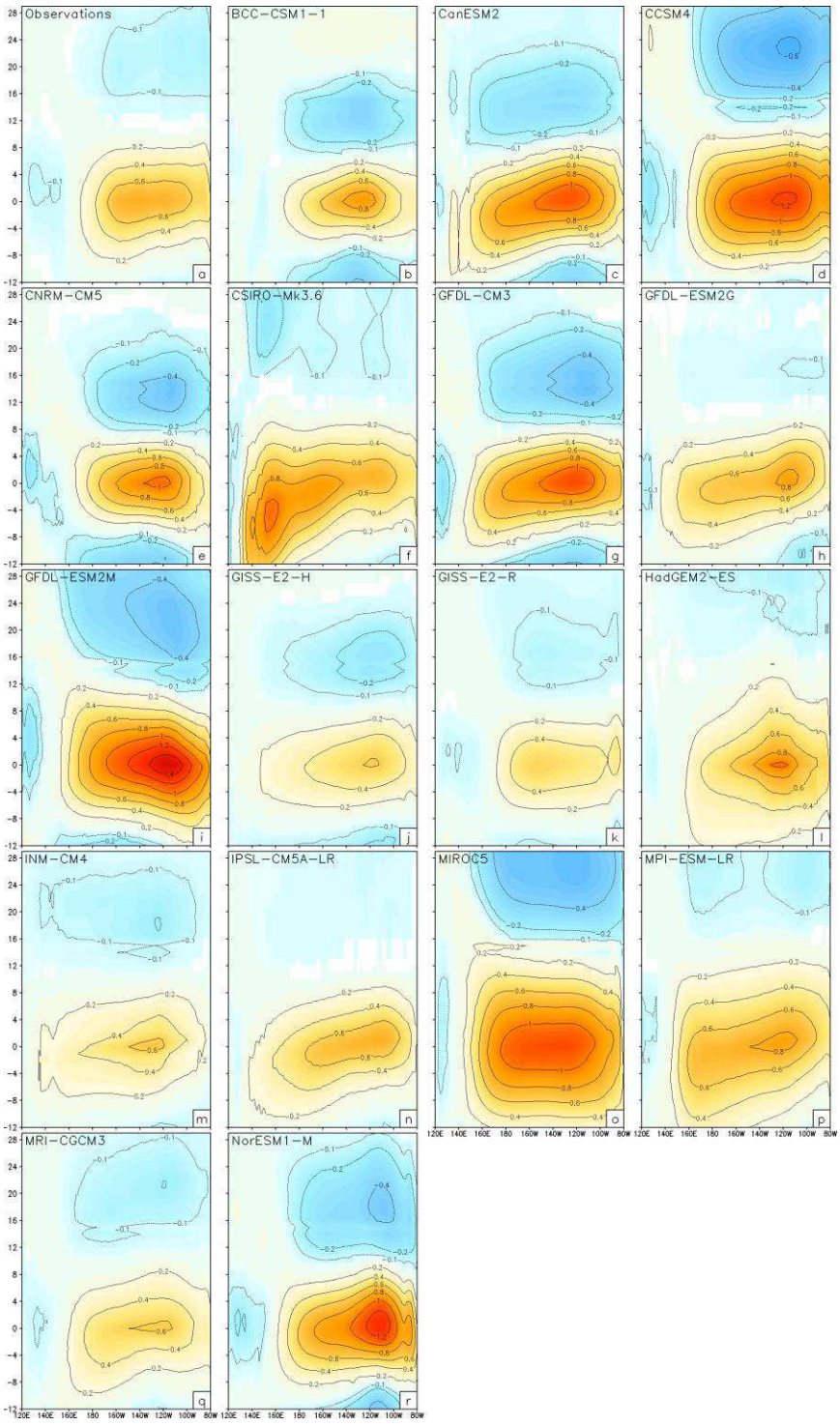

Figure A4: Lead/lag Niño 3.4 SST regression on equatorial Pacific SST normalized by the standard deviation of the Niño 3.4 SSTs for 17 models. Observations are from ERSSTv3b.

A.3.4 Remote ENSO forcing over the Southeastern United States

The mid-latitude response as a result of atmospheric waves traversing the great circle from anomalous convection in the equatorial oceans as a consequence of in-situ anomalous warm SST has been found to constitute over 50% of the variance of the boreal winter 200hPa geopotential heights in some areas of North America (Wallace and Gutzler 1981; Straus and Shukla 2000) including the southeastern United States. As a result of this teleconnection El Niño (La Niña) winters are typically characterized by wet and cold (warm and dry) winter climate in the southeastern United States.

Fig. 5a shows the Boreal winter (DJF) regression of Niño3.4 SST anomalies on precipitation from NCEP-NCAR reanalysis; likewise, Fig. 6a shows the Boreal winter (DJF) regression of Niño3.4 SST anomalies on 200hPa geopotential height. The observed geopotential height anomaly shows the typical high (over tropical Pacific), low (over Alaska), high (over Canadian Prairies) and low (over southeast US) associated with warm Niño3.4 SST anomalies that dictate the precipitation pattern of the Southeastern US in Fig. 5a. Most of the CMIP5 models (Figs. 5b-r, 6b-r) are able to capture some form of this ENSO teleconnection over the Southeastern US. There are large variations in the response, with several being rather weak (GISS-E2-H [Figs. 5j, 6j], GISS-E2-R [Fig. 5k, 6k]) or non-existent (GFDL-ESM2G [Figs. 5h, 6h], GFDL-ESM2M [Fig. 5i]), several being relatively strong (CanESM2 [Figs. 5c, 5c], CCSM [Figs. 5d, 6d], CSIRO-Mk3-6 [Fig. 5g], NorESM1-M [Fig. 6r]), and one (INM-CM4 [Fig. 5m, 6m]) being excessively zonally symmetric.

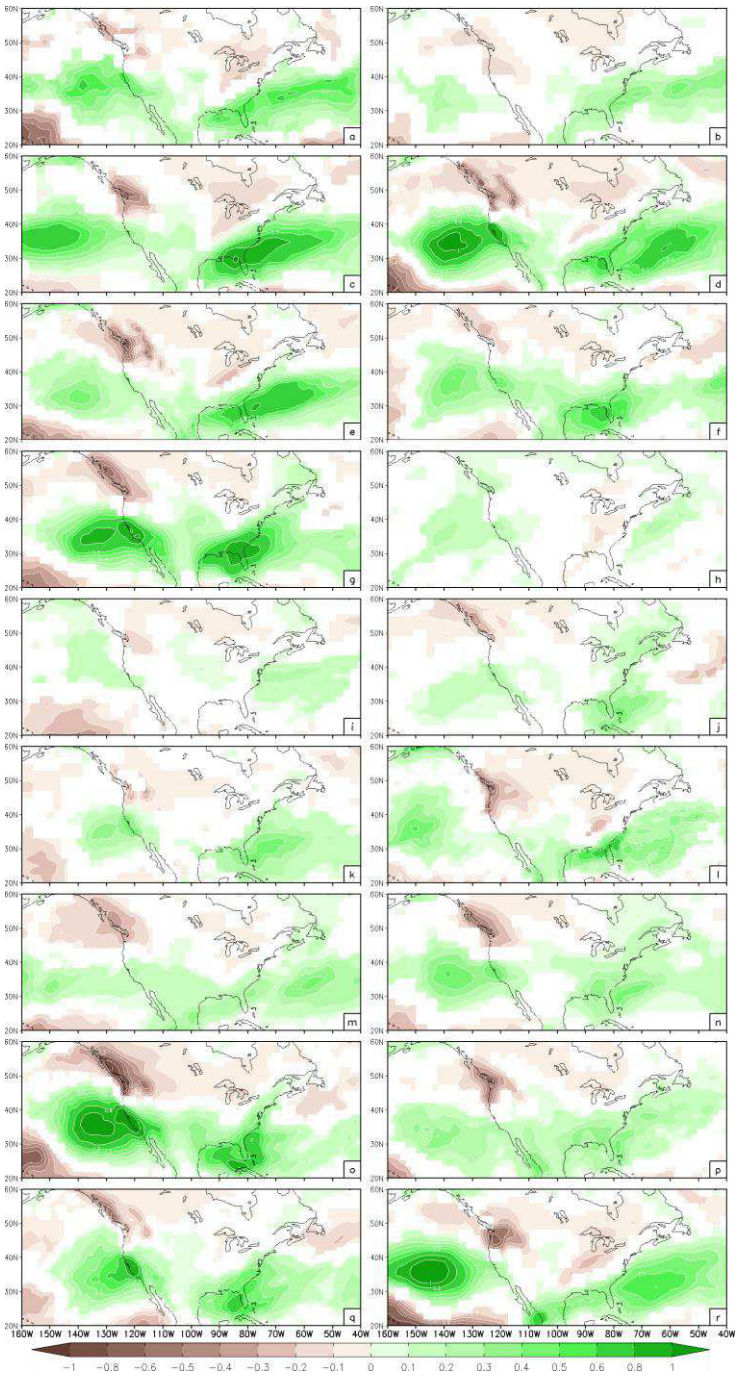

Figure A5: Regression of Niño 3.4 SST on DJF precipitation anomalies normalized by the standard deviation of the Niño 3.4 SST anomalies for a) Observations (NCEP-NCAR), b) BCC-CSM1-1, c) CanESM2, d) CCSM4, e) CNRM-CM5, f) CSIRO-Mk3-6, g) GFDL-CM3, h) GFDL-ESM2G, i) GFDL-ESM2M, j) GISS-E2-H, k) GISS-E2-R, l) HadGEM2-ES, m) INM-CM4, n) IPSL-CM5A-LR, o) MIROC5, p) MPI-ESM-LR, q) MRI-CGCM3, and r) NorESM1-M. Units are in millimeters per day.

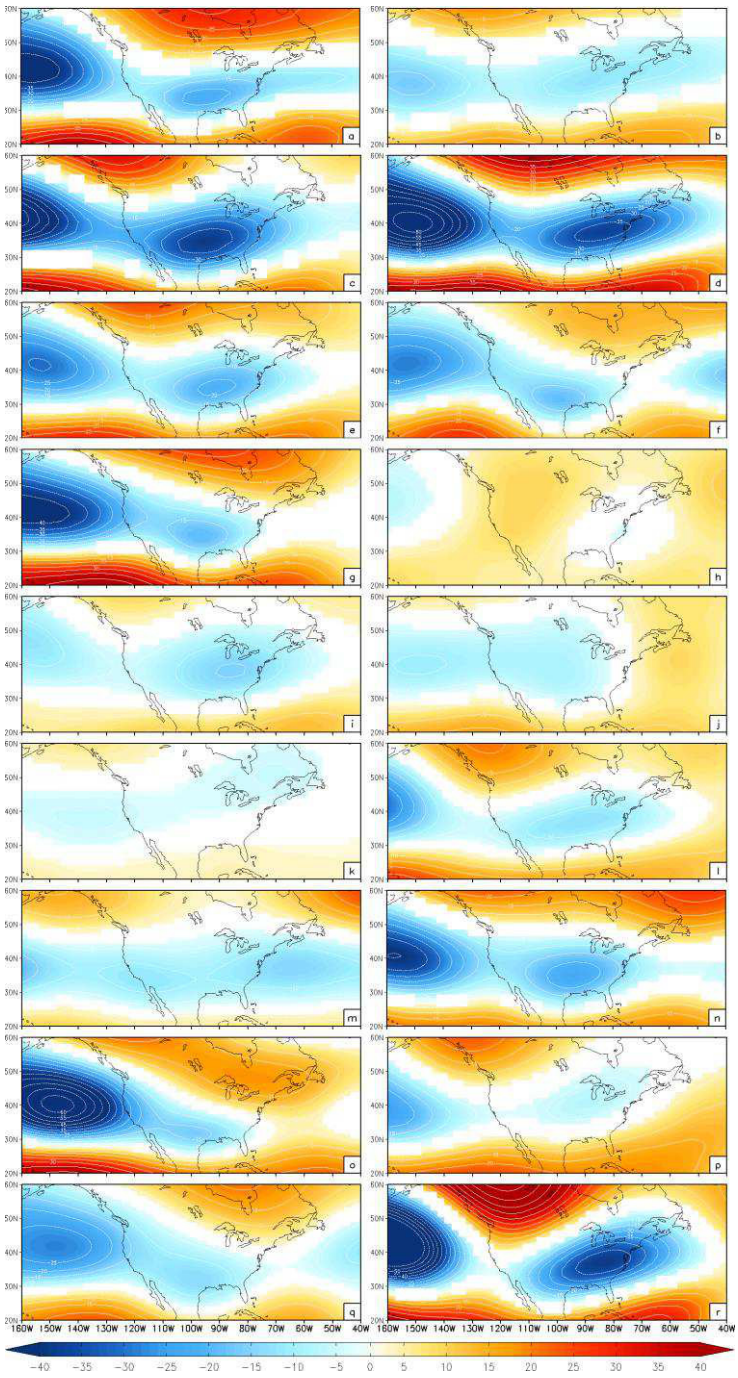

Figure A6: Linear regression of Niño 3 SST on 200hPa geopotential height for a) Observations (ERSSTv3b and NCEP-NCAR reanalysis), b) BCC-CSM1-1, c) CanESM2, d) CCSM4, e) CNRM-CM5, f) CSIRO-Mk3-6, g) GFDL-CM3, h) GFDL-ESM2G, i) GFDL-ESM2M, j) GISS-E2-H, k) GISS-E2-R, l) HadGEM2-ES, m) INM-CM4, n) IPSL-CM5A-LR, o) MIROC5, p) MPI-ESM-LR, q) MRI-CGCM3, and r) NorESM1-M. Units are in meters.

A.4 Summary and Conclusion

The teleconnection of ENSO variation with the southeast climate variations cannot be understated. It is one of the most well-known teleconnection patterns with warm (cold) ENSO events typically resulting in wet (dry) and cold (warm) winters in the southeastern US. By analyzing the ENSO fidelity in the 20th century simulations of these CMIP5 models, we are providing some guidance to potential users who would want to use the results of these models to assess the future impact of climate change on say the local ecology, hydrology, crop yield in the southeastern US.

In this paper, we have examined the surface and sub-surface oceanic variables, the coupled feedbacks and the atmospheric response associated with ENSO variations in the centennial integrations forced with the time varying 20th century emissions of the CMIP5 historical runs. CMIP5 models are part of the latest generation of models that will be extensively analyzed for the forthcoming IPCC AR5. CMIP5 has followed from the CMIP3 suite of models, which was used in IPCC AR4 (Solomon et al. 2007); there was no CMIP4.

Our analysis shows that majority of the CMIP5 models continues to display an erroneous split ITCZ feature, with a cold SST bias in the equatorial oceans and an overtly active ITCZ just south of the equator. The warm bias in the stratiform regions of the eastern subtropical oceans is also prominent. There was significant diversity in the CMIP5 historical simulation of the ENSO teleconnection with the southeastern US climate. Many models produced geopotential height patterns over North America uncharacteristic of the observed ENSO teleconnection, with either the forced variability being too strong or too weak. These erroneous teleconnections were also reflected in the corresponding ENSO forced rainfall anomalies over southeastern US. However, the simulation of the mean state of the equatorial Pacific thermocline is most of the time well represented in the CMIP5 models, with its zonal slope comparable to the GODAS reanalysis. The seasonal cycle of equatorial Pacific SST is also well captured whereas the seasonality of the coupled feedback between the zonal wind stress and SST shows apparent issues with the models. There is some modest improvement in the power spectrum of the Niño3.4 SST variations from the CMIP3 models, in that the power has increased in several of the models. There are however only a minority of CMIP5 models whose ENSO power spectrum is comparable to the observed spectrum. Lastly, in a majority of the CMIP5 models, the mid-latitude atmospheric response to ENSO is quite robust and comparable to observations.

As stated earlier, improving climate models is not a straight forward task especially in a non-linear model with complex feedbacks. Furthermore, many of the CMIP5 models have introduced more complexities relative to CMIP3 by including interactions with aerosols; land ice, biogeochemical cycle, ecosystem models that could further accentuate the bias through coupled feedbacks. The resolution of the CMIP5 models for these centennial integrations has been raised by at least two fold (and even more in some cases) compared to the CMIP3 models. The community has invested significantly over the years, targeting many of these systematic errors through concerted efforts like the climate process teams in the United States and is actively considering a joint world effort (http://wcrp.ipsl.jussieu.fr/Workshops/ModellingSummit/Documents/FinalSummitStat_6_6.pdf; Shukla et al. 2009) to improve climate models. Validating ENSO simulation seems to be one of the best ways to test the fidelity of a climate simulation given that is relatively well understood phenomenon, has good observational analysis and there is a rich historical tradition of verifying ENSO, which has enabled a better understanding of model physics. More specifically, the strong teleconnections of ENSO with the southeastern United States climate makes this validation exercise highly relevant for understanding the reliability of the CMIP5 projections of future southeastern United States climate.

APPENDIX B

TABLES AND FIGURES

B.1 Tables

Table B.1: Reference list of all acronyms used in this document

Acronym	Description
CCSM4	Community Climate System Model version 4
CFS	Climate Forecast System model
CGCM	Coupled general circulation model
CMAP	Climate Prediction Center Merged Precipitation Analysis
CMIP	Coupled Model Intercomparison Project
CMIP5	Phase 5 of the CMIP project
COLA	Center for Ocean-Land-Atmosphere Studies
CPW	Central Pacific warming
CTB	Climate Test Bed
CZ	Cane-Zebiak model
DOT	delayed oscillator theory
ECMWF	Centre for Medium-Range Weather Forecasting
ENSO	El Niño and the Southern Oscillation
EOF	empirical orthogonal functions
GMAO	NASA Global Modeling and Assimilation Office
GODAS	Global Ocean Data Assimilation
IOD	Indian Ocean Dipole
IRI	International Research Institute for Climate and Society
JMA	Japanese Meteorological Agency
LIM	linear inverse model

Table B.1 – continued

Acronym	Description
LOAM	Linearized Ocean Atmosphere Model
MOO	Multi-Objective Optimization
MSD	mean squared difference
MSE	mean squared error
NCEP	National Centers for Environmental Prediction
CPC	Climate Prediction Center
NMME	North American Multi-Model Ensemble
NOAA	National Oceanic and Atmospheric Administration
NWP	numerical weather prediction
ODA	Ocean Data Assimilation
OND	October-November-December
RMS	root mean squared
ROC	relative operating characteristic curve
SAM	Southern Annular Mode
SIA	sea ice area
SOI	Southern Oscillation Index
SST	Sea surface temperature
SSTA	Sea Surface Temperature Anomalies
TC	tropical cyclone
UCAR	University Corporation for Atmospheric Research
WCT	Wall Clock Time

Table B.2: List of the CMIP5 models analyzed for ENSO

Model Name	Institution	Simulation Years
BCC-CSM1-1	Beijing Climate Center, China Meteorological Administration	1850-2012
CanESM2	Canadian Centre for Climate Modelling and Analysis	1850-2005
CCSM4	National Center for Atmospheric Research	1850-2005
CNRM-CM5	Centre National de Recherches Meteorologiques / Centre Europeen de Recherche et Formation Avancees en Calcul Scientifique	1850-2005
CSIRO-Mk3-6	CSIRO (Commonwealth Scientific and Industrial Research Organisation, Australia), and BOM (Bureau of Meteorology, Australia)	1850-2005
GFDL-CM3	Geophysical Fluid Dynamics Laboratory	1860-2005
GFDL-ESM2G	Geophysical Fluid Dynamics Laboratory	1861-2005
GFDL-ESM2M	Geophysical Fluid Dynamics Laboratory	1861-2005
GISS-E2-H	NASA Goddard Institute for Space Studies	1850-2005
GISS-E2-R	NASA Goddard Institute for Space Studies	1850-2005
HadGEM2-ES	Met Office Hadley Centre	1860-2005
INM-CM4	Institute for Numerical Mathematics	1850-2005
IPSL-CM5A-LR	Institut Pierre-Simon Laplace	1850-2005
MIROC5	Atmosphere and Ocean Research Institute (The University of Tokyo), National Institute for Environmental Studies, and Japan Agency for Marine-Earth Science and Technology	1850-2012
MPI-ESM-LR	Max Planck Institute for Meteorology (MPI-M)	1850-2005
MRI-CGCM3	Meteorological Research Institute	1850-2005
NorESM1-M	Norwegian Climate Centre	1850-2005

Table B.3: Brief outline of observationally based datasets used in this study

Name	Dataset	Time Period	Grid Spacing	Reference
GODAS	Global Ocean Data Assimilation System	1980-2011	0.33x1	Behringer and Xue 2004
ERSSTv3b	Extended Reconstructed SST version 3b	1854-2011	2x2	Smith et al. 2008
CMAP	CPC Merged Analysis of Precipitation	1979-2008	2.5x2.5	Xie and Arkin 1997
NCEP-NCAR	NCEP-NCAR Reanalysis	1948-2010	2.5x2.5	Kalnay et al. 1996

Table B.4: Annual SST Errors in three regions of the Pacific Ocean: the cold tongue region, the warm pool region and the Southeast Pacific region for the CMIP5 models.

Model	6S-6N, 180-90W (Cold tongue region of equatorial Pacific)	20S-20N, 120E-180 (Warm pool region of the tropical western Pacific)	30S-10S, 120W-80W (Southeast Pacific region of stratocumulus)
BCC-CSM1-1	-0.360	-0.985	1.172
CanESM2	-0.102	-0.258	-0.307
CCSM4	0.061	-0.157	0.736
CNRM-CM5	0.003	-0.938	0.954
CSIRO-Mk3-6	-1.862	-0.908	-1.249
GFDL-CM3	-0.449	-0.550	0.231
GFDL-ESM2G	-1.367	-0.755	-0.755
GFDL-ESM2M	-0.608	-0.624	0.901
GISS-E2-H	0.985	-0.315	2.227
GISS-E2-R	1.396	0.094	1.506
HadGEM2-ES	-0.528	-0.419	0.101
INM-CM4	-0.252	-0.477	0.992
IPSL-CM5A-LR	-1.083	-1.085	-0.671
MIROC5	-0.413	-0.810	0.082
MPI-ESM-LR	-1.155	-0.697	0.376
MRI-CGCM3	-0.014	-0.345	2.421
NorESM1-M	-0.337	-0.843	-0.097

Table B.5 Mean Equatorial Pacific thermocline depth maxima and minima for selected CMIP5 models.

Model	Maximum (m)	Minimum (m)
Observations	168.744	39.8756
CanESM2	178.918	35.5945
CCSM4	163.71	43.1998
CSIRO-Mk3.6	175.461	35.3487
GFDL-CM3	191.66	32.3429
GFDL-ESM2G	181.61	20.2436
GFDL-ESM2M	183.295	30.3786
GISS-E2-H	159.438	53.291
GISS-E2-R	127.997	52.5621
HadGEM2-ES	160.808	35.1387
IPSL-CM5A-LR	163.746	23.3442
MPI-ESM-LR	191.639	31.0563
MRI-CGCM3	176.551	40.0751
NorESM1-M	169.927	36.4064

B.2 Figures

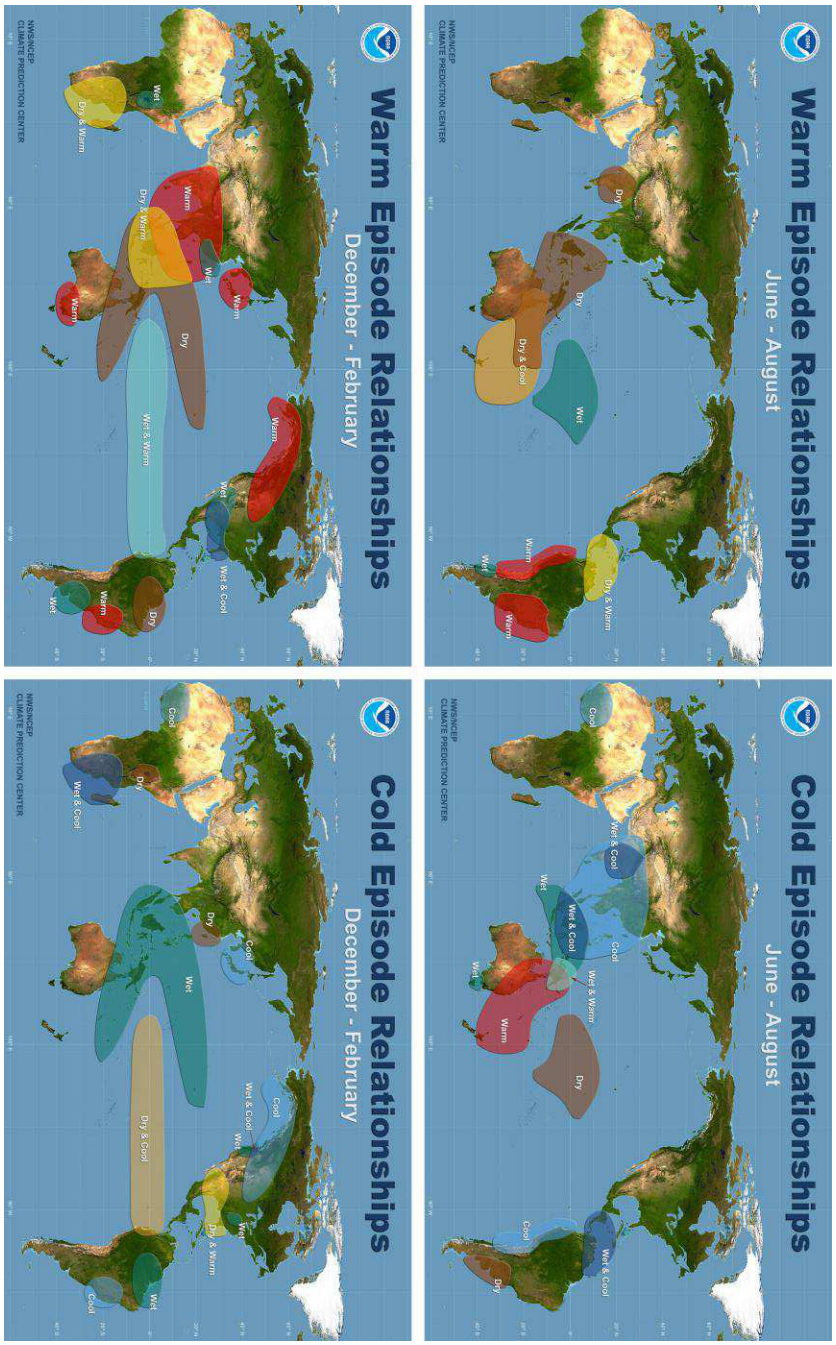

Figure B.1 Summer and winter weather patterns associated with warm and cold ENSO events. Image courtesy of NOAA/CPC available at http://www.cpc.ncep.noaa.gov/products/precip/CWlink/ENSO/ENSO-Global-Impacts/.

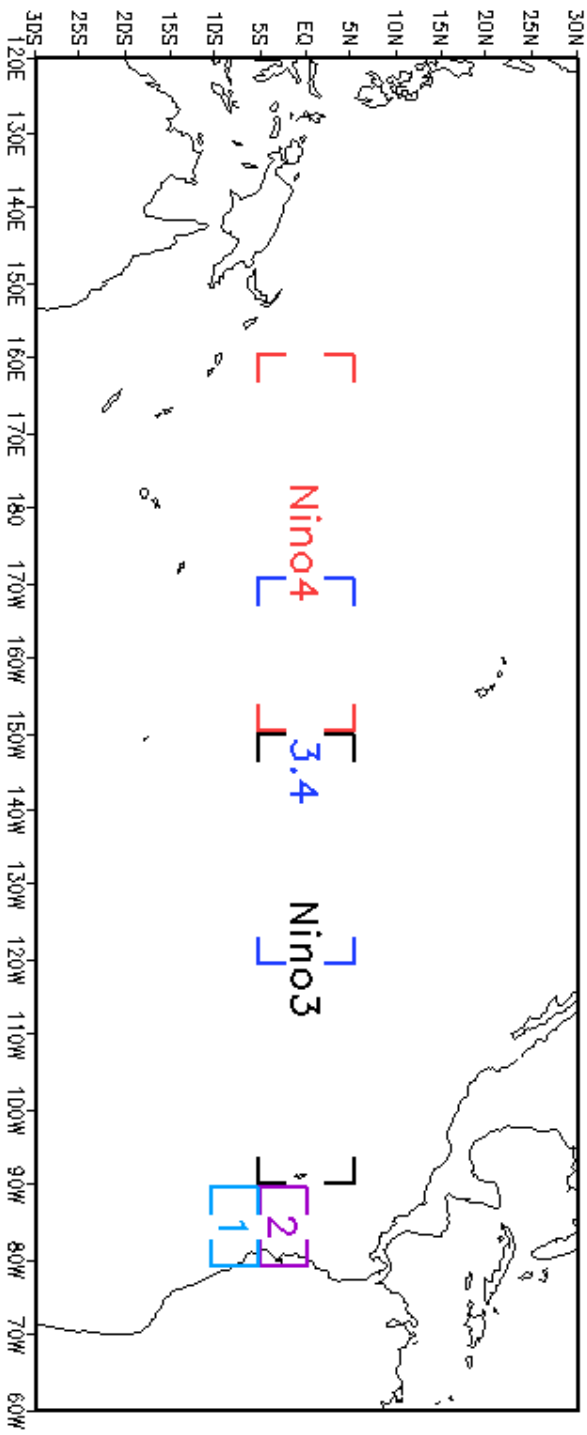

Figure B.2: Map of Niño Regions 1+2, 3, 3.4, and 4. These regions are used to calculate Phase and Strength of ENSO.

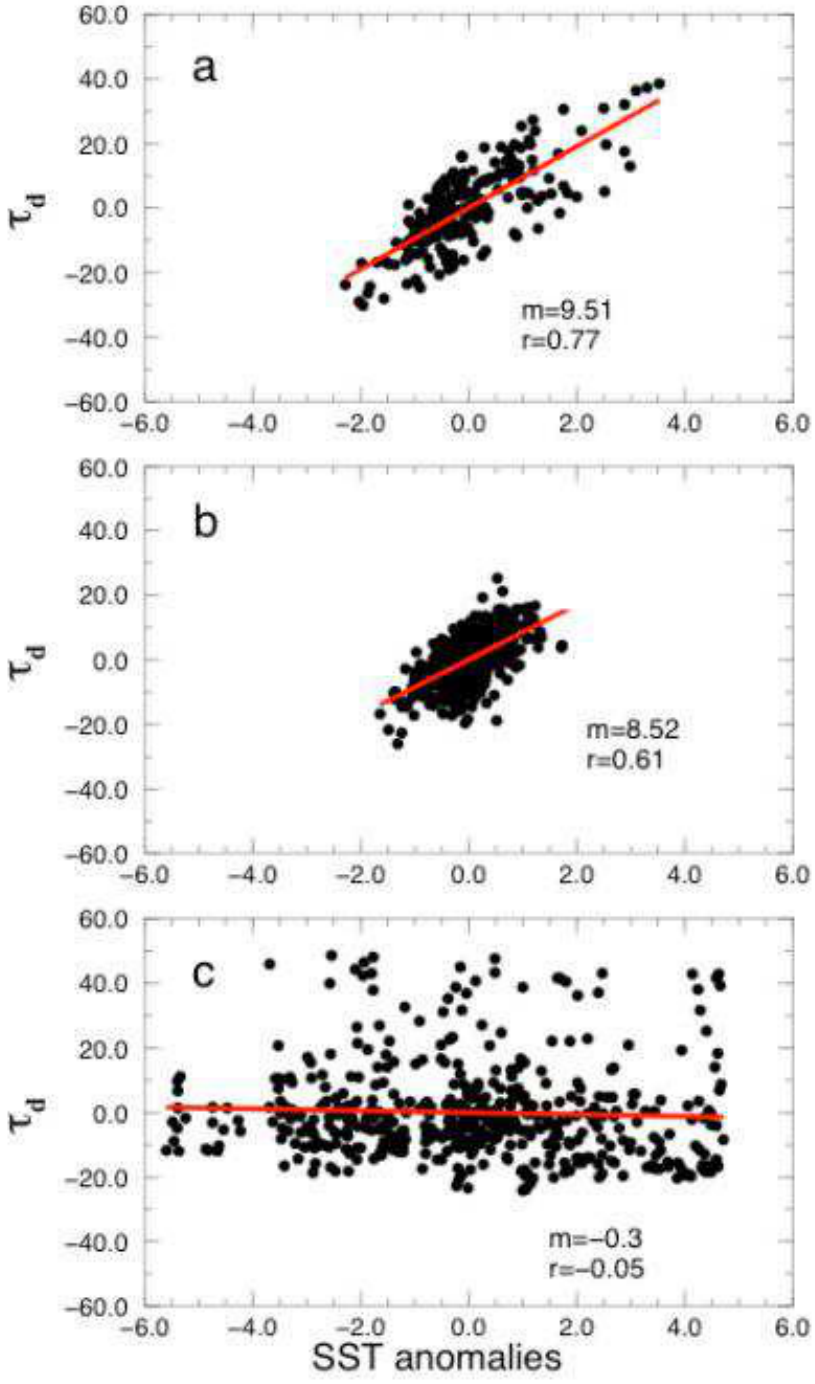

Figure B.3: Scatter of Niño3 SST anomalies (abscissa) and the corresponding anomalous depth of the 20°C isotherm at zero lag from a) GFDL ODA (Rosati et al. 1997), b) retrospective prediction, and c) multidecadal simulation from the same coupled (COLA) climate model (Misra et al. 2008)

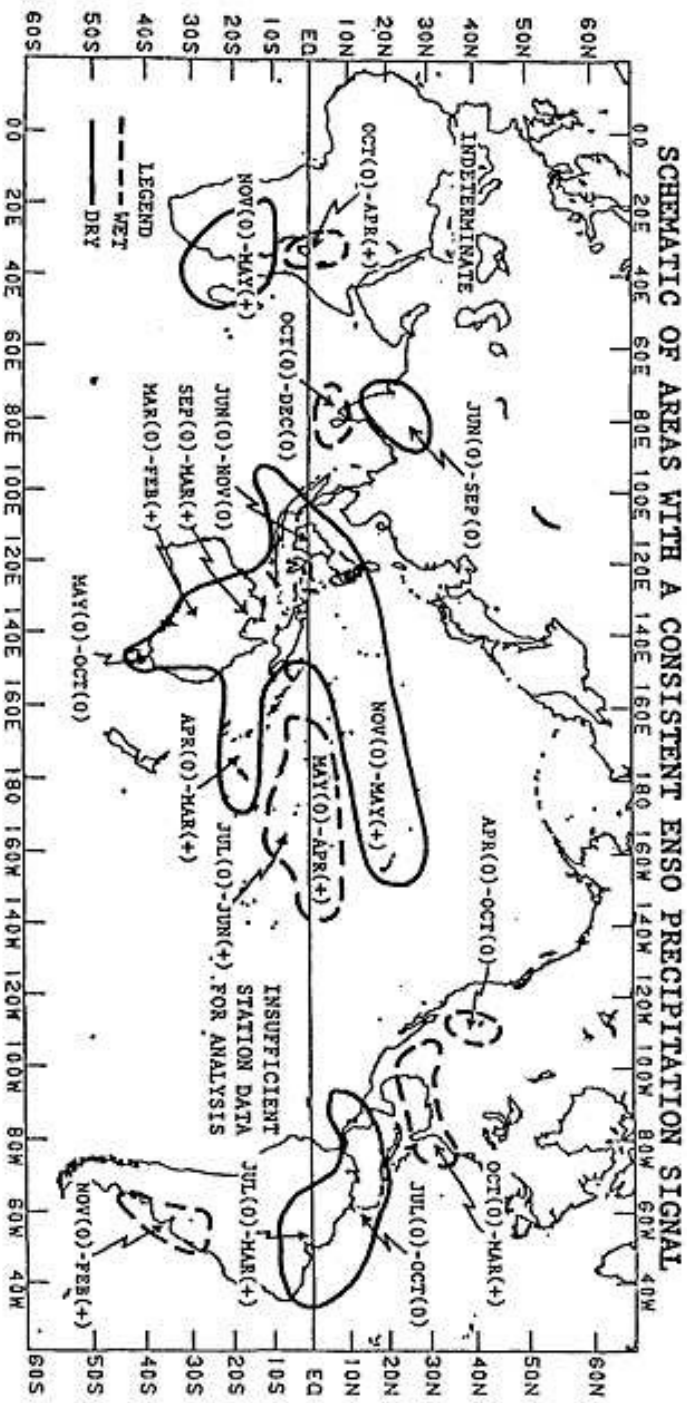

Figure B.4: Schematic representation of the principal ENSO precipitation teleconnections. From Ropelewski and Halpert (1987).

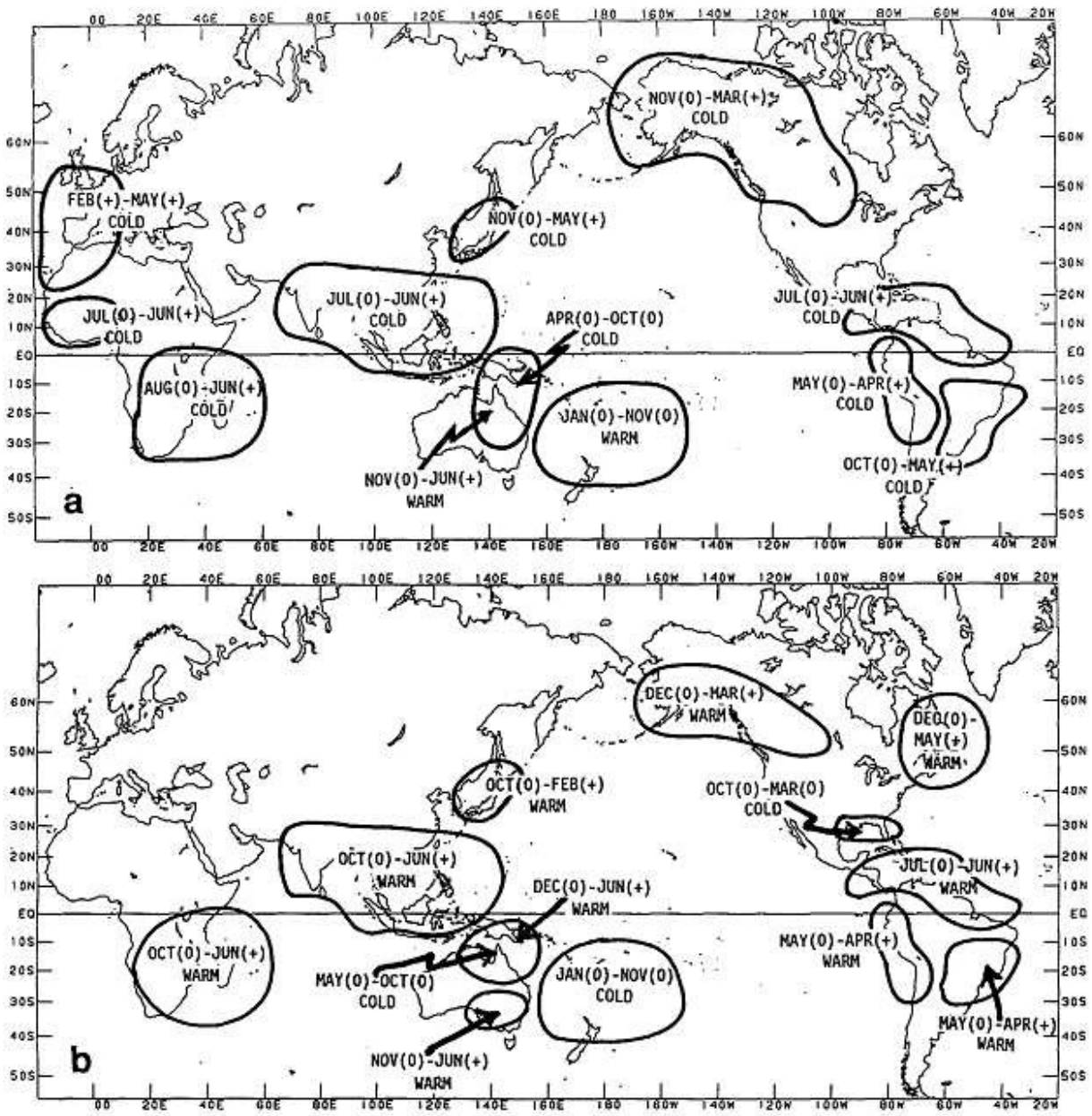

Figure B.5 Schematic representations of the principal ENSO temperature teleconnections for (a) cold phase (b) warm phase (Halpert and Ropelewski 1992).

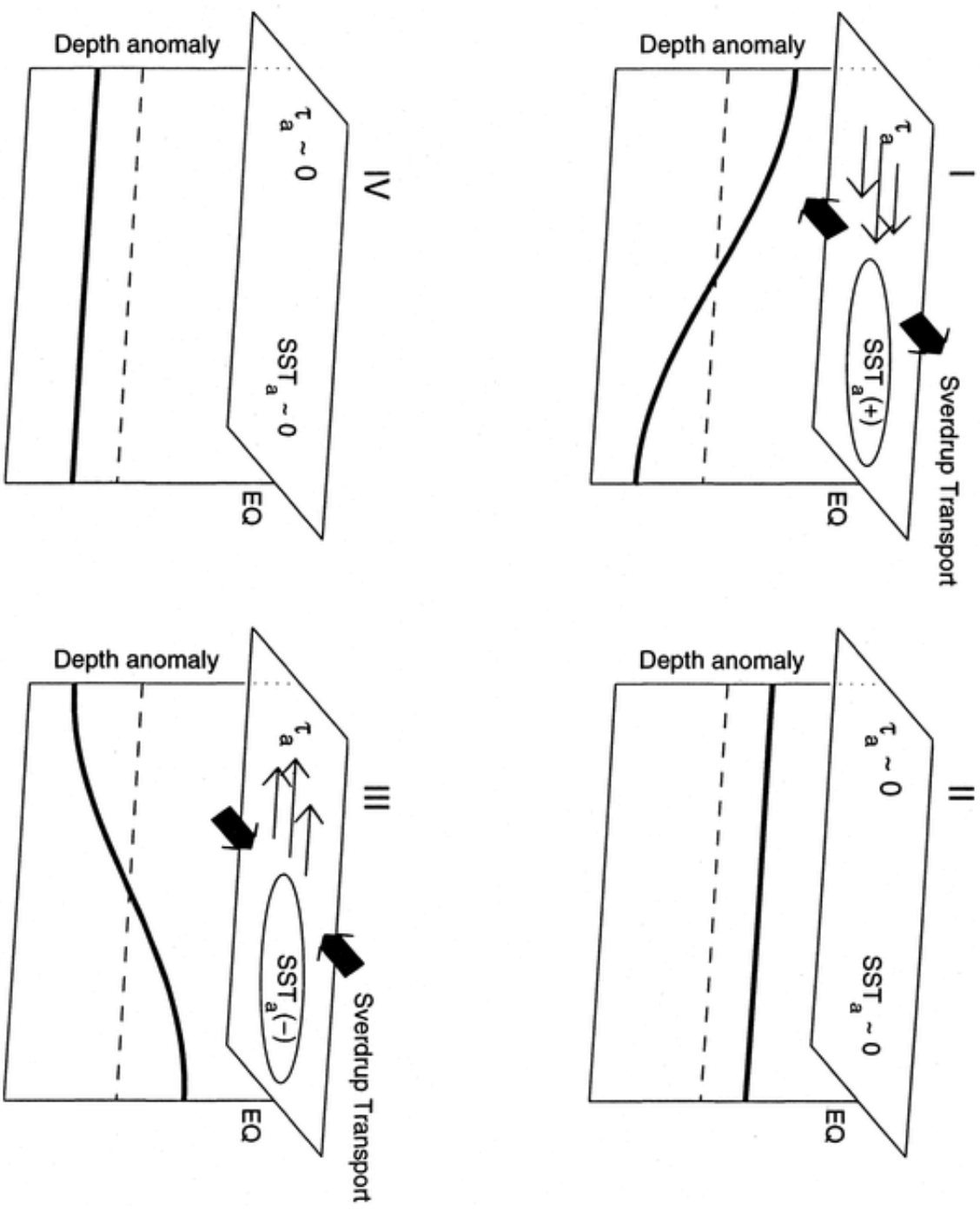

Figure B.6 Schematic of the four phases of the recharge-discharge oscillation I) warm phase II) warm to cold transition phase III) cold phase and IV) cold to warm transition phase. The rectangular box represents the Equatorial Pacific basin; the elliptical circle represents the SST anomaly; the thin arrows represent wind stress anomaly; the thick arrows represent the recharge/discharge of equatorial heat content; the heavy black line in each box shows the thermocline depth anomaly (Meinen and McPhaden 2000).

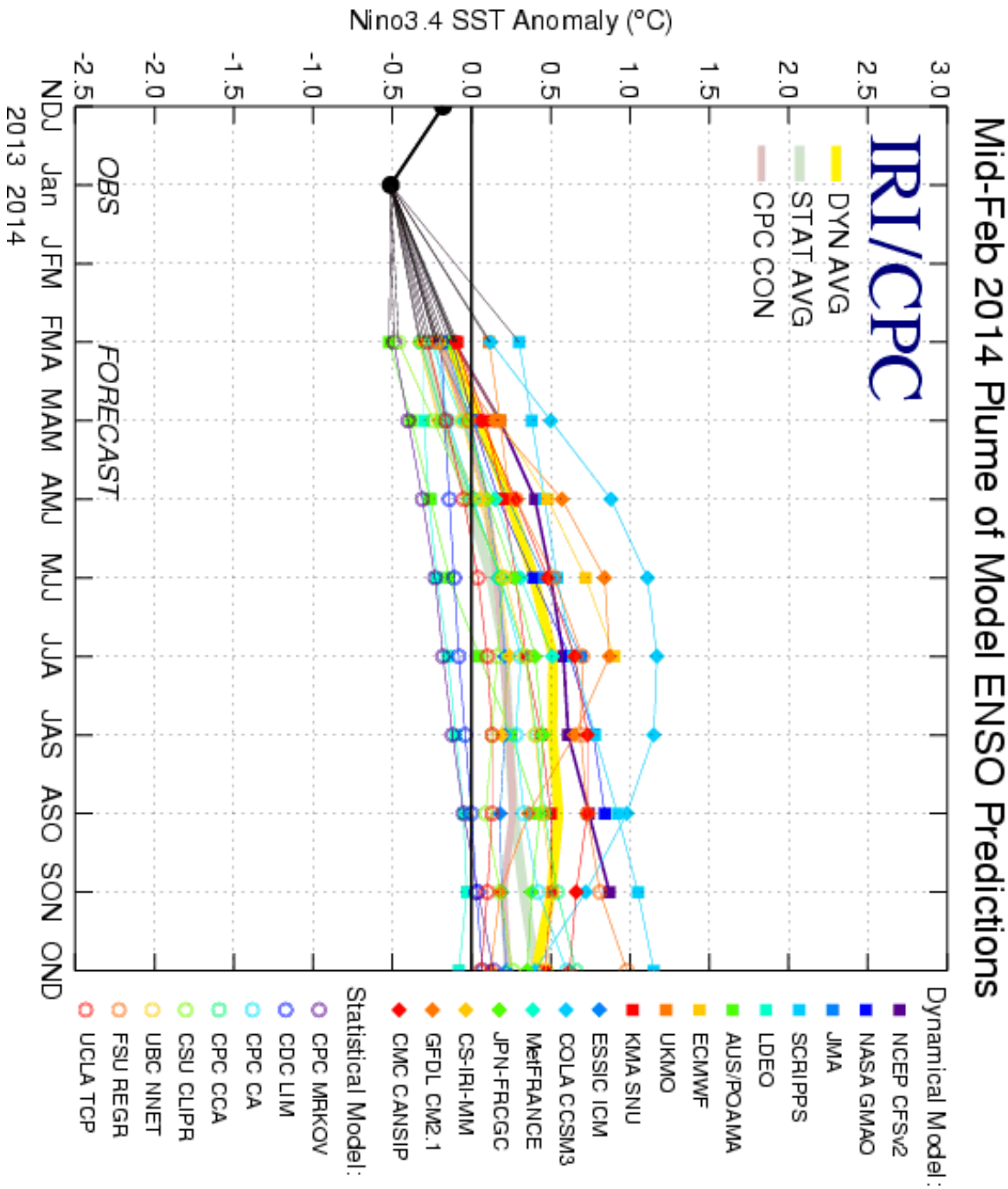

Figure B.7 February 2014 ENSO Plume forecast of Niño3.4 SSTs constructed from Statistical and Dynamical Models of ENSO. This forecast product is produced by the International Research Institute for Climate and Society in collaboration with CPC.

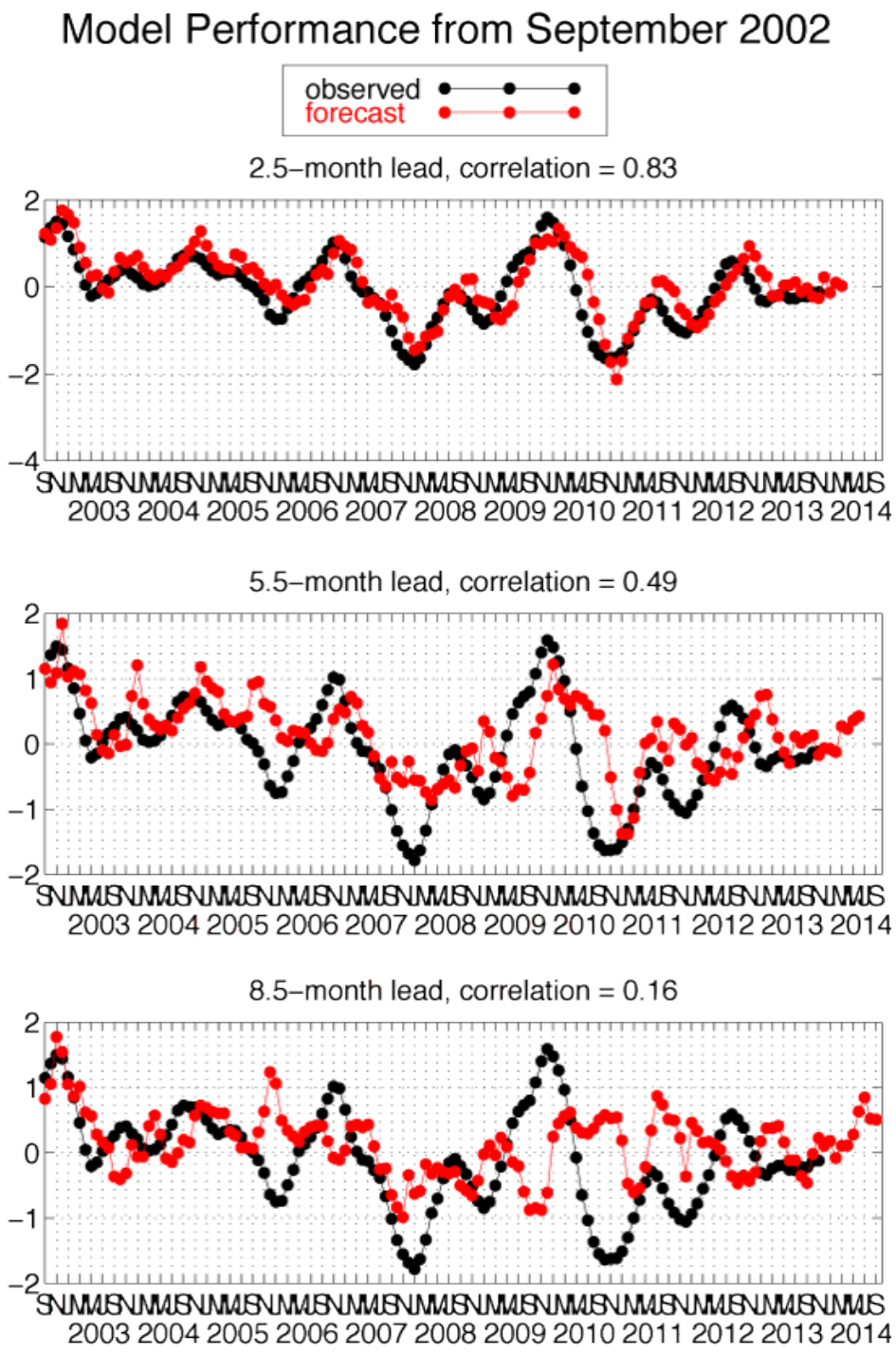

Figure B.8 Model performance of the Clarke and Van Gorder (2003) ENSO forecast model for 2.5-month lead, 5.5 month lead, and 8.5 month lead times from 2002 to 2014. Retrieved January 2014, from http://eoas.fsu.edu/elnino_forecast.

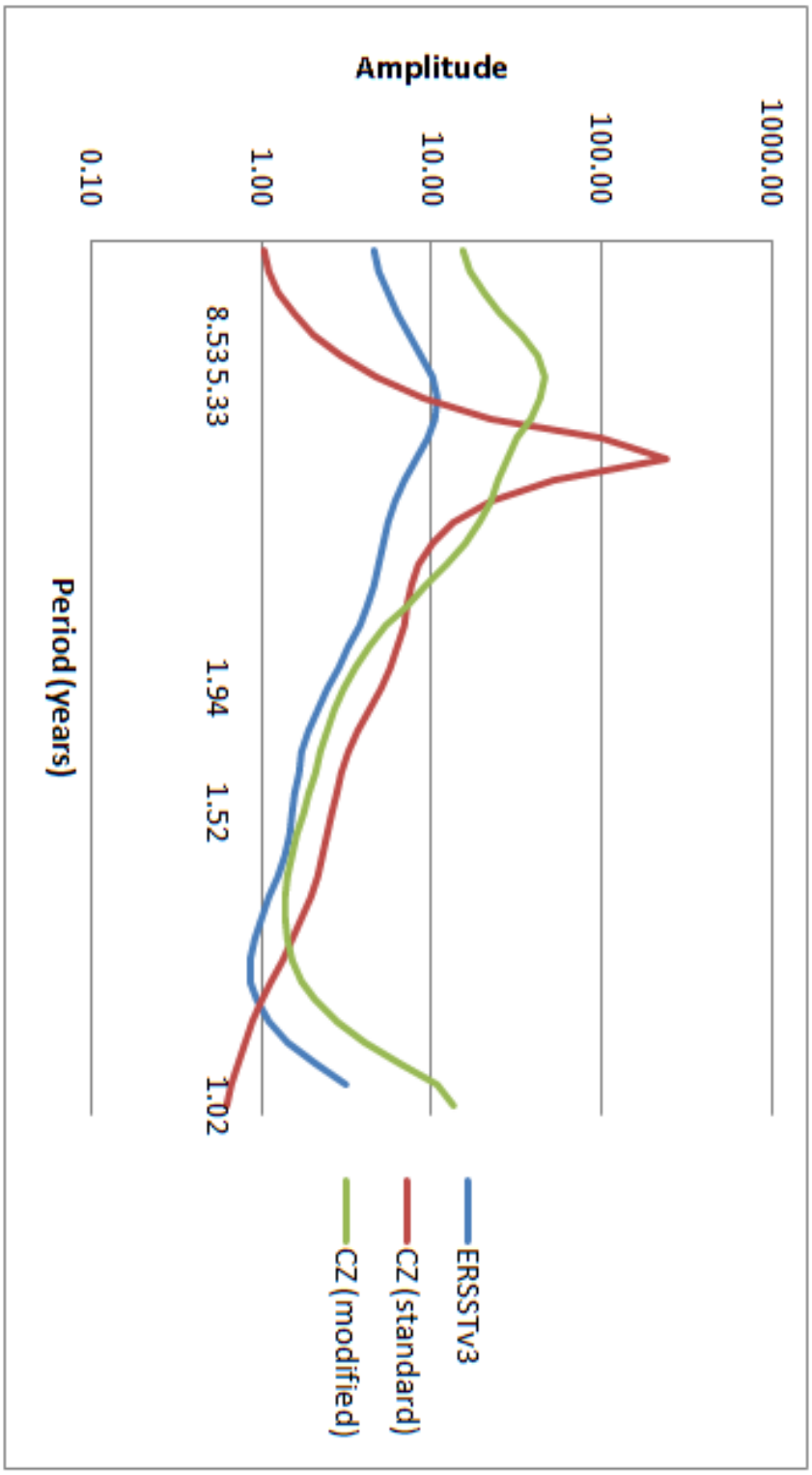

Figure B.9 The sample spectrum of ERSSTv3 (blue), Cane-Zebiak (red), and modified Cane-Zebiak model (green).

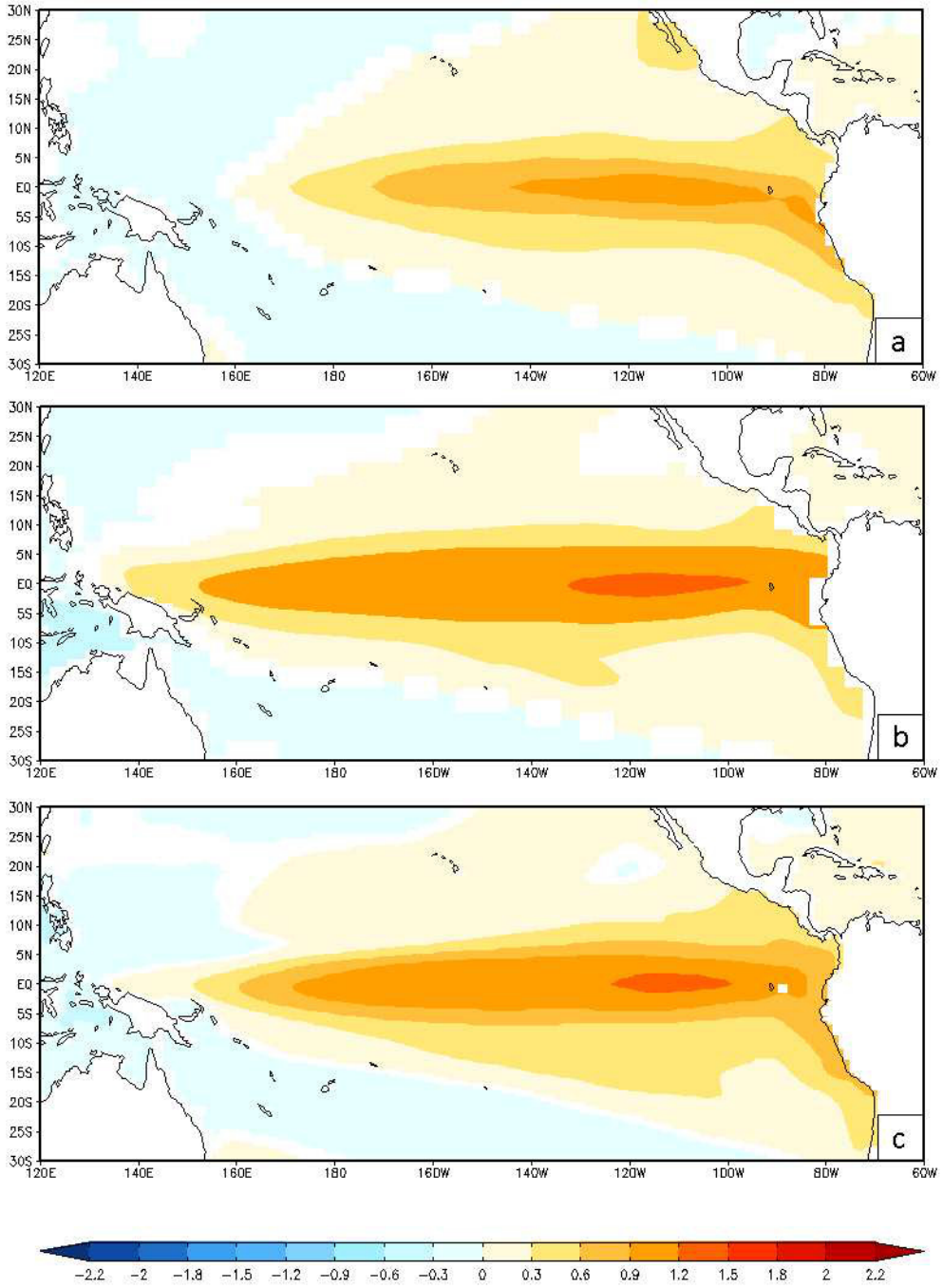

Figure B.10 Regression of Niño 3.4 SST on Tropical Pacific SST for (a) observations from ERSSTv3b, (b) CCSM4 low resolution, and (c) CCSM4 medium resolution. Significant values at 95% confidence plotted.

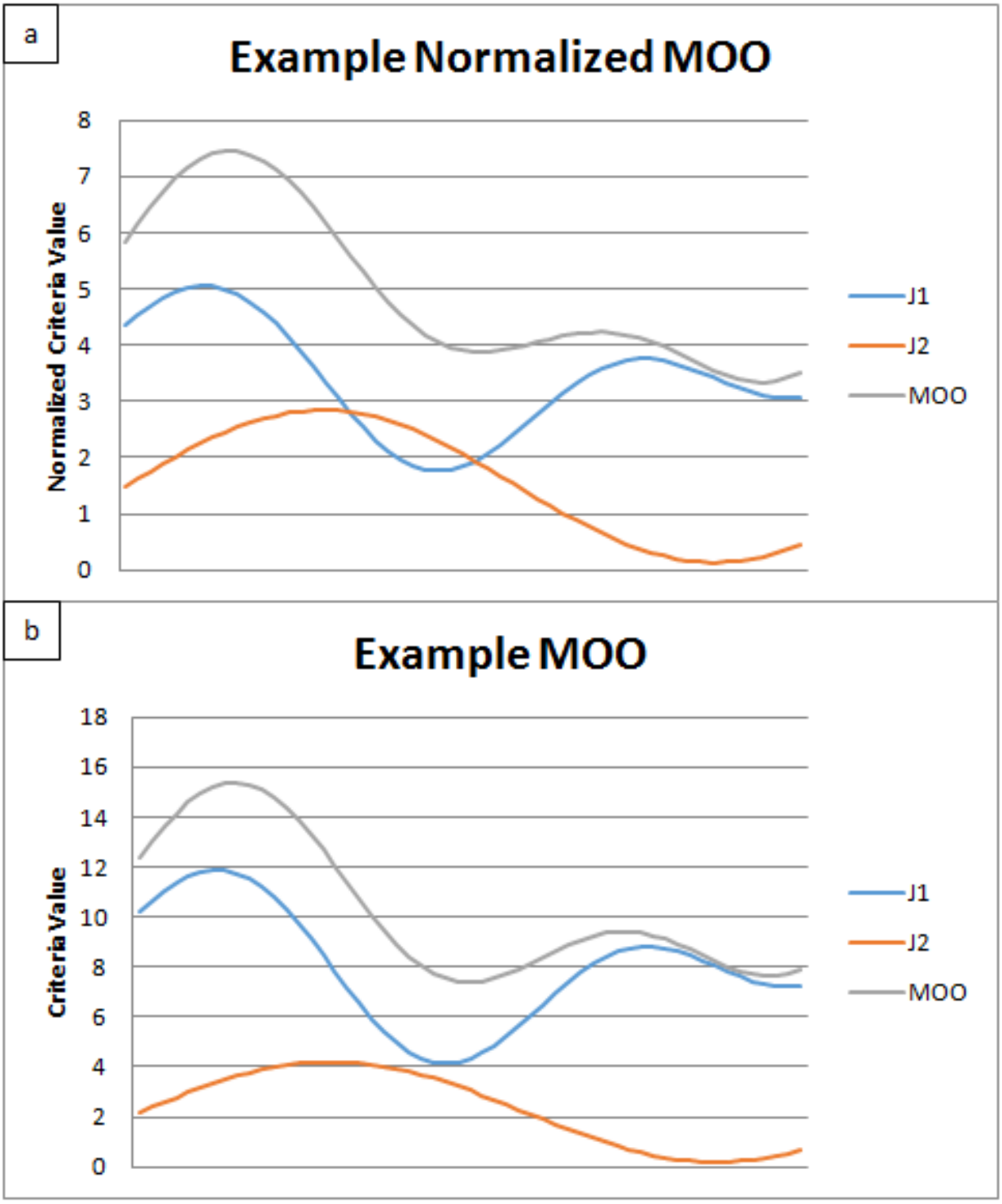

Figure B.11 Schematic examples of (a) two jointly optimized curves using multiobjective optimization and (b) two incorrectly optimized curves

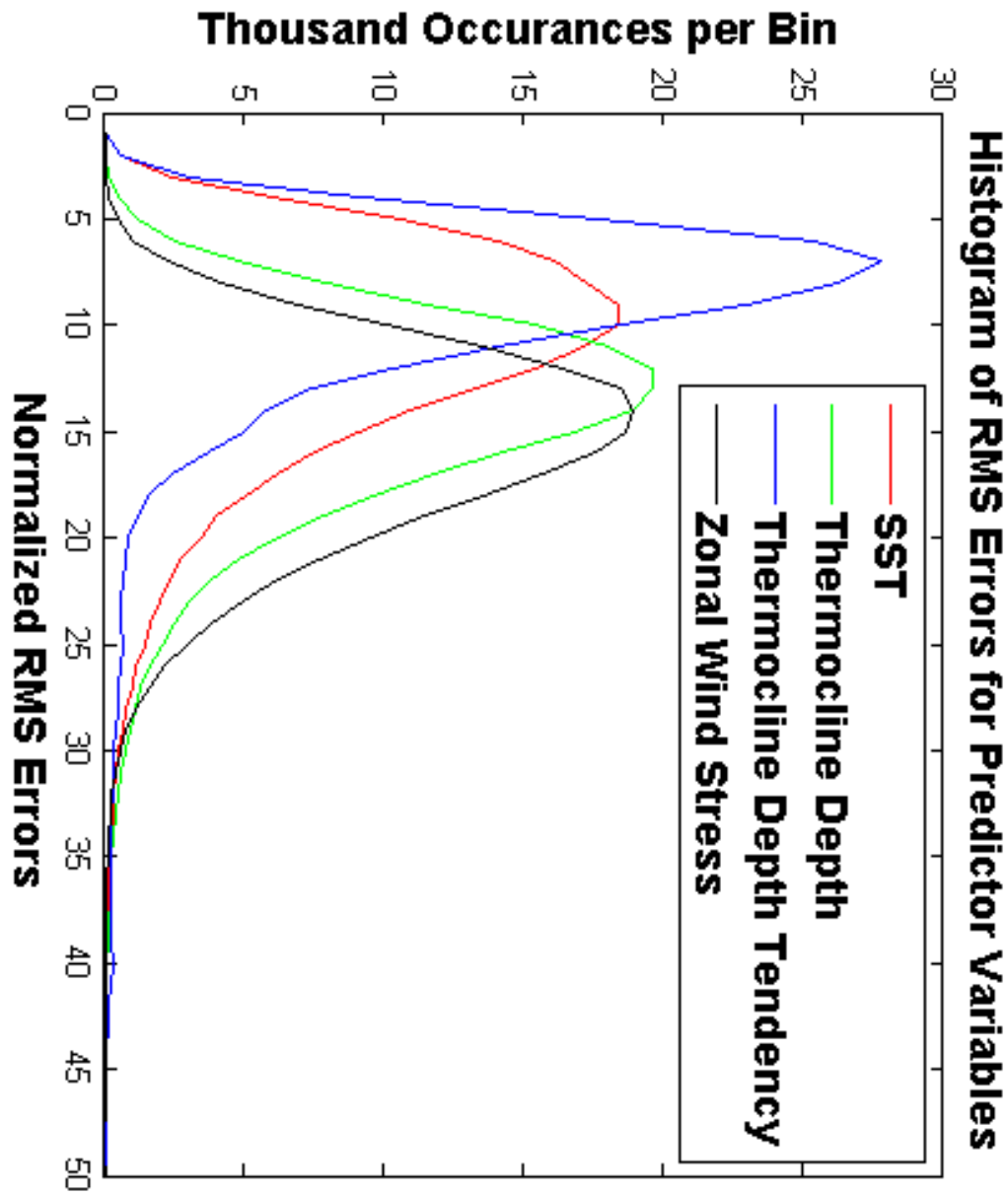

Figure B.12 Histogram of the RMS differences for all four predictor variables

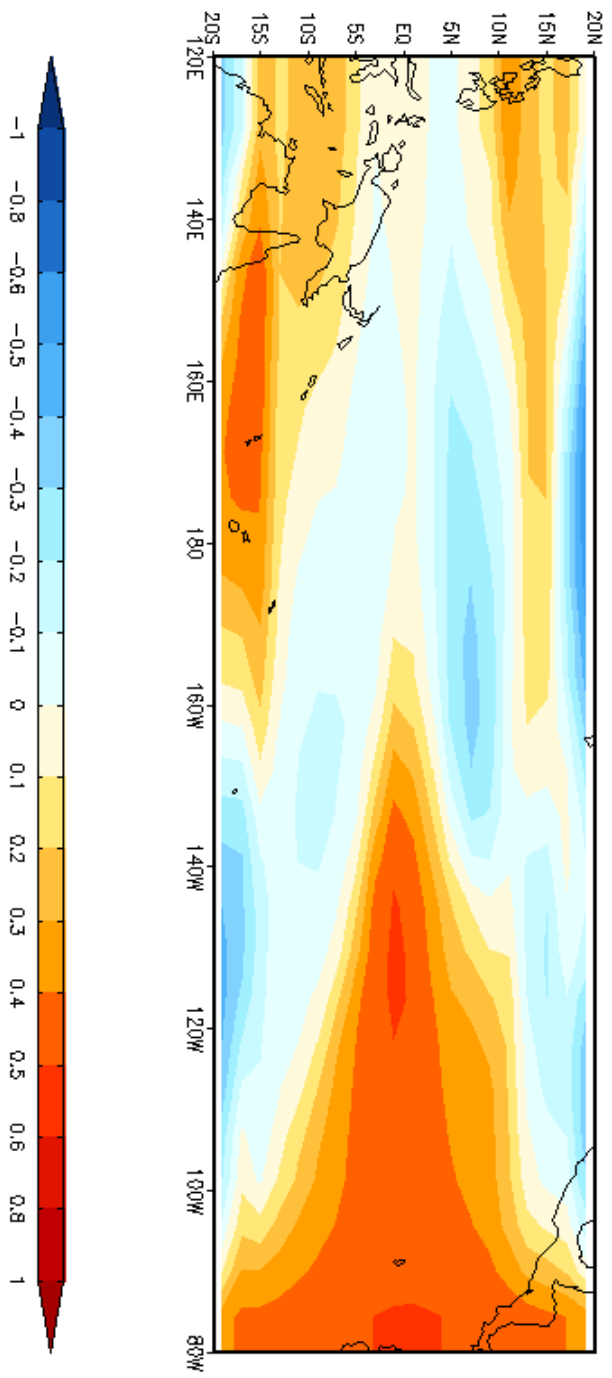

Figure B.13 As example, the 9-month lag correlation of thermocline depth (leads) with Niño 3 SST (lags) in the CZ model. In total, 432 similar plots were used to determine the subdomains of analogue matching used.

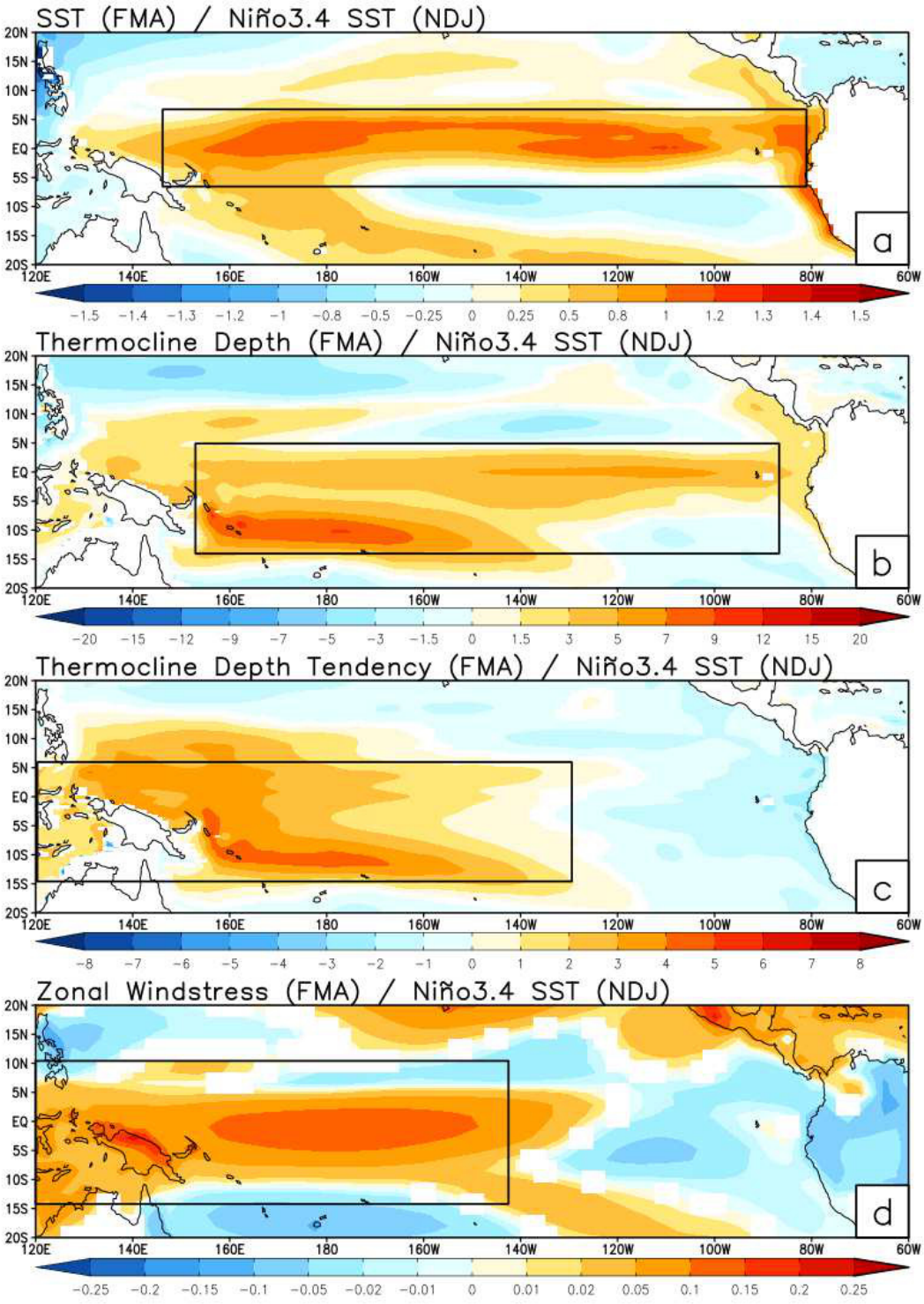

Figure B.14 9-month lag regression of December Niño 3.4 SST with (a) Tropical Pacific SSTs, (b) thermocline depth, (c) thermocline depth tendency, and (d) zonal wind stress for 600 years of CCSM4. Subregion boxes are overlaid in each panel.

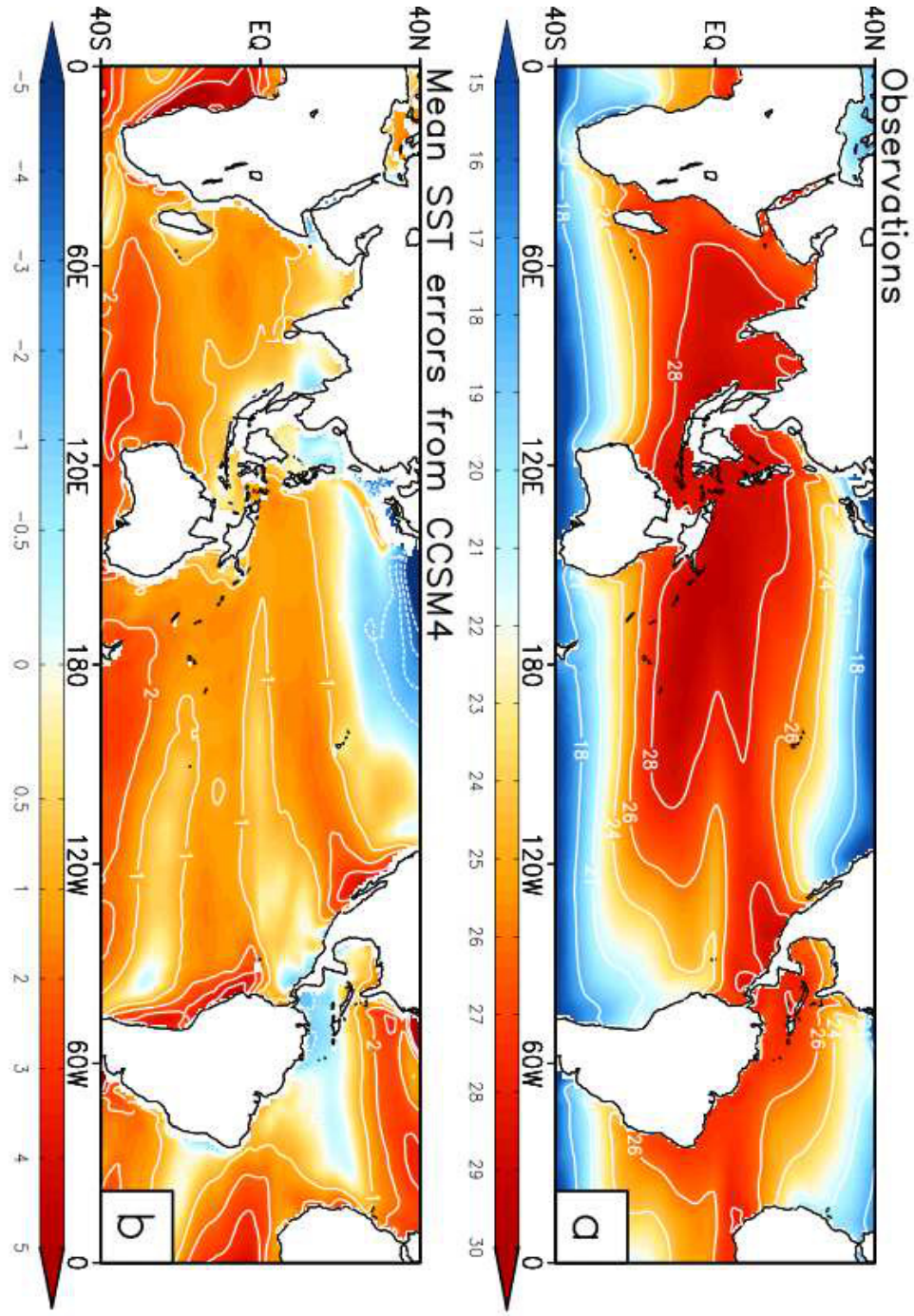

Figure B.15 Average SST over the Tropical Pacific for (a) GODAS and (b) the average SST errors of CCSM4

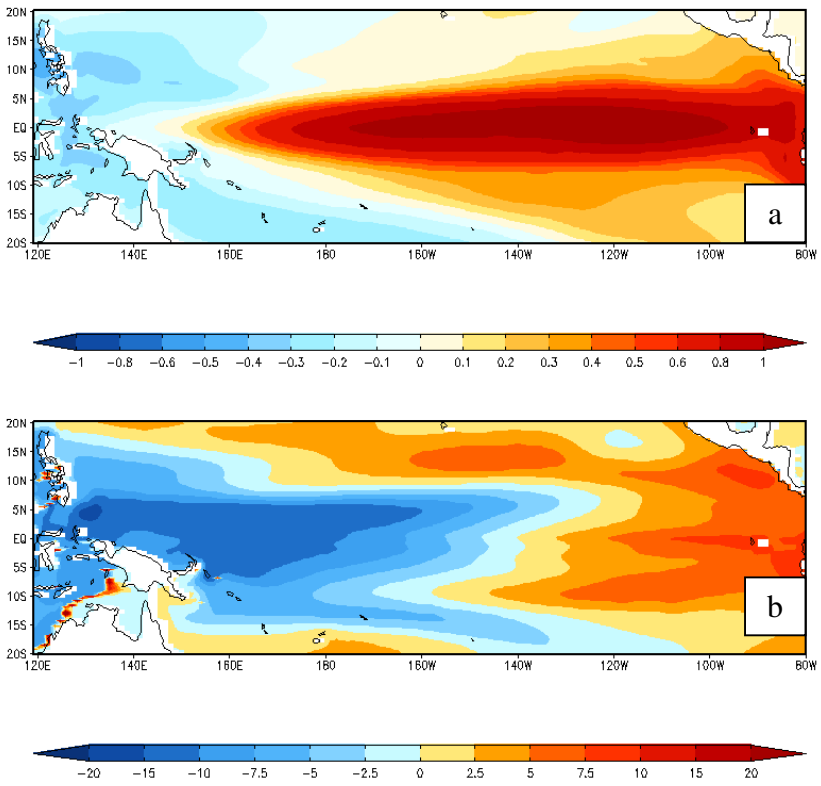

Figure B.16 The first mode of (a) SST and (b) thermocline depth for 600 years of CCSM4

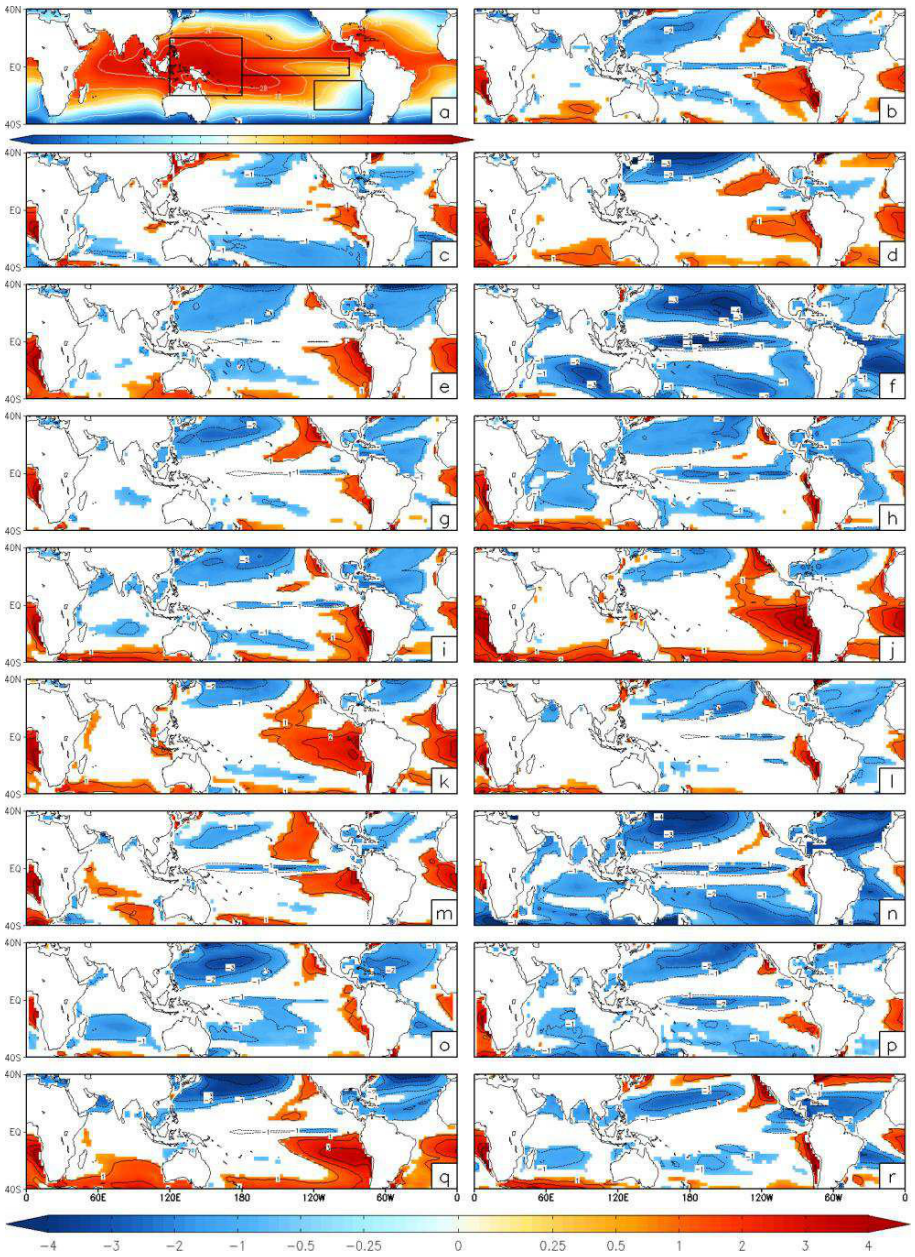

Figure B.17 a) Observed annual mean SST (ERSSTv3b) and annual mean SST errors from, b) BCC-CSM1-1, c) CanESM2, d) CCSM4, e) CNRM-CM5, f) CSIRO-Mk3-6, g) GFDL-CM3, h) GFDL-ESM2G, i) GFDL-ESM2M, j) GISS-E2-H, k) GISS-E2-R, l) HadGEM2-ES, m) INM-CM4, n) IPSL-CM5A-LR, o) MIROC5, p) MPI-ESM-LR, q) MRI-CGCM3, and r) NorESM1-M The units are in °C. Values significant at 95% confidence limit are shaded.

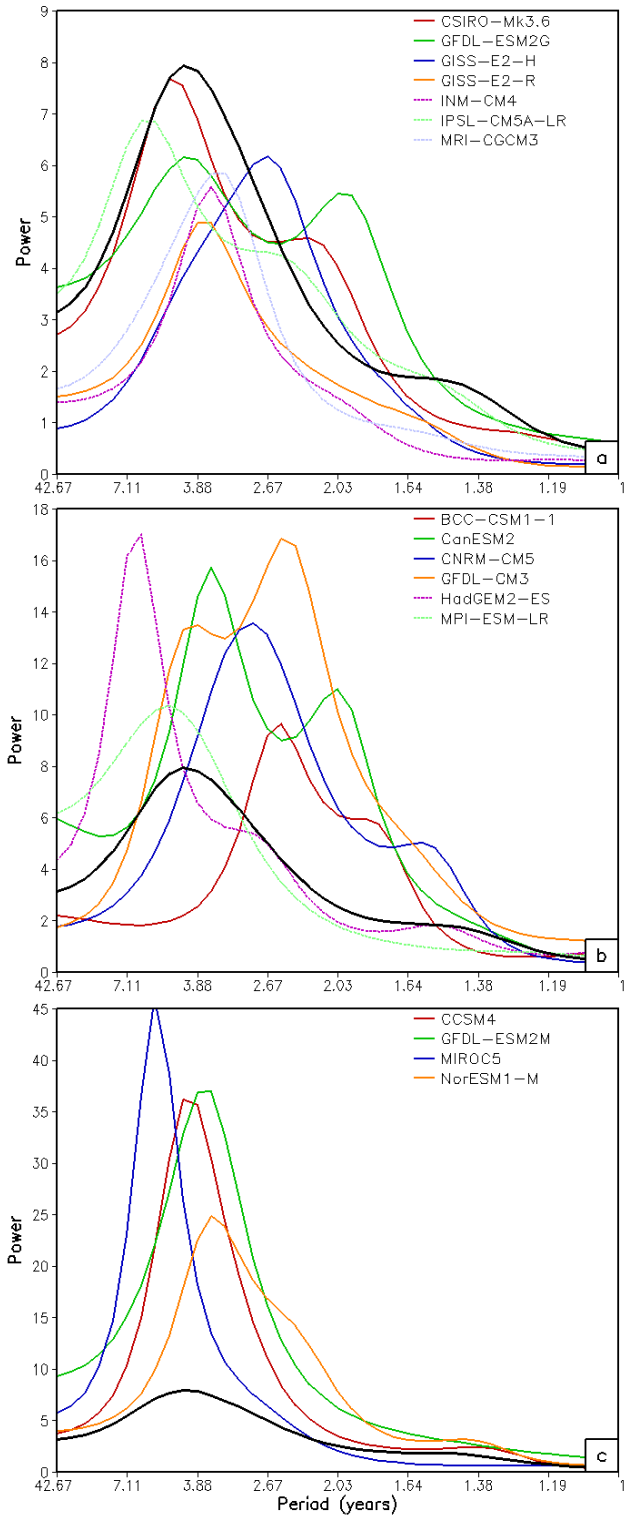

Figure B.18 Sample spectra of the Niño3 SST for selected CMIP5 models and ERSSTv3b (black).

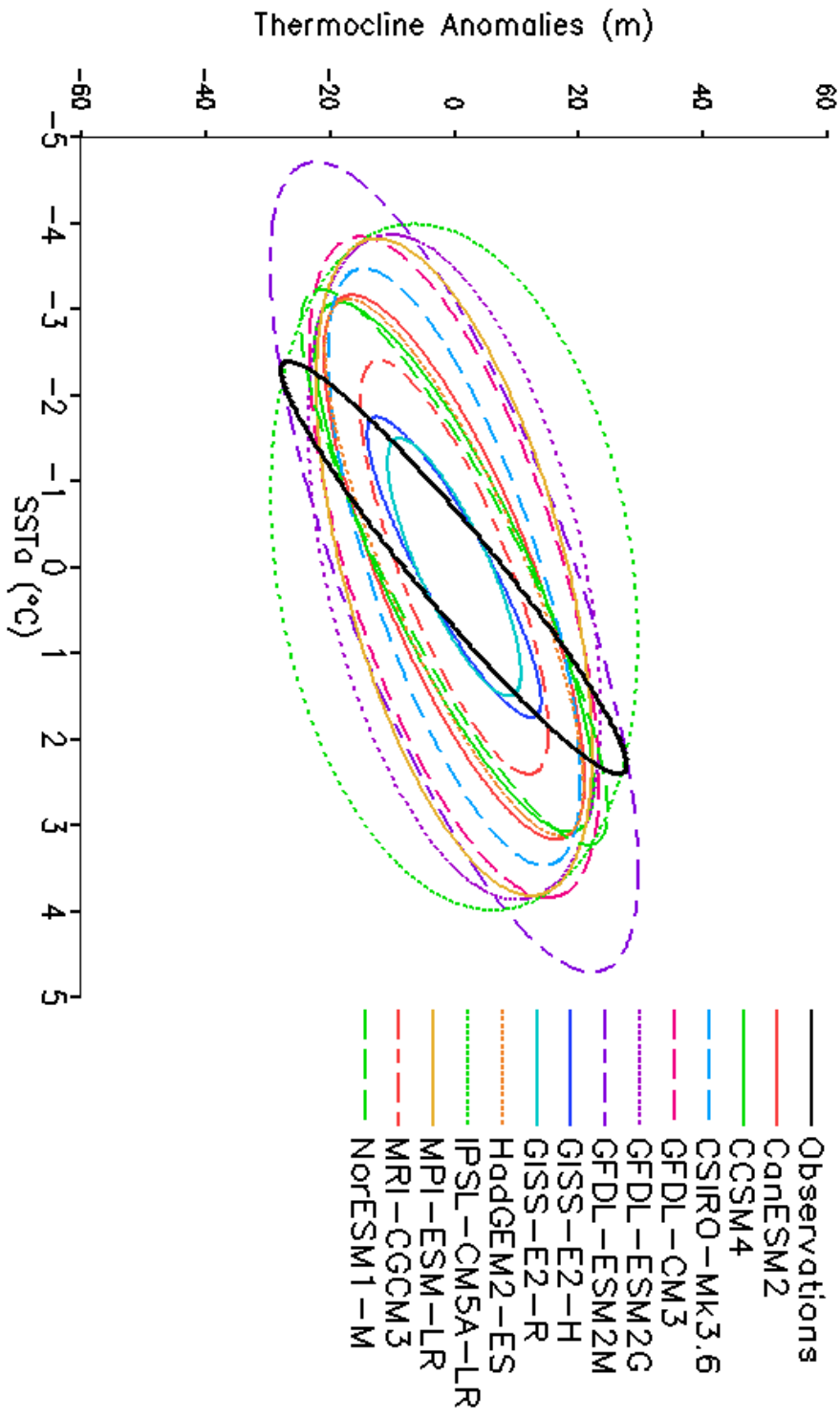

Figure B.19 Ellipse representing the 95 percentile of the scatter of thermocline depth anomalies and SST anomalies over the Niño 3 region overlaid with observations (black).

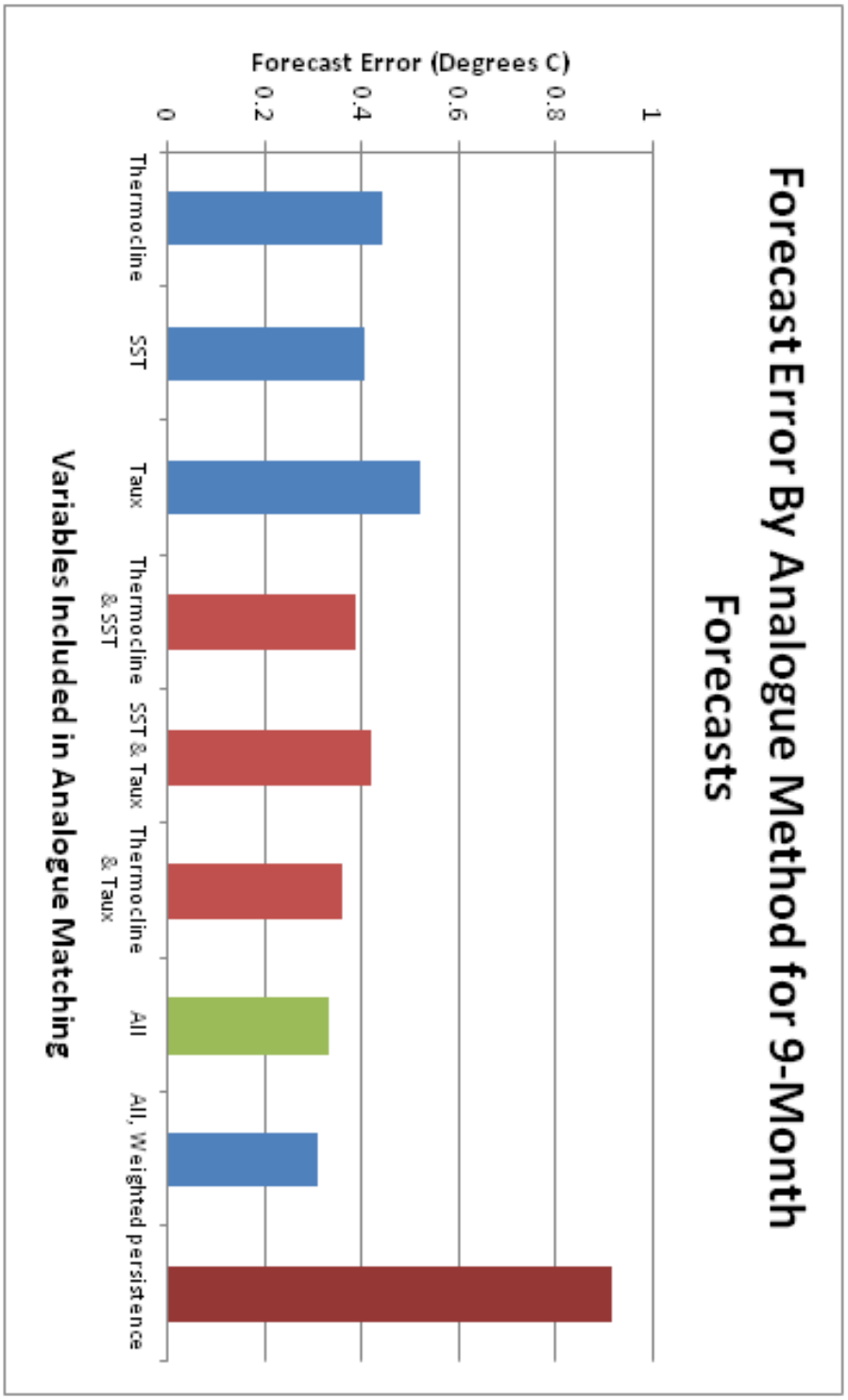

Figure B.20 The forecast error for 9-month Niño3 SSTs using various variable combinations for matching.

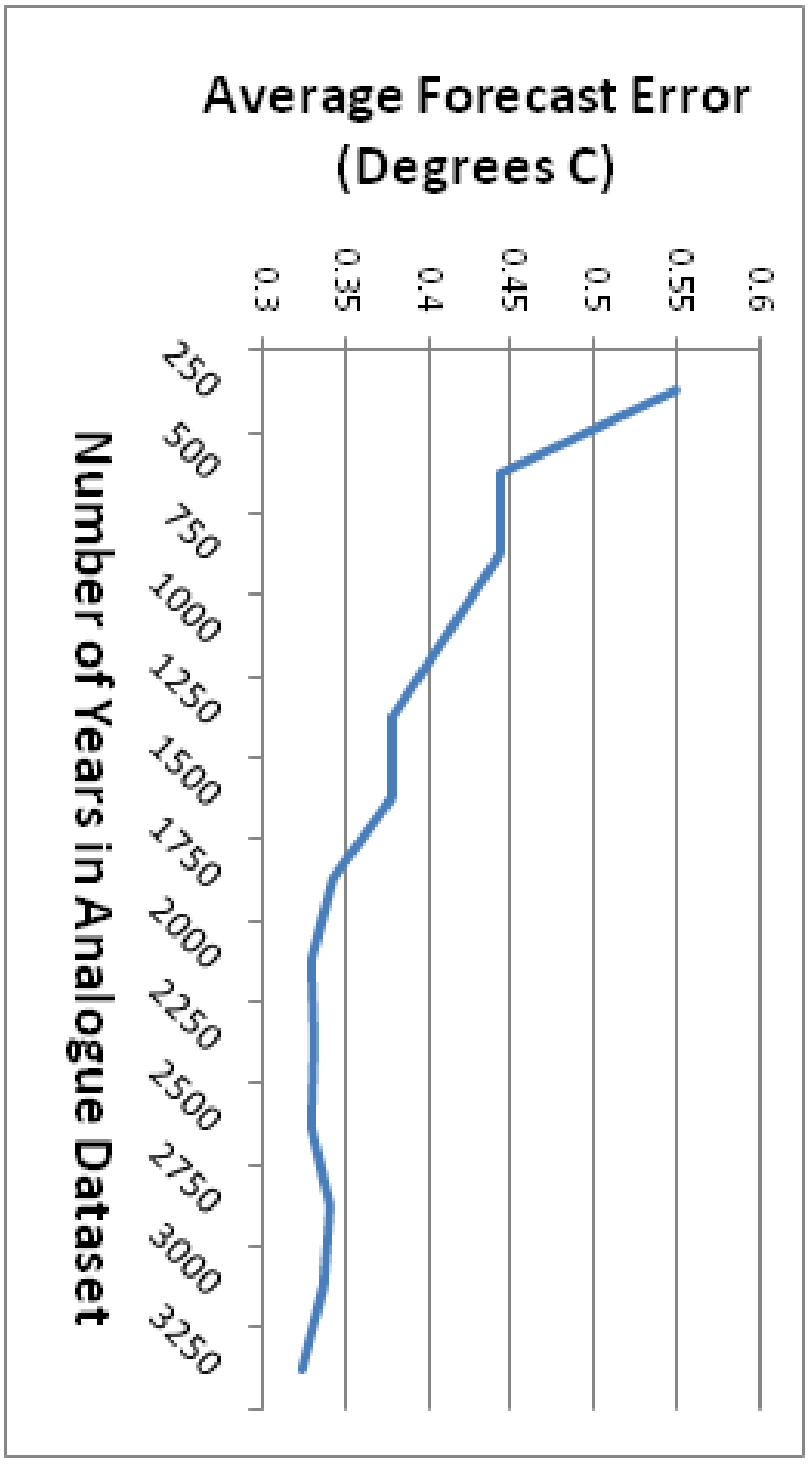

Figure B.21 9-month forecast errors in Niño3 SST by library length for one analogue matching technique.

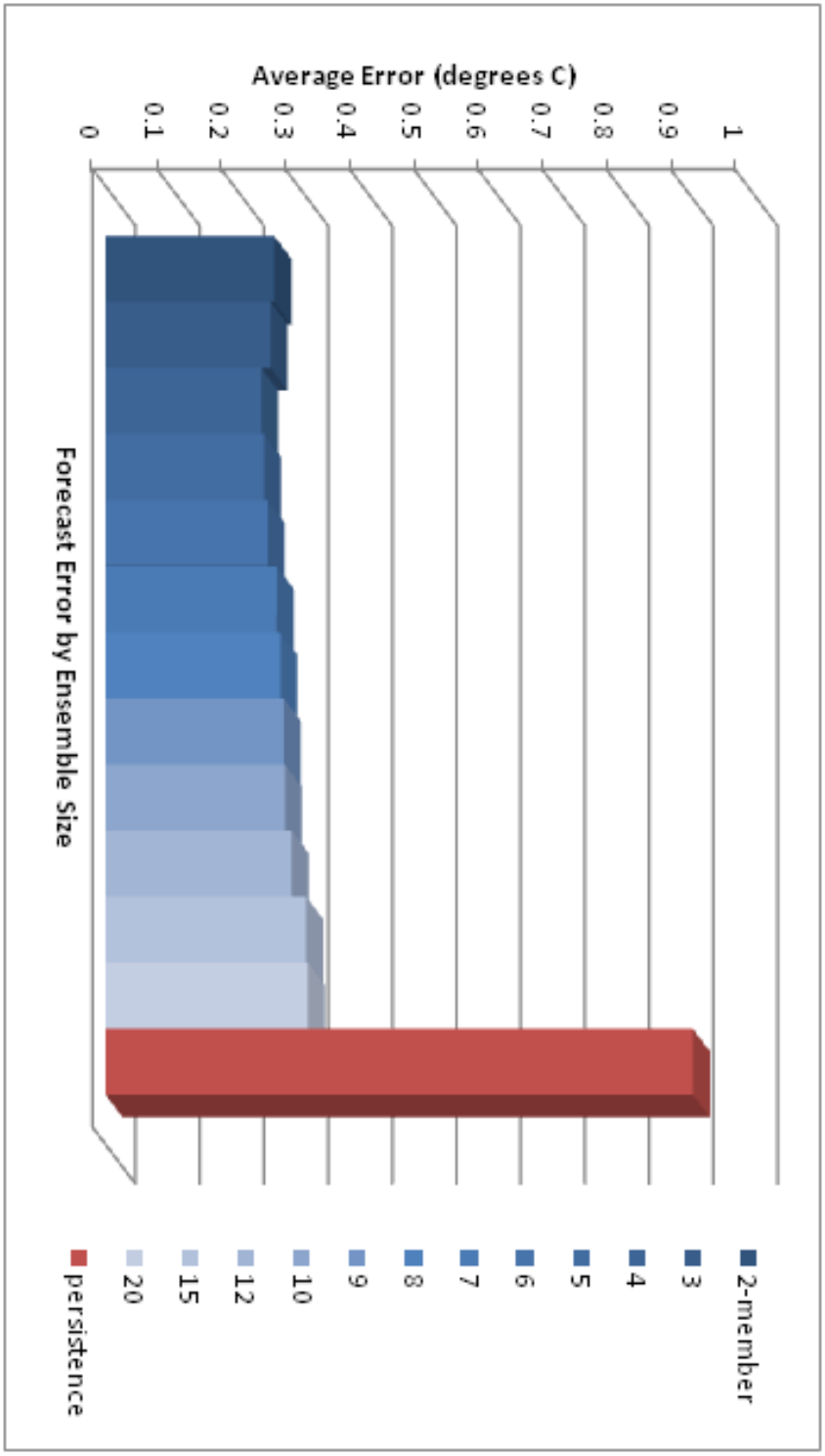

Figure B.22 Forecast error of Niño3 SST for varying ensemble size based on a 5000-year library and a 9-month forecast.

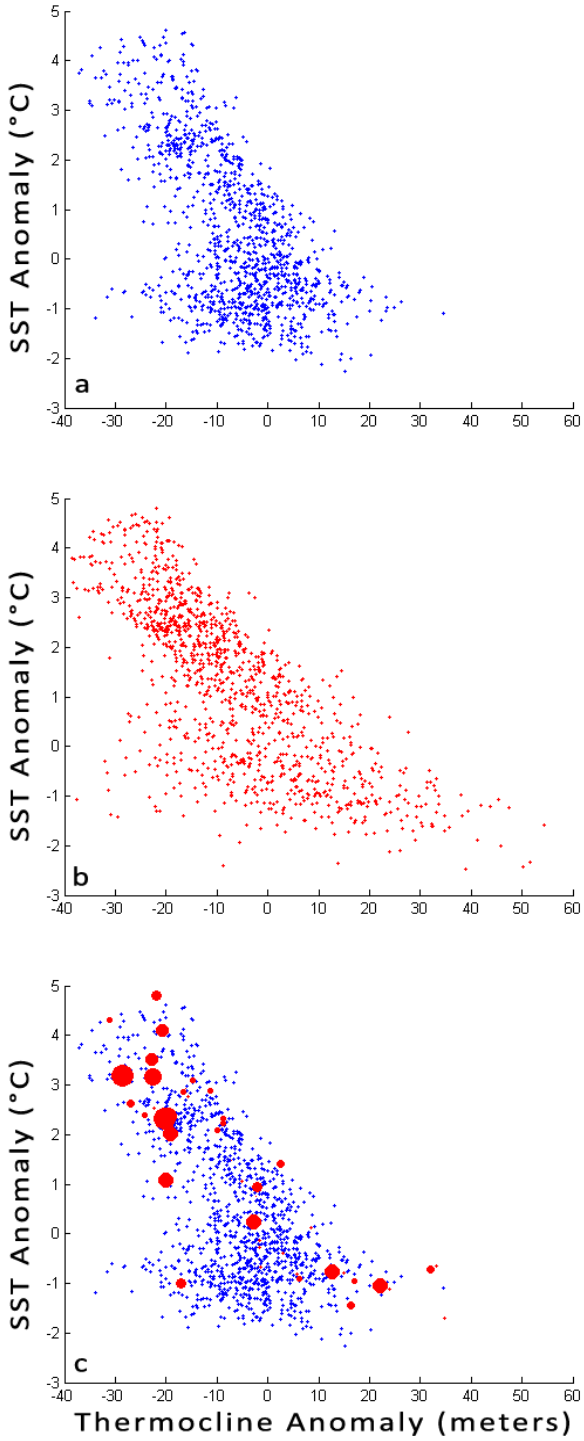

Figure B.23 The distribution of modeled states for a) the Cane-Zebiak model b) the modified Cane-Zebiak model represented by the scatter of Thermocline depth and SST in the Niño 3 region. Panel c) the forecast error for 50 forecasts (larger red circles = greater error) overlaid on panel ‘a’.

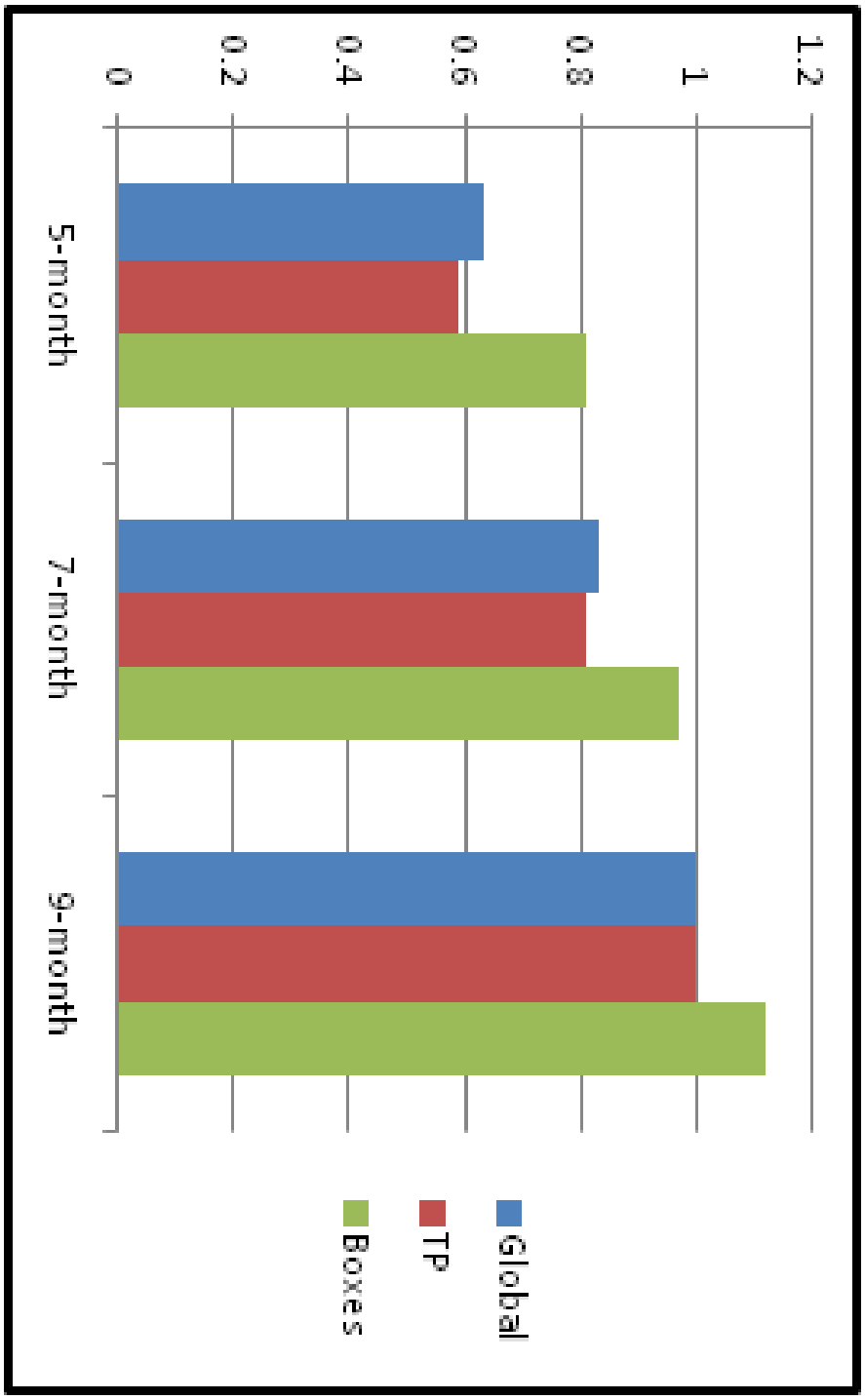

Figure B.24 The average error of 32 hindcasts when analogue matching is done over three different domains: globally (blue), tropical Pacific (red), and small sub regions (green).

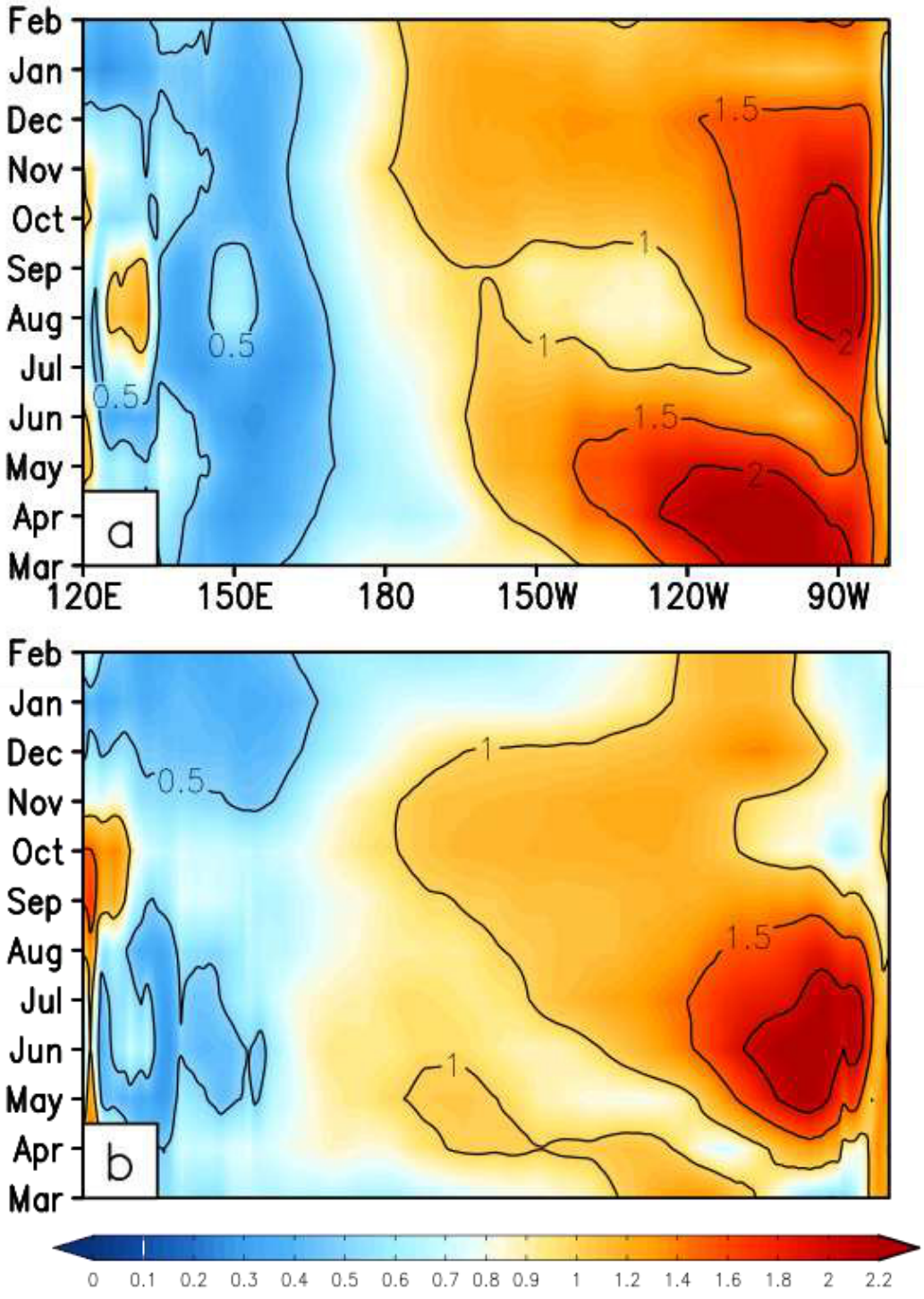

Figure B.25 The seasonal cycle of standard deviation of SSTs for (a) GODAS and (B) CCSM4 hindcast results.

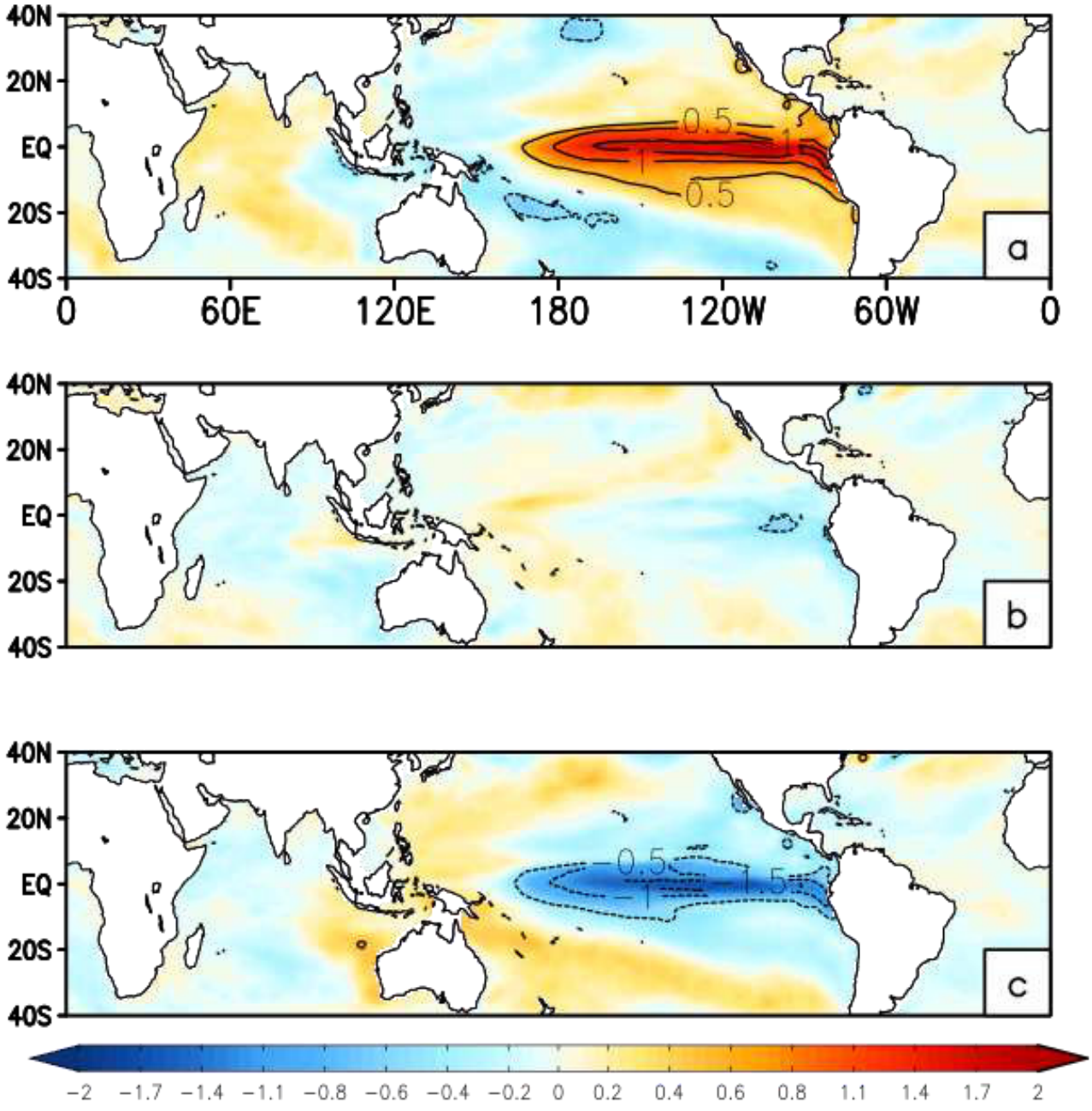

Figure B.26 The composite October-November-December (OND) SSTs of (a) warm, (b) neutral, and (c) cold events from 32 years of GODAS, observational proxy dataset.

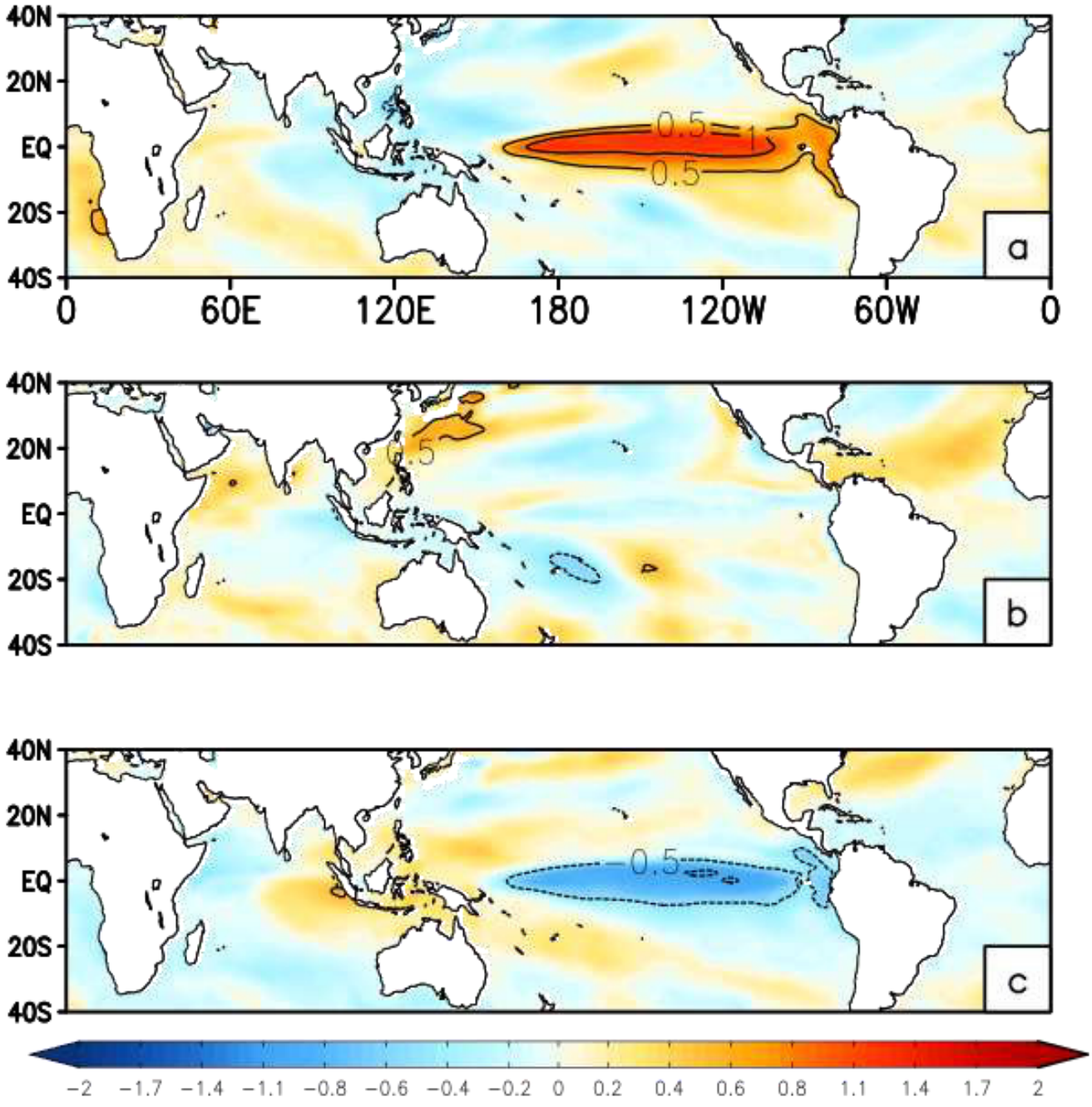

Figure B.27 The composite OND SSTs of (a) warm, (b) neutral, and (c) cold events from 32 years of the 9-month hindcasts.

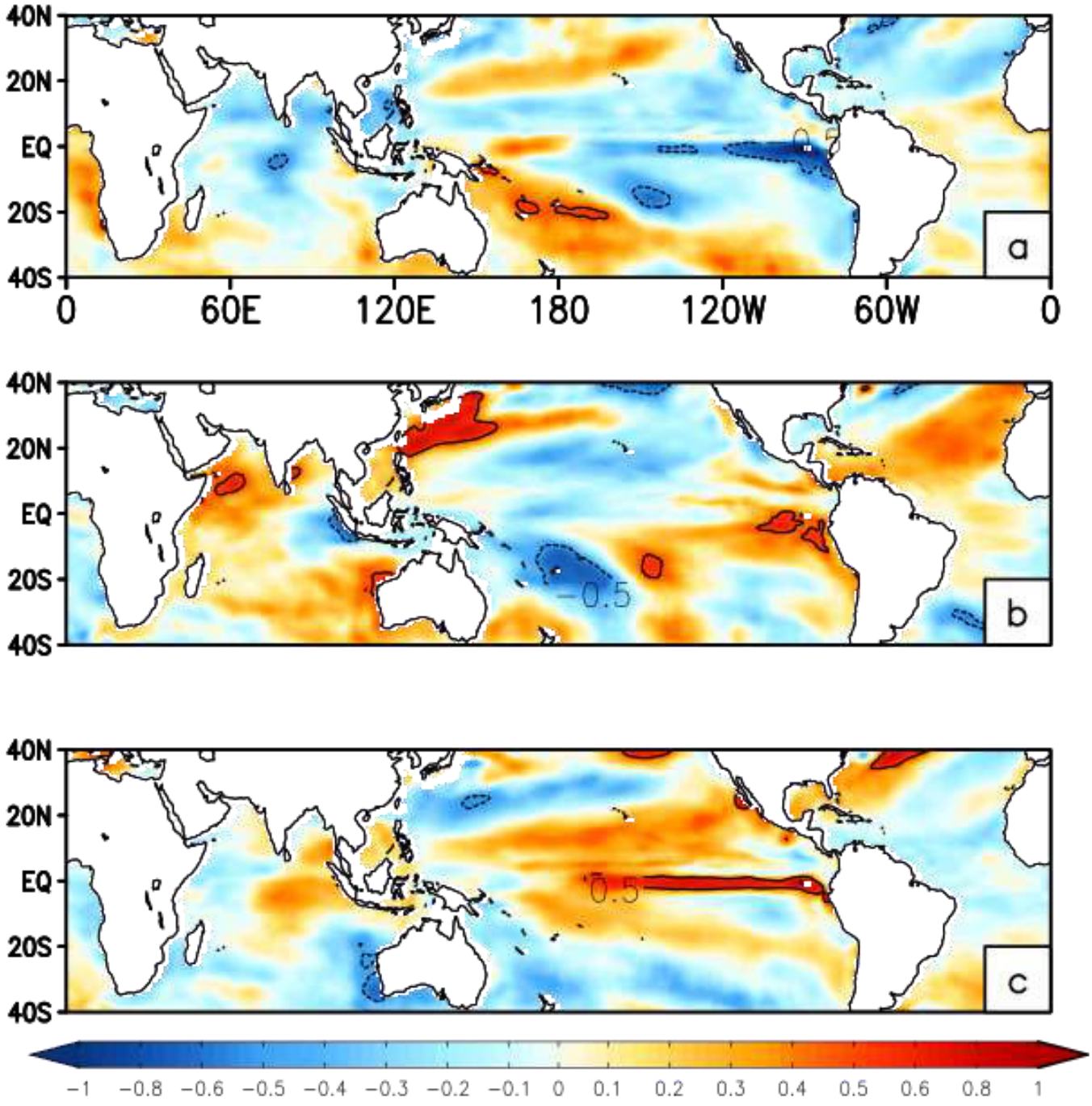

Figure B.28 The OND SSTs differences of (a) warm, (b) neutral, and (c) cold events for hindcasts and GODAS (Figure B.28 minus Figure B.27).

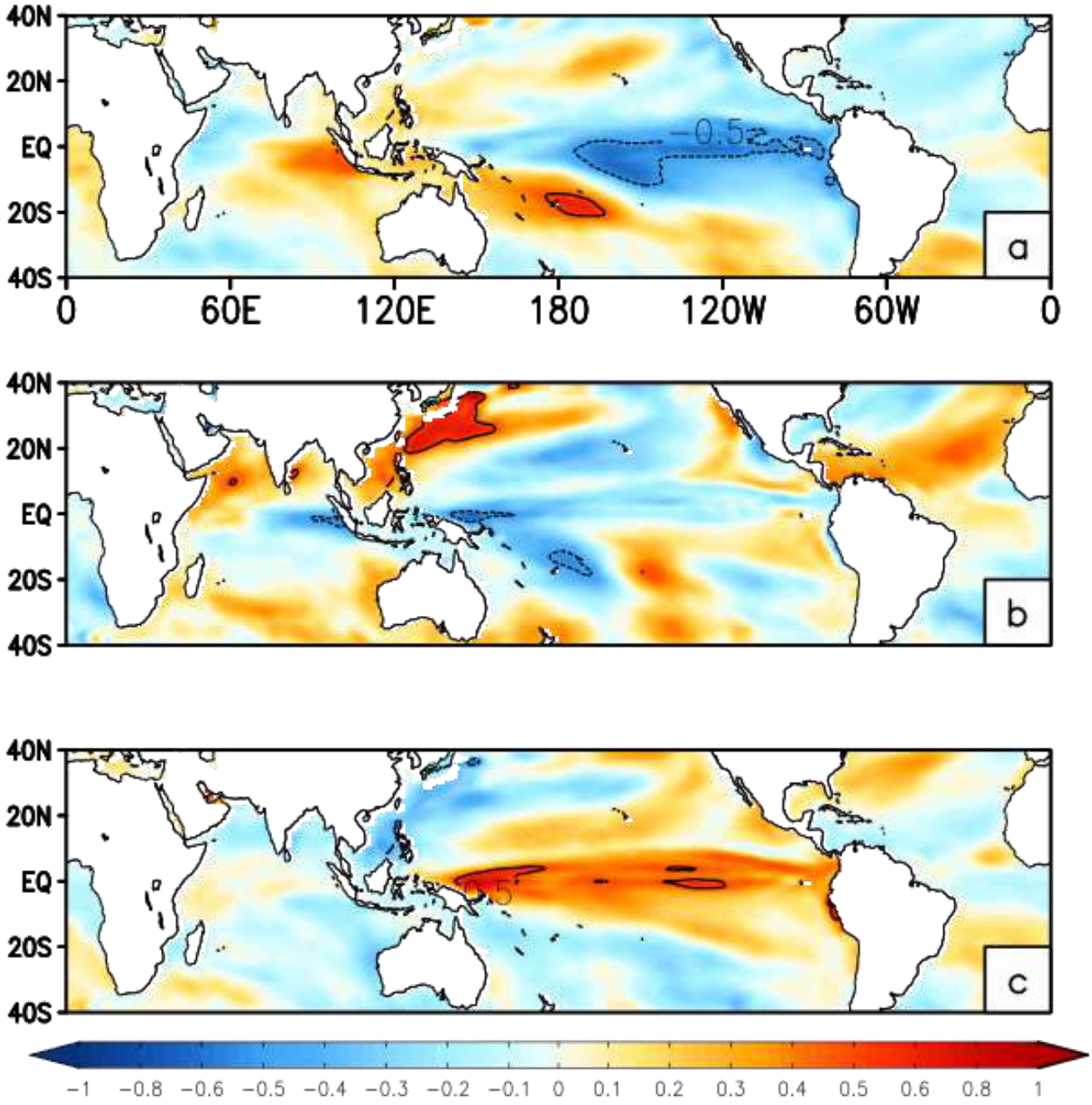

Figure B.29 Same as Figure B.28 but for hindcasts minus simulation. The OND SSTs differences of (a) warm, (b) neutral, and (c) cold events for hindcasts and 600 years of CCSM4 simulation.

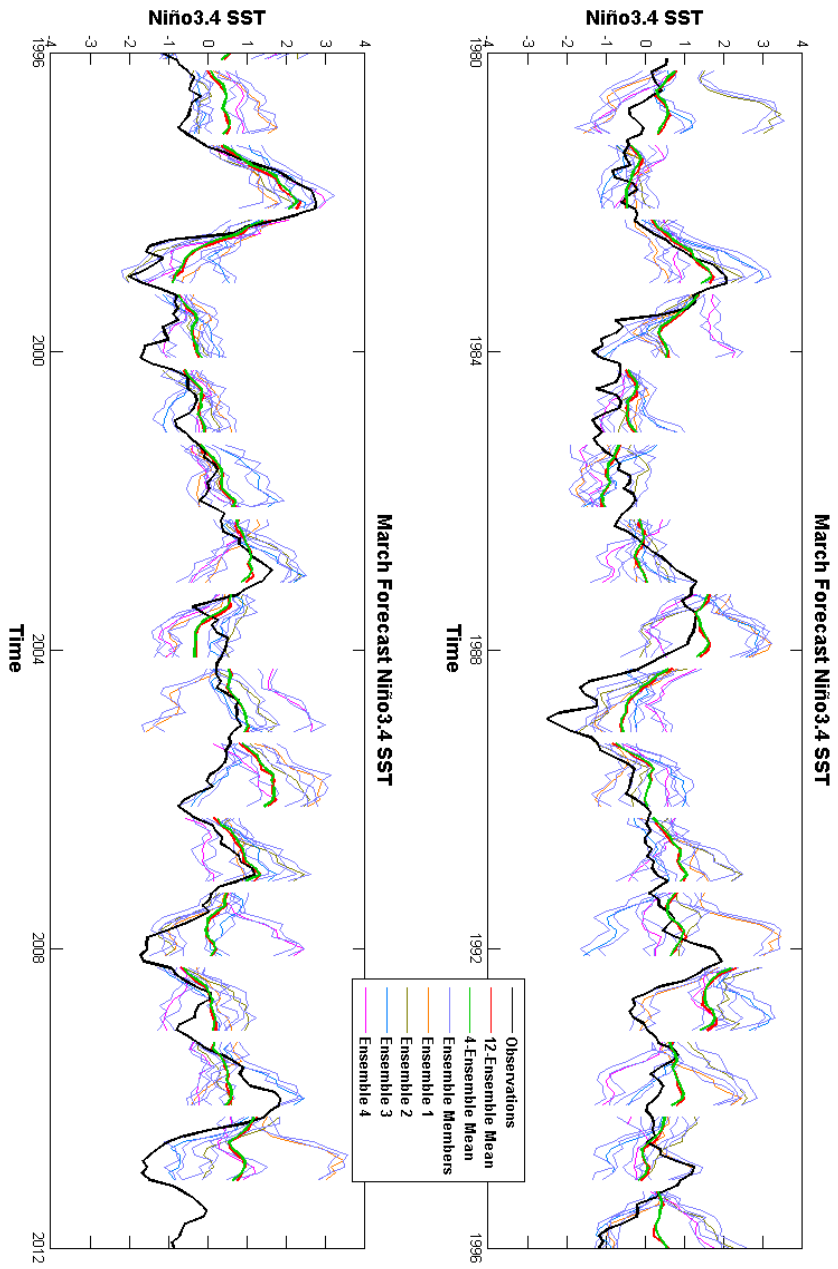

Figure B.30 The observed SSTA of the Niño3.4 region (black) with the hindcast ensemble mean (green), the 4-member mean (red) and the 12 ensemble realizations

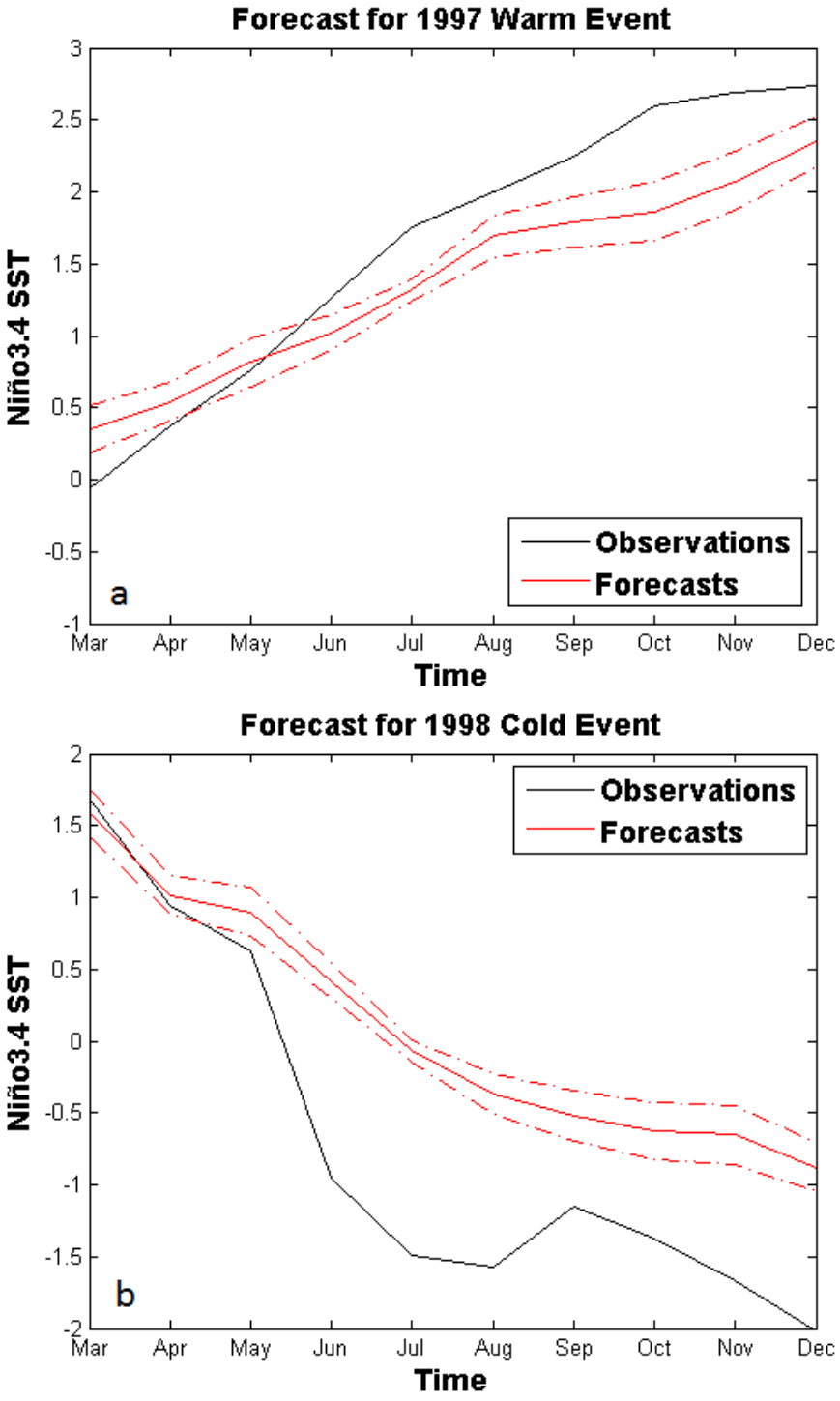

Figure B.31 Observed evolution of the 1997 warm event (a) and the 1998 cold event (b) as observed (black) with 12-member hindcast ensemble (red) and ensemble spread (red, dashed).

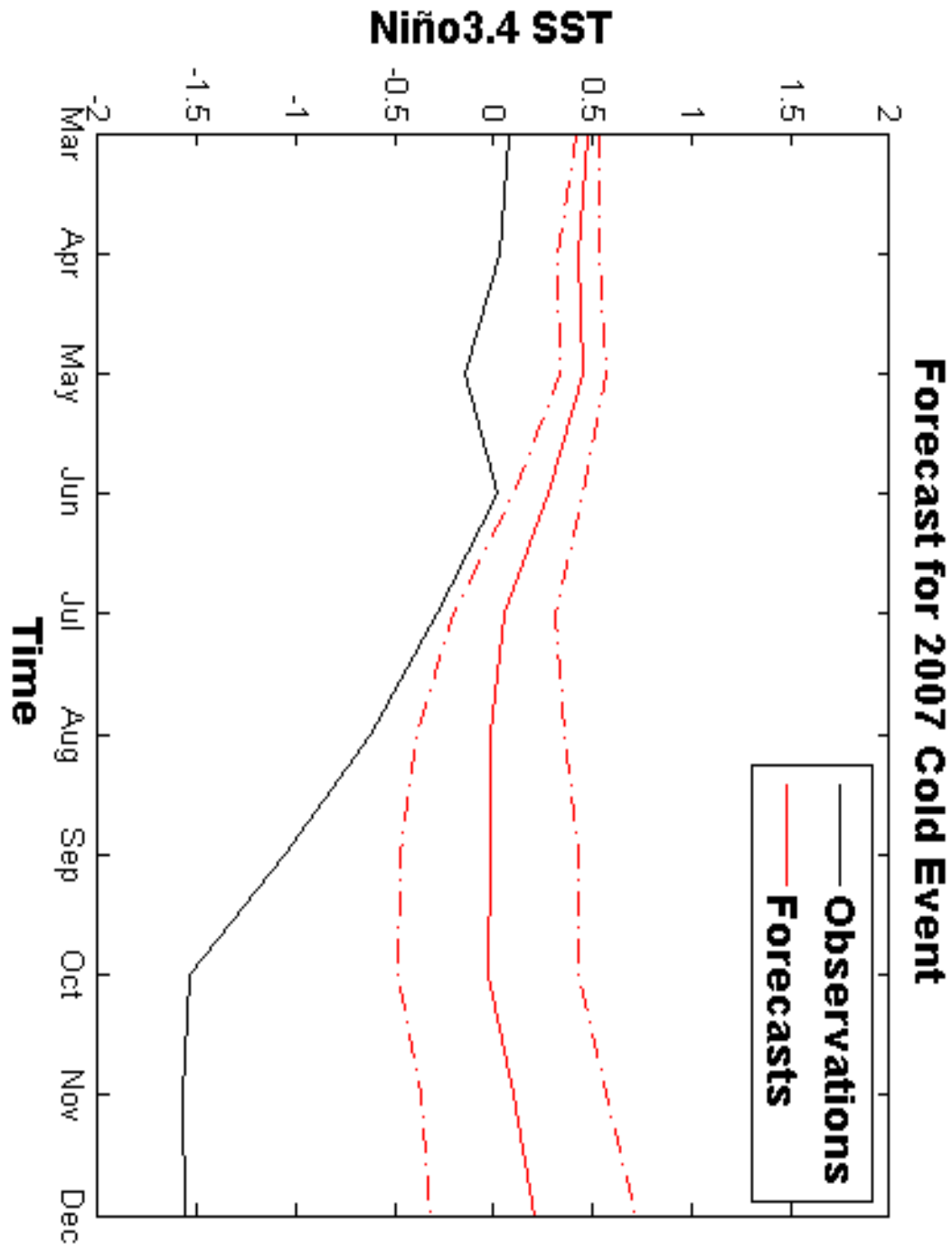

Figure B.32 Observed evolution of the 2007 cold event as observed (black) with 12-member hindcast ensemble (red) and ensemble spread (red, dashed).

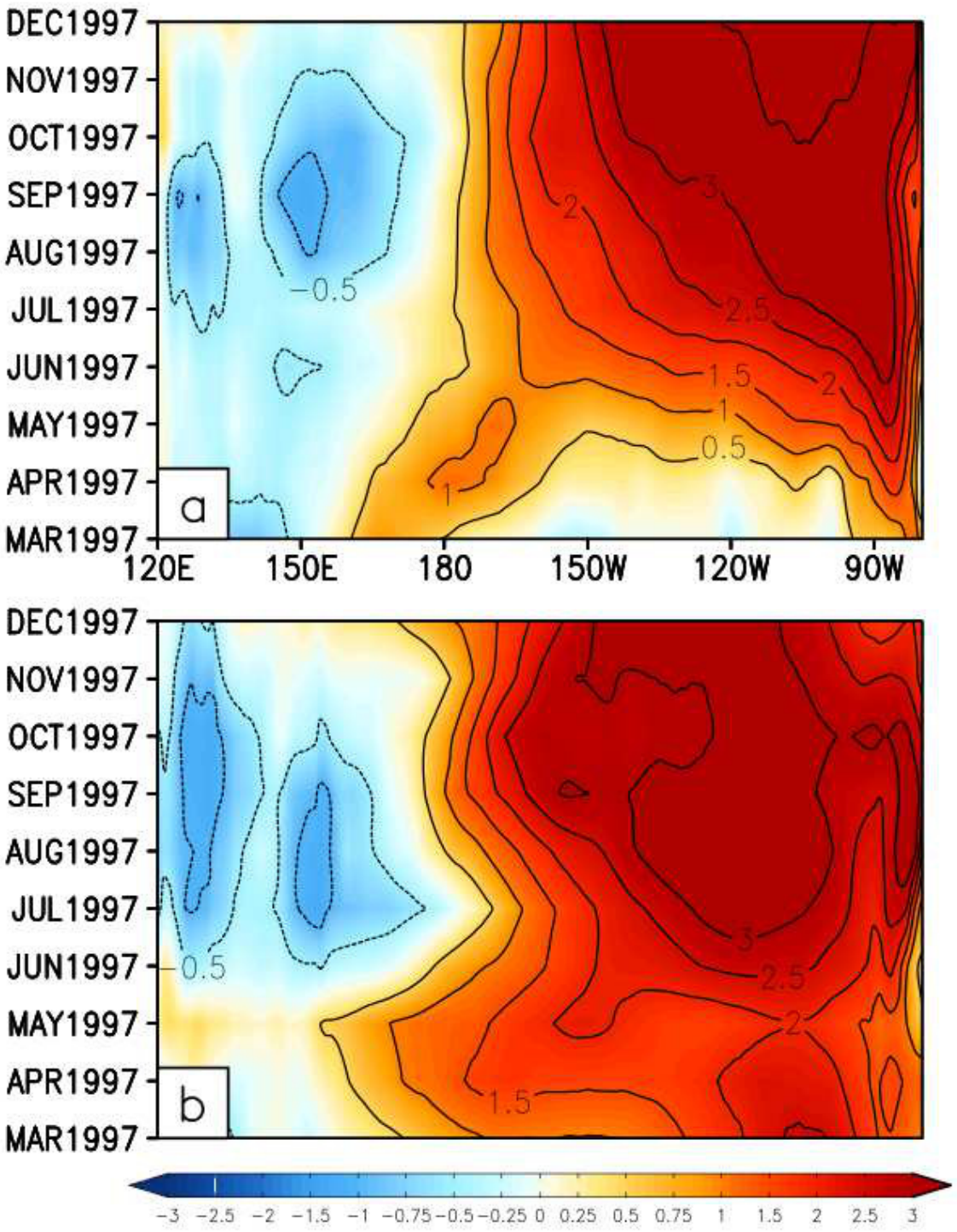

Figure B.33 Time-evolution of the equatorial (5S-5N) Pacific SSTAs of the 1997 warm event as depicted by (a) GODAS and (b) CCSM4 hindcasts

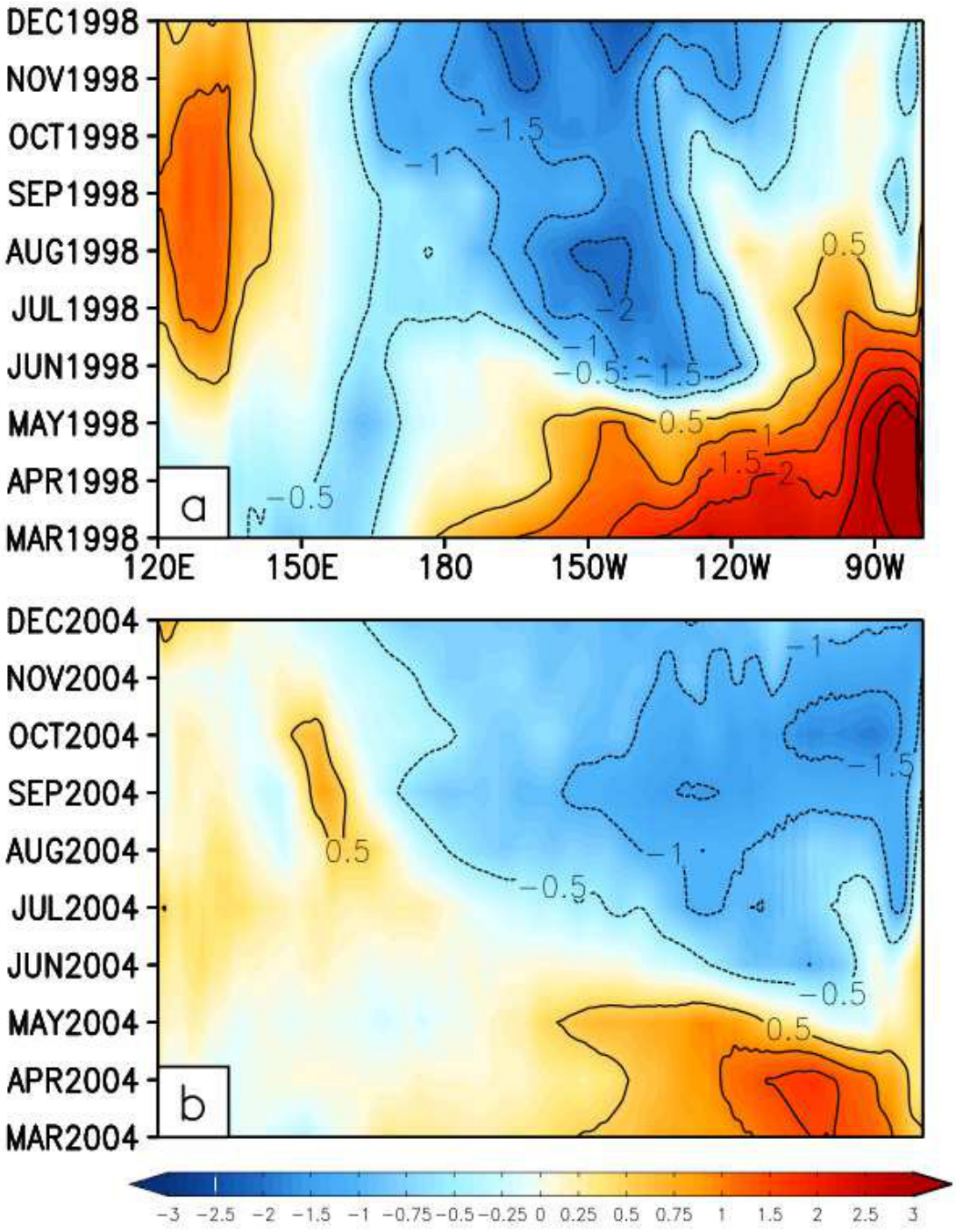

Figure B.34 Time-evolution of the equatorial (5S-5N) Pacific SSTAs of the 1998 cold event as depicted by (a) GODAS and (b) CCSM4 hindcasts

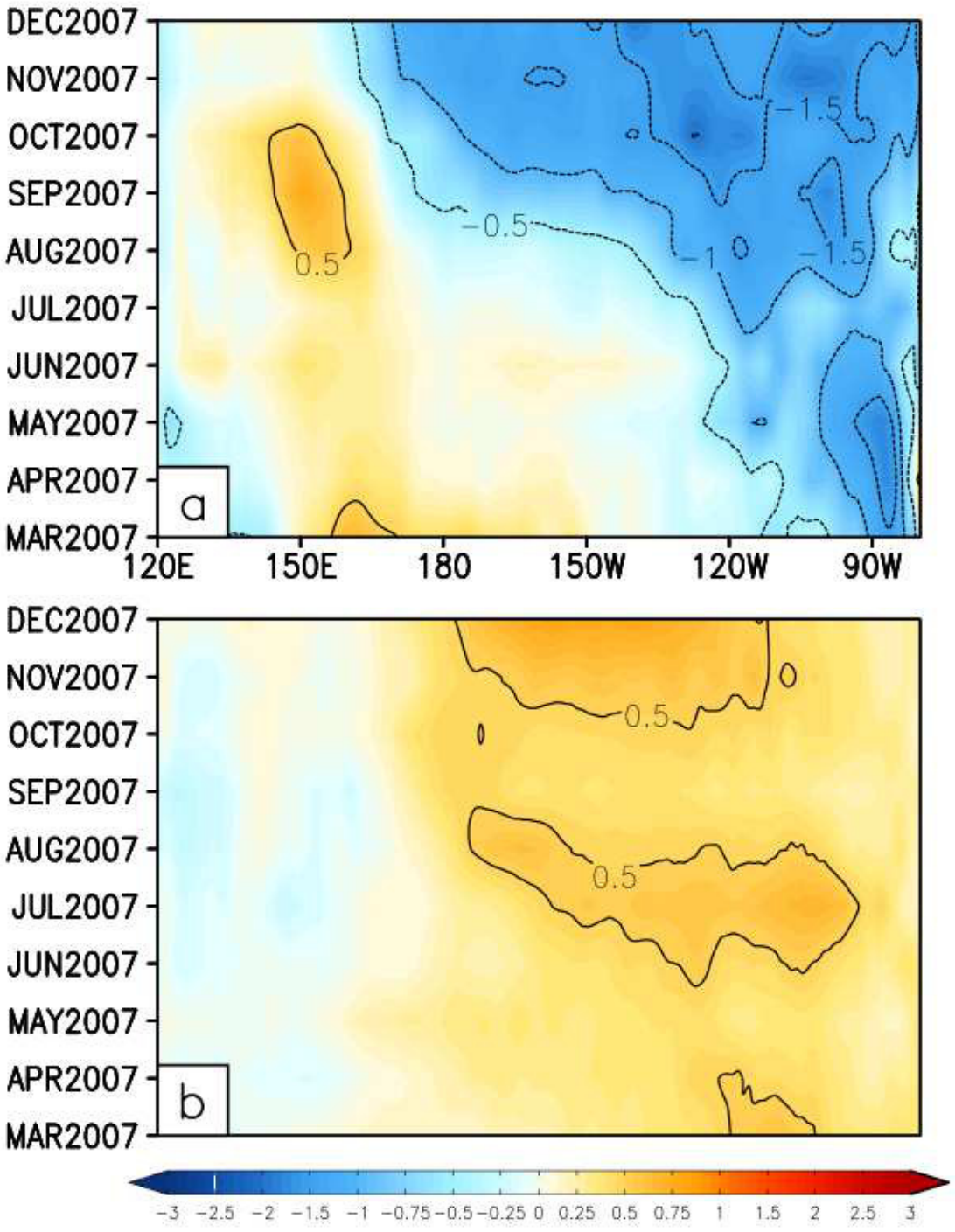

Figure B.35 Time-evolution of the equatorial (5S-5N) Pacific SSTAs of the 2007 cold event as depicted by (a) GODAS and (b) CCSM4 hindcasts

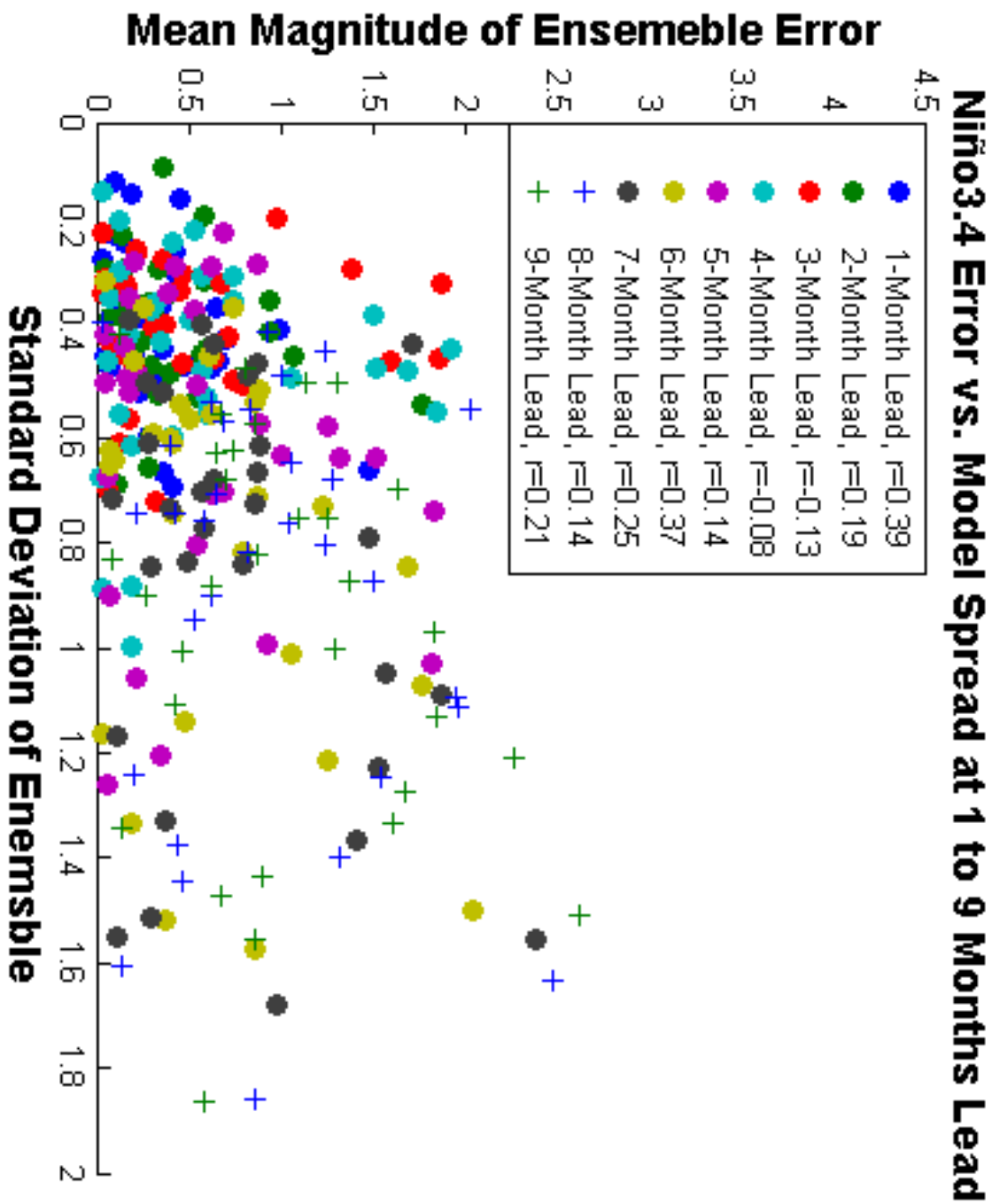

Figure B.36 Scatter of Standard deviation of the 12-member ensemble versus the mean magnitude of ensemble error

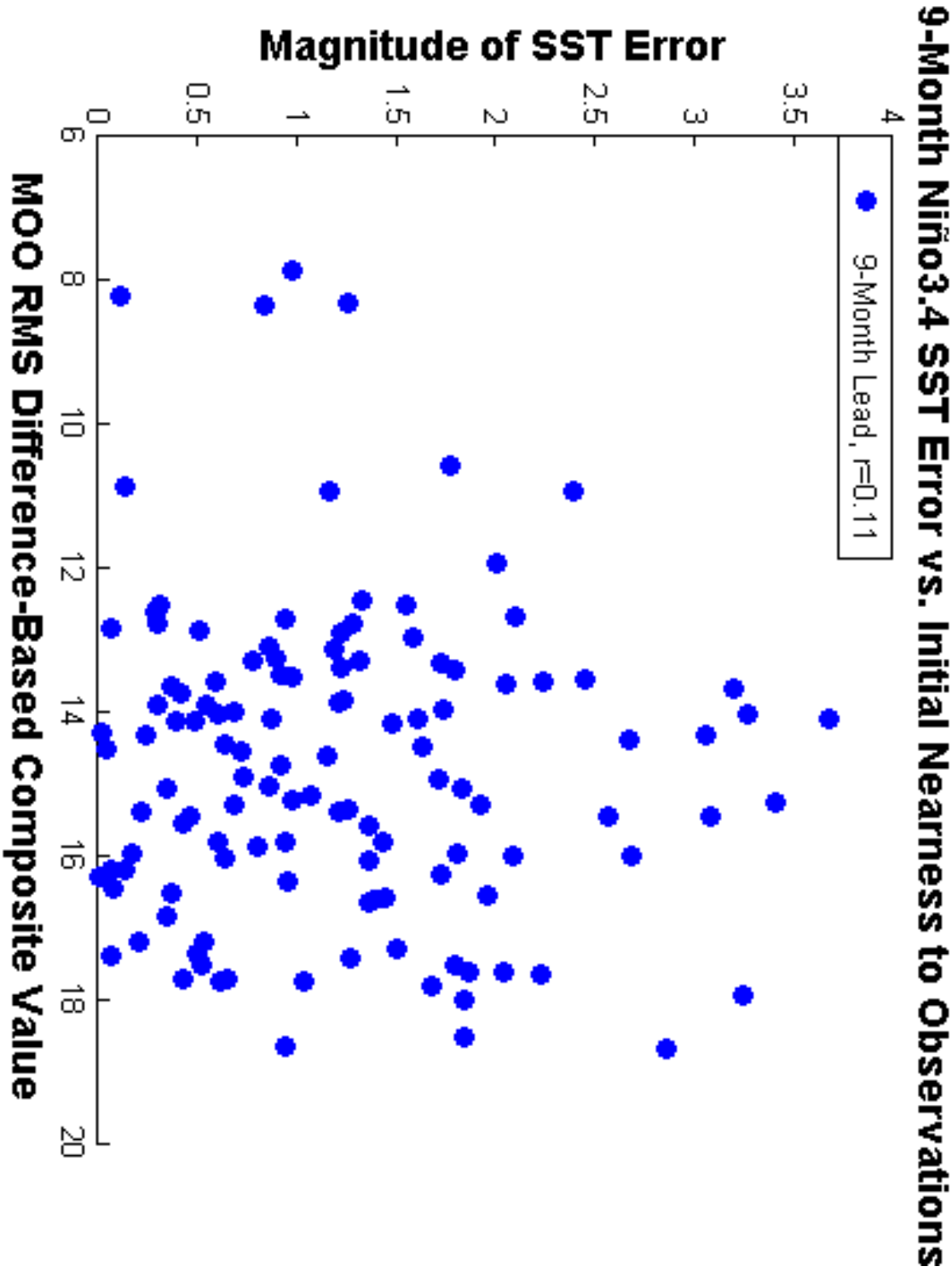

Figure B.37 Scatter diagram of the MOO values of the chosen analogues versus the magnitude of the 9-month forecast errors

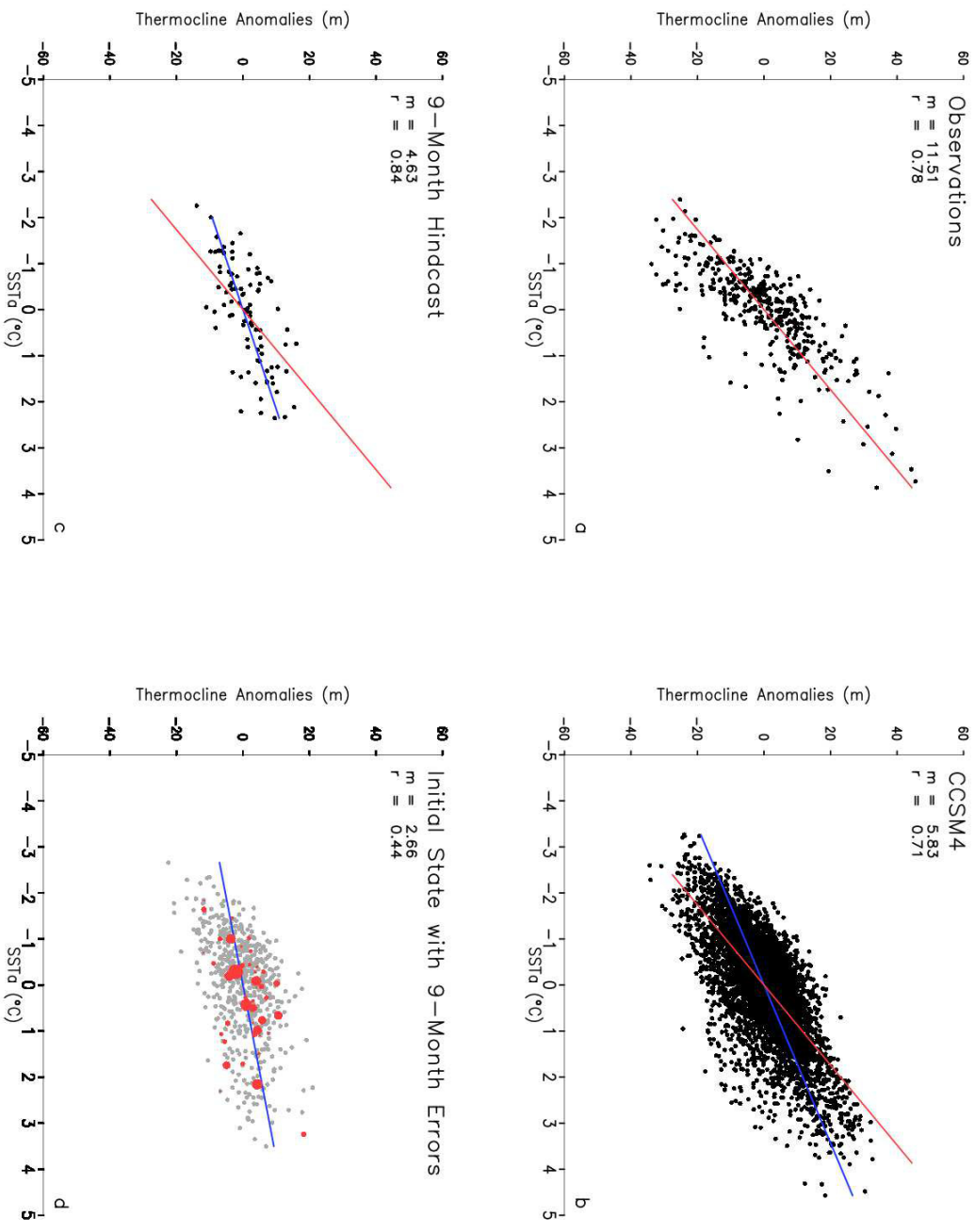

Figure B.38 Scatter plot of the thermocline depth anomalies versus the SSTAs for (a) GODAS, (b) 600-year simulation of CCSM4, (c) 9-month hindcasts (d) 600-year march states with hindcast initial states in red (error of hindcast represented by size of red dot).

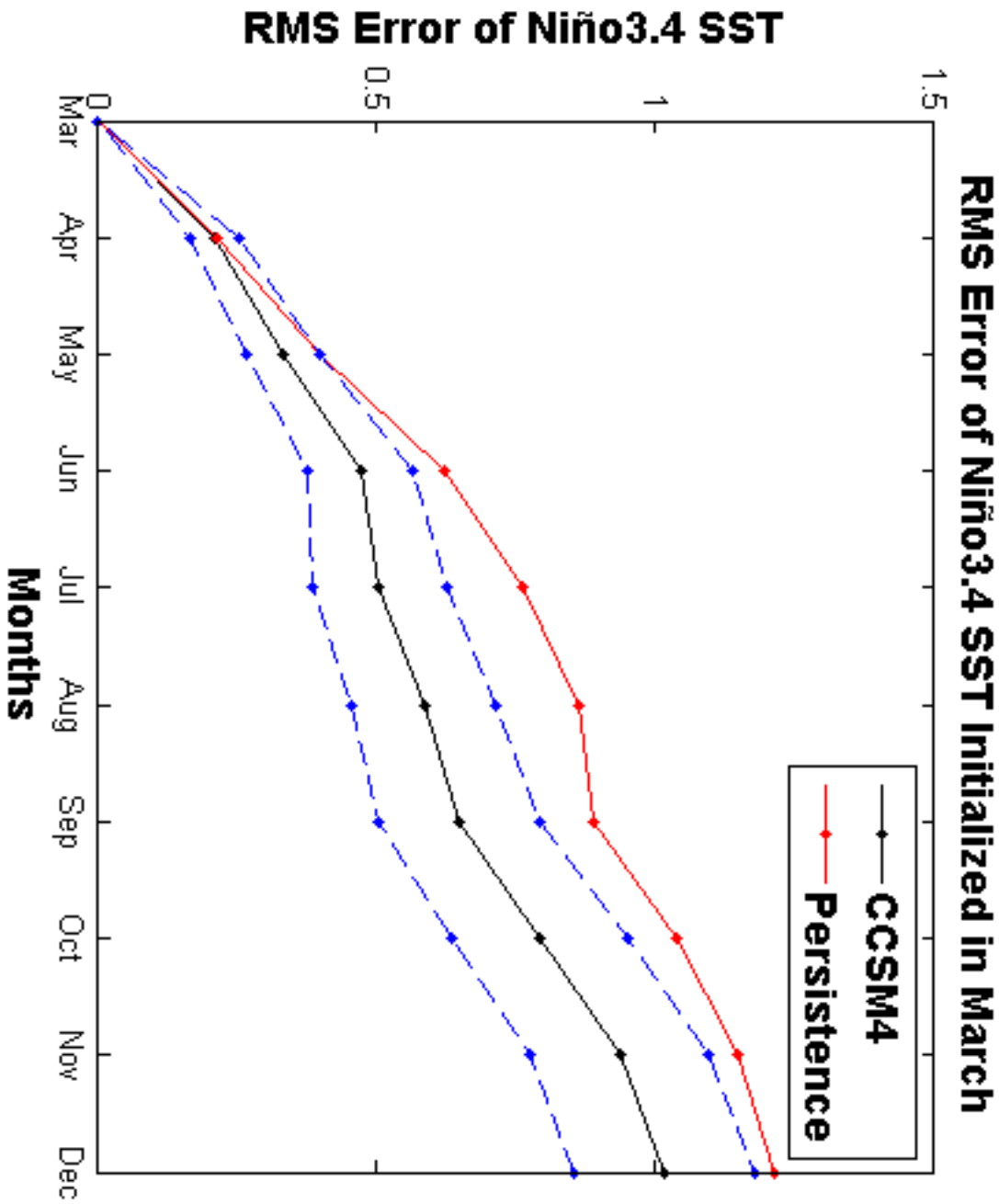

Figure B.39 RMS Errors of Niño 3.4 SST for the 12-member ensemble hindcast initialized in March.

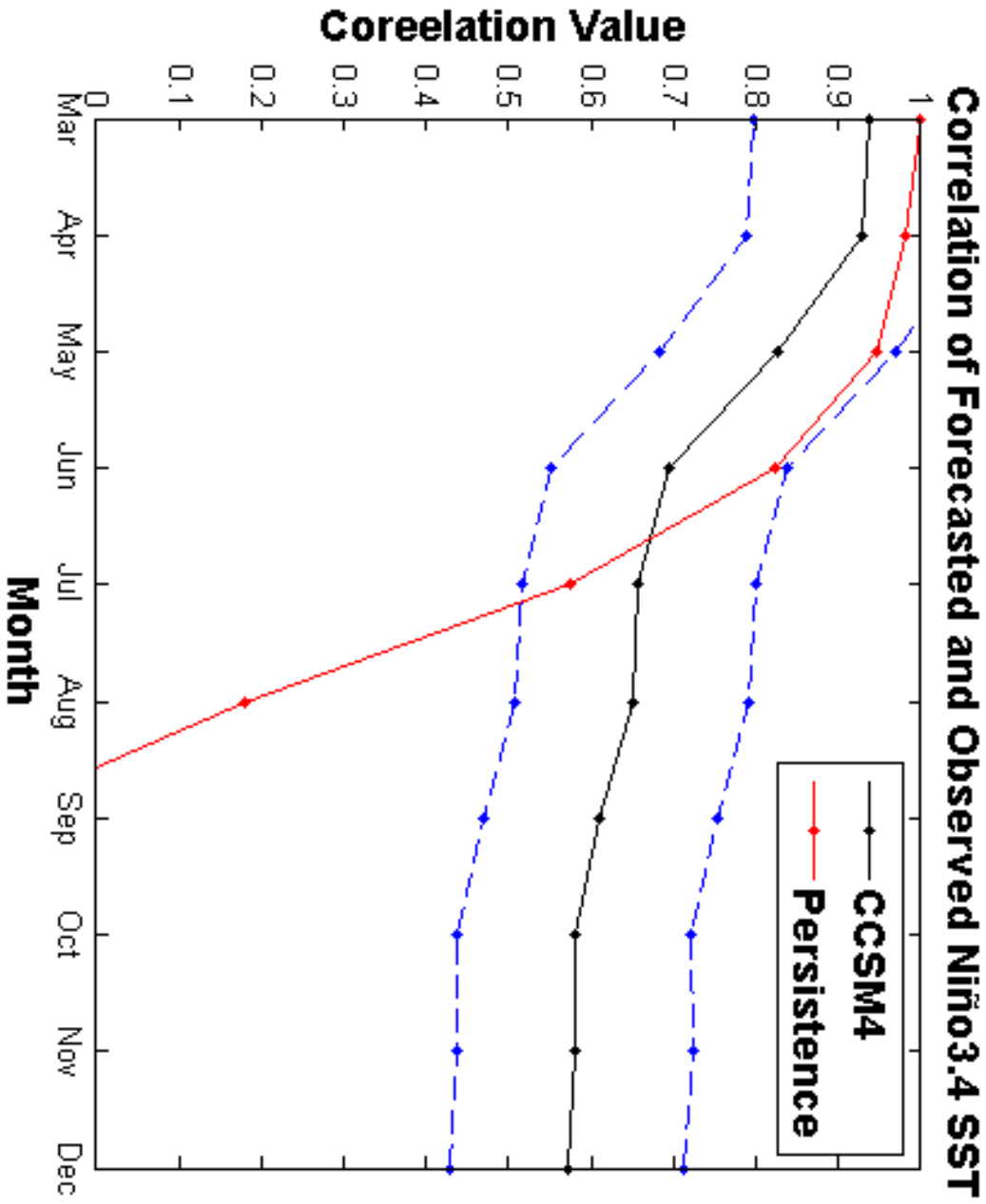

Figure B.40 Correlation of Niño 3.4 SST from 12-member ensemble initialized in March with GODAS.

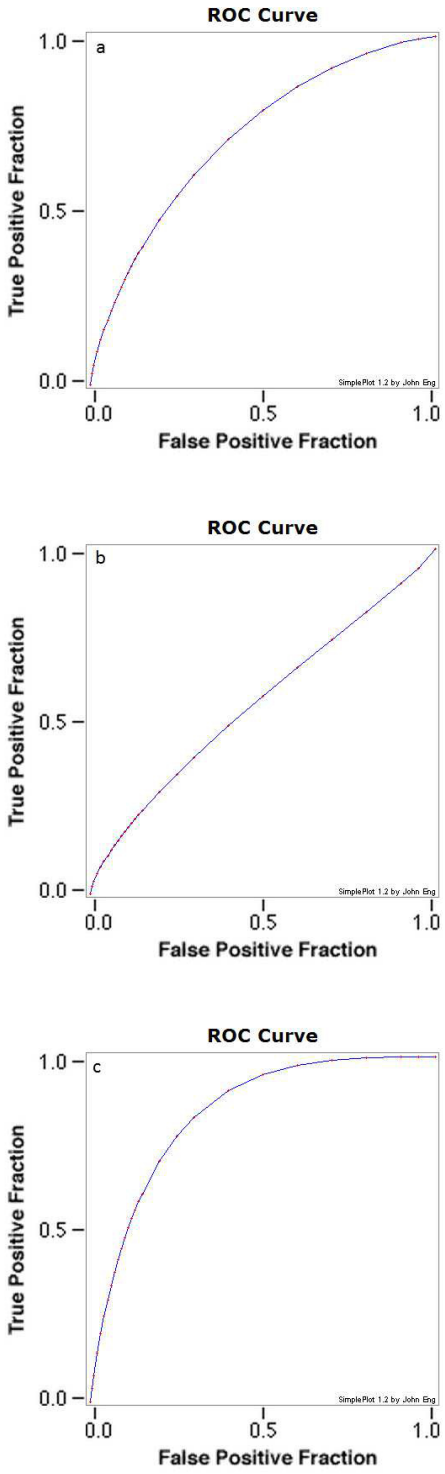

Figure B.41 Relative Operative Characteristic Curves for (a) warm, (b) neutral, and (c) cold events for forecasts initialized in March.

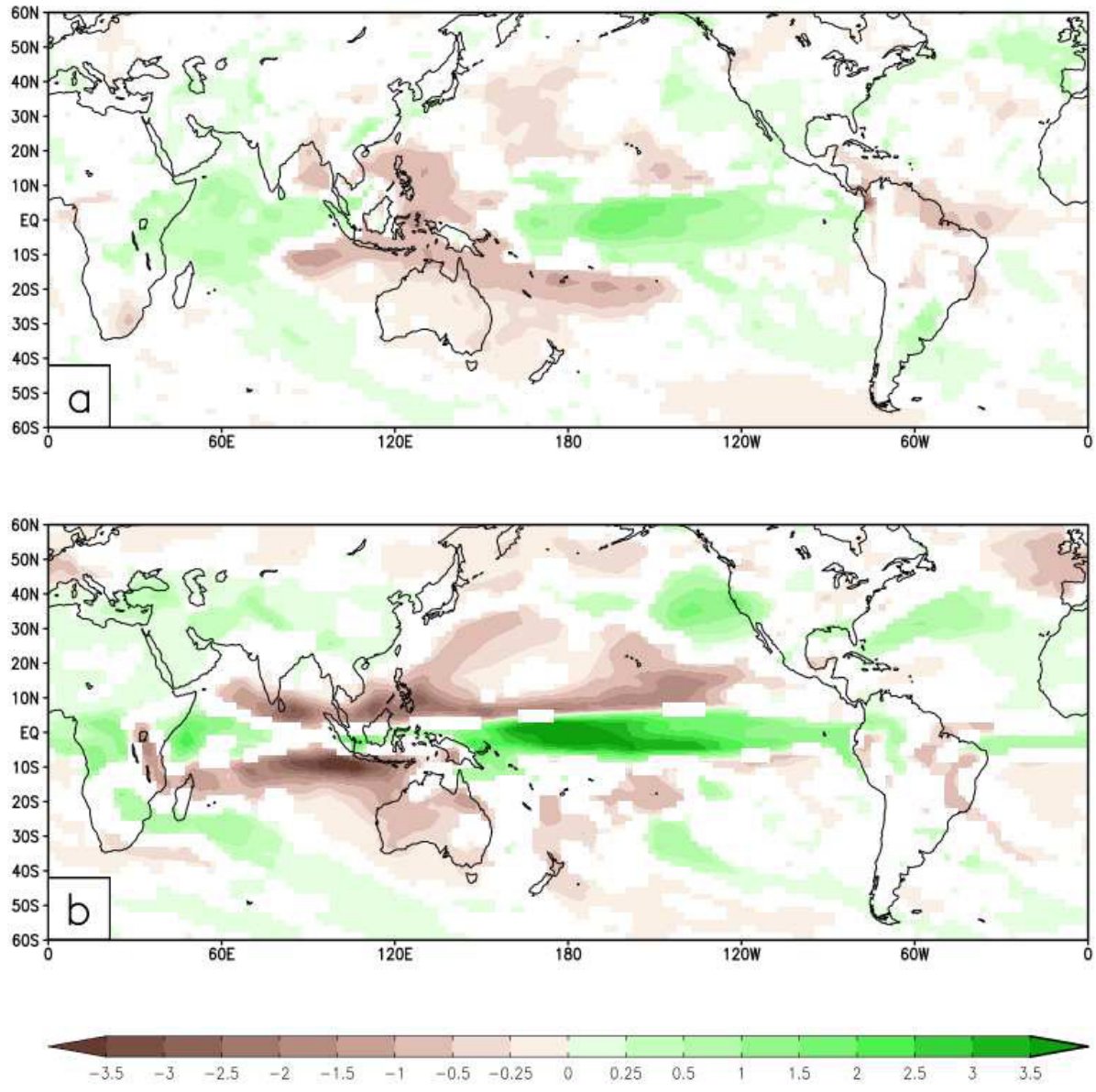

Figure B.42 Regression of Niño3.4 SST with the global precipitation rates for OND from (a) CMAP and (b) hindcast.

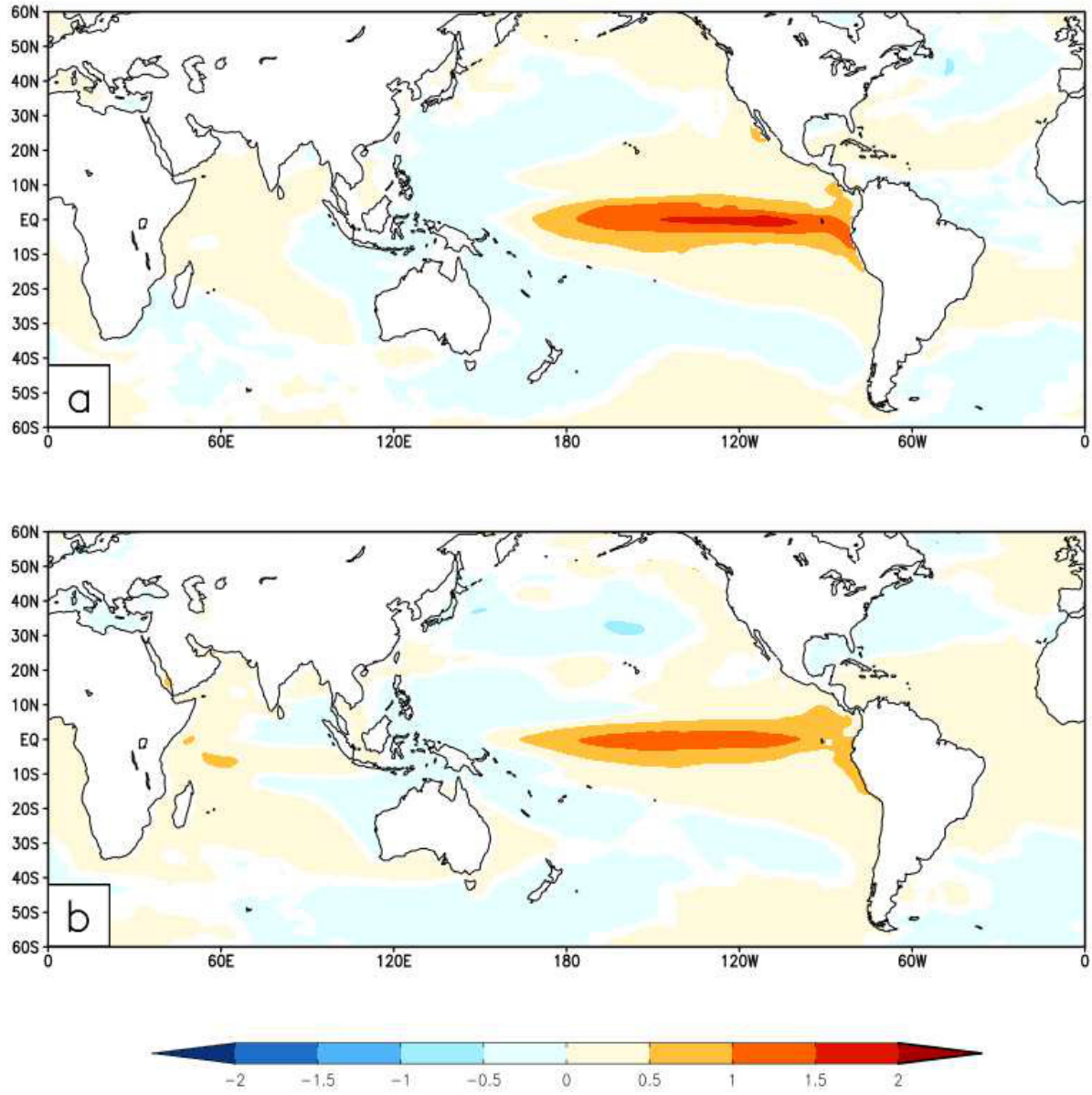

Figure B.43 Regression of Niño3.4 SST with the global ocean temperatures for OND from (a) GODAS and (b) hindcast.

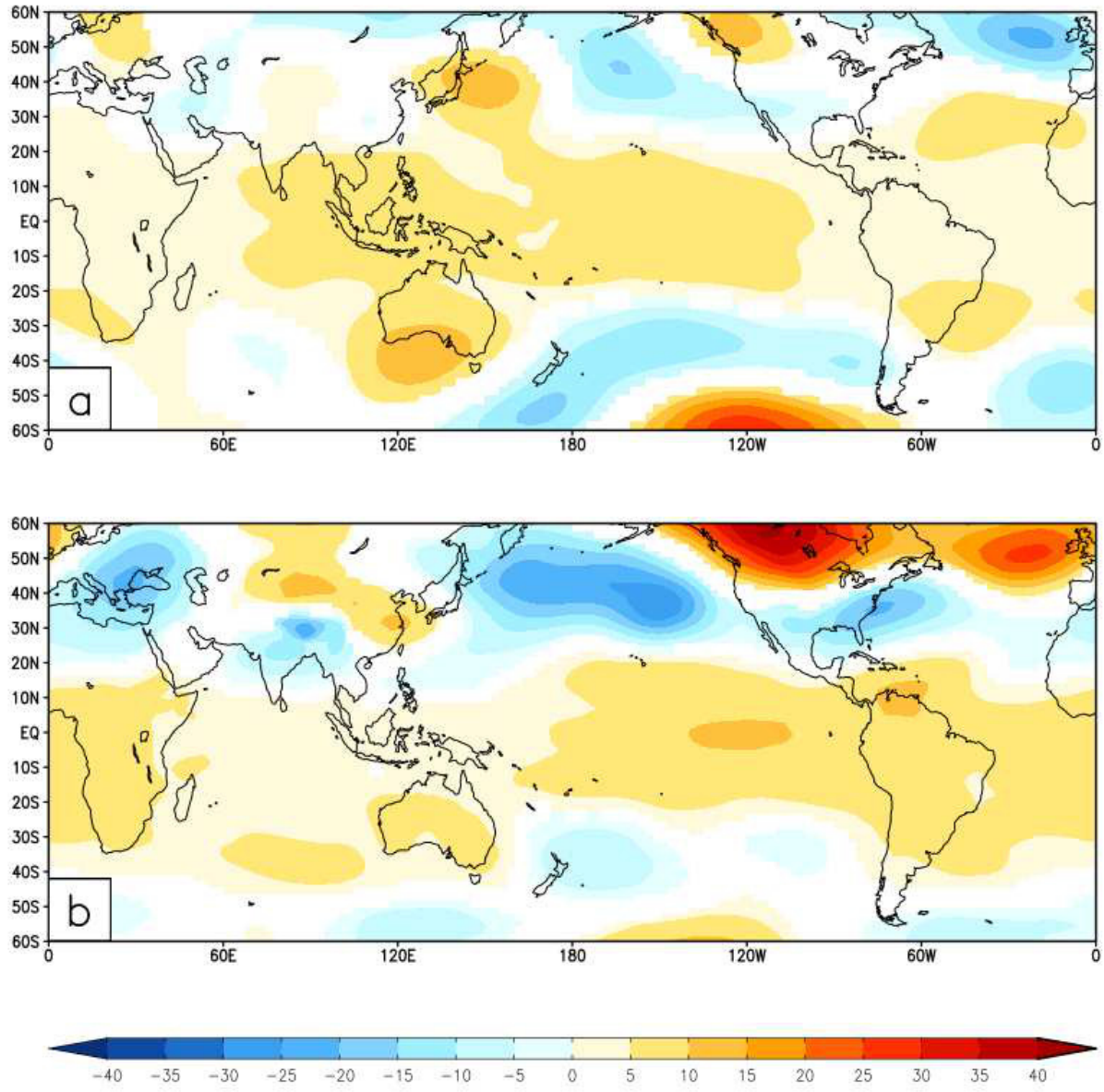

Figure B.44 Regression of Niño3.4 SST with the 500mb global geopotential heights for OND from (a) NCEP-R2 and (b) hindcast.

REFERENCES

AchutaRao, K., K. R. Sperber, 2006: ENSO simulation in coupled ocean-atmosphere models: are the currentmodels better? *Clim. Dyn.,* **27**, 1-15, doi:10.1007/s00382-006-0119-7.

Ashok, K., S. Behera, S. A. Rao, H. Weng, and T. Yamagata, 2007: El Niño Modoki and its possible teleconnection. *J. Geophys. Res*., **112**, C11 007

Awange, J. L., R. Anyah, N. Agola, E. Forootan, and P. Omondi, 2013: Potential impacts of climate and environmental change on the stored water of Lake Victoria Basin and economic implications. *Water Resour.Res.,* , n/a-n/a, doi:10.1002/2013WR014350.

Barber, R. T. and F. P. Chavez, 1983: Biological Consequences of El Niño. *Science*, **222**, 1203–1210.

Barnett, T. P., R. W. Preisendorfer, 1978: Multifield Analog Prediction of Short-Term Climate Fluctuations Using a Climate State vector. *J.Atmos.Sci.,* **35**, 1771-1787, doi:10.1175/1520-0469(1978)035<1771:MAPOST>2.0.CO;2.

Barnston, A. G., M. Chelliah, and S. B. Goldenberg, 1997: Documentation of a highly ENSO-related SST region in the equatorial Pacific. *Atmosphere-Ocean,* **35**, 367-383.

Barnston, A. G., A. Leetmaa, V. E. Kousky, R. E. Livezey, E. O'Lenic, H. Van den Dool, A. J. Wagner, and D. A. Unger, 1999: NCEP forecasts of the El Niño of 1997-98 and its US Impacts. *Bull.Am.Meteorol.Soc.,* **80**, 1829-1852.

Bastola, S., V. Misra, 2014: Evaluation of dynamically downscaled reanalysis precipitation data for hydrological application. *Hydrol.Process.,* **28**, 1989-2002, doi:10.1002/hyp.9734.

Battisti, D.S., 1988: The Dynamics and Thermodynamics of a Warming Event in a Coupled Tropical Atmosphere/Ocean Model." J. Atmos. Sci., 45, 2889-2919.

Battisti, D.S. and A.C. Hirst, 1989: Interannual Variability in the Tropical Atmosphere-Ocean System: Influences of the Basic State, Ocean Geometry and Nonlinearity. J. Atmos. Sci., 46, 1687-1712.

Behringer, D., Y. Xue, 2004: Evaluation of the global ocean data assimilation system at NCEP: The Pacific Ocean. *Proc. Proc. Eighth Symp. on Integrated Observing and Assimilation Systems for Atmosphere, Oceans, and Land Surface,* .

Bellenger, H., E. Guilyardi, J. Leloup, M. Lengaigne, and J. Vialard, 2013: ENSO representation in climate models: from CMIP3 to CMIP5. *Clim.Dyn.,* , 1-20, doi:10.1007/s00382-013-1783-z.

Bergen, R. E., R. P. Harnack, 1982: Long-Range Temperature Predictions Using a Simple Analog Approach. *Mon. Wea. Rev.,* **110**, 1083-1099.

Bjerknes, J. 1969: Atmospheric Teleconnections From the Equatorial Pacific. *Mon. Wea. Rev.*, **97**, 163-172

Brunet, G., M. Shapiro, B. Hoskins, M. Moncrieff, R. Dole, G. Kiladis, B. Kirtman, A. Lorenc, B. Mills, and R. Morss, 2009: Toward a seamless process for the prediction of weather and climate: The advancement of sub-seasonal to seasonal prediction. *Bull.Am.Meteorol.Soc.,* .

Buizza, R., T. Palmer, 1995: The singular-vector structure of the atmospheric global circulation. *J.Atmos.Sci.,* **52**, 1434-1456.

Bunge, L., A. J. Clarke, 2014: On the warm water volume and its changing relationship with ENSO. *J.Phys.Oceanogr.,* , doi:10.1175/JPO-D-13-062.1.

Cai, M., E. Kalnay, and E. Toth, 2003: Bred vectors of the zebiak–cane model and their potential application to enso predictions. *J. Climate,* **16**, 40-56, doi: http://dx.doi.org/10.1175/1520-0442(2003)016<0040:BVOTZC>2.0.CO;2.

Cane, M. A., 1983: Oceanographic Events During El Niño. *Science,* **222**, 1189-1194.

Cane, M. A., S. E. Zebiak, 1985: A Theory for El Niño and the Southern Oscillation. *Science,* **228**, 1085-1087, doi: 10.1126/science.228.4703.1085.

Cane, M. A., M. Münnich, and S. F. Zebiak, 1990: A Study of Self-Excited Oscillations of the Tropical Ocean-Atmosphere System. Part I: Linear Analysis. *J.Atmos.Sci.,* **47**, 1562-1577, doi:10.1175/1520-0469(1990)047<1562:ASOSEO>2.0.CO;2.

Capotondi, A., 2013: ENSO diversity in the NCAR CCSM4 climate model. *Journal of Geophysical Research: Oceans,* , n/a-n/a, doi:10.1002/jgrc.20335.

Chang, Ping, Link Ji, Bin Wang, Tim Li, 1995: Interactions between the Seasonal Cycle and El Niño-Southern Oscillation in an Intermediate Coupled Ocean-Atmosphere Model. J. Atmos. Sci., 52, 2353–2372.

Chou, C., L. Huang, J. Tu, L. Tseng, and Y. Hsueh, 2009: El NiÃ±o Impacts on Precipitation in the Western North Pacificâ€“East Asian Sector. *J.Climate,* **22**, 2039-2057, doi:10.1175/2008JCLI2649.1.

Chowdary, J. S., A. Parekh, C. Gnanaseelan, and P. Sreenivas, 2013: Inter-decadal modulation of ENSO teleconnections to the Indian Ocean in a coupled model: Special on decay phase of El Niño. *Global Planet.Change,* , doi:http://dx.doi.org/10.1016/j.gloplacha.2013.11.003.

Clarke, A. J., S. Van Gorder, 2003: Improving El Niño prediction using a space-time integration of Indo-Pacific winds and equatorial Pacific upper ocean heat content. *Geophys.Res.Lett.,* **30**.

Clarke, A. J., 2008: *Dynamics of El Niño & the Southern Oscillation.* Academic Press, 308 pp.

Davey, M. D., and coauthors, 2002: STOIC: a study of coupled model climatology, and variability in tropical ocean regions. *Climate Dyn.*, **18**, 403-420 doi: 10.1007/s00382-001-0188-6.

Deb, P., M. K. Dash, and P. C. Pandey, Effect of Pacific warm and cold events on the sea ice behavior in the Indian sector of the Southern Ocean. *Deep Sea Research Part I: Oceanographic Research Papers,* , doi:http://dx.doi.org/10.1016/j.dsr.2013.10.002.

Deser, C., A. S. Phillips, R. A. Tomas, Y. M. Okumura, M. A. Alexander, A. Capotondi, J. D. Scott, Y. Kwon, and M. Ohba, 2012: ENSO and Pacific Decadal Variability in the Community Climate System Model Version 4. *J.Climate,* **25**, 2622-2651, doi:10.1175/JCLI-D-11-00301.1.

Diaz, H. F., V. Markgraf, 1992: *El Niño: historical and paleoclimatic aspects of the Southern Oscillation.* Cambridge University Press.

Diaz, H. F., M. P. Hoerling, and J. K. Eischeid, 2001: ENSO Variability, Teleconnections and Climate Change. *Int. J. Climatol.*, **21**, 1845–1862.

Doblas-Reyes, F. J., R. HAGEDORN, and T. N. PALMER, 2005: The rationale behind the success of multi-model ensembles in seasonal forecasting ? II. Calibration and combination. *Tellus A,* **57**, 234-252, doi:10.1111/j.1600-0870.2005.00104.x.

Eng, J., 2013: ROC analysis: web-based calculator for ROC curves. [Available online at http://www.jrocfit.org].

Enos, P., 2013: El Niño Southern Oscillation (enso) Evolution And Its Influence On Rainfall Over Tanzania. , doi:http://erepository.uonbi.ac.ke:8080/xmlui/handle/123456789/57927.

Errico, R. M., 1997: What is an adjoint model? *Bull.Am.Meteorol.Soc.,* **78**, 2577-2591.

Fang, G., Z. Wei, B. Choi, K. Wang, Y. Fang, and W. Li, 2003: Interbasin freshwater, heat and salt transport through the boundaries of the East and South China Seas from a variable-grid global ocean circulation model. *Science in China Series D: Earth Sciences,* **46**, 149-161, doi:10.1360/03yd9014.

Fedorov, A.V. and Philander, S.G., 2001: A stability analysis of the tropical ocean-atmosphere interactions: Bridging Measurements of, and Theory for El Niño. J. Climate 14, 3086-3101.

Feng, J., W. Chen, 2013: Influence of the IOD on the relationship between El Niño Modoki and the East Asian-western North Pacific summer monsoon. *Int.J.Climatol.,* , n/a-n/a, doi:10.1002/joc.3790.

Gadgil, S., P. V. Joseph, and N. V. Joshi, 1984: Ocean-Atmosphere coupling over monsoon regions. *Nature,* **312**, 141-143.

Gadgil, S., P. N. Vinayachandran, P. A. Francis, and S. Gadgil, 2004: Extremes of the Indian summer monsoon rainfall, ENSO and equatorial Indian Ocean oscillation. *Geophys.Res.Lett.,* **31**, - L12213, doi:10.1029/2004GL019733.

Gelati, E., H. Madsen, and D. Rosbjerg, 2013: Reservoir operation using El Niño forecasts – case study of Daule Peripa and Baba, Ecuador. *Hydrological Sciences Journal,* , doi:10.1080/02626667.2013.831978.

Gent, P. R., S. G. Yeager, R. B. Neale, S. Levis, and D. A. Bailey, 2010: Improvements in a half degree atmosphere/land version of the CCSM. Clim. Dyn., 34, 819-833. Doi:10.1007/s00382-009-0614-8.

Ghil, A., et al. (2002), Advanced spectral methods for climatic time series, Rev. Geophys., 40(1), 1003, doi:10.1029/2000RG000092.

Gille, S. T., 2002: Warming of the Southern Ocean since the 1950s. *Science,* **295**, 1275-1277, doi:10.1126/science.1065863.

Gissila, T., E. Black, D. I. F. Grimes, and J. M. Slingo, 2004: Seasonal forecasting of the Ethiopian summer rains. *Int.J.Climatol.,* **24**, 1345-1358, doi:10.1002/joc.1078.

Glantz, M. H., 2000*: Currents of Change: Impacts of El Niño and La Niña on Climate and Society*. 2d ed., Cambridge University Press, 266 pp.

Goddard, L., M. Dilley, 2005: El Niño: Catastrophe or opportunity. *J.Clim.,* **18**.

Gray, W. M., 1984: Atlantic hurricane frequency. Part I: El Niño and 30mb Quasi-biennial Oscillation Influences. *Mon. Wea. Rev.*, **112**, 1649–1668.

Green, D. M., J. A. Swets, 1966: *Signal detection theory and psychophysics.* Vol. 1, Wiley New York.

Guilyiardi, E., 2006: El Niño mean state-seasonal cycle interactions in a multimodel ensemble. Clim. Dyn., 26, 329-348.

Guilyardi, E., A. Wittenberg, A. Fedorov, M. Collins, C. Wang, A. Capotondi, G. J. van Oldenborgh, and T. Stockdale, 2009: Understanding El Niño in ocean-atmosphere general circulation models. Bull. Amer. Meteor. Soc., 90, 325–339.

Guilyardi, E., H. Bellenger, M. Collins, S. Ferrett, W. Cai, A. Wittenberg, 2012: A First look at ENSO in CMIP5. Clivar exchanges 59, 17, 29-32.

Gutiérrez, F., J. A. Dracup, 2001: An analysis of the feasibility of long-range streamflow forecasting for Colombia using El Niño–Southern Oscillation indicators. *Journal of Hydrology,* **246**, 181-196, doi:http://dx.doi.org/10.1016/S0022-1694(01)00373-0.

Gutzler, D. S., J. Shukla, 1984: Analogs in the Wintertime 500 mb Height Field. *J. Atm Sci,* **41**, 177-189.

Hagedorn, R., F. J. DOBLAS-REYES, and T. Palmer, 2005: The rationale behind the success of multi-model ensembles in seasonal forecasting–I. Basic concept. *Tellus A,* **57**, 219-233.

Halpert, M. S. and C. F. Ropelewski, 1992: Surface Temperature Patterns Associated with the Southern Oscillation. *J. Climate*, **5**, 577–593.

Hanley, D. E., M. A. Bourassa, J. J. O'Brien, S. R. Smith, and E. R. Spade, 2003: A quantitative evaluation of ENSO indices. *J.Clim.,* **16**.

Hansen, J. W., A. W. Hodges, and J. W. Jones, 1998: ENSO Influences on Agriculture in the Southeastern United States. *J. Climate,* **11**, 404-411.

Hansen, J. W., J. W. Jones, C. F. Kiker, and C. F. Hodges, 1999: El Niño-Southern Oscillation Impacts on Winter Vegetable Production in Florida. *J. Climate,* **12**, 92-102.

Harley, G. L., H. Grissino-Mayer, S. P. Horn, and C. Bergh, 2013: Fire Synchrony and the Influence of Pacific Climate Variability on Wildfires in the Florida Keys, United States. *Ann.Assoc.Am.Geogr.,* , doi:10.1080/00045608.2013.843432.

Harrison, D. E., G. A. Vecchi, 1999: On the termination of El Niño. *Geophys.Res.Lett.,* **26**, 1593-1596, doi:10.1029/1999GL900316.

Hayes, S. P., L. J. Mangum, J. Picaut, A. Sumi, K. Takeuchi, TAO: A moored array for real-time measurements in the tropical Pacific Ocean, *Bull. Am. Meteorol. Soc.*, **72**, 339–347, 1991.

Hoerling, M. P., A. Kumar, 2002: Atmospheric response patterns associated with tropical forcing. *J. Climate,* **15**, 2184-2203.

Hurrell, J., G. A. Meehl, D. Bader, T. L. Delworth, B. Kirtman, and B. Wielicki, 2009: A unified modeling approach to climate system prediction. *Bull.Am.Meteorol.Soc.,* **90**, 1819-1832.

Hwang, C. L., A. S. M. Masud, 1979: *Multiple objective decision making, methods and applications: a state-of-the-art survey.* Springer-Verlag, .

Jiang, X., S. Yang, Y. Li, A. Kumar, X. Liu, Z. Zuo, and B. Jha, 2013: Seasonal-to-Interannual Prediction of the Asian Summer Monsoon in the NCEP Climate Forecast System Version 2. *J.Clim.,* **26**.

Jihoon, S., W. Choi, D. Youn, D. R. Park, and J. Y. Kim, 2013: Relationship between the stratospheric quasi-biennial oscillation and spring rainfall in the western North Pacific. *Geophys.Res.Lett.,* , - 2013GL058266, doi:10.1002/2013GL058266.

Jin, F.-F., 1997: An Equatorial Ocean Recharge Paradigm for ENSO. Part I: Conceptual Model. *J. Atmos. Sci.*, **54**, 811–829.

Joseph, R. and S. Nigam, 2006: ENSO Evolution and Teleconnections in IPCCs Twentieth-Century Climate Simulations: Realistic Representation? J. Climate, 19, 4360–4377.

Kalnay, E., M. Kanamitsu, R. Kistler, W. Collins, D. Deaven, and co authors, 1996: The NCEP/NCAR 40-year reanalysis project. Bull. Amer. Soc., 77, 437-471.

Kanamitsu, M., W. Ebisuzaki, J. Woollen, S-K Yang, J.J. Hnilo, M. Fiorino, and G. L. Potter, 2002: NCEP-DOE AMIP-II Reanalysis (R-2). Bull. Amer. Met. Soc., 83, 1631-1643.

Karnauskas, K. B., 2013: Can we distinguish canonical El Niño from Modoki? *Geophys.Res.Lett.,* **40**, 5246-5251, doi:10.1002/grl.51007.

Kiladis, G. N. and H.F. Diaz, 1989: Global Climatic Anomalies Associated with Extremes in the Southern Oscillation. J. Climate, 2, 1069-1090.

Kinter III, J., K. Miyakoda, and S. Yang, 2002: Recent change in the connection from the Asian monsoon to ENSO. *J.Clim.,* **15**.

Kim, H.-M., P. J. Webster, and J. A. Curry, 2009: Impact of Shifting Patterns of Pacific Ocean Warming on North Atlantic Tropical Cyclones. *Science*, **325**, 77–80.

Kim, S. T., J. Yu, 2012: The two types of ENSO in CMIP5 models. *Geophys.Res.Lett.,* **39**, -L11704, doi:10.1029/2012GL052006.

Kirtman, B. P. 1997. Oceanic rossby wave dynamics and the ENSO period in a coupled model. J. Climate 10, 1690–1704.

Kirtman, B. P., 2003: The COLA anomaly coupled model: Ensemble ENSO prediction.*Monthly Weather Review*, **131,** 2324-2341.

Kirtman, B. P., D. Min, 2009: Multimodel Ensemble ENSO Prediction with CCSM and CFS. *Mon. Wea. Rev.,* **137**, 2908-2930, doi: http://dx.doi.org/10.1175/2009MWR2672.1.

Kirtman, B. P., D. Min, J. M. Infanti, J. L. Kinter III, D. A. Paolino, Q. Zhang, H. van den Dool, S. Saha, M. P. Mendez, and E. Becker, 2013: The North American Multi-Model Ensemble (NMME): Phase-1 seasonal to interannual prediction, phase-2 toward developing intra-seasonal prediction. *Bull.Am.Meteorol.Soc.,* .

Korecha, D., A. Sorteberg, 2013: Validation of operational seasonal rainfall forecast in Ethiopia. *Water Resour.Res.,* , n/a-n/a, doi:10.1002/2013WR013760.

Krishnamurthy, V., and B. P. Kirtman, 2003: Variability of the Indian Ocean: Relation to monsoon and ENSO. *Quarterly Journal of the Royal Meteorological Society*, **129,**1623-1646.

Kug, J., Y. Ham, M. Kimoto, F. Jin, and I. Kang, 2010: New approach for optimal perturbation method in ensemble climate prediction with empirical singular vector. *Clim. Dyn.,* **35**, 331-340, doi:Jong-Seong Kug Yoo-Geun Ham Masahide Kimoto Fei-Fei Jin In-Sik Kang.

Kug, J., J. Choi, S. An, F. Jin, and A. T. Wittenberg, 2010: Warm Pool and Cold Tongue El Niño Events as Simulated by the GFDL 2.1 Coupled GCM. *J.Climate,* **23**, 1226-1239, doi:10.1175/2009JCLI3293.1.

Kug, J., S. An, Y. Ham, and I. Kang, 2010: Changes in El Niño and La Niña teleconnections over North Pacific-America in the global warming simulations. *Theoretical and Applied Climatology,* **100**, 275-282, doi:10.1007/s00704-009-0183-0.

Kug, J., Y. Ham, E. Lee, and I. Kang, 2011: Empirical singular vector method for ensemble El Niño–Southern Oscillation prediction with a coupled general circulation model. *J.Geophy Res,* **116**, C08029, doi:10.1029/2010JC006851.

Kumar, K. K., B. Rajagopalan, and M. A. Cane, 1999: On the Weakening Relationship Between the Indian Monsoon and ENSO. *Science,* **284**, 2156-2159, doi:10.1126/science.284.5423.2156.

Latif, M., E. Roeckner, U. Mikolajewicz, and R. Voss, 2000: Tropical stabilization of the thermohaline circulation in a greenhouse warming simulation. *J.Clim.,* **13**.

Larkin, N. K., D. E. Harrison, 2005a: Global seasonal temperature and precipitation anomalies during El Niño autumn and winter. *Geophys.Res.Lett.,* **32**, - L16705, doi:10.1029/2005GL022860.

Larkin, N. K., D. E. Harrison, 2005b: On the definition of El Niño and associated seasonal average U.S. weather anomalies. *Geophys.Res.Lett.,* **32**, - L13705, doi:10.1029/2005GL022738.

Lengaigne, M., E. Guilyardi, J. Boulanger, C. Menkes, P. Delecluse, P. Inness, J. Cole, and J. Slingo, 2004: Triggering of El NiÃ±o by westerly wind events in a coupled general circulation model. *Clim.Dyn.,* **23**, 601-620, doi:10.1007/s00382-004-0457-2.

Letnic, M., C. Dickman, 2006: Boom means bust: interactions between the El NiÃ±o/Southern Oscillation (ENSO), rainfall and the processes threatening mammal species in arid Australia. *Biodiversity & Conservation,* **15**, 3847-3880, doi:10.1007/s10531-005-0601-2.

Li, H., J. Sheffield, and E. F. Wood, 2010: Bias correction of monthly precipitation and temperature fields from Intergovernmental Panel on Climate Change AR4 models using equidistant quantile matching. *Journal of Geophysical Research: Atmospheres,* **115**, - D10101, doi:10.1029/2009JD012882.

Li, T. and G. Philander, 1996: On the annual cycle in the eastern equatorial pacific. J. Clim., 9, 2986-2998.

Li, H., J. Sheffield, and E. F. Wood, 2010: Bias correction of monthly precipitation and temperature fields from Intergovernmental Panel on Climate Change AR4 models using equidistant quantile matching, *J. Geophys. Res.*, **115**, D10101, doi:10.1029/2009JD012882.

Lin, J. L., 2007: Interdecadal variability of ENSO in 21 IPCC AR4 coupled GCMs. *Geophysical Research Letters,* **34**, L12702, doi:10.1029/2006GL028937.

López Pérez, N., 2012: The influence of El Niño Southern Oscillation episodes (ENSO) on rainfall in Monagas State, Venezuela. *Revista Científica UDO Agrícola,* **12**, 400-406.

Lorenz, E. N., 1969: Atmospheric Predictablity as Revealed by Naturally Occurring Analogues. *J. Atm Sci,* **26**, 636-646.

Magnusson, L., M. Alonso-Balmaseda, and F. Molteni, 2013: On the dependence of ENSO simulation on the coupled model mean state. *Clim.Dyn.,* **41**, 1509-1525, doi:10.1007/s00382-012-1574-y.

Maraun, D., 2013: Bias Correction, Quantile Mapping, and Downscaling: Revisiting the Inflation Issue. *J.Clim.,* **26**, 2137-2143, doi:10.1175/JCLI-D-12-00821.1.

Mason, S. J., L. Goddard, N. E. Graham, E. Yulaeva, L. Sun, and P. A. Arkin, 1999: The IRI seasonal climate prediction system and the 1997/98 El Niño event. *Bull.Am.Meteorol.Soc.,* **80**, 1853-1873.

McPhaden, M. J., 1999: Genesis and evolution of the 1997-98 El Nino. *Science,* **283**, 950-954Mechoso, C. R., and Coauthors, 1995: The seasonal cycle over the tropical Pacific in coupled ocean–atmosphere general circulation models. Mon. Wea. Rev., 123, 2825–2838.

Meehl, G. A., P. R. Gent, J. M. Arblaster, B. Otto-Bliesner, E. Brady, and A. Craig, 2001: Factors that affect the amplitude of El Niño in global coupled climate models. Clim. Dyn., 17, 515-526.

Meinen, C. S., M. J. McPhaden, 2000: Observations of warm water volume changes in the equatorial Pacific and their relationship to El Niño and La Niña. *J.Clim.,* **13**.

Michael, J. -., V. Misra, and E. Chassignet, 2013: The El Niño and Southern Oscillation in the historical centennial integrations of the new generation of climate models. *Regional Environmental Change,* **13**, 121-130, doi:10.1007/s10113-013-0452-4.

Misra, Vasubandhu, 2008: Coupled Interactions of the Monsoons. Geophys. Res. Lett. 35, L12705, doi:10.1029/2008GL033562

Misra, Vasubandhu, L. Marx, J. L. KinterIII, B. P. Kirtman, Z. Guo, D. Min, M. Fennessy, P. A. Dirmeyer, R. Kallumal and D. M. Straus, 2007: Validating and Understanding ENSO in Two Coupled Climate Models. Tellus, 58A, 292-308.

Misra, V., L. Marx, M. Fennessy, B. Kirtman, and J. L. I. Kinter, 2008: A Comparison of Climate Prediction and Simulation over the Tropical Pacific. *J. Climate,* **21**, 3601-3611, doi: http://dx.doi.org/10.1175/2008JCLI1932.1.

Misra, V., & DiNapoli, S. M. (2012). Understanding wet season variations over Florida. *Climate Dynamics*, 34. doi:10.1007/s00382-012-1382-4

Molteni, F., R. Buizza, T. N. Palmer, and T. Petroliagis, 1996: The ECMWF ensemble prediction system: Methodology and validation. *Q.J.R.Meteorol.Soc.,* **122**, 73-119.

Moore, A. M., R. Kleeman. (1997) The singular vectors of a coupled ocean-atmosphere model of Enso. I: Thermodynamics, energetics and error growth. *Quarterly Journal of the Royal Meteorological Society* **123**:540, 953-981

Mukhin, D., D. Kondrashov, E. Loskutov, A. Feigin, and M. Ghil.2012: Empirical modeling of ENSO dynamics: Predicting critical transitions in spatially distributed systems. *EGU General Assembly Conference Abstracts*. Vol. 14.

Nicholls, N., & Lavery, B., 1992: Australian rainfall trends during the twentieth century. *International journal of climatology*, *12*(2), 153-163.

Neale, R. B., J. H. Richter, and M. Jochum, 2008: The Impact of Convection on ENSO: From a Delayed Oscillator to a Series of Events. *J.Climate,* **21**, 5904-5924, doi:10.1175/2008JCLI2244.1.

Neelin, J. D., 1991: The Slow Sea Surface Temperature Mode and the Fast-Wave Limit: Analytic Theory for Tropical Interannual Oscillations and Experiments in a Hybrid Coupled Model. *J. Atm Sci,* **48**, 548-606.
Pauly, D., P. Muck, J. Mendo, and I. Tsukayama, 1989: *The Peruvian Upwelling Ecosystem: Dynamics and Interactions.* Deutsche Gesellschaft fur Technische Zusammenarbeit and International Center for Living Aquatic Resources Managment, 438 pp.

Nigam, S., 1994: On the Dynamical Basis for the Asian Summer Monsoon Rainfall-El NiÃ±o Relationship. *J.Climate,* **7**, 1750-1771, doi:10.1175/1520-0442(1994)007<1750:OTDBFT>2.0.CO;2.

NRC, 2010: *Assessment of intraseasonal to interannual climate prediction and predictability.* The National Academies Press.

Philander, S. G., 1990: *El Niño, La Niña, and the Southern Oscillation*, International Geophysical Series, Vol. 46. Academic Press, 293pp.

Oliveira, F. N., Carvalho, L., & Ambrizzi, T. (2013). A new climatology for Southern Hemisphere blockings in the winter and the combined effect of ENSO and SAM phases. *International Journal of Climatology*.

Ortíz, W. R., & Sánchez, J. S. (2011). IMPACT OF ENSO AND THE OPTIMUM USE OF YELLOWFIN TUNA (THUNUS ALBACARES) IN THE EASTERN PACIFIC OCEAN REGION. *Ingeniería de Recursos Naturales y del Ambiente*, (10).

Palmer, T. N., Branković, Č., & Richardson, D. S. (2000). A probability and decision-model analysis of PROVOST seasonal multi-model ensemble integrations. *Quarterly Journal of the Royal Meteorological Society*, *126*(567), 2013-2033.

Palmer, T. N., and Coauthors, 2004: DEVELOPMENT OF A EUROPEAN MULTIMODEL ENSEMBLE SYSTEM FOR SEASONAL-TO-INTERANNUAL PREDICTION (DEMETER). *Bull. Amer. Meteor. Soc.*, **85**, 853–872. doi: http://dx.doi.org/10.1175/BAMS-85-6-853

Palmer, T. N., F. J. Doblas-Reyes, A. Weisheimer, M. J. Rodwell, 2008: Toward Seamless Prediction: Calibration of Climate Change Projections Using Seasonal Forecasts. *Bull. Amer. Meteor. Soc.*, **89**, 459–470. doi: http://dx.doi.org/10.1175/BAMS-89-4-459

Pauly, D., P. Muck, J. Mendo, and I. Tsukayama, 1989: The Peruvian upwelling ecosystem: dynamics and interactions. Deutsche Gesellschaft fur Technische Zusammenarbeit and International Center for Living Aquatic Resources Managment, 438 pp.

Penland, C., & Magorian, T. (1993). Prediction of Nino 3 sea surface temperatures using linear inverse modeling. *Journal of Climate*, *6*(6), 1067-1076.

Philander, S. G., 1990: El Ni˜no, La Ni˜na, and the Southern Oscillation, International

Geophysical Series, Vol. 46. Academic Press, 293 pp.Power, S., F. Delage, C. Chung, G. Kociuba, and K. Keay, 2013: Robust twenty-first-century projections of El Niño and related precipitation variability. *Nature,* **advance online publication**.

Qu, T., Y. Y. Kim, M. Yaremchuk, T. Tozuka, A. Ishida, and T. Yamagata, 2004: Can Luzon Strait Transport Play a Role in Conveying the Impact of ENSO to the South China Sea?*. *J.Climate,* **17**, 3644-3657, doi:10.1175/1520-0442(2004)017<3644:CLSTPA>2.0.CO;2.

Radinovic, D., 1975: An Analog Method for Weather Forecasting Using the 500-1000mb Relative Topography. *Mon. Wea. Rev.,* **103**, 639-649.

Randall, D., Krueger, S., Bretherton, C., Curry, J., Duynkerke, P., Moncrieff, M., ... & Wielicki, B. (2003). Confronting models with data: The GEWEX cloud systems study. *Bulletin of the American Meteorological Society*, *84*(4).

Rasmusson, E. M., T. H. Carpenter, 1982: Variations in Tropical Sea Surface Temperature and Surface Wind Fields Associated with the Southern Oscillation/El Niño. *Mon. Wea. Rev.,* **110**, 354-384, doi: http://dx.doi.org/10.1175/1520-0493(1982)110<0354:VITSST>2.0.CO;2.

Roberts, W. H., Battisti, D. S., & Tudhope, A. W. (2014). ENSO in the Mid-Holocene according to CSM and HadCM3. *Journal of Climate*, *27*(3).

Rodwell, M. J. and Palmer, T. N. (2007), Using numerical weather prediction to assess climate models. Q.J.R. Meteorol. Soc., 133: 129–146. doi: 10.1002/qj.23

Ropelewski, C. F. and M. S. Halpert, 1987: Global and Regional Scale Precipitation Patterns Associated With the El Niño/Southern Oscillation. *Mon. Wea. Rev.*, **115**, 1606–1626.

Rosati, A., R. Gudgel, and K. Miyakoda, 1997: The impact of ocean initial conditions on ENSO forecasting with a coupled model. Mon. Wea. Rev., 125, 754–772

Saha, S., et al., 2006: The NCEP climate forecast system. J. Climate, 19, 3483–3517.

Saha, S., Moorthi, S., Pan, H. L., Wu, X., Wang, J., Nadiga, S., Iredell, M. (2010). The NCEP climate forecast system reanalysis. *Bulletin of the American Meteorological Society*, *91*(8).

Schneider, E. K., B. Huang, and J. Shukla, 1995: Ocean wave dynamics and El Niño. *J. Climate,* **8**, 2415-2439.

Shukla, J., D. A. Paolino, 1983: The Southern Oscillation and Long-Range Forecasting of the Summer Monsoon Rainfall over India. *Mon.Wea.Rev.,* **111**, 1830-1837, doi:10.1175/1520-0493(1983)111<1830:TSOALR>2.0.CO;2.

Shukla, J. (1995). Predictability of the tropical atmosphere, the tropical oceans and TOGA. *WORLD METEOROLOGICAL ORGANIZATION-PUBLICATIONS-WMO TD*, 725-730.

Shukla, J., R. Hagedorn, M. Miller, T. N. Palmer, B. Hoskins, J. Kinter, J. Marotzke, J. Slingo, 2009: Strategies: Revolution in Climate Prediction is Both Necessary and Possible: A Declaration at the World Modelling Summit for Climate Prediction. *Bull. Amer. Meteor. Soc.*, **90**, 175–178. doi: http://dx.doi.org/10.1175/2008BAMS2759.1

Sikka, D. R., 1980: Some aspects of the large scale fluctuations of summer monsoon rainfall over India in relation to fluctuations in the planetary and regional scale circulation parameters. *Proceedings of the Indian Academy of Sciences - Earth and Planetary Sciences,* **89**, 179-195, doi:10.1007/BF02913749.

Simpson, H. J., M. A. Cane, S. K. Lin, S. E. Zebiak, and A. L. Herczeg, 1993: Forecasting Annual Discharge of River Murray, Australia, from a Geophysical Model of ENSO. *J.Climate,* **6**, 386-390, doi:10.1175/1520-0442(1993)006<0386:FADORM>2.0.CO;2.

Smith, T., R. Reynolds, T. C. Peterson, and J. Lawrimore, 2008: Improvements to NOAA's Historical Merged Land-Ocean Surface Temperature Analysis (1880-2006). J. Climate, 21, 2283–2296.

Solís, D., & Letson, D. (2013). Assessing the value of climate information and forecasts for the agricultural sector in the Southeastern United States: multi-output stochastic frontier approach. *Regional Environmental Change*, *13*(1), 5-14.

Solomon, S., and co-authors, 2007: Global climate projections. Cambridge Univ. Press, Cambridge, UK, and New York

Song, W., J. Lan, Q. Liu, and D. Wang, 2013: Decadal variability of heat content in South China Sea inferred from observation data and an ocean data assimilation product. **10**, 1329-1342, doi:10.5194/osd-10-1329-2013.

Suarez, M. J. and P. S. Schopf, 1988: A Delayed Action Oscillator for ENSO. *J. Atmos. Sci.*, **45**, 3283–3287.

Takahashi, K., Montecinos, A., Goubanova, K., & Dewitte, B. (2011). ENSO regimes: Reinterpreting the canonical and Modoki El Niño. *Geophysical Research Letters*, *38*(10).

Tang, B. H., J. D. Neelin, 2004: ENSO Influence on Atlantic hurricanes via tropospheric warming. *Geophysical Research Letters,* **31**, L24204, doi:10.1029/2004GL021072.

Taylor, K. E., R. J. Stouffer, and G. A. Meehl, 2012: An Overview of CMIP5 and the Experiment Design. *Bull.Amer.Meteor.Soc.,* **93**, 485-498, doi:10.1175/BAMS-D-11-00094.1.

Thompson, C. J. (1998). Initial Conditions for Optimal Growth in a Coupled Ocean-Atmosphere Model of ENSO*. *Journal of the atmospheric sciences*,*55*(4), 537-557.

Toth, Z., & Kalnay, E. (1993). Ensemble forecasting at NMC: The generation of perturbations. *Bulletin of the American Meteorological Society*, *74*(12), 2317-2330.

Toth, Z., & Kalnay, E. (1997). Ensemble forecasting at NCEP and the breeding method. *Monthly Weather Review*, *125*(12).

Trenberth, K. E., G. W. Branstator, D. Karoly, A. Kumar, N.-C. Lau, and C. Ropelewski (1998),Progress during TOGA in understanding and modeling global teleconnections associated with tropical sea surface temperatures, J. Geophys. Res., 103(C7), 14291–14324, doi:10.1029/97JC01444.

Troccoli A Rienecker M, Keppenne CL and Johnson GC (2003) “Temperature data assimilation with salinity corrections: validation for the NSIPP ocean data assimilation system in the tropical Pacific ocean”, NASA Tech. Memo N. 104606, vol. 24.

Tziperman, E., M. A. Cane, and S. E. Zebiak, 1995: Irregularity and locking to the seasonal cycle in an ENSO prediction model as explained by the quasi periodicity route to chaos. J. Atmos. Sci., 52, 293-306

Tziperman, E., S. E. Zebiak, M. A. Cane, 1997: Mechanisms of Seasonal – ENSO Interaction. J. Atmos. Sci., 54, 61–71

van den Dool, H. M., 1987: A Bias Skill in Forecasts Based on Analogues and Antilogues. *J. Climate,* **26**, 1278-1281.

van den Dool, H. M., 1989: A New Look at Weather Forecasting through Analogues. *Mon. Wea. Rev.,* **117**, 2230-2247.

van den Dool, H. M., 1994: Searching analogues, how long must we wait? *Tellus,* **46**, 314-324.

van den Dool, H., J. Huang, and Y. Fan, 2003: Performance and analysis of the constructed analogue method applied to U.S. soil moisture over 1981-2001. *Journal of Geophysical Research: Atmospheres,* **108**, - 8617, doi:10.1029/2002JD003114.

Van Oldenborgh, G. J., Philip, S., & Collins, M. (2005). El Niño in a changing climate: A multi-model study. *Ocean Science Discussions*, *2*(3).

Vannière, B., E. Guilyardi, G. Madec, F. Doblas-Reyes, S. Woolnough, 2012: Using seasonal hindcasts to understand the origin of the equatorial cold tongue bias in CGCMs, and its impact on ENSO. *Climate Dynamics*, 1-19, doi: 10.1007/s00382-012-1429-6

Vikhliaev, Y., B. Kirtman, and P. Schopf, 2007: Decadal North Pacific Bred Vectors in a Coupled GCM. *J. Climate,* **20**, 5744-5764, doi: http://dx.doi.org/10.1175/2007JCLI1620.1.

Wallace, J. M., & Gutzler, D. S. (1981). Teleconnections in the geopotential height field during the Northern Hemisphere winter. *Monthly Weather Review*,*109*(4), 784-812.

Wang, B., R. Wu, and X. Fu, 2000: Pacificâ€“East Asian Teleconnection: How Does ENSO Affect East Asian Climate? *J.Climate,* **13**, 1517-1536, doi:10.1175/1520-0442(2000)013<1517:PEATHD>2.0.CO;2.

Wang, C., 2001: A Unified Oscillator Model for the El Niño–Southern Oscillation. *J. Climate,* **14**, 98-115, doi: http://dx.doi.org/10.1175/1520-0442(2001)014<0098:AUOMFT>2.0.CO;2.

Wang, L., Z. Guo, 2013: Modulation of tropical cyclogenesis over the South China Sea by ENSO Modoki during boreal summer. *Journal of Ocean University of China,* , 1-13, doi:10.1007/s11802-014-2128-1.

Webster, P. J., S. Yang, 1992: Monsoon and Enso: Selectively Interactive Systems. *Q.J.R.Meteorol.Soc.,* **118**, 877-926, doi:10.1002/qj.49711850705.

Wei, W., Y. Chang, and Z. Dai, Streamflow changes of the Changjiang (Yangtze) River in the recent 60 years: Impacts of the East Asian summer monsoon, ENSO, and human activities. *Quaternary International,* , doi:http://dx.doi.org/10.1016/j.quaint.2013.10.064.

Wittenberg, A. T., A. Rosati, N.-C. Lau, and J. J. Ploshay, 2006: GFDL's CM2 global coupled climate models. Part III: Tropical Pacific climate and ENSO. J. Climate, 19, 698-722.

Wu, R., Z. Hu, and B. P. Kirtman, 2003: Evolution of ENSO-Related Rainfall Anomalies in East Asia. *J.Climate,* **16**, 3742-3758, doi:10.1175/1520-0442(2003)016<3742:EOERAI>2.0.CO;2.

Wu, R. and B. P. Kirtman, 2005: Near-annual SST variability in the equatorial Pacific in a coupled general circulation model. J. Climate, 18, 4454-4473.

Xie P., and P. A. Arkin, 1997: Global precipitation: a 17-year monthly analysis based on gauge observations, satellite estimates, and numerical model outputs. Bull. Amer. Meteor. Soc., 78, 2539-2558.

Xue, Y., Cane, M. A., Zebiak, S. E., & Blumenthal, M. B. (1994). On the prediction of ENSO: a study with a low-order Markov model*. Tellus A, 46(4), 512-528.

Xue, Y., Cane, M. A., Zebiak, S. E., & Blumenthal, M. B. (1997). Predictability of a coupled model of ENSO using singular vector analysis. Part I: Optimal growth in seasonal background and ENSO cycles. Monthly weather review, 125(9), 2043-2056.

Xue, Y., Leetmaa, A., & Ji, M. (2000). ENSO prediction with Markov models: The impact of sea level. Journal of Climate, 13(4), 849-871.

Yonekura, E., T. M. Hall, 2013: ENSO Effect on East Asian Tropical Cyclone Landfall via Changes in Tracks and Genesis in a Statistical Model. *J.Appl.Meteor.Climatol.,* , doi:10.1175/JAMC-D-12-0240.1.

Zebiak S. and Cane M.A. (1987) A model for El Niño Southern Oscillation. Mon Wea Rev 115:2262-78.

Zelle, H., G. Appeldoorn, G. Burgers, G. J. van Oldenborgh, 2004: The Relationship between Sea Surface Temperature and Thermocline Depth in the Eastern Equatorial Pacific. J. Phy. Ocn., 34, 643-655.

Zhang, W., F. Jin, 2012: Improvements in the CMIP5 simulations of ENSO-SSTA meridional width. *Geophys.Res.Lett.,* **39**, - L23704, doi:10.1029/2012GL053588.

Zhu, J., B. Huang, M. Balmaseda, Kinter,JamesL., I.,II, P. Peng, Z. Hu, and L. Marx, 2013: Improved reliability of ENSO hindcasts with multi-ocean analyses ensemble initialization. *Clim.Dyn.,* **41**, 2785-2795, doi:10.1007/s00382-013-1965-8.

BIOGRAPHICAL SKETCH

J-P Michael received his meteorological education at Florida State University.