SEARCHING FOR A NEW ENSO SST INDEX: THE TWO-BOX METHOD

A Thesis

by

SCOTT EVAN MEYER

Submitted to the Office of Graduate and Professional Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

Chair of Committee,	John Nielsen-Gammon
Committee Members,	Andrew Dessler
	Ping Chang
Head of Department,	R. Saravanan

December 2020

Major Subject: Atmospheric Sciences

ABSTRACT

Operational Niño indices in the Eastern Pacific region use a one-box method to calculate sea surface temperature anomalies for identifying El Niño and La Niña events. A new sea surface temperature index method is presented here, which we call the Niño Difference Index. The definition calls for an additional sea surface temperature box to be placed in the Maritime Continent to use in conjunction with an Eastern Pacific sea surface temperature box. Our two-box method measures the sea surface temperature gradient in the tropical Pacific region, a hallmark feature of ENSO events since the sea surface temperature gradient weakens (strengthens) during an El Niño (La Niña) event. The definition of the Niño Difference Index has a more fundamental connection to ENSO since Niño indices only measure the sea surface temperature anomalies in a localized area of the tropical Pacific which have a strong sea surface temperature response during El Niño/La Niña events. The Niño Difference Index also relates to the shift in the locations of strong atmospheric convection in the tropical Pacific because the it is a measure of where convection migrates to/away from during an ENSO event. Niño Difference Index definitions are searched for using the local and remote response of precipitation to ENSO to use this measure of the atmospheric response and see if different regions of the globe find different index definitions that are optimal. Southern Oscillation Index data is also incorporated in our search process to search for Niño Difference Index options. Once Niño Difference Index options are narrowed down to a small subset, a series of final correlation tests is done using different ENSO metrics. This will determine the strength of the final Niño Difference Index options relationship with ENSO compared to the operational Niño indices.

ACKNOWLEDGEMENTS

I would like to thank my committee chair, John Nielsen-Gammon, for his guidance and support throughout the course of this research and in graduate school in general. You have been so kind to me during my years at Texas A&M and your support helped me to see my masters degree through to completion.

CONTRIBUTORS AND FUNDING SOURCES

Contributors

This work was supervised by a thesis committee consisting of Andrew Dessler and Ping Chang and R. Saravanan of the Department of Atmospheric Sciences. All work conducted for the thesis was completed by the student independently with guidance from the committee chair, John Nielsen-Gammon.

Funding Sources

This research was funded by the Climate Program Office, Office of Oceanic and Atmospheric Research, National Oceanic and Atmospheric Administration, under award number NA17OAR4310157.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1. INTRODUCTION

1.1. ENSO background

The term El Niño Southern Oscillation (ENSO) identifies a coupled ocean-atmosphere phenomenon that occurs over the equatorial Pacific. El Niño relates to the oceanic component of warming sea surface temperatures while the Southern Oscillation relates to the associated Walker Circulation change. ENSO is also the leading mode of inter-annual global climate variability (Wu et al. 2019). The purpose of this research is to better quantify the magnitude of ENSO and its planetary consequences in the context of changing climate conditions.

A common way to characterize the development of an ENSO event is through the Bjerknes feedback process (Cai et al. 2015). The Walker Circulation's mean flow at the surface is from east to west (easterly), which locally amplifies the trade winds and directly correlates with the direction of the mean SST gradient in the equatorial Pacific. Under normal conditions, winds blow from higher pressure in the eastern Pacific (EP) where cooler SSTs reside to lower pressure where warmer SSTs are in the western Pacific. However, shifts in the normal convective zones in the equatorial tropical Pacific and SST gradient changes during an ENSO event lead to a change in the strength or even the direction of the Walker Circulation. During an El Niño, warmer sea surface temperatures (SST) extend eastward. As a result, the strength of the Walker Circulation decreases in an El Niño event due to the weakening of the SST gradient, allowing the main area of convection in the equatorial Pacific to extend eastward into the central Pacific (CP) region. During a La Niña, anomalously cool SSTs develop in the western tropical Pacific. This causes the main area of convection to be constrained westward, deeper in the

Maritime Continent (MC) region (Neelin 2011, pp. 103-107). As a result, there is enhanced easterly trade wind intensity due to a stronger than normal SST gradient. The shifts in the atmospheric circulation in the equatorial Pacific during an ENSO event lead to changes in circulation and weather patterns across the globe.

1.2. Index Definitions

Multiple indices already exist to characterize and quantify ENSO because multiple atmospheric and oceanic fields change during an occurrence of an ENSO event. The different indices in public or private use by meteorologists can focus on changes in different atmospheric and oceanic fields. The Southern Oscillation Index (SOI), the Niño 3 index, the Oceanic Niño index (ONI), the MEI, the OLR El Niño and OLR La Niña are examples of indices that have been created to identify different phases of ENSO. Our research is about creating a new index; to illustrate the need, we will now review other indices in operational use or use in research.

The first definition to reach operational use was the SOI. Walker and Bliss (1932) created the original SOI using seasonal pressure, temperature, and rainfall measurements from carefully selected individual stations across the globe to develop a single number. Each season had different measurements from different locations to best represent the variability due to ENSO. However, the most common present-day form of the SOI calculates the difference between atmospheric pressure measurements from two locations in the tropical Pacific, Tahiti and Darwin (Allan et al. 1991).

The National Oceanic and Atmospheric Administration (NOAA) uses sea surface temperature anomalies (SSTa) over the Niño 3.4 region (5°N to 5°S, 120°W to 170°W) to classify El Niño or La Niña conditions via the ONI or the Niño 3.4 index. Both use the Niño 3.4

region for SST averaging and a three-month running mean of SSTa's for each individual index value (Lindsey 2013; Kousky and Higgins 2007). For example, a Niño 3.4 index and an ONI value for January of 1980 would be the average SSTa for December 1979 through February 1980 (DJF). The Niño 3.4 index uses 1981-2010 as climatology to calculate the SSTa's for index values while the ONI uses thirty year centered averages for each five-year period (Huang et al. 2016). For example, the ONI calculates SSTa's for the years 1966-1970 using the base SST average from 1951-1980. According to NOAA, a El Niño or La Niña event has occurred when SSTa's exceed the threshold of +/- 0.5°C for five consecutive monthly index values (Kousky and Higgins 2007). The Japan Meteorological Agency (JMA) calculates the Niño 3 index using a five month running mean of SST's in the Niño 3 region (5°N to 5°S, 90°W to 150°W) to prevent intra-seasonal variations in SST's from yielding false signals in identifying ENSO events (Trenberth 1997). For example, an index value for March of 2019 uses the SST mean from January 2019 to May 2019. Then they use a sliding thirty-year period as climatology to calculate SSTa's for final index values. The JMA declares an ENSO event when SSTa's exceed the threshold of +/- 0.5°C for six consecutive monthly index values or more (Trenberth 1997).

The process-oriented El Niño index (PEI) separates the four Niño indices (Niño 1+2, Niño 3, Niño 3.4, Niño 4) and specifies classification of events based on the evolution of ENSO events (Song et al. 2016). EP El Niño events are declared if SSTa's in the Niño 1+2 or 3 regions are equal to or greater than 0.5 standard deviation of monthly SSTa's for four consecutive months. CP El Niño's are declared if Niño 3.4 or Niño 4 anomalies are equal to or greater than a 0.5 standard deviation of monthly SSTa's for four consecutive months (Song et al. 2016).

The Multivariate ENSO Index (MEI) includes the atmospheric and oceanic responses associated with ENSO based on a combined principal component (PC) analysis of six observed

fields: sea level pressure (SLP), zonal and meridional wind, SST, near surface air temperature, and total cloudiness (Wolter and Timlin 1993). The MEI approach calculates the first PC of the combined observed fields to identify and quantify the variability due to ENSO in the leading mode. A new version of the MEI is based on two fields rather than six, which are the SST and SLP fields (Wolter and Timlin 2011). This allows for their analysis to go back farther in time since they have more temporal data for those two fields.

The outgoing long-wave radiation (OLR) method uses two separate indices for identifying El Niño's and La Niña's (Chiodi and Harrison 2010;2013;2015). The OLR El Niño index uses OLR anomalies over the region 5°S to 5°N and 160°W to 110°W (Chiodi and Harrison 2013). On the other hand, the OLR La Niña index uses the same latitude bounds, but shifts the box westward to 150°E to 180°E (Chiodi and Harrison 2015). These boxes are set to the regions where atmospheric convection is known to shift in the equatorial tropical Pacific during an El Niño and La Niña, respectively. OLR anomalies in the regions where anomalous convective activity occurs due to ENSO reflect atmospheric heating anomalies, which drive changes in extratropical circulation patterns during ENSO events (Chiodi and Harrison 2010;2015). Additionally, these indices yield a distribution that exhibit a more event-like pattern, where extreme events deviate from the Gaussian distribution of index values for more moderate events or normal conditions. Indices such as the ONI yield a distribution that closely resembles a Gaussian-type distribution (Chiodi and Harrison 2010). We used Monthly CP OLR data in our research, which is the average OLR in a box bounded by 170°E-140°W,5°S-5°N, in our research because we could not find the calculated index values for the original Chiodi and Harrison definitions.

1.3. Strengths and weaknesses of each ENSO index

Climate change may affect some of the current ENSO indices because of warming SST's in the tropical Pacific. However, there are uncertainties regarding the ENSO signal and teleconnections changing in a warming climate (Drouard 2019). Changes in ENSO SST patterns also remain uncertain in a warming climate (Collins 2000; Meehl et al. 2006, Cai et al. 2015). Warming may be non-uniform across the tropical Pacific basin, which would lead to changes in the zonal mean SST gradient. This would create issues with current Niño indices since a different zonal mean SST gradient would cause a need for different index thresholds than currently in use to identify ENSO events. Additionally, there have been more CP El Niño's since the early 1990's (Song et al. 2016). If this trend continues, it would hurt the viability of the Niño 3 index, which is centered in the cold tongue region of the EP. The Niño 3.4 index catches some of the CP El Niño's variability.

The Niño 3 index and the ONI are currently in operational use to monitor ENSO events by the JMA and NOAA, respectively. Both use shifting thirty-year base average periods to calculate SSTa's for monthly index values. This means that the most recent ten to fifteen years of index values must use lagged averages (Lindsey 2013; Kousky and Higgins 2007), which will create inhomogeneity in the indices most recent values as the averages change. For example, the ONI uses the 1986-2015 SST average as climatology for the period of 2011-2015.

The MEI may be affected because a warmer climate could yield different spatial Empirical Orthogonal Function (EOF) patterns for ENSO event responses than in the past (Wolter and Timlin 2011). A change in the spatial pattern would lead to a weaker correspondence to the ENSO signal, the leading mode of variability across the globe.

A need to change the pressure stations used for SOI may occur under climate change because these stations are in the western (Darwin) and central (Tahiti) tropical Pacific regions. Their pressure difference reflects pressure changes in different regions of the tropical Pacific where high variability due to ENSO currently occurs. However, it measures the atmospheric response, not the initial SST gradient change in the tropical Pacific associated with ENSO. If the variability pattern of pressure shifts eastward, for example, then Tahiti would no longer be in the region where high variability occurs and would no longer be a good location to use for the SOI.

The OLR method may be slightly affected by climate change due to a shift in where the main convective area lies during an El Niño or a La Niña. Thus, in a changing climate, the box bounds that monitor El Niño or La Niña may need adjustment. This is the case with SOI as well because it is an atmospheric response to the changing SST gradient during ENSO events in the tropical Pacific. Additionally, the period of record for the OLR is short since it spans the length of the satellite-era. This and the fact that is based on OLR would make it hard to relate it to model output.

Different model-based studies use different definitions of ENSO. Ham et al. 2015a uses the Niño 3 and Niño 4 indices to identify EP and Central Pacific (CP) El Niño's. Their definition of EP (CP) El Niño events call for the Niño 3 index (Niño 4 index) in DJF exceeds one standard deviation of the mean index value. Ham et al. 2015b uses the Niño 3 index to investigate the behavior of ENSO amplitude in a warming climate, using RCP 4.5 runs, versus historical climate runs. Oh et al. 2014 also uses the Niño 3 index, and examines the role of ENSO on precipitation anomalies over the southeastern United States in DJF using the ENSO event definition from JMA. Gonzalez et al. 2015 uses the Niño 3.4 index to examine the skill of the index to detect ENSO events across different lead times for different climate models. Rashid et al. 2016 uses the

leading EOF and PC of SSTa's over the tropical Pacific basin to define ENSO. The lack of consensus between model-based studies on how to define ENSO events makes it difficult to compare results with observation-based studies.

Other papers create ENSO indices that solve some of the issues of other indices but are not in operational use. One such index is the ENSO longitude index (ELI) (Williams and Patricola 2018). The ELI uses the longitude of the primary warm pool of SST's in the tropical Pacific as the proxy to define El Niño and La Niña events. Williams and Patricola (2018) compares the ONI to the ELI and finds that the ELI is more fundamentally intertwined with ENSO. It is based on the direct connection between convective heating and where the maximum of SST's on the longitude axis exist during an ENSO event (Williams and Patricola 2018). ELI performance would not be affected in a warmer climate since it measures the longitudinal axis where the largest SST's reside, which may change for common ENSO events in a future climate.

1.4. Motivation

An ideal ENSO index would be able to measure the non-linear aspects of ENSO. It would also be able to summarize these non-linear aspects in a way that enables the scientific community and community at large to convey information about the current state of ENSO. Also, an index definition that is unambiguous is desirable to monitor ENSO since it would likely not need to be adapted due to a changing climate. This summary directly relates to the projected seasonal forecast for extratropical regions susceptible to ENSO.

The Niño Difference Index (NDI) employs a two-box method, where we take the difference of monthly average SSTs between where deep convection normally occurs and where it migrates to (or away from) during ENSO events. A hallmark feature of El Niño events is the

reduction or reversal of the zonal SST gradient in the tropical Pacific. Thus, this method is meant to encapsulate this feature into its index values. Also, the NDI is intended to be simple to use and measure what leads to the atmospheric response. This would make the NDI a strong candidate to be utilized as a universal tool for ENSO monitoring and forecasting.

The OLR method will be our primary guide for the creation of the NDI because it is a measure of where high convective activity shifts to in the tropical Pacific during ENSO events, which directly relates to our method to create an ENSO index definition. Each of the definitions listed above has merit in representing the coupled ocean-atmosphere response due to ENSO. However, an issue that arises from this is the lack of consensus among the scientific community on which method to use to monitor ENSO. The global atmospheric response to ENSO, which is driven by a change in the zonal SST gradient in the equatorial tropical Pacific, is nonlinear. The desirability of an ENSO index that directly relates to the change of the zonal SST gradient in the tropical Pacific, which is a fundamental feature of ENSO, is the motivating purpose for the creation of the NDI.

In section 2, I will discuss the data sets and methods used for NDI testing and creation. Sections 3 through 8 will cover and explain the results from NDI testing. Section 9 summarizes our findings and identifies avenues for future work.

2. DATA AND METHODS

2.1. Data

The data for this project includes gridded sea surface temperature data, various ENSO indices, and global precipitation data. Our primary SST dataset is the ERSSTv5 dataset, which has a 2° by 2° grid resolution, from the National Centers for Environmental Information (Source: https://www.ncdc.noaa.gov/data-access/marineocean-data/extended-reconstructed-sea-surface-temperature-ersst-v5 Huang et al. 2018) For sensitivity testing, we use the Centennial in situ Observation-Based Estimates (COBE) SST dataset from the Japanese Meteorological Agency (https://ds.data.jma.go.jp/tcc/tcc/products/elnino/cobesst/cobe-sst.html Ishii et al. 2005) and the Hadley Centre Sea Ice and Sea Surface Temperature (HadISST1) dataset from the United Kingdom Met Office (https://www.metoffice.gov.uk/hadobs/hadisst/data/download.html Rayner et al. 2003), which both have a 1° by 1° grid resolution. We choose to use monthly data from all three datasets with a temporal coverage of January 1900 to February 2019 to have 119 years of data for each season.

We use three different reconstructed SST datasets to verify that our results for the optimal NDI are robust to choice of dataset. All three reconstructed SST datasets use the International Comprehensive Ocean-Atmosphere dataset (ICOADS) for in-situ SST observations (https://climatedataguide.ucar.edu/climate-data/sst-data-hadisst-v11) (Huang et al. 2017). However, each dataset uses a different version of the ICOADS due to the different release dates of the three SST datasets. Additionally, ERSSTv5 uses ARGO float (Destin 2014) ocean temperature data within five meters of the sea surface for additional in-situ SST data to track temporal SST changes across ocean basins (Huang et al 2018). Near-surface ARGO float

observations are similar to ship observations, thus it's an additional observational data source ERSSTv5 utilizes to make the final SST reanalysis product. However, ERSSTv5 only uses ARGO float observations after the turn of the twentieth century because they do not become statistically useful until then (Huang et al. 2017). HadISST1 and COBE do not include ARGO float observational data. Differences between each SST reconstruction dataset also stem from the time period they choose to use as the mean climatology, methods they use to adjust raw biases in in-situ SST observations, and the methods for reconstruction. Climatology base period is not an important difference since all three use overlapping time periods; however, it is beneficial to know what the base period is for final SST value calculations.

Bias-adjustments to in-situ SST data differ amongst the three datasets used. The ERSSTv5 uses nighttime marine air temperature data that extends back to the late 1800's to make bias adjustments to the sea-surface temperature observations (Huang et al. 2016). The HadISST1 only makes bias adjustments for in-situ SST data prior to World War II (Rayner et al. 2003), using linear equations to take into account the annual amount of slow/fast moving ships and wooden/canvas bucket measurements leading up to 1941. These equations use the period 1951-1980 to compare the adjustment calculations (Folland and Parker 1995). COBE follows the same bias-adjustment procedure as HadISST1 for pre-World War II in-situ SST data except for engine room intake measurements (Hirahara 2014).

The HadISST1 converts in-situ observational SST values to gridded SSTa's using the Reduced Space Optimal Interpolation (RSOI) technique (Rayner et al. 2003). RSOI uses EOF's that are as large or larger than the size of ocean basins to reconstruct SSTa fields, especially in regions/periods of time where data is sparse (Jones et al. 2001). With this method, 1961-1990 is used as the climatology base period. Since such large-scale procedures lead to dampened local

variance, the non-interpolated gridded in-situ SST data was blended with the reconstruction data to restore local variance. Thus, HadISST1 fields are a blend of observations with reconstruction fields (Rayner et al. 2003).

COBE uses the base period of 1961-2005 as climatology. They use bias-corrected in situ and satellite SST observations to create a SST analysis of their defined SST field mean state. They call their reconstructed SST field analysis the "multi-time-scale analysis" (Hirahara 2014). First, they calculate the secular trend, which is the leading EOF of year-to-year mean SSTa's for 5° longitude by 5° latitude grid boxes of satellite and in situ observations that spans the length of the data set. Then, they compute EOFs to explain ~90% of the SST variance for time periods ranging from interannual to interdecadal using the SST field mean state from the aforementioned climatology period. They use this information to reconstruct the SST fields in areas with a lack of observations in space and time. They also include daily SST information in the reanalysis product to incorporate localized SST variance where data is temporally and spatially high. The SST fields are then converted to monthly, gridded data with a grid resolution of 1° longitude by 1° latitude (Hirahara 2014).

ERSSTv5 utilizes the SST climatology base period of 1971-2000 to convert in-situ SST observations that pass quality control tests into monthly SSTa's (Huang et al. 2017). The calculation of monthly averages for each 2° by 2° grid box is done after compiling SSTa's from observations into their respective grid boxes. They execute this by first filtering the in-situ SST observations to create low frequency gridded SSTa's. Then, the high frequency component of SSTa's, defined by the difference of the raw and low frequency SSTa's, are fit to 130 Empirical Orthogonal Teleconnections (EOT). EOT's are regional EOFs that have a maximum spatial scale of 5000 km in longitude and 3000 km in latitude (45° longitude by 27° latitude size box). The

EOT spatial patterns are derived from the Optimum Interpolation SST (OISST) version 2 monthly data using the time period of 1982 to 2011. The final reconstruction of SST's merges low frequency and high frequency SSTA's and then adds this field to the monthly climatology (Huang et al. 2016).

Since ENSO drives seasonal precipitation anomalies across the globe, and precipitation anomalies are the primary driver of societal impacts due to ENSO, we measure the global ENSO response using patterns of precipitation variability. We use the Global Precipitation Climatology Project (GPCP) version 2.3, which sources satellite and rain gauge data to integrate into a monthly gridded precipitation field with 2.5° by 2.5° grid resolution (Adler et al. 2018) (Source: https://www.esrl.noaa.gov/psd/data/gridded/data.gpcp.html). GPCP data becomes available in 1979, the beginning of the satellite era, since the methods used to create the global precipitation product rely on satellite observations. We use data that spans from January of 1979 to February of 2019 so we have 40 years of precipitation data for all seasons to use for our analysis.

Two different methods are used to create the final precipitation product based on the availability of microwave estimates. For all data after 1987, they first use multiple microwave and infrared satellites to produce large-scale (5° by 5° grid) average precipitation values. Then, they use large-scale averages from land gauges where they are available to adjust the merged large-scale satellite analysis to better match the regional-scale (2.5° by 2.5° grid) gauge analysis. The large-scale gauge adjusted satellite analysis is merged with the post-1987 gauge analysis to create the final analysis for this time period (Adler et al. 2003). However, no microwave-estimates are available to create monthly gridded precipitation data prior to 1987. Instead, they use the OLR precipitation index (OPI) to create the beginning portion of the GPCP dataset. The fundamental principle of this technique is that observations show OLR anomalies have a defined

negative correlation with precipitation anomalies worldwide. First, they calibrate a global linear function for the OPI using OLR and local precipitation anomalies from the merged GPCP analysis data spanning from 1988-1997. Then, they use OLR anomalies from 1979-1987 to calculate monthly OPI-estimated precipitation anomalies for that time period. The precipitation anomaly values are then added to the mean precipitation field to create monthly precipitation estimates. This field is merged with the gauge analysis for 1979-1987 to create the final analysis for this time period. The monthly gridded precipitation fields for each time period are combined to create the final GPCP product that spans from the beginning of 1979 to 2019 (Adler et al. 2003).

Four different indices incorporated in our analysis each provide quantified measures of the responses to ENSO using different oceanic and atmospheric parameters. The primary SST index used in our analysis is the monthly Niño 3.4 index data from the Earth System Research Laboratory (Rayner et al. 2003 https://www.esrl.noaa.gov/psd/gcos_wgsp/Timeseries/Nino34/). The Niño 3.4 index uses HadISST1 data to calculate monthly index (SSTa) values based on the 1981-2010 climatological SST mean for three-month running means of SST's in the Niño 3.4 region. Niño 3.4 index data is available for the full temporal length of the three SST datasets used for analysis (January 1900-February 2019) while ONI data is only available dating back to 1950. Since the ONI is the current operational ENSO index in use by NOAA, we correlate ONI data, calculated using ERSSTv5 SST data, from the Climate Prediction Center (https://origin.cpc.ncep.noaa.gov/products/analysis_monitoring/ensostuff/ONI_v5.php) with Niño 3.4 index data for each standard season to verify that these Niño 3.4 region indices are interchangeable. The strong correlations for all seasons indicate that the indices are nearly identical, thus confirming we can use the Niño 3.4 index rather than the ONI to measure the SST

response in the Niño 3.4 region due to ENSO. We include Niño 3 index data from the Tokyo Climate Center (http://ds.data.jma.go.jp/tcc/tcc/products/elnino/index/), calculated using the coupled ocean-atmosphere JMA model (MRI-CPS2), to compare our NDI choices with another operational SST index in use to monitor ENSO. The JMA calculates index (SSTa) values using a sliding thirty–year period as climatology for five month running means. Southern Oscillation Index (SOI) data from the University of East Anglia's Climate Research Unit (Ropelewski and Jones 1987 https://crudata.uea.ac.uk/cru/data/soi/) is also used, with the same temporal length as the Niño 3.4 index. The Climate Research Unit calculates the SOI using a two-step normalization procedure with raw monthly SLP data. 1951-1980 is used as the climatological mean base period to generate the index data (Ropelewski and Jones 1987). We also use the CP OLR index, which has a spatial coverage of 170°E to 140°W and 5°S to 5°N (https://www.cpc.ncep.noaa.gov/data/indices/cpolr.mth.81-10.ascii from https://www.cpc.ncep.noaa.gov/data/indices/). Although the CP OLR index does not match the El Niño and La Niña OLR indices used by Chiodi and Harrison 2013 and 2015, it is a single index rather than two separate indices based on ENSO phase. We use to evaluate the relationship between the ENSO induced atmospheric OLR response in the tropical Pacific and our two-box SST method.

2.2. Methods

Our first objective is to identify the combination of MC and EP regions whose mean monthly sea surface temperatures, when differenced, are highly correlated to the global ENSO response in all seasons. For simplicity and consistency with the traditional Niño regions (Barnston et al.1997), we constrain each region to be rectangular, bounded by meridians and

parallels. Each such box, as we shall henceforth refer to them, is uniquely defined by four parameters, which we choose to be the box width, box height, and the latitude and longitude position of the box center. This yields a total of eight degrees of freedom to investigate when choosing the optimal MC and EP boxes for the NDI. After a global exploration of the full eight-parameter space, we perform a series of tests that isolate different degrees of freedom to optimize our MC and EP box sizes and locations for our final NDI options. The primary measure of optimization is the correlations between the resulting NDI and the time series measuring the atmospheric responses to ENSO.

We use Principal Component (PC) analysis to calculate the six leading modes of precipitation variability for all standard seasons to confirm that the leading mode is the one corresponding to the planetary ENSO response. To do this, we first extract the monthly values for each standard season using the gridded global precipitation data. We detrend the seasonal precipitation data by removing the mean trend over time at each grid point and then calculate the annual mean of each season at each grid point. We perform this calculation over four differing spatial domains because our interest is to see whether precipitation responses in different parts of the globe require different definitions of the NDI. Note that these definitions focus on separating the regions where the local and remote responses of precipitation patterns occur due to ENSO. The regions in the tropical Pacific where local shifts in convection patterns occur due to the zonal SST gradient changes associated with ENSO are the EP and MC regions. Here we define EP region latitude and meridian bounds to be 10°S to 10°N, and 160°E to 80°W while the MC region includes the area between 10°S and 15°N, and 110°E to 160°E. In figures 1 through 4, normalized amplitude refers to the method we use to standardize each PC and the corresponding spatial EOF data, which is the standard deviation of the original PC. We divide the raw PC time

series by the standard deviation of that PC time series to generate the time series used in analysis. Then, we multiply the respective spatial EOF field by the standard deviation of the raw PC time series. The first domain is Extratropical Precipitation (EXP), which includes all global precipitation data except for the defined EP and MC regions (Figure 1). The second is Northern Hemisphere Precipitaton (NHP), which includes precipitation data north of the equator and outside the EP and MC regions (Figure 2). The third is Tropical Precipitation (TROP), which includes precipitation data only in the defined EP and MC regions (Figure 3). Global Precipitation (GP) is included in our analysis for comparison purposes and final correlation tests (Figure 4).

Precipitation EOFs are calculated to show the spatial patterns and retrieve the resulting PC time series for the leading modes of variability across the globe. Note that the standardized PC time series is what we use for analysis discussed in sections 3 through 7. The sign of the PC is chosen so that the slope of the time series against the Niño 3.4 index is always positive since sign is arbitrary in a PC time series. This is applied to SOI data as well for consistency purposes. We only display the two leading spatial EOF modes for all four cases. We choose to display DJF fields since the strongest atmospheric impacts due to ENSO appear in this season (Neelin, pp. 132). In figures 1, 2 and 4, the high spatial variability of precipitation over continental regions in the leading mode (mode 1) match the DJF precipitation variability due to El Niño from past observations (Neelin, pp. 132). The remote and global precipitation EOFs (Figures 1, 2, and 4) have similar spatial variability patterns in the leading mode, making them not as unique as anticipated. However, the TROP (Figure 3) domain covers only the EP and MC regions, thus making it unique since it isolates the local response of convection due to SST gradient changes associated with ENSO.

Tables 2 and 3 display a big gap in seasonal correlation values between the first and second PC of GP and the Niño 3.4 index/SOI. In NDJ and DJF, the second PC is nearly uncorrelated with the SOI and Niño 3.4 while the first PC is strongly correlated with them. Although ENSO is the leading mode of climate variability, there is evidence that the second spatial EOF mode captures climate variability of a different strain of ENSO, the central Pacific El Niño (Ashok et al. 2007). Ashok 2007 calculates and plots the leading spatial EOF modes of SSTa's in the Pacific basin from 1979-2004. The first spatial EOF mode clearly identifies the traditional flavor of ENSO, the eastern Pacific El Niño (Figure 5). However, the second spatial EOF mode displays anomalously warm SST's in the central Pacific, shifted westward to the dateline compared to EP El Niño's (Figure 5). This mode captures the spatial SSTa pattern observed in years when CP El Niño's occur (Ashok et al. 2007) (Figure 6). For our analysis, we only use the first PC of precipitation for all four cases to measure the local and remote precipitation responses to ENSO since the strong correlations with the SOI and the Niño 3.4 index indicate that it captures the dominant ENSO signal.

The SOI has consistently lower values for correlations with the first PC of GP than the Niño 3.4 index. This does not mean the SOI is not useful in operation to monitor ENSO. Correlations are still high, above 0.8 for most seasons except for MAM and MJJ (Table 2). However, the Niño 3.4 index has a stronger relationship to the response of global precipitation patterns due to ENSO than the SOI. This is expected since changes in the location of the SST warm pool in the tropical Pacific region drive the changes the atmospheric pressure field and regions of strong convection during ENSO.

3. OVERVIEW OF RESULTS

We employ a multi-step selection process to narrow down the multitude of NDI options available. Correlation calculations of the NDI and EP/MC box options for the NDI with the PC time series of seasonal precipitation and other ENSO indices use SST data that covers the standard and one-month lead seasons. In section 4, we review key papers that lead to the selection of the Niño 3.4 region to monitor ENSO. In section 5, the search of all possible combinations of EP and MC boxes varies the latitude and longitude of the box centroids as well as the size of the box: eight degrees of freedom in all. Here the exploration of ideal NDI box options with comparison to the SOI and EXP PC 1 will be analyzed. Also, the limitation of the degrees of freedom will allow us to isolate the search to the MC region first, then the EP region once MC box candidates are chosen in section 6. Then, in section 7, correlations of the NDI options, the Niño 3 index, and the Niño 3.4 index with atmospheric ENSO metrics will determine whether the two-box SST method improves upon the one-box SST methods in operational use to monitor ENSO and which NDI options will be recommended for future work. Finally in section 8, the ability of the recommended NDI options from section 7 and the Niño 3.4 index to identify moderate vs. strong ENSO events will be evaluated using defined OLR vs. non-OLR events from Chiodi and Harrison 2015.

4. THE REBIRTH OF THE NIÑO 3.4 SST REGION FOR ENSO MONITORING

Trenberth 1997 and Barnston et al. 1997 were the pioneers of developing the Niño 3.4 index and the ONI, the latter being the operational index in use by NOAA. The rest of this section reviews these papers and reevaluates their results in the context of global precipitation, with a focus on Barnston et al. 1997. Since our limited degrees of freedom method to search for NDI options in section 3.4 incorporates the Niño 3.4 index, it is important to discuss how the Niño 3.4 region was chosen as a viable location in the EP region to monitor ENSO.

Trenberth 1997 compared the Niño 3 region to the Niño 3.4 region. Trenberth also looked at histograms of SSTa's from 1950 to March of 1997 for Niño 3 and Niño 3.4 (Figure 7). The Niño 3 histogram is strongly skewed in the negative direction. On the other hand, the Niño 3.4 histogram shows a more bimodal pattern, which is more representative of the ENSO phenomenon (Trenberth 1997). This indicates that the Niño 3.4 region is more representative of where warmer SST's in the tropical Pacific migrate to (away from) during El Niño (La Niña) events, which suggests that the Niño 3.4 region has a stronger connection to ENSO than the Niño 3 region. Although the paper shows some advantages of using the Niño 3.4 region rather than the Niño 3 region to monitor ENSO, Trenberth acknowledged that it may be suitable for different countries to use different ENSO indices based off of remote impacts of relevance. The objective of our research is to determine a NDI candidate to recommend as a universal tool for ENSO event identification.

Barnston et al. 1997 searched for the SST region that has the strongest relationship to ENSO by evaluating the relationship between the SOI and tropical Pacific SSTs. Using data from 1950-1979, Barnston found that the highest correlations between the SOI and local SSTs

reside between the dateline and 120°W, centered near the equator (Figure 8). Barnston et al. 1997 also uses Canonical Correlation Analysis (CCA) to predict JFM tropical Pacific SSTs with subsurface sea temperature, global sea level pressure and tropical SSTs being the predictors for zero, one, and two season leads (Figure 9). They find that the highest correlation values (>0.8) reside south of the equator (0° to 10°S) between 170°W and 120°W. The strongest correlations in the tropical Pacific regions shown in figures 6 and 7 closely align with where the Niño 3.4 region is located.

Barnston et al. 1997 also calculated spatial correlations between local SST and SOI for the season DJF using data from 1955-1994. Figure 10 shows that for DJF in the tropical Pacific region, the correlation maximum is near the longitudinal center of the Niño 3.4 region, which is 145°W; however, the latitude center is located around 5°S (Barnston et al. 1997). Similar to what is shown in figure 8 and 9 using CCA, the strongest area of correlations, outlined by the -80 contour, encompasses a large part of the Niño 3.4 region in figure 10. To compare our data to figure 10, we present spatial grid point correlation maps that use DJF ERSSTv5 SST and SOI data for two different time periods.

The correlations of EP region grid point SST's and the SOI for 1900-2019 display a similar spatial pattern to figure 10 from Barnston et al. 1997 (Figure 11). The region of strongest correlations (<-0.8) both extend from ~170°W to ~120°W, the longitude bounds for the Niño 3.4 region. The most important difference to note is that the area of SSTs with the strongest correlations to the SOI (<-0.8) includes a larger band of latitudes, extending northward to 5°N, using SST and SOI data from 1900-2019 compared to 1955-1994. Using a longer time series encompasses a larger portion of the Niño 3.4 region compared to the same analysis Barnston et al. 1997 did with a shorter period of data. A potential cause for this minor difference is the SST

dataset Barnston et al. 1997 used for their results. Barnston et al. 1997 did not mention what SST data they used and ERSST data was not available at the time of this publication, which means differences within the SST datasets may cause the differences seen between figures 10 and 11. Using satellite era data from 1979 to 2019, the region of strongest correlations (<-0.8) covers the same area as data that begins in 1900 (Figures 11 and 12). Smoothing the correlation field by using a 50w by 10h EP box using data from 1979 to 2019, the same box size used to calculate the ONI and Niño 3.4 index, still robustly identifies the Niño 3.4 region as the optimal place for an SST box (Figures 13). A similar pattern is observed using data from 1900-2019 for 50w by 10h boxes. This verifies what influenced the choice to use the Niño 3.4 region for an operational oceanic ENSO index by NOAA. The Niño 3.4 region, which lies between the Niño 3 and Niño 4 regions, was chosen based on the strong correlations in DJF between SSTs in that area and the SOI (Trenberth 1997).

Our results build upon what Barnston and Trenberth did. Our analysis uses improved datasets that were not viable to use or available when Barnston and Trenberth did their EP and Niño 3.4 SST region tests. First, there were only ~15 years of satellite-era data, not near enough to be statistically relevant. This would eliminate the potential to use GPCP data in their analysis. Also, there were no satisfactory reconstructed SST datasets available when Trenberth and Barnston wrote their papers (e.g., Smith and Reynolds 2003). Zero-month and one-month leads are both employed in our methods since the oceanic response leads the atmospheric response (Trenberth 1997; Barnston et al. 1997).

5. EIGHT DEGREES OF FREEDOM SEARCH METHOD FOR NDI OPTIONS

Eight degrees of freedom represents the full available amount to scan for a new two-box SST index. The following sub-section uses the full eight degrees of freedom to search for the best candidate box definitions for a new index, using SOI and EXP PC 1 data that spans the length of the satellite era (1979-2019).

In this process all degrees of freedom are allowed to change to analyze which boxes are most suitable for the NDI. The degrees of freedom are box height, width, longitude, and latitude. There are four degrees of freedom for each region, making a total of eight. We define the MC region to be bound by the meridians 80°E/180° and 25°S/29°N parallels and the EP region to be bound by the meridians 160°E/80°W and 13°S/13°N parallels for this search. The box centroid restrictions have odd parallel and even meridian bounds in this test is because of the ERSSTv5 SST dataset 2° by 2° grid boxes center on the odd grid points for latitude and the even grid points for longitude. The maximum allowed height (north-south extent) for the box in the EP region is 12° while the maximum height of the box in the MC region is 28°. The minimum allowed height for boxes in both regions is 4°. Although the longitude bounds for the possible MC and EP regions slightly overlap, we exclude all MC/EP region box pairs that overlap from this search but do not otherwise restrict box width (east-west extent) within this region. We find the 10,000 NDI options whose correlation is the strongest with the SOI for each one-month lead and standard season. Color shading identifies the grid point locations of EP/MC box options for NDI options where the correlations with the SOI are the strongest. Lighter colors represent a high number of grid points for individual EP/MC region NDI box sizes and locations that share that same grid point. The darker colors represent a low number of NDI options that share that same grid point.

Figures 14 shows heat maps: the number of times each grid point falls within one of the 10,000 EP and MC box combinations that produce the largest seasonal correlations between the NDI and the SOI. Both the latitude and longitude of the most frequently selected EP and MC grid points vary for each season. This relates to the changing seasonal location of maximum SSTs in the tropical Pacific during the maturation and decay of ENSO events.

MJJ has a preferred location for the MC box near the equator and ~125°E (Figure 14a). The preferred location for the EP box is centered near 160°W and the equator. The light colored heat map signature for EP box options covers most of the longitudinal EP region domain, becoming weaker east of ~120°W. This implies that short and wide EP SST boxes are more frequently selected. The EP high hits concentration spans ~10° latitude, centered just south of the equator. JJA has a core for MC heat map values at ~5°S and 130°E (Figure 14b). The EP region has a core near the dateline centered at ~5°S that covers the western half of the region (~170°E to ~140°W). The preferred location for the EP box in ASO is constrained farther westward than JJA, covering longitudes 160°E to 160°W, compared to JJA whereas the preferred location for the MC box is near the heart of the MC region and as in JJA is centered at ~10°S and ~130°E (Figures 14c). The preferred box options for this plot overlap near the dateline for the EP and MC choices. SON has an EP box heat core centered near 10°S that extends farther eastward into the EP region compared to ASO, covering the longitude band of ~170°E to 140°W (Figure 14d). The preferred MC region has a short and wide core centered along the EP region at ~10°S, spanning the breadth of longitudes from ~110°E to 150°E.

NDJ exhibits broad and diffused heat cores in the MC and EP regions, which are both centered near the equator (Figure 14e). This implies that the correlations with SOI are not very sensitive to the EP and MC box locations. However, DJF exhibits a different pattern. From NDJ

to DJF, the heat cores decrease in size and the locations of the heat core centers shift northward into the northern areas of the EP and MC regions (Figure 14f). It is interesting that the preferred areas for EP and MC region SST boxes shift so far northward with only a one-month difference in seasonal averages. The FMA box options heat core in the MC region is centered at roughly 135°E and 5°N, whereas the EP region does not have a distinct core (Figure 14g). Rather, box hits are spread across the EP region and extend across the dateline. On the other hand, in MAM the MC region does not have a distinct heat core (Figure 14h) and the EP heat core extends from ~120°W across the dateline to about 170°E along the latitudinal center of 10°N.

A noticeable feature of the aforementioned plots is that the breadth of the heat cores in the EP region shrinks westward from MJJ to ASO. Then, it expands slightly eastward in SON before shifting eastward into the heart of the EP region in NDJ and DJF. These changes likely reflect the spatial evolution of SSTs and the corresponding circulation response due to ENSO.

Similar to the SOI, using EXP PC 1 as the atmospheric ENSO metric to find optimal NDI options leads to distinct seasonal differences in the preferred areas of EP and MC box locations (Figure 15a-h). This confirms that the box locations for the NDI options with the strongest correlations to the SOI and EXP PC 1 are highly sensitive to season using this method. Also, the locations of the highest hits in the EP and MC region for each one-month lead and standard season differs from the SOI. This shows that optimal box locations of NDI candidates are sensitive to the metric used to measure the atmospheric response to ENSO.

There are three issues that arise from this search using eight degrees of freedom. First, different seasons identify different areas in the EP/MC region where there is high overlap of NDI boxes using both SOI and EXP PC 1. This shows that the eight degrees of freedom method does not robustly identify a specific MC and EP region to place boxes for the NDI across all seasons.

Second, the strongest EP/MC box options heat map signatures blend together for some one-month lead and standard seasons, making it hard to distinguish the boundary between the EP and MC region for the placement of boxes. We do not expect optimal EP and MC boxes for the NDI that nearly overlap because they would not sample much of the SST gradient in the tropical Pacific. Third, the NDI options we plot as heat maps for each one-month lead and standard season account for 0.18% (10,000 out of ~5,500,000) of the total amount of options available when allowing for all eight degrees of freedom to vary. The heat maps show a statistically indistinguishable sample size of the total data, even though these are the NDI options with the strongest correlations to the SOI or EXP PC 1.

The highest NDI correlations with the SOI (EXP PC 1) from the eight degrees of freedom search are compared to the correlations of the SOI (EXP PC 1) with the Niño 3.4 index (Table 4-5). Using the SOI, the strongest NDI candidate has slightly higher correlations than the Niño 3.4 index for JJA and SON; however, the Niño 3.4 index has stronger correlations for DJF and MAM (Table 4). Using EXP PC 1, the top NDI candidate has a much stronger correlation than the Niño 3.4 index in JJA and SON. In DJF, the top NDI candidate has a slightly stronger correlation than the Niño 3.4 index. In MAM, the top NDI candidate has a slightly weaker correlation than the Niño 3.4 index (Table 5). This implies that the leading PC's of precipitation may be better ENSO metrics to use for our search of optimal NDI options than the SOI. The top NDI candidate using EXP PC 1 has a much stronger correlation in JJA and SON with this ENSO metric than the Niño 3.4 index whereas

Indeed, allowing all eight degrees of freedom to vary is the most straightforward way to search for a two-box SST index method; however, it does not identify a robust location in the EP or MC regions to place boxes for the NDI. In the next section, we limit the degrees of freedom

in our search for NDI options. This approach narrows down the most important aspects to investigate when searching for NDI options, which is the box centroid location in both the MC and EP region.

6. TWO DEGREES OF FREEDOM SEARCH METHOD FOR NDI OPTIONS

6.1. Overview of process to select NDI candidates

Searching across all eight degrees of freedom did not isolate potential EP or MC box options to use for the NDI. Each seasonal heat map had unique spatial variability, and only samples less than one percent of the available 5,500,000 total NDI options. To constrain our search for NDI boxes to a manageable amount of options to investigate, we limit the search to two degrees of freedom. This method inhibits us from searching for EP and MC boxes for the NDI simultaneously, as we were able to do in the prior section when allowing all eight degrees of freedom to vary. Thus, we first search for boxes within the MC region domain by using the Niño 3.4 index for the EP region. Once MC boxes are selected, the MC region is then held constant by using each selected MC box to search for optimal boxes within the EP domain to create final NDI options. The MC region includes the area between the 80°E/180° meridians, and the 21°S/21°N parallels while the EP region includes the area between the 160°E/80°W meridians and the 15°S/15°N parallels (Figure 16). These region definitions are slightly modified from the ones used in section 5. We constrain the latitudinal extent of the MC region because we want to limit the area of large land masses (Southeast Asia and Australia) included in the search region so that we analyze MC boxes that mostly reside over the ocean. We also slightly increase the latitudinal extent of the EP region to evaluate whether boxes centered in the northern and southern EP will be suitable to use for the NDI. We only include boxes in these searches that fully fit within the defined domain for each respective region. The following sub-sections explain the methods for the MC and EP regions box search, discuss the criteria used for the selection of

boxes in each region, and walk through the selection process for boxes in each region for our final subset of NDI options.

6.2. Selection of MC box options

The Niño 3.4 box is the SST region in use by NOAA to quantify ENSO. Linear regression of the PC/SOI time series with the Niño 3.4 index holds fixed all four degrees of freedom in the EP region. A linearly regressed time series uses the Niño 3.4 index as the predictor for five measures of atmospheric response: NHP PC 1, EXP PC 1, TROP PC 1, and the SOI for the time period that matches the temporal period of GPCP data (1979-2019) as well as for SOI data that matches the full length of our three SST datasets (1900-2019). The residual time series calculation subtracts the linearly regressed response time series from the original response time series for all standard and one-month lead seasons.

Holding MC box size constant further narrows the search down from four degrees of freedom to two after taking into account the EP region. Six possible MC box sizes are considered in this search. The MC region box search utilizes six different box sizes: 10° longitude (lon) by 10° latitude (lat), 30° lon by 10° lat, 10° lon by 30° lat, 30° lon by 18 ° lat, 60° lon by 10° lat and 60° lon by 30° lat. A 10° lat by 10° lon box is a small box size that suppresses the random variability in having just a single grid point correlation. A 60° lon by 10° lat box is the largest short and wide box that yields a fair number of box option given our MC region definition. A 30° lon by 10° lat box cuts the width of the prior box in half, so it represents a more "reasonable" short and wide box option relative to the area of the MC itself. The 10° lon by 30° lat uses a tall and narrow box size that will limit its northward and southward extent of results. The 60° lon by 30° lat box is the largest box one can use before it overlaps too much with the EP region. A 30°

lon by 18° lat box is the aspect ratio relative to the definition of the MC region that we hypothesize will have the strongest correlations with the residuals. These choices incorporate different MC box aspect ratios to determine the effect each has on MC box location preferences. This includes a range of box sizes as well, which will help determine whether a certain box size yields the best MC correlations.

We then compute the correlations between the residual time series and all possible MC box time series' within our defined MC region for all cases one-month lead/standard seasons. Since the PC's are calculated using detrended GPCP data, we use detrended ERSSTv5 SST data for consistency. The correlation values of MC box options with the respective residual time series are assigned to their respective box centers and contoured on a map of the defined MC region for each individual box size. The objective is to identify MC box candidates to use in conjunction with an EP box to improve the NDI's (two-box method) relationship to ENSO induced atmospheric responses over the single-box method like the Niño 3.4 index. The ultimate test of this hypothesis will come later when we compare the NDI correlations with atmospheric ENSO metrics against the Niño 3.4 index.

Kug and Kang (2006) show that Indian Ocean SST's play a major role in the rapid transition to La Niña conditions in the tropical Pacific after strong El Niño events. They look at the relative contributions of the western Indian Ocean SST index (WISST), which uses SSTa's over the western Indian Ocean region (10°S-10°N, 55°E to 75°E), and Niño 3.4 region SSTa's to changes in tropical Pacific SST's over the course of an ENSO event. To do this, they calculate partial correlations of each respective SST region by first removing the affect of the other SST region with zonal surface wind and equatorial surface SST's spanning from 5°S to 5°N. The partial lead/lag correlations of Niño 3.4 region SST's based on NDJ show that the Niño 3.4

region SST's affect MC region SST cooling (negative partial correlations) during the mature phase of El Niño, but that the WISST affects MC warming (positive partial correlations) and EP cooling in MAM and JJA following the mature phase of an El Niño (Figure 17). This confirms that the western Indian Ocean SST's influence the transition and onset of La Niña events following an El Niño event (Kug and Kang 2006). The changes in the SSTa patterns over the course of an ENSO event coinciding with high western Indian Ocean SSTa's (WISST cases) in NDJ are presented in the left panels of figure 18. Based on the seasonal SSTa patterns, it is clear that El Niño events that correspond with anomalously warm WISST are strong in DJF and dissipate quickly in MAM. Compared to WISST El Niño events, the SSTa's in the EP region dissipate slower after the mature phase of El Niño in DJF for non-WISST El Niño events (Figure 18). This provides evidence that anomalously warm western Indian Ocean SST's drive the fast dissipation of strong El Niño's from DJF to MAM. Although it does not directly explain our results for this box search, it is useful to know that western Indian Ocean SST's can influence changes in MC and EP region SST's during the transition from El Niño to La Niña conditions.

Figure 14a shows that the best-performing area for a MC box using the SOI in MJJ is in the center of the MC region, which aligns with the region indicated as the best place to put a MC box for the two degrees of freedom search in figure 19b. Also, in figure 15a and 15b using EXP PC 1 the best-performing area for a MC box is centered over SE Asia, similar to figure 20 and 21 in MJJ and JJA using EXP PC 1 for the two degrees of freedom search. This shows that that the eight degrees of freedom search identified areas in the MC region for a given season where a MC box adds information to the EP box.

Seasonal correlation patterns differ for each case. For example, the NHP case has a strong correlation minimum in JJA for box centers that extend from 95°E-105°E and 8°N-16°N for a

10° lon by 10° lat box (Figure 22b) while the TROP case has a weaker correlation minimum farther eastward at ~140°E and 14°N for the same MC box size (Figures 23b). Also, the correlation patterns tend to be stronger farther north in the NHP where they are stronger farther south in the SOI. In MJJ, a strong correlation minimum exists from ~105°E to 135°E and ~8°S to 0° (Figure 19), farther south compared to MJJ in figure 22a where the correlation minimum is north of the equator. This is consistent with the hemispheres of the predictands. Another example is for the DJF TROP case using a 10° lon by 10° lat box size, a strong correlation minimum is centered just south of the equator along ~160°E (Figure 23) while in the NDJ/DJF EXP case for the same box size the correlation minimum is shifted westward to ~100°E (Figure 21). Different precipitation response regions can be sensitive to different ENSO SST patterns, although sampling error will also contribute to differences.

The changing correlation patterns in the MC region may be an artifact of the relatively short temporal coverage of the GPCP. In other words, the correlation pattern variability from season to season may be smaller when there is a longer time period for analysis. Figure 24 presents the MC box correlations using SOI data spanning from 1900 to 2019 for a 30° lon by 10° lat box. Compared to figure 19 (SOI 1979-2019), the largest negative correlations shift slightly northward in MJJ (Figure 24a). ASO and NDJ also have similar correlation patterns, where NDJ has a correlation maximum that is shifted southeastward to ~6°S 125°E in SOI 1979-2019 compared to SOI 1900-2019. In FMA, the spatial correlation patterns differ between the subplots of the different time periods. As another test of robustness, figures 25 and 26 show the standard season correlation patterns for SOI spanning from 1900-2019 for even and odd years. The correlation patterns are similar for JJA, SON, and DJF; however, they differ in MAM. ENSO events typically decay in MAM, which may account for the lack of agreement in MAM

between figures 25 and 26 in MAM and in FMA between figures 19 and 24a. Also, feedbacks due to ENSO occur in MAM such as changes in trade winds and pressure fields in the tropical Pacific. Otherwise, this shows that the correlation patterns are robust using either time period for this data set. We did not perform this analysis for GPCP data due to the short temporal length of the data set.

We choose box options that have robust negative correlation extrema in both the one-month lead and the zero month lead seasons. That way, we don't have to worry about sensitivity in the one-month lead versus the zero month lead season. We look for areas in the MC region where there will be additional information added for monitoring ENSO using a two-box SST method rather than just using the Niño 3.4 index. Thus, negative correlation extremes are what we look for since the definition of the NDI calls for the MC box to have the opposite sign of the EP box. An example of a box not chosen is the correlation maximum that appears at ~130°E and 8°S in NDJ and DJF using NHP PC 1 (Figures 22,27-28). Positive correlation maxima of the residuals and the MC SST box options do not add any additional information to the NDI since the sign of that MC box is the same as the Niño 3.4 index.

We select six MC box options for further analysis. For our first box candidate, there is a correlation signal that appears in NDJ/DJF EXP case for MC box sizes 10° lon by 10° lat and 30° lon by 10° lat, which leads to our first MC box choice. For each MC box size (Figures 20-21), a moderate correlation minimum appears for MC box centers that vary between ~4°S-8°S/95°E-110°E. For our first candidate, we choose the MC box center location to be at 4°S and 101°E with dimensions 10° lon by 10° lat since the correlation minimum is well defined and the strongest compared to the 30° lon by 10°lat box size. A robust signal appears in a different area in ASO and SON using NHP PC 1 for MC box sizes of 10° lon by 10° lat, 30° lon by 10° lat, and

30° lon by 18° lat, which leads to our second MC box choice. For each MC box size (Figures 22, 27-28) strong correlation minima appear for MC box centers that vary between ~0°-6°S/105°E-115°E. For our second candidate, we choose the MC box center location to be at 0° (equator) and 113°E with dimensions 30° lon by 18° lat. Two reasons exist as to why we use the 30° lon by 18° lat dimensions for the second, third and sixth MC box selections. First, regions of strong correlations for the 30° lon by 18° lat MC box size match up with other MC box sizes for season(s) that contribute to these three MC box selections. Second, the aspect ratio of this MC box size best represents the aspect ratio of our MC region definition out of the box sizes used in this search. For our third and fourth box candidates, a strong correlation minimum exists for the 10° lon by 10° lat and 30° lon by 18° lat box sizes in ASO/SON, NDJ/DJF, and FMA/MAM using TROP PC 1 (Figures 23,29). The minimum is located at ~155°-175°E and ~5°S-5°N for the 10° lon by 10° lat box size and at ~155°E-165°E across different latitudinal bands for each season for the 30° lon by 18° lat box size. For our third and fourth MC box options, the chosen MC box centers are 0°,149°E with box dimensions 30 ° lon by 18° lat and 4°S,149°E with dimensions 10° lon by 10° lat. Both MC box options are selected using the ASO/SON and NDJ/DJF subplots. The box centroids are both moved westward to give the MC box some extra space from the EP region because placing the box center on this longitude axis is still within the area of strong negative correlations in ASO/SON and NDJ/DJF in figures 23 and 29. The fifth box choice, with a box center of 4°N and 130°E, is placed in the heart of the MC region due to its size of 60° lon and 30° lat and the lack of box options available to analyze when using this MC box size for this search. For our sixth candidate, a large correlation minimum appears in the 30° lon by 18° lat, 30° lon by 10° lat, and 10° lon by 10° lat cases in MJJ/JJA for SOI (1900-

2019) between ~125-140°E and 0°-8°S (Figures 24, 30-31). For our sixth MC box option, we choose the MC box center 4°S by 131°E with box dimensions 30° lon by 18° lat.

6.3. EP box selection for MC box candidates: final NDI options

Next, we calculate the NDI with EP boxes for all six MC box options across our defined EP domain and correlate it with the leading PC's of precipitation and the SOI. The criteria for choosing an EP box for each MC box option is to have a correlation maximum when the EP box is in the Niño region. Since the definition of the NDI calls for the EP box to be subtracted by the MC box, higher index values favor El Niño conditions and lower index values favor La Niña conditions like the Niño 3.4 index. The correlations differ than those in section 4 because those are the correlations of EP boxes and the SOI, not the NDI. The same procedure of searching over two degrees of freedom in the MC region is done for the EP region. The four degrees of freedom in the MC region are held constant by using the six MC box options selected in the previous section. Two degrees of freedom in the EP region are accounted for by holding EP box size constant. All EP boxes for the respective MC box choices are constrained to have dimensions 50° lon by 10° lat since the Niño 3.4 index, which uses a 50° lon by 10° lat EP box, holds the EP regions degrees of freedom constant for our MC box selection process.

All MC boxes have EP box centroids that mirror the Niño 3.4 region or are very close to that. All EP boxes chosen use the same measure of atmospheric response as was used to choose the respective MC box option. To illustrate our EP box sensitivity procedure to look for EP boxes that optimally correspond to for our MC box locations, MC box options one through three will be used as examples. In the 50° lon by 10° lat EP box size for MC box option one, the broad area of high correlations in EXP NDJ appears (Figure 32a). However, the correlation maximum

in DJF is smaller and constrained to the western EP region (Figure 32b). Although the correlation maximum in NDJ/DJF is not centered on the Niño 3.4 region box center (0°,145°W), the correlation maximum encompasses that EP box centroid. Since the Niño 3.4 index is used to hold the EP region constant in the MC box search, we choose the EP box center to be 0°,145°W for MC box option one, which mirrors the Niño 3.4 region SST box. For the second MC box, the area of highest correlations in NHP NDJ (Figure 33a) is small and is centered just south of the equator along ~145°W. The correlation maximum broadens in NHP DJF (Figure 33b) and includes the area of highest correlations seen in NHP NDJ. Box option two uses 2°S,145°W as its EP box. This is centered along the Niño 3.4 box meridian center but lies 2°S from its latitudinal center. Box option three has a broad area of high correlations in TROP JJA, ASO/SON, and NDJ/DJF that appears in the heart of the EP region (Figure 34). Box option three uses an EP box centered at 0°,145°E since the area of the highest positive correlations includes the Niño 3.4 region SST box. The seasonal EP region plots using TROP PC for box option four look similar to box option three. Thus, box option four also uses an EP box that mirrors the Niño 3.4 region SST box. Box options five and six use the same EP box center as box options one, three, and four. Given that the highest correlations tend to be similar to the Niño 3.4 EP box correlations, we choose the Niño 3.4 EP box for almost all of our MC box options because this box is widely used and to facilitate comparisons with the MC box sensitivity test which used the Niño 3.4 index to hold the EP region degrees of freedom constant. The coordinates of the final NDI options and a map showing the MC boxes locations are presented in table 6 and figure 35, respectively.

6.4. Supplemental analysis of the two degrees of freedom search

Correlation maps similar to the figures shown in section 6b. are presented in figures 66 and 67, where the only difference is that the box search domain has been expanded eastward into the EP region. The correlation maps give us a better understanding of how the correlation gradients behave outside of the MC region.

The correlations for SST box sizes 10° lon by 10° lat and 30° lon by 18° lat are calculated using the residuals from linear regression of GP PC 1 with the Niño 3.4 index. Slope values of the respective box options and residuals are overlaid. The figures show a correlation and slope gradient increasing to zero from west to east towards and past the dateline (Figures 36-37). The box options become uncorrelated with the residuals near or in the Niño 3.4 SST region in JJA, SON and DJF, which is what we expect since the boxes begin to overlap with the Niño 3.4 SST region. The western MC region is negatively correlated with the residuals in JJA and SON; however, it is positively correlated with the residuals in DJF and MAM (Figures 36-37). This indicates that the western MC region isn't a suitable area to place an additional SST box to monitor ENSO in DJF and MAM since the boxes are positively correlated with the residuals, as outlined in our box selection methods in section 6b. This suggests that the eastern MC region is most suitable to place a SST box to use in conjunction with the Niño 3.4 SST box to monitor ENSO since the correlations with the residuals are negative in all four standard seasons (Figures 36-37). Also, this shows what we expect by implementing linear regression used to find suitable MC boxes for our NDI candidates east of our defined MC region.

Next, raw correlations of a 30° lon by 18° lat box size with the Niño 3.4 SST box in the EP and MC regions show a region of high positive correlations east of the dateline for all

standard seasons, as expected (Figure 38). However, a region of negative correlations becomes straddled between regions of high positive correlations in DJF and MAM covering all latitudes for box centers between longitudes 140°E and 160°E. This further supports that it is likely better to place a MC box in the eastern MC region since we want a MC box that has the opposite sign of the EP box in all seasons for the NDI. Thus, we hypothesize that NDI options three and four, chosen using TROP PC 1, will be the best NDI candidates since their MC boxes are centered in the eastern portion of the MC region at 149°E. This observation will be tested in the next section, where we correlate all NDI options with different atmospheric ENSO metrics and compare it to the Niño 3 and Niño 3.4 index to determine whether any NDI options improve upon the one-box SST method.

7. CORRELATIONS OF FINAL NDI OPTIONS WITH ENSO INDICES AND METRICS

Next, we calculate correlations of the final six NDI options and the Niño 3/3.4 index with different ENSO metrics to determine whether any of the NDI options have a stronger relationship with atmospheric responses to ENSO than the Niño indices. All three SST datasets (ERSSTv5, COBE, HadISST) are included to determine whether these correlations are robust.

One metric we use to determine whether any of the six NDI options is an improvement over a single box method is correlations with GP PC 1 (Tables 7-9). It is noteworthy that the Niño 3 index has weaker correlations with GP PC 1 than the Niño 3.4 index for all seasons. Also, Barnston et al. 1997 also shows that the Niño 3.4 index has stronger correlations with the SOI than the Niño 3 index for almost all rolling seasons (Table 10). Since the NDI choices evaluated in this section use the Niño 3.4 index in the box selection process, we choose to compare our NDI options with only the Niño 3.4 index from this point forward. The correlation of GP PC 1 and the NDI options robustly improve upon the Niño 3.4 index only in SON (Tables 7-9). NDI options one and two have weaker correlations than the Niño 3.4 index in JJA, DJF, and especially MAM. NDI option three has a slightly stronger correlation with GP PC 1 in JJA, while having a slightly weaker correlation in DJF. In MAM, this NDI option has a weaker correlation than the Niño 3.4 index using ERSST and COBE SST data; however, the correlations are nearly identical using HadISST data. NDI option four has nearly identical correlations with the Niño 3.4 index in JJA. In DJF, the correlations are slightly weaker than the Niño 3.4 index. In MAM the correlations are weaker than the Niño 3.4 index. NDI options five and six have slightly stronger correlations than the Niño 3.4 index using all three SST datasets in JJA; however, they have weaker correlations than the Niño 3.4 index and NDI options three and four in DJF and MAM.

Variance explained is compared between the Niño 3.4 and the six NDI options with the CP OLR index data (Tables 11-14). All NDI options have higher variance explained in SON compared to the Niño 3.4 index, with NDI options three and four having the largest difference. NDI options one and two explain less of the variance in the CP OLR index data than the Niño 3.4 index in JJA, DJF and MAM for all SST datasets. NDI options three and four have a higher variance explained for all SST datasets than the Niño 3.4 in JJA, and DJF. In MAM, the variance explained is roughly equivalent to the Niño 3.4 index. NDI options five and six have roughly equivalent variance explained in JJA and DJF months for all SST datasets; however, in MAM it has a weaker relationship to the CP OLR index data compared to the Niño 3.4 index.

Correlations of SOI GPCP length (1979-2019) data with the Niño 3.4 index and final NDI options are also compared (Tables 15-17). Since the correlations between all NDI options and the Niño 3.4 index are negative, we interpret NDI options correlations with a higher absolute value as ones that have a stronger relationship to this ENSO index. In JJA, SON, and DJF almost all NDI options have a stronger correlation with the SOI GPCP length than the Niño 3.4 index for all SST datasets. The only exception is NDI option one, where in SON the correlation strength is nearly identical compared to the Niño 3.4 index for all SST datasets and in MAM it has a weaker correlation using ERSSTv5 data. For DJF, NDI options two through six have nearly identical correlation strength compared to the Niño 3.4 index for all SST datasets.

Correlations of SOI data spanning from 1900-2019 with the Niño 3.4 index and the final NDI options are included. This helps us to further evaluate how our final NDI options perform over a longer time period since most of our NDI options were chosen using data that spans the satellite era. Since there are only minor differences in the correlation values for each SST dataset using SOI GPCP length data and SOI full length, only correlations using ERSSTv5 are included

(Table 18). In JJA, all NDI options except for option one have stronger correlations with the SOI than the Niño 3.4 index. For SON, NDI option one has a weaker correlation while NDI option six has a stronger correlation than Niño 3.4 index data. NDI options two through five have roughly similar correlations. For DJF, NDI option one has a much weaker correlation than the Niño 3.4 index while NDI options two through six have slightly weaker correlations. In MAM, all NDI options have weaker correlations with the Niño 3.4 index. NDI options three through six have the strongest correlations out of all NDI options with SOI 1900-2019 data for this season.

Looking at GP PC 1 and CP OLR correlations/variance tests, NDI options one and two are the weakest out of all the options compared to the Niño 3.4 index, especially in MAM. The MC boxes for NDI options one and two were chosen using the EXP and NHP PC 1, respectively. It is possible that since these NDI options were chosen using remote precipitation impacts, they have a statistically weaker relationship to these metrics than the Niño 3.4 index since these PC's are based on second order impacts due to ENSO. On the other hand, NDI options three through six have similar correlations compared to the Niño 3.4 index for JJA, SON and DJF; however, in MAM they all have weaker correlations than the Niño 3.4 index (Tables 7-9). NDI options three and four have higher variance explained than Niño 3.4 index in JJA, SON and DJF (Tables 11-14). In MAM, the variance explained is nearly identical to the Niño 3.4 index. NDI options five and six explain nearly the same amount of variance as the Niño 3.4 index in JJA, SON, and DJF using CP OLR data (Tables 11-14). In MAM, the Niño 3.4 index outperforms NDI options five and six. Since NDI options three and four were chosen using TROP PC 1, where the first order response of precipitation occurs due to ENSO, it's expected that these options have the strongest relationship to CP OLR data and outperform the Niño 3.4 index in most seasons.

The correlations of the NDI options and the Niño 3.4 index with the SOI weaken when

using the longer time period. Two potential reasons may explain the reduction in correlation strength. First, bias in the early time period (pre-WWII) of the three reconstructed SST datasets used to calculate the NDI options time series may affect correlation strength with the SOI. Since the Niño 3.4 index uses HadISST1 data to generate index values, this potential explanation applies to it as well. Also, it is possible that the signal for ENSO within the NDI options, the Niño 3.4 index, and the SOI has changed over the course of the 1900's to the present day due to global warming.

Using CP OLR data, NDI options three and four distinguish themselves as the top NDI candidates. This indicates that creating a NDI option using the local response of convection to ENSO (TROP) measures the overall global convective response to ENSO better than using the remote response regions (EXP and NHP).

8. EVALUATING ENSO EVENT IDENTIFICATION OF SEASONAL PC TIME SERIES AND NDI OPTIONS VERSUS NIÑO 3.4 INDEX

As stated in section two, we use the seasonal PC 1 time series of precipitation in our selection process for NDI options because it captures the dominant ENSO signal. Figures x through y show the SON and DJF plots for EXP PC 1 and TROP PC 1 to compare the tropical vs. non-tropical PC behavior. Here, strong ENSO events are classified if the normalized PC value exceeds one. Note that this is just a reference point to analyze and compare the ability of each PC to identify OLR ENSO events since the Chiodi and Harrison papers influence our choice to use this metric in the NDI selection process. Additionally, these seasonal PC's are the ones that were used to select NDI options one, three and four.

In SON and DJF of TROP PC 1, the El Niño events in 1982/1983, 1997/1998, and 2015/2016 are clearly the strongest, all three of which were OLR El Niño events (Figures 39-40). The 1991/1992 OLR El Niño event was just below the threshold set to define strong El Niño events in SON, but exceeds the threshold in DJF. This emphasizes that the reference line is by no means the criteria to define a strong ENSO event, but is just a proxy to evaluate the PC's ability to identify strong ENSO events. Both seasons classify the four OLR La Niña events (1988/1989, 1998/1999,1999/2000,2010/2011) with the exception of SON for the 1999/2000 event, which had a value that was near the threshold but did not exceed it (Figures 39-40). Also, some ENSO events identified by the ONI that are not OLR events are classified as strong ENSO events. For example, 2002/2003 is defined as a strong El Niño event although Chiodi and Harrison 2015 does not classify it as an OLR event.

In SON and DJF of EXP PC 1, the 1982/1983, 1997/1998, and 2015/2016 OLR El Niño's

are not clearly defined as the strongest events as they are in TROP PC 1 (Figures 41-42). Also, more El Niño events are classified as strong than in TROP PC 1, although that can be partly attributed to the larger amplitude of each seasonal time series compared to TROP PC 1. In general, TROP PC 1 exhibits more of a non-gaussian behavior than EXP PC 1 in SON and DJF, (Figures 39-42). This makes sense since the tropical region experiences the strongest impacts during ENSO events. The region definition used for TROP PC 1 encompasses the area where the coupled ocean-atmosphere interaction occurs.

We compare NDI options three and four, which were chosen using TROP PC 1, with the Niño 3.4 index to classify ENSO events. This analysis only includes standard season averages of the index values to smooth the data and to make the analysis more straightforward. MAM is not included since ENSO events typically decay during this season and we're interested here in the NDI and Niño 3.4's ability to identify ENSO conditions before or when the strongest extratropical impacts occur in DJF. This only gives an idea of the viability of our recommended NDI options for ENSO event identification using satellite-era data. Further research is needed with model data, which provides a much longer time period for analysis, to verify the utility of the NDI for ENSO monitoring and to set a concrete definition for event identification. The NDI thresholds used in this section to define ENSO events are based on the operational definition of the ONI, which is +/- 0.5°C for five consecutive index values. Here, an El Niño event is defined if the mean seasonal index value exceeds -2°C and a La Niña event is defined if the mean seasonal index value is less than -3°C. ENSO neutral conditions, which ranges from -2°C to -3°C under this definition, matches the SSTa range for ENSO neutral conditions using the ONI definition (1°C). Additionally, we set index thresholds for the NDI and Niño 3.4 index to classify strong ENSO events as an exercise to explore the NDI's and Niño 3.4's ability to distinguish

between OLR vs. non-OLR events. For the NDI, a strong El Niño event is defined if the mean seasonal index value exceeds -1°C and a strong La Niña event is defined if the mean seasonal index value is less than -4°C. For the Niño 3.4 index, a strong El Niño/La Niña event is defined if the mean seasonal index value exceeds +/- 1°C, respectively.

For this analysis, we only include data that spans from 1979 to 2016 since this is the most recent time Chiodi and Harrison identified OLR and non-OLR events, which they define as years that were defined as ENSO events by the ONI but did not meet the threshold to be defined as OLR events, using their ENSO phase based OLR index. Although Chiodi and Harrison 2015 did not classify the 2015/2016 as an OLR El Niño event, it probably was one based on the high seasonal index values in SON and DJF for all three indices (Figures 44-45, 47-48, 50-51). Both NDI options identified all ENSO events that the ONI classified for this time period (Tables 19-24). All indices identified the OLR La Niña events as strong events in 1988/1989, 1998/1999, 1999/2000 and 2010/2011 (Tables 22-24). Also, the 1982/1983, 1997/1998, and 2015/2016 OLR events stand out as the strongest El Niño events in SON and DJF for both NDI options, similar to in TROP PC 1, as well as the Niño 3.4 index (Figures 44-45, 47-48, 50-51). From SON 1991 to DJF 1991/1992, the index value increases past the threshold to define strong El Niño events for all three indices and is one of the strongest El Niño's using each respective index metric, which was classified as an OLR El Niño year by Chiodi and Harrison 2015. An interesting case is the 1986/1987 OLR El Niño, which was followed by a non-OLR El Niño the following year. Based on the set index thresholds we use to define strong El Niño events, the event was classified as a non-OLR event in JJA and SON 1986 for all three indices (Tables 19-21). In DJF 1986/1987, the event becomes classified as a strong El Niño event for NDI option 3 and the Niño 3.4 index; however, it barely exceeds the threshold defined to categorize strong El Niño events (Figures 45,

51). JJA of 1987 has one of the highest values for the full length of the dataset for all three indices (Figures 43, 46, 49), but by DJF of 1987/1988 it becomes categorized as a moderate El Niño for all three indices (Figures 45, 48, 51 and Tables 19-21).

Each event can be categorized based on ENSO "flavor", which encapsulates the typical lifetime of an event and the associated location of the SST warm pool during the mature phase. However, the above discussion highlights that each ENSO event is unique, which can make it difficult to categorize certain ENSO events, like the El Niño event that persisted from 1986 to DJF of 1987/1988.

Some differences exist in ENSO event identification in NDI options three and four compared to the Niño 3.4 index. For example, 1992/1993 is identified as an El Niño event in each NDI option, but not in the Niño 3.4 index (Tables 19-21). Since the NDI uses a different method to measure SST pattern changes in the tropical Pacific associated with ENSO, we expect there to be minor differences in what years are defined as ENSO events. Also, the NDI options indicate El Niño conditions in 2004 (2006) JJA and SON for the 2004/2005 (2006/2007) non-OLR El Niño event; however, it is not identified in DJF 2004/2005 (2006/2007) like it is in the Niño 3.4 index (Tables 19-21). Additionally, NDI option four identifies the non-OLR La Niña events in 2000/2001, 2005/2006, and 2011/2012 as strong La Niña events, while NDI option four and the Niño 3.4 index do not (Tables 22-24). However, since this is a statistically small sample size and operational ENSO SST indices do not have thresholds to distinguish between moderate and strong ENSO events, this does not provide an argument that NDI option three is more representative of ENSO conditions than NDI option four.

Using satellite-era data, the NDI options perform as well as the Niño 3.4 index in ENSO event identification; however, it does not provide an argument that the NDI is a better ENSO

SST index. SST indices only capture the changing SSTa's, or the changing SST gradient in the case of the NDI, in certain regions of the tropical Pacific. Although these indices have been proven to be useful in monitoring ENSO, the Niño indices and the NDI are limited in utility by the geographical location (locations) of the SST box (boxes). Each provides a metric that is a proxy for the overall SST pattern change in the tropical Pacific during an ENSO event. The associated changes in the main locations of convective activity in the tropical Pacific during ENSO events can vary due to differences in the location of the SST warm pool.

Based on this analysis, a hypothetical El Niño (La Niña) event definition for the NDI is for monthly index values to be higher (lower) than -2°C (-3°C) for four consecutive months. A plausible way to identify strong vs. moderate ENSO events in the future is to use the NDI in conjunction with the OLR El Niño and La Niña indices used in Chiodi and Harrison 2015. A strong El Niño or La Niña event would be identified when an ENSO event is already defined using the NDI and the OLR El Niño/La Niña indices from Chiodi and Harrison 2015 identify an OLR event under their definition.

9. SUMMARY AND CONCLUSIONS

ENSO causes changes to global circulation and weather patterns due to shifts in the location of warm SST's in the tropical Pacific. Multiple indices exist to monitor ENSO using oceanic and atmospheric fields that are known to change due to the phenomenon. These indices measure the oceanic and/or atmospheric responses to ENSO in locations/areas where there is high variability of the respective field used for each index. A fundamental future of the evolution of ENSO events is a change in the mean zonal SST gradient. During an El Niño, the SST gradient weakens or reverses, whereas during a La Niña the SST gradient strengthens in the tropical Pacific. Warming SST's due to climate change will likely cause changes to the mean SST gradient in the tropical Pacific. The NDI, the acronym for our proposed ENSO SST index, would be less sensitive to these changes in a future climate since it uses a two-box method to measure the SST gradient in the tropical Pacific. However, one-box SST methods in the Niño regions, such as the ONI, will need new index thresholds to define El Niño or La Niña conditions since different regional SST anomaly magnitudes will reflect active ENSO events due to a different tropical Pacific mean SST gradient.

To narrow down the plethora of NDI options to a small subset, we used the Niño 3.4 index and the SOI as a measure of the SST response in the EP, also known as the eastern Pacific, region and the atmospheric pressure response due to ENSO, respectively. Although the Niño 3.4 index is not the operational Nino 3.4 region SST index in use to monitor ENSO, we used it in our research since it has a longer temporal period of data and its correlation with the ONI indicate that they're nearly identical. We use these indices in tandem with the first PC for Northern Hemisphere precipitation, Extratropical precipitation, and Tropical precipitation to identify NDI

options that have a strong connection to the local and/or remote response of precipitation patterns due to ENSO. ERSSTv5 SST data is used to calculate all potential EP and MC, which is the Maritime Continent region, boxes to select NDI options. First, we search for NDI options across the EP and MC regions using all eight degrees of freedom. Then, we constrain the search to two degrees of freedom to select NDI options. For each selected NDI option, correlations with the first PC of Global precipitation, the central Pacific OLR index, and the SOI are compared to the Niño 3.4 index. COBE and HadISST1 SST data along with the ERSSTv5 are used to calculate the NDI options to test the robustness of the NDI correlations with these ENSO metrics.

Since our methods use the Niño 3.4 index as the metric to gauge the strength of each NDI options relationship to ENSO, we first review important papers that analyze and discuss the viability of the Niño 3.4 region for use as an ENSO index. Grid point correlations of SST's with SOI 1900-2019 and SOI 1979-2019 data in the EP region indicate that the Niño 3.4 region is an optimal area to place a SST box to monitor ENSO. This confirms what Barnston found using SST and SOI data spanning from 1955-1994.

Allowing all eight degrees of freedom to vary in the EP and MC region, we select the ten thousand NDI options that have the strongest correlation with the SOI and first PC of Extratropical precipitation for each one-month lead and standard season. After plotting the data as heat maps, there was no clear consensus for satisfactory NDI options. The areas with the highest overlap between different EP and MC box grid points for the NDI would change locations amongst the one-month and standard seasons. However, most one-month lead and standard seasons clearly placed our second box, which we hypothesized to be located in the MC region, in different areas of this region to use in conjunction with an EP box. This confirms that the MC region is a good location to place a second box to measure the SST gradient in the

tropical Pacific. A shortcoming of the heat maps is that they did not provide any information on the sizes of NDI boxes plotted, making it a sub-optimal method to select strong NDI candidates for further testing.

We then narrow the search down to two degrees of freedom, allowing latitude and longitude to vary for the MC or EP region. This allows us to plot spatial correlation maps to easily identify strong MC/EP box candidates. First, MC box options are selected. For this search, two degrees of freedom are held constant in the MC region by using different prescribed box sizes and all four degrees of freedom are held constant in the EP region by using the Niño 3.4 index. Linear regression of the Niño 3.4 index with four cases, the first PC's of Extratropical precipitation, northern hemisphere precipitation, and tropical precipitation, as well as the SOI, identify the relationship between this ENSO SST index and the atmospheric responses to ENSO. Correlations of the residuals, which is the difference between the original PC/SOI time series and the linearly regressed PC/SOI time series, with all SST boxes in the MC region are plotted. This allowed us to evaluate which MC box locations for a two box method may have a stronger connection to the atmospheric response due to ENSO compared to the Niño 3.4 index, a one box method.

After the MC box options are chosen, we search across the EP region and select EP boxes for each respective MC box to finalize our small set of NDI options for further testing. All four degrees of freedom in the MC region are held constant by using the previously selected MC box options. Since we use the Niño 3.4 index to help select our MC boxes, we constrain the EP box size to 50° lon by 10° lat, the same box size used to calculate the Niño 3.4 index, to hold two degrees of freedom constant in the EP region. Since the purpose of the NDI is to measure the tropical Pacific SST gradient by taking the difference between an EP box and a MC box,

correlations of the NDI options with each case used in the MC box search are calculated. As expected, the final NDI options chosen use an EP box that mirrors or is very close to the Niño 3.4 region.

We then verify the eastern meridian used to define our MC region for our NDI search is sensible. The MC region used for our MC box option selection process is expanded eastward from the dateline to 140°W, overlapping with the Niño 3.4 region. Using the same type of analysis as was done in our two degrees of freedom MC box selection process, the correlations of the residuals with the box option candidates trend to zero east of the dateline for all standard seasons. This shows what we expected since the Niño 3.4 index is used to hold the EP region constant for this analysis. Also, it verifies that our original MC region definition is viable.

To determine whether any of the selected NDI candidates will be recommended for future work, we compare the Niño 3.4 index and each candidates correlations with different metrics that measure the atmospheric response due to ENSO. Using global precipitation PC 1 and the central Pacific OLR index correlations, NDI options three and four perform the best. In JJA, SON and DJF, these options have stronger or roughly equivalent correlations than the Niño 3.4 index; however, in MAM, the Niño 3.4 index has stronger correlations. There is no clear NDI candidate looking at SOI correlations, although it is interesting to note that NDI options two through six perform better than the Niño 3.4 index in all seasons during the satellite-era (1979-2019) except in DJF, where correlations are roughly equivalent. On the other hand, the Niño 3.4 index performs better than the NDI options in DJF and MAM using SOI data going back to the beginning of the twentieth century (1900-2019). In JJA, NDI options two through six have stronger correlations than the Niño 3.4 index and roughly equivalent correlations in the SON.

Since we use the convective response to ENSO, which was the motivation of creating the OLR El Niño and La Niña index in the Chiodi and Harrison papers, we recommend NDI options three and four for future work. NDI option three has an MC box center located at 0° and 149°E that is 30° longitude in width and 18° latitude in height. The EP box center is located at 0°, 145°W and is 50° longitude in width and 10° latitude in height, which is the same box size and location used to calculate the ONI. NDI option four has an MC box center located at 4°S and 149°E that is 10° longitude in width and height with the same EP box as NDI option four. While NDI options three and four's correlations with global precipitation PC 1 had statistically negligible differences compared the Niño 3.4 index, their correlations with the central Pacific OLR index identify them as a potentially stronger ENSO index than the Niño 3.4 index. We used variance to calculate the difference between the NDI options and the Niño 3.4 index. In JJA and SON, NDI options three and four explained seven to eleven percent more of the variance in central Pacific OLR index values than the Niño 3.4 index using all three SST datasets. In DJF, these NDI options explained roughly three to four percent more variance than the Niño 3.4 index across all SST datasets while in MAM their variance explained was roughly equivalent.

To analyze the ability of the recommended NDI options to identify ENSO events leading up to and during the traditional mature phase of ENSO in DJF, standard season time series of NDI options three, four and the Niño 3.4 index are generated and evaluated. The ENSO event threshold definition for the ONI is used for the Niño 3.4 index, and an ENSO event definition is created for the NDI based on the definition of the ONI. Additionally, an additional threshold is set to identify strong El Niño's/La Niña's since we compare moderate vs. strong ENSO events under our definitions identified for these indices with OLR vs. non-OLR ENSO events identified by Chiodi and Harrison 2015. Both NDI options and the Niño 3.4 index identified all OLR/non-

OLR ENSO events from Chiodi and Harrison 2015. Although there are some differences between the classification of moderate vs. strong ENSO events for given seasons between the Niño 3.4 index and the NDI options, the analysis displayed that the NDI performs as well as the Niño 3.4 index; however, it does not provide any evidence that the NDI is a better ENSO SST index than the Niño 3.4. Additionally, only satellite-era data is used for this analysis since reliable OLR data becomes available in 1979. This is a statistically small sample size of data, thus it only provides an idea of how the NDI performs compared to the Niño 3.4 index.

Future work will include working with model output to test the robustness of these recommended NDI options chosen in past and future climates. Over the next two years we will use different climate models and model runs from CMIP5 or CMIP6 to determine if the Niño 3.4 SST box performs better than the recommended NDI options for a much larger temporal dataset than what is used in this paper. This will also help determine if the Niño 3.4 SST box and the recommended NDI options behavior changes in a future climate.

REFERENCES

Adler, R.F., and Coauthors, 2003: The Version-2 Global Precipitation Climatology Project (GPCP) Monthly Precipitation Analysis (1979-Present). *Journal of Hydrometeorology*, 4, 1147-1167.

— —, 2018: The Global Precipitation Climatology Project (GPCP) Monthly Analysis (New Version 2.3) and a Review of 2017 Global Precipitation. *Atmosphere (Basel)*, **9**, 1-14.

Allan, R.J., Nicholls, N., Jones, P.D. and Butterworth, I.J., 1991: A further extension of the Tahiti-Darwin SOI, early SOI results and Darwin pressure. *J. Climate*, **4**, 743-749.

Barnston, A.G., M. Chelliah, and S.B. Goldenberg, 1997: Documentation of a Highly ENSO-related SST Region in the Equatorial Pacific: Research Note. *Atmosphere-Ocean Journal*, **35**, 367-383.

Cai, W., and Coauthors 2015: ENSO and greenhouse warming. Nat. Climate Change, **5**, 849.

Chiodi, A.M., and D.E. Harrison, 2010. Characterizing Warm-ENSO variability in the Equatorial Pacific. An OLR perspective. *Journal of Climate,* **23**, 2428-2439.

— —, 2013: El Niño Impacts on Seasonal U.S. Atmospheric Circulation, Temperature, and Precipitation Anomalies: The OLR-Event Perspective. *Journal of Climate*, **28**, 822-837.

— —, 2015: Global Seasonal Precipitation Anomalies Robustly Associated with El Niño and La Niña events- An OLR Perspective. *Journal of Climate*, **28**, 6133-6159.

Collins, M., 2000: Understanding uncertainties in their response of ENSO to greenhouse warming. Geophys. Res. Lett., **27**, 3509-3512.

Drouard, M. and C. Cassou, 2019: A Modeling- and Process-Oriented Study to Investigate the Projected Change of ENSO-Forced Wintertime Teleconnectivity in a Warmer World, *Journal of Climate*, **32**, 8047-8068.

Folland, C.K, and D.E. Parker, 1995: Correction of instrumental biases in historical sea surface temperature data. *Quarterly Journal of the Royal Meteorological Society*, **121,** 319-367.

Gonzalez, P. L. M., L. Goddard, 2015: Long-lead ENSO predictability from CMIP5 decadal hindcasts. *Climate Dynamics*, **46**, 3127-3147.

Ham, Y., Y. Jeong and J. Kug, 2015: Changes in Independency between Two Types of El Niño Events under a Greenhouse warming Scenario in CMIP5 Models. *Journal of Climate*, **28**, 7561-7575.

Ham. Y, J. Kug, 2015: ENSO amplitude changes due to greenhouse warming in CMIP5: Role of mean tropical precipitation in the twentieth Century, *Geophysical Research Letters*, **43**, 422-430.

Hirahara, S, 2014.: Centennial-Scale Sea Surface Temperature Analysis and Its Uncertainty. *Journal of Climate*, **27**, 57-75.

Huang, B, 2016: Further Exploring and Quantifying Uncertainties for Extended Reconstructed Sea Surface Temperature (ERSST) Version 4 (v4). *Journal of Climate*, **29**, 3119-3142.

Huang, B. and Coauthors, 2017: Extended Reconstructed Sea Surface Temperature, Version 5 (ERSSTv5): Upgrades, Validations, and Intercomparisons. *Journal of Climate*, **30**, 8179-8205.

Huang, B., C. Liu, G. Ren, H.M. Zhang, and L. Zhang, 2018: The role of buoy and Argo observations in two SST analyses in the global and tropical Pacific oceans. *J. Climate*, **32**, 2517-2535.

Ishii, M., A. Shouji, S. Sugimoto, T. Matsumoto, 2005: Objective analyses of sea-surface temperature and marine meteorological variables for the 20th century using ICOADS and the Kobe Collection. *International Journal of Climatology*, **25**,865-879.

Jones, P.D.,T.J. Osborn, and N.A. Rayner, 2001: Adjusting for sampling density in grid box land and ocean surface temperature time series. *Journal of Geophysical Research*, **106**, 3371-3380.

Kousky, V.E. and R. W. Higgins 2007: An Alert Classification System for Monitoring and Assessing the ENSO cycle. *Weather and Forecasting*, **22**, 353-371.

Kug, J.S and I.S. Kang 2006. Interactive Feedback between ENSO and the Indian Ocean. *Journal of Climate*, **19**, 1784-1801.

Lindsey R., 2013: In Watching for El Niño and La Niña, NOAA Adapts to Global Warming. *Climate.gov*, National Oceanic and Atmospheric Administration.

Meehl, G. A., H. Teng, and G. Branstator, 2006: Future changes of El Niño in two global coupled climate models. Climate Dyn., **26**, 549-566.

Neelin, J.D., 2011: El Niño and year-to-year climate prediction. *Climate Change and Climate Modeling*, Cambridge University Press, 103-144.

Oh, J., D.W. Shin, S. Cocke., G. Baigorria,: 2014: ENSO Teleconnection Pattern Changes over the Southeastern United States under a Climate Change Scenario in CMIP5 Models, **2014**, 16 pages.

Rashid, H.A., and Coauthors, 2016: An atmospheric mechanism for ENSO amplitude changes under an abrupt quadrupling of CO2 concentration in CMIP5 models. *Geophysical Research Letters*, **43**, 1687-1694.

Rayner, N.A. and Coauthors, 2003: Global analyses of sea surface temperature,sea ice, and night time marine air temperature since the late nineteenth century. *Journal of Geophysical Research,* **108**, 1-29.

Ropelewski, C.F, and P.D. Jones, 1987: An Extension of the Tahiti-Darwin Southern Oscillation Index. *Monthly Weather Review*, **115**, 2161-2165.

Smith, T.M., R.W. Reynolds, 2003: Extended Reconstruction of Global Sea Surface Temperatures Based on COADS Data (1854-1997). *Journal of Climate*, **16**,1495-1510.

Song, W., Q. Dong, and C. Xue, 2016: A classified El Niño index using AVHRR remote-sensing SST data. *International Journal of Remote Sensing*, **37**, 403-417.

Trenberth, K.E., 1997: The definition of El Niño. *Bulletin of the American Meteorological Society*, **78**, 2771-2777.

Walker,G.T., and E.W. Bliss, 1932: World Weather V. *Memoirs of the Royal Meteorological Society*, **4**, 53-84.

Williams, I., C.M. Patricola, 2018: Diversity of ENSO events Unified by Convective Threshold Sea Surface Temperature: A Nonlinear ENSO Index. *Geophysical Research Letters,* **45**, 9236-9244.

Wolter, K., and M.S. Timlin, 1993: Monitoring ENSO in COADS with a seasonally adjusted principal component index. *17th Climate Diagnostics Workshop*, Norman OK, Climate Analysis Center, National Meteorological Center, National Oceanic and Atmospheric Administration, 52-57.

— —, 2011: El Niño/Southern Oscillation behavior since 1871 as diagnosed in an extended multivariate ENSO index (MEI.ext), *International Journal of Climatology*, **31**, 1074-1087.

Wu, Xian, Y.M. Okumura, P. N. DiNezio, 2019: What Controls the Duration of El Niño and La Niña Events. *Journal of Climate*, **31**,5941-5965.

Correlations of Niño 3.4 index and ONI	
JJA	0.978
SON	0.990
DJF	0.996
MAM	0.980

Table 1. Standard season correlations of the Niño 3.4 index and ONI for the period of January 1979-February 2019.

Zero Month Leads	PC 1	JJA	SON	DJF	MAM
	SOI	0.87	0.82	0.90	0.69
	Niño 3.4 index	0.90	0.94	0.95	0.90
	PC 2	JJA	SON	DJF	MAM
	SOI	0.03	0.21	0.05	0.32
	Niño 3.4 index	0.19	0.08	0.02	0.27

Table 2. Correlations between SOI/Niño 3.4 and PC 1 and 2 of global precipitation for zero month leads.

One Month Leads	PC 1	MJJ	ASO	NDJ	FMA
	SOI	0.79	0.90	0.86	0.84
	Niño 3.4 index	0.82	0.94	0.95	0.89
	PC 2	MJJ	ASO	NDJ	FMA
	SOI	0.13	0.17	0.05	0.18
	Niño 3.4 index	0.37	0.05	0.05	0.19

Table 3. Same as in Table 2 but for one month leads.

Correlations	Top NDI option vs. SOI	Niño 3.4 index vs. SOI
JJA	-0.83	-0.81
SON	-0.85	-0.84
DJF	-0.88	-0.90
MAM	-0.69	-0.80

Table 4. Comparison of the highest correlation for the NDI in each season calculated using the eight degrees of freedom method and SOI vs. Niño 3.4 index.

Correlations	Top NDI option vs. EXP PC 1	Niño 3.4 index vs. EXP PC 1
JJA	0.89	0.58
SON	0.92	0.82
DJF	0.92	0.89
MAM	0.83	0.84

Table 5. Comparison of the highest correlation for the NDI in each season calculated using the eight degrees of freedom method and EXP PC 1 vs. Niño 3.4 index.

	MC box centroid	MC box size	EP box centroid	EP box size
NDI Option 1	4°S,101°E	10° lon by 10° lat	0°,145°W	50° lon by 10° lat
NDI Option 2	0°S,113°E	30° lon by 18° lat	2°S,145°W	50° lon by 10° lat
NDI Option 3	0°S,149°E	30° lon by 18° lat	0°,145°W	50° lon by 10° lat
NDI Option 4	4°S,149°E	10° lon by 10° lat	0°,145°W	50° lon by 10° lat
NDI Option 5	4°N,130°E	60° lon by 30° lat	0°,145°W	50° lon by 10° lat
NDI Option 6	4°S,131°E	30° lon by 18° lat	0°,145°W	50° lon by 10° lat

Table 6. Final NDI options coordinate details.

ERSSTv5 corr GP vs. NDI Options/ Niño 3.4 index	JJA	SON	DJF	MAM
NDI OPTION 1	0.82	0.95	0.90	0.47
NDI OPTION 2	0.88	0.96	0.91	0.69
NDI OPTION 3	0.91	0.95	0.95	0.83
NDI OPTION 4	0.90	0.95	0.94	0.80
NDI OPTION 5	0.92	0.95	0.93	0.77
NDI OPTION 6	0.92	0.96	0.92	0.74
Niño 3.4 index	0.90	0.94	0.95	0.90
Niño 3 index	0.86	0.93	0.92	0.88

Table 7. Standard season correlations between first PC of global precipitation and final NDI options, with Niño 3.4 as a comparison. ERSSTv5 SST data is used to calculate time series for NDI options.

COBE Corr GP vs. NDI options/ Niño 3.4 index	JJA	SON	DJF	MAM
NDI OPTION 1	0.85	0.96	0.89	0.63
NDI OPTION 2	0.88	0.96	0.91	0.77
NDI OPTION 3	0.92	0.95	0.95	0.88
NDI OPTION 4	0.90	0.95	0.94	0.85
NDI OPTION 5	0.92	0.95	0.93	0.82
NDI OPTION 6	0.91	0.96	0.93	0.79
Niño 3.4 index	0.90	0.94	0.95	0.90
Niño 3 index	0.86	0.93	0.92	0.88

Table 8. Same as in Table 8 except using COBE SST data.

HadISST Corr GP vs. NDI options/ Niño 3.4 index	JJA	SON	DJF	MAM
NDI OPTION 1	0.87	0.96	0.92	0.72
NDI OPTION 2	0.89	0.95	0.92	0.81
NDI OPTION 3	0.91	0.95	0.95	0.89
NDI OPTION 4	0.90	0.94	0.94	0.85
NDI OPTION 5	0.93	0.95	0.94	0.84
NDI OPTION 6	0.92	0.95	0.93	0.82
Niño 3.4 index	0.90	0.94	0.95	0.90
Niño 3 index	0.86	0.93	0.92	0.88

Table 9. Same as in Table 8 except using HadISST data.

Season	Correlation between SOI and Niño 3	Correlation between SOI and Niño 3.4	Significance of difference	Correlation between Niño 3 and Niño 3.4
Dec.–Jan.–Feb.	−0.83	−0.84	0.73	0.91
Jan.–Feb.–Mar.	−0.85	−0.85	1.00	0.89
Feb.–Mar.–Apr.	−0.80	−0.82	0.67	0.87
Mar.–Apr.–May.	−0.70	−0.78	0.26	0.84
Apr.–May.–Jun.	−0.67	−0.79	0.08	0.85
May.–Jun.–Jul.	−0.66	−0.79	0.08	0.86
Jun.–Jul.–Aug.	−0.69	−0.82	**0.046**	0.84
Jul.–Aug.–Sep.	−0.77	−0.87	**0.041**	0.84
Aug.–Sep.–Oct.	−0.80	−0.87	0.13	0.88
Sep.–Oct.–Nov.	−0.84	−0.87	0.49	0.91
Oct.–Nov.–Dec.	−0.80	−0.81	0.85	0.92
Nov.–Dec.–Jan.	−0.80	−0.81	0.94	0.93
Annual	−0.77	−0.83	0.25	0.88

Table 10. Reprinted from Barnston et al. 1997. "Comparison of correlations between Niño 3 and the SOI and Niño 3.4 and SOI, over the 1950-96 period for running 3-month seasons. The probability that the correlation difference could occur by chance is given the third column, using a 2-sided test based on Fisher r-to-Z transformation. The effective sample size is 47 for the seasonal analysis and 56 for the annual. Significant results at the 0.05 level or better are shown in bold. The Niño 3 versus Niño 3.4 SST correlation is also shown."

JJA	NDI Option 1	NDI Option 2	NDI Option 3	NDI Option 4	NDI Option 5	NDI Option 6
ERSSTv5	-0.21	-0.11	0.10	0.09	0.01	-0.02
COBE	-0.18	-0.09	0.09	0.09	-0.01	-0.02
HadISST	-0.14	-0.05	0.10	0.10	0.04	0.01

Table 11. Variance explained difference between the NDI and the Niño 3.4 index using CP OLR for JJA. Positive values represent where the NDI option has higher variance explained than the Niño 3.4 index.

SON	NDI Option 1	NDI Option 2	NDI Option 3	NDI Option 4	NDI Option 5	NDI Option 6
ERSSTv5	0.04	0.04	0.10	0.09	0.05	0.05
COBE	0.05	0.05	0.11	0.09	0.06	0.06
HadISST	0.04	0.06	0.09	0.07	0.06	0.05

Table 12. Same as Table 11 except for SON

DJF	NDI Option 1	NDI Option 2	NDI Option 3	NDI Option 4	NDI Option 5	NDI Option 6
ERSSTv5	-0.08	-0.04	0.04	0.03	-0.01	-0.02
COBE	-0.08	-0.05	0.04	0.03	-0.01	-0.02
HadISST	-0.05	-0.05	0.04	0.02	-0.01	-0.02

Table 13. Same as Table 11 except for the DJF

MAM	NDI Option 1	NDI Option 2	NDI Option 3	NDI Option 4	NDI Option 5	NDI Option 6
ERSSTv5	-0.45	-0.24	-0.01	-0.01	-0.10	-0.13
COBE	-0.30	-0.16	0.01	0.02	-0.06	-0.10
HadISST	-0.17	-0.05	0.05	0.01	0.01	-0.04

Table 14. Same as in Table 11 except for the MAM

ERSST corr NDI Options/ Niño 3.4 index and SOI	JJA	SON	DJF	MAM
NDI OPTION 1	-0.85	-0.82	-0.86	-0.73
NDI OPTION 2	-0.89	-0.87	-0.87	-0.84
NDI OPTION 3	-0.88	-0.87	-0.90	-0.82
NDI OPTION 4	-0.89	-0.88	-0.90	-0.83
NDI OPTION 5	-0.90	-0.88	-0.89	-0.83
NDI OPTION 6	-0.91	-0.88	-0.87	-0.85
Niño 3.4 index	-0.81	-0.84	-0.90	-0.79

Table 15. Standard season correlations between SOI GPCP length (1979-2019) and final NDI options, with Niño 3.4 as a comparison. ERSSTv5 SST data is used to calculate time series for NDI options.

COBE corr NDI Options/ Niño 3.4 index and SOI	JJA	SON	DJF	MAM
NDI OPTION 1	-0.85	-0.84	-0.87	-0.80
NDI OPTION 2	-0.87	-0.88	-0.89	-0.83
NDI OPTION 3	-0.87	-0.85	-0.90	-0.82
NDI OPTION 4	-0.88	-0.86	-0.90	-0.83
NDI OPTION 5	-0.87	-0.88	-0.90	-0.84
NDI OPTION 6	-0.89	-0.88	-0.90	-0.85
Niño 3.4 index	-0.81	-0.84	-0.90	-0.79

Table 16. Same as in Table 15 except using COBE SST data

HadISST corr NDI Options/ Niño 3.4 index and SOI	JJA	SON	DJF	MAM
NDI OPTION 1	-0.86	-0.83	-0.86	-0.82
NDI OPTION 2	-0.88	-0.87	-0.89	-0.84
NDI OPTION 3	-0.88	-0.86	-0.91	-0.81
NDI OPTION 4	-0.88	-0.88	-0.91	-0.81
NDI OPTION 5	-0.88	-0.88	-0.90	-0.82
NDI OPTION 6	-0.90	-0.88	-0.89	-0.84
Niño 3.4 index	-0.81	-0.84	-0.90	-0.79

Table 17. Same as in Table 15 except using HadISST SST data.

ERSSTv5 corr NDI Options/ Niño 3.4 index and SOI	JJA	SON	DJF	MAM
NDI OPTION 1	-0.64	-0.72	-0.76	-0.50
NDI OPTION 2	-0.71	-0.76	-0.80	-0.57
NDI OPTION 3	-0.69	-0.75	-0.83	-0.59
NDI OPTION 4	-0.70	-0.76	-0.82	-0.56
NDI OPTION 5	-0.72	-0.77	-0.83	-0.61
NDI OPTION 6	-0.76	-0.78	-0.82	-0.61
Nino 3.4 index	-0.66	-0.76	-0.85	-0.66

Table 18. Correlations between SOI full length (1900-2019) and final NDI options, with Niño 3.4 index as a comparison. ERSSTv5 SST data is used to calculate time series for NDI options.

NDI Option 3 OLR/Non-OLR EN			
	JJA	SON	DJF
	1979	1979	1979/1980
	1980		
	1982	1982	1982/1983
	1983		
	1984		
	1986**	1986**	1986/1987
	1987**	1987**	1987/1988
	1990		1990/1991
	1991	1991**	1991/1992
	1992		1992/1993
	1993	1993	
	1994	1994	1994/1995
	1997	1997	1997/1998
	2001		
	2002**	2002	2002/2003
	2003		
	2004	2004	
	2005		
	2006	2006	
	2009	2009	2009/2010**
	2012		
	2015	2015	2015/2016

Table 19. List of years where the mean index values for a given season meet the criteria to be defined as an El Niño event for NDI option 3. A moderate El Niño event is defined when the mean index values for a given season exceeds -2°C and a strong El Niño event is defined when the mean index values for a given season exceeds -1°C. Red text (gray highlighted boxes) indicate defined OLR (non-OLR) events from Chiodi and Harrison 2015. Black (red) stars indicate when a defined OLR (non-OLR) El Niño event was labeled as a moderate (strong) El Niño event based on our NDI event definition. Included years that are plain black text and don't have a highlighted box are years that meet the event definition criteria for the NDI but are not OLR or non-OLR events from Chiodi and Harrison 2015.

NDI Option 4 OLR/Non-OLR EN			
	JJA	SON	DJF
	1979	1979	
	1980		
	1981		
	1982	1982	1982/1983
	1983		
	1984		
	1986**	1986**	1986/1987**
	1987**	1987**	1987/1988
	1990	1990	
	1991	1991**	1991/1992
	1992	1992	1992/1993
	1993	1993	
	1994**	1994	1994/1995
	1997	1997	1997/1998
	2001		
	2002**	2002	2002/2003
	2003		
	2004**	2004	
	2005		
	2006	2006	
	2008		
	2009	2009	2009/2010
	2012		
	2014		
	2015	2015	2015/2016

Table 20. Same as in Table 19 except for NDI option 4.

Niño 3.4 index OLR vs. Non-OLR EN			
	JJA	SON	DJF
			1979/1980
	1982**	1982	1982/1983
		1986**	1986/1987
	1987**	1987**	1987/1988
	1991**	1991**	1991/1992
		1994	1994/1995**
	1997	1997	1997/1998
	2002	2002**	2002/2003**
		2004	2004/2005
		2006	2006/2007
	2009	2009**	2009/2010**
		2014	2014/2015
	2015	2015	2015/2016

Table 21. Same as in Table 19 except for Niño 3.4 index.

NDI Option 3 OLR/Non-OLR LN			
	JJA	SON	DJF
		1983	1983/1984
		1984	1984/1985
			1985/1986
	1988**	1988	1988/1989
		1989	
		1995	1995/1996
		1996	1996/1997
		1998	1998/1999
		1999**	1999/2000
		2000	2000/2001
		2005	2005/2006
		2007**	2007/2008**
		2008	2008/2009
	2010**	2010	2010/2011
		2011	2011/2012

Table 22. List of years where the mean index values for a given season meet the criteria to be defined as a La Niña event for NDI option 3. A moderate La Niña event is defined when the mean index values for a given season is lower than -3°C and a strong La Niña event is defined when the mean index values for a given season is lower than -4°C. Blue text (gray highlighted boxes) indicate defined OLR (non-OLR) events from Chiodi and Harrison 2015. Black (blue) stars indicate when a defined OLR (non-OLR) La Niña event was labeled as a moderate (strong) La Niña event based on our NDI event definition. Included years that are plain black text and don't have a highlighted box are years that meet the event definition criteria for the NDI but are not OLR or non-OLR events from Chiodi and Harrison 2015.

NDI Option 4 OLR/Non-OLR LN			
	JJA	SON	DJF
			1980/1981
			1983/1984
			1984/1985
			1985/1986
	1988**	1988	1988/1989
		1995	1995/1996**
		1996	1996/1997
		1998	1998/1999
		1999**	1999/2000
		2000	2000/2001**
			2005/2006**
		2007**	2007/2008**
		2008	2008/2009
	2010**	2010	2010/2011
		2011	2011/2012**

Table 23. Same as in Table 22 except for NDI option 4.

Niño 3.4 index OLR vs. Non-OLR LN			
	JJA	SON	DJF
		1983	1983/1984
	1984	1984	1984/1985**
	1985	1985	1985/1986
	1988	1988	1988/1989
	1989		
		1995	1995/1996
	1998**	1998	1998/1999
	1999**	1999	1999/2000
	2000	2000	2000/2001
			2005/2006
		2007**	2007/2008**
			2008/2009
	2010**	2010	2010/2011
		2011	2011/2012

Table 24. Same as in Table 22 except for Niño 3.4 index.

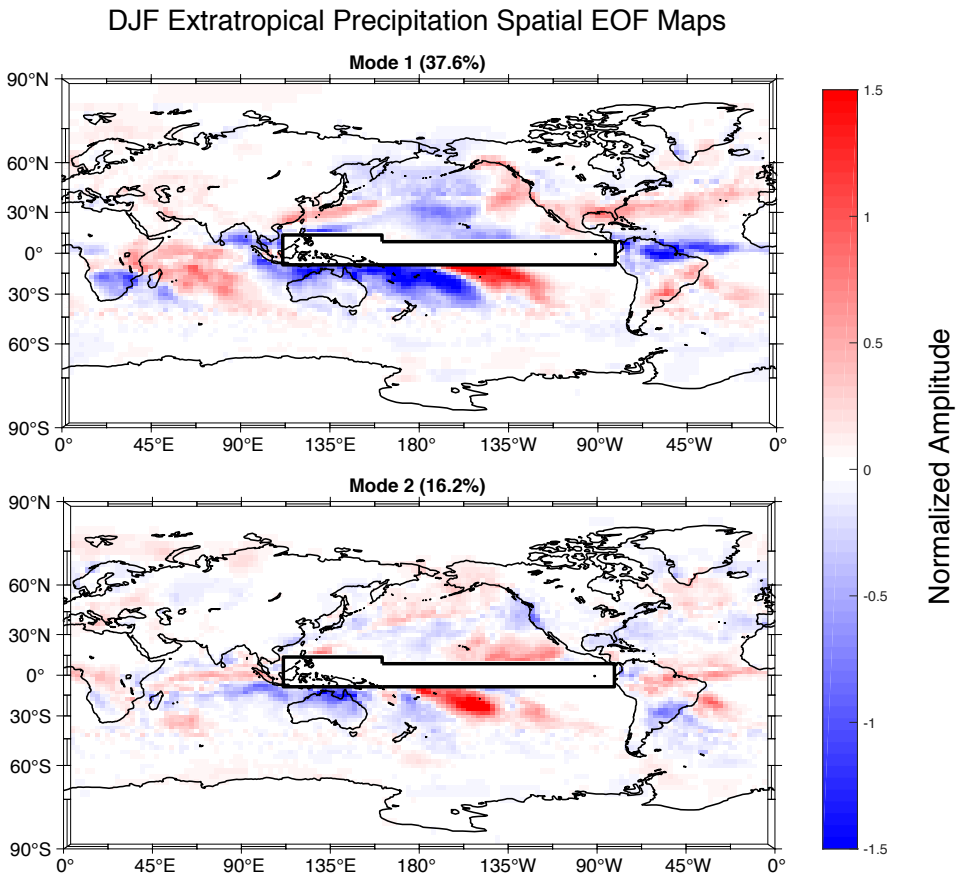


Figure 1. DJF spatial variability map of the leading two modes of variability for EXP (Remote Precipitation) using PC analysis.

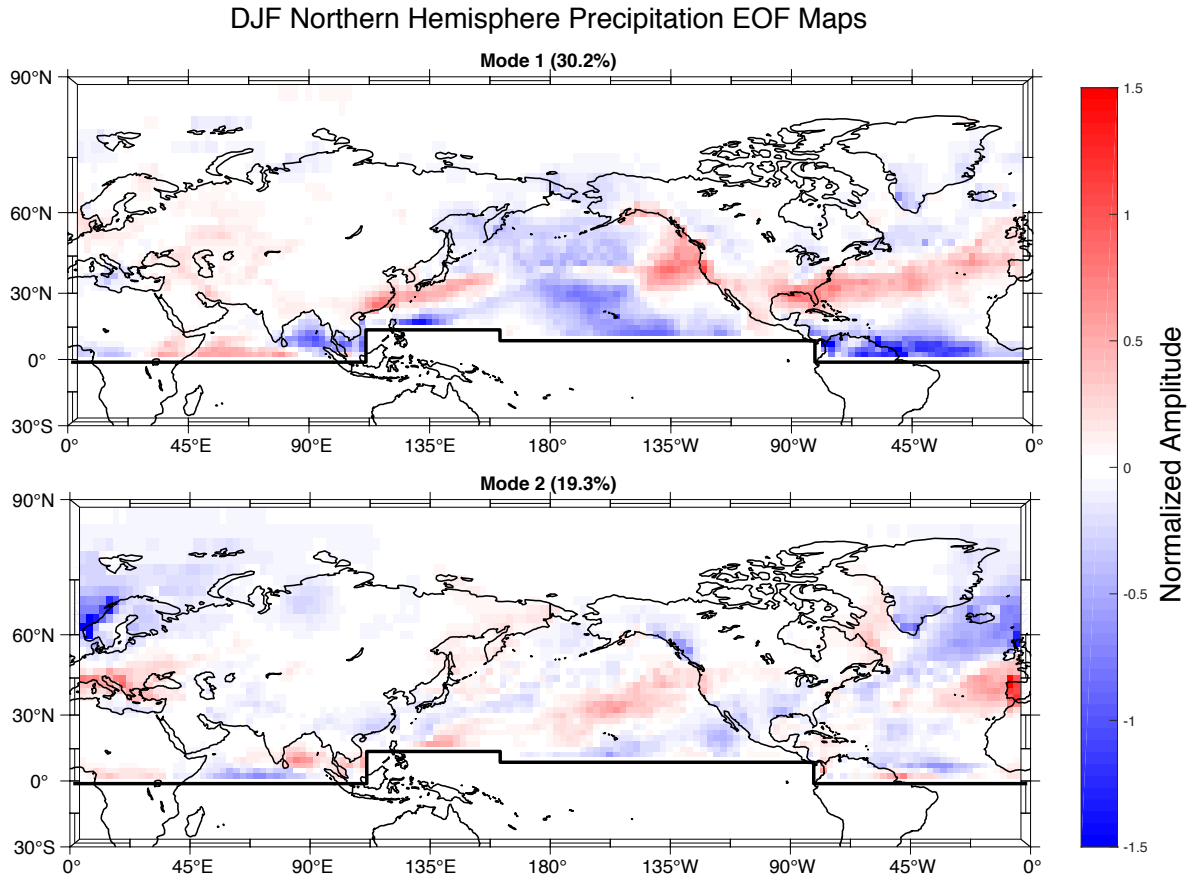


Figure 2. DJF spatial variability map of the leading two modes of variability for NHP (Remote Northern Hemisphere Precipitation) using PC analysis.

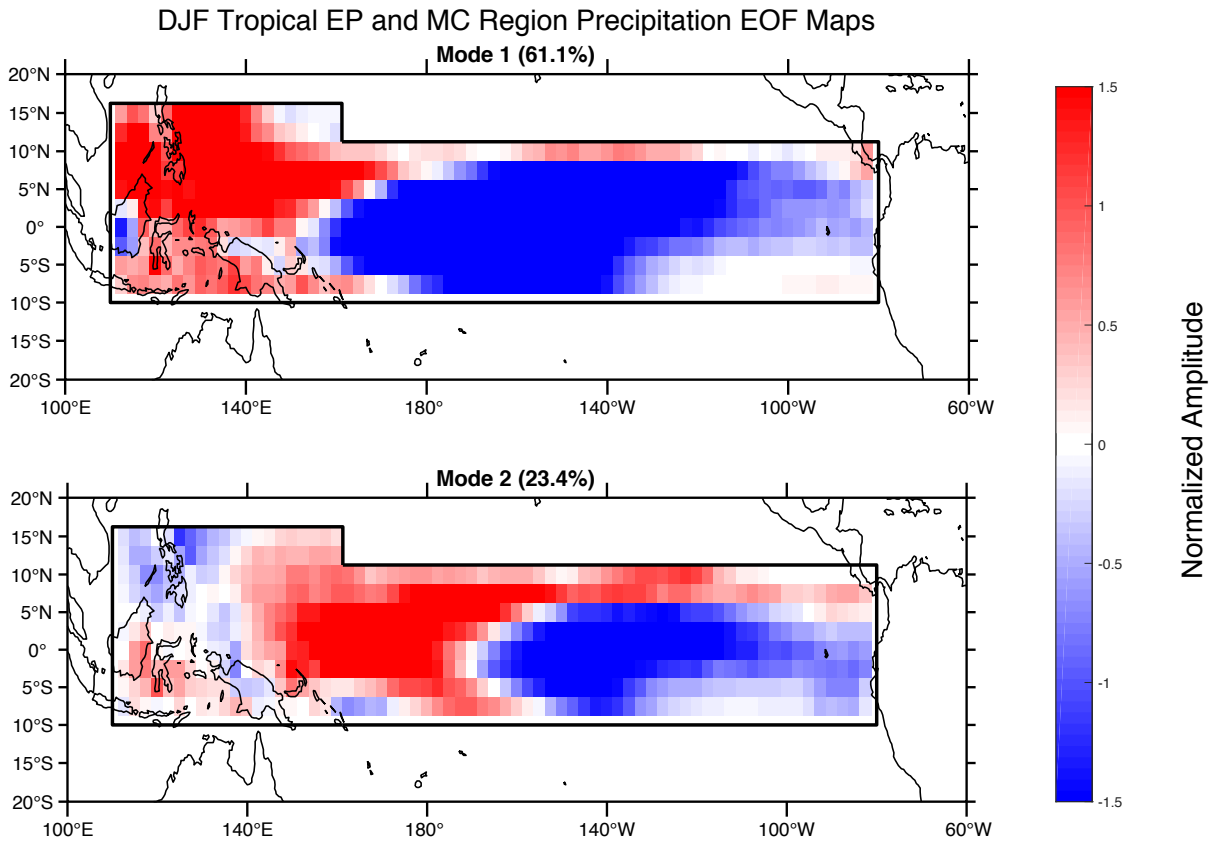


Figure 3. DJF spatial variability map of the leading two modes of variability for TROP (Local Precipitation) using PC analysis.

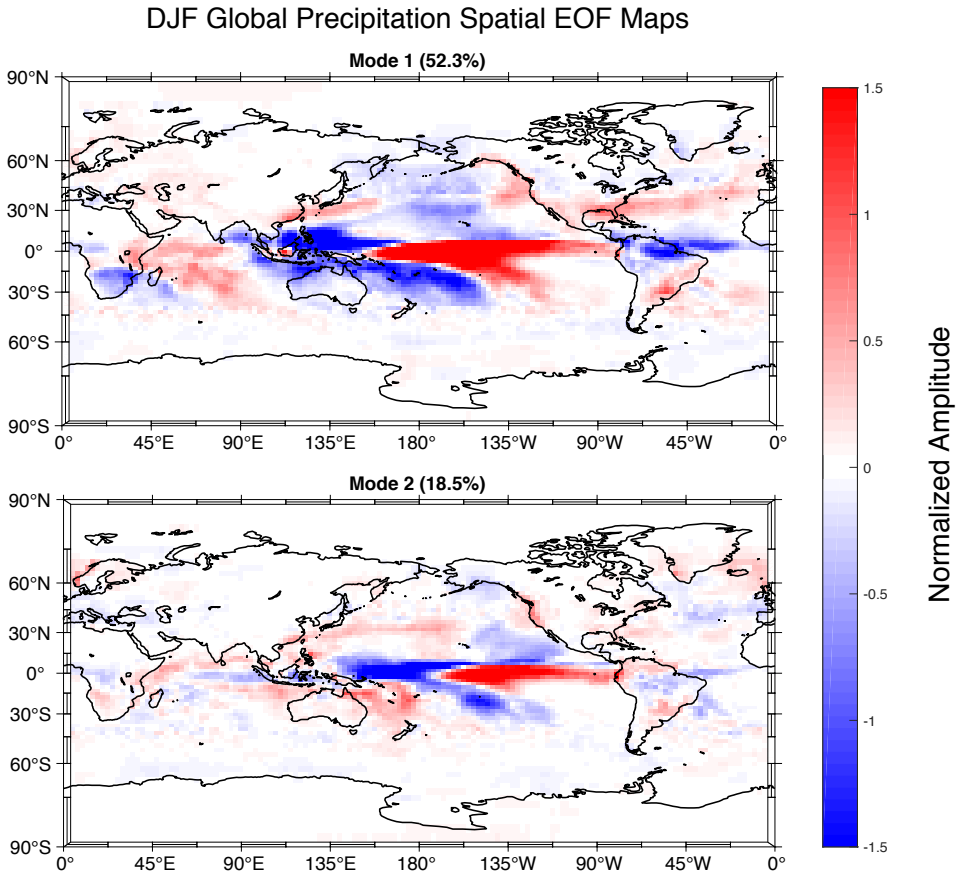


Figure 4. DJF spatial variability map of the leading two modes of variability for GP using PC analysis.

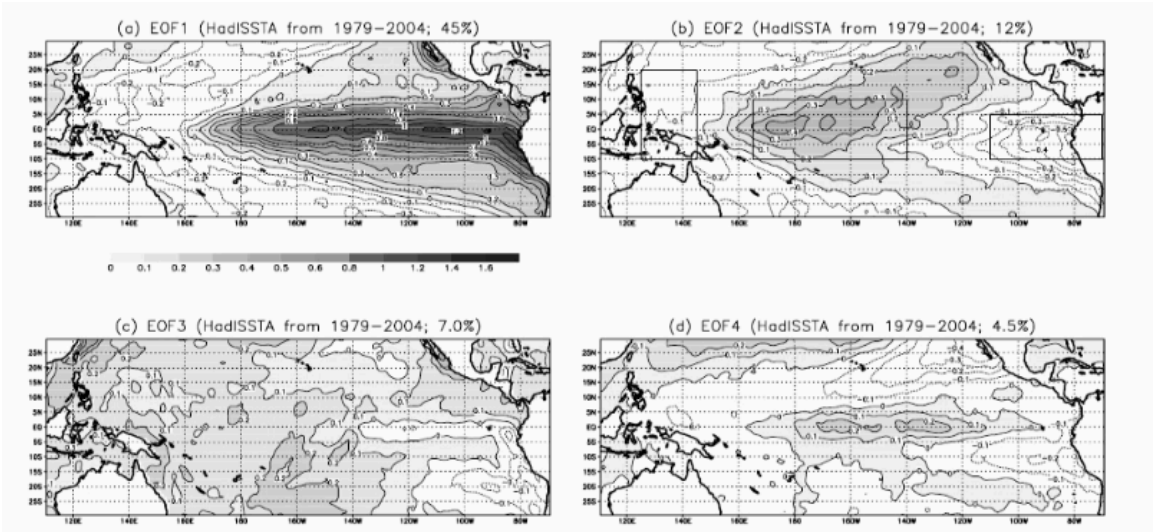


Figure 5. Reprinted from Ashok et al. 2007. "Top four EOF modes of tropical Pacific SSTA (1979-2004) multiplied by respective standard deviations of the principal components; units in °C."

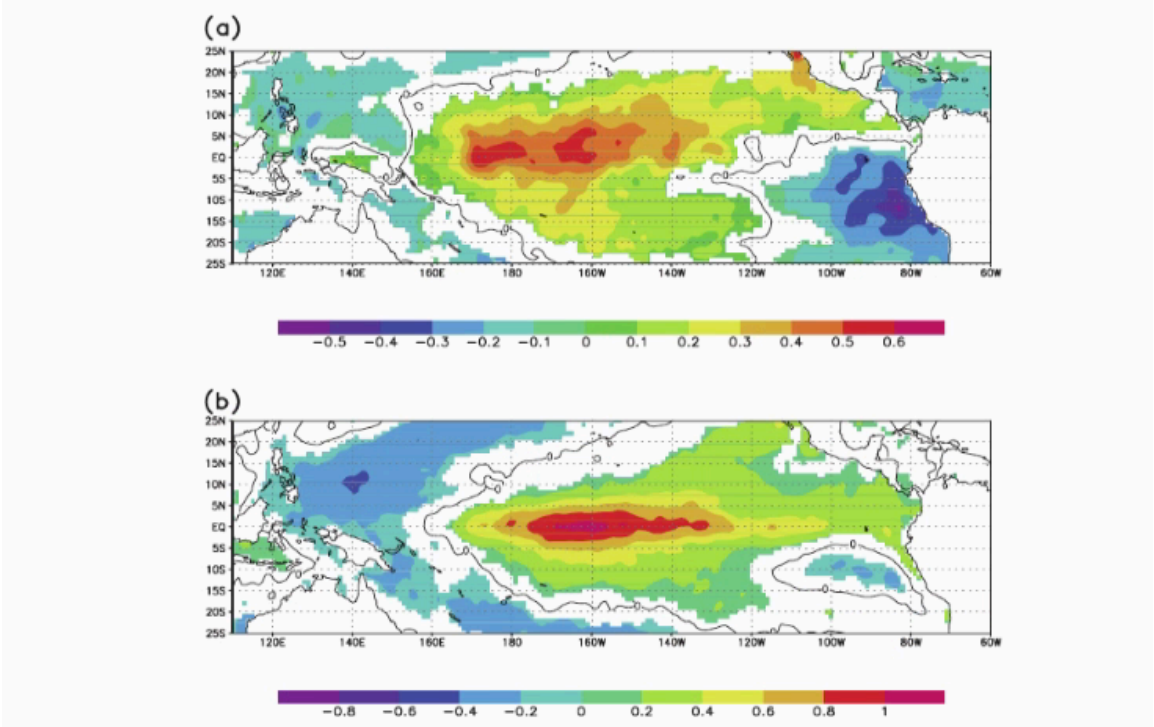


Figure 6. Reprinted from Ashok et al. 2007. "Composite SSTA in °C during strong positive El Niño Modoki events averaged over (a) seven boreal summers, namely, JJAS seasons of 1986, 1990, 1991, 1992, 1994, 2002, and 2004, and (b) 8 boreal winters, namely, DJF seasons of 1979-1980, 1986-1987, 1990-1991, 1991-1992, 1992-1993, 1994-1995, 2002-2003, and 2004-2005. Significant values above 95% confidence level from a two-tailed Student's t test are shaded."

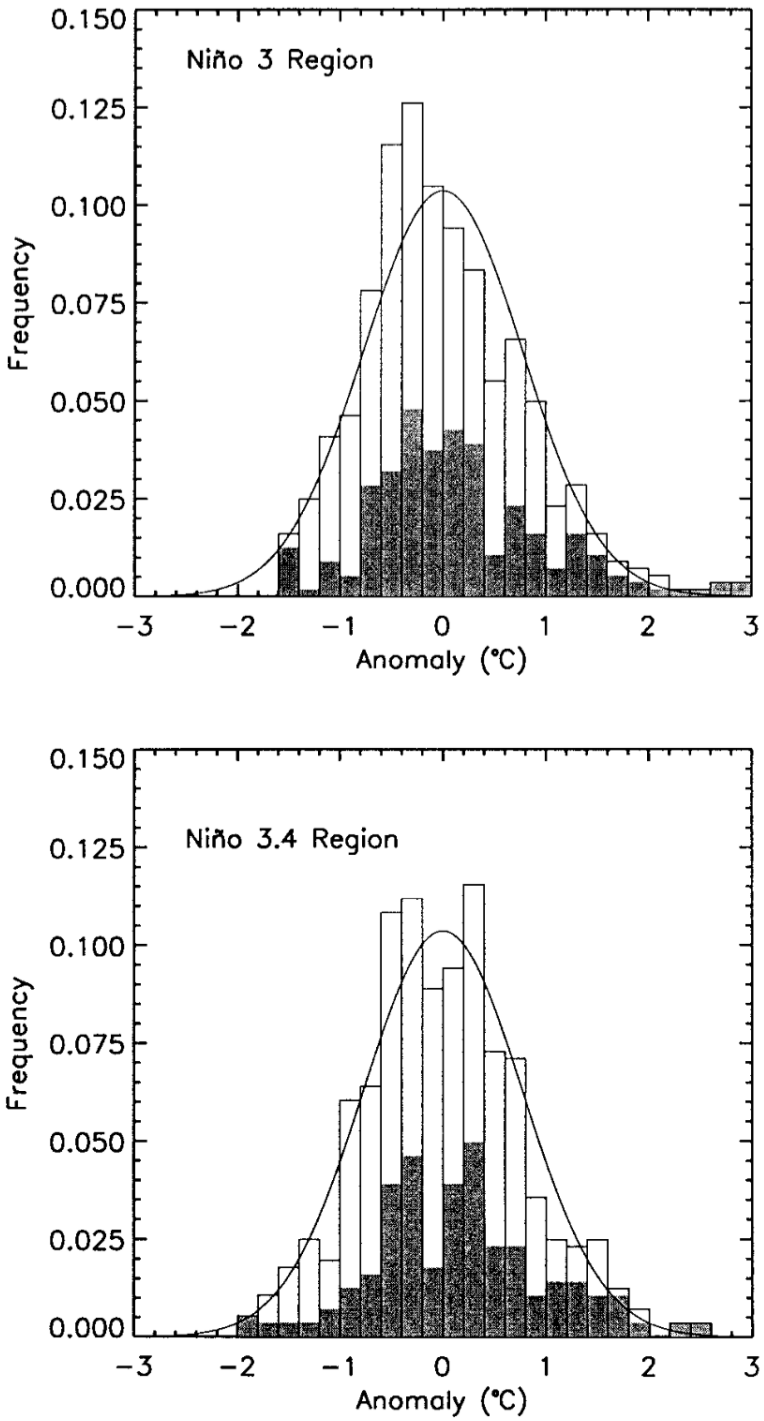


Figure 7. Reprinted from Trenberth 1997. "Histograms of the distribution of SST anomalies for Niño 3 and 3.4 from 1950 to March 1997 relative to the mean for the entire period. The contribution from the post-1979 period is shown by the stippled areas. Also given is the corresponding normal distribution with the same variance."

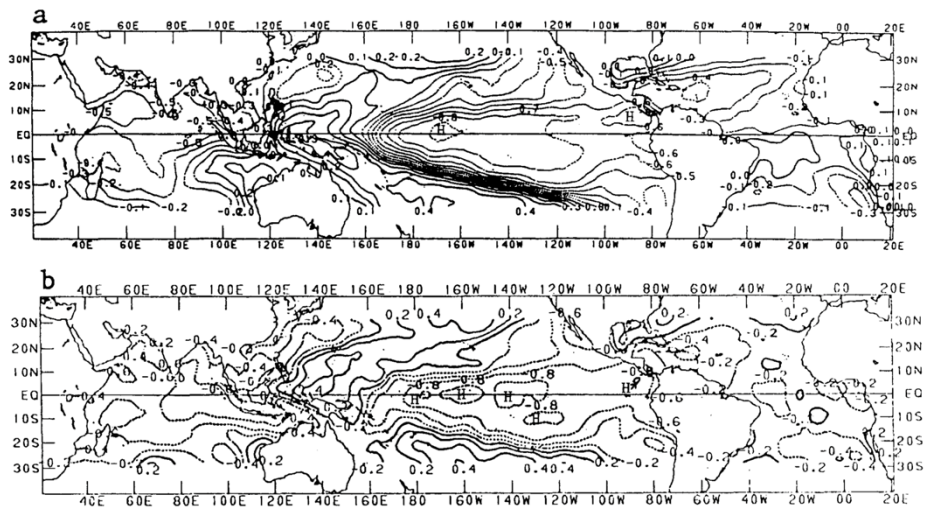


Figure 8. Reprinted from Barnston et al. 1997. "(a) The field of correlation between local SST and the SOI for all four regular 3-month seasons for 1950-1979. (b) As in (a), except for January-February periods (part of the northern winter-to-spring period when mature ENSO episodes often occur) for 1950-1979."

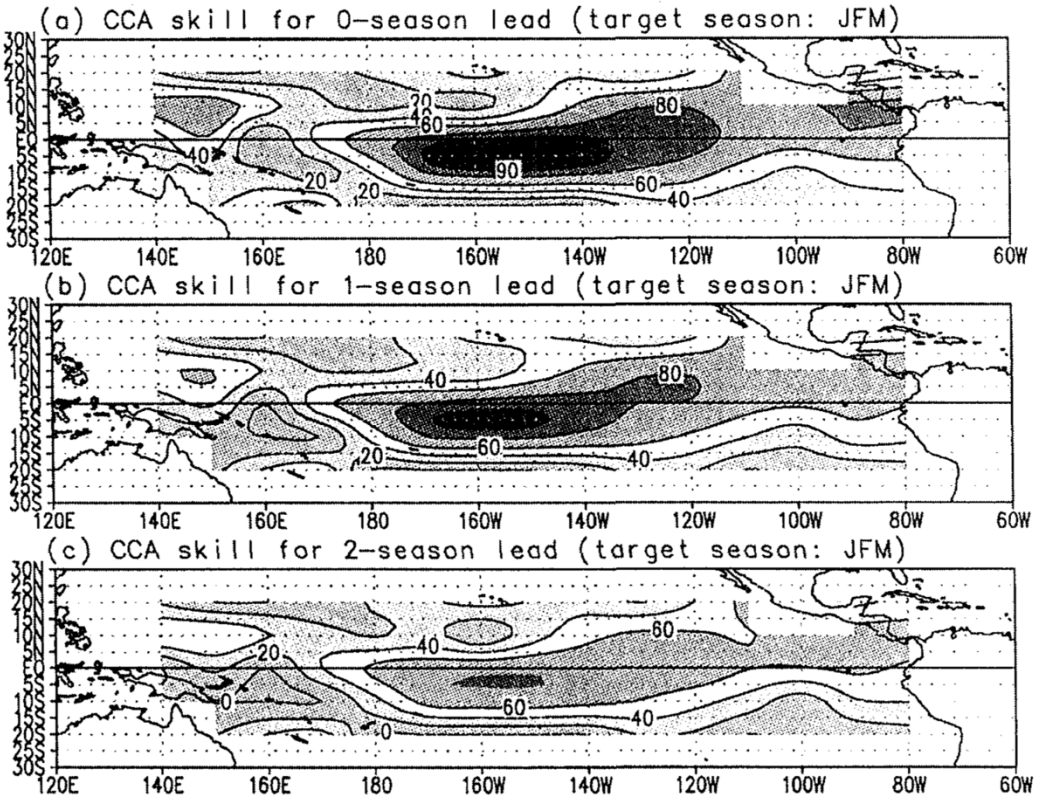


Figure 9. Reprinted from Barnston et al. 1997. "Correlation skill field for CCA predictions of Jan.-Feb.-Mar. tropical Pacific SST using global sea level pressure, tropical Pacific SST and subsurface sea temperatures as predictors. Forecast skill is shown for lead times of (a) 0, (b) 1 and (c) 2 seasons, where 0 season lead implies a 3-month target season that begins at the time the forecast is made (i.e., using data through the end of December for a Jan.-Feb.-Mar. forecast). Predictions for the 1957-1994 period are included."

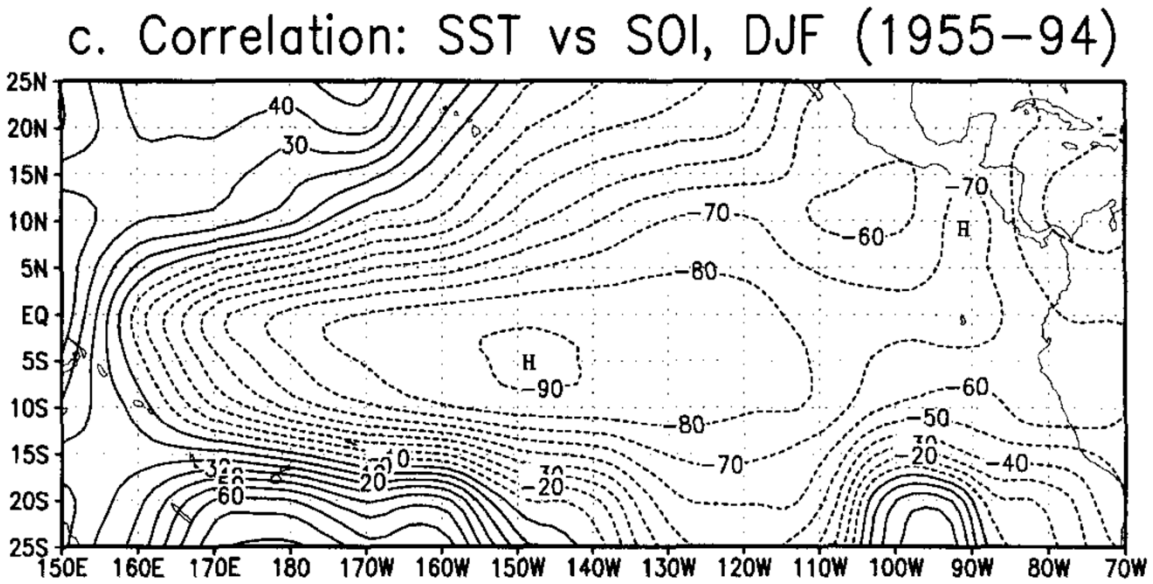


Figure 10. Reprinted from Barnston et al. 1997. Grid point correlations of DJF local SST's and DJF SOI for the time period 1955-1994.

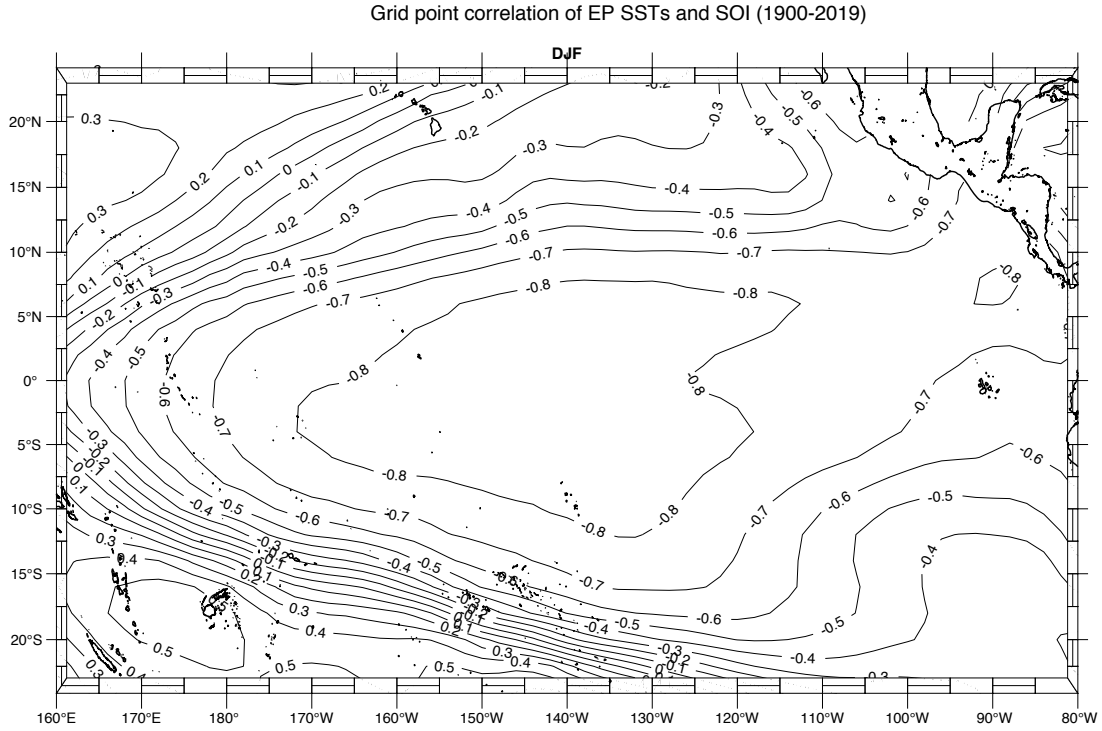


Figure 11. DJF grid point correlations of ERSSTv5 SST and SOI data that spans from 1900-2019.

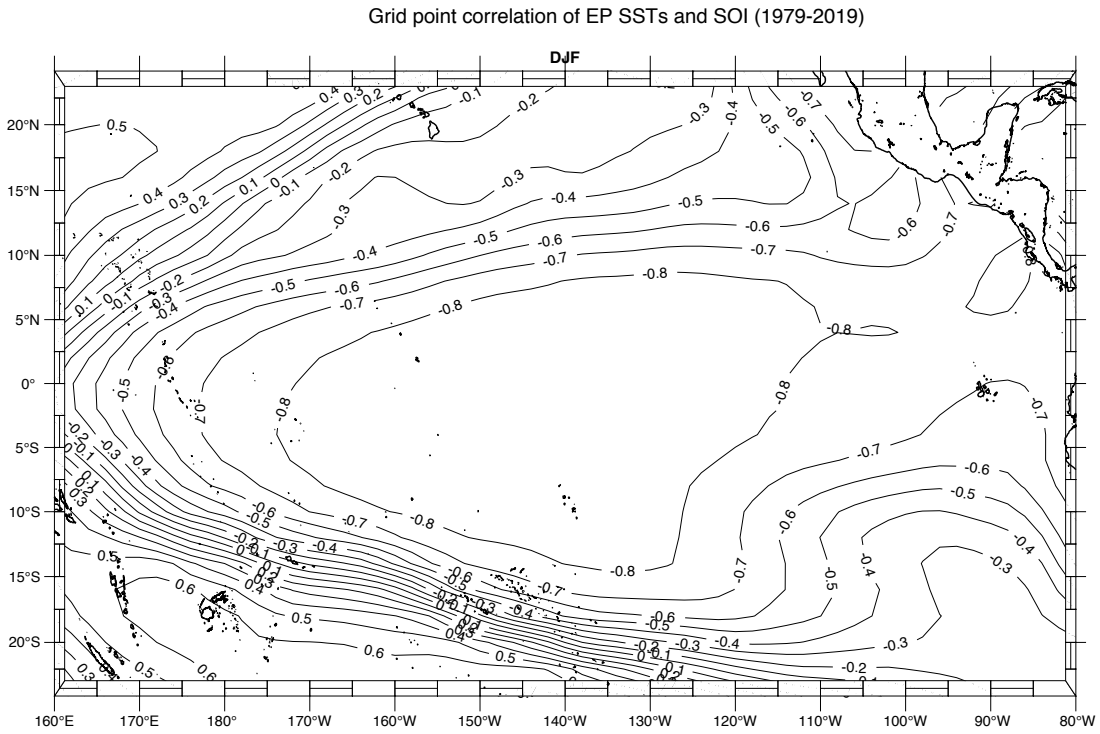


Figure 12. DJF grid point correlations of ERSSTv5 SST and SOI data that spans from 1979-2019, the length of the satellite era.

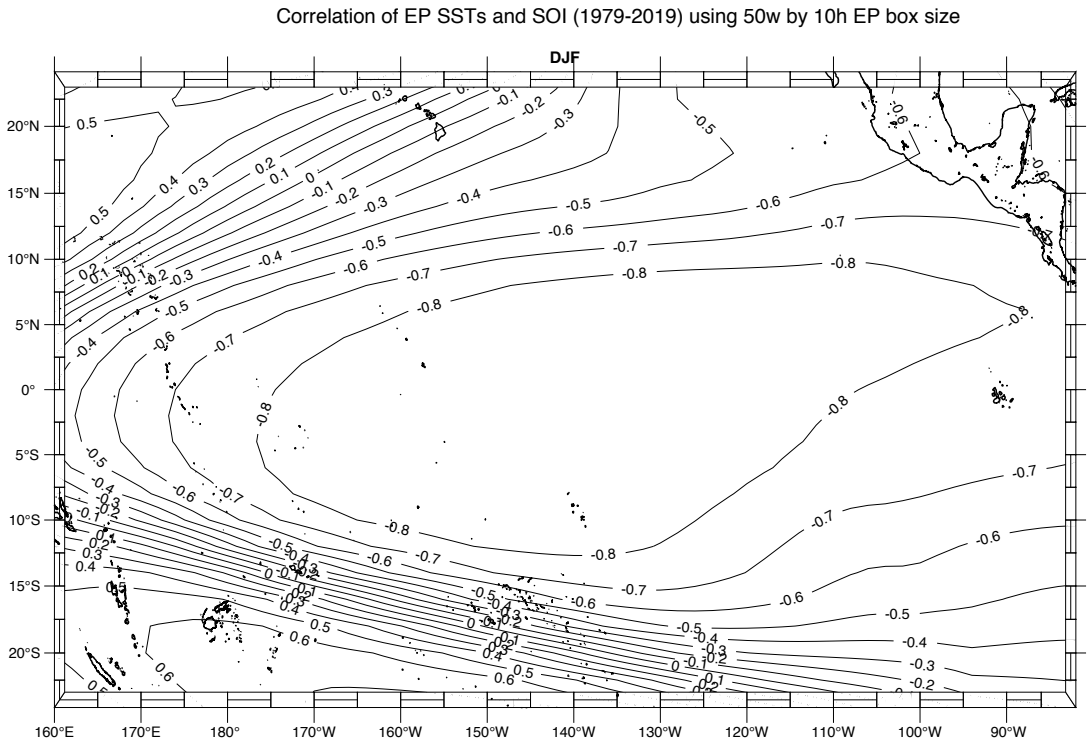
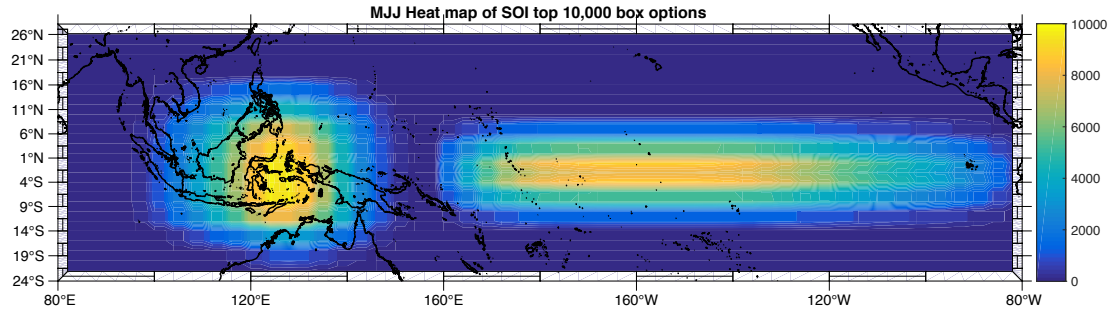
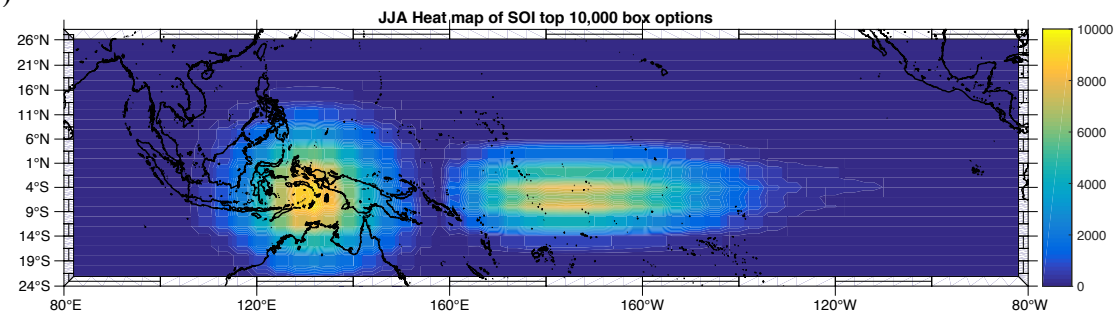


Figure 13. DJF correlations of ERSSTv5 SST and SOI data using a 50w by 10h SST box in the EP region for the time period of 1979-2019.

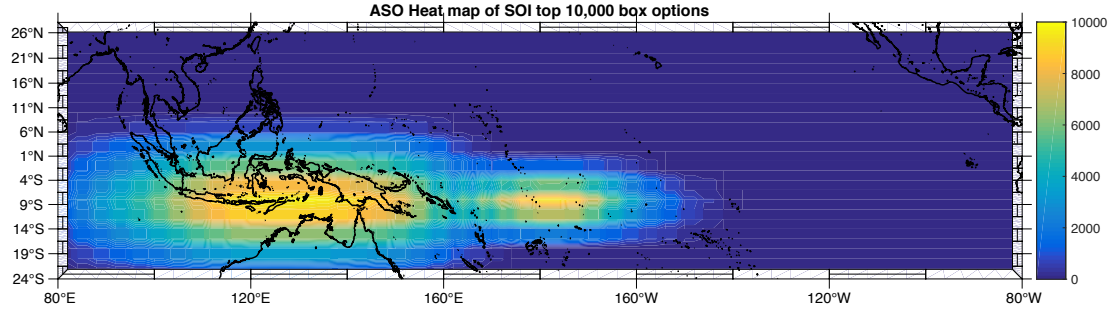
a)



b)



c)



d)

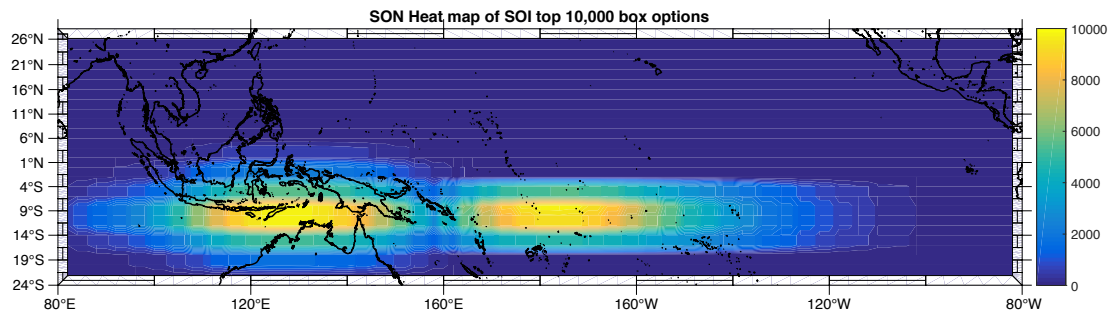
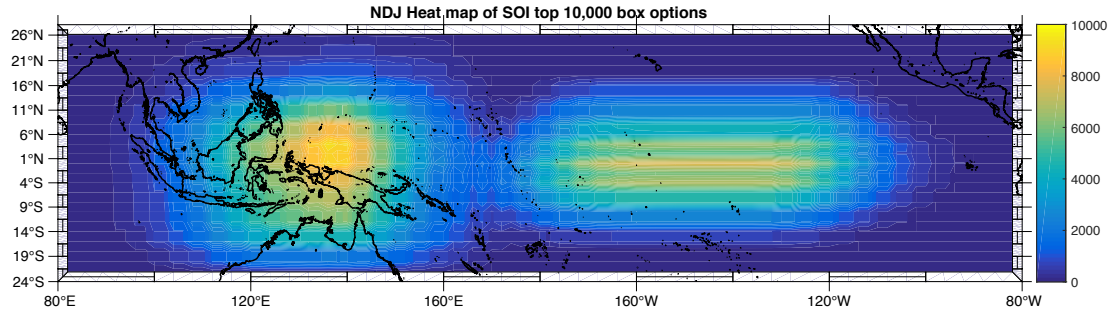
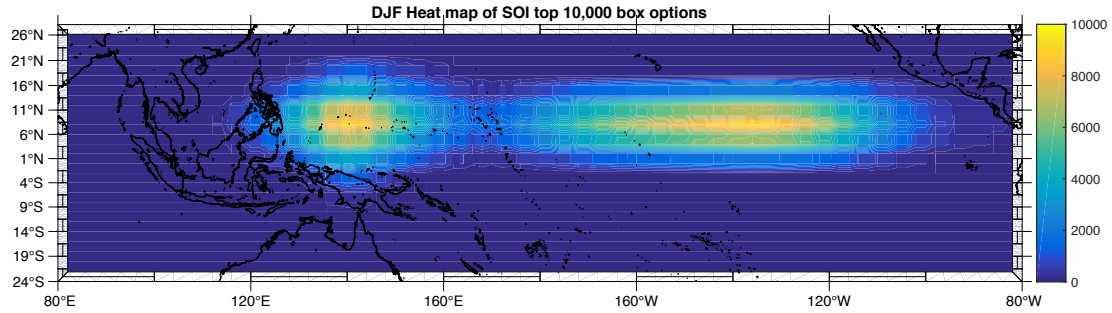


Figure 14. Heat maps of preferred boxes for EP and MC regions for the NDI in MJJ-MAM. Color shading shows the relative frequency that certain latitude-longitude grid points were included in the 10,000 NDI definitions having the strongest negative correlations to the SOI. Panels a and b: One-month and zero-month lead correlations with JJA SOI. Panels c and d: Correlations with SON SOI. Panels e and f: Correlations with DJF SOI. Panels g and h: Correlations with MAM SOI.

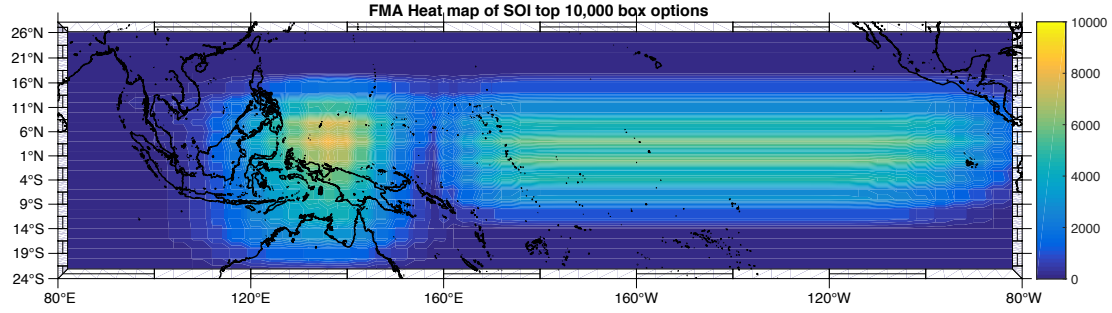
e)



f)



g)



h)

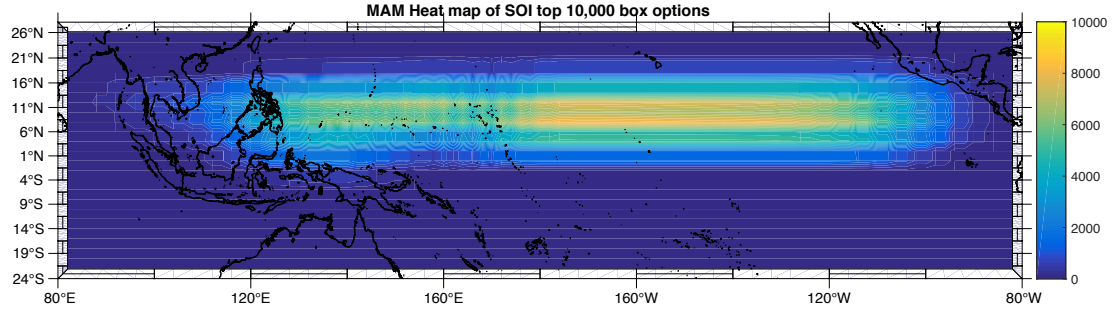
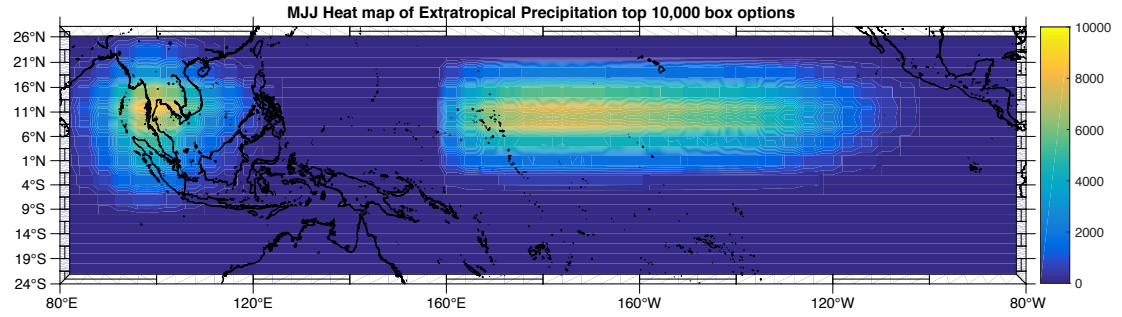
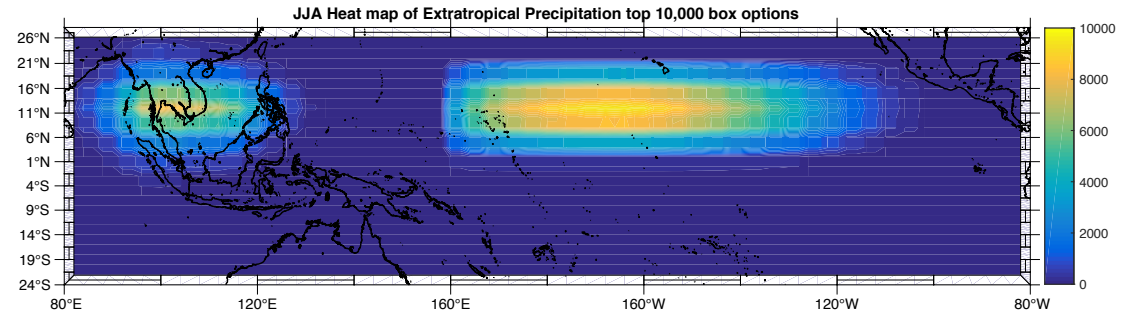


Figure 14. (continued)

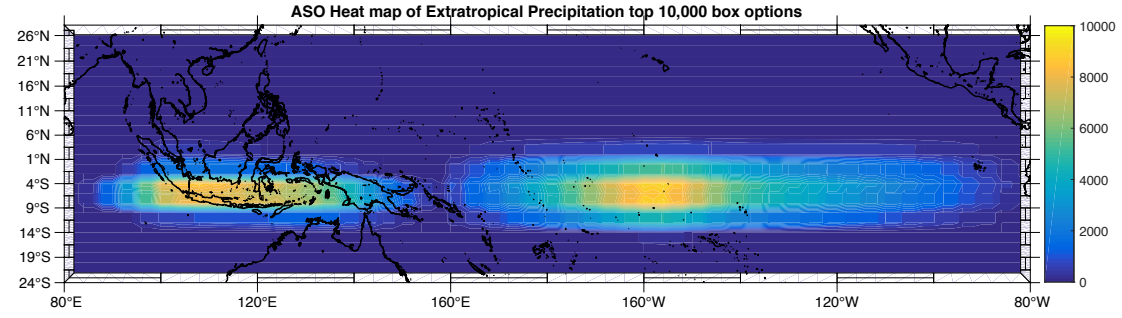
a)



b)



c)



d)

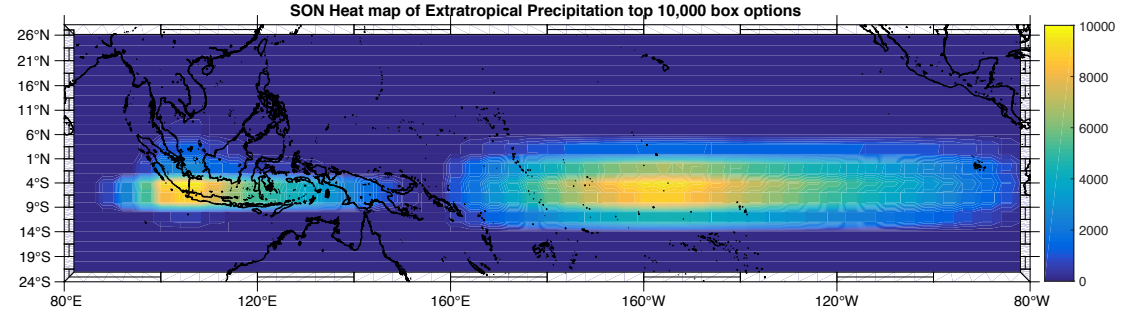
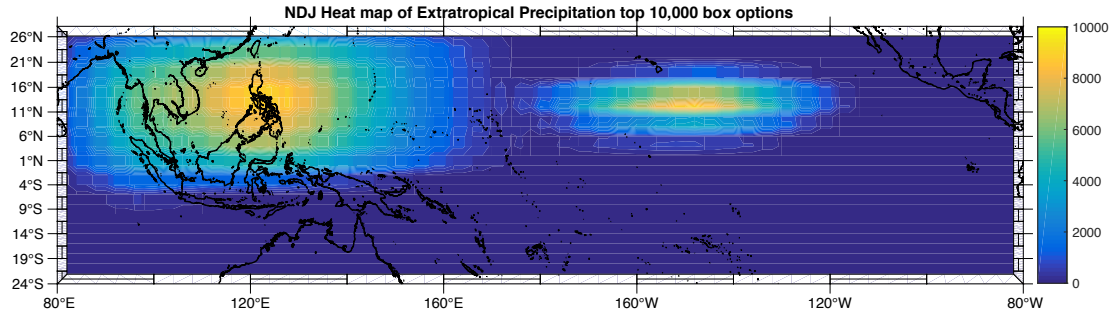
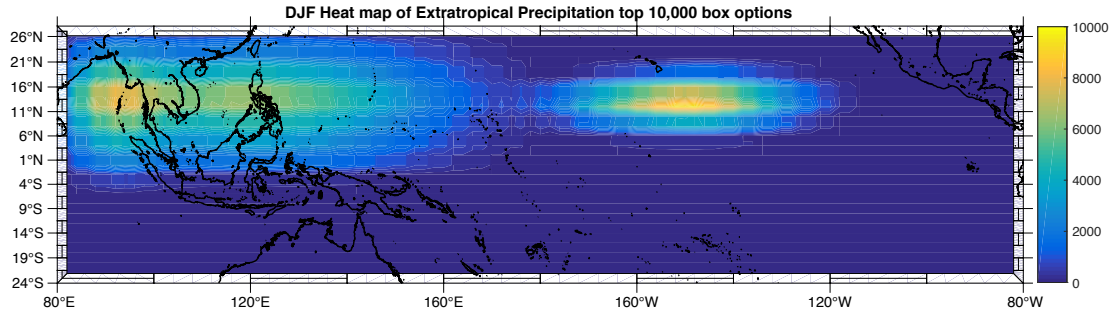


Figure 15. Same as in Figure 14 except for using EXP PC 1.

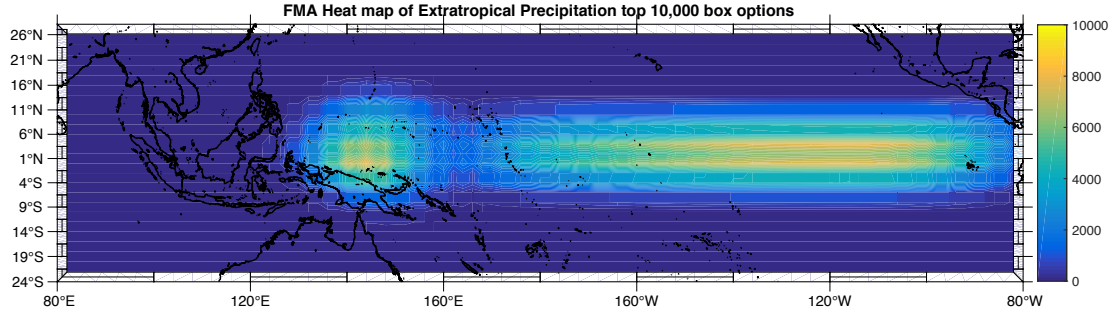
e)



f)



g)



h)

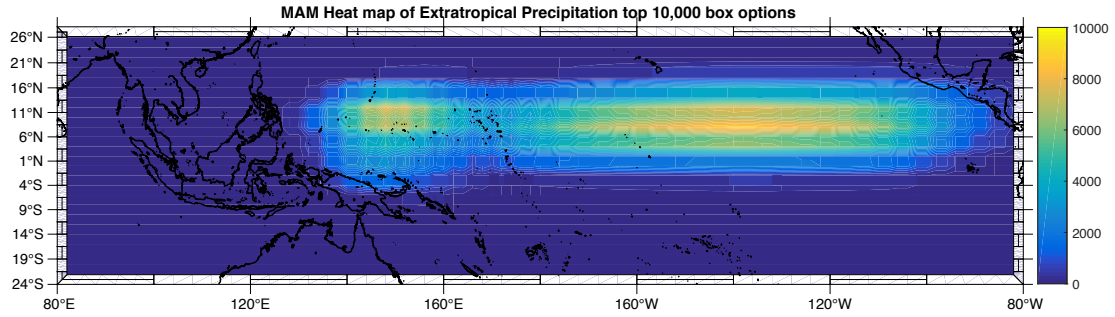


Figure 15. (continued)

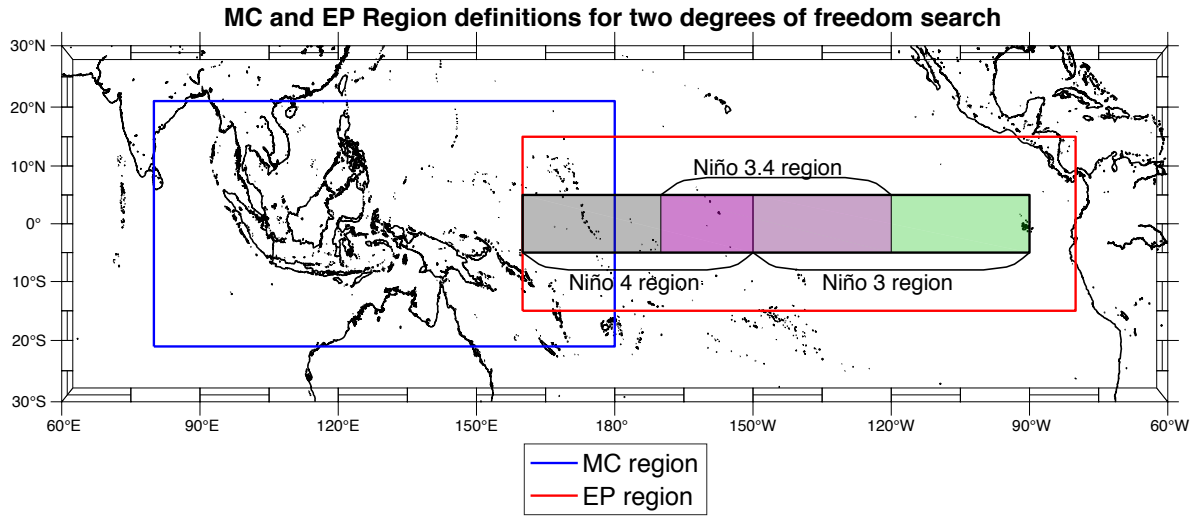


Figure 16. MC and EP region definitions used for the two degrees of freedom search. Only boxes that fully fit within each defined region are included in the search process. The Niño SST regions are included in this figure for reference.

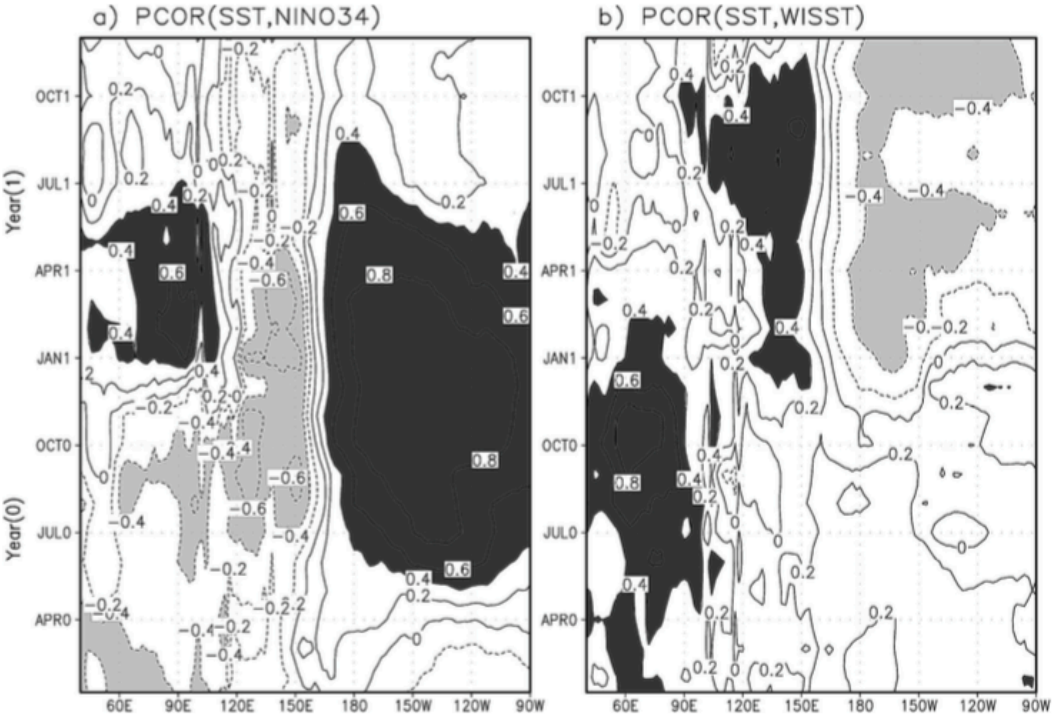


Figure 17. Reprinted from Kug and Kang 2006. "(a) The partial correlation of the equatorial SST (5°S-5°N) on the Niño-3.4 SST during November-January, after accounting for the effect of $WISST_{ON}$. (b) Same as in (a), but for the partial correlation of $WISST_{ON}$ after accounting for the effect of the Niño-3.4 SST during November-January. The y axis denotes the calendar months of the ENSO developing year (year 0) and decaying year (year 1)."

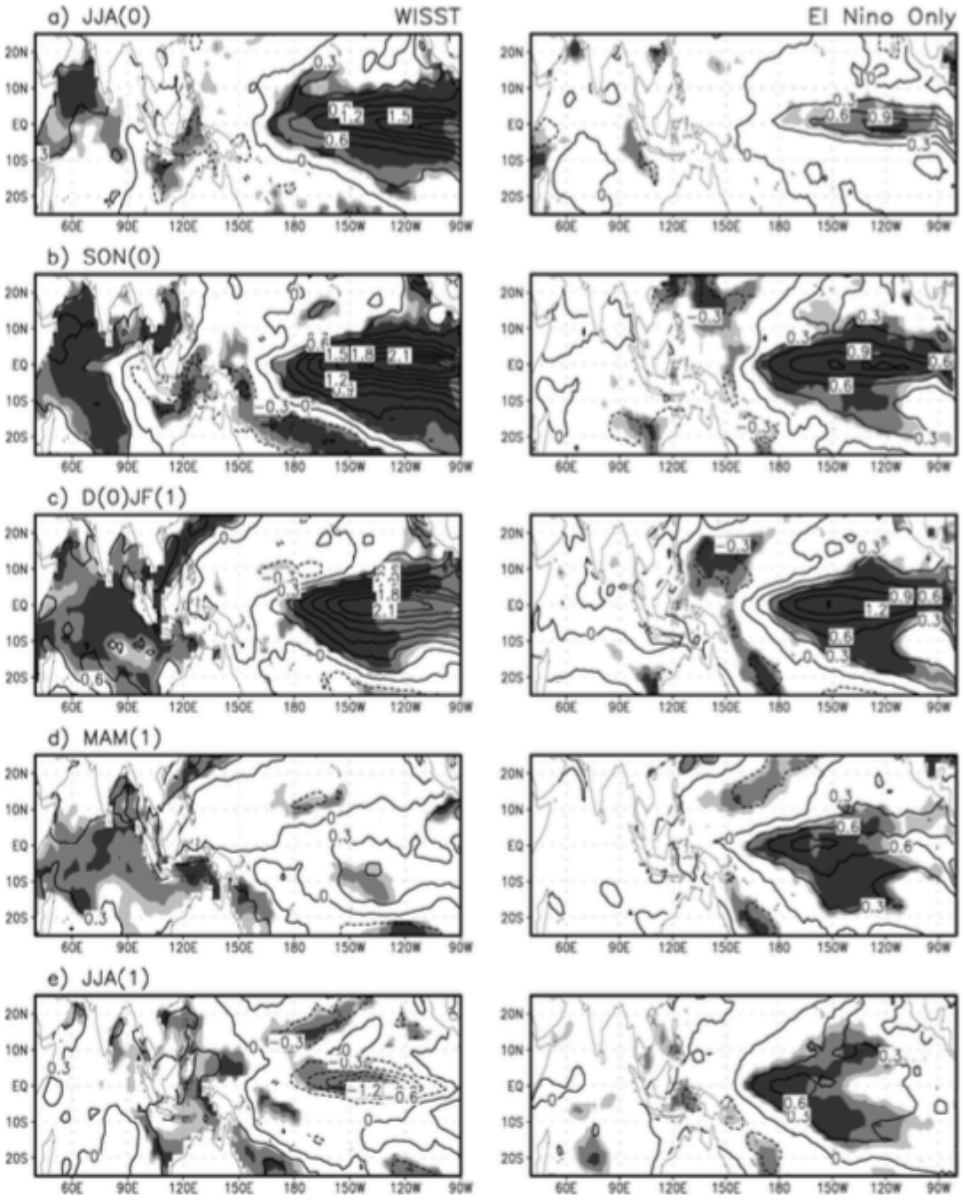


Figure 18. Reprinted from Kug and Kang 2006. "Composites of SST anomalies for (left) the WISST and (right) the El Niño-only cases. Light, medium, and dark shading represent the 90%, 95%, and 99% confidence levels, respectively. Year 0 and year 1 denote the year where an El Niño develops and the following year, respectively."

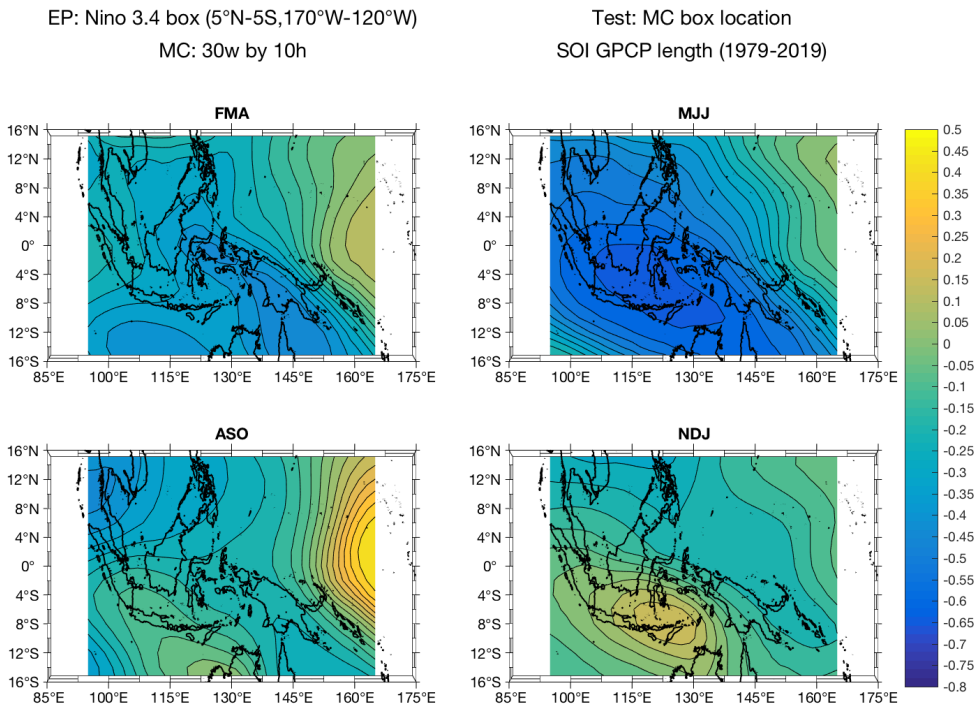
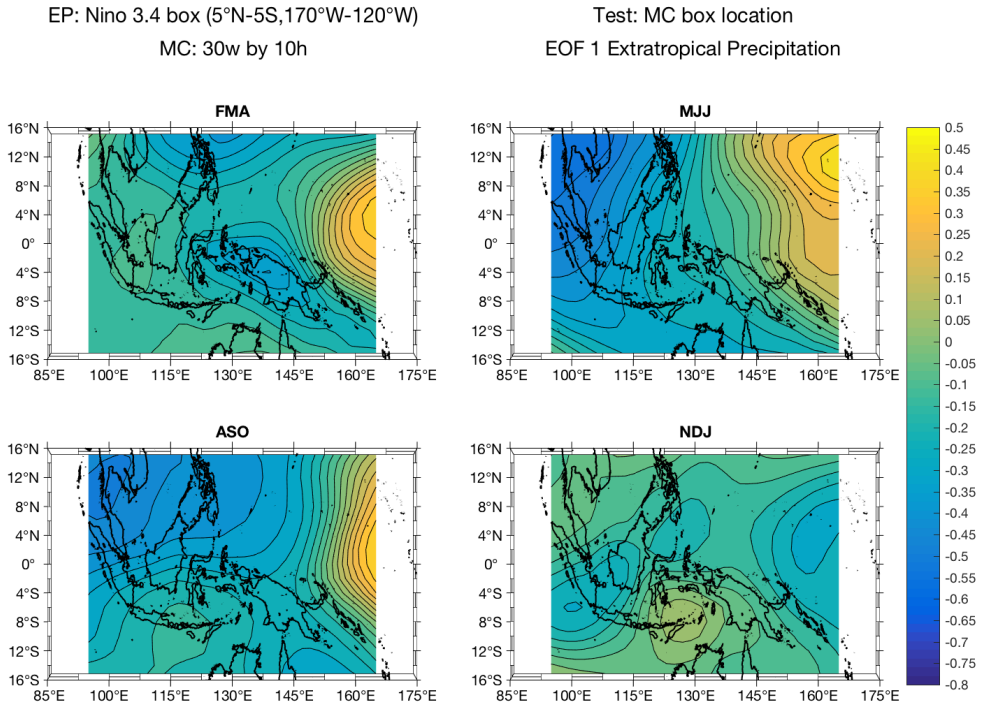


Figure 19. Linear regression of the Niño 3.4 index with SOI (1979-2019) for box options that are 30w by 10h in size. Correlations are between the MC box centers that fully fit within the defined MC region domain and the residuals for one-month lead seasons. Residuals represent what is not correlated with the Niño 3.4 index using SOI (1979-2019). The regions of negative correlation values indicate MC boxes with more information about the SOI than explained by just the Niño 3.4 index.

a)



b)

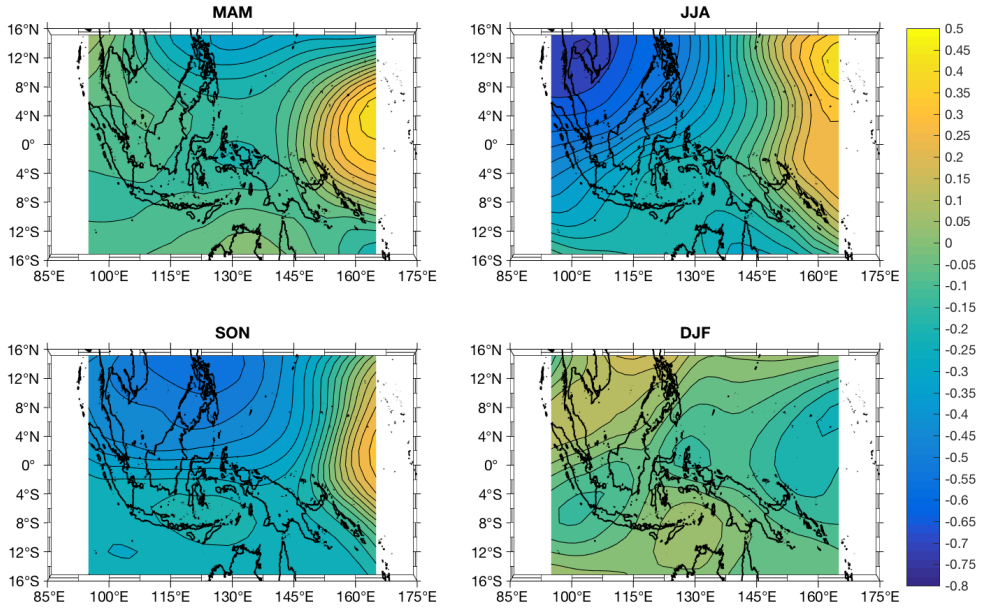
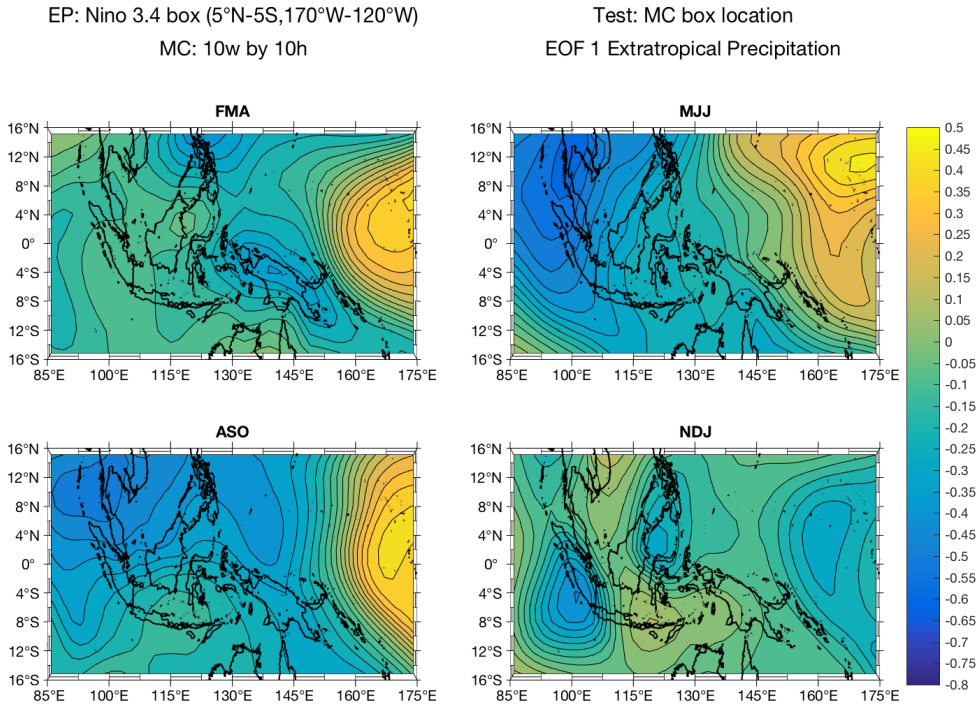


Figure 20. Linear regression of the Niño 3.4 index with EXP PC 1 for box options that are 30w by 10h in size. Correlations are between the MC box centers that fully fit within the defined MC region domain and the residuals for one-month lead seasons. Residuals represent what is not correlated with the Niño 3.4 index using EXP PC 1. The regions of negative correlation values indicate MC boxes with more information about EXP PC 1 than explained by just the Niño 3.4 index for a) one-month lead seasons and b) standard seasons.

a)



b)

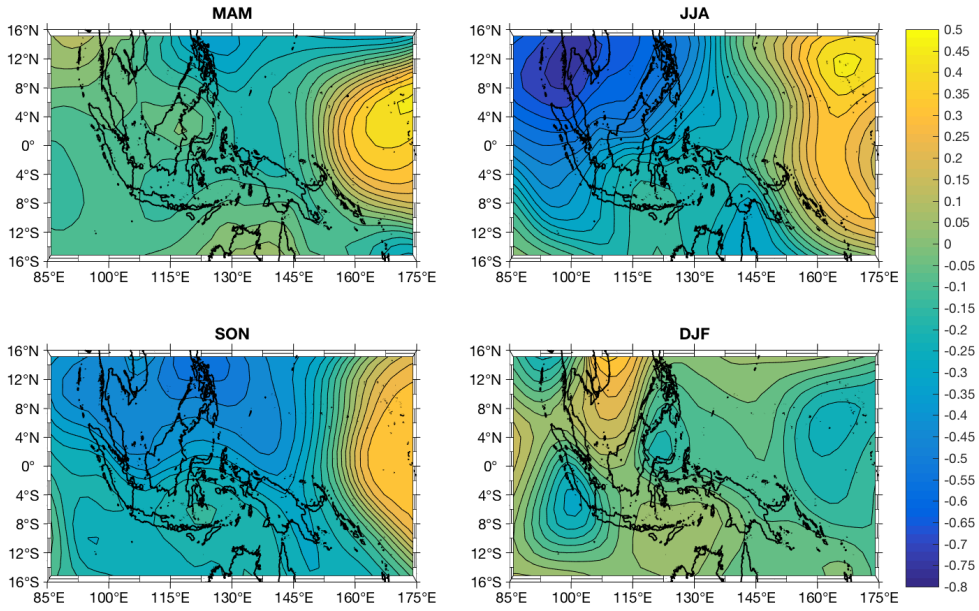
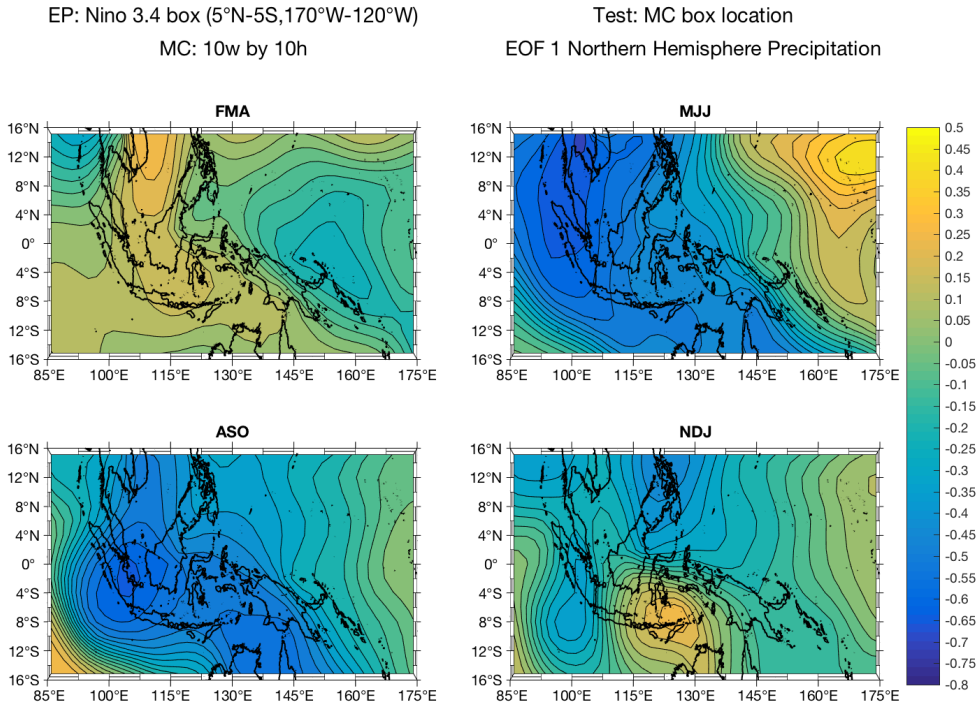


Figure 21. Same as in figure 20 except for box options that are 10w by 10h.

a)



b)

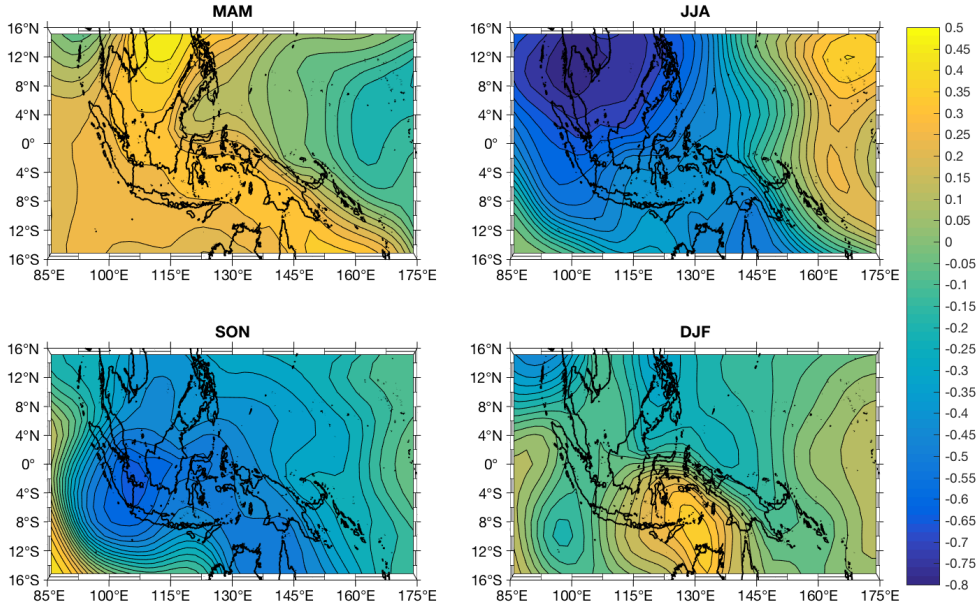
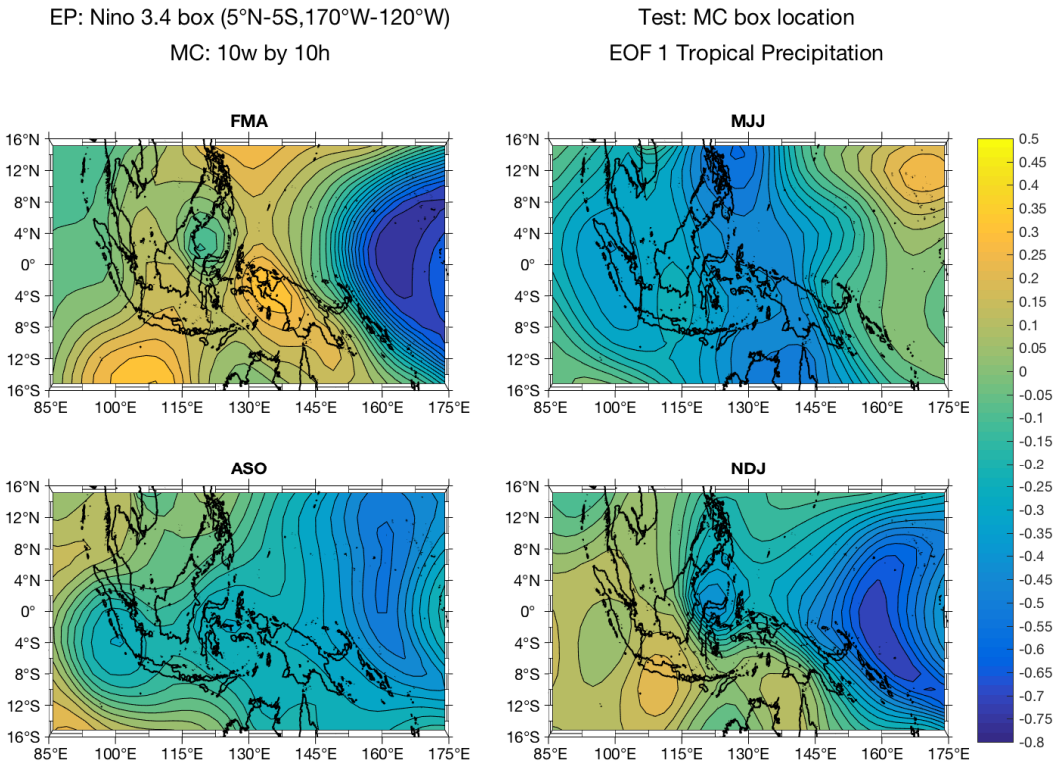


Figure 22. Same as in figure 20 except for using NHP PC 1 for box options that are 10w by 10h.

a)



b)

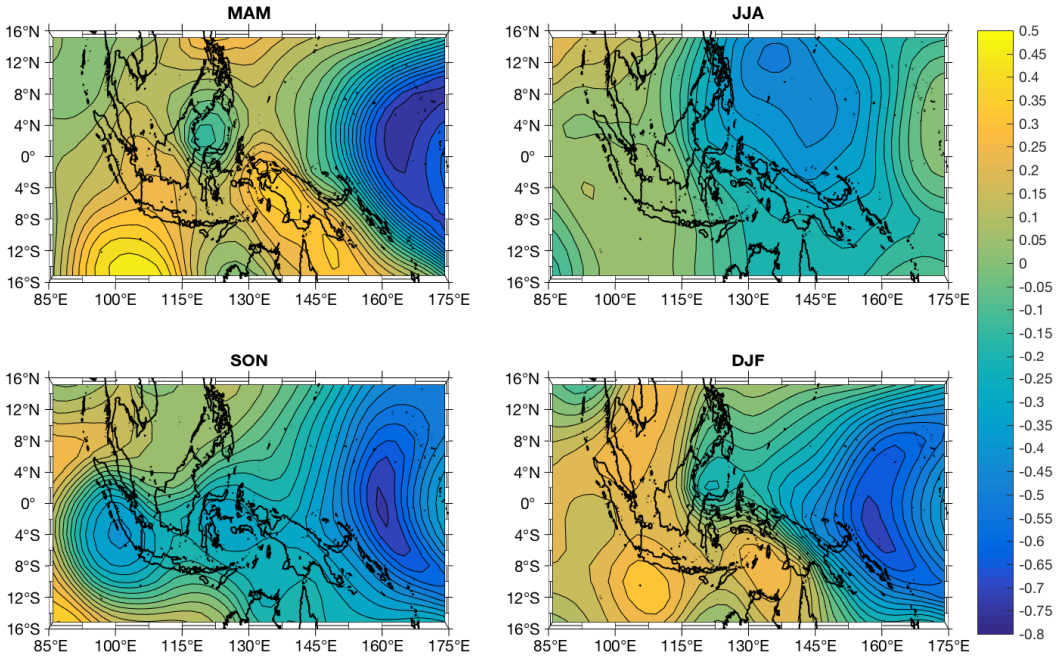
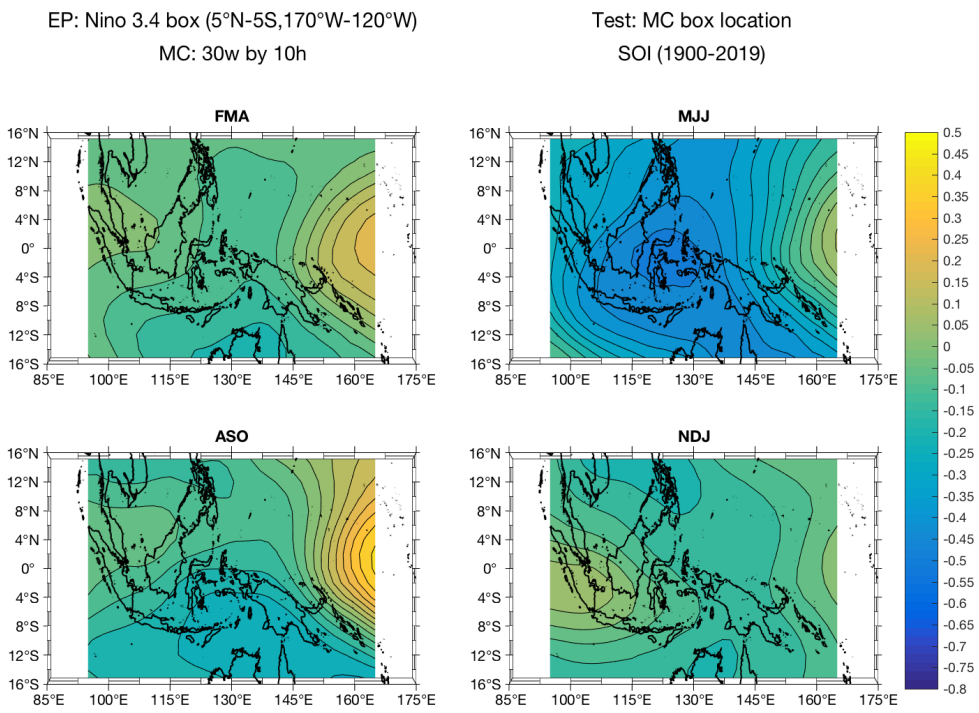


Figure 23. Same as in figure 20 except for using TROP PC 1 for box options that are 10w by 10h.

a)



b)

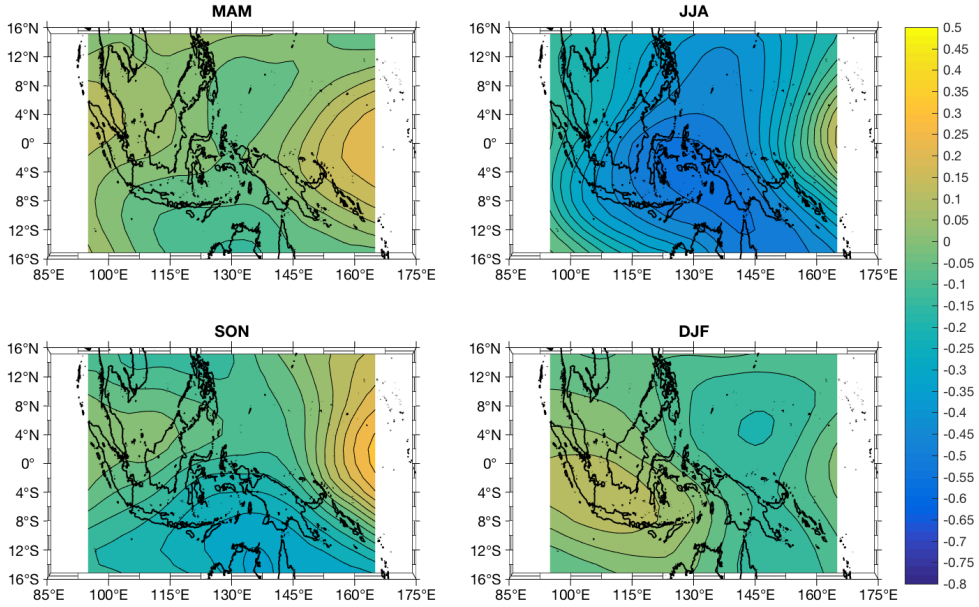


Figure 24. Same as in figure 20 except for using SOI (1900-2019) for box options that are 30w by 10h.

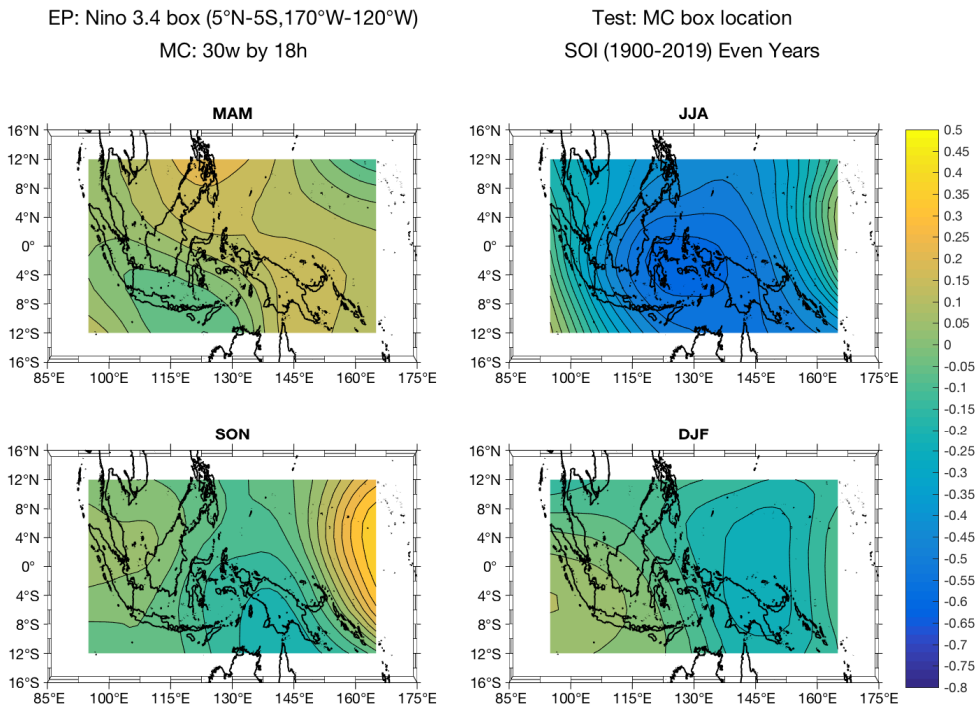


Figure 25. Same as in figure 20 except for using the even years of SOI (1900-2019) for box options that are 30w by 18h for zero month leads.

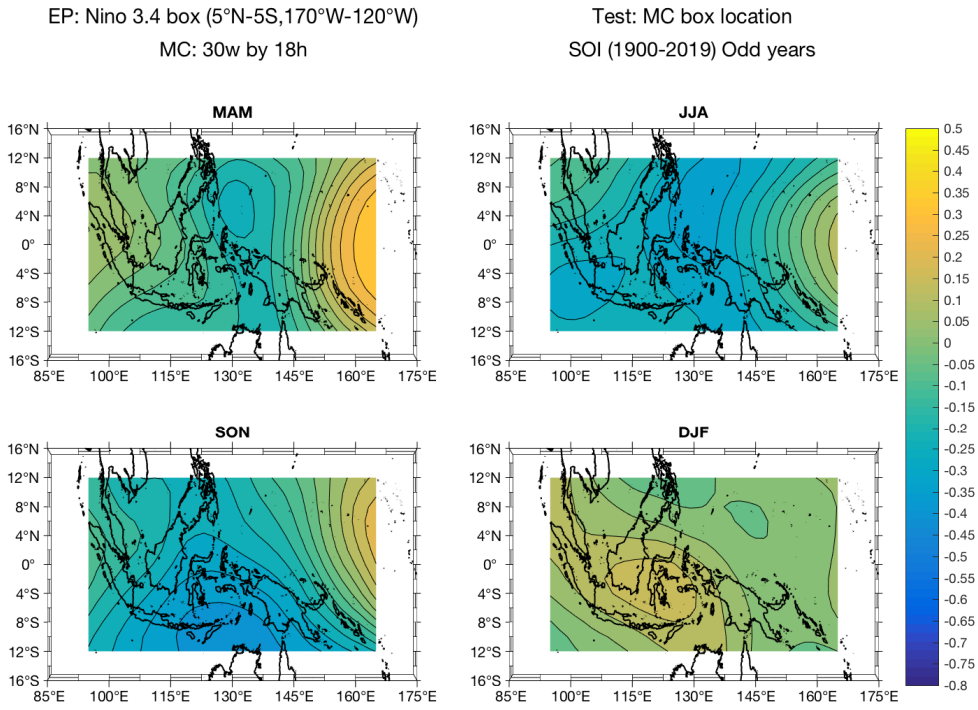


Figure 26. Same as in figure 20 except for using the odd years of SOI (1900-2019) for box options that are 30w by 18h for zero month leads.

a)

b)

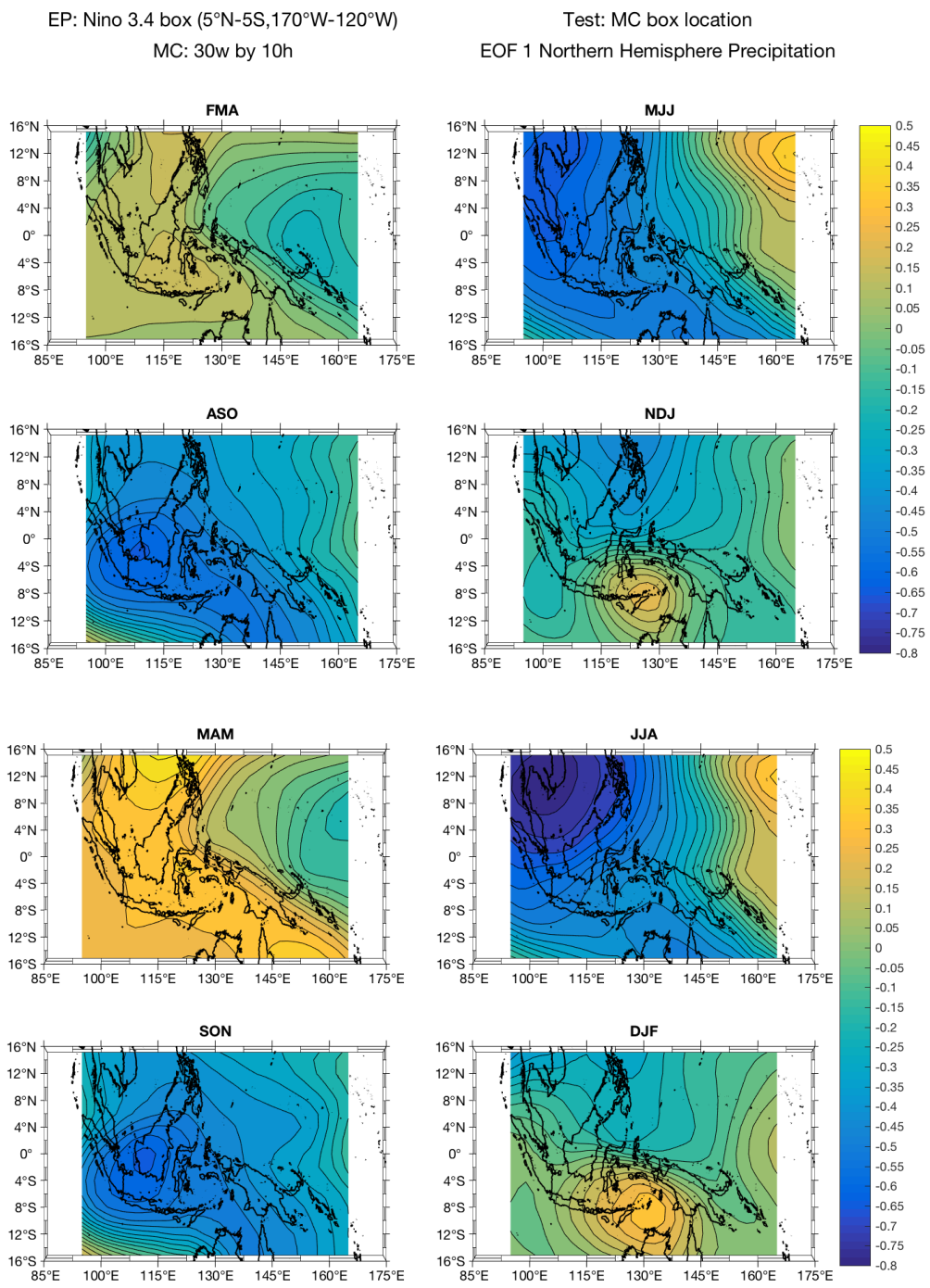
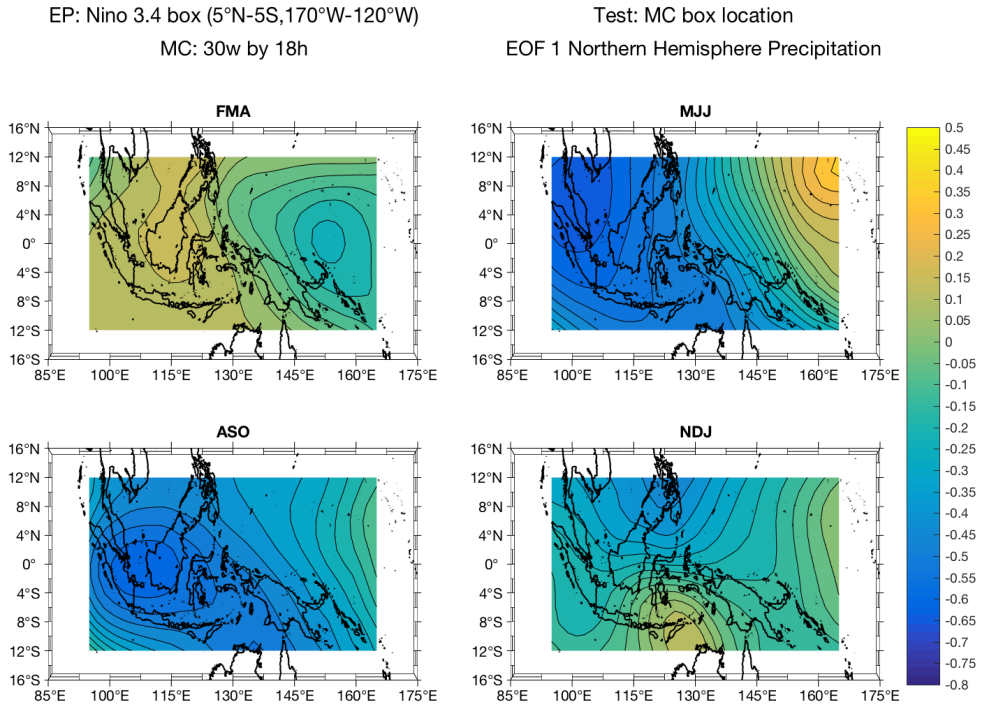


Figure 27. Same as in figure 20 except for using NHP PC 1 for box options that are 30w by 10h.

a)



b)

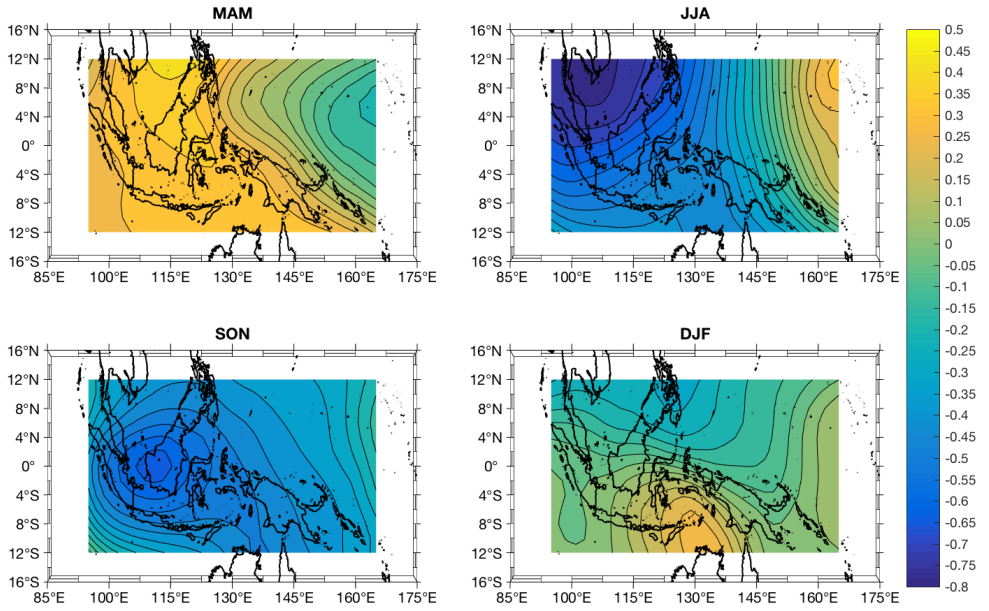
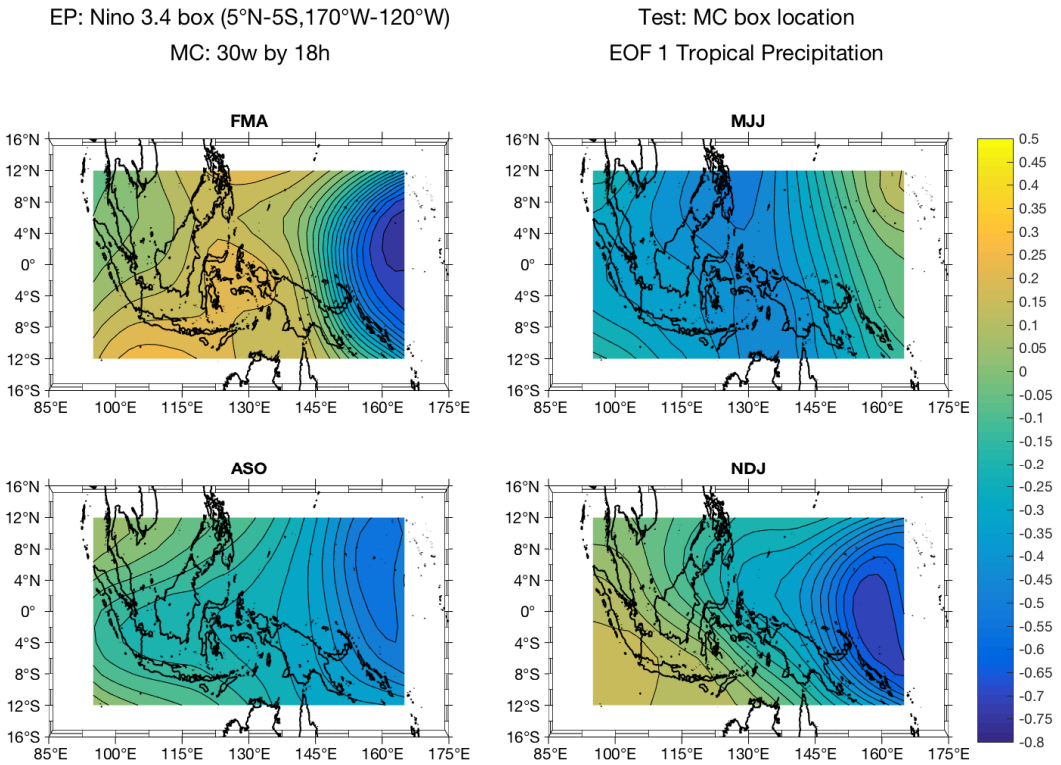


Figure 28. Same as in figure 20 except for using NHP PC 1 for box options that are 30w by 18h.

a)



b)

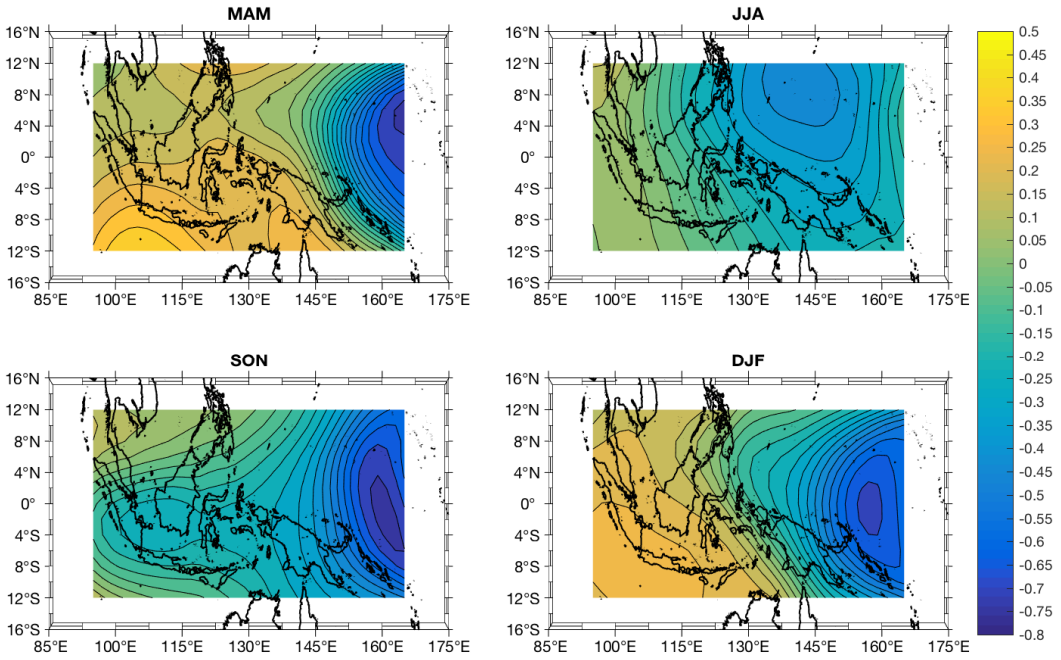
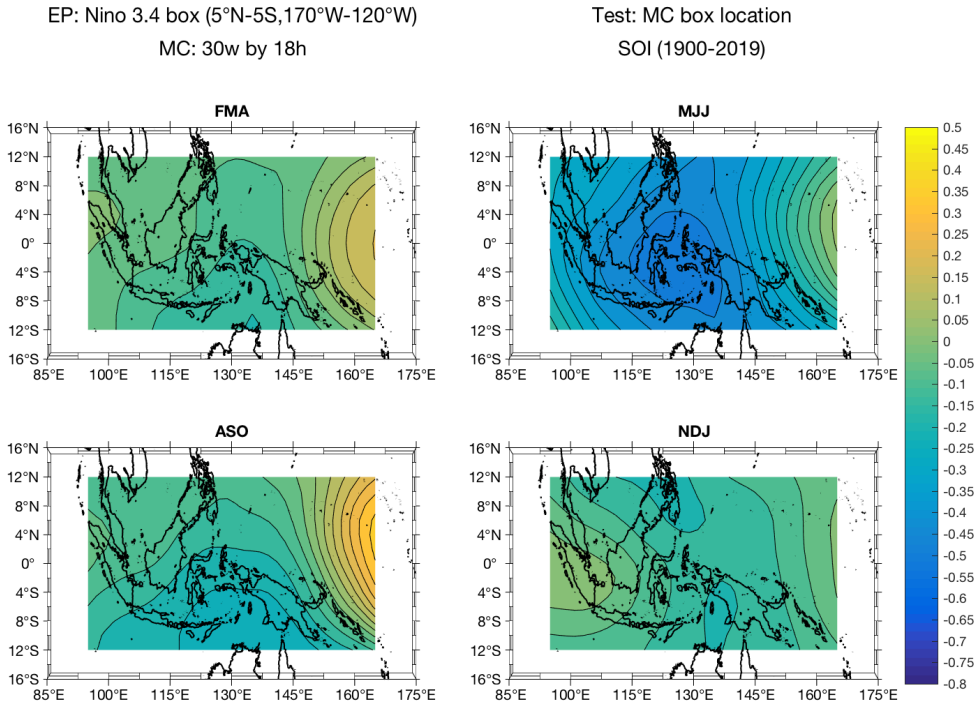


Figure 29. Same as in figure 20 except for using TROP PC 1 for box options that are 30w by 18h.

a)



b)

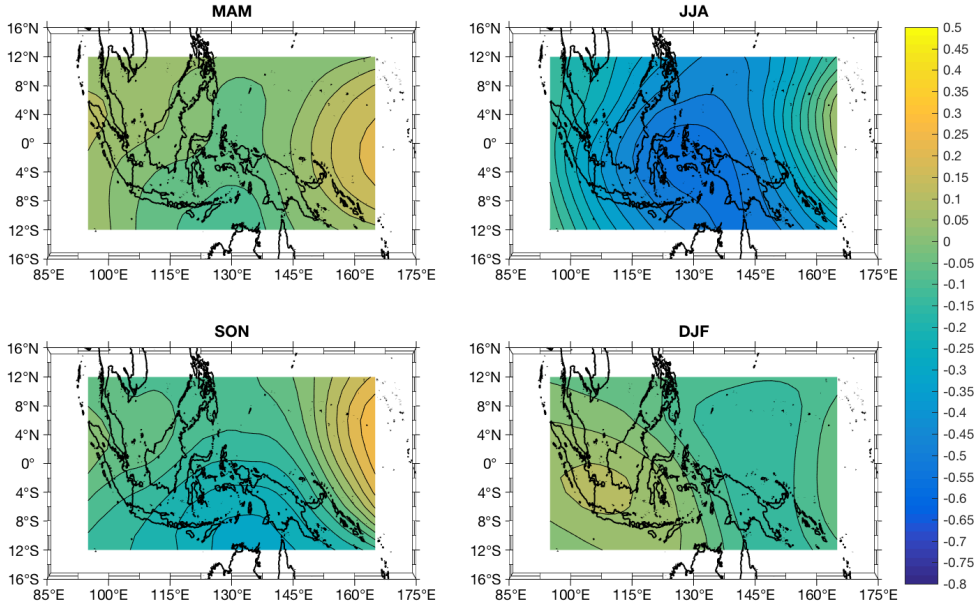
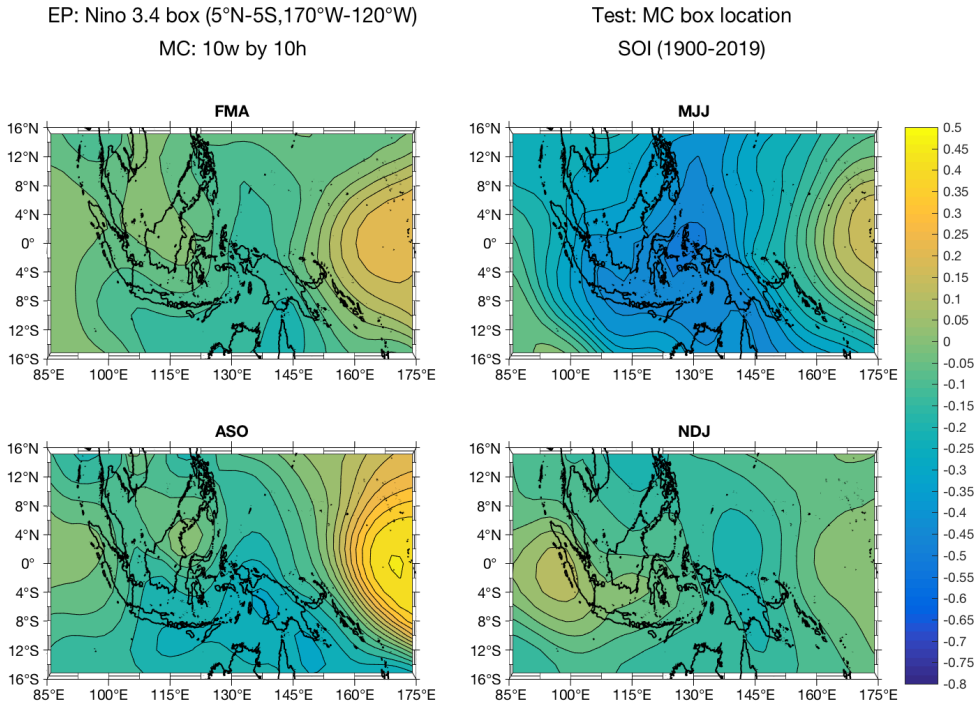


Figure 30. Same as in figure 20 except for using SOI (1900-2019) for box options that are 30w by 18h.

a)



b)

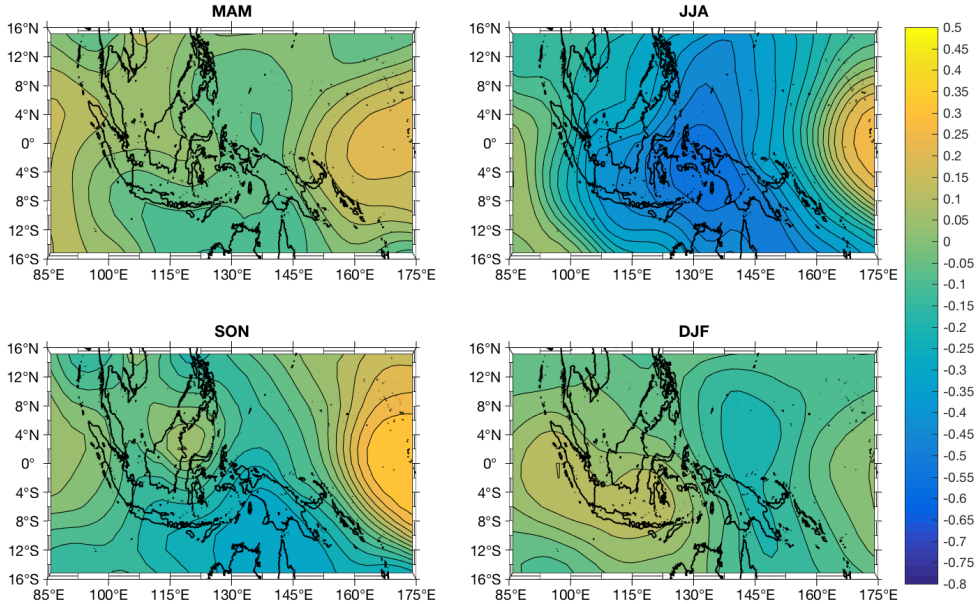


Figure 31. Same as in figure 20 except for using SOI (1900-2019) for box options that are 10w by 10h.

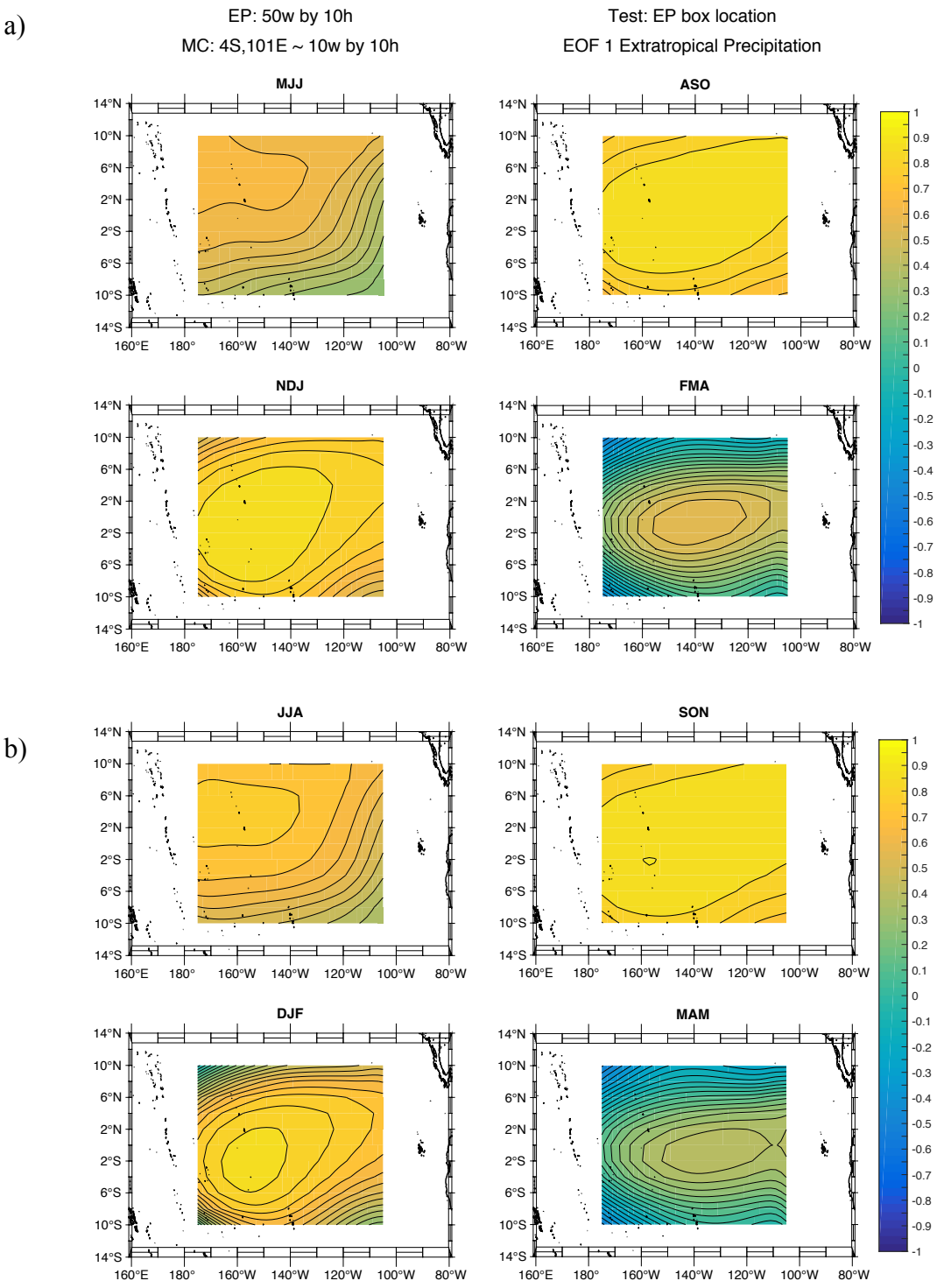
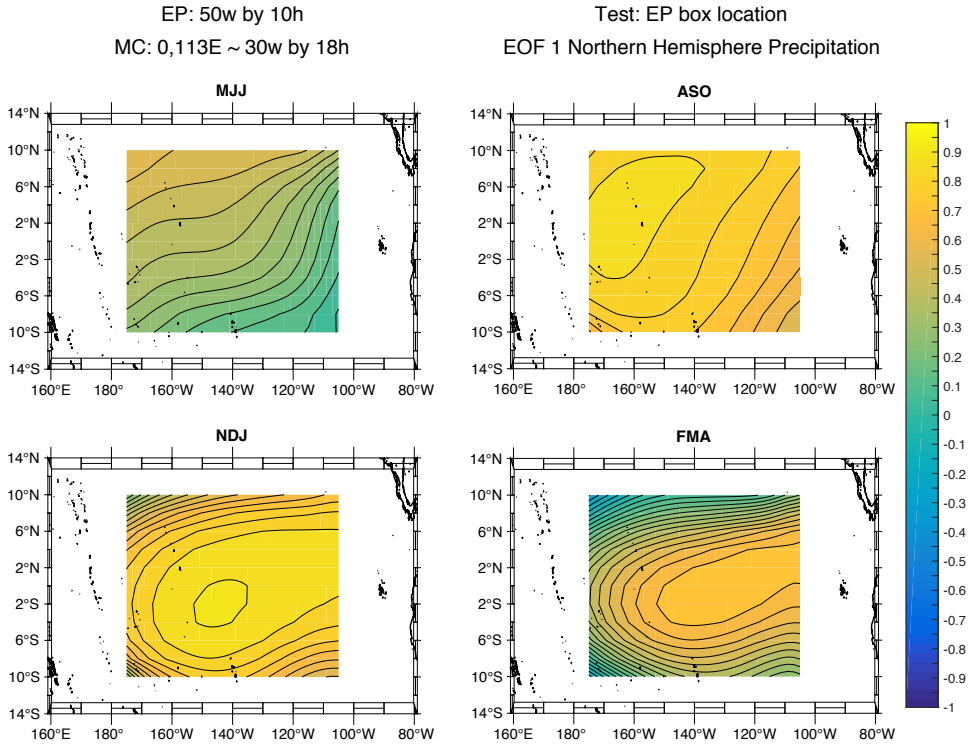


Figure 32. EP box location sensitivity testing using EXP PC 1 performed by using MC box center 4°S, 101°E that is 10w by 10h to hold the MC region fixed. Search across 2 degrees of freedom by holding EP box size constant at 50w by 10h for a) one-month lead seasons and b) standard seasons. Correlation values are between the NDI with each EP box that fully fits within the defined EP region domain and EXP PC 1. The high positive correlations represent the best region to place an EP box.

a)



b)

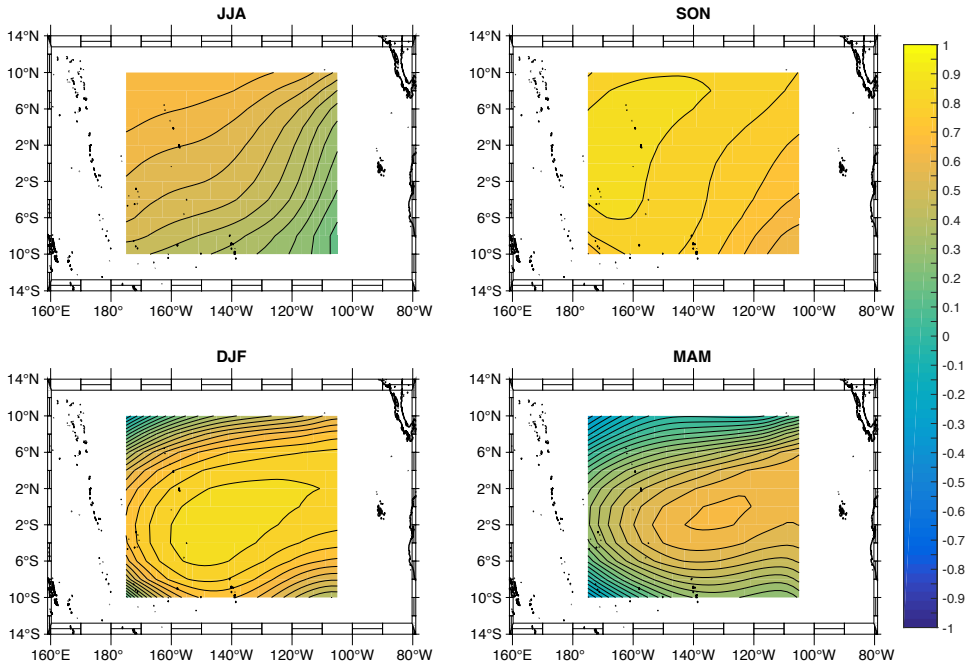


Figure 33. Same as figure 32 except for using NHP PC 1 for MC box center 0°,113°E that is 30w by 18h to hold the MC region fixed.

a)

b)

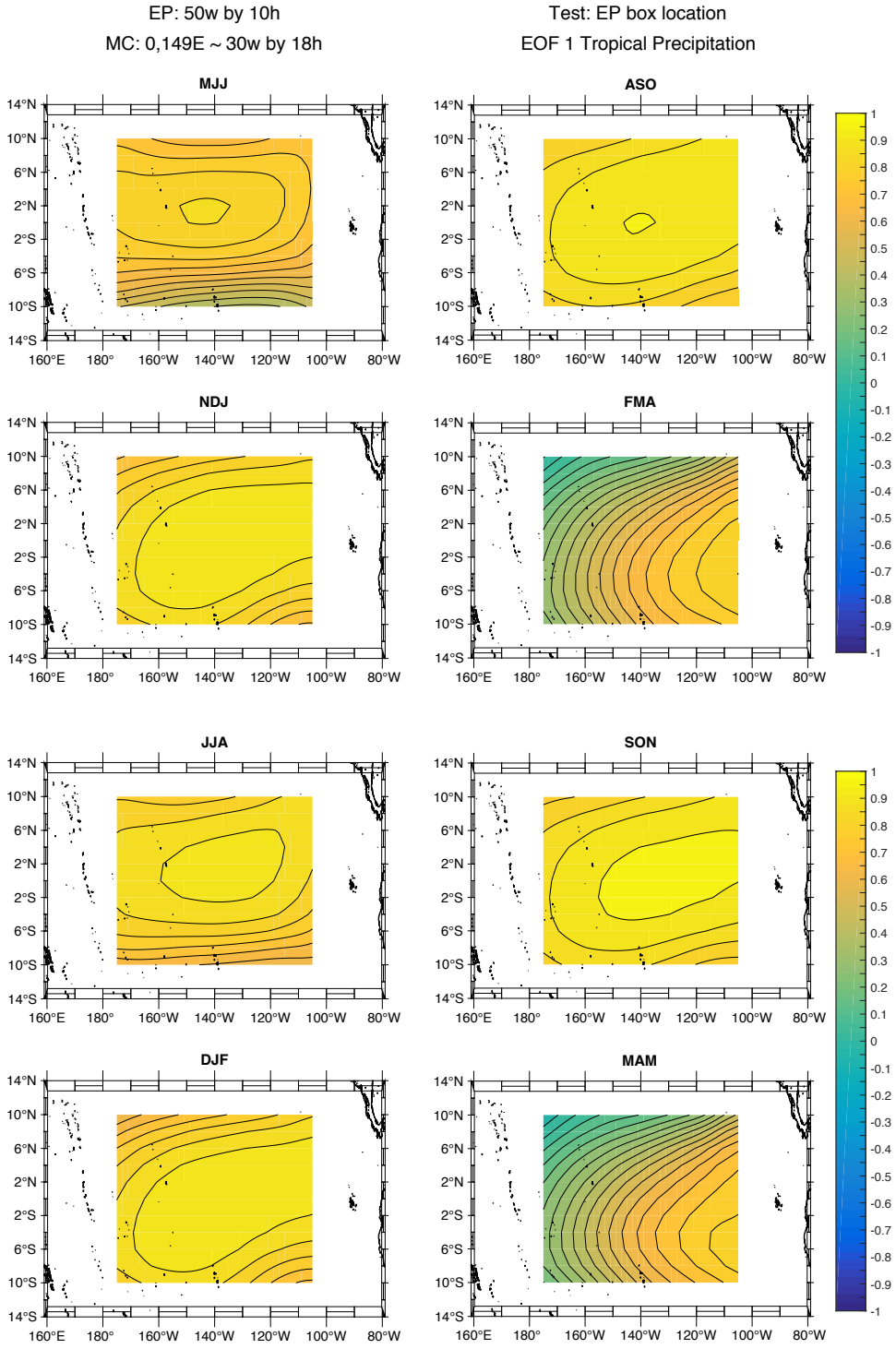


Figure 34. Same as figure 32 except for using TROP PC 1 for MC box center 0°,149°E that is 30w by 18h to hold the MC region fixed.

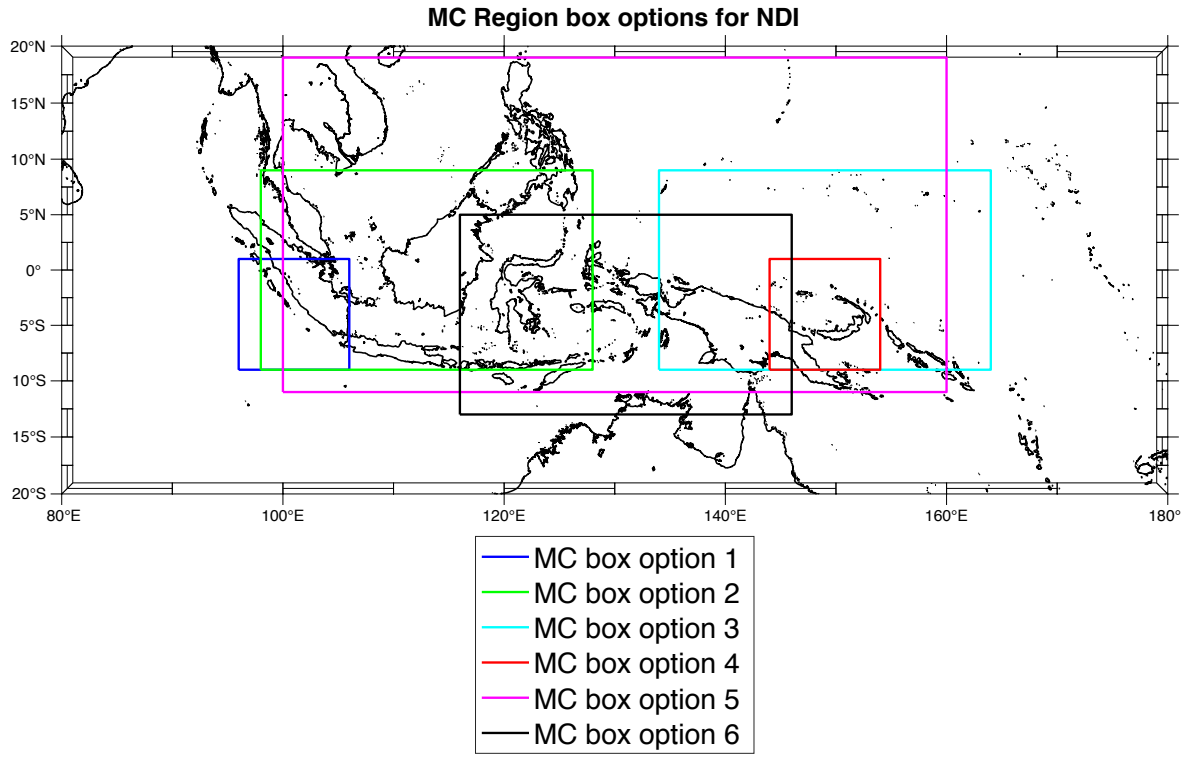


Figure 35. Six MC box option locations for the final NDI options

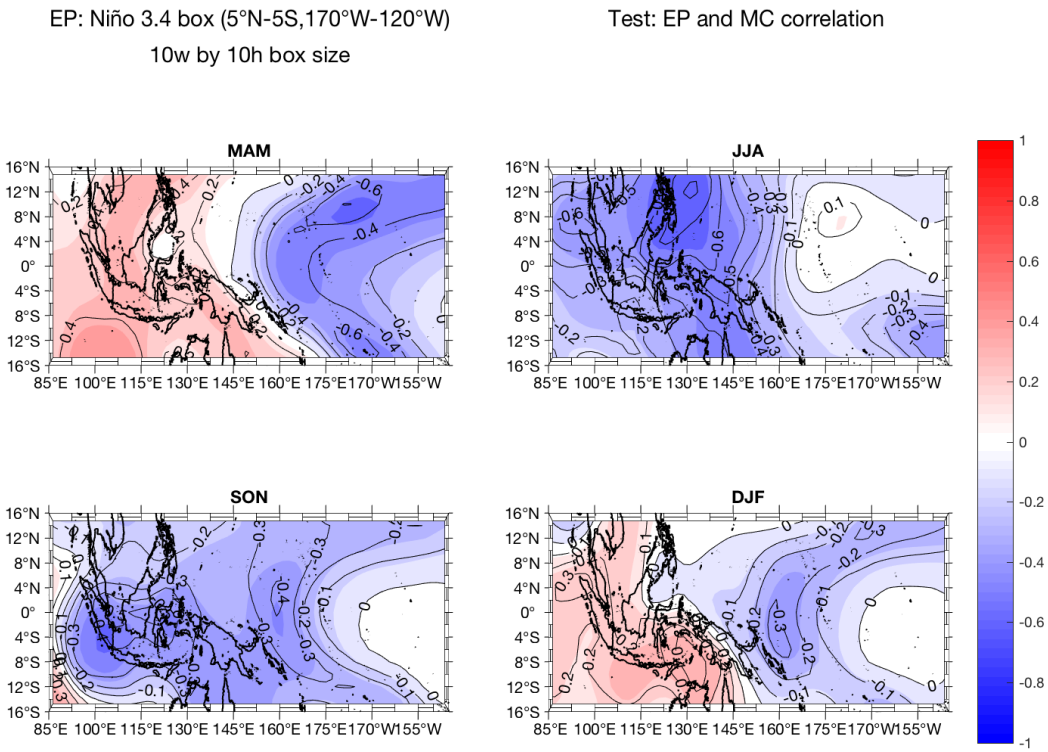


Figure 36. Linear regression of the Niño 3.4 index with GP PC 1. Correlations of the residuals and boxes that are 10w by 10h in size for standard seasons with a region domain that is extended eastward into the EP region. Slope values of the MC box options and the residuals are overlaid.

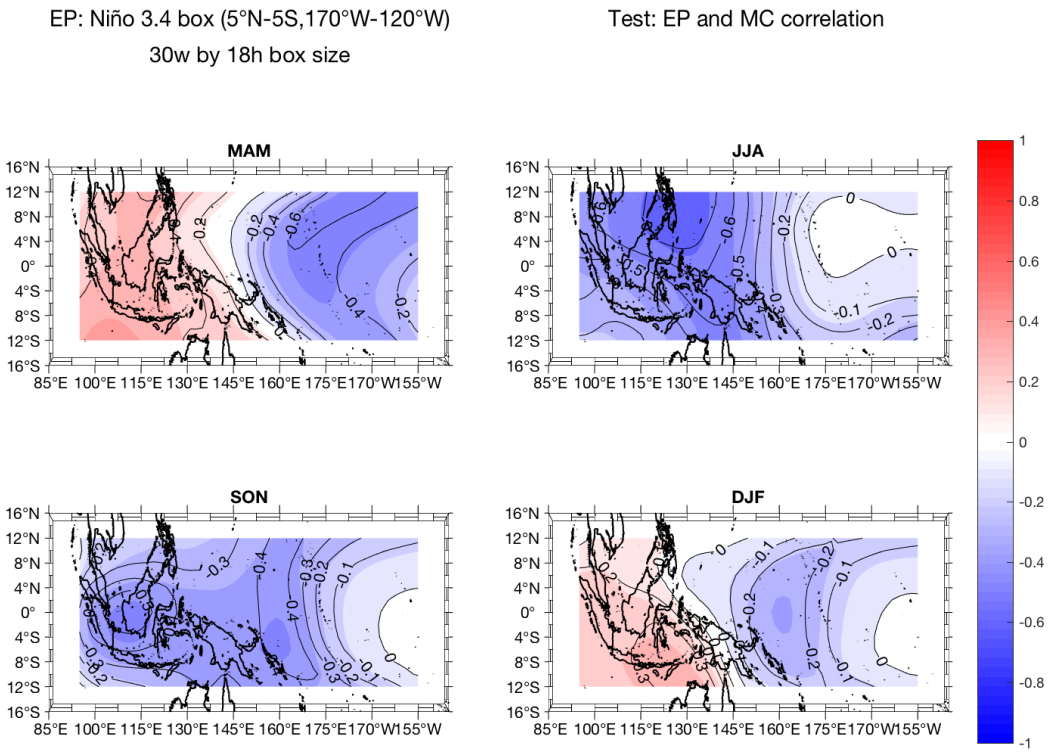


Figure 37. Linear regression of the Niño 3.4 index with GP PC 1. Correlations of the residuals and boxes that are 30w by 18h in size for standard seasons with a region domain that is extended eastward into the EP region. Slope values of the MC box options and the residuals are overlaid.

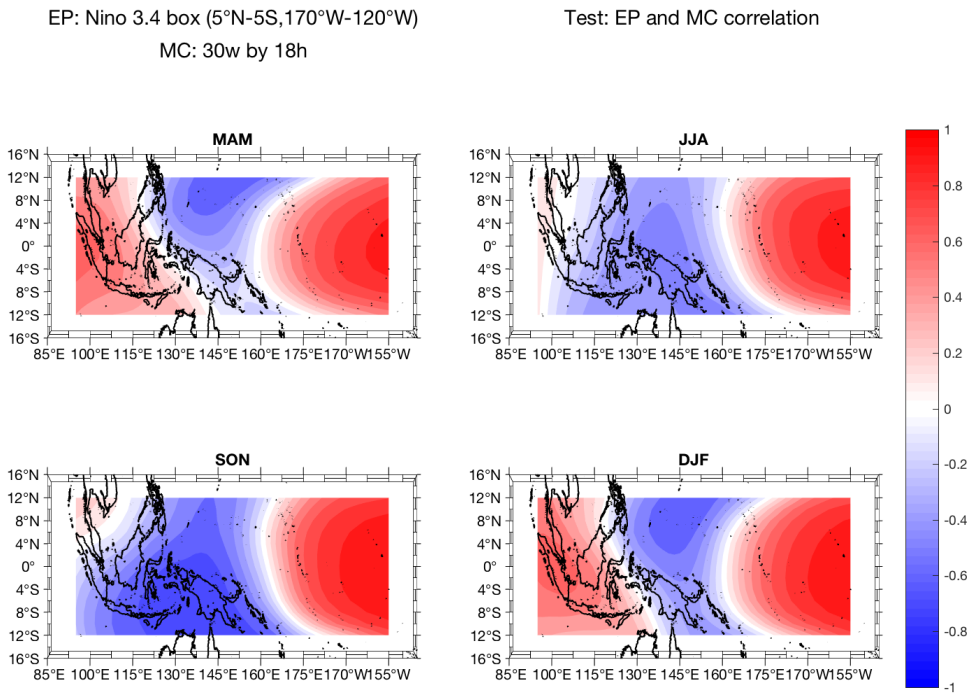


Figure 38. Standard season correlations of Niño 3.4 box SST's with SST boxes that are 30w by 18h in size.

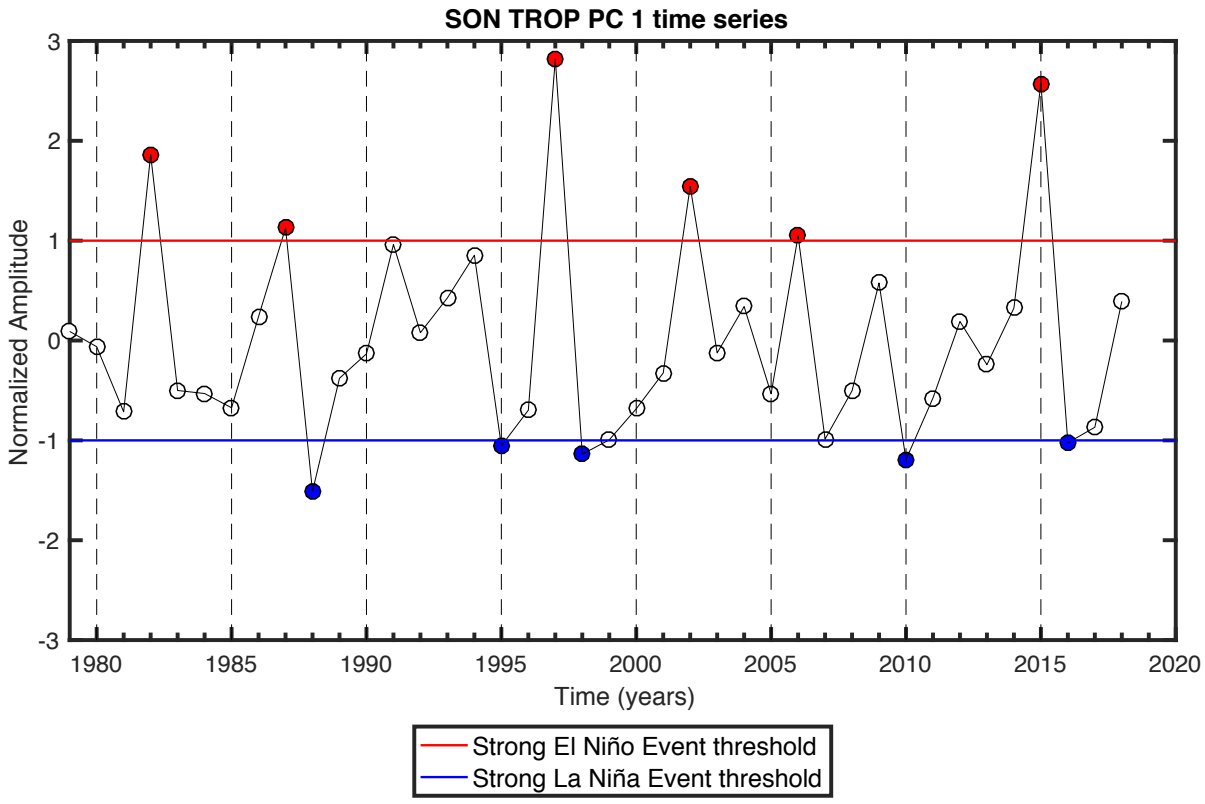


Figure 39. SON time series of TROP PC 1. Strong El Niño/La Niña events are classified when the annual PC time series value exceeds +/-1, respectively. Strong El Niño (La Niña) events are labeled as filled red (blue) circles.

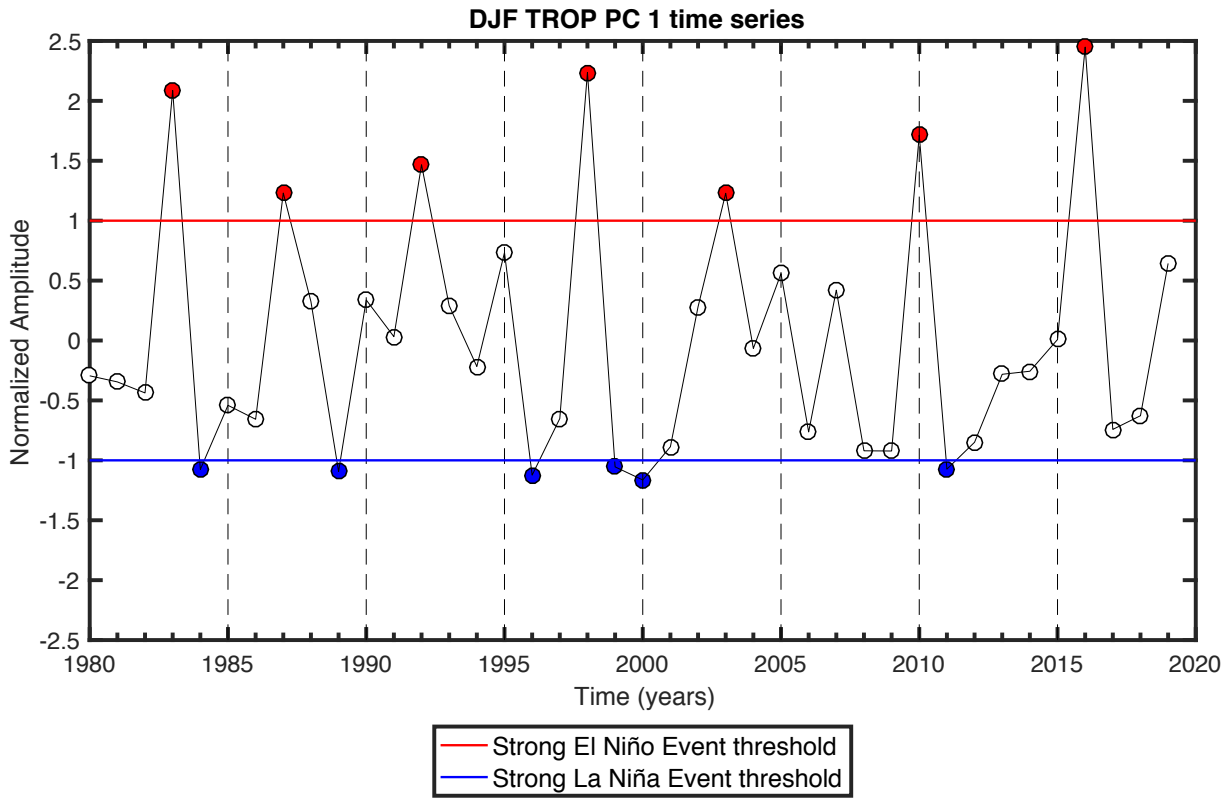


Figure 40. DJF time series of TROP PC 1. Strong El Niño/La Niña events are classified when the annual PC time series value exceeds +/-1, respectively. Strong El Niño (La Niña) events are labeled as filled red (blue) circles. For reference, DJF of 1979/1980 is labeled as the year 1980 (x-axis).

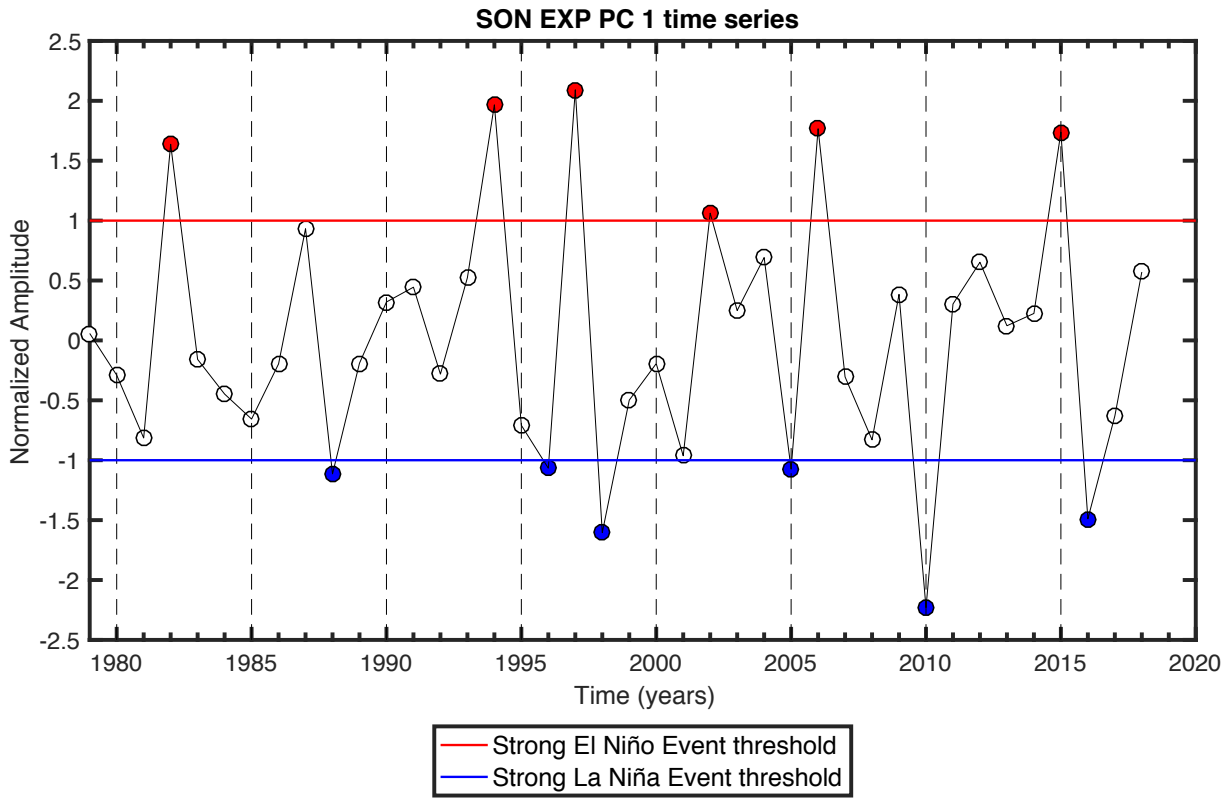


Figure 41. SON time series of EXP PC 1. Strong El Niño/La Niña events are classified when the annual PC time series value exceeds +/-1, respectively. Strong El Niño (La Niña) events are labeled as filled red (blue) circles.

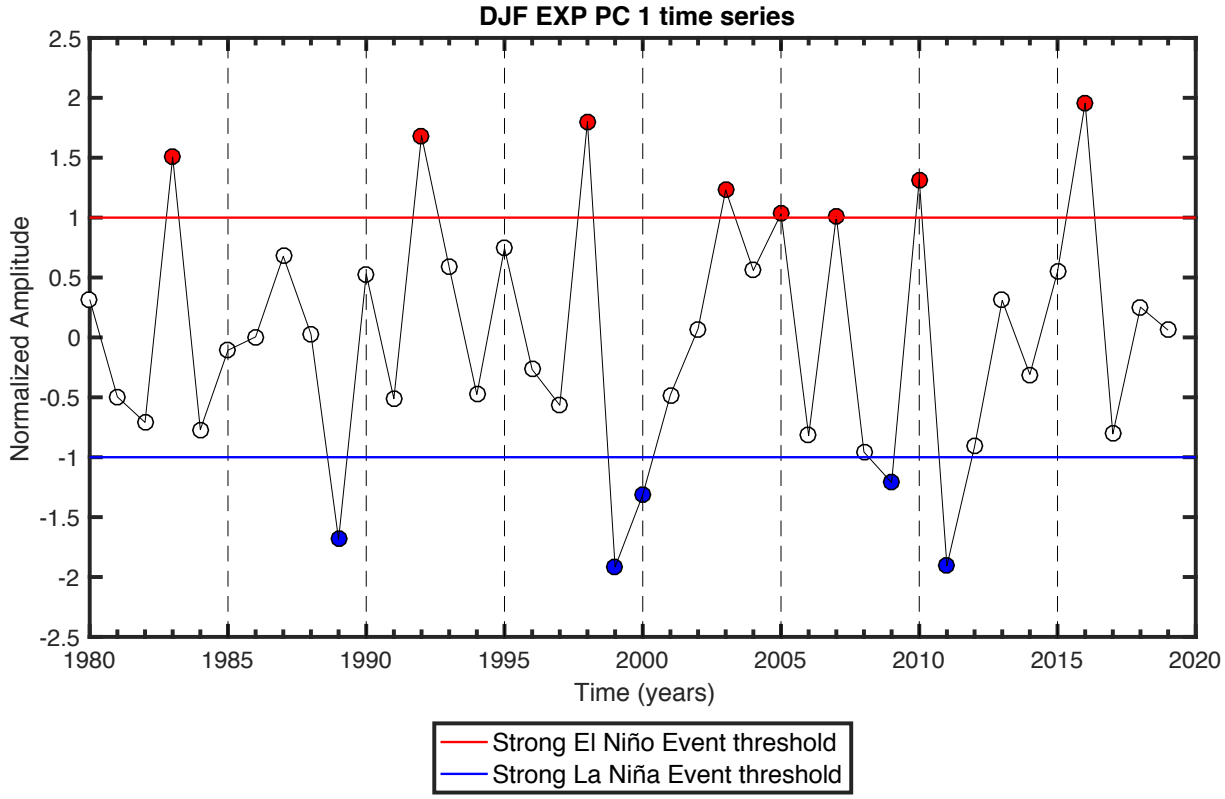


Figure 42. DJF time series of EXP PC 1. Strong El Niño/La Niña events are classified when the annual PC time series value exceeds +/-1, respectively. Strong El Niño (La Niña) events are labeled as filled red (blue) circles. For reference, DJF of 1979/1980 is labeled as the year 1980 (x-axis).

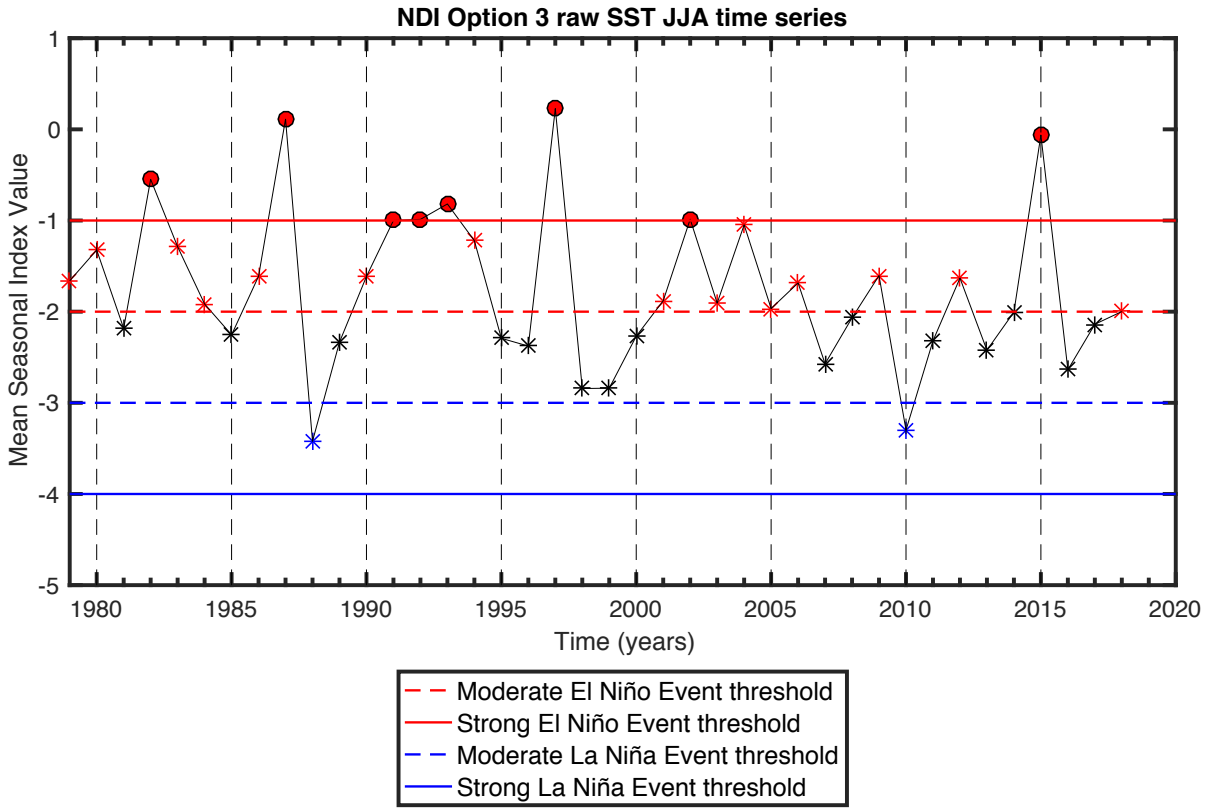


Figure 43. NDI option 3 annual mean value time series for JJA. Years that meet the moderate El Niño (La Niña) event definition criteria are denoted as red (blue) stars and years that meet the strong El Niño (La Niña) event definition criteria are denoted as filled red (blue) circles.

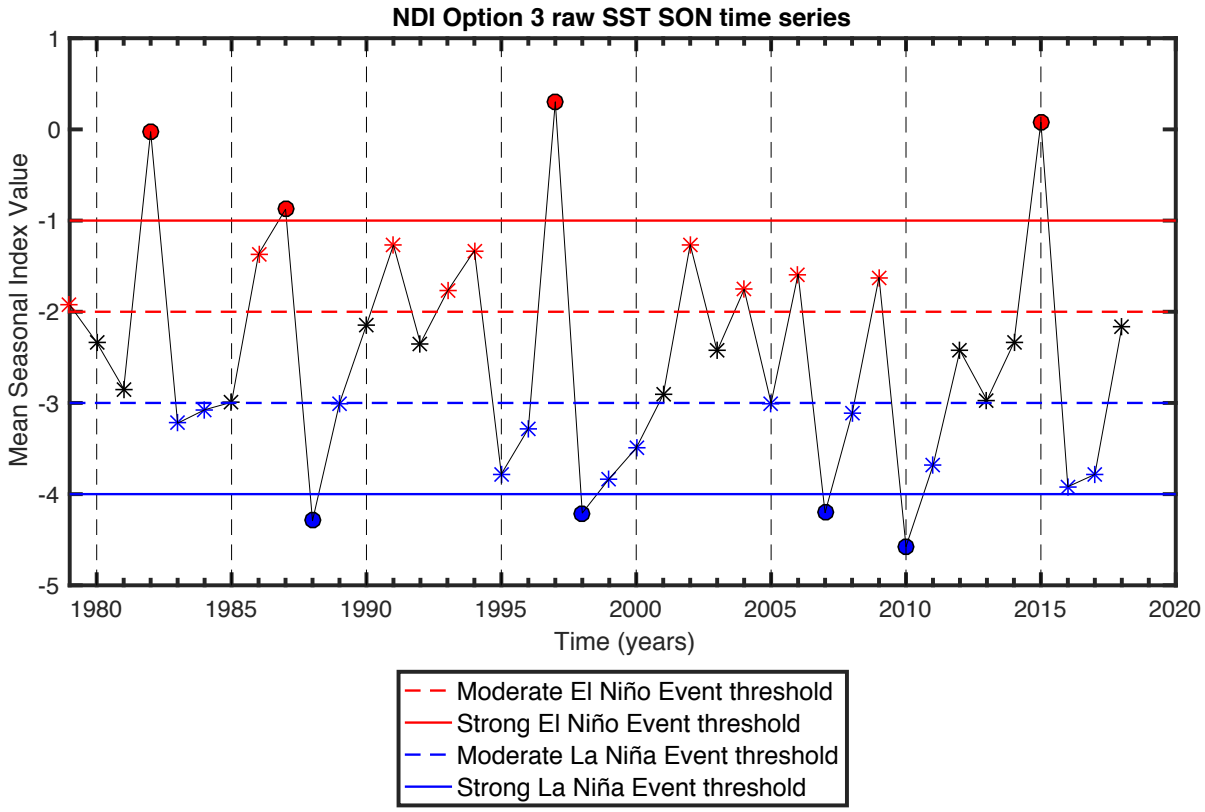


Figure 44. NDI option 3 annual mean value time series for SON. Years that meet the moderate El Niño (La Niña) event definition criteria are denoted as red (blue) stars and years that meet the strong El Niño (La Niña) event definition criteria are denoted as filled red (blue) circles.

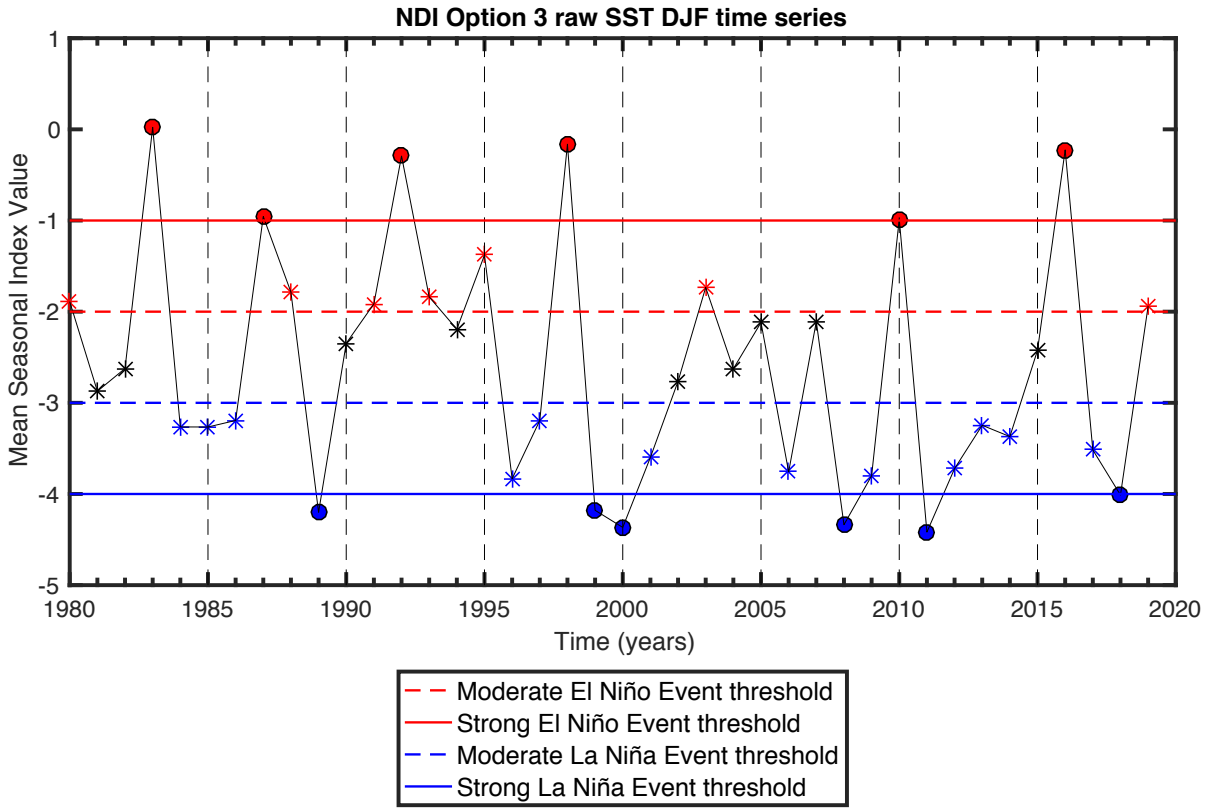


Figure 45. NDI option 3 annual mean value time series for DJF. Years that meet the moderate El Niño (La Niña) event definition criteria are denoted as red (blue) stars and years that meet the strong El Niño (La Niña) event definition criteria are denoted as filled red (blue) circles. For reference, DJF of 1979/1980 is labeled as the year 1980 (x-axis).

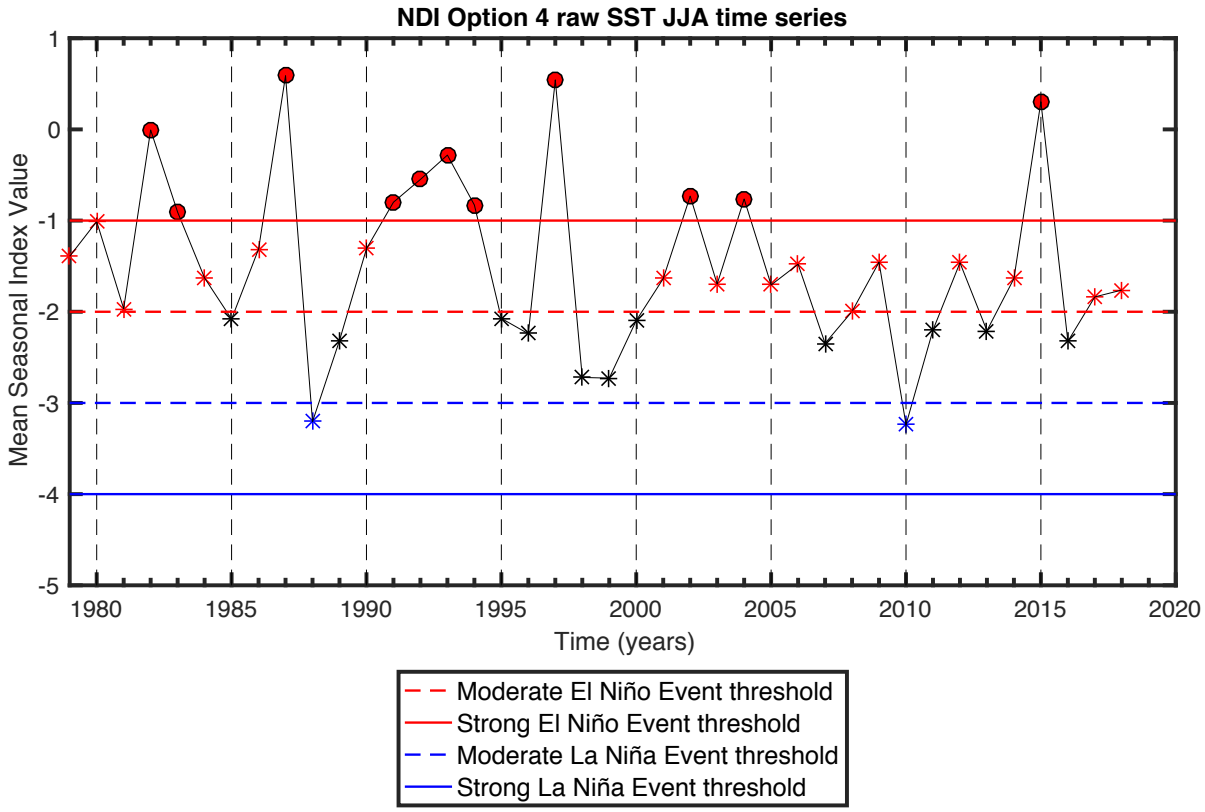


Figure 46. NDI option 4 annual mean value time series for JJA. Years that meet the moderate El Niño (La Niña) event definition criteria are denoted as red (blue) stars and years that meet the strong El Niño (La Niña) event definition criteria are denoted as filled red (blue) circles.

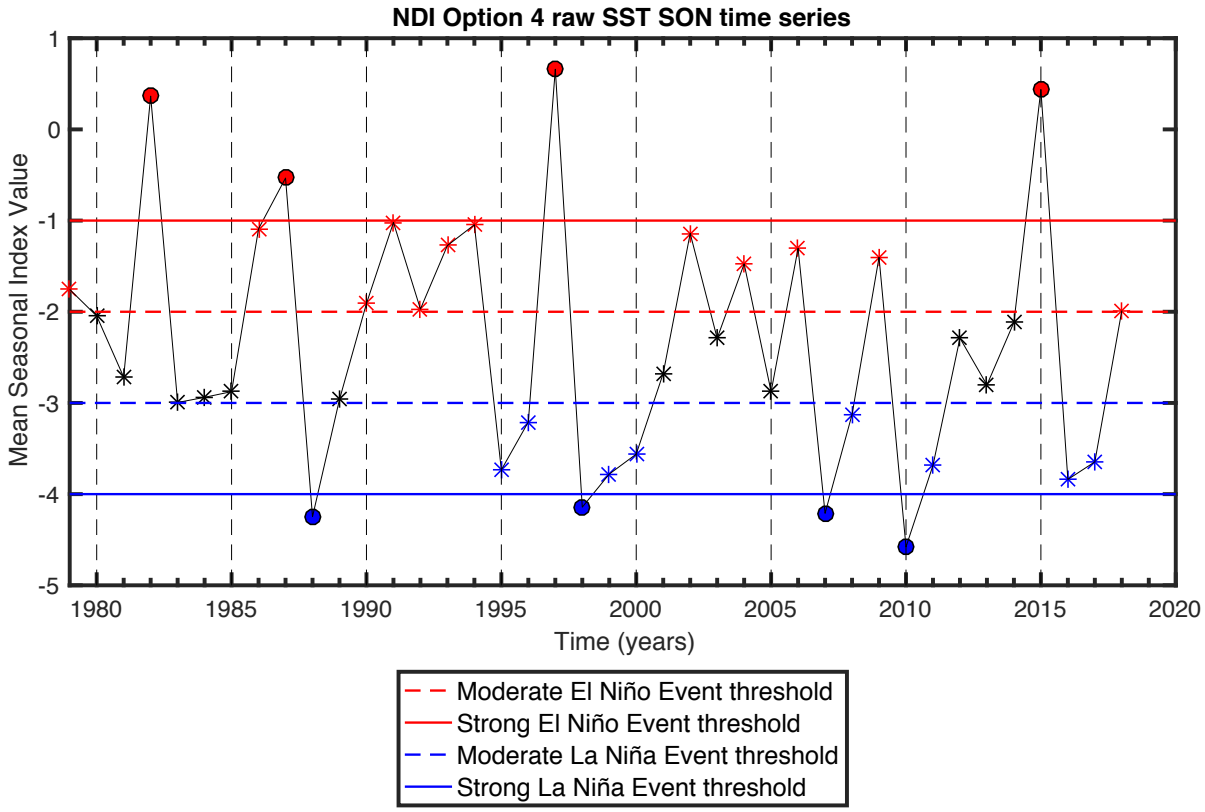


Figure 47. NDI option 4 annual mean value time series for SON. Years that meet the moderate El Niño (La Niña) event definition criteria are denoted as red (blue) stars and years that meet the strong El Niño (La Niña) event definition criteria are denoted as filled red (blue) circles.

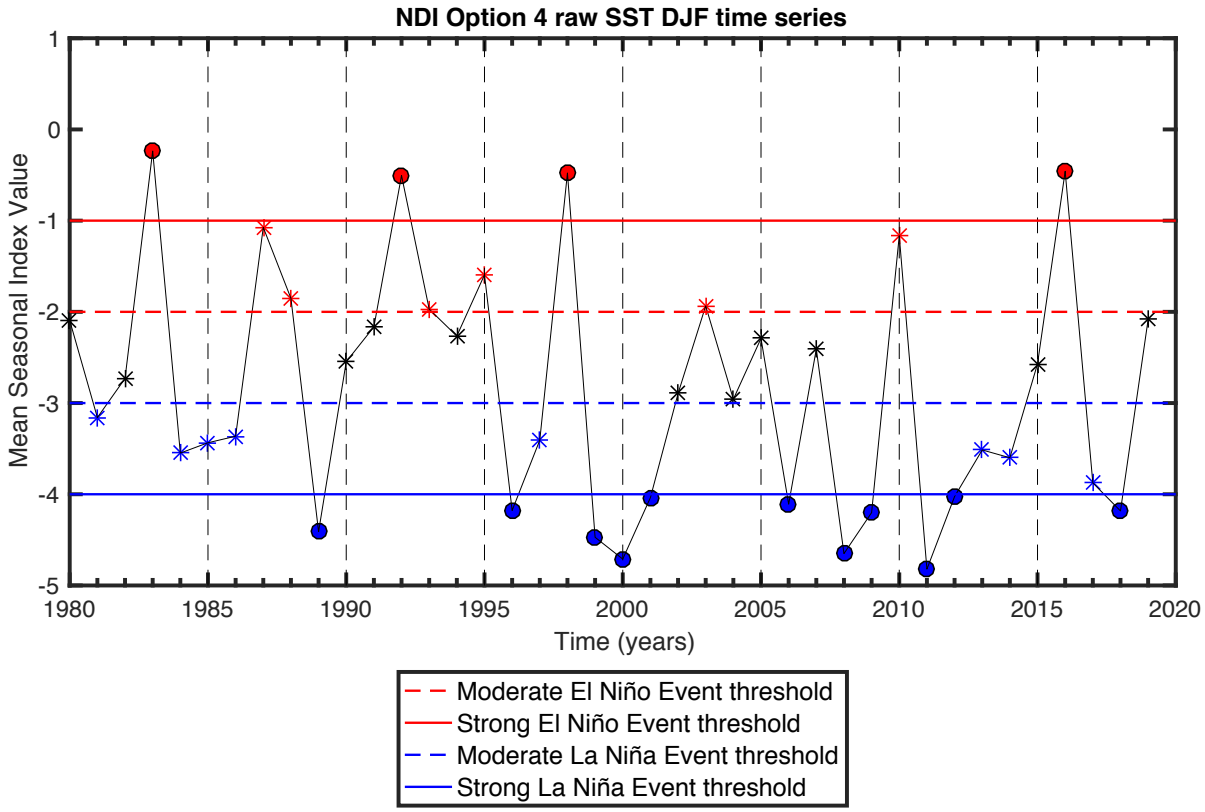


Figure 48. NDI option 4 annual mean value time series for DJF. Years that meet the moderate El Niño (La Niña) event definition criteria are denoted as red (blue) stars and years that meet the strong El Niño (La Niña) event definition criteria are denoted as filled red (blue) circles. For reference, DJF of 1979/1980 is labeled as the year 1980 (x-axis).

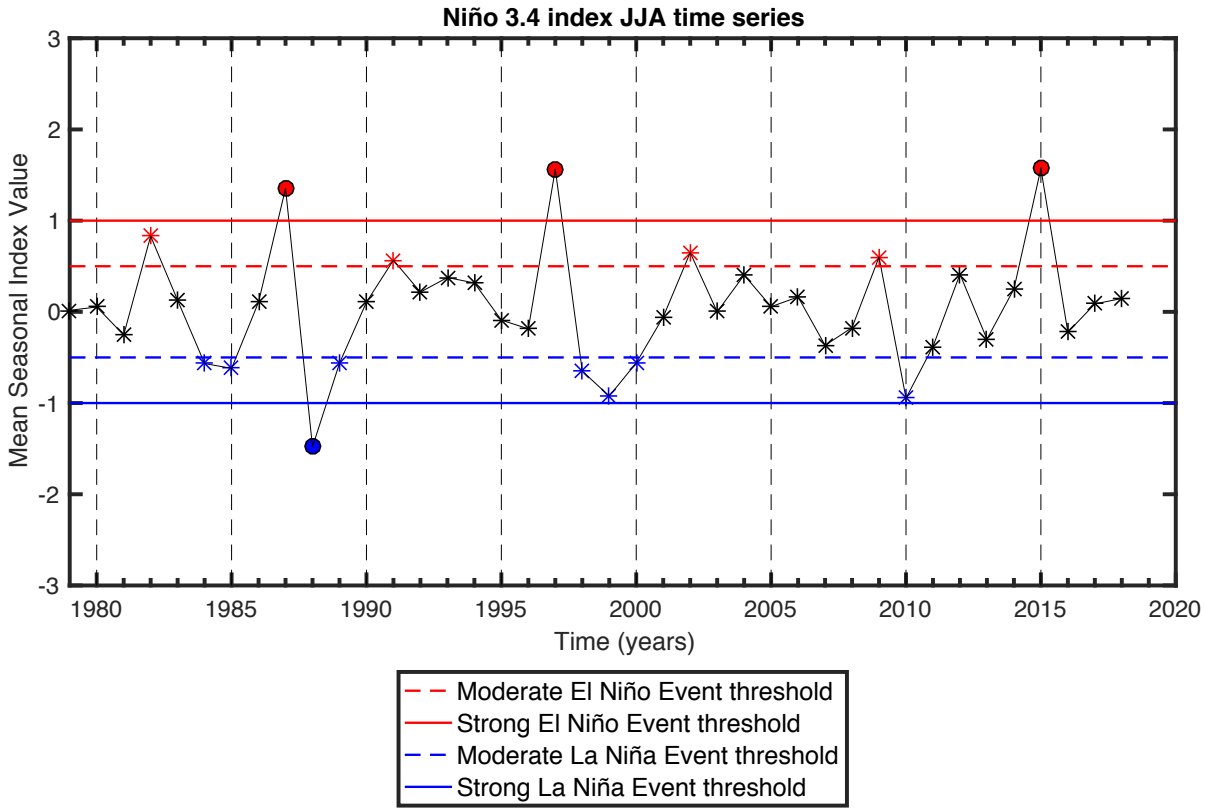


Figure 49. Niño 3.4 index annual mean value time series for JJA. Years that meet the moderate El Niño (La Niña) event definition criteria are denoted as red (blue) stars and years that meet the strong El Niño (La Niña) event definition criteria are denoted as filled red (blue) circles.

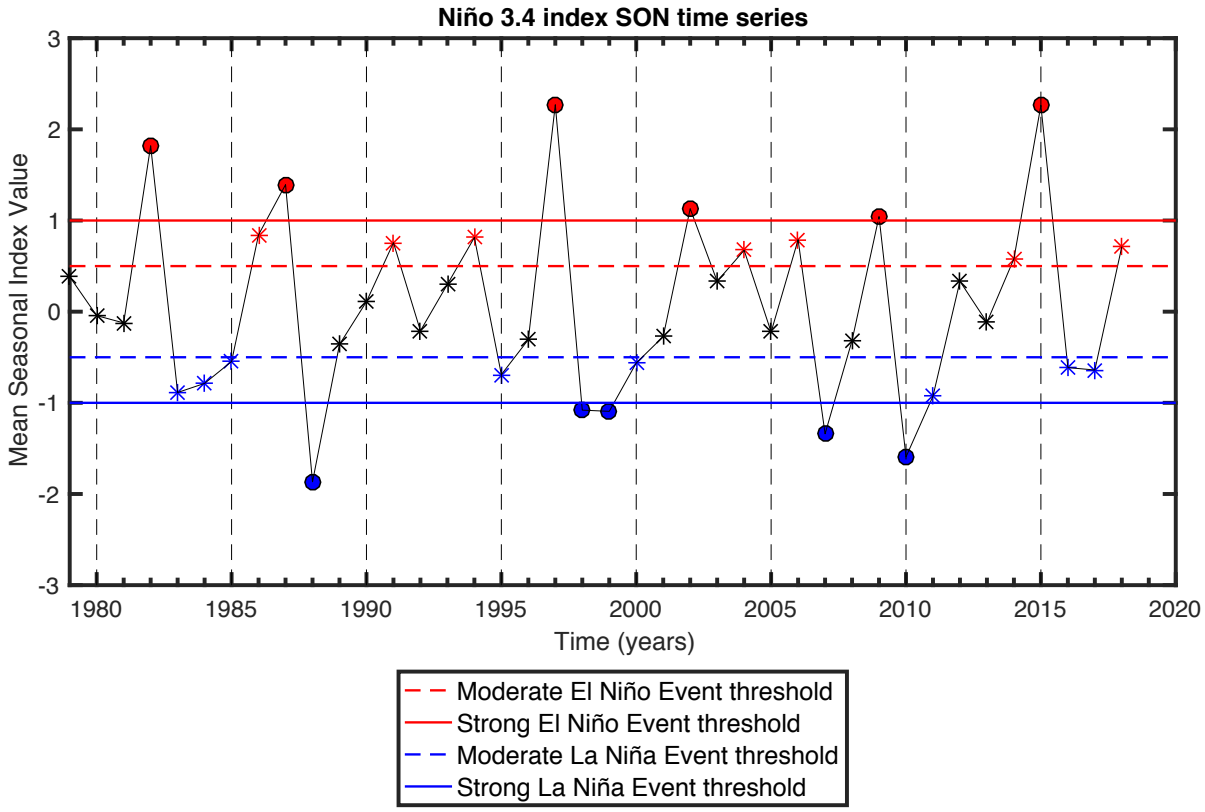


Figure 50. Niño 3.4 index annual mean value time series for SON. Years that meet the moderate El Niño (La Niña) event definition criteria are denoted as red (blue) stars and years that meet the strong El Niño (La Niña) event definition criteria are denoted as filled red (blue) circles.

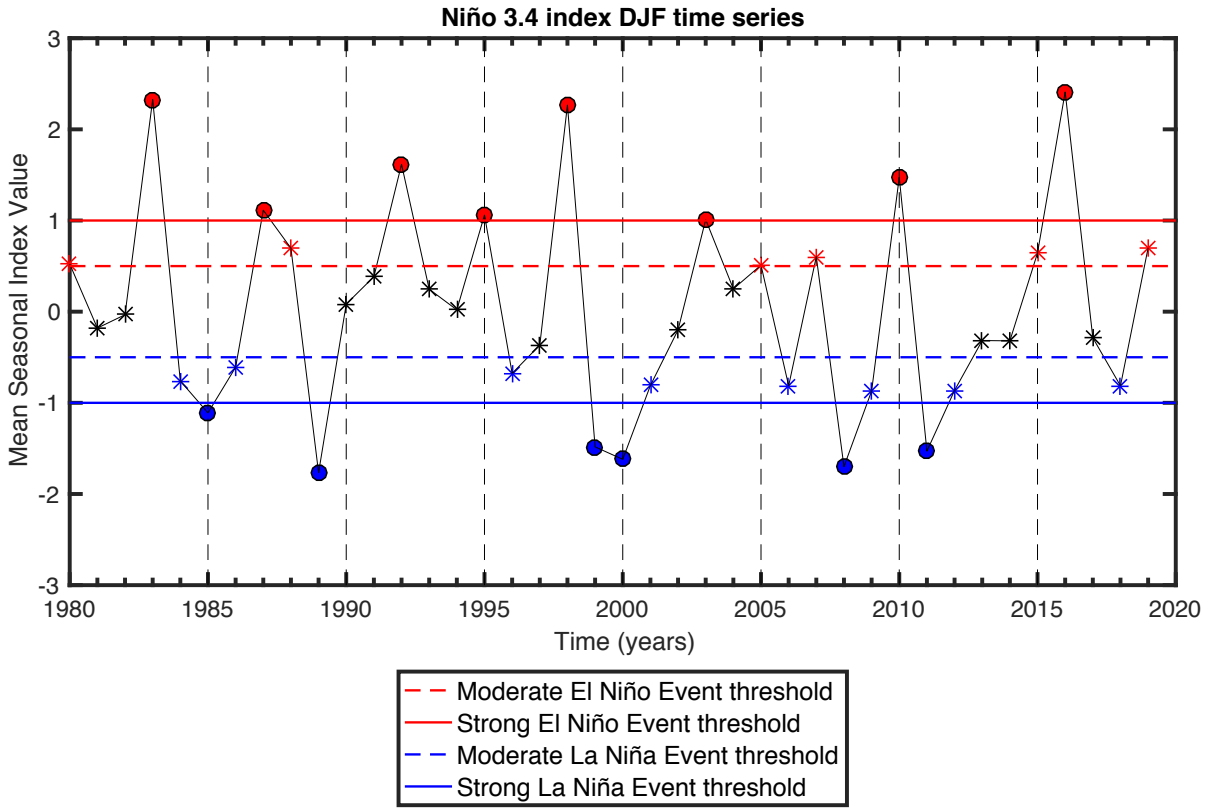


Figure 51. Niño 3.4 index annual mean value time series for DJF. Years that meet the moderate El Niño (La Niña) event definition criteria are denoted as red (blue) stars and years that meet the strong El Niño (La Niña) event definition criteria are denoted as filled red (blue) circles. For reference, DJF of 1979/1980 is labeled as the year 1980 (x-axis).