Climate Dynamics

# Predictability of Phases and Magnitudes of Natural Decadal Climate Variability Phenomena in CMIP5 Experiments with the UKMO HadCM3, GFDL-CM2.1, NCAR-CCSM4, and MIROC5 Global Earth System Models

--Manuscript Draft--

| | |
|---|---|
| **Manuscript Number:** | CLDY-D-17-00230 |
| **Full Title:** | Predictability of Phases and Magnitudes of Natural Decadal Climate Variability Phenomena in CMIP5 Experiments with the UKMO HadCM3, GFDL-CM2.1, NCAR-CCSM4, and MIROC5 Global Earth System Models |
| **Article Type:** | Original Article |
| **Keywords:** | Decadal climate variability; climate predictability; Pacific Decadal Oscillation; volcanic eruptions |
| **Corresponding Author:** | Vikram Mehta, Ph.D.<br>The Center for Research on the Changing Earth System<br>Catonsville, Maryland UNITED STATES |
| **Corresponding Author Secondary Information:** | |
| **Corresponding Author's Institution:** | The Center for Research on the Changing Earth System |
| **Corresponding Author's Secondary Institution:** | |
| **First Author:** | Vikram Mehta, Ph.D. |
| **First Author Secondary Information:** | |
| **Order of Authors:** | Vikram Mehta, Ph.D. |
| | Hui Wang, Ph.D. |
| | Katherin Mendoza, M.S. |
| **Order of Authors Secondary Information:** | |
| **Funding Information:** | National Institute of Food and Agriculture (2011-67003-30213) — Dr. Vikram Mehta<br>National Aeronautics and Space Administration (NNX15AD18A) — Dr. Vikram Mehta<br>U.S. Army Corps of Engineers (W912HQ-15-P-0056) — Dr. Vikram Mehta |
| **Abstract:** | Data from decadal hindcast experiments conducted under CMIP5 were used to assess the ability of CM2.1, HadCM3, MIROC5, and CCSM4 Earth System Models (ESMs) to hindcast sea-surface temperature (SST) indices of three decadal climate variability phenomena - the Pacific Decadal Oscillation (PDO), the tropical Atlantic SST gradient (TAG) variability, and the West Pacific Warm Pool (WPWP) SST variability - from 1961 to 2010. Aerosol optical depth (AOD) and other external forcings were specified in these experiments, and the ESMs were initialized at specific times with observed data to make ten- and 30-year hindcasts/forecasts.<br>All ESMs hindcast occurrence frequencies of positive and negative phases of the indices, and probabilities of same-phase transitions from one year to the next reasonably well. Except for the PDO in the 1980s, no one of the decade-average hindcasts show significant skill. Major volcanic eruptions are associated with phase transitions of indices in observed data and in some of the ensemble-average hindcasts. Some phase transitions associated with volcanic eruptions are also present in non-initialized simulations with these ESMs. Hindcasts from some of the ESMs show correct phase transitions in the absence of AOD changes also, implying that initializations with observed data are beneficial in predicting phase transitions. The best-performing ESM, MIROC5, predicts PDO and WPWP indices to decrease from maxima in 2016 to minima in 2018-19. The skills of PDO and WPWP indices' phase |

| | |
|---|---|
| | prediction up to at least two years in advance, and perhaps longer, can be used to inform societal impacts management decisions.<br><br>Key words:  Decadal climate variability;climate predictability;Pacific Decadal Oscillation;volcanic eruptions |
| **Suggested Reviewers:** | Holger Pohlmann<br>Max-Planck-Institut fur Meteorologie<br>holger.pohlmann@mpimet.mpg.de<br>Dr. Pohlmann is a pioneer in decadal climate predictability, the topic of this manuscript. |
| | Georgiy Stenchikov<br>Georgiy.Stenchikov@kaust.edu.sa<br>DR. Stenchikov is an expert on climate effects of volcanic eruptions, a topic of this manuscript. |
| | Arun Kumar<br>Arun.Kumar@noaa.gov<br>Dr. Kumar is an expert on climate predictability and prediction, a topic of this manuscript. |
| | Noel Keenlyside<br>Noel.Keenlyside@uib.no<br>Dr. Keenlyside is a pioneer in decadal climate predictability, the main topic of this manuscript. |

# Predictability of Phases and Magnitudes of Natural Decadal Climate Variability Phenomena in CMIP5 Experiments with the UKMO HadCM3, GFDL-CM2.1, NCAR-CCSM4, and MIROC5 Global Earth System Models

**Vikram M. Mehta•, Katherin Mendoza, and Hui Wang#**
Center for Research on the Changing Earth System
Catonsville, Maryland 21228

Submitted for publication in *Climate Dynamics*

- Corresponding author address: Vikram M. Mehta, Center for Research on the Changing Earth System, 5523 Research Park Drive, Suite 205, Catonsville, Maryland 21228, U.S.A.
- E-mail address: vikram@crces.org; Phone: 443-543-5493

# Current affiliation: NOAA/Climate Prediction Center, College Park, Maryland.

## Abstract

Data from decadal hindcast experiments conducted under CMIP5 were used to assess the ability of CM2.1, HadCM3, MIROC5, and CCSM4 Earth System Models (ESMs) to hindcast sea-surface temperature (SST) indices of three decadal climate variability phenomena – the Pacific Decadal Oscillation (PDO), the tropical Atlantic SST gradient (TAG) variability, and the West Pacific Warm Pool (WPWP) SST variability – from 1961 to 2010. Aerosol optical depth (AOD) and other external forcings were specified in these experiments, and the ESMs were initialized at specific times with observed data to make ten- and 30-year hindcasts/forecasts.

All ESMs hindcast occurrence frequencies of positive and negative phases of the indices, and probabilities of same-phase transitions from one year to the next reasonably well. Except for the PDO in the 1980s, no one of the decade-average hindcasts show significant skill.  Major volcanic eruptions are associated with phase transitions of indices in observed data and in some of the ensemble-average hindcasts.  Some phase transitions associated with volcanic eruptions are also present in non-initialized simulations with these ESMs.  Hindcasts from some of the ESMs show correct phase transitions in the absence of AOD changes also, implying that initializations with observed data are beneficial in predicting phase transitions.  The best-performing ESM, MIROC5, predicts PDO and WPWP indices to decrease from maxima in 2016 to minima in 2018-19.  The skills of PDO and WPWP indices' phase prediction up to at least two years in advance, and perhaps longer, can be used to inform societal impacts management decisions.

**Key words:** Decadal climate variability;climate predictability;Pacific Decadal Oscillation;volcanic eruptions

## 1. Introduction

Societies have sought skillful climate prediction at monthly to decadal lead times for centuries, primarily for use in management of water resources and in planning agricultural activities. It continues to be increasingly recognized now that skillful decadal climate predictions can greatly benefit planning in many societal sectors, such as agriculture, reservoir operations, municipal water supply and drainage systems, hydro-electricity generation, transportation, fisheries and wildlife habitat manitenance, thermal and nuclear power plant operations, river- and reservoir-based recreation industry, forest fires, and state and national government decisions (Mehta et al., 2013a; Meehl et al., 2014; Mehta, 2017). In addition to the importance of decadal climate prediction for societal impacts prediction and planning, it is also important for understanding and attribution of past, current, and future climate to natural decadal climate variability (DCV) or anthropogenic climate change. In order for stakeholders and policymakers to use decadal climate predictions, it is very important to establish a prediction skill record by using prediction models and past, observed climate data – both for model initialization as well as for prediction verification – to make retrospective predictions, or "hindcasts", of past climate as envisaged in the World Climate Research Program's Coupled Model Intercomparison Project (CMIP) 5 and follow-on Projects. It is also very important to assess climate information needs of stakeholders and policymakers, and orient prediction research towards satisfying those needs as envisaged in the World Meteorological Organization's Global Framework for Climate Services Vision[1] "To enable better management of the risks of climate variability and change and adaptation to climate change, through the development and incorporation of science-based climate

[1] http://gfcs-climate.org/

information and prediction into planning, policy and practice on the global, regional and national scale."

The climate during a period of one or two decades consists of several interacting components, therefore prospects for decadal climate prediction depend on prospects for skillful predictions/projections of interannual variability such as El Niño-Southern Oscillation (ENSO); natural DCV, including climate system responses to variations in solar particulate and radiative emissions, and to volcanic eruptions; and responses to human-induced changes in land use-cover and atmospheric constituents. The present study focuses on one of these components - namely, natural DCV. As the most recent report of the Inter-governmental Panel on Climate Change states (IPCC, 2013), “Natural internal variability will continue to be a major influence on climate, particularly in the near-term and at the regional scale. By the mid-21st century the magnitudes of the projected changes are substantially affected by the choice of emissions scenario.” Thus, for the next 30 to 40 years, natural climate variability will continue to be more important than climate change. After 40 years also, natural climate variability will still contribute substantially to the totality of climate impacts.

Among natural DCV phenomena, the Pacific climate variability generally known as the Pacific Decadal Oscillation (PDO; Mantua et al., 1997) or the Inter-decadal Pacific Oscillation (IPO; Power et al., 1999), the tropical Atlantic sea surface temperature (SST) gradient (TAG; Hastenrath, 1990; Houghton and Tourre, 1992; Mehta and Delworth, 1995; Mehta, 1998; Rajagopalan et al., 1998), and variability of the West Pacific Warm Pool (WPWP) SST (Wang and Mehta, 2008), and their impacts on global climate are attracting increasing attention in predictability and prediction studies because of their impacts on water resources, agriculture, hydro-electricity generation, inland water-borne transportation, and fish and crustacean stocks and captures (Mehta, 2017). Analyses of associations between SST indices of these three natural DCV phenomena; and decadal – multidecadal variability of global precipitation, temperatures, and the Palmer Drought Severity Index (PDSI) show that approximately 60 – 90% variance in these three hydro-meteorological variables on land is explained by the PDO, the TAG SST variability, and the WPWP SST variability (see, for example, Mehta (2017)).

The present study is a part of a program to develop a decadal climate and impacts simulation and prediction system for the Missouri River Basin (MRB)[2], to develop adaptation options for water and agriculture sectors in the MRB using decadal climate and impacts information, and to develop a methodology to estimate the value of decadal climate and impacts information to the agriculture sector. Global Earth System Models (ESMs) and a very high-resolution land use – hydrology – crop model are being used in this program. From this program, preliminary results on decadal predictability of ocean basin averaged SSTs in decadal hindcast experiments with the Geophysical Fluid Dynamics Laboratory CM2.1, the U.K. Meteorological Office HadCM3, the Japanese Model for Interdisciplinary Research On Climate 5 (MIROC5), and the National Center for Atmospheric Research–CCSM4 ESMs in CMIP5 were reported in Mehta et al. (2013b); and a dynamical–statistical technique for decadal hydro-meteorological predictions being developed–applied to southern Africa as a test case - was reported in Mehta et al. (2014). Research designed to simulate impacts of DCV phenomena on surface and ground water in the MRB is reported in Daggupati et al. (2016) and Mehta et al. (2016), and on wheat yields in the MRB is reported in Mehta et al. (2017a). The value of decadal climate information to the agriculture sector in the MRB is estimated by Fernandez et al. (2016). The ability of the CM2.1,

[2] The MRB is the largest river basin in the U.S.; and is a major “bread basket” of the U.S. and the world, producing approximately 45% of wheat, 20% of grain corn, and 33% of cattle produced in the U.S..

HadCM3, MIROC5, and CCSM4 ESMs in CMIP5 to simulate major attributes of the PDO, the TAG variability, and the WPWP variability is described in a companion paper (Mehta et al., 2017b). The ability of these four ESMs to retrospectively forecast (or, hindcast) the three DCV phenomena is addressed in the present paper. These four ESMs were selected because it is important to assess simulation and hindcast skills of the same ESMs in the same experimental framework. The modeling groups who have developed these four ESMs conducted CMIP5 experiments with generally the same model configurations. Also, decadal hindcast/forecast experiments with these four ESMs were run in CMIP5 in the ensemble mode with up to 10 members in each ensemble.

### 1.1 Review of Previous Research

Perhaps the earliest recorded instance of prediction of impacts of decadal climate anomalies was by Sir William Herschel, a noted German – British astronomer and music composer. Having observed variations in sunspots, Herschel (1801) hypothesized that variations in sunspot numbers implied variations in solar irradiance which might cause variations in atmospheric heating, rainfall, and temperature, and thereby influence the price of wheat in London. Herschel's initial and controversial investigation, motivated by the desire for prediction of agricultural productions and prices, was followed by a subsequent investigation by Carrington (1863). Jevons (1879) found a correlation between sunspot variation and wheat price in India. Poynting (1884) found correlations between sunspot variation and wheat price, and cotton and silk imports into Great Britain. Since Schwabe (1884)'s discovery of the 11-year sunspot cycle, the sunspot–terrestrial climate–societal impacts investigations were essentially focused on externally-forced decadal climate and impacts prediction. Subsequent analyses of correlations between sunspot numbers, and a wide variety of natural phenomena and production of food and wealth -

and predictions based on these correlations - have continued into the $21^{st}$ Century Current Era (CE) (see, for example, Proctor (1880), Chambers (1886), Currie (1974), King et al. (1974), Meadows (1975), Harrison (1976), Vines (1977), Currie and Fairbridge (1985), Currie et al. (1993), Mehta and Lau (1997), Garnett et al. (2006), Pustil'nik and Yom Din (2004a, 2004b, 2009, 2013), and Love (2013)). There is also a voluminous published literature on associations between the 18.6-year lunar nodal cycle and a variety of hydro-meteorological and oceanographic variables, their impacts on several societal sectors, and their prediction. Thus, the field of externally-forced DCV, its impacts, and their prediction is over two centuries old.

The availability of archives of multidecades-long oceanic observations since the end of the Cold War in the early 1990s CE, quality-checked and model-assimilated global atmospheric observations, and the development of climate models incorporating increasingly realistic descriptions of physical processes has resulted in , a substantial body of research in the last two decades. This research is focused on understanding causes and mechanisms of DCV and putting seasonal to interannual climate prediction experience (McPhaden et al., 2010) to use in decadal climate prediction despite fundamental and substantial problems in using the seasonal to interannual climate prediction methodology for decadal climate prediction. Some major problems are (Meehl et al., 2009, 2014; Mehta et al., 2011a): (1) relatively short time series of instrument-based global ocean observations, especially sub-surface observations, for understanding, model initialization, and comparison with prediction; (2) an insufficient understanding of fundamental physics of DCV; (3) an insufficient theoretical understanding of possible behaviors of geographically-varying, complex and non-linear dynamical systems with mixed initial and boundary values; (4) global climate models displaying less than satisfactory skill in simulating climate in general and DCV in particular; and (5) insufficient guidance from stakeholders and policymakers as to which DCV-related climate, weather, and impacts information would be useful for applications to societal impacts of DCV if predicted. As a result, much of the experimental decadal climate prediction work so far is empirical and ad hoc, based on experimentation with various model configurations, prediction initialization schemes, ensemble sizes, forcing fields, and other aspects of numerical climate prediction. In spite of these problems, however, there have been many encouraging decadal prediction studies with ESMs, beginning with pioneering research by Smith et al. (2007), Keenlyside et al. (2008), and Pohlmann et al. (2009). In these three studies, ESMs were initialized from observed data - as in weather and seasonal climate forecasting - and natural and anthropogenic changes in aerosol optical depth (AOD) were prescribed from observations-based estimates (or scenarios) - as in anthropogenic climate change projection experiments. Smith et al. (2007) showed that skillful decadal prediction of global-average temperature may be possible. Keenlyside et al. (2008)'s and Pohlmann et al. (2009)'s results showed that skillful prediction of decadal, North Atlantic SSTs may be possible. Building on these studies, Yang et al. (2012) found that an inter-hemispheric, multidecadal SST pattern in the Atlantic may be predictable 4 to 10 years in advance.

Concurrently with these initial decadal climate predictability studies with ESMs, the World Climate Research Program organized the CMIP5 project to use ESMs to aid potential climate change assessments by the Inter-governmental Panel on Climate Change. CMIP5 also included experimental decadal hindcasts and forecasts (Taylor et al., 2012). Meehl et al. (2014) have described results from hitherto published CMIP5 and other decadal hindcasting experiments, so only major results pertaining to indices of decadal SST variability, and precipitation and surface air temperature on land areas are briefly summarized here.

There have been two types of assessments of prediction skill of the PDO index; one, correlation coefficient between observed and predicted indices or area-average SSTs over several decades, and two, prediction skill of specific warm or cold events. An example of the former type is a skill assessment of decadal hindcasts of the PDO index in five ESMs participating in CMIP5 by Kim et al. (2012) who found that there was a reasonably significant prediction skill for up to 5 years after prediction initialization, but that this skill was less than that derived from persistance of the PDO index. An example of the latter type is the improved prediction skill of the mid to late 1970s CE change in the PDO phase from negative (cold) to positive (warm) (described as climate shift by some researchers) in combined initial and boundary value experiments with several CMIP5 and other ESMs by Meehl and Teng (2012, 2014) compared to uninitialized experiments or simulations as boundary value experiments. Kim et al. (2012) also showed that the AMO index has a reasonably high prediction skill up to 7 years compared to the skill of persistance in five CMIP5 ESMs. As mentioned earlier and described in detail by Meehl et al. (2014), reasonably high skill of area-average North Atlantic SSTs is shown by several ESMs (see, for example, Keenlyside et al. (2008), Pohlmann et al. (2009), van Oldenborgh et al. (2012), Yang et al. (2012), Hazeleger et al. (2013), Ham et al. (2014), and others). Using decadal hindcast data from four CMIP5 ESMs, Mehta et al. (2013b) found that there was significant, but variable, decadal hindcast skill of global- and tropical ocean basin-average SSTs, among them the PDO region in the Pacific, during 1961 to 2010 CE. The skill varied by averaging region and decade. It was also found that volcanic eruptions influence SSTs and are one of the sources of decadal SST hindcast skill when significantly large eruptions occurred. In the four ESMs, decadal hindcast skills of SST anomalies over ocean basin size averaging regions generally improved due to model initialization with observed data.

These prediction skills of SSTs do not translate to comparable skills of precipitation and surface air temperature on land areas as shown by Doblas-Reyes et al. (2013). There is some skill, however, in northern Canada, northeast north America, and Greenland; southeast South America; some regions in sub-Saharan Africa; and scattered regions in central, south, north, and southeast Asia. Using decadal hindcast DCV indices (PDO, TAG, WPWP, and Niño 3.4) from CMIP5 experiments with the MIROC5 ESM in a regression-based statistical model, Mehta et al. (2014) also reported low to moderate decadal predictability of decadal hydrologic cycles, as represented by the PDSI, in seven countries of southern Africa from 1961 to 2010 CE. Kirtman et al. (2013) summarize conclusions about decadal prediction that "Predictions for averages of temperature, over large regions of the planet and for the global mean, exhibit positive skill when verified against observations for forecast periods up to ten years." They also conclude that "Predictions of precipitation over some land areas also exhibit positive skill." Thus, there is slow and incremental, but definite, progress in making skillful decadal climate predictions.

## 1.2 Objectives of the Present Study

Following seasonal to interannual climate prediction, the contemporary field of decadal climate prediction using dynamical models has also adopted the traditional numerical weather prediction approach. Specifically, prediction skill of a (or, the) final state of a variable, say the SST, is evaluated with respect to observations in terms of correlations and root-mean-square errors. Ensembles of multiple members are used to isolate a climate signal from noise arising from the non-linear model's chaotic behavior. It is believed that the goal should be to skillfully predict the final state, in this case a specific month or season ten years after starting the prediction experiment. But, new approaches need to be evolved for decadal climate prediction from the points of view of what is important for users of decadal climate information – stakeholders and policymakers – if the predicted information is to be useful for application. Although impacts of quantitative changes in DCV indices on hydro-meteorology (and, consequently, on water resources and agriculture) have not attracted much attention from researchers, impacts of DCV phases – positive and negative – are known much better via analyses of empirical data and via experiments with numerical models of the global atmosphere (e.g., Schubert et al., 2004a, 2004b). For example, data and information such as phase (positive or negative) of average anomaly in precipitation and temperature, stream/river flow, drought index, and other quantities over the next two to ten years can be very useful for management decisions in water and agriculture sectors if the data and information are provided at the spatial resolution required for each sector (Mehta et al., 2013a; Mehta, 2017). Therefore, understanding and prediction of DCV phase transitions sustained for several months to an year or longer can be useful in understanding and prediction of DCV impacts. Understanding and prediction of DCV phases is also important for attribution of DCV phase transitions to internal ocean-atmosphere processes or changes in external forcings.

A study of the value of decadal climate information to the agriculture sector in the MRB with a water and crop choices model showed that the correct prediction of important DCV phenomena that impact MRB agriculture one year in advance can be worth approximately $80 million per year (Fernandez et al., 2016). This study also showed that the correct prediction of even the phase of the important DCV phenomena next year, based on the phase in the current year, can realize a sizeable fraction of this monetary value. Moreover, it is also important to evolve combined dynamical – statistical prediction approaches for variables important to users that would translate useful skill in slower variables such as, for example, the PDO SST index, into applicable information about precipitation or drought index over one, two, five, or ten years.

Another reason to evolve different approaches for decadal climate prediction is that, unlike in weather prediction, variations/changes in external or boundary forcings such as solar radiations, volcanic and anthropogenic aerosols, anthropogenic greenhouse gases, and land use – land cover also influence/impact climate at the multiyear to decadal timescales. Since decadal predictions using dynamical models are made as a mixed initial – boundary value problem, contributions of both model initialization and boundary forcings in decadal prediction skill should be evaluated. Therefore, comparison of initialized predictions with uninitialized simulations with the same models is very important, especially the respective roles of boundary and initial conditions in phase transitions of DCV phenomena.

Based on the foregoing rationale, the objectives of this study are: (1) to assess transition probabilities of phases of the PDO, TAG, and WPWP indices, individually as well as in combinations of indices, in decadal hindcast experiments with the four selected ESMs and compare them with transition probabilities of observed indices; (2) to assess the skill of these ESMs to hindcast the phase and magnitude of the three DCV indices one and two years in advance; (3) to assess hindcast skill of the DCV indices over individual decades; (4) to understand the role of external forcings and internal ocean-atmosphere variability in phase transitions of DCV indices; and (5) to assess the impacts, if any, of initialization on hindcast skill. These objectives are addressed and results are interpreted in light of the fact that lead times of hindcasts vary from one to ten years in CMIP5 experiments with these four ESMs.

# 2. Materials and Methods

## 2.1 CMIP5 and Observational Data sets

Two sets of core decadal prediction experiments have been conducted under CMIP5 (Meehl et al., 2009). The first set is a series of 10-year hindcasts starting approximately in 1960, 1970, 1980, 1990, and 2000 CE. The second is a series of 30-year hindcasts starting in 1960, 1980, and 2005 CE, the last a combined hindcast-forecast. In both sets, AODs (including those due to volcanic eruptions) and solar radiation are prescribed from past observations. Each experiment has a minimum ensemble size of three members. These experiments are somewhat idealistic and exploratory, especially in view of the well-known difficulty of predicting volcanic eruptions well in advance.

We used SST and AOD data from the HadCM3, CM2.1, CCSM4, and MIROC5 ESMs. Table 1 summarizes major attributes of these models and the CMIP5 decadal hindcast experiments carried out with them. In the CMIP5 hindcast experiments, the CM2.1 used a fully-coupled initialization scheme (Zhang et al., 2007), the MIROC5 used an ocean-only initialization scheme (Tatebe et al., 2012), the CCSM4 used ocean and sea ice initial conditions from a historical forced experiment (Yeager et al., 2012), and the HadCM3 was initialized by relaxation to analyzed ocean and atmosphere observations (Smith et al., 2007). In all CMIP5 experiments, Northern Hemisphere and Southern Hemisphere time series of AOD, based on observations (Ammann et al. (2003) in the NCAR ESM, and Sato et al. (1993) and Hansen et al. (2002) in the other three ESMs), were specified. These data sets provide zonal-average, vertically-resolved AOD for visible wavelengths and column-average effective radii of aerosols (Stenchikov et al., 2006). We also combined hindcast data from the four ESMs as a multi-model ensemble (MME; Krishnamurti et al. (2000)). The MME in this study is the average of the ensemble-average data from each ESM. In this way, each ESM is treated equally in the MME. We used the Extended Reconstructed SSTs (ERSST; Reynolds et al., 2002) from 1961 to 2010 for comparison with hindcast SSTs.

## 2.2 Analysis Techniques

We calculated the PDO index from each decadal hindcast experiment by projecting hindcast SSTs from each ESM on the PDO patterns isolated from simulation runs with that ESM (Mehta et al., 2017b) to quantify the evolution of the PDO patterns during each 10-year hindcast period. The assumption was that the basic character of the PDO patterns is generally the same in simulation and hindcast experiments conducted with a particular ESM. The TAG and WPWP indices were calculated directly from the hindcast SSTs. These SST indices were calculated by averaging SST in the WPWP (20$^o$S to 20$^o$N, 90$^o$E to 180$^o$) for the WPWP index and in the tropical North (5$^o$ to 20$^o$N, 30$^o$ to 60$^o$W) and South (0$^o$ to 20$^o$S, 30$^o$W to 10$^o$E) Atlantic with the difference between the two for the TAG index.

Probabilities of transition of a DCV index from one phase to another phase (for example, from positive phase PDO$^+$ to negative phase PDO$^-$) were calculated by counting the number of times each phase transition occurred in a given seasonal or annual index time series and then by expressing the number as a percentage of the total number of data points in the index time series. The same approach was followed in calculating transition probabilities of simultaneous phases of more than one DCV phenomena (for example, from the (PDO$^+$, TAG$^+$) combination to the (PDO$^+$, TAG$^-$) combination). For the purpose of assessing hindcast skill of magnitudes of DCV indices, following the definitions of Niño3.4 phases (see, for example, Trenberth (1997)), we defined three states of each index – largest negative value to -0.5 times standard deviation (negative), -0.5 times standard deviation to +0.5 times standard deviation (neutral), and greater than +0.5 times standard deviation (positive). All index time series were normalized by subtracting the long-term average of annual cycles and dividing by the standard deviation of the time series before calculating states.

Following Smith et al. (2007), Keenlyside et al. (2008), and Pohlmann et al. (2009), we estimated decadal hindcast skill in the form of root-mean-square (RMS) hindcast errors, and correlation coefficients between hindcast and observed variables. The skill estimates were evaluated based on the ensemble-average, monthly average data from each ESM and also the data from the MME. Prior to calculating correlation coefficients, all data were detrended over the 1961–2010 CE period. The Monte Carlo technique (see, for example, Wilks (1995)) was used to estimate statistical significance of correlation coefficients. Correlation coefficients equal to or greater than 95% confidence limit are referred to as statistically significant in this paper. Also, negative correlation coefficients are referred to as no skill.

# 3. Results

## 3.1 Transition Probabilities

We begin the description of results with statistics of occurrence of each DCV phase and of combinations of phases of three DCV phenomena in observed and hindcast DCV indices. Then, observed and hindcast probabilities of transition between positive and negative phases of each DCV phenomenon, and among combinations of phases of three DCV phenomena are described.

The occurrence of each phase, as percent of total number of years, are shown in Table 2 for annual observed DCV indices from 1961 to 2010 CE. Occurrences of individual phases and combinations of phases in ensemble-average indices and the range (minimum to maximum within an ensemble) of occurrences within each ensemble of the four ESMs for the 1961 to 2010 CE period are also shown in Table 2. Please note that the phase occurrences in ensemble-average DCV indices are not the average of the occurrences in individual members of an ensemble. If it is assumed that both phases of a DCV index over a multidecadal period have equal probabilities of occurring, then the average occurrence of each phase would be 50% of the period. As Table 2 shows, the occurrence rate is almost 50% for the ERSST PDO, TAG, and WPWP indices, with small departures from the expected occurrence attributable perhaps to a relatively small sample size (50 years). Phase occurrences in three-month average index (December – January – February, DJF; March – April – May, MAM; June - July – August, JJA; September – October – November, SON) data are generally similar (not shown), except that the $WPWP^+$ and $WPWP^-$ phases occur 40% and 60% of the total years, respectively, in DJF; and the $TAG^+$ and $TAG^-$ phases occur 56% and 44% of the total years, respectively, in SON.

The corresponding occurrence rates for the ESM hindcast data in Table 2 show that while the PDO phase occurrence rates in the ensemble-average hindcast data from the CCSM4, CM2.1, and HadCM3 ESMs are generally similar, the MIROC5 ensemble-average results for PDO$^{+}$ and PDO$^{-}$ phases are 40% and 60%, respectively, for the annual, MAM, JJA, and SON data. The TAG$^{+}$ and TAG$^{-}$ occurrence rates are almost 60% and 40%, respectively, for the CM2.1 ensemble in DJF and SON data. The WPWP$^{+}$ and WPWP$^{-}$ occurrence rates in the CM2.1 ESM are 42% and 58%, respectively, in MAM and JJA averages. The WPWP$^{+}$ and WPWP$^{-}$ occurrence rates are 43% and 57%, respectively, in DJF in HadCM3; and 42% and 58%, respectively, in JJA in CCSM4. In the MME, the WPWP$^{+}$ and WPWP$^{-}$ occurrence rates are 40% and 60%, respectively, in the annual data. These results imply that ensemble hindcasts of the three DCV indices made with the four ESMs have generally comparable occurrence rates of the three indices with respect to the observed occurrence rates. Ranges of occurrence rates for each ESM's hindcast ensemble are also shown in Table 2. The ranges straddle the corresponding ensemble-averages in all except two cases (PDO$^{+}$ and PDO$^{-}$) in MIROC5 hindcasts. Also, there are no extraordinary outlier occurrence values. Thus, Table 2 shows that all four ESMs hindcast individual DCV phase occurrence rates reasonably accurately.

Some phase combinations of two or all three of the PDO, TAG, and WPWP indices are known to be associated with hydro-meteorological (see, for example, Schubert et al. (2004a, 2004b), Mehta et al. (2011b, 2016)) and agricultural (Mehta et al., 2012; 2017a) impacts in the U.S. Great Plains; impacts on hydro-meteorology, river flows, agriculture, inland water-borne transportation, and hydro-electricity generation in North America (Mehta, 2017); and worldwide impacts on hydro-meteorology, river flows, agriculture, fish captures, and other societal impacts

(Mehta, 2017). Therefore, it is important to estimate predictability of these phase combinations and their transitions to other combinations. There are eight such combinations (2 phases and 3 DCV indices; $2^3$=8) and the theoretical occurrence rate for each phase combination of the three DCV phenomena would be 12.5% if probabilities of all combinations were equal. These eight combinations are (PDO$^+$, TAG$^+$, WPWP$^+$), (PDO$^-$, TAG$^-$, WPWP$^-$), (PDO$^+$, TAG$^-$, WPWP$^+$), (PDO$^+$, TAG$^-$, WPWP$^-$), (PDO$^-$, TAG$^+$, WPWP$^+$), (PDO$^-$, TAG$^+$, WPWP$^-$), (PDO$^+$, TAG$^+$, WPWP$^-$), and (PDO$^-$, TAG$^-$, WPWP$^+$). In subsequent description of the simultaneous occurrence of two or more DCV phenomena, PDO, TAG, and WPWP are referred to as P, T, and W, respectively, with phases indicated by + or – sign as a superscript. Also, these three DCV indices are treated as independent since the simultaneous correlations among them are indistinuguishable from zero.

Table 2 shows that, in ERSST data, the (P$^+$, T$^+$, W$^+$), (P$^+$, T$^-$, W$^+$), and (P$^-$, T$^+$, W$^-$) combinations have much lower occurrence rates, whereas the (P$^-$, T$^+$, W$^+$), (P$^-$, T$^-$, W$^+$), and (P$^+$, T$^+$, W$^-$) combinations have a few percent higher occurrence rates. The occurrence rates for three-month average ERSST data are generally similar to the results for annual data shown in Table 2, excpt that the (P$^-$, T$^+$, W$^+$) and (P$^+$, T$^-$, W$^-$) combinations have much higher (25%) occurrence rates in SON. The corresponding occurrence rates for three-DCV combinations in the ESM hindcasts are also shown in Table 2. The CCSM4 ensemble-average hindcasts have both much above- and much below-average outliers; the (P$^-$, T$^-$, W$^+$) and (P$^+$, T$^+$, W$^-$) combinations have 26% and 20% occurrence rates, respectively, and the (P$^+$, T$^-$, W$^+$) and (P$^-$, T$^+$, W$^-$) combinations have 2% and 6% occurrence rates, respectively. It is interesting to note that the occurrence rates for the former two combinations are above average outliers in the ERSST data also, and the rates for the latter two combinations are below average outliers in the ERSST data also. Occurrence rates in other ESMs' hindcasts are generally close to the expected average rate. In the MME, however, the ensemble-average occurrence rates are substantially different from the rates in ERSST data. Ranges of occurrence rates within ensembles of hindcasts (Table 2) generally straddle the corresponding occurrence rates of ensemble-average hindcasts; hindcasts by each ESM, however, show a few DCV index combinations in which the occurrence rate from the ensemble-average hindcast lies outside the range of rates within that hindcast ensemble. Thus, the occurrence rates of individual and multiple DCV phases in ERSST observations and ensemble-average ESM hindcasts were found to be generally similar, establishing that the ESM hindcasts represent combinations of DCV phases reasonably well.

Next, the probabilities of transition from the phase in one year to either of the two possible phases of individual DCV indices in the next year in the observed and hindcast annual data were estimated and are shown in Figure 1. Ranges of within-ensemble transition probabilities in the ESM hindcasts are also shown in Figure 1 as vertical black bars, superimposed on each color bar, with horizontal black lines at minimum and maximum values. These ranges were calculated from individual ensemble members for each ESM and the MME. For the PDO phases (Figure 1a), the probabilities of transitions from $P^+$ to $P^+$ and to $P^-$ in the ERSST data are 72% and 27%, respectively. The transition probabilities from $P^-$ to $P^+$ and to $P^-$ are 20% and 80%, respectively. These results show an overwhelming tendency for same-phase transitions, or persistance, of PDO from one year to the next. Ensemble-average hindcasts by all ESMs and the MME generally show this tendency in Figure 1a. Even including the ranges of probabilities for each ESM in the comparison, the higher probabilities of same-phase transitions are clearly evident; the CM2.1 hindcast ranges, however, overlap. There are some seasonal variations in probabilities in the

ERSST, ESM, and MME data, with the same-phase PDO transitions most probable (approximately 80%) in June–July–August.

The transition probability of the TAG phases (Figure 1b) in ERSST annual data is largest (55%) for the $T^+$ to $T^+$ transition, but is considerably lower than the corresponding PDO same-phase transition. The $T^-$ to $T^-$ transition probability is even lower (53%). The opposite-phase transition probabilities are approximately 40-45%. Thus, TAG phases are less persistent than PDO phases in observed data and their transition probabilities are approximately equal, although same-phase transitions have higher probabilities. TAG phases in the four ESMs and the MME are more persistent as indicated by considerably larger same-phase transition probabilities for annual data in Figure 1b – 70% to 80% probabilities in CCSM4, CM2.1, and the MME, and 65% to 70% probabilities in the HadCM3 and MIROC5 ESMs even when their respective probability ranges are included. Consequently, opposite-phase transition probabilities are much lower in the individual ESM and MME hindcasts.

As for the PDO and TAG phases, same-phase transition probabilities of WPWP phases in the observed annual data (Figure 1c) are much higher (approximately 70%) compared to the opposite phase transition probabilities (approximately 30%). The same-phase transition probabilities in ensemble-average annual data from the four ESMs and the MME (Figure 1c) are at least as high as the probabilities in the observed data even when the within-ensemble ranges are included in the comparison. Consequently, opposite phase transition probabilities in the four ESMs and the MME are equal to or lower than those in the observed data. In the seasonal observed data, the probabilities of transition from any WPWP phase to any phase are approximately equal (approximately 50%) in MAM, JJA, and SON. In DJF, the same phase transition probabilities are 70 to 80% and the opposite phase probabilities are consequently approximately 20 to 30%. In the seasonal hindcast data from the ESMs, same phase transition probabilities are much higher than the opposite phase transition probabilities in all seasons unlike the probabilities in the observed data. Thus, Figure 1 shows that probabilities of same-phase transitions from one year to the next are considerably larger than opposite-phase transitions for PDO and WPWP phases in ERSST data and ensemble-average ESM and the MME hindcasts, except in the CM2.1 hindcasts where the differences among probabilities of PDO phase transitions are much smaller. Probabilities for TAG phases are almost the same in the ERSST data, but in the ensemble-average ESM and MME hindcasts the same-phase transition probabilities are much larger than the opposite-phase probabilities.

Next, we consider transition probabilities among combinations of phase of two DCV phenomena, the PDO and TAG variability. There are four possible combinations of phenomena and phases – $(P^+, T^+)$, $(P^-, T^-)$, $(P^+, T^-)$, and $(P^-, T^+)$ - and the theoretical transition probability for each transition would be 25% if the transitions occur randomly; that is, there would be equal probabilities of a transition to any of the four combinations. The actual transition probabilities of combined PDO and TAG phases are shown in Figure 2 as four color bars, one for each phase combination, for observed and ESM – including MME - data sets. Ranges of within-ensemble transition probabilities in the ESM hindcasts are also shown in Figure 2 as vertical black bars superimposed on each color bar with horizontal black lines at minimum and maximum values.

For the combination $(P^+, T^+)$, the calculated transition probability in the ERSST data (Figure 2a) is highest (45%) for transition to the same combination from one year to the next, followed by the transition to $(P^+, T^-)$ (30%). The probabilities are from approximately 7% to 15% for the other two combinations. Hindcasts with all ESMs, except HadCM3, and the MME appear to replicate the highest probability of the $(P^+, T^+)$ same-combination transition. The persistence of this combination from one year to the next is highest (approximately 75%) in the CCSM4 ESM, followed by the MME (approximately 65%). It is interesting to observe that the transition to ($P^+$, $T^-$) combination is much lower than in observed data in all ESMs except HadCM3, with zero probability in the MME ensemble-average hindcasts. Figures for seasonal data indicate (not shown) that the probability of ($P^+$, $T^+$) same-combination transition in observed data is considerably lower in DJF and SON, with the latter season having nearly equal probability of transition to any of the four possible combinations. Although all four ESMs generally have the highest probability of same-combination transition to ($P^+$, $T^+$) in all seasons, details vary among the models. HadCM3 is unique in that the transition probability of its ensemble-average hindcast to the ($P^-$, $T^-$) combination is nearly zero in all seasons and annual data.

In the case of ($P^+$, $T^-$) transitions (Figure 2b), the highest probabilities are for transitions to ($P^+$, $T^-$) and ($P^+$, $T^+$) combinations, both approximately 33%, in observed data. Probabilities for the other two transitions are 10 to 22%. Ensemble-average data from CCSM4 hindcasts nearly replicate the two highest probability transitions, but with somewhat higher (40%) probabilities. Ensemble-average data from CM2.1, HadCM3, and the MME show the highest probability of same-combination transition, but with almost twice as high a probability (60 to 65%) as the observed data. Both these ESMs and the MME show very low probabilities of other transitions from the ($P^+$, $T^-$) combination. Ensemble-average data from MIROC5 hindcasts show moderate probabilities of transitions to ($P^+$, $T^+$) and ($P^+$, $T^-$), and small to zero probabilities of transitions to the other two combinations. In DJF and SON, the same-combination transition probability from one year to the next is highest of all possible transitions in observed data. All ESMs generally have comparable probabilities of same-combination transition to ($P^+$, $T^-$), although transitions to the ($P^+$, $T^+$) combination also have moderate to high probabilities.

The ($P^-$, $T^+$) combination (Figure 2c) has the highest transition probabilities (approximately 35%) in the observed annual data for transitions to ($P^-$, $T^+$) and ($P^-$, $T^-$) combinations. Transitions to ($P^+$, $T^+$) and ($P^+$, $T^-$) combinations have approximately 15% probability. Hindcast annual data from the four ESMs and the MME show even higher probability of same-combination transition of ($P^-$, $T^+$) as the observed data. The next highest probability in the ESM and the MME data is of transition to ($P^-$, $T^-$) combination. Both in the observed data and ESM hindcast data, the third highest probability of ($P^-$, $T^+$) combination is for the ($P^+$, $T^+$) combination, followed by the probability of transition to the ($P^+$, $T^-$) combination. In the observed and hindcast seasonal data, the highest probability is of transition from ($P^-$, $T^+$) to the same combination.

Lastly, for the ($P^-$, $T^-$) combination (Figure 2d), the highest probability in observed (60%) and annual hindcast (40 to 80%) data is for transition to the same combination; the next highest probability is for transition to the ($P^-$, $T^+$) combination, except in the MME ensemble-average hindcast. This order of probabilities holds in observed seasonal data also. In the hindcast seasonal data from CM2.1, the transition to ($P^-$, $T^+$) has a higher probability than the same-combination probability in DJF, MAM, and JJA seasons.

Thus, a general tendency of all four combinations in the ERSST and ensemble-average ESM and MME indices to remain in the same combination is obvious, including when the ranges of ensemble member results are included, although there are cases in which probabilities are higher for transitions to other combinations (for example, ($P^+$, $T^+$) in CM2.1 and HadCM3). This general observation implies that ensemble-average results may be reliable enough for actual prediction of phase combinations at one to two years lead times. Details show, however, that there are very large ranges of transition probabilities for some combinations, pointing to the need for ensembles and ensemble averaging, including MME averaging, to increase the signal to noise ratio.

### 3.2 Skills of Phase and Magnitude Hindcasts

After comparing the occurrence statistics and transition probabilities of various phases of DCV indices and their combinations, we now describe skills of the four ESMs and the MME in predicting the phase and magnitudes of the PDO, TAG, and WPWP indices.

The percent of total numbers of years in which the ensemble-average ESM and the MME hindcasts accurately predicted the PDO phase is shown in Figure 3a. Since there are two possible phases to predict, the theoretical skill would be 50% if both phases are equally likely; that is, there would be an equal probability of predicting either phase. There would be skill if the actual probability exceeds 50%. The annual average hindcasts from the CCSM4, CM2.1, HadCM3, and MIROC5 ESMs, and the MME predicted the negative PDO phase correctly approximately 57%, 70%, 60%, 60%, and 65%, respectively, of the 28 years in which the PDO was in the negative phase. The four ESMs and the MME predicted the positive phase correctly approximately 45%, 56%, 60%, 40%, and 55%, respectively, of the 22 years in which the PDO was in the positive phase. These results imply that the ensemble-average hindcasts have skill above the nominal thershold in all but the CCSM4 and MIROC5 hindcasts of positive PDO phase. Figure 3a also shows that there are some members in each ensemble with higher skill than the skill of ensemble-average hindcasts. In seasonal hindcast data, the highest skill of prediction of both PDO phases is in DJF and SON, and the lowest skill in JJA. Even the lowest skill, however, is approximately 40%. The percent of total numbers of years in which TAG phases are predicted correctly by the ensemble-average hindcasts is shown in Figure 3b. Both TAG phases are predicted between approximately 50% and 65% of times correctly by ensemble-average hindcasts made by all four ESMs and the MME. Variations of skill in the seasonal data are also within this range for all four ESMs and the MME. In the case of WPWP phases, annual, ensemble-average hindcast data from all four ESMs predict both phases correctly between 40% and 62% of the times each phase occurs (Fig. 3c). Variations of skill in seasonal data are from 50% to 70%. Thus, Figure 3 indicates that ensemble-average hindcast data from all four ESMs and the MME show, with some exceptions, skills exceeding the nominal threshold in predicting phases of the three DCV phenomena. It must be re-iterated here that the ensemble-average data are from decadal hindcasts that are initialized once every ten years, so the prediction lead times vary from one year to ten years.

To evaluate the prediction skill for magnitudes of the DCV indices, three states of each index were defined as negative, neutral, and positive as described in Section 2.2. The percent of total years of each state in which each ESM and the MME correctly predicted the state is shown in Figure 4. The horizontal dashed line at 33.3% in each panel of Figure 4 shows the nominal skill threshold that would be expected if all three states were equally probable; probabilities above this threshold are considered significant skill in this study. The negative PDO state was predicted correctly in at least 42% of the 16 years in which it occurred, the neutral PDO state was predicted correctly at least 32% of the 21 years, and the positive PDO state was predicted correctly at least 15% of the 13 years as shown in Figure 4a. So, with ensemble-average hindcast data, CCSM4 shows skill above the threshold for the negative state, CM2.1 shows skill for negative and neutral states, HadCM3 shows skill for negative and positive states, MIROC5 shows significant skill for negative and neutral states, and the MME shows skill for all three states. Overall, the MME is the best for all three states, followed by the CM2.1, HadCM3, and MIROC5 ESMs over the 50 years of the hindcast period. Seasonal-average, ensemble-average data show that almost all four ESMs have the highest skill in predicting the negative state of the PDO. The seasonal data also show that, in SON, all four ESMs have prediction skill above the theoretical probability for all three PDO states.

Prediction skill for the three states of ensemble-average, annual hindcast of the TAG index is shown in Figure 4b. CCSM4 has skill above 33% only for the neutral state, and HadCM3 and MIROC5 have skill for the negative and positive states; hindcast data from CM2.1 do not show skill of any state. The MME shows skill for the neutral and positive states. From the seasonal-average hindcast data, all except HadCM3 in MAM and CCSM4 in SON show significant skill for prediction of the neutral TAG state. The skill for the other two states vary among the four ESMs and the MME in all four seasons. Thus, the overall skill of TAG prediction appears to be the best in the MME.

Skillful prediction of the three WPWP states is shown (Figure 4c) by ensemble-average annual data from all ESMs and the MME except that of the negative state by HadCM3 and the MME and of the neutral state by MIROC5. In the seasonal-average hindcast data, there is significant skill for all three states in all ESMs except the negative state in CCSM4 and MIROC5 in MAM; the negative state in CCSM4, HadCM3, and MIROC5 in JJA; and the neutral and positive states in HadCM3 and the positive state in MIROC5 in SON. Overall, the CCSM4 ensemble-average hindcasts of the three states appear to be the best, followed by the MME and MIROC5 hindcasts.

As mentioned in Section 2.1, the decadal hindcast experiments were initialized once (in the $0^{th}$ year - 1960, 1970, etc.) every ten years. The phase hindcast skills for the PDO, TAG, and the WPWP indices in the second year after initialization are described here following the description of the skills in the first year. For both the first and second years, we analyzed the accuracy of phase hindcast using data from annual-average and ensemble-average hindcasts as well as from all individual members of each ensemble. The results for both first and second years are shown in Table 3 for the PDO. The ensemble averages from the ESMs and the MME hindcast the PDO phase in the first year after initialization correctly in all five decades, except for CM2.1 in 1961.The second year phase hindcast by ensemble averages was correct for three ESMs (CCSM4, CM2.1, and MIROC5) and the MME in 1982. In other decades, however, fewer ensemble averages from individual ESMs hindcast the PDO phase correctly. The ensemble-average MME hindcast of the PDO phase in the second year was correct in 1982, 1992, and 2002. Table 3 also shows that first year phase hindcasts of the PDO index by individual members of each ensemble were correct for the largest number of members of CCSM4 ensembles in all five decades, followed by MIROC5 and the MME. It is obvious that the success rate or skill of phase prediction decreases from first year to second year for CCSM4, CM2.1, and HadCM3, but the second-year phase prediction skill of MIROC5 hindcasts is 100% in four of the five decades. It is also interesting to note that a correct hindcast of first-year PDO phase appears to be necessary for a correct hindcast of second-year phase, but it is not a sufficient condition.

As for the PDO index, MIROC5 performs better than the other three ESMs and the MME for the second year prediction of the TAG index also (Table 4) with correct phase prediction in four out of five decades. CCSM4, CM2.1, and the MME are next with three correct predictions of second-year TAG phase out of five decades, and HadCM3 has correct prediction of second-year phase in two out of five decades. Unlike for PDO predictions, however, a correct first-year prediction of the TAG phase does not appear to be a pre-requisite for a correct second-year phase prediction. Of the three DCV indices, first- and second-year hindcasts of the WPWP index are correct in the majority of the ESM-decade combinations (Table 5). In 1961, 1981, and 2001, ensemble-average WPWP index hindcasts by all four ESMs and the MME are correct for the first year after initialization. In 1962, 1992, and 2002, second-year phase hindcasts are also correctly made by ensemble-average WPWP indices by all four ESMs and the MME. It is also remarkable that when the first/second year phase of the WPWP index is correctly hindcast by the ESMs and the MME, almost all members of the corresponding ensembles also hindcast the phase correctly. This general success in hindcasting the WPWP index phase for two years is further addressed in the next Section where decade-average hindcast skills are described.

### 3.3 Decade-average Hindcast Skills

The next step in the journey to assess prediction skills of the PDO, TAG, and WPWP is the average skill over each decade of the decadal hindcast experiments, starting with the overall skill over the 1961 to 2010 CE period. Figure 5a shows correlation coefficients, using seasonal-average data, between the three observed and hindcast DCV indices over the 1961 to 2010 CE period. These coefficients were calculated with ensemble-average data from the four ESMs and the MME. In the cases of the PDO and the TAG indices, no one of the ESMs or the MME shows significant skill. The WPWP has small but substantial and significant skill in all ESMs except MIROC5, approaching 0.4 correlation coefficients.

Looking at the skill decade by decade after removing linear trends from the ERSST and ESM indices, Figure 5b shows that only PDO hindcasts by HadCM3 and the MME have substantial and significant skill in the 1980s CE. There is no significant hindcast skill of TAG (Fig. 5c) and WPWP (Fig. 5d) indices in any decade even though correlation coefficients are moderately large in some decades. Incidentally, the MIROC5 ESM's decadal hindcast data were used in a statistical prediction system for the PDSI in southern Africa (Mehta et al., 2014) because of the moderately large, but not statistically significant, skill of decadal hindcasts of the DCV indices with this ESM.

RMS errors (RMSEs) of the decadal hindcasts of the DCV indices, compared to the ERSST indices, are shown in Figure 6. Over the 1961 to 2000 CE period, WPWP index hindcasts with all ESMs and the MME have approximately the same RMSE (Fig. 6a). It is interesting to note that RMSE of PDO hindcasts (Fig. 6b) vary among decades and ESMs, but it is the smallest in all ESMs and the MME in the 1980s when PDO hindcast skills are the highest (Fig. 5b).

### 3.4 Roles of External Forcing and Internal Variability in Phase Transitions

As mentioned in Section 1.2, understanding and prediction of DCV phase transitions sustained for several months to an year or longer can be useful in understanding and prediction of DCV impacts. Understanding and prediction of DCV phases is also important for attribution of DCV phase transitions to internal ocean-atmosphere processes or changes in external forcings. Therefore, sustained transitions in phases of the PDO, and the TAG and WPWP SST variabilities in observed and ensemble-average hindcast indices of these DCV phenomena were visually identified. The phase transitions occurred over many months to 1 to 3 years and there is some subjectivity in the choice of selected transitions. The observed and hindcast phase transitions were also compared with major volcanic eruptions at low latitudes as represented in AOD time series andother publicly available information. The following questions were addressed to visually identify roles of external forcing and internal variability in DCV phase transitions.

- Are there phase transitions in observed and hindcast DCV indices which are physically consistent with external forcing changes as represented in AOD changes?
- Are there phase transitions in observed DCV indices which are also hindcast by the ESMs, but are not associated with AOD changes?

- Are there phase transitions in observed DCV indices which are in simulations and initialized hindcasts? Are they associated with AOD changes?
- What is the impact, if any, of initialization on phase transition events and on overall hindcasts?

In the following description of results, positive to negative phase transitions are referred to as PTN and negative to positive phase transitions are referred to as NTP.

### 3.4.1 Pacific Decadal Oscillation Phase Transitions

There were 14 PDO phase transitions between 1961 and 2010 CE in the ERSST data, with each phase persisting for many months to many years. Table 6 shows transitions in the observed PDO index; and in the ensemble-average, hindcast index in each of the four ESMs and the MME. Times (months and years) and locations of major (Volcanic Explosivity Index (VEI) ≥ 4; Newhall and Self (1982)), low-latitude volcanic eruptions are also mentioned in Table 6. As is evident, there are two types of phase transitions in the observed PDO index - transitions associated with internal ocean-atmosphere dynamics and those associated with AOD changes associated with volcanic eruptions. Three of the four major eruptions during the 1961 to 2010 CE period – Mount Agung in 1963 CE, Volcan de Fuego in 1974 CE, and Mount Pinatubo in 1991 CE – were associated with a phase transition in observed and hindcast PDO indices. The El Chichón eruption in Mexico, even though it was very explosive (VEI 5), was not associated with a phase transition in PDO hindcasts, but only with a phase transition in observed PDO index. It was, however, associated with phase transition in PDO simulations with the ESMs as discussed in Mehta et al. (2017b).

It is also evident in Table 6 that out of the 10 observed phase transitions not associated with a volcanic eruption, no ESM hindcast showed the correct phase transition in four such events (1961-62, 1988-90, 1995-97, and 2005 CE). The 1993-94 CE PTN transition is the only instance of hindcasts with all four ESMs and the MME showing the transition in the correct direction and magnitude without a volcanic eruption associated with it. Hindcasts with the CM2.1 show the correct phase transitions six times (4 NTP), CCSM4 three times (2 NTP), HadCM3 two times (both PTN), MIROC5 two times (1 each PTN and NTP), and MME three times (2 NTP). Thus, out of the 14 phase transitions in Table 6, the CM2.1 was successful in hindcasting 9, including 3 associated with volcanic eruptions; CCSM4 and MME in 6, including 2 in each associated with volcanic eruptions; and HadCM3 and MIROC5 in 5 phase transitions, including 3 in each associated with volcanic eruptions.

To gain further insight, the numbers of NTP and PTN phase transition events were identified from Table 6 and their possible attribution to external forcing or internal variability was identified. There are 6 events in the NTP and 8 events in the PTN category. Also, there is one major volcanic eruption during the former and three during the latter category. Thus, there are 5 other – "non-volcanic" – events in each category. The one NTP event during the 1991-92 CE Mount Pinatubo eruption was hindcast correctly by all four ESMs, but, surprisingly, not by the MME. Out of the 3 PTN transition events during volcanic eruptions, the correct hindcasts were 1 by CCSM4; and 2 each by the other 3 ESMs and the MME. Thus, in this relatively small sample size, almost all ESM hindcasts responded to AOD changes associated with volcanic eruptions. This result is very encouraging because, while it is well known that it is (almost) impossible to predict volcanic eruptions of any explosivity months to years in advance, the generally correct responses of the ESMs and the MME indicate that they can be used to predict post-eruption evolution of the ocean-atmosphere system reasonably accurately, at least qualitatively, for perhaps two to three or more years. Finally, there were four phase transition events from 1995-97 to 2006-07 CE; there were no major volcanic eruptions during this period. As Table 6 shows, not one of the ESMs or the MME hindcast these events correctly, with the 1997-99 CE event in HadCM3 being the lone exception to some extent. These four events occurred several years after the hindcasts were initialized in 1990 CE and 2000 CE for 10 years each, so it is reasonable to speculate that perhaps the initial condition effects were "forgotten" by the ESMs by the time these four phase transitions occurred.

Thus, as the foregoing shows, these ESMs were able to hindcast some of the PDO phase transitions caused by major volcanic eruptions and some caused by internal ocean-atmosphere dynamics. A comparison with PDO phase transitions in simulations with the same ESMs (Mehta et al., 2017b) shows that a correct response of the simulated PDO to a major volcanic eruption is not a pre-requisite for a successful hindcast of PDO phase transition after the same volcanic eruption. For example, only MIROC5 both simulated and hindcast the 1963 PDO phase transition in response to the Mount Agung (Bali), Indonesia, eruption. The other three ESMs and the MME did not simulate this phase change, but hindcast the change successfully. On the other hand, all except CCSM4 were able to simuate as well as hindcast the PDO phase change in response to the 1974-75 Volcan de Fuego, Guatemala, eruption. As mentioned earlier, all ESMs and the MME simulated the PDO phase transition in response to the 1981-82 El Chichón, Mexico, eruption, but no one of the five was able to hindcast the transition correctly. Other than this event, only MIROC5 was able to both simulate and hindcast the remaining three PDO phase transitions successfully.

From these results based on visual inspections, summary answers to the questions posed are: (1) There are 3 PDO phase transitions during the 1961 to 2010 CE period which are associated with AOD changes in both observed and hindcast indices in all ESMs and the MME, except for the 1974-75 PTN transition in CCSM4; (2) All ESMs' hindcasts capture phase transitions not associated with AOD changes in varying numbers, such correct transitions in an ESM's hindcast vary from two to six; (3) The 1963 CE and 1991-92 CE transitions associated with AOD changes due to volcanic eruptions are in simulations with all four ESMs and the MME also, but the sizes of the simulated changes vary among the ESMs and the MME (Mehta et al., 2017b); (4) The 1976-77 CE NTP transition is simulated by CM2.1, HadCM3, and CCSM4 to some extent, which suggests the intriguing possibility that perhaps coupled ocean-atmosphere response to the 1974-75 CE Volcan de Fuego volcanic eruption resulted in the 1976-77 CE NTP transition; this transition is present, but does not have the full range of PDO index, only in ensemble-average hindcasts by CM2.1 and the MME initialized in 1970 CE. Thus, initialization appears to have interfered with this NTP transition in HadCM3 and CCSM4 ESMs if indeed it was caused as a response to the 1974-75 CE volcanic eruption; and (5) a correctly simulated response to external forcing changes does not appear to be a pre-requisite for an ESM to successfully hindcast the PDO response to the same forcing change.

#### 3.4.2 Tropical Atlantic SST Gradient Phase Changes

There were 9 TAG phase transitions between 1961 and 2010 CE in the ERSST data, each of which persisted in positive or negative phase for many months to many years. Table 7 shows transitions in the observed TAG index; and in the ensemble-average, hindcast index in each of the four ESMs and the MME. Times (months and years) and locations of major low-latitude volcanic eruptions are also shown in Table 7. As is evident, there are two types of phase transitions in the observed TAG index - one group associated with internal ocean-atmosphere dynamics and the other associated with radiative forcings associated with volcanic eruptions. Three of the four major eruptions during the 1961 to 2010 CE period – Mount Agung in 1963 CE, Volcan de Fuego in 1974 CE, and El Chichón in 1982 CE – were associated with a positive (or approximately zero) to negative phase transition in observed TAG index. The Mount Pinatubo eruption in Phillipines in 1991 CE was associated with an NTP phase transition in observed TAG index. No one of the four

ESMs could hindcast these four TAG phase transitions correctly. It is also evident in Table 7 that no one of the remaining seven TAG phase transitions were correctly hindcast by any of the four ESMs. It is intriguing why no one of the 9 TAG phase transitions in the ERSST data are present in the ESM and MME hindcasts. On the other hand, as described in Mehta et al. (2017b), all ESMs and the MME correctly simulated some of the TAG phase changes associated with major volcanic eruptions. The 1963 TAG phase change was correctly simulated by CM2.1, HadCM3, CCSM4, and the MME; the 1974-75 TAG phase change was correctly simulated by CM2.1, MIROC5, and the MME; the 1981-82 phase change was correctly simulated by CM2.1 and MIROC5; and the 1991-92 TAG phase change was correctly simulated by HadCM3, CCSM4, and the MME. So, initialization appears to have interfered with TAG phase changes even when they were correctly simulated by an ESM. It is possible, as Swingedouw et al. (2015) found, that there is a multiyear to decade delayed response of some ESMs to Mount Agung-like eruptions on North Atlantic Ocean circulation and temperature. Possible effects of a delayed response of the TAG index to volcanic eruptions should be further investigated with controlled experiments with an ESM in simulation and hindcast modes.

From these results based on visual inspections, summary answers to the questions posed are: (1) There are no TAG phase transitions in hindcast data which are also in observed data, either associated with AOD changes or due to internal ocean-atmosphere interactions; (2) some of the TAG phase changes which are in observed data are simulated by some of the ESMs and the MME, but they are not hindcast by any ESM; and (3) initialization appears to have interfered with the ESMs' hindcasting the correct response to major volcanic eruptions.

### 3.4.3 West Pacific Warm Pool Variability Phase Transitions

There were nine phase transitions in the WPWP SST index from 1961 to 2010 CE in the ERSST data, with each phase persisting for many months to many years. Table 8 shows transitions in the observed WPWP index, and in the ensemble-average, hindcast index in each of the four ESMs and the MME. Times (months and years) and locations of major, low-latitude volcanic eruptions are also shown in Table 8. As in the cases of PDO and TAG phase transitions, there are two types of transitions in WPWP index; one group associated with internal ocean-atmosphere dynamics and the other associated with AOD changes associated with volcanic eruptions. There is a cooling trend from PTN phase associated with three volcanic eruptions – Mount Agung, Volcan de Fuego, and Mount Pinatubo – in the ERSST and hindcast indices. Out of the other six phase changes, the observed transitions in 1981-82 CE (PTN), 1993-95 CE (NTP), and 1994-96 CE (NTP) are hindcast, to some extent, by all four ESMs and the MME. The observed NTP transitions in 1967-68 CE and 1997-98 CE, and the PTN transition in 1996-97 CE are not hindcast by any of the ESM or the MME. Thus, out of the nine phase transitions, six are hindcast to some extent by all ESMs and the MME. A comparison with simulated responses of the WPWP index in these four ESMs (Mehta et al., 2017b) shows that the 1963, 1981-82, and 1991-92 phase changes associated with volcanic eruptions were correctly simulated by all ESMs and the MME. The 1974-75 WPWP phase change associated with the Volcan de Fuego, Guatemala, eruption was correctly simulated only by MIROC5 and HadCM3. It is also evident in Table 8 that out of the five phase transitions not associated with a volcanic eruption, all ESMs' and the MME's hindcasts showed the correct phase transition in two such events (1993-94 and 1994-96 CE); both of these were NTP transitions and both appeared as warming trends. The remaining three phase transitions (1967-68, 1996-97, and 1997-98 CE) were not hindcast correctly by any of the ESMs or the MME.

From these results based on visual inspections and a comparison with simulations by these four ESMs (Mehta et al., 2017b), summary answers to the questions posed are: (1) There are four phase transitions associated with AOD changes in the observed WPWP index which were generally correctly hindcast by the four ESMs and the MME; (2) all ESMs' ensemble-average hindcasts capture transitions not associated with AOD changes in 1993-94 and 1994-96 CE to varying degrees; (3) simulations with all four ESMs and the MME capture the 1963-64, 1981-82, and 1991-93 CE phase transitions associated with AOD changes in the WPWP SST index. Simulations with MIROC5 and HadCM3 capture the 1973-76 CE phase transition associated with AOD changes. Sizes of simulated transitions vary among the ESMs and the MME; (4) The impact of initialization appears to be reinforcement of the four transitions associated with AOD changes and correct hindcasts of two additional transitions not associated with AOD changes. The latter two, however, are also present in simulations with all four ESMs and the MME, so perhaps there is another radiative forcing (not AOD changes) driving these two transitions. It is also interesting to note that simulations show warming trend in the WPWP SST index continuing after 1996 CE which is not captured by any of the ESMs' hindcasts.

## 4. Summary and Discussion

We analyzed positive/negative phase occurrence rates, phase transition probabilities, and one-year and two-year phase and state predictability of the PDO, the TAG SST variability, and the WPWP SST variability in observations and ensembles of decadal hindcasts made with the CCSM4, CM2.1, HadCM3, and MIROC5 ESMs - and the MME formed from these ESM hindcasts - from 1961 to 2010 CE. The hindcasts were initialized every ten years. We also analyzed hindcast skills of these DCV phenomena over this 50 years period and in individual decades; and conducted case studies of their individual, sustained, phase transitions in the ensembles of decadal hindcasts in order to attribute the phase transitions to external forcing or initialized internal variability. Major results are:

- ⦿ Ensemble-average hindcasts of the three DCV indices made with the four ESMs and the MME have generally comparable phase occurrence rates with respect to observed rates.
- ⦿ There is a moderate to high probability (70%) of phase persistance or same-phase transitions of PDO and WPWP phases from one year to the next in observed data and also generally in the ensemble-average ESM hindcasts, whereas the same-phase transition probability of TAG phases is moderate (55%).
- ⦿ In observed data, out of the eight possible combinations of phases of the three DCV indices, the ($P^-$, $T^+$, $W^+$), ($P^-$, $T^-$, $W^+$), and ($P^+$, $T^+$, $W^-$) combinations have the highest occurrence rates, whereas ($P^+$, $T^+$, $W^+$), ($P^+$, $T^-$, $W^+$), and ($P^-$, $T^+$, $W^-$) combinations have the lowest occurrence rates; the other two combinations have intermediate occurrence rates.
- ⦿ There is a general tendency of all four combinations of PDO and TAG phases in the ERSST and ensemble-average ESM indices to remain in the same combination for at least two years, including when the ranges of ensemble member results are included, although there are cases in which probabilities are higher for transitions to other combinations (for example, ($P^+$, $T^+$) in CM2.1 and HadCM3).
- ⦿ Annual-average hindcasts from the four ESMs and the MME predicted the negative PDO phase correctly nearly 60% to 70% times of the 28 years in which the PDO was in the negative phase. These four ESMs and the MME predicted the positive phase correctly nearly 40% to 65% times of the 22 years in which the PDO was in the positive phase. Both TAG phases were predicted between approximately 50% and 65% times correctly by all four ESMs and the MME. In the case of WPWP phases, annual, ensemble-average hindcast data from all four ESMs predicted both phases correctly between 40% and 62% of the times each phase occured; the MME predicted negative and positive phases 45% and 65% of the times correctly. Thus, ensemble-average

hindcast data from all four ESMs and the MME show some skill in predicting phases of the three DCV phenomena above the 50% threshold if both phases were equally probable.

- The negative PDO state was hindcast correctly in at least 42% of the 16 years in which it occurred, the neutral PDO state was hindcast correctly at least 32% of the 21 years, and the positive PDO state was hindcast correctly at least 15% of the 13 years. So, with ensemble-average hindcast data, CCSM4 shows significant skill above the 33.3% threshold for the negative state, CM2.1 shows significant skill for negative and neutral states, HadCM3 shows significant skill for negative and positive states, MIROC5 shows significant skill for negative and neutral states, and the MME shows significant skill for all three states.
- For TAG states, CCSM4 has hindcast skill above the 33.3% threshold only for the neutral state, and HadCM3 and MIROC5 have skill for the negative and positive states; hindcast data from CM2.1 do not show hindcast skill of any state. The MME shows significant skill for neutral and positive TAG states.
- Skillful hindcast of all three WPWP states is shown by ensemble-average annual data from all ESMs except that of the negative state by HadCM3 and of the neutral state by MIROC5. The MME shows significant skill for neutral and positive WPWP states.
- Ensemble-average and most of ensemble members of MIROC5 hindcasts correctly predict PDO phases one and two years after initialization in all five decades. Prediction success rate decreases from the first year to the second in CCSM4, CM2.1, and HadCM3 hindcasts. Ensemble-average and most of ensemble members of the MME hindcasts correctly predict PDO phases one and two years after initialization after 1980; they correctly predict only the first-year PDO phase in 1960s and 1970s.

- Over the entire 1961 to 2010 CE period, no one of the four ESMs shows significant, 50-year average skill of PDO and TAG indices hindcasts. All individual ESMs except MIROC5, and the MME, show significant average skill of WPWP index hindcast over the 1961 to 2010 CE period.
- Decade-average hindcast skills of all three DCV indices vary from decade to decade, with only PDO index hindcasts by HadCM3 and the MME showing substantial and significant skill in the 1980s decade. There is no significant skill of TAG and WPWP indices hindcasts in any ESM or the MME in any of the five decades.
- Major, low-latitude volcanic eruptions - as represented in AOD changes - in 1963 (Mount Agung), 1974-75 (Volcan de Fuego), 1981-82 (El Chichón), and 1991-92 (Mount Pinatubo) are associated with sustained phase transitions of DCV indices in observed data and in some of the ensemble-average decadal hindcasts of the indices with the four ESMs and the MME. Three of the four major volcanic eruptions were associated with PDO phase changes in observed data and almost all hindcasts. The WPWP index phase changes associated with all four eruptions were hindcast by all ESMs and the MME. In contarst, no one of the 9 TAG phase transitions in observed data were present in the ESM and MME hindcasts. Hindcasts from some of the ESMs and the MME show approximately correct phase transitions in the absence of AOD changes also, implying that the initialization of the ESM hindcasts with observed data is beneficial in predicting phase transitions of DCV indices.

Before these results are discussed further, it must be mentioned that there are several shortcomings of these ESMs and decadal hindcast/forecast experiments conducted with them as mentioned in Section 1.2. Additionally, the four ESMs selected for the present study were initialized with different techniques and the decadal hindcasts were initialized every ten years. In spite of these and other shortcomings such as the inclusion of future volcanic eruptions in decadal hindcsts, the results of the analyses presented in Section 3 shed considerable light on prospects for future predictions of DCV indices and their usability for impacts prediction.

It is very encouraging that decadal hindcasts of the three DCV indices by the four ESMs and the MME have generally the same phase occurrence rates as the observed data. This similarity also carries over to probabilities of same-phase transitions of the PDO and WPWP indices from one year to the next in the observed and hindcast data. Another encouraging result is that there is some skill (above the 50% threshold) of annual-average PDO phase prediction in all four ESMs and the MME hindcasts. These results provide grounds for guarded optimism that there may be useable skill in phase prediction of the three DCV phenomena at least one year in advance and up to at least two years in advance for the PDO index. There is less confidence about magnitude prediction skill.

Although it is (almost) impossible to predict volcanic eruptions of any explosivity, it is instructive that AOD changes associated with major volcanic eruptions were included in the CMIP5 hindcast experiments. As the results show, the four ESMs and the MME appear to respond accurately to varying degrees to the eruption-associated AOD changes, and the hindcasts of the PDO and WPWP indices show phase transitions and subsequent evolutions of the DCV indices comparable to those in observed indices for several months to several years in some cases. Therefore, these hindcast results give encouragement for the use of these and other ESMs for multi-year prediction initialized soon after a major volcanic eruption occurs. As described earlier, AOD changes appear to cause damped oscillations in the DCV indices in some cases over several years, which might extend predictability of these indices beyond the immediate effects of AOD changes. These impacts of eruption-associated AOD changes on DCV indices imply that volcanic eruptions can influence global atmospheric dynamics and climate not only directly via interactions between ejected material in the atmosphere and short- and long-wave radiations, but also via influencing DCV phenomena's impacts on global climate.

Table 6 shows intriguing associations between PDO phase changes and volcanic eruptions. Positive to negative phase changes are associated with eruptions in 1963, 1974-75, and 1981-82 CE; but, a negative to positive phase change is associated with the Mount Pinatubo eruption in 1991-92 CE. The ejected material from a volcano can "shield" the underlying ocean or land surface if the material is ejected into the upper troposphere or stratosphere, reducing the incoming visible solar radiation and cooling the underlying surface. But, how can an eruption warm the tropical-subtropical central and eastern Pacific Ocean SSTs as is implied by the negative to positive PDO phase change? Based on the location of the eruption (Mount Pinatubo in Philippines), it can be hypothesized that the material ejected from the eruption can cool the WPWP, thereby decreasing the east-west SST difference in the tropical Pacific. This decreased SST difference can weaken easterly winds near the ocean surface, which, in turn, would reduce coastal and equatorial upwelling in eastern and equatorial central Pacific, respectively, and thereby warm central and eastern Pacific and change the PDO phase from negative to positive. This hypothesis can be and should be tested with ESM experiments.

The analyses presented in this paper are entirely of decadal hindcasts from 1961 to 2010 CE. But, as mentioned in Section 2.1, CMIP5 also has a set of 30-year hindcast/forecast experiments, the last of which was initialized with data from January 2006. How do these experiments perform with respect to observations since 2010 CE and what do they indicate about future evolutions of the DCV indices? All four ESMs and the MME perform poorly in hindcasting/forecasting the TAG index after 2010 CE. The best performance in the 2011 to 2015 period of verification by independent observed data is by MIROC5 for the PDO and the WPWP indices. Figure 7 shows the observed and hindcast evolutions of these two indices from 1961 to 2010 CE, the observed evolutions from 2011 to 2015 CE, and forecast evolutions from 2011 to 2020 CE; thus, there is a five-year overlap between independent observed data and forecast. In addition to ensemble-average hindcast/forecast indices, Figure 7 also shows the ± one standard deviation range of hindcasts/forecasts by ensemble members; as mentioned in Table 1, the MIROC5 hindcast/forecast ensembles have six members. Figure 7a shows that there is some similarity between observed and hindcast/forecast PDO indices from 2006 to 2015 CE, especially in the general shapes of the time series since 2011-2012 when the observed PDO index was within ± one standard deviation of forecast index. Figure 7b shows that there is a reasonable similarity between observed and hindcast/forecast WPWP indices from 2008-2009 to 2014 CE during which period the observed WPWP index was within ± one standard deviation of forecast index. Figures 7a and 7b also show a confirmation of the phase hindcast skill one and two years after initialization of MIROC5, especially since the 1970s, which was described and discussed in Section 3.2. This reasonably encouraging performance of MIROC5 in hindcasting the PDO and WPWP indices over the 1961 to 2010 CE period was the reason for using the MIROC5 data to hindcast decadal hydrologic cycles in seven countries of southern Africa by Mehta et al. (2014). It will be interesting to see if the PDO and WPWP indices indeed reached relative maxima in 2015-2016 CE, begin to decrease now, and reach relative minima in 2018-2019 CE as predicted by MIROC5. Such future evolutions of these indices would have very substantial, worldwide societal impacts as described by Mehta (2017).

The results presented in this paper indicate that the persistance and phase transition probability statistics of DCV indices and their predictability by the ESMs, and also perhaps long-term evolutions, can be exploited for prediction of these indices' possible impacts on hydro-meteorology, streamflows, agriculture, and other societal sectors. The importance and usefulness of such impacts predictions were mentioned in Section 1. Simulations of the three DCV phenomena with the same four ESMs and the MME, described in Mehta et al. (2017b), however, show that while these ESMs simulate the PDO's attributes (spatial pattern, annual cycle, and variability timescales) reasonably well, the ESMs only simulate the annual cycle and variability timescales of the WPWP SST variability reasonably well and the WPWP's spatial pattern is very poorly simulated by the ESMs and the MME. In the case of the TAG SST variability, simulation results show that while the spatial pattern simulation by the ESMs and the MME is approximately correct, the annual cycle and variability timescales are simulated very poorly. These incorrect simulations have serious implications not only for the prediction of impacts of these phenomena on global climate and society, but also about the simulation and prediction/projection of future climate change and its impacts. This is especially true about the WPWP since it is the largest heat source for driving global atmospheric circulations. Therefore, using the DCV indices' prediction from ESMs in statistical models to predict societal impacts may be a safer alternative, at least until the ESMs' simulation of these phenomena can be improved sufficiently to use climate and hydro-meteorological predictions/projections made by the ESMs directly as shown by Mehtal et al. (2014). Despite of these problems, the day may not be very far in the future when some aspects of DCV information are skillfully predicted and routinely used in agriculture and water resource managements, and other societal sectors.

*Acknowledgements* This research is supported by the U.S. Department of Agriculture-National Institute of Food and Agriculture under grant 2011-67003-30213 in the NSF – USDA – DOE Earth System Modelling Program, the NASA – Physical Oceanography Program under grant NNX15AD18A, and the U.S. Army Corps of Engineers – Institute for Water Resources under

Contract W912HQ-15-P-0056. We acknowledge the World Climate Research Programme's Working Group on Coupled Modelling, which is responsible for CMIP, and we thank the climate modeling groups (listed in Table 1 of this paper) for producing and making available their model output. For CMIP the U.S. Department of Energy's Program for Climate Model Diagnosis and Intercomparison provides coordinating support and led development of software infrastructure in partnership with the Global Organization for Earth System Science Portals.

## References

Ammann, C., G. Meehl, W. Washington, and C. Zender (2003), A monthly and latitudinally varying forcing data set in simulations of 20th century climate, Geophys. Res. Lett., 30(12), 1657, doi:10.1029/2003GL016875.

Carrington, R. C. (1863), Observations of the Spots on the Sun, 1–248 pp., Williams and Norgate, London, U. K.

Chambers, F. (1886), Sunspots and prices of Indian food-grains, Nature, 34, 100–104.

Currie, R. G. (1974), Solar cycle signal in surface air temperature, J. Geophys. Res., 79, 5657–5660.

Currie, R. G., and R. W. Fairbridge (1985), Periodic 18.6-year and cyclic 11-year induced drought and flood in northeastern China and some global implications, Quatern. Sci. Rev., 4, 109–134.

Currie, R. G., T. Wyatt, and D. P. O'Brien (1993), Deterministic signals in European fish catches, wine harvests, and sea-level, and further experiments, Int. J. Climatol., 13, 665–687, doi:10.1002/joc.3370130607.

Daggupati., P., D. Deb, R. Srinivasan, D. Yeganantham, V. M. Mehta, and N. J.

Rosenberg (2016), Spatial calibration of hydrology and crop yields through parameter regionalization for a large river basin. *Journal of the American Water Resources Association*, 52, 648 - 666.

Doblas-Reyes, F. J., and co-authors (2013), Initialized near-term regional climate change prediction, Nature Commun., 4, 1715, doi:10.1038/ncomms2704.

Fernandez, M., P. Huang, B. McCarl, and V.M. Mehta (2016), Value of decadal climate variability information for agriculture in the Missouri River Basin. *Climatic Change*, in press, DOI 10.1007/s10584-016-1807-x.

Garnett, R., N. Nirupama, C. E. Haque, and T. S. Murty (2006), Correlates of Canadian prairie summer rainfall: Implications for crop yields, Climate Res., 32, 25–33.

Ham, Y.-G., M. M. Rienecker, M. J. Suarez, Y. Vikhliaev, B. Zhao, J. Marshak, G. Vernieres, and S. D. Schubert (2014), Decadal prediction skill in the GEOS-5 forecast system, Climate Dynamics, 42, 1–20.

Hansen, J. E., et al. (2002), Climate forcing in Goddard Institute for Space Studies SI2000 simulations, J. Geophys. Res., 107(D18), 4347, doi:10.1029/2001JD001143.

Harrison, V. L. (1976), Do Sunspot Cycles Affect Crop Yields?, Agriculture Econ. Rep., 327, 1–23 pp., U.S. Dep. Agriculture,Washington, D. C.

Hastenrath, S. (1990), Decadal-scale changes of the circulation in the tropical Atlantic sector associated with Sahel drought, Int. J. Climatol., 10, 459–472.

Hazeleger, W., and co-authors (2013), Predicting multiyear North Atlantic Ocean variability, J. Geophys. Res., 118, 1087–1098, doi:10.1002/jgrc.20117.

Herschel, W. (1801), Observations tending to investigate the nature of the Sun, in order to find the causes or symptoms of its variable emission of light and heat; with remarks on the use that may possibly be drawn from solar observations, Phil. Trans. R. Soc. London, 91, 265–318.

Houghton, R. W., and Y. M. Tourre (1992), Characteristics of low frequency sea surface temperature fluctuations in the tropical Atlantic, J. Climate, 5, 765–771.

IPCC, 2013: Summary for Policymakers. In: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Stocker, T.F., D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex and P.M. Midgley (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Jevons, W. S. (1879), Sunspots and commercial crises, Nature, 19, 588–590.

Keenlyside, N., M. Latif, J. Jungclaus, L. Kornblueh, E. Roeckner (2008), Advancing decadal-scale climate prediction in the North Atlantic sector, Nature, 453, 84-88.

Kim, H.-M., P. J. Webster, and J. A. Curry (2012), Evaluation of short-term climate change prediction in multi-model CMIP5 decadal hindcasts, Geophys. Res. Lett., 39, L10701, doi:10.1029/2012GL051644.

King, J. W., E. Hurst, A. J. Slater, P. A. Smith, and B. Tamkin (1974), Agriculture and sunspots, Nature, 252, 2–3.

Kirtman, B., and co-authors (2013), Near-term Climate Change: Projections and Predictability. In: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Stocker, T.F., D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex and P.M Midgley (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Krishnamurti, T. N., C. M. Kishtawal, Z. Zhang, T. LaRow, D. Bachiochi, E. Williford, S. Gadgil, and S. Surendran (2000), Multimodel ensemble forecasts for weather and seasonal climate, J. Climate, 13**,** 4196–4216.

Love, J.J. (2013), On the insignificance of Herschel's sunspot correlation, Geophys. Res. Lett., 40, 4171 – 4176, doi: 10.1002/grl.50846.

Mantua, N.J., S.R. Hare, Y. Zhang, J.M. Wallace, and R.C. Francis (1997), A Pacific interdecadal climate oscillation with impacts on salmon production, Bulletin of the American Meteorological Society, 78, 1069 – 1079.

McPhaden, M.J., A.J. Busalacchi, and D.L.T. Anderson, 2010: A TOGA retrospective, Oceanography, 23, 86 – 103.

Meadows, A. J. (1975), A hundred years of controversy over sunspots and weather, Nature, 256, 95–97.

Meehl, G. A., L. Goddard, J. Murphy, R. J. Stouffer, G. Boer, G. Danabasoglu, K. Dixon, M. A. Giorgetta, A. Greene, E. Hawkins, G. Hegerl, D. Karoly, N. Keenlyside, M. Kimoto, B. Kirtman, A. Navarra, R. Pulwarty, D. Smith, D. Stammer, and T. Stockdale (2009), Decadal Prediction: Can it be skillful? Bulletin of the American Meteorological Society, 90**,** 1467, doi: 10.1175/2009BAMS2778.1.

Meehl, G.A., and co-authors (2014), Decadal climate prediction: An update from the trenches, Bulletin of the American Meteorological Society, 243 – 267.

Meehl, G.A., and H. Teng (2012), Case studies for initialized decadal hindcasts and predictions for the Pacific region, Geophys. Res. Lett., 39, L22705, doi:10.1029/2012GL053423.

Meehl, G.A., and H. Teng (2014), CMIP5 multi-model initialized decadal hindcasts for the mid-1970s shift and early-2000s hiatus and predictions for 2016–2035, Geophys. Res. Lett., 10.1002/2014GL059256.

Mehta, V.M., and T. Delworth (1995), Decadal variability of the tropical Atlantic Ocean surface temperature in shipboard measurements and in a global ocean-atmosphere model, J. Climate, 8, 172-190.

Mehta, V.M., and K.-M. Lau (1997), Influence of solar irradiance on the Indian monsoon-ENSO relationship at decadal-multidecadal time scales, Geophys. Res. Lett., 24, 159-162.

Mehta, V.M. (1998), Variability of the tropical ocean surface temperatures at decadal-multidecadal timescales, Part I: The Atlantic Ocean, J. Climate, 11, 2351-2375.

Mehta, V.M., G. Meehl, L. Goddard, J. Knight, A. Kumar, M. Latif, T. Lee, A. Rosati , and D. Stammer (2011a), Decadal Climate Predictability and Prediction: Where Are We? Bull. Amer. Meteorol. Soc., 92, 637-640.

Mehta, V.M., N. J. Rosenberg, and K. Mendoza (2011b), Simulated Impacts of Three Decadal Climate Variability Phenomena on Water Yields in the Missouri River Basin, Journal of the American Water Resources Association, 47, 126-135.

Mehta, V.M., N. J. Rosenberg, and K. Mendoza (2012), Simulated Impacts of Three Decadal Climate Variability Phenomena on Dryland Corn and Wheat Yields in the Missouri River Basin, Agricultural and Forest Meteorology, 152, 109-124.

Mehta, V.M., C. L. Knutson, N. J. Rosenberg, J. R. Olsen, N. A. Wall, T. K. Bernadt, and M. J. Hayes (2013a), Decadal Climate Information Needs of Stakeholders for Decision

Support in Water and Agriculture Production Sectors: A Case Study in the Missouri River Basin, Weather, Climate, and Society, 5, 27-42.

Mehta, V.M., H. Wang, and K. Mendoza (2013b), Decadal predictability of tropical basin-average and global-average sea-surface temperatures in CMIP5 experiments with the HadCM3, GFDL-CM2.1, NCAR-CCSM4, and MIROC5 global earth system models, Geophysical Research Letters, 40, doi:10.1002/grl.50236.

Mehta, V.M., H. Wang, K. Mendoza, and N.J. Rosenberg (2014), Predictability and Prediction of Decadal Hydrologic Cycles: A Case Study in Southern Africa, Weather and Climate Extremes, 3, 47-53.

Mehta, V.M., K. Mendoza, P. Daggupati, R. Srinivasan, N. J. Rosenberg, and D. Deb (2016), High-resolution Simulations of Decadal Climate Variability Impacts on Water Yield in the Missouri River Basin with the Soil and Water Assessment Tool (SWAT), *J. Hydrometeorology*, **17**, 2455 - 2476.

Mehta, V.M., K. Mendoza, P. Daggupati, R. Srinivasan, and N. J. Rosenberg (2017a), High-resolution Simulations of Decadal Climate Variability Impacts on Dryland Spring and Winter Wheat Yields in the Missouri River Basin with the Soil and Water Assessment Tool (SWAT), Agricultural and Forest Meteorology, in review.

Mehta, V.M., H. Wang, and K. Mendoza (2017b), Simulation of Three Natural Decadal Climate Variability Phenomena in CMIP5 Experiments with the UKMO-HadCM3, GFDL-CM2.1, NCAR-CCSM4, and MIROC5 Global Earth System Models, Climate Dynamics, in review.

Mehta, V.M., 2017: *Decadal Climate Variability: Societal Impacts.* CRC Press (Taylor & Francis), 325 pp.

Newhall, C. G., and S. Self (1982), The Volcanic Explosivity Index (VEI): An estimate of explosive magnitude for historical volcanism, J. Geophys. Res., 87, 1231–1238.

Pohlmann, H., J. H. Jungclaus, A. Kohl, D. Stammer, and J. Marotzke (2009), Initializing decadal climate predictions with the GECCO oceanic synthesis: Effects on the North Atlantic, J. Climate, 22, 3926–3938.

Power, S., T. Casey, C. Folland, A. Colman, and V.M. Mehta (1999), Interdecadal modulation of the impact of ENSO on Australia, Climate Dynamics, 15, 319-324.

Poynting, J. H. (1884), A comparison of the fluctuations in the price of wheat and in the cotton and silk imports into Great Britain, J. Stat. Soc. London, 47, 34–74.

Proctor, R. A. (1880), Sun-spots and financial panics, Scribner's Monthly, 20, 170–178.

Pustil'nik, L. A., and G. Yom Din (2004a), Influence of solar activity on the state of the wheat market in medieval Europe, Solar Phys., 223, 335–356.

Pustil'nik, L. A., and G. Yom Din (2004b), Space climate manifestations in Earth prices – from medieval England up to modern U.S.A., Solar Phys., 224, 473–481.

Pustil'nik, L. A., and G. Yom Din (2009), Possible space weather influence on the Earth wheat markets, Sun Geosphere, 4, 35–41.

Pustil'nik, L. A., and G. Yom Din (2013), On possible influence of space weather on agricultural markets: Necessary conditions and probable scenarios, Astrophys. Bull., 68, 107–124.

Rajagopalan, B., Y. Kushnir, and Y.M. Tourre (1998), Observed decadal mid-latitude and tropical Atlantic climate variability, Geophys. Res. Lett., 25, 367 – 370.

Reynolds, R. W., N. A. Rayner, T. M. Smith, D. C. Stokes, and W. Wang (2002), An improved in situ and satellite SST analysis for climate, J. Climate, 15, 1609–1625.

Sato, M., J. Hansen, M.P. McCormick, and J. Pollack (1993), Stratospheric aerosol optical depth, 1850-1990, J. Geophys. Res., 98, 22,987-22,994.

Schubert, S.D., M.J. Suarez, P.J. Pegion, R.D. Koster, and J.T. Bacmeister (2004a), On the cause of the 1930s Dust Bowl, Science, 303, 1855–1859.

Schubert, S.D., M.J. Suarez, P.J. Pegion, R.D. Koster, and J.T. Bacmeister (2004b), Causes of long-term drought in the US Great Plains, J. Climate, 17, 485–503.

Schwabe, H. (1844), Sonnen-Beobachtungen im Jahre 1843, Astron. Nachr., 21, 233 – 236.

Smith, D., S. Cusack, A. Colman, A. Folland, G. Harris, J. Murphy (2007), Improved surface temperature prediction for the coming decade from a global circulation model, Science, 317, 796-799.

Stenchikov, G., K. Hamilton, R.J. Stouffer, A. Robock, V. Ramaswamy, B. Santer, and H.-F. Graf (2006), Arctic Oscillation response to volcanic eruptions in the IPCC AR4 climate models, J. Geophys. Res., 111, D07107, doi: 10.1029/2005JD006286.

Swingedouw, D., P. Ortega, J. Mignot, E. Guilyardi, V. Masson-Delmotte, P. G. Butler, M. Khodri, and R. Séférian (2015), Bidecadal North Atlantic ocean circulation variability controlled by timing of volcanic eruptions**.** Nature Communications, 6, DOI: 10.1038/ncomms7545.

Tatebe, H. and coauthors (2012), Initialization of the climate model MIROC for decadal prediction with hydographic data assimilation. JMSJ Special issue on the recent development on climate models and future climate projections. JMSJ Special Issue on Recent Development on Climate Models and Future Climate Projections, 90A, 275-294.

Taylor, K. E., R. J. Stouffer, and G. A. Meehl (2012), An overview of CMIP5 and the experiment design, Bulletin of the American Meteorological Society, 93, 485–498.

Trenberth, K. E. (1997), The definition of El Niño, Bull. Amer. Meteorol. Soc., 78, 2771–2777.

van Oldenborgh, G., F. Doblas Reyes, B. Wouters, and W. Hazeleger (2012), Decadal prediction skill in a multi-model ensemble, Climate Dynamics, 38, 1263–1280.

Vines, R. G. (1977), Possible relationships between rainfall, crop yields, and the sunspot cycle, J. Austral. Inst. Agric. Sci., 43, 3–13.

Wang, H., and V.M. Mehta (2008), Decadal variability of the Indo-Pacific Warm Pool and its association with atmospheric and oceanic variability in the NCEP–NCAR and SODA reanalyses, J. Climate, 21, 5545-5565.

Wilks, D. S. (1995), *Statistical Methods in the Atmospheric Sciences*, Academic Press, 467 pp.

Yang, X., and co-authors (2012), A predictable AMO-like pattern in the GFDL fully coupled ensemble initialization and decadal forecasting system, J. Climate, 26, 650 – 661.

Yeager, S., A. Karspeck, G. Danabasoglu, J. Tribbia, and H. Teng (2012), A decadal prediction case study: Late twentieth-century North Atlantic Ocean heat content, J. Clim., 25, 5173–5189, doi:10.1175/JCLI-D-11-00595.1.

Zhang, S., M. J. Harrison, A. Rosati, and A. Wittenberg (2007), System design and evaluation of coupled ensemble data assimilation for global oceanic climate studies, Mon. Weather Rev., 135, 3541–3564, doi:10.1175/MWR3466.1.

## Figure Captions

**Figure 1:** Probabilities of transitions among phases of (a) the Pacific Decadal Oscillation,(b) the tropical Atlantic SST gradient variability, and (c) the West Pacific Warm Pool SST variability from 1961 to 2010 in ERSST data, and in decadal hindcasts made with CCSM4, CM2.1, HadCM3, and MIROC5 Earth System Models, and the Multi-Model Ensemble (MME). For the model data, color bars show probabilities derived from ensemble-average data and black bars show the range of probability derived from ensemble members. Please refer to the text for more details.

**Figure 2:** Probabilities of transitions among combined phases of the Pacific Decadal Oscillation (PDO) and the tropical Atlantic SST gradient (TAG) variability from 1961 to 2010 in ERSST data, and in decadal hindcasts made with CCSM4, CM2.1, HadCM3, and MIROC5 Earth System Models, and the Multi-Model Ensemble (MME). (a) PDO$^+$, TAG$^+$; (b) PDO$^+$, TAG$^-$; ( c) PDO$^-$, TAG$^+$; and (d) PDO$^-$, TAG$^-$. For the model data, color bars show probabilities derived from ensemble-average data and black bars show the range of probability derived from ensemble members. See text for more details.

**Figure 3:** Probabilities of correct prediction of phases of the Pacific Decadal Oscillation (PDO), the tropical Atlantic SST gradient (TAG) variability, and the West Pacific Warm Pool (WPWP) SST variability from 1961 to 2010 in ERSST data, and in decadal hindcasts made with CCSM4, CM2.1, HadCM3, and MIROC5 Earth System Models, and the Multi-Model Ensemble (MME). (a) PDO, (b) TAG, (c) WPWP. For the model data, color bars show probabilities derived from ensemble-average data and black bars show the range of probability derived from ensemble members. The numbers of years in positive and negative phases of each index are given above each box. See text for more details.

**Figure 4:** Probabilities of correct prediction of states of the Pacific Decadal Oscillation (PDO), the tropical Atlantic SST gradient (TAG) variability, and the West Pacific Warm Pool (WPWP) SST variability from 1961 to 2010 in ERSST data, and in decadal hindcasts made with CCSM4, CM2.1, HadCM3, and MIROC5 Earth System Models, and the Multi-Model Ensemble (MME). For the model data, color bars show probabilities derived from ensemble-average data and black bars show the range of probability derived from ensemble members. (a) PDO, (b) TAG, (c) WPWP. See text for more details.

**Figure 5:** Correlation coefficients between ERSST and hindcast indices of the Pacific Decadal Oscillation (PDO), the tropical Atlantic SST gradient (TAG) variability, and the West Pacific Warm Pool (WPWP) SST variability from 1961 to 2010 in decadal hindcasts made with CCSM4, CM2.1, HadCM3, and MIROC5 Earth System Models, and the Multi-Model Ensemble (MME). Color bars show correlation coefficients derived from ensemble-average data and black bars show the range of coefficients derived from ensemble members. (a) 1961 to 2010, (b) PDO, (c) TAG, and (d) WPWP. See text for more details.

**Figure 6:** Root-mean-square error (RMSE) between ERSST and hindcast indices of the Pacific Decadal Oscillation (PDO), the tropical Atlantic SST gradient (TAG) variability, and the West Pacific Warm Pool (WPWP) SST variability from 1961 to 2010 in decadal hindcasts made with CCSM4, CM2.1, HadCM3, and MIROC5 Earth System Models, and the Multi-Model Ensemble (MME). Color bars show RMSE derived from ensemble-average data and black bars show the range of RMSE derived from ensemble members. (a) 1960 to 2010, (b) PDO, (c) TAG, and (d) WPWP. See text for more details.

**Figure 7:** Observed (black line, 1961 to 2015), hindcast (red line, 1961 to 2010), and forecast (blue line, 2011 to 2020) indices of the Pacific Decadal Oscillation (PDO) and the West Pacific Warm Pool (WPWP) sea-surface temperature. The observed indices are from the ERSST data, and the ensemble-average hindcast and forecast indices are from the MIROC5 Earth System

Model. Cross hatching shows the ± one standard deviation range of hindcasts and forecast members of each ensemble. Vertical dashed lines show when each decadal hindcast ensemble was initialized; the forecast ensemble was initialized in January 2006. (a) PDO, and (b) WPWP. See text for more details.

Figure 1 Click here to download Figure Fig1_final.png

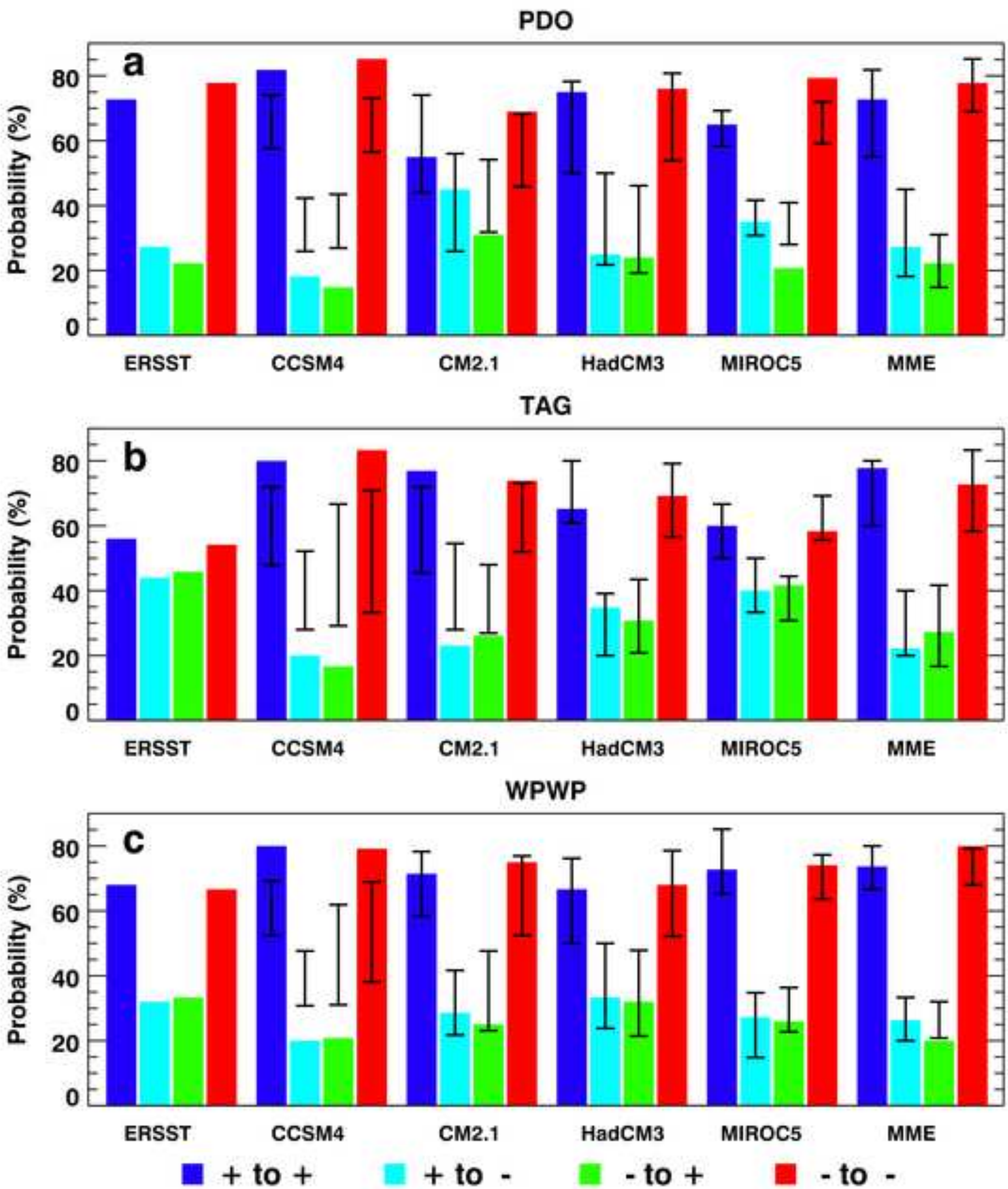

Figure 2

Click here to download Figure Fig2_final.png

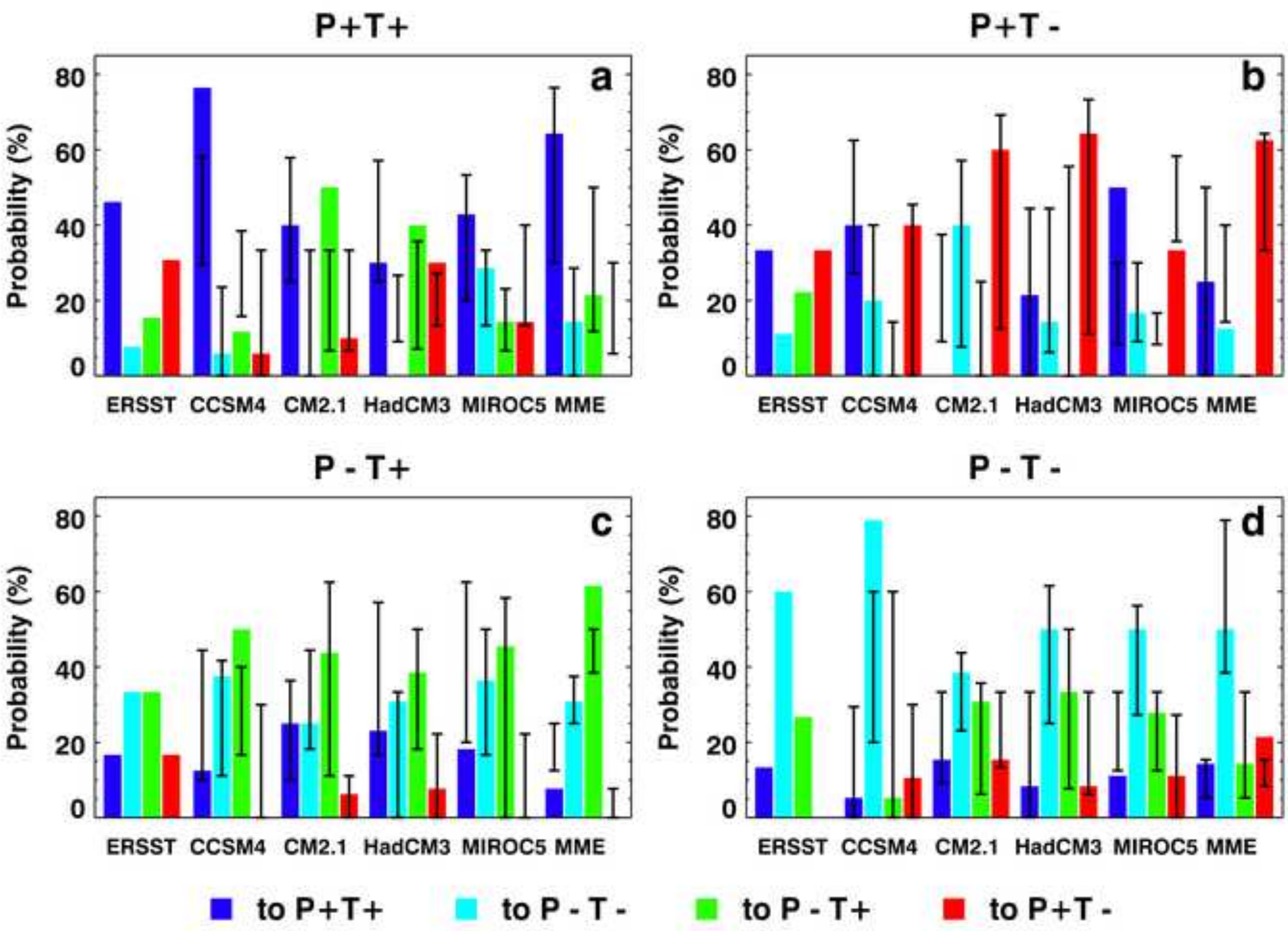

Figure 3



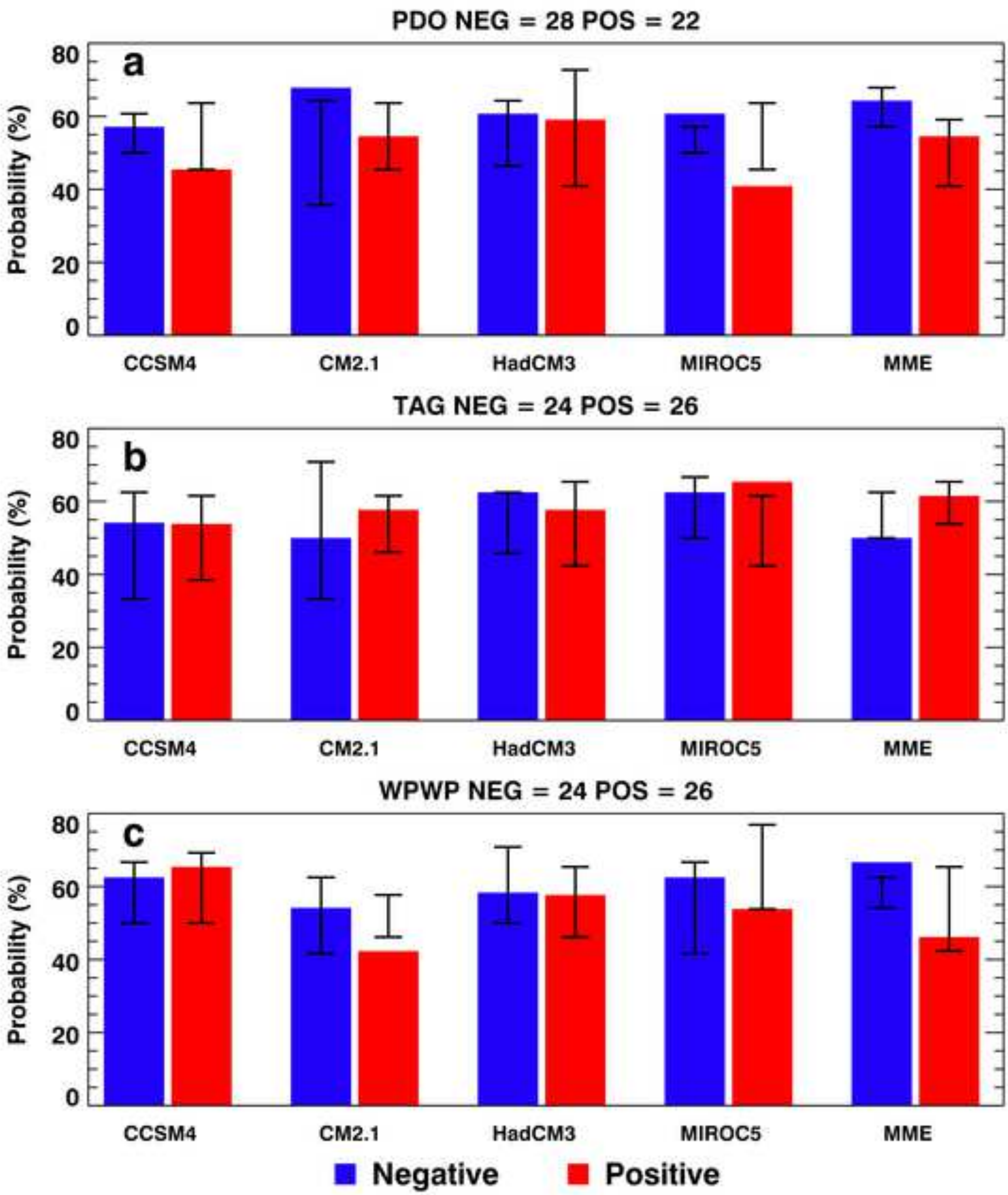
PDO NEG = 28 POS = 22
a
TAG NEG = 24 POS = 26
b
WPWP NEG = 24 POS = 26
c
Probability (%)
0
20
40
60
80
CCSM4
CM2.1
HadCM3
MIROC5
MME
Negative
Positive

Figure 4



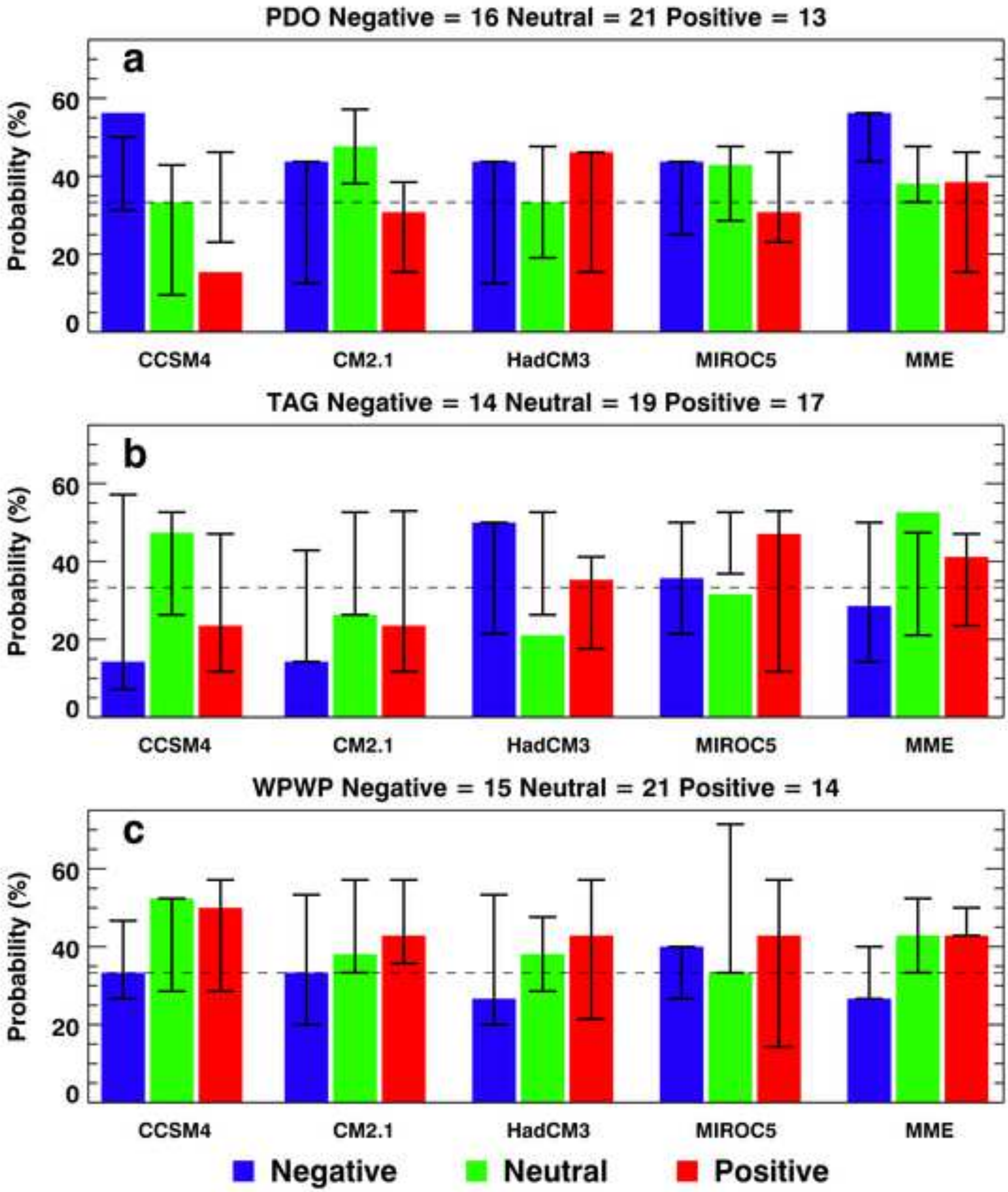
PDO Negative = 16 Neutral = 21 Positive = 13
a
Probability (%)
0
20
40
60
CCSM4
CM2.1
HadCM3
MIROC5
MME
TAG Negative = 14 Neutral = 19 Positive = 17
b
Probability (%)
0
20
40
60
CCSM4
CM2.1
HadCM3
MIROC5
MME
WPWP Negative = 15 Neutral = 21 Positive = 14
c
Probability (%)
0
20
40
60
CCSM4
CM2.1
HadCM3
MIROC5
MME
Negative
Neutral
Positive

Figure 5

Click here to download Figure Fig5_final.png

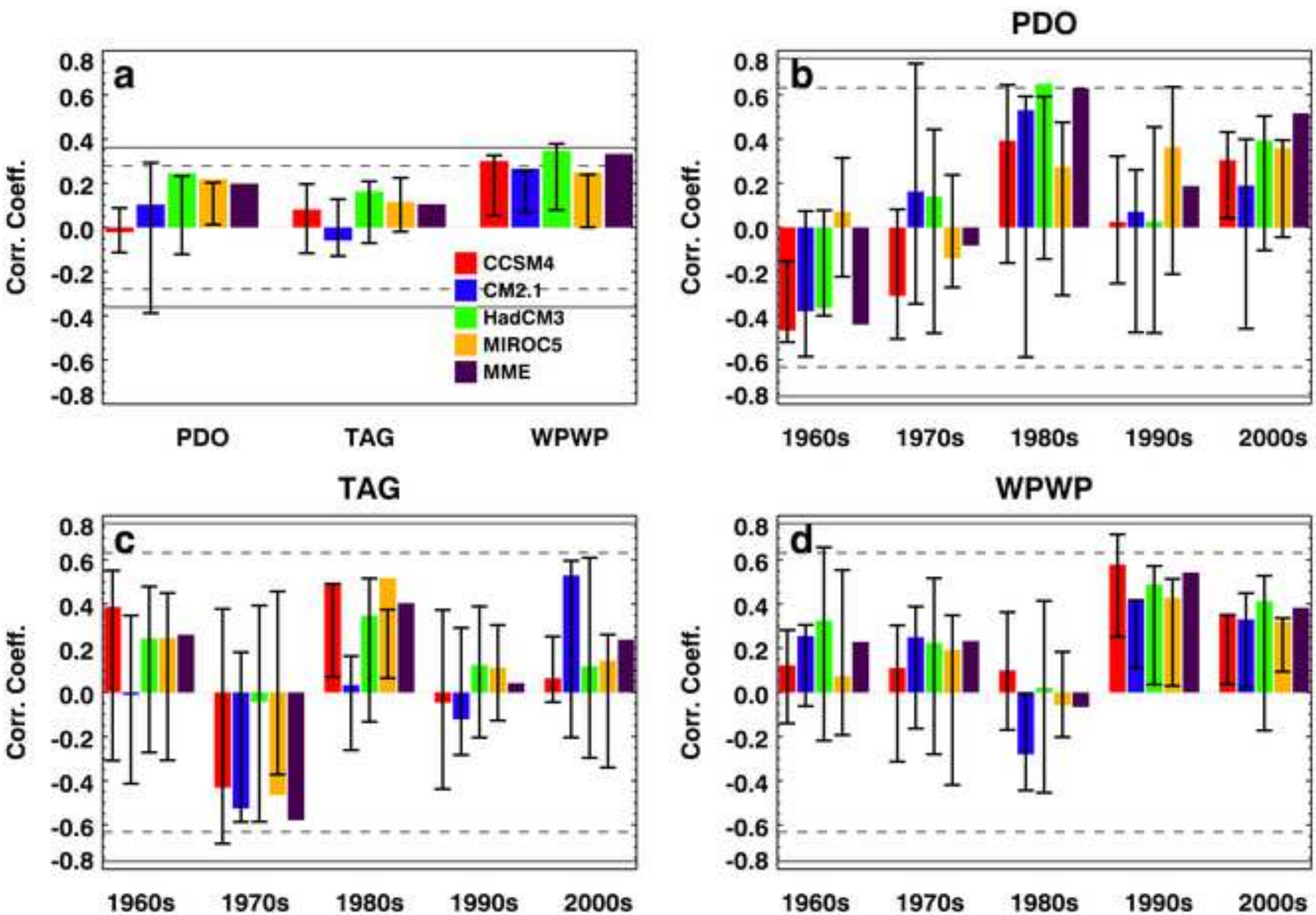

Figure 6

Click here to download Figure Fig6_final.png

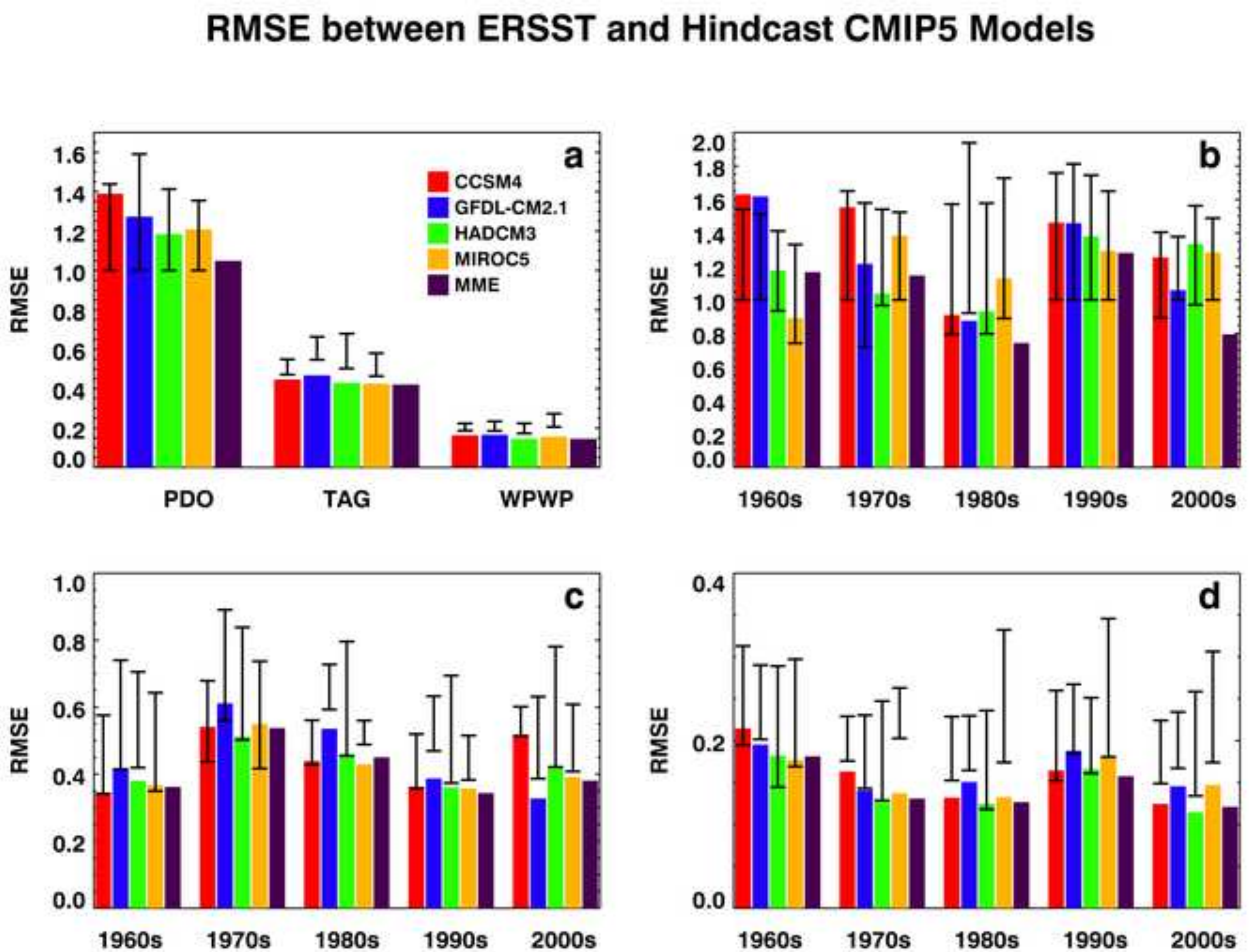

Figure 7

Click here to download Figure Fig7_final.png

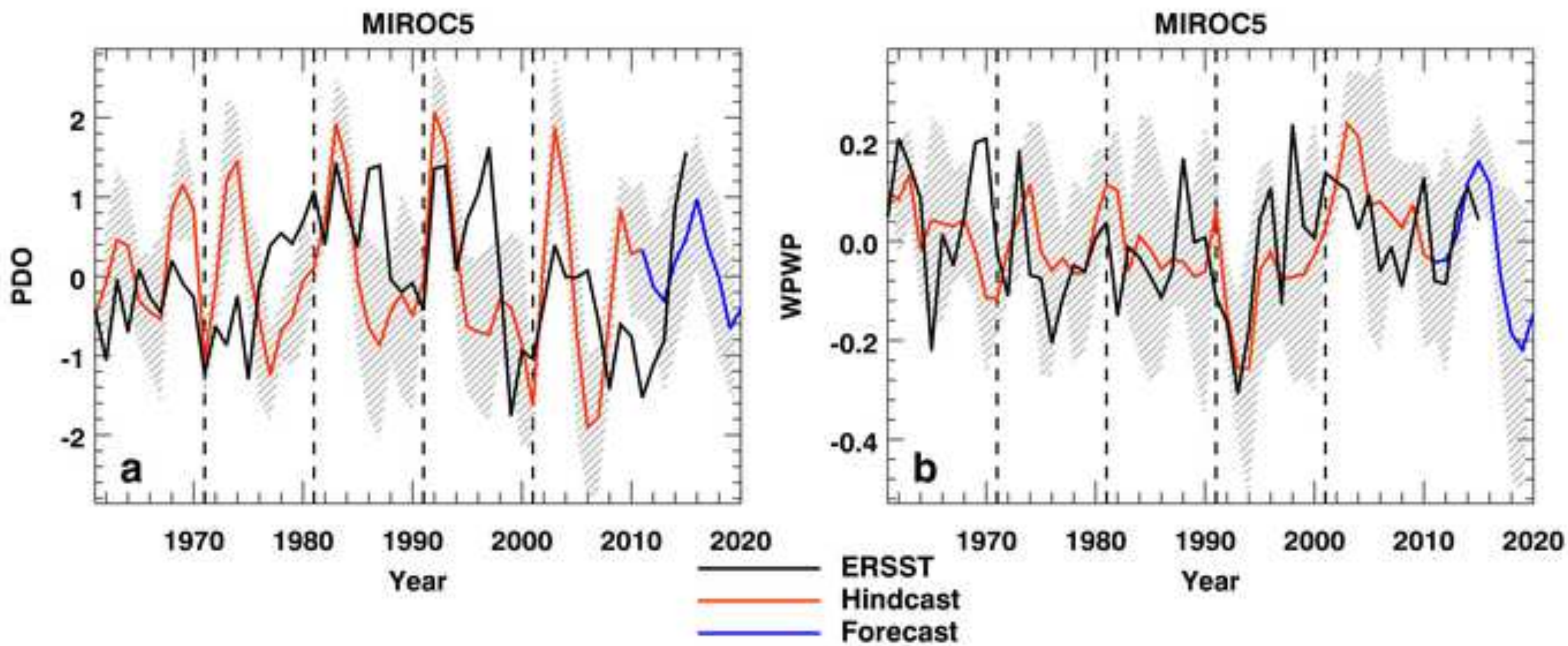

**Table 1**: CMIP5 hindcast experiments with Earth System Models used in this study.

| Model | Institute | Experiment | Ensemble members | SST resolution |
|---|---|---|---|---|
| CM2.1 | NOAA Geophysical Fluid Dynamics Laboratory, U.S.A. | Decadal hindcast (1960, 1970, 1980, 1990, 2000) | 10 | $1^o$ (lon.) × $0.34^o$ (lat.) at Eq., and $1^o$ (lat.) at $28^o$ and poleward |
| HadCM3 | Hadley Centre, U.K. | Decadal hindcast (1060, 1970, 1980, 1990, 2000) | 10 | $1.25^o$ × $1.25^o$ |
| MIROC5 | Atmosphere and Ocean Research Institute (Univ. of Tokyo), National Institute for Environmental Studies, and Japan Agency for Marine-Earth Science and Technology, Japan | Decadal hindcast (1960, 1970, 1980, 1990, 2000) | 6 | Rotated pole grid ~ $1.41^o$ (lon.) × $0.79^o$ (lat.) |
| CCSM4 | National Center for Atmospheric Research, U.S.A. | Decadal hindcast (1960, 1970, 1980, 1990, 2000) | 10 | $1.25^o$ × $1.25^o$ |

Table 2

**Table 2:** Occurrences (% of total number of years) of individual and combination phases of decadal climate variability indices from 1961 to 2010 CE in hindcasts with individual Earth System Models and the Multi-Model Ensemble.

| **DCV Phases** | **ERSST** | **CCSM4** | | **CM2.1** | | **HadCM3** | | **MIROC5** | | **MME** | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Ens.-ave. | Member range | Ens.-ave. | Member range | Ens.-ave. | Member range | Ens.-ave. | Member range | Ens.-ave. | Member range |
| PDO$^{+}$ | 44 | 44 | 44 - 54 | 42 | 42 - 56 | 48 | 46 - 62 | 40 | 44 - 54 | 44 | 40 - 48 |
| PDO$^{-}$ | 56 | 56 | 46 - 56 | 58 | 44 - 58 | 52 | 38 - 54 | 60 | 46 - 56 | 56 | 52 - 60 |
| TAG$^{+}$ | 52 | 50 | 44 - 64 | 54 | 44 - 58 | 48 | 44 - 54 | 52 | 44 - 56 | 56 | 48 - 54 |
| TAG$^{-}$ | 48 | 50 | 36 - 56 | 46 | 42 - 56 | 52 | 46 - 56 | 48 | 44 - 56 | 44 | 46 - 52 |
| WPWP$^{+}$ | 52 | 52 | 42 - 58 | 44 | 44 - 58 | 50 | 40 - 54 | 46 | 46 - 58 | 40 | 44 - 52 |
| WPWP$^{-}$ | 48 | 48 | 42 - 58 | 56 | 42 - 56 | 50 | 46 - 60 | 54 | 42 - 54 | 60 | 48 - 56 |
| $P^{+}T^{+}W^{+}$ | 8 | 14 | 8 - 24 | 14 | 14 - 26 | 8 | 10 - 26 | 14 | 12 - 24 | 12 | 8 - 14 |
| $P^{-}T^{-}W^{-}$ | 12 | 14 | 8 - 22 | 16 | 10 - 24 | 10 | 10 - 26 | 24 | 14 - 22 | 18 | 10 - 24 |
| $P^{-}T^{+}W^{+}$ | 20 | 10 | 2 - 16 | 8 | 4 - 14 | 16 | 4 - 16 | 12 | 4 - 8 | 12 | 8 - 16 |
| $P^{+}T^{-}W^{-}$ | 12 | 8 | 6 - 14 | 8 | 4 - 16 | 16 | 6 - 22 | 4 | 4 - 12 | 10 | 4 - 16 |
| $P^{-}T^{-}W^{+}$ | 18 | 26 | 10 - 24 | 10 | 4 - 18 | 14 | 6 - 18 | 12 | 8 - 20 | 10 | 10 - 26 |
| $P^{+}T^{+}W^{-}$ | 18 | 20 | 10 - 18 | 8 | 6 - 16 | 12 | 8 - 14 | 14 | 6 - 10 | 16 | 8 - 20 |
| $P^{+}T^{-}W^{+}$ | 6 | 2 | 4 - 16 | 12 | 6 - 22 | 12 | 4 - 20 | 8 | 10 - 20 | 6 | 2 - 12 |
| $P^{-}T^{+}W^{-}$ | 6 | 6 | 6 - 16 | 24 | 4 - 18 | 12 | 6 - 20 | 12 | 8 - 22 | 16 | 6 - 24 |

**Table 3:** One- and two-year phase prediction skill in decadal hindcasts of the Pacific Decadal Oscillation (PDO) in each decade from 1961 to 2010. In parentheses after the Earth System Model (ESM) name are shown the number of ensemble members for each ESM. The phase of the observed PDO index (-/+) in first and second year of each decade is shown in parentheses after each year. Bold numbers denote correct phase prediction by the annual-average, ensemble-average hindcast by each ESM and the Multi-Model Ensemble (MME), and the numbers denote how many members of each ensemble also hindcast the phase correctly.

| **Earth System Model (Ensemble members)** | **1961 - 1970** | | **1971 - 1980** | | **1981 - 1990** | | **1991 - 2000** | | **2001 - 2010** | |
|---|---|---|---|---|---|---|---|---|---|---|
| | 1961(-) | 1962(-) | 1971(-) | 1972(-) | 1981(+) | 1982(+) | 1991(-) | 1992(+) | 2001(-) | 2002(-) |
| CCSM4 (10) | **10** | 0 | **9** | 0 | **9** | **10** | **10** | **7** | **9** | **8** |
| CM2.1 (10) | 3 | 0 | **5** | **6** | **8** | **8** | **5** | 7 | **4** | 3 |
| HadCM3 (10) | **3** | 2 | **1** | 5 | **5** | 6 | **6** | 3 | **10** | **7** |
| MIROC5 (6) | **4** | **5** | **6** | **3** | **4** | **5** | **3** | **6** | **6** | 2 |
| MME (36) | **20** | 7 | **21** | 14 | **26** | **29** | **24** | **23** | **29** | **20** |

**Table 4:** One- and two-year phase prediction skill in decadal hindcasts of the tropical Atlantic sea-surface temperature gradient (TAG) index in each decade from 1961 to 2010. In parentheses after the Earth System Model (ESM) name are shown the number of ensemble members for each ESM. The phase of the observed TAG index (-/+) in first and second year of each decade is shown in parentheses after each year. Bold numbers denote correct phase prediction by the annual-average, ensemble-average hindcast by each ESM and the Multi-Model Ensemble (MME), and the numbers denote how many members of each ensemble also hindcast the phase correctly.

| **ESM (Ensemble members)** | **1961 - 1970** | | **1971 - 1980** | | **1981 - 1990** | | **1991 - 2000** | | **2001 - 2010** | |
|---|---|---|---|---|---|---|---|---|---|---|
| | 1961(+) | 1962(+) | 1971(-) | 1972(-) | 1981(+) | 1982(+) | 1991(-) | 1992(+) | 2001(-) | 2002(-) |
| CCSM4 (10) | 4 | **10** | 0 | 5 | **10** | 4 | 0 | **6** | **10** | **9** |
| CM2.1 (10) | **7** | **10** | 1 | 4 | 2 | 2 | **6** | **7** | **7** | **6** |
| HadCM3 (10) | **8** | **7** | 1 | 4 | **6** | 2 | 1 | 5 | **10** | **5** |
| MIROC5 (6) | **4** | **6** | **4** | **3** | **5** | 0 | 2 | **4** | **6** | **4** |
| MME (36) | **23** | **33** | 6 | 16 | **23** | 8 | 9 | **22** | **33** | **24** |

**Table 5:** One- and two-year phase prediction skill in decadal hindcasts of the West Pacific Warm Pool (WPWP) sea-surface temperature index in each decade from 1961 to 2010. In parentheses after the Earth System Model (ESM) name are shown the number of ensemble members for each ESM. The phase of the observed WPWP index (-/+) in first and second year of each decade is shown in parentheses after each year. Bold numbers denote correct phase prediction by the annual-average, ensemble-average hindcast by each ESM and the Multi-Model Ensemble (MME), and the numbers denote how many members of each ensemble also hindcast the phase correctly. Observed and hindcast WPWP indices were detrended before calculation of prediction skill.

| **ESM (Ensemble members)** | **1961 - 1970** | | **1971 - 1980** | | **1981 - 1990** | | **1991 - 2000** | | **2001 - 2010** | |
|---|---|---|---|---|---|---|---|---|---|---|
| | 1961(+) | 1962(+) | 1971(-) | 1972(-) | 1981(+) | 1982(-) | 1991(-) | 1992(-) | 2001(+) | 2002(+) |
| CCSM4 (10) | **10** | **10** | 2 | 0 | **9** | 2 | **10** | **10** | **7** | **10** |
| CM2.1 (10) | **10** | **10** | **8** | 0 | **10** | 1 | 6 | **10** | **8** | **10** |
| HadCM3 (10) | **10** | **10** | **9** | 3 | **9** | **5** | 7 | **10** | **7** | **8** |
| MIROC5 (6) | **5** | **5** | **6** | **3** | **6** | 1 | 2 | **6** | **5** | **6** |
| MME (36) | **35** | **35** | **25** | 6 | **34** | 9 | **25** | **36** | **27** | **34** |

**Table 6:** Phase Transitions in Ensemble-average Decadal Hindcasts, 1961 – 2010 CE. The Pacific Decadal Oscillation

| Transition Years in ERSST | Transition States | CCSM4 Hindcasts | CM2.1 Hindcasts | HadCM3 Hindcasts | MIROC5 Hindcasts | MME Hindcasts | Volcanic activity |
|---|---|---|---|---|---|---|---|
| 1961-62 | +0.75 to -1.75 | Negative to positive | Negative to positive | Negative to positive | Negative; no change | Negative to positive | - |
| 1963 Feb - May | +0.5 to -0.5 | Positive; coincide. decrease | Positive; coincide. decrease | Positive; coincide. decrease | Positive; coincide. decrease | Positive; coincide. decrease | Mount Agung, Bali; VEI 5 |
| 1964-65 | -1.5 to +1.0 | Similar to observed | Similar to observed | Positive | Positive to negative | Positive | - |
| 1974-75 Oct – Dec 1974 | +0.75 to -2.0 | Positive | Coincide. decrease | Coincide. decrease | Coincide. decrease | Coincide. decrease | Volcan de Fuego, Guatemala; VEI 4 |
| 1976-77 | -1.4 to +1.0 | Positive; no change | Negative to positive | Positive; no change | Negative; no change | Small negative to small positive | - |
| 1981-82 Mar – Apr 1982 | +0.75 to -0.25 | Decrease | Increase | Increase | Increase | Increase | El Chichón, Mexico; VEI 5 |
| 1982-83 | -0.25 to +2.0 | Similar to observed | Similar to observed | Positive | Similar to observed | Similar to observed | - |
| 1988-90 | +1.75 to -1.5 | Negative; no change | Negative; no change | Positive; no change | Negative; no change | Negative; no change | - |
| 1991-92 Jun 1991 | -1.8 to +2.2 | Delayed negative to positive | Delayed negative to positive | Delayed negative to positive | Negative to positive | Negative; no change | Mount Pinatubo, Philippines; VEI 6 |
| 1993-94 | +2.0 to -1.5 | Similar to observed | Similar to observed | Similar to observed | Similar to observed | Similar to observed | - |
| 1995-97 | -1.5 to +2.8 | Negative | Negative | Negative | Negative | Negative | - |
| 1997-99 | +2.5 to -2.2 | Negative; no change | Negative; no change | Delayed small positive to small negative | Negative; no change | Negative; no change | - |
| 2005 | -1.5 to +0.5 | Negative | Positive | Negative | Negative | Negative | - |
| 2006-07 | +0.4 to -2.0 | Negative | Similar to observed | Negative | Negative | Negative | - |

**Table 7:** Multiyear Phase Transitions in Ensemble-average Decadal Hindcasts, 1961 - 2010: The Tropical Atlantic Sea-surface Temperature Gradient Variability

| Transition Years in ERSST | Transition States | Hindcast in CCSM4 | Hindcast in GFDL CM2.1 | Hindcast in HadCM3 | Hindcast in MIROC5 | Hindcast in MME | Volcanic or other forcing activity |
|---|---|---|---|---|---|---|---|
| 1963 | +0.2 to -0.2 | Positive | Positive | Positive | Positive | Positive | Mount Agung, Bali |
| 1968-69 | -0.3 to +0.6 | Small negative to small positive | Negative | Small negative to small positive | Small negative to small positive | Small negative to small positive | |
| 1971-72 | -0.7 to +0.5 | Slow trend from positive towards negative | Slow trend from positive towards negative | Slow trend from positive towards negative | Negative | Slow trend from positive towards negative | |
| 1974 | 0 to -0.5 | Indifferent | Positive to negative trend | Indifferent | Indifferent | Indifferent | Volcan de Fuego, Guatemala |
| 1982 | +0.6 to -0.6 | Increasing trend | Increasing trend | Increasing trend | Increasing trend | Increasing trend | El Chichón, Mexico |
| 1983-84 | +0.8 to -0.8 | Negative | Negative | Negative | Delayed positive to negative | Negative | |
| 1991-92 | -0.5 to +0.6 | Positive | Positive | Positive | Positive | Positive | Mount Pinatubo, Philippines |
| 1992-94 | +0.6 to -0.8 | Fluctuating around zero | Fluctuating around zero | Fluctuating around zero | Fluctuating around zero | Fluctuating around zero | |
| 2003-04 | -0.5 to +1.0 | Negative | Negative | Negative | Small negative to small positive | Negative | |

**Table 8:** Phase Transitions in Ensemble-average Decadal Hindcasts, 1961 – 2010 CE. The West Pacific Warm Pool Sea-surface Temperature Variability

| Transition Years in ERSST | Transition States | CCSM4 Hindcasts | CM2.1 Hindcasts | HadCM3 Hindcasts | MIROC5 Hindcasts | MME Hindcasts | Volcanic activity |
|---|---|---|---|---|---|---|---|
| 1963-64 Feb - May | +0.2 to -0.2 | Small positive to -0.2 | Small positive to -0.2 | Small positive to -0.2 | Small positive to -0.2 | Small positive to -0.2 | Mount Agung, Bali; VEI 5 |
| 1967-68 | -0.1 to +0.25 | Negative; fluctuating | Negative; fluctuating | Negative; fluctuating | Positive; fluctuating | Negative; fluctuating | |
| 1973-76 Oct – Dec 1974 | +0.25 to -0.3 | Small positive to negative | Small positive to negative | Small positive to negative | Small positive to negative | Small positive to negative | Volcan de Fuego, Guatemala; VEI 4 |
| 1981-82 | +0.1 to -0.35 | Slow downward trend from positive to negative | Slow downward trend from positive to negative | Slow downward trend from positive to negative | Slow downward trend from positive to negative | Slow downward trend from positive to negative | El Chichón, Mexico; VEI 5 |
| 1991-93 June 1991 | >0 to -0.5 | >=0 to negative; fluctuating | >=0 to negative; fluctuating | >=0 to negative; fluctuating | >=0 to negative; fluctuating | >=0 to negative; fluctuating | Mount Pinatubo, Philippines; VEI 6 |
| 1993-94 | -0.5 to 0.2 | Warming trend | Warming trend | Warming trend | Warming trend | Warming trend | |
| 1994-96 | -0.2 to +0.2 | Warming trend | Warming trend | Warming trend | Warming trend | Warming trend | |
| 1996-97 | +0.3 to -0.35 | Steady around zero | Steady around zero | Steady around zero | Steady around zero | Steady around zero | |
| 1997-98 | -0.35 to +0.4 | Steady around zero | Steady around zero | Steady around zero | Steady around zero | Steady around zero | |