Understanding the Control of Extratropical Atmospheric Variability on ENSO Using a Coupled Data Assimilation Approach[1]

Feiyu Lu[2]

Nelson Institute Center for Climatic Research and the Department of Atmospheric and Oceanic Sciences, University of Wisconsin-Madison, Madison, Wisconsin, USA

Zhengyu Liu[2]

Nelson Institute Center for Climatic Research and the Department of Atmospheric and Oceanic Sciences, University of Wisconsin-Madison, Madison, Wisconsin, USA and Laboratory for Climate, Ocean and Atmosphere Studies, Peking University, Beijing, China

Yun Liu

Department of Atmospheric and Oceanic Science and Earth System Science Interdisciplinary Center, University of Maryland, College Park, Maryland, USA

Shaoqing Zhang

Geophysical Fluid Dynamics Laboratory, NOAA, Princeton, USA

Robert Jacob

Mathematics and Computer Science Division, Argonne National Laboratory, Argonne, Illinois, USA

Prepared for Climate Dynamics

Revision date: April 21, 2016

[1] Center for Climatic Research Contribution Number

[2] Corresponding author address: Center for Climatic Research, 1225 W. Dayton St., Madison, Wisconsin 53706, USA, E-mail: flu7@wisc.edu, zliu3@wisc.edu

Abstract

The control of extratropical atmospheric variability on ENSO variability is studied in a coupled general circulation model (CGCM) utilizing an ensemble-based coupled data assimilation (CDA) method in the perfect-model framework. Assimilation is limited to the desired model components (e.g. atmosphere) and spatial areas (e.g. the extratropics) to study the ensemble-mean model response (e.g. tropical response to “observed” extratropical atmospheric variability). The CDA provides continuously "corrected" extratropical atmospheric forcing and boundary conditions for the tropics and the use of ensemble optimizes the observational forcing signal over internal variability in the model component or region without assimilation.

The experiments demonstrate significant control of extratropical atmospheric forcing on ENSO variability in the CGCM. When atmospheric “observations” are assimilated only poleward of 20° in both hemispheres, most ENSO events in the “observation” are reproduced and the error of the Nino3.4 index is reduced by over 40% compared to the ensemble control experiment that does not assimilate any observations. Further experiments with the assimilation in each hemisphere show that the forced ENSO variability is contributed roughly equally and independently by the Southern and Northern Hemisphere extratropical atmosphere. Further analyses of the ENSO events in the southern hemisphere forcing experiment reveal robust precursors in both the extratropical atmosphere over southeastern Pacific and equatorial Pacific thermocline, consistent with previous studies of the South Pacific Meridional Mode and the discharge-recharge paradigm, respectively. However, composite analyses based on each precursor show that neither precursor alone is sufficient to trigger ENSO onset by itself and therefore neither alone could serve as a reliable predictor. Additional experiments with northern hemisphere forcing, ocean assimilation or different latitudes are also performed.

1. Introduction

The dynamics of the El Nino-Southern Oscillation (ENSO) phenomenon has been studied extensively in the past 50 years. It has been recognized that the genesis of ENSO events depends critically on coupled ocean-atmosphere dynamic processes in the tropical Pacific (e.g. Philander 1990; Neelin et al. 1998). In the meantime, more recent studies suggest that ENSO could also be triggered by extratropical atmospheric variability from the North Pacific through the "seasonal footprinting mechanism" (SFM, Vimont et al. 2001, 2003a,b) or the North Pacific Meridional Model (NPMM, Chiang and Vimont 2004; Chang et al. 2007), and from the South Pacific through the South Pacific Meridional Model (SPMM, Zhang et al. 2014), with the extratropical influence on the equatorial Pacific accomplished by the equatorward penetration of coupled ocean-atmosphere disturbances via the Wind-Evaporation-SST (WES) feedback (Liu and Xie 1994; Liu 1996; Vimont 2010). As such, extratropical climate variability and the associated tropical Pacific climate variability, such as NPMM, has also been suggested as a precursor for the onset of ENSO (Anderson 2007; Chang et al. 2007; Larson and Kirtman 2013, 2014).

Since ENSO is well known to exert a strong impact on extratropical climate over the Pacific (Alexander et al. 2002; Liu and Alexander 2007) and therefore can be closely coupled with extratropical climate variability, the study of the extratropical impact on ENSO requires first to separate the triggering extratropical variability from the tropical ENSO variability itself before the extratropical impact on ENSO could be studied. The impact of extratropical climate variability on ENSO has been studied mainly in two approaches. In the first approach, the extratropical climate variability and its equatorial impact is statistically extracted from observations or a control simulation in a fully coupled general circulation model (CGCM) using linear statistical methods (e.g. Vimont et al. 2001, 2003a,b; Anderson 2007; Yu et al. 2010; Larson and Kirtman 2013, 2014). However, ENSO is known to exert a significant impact on extratropical climate variability, and its impact may not be filtered cleanly in the observation statistically (Compo and Sardeshmukh 2010). In addition, some key extratropical impact signals, such as the NPMM (Chiang and Vimont 2004) and SPMM (Zhang et al. 2014), have significant imprint all the way into the deep tropics. Therefore the extratropical impact on ENSO may not be fully represented in this approach. In the second approach, the extratropical variability is simulated dynamically or specific extratropical forcing is prescribed in a coupled model that filters out ENSO by employing a slab ocean (Vimont

et al. 2009; Newman et al. 2011; Zhang et al. 2014), and then its impact on ENSO can be studied using either statistical methods or another coupled model that incorporates ENSO dynamics. This approach can isolate the extratropical impact from ENSO variability, but the dynamics of the impact may not be well represented because of the use of different models. In either approach, the extratropical impact on ENSO is studied more from a statistical perspective, which makes it difficult to assess the role of extratropical impact for any specific ENSO event.

As a complementary approach to the previous work, we attempt to study the extratropical impact on tropical coupled climate variability explicitly using an approach derived from coupled data assimilation (CDA). We will prescribe the extratropical variability using an ensemble CDA scheme in a CGCM and then study the tropical responses, in particular ENSO response, of the coupled model to the extratropical forcing. The CDA, especially the ensemble Kalman filter (EnKF) and its variations, has been used as a comprehensive strategy for generating climate reanalyses and initial conditions for prediction in the coupled climate system (Zhang et al. 2007; Saha et al. 2010; Raeder et al. 2012; Karspeck et al. 2013). Here, we will use the CDA as a tool for the understanding of climate dynamics, specifically the extratropical impact on ENSO in a GCCM. Ensemble based data assimilation has been used in the study of atmospheric dynamics (e.g. Hakim and Torn 2008; Torn and Hakim 2008; Liu and Kalnay 2008; Kalnay et al. 2012; Kunii et al. 2012), but seems to have not been used to investigate the coupled dynamics in a CGCM. An ensemble approach has been adopted in some previous studies mainly to reduce the atmospheric noises (e.g. Kirtman and Shukla 2002; Vimont et al. 2009).

As a pilot study, we will use CDA to investigate coupled model dynamics and teleconnections systematically in a CGCM in the perfect-model framework. More specifically, active data assimilation of model-generated observations is performed in the extratropics and the forced climate variability in the tropics is studied. The use of model ensemble minimizes the noise from natural variability in the experiments. Our study shows that extratropical atmospheric variability in our CGCM can indeed exert significant control on ENSO and therefore serve as a precursor for ENSO onset. This paper is organized as follows. Section 2 describes the CGCM, our CDA system and the design of the experiments. Section 3 includes a summary of all experiments and detailed analysis of 2 cases. Section 4 summarizes the results and discusses the implications.

2. Model and Methods

a. Model description

The CGCM that we used is the Fast Ocean Atmospheric Model (FOAM, version 1.5), which has been used in several studies about coupled data assimilation (Liu et al. 2014a,b; Lu et al. 2015). FOAM is a fully coupled global atmosphere-ocean model (Jacob 1997). The atmosphere component (PCCM3-UW, Drake et al. 1995) is a spectral model with a R15 horizontal resolution (equivalent to *7.5°×4.5°*) and a hybrid vertical coordinate with 18 levels. The ocean component (OM3) is based on the Modular Ocean Model (MOM, Cox 1984) created by the Geophysical Fluid Dynamics Laboratory (GFDL). It has a horizontal resolution of *2.8°×1.4°* and 24 vertical levels. The land surface and sea ice models are based on those of Community Climate Model 2 (CCM2, Hack et al. 1993). Without flux adjustment, a 6000-model-year simulation of FOAM shows no apparent drift in tropical climate (Liu et al. 2007). FOAM is able to capture most major features of the observed global climatology and climate variability as in some more advanced CGCMs.

b. ENSO in FOAM

The current version of FOAM incorporates a parameterization of the solar penetration depth (SPD), which can influence the tropical climate (Lewis et al. 1990; Schneider and Zhu 1998; Murtugudde et al. 2002). FOAM simulates a reasonable tropical climatology, although the model still exhibits a tendency of strong Cold Tongue and double ITCZ (Fig. 1a) as in most state-of-the-art CGCMs (e.g. Lin 2007). The model equatorial eastern Pacific SST is also dominated by an annual cycle (not shown).

The model ENSO is dominated by variability with the frequency of 2-6 years, as shown by the power spectrum of the Nino3.4 index (average SST anomaly in the region of 5°S-5°N, 120°W-170°W) in Fig. 1b, similar to previous work (Liu et al. 2000). The evolution of the equatorial (3°S-3°N meridional average) upper ocean temperature associated with ENSO can be seen in its regression on the normalized Nino3.4 index for different leads and lags. A weak subsurface warming develops in the western Pacific 12 months prior to the peak of ENSO (Fig. 2a), expanding eastward along the thermocline and filling the entire equatorial upper Pacific with an anomalously high heat content (Fig. 2b); the eastern Pacific SST (and therefore ENSO onset) then starts to develop (Fig. 2c), growing in amplitude (Fig. 2d) and spread westward across the surface of equatorial Pacific (Fig. 2e); in the meantime, a weak cooling develops in the subsurface western Pacific, growing in amplitude (Fig. 2f) and spreads across the thermocline eastward (Fig. 2g),

leading to the following cold event of La Nina (Fig. 2g, h, i), which resembles the preceding warm event (Fig. 2a, b, c), albeit with a smaller amplitude. The similar patterns, but of opposite sign, between the 9-month lead (Fig. 2b) and 9-month lag (Fig. 2h) regressions imply a prevailing period of about 3 years. The oceanic process displayed in Fig. 2, along with the active Bjerknes feedback (not shown), follows the classical delayed recharge oscillator theory (Cane and Zebiak 1985; Cane et al. 1986; Jin 1997).

c. CDA and observing system

The CDA system in FOAM (Lu et al. 2015; Liu et al. 2014a, 2014b) uses the Ensemble Adjustment Kalman Filter (EAKF, Anderson 2001, 2003) and includes both atmosphere data assimilation (ADA) and ocean data assimilation (ODA). All experiments in this study use an ensemble size of 16, similar to previous studies (Lu et al. 2015; Liu et al. 2014a). Considering the moderate ensemble size, a relax-to-prior scheme (Zhang et al. 2004) is used for all variable updates with a relaxation factor of 0.5. More details about the CDA system are described in Lu et al. (2015).

A perfect-model framework is adopted, and the output of a 50-year control simulation is used as the "truth". The observations are constructed by adding Gaussian white noise onto the "truth". The observations are therefore gridded data at the same gridpoints as the model state variables. The available observations include 5-day-mean SST with an error scale (standard deviation) of 0.5 K and daily-mean atmosphere temperature (T) and wind components (U, V) with error scales of 1 K and 1 m/s, respectively. These observational errors and frequencies represent typical conditions for such observed variables as in previous studies (Zhang et al. 2007; Liu et al. 2014a; Lu et al. 2015). The details of the CDA is not critical for the purpose of this study, because the CDA is used only as a sophisticated way of "nudging" the model ensemble atmosphere towards the "observation". Nevertheless, the CDA likely minimizes the shocks of "observational" constraints and provides a more accurate reconstruction of atmospheric variability in the assimilation region. Furthermore, the ensemble aspect is of critical importance for the suppression of noise and the interpretation of the results.

d. Experiment design

The CDA system serves as the foundation of our proposed new approach. Because of the flexibility of the ensemble-based Kalman filter, the data assimilation can be selectively activated for any model variable in any region. In other words, the system could proceed with only a subset of the

observations described in Section 2c. If both ADA and ODA are active over the entire globe, the assimilation product is like a "reanalysis" in the CGCM and thus the model is closely constrained by the observations everywhere (e.g. Zhang et al. 2007). If the data assimilation is active only in a limited region, the product shows the impact of the observations in the chosen region on the evolution of the coupled model outside the assimilation region, similar to the observation sensitivity experiments (e.g. Kunii et al. 2012).

In this paper, we apply this approach to investigate the influence of extratropical climate variability on tropical Pacific climate variability with the focus on ENSO. In particular, we will focus on the forcing role of extratropical atmospheric variability on ENSO, so that most of our experiments apply only ADA in the coupled model. The sole use of ADA ensures that the atmosphere is the only source of observational information, while the use of a coupled model provides the full ocean dynamics over the globe and fully coupled ocean-atmosphere dynamics both inside and outside the assimilation region. The major experiments consist of those with the ADA activated at all latitudes (*ada_all*), poleward of 20° in both hemispheres (*ada_20*), north of 20°N (*ada_north20*), and south of 20°S (*ada_south20*). These major experiments, along with other supplementary experiments, are summarized in Table 1: *CTRL* represents the ensemble control experiment without data assimilation; *ADA* (*ODA*) experiments employ only ADA (ODA) in the coupled model; *CDA* experiments employ both ADA and ODA simultaneously in the coupled model. For *CDA* and *ODA*, the assimilation is active either at all latitudes (*cda_all* and *oda_all*) or poleward of 20° (*cda_20* and *oda_20*) in both hemispheres. Besides, additional *ADA* experiments are performed with active ADA poleward of 10° (*ada_10*), 30° (*ada_30*) and 40° (*ada_40*) in both hemispheres, north of 30°N (*ada_north30*), or south of 30°S (*ada_south30*).

More specifics of the experiment design can be illustrated using the *ada_20* experiment as the example. The ADA is activated only in the extratropics (poleward of 20°) and therefore the atmosphere resembles closely to the observation in the extratropics. This observed atmospheric variability in the extratropics directly forces the underlying extratropical ocean via buoyancy, heat and momentum fluxes through the coupler, as well as the tropical climate system through atmospheric and coupled dynamics. In the tropics, the model is constrained along the boundaries at 20°S and 20°N in the atmosphere. Because the use of ensemble-based data assimilation, the model tropics are forced by slightly different extratropical variability in each ensemble member, so each model member's tropical coupled climate variability tends to develop differently due to

its chaotic natural variability. The averaged tropical variability of all ensemble members, however, can eliminate the influence of the chaotic natural variability, as later shown by the results of the *CTRL* experiment and similar to the ensemble approach in Kirtman and Shukla (2002). However, compared to Kirtman and Shukla (2002) where ENSO variability could be caused by both tropical internal variability and extratropical forcing, the current approach can isolate the extratropical control on ENSO variability such that any significant tropical variability in the ensemble average should be generated by the common extratropical atmospheric forcing to all ensemble members. Another main difference between our approach and previous ensemble experiments (Kirtman and Shukla 2002; Vimont et al. 2009; Larson and Kirtman 2015a) is that the extratropical atmospheric variability, as well as its forcing on the ocean and the tropical atmosphere, are prescribed continuously as in the observations such that the coupled tropical climate variability is no longer a pure initial value ensemble forecast. Instead, our experiments represent the tropical climate variability forced by the observed extratropical atmospheric variability. Furthermore, our CDA using EnKF should provide a more accurate analysis of the atmospheric/oceanic states in the assimilation region than, for example, a simple nudging scheme.

Each experiment runs for 52 years, starting from an ensemble of initial conditions of 16 consecutive years within the long control simulation. The data assimilation is activated after 2 years of spin-up and lasts for 50 years, which is the length of the observation. The ensemble-mean output is calculated by averaging the monthly outputs from all ensemble members, and the anomalies of all variables are then calculated by subtracting the corresponding seasonal cycles. All experiments are repeated with different sets of initial conditions and observational errors, and the results are robust. Therefore, we will base our analysis on one set of experiments unless otherwise specified.

e. Performance of CDA system

Fig. 3 displays the quality of the atmospheric (T and U) and oceanic (SST) analyses over the Pacific (120°E-80°W) for experiments *ada_all*, *ada_20*, *ada_north20* and *ada_south20*, all normalized by *CTRL*. The RMSE (root mean square error) is calculated from the differences between monthly ensemble-mean analysis and the "truth" at each gridpoint, and Fig. 3 shows the zonally-averaged RMSE. The quality of V analysis is quantitively similar to T and U. When the ADA is employed across all latitudes (Fig. 3a), the RMSE of both T and U are reduced by 70-80% across all latitudes

and heights compared to *CTRL*. Because the analyzed atmosphere provides fairly accurate surface boundary condition to the ocean, the SST RMSE is also reduced by over 70% in the tropics and 40-60% in the mid-latitudes. When the ADA inside 20° is removed in *ada_20* (Fig. 3b), the quality of analyses is largely maintained outside of 20°, but deteriorates rapidly from the 20° boundaries (dash-dot lines) equatorward. However, the tropical atmosphere and SST are still partially constrained because of the accurate boundary conditions provided by ADA outside of 20°. In the deep tropics, RMSEs of T, U and SST are still 30-40% smaller than *CTRL*. Similarly for *ada_north20* (Fig. 3c) and *ada_south20* (Fig. 3d), the analyses are well constrained where assimilation is active, and deteriorates rapidly beyond the assimilation boundaries. The RMSE ratio of SST is essentially the same as that of the atmospheric surface temperature in all of Fig. 3. The ADA outside of 20° in one hemisphere could affect the equatorial region, but shows little influence on the other hemisphere, where the RMSE remains the same as *CTRL*. At the equator, the RMSEs are reduce by about 5-10% in *ada_north20* and *ada_south20*.

3. Extratropical Atmospheric Forcing of ENSO

a. General Assessment

Overall, our experiments show a significant control of extratropical atmospheric variability on ENSO. Table 1 summarizes all experiments' ability to reproduce the "true" ENSO variability by showing the RMSE of their ensemble-mean Nino3.4 indices in regards to the "true" Nino3.4 index. The ensemble spread of the Nino3.4 indices and the RMSE as a percentage of that of *CTRL* are also shown for each experiment. The ensemble spread represents the uncertainty in the Nino3.4 indices of the model ensemble, and is comparable to the corresponding RMSE, which is usually the case for a well-behaved CDA system in the perfect model framework (Anderson 2001; Zhang et al. 2007). Fig. 4 shows the Nino3.4 time series of several experiments (*CTRL*, *ada_all*, *ada_20*, *ada_north20*, *ada_north20A*, *ada_south20*, and *ada_south20A*), including those of each ensemble member, the ensemble mean and the "truth".

- The RMSE of 0.617 in *CTRL* is very close to the standard deviation of the "true" Nino3.4 index because the ensemble average of 16 control simulations stays close to 0 at all times (Fig. 4a). Even though each ensemble member has its own natural variability, the ensemble mean is affected little by the variability of any single member.

- Nino3.4 index closely resembles the "truth" when ODA is active in the tropics. This should be expected since the tropical SST is directly adjusted by observations in *oda_all*. Meanwhile, extratropical ODA has no effect on Nino3.4, as the RMSE of *oda_20* is comparable to that of *CTRL*. It should be noted that *oda_20* only assimilates SST observations poleward of 20°. The failure of *oda_20* to reproduce the tropical variability in the observation implies the ineffectivenss of extratropical SST variability alone in forcing tropical climate variability.

- Among the *ada* experiments, the RMSE increases, as expected, when the boundaries of active ADA moves poleward. However, extratropical ADA shows significant influence on ENSO variability, reducing the RMSE of Nino3.4 by over 40% in *ada_20* and over 30% in *ada_30* compared to *CTRL*. The *ada_all* experiment, in which the ocean is forced by close-to-observation atmosphere at all latitudes, could accurately reproduce the “true” Nino3.4 index with minimal ensemble spread (Fig. 4b). The *ada_20* experiment produces both larger RMSE and ensemble spread compared to *ada_all*, but its ensemble-mean output still captures most of the major ENSO events in the "truth" (Fig. 4c). The fact that most ENSO events in *ada_20* are significant among all ensemble members indicates that extratropical atmosphere alone could generate consistent ENSO signals.

- When the ADA is limited to only one hemisphere, the resulting Nino3.4 variability is significantly worse than when ADA is active in both hemispheres with the same latitudal boundary (87.0% and 80.7% vs. 57.0% for 20° cases). As shown by Fig. 4d and Fig. 4f, not only are much fewer ENSO events correctly produced in *ada_south20* and *ada_north20*, the ensemble-mean magnitudes are also much smaller, and the ensemble spread much larger.

- The ENSO variability in *ada_south20* and *ada_north20* is indeed forced by the extratropical atmosphere, confirmed by parallel experiments *ada_north20A* and *ada_south20A* (Table 1 and Fig. 4), which are the same as *ada_north20* and *ada_south20*, respectively, except for different initial conditions and random observational errors. Each pair of experiments (e.g. *ada_north20* and *ada_north20A*) have almost identical RMSEs of Nino3.4 indices and reproduce the ENSO events at roughly the same times and with similar magnitudes.

The results in Table 1, along with Fig. 4, clearly demonstrate that the extratropical atmosphere has significant impact on ENSO variability. Fig. 4 also shows the necessity of using ensembles. For *ada_north20* and *ada_south20*, or even some ENSO events in *ada_20*, the Nino3.4 index differs

greatly among ensemble members. Due to the chaotic natural variability, each individual member often fails to capture the "true" ENSO events, while also generates erroneous events that do not exist in the "truth". Clearly, the impact of natural variability is minimized by taking the ensemble average from the CDA scheme, as shown by the near-zero Nino3.4 index of *CTRL*.

We note that the *CTRL* experiment can be used conveniently as a benchmark to evaluate the significance of other experiments. For every variable, the distribution of its ensemble-mean anomaly from *CTRL* specifies the magnitude of its natural variability without any input observations. In the following sections, the standard deviations of the ensemble-mean anomalies from *CTRL* (CTRL_SD hereafter) will be used to evaluate the significance of all variables. For example, the Nino3.4 index of *CTRL* has a standard deviation of 0.17°C and a maximum value of 0.55°C. In fact, there are only two instances when the Nino3.4 index exceeds 0.5°C over the 50 years of *CTRL*, and each lasts only one month. This means that the ensemble-mean ENSO events in *ada_20*, *ada_north20* or *ada_south20* that exceed peak Nino3.4 value of 0.5°C are very unlikely to be caused by natural variability. Rather, they are caused by the assimilation of extratropical atmospheric observations, which provides the same signal across all ensemble members. Therefore for the analysis of *ada_south20*, an ENSO event is counted everytime the monthly Nino3.4 index exceeds 0.5°C, regardless of the duration. A different criterion like Nino3.4 over 0.4°C for at least 3 consecutive months gives almost the same events[3]. The identified ENSO events in *ada_south20* are indicated by the red (El Nino) and blue (La Nina) circles in Fig. 4f. Note that there are a few instances where two ENSO events are very close to each other. Because the weak phase-locking of ENSO events to the seasonal cycle, we will keep the identified events that have peak values less than 1 year but more than 6 months apart from each other.

b. Extratropical Control of ENSO

Now, we examine the control of extratropical atmosphere on ENSO variability in detail with the focus on the three main experiments, *ada_20, ada_south20* and *ada_north20*. First, the extratropical atmospheric variability from both hemispheres can generate most of the "true" ENSO variability, while that from each hemisphere generates less. In Fig. 4d and Fig. 4f, although both

[3] The ENSO events in *ada_south20* are identified with lower SST threshold and shorter duration compared to observational standard because first, the overall weaker ENSO variability in FOAM (Nino3.4 standard deviation 0.65°C) than in observation (around 1.0°C), and second, the even weaker ENSO variability of *ada_south20* than the "truth" due to ensemble averaging.

ada_north20 and *ada_south20* could reproduce some of the ENSO events from the "truth", the magnitudes are mostly smaller than the "truth" or *ada_20*, and the overall variances of the Nino3.4 indices are also smaller. These features are displayed more clearly in Fig. 5a, where the Nino3.4 indices of *ada_north20* (blue) and *ada_south20* (red) are scattered against the "truth". The markers are less scattered vertically than horizontally, indicating smaller variances of Nino3.4 in *ada_north20* and *ada_south20* than the "truth". More specifically, the standard deviation of the Nino3.4 index is 0.63 for the "truth", 0.58 for *ada_20*, 0.39 for *ada_north20*, and 0.44 for *ada_south20*.

Although *ada_north20* and *ada_south20* each inherits, in a statistical sense, about half of the Nino3.4 variance from *ada_20*, the relationship among *ada_20*, *ada_north20* and *ada_south20* for individual ENSO events is complicated, as seen in the time series of Fig. 4c, d and f. Some events in *ada_20* are produced in *ada_north20*, but not in *ada_south20* (around year 11, 27, 42); some are the opposite, produced in *ada_south20* but not in *ada_north20* (around year 5, 9, 36); some are partially produced in both *ada_north20* and *ada_south20*, but in smaller magnitudes (around year 4, 19). Overall *ada_north20* and *ada_south20* are reproducing mostly different ENSO events, indicating independent forcing from NH and SH. The independence of NH and SH extratropical control is supported by the plot between the Nino3.4 indices of *ada_north20* and *ada_south20* in Fig. 5b, which has a weak correlation of 0.193 but an adjusted R^2 of only 0.033. Therefore, in the following sections, we will analyze *ada_north20* and *ada_south20* separately to better understand the role of extratropical atmospheric forcing on ENSO. Interestingly, the sum of the Nino3.4 indices of *ada_north20* and *ada_south20* closely resembles that of *ada_20* with a correlation of 0.81, and the standard deviation of the sum is 0.61.

The extratropical control of ENSO and the comparable contribution from NH and SH can also been seen in the seasonality of ENSO. Fig. 6 plots the standard deviations by calendar month of the Nino3.4 indices from "truth", *ada_20*, *ada_north20* and *ada_south20*. The seasonality of Nino3.4 variability in *ada_20* is similar to the observation, albeit slightly smaller in magnitude for all months. The peak variance of the Nino3.4 index occurs in August and September, followed by October and November, which is shifted early to boreal fall from the boreal winter in the real world. In the single-hemisphere forcing experiments, the overall variance decreases significantly and the seasonality changes. Relatively speaking, there are higher spring variability (April to June) in *ada_north20* and lower spring variability in *ada_south20*. These changes of seasonality should

be related to their respective extratropical atmospheric forcing and will be discussed in the following sections.

4. ENSO and Precursors

Now, we examine in detail the extratropical control of ENSO, with the emphasis on the precursors for ENSO events. We will focus on the SH forcing in *ada_south20*, because the NH influence can be discussed in a similar fashion. We first discuss the potential precursors in equatorial heat content and extratropical atmosphere based on the composite of ENSO events. We will then examine if each potential precursor is sufficient to trigger ENSO on its own based on the composite of anomalous events on each precursor. Finally, we will discuss a potentially necessary and sufficient precursor by combining the heat content and extratropical atmosphere precursors together.

a. ENSO Composites and Potential Precursors

ENSO evolution and its potential precursors of ENSO in *ada_south20* can be seen in the composite of ENSO events in Fig. 7, which shows the composite of the Nino3.4 index (Fig. 7a), the zonal propagation of equatorial (5°S-5°N meridional average) Pacific SST (Fig. 7b) for the 29 ENSO events in *ada_south20* (La Nina events are included with reversed sign). These ENSO events, as indicated by the circles in Fig. 4f, have peak Nino3.4 magnitudes that exceed 0.5°C, which is very close to the maximum value (0.55°C) and 3 times the standard deviation (0.17°C, black dashed lines in Fig. 7a) of the Nino3.4 index from CTRL. In other words, all these ENSO events in *ada_south20* are very unlikely to occur due to natural variability without the assimilated extratropical observations. The Nino3.4 composite shows a warming that starts 6-7 months prior to the peak, rises rapidly 2 months prior to the peak, and slowly decays afterwards over a span of 10 months. Although there is no obvious cycle, the Nino3.4 composite is slightly negative 1-2 years prior and after the peak, indicating the succession between warm and cold events. Fig. 7b shows that the SST anomalies originate mostly from the eastern equatorial Pacific, and propagate westward from 100°W to the dateline in 3-4 months. In sum, the composite ENSO events in Fig. 7 are highly significant compared to CTRL, indicating a robust tropical response to SH extratropical atmospheric forcing.

Next we examine the evolution of surface (Fig. 8) and subsurface (Fig. 9) conditions prior to the ENSO events in *ada_south20*. Composites of SST, low-level wind, and latent heat flux (LHF

hereafter) anomalies are created in the equatorial and south Pacific for the 5 months prior to the Nino3.4 peak in *ada_south20* (Fig. 8). The low-level wind is averaged over the bottom 3 levels (>900mb). Positive LHF anomaly (blue contours) indicates more LHF out of the ocean or a cooling effect on the ocean. The SST and wind anomalies are only visible where they exceed their CTRL_SD. The NH extratropics is excluded from Fig. 8 because no robust ensemble-mean signals exist there for any variable. Only the atmospheric (T, U, V) observations south of 20°S are assimilated, so any robust signals in the ensemble-mean anomalies of *ada_south20* should be attributed, ultimately, to SH extratropical atmospheric variability.

The propagation of SH extratropical variability into the equatorial Pacific preceding ENSO in *ada_south20* resembles the SPMM described in Zhang et al. (2014). At 5 months prior to the Nino3.4 peak (Fig. 8a), there is already positive SST anomaly in the eastern equatorial Pacific. Meanwhile, a low pressure centers at 40°S, 100°W as in the observation. The anomalous northwesterlies to the northeast of the low-pressure center reduces the climatological southeasterly trade winds and, in turn, the LHF out of the ocean. One month later (Fig. 8b), the low-pressure center moves equatorward to 30°S, 100°W with the wind anomalies intensified. The wind anomalies that extend beyond the assimilation region to north of 20°S, where the atmosphere and ocean are fully coupled and active, could be caused by atmospheric dynamic adjustment (Pierce et al. 2000). Furthermore, the reduced LHF cooling continues to warm the SST in the southeast subtropical Pacific. This subtropical warming differs from the simultaneous equatorial warming in that the LHF anomaly tends to intensify, instead of damp, the SST anomaly in the former. As such, the coupled wind-SST anomaly can propagate equatorward through the WES feedback as suggested by Liu and Xie (1994). The warm SST anomaly would induce anomalous westerlies on the equatorial side, which further reduces the mean easterly trade wind and LHF there, allowing the coupled anomaly to propagate equatorward. In the following 2 months (Fig. 8c, d), the subtropical SST anomaly moves from 20°S to 10-15°S and becomes connected to the pre-existing equatorial SST anomaly. The anomalous northwesterlies still persist in the eastern equatorial and southeast tropical Pacific. Meanwhile, anomalous westerly wind expands westward and intensifies along the equator from 100°W all the way to 160°W. In the following month (Fig. 8e), the SST and wind anomalies start to move off the coast. In the meantime, the off-equator northerly wind anomalies in southeast Pacific have mostly reversed to southerly, and the LHF anomalies have

changed sign to positive across the entire tropical and subtropical eastern Pacific. Overall, the subtropics-to-tropics process in Fig. 8 is similar to the SPMM process found in CMIP3 (Coupled Model Intercomparison Project phase 3) AGCM-slab ocean and fully coupled models, as well as in the real world (Zhang et al. 2014).

The accompanied evolution of the equatorial (3°S-3°N average) upper ocean temperature can be seen for 8 and 4 months prior to the Nino3.4 peak in the composite of Fig. 9. There is significant subsurface warming before the ENSO events in *ada_south20*, similar to the regressed upper ocean temperature evolution of the 500-year control simulation in Fig. 2 (thus only two snapshots in Fig. 9). Both Fig. 2 and Fig. 9 show an eastward and upward propagation and the subsequent ENSO onset in the eastern equatorial Pacific.

Based on Fig. 2 and Fig. 9, we use the heat content anomaly in the equatorial western Pacific as the subsurface precursor for ENSO, with the heat content anomaly (HC_EW hereafter) calculated as the 3-month running-averaged anomalous temperature integrated in the western equatorial Pacific (3°S-3°N, 120°E-180°, 40-240m). We will also use the 3-month running-averaged low-level wind speed anomaly (WS_SE hereafter) in the southeast subtropical Pacific (15°S-25°S, 80°W-100°W, bottom 3 levels) as the extratropical atmospheric precursor. We use wind speed instead of sea level pressure (Anderson 2007; Deser et al. 2012) or SST (Zhang et al. 2014) because wind is directly constrained by the ADA in our experiments and better represents the prescribed extratropical atmospheric variability due to assimilation. As shown in Fig. 8, the anomalous wind speed in this region is the largest 3-5 months prior to the Nino3.4 peak and leads to reduced LHF cooling, SST warming and subsequent WES propagation.

The composites of WS_SE and HC_EW series based on the 29 ENSO events in Fig. 7a are plotted in Fig. 7c and Fig. 7d, respectively. The WS_SE composite shows significant negative values 3-5 months prior to the Nino3.4 peak, where it exceeds twice its CTRL_SD. There is also a sign change for WS_SE right before the Nino3.4 peak, consistent with the reversal of trade wind anomalies after the ENSO onset in Fig. 8e. The HC_EW composite shows consistently positive values in the 1.5 years leading to the Nino3.4 peak and significant negative values after the Nino3.4 peak. The peak value of HC_EW composite preceding ENSO is not as significant as WS_SE, mainly because the time between HC_EW and Nino3.4 peaks varies greatly among ENSO events. Based on the preceding WS_SE and HC_EW, the 29 individual events in Fig. 7 are color coded as follows: 18 “purple” events when any negative WS_SE value exceeds its

CTRL_SD in the range of 3-5 months prior to Nino3.4 peak *AND* any positive HC_EW value exceeds its CTRL_SD in the range of 6-18 months prior to Nino3.4 peak; 6 "green " events when *only* the WS_SE criterion is met; 2 "yellow" events when *only* the HC_EW criterion is met; and 3 "blue" events when *neither* criterion is met. In sum, most of the ENSO events are preceded by significant anomalies in the wind and heat content precursors, given our wide range of leading times.

Fig. 7, Fig. 8 and Fig. 9 together confirm that ENSO events in *ada_south20* are preceded by significant anomalous signals, or precursors, in both SH extratropical atmosphere/ocean and equatorial subsurface ocean. This is consistent with previous observational and modeling analyses, which showed that both the forcing of extratropical atmospheric variability and the preconditioning of the equatorial ocean heat content could play critical roles in the ENSO onset (e.g. Anderson 2007; Zhang et al. 2009; Vimont et al. 2009; Deser et al. 2012; Larson and Kirtman 2013, 2014). As pointed out by Sarah and Kirtman (2014), however, this type of composite analysis is based on subsampling only the years with ENSO events, therefore the identified precursors, such as WS_SE and HC_EW here, are potentially necessary precursors, but may be insufficient to trigger ENSO in the forecast sense or even produce false alarms. In the following, we further determine the necessity and sufficiency of either precursor, the anomalous equatorial heat content or extratropical atmosphere, for triggering ENSO events in the *ada_south20* experiment.

b. Subsurface Ocean Preconditioning

To focus on the role of the equatorial ocean heat content as a precursor, the 24 strongest anomalous events[4] in the HC_EW time series are composited (Fig. 10a), accompanied by the corresponding composites of Nino3.4 (Fig. 10b) and WS_SE (Fig. 10c) that are aligned by the HC_EW peaks in Fig. 10a. An individual event is marked as blue or red dashed line, while the average of all events as the black solid line. The horizontal black dashed lines in each plot indicate the corresponding CTRL_SD. Although all the HC_EW events have significant heat content anomaly in the western Pacific, their composite Nino3.4 index following the HC_EW peaks is close to 0, indistinguishable

[4] The threshold is about twice the corresponding CTRL_SD and is chosen such that there are a comparable number of HC_EW events as ENSO events. The negative anomalous events are again included by reversing the sign. The Nino3.4 and WS_SE series of those negative HC_EW events are also reversed accordingly.

from the natural variability of CTRL. Note the large spread among the Nino3.4 responses, we select those 12 HC_EW events that are followed by ENSO events (Nino3.4 over 0.5°C) of the same sign in the following 16 months and mark them red. The "red" events account for about half of the ENSO events in *ada_south20*. The remaining 12 HC_EW events are marked blue, and the averages of "red" and "blue" events are shown by the solid red and blue lines, respectively, in all of Fig. 10. The 6 strongest HC_EW events are indeed "red", but most other "red" and "blue" events are indistinguishable in terms of HC_EW magnitude. The "red" Nino3.4 composite hovers around only 0.2-0.4°C because the response time from the HC_EW peak to the Nino3.4 peak varies considerably. For the same reason of various Nino3.4 response time, there is no clear negative peak in the wind WS_SE composite for the "red" events (Fig. 10c). However, the "red" WS_SE composite does stay negative in 6 of the 9 months following the HC_EW peak, while consistent and significant positive values of the "blue" WS_SE composite following the HC_EW peak are more notable (Fig. 10c), indicating that intensified trade wind in southeast subtropical Pacific could hamper the possible ENSO onset induced by the heat content anomaly. It is also notable that very robust negative Nino3.4 anomalies precede the HC_EW events (Fig. 10b), indicating the overshoot effect of previous ENSO events on subsequent equatorial heat content. The composite analyses based on the HC_EW events show that the heat content precursor is not a reliable predictor for ENSO onset (12 of 24 are followed by ENSO events), although its reliability could be increased (decreased) by subsequent favorable (unfavorable) extratropical atmospheric conditions.

The spatial pattern of the difference between the "red" and "blue" HC_EW events can be seen in the composites of the equatorial subsurface ocean (Fig. 11). At the time of HC_EW peak (Fig. 11a, d), the ocean temperature anomalies are almost identical, with cold anomalies at the surface and in the eastern Pacific and strong warm anomalies in the western Pacific thermocline. On average, the eastern Pacific is slightly cooler in "blue" events than in "red" ones. The warm temperature anomalies propagate eastward for both "red" and "blue" events, however, the anomalies decay considerably for the "blue" events such that the magnitude and extent of the heat content anomaly in the eastern Pacific are much smaller in the "blue" events than in the "red" events (Fig. 11e vs. b). At 6 months after the HC_EW peak, the "red" events have displayed considerable warming in the eastern Pacific SST (Fig. 11c), while the warm anomalies in the "blue" events have disappeared and turned into substantial cooling in the eastern Pacific. The comparison of the two evolutions is

quite striking: almost identical thermocline anomalies in the western Pacific lead to completely different ENSO responses. Different from previous analyses of the observation or a fully coupled control simulation (e.g. Anderson 2007; Deser et al. 2012; Larson and Kirtman 2013, 2014), the design of the our experiment ensures that this difference is caused, ultimately, by the information from the extratropical atmosphere in the SH, rather than that in the NH or the initial conditions in the tropics.

The difference between the "red" and "blue" events can also be seen in the composites of the surface climate (Fig. 12) for SST, wind and LHF composites at several lags after the HC_EW peak. For the "red" events, the tropical warming following the HC_EW events is the direct surfacing of the subsurface anomalies in Fig. 11, but it is not preceded by any significant extratropical anomalies as in Fig. 8. Because the timing from HC_EW peaks to the triggering of ENSO events varies considerably for different events, the short extratropical signals would not result in robust composites in Fig. 12a-c. Conversely and more notably for the "blue" events, cold SST anomalies and intensified trade winds persist in the southeast tropical and subtropical Pacific following the HC_EW peak, which could offset the positive temperature anomalies propagating eastward (Fig. 11d-f) and hamper the possible onset of ENSO events. The discussion here suggests that equatorial heat content alone is insufficient to trigger ENSO and is therefore not a sufficient predictor for ENSO. This is consistent with an real-world observational analysis by Anderson (2007).

c. Extratropical Atmospheric Forcing

Next, we assess the extratropical atmospheric variability, in particular WS_SE, as a precursor for ENSO events in *ada_south20*. Fig. 13 shows the composite analyses based on the 24 strongest anomalous WS_SE events[5] from *ada_south20* in the same way as the heat content in Fig. 10. The composite in Fig. 13a shows that all the WS_SE events have highly significant 1-month peaks (the width is enlarged due to the 3-month running average), reflecting the short timescale of the extratropical atmospheric internal variability. Most of these WS_SE events occur from May through August, since the model SH extratropical atmosphere is the most energetic in boreal summer. At the peak of the WS_SE composite, the SST, wind and LHF anomalies (Fig. 13b) resemble the precursor composite based on ENSO events in Fig. 8b, albeit with much stronger

[5] The threshold for WS_SE events is 1m/s and it is 7 times the corresponding CTRL_SD. The much stronger ensemble-mean wind variability compared to *CTRL* is due to the data assimilation in the WS_SE region.

magnitudes in wind and weaker magnitudes in the equatorial SST warming. Following the WES feedback as in Fig. 8, the coupled anomalies of reduced wind speed, weaker LHF cooling, and warm SST could propagate into the equatorial region and trigger ENSO onset in about 2 months. However, on average, the Nino3.4 composite reaches the maximum of only 0.28°C and stays within or very close to CTRL_SD throughout the entire period following the WS_SE peaks (black solid, Fig. 13c). Among the 24 WS_SE events, only 8 are followed by ENSO events (Nino3.4 over 0.5°C) of the correct sign in the following 2-6 months (Fig. 13c). Again we mark these 9 events in red (dashed) and the other 16 in blue (dashed), and plot their averages as the solid red and blue lined, respectively. In terms of the magnitude of the WS_SE events, the "red" average is indistinguishable from the "blue" average (Fig. 13a), but the subsequent responses in Nino3.4 are notably different: the "red" average exhibits a significant warming 4 months after the negative peak of WS_SE, while the "blue" average exhibits no warming after the wind peak (Fig. 13c). The HC_EW composite based on the WS_SE events is plotted in Fig. 13d. The 3 WS_SE events that are preceded by large positive HC_EW values are indeed "red", while the other 6 "red" events do not have significant leading HC_EW events. On average, the positive "red" HC_EW composite is primarily caused by the 3 events with large HC_EW values. Similar to the heat content precursor, the composite analyses based on the WS_SE events show that the extratropical atmospheric precursor alone is not a reliable predictor either (9 of 24 are followed by ENSO events), while favorable equatoral preconditioning could increase its skill.

The preconditioning role of the equatorial ocean can also be seen by comparing the preceding patterns of SST and heat content between the "red" and "blue" events. The surface conditions at the peak of the WS_SE composite are plotted for "red" and "blue" events separately in Fig. 14. While the SH extratropical anomalies remain the same as in the total composite in Fig. 13b, there is a dramatic difference in the equatorial eastern Pacific. The "red" events have the same equatorial warming as in Fig. 8, while the "blue" events have slight cooling in the equatorial Pacific. This cold anomaly does not favor the onset of ENSO events, and is likely to prevent the extratropical atmospheric variability from triggering ENSO. Indeed, the difference in equatorial SST between the "red" and "blue" events can be traced back to the subsurface conditions prior to the WS_SE events. Fig. 15 shows the composites of ocean temperature anomaly 8 and 4 months prior to, and at simultaneous time of WS_SE events for the "red" and "blue" events, separately. The subsurface evolution of the "red" events (Fig. 15a-c) matches the development in the composites of Fig. 9 and

Fig. 11a-c. In contrast, the "blue" events exhibits weak to moderate subsurface cooling across the entire Pacific thermocline (Fig. 15d-f), which does not provide the favorable preconditioning for the onset of ENSO events. The discussion here suggests that, similar to the heat content, the extratropical atmospheric variability alone is insufficient to trigger ENSO and therefore is not a sufficient precursor, either.

d. The Combined Effect: Equatorial Heat Content and Extratropical Atmosphere

The discussions above show that both WS_SE and HC_EW demonstrate robust signals prior to ENSO events in *ada_south20* and therefore can be considered as potentially necessary precursors for ENSO onset. However, neither is sufficient to trigger ENSO by itself and be considered as a reliable predictor. This is consistent with previous studies of observational and model analyses (Anderson 2007; Deser et al. 2012; Larson and Kirtman 2013, 2014, 2015b). These previous works further suggested that the combined effect of both equatorial preconditioning and extratropical atmospheric variability is more likely to trigger ENSO. This point seems to be consistent with our study here (Fig. 10 through Fig. 15).

The combined effect of both the WS_SE and HC_EW precursors can be illustrated more clearly in Fig. 16, which is the scatterplot of the peak Nino3.4 values of all 29 ENSO events in *ada_south20* based on the maximum WS_SE (opposite-sign) in the range of 3-5 months prior to Nino3.4 peak and the maximum HC_EW value (same-sign) in the range of 6-18 months prior. Most El Nino events are in the bottom right quadrant, preceded by significantly negative WS_SE *and* positive HC_EW values. Conversely, the upper left quadrant contains most La Nina events, which are preceded by significantly positive WS_SE *and* negative HC_EQ values. Based on Fig. 16, it may seem that WS_SE and HC_EW is negatively correlated. However, this "correlation" exists because only the values preceding ENSO events are chosen and more importantly, each precursor is chosen from a range of leading times to highlight its impact on ENSO variability. The entire WS_SE and HC_EW series in the *ada_south20* experiment vary independently with a correlation of -0.01.

The combined effect of WS_SE and HC_EW is further demonstrated in Fig. 17, which is similar to the Fig. 3 of Anderson (2007) and Fig. 19 of Deser et al. (2012). In Fig. 17, the correlation between April-July averaged WS_SE or June (previous year)-May averaged HC_EW and August-October averaged Nino3.4 of all years in *ada_south20* is separately estimated depending on

whether the leading averaged WS_SE and HC_EW have the opposite or same sign. When the averaged WS_SE and HC_EW have the opposite sign, the two precursors may work together to trigger ENSO onset; when they have the same sign, they may work against each other and fail to trigger ENSO onset. Fig. 17 shows that for averaged WS_SE (HC_EW), its negative (positive) correlation with the subsequent Nino3.4 is larger in the opposite-sign case (Fig. 17a, c) than the same-sign case (Fig. 17b, d). The inability of the two precursors to explain all the ENSO events is also reasonable, since ENSO variability can be forced by atmospheric variability features not considered here, such as the atmospheric variability in the central and western subtropical Pacific. The analyses on the SH impact in *ada_south20* can also be carried out similarly for the NH impact in *ada_north20*, a point to return to later.

The effectiveness of WS_SE and HC_EW together as a ENSO precursor (Fig. 16) and their relation with each other (Fig. 17) are consistent qualitatively with the previous works on ENSO precursors in the real world (Anderson 2007) and climate models (Deser et al. 2012; Larson and Kirtman 2013, 2014). However, it should be pointed out that there is an important difference between our study and previous studies. By experimental design, all the significant ensemble-mean tropical responses in *ada_south20* or *ada_north20*, including the ENSO events, are ultimately forced by SH or NH extratropical atmospheric variability. In contrast, the ENSO events in the observation or a fully coupled model simulation could be simultaneously related to extratropical variability from both hemispheres as well as internal climate variability within the tropics. The exclusively forced tropical response in our model assures that the equatorial heat content anomaly precursor is also forced, somehow, by the extratropical atmosphere. In contrast, the ocean preconditioning in previous observational or modeling studies could be caused, again by extratropical variability from both hemispheres and tropical natural variability. It remains to be studied how the equatorial subsurface ocean is preconditioned by extratropical atmospheric variability. Tentatively, we speculate the heat content anomaly can be induced by the oceanic teleconnection (Matei et al. 2008), with the extratropical atmospheric variability directly forcing subtropical oceanic anomaly, which then propagations into the equatorial ocean via oceanic Rossby wave and the subsequent tropical thermocline adjustment (Kirtman 1997; Anderson et al. 2013; Anderson and Perez 2015), or via thermocline subduction (Liu et al. 1994; Schott et al. 2004); it can also be forced directly by the atmospheric adjustment from the subtropics (Pierce et al. 2000).

5. Summary and Discussions

This paper studies the control of extratropical atmospheric variability on ENSO variability in a CGCM using a limited-domain CDA system, in which the active assimilation is confined to the extratropics. The extratropical atmospheric variability is shown to have significant impact on ENSO variability, while the extratropical SST has no influence on ENSO at the timescale considered here. When atmospheric observations are assimilated only poleward of 20° in both hemispheres, most of the strong ENSO events in observation are reproduced in *ada_20* and the RMSE of the Nino3.4 index is reduced by over 40% compared to *CTRL* with no assimilation. The comparison with *CTRL* also indicates that any robust ensemble-mean tropical variability must be attributed to the assimilated extratropical atmosphere ultimately. Furthermore, the forced ENSO variability is contributed independently and roughly equally by the atmospheric forcing from the NH and SH extratropical atmosphere, as shown by *ada_south20* and *ada_north20*.

Composite analyses of the ENSO events in *ada_south20* reveal robust signals in both extratropical atmosphere and equatorial subsurface Pacific, following previous studies of SPMM and the discharge-recharge paradigm. Specifically, most *ada_south20* ENSO events are preceded by corresponding precursors of wind speed, LHF and SST anomalies in the southeast subtropical Pacific similar to the SPMM and heat content anomalies in the equatorial western Pacific. Furthermore, the ability of each precursor to act as a predictor is investigated based on their own composites in *ada_south20*. It is found that neither the extratropical atmospheric nor the tropical oceanic precursor alone is sufficient to trigger ENSO onset. The combined effect of the two precursors is also considered: when one precursor strongly favors ENSO onset, the other should also be favorable, or at least neutral for the emergence of ENSO. The existence of ENSO events without either precursor indicates the possibility of other connections from SH extratropical atmosphere to ENSO variability.

Detailed analysis of the *ada_north20* experiment, which is performed in the same way as *ada_south20*, is not shown in this paper. The major conclusions from *ada_south20* qualitatively hold for *ada_north20*, although the results are somewhat more complex for *ada_north20*. In *ada_north20*, most ENSO events are preceded by wind, LHF and SST anomalies in the northeast and north-central subtropical and tropical Pacific, while only some are preceded by heat content anomalies in the equatorial western Pacific. The extratropical signals in *ada_north20* spread over

637 a larger region compared to *ada_south20* and resemble the key patterns of the NPMM. However,
638 a few notable differences should be pointed out. First, *ada_north20* appears to include both the
639 Central-Pacific (CP) and Eastern-Pacific (EP) types of ENSO events (Yu and Kao 2007; Kao and
640 Yu 2009), while *ada_south20* includes the EP type exclusively. The EP type in *ada_north20* is
641 preceded by significant HC_EW precursor similar to *ada_south20*, while the CP type is not.
642 Furthermore, the CP type in *ada_north20* is more closely linked to preceding SST anomalies in
643 the northeast subtropical Pacific and NH trade wind anomalies, similar to the "footprinting"
644 mechanism (Vimont et al. 2001, 2003a,b) and the trade wind charging (Anderson et al. 2013;
645 Anderson and Perez 2015). The relation between NPMM or SPMM and different types of ENSO
646 events has been investigated by several studies (Yu et al. 2010; Zhang et al. 2014; Vimont et al.
647 2014; Yeh et al. 2015), and most of which agree that NPMM is related to both types while SPMM
648 leads to mostly the EP type. The different types of ENSO events and their respective mechanism
649 in *ada_north20* will be explored in a future paper.

650 It should be pointed out that the quantitative strength of the extratropical control on ENSO
651 variability depends on the dynamics of the model. The same *ada_20* experiment is also performed
652 in two other versions of FOAM with modified model parameters and shows significantly different
653 strength of extratropical control. In one version of FOAM, ENSO becomes much stronger than the
654 default version and is dominated by a biennial oscillation. In this version, the extratropical
655 atmospheric variability has little impact on ENSO, suggesting the ENSO is generated
656 predominantly by the tropical coupled system. In another version of FOAM, ENSO becomes
657 weaker and its spectrum appears much "redder". In this version, the extratropical atmosphere
658 exerts an even stronger control on ENSO than in the default version. Our diagnosis of the control
659 simulations of the various versions of FOAM suggests that the default version is the closest to the
660 real world. In addition, the almost equal strength of extratropical control from both hemispheres
661 could also be model-specific. Nevertheless, we caution the direct application of the quantitative
662 results in this study to other CGCMs and the real world.

663 There are still many issues to be further explored on the current study. As previously discussed,
664 the coupled dynamics in the eastern Pacific should only be one of the mechanisms by which
665 extratropical atmosphere could influence ENSO variability. In addition, it remains unclear how
666 the ocean preconditioning is caused by extratropical atmospheric variability. More analyses are
667 needed to investigate the roles of atmospheric/oceanic/coupled processes in transferring

extratropical variability into the tropics. This pilot study also serves to demonstrate of the utility of the limited-domain CDA method, or more generally the use of coupled data assimilation in studying climate dynamics. The real potential of this method would be its application to the understanding of the real world ENSO events using a state-of-the-art CGCM. With this new method based on the CDA system, it is possible to investigate the specific atmospheric noise forcing and related coupled dynamics that contribute to each real world ENSO event.

Acknowledgements

We thank Drs. Shu Wu and Xinyao Rong for their work during the early development of our data assimilation code. This research is sponsored by NSF AGS-0968383 and Chinese MOST 2012CB955200.

References

Alexander, M. a., and Coauthors, 2002: The atmospheric bridge: The influence of ENSO teleconnections on air-sea interaction over the global oceans. *J. Clim.*, **15**, 2205–2231, doi:10.1175/1520-0442(2002)015<2205:TABTIO>2.0.CO;2.

Anderson, B. T., 2007: On the joint role of subtropical atmospheric variability and equatorial subsurface heat content anomalies in initiating the onset of ENSO events. *J. Clim.*, **20**, 1593–1599, doi:10.1175/JCLI4075.1.

——, and R. C. Perez, 2015: ENSO and non-ENSO induced charging and discharging of the equatorial Pacific. *Clim. Dyn.*, **45**, 2309–2327, doi:10.1007/s00382-015-2472-x.

——, ——, and A. Karspeck, 2013: Triggering of El Niño onset through trade wind-induced charging of the equatorial Pacific. *Geophys. Res. Lett.*, **40**, 1212–1216, doi:10.1002/grl.50200.

Anderson, J. L., 2001: An ensemble adjustment Kalman filter for data assimilation. *Mon. Wea. Rev.*, **129**, 2884–2903.

Cane, M. A., and S. E. Zebiak, 1985: A theory for el nino and the southern oscillation. *Science*, **228**, 1085–1087, doi:10.1126/science.228.4703.1085.

——, ——, and S. C. Dolan, 1986: Experimental forecasts of El Niño. *Nature*, **321**, 827–832, doi:10.1038/321827a0.

Chang, P., L. Zhang, R. Saravanan, D. J. Vimont, J. C. H. Chiang, L. Ji, H. Seidel, and M. K. Tippett, 2007: Pacific meridional mode and El Niño-Southern Oscillation. *Geophys. Res. Lett.*, **34**, n/a – n/a, doi:10.1029/2007GL030302.

Chiang, J. C. H., and D. J. Vimont, 2004: Analogous Pacific and Atlantic Meridional Modes of Tropical Atmosphere–Ocean Variability*. *J. Clim.*, **17**, 4143–4158, doi:10.1175/JCLI4953.1.

Compo, G. P., and P. D. Sardeshmukh, 2010: Removing ENSO-related variations from the climate record. *J. Clim.*, **23**, 1957–1978, doi:10.1175/2009JCLI2735.1.

Cox, M. D., 1984: *A primitive equation 3-dimensional model of the ocean*. GFDL–Princeton University, 143 pp.

Deser, C., and Coauthors, 2012: ENSO and pacific decadal variability in the community climate system model version 4. *J. Clim.*, **25**, 2622–2651, doi:10.1175/JCLI-D-11-00301.1.

Drake, J., I. Foster, J. Michalakes, B. Toonen, and P. Worley, 1995: Design and performance of a scalable parallel community climate model. *Parallel Comput.*, **21**, 1571–1591, doi:10.1016/0167-8191(96)80001-9.

Hack, J. J., B. A. Boville, B. P. Briegleb, J. T. Kiehl, P. J. Rasch, and D. Williamson, 1993: *Description of the NCAR community climate model (CCM2)*. Climate and Global Dynamics Division, NCAR, 108 pp.

Hakim, G. J., and R. D. Torn, 2008: Ensemble Synoptic Analysis. *Meteorol. Monogr.*, **33**, 147–162, doi:10.1175/0065-9401-33.55.147.

715 Jacob, R., 1997: Low Frequency Variability in a Simulated Atmosphere Ocean System. University of
716 Wisconsin-Madison, 159 pp.

717 Jin, F. F., 1997: An equatorial ocean recharge paradigm for ENSO. Part 1: Conceptual model. *J. Atmos.*
718 *Sci.*, **54**, 811–829, doi:Doi 10.1175/1520-0469(1997)054<0811:Aeorpf>2.0.Co;2.

719 Kalnay, E., Y. Ota, T. Miyoshi, and J. Liu, 2012: A simpler formulation of forecast sensitivity to
720 observations: application to ensemble Kalman filters. *Tellus A*, **64**, doi:10.3402/tellusa.v64i0.18462.

721 Kao, H. Y., and J. Y. Yu, 2009: Contrasting Eastern-Pacific and Central-Pacific types of ENSO. *J. Clim.*,
722 **22**, 615–632, doi:10.1175/2008JCLI2309.1.

723 Karspeck, A. R., S. Yeager, G. Danabasoglu, T. Hoar, N. Collins, K. Raeder, J. L. Anderson, and J.
724 Tribbia, 2013: An Ensemble Adjustment Kalman Filter for the CCSM4 Ocean Component. *J. Clim.*,
725 **26**, 7392–7413, doi:Doi 10.1175/Jcli-D-12-00402.1.

726 Kirtman, B. P., and J. Shukla, 2002: Interactive coupled ensemble: A new coupling strategy for CGCMs.
727 *Geophys. Res. Lett.*, **29**, 5–1 – 5–4, doi:10.1029/2002GL014834.

728 Kunii, M., T. Miyoshi, and E. Kalnay, 2012: Estimating the Impact of Real Observations in Regional
729 Numerical Weather Prediction Using an Ensemble Kalman Filter. *Mon. Wea. Rev.*, **140**, 1975–1987,
730 doi:Doi 10.1175/Mwr-D-11-00205.1.

731 Larson, S. M., and B. P. Kirtman, 2013: The Pacific Meridional Mode as a trigger for ENSO in a high-
732 resolution coupled model. *Geophys. Res. Lett.*, **40**, 3189–3194, doi:10.1002/grl.50571.

733 ——, and ——, 2014: The Pacific Meridional Mode as an ENSO Precursor and Predictor in the North
734 American Multimodel Ensemble. *J. Clim.*, **27**, 7018–7032, doi:10.1175/JCLI-D-14-00055.1.

735 ——, and ——, 2015a: An alternate approach to ensemble ENSO forecast spread: Application to the
736 2014 forecast. *Geophys. Res. Lett.*, **42**, 9411–9415, doi:10.1002/2015GL066173.

737 ——, and ——, 2015b: Revisiting ENSO Coupled Instability Theory and SST Error Growth in a Fully
738 Coupled Model. *J. Clim.*, **28**, 4724–4742, doi:10.1175/JCLI-D-14-00731.1.

739 Lewis, M., M. Carr, G. Feldman, W. Esaias, and C. McClain, 1990: Influence of penetrating solar
740 radiation on the heat budget of the equatorial Pacific Ocean. *Nature*, **347**, 543–545.

741 Lin, J. L., 2007: The double-ITCZ problem in IPCC AR4 coupled GCMs: Ocean-atmosphere feedback
742 analysis. *J. Clim.*, **20**, 4497–4525, doi:Doi 10.1175/Jcli4272.1.

743 Liu, J., and E. Kalnay, 2008: Estimating observation impact without adjoint model in an ensemble
744 Kalman filter. *Q. J. R. Meteorol. Soc.*, **134**, 1327–1335, doi:10.1002/qj.280.

745 Liu, Y., Z. Liu, S. Zhang, R. Jacob, F. Lu, X. Rong, and S. Wu, 2014a: Ensemble-Based Parameter
746 Estimation in a Coupled General Circulation Model. *J. Clim.*, **27**, 7151–7162, doi:10.1175/JCLI-D-
747 13-00406.1.

748 ——, ——, ——, X. Rong, R. Jacob, S. Wu, and F. Lu, 2014b: Ensemble-based parameter estimation in
749 a coupled GCM using the adaptive spatial average method. *J. Clim.*, **27**, 4002–4014,

doi:10.1175/JCLI-D-13-00091.1.

Liu, Z., 1996: Modeling Equatorial Annual Cycle with a Linear Coupled Model. *J. Clim.*, **9**, 2376–2385, doi:10.1175/1520-0442(1996)009<2376:MEACWA>2.0.CO;2.

——, and S. S.-P. Xie, 1994: Equatorward Propagation of Coupled Air–Sea Disturbances with Application to the Annual Cycle of the Eastern Tropical Pacific. *J. Atmos. Sci.*, **51**, 3807–3822, doi:10.1175/1520-0469(1994)051<3807:EPOCAD>2.0.CO;2.

——, and M. Alexander, 2007: Atmospheric bridge, oceanic tunnel, and global climatic teleconnections. *Rev. Geophys.*, **45**, RG2005, doi:10.1029/2005RG000172.

——, S. G. H. Philander, and R. C. Pacanowski, 1994: A GCM Study of Tropical–Subtropical Upper-Ocean Water Exchange. *J. Phys. Oceanogr.*, **24**, 2606–2623, doi:10.1175/1520-0485(1994)024<2606:AGSOTU>2.0.CO;2.

——, J. Kutzbach, and L. X. Wu, 2000: Modeling climate shift of El Nino variability in the Holocene. *Geophys. Res. Lett.*, **27**, 2265–2268, doi:Doi 10.1029/2000gl011452.

——, and Coauthors, 2007: Simulating the transient evolution and abrupt change of Northern Africa atmosphere–ocean–terrestrial ecosystem in the Holocene. *Quat. Sci. Rev.*, **26**, 1818–1837, doi:10.1016/j.quascirev.2007.03.002.

Lu, F., Z. Liu, S. Zhang, Y. Liu, and R. Jacob, 2015: Strongly Coupled Data Assimilation Using Leading Averaged Coupled Covariance (LACC). Part II: CGCM Experiments*. *Mon. Wea. Rev.*, **143**, 4645–4659, doi:10.1175/MWR-D-15-0088.1.

Matei, D., N. Keenlyside, M. Latif, and J. Jungclaus, 2008: Subtropical Forcing of Tropical Pacific Climate and Decadal ENSO Modulation. *J. Clim.*, **21**, 4691–4709, doi:10.1175/2008JCLI2075.1.

Murtugudde, R., J. Beauchamp, C. R. McClain, M. Lewis, and A. J. Busalacchi, 2002: Effects of Penetrative Radiation on the Upper Tropical Ocean Circulation. *J. Clim.*, **15**, 470–486, doi:10.1175/1520-0442(2002)015<0470:EOPROT>2.0.CO;2.

Neelin, J. D., D. S. Battisti, A. C. Hirst, F.-F. Jin, Y. Wakata, T. Yamagata, and S. E. Zebiak, 1998: ENSO theory. *J. Geophys. Res.*, **103**, 14261, doi:10.1029/97JC03424.

Newman, M., M. a. Alexander, and J. D. Scott, 2011: An empirical model of tropical ocean dynamics. *Clim. Dyn.*, **37**, 1823–1841, doi:10.1007/s00382-011-1034-0.

Philander, S. G., 1990: *El Niño, La Niña, and the Southern Oscillation*. Elsevier,.

Pierce, D. W., T. P. Barnett, and M. Latif, 2000: Connections between the Pacific Ocean Tropics and Midlatitudes on Decadal Timescales. *J. Clim.*, **13**, 1173–1194, doi:10.1175/1520-0442(2000)013<1173:CBTPOT>2.0.CO;2.

Raeder, K., J. L. Anderson, N. Collins, T. J. Hoar, J. E. Kay, P. H. Lauritzen, and R. Pincus, 2012: DART/CAM: An Ensemble Data Assimilation System for CESM Atmospheric Models. *J. Clim.*, **25**, 6304–6317, doi:Doi 10.1175/Jcli-D-11-00395.1.

785 Saha, S., and Coauthors, 2010: The NCEP Climate Forecast System Reanalysis. *Bull. Am. Meteorol. Soc.*,
786 **91**, 1015–1057, doi:10.1175/2010BAMS3001.1.

787 Schneider, E. K., and Z. Zhu, 1998: Sensitivity of the Simulated Annual Cycle of Sea Surface
788 Temperature in the Equatorial Pacific to Sunlight Penetration. *J. Clim.*, **11**, 1932–1950,
789 doi:10.1175/1520-0442(1998)011<1932:SOTSAC>2.0.CO;2.

790 Schott, F. a, J. P. Mccreary, and G. C. Johnson, 2004: Shallow Overturning Circulations of the Tropical-
791 Subtropical Oceans. *Earth's Climate*, 261–304.

792 Torn, R. D., and G. J. Hakim, 2008: Ensemble-Based Sensitivity Analysis. *Mon. Wea. Rev.*, **136**, 663–
793 677, doi:10.1175/2007MWR2132.1.

794 Vimont, D. J., 2010: Transient growth of thermodynamically coupled variations in the tropics under an
795 equatorially symmetric mean state. *J. Clim.*, **23**, 5771–5789, doi:10.1175/2010JCLI3532.1.

796 ——, D. S. Battisti, and A. C. Hirst, 2001: Footprinting: A seasonal connection between the tropics and
797 mid-latitudes. *Geophys. Res. Lett.*, **28**, 3923–3926, doi:10.1029/2001GL013435.

798 ——, ——, and ——, 2003a: The Seasonal Footprinting Mechanism in the CSIRO General Circulation
799 Models*. *J. Clim.*, **16**, 2653–2667, doi:10.1175/1520-0442(2003)016<2653:TSFMIT>2.0.CO;2.

800 ——, J. M. Wallace, and D. S. Battisti, 2003b: The Seasonal Footprinting Mechanism in the Pacific:
801 Implications for ENSO*. *J. Clim.*, **16**, 2668–2675, doi:10.1175/1520-
802 0442(2003)016<2668:TSFMIT>2.0.CO;2.

803 ——, M. a. Alexander, and A. Fontaine, 2009: Midlatitude Excitation of Tropical Variability in the
804 Pacific: The Role of Thermodynamic Coupling and Seasonality*. *J. Clim.*, **22**, 518–534,
805 doi:10.1175/2008JCLI2220.1.

806 ——, ——, and M. Newman, 2014: Optimal growth of Central and East Pacific ENSO events. *Geophys.*
807 *Res. Lett.*, **41**, 4027–4034, doi:10.1002/2014GL059997.

808 Yeh, S.-W., X. Wang, C. Wang, and B. Dewitte, 2015: On the Relationship between the North Pacific
809 Climate Variability and the Central Pacific El Niño. *J. Clim.*, **28**, 663–677, doi:10.1175/JCLI-D-14-
810 00137.1.

811 Yu, J.-Y., and H.-Y. Kao, 2007: Decadal changes of ENSO persistence barrier in SST and ocean heat
812 content indices: 1958-2001. *J. Geophys. Res. Atmos.*, **112**, n/a – n/a, doi:10.1029/2006JD007654.

813 ——, ——, and T. Lee, 2010: Subtropics-Related Interannual Sea Surface Temperature Variability in the
814 Central Equatorial Pacific. *J. Clim.*, **23**, 2869–2884, doi:10.1175/2010JCLI3171.1.

815 Zhang, F., C. Snyder, and J. Z. Sun, 2004: Impacts of initial estimate and observation availability on
816 convective-scale data assimilation with an ensemble Kalman filter. *Mon. Wea. Rev.*, **132**, 1238–
817 1253.

818 Zhang, H., A. Clement, and P. Di Nezio, 2014: The South Pacific Meridional Mode: A Mechanism for
819 ENSO-like Variability. *J. Clim.*, **27**, 769–783, doi:10.1175/JCLI-D-13-00082.1.

820 Zhang, L., P. Chang, and L. Ji, 2009: Linking the Pacific Meridional Mode to ENSO: Coupled Model
821 Analysis. *J. Clim.*, **22**, 3488–3505, doi:10.1175/2008JCLI2473.1.

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

825

826 Table 1 A summary of all experiments in this study.

827 Fig. 1 (a) Annual-mean tropical SST climatology from a 500-year control simulation. (b) The
828 power spectrum of the Nino3.4 index from the same 500-year model control simulation.

829 Fig. 2 Regression of equatorial upper ocean temperature (3°S-3°N meridional average) on
830 normalized Nino3.4 index from the 500-year control simulation. The relative times between
831 ocean temperature and Nino3.4 index are -12, -9, -6, -3, 0, 3, 6, 9 and 12 months (positive when
832 Nino3.4 index leads) from left to right and top to bottom.

833 Fig. 3 Pacific (120°E-80°W) zonal-mean RMSE (normalized by *CTRL* experiment) of ensemble-
834 mean atmospheric temperature (shadings), zonal wind (dashed lines) and SST (lower panel) for
835 (a) *ada_all*, (b) *ada_20*, (c) *ada_north20*, and (d) *ada_south20*. The thick dash-dot lines indicate
836 the boundaries of data assimilation where necessary.

837 Fig. 4 The time series of Nino3.4 index from (a) *CTRL*, (b) *ada_all*, (c) *ada_20*, (d)
838 *ada_south20*, and (e) *ada_north20*. Red lines indicate the "truth" (same for all panels), black
839 lines the ensemble mean, and grey lines all 16 ensemble members.

840 Fig. 5 Scatter plots with linear regression slope coefficient and adjusted R^2 for the Nino3.4
841 indices of (a) *ada_north20* (blue) and *ada_south20* (red) vs. truth, (b) *ada_south20* vs.
842 *ada_north20*.

843 Fig. 6 Standard deviations of the Nino3.4 index by calendar month from the *truth*, *ada_20*,
844 *ada_north20* and *ada_south20*.

845 Fig. 7 For *ada_south20*: (a) Composite of Nino3.4 index of 29 ENSO events exceeding 0.5°C.
846 La Nina events are included with reversed sign. Dashed lines indicate CTRL_SD, the standard
847 deviations of the ensemble-mean anomalies from *CTRL*. (b) Composite of 5°S-5°N averaged
848 Pacific SST based on the Nino3.4 peaks (shadings) and its ratio to the corresponding CTRL_SD
849 (contours). (c) Same as (a) but for the WS_SE series. (d) Same as (a) but for the HC_EW series.

850 Fig. 8 Composites of anomalous SST (shadings, K), LHF (contours, 4 W/m^2 interval and zero
851 line omitted) and wind (arrows) for the 5 months prior to the peak of the Nino3.4 composite in
852 *ada_south20*. Blue (red) contours indicate more (less) upward LHF. Black dotted lines indicate

853 the boundary of data assimilation. SST and wind anomalies are only shown where the composite
854 exceed CTRL_SD.

855 Fig. 9 Composite of anomalous 3°S-3°N averaged Pacific upper ocean temperature (a) 8 months
856 and (b) 4 months prior to the peak of the Nino3.4 composite in *ada_south20*. The contours
857 indicate the ratios of the composite anomalies to CTRL_SD.

858 Fig. 10 (a) Composite of HC_EW series for the 24 strongest HC_EW events. Negative
859 anomalies are included with reversed sign. The HC_EW events that are followed by ENSO
860 events of the same sign within 16 months are marked red, and the rest are marked blue. The blue
861 (red) solid line is the average of all blue (red) dashed lines. (b) Composite of Nino3.4 index in
862 *ada_south20* for the same events as (a); (c) Same as (b), but of WS_SE.

863 Fig. 11 Same as Fig. 9, but for the "red" (left column, a-c) and "blue" (right column, d-f)
864 HC_EW events from Fig. 10, respectively. The composites are at the same time as the peak of
865 HC_EW composite (a, d), 3 months after (b, e), and 6 months after (c, f).

866 Fig. 12 Same as Fig. 8, but for the "red" (left column, a-c) and "blue" (right column, d-f)
867 HC_EW events from Fig. 10, respectively. The composites are at 3 months (a, d), 6 months (b, e)
868 and 9 months (c, f) after the peak of HC_EW composite.

869 Fig. 13 (a) Composite of WS_SE for the 24 strongest WS_SE events. Positive anomalies are
870 included with reversed sign. The WS_SE events that are followed by ENSO events of opposite
871 sign within 6 months are marked red, and the rest are marked blue. The blue (red) solid line is
872 the average of all blue (red) dashed lines. (b) Same as Fig. 8, but for all the WS_SE events at the
873 peak. (c) Composite of Nino3.4 index in *ada_south20* for the same events as (a); (d) Same as (c),
874 but of HC_EW.

875 Fig. 14 Same as Fig. 13b, but for a) "red" and b) "blue" events, respectively.

876 Fig. 15 Same as Fig. 9, but for the "red" (left column, a-c) and "blue" (right column, d-f) WS_SE
877 events from Fig. 13a, respectively. The composites are 8 months (a, d), 4 months (b, e) prior to,
878 and at the same month (c, f) as the peak of the WS_SE composite.

Fig. 16 Scatterplot of all 29 ENSO events in *ada_south20* based on the minimum opposite-sign value of 3-5 months leading WS_SE (y-direction) and maximum same-sign value of 6-18 months leading HC_EW (x-direction). The color and size of the markers together indicate the peak Nino3.4 values. The dashed lines indicate the corresponding CTRL_SD.

Fig. 17 (a) Scatterplot of the April-July averaged WS_SE with August-October averaged Nino3.4 index of the same year, plotted for the years when the averaged WS_SE has the opposite sign as the June (previous year)-May (same year) averaged HC_EW. (b) Same as (a), but for the years when the averaged WS_SE has the same sign as the averaged HC_EW. (c) Scatterplot of the June (previous year)-May (same year) averaged HC_EW with August-October averaged Nino3.4 index, plotted for the years when the April-July averaged WS_SE has the opposite sign as the averaged HC_EW. (d) Same as (c), but for the years when the averaged WS_SE has the same sign as the averaged HC_EW.

Figure

Assimilation	Assimilated Latitudes	Experiment Name	RMSE of EnsMean Nino3.4	Ensemble Spread of Nino3.4	% of CTRL
None	None	*CTRL*	0.641	0.637	100%
ADA	All	*ada_all*	0.114	0.085	17.8%
	>10°N & <10°S	*ada_10*	0.206	0.169	32.2%
	>20°N & <20°S	*ada_20*	0.366	0.309	57.0%
	>30°N & <30°S	*ada_30*	0.439	0.384	68.4%
	>20°N	*ada_north20*	0.558	0.533	87.0%
		ada_north20A	0.551	0.535	86.0%
	<20°S	*ada_south20*	0.517	0.506	80.7%
		ada_south20A	0.524	0.507	81.8%
	>30°N	*ada_north30*	0.581	0.542	90.6%
	<30°S	*ada_south30*	0.600	0.605	93.5%
ODA	All	*oda_all*	0.125	0.105	19.5%
	>20°N & <20°S	*oda_20*	0.654	0.627	101.9%
CDA	All	*cda_all*	0.119	0.041	18.5%
	>20°N & <20°S	*cda_20*	0.347	0.319	54.1%

Table 1 A summary of all experiments in this study.

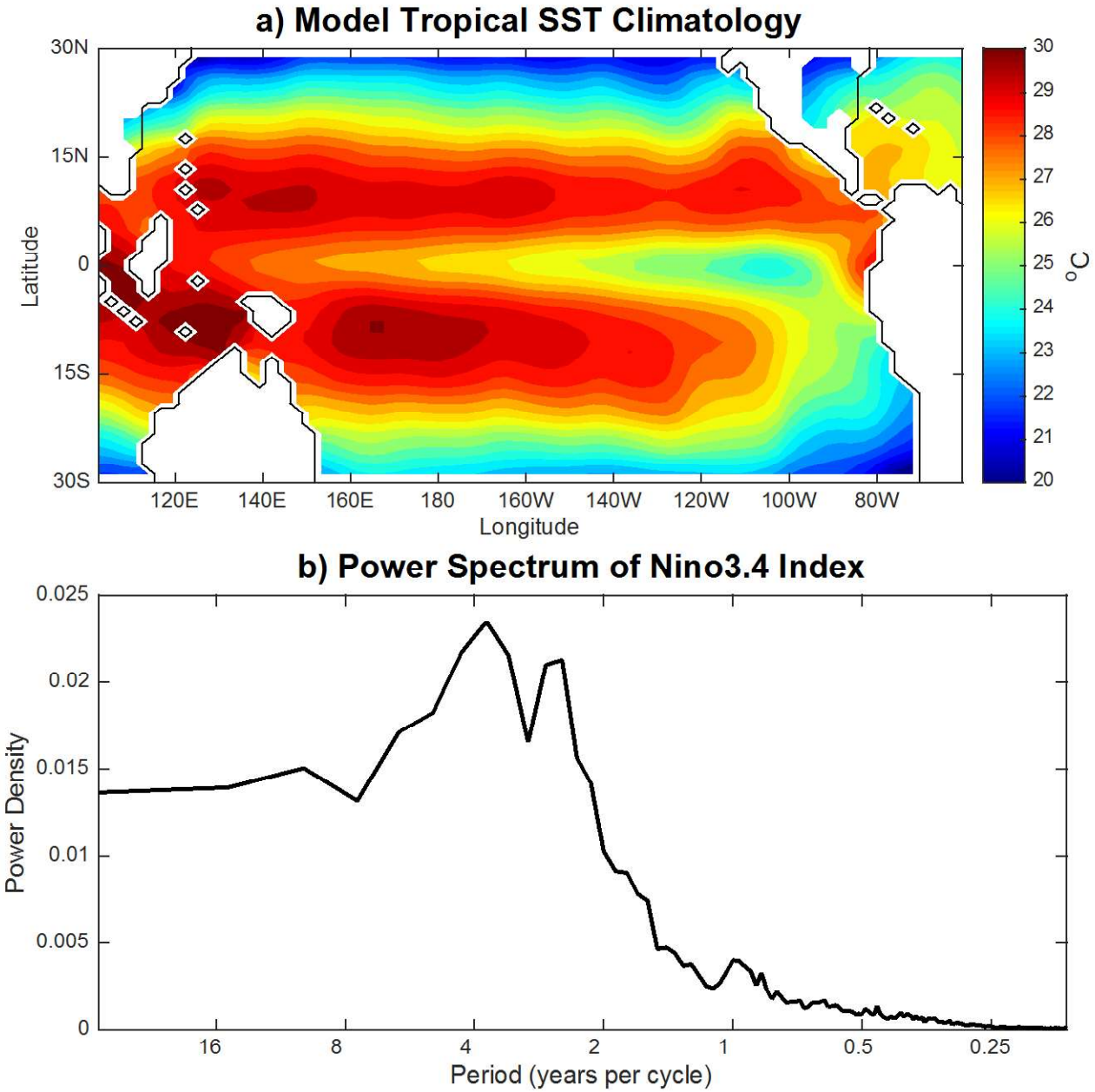

Fig. 1 (a) Annual-mean tropical SST climatology from a 500-year control simulation. (b) The power spectrum of the Nino3.4 index from the same 500-year model control simulation.

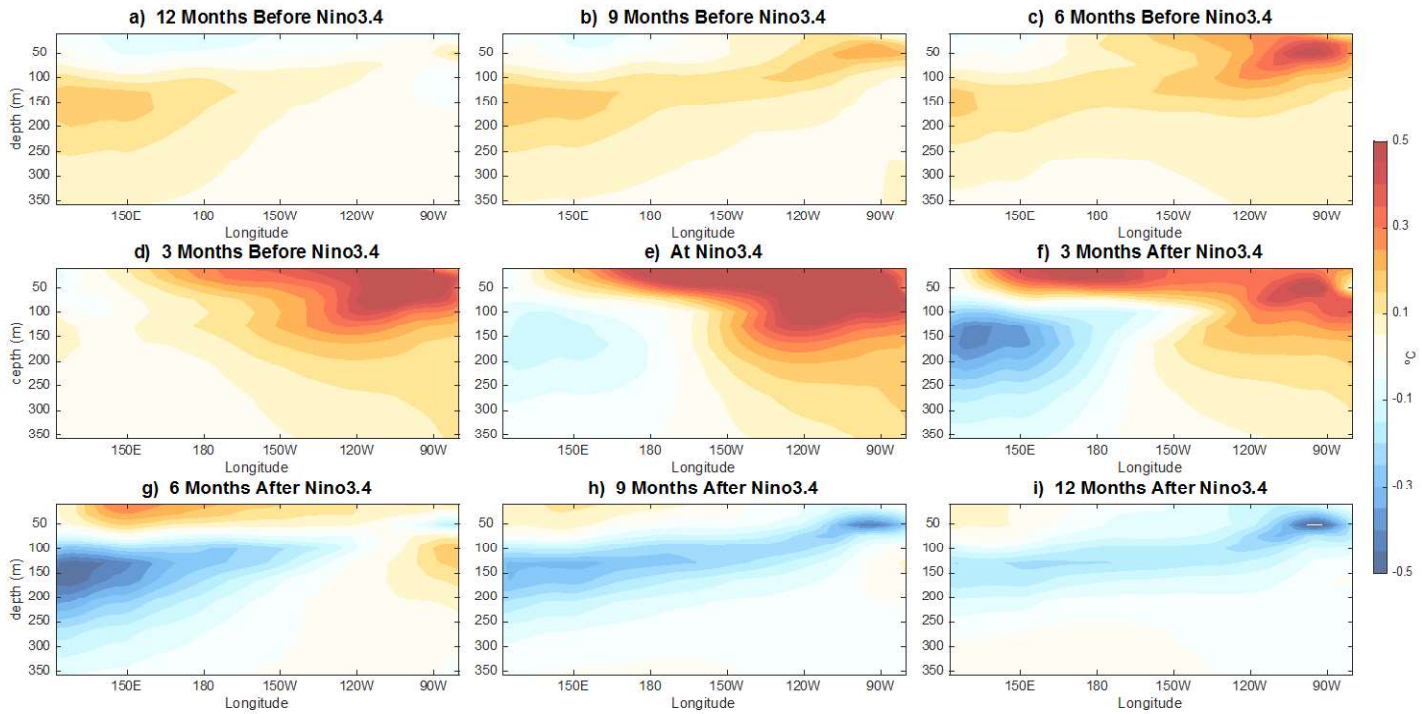

Fig. 2 Regression of equatorial upper ocean temperature (3°S-3°N meridional average) on normalized Nino3.4 index from the 500-year control simulation. The relative times between ocean temperature and Nino3.4 index are -12, -9, -6, -3, 0, 3, 6, 9 and 12 months (positive when Nino3.4 index leads) from left to right and top to bottom.

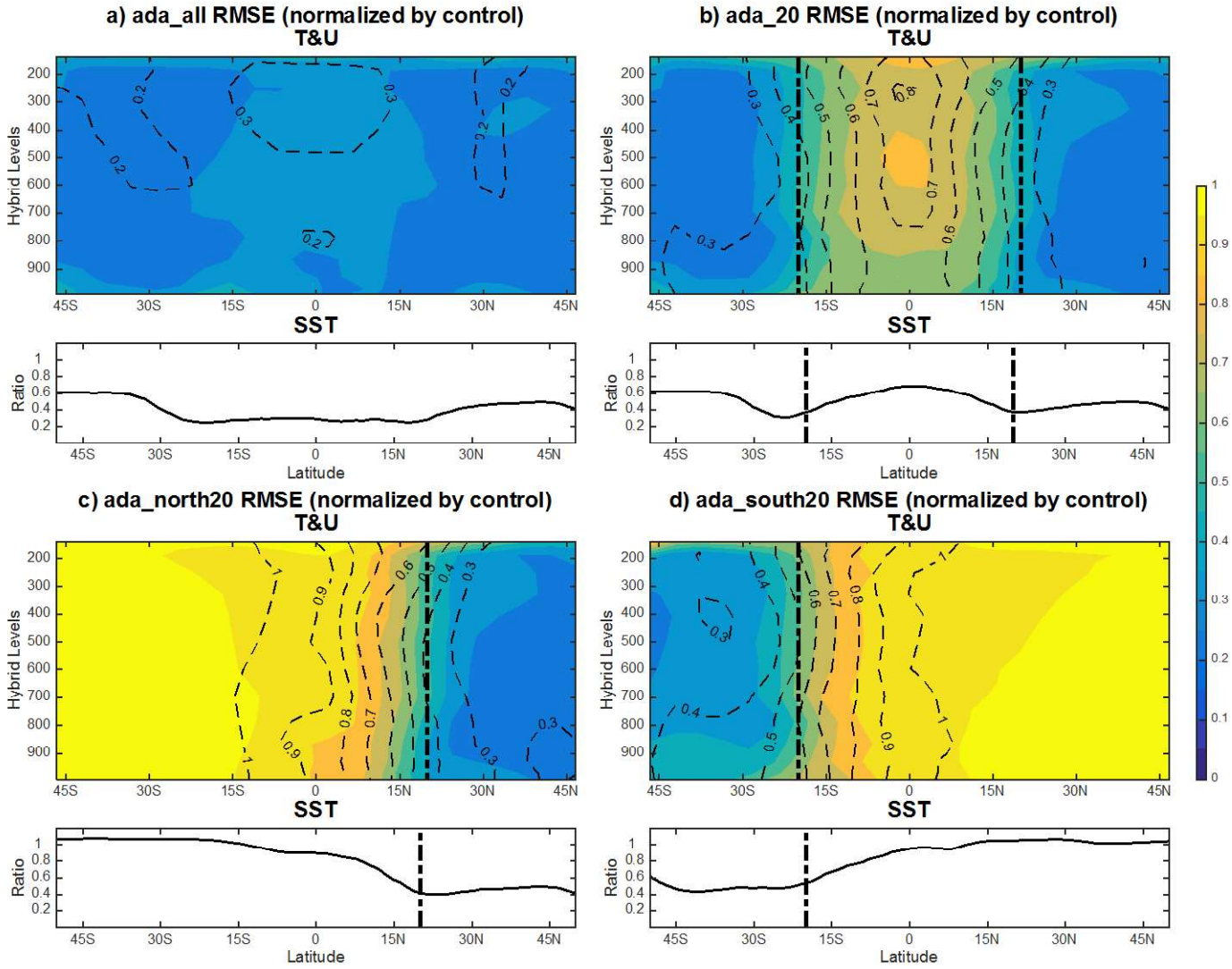

Fig. 3 Pacific (120°E-80°W) zonal-mean RMSE (normalized by *CTRL*) of ensemble-mean atmospheric temperature (shadings), zonal wind (dashed lines) and SST (lower panel) for (a) *ada_all*, (b) *ada_20*, (c) *ada_north20*, and (d) *ada_south20*. The thick dash-dot lines indicate the boundaries of data assimilation where necessary.

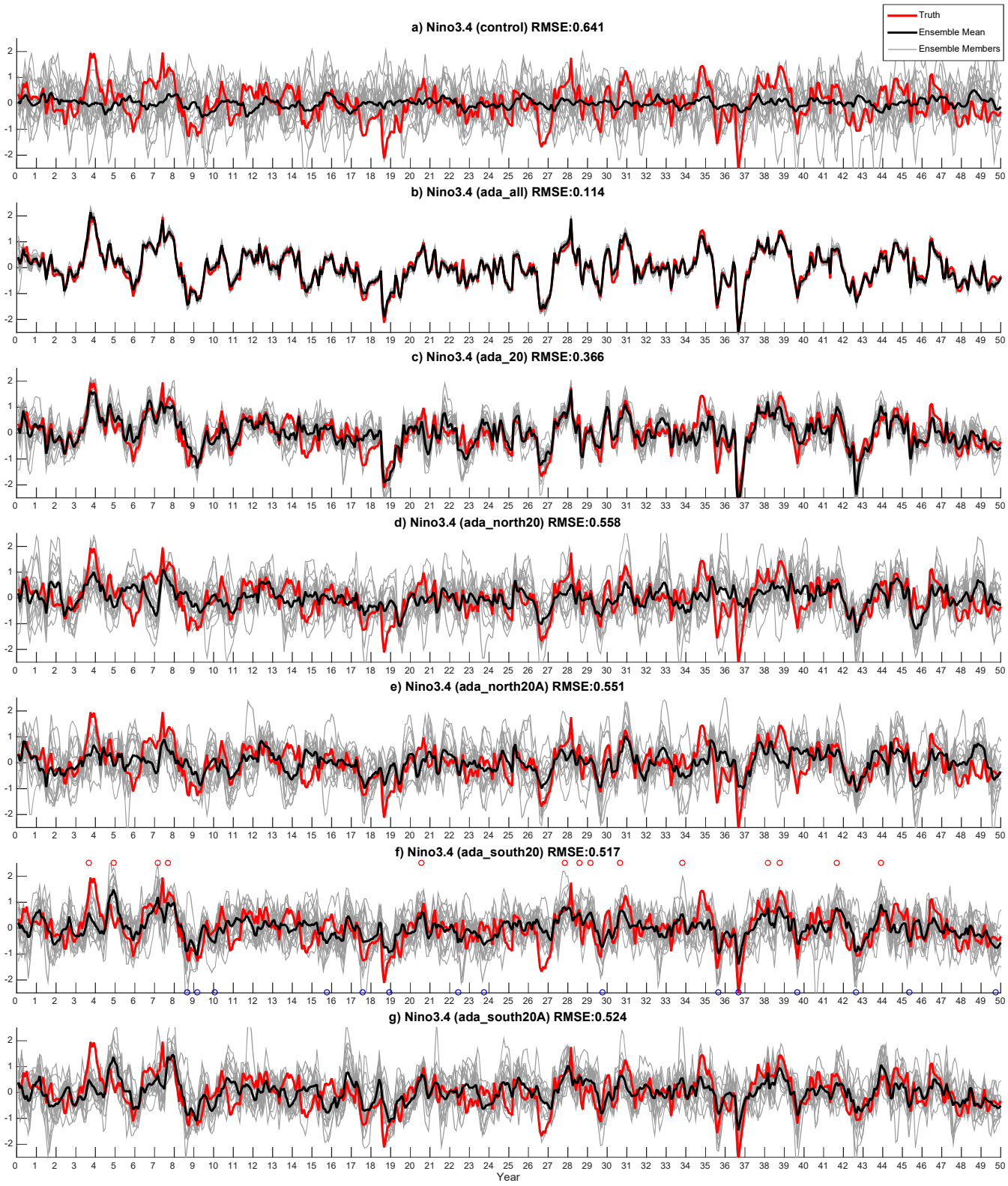

Fig. 4 The time series of Nino3.4 index from (a) *CTRL*, (b) *ada_all*, (c) *ada_20*, (d) *ada_north20*, (e) *ada_north20A*, (f) *ada_south20*, and (g) *ada_south20A*. Red lines indicate the "truth" (same for all panels), black lines the ensemble mean, and grey lines all 16 ensemble members.

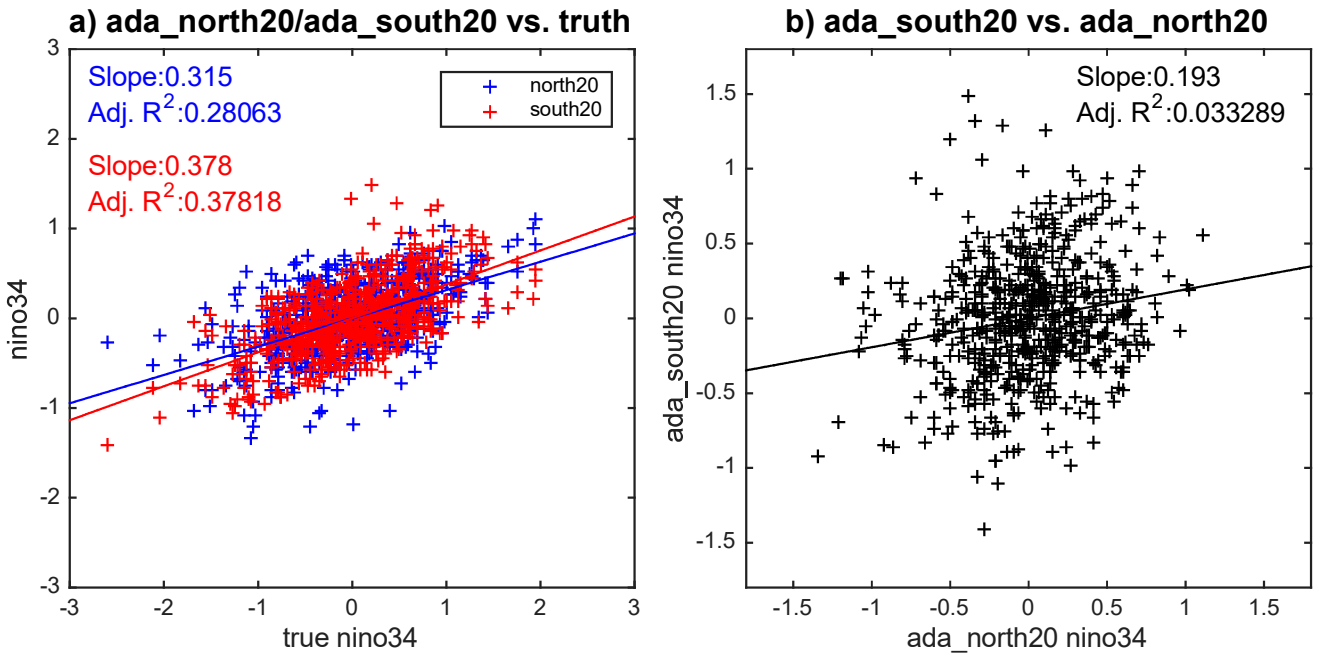

Fig. 5 Scatter plots with linear regression slope coefficient and adjusted R^2 for the Nino3.4 indices of (a) *ada_north20* (blue) and *ada_south20* (red) vs. truth, (b) *ada_south20* vs. *ada_north20*.

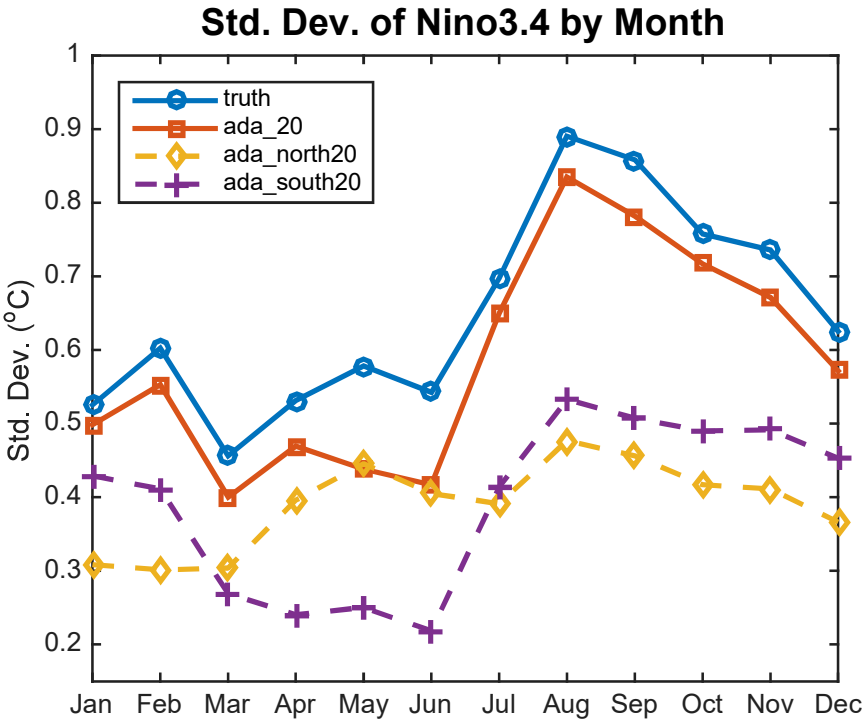

Fig. 6 Standard deviations of the Nino3.4 index by calendar month from the *truth*, *ada_20*, *ada_north20* and *ada_south20*.

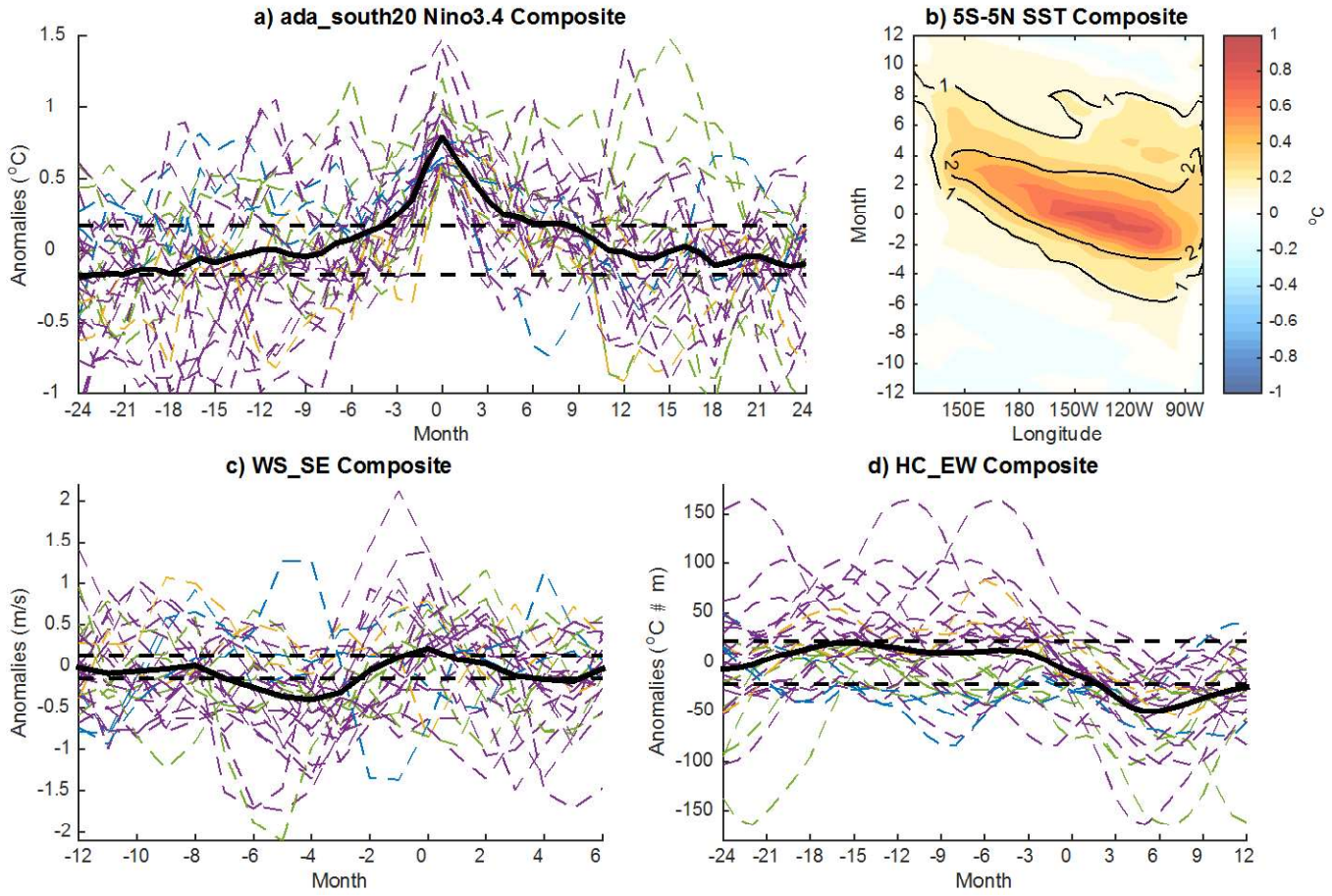

Fig. 7 For *ada_south20*: (a) Composite of Nino3.4 index of 29 ENSO events exceeding 0.5°C. La Nina events are included with reversed sign. Dashed lines indicate indicate CTRL_SD, the standard deviations of the ensemble-mean anomalies from *CTRL*. (b) Composite of 5°S-5°N averaged Pacific SST based on the Nino3.4 peaks (shadings) and its ratio to the corresponding CTRL_SD (contours). (c) Same as Fig. 7a but for the WS_SE series. (d) Same as Fig. 7a but for the HC_EW series.

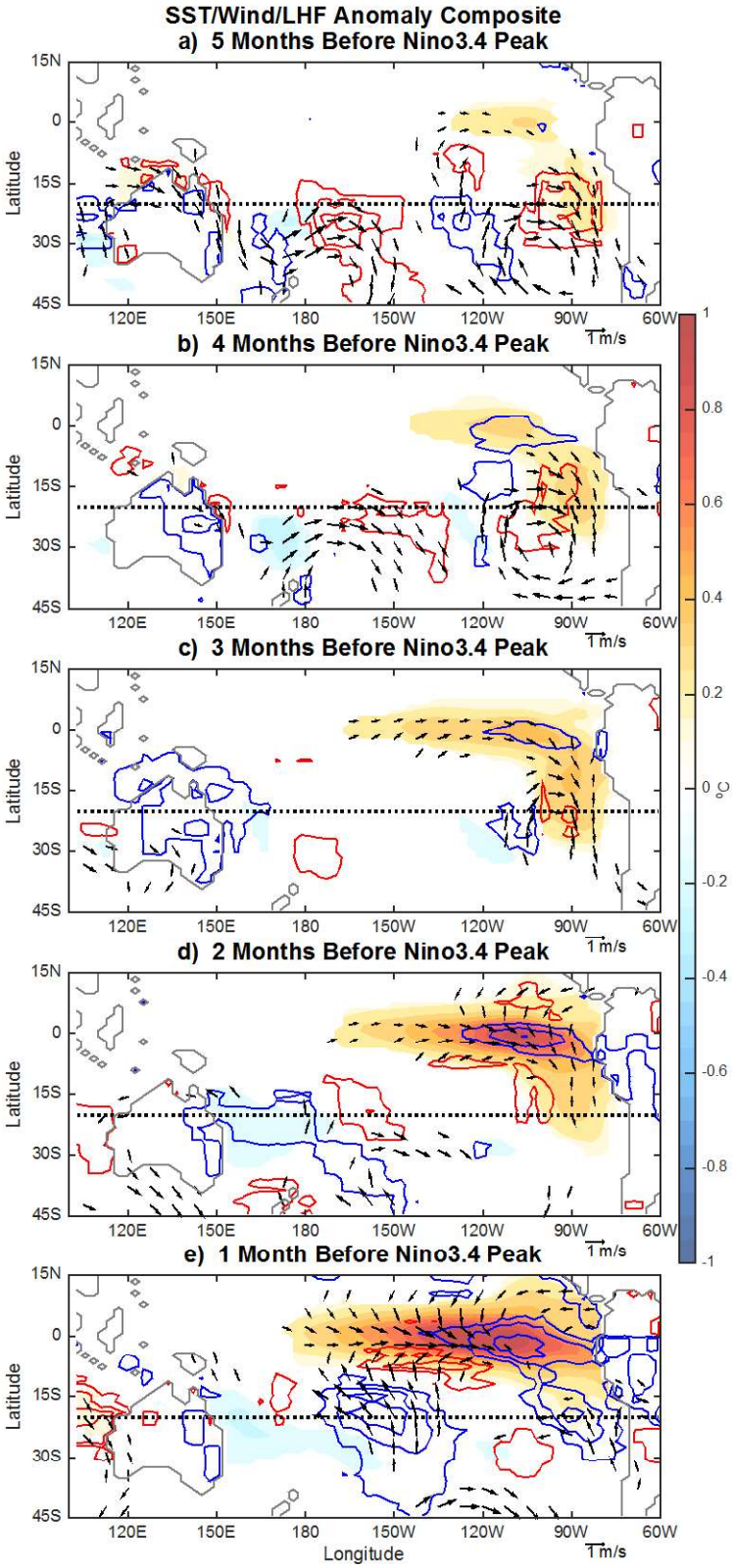

Fig. 8 Composites of anomalous SST (shadings, K), LHF (contours, 4 W/m^2 interval and zero line omitted) and wind (arrows) for the 5 months prior to the peak of the Nino3.4 composite in *ada_south20*. Blue (red) contours indicate more (less) upward LHF. Black dotted lines indicate the boundary of data assimilation. SST and wind anomalies are only shown where the composite exceed CTRL_SD.

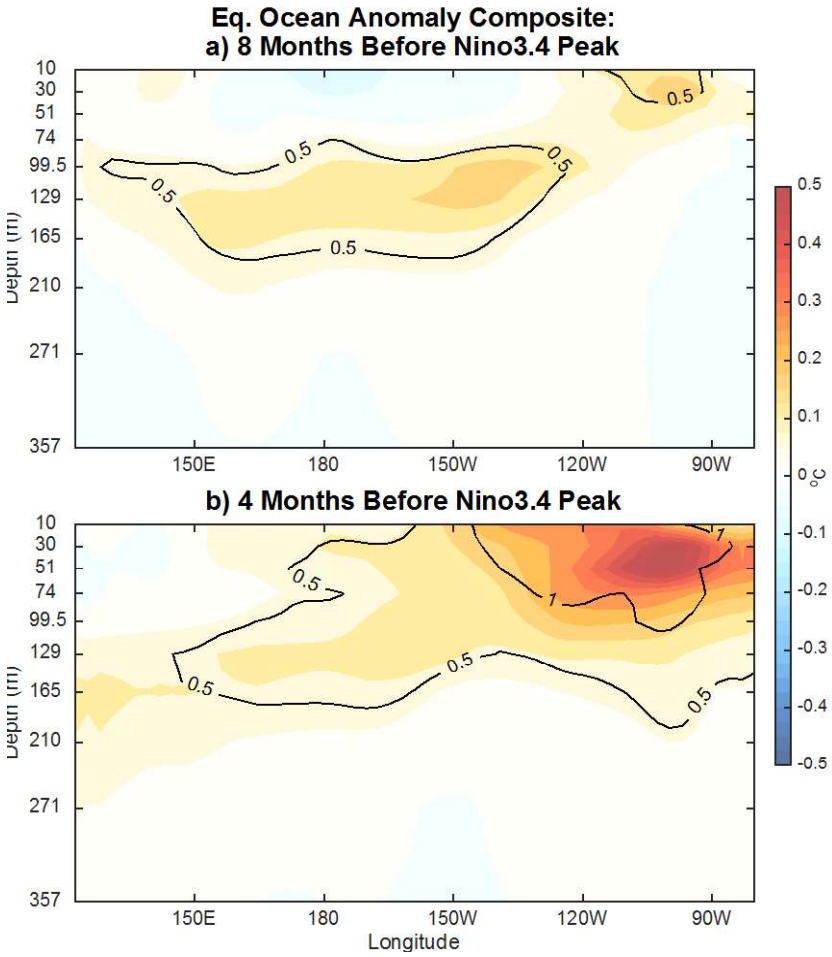

Fig. 9 Composite of anomalous 3°S-3°N averaged Pacific upper ocean temperature (a) 8 months and (b) 4 months prior to the peak of the Nino3.4 composite in *ada_south20*. The contours indicate the ratios of the composite anomalies to CTRL_SD.

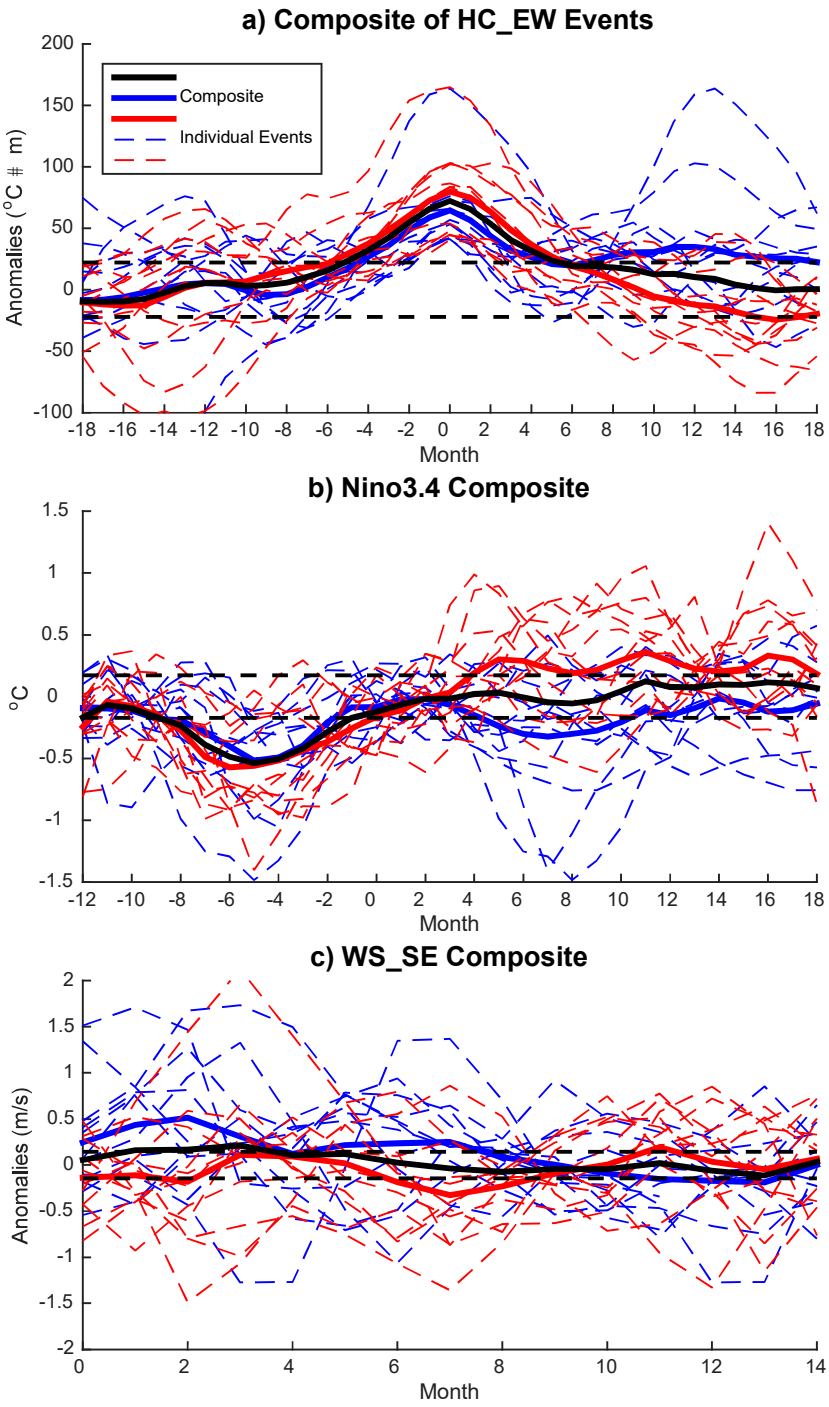

Fig. 10 (a) Composite of HC_EW series for the 24 strongest HC_EW events. Negative anomalies are included with reversed sign. The HC_EW events that are followed by ENSO events of the same sign within 16 months are marked red, and the rest are marked blue. The blue (red) solid line is the average of all blue (red) dashed lines. (b) Composite of Nino3.4 index in *ada_south20* for the same events as (a); (c) Same as (b), but of WS_SE.

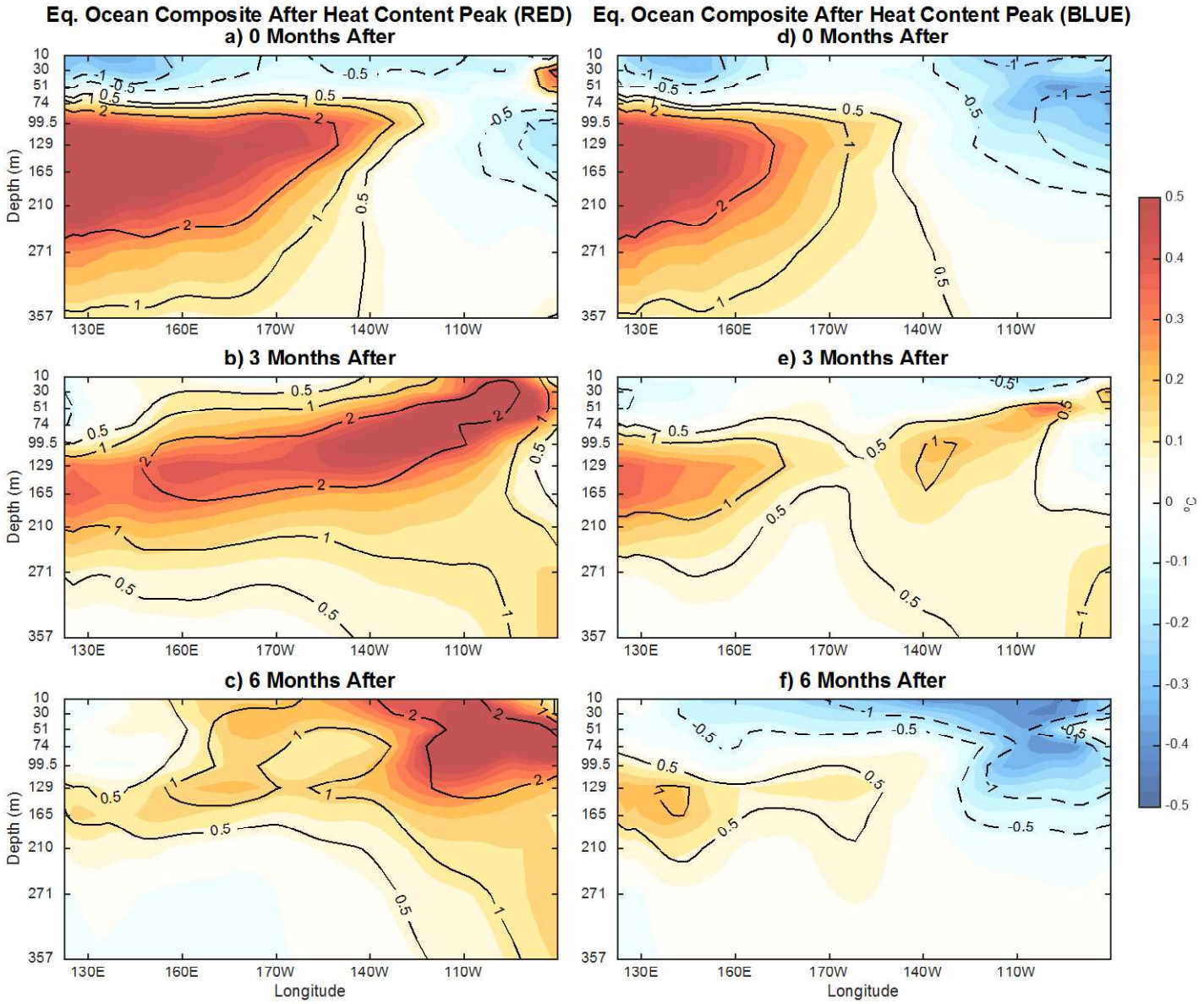

Fig. 11 Same as Fig. 9, but for the "red" (left column, a-c) and "blue" (right column, d-f) HC_EW events from Fig. 10, respectively. The composites are at the same time as the peak of HC_EW composite (a, d), 3 months after (b, e), and 6 months after (c, f).

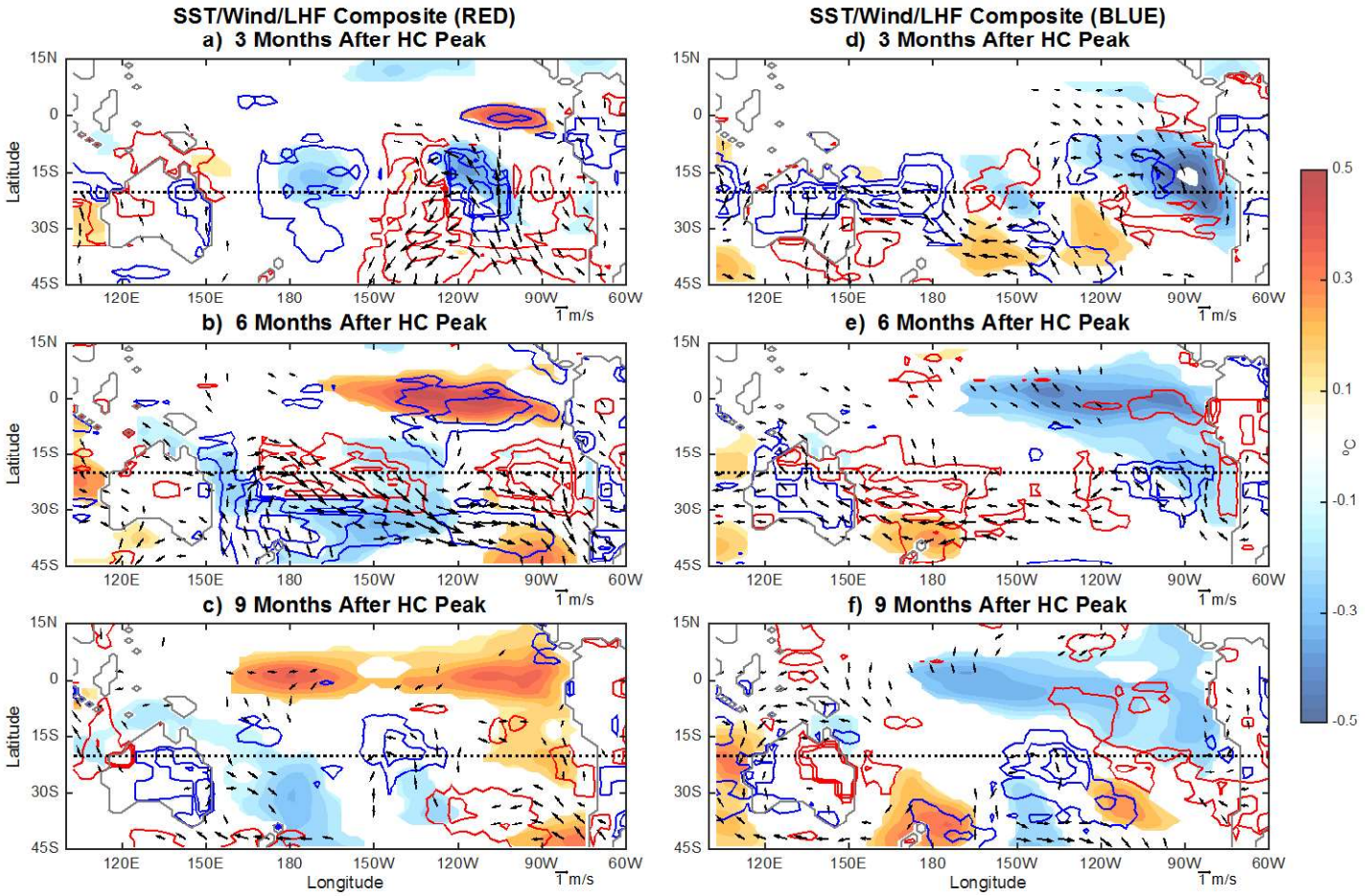

Fig. 12 Same as Fig. 8, but for the "red" (left column, a-c) and "blue" (right column, d-f) HC_EW events from Fig. 10, respectively. The composites are at 3 months (a, d), 6 months (b, e) and 9 months (c, f) after the peak of HC_EW composite.

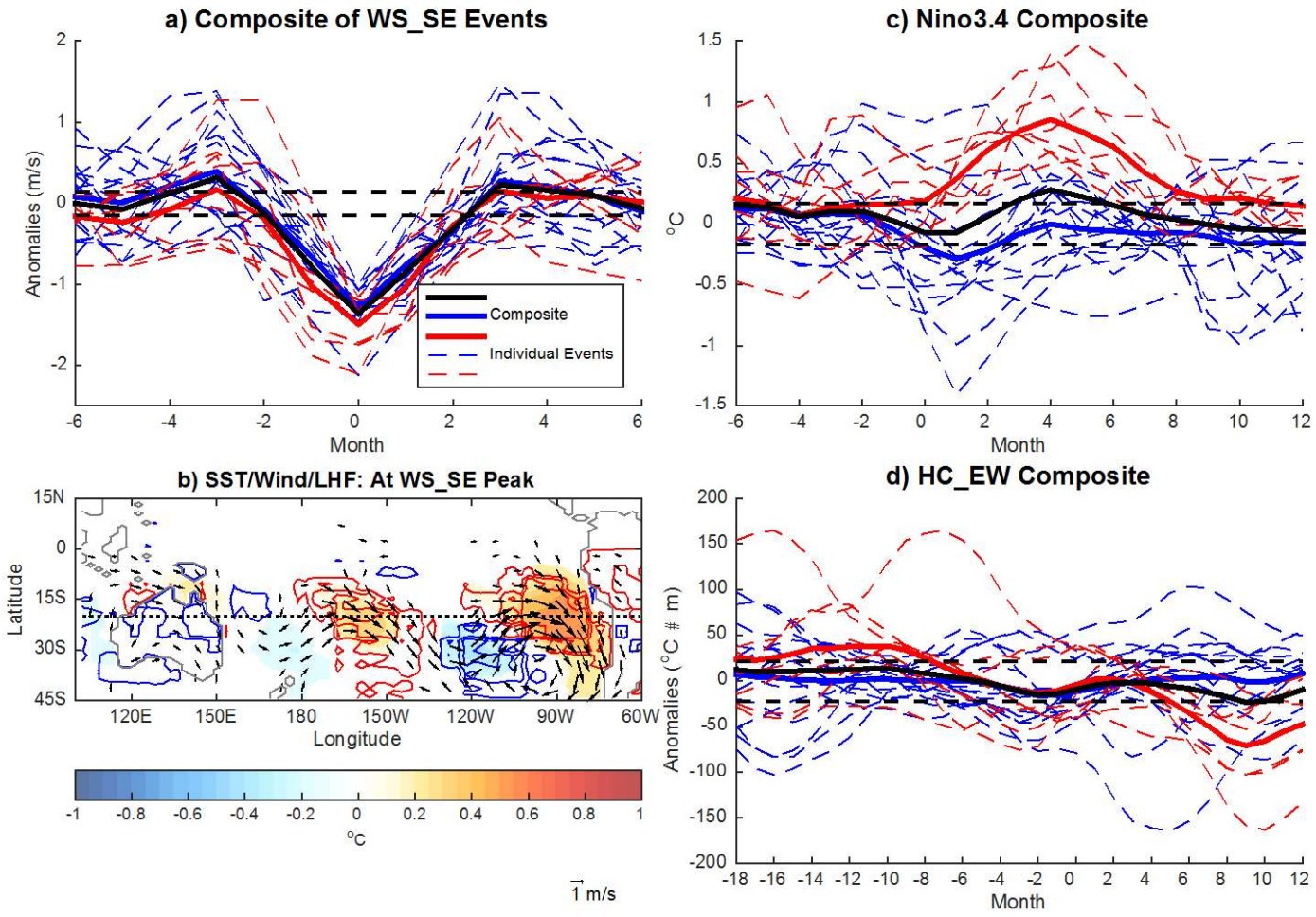

Fig. 13 (a) Composite of WS_SE anomalies for the 24 strongest WS_SE events. Positive anomalies are included with reversed sign. The WS_SE events that are followed by ENSO events of opposite sign within 6 months are marked red, and the rest are marked blue. The blue (red) solid line is the average of all blue (red) dashed lines. (b) Same as Fig. 8, but for all the WS_SE events at the peak. (c) Composite of Nino3.4 index in *ada_south20* for the same events as (a); (d) Same as (c), but of HC_EW.

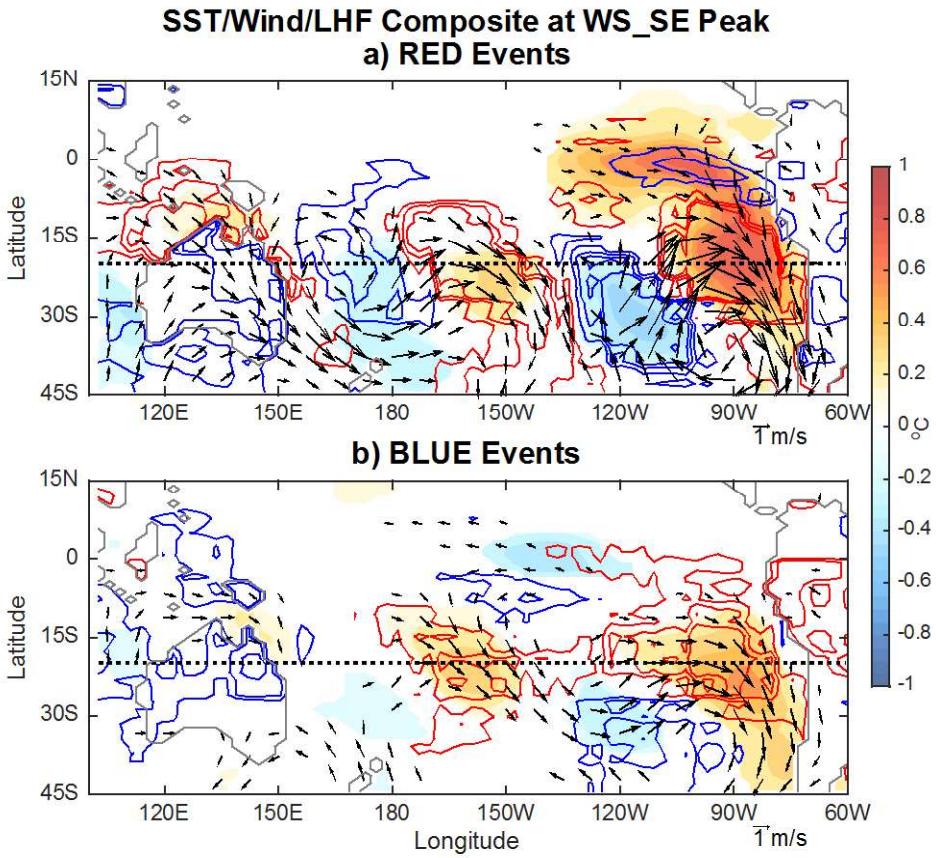

Fig. 14 Same as Fig. 13b, but for (a) “red” and (b) “blue” events, respectively.

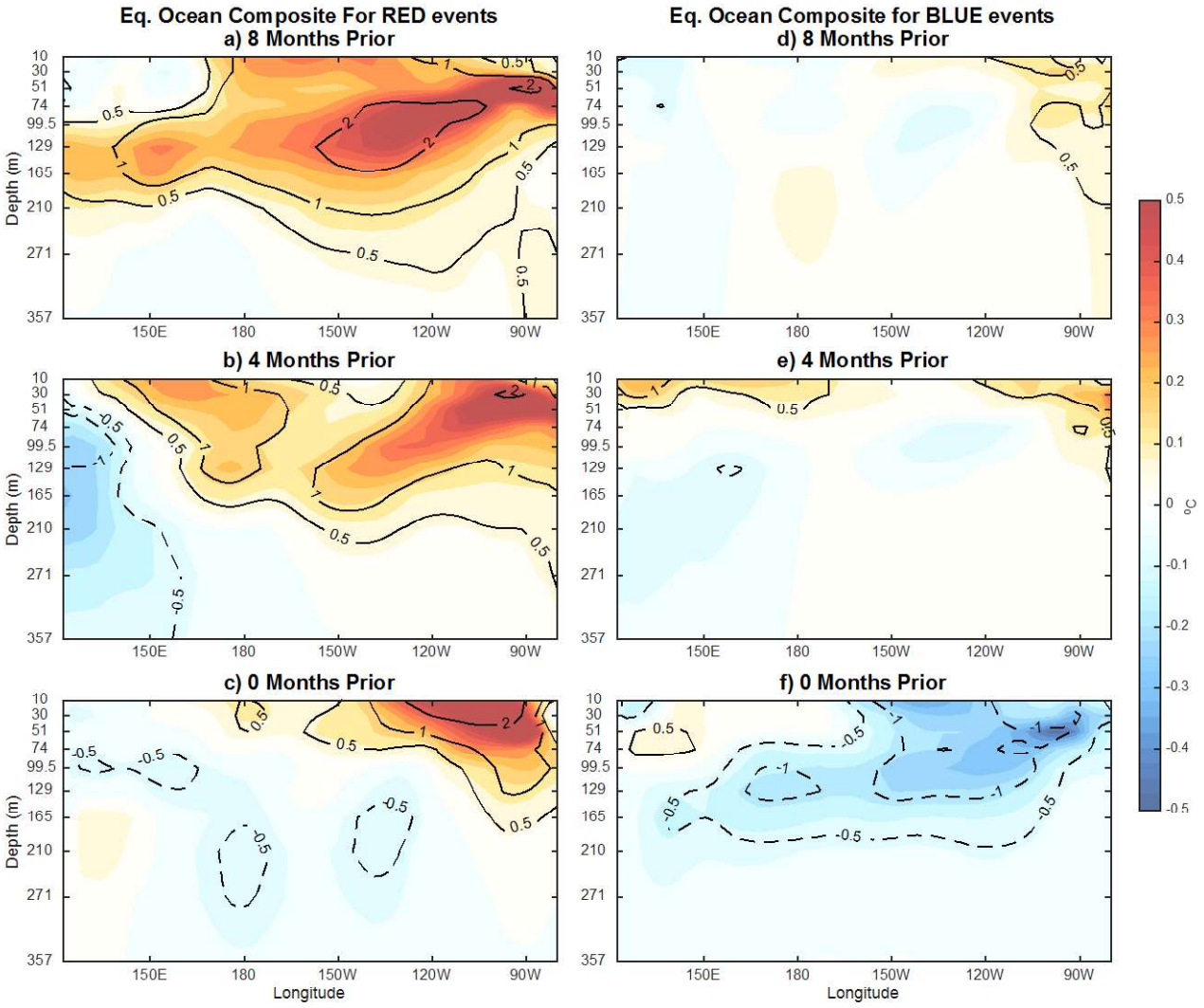

Fig. 15 Same as Fig. 9, but for the "red" (left column, a-c) and "blue" (right column, d-f) WS_SE events from Fig. 13a, respectively. The composites are 8 months (a, d), 4 months (b, e) prior to, and at the same month (c, f) as the peak of the WS_SE composite.

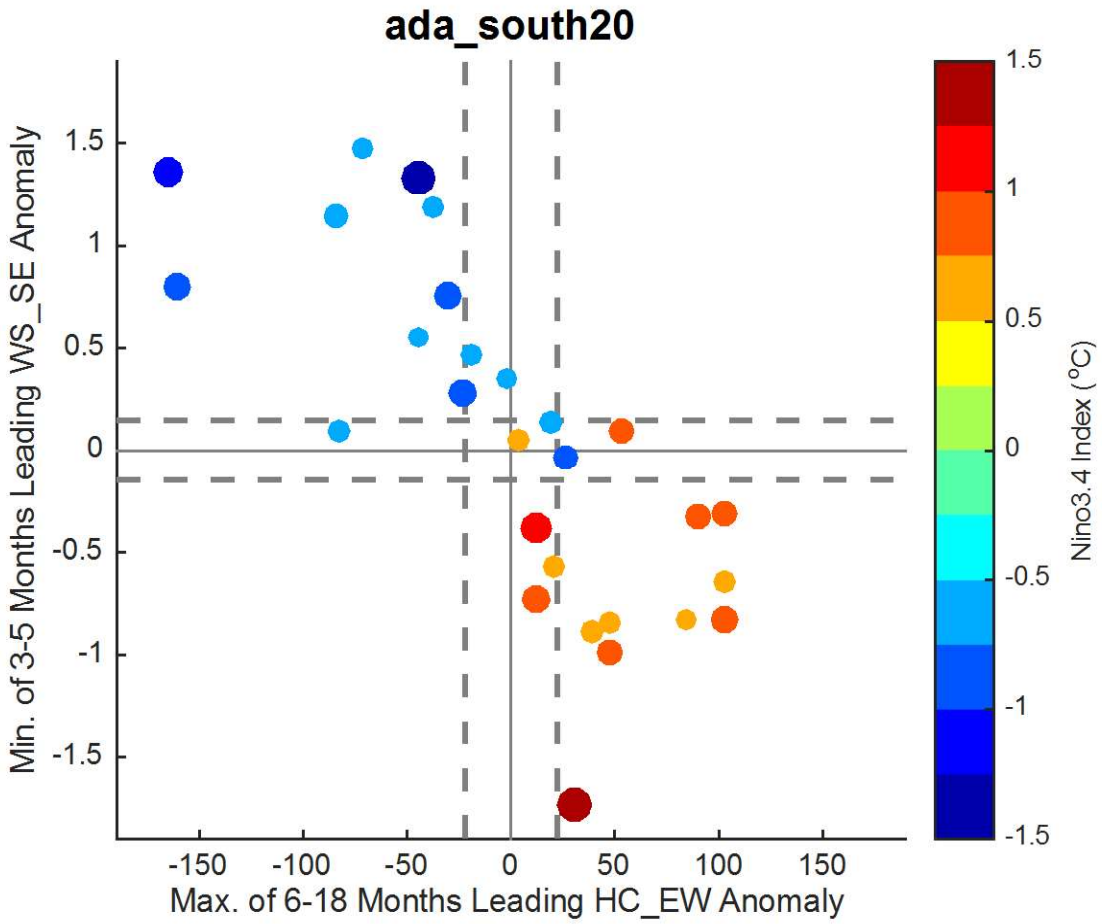

Fig. 16 Scatterplot of all 29 ENSO events in *ada_south20* based on the minimum opposite-sign value of 3-5 months leading WS_SE (y-direction) and maximum same-sign value of 6-18 months leading HC_EW (x-direction). The color and size of the markers together indicate the peak Nino3.4 values. The dashed lines indicate the corresponding CTRL_SD.

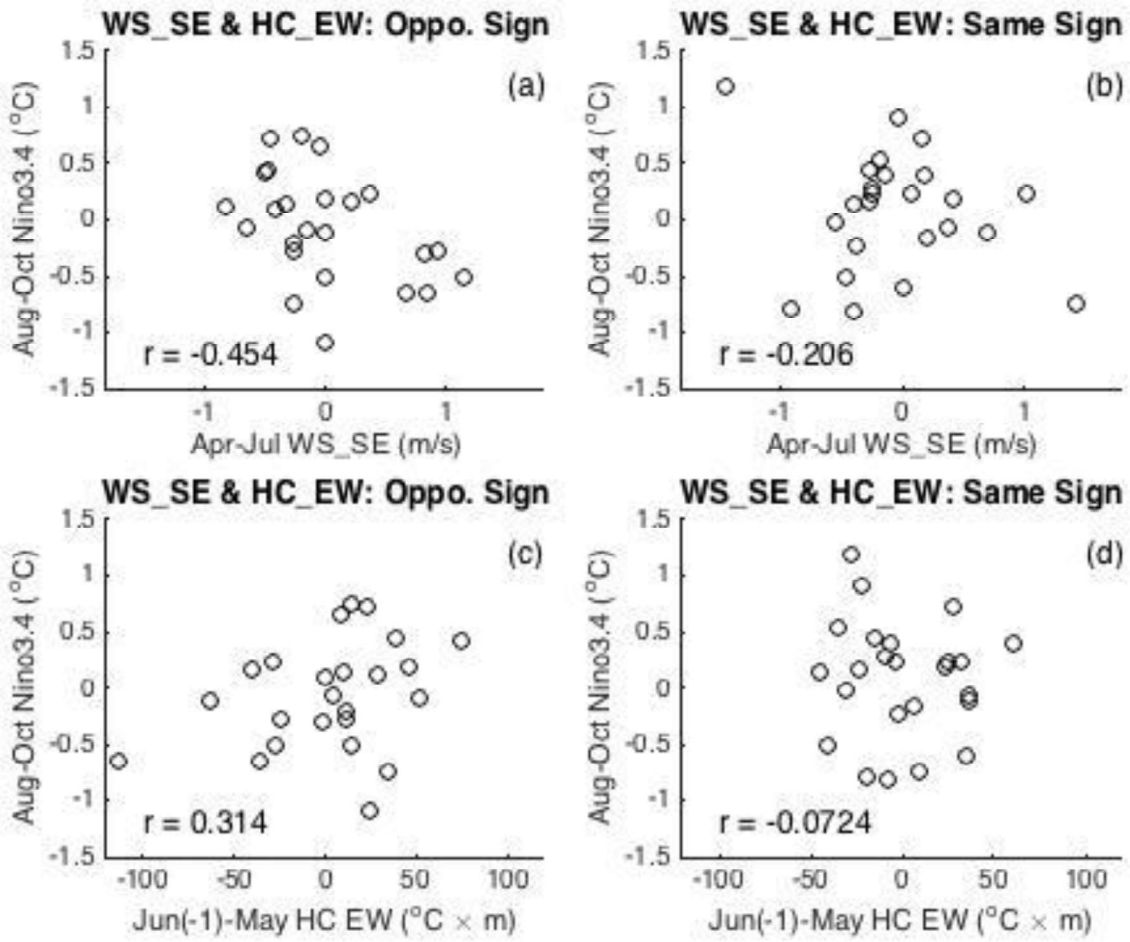

Fig. 17 (a) Scatterplot of the April-July averaged WS_SE with August-October averaged Nino3.4 index of the same year, plotted for the years when the averaged WS_SE has the opposite sign as the June (previous year)-May (same year) averaged HC_EW. (b) Same as (a), but for the years when the averaged WS_SE has the same sign as the averaged HC_EW. (c) Scatterplot of the June (previous year)-May (same year) averaged HC_EW with August-October averaged Nino3.4 index, plotted for the years when the April-July averaged WS_SE has the opposite sign as the averaged HC_EW. (d) Same as (c), but for the years when the averaged WS_SE has the same sign as the averaged HC_EW.