

RESEARCH ARTICLE

Model errors of an intermediate model and their effects on realistic predictions of El Niño diversity

Lingjiang Tao^{1,2} | Wansuo Duan^{2,3}  | Lin Jiang^{2,3}

¹Department of Atmospheric and Oceanic Sciences & Institute of Atmospheric Sciences, Fudan University, Shanghai, China

²LASG, Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing, China

³University of Chinese Academy of Sciences, Beijing, China

Correspondence

Wansuo Duan, LASG, Institute of Atmospheric Physics, Chinese Academy of Sciences, 100029 Beijing, China.
Email: duanws@lasg.iap.ac.cn

Funding information

China Postdoctoral Science Foundation, Grant/Award Number: 2020M681155; National Natural Science Foundation of China, Grant/Award Numbers: 41525017, 41930971

Abstract

In the present study, a nonlinear forcing singular vector (NFSV)-based data assimilation approach is adopted to quantify the model errors of an intermediate coupled model (ICM) and their effects on El Niño predictions. Then, the tendency errors of the NFSV structure (NFSV-TEs) that represent the combined effect of different kinds of model errors are determined in terms of the sea surface temperature (SST) anomaly component in the El Niño predictions. The NFSV-TEs exhibit large values over the eastern equatorial Pacific and on model boundaries, indicating that large model errors exist there. In addition, two dominant NFSV-TEs are found: one is E-type that NFSV-TEs are mostly located in the far eastern Pacific, and the other is the D-type that presents positive anomalies in the eastern equatorial Pacific (EQeast-TEs) and negative anomalies in the central equatorial Pacific (EQcenter-TEs). The D-type NFSV-TEs often occur during realistic predictions of El Niño events. Simulations using the ICM equipped with the NFSV-TEs are then implemented to study the effect of NFSV-TEs on the ENSO predictions. It is shown that the ICM forced by the EQeast-TEs shows better performance in reducing prediction errors and systematic bias, while the EQcenter-TEs forcings are superior to the EQeast-TEs forcings in reproducing the horizontal distribution of the SST anomaly, especially in identifying the difference between the central tropical (CP) and eastern tropical (EP) El Niño. This is because EQcenter-TEs forcings can adjust not only the wind but also the ocean processes to yield realistic air-sea conditions favouring CP-El Niño formulations. Therefore, to make a better prediction of CP-El Niño, the model uncertainties occurring in the central tropical Pacific should be considered preferentially and finally removed in realistic predictions of El Niño diversity.

KEYWORDS

data assimilation, ENSO diversity, intermediate coupled model, nonlinear forcing singular vector

1 | INTRODUCTION

The El Niño-Southern Oscillation (ENSO) is widely known as a short-term climate phenomenon with high

global climate impacts (Philander, 1983). During El Niño years, the anomalous warming sea surface temperature (SST) in the tropical Pacific typically changes the atmospheric circulation and redistributes precipitation

(Ropelewski and Halpert, 1987), thereby tremendously influencing the social economy (Hansen *et al.*, 1998; Shuai *et al.*, 2013). The importance of understanding and predicting ENSO phenomena is therefore self-evident (Solow *et al.*, 1998). Fortunately, since El Niño was first successfully predicted by a dynamic model during the 1980s (Zebiak and Cane, 1987), an increasing number of general coupled climate models with various complexities have been developed to explore the Earth system and predict ENSO events (see website at <https://iri.columbia.edu/our-expertise/climate/forecasts/enso/current/>).

Although great efforts have been made to perfect ENSO-related climate models, unsatisfactory problems still exist in realistic predictions of ENSO. One well-known issue is the so-called ‘spring predictability barrier (SPB)’, in which almost all models show the poorest grades in ENSO predictions across boreal spring (Duan and Mu, 2018). Previous studies projected the SPB problem onto the initial conditions and stochastic processes that the climate models unsolved (Mu *et al.*, 2007; Duan and Hu, 2016; Hou *et al.*, 2019; Tan *et al.*, 2020). Hence, advanced data assimilation approaches are widely adopted to equip climate models to yield optimal initial fields (Chen *et al.*, 2004; Zheng *et al.*, 2007; Gao *et al.*, 2016; O’Kane *et al.*, 2019). By generating different samples, the ensemble mean is usually performed to remove unpredictable processes as well as the effect of stochastic processes (Liu *et al.*, 2019).

Since the former efforts are made at the level of the individual model, the improvement of ENSO prediction skill is largely controlled by the model itself. Although intermediate coupled models capture the main processes of ENSO and have skills in ENSO predictions, the improvement of ENSO prediction is limited due to complex model errors. For general coupled models, undeniable biases still exist in the simulation of climatology with current models (Wang *et al.*, 2014), which could contaminate the interannual simulation, for example, ENSO phenomena. For instance, the negative SST biases that are evident along the central Pacific make models hardly simulate the central tropical Pacific El Niño (CP-El Niño) events in which the SST anomalies are centred in the central tropical Pacific (Leloup *et al.*, 2008; Ham and Kug, 2012). Most models can describe only one variation of El Niño, that is, the eastern tropical Pacific El Niño events (EP-El Niño), of which the warming peak is located in the eastern equatorial Pacific. One possible reason is that the climatological zonal SST gradient over the western central tropical Pacific simulated by the models is not as strong as in observations, giving rise to weak zonal advection feedback (Feng *et al.*, 2020). Note that zonal advection feedback is an important ocean process in central tropical Pacific warming, and the CP-El

Niño, therefore, fails to be present in the models with weak zonal advection. It is conceivable that the prediction of CP-El Niño is unsatisfactory using these models (Xue *et al.*, 2013; Ren *et al.*, 2018). Thus, models usually lose skills in predicting the type of El Niño that will occur when the prediction time is larger than 3 months (Tao *et al.*, 2020). This is another issue in realistic predictions of ENSO, save for the previously mentioned SPB phenomena.

Compared with the initial condition errors, one that should be highlighted is that the role of model errors is equally even much more important in ENSO predictions as the lead time increases (Zheng *et al.*, 2009; Tao and Duan, 2019). It is therefore crucial to study the relationship between model errors and prediction error growth. Focusing on the uncertainties in model parameters, Mu *et al.* (2010) introduced a conditional nonlinear optimal parametric perturbation (CNOP-P) approach to identify the most unstable parametric perturbation in prediction. It considers the nonlinearity of climate events and thus has the ability to quantify the nonlinear effect of model parameter uncertainties on error growth in prediction. Based on the CNOP-P approach, subsequent studies pointed out that the SPB issue is also attributable to the uncertainties in the model parameters (Tao *et al.*, 2019) and thus theoretically proved the fact that optimizing the model parameters can reduce the SPB phenomena and prolong the valid prediction time (Wu *et al.*, 2016).

Not just originated from the parameter errors, uncertainties in models are of multiple origins and interacting with each other. It is therefore unrealistic to study the impacts of each source of model error. To better untangle the model error effects, one feasible approach is to consider the combined effect of model errors with various sources (Zheng and Zhu, 2016). The nonlinear forcing singular vector (NFSV) approach proposed by Duan and Zhou (2013) is the method. The NFSV approach is used to detect the most growing type of the model tendency perturbation that can induce the largest prediction error growth, where the model tendency perturbation takes all responsibility that causes model uncertainties. The NFSV approach has been successfully applied to weather and climate events to study their predictability (Duan *et al.*, 2016; Tao and Duan, 2019; Tao *et al.*, 2020; Yao *et al.*, 2021). For instance, with an ENSO model [i.e. developed by Zebiak and Cane, 1987], Duan *et al.* (2016) found that NFSV-related model tendency perturbations (hereafter referred to as NFSV-TPs) of the SST model that present negative values in the eastern equatorial Pacific and positive values in the central Pacific could destroy the El Niño prediction to the largest extent. The prediction errors caused by those NFSV-TPs are of prominent seasonality with the fastest error growth in spring, indicating

that reducing NFSV-like model errors is the key to solving the SPB phenomena during ENSO prediction. NFSV-TPs are often concentrated in a specific area (i.e. the central-eastern tropical Pacific), which may suggest sensitive areas of ENSO prediction associated with model tendency errors. That is, the model tendency uncertainties in the central-eastern tropical Pacific play dominant roles in ENSO prediction errors as well as SPB-like error growth. The study with the NFSV approach opens up a new path for improving ENSO predictions (e.g. Tao and Duan, 2019).

However, the shortcoming of the former study is that the NFSV-TPs derived are inconstant and stem from an idea experiment where the true state is generated by the model alone. Whether this kind of NFSV-TP exists and influences realistic predictions is still unknown. In addition, one may question whether the model tendency errors in the central-eastern tropical Pacific are truly an important source of prediction errors in practice, especially for different types of El Niño events. To this end, in the present study, a series of sensitivity experiments regarding ENSO were performed during the period from 1950 to 2010 to demonstrate the impacts of the tendency errors. To quantify the model tendency errors in ENSO realistic predictions, first, the NFSV-based assimilation that deals with the model errors by assimilating observation data was adopted to obtain the tendency errors that represent the total effect of model errors with various sources (Duan *et al.*, 2014; Tao *et al.*, 2020). We find that the tendency errors are usually located in the central and eastern equatorial Pacific, of which type is quite similar to the results from Duan *et al.* (2016). Then, the main scientific issues that will be answered in the present study are as follows: What are the model errors look like in realistic predictions? What are the different effects of local tendency errors on El Niño diversity prediction? How do they influence the El Niño predictions?

The rest of this article is organized as follows. Section 2 gives descriptions of the ENSO prediction model and the NFSV-based assimilation approach. Section 3 illustrates the NFSV assimilation-determined tendency errors (NFSV-TEs) with the ENSO model. In Section 4, the impact of local tendency errors on the predictions of the ENSO-related SST anomaly and different types of El Niño events are demonstrated. Section 5 is narrowed down to reveal the dynamics involved with the evolution of the predicted CP-El Niño. The former results are concluded and discussed in Section 6.

2 | MODELS AND METHODS

In this section, the ENSO prediction model used in this study is first reviewed briefly. Then, the NFSV-based data

assimilation approach that is used to quantify the model uncertainties in ENSO predictions is introduced.

2.1 | ENSO prediction model and initialization scheme for ENSO prediction

The ENSO model adopted in the present study is an intermediate complexity model (ICM) developed by Zhang *et al.* (2003). It is an air–sea coupling regional model that covers only the tropical Pacific. Three main submodels make up the ICM: a dynamical intermediate ocean model (IOM), an empirically determined wind anomaly model, and an SST anomaly model. The wind model is constructed based on the highly coupled SST and wind fields (denoted as τ), which can be expressed as

$$\tau = \mathbf{g}(\mathbf{x}), \quad (1)$$

in which $\mathbf{x}$ is the SST anomaly and $\mathbf{g}$ denotes the empirical relationship between the SST anomaly and τ based on historical data. The SST anomaly model representing the thermodynamics over the surface mixed layer is a fully nonlinear model, which can be symbolically written as

$$\mathbf{x}_t = \mathbf{F}(\mathbf{x}_0, \mathbf{m}, t), \quad (2)$$

where $\mathbf{m}$ includes the dynamic ocean field (i.e. ocean current in the mixed layer and sea level anomaly) from the IOM and the wind anomalies derived from Equation (1), and $\mathbf{F}$ denotes the nonlinear propagator of the government equation. With the ocean current anomalies offered by the IOM and the wind stress anomaly provided by the wind model, the SST anomaly can be updated in each step with the SST anomaly model. In the ICM, the simulated SST anomaly evolutions can be well-explained by the Bjerknes feedback and the charge–discharge mechanisms. Therefore, the ICM has skills in ENSO prediction in practice (Zhang and Gao, 2016). The ICM, termed IOCAS ICM, has recently been included by the International Research Institute for Climate and Society (IRI) as one member of real-time ENSO prediction products (see website at <https://iri.columbia.edu/our-expertise/climate/forecasts/enso/current/>). For convenience, more details with regard to the IOM and ICM are provided in Keenlyside and Kleeman (2002) and Zhang *et al.* (2005), respectively.

In the experiments, we adopt the same initialization scheme as Zhang *et al.* (2005) to initialize the ICM before the prediction is made. First, the observed monthly SST anomalies are utilized to generate the wind field by Equation (1) from 1854 to the time when the prediction is started. Then, the reconstructed wind field forces the

IOM and the SST anomaly model to generate the initial conditions of the dynamic and thermodynamic fields. Finally, the observed SST anomaly at the start time is brought into the SST model to generate more realistic initial conditions. Since this procedure effectively avoids the ‘initial shock’ and balances the model and observations, the ICM shows high performance in ENSO predictions within several-month lead times (Zhang *et al.*, 2005). In the present study, the observed monthly SST field, from the third version of the National Oceanic and Atmospheric Administration (NOAA) Extended Reconstructed SST (ERSSTv3; Smith *et al.*, 2008), is the only data used to initialize the ICM and estimate the prediction skills.

2.2 | NFSV-based assimilation and its application to the ICM

2.2.1 | NFSV and NFSV-assimilation

The NFSV approach first proposed by Duan and Zhou (2013) is a nonlinear extension of forcing singular vector. The basic idea of NFSV is to search an optimal tendency perturbation (referred to as NFSV-TP) that can make the prediction depart the most from the reference state in a nonlinear climate/weather system under the assumption of the perfect initial condition. Mathematically, the calculation of NFSV-TP $\mathbf{f}^*$ is equally to solving the following maximum problems:

$$J(\mathbf{f}^*) = \max_{\mathbf{f} \in \Omega} \|\mathbf{M}_t(\mathbf{x}_0, \mathbf{f}) - \mathbf{M}_t(\mathbf{x}_0, 0)\|, \quad (3)$$

where $\mathbf{M}_t(\mathbf{x}_0, 0)$ is the reference states generated by the original model without tendency perturbation at t integral time from initial field $\mathbf{x}_0$, and $\mathbf{M}_t(\mathbf{x}_0, \mathbf{f})$ is the perturbed result by the model tendency perturbation $\mathbf{f}$, and $\mathbf{f} \in \Omega$ denotes the constraint condition of the tendency perturbation; $\|\cdot\|$ is an L2 norm square to measure the tendency perturbation-induced error growth. Then, the object function $J(\mathbf{f}^*)$ quantifies the second predictability of the climate/weather event and $\mathbf{f}^*$ denotes the most unstable tendency perturbation when Equation (3) is satisfied.

In practice, the original model we used is an analogue of the climate system that cannot completely describe the earth system. Thus, model uncertainties exist in realistic simulations and predictions. As mentioned in introduction, Duan *et al.* (2014) contributed all model errors to the model tendency errors and proposed an optimal forcing vector (OFV) approach to correct the model and make the simulation closest to the observation. In short, the OFV is obtained by solving the minimum problem:

$$J(\mathbf{f}_0^*, \mathbf{f}_1^*, \dots, \mathbf{f}_t^*) = \min \sum_t \|\mathbf{M}_t(\mathbf{x}_0, \mathbf{f}_t) - \mathbf{x}_{\text{obs}}\|, \quad (4)$$

where $\mathbf{x}_{\text{obs}}$ is the observation. When the object function J is close to 0, the model with the tendency forcings $(\mathbf{f}_0^*, \mathbf{f}_1^*, \dots, \mathbf{f}_t^*)$ can simulate the evolution of the climate in agreement with the observations. In this case, the corrected model $\mathbf{M}_t(\mathbf{x}_0, \mathbf{f}_t)$ can be a ‘perfect’ analogue of the realistic climate system.

Clearly, the OFV approach is an extension application of the NFSV, and is similar to the variational data assimilation technique where the observations are assimilated into the model to obtain the model tendency perturbation and correct the model. In particular, the OFV is another form of the NFSV in mathematics [see details in Tao and Duan, 2019], which means that the Equation (4) can be changed as a form of NFSV:

$$J(\mathbf{f}_0^*, \mathbf{f}_1^*, \dots, \mathbf{f}_t^*) = \max_{\mathbf{f} \in \Omega} \sum_t \|\mathbf{M}_t(\mathbf{x}_0, \mathbf{f}_t) - \mathbf{M}_t(\mathbf{x}_0, 0)\|, \quad (5)$$

where the tendency forcing is constraint in the space of

$$\begin{aligned} \Omega &= \left\{ (\mathbf{f}_0^*, \mathbf{f}_1^*, \dots, \mathbf{f}_t^*) \mid J(\mathbf{f}_0^*, \mathbf{f}_1^*, \dots, \mathbf{f}_t^*) \right. \\ &\quad \left. = \min \sum_t \|\mathbf{M}_t(\mathbf{x}_0, \mathbf{f}_t) - \mathbf{x}_{\text{obs}}\| \right\}, \end{aligned} \quad (6)$$

Thus, for unity, the OFV approach is termed NFSV-assimilation in Tao and Duan (2019). Physically, the NFSV-assimilation is to search for a kind of tendency perturbation that is constrained by the observations and can make the perturbed state departure from the reference state at the greatest extent but be closest to the observation.

The computational procedure of the NFSV and NFSV-assimilation are also quite similar. Figure 1 shows the procedures of NFSV and NFSV-assimilation analyses. As for the NFSV analysis, first, given the initial field $\mathbf{x}_0$, we integrate the model and obtain the reference state $\mathbf{M}_t(\mathbf{x}_0, 0)$. In the meantime, we superimpose the first tendency guess $\mathbf{f}$ that has been adjusted to satisfy the constraint condition of the tendency perturbation into the tendency equation of the model, and forward integrate the perturbed model with the initial field $\mathbf{x}_0$ to generate the perturbed result $\mathbf{M}_t(\mathbf{x}_0, \mathbf{f})$. Then, the object function J is determined and is input into the optimization algorithm (e.g. artificial intelligence, gradient descent method related algorithm). If not meeting the requirement of accuracy, a new tendency $\mathbf{f}$ is generated by the optimization algorithm, which then enters the next loop to calculate the object function again until satisfying the accuracy or exceeding the numbers of pre-set iteration

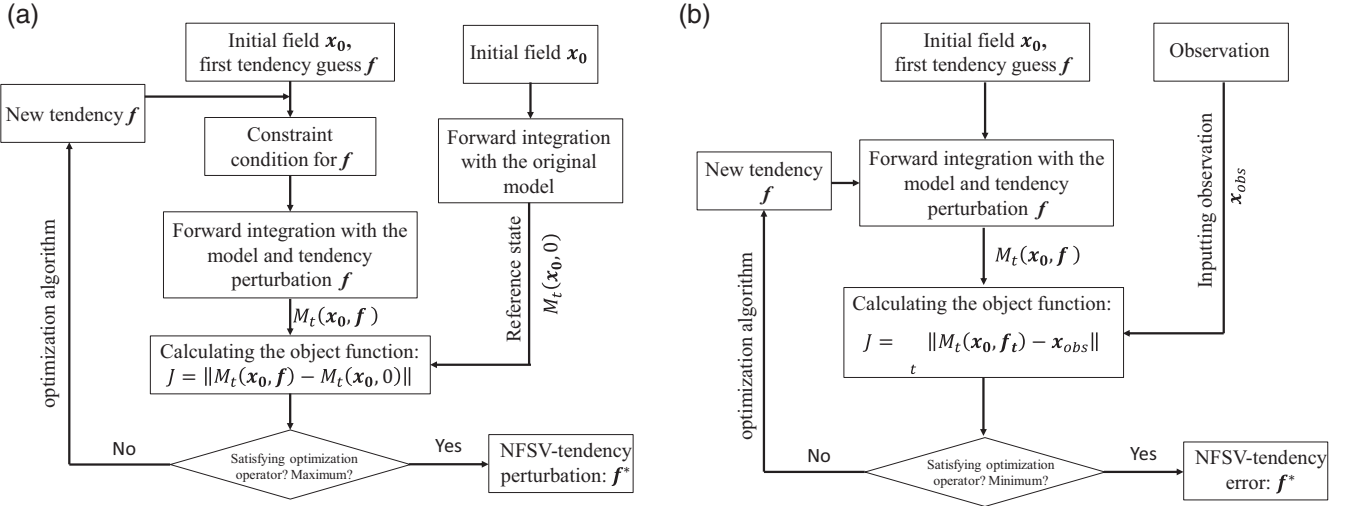


FIGURE 1 A schematic diagram illustrating the (a) NFSV and (b) NFSV-assimilation analyses

steps. As for the NFSV-assimilation analysis, the main difference from the NFSV analysis is that the reference state is replaced by the observation. In addition, the optimization algorithm is used to minimize the object function.

2.2.2 | ICM equipped with the NFSV-assimilation

Section 2.1 shows that the ICM is an artificial system to represent the ENSO-related climate system and depicts only the interannual variability of SST in the tropical Pacific. Thus, inevitable uncertainties exist in the model itself during the prediction due to the high simplification. Here, we apply the previously mentioned NFSV-based assimilation approach to the ICM to determine the model tendency errors in the hindcast of ENSO during the period from 1950 to 2010.

According to the Equation (4), the NFSV-assimilation can be mathematically written as a form of data assimilation scheme:

$$J(\mathbf{f}_0^*, \mathbf{f}_1^*, \dots, \mathbf{f}_t^*) = \min \sum_t (\mathbf{H}\mathbf{x}(\mathbf{f}_t, t) - \mathbf{x}_{\text{obs}}(t))^T \mathbf{R}^{-1} (\mathbf{H}\mathbf{x}(\mathbf{f}_t, t) - \mathbf{x}_{\text{obs}}(t)), \quad (7)$$

in which $\mathbf{H}$ and $\mathbf{R}$ denote the observation operator and observation error covariance, respectively. Here, the $\mathbf{H}$ and $\mathbf{R}$ are identity operators for convenience. $\mathbf{x}(\mathbf{f}_t, t) = \mathbf{M}_t(\mathbf{x}_0, \mathbf{f}_t)$ is the predicted SST anomaly that is the forward integral result of the model with the tendency perturbation $\mathbf{f}_t$, which can also be written as:

$$\mathbf{x}(\mathbf{f}_t, t) = \mathbf{F}_t \cdot \mathbf{x}_{t-\delta t} + \mathbf{f}_t^* \cdot \delta t. \quad (8)$$

where $\mathbf{F}_t$ is the integral operator of the ICM. Clearly, $(\mathbf{f}_0^*, \mathbf{f}_1^*, \dots, \mathbf{f}_t^*)$ is a series of optimal perturbations that can make the simulation closest to the observation when the object function J is close to 0. If the initial condition is perfect, these tendency perturbations $(\mathbf{f}_0^*, \mathbf{f}_1^*, \dots, \mathbf{f}_t^*)$ therefore can describe the model unsolved processes revealed in the tendency equation to some extent when simulating the time series of the observation $\mathbf{x}_{\text{obs}}(t)$. In a sense, the model with the optimized tendency perturbation, that is, $\mathbf{x}(\mathbf{f}_t, t) = \mathbf{F}_t \cdot \mathbf{x}_{t-\delta t} + \mathbf{f}_t^* \cdot \delta t$, can be assumed to be a nearly ‘perfect’ model. Then, from another approach, the negative $(\mathbf{f}_0^*, \mathbf{f}_1^*, \dots, \mathbf{f}_t^*)$ can be regarded as the model tendency errors when a prediction is made with the original model without model tendency perturbation, that is, $\mathbf{x}(0, t) = \mathbf{F}_t \cdot \mathbf{x}_{t-\delta t}$. For convenience, the NFSV assimilation-optimized tendency perturbations are termed NFSV-tendency errors (NFSV-TEs).

In the present study, the tendency perturbations of the SST anomaly model are only considered since we only focus on the SST prediction. It means that $\mathbf{f}_t$ is only added to the SST tendency equation. In addition, we assume that the tendency errors are constant within 1 month both to reduce the computational complexity but also because our main purpose is related to the prediction of the interannual variability. Within a 1-year assimilation window, we can obtain 12 members of optimal tendency forcings for 12 calendar months using the NFSV-assimilation. In practice, these tendency forcings are the mixed production of the initial and model errors. To this end, we performed a composite analysis on these

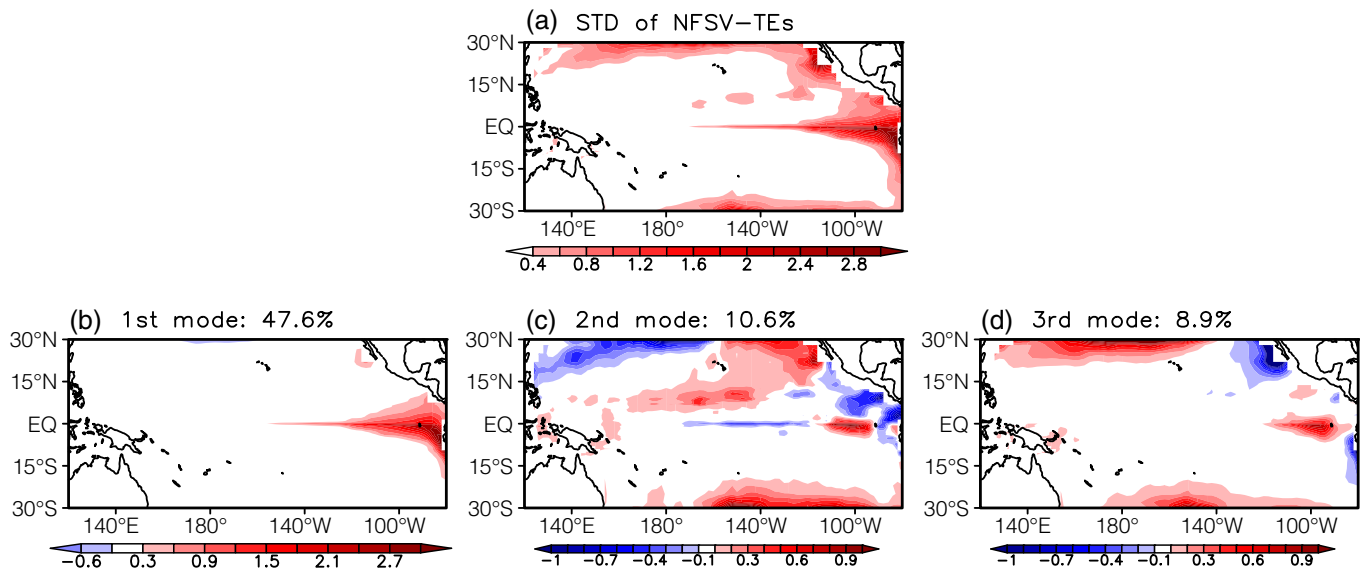


FIGURE 2 The standard deviation of the NFSV-TEs and corresponding three leading EOF modes

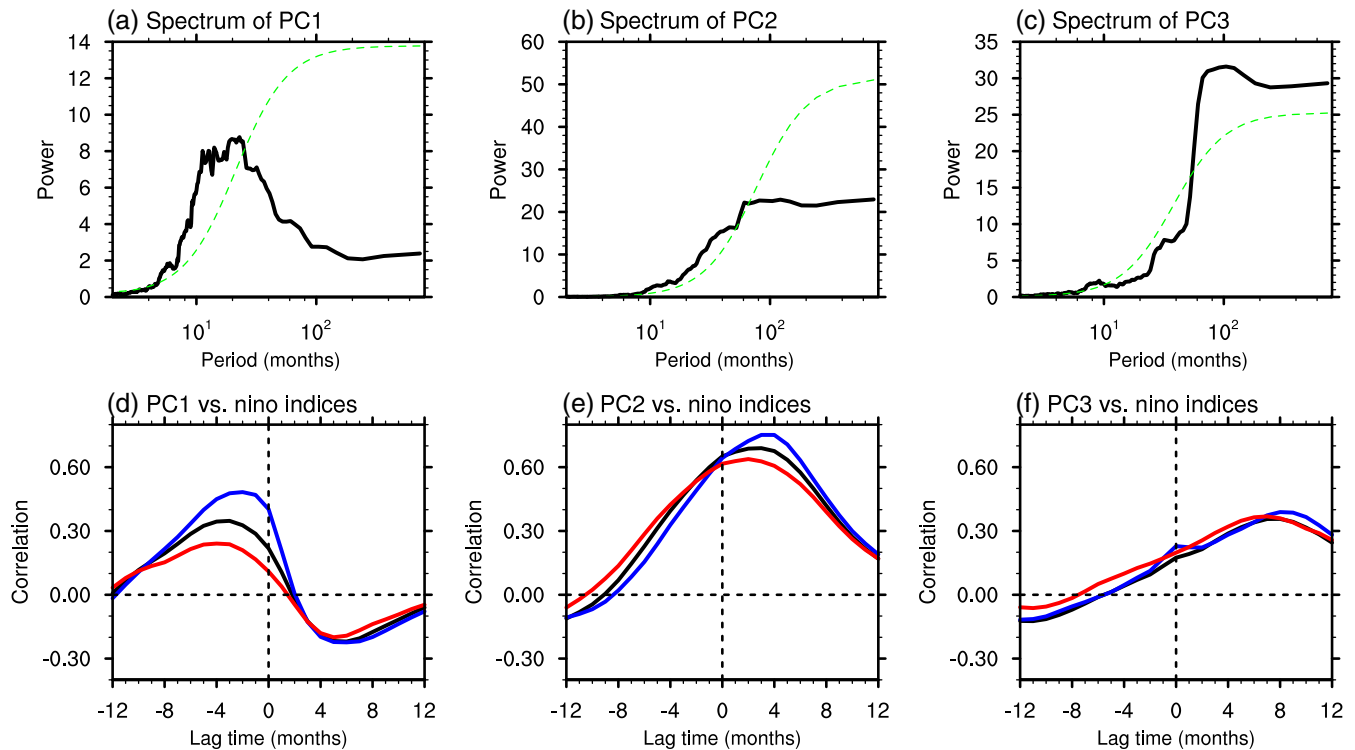


FIGURE 3 Power spectra of the principal components of the NFSV-TEs and their lag-correlations with the Niño indices. The green dotted curves in upper panels indicate the 90% significance level. The black, blue and red curves in bottom panels are lag-correlations with the Niño3.4, Niño3 and Niño4 indices, respectively

tendency forcings whose months overlap in different assimilation windows to filter the effect of initial errors and obtain the ‘pure’ errors from the model. For example, to obtain the NFSV-TEs in December 2000, the

NFSV-assimilations are implemented with various assimilation windows (e.g. 2000.1–2000.12, 2000.2–2001.1, ..., 2000.12–2001.11). Then the relative pure model errors, that is, NFSV-TEs in December 2000, is obtained by a

TABLE 1 Control and four sets of sensitivity experiments during 1950–2010

Experiments	Descriptions of the prediction model used
Control	ICM without tendency perturbation
Sensitivity experiments	
NFSV-TEs	ICM forced by the NFSV-TEs over the whole model grids
EQ	ICM forced by the NFSV-TEs over the central and eastern tropical Pacific (170°W–80°W, 10°S–10°N)
EQeast	ICM forced by the NFSV-TEs over the eastern tropical Pacific (120°W–80°W, 10°S–10°N)
EQcenter	ICM forced by the NFSV-TEs over the central tropical Pacific (170°W–120°W, 10°S–10°N)

Note: All experiments are started in January and performed to predict the SST evolutions in the following 12 months. The sensitivity experiments, denoting the prediction made with the ICM equipped with NFSV-TEs, consist of NFSV-TEs, EQ, EQeast and EQcenter experiments.

composite of those optimal tendency forcings at the same time.

Finally, by assimilating the ERSSTv3 data from 1950 to 2010, a total of 61×12 samples of the monthly NFSV-TEs during this period can be obtained via the NFSV-based assimilation approach (shown in Figure 2). The acquired NFSV-TEs make sense in representing the total effect of unsolved mechanisms or model errors while depicting realistic SST evolution. When the NFSV-TEs are added to the tendency equation of the SST model, the NFSV-TEs-forced ICM can successfully reproduce the ENSO evolution (Figure 4).

3 | NFSV-TEs IN THE ICM

Figure 2a displays the standard deviation of the NFSV-TEs during the assimilation periods from 1950 to 2010. The NFSV-TEs of the ICM feature the largest errors in the central and eastern tropical Pacific, indicating that significant model uncertainties exist in the SST simulation and prediction in these regions. In addition, large values of the NFSV-TEs are also found over the northern and southern boundaries of the ICM. This shortcoming of the model probably lies in the fact that the ICM is a regional model. Thus, the ICM fails to describe the SST variability near the boundary of the model. The sources of equatorial NFSV-TEs are complex and relate not only to tropical thermodynamics but also originate from extratropical processes that are indescribable in the ICM (see the discussion section of the paper).

An empirical orthogonal function (EOF) is implemented with the total NFSV-TEs to better show the temporal and spatial variability of the NFSV-TEs. Results are shown in Figure 2b–d. The first leading EOF mode of the NFSV-TEs (denoted as EOF1; the second mode is EOF2, and so on) that explains 47.7% of the variance is mainly concentrated on the far eastern tropical Pacific and extends to the central-eastern equatorial Pacific. The pattern of EOF1 is quite similar to the E-type of El Niño (Takahashi *et al.*, 2011), which means that the ICM has poor skills and large model uncertainties in SST prediction in the far eastern tropical Pacific where the SST variability and predictability are dependent much on the nonlinear processes (Takahashi and Dewitte, 2016; Ding *et al.*, 2018; Ying *et al.*, 2019). Thus, here, the EOF1 is termed as E-type NFSV-TEs. On the other hand, the EOF1 is of high frequency in time (less than 1-year, see Figure 3a), indicating that the model errors of ICM are also related to the poor performance in seasonal and sub-seasonal processes. For the third leading mode of the NFSV-TEs, EOF3 mainly describes the model uncertainties near the model boundary and is of low frequency in time (about 10-year, see Figure 3c). In contrast to EOF1 and EOF3, EOF2 exhibits a strong signal both in the subtropical Pacific and along the equator. In particular, from the lag-correlation between the principal component of EOF modes and the Niño indices (Figure 3d–f), the variability of EOF2 is noticeably related to the Niño 3.4 index. It is suggested that EOF2-type tendency errors are usually prominent in realistic predictions of El Niño and La Niña events. One can find that EOF2 of the NFSV-TEs presents a dipole pattern along the equator with a positive value in the eastern Pacific and a negative value in the central tropical Pacific (referred to as D-type), which is remarkably similar to the out-phase NFSV-TPs from the ideal study by Duan *et al.* (2016). This result exactly confirms the fact that NFSV-TPs exist in realistic ENSO predictions. By adding the NFSV-TEs to the tendency equation of the SST model, the forced ICM can filter the NFSV-TPs-like model errors and limit the model error-induced prediction errors so that it can make a more realistic ENSO prediction.

To prove that the obtained NFSV-TEs can represent the model uncertainties from various sources, two experiments are performed during the period from 1950 to 2010 (Table 1): one is the control experiment where the experiment is made using the ICM without NFSV-TEs, and the other is using the ICM forced by the NFSV-TEs over the whole model grids. As shown in Figure 4, the ICM does have the ability to predict the SST anomaly over the Niño 3.4 region, while SST prediction errors are outstanding in the eastern tropical Pacific and extend to the central tropical Pacific as the lead time increases due to the model

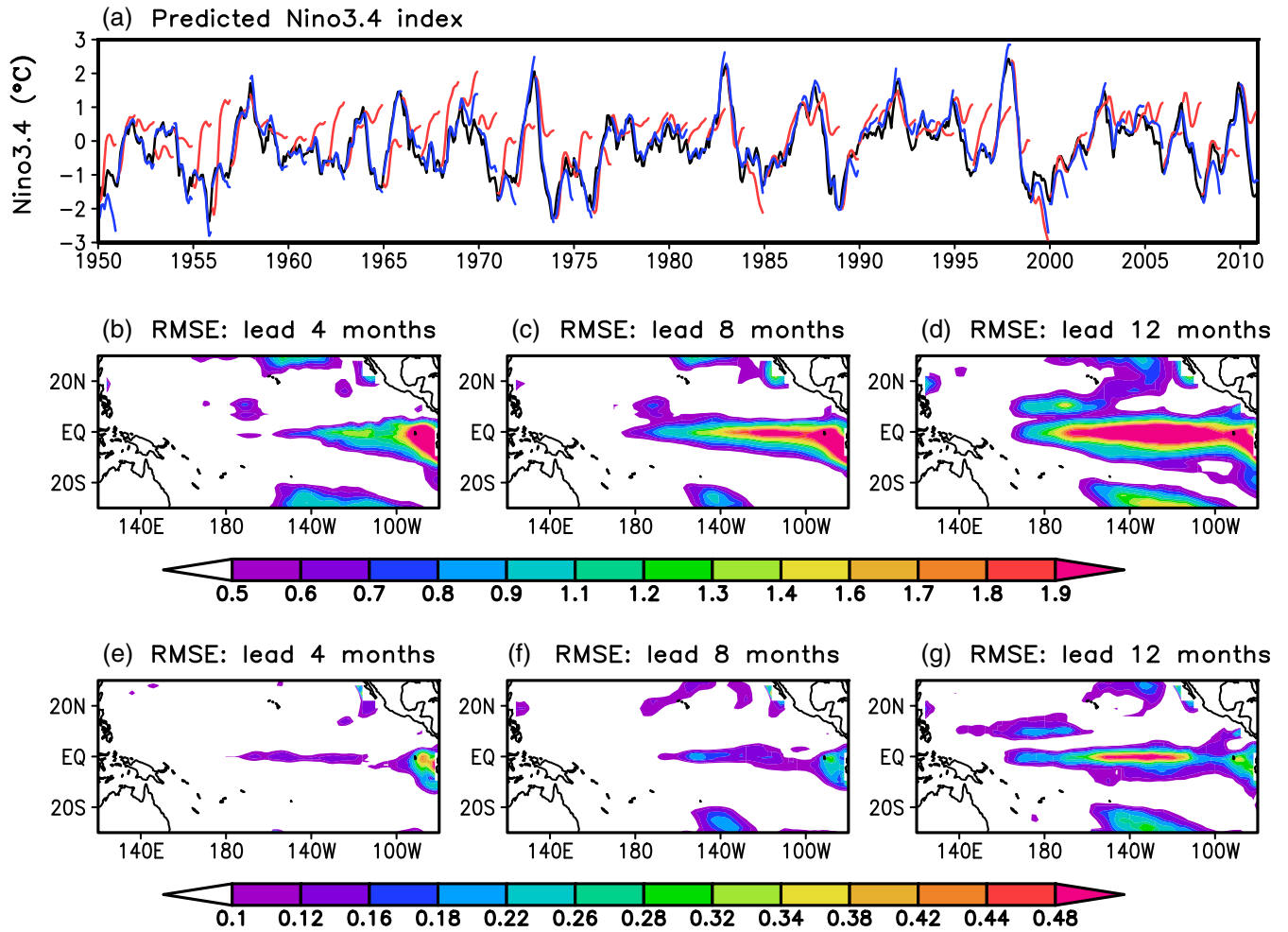


FIGURE 4 (a) Time series of the observed and predicted Niño 3.4 indices and prediction errors (unit: °C) in the (b–d) control and (e–g) NFSV-TEs experiments at 4-, 8- and 12-month lead times. The red curves and blue curves in (a) are from the control and NFSV-TEs experiments, respectively, and the black curve is the observation

and initial condition errors. The NFSV-TEs forcings can fix the ICM and filter the impact of model uncertainties on the error growth to a large extent. In the experiment with the NFSV-TEs perturbed, the root mean square errors (RMSEs) of the SST anomaly are reduced to less than 0.5°C, which is mainly linked to the initial condition errors. Based on the former results, the ICM forced by the NFSV-TEs can be served as a ‘true’ climate model to describe the realistic ENSO evolution. It is also demonstrated that the NFSV-TEs can be regarded as the model uncertainties of the original ICM in realistic prediction.

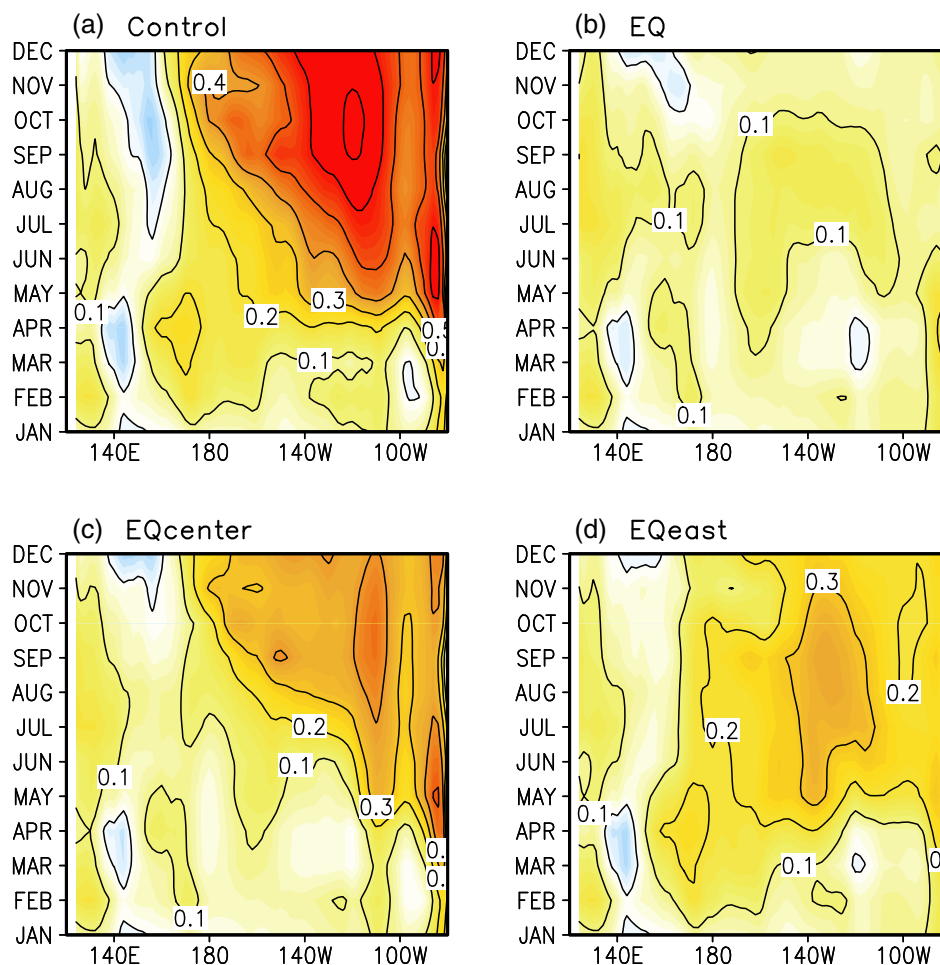
4 | EFFECTS OF THE EQUATORIAL NFSV-TEs ON ENSO PREDICTION

The last section has shown the characteristics of NFSV-TEs, which are representatives of model uncertainties.

Although the NFSV-TEs exhibit a large signal near the boundaries of the ICM, their effects on the error growth of the ENSO prediction are much weaker relative to the NFSV-TEs along the equator (data not shown). Considering not only the characteristics of the NFSV-TEs (Figure 2) but also the crucial role of the equatorial NFSV-TEs (Duan *et al.*, 2016), in this section, we will confine our attention to examining the impacts of the local NFSV-TEs along the equator on the realistic prediction of ENSO.

To better show the impacts of model errors on the SST predictions, the simulation using the ICM with NFSV-TEs is served as ‘observation’ (i.e. the experiment with the whole NFSV-TEs perturbed), the original ICM and that equipped with the equatorial component of the NFSV-TEs are the prediction models and used to predict the pseudo-true ‘observation’. We therefore perform three additional sensitivity experiments using the ICM equipped with the equatorial component of the NFSV-

FIGURE 5 Systematic biases along the equator in the (a) control, (b) EQ, (c) EQcenter and (d) EQeast experiments. The systematic biases are defined as the mean deviations of the annual mean from the observation. The contour interval is 0.1°C



TEs (Table 1): EQ, EQeast, and EQcenter experiments. The EQ experiment is performed using the ICM forced by the NFSV-TEs over the central and eastern equatorial Pacific. The EQeast (EQcenter) experiment is made by the ICM with NFSV-TEs only over the eastern (central) equatorial Pacific. All these experiments are initiated in January of each year with a lead time ranging from 1 to 12 months. Note that the initial conditions are kept the same among these experiments; the different experiments are totally attributable to the imposed NFSV-TEs. Thus, by comparing the prediction skills revealed by the control and sensitivity experiments, we can study the impacts of NFSV-TEs on ENSO predictions in terms of systematic bias, interannual variability, and El Niño types.

4.1 | Systematic bias

The climate drift phenomenon is common among coupled models, which induces a systematic bias even larger than the SST variability (Schneider *et al.*, 2003; Jin *et al.*, 2008). Although the systematic bias is smaller than that in some general circulation models, it becomes large as the prediction time is increased in the ICM

(Figure 5a). There are two paths leading to the increased systematic bias: one originates from the eastward bias from the western tropical Pacific, and the other relates to the bias in the eastern tropical Pacific. In seasonality, the systematic biases associated with both paths increase rapidly from March to June, which may favour the cause of the SPB phenomenon during the ENSO predictions. In the EQ experiment, forced by the NFSV-TEs along the equator (hereafter termed EQ-TEs for convenience), the bias propagation paths are so obscured that the systematic bias fails to develop, where the bias is less than 0.1°C even at a 12-month lead time (Figure 5b). The demise of systematic bias indicates that the systematic bias in the ICM is mainly due to model tendency errors over the equatorial Pacific. That is, a lower model uncertainty of the model in the eastern and central equatorial Pacific tends to exhibit a less systematic bias.

The western and eastern parts of the EQ-TEs (the former is referred to as EQcenter-TEs, and the latter are EQeast-TEs) play distinct roles in weakening the systematic biases. As shown in Figure 5c,d, both EQcenter-TEs and EQeast-TEs can reduce the systematic bias in the local region where the NFSV-TEs are superimposed but also in the region where the added NFSV-TEs are not

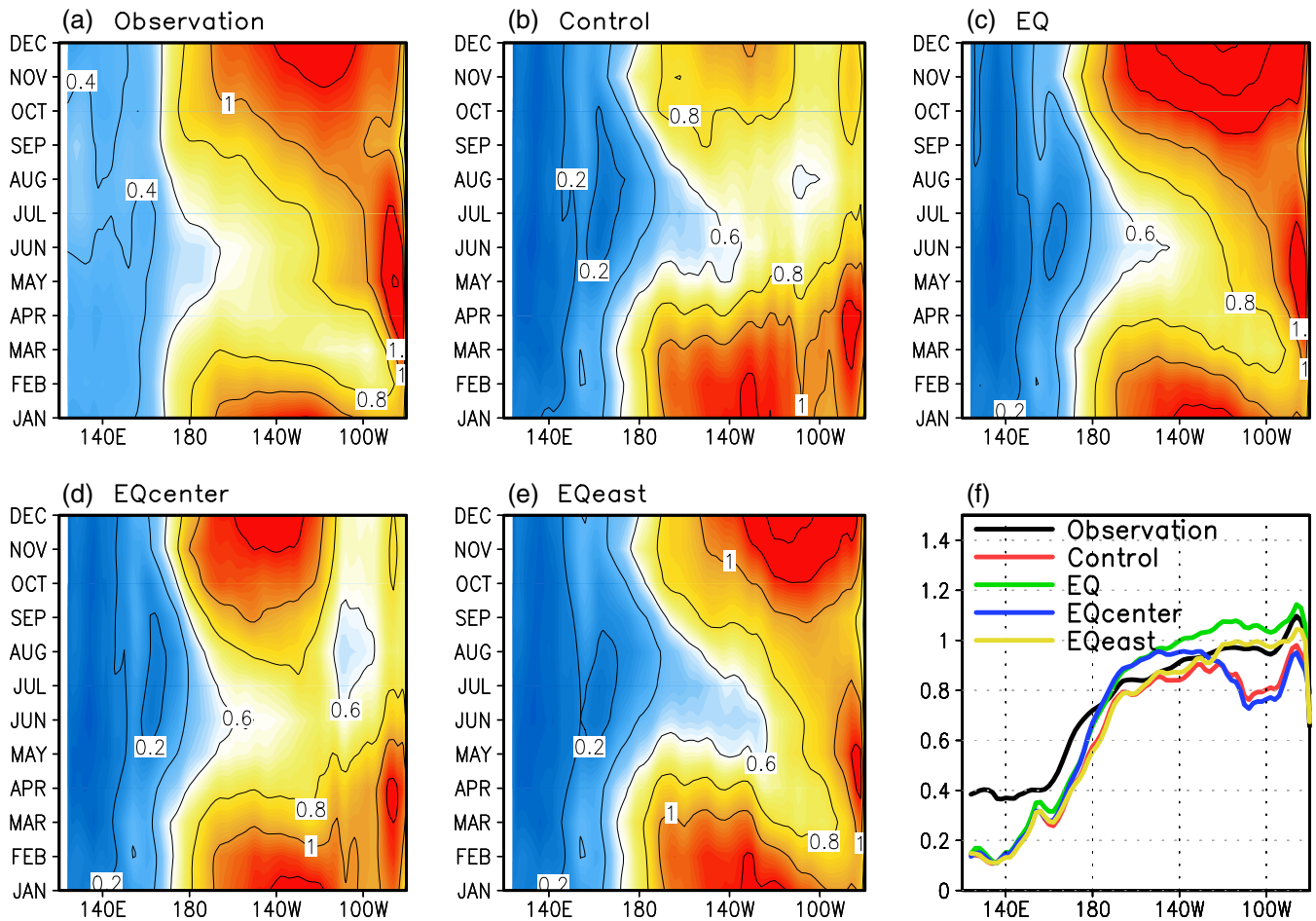


FIGURE 6 The standard deviation of the observed and predicted SST anomalies along the equator. Results are calculated based on the (a) observation, (b) control, (c) EQ, (d) EQcenter and (e) EQeast sensitivity experiments along the equator (meridionally averaged across 5°S – 5°N) as a function of calendar month. Results in (f) is calculated without considering the seasonality. The contour interval is 0.2°C in (a)–(e). In (f), the red line is from the control experiment; the green line is from the EQ experiment; the blue and yellow lines denote the interannual variability of SST derived from the EQcenter and EQeast experiments, respectively

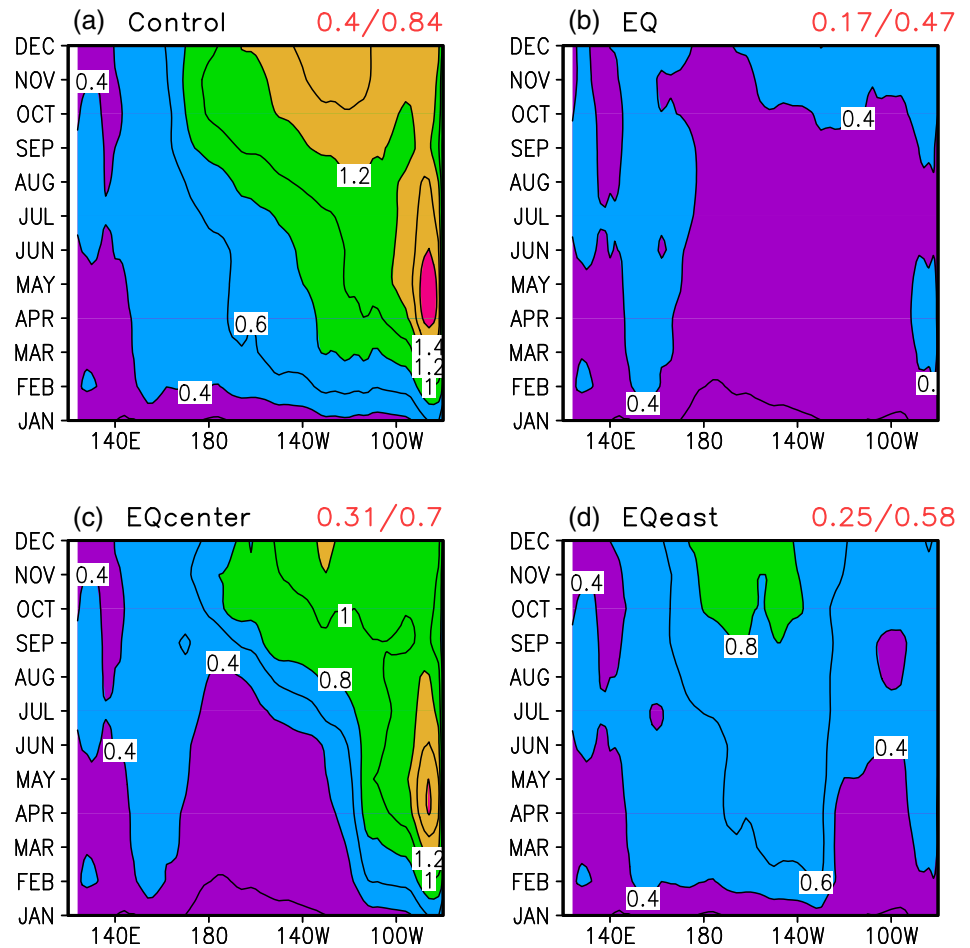
covered. However, EQeast-TEs show better performance than EQcenter-TEs in the reduction of systematic bias. For example, the bias in the EQcenter experiment is larger than 0.4 in the central-eastern tropical Pacific, while it fluctuates at approximately 0.3°C in the EQeast experiment. This result suggests that the remote effect of the EQeast-TEs on weakening the bias in the central-eastern Pacific is equal to or better than the local effect of the EQcenter-TEs. From this point, the model performance in the eastern tropical Pacific should be improved to effectively reduce the systematic bias.

4.2 | Interannual variability

The interannual variability is the standard deviation of the monthly SST anomaly, which usually represents the strength of the SST variability related to the ENSO.

Figure 6f shows the interannual variabilities of SST anomalies from the observation and model experiments. All hindcast experiments present high performance in the strength and distribution of the interannual variability that displays a strong variance in the eastern equatorial Pacific and the weakest variance in the west. However, the seasonality of the SST variance derived from the ICM is entirely different from the observations. In the observations, the SST variability in the central-eastern (near 130°W) tropical Pacific is strongest in winter (from November to January) and weakest in spring, while the SST variability near the coast of South America (near 90°W) is strongest in spring (Figure 6a). In the control experiment, the predicted SST in the central and eastern tropical Pacific features the strongest variance during late winter and early spring (from January to April) and the weakest variance in summer (Figure 6b). In the EQ experiment where the ICM is forced by the equatorial

FIGURE 7 RMSEs of the SST anomalies along the equator in the (a) control, (b) EQ, (c) EQcenter and (d) EQeast sensitivity experiments. The ratio of error growth in FMAM and the total error growth in prediction period over the whole tropical Pacific is noted in the top right corner of each panel. The contour interval is 0.2°C



NFSV-TEs (i.e. EQ-TEs), the SST variance is amplified during October–December and reduced in spring, giving rise to a much more realistic seasonality (Figure 6c). When only the central parts of the equatorial NFSV-TEs (i.e. EQcenter-TEs) are used to force the ICM, the seasonality of the predicted SST variability is only corrected over the central tropical Pacific (Figure 6d). By comparison, the eastern parts of the equatorial NFSV-TEs (i.e. EQeast-TEs) can improve the interannual variability in SST not only over the eastern tropical Pacific but also over the central tropical Pacific (Figure 6e). Hence, reducing the model uncertainties in the eastern tropical Pacific is an effective approach to enhance the seasonality of the SST variance in the whole tropical Pacific during ENSO prediction.

The impacts of the NFSV-TEs on the interannual variability are also reflected in the SST prediction errors (Figure 7). The prediction errors here are defined as the RMSEs that measure the agreement in the amplitude of the interannual variability of the SST. The weak SST variability in the ICM causes a large RMSE in the eastern tropical Pacific that is up to 1.4°C . The prediction errors can be largely reduced at the basin scale in the EQ and EQeast experiments. For the EQcenter experiment,

although the model tendency errors are removed in the central-eastern equatorial region, substantial prediction errors exist due to the westward errors from the far eastern tropical Pacific (see Figure 7c after May). As for the SPB phenomenon, the ICM also shows the largest error growth in spring (i.e. February–May), where the error growth is 0.84°C totally but 0.4°C during spring. With the forcing of the NFSV-TEs, the error growth in spring is much decreased. Not accidentally, this echoes the previous studies on the SPB phenomenon that SPB of ENSO prediction is partly originated from the model errors (Mu *et al.*, 2007; Duan and Hu, 2016; Hou *et al.*, 2019; Tan *et al.*, 2020).

From the above, reducing the model uncertainties in the eastern equatorial Pacific is more important than reducing those in the central equatorial Pacific in the improvement of the ENSO predictions in terms of the SST variability and RMSEs.

4.3 | CP- and EP-El Niño predictions

Predicting only the Niño index or the strength of the ENSO cannot meet our needs since many evidences show

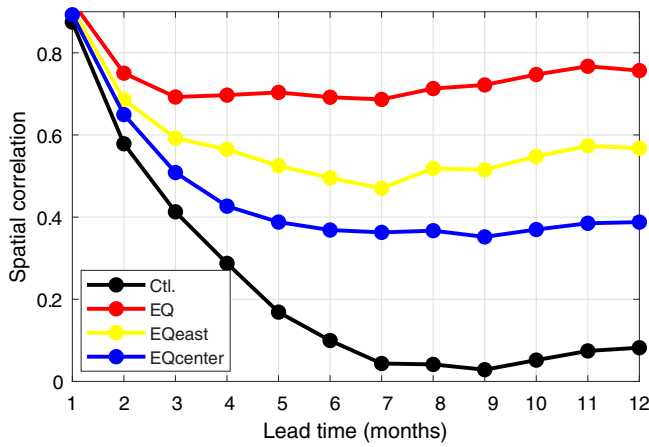


FIGURE 8 Spatial correlations between the observed and predicted SST anomalies in the tropical Pacific as a function of lead time

that ENSO-induced climate anomalies are critically dependent on the horizontal distributions of the ENSO themselves (Timmermann *et al.*, 2018). Thus, this section will examine the prediction skills with respect to the horizontal distributions of the SST anomalies.

A spatial correlation coefficient (SCC) is defined to quantify the pattern prediction skills, which is written as

$$SCC = \frac{\sum_{i,j} x_{i,j} \cdot y_{i,j}}{\sqrt{\sum_{i,j} x_{i,j}^2 \cdot \sum_{i,j} y_{i,j}^2}}, \quad (9)$$

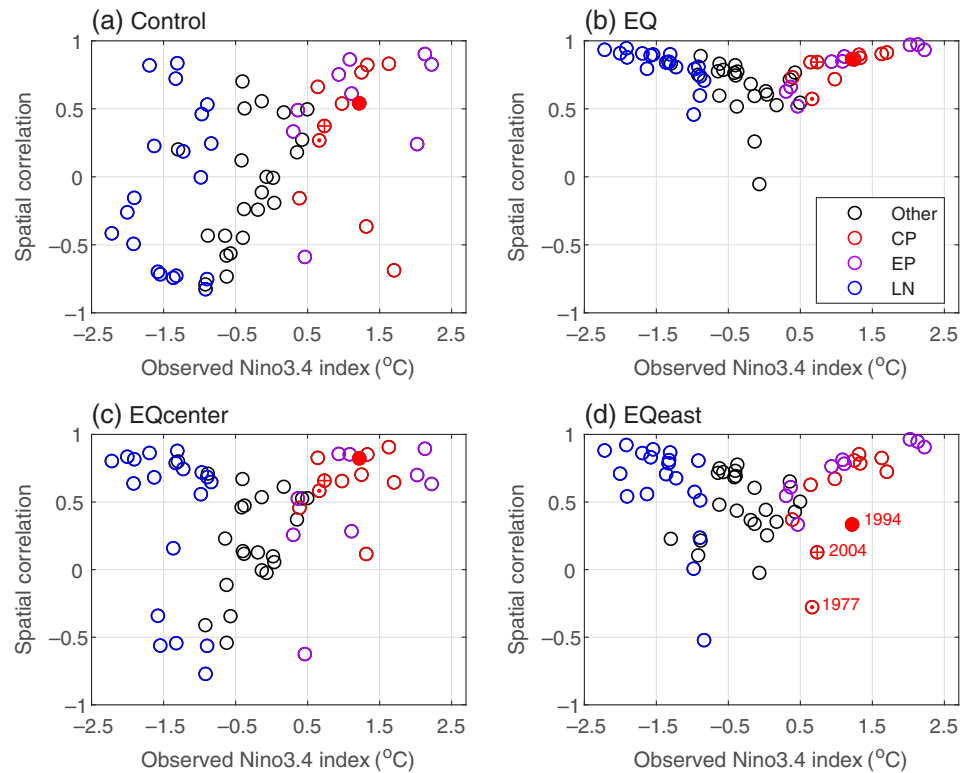
where $x_{i,j}$ and $y_{i,j}$ are the predicted and observed SST anomalies in the model grid (i,j) . The pattern prediction skills in different experiments are shown in Figure 8. The ICM achieves poor grades in predicting the horizontal characteristics of the SST anomalies in the tropical Pacific, where the SCC is lower than 0.4 after one season. When correcting the model tendency equation with the EQ-TEs, the prediction skill towards the horizontal SST anomalies is significantly improved with an SCC greater than 0.6 within a 12-month lead time. The improvement is reduced when only correcting the model in the central equatorial Pacific (i.e. adding the EQcenter-TEs to the tendency equation). The SCC remains at approximately 0.4 at a long lead time in the EQcenter experiment. Compared to the EQcenter experiment, the EQeast experiment obtains higher scores in SCC at all lead time predictions.

Previous results have clearly shown the impact of the local NFSV-TEs on the pattern prediction skills of all ENSO events (including La Niña, El Niño, and neutral events). We now turn to look at the events individually, with a specific focus on the CP and EP-El Niño events. In

Figure 9, the SCCs of each event in December are shown as a function of the observed Niño 3.4 index. There exists some capability of spatial predictions for El Niño events among all experiments. However, the SCC involved with the neutral and La Niña event predictions are not as satisfactory as the El Niño predictions using the ICM, where the SCC is even close to -1 during La Niña predictions (Figure 9a). Forced by the EQ-TEs or EQeast-TEs, the predictions of the horizontal distributions of the neutral and La Niña events are significantly improved. By comparison, the EQcenter-TEs have little effect on the improvement of the La Niña predictions (Figure 9c), thus leading to a lower mean SCC in the EQcenter experiment than in the EQeast experiment (Figure 8).

For the two types of El Niño events, the EQcenter-TEs are superior to the EQeast-TEs with respect to the horizontal SST predictions. Figure 10 displays the composites of the CP- and EP-El Niño events from the observation and sensitivity experiments, where the types of El Niño events are classified based on the method proposed by Ashok *et al.* (2007) (Table 2). In observation, the SST anomaly of EP-El Niño peaks in the far eastern equatorial Pacific while the CP-El Niño peaks near 160°W (Figure 10a1,b1). Thus, the difference between EP-El Niño and CP-El Niño shows negative SST anomaly in the central tropical Pacific and positive SST anomaly in the eastern tropical Pacific (Figure 10c1). One can find that the original model not only tends to predict a warm event that is weaker than observations for both events but also has no skill to predict the difference between CP- and EP-El Niño events as the observation. For example, the maximum SST anomalies that are predicted by the ICM are both located near the 150°W for the EP- and CP-El Niño predictions. In the EQ experiment, the ICM forced by the EQ-TEs can capture the main difference of the El Niño types, and successfully predict the locations of the maximum SST anomaly for EP- and CP-El Niño. Only forced by the eastern part of the EQ-TEs, the model seems to not work well in reproducing the peak location of the CP-El Niño event. As shown in Figure 10b4, we can see that an EP-El Niño is usually produced during the CP-El Niño prediction. The predicted SST anomalies both for the EP- and CP-El Niño peak near 120°W . Thus, the EQeast experiment fails to identify the type of El Niño in advance. What is worse is that the prediction results for the CP-El Niño events sometimes are worse in the EQeast experiment than in the control experiment, such as 1997, 1994 and 2004 CP-El Niño events (Figure 9). Taking the CP-El Niño event that occurred in 1977 as an example, the predicted SST pattern is opposite to the observation with an SCC less than -0.3 in the EQeast experiment. In contrast, a noticeable improvement in the CP-El Niño prediction is seen in the EQcenter

FIGURE 9 Scatter plot of the spatial correlations of the December SST anomaly as a function of the observed Niño 3.4 index in the (a) control, (b) EQ, (c) EQcenter and (d) EQeast experiments. The black circles are the neutral events, and the blue circles are the La Niña events. The red and purple circles denote the CP- and EP-El Niño events, respectively. The circles with dots, shading and pluses refer specifically to the events in 1977, 1994 and 2004, respectively



experiment where the EQcenter-TEs are added to eliminate the tendency uncertainties of the model over the central equatorial Pacific (Figures 9c and 10b5). Same as observations (Figure 10a1,b1), the predicted SST anomalies for EP-El Niño and CP-El Niño peak near 120°W and 160°W , respectively (Figure 10a5,b5).

In addition, the EQcenter-TEs are also superior to the EQeast-TEs in the prediction towards the strength of the CP-El Niño event. Table 3 shows the strengths of EP- and CP-El Niño in observation and model experiments. The ICM usually underestimates the amplitudes of both EP- and CP-El Niño. Both EQeast-TEs and EQcenter-TEs can enhance the El Niño amplitudes but to varying degrees. The EQeast-TEs can better improve the EP-El Niño prediction while the EQcenter-TEs improve the CP-El Niño better. Combining the effects of NFSV-TEs on the predictions towards types and strengths of El Niño, it is concluded that the model uncertainties both in the central and eastern equatorial Pacific are vital to enhancing the ENSO predictions, but fixing the east alone could lead to worsen results for CP-El Niño predictions.

5 | DYNAMIC ROLES OF THE NFSV-TEs IN DIFFERENT REGIONS IN CP-EL NIÑO PREDICTIONS

The previous section clearly demonstrates that the central and eastern equatorial NFSV-TEs exhibit different

effects on the horizontal distribution of SST during the CP-El Niño predictions. This naturally leads us to explore their involved dynamics to answer the question of why EQcenter-TEs are important in CP-El Niño predictions. Here, a strong CP-El Niño event that occurred in 1994 is chosen as a case study. Note that the following results are similar to the predictions of other CP-El Niño events.

The evolutions of the CP-El Niño event in the observation and those predicted by the ICMs with parts of the NFSV-TEs are shown in Figure 11. As indicated in Figure 11a,d, the observed SST anomaly in the central tropical Pacific is rapidly developed after spring, while the strong positive SST anomaly in January decays unrealistically with time in the control experiment. As a result, a neutral event of which the SST anomaly is less than 0.5°C is incorrectly predicted, although the horizontal pattern of the SST anomalies is similar to the observation where the SCC is approximately 0.5 (shown in Figure 9a). The case is worse in the EQeast experiment. In the EQeast experiment (Figure 11b,e), the predicted negative SST anomaly in the eastern tropical Pacific does present little departure from the observation, but the predicted SST anomaly in the central tropical Pacific is gradually different from the observation over time. This is because the positive SST anomaly in the central tropical Pacific is gradually replaced by the westward negative SST anomaly from the eastern tropical Pacific. Finally, the EQeast-TEs make the ICM give a failed prediction that a negative SST anomaly is predicted in the central

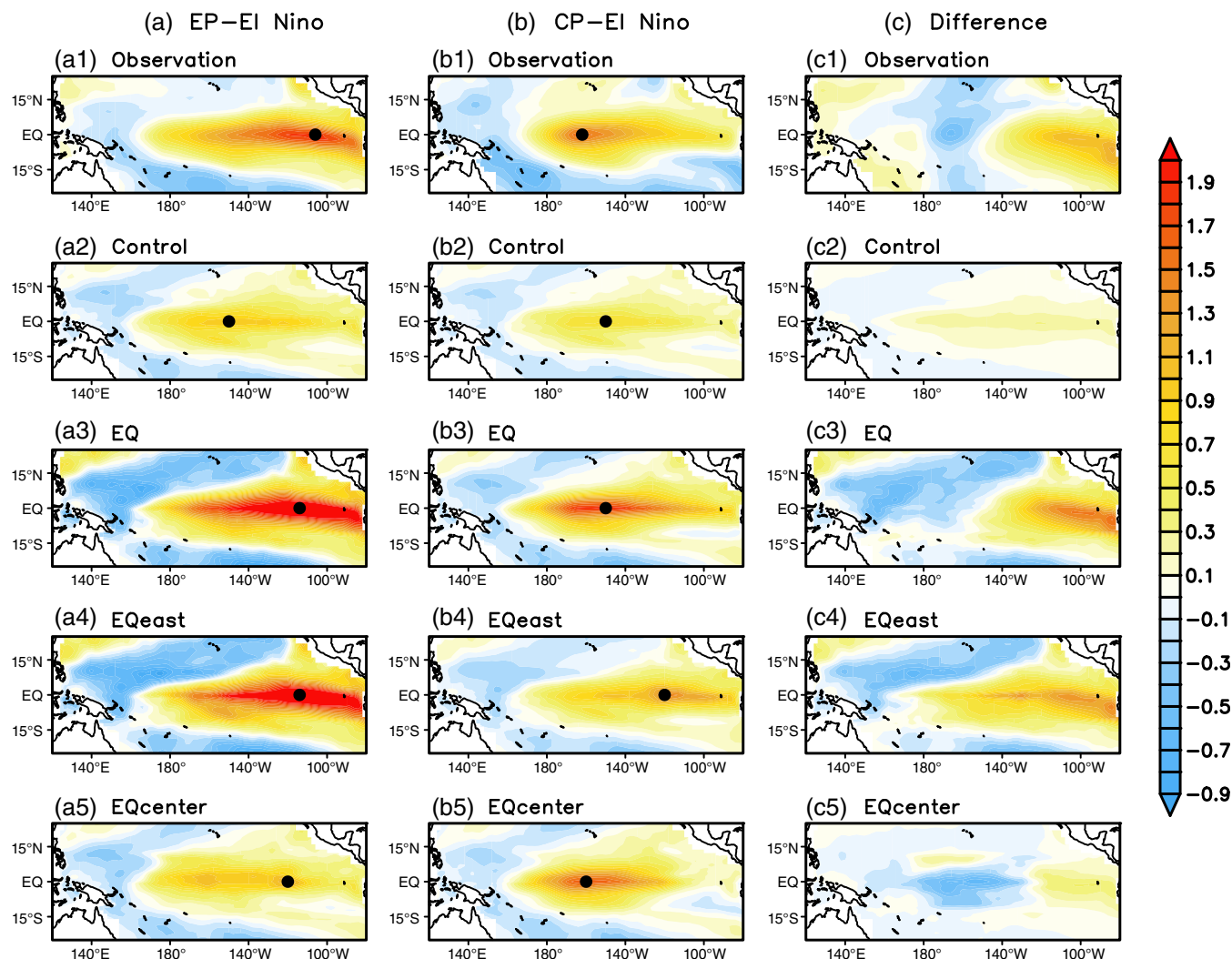


FIGURE 10 Composites of the (a) EP- and (b) CP-El Niño events and (c) their differences derived from the observation and sensitivity experiments. The black solid circles over the left two columns indicate the location of the maximum SST anomaly along the equator

TABLE 2 La Niña years and El Niño years with respect to two types

Events	The year when the event peaks
EP El Niño	1951, 1969, 1972, 1976, 1982, 1986, 1987, 1997, 2006
CP El Niño	1953, 1957, 1963, 1965, 1968, 1977, 1991, 1994, 2002, 2004, 2009
La Nina	1954, 1955, 1964, 1970, 1971, 1973, 1974, 1975, 1983, 1984, 1988, 1995, 1998, 1999, 2000, 2005, 2007, 2008, 2010

Note: The type of El Niño is classified based on the method proposed by Ashok et al. (2007).

tropical Pacific. When adding the NFSV-TEs to the SST model over the central equatorial Pacific, an obvious rapid warming in the central tropical Pacific in summer

can be reproduced in the EQcenter experiment (Figure 11c,f). A CP-El Niño is thus successfully predicted with a strong positive SST anomaly in the central tropical Pacific at the end of the year.

Figure 12 further demonstrates the validity of the EQcenter-TEs from the perspective of air-sea interactions. The air-sea states are basically the same among different sensitivity experiments at the beginning of the year, although differences appear later in the year. As seen in Figure 12a, for the experiment made by the ICM, the demise of the SST anomaly is dynamically associated with the decrease in the wind anomaly during spring and summer. Without persistent atmospheric forcing on the tropical ocean, the tilt of the thermocline indicated by the sea level anomaly is difficult to maintain, thus showing a weak air-sea coupling system in the control experiment. In the EQeast experiment, the air-sea coupling is

TABLE 3 The strengths of composite El Niño events in observation and sensitivity simulations

	Obs.	Control	EQ	EQeast	EQcenter
Strength of EP-El Niño	1.42	0.65	1.93	1.91	0.72
Strength of CP-El Niño	0.77	0.40	0.97	0.43	0.93

Note: The value of Niño3 and Niño4 indices are used to quantify the strengths of EP-El Niño and CP-El Niño, respectively.

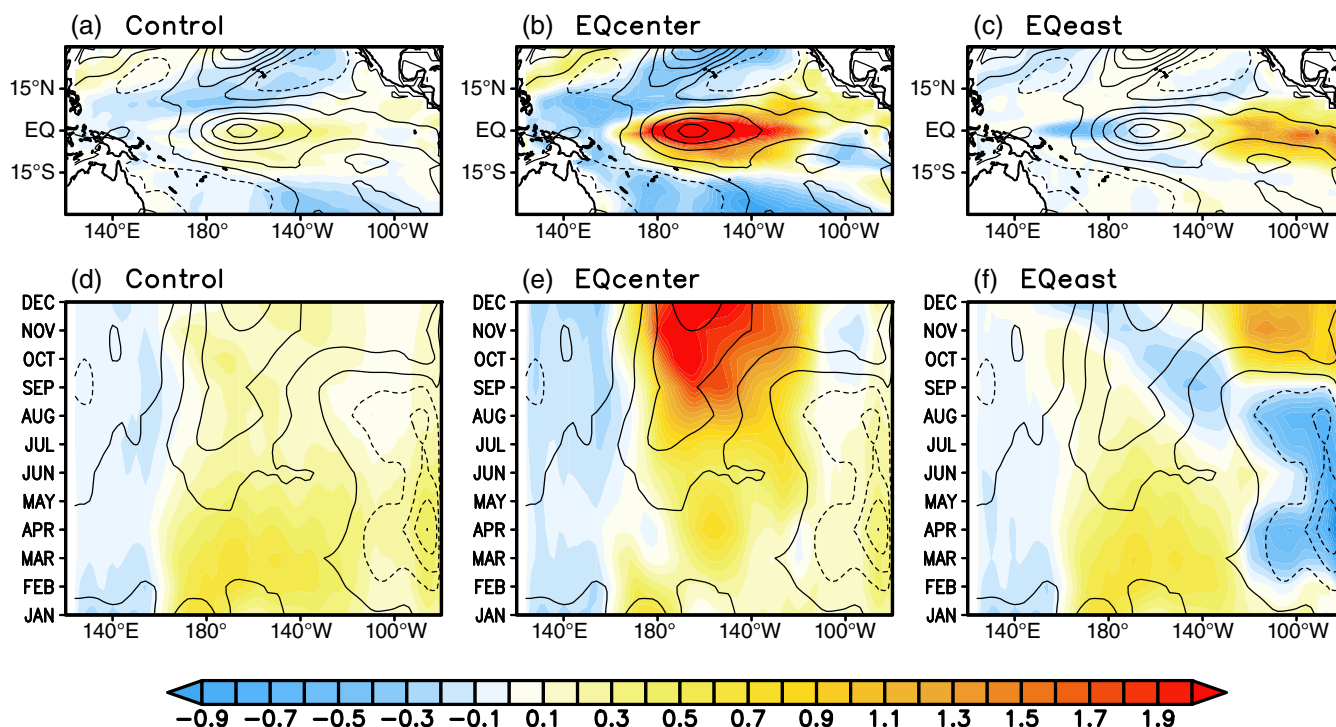


FIGURE 11 (a–c) Horizontal distributions of the predicted SST anomalies in December 1994 and (d–f) evolutions of predicted SST anomalies along with the equator during 1994. The shading is from the sensitivity experiments. The contoured lines with an interval of 0.5°C are the observations

the same as that in the control experiment, and the westerly wind anomaly also decreases across spring. Due to the strong effect of EQeast-TE forcing, the air–sea interaction gradually becomes strong over the eastern tropical Pacific, as indicated by the strong sea level anomaly and wind anomaly to the east of 140°W (Figure 12c,f). The east wind anomaly interacts with the triggered Rossby waves and therefore carries the negative SST anomaly into the central tropical Pacific at the end of the year. Therefore, the CP-El Niño is difficult to predict in the EQeast experiment. When adding the EQcenter-TEs to the SST tendency, the west wind anomaly does not disappear across spring, which is accompanied by the east wind anomaly in the eastern equatorial Pacific acting to deepen the thermocline in the central tropical Pacific. Heated by the subsurface, the positive SST anomaly in turn feeds back to the atmosphere and tends to further enlarge the westerly anomaly. Under the circumstances

of strong positive feedback in the central tropical Pacific, a CP-El Niño is easily formulated in the EQcenter experiment.

From the above, the key to predicting the CP-El Niño is dependent on the model performance in predicting the rapid warming from spring to summer and the convergence wind anomaly. Both phenomena are captured in the EQcenter experiment; thus, the horizontal distribution of CP-El Niño is successfully predicted. We know that the differences in air–sea coupling among different experiments are merely due to the differences in the prediction models themselves. Obviously, EQcenter-TEs do exert in the CP-El Niño prediction. Then, what roles do EQcenter-TEs play in predicting the anomalous wind convergence in spring and the subsequent rapid warming?

A mathematical derivation of the effect of the NFSV-TEs across spring shows that the change in the wind

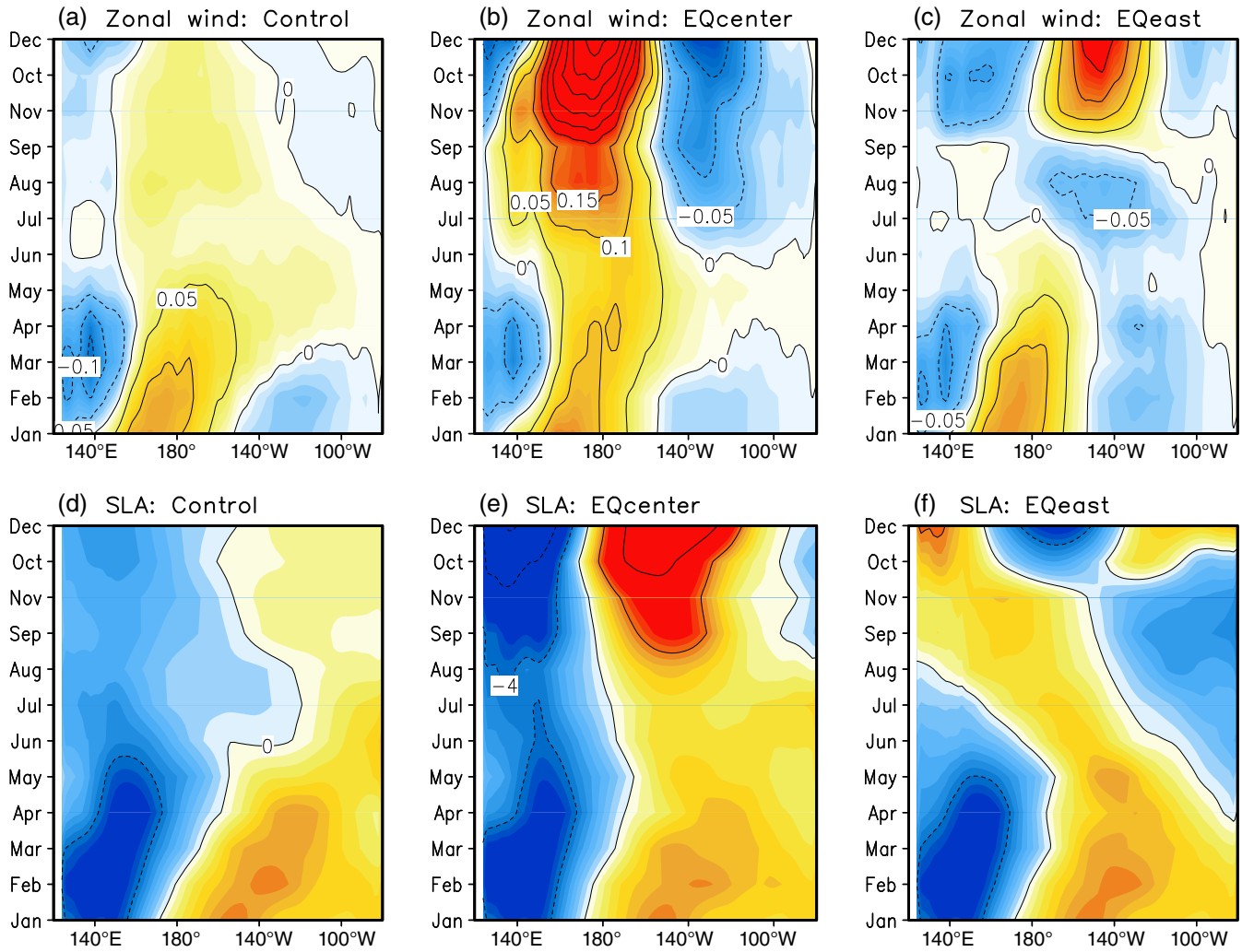


FIGURE 12 Longitude-time sections of the predicted (top panels) zonal wind stress anomalies and (bottom panels) sea level anomalies. The contour intervals in the top and bottom panels are $0.05 \text{ dyn}\cdot\text{cm}^{-2}$ and 4 cm, respectively

tendencies is a function of the superimposed NFSV-TEs (see the derivation in the Appendix), which can be represented as

$$\frac{\Delta \tau_2}{\Delta t} - \frac{\Delta \tau_1}{\Delta t} = \mathbf{G} \cdot \mathbf{f}_{\text{part}}^* \quad (10)$$

in which $\mathbf{G}$ is the derivative of the wind model indicated by Equation (1), $\mathbf{f}_{\text{part}}^*$ denotes the part of the NFSV-TEs (i.e. EQ-TEs, EQeast-TEs and EQcenter-TEs); $\frac{\Delta \tau_2}{\Delta t}$ and $\frac{\Delta \tau_1}{\Delta t}$ present the wind tendencies from April to June in the sensitivity and control experiments, respectively. That is, NFSV-TEs can directly project onto the tendency of the wind anomaly through the operator $\mathbf{G}$, thereby influencing the wind anomaly simulation and the following evolution of the SST anomaly in the tropical Pacific.

Figure 13 clearly shows the effect of the NFSV-TEs on the changes in the wind tendencies. The equatorial

NFSV-TEs for the CP-El Niño prediction exhibit negative values to the east and relatively weak positive values to the west, which also implies that large positive model errors exist in the ICM over the eastern tropical Pacific in spring. As revealed in Figure 13a, the EQ-TEs can force the ICM to yield the tendency changes of the zonal wind towards the central equatorial Pacific during spring, the condition of which is dynamically favoured to generate a CP-El Niño event. When considering only the eastern part of the EQ-TEs, one can find that the EQeast-TEs do have a role in the changes in the meridional wind tendencies but have little effect on the zonal wind tendencies in the western and central equatorial Pacific (Figure 13c). Therefore, the predicted zonal wind anomalies in spring are almost identical among the control and EQeast experiments (Figure 12a,c). In contrast, although the strength of the EQcenter-TEs is weaker than that of the EQeast-TEs, the former has the ability to change the

FIGURE 13 NFSV-TEs and their induced changes in the wind tendencies in the (a) EQ, (b) EQcenter and (c) EQeast experiments compared to the control experiment

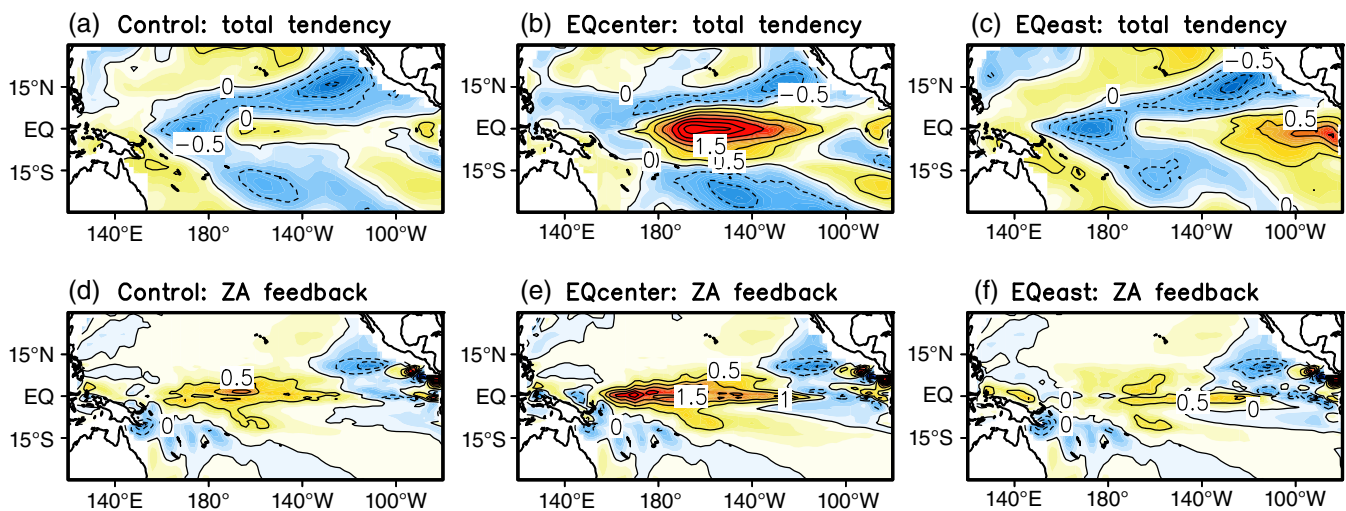
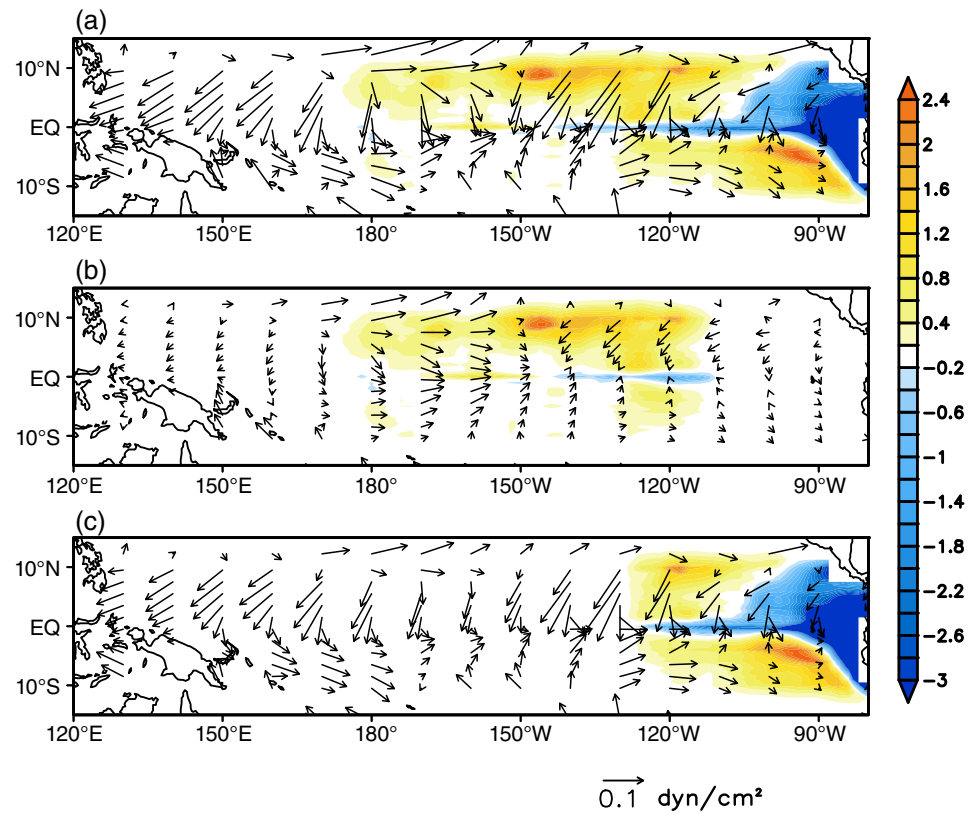


FIGURE 14 Total SST tendency and zonal advection feedback from January 1994 to December 1994. The contour interval is 0.5°C/year

zonal wind tendency and make the model tend to generate a westerly wind anomaly in the western tropical Pacific and an easterly wind anomaly in the eastern tropical Pacific (Figure 13b).

In addition to the wind, the ocean processes are also modified by the NFSV-TEs to make the predictions of CP-El Niño different. As shown in Figure 14, the zonal advection feedback term throughout 1994 features a positive value in the central tropical Pacific but with various

amplitudes among different experiments, hinting at different warming effects on the SST prediction. Comparing the zonal advection term to the total tendency of the SST in the individual experiments, one can find that the zonal advection can explain the total SST tendency in the EQcenter experiment only. In other sensitivity experiments, the tendencies of the SST anomaly are unrelated to the zonal advection feedback. This indicates that the SST evolutions in the control and EQeast experiments are

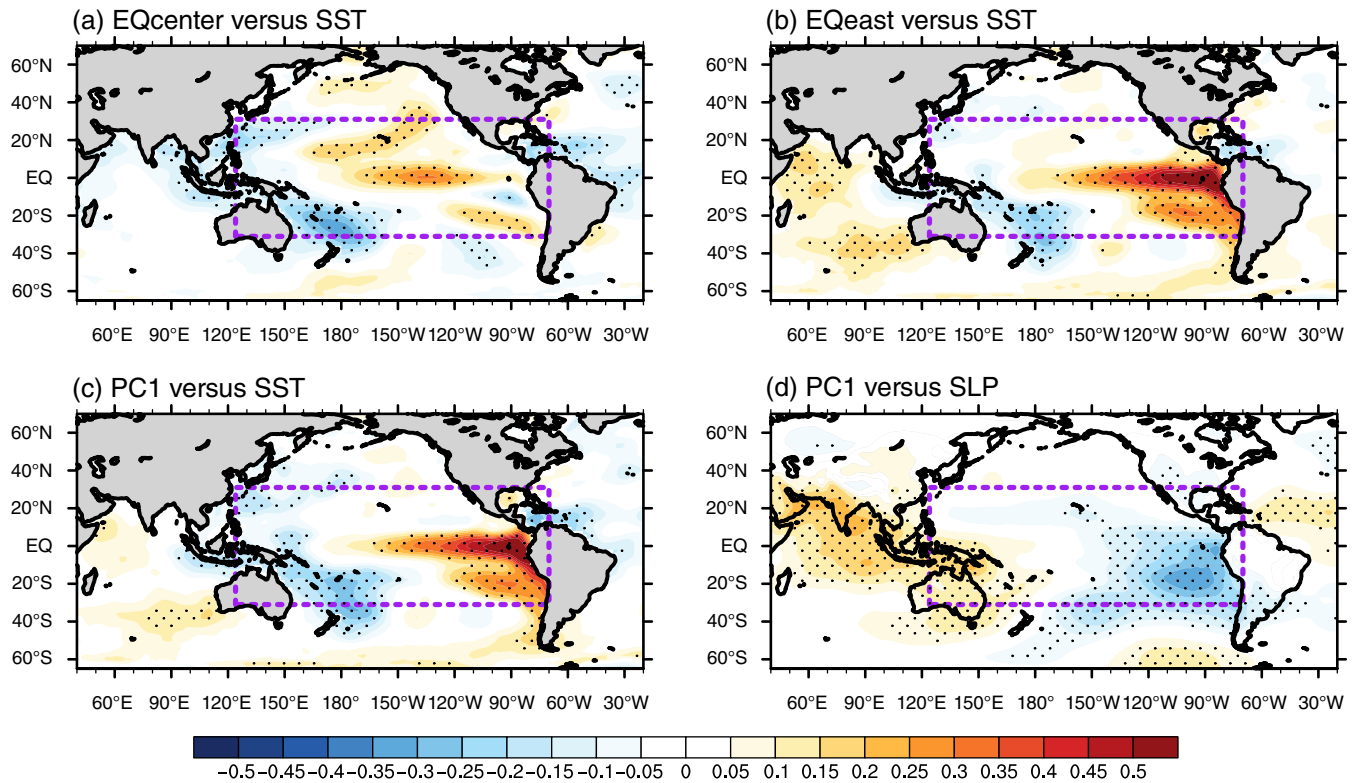


FIGURE 15 Correlations of the NFSV-TEs to the SST and SLP anomalies globally. Results in (a)–(c) are the SST correlations to the EQcenter-TEs, EQeast-TEs and PC1 of NFSV-TEs, respectively. Results in (d) are the SLP correlations to the PC1 of NFSV-TEs, respectively. The region that the SST model of the ICM covers is framed by the purple dashed lines. Dots are statistically significant at the 90% significance level using the *t*-test

not dominated by zonal advection feedback, but the SST evolutions in the EQcenter experiment are. Note that previous studies have demonstrated the important role of zonal advection feedback in the development of CP-El Niño events; thus, we reveal in the present study that EQeast-TEs have the ability to adjust ocean processes more realistically and amplify the effect of zonal advection feedback so that CP-El Niño events can be successfully predicted.

6 | DISCUSSION AND CONCLUSION

Prediction uncertainties for ENSO are attributed not only to erroneous initial conditions but also to imperfect models and inherent stochastic processes. Such predictability issues are classified into two distinct problems by Lorenz (1975): the first involves the initial error problem, while the second involves the model uncertainty problem. In this article, focusing on the second predictability problem, we investigate the impacts of the model uncertainties on the predictions of ENSO by reducing the model tendency errors in the SST model.

To quantify the model errors in realistic predictions, the NFSV-based data assimilation is applied to an ENSO forecasting model (i.e. ICM) to identify the optimal model tendency error occurring in the ENSO hindcasts during 1950–2010. By assimilating the SST data from 1950 to 2010, we obtain a total of 61×12 samples of the monthly optimal tendency perturbations (i.e. NFSV-TEs) that characterize the combined effect of the model uncertainties with various sources such as the missing processes from the extratropical Pacific, parameterizations, and so forth. When these NFSV-TEs are re-added to the tendency of the SST model and force the ICM, the historical evolution of the SST anomaly in the tropical Pacific can be successfully simulated with much smaller errors that can be attributed mainly to the initial condition errors. Through analysis of the temporal and spatial NFSV-TEs, it is found that the ICM presents large model tendency errors over the eastern equatorial Pacific and the model boundaries, as revealed by the NFSV-TEs concentrated in these regions. Two dominant NFSV-TEs are found in the ICM (Figure 2): one is E-type that mostly located in the far eastern Pacific, the other is the D-type that shows a dipole pattern of the NFSV-TEs along the equator. The D-type usually occurs during the predictions

of El Niño events, confirming the fact that NFSV-TPs exist in realistic ENSO predictions. Thus, by adding the NFSV-TEs to the model, we can effectively filter those optimally growing patterns of the tendency errors (i.e. NFSV-TPs) and limit the error growth induced by model errors, thereby effectively reproducing the ENSO events.

For the large model tendency errors in the central and eastern equatorial Pacific, the impact of the eastern (i.e. EQeast-TEs) and central (i.e. EQcenter-TEs) parts of the equatorial NFSV-TEs on the ENSO hindcasts are examined in terms of the systematic bias, interannual variability, and El Niño types. We find that the EQeast-TEs show better performance than the EQcenter-TEs in the reduction of the systematic bias and the prediction of the seasonality of interannual variability. This is likely because the large prediction errors in the eastern tropical Pacific tend to propagate westward and contaminate the prediction over the central tropical Pacific even though EQcenter-TEs are added to the tendency of the ICM. It is therefore suggested that the model performance in the far eastern tropical Pacific should be improved to effectively enhance the SST prediction at a basin scale.

However, the EQcenter-TEs are found to be superior to the EQeast-TEs when reproducing the horizontal distribution of the sea surface temperature (SST) anomaly, especially for the difference between the CP- and EP-El Niño events. When using the ICM forced by the EQeast-TEs to reproduce the CP- and EP-El Niño events, the horizontal distributions of the reproduced SST anomalies are the same, with the positive value concentrated on the eastern tropical Pacific. The failure of reproducing CP-El Niño is dynamically due to the simulation that fails to capture the convergence of the wind anomaly in the central tropical Pacific and the rapid SST warming from spring to summer. When the EQcenter-TEs are superimposed on the SST tendency of the ICM, the reproduced CP-El Niño can be well-identified since the EQcenter-TEs effectively modify both the wind tendency and the zonal advection feedback to yield a more realistic air-sea condition that is favoured to the CP-El Niño formulation.

One key hypothesis in the present study is that NFSV-TEs are analogues of the combined effect of various model errors on the SST tendency equation. NFSV-TEs do make sense physically, even though they are mathematical products. In practice, most models present poorest skills in the far eastern equatorial Pacific as the region is where the SST variability has strong non-linearity (Takahashi and Dewitte, 2016; Ding *et al.*, 2018; Ying *et al.*, 2019). Thus, the E-type NFSV-TEs here may results from the underestimate nonlinearity processes in the eastern tropical Pacific. Given the results of the different impacts of EQeast-TEs and EQcenter-TEs, it is

therefore meaningful to study their sources. Hereon, correlation analyses are performed between the focused NFSV-TEs and SST anomalies in the globe to snoop on the potential physical factors. As shown in Figure 15b, the EQeast-TEs are significantly related to the SST in the tropical Pacific and the Indian Ocean, implying that the model uncertainties of the ICM over the eastern tropical Pacific are due to the missed interaction between the tropical Pacific and the Indian Ocean. Because the EQeast-TEs almost explained the dominant mode of the NFSV-TEs, the E-type NFSV-TEs are also significantly related to the SST anomaly in the eastern tropical Pacific (Figure 15c). The strong E-type NFSV-TEs tend to influence the basin-scale zonal circulation and lead to a dipole error pattern of sea level pressure (SLP) over the tropical Pacific, as revealed in the correlation map between the PC of E-type and SLP anomaly (Figure 15d). The sources of the EQcenter-TEs are slightly more involved than those of EQeast-TEs (Figure 15a). The EQcenter-TEs are more likely associated with the teleconnection of the extratropical Pacific. For example, the EQcenter-TEs are significantly related to the SST anomalies in the North Pacific, where the correlation map presents a Victoria Mode (VM)-like pattern (Bond *et al.*, 2003). In a sense, the model uncertainties in the central-eastern tropical Pacific are a result of the effect of the VM. Combined with previous studies of the VM (Ding *et al.*, 2015; Chen *et al.*, 2021), we conclude that improving the simulation of the VM is effective in reducing the model uncertainties in the central-eastern tropical Pacific to effectively improve the predictions concerning two types of El Niño events.

AUTHOR CONTRIBUTIONS

Lingjiang Tao: Conceptualization; data curation; formal analysis; investigation; methodology; resources; software; validation; visualization; writing – original draft; writing – review and editing. **Wansuo Duan:** Conceptualization; funding acquisition; methodology; project administration; resources; supervision; writing – review and editing. **Lin Jiang:** Resources; writing – review and editing.

ACKNOWLEDGEMENTS

The authors thank the anonymous reviewers for their constructive comments and suggestions on improving the manuscript. This work was supported by the National Natural Science Foundation of China (Grant Nos. 41525017 and 41930971) and the China Postdoctoral Science Foundation (Grant No. 2020M681155).

DATA AVAILABILITY STATEMENT

The SST reanalysis data is from the third version of the National Oceanic and Atmospheric Administration

(NOAA) Extended Reconstructed SST, which can be freely downloaded at the website <http://iridl.ldeo.columbia.edu/SOURCES/.NOAA/.NCDC/.ERSST/.version3b/.sst/>.

ORCID

Wansuo Duan  <https://orcid.org/0000-0002-0122-2794>

REFERENCES

- Ashok, K., Behera, S.K., Rao, S.A., Weng, H.Y. and Yamagata, T. (2007) El Niño Modoki and its possible teleconnection. *Journal of Geophysical Research, Oceans*, 112, C11007.
- Bond, N.A., Overland, J.E., Spillane, M. and Stabenro, P. (2003) Recent shifts in the state of the North Pacific. *Geophysical Research Letters*, 30, 2183.
- Chen, D., Cane, M.A., Kaplan, A., Zebiak, S.E. and Huang, D.J. (2004) Predictability of El Niño over the past 148 years. *Nature*, 428, 733–736.
- Chen, S.F., Yu, B., Wu, R.G., Chen, W. and Song, L.Y. (2021) The dominant North Pacific atmospheric circulation patterns and their relations to Pacific SSTs: historical simulations and future projections in the IPCC AR6 models. *Climate Dynamics*, 56, 701–725.
- Ding, H., Newman, M., Alexander, M.A. and Wittenberg, A.T. (2018) Skillful climate forecasts of the tropical Indo-Pacific Ocean using model-analogs. *Journal of Climate*, 31(14), 5437–5459.
- Ding, R.Q., Li, J.P., Tseng, Y.H., Sun, C. and Guo, Y.P. (2015) The Victoria mode in the North Pacific linking extratropical sea level pressure variations to ENSO. *Journal of Geophysical Research-Atmospheres*, 120, 27–45.
- Duan, W. and Mu, M. (2018) Predictability of El Niño-southern oscillation events. *Oxford Research Encyclopedia of Climate Science*. <https://oxfordre.com/climatescience/view/10.1093/acrefore/9780190228620.001.0001/acrefore-9780190228620-e-80>
- Duan, W.S. and Hu, J.Y. (2016) The initial errors that induce a significant “spring predictability barrier” for El Niño events and their implications for target observation: results from an earth system model. *Climate Dynamics*, 46, 3599–3615.
- Duan, W.S., Tian, B. and Xu, H. (2014) Simulations of two types of El Niño events by an optimal forcing vector approach. *Climate Dynamics*, 43, 1677–1692.
- Duan, W.S., Zhao, P., Hu, J.Y. and Xu, H. (2016) The role of nonlinear forcing singular vector tendency error in causing the “spring predictability barrier” for ENSO. *Journal of Meteorological Research*, 30, 853–866.
- Duan, W.S. and Zhou, F.F. (2013) Non-linear forcing singular vector of a two-dimensional quasi-geostrophic model. *Tellus A*, 65, 256.
- Feng, J., Lian, T., Ying, J., Li, J.D. and Li, G. (2020) Do CMIP5 models show El Niño diversity? *Journal of Climate*, 33, 1619–1641.
- Gao, C., Wu, X.R. and Zhang, R.H. (2016) Testing a four-dimensional variational data assimilation method using an improved intermediate coupled model for ENSO analysis and prediction. *Advances in Atmospheric Sciences*, 33, 875–888.
- Ham, Y.G. and Kug, J.S. (2012) How well do current climate models simulate two types of El Niño? *Climate Dynamics*, 39, 383–398.
- Hansen, J.W., Hodges, A.W. and Jones, J.W. (1998) ENSO influences on agriculture in the southeastern United States. *Journal of Climate*, 11, 404–411.
- Hou, M.Y., Duan, W.S. and Zhi, X.F. (2019) Season-dependent predictability barrier for two types of El Niño revealed by an approach to data analysis for predictability. *Climate Dynamics*, 53, 5561–5581.
- Jin, E.K., Kinter, J.L., III, Wang, B., Park, C.-K., Kang, I.-S., Kirtman, B.P., Kug, J.-S., Kumar, A., Luo, J.-J., Schemm, J., Shukla, J. and Yamagata, T. (2008) Current status of ENSO prediction skill in coupled ocean-atmosphere models. *Climate Dynamics*, 31, 647–664.
- Keenlyside, N. and Kleeman, R. (2002) Annual cycle of equatorial zonal currents in the Pacific. *Journal of Geophysical Research, Oceans*, 107, 8-1–8-13.
- Leloup, J., Lengaigne, M. and Boulanger, J.P. (2008) Twentieth century ENSO characteristics in the IPCC database. *Climate Dynamics*, 30, 277–291.
- Liu, T., Tang, Y., Yang, D., Cheng, Y., Song, X., Hou, Z., Shen, Z., Gao, Y., Wu, Y., Li, X. and Zhang, B. (2019) The relationship among probabilistic, deterministic and potential skills in predicting the ENSO for the past 161 years. *Climate Dynamics*, 53, 6947–6960.
- Lorenz, E.N. (1975). Climatic predictability in the physical basis of climate and climate modeling. International Study Conference (29 July–10 August 1974; Stockholm, Sweden), WMO GARP Publications Series No. 16, World Meteorological Organization (WMO), International Council of Scientific Unions, pp. 132–136.
- Mu, M., Duan, W., Wang, Q. and Zhang, R. (2010) An extension of conditional nonlinear optimal perturbation approach and its applications. *Nonlinear Processes in Geophysics*, 17, 211–220.
- Mu, M., Duan, W.S. and Wang, B. (2007) Season-dependent dynamics of nonlinear optimal error growth and El Niño-southern oscillation predictability in a theoretical model. *Journal of Geophysical Research-Atmospheres*, 112, D10113.
- O’Kane, T.J., Sandery, P.A., Monselesan, D.P., Sakov, P., Chamberlain, M.A., Matear, R.J., Collier, M.A., Squire, D.T. and Stevens, L. (2019) Coupled data assimilation and ensemble initialization with application to multiyear ENSO prediction. *Journal of Climate*, 32, 997–1024.
- Philander, S.G.H. (1983) El-Niño southern oscillation phenomena. *Nature*, 302, 295–301.
- Ren, H.-L., Scaife, A.A., Dunstone, N., Tian, B., Liu, Y., Ineson, S., Lee, J.-Y., Smith, D., Liu, C., Thompson, V., Vellinga, M. and MacLachlan, C. (2018) Seasonal predictability of winter ENSO types in operational dynamical model predictions. *Climate Dynamics*, 52, 3869–3890.
- Ropelewski, C.F. and Halpert, M.S. (1987) Global and regional scale precipitation patterns associated with the El-Niño southern oscillation. *Monthly Weather Review*, 115, 1606–1626.
- Schneider, E.K., DeWitt, D.G., Rosati, A., Kirtman, B.P., Ji, L. and Tribbia, J.J. (2003) Retrospective ENSO forecasts: sensitivity to atmospheric model and ocean resolution. *Monthly Weather Review*, 131, 3038–3060.
- Shuai, J.B., Zhang, Z., Sun, D.Z., Tao, F.L. and Shi, P.J. (2013) ENSO, climate variability and crop yields in China. *Climate Research*, 58, 133–148.

Smith, T.M., Reynolds, R.W., Peterson, T.C. and Lawrimore, J. (2008) Improvements to NOAA's historical merged land-ocean surface temperature analysis (1880-2006). *Journal of Climate*, 21, 2283–2296.

Solow, A.R., Adams, R.F., Bryant, K.J., Legler, D.M., O'Brien, J.J., McCarl, B.A., Nayda, W. and Weiher, R. (1998) The value of improved enso prediction to US agriculture. *Climatic Change*, 39, 47–60.

Takahashi, K. and Dewitte, B. (2016) Strong and moderate nonlinear El Nio regimes. *Climate Dynamics*, 46, 1627–1645.

Takahashi, K., Montecino, A., Goubanova, K. and Dewitte, B. (2011) ENSO regimes: reinterpreting the canonical and Modoki El Niño. *Geophysical Research Letters*, 38(10), L10707.

Tan, X.X., Tang, Y.M., Lian, T., Zhang, S.W., Liu, T. and Chen, D. K. (2020) Effects of Semistochastic westerly wind bursts on ENSO predictability. *Geophysical Research Letters*, 47, e2019GL086828.

Tao, L.J. and Duan, W.S. (2019) Using a nonlinear forcing singular vector approach to reduce model error effects in ENSO forecasting. *Weather and Forecasting*, 34, 1321–1342.

Tao, L.J., Duan, W.S. and Vannitsem, S. (2020) Improving forecasts of El Niño diversity: a nonlinear forcing singular vector approach. *Climate Dynamics*, 55, 739–754.

Tao, L.J., Gao, C. and Zhang, R.H. (2019) Model parameter-related optimal perturbations and their contributions to El Niño prediction errors. *Climate Dynamics*, 52, 1425–1441.

Timmermann, A., An, S.I., Kug, J.S., Jin, F.F., Cai, W., Capotondi, A., Cobb, K.M., Lengaigne, M., McPhaden, M.J., Stuecker, M.F., Stein, K., Wittenberg, A.T., Yun, K.S., Bayr, T., Chen, H.C., Chikamoto, Y., Dewitte, B., Dommenges, D., Grothe, P., Guilyardi, E., et al. (2018) El Niño-southern oscillation complexity. *Nature*, 559, 535–545.

Wang, C.Z., Zhang, L.P., Lee, S.K., Wu, L.X. and Mechoso, C.R. (2014) A global perspective on CMIP5 climate model biases. *Nature Climate Change*, 4, 201–205.

Wu, X.R., Han, G.J., Zhang, S.Q. and Liu, Z.Y. (2016) A study of the impact of parameter optimization on ENSO predictability with an intermediate coupled model. *Climate Dynamics*, 46, 711–727.

Xue, Y., Chen, M.Y., Kumar, A., Hu, Z.Z. and Wang, W.Q. (2013) Prediction skill and bias of tropical Pacific Sea surface temperatures in the NCEP climate forecast system version 2. *Journal of Climate*, 26, 5358–5378.

Yao, J.W., Duan, W.S. and Qin, X.H. (2021) Which features of the SST forcing error Most likely disturb the simulated intensity of tropical cyclones? *Advances in Atmospheric Sciences*, 38, 581–602.

Ying, J., Huang, P., Lian, T. and Tan, H. (2019) Understanding the effect of an excessive cold tongue bias on projecting the tropical pacific sst warming pattern in cmip5 models. *Climate Dynamics*, 52, 1805–1818.

Zebiak, S.E. and Cane, M.A. (1987) A model El-Niño southern oscillation. *Monthly Weather Review*, 115, 2262–2278.

Zhang, R.H. and Gao, C. (2016) The IOCAS intermediate coupled model (IOCAS ICM) and its real-time predictions of the 2015–2016 El Niño event. *Scientific Bulletin*, 61, 1061–1070.

Zhang, R.H., Zebiak, S.E., Kleeman, R. and Keenlyside, N. (2003) A new intermediate coupled model for El Niño simulation and prediction. *Geophysical Research Letters*, 30, 2012.

Zhang, R.H., Zebiak, S.E., Kleeman, R. and Keenlyside, N. (2005) Retrospective El Niño forecasts using an improved intermediate coupled model. *Monthly Weather Review*, 133, 2777–2802.

Zheng, F., Wang, H. and Zhu, J. (2009) ENSO ensemble prediction: initial error perturbations vs. model error perturbations. *Chinese Science Bulletin*, 54, 2516–2523.

Zheng, F. and Zhu, J. (2016) Improved ensemble-mean forecasting of ENSO events by a zero-mean stochastic error model of an intermediate coupled model. *Climate Dynamics*, 47, 3901–3915.

Zheng, F., Zhu, J. and Zhang, R.H. (2007) Impact of altimetry data on ENSO ensemble initializations and predictions. *Geophysical Research Letters*, 34(13), 256–260.

How to cite this article: Tao, L., Duan, W., & Jiang, L. (2022). Model errors of an intermediate model and their effects on realistic predictions of El Niño diversity. *International Journal of Climatology*, 1–22. <https://doi.org/10.1002/joc.7656>

APPENDIX

For an air–sea coupling system, a relationship exists between wind and SST, which can be represented as

$$\tau = g(\mathbf{x}), \quad (\text{A1})$$

in which $\mathbf{x}$ and τ are the SST anomaly and wind anomaly, respectively. The SST anomaly $\mathbf{x}$ is controlled by a type of partial differential equation (i.e. SST model) that can be symbolically written as

$$\frac{\partial \mathbf{x}}{\partial t} = \mathbf{F}(\mathbf{x}), \quad (\text{A2})$$

where $\mathbf{F}$ is the nonlinear operator of the SST model.

Adding a perturbation of SST anomaly, we have

$$\tau + \Delta\tau = g(\mathbf{x} + \Delta\mathbf{x}). \quad (\text{A3})$$

When the $\Delta\mathbf{x}$ is small enough, with Equation (A1) and Equation (A3), we obtain

$$\begin{aligned} \Delta\tau &= g(\mathbf{x} + \Delta\mathbf{x}) - g(\mathbf{x}) \\ &= \frac{\partial g}{\partial \mathbf{x}} \cdot \Delta\mathbf{x} + O(\Delta\mathbf{x}^2). \\ &\approx \frac{\partial g}{\partial \mathbf{x}} \cdot \Delta\mathbf{x} \end{aligned} \quad (\text{A4})$$

Given the time interval, we then have

$$\frac{\Delta \tau}{\Delta t} = \mathbf{G} \cdot \frac{\Delta \mathbf{x}}{\Delta t}, \quad (\text{A5})$$

$$\frac{\Delta \tau_n}{\Delta t} = \mathbf{G} \cdot \frac{\Delta \mathbf{x}_n}{\Delta t}. \quad (\text{A7})$$

where $\mathbf{G} = \frac{\partial \mathbf{g}}{\partial \mathbf{x}}$. That is, the relationship between the tendencies of the wind and SST anomalies is determined by $\mathbf{G}$.

When a forcing vector $\mathbf{f}$ is added to Equation (A2), a new SST model can be obtained as

$$\frac{\partial \mathbf{x}_n}{\partial t} = \mathbf{F}(\mathbf{x}_n) + \mathbf{f}. \quad (\text{A6})$$

Since the SST-wind relationship is not changed, the tendency relationship in SST and wind is also the same as Equation (A5), here denoted as

Subtracting Equation (A7) from Equation (A5), finally we get

$$\frac{\Delta \tau_n}{\Delta t} - \frac{\Delta \tau}{\Delta t} \approx \mathbf{G} \cdot \mathbf{f}. \quad (\text{A8})$$

This indicates that, for a small perturbation of the SST anomaly, the tendency difference of the wind anomaly is approximately related to the forcing vector $\mathbf{f}$, where the relationship is determined by the derivative of the wind model (i.e. $\mathbf{G}$).