

Using a Nonlinear Forcing Singular Vector Approach to Reduce Model Error Effects in ENSO Forecasting

TAO LINGJIANG AND DUAN WANSUO

LASG, Institute of Atmospheric Physics, Chinese Academy of Sciences, and University of Chinese Academy of Sciences, Beijing, China

(Manuscript received 8 March 2019, in final form 30 May 2019)

ABSTRACT

Nonlinear forcing singular vector (NFSV)-based assimilation is adopted to determine the model tendency errors that represent the combined effect of different kinds of model errors; then, an NFSV-tendency error forecast model is formulated. This error forecast model is coupled with an intermediate complex model (ICM) and makes the ICM output closer to the observations; finally, an NFSV-ICM forecast model for ENSO is constructed. The competing aspect of the NFSV-ICM is to consider not only model errors but also the interaction between model errors and initial errors because of the mathematical nature of the NFSV-tendency errors. Based on the prediction experiments for tropical SSTAs during either the training period (1960–96; i.e., when the NFSV-ICM is formulated) or the cross-validation period (1997–2016), the NFSV-ICM is determined to have a much higher forecast skill in predicting ENSO that, specifically, extends the skillful predictions of ENSO from a lead time of 6 months in the original ICM to a lead time of 12 months. The higher skill of the NFSV-ICM is especially reflected in the predictions of SSTAs in the central and western Pacific. For the well-known spring predictability barrier (SPB) phenomenon that greatly limits ENSO forecasting skill, the NFSV-ICM also shows great abilities in suppressing its negative effect on ENSO predictions. Although the NFSV-ICM is presently only involved with the NFSV-related assimilation of SSTs, it has shown its usefulness in predicting ENSO. It is clear that the NFSV-based assimilation approach is effective in dealing with the effect of model errors on ENSO forecasts.

1. Introduction

El Niño–Southern Oscillation (ENSO), which is known as the dominant interannual mode in the tropical Pacific, has been the focus of scientists and the public over the past several decades because of its global impact on climate/weather (McPhaden et al. 2006). Through efforts over the decades, there has been significant progress toward observing, understanding, and simulating ENSO (Zebiak and Cane 1987; Jin 1997a,b; McPhaden et al. 1998; Timmermann et al. 2018). Consequently, the skillful predictions of Niño-3.4 SST anomalies related to ENSO have been achieved 6–12 months in advance (even up to 2 years) in hindcast experiments (Chen and Cane 2008; Luo et al. 2008).

To date, more than 20 climate models have been used to routinely predict ENSO in real time (see the website at <https://iri.columbia.edu/our-expertise/climate/enso/>). However, the skillful and realistic forecasting of ENSO can only be made with, at most, a 6-month lead (Barnston et al. 2012).

Initial errors have been recognized as one of the main factors that limit ENSO forecast skills. To reduce the impact of initial errors, great efforts have been made to optimize initial conditions under model constraints and observations by data assimilation (e.g., Chen et al. 1997; Behringer et al. 1998; Sugiura et al. 2008; Gao et al. 2016; etc.). In addition, intensified observations in some key areas (i.e., targeted observations) were suggested to provide the most useful observations for data assimilation to suppress initial error growth and provide a

 Denotes content that is immediately available upon publication as open access.

Corresponding author: Wansuo Duan, duanws@lasg.iap.ac.cn



This article is licensed under a [Creative Commons Attribution 4.0 license](http://creativecommons.org/licenses/by/4.0/) (<http://creativecommons.org/licenses/by/4.0/>).

DOI: 10.1175/WAF-D-19-0050.1

© 2019 American Meteorological Society

skillful prediction of ENSO (Mu et al. 2015; Duan and Hu 2016; Hu and Duan 2016; Duan and Feng 2017; Tao et al. 2017, 2018). For example, Morss and Battisti (2004a,b), based on the observation system simulation experiment (OSSE), suggested that for ENSO forecasting longer than a few months, the most important area for observations is the eastern equatorial Pacific and its south; the secondary region of importance is the western equatorial Pacific. By using sequential importance sampling assimilation method, Kramer and Dijkstra (2013) also showed that the optimal observation locations for SST are located in the eastern tropical Pacific for minimizing the uncertainty in the Niño-3 index.

In addition to initial errors, an increasing number of studies have also emphasized the importance of model errors in yielding ENSO prediction uncertainties (Latif et al. 2001; Stainforth et al. 2005; Zheng et al. 2009a; Wu et al. 2016; Tao et al. 2019). For instance, Jin et al. (2008) noted that the accuracy of ENSO variability is related to the simulated climatology as a result of a poor prediction skill when using models with climatological biases.

As one source of model errors, the model parametric errors (MPEs) were also found to have a role in yielding model systematic errors and influencing ENSO variability (Bejarano and Jin 2008; Macmynowski and Tziperman 2008; Zhu and Zhang 2018). Tao et al. (2019) investigated the impact of MPEs on ENSO predictions from the perspective of optimal error growth and indicated that MPEs have the potential to influence the strength of the Bjerknes feedback, which is crucial to the development of SST anomalies, thus disturbing ENSO predictions. In addition, they also demonstrated that the spring predictability barrier (SPB) phenomenon of ENSO can also be caused by MPEs besides initial errors. To reduce the effect of the model errors caused by the MPEs, Zhang et al. (2018) made a skillful hindcast of the strong El Niño event in 2015 by optimizing two key parameters associated with Bjerknes feedback. In particular, Wu et al. (2016) reduced the SPB phenomenon of ENSO and extended the valid lead time by considering multiple physical parameters determined by the ensemble adjustment of the Kalman filter.

Missing some processes in models also tends to induce model errors and affect ENSO simulations and predictions. For example, westerly wind bursts (WWBs) have been shown to have the ability to excite the onset of El Niño events (Chen et al. 2015). Lopez and Kirtman (2014) demonstrated that the inclusivity of state-dependent WWBs in an ENSO model can greatly improve the ENSO prediction skill. Therefore, when the model lacks the WWB effect and yields model errors,

the ENSO forecast skill can be largely influenced. Yu et al. (2003) suggested that the characteristics of WWBs depend on the large-scale SST field and therefore indicated that the model errors induced by the lack of WWBs may be of a certain structure and significantly disturb the predictions of ENSO. Of course, model errors are not only caused by some physical processes that do not appear in models but also due to those that are highly simplified. For an intermediate coupled model, model errors are significant and from different sources (Qi et al. 2017); moreover, it is hard to distinguish their effects in predictions. Therefore, improving some processes in models may fail to enhance the ENSO prediction skills due to the interaction of other model errors. Finding a way to obtain or filter the combined effect of kinds of model errors is promising to improve the ENSO prediction.

To this end, Zheng and Zhu (2016) considered synthetically the model uncertainties and developed a first-order Markov stochastic model that is added to the tendency equation for SST in the intermediate complex model (ICM) in an attempt to depict the approximately combined effect of model uncertainties. Such an approach has the ability to filter unpredictable stochastic processes and capture more realistic ENSO evolutions (Zheng et al. 2009a). But there are still large systematic biases since the Markov stochastic model perturbation cannot approximate the effect of all model errors to the greatest extent (Qi et al. 2017). Therefore, an optimal method was suggested to represent the combined effect of different kinds of model errors. Duan and Zhou (2013) extended the (linear) forcing singular vector (Barkmeijer et al. 2003) to a nonlinear field and proposed the approach of a nonlinear forcing singular vector (NFSV), which depicts the model tendency errors that have the largest effect on prediction uncertainty (Duan et al. 2016; Qi et al. 2017). Then with the idea of data assimilation, Duan et al. (2014) improved the NFSV approach so that it can be used to extract the comprehensive tendency errors (see section 2). Using the NFSV-related assimilation, they obtained the tendency errors of the Zebiak–Cane model (Zebiak and Cane 1987), which was in turn forced on the SST equation of the Zebiak–Cane model in an attempt to offset the effects of the model errors. As a result, they successfully simulated realistic ENSO evolutions using the Zebiak–Cane model equipped with the NFSV-related assimilation.

The impressive effect of NFSV on correcting model stimulates us to apply the NFSV to improve the ENSO prediction. It should be pointed out that the model errors are time dependent and the NFSV-tendency errors during predictions are case dependent. Thus, to improve

the prediction, we have to obtain the mode tendency errors in advance. The observation data are available so that we can obtain the NFSV-tendency errors during the assimilation period. However, the NFSV-tendency errors beyond the assimilation period are unavailable due to unavailable observations. Therefore, can we use the NFSV-related assimilation for predictions? If yes, how do we do it? In doing so, can the prediction skill be greatly improved by using the NFSV-related assimilation?

To address these above questions, we develop an approach to apply the NFSV-related assimilation in ENSO predictions and establish a new ENSO forecast system by using an intermediate coupled model. The study is classified into two parts: the first part introduces the new ENSO forecast system, with the NFSV-related assimilation, and its performance in ENSO predictions; the second part examines the prediction skill in distinguishing El Niño types (i.e., the EP and CP El Niño) to explain the physical reasons for improving the prediction skill associated with the various types of ENSO. In the present article, we focus on the first part (i.e., the new ENSO forecast system and its performance in ENSO predictions).

The remainder of this paper is organized as follows. Section 2 provides a description of the NFSV approach and its related assimilation. Section 3 describes the intermediate coupled model for ENSO (ICM) used in the present study. In section 4, we develop the new ENSO forecast system based on the NFSV-related assimilation and the ICM. In section 5, the performance of the NFSV-ICM is examined. Finally, we present a summary and a discussion in section 6.

2. Methods

In the present study, we use the NFSV-related assimilation approach to correct the ENSO model and establish the ENSO forecast system. For convenience, we briefly describe the idea behind the NFSV and NFSV-related assimilation.

a. The NFSV approach

The NFSV is a nonlinear extension of the (linear) forcing singular vector proposed by Barkmeijer et al. (2003). It considers the combined effect of different kinds of model errors and represents the tendency error that causes the largest prediction error at the prediction time (Duan and Zhou 2013; Duan and Zhao 2015). The NFSV \mathbf{f}^* can be derived from the maximization problem shown in Eq. (1):

$$J(\mathbf{f}^*) = \max_{\mathbf{f} \in \Omega} \|\mathbf{M}_\tau(\mathbf{f})(\mathbf{u}_0) - \mathbf{M}_\tau(0)(\mathbf{u}_0)\|, \quad (1)$$

where $\mathbf{M}_\tau(0)$ represents the propagator of a nonlinear model [Eq. (2)] and $\mathbf{M}_\tau(\mathbf{f})$ represents that of the

nonlinear model but with a tendency perturbation \mathbf{f} [see Eq. (3)]; \mathbf{u} represents the state variable and \mathbf{u}_0 denotes its initial state. The term J represents the cost function that measures the deviation in the state from the reference state $\mathbf{M}_\tau(0)(\mathbf{u}_0)$ in terms of the norm $\|\cdot\|$ due to the effect of the tendency perturbation \mathbf{f} . Here, $\mathbf{f} \in \Omega$ tells the constraint of the tendency perturbation. Usually, we define $\|\mathbf{f}\| \leq \delta$, which means the tendency error is less than δ [δ is a positive number; the details can be seen in Duan and Zhou (2013) and Duan et al. (2016)]:

$$\begin{cases} \frac{\partial \mathbf{u}}{\partial t} = F(\mathbf{u}, t) \\ \mathbf{u}|_{t=0} = \mathbf{u}_0 \end{cases}, \quad (2)$$

$$\begin{cases} \frac{\partial \mathbf{u}}{\partial t} = F(\mathbf{u}, t) + \mathbf{f} \\ \mathbf{u}|_{t=0} = \mathbf{u}_0 \end{cases}. \quad (3)$$

From Eq. (1), it is easily known that if $\mathbf{M}_\tau(0)(\mathbf{u}_0)$ is a control forecast, the NFSV represents one kind of optimal tendency perturbation of the control forecast and has the potential to cause the perturbed forecast largest departure from the control forecast at the prediction time. That is, the control forecast $\mathbf{M}_\tau(0)(\mathbf{u}_0)$ is most sensitive to the NFSV-tendency perturbation, which may provide information on correcting the control forecast skill. Of course, we can also modify the cost function of the NFSV according to a particular physical problem. For example, to obtain a tendency perturbation that can induce the largest uncertainties within the prediction period from t_0 to τ , the cost function can be modified as

$$J(\mathbf{f}^*) = \max_{\mathbf{f} \in \Omega} \sum_{t=t_0}^{\tau} \|\mathbf{M}_t(\mathbf{f})(\mathbf{u}_0) - \mathbf{M}_t(0)(\mathbf{u}_0)\|. \quad (4)$$

In this case, the cost function J gives the largest accumulated departure from the control forecast that is merely induced by the tendency error.

b. The NFSV-related assimilation

Based on the NFSV approach, Duan et al. (2014) proposed an optimal forcing vector (OFV) approach to offset the model errors and improve the model simulation ability, in which they simply modified the cost function J of the NFSV approach [see Eq. (5)]. The OFV represents total tendency perturbation, which is superimposed on the tendency equation of the model and makes the model simulation closest to the observations. The OFV, which is denoted by \mathbf{f}^o , is related to a minimization (or assimilation) problem:

$$J(\mathbf{f}^o) = \min_{\mathbf{f}} \sum_{t=t_0}^{\tau} \|\mathbf{M}_t(\mathbf{f})(\mathbf{u}_0) - \mathbf{u}^{\text{obs}}(\mathbf{x}, t)\|, \quad (5)$$

where $\mathbf{u}^{\text{obs}}(\mathbf{x}, t)$ denotes the time series of the observation data and $\mathbf{M}_t(\mathbf{f})(\mathbf{u}_0)$ represents the model simulation result at the t month, which is obtained by integrating Eq. (3) with an initial state \mathbf{u}_0 .

The assimilation problem Eq. (5) is identical to the following optimization problem:

$$J(\mathbf{f}^o) = \max_{\mathbf{f} \in \Omega} \sum_{t=t_0}^{\tau} [\|\mathbf{M}_t(\mathbf{f})(\mathbf{u}_0) - \mathbf{M}_t(0)(\mathbf{u}_0)\|],$$

$$\Omega = \left\{ \mathbf{f} \left| J(\mathbf{f}) = \min_{\mathbf{f}} \sum_{t=t_0}^{\tau} [\|\mathbf{M}_t(\mathbf{f})(\mathbf{u}_0) - \mathbf{u}^{\text{obs}}(\mathbf{x}, t)\|] \right. \right\}. \quad (6)$$

That is, the OFV derived by Eq. (5) satisfies Eq. (6), while Eq. (6) bears greatly resemblance to Eq. (4) for the NFSV approach. In this situation, the OFV can be understood as a tendency perturbation that is constrained by $\mathbf{f} \in \Omega$, which makes the perturbed forecast depart from the control forecast (i.e., the unperturbed forecast) at the greatest extent but remain closest to the observation due to the constraint condition. In this sense, the OFV is mathematically consistent with the idea of the NFSV. For convenience, here, we rename the ‘‘OFV’’ as the ‘‘NFSV’’, and the related calculation is called the ‘‘NFSV-related assimilation.’’ The NFSV-related assimilation (which is used to treat model errors) can show its difference from the initial value assimilation (which is being used to deal with the initial errors). In the present study, we use the NFSV-related assimilation to correct the ICM for ENSO and establish a new ENSO forecast system.

3. ENSO model and data

The ENSO model adopted here is an ICM developed by Zhang et al. (2003), and its forecast results have been presented on the International Research Institute for Climate and Society (IRI) web page, showing real-time ENSO forecast results generated by more than 20 models across the globe (see https://iri.columbia.edu/our-expertise/climate/forecasts/enso/current/?enso_tab=enso-sst_table). The ICM is an air–sea coupled model that consists of a statistical wind stress model, an intermediate dynamic ocean model (Keenlyside and Kleeman 2002), and an SST anomaly model that represents surface ocean thermodynamics. The wind stress model is constructed from the singular value decomposition (SVD) approach that determines the relationship between the SST and the wind field from

1963 to 1996. To represent the effect of thermocline fluctuations on SST variability, an empirical model is developed to parameterize the temperature of the subsurface water entrained into the mixed layer (T_e) from the sea surface height. As a result, the SST anomaly model with the T_e model can capture the realistic features of ENSO evolution, including the period and amplitude (Zhang et al. 2005). However, since some processes are missing in this model, such as the effects of freshwater flux and salinity, which were found to have a role in the amplitude of ENSO (e.g., Zhang et al. 2012), the ICM fails to well capture the strength of ENSO in realistic predictions (Zhang and Gao 2016; Zheng and Zhu 2016). In addition, the ICM only represents the air–sea interaction in the tropical Pacific and ignores the effect of the extratropical Pacific; therefore, considerable model errors still exist in the model itself for realistic prediction. Zheng et al. (2009a) found that initial perturbations-related ensemble predictions have small effect on improving ENSO predictions using the ICM, while model perturbations-based ensemble predictions show large improvement in ENSO predictions. Especially, Qi et al. (2017) showed that model errors in the ICM are more important than initial errors in realistic prediction. All these studies encourage us to use the NFSV approach to optimally capture the model errors so as to improve the ENSO prediction skills.

Observational and/or reanalysis data are required to initialize and evaluate the ICM in ENSO forecasting. In the present study, we follow Zhang et al. (2005) and adopt a simple nudging procedure to initialize the ICM by using SST observations (Barnett et al. 1993). The wind stress anomaly is reconstructed from the SST field during the period from 1854 to the start time of the forecast via the SVD-based historical SST–wind relation. Then, the reconstructed wind field is used to force the ocean model to initialize the ocean dynamic states. In addition, the observed SST anomalies are nudged into the SST anomaly model to generate the initial SST field. Here, the observed monthly SST field is from the National Oceanic and Atmospheric Administration (NOAA) Extended Reconstructed SST, version 3b dataset [ERSSTv3b; (Smith et al. 2008)], and the monthly wind stress field is from the National Centers for Environmental Prediction–National Center for Atmospheric Research (NCEP–NCAR) reanalysis (Kalnay et al. 1996). The former observed SST data span from 1854 to present, and the latter wind stress field spans from 1949 to the present.

As for other configurations of the ICM, the readers can refer to Zhang et al. (2005).

4. ENSO forecast system of the NFSV-related assimilation

Based on the ICM and its initial assimilation described in section 3, we develop the ENSO forecast system using the NFSV-related assimilation associated with the model error correction. The main idea is as follows. An NFSV-tendency error forecast model is built up and then coupled with the ICM, with the initial assimilation in section 3, to correct the ICM and achieve a useful skill that is substantially higher than that of the ICM that only has the initial assimilation. The ENSO forecast system with the NFSV-tendency error model can be achieved through the following three steps: (i) reveal the NFSV-type tendency error that makes the model simulation for SST anomalies closest to those that are observed by using the NFSV-related assimilation approach, (ii) derive a tendency error forecast model by considering a function of the pre-determined NFSV-tendency errors as the observed initial SST, and (iii) couple the NFSV-type tendency error forecast model with the ICM and finally establish the ENSO forecast system, which not only predicts the SSTA but also estimates the future tendency error.

a. The NFSV-type tendency error

Focusing on the SST predictions, we attribute the combined effect of the model errors to the SST tendency errors. The NFSV-related assimilation problem is constructed as follows:

$$J(\mathbf{f}_t^*) = \min_{t=t_0}^{t_0+n\Delta t} \|X(\mathbf{f}_t, \mathbf{u}_0) - X^{\text{obs}}(t)\|, \quad (7)$$

where X^{obs} and X denote the observed and simulated monthly SST anomalies within a 1-yr assimilation window $[t_0, t_0 + n\Delta t]$ ($\Delta t = 1$ month, $n = 12$) and \mathbf{u}_0 represents the initial analysis obtained by the initial value assimilation. Here \mathbf{f}_t^* represents the NFSV-type tendency error that can lead the model to output accurate SST anomalies. Here, the tendency errors are monthly dependent in the 1-yr assimilation window. That is, the tendency error is constant within one month to guarantee the mutually adjustment of ocean and atmospheric variables (Duan et al. 2014). Thus, according to Eq. (7), we can obtain a set of NFSV-type tendency errors with twelve components [i.e., $\mathbf{f}_{t_0, 12\Delta t}^* = (\mathbf{f}_{t_0+\Delta t}^* \dots, \mathbf{f}_{t_0+12\Delta t}^*)$] in the 1-yr assimilation window, namely, each month has one tendency error. Obviously, the NFSV-type tendency errors are somewhat different due to the different assimilation windows and initial times. For the period from 1854 to 2017,

we take each month as the initial month of the 1-yr assimilation window and calculate the NFSV-type tendency error based on the NFSV-related assimilation. Figure 1 shows a sketch diagram of the NFSV-related assimilation for different initial months. Then, we obtain a set of monthly dependent NFSV-type tendency errors with respect to the 12 components. For the NFSV-type tendency errors of different assimilation windows, we perform a composite analysis on the components whose months overlap in different assimilation windows. Then, we can obtain 164 monthly tendency errors during the period from 1854 to 2017.

A snapshot of the winter components of the NFSV-type tendency errors during El Niño episodes is displayed in Fig. 2. It is shown that large NFSV-type tendency errors are found near the north and south boundaries of the model. This is probably because the ICM is a regional model that has low skill in simulating the climate state near the model boundaries. It is also found that large tendency errors arise in the eastern Pacific cold tongue, which indicates that the intensity of the El Niño event simulated by the ICM has large errors. When further examining the relationship between the evolution of the observed SST anomalies and the corresponding NFSV-type tendency errors along the equator (see Fig. 3), it is illustrated that the NFSV-tendency errors exhibit an ENSO-like oscillation, of which the tendency errors are mainly located east of 160°W; furthermore, these errors are positive (negative) when the observed SSTs are warming (cooling). From the definition of the NFSV-type tendency errors in Eq. (7), it is inferred that the ICM tends to underestimate both El Niño and La Niña events in terms of their amplitudes due to the effects of model errors, while the NFSV-tendency errors can offset such effects (see Duan et al. 2014).

b. The NFSV-type tendency error forecast model

In step i, we obtain the NFSV-type tendency errors during different assimilation windows, which can correct the ICM to simulate the observed ENSO cycle. Note that the time-dependent observations adopted to determine the NFSV-tendency errors are available during the assimilation windows. However, for predictions, we do not have access to observations during the prediction time period and, thus, cannot obtain the corresponding NFSV-tendency errors by the approach in step i. By realizing the usefulness of the NFSV-related assimilation in correcting the model errors, we hope this assimilation can be used for the predictions. Therefore, how do we make the predictions via this type of assimilation?

According to Eq. (7), it is known that NFSV-tendency errors are dependent on the known observed SST

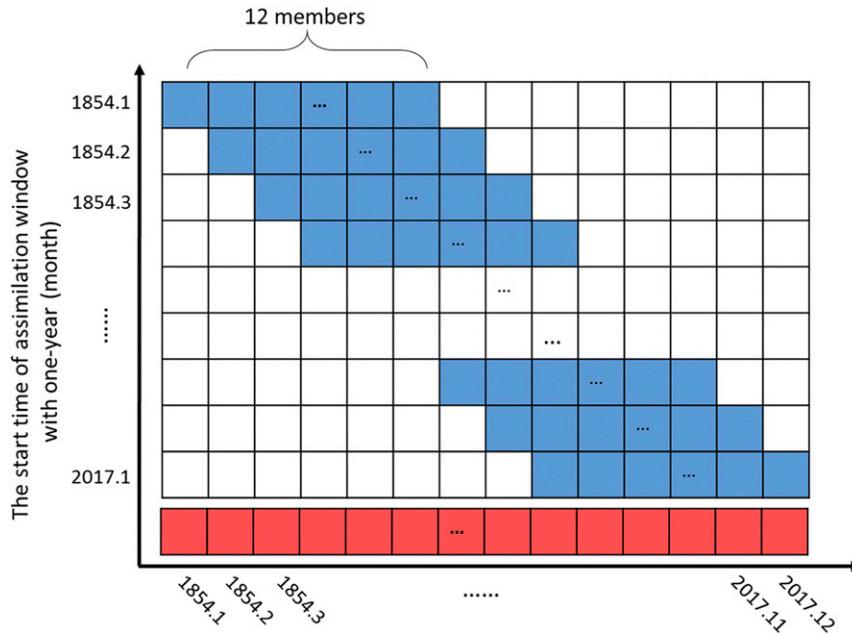


FIG. 1. Schematic diagram illustrating the NFSV-related assimilation windows and the strategy of the composite NFSV-tendency errors. The twelve blue shaded boxes in each row cover one assimilation window, with the initial month marked on the vertical axis, which denote the 12 members of the NFSV-tendency errors during the 1-yr assimilation window. The red shaded boxes on the last row represent the 164 monthly tendency errors during the period from 1854 to 2017, which are obtained by taking the ensemble mean of the members whose months overlap in different assimilation windows.

anomalies. Furthermore, we find that the related observation series during different assimilation windows correspond to different NFSV-tendency errors. That is, the NFSV-tendency errors are flow dependent. [Figure 3](#)

shows that a certain flow-dependent relation exists between the NFSV-type tendency error and the observed SST anomaly (see [section 4a](#)). That is, the NFSV-type tendency errors are positive (negative) in the eastern

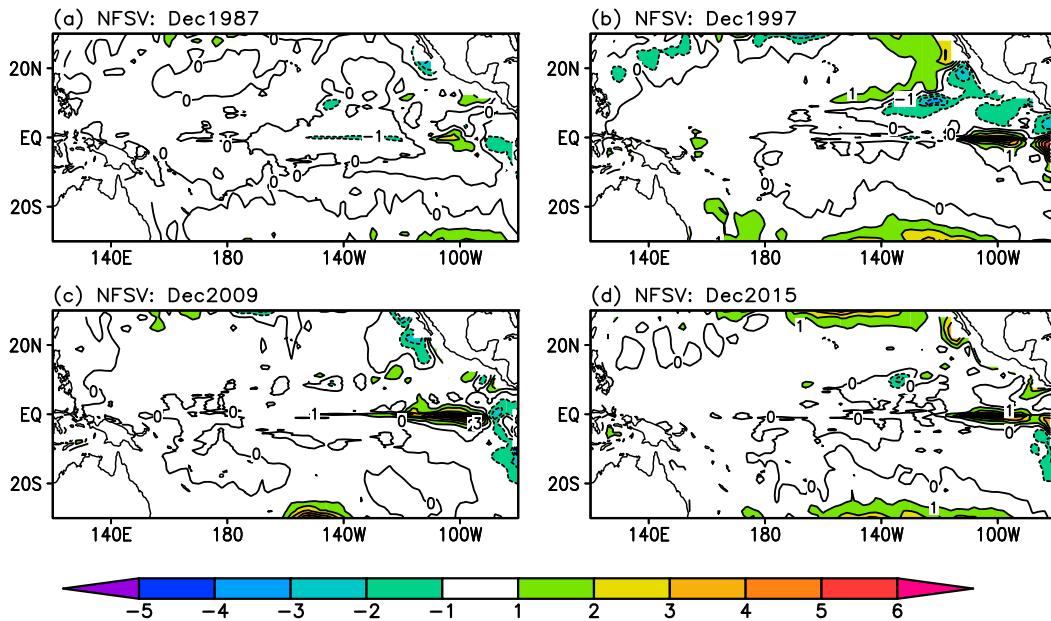


FIG. 2. Horizontal distributions of the NFSV-type tendency errors (10^{-6}C s^{-1}) during December of (a) the 1987 El Niño year, (b) the 1997 El Niño year, (c) the 2009 El Niño year, and (d) the 2015 El Niño year.

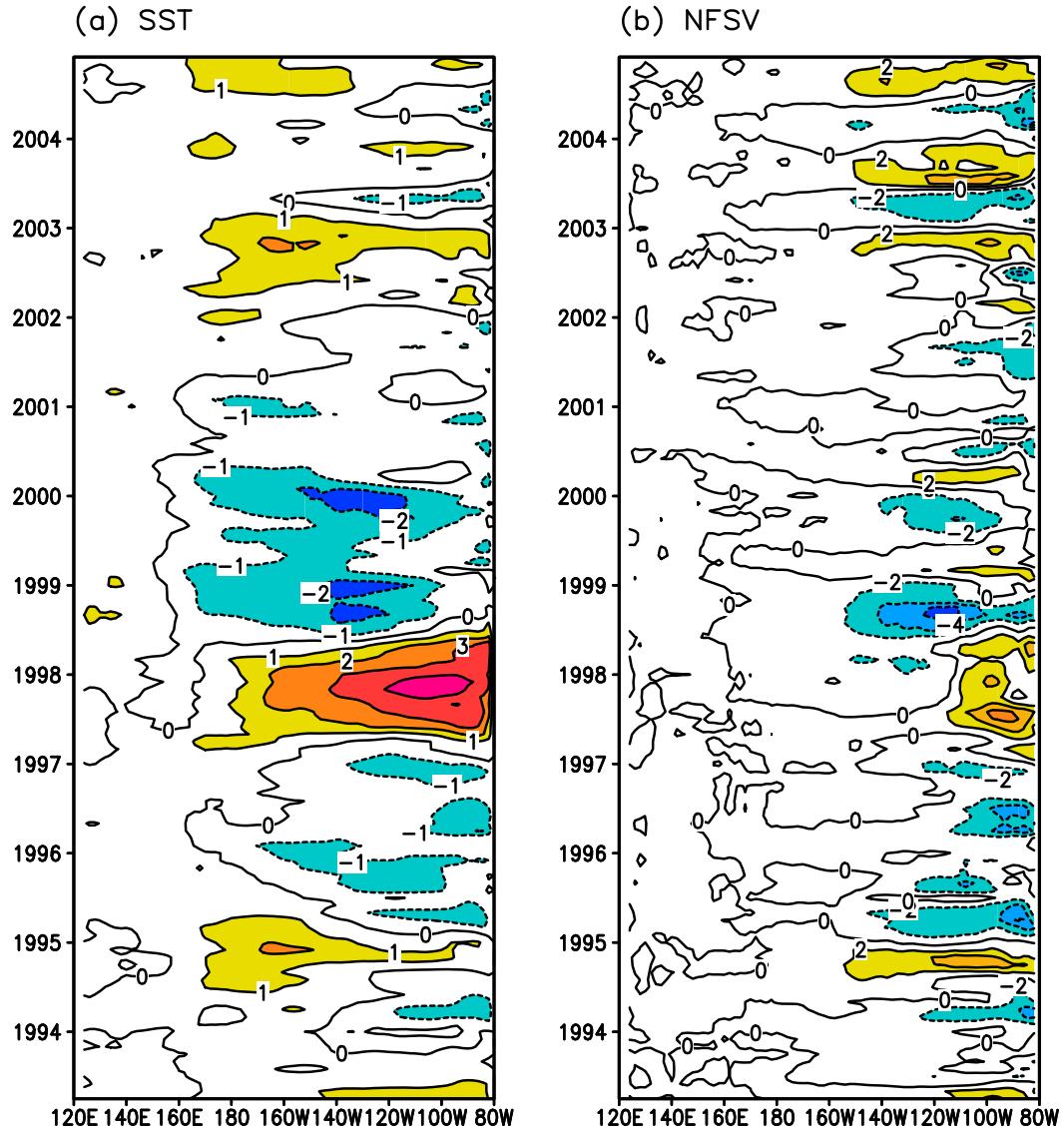


FIG. 3. Time-dependent section along the equator of the (a) observed SST anomalies ($^{\circ}\text{C}$) and (b) NFSV-tendency errors (10^{-6}C s^{-1}). The NFSV-tendency errors show an ENSO-like oscillation for the SST component.

tropical Pacific when the observed SSTs are warming (cooling). This relation encourages us to construct an equation that addresses the dependence of the NFSV-tendency errors on the observed SST anomalies. Then, we can use this equation to forecast future NFSV-tendency errors according to known observations.

The SVD approach is used to clarify the flow-dependent relation between the NSFV-type tendency errors and the observed SSTs. We can build up a lead-lag relationship between the observed SST anomalies and the NFSV-tendency errors by the SVD. That is, we develop an equation to describe the sensitivity of the lagged NFSV-tendency errors to the lead observations,

which allows for the estimation of the NFSV-tendency errors during the forecast period using the current observations. To achieve this, the covariance \mathbf{C} in the SVD analysis is calculated from the matrix including the observed SST anomalies and the lag NFSV-tendency errors [see Eq. (8)]:

$$\mathbf{C}_l(i, j) = \frac{1}{N-1} \sum_{t=t_1}^{t_2} \text{SST}(t, i, j) \text{NFSV}(t+l, i, j), \quad (8)$$

where (i, j) represents the model grid, l represents the lagged months of the NFSV-type tendency errors relative to the SST anomalies, and N represents the time

length for computing the covariance \mathbf{C} . Then, the lead–lag relation is obtained by performing the SVD technique on the covariance matrix \mathbf{C} (Bretherton et al. 1992), which can be written as

$$\mathbf{TE}_l = \alpha \mathbf{F}_l(\text{SST}), \quad (9)$$

where \mathbf{F}_l describes the relationship between the lagged NFSV and the l month lead SST anomaly and \mathbf{TE} represents the NFSV-tendency error estimated by Eq. (9) with the l month lead observed by the SST anomaly field. Such a lead–lag relationship provides the possibility of estimating NFSV-tendency tendency errors in advance. That is, when the prediction is initialized at one month, the future NFSV-tendency errors can be forecasted through this lead–lag relation by inputting the known initial observations. Thus, an NFSV-tendency error forecast model can be constructed.

The NFSV-tendency error forecast model, as shown in the last paragraph, can be constructed using the known SST observations and predetermined NFSVs. To examine its validity, we take the period from 1960 to 1996 as the training period to determine the NFSV-tendency error forecast model and the period 1997–2017 as the cross-validation period. In the training period, 10 leading SVD modes are used to construct \mathbf{F}_l in Eq. (9), while the remaining SVD modes are discarded due to their nearly stochastic and unrelated properties. Since the truncation of the SVD modes may reduce the variance in the NFSV-tendency errors, a scalar coefficient α is introduced to scale the strength of the tendency errors. In the present study, the coefficient α is taken as 0.6 and verified to be more applicable than other values in predicting ENSO. The experiments for determining α and SVD modes are referred to in the appendix.

Now, we use the NFSV-tendency error forecast model to predict the NFSV-type tendency errors during the cross-validation period. As mentioned above, tendency errors are significant along the equator and northern boundary of the model. Thus, to demonstrate the effectiveness of the NFSV-tendency error forecast model, we present the predicted NFSV-tendency errors along the equator (Fig. 4a) and northern boundary of the model (Fig. 4b). In particular, Fig. 4 displays the predicted NFSV-tendency errors at the 6-month lead time during the period 1990–2001, which includes 1990–96 as part of the training period and 1997–2001 as part of the cross-validation period. It is shown that the predicted NFSV-type tendency errors during the training period 1990–96 are almost identical to the NFSV-type tendency errors predetermined by the NFSV assimilation along

both the equator and northern boundary of the model, suggesting that the lead–lag relationship between the observed SST and NFSV-type tendency errors is well captured by the function \mathbf{F}_l in Eq. (9). For the cross-validation period 1996–2001, it is also found that the predicted NFSV-tendency errors are well approximated compared to the predetermined NFSV-tendency errors, including the amplitudes and locations. In particular, the phase change in the NFSV-tendency errors fits the predetermined errors very well. All of these results suggest that the constructed NFSV-SST relation [i.e., Eq. (9)] has the ability to predict future NFSV-tendency errors from the current SST field. Therefore, it can be confidently said that the NFSV-tendency error can reduce the model uncertainties and make a better prediction when its related tendency error forecast model is coupled with the ICM.

c. The ENSO forecast system of the NFSV assimilation

We combine the original ICM with its initial assimilation and the NFSV-tendency error forecast model determined by the NFSV-assimilation approach and finally formulate a new ENSO forecast system by superimposing the predicted NFSV-tendency error on the SST tendency of the ICM. Figure 5 shows a schematic diagram of the new ENSO forecast system (hereafter NFSV-ICM). The initialization scheme in the NFSV-ICM is the same as that in ICM (section 3). Specifically, we run the oceanic component of the ICM forced by the reconstructed wind before the prediction begins, nudge the observed SST to the model, and then finally obtain the initial states of the predictions. Using these initial states, we integrate the NFSV-ICM and make the ENSO predictions. The predicted NFSV-tendency errors are superimposed on the total tendency of the SST equation in the ICM with the initial value assimilation, which then perturbs the predicted SST anomalies at each time step of the model integral. As such, the predicted NFSV-tendency errors also suppress the effects of the initial errors on the prediction uncertainties. Theoretically, the NFSV-tendency errors consider the interaction between the model errors and initial errors.

5. The performance of the NFSV-ICM

The predictions of SST anomalies (SSTAs) associated with ENSO are made with 1-, 2-, 3-, . . . , 12-month lead times for the period 1960–2016. For comparison, the prediction results from both the original ICM and NFSV-ICM are output, and their skill scores are

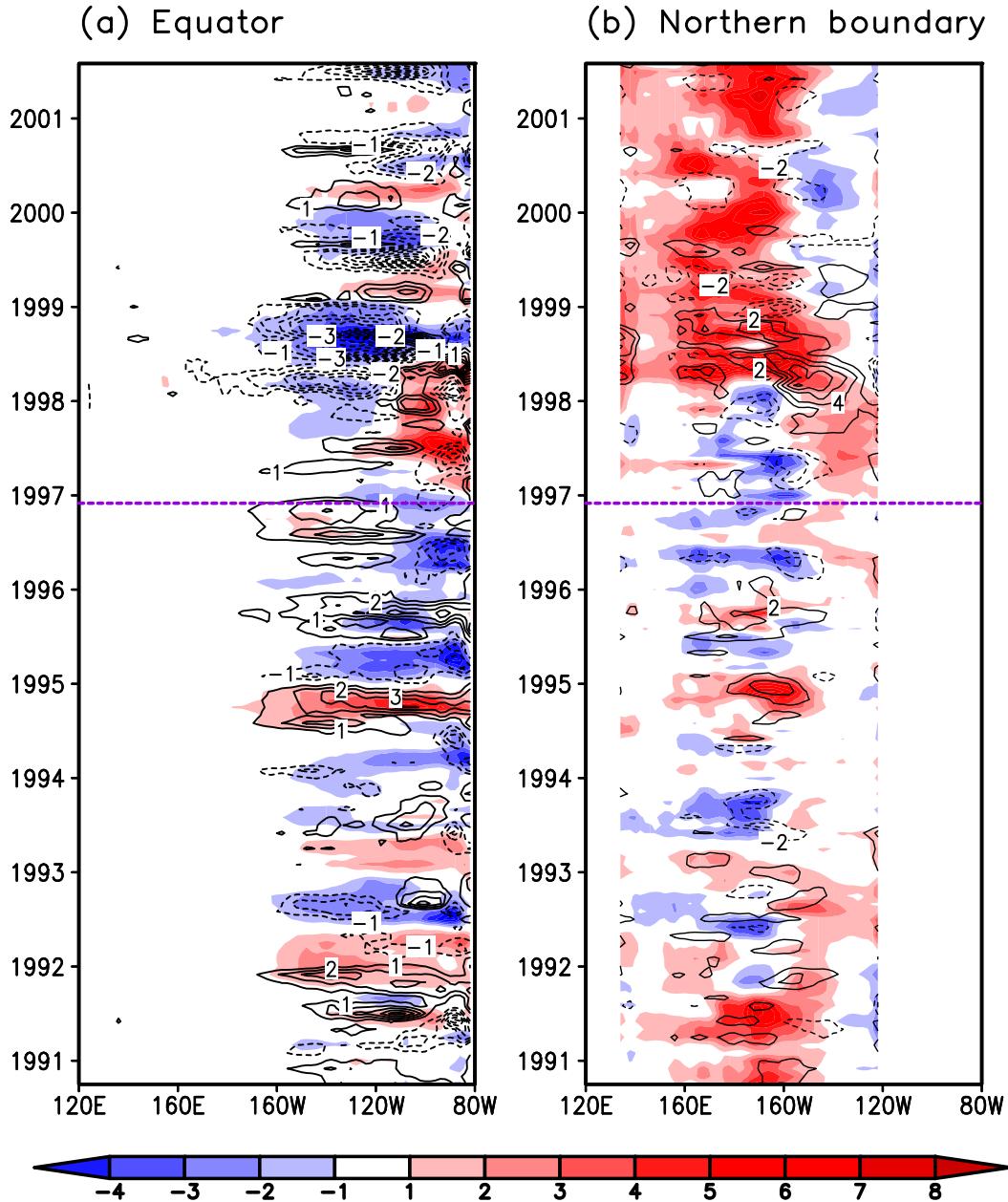


FIG. 4. Time-dependent NFSV-tendency errors determined by the NFSV-related assimilation (shaded) and those estimated from the NFSV-tendency error forecast model with a 6-month lead time (contours) along (a) the equator and (b) the northern boundary of the model. The purple dashed lines distinguish the training period (1991–96) from the cross-validation period (1997–2001). The shaded regions either during the training period or cross-validation period coincide well with those marked by contours, which indicates that the NFSV-tendency error forecast model is valid in estimating future NFSV-tendency errors. The contour interval is $2 \times 10^{-6} \text{ } ^\circ\text{C s}^{-1}$.

evaluated against the observed monthly mean SST anomalies. In addition, we note that the NFSV-tendency error forecast model embedded in the NFSV-ICM is obtained by the NFSV-related assimilation for the observed SST during the training period of 1960–96. Hence, the NFSV-ICM should be first validated by

predicting ENSO during the training period of 1960–96 and then try to show whether the feedback among state variables is reasonable. Then, the model should be tested by forecasting ENSO during the cross-validation period of 1997–2016 and examining its accountability in more realistic predictions.

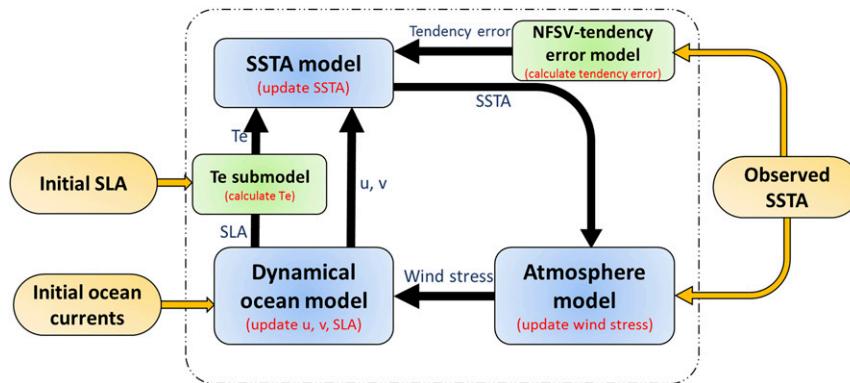


FIG. 5. A schematic diagram showing the new ENSO forecast model NFSV-ICM. The NFSV-ICM is obtained by the ICM coupled with an NFSV-tendency error forecast model. The ICM is a tropical Pacific coupled model and consists of a dynamical ocean model, a statistical atmospheric model, and an SST anomaly model embedded with a T_e model describing the temperature of the subsurface water entrained into the mixed layer.

a. Predictions of ENSO during the training period of 1960–96

To examine the prediction skill of SSTAs, two frequently used measurements are selected. One is the root-mean-square error (RMSE), which represents the deviation in the predictions from the observations, and the other is the anomaly correlation coefficient (ACC) that measures how parallel predictions and observations reach. Figures 6 and 7 illustrate the spatial distributions of the ACC and RMSE for predicted SST anomalies against those for the observed SSTAs in the tropical Pacific, respectively. It is shown that, for all lead times, the NFSV-ICM shows a much higher prediction skill (either in the ACC or RMSE) in the central tropical Pacific than the ICM. The NFSV-ICM also demonstrates skillful predictions of SSTAs in the far western tropical Pacific, while the ICM fails to do so. In the meridional direction, the skillful predictions generated by the NFSV-ICM cover almost the whole tropical Pacific, especially for short lead times. However, the ICM only shows skill in the region between 10°N and 10°S ; furthermore, with increasing lead times, the skills drop quickly in the eastern tropical Pacific and near the equator. When the lead time is up to 9 months, the ICM predictions lose useful skill in almost the whole tropical Pacific and even show a negative ACC off the equator (Fig. 6c1). Comparatively, the ACC of the NFSV-ICM predictions is still greater than 0.6 in the central tropical Pacific at this lead time (Fig. 6c2).

From the perspective of the RMSE, it can be seen that the prediction errors are much larger along the equator when the predictions are made using the ICM (left panels of Fig. 7). At short lead times, the prediction errors are mainly concentrated near the coast of Peru and, with

increasing lead times, such prediction errors in the eastern Pacific become large and extend toward the west. In addition, large prediction errors are also found near the model boundaries; these errors propagate to the central equatorial Pacific with time and increase to 1.2°C at the 12-month lead time. However, when using the NFSV-ICM, all of these prediction errors are significantly decreased, especially over regions near the equator. It is obvious that the NFSV-ICM possesses smaller prediction errors than the ICM in predicting SSTAs in the tropical Pacific.

The anomaly correlations between the observed and predicted SSTAs in the Niño-3.4 area are plotted in Fig. 8a1 as a function of the lead time. The prediction skill of the ICM declines faster than that of the NFSV-ICM, and the differences between the ICM and NFSV-ICM in the ACCs gradually increase from 0.1 at the 3-month lead time to 0.3 at the 12-month lead time, with the NFSV-ICM showing a much larger ACC. Furthermore, it is at all lead times that the skill of NFSV-ICM defeats that of ICM and persistence prediction. If the skillful predictions are regarded as the ACC being larger than 0.6, the predictable time length for the Niño-3.4 index can be increased from 6 months using the ICM to 12 months using the NFSV-ICM. A similar improvement is also shown in the RMSE (see Fig. 8b1), especially for predictions with a 12-month lead time. These results indicate that the NFSV-ICM is more skillful than the ICM in predicting Niño-3.4 index.

To reveal the season-dependent prediction skill, the anomaly correlations for the Niño-3.4 index are also calculated as a function of both start months and lead times. The results for the ICM and NFSV-ICM are shown in Figs. 9a1 and b1, respectively. It is clearly

Correlations between observed and predicted SSTA

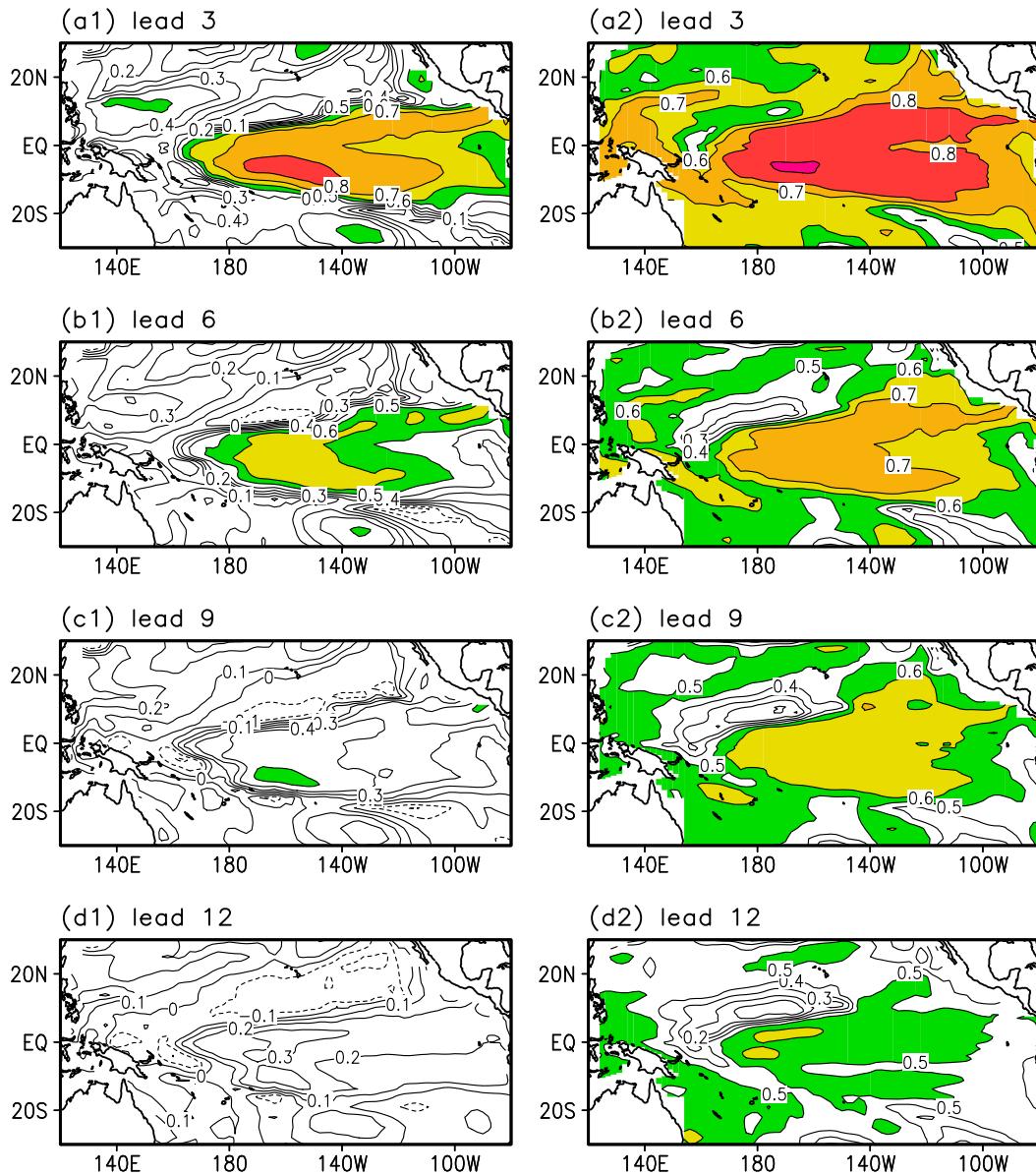


FIG. 6. Horizontal distributions of the ACC between the observed and predicted SSTA at (a) 3-, (b) 6-, (c) 9-, and (d) 12-month lead times. (left) The ICM and (right) the NFSV-ICM during the training period of 1960–97. The contour interval is 0.1.

shown that the prediction skill is significantly dependent on the season. Specifically, both the ICM and NFSV-ICM predictions show high skill in boreal winter and low skill in spring. However, in either winter or spring, the former predictions always present lower skills than the latter predictions in predicting the Niño-3.4 index. The low prediction skill in spring is generally referred to as the well-known SPB phenomenon. The SPB is defined as a rapid decline in the anomaly correlation coefficient when the prediction is made across boreal spring. From

the above results, it can be deduced that although both the ICM and NFSV-ICM suffer from the SPB phenomenon, the NFSV-ICM notably reduces the effect of the SPB in predicting tropical Pacific SSTA and gives a more accurate prediction than the ICM.

In summary, we demonstrate here that the NFSV-ICM has greater abilities than the ICM in predicting SSTA anomalies associated with ENSO during the training period. In particular, the SPB phenomenon is weakened in the NFSV-ICM, and the related prediction skills are

Root mean square error of predicted SSTA

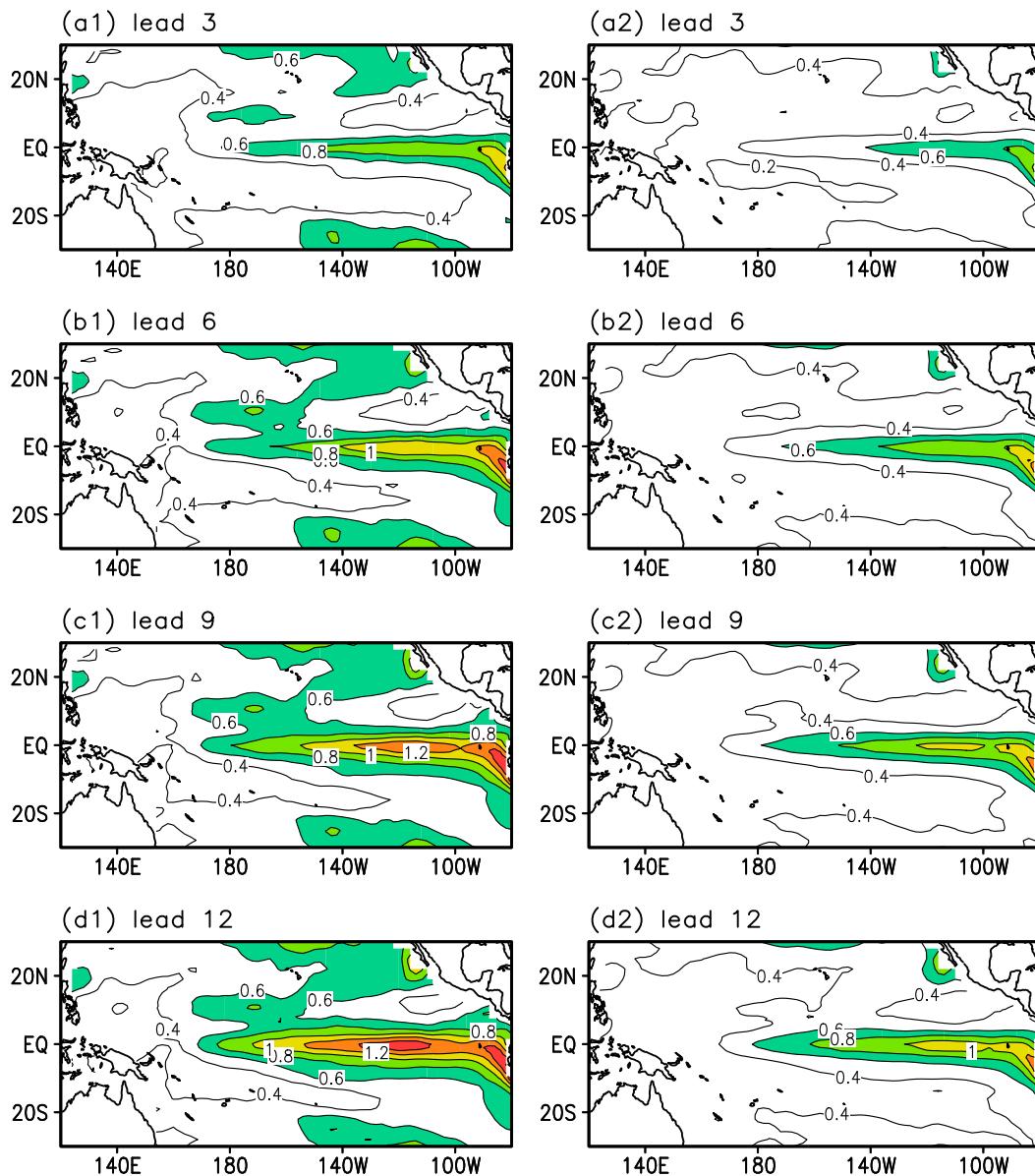


FIG. 7. As in Fig. 6, but for the RMSE. The contour interval is 0.2°C .

highly improved in the central and western tropical Pacific. These results indicate that the NFSV-ICM describes the dynamical and physical feedback associated with ENSO.

b. Predictions of ENSO during the cross-validation period of 1997–2016

In the last section, we have validated the NFSV-ICM in improving ENSO prediction during the training period from 1960 to 1996. The NFSV-tendency error forecast model is obtained by assimilating the observation

information during this period. It is therefore understandable that the NFSV-ICM provides a significant improvement in ENSO predictions during 1960–96. In fact, a cross-validation experiment is much more realistic for examining the validity of the NFSV-ICM in predicting ENSO. The cross-validation experiment here refers to the fact that the NFSV-tendency error forecast model obtained during 1960–96 is inserted into the ICM to predict the SSTAs during the period of 1997–2016. That means we use the same NFSV-ICM in the last section to examine the improvement of the

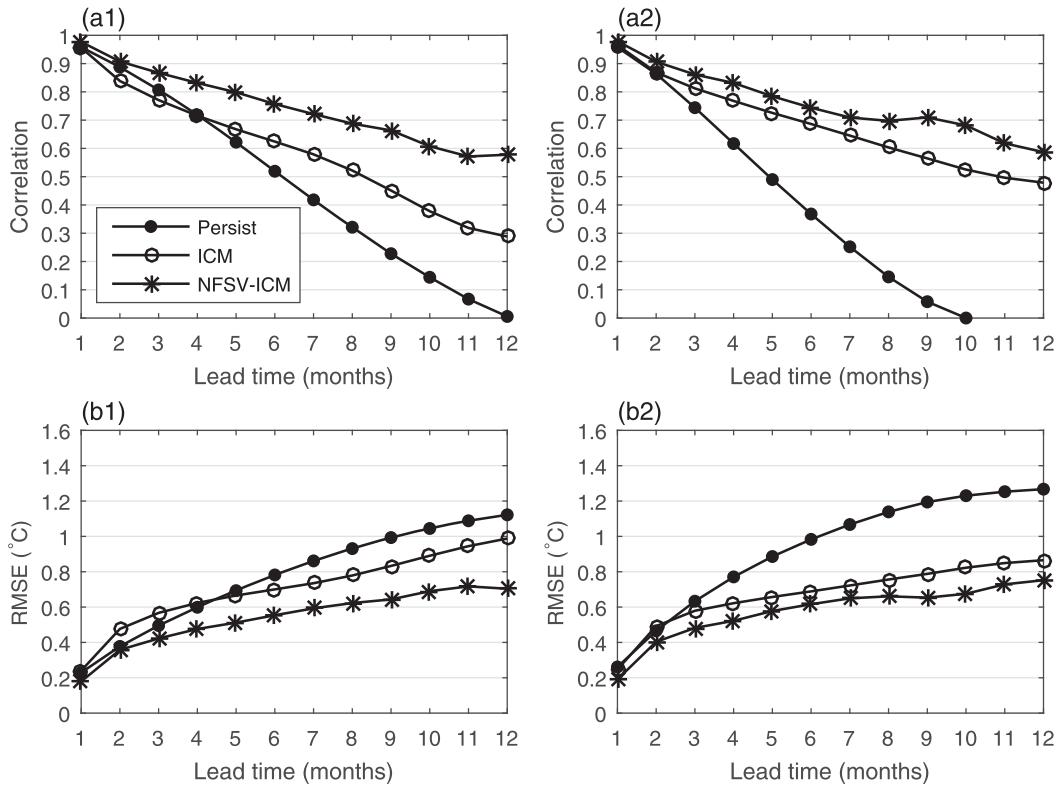


FIG. 8. (a) ACC and (b) RMSE of the predicted Niño-3.4 SST anomalies with respect to the observed SST anomalies as a function of lead time. (left) The training period (1960–96) and (right) the cross-validation period (1997–2016). The lines with open (filled) circles are for the ICM (persistence), and those with asterisks are for the NFSV-ICM.

prediction skill of the NFSV-ICM against the ICM in predicting ENSO.

The ACC and RMSE for the prediction experiments are presented in Figs. 10 and 11, respectively. It is shown that the anomaly correlation obtained from the NFSV-ICM is obviously larger than that from the ICM at lead times of 1, 2, 3, . . . , 12 months, and the RMSE is much smaller than that of the ICM. This indicates that the NFSV-ICM outperforms the ICM in predicting SST anomalies. In particular, the NFSV-ICM still provides useful skill at a 12-month lead time, with an ACC larger than 0.6 (Fig. 8a2). Although the SPB phenomenon also occurred in the cross-validation period of 1997–2016, it was weaker in the NFSV-ICM than the ICM (Figs. 9a2 and 9b2). Similar to the training period, the prediction skills for SSTAs in the central and western tropical Pacific are also significantly improved in the NFSV-ICM compared with the original ICM (Fig. 10). These results show that the NFSV-ICM is valid for reaching high prediction skills of SSTAs during the cross-validation period. Therefore, the NFSV-ICM can be a useful forecast system for realistic ENSO events.

c. Prediction skills for two types of ENSO

The fact that the NFSV-ICM shows high performance in central Pacific implies that the new model can well predict the central Pacific warming events known as CP El Niño. Such warming events that occur frequently in recent decades are found to show different climate effect compared with the traditional El Niño (denoted as EP El Niño) (Ashok et al. 2007). So, predicting the space structures, especially distinguishing two kinds of El Niño events are also important.

Predictions for two kinds of El Niño events are shown in Figs. 12 and 13, respectively. Although the ICM has the ability to predict the amplitude of EP El Niño, the ICM tends to predict a cooler-than-normal SST anomaly in subtropical Pacific. By contrast, the NFSV-ICM cannot only predict the amplitude but also capture the space structures of the EP El Niño events. Furthermore, an evident improvement is found in CP El Niño prediction using the NFSV-ICM. The ICM tends to lose skill in predicting CP El Niño at 6-month lead time and usually predicts a cooling event. While the NFSV-ICM still has skills in predicting the spatial distributions of CP

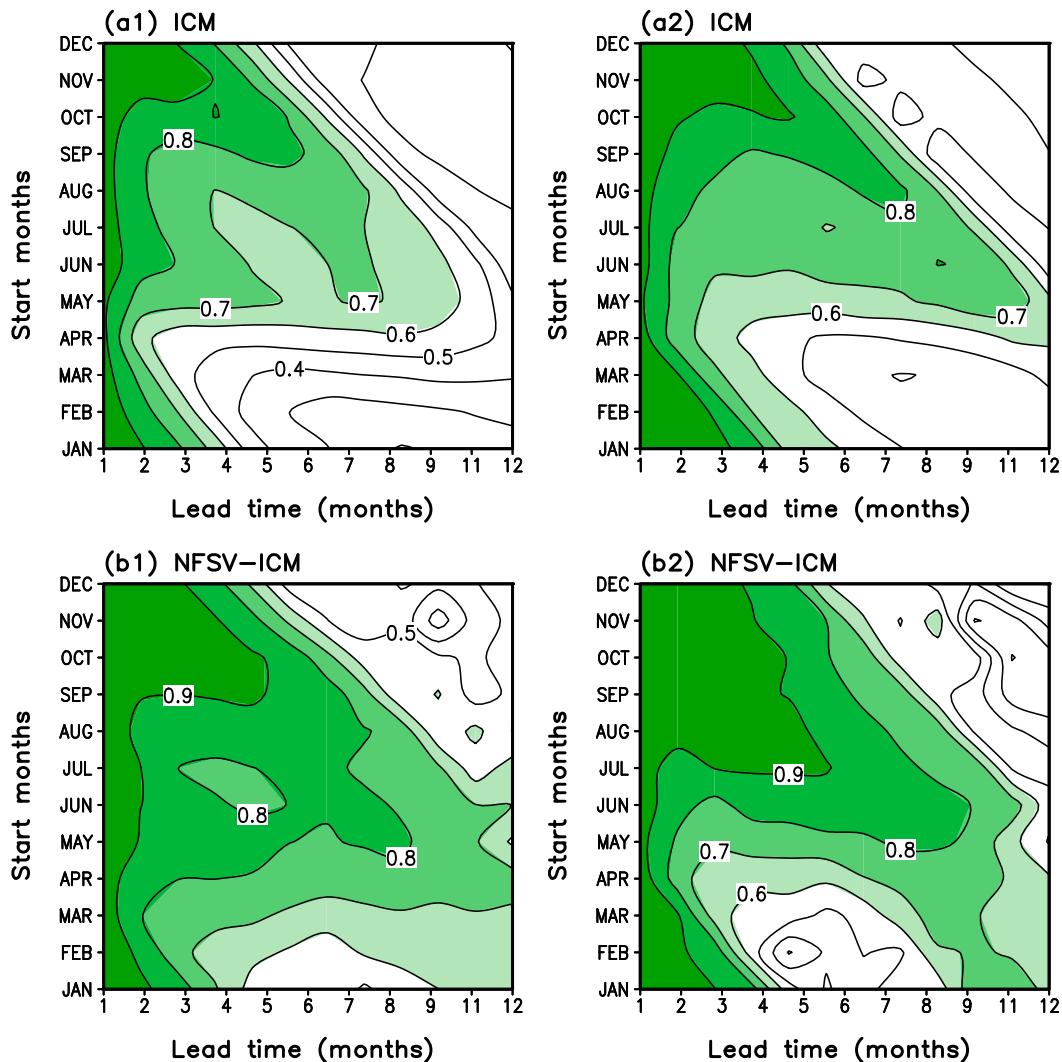


FIG. 9. ACC of the predicted Niño-3.4 SST anomalies as a function of start month and lead time. The predictions are made by the (a) ICM and (b) NFSV-ICM during the (left) training period (1960–96) and (right) validation period (1997–2016). The contour interval is 0.1.

El Niño events. From the above, it is indicated that the ICM equipped with the NFSV-assimilation is likely to have advantage in discerning and predicting different types of El Niño events.

6. Conclusions and discussion

The predictions of ENSO events are generally influenced by both initial errors and model uncertainties. In particular, an ICM usually neglects or simplifies some physical processes, which induces large model errors and influences the accuracy of the predictions of ENSO (Qi et al. 2017). In the present study, we focus on the model errors and develop a new ENSO forecast system (NFSV-ICM) consisting of an ICM and an NFSV-tendency error

forecast model that is used to estimate the combined effect of the model errors. The prediction experiments are performed for tropical SSTAs during both the training period (1960–1996) and the cross-validation period (1997–2016). The results commonly show that the NFSV-ICM tends to possess a much higher forecast skill compared with the original ICM. In particular, a considerable improvement in the forecast skill is reflected in the central and western tropical Pacific. Furthermore, the well-known SPB phenomenon is also obviously weakened in the NFSV-ICM. The NFSV-ICM shows useful skill in ENSO forecasting and can be a promising ENSO forecast system.

The high skill of the NFSV-ICM in predicting ENSO is mainly due to the embedded NFSV-tendency error

Correlations between observed and predicted SSTA

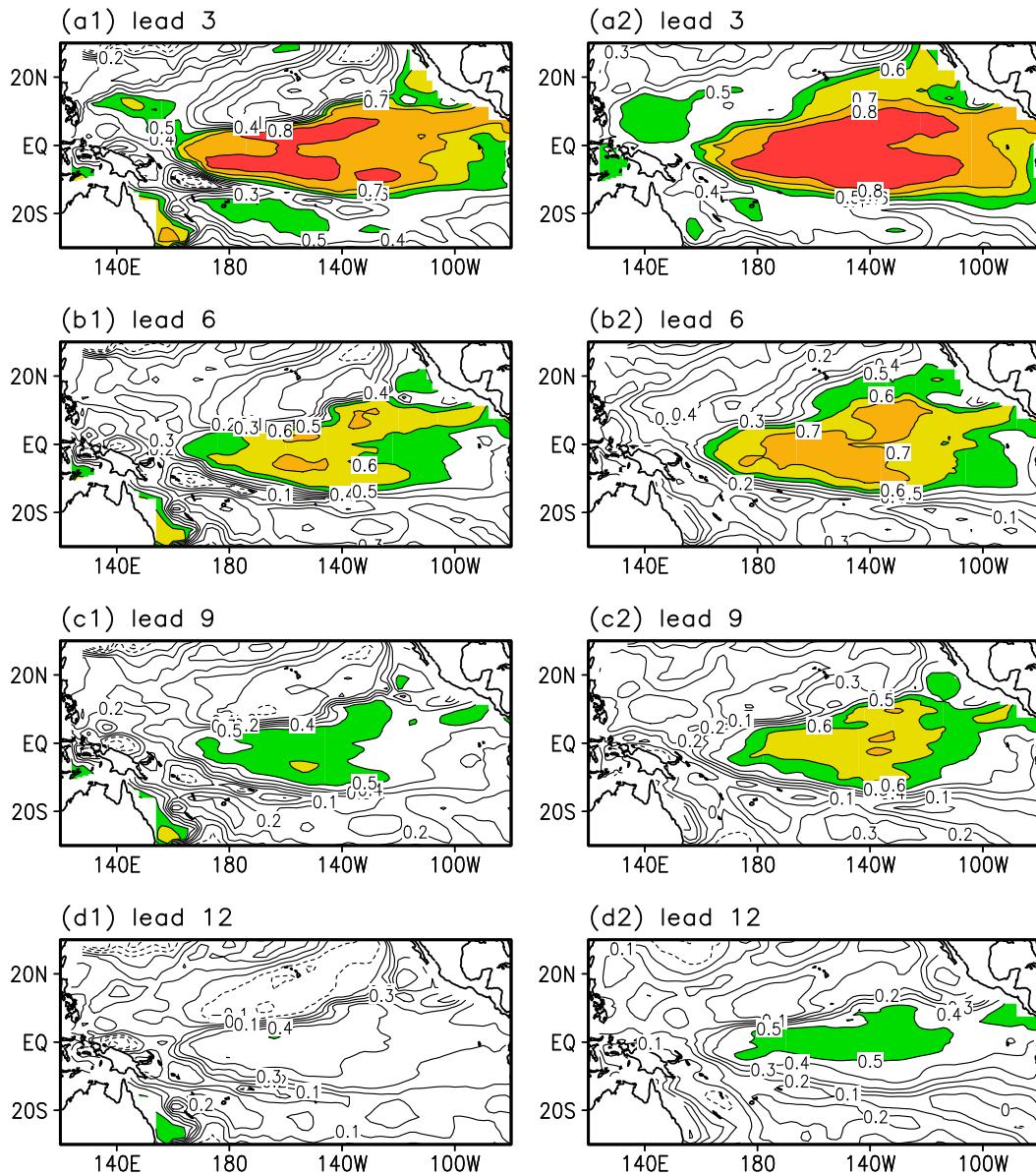


FIG. 10. As in Fig. 6, but for the cross-validation period (1997–2016).

forecast model. Since the model errors are from different model error sources and their effects are mixed in the prediction uncertainties (Jin et al. 2008; Zheng and Zhu 2016), it is difficult to distinguish them and study them separately. The NFSV-tendency error forecast model considers the effect of model errors from a macro perspective and proposes the approach of describing the combined effect of different model errors by the NFSV-tendency perturbation. The NFSV-tendency perturbation is superimposed on each time step of the model integrals and therefore

also suppresses the effect of the initial errors. The NFSV-tendency perturbation tries to reach the optimal tendency error by using an NFSV-related assimilation and is sufficient for estimating the interaction between model errors and initial errors, which therefore corrects the ICM forecasts to the greatest extent and makes the NFSV-ICM obtain high skill in predicting ENSO. The NFSV-ICM can consider not only the effect of model errors but also the effect of the initial errors. This is the competing aspect of the NFSV-ICM compared with the traditional initial value

Root mean square error of predicted SSTA

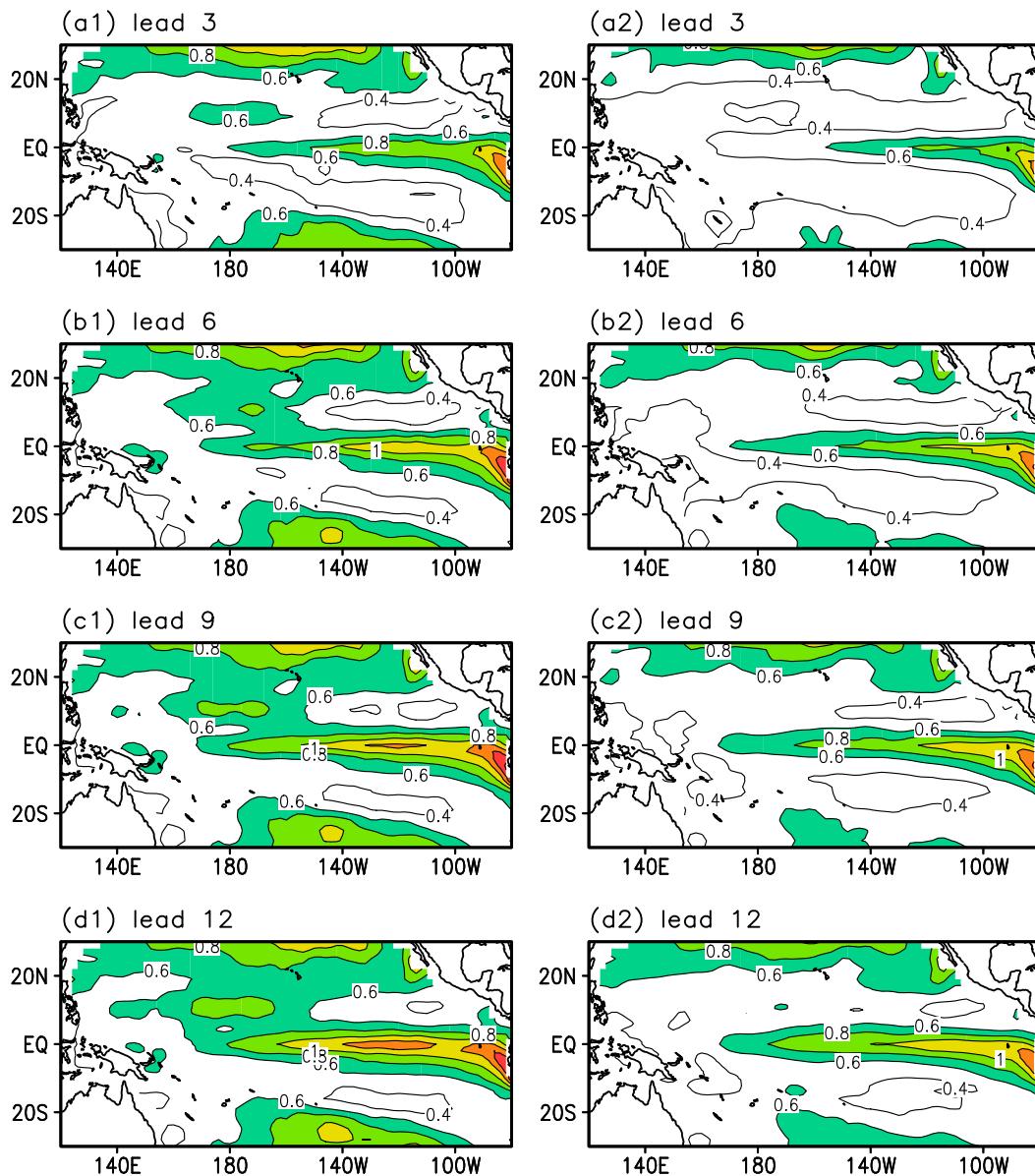


FIG. 11. As in Fig. 7, but for the cross-validation period (1997–2016).

assimilation. Because the NFSV-tendency error forecast model is constructed based on the SVD relationship between SST and NFSV. An undeniable fact is that the trained SST–NFSV relation is dependent on the historical data and the uncertainties of the SVD analyses. Therefore, sensitivity experiments regarding to the NFSV-tendency error model are implemented to explore the prediction skills influenced by the uncertainties of the SST–NFSV relation (see the appendix). The fact that the SST–NFSV relation is essentially nonlinear implies the uncertainties of the relation

constructed by the linear SVD approach. More intelligent and advanced methods (e.g., machine learning) are expected to be adopted to make the SST–NFSV relation more robust (Reichstein et al. 2019).

It is pointed out that the ICM-based ensemble prediction system (EPS-ICM; see Zheng et al. 2009b) shows skill almost equivalent to that in the NFSV-ICM for SST predictions. Note that the EPS-ICM is constructed not only using ensemble Kalman filter (EnKF) data assimilation with SST fields to generate the initial ensemble conditions but also a model-error model to characterize

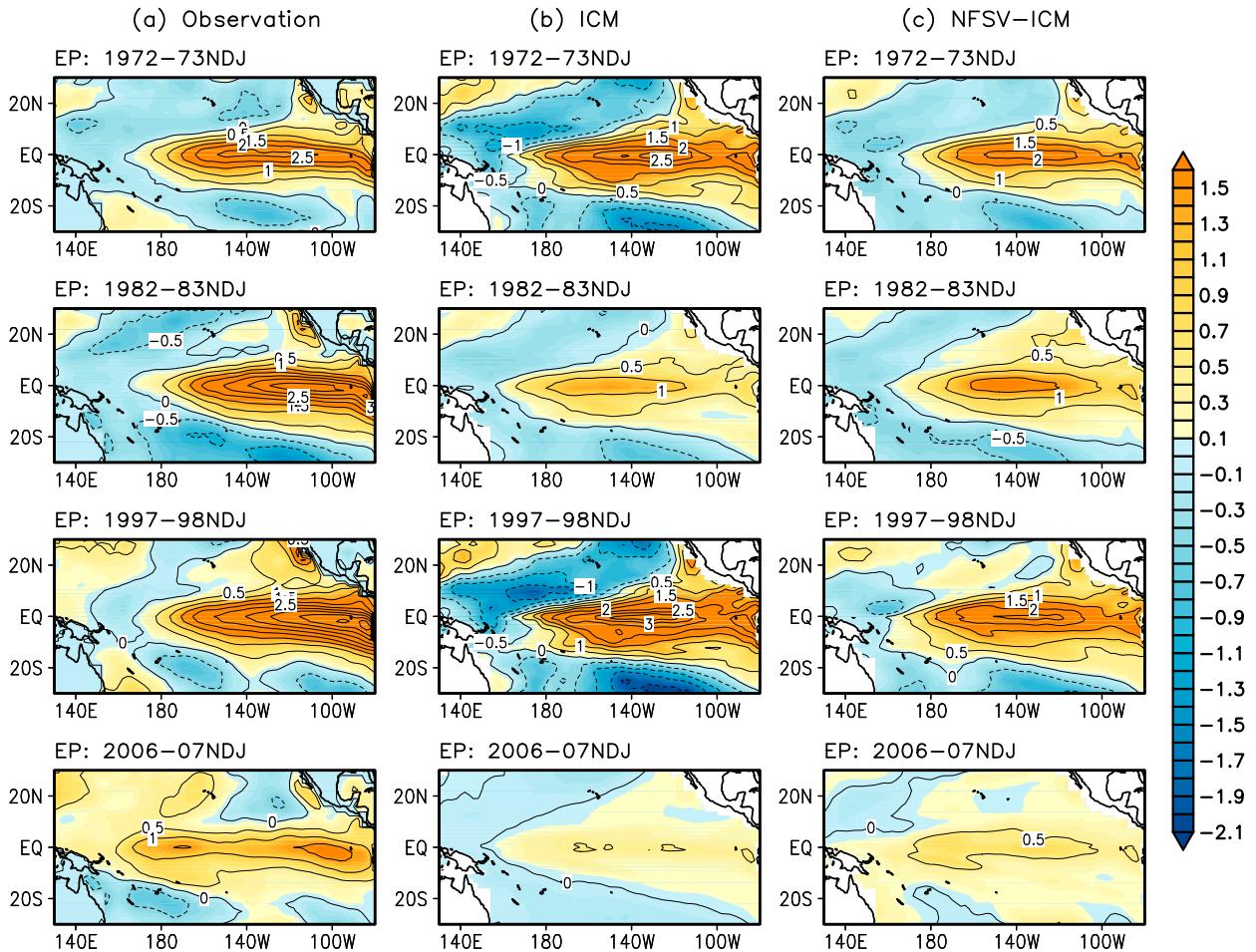


FIG. 12. The horizontal distributions of SST anomalies in (a) observations, (b) ICM prediction, and (c) NFSV-ICM prediction at 6-month lead time predictions for the mature phase of EP-type El Niño events.

the model uncertainties. The model-error model is a zero-mean first-order Markov stochastic model that is developed by analyzing historical model errors. That is, the EPS-ICM is an ensemble forecast system involving initial and model errors and using a relatively advance data assimilation. However, the NFSV-ICM is only involved with a deterministic prediction system, where the SST observations are only used to initialize the model with simple data assimilation (i.e., the nudging method). Obviously, the NFSV-tendency error forecast model is the one that plays an important role in improving the level of predictions generated by the NFSV-ICM. The NFSV-tendency error forecast model, due to its optimal NFSV-tendency perturbation, possesses more possibilities in correcting the model and greatly improving the prediction skill. On the other hand, it is worth mentioning that [Zheng and Zhu \(2016\)](#) considered the coupling of atmosphere and ocean in initialization of the EPS-ICM and used an advanced EnKF assimilation

approach to initialize the model, finally achieving much higher forecast skill of ENSO. This encourages us to equip the NFSV-ICM with EnKF and consider the air-sea coupling in initialization, and then enhance greatly the forecast skill of the NFSV-ICM with respect to the types of ENSO. Exactly, such an idea is under investigation.

In addition, after the 1990s, a new flavor of El Niño, with its warm center in the central tropical Pacific (known as CP El Niño) in comparison with the traditional El Niño (EP El Niño), which has a warm center in the eastern tropical Pacific, occurred frequently, giving rise to additional model errors and proposing new challenges to the simulations and predictions of ENSO ([Kim et al. 2012](#); [Tian and Duan 2016](#); [Duan et al. 2018](#)). Although some models, including the ICM used in the present study, are equipped with advanced assimilation techniques to optimize the initial fields, they still showed low forecast skill for CP ENSO due to model error

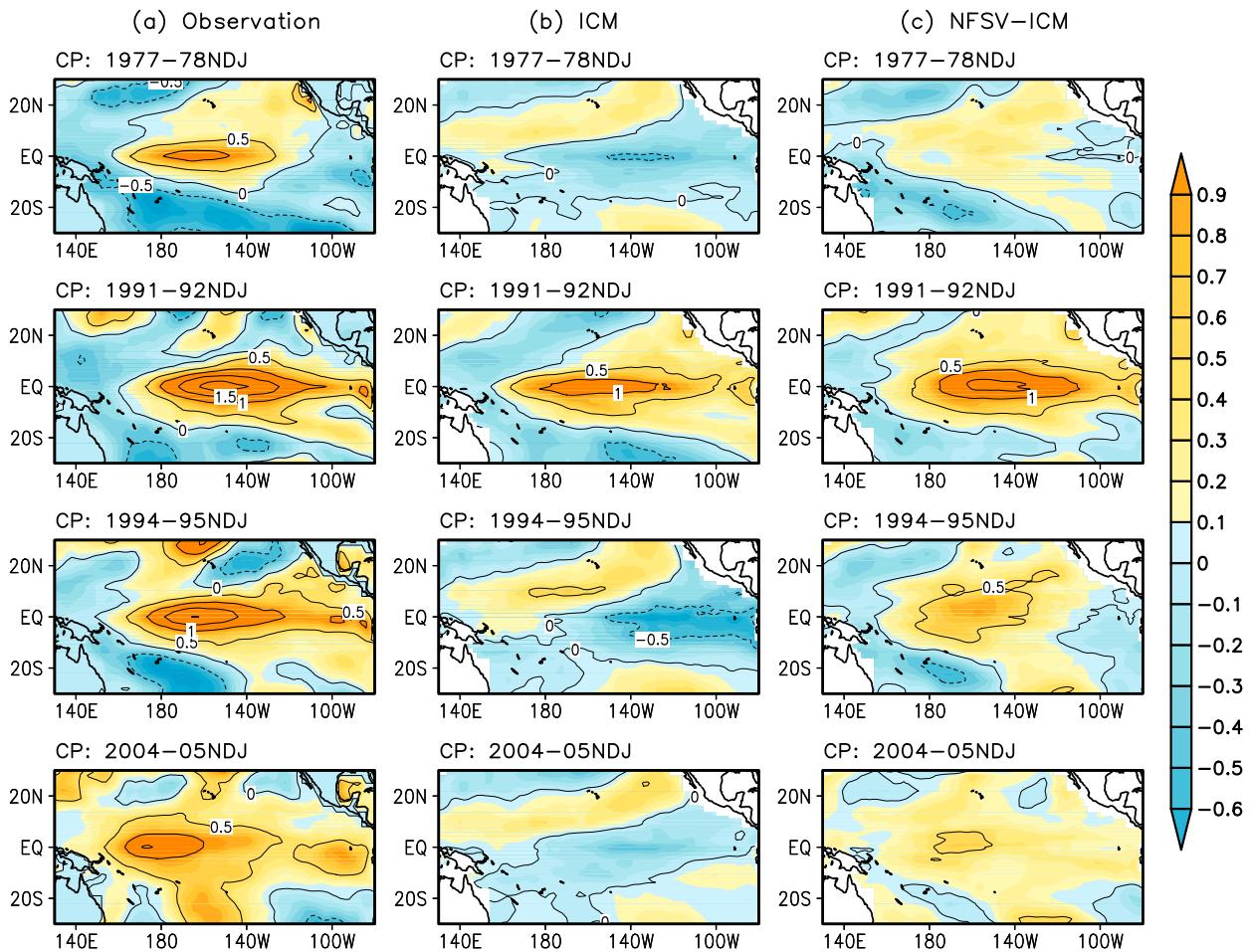


FIG. 13. As in Fig. 12, but for the mature phase of CP-type El Niño events.

effects (Hendon et al. 2009; Duan et al. 2014). In the present study, we have shown that the NFSV-ICM shows great improvement in the prediction skills for SSTs in the central and western tropical Pacific. And we find that the ICM equipped with the NFSV assimilation greatly improve the capacity of discernments and predictions for two kinds of El Niño events. Thus, the NFSV-ICM may provide a promising way to study predictability in terms of the EP and CP El Niño events. In addition, we have not addressed why and how the NFSV-tendency errors work in the SST predictions, especially for the high prediction skill in the central Pacific. These issues will be focused on and addressed in the next paper.

Acknowledgments. The authors wish to appreciate anonymous reviewers for their very insightful comments and suggestions. This work was supported by the National Natural Science Foundation of China (Grants 41525017 and 41690124).

APPENDIX

Sensitivity of NFSV-ICM to the Error Forecast Model

The success of the NFSV-ICM is sensitive to the constructed NFSV-tendency error model. In addition to the training periods, the error model is also dependent on the intensity of the error forcing and the statistical method [Eq. (9)]. As mentioned above, since only some leading SVD modes are retained, part of the information and variance is lost. Thus, it is very important to obtain a reasonable α that represents the strength of the tendency error and limited SVD modes to facilitate the error model. On the one hand, a high α may exorbitantly estimate the model errors against the ENSO prediction, while a too low α is not enough to produce tendency errors that offset the model uncertainties in predictions. On the other hand, if too many SVD modes are retained in statistical relation, the noise that has nothing to do with

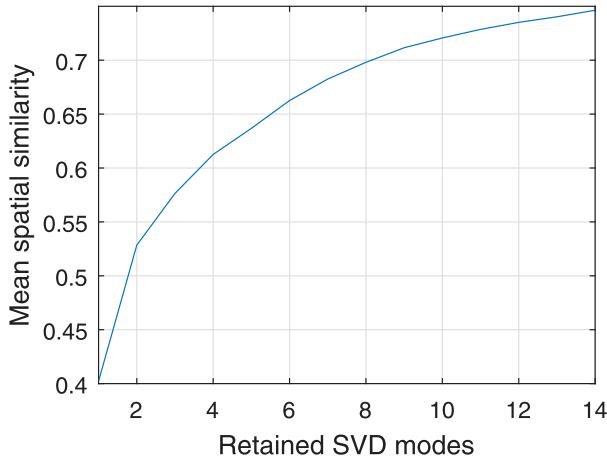


FIG. A1. The mean spatial similarity between NFSVs and statistical determined NFSVs as a function of retained SVD modes.

the SST–NFSV relation will be included. As a result, the constructed error model has low skill in estimating the tendency errors during the period that is not overlapping the training period. Thus, this section will further examine the sensitivity of the NFSV-ICM to the strength of the error forcing and the retained SVD modes.

The structure of the tendency errors that are estimated using the SST anomaly is highly dependent on how many

SVD modes are retained. To estimate the error model in terms of the spatial structure during the training period (i.e., 1960–97), a mean spatial similarity (MSS) between predicted tendency errors (NFSV^p) and NFSVs is used defined as $MSS = (1/N) \sum_{t=t_1}^N [NFSV(t)NFSV^p(t)] / [NFSV(t)||NFSV^p(t)]$, in which N is 37×12 (months). The result is presented in Fig. A1. It is clear that more SVD modes retained in constructing the NFSV-error model has a better ability to capture the patterns of the tendency errors. Particularly, the skill is improved gently when more than 6 leading modes are reserved. That is, the high-order modes have small role in the SST–NFSV relation. Besides, retaining high-order modes will contain the noise that harms the estimation of the tendency error using the SST information. In this sense, retaining reasonable SVD modes can filter out noise and extract the robust relation for SST–NFSV.

The truncation of the SVD modes necessarily changes the strength of the tendency errors estimated. The sensitivity to the SVD modes and the corresponding strength of the error forcing to the ICM is analyzed from the perspective of the predictions for the SST anomaly. The anomaly correlations between predicted and observed SST anomaly as a function of lead time during the period 1960–1996 are shown in Fig. A2. In short time

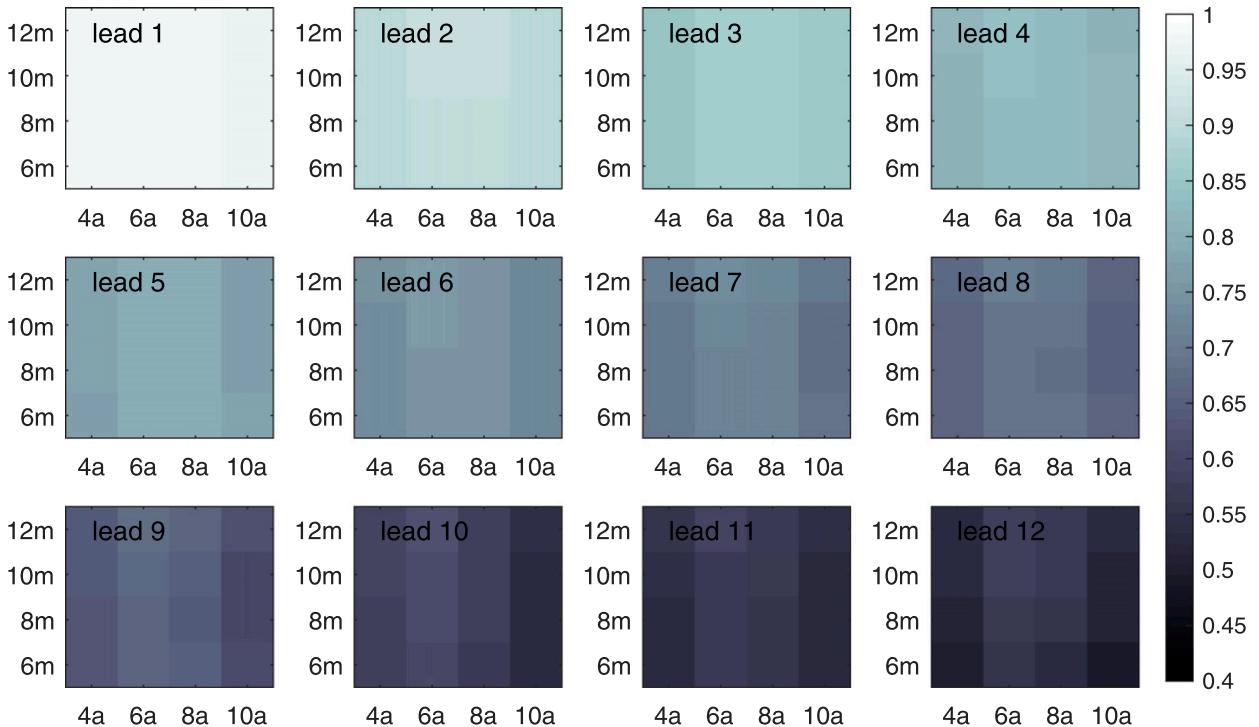


FIG. A2. Correlations of the Niño-3.4 SST anomaly during the period 1960–97, as a function of retained SVD modes and intensity of the error forcing (i.e., α). Each panel denotes the result at a certain lead time. The x coordinate is the value of the α (e.g., 6a denotes $\alpha = 0.6$). The y coordinate is the number of the leading SVD modes retained in the NFSV-tendency error model (e.g., 10m denotes 10 modes).

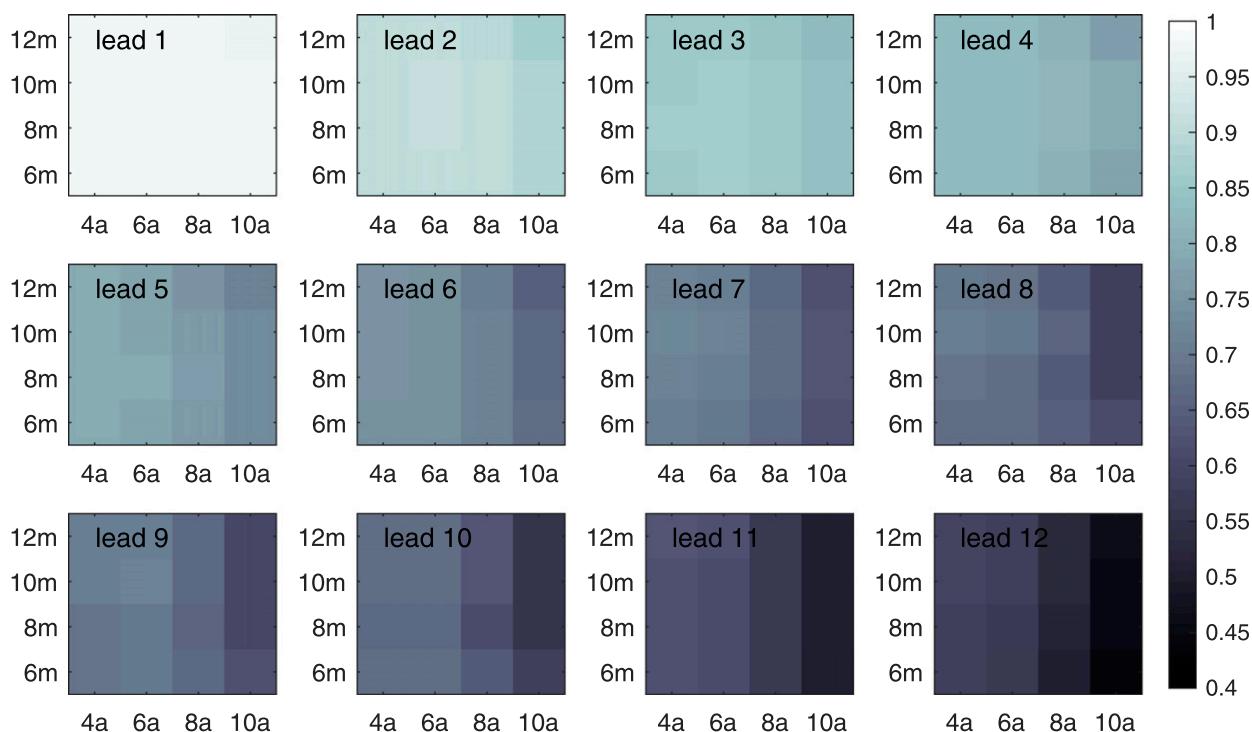


FIG. A3. As in Fig. A2, but for the period 1997–2016.

(e.g., 1-month lead time) predictions, the predicted results are insensitive to the SVD modes and the error strength since the model errors play small role in short time predictions. As the prediction length is increased, the performance of the NFSV-ICM is dependent on the SVD modes and α . As shown in Fig. A2, the NFSV-ICM possesses largest skill when $\alpha = 0.6$ or $\alpha = 0.8$, where the correlation is larger than 0.65 even predicting 12 months. In addition, the dependence of the model skill on SVD modes is prominent with lead time increased. Consistent with Fig. A1, the NFSV-ICM with higher modes retained tends to show higher skill in SST predictions. But the skill using 10 SVD modes are identical to that using 12 SVD modes.

As discussed above, a success of the NFSV-ICM cannot only improve the prediction during the training period but also can make sense in other periods without overlapping the training period. Therefore, the sensitivity of the model skill to the SVD modes and α during the period 1997–2016 is shown in Fig. A3. Similar to Fig. A2, limited SVD modes and a certain α allow the improvement of the ICM equipped with the NFSV-tendency error model. The skill of the NFSV-ICM reach the peak when the 10 SVD modes are retained and $\alpha = 0.6$ or $\alpha = 0.4$. Also, significant decrease of the prediction error is achieved (not shown).

From the above, the new ENSO forecast system has the largest skill in ENSO prediction when the NFSV-tendency error model is determined by 10 leading SVD modes and $\alpha = 0.6$.

REFERENCES

- Ashok, K., S. K. Behera, S. A. Rao, H. Y. Weng, and T. Yamagata, 2007: El Niño Modoki and its possible teleconnection. *J. Geophys. Res.*, **112**, C11007, <https://doi.org/10.1029/2006JC003798>.
- Barkmeijer, J., T. Iversen, and T. N. Palmer, 2003: Forcing singular vector and other sensitivity model structures. *Quart. J. Roy. Meteor. Soc.*, **129**, 2401–2423, <https://doi.org/10.1256/qj.02.126>.
- Barnett, T. P., M. Latif, N. Graham, M. Flugel, S. Pazan, and W. White, 1993: ENSO and ENSO-related predictability. Part I: Prediction of equatorial Pacific sea surface temperature with a hybrid coupled ocean–atmosphere model. *J. Climate*, **6**, 1545–1566, [https://doi.org/10.1175/1520-0442\(1993\)006<1545:EAERPP>2.0.CO;2](https://doi.org/10.1175/1520-0442(1993)006<1545:EAERPP>2.0.CO;2).
- Barnston, A. G., M. K. Tippett, M. L. L’Heureux, S. H. Li, and D. G. DeWitt, 2012: Skill of real-time seasonal ENSO model predictions during 2002–11: Is our capability increasing? *Bull. Amer. Meteor. Soc.*, **93**, 631–651, <https://doi.org/10.1175/BAMS-D-11-00111.1>.
- Behringer, D. W., M. Ji, and A. Leetmaa, 1998: An improved coupled model for ENSO prediction and implications for ocean initialization. Part I: The Ocean Data Assimilation System. *Mon. Wea. Rev.*, **126**, 1013–1021, [https://doi.org/10.1175/1520-0493\(1998\)126<1013:AICMFE>2.0.CO;2](https://doi.org/10.1175/1520-0493(1998)126<1013:AICMFE>2.0.CO;2).

- Bejarano, L., and F. F. Jin, 2008: Coexistence of equatorial coupled modes of ENSO. *J. Climate*, **21**, 3051–3067, <https://doi.org/10.1175/2007JCLI1679.1>.
- Bretherton, C. S., C. Smith, and J. M. Wallace, 1992: An intercomparison of methods for finding coupled patterns in climate data. *J. Climate*, **5**, 541–560, [https://doi.org/10.1175/1520-0442\(1992\)005<0541:AIOMFF>2.0.CO;2](https://doi.org/10.1175/1520-0442(1992)005<0541:AIOMFF>2.0.CO;2).
- Chen, D., and M. A. Cane, 2008: El Niño prediction and predictability. *J. Comput. Phys.*, **227**, 3625–3640, <https://doi.org/10.1016/j.jcp.2007.05.014>.
- , S. E. Zebiak, M. A. Cane, and A. J. Busalacchi, 1997: Initialization and predictability of a coupled ENSO forecast model. *Mon. Wea. Rev.*, **125**, 773–788, [https://doi.org/10.1175/1520-0493\(1997\)125<0773:IAPOAC>2.0.CO;2](https://doi.org/10.1175/1520-0493(1997)125<0773:IAPOAC>2.0.CO;2).
- , and Coauthors, 2015: Strong influence of westerly wind bursts on El Niño diversity. *Nat. Geosci.*, **8**, 339–345, <https://doi.org/10.1038/ngeo2399>.
- Duan, W. S., and F. F. Zhou, 2013: Non-linear forcing singular vector of a two-dimensional quasi-geostrophic model. *Tellus*, **65A**, 18452, <https://doi.org/10.3402/tellusa.v65i0.18452>.
- , and P. Zhao, 2015: Revealing the most disturbing tendency error of Zebiak–Cane model associated with El Niño predictions by nonlinear forcing singular vector approach. *Climate Dyn.*, **44**, 2351–2367, <https://doi.org/10.1007/s00382-014-2369-0>.
- , and J. Y. Hu, 2016: The initial errors that induce a significant “spring predictability barrier” for El Niño events and their implications for target observations: Results from an earth system model. *Climate Dyn.*, **46**, 3599–3615, <https://doi.org/10.1007/s00382-015-2789-5>.
- , and R. Feng, 2017: Reducing the prediction uncertainties of high-impact weather and climate events: An overview of studies at LASG. *J. Meteor. Res.*, **31**, 224–235, <https://doi.org/10.1007/s13351-016-6099-6>.
- , B. Tian, and H. Xu, 2014: Simulations of two types of El Niño events by an optimal forcing vector approach. *Climate Dyn.*, **43**, 1677–1692, <https://doi.org/10.1007/s00382-013-1993-4>.
- , P. Zhao, J. Y. Hu, and H. Xu, 2016: The role of nonlinear forcing singular vector tendency error in causing the “spring predictability barrier” for ENSO. *J. Meteor. Res.*, **30**, 853–866, <https://doi.org/10.1007/s13351-016-6011-4>.
- , X. Q. Li, and B. Tian, 2018: Towards optimal observational array for dealing with challenges of El Niño–Southern Oscillation predictions due to diversities of El Niño. *Climate Dyn.*, **51**, 3351–3368, <https://doi.org/10.1007/s00382-018-4082-x>.
- Gao, C., X. R. Wu, and R. H. Zhang, 2016: Testing a four-dimensional variational data assimilation method using an improved intermediate coupled model for ENSO analysis and prediction. *Adv. Atmos. Sci.*, **33**, 875–888, <https://doi.org/10.1007/s00376-016-5249-1>.
- Hendon, H. H., E. Lim, G. M. Wang, O. Alves, and D. Hudson, 2009: Prospects for predicting two flavors of El Niño. *Geophys. Res. Lett.*, **36**, L19713, <https://doi.org/10.1029/2009GL040100>.
- Hu, J. Y., and W. S. Duan, 2016: Relationship between optimal precursory disturbances and optimally growing initial errors associated with ENSO events: Implications to target observations for ENSO prediction. *J. Geophys. Res. Oceans*, **121**, 2901–2917, <https://doi.org/10.1002/2015JC011386>.
- Jin, E. K., and Coauthors, 2008: Current status of ENSO prediction skill in coupled ocean–atmosphere models. *Climate Dyn.*, **31**, 647–664, <https://doi.org/10.1007/s00382-008-0397-3>.
- Jin, F. F., 1997a: An equatorial ocean recharge paradigm for ENSO. Part I: Conceptual model. *J. Atmos. Sci.*, **54**, 811–829, [https://doi.org/10.1175/1520-0469\(1997\)054<0811:AEORPF>2.0.CO;2](https://doi.org/10.1175/1520-0469(1997)054<0811:AEORPF>2.0.CO;2).
- , 1997b: An equatorial ocean recharge paradigm for ENSO. Part II: A stripped-down coupled model. *J. Atmos. Sci.*, **54**, 830–847, [https://doi.org/10.1175/1520-0469\(1997\)054<0830:AEORPF>2.0.CO;2](https://doi.org/10.1175/1520-0469(1997)054<0830:AEORPF>2.0.CO;2).
- Kalnay, E., and Coauthors, 1996: The NCEP/NCAR 40-Year Reanalysis Project. *Bull. Amer. Meteor. Soc.*, **77**, 437–471, [https://doi.org/10.1175/1520-0477\(1996\)077<0437:TNYRP>2.0.CO;2](https://doi.org/10.1175/1520-0477(1996)077<0437:TNYRP>2.0.CO;2).
- Keenlyside, N., and R. Kleeman, 2002: Annual cycle of equatorial zonal currents in the Pacific. *J. Geophys. Res.*, **107**, 3093, <https://doi.org/10.1029/2000jc000711>.
- Kim, S. T., J. Y. Yu, A. Kumar, and H. Wang, 2012: Examination of the two types of ENSO in the NCEP CFS model and its extratropical associations. *Mon. Wea. Rev.*, **140**, 1908–1923, <https://doi.org/10.1175/MWR-D-11-00300.1>.
- Kramer, K., and H. A. Dijkstra, 2013: Optimal localized observations for advancing beyond the ENSO predictability barrier. *Nonlinear Processes Geophys.*, **20**, 221–230, <https://doi.org/10.5194/npg-20-221-2013>.
- Latif, M., and Coauthors, 2001: ENSIP: The El Niño simulation intercomparison project. *Climate Dyn.*, **18**, 255–276, <https://doi.org/10.1007/s003820100174>.
- Lopez, H., and B. P. Kirtman, 2014: WWBS, ENSO predictability, the spring barrier and extreme events. *J. Geophys. Res. Atmos.*, **119**, 10114–10138, <https://doi.org/10.1002/2014JD021908>.
- Luo, J. J., S. Masson, S. K. Behera, and T. Yamagata, 2008: Extended ENSO predictions using a fully coupled ocean–atmosphere model. *J. Climate*, **21**, 84–93, <https://doi.org/10.1175/2007JCLI1412.1>.
- Macmynowski, D. G., and E. Tziperman, 2008: Factors affecting ENSO’s period. *J. Atmos. Sci.*, **65**, 1570–1586, <https://doi.org/10.1175/2007JAS2520.1>.
- McPhaden, M. J., and Coauthors, 1998: The Tropical Ocean–Global Atmosphere observing system: A decade of progress. *J. Geophys. Res.*, **103**, 14 169–14 240, <https://doi.org/10.1029/97JC02906>.
- , S. E. Zebiak, and M. H. Glantz, 2006: ENSO as an integrating concept in Earth science. *Science*, **314**, 1740–1745, <https://doi.org/10.1126/science.1132588>.
- Morss, R. E., and D. S. Battisti, 2004a: Evaluating observing requirements for ENSO prediction: experiments with an intermediate coupled model. *J. Climate*, **17**, 3057–3073, [https://doi.org/10.1175/1520-0442\(2004\)017<3057:EORFEP>2.0.CO;2](https://doi.org/10.1175/1520-0442(2004)017<3057:EORFEP>2.0.CO;2).
- , and —, 2004b: Designing efficient observing networks for ENSO prediction. *J. Climate*, **17**, 3074–3089, [https://doi.org/10.1175/1520-0442\(2004\)017<3074:DEONFE>2.0.CO;2](https://doi.org/10.1175/1520-0442(2004)017<3074:DEONFE>2.0.CO;2).
- Mu, M., W. S. Duan, D. K. Chen, and W. D. Yu, 2015: Target observations for improving initialization of high-impact ocean–atmospheric environmental events forecasting. *Natl. Sci. Rev.*, **2**, 226–236, <https://doi.org/10.1093/nsr/nwv021>.
- Qi, Q. Q., W. S. Duan, F. Zheng, and Y. M. Tang, 2017: On the “spring predictability barrier” for strong El Niño events as derived from an intermediate coupled model ensemble prediction system. *Sci. China Earth Sci.*, **60**, 1614–1631, <https://doi.org/10.1007/s11430-017-9087-2>.
- Reichstein, M., C. V. Gustau, B. Stevens, M. Jung, J. Denzler, N. Carvalhais, and Prabhat, 2019: Deep learning and process understanding for data-driven Earth system science. *Nature*, **566**, 195–204, <https://doi.org/10.1038/s41586-019-0912-1>.
- Smith, T. M., R. W. Reynolds, T. C. Peterson, and J. Lawrimore, 2008: Improvements to NOAA’s historical merged land–ocean

- surface temperature analysis (1880–2006). *J. Climate*, **21**, 2283–2296, <https://doi.org/10.1175/2007JCLI2100.1>.
- Stainforth, D. A., and Coauthors, 2005: Uncertainty in predictions of the climate response to rising levels of greenhouse gases. *Nature*, **433**, 403–406, <https://doi.org/10.1038/nature03301>.
- Sugiura, N., T. Awaji, S. Masuda, T. Mochizuki, T. Toyoda, T. Miyama, H. Igarashi, and Y. Ishikawa, 2008: Development of a four-dimensional variational coupled data assimilation system for enhanced analysis and prediction of seasonal to interannual climate variations. *J. Geophys. Res.*, **113**, C10017, <https://doi.org/10.1029/2008JC004741>.
- Tao, L. J., R. H. Zhang, and C. Gao, 2017: Initial error-induced optimal perturbations in ENSO predictions, as derived from an intermediate coupled model. *Adv. Atmos. Sci.*, **34**, 791–803, <https://doi.org/10.1007/s00376-017-6266-4>.
- , C. Gao, and R. H. Zhang, 2018: ENSO predictions in an intermediate coupled model influenced by removing initial condition errors in sensitive areas: A target observation perspective. *Adv. Atmos. Sci.*, **35**, 853–867, <https://doi.org/10.1007/s00376-017-7138-7>.
- , —, and —, 2019: Model parameter-related optimal perturbations and their contributions to El Niño prediction errors. *Climate Dyn.*, **52**, 1425–1441, <https://doi.org/10.1007/s00382-018-4202-7>.
- Tian, B., and W. S. Duan, 2016: Comparison of constant and time-variant optimal forcing approaches in El Niño simulations by using the Zebiak–Cane model. *Adv. Atmos. Sci.*, **33**, 685–694, <https://doi.org/10.1007/s00376-015-5174-8>.
- Timmermann, A., and Coauthors, 2018: El Niño–Southern Oscillation complexity. *Nature*, **559**, 535–545, <https://doi.org/10.1038/s41586-018-0252-6>.
- Wu, X. R., G. J. Han, S. Q. Zhang, and Z. Y. Liu, 2016: A study of the impact of parameter optimization on ENSO predictability with an intermediate coupled model. *Climate Dyn.*, **46**, 711–727, <https://doi.org/10.1007/s00382-015-2608-z>.
- Yu, L., R. A. Weller, and W. T. Liu, 2003: Case analysis of a role of ENSO in regulating the generation of westerly wind bursts in the Western Equatorial Pacific. *J. Geophys. Res.*, **108**, 3128, <https://doi.org/10.1029/2002JC001498>.
- Zebiak, S. E., and M. A. Cane, 1987: A Model El Niño–Southern Oscillation. *Mon. Wea. Rev.*, **115**, 2262–2278, [https://doi.org/10.1175/1520-0493\(1987\)115<2262:AMENO>2.0.CO;2](https://doi.org/10.1175/1520-0493(1987)115<2262:AMENO>2.0.CO;2).
- Zhang, R. H., S. E. Zebiak, R. Kleeman, and N. Keenlyside, 2005: Retrospective El Niño forecasts using an improved intermediate coupled model. *Mon. Wea. Rev.*, **133**, 2777–2802, <https://doi.org/10.1175/MWR3000.1>.
- , and C. Gao, 2016: The IOCAS intermediate coupled model (IOCAS ICM) and its real-time predictions of the 2015–2016 El Niño event. *Sci. Bull. (Beijing)*, **61**, 1061–1070, <https://doi.org/10.1007/s11434-016-1064-4>.
- , S. E. Zebiak, R. Kleeman, and N. Keenlyside, 2003: A new intermediate coupled model for El Niño simulation and prediction. *Geophys. Res. Lett.*, **30**, 2012, <https://doi.org/10.1029/2003GL018010>.
- , F. Zheng, J. Zhu, Y. Pei, Q. Zheng, and Z. Wang, 2012: Modulation of El Niño–Southern Oscillation by freshwater flux and salinity variability in the tropical Pacific. *Adv. Atmos. Sci.*, **29**, 647–660, <https://doi.org/10.1007/s00376-012-1235-4>.
- , L. J. Tao, and C. Gao, 2018: An improved simulation of the 2015 El Niño event by optimally correcting the initial conditions and model parameters in an intermediate coupled model. *Climate Dyn.*, **51**, 269–282, <https://doi.org/10.1007/s00382-017-3919-z>.
- Zheng, F., and J. Zhu, 2016: Improved ensemble-mean forecasting of ENSO events by a zero-mean stochastic error model of an intermediate coupled model. *Climate Dyn.*, **47**, 3901–3915, <https://doi.org/10.1007/s00382-016-3048-0>.
- , H. Wang, and J. Zhu, 2009a: ENSO ensemble prediction: Initial error perturbations vs. model error perturbations. *Chin. Sci. Bull.*, **54**, 2516–2523, <https://doi.org/10.1007/s11434-009-0179-2>.
- , J. Zhu, H. Wang, and R. H. Zhang, 2009b: Ensemble hindcasts of ENSO events over the past 120 years using a large number of ensembles. *Adv. Atmos. Sci.*, **26**, 359–372, <https://doi.org/10.1007/s00376-009-0359-7>.
- Zhu, Y. C., and R. H. Zhang, 2018: An Argo-derived background diffusivity parameterization for improved ocean simulations in the tropical Pacific. *Geophys. Res. Lett.*, **45**, 1509–1517, <https://doi.org/10.1002/2017GL076269>.