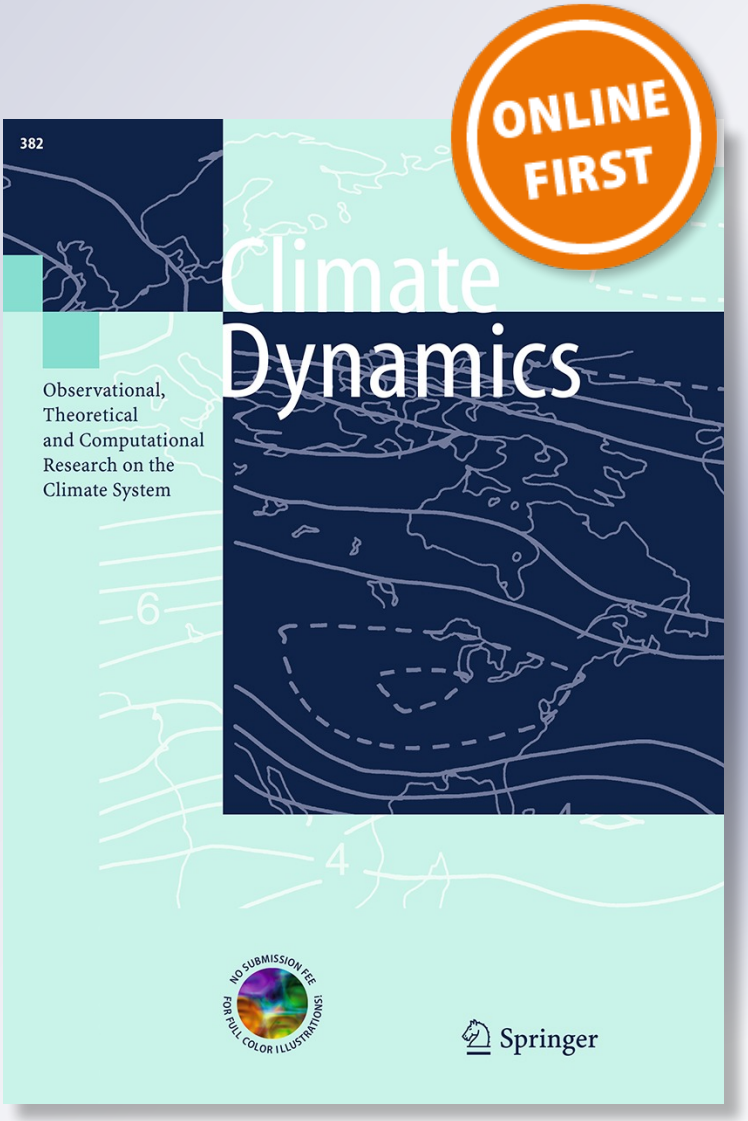
Optimal error growth of South Asian monsoon forecast associated with the uncertainties in the sea surface temperature

Siraj Ul Islam, Youmin Tang & Peter L. Jackson

Climate Dynamics
Observational, Theoretical and Computational Research on the Climate System

ISSN 0930-7575

Clim Dyn
DOI 10.1007/s00382-015-2686-y



Springer

Clim Dyn
DOI 10.1007/s00382-015-2686-y

Optimal error growth of South Asian monsoon forecast associated with the uncertainties in the sea surface temperature

Siraj Ul Islam[1] · **Youmin Tang**[1] · **Peter L. Jackson**[1]

Received: 28 December 2014 / Accepted: 23 May 2015

Abstract A recently developed method of climatic relevant singular vectors (CSVs) is applied to an atmospheric general circulation model (CAM4) to investigate the optimal error growth of the South Asian monsoon (SAM) seasonal prediction due to uncertainties in initial conditions of the sea surface temperature (SST). Emphasis is placed on the investigation of the optimal error growth of SAM seasonal forecast due to the SST uncertainties in the Indian and the equatorial Pacific Oceans. It is found that the uncertainties in the Indian Ocean can result in much larger error growth of SAM seasonal prediction than those in equatorial Pacific Ocean. Most of the CSV patterns over the Indian Ocean resembled a dipole-like structure with opposite signs spanning the northern and southern Indian Ocean. It is seen that the CSVs error growth rate changes significantly depending on the initial states whereas the CSVs patterns are insensitive to the initial conditions. The CSV patterns and error growth rates, calculated using CAM4, are also compared against those using coupled model CCSM4, indicating that the CSVs patterns from CAM4 are similar to those from CCSM4 coupled model while the error growth rate is lower in CAM4 than in CCSM4. Ensemble summer hindcasts, for the period from 2000 to 2009 over Indian Ocean, are constructed by the CSVs. For the purpose of comparison, ensemble forecasts constructed by the time lag ensemble (TLE) method are also conducted. It is found that the ensemble mean prediction by CSVs has a better skill than both the prediction by TLE and by the control run, in particular for predictions longer than 3 months, indicating the merit of CSV for SAM ensemble forecast.

Keywords Predictability and optimal error growth · South Asian monsoon · The tropical oceans · Ensemble prediction

1 Introduction

The chaotic nature of the atmospheric circulation imposes a deterministic limit on long-term weather forecasts (Lorenz 1963) but some large-scale atmospheric features such as El Nino/Southern Oscillation (ENSO) are potentially predictable beyond this limit (Shukla 1981). This was initially indicated by general circulation model (GCM) results which showed that a large part of the tropical variability is determined by slowly varying boundary conditions of sea surface temperature (SST), soil moisture and snow cover (Charney and Shukla 1981; Shukla 1998). Based on these indications, significant progress has been made over the past decades in understanding and predicting ENSO and large-scale tropical features associated with its variation. Although the correlation skill of ENSO forecasts have improved remarkably (Wang et al. 2009; Jin et al. 2008), seasonal prediction of other climatic features such as the Asian monsoon still needs substantial improvement. Being an integral part of global climate system having strong modulation due to the SST in Indian and Pacific Oceans (Yang and Lau 2006), the skillful monsoon forecasting becomes an issue of immense importance.

Several attempts have been made for the forecast of the Asian monsoon particularly the South Asian monsoon (SAM or Indian monsoon hereafter). Many studies have shown that the current skill of Asian monsoon forecasts

✉ Youmin Tang
ytang@unbc.ca

[1] Environmental Science and Engineering, University of Northern British Columbia, 3333 University Way, Prince George, BC V2N 4Z9, Canada

Published online: 16 June 2015

using dynamical models, is poor on the seasonal time scale (Drbohlav and Krishnamurthy 2010; Chowdary et al. 2010; Sohn et al. 2012) as well as on the sub-seasonal time scale (Fu et al. 2009, 2011). A recent work by Jiang et al. (2013) showed, using NCEP Climate Forecast System (CFS) version 2, that the new version has lower skill in predicting the SAM as compared to the Southeast Asian monsoon. Studies such as Acharya et al. (2011) and Kulkarni et al. (2011) have also highlighted that the GCMs have limited skill in predicting SAM rainfall. Singh et al. 2012 has reported that GCMs have large biases in simulating the observed teleconnection pattern, which lower the skill of dynamical seasonal prediction. Beside dynamical models, many statistical models have shown inconsistent skill in predicting SAM rainfall (Gadgil et al. 2005).

In general, the lower SAM forecast skill is mainly due to the our incomplete understanding of monsoon dynamics which lead to a poor formulation of model physics, large model biases (Islam et al. 2013) and the uncertainties involved in specifying the forecast initial states. This means that there is potential to increase SAM forecast skill with (1) better climate models having more accurate representation of intraseasonal variability and better parameterizations schemes and (2) improved ensemble methodologies used in forecast initialization. Together (1) and (2) can ensure improved SAM forecast skill.

Although climate models are continuously developing with time, improving forecast ensemble strategies remain a long-term challenge to advance SAM forecast. To make a reliable ensemble, it is important to explore the impact of uncertainties in initial conditions on SAM seasonal predictions (Palmer 2000). Among many initial uncertainties involved in monsoon forecast, the uncertainties in SST can greatly impact its forecast skill. It has been argued that the essential source of SAM predictability at seasonal time scales mainly comes through the tropical SST forcings, especially ENSO (Sikka 1980; Charney and Shukla 1981; Shukla and Paolino 1983; Nigam 1994; Slingo and Annamalai 2000; Meehl and Arblaster 2002; Annamalai and Liu 2005). Along with its teleconnection with ENSO forcing, SAM has close interaction with SST in the tropical Indian Ocean (Wang et al. 2009; Chowdary et al. 2010; Kosaka et al. 2012). It is interesting to investigate how SST uncertainties impact SAM forecast error growth. This is important in its own right from the point of view of error dynamics. Additionally, a direct application of the uncertainties study is to construct the optimal ensemble.

It is well known that the ensemble mean has usually more reliable and greater skill than a single deterministic forecast (Leith 1974). The importance of ensemble forecast has been greatly acknowledged, with many approaches proposed and used to construct optimal ensembles. These methods include time lag ensemble (TLE), breed vectors, singular vectors and ensemble Kalman filter etc. (e.g., Toth and Kalnay 1993, 1997; Molteni and Palmer 1993; Moore and Kleeman 1996; Chen et al. 2004; Tang et al. 2006). However, these studies mainly focus on weather and medium range seasonal forecast on a global domain or ENSO relevant climate prediction. There is little research on how to construct the optimal ensemble for SAM seasonal prediction in the literature.

In this study, we use singular vector (hereafter named as SV) to explore the impact of initial SST uncertainty on SAM seasonal prediction and to construct the optimal ensemble. The SV is a widely used method in studying the optimal growth of initial perturbations (Moore and Mariano 1999; Palmer and Zanna 2013). The aim of using SV is to find those perturbations, superimposed on a given initial state, which grow most rapidly under the assumption that the perturbations grow linearly in time. As the fast-growing mode of prediction error is often dominated by weather-scale instabilities in a full GCM, it is problematic to use the traditional SV method to characterize the fast-growing modes associated with long-term timescale variability i.e. at longer lead time. Therefore, when applying the SV method in GCMs to study climatologically relevant problems, particular care needs to be taken to filter out the fast-growing modes of weather instabilities (Kleeman et al. 2003).

This study focuses on the implementation of the climatically-relevant SV (hereafter named as CSV) method for SAM seasonal forecasts. The CSV method was introduced by Kleeman et al. (2003) which was later applied to realistic coupled models (Tang et al. 2006). It addresses the fast error growth due to climatically relevant instabilities by running an ensemble to average out the weather noise, thereby, filtering out the atmospheric noise but retaining the climatic response. This is especially important for seasonal climate forecasts. In addition, the CSV is cost-efficient since it does not require tangent linear and adjoint models, which are expensive in computation and complicated in technique for a GCM (Tang et al. 2005). Previous studies have applied the CSV method to estimate CSVs for seasonal forecasting of SST in the Pacific Ocean such as ENSO or decadal forecasting of the North Atlantic Ocean (Hawkins and Sutton 2010). It has not been used for SAM seasonal forecasts. This study is therefore the attempt to apply CSV to investigate the optimal error growth of SAM seasonal prediction due to uncertainties in SST, using a general circulation model. The overall motivation is to use CSV to explore the error dynamics of SAM forecast, and the optimal construction of SAM ensemble forecast.

The paper is structured as follows. Section 2 discusses the CSV method and its implementation, the estimation

of the linear operator R, model description and experimental setup and choice of variables. Section 3 highlights characteristic of leading CSVs and corresponding final patterns for the Indian and Pacific Oceans. Section 4 discusses validity of CSVs and its sensitivity to varying number of EOFs and ensemble members. Ensemble SAM forecasts, constructed using CSVs, are discussed in Sect. 5. Section 6 is a brief summary and conclusion.

2 Method and models

2.1 Estimation of singular vectors

As discussed in the previous section, this study utilizes the CSV method for extracting optimum perturbation patterns. The mathematical formulation of the CSV method, based on Kleeman et al. (2003) and Tang et al. (2006), is as follows:

A general dynamical system may be written compactly as

$$X(t) = F[X(t')] \tag{1}$$

where $X(t)$ is a vector representing the system state and F is a nonlinear operator. For a small perturbation Φ, Eq. (1) can be written as

$$X(t) + \Delta X(t) = F[X(t') + \Phi] \tag{2}$$

Subtracting Eq. (1) from Eq. (2), we have

$$\Delta X(t) = R\Phi \tag{3}$$

where the linear operator R in Eq. (3) is the first-order derivative of F with respect to X (at the time of t'). It is often called the propagator of Eq. (1) and gives the time evolution of the dynamical system by representing perturbation growth matrices.

The singular vectors of the system, which are the perturbations that amplify maximally over the time period $(t - t')$, are the eigenvectors (E) of $R^T R$ with the largest real part (e.g., Buizza and Palmer 1995), where R^T is the transpose of R. Thus the SV can be obtained by two methods: the eigenvector analysis of the $R^T R$ matrix or singular value decomposition (SVD) analysis of R. Mathematically it can be shown that:

$$R(t, t')E_1 = \lambda_1 S_1 \tag{4}$$

where λ_1 represents first singular value and S_1 is first SV. Hence $\lambda_1 S_1$ can be derived by applying the propagator $R(t, t')$ to the initial pattern E_1 as shown by (4). $\lambda_1 S_1$ is called the final pattern. Here S_1 is the left vector of SVD of R. The calculation of R is discussed below.

2.2 Calculation of R and application to the SAM forecast system

For calculating R of the dynamical system (1), a perturbation variable (denoted by T_p) at the initial time and the target variable (denoted by A) used to measure forecast errors, are selected. By definition, the leading SV of A indicates what kind of uncertainty in T_p can lead to the fastest error growth for predicting A. The detailed procedure for obtaining R is similar to that described in Kleeman et al. (2003) and is summarized as:

1. An ensemble of ten forecasts with lead time of 4 months (June, July, August and September, named as JJAS) is constructed by randomly perturbing the initial T_p field with ten "very small" random patterns. The ensemble mean of A is denoted by $\overline{\Psi_0(t)}$.
2. Each of the leading three correlation EOF modes e_i of $T_p (i = 1, 2, 3)$ is added (with a multiplication factor of 0.1 to ensure linearity) to the initial condition described in step (1) and a new ensemble of 10 forecasts is produced. The corresponding ensemble mean of A is denoted by $\overline{\Psi_i(t)}$.
3. A reduced-state space matrix version r_{ij} of the propagator R is then obtained taking the difference of both the ensembles. Mathematically it can be represented as

$$\begin{aligned} \mathrm{Re}_i &= \overline{\Delta\Psi(t)} = \overline{\Psi_i}(t) - \overline{\Psi_o}(t) \\ &= \sum_{j=1}^{3} r_{ij} e_j + residual \end{aligned} \tag{5}$$

The residual in Eq. (5) is generally very small and can be ignored (Kleeman et al. 2003). The climatically relevant singular vectors are thus obtained by SVD analysis of R as aforementioned. These singular vectors are then projected back to real T_p space using the EOF basis vector expansion.

In implementing this method, it is important to note that ensemble experiments are carried out to filter the weather noise by averaging in order to extract fast error growth due to climatically relevant instabilities. Another advantage of the CSV method is that the choice of the analysis domain and optimization lead time can be made after the ensembles have been completed which allows one to explore the sensitivity to different choices without further model experiments.

2.3 CAM4 model

The model used in our seasonal forecast system is Community Atmosphere Model version 4 (CAM4) of the US National Center for Atmospheric Research (NCAR). This model is a state of the art research tool for the investigation

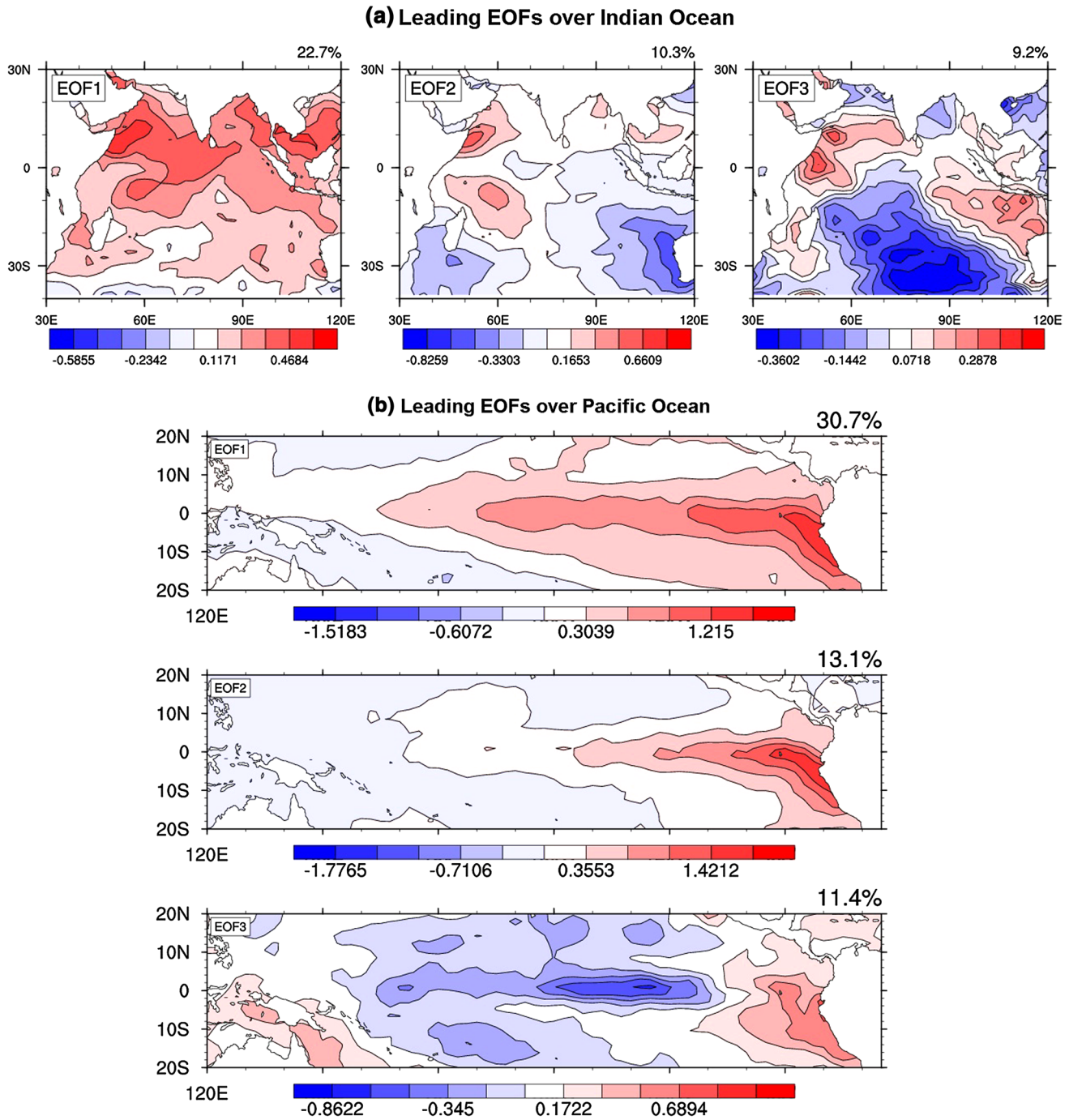


Fig. 1 First three leading correlation EOFs patterns of the June SST over **a** Indian Ocean and **b** Pacific Ocean. These patterns are calculated from monthly observed SST for time period 1980–2009, and are added as perturbation in the instantaneous initial states of persistent SST (see text for detail)

of climate variability and change. For the convenience of readers we briefly introduce CAM4 model, but for full details of the model consult the references (Neale et al. 2010; Gent et al. 2011).

CAM4 (Neale et al. 2010) is developed from CAM3 (Collins et al. 2006) with modifications to the deep convection (Neale et al. 2008), polar filtering (Anderson et al. 2009), and the polar cloud fraction in extremely cold condition parameterization schemes (Vavrus and Waliser 2008). It uses an updated convection parameterization scheme (Neale et al. 2008; Richter and Rasch 2008). This model can be used with three different dynamic schemes

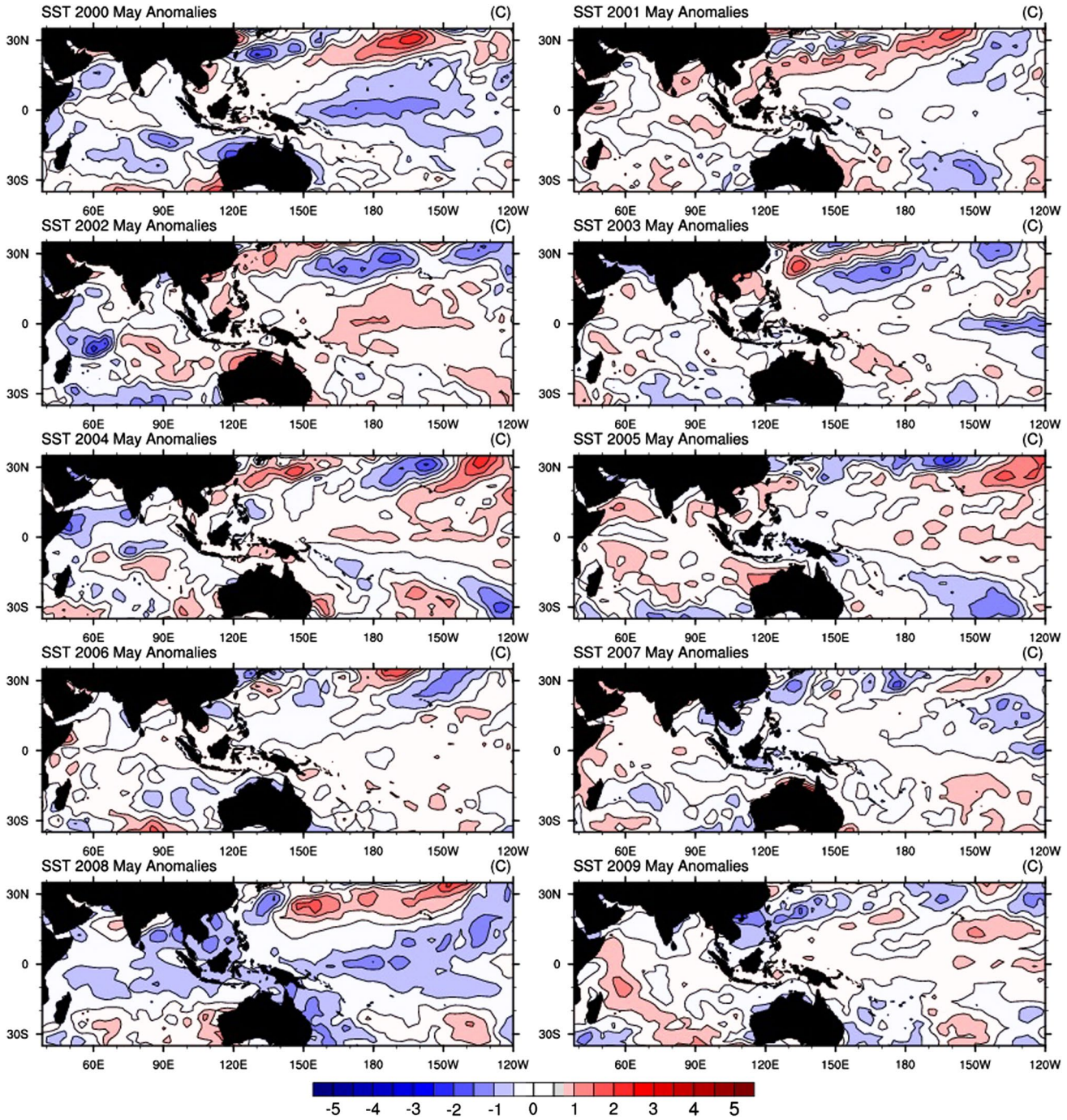


Fig. 2 SST anomalies for May of each forecast year. Each anomaly pattern is calculated based on the mean SST climatology of 1980–2009 and is kept persistent throughout the forecast. Units are in °C

along with different resolution settings. In this study, all the experiments are performed using 1.9° (latitude) × 2.5° (longitude) horizontal resolution with 26 vertical levels, a finite volume dynamical core and a hybrid terrain-following coordinate system. In its coupling mode, the coupled version of CAM4 model becomes Community Climate System Model version 4 (CCSM4, Gent et al. 2011). CCSM4 contains a coupler that exchanges fluxes and state information among all the embedded models i.e. the CAM4 atmospheric model, the Community Land Model (CLM4), the Los Alamos Parallel Ocean Program ocean model version 2.2 (POP 2.2) (Smith and Gent 2002) and the Community Ice Code version 4 (CICE4) sea ice model (Hunke and Lipscomb 2008).

2.4 Experimental setups

The ability of CAM4 to simulate the summer monsoon has been explored in our recent work (Islam et al. 2013). It is found that CAM4 overestimates the monsoon rainfall over most of the SAM region when compared to observation and CCSM4. The overall simulations of CAM4 reasonably captured the monsoon mean climatology (Islam et al. 2013), allowing its use to study predictability.

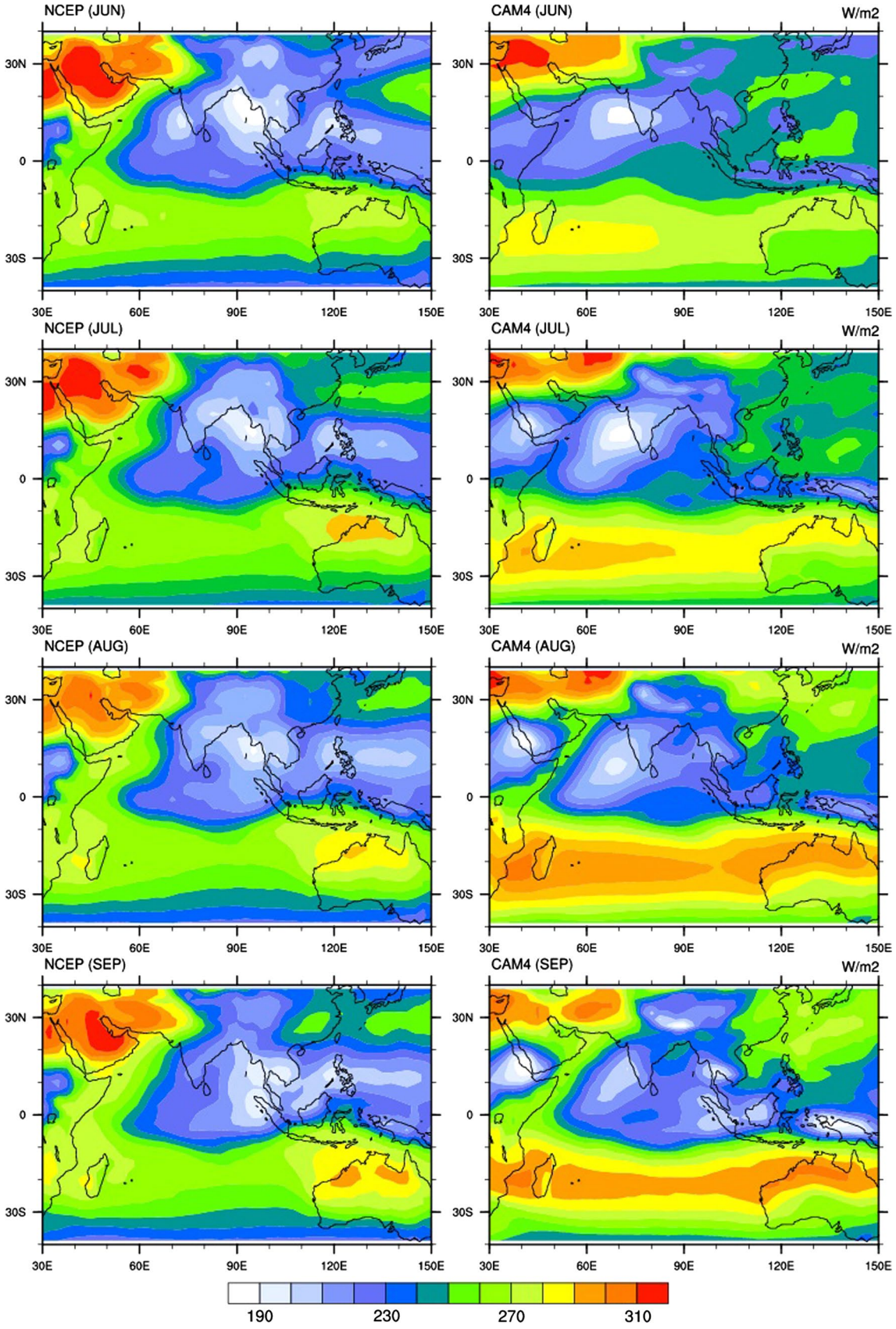
As discussed above, to implement the CSV method, three leading modes of the correlation EOF are used to perturb the initial conditions. Figure 1a, b shows the first three correlation EOF modes for SST over the Indian and Pacific Oceans. The major reason for SV analysis over the two Ocean domains is to explore the individual effect of the Indian and Pacific Ocean on the error growth of SAM, respectively. As reported in many studies, both of these Oceans can significantly affect SAM and play a key role in SAM variability (e.g., Meehl and Arblaster 2002; Annamalai and Liu 2005; Wang et al. 2009; Chowdary et al. 2010; Kosaka et al. 2012). Leading EOFs are calculated for the month of June using the observed SST data (Hadley Centre SST, Rayner et al. 2003) for the period 1980–2009. In the case of the Indian Ocean (Fig. 1a), 1st and 2nd leading EOF mode account for 39 and 12 % of the total variance. EOF1 shows the Indian Ocean Basin (IOB) Mode and EOF2 depicts the east–west Indian Ocean dipole (Saji et al. 1999, IOD) mode. The former is closely linked to ENSO, whereas the latter occurs with and without ENSO (Pokhrel et al. 2012). The two modes differ not only in spatial structure, but also in their seasonal dependency. In Fig. 1b, the 1st leading EOF over the Pacific domain represents the ENSO mode. The leading mode shows 30.7 % explained variance over the Pacific Ocean. The spatial pattern associated with the warm phase of ENSO consists of positive SST anomalies across the eastern equatorial Pacific Ocean and weaker

Fig. 3 Mean climatology of June, July, August and September for outgoing longwave radiation (OLR). NCEP OLR is shown in *left column* and CAM4 control run is in *right column*. Mean is calculated over the period 1980–2009. The *shading* corresponds to OLR values in W/m^2

Table 1 List of experiments used for CSV analysis including detail of perturbation domain, number of perturbation EOFs, number of ensembles and forecast years

Exp. name	Description	Perturbation domain	Time span	Initial conditions (IC)	No of EOFs used	Ensemble members for each EOF
IO30	30 years perturb run using EOFs as perturbations	Indian Ocean	1980–2009	Control run IC created with prescribed SST	3	10
IO10	10 years perturb run using EOFs as perturbations	Indian Ocean	2000–2009	Observed IC from NCAR's DART system	3 and 10	10 and 20
PO10	10 years perturb run using EOFs as perturbations	Pacific Ocean	2000–2009	Observed IC from NCAR's DART system	3	10
IO_CCSM4	4 years coupled run using EOFs as perturbations	Indian Ocean	2000, 2004, 2008 and 2009	Observed IC from NCAR's DART system + 100 year control run SST	3	10
TLE	Time lag ensemble forecast	–	2000–2009	6 h lag observed IC from NCAR's DART system	–	6
CSVp	Ensemble forecast using +ve CSV as perturbations	Indian Ocean	2000–2009	Observed IC from NCAR's DART system	–	10
CSVn	Ensemble forecast using −ve CSV as perturbations	Indian Ocean	2000–2009	Observed IC from NCAR's DART system	–	10
CSV30p	Ensemble forecast using +ve CSV as perturbations	Indian Ocean	1980–2009	Control run IC created with prescribed SST	–	10
Control forecast	Ensemble forecast without CSV perturbation	–	2000–2009 and 1980–2009	Observed IC from NCAR's DART system	–	10

All experiments start from Jun 1st of every year with lead time of 4 months (June, July, August and September, JJAS). In CAM4 experiments, persistence SST is used as boundary conditions (BC) whereas in CCSM4, SST is used from the multiyear year control run of ocean model



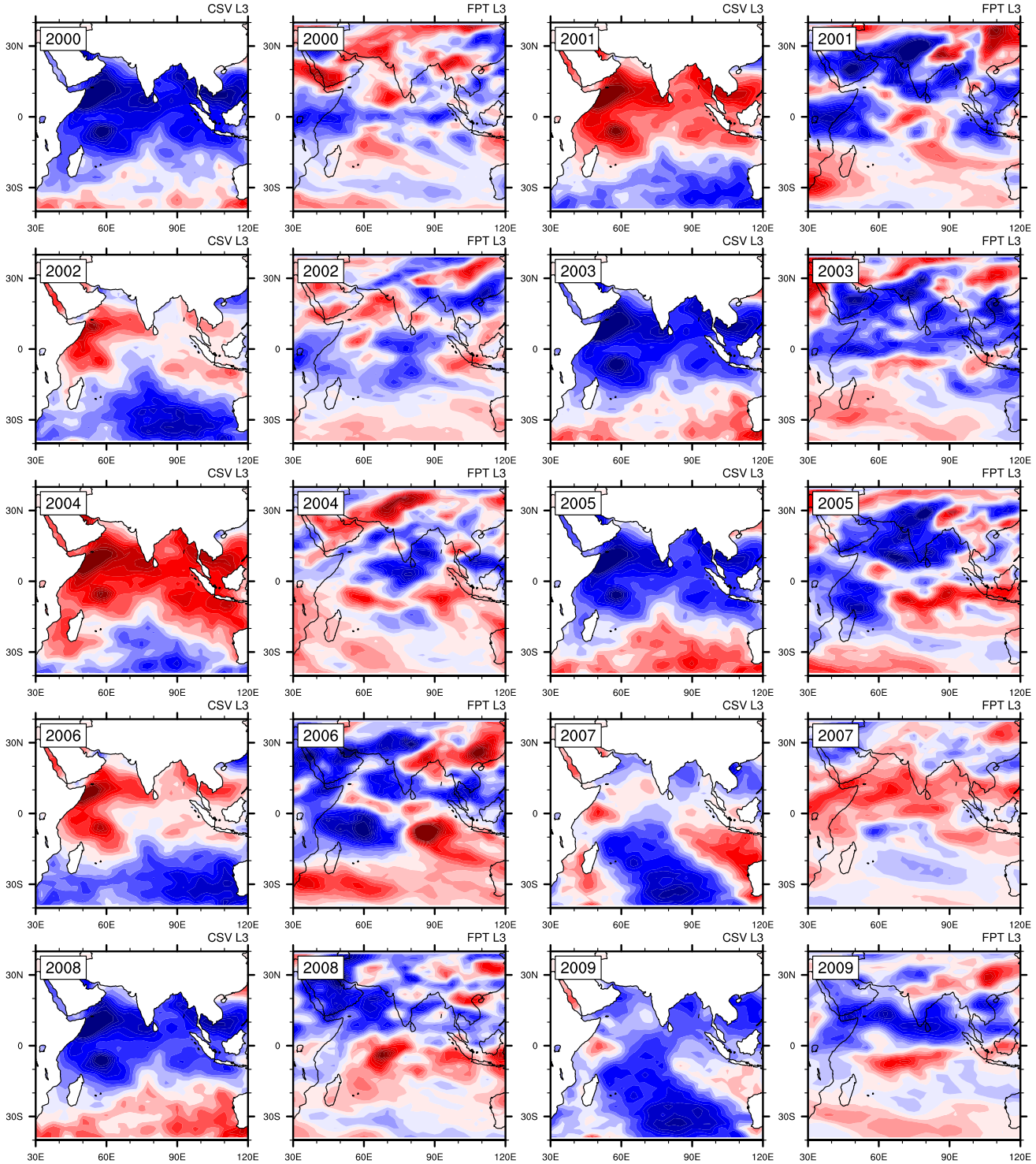


Fig. 4 The optimal leading climatology relevant singular vectors (CSVs) of SST and corresponding OLR final patterns (FPs) of, optimized for 4 months (L3 ~ lead time 3) over the Indian Ocean domain (see text for detail). Each individual CSV and FP is shown starting from year 2000 to 2009. The SST units are °C and OLR is in W/m^2

negative anomalies over the western tropical Pacific. In our CSV implementation, each of these EOFs is used as a perturbation added to the initial SST state.

Using the procedure mentioned in Sect. 2.2, for each initial condition, Eq. (5) can be evaluated by averaging ten different ensemble members. The initial conditions

used in experiments are from NCAR's Data Assimilation Research Tool (DART) system (Anderson et al. 2009). DART employs an ensemble Kalman Filter (Houtekamer et al. 2005) which nudges the underlying models toward a state that is more consistent with information from a set of observations. Overall, these initial conditions are quite close to the National Center for Environmental Prediction (NCEP) reanalysis data sets, but are more consistent with the CAM4 model. In coupled model experiments (i.e. CCSM4), the atmospheric (CAM4) and land (CLM4) counterparts are initialized using the same initial conditions as used in CAM4 initialization. To generate initial condition for ocean (POP2.2) and sea ice (CICE4) models, a multi years ocean control hindcast experiment is performed using the CORE II (Coordinated Ocean-ice Reference Experiments—Phase II, Griffies et al. 2012) input data set as boundary conditions. The CSV implementation for CCSM4 model is the same as discussed for the CAM4 model. The three leading EOF patterns in this case are calculated using the CCSM4 predicted SST instead of observations. In most of our analysis, we mainly focus on CAM4 simulations (uncoupled) and its analysis. The results from the coupled model are only used for the purpose of comparison between uncoupled and coupled CSVs.

The forecast experiments are performed for the time period from 2000 to 2009 by initializing both atmospheric and land models (embedded in CAM4) simultaneously on June 1st using DART initial conditions. The prediction lasts 4 months from June to September during which the monthly mean SST and sea ice (Hadley Centre SST, Rayner et al. 2003) at the initial time (May) are kept persistent (i.e. 1 month prior to the forecast start time). Figure 2 shows SST anomalies of month May from 2000 to 2009. These anomalies vary each year, which can influence the growth of forecast error when perturbations are applied.

To get a more robust statistical analysis, another set of ensemble forecasts is performed for the 1980–2009 time period using persistent SST as boundary forcing. The initial conditions in this case are form CAM4 control run forced with observed SST. The CSV method is implemented in the same way as in the simulations using DART initial conditions.

A summary of all the experiments using different domains, models, time periods, number of EOFs and ensemble size is given in Table 1.

2.5 Target variable

In this study, the forecast variable is chosen to be outgoing longwave radiation (OLR) and optimum interval (i.e. lead time) is 4 months. The use of OLR, instead of precipitation, is because it is comparatively better predicted than actual precipitation. On the other hand, it has been widely used to estimate SAM precipitation often as an index related to convective activity as lower OLR values are associated with higher cloud top heights in well-developed clouds (Schmetz and Liu 1988; Kousky and Kayano 1994; Moron 1995).

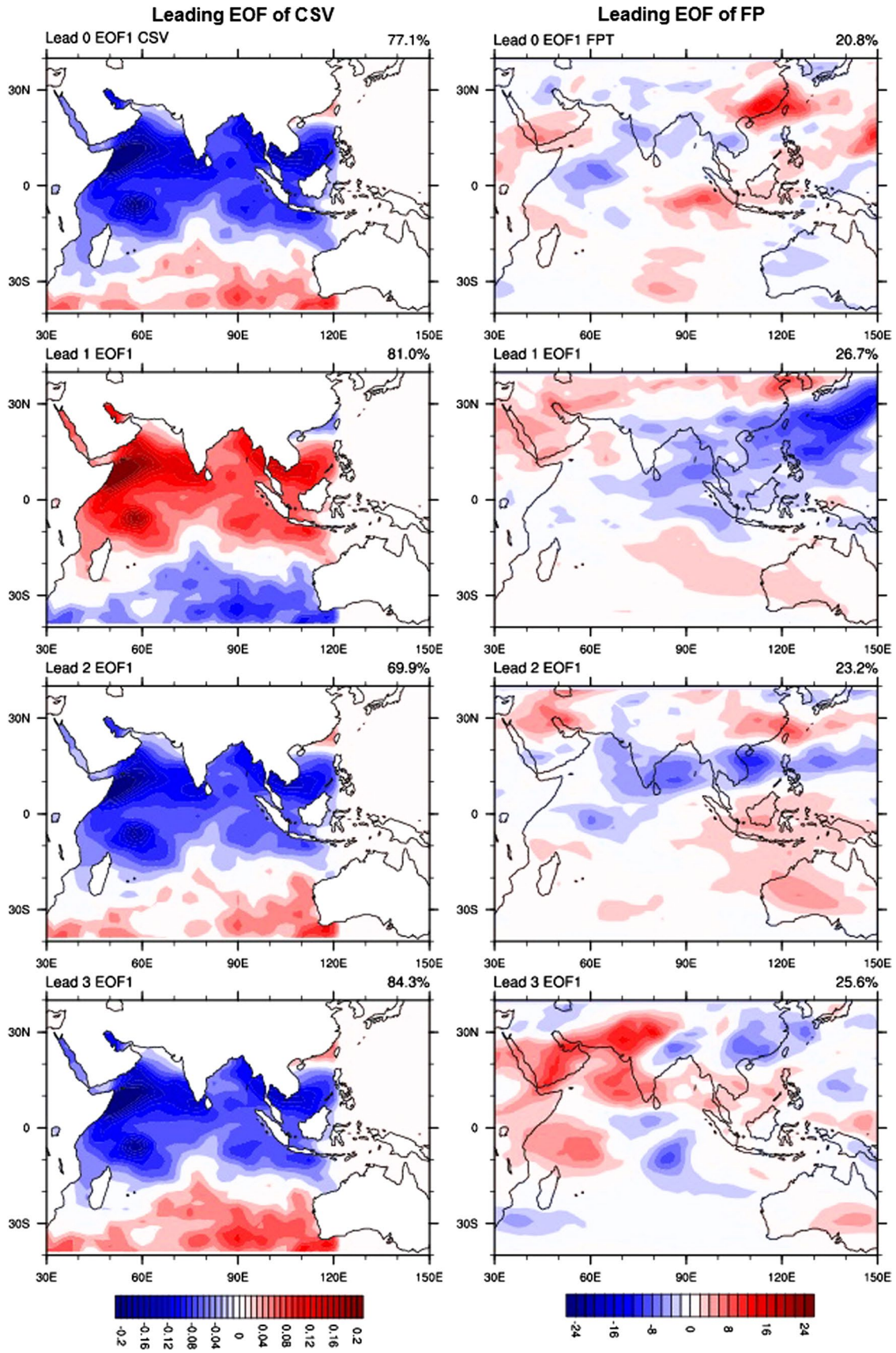
In summary, we use SST as perturbation variable (T_p) at the time $t = 0$ and investigate OLR as target variable (A) at $t = t'$. As discussed before, the main reason to select SST as perturbation variable, over the Indian and Pacific Oceans, is due to the fact that SAM variability, at the seasonal time scale, is strongly influenced by the SST in central Pacific and Indian Ocean.

3 The optimal error due to the SST uncertainties in Indian and Pacific Oceans

The precipitation over the SAM region in CAM4 simulation is validated in Islam et al. (2013) and it is found that CAM4 can well capture the main climatological features of the SAM summer precipitation. Figure 3 shows the 30 year (1980–2009) mean OLR climatology from NCEP reanalysis data (left column) and CAM4 control simulation (right column). Generally, negative (positive) values of OLR indicate more (less) convection and hence more (less) cloud coverage. In observations, for all four monsoon months, there is a widespread area of heavy convection around the northern Indian Ocean and Indian continent reflecting heavier rainfall. In contrast, most of the equatorial Indian Ocean has higher OLR values showing less convection. In CAM4, these observed OLR feature are well captured except in some regions such as suppressed rainfall over the Southern China Sea and excessive rainfall over the western Arabian Sea. Overall, the OLR simulated by CAM4 compares well with the observed OLR climatology.

3.1 Results during the 2000–2009 period

We now examine how error growth varies with different initial conditions and lead time under different scenarios. In our CSV analysis, the first singular value is significantly larger than the remaining singular values. We therefore only discuss the leading CSV and its optimum growth i.e. the final pattern (FP hereafter) throughout the text. Over the Indian Ocean, the leading CSVs and corresponding FPs, optimized for time interval of 4 months (lead of 3 months), are shown in Fig. 4 for each individual forecast during 2000–2009. In case of leading optimized perturbation (i.e. CSV), a distinct north–south dipole pattern can be identified in most of the years (while ignoring the arbitrary sign of CSVs). In the years having the dipole-like CSV pattern, the corresponding FP shows error growth mainly over most of the Indian Ocean particularly over the north



◀**Fig. 5** The first leading EOF patterns of the CSVs (*left*) and FPs (*right*) computed over ten individual CSV patterns for all the four lead times. CSVs are extracted over Indian Ocean. The SST units are in °C and OLR is in W/m^2

Indian Ocean and the central Indian subcontinent. Although most of the CSVs in individual years are similar, the FPs are quite different in each year. Large variation of OLR FPs within different years could be due to the low-frequency processes causing asymmetry between the responses of OLR to the SST. This could be also due to the different initial SST state anomalies which may significantly change the CSV growth. Indeed these SST anomalies, as the background state, play a crucial role in modifying the OLR spatial variation. This can be seen in some observed summer OLR anomalies over the SAM region for different years. For example, Charabi and Abdul-Wahab (2009) found that the July anomaly of OLR in ENSO years such as 2000 and 2002 are significantly different compared to other years. They have seen that the geographical variations of OLR anomalies are positive in the west Indian Ocean and negative in the East Indian Ocean in 2000 and 2002 whereas the patterns for years such as 1997 (strong El Nino year) and 1994 (strong IOD year) are the opposite.

To extract the general characteristic of these CSVs and FPs, we perform the EOF analysis for the 10 individual years for each lead time. The first leading EOF mode of CSVs (left column) and FPs (right column) is shown for all lead times in Fig. 5. As can been seen, the CSV patterns now more prominently resemble the dipole structure, which is not dependent on the lead time if the arbitrary sign is ignored. The FP has differences for different lead times, due to the impact of initial stated on perturbation growth as discussed before. During the 4 month optimization interval, the error growth gets concentrated over the northern Indian Ocean and Indian subcontinent. Therefore the maximum perturbation growth is over those areas which are usually used to measure SAM intensity and variability i.e. the northern Indian Ocean and Indian subcontinent.

The observed OLR climatology presented in Fig. 3 shows convective center over the Bay of Bengal whereas, due to systematic bias, this convection is shifted to the west in CAM4 simulation. Such a systematic bias may have the impact on error growth over that region (Fig. 5) due to initial perturbation. However we believe that the error growth shown in Fig. 5 is mainly due to the uncertainties of SSTs, as the CSV method focuses on the difference between two models run, by which the systematic bias should be greatly reduced. Although a strong nonlinear interaction between bias and initial uncertainty could impact the error growth, the error growth calculated here is based on the linear theory framework. In any case, it should be appropriated to state that the high error growth over the head Bay of Bengal is due to the high sensitivity of the OLR over this region to the uncertainty of SST in CAM4.

We now discuss the Pacific Ocean case, where CSVs are extracted over the ENSO region and the prediction target is still the OLR of SAM region. The results show that the CSVs are not sensitive to the initial conditions as in the Indian Ocean case, although the FPs vary significantly (not shown here). Figure 6 shows the leading EOFs of CSVs and FPs for all initial conditions at different lead times. A prominent structure in CSVs is the equatorial ENSO-like mode, representing a large-scale structure with major weighting in the equatorial eastern Pacific and less weighting in the western Pacific. The CSVs in this case are similar to SVs analyzed in many intermediate models when targeting ENSO forecasts (e.g., Chen et al. 1997). This resemblance indicates the link between the error growth of ENSO prediction and that of SAM, suggesting ENSO as a major source of SAM predictability. The FPs (Fig. 6) in the Pacific Ocean case has close similarities to the FPs calculated in the Indian Ocean case (Fig. 5) except in some regions.

The above CSV analysis considers the individual impact of the Indian and Pacific Oceans on the error growth of SAM prediction where the Indian and Pacific Oceans are perturbed separately. We will next explore the total impact of the two Oceans on the error growth of SAM prediction. For this purpose, we analyze CSVs for the Indian-Pacific Ocean. In this case, the spatial structure of CSVs over the Indian and Pacific Oceans are the same as seen in the individual oceans case (not shown here). Figure 7 shows the leading EOF computed using 10 individual FPs from all years for each lead time. The interesting outcome from the joint analysis of the two oceans is the damped error growth over most of the Indian subcontinent and the Indian Ocean. The SST uncertainties from two different oceans partially cancel their growth with increasing lead time. This means that domain size choice is crucial, as uncertainties over a large region can suppress its growth and contribute less in error optimization. This was also confirmed by running an ensemble forecast using the CSVs over the whole Indian and Pacific Oceans which shows lower skill compared to the forecast perturbed with CSVs over Indian Ocean only. We will discuss the SAM ensemble forecast later.

The above discussion of error growth can be further investigated by analyzing singular values of CSVs over the Indian Ocean and over the Pacific Ocean. These values measure the fastest growth rate of perturbation during an optimum interval. The annual variation of singular values is shown in Fig. 8 for the Indian (Fig. 8a) and Pacific (Fig. 8b) Oceans. The singular values of all the lead times are averaged each year to show an overall error growth. In the case of the Indian Ocean only, the average singular values are higher than those of the Pacific Ocean. In Fig. 8d, the lead

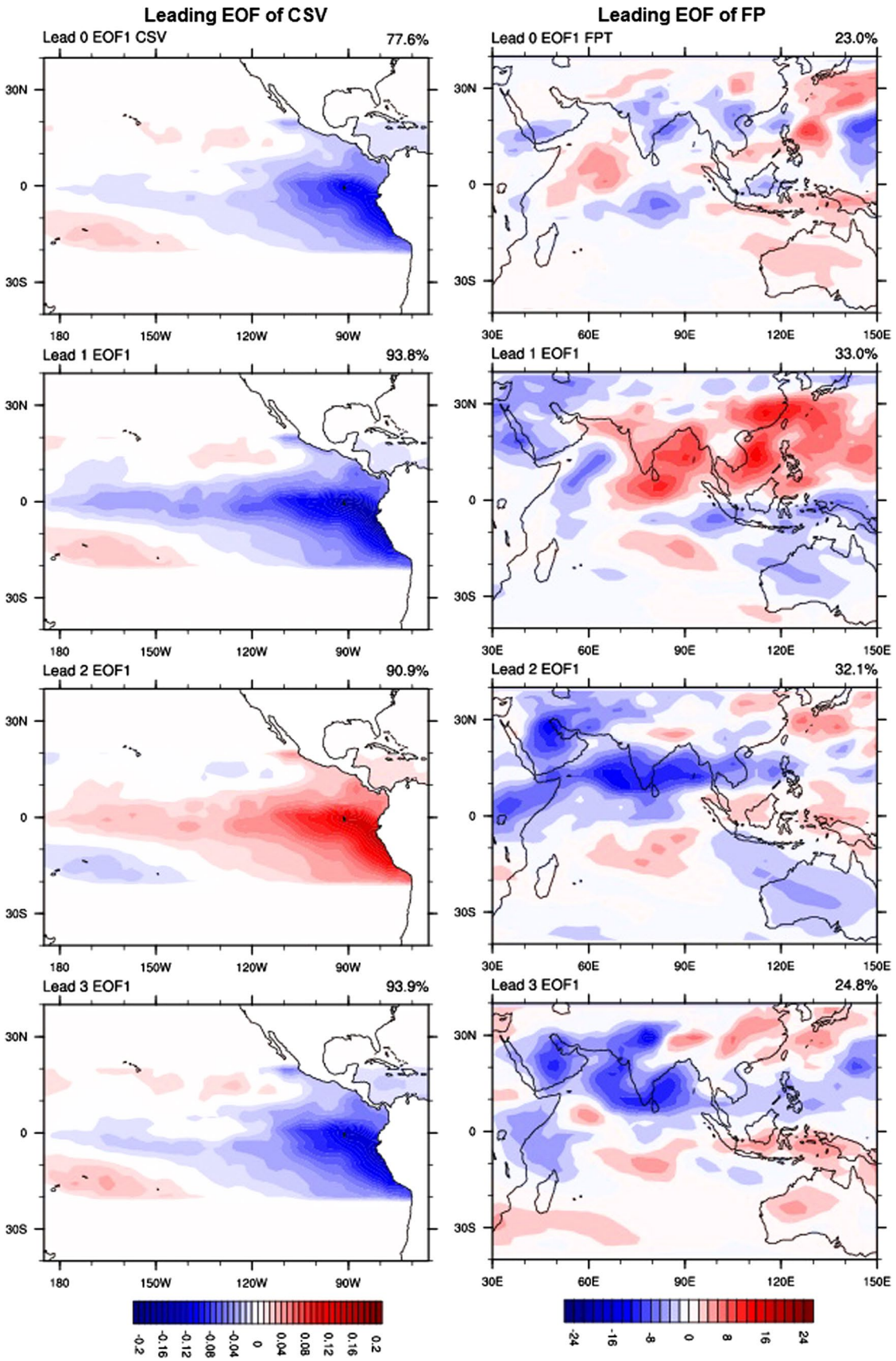


Fig. 6 Same as Fig. 5 but for CSVs calculated over Pacific Ocean and corresponding FPs over SAM region. The SST units are in °C and OLR is in W/m^2

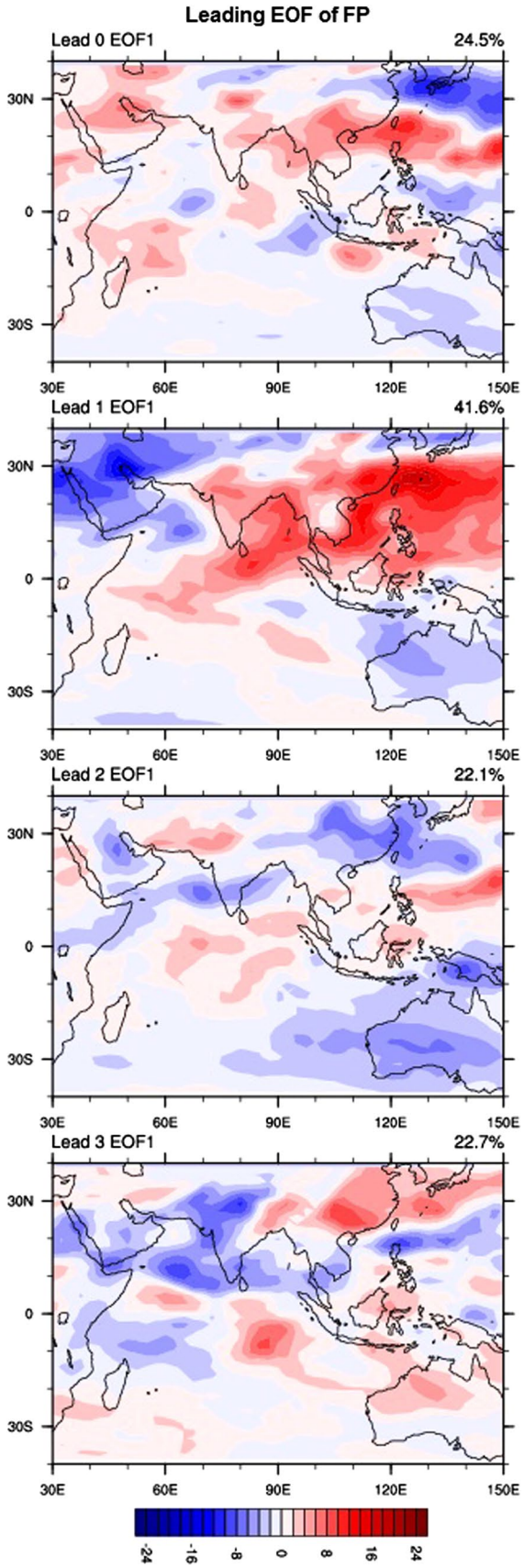


Fig. 7 Same as Fig. 6 but for CSV calculated over combined Indian-▶ Pacific Ocean. Only FPs are shown here

time dependence of singular values is shown by averaging all the individual years. The gradual increase of singular values from lead times of 0 to lead times of 3 months is more prominent in the Indian Ocean case (black line) compared to the Pacific Ocean case (red line).

3.2 Results during the 1980–2009 period

Due to the availability of DART initial conditions, based on real observation, we only used 10 years (2000–2009) of independent experiments in Sect. 3.1. For more rigorous statistics, in this section we extend the CSV experiment to the period 1980–2009 (30 individual SAM summer forecasts, IO30 hereafter), however the initial condition in this case are not from real observation. All settings in this experiment are the same as those of 10-year previous analyses except the initial conditions, which were produced from the control simulation of CAM4 forced by observed SST. As mentioned in the singular values discussion in Sect. 3.1, the Indian Ocean has a more significant impact on the SAM seasonal prediction than the Pacific Ocean. We therefore emphasize results only over the Indian Ocean in further analyses. For simplicity, we only focus on analysis at the seasonal time scale (i.e. 3 month lead time).

The leading EOF mode obtained using the 30 individual CSVs and FPs is shown in Fig. 9 at a lead time of 3 months. The spatial structures of CSV found in IO30 experiment are very similar to CSVs found in the 10 year analysis (Fig. 5) confirming the robustness of CSVs obtained in the analysis of Sect. 3.1. The FP has some regional differences along with significant year to year variation (not shown here) mainly due to the impact of initial state superimposed by CSV as discussed in Sect. 3.1. The annual variation of IO30 singular values (Fig. 8c) and its lead time variation (Fig. 8d, blue line) are quite consistent in magnitude with the 10 year CSV values.

It has been also found in previous studies that the SVs are insensitive to initial conditions in many models (e.g., Chen et al. 1997; Xue et al. 1997; Zhou et al. 2007; Cheng et al. 2010). To explore the sensitivity of CSVs to initial conditions, we computed the spatial correlation between the EOF_c and each individual CSV pattern, as in Tang and Deng (2011), where the EOF_c is the first EOF mode obtained by all CSVs.

If the EOF_c and the CSV_i are denoted by the normalized one-dimensional vectors eof_c and csv_i, respectively, the spatial correlation R_{sp} is calculated as below,

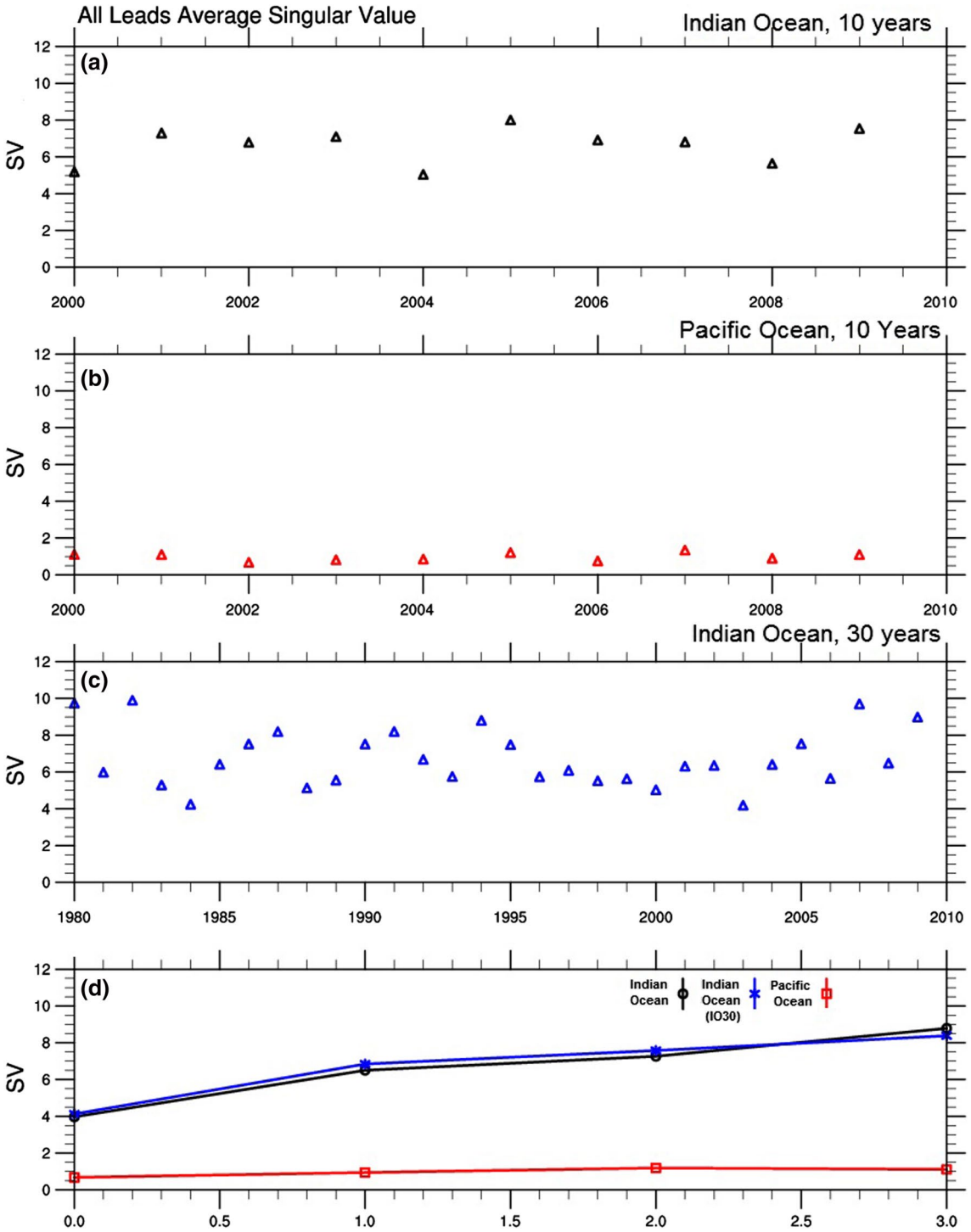


Fig. 8 Annual variation of singular values averaged over all the lead times for **a** Indian Ocean, **b** Pacific Ocean and **c** IO30 experiments. The lead time variation of **a**, **b** and **c** singular values is in shown in **d** (see text for detail). SV are scaled dividing by 100

$$R^{i}_{sp} = \frac{1}{NG-1}\sum_{G=1}^{NG} eof_c(G) * csv_i(G) \tag{6}$$

where NG is the number of total model grids over the model domain and $i = 1, 2, \ldots 10$. The spatial correlation values are calculated for Indian Ocean and Pacific Ocean cases for each individual year and are shown in Fig. 10a, b. For most years (over 90 %), the absolute value of the spatial correlation coefficient is over 0.90, for both Indian Ocean (Fig. 10a) and Pacific Ocean (Fig. 10b) cases. The

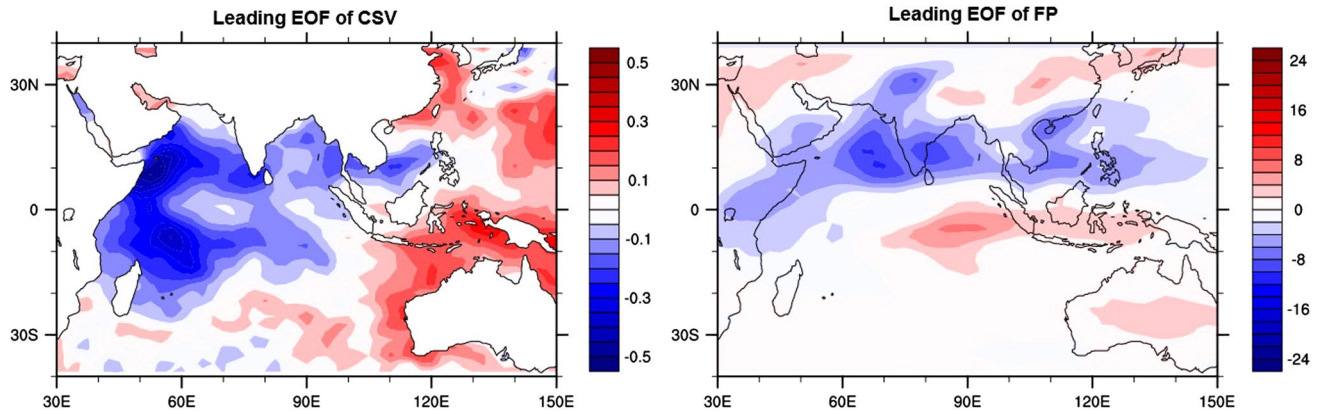


Fig. 9 Same as Fig. 5 but for CSVs and FPs calculated over Indian Ocean using IO30 experiment. 30 years CSVs and FPs are used for EOF analysis. The initial conditions for IO30 experiment are generated using CAM4 prescribed SST run

positive phase of the same pattern should be regarded as equivalent to the negative phase, at least from a linear system perspective. In addition, the pattern sign can be arbitrary from the view of EOF analysis. Thus, the preliminary features of CSVs, as represented by the first EOF mode (equivalent to the average), are insensitive to initial conditions. This conclusion is further confirmed by using CSVs from the IO30 experiment by increasing the individual years to 30 i.e. $i = 1, 2, \ldots 30$ as shown in Fig. 10c. Higher values of individual correlations indicate independence of the CSVs from initial conditions.

3.3 The results from the coupled model CCSM4

The CSVs and FPs discussed above are based on atmospheric model simulation with persistent SST boundary conditions. How do the CSVs change in a coupled model? As argued in the introduction, the air-sea interaction plays a crucial role in SAM simulation and prediction. To explore this issue, we calculate the CSVs for CCSM4. The leading CSVs and FPs of CCSM4 over the Indian Ocean for all lead times are shown in Fig. 11. We only show one particular year here, i.e. 2000, as the CSVs for other years are similar. Interestingly, the spatial structure of the CCSM4 CSVs is similar to CAM4 CSVs, but the growth rate shown in Fig. 11 has differences with the CAM4 counterpart in Fig. 8. The growth rate of CCSM4 varies more significantly with lead times than that of CAM4, so that the former is larger than the latter for longer lead time beyond 2 months. This suggests that the air-sea coupling could act as an amplifier by positive feedbacks to strengthen the perturbation growth for a long lead time. Due to the difference in the growth rate, the CCSM4 show stronger anomalies than CAM4 in their FPs, as evidenced in Figs. 8 and 11.

The similarities of CSVs between CAM4 and CSAM4 suggest that one may be able to use the CAM4 to construct ensemble prediction for the CCSM4, which is more computational efficient and has potential in operational forecast. We leave this topic for the future when we will investigate the nature of coupled CSVs and optimal ensemble SAM forecast by CCSM4.

4 Sensitivity experiments of CSV to some parameters

Before using CSVs to construct ensemble forecasts, it is necessary to evaluate CSV robustness to the various choices made in the implementation of the methodology. In this section we therefore discuss the robustness of CSVs to the number of EOFs used as initial perturbations, and the number of ensemble members used in calculating means over the Indian Ocean with help of sensitivity experiments.

4.1 Convergence of CSV with number of EOFs

In this study, the first three EOF modes are used to generate the CSVs. In the sensitivity experiment, we explore the CSVs with the number of EOFs used for initial perturbations. For simplicity, we randomly choose an initial condition of an individual year for the sensitivity analysis. Figure 12 shows CSVs computed using different numbers of EOFs. The lead time for CSV is 4 months. As we increase the number of EOF perturbations from 3 to 10 EOFs, the CSV structure does not change significantly. The corresponding FP gets optimized even with a relatively small number of EOFs i.e. 3–5 EOFs. This suggests that the dominant features of the CSVs can be captured by the leading 3 EOF modes. Indeed, increasing the number of EOF perturbations means including more noise in the initial conditions which can, beyond a certain limit, start degrading the optimization of perturbation. This analysis suggests that our

choice of using the three leading SST EOFs in construction of CSVs is appropriate. We repeated the same procedure for other years with similar results.

4.2 Convergence of CSV with number of ensemble members

Figure 13 shows how the first singular value changes for different numbers of ensemble members, when using three EOFs as perturbations. If only ten ensemble members are used, singular values converge to its minimum values for lead time 3 months and for all individual years (colored lines). Beyond 10–15 ensemble members, the singular values remain quite similar. This means that the true mean can

Fig. 11 (*Spatial*) The optimal leading SST CSVs and OLR FPs for lead time 0, 1, 2 and 3 obtained using CCSM4 model. The domain of perturbation is Indian Ocean. The SST units are °C and OLR is in W/m^2. (*Line*) Lead time variation of CCSM4 singular values over Indian Ocean. SV are scaled dividing by 100

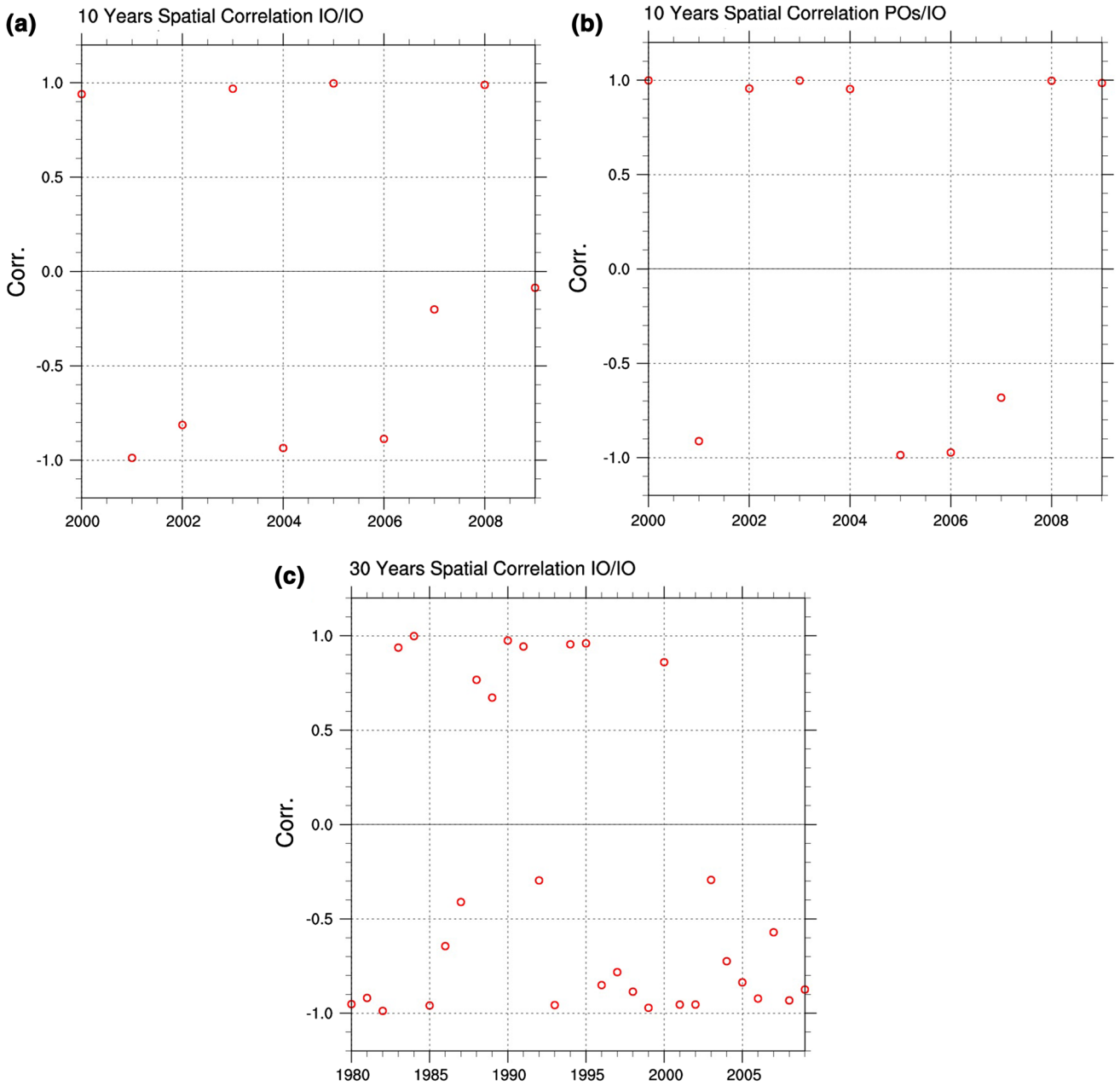
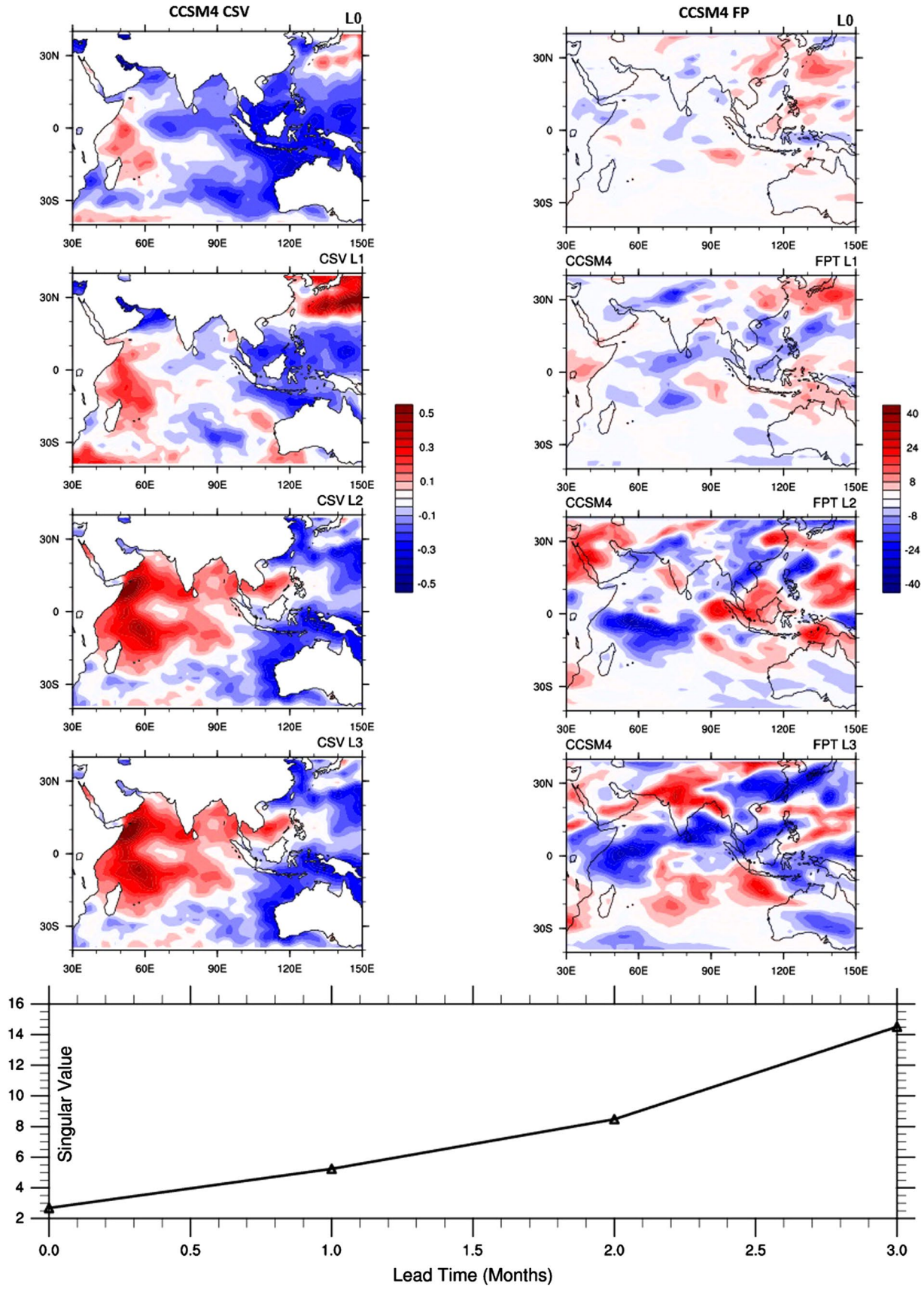
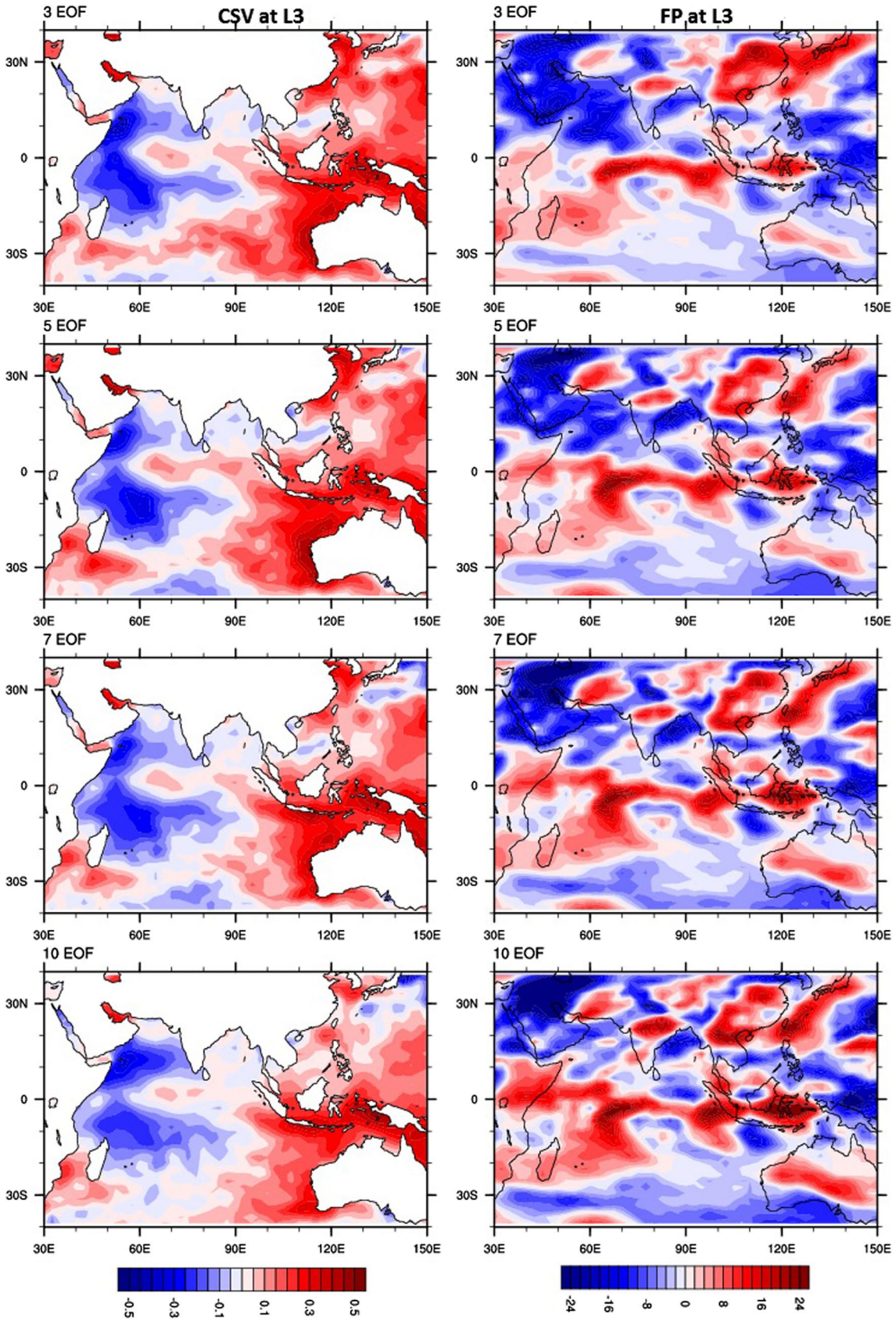


Fig. 10 The spatial correlation for **a** Indian Ocean, **b** Pacific Ocean and **c** IO30 experiments. It is computed between the first EOF mode (obtained by all CSVs) and each individual CSV pattern, same as in Tang and Deng (2011)





◀**Fig. 12** Spatial patterns of the leading CSV (*left column*) based on the increasing number of EOFs used in the perturbation. Ten leading SST correlation EOFs are used by increasing EOFs from 3 to 10 in each CSV and FPs calculation. The optimization time is 4 months. *Right column* shows associated FPs. The SST units are in °C and OLR is in W/m^2

be estimated with the use of only ten ensemble members. A similar convergence is seen for CSV spatial patterns suggesting the use of ten ensemble members being reliable for their calculation (not shown here).

4.3 Validity of CSV

The CSVs and FPs, shown in Fig. 4, are obtained by a linear approximation approach for the nonlinear model CAM4. It is therefore important to inspect whether the CSVs can result in resembled results that are directly obtained by the nonlinear model. This validation can be conducted by integrating the CSM4 twice, one from the initial condition superimposed by the CSVs (scaled by using a multiplication factor of 0.1), and the other from the initial condition only. The difference between two integrations, denoted as D_non, indicate the perturbation growth of the CSVs obtained directly by the nonlinear model, which can be compared with the FPs and served as the purpose of the validation. We conduct this experiment for each year during the period from 2000 to 2009. Figure 14a shows the leading EOF of D_non over the entire period at 3 months lead time. The leading EOF pattern of the FP directly from the CSV method, i.e. applying the propagator R to the CSVs, is also shown in Fig. 14b for comparison. It can be seen that the leading EOF of D_non (Fig. 14a) is very similar to the FP from the linear approximation (Fig. 14b) although there are some discrepancies. Despite these discrepancies, the overall features very resemble each other, indicating that the estimates of the CSVs are robust and correct.

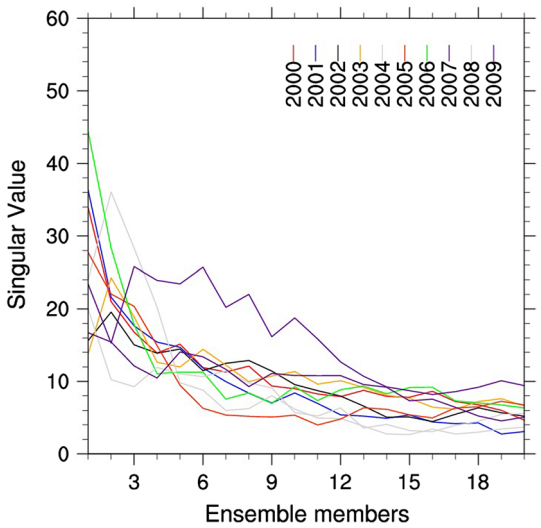


Fig. 13 The convergence of the leading singular values with increasing number of ensemble members, optimized for 4 month time interval. *Each color line* represent individual year. In each year, 20 singular values are calculated by increasing the ensemble size from 1 to 20 in the CSV method. SV are scaled dividing by 100

5 SAM ensemble forecast using CSVs

In the previous sections, we mainly focused on the general characteristics and validation of CSV and the FP. We now apply the CSVs to generate ensemble predictions. It is expected that such an ensemble prediction should be optimal and more skillful than single predictions and

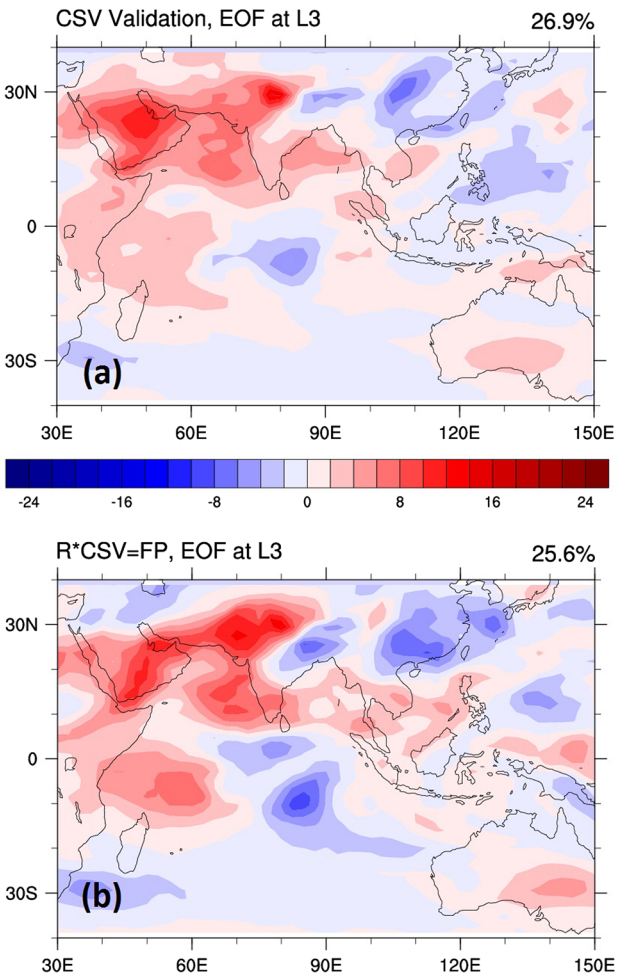


Fig. 14 Leading EOF of OLR FPs, at lead time 3, **a** calculated using CSV as perturbation and **b** estimated by applying linear propagator R to CSV i.e. by linear approximation. For **a**, the difference of perturbed and control forecast is calculated each year and leading EOF is extracted using all the differences. In **b**, EOF pattern is calculated using estimated FPs from linear approximation for all the 10 years (2000–2009)

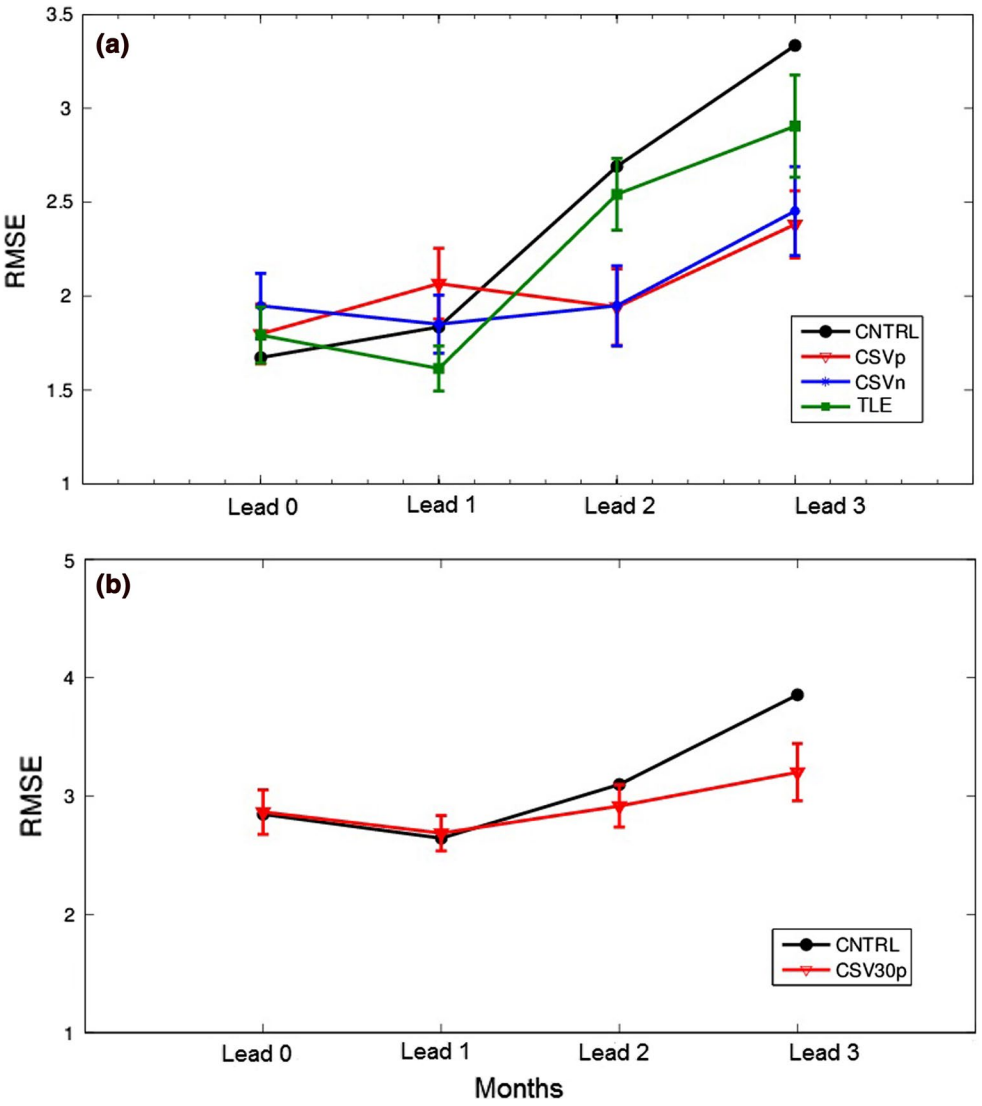


Fig. 15 Root mean square error (RMSE) of the SAM ensemble forecast generated using CSV perturbation. **a** RMSE calculated over 10 years (2000–2009 using observed initial conditions) whereas in **b** RMSE is for 30 years forecast (1980–2009 using CAM4 control initial conditions). *Error bar* are drawn using bootstrap method

non-optimal ensemble predictions. For comparison, we also use the time lag ensemble (TLE) method to construct ensemble predictions. The TLE is achieved using different initial conditions with the lag of 6 h to initialize six forecasts, starting at 1st June of each year.

We construct ensemble predictions by the CSVs extracted over the Indian Oceans. Each individual CSV is multiplied by random noise and is superimposed onto the initial conditions to generate an ensemble forecast. The ensemble size is 20, by using different random numbers. In practice, we use ten random numbers multiplied by a positive CSV and ten random number multiplied by a negative CSV, since the sign of CSVs are arbitrary. The ensemble mean of +ve and −ve ensembles is labeled as CSVp and CSVn in discussions below, respectively.

The ensemble prediction is run for the period from 2000 to 2009 and from 1980 to 2009, respectively. The RMSE values from the period from 2000 to 2009 are shown in Fig. 15a, for a single control run prediction without perturbation (black line), the ensemble of CSVp and CSVn (blue and red line), and ensemble of TLE (green line). For each forecast, error bar are drawn using bootstrap method (Efron and Tibshirani 1993) for its significance. The bootstrap method is obtained as follows. (1) At a given forecast lead month, the observation and forecast are paired together, based on same target month. In this way, the forecast and observation sampled are paired together for the whole period (2000–2009). (2) The 95 % of the paired sample is then chosen randomly and used for RMSE calculation. (3) The step (2) is repeated 1000 times to obtain 1000 RMSE. (4) The standard deviation of these 1000 correlation coefficients at each lead month is drawn as error bars.

As shown in Fig. 15a, the ensemble mean RMSE of CSVp and CSVn forecasts are significantly better than the Control and TLE forecasts at lead times beyond 2 months. The probable reason for the significant difference in the

last 2 months is that the CSV used to construct ensemble prediction is based on the analysis of optimal interval of 4 months. In other words, it is due to the fact that the CSVs used for perturbation are optimized for a 4 month period (i.e. from June to September each year).

To investigate the significance on larger sample size, we also perform 30 year forecasts (CSV30p) from 1980 to 2009 using the CSVs extracted from the IO30 experiment (CSVp only). This forecast is compared against IO30 control forecast and RMSE is shown in Fig. 15b. The bootstrap methodology is repeated for IO30 forecasts, the same as for the 10 year forecasts. The improvement in RMSE for lead time 3 and 4 is significant in this case as well. This ensemble forecast increase the reliability of RMSE results revealed for 10 years forecast (Fig. 15a).

6 Summary and conclusion

In this study we apply a recently developed technique of climatologically relevant singular vector (CSV) to examine the error growth of OLR prediction over South Asian monsoon (SAM) region, at monthly time scales using both CAM4 and the CCSM4 model. Different perturbation domains and optimal growth intervals are chosen to identify the perturbation structures most favorable for the error growth of SAM prediction. The CSVs are computed using both uncoupled and coupled models to investigate the role of coupling in the error growth of SAM prediction. The robustness of CSVs is demonstrated with the help of many sensitivity experiments. The CSVs over Indian Ocean are extracted to generate SAM ensemble forecasts.

It is seen that the CSV resembles a dipole-like structure over the Indian Oceans and the ENSO-like pattern over the Pacific Ocean. The magnitude of error growth (singular value) is different over the Indian Ocean and over the Pacific Ocean. When the CSVs are extracted over Indian Ocean, their growth rates are found to be more consistent with the increase of lead time and generally larger than the counterparts over the Pacific Ocean. Different parameters such as the number of the EOF modes used for initial perturbation and the number of ensemble members are tested to evaluate CSV robustness over Indian Ocean.

Ensemble forecasts constructed using the negative and positive CSVs over the Indian Ocean are compared with forecasts using the time lag ensemble (TLE) method and the single control forecast for the period of 10 year from 2000 to 2009. It is seen that the ensemble forecast generated by CSV perturbations has a more reliable ensemble mean compared to both the TLE mean and the control forecast, and its RMSE is found significantly better than TLE and control forecast at lead times beyond 2 months. To explore consistency and robustness of the hindcast results, we also extend the hindcast to the period of 30 years from 1980 to 2009, and obtained similar conclusions, namely that, the ensemble prediction by the CSV is better than that by TLE and the control forecast, further confirming the merits of CSV method for SAM prediction.

While this study facilitates the investigation of the SAM optimal error growth by using the CSV method, some cautions should be mentioned. First, the main model used in this study is an atmospheric general model (CAM4), which lacks some key dynamics related to the air-sea coupling. The validation is conducted using the coupled model (CCSM4) but seems not in a robust and systematic manner. In the near future, we will use CCSM4 and other models to systematically conduct the analysis of error dynamics using the CSV method. It would be interesting to explore the robustness of the results of this study with other models and important monsoon variables. Nevertheless this work seems the first to explore the optimal error growth of SAM seasonal prediction, and the results reported here offer valuable insight to SAM predictability and have practical significance to ensemble prediction. These findings have implications for the SAM seasonal forecast in both the construction of the ensemble forecast system and the detection of key oceanic areas that impact SAM forecast, where the uncertainties can be reduced by adaptive observations and data assimilation.

Acknowledgments This work is supported by NSERC Discovery Grant to Y. Tang. The authors are thankful to NCAR for providing models and their input data sets as well as providing assistance during simulations.

References

Acharya N, Kar SC, Mohanty UC, Kulkarni MA, Dash SK (2011) Performance of GCMs for seasonal prediction over India—a case study for 2009 monsoon. Theor Appl Climatol 105:505–520

Anderson J, Hoar T, Raeder K, Liu H, Collins N, Torn R, Avellano A (2009) The data assimilation research testbed: a community facility. Bull Am Meteorol Soc 90:1283–1296

Annamalai H, Liu P (2005) Response of the Asian summer monsoon to changes in El Niño properties. Q J R Meteorol Soc 131:805–831

Buizza R, Palmer TN (1995) The singular vector structure of the atmosphere global circulation. J Atmos Sci 52:1434–1456

Charabi Y, Abdul-Wahab SA (2009) Synoptic aspects of the summer monsoon of southern Oman and its global teleconnections. J Geophys Res 114:D07107. doi:10.1029/2008JD010234

Charney JG, Shukla J (1981) Predictability of monsoons. In: Lighthill J, Pearce RP (eds) Monsoon dynamics. Cambridge University Press, pp 99–109

Chen Y-Q, Battisti DS, Palmer TN, Barsugli J, Sarachik ES (1997) A study of the predictability of tropical Pacific SST in a coupled atmosphere–ocean model using singular vector analysis: the role of the annual cycle and the ENSO cycle. Mon Weather Rev 125:831–845

Chen D, Cane MA, Kaplan A, Zebiak SE, Huang D (2004) Predictability of El Niño over the past 148 years. Nature 428:733–736

Cheng Y, Tang Y, Zhou X, Jackson P, Chen D (2010) Further analysis of singular vector and ENSO predictability in the Lamont model—part I: singular vector and the control factors. Clim Dyn 35:807–826. doi:10.1007/s00382-009-0595-7

Chowdary J, Xie S-P, Lee JY, Kosaka Y, Wang B (2010) Predictability of summer Northwest Pacific climate in eleven coupled model hindcasts: local and remote forcing. J Geophys Res 115:D22121. doi:10.1029/2010JD014595

Collins WD et al (2006) The community climate system model version 3 (CCSM3). J Clim 19:2122–2143

Efron B, Tibshirani RJ (1993) An introduction to the bootstrap. Chapman and Hall, London. ISBN 0412042312

Fu X, Wang B, Bao Q, Liu P, Lee J-Y (2009) Impacts of initial conditions on monsoon intraseasonal forecasting. GRL 36:L08801

Fu X, Wang B, Lee J-Y, Wang W, Gao L (2011) Sensitivity of dynamical intraseasonal prediction skills to different initial conditions. Mon Weather Rev 139:2572–2592

Gadgil S, Rajeevan M, Nanjundiah R (2005) Monsoon prediction—Why yet another failure? Curr Sci 88:1389–1400

Gent P, Danabasoglu G, Donner L, Holland M, Hunke E, Jayne S, Lawrence D, Neale R, Rasch P, Vertenstein M, Worley P, Yang ZL, Zhang M (2011) The community climate system model version 4. J Clim 24:4973–4991

Griffies SM, Winton M, Samuels B, Danabasoglu G, Yeager S, Marsland S, Drange H, Bentsen M (2012) Datasets and protocol for the CLIVAR WGOMD coordinated Ocean-sea ice reference experiments (COREs), WCRP report no. 21/2012, p 21

Hawkins ED, Sutton R (2010) Estimating climatically relevant singular vectors for decadal predictions of the Atlantic Ocean. J Clim 24:109–123. doi:10.1175/2010JCLI3579.1

Houtekamer PL, Mitchell Herschel L, Pellerin Gérard, Buehner Mark, Charron Martin, Spacek Lubos, Hansen Bjarne (2005) Atmospheric data assimilation with an ensemble Kalman filter: results with real observations. Mon Weather Rev 133:604–620. doi:10.1175/MWR-2864.1

Hunke EC, Lipscomb WH (2008) CICE: the Los Alamos sea ice model user's manual, version 4. Los Alamos National Laboratory Tech. Report LA-CC-06-012

Islam S, Tang Y, Jackson P (2013) Asian monsoon simulations by community climate models CAM4 and CCSM4. Clim Dyn. doi:10.1007/s00382-013-1752-6

Jiang X, Yang S, Li Y, Kumar A, Liu X, Zuo Z, Jha B (2013) Seasonal-to-interannual prediction of the Asian summer monsoon in the NCEP Climate Forecast System version 2. J Clim 26:3708–3727. doi:10.1175/JCLI-D-12-00437.1

Jin EK, Kinter JL, Wang B, Park CK, Kang IS, Kirtman BP, Kug JS, Kumar A, Luo JJ, Schemm J, Shukla J, Yamagata T (2008) Current status of ENSO prediction skill in coupled ocean-atmosphere models. Clim Dyn 31(6):647–664. doi:10.1007/s00382-008-0397-3

Kleeman R, Tang Y, Moore AM (2003) The calculation of climatically relevant singular vectors in the presence of weather noise as applied to the ENSO problem. J Atmos Sci 60:2856–2868

Kosaka Y, Chowdary JS, Xie S-P, Min Y-M, Lee J-Y (2012) Limitations of seasonal predictability for summer climate over East Asia and the Northwestern Pacific. J Clim 25:7574–7589

Kousky VE, Kayano MT (1994) Principal modes of outgoing longwave radiation and 250 mb circulation for the South American sector. J Clim 7:1131–1143

Kulkarni MA, Acharya N, Kar SC, Mohanty UC, Tippett MK, Robertson AW, Luo JJ, Yamagata Toshio (2011) Probabilistic prediction of indian summer monsoon rainfall using global climate models. Theor Appl Climatol 107(3–4):441–450. doi:10.1007/s00704-011-0493-x

Leith CE (1974) Theoretical skill of Montecarlo forecasts. Mon Weather Rev 102:409–418

Lorenz EN (1963) Deterministic nonperiodic flow. J Atmos Sci 20(2):130–141

Meehl GA, Arblaster JM (2002) Indian monsoon GCM sensitivity experiments testing tropospheric biennial oscillation transition conditions. J Clim 15:923–944

Molteni F, Palmer TN (1993) Predictability and finite time instability of the northern winter circulation. Q J R Meteorol Soc 119:269–298

Moore AM, Kleeman R (1996) The dynamics of error growth and predictability in a coupled model of ENSO. Q J R Meteorol Soc 122:1405–1446

Moore AM, Mariano AJ (1999) The dynamics of error growth and predictability in a model of the Gulf Stream. Part I: Singular vector analysis. J Phys Oceanogr 29:158–176

Moron V (1995) Variability of the African convection centre as viewed by outgoing longwave radiation records and relationships with sea surface temperature patterns. Int J Climatol 15:25–34

Neale RB, Jochum M, Richter JH (2008) The impact of convection on ENSO: from a delayed oscillator to a series of events. J Clim 21:5904–5924

Neale RB, Richter JH, Andrew JC, Park S, Lauritzen PH, Gettelman A, Williamson DL, Rasch PJ, Vavrus SJ, Taylor MA, Collins WD, Zhang M, Lin S (2010) Description of the NCAR community atmosphere model (CAM 4.0). NCAR Technical Note, NCAR/TN-485+STR. National Center for Atmospheric Research, Boulder, Colorado

Nigam S (1994) On the dynamical basis for the Asian summer monsoon rainfall-El Niño relationship. J Clim 7:1750–1771

Palmer TN (2000) Predicting uncertainty in forecasts of weather and climate. Rep Prog Phys 63:71–116

Palmer TN, Zanna L (2013) Singular vectors, predictability and ensemble forecasting for weather and climate. J Phys A Math Theor 46:254018. doi:10.1088/1751-8113/46/25/254018

Pokhrel S, Chaudhari HS, Saha SK, Dhakata A, Yadav RK, Salenke K, Mahaptra S, Rao SA (2012) ENSO. Clim Dyn, IOD and Indian summer monsoon in the NCEP climate forecast system. doi:10.1007/s00382-012-1349-5

Rayner NA, Parker DE, Horton EB, Folland CK, Alexander LV, Rowell DP, Kent EC, Kaplan A (2003) Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century. J Geophys Res 108:4407. doi:10.1029/2002JD002670

Richter J, Rasch P (2008) Effects of convective momentum transport on the atmospheric circulation in the Community Atmosphere Model, version 3. J Clim 21:1487–1499

Saji NH, Goswami BN, Vinayachandran PN, Yamagata T (1999) A dipole mode in the tropical Indian Ocean. Nature 401:360–363

Schmetz J, Liu Q (1988) Outgoing longwave radiation and its diurnal variation at regional scales derived from Meteosat. J Geophys Res 93(D9):11192–11204

Shukla J (1981) Dynamical predictability of monthly means. J Atmos Sci 38:2547–2572

Shukla J (1998) Predictability in the midst of chaos: a scientific basis for climate forecasting. Science 282:728–731

Shukla J, Paolino DA (1983) The southern oscillation and long-range forecasting of the summer monsoon rainfall over India. Mon Weather Rev 111:1830–1837

Sikka (1980) Some aspects of the large scale fluctuations of summer monsoon rainfall over India in relation to fluctuations in the planetary and regional scale circulation parameters. Proc Indian Acad Sci (E&P Sci) 89(2):179–195

Singh A, Acharya N, Mohanty UC, Robertson AW, Mishra G (2012) On the predictability of Indian Summer Monsoon Rainfall in general circulation model at different lead time. Dyn Atmos Oceans 58:108–127

Slingo JM, Annamalai H (2000) 1997: the El Niño of the century and the response of the Indian summer monsoon. Mon Weather Rev 128:1778–1797

Smith RD, Gent P (2002) Reference manual for the parallel ocean program (POP). Los Alamos unclassified report LA-UR-02-2484

Sohn SJ, Tam CY, Ashok K, Ahn JB (2012) Quantifying the reliability of precipitation datasets for monitoring large-scale East Asian precipitation variations. Int J Climatol. doi:10.1002/joc.2380

Tang Y, Deng Z (2011) ENSO bred vectors and predictability in a hybrid coupled model from 1881–2000. J Clim 24(1):298–314

Tang Y, Kleeman R, Moore A (2005) On the reliability of ENSO dynamical predictions. J Atmos Sci 62(6):1770–1791

Tang Y, Kleeman R, Miller S (2006) ENSO predictability of a fully coupled GCM model using singular vector analysis. J Clim 19:3361–3377

Toth Z, Kalnay E (1993) Ensemble forecasting and NMC: the generation of perturbations. Bull Am Meteorol Soc 74:2317–2330

Toth Z, Kalnay E (1997) Ensemble forecasting at NCEP and the breeding method. Mon Weather Rev 125:3297–3319

Vavrus S, Waliser D (2008) An improved parameterization for simulating Arctic cloud amount in the CCSM3 climate model. J Clim 21:5673–5687

Wang B et al (2009) Advance and prospectus of seasonal prediction: assessment of the APCC/CliPAS 14-model ensemble retrospective seasonal prediction (1980–2004). Clim Dyn 33:93–117. doi:10.1007/s00382-008-0460-0

Xue Y, Cane MA, Zebiak SE (1997) Predictability of a coupled model of ENSO using singular vector analysis. Part I: optimal growth in seasonal background and ENSO cycles. Mon Weather Rev 125:2043–2056

Yang S, Lau K-M (2006) Interannual variability of the Asian monsoon. In: Wang B (ed) the Asian monsoon. Praxis, Moscow, pp 259–293

Zhou X, Tang Y, Deng Z (2007) The impact of nonlinear atmosphere on the fastest error growth of ENSO prediction. Clim Dyn. doi:10.1007/s00382-007-0302-5