Chinese Journal of Oceanology and Limnology
Vol. 30 No. 6, P. 1093-1104, 2012
http://dx.doi.org/10.1007/s00343-012-1271-z

The spring prediction barrier in ENSO hindcast experiments using the FGOALS-g model*

YAN Li (严厉)[1, 2, **], YU Yongqiang (俞永强)[2]

[1] *State Key Laboratory of Tropical Oceanography*, *South China Sea Institute of Oceanology*, *Chinese Academy of Sciences*, *Guangzhou 510301*, *China*

[2] *State Key Laboratory of Atmospheric Sciences and Geophysical Fluid Dynamics*, *Institute of Atmospheric Physics*, *Chinese Academy of Sciences*, *Beijing 100029*, *China*

Received Dec. 26, 2011; accepted in principle Feb. 7, 2012; accepted for publication Mar. 19, 2012

Abstract The Flexible Global Ocean-Atmosphere-Land System Model-gamil (FGOALS-g) was used to study the spring prediction barrier (SPB) in an ensemble system. This coupled model was developed and maintained at the State Key Laboratory of Atmospheric Sciences and Geophysical Fluid Dynamics (LASG). There are two steps in our hindcast experiments. The first is to integrate the coupled model continuously with sea surface temperature (SST) nudging, from 1971 to 2006. The second is to carry out a series of one-year hindcasts without SST nudging, by adopting initial values from the first step on January 1st, April 1st, July 1st, and October 1st, from 1982 to 2005. We generate 10 ensemble members for a particular start date (1st) by choosing different atmospheric and land conditions around the hindcast start date (1st through 10th). To estimate the predicted SST, two methods are used: (1) Anomaly Correlation Coefficient and its rate of decrease; and (2) Talagrand distribution and its standard deviation. Results show that FGOALS-g offers a reliable ensemble system with realistic initial atmospheric and oceanic conditions, and high anomaly correlation (>0.5) within 6 month lead time. Further, the ensemble approach is effective, in that the anomaly correlation of ensemble mean is much higher than that of most individual ensemble members. The SPB exists in the FGOALS-g ensemble system, as shown by anomaly correlation and equal likelihood. Nevertheless, the role of the ensemble mean in reducing the SPB of ENSO prediction is significant. The rate of decrease of the ensemble mean is smaller than the largest deviations by 0.04–0.14. At the same time, the ensemble system "equal likelihood" declines during spring. An ensemble mean helps give a correct prediction direction, departing from largely-deviated ensemble members.

Keyword: spring prediction barrier; ensemble; ENSO; hindcast experiments; equal likelihood

1 INTRODUCTION

El Niño/Southern Oscillation (ENSO) is an important tropical air-sea interaction at the Earth's surface. An interesting issue is that during boreal spring, model predictive skill drops sharply, regardless of the month in which prediction begins. This is called the "spring prediction barrier" (SPB hereafter), first suggested by Webster and Yang (1992). A series of works has proven that SPB exists in models, from simple, dynamical, statistical, or hybrid ones to fully coupled ones (Balmaseda et al., 1994; Barnston et al., 1994; Xue et al., 1994; Scheider et al., 2003; Dewitt, 2005; Jin et al., 2008).

The SPB is an open question in ENSO predictability studies. Some explanations could be merged into one. In boreal spring, ENSO signal is weakest and, consequently, signal-to-noise ratio is lowest (Torrence and Webster, 1998; Clarke and van Gorder, 1999). Another plausible explanation is strong spring ocean-

* Supported by the National Basic Research Program of China (973 Program) (No. 2007CB411806), the Knowledge Innovation Program of Chinese Academy of Sciences (Nos. KZCX2-YW-Q11-02, XDA05090404), the National Natural Science Foundation of China (No. 40975065), and the National High Technology Research and Development Program of China (863 Program) (No. 2010AA012304)
** Corresponding author: yanl@scsio.ac.cn

atmosphere coupled instability (Mu et al., 2007; Duan et al., 2009; Yu et al., 2009). There are other explanations, such as weakest east-west SST gradient and ocean-atmosphere coupling in spring (Webster and Yang, 1992; Webster, 1995).

Recently, Duan et al. (2009) demonstrated that SPB occurrence depends on the particular initial error mode. That is, there is the possibility that certain types of initial error cause the SPB, but other types fail to cause it. Similar results were forwarded by Yu et al. (2009). They found that if there are certain types of initial error in realistic ENSO predictions and if a target method or data assimilation approach can filter them, ENSO forecast ability can be improved. In addition to these intermediate-complexity Zebiak-Cane model results, Wei and Duan (2010) used coupled model (FGOALS-g) outputs for actual El Niño and La Niña events. They found that the SPB tends to be more prominent in El Niño than in La Niña events, and is more significant in the growth phase than in decay phase. Similar results were put forth by Zhang et al. (2012). In addition to FGOALS-g, they used retrospective results of three other models, containing both coupled models (Geophysical Fluid Dynamics Laboratory, GFDL, and University of Hawaii, UH) and a statistical-dynamical model (Seoul National University, SNU). All these results support the viewpoints of Chen et al. (1995, 2004) that SPB may be suppressed through assimilation. In addition, Zheng and Zhu (2010) suggested that reasonable consideration of model errors during the ensemble forecasting process can alleviate the SPB effect.

All these previous studies emphasized the role of accuracy of initial conditions. However, most of these results were obtained by using theoretical El Niño events generated by the Zebiak-Cane model, or ENSO events by another intermediate-complexity model. Therefore, further work to estimate SPB based on coupled models is natural. New studies have used actual ENSO events generated by a coupled FGOALS-g model (Wei and Duan, 2010), or still other climate models (Zhang et al., 2012); and the prediction errors in their study are estimated by the slope of L2 norm. In this article, we also use coupled FGOALS-g model results to investigate the SPB, but the main difference lies in the statistical methods to estimate prediction errors. We apply the Talagrand diagram to evaluate equal likelihood in an ensemble system. Furthermore, the function of ensemble mean is explicitly discussed.

This article presents the role of ensemble mean in actual ENSO predictions. The remainder of this paper is organized as follows: Section 2 describes the model, experimental design and method. Section 3 exhibits the SPB phenomenon of ensemble mean and ensemble members. Finally, main results are concluded and discussed in Section 4.

2 MODEL DESCRIPTION, EXPERIMENT DESIGN AND METHOD

2.1 Flexible Global Ocean Atmosphere Land System-gamil (FGOALS-g)

The coupled model used is the Flexible Global Ocean Atmosphere Land System-gamil (FGOALS-g), version 1.11, which originated from FGCM-1.0 (Yu et al., 2002, 2004, 2007) and was developed at the State Key Laboratory of Numerical Modeling for Atmospheric Sciences and Geophysical Fluid Dynamics (LASG) at the Institute of Atmospheric Physics (IAP). FGOALS-g has been applied to ENSO-related research (Yu et al., 2007; Zheng and Yu, 2007), the Indian Ocean Dipole (IOD) (Yu and Liu, 2004), closure of the Indonesian Seaway (Yu et al., 2003), paleoclimate (Zheng et al., 2008), 20th century global warming (Zhou and Yu, 2006), scenario projection of future climate change (Li et al., 2010, 2011), and others (Zhou et al., 2007; Yu et al., 2008). This model couples atmospheric, oceanic, land, and sea ice component models with the National Center for Atmospheric Research (NCAR) flux coupler. The frequency of coupling is one day for the oceanic model and one hour for the atmospheric, land and sea ice models.

The atmospheric component is a Grid-point Atmospheric Model of IAP/LASG (GAMIL), version 1.1.0. This model is based on a new dynamic core (Wang et al., 2004) and physical parameterizations of the Community Atmospheric Model Version 2 (CAM2) of NCAR (Kiehl et al., 1996), except for a modified Tiedtke convective scheme (Li et al., 2007). The model uses a hybrid horizontal grid, with a 2.8° Gaussian grid between 65.58ºN and 65.58ºS, and a weighted, even-area grid elsewhere (Wang et al., 2004). Vertically, there are 26 σ-layers from the surface to 2.194 hPa.

The oceanic component model is the LASG/IAP Climate system Ocean Model (LICOM), version 1.0 (Zhang et al., 2003; Liu et al., 2004). It has thirty vertical layers, with twelve equal levels in the upper 300 meters. Horizontal resolution was prescribed at

1°×1°. The oceanic model domain covers 75°S to 88°N, and the North Pole is treated as an isolated island. More detailed description of the ocean model can be found in Zhang et al. (2003).

The land component model is known as Community Land Model Version 2 (CLM2) (Bonan et al., 2002). CLM2 uses ground and vegetation datasets derived from satellite data products. There are 10 vertical layers for soil temperature and soil water, with explicit treatment of liquid water and ice, along with a multilayer snowpack and runoff model.

The sea ice component is a thermodynamic-dynamic model known as Community Sea Ice Model version 4 (CSIM4) (Bettege et al., 1996). The model has five layers in the vertical. In this model, thermodynamic procedures take advantage of an energy-conserving treatment (Bitz and Lipscomb, 1999), and dynamic procedures are based on the elastic-viscous-plastic principle (Hunke and Dukowicz, 1997).

2.2 Ensemble hindcast experiment design

There were two steps to our ENSO hindcast experiments. The first was integration of the coupled GCM with sea surface temperature (SST) nudged to observed from 1971 to 2006, to achieve continuous and compatible initial conditions for ocean and atmosphere. The SST nudging approach for generating initial conditions has been used in previous studies (Chen et al., 1997; Oberhuber et al., 1998; Luo et al., 2005), and has been proven to effectively produce realistic thermocline and wind stress in the equatorial Pacific (Luo et al., 2005). The time scale for the SST nudging term (also called SST-relaxed or SST-restored) is two days. For the period January 1971 to October 1981, SSTs were strongly nudged toward monthly Global sea-Ice and Sea Surface Temperature (GISST) observations (Rayner et al., 2006). For the period October 1981 to December 2005, SSTs were strongly nudged toward weekly National Oceanic and Atmospheric Administration (NOAA) Optimum Interpolation Sea Surface Temperature (OISST) observations (Reynolds et al., 2002). Both GISST and OISST were linearly interpolated to a daily mean before the SST nudging experiment. Results from the first step were a set of balanced oceanic, atmospheric, land, ice, and coupler initial conditions.

The second step was to use the initial values from the SST nudging experiment on January 1st, April 1st, July 1st, and October 1st, from 1982 to 2005. In addition, there were 10 ensemble members for each start date. Ensemble members were generated by choosing different atmospheric and land conditions around the hindcast start date, to describe stochastic information. Meanwhile, oceanic and sea ice initial conditions were kept unchanged, to describe slowly varying lower boundaries. The coupled GCM was integrated for one year without SST nudging. In total, there were 960 hindcast experiments executed in the second step; each was integrated for 12 months.

2.3 Method

2.3.1 ACC score and rate of decrease

ACC (Anomaly Correlation Coefficient) score has been widely used for assessing model ability to predict ENSO (e.g., Zhou et al., 1998; Palmer et al., 2004; Luo et al., 2005). Using this score, we can evaluate correlation between predicted and observed SST anomalies. To compute ACC, we use the following:

$$\mathrm{ACC}=\frac{\sum_{i=1}^{24} P_i' O_i'}{\sqrt{\sum_{i=1}^{24} P_i'^2}\sqrt{\sum_{i=1}^{24} O_i'}}, \tag{1}$$

where i represents the time index and P_i' (O_i') the predicted (observed) SST anomalies. ACC of the Niño 3.4 SST index is the average ACC over the Niño 3.4 area. For each lead month and start season, there are 24 time points, corresponding to 24 total years (from 1982 to 2005). For each start season, we subtract a particular season's climatology to get an anomalous series. Observed SSTs are from monthly NOAA Climate Diagnostic Center (CDC) data (Reynolds et al., 2002).

To evaluate the ACC score decrease during spring, we introduce the rate of decrease. It is defined by the following:

$$D_{\_\mathrm{rate}} = -(\mathrm{ACC(June)}-\mathrm{ACC(April)}). \tag{2}$$

2.3.2 Talagrand diagram and its standard deviation

The Talagrand diagram is, in fact, a frequency distribution. It is an important evaluation standard for equal likelihood. If equal likelihood is adequate, then all member forecasts have a consistent tendency, and prediction by the ensemble system is more reliable (Kalnay, 2003).

The Talagrand diagram algorithm was first described by Talagrand (1997). He showed that equal likelihood is perfect when hindcast results and verifying observations are sampled from the same probability distribution (e.g., Buizza and Palmer,

1998; Hamil, 2001). Talagrand diagrams are generated by computing the frequency of observation within intervals spaced by ensemble members, from smallest to largest. The process for attaining the Talagrand distribution can be described as follows.

Suppose the total number of ensemble members is N (N=10), and in the proof-test area there are K grids, covering periods involving T time points. At each grid point j (j = 1, 2, 3, … , K) at a certain time t (t = 1, 2, 3, … , T), the ensemble member forecast value is denoted $x^f_{i,j,t}$ and verifying observation $x^o_{j,t}$, where superscript f represents "forecast," o represents "observation," and subscript i (i = 1, 2, 3, … , N) represents each ensemble member. With regard to a fixed grid and fixed time, hindcast values of 10 ensemble members are ordered from smallest to largest,

$$x^f_{1,j,t} \le x^f_{2,j,t} \le x^f_{3,j,t} \le \cdots \le x^f_{N,j,t}$$

$$d_{i+1} = x^f_{i+1,j,t} - x^f_{i,j,t} \ge 0$$

$$d_1 : < x^f_{1,j,t}$$

$$d_{N+1} : > x^f_{N,j,t} \quad .$$

$$i = 1, 2, 3, \ldots, N-1 \qquad (3)$$

Then, according to Talagrand, the observation value must fall into one of intervals d_i (i = 1, 2, 3, … , N+1). There are N members, with effective sample M=K×T; suppose that the time the verifying observation falls into a certain interval d_i is S_i, and its expected value is $E = M \times N/(N+1)$. Then, the probability distribution is $P_i = S_i/E$, so $\sum_{i=1}^{N+1} P_i = \left(\sum_{i=1}^{N+1} \mathrm{S}_i\right) \Big/ E = N+1$; let frequency $F_i = P_i/(N+1)$ to make $\sum_{i=1}^{N+1} F_i = 1$. Through the F_i, we obtain the Talagrand diagram.

The flatness of frequency distribution represents the level of equal likelihood. This flatness can be evaluated by standard deviation. We calculate the standard deviation D_{EVI} of F_i^{nor} with category index between 2 and 10 (i=2–10), represented by

$$D_{\mathrm{EVI}} = \sqrt{\frac{1}{N-2}\sum_{i=2}^{N-1}(F_i^{\mathrm{nor}} - 1)^2} \ , \qquad (4)$$

where i represents the category index, N (=11) designates the total category numbers; $F_i^{\mathrm{nor}} = \dfrac{F_i}{\dfrac{1}{N-2}\sum_{i=2}^{N-1} F_i}$ represents normalized frequency; and 1.0 is just the average value of F_i^{nor}.

3 RESULT

3.1 Initial conditions

Thermocline depth variation is a foregleam of ENSO. Thus, for validation of oceanic initial conditions, 20°C isotherm depth anomalies in the equatorial Pacific (2°S–2°N) from both Global Ocean Data Assimilation System (GODAS: Xue et al., 2008) and model results are shown in Fig.1. The model can realistically reproduce eastward propagation of warm subsurface ocean signals during positive ENSO events, albeit with smaller-amplitude thermocline variations. In comparison to the GODAS assimilation, thermocline fluctuations associated with interannual ENSO events over the past 20 years were well simulated. In this SST-nudging experiment, the SSTs influence wind stress through surface heat flux, and wind stress in turn influences thermocline depth via upwelling. Realistic subsurface signal simulation was produced by good simulation of wind stress (Fig.2). Both strong anomalous westerly signals (1982/1983, 1997/1998) and weak ones (1986/1987, 1991/1992, 2002/2003) were captured by the simulation, without magnitude decay. Negative signals related to La Niña events, as well as positive signal propagation, were also well simulated.

As mentioned by Luo et al. (2005), when the AGCM component is forced by such generated SSTs, which are very close to observed values, the model tends to produce realistic wind stress, heat and water fluxes. Furthermore, the OGCM driven by the wind stress tends to produce realistic thermocline variations in the equatorial Pacific. Therefore, the success of the simple coupled SST-nudging scheme for initialization crucially depends on performance of the coupled GCM. The initial condition results show that the FGOALS-g 1.11 coupled model is able to reproduce realistic oceanic memory for ENSO prediction, using the simple SST-nudging scheme. This coupled model can simulate tropical air-sea interactions well, and gives a basis for a credible ensemble system.

3.2 SPB in ENSO ensemble hindcast experiments

Three important characteristics of the ENSO ensemble experiments were demonstrated through the ACC score (Fig.3): (1) Performance of the coupled model is credible. Regardless of ensemble mean or ensemble members, most predictive skills (ACC scores) were higher than 0.5 within 6-month lead time. This means that the model produced useful predictions (ACC scores ≥0.6) or near useful. Further,

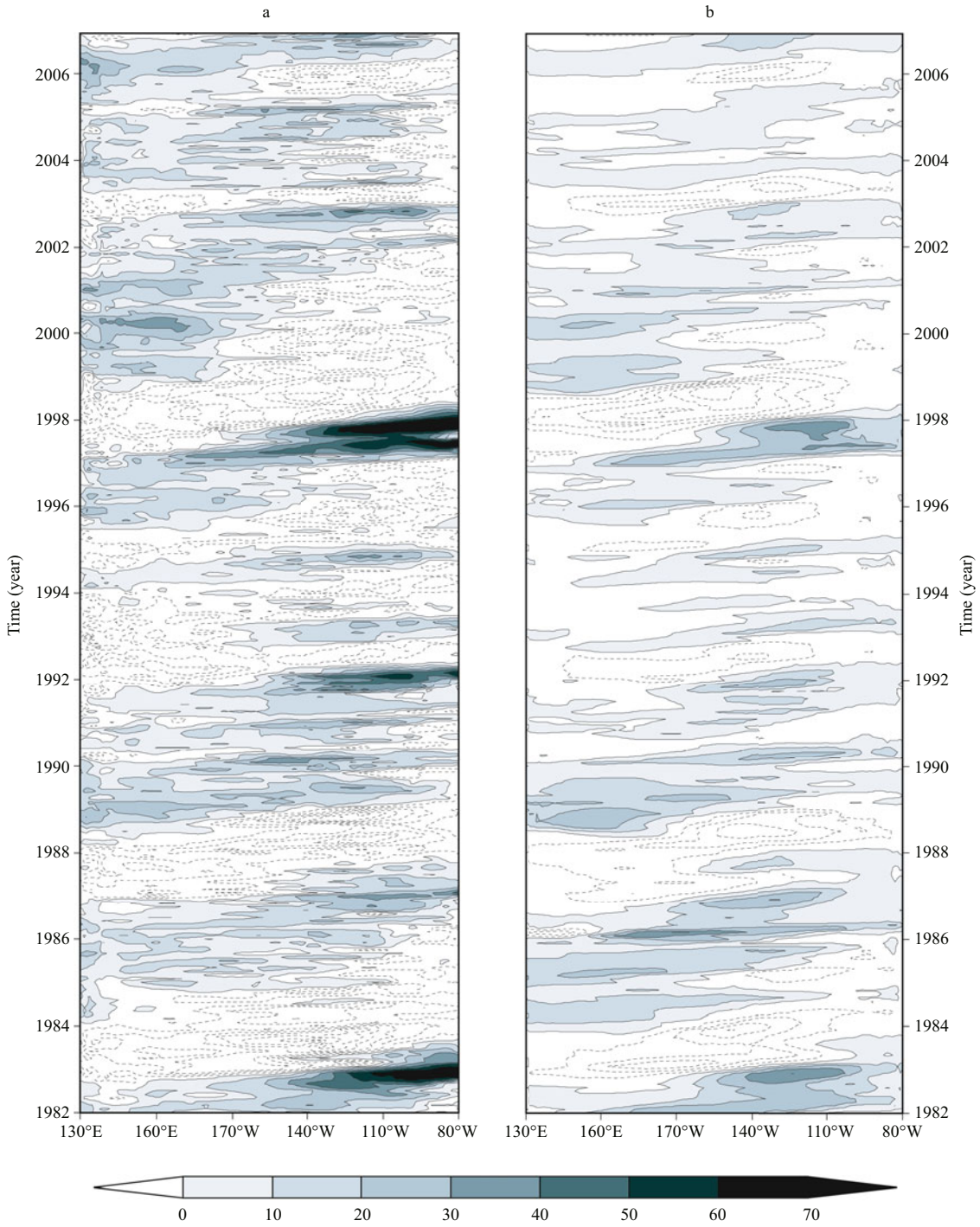

Fig.1 Depth anomalies of 20°C isotherm (contour: 10 m; units: m) along equatorial Pacific (2°S–2°N) from (a) GODAS, and (b) model results, based on SST nudging scheme

Regions with positive values are shaded.

the model is tenable for experiments beginning in different seasons. A model with such skill provides the basis for a perfect ensemble system; (2) The ensemble scheme is effective, because the ensemble mean enhances predictive skill. ACC scores of the ensemble mean were much higher than those of most ensemble members, especially for long lead time. Although the ACC score of the ensemble mean is at times not the highest, it is very close to that of the highest member; (3) After the ensemble, the SPB still exists. For the ensemble mean in boreal spring, predictive ability declines more rapidly than in other

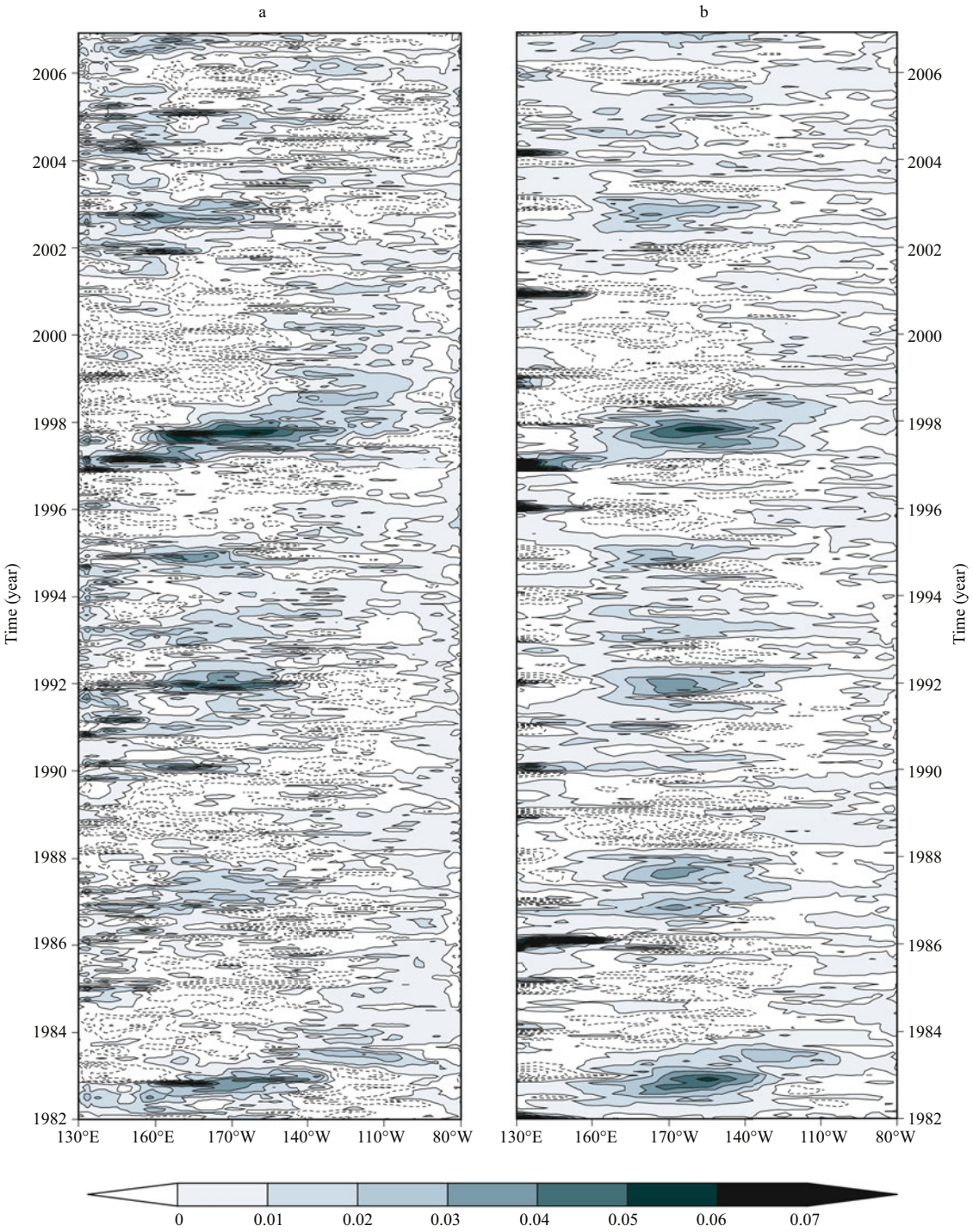

Fig.2 Zonal wind stress anomalies (contour: 0.01 N/m^2; units: N/m^2) in the equatorial Pacific (2°S–2°N) from (a) GODAS, and (b) model results, based on SST-nudging scheme

Regions with positive values are shaded.

seasons. In addition, it is tenable for experiments starting in all 4 seasons. This suggests that the ensemble mean can remove stochastic errors for improvement of predictive skill. Nevertheless, the ensemble cannot remove the SPB. However, the ensemble method is still functional—its function lies in avoiding the largest deviated prediction. If we adopt certain ensemble members (such as member 01) in place of the ensemble mean, then in the experiments beginning in April (Fig.3b), the ACC scores are very low with lead time beyond 6 months (lower than the ensemble mean by about 0.2).

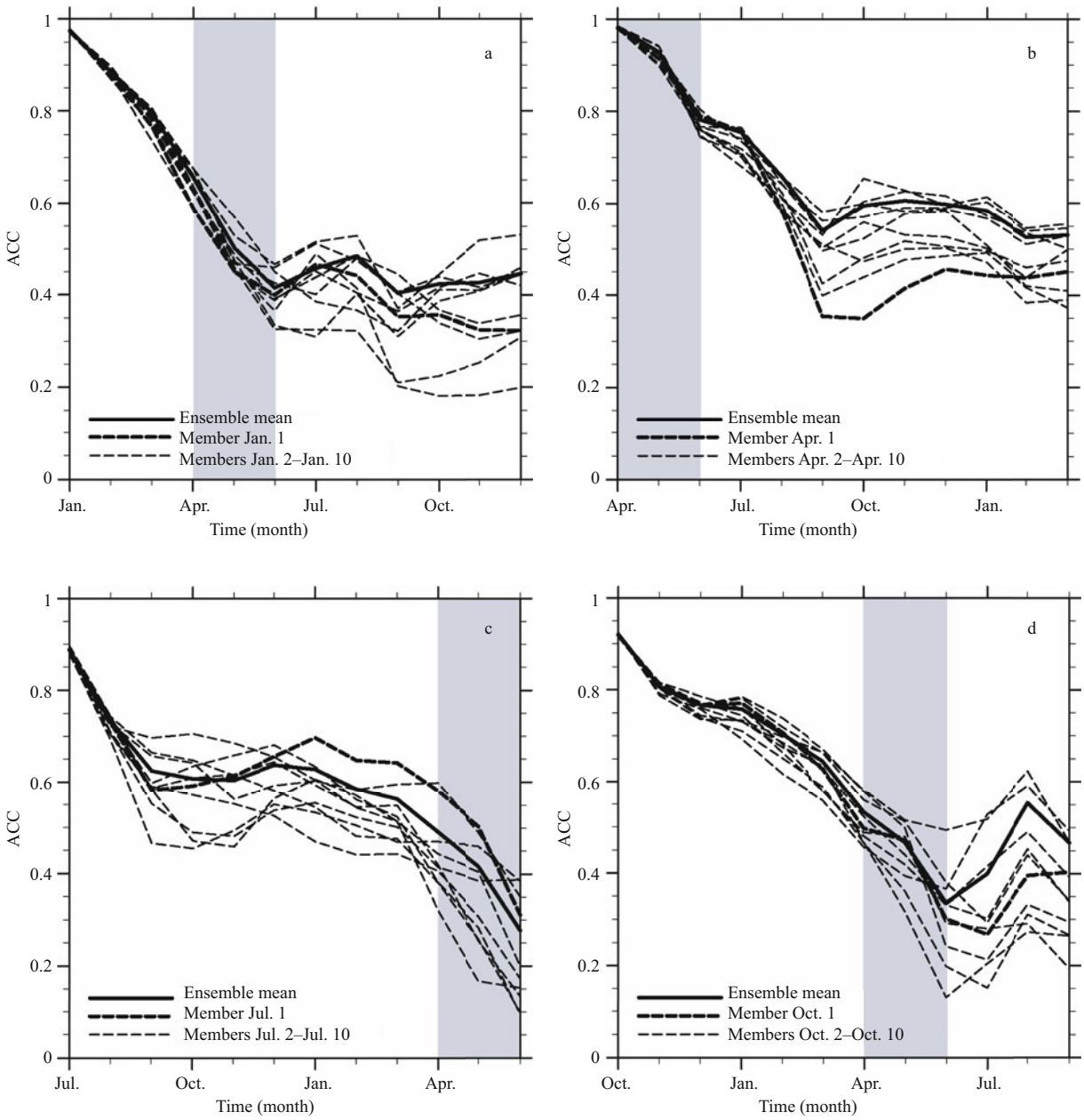

Fig.3 ACC (*y*-axis) as functions of lead time for Niño 3.4 anomalies, for model ensemble mean and for ten member hindcasts starting in January (a), April (b), July (c), October (d)

To quantify SPB in ENSO ensemble hindcast experiments, we computed the difference in ACC scores between June and April (i.e., the rate of decrease). Table 1 shows the following: (1) The SPB exists after the ensemble. Every "enb" (ensemble mean) exceeds 0.20 during spring, regardless of the month hindcast experiments start. The "enb" values were 0.24, 0.20, 0.21, and 0.20. In all four-group experiments, such sharp decreases never occurred during other seasons; (2) Adopting the ensemble method helps avoid the largest SPB. First, the largest

Table 1 Rate of decrease and better size

	Hindcast experiments start in				Ave.
	January	April	July	October	
Biggest_D_rate	0.34	0.24	0.29	0.34	0.30
Enb_D_rate	0.24	0.20	0.21	0.20	0.21
Better_size	3	7	7	3	5

In the left column are: Maximum difference of ensemble member, denoted by "biggest_D_rate"; ensemble mean, denoted by "enb_D_rate"; and total size for which ensemble mean is better than ensemble member, denoted by "better_size." "Ave" denotes average value of four group hindcast experiments.

rates of decrease were 0.34, 0.24, 0.29, and 0.34. They are larger than corresponding "enb" by 0.10, 0.04, 0.08, and 0.14, respectively. On average, the extent ensemble mean (0.21) was better than the largest (0.30) by 0.09. The improvement is remarkable. Second, compared with all 10 ensemble members in each experiment group, "enb" was better than 3, 7, 7, 3 ensemble members in total, respectively. On average, the ensemble mean performed better than 5 ensemble members. Until now, we could not select the most accurate ensemble member. The ensemble mean is the best choice, which is tenable for reducing the SPB.

3.3 Equal likelihood SPB in ENSO ensemble hindcast experiments

For the full ensemble with 10 (N=10) members, there are 11 intervals, and the value of the verifying observation then falls into one of the 11 categories. The Talagrand diagram for the Niño 3.4 SST anomalies (Fig.4) show that the frequency is a function of category index. This suggests that: (1) Regardless of when hindcast experiments start in the a–d series, the distribution for SST anomalies is even, although two extreme categories are significantly higher than adjacent ones. Such frequency distributions indicate that the probability distribution of observations is well represented by our ensemble approach; (2) Regardless of when hindcast experiments start, for the target month of June (Fig.4b, 4e, 4l, 4o), the probability distribution is more uneven than in other months. Since June is the end of spring, we suggest that the equal likelihood drops sharply during boreal spring (Fig.4).

Differences among ensemble members increase sharply from the beginning of spring (March) to its

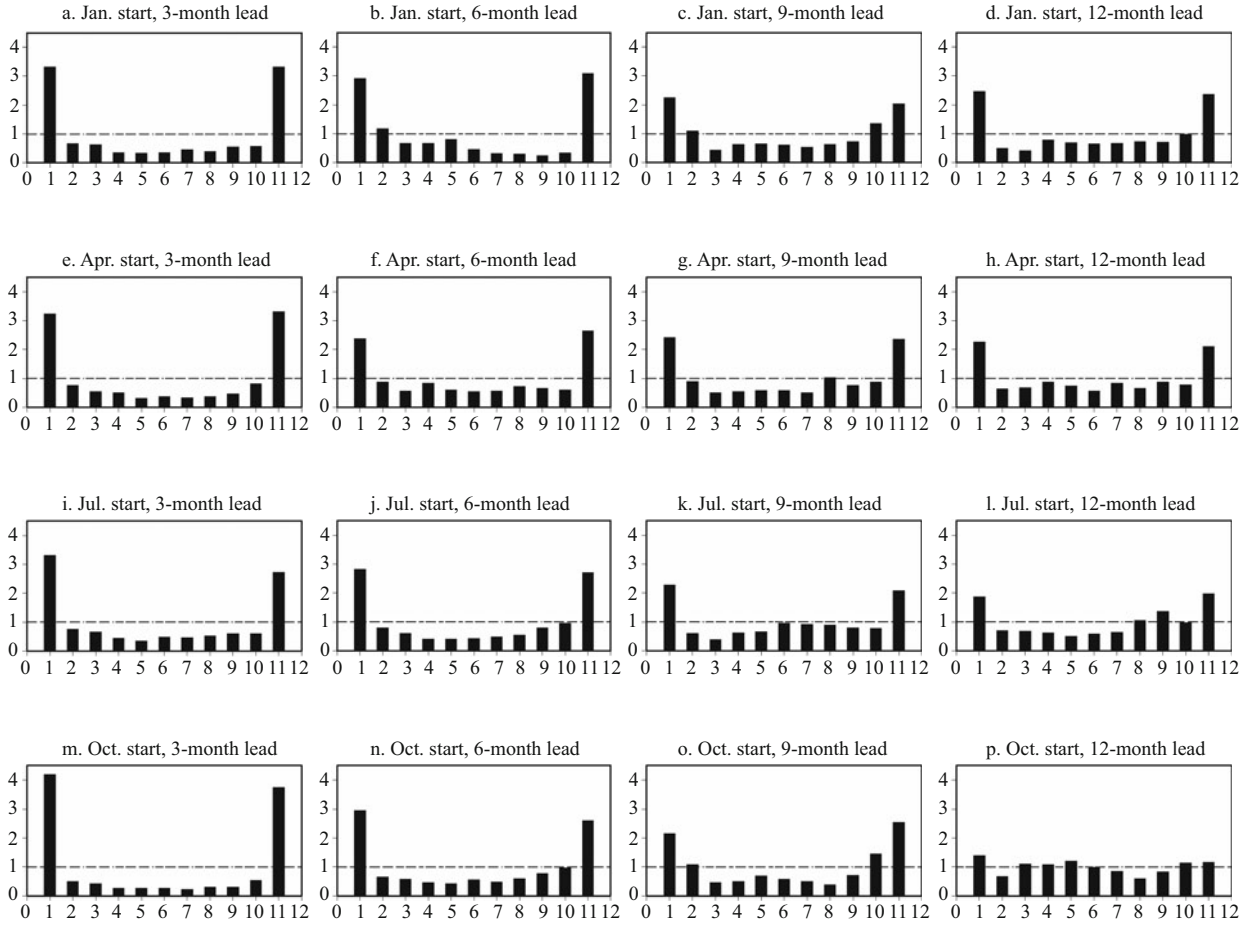

Fig.4 Talagrand diagram, showing frequencies of verifying observation within 11 categories, defined by 10 ordered ensemble members at each grid point for Niño 3.4 anomalies of ensemble hindcasts, starting in January (a–d), April (e–h), July (i–l), October (m–p), with 3, 6, 9, and 12 month lead times, respectively

Dashed line marks theoretical frequency (1/(N+1)), where N is number of ensemble members for a perfectly reliable ensemble prediction system.

end (near June). It is obvious (from Fig.5) that standard deviation increases sharply from March through June (Fig.5b, 5d) or near June (Fig.5a, 5c), regardless of prediction start month. Moreover, the standard deviations peak in June or thereabouts. This is quantitative evidence for the increase of standard deviation during spring, so that consistency of prediction of ensemble members decreases. Therefore, at least some ensemble members produce a prediction that deviates considerably from reality. The standard

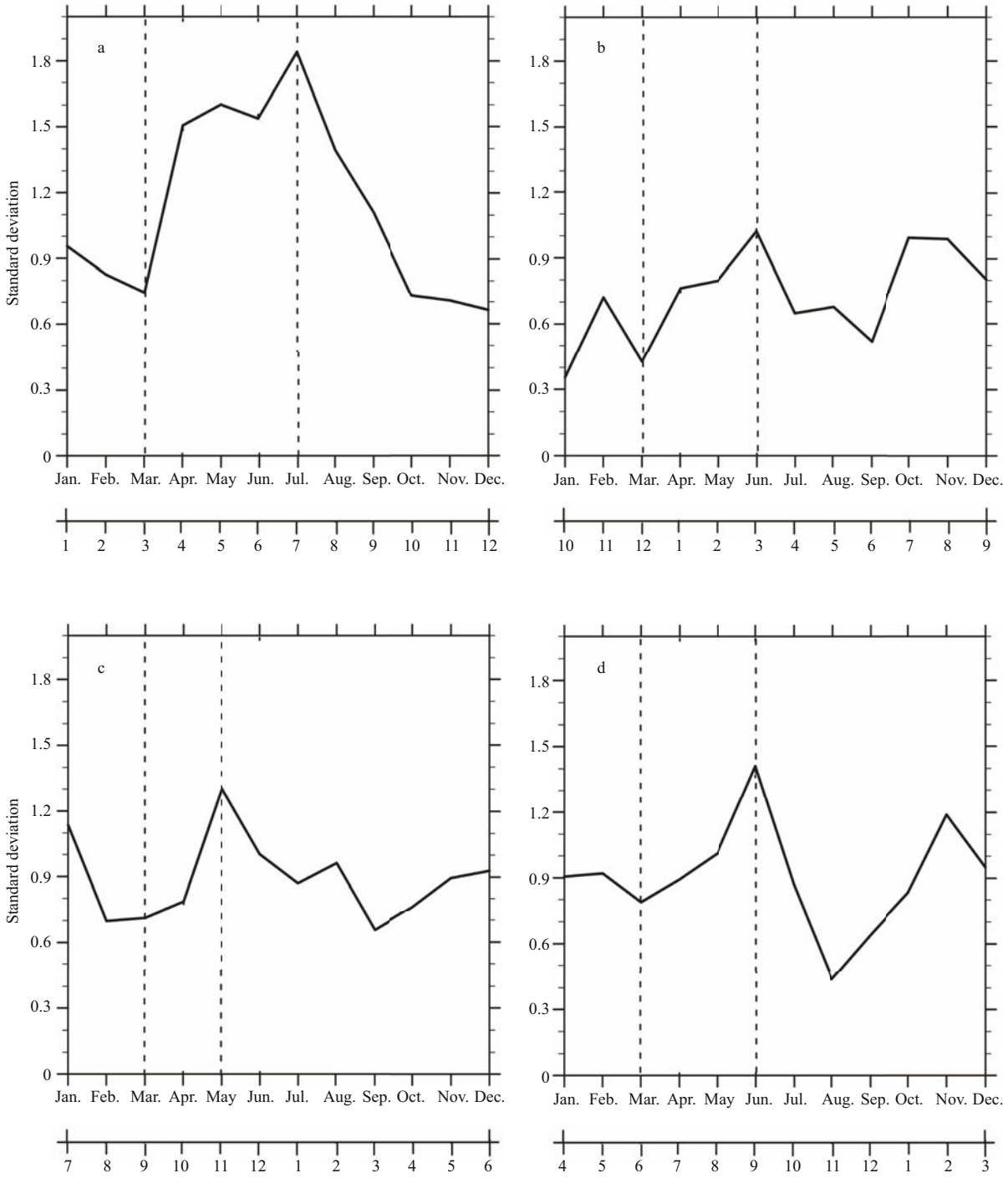

Fig.5 "Standard deviation" (*y*-axis) of Talagrand distribution, as function of calendar time/lead month. For hindcasts starting in January (a), April (b), July (c), and October (d)

Upper *x*-axis denotes target month, and lower *x*-axis lead month. Dashed lines show March and May/June/July.

deviation also ascends sharply from September to November (Fig.5b, 5d), but does not increase in this period with other start months (Fig.5a, 5c). This phenomenon is therefore not as common as the SPB, and there may be other factors causing this ascent in autumn.

4 DISCUSSION AND CONCLUSION

We discussed SPB in a FGOALS-g ensemble system from the standpoint of anomaly correlation and equal likelihood. The main results are as follows:

Coupled model FGOALS-g offers a reliable ensemble system. It can reproduce realistic atmospheric and oceanic initial conditions for seasonal prediction experiments, using the SST-nudging approach (Figs.1, 2). The model also shows a high Anomaly Correlation Coefficient (ACC) score (>0.5) within 6-month lead time for SST in the Niño 3.4 region (Fig.3). In addition, FGOALS predicts ENSO with a 6–9 month lead time, consistent with most state-of-the-art models (Yan et al., 2009). Furthermore, the ensemble approach is effective. The ACC score of the ensemble mean is much higher than that of most individual ensemble members, especially for long lead time (Fig.3). Although the ACC score of the ensemble mean is sometimes not the highest, it is very close to that of the highest member. The ensemble mean filters out uncertain elements (ensemble members differ from each other) to retain the consistent tendency portion of all members. In addition, equal likelihood is well represented by our ensemble approach, because the Talagrand distribution is substantially flat (Fig.4).

The SPB exists in the FGOALS-g ensemble forecast system, from the viewpoint of anomaly correlation and equal likelihood. From the perspective of anomaly correlation, the SPB exists for both ensemble members and ensemble mean (Fig.3). During boreal spring, predictive skill declines very rapidly for all groups of hindcast experiments. During other target seasons, this skill decreases rapidly for only one or two particular group hindcast experiments.

From the standpoint of equal likelihood, an obvious decline during spring was observed (Figs.4, 5). This means that during spring, the difference among ensemble members increases. If the difference becomes sufficiently large, it implies that some ensemble members make deviated predictions. The standard deviation of Talagrand distribution frequencies increases during spring, and peaks at the end of this season (Fig.5). This figure shows a quantitative increase of difference among members (i.e., a reduction of equal likelihood).

Nevertheless, the ensemble method can reduce the SPB (Table 1). Here, “reduce” means avoiding the largest deviated prediction. On average, the ensemble mean improves ACC scores by 0.09 during spring. Thus, the ensemble mean can greatly diminish the SPB, from the viewpoint that one never knows in advance the most accurate initial condition. With use of the most inaccurate initial condition as one member to make a prediction, the SPB will be more serious. In contrast, the ensemble mean can provide a correct tendency.

The ensemble mean effectively reduces the SPB by improving predictive skill. Since the SPB is very evident in the ensemble forecast system, the role of ensemble mean is significant. Further, this mean is helpful in providing a correct prediction direction, avoiding deviated ensemble members. Therefore, it is very important in ENSO predictions. To further reduce the SPB, adding the total number of ensemble members can be attempted. Although this has been done using an intermediate coupled model (Zheng et al., 2009), it requires further investigation based on global ocean-atmosphere models.

References

Balmaseda M A, Anderson D L T, Davey M K. 1994. ENSO prediction using a dynamical ocean model coupled to statistical atmospheres. *Tellus*, **46A**: 497-511.

Barnston A G, Dool V D, Rodenhuis H M et al. 1994. Long-lead seasonal forecasts—where do we stand? *Bull*. *Amer*. *Meteor*. *Soc*., **75**: 2 097-2 114.

Bettge T W, Weatherly J W, Washington W M et al. 1996. “The CSM Sea Ice Model.” NCAR Technical Note NCAR/TN-425+STR. National Center for Atmospheric Research, Boulder, Colorado.

Biz C M, Lipscomb W H. 1999. An energy-conserving thermodynamic model of sea ice. *J. Geophy*. *Res*., **104**: 15 669-15 677.

Bonan G B, Oleson K W, Vertenstein M et al. 2002. The land surface climatology of the community land model coupled to the NCAR community climate model. *J. Climate*, **15**: 3 123-3 149.

Buizza R, Palmer T N. 1998. Impact of ensemble size on ensemble prediction. *Mon*. *Wea*. *Rev*., **129**: 550-560.

Chen D, Cane M A, Kaplan A et al. 2004. Predictability of El Niño over the past 148 years. *Nature*, **428**: 733-736.

Chen D, Zebiak S E, Busalacchi A J et al. 1995. An improved procedure for El Niño forecasting: implications for predictability. *Science*, **269**: 1 699-1 702.

Chen D, Zebiak S E, Cane M A et al. 1997. Initialization and predictability of a coupled ENSO forecast model. *Mon. Wea. Rev*., **125**: 773-788.

Clarke A J, van Gorder S. 1999. The connection between the boreal spring Southern Oscillation persistence barrier and biennial variability. *J. Climate*, **12**: 610-620.

DeWitt D G. 2005. Retrospective Forecasts of Interannual sea surface temperature anomalies from 1982 to present using a directly coupled atmospheric-ocean general circulation model. *Mon. Wea. Rev.*, **133**: 2 972-2 995.

Duan W, Liu X, Zhu K et al. 2009. Exploring the initial errors that cause a significant "spring predictability barrier" for El Niño events, *J. Geophys. Res.*, **114**: C04022, http://dx.doi.org/10.1029/2008JC004925.

Hamill T M. 2001. Interpretation of rank histograms for verifying ensemble forecasts. *Mon. Wea. Rev.*, **129**: 550-560.

Hunke E C, Dukowicz J K. 1997. An elastic-viscous-plastic model for sea ice dynamics. *J. Phys. Oceanogr.*, **27**: 1 849-1 867.

Jin E K, Kinter J L, Wang B et al. 2008. Current status of ENSO prediction skill in coupled ocean-atmosphere models. *Clim. Dyn.*, **31**: 647-664.

Kalnay E. 2003. Atmospheric modeling, data assimilation and predictability. Cambridge University Press. 341p.

Kiehl J T, Hack J, Bonan G et al. 1996. Description of the NCAR Community Climate Model (CCM3), Technical Report NCAR/TN-420+STR. National Center for Atmospheric Research, Boulder, Colorado, USA. 152p.

Li L, Wang B, Wang Y et al. 2007. Improvements in climate simulation with modifications to the Tiedtke convective parameterization in the grid-point atmospheric model of IAP LASG (GAMIL). *Adv. Atmos. Sci.*, **24**: 323-335.

Li H M, Feng L, Zhou T J. 2010. Multi-model projection of July-August climate extreme changes over China under CO_2 doubling. Part II: temperature. *Adv. Atmos. Sci.*, **28**: 448-463.

Li H M, Feng L, Zhou T J. 2011. Multi-model projection of July-August climate extreme changes over China under CO_2 doubling. Part I: precipitation. *Adv. Atmos. Sci.*, **28**: 433-447.

Liu H L, Zhang X H, Li W et al. 2004. An eddy-permitting oceanic general circulation model and its preliminary evaluation. *Adv. Atmos. Sci.*, **21**: 675-690.

Luo J J, Masson S, Behera S K et al. 2005. Seasonal climate predictability in a coupled OAGCM using a different approach for ensemble forecasts. *J. Climate*, **18**: 4 474-4 497.

Mu M, Duan W S, Wang B. 2007. Season-dependent dynamics of nonlinear optimal error growth and El Niño-Southern Oscillation predictability in a theoretical model. *J. Geophys. Res.*, **112**: D10113, http://dx.doi.org/10.1029/2005JD006981.

Oberhuber J M, Christoph E R, Esch M et al. 1998. Predicting the '97 El Niño event with a global climate model. *Geophys. Res. Lett.*, **25**: 2 273-2 276.

Palmer T N, Alessandri A, Andersen U et al. 2004. Development of a European multimodel ensemble system for seasonal-to-interannual prediction (DEMETER). *Bull. Amer. Meteor. Soc.*, **85**: 853-872.

Rayner N A, Parker D E, Horton E B et al. 2006. UKMO - GISST/MOHMATN4/MOHSST6 - Global Ice coverage and SST (1856-2006), [Internet]. UK Meteorological Office, available from http://badc.nerc.ac.uk/data/gisst/.

Reynolds R W, Rayner N A, Smith T M et al. 2002. An improved in situ and satellite SST analysis for climate. *J. Climate*, **15**: 1 609-1 625.

Schneider E K, DeWitt D G, Rosati A et al. 2003. Retrospective ENSO forecasts: sensitivity to atmospheric model and ocean resolution. *Mon. Wea. Rev.*, **131**: 3 038-3 060.

Talagrand O, Vautard R, Strauss B. 1997. Evaluation of probabilistic prediction systems. ECMWF Workshop on Predictability.

Torrence C, Webster P J. 1998. The annual cycle of persistence in the El Niño/Southern Oscillation. *Q. J. R. Meteorol. Soc.*, **124**: 1 985-2 004.

Wang B, Wan H, Ji Z Z et al. 2004. Design of a new dynamical core for global atmospheric models based on some efficient numerical methods. *Science in China*, **47**: 4-21.

Webster P J. 1995. The annual cycle and the predictability of the tropical coupled ocean-atmosphere system. *Meteor. Atmos. Phys.*, **56**: 33-55.

Webster P J, Yang S. 1992. Monsoon and ENSO: selectively interactive systems. *Q. J. R. Meteor. Soc.*, **118**: 825-877.

Wei C, Duan W S. 2010. The "Spring Predictability Barrier" phenomenon of ENSO predictions generated with the FGOALS-g Model. *Atmos. Oceanic Sci. Lett.*, **3**: 87-92.

Xue Y, Cane M A, Zebiak S E et al. 1994. On the prediction of ENSO: a study with a low order Markov model. *Tellus*, **46A**: 512-528.

Xue Y, Behringer D W, Huang B. 2008. introduction_godas_web.pdf, [Internet]. NOAA Climate Prediction Center, available from http://www.cpc.ncep.noaa.gov/products/GODAS/pl/.

Yan L, Yu Y Q, Wang B et al. 2009. ENSO hindcast experiments using a coupled GCM. *Atmos. Oceanic Sci. Lett.*, **2**: 7-13.

Yu Y Q, Yu R C, Zhang X H et al. 2002. A flexible coupled ocean-atmosphere general circulation model. *Adv. Atmos. Sci.*, **19**: 169-190.

Yu Y Q, Zhou Z Y, Zhang X H. 2003. Impact of the closure of Indonesian seaway on climate: a numerical modeling study. *Chinese Science Bulletin*, **48 Supp.II**: 88-93.

Yu Y Q, Zhang X H, Guo Y F. 2004. Global coupled ocean-atmosphere general circulation models in LASG/IAP. *Adv. Atmos. Sci.*, **21**: 444-455.

Yu Y Q, Liu X Y. 2004. ENSO and Indian Ocean Dipole mode in three coupled GCMs. *Acta Oceanologica Sinica*, **23**: 581-595.

Yu Y Q, Zheng W P, Liu H L et al. 2007. The LASG Coupled Climate System Model FGCM-1.0. *Chinese Journal of Geophysics*, **50**: 1 454-1 465.

Yu Y Q, Zhi H, Wang B et al. 2008. Coupled model simulations of climate changes in the 20th century and beyond. *Adv. Atmos. Sci.*, **25**(4): 641-654.

Yu Y S, Duan W S, Xu H et al. 2009. Dynamics of nonlinear error growth and season-dependent predictability of El Niño events in the Zebiak-Cane model. *Quart. J. Roy. Meteor. Soc.* **135(**645): 2 146-2 160.

Zhang X H, Yu Y Q, Liu H L. 2003. The development and application of the oceanic general circulation models. Part I. The global oceanic general circulation models. *Chinese J. of Atmos*. *Sci*., **27**: 607-617. (in Chinese with English abstract)

Zhang Y L, Yu Y Q, Duan W S. 2012. The spring prediction barrier of ENSO in retrospective prediction experiments from coupled ocean-atmosphere models. *Acta Meteorologica Sinica*, in press. (in Chinese with English abstract)

Zheng F, Zhu J. 2010. Spring predictability barrier of ENSO events from the perspective of an ensemble prediction system. *Global and Planetary Change*, **72**: 108-117.

Zheng F, Zhu J, Wang H et al. 2009. Ensemble hindcasts of ENSO events over the past 120 years using a large number of ensembles. *Adv*. *Atmos*. *Sci*., **26**: 359-372.

Zheng W P, Braconnot P, Guilyardi E et al. 2008. ENSO at 6ka and 21ka from ocean-atmosphere coupled model simulations. *Clim*. *Dyn*., **30**: 745-762.

Zheng W P, Yu Y Q. 2007. ENSO phase-locking in an ocean-atmosphere coupled model FGCM-1.0. *Adv*. *Atmos*. *Sci*., **24**: 833-844.

Zhou G Q, Li X, Zeng Q C. 1998. A coupled ocean-atmosphere general circulation model for ENSO prediction and 1997/1998 ENSO forecast. *Climatic and Environmental Research*, **3**: 349-357. (in Chinese with English abstract)

Zhou T J, Yu R C. 2006. Twentieth century surface air temperature over China and the globe simulated by Coupled Climate Models. *J. Climate*, **19**(22): 5 843-5 858.

Zhou T J, Yu Y Q, Liu H L et al. 2007. Progress in the development and application of Climate Ocean Models and Ocean-atmosphere Coupled Models in China. *Adv*. *Atmos*. *Sci*., **24**(6): 729-738.