Clim Dyn
DOI 10.1007/s00382-009-0582-z

Optimal initial perturbations for El Nino ensemble prediction with ensemble Kalman Þlter

Yoo-Geun Ham · Jong-Seong Kug · In-Sik Kang

Received: 2 June 2008 / Accepted: 14 April 2009

Abstract A method for selecting optimal initial perturbations is developed within the framework of an ensemble Kalman Þlter (EnKF). Among the initial conditions generated by EnKF, ensemble members with fast growing perturbations are selected to optimize the ENSO seasonal forecast skills. Seasonal forecast experiments show that the forecast skills with the selected ensemble members are significantly improved compared with other ensemble members for up to 1-year lead forecasts. In addition, it is found that there is a strong relationship between the forecast skill improvements and ßow-dependent instability. That is, correlation skills are significantly improved over the region where the predictable signal is relatively small (i.e. an inverse relationship). It is also shown that forecast skills are significantly improved during ENSO onset and decay phases, which are the most unpredictable periods among the ENSO events.

Keywords Ensemble Kalman Þlter · Seasonal prediction · Optimal initial perturbation · Ensemble prediction

1 Introduction

During the last few decades, many studies have been devoted to modeling and predicting climate variability, in accordance with increased scientiÞc and economic interest in seasonal climate prediction and predictability. To keep pace with these demands, many climate centers have developed and tried to improve seasonal forecast systems to enhance societyÕs ability to cope with climate variability and therefore reduce its vulnerability by providing useful information. In order to obtain useful seasonal predictions, it is essential to have a reliable prediction model such as a coupled global climate model (GCM). It is also widely believed that good initial condition is an essential factor to guarantee skillful forecasts, since seasonal prediction is fundamentally an initial value problem in a tier-1 prediction system.

A good initial condition should be accurate and well-balanced within the physical constraints of the prediction model. If an initial condition is not balanced, it causes a serious initial drift as soon as prediction starts, so the forecast skill is rapidly degraded. Therefore, many assimilation techniques have been developed in order to get accurate and well-balanced initial conditions (Derber and Rosati 1989; Rosati et al. 1997; Chen et al. 1995, 1997a, b; Galanti et al. 2003).

Due to its simplicity, the nudging and 3DVAR is widely used for assimilating observational data into complex coupled GCMs (Kirtman and Zebiak 1997; Luo et al. 2005; Palmer et al. 1994; Kug et al. 2008, 2009). However, the weighting coefÞcients in these schemes are constant in time, therefore, it can generate inaccurate initial conditions because the model uncertainty is ßow-dependent. To get over the problems, there are attempts to apply more complicated assimilation systems, such as the extended Kalman Þlter (EKF), ensemble Kalman Þlter (EnKF) and 4DVAR (Evensen 1994; Houtekamer and Mitchell 1998; Evensen and Van Leeuwen 2000; Evensen 2003; Zhang et al. 2005; Leeuwenburgh 2007).

Y.-G. Ham · I.-S. Kang (&)
School of Earth and Environment Sciences,
Seoul National University, Seoul 151-742, Korea
e-mail: kang@climate.snu.ac.kr

J.-S. Kug (&)
Korea Ocean Research and Development Institute, Ansan, Korea
e-mail: jskug@kordi.re.kr

As coupled GCMs evolve using high-resolution dynamics and complex physics, it becomes very difficult to develop and solve the linearized coupled GCM. In this sense, instead of 4DVAR and EKF which requires a linearized version of a prediction model (or adjoint model), the EnKF assimilation system is affordable as an initialization technique for a coupled GCM. Instead of the linearized model, the EnKF requires many ensemble members. It was highly problematic due to a huge amount of computational time, however, it has gradually become feasible to generate large numbers of ensemble members for EnKF assimilation as computing power has rapidly increased. As a result, the EnKF has recently been paid a lot of attention as a future framework for initialization of seasonal predictions.

However, although state-of-the-art initialization processes are used for seasonal predictions, it is unavoidable that every initial condition contains errors. These errors grow as the model integrates, and degrades the accuracy of the prediction. Because no prior information is given about these initial errors, the uncertainties arising from them are inevitable. Therefore, to reduce this initial uncertainty, ensemble prediction is introduced (Molteni and Palmer 1993; Buizza et al. 1998). As the number of ensemble members is increased, the initial uncertainty is greatly reduced. However, due to limited computing capacity, it is important to generate initial perturbations correctly (i.e. an appropriate initial ensemble spread), in order to reduce the initial uncertainty effectively.

According to studies about predictability, it is known that the prediction error is much reduced when a pair of initial ensemble members filters out the growing errors by containing fast-growing errors, because fast-growing errors dominate the total errors right after the beginning of the prediction, while the non-growing errors do not contribute much to the prediction errors (Toth and Kalnay 1997). It implies that a fast-growing mode in the prediction system is an optimal perturbation for ensemble prediction. To detect the fast growing mode, two methods are well known and commonly used. One of them is the singular vector (SV) method (Xue et al. 1997a, b; Chen et al. 1997a, b; Tang et al. 2006; Fan et al. 2000; Farrell 1989; Palmer et al. 1994; Molteni et al. 1996). In the SV method, unstable perturbations are calculated from the SVs of the linear propagator of the primitive equations. The other is the breeding method (Toth and Kalnay 1993, 1997; Cai et al. 2003; Corazza et al. 2003; Yang et al. 2006). Toth and Kalnay suggest breeding methods first to catch the fastest growing perturbations in the weather forecast system. After their first attempts, Cai et al. (2003) show that bred vectors improve the ENSO prediction skill with their Zebiak-Cane model. Yang et al. (2006) show that the breeding method can be applied to detect the coupled ENSO instability and isolate it from other short-term instabilities in coupled GCMs.

Though both methods are verified sufficiently by a number of researchers, they can cause initial drifts due to dynamical imbalances between the variables of the initial conditions, when the initial perturbation is added to the assimilated initial states. Because the relationship between ocean variables is nonlinear, final initial conditions lose the constraints of model physics, even though initial states and initial perturbations are well balanced individually. Also, it takes extra computer resources to generate initial perturbations to match the number of initial states generated. Therefore, it would be more effective if the initial perturbation generation method is merged into the assimilation system.

Compared with other techniques, using the EnKF assimilation system, it is easier to merge the fast growing initial perturbation generation method into the assimilation techniques. This is because the EnKF assimilation procedure is similar to that of the breeding method, which is a sequential method implying free integration of coupled GCMs during a fixed time periods. The only difference between the breeding method and the EnKF is the rescaling procedure, in which the analysis equation is solved in EnKF. Therefore, fast growing perturbations would be dynamically generated from EnKF for at least some of the ensemble members, like a perturbation with the breeding method. Thus, the method for generating fast growing initial perturbations is developed in this study by simply selecting fast-growing initial perturbations among a large number of initial conditions from EnKF. The authors apply this concept to a hybrid coupled model.

In this study, the EnKF assimilation system is applied in the context of a perfect model (Leeuwenburgh et al. 2005; Houtekamer et al. 2005). This means that an arbitrary model simulation is considered as the true state, and then observation values are obtained by adding random perturbations to this true state. There are two reasons for applying this perfect model assumption. The first one is that the performance of the EnKF technique is hard to validate in the assimilation using real observations, because the exact true states are never known. Secondly, wrong schemes for model error consideration excessively degrade the assimilation or the seasonal prediction results.

The paper is organized as follows. In Sect. 2, descriptions and the ENSO characteristics of the hybrid coupled model are mentioned. Data assimilation using EnKF techniques and a method to generate optimal ensemble perturbations are introduced in Sect. 3. Section 4 describes the superiority of the selection method, and the mechanisms of this superiority. A brief summary and discussion are included in Sect. 5.

2 Model descriptions

In this study, a hybrid coupled model is developed and used to apply EnKF, and then ENSO prediction is carried out with this model. The oceanic component of the hybrid coupled model is based on an intermediate ocean model similar to the Cane–Zebiak (CZ) model (Zebiak and Cane 1987). The difference of the present ocean model from the original one is a parameterization of subsurface temperature. The parameterization of subsurface temperature is calculated using a statistical relationship constructed on the basis of SVD SVs of the 20°C isothermal depth and the ocean temperature at 45 m depth from the NCEP ocean assimilation data (Kang and Kug 2000). Kang and Kug (2000) showed that the model with the present parameterization has better performance in the overall Pacific domain for SST predictions. Note that the time step of the oceanic model is 10 days.

The atmospheric component of the hybrid model is SPEEDY AGCM (Simplified Parameterizations, primitivE-Equation DYnamics, Molteni 2003). The goal of this model is to achieve computational efficiency while maintaining characteristics similar to the state-of-the-art AGCMs with complex physics. The resolution of the model is T42L10 (horizontal spectral truncation of 42 wave numbers and ten vertical levels). According to Molteni (2003), the SPEEDY model simulates the general structure of global atmospheric circulation fairly well, and some aspects of the systematic errors are similar to many AGCMs. The SPEEDY model includes basic components of physical parameterizations, such as convection, large-scale condensation, clouds, short-wave radiation, long-wave radiation, surface fluxes of momentum and energy, and vertical diffusion.

The coupling strategies between pairs of components are as follows. First, the air-sea coupling interval is 10 days. The oceanic (atmospheric) model provides its anomalous SST (anomalous wind stress), and receives anomalous zonal and meridional wind stresses (anomalous SST), whose value is 10-day averaged. Note that the oceanic part of the hybrid coupled model calculates only an anomaly. Therefore, climatological wind stress (SST) should be subtracted (added) to the total wind stress (SST anomalies) output before exchanging the properties. The climatological wind stress is obtained from 10-years integration of SPEEDY AGCM data with the prescribed climatological SST from NCEP data (Reynolds and Smith 1994). The SST data used for calculating climatology is from 1980 to 2000. The coupling with the CZ-type ocean model is done only over the Pacific domains (124°E–90°W, 20°S–20°N), and SST anomalies over the rest of the regions are calculated using a simple mixed layer model. In addition, because surface winds of the SPEEDY are noisy and relatively inaccurate due to the crude PBL processes, wind stresses are calculated from 850 hPa winds by using the simple bulk formula (Huang and Shukla 1997; Kirtman and Schneider 1996). The drag coefficient (C_d) for calculating wind stress is set to 0.6×10^{-3}.

Figure 1 shows the NINO3.4 SST from 100-years integration of the model. The standard deviation of NINO3.4 SST is 1.39, which is a value comparable with observations. The ENSO periods of the model are between 2 and 7 years, which are also comparable with observations. In addition, the power spectrum indicates that the model simulates an irregular ENSO period, similar to the observational one. Note that the period and amplitude of ENSO in the model is quite irregular compared with the original CZ model, which is mainly due to the atmospheric noises in the SPEEDY. Note that the irregular characteristics of ENSO are also observed and investigated in many studies (Ecker and Latif 1997; Blanke et al. 1997; Wittenberg et al. 2006). Moreover, ENSO shows a decadal modulation. For example, ENSO variability is relatively small from year 15 to 40, but is stronger from year 60 to 80.

3 EnKF algorithm and fast growing member selection method

Using the hybrid coupled model, the EnKF is applied for 50 years in the perfect model context. The SST, Kelvin and Rossby-wave components of thermocline depth are assimilated into the ocean model. Note that the thermocline depth is decomposed by Kelvin and Rossby wave components separately in the CZ-type model (Zebiak and Cane 1987; Cane and Patton 1984). The true state is obtained from year 1 to 50 of the 100-year model control simulation (see Fig. 1). Observational states are produced by adding random perturbations to the true state. The standard deviations of random perturbations for SST and thermocline depth are 0.4 C and 7 m, respectively. The prescribed magnitude of observational error is similar to that of Ballabrera-Poy et al. (2001). The EnKF assimilation starts from an arbitrary state, and then successive assimilation steps are generated for each 1-month period as the model integrates.

The analysis equation of the EnKF is as follows:

$$X_a = X_b + BH^T\left(HBH^T + R\right)^{-1}\{y - H(X_b)\} \quad (1)$$

X_a, X_b, y, B, R, and H denote the analysis values, background values (model forecasts prior to the analysis), observations, background error covariance, observational error covariance, and transform matrix, respectively. Simply put, this assimilation procedure determines the analysis values as a weighted linear combination of the background values (model forecasts prior to the analysis) and the observations. The weighting coefficient for model forecasts

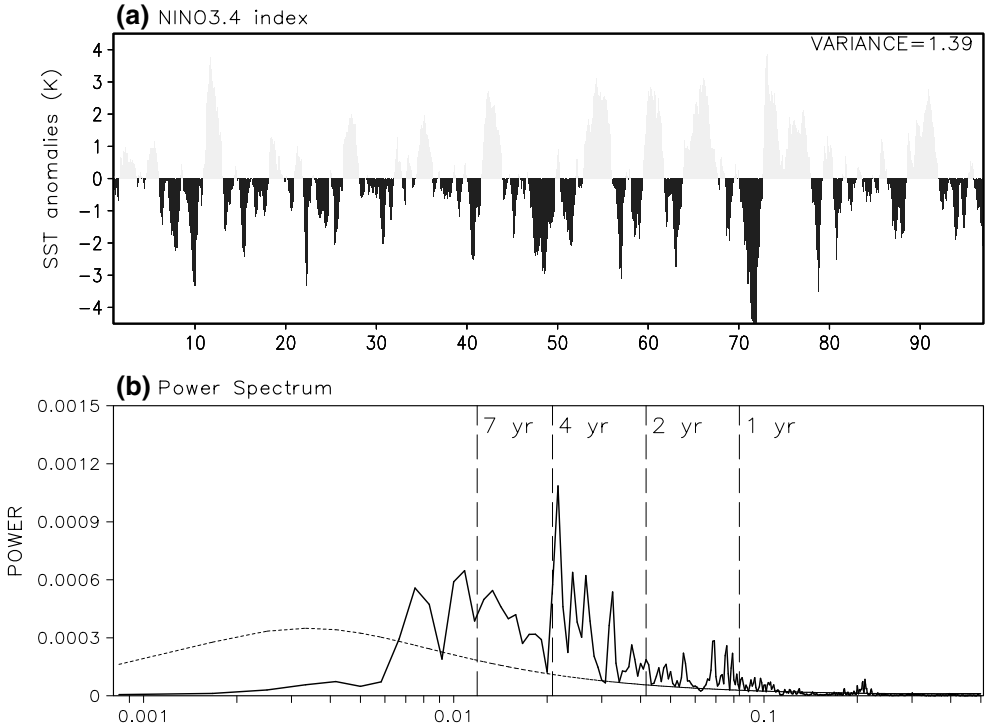

Fig. 1 a Time series of NINO3.4 SST anomalies of 100-years CZ-SPEEDY simulation. b Power-spectrum of the index

is inversely proportional to the magnitude of background error covariance. The background error covariance (B) is defined as the ensemble spread using background values for the EnKF.

In this study, 16 ensemble members are used for the EnKF assimilation. As is well known, due to computational limitations, the use of a limited size of ensemble members to represent forecast probability is unavoidable. With finite-sized ensemble members are used, the limitation is that only N directions of the forecast uncertainty can be sampled. This may cause an insufficient ensemble spread for the true forecast uncertainty distributions. However, the degrees of freedom of the hybrid coupled model are relatively small, due to the simplified physics, compared to those of other complex GCMs, and so 16 ensemble members would be enough to sample the solution space. For example, five dominant eigen-vectors explain over 90% of the total SST variance of the hybrid coupled model.

In addition, to reduce the deficiency originating from the limited size of the ensemble members, local analysis techniques have been used (Ott et al. 2004). With this technique, relatively small ensemble member sizes are sufficient to sample the forecast uncertainties, because the degrees of freedom in the solution matrix are reduced significantly by dividing the model dimension locally. A brief description of EnKF procedures is as follows. First, each ensemble member of the model is integrated synchronously for 1 month with slightly different initial conditions. The initial condition is selected randomly from 100 year integration, and initial perturbation is given in SST and thermocline depth anomaly whose magnitude is same as observational errors. As the model integrates, the ensemble spread will grow. The 1-month-integrated forecast values are defined as background values (X_b), and background error covariance (B) is estimated as the ensemble spread around the ensemble mean by using these forecast values. After that, by using given observations and observational error covariance, analysis values are obtained for each ensemble member followed by Eq. 1. Therefore, each ensemble member has an analysis value, and their ensemble mean represents a final analysis value. This is one cycle of the EnKF procedure. Then, another EnKF cycle starts by integrating the model over a 1-month period again using initial conditions which are the previous analysis values of the individual members. After the 1-month integration, the values are considered as background values again, and so the next analysis equation is solved with these background values. Likewise, the EnKF procedure is repeated until the observational data are not available.

Figure 2 shows an example of the assimilated thermocline depth with the EnKF at 110° W, 0° N. For comparison, true states and observations are also shown in Fig. 2. Because the random errors are added to the true states to generate the observational values in the perfect model context, the temporal patterns of observational errors are noisy and random. On the other hand, assimilation values

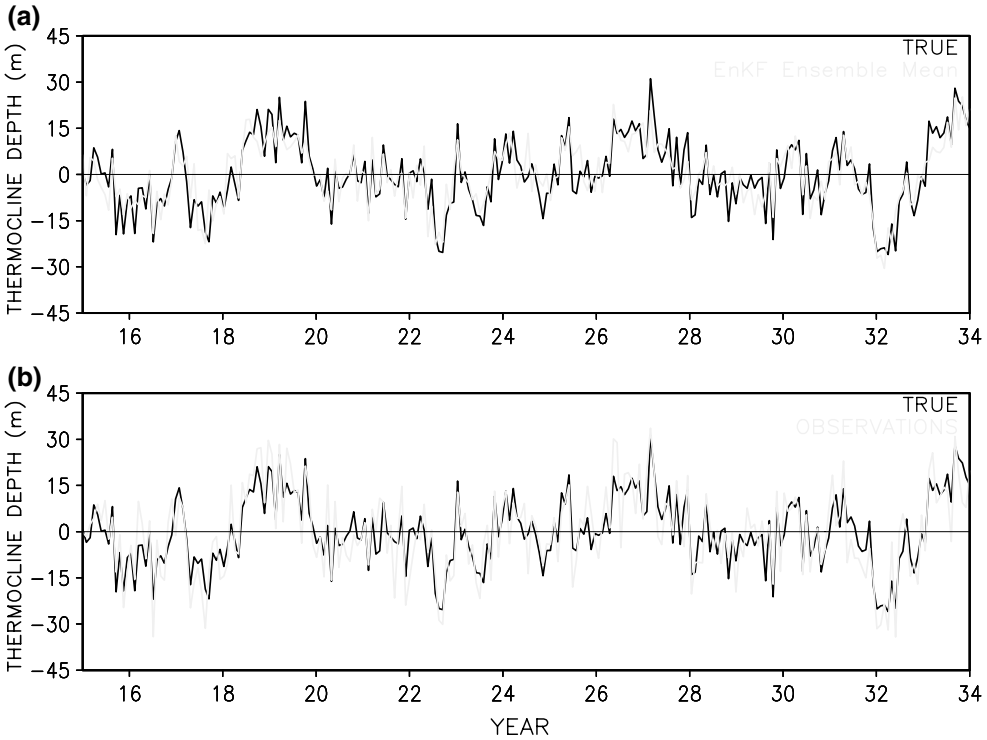

Fig. 2 True states (*black lines*), EnKF analysis values (*red line* in the *upper panel*), and observations (*red line* in the *lower panel*) of thermocline depth at 110° W, 0 N from year 15 to 34

with EnKF tend to be relatively smoother and systematic. This is because the model forecast values, which are constrained by the model equations, are mixed with the observations by using the analysis equations. In addition, RMS error with EnKF scheme is 6.07(m), which is smaller than the value for the observations. As mentioned before, the prescribed magnitudes of observational errors are 7(m). It appears that there is an effect of about 15% RMS error reduction after the EnKF method is applied.

Using the assimilated initial conditions with EnKF, an experimental ENSO forecast is performed. Usually, the ensemble mean of each analysis value is used as an initial condition for the ENSO forecast in the present EnKF method. Then, how can we make initial perturbations for the ensemble forecast? It is possible that analysis of each member could be used as an initial condition for the ensemble prediction. If we ignore computational costs, it is good to use whole ensemble members in the EnKF for ensemble prediction again. However, due to computational limitations, it is not affordable to use all the ensemble members, considering their huge number.

Note that computing cost is several times larger in the prediction than in assimilation if the ensemble member is same because the integration period is longer in the prediction. For example, in operational seasonal prediction system, time integration period for the assimilation is only 1 month when seasonal prediction is performed in every month (e.g. assimilation will start from the previous month). On the other hand, for prediction, 12-month integration should be carried out for 12-month lead forecast. Therefore, the prediction procedure mostly spent 12 times bigger computing resources than that of the assimilation procedure. This implies that reducing the integration load in the prediction by reducing ensemble member is critical in operational sense. Therefore, we have to select few members from the whole ensemble members. Thus, the selection process is important for effective ensemble prediction. In this study, we develop a simple method for selecting members for ensemble prediction.

In this study, we select ensemble member whose perturbation with respect to ensemble mean (which is best estimates in EnKF analysis) is fast-growing, because the fast growing mode is known to be an optimal initial perturbation for ensemble prediction. Note that each ensemble member is likely to contain same amount of initial uncertainty in the EnKF procedure, indicating that we can not choose a specified member in terms of the initial error. However, initial errors of some ensemble members can closely represent a fast-growing mode in the model. Therefore, they have relatively faster-growing perturbation from ensemble mean at some time, indicating that they will effectively reduce initial uncertainty in ensemble prediction. For example, growth rate based on single assimilation interval would be much different from ensemble member

to member, even though the time-averaged growth rate with respect to ensemble member is similar.

Then, how can we select flow-dependent fast growing perturbation in the present model? The fast growing mode can be obtained from the SV method or the Breeding method (Toth and Kalnay 1993; Xue et al. 1997a, b). However, they require an additional initialization system and huge computational cost. Instead of using an additional system, we developed a method to select a relatively fast growing mode with the EnKF procedure. In the breeding method, the fast growing mode is captured by repeating breeding (free integration) and rescaling. This procedure is quite similar to the present EnKF procedure which repeats 1-month integration and analysis, though the rescaling method is somewhat different. Therefore, it is possible to pick up a relatively fast growing mode with the EnKF procedure.

The detailed methodology for selecting ensemble members for seasonal predictions is as follows. Assume that seasonal prediction is started at time T. Firstly, the fastest growing member, whose growing time is from $T-1$ to T, is selected from among 16 ensemble members. To do this, a growth rate for individual members is defined as follows:

$$\text{Growth rate}_i = \frac{\sqrt{\int\int \left[SST(T)_i - SST(T)_{mean}\right]^2 dx dy}}{\sqrt{\int\int \left[SST(T-1)_i - SST(T-1)_{mean}\right]^2 dx dy}} \quad (2)$$

where i and mean indicate the individual ensemble and their ensemble mean, respectively. The growth rate represents how much the perturbation (departure from ensemble mean) grows for an 1-month integration. The area for calculating growth rate is over the central and eastern Pacific (180–90W, 10 S–10N). Based on this definition, the member with the largest growth rate is regarded as the fastest growing member, and its perturbation is saved as P_{fast}. Note that P_{fast} is a fastest growing perturbation among ensemble perturbations in EnKF. That is, true fastest growing perturbations can be different from the spatial pattern of P_{fast}. However, this method is beneficial that extra fast-growing perturbation generation method is not needed, therefore, it is easy to extract optimal perturbations.

Next, the analysis of the EnKF is applied, followed by Eq. 1. Note that the perturbation pattern of each individual ensemble member is changed after the analysis. Therefore, among 16 members, two ensemble members whose spatial patterns of ensemble perturbations are most similar to P_{fast} are finally selected for seasonal predictions based on pattern correlation. Because of the growth rate is computed based on a single assimilation interval, the shape of P_{fast} is different in every month. In addition to these two selected ensemble members, two mirror members are generated by multiplying -1 onto each ensemble perturbation. Finally, these four ensemble members are used for the ensemble ENSO prediction.

Figure 3 shows a dominant pattern of fast growing perturbation, P_{fast}, indicated by their first and second EOF eigen vector. Note that EOF analysis is based on 20 year assimilation. The first EOF eigen vector shows an east–west contrast pattern. This spatial pattern is similar to the ENSO peak pattern of the equatorial heat contents in the Recharge Oscillator. The spatial pattern of SST shows dominant anomalies over the eastern Pacific. This result is similar to that of Cai et al. (2003) who applied the breeding method to the Cane–Zebiak model. The second EOF vector shows maxima over the central Pacific, and this pattern is also regressed onto ENSO-like SST anomalies. This gives rise to the two dominant modes of fast growing perturbation which most represent the ENSO-related mode. In addition, each EOF mode of fast growing perturbation is well correlated to the dominant EOF pattern of initial errors, whose pattern correlation coefficient is around 0.9 (not shown).

4 ENSO prediction results

Within the 50-year period of EnKF analysis, the 20-year period from year 15 to 34 is used for seasonal prediction. For each case, a 12-month lead forecast is performed, which starts from 1st March, June, September, and December. Hereafter, forecasts with selected ensemble members are denoted as ÔÔselected predictionÕÕ. To compare the forecast skills of the selected ensemble members, the averaged forecast skill of all possible combinations of four ensemble members among 16 ensemble members (a total of 1,820 cases) is computed. This skill represents the mean correlation skill of the ensemble mean of four members. Hereafter, we will denote this forecast as the ÔÔcontrol predictionÕÕ. The skill of the control prediction will be compared with that of the selected prediction (the prediction based on four selected ensemble members with fast growing perturbations). Furthermore, the forecast skill of the ensemble mean of a 16-member ensemble in the EnKF (i.e. the prediction with all possible ensemble members obtained by the EnKF) is also calculated. The correlation skill and RMS errors are used to evaluate the performance of the seasonal forecasts.

Then, how much improvement in forecasts skills is achieved with selected fast growing perturbations? Prior to comparing ensemble mean predictions, we checked the forecast skill of individual members that consist of the control and selected predictions. We can define small and

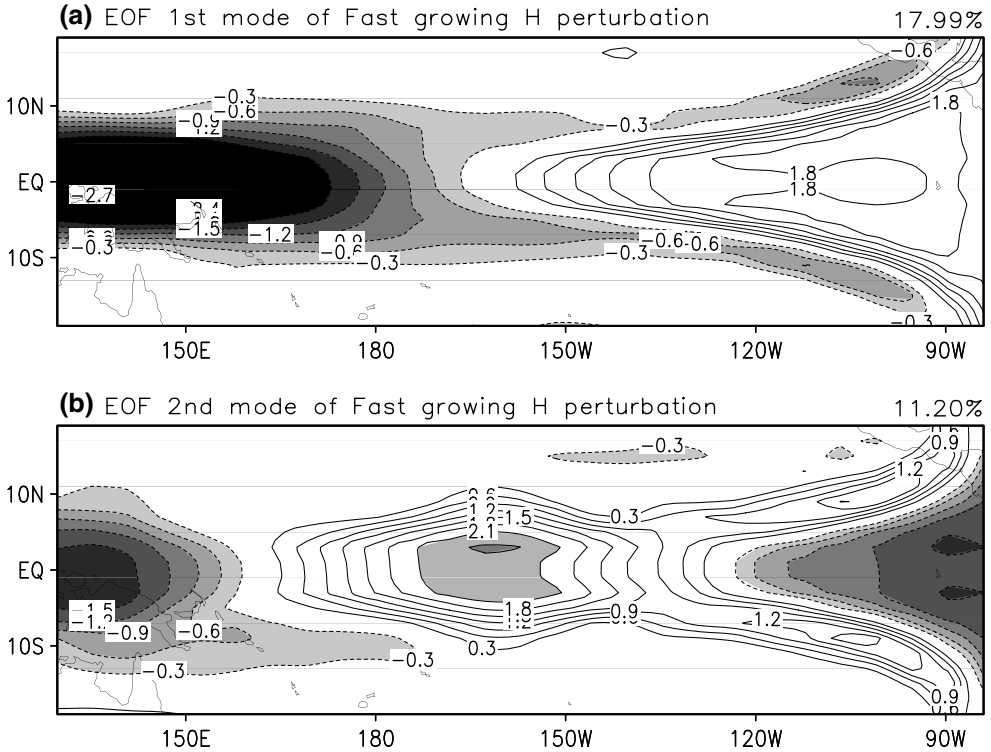

Fig. 3 EOF a Þrst and b second eigen-vector of fastest thermocline depth perturbations

large initial ensemble errors between one pair of ensemble members for each forecast case. According to studies about ensemble forecasts, ensemble prediction guarantees good forecast skills, because the reduction in forecast error for small initial ensemble error is bigger than the increase in forecast error for large initial ensemble error (Toth and Kalnay 1997). At the same time, among the small initial error ensemble members, the forecast error for initial conditions with fast-growing perturbations is considerably smaller than for those with other initial conditions, because the proportion of fast growing errors is mostly excluded. On the other hand, among large initial error ensemble members, the forecast error for initial conditions with fast-growing perturbations is almost same as that for those with other initial conditions, because there are mature stages of forecast error regardless of the proportion of fast growing forecast error (Cai et al. 2003).

To investigate whether this theory is applicable for seasonal predictions with a complex model, ensemble members with small initial errors are separated from those with large initial errors for each ensemble pair (small and large initial error groups have 16 ensemble members for each prediction case). We Þrst categorize averaged SST RMS errors of all ensemble members (E_{AVE}), that of small initial error groups (E_{AVE_SMALL}), and that of large initial error groups (E_{AVE_LARGE}). Then, among small and large initial error groups, selected ensemble member groups are further categorized as E_{FAST_SMALL}, and E_{FAST_LARGE}.

Figure 4 shows differences between E_{AVE} and E_{FAST_SMALL} (E_{AVE} and E_{FAST_LARGE}) and differences between E_{AVE_SMALL} (E_{AVE_LARGE}) and E_{FAST_SMALL} (E_{FAST_LARGE}). It is clear that the average forecast skill of all ensemble pairs would be smaller than that of single forecasts, because the RMSE reduction for a small initial error ensemble (Fig. 4a) is bigger than the RMSE increase for a large initial error ensemble (Fig. 4b). At the same time, the RMSE reduction of small initial errors with fast growing perturbations is much more signiÞcant than those for the small initial error groups (Fig. 4c). This implies that the RMSE reduction of ensemble forecasts mostly originates from small initial error ensembles, as many previous studies have mentioned.

In the ensemble prediction, forecast skill improvements for the selected prediction are signiÞcant. Figure 5 shows the correlation skills and RMS errors of four predicted ensemble-member mean NINO3.4 SST anomalies of the selected prediction and the control prediction. For comparison, the ensemble mean prediction for 16 ensemble members, which include all the ensemble members obtained by EnKF, is also calculated. The selected prediction demonstrates a greater accuracy than the control

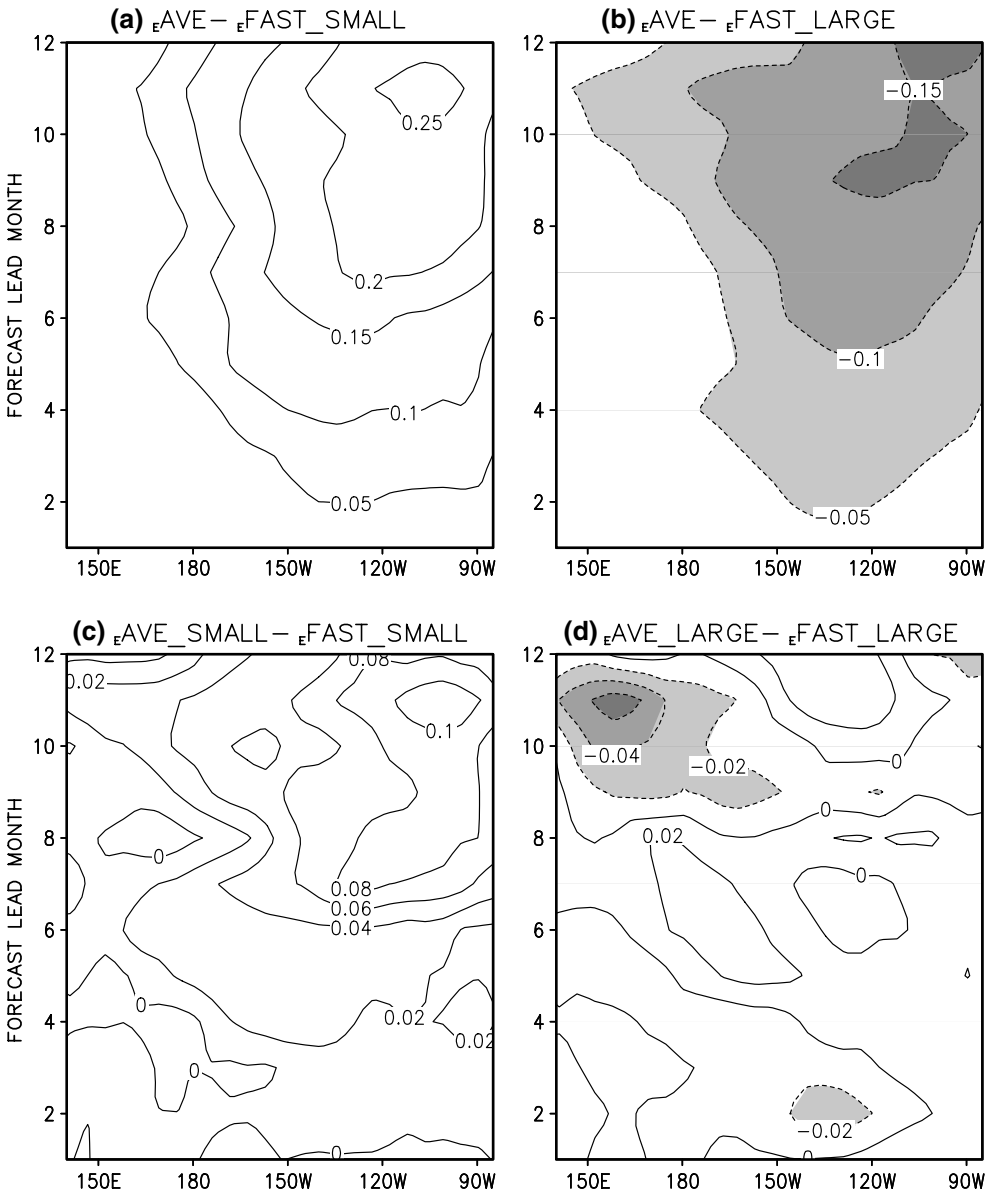

Fig. 4 Upper panels show difference of averaged SST RMS error of all ensemble members from a that of small initial error ensembles (εAVE−εFAST_SMALL), b that of large initial error ensembles among selected members (εAVE−εFAST_LARGE). *Lower panels* show difference of averaged SST RMS errors of c small initial error group from small initial error ensembles among selected ensemble members (εAVE_SMALL−εFAST_SMALL), d large initial error group from large initial error ensembles among selected ensemble members (εAVE_LARGE−εFAST_LARGE)

prediction over all forecast lead times. In particular, the forecast skill improvement is distinctive for a 6–8 month forecast lead. Also, it is interesting to note that the selected prediction shows an even higher skill than the control prediction with 16 ensemble members and a 6–8 month forecast lead. As the forecast lead period gets longer, the forecast skills of the selected prediction converge to that of the control prediction. However, note that the correlation skills of the selected ensemble members are still better up to a 12-month forecast lead.

One may wonder why the forecast skills are improved significantly for a 6–8 month forecast lead. According to Lorenz's simple error model results, the error reduction effect by performing ensemble prediction is significant when nonlinear forecast error growth is maxima. The forecast error growth can be roughly estimated by calculating the ensemble spread with respect to forecast lead time. The ensemble spread growth in ENSO prediction of the hybrid model is robust at a 6–8 month forecast lead (not shown). Therefore, with the aid of the Lorenz's

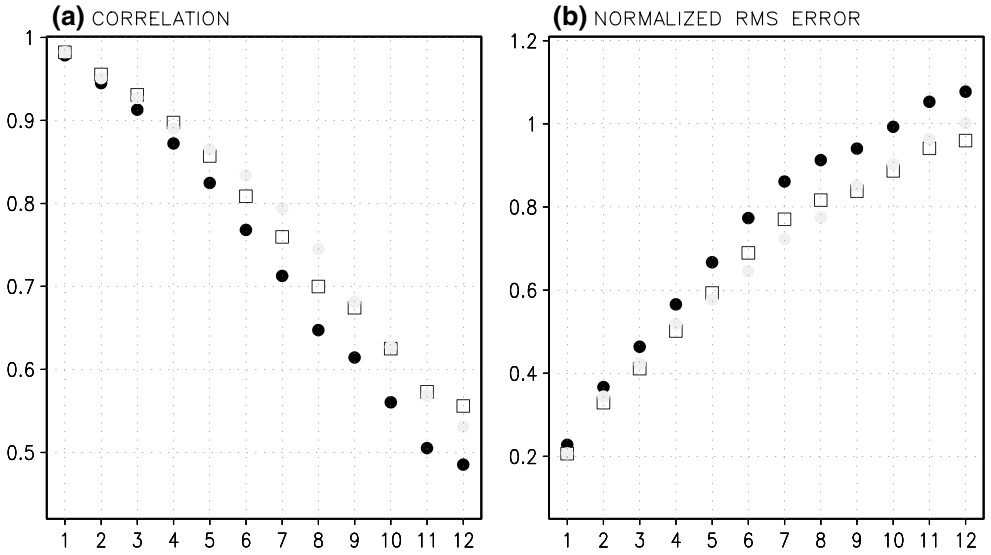

Fig. 5 a Correlation skills and b RMS errors of four ensemble-member mean NINO3.4 SST anomalies of the selected prediction (*red circle*), the control prediction (*black circle*), and that of 16 ensemble-member mean NINO3.4 SST anomalies (whole ensemble member obtained by EnKF, *blue squares*)

simple error model results, it is expected that the maximum error reduction will occur around a 6–8 month forecast lead.

Figure 6 shows the spatial patterns of correlation skill of the control prediction for SST anomalies (contour), and the difference between the control and selected prediction is shown (shading). Note that the positive sign denotes that forecast skills of the selected prediction have improved. In the control prediction, the correlation coefficients over the eastern Pacific are about 0.9 at a 3-month forecast lead, and it still reaches over 0.5 up to a 12-month forecast lead. Spatially, the correlation skill over the eastern Pacific is relatively higher than that over the western Pacific. Correlation skills of the control prediction start to decrease from the equatorial central Pacific, and the correlation skills are relatively higher over the off-equatorial central-eastern Pacific after a 6-month forecast lead. Because SST variability is higher over the eastern Pacific, RMS errors and their error reductions are also concentrated over that region.

It is worthwhile noting that the improvement in correlation skill is robust over the region where the correlation skill is relatively low. For example, at a 6-month forecast lead, improvements of forecast skills are significant over the central Pacific and far north-eastern Pacific, where the forecast skills of the control prediction drop rapidly. In addition, with a 9-month forecast lead, the improvements are significant over the central-southern Pacific where the correlation of the control prediction is only around 0.6. According to Lorenz, the forecast skill of the atmospheric model depends not only on the accuracy of the initial conditions but also on the instability of the flow itself[1,2] in the perfect model context (Lorenz 1963a, b, 1965). This means that the forecast skill is degraded when the errors in the initial conditions are large, or when the initial errors grow fast even though the errors in the initial conditions are relatively small. First, we will investigate how much the magnitudes of the initial errors affect the forecast skill. The averaged initial errors of the control prediction are large over the fast eastern Pacific, whose pattern is quite similar to that of the RMS errors of the control prediction (not shown). However, spatial patterns of the correlation skills of the control prediction are different from those of the initial errors. For example, at 3- and 6-month forecast leads, the correlation skills of the control prediction are still high over the far eastern Pacific. This implies that the correlation skill of the control prediction is also controlled by the instability of the flow (i.e. air–sea coupled instability). Because the flow-dependent instability is closely related to the limit of seasonal predictability, it is possible that the correlation skill of the control prediction determines the seasonal predictability. Therefore, if the contribution of coupled instability to the correlation skill is large, this means that there is an inverse relationship between predictability and skill improvements with fast growing initial perturbations.

[1] In the coupled model, especially over the tropics, the instability of the flow mainly comes from the air-sea coupled process (i.e. air–sea coupled instability).

[2] In the real case, uncertainties from model imperfections due to lack of model physical parameterization and numerical errors are other important sources for determining the forecast skills.

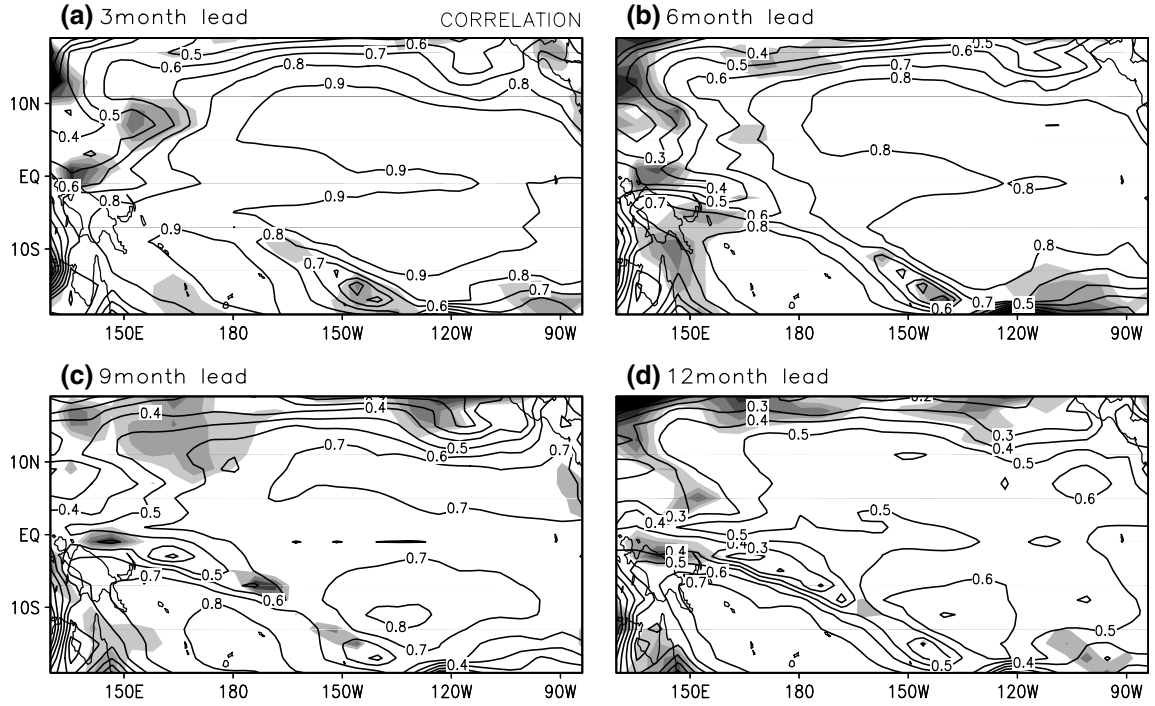

Fig. 6 Spatial patterns of correlation skill of the control prediction for SST anomalies (*contour*), and the difference between the control and the selected prediction is shown (*shading*). Note that positive sign denotes that forecast skills of the selected prediction are improved. The interval of shading is 0.02

In order to measure the predictability clearly and investigate the relationship between predictability and forecast skill improvements, the signal-to-noise ratio is introduced as follows (Rowell 1998; Kang and Shukla 2006).

$$\text{Ratio} = \frac{\sigma^2_{\text{signal}}}{\sigma^2_{\text{noise}}} \quad (3)$$

where,

$$\sigma^2_{\text{signal}} = \frac{1}{N-1}\sum_{i=1}^{N}\left(\bar{x}_i - \bar{\bar{x}}\right)^2$$

$$\sigma^2_{\text{noise}} = \frac{1}{N(n-1)}\sum_{i=1}^{N}\sum_{j=1}^{n}\left(x_{ij} - \bar{x}_i\right)^2$$

i indicates the individual prediction year ($N = 20$), j is the ensemble member, and ($n = 16$). $\bar{\bar{x}}$, $\bar{x}_i$ are the climatology of the ensemble mean and ensemble mean itself, respectively. Originally, a signal-to-noise ratio is introduced to examine the limit of seasonal predictability with atmospheric GCM. According to the theory, model prediction is decomposed into the signal (external) part, due to predictable SST forcing, and the noise (internal) part due to unpredictable atmospheric non-linear internal dynamics. Basically, the forecast is deterministic (stochastic) where the signal-to-noise ratio is large (small). Also, the region becomes relatively predictable (unpredictable) where the ratio is large (small).

Figure 7 shows the signal-to-noise ratio of SST anomalies with respect to the forecast lead time. It is found that the improvement of correlation skill is more significant where the signal-to-noise ratio is lower. For example, improvements of forecast skill are significant over the central Pacific and far north-eastern Pacific, where the STN ratio is lower than domain average at a 6-month forecast lead. The pattern correlation coefficient between STN ratio and correlation improvement over the domain (120E–90 W, 15 S–15N) is −0.45. Off-equatorial and far-western Pacific region are excluded, because SST norm is defined over equatorial eastern Pacific region. As mentioned before, because we use perfect model experiments, any systematic model bias cannot contribute to the signal-to-noise ratio (predictability). This means fast growing perturbation works more effectively over the unpredictable region where the forecast skill is sensitive to the uncertainty of the initial condition. On the other hand, the forecast skill is not sensitive to the pattern of ensemble perturbations where the signal-to-noise ratio is high. It means that the initial perturbation does not grow much, and therefore every ensemble member evolves to similar states regardless of the pattern of ensemble perturbations over those regions.

To show the relationship between the correlation improvement and signal-to-noise ratio clearly, a scatter diagram between them over the 180–90W, 10 S–10N

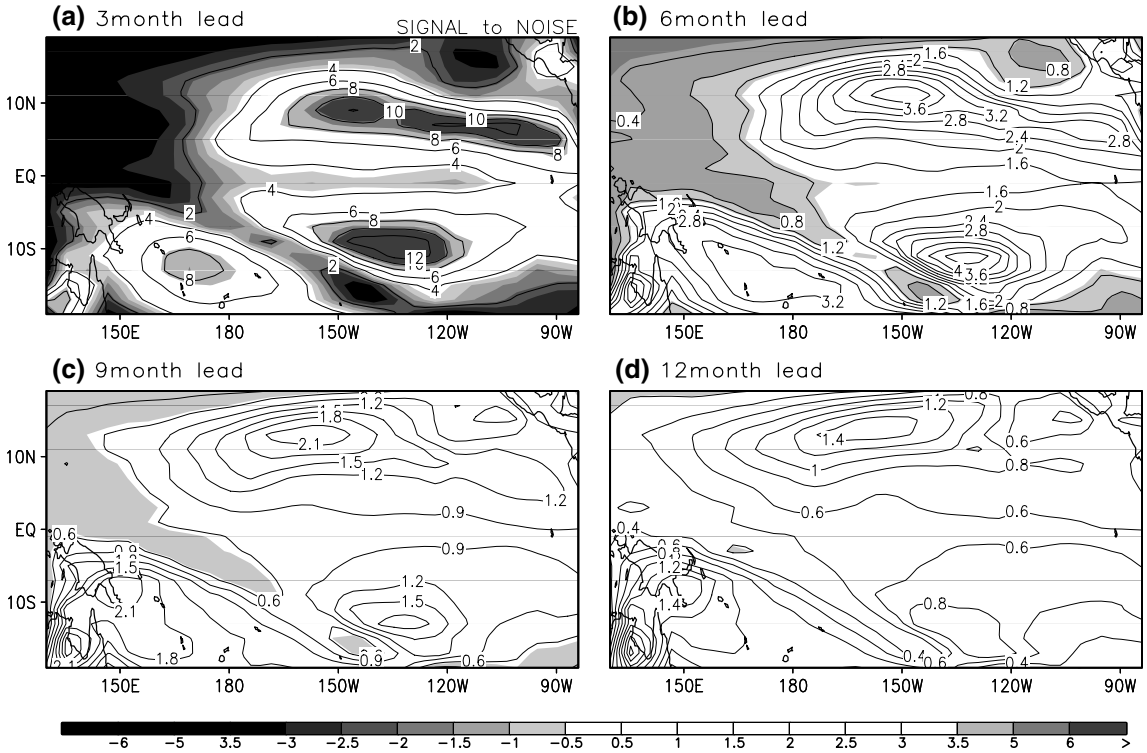

Fig. 7 Signal-to-noise (*STN*) ratio of SST anomalies with respect to the forecast lead month. Shading denotes the deviation of STN ratio from the domain average

region at a 6-month forecast lead is shown in Fig. 8. It is clear that the correlation improvement is inversely proportional to the signal-to-noise ratio. The inverse relationship is robust at the 6–9 month forecast lead, but the relationship is also observed at other forecast leads (not shown).

In addition, it is worthwhile investigating whether the inverse relationship between forecast skill improvement and signal-to-noise ratio is applicable temporally, as well as spatially. In particular, forecast skills and predictabilities with respect to the ENSO phase have been given attention and have been investigated by many researchers. Among them, Chen et al. (1997a, b) and Xue et al. (1997a, b) show that initial perturbation growth is robust during the onset and the breakdown of an ENSO event, using the SV method. Cai et al. (2003) support the phenomena by measuring initial perturbation growth using the breeding method. Thus, in accordance with the discussions mentioned in previous paragraphs, the forecast skill improvements made by selecting fast growing ensemble members would be significant at the ENSO onset and decaying phases.

To investigate the forecast skill improvement with respect to ENSO phase, reduction of RMS errors by selecting fast growing perturbations (thick bar), ensemble spread of 16 ensemble members (thin bar), and NINO3.4 SST anomalies (solid line) of five biggest El Nino events are shown in Fig. 9. The positive (negative) sign on the x-axis denotes the El Nino decaying (developing) phase. In this composite analysis, the magnitude of the ensemble spread is used to measure the predictability instead of the signal-to-noise ratio, because we assume the signal is constant in the same region with the same forecast lead time. Note that the figure is obtained using forecast values at 6–8 month forecast lead. For example, values for September, October, and November are obtained, respectively, from forecast values at 6-, 7-, and 8-month forecast leads starting from March. Likewise, values for December are from the 6-month forecast values starting from June.

This results is consistent to several studies with bred vectors and SVs (Xue et al. 1997a, b; Cai et al. 2003; Yang et al. 2006; Kug et al. 2008, 2009) that bred vectors (or SVs) grow fastest during ENSO transition. Especially, Cai et al. (2003) explain the dynamical relationship between the growth rate of the fast growing modes and the ENSO phase by using simple delayed oscillator model. In addition, Yang et al. (2006) show that the growth rate of the coupled bred vectors peaks about 3 months before an ENSO event.

The reduction of RMS errors is significant during the El Nino onset phase (10–16 month prior to the El Nino peak), the decaying phase (9–11 month after the El Nino peak),

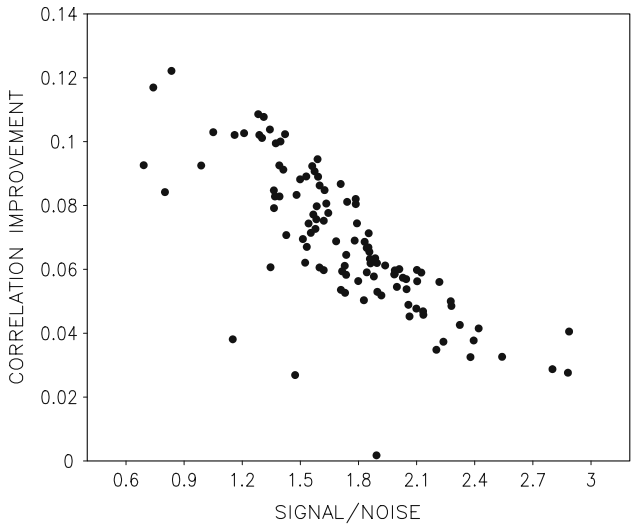

Fig. 8 Scatter diagram between signal-to-noise ratio (*x*-axis) and correlation improvement (*y*-axis) over eastern Paciþc regions (180Ð90 W, 10 SÐ10N)

and several months prior to the El Nino peak phase (2Ð4 months prior to the El Nino peak). Note that the most signiþcant improvements are during the El Nino onset and decaying phases. Then, is the predictability measured by the ensemble spread also small during the El Nino onset and decaying phases as expected? It is clear that the ensemble spread is also large during the El Nino onset and decaying phases. This supports the results in many studies about the initial perturbation growth which were mentioned before.

The large ensemble spread during the El Nino onset and decaying phases is related to the fact that small initial perturbations can lead to totally different ENSO evolutions during that period. For example, we may assume that an ENSO episode is on neutral phase with slight negative SST anomalies over the eastern Paciþc. In this case, small positive SST perturbations can change the sign of the SST anomalies to positive, and then the positive SST anomalies can be enhanced during the forecast time by Bjerknes feedback. Several months later, a case with positive initial perturbations may show totally different ENSO evolution from that without the perturbations. However, if an episode is on the mature phase, small SST perturbations are unable to change the sign of the SST anomalies. This implies that it is more crucial to treat the initial uncertainty effectively during the El Nino onset or decaying phases than it is for the mature phase. Because the ensemble member selection method developed in this study reduces uncertainties in the initial conditions, results showing that forecast skill improvements are signiþcant during the El Nino onset and decaying phases are acceptable.

5 Summary and discussion

In this study, the EnKF technique is applied to a hybrid coupled model, in order to generate initial conditions for ensemble seasonal prediction in the perfect model context. In this framework, a new approach is introduced to make initial perturbations for ensemble prediction. Among the initial conditions generated by the EnKF analysis, ensemble members with fast growing perturbations are selected for seasonal prediction.

With the selected ensemble members, seasonal forecast experiments are performed to investigate whether initial conditions with fast growing perturbations guarantee improved forecast skills. In the case of ensemble forecasts

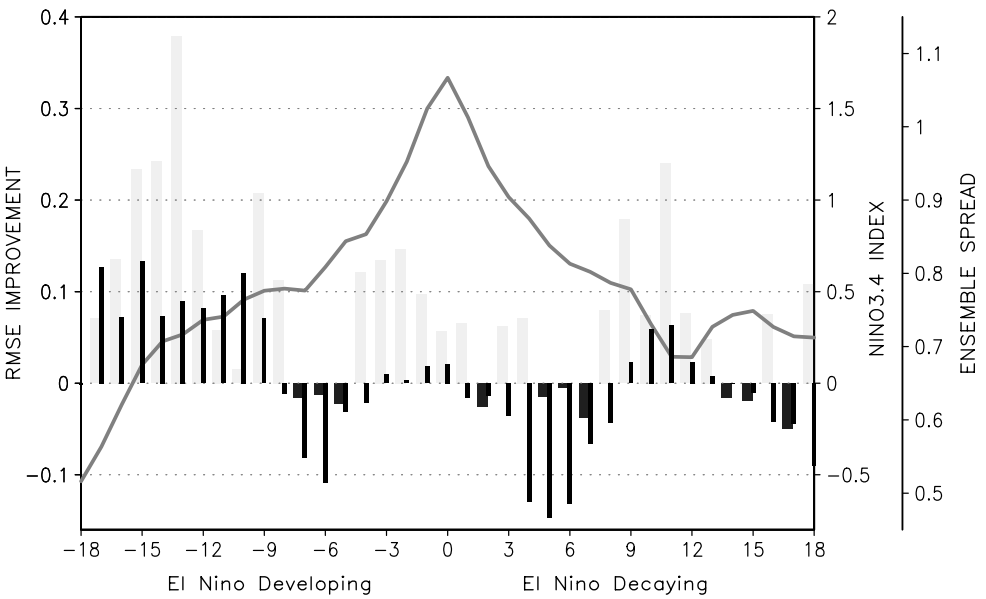

Fig. 9 Reduction of RMS errors by selecting fast growing perturbation (*thick bar*), ensemble spread of 16 ensemble members (*thin bar*), and NINO3.4 SST anomalies (*solid line*) of Þve biggest El Nino events

with four ensemble members, the correlation skills of the selected ensemble members is improved by a factor of about 0.1 more than those of other ensemble members at 6-7 month forecast leads. With this forecast lead, ensemble prediction with four selected ensemble members shows improved forecast skills even beyond predictions with 16 ensemble members (whole ensemble member obtained by EnKF). As the forecast lead gets longer, the forecast skills of selected ensemble members become similar to that of other ensemble members, but the skill improvement is still maintained up to a 12-month forecast lead.

In addition, it is found that there is an inverse relationship between forecast skill improvements and flow-dependent instability. That is, at the 6- and 9-month forecast leads, the region where correlation improvement is most significant is coincident with the region of low signal-to-noise ratio. In addition, this relationship applies to the case of time evolution of ENSO events. The predictability measured by the ensemble spread shows that the ENSO onset and decaying phases are the most unpredictable periods of the whole event. Being consistent with low predictability (large ensemble spread), the RMS error reduction in predictions using fast growing perturbations is also significant during that period.

This study has laid a foundation for the application of the optimal perturbation selection method to real predictions with more complex coupled GCMs. Then, what are the further points to investigate to apply the selection method developed in this study to complex coupled GCMs? Firstly, the initial amplitude of ensemble perturbations and the analysis time interval should be chosen carefully to generate the ensemble perturbations which guarantee skillful seasonal forecasts. Toth and Kalnay (1993) mentioned that if the initial perturbation amplitude is chosen to be much smaller than the estimated analysis errors, then an unwanted perturbation can be generated. Fortunately, the rough magnitude of the analysis error (it may be between twice and ten times the observed errors) is known, and we are therefore confident we can handle this point.

On the matter of the analysis time interval, more caution is needed. Usually, the analysis time interval of EnKF is chosen as 5 or 10 days, in order to use all the accessible observations in many studies using oceanic or coupled GCMs (Leeuwenburgh et al. 2005; Leeuwenburgh 2007; Zhang et al. 2005; Haugen and Evensen 2002). However, some argue that an analysis time interval of 5 or 10 days is rather short for smoothing out perturbations related to the air–sea coupled instability which is beneficial to the seasonal forecasts. For example, Yang et al. (2006) mentioned that the rescaling period would be longer than 2 weeks to separate fast weather instabilities from the slowly evolving ENSO instability. Therefore, in their study, the time interval for breeding is set to 1 month, in order to extract the ENSO instability. This implies that further work should be carried out to apply the ensemble selection method to complex GCMs, to investigate the appropriate analysis time interval to obtain both the optimal ensemble mean values and the ensemble perturbations simultaneously.

The observing network and local analysis strategy used by this study may degrade the necessity of EnKF system, because one of main purpose of assimilation system is to propagate the observational signals over areas with sparse data. However, some of ensemble perturbations would be relatively fast-growing than other ensemble perturbations with coarse observing network, because generation of ensemble perturbation largely depends on model dynamics with given background flows. It means observing system is less sensitive to generation of the ensemble perturbation. Therefore, our arguments are still rigorous if coarser observing network is applied.

Because this study has focused on ENSO prediction, the norm for measuring ensemble perturbation growth is defined as the area-averaged RMS magnitudes of SST perturbation (the SST difference between the ensemble mean value and each ensemble member) over the equatorial eastern Pacific region. This has to do with the fact that forecast skills are significantly better over the equatorial regions than in other regions. Therefore, if the main target of seasonal prediction is region-dependent, the definition of the norm should be changed. Some may also ask whether the variable which defines the norm is important for determining the forecast skills with selected ensemble members. On this subject, several studies about bred vectors have already mentioned that the choice of the norm-defining variable is not sensitive to the spatial patterns of perturbations (Cai et al. 2003; Toth and Kalnay 1993, 1997). However, studies about the role of the choice of the norm-defining variable have not been investigated at all with reference to complex coupled GCMs, and therefore remain worthy objectives for the future.

This study is focused on reducing prediction error by selecting fast-growing perturbation. However, because fast-growing perturbations act to widen the distribution of ensemble PDF, this method can be also applied to solve covariance underestimation problem, which is common in many EnKF systems (Anderson and Anderson 1999; Houtekamer and Mitchell 1998; Whitaker and Hamill 2002). To expand our study to assimilation procedure, several problems should be carefully considered. One of them is that fast-growing perturbation can overestimate the forecast error magnitude in assimilation procedure. The other is that, when the present method is applied to the assimilation procedure, it is possible every ensemble member shows a similar spatial pattern, which leads severe rank deficiency problem. Due to these problems, combine fast-growing perturbation generation method into EnKF is

a still controversial issue, even though several studies have already tried to apply fast-growing perturbation method into EnKF (Cohn and Dee 1988; Bishop et al. 2003; Wang and Bishop 2003).

The ensemble member selection method developed in this study emphasizes the importance of optimal initial perturbations on seasonal prediction. Though many studies have investigated the effect of initial perturbations on medium range forecasts with atmospheric GCM (Thompson 1957; Lorenz 1965; Charney et al. 1966; Epstein 1969; Leith 1974; Molteni et al. 1996), studies about optimal initial perturbations for seasonal prediction with coupled GCMs are just at the beginning stages. This is because seasonal prediction experiments with coupled GCMs require intensive computational powers, and so it is expensive to run coupled GCMs with large numbers of slightly different initial conditions to investigate the effect of those initial perturbations. For that reason, many operational seasonal prediction centers apply the simple Lagged Averaged Forecast method (LAF) for generating ensemble perturbations for coupled GCMs. As a first step, this study suggests how to generate the optimal initial perturbations for seasonal predictions within the EnKF initialization system. As mentioned in the introduction, the selection method developed in this study is easily applicable as long as the EnKF analysis system is already prepared. Therefore, as the scientific interests in, and operational demands on, the EnKF are increased, the ensemble member selection method is worthwhile as a process for generating optimal initial perturbations.

Acknowledgments This work supported by the Korea Meteorological Administration Research and Development Program under Grant CATER_2006-4206 and the second stage of the Brain Korea 21.

References

Anderson JL, Anderson SL (1999) A Monte Carlo implementation of the nonlinear filtering problem to produce ensemble assimilation and forecasts. Mon Weather Rev 127:2741–2758. doi: 10.1175/1520-0493(1999)127<2741:AMCIOT>2.0.CO;2

Ballabrera-Poy J, Busalacchi AJ, Murtugudde R (2001) Application of a reduced-order Kalman filter to initialize a coupled atmosphere–ocean model: impact on the prediction of El Nino. J Clim 14:1720–1737. doi: 10.1175/1520-0442(2001)014<1720:AOARKO>2.0.CO;2

Bishop CH, Reynolds CA, Tippett MK (2003) Optimization of the fixed global observing network in a simple model. J Atmos Sci 60:1471–1489. doi: 10.1175/1520-0469(2003)060<1471:OOTFGO>2.0.CO;2

Blanke B, Neelin JD, Gutzler D (1997) Estimating the effect of stochastic wind stress forcing on ENSO irregularity. J Clim 10:1473–1486. doi: 10.1175/1520-0442(1997)010<1473:ETEOSW>2.0.CO;2

Buizza R, Petroliagis T, Palmer TN, Hamrud M, Hollingsworth A, Simmons A, Wedi N (1998) Impact of model resolution and ensemble size on the performance of an ensemble prediction system. Q J R Meteorol Soc 124:1935–1960. doi: 10.1002/qj.49712455008

Cai M, Kalnay E, Toth Z (2003) Bred vectors of the Zebiak–Cane model and their potential application to ENSO predictions. J Clim 16:40–56. doi: 10.1175/1520-0442(2003)016<0040:BVOTZC>2.0.CO;2

Cane MA, Patton RJ (1984) A numerical model for low-frequency equatorial dynamics. J Geophys Res 99:25725–25738

Charney JG, Fleagle RG, Riehl H, Lally VE, Wark DQ (1966) The feasibility of a global observation and analysis experiment. Bull Am Meteorol Soc 47:200–220

Chen D, Zebiak SE, Busalacchi AJ, Cane MA (1995) An improved procedure for El Nino forecasting: implications for predictability. Science 269:1699–1702. doi: 10.1126/science.269.5231.1699

Chen D, Zebiak SE, Cane MA, Busalacchi AJ (1997a) Initialization and predictability of a coupled ENSO forecast model. Mon Weather Rev 125:773–788. doi: 10.1175/1520-0493(1997)125<0773:IAPOAC>2.0.CO;2

Chen Y-Q, Battisti DS, Palmer TN, Barsugli J, Sarachik ES (1997b) 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. doi: 10.1175/1520-0493(1997)125<0831:ASOTPO>2.0.CO;2

Cohn SE, Dee DP (1988) Observability of discretized partial differential equations. Sian J Numer Anal 25:586–617

Corazza M, Kalnay E, Patil DJ, Yang S-C, Morss R, Cai M, Szunyogh I, Hunt BR, Yorke JA (2003) Use of the breeding technique to estimate the structure of the analysis ÔÔerrors of the dayÕÕ. Nonlinear Processes Geophys 10:1–11

Derber J, Rosati A (1989) A global oceanic data assimilation system. J Phys Oceanogr 33:1877–1888

Ecker C, Latif M (1997) Predictability of a stochastically forced hybrid coupled model of El Nino. J Clim 10:1488–1504. doi: 10.1175/1520-0442(1997)010<1488:POASFH>2.0.CO;2

Epstein ES (1969) Stochastic dynamic prediction. Tellus 21:739–759

Evensen G (1994) Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. J Geophys Res 99:10143–10162. doi: 10.1029/94JC00572

Evensen G (2003) The ensemble Kalman filter: theoretical formulation and practical implementation. Ocean Dyn 53:343–367. doi: 10.1007/s10236-003-0036-9

Evensen G, van Leeuwen PJ (2000) An ensemble Kalman smoother for nonlinear dynamics. Mon Weather Rev 128:1852–1867. doi: 10.1175/1520-0493(2000)128<1852:AEKSFN>2.0.CO;2

Fan Y, Allen MR, Anderson DLT, Balmaseda MA (2000) How predictability depends on the nature of uncertainty in initial conditions in a coupled model of ENSO. J Clim 13:3298–3313. doi: 10.1175/1520-0442(2000)013<3298:HPDOTN>2.0.CO;2

Farrell BF (1989) Optimal excitation of baroclinic waves. J Atmos Sci 46:1193–1206. doi: 10.1175/1520-0469(1989)046<1193:OEOBW>2.0.CO;2

Galanti E, Tziperman E, Harrision M, Rosati A, Sirkes Z (2003) A study of ENSO prediction using a hybrid coupled model and the adjoint method for data assimilation. Mon Weather Rev 131:2748–2764. doi: 10.1175/1520-0493(2003)131<2748:ASOEPU>2.0.CO;2

Haugen VE, Evensen G (2002) Assimilation of SLA and SST data into an OGCM for the Indian Ocean. Ocean Dyn 52:133–151. doi: 10.1007/s10236-002-0014-7

Houtekamer PL, Mitchell HL (1998) Data assimilation using an ensemble Kalman filter technique. Mon Weather Rev 126:796–811. doi: 10.1175/1520-0493(1998)126<0796:DAUAEK>2.0.CO;2

Houtekamer PL, Herschel L, Pellerin G, Buenhner M, Charron M, Spacek L, Hansen B (2005) Atmospheric data assimilation with an ensemble Kalman Þlter. Mon Weather Rev 133:604Ð620. doi: 10.1175/MWR-2864.1

Huang B, Shukla J (1997) An examination of the AGCM simulated surface wind stress and low-level winds over the tropical Paciþc Ocean. Mon Weather Rev 125:985Ð998. doi: 10.1175/1520-0493(1997)125<0985:AEOTAS>2.0.CO;2

Kang I-S, Kug J-S (2000) An El-Nino prediction system using an intermediate ocean and a statistical atmosphere. Geophys Res Lett 15:1167Ð1170. doi: 10.1029/1999GL011023

Kang I-S, Shukla J (2006) Dynamic seasonal prediction and predictability of the monsoon. In: The Asian monsoon, chap 15. Springer-Praxis, Chichester

Kirtman BP, Schneider EK (1996) Model based estimates of equatorial zonal wind stress. J Clim 9:1077Ð1091. doi: 10.1175/1520-0442(1996)009<1077:MBEOEW>2.0.CO;2

Kirtman BP, Zebiak SE (1997) ENSO simulation and prediction with a hybrid coupled model. Mon Weather Rev 125:2620Ð2641. doi: 10.1175/1520-0493(1997)125<2620:ESAPWA>2.0.CO;2

Kug J-S, Kang I-S, Choi D-H (2008) Seasonal climate predictability with tier-one and tier-two prediction systems. Clim Dyn 31. doi: 10.1007/s00382-007-0264-7

Kug J-S, Ham Y-G, Kimoto M, Jin F-F, Kang I-S (2009) New approach on the optimal perturbation method for ensemble climate prediction. J Clim (submitted)

Leeuwenburgh O (2007) Validation of an EnKF system for OGCM initialization assimilating temperature, salinity, and surface height measurements. Mon Weather Rev 126:125Ð139. doi: 10.1175/MWR3272.1

Leeuwenburgh O, Evensen G, Bertino L (2005) The impact of ensemble Þlter deÞnition on the assimilation of temperature proÞles in the tropical Paciþc. Q J R Meteorol Soc 131:3291Ð 3300. doi: 10.1256/qj.05.90

Leith CE (1974) Theoretical skill of Monte Carlo forecasts. Mon Weather Rev 102:409Ð418. doi: 10.1175/1520-0493(1974)102<0409:TSOMCF>2.0.CO;2

Lorenz EN (1963a) Deterministic non-periodic ßow. J Atmos Sci 20:130Ð141. doi: 10.1175/1520-0469(1963)020<0130:DNF>2.0.CO;2

Lorenz EN (1963b) The predictability of hydrodynamic ßow. Trans NY Acad Sci Ser II 25:409Ð432

Lorenz EN (1965) A study of the predictability of a 28-variable atmospheric model. Tellus 17:321Ð333

Luo J-J, Masson S, Behera S (2005) Seasonal climate predictability in a coupled OAGCM using a different approach for ensemble forecasts. J Clim 18:4474Ð4497. doi: 10.1175/JCLI3526.1

Molteni F (2003) Atmospheric simulations using a GCM with simpliÞed physical parameterizations. I: model climatology and variability in multi-decadal experiments. Clim Dyn 20:175Ð191

Molteni F, Palmer TN (1993) Predictability and Þnite time instability of the northern winter circulation. Q J R Meteorol Soc 119:269Ð 298. doi: 10.1002/qj.49711951004

Molteni F, Buizza R, Palmer TN, Petroliagis T (1996) The ECMWF ensemble prediction system: methodology and validation. Q J R Meteorol Soc 122:73Ð119. doi: 10.1002/qj.49712252905

Ott E, Hunt BR, Szunyogh I, Zimin AV, Kostelich EJ, Corazza M, Kalnay E, Patil DJ, Yorke JA (2004) A local ensemble Kalman Þlter for atmospheric data assimilation. Tellus 56A:415Ð428

Palmer TN, Buizza R, Molteni E, Chen Y-Q, Corti S (1994) Singular vectors and the predictability of weather and climate. Philos Trans R Soc Lond 348:459Ð475. doi: 10.1098/rsta.1994.0105

Reynolds RW, Smith TM (1994) Improved global sea surface temperature analyses using optimum interpolation. J Clim 7:929Ð948. doi: 10.1175/1520-0442(1994)007<0929:IGSSTA>2.0.CO;2

Rosati A, Miyakoda K, Gudgel R (1997) The impact of ocean initial conditions on ENSO forecasting with a coupled model. Mon Weather Rev 125:754Ð772. doi: 10.1175/1520-0493(1997)125<0754:TIOOIC>2.0.CO;2

Rowell DP (1998) Accessing potential seasonal predictability with an ensemble of multidecadal GCM simulations. J Clim 11:109Ð120. doi: 10.1175/1520-0442(1998)011<0109:APSPWA>2.0.CO;2

Tang Y, Kleeman R, Miller S (2006) ENSO predictability of a fully coupled GCM model using singular vector analysis. J Clim 19:3361Ð3377. doi: 10.1175/JCLI3771.1

Thompson PD (1957) Uncertainty of initial state as a factor in the predictability of large scale atmospheric ßow patterns. Tellus 9:275Ð295

Toth Z, Kalnay E (1993) Ensemble forecasting and NMC: the generation of perturbations. Bull Am Meteorol Soc 74:2317Ð 2330. doi: 10.1175/1520-0477(1993)074<2317:EFANTG>2.0.CO;2

Toth Z, Kalnay E (1997) Ensemble forecasting at NCEP and the breeding method. Mon Weather Rev 125:3297Ð3318. doi: 10.1175/1520-0493(1997)125<3297:EFANAT>2.0.CO;2

Wang X, Bishop CH (2003) A comparison of breeding and ensemble transform Kalman Þlter ensemble forecast schemes. J Atmos Sci 60:1140Ð1158. doi: 10.1175/1520-0469(2003)060<1140:ACOBAE>2.0.CO;2

Whitaker JS, Hamill TM (2002) Ensemble data assimilation without perturbed observations. Mon Weather Rev 130:1913Ð1924. doi: 10.1175/1520-0493(2002)130<1913:EDAWPO>2.0.CO;2

Wittenberg AT, Rosati A, Lau N-C, Ploshay J (2006) GFDLÕs CM2 global coupled climate models. Part III: tropical Paciþc climate and ENSO. J Clim 19:698Ð722. doi: 10.1175/JCLI3631.1

Xue Y, Cane MA, Zebiak SE (1997a) Predictability of a coupled model of ENSO using singular vector analysis. Part I: optimal growth in seasonal background and ENSO cycles. Mon Weather Rev 125:2043Ð2056. doi: 10.1175/1520-0493(1997)125<2043:POACMO>2.0.CO;2

Xue Y, Cane MA, Zebiak SE (1997b) Predictability of a coupled model of ENSO using singular vector analysis. Part II: optimal growth and forecast skill. Mon Weather Rev 125:2057Ð2073. doi: 10.1175/1520-0493(1997)125<2057:POACMO>2.0.CO;2

Yang S-C, Cai M, Kalnay E, Rienecker M, Yuan G, Toth Z (2006) ENSO bred vectors in coupled oceanÐatmosphere general circulation models. J Clim 19:1422Ð1436. doi: 10.1175/JCLI3696.1

Zebiak SE, Cane MA (1987) A model for El NinoÑsouthern oscillation. Mon Weather Rev 115:2262Ð2278. doi: 10.1175/1520-0493(1987)115<2262:AMENO>2.0.CO;2

Zhang S, Harrision MJ, Wittenberg AT, Rosati A (2005) Initialization of an ENSO forecast system using a parallelized ensemble Þlter. Mon Weather Rev 133:3176Ð3201. doi: 10.1175/MWR3024.1