The retrospective prediction of ENSO from 1881-2000 by a hybrid coupled model – (I) SST Assimilation with Ensemble Kalman Filter

Ziwang Deng, Youmin Tang* and Xiaobing Zhou

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

Submitted to Climate Dynamics

Corresponding author: Youmin Tang, E-mail: ytang@unbc.ca

ABSTRACT

In this study, we assimilated sea surface temperature (SST) data of the past 120 years into an oceanic general circulation model (OGCM) for El Niño-Southern Oscillation (ENSO) retrospective predictions using Ensemble Kalman Filter (EnKF). It was found that the ensemble covariance matrix in EnKF can act as a time-variant transfer operator to project the SST corrections onto the subsurface temperatures effectively when initial perturbations of ensemble were constructed using vertically-coherent random fields. As such the increments of subsurface temperatures can be obtained via the transfer operator during assimilation cycles.

The results show that the SST assimilation improves the model simulation skills significantly, not only for the SST anomalies over the whole assimilated domain, but for the subsurface temperature anomalies of the upper 100m over the tropical Pacific off the equator. Along the equator, the improvement of the assimilation is confined within the mixing layer because strong upwelling motions there prevent the downward transfer of SST information. The retrospective prediction skills of ENSO over the past 120 years from 1881-2000 were significantly improved by the SST assimilation at all leads of 1-12 months, especially for the 3-6 months leads, compared with those initialized by the control run without assimilation. The skilful predictions by the assimilation allow us to further study ENSO predictability using this coupled model.

1. Introduction

As a critical component of global change, climate variability is potentially one of the most serious environmental issues we face today. A prominent example is the famous ENSO phenomenon, the Earth's strongest source of interannual climate variability. ENSO is centered in the tropical Pacific but has great impact worldwide. It has been intensively studied and significant progress in prediction has been made over the past few decades (e.g., Latif et al. 1998; Shukla et. al. 2000; Goddard et. al. 2001).

Some important and challenging issues still remain for the ENSO prediction community. One specific issue is to estimate the uncertainties of ENSO prediction and predictions of ENSO-induced climate variability. Currently there are a few ENSO prediction models to issue routine predictions (see ENSO prediction plume by the International Research Institute for Climate and Society: http://iri.columbia.edu/climate/ENSO/currentinfo/update.html). These models differ in terms of their initialization schemes and complexity in physics and dynamics, ranging from simple statistical models to fully coupled general circulation models. Though some models display very good hindcast skill (e.g., Luo et al. 2005; Kirtman et al. 2007; Chen et al. 2004), considerable models have a similar predictive skill level and display no statistically significant difference in the predictive capability. Also we have to admit that the advancement in improving ENSO predictive skills in recent years is limited although a large number of efforts were made. These facts suggest that ENSO predictive skills might have reached an upper limit.

A widely used strategy for the study of ENSO predictability is to perform retrospective predictions using dynamical models, by which the variations of predictability and the causes responsible for the variations are examined. So far, most studies of ENSO predictability typically performed hindcasts for a period around 20-30 years due to the limited length of observations (e.g., Moore et al. 1998; 2006; Tang et al 2005). The period available to test predictability covers few ENSO cycles, roughly 10-15.

Using only data from this short period limits confidence about conclusions of such studies. Chen et al. (2004) performed a retrospective prediction of the past 148 years from 1856-2003 using the Lamont model. To date, that is the only work which has explored ENSO predictability for over the last 100 years.

The motivation for the present study stems from the fact that ENSO prediction skill is model dependent. Thus an important step towards fully understanding ENSO predictability is to perform long-term retrospective predictions using more models. In this study, we developed a hybrid couple model to perform long-term ENSO predictions. A very important task in ENSO predictions is to determine the oceanic initial conditions as accurately as possible. They are usually incorporated using data assimilation. It has been found that the assimilation of subsurface in situ observations and satellite altimetry can significantly improve model skills (e.g., Ji. et al 2000; Tang and Kleeman 2002; Tang and Hsieh 2003; Tang et al. 2004). The oceanic subsurface in-situ observations and satellite altimetry are both short. However, records of SST are much longer and thus may be used to initialize ocean models for long-term ENSO hindcast.

This paper presents the first part of our predictability study. In this paper, we will address how to assimilate SST into an OGCM to benefit ENSO simulation and ENSO prediction. Usually SST is a prognostic variable in ocean models, thus the general procedure of SST assimilation is to optimally insert it into models (e.g. Chen et. al. 1997; Rosati et al. 1997; Syu and Neelin 2000b; Tang and Hsieh 2003). However this strategy could lead to serious imbalances between the model SST field and other fields during the assimilation process (Tang and Kleeman 2002), leading to failure in initializing ENSO predictions. Therefore, a key issue for SST assimilation is to alleviate the imbalances. By using either statistical methods or well-tuning nudging parameters, some studies showed significant improvement of ENSO prediction skill by SST assimilation (e.g., Oberhuber 1998; Tang and Kleeman

2002; Tang et al. 2004; Chen et al 2004; Keenlyside and Latif et al. 2005; Luo et al. 2005; Zheng et al 2006). However, either nudging or statistical schemes are highly model and tuning-parameters dependent. For example, Tang and Kleeman (2002) and Tang et al (2004) used a statistical relationship to construct the transfer function between the SST and the subsurface temperatures as in the assimilation of sea level height (Carton et al. 1996; Ji et al. 2000). Such statistical relationships are usually derived from either model control runs or observations. Thus the transfer function, which is invariant during all assimilation cycles, is highly dependent on the model and the period used for constructing the statistical relationship. The invariant nature of the transfer function in time is a problem since it should depend on the evolution of model states and assimilation cycles.

The study of SST assimilation also has practical significance in other aspects. The availability of data is probably the most important issue in ocean data assimilation. Until recently the oceanic observations, especially subsurface observations, were considered spatially sparse and temporally sporadic (McPhaden et al. 1998). At the end of 2000 a new global oceanic observing system called Argo, was lunched to more completely sample the upper ocean. However Argo is still at an early stage and it can not yet be used to generate long-term oceanic analyses. Satellites have provided high spatial and temporal resolution SST data for over 25 years and in situ SST analyses are available for extended periods, so it is important to explore efficient schemes to analyze the oceanic subsurface by assimilating SST observations.

In this study, we assimilated SST historic data for the period 1881-2000 into an OGCM using Ensemble Kalman Filter (EnKF). The emphasis will be placed on developing a general method to construct the transfer function between the correction in SST and that in subsurface temperature. As argued above, the key point for SST assimilation is to effectively correct subsurface temperatures by the transfer function. In later sections we will show how EnKF can be used to construct a time-variant transfer function by using ensemble members at each assimilation cycle.

This paper is organized as follows: the data and model used is described in section 2. The EnKF algorithm is presented in section 3. The oceanic analysis from the SST assimilation is discussed in section 4. Hindcast experiments for the 120 years from 1881-2000, initialized by the SST assimilation, are reported in section 5, followed by a summary and discussion.

2. Data and Model

2.1 Data

Presently there are at least 4 long-term (over 100 years) global SST datasets available: including the United Kingdom Met Office (UKMO) Hadley Centre Sea Ice and SST dataset (HadISST; Rayner et al. 2003); the Lamont-Doherty Earth Observatory reconstruction by Kaplan et al.(1998) , the Extended Reconstruction SST versions 1(ERSST.v1) and version2(ERSST.v2) by the National Oceanic and Atmospheric Administration (NOAA, Smith and Reynolds 2002; 2003; 2004). In this study we used the ERSST.v2 SST data from 1878-2002, with 2^o lat. by 2^o lon. resolution. This bias corrected dataset has been used for studying climate variation and prediction (e.g., Xue et al. 2003; Nakaegawa et al. 2004; Monahan 2001; Monahan and Dai 2004). Due to the relatively poor quality of the data set prior to 1878, we focus on the period from 1881 - 2001 in this study. For the validation of SST assimilation, the NCEP reanalysis subsurface temperature from 1981-2000 is also used in this study (Behringer et al. 1998; referred to as NCEP data hereafter). The data domain is confined in the tropical Pacific Ocean.

The monthly zonal wind U_{10} and meridional wind V_{10} at the 10 meter level, on the T62 Gaussian grid during the period from January 1948 to December 2002, is from the NCEP-NCAR reanalysis data set (Kalnay et al. 1996;Kistler et al. 2001). The monthly wind stresses were calculated by a method similar to that in Wittenberg (2004), as below:

$$\tau = \rho C_D \| \mathbf{U_{10}} \| \mathbf{U_{10}} \qquad (1)$$

where $\mathbf{U_{10}}$ is the vector wind at the 10m above the sea surface, air density $\rho = 1.2 kg/m^3$, and drag coefficient C_D=1.53×10^{-3}. To perform a long-term control run with the ocean model, the wind stresses prior to 1948 were reconstructed using a statistical model. The empirical atmospheric model using SST to predict wind has been widely used in ENSO studies (e.g., Syu and Neelin 2000a; Barnett et al. 1993; Tang and Hsieh 2002). It has been found that the statistically estimated wind by SST can produce better prediction skills than the observed wind when both are used to initialize predictions (Syu and Neelin 2000b; Tang and Hsieh 2002). Thus, the long-term wind stress dataset is reconstructed with historical SST dataset using a singular vector decomposition (SVD) model (Deng et al. 2008).

To keep the wind stress used for the OGCM consistent through the whole period from 1881-2000, the wind stress from 1948-2000 was also reconstructed by means of a cross- validation scheme (Deng at al. 2008, see also sec. 2.3). The monthly Niño3, Niño3.4 and Niño4 sea temperature anomaly (STA) indices are defined as the STA averaged over 5°S-5°N, 90-150°W, 5°S-5°N, 120-170°W and 5°S-5°N, 160°E-150°W respectively. Forced by the reconstructed wind stress, the ocean model generated a very good SSTA simulation over the tropical Pacific ocean for the period from 1881-2000, with the correlation coefficient over 0.7 between modeled and observed Niño3.4 SSTA indices (Deng et al. 2007).

2.2 Ocean model

The ocean model is the latest version of Ocean Parallelise (OPA9.0), a primitive equation OGCM widely used in oceanic studies (e.g., Tang et al. 2004; Luo et al. 2005; Moore et al. 2006). Detailed formulation of the ocean model is described in Madec et al. (1998). The domain of the model used here is the Pacific Ocean between 30°N and 30°S and between 122°E and 70°W, for a total of 85×49 horizontal grid points. The horizontal resolution in the zonal direction is 2°, while the resolution in the meridional direction is 0.5° within 5° of the equator, smoothly changing up to 2.0° at 30°N and 30°S. There are 31 vertical levels, with 17 concentrated in the top 250 m of the ocean. The model is integrated with closed boundaries using a time step of 1.5 hours.

The model was first spun up for 500 years using climatological wind stress derived from the 50-year NCEP Reanalysis wind stress and the heat flux Q_s

$$Q_s = Q_0 + \lambda(T - T_0) \qquad (2)$$

where Q_0 is the climatological heat flux, obtained from the European Centre for Medium-Range Weather Forecasts (ECMWF) reanalysis project for the base period 1971-2000. T is the model SST, T_0 is Levitus' observed climatological SST (Levitus 1998), and λ is the relaxation rate, set to $-40Wm^{-2}K^{-1}$ (Tang et al. 2004; Moore et al. 2006). For a 50 m mixed-layer depth, this value corresponds to a relaxation time scale of two months (Madec et al. 1998).

2.3 Atmosphere model

The atmosphere component of the hybrid coupled model (HCM) is constructed by the SVD method using observed SST and wind stress data from 1948-2002. The cross-validation scheme was used in the construction of the atmospheric model to ensure that the data used for validation was never used in training the model. In details, three atmosphere models, denoted by AM1, AM2 and AM3, were respectively constructed using the data from 1948-1975, 1977-2002 and 1949-2002. The AM1 was used for the hindcast for 1976-2000, the AM2 for 1948-1975, and AM3 for 1881-1947. The corresponding HCMs are denoted by HCM1, HCM2 and HCM3 respectively. First three SVD modes were used for constructing the statistical atmospheric model, similar to Harison et al. (2002), Galanti et al (2003), Tang et al. (2004), and Moor et al. (2006). The HCM can simulate the primary characteristics of ENSO and has a useful prediction skill as shown below.

2.4 Coupled model and ENSO simulation

After an atmospheric model has been constructed, coupling between the atmosphere and the ocean can be implemented. The SST anomalies from the ocean model are projected onto the first three SVD modes to extract the predictors. An atmospheric model (AM1, AM2 or AM3) is then used to reconstruct the zonal wind stress, and a similar one is used for the meridional component. As common with statistical models, the variance of the predicted variables are lower than the variance in the observed variables, hence the wind-stress estimates are scaled up to their observed variance by an adjusting scale factor (i. e, coupling coefficient), as in some previous work (e.g., Barnett et al. 1993; Tang et al. 2002). The reconstructed wind stress anomalies are then added to the climatological wind stress field to force the ocean model. The coupling interval is one day.

Fig. 1 shows a 120-year simulation of Nino3 SSTA index, and its local and global wavelet power spectrum (Torrence and Compo, 1998), derived from a 130-year coupling run with HCM1 starting from a randomly chosen time. The first 10-years of the coupled run were removed here because the coupled model requires several years to reach equilibrium, and the impact of initial condition should be removed. As can be seen, the model exhibits an apparent ENSO-like interannual variability with a significant period of around 33 months. This coupled model has a realistic phase-locking phenomena and the peaks of oscillation tend to occur during late fall and winter (not shown). Compared with observed ENSO oscillation, the model has a relatively weaker amplitude and shorter period. In addition, the period seems too regular during the whole period of the 120-year simulation as shown in Fig. 1b. These unrealistic features were also found in other HCMs (e.g., Kirtman and Zebiak

1997; Syu and Neelin 2000a; Tang 2002), probably due to the simplification of the atmospheric model.

The coupling experiments were also performed for HCM2 and HCM3, and similar results were obtained (not shown).

3. Ensemble Kalman Filter (EnKF) SST assimilation scheme

3.1 EnKF Algorithm

As a novel and flexible technique of data assimilation, EnKF requires no derivation of a tangent linear operator or adjoint equations, no integrations backward in time, and the computational requirements are affordable and comparable with other popular sophisticated assimilation methods (Evensen, 2003; 2004). Since it was first proposed by Evensen (1994), EnKF has been widely used for atmospheric and oceanic data assimilation (e.g., Anderson,2001; Whitaker and Hamill, 2002;Snyder and Zhang, 2003; Keppenne and Rienecker, 2003; Annan et al. 2005; Hargreaves et al. 2004; Houtekamer et al. 2005; Zhang et al. 2005,2006; Bowler, 2006; Zheng et al. 2006).

The standard EnKF analysis scheme can be expressed as follows (Evensen, 2003; 2004; Kepert, 2004):

$$\mathbf{X}^a = \mathbf{X}^f + \mathbf{P}_e^f \mathbf{H}^T (\mathbf{H}\mathbf{P}_e^f \mathbf{H}^T + \mathbf{R}_e)^{-1} (\mathbf{d} - \mathbf{H}\mathbf{X}^f) \quad (3)$$

Here **X** denotes the sea temperature anomalies of the upper 17 layers, the superscript *a* is for the analysis, and superscript *f* for first guess, **H** is measurement operator, **d** is a L-dimension vector to denote observations (L=71×16, the number of grids of SST observations), $\mathbf{R}_e$ is a $L \times L$ observation error covariance matrix, $\mathbf{P}_e^f$ is the ensemble covariance matrix.

Theoretically $\mathbf{R}_e$ could be estimated using ensemble members if the ensemble size is sufficiently large and comparable with L (Evensen 2004). In practice such a large ensemble size is often unaffordable. In this study, we set the ensemble size as 26, far less than L. With such limited ensemble members, the estimated $\mathbf{R}_e$ is of reduced rank and $(\mathbf{H}\mathbf{P}_e^f\mathbf{H}^T+\mathbf{R}_e)$ becomes a singular matrix. This makes Eq (3) unsolvable (Kepert 2004). Thus we will not use ensemble members to obtain $\mathbf{R}_e$ here. Instead we assume that the observation errors of different grids are independent as in Derber and Rosati (1989), Ji et al. (2000) and Tang et al. (2004) etc.. With this assumption, the $\mathbf{R}_e$ is a diagonal matrix with diagonal elements being the variance of observation error and zero everywhere else[1]. The variance of observation error of diagonal elements was set to a constant for simplicity of computation as in some other works (e.g., Derber and Rosati 1989; Ji. et al 2000; Tang et al. al 2003; 2004).

Motivated by the fact that the ocean model climatology is somewhat different from that of observed data, direct assimilation of SST may produce a large shift in the model climatology (e.g., Chen et al. 1995; Galanti et al. 2003). On the other hand, the data assimilation theoretically assumes model and observation errors to be unbiased, implying that the quantity assimilated is observation anomaly. Hence, in this study we actually assimilate observed SST anomalies into the ocean model, although the term of “SST assimilation” is also used alternatively in this paper. Since there are no long-term subsurface temperature observations, special care is taken to effectively transfer corrections in SST to the subsurface temperatures. This can be implemented using the ensemble covariance matrix $\mathbf{P}_e^f$ that is gained from ensemble members (Eq A1). In the Appendix, we find that the $\mathbf{P}_e^f$ can act as a time-variant transfer matrix to project corrections of SST onto the subsurface temperatures.

---

[1] In this study, it is set to $(0.25℃)^2$. It should be noted that the error variance is often assumed to be proportional to the signal variance. Thus it might be more reasonable to produce the variance of observation error based on the signal variance, as shown in Fig. 1.

The assimilation domain covers the tropical Pacific from 140°E to 80°W and 15.85°S to 15.85°N in horizontal and in the upper 17 levels (250 m), with a total number of 71×31×17 grid points. The assimilation was performed by every 5 days.

3.2 Perturbation construction

The success of EnKF highly depends on the quality of ensemble members produced by initial perturbations. Since it is impractical to represent all possible types of errors within the ensemble because of the computational cost, the method of generating initial perturbations must be chosen judiciously. Usually the following two factors are considered when selecting the amplitude of initial perturbations: the amplitude of observation error and the amplitude of model errors induced by model parameters and uncertainty in model physics. If a model is perfect, the amplitude of the perturbations should be the same as the amplitude of observation errors. However the model used in this study has larger errors in the eastern tropical Pacific along the equator (Tang et al. 2004) similar to other OGCMs, thus we add the model error to the observation errors, as in Fukumori (2001) and Hargreaves and Annan (2004). This combined error is used to determine the amplitude of perturbations.

When the perturbation amplitude is determined, the practical initial perturbation field generating each ensemble member could be constructed by a normalized pseudorandom field multiplied by the prescribed amplitude. Considering the spatial coherence, the pseudorandom field is red-noise as proposed by Evensen (2003). We also considered the vertical coherence of pseudorandom fields between adjacent levels using the below method:

The pseudorandom field at the *kth* level $\boldsymbol{\varepsilon}_k$ (k=1,2,3,…,L) is constructed by following equation:

$$\boldsymbol{\varepsilon}_k = \alpha \boldsymbol{\varepsilon}_{k-1} + \sqrt{1-\alpha^2}\mathbf{W}_k \qquad (4)$$

where $\mathbf{W}_k$ (k=1,…,L) is the pseudorandom field at the *kth* level without considering vertical coherence, constructed using the method of Evensen (2003). Initially, for the surface perturbation (k=1), the vertical coherence is not considered, i.e., $\alpha = 0$, since $\boldsymbol{\varepsilon}_{k-1}$ does not exist. Eq (4) indicates that a practical pseudorandom at the *kth* level ($\boldsymbol{\varepsilon}_k$) is composed of $\mathbf{W}_k$ and $\boldsymbol{\varepsilon}_{k-1}$. As such the $\boldsymbol{\varepsilon}_k$ is correlated with $\boldsymbol{\varepsilon}_{k-1}$, i.e., the practical pseudorandom fields of two adjacent levels ($\boldsymbol{\varepsilon}_{k-1}$ and $\boldsymbol{\varepsilon}_k$) are coherent with each other. Their correlation or coherent structure is determined by the coefficient $\alpha \in [0,1]$. Eq (4) generates a sequence that is white in the vertical direction if $\alpha = 0$ (i.e., $\boldsymbol{\varepsilon}_k = \mathbf{W}_k$), but a sequence that is perfect correlated in vertical if $\alpha = 1$ (i.e., $\boldsymbol{\varepsilon}_k = \boldsymbol{\varepsilon}_{k-1}$). As indicated in Tang (2004), the sea temperature variation in two neighbor levels is highly correlated for the upper ocean region. With the correlation analysis of the sea temperature from NCEP reanalysis dataset and some sensitivity experiments with the ocean model used in this paper, we set $\alpha = 0.95$ in this study for the top 5 levels (around the depth of 50m). Thus in this study, the coherence in both horizontal and vertical is considered in the construction of initial perturbations.

The covariance matrix $\mathbf{P}_e^f$ was estimated using 26 ensemble members, obtained from 13 pairs of positive-negative perturbations of initial conditions. A 5-day time integration was used for the ensemble because the assimilation was performed every 5 days. Such a scheme of ensemble generation has been used in both ECMWF (Molteni et al. 1996) and NCEP (Toth and Kalnay 1993, 1997) for their operational ensemble predictions.

One important question is: how many levels should be perturbed initially to construct ensemble members? If the number of levels is too small, the constructed ensemble members might not be able to generate an effective connection between the different levels, leading to a very small

$\mathbf{P}_e^f$, whereas if the number of levels is too large, perturbations onto deep levels would probably lead to unrealistic model states since the ocean hardly adjusts to a stable state in an assimilation cycle of 5 days. This situation arises especially when we do not have deep level observations to adjust ocean states. With sensitivity experiments of the number from 1 to 17, we found that the perturbation onto the top 5 levels produced the best results for ENSO prediction skills. Thus, we perturbed the top 5 levels in this study. The depth of the top 5 levels covers the mixed layer of the equatorial Pacific, allowing the model state to quickly respond to any perturbations.

4. Oceanic analysis

In this section, we will examine the oceanic analysis from the SST assimilation. The central issue addressed here is how the subsurface temperature responds to the SST assimilation. Prior to a detailed investigation of subsurface oceanic analysis, we will first look at the surface temperature analysis.

4.1 SSTA Analysis

Figure 2 shows anomaly correlation and RMSE skills between observed and modeled SSTA for the period from 1881-2000, obtained respectively from the control run and the assimilation run. As can be seen, the control run has a good SSTA simulation in the central and the eastern equatorial Pacific. The large RMSE resides in the equatorial eastern Pacific due to large skewness and variance there. As the performances of some OGCMs, forced by observational wind stress, the OPA9.1 always underestimates the amplitude of SSTA in the central eastern Pacific, and it is more serious in boreal fall-winter when the SSTA variance is large. The comparison of simulated and observed indices (nino3, nino3.4 and nino4) clearly indicates this phenomenon (not shown). The ratios between the standard deviations of observed and simulated nino3, nino3.4 and nino4 are 2.39, 1.76 and 1.45 respectively. Comparing the control run with the assimilation reveals that the assimilation leads to much better SSTA simulation for the whole tropical Pacific. For example, the RMSE attained by the assimilation is only about half of that by the control run. Also the maximum RMSE center near the eastern boundary in the control run disappears in the assimilation. A good skill of SSTA simulation attained by the assimilation is not surprising since we measure the simulation skill using SSTA which was assimilated into the ocean model. As discussed in the introduction, the critical issue in the SSTA assimilation is the correction in subsurface temperature, which we will explore below.

4.2 Subsurface temperature Analysis from 1981-2000

In this section, we will examine the impact of SST assimilation on the subsurface temperatures. The NCEP reanalysis subsurface temperature from 1981-2000 is used in lieu of subsurface in situ observations (e.g., TAO and XBT) for verification. This is because NCEP reanalysis product is much easier and more convenient to use compared with sparse and sporadic subsurface in situ observations (McPhaden et al. 1998). Also it is able to effectively improve the predictions for Niño3 SSTA at all lead times and generate as good oceanic analyses as in situ observations (Tang et al. 2003). For the purpose of examining the impact of SST assimilation on the subsurface temperature, we will address here the difference between the assimilation and control runs.

The correlation and RMSE skills of several El Niño indices, as a function of depth, are obtained for control and assimilation runs respectively, both against the NCEP reanalysis dataset, as shown in Fig. 3. The vertical bars in Fig. 3 measure the extent to which the correlation and RMSE might change due to sampling errors, which were obtained using the bootstrap method (von Storch and Zwiers, 1999). The length of the bar is the twice of standard deviation of bootstrapped samples of correlation and RMSE, representing an upper bound of sampling errors. As can be seen, the improvement in correlation and RMSE skills by the assimilation decreases with the depth and mainly occurs in the upper ocean of 100-150 m, below which the effect of assimilation is hardly visible. Comparing the skills of different indices reveals that the assimilation most significantly improves the sea temperature simulation over Niño4 region, followed by Niño3.4 and Niño3. The improvement in RMSE skill is more significant than the improvement in correlation skill.

To evaluate the impacts of the SST assimilation on temperature of subsurface (2-17 levels), the correlation and RMSE skills measured against the NCEP reanalysis data for the 20 yr period from 1981 to 2000 were calculated. By subtracting the correlation skill of the control run from that of the assimilation run, the differences of correlation skill between the assimilation and control runs at different levels were obtained. A positive value represents a better skill from data assimilation. From the difference in spatial distribution (not shown), one can see that, at some depths, there are some negative differences, but the positive values dominate over the equatorial Pacific Ocean, especially on the upper levels. If we consider the difference great than 0.2 as significant, based on uncertainty due to sampling errors estimated by the bootstrap method, it is found that the significant improvements appear over regions off the equator where the control run has poor simulation skill. It is clear that the areas of positive difference decrease with the depth, indicating that the downward transfer of SST increments weakens with the depth as shown in Fig. 3. At the depth below 117m, the improvement is weak.

For RMSE skill, similarly, the difference of RMSE between the assimilation run and control run (not shown) indicates the extent to which the skill is improved by the assimilation. Note that a large RMSE means a poor skill, thus a negative value of the difference indicates the improvement of the skill by the assimilation. Similar to the correlation skill, the improvement of RMSE mainly occurs at or near the surface, and weakens with the depth. At the depth of 117m or below, the improvement is small, and mainly in the western equatorial Pacific.

To further explore the impact of SST assimilation over the upper ocean, the total heat content anomaly (HCA) over the upper 10 levels (roughly 100m) is examined. Shown in figure 4 are the correlation and RMSE skill of HCA, against the NCEP reanalysis data, for the assimilation run and the control run, as well their differences. As can be seen, both the control and assimilation runs have a high correlation skill within a narrow belt along the equator. In the eastern tropical Pacific the high correlation extends to higher latitudes on both sides of the equator. Clearly the SST assimilation leads to higher correlation skill in the central tropical Pacific, and smaller RMSE in the central eastern Pacific as shown in Figure 4e and f. Comparing Fig.4b and 4f reveals that the assimilation reduces the RMSE of control run by 15-25% over the equatorial eastern Pacific.

Figure 5 shows the time-longitude diagram of the averaged HCA of the upper 100m over the equator belt from 4°S to 4°N during 1981-2000, displaying that the assimilation significantly improves the HCA simulation. As can be seen, the relatively weak amplitude of HCA in the control run was greatly increased by the assimilation, especially in El Niño and La Niña events.

To further explain the improvement, we will examine the subsurface temperature anomalies for two specific cases: the mature phase of 1988-1989 La Niña (January 1989) and the mature phase of 1997-1998 El Niño (January 1998). For the following discussion, we define a statistical metric at each model grid to quantify the contribution of assimilation (CA) at this grid as follows:

$$CA = \left|STA_{NCEP} - STA_{ASM}\right| - \left|STA_{NCEP} - STA_{CTL}\right| \qquad (5)$$

where STA_{NCEP}, STA_{ASM} and STA_{CTL} are respectively the NCEP reanalysis temperature anomaly, the analyzed temperature anomaly of assimilation run and control run. The CA measures the difference of the absolute error of sea temperature anomalies between the assimilation and the control, against the NCEP reanalysis. A negative value of CA indicates that the analyzed temperature anomalies of the assimilation are closer to the observation than the temperature anomalies from the control run, and vice versa.

The results (not shown) show that, at January 1989 and January 1998, compared with NCEP reanalysis data, the STA simulation is considerably better in the assimilation run than in the control run in the upper ocean and in the central and the eastern Pacific. The amplitude of STA of the upper ocean 100m is much larger and more realistic in the assimilation than in the control run. In both cases, for the deeper ocean below 106m, the assimilation could only lead to slight improvement in STA simulation, mainly in the region off the equator.

4.3 Subsurface temperature Analysis from 1881-2000

In this subsection, we will explore the impact of the SST assimilation on STA over the whole period of 120 years from 1881-2000. Since we do not have the subsurface observations for such a long period, emphasis will be placed on the difference in STA between the control run and the assimilation run. We use the root mean square of the difference (RMSD) to measure the 120-yrs average impact of data assimilation on STA simulation, defined as below:

$$RMSD = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(STA_i^{CTL} - STA_i^{ASM})^2} \qquad (6)$$

Where STA_i^{CTL} is the STA from the control run without assimilation, STA_i^{ASM} is the analyzed STA with the assimilation, and the N=480 is the sample size used (January, April, July and October of each year). The *RMSD* reflects the impacts of the assimilation on STA simulation. The spatial distribution of RMSD at several chosen depths is shown in Figure 6. As can been seen, the prominent differences in STA exist between the assimilation and the control run from the surface to the depth of around 200 m, and the regions where the RMSD is larger than 0.5℃ mainly appear in top levels. The amplitude of RMSD varies with depth with the maximum value appearing at 25m.

Z_{mix}

Z_{20}

$$SST - T(Z_{mix}) = 0.5°C \qquad (7)$$

Fig. 7 manifests the features shown in Fig. 6. For example, the maximum RMSD generally appears between the mixing depth and thermocline depth in the eastern Pacific. Above the mixing depth or below the thermocline depth, the RMSD is relatively small. One reason for this is that only the upper 5 model levels are perturbed in this assimilation, as noted in the section 3.2. Below that, adjustments are due to a combination of assimilation vertical covariance and model vertical dynamic adjustments. As shown by Fig. 6 and 7, in some regions the influences of assimilation reaches deeper than in others. To explore this further, we examine the vertical motion on the equator and off the equator. Figure 8 shows the model climatological vertical velocity along 4^{o}S, the equator and 4^{o}N, averaged over 120 years from 1881-2000 of the assimilation run. Fig. 8b indicates that the average vertical advection is dominated by an upwelling motion along the equator, mainly from 160^{o}E to 90^{o}W, at the depths from the surface to 200 m. This is similar to that obtained in Zelle (2004) with the Hamberg Ocean Primitive Equation Model. As shown in Fig. 8b, the upwelling is the strongest between the mixing and the thermocline depth. The dominant upwelling process limits the downward transfer of the observed SST information into the subsurface in the assimilation. As shown in Eqs. (3) and A7, the downward transfer of SST increment information is via the covariance matrix $\mathbf{C}_{k,1}$ constructed by the perturbations of different levels. The $\mathbf{C}_{k,1}$ depend on two terms, the perturbation of surface and the perturbation of kth level, the amplitude of the former is mostly associated with the amplitude of initial perturbations and the later is determined by both the former and the downward information transfer in the vertical direction. Along the equator, in the mixed layer, mixing is the strongest process whereas between the mixed layer and the thermocline the upwelling is the strongest process (Zelle 2004). Due to the strong upwelling motion, the perturbation on the surface can not effectively transfer to the subsurface, thus the matrix $\mathbf{C}_{k,1}$ is very small, which may explain why there are relatively small values of RMSD troughs along the equator in Figure 6d-f. On the other hand, downward Rossby waves in the off-equatorial region can help the downward transfer of the perturbation acted on the surface, leading to the adjustment of subsurface temperature during the assimilation.

In summary, the SST assimilation with EnKF can transfer the increments in SST into the subsurface by means of the covariance matrix that connects perturbations of different levels. Such a downward transfer of information is most effective off the equator where downward Rossby waves are important and downwelling dominates. Along the equator upwelling limits the downward propagation of information, limiting improvements of equatorial STA from assimilation to roughly the upper 30-50m. Thus the SST assimilation with EnKF mainly benefits the improvement in the subsurface temperature off the equator due to the important role of vertical motion in the SST assimilation.

5. Hindcast experiments

The primary purpose of SST assimilation in this study is to provide good initial conditions for model predictions. The prediction performance is thus a valid test bed to evaluate the assimilation scheme. In this section, we will examine ENSO predictions initialized by the SST assimilation. For comparison, the predictions initialized by the control run and persistence are also evaluated. The whole period of 1881-2000 was used for prediction experiments. A total of 480 hindcast cases of 12 months duration were made from January 1881 to December 2000, starting every 3 months from 1 January, April 1, July 1 and October 1. As stated in section 2.3, atmospheric models used for wind stress reconstruction and coupled with the ocean model were obtained using a cross-validation scheme. As such, the data used for hindcast was never used in training these models. However, , there are still probably artificial components in prediction skills since several model parameters are turned without considering cross-validation such as the coupling coefficient as mentioned in section 2.4 and the number of perturbation level in data assimilation.

Fig. 9a shows correlation skills of the predictions initialized by the control run, the assimilation and the persistence, where the predicted Nino3.4 SSTA index is compared against the observed value. The predictions initialized by the assimilation run beat the persistent predictions beginning the second month whereas the predictions initialized by the control run beat persistent preditions beginning the seventh month. The prediction skills from the control run are far lower than those from the assimilation up to lead times of 10 months, indicating the importance of initial conditions on ENSO predictions. The RMSE skill shows similar characteristics (Fig.9b).

Compared with the control run, the assimilation improved the prediction skills for lead times of 1-9 months. The most obvious improvement occurs for the prediction of lead times of 3-6 months, with high correlation skills around 0.5-0.8. Such correlation skills of hindcast over a 120 year period are comparable with those in the Lamont prediction model (Chen et al. 2004; Tang et al. 2007).

Likewise, the SST assimilation also displays significant improvements in RMSE for all lead times (Fig. 7b), in particular for the lead times of 3-6 months. These results are also obtained using the Nino3 SSTA index to measure ENSO prediction skills (not shown). The hindcast skill achieved here justifies further ENSO predictability studies and shows that the SST assimilation has practical significance.

It is interesting to compare our model prediction skill with those of some operational forecast models. With the period from 1981-2000 as the comparison bed, our model has an around 0.6 correlation skill and 0.75℃ RMSE skill at the lead time of 6 months. Compared with some operational forecast models such as the LDEO5 model (Chen et al. 2004), COLA model (Kirtman 2003) and ECMWF model (Palmer and Hagedorn 2006), our model has a comparable correlation skill but a worse RMSE skill. At the 6-month leads, these operational prediction models have a RMSE skill of 0.4-0.6℃.

Fig. 10 shows the correlation and RMSE skills of the predicted SSTA against the observed SSTA at the 6 month lead for the period 1881-2000. The prediction skills from the control run (upper panels) and the assimilation (bottom panels) share many features. For example, the good correlation skills mainly reside in the central and eastern Pacific and decrease toward the west and higher latitudes, as in other ENSO models (e.g. Barnett et al. 1993; Chen et al. 1997). However, the prediction skills initialized from the assimilation are considerably better than those from the control run, especially for correlation skill.

The relationship between the initial conditions of predictions and the predicted SSTA are next evaluated. This relationship could shed lights on the possible mechanism responsible for the improvement of prediction skills by the assimilation. We correlated respectively the predicted Niño3 SSTA index of the 6-month lead time with the STA of the initial conditions from the assimilation and the control run. Fig. 11c is obtained by correlating the observed STA of the initial time of predictions with the observed Niño3 SSTA index of 6-month later.

Fig. 11a and 11b were obtained using 480 predictions from 1881-2000 whereas Fig. 11c was derived from NCEP data from 1981-2000. Comparing Fig. 10a, b and c reveals that the correlation pattern from the assimilation is closer to that from the observation. At the depth from 5m to 105m, the assimilation and observation generate a relatively broad positive correlation area in the central and eastern Pacific whereas the positive correlation area is narrow and extended to the western Pacific in the control run. The strong positive correlation of the western Pacific in the control run is absent in both the assimilation and observation whereas strong positive correlation near the eastern boundary in the assimilation and observation is also absent in the control run. Thus the impact of initial conditions on the predictions is more realistic in the assimilation than in the control run, which might explain why the SST assimilation can improve predictions.

6. Summary and conclusions

A widely used strategy to study ENSO predictability is to perform skilful long-term retrospective predictions. To achieve this, one needs to produce oceanic initial conditions as accurately as possible by data assimilation. In the first part of the ENSO predictability study, we focused on the assimilation of SST that is the only available oceanic data over 100 years to improve ENSO predictions in this paper. This provided a basic foundation to further study ENSO predictability, which will be presented in the second part of our ENSO predictability study (Deng et al. 2008).

One key issue in the SST assimilation is how to transfer SST information to the subsurface since the downward transfer of observation information by the adjustment of the model itself is not sufficient (e.g., Tang and Kleeman 2002; Tang et al. 2003). The previous studies on SST assimilation solved this problem by using either statistical relationships to speed up the downward transfer or by using nudging schemes of empirically-determined tuning parameters. Both methods are highly model and time-period dependent and difficult to use for improving long-term prediction skills over a 100 years.

Theoretically measurement operators can project observations onto any model variables in all algorithms of data assimilation if their relations are known. Unfortunately it is impractical in many cases to know the relationship. For EnKF however, the relationship between observations and model variables can be optimally estimated by ensemble members that are generated by well-designed perturbations, as demonstrated in this paper.

Another advantage of EnKF is that it can produce a dynamical transfer matrix to project SST correction onto the subsurface, since different assimilation cycles have different ensemble members, leading to a time-variant transfer matrix. Compared with the stationary transfer operators used in previous work (e.g., Tang et al. 2002; 2003; 2004), a dynamical transfer operator is much more realistic.

Using the designed EnKF, we assimilated the historical SST data of 120 years into an OGCM. The results show that the assimilation improves the simulation skills, not only for the SSTA over the whole assimilated domain, but for the subsurface temperature anomalies of the upper 100m over the tropical Pacific off the equator. Along the equator, the improvement of the assimilation is confined within the mixing layer because the strong upwelling motion below the mixing depth prevents downward transfer of the SST corrections. The downward transfer of SST corrections decays with the depth, leading to a decrease in improvement with depth. In the surface and near surface, the maximum differences of assimilation and control run appear at or near the equator, whereas in the subsurface the maximum differences appear at 3-5° off the equator and move westward with the depth.

One important result of the SST assimilation in this study was the production of skilful retrospective predictions for the period 1881-2000. The prediction skills initialized by the assimilation were significantly improved compared with those initialized by the control run without assimilation. Verified by predictions of 20 years from 1981-2000, the correlation and RMSE skill from the SST assimilation is, in fact, very comparable with the skill from the joint assimilation of surface and subsurface data (not shown). The overall skill of the hindcast from 1881-2000 is similar to Chen et al. (2004). The skilful predictions by the assimilation justify further investigation of ENSO variability and predictability.

Another important result of this work is the demonstration of the practical value of assimilating SST. The satellites consistently provide complete and relatively long surface temperature observations at a very high resolution. It is therefore important to explore efficient schemes to analyze oceanic subsurface states using surface data. The current study provides a useful scheme for assimilating SST into oceanic models. It also suggests an effective method for the assimilation of the sea-surface altimetry data. The information of sea-level height can be directly transferred to model temperatures by the derived dynamical transfer matrix by ensemble members as discussed in this study. This could improve the assimilation of sea-surface altimetry data.

One key point in designing EnKF is the construction of initial perturbations. In this study, we perturbed the sea temperatures of the top 5 levels. We also tested perturbations over less/more levels such as the top 3, 10 and 17 levels. These sensitivity experiments show the perturbation of the top 5 levels, which cover the mixing layer, leads to the best ENSO prediction skill. In addition, the wind stress, which drives the ocean and determines the model’s vertical structure, plays an important role in SST assimilation.

In this study, we only considered the downward transfer from the surface to subsurface, and ignored the transfer between neighbouring levels as well the upward transfer from subsurface to surface. To consider all possible vertical transfers, further EnKF studies are ongoing.

Acknowledgements

We are grateful to Dr.Tom Smith for detailed comments on the original manuscript. This work was supported by Canadian Foundation for Climate and Atmospheric Sciences (CFCAS) Grant GR-523 and the Discovery grant of Natural Sciences and Engineering Research Council (NSERC) of Canada.

APPENDIX: Estimation of the linear increment transfer operator

In this study, we consider the sea temperatures of the top 17 model levels. Thus $\mathbf{X}$ in (3) can be written $\mathbf{X}=[\mathbf{x}_1\mathbf{x}_2...\mathbf{x}_\mathbf{k}...\mathbf{x}_{17}]'$, where $\mathbf{x}_\mathbf{k}$ is the *kth l*evel's temperature field. Denoted by m $(m=71\times31)$ the number of spatial grids of assimilation domain at a level, $\mathbf{x}_\mathbf{k}$ is an *m*-dimensional vector. $\mathbf{P}_e^f$ in (3) is a covariance matrix between grids, which is an $M\times M$ ($M=17\times m$) 2-dimentional matrix. We write $\mathbf{P}_e^f$ using sub-matrices, the covariance between grids of two levels and the covariance between grids of the same level, namely,

$$\mathbf{P}_e^f=\begin{bmatrix} \mathbf{C}_{1,1} & \mathbf{C}_{1,2} & \cdots & \mathbf{C}_{1,k2} & \cdots & \mathbf{C}_{1,K-1} & \mathbf{C}_{1,K} \\ \mathbf{C}_{2,1} & \mathbf{C}_{2,2} & \cdots & \mathbf{C}_{2,k2} & \cdots & \cdots & \cdots \\ \cdots & \cdots & & & & \cdots & \\ & & & & & & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & & \\ & & & & & & \cdots \\ \mathbf{C}_{k1,1} & \mathbf{C}_{k1,2} & \cdots & \mathbf{C}_{k1,k2} & \cdots & \mathbf{C}_{k1,K-1} & \\ & & & & & & \cdots \\ \cdots & \cdots & & & & \cdots & \\ \mathbf{C}_{K-1,1} & \cdots & & \cdots & & \mathbf{C}_{K-1,K} & \mathbf{C}_{K-1,K-1} \\ \mathbf{C}_{K,1} & \mathbf{C}_{K,1} & & \mathbf{C}_{K,j} & & \mathbf{C}_{K,K-1} & \mathbf{C}_{K,K} \end{bmatrix} \qquad (A1)$$

where sub-matrix $\mathbf{C}_{k1,k2}$ is an $m\times m$ covariance matrix between grids of the k_1th and grids of the k_2th level. When $k_1=k_2$, the sub-matrix denotes the covariance of grids of a same level (diagonal element in (A1). *K*=17 is the number of all levels.

The measurement operator $\mathbf{H}$ in (3) converts the model state variables to the observational variables. Its simplest form is a linear function to interpolate from the model grid to the observational stations, which is often used when the observed variable is the same as the analyzed state variable. In this study, the observed variable is SSTA and the analyzed state variable is STA at 17 levels including SSTA. For a full presentation, $\mathbf{H} = \begin{bmatrix} \mathbf{H}_1 & \mathbf{H}_2 & \dots & \mathbf{H}_k & \mathbf{H}_{k+1} & \dots & \mathbf{H}_K \end{bmatrix}$ is a $L \times M$.matrix where $\mathbf{H}_1$ is a $L \times m$ matrix which interpolates model SSTA from model grids to SSTA on observation grids (L denoting the number of observation grids). $\mathbf{H}_2, \mathbf{H}_3, \dots, \mathbf{H}_K$ are assumed, for simplicity, to be zero matrixes with the same dimension as matrix $\mathbf{H}_1$. Hence we can simplify terms in the equation (3) as follows:

$$\mathbf{P}_e^f \mathbf{H}^T = \begin{bmatrix} \mathbf{C}_{1,1}\mathbf{H}_1 & \mathbf{C}_{2,1}\mathbf{H}_1 & \dots & \mathbf{C}_{K,1}\mathbf{H}_1 \end{bmatrix}^T \qquad (A2)$$

$$\mathbf{H}\mathbf{P}_e^f \mathbf{H}^T = \begin{bmatrix} \mathbf{H}_1\mathbf{C}_{1,1} & \mathbf{H}_1\mathbf{C}_{1,2} & \dots & \mathbf{H}_1\mathbf{C}_{1,K} \end{bmatrix} \mathbf{H}^T = \mathbf{H}_1\mathbf{C}_{1,1}\mathbf{H}_1^T \qquad (A3)$$

$$(\mathbf{d} - \mathbf{H}\mathbf{X}) = (\mathbf{d} - \mathbf{H}_1\mathbf{x}_1) \qquad (A4)$$

Equation (3) could be rewritten as below:

$$\begin{bmatrix} \mathbf{x}_1^a \\ \mathbf{x}_2^a \\ \dots \\ \mathbf{x}_k^a \\ \dots \\ \mathbf{x}_{16}^a \\ \mathbf{x}_{17}^a \end{bmatrix} = \begin{bmatrix} \mathbf{x}_1^f \\ \mathbf{x}_2^f \\ \dots \\ \mathbf{x}_k^f \\ \dots \\ \mathbf{x}_{16}^f \\ \mathbf{x}_{17}^f \end{bmatrix} + \begin{bmatrix} \mathbf{C}_{1,1} \\ \mathbf{C}_{2,1} \\ \dots \\ \mathbf{C}_{k,1} \\ \dots \\ \mathbf{C}_{16,1} \\ \mathbf{C}_{17,1} \end{bmatrix} \mathbf{H}_1^T (\mathbf{H}_1\mathbf{C}_{1,1}\mathbf{H}_1^T + \mathbf{R}_e)^{-1} (\mathbf{d} - \mathbf{H}_1\mathbf{x}_1^f) \qquad (A5)$$

If $\mathbf{C}_{1,1}$ is non-singular, the equation *(A5)* could then be written:

$$\begin{bmatrix} \mathbf{x}_1^a \\ \mathbf{x}_2^a \\ \dots \\ \mathbf{x}_k^a \\ \dots \\ \mathbf{x}_{16}^a \\ \mathbf{x}_{17}^a \end{bmatrix} = \begin{bmatrix} \mathbf{x}_1^f \\ \mathbf{x}_2^f \\ \dots \\ \mathbf{x}_k^f \\ \dots \\ \mathbf{x}_{16}^f \\ \mathbf{x}_{17}^f \end{bmatrix} + \begin{bmatrix} \mathbf{I} \\ \mathbf{B}_{2,1} \\ \dots \\ \mathbf{B}_{k,1} \\ \dots \\ \mathbf{B}_{16,1} \\ \mathbf{B}_{17,1} \end{bmatrix} \mathbf{C}_{1,1}\mathbf{H}_1^T (\mathbf{H}_1\mathbf{C}_{1,1}\mathbf{H}_1^T + \mathbf{R}_e)^{-1} (\mathbf{d} - \mathbf{H}_1\mathbf{x}_1^f) \qquad (A6)$$

Where $\mathbf{I}$ is an identity matrix, $\mathbf{B}_{k,1} = \mathbf{C}_{k,1}\mathbf{C}_{1,1}^{-1} \qquad (A7)$

From (A6), we have

$$\mathbf{x}_1^a = \mathbf{x}_1^f + \mathbf{C}_{1,1}\mathbf{H}_1^T(\mathbf{H}_1\mathbf{C}_{1,1}\mathbf{H}_1^T + \mathbf{R}_e)^{-1}(\mathbf{d} - \mathbf{H}_1\mathbf{x}_1^f) = \mathbf{x}_1^f + \Delta\mathbf{x}_1 \qquad (A8)$$

$$\mathbf{x}_k^a = \mathbf{x}_k^f + \mathbf{B}_{k,1}\,\Delta\mathbf{x}_1 \qquad (A9)$$

Equation (A9) shows that $\mathbf{B}_{k,1}$ is a transfer matrix to project the corrections of SST ($\Delta\mathbf{x}_1$) to the subsurface of the *kth* level to correct its temperatures by $\mathbf{B}_{k,1}\Delta\mathbf{x}_1$.

At each assimilation step, the transfer operator $\mathbf{B}_{k,1}$ can be calculated by ensemble members with equation (*A7*), thus it is a time-variant matrix. Theoretically, when the SST is perturbed with a random perturbation, $\mathbf{B}_{k,1}$ could be obtained. However it was found that if we only perturbed SST, the $\mathbf{C}_{k,1}$ would be very small due to a very small variation of the kth level temperature, leading to very small $\mathbf{B}_{k,1}$. Thus in this study, the initial perturbations are exerted on not only SST but also the temperatures of other upper 4 model levels. From these initial perturbations, the ensemble is constructed. $\mathbf{B}_{k,1}$ that is obtained by ensemble members of the prediction model, characterizes the physical relationship between SST and the subsurface temperature.

Usually the dimension of $\mathbf{C}_{1,1}$ is far larger than the limited ensemble size, easily leading to a singular $\mathbf{C}_{1,1}$. Theoretically it is still possible to compute the pseudo-inverse matrix of $\mathbf{C}_{1,1}$. In practice, we directly use (A6) to avoid computing the inverse matrix of $\mathbf{C}_{1,1}$.

REFERENCES

Anderson, J. L., 2001: An ensemble adjustment Kalman filter for data assimilation. *Mon. Wea. Rev.,* 129, 2884–2903.

Annan, J. D., J. C. Hargreaves, N. R. Edwards & R. Marsh, 2005: Parameter estimation in an intermediate complexity Earth System Model using an ensemble Kalman filter, Ocean Modelling, 8(1–2), 135–154.

Barnett, T.P., M. Latif, N. Graham, M. Flugel, S. Pazan and W. White, 1993: ENSO and ENSO-related predictability: Part 1 - Prediction of equatorial Pacific sea surface temperatures with a hybrid coupled ocean-atmosphere model. J. Climate, 6, 1545-1566.

Behringer, D. W., M. Ji and A. Leetmaa, 1998: An improved coupled model for ENSO prediction and implications forocean initialization. Part I: The ocean data assimilation system. Mon. Wea. Rev., 126, 1013-1021.

Bowler, Neill E. 2006: "Comparison of error breeding, singular vectors, random perturbations and ensemble Kalman filter perturbation strategies on a simple model" Tellus, 58A, 538-548.

Carton JA, Cao X, Giese BS, da Silva AM (1996) Decadal and interannual SST variability in the tropical Atlantic Ocean. Phys Oceanogr 26: 1165–1175

Chen, D., S. E. Zebiak, A. J. Busalacchi, and M. A. Cane, 1995:An improved procedure for El Nin˜o forecasting, Science, 269, 1699–1702.

—, —and M. A. Cane, 1997: Initialization and predictability of a coupled ENSO forecast model. Mon. Wea.Rev., 125, 773–788.

—, M. A. Cane,A. Kaplan, S. E. Zebiak, and D. Huang, 2004: Predictability of El Niño over the past 148 years. Nature, 428,733–736.

Deng Z. and Tang Y. 2008: The retrospective prediction of ENSO from 1881-2000 by a hybrid coupled model – (II) Interdecadal and decadal variations in predictability. Climate Dynamics, submitted.

— , — and Zhou X., 2007: Reconstruction of historic wind stress over the tropical Pacific Ocean for 1856-1948, to be submitted.

Derber, J. and A. Rosati, 1989: A global oceanic data assimilation system. J. Phys. Oceanogr., 19, 1333-1347.

Evensen, G., 1994: Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte-Carlo methods to forecast error statistics. J. Geophys. Res., 99 (C5), 10143–10162..

—, 2003: The ensemble Kalman filter: Theoretical formulation and practical implementation. Ocean Dyn., 53, 343–367.

—, 2004: Sampling strategies and square root analysis schemes for the EnKF. Ocean Dynamics, 54, 39–560.

Fukumori I ,2001: Data assimilation by models Chapter 5. Academic, pp 237-265.

Galanti, E., and E. Tziperman, 2003: A study of ENSO prediction using a hybrid coupled model and the adjoint method for data assimilation. Mon. Wea. Rev.,131,2748-2764.

Goddard, L., S. J. Mason, S. E. Zebiak, C. F. Ropelewski, R. Basher, and M. A. Cane, 2001: Current approaches to seasonal to interannual climate predictions. Int. J. Climatol., 21, 1111-1152.

Hargreaves, J. C., J. D. Annan, N. R. Edwards & R. Marsh,2004: Climate forecasting using an intermediate complexity Earth System Model and the ensemble Kalman filter,Climate Dynamics, 23(7–8), 745–760.

Harrison, M. J., A. Rosati, B. J. Soden, E. Galanti, and E. Tziperman 2002: An evaluation of air-sea flux products for ENSO simulation and prediction, Mon. Weather Rev., 130, 723– 732.

Houtekamer, P. L., H. L. Mitchell, G. Pellerin, M. Buehner, M. Charron, L. Spacek, and B.Hansen, 2005: Atmospheric data assimilation with an ensemble Kalman filter: Results with real observations. Mon. Wea. Rev., 133, 604-620.

Ji, M., R. W. Reynolds, and D. W. Behringer, 2000: Use of TOPEX/Poseidon sea level data for ocean analyses and ENSO prediction: Some early results. J. Climate, 13, 216–231.

Kalnay et al.,The NCEP/NCAR 40-year reanalysis project, Bull. Amer. Meteor. Soc., 77, 437-470, 1996.

Kaplan, A. et al. 1998: Analysis of global sea surface temperature 1856−1991. J. Geophys. Res. 103, 18567−18589.

Keenlyside, N., M. Latif, M. Botzet, J. Jungclaus, and U. Schulzweida, 2005: A coupled method for initializing El Niño–Southern Oscillation forecasts using sea surface temperature.Tellus, 57A, 340–356.

Keppenne C. L, M. M. Rienecker, 2003: Assimilation of Temperature into an isopycnal oceangeneral circulation model using a parallel Ensemble Kalman Filter. J. Mar. Sys., 40-41,363-380.

Kepert J.D. (2004) On ensemble representation of the observation –error convariance in the Ensemble Kalman Filter. Ocean Dynamics 6: 561–569

Kessler, W.S, 2006: The circulation of the eastern tropical Pacific: A review. Progress in Oceanography. 69,181-217.

Kirtman, B.P., 2003: The COLA Anomaly Coupled Model: Ensemble ENSO Prediction. Mon. Weather Rev. 131,2323-2341.

--- and S. E. Zebiak, 1997: ENSO simulation and prediction with a hybrid coupled model. Mon. Wea. Rev., 125, 2620–2641.

--- and Dughong Min, 2007: US National Multi-Model ENSO Prediction with CFS and CCSM3. NOAA CTB – COLA Joint Seminar, Oct. 10, 2007.

Kistler R., Coauthors, 2001: The NCEP–NCAR 50-year reanalysis: Monthly means CD-ROM and documentation. Bull. Amer. Meteor. Soc, 82, 247–268.

Latif M, Anderson D, Barnett T, Cane M, Kleeman R, Leetmaa A, O'Brien J, Rosati A, Schneider E. 1998. A review of the predictability and prediction of ENSO. Journal of Geophysical Research 103: 14 375-14 393.

Levitus, S. and T. Boyer,1998: NOAA/OAR/ESRL PSD, Boulder, Colorado, USA,http://www.cdc.noaa.gov.

Luo J.-J., S. Masson, S. Behera, S. Shingu and T. Yamagata. 2005: Seasonal climate predictability in a Coupled OAGCM Using a different approach for ensemble forecasts. J. Climate, 18:4474-4495.

Madec, G., P. Delecluse, M. Imbard, and C. Levy, 1998: OPA 8.1 ocean general circulation model reference manual. Institut Pierre-Simon Laplace des Sciences l'Environnement Global, LODYC, Université Pierre et Marie Curie, Paris, France, 97pp.

McPhaden, M.J. and Coauthors, 1998: The Tropical Ocean Global Atmosphere (TOGA) observing system: A decade of progress. J. Geophys. Res., 103 (C7), 14169-14240.

Molteni, F., R. Buizza, T. N. Palmer, and T. Petroliagis, 1996: The ECMWF ensemble prediction system: Methodology and validation. Quart. J. Roy. Meteor. Soc., 122, 73–119.

Monahan, A. H., 2001: Nonlinear principal component analysis: Tropical Indo-Pacific sea surface temperature and sea level pressure. J. Climate, 14, 219–233.

—, and A. Dai, 2004:The spatial and temporal structure of ENSO nonlinearity . J. Climate, 17 3026-3036.

Moore, A. M., and R. Kleeman, 1998: Skill assessment for ENSO using ensemble prediction. Quart. J. Roy. Meteor. Soc., 124,557–584.

—, J. Zavala-Garay, Y. Tang, R. Kleeman, J. Vialard, A. Weaver, K. Sahami, D. L. T. Anderson and M. Fisher 2006: Optimal Forcing Patterns for Coupled Models of ENSO J. Climate, 19, 4683-4699.

Nakaegawa, T., Kanamttsu, M. and Smith, T. M. 2004: Interdecadal Trend of Prediction Skill in an Ensemble AMIP-Type Experiment. J. Climate, 15: 2881-2889.

Oberhuber, J. M., E. Roeckner, M. Christoph, M. Esch, and M. Latif,1998: Predicting the '97 El Nin˜o event with a global climate model. Geophys. Res. Lett., 25, 2273–2276.

Palmer, T. and R. Hagedorn, 2006: Predictability of weather and climate. Cambridge University Press, 514-560.

Rayner, N., D. Parker, E. Horton, C. Folland, L. Alexander, D. Rowell, E. Kent, and A. Kaplan, 2003: Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century. J. Geophys. Res., 108, 4407, doi:10.1029/2002JD002670.

Rosati, A., K. Miyakoda, and R. Gudgel, 1997: The impact of ocean initial conditions on ENSO forecasting with a coupled model.Mon. Wea. Rev., 125, 754–772.

Smith, T. M., and R. W. Reynolds, 2002: Bias corrections for historical sea surface temperatures based on marine air temperatures. J. Climate,15,73-87.

——, ——, 2003: Extended reconstruction of global sea surface temperatures based on COADS data (1854–1997). J. Climate, 16, 1495–1510.

——, and ——, 2004: Improved extended reconstruction of SST (1854–1997). J. Climate, 17, 2466–2477.

Shukla, J., and Coauthors, 2000: Dynamical seasonal prediction. Bull. Amer. Meteor. Soc., 81, 2593-2606.

Snyder, C., and F. Zhang, 2003: Assimilation of simulated Doppler radar observations with anensemble Kalman filter. Mon. Wea. Rev., 131, 1663-1677.

Syu, H-H and D. Neelin, 2000a: ENSO in a hybrid coupled model. Part I: Sensitivity to physical parametrizations. Clim. Dyn, 16: 19-34.

Syu, H-H and D. Neelin, 2000b: ENSO in a hybrid coupled model. Part II: Prediction with piggyback data assimilation. Clim. Dyn., 16, 35-48.

Tang, Y, 2002.: Hybrid coupled models of the tropical Pacific – Interannual variability. Climate

Dynamics, 19, 331-342.

Tang, Y., and Hsieh, W. W, 2002: Hybrid coupled models of the tropical Pacific -- ENSO prediction . Climate Dynamics, 19, 343-353.

— and —, 2003: ENSO simulation and predictions using a hybrid coupled model with data assimilation J. of Japan Met. Soc., 81, 1-19.

— and Kleeman, R. 2002: A new strategy for SST assimilatinon for ENSO prediction, Geophysical Research Letters, 10.1029

—,— and Moore, A., 2004; SST assimilation experiments in a tropical Pacific Ocean model, J Phys. Oceangr. 34, 623-642.

—,—and—, 2005: On the reliability of ENSO dynamical predictions. J. Atmos. Bci., 62, 1770-1791.

—, Z. Deng, X. Zhou, Y. Cheng and D. Chen, 2007: Interdecadal Variation of ENSO Predictability in Multiple Models. J. Climate (submitted).

Torrence, C. and G. P. Compo, 1998: A Practical Guide to Wavelet Analysis. Bull. Amer. Meteor. Soc., 79, 61-78.

Toth, Z., and E. Kalnay, 1993: Ensemble forecasting at NMC: The generation of perturbations. Bull. Amer. Meteor. Soc., 74, 2317–2330.

—and—, 1997: Ensemble forecasting at NCEP and the breeding method. Mon. Wea. Rev., 125, 3297–3319.

von Storch, H., and F.W. Zwiers, 1999: Statistical Analysis in Climate Research. Cambridge University Press, 499 pp.

Whitaker J. S., and T. M. Hamill, 2002: Ensemble data assimilation without perturbed observations. Mon. Wea. Rev., 130, 1913-1924.

Whittenberg A. T., 2004: Extended wind stress analyses for ENSO. J. Climate, 17,1526-2540.

Xue, Y., Smith, T. M., and Reynolds, R. W. 2003: Interdecadal Changes of 30-Yr SST Normals during 1871–2000. J. Climate, 15, 1601-1612.

Zelle, H., 2004: The relation between sea surface temperature and thermocline depth in the eastern equatorial Pacific. J. Phys. Oceangr. 34, 643-655.

Zhang, F., 2005: Dynamics and structure of mesoscale error covariance of a winter cycloneestimated through short-range ensemble forecasts. Mon. Wea. Rev., 133, 2876-2893.

—, Z. Meng, and A. Aksoy, 2006: Tests of an ensemble Kalman filter for mesoscale and regional-scale data assimilation, Part I: Perfect model experiments. Mon. Wea. Rev., 134, 722-736.

Zheng, F., and J. Zhang, R. Zhang and G. Zhou, 2006: Ensemble hindcasts of SST anomalies in the tropical Pacific using an intermediate coupled model. Geophys. Res. Lett,Vol. 33, L19604.

Figure Caption

FIG. 1. (a) Time series of the simulated Nino3 SSTA index for a coupling run of 1000 year with HCM1, and (b) its local wavelet power spectrum. The contour levels are chosen so that 75%, 50%, 25%, or 5% of the wavelet power is above each level respectively. Black contour is the 10% significance level, using a red-noise (autoregressive lag(1)) background spectrum. Cross-hatched regions on either end indicate the "cone of influence", where the edge effects become important. The corresponding global wavelet power spectrum is shown as the black line in (c). The dashed line is the significance for the global wavelet spectrum, assuming the same significance level and background spectrum as in local wavelet spectrum. The right up figure is the mother function used.

FIG.2. Correlation and RMSE of SSTA from the control run (a,b) and assimilation run (c,d) relative to the observed SSTA during Jan 1881-Dec 2000. The contour interval for correlation is 0.2 and for RMSE is 0.2℃.The areas with correlation over 0.5 and RMSE over 0.6℃ are shaded.

FIG.3. Correlation (left panels) and RMSE (right panels) skills for several sea temperature anomaly indices as a function of depth, for both assimilation run (solid line) and control run (dashed line), against NCEP reanalysis dataset. The skills are computed for the period from 1981-2000. The vertical bar represents the uncertainty of the computed skill due to finite sampling, estimated by the bootstrap method.

FIG.4. The correlation (left panel) and RMSE (right panel) of HCA (defined as the depth integral of sea temperature anomalies over the top 10 levels) from the control run (top panel) and assimilation run (middle panel) respectively, both against the NCEP

FIG.5 Time-longitude plots of averaged STA over top 100m and the region of 4°S to 4°N during 1981-2000 from (a) the control run, (b) the assimilation run, and (c) NCEP reanalysis data. Contour interval is 1.0℃. The absolute values over 1.0℃ are shaded.

FIG.6. RMSD (see Eq. 6 for definition) at several depths for the period from 1881-2000. The contour interval is 0.2℃ and the values larger than 0.5℃ are shaded.

FIG.7. Averaged depth-longitude diagrams of RMSD over 120 years from 1881-2000, along (a) 4°S, (b) the equator and (c) 4°N. The dashed line is the model climatological mixing layer depth, and thick line is the thermocline depth.

FIG.8. Vertical profile of averaged vertical velocity over 120 years from 1881-2000, along (a) the 4°S, (b) equator and (c) 4°N, derived from model climatology. The unit of upweling (downweling) velocity is 10^{-6} m/s. The dashed line is the model climatological mixing layer depth, and thick line is the thermocline depth.

FIG.9. (a) Correlation and (b) RMSE skill between observed and predicted Niño3.4 SSTA indices, as a function of lead time. Shown are the persistence skill (thick gray line), prediction skill initialized from control run (dashed line) and the skill initialized from the assimilation (thick line). All skills were obtained using the predictions from 1881-2000. The vertical bar represents the error bar due to finite sampling, estimated by a bootstrap method.

FIG.10. Correlation (left panel) and RMSE (right panel) skill between predicted SSTA at 6-month lead against the observation for period from 1881-2000. The skills of predictions initialized from control run (upper panel) and from the assimilation (bottom

FIG.11. Correlation between the predicted Nino3 SSTA of 6-month lead and the STA of initial fields of several depths, for control run (a) and assimilation run (b). (c) is the correlation between observed STA at the initial time of prediction with observed Nino3.4 SSTA of 6 months later. The contour interval is 0.3 and the absolute values over 0.3 are shaded. The correlations were calculated for the period from 1881-2000 for (a) and (b), and 1981-2000 for (c).

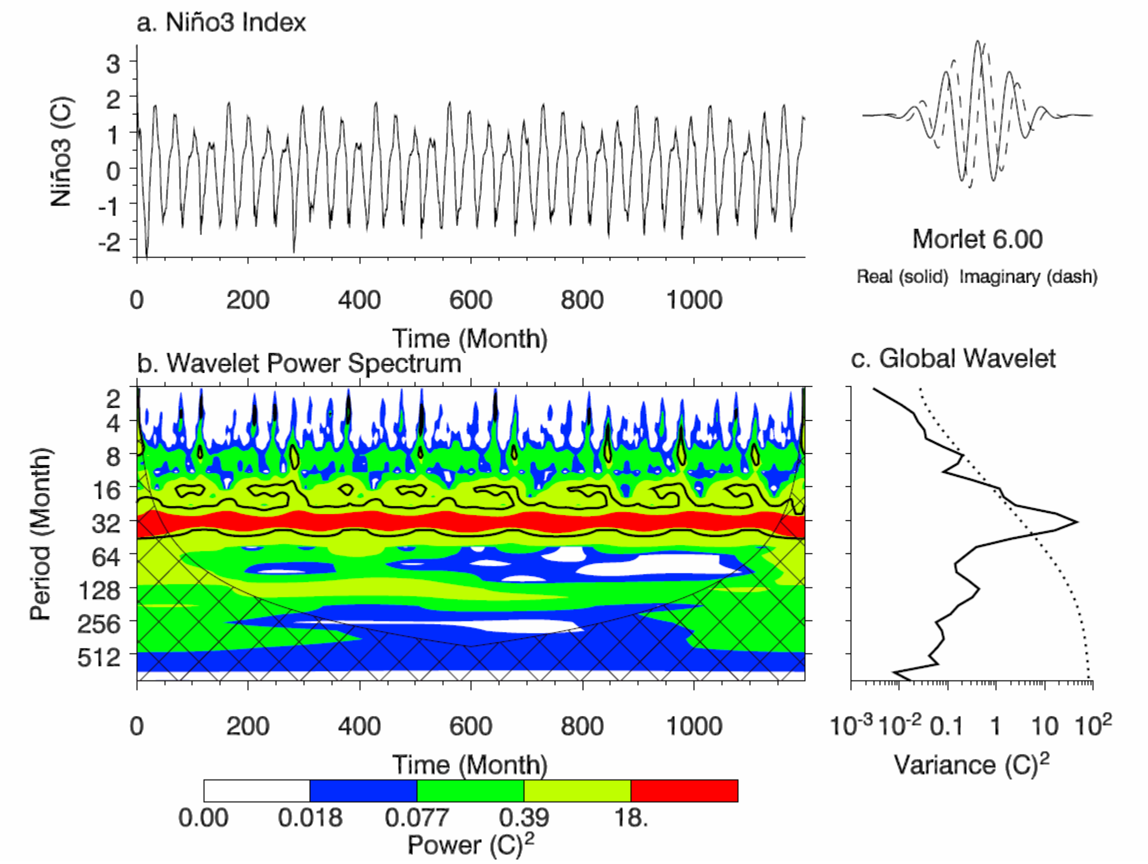

FIG. 1. (a) Time series of the simulated Nino3 SSTA index for a coupling run of 1000 year with HCM1, and (b) its local wavelet power spectrum. The contour levels are chosen so that 75%, 50%, 25%, or 5% of the wavelet power is above each level respectively. Black contour is the 10% significance level, using a red-noise (autoregressive lag(1)) background spectrum. Cross-hatched regions on either end indicate the "cone of influence", where the edge effects become important. The corresponding global wavelet power spectrum is shown as the black line in (c). The dashed line is the significance for the global wavelet spectrum, assuming the same significance level and background spectrum as in local wavelet spectrum. The right up figure is the mother function used.

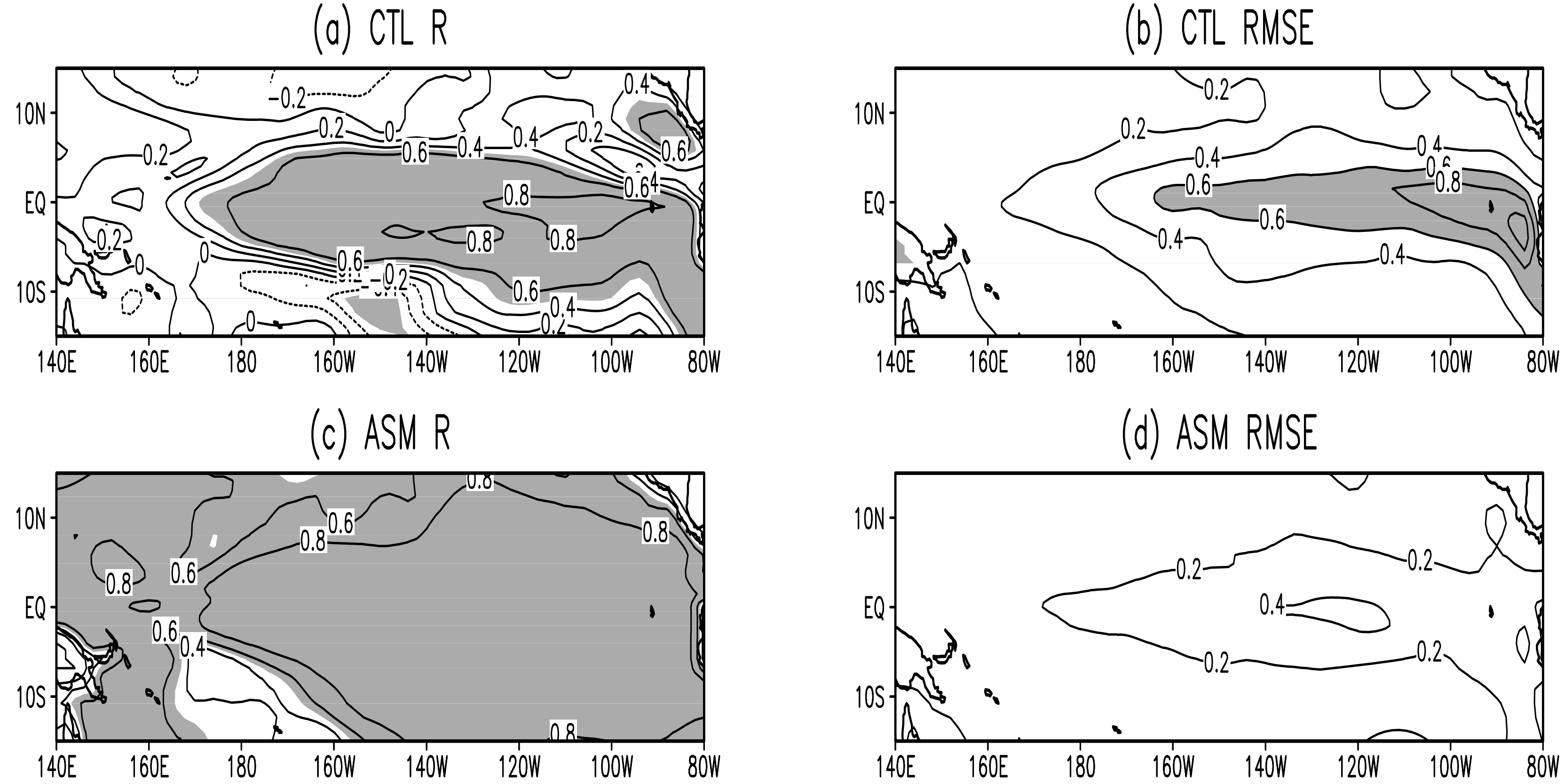

FIG.2. Correlation and RMSE of SSTA from the control run (a,b) and assimilation run (c,d) relative to the observed SSTA during Jan 1881-Dec 2000. The contour interval for correlation is 0.2 and for RMSE is 0.2℃.The areas with correlation over 0.5 and RMSE over 0.6℃ are shaded.

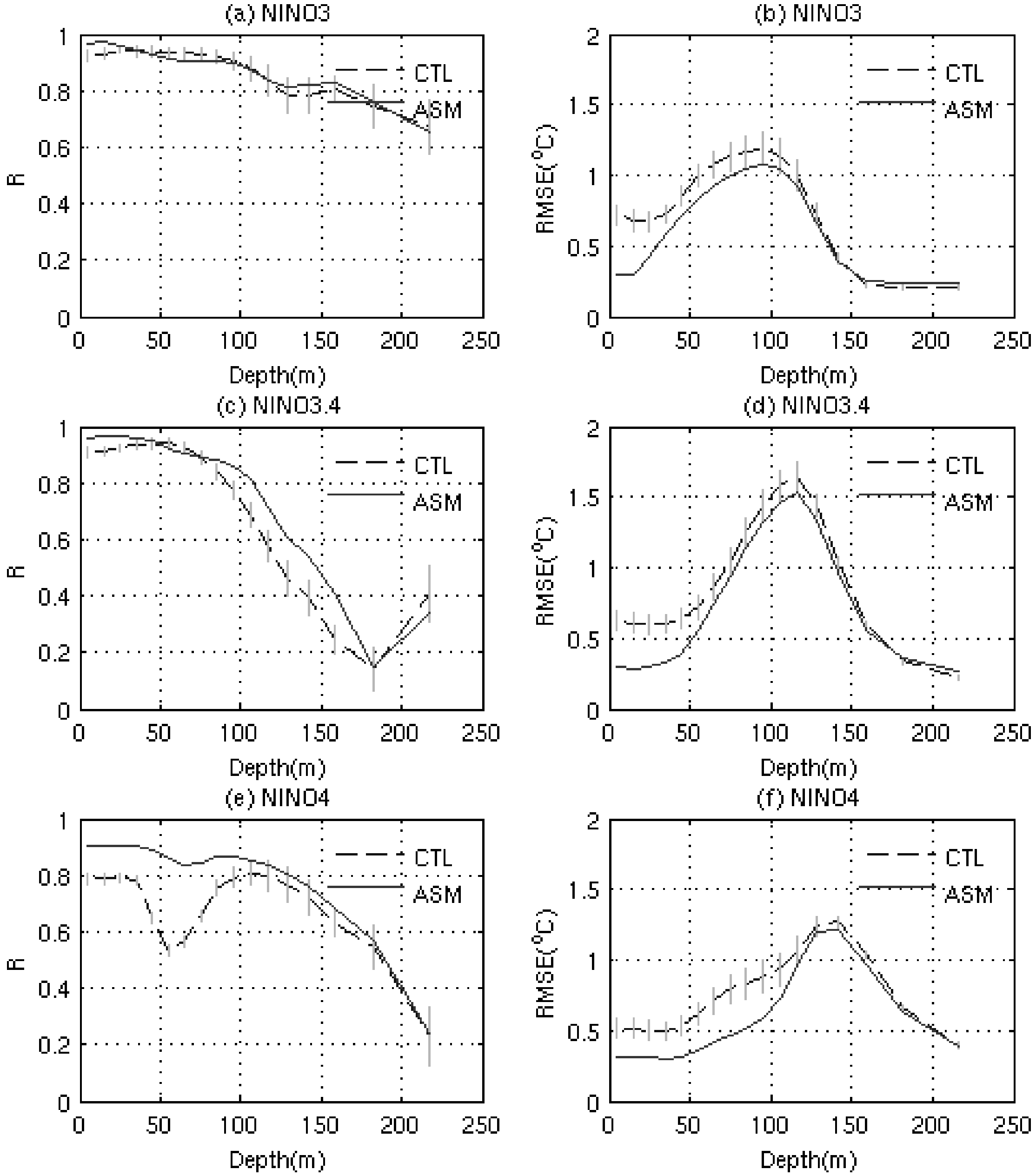

FIG.3. Correlation (left panels) and RMSE (right panels) skills for several sea temperature anomaly indices as a function of depth, for both assimilation run (solid line) and control run (dashed line), against NCEP reanalysis dataset. The skills are computed for the period from 1981-2000. The vertical bar represents the uncertainty of the computed skill due to finite sampling, estimated by the bootstrap method.

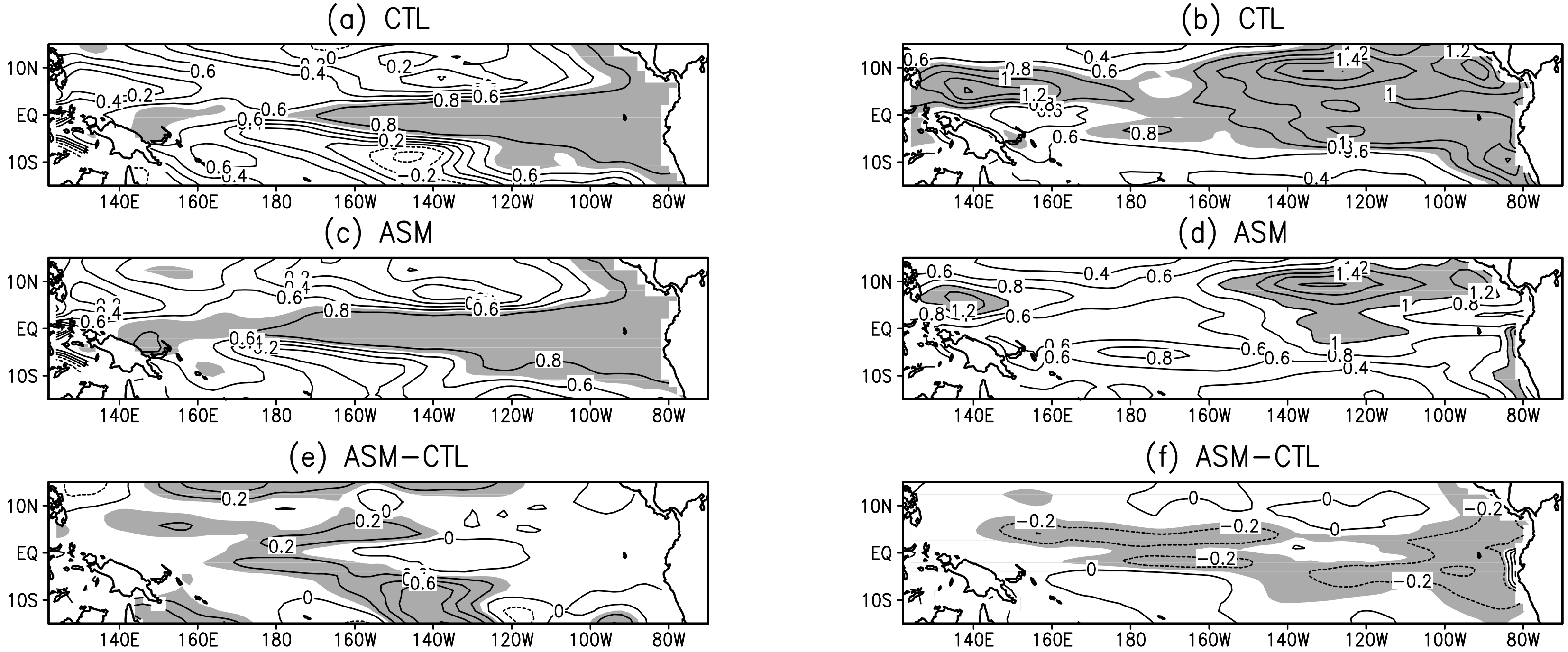

FIG.4. The correlation (left panel) and RMSE (right panel) of HCA (defined as the depth integral of sea temperature anomalies over the top 10 levels) from the control run (top panel) and assimilation run (middle panel) respectively, both against the NCEP reanalysis HCA. Their differences are shown in the bottom panel. Shaded are the values over 0.7 in (a) and (c), over 0.1 in (e), over 1.0℃ in (b) and (d), and under -0.1℃ in (f).

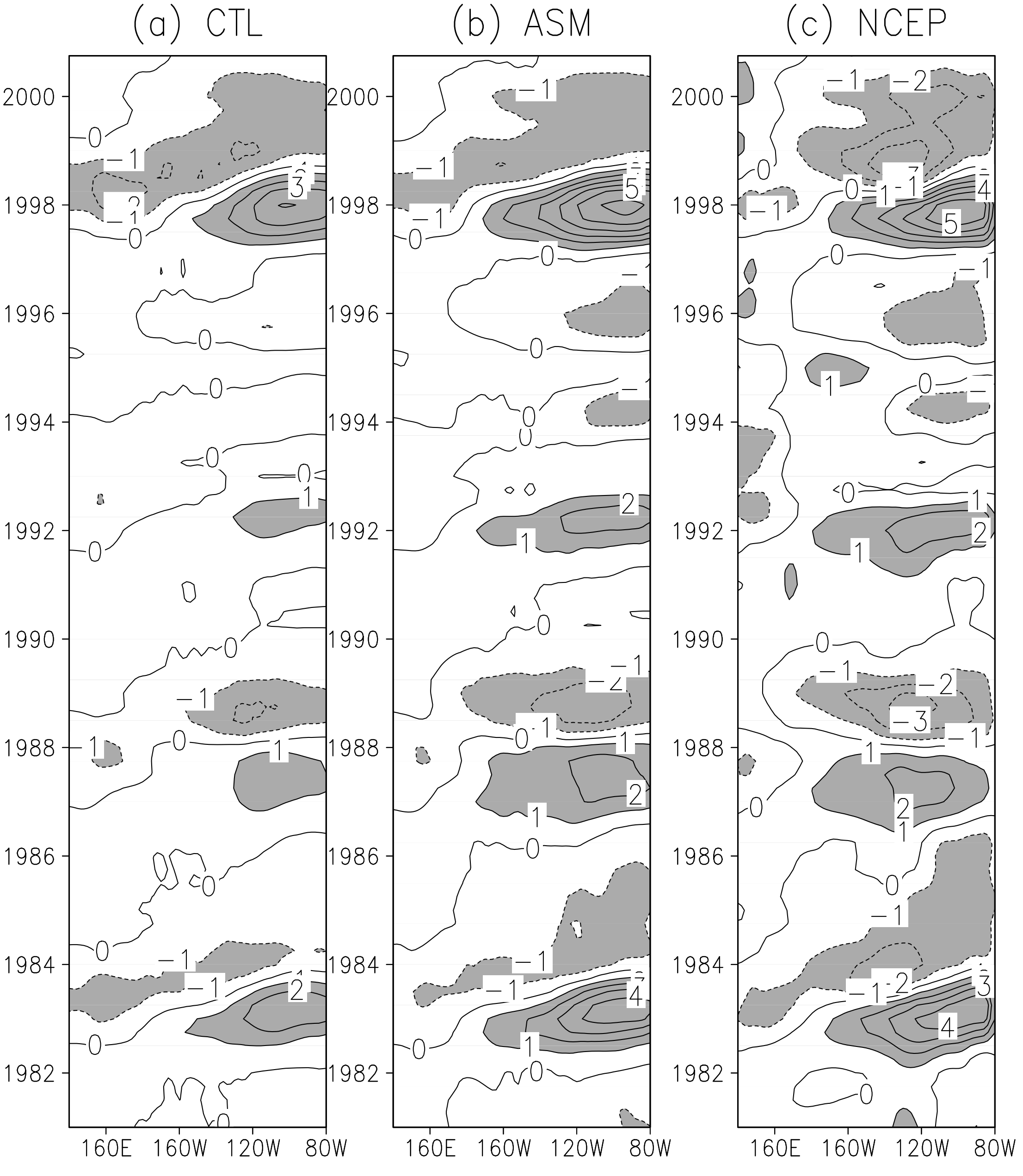

FIG.5. Time-longitude plots of averaged STA over top 100m and the region of 4°S to 4°N during 1981-2000 from (a) the control run, (b) the assimilation run, and (c) NCEP reanalysis data. Contour interval is 1.0℃. The absolute values over 1.0℃ are shaded.

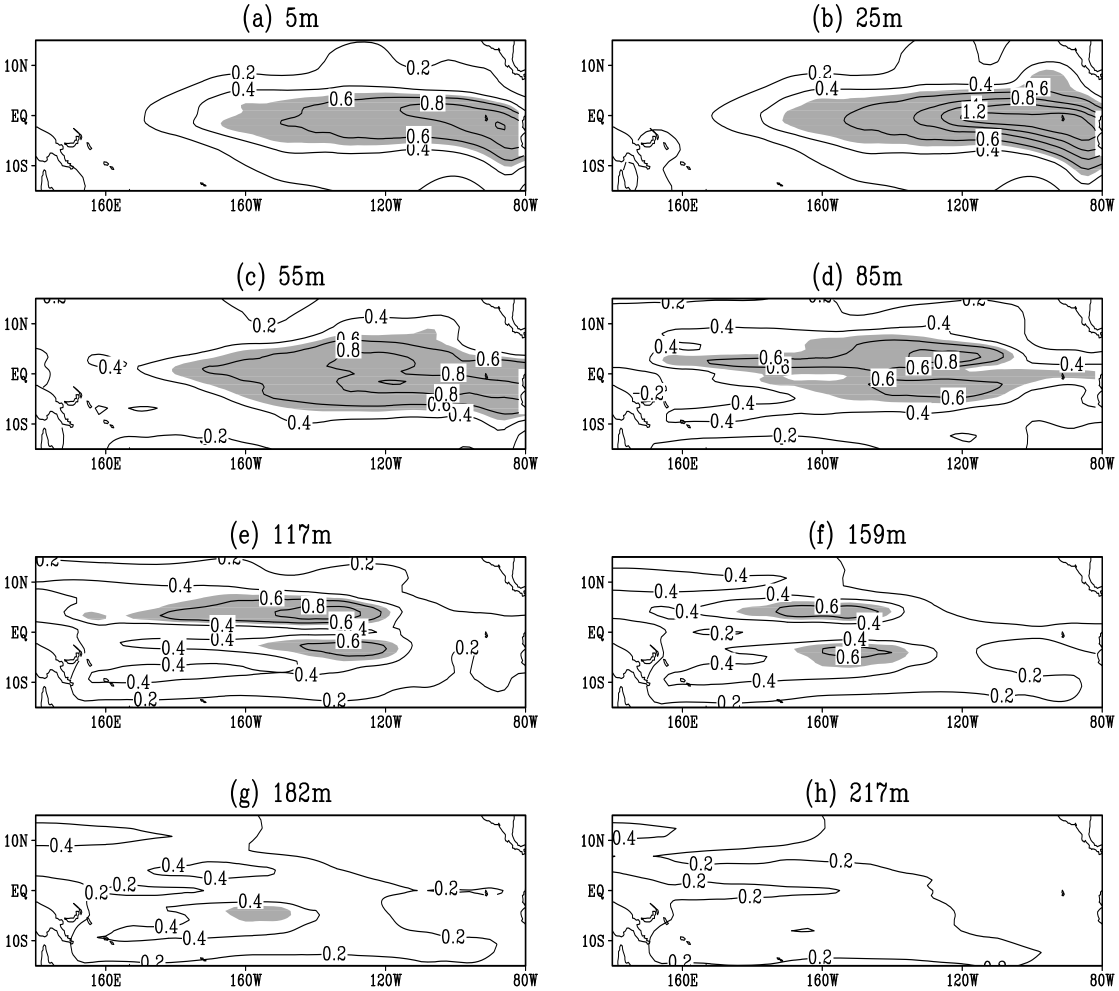

FIG.6. RMSD (see Eq. 6 for definition) at several depths for the period from 1881-2000. The contour interval is 0.2℃ and the values larger than 0.5℃ are shaded.

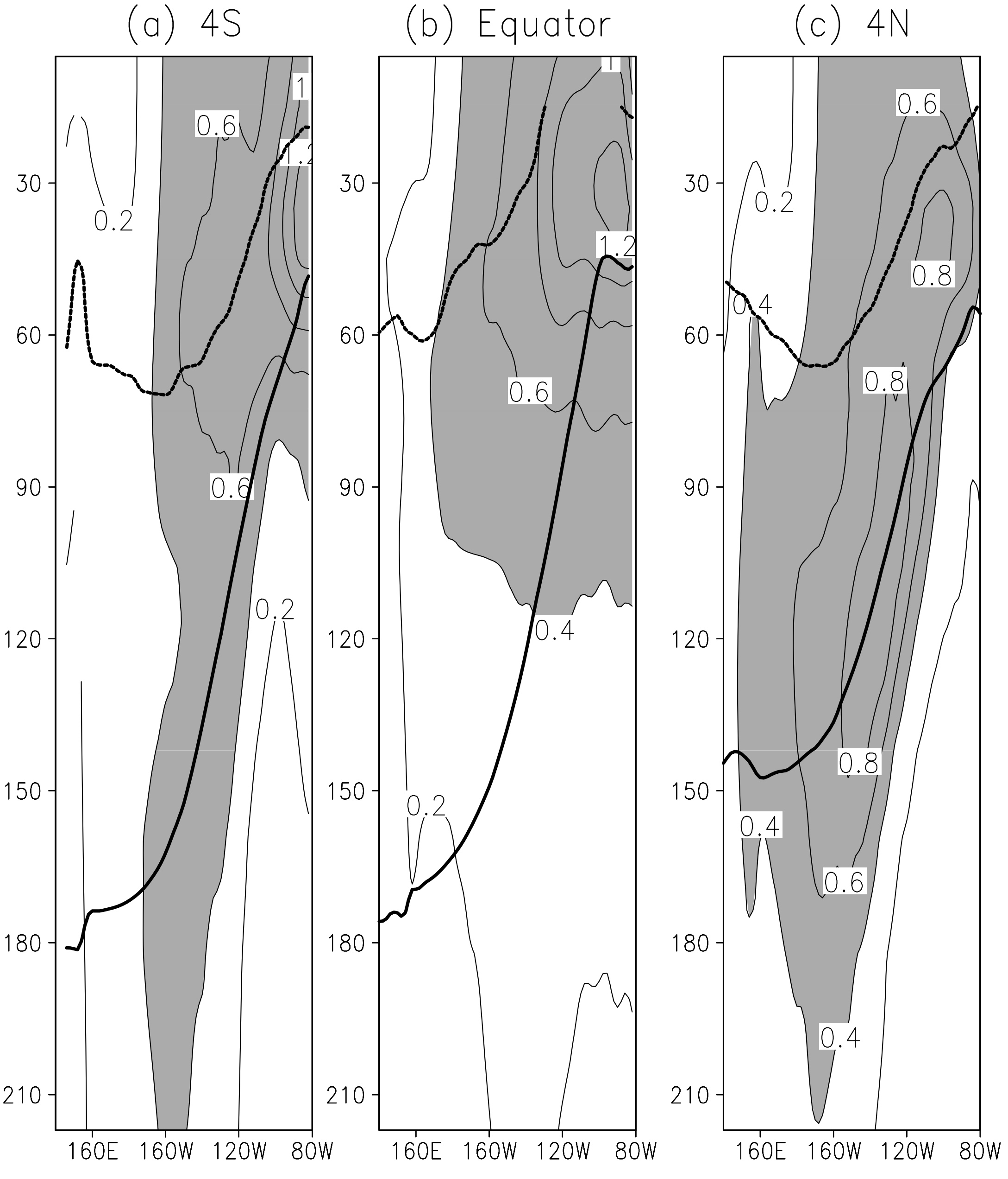

FIG.7. Averaged depth-longitude diagrams of RMSD over 120 years from 1881-2000, along (a) 4°S, (b) the equator and (c) 4°N. The dashed line is the model climatological mixing layer depth, and thick line is the thermocline depth.

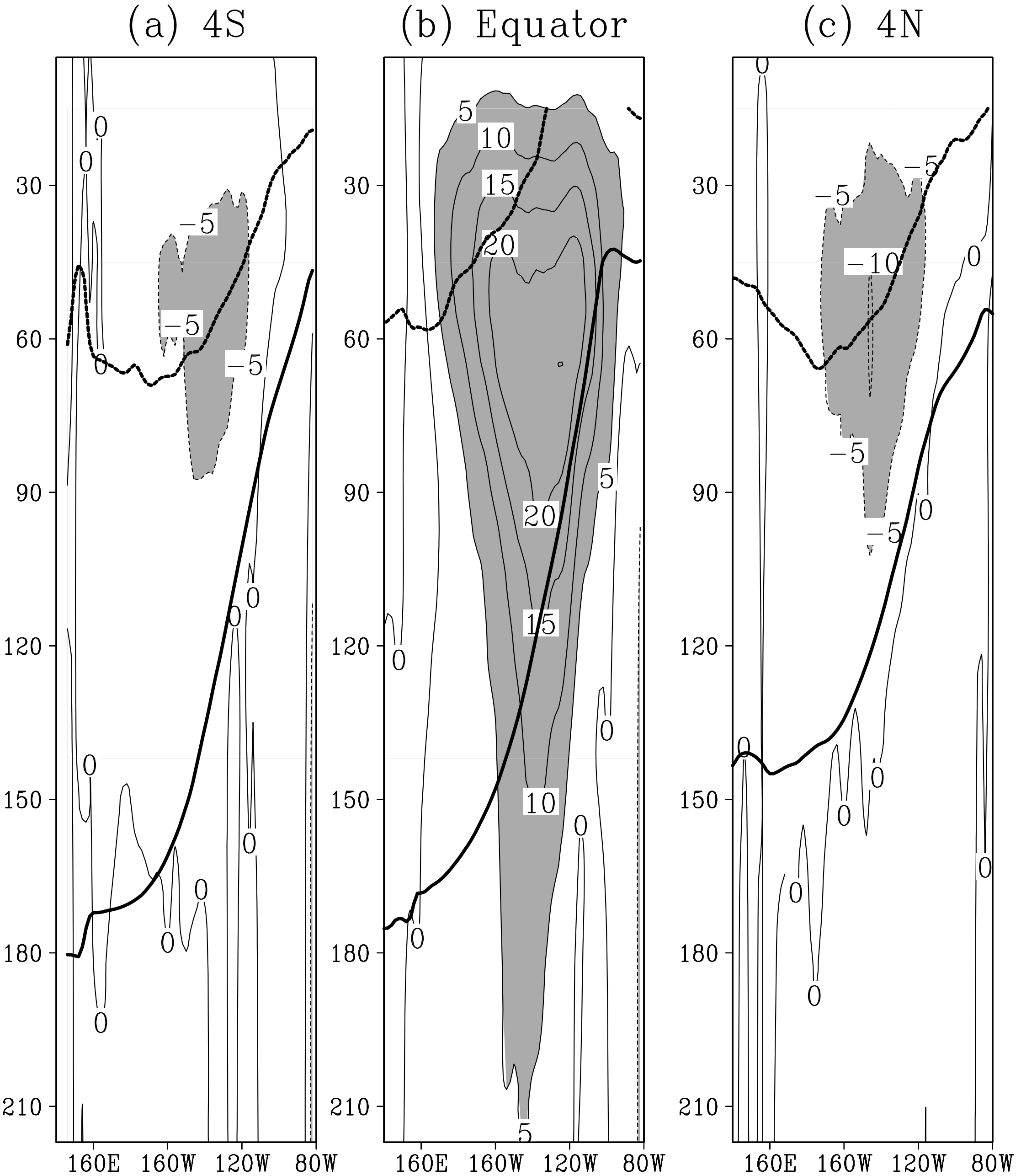

FIG.8. Vertical profile of averaged vertical velocity over 120 years from 1881-2000, along (a) the 4^{o}S, (b) equator and (c) 4^{o}N, derived from model climatology. The unit of upweling (downweling) velocity is 10^{-6} m/s. The dashed line is the model climatological mixing layer depth, and thick line is the thermocline depth.

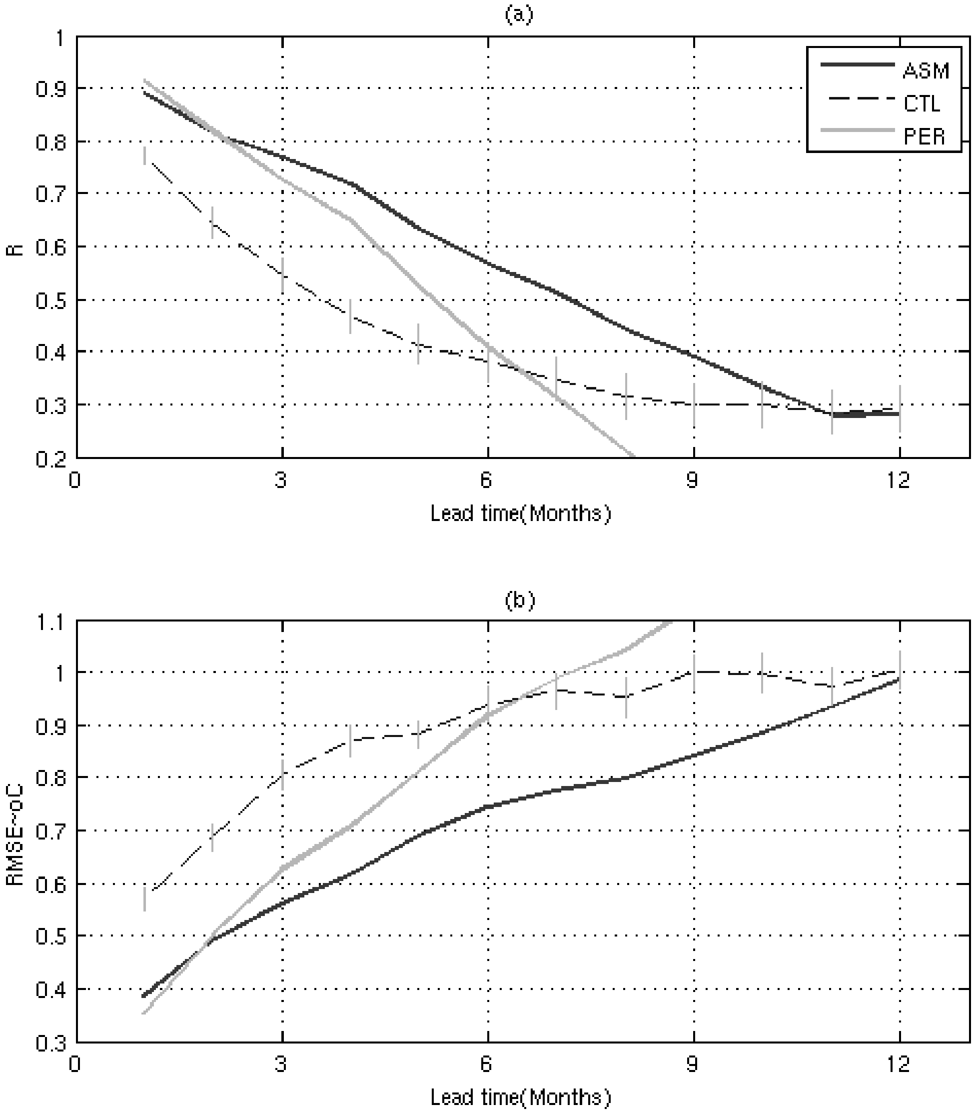

FIG.9. (a) Correlation and (b) RMSE skill between observed and predicted Niño3.4 SSTA indices, as a function of lead time. Shown are the persistence skill (thick gray line), prediction skill initialized from control run (dashed line) and the skill initialized from the assimilation (thick line). All skills were obtained using the predictions from 1881-2000. The vertical bar represents the error bar due to finite sampling, estimated by a bootstrap method.

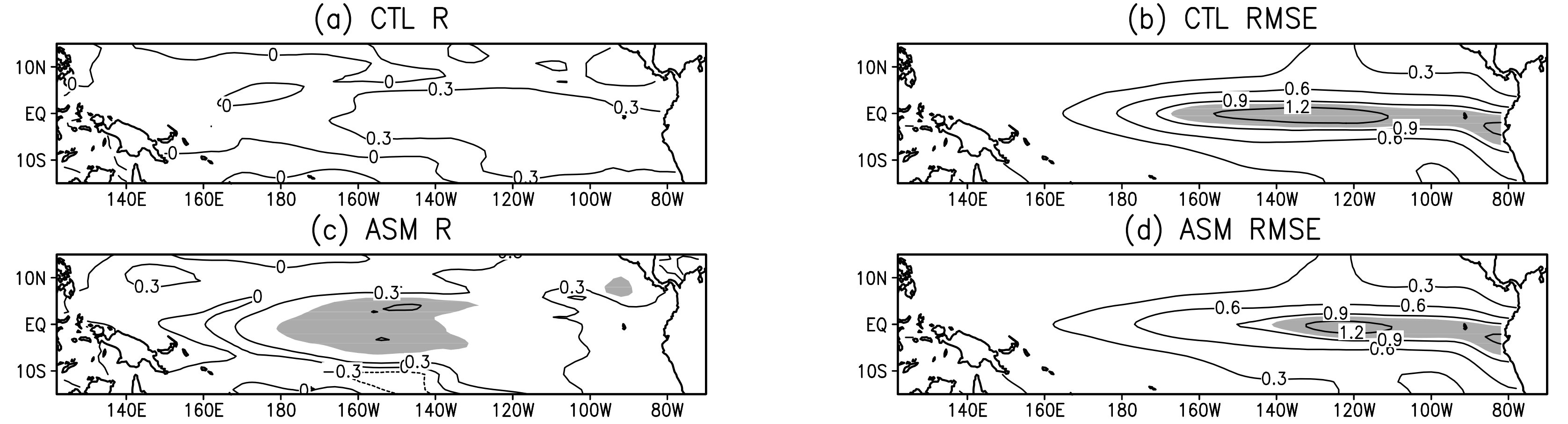

FIG.10. Correlation (left panel) and RMSE (right panel) skill between predicted SSTA at 6-month lead against the observation for period from 1881-2000. The skills of predictions initialized from control run (upper panel) and from the assimilation (bottom pane) were shown here. The contour interval is 0.3. The unit in (b) and (d) is ℃. Shaded are the values over 0.5 in (a) and (c) and over 1.0 ℃ in(b) and (d).

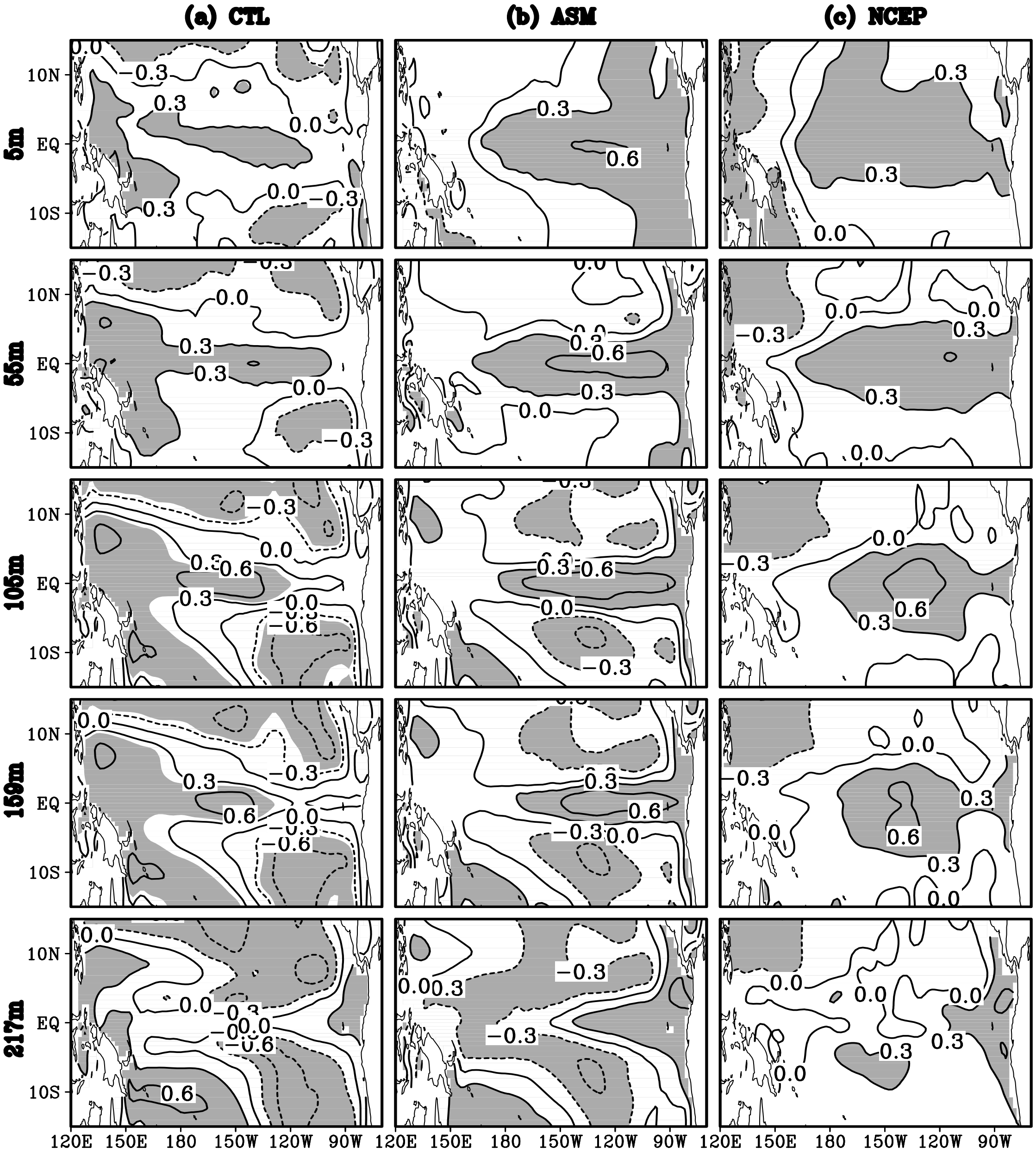

FIG.11. Correlation between the predicted Nino3 SSTA of 6-month lead and the STA of initial fields of several depths, for control run (a) and assimilation run (b). (c) is the correlation between observed STA at the initial time of prediction with observed Nino3.4 SSTA of 6 months later. The contour interval is 0.3 and the absolute values over 0.3 are shaded. The correlations were calculated for the period from 1881-2000 for (a) and (b), and 1981-2000 for (c).