Evaluation of several model error schemes in the EnKF assimilation: Applied to Argo profiles in the Pacific Ocean

Ziwang Deng,[1] Youmin Tang,[1] and Howard J. Freeland[2]

Received 5 January 2011; revised 22 June 2011; accepted 6 July 2011; published 22 September 2011.

[1] The efficacy of several model error schemes in the Ensemble Kalman Filter (EnKF) data assimilation is investigated through a series of sensitivity experiments, in which the Argo and other in situ temperature and salinity profiles are assimilated into an ocean general circulation model (OGCM) for the Pacific Ocean. Different schemes for combining the additive inflation, multiplicative inflation, one-step bias correction and two-stage bias correction are evaluated in the framework of the EnKF. Experimental results indicate that the additive inflation is the key technique that can maintain ensemble spread in an adequate range. When sufficient observations are available, the assimilation system with additive inflation scheme can efficiently reduce both model bias and random errors. The combination of additive inflation and multiplicative inflation can further improve the performance of the assimilation system, in particular when the additive inflation underestimates model error. The bias correction schemes, the one-step method and the persistent bias method are effective in reducing the model bias only within a relatively short initial assimilation period and in some regions. Further improvement from the bias correction schemes is not evident as the assimilation period increases.

Citation: Deng, Z., Y. Tang, and H. J. Freeland (2011), Evaluation of several model error schemes in the EnKF assimilation: Applied to Argo profiles in the Pacific Ocean, *J. Geophys. Res.*, *116*, C09027, doi:10.1029/2011JC006942.

1. Introduction

[2] Ocean models, including complex OGCMs, always contain, more or less, systematic biases and random errors, which often undermine ocean simulation and prediction. Developing realistic representations of these errors is a major challenge for data assimilation researchers [*Dee and da Silva*, 1998; *Zupanski and Zupanski*, 2006]. Some schemes, such as the "additive inflation" and "multiplicative inflation," have been proposed in the EnKF to deal with these errors. The "additive inflation" was used to represent model error by adding random perturbations of known statistics to the analysis or forecast ensemble [e.g., *Evensen*, 2003; *Hamill and Whitaker*, 2005; *Houtekamer and Mitchell*, 2005; *Hamill*, 2006; *Whitaker et al.*, 2008; *Bonavita et al.*, 2010] and the "multiplicative inflation" was used to inflate the forecast ensemble covariance through multiplication of a suitable temporally varying parameter [*Anderson and Anderson*, 1999; *Li et al.* 2009].

[3] It was found that a standard Kalman filter with appropriate inflation of background error covariance is able to provide an efficient analysis where measurements are available [*Drécourt et al.*, 2006]. Further, the use of the incremental analysis update (IAU) can enhance the impact of inflation [*Dee and Todling*, 2000]. Increasing the covariance matrix spread by additive inflation and multiplicative inflation reduces the bias but increases the random component of the analysis error [e.g., *Dee and Todling*, 2000]. However, in these bias-blind data assimilations (the bias-blind assimilation assumes the model to be perfect, ignoring systematic bias and treating it as a random error), that are only designed to correct random model errors based on unbiased observations, the model biases are ignored or treated as random model error. They are actually implicitly corrected because the innovation (observation-minus-forecast) includes both bias and random error. The bias correction will be achieved in a suboptimal way because it is highly unlikely that the covariance structure of forecast errors will match the spatial structure of the bias [*Thacker and Esenkov*, 2002; *Thompson et al.*, 2006] due to different characteristics between the model bias and random model errors.

[4] Some specific techniques that deal with the model bias have been developed and incorporated into standard data assimilation methods [*Thiébaux and Morone*, 1990; *Dee*, 1995, 2005; *Dee and da Silva*, 1998; *Martin et al.*, 2002; *Kalnay*, 2002; *Nichols*, 2003; *Chepurin et al.*, 2005; *Drécourt et al.*, 2006; *Deng et al.*, 2010]. An assimilation system with a bias correction scheme is often called a bias-aware assimilation system. A scheme referred to as "two-stage estimation" and its simplified version "one-step" method are widely used in bias-aware data assimilation [*Dee*, 1995; *Dee and da Silva*, 1998; *Kalnay*, 2002; *Nichols*, 2003; *Deng et al.*, 2010]. In two-stage estimation we assume that a reasonable estimate of the bias may be made prior to estimating the state of the system itself, thus allowing the estimation procedures for the

[1] Environmental Science and Engineering, University of Northern British Columbia, Prince George, British Columbia, Canada.
[2] Institute of Ocean Sciences, Fisheries and Oceans Canada, Sidney, British Columbia, Canada.

0148-0227/11/2011JC006942

bias and the state to be carried out successively. Hence it requires two analysis steps: one for the bias estimation and a second for the state vector estimation [*Chepurin et al.*, 2005]. On the other hand, assuming that the bias is nearly constant in time, and the bias error covariance matrix is proportional to the forecast error covariance matrix, with the proportionality constant smaller than one, the "two-stage estimation" can be simplified by the "one-step" method.

[5] An important issue in a bias correction data assimilation system is the forecasting of the bias. A number of bias forecast models were developed, including the persistence model, which forecast the bias using the bias analysis value from the previous step [*Dee*, 2005]; the spatial average of analysis error (analysis-minus-observation) of the previous step [*Chu et al.*, 2004; *Frank and Colby*, 1997; *Deng et al.*, 2010]; the constant bias model, which assumes the bias is constant or very-slowly changing compared to the model dynamics [*Drécourt et al.*, 2006]; and the low-dimensional bias forecast model (LDM) [*Danforth et al.*, 2007].

[6] It has been of great interest and practical significance to compare such data assimilation techniques in terms of the reduction of total analysis error as the bias forecast models, the inflation methods, localization, etc. For example, *Li et al.* [2009] compared the performance of several model error schemes in an EnKF system that assimilates NCEP-NCAR reanalysis fields into a low-resolution AGCM. Their results show that the "one-step" scheme is unable to beat the multiplicative or additive inflation schemes. *Chepurin et al.* [2005] tested several relatively simple empirical bias forecast models, which are the combination of constant (in time) bias, periodic bias and ENSO-related bias in the tropical Pacific and found that the constant or periodic bias forecast may overestimate the bias as the assimilation steps increase, because it ignores the fact that the bias should decrease as more and more observations are assimilated.

[7] In this study, we will compare the impacts of different model error schemes on ocean data assimilation, especially on the assimilation of Argo data in the framework of the EnKF that was weakly addressed before. The use of a realistic OGCM and Argo data for assimilation experiments allows this study not only to shed light on data assimilation methodology itself but also to have practical significance for Argo data assimilation, which has been an intensive research field in recent years [e.g., *Huang et al.*, 2008; *Smith and Haines*, 2009; *Deng et al.*, 2010].

[8] This paper is organized as follows. The data assimilation system is described in section 2. The experimental design is presented in section 3. The metrics for quantifying the performance of data assimilation are described in section 4. In section 5 we evaluate the bias in the OGCM. The validations and comparisons among the experiments are shown in section 6 and 7. Finally, section 8 summarizes results and explores implications for future work.

2. The Ocean Data Assimilation System

2.1. Data

[9] The data assimilation system used in this study is similar to that used by *Deng et al.* [2010] except for some specific settings in some of the experiments. Here we briefly review the common characteristics of the system and leave the specific settings to section 4 where we will describe the design of the experiments.

[10] The data sets assimilated include the delayed-mode Argo profiles (Argo) [*Carval et al.*, 2006], Expendable Bathythermographs (XBT), Conductivity-Temperature-Depth (CTD) [*Bellucci et al.*, 2007] and Tropical Atmosphere and Ocean-Triangle Trans-Ocean buoy Network (TAO-TRITON) [*McPhaden*, 1995] for the period from 2005 to 2007. All these data sets are available online at (http://www.nodc.noaa.gov/GTSPP/). Some independent data including the Ocean Surface Current Analysis-Real time [*Bonjean and Lagerloef*, 2002; *Helber et al.*, 2007], withheld Argo profiles that are not assimilated and the satellite sea level anomaly [*Dibarboure et al.*, 2009] which is a time delayed low-resolution (1° × 1°, Mercator grid) gridded product (http://www.aviso.oceanobs.com), are used for validation of experiments.

2.2. The Ocean General Circulation Model

[11] The ocean general circulation model (OGCM) is the ocean component of the Nucleus for European Modeling of the Ocean version 2.3 (NEMO) [*Madec*, 2008]. The domain of the model used here is the Pacific Ocean between 60°N and 60°S and between 116°E and 66°W, for a total of 90 × 95 horizontal grid points. Closed lateral boundary conditions, a free surface and a z-coordinate with a partial step (to match the bathymetry) are used in this study. 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 and then changing down to 1.0° at 60°N and 60°S. There are 31 unevenly spaced levels along the vertical with 24 levels concentrated in the upper 2000 m. The thickness of the levels varies from 10 m at the surface (within the first 105 m) to 500 m below the 3000 m level. The maximum depth is set to 5000 m and a realistic topography based on the ETOPO5 global atlas is used [*Ferry et al.*, 2007].

[12] The model is forced for 200 years with the NCEP monthly climatological mean wind stress, derived from the 50-year NCEP reanalysis wind stress and the heat flux Q_s to get an initial state. From this initial condition, the model is forced by the actual monthly NCEP wind stress to simulate conditions for the period 1981–2004. The ocean state at the end of 2004 provided the initial condition for the control run starting on January 1 2005. The control run for the period 2005–2007, forced by NCEP reanalysis wind stress without data assimilation, provides a basis for comparison. The initial conditions for the ensemble members are provided by a 3-month spin-up run starting from the end of September 2004. Considering uncertainties in the atmospheric forcing (wind), the spin-up run is forced by perturbed NCEP wind stress, with a perturbation defined by

$$\delta\tau_{x,y} = \gamma\{\tau'_{x,y}[t_{k1}] - \tau'_{x,y}[t_{k2}]\} \quad (1)$$

where $\tau'_{x,y}[t_{k1}]$ and $\tau'_{x,y}[t_{k2}]$ are the NCEP reanalysis wind stress anomalies at two randomly chosen months t_{k1} and t_{k2} during the period from 1980 to 2004. The parameter $\gamma = 0.1$ controls the amplitude of the perturbation. This method is similar to the scheme proposed by *Keppenne et al.* [2008].

The perturbation of wind is also applied to the sequential assimilation run. With such perturbations, there is an ensemble of spin-up, providing initial conditions for all of the assimilation runs. In this study, the ensemble size is set to 31.

[13] The heat flux Q_s is given by

$$Q_s = Q_0 + \lambda(T - T_0) \tag{2}$$

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 and Boyer*, 1998], and λ is the relaxation rate, set to $-40\ \mathrm{Wm^{-2}K^{-1}}$ [*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*, 2008].

2.3. The Data Assimilation System

[14] The data assimilation system is based on the localized EnKF, which addresses computational efficiency and the spurious correlation between distant locations in the background covariance matrix [e.g., *Gaspari and Cohn*, 1999; *Houtekamer and Mitchell*, 2001; *Oke et al.*, 2005, 2008; *Hunt et al.*, 2007; *Deng et al.*, 2010]. It assimilates all observations that may affect the analysis at a given grid point simultaneously and obtains the analysis independently for each model grid point as with *Ott et al.* [2004] and *Szunyogh et al.* [2008]. In detail, we perform local analyses grid by grid for the top 24 model levels. First, a box, centered at the grid point that is analyzed (called the analysis grid) is defined and then all observations within this box are used to generate the analysis. In this study, we set this box to 6000 km × cos(θ) in longitude, 3000 km × cos(θ) in latitude and 200 m in the top 250 m and 5 levels (>200 m) for deeper levels in the vertical, where θ is latitude. After all grids over the model domain are analyzed, a final analysis ensemble can be obtained. The box used in localization smoothly shifts with the analysis grid. With the local analysis strategy, this size is probably sufficient in estimating the error covariance matrix. Our sensitivity experiments show that the use of ensemble with size larger than 31 resulted in marginal improvements in the estimation of state.

[15] The forecast ensemble member $X^f_{k,i}$ is updated by

$$X^f_{k,i} = M(X^a_{k-1,i}) + \lambda_1 q_{k,i} \tag{3}$$

where M is the nonlinear model, $X^a_{k-1,i}$ is the analysis of ith ensemble member at step k-1, $q_{k,i}$ is a pseudo random spatial coherent model error obtained using the scheme of *Evensen* [2003] and λ_1 is a parameter that controls the amplitude of the random model error. Using the forecast ensemble, the model background error covariance matrix can be approximated by

$$P^p_k = \overline{\left(X^f_k - \overline{X^f_k}\right)\left(X^f_k - \overline{X^f_k}\right)^T} \tag{4}$$

where the overline denotes an average over the ensemble. A covariance localization function is used to reduce spurious correlations between distant locations due to the limited ensemble size [e.g., *Gaspari and Cohn*, 1999; *Houtekamer and Mitchell*, 2001; *Oke et al.*, 2005; *Deng et al.*, 2010]. In addition, an adjustable inflator λ_2 is used to improve estimates of the ensemble error covariance matrix, i.e.,

$$P^f_k \approx (1 + \lambda_2)\left(\rho \circ P^p_k\right) \tag{5}$$

[16] As with *Mitchell and Houtekamer* [2000], the inflator λ_2 is estimated by

$$\lambda_2 = \frac{tr(\langle \nu\nu^T \rangle) - tr\left(H_k P^p_k H^T_k\right) - tr(R_k)}{tr\left(H_k P^p_k H^T_k\right)} \tag{6}$$

The inflator estimated by (6) is determined by the three items: the traces of the observation-minus-forecast covariance $\langle \nu\nu^T \rangle$, the covariance matrix of forecast ensemble in observation space ($\mathrm{H_k P^p_k H^T_k}$) and the observation error covariance matrix ($\mathrm{R_k}$). Here H_k is the observation operator. It should be noted that the parameter λ_2 is set to zero if its actual value in (6) is less than 0.

[17] The standard Kalman gain can be written as below [*Evensen*, 2003; *Deng et al.*, 2010]

$$K_k = P^f_k H^T_k \left(H_k P^f_k H^T_k + R_k\right)^{-1} \tag{7}$$

The bias-aware analysis for the state variables is obtained using the standard analysis procedure [*Balmaseda et al.*, 2005]

$$X^a_{k,i} = \left(X^f_{k,i} - \lambda_3 \beta^a_k\right) + K_k\left[y_{k,i} - H\left(X^f_{k,i} - \lambda_3 \beta^a_k\right)\right] \tag{8}$$

where $(X^f_{k,i} - \lambda_3 \beta^a_k)$ is the bias-corrected model state vector ($\lambda_3 = 1$) or the original model state vector without bias correction ($\lambda_3 = 0$), β^a_k is the analysis of model bias (see below), $y_{k,i}$ is the perturbed observation. In this study the state vector $X = [T\ S]^T$ includes temperature (T) and salinity (S).

[18] The definition of the observation error covariance matrix ($\mathrm{R_k}$) is the same as that given by *Deng et al.* [2010]. The localization function (ρ) is a Gaussian function defined by

$$\rho(x, y, z) = e^{-\left\{\left(\frac{dx}{Lx}\right)^2 + \left(\frac{dy}{Ly}\right)^2 + \left(\frac{dz}{Lz}\right)^2\right\}} \tag{9}$$

where dx, dy and dz are distances between the analysis grid and its surrounding grids in the zonal, meridional and vertical, respectively. The localization length scales are set to Lx = 3000 km × cos(θ), Ly = 1500 km × cos(θ) and Lz = 100 m in the three directions, just a half of the size of the box used in localization.

[19] The incremental analysis update (IAU) proposed by *Bloom et al.* [1996] is also used to avoid model dynamical instability. The principle of this method is to incorporate the analysis increment directly into prognostic equations of the model as an additional forcing term [*Balmaseda et al.*, 2007; *Castruccio et al.*, 2008]. Like *Keppenne et al.* [2005], the analysis increments are added to model states gradually over the time steps from k to $k + 1$. In this study, there are a total

Table 1. The Parameters for the Six Experiments

Experiment	λ_1 (Additive)	λ_2 (Multiplicative)	λ_3 (Bias)	Bias Correction
Exp1	0	0	0	No
Exp2	0	Adaptive	0	No
Exp3	1	0	0	No
Exp4	1	Adaptive	0	No
Exp5	1	Adaptive	1	Dee's
Exp6	1	Adaptive	1	Persistent

number of 120 integration steps from k to $k + 1$, thus each step only accounts for 1/120 of increments.

3. Experimental Design

[20] Six sensitivity experiments are designed to explore the impact of different schemes on the performance of the data assimilation system, as summarized in Table 1. In the first four experiments, model bias is not considered whereas in the last two experiments model bias is corrected using bias correction methods. A detailed introduction for each of the experiments is given as below.

3.1. Perfect Model (Exp1)

[21] In this experiment, we assume the model is perfect, i.e., the data assimilation system only considers errors in initial conditions without considering random model error ($\lambda_1 = 0$) and bias. It is an idealized case, as a reference for the comparisons against other experiments. In this case, there is no additive random error added to the forecast ensemble and no inflation used to tune the background error covariance matrix. As can be seen in the following sections, the spread of the analysis and forecast ensembles will not have to decrease with assimilation steps since the random error in observation and the uncertainty in wind stress are considered for ensembles during the analysis cycle.

3.2. Adaptive Multiplicative Inflation (Exp2)

[22] The inflation method has been used in many studies to avoid underestimation of model background error covariance ($\lambda_2 \geq 0$). This experiment is the same as Exp1 but an adaptive inflator estimated using equation (6) is used to inflate the model background error covariance matrix according to equation (5). This experiment can be used to explore the role of the multiplicative inflation method in data assimilation.

3.3. Additive Inflation (Exp3)

[23] This experiment is the same as Exp1 but a pseudo-random perturbation is added to each forecast ensemble member prior to the calculation of model background error covariance matrix ($\lambda_1 = 1$). Comparison between this experiment and Exp1 can help us assess the impact of the additive inflation. The construction of the 3-dimensional pseudorandom perturbation ε_j (j = 1,2,..,L) is based on 2-dimensional pseudo-random field series generated with the method proposed by *Evensen* [2003], namely,

$$\begin{aligned} \varepsilon_1 &= W_1 \\ \varepsilon_j &= \alpha_j \varepsilon_{j-1} + \sqrt{1-\alpha_j^2} W_j \end{aligned} \tag{10}$$

where W_j (j = 1,2…,L) is a two-dimensional pseudorandom field at the *j-th* model level, $a_j \in [0\ 1]$ is a parameter to control the coherent structure between level *j-1* and level *j*. Here, we approximate it using the correlation of model variable anomalies (temperature and salinity) between the two levels, which can be derived from the control run during the period of 1985–2004.

[24] Different from *Deng et al.* [2010], the amplitude of the model error (q) is always rescaled to a random number $\kappa \in N(0, 0.025\sigma_j^2)$, where σ_j^2 is the mean square of temperature or salinity anomalies at level j averaged over model domain obtained using the control run for the period 1985–2004. Comparisons between Exp3 and Exp2 can also help us to understand the relative roles of the additive inflation and the multiplicative inflation in reducing analysis error.

3.4. Combination of the Additive Inflation and Multiplicative Inflation (Exp4)

[25] This experiment is a mix of Exp2 and Exp3, i.e., using both the additive inflation and multiplicative inflation simultaneously ($\lambda_1 = 1, \lambda_2 \geq 0$). It can help us explore the joint contribution of the additive inflation and multiplicative inflation on the data assimilation performance.

3.5. One-Step Bias Correction Scheme (Exp5)

[26] In the previous four experiments, there is no bias correction method incorporated into the data assimilation system. This experiment is the same as Exp4 except that it includes a simplified bias correction scheme proposed by *Dee* [2005], in which the model bias is assumed to be the analysis bias of the previous step (i.e. $\beta_k^f = \beta_{k-1}^a$). The bias is subtracted from each forecast ensemble member before the analysis of the model state. At the first data assimilation step, the forecast of bias is set $\beta_1^f = 0$. Following *Dee* [2005], the state analysis is given by

$$X_{k,i}^a = \left(X_{k,i}^f - \beta_{k-1}^a\right) + K_k\left[y_{k,i} - H\left(X_{k,i}^f - \beta_{k-1}^a\right)\right] \tag{11}$$

The analysis of the model bias, which will be used as the bias forecast for the next step, is calculated using the following formula [*Dee*, 2005]:

$$\beta_k^a = \beta_{k-1}^a - \gamma_2 K_k\left[\overline{y_k} - H_k\left(\overline{X_k^f} - \beta_{k-1}^a\right)\right] \tag{12}$$

where $\gamma_2 (\ll 1)$, a tuneable coefficient, is set as a constant of 0.01 in this study; $\overline{y_k}$ is the original observation without perturbation.

3.6. Persistent Bias Correction Scheme (Exp6)

[27] For this method we assume that the model bias is a large scale error and can be approximated by the mean of the differences between simulation and observation averaged over a large spatial area [*Chu et al.*, 2004; *Frank and Colby*, 1997; *Deng et al.*, 2010], i.e.,

$$\beta_k^f = \left[\frac{1}{N_{k-1}} \sum_{|dx|<l_x, |dy|<l_y} \left(H_{k-1}\overline{X_{k-1}^a} - \overline{y_{k-1}}\right)\right] \tag{13}$$

where $|dx|$ and $|dy|$ are the distances between the model grid point and the location of the observation in the zonal and meridional directions respectively; $|dx| < l_x$, $|dy| < l_y$ defines a rectangle region of l_x and l_y, centered at the model grid point. We set $l_x = 3000$ km $\times \cos(\theta)$ and $l_y = 1500$ km $\times \cos(\theta)$ here, which are the same as those used in localization. N_{k-1} is the number of the observations within the rectangular region, H_{k-1} is the measurement operator, $\overline{X^a_{k-1}}$ is the mean of the analyzed state ensemble X^a_{k-1}, $\overline{y_{k-1}}$ is the original observation vector and also the mean of the observation ensemble. Similar to Exp5, the Kalman gain used for the bias analysis is assumed to be a small part of the Kalman gain used for state estimation and the bias forecast at the first step is set to 0. Following the standard "two-stage" procedure, the analysis is as follows [*Dee*, 2005]:

$$\beta^a_k = \beta^f_k - \gamma_2 K_k \left[\overline{y_k} - H_k\left(\overline{X^f_k} - \beta^f_k\right)\right] \tag{14}$$

$$X^a_{k,i} = \left(X^f_{k,i} - \beta^a_k\right) + K_k\left[\overline{y_k} - H\left(X^f_{k,i} - \beta^a_k\right)\right] \tag{15}$$

[28] The above experiments are summarized in Table 1. These experiments are all conducted for the 3-year period from January 1 2005 to December 31 2007. For validation, only 90% randomly chosen Argo profiles and the other in situ observations are assimilated and the remaining 10%, not assimilated, will be used for validation. A control run (CTL) without data assimilation is also performed as the reference for comparison.

4. Metrics Used for Evaluation

[29] With the assumption that observations are unbiased, the model bias can be detected using time- and space-averaged differences between observation and model background. Nonzero mean difference (MD) indicates the presence of bias. The root-mean square error (RMSE) represents the total error in forecast and/or analysis [*Dee*, 2005]. Following this convention, we use the MD and RMSE to represent bias and total error, respectively, as defined below

$$MD = \frac{1}{NM}\sum_{k=1}^{M}\sum_{i=1}^{N}\left(Y^m_{i,k} - Y^o_{i,k}\right) \tag{16}$$

$$RMSE = \sqrt{\frac{1}{NM}\sum_{k=1}^{M}\sum_{i=1}^{N}\left(Y^m_{i,k} - Y^o_{i,k}\right)^2} \tag{17}$$

where Y is the variable of evaluation; the superscript "o" denotes the observations and "m" denotes the model; the subscript i is the index of observation and N is the total number of the observations within a region or grid cell of evaluation; k is the assimilation step and M is the number of total steps.

[30] The forecast and analysis ensemble spread s^f_k and s^a_k are also used in following discussions:

$$s^f_k = \sqrt{\frac{1}{MN_k - 1}\sum_{i=1}^{M}\sum_{j=1}^{N_k} Y^f_{i,j}} \tag{18}$$

$$s^a_k = \sqrt{\frac{1}{MN_k - 1}\sum_{i=1}^{M}\sum_{j=1}^{N_k} Y^a_{i,j}} \tag{19}$$

where M is the ensemble size, N_k is the number of the model grid in the region of evaluation.

[31] We performed several different calibration experiments using different observations, including all Argo profiles, the 10% withheld Argo profiles, satellite surface currents and satellite mean sea level height anomalies. When the Argo profiles are used for calibration, the modeled variables are linearly interpolated to the locations of the Argo observations. To ensure that sufficient Argo profiles are available, all Argo profiles within a grid cell of 5° (lat.) × 5° (lon.) centered at the grid of evaluation are used. For consistency and simplicity, the comparisons between the high spatial resolution satellite data and the model data are also performed on the 5° (lat.) × 5° (lon.) grids.

5. Evaluation of Model Bias

[32] We first diagnose model bias by comparing the CTL against the observations. The CTL is also used as a basic reference in the discussions of assimilation experiments below.

[33] Figure 1 shows the MD between modeled and observed temperature and salinity at the depths of 45 m, 105 m and 180 m based on the 3-year data, respectively. These levels are selected due to either the large model bias or their importance for seasonal climate prediction. For example, the depths of 45 m and 180 m have large cold and warm biases respectively whereas the 105 m depth is within the thermocline at the equator, a very important region for ENSO forecasting. Figure 1 was obtained using all Argo profiles for the 3 year interval. As can be seen, there are large-scale cold biases at 45 m and 105 m and a warm bias at 180 m in the northern Pacific and the subtropical southern Pacific Ocean. For salinity, its MD shows a similar spatial structure in all of the three levels with positive bias (saltier water) in the northern Pacific and negative bias (fresher water) in the central to eastern equatorial Pacific. The spatial distribution of the MD dominates the spatial distribution of RMSE as shown in Figure 2. Corresponding to the large temperature MDs shown in Figure 1, the RMSEs of temperature are larger than 1°C in most parts of the Pacific Ocean at 105 m and 180 m and most parts of the North Pacific Ocean at 45 m. It should be noted that Figures 1 and 2 are based on the data for the period 2005–2007, during which there were two El Niño events (2004–2005, 2006–2007) and two La Niña events (early 2006 and 2007–2008). The cooler upper ocean temperature in the CTL might be due to the model's inability to simulate the large positive anomalies corresponding to these El Niño events. Similarly, the large RMSE regions correspond to large MD regions for salinity (e.g., the northeast Pacific). The consistent distribution between MD and RMSE indicates that the model mean bias probably dominates the model RMSE skill.

[34] To examine the time dependence of the MD, we plotted the time-longitude diagram of the temperature MD along the equator at the three depths, as shown in Figure 3. It can be seen that the large temperature MD has strong persistence. For example, at the depths of 45 m and 105 m,

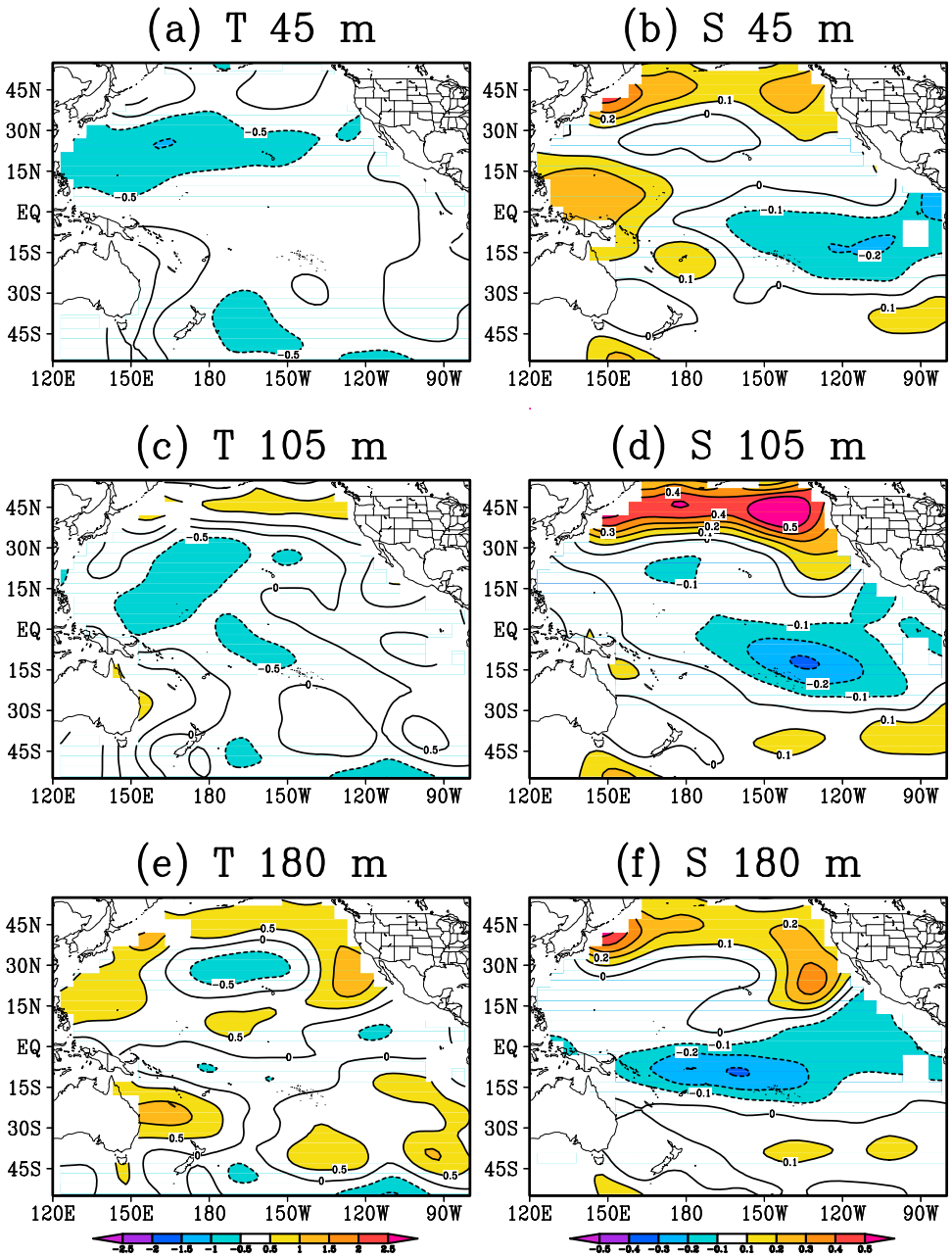

Figure 1. Temperature (Figures 1a, 1c, and 1e) and salinity (Figures 1b, 1d, and 1f) differences between model and observation averaged over the 3-year period at (a and b) 45 m, (c and d) 105 m and (e and f) 180 m for the control run. The contour interval is 0.2°C for temperature and 0.1 psu for salinity.

the warm water mainly lies in the eastern tropical Pacific Ocean while the cold water occupies the region from the west to central tropical Pacific. There are always lower temperatures along the equator at the 180 m. These large spatial-temporal scale warmer (or cooler) waters, which can be consider to be model biases, persist for relatively long periods with a varying amplitude with time.

6. Validation of Assimilation Experiments

[35] In this section, we evaluate the six assimilation experiments described in section 3. Emphasis is placed on the comparison among these experiments, revealing the impact of different error schemes (model biases and random errors) on assimilation performance.

6.1. Temperature and Salinity

[36] As the 10% withheld Argo profiles are not enough to produce spatial distribution maps of MD and RMSE, all Argo profiles for the 3 years are used for calculating the MD and RMSE within each grid cell. Good skill is expected when validating against the same observations that were directly assimilated. To eliminate this artificial skill, we used the 5-day forecast ensemble mean instead of the oceanic analysis for validation. The observations within a time window of 5 days centered at the forecast date are used. Figure 4 shows

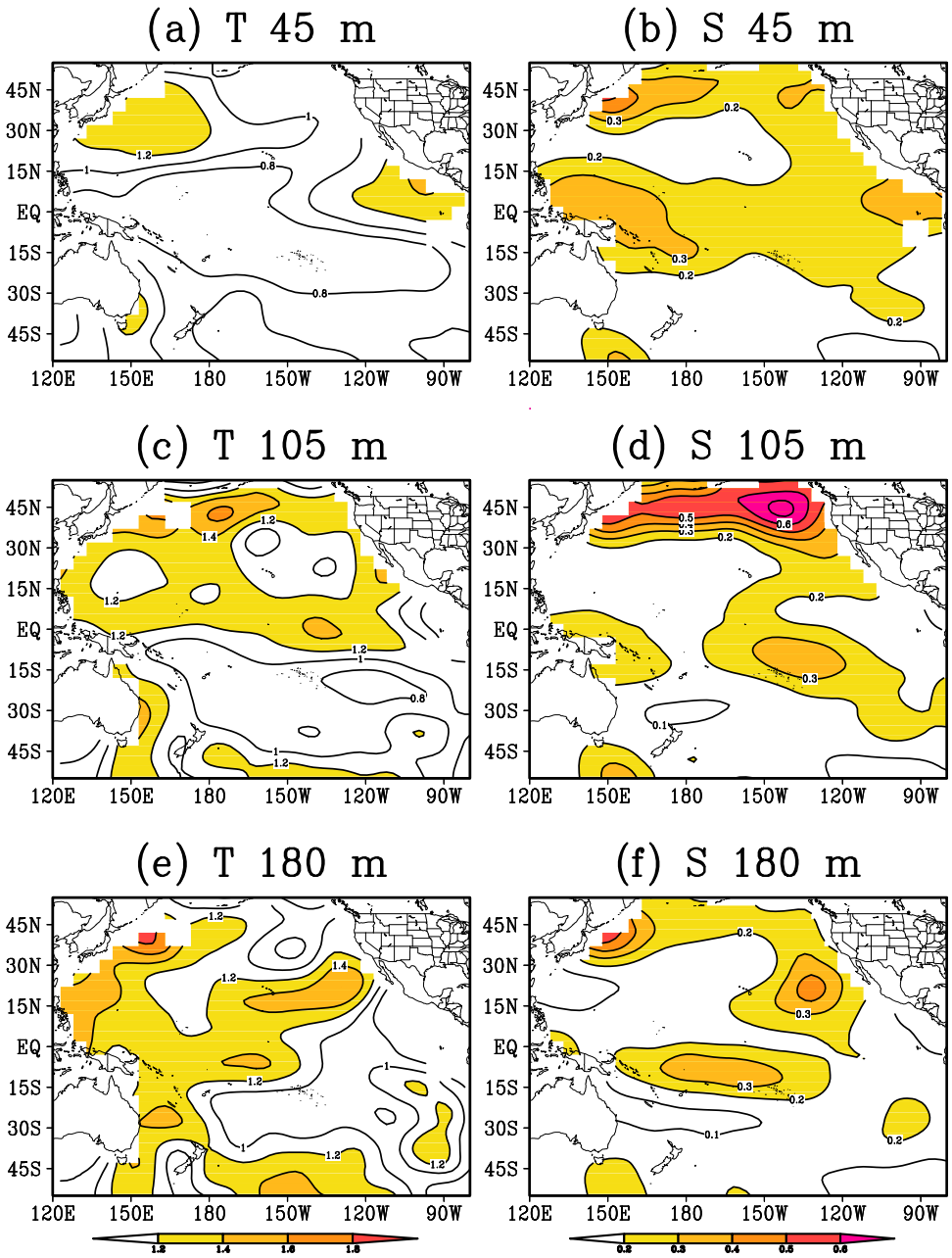

Figure 2. Temperature (Figures 2a, 2c, and 2e) and salinity (Figures 2b, 2d, and 2f) RMSEs between model and observation averaged over the 3-year period at at (a and b) 45 m, (c and d) 105 m and (e and f) 180 m for the control run. The contour interval is 0.5°C for temperature and 0.1 psu for salinity.

the spatial distributions of temperature MD of the six experiments at the depth of 105 m for the period from 2005 to 2007. As illustrated in Figure 4a, though the model is assumed to be perfect, without considering either the biases or the random model errors in Exp1, the large temperature MD values in CTL (Figure 1c) are greatly reduced (under 0.05°C) in most areas of the Pacific Ocean by assimilation. The relatively large MD values, which are still much smaller than those in CTL, mainly occur in the northern subtropical Pacific, around Australia and the eastern equatorial Pacific where there are large biases as shown in CTL (Figure 1c).

[37] Figure 4b shows the spatial distribution of temperature MD of Exp2. The comparison of Figure 4b with Figure 4a indicates that the multiplicative inflation of the model background error covariance can reduce the MD. For example, the large MD values in the eastern equatorial Pacific Ocean, the northern Pacific and around Australia are further reduced. Figure 4c shows the spatial distribution of temperature MD of Exp3, in which additive spatially coherent Gaussian noise are added to each ensemble during the data assimilation. The comparison of Figure 4c and Figure 4a indicates the positive effects of the additive inflation on the reduction of temperature MD in data assimilation. For example, relatively large MDs shown in Figure 4a in the eastern equatorial Pacific Ocean and in the northern Pacific are further reduced. The comparison of Figure 4c with Figure 4b indicates that the

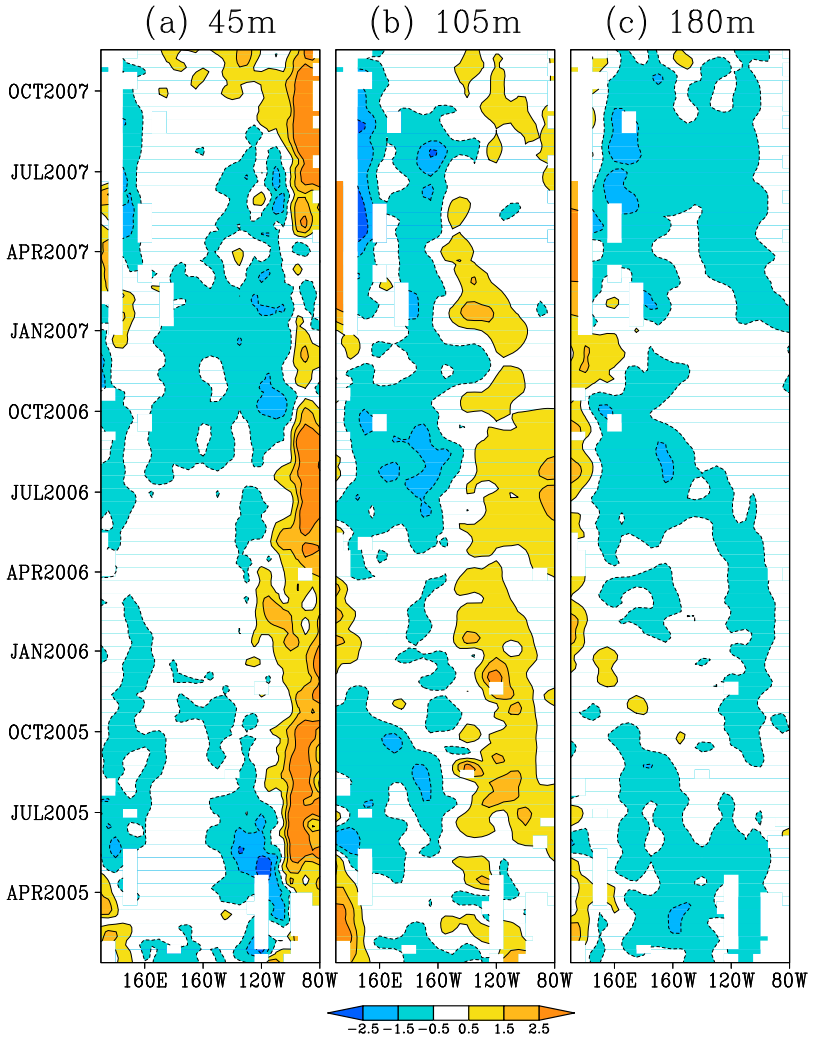

Figure 3. Meridional averaged tropical (10°S–10°N) temperature bias as a function of longitude and time: (a) 45 m, (b) 105 m and (c) 180 m. The contour interval is 1.0°C. Areas with absolute value over 0.5°C are shaded.

additive inflation (Exp3) is more effective than multiplicative inflation (Exp2) in reduction of the MD. Figure 4d shows the MD generated by Exp4, in which a combination of the additive inflation and the multiplicative inflation are used. Comparing Figure 4d with Figures 4b and 4c shows that the combination of these two methods can more effectively reduce MD than using only a single method in some areas, for example in the southwestern Pacific Ocean. The advantage of the combination has also been confirmed by other studies [*Bonavita et al.*, 2010]. The impacts of the two bias correction methods used in Exp5 and Exp6 are shown in Figures 4e and 4f. Actually these two experiments combine the additive inflation, multiplicative inflation and bias correction methods into the data assimilation. Because the MD values have already been significantly reduced in Exp4, the impacts of the two bias correction methods are not evident through comparing Figures 4e, 4f, and 4d. At a model grid when there are sufficient observations in vicinity and the background error covariance is adequately estimated, the traditional ensemble Kalman filter (e.g., Exp4) can effectively correct errors (bias and random error) and produce a sufficiently accurate analysis [*Dee and Todling*, 2000]. The use of IAU enhances the positive impacts of traditional EnKF on the reduction of biases because the IAU procedure consists of an integration of the model forced by the analysis increment including the bias tendency [*Dee and Todling*, 2000]. This is also the reason why experiments 1~4 significantly reduce the MD even though these experiments were not designed to treat bias. After the biases have already been significantly reduced, the effect of bias correction schemes (Exp5 and Exp6) might be lessened. Very small differences between Exp5, Exp6 and Exp4 might be due to the reason that we used a very small γ_2 in equation (12) (1%). In the following discussions, we will show that the bias-aware data assimilation experiments (Exp5 and Exp6) are actually more effective than the bias-blind experiment (Exp4) in reducing RMSE in the first five months (30 assimilation steps), in particular for some model levels.

[38] Figure 5a shows that the assimilation significantly reduces the temperature RMSE in most areas of the Pacific Ocean relative to CTL (Figure 2c). However, in the areas between the equator and 45°N the RMSEs are still large and

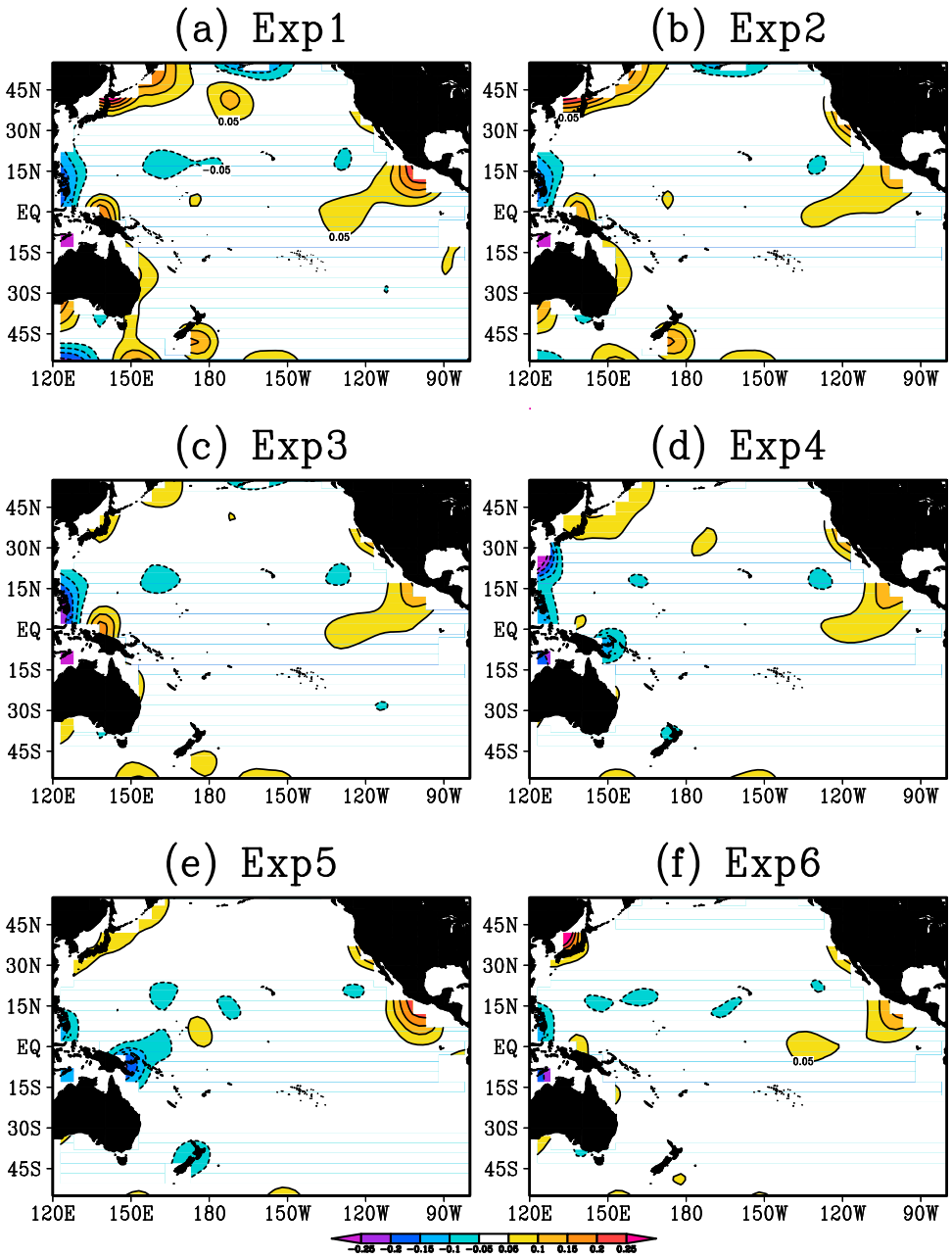

Figure 4. Temperature differences between model and observation averaged over the 3-year period at 105 m for the 6 data assimilation experiments. The contour interval is 0.05°C.

the reduction in RMSE by assimilation is limited. The positive contributions of the additive inflation, multiplicative inflation and their combination in the reduction of temperature RMSE are evident by comparing Figures 5b–5d with Figure 5a. As in the reduction of MD, the contributions of the bias correction schemes in the reduction of RMSE are not very evident as well based on the 3-year data (Figures 5e and 5f).

[39] Figure 6a shows that Exp1 can reduce the salinity MD at the depth of 105 m in the Pacific Ocean as well. The use of multiplicative inflation (Figure 6b) and additive inflation (Figure 6c) can further reduce salinity MD. The combination of the two techniques (Figure 6d) can obtain smaller MD values than only using one method. Comparing Figures 6e and 6f with Figure 6d shows somewhat difference between them indicating the impact of the bias correction scheme. However, we find the difference among the three figures is not evident except that the Exp5 generates relative large MD values in the southeast Pacific Ocean. Similarly, the assimilation system can significantly reduce salinity RMSE (Figure 7a) and the contribution of the multiplicative inflation (Figure 7b), additive inflation (Figure 7c) and the combinations of the two methods (Figure 7d) in reduction of salinity RMSE is evident. Comparing Figures 7e and 7f with Figure 7d, reveals that after the RMSE has already been significantly reduced in Exp4, the contributions of the bias correction schemes in reduction of RMSE are not evident.

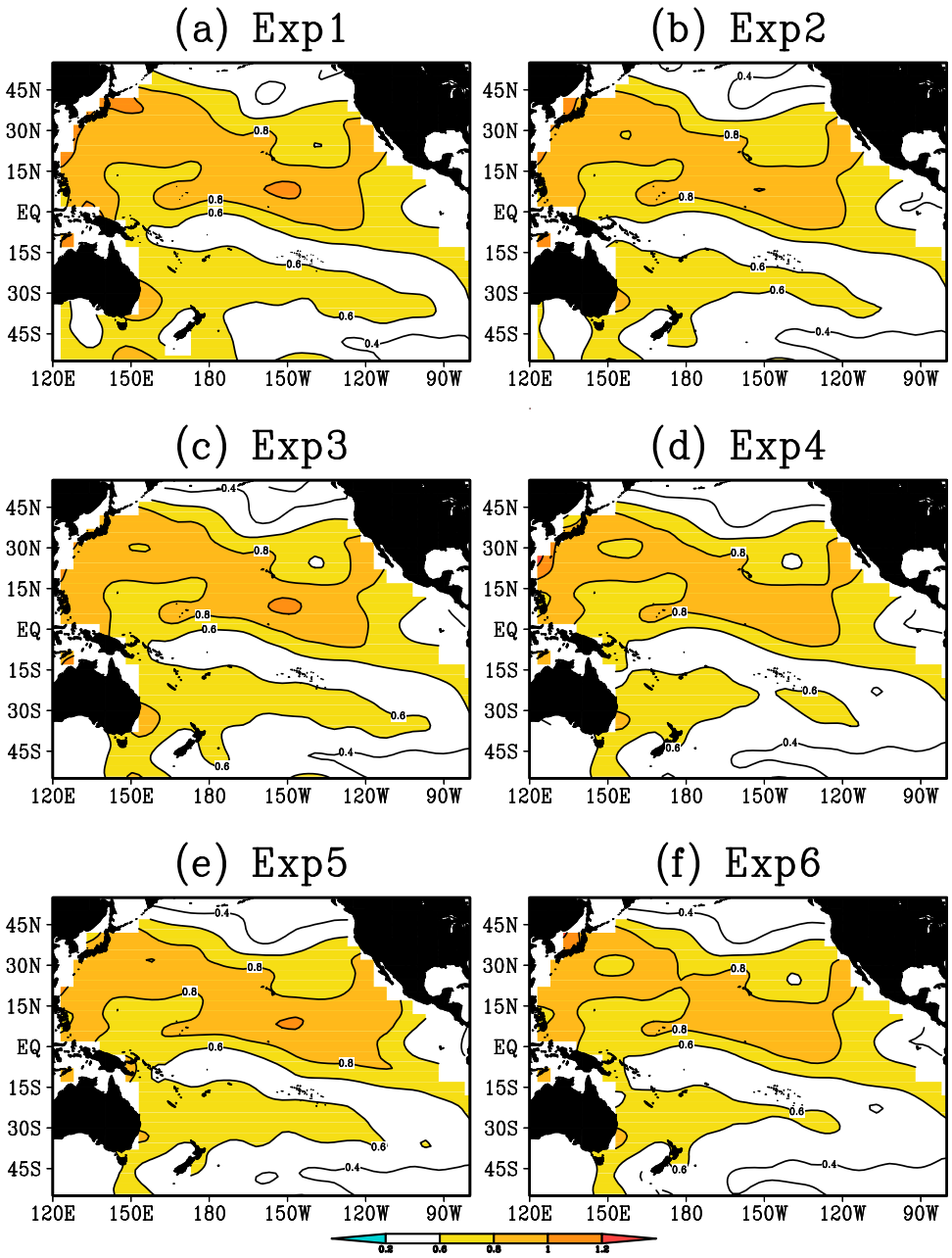

Figure 5. Temperature RMSEs between model and observation averaged over the 3-year period at 105 m for the 6 data assimilation experiments. The contour interval is 0.2°C.

[40] As mentioned above, the impacts of bias correction schemes used in Exp5 and Exp6 are not evident relative to Exp4 when comparing the MDs and RMSEs at 105 m. Because the model error will decrease as the assimilation proceeds, the metrics based on long-term evaluation might muddle the differences among different methods. Here, we only use the first year (2005) for a further comparison. The vertical distributions of temperature and salinity MDs averaged over the year of 2005 for the whole model domain based on the withheld Argo profiles are shown in Figure 8. Compared to CTL, all the experiments generate much smaller temperature and salinity MDs. For temperature (Figure 8a), the Exp3~6 generate similar small MDs in depths between 80 and 1000 m. It is evident that the bias-aware experiments (Exp5~6) generate smaller MDs than the bias-blind experiments (Exp1~4) near the surface and the deep level below 1000 m. The use of inflation methods (Exp2~Exp4) can generate smaller MDs than the traditional method (Exp1). For salinity (Figure 8b), the six experiments also generate similar small MDs between 80 and 1000 m. The bias aware experiments (Exp5~6) generate smaller MDs than the other experiments (Exp1~4) near surface. However, the differences between Exp3~6 are almost undetectable when all withheld data during the 3-year are used in the validation (not shown).

[41] The average over the entire Pacific Ocean might cancel out the regional differences of MD. To explore the spatial differences, the MD and RMSE are also computed for several different regions, including region I (35–50°N, 150–240°E),

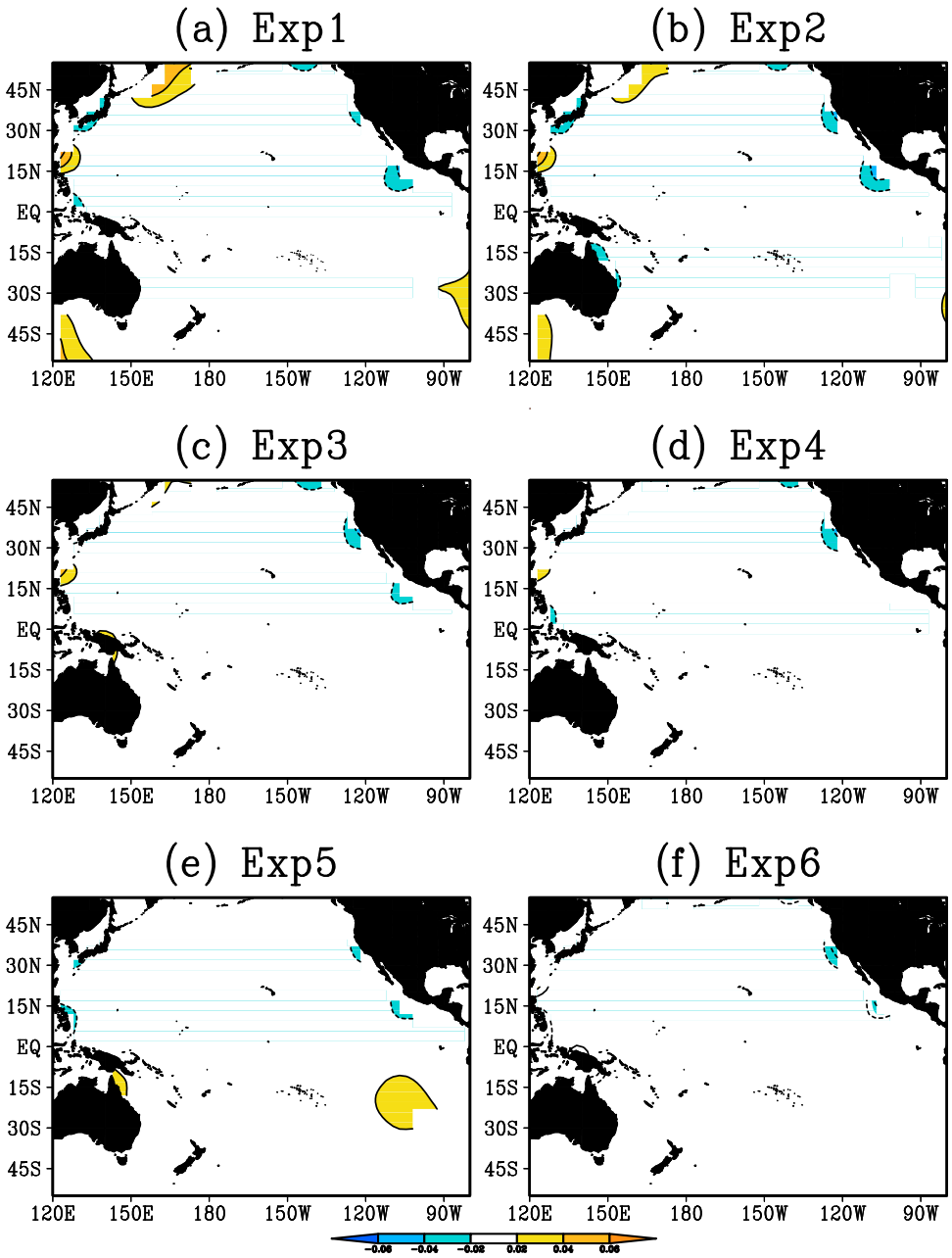

Figure 6. Salinity differences between model and observation averaged over the 3-year period at 105 m for the 6 data assimilation experiments. The contour interval is 0.02 psu.

region II (10°S–10°N, 120–190°E), region III (10°S–10°N, 190–240°E) and region IV (30–45°S, 120–280°E). The regions I-IV were selected to typify the N. Pacific, the western equatorial Pacific, the eastern equatorial Pacific and the S. Pacific, respectively. The regions II and III are also of particular interest for seasonal climate prediction. Figure 9 shows the averaged MD of temperature for the region I~IV. The performance of the experiments with additive inflation (E3~E6) are much better than the perfect model experiments (E1 and E2) for all regions. This result further confirms the finding that ignorance of random model error may not be a viable approach when working with a high-dimensional state vector [*Keppenne et al.*, 2005]. The biases in region II (the western equatorial Pacific) and IV (the southern Pacific) are significantly reduced to very small values at all levels in Exp3~6. However, in region I (the northern Pacific) and III (the eastern equatorial Pacific) there are still relatively large temperature MD values at most levels possibly due to the fact that the observations are sparse in region I and III. The opposite sign of MDs between CTL and the experiments at depths between 40 and 100 m of region III indicates that the experiments may over-adjust the temperature bias. Figure 10 shows the vertical distribution of salinity MD for the four regions. The advantages of the one-step bias correction scheme (Exp5) and the persistent bias correction scheme (Exp6) are evident in region I and III and not very clear in other regions. The vertical distributions of RMSE for the four regions are also calculated. The differences among

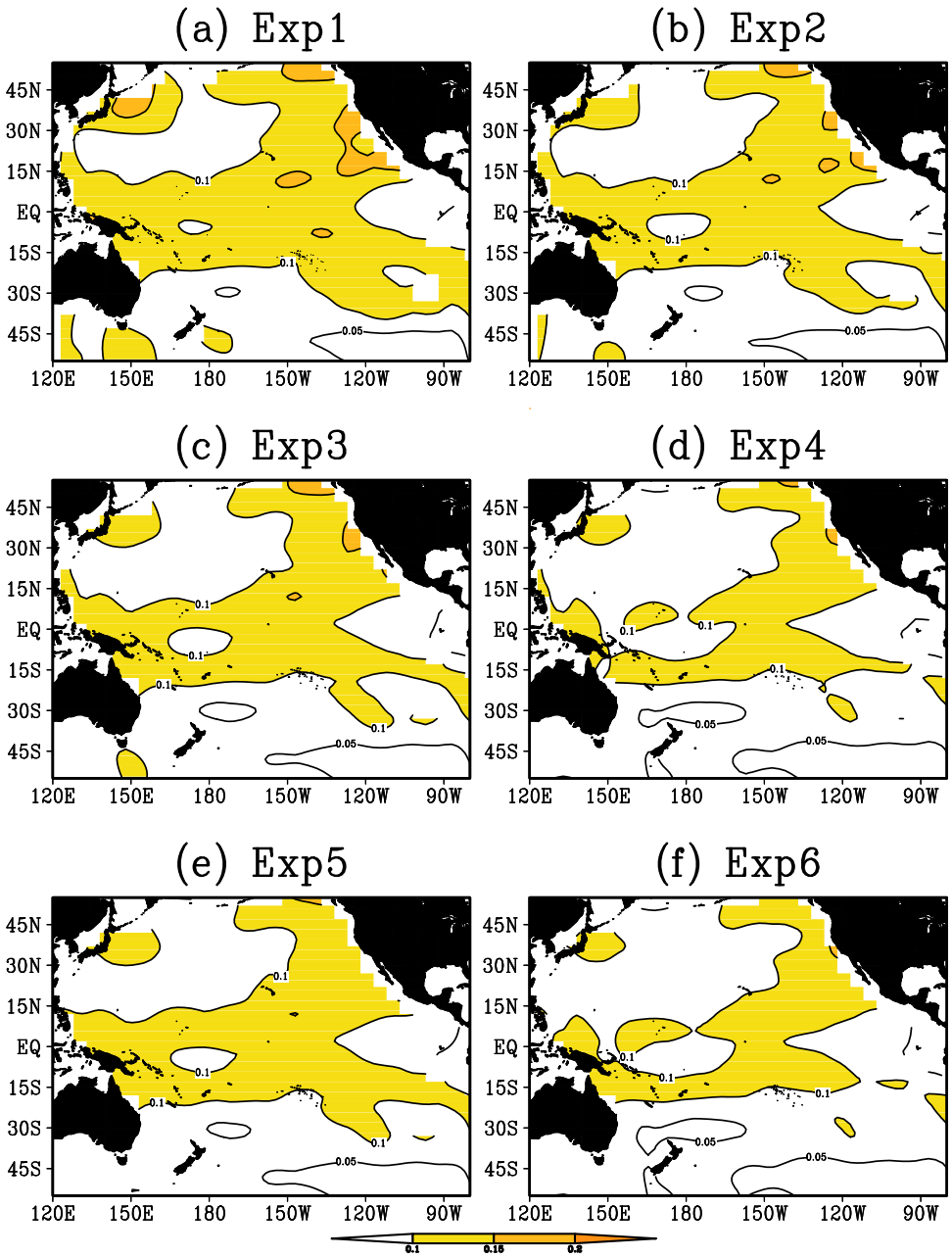

Figure 7. Salinity RMSE between model and observation averaged over the 3-year period at 105 m for the 6 data assimilation experiments. The contour interval is 0.05 psu.

Exp3~6 are evident in region I but not obvious in other regions. Figure 11 shows the vertical distribution of RMSE for the region I. It is clear that Exp6 can gain smaller temperature RMSE and smaller salinity RMSE than Exp4, and Exp5 can generate smaller RMSEs than Exp4 at some model levels, indicating positive impacts of the bias correction schemes.

[42] The above results indicate that the adaptive multiplicative inflation (Exp2) is able to improve the performance of assimilation. The additive inflation (Exp3) results in much smaller forecast errors than Exp1 and Exp2 at most model levels. Combining the additive inflation and the adaptive multiplicative inflation can further improve the performance of the data assimilation system. The incorporation of the bias correction schemes into the assimilation system (Exp5 and Exp6) can further reduce the forecast error relative to Exp4 at some model levels. The impacts of the bias correction schemes are most evident in the northern Pacific Ocean where the observations are relatively sparse and the model has large warmer and saltier biases. However, in regions where the observations are dense, the contributions of bias correction are not evident because the standard assimilation scheme is capable of gaining sufficiently accurate analysis when observations are dense [*Dee and Todling*, 2000].

6.2. Error Evolution With Assimilation Cycles

[43] Insight into whether the assimilation reduces or eliminates systematic errors may be gained by examining spatially

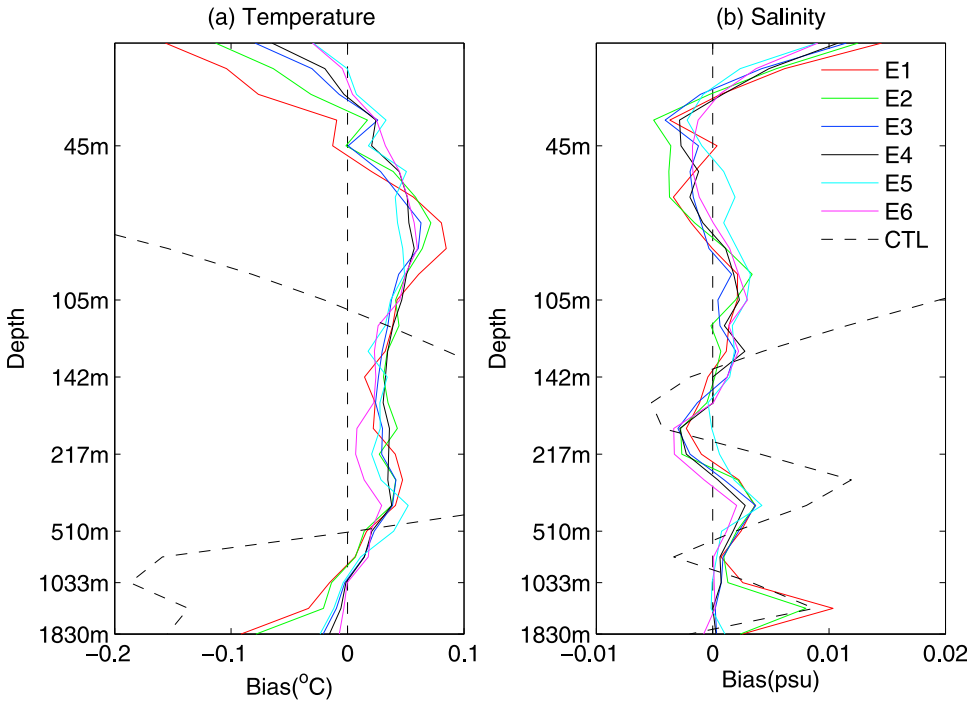

Figure 8. Biases of (a) temperature and (b) salinity as a function of depth, from the control experiment (black dash line), Exp1 (E1:red), Exp2 (E2:green), Exp3 (E3:blue), Exp4 (E4:black solid line), Exp5 (E5:cyan) and Exp6 (E6:magenta) for Pacific Ocean, where the observations are the withheld 10% Argo profiles for the year of 2005.

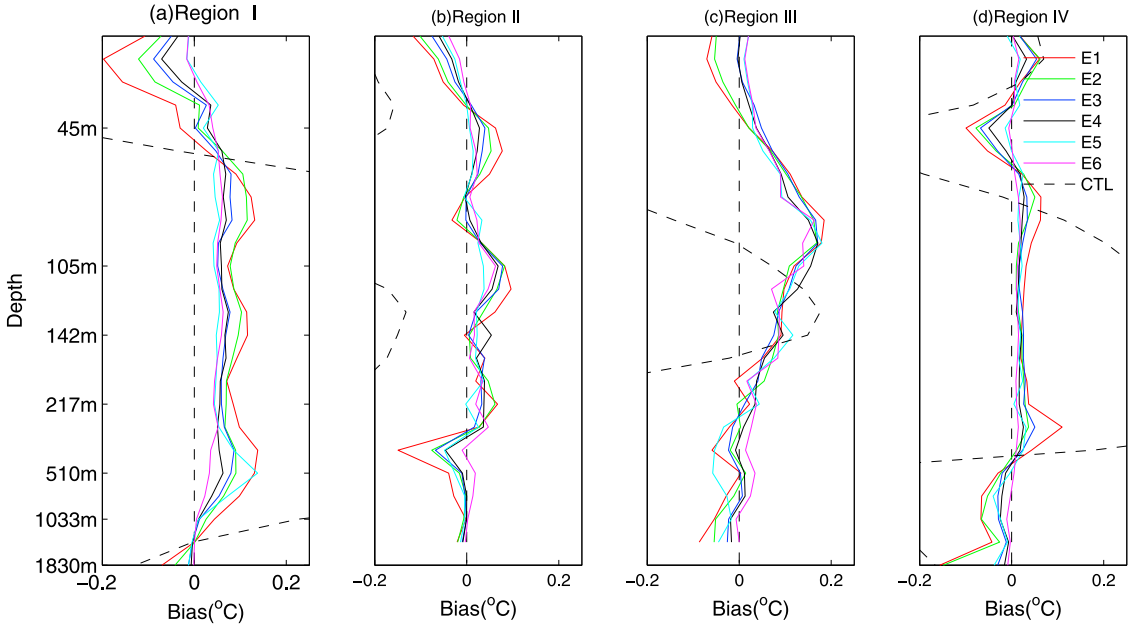

Figure 9. Biases of temperature as a function of depth, from the control experiment (black dash line), Exp1 (E1:red), Exp2 (E2:green), Exp3 (E3:blue), Exp4 (E4:black solid line), Exp5 (E5:cyan) and Exp6 (E6: magenta) for region (a) I, (b) II, (c) III and (d) IV, where the observations are Argo profiles for the year of 2005.

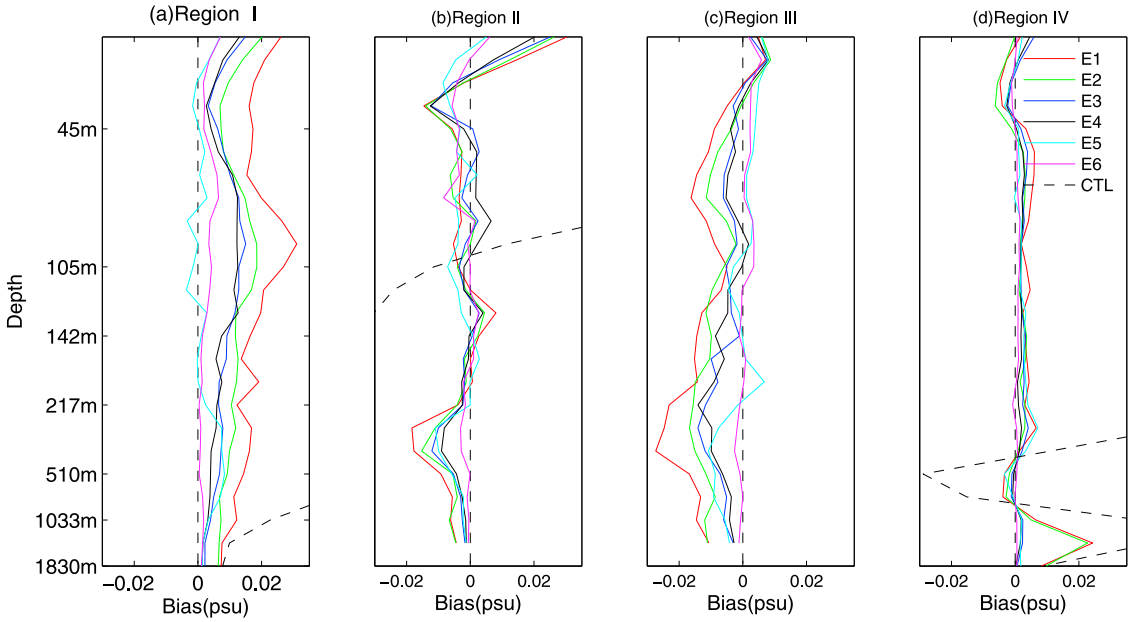

Figure 10. As with Figure 9 but for salinity.

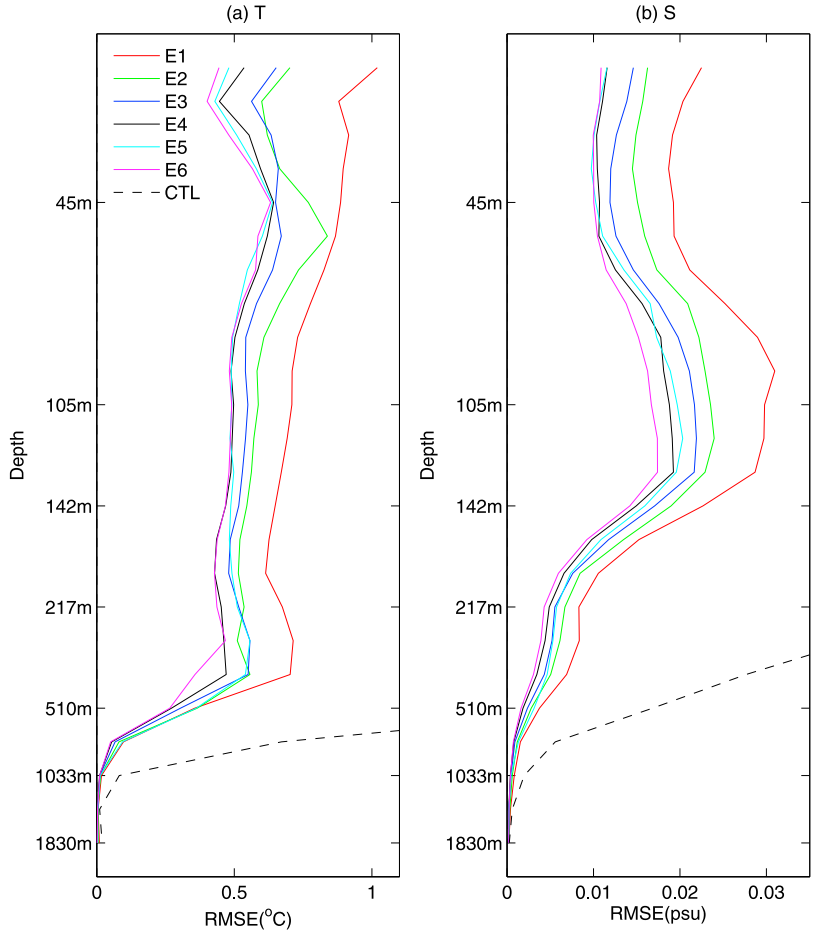

Figure 11. RMSEs of (a) temperature and (b) salinity, as a function of depth, from analyses of the control experiment (black dash line), Exp1 (red), Exp2 (green), Exp3 (blue), Exp4 (black solid line), Exp5 (cyan) and Exp6 (magenta) for the north Pacific Ocean(35–50°N,150–240°E), where the observations are Argo profiles for the year of 2005.

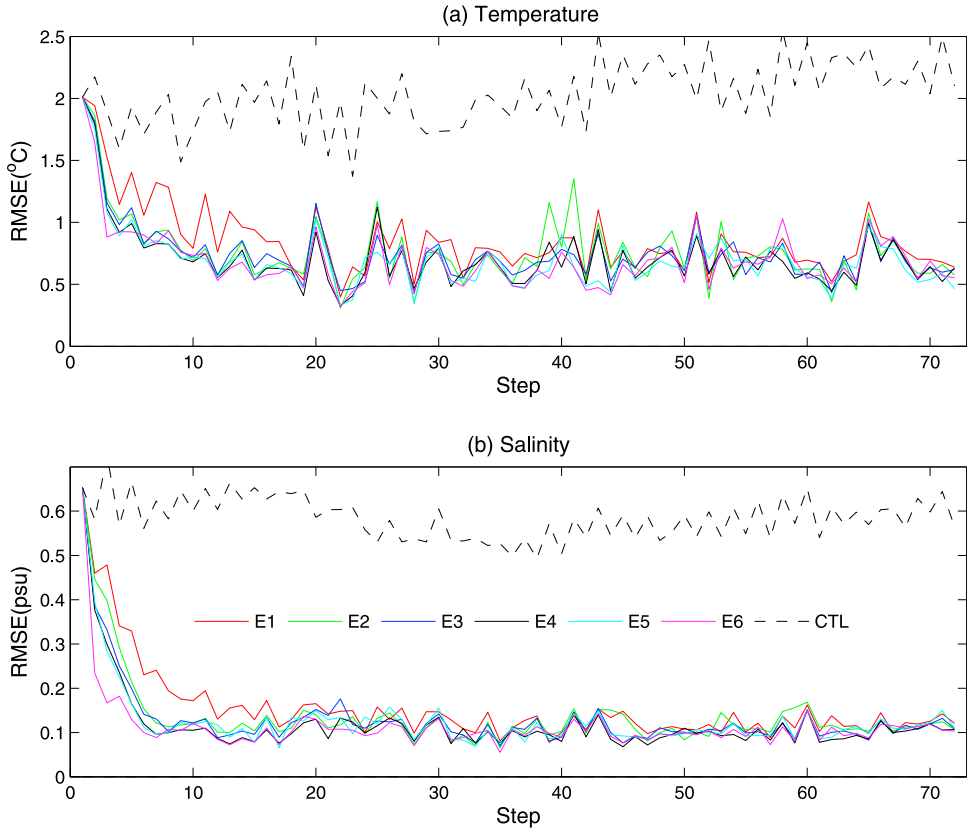

Figure 12. Time evolution of the RMSEs of (a) temperature and (b) salinity at 105 m, as a function of data assimilation step, from analyses of the control experiment (black dash line), Exp1 (red), Exp2 (green), Exp3 (blue), Exp4 (black solid line), Exp5 (cyan) and Exp6 (magenta) for the north Pacific Ocean (35–50°N,150–240°E), where the observations are Argo profiles for the year of 2005.

averaged errors (observation-minus-forecast). In this section, we will explore the temporal evolution of the errors. For convenience, we confine the analysis to the RMSE averaged over the northern Pacific Ocean (i.e., region I) defined using equation (17).

[44] As shown in preceding section, the MD dominates the spatial pattern of RMSE. To show the differences between the 6 experiments clearly, we calculate RMSEs for each data assimilation cycle for the six experiments respectively. Figure 12 only shows the RMSEs at 105 m for region I for the period of 2005, because the differences are not evident among Exp3~6 after 2005. As shown in Figure 12, the RMSEs of temperature and salinity are quickly reduced during the first few assimilation steps for all of the six experiments. The additive inflation scheme is very effective in reducing RMSE as demonstrated by the small RMSE with it (E3~E6) against large RMSE without it (E1~E2). It is also found that the bias correction schemes (E5 and E6) are not very effective in reducing the model bias after around 5 months (30 assimilation steps). For other regions, there are features similar to those in Figure 12 (not shown).

[45] In the EnKF method, the ensemble spread is an important parameter that can be used to explain the performance of the system. Figure 13 shows the temporal variation of the spread of the forecast and analysis ensembles at 105 m in the region I. As can be seen, the ensemble spread presents a typical feature of a good ensemble system in all experiments, i.e., increasing until saturation after some assimilation steps. This increasing ensemble spread allows the forecast to draw to the observations and thus do not encounter filter divergence, which is where bias can really increase. This is an underlying reason why exp 1–4 do reasonably well in reducing MD. The continual increase in ensemble spread is due to the continual perturbations in the wind stress during the assimilation. In similar experiments without the perturbation of wind, the ensemble spread is quickly damped after 5–10 steps (not shown). The saturation is probably a consequence of the adjustment of model dynamics. The spreads of Exp3–6 retain relatively large values and are closer to the RMSE (Figure 12), compared with those of E1~2. A good ensemble system should show an ensemble spread close to the RMSE [e.g., *Cheng et al.*, 2010]. Thus, Exp3–6 is better than E1~2, showing the advantage of the inflation method.

[46] The standard implementation of the Kalman filter is able to provide an efficient analysis where measurements are available [*Drécourt et al.*, 2006]. This is true when the amplitude of the model background error covariance can reflect the amplitude of total model errors including both bias and random model error. When we estimate the parameter λ_2 through equation (6), the bias and random model error are combined because the observation-minus-forecast covariance consists of these two components. So, the Exp4 will always generate results similar to Exp5–Exp6. Comparing

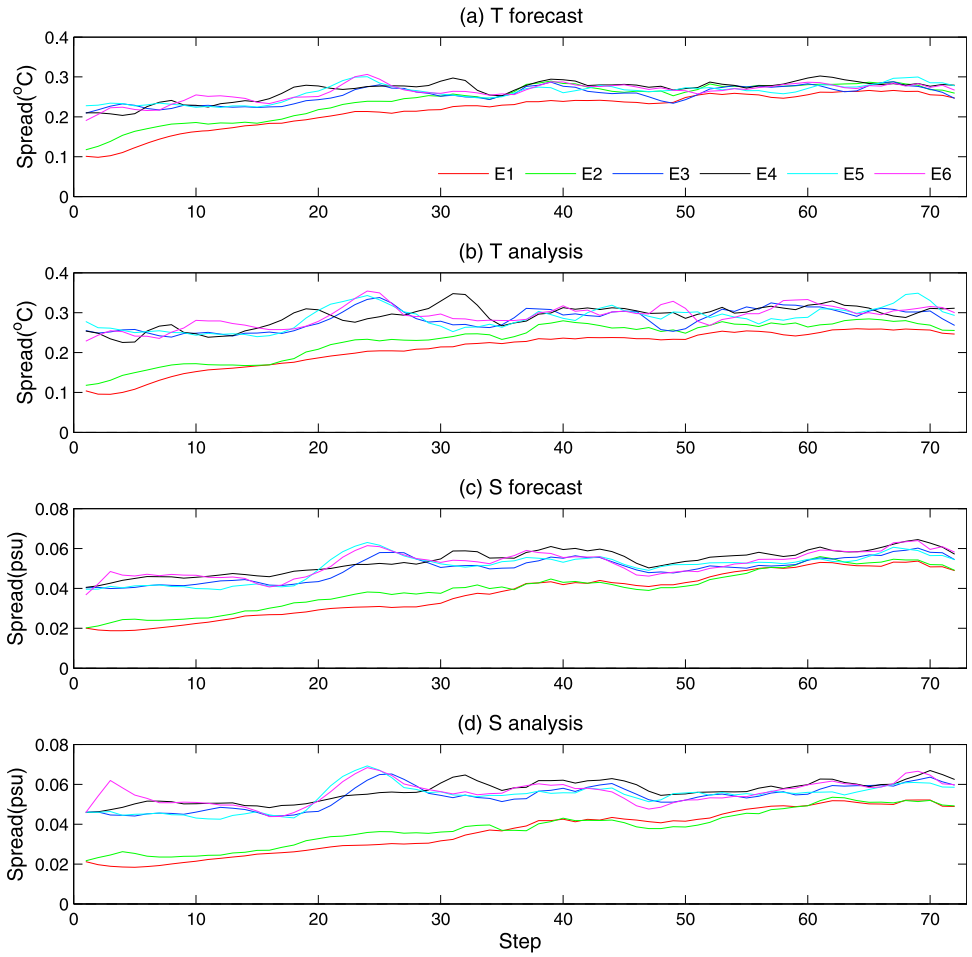

Figure 13. Time evolution of the forecast ensemble spreads of (a) temperature and (c) salinity and analysis ensemble spreads of (b) temperature and (d) salinity at 105 m, as a function of data assimilation step, from analyses of the control experiment (black dash line), Exp(red), Exp2 (green), Exp3 (blue), Exp4 (black solid line), Exp5 (cyan) and Exp6 (magenta) for the north Pacific Ocean(35–50°N,150–240°E), where the observations are Argo profiles for the year of 2005.

the Kalman gains of the six experiments listed in the appendix, we can see that the gains of Exp4–6 have the same pattern. After model biases are significantly reduced, the results of these three experiments should be similar especially when we validate them against long-term observational data.

7. Further Validation Using Independent Data

[47] The surface current data from the Ocean Surface Current Analyses-Real time [*Helber et al.*, 2007] and the sea level height anomaly data (SLA) from (M)SLA and (M)ADT near-real time and delayed time products [*Dibarboure et al.*, 2009] are used for further validation. The correlation coefficients and RMSEs between the observed and simulated zonal currents generated by CTL and the six experiments are calculated on 5° grids. To show the impacts of different schemes, the differences in correlation and RMSEs between the six experiments and CTL are also computed for a period of 3 years. Figure 14 shows the correlation coefficients of CTL (Figure 14a), and the correlation differences between the assimilation experiments and CTL. In Figures 14b–14g positive values show where assimilation experiments improved the simulation of currents, and negative values show where they did a worse job. Figure 14a illustrates a good simulation of surface current in the tropical Pacific regions by CTL but a poor simulation between 15°N–30°N. In contrast, the assimilation experiments significantly improve the surface current simulation in most areas between 15°S–30°N (Figures 14b–14g). At higher latitudes and in the tropical western Pacific, the correlation skill is lower from the assimilation than from CTL, probably due to the small number of Argo profiles in these regions. Comparing Figure 14 with the number of Argo profiles as shown in Figure 1 of *Deng et al.* [2010], one can find that these relative poor skills in assimilation coincide with areas having few Argo profiles. Dynamical adjustment plays a critical role in the surface current analyses when assimilating the temperature (T) and salinity (S) profiles. The assimilation of insufficient Argo T-S profiles actually could greatly degrade the dynamical balance and make dynamical adjustment harder, leading to skill even poorer than the control run. Another possible reason for poor skills at higher latitudes is probably related to the design of the localization box. As the box is shifted poleward from the equator it would begin to

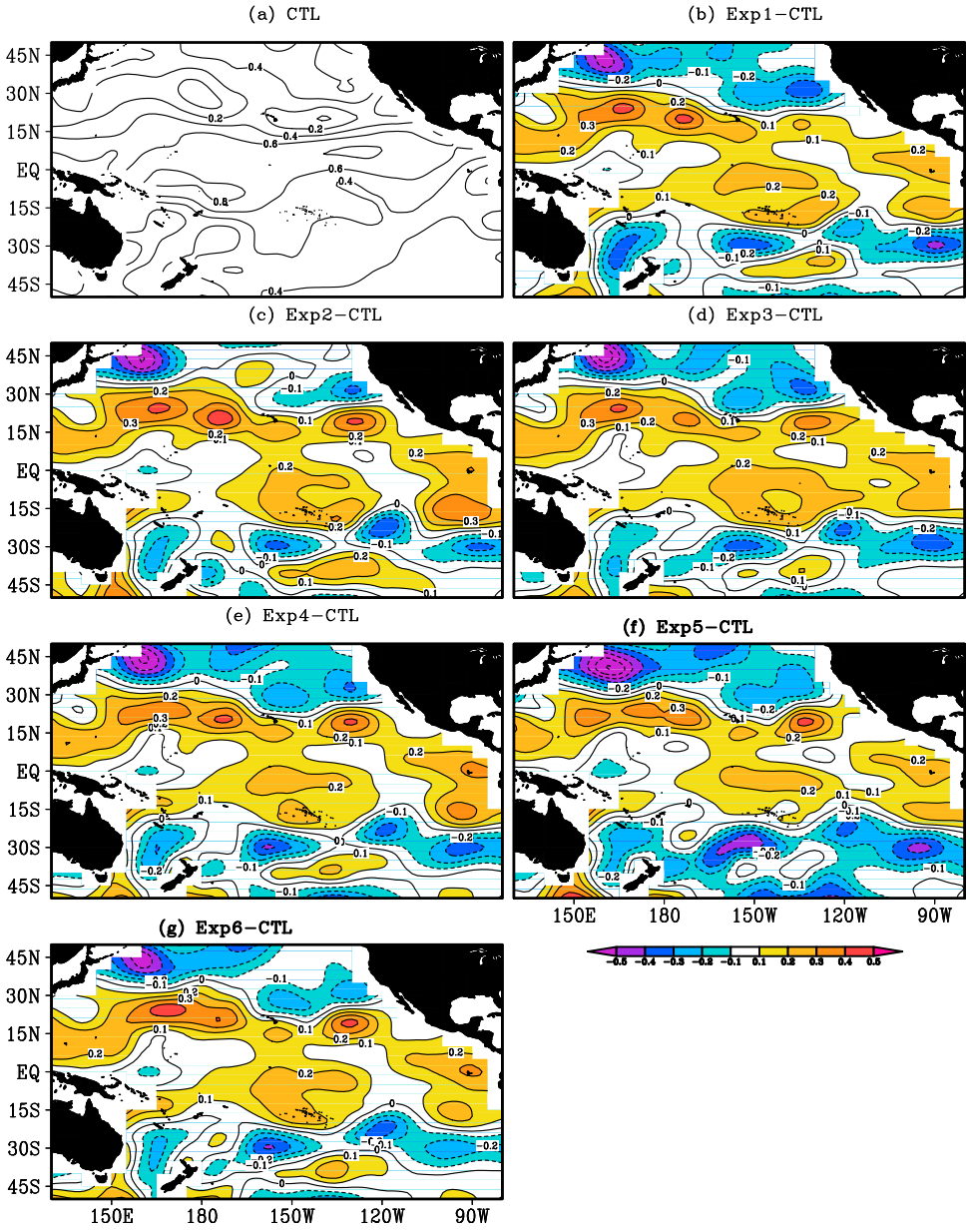

Figure 14. (a) Correlation between the observed zonal surface current and the CTL. (b–g) The differences between the correlations of the 6 experiments and the CTL (R_{exps}-R_{CTL}). Contour interval is 0.2 in Figure 14a and 0.1 in Figures 14b–14g. Areas with absolute value over 0.1 are shaded in Figures 14b–14g.

span parts of the ocean with very different vertical structures. In the Pacific Ocean, the thermocline deepens and broadens from the equator reaching its deepest point at about 30 N, and then rises to break the surface farther north. The complex vertical structure is a great challenge in the design of localization. In addition, relative small mean and variability in the currents in these regions (about 10~20% of the values of the currents in the tropics) make them very sensitive to the analysis increments, easily leading to poor skill in correlation and RMSE.

[48] The most evident improvement of RMSE skill due to assimilation occurs in the equatorial central and the eastern Pacific, the area with relatively more Argo profiles and relatively stronger equatorial currents, as shown in Figure 15. In other regions, the surface currents are weak, associated with a small RMSE in all experiments, leading to little evident impact of assimilation on RMSE skill. This is the reason why the six experiments have small differences in RMSE skill. However, a comparison between Figures 14c–14g and Figure 14b reveals that the experiments with inflation methods can improve the simulation of surface currents in most areas over experiments without them.

[49] Correlations and RMSEs between the observed and simulated sea level height are also calculated grid by grid. Figure 16a shows the correlation between the observed and simulated sea level height anomalies by CTL for the three year interval. The CTL experiment generates a good simulation of SLA in the tropical Pacific but does badly in the region between 15° and 30° in both the hemispheres. All of the six experiments (Figures 16b–16g) perform better in the

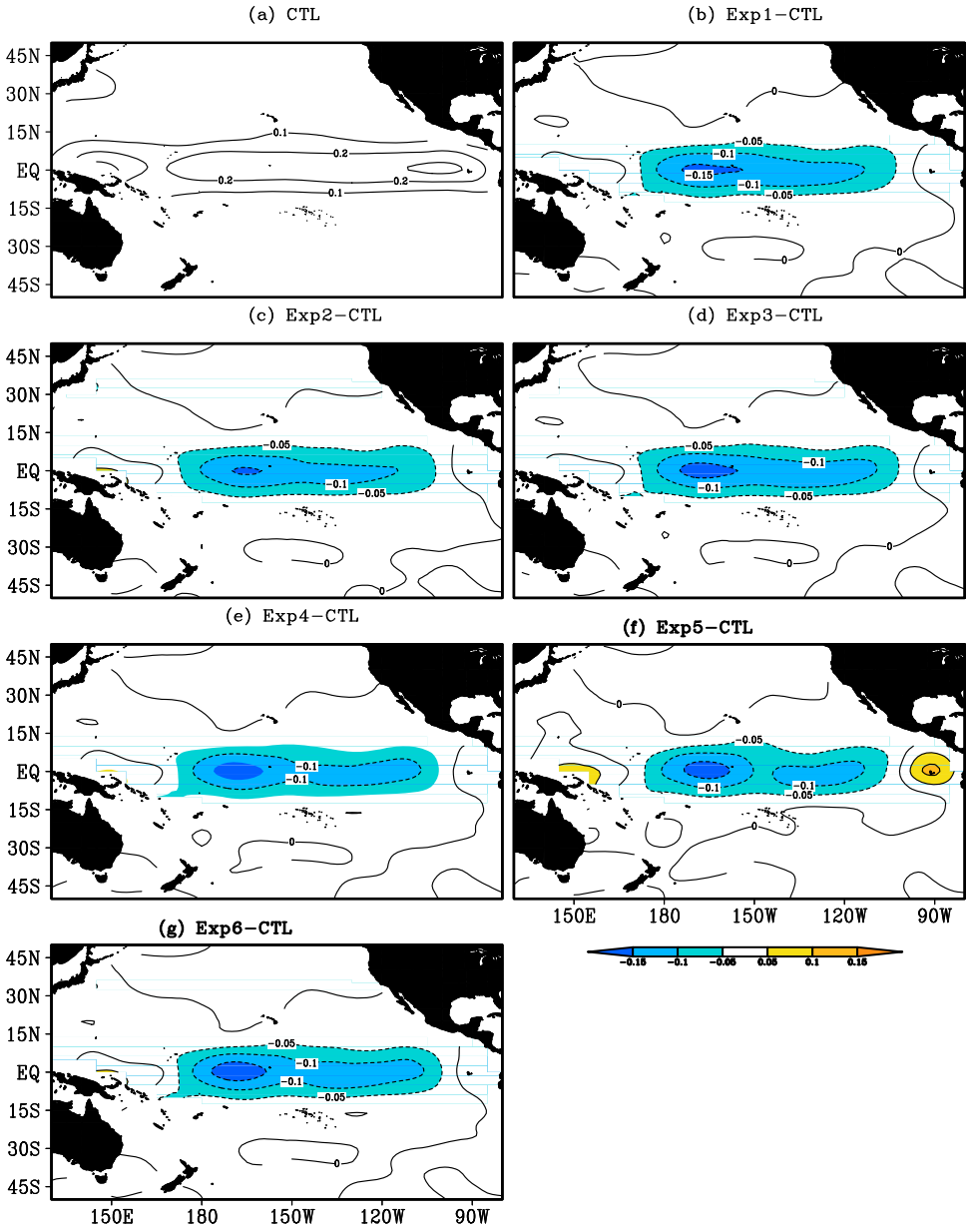

Figure 15. (a) RMSE between the observed zonal surface current and the CTL. (b–g) The differences between the RMSEs of the 6 experiments and the CTL ($RMSE_{exps}$-$RMSE_{CTL}$). The interval in Figure 15a is 0.1 m/s and 0.05 m/s in Figures 15b–15g. Areas with absolute value over 0.05 m/s are shaded in Figures 15b–15g.

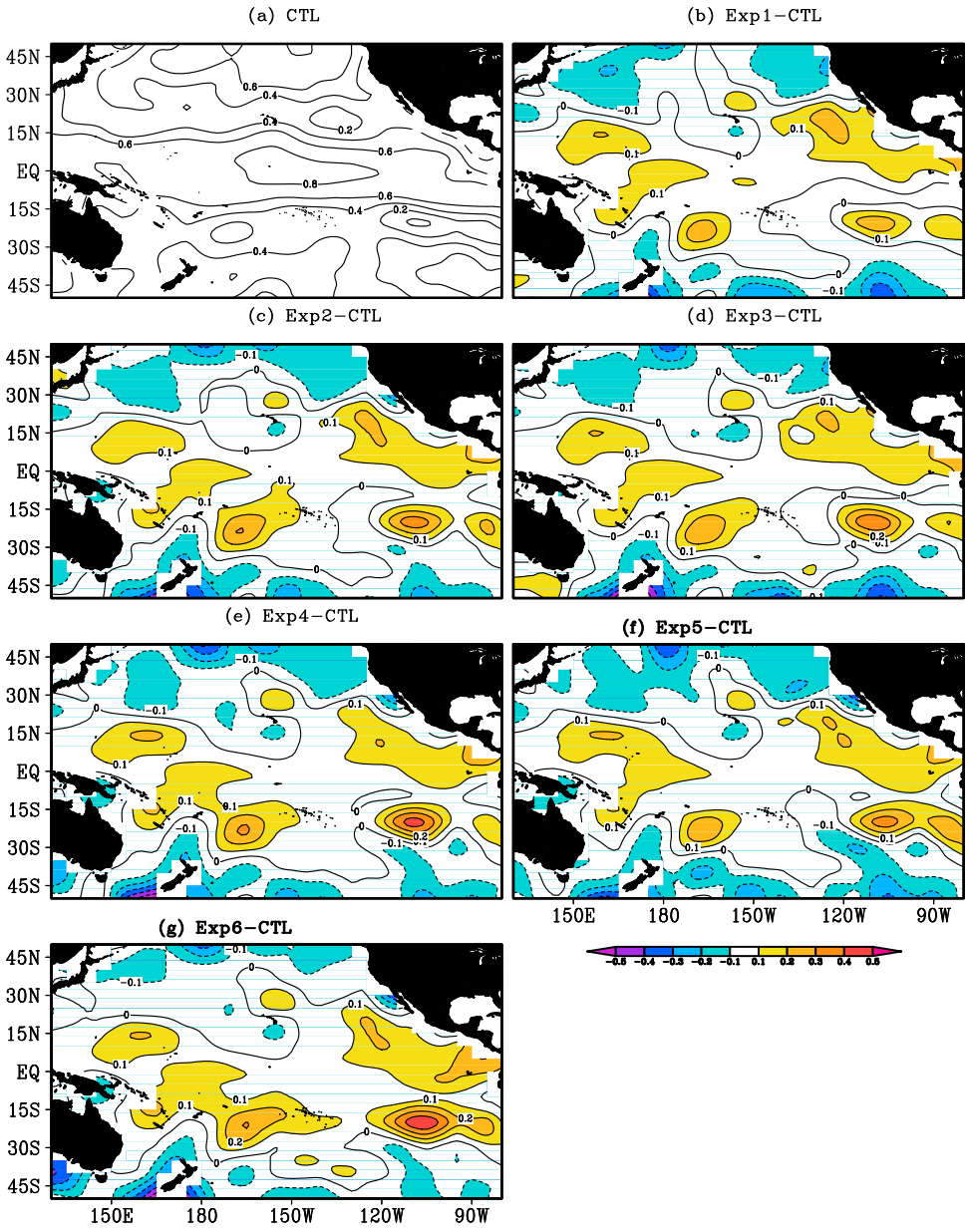

Figure 16. (a) Correlation between mean sea level anomalies of observation and the CTL. (b–g) The differences between the correlations of the 6 experiments and the CTL (R_{exps}-R_{CTL}). Contour interval is 0.2 in Figure 16a and 0.1 in Figures 16b–16g. Areas of absolute value over 0.1 are shaded in Figures 16b–16g.

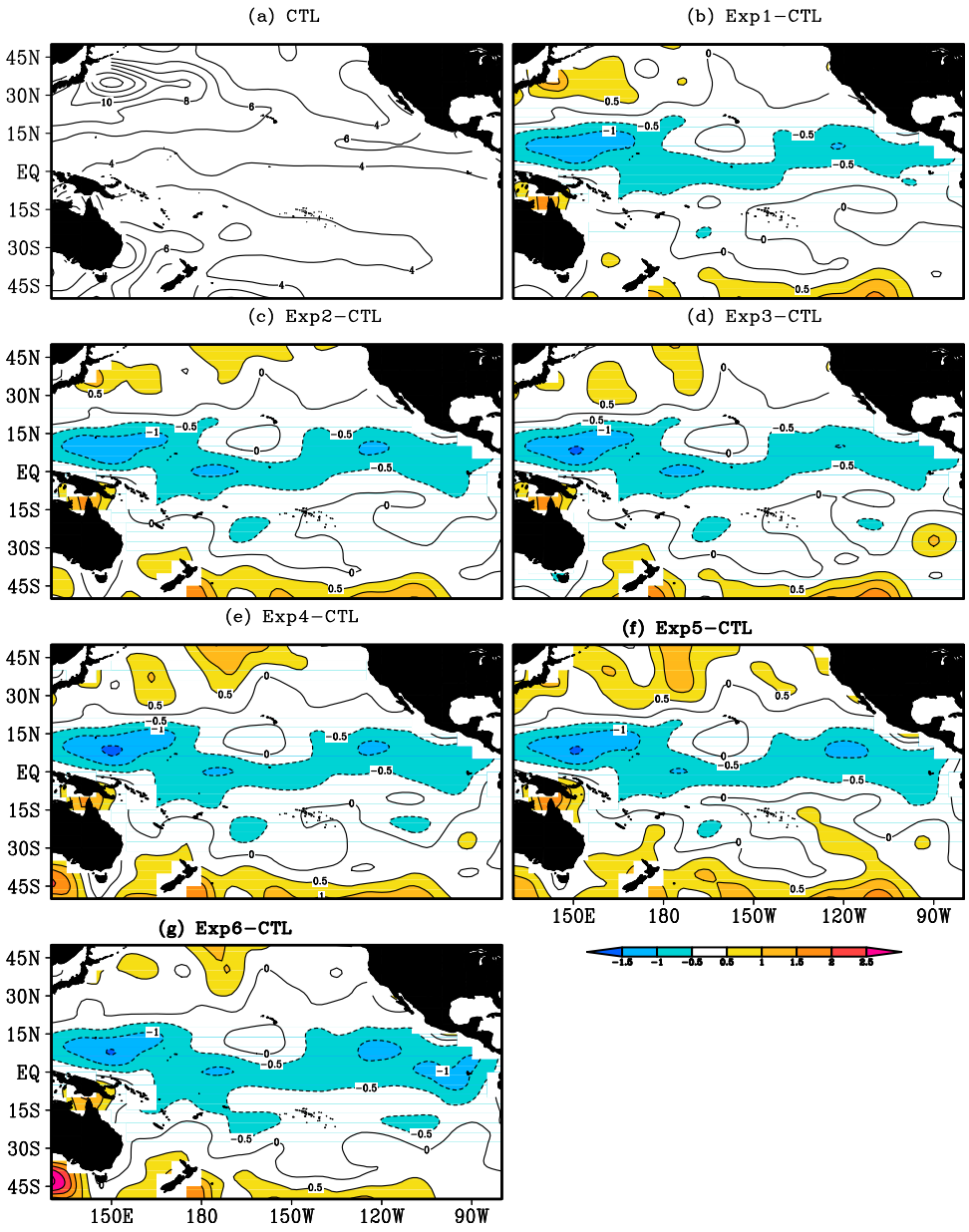

Figure 17. (a) RMSE of mean sea level anomalies (cm) between the observation and the CTL. (b–g) The differences between the RMSEs of the 6 experiments and the CTL ($RMSE_{exps}$-$RMSE_{CTL}$). Areas with absolute value over 0.5 cm are shaded in Figures 17b–17g.

tropical Pacific between 30°S–30°N than CTL, which is important for ENSO prediction. Exp2~6 produce somewhat better results than Exp1 in most areas. Corresponding to the improvements in correlation, the six experiments (shown in Figures 17b–17g) generate smaller RMSEs than CTL (Figure 17a) between 20°S–20°N but larger RMSEs at higher latitude, which is most probably due to the paucity of Argo profiles.

[50] An interesting feature in Figures 16 and 17 is an evident improvement of SLA in all assimilation experiments in the tropical western Pacific where Argo profiles are sparse. This is different from the result for surface currents as discussed above, and is not well understood. A possible reason is that the SLA is mainly controlled by thermodynamics and surface pressure, which are highly related to temperature and salinity. The propagation of the equatorial Kelvin waves and the off-equator Rossby waves can quickly bring thermodynamic information (e.g., thermocline adjustment) from the central and the east to the west.

8. Summary

[51] An adequate estimate of random model error and model bias is an important issue in data assimilation. In this paper, six experiments were performed to examine the capability of different model error schemes to reduce these model errors for the Pacific Ocean. The delayed-mode Argo, XBT, CTD, and TAO-TRITON profiles for the period

2005–2007 were assimilated for this purpose. The validation and comparisons of these experiments indicated that all of the six assimilation experiments can obtain better simulation of temperature and salinity than the control run without assimilation. Assimilation of temperature and salinity also improved the simulation of surface zonal currents and SLA in the tropical Pacific but degraded the simulation at higher latitudes. However, the degradations of zonal surface current and SLA probably were due to a relatively sparse distribution of Argo profiles. It is interesting that all the assimilations significantly improved simulation of the tropical Pacific Ocean, which can benefit ENSO prediction. The additive inflation was found to be important in maintaining ensemble spread amplitude in an appropriate range thus alleviating the underestimation of the background error covariance. The advantage of additive inflation is most evident near the surface and in the deep ocean. The multiplicative inflation also contributed positively to improvement in the assimilations. The combination of additive inflation and adaptive multiplicative inflation further improved the performance of assimilation, and was better than using either one. The combination with the additive inflation, adaptive multiplicative inflation and the bias correction schemes is appropriate only for a relatively short initial period of assimilation. With any increase of assimilation cycles, the advantages of bias correction schemes gradually disappeared since the errors had become small after only a few assimilation steps.

[52] Several other sensitivity experiments were also conducted as part of this study. For example, we used bias correction schemes without additive inflation and/or multiplicative inflation. The results are a little better than Exp1 and Exp2 but worse than Exp3–6. We also found that if the amplitude of the additive model error is significantly increased, over-fitting occurs after the first few assimilation cycles. In the bias correction schemes, Kalman gains for bias analysis are assumed to have the same structure but much small amplitude, as those for the state analysis. The misrepresentation of bias gain might affect the performance of the assimilation system.

[53] This paper explores the impacts of different model error schemes on ocean assimilation, especially on the assimilation of Argo data, in the framework of the EnKF. Some similar comparisons have been made in atmospheric data assimilation, but this kind of work is new in ocean data assimilation, in particular within the Argo data assimilation community. The use of a realistic OGCM and the Argo data for assimilation experiments makes this study not only useful in selecting appropriate error schemes in oceanic data assimilation but also has practical significance for Argo data assimilation.

[54] In this study, only the Argo and other in situ profiles are assimilated to explore the impact of several different model error schemes on ocean assimilation. Neglect of other observations such as satellite SST and SLA, as well as oceanic currents etc. may impact ocean state estimation. The ensemble size in this study is not very large and that probably affects the assimilation performance despite the use of a local analysis strategy that can reduce the impact of the small ensemble size on the efficiency of the EnKF. Also, the localization scale and the additive model error amplitude are chosen somewhat subjectively. All of these issues may impact the assimilation performance and need to be addressed in future studies.

Appendix A: The Kalman Gains of the Six Experiments

[55] Given $P = \rho \circ P^f$, the Kalman gains of the six experiments can be written as

$$K_{E1} = PH^T\left(HPH^T + R\right)^{-1} \tag{A1}$$

$$K_{E2} = \left(PH^T + \lambda_2 PH^T\right)\left[\left(HPH^T + R + \lambda_2 HPH^T\right)^{-1}\right] \tag{A2}$$

$$K_{E3} = \left(PH^T + QH^T\right)\left[\left(HPH^T + R\right) + HQH^T\right]^{-1} \tag{A3}$$

$$K_{E4} = \left[PH^T + (Q + \lambda_2 P + \lambda_2 Q)H^T\right] \cdot \left[\left(HPH^T + R\right) + H(Q + \lambda_2 P + \lambda_2 Q)H^T\right]^{-1} \tag{A4}$$

$$K_{E5} = \left[PH^T + (Q + \lambda_2 P + \lambda_2 Q)H^T\right] \cdot \left[\left(HPH^T + R\right) + H(Q + \lambda_2 P + \lambda_2 Q)H^T\right]^{-1} \tag{A5}$$

$$K_{E6} = \left[PH^T + (Q + \lambda_2 P + \lambda_2 Q)H^T\right] \cdot \left[\left(HPH^T + R\right) + H(Q + \lambda_2 P + \lambda_2 Q)H^T\right]^{-1} \tag{A6}$$

[56] **Acknowledgments.** This work is supported by BC-China Innovation and Commercialization Strategic Development Program (ICSD-2007-Tang-Y). We thank Tom Hamill and one anonymous reviewer for their constructive comments, which greatly improved this manuscript.

References

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

Balmaseda, M., D. Dee, A. Vidard, and D. Anderson (2005), A multivariate treatment of bias for sequential data assimilation: Application to the tropical oceans, *Q. J. R. Meteorol. Soc.*, *128*, 1–20.

Balmaseda, M., D. Anderson, and A. Vidard (2007), Impact of Argo on analyses of the global ocean, *Geophys. Res. Lett.*, *34*, L16605, doi:10.1029/2007GL030452.

Bellucci, A., S. Masina, P. Di Pietro, and A. Navarra (2007), Using temperature-salinity relations in a global ocean implementation of a multivariate data assimilation scheme, *Mon. Weather Rev.*, *135*, 3785–3807, doi:10.1175/2007MWR1821.1.

Bloom, S. C., L. L. Takas, A. M. da Silva, and D. Ledvina (1996), Data assimilation using incremental analysis updates, *Mon. Weather Rev.*, *124*, 1256–1271, doi:10.1175/1520-0493(1996)124<1256:DAUIAU>2.0.CO;2.

Bonavita, M., L. Torrisi, and F. Marcucci (2010), Ensemble data assimilation with the CNMCA regional forecasting system, *Q. J. R. Meteorol. Soc.*, *136*, 132–145, doi:10.1002/qj.553.

Bonjean, F., and G. S. E. Lagerloef (2002), Diagnostic model and analysis of the surface currents in the tropical Pacific Ocean, *J. Phys. Oceanogr.*, *32*, 2938–2954, doi:10.1175/1520-0485(2002)032<2938:DMAAOT>2.0.CO;2.

Carval, T., et al. (2006), Argo data management, user's manual, version 2.1, *Rep. ar-um-02–01*, Argo, Ramonville, France.

Castruccio, F., J. Verron, L. Gourdeau, J. M. Brankart, and P. Brasseur (2008), Joint altimetric and in-situ data assimilation using the GRACE mean dynamic topography: A 1993–1998 hindcast experiment in the tropical Pacific Ocean, *Ocean Dyn.*, *58*, 43–63, doi:10.1007/s10236-007-0131-4.

Cheng, Y., Y. Tang, P. Jackson, D. Chen, and Z. Deng (2010), Ensemble construction and verification of the probabilistic ENSO prediction in the LDEO5 Model, *J. Clim.*, *23*, 5476–5497, doi:10.1175/2010JCLI3453.1.

Chepurin, G., J. Carton, and D. Dee (2005), Forecast model bias correction in ocean data assimilation, *Mon. Weather Rev.*, *133*, 1328–1342, doi:10.1175/MWR2920.1.

Chu, C., G. Wang, and C. Fan (2004), Evaluation of the U.S. Navy's modular ocean data assimilation system (MODAS) using South China Sea monsoon experiment (SCSMEX) data, *J. Oceanogr.*, *60*, 1007–1021, doi:10.1007/s10872-005-0009-3.

Danforth, C. M., E. Kalnay, and T. Miyoshi (2007), Estimating and correcting global weather model error, *Mon. Weather Rev.*, *135*, 281–299, doi:10.1175/MWR3289.1.

Dee, D. (1995), On-line estimation of error covariance parameters for atmospheric data assimilation, *Mon. Weather Rev.*, *123*, 1128–1145, doi:10.1175/1520-0493(1995)123<1128:OLEOEC>2.0.CO;2.

Dee, D. (2005), Bias and data assimilation, *Q. J. R. Meteorol. Soc.*, *131*, 3323–3343, doi:10.1256/qj.05.137.

Dee, D., and A. M. da Silva (1998), Data assimilation in the presence of forecast bias, *Q. J. R. Meteorol. Soc.*, *124*, 269–295, doi:10.1002/qj.49712454512.

Dee, D., and R. Todling (2000), Data assimilation in the presence of forecast bias: The GEOS moisture analysis, *Mon. Weather Rev.*, *128*, 3268–3282, doi:10.1175/1520-0493(2000)128<3268:DAITPO>2.0.CO;2.

Deng, Z., Y. Tang, and G. Wang (2010), Assimilation of Argo temperature and salinity profiles using a bias-aware localized EnKF system for the Pacific Ocean, *Ocean Modell.*, *35*, 187–205, doi:10.1016/j.ocemod.2010.07.007.

Dibarboure, G., O. Lauret, F. Mertz, V. Rosmorduc, and C. Maheu (2009), SSALTO/DUACS user handbook: (M)SLA and (M)ADT near-real time and delayed time products, *Rep. CLS-DOS-NT-06.034*, Cent. Natl. d'Etud. Spaciales, Toulouse, France.

Drécourt, J.-P., H. Madson, and D. Rosbjerg (2006), Bias aware Kalman filters: Comparison and improvements, *Adv. Water Resour.*, *29*, 707–718, doi:10.1016/j.advwatres.2005.07.006.

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

Ferry, N., E. Rémy, P. Brasseur, and C. Maes (2007), The Mercator global ocean operational analysis system: Assessment and validation of an 11-year reanalysis, *J. Mar. Syst.*, *65*, 540–560, doi:10.1016/j.jmarsys.2005.08.004.

Frank, P., and J. R. Colby (1997), A preliminary investigation of temperature errors in operational forecasting models, *Weather Forecasting*, *13*, 187–205.

Gaspari, G., and S. E. Cohn (1999), Construction of correlation functions in two and three dimensions, *Q. J. R. Meteorol. Soc.*, *125*, 723–757, doi:10.1002/qj.49712555417.

Hamill, T. M. (2006), Ensemble-based atmospheric data assimilation, in *Predictability of Weather and Climate*, edited by T. Palmer and R. Hagedorn, chap. 6, pp. 124–156, Cambridge Univ. Press, Cambridge, U. K.

Hamill, T. M., and J. S. Whitaker (2005), Accounting for the error due to unresolved scales in ensemble data assimilation: A comparison of different approaches, *Mon. Weather Rev.*, *133*, 3132–3147, doi:10.1175/MWR3020.1.

Helber, R. W., R. H. Weisberg, F. Bonjean, E. S. Johnson, and G. S. E. Lagerloef (2007), Satellite-derived surface current divergence in relation to tropical Atlantic SST and wind, *J. Phys. Oceanogr.*, *37*, 1357–1375, doi:10.1175/JPO3052.1.

Houtekamer, P. L., and H. L. Mitchell (2001), A sequential ensemble Kalman filter for atmospheric data assimilation, *Mon. Weather Rev.*, *129*, 123–137, doi:10.1175/1520-0493(2001)129<0123:ASEKFF>2.0.CO;2.

Houtekamer, P. L., and H. L. Mitchell (2005), Ensemble Kalman filtering, *Q. J. R. Meteorol. Soc.*, *131*, 3269–3289, doi:10.1256/qj.05.135.

Huang, B., Y. Xue, and D. Behringer (2008), Impacts of Argo salinity in NCEP Global Ocean Data Assimilation System: The tropical Indian Ocean, *J. Geophys. Res.*, *113*, C08002, doi:10.1029/2007JC004388.

Hunt, B. R., E. J. Kostelich, and I. Szunyogh (2007), Efficient data assimilation for spatiotemporal chaos: A local ensemble transform Kalman filter, *Phys. D*, *230*, 112–126, doi:10.1016/j.physd.2006.11.008.

Kalnay, E. (2002), *Atmospheric Modeling, Data Assimilation and Predictability*, 341 pp., Cambridge Univ. Press, New York.

Keppenne, C. L., M. M. Rienecher, N. P. Kurkowski, and D. A. Adamec (2005), Ensemble Kalman filter assimilation of temperature and altimeter data with bias correction and application to seasonal prediction, *Nonlinear Processes Geophys.*, *12*, 491–503, doi:10.5194/npg-12-491-2005.

Keppenne, C. L., M. M. Rienecher, J. P. Jacob, and R. Kovach (2008), Error covariance modeling in the GMAO ocean ensemble Kalman filter, *Mon. Weather Rev.*, *136*, 2964–2982, doi:10.1175/2007MWR2243.1.

Levitus, S., and T. Boyer (1998), NODC (Levitus) World Ocean Atlas 1998, http://www.esrl.noaa.gov/psd/data/gridded/data.nodc.woa98.html, Phys. Sci. Div., Earth Syst. Res. Lab., Boulder, Colo.

Li, H., E. Kalnay, and T. Miyoshi (2009), Simultaneous estimation of covariance inflation and observation errors within an ensemble Kalman filter, *Q. J. R. Meteorol. Soc.*, *135*, 523–533, doi:10.1002/qj.371.

Madec, G. (2008), NEMO ocean engine, *Note Pôle Modél. 27*, Inst. Pierre-Simon Laplace, Paris.

Martin, M. J., M. J. Bell, and N. K. Nichols (2002), Estimation of systematic error in an equatorial ocean model using data assimilation, *Int. J. Numer. Methods Fluids*, *40*, 435–444, doi:10.1002/fld.298.

McPhaden, M. J. (1995), The Tropical Atmosphere Ocean array is completed, *Bull. Am. Meteorol. Soc.*, *76*, 739–741.

Mitchell, H. L., and P. L. Houtekamer (2000), An adaptive ensemble Kalman filter, *Mon. Weather Rev.*, *128*, 416–433, doi:10.1175/1520-0493(2000)128<0416:AAEKF>2.0.CO;2.

Moore, A. M., 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. Clim.*, *19*, 4683–4699, doi:10.1175/JCLI3870.1.

Nichols, N. K. (2003), Treating model error in 3-D and 4-D data assimilation, in *Data Assimilation for the Earth System*, edited by R. Swinbank, V. Shutyaev, and W. A. Lahoz, pp. 127–135, Kluwer Acad., Dordrecht, Netherlands.

Oke, P. R., A. Schiller, D. A. Griffin, and G. B. Brassington (2005), Ensemble data assimilation for an eddy-resolving ocean model of the Australian region, *Q. J. R. Meteorol. Soc.*, *131*, 3301–3311, doi:10.1256/qj.05.95.

Oke, P. R., G. B. Brassington, D. A. Griffin, and A. Schiller (2008), The Bluelink ocean data assimilation system (BODAS), *Ocean Modell.*, *21*, 46–70, doi:10.1016/j.ocemod.2007.11.002.

Ott, E., B. H. Hunt, I. Szunyogh, Z. V. Zimin, E. J. Kostelich, M. Corazza, E. Kalnay, D. J. Patil, and J. A. Yorke (2004), A local ensemble Kalman filter for atmospheric data assimilation, *Tellus*, *56A*, 415–428.

Smith, G. C., and K. Haines (2009), Evaluation of the S(T) assimilation method with Argo dataset, *Q. J. R. Meteorol. Soc.*, *135*, 739–756, doi:10.1002/qj.395.

Szunyogh, I., E. J. Kostelich, G. Gyarmati, E. Kalnay, B. R. Hunt, E. Ott, E. Satterfield, and J. A. Yorke (2008), A local ensemble transform Kalman filter data assimilation system for the NCEP global model, *Tellus*, *60A*, 113–130.

Tang, Y., R. Kleeman, and A. M. Moore (2004), SST assimilation experiments in a tropical Pacific Ocean model, *J. Phys. Oceanogr.*, *34*, 623–642, doi:10.1175/3518.1.

Thacker, W. C., and O. E. Esenkov (2002), Assimilating XBT data into HYCOM, *J. Atmos. Oceanic Technol.*, *19*, 709–724.

Thiébaux, H. J., and L. L. Morone (1990), Short-term systematic errors in global forecasts: Their estimation and removal, *Tellus*, *42A*, 209–229.

Thompson, K. R., D. G. Wright, Y. Lu, and E. Demirov (2006), A simple method for reducing seasonal bias and drift in eddy resolving ocean models, *Ocean Modell.*, *13*, 109–125, doi:10.1016/j.ocemod.2005.11.003.

Whitaker, J. S., T. M. Hamill, X. Wei, Y. Song, and Z. Toth (2008), Ensemble data assimilation with the NCEP global forecast system, *Mon. Weather Rev.*, *136*, 463–482, doi:10.1175/2007MWR2018.1.

Zupanski, D., and M. Zupanski (2006), Model error estimation employing ensemble data assimilation approach, *Mon. Weather Rev.*, *134*, 1337–1354, doi:10.1175/MWR3125.1.

Z. Deng and Y. Tang, Environmental Science and Engineering, University of Northern British Columbia, 3333 University Way, Prince George, BC V2N 4Z9, Canada. (ytang@unbc.ca)

H. J. Freeland, Institute of Ocean Sciences, Fisheries and Oceans Canada, PO Box 6000, Sidney, BC V8L 4B2, Canada.