

# A “Dressed” Ensemble Kalman Filter Using the Hybrid Coordinate Ocean Model in the Pacific

WAN Liying<sup>\*1</sup> (万莉颖), ZHU Jiang<sup>2</sup> (朱江), WANG Hui<sup>1</sup> (王辉),  
YAN Changxiang<sup>2</sup> (闫长香), and Laurent BERTINO<sup>3</sup>

<sup>1</sup>*National Marine Environmental Forecasting Center, Beijing 100086*

<sup>2</sup>*Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing 100029*

<sup>3</sup>*Mohn-Sverdrup Center/Nansen Environmental and Remote Sensing Center, Bergen N-5006, Norway*

(Received 4 April 2008; revised 28 October 2008)

## ABSTRACT

The computational cost required by the Ensemble Kalman Filter (EnKF) is much larger than that of some simpler assimilation schemes, such as Optimal Interpolation (OI) or three-dimension variational assimilation (3DVAR). Ensemble optimal interpolation (EnOI), a crudely simplified implementation of EnKF, is sometimes used as a substitute in some oceanic applications and requires much less computational time than EnKF. In this paper, to compromise between computational cost and dynamic covariance, we use the idea of “dressing” a small size dynamical ensemble with a larger number of static ensembles in order to form an approximate dynamic covariance. The term “dressing” means that a dynamical ensemble seed from model runs is perturbed by adding the anomalies of some static ensembles. This dressing EnKF (DrEnKF for short) scheme is tested in assimilation of real altimetry data in the Pacific using the HYbrid Coordinate Ocean Model (HYCOM) over a four-year period. Ten dynamical ensemble seeds are each dressed by 10 static ensemble members selected from a 100-member static ensemble. Results are compared to two EnKF assimilation runs that use 10 and 100 dynamical ensemble members. Both temperature and salinity fields from the DrEnKF and the EnKF are compared to observations from Argo floats and an OI SST dataset. The results show that the DrEnKF and the 100-member EnKF yield similar root mean square errors (RMSE) at every model level. Error covariance matrices from the DrEnKF and the 100-member EnKF are also compared and show good agreement.

**Key words:** Dressing Ensemble Kalman Filter (DrEnKF), HYbrid Coordinate Ocean Model, root mean square errors

**Citation:** Wan, L. Y., J. Zhu, H. Wang, C. X. Yan, and L. Bertino, 2009: A “dressed” ensemble Kalman filter using the HYbrid Coordinate Ocean Model in the Pacific. *Adv. Atmos. Sci.*, **26**(5), 1042–1052, doi:10.1007/s00376-009-7208-6.

## 1. Introduction

Ensemble-based data assimilation uses Monte Carlo model simulations in order to derive flow-dependent forecast statistics and is applied, for instance, in operational weather prediction by the use of an ensemble forecast. Since the first introduction of the ensemble Kalman filter in an ocean model by Evensen (1994), the method has gained considerable attention and has also been successfully applied in meteorology and other research fields (Hamill et

al., 2001; Houtekamer and Mitchell, 2001; Keppenne and Rienecker, 2002, 2003; Natvik and Evensen, 2003; Leeuwenburgh, 2005).

Previous studies showed that the ensemble size is one of the main factors influencing the quality of covariance estimates. Natvik and Evensen (2003) emphasized that too few members, e.g., less than 60–80, degraded the performance of the EnKF. Natvik and Evensen (2003) tried various ensemble sizes from 20 to 100 in a high-dimensional non-linear system. An ensemble of 100 members was sufficient in their exper-

---

\*Corresponding author: WAN Liying, wanly@nmefc.gov.cn

iments and in other applications, e.g., Lisæter et al. (2003). However, running 100 members multiplies the computational costs by a factor of 100 with respect to that of a single model run, which prevents use in applications that are large-scale and/or high-resolution. To reduce computational costs, we consider the reduction of the number of ensemble members based on a hybridization of the Optimal Interpolation (OI) and Ensemble Kalman Filter (EnKF) approaches.

Evensen (2003) proposed a sub-optimal EnKF approach similar to OI and called this Ensemble Optimal Interpolation (EnOI). EnOI and EnKF experiments have been discussed in Wan (2006), which agree with the findings of the twin experiment by Leeuwenburgh (2005) and the results of Durand et al. (2003) for the ability of satellite data assimilation to correct ocean fields below the mixed layer.

A dynamical ensemble quantifies the model uncertainty but is inevitably limited by a finite ensemble size, by the approximate definition of model errors (often limited to errors in the model forcing) and the initial state errors. The ensemble is thus confined into a subspace of small dimensions and subject to bias. One way to compensate for bias is to add perturbations from a subspace of larger dimension to each dynamical ensemble member during the model integration. In the context of ensemble forecasting, Roulston and Smith (2003) bring forward the so-called “best member method”, a method for producing ensemble by “dressing” a single forecast with historical error statistics. The skills of the resulting hybrid ensembles in probabilistic forecasts are superior to those of ensembles generated statistically around the best guess forecast. The best member method is applicable to all types of ensembles, and in particular to ensemble-based data assimilation techniques.

In this work, we make an initial attempt to apply the “dressing” idea outside of the numerical weather ensemble forecasting context, to test its performance in assimilating realistic data in the Pacific Ocean using a HYbrid Coordinate Ocean Model (HYCOM). In the dressing approach following Roulston and Smith (2003) and Wang and Bishop (2005), the dressing is done by a dynamical ensemble of the model state and a static ensemble of observations of the atmosphere. Due to the lack of comprehensive observations in the ocean (mostly temperature sections before 1992, sea surface height anomalies after 1992, and temperature and salinity profiles after 2000), we rather use a collection of states from a long model integration as the static ensemble.

Our main motivation is to explore the feasibility, efficiency and reliability of such an ocean data assimilation scheme. The Dressed EnKF (DrEnKF for short)

experiment is compared to two standard EnKF experiments that use 10 and 100 members, and to the free running model. The paper is arranged as follows. In section 2, the DrEnKF method is described. Section 3 gives a brief introduction to the HYCOM model and the experimental design. In section 4, we discuss the experimental results. Section 5 gives the conclusions.

## 2. Dressing the ensemble Kalman filter

### 2.1 Dressed ensemble generation

Dynamical ensembles are constructed using a forward model integration to quantify the state dependence of the prediction errors. The atmosphere is so chaotic that it is often sufficient to add small perturbations to the observations assimilated to ensure divergence of the ensemble forecast. In ocean physics, the temporal scales are much longer and we rather assume that inaccuracy in an initial snapshot can be attributed to a displacement of the isopycnal vertical coordinates. We sample the initial ensemble by perturbing the layer interfaces from a model snapshot in equilibrium and the initial ensemble is then integrated forward with model errors in the atmospheric forcing fields to build up the multivariate covariances. The method for generating perturbations is reported in Evensen (2003) and Natvik and Evensen (2003).

The seed dynamical ensemble state vectors are stored in a matrix, also known as the “father” ensemble matrix, with each member denoted as  $\mathbf{x}_i \in \mathfrak{X}$ ,

$$\mathbf{F} = (x_1, x_2, \dots, x_M) \in \mathfrak{X}^{n \times M}, \quad (1)$$

Here,  $M$  is the size of the dynamical ensemble and  $n$  is the model state dimension. We write the ensemble mean and ensemble anomalies as

$$F_j = \bar{F} + F'_j, \quad j = 1, \dots, M. \quad (2)$$

The static or stationary ensemble is sampled from snapshots of a long model run. The static ensemble matrix is defined as

$$\mathbf{S} = (\psi_1, \psi_2, \dots, \psi_N) \in \mathfrak{X}^{n \times N}, \quad (3)$$

where  $N$  is the number of model snapshots of the model states.  $N$  is an integer multiple of  $M$  (i.e.,  $N = pM$ ,  $p > 1$ ). The mean and anomaly of the static ensemble are given as

$$S_i = \bar{S} + S'_i, \quad j = 1, \dots, N. \quad (4)$$

Then we can define the dressed ensemble by adding static anomalies to a dynamical ensemble member:

$$\mathbf{A} = [\mathbf{F}', \mathbf{S}' + \mathbf{F}'^*] \in \mathfrak{X}^{n \times (N+M)}. \quad (5)$$

Here,  $\mathbf{F}'^* \in \mathfrak{X}^{n \times N}$  is the result of  $p$  concatenations of the matrix  $\mathbf{F}'$ , with same total size as  $\mathbf{S}'$ . The ensemble  $\mathbf{S}' + \mathbf{F}'^*$  is called the “daughter” ensemble here.

The relative weights of the “father” and “daughter” ensembles are therefore equivalent in the resulting covariance matrix. This has been chosen arbitrarily to ensure their respective variance of Sea Level Anomaly (SLA) is similar.

## 2.2 Forecast covariance and analysis scheme

The covariance matrix  $\mathbf{P}_e \in \mathfrak{R}^{n \times n}$  is calculated using the  $M + N$  ensembles as:

$$\begin{aligned} \mathbf{P}'_e &= \frac{\mathbf{A}'(\mathbf{A})^T}{M + N - 1} \\ &= \frac{\mathbf{F}'\mathbf{F}'^T + (\mathbf{S}' + \mathbf{F}'^*)(\mathbf{S}' + \mathbf{F}'^*)^T}{M + N - 1} \\ &= \frac{N + M}{(M + N - 1)M} \mathbf{F}'\mathbf{F}'^T + \frac{1}{M + N - 1} \mathbf{S}'\mathbf{S}'^T \\ &\quad + \frac{1}{M + N + 1} (\mathbf{F}'^* \mathbf{S}'^T + \mathbf{S}'^* \mathbf{F}'^T). \end{aligned} \quad (6)$$

If the dynamic and static anomalies  $\mathbf{F}$  and  $\mathbf{S}$  are uncorrelated, we obtain

$$\mathbf{P}_e = \left(1 - \frac{N}{M(M + N - 1)}\right) \mathbf{P}_e^F + \frac{N - 1}{M + N - 1} \mathbf{P}_e^S. \quad (7)$$

Here,  $\mathbf{P}_e^F$  and  $\mathbf{P}_e^S$  are, respectively, the ensemble covariance matrices for the “father” and “daughter” ensembles. We expect, therefore, that the DrEnKF has a larger spread than the small size dynamical ensemble, since a small ensemble tends to underestimate the variance. The comparison with the large size EnKF is not a priori. The impacts of cross-terms on the multivariate covariances are not obvious.

The observations are perturbed to form an ensemble of observations, with specified observation error statistics. There are two goals: one is to to the standard EnKF ensemble update algorithm (Burgers et al., 1998), the second is to consider the storage in square-root form of a large observation error covariance matrix  $\mathbf{R}$  (of size with  $m \times m$ , the number of observations  $m$  being of the order of 200 000 in our case). We note two matrices of perturbed observations. The first one has the dimension of the “father” ensemble

$$\mathbf{D}^F = (d_1, d_2, \dots, d_M) \in \mathfrak{R}^{m \times M}. \quad (8)$$

The second matrix is the ensemble representation of  $\mathbf{R}$

$$\boldsymbol{\gamma} = (\varepsilon_1, \varepsilon_2, \dots, \varepsilon_{M+N}) \in \mathfrak{R}^{m \times (M+N)}. \quad (9)$$

The size of this ensemble is arbitrary, and we have chosen the same size as the “daughter” ensemble for convenience. The approximate observation error covariance matrix is

$$\mathbf{R}_e = \frac{\boldsymbol{\gamma}\boldsymbol{\gamma}^T}{M + N - 1}. \quad (10)$$

We use the standard analysis equation to update the “father” ensemble only:

$$\begin{aligned} F_j^a &= F_j^f + \mathbf{A}'\mathbf{A}'^T H^T (H\mathbf{A}'\mathbf{A}'^T H^T + \boldsymbol{\gamma}\boldsymbol{\gamma}^T)^{-1} \\ &\quad (D_F - HF_j), \quad j = 1, 2, \dots, M. \end{aligned} \quad (11)$$

In the DrEnKF scheme, only the “father” ensemble  $\mathbf{F}$  is updated. The static ensemble  $\mathbf{S}$  is only used to calculate the forecast covariance and is not updated. Here,  $H$  is the observation operator. The Kalman gain is defined as

$$K_{\text{DrEnKF}} = \mathbf{A}'\mathbf{A}'^T H^T (H\mathbf{A}'\mathbf{A}'^T H^T + \boldsymbol{\gamma}\boldsymbol{\gamma}^T)^{-1}. \quad (12)$$

The detailed algorithm of the Kalman gain is following the formulation referred as “analysis2” in Evensen (2004). The perturbed observations introduce sampling errors, particularly for small ensembles (Whitaker and Hamill, 2002). There are number of square-root implementations proposed (Anderson, 2001; Evensen, 2004; Sakov and Oke, 2008). In this study, we do not experience ensemble collapse even though the “father” ensemble has very small size ( $N = 10$ ). The use of a static ensemble has possibly prevented the ensemble from collapse. The implementation of a square root scheme would have been similarly simple and should be evaluated in future work.

## 2.3 Computational cost

The DrEnKF is implemented here in a way that saves considerable computational cost. The size of the “daughter” ensemble  $N$  is similar to usual EnKF applications, but the size of the “father” ensemble  $M$  is smaller, and only the latter is integrated into the model. The propagation is thus cheaper in the present implementation of the DrEnKF. Using the analysis scheme from Burgers et al. (1998) the computational cost of the analysis is approximately  $nNN$  for an  $N$ -member EnKF approach. The update cost of DrEnKF ( $M$  “father” ensemble members,  $N$  “daughter” ensemble members) is  $nM(M + N)$ . If  $M$  is much smaller than  $N$ , the update cost of DrEnKF is also reduced. Hence, from both the model integration and update aspects, DrEnKF can save computational cost dramatically if we can keep the size of the “father” ensemble small.

## 3. Experimental design and preparation

### 3.1 Model

The data assimilation experiments use a Pacific Ocean implementation of HYCOM (Bleck, 2002), a primitive equation, general circulation model with hybrid vertical coordinate.

The model domain covers the area from 28°S–52°N, 95°E–70°W. The grid is generated by conformal mapping (Bentsen et al., 1999), and has horizontal resolution between 42 km and 72 km (roughly half a degree). The present implementation uses 22 hybrid isopycnal and z-level coordinates, the 10 top layers using a linear increase of density and the last 12 layers an exponential increase. The range of the potential densities ( $\sigma$  is the potential density in each layer) varies from 18.00 kg m<sup>-3</sup> to 27.84 kg m<sup>3</sup>. The model prognostic variables include 3-dimensional temperature, salinity, layer thickness, and velocity, as well as the barotropic velocity and pressure fields. All prognostic variables are updated within the assimilation process.

The model is forced by atmospheric forcing fields from the European Center for Medium-Range Weather Forecasting (ECMWF) reanalysis. The high-frequency synoptic forcings used were temperature, winds, and relative humidity determined from dew-point temperatures. Clouds and precipitations are based on the climatology of the Comprehensive Ocean and Atmosphere Dataset (COADS) and precipitation from Legates et al. (1990). “Closed” boundaries in the north and south have their salinities and temperatures relaxed to the Generalized Digital Environmental Model (GDEM) climatology (Teague et al., 1990), with a common relaxation time scale of 100 days over 15 grids cells. River input is modeled as a negative salinity flux, and there are 13 river sources used in the Pacific simulation.

### 3.2 Experiment design

We carry out three data assimilation experiments using the DrEnKF, the 10-member EnKF, and the 100-member EnKF. Both schemes assimilate sea surface height (SSH) over the Pacific Ocean.

The static ensemble of model snapshots is based on a control run from 1 January 1981 to 31 December 2001, of size  $N = 100$ . The seasonality has been removed from the static ensemble. The dynamical ensemble is initialized following the optimal sample of Wan et al. (2008), with 10 and 100 members for EnKF experiments and 10 for the DrEnKF. We randomly select 10 members of the static ensemble anomalies and add them to each dynamical member.

We simulate the model errors affecting the evolution of the ensemble members by generating random fields added to the ECMWF atmospheric forcing fields as outlined in Evensen (2003). The spatial correlation depends on the given spatial scale (also called decorrelation scale) and typical grid size, while the temporal correlation is formed by a Markov process with a sequence of time-correlated pseudorandom fields with

mean equal to zero and variance equal to 1. The set up of these errors follows Wan (2006) and detailed parameters are given in Table 1.

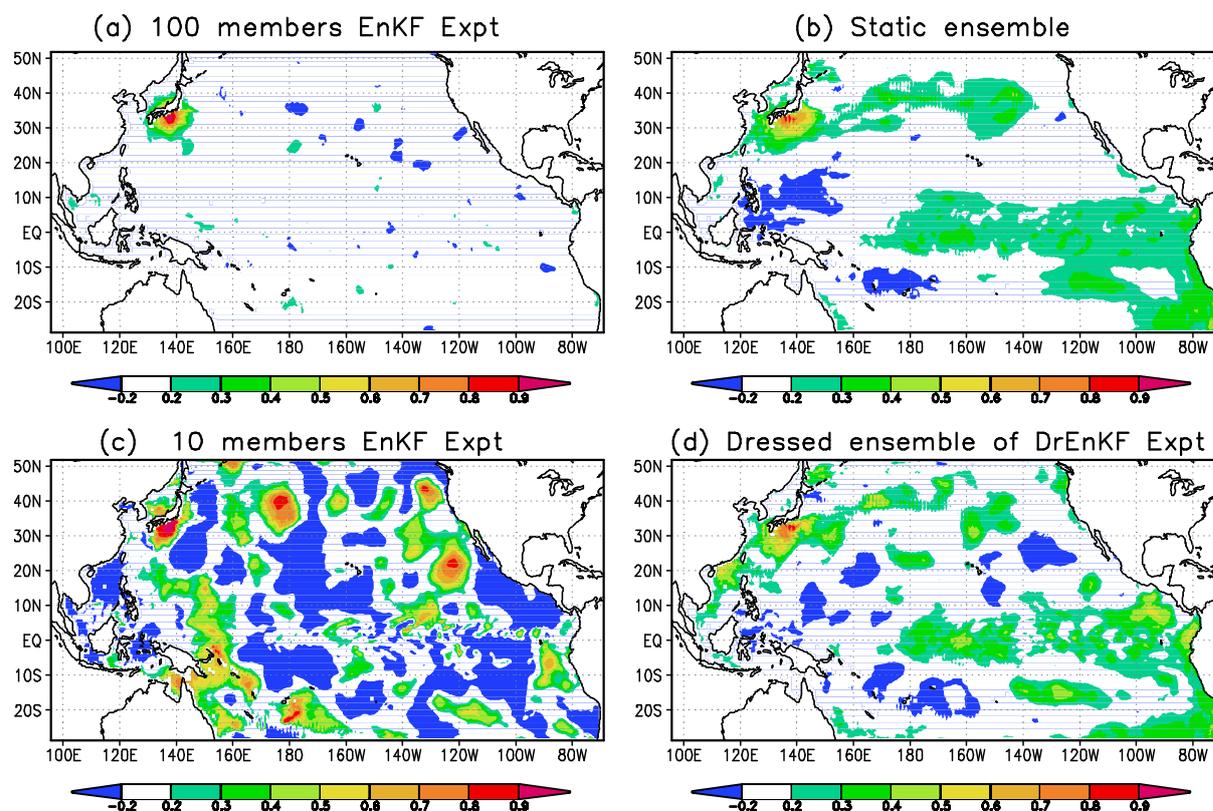
Figure 1 shows the correlation coefficients of SST relative to the grid point (33°N, 137°E) in the main axis of the Kuroshio Current sampled from the dynamical ensemble created by the EnKF experiment (100 members), the EnKF experiment (10 members), static ensemble (100 members), and dressed ensemble (110 members) created by the DrEnKF experiment, respectively. There are very weak long range correlations from the EnKF experiment using 100 members. The static ensemble generates more long range correlations than the EnKF. For the 10-member EnKF, because of the small size of its ensemble, high pseudo correlations exist in many regions. Once a dressed ensemble was created, many pseudo correlations disappeared, and the magnitude of pseudo correlation becomes smaller. However for realistic applications such as this one, the localization can help to remove remaining long range correlations. More discussions of localization and filtering of long range correlations are reported by Hamill et al. (2001), Mitchell et al. (2002) and Oke et al. (2006).

It is necessary to use localization to avoid the problem caused by obvious long range correlations. It uses only observations within certain circle around a grid to update the model state at a given grid point. We use a local analysis scheme that makes analysis of a grid point using a maximum of 49 observations within 700 km distance, and this has been applied in both of the two experiments. The time window of the two experiments is the same, i.e., from 1 January 2002 to 31 December 2005.

### 3.3 Data

Observational data used in the assimilation experiments are SSH satellite data sets merged from different satellites, e.g., TOPEX/Poseidon, Jason-1, ENVISAT, etc. The spatial resolution of the data sets is (1/3)° and the temporal resolution is one week. The data sets are provided by Collecte Localisation Satellites (CLS).

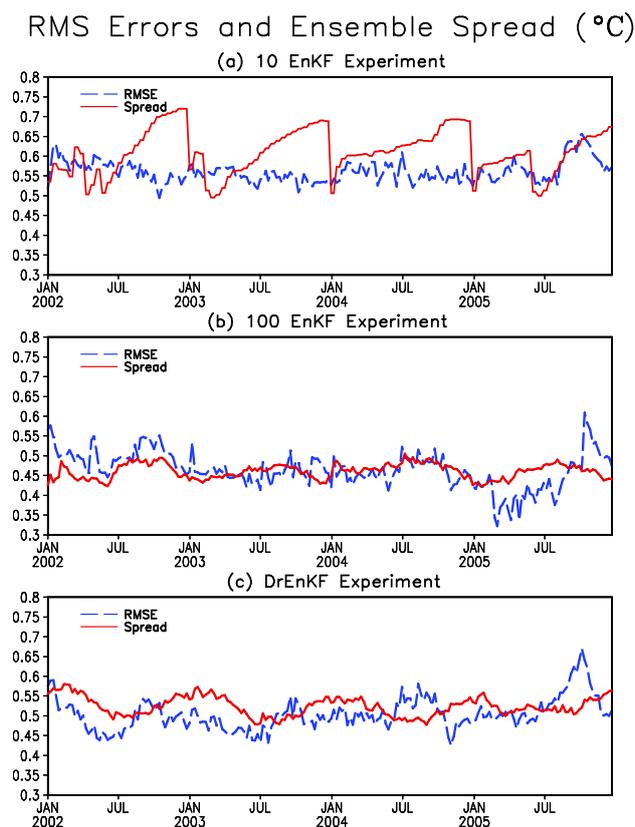
The temperature and salinity profiles from Argo floats and Optimum Interpolation Sea Surface Temperature (OISST) (Reynolds et al., 2002) are used to validate the performances of the two experiments. The OISST data set is produced at the National Oceanic and Atmospheric Administration (NOAA) using both *in situ* and satellite data, which is weekly on a one-degree grid. The Argo floats carry out continuous monitoring of the temperature and salinity from the sea surface down to 1000 m or 2000 m Carval et al. (2006). We used all available (17354) temperature/salinity profiles in delayed mode within the model



**Fig. 1.** Sea surface temperature correlation with point ( $33^{\circ}\text{N}$ ,  $137^{\circ}\text{E}$ ) on 18 August 2004. (a) 100 members EnKF experiment, (b) static ensemble of DrEnKF, (c) 10-member EnKF experiment, (d) dressed ensemble of DrEnKF.

**Table 1.** Parameters in assimilation experiments.

Parameters	Description	Value
$N$	Number of static members in DrEnKF Experiment; Number of dynamical members in EnKF Experiment	100 100
$M$	Number of dynamical members in DrEnKF Experiment	10
$m$	Maximum number of local observations	49
$r_0$	Observations radius of influence	700 km
$\sigma_e$	Observations error standard deviation	0.05 km
$r_e$	Observations error decorrelation length	10 km
$r_h$	Initial error decorrelation length	1000 km
$\sigma_d$	Layer thickness logarithmic standard deviation	10%
$\sigma_T$	Temperature logarithmic standard deviation in mixed layer	5%
$r_{vd}$	Vertical correlation coefficient of layer thickness	2.0
$r_{vT}$	Vertical correlation coefficient of temperature	3.0
$r_{\beta}$	Decorrelation length of random forcing	10 grid cells
$\sigma_T$	Atmospheric temperature standard deviation	3 K
$\sigma_{\tau}$	Atmospheric wind stress standard deviation	$0.03 \text{ N m}^{-2}$
$\sigma_w$	Wind speed standard deviation	$1.6 \text{ m s}^{-1}$
$\sigma_r$	Radiative flux standard deviation	$0.2 \text{ W m}^{-2}$
$r_t$	Time correlation coefficient parameter of random forcing	3 days



**Fig. 2.** RMS errors against OISST and the ensemble spread (units:  $^{\circ}\text{C}$ ). (a) 10-member EnKF experiment, (b) 100-member EnKF experiment, (c) DrEnKF experiment. Dashed lines are RMS errors against OISST and solid lines are the ensemble spread.

domain and the assimilation period. The Argo data is downloaded from the Argo Data Center. (<http://www.coriolis.eu.org/cdc/dataSelection/cdcDataSelections.asp>)

## 4. Results

### 4.1 Overall performance

In this part, we first compare the EnKF error estimate against OISST to monitor the evolution of the ensemble with the Root Mean Square Errors (RMSE). Figure 2 displays them over the four-year experiment. For both experiments, the ensembles are time-dependent and the spread is fluctuating. The curves show no visible trend over four years—neither growing nor shrinking. The three time series of EnKF and DrEnKF show a faint seasonal cycle, the error being slightly larger during winter than summer. The spread of the 10-member EnKF looks obviously divergent, which leads to an overestimate of the error background. This cycle appears both in the error

estimate and in the RMS errors, indicating that the evolution of model errors is realistic. Since the heat fluxes are dependent on winds, it is also logical that the model errors are more efficient, and therefore the ensemble spread of DrEnKF is larger than that of the 100-member EnKF to avoid error underestimate by the small dynamical ensemble.

The overall performances of the 100-member EnKF and the DrEnKF are summarized in Table 2. Root mean square (RMS) errors over the model domain and experiment time window for the temperature and salinity fields from the EnKF and the DrEnKF at different model levels are calculated by comparing them to Argo observations, and the SST is compared to OISST. Argo floats are inhomogeneous in their spatial and temporal sampling. So, we interpolated the model results to the locations of Argo floats. The comparisons are only computed for the top 14 model layers because of insufficient observations in deep waters. While for OISST was originally mapped weekly on a one-degree grid, we interpolated its grid to the model grid. The RMS errors of the free running model (without assimilation) are also shown.

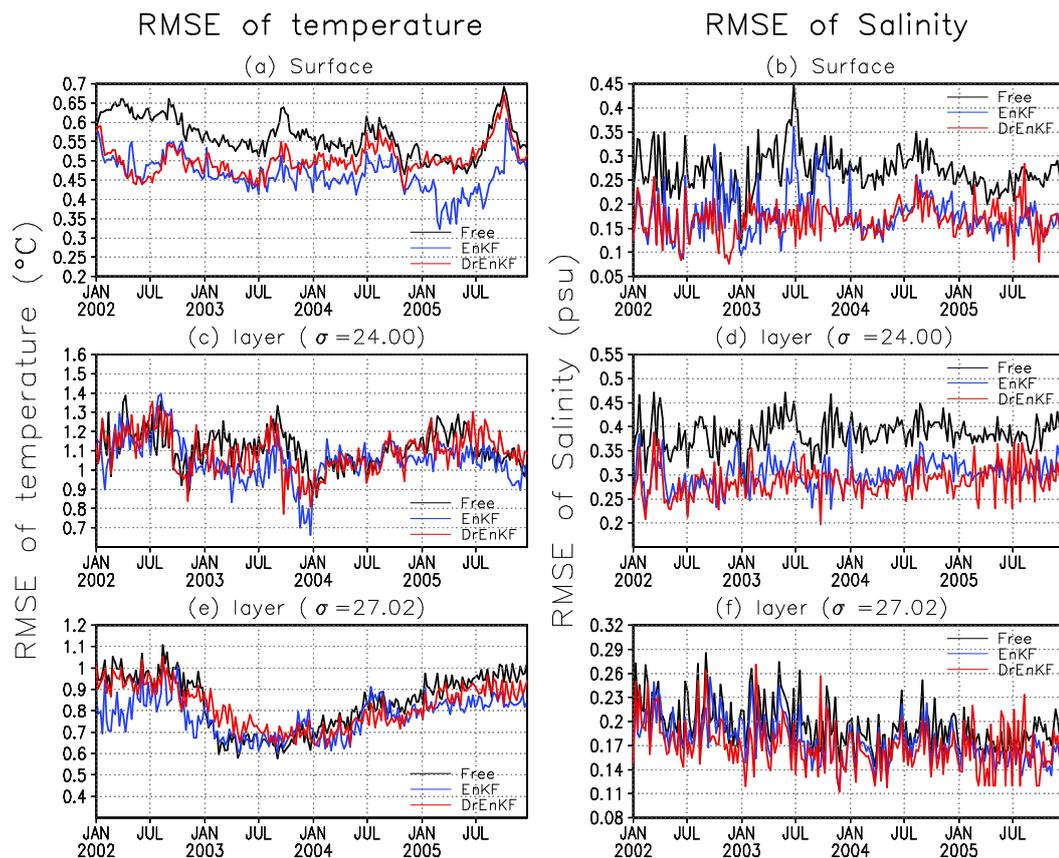
In the 100-member EnKF experiment, the RMS error of temperature is reduced by  $0.08^{\circ}\text{C}$  and that of salinity by  $0.091$  psu at the surface, while in the DrEnKF experiment, RMS error of salinity is reduced by  $0.06^{\circ}\text{C}$  and  $0.094$  psu at the surface. In the surface layer, the EnKF reduces the errors by 14.2% for temperature and 33.3% for salinity, while the reduction by the DrEnKF is 10.3% for temperature and 34.8% for salinity. In the mixed layer, a reduction of temperature and salinity errors is observed in the both experiments. Weaker but non-negligible reductions of RMS errors are observed in the thermocline for both experiments, indicating that none of the assimilation experiments have a strong impact in the thermocline, but still tend to improve the results. It is also comforting that the assimilation imbalances do not severely affect the stability of the water column and do interfere with deep water mass properties. The results of the DrEnKF are surprisingly close to those of the EnKF.

The evolution of the Root Mean Square Error (RMSE) over four years of the DrEnKF experiment and the 100-member EnKF experiment is shown in Fig. 3. We select three layers: the surface layer (top 3 meters in the model), the model layer with potential density  $24.00\text{ kg m}^{-3}$  (representative of the bottom of the mixed layer), and the layer with potential density  $27.02\text{ kg m}^{-3}$  (located below the thermocline). Figures 3a, c, and e show RMS errors of temperature in different layers, while Figs. 3b, d, and f are those of salinity. It is obvious that the RMS errors of temperature and salinity in the two assimilation experiments

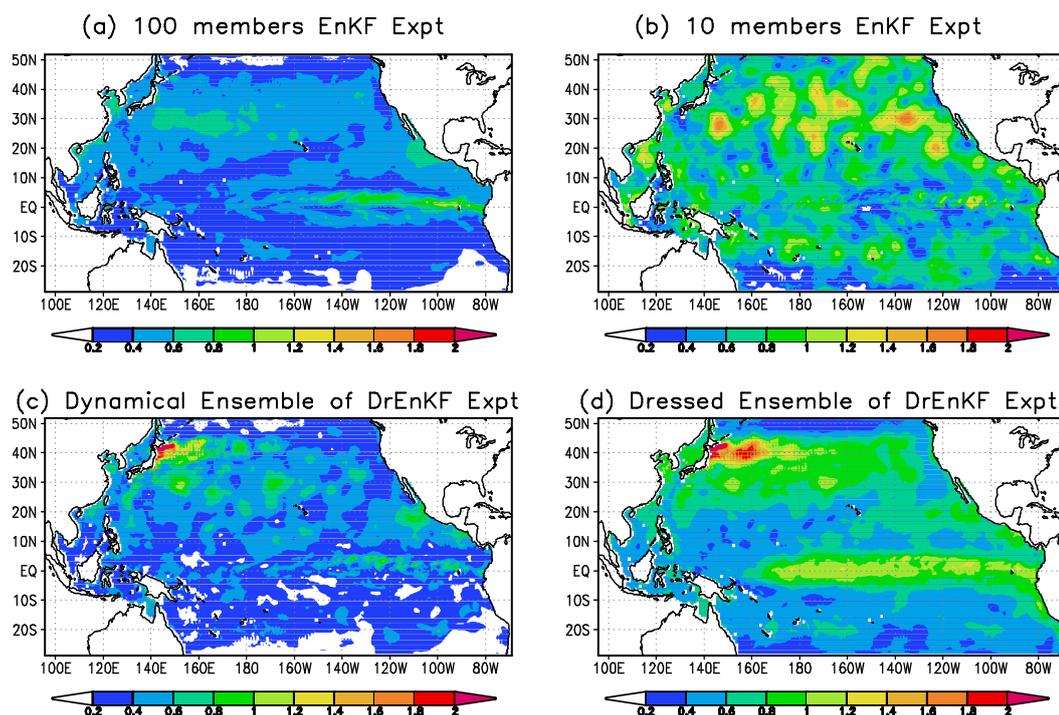
**Table 2.** RMS errors of temperature and salinity in different layers.

Layer	Temperature ( $^{\circ}\text{C}$ )			Salinity (psu)		
	Model	EnKF	DrEnKF	Model	EnKF	DrEnKF
1*	0.564	0.484	0.506	0.270	0.179	0.176
2	0.982	0.930	0.935	0.439	0.391	0.393
3	0.987	0.936	0.936	0.437	0.391	0.393
4	1.001	0.946	0.948	0.439	0.391	0.394
5	1.011	0.967	0.969	0.439	0.393	0.394
6	1.046	1.009	1.008	0.441	0.408	0.408
7	1.124	1.088	1.100	0.456	0.433	0.437
8	1.172	1.141	1.161	0.460	0.443	0.440
9	1.173	1.170	1.172	0.433	0.414	0.417
10	1.215	1.205	1.209	0.360	0.321	0.318
11	1.194	1.181	1.182	0.412	0.369	0.376
12	1.419	1.412	1.416	0.348	0.287	0.296
13	1.341	1.271	1.287	0.248	0.218	0.208
14	0.725	0.674	0.694	0.190	0.176	0.179

Note: \*—RMS errors of sea surface temperature are computed against OISST, while other RMS errors of temperature and salinity are calculated against Argo data.



**Fig. 3.** RMSE of temperature and salinity with time (unit:  $^{\circ}\text{C}$  and psu) (a) surface temperature, (b) surface salinity, (c) temperature in the layer ( $\sigma = 24.00 \text{ kg m}^{-3}$ ), (d) salinity in the layer ( $\sigma = 24.00 \text{ kg m}^{-3}$ ), (e) temperature in the layer ( $\sigma = 27.02 \text{ kg m}^{-3}$ ), (f) salinity in the layer ( $\sigma = 27.02 \text{ kg m}^{-3}$ ). Black lines are from the free-running experiment, blue lines are from DrEnKF experiment, and red lines are from EnKF experiment.



**Fig. 4.** Standard deviation of SST in EnKF experiment and DrEnKF experiment. The blank area in the South Pacific is below  $0.2^{\circ}\text{C}$  (data from 18 August 2004). (a) 100-member EnKF experiment, (b) 10-member EnKF experiment, (c) DrEnKF experiment (dynamical ensemble), (d) DrEnKF experiment (dressed ensemble).

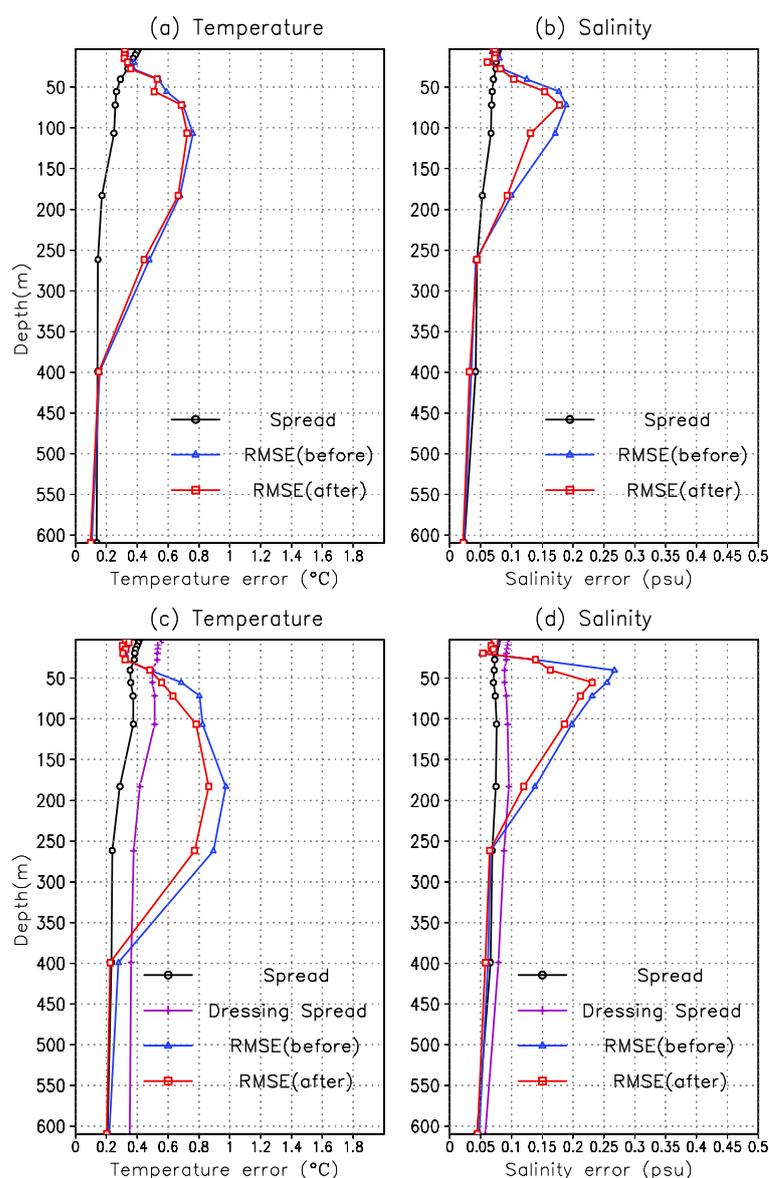
are smaller than those of the free running experiment in the three layers. In the surface and the intermediate layer ( $\sigma = 24.00 \text{ kg m}^{-3}$ ), the DrEnKF has RMS errors close to those of the EnKF. During a few periods, DrEnKF performs even better than EnKF. In the deep layer ( $\sigma = 27.02 \text{ kg m}^{-3}$ ), the EnKF and DrEnKF yield the similar RMS errors.

#### 4.2 Specific aspects of the DrEnKF

Figure 4 shows the standard deviation of surface temperature in different experiments. Figures 4a and b show the snapshot of the standard deviation of surface temperature in the 100-member and 10-member EnKF experiment on 18 August, 2004. Figures 4c and d show the snapshots of the standard deviation of surface temperature in the DrEnKF experiment before and after dressing (Fig. 4c is for the dynamical ensemble, Fig. 4d is for the dressing ensemble). In Fig. 4b, the large standard deviation located in more areas in the North Pacific, which brings an overestimate of the background errors in these areas. In Fig. 4c, there is large standard deviation located in Northwest Pacific and East Equatorial Pacific, while the results of the “father” ensemble in DrEnKF only show a large spread only in the Northwest Pacific. After dressing, the location of the maximum agrees better, especially along the East Equatorial Pacific. Moreover, the magnitude

of the ensemble spread is increased. For temperature and salinity in deep layers, we have also compared the standard deviation (figures not shown) of the two assimilation experiments, and found the DrEnKF can effectively account for residual errors and realize the time-dependence of error statistics. The DrEnKF reproduces the large scale correlation structures of SST. The assimilation of altimeter data improves the water mass properties only if the multivariate correlations are physically consistent. The validation we present in Fig. 4 has not been reported before to our knowledge.

Near the thermocline, Fig. 5 gives a snapshot of the comparison between ensemble spread and the difference between ensemble mean and observation. Because of the sparseness of Argo floats, we select a small region ( $10^{\circ}\text{S}$ – $20^{\circ}\text{N}$ ,  $128^{\circ}\text{E}$ – $160^{\circ}\text{W}$ , which is similar to the region3 in Wan, 2006) on a typical date (25 August 2004) to calculate the RMS errors before and after assimilation in the 100-member EnKF and DrEnKF experiments. Ideally, the ensemble spread should be comparable to the RMS errors. Since the model errors are only affecting the sea surface, the errors introduced initially in the interior ocean are difficult to sustain. Near the surface, the ensemble spread is comparable to the RMS errors, but between the bottom of the mixed layer and the thermocline, it is underestimated



**Fig. 5.** Comparison of the ensemble spread and RMS errors before and after assimilation (25 August 2004). (a) Temperature and (b) Salinity for EnKF experiment, (c) temperature and (d) salinity for Dressing EnKF experiment. The black lines are ensemble spread, the purple lines are ensemble spread after dressing, the blue lines are RMS errors against Argo data before assimilation, and the red lines are RMS errors against Argo data after assimilation.

for all methods. The spread is larger in the DrEnKF than in the EnKF through all the depths, mostly due to the addition of the spread of the static ensemble. Therefore, DrEnKF performs with results close to the EnKF and better than the EnKF sometime. However, treatment of model errors below the mixed layer is still a key issue to be focused on in subsequent model revisions.

## 5. Conclusions

Two data assimilation experiments (EnKF and DrEnKF) are compared and assessed against a free running model experiment. The DrEnKF is a new attempt stemming from state-of-the-art methodology in ensemble atmospheric forecasting, intended to save computing costs and add residual errors. Both methods assimilate real altimetry data into a HYCOM

model of the Pacific. The experiments have processed four years from 1 January 2002 to 31 December 2005. The results are compared and validated with independent observations from remote sensing OISST and Argo floats.

The idea of a “dressing” ensemble is the combination of a dynamical ensemble with a small size (here 10 members) and a static ensemble. The aim of the paper is to evaluate whether this idea could be a way to reduce the computational costs of the large number of members required for EnKF. We show that the DrEnKF is stable over an integration period of 4 years and reproduces the temporal variations of the error similarly to the EnKF. It is efficiently projecting the influence of the assimilated sea surface heights into the interior of the ocean, although with slightly less success than a large-membered EnKF. The DrEnKF avoids the underestimation of errors as expected with a dynamical ensemble of small size. The analysis of ensemble spread shows it is able to locate the horizontal patterns of the error variance as in a large-membered EnKF, and produces larger ensemble spread than the large-membered EnKF across the whole basin and at all observed depths of the ocean. The EnKF and DrEnKF, however, both underestimate the errors between the bottom of the mixed layer and the thermocline, which indicates that the addition of a static ensemble does not necessarily compensate for insufficient model accuracy, and neither does it compensate for possible model biases. Further research should therefore pursue improvement of model errors and robustness against model biases.

In the present work, the dynamical and static ensembles have approximately the same weight and we have not attempted to tune their respective contributions. There are, however, several ways to generate a dressing ensemble such that the ensemble spread matches closely some targeted error statistics. Wang and Bishop (2005) propose a refinement of the method based on seasonal statistics that could also be used to refine this study.

**Acknowledgements.** This study is supported by the Knowledge Innovation Program of Chinese Academy of Sciences (Grant No. KZCX1-YW-12-03), National Basic Research Program of China (2006CB403600), Project of Young Scientists Fund by National Natural Sciences Foundation of China (Grant No. 40606008) and National Science and Technology Infrastructure Program (2006BAC03B04). Hui Wang is supported by National Natural Sciences Foundation of China (Grant No. 40531006). Laurent Bertino is supported by a private donation from Trond Mohn c/o Frank Mohn AS, Bergen and the MERSEA project from the European Commission

(Grant No. SIP3-CT-2003-502885).

## REFERENCES

- Anderson, J. L., 2001: An ensemble adjustment Kalman filter for data assimilation. *Mon. Wea. Rev.*, **129**, 2884–2903.
- Bentsen, M., G. Evensen, H. Drange, and A.D. Jenkins, 1999: Coordinate transform on a sphere using conformal Mapping. *Mon. Wea. Rev.*, **127**, 2733–2740.
- Bleck, R., 2002: An oceanic general circulation model framed in hybrid isopycnic-cartesian coordinates. *Ocean Modelling*, **4**, 55–88.
- Burgers, G., P. J. van Leeuwen, and G. Evensen, 1998: Analysis scheme in the ensemble Kalman filter. *Mon. Wea. Rev.*, **126**, 1719–1724.
- Carval, T., and Coauthors, 2006: *ARGO Data User’s Manual, Version 2.1*. [Available online from [www.usgoda.org/argo/argo-dm-user-manual.pdf](http://www.usgoda.org/argo/argo-dm-user-manual.pdf)].
- Durand, F., L. Gourdeau, T. Delcroix, and J. Verron, 2003: Can we improve the representation of modeled ocean mixed layer by assimilating surface-only satellite-derived data? A case study for the tropical Pacific during the 1997–1998 El Niño. *J. Geophys. Res.*, **108**(C6), 3200, doi:10.1029/2002JC001603.
- 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.
- Evensen, G., 2003: The ensemble Kalman filter: Theoretical formulation and practical implementation. *Ocean Dynamic*, **53**, 343–367.
- Evensen, G., 2004: Sampling strategies and square root analysis schemes for the EnKF. *Ocean Dynamics*, **54**, 539–560.
- Hamill, T. M., J. S. Whitaker, and C. Snyder, 2001: Distance dependent filtering of background error covariance estimates in an ensemble Kalman filter. *Mon. Wea. Rev.*, **129**, 2776–2790.
- Houtekamer, P. L., and H. L. Mitchell, 2001: A sequential ensemble Kalman filter for Atmospheric data assimilation. *Mon. Wea. Rev.*, **129**, 123–137.
- Keppenne, C. L., and M. M. Rienecker, 2002: Initial testing of a massively parallel ensemble Kalman filter with the Poseidon isopycnal ocean general circulation model. *Mon. Wea. Rev.*, **130**, 2951–2965.
- Keppenne, C. L., and M. M. Rienecker, 2003: Assimilation of temperature into an isopycnal ocean general circulation model using a parallel ensemble Kalman filter. *Journal of Marine Systems*, **40–41**, 363–380.
- Leeuwenburgh, O., 2005: Assimilation of along-track altimeter data in the tropical Pacific region of a global OGCM ensemble. *Quart. J. Roy. Meteor. Soc.*, **131**, 2455–2472.
- Legates, D. R., and C. J. Willmott, 1990: Mean seasonal and spatial variability in gauge-corrected, global precipitation. *International Journal of Climatology*, **10**, 111–127.
- Lisæter, K. A., J. Rosanova, and G. Evensen, 2003: As-

- simulation of ice concentration in a coupled ice-ocean model, using the ensemble Kalman Filter. *Ocean Dynamics*, **53**, 368–388.
- Mitchell, H. L., P. L. Houtekamer, and G. Pellerin, 2002: Ensemble size, and model-error representation in an ensemble Kalman filter. *Mon. Wea. Rev.*, **130**, 2791–2808.
- Natvik, L. J., and G. Evensen, 2003: Assimilation of ocean colour data into a biochemical model of the North Atlantic: Part 1. Data assimilation experiments. *Journal of Marine Systems*, **40–41**, 127–153.
- Oke, P. R., P. Sakov, and S. P. Corney, 2006: Impacts of localisation in the EnKF and EnOI: Experiments with a small model. *Ocean Dynamics*, **57**, 32–45.
- Reynolds, R. W., and Coauthors, 2002: An improved *in situ* and satellite SST analysis for climate. *J. Climate*, **15**, 1609–1624.
- Roulston, M. S., and L. A. Smith, 2003: Combining dynamical and statistical ensembles. *Tellus*, **55A**, 16–30.
- Sakov, P., and P. R. Oke, 2008: A deterministic ensemble Kalman filter. *Tellus*, **60A**, 361–371.
- Teague, W. J., M. J. Carron, and P. J. Hogan, 1990: A comparison between the generalized digital environmental model and Levitus climatologies. *J. Geophys. Res.*, **95**, 7167–7183.
- Wan, L., 2006: Ensemble methods and applications to altimetry data assimilation in the Pacific. Ph. D. dissertation, Institute of Atmospheric Physics, Chinese Academy of Science, China, 117pp. (in Chinese)
- Wan, L., J. Zhu, B. Laurent, and H. Wang, 2008: Initial ensemble generation and validation for ocean data assimilation using HYCOM in the Pacific. *Ocean Dynamics*, **58**, 81–99.
- Wang, X. G., and C. H. Bishop, 2005: Improvement of Ensemble Reliability with a New Dressing Kernel. *Quart. J. Royal. Meteor. Soc.*, **131**, 965–986.
- Whitaker, J. S., and T. M. Hamill, 2002: Ensemble data assimilation without perturbed observations. *Mon. Wea. Rev.*, **130**, 1913–1924.