Bred vectors and forecast errors in the NASA coupled general circulation model

Shu-Chih Yang\textsuperscript{1,2,4}, Eugenia Kalnay\textsuperscript{2}, Ming Cai\textsuperscript{3} and Michele Rienecker\textsuperscript{4}

\textsuperscript{1} Earth System Science Interdisciplinary Center, University of Maryland, College Park, Maryland, 20742
\textsuperscript{2} Department of Atmospheric and Oceanic Science, University of Maryland, College Park, Maryland, 20742
\textsuperscript{3} Department of Meteorology, Florida State University, Tallahassee, Florida, 32306
\textsuperscript{4} Global Modeling Assimilation Office, NASA/ Goddard Space Flight Center, Greenbelt, Maryland, 20770

Corresponding author address: Dr. Shu-Chih Yang, Earth System Science Interdisciplinary Center/Department of Atmospheric and Oceanic Science. University of Maryland, College Park, MD, 20742.

Email: scyang@atmos.umd.edu
Abstract

The breeding method has been implemented in the NASA Global Modeling and Assimilation Office (GMAO) coupled general circulation model (CGCM) in its operational configuration where ocean data assimilation is used to initialize the coupled forecasts. Bred vectors (BV), designed to capture the dominant growing errors in the atmosphere-ocean coupled system, are applied as initial ensemble perturbations. We investigate the potential improvement for ensemble prediction by comparing BVs with the oceanic growing errors, estimated by the one-month forecast error from the non-perturbed forecast.

Our results show that one-month forecast errors and BVs from the NASA CGCM share very similar features: BVs are clearly related to forecast errors in both SST and equatorial subsurface temperature, particularly when the BV growth rate is large. Both the forecast errors and the BVs in the subsurface are dominated by large-scale structures near the thermocline. Our results suggest that the forecast errors are dominated by dynamically evolving structures related to the variations of the background anomalous state, and that their shapes can be captured by BVs, especially during the strong 1997-1998 El Niño. Hindcast experiments starting from January 1997 with one pair of BVs achieve a significant improvement compared to the control (unperturbed) hindcast by capturing many important features of this event, including the westerly wind burst in early 1997.
1. Introduction

During the last decade, the ability to simulate and predict El Niño-Southern Oscillation (ENSO) phenomena has greatly increased because of (1) the establishment of new types of observing systems contributing to a better understanding of the dominant physical process and improved initial conditions, and (2) the improvement of coupled ocean-atmosphere modeling (Hayes et al. 1991; Wallace et al. 1998; McPhaden et al., 1998, Saha et al. 2006). However, there are still limitations in advancing ENSO forecast skill to several seasons ahead (Latif et al. 1993, Chang et al. 1996; Fedorov et al. 2003; Chen et al. 2004, and the special volume of El Niño from JGR, 1998). These limitations arise from (1) errors in the initial conditions, (2) systematic errors/biases in the ocean-atmosphere coupled models, (3) problems with the initialization process, e.g. model shocks in the initial coupling, and (4) the physical processes in coupled ocean-atmosphere models used to represent ENSO, including feedbacks from variabilities of different time scales (e.g. Madden-Julian Oscillation, Slingo et al. 1999).

Ensemble forecasting and initialization using ocean data assimilation are, at the present, the most straightforward solutions for improving ENSO prediction (Stockdale et al. 1998; Kirtman et al. 2003; Chen et al. 1998; Ji et al. 1997; Rosati et al., 1997; Saha et al. 2006). Ensemble forecasting systems are designed to use a set of initial perturbations to represent the uncertainties and sensitivities of the state. The ensemble mean from a well-designed ensemble forecast system is expected to improve the control forecast (without perturbations) assuming the true state can be encompassed within a set of
ensemble evolutions (Toth and Kalnay, 1997). In addition, the ensemble spread can be used to assess the forecasting skill (Moore and Kleeman, 1998). However, at this stage, the ensemble forecasts implemented for seasonal to interannual prediction at operational centers still face the difficulty that the ensemble perturbations from a single coupled ocean-atmosphere model have limited growth at early forecast leads relative to the amplitude of mean error (Vialard et al. 2003; Palmer et al. 2004 and Saha et al. 2006). This indicates that current techniques to initialize ensemble members may not be optimal and not sensitive enough to low-frequency variations, such as the seasonal-to-interannual variability. Experiments from Moore and Kleeman (1996) showed a connection between ensemble spread and forecast skill when the ensemble perturbations are designed to increase the low-frequency ENSO variability. However, low ensemble spread would occur if random perturbations are used for the initial ensemble. Toth and Kalnay (1996) also point out that initial coupled ensemble perturbations need to be generated in such a way that they carry the slowly growing coupled instabilities for ENSO prediction. These studies illustrate the concept that the initial ensemble perturbations and their evolution should project on the dominant modes of variability of the coupled system.

Since the upper ocean circulation is mostly a wind-driven process, the changes in sea surface temperature (SST) anomalies result from coupled atmosphere-ocean processes. Most of the current methods for generating ensemble perturbations for coupled models are intended to perturb the wind to assess the uncertainties in SST fields. For
example, in a Tier 2 system*, the atmospheric ensemble is generated under the influence of only a single SST field (Bengtsson et al. 1993). Tier 1 systems generate the ensemble members via a coupled ocean-atmospheric model in order to have the perturbation grow under a coupled configuration (e.g. Stockdale et al. 1998; Saha et al. 2006). This single-stage configuration is now widely used for operational ENSO prediction but the initial perturbations are introduced only in either the atmospheric or the ocean component. As a result, ensemble perturbations are not optimally coupled at the initial time and a spinup time is needed to couple the oceanic and atmospheric perturbations. Such a lack of coupling in the initial ensemble perturbations would limit the initial growth of the ensemble members and their ability to depict the uncertainties in the forecasted state.

The coupled breeding technique proposed by Toth and Kalnay (1996) provides a method to derive the slowly varying, coupled instability for initial ensemble perturbations. By repeatedly rescaling and evolving the coupled bred perturbations upon the background flow with the full nonlinear coupled model, the breeding cycle is aimed at filtering out fast unwanted instabilities (like fast weather noise) and retaining the growing mode we are interested in (slowly varying coupled instabilities). A bred vector (BV) is defined as the difference between two nonlinear runs (one with bred perturbation and the other without any perturbation). By taking advantage of the fact that instabilities from a complex coupled system are characterized by different time scales, and that their amplitudes saturate at different levels, nonlinear integrations allow the unwanted fast growing instabilities to saturate early, and leave the slow perturbations dominating the

---

* In a Tier 2 system, a single SST, generated from the coupled system, is used to initialize the atmospheric ensemble members for the stand-alone AGCM forecasts. In a Tier 1 system, both atmospheric and oceanic components are perturbed at the same time and evolve with the coupled model.
bred perturbations. The key to selecting the instability of interest is to use physically meaningful breeding parameters, i.e., the rescaling amplitude and rescaling time interval (Toth and Kalnay, 1993, 1997 and Peña and Kalnay, 2004). Peña and Kalnay (2004) derived bred vectors, singular vectors (SV) and Lyapunov vectors (LV) with experiments using fast and slow coupled Lorenz 3-variable models. Their idealized coupled system was designed to mimic a coupled system including a tropical ocean, and tropical and extratropical atmospheres. They demonstrated that breeding can be used to isolate the slow modes when choosing a slow variable to measure the growth of the perturbation and when using a long interval for the rescaling, allowing the fast extratropical weather noise to saturate. In contrast, SV and LV are swamped by the fast growing instabilities, since by being linear they are designed to capture the fastest growing instability of the system.

The coupled breeding experiments of Cai et al. (2003) with a Cane-Zebiak intermediate coupled model, and of Yang et al. (2006) with the coupled general circulation model (CGCM) developed at NASA Goddard Space Flight Center gave very similar robust conclusions. In these breeding experiments, all the dynamical variables in both atmospheric and oceanic components are perturbed so that coupled nonlinear processes are naturally included. Both studies show that ENSO bred vectors and their growth rate are sensitive to the background ENSO evolution, and that the dominant fast growing mode in the SST appears in the eastern equatorial Pacific. Moreover, Yang et al. (2006) confirmed that the characteristics of bred vectors from two independently developed CGCMs (NASA and NOAA/NCEP) have a similar structure, not only in the tropical Pacific basin but also in the atmospheric ENSO-teleconnected region. These
findings suggest that the perturbations derived from breeding runs are representative of the coupled atmospheric and oceanic changes associated with the ENSO variability. Cai et al. (2003) also showed that using a pair of positive/negative BVs as the ensemble perturbations, the “spring barrier” in ENSO forecast skill could be substantially reduced.

Although the results of these studies were encouraging, the coupled breeding experiments were performed under a perfect model scenario. In this paper we extend the coupled breeding technique to a realistic scenario: performing breeding cycles in the NASA operational CGCM with real, noisy observations assimilated into the ocean component. The first goal in this study is to test whether BVs (representing the leading coupled instabilities) can be used to represent the structure of the short-range (about one month) coupled forecast error. If BVs turn out to be similar to forecast errors then they have potential for use in ensemble forecasting. Another goal in our study is to explore the potential to incorporate bred vectors with operational ocean data assimilation for better use of the observations. The current ocean data assimilation in the NASA operational system uses the optimal interpolation (OI) scheme. The background error covariance used in the OI scheme is a statistical estimation of the forecast error structure averaged in time and therefore it is flow-independent. A data assimilation scheme with a time-independent background error covariance may underestimate sudden changes of the background flow due to the dynamical instabilities. Advanced oceanic data assimilation schemes carrying flow-dependent error statistics, like ensemble ocean data assimilation (Evensen, 2003, Keppenne et al. 2002, 2003, 2005) or 4-dimensional variational analysis (4D-Var)

† “Spring Barrier” is a common feature of a sudden drop of forecast skill for the forecasts that begin in the spring and pass through the summer.
(Weaver et al., 2003, Vialard et al. 2003), have shown their advantages in the tropical Pacific and have a positive impact on the ENSO forecasts. However, even in these advanced methods, it is difficult to include the impact from coupled processes needed for better initialization for ENSO prediction when constructing the background error covariance (e.g., uncertainties in the wind forcing). Moreover, the computational cost of these advanced techniques is still a major concern. Therefore, we propose to test the use of bred vectors obtained from the CGCM to provide information related to coupled error structures at a very low computational cost. If indeed there are similarities between bred vectors and forecast errors (which we will determine according to our first goal), BVs have the potential to provide information related to the slowly varying, coupled error structure in the OI scheme by augmenting the time-independent background structure with flow dependent structure (Corazza et al. 2002 and Yang et al, 2006). Alternatively, the BVs can be used in conjunction with the background error structure derived from the ocean EnKF of Keppenne et al (2002) for better description of background uncertainties and to alleviate the need for a large ensemble. Therefore, in this study we would like to examine how well the structure of the forecast “errors of the month” is represented in the CGCM with bred vectors. We will examine the characteristics of the one-month forecast error and bred vectors against the evolving background state. If an apparent link can be diagnosed among these quantities, the structures captured by bred vectors can be applied as the ensemble perturbations not only for projecting on the dominant seasonal-to-interannual associated variability but also for an augmentation of the error covariance structure in oceanic data assimilation scheme.
Our paper is organized as follows. In section 2, we briefly introduce the NASA coupled general circulation model (GCM) used for experimental forecasts and the setup of the coupled breeding experiments. Section 3 describes the relationship between bred vectors and forecast errors. The application of bred vectors in ocean data assimilation scheme is discussed in section 4. The summary and discussion are given in section 5.

2. Description of the breeding experiments in the GMAO operational system

2.1. The NASA GMAO coupled GCM

The GMAO coupled model is a fully coupled global ocean-atmosphere-land system developed at NASA Goddard Space Flight Center (GSFC) (Vintzileos et al. 2003). It is comprised of the NSIPP-1 atmospheric general circulation model (AGCM, Suarez 1996, Bacmeister and Suarez 2002; Bacmeister et al. 2000), the Poseidon ocean model (OGCM, Schopf and Loughe 1995, Yang et al. 1999), and the Mosaic land surface model (LSM, Koster and Suarez 1992) all developed for the purpose of improving seasonal to interannual prediction.

The model variables of the ocean GCM are arranged on isopycnal layers; they are thickness, temperature, zonal and meridional current velocity and salinity. The oceanic grid resolution is 5/8° in longitude, 1/3° in latitude and 27 layers in vertical. The model variables of the atmospheric components are prescribed in sigma coordinates; they are surface pressure, zonal and meridional wind, potential temperature and specific humidity.
The atmospheric grid resolution is 2.5° in longitude, 2.0° in latitude and 34 levels in vertical. Detailed descriptions of the coupled model and the model performances for climate variabilities can be found under [http://gmao.gsfc.nasa.gov/cgi-bin/products/climateforecasts/index.cgi](http://gmao.gsfc.nasa.gov/cgi-bin/products/climateforecasts/index.cgi).

In this study, breeding experiments are performed in the operational configuration, where the ocean is initialized with a univariate OI analysis scheme, assimilating daily subsurface temperature observations from the global XBT data (Troccoli et al. 2003). In order to maintain the relationship between temperature and salinity to conserve water mass properties, the local salinity profile from the model is then adjusted by the temperature analysis increment according to the method of Troccoli and Haines (1999). During the assimilation process, the sea surface temperature (SST) is strongly relaxed to the SST of Reynolds et al. (2002), referred to hereafter as “Reynolds SST”.

For the coupled forecasts, the atmospheric and land states are initialized independently from the ocean, and from NSIPP AMIP-style (Atmospheric Model Intercomparison Project) runs (Gates, 1992) whose fields are obtained with the Reynolds SST specified as the boundary condition. Prior to coupling, the atmospheric state spins up for one day from the AMIP-style initial fields. The ocean is then initialized by this daily averaged wind stress, which allows the ocean to include the memory of past wind, and is important for maintaining the subsurface structure and reducing climate drift and initial
shock (Latif et al. 1993, Schneider et al. 1999). From the results of coupled forecasts, one of the benefits of updating the ocean initial condition with the OI scheme is a reduction of the cold seasonal drift in the early forecast months, particular for the tropical central to western Pacific.

In the operational forecasting system with six ensemble members, the NSIPP coupled GCM is used for forecasts of 12-month duration. However, the forecast skill has a strong dependence on the seasonal cycle. For example, the prediction of the Niño3 SST anomaly index has the best prediction skill (in terms of anomaly correlation and mean error) when starting from the cold phase of the annual cycle in the tropical Eastern Pacific, especially when starting from September. On the other hand, a large mean forecast SST error is observed at early forecast leads when the coupled GCM starts from May-June, a time in which SST is rapidly cooling and before the strong cold equatorial tongue is established. In addition, except for the cases starting from September, the operational ensemble exhibits rather small spread at early forecast leads and is insensitive to season. The forecast of heat content exhibits skill similar to that of SST in the tropical Eastern Pacific. The forecasted heat content in the Western Pacific shows good skill in the early forecast months starting from February to June. However, the ensemble spread is still very limited in the subsurface in equatorial region and is far from the level of the mean anomaly forecast error. In addition, and in contrast to the real evolution, the forecast state in the Eastern Pacific does not seem to benefit from the memory of the Western Pacific evolution, with SST anomalies dominated by local processes for

‡ Despite the one day coupling, the introduction of a different atmosphere produces a short-lived “coupling shock”. In this work we assume that since the BVs are rescaled with a period long enough to recover from the coupling shock, they are not fundamentally affected by it.
February – June starting months, and minimal coupling between the surface and the thermocline (Galanti et al. 2002). These observations also indicate the need for ensemble perturbations that reflect seasonal uncertainties and take advantage of the full use of the physics of the mixed layer.

Based on these considerations, the main comparisons between bred vectors and forecast errors will first focus on their evolution within the oceanic component, including the surface and subsurface variations. The oceanic growing error is measured by the first month forecast error, which is estimated as the difference between the first month control forecast and the corresponding analysis (both available as daily snapshots). The control forecast is initialized from the unperturbed initial condition.

2.2. Breeding experiment

As discussed in Section 1, initial perturbations generated through coupled breeding are designed to include the growing errors associated with the slowly varying, coupled instabilities in order to capture uncertainties when predicting seasonal to interannual variations. In the breeding cycle, there are two parameters allowing us to select the growing instability of interest: (1) the rescaling period and (2) the size of rescaling norm (Toth and Kalnay, 1996, Peña and Kalnay, 2004, Yang et al. 2006 and Vikhliaev et al. 2007). The choice of these parameters needs to be based on physical considerations (see the discussion related to Figure1 in Yang et al. (2006)). For our purpose, a length of the rescaling period of at least two weeks is required to allow the fast growing weather
signals to saturate and, at the same time, keep the slow, coupled instability in the bred perturbations. Vikhliaev et al. (2006) obtained similar results rescaling the bred perturbations from the COLA coupled GCM every six months with a norm measuring the growth of upper ocean heat content.

The procedure we used to perform the breeding cycle in the GMAO operational system is the same as described in Yang et al (2006), with a rescaling period of one month. The period used for experimental breeding and control (non-perturbed) forecasts is January 1993 to November 1998. The choice of a one-month rescaling period is fairly reasonable according to the error statistics from this coupled GCM. SST errors from operational ensemble forecasts have their fastest growth on the first month and grow much more slowly after two months. Climate drift has little impact on the relative error growth rate in the early forecast and removing it mainly reduces the magnitude of the mean error.

During the coupled breeding cycle, all the model variables in the ocean and atmospheric components are perturbed. The oceanic-bred perturbations are added to oceanic analysis fields, and atmospheric bred perturbations are added to the AMIP-style atmospheric restart fields. After one month of coupled integration, the coupled bred perturbation is computed as the difference between the one-month perturbed and unperturbed (control) forecasts, and is rescaled to its initial amplitude according the chosen rescaling norm. The rescaled perturbation is then added to the next oceanic and atmospheric initial conditions. In our standard breeding cycle, the rescaling factor is the
inverse of the growth of the RMS SST perturbation in the Niño3 region [150°W-90°W, 5°S-5°N] after one month, and is applied globally to all oceanic and atmospheric variables. The process is repeated throughout the breeding period (January 1993 to November 1998). Bred vectors (BV) are the difference between the one-month perturbed and unperturbed (control) forecasts and they provide an estimate of the structure of growing instabilities, which presumably are also generating forecast errors. It is important to keep in mind that bred vectors are generated without any information about the observations used to estimate the new analysis state.

We have performed breeding experiments with three types of rescaling norms with the amplitude chosen to be about 10% of climate variability (i.e. the background anomaly variance). The first two experiments use global perturbations for both ocean and atmosphere states and their rescaling factors are measured by (1) the SST bred perturbations in the Niño3 region with an amplitude of 0.1°C (the standard run), (2) the perturbation of depth of the 20°C isotherm (Z20) in the tropical Pacific with an amplitude of 0.2 m. In the third experiment the oceanic and atmospheric perturbations are added only in the tropical regions and the rescaling factor is also determined by the SST bred perturbation in Niño3 region. Since the BV structures obtained from different rescaling norms are dominated by similar (but not identical) growing instabilities, comparisons between the structures of forecast error and bred vectors are focused on the run rescaled with the Niño3 SST perturbations norm in the next section.
3. **Observed relationship between one month forecast error, bred vectors and background anomaly**

   In this section, we illustrate the characteristics of forecast error and bred vector from this coupled system and how they evolve with the slowly varying background state (i.e., the background ENSO events). The one-month forecast error is approximated by the difference between the one-month control forecast and the analysis state verifying at the same time. This quantity represents the optimized correction to the forecast state after combining it with observational information and as such is a proxy for *forecast error*. The bred vector is derived as the difference between the perturbed forecast and the non-perturbed forecast. We will first assess whether ENSO dynamical processes dominate the evolution of both one-month forecast errors and bred vectors.

   In ENSO variability, the slow dynamical processes in the ocean component play an important role in characterizing an ENSO event and, therefore, such processes should also determine the characteristics of the growing errors. This can be verified by examining the structures of forecast error and bred vector as they evolve with the background ENSO events. Several studies have indicated that the origin of an ENSO event can be traced back to *perturbations in the subsurface* of the western Pacific, which is excited by variations of the trade wind. During an El Nino event, the downwelling perturbation travels to the eastern Pacific through equatorial wave propagation, deepening the thermocline, allowing the anomalous warm water to establish, and reducing the zonal temperature gradient. The subsequent positive feedbacks from air-sea
interaction then contribute to the growth of a warm episode (McPhaden et al. 1997, McPhaden 1999, Wallace et al. 1997 and Boulanger and Menkes 1995). Thus, dynamical coupled errors in the surface layer of the eastern Pacific will not only be affected by the local processes (e.g. upwelling) but also be closely related to the errors in the subsurface state of the western Pacific, where it is also connected to the atmospheric conditions.


The 1997-98 El Niño is a strong warm event whose background anomaly state is characterized by distinct features of equatorial wave propagation (McPhaden 1999, McPhaden and Yu 1999). The event was triggered by a strong westerly wind burst in the western Pacific; the extreme intensity of this event has been attributed to the nonlinear interaction between ocean and atmosphere (Zhang and McPhaden 2006, Vecchi et al. 2006).

For the following analysis we select four months representing the different stages of the 1997 El Niño event (prior, developing, mature, and post stages) in order to illustrate how the forecast errors and bred vectors evolve during ENSO development. Figure 1 shows the Niño3 index for the period October 1996 to October 1998 covering this major El Niño event that peaked at the end of 1997. Figures 1a-d are snapshots of vertical cross sections along the equator of forecast error (color shading) and bred vectors (contour) of temperature corresponding to prior, developing, mature and post El Niño
stages. These snapshots are plotted in model coordinates (layers), where the mean thermocline depth is located near the 10th-15th layers. As shown in Figure 1a (the prior stage), before the warming commences the forecast error was mainly located in the subsurface of the western Pacific with some error also near the eastern coast. During the developing stage, the forecast error shifted eastward and extended vertically as the warm anomaly built up in the eastern Pacific (Figure 1b). At the mature stage of the event, the forecast errors were smallest near the eastern coast and upper levels near the surface (Figure 1c). When the warm anomaly diminished and the background state returned to normal conditions, the forecast error accumulated mostly in the eastern Pacific (Figure 1d). It is remarkable that very similar longitude-vertical variations appear as well in the bred vectors shown with contour lines in the same figures. The bred vector evolution is characterized by an eastward propagation well synchronized with the forecast error movement and BV shapes tend to capture the large-scale forecast error, as indicated by the collocation of forecast error and bred vector. These results support the hypothesis that the bred vectors obtained from such a complicated CGCM with data assimilation are able carry the information related to realistic ENSO development and that they indicate where the forecast errors will be located. Even without access to the observations used in estimating the analysis state and the forecast errors, the results show that the bred vector contains the dynamical instabilities that lead to the forecast errors corrected in the data assimilation. The good agreement between bred vector and forecast error in the model coordinate suggests their potential application incorporating information on the “errors of the month” into the data assimilation process. More details and statistical diagnostics are provided in section 4.
The similarity between bred vectors and forecast errors remains strong after converting the bred vector and forecast error from model layers to physical ocean depth by combining the variations of model layer thickness and temperature. As could be expected, both the forecast error and bred vector maxima are located near the thermocline, where sharp temperature gradients can easily introduce instabilities (figures not shown). The structure of the bred vector suggests its potential impact when used as an initial ensemble perturbation for ensemble forecasting since it projects strongly on background instabilities leading to forecast error growth.

3.2. The temporal and spatial relationships between bred vector, forecast error and background variability

An important conclusion from Yang et al. (2006) and Cai et al. (2003) was that there is a relationship between the ENSO phase and the bred vector growth rate, with the maximum BV growth occurring between the two extreme phases of ENSO episodes. BV growth rate is measured as the amplitude of the BV SST at the rescaling time (i.e. one-month after coupled integration) relative to the initial amplitude and it has a non-dimensional unit per month. This relationship remains valid for the BV growth rate obtained from the operational system with data assimilation. Figure 2a is the background Niño3 index computed using the SST to indicate the temporal evolution of the ENSO episodes from 1994 to 1998. Figure 2b shows that the BV growth rates are large before and after an ENSO event and smallest at the mature state of an ENSO event. We found
that the growth rate is also strongly correlated to the root mean square (RMS) SST forecast error measured in the Niño3 region (Figure 2c vs. 2b). As already apparent in Figure 1, the amplitude of the forecast error is also very sensitive to the phase of ENSO and is related to background variability. During the decaying phase of the event, with rapid cooling, the BV growth rate shows a sharp increase and so does the forecast error. The relationship between these two time series suggests that the bred vector has similar behavior to the forecast error field and that bred vector growth rates can provide information related to large forecast errors, particularly when the background SST is neutral and after an ENSO event. After the anomalous ENSO warming is established in the background, error growth is inhibited because the background has reached a fairly stable state. It has been confirmed that the forecasts made by the same CGCM starting from August, 1997 are very successful in capturing the subsequent temporal and spatial SST variations up to 9 months in terms of both timing and intensity (not shown).

In order to quantify the relationship between bred vectors (BV), forecast error and background state, we used the pattern correlation between the bred vector and the forecast error in the Niño3 region to determine to what extent the breeding method could capture such dynamical error throughout the 6 years. The monthly pattern correlations are grouped into bins classified by BV growth rate (< 2.5, 2.5–3.5, 3.5–4.5 … and > 8.5 per month). We then calculate the mean of the absolute value of pattern correlation in each group. For comparison, the corresponding mean absolute value of the Niño3 SST anomaly is also calculated. The absolute value is used to emphasize the phase of the ENSO event, i.e. the anomalous state of the background. Figure 3 shows that the
absolute value of the Niño3 index is high when the BV growth rate is small, and close to zero when the growth rate is large, as discussed above, confirming the results of Cai et al. (2003) obtained with the ZC model. The mean pattern correlation between BVs and forecast error in the last two groups with large growth rates are much higher than for lower growth rates. This suggests that when the background is in a near-neutral state, BVs tend to have larger projection on the forecast error. The relatively small spatial correlation (~0.4) is due to the fact that such a complex system includes not only ENSO instabilities but also instabilities associated with shorter space and time scales and with the introduction of AMIP atmospheric initial conditions, noisy processes that could not be captured with breeding.

Combining the results from Figures 2 and 3, we can conclude that bred vectors are naturally influenced by the background instabilities, particularly the dominant large-scale, low frequency variabilities. The shapes of forecast errors are also affected by these same instabilities. Therefore, the bred vector can detect the rapidly changing stage of the background occurring in the eastern Pacific and represent the shape of the corresponding forecast errors.

The spatial correlation between bred vectors and forecast error suggests bred vectors can be used to identify the shapes of the fast growing instabilities that dominate the forecast error. We have examined the agreement between the bred vectors and forecast error by fitting the zonal and meridional error covariance structures with a Gaussian (see the Appendix). The correlation length derived from forecast error and bred
vector illustrates the characteristic length of the “error of the month” associated with the coupled instability. Table 1 summarizes the mean value of the correlation lengths estimated from one-month forecast error and bred vector near the equator for all months. It is seen that the ranges of the mean correlation lengths and their standard deviations are very similar for the forecast error and bred vector, suggesting that the bred vector is able to provide information on the structure of the forecast error. The characteristic scale of error of the month is much shorter than what is currently used in the OI scheme (~1500km). In addition, the fitted variance from the forecast error and bred vector both indicate larger amplitudes near the equator, particularly in the bred vector field. Results also showed that both the meridional correlation scales estimated from forecast error and bred vector are similar to the one used in the current OI scheme (~400km).

Table 1 The error zonal correlation lengths and mean error variance of SST, obtained by fitting the bred vector and forecast error with a Gaussian function

|         | Forecast error (Km) | Bred vector (Km) | STD of Correlation length (Km) | Forecast error (Km) | Bred vector (Km) | Fitted variance |
|---------|---------------------|------------------|-------------------------------|---------------------|------------------|-----------------
| 2.5°N -7.5°N | 514            | 540               | 272                           | 269                | 0.63             | 0.05            |
| 2.5°S - 2.5°N | 575            | 505               | 251                           | 226                | 0.99             | 0.14            |
| 7.5°S - 2.5°S | 445            | 416               | 217                           | 199                | 0.83             | 0.07            |

3.3. Dynamic error structure in the equatorial subsurface ocean

In the previous subsections, we have shown that in the eastern Pacific both the growth of BV SST and the forecast error are dominated by similar dynamic errors. The growth of SST perturbations in the western Pacific is limited by its low natural variability due to the deep well-mixed surface warm water. As the result, the growth of temperature
perturbations is less detectable near the surface layer in the western-central Pacific. However, *subsurface* perturbations can carry the information of coupled instabilities driven by the windstress. Moreover, the warm pool in this region plays an essential role in the interaction between the atmosphere and the ocean through heat exchange near the surface. These observations suggest that a good ensemble forecast system for ENSO prediction should be able to generate perturbations to include the uncertainties related to these effects. We notice that in the current operational ensemble perturbations, the ensemble spread is very limited near the thermocline in the western Pacific while the temperature error is large in this region.

It has been recognized that interannual variability of equatorial subsurface temperature has large variance near the mean thermocline, i.e. at a depth of about 175 m in the west Pacific and 50 m in the east. Figures 4a-b show that the first two leading EOF modes of the background anomaly oceanic subsurface temperature fields (computed using the vertical cross-section of temperature analysis fields at the Equator) are strongly related to the development of an ENSO event and explain more than 60% of the total variance of the equatorial subsurface temperature. The leading mode has large variability located in the eastern Pacific (mature El Niño) and its corresponding principal component is in phase with the variations of Niño3 index. The second mode, related to the early development of the event, exhibits large variability that extends from the subsurface of the western Pacific and peaks in the central Pacific. Based on the temperature analysis fields, this mode leads the first one by about 7 months. It is clear that for ENSO prediction, the ability to describe thermal variations in subsurface conditions should
strongly contribute to how well the SST forecast would perform. Simulated subsurface temperature variations from this CGCM have a weak magnitude compared to the observations (not shown). Such a deficiency is a common feature of many CGCMs and will crucially influence the ability to forecast SST anomalies in the equatorial eastern Pacific.

The same EOF analysis is applied to the forecast error and the bred vector temperature cross-section at the Equator. Figures 5a-c are the first three EOF modes of forecast error and their explained variances are 19%, 11% and 5%, respectively. The first two modes of the forecast error have somewhat similar locations as the dominant ENSO-related modes, and they all show large amplitudes along the thermocline. The first mode has large variabilities in both the eastern and western Pacific with comparable amplitudes. The second mode is mainly located in the central Pacific. This implies that the subsurface growing error is dominated by large-scale variations, related to the zonal tilt of the thermocline. Comparing with Figure 4, these patterns shown in Figure 5 strongly project on the space of ENSO-related variability. Consequently, such growing errors will develop variations like El Niño/La Niña.

We then applied the same EOF analysis to the bred vector equatorial temperature. Here, we use the bred vectors before rescaling, i.e. the bred vectors are weighted by their growth rates, in order to find the patterns that dominate the growing components. Figure 6a-c are the first three EOF modes having similar amounts of explained variances (7%, 8% and 9%). The EOF modes have a striking resemblance to the EOF modes of the
forecast error, though their order was changed to emphasize their relation with the forecast error. As discussed before, the relatively low explained variance results from the fact that, in addition to ENSO instabilities, there are also other instabilities and processes with very different space and time scales present in such a complex system. Despite this, the dominant growing pattern is still associated with the large-scale variations. This supports the hypothesis that bred vectors can capture the shape of the related dynamic error that dominates the forecast error. Currently, the background error covariance in the univariate Optimal Interpolation analysis uses a Gaussian shape in the horizontal and vertical directions. By contrast, the patterns of EOF modes suggest that the background error in the subsurface should have considered large-scale shapes expanding near the thermocline in order to ensure the corrections from the data assimilation scheme project on the ENSO-related low dimensional space.

We also notice that in the BV EOF modes, there is a feature locally trapped in the extreme eastern equatorial Pacific near surface. This feature is absent in the EOF modes of the forecast error. This may be due to the rapid dynamical adjustment of Ekman pumping in the shallow region off the South America coast when adding the bred perturbation to the unperturbed control background, i.e. the analysis field.


In order to explore the impact of using coupled BV for ensemble forecasting, we performed a two-sided breeding cycle (Toth and Kalnay 1997) to generate initial
ensemble perturbations and study the impact of predicting the strong 1997-1998 El Niño event. In a two-sided breeding cycle, pairs of positive and negative bred perturbations are used to perturb the initial state. The rescaling factor is then measured by the difference between this pair of perturbations after one month, calculated according to Equation (3):

$$factor = \frac{\|\delta x_0\|}{\|\delta x(t)\|} = \frac{\|\delta x_0\|}{0.5(\|x_p(t)\| - \|x_n(t)\|)}.$$  

(3)

In Equation (3), $\|\delta x_0\|$ is the initial size of the perturbation (e.g., the RMS of Niño3 BV SST chosen to be 0.1°C) and $\|\delta x(t)\|$ is the evolved (unrescaled) bred perturbation after one-month, $x_p(t)$ is the positively perturbed one-month forecast and $x_n(t)$ is the negatively perturbed one-month forecast. This constant factor is then applied to rescale the evolved bred perturbation. This rescaled perturbation is added and subtracted from the control as the next initial condition for the next breeding cycle.

Studies have shown that the trigger and development of this event is closely related to a series of westerly wind bursts with long fetch along the equator, which weakened and reversed the trade winds in early 1997 (McPhaden, 1999). The wind burst forced a downwelling Kelvin wave that deepened the thermocline in the eastern equatorial Pacific and contributed to the development of the warm SST anomaly. The reduced zonal SST gradient helped to strengthen the westerly wind anomalies. The anomalous states were further intensified due to the nonlinear interactions between the atmosphere-ocean coupling, including phase-locking with the seasonal variability in the equatorial Pacific.
The hindcast experiment with the operational ensemble is able to predict the phase and intensity of this event after it has already developed (e.g., starting on May 1997). The hindcast initialized in January 1997 forecasts a warm event, but underestimates the intensity. Using forecasts started from a pair of positive and negative BV perturbations, we found that the mean from this pair of ensemble members is able to substantially improve upon the control (unperturbed) forecast starting from January 1997. In the following analysis, we use this case to discuss the impact of applying coupled BV perturbations.

Starting from January 1997, we found that the ensemble member initialized with the positive BV captures very well the intensity of this strong El Niño event and that the anomalous phase is more accurate, as shown by the red line in Figure 7. The observed Niño3 index (blue dashed line) is computed from the Reynolds SST and is used for verification. The one initialized with the negative BV (red dash-dotted line) shows a behavior rather similar to that of the control forecast (black line). For convenience, we will refer to the positive BV as PBV, and the negative BV as NBV and the control forecast as CNT. In order to test whether the best forecast in this case can only be achieved when both the atmosphere and ocean are perturbed, we performed two additional forecast experiments that use only the ocean component of the PBV as the initial perturbation (PBVO) and only the atmospheric component (PBVA). Neither PBVO (green line) nor PBVA (magenta line) are able to reproduce the intensity of the anomalous warm state that PBV predicted, even though PBVA shows a larger impact.
than perturbing the ocean alone. Our results suggest that only when the coupled bred perturbation is applied is the evolved perturbation able to maximize the coupling effect and predict the event.

The initial SST perturbation of PBV, shown in Figure 8, has a warm perturbation located in the eastern Pacific (140°W-100°W) and a slight cool perturbation in the west. On the western edge of the warm perturbations, there are corresponding westerly wind perturbations. As we will discuss later, the westerly wind perturbation is able to expand further to the east and makes the downwelling (deepening) effect more easily amplified in the shallow region. Although the details of the appearance of the wind burst in PBV are not completely consistent with observations, the development of the warm anomaly in PBV has a similar triggering effect and results in a better prediction of the strong El Niño event.

We examine the Hovmoller diagram of zonal wind stress from these experiments and notice that only PBV (Figure 9b) has the westerly wind burst that took place in mid-February until early March, slightly earlier than the observations (Figure 9f). Also, the duration of this wind burst in Figure 9b is long and it extends to 140°W, while other forecast members show relatively neutral or easterly wind anomalies during March. In addition, only PBV shows the consistent strong positive wind anomaly after the El Niño event has been established in May 1997, suggesting its role is tightly coupled with the warm anomaly in the upper ocean. This can also be observed in the anomalous changes of the subsurface structure. Figure 10 shows the monthly mean heat content in the
equatorial upper ocean above 300m. The analysis state derived from the data assimilation scheme is used to represent the best estimate of the true state. As shown in Figure 10a, a seesaw pattern (negative-positive anomaly to represent the flattened thermocline) is evident during the intensification of this warm event because of the positive feedback from air-sea interaction. Overall, the forecast skill from all the ensembles is coming from the subsurface western Pacific, where all the forecast states have already indicated a deepening anomaly and are able to describe its eastward propagation. However, only PBV captures the dominant seesaw pattern in the heat content anomaly, though the mature phase is about one month ahead of the verification. In the early forecast months, all of the forecast states (including PBV) show a too shallow thermocline in the eastern basin and they stay in the cold phase. Such results indicate that a much better depiction of the equatorial wave propagations could be obtained if the perturbation is able to effectively modulate the wind forcing. In addition, except for PBV, the amplitude of the deepening anomaly decays prematurely and the deepening feature near the eastern end of the Pacific basin becomes neutral after September.

We notice that when the PBV atmospheric component is used (PBVA), it has some important features shown in PBV. For example, the deepening feature is able to develop in the eastern Pacific even though the amplitude is smaller, and it includes the shallowing anomaly in the western Pacific. This is evidence that the atmospheric perturbation carries information that is able to drive the tilting of the thermocline.
5. Summary and discussion

The coupled breeding method has been implemented into the NASA operational coupled GCM, in which the ocean component assimilates the real temperature observations through an OI scheme. Breeding experiments are then performed under this imperfect model scenario, so that growing errors are much more complicated than in the perfect model scenario explored in Yang et al. (2006). They include, in addition to those dominated by the coupled ENSO dynamics, other errors such as initial errors due to an imperfect data assimilation scheme, noisy observation errors, and model errors (as well as those due to the spin-up of the AMIP atmospheric initialization). We explore the possibility that bred vectors could still be used as a proxy to represent the structure of the one month forecast errors despite the presence of these complications in a real operational system. Potential applications for this work are the improvement of ENSO ensemble forecasting system and the better use of observations in the data assimilation scheme by including evolving “errors of the month”.

For the ENSO prediction, the forecast SST in the tropical Pacific is the most crucial variable directly depicting air-sea interaction and determining the performance of the coupled GCM. As a first step, we analyze the temperature structure characteristics of the one-month forecast errors and the bred vectors. One month forecast errors show dynamical structures evolving with the variations of the background flow and bred vectors strongly project onto these forecast errors. Both the amplitude of forecast error and the bred vector growth rate are dominated by low-frequency variations with a very
similar trend: they are small when the background anomalies are in the mature phase of ENSO events and are large when the background anomalies are near neutral, especially during the sharp transition that occurred from the 1997-8 strong El Niño to the 1999 La Niña event. A statistical analysis suggests that the agreement between bred vector and forecast error is particularly good when the BV growth rate is large and when the background anomaly is near neutral.

For the 1997 El Niño event, the forecast error of the temperature cross-section along the equator large error appeared first in the subsurface of the western Pacific, propagated eastward and amplified off the east coast. The operational forecast error is smallest near the surface at the mature stage but the subsurface errors show a wide zonal expansion. The evolution of the one-month subsurface forecast error structure shows that the errors reflect the large-scale displacement of the thermocline associated with the anomalous background variations. The bred vectors at the relevant times captured most of the shapes of the one-month forecast error, including the eastward propagation and vertical and zonal expansion, even though they were computed without using observations. In addition, our results suggest that the leading EOF modes of subsurface structure of one month forecast errors and bred vectors are both dominated by similar large-scale features with peak variations near the thermocline.

The potential application of using bred vectors in the ocean data assimilation is also explored in this study. Due to the flow-dependent error structure, the correlation scale or covariance structure should carry information on the “errors of the month” and
allow for better use of the observations. For a complex system like a coupled GCM, it is far too costly and practically impossible to retrieve the full dimension, atmospheric-ocean error covariance structure. However, it is possible to create a shortcut, representing the dominant structures with the bred vectors derived here to augment the time-independent covariance structure. We have demonstrated that bred vectors are characterized by similar correlation length as the ones derived from the forecast error, even though the bred vector is calculated only from the difference between perturbed and unperturbed forecasts and has no information about the analysis used to calculate the forecast errors. The horizontal correlation length associated with “errors of the month” is found to be shorter than the one used in OI scheme.

Finally, we include the coupled bred vectors with two-sided breeding in the ensemble members to explore the impact of using coupled initial perturbations. Our results show that the mean from the pair of BV ensemble members substantially improves the control forecast starting from January 1997 because one of the BV members captures the most of the important features in the 1997-1998 El Niño event. We found that such result cannot be achieved if only the atmospheric or the oceanic component of the bred perturbation is retained. Only when the coupled perturbation is used, are the amplitude and phase of the forecasted warm event close to the observed evolution. The successful member initialized with the positive BV is the only one that developed westerly wind bursts that extended eastward of the dateline in early 1997. We believe this precursor contributes to the subsequent well-developed positive feedback of this event,
including the deepening thermocline in the eastern Pacific and the consistent positive zonal wind stress anomaly that prevailed in the central basin.

For the next phase of our studies, we will extend the coupled bred vectors with two-sided breeding in the ensemble members for a much longer hindcast experiment period so that we can explore their impact on different starting months. Additionally, based on our results of error structure in subsurface temperature, we will explore the possibility of including the time-varying error structures derived from the BV in the background error covariance within a data assimilation scheme like OI.

Appendix

In this section, we briefly describe the procedures that we use to obtain the fitting parameters of a Gaussian function (Equation (1)), i.e. the fitted error variance and the correlation length.

\[ f(r) = C_0 \exp\left(-\frac{r^2}{L^2}\right) \]  

(1)

The Levenberg-Marquardt algorithm is first applied to estimate these parameters and the standard deviation of the fitting points are required for nonlinear fitting. In our results, the standard deviations from both the one-month forecast error and bred vector have larger values at short ranges and smaller values near the tail. Therefore, this scheme will try to overfit the tail. In order to avoid overemphasizing the small covariances at large separations (which is dominated by sampling errors), we use the e-folding value of the
average of the first two fitting points to determine how many data points will be used for fitting and a corresponding length, $L_e$, is chosen.

Once $(C_0, L)$ are determined, we repeat the procedure by increasing the number of points used for the fitting and calculate the rms error within $L_e$. The $(C_0, L)$, which has the smallest rms error, is assigned to our final result. If the e-folding range cannot be found at the last point of the data, it suggests a very large-scale linear behavior of the error covariance. We then switch to a simple linear regression form (Equation (2)) to avoid failure in fitting the nonlinear scheme:

$$\ln(f(r)) = \ln C_0 - \frac{r^2}{L^2}. \quad (2)$$

Acknowledgements

We are very grateful for much technical support from NASA global modeling assimilation office support staff, especially to Sonya Miller. We also greatly appreciate many valuable discussions with Joaquim Ballabrera. The GMAO component of this work has been supported by grants from NASA Seasonal-to-Interannual Prediction Project (NASA-NAG-55825 and NASA-NAG-12418).

References


Cai M., E. Kalnay, and Z. Toth, 2003: Bred vectors of the Zebiak-Cane model and their


Schneider, E., B. Huang, Z. Zhu, D. DeWitt, J. Kinter III, B. Kirtman and J. Shukla.


____, ____ , and ____, 1997b: Predictability of a coupled model of ENSO using singular


TABLE CAPTION

Table 1 The error zonal correlation lengths and mean error variance of SST, obtained by fitting the bred vector and forecast error with a Gaussian function. The operational OI system assumes a zonal error correlation length of about 1500 Km.

FIGURE CAPTION

Figure 1 (a) Background Niño3 index (°C) and vertical cross-section of temperature analysis increment (°C, color) and BV temperature (°C , contour) corresponding to (a) October 1996, prior to ENSO stage, before warming developed (b) April 1997, developing stage, when warming started (c) November 1997, peak stage, when warming is strongest, and (d) April 1998, post ENSO stage, when the warm anomaly diminished. The contour interval is 0.5°C and the zero contour is not plotted.

Figure 2 (a) Background Niño3 index (°C), (b) bred vector growth rate (per month) and (c) root mean square of the difference between SST analysis and one-month forecast in the Niño3 region (°C).

Figure 3 Mean value of the pattern correlation (the blue line) and the Niño3 index (red line) in bins defined by the BV growth rate. Pattern correlation and the Niño 3 index are grouped based on their corresponding growth rate located within the ranges of < 2.5, 2.5-3.5, 3.5-4.5,… and > 8.5. Pattern correlation is defined as the spatial correlation
between the bred vector and analysis increment in the Niño3 region and the absolute value is used for both the pattern correlation and the Niño3 index.

**Figure 4** (a) The first EOF mode of the equatorial temperature anomaly and (b) the second EOF mode. The thick dashed line is the depth of the mean thermocline. EOF modes are normalized, and their explained variance is in parenthesis.

**Figure 5** (a) The first EOF mode of equatorial temperature of analysis increment, and (b) the second EOF mode and (c) the third EOF mode.

**Figure 6** (a) The third EOF mode of equatorial temperature of the unrescaled bred vector, and (b) the second EOF mode and (c) the first EOF mode.

**Figure 7** Forecasted Niño3 indices from different initial condition starting from January 1997. The observed Niño3 index (from Reynolds SST) is denoted with blue dashed line. The control forecast without perturbation is denoted as the black line. The red solid line is the forecast initialized with PBV, the red dot-solid line is the one initialized with NBV, the green solid line is the one initialized with perturbed ocean only from PBV and the line is initially perturbed atmosphere only from PBV. Please see the detail discussion in Section 4.1.

**Figure 8** Bred perturbation used to initialize PBV: color shading: SST perturbation. Vector: wind perturbation averaged from 1000 hpa to 850 hpa.
Figure 9 Forecasted zonal wind stress anomaly in the upper ocean from January 1997 to December 1997 from (a) the control forecast, (b) the perturbed forecast initialized with PBV, (c) the perturbed forecast initialized with the NBV, (d) the perturbed forecast initialized with PBVO (ocean only), (e) the perturbed forecast initialized with PBVA (atmosphere only) and (f) observations, obtained from SSM/I. All fields are daily data with 1 week running mean.

Figure 10 Monthly mean equatorial heat content anomaly in the upper ocean (defined as the depth of 300m) from January 1997 to December 1997 from (a) the control forecast, (b) the perturbed forecast initialized with PBV, (c) the perturbed forecast initialized with the NBV, (d) the perturbed forecast initialized with PBVO (ocean only) and (e) the perturbed forecast initialized with PBVA (atmosphere only) and (f) the analysis data. The analysis data used for validation is derived from the OI scheme implemented in the same ocean GCM to assimilate temperature and synthetic salinity.
Figure 1 Background Niño3 SST index (°C) (left panel) and vertical cross-section of the temperature analysis increment (°C, color shades) and BV temperature (°C, contours) on model layers corresponding to (a) October 1996, prior to ENSO stage, before warming developed, (b) April 1997, the developing stage, when warming started (c) November 1997, the mature stage, when warming is strongest, and (d) April 1998, post ENSO stage, when the warm anomaly diminished. The contour interval is 0.5°C and the zero contour is not plotted.
Figure 2 (a) Background Niño3 index (°C), (b) bred vector growth rate (per month) and (c) root mean square of the difference between the SST analysis and one-month forecast in the Niño3 region (°C).
Figure 3 Mean value of the pattern correlation (the blue line) and the Niño3 index (red line) in bins defined by the BV growth rate. The pattern correlation and Niño 3 index are grouped based on their corresponding growth rate located within the ranges of < 2.5, 2.5-3.5, 3.5-4.5,..... and > 8.5. Pattern correlation is defined as the spatial correlation between the bred vector and analysis increment in the Niño3 region. The absolute value is used for both the pattern correlation and the Niño3 index.
Figure 4 (a) The first EOF mode of the equatorial temperature anomaly and (b) the second EOF mode. The thick dashed line is the depth of the mean thermocline. EOF modes are normalized, and their explained variance is in parenthesis.

Figure 5 (a) The first EOF mode of equatorial temperature of forecast error, and (b) the second EOF mode and (c) the third EOF mode. (a) EOF1 (19%) (b) EOF2 (11%) (c) EOF3 (6%)

Figure 6 (a) The third EOF mode of equatorial temperature of the unrescaled bred vector, and (b) the second EOF mode and (c) the first EOF mode. (a) EOF3 (7%) (b) EOF2 (8%) (c) EOF1 (9%)
Figure 7 Forecasted Niño3 indices from different initial conditions starting from January 1997. The observed Niño3 index (from Reynolds SST) is denoted with blue dashed line. The control forecast without perturbation is denoted as the black line. The red solid line is the forecast initialized with PBV, the red dot-solid line is the one initialized with NBV, the green solid line is the one initialized with perturbed ocean only from PBV and the line is initially perturbed atmosphere only from PBV. Please see the detailed discussion in Section 4.1.
Figure 8 Bred perturbation used to initialize PBV: color shading: SST perturbation. Vector: wind perturbation averaged from 1000 hpa to 850 hpa.
Figure 9 Forecasted zonal wind stress anomaly in the upper ocean from January 1997 to December 1997 from (a) the control forecast, (b) the perturbed forecast initialized with PBV, (c) the perturbed forecast initialized with the NBV, (d) the perturbed forecast initialized with PBVO (ocean only), (e) the perturbed forecast initialized with PBVA (atmosphere only) and (f) observations, obtained from SSM/I. All fields are daily data with 1 week running mean.
Figure 10 Monthly mean equatorial heat content anomaly in the upper ocean (defined as the depth of 300m) from January 1997 to December 1997 from (a) the control forecast, (b) the perturbed forecast initialized with PBV, (c) the perturbed forecast initialized with the NBV, (d) the perturbed forecast initialized with PBVO (ocean only) and (e) the perturbed forecast initialized with PBVA (atmosphere only) and (f) the analysis data. The analysis data used for validation is derived from the OI scheme implemented in the same ocean GCM to assimilate temperature and synthetic salinity.