Improving SST Anomaly Simulations in a Layer Ocean Model with an Embedded Entrainment Temperature Submodel

RONG-HUA ZHANG, ANTONIO J. BUSALACCHI, AND RAGHURAM G. MURTUGUDDE

Earth System Science Interdisciplinary Center, University of Maryland, College Park, College Park, Maryland

(Manuscript received 19 November 2004, in final form 2 November 2005)

ABSTRACT

In this study, an improved sea surface temperature (SST) anomaly (SSTA) solution for the tropical Pacific is presented by explicitly embedding into a layer ocean general circulation model (OGCM) a separate SSTA submodel with an empirical parameterization for the temperature of subsurface water entrained into the ocean mixed layer ($T_e$). Instead of using subsurface temperature directly from the OGCM, $T_e$ anomalies for the embedded SSTA submodel are calculated from a historical data-based empirical procedure in terms of sea level (SL) anomalies simulated from the OGCM. An inverse modeling approach is first adopted to estimate $T_e$ anomalies from the SSTA equation using observed SST and simulated upper-ocean currents from the OGCM. A relationship between $T_e$ and SL anomalies is then obtained by utilizing an empirical orthogonal function (EOF) analysis technique. The empirical $T_e$ parameterization optimally leads to a better balanced depiction of the subsurface effect on SST variability by the mean upwelling of anomalous subsurface temperature and vertical mixing in the equatorial Pacific. As compared with a standard OGCM simulation, SSTA simulations from the embedded submodel exhibit more realistic variability, with significantly increased correlation and reduced SSTA errors due to the optimized empirical $T_e$ parameterization. In the Niño-3 region (5°S–5°N, 150°–90°W), the anomaly correlation and root-mean-square (RMS) error of the simulated SST anomalies for the period 1963–96 from the standard OGCM are 0.74° and 0.58°C, while from the embedded SSTA submodel they are 0.94° and 0.29°C in the $T_e$-dependent experiment, and 0.86° and 0.41°C in the experiment with one-dependent-year data excluded, respectively. Cross validation and sensitivity experiments to training periods for building the $T_e$ parameterization are made to illustrate the robustness and effectiveness of the approach. Moreover, the impact on simulations of SST anomalies and El Niño are examined in hybrid coupled atmosphere–ocean models (HCMs) consisting of the OGCM and a statistical atmospheric wind stress anomaly model that is constructed from a singular value decomposition (SVD) analysis. Results from coupled runs with and without embedding the SSTA submodel are compared. It is demonstrated that incorporating the embedded SSTA submodel in the context of an OGCM has a significant impact on performance of the HCMs and the behavior of the coupled system, with more realistic simulations of interannual SST anomalies (e.g., the amplitude and structure) in the tropical Pacific.

1. Introduction

Sea surface temperature (SST) in the tropical Pacific has significant effects on seasonal-to-interannual climate variability worldwide (e.g., Cane and Zebiak 1985; Barnett et al. 1993). In past decades modeling and predicting SST variability has made great progress in the tropical Pacific. Various ocean and atmosphere–ocean coupled models with varying levels of complexity have been developed to better describe and simulate SST variability in the region and its essential physics (e.g., McCreary and Anderson 1991; Neelin et al. 1992; Latif et al. 2001); for classifications of intermediate and hybrid coupled atmosphere–ocean models (HCMs), see McCreary and Anderson (1991). Nevertheless, SST anomaly (SSTA) simulations in the tropical Pacific Ocean still remain inadequate for many climate studies, with relatively large systematic errors in most state-of-the-art numerical models (e.g., Stockdale et al. 1998; Ji et al. 1998; Latif et al. 2001). Furthermore, the skill of SSTA simulation and prediction in the equatorial Pacific is strongly model dependent and widely divergent.
across various coupled prediction systems (e.g., Latif et al. 1998; Barnston et al. 1999).

It has been known that subsurface processes at the base of the ocean mixed layer (entrainment and mixing) are important in controlling SST variability in the central and eastern equatorial Pacific where the thermocline is shallow and the mean upwelling is strong (e.g., Zebiak and Cane 1987, hereafter ZC87; Kleeman 1993; Wang and McPhaden 2000; Keenlyside 2001). Extensive modeling studies indicate that the temperature of subsurface water entrained into the mixed layer \( T_e \) plays an important role in producing and sustaining SST anomalies associated with the El Niño–Southern Oscillation (ENSO) in the tropical Pacific. Thus, a crucial requirement for realistic SSTAs simulations in ocean and coupled ocean–atmosphere models is the parameterization of entrainment and vertical mixing processes associated with the subsurface temperature anomalies. Since \( T_e \) is a major factor determining SST variability, realistic estimates of \( T_e \) can effectively improve SSTAs simulations and predictions in numerical models. On the other hand, errors in \( T_e \) can inevitably be a major error source for SSTAs simulations and predictions as well.

To improve SSTAs simulations, we focus on better estimates of the \( T_e \) fields for use in ocean models. However, determination of entrainment temperature beneath the mixed layer is a great challenge since the details of the temperature and density structure below the mixed layer and the mixed layer–thermocline interactions are not well known due to a lack of high vertical resolution observations in this region (e.g., Wang and McPhaden 2000). Moreover, \( T_e \) is not directly observed, and therefore an algorithm is needed to determine the mixed layer depth and then to estimate the \( T_e \) (e.g., Kara et al. 2000). In this context, an empirical \( T_e \) model has been developed that has the potential to significantly improve SSTAs simulation and prediction (Zhang et al. 2003, 2005). First, an inverse modeling approach is adopted to estimate \( T_e \) anomalies using an SST equation, observed SST, and simulated mean and anomaly currents. In so doing, for a given SST equation, the inverted \( T_e \) anomalies yield an optimized estimate of \( T_e \) for use in simulating SSTAs by balancing various terms in the heat budget of the mixed layer. Then, a relationship between the so determined \( T_e \) anomalies and simulated sea level (SL) anomalies (a diagnostic or a prognostic field from an ocean model) is obtained by using an empirical orthogonal function (EOF) analysis technique. (This is based on the fact that thermocline fluctuations in response to winds are a primary source of interannual variability of temperature throughout the upper ocean in the equatorial Pacific.) Using this empirical approach, SSTAs simulations can be significantly improved in the equatorial Pacific because the optimized \( T_e \) parameterization leads to a balanced depiction of the subsurface effect on SST variability by the mean upwelling of anomalous subsurface temperature and vertical mixing (Zhang et al. 2005). This approach has been successfully tested and implemented in an intermediate coupled model (ICM) with demonstrable improvements in its ENSO simulation and prediction in the tropical Pacific (Zhang et al. 2003).

In the context of ocean general circulation models (OGCMs), systematic biases are still a major challenge in simulations of both mean climatology and interannual variability. Generally, in OGCMs, the thermocline is too diffuse, that is, the vertical gradient of temperature is much weaker than observed (e.g., Rosati and Miyakoda 1988; Stockdale et al. 1998). While observed interannual changes in SST are largest over the eastern basin, those simulated from OGCMs are commonly centered in the central equatorial Pacific. Typical SST errors of OGCM simulations forced by observed atmospheric data are above 0.6°C in the Niño-3 region. Furthermore, OGCM-based coupled models commonly have weak interannual variability and especially underestimate SST variability in the eastern equatorial Pacific (e.g., Meehl et al. 2001). Although the observed ENSO system is characterized by interannual oscillations with a main period of 3–4 yr, many OGCM-based coupled models tend to produce shorter ENSO periods, favoring a quasi-biennial oscillation in particular (e.g., Barnett et al. 1993; Syu et al. 1995). As suggested by a number of modeling studies (e.g., Rosati and Miyakoda 1988; Syu et al. 1995; Meehl et al. 2001; Zhang and Zebiak 2004), these biases and systematic errors common in OGCM-based coupled models can be partially ascribed to the parameterizations of entrainment and vertical mixing in OGCMs; for purported impact of biological feedbacks see Murtugudde et al. (2002) and Rochford et al. (2001). For example, the weakened thermocline effects on SSTAs in the eastern equatorial Pacific can cause the surface horizontal advection to play a more dominant role in determining SST variability than the vertical processes, leading to the predominance of the westward propagation of SSTAs on the equator (Zhang and Zebiak 2004). Without strong enough thermocline feedback, SST anomalies in such coupled models are weak and short-lived over the eastern equatorial Pacific and along the coast of South America.

Substantial progress has been made in understanding and parameterizing subsurface vertical processes in OGCMs. Previously, numerous studies have focused on the determination of entrainment velocity across the
base of the mixed layer in layer OGCMs (e.g., Chen et al. 1994; Schopf and Loughe 1995; Wallcraft et al. 2003) and/or the coefficient of vertical diffusivity in level OGCMs (e.g., Large et al. 1997; Large and Gent 1999) for better simulations of upper-ocean temperature variability. Since the effects of subsurface vertical processes on SST associated with entrainment and vertical mixing also depend on subsurface temperature fields, an alternative way to improve SSTA simulations would be to focus on \( T_e \). As has been demonstrated (e.g., ZC87; Kiehlman 1993; Keenlyside 2001), \( T_e \) can have significant impact on SST and ENSO simulations in ocean and coupled models. In particular, an empirical procedure to optimize the \( T_e \) effect on SST variability leads to significantly improved SSTA simulation and prediction in intermediate ocean models (Zhang et al. 2003, 2005). However, direct application of this kind of empirical \( T_e \) parameterization scheme to improving SSTA simulations in an OGCM is quite difficult because of the complexity of an OGCM that attempts to simulate both mean climatology and interannual anomalies.

To take advantage of the empirical \( T_e \) parameterization for improving SSTA simulations of the tropical Pacific in the context of OGCMs, an embedding approach is tested and implemented in this paper: a separate SSTA submodel is explicitly embedded into a layer OGCM. The OGCM used is the Gent–Cane OGCM (Gent and Cane 1989; Murtaghude and Busalacchi 1998), which is an efficient layer model with an explicit bulk mixed layer. The SST variability in the embedded SSTA submodel is controlled by subsurface temperature anomalies associated with entrainment and vertical mixing, as well as ocean horizontal advection, whose fields are provided by the OGCM calculation. Although \( T_e \) anomalies that can be used in the SSTA submodel are also available directly from the OGCM (i.e., the temperature of the second model layer, \( T_2 \)), systematic errors are evident in the OGCM simulated subsurface temperature fields and thus can cause errors in SSTA simulations. (The possible sources of the systematic errors in the simulated subsurface temperature fields include the difficulty in representing the upper-layer thermal structure, upwelling and/or mixing processes.) Therefore, instead of using subsurface temperature directly from the OGCM, \( T_e \) anomalies for the embedded SSTA submodel are calculated from an empirical procedure in terms of SL anomalies simulated by the OGCM.

This embedded system (an OGCM plus an additional SSTA submodel) has potential advantages in improving SSTA and ENSO simulations in ocean and coupled models. Via the embedded SSTA submodel, only perturbation SST fields (relative to mean seasonal cycle) are explicitly calculated, with mean seasonal climatology being prescribed directly from observations (such as SST and the vertical temperature gradient), which allows realistic representation of anomalous SST tendencies associated with mean thermal structure. Another advantage of the embedded SSTA submodel is that with the anomaly formulation, physical parameterizations can be easily handled and SSTA simulations can be more easily controlled (tuned) to a regime that mimics nature. For example, via an empirical \( T_e \) parameterization, the amplitude and structure of SST variability in the equatorial Pacific can be much better controlled in the embedded SSTA submodel than in an OGCM. As a result, the SSTA simulations from the embedded submodel can be significantly improved compared to OGCM simulations. While this empirical approach to \( T_e \) may appear to be just a time- and space-varying bias correction, the real benefits are realized in a coupled atmosphere–ocean modeling context. The role of implementing this embedding in a coupled model should be fairly obvious, since the improved SSTA simulation from the embedded submodel will directly impact the atmospheric response and feedback to the oceanic component of the coupled system.

The paper is organized as follows. Section 2 describes briefly the models and some observational data used in this work. Section 3 examines the performance of SSTA simulations from a standard OGCM run. SSTA simulations from its embedded SSTA submodel are presented in section 4. Section 5 deals with some validation and sensitivity experiments of the proposed \( T_e \) parameterization scheme. The impacts on coupled atmosphere–ocean models are demonstrated in section 6. Conclusions are given in section 7. Further applications to a level OGCM is briefly presented in the appendix.

2. Model description, and observational and model-based data

We have developed an embedded model system (Fig. 1) consisting of an OGCM, an SSTA submodel, and an empirical \( T_e \) model, which can be coupled to an atmospheric wind stress anomaly model. These are briefly described in this section, as well as observational and model-based datasets used for constructing the empirical \( T_e \) and wind stress models and for verifying model simulations.

a. The OGCM

The ocean GCM used is based on a reduced gravity, primitive equation, sigma coordinate model of Gent and Cane (1989), which is developed specifically for studying the coupling between the dynamics and the thermodynamics of the upper ocean. The vertical struc-
ture of the model ocean consists of a mixed layer and a number of layers below specified according to a sigma coordinate. The mixed layer depth and the thickness of the last sigma layer are computed prognostically, and the remaining layers are computed diagnostically such that the ratio of each sigma layer to the total depth below the mixed layer is held to its prescribed value.

Several related efforts have improved this ocean model further. For example, Chen et al. (1994) developed and embedded a hybrid mixed layer model and studied the effects of vertical mixing, solar radiation, and wind stresses on the seasonal cycle of SSTs in the tropical Pacific. Murtugudde et al. (1996) coupled the OGCM to an advective atmospheric mixed layer (AML) model to estimate sea surface heat fluxes and showed the nonlocal effects of the atmospheric boundary layer on SST in all three tropical oceans. This heat flux parameterization allows a realistic representation of the feedbacks between mixed layer depths, SSTs, and the heat fluxes (e.g., Seager et al. 1995; Murtugudde et al. 1996). Complete hydrology has also been added to the model with freshwater flux treated as a natural boundary condition (Murtugudde and Busalacchi 1998).

More recently, the effect of penetrative radiation on the upper tropical ocean circulation has been taken into account, with attenuation depths derived from remotely sensed ocean color data (Murtugudde et al. 2002). As a result, the model can properly represent the subsurface heat source (biological heating) associated with the subsurface chlorophyll maximum that occurs in the cold tongue region during the boreal spring, resulting in a slight weakening of the stratification below the mixed layer and hence deeper mixed layers leading to a reduced surface divergence and hence reduced upwelling. This process improves simulations of not only the SSTs but also the upper water column stratification in wind-forced OGCM calculation, reducing the systematic cold bias in the central-eastern equatorial Pacific.

The model details can be found in Murtugudde et al. (1996), Murtugudde and Busalacchi (1998), and Murtugudde et al. (2002). The OGCM domain covers the tropical Pacific basin from 25°S to 25°N and from 124°E to 76°W, with horizontal resolution of 0.5° latitude by 1° longitude, and 31 layers in the vertical. Near the model’s southern and northern boundaries (poleward of 20°S/N), sponge layers are introduced, that is, a Newtonian term is added to the temperature and salinity equations, which damps the model solution back to observational temperature and salinity data from the World Database 2001 (see http://www.nodc.noaa.gov/OC5/WOA01/pr_woa01.html; Levitus et al. 2005). The OGCM, initiated from the Levitus temperature and salinity datasets, is integrated for 20 yr using climatological forcing fields. The mean atmospheric forcing data
for the OGCM spinup include wind stress from the European Centre for Medium-Range Weather Forecasts (ECMWF) averaged from 1985 to 1998 (e.g., Hackert et al. 2001), precipitation from Xie and Arkin (1995), solar radiation from the Earth Radiation Budget Experiment (ERBE; Harrison et al. 1993), and cloudiness from the International Satellite Cloud Climatology Project (ISCCP; Schiffer and Ros sow 1985), respectively.

b. An embedded SST anomaly submodel

Mathematically, the governing equation determining the evolution of interannual SST anomalies in a surface mixed layer can be written as (e.g., Keenlyside 2001; ZC87)

\[
\frac{\partial T}{\partial t} = -u' \frac{\partial T}{\partial x} + (\overline{u} + u') \frac{\partial T}{\partial x} - v' \frac{\partial T}{\partial y} + (\overline{v} + v') \frac{\partial T}{\partial y} \\
- [(\overline{w} + w')M(-\overline{w} - w') - wM(-\overline{w})] \frac{T_e - T}{H} \\
- (\overline{w} + w')M(-\overline{w} - w') \frac{(T_e - T)}{H} - \alpha_e T' \\
+ \frac{\kappa_b}{H} \nabla_h \cdot (H \nabla_h T') + \frac{2 \kappa_e}{H(H + H_2)}(T_e - T').
\]

Here, \( T' \) and \( T_e \) are anomalies of SST and the temperature of subsurface water entrained into the mixed layer, respectively; \( T \) and \( T_e \) are the prescribed seasonally varying mean SST and \( T_e \) from observations; \( \overline{u} \) and \( \overline{v} \) are the prescribed seasonally varying mean zonal and meridional currents in the mixed layer; and \( \overline{w} \) is the prescribed seasonally varying mean entrainment velocity at the base of mixed layer, which are all obtained from the OGCM; \( u' \), \( v' \), and \( w' \) are the corresponding anomaly fields; \( H \) is the depth of the mixed layer; \( H + H_2 \) is a constant (125 m); \( M(\delta) \) is the Heaviside step function [i.e., \( M(\delta) = 0 \) if \( \delta \) is positive and \( M(\delta) = 0 \) if \( \delta \) is negative]; \( \kappa_b \) (2.5 \times 10^7 \text{ m}^2 \text{s}^{-1}) meridionally and 2.5 \times 10^4 \text{ m}^2 \text{s}^{-1} \) zonally and \( \kappa_e \) (10^{-3} \text{ m}^2 \text{s}^{-1} \) are the horizontal and vertical diffusion coefficients; \( \alpha_e \) is the thermal damping coefficient. As expressed, the local rate of SST change (tendency) is controlled by horizontal advection, entrainment [(the \( M(\delta) \) terms), anomalous heat flux (\( -\alpha_e T' \)), horizontal diffusion, and vertical mixing, respectively. The surface heat flux is parameterized as being negatively proportional to the local SST (\( -\alpha_e T' \)) with \( \alpha_e = (100 \text{ day})^{-1} \) estimated from observations (e.g., ZC87). The surface mixed layer depth is prescribed from the Monterey and Levitus (1997) annual mean data (a fixed density criterion). Other mixed layer depth data with different criteria or different sources (e.g., Kara et al. 2003) could be used, which exhibit substantial differences particularly in the off-equatorial regions. Their sensitivity to SSTA simulations needs to be examined in the future.

Note that \( T_e \) is associated with two vertical processes: entrainment across the base of the mixed layer and mixing between the surface mixed layer and subsurface layer. The function \( M(\delta) \) accounts for the fact that SSTA's are affected by vertical advection only when subsurface water is entrained into the mixed layer; but SSTA's can always be influenced by subsurface temperature variability through vertical mixing (the last term).

c. An empirical \( T_e \) parameterization

An empirical relationship between \( T_e \) and SL anomalies has been developed to parameterize \( T_e \) for use in the embedded SSTA submodel (Zhang et al. 2005). In this subsection we briefly describe the scheme that is developed in two steps. First, an inverse modeling method is adopted to estimate \( T_e \) anomalies, since its basinwide geographic distribution and temporal evolution are not available directly from observations. Second, an EOF technique is adopted to calculate \( T_e \) anomalies in terms of SL anomalies that are calculated directly from the OGCM.

1) An inverse modeling approach to estimating \( T_e \)

In the SSTA equation [Eq. (1)], \( T_e \) is associated with two terms: entrainment by upwelling and the vertical mixing between the surface mixed layer and subsurface layer. Since the SST tendency on the left side can be estimated from observational data, it is possible to determine \( T_e \) anomalies by inverting the SSTA equation using observations and model simulations. As such, \( T_e \) anomalies calculated by this inverse method are, by definition, exactly those required by the SSTA submodel to perfectly simulate observed SSTA's.

2) An empirical relationship between \( T_e \) and SL variations

Observations and modeling studies indicate that variations in SL and the thermocline are well correlated over the equatorial Pacific (e.g., Wyrtki 1975; Wang and McPhaden 2000). As such, subsurface temperature anomalies associated with thermocline displacements are closely related with SL variability, which provides a statistical basis for inferring subsurface temperature anomalies from SL variability. Since \( T_e \) is not a directly observed quantity (but SL is), utilizing SL to parameterize \( T_e \) is extremely well suited to test the empirical parameterization in the context of the two-layer ap-
proximation that holds at low latitudes and the large
expanse of the tropical Pacific Ocean.

Thus, an empirical relation between \( T_e \) and SL
anomalies can be developed to estimate \( T_e \) anomalies in
terms of SL anomalies. Such statistical approaches have
been used successfully to construct wind stress anoma-
lies from a given SSTA in many tropical coupled atmo-
sphere–ocean models (e.g., Barnett et al. 1993; Syu et
al. 1995; Chang et al. 2001). To determine statistically
optimized empirical modes of interannual variability
between \( T_e \) and SL, an EOF analysis technique is
adopted to establish the relationship between SL and
\( T_e \) variations (e.g., Barnett et al. 1993).

More specifically, the procedure with the EOF-based
statistical scheme is as follows (Fig. 2). Monthly \( T_e \)
and SL anomaly data are first normalized by dividing their
spatially averaged standard deviation to form the vari-
ance matrix with which an EOF decomposition is made
into dominant spatial modes \((e_n, p_n)\) and the corre-
sponding time series (principal components; \( \alpha_n \) and \( \beta_n \)).
The latter are then used to obtain a matrix of regression
coefficients relating the two fields \((\gamma_{nm})\). Thus, a given
SL anomaly pattern can be converted into a \( T_e \) anomaly
using the derived spatial EOF modes \((e_m)\) and temporal
regression coefficients \((\alpha_m)\). Two EOF calculations are
possible. One is called the seasonally invariant version
(annual model): the EOF analysis is performed on all
time series data irrespective of season. Another is
called the seasonally varying version (monthly model): the
EOF analysis is performed separately for each cal-
endar month to construct seasonally dependent models,
giving 12 \( T_e \) models, one for each calendar month. In
this paper, the latter is adopted.

d. An empirical atmospheric wind stress (\( \tau \)) model

To examine the impacts of the embedding on the
performance of a coupled atmosphere–ocean model,
we have constructed an HCM consisting of the OGCM
and a simple atmospheric wind stress model. The atmo-
spheric model adopted in this work is statistical, spe-
cifically relating wind stress \((\tau)\) and SST anomaly fields.
The \( \tau \) model is constructed from an SVD of the co-
variance matrix that is calculated from time series of
monthly mean SST and \( \tau \) fields (e.g., Syu et al. 1995;
Chang et al. 2001). In this work, we perform a com-
bined SVD analysis of the covariance among anomalies
of SST and zonal and meridional wind stress compo-
nents.

e. The embedded system

An embedded ocean model system is developed in
which a separate SST anomaly submodel is explicitly
added to the OGCM. Figure 1 shows the embedded
system consisting of the OGCM, the SSTA submodel
with the empirical \( T_e \) parameterization. The embedding
is implemented as follows. At each time step, the
OGCM calculates SL and currents in the surface mixed
layer. The corresponding SL and current anomalies are
then obtained relative to its mean climatological fields
that are predetermined from the OGCM-only run

Fig. 2. A schematic showing a statistical method to estimate \( T_e \) anomalies from a given SL anomaly by an EOF analysis technique. In this work, simulated SL anomalies from the OGCM and estimated \( T_e \) anomalies from the inverse modeling of the SSTA equation are used in the analysis.
forced by observed atmospheric fields. Instead of using subsurface temperature directly from the OGCM, \( T_e \) anomalies for the embedded SSTA submodel are calculated (and corrected) using the empirical parameterization procedure in terms of the SL anomalies simulated from the OGCM. These interannual anomalies, together with prescribed mean climatological fields, are passed to the embedded SSTA submodel to calculate its own space–time evolution. The three components exchange anomaly information every day, with the \( T_e \) anomalies parameterized from the empirical scheme serving as the interface between the OGCM and the embedded SSTA model.

The SSTA submodel domain extends from 33.5°S to 33.5°N and from 124° to 78°E in the tropical Pacific with a realistic representation of continents. Because of computational limitations in performing EOF or SVD analyses for deriving empirical models, the horizontal grid of the embedded SSTA submodel, the \( T_e \) model, and the statistical atmospheric wind stress model is different from the high-resolution OGCM, which is chosen to have a 2° zonal grid spacing and a meridional grid stretching from 0.5° within 10° of the equator to 3° at the meridional northern and southern boundaries, respectively. With this resolution, the dimension of the matrix for the EOF (\( T_e \) model) and SVD (\( \tau \) model) analyses is about \( 84 \times 59 \times 408 \) (zonal and meridional grid points in the tropical Pacific with 34-yr temporal sampling from 1963 to 1996). The time step for the SSTA model integration is 4800 s.

f. Interannual anomaly data from observations and models

Various observational and model-based data are used for constructing the empirical relationships as well as for verifying model simulations. The observed SSTAs are from Reynolds et al. (2002). Interannual anomalies of wind stresses used to force the OGCM are from the National Centers for Environmental Prediction–National Center for Atmospheric Research (NCEP–NCAR) reanalysis (Kalnay et al. 1996). Interannual variations in solar radiation, cloudiness, and precipitation are not taken into account in the heat flux computation, while wind speed anomalies are estimated from the corresponding wind stress anomalies and used for heat flux calculations in the Seager et al. (1995) model.

For atmosphere–ocean coupling experiments, wind stress data used to construct the \( \tau \) model are the ensemble mean of a 24-member ECHAM 4.5 simulation during the period 1950–99, forced by observed SST anomalies (M. Tippett 2002, personal communication).

Using the ensemble mean data is an attempt to enhance the SST-forced signal by reducing atmospheric noise.

In addition, the monthly Simple Ocean Data Assimilation (SODA) reanalysis products of Carton et al. (2000a,b; see Web site at http://www.atmos.umd.edu/~ocean), covering the period 1950–99, are used to examine the relationship among anomalies of SST, \( T_e \), and sea surface height, and the accuracy of parameterized entrainment temperature. The ocean reanalysis uses all available temperature and salinity observations from the World Database 2001, satellite altimetry, and SST observations to constrain the GFDL OGCM Modular Ocean Model (MOM 2). We use the results from a moderate-resolution version (beta7), which has 2.0 × 2.5 latitude–longitude horizontal resolution in midlatitudes, reducing to 0.5 × 2.5 resolution in the Tropics, and 20 levels in the vertical. Figure 3 illustrates the anomalies of SST, \( T_e \), and sea surface height along the equator from the SODA reanalysis. Here, \( T_e \) is defined as the temperature at the base of the surface mixed layer whose depth is determined from Monterey and Levitus (1997) analysis.

3. The standard OGCM simulations

The OGCM is first forced by seasonally varying atmospheric climatological fields for 20 yr (climatology run); then it is forced by the anomalous wind stress from the NCEP–NCAR reanalysis for the period January 1961 to December 1999 (interannual run). From this interannual run, sea level and current anomalies can be obtained and will be used to construct an empirical \( T_e \) model. Since the OGCM performance has been examined extensively in simulating mean and interannual variability in the tropical Pacific Ocean (e.g., Murtugudde et al. 1996; Rothstein et al. 1998; Hackert et al. 2001), we only show some relevant results here.

Figure 4b illustrates the simulated SSTAs along the equator from the OGCM interannual run for year 1980 through year 1990. The model simulates interannual SST variability quite well in association with warm and cold ENSO events. Nevertheless, there are some clear differences in the amplitude and structure of simulated (Fig. 4b) and observed (Fig. 4a) SST anomalies. The OGCM tends to have weak variability in the eastern equatorial and coastal regions. This is in contrast to the corresponding observations where the largest variability is in the eastern basin.

Model performance is further evaluated in terms of anomaly correlation and root-mean-square (RMS) error between modeled and observed SST anomalies. To allow a direct comparison of our results with those in Zhang et al. (2005), we follow their analysis procedure: the same analysis periods are adopted and a 5-month
running mean of SSTAs is carried out before calculating correlation and RMS error. Figure 5 shows the correlation and RMS error between simulated and observed SSTAs for the period 1963–96. High correlation regions (Fig. 5a) are located in the central and eastern equatorial Pacific centered at 180°–120°W, but large errors in the SST simulation are evident in the eastern basin (Fig. 5b). In the Niño-3 region, the anomaly correlation and RMS error of the simulated SST anomalies are 0.74° and 0.58°C, respectively.

To illustrate the possible relationship between SST and subsurface temperature variability, we show in Fig. 4c the temperature anomalies of model layer 2 ($T_2$), which directly affect SST variability through vertical advection and mixing. As shown, variations in SST follow those in $T_2$ very closely, indicating that $T_2$ is a major factor in determining SST variability. This is further illustrated in Fig. 6a by the correlation between SST and $T_2$ anomalies calculated during the period 1963–96. High correlation exists between SST and $T_e$ anomalies in the entire equatorial Pacific, indicating that $T_e$ is a major source for SST variability in the tropical Pacific Ocean. This can be also confirmed from the SODA reanalysis as shown in Fig. 6b for the corresponding correlation.

To examine the possible contributions of systematic
errors in $T_2$ to those in SSTA simulations, $T_e$ anomalies can be estimated by inverting the SSTA equation [Eq. (1)] using observed SST from Reynolds et al. (2002) and the model mean and anomalous currents from the OGCM simulations. Figure 7a illustrates the $T_e$ anomalies along the equator obtained in this manner. Note that this inverse approach incorporates observations and models in a consistent way to allow the best estimate of $T_e$ possible; best in the sense that it is exactly the $T_e$ field required by the SSTA submodel to perfectly simulate the observed SST. As can be seen, the $T_e$ anomalies indeed reflect the observed SST variability (Fig. 4a) very well.

Clear differences exist between the subsurface temperature anomalies simulated directly from the OGCM (Fig. 4c) and those estimated from the inverse modeling (Fig. 7a). For example, the amplitude of $T_2$ directly from the OGCM is clearly underestimated over the central-eastern equatorial Pacific. As a result, the relative importance of subsurface temperature perturbations versus mixed layer current perturbations in determining SST variability will be underestimated. Thus, SST variability in the OGCM simulation is weak over the eastern equatorial Pacific (Fig. 4b). These characteristics suggest that the simulated $T_2$ anomaly fields and its effects on SST are deficient in the model, being
responsible for systematic errors in the SSTA simulations. Experiments with many tuning efforts in parameters and physics could not rectify this bias, nor significantly improve SSTA simulations. It should be noted that the deficiencies of the OGCM at underestimating SST variability in the eastern Pacific are typical of most OGCMs (e.g., Barnett et al. 1993; Meehl et al. 2001; Zhang and Zebiak 2004).

The simple diagnostic analysis above suggests that the long-known deficiency of some OGCMs in their parameterization of $T_e$ and its interactions with SST may contribute to unrealistic SST variability (amplitude and structure) in OGCMs. Apparently, improvement is needed in the OGCMs for better representing $T_e$ anomaly fields. The preferred solution would be to improve model physics in a comprehensive and consistent manner, but this has proven to be very difficult (e.g., Large and Gent 1999; Zhang and Zebiak 2004). Indeed, numerous previous studies have focused on the determination of entrainment velocity across the base of the mixed layer in layer OGCMs (e.g., Chen et al. 1994; Wallace et al. 2003) and/or the coefficient of vertical diffusivity in level OGCMs (e.g., Large et al. 1997; Large and Gent 1999; Meehl et al. 2001) for better simulations of upper-ocean temperature variability. Yet systematic errors in SSTA simulations are still quite large in the current state-of-the-art ocean models. Since the effects of subsurface vertical processes (entrainment and vertical mixing) on SST also depend on subsurface temperature fields, an alternative way to improve SSTA simulations would be to focus directly on $T_e$. Thus far, the role of $T_e$ in improving SSTA simulations has been emphasized in intermediate ocean models (e.g., ZC87; Zhang et al. 2003), but has not been done in the context of OGCMs.

4. The embedded OGCM simulations

To improve SSTA simulations, we embed into the OGCM a separate SSTA submodel with the empirical $T_e$ parameterization. The purpose of such a modeling exercise is to improve SSTA simulations through the embedded SSTA submodel, which can be accomplished by optimizing $T_e$ anomalies in terms of SL anomalies, as demonstrated in intermediate ocean models (e.g., Zhang et al. 2003, 2005).

a. The $T_e$ anomaly simulations

As described above, $T_e$ anomalies can be inferred through the SSTA equation [Eq. (1)] from observed SST fields and simulated anomaly current fields. This
The inverse modeling approach is conceived as a means for calculating the entrainment temperature in a balanced way, adjusting the various terms of heat budget in the surface mixed layer. Note that in the present study for the tropical Pacific, the surface heat flux is parameterized as having a damping effect on SST variability (i.e., $-\alpha T'$). Then, SL anomalies directly available from the OGCM are used to estimate $T_e$ anomalies by an EOF analysis technique. The EOF analysis is performed for the period 1963–96 (a total of 34 yr of data) using SL anomalies (Fig. 7b) from the interannual run forced by the NCEP–NCAR winds and $T_e$ anomalies estimated from the same simulation (Fig. 7a). Seasonally dependent relations between SL and $T_e$ anomalies are constructed for each calendar month. Based on the sequence of the variance and the amplitude of the SL-based reconstruction of $T_e$ anomalies, the first five EOF modes are retained in estimating $T_e$ anomaly fields.

Figure 7c shows the parameterized $T_e$ anomalies along the equator for the period 1980–90. Compared with the inverse modeling results (Fig. 7a), the estimated $T_e$ anomaly fields faithfully reproduce the large-scale interannual variability associated with warm and cold ENSO events. The amplitude of the reconstructed $T_e$ anomalies is comparable to the original field (Fig. 7a), indicating that the first five EOF modes are sufficient for recovering the realistic strength of the observed variability. However, the anomalies are somewhat smoothed and less noisy, indicating that the selected EOF modes effectively act as a low-pass filter.

Figure 6. Correlation for the period 1963–96 (a) between simulated SSTAs and temperature anomalies of model layer 2 ($T_2$) from the OGCM simulations, and (b) between SSTAs and $T_e$ anomalies estimated from the SODA reanalysis, respectively. The time series at each point have been smoothed by a 5-month running mean filter. The contour interval is 0.1.

A comparison can be made among the $T_2$ anomalies directly from the OGCM (Fig. 4c), $T_e$ anomalies empirically parameterized from sea level (Fig. 7c), and estimated from the SODA reanalysis (Fig. 3b), respectively. Although relatively weak, the amplitudes and structure of $T_2$ from the OGCM (Fig. 4c) and $T_e$ parameterized from the sea level (Fig. 7c) are comparable to the SODA reanalysis (Fig. 3b). The qualitative comparison is shown in Fig. 8 for their anomaly correlations calculated for the period 1963–96. Clearly, the parameterized $T_e$ anomalies from the SL anomalies are better correlated with those from the SODA reanalysis in the tropical Pacific.

b. SSTA simulations from the embedded submodel

Using the empirical $T_e$ parameterization scheme, SSTA simulations from the embedded SSTA submodel are now presented. Note that in this subsection we use the $T_e$ model that is trained during the period 1963–96 and is used for SSTA simulation during the same pe-
period. As such, the skill for SSTA simulations (e.g., as measured by the anomaly correlation) can be artificial, because of observational information of $T_e$ and SST variability covering the simulation period being already included in the training period. Various validation studies will be presented in the following section.

Figure 9 presents the longitude–time sections of simulated SSTAs along the equator from the embedded SSTA submodel for the 1980–90 period. Clear improvements can be seen in SSTA simulations compared with those from the standard OGCM (Fig. 4b). It is evident that the simulated SST interannual variability from the embedded SSTA submodel is in good agreement with the observations (Fig. 4a). In particular, the amplitude of the simulated SSTAs has increased over the eastern equatorial Pacific due to the optimized $T_e$ variability effect. Figure 10 demonstrates the anomaly correlation and RMS error calculated for the period 1963–96. When compared to the standard OGCM simulations (Fig. 5), a basinwide improvement is clearly evident in terms of both measures. It is quite striking that the correlation values exceeding 0.80 cover a significantly broader region over the central and eastern equatorial Pacific from 170°E to the coastal region of South America, with greatly reduced systematic SSTA errors over the central basin. In the Niño-3 region, the
anomaly correlation and RMS error of the simulated SST anomalies are 0.94° and 0.29°C, respectively.

5. Validation studies

The historical data-based $T_e$ parameterization scheme presented above is empirical in nature, and as inherent to any statistical method, its performance can be sensitive to a variety of factors, including the period used to construct the $T_e$ model (the training period), and the period used to test the model (the application period). It remains to be demonstrated if the approach is strongly sensitive to the data periods selected for developing the $T_e$ model and for applying it to SSTA simulations. In this section we will examine these sensitivity issues to evaluate the effectiveness and robustness of the scheme for improving SSTA simulations.

a. Cross-validation experiments

Cross-validation experiments are performed by dividing the total analysis period (1963–96), into two subperiods from 1963 to 1979 and from 1980 to 1996, from which the two $T_e^{63–79}$ and $T_e^{80–96}$ models are separately constructed and are then cross-tested to simulate SSTAs independently for the two subperiods (i.e., the $T_e^{63–79}$ model is used for the period 1980–96 and the $T_e^{80–96}$ model for the period 1963–79, respectively). Correlation and RMS error between simulated and observed SSTAs during the two cross-validation periods (both pairs of dependent and independent cases) have been calculated using the seasonally varying version of the EOF-based $T_e$ model. An experiment is termed dependent (independent) when the $T_e$ model is constructed using data covering (not covering) the same period. One expects that the performance will be relatively good and the simulation skill in the SSTA submodel will be relatively high if the application period overlaps the training period, but may not be so when the application period does not overlap the training period. Thus, the cross-validated results provide a more reliable estimate of skill improvement in SSTA simulations. Figure 9 shows examples of the simulated SSTAs along the equator from the embedded SSTA submodel using the dependent (Fig. 9a) and independent (Fig. 9b) $T_e$ models. The anomaly correlation and RMS error from the cross validation experiments for both periods are shown in Figs. 11–12, together with those from the standard OGCM simulations.
As expected, the correlation values for the two dependent periods (Figs. 11b,e) are higher than those for the two independent periods (Figs. 11c,f). Correspondingly, the RMS errors for the two dependent periods (Figs. 12b,e) are lower than those for the two independent periods (Figs. 12c,f). The performance for the two independent periods (Figs. 11c,f and 12c,f) is poorer in the far eastern equatorial Pacific, but over much of the central equatorial Pacific the correlation does not drop very much. Compared with those from the standard OGCM simulations (Figs. 11a,d), the correlation obtained from the two independent cases (Figs. 11c,f) is still measurably higher, with the values over 0.8 covering a broad area in the central and eastern equatorial basin. This is also reflected in the RMS errors (Fig. 12) with error reductions in the central basin. So, even with an independent $T_e$ model (Figs. 12c,f), SSTA errors exceeding 0.7°C are located in a very narrow region in the far eastern equatorial Pacific, smaller than those in the standard OGCM simulations (Figs. 12a,d).

Note that the cross-validation experiments we designed above are the strictest possible test of the $T_e$ parameterization scheme since totally independent historical data are used in constructing $T_e$ models for use in simulating SSTAs. Thus, the results obtained during the two independent periods present the worst cases in terms of SSTA simulation skill. However, given the fact that the real ocean states between these two periods

---

**Fig. 9.** Simulated SST anomalies along the equator during period 1980–90 from the embedded SSTA submodel using (a) the $T_e^{63–96}$ model, (b) the $T_e^{63–79}$ model, and (c) the $T_e$ model constructed with one-year-out data, respectively. The contour interval is 0.5°C.
experienced significant decadal changes in the tropical Pacific, including subsurface thermal structure and SST (e.g., Zhang and Levitus 1997), these results are encouraging. This suggests that the performance of the empirical $T_e$ parameterization is not unduly sensitive to the training period selected nor to the application period. Although the periods selected for cross validation are too short to produce stable statistics for the EOF analysis, the procedure appears to be quite successful in improving SSTA simulations using the embedded SSTA submodel with the empirical $T_e$ parameterization, as compared with the standard OGCM simulation.

b. One-dependent-year data excluding experiments

To prevent the information from entering into the application period, the $T_e$ model is constructed in such a way that one-year-dependent data are excluded in both $T_e$ and SL fields when performing the EOF analysis for constructing the $T_e$ model. For example, when year 1982 is being considered, the year 1982 data are taken out before performing the EOF analysis; that is, the EOF-based $T_e$ model is constructed only with data during the periods 1963–81 and 1983–96, respectively. Since no directly dependent data from the year have been entered into the analysis procedure through the historically constructed $T_e$ model, this can provide a reasonable estimate of simulation skill even though the year eliminated is not entirely independent of the preceding or following years.

Figure 9c shows the simulated SST anomalies along the equator from the embedded SSTA submodel using the $T_e$ model constructed with the one-year-out data experiment. The corresponding maps for the correlation and RMS error are shown in Fig. 13 for the period 1963–96, which can be compared to the corresponding results for the standard OGCM simulation (Fig. 5) and for dependent $T_e$ model results (Fig. 10), respectively. In the Niño-3 region, the anomaly correlation and RMS error of the simulated SST anomalies are 0.86° and 0.41°C, respectively. Furthermore, the results for the two subperiods 1963–79 and 1980–96 separately calculated are shown in Fig. 14. As expected, the correlation values lie between the dependent (Figs. 11b,e) and independent cases (Figs. 11c,f). The exclusion of one-dependent-year data that overlap the simulation year causes a reduction in skill but does not significantly degrade the simulations. In particular, the model is quite successful in simulating SSTA in the central equatorial Pacific. Clear improvements are evident when compared with those from the standard OGCM simulations (Figs. 5, 11a,d, and 12a,d, respectively).
6. Experiments with HCMs

Finally, we present some preliminary results from hybrid coupled atmosphere–ocean models of the tropical Pacific climate system. The layer OGCM with or without the embedded SSTA submodel is coupled to a statistical wind stress anomaly ($\tau_e$) model, which is derived from an SVD analysis of the covariance matrix that is calculated from time series of monthly mean SST and wind stress fields (e.g., Syu et al. 1995; Chang et al. 2001; Zhang and Zebiak 2004). The $\tau$ model is constructed from the ECHAM4.5 ensemble simulations and observed SST during the period 1963–96 (Zhang et al. 2003). Figure 1 shows the coupled atmosphere–ocean system. With standard coupling procedures (e.g., Barnett et al. 1993; Syu et al. 1995; Chang et al. 2001), the coupling between the atmospheric $\tau$ model and the OGCM is as follows: the OGCM provides SST anomalies (relative to its uncoupled mean climatology), which are used to calculate wind anomalies, which in turn are added to the observed mean seasonal climatology of wind stress to drive the OGCM. Furthermore, an embedded ocean–atmosphere coupling is implemented by explicitly adding into the layer OGCM the separate SSTA submodel [Eq. (1)] with an empirical $\tau_e$ parameterization. Note that with this embedded system, two SST anomaly fields are available: one directly from the embedded SSTA submodel, and the other from the OGCM (calculated as departures from its uncoupled mean climatology). Since SST anomalies from the

Fig. 11. Correlation between simulated and observed SST anomalies calculated during the periods (left) 1963–79 and (right) 1980–96, respectively. The simulated SST anomalies are from (a), (d) the OGCM and from the embedded SSTA submodel with the empirical $\tau_e$ parameterization (b), (e) dependently and (c), (f) independently constructed, respectively. The time series at each point have been smoothed by a 5-month running mean filter. The contour interval is 0.1.
former have more reasonable structure and amplitude over the central and eastern equatorial Pacific, they are used for determining wind stress anomalies via the statistical atmospheric \( \tau \) model.

We have then performed standard and embedded coupled atmosphere–ocean experiments using the identical OGCM that is coupled to the same SVD-based atmospheric \( \tau \) model, which are initiated with an imposed westerly wind anomaly for eight months. Evolution of anomalous conditions thereafter is determined solely by coupled interaction in the system. As examined previously by numerous studies (e.g., Barnett et al. 1993; Syu et al. 1995), coupled behavior depends on the so-called relative coupling coefficient \( \alpha \), that is, the wind stress anomalies from the atmospheric \( \tau \) model can be further multiplied by a scalar parameter before being added to the climatological wind stress fields to drive the OGCM. Several tuning experiments are performed with different values of \( \alpha \) to examine the coupled interannual variability. For the standard coupling, a larger \( \alpha \) (more than 1) is necessary to produce a sustained interannual variability, while with the embedded coupling, we find that a relatively large \( \alpha \) is not necessary for producing interannual oscillations. Figures 15–16 show examples of the simulated SST and its anomalies along the equator from the standard coupled run (taking \( \alpha = 1.3 \)) and the embedded coupled run (\( \alpha = 0.80 \)), respectively. For the embedded simulations, the SST anomalies are outputs directly from the embedded SSTA submodel, while for the standard simulations, they are calculated as departures from the uncoupled climatological simulation of the OGCM. Basically, the HCMs can produce interannual oscillations with a 4–5 yr period quite well. The overall time scale,
the variability structure, and coherent phase relationships among various atmosphere–ocean anomalies are consistent with corresponding observations (e.g., Zhang and Levitus 1997).

In the standard coupled run (Figs. 15a and 16a), several unrealistic features are evident. The zonal distributions of SST variability are not simulated very realistically. The amplitudes of SST anomalies are too weak in the eastern equatorial Pacific, but strong in the central and western basin. The excessive SST anomalies in the western equatorial Pacific are particularly notable. On the other hand, in the eastern Pacific, the weak SST anomalies seem to be associated with the systematic weak anomalies of $T_2$ in the uncoupled OGCM simulation (Fig. 4c). Clearly, the weakened SST anomalies from the OGCM will result in a weak thermocline feedback among anomalies of SST, thermocline, and wind stress, which will in turn cause an unrealistic variability pattern of SST in the equatorial Pacific. For example, during El Niño events, warm waters do not extend far enough to the east, with the 27°C isotherm of SST being located west of 140°W (Fig. 16a).

These systematic biases can also be found in other model simulations (e.g., Syu et al. 1995; Barnett et al. 1993; Latif et al. 2001; Zhang and Zebiak 2004), remaining a major challenge to short-term climate studies. In particular, OGCM-based coupled models often significantly underestimate SST variability in the eastern equatorial Pacific. As shown above, the embedding approach attempts to optimize the $T_e$ effects on SST variability in the equatorial Pacific, with the expectation that SSTA simulations in the embedded SSTA submodel can be improved through the empirical $T_e$ parameterization.

The embedding approach has a significant impact on the coupled variability of the tropical Pacific climate system, as shown in Fig. 15b and Fig. 16b. The amplitudes and positions of the maximum SST anomalies simulated directly from the embedded submodel (Fig. 15b) are now very realistic as compared to those from the standard coupling (Fig. 15a) and observations (e.g., Fig. 4a). The SST variability in the central-eastern basin is increased significantly, while the excessive cold SST anomalies disappear in the western Pacific. It seems that the enhanced SST anomalies in the eastern Pacific have improved the model depiction of the interactions among SST, thermocline, and wind stress in the equatorial Pacific in such a way that the enhanced SST variability leads to improved wind stress feedbacks that in turn reinforce the SST anomalies in the east. For example, at the mutual phase of El Niño, large warm SST anomalies (≈3°C) cover the whole eastern basin, with large wind anomalies being able to extend far to the east along the equator (figures not shown). Further-
more, although the OGCM itself has not been modified at all, the simulated ocean states from the OGCM in the embedded coupling are strikingly improved compared to those in the standard coupling. Examples are shown in Fig. 16b for the total SST fields from the output of the OGCM. Now, the seasonal cycle of SST has been enhanced in the eastern equatorial Pacific; the longitudinal displacements of the warm pool in the west and cold tongue in the east are more pronounced in association with ENSO events (Fig. 16b). During El Niño developments, the isotherm of 27°C can extend all the way to the coastal region (Fig. 16b). It is expected that the improved simulations of ocean states (i.e., the output of the OGCM) in the embedded coupled system (such as total SST and upwelling fields) can have significant impacts upon simulations of other physical and biogeochemical parameters. Detailed analyses of the standard and embedded coupled simulations will be presented elsewhere.

7. Discussion and conclusions

Standard OGCM simulations indicate that the simulated SST variability is underestimated in the eastern equatorial Pacific and along the coast of South America, but overestimated in the western-central basin. Analyses suggest that these can be partially associated with the problem of calculating subsurface entrainment temperature anomalies, due to the model deficiencies in the parameterization of entrainment and/or vertical mixing. Since $T_e$ is a major factor controlling interannual variability of SST in the equatorial Pacific, it can also be one of the major error sources for SSTA simulations and predictions in ocean and coupled ocean–atmosphere models.

To improve SSTA simulations, an empirical $T_e$ parameterization is proposed for use in an embedded SSTA submodel. The scheme is developed in two steps. First, an inverse modeling approach is adopted to estimate $T_e$ from the SSTA equation, using observations of SST fields and their tendencies, and mean and anomaly currents from the OGCM simulations. This approach combines the observations and the SSTA submodel in a consistent way to allow the best estimate of $T_e$ possible—best in the sense that it is exactly the $T_e$ field required by the SSTA submodel to perfectly simulate the observed SSTAs. Then, an EOF technique is used to build a relationship between SL and $T_e$ anomalies, allowing the major features of interannual variability associated with El Niño and La Niña events to be cap-
Fig. 15. The longitude–time sections along the equator of simulated SST anomalies from the (a) standard coupled run and (b) embedded coupled run, respectively. Note that the SST anomalies in (a) are a departure of the OGCM simulation relative to its uncoupled mean climatology, while those in (b) are directly from the embedded SSTA submodel; these are exactly the SST anomalies used to calculate wind stress anomalies in coupled simulations. The contour interval is 0.5°C.
tured. This relationship is justified, since there exists a good relationship between \( T_e \) and SL anomalies in the equatorial Pacific on interannual time scales. This procedure is able to optimize \( T_e \) anomaly fields, leading to a balanced depiction of various terms in the heat budget of the surface mixed layer.

To take advantage of the empirical \( T_e \) parameterization for improving SSTA simulations in the context of...
OGCMs, we developed an embedded system: a separate SSTA submodel with the empirical $T_e$ parameterization is explicitly embedded into a layer Gent–Cane OGCM. The governing equation of the embedded SSTA submodel describes the evolution of temperature anomalies in the surface mixed layer, driven by ocean horizontal and vertical advections associated with both mean and anomalous currents, which are calculated from the OGCM. This is intended to better capture the relation between thermocline variability and entrainment temperature in the embedded SSTA submodel through the empirical $T_e$ parameterization, a relation that is important for the simulation of interannual SST variability in the equatorial Pacific Ocean. Extensive experiments have been performed to demonstrate the robustness and effectiveness of the empirical $T_e$ parameterization to reduce model errors in SSTA simulations. By optimizing the subsurface effects on SST in the equatorial Pacific, various processes affecting SST variability can be balanced in the embedded SSTA submodel so that it produces good SSTA simulations, as compared to the standard OGCM simulations. In addition, since the embedded SSTA submodel is just one equation with various terms diagnostically estimated from the OGCM, it can be easily implemented in any ocean model without additional computational cost. For example, the computer CPU time is increased only by about 1% in the embedded simulation as compared to the standard simulation that does not include the embedded SSTA submodel. So, the additional overhead paid to add a separate SSTA submodel into an ocean model is totally negligible.

To date, various approaches have been introduced to improve model simulation and prediction of tropical Pacific SSTAs. For example, model output statistics (MOS) corrections have been tested and implemented in postprocessing model analyses (e.g., Barnett et al. 1993). In this regard, the empirical procedure to improve SSTA simulations we develop here can be viewed as an interactive correction approach, balancing the relative role of $T_e$ perturbations versus surface current perturbations in determining SST variability. But one important difference from the MOS correction is that the $T_e$ approach we developed here corrects the major error source for SSTA simulations in the embedded SSTA submodel, rather than the already produced SSTAs themselves.

The embedding approach taken in this study has not directly modified OGCM itself, but simply embedded into it an SSTA submodel with the optimized $T_e$ parameterization. While this approach may appear to be just a time- and space-varying statistical bias correction, the real benefits are realized in a coupled atmosphere–ocean modeling context where improved SSTA simulations can feed back to the ocean through atmospheric responses. Through the standard and embedded coupled atmosphere–ocean experiments with the identical OGCM that is coupled to the same empirical atmospheric wind stress anomaly model, we have demonstrated that the embedding has significant impacts on SSTA and El Niño simulations of the coupled tropical Pacific climate system. Not only do the simulated SSTAs directly from the embedded SST submodel exhibit more realistic variability in the equatorial Pacific (the amplitude and spatial structure), but also the mean ocean states from the output of the OGCM in the embedded coupling.

Further improvements and applications of the embedded coupled system are underway. In the present study, for instance, SL anomalies were taken from an OGCM simulation forced by observed atmospheric winds and used to construct the empirical $T_e$ model. However, model SL can have systematic errors due to erroneous atmospheric forcing and/or model physics. Using possibly erroneous SL information to construct the $T_e$ model can thus distort the $T_e$ parameterization and degrade SSTA simulations in the embedded SSTA submodel. This may partially explain the low skill of the SSTA simulations in the eastern equatorial Pacific. The advance of space-based observations (e.g., Busalacchi 1997) provides an unprecedented basinwide coverage of data over the ocean, with global sea level data in particular being currently available from satellite observations. In fact, we have already proposed a new approach to improved SSTA simulations by constructing an empirical $T_e$ model from Topex/Poseidon/Jason-1 (T/P/J) sea level data that is used in an intermediate ocean model (Zhang et al. 2004). It is demonstrated that using the $T_e$ model constructed from the observed sea level data allows observed information about dynamic ocean variability to be built directly into the empirical $T_e$ parameterization, having significant thermodynamic effect on SST in the equatorial Pacific. As a result, the use of T/P/J SL data leads to better SSTA simulations, with the largest coherent area of improvement in the eastern equatorial basin. This approach will be tested in the OGCM-based embedded system for further improvements in SSTA simulations.

Previously, the empirical $T_e$ parameterization has been successfully used to improve SSTA simulations in an intermediate ocean model (Keenlyside 2001; Zhang et al. 2003, 2005). In the present work, it is demonstrated that such improvements can be achieved in a layered OGCM also. Based on these and further demonstration in the appendix, we are confident that the embedding approach can be easily extended to other
types of ocean models, such as z-coordinate OGCMs (e.g., Zhang and Zebiak 2004), hybrid vertical coordinate OGCMs [e.g., Hybrid Coordinate Ocean Model (HYCOM); Shaji et al. 2005], and isopycnal ocean models (e.g., Schopf and Loughe 1995; Wallcraft et al. 2003). In addition, the intent of the present study is limited to improving SSTA simulation in the tropical Pacific where $T_e$ is a major forcing factor to SST variability in association with ENSO events. We believe that the embedding approach with some further improvements can be applicable to other tropical oceans where upwelling and vertical mixing play an important role in the SST variability. For example, since the heat flux in the tropical Atlantic Ocean plays an active role in the heat budget of the surface mixed layer (e.g., Chang et al. 2001), its effect needs to be parameterized realistically and taken into account in the inverse calculation of $T_e$, rather than the damping effect approximated for the tropical Pacific.

Finally, the demonstrated improvements of the embedded coupled system in the SST and El Niño simulations within the OGCM context are promising for improved El Niño prediction since this is a key region for seasonal-to-interannual climate forecasting. Experiments to explore this issue are being undertaken and will be addressed elsewhere.

Acknowledgments. We thank Drs. N. Keenlyside, S. E. Zebiak, and Dave DeWitt for their comments. The authors wish to thank anonymous reviewers for their numerous comments that helped to improve the original manuscript. Some model computations were performed at the International Research Institute for Climate and Society (IRI)/Columbia University. The SODA data are provided by Prof. J Carton, and the ECHAM ensemble mean wind stress data are provided by Dr. M. Tippett at IRI, both of whom are greatly appreciated. This research is supported by NASA Grant NCC5374 for funding of the Ocean Surface Topography Science Team (OST/ST), NASA Grant NAG512246 for funding of the Pacific Subtropical Cells, and NOAA PACS Grant NA17EC1483.

APPENDIX

Further Demonstration Using the GFDL OGCM (MOM 3)

As a further application for improving SSTA simulations in other types of OGCMs, we have tested the submodel approach in the National Oceanic and Atmospheric Administration/Geophysical Fluid Dynamics Laboratory (NOAA/GFDL) MOM 3 (Pacanowski and Griffies 1998), a comprehensive level OGCM. In this appendix, experiment results with the GFDL MOM 3 are briefly presented.

The MOM 3 domain in this application covers the entire Pacific basin from 55.5°S to 65°N, 107°E to 70°W with horizontal resolution of 1° latitude by 1° longitude (but 0.33° latitude between 10°S and 10°N). It has 40 vertical levels with a constant 10-m resolution in the upper 210 m. The model incorporates realistic continents and bottom topography; sponge conditions are adopted in the South Pacific poleward of 44.5°S, with model temperature and salinity restored to monthly climatology from the World Ocean Atlas 2001 (WOA2001). The OGCM is coupled with an advective atmospheric boundary layer model to estimate surface heat fluxes (Seager et al. 1995), allowing a realistic representation of the feedbacks between SST and the heat flux. The freshwater flux in the model includes two terms. One term is concerned with the differences in evaporation and precipitation ($E - P$). The second term is essentially a restoring boundary condition on sea surface salinity, by which the model top-level salinity is restored to the WOA2001 seasonally varying climatology. [Details of the model configuration and parameters can be found in Zhang et al. (2001).]

Datasets used for forcing the model and constructing empirical $T_e$ parameterization include observed SST data from Reynolds et al. (2002) and monthly atmospheric forcing data from the NCEP-NCAR reanalysis (Kalnay et al. 1996). The model, initiated from the WOA2001 temperature and salinity climatologies, is integrated for 40 yr with the NCEP-NCAR reanalysis climatological forcing fields (climatology run). The model is then integrated further with the reanalysis monthly forcing fields from January 1948 to April 2000 (interannual run). Monthly anomalies are obtained relative to corresponding model climatology that is derived from the interannual run for the period 1948–99.

Figure A1 shows the correlation and RMS error between simulated and observed SSTAs directly from the OGCM simulation for the period 1963–96. High correlation regions are located in the central and eastern equatorial Pacific. In the Niño-3 region, the anomaly correlation and RMS error of the simulated SST anomalies are 0.75° and 0.64°C, respectively. Large errors can be seen in the central basin around 150°W (Fig. A1b). The simulation skills are comparable to those from the layer OGCM simulations (Fig. 5).

We adopt the embedded submodel method to further improve SSTA simulations. A separate SSTA submodel with an empirical $T_e$ parameterization is embedded into the OGCM. As detailed above, the purpose of such a modeling exercise is to improve SSTA simula-
tions through the embedded SSTA submodel, which can be accomplished by optimizing the $T_e$ parameterization in terms of SL anomalies. The procedure to construct the empirical $T_e$ parameterization in the levelOGCM is the same as in the layer OGCM. From the interannual run, horizontal current anomalies needed in the embedded SSTA submodel are estimated by averaging the OGCM’s velocity fields within the mixed layer; vertical velocity anomalies are taken from those at the base of the mixed layer. The $T_e$ anomalies can then be estimated via the SSTA equation from observed SST fields and simulated anomaly and mean current fields from the OGCM. An empirical relationship between $T_e$ and SL anomalies can be correspondingly constructed by an EOF analysis for each calendar month. Thus, given an SL anomaly field directly calculated from the OGCM, $T_e$ anomalies can be parameterized and used for the embedded SSTA submodel.

Figure A2 illustrates the correlation and RMS errors for the period 1963–96 from the embedded SSTA submodel using the $T_e$ parameterization constructed with the one-year-out data experiment. Similar to the layer OGCM experiment (Fig. 13), the $T_e$ model is constructed in such a way that one-year-dependent data are excluded in both $T_e$ and SL fields when performing the EOF analysis. Since no directly dependent data from the year have been entered into the analysis procedure through the historically constructed $T_e$ model, this can provide a reasonable estimate of simulation skill.

Now, two SSTA results are available and can be compared with each other, one with the standard OGCM simulation, and another with embedded submodel simulation. The impact of the methodology on improving SST anomaly simulations can be clearly seen in Figures A1 and A2. As compared with the standard OGCM simulations, the SSTA simulation from the embedded submodel exhibits more realistic interannual variability in the tropical Pacific, with higher correlations and reduced errors in most regions. For example, in the Niño-3 region, the anomaly correlation and RMS of the simulated SST anomalies are 0.83° and 0.46°C, respectively. The level of the achieved improvement is comparable with that in the layer OGCM.

So, the submodel approach has been applied to two types of OGCMs with significantly improved SSTA simulations, one for the Gent–Cane ocean model and another for the GFDL MOM3. The former is a sigma-coordinate layer model with an explicit bulk mixed layer, while the latter is a z-coordinate level model with advanced K-profile parameterization (KPP) vertical mixing scheme (e.g., Large et al. 1997). These two mod-
els differ significantly in the vertical coordinate system and the parameterizations of the surface mixed layers. Using two different OGCMs can thus provide independent validations and help to identify possible limitations of the methodology. When dealing with subtle ocean modeling issues like the thermocline–mixed layer interactions and SSTA simulations, it is necessary to use differently formulated models in order for the results to be deemed robust.

REFERENCES


Kara, A. B., P. A. Rochford, and H. E. Hurlburt, 2000: An opti-

![Fig. A2. (a) Correlation and (b) RMS error for the period 1963–1996 between SST anomalies observed and simulated from the embedded SST submodel with the empirical $T_e$ parameterization constructed from one-year-out data experiment. The time series at each point have been smoothed by a five-month running mean filter. The contour interval is 0.1.](image-url)