JOURNAL OF GEOPHYSICAL RESEARCH, VOL. 112, XXXXXX, doi: 10.1029/2007JC004250, 2007

Performance of mixed layer models in simulating SST in the equatorial pacific ocean

A. B. Kara,[1] A. J. Wallcraft,[1] P. J. Martin,[1] and E. P. Chassignet[2]

Received 1 April 2007; revised 27 August 2007; accepted 8 October 2007; published XX Month 2007.

[1] This paper examines the ability of three ocean mixed layer submodels to depict inter-annual variations of sea surface temperature (SST) in a global configuration of the HYbrid Coordinate Ocean Model (HYCOM). The mixed layer submodels are (1) the K-Profile Parameterization (KPP), (2) the NASA Goddard Institute for Space Studies (GISS) turbulence closure, and (3) the Mellor-Yamada Level 2.5 (MY) turbulence closure. Accuracy of SSTs from the submodels is investigated during 1996–2001, which includes the onset of the strong 1998 La Niña event, when a record cold SST anomaly in the eastern equatorial Pacific occurred. The model simulations (with no ocean data assimilation or relaxation to SST climatology) reveal that all three submodels generally capture the westward extent of the SST cooling within the eastern equatorial Pacific during the transition period from the 1997 El Niño to the 1998 La Niña, one of the largest short term events ever observed (7°C change in SST from May to June 1998). During the six-month period after the transition, the daily SST from the submodels is $\approx$ 2°C warmer than the buoy SSTs obtained from the Tropical Atmosphere Ocean (TAO) array. Some of these biases are due to deficiencies in the net shortwave radiation and near-surface air temperature used for the simulations. Finally, comparisons with 166 yearlong daily SST time series from many buoys over various regions of the global ocean, including mostly equatorial Pacific, give median RMS differences of 0.65°, 0.70°, and 0.78°C for KPP, GISS, and MY, respectively, during 1996–2001.

Citation: Kara, A. B., A. J. Wallcraft, P. J. Martin, and E. P. Chassignet (2007), Performance of mixed layer models in simulating SST in the equatorial pacific ocean, *J. Geophys. Res.*, *112*, XXXXXX, doi:10.1029/2007JC004250.

1. Introduction and Motivation

[2] Sea surface temperature (SST) plays an important role in atmosphere-ocean interactions. Therefore an accurate determination of SST is essential for various types of applications over the global ocean [e.g., *Latif and Barnett*, 1996; *Schneider et al.*, 1999; *Bond et al.*, 2003], but especially over the tropical Pacific [e.g., *Cronin and McPhaden*, 1997; *Shinoda*, 2005]. This is true on short (e.g., daily and monthly) as well as on longer (e.g., inter-annual) timescales since climate patterns involve atmosphere-ocean feedbacks on all timescales [*Enfield and Mayer*, 1997; *Sutton and Allen*, 1997].

[3] A realistic mixed layer submodel (MLS) in ocean general circulation models (OGCMs) is a prerequisite in order to be able to depict realistic SST variations on a wide variety of temporal and spatial scales in the equatorial Pacific [e.g., *Swenson and Hansen*, 1999]. Several MLSs have become increasingly popular for use in OGCM studies because of their conceptual appeal and promising accuracy for the treatment of turbulent processes. These submodels are as follows: (1) the K-Profile Parameterization (KPP) model [*Large et al.*, 1997], (2) the NASA Goddard Institute for Space Studies (GISS) model [*Canuto et al.*, 2002], and (3) the Mellor-Yamada (MY) model [*Mellor and Yamada*, 1982]. Because of their extensive use in OGCMs, it is important to evaluate performance of each one globally as well as for specific regions and events.

[4] In this paper, the focus is on the tropical Pacific Ocean region and the strong 1997-98 El-Niño-Southern Oscillation (ENSO) event. We present qualitative and quantitative analysis of simulated SST using the above three commonly-used MLSs. More specifically, the model evaluation is performed using extensive sets of observational data sets, including daily time series of mooring buoy SSTs. The evaluation of the MLSs is based on the accuracy of the model to reproduce the SST variability on various timescales (from daily to inter-annual) without any assimilation of oceanic temperature. In addition to the evaluation of MLSs in predicting SST at locations where the SST is dominated by local forcing and vertical mixing, we examine the performances of the MLSs during strong events in the tropical Pacific such as the marked shift in the equatorial Pacific Ocean SST anomalies that occurs between the warm (El Niño) and cold (La Niña) phases of ENSO [*McPhaden*,

[1] Oceanography Division, Naval Research Laboratory, Stennis Space Center, Mississippi, USA.

[2] Center for Ocean-Atmospheric Prediction Studies, and Department of Oceanography, Florida State University, Tallahassee, Florida, USA.

1999]. Given that both phases have considerable impact on the global climate [*Hendon*, 2003] and result in potential socio-economic damages [*Elsner and Kara*, 1999], reliable determination of SST from dynamical models is of importance during an ENSO event.

[5] The strong 1997-98 ENSO event is of particular interest, since it developed very rapidly, with a record high SST drop (≈7°C) occurring in the eastern equatorial Pacific [*Harrison and Vecchi*, 2001]. Thus simulation of the SST evolution during this event presents an excellent test case for numerical models. *Barnston et al.* [1999] presented inter-comparisons of 8 dynamical (coupled atmosphere-ocean) and 7 statistical models and showed that none of them was able to properly forecast the extent of the El Niño. They concluded that significant progress and evaluation were needed to better represent ENSO events. Thus for the model-data comparisons of SSTs obtained from MLSs, we make extensive use of buoy time series. While our main focus is to examine MLS performance using buoy measurements, SSTs from a satellite-based product and an archived numerical weather prediction model are also used for global validation.

[6] This paper is organized as follows. The OGCM and the MLSs used in this study are introduced in section 2, The statistical metrics used for evaluating the SSTs from the MLSs are then described in section 3. The overall global performance of the MLSs over the 1996-2001 time frame is discussed in section 4. The model's ability in simulating the 1997 El Niño and 1998 La Niña events, including the transition period is investigated in section 5. The impact of wind errors on the modeled SSTs is discussed in section 6. Finally, the results are summarized in the concluding section.

2. The Ocean Model

2.1. HYCOM General Features

[7] The OGCM used in this study is the HYbrid Coordinate Ocean Model (HYCOM) [*Bleck*, 2002, Chassignet et al., 2003]. HYCOM contains five prognostic equations: two for the horizontal velocity components, a mass continuity or layer thickness tendency equation, and two conservation equations for the thermodynamic variables, which can either be salt and potential temperature or salt and potential density. The model behaves like a conventional σ (terrain-following) model in very shallow oceanic regions, like a z-level (fixed-depth) coordinate model in the mixed layer or other unstratified regions, and like an isopycnic-coordinate model in stratified regions [*Chassignet et al.*, 2003, 2006]. HYCOM uses the layered continuity equation to make a dynamically smooth transition to z-levels in the unstratified surface mixed layer and to σ-layers in shallow water. The optimal coordinate is chosen every time step using a hybrid coordinate generator [Halliwell, 2004] with further improvements [*Kara et al.*, 2005a]. The model automatically generates the lighter isopycnal layers that are often needed for the pycnocline when the ocean mixed layer is very shallow, as it commonly occurs in the eastern equatorial Pacific [e.g., *Kara et al.*, 2003].

2.2. HYCOM Global Configuration

[8] The model grid used in this study spans the global ocean from 78°S to 90°N. The grid is a 0.72° equatorial resolution Mercator grid between 78°S-47°N with a bipolar Arctic patch north of 47°N, i.e., a tripole grid [*Murray*, 1996]. The average zonal (longitudinal) resolution for the 0.72° global grid varies from ≈80 km at the equator to ≈60 km at midlatitudes (e.g., at 40°N). The meridional (latitudinal) grid resolution is doubled to 0.36° near the equator to better resolve the equatorial wave-guide and is halved in the Antarctic for computational efficiency. Hereinafter, the model resolution will be referred to as 0.72° for simplicity. The model's land-sea boundary is at the 50-m isobath (with a closed Bering Strait) so it never uses a terrain-following vertical coordinate. The bottom topography was constructed from the NRL Digital Bathymetry Database (DBDB2) bathymetry database, which has a resolution of 2-min and is available online at http://www7320.nrlssc.navy.mil/DBDB2_WWW/.

[9] There are 26 hybrid layers in the vertical in the model. The target density values for the isopycnals and the decreasing change in density with depth between isopycnal coordinate surfaces are based on the 1/4° Generalized Digital Environmental Model (GDEM) climatology [*NAVOCEANO*, 2003]. The density difference values were chosen, so that the layers tend to become thicker with increasing depth, with the lowest abyssal layer being the thickest. The minimum thickness of the top layer is 3 m, and this minimum increases 1.125× per layer up to a maximum at 12 m, and target densities are chosen such that at least the top four layers are always in z-level coordinates.

2.3. Mixed Layer Submodels

[10] Three of the MLSs available in HYCOM are based on solving for Laplacian vertical diffusion over the full water column with a variable diffusion coefficient (K). Among these, KPP is a level 1 turbulence closure, which parameterizes the influence of a large suite of physical processes. GISS is a level 2 turbulence closure, which includes both large- and small-scale vertical shear. MY is a level 2.5 turbulence closure, which is an improvement over the MY level 2 closure [*Smith and Hess*, 1993], since the former includes the advection and diffusion of turbulent kinetic energy.

2.4. Atmospheric Forcing

[11] We use the atmospheric forcing data from the European Centre for Medium-Range Weather Forecasts (ECMWF) 40-year Re-Analyses (ERA-40) [*Kållberg et al.*, 2004] for climatological simulations and operational ECMWF data sets [*Gibson et al.*, 1999] for inter-annual simulations. The atmospheric forcing includes wind stress at the sea surface, wind speed at 10 m above the sea surface and scalar fields (net shortwave and longwave radiation fluxes at the sea surface, air temperature and air mixing ratio at 10 m above the sea surface). The components of the surface heat flux, the net longwave and latent and sensible fluxes, were computed with bulk formulations using the model SST and the input ECMWF air temperature and mixing ratio at 10 m above the sea surface [*Kara et al.*, 2005b]. The evaporation was derived from the computed latent heat flux.

[12] For the model spin-up, the years 1979–2002 from ERA-40 are averaged to form a climatological monthly mean atmospheric forcing. The years prior to 1979 were not used in the average since there were not many data used in

the assimilation of the ERA-40 Re-Analyses. 6-hourly submonthly wind anomalies from operational ECMWF over the period September 1994 to September 1995 are then added to the 12 monthly averages. Choosing another time period for the 6-hourly wind anomalies (other than 1994-95) did not have any significant impact on the model SST.

[13] There is no explicit relaxation of the HYCOM SST. However, including air temperature from ECMWF in the formulations for latent and sensible heat flux automatically provides a physically realistic tendency toward the correct SST. There is a relaxation of the HYCOM sea surface salinity (SSS) to the monthly climatology of the Polar Science Center Hydrographic Climatology (PHC). The PHC climatology was chosen for its accuracy in the Arctic region [*Steele et al.*, 2001]. The SSS relaxation has a constant coefficient of relaxation. The actual e-folding time depends on the mixed layer depth (MLD), expressed as 30 days $\times$ 30 m/MLD, i.e., it is more rapid when the MLD is shallow. Here, MLD is in meters. A relaxation of the SSS is necessary to prevent long-term drift, and it is in addition to the evaporation and precipitation surface fluxes [e.g., *Kara et al.*, 2005c].

[14] Additional forcing parameters read into the model are monthly mean climatologies of satellite-based attenuation coefficient for Photosynthetically Active Radiation (k_{PAR} in 1/m) and river discharge values. The shortwave radiation at depth is calculated using a spatially varying monthly k_{PAR} climatology [e.g., *Kara et al.*, 2005a]. Thus using ocean color data, the effects of water turbidity are included in the model simulations through the attenuation depth ($1/k_{\mathrm{PAR}}$) for the shortwave radiation [*Kara et al.*, 2004]. The rate of heating/cooling of model layers in the upper ocean is obtained from the net heat flux absorbed from the sea surface down to a depth, including water turbidity effects. The model also treats rivers as a runoff addition to the surface precipitation field.

2.5. Initialization and Spinup

[15] The simulations were initialized from the monthly mean temperature and salinity for August from the GDEM climatology. Model simulations are performed for each MLSs, i.e., of KPP, GISS and MY, respectively. Each model simulation is first spun up for 8 a (statistical equilibrium is reached in $\approx$5 years) using the ERA-40 climatological, monthly mean thermal atmospheric forcing from with 6-h wind forcing as described in the previous section. A linear regression analysis was performed for domain-averaged quantities (temperature, salinity, potential and kinetic energy, etc.) to determine the statistical equilibrium in each model layer, which is expressed numerically as % change per decade. The model simulations were deemed to be in statistical equilibrium when the rate of potential energy change was acceptably small (e.g., <1% in 5 years) in all layers. After the 8-year spin up, the HYCOM simulations were extended inter-annually from 1995 to 2001 using the 6 hourly wind/thermal surface forcing from the ECMWF operational data set introduced in section 2.4.

3. Validation Data and Statistical Metrics

[16] Satellite SSTs as well as daily SST time series from buoys will be used to evaluate the modeled HYCOM SST obtained from simulations using the KPP, GISS, and MY MLSs, respectively. Our goal is to provide quantitative model-data comparisons of SST for each MLS. The statistical metrics used for comparing the SST time series from the models and observations are mean error (ME), root-mean-square error (RMS), correlation coefficient (R) and non-dimensional skill score (SS). Let X_i (i = 1, 2,…,n) be the set of n observations (reference), and let Y_i (i = 1, 2,…,n) be the set of corresponding model estimates. Also let $\overline{X}$ ($\overline{Y}$) and σ_X (σ_Y) be the mean and standard deviations of the reference (estimate) values, respectively. Following *Murphy* [1995], the preceding statistical measures can be expressed as follows:

$$\mathrm{ME} = \overline{Y} - \overline{X}, \tag{1}$$

$$\mathrm{RMS} = \left[\frac{1}{n}\sum_{i=1}^{n}(Y_i - X_i)^2\right]^{1/2}, \tag{2}$$

$$R = \frac{1}{n}\sum_{i=1}^{n}\left(X_i - \overline{X}\right)\left(Y_i - \overline{Y}\right)/(\sigma_X\sigma_Y), \tag{3}$$

$$\mathrm{SS} = 1 - \mathrm{RMS}^2/\sigma_X^2 \tag{4}$$

ME is the mean difference between the HYCOM and observed values over the time series. The RMS error can be considered as an absolute measure of the difference between the observed and modeled time series and a useful absolute measure of the accuracy of the model hindcasts. The R value is a measure of the degree of linear association between the observed and modeled time series. SS takes both RMS and σ_X into account, thereby providing a normalization when the SST standard deviation is quite different at two different locations. Values of SS range from 1.0 for the best result to negative values for the worst.

4. Evaluation of Climatological SST Over the Global Ocean

[17] We first examine if the MLSs are able to reproduce the monthly mean climatological SST over the global ocean when using the monthly climatological ERA-40 atmospheric forcing introduced earlier. For that purpose, monthly climatological mean HYCOM SSTs were formed from the SSTs of the last four years of the spin up (i.e., model years 5 to 8) of the climatologically forced simulations. The 4-year averaging period was considered sufficient given the climatological atmospheric forcing (no inter-annual variability) and lack of eddy activity. The modeled SSTs are then compared to the NOAA monthly SST climatology [*Reynolds et al.*, 2002] which is derived using an optimal interpolation (OI) of in situ and satellite SSTs from 1971 to 2000. The horizontal resolution of the NOAA climatology (1° $\times$ 1°) is close to that of HYCOM (0.72° $\times$ 0.72° cos (lat)) and was interpolated to the global HYCOM grid for comparisons with the modeled SSTs.

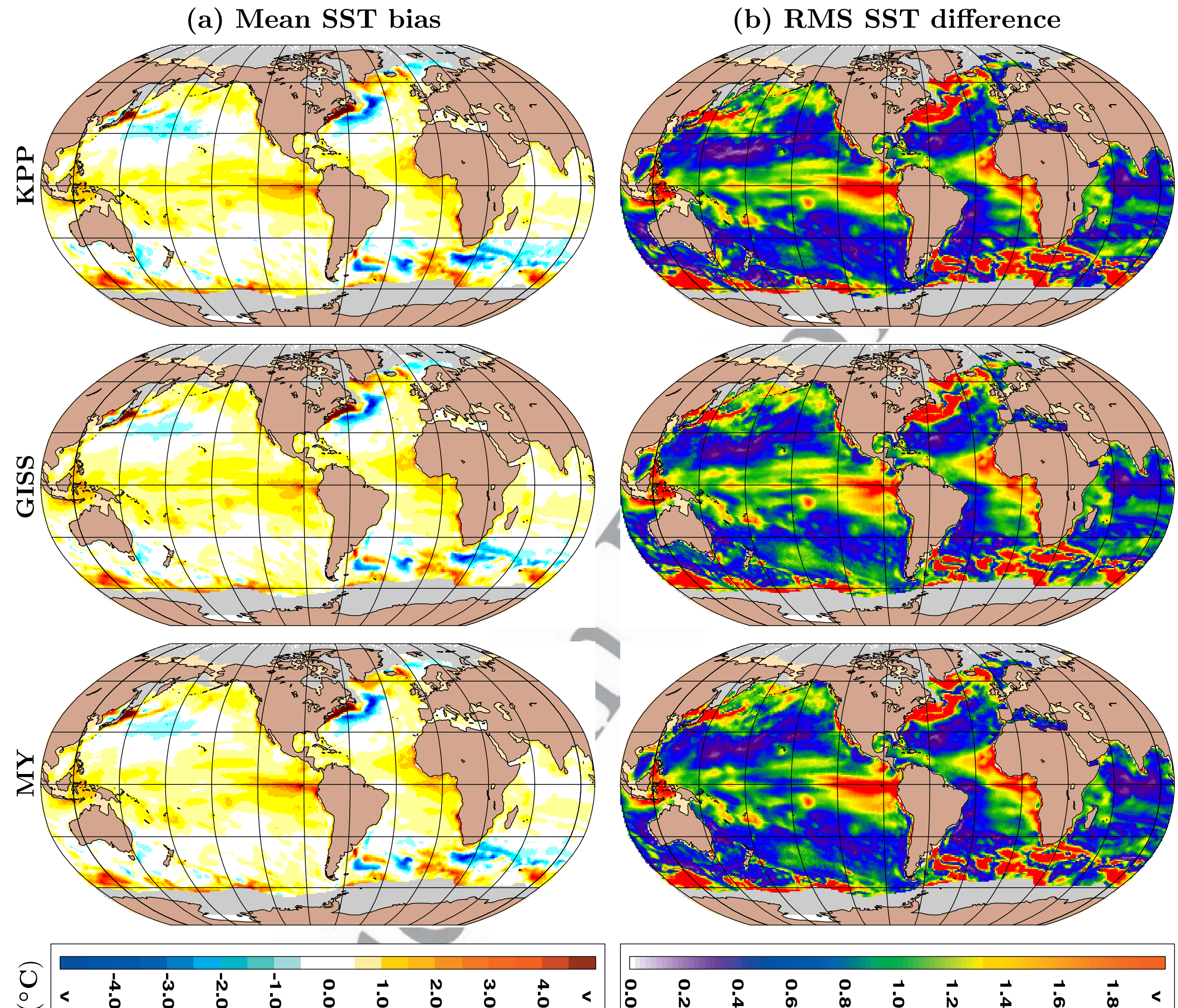

Figure 1. Validations of monthly mean SSTs obtained from climatologically-forced HYCOM simulations with those from the NOAA SST climatology. Comparisons are performed when SSTs from HYCOM simulations are obtained from three mixed layer submodels, separately. Both mean bias and RMS SST differences are calculated over the seasonal cycle at each grid point over the global ocean. Global average of annual SST mean error is 0.2°, 0.3°, and 0.1°C when using KPP, GISS and MY in HYCOM, respectively. The global RMS difference is 0.7C for all cases. Performing a 1-year 0.72° global HYCOM simulation requires ≈15 wall-clock hours on 64 HP/Compaq SC45 processors. The overall model run time is approximately the same with KPP and GISS, but is 1.5 times longer with MY (primarily because of its additional prognostic fields).

[18] We use the statistical metrics introduced in section 3 with n = 12 (from January through December) in the time series comparisons. In other words, we let X_i (i = 1, 2,. . .,12) be the set of monthly mean NOAA reference (observed) SST values from January to December, and Y_i (i = 1, 2,. . .,12) be the set of corresponding HYCOM estimates at a model grid point. Statistical values over the seasonal cycle were then calculated. The resulting ME and RMS fields clearly indicate that all MLSs result in similar errors over the global ocean. Mean SST bias with respect to the NOAA climatology is typically within ±0.5° (Figure 1a). However, there are relatively large errors in the regions where the strong Kuroshio and Gulf Stream current systems are located. These current systems are not well resolved in the coarse resolution (0.72°) version of HYCOM, as used in this paper, resulting in such errors. Similar to mean bias, the corresponding RMS SST differences calculated over the seasonal cycle are also similar when using either of the MLSs (Figure 1b), giving a globally-averaged RMS value of ≈ 0.7°C for KPP, GISS and MY.

[19] Correlation and skill score fields are also computed for HYCOM versus NOAA SSTs over the global ocean (not shown). The global average of correlation is high with a value of 0.88 for all MLSs. This shows that all MLSs can reproduce SST seasonal cycle accurately. Similarly, overall MLS success in simulating SST is evident from the rela-

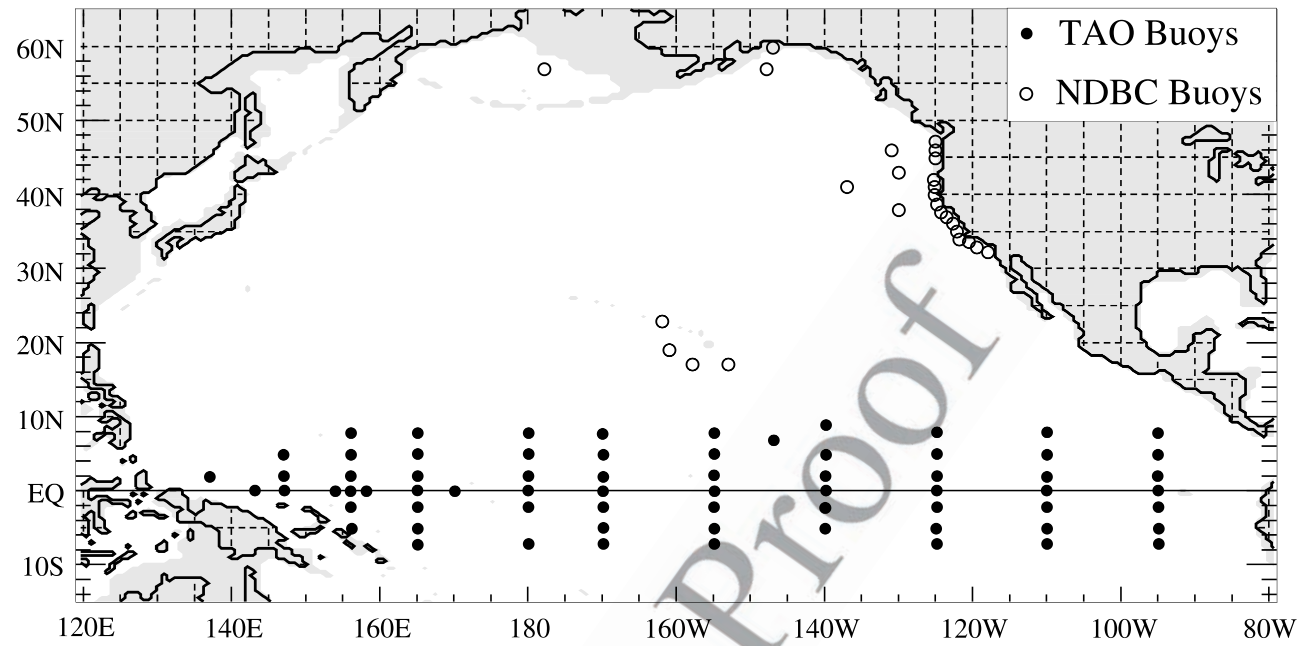

Figure 2. TAO array and NDBC buoys used for validating ocean model SSTs used in this paper. The TAO buoys are located in the equatorial Pacific Ocean. The NDBC buoys provide sampling outside the equatorial Ocean and are located off the coasts of the continental U.S., Hawaii and Alaska.

tively large and positive skill scores whose global averages are typically identical to each other, 0.41 for KPP, 0.37 for GISS and 0.38 for MY.

5. SST Variability in the Pacific Ocean During 1996–2001

[20] The climatologically-forced simulations were extended inter-annually from 1995 to 2001 using the 6-hourly wind/thermal surface forcing from the ECMWF operational data set (section 2.4) in order to evaluate the MLSs in locations where the SST is dominated by local forcing and vertical mixing. The first year of the simulations (i.e., 1995) is considered to be a spin up period and only the modeled SSTs from 1996 through 2001 will be considered for the analyses. All the simulations are identical, except for the MLS used.

[21] Observed SST time series from the Tropical Atmosphere Ocean (TAO) array [*McPhaden*, 1999] located in the equatorial Pacific Ocean. In addition, we use National Data Buoy Center (NDBC) buoys located offshore of various parts of the U.S. coast, including Hawaii and Alaska (Figure 2) are used for the model-data comparisons. The latter are available from the National Oceanographic Data Center (NODC).

[22] The positions of moored buoys can change over the course of a few days to a week, depending on the current regime, by up to $\approx$ 3 km. This is the typical diameter of the watch circle within which the buoys move. Since each mooring moves in time and space about its deployment location, we calculated the average position based on the historical latitude and longitude data for each buoy. The modeled SSTs were then extracted at these locations from each HYCOM simulation using KPP, GISS, and MY simulations. For ease of notation, the nearest integer values of the average latitude and longitude for buoy locations are used in the text. One challenge is how best to compare intermittent time series of different lengths and covering different time intervals, while allowing inter-annual comparison of the verification statistics at the same location and comparison of statistics at different locations over the same time interval.

[23] In the following subsections, we seek answers to the following questions: (1) do KPP, GISS, and MY exhibit average large differences during the 1996–2001 time period?, (2) how do KPP, GISS, and MY compare in simulating the SST during the impressive 1997 El Niño onset, i.e., how does the relative performance of the parameterizations apply to the warming phase (i.e., El Niño)?, and (3) how is the performance of KPP, GISS, and MY in simulating the monthly and daily SST during the transition period from the 1997 El Niño to the 1998 La Niña?

5.1. Evaluation Over the 1996–2001 Period

[24] The first question is addressed by performing model-data comparisons at all the available TAO and NDBC buoys locations from 1996 through 2001. Two examples of SST time series at two NDBC locations in 2001 are shown in Figure 3. All three MLSs are able to reproduce well the daily SST variations, including its seasonal variations. For example, in comparison to buoy SSTs, RMS differences at (23°N, 162°W) are almost identical with values of 0.32°, 0.37° and 0.30°C for KPP, GISS and MY simulations, respectively. Similarly, RMS SST differences are 0.86°, 1.01° and 0.82°C at (41°N, 137°W). Statistical values are also calculated between daily modeled and observed SSTs at all buoy locations in 2001. They are provided for the TAO buoys in Table 1 and in Table 2 for the NDBC buoys. As described in section 3, there are n = 365 values in the comparisons.

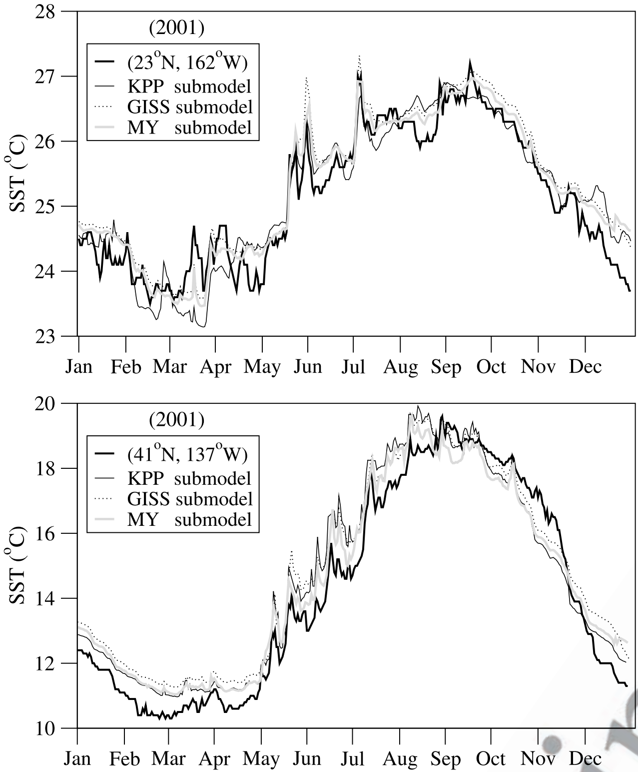

Figure 3. Comparisons of daily SST time series from HYCOM simulations using three mixed layer submodels at two NDBC buoy locations in 2001.

[25] Similar statistical calculations, as in Tables 1 and 2, were then repeated at all available buoys for the other years from 1996 through 2001. Distribution of the RMS SST differences and R values based on the number of buoys is shown in Figure 4. These class intervals for statistical metrics are determined by combining the values for each yearly statistic. There is a total of 166 yearlong daily SST buoy time series from all the TAO and NDBC buoys during 1996–2001. Most of the RMS SST differences with respect to the buoy SST are <1°C when considering all locations. For example, there are 45, 37, and 32 buoys where the SST RMS differences between submodels and buoys are ≥0.4°C but <0.6°C for KPP, GISS, and MY, respectively (Figure 4a).

[26] Similarly, there are 42, 40, and 36 buoys where the SST RMS differences are ≥0.6°C but <0.8°C. Comparisons with 166 yearlong daily SST buoy time series indicate that KPP and GISS give median RMS differences of 0.65°C and 0.70°C, respectively, and MY gives median RMS difference of 0.78°C (Table 3). Obviously, these median RMS differences can be considered identical to each other since a 0.08°C RMS difference value between KPP and MY is negligible. Median R values are identical for all three submodels, with a value of 0.93. All the submodels are also able to simulate the SST seasonal cycle well. This is evident from R values >0.9 at most of the locations (Figure 4b). There are 65 (KPP), 69 (GISS), and 64 (MY) buoys out of 166 where correlations are >0.95, corresponding to 39%, 42%, and 39% of all buoys. Thus the SST seasonal cycle for nearly half of the yearlong time series from the TAO and NDBC buoys are predicted very accurately by the MLSs.

Table 1. Statistical Verification of Daily SST Simulated by HYCOM[a]

TAO Buoy	Model	RMS	ME	σ_{BUOY}	σ_{HYCOM}	R	SS
(2°S, 125°W)	KPP	0.70	−0.19	1.32	1.26	0.86	0.72
	GISS	0.83	−0.34	1.32	1.37	0.84	0.61
	MY	0.73	−0.20	1.32	1.15	0.85	0.70
(5°N, 155°W)	KPP	0.59	−0.26	0.67	0.67	0.68	0.22
	GISS	0.58	−0.21	0.67	0.56	0.63	0.25
	MY	0.89	−0.61	0.67	0.71	0.56	−0.72
(5°S, 110°W)	KPP	0.64	0.37	1.35	1.22	0.92	0.78
	GISS	0.45	0.15	1.35	1.28	0.95	0.89
	MY	0.61	0.44	1.35	1.14	0.96	0.80
(5°S, 140°W)	KPP	0.40	−0.16	0.85	0.75	0.90	0.77
	GISS	0.42	−0.15	0.85	0.79	0.89	0.75
	MY	0.48	−0.25	0.85	0.77	0.88	0.69
(5°S, 155°W)	KPP	0.38	−0.13	0.57	0.31	0.84	0.56
	GISS	0.39	−0.15	0.57	0.27	0.87	0.54
	MY	0.39	−0.21	0.57	0.32	0.87	0.53
(8°N, 170°W)	KPP	0.58	−0.38	0.72	0.39	0.84	0.34
	GISS	0.52	−0.30	0.72	0.43	0.84	0.48
	MY	0.59	−0.40	0.72	0.47	0.83	0.34
(8°S, 110°W)	KPP	0.62	0.36	1.40	0.98	0.97	0.80
	GISS	0.52	0.35	1.40	1.12	0.98	0.86
	MY	0.64	0.31	1.40	0.92	0.97	0.80
(8°S, 155°W)	KPP	0.34	−0.11	0.40	0.25	0.59	0.27
	GISS	0.33	−0.10	0.40	0.25	0.63	0.33
	MY	0.35	−0.11	0.40	0.25	0.59	0.27

[a]HYCOM uses KPP, GISS and MY mixed layer submodels, and the validation is performed with respect to SSTs from TAO buoys in the equatorial Pacific Ocean. Comparisons are made based on daily time series (n = 365 days) in 2001. There is no assimilation of any oceanic data including buoy SSTs and no relaxation to any SST climatology in model simulations. The nondimensional skill score takes bias into account, something not done by the correlation coefficient. In the table σ refers to standard deviation of daily time series over a year. See section 3 for a detailed explanation of the statistical metrics and their calculations.

[27] In addition to the RMS and R values, we also calculated median statistics for other statistical metrics to provide a comprehensive summary for the MLSs in predicting daily SSTs. Median biases based on 166 yearlong buoy SST time series are almost zero with values of −0.01, 0.03, and −0.03°C for KPP, GISS, and MY, respectively. This indicates that all the submodels had no problem predicting the annual mean SST. Similarly, the median SST standard

Table 2. Same as Table 1 but for NDBC Buoys

NDBC Buoy	Model	RMS	ME	σ_{BUOY}	σ_{HYCOM}	R	SS
(17°N, 158°W)	KPP	0.26	0.17	0.79	0.82	0.97	0.89
	GISS	0.30	0.21	0.79	0.85	0.97	0.86
	MY	0.27	0.21	0.79	0.80	0.98	0.89
(23°N, 162°W)	KPP	0.32	0.20	1.05	1.01	0.97	0.91
	GISS	0.37	0.24	1.05	1.06	0.97	0.88
	MY	0.30	0.14	1.05	1.07	0.97	0.92
(41°N, 137°W)	KPP	0.86	0.58	3.11	2.65	0.99	0.92
	GISS	1.01	0.81	3.11	2.81	0.98	0.90
	MY	0.82	0.52	3.11	2.80	0.98	0.93
(43°N, 130°W)	KPP	1.07	0.93	2.68	2.70	0.98	0.84
	GISS	1.38	1.19	2.68	2.79	0.97	0.73
	MY	1.20	1.02	2.68	2.77	0.97	0.80
(46°N, 131°W)	KPP	1.01	0.69	2.68	2.74	0.96	0.86
	GISS	1.26	0.90	2.68	2.95	0.96	0.78
	MY	1.15	0.86	2.68	2.83	0.96	0.82
(56°N, 148°W)	KPP	0.56	0.02	3.39	3.58	0.99	0.97
	GISS	0.87	0.01	3.39	3.93	0.98	0.93
	MY	0.76	0.14	3.39	3.66	0.98	0.95

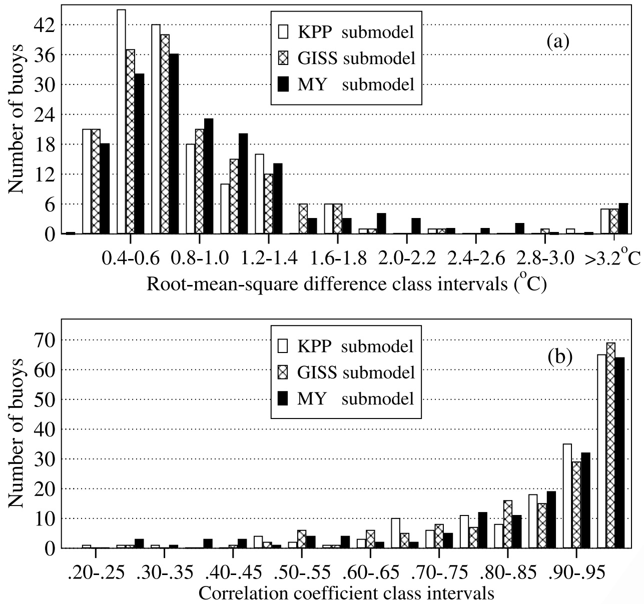

Figure 4. Number of buoys for class intervals of each statistical metric based on daily SST comparisons between HYCOM and buoy SST (both TAO and NDBC) from 1996 through 2001. Results are based on 161 yearlong SST time series.

deviation values of 1.22°C (KPP), 1.24°C (GISS), and 1.15°C (MY) are very close to that of 1.35°C (buoys). Finally, the median SST skill values of 0.72, 0.71, and 0.66 clearly confirm the success of these submodels in simulating SST. There are 23, 21, and 29 yearlong SST time series during 1996–2001 (out of 166) for KPP, GISS, and MY, respectively, for which HYCOM gave negative skill scores. Overall, this indicates the model failure rate at ≈ 13%, 14%, and 17% of all buoys since positive skill is considered as acceptable HYCOM SST simulation for a given buoy (see section 3).

5.2. Daily SST Comparisons During 1997

[28] We address the second question by analyzing SST time series during 1997 when the El Niño was starting in April. The presence of unusually warm water observed in the Pacific Ocean during the El Niño phase is clearly evident from the TAO buoy at (0°N, 140°W). *Picaut et al.* [2002] explained that the westerly wind bursts excited equatorial downwelling Kelvin waves and advected the eastern edge of the warm pool eastward. This resulted in a distinct warming over the central and eastern parts of the equatorial basin. All the MLSs are able to simulate this warming in SST well as evident from comparisons against the buoy SST time series at five different locations at the eastern and central equatorial Pacific (Figure 5). The biases in the SST in comparison to the buoy time series are generally very small (<0.5°C), and this is true at all the buoy locations (Table 4). All the MLSs tend to however give a cold bias. In general, skill of the simulated SST is always positive, which demonstrates HYCOM's ability to simulate SST over the seasonal cycle regardless of which submodel is used. These results clearly show that all three MLSs in HYCOM are able to simulate the SST warming starting in April (i.e., at the beginning of the onset of the 1997 La Niña) and at later time periods with high accuracies.

[29] The performance of the MLSs is also tested at a NDBC buoy located at (57°N, 178°W) outside the equatorial ocean in the same year (Figure 6). There are not any significant differences among the models at this location, either, and the SSTs from all the submodels agree with those from the buoy reasonably well. The MLSs are typically warmer than observed by the NDBC buoy (by ≈ 1°C) from July to September, but they are almost identical before and after this period. This suggests that the bias in HYCOM during this period is not due to the particular MLS used. Known biases in the radiation and wind fields are possible sources for the errors (see section 6 for a discussion of wind errors).

Table 3. Median SST Error Statistics for HYCOM MLSs[a]

HYCOM	RMS, (°C)	ME, (°C)	R	SS	Std. dev. σ_{BUOY}	(°C) σ_{HYCOM}
KPP	0.65	−0.01	0.93	0.72	1.35	1.22
GISS	0.70	0.03	0.93	0.71	1.35	1.24
MY	0.78	−0.03	0.93	0.66	1.35	1.15

[a]Median values are calculated based on 166 yearlong daily time series from TAO and NDBC buoys over the time frame 1996–2001.

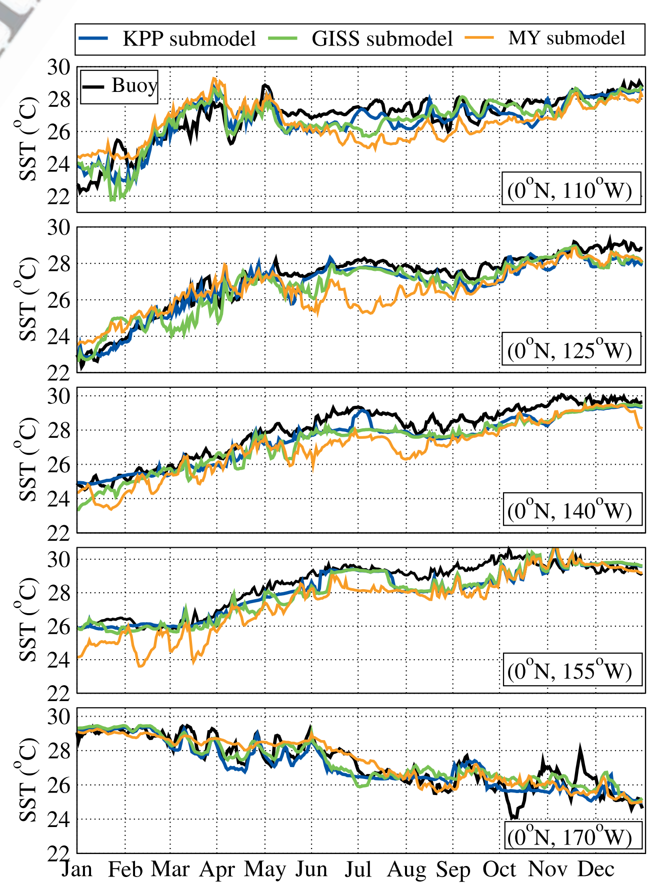

Figure 5. Daily SST time series from TAO buoys and those obtained from HYCOM simulations performed with three mixed layer submodels in 1997 when the El Niño started in April. HYCOM is forced using ECMWF wind and thermal forcing. No data, including SST, are assimilated by HYCOM for the simulations.

Table 4. Validation of Daily HYCOM SST in the Equatorial Pacific in 1997[a]

TAO buoy	Model	RMS	ME	σ_{BUOY}	σ_{HYCOM}	R	SS
(0°N, 110°W)	KPP	0.79	−0.23	1.45	1.42	0.86	0.70
	GISS	0.89	−0.24	1.45	1.47	0.83	0.63
	MY	1.22	−0.39	1.45	1.18	0.63	0.30
(0°N, 125°W)	KPP	0.70	−0.36	1.64	1.48	0.93	0.82
	GISS	0.57	−0.34	1.64	1.49	0.96	0.88
	MY	1.19	−0.60	1.64	1.30	0.78	0.48
(0°N, 140°W)	KPP	0.65	−0.52	1.57	1.38	0.97	0.83
	GISS	0.79	−0.67	1.57	1.51	0.96	0.75
	MY	1.10	−0.96	1.57	1.57	0.94	0.51
(0°N, 155°W)	KPP	0.62	−0.35	1.44	1.41	0.94	0.82
	GISS	0.70	−0.43	1.44	1.48	0.93	0.77
	MY	1.16	−0.93	1.44	1.78	0.93	0.36
(0°N, 170°W)	KPP	0.58	−0.17	0.89	1.13	0.88	0.58
	GISS	0.82	−0.37	0.89	1.37	0.88	0.15
	MY	0.86	−0.46	0.89	1.39	0.86	0.07

[a]Statistical metrics are calculated using daily SST time series (n = 365 days) from available TAO buoy in 1997 when the strong El Niño event was starting.

5.3. SST Variability During the ENSO Transition Period

[30] In this section, the performance of KPP, GISS, and MY in simulating the monthly and daily SST is investigated during the transition period from the 1997 El Niño to the 1998 La Niña. We first determine the time period when the 1997 El Niño was transitioning to the 1998 La Niña. There is no universally accepted definition of the warm and cold phases of ENSO events [e.g., *Hanley et al.*, 2003]. For the present analysis, ENSO phases are classified according to several indices. The use of various indices is motivated by concerns about possible discrepancies that may exist in the data sources used for creating the indices, and we would like to have an independent assessment of the transition period duration from El Niño to La Niña.

[31] The ENSO indices analyzed here are based on regional SST anomalies (Figure 7a) and surface atmospheric pressure (Figure 7b). Detailed definitions of these indices are discussed by *Trenberth* [1997]. The SST-based indices are defined using area-mean SSTs within different regions of the equatorial Pacific. The pressure-based indices are the conventional Southern Oscillation Index (SOI), which is the difference between the normalized Darwin and Tahiti SLP anomaly time series, the eastern equatorial Pacific SOI index (SOI east), and the equatorial SOI index (SOI equation) as well. Also included in Figure 7a is the Multivariate ENSO Index (MEI) calculated over the tropical Pacific [*Wolter and Timlin*, 1998].

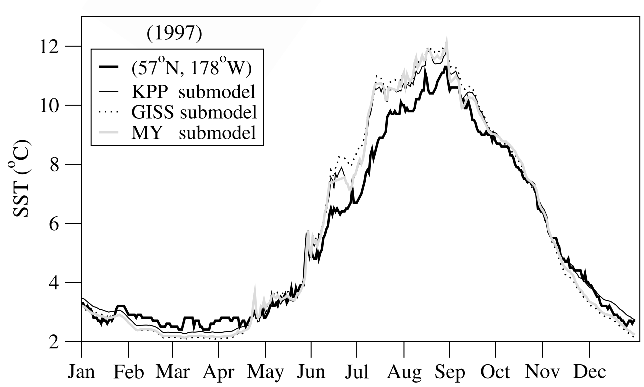

Figure 6. Daily SST time series from a NDBC buoy at (57°N, 178°W) and HYCOM simulations using three mixed layer submodels in 1997.

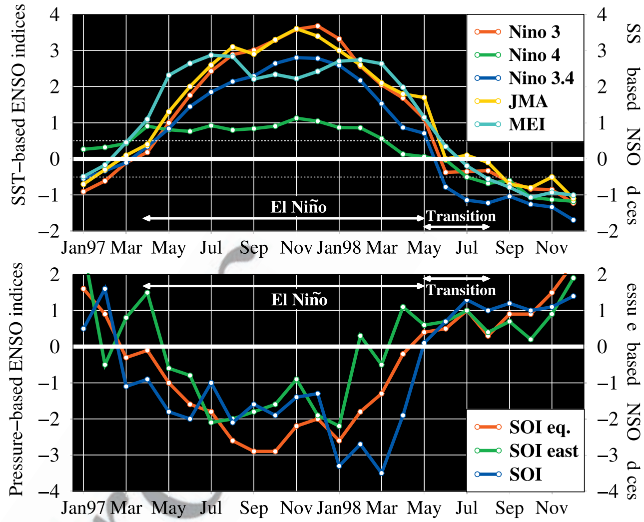

Figure 7. Time series of monthly mean ENSO indices from the beginning of 1997 to the end of 1998: (a) SST-based indices (°C), including the standardized MEI values, and (b) standardized surface atmospheric pressure-based indices. Monthly mean time series for the indices are obtained from National Oceanic Atmospheric Administration (NOAA) Climate Diagnostics Center (http://www.cdc.noaa.gov/ClimateIndices/List/). The original MEI time series is bi-monthly, and they were interpolated to monthly means to be consistent with other indices. The Niño 1 + 2 index, which represents the extreme eastern tropical Pacific SST (0°–10°S, 90°W–80°W) is not shown because we would like to examine the westward extent of SST cooling in the central equatorial Pacific.

[32] The methods and threshold values for identifying the occurrence of a warm or cold phase vary for each index. For example, the Niño 3 index (bounded by the region 5°N-5°S and 90°W–150°W) uses monthly SST anomalies based on a 5-month running mean, and the threshold value is ±0.5°C (i.e., ≥0.5°C for the warm phase and −0.5°C for the cold phase). Similarly, periods of negative (positive) SOI values coincide with typical warm (cold) phases. Our primary interest is to detect the timing and duration of the transition between El Niño to La Niña based on the sign change of ENSO indices.

[33] As shown by the white arrows in Figure 7, all the indices generally agree that the El Niño had transitioned to La Niña by early summer 1998. This is the transition period (May, June, July, and August 1998) that is considered in the OGCM analysis. Daily and monthly mean SST obtained from the model simulations are then evaluated against observations during the ENSO transition period. The SSTs are first evaluated for the May to October 1998 period (Figure 8) which covers the end of the 1997 El Niño through the development of the 1998 La Niña (Figure 7).

[34] Monthly mean SSTs from the MLSs are first formed from the daily values before being compared to the obser-

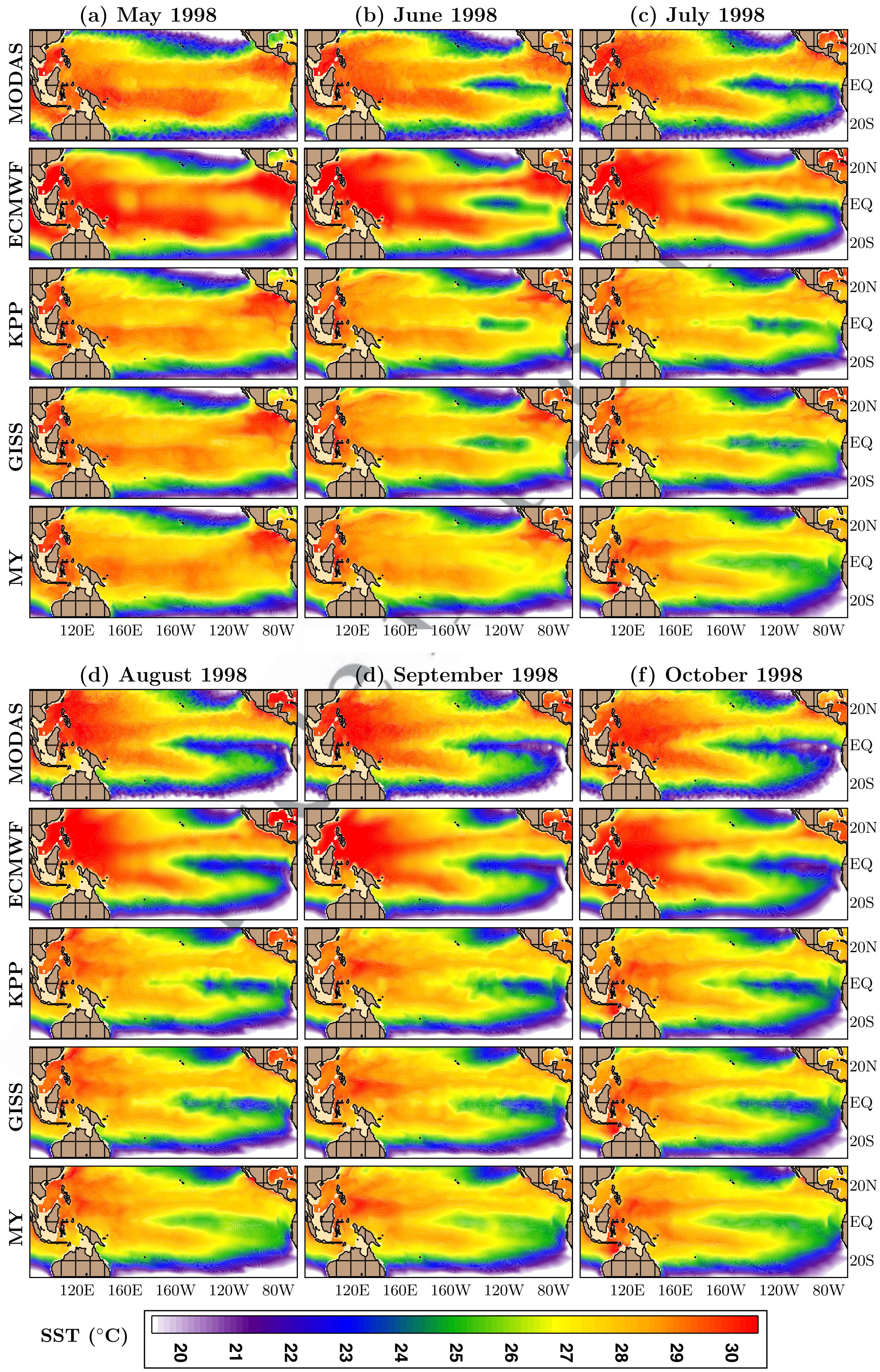

Figure 8. Monthly mean SST as constructed from MODAS and ECMWF and HYCOM simulations using the KPP, GISS and MY from May through October, 1998.

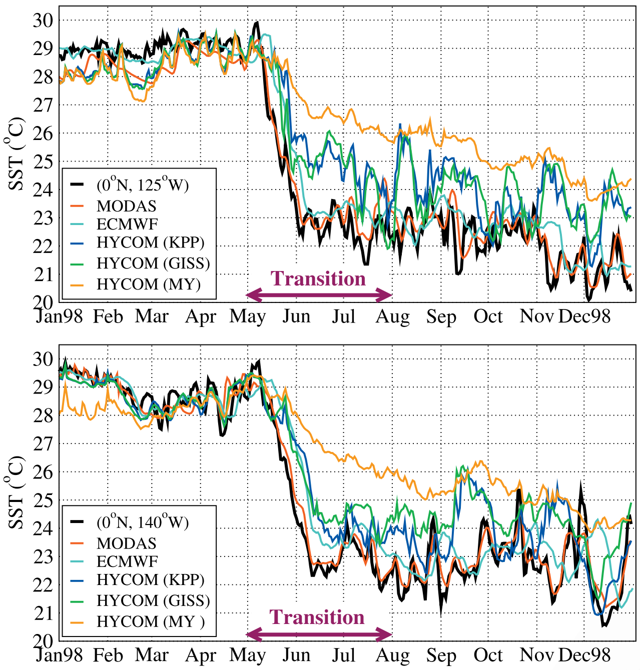

Figure 9. Daily averaged SST and air temperature at 3 m above the sea surface from two Tropical Atmosphere Ocean (TAO) buoys at (0°N, 125°W) and (0°N, 140°W) in 1998. Air temperatures from ECMWF operational analyses at 2 m used in HYCOM simulations are also included. The *x* axis is labeled starting from the beginning of each month.

vations (Figure 8). The observational SST fields are monthly averages of the daily Modular Ocean Data Analysis System (MODAS) SST re-analyses [*Barron and Kara*, 2006]. The original MODAS SST fields are on a 1/8° global global grid, and were interpolated to the HYCOM grid for these comparisons Each daily MODAS SST is produced by an optimal interpolation of Advanced Very-High Resolution Radiometer (AVHRR) Multi-Channel SST (MCSST) data.

[35] As evident from Figure 8, the MODAS SST drops substantially (by ≈7°C) in the eastern equatorial Pacific in only one month (from May to June), while such a drop in SST appears in the simulations only when KPP and GISS are used. In June, the cold tongue of the SST cooling in all three simulations has spread from 80°W to 160°W, a pattern that is consistent with MODAS. Cooling of the MODAS SST continues gradually, even dropping below 20°C in the eastern equatorial Pacific during August-October. This cooling is generally evident in the simulations using KPP and GISS, but with a warm SST bias of <2°C. SSTs from ECMWF also agree with those from MODAS. In general, all three mixed layer models reproduce the areal extent of the SST cooling reasonably well during the strong transition period. Unlike KPP and GISS, however, MY usually yields warmer SSTs in the cold tongue.

[36] Model-data comparisons are also performed on short (daily) timescales to further evaluate the performance of KPP, GISS, and MY in simulating SST (Figure 9). Daily averaged buoy SSTs obtained from the TAO array (http://www.pmel.noaa.gov/tao/data_deliv/) at two locations, (0°N, 125°W) and (0°N, 140°W), clearly indicate SST cooling as large as ≈8°C, occurring from May to June when the rapid phase of the 1998 transition from El Niño to La Niña is in progress. Undetected by the monthly mean SST analysis (Figure 8), the daily time series show that the MLSs lag in producing the rapid drop in SST from May to June, although all performed well before the transition started. In particular, SSTs from KPP and GISS are ≈ 2°C and from MY ≈ 4°C warmer than those from the TAO buoys during May–June, results that are investigated further in section 4.

[37] Another striking feature of the daily SST is that while MODAS captures the large SST drop during May–June, ECMWF has a phase lag during the same period. The satellite SST is available at least daily, except for cloud cover, and the MODAS re-analyses include satellite data from 1 day before to 1 day after the analysis time, thereby capturing the large SST drop of ≈ 8°C. The SST from ECMWF is produced using a 7-day running mean analysis window, which for a real-time system inevitably gives a lag of ≈ 10 days (Tim Stockdale of ECMWF, personal communication) even though (unlike MODAS) the TAO SST measurements are assimilated. The phase lag is also evident from HYCOM simulations using all MLSs. This is because air temperature from ECMWF is also lagging the SST as noted at both buoy locations (Figure 10).

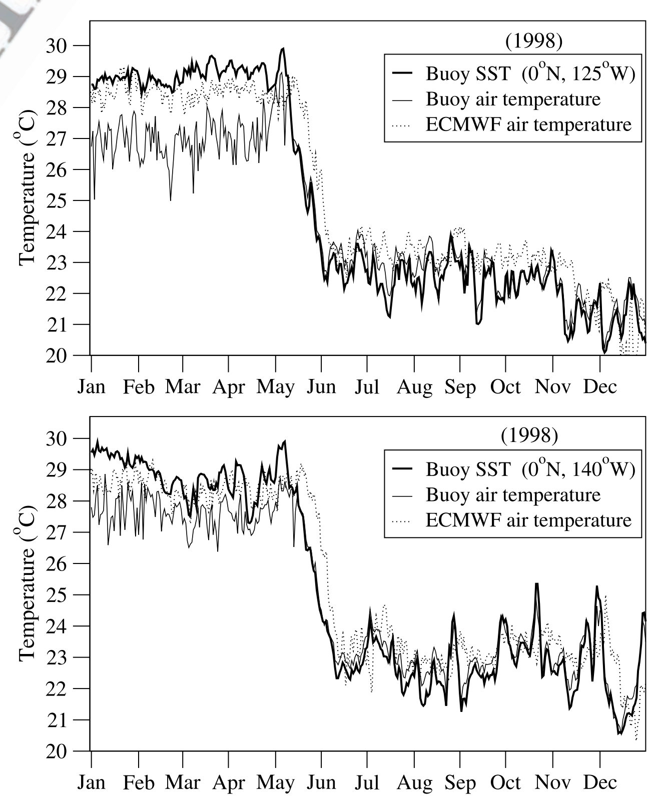

Figure 10. Daily-averaged SST from two Tropical Atmosphere Ocean (TAO) buoys at (0°N, 125°W) and (0°N, 140°W) in the eastern equatorial Pacific. Also included is the daily-averaged SST from the ECMWF operational product, the MODAS SST and the 0.72° resolution global HYCOM using KPP, GISS and MY mixed layer submodels. The *x*-axis is labeled starting from the beginning of each month, and the 1998 transition period from El Niño to La Niña is marked.

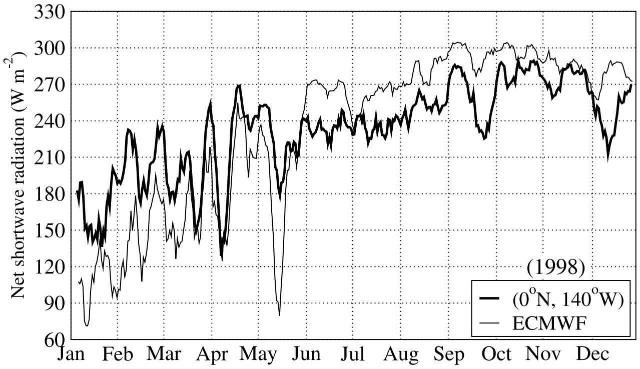

Figure 11. Time series of shortwave radiation entering the sea surface from ECMWF and the TAO buoy at (0°N 140°W). Note that the TAO buoy measures shortwave radiation above the sea surface, has been multiplied by 0.94 (albedo of seawater) to be consistent with the shortwave radiation from ECMWF. A 7-day running mean is applied to the daily shortwave radiation time series to filter out high-frequency variations purposes.

[38] ECMWF real-time operations has been using the 0.5° resolution Real-Time Global (RTG) analysis of SST [*Thiébaux et al.*, 2003], which results in some time lag. However, the main problem is that the TAO moorings are points located on the equator, whereas the ECMWF SST analysis is a gridded product (1° grid with grid boxes centered on either side of the equator), created by an analysis method that introduces some additional smoothing. The interpolation to the atmospheric model grid will introduce additional smoothing. All this smoothing matters because the physical process in which the rapid cooling in 1998 occurred seems to have been strong upwelling/mixing at the equator, which then spread out meridionally. Thus a 1° gridded, slightly smoothed analysis field is going to show the cooling delayed at the equator compared to point measurements at the TAO buoys.

[39] One other issue to emphasize here is that even KPP and GISS did not get as cold as the MODAS SST. This is partly related to the fact that there are also a few shortcomings of the atmospheric forcing used for the model simulations that affect the accuracy of the submodels. In our case, the shortwave radiation from ECMWF used in HYCOM simulations introduces some error ($\approx$50 W m^2) relative to the shortwave radiation measured by the TAO buoys during the transition period (Figure 11). In particular, ECMWF is $\approx$30 W m^2 larger in comparison to that provided by the buoy. Note that shortwave radiation at TAO buoy locations is measured at a height of 3.5 m above mean sea level. Buoy measurements are therefore multiplied by the albedo of sea to find shortwave radiation, entering sea surface, so that they can be consistent with ECMWF values. Differences in shortwave radiation between the TAO buoys and ECMWF can be as large as $\approx$100 W m^2, especially during May (Figure 11). The bias in shortwave radiation tends to cause excessive warming of the model SST.

[40] Both KPP and GISS give a mean bias of $\approx$1.5°C at (0°N, 125°W) and slightly smaller at (0°N, 140°W) (Table 5). In contrast, MY gives larger SST error, with a bias (RMS) of 2.14°C (2.92°C) at (0°N, 125°W). This is mainly because MY gave a much more gradual SST decrease than observed during the development of the 1998 La Niña. The SST seasonal cycle is successfully produced by all three models with a linear correlation coefficient generally >0.8. The non-dimensional skill score, calculated using the RMS SST difference and the standard deviation of the buoy SST, demonstrates that KPP and GISS performed relatively better in HYCOM than MY when simulating the daily SST during the 1998 transition period.

[41] A zonal temperature cross-section analysis (Figure 12) is presented along the equator to provide some physical insight as to the reason for the differences among the MLSs in the HYCOM simulations of the 1998 La Niña onset and development. Before the La Niña event started (March 1998), the SST is similar for all the MLSs (Figure 12a). The thick white line in the figure is the diagnosed MLD calculated as the first depth at which the density increase with respect to the surface is the equivalent of 0.2°C, and this MLD is also similar for all three MLSs. However, in the eastern equatorial Pacific between $\approx$20 and 60 m ($\approx$80 and 100 m) depth, MY gives cooler (warmer) temperatures than GISS or KPP. This diffusion of the thermocline with MY occurs because MY has much higher diffusion coefficients than GISS or KPP in this depth range. This reduces the surface cooling caused by upwelling during the La Niña development when MY is used (Figure 12b). Previously, *Halpern et al.* [1995] reported that MY may result in a relatively deep thermocline. Such a feature is also evident from the HYCOM simulation in the eastern equatorial Pacific during June of 1998 (e.g., see the 22°C isotherm in Figure 12b). However, except for this particular case, MY generally performs well as we already analyzed SSTs at any buoy locations.

6. The Impact of Wind Errors on the Representation of SSTs

[42] A significantly large bias in wind speeds from ECMWF used for the HYCOM simulations is a possible source for the relatively warm model-simulated SSTs in comparison to buoy SSTs during the 1998 transition period discussed in the preceding section. An evaluation of monthly wind speeds from ECMWF with those from the satellite-based Special Sensor Microwave/Imager (SSM/I) clearly

Table 5. SST Validation From May Through July in 1998[a]

TAO buoy	Product	RMS, (°C)	ME, (°C)	R	SS	Std. dev. σ_X	(°C) σ_Y
(0°N, 125°W)	MODAS	0.58	0.28	0.97	0.93	2.18	2.04
	ECMWF	1.12	0.64	0.91	0.73	2.18	2.24
	KPP	2.16	1.69	0.88	0.02	2.18	1.78
	GISS	1.90	1.51	0.85	0.24	2.18	1.96
	MY	2.92	2.14	0.87	−1.47	2.18	1.05
(0°N, 140°W)	MODAS	0.43	0.18	0.99	0.97	2.41	2.36
	ECMWF	1.13	0.74	0.94	0.78	2.41	2.42
	KPP	1.45	1.11	0.92	0.64	2.41	2.12
	GISS	1.70	1.47	0.96	0.51	2.41	1.84
	MY	2.76	2.05	0.89	−0.72	2.41	1.33

[a]Statistical values are based on daily SSTs. As before, σ_X refers to the standard deviation of buoy SSTs, and σ_Y refers to that of ECMWF and three MLSs.

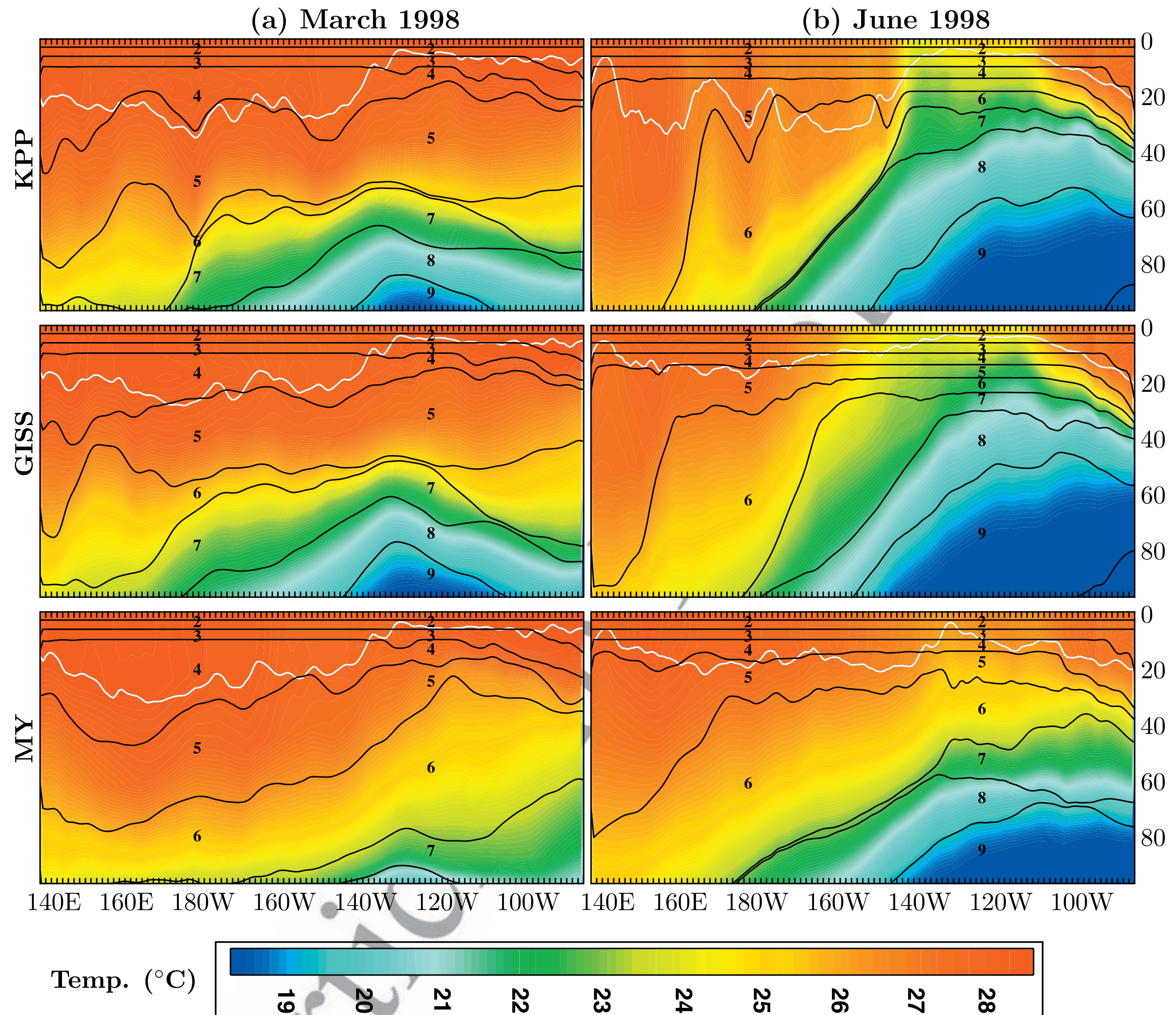

Figure 12. Cross-section of mean temperature along equator from HYCOM using the KPP, GISS and MY mixed layer submodels during March and June of 1998. The thick white line represents the mixed layer depth (MLD), which is a diagnostic quantity in HYCOM. The model layers (separated by black lines) are numbered in each panel, demonstrating significant differences before the La Niña started (e.g., March 1998) and during the transition period (e.g., June 1998), especially in the eastern equatorial Pacific.

reveals noteworthy differences (Figure 13). Here, a radiometer (used for SSM/I measurements) measures polarization mixing and sea foam emission, and considered as the truth though they have their unique errors. SSM/I winds are already calibrated to equivalent neutral conditions [*Meissner et al.*, 2001]. For comparisons, neutral SSM/I winds were converted to stability-dependent 10 m winds (i.e., winds that would be locally observed) using near-surface atmospheric variables (i.e., air temperature, relative humidity) from ECMWF globally.

[43] While the spatial patterns of wind speeds generally agree each other, ECMWF winds generally are too strong (>2 m s^{-1}) at the eastern Pacific cold tongue (Figure 13). This becomes evident just after the 1998 transition started (i.e., in June). A similar bias remains afterward. Wind speeds from other numerical weather predictions products, the National Center Environmental (NCEP) and the Navy Operational Global Atmospheric Prediction System (NOGAPS), also had similar biases during the same time period (not shown). Strong winds are an indication of deeper mixed layer depths in the HYCOM simulations, indication of warmer SSTs than expected. In other words, the deep ML gives a warmer SST when heat is lost from the mixed layer. If the mixed layer is deeper, then it will cool more slowly. For a heating case, the opposite would happen, i.e., a shallower mixed layer would warm up more rapidly if it were being heated.

[44] The explanation for the warm HYCOM SSTs in the preceding paragraph is based on that fact that winds from SSM/I are weaker in the cold tongue region. Thus the question arises, "are SSM/I winds actually correct". To

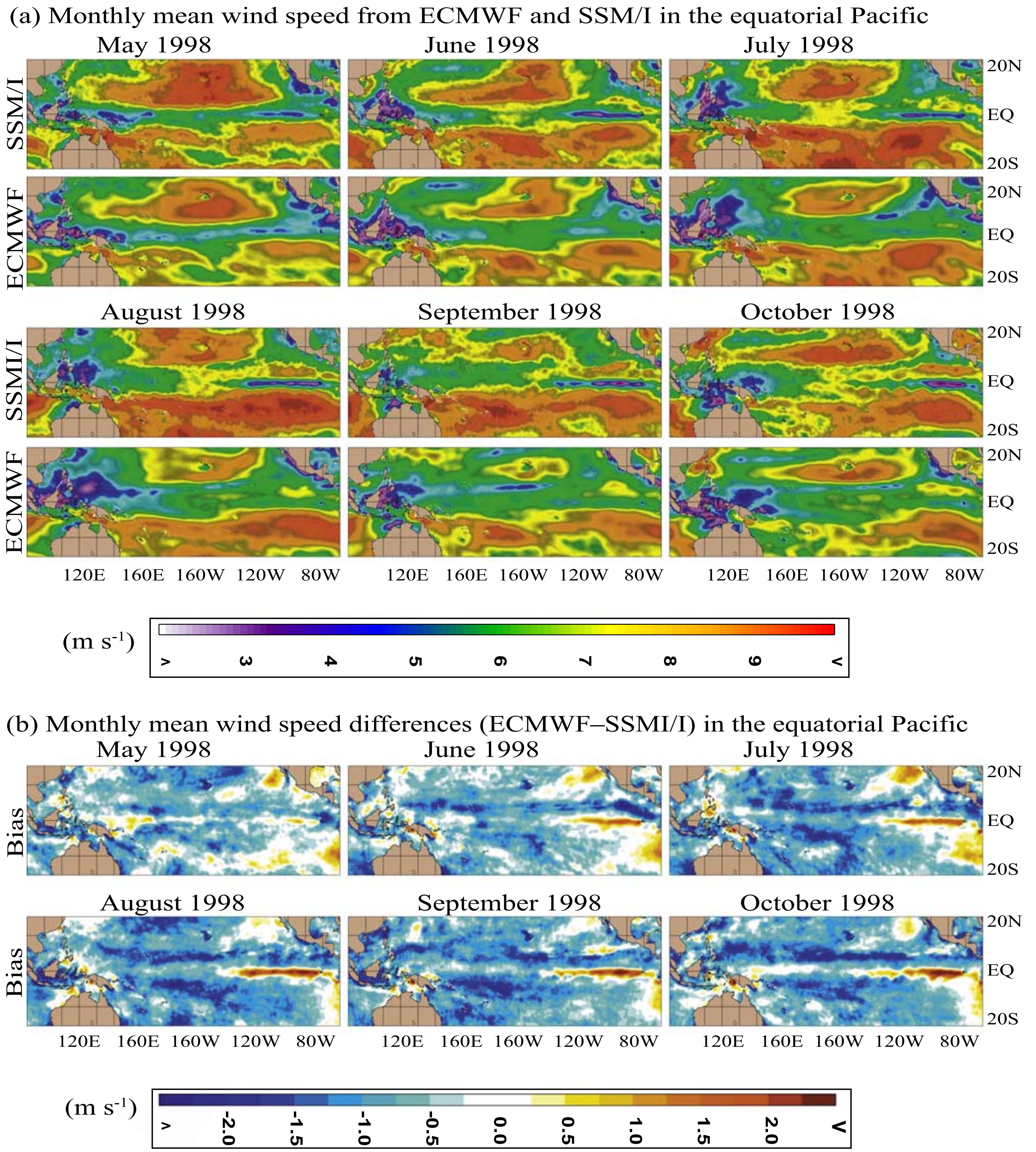

Figure 13. An evaluation of monthly wind speed at 10 m above the sea surface between ECMWF and SSM/I: (a) Spatial variability of wind speed in the equatorial Pacific Ocean from May through October 1998, and (b) differences between the two. In the latter, the red color demonstrates regions where wind speeds from ECMWF are stronger than those from SSM/I.

answer this question, we formed monthly mean wind speeds from three TAO buoy locations. Daily wind speeds reported at 4 m above the sea surface from buoys were first adjusted to 10 m using the COARE3.0 algorithm [*Fairall et al.*, 2003], and monthly means were formed. We then compared the winds from the buoys with those from SSM/I and ECMWF (Figure 14). Daily wind speed measurements from the TAO buoy at (0°N, 125°W), where the SST time series from the MLSs are evaluated (Figure 9), were not available during the entire year, so we use the closest location (5°S, 125°W) instead, where wind measurements are available. In

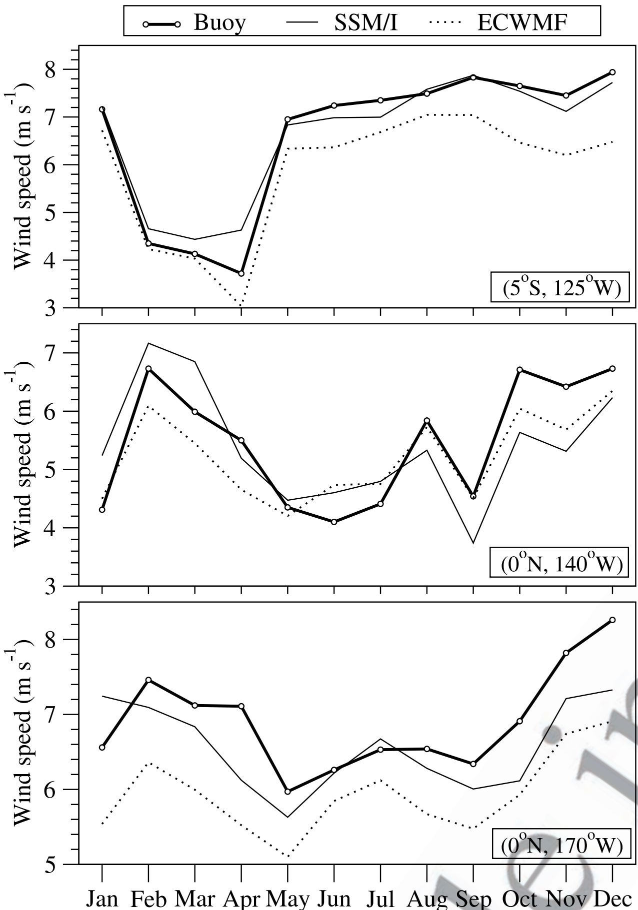

Figure 14. Time series of monthly wind speed at 10 m above the sea surface from TAO buoy measurements, ECMWF and SSM/I at three locations in 1998. Wind speeds at 4 m above the sea surface from the buoys were adjusted to 10 m, and the neutral equilibrium SSM/I winds were converted to actual winds for these comparisons as explained in the text. The 1998 transition from the El Niño to the La Niña is from May through October, as explained in section 4a. Note that the *y*-axis range is different for each panel.

addition, evaluations are also presented (0°S, 170°W) for comparisons in the central Pacific Ocean.

[45] Winds from SSM/I generally agree with the observations better than those from ECMWF. There is almost no difference between the ECMWF and buoy wind at (0°N, 140°W) from May to November, 1998, while SSTs from HYCOM are still too warm (Figure 9). Thus wind-forcing is not the primary reason for the model failure. In fact, it is already shown that errors in shortwave radiation are also major contributor to model SST bias at this particular location (Figure 11). One possible reason is that extensive cloudiness, which can be expected during such strong mesoscale events, may have affected the accuracy of shortwave radiation predicted by ECWMF. Unfortunately, there are no daily cloud cover observations to confirm this statement, but relatively low outgoing longwave radiation (OLR) values from ECMWF and NOAA in the central equatorial Pacific confirm the existence of cloudiness (not shown). In addition, while the wind speed from ECMWF is reliable, in general, one should note that the transition is not a local event. This means other external effects (i.e., large scale events), such as Rossby wave propagation across the equatorial Pacific, can have significant influences on the SST variability.

7. Summary and Conclusions

[46] Overall, the three mixed layer submodels (KPP, GISS and MY) perform similarly in simulating SST over the global ocean. This is the case on both climatological and inter-annual timescales. The simulations presented in this paper did not include direct assimilation of SST or other date-specific data and there is no relaxation to SST climatology. Hence we were able to examine the first-order behavior of each individual MLS in simulating SSTs. Daily SST time series from a large number of buoys are used for the validation. In addition, daily SSTs from a the satellite-based MODAS re-analyses and a numerical weather prediction product (ECMWF) were included in our analysis as reference data sets.

[47] We specifically examined the SST variability during the onset of the 1998 La Niña event since (1) this event is one of the largest short-term SST events on record (≈8°C change over 30 days), and therefore (2) simulating the westward propagation of the SST cooling during this event presents a challenge and a useful test for the evaluation of mixed layer models. We first properly identify the transition period from the 1997 El Niño to 1998 La Niña using various indices, then simulate the SST during this period with HYCOM using three MLSs, and finally determine the ability of the models to reproduce the La Niña event. Evaluation is also extended to other more normal years to further examine differences among the mixed layer submodels. The behavior of the three submodels are considered under particular wind and thermal forcing, which are from ECMWF.

[48] Based on the results, HYCOM is able to represent not only the extent of the SST cooling but also its magnitude (a warm annual mean bias of ≈1°C in comparison to observations) during the 1998 transition from El Niño to La Niña. KPP, GISS, and MY performed similarly in making this transition, while the MY simulation gave a slightly diffusive thermocline, resulting in an underestimate of the SST cooling. Overall, performance of MLSs examined at locations outside the equatorial Pacific during other time periods, from 1996 through 2001, further confirmed the accuracy of HYCOM SSTs when using KPP, GISS or MY. All the MLSs gave nearly identical results in generating climatological mean SSTs over the global Ocean. In particular, the MLSs gave a global mean RMS SST difference of ≈0.7°C in comparison to the NOAA climatology, based on satellite and in situ SSTs, over the seasonal cycle.

[49] In the paper, we also demonstrated that the ECMWF SSTs may not be quite accurate when there are strong trends in the SST with time (e.g., during the transition period) due to the time-lagged average used in their analyses. On the other hand, the MODAS SST did not have such a problem and accurately reproduced the observed daily SST variability. During the transition period, SST from ECMWF, an operational gridded model product, has a time lag for the

cooling at the equator of more than one week. This is because the TAO moorings are points located at the equator, whereas the SST used in the ECMWF analyses is a gridded product ($1.125° \times 1.125°$), created by an analysis method. Daily SSTs from the MODAS SST re-analyses captures the magnitude and timing of the large SST drop of $\approx 7°C$ since it includes satellite data centered on the analysis time (i.e., from both 1 day before and 1 day after).

[50] Finally, performance of KPP, GISS and MY explored in this paper is based on a particular OGCM (i.e., HYCOM) which use atmospheric forcing from a given operational weather center (i.e., ECMWF). Further studies using various other OGCMs and atmospheric forcing products will provide more information about the reliability of these MLSs. In general, results based on an extensive set of buoy SST time series, as presented in this paper, clearly demonstrate the similar success of all MLSs in simulating daily and monthly SST. We also note that HYCOM includes additional mixed layer models, which are not presented in this study.

[51] **Acknowledgments.** This work was funded by the Office of Naval Research (ONR) under the 6.1 project, Global Remote Littoral Forcing via Deep Water Pathways and by the National Ocean Partnership Program (NOPP). This paper is contribution NRL/JA/7320/05/5166 and has been approved for public release. Appreciation is extended to two anonymous reviewers for their helpful comments. The help of G. Halliwell (University of Miami) in implementing the mixed layer models in HYCOM is greatly appreciated. Appreciation is also extended to M. McPhaden of the TAO project office for providing buoy data from the TAO array. The HYCOM simulations were performed under the Department of Defense High Performance Computing Modernization Program on an IBM SP POWER3 at the Naval Oceanographic Office (NAVOCEANO), Stennis Space Center, Mississippi and on a HP/COMPAQ SC45 at the United States Army Engineer Research and Development Center (ERDC), Vicksburg, Mississippi.

References

Barnston, A. G., M. Glantz, and Y. He (1999), Predictive skill of statistical and dynamical climate models in SST forecasts during the 1997-98 El Niño and the 1998 La Niña onset, *Bull. Am. Meteorol. Soc.*, *60*, 217–243.

Barron, C. N., and A. B. Kara (2006), Satellite-based daily SSTs over the global ocean, *Geophys. Res. Lett.*, *33*, L15603, doi:10.1029/2006GL026356.

Bleck, R. (2002), An oceanic general circulation model framed in hybrid isopycnic-cartesian coordinates, *Ocean Modelling*, *4*, 55–88.

Bond, N. A., J. E. Overland, M. Spillane, and P. J. Stabeno (2003), Recent shifts in the state of the North Pacific, *Geophys. Res. Lett.*, *30*(23), 2183, doi: 10.1029/2003GL018597.

Canuto, V. M., A. Howard, Y. Cheng, and M. S. Dubovikov (2002), Ocean turbulence, Part II: Vertical diffusivities of momentum, heat, salt, mass, and passive scalars, *J. Phys. Oceanogr.*, *32*, 240–264.

Cassou, C., and C. Perigaud (2000), ENSO simulated by intermediate coupled models and evaluated with observations over 1970-98, Part II: Role of the off-equatorial ocean and meridional winds, *J. Clim.*, *13*, 1635–1663.

Chassignet, E. P., H. E. Hurlburt, O. M. Smedstad, G. R. Halliwell, A. J. Wallcraft, E. J. Metzger, B. O. Blanton, C. Lozano, D. B. Rao, P. J. Hogan, and A. Srinivasan (2006), Generalized vertical coordinates for eddy-resolving global and coastal ocean forecasts, *Oceanography*, *19*, 20–31.

Cronin, M. F., and M. J. McPhaden (1997), The upper ocean heat balance in the western equatorial Pacific warm pool during September-December 1992, *J. Geophys. Res.*, *102*, 8533–8553.

Elsner, J. B., and A. B. Kara (1999), *Hurricanes of the North Atlantic: Climate and Society*, 496 pp, Oxford Univ. Press, New York.

Enfield, D. B., and D. A. Mayer (1997), Tropical Atlantic sea surface temperature variability and its relation to El-Niño-Southern Oscillation, *J. Geophys. Res.*, *102*, 929–945.

Fairall, C. W., E. F. Bradley, J. E. Hare, A. A. Grachev, and J. B. Edson (2003), Bulk parameterization of air-sea fluxes: Updates and verification for the COARE algorithm, *J. Clim.*, *16*, 571–591.

Gibson, J. K., P. Kllberg, S. Uppala, A. Hernandez, A. Nomura, and E. Serrano (1999), ECMWF Re-Analysis Project Report Series: 1. ERA description (Version 2), 74 pp, U. K..

Halpern, D., Y. Chao, C.-C. Ma, and C. R. Mechoso (1995), Comparison of tropical Pacific temperature and current simulations with two vertical mixing schemes embedded in an ocean general circulation model and reference to observations, *J. Geophys. Res.*, *100*, 2515–2522.

Hanley, D. E., M. A. Bourassa, J. J. O'Brien, S. R. Smith, and E. R. Spade (2003), A quantitative evaluation of ENSO indices, *J. Clim.*, *16*, 1249–1258.

Harrison, D. E., and G. A. Vecchi (2001), El Niño and La Niña equatorial Pacific thermocline depth and sea surface temperature anomalies, 1986-1998, *Geophys. Res. Lett.*, *28*, 1051–1054.

Hendon, H. H. (2003), Indonesian rainfall variability: Impacts of ENSO and local air-sea interaction, *J. Clim.*, *16*, 1775–1790.

Kållberg, P., A. Simmons, S. Uppala, and M. Fuentes (2004), ERA-40 Project Report Series, the ERA-40 archive 17, 31 pp, U. K..

Kara, A. B., P. A. Rochford, and H. E. Hurlburt (2003), Mixed layer depth variability over the global ocean, *J. Geophys. Res.*, *108*(C3), 3079, doi:10.1029/2000JC000736.

Kara, A. B., H. E. Hurlburt, P. A. Rochford, and J. J. O'Brien (2004), The impact of water turbidity on the inter-annual sea surface temperature simulations in a layered global ocean model, *J. Phys. Oceanogr.*, *34*, 345–359.

Kara, A. B., A. J. Wallcraft, and H. E. Hurlburt (2005a), A new solar radiation penetration scheme for use in ocean mixed layer studies: An application to the Black Sea using a fine resolution HYbrid Coordinate Ocean Model (HYCOM), *J. Phys. Oceanogr.*, *35*, 13–32.

Kara, A. B., H. E. Hurlburt, and A. J. Wallcraft (2005b), Stability-dependent exchange coefficients for air-sea fluxes, *J. Atmos. Oceanic Technol.*, *22*, 1080–1094.

Kara, A. B., H. E. Hurlburt, A. J. Wallcraft, and M. A. Bourassa (2005c), Black Sea mixed layer sensitivity to various wind and thermal forcing products on climatological time scales, *J. Clim.*, *18*, 5266–5293.

Kara, A. B., A. J. Wallcraft, and H. E. Hurlburt (2005d), Sea surface temperature sensitivity to water turbidity from simulations of the turbid Black Sea using HYCOM, *J. Phys. Oceanogr.*, *35*, 33–54.

Large, W. G., G. Danabasoglu, S. C. Doney, and J. C. McWilliams (1997), Sensitivity to surface forcing and boundary layer mixing in a global ocean model: Annual-mean climatology, *J. Phys. Oceanogr.*, *27*, 2418–2447.

Latif, M., and T. P. Barnett (1996), Decadal climate variability over the North Pacific and North America: dynamics and predictability, *J. Clim.*, *9*, 2407–2423.

McPhaden, M. J. (1999), Genesis and evolution of the 1997-98 El Niño, *Science*, *283*, 950–954.

Meissner, T., D. Smith, and F. J. Wentz (2001), A 10-year inter-comparison between collocated SSM/I oceanic surface wind speed retrievals and global analyses, *J. Geophys. Res.*, *106*, 11,731–11,742.

Mellor, G. L., and T. Yamada (1982), Development of a turbulence closure model for geophysical fluid problems, *Rev. Geophys.*, *20*, 851–875.

Murray, R. J. (1996), Explicit generation of orthogonal grids for ocean models, *J. Comp. Phys.*, *126*, 251–273.

NAVOCEANO (2003), Database description for the generalized digital environmental model (GDEM-V) Version 3.0, OAML-DBD-72, 34 pp, Stennis Space Center, MS.

Pacanowski, R., and S. G. H. Philander (1981), Parameterization of vertical mixing in numerical models of tropical oceans, *J. Phys. Oceanogr.*, *11*, 1443–1451.

Picaut, J., E. Hackert, A. J. Busalacchi, R. Murtugudde, and G. S. E. Lagerloef (2002), Mechanisms of the 1997-1998 El Niño-La Niña, as inferred from space-based observations, *J. Geophys. Res.*, *107*(C5), 3037, doi:10.1029/2001JC000850.

Reynolds, R. W., N. A. Rayner, T. M. Smith, and D. C. Stokes (2002), An improved in-situ and satellite SST analysis for climate, *J. Clim.*, *15*, 1609–1625.

Schneider, E. K., B. Huang, Z. Zhu, D. G. DeWitt, J. L. Kinter III, B. Kirtman, and J. Shukla (1999), Ocean data assimilation, initialization, and predictions of ENSO with a coupled GCM, *Mon. Weather Rev.*, *127*, 1187–1207.

Shinoda, T. (2005), Impact of diurnal cycle of solar radiation on intraseasonal SST variability in the western equatorial Pacific, *J. Clim.*, *18*, 2628–2636.

Smith, N. R., and G. D. Hess (1993), A comparison of vertical eddy mixing parameterizations for equatorial ocean model, *J. Phys. Oceanogr.*, *23*, 1823–1830.

Steele, M., R. Morley, and W. Ermold (2001), PHC: A global ocean hydrography with a high quality Arctic Ocean, *J. Clim.*, *14*, 2079–2087.

Sutton, R. T., and M. R. Allen (1997), Decadal predictability in North Atlantic sea surface temperature and climate, *Nature*, *388*, 563–567.

Swenson, M. S., and D. V. Hansen (1999), Tropical Pacific Ocean mixed layer heat budget: The Pacific cold tongue, *J. Phys. Oceanogr.*, *29*, 69–81.

Thiébaux, J., E. Rogers, W. Wang, and B. Katz (2003), A new high-resolution blended real-time global sea surface temperature analysis, *Bull. Am. Meteorol. Soc.*, *84*, 645–656.

Trenberth, K. E. (1997), The definition of El Niño, *Bull. Am. Meteorol. Soc.*, *78*, 2771–2777.

Wolter, K., and M. S. Timlin (1998), Measuring the strength of ENSO-how does 1997/98 rank?, *Weather*, *53*, 315–324.

E. P. Chassignet, Center for Ocean-Atmospheric Prediction Studies, and Department of Oceanography, Florida State University, Tallahassee, FL, USA.

A. B. Kara, P. J. Martin, and A. J. Wallcraft, Naval Research Laboratory, Oceanography Division, Code 7320, Bldg. 1009, Stennis Space Center, MS 39529, USA. (kara@nrlssc.navy.mil)

American Geophysical Union
Author Query Form

Journal: **Journal of Geophysical Research – Oceans**
Article Name: Kara (2007JC004250)

Please answer all author queries.

1. Citations of the following references were changed to match the reference list. Please check if this is appropriate

   - Canuto et al., 2001 to **Canuto et al., 2002**
   - McPhaden, 1995 to **McPhaden, 1999**

2. The following references were cited in your paper, but were not provided in the reference list. Please provide the appropriate information for the references or delete the citations from the text.

   - *Chassignet et al.,* 2003
   - *Halliwell*, 2004
   - *Murphy,* 1995

3. The following references were present in the reference list but not cited in your paper. Kindly provide citation for the said references or remove them from the reference list.

   - Cassou and Perigaud, 2000
   - Kara et al., 2005d
   - Pacanowski and Philander, 1981