

## Using Data and Intermediate Coupled Models for Seasonal-to-Interannual Forecasts

CLAIRE M. PERIGAUD, CHRISTOPHE CASSOU, BORIS DEWITTE, AND LEE-LUENG FU

*Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California*

J. DAVID NEELIN

*Department of Atmospheric Sciences and Institute of Geophysics and Planetary Physics,  
University of California, Los Angeles, Los Angeles, California*

(Manuscript received 26 March 1999, in final form 1 December 1999)

### ABSTRACT

This paper provides a detailed illustration that it is beneficial for ENSO forecasting to improve in priority the model parameterizations, instead of increasing only the consistency of the initial conditions with the coupled model. Moreover it is shown that the latter can lead to misleading results. Using sea level data in addition to wind to initialize the Cane and Zebiak model does not improve El Niño forecasts. Nudging the observed wind to the model one to initialize the forecasts as proposed by Chen et al. also fails to correct the model deficiencies and degrades the initial conditions of the ocean and atmosphere. These failures are explained by large model errors in the off-equatorial sea level and wind anomalies that play a key role in the coupled behavior. The use of data to estimate new model parameterizations allows for significantly improving both the initial conditions and the forecasts up to 6-month lead time. This success holds for all the various initialization procedures used in this study. Because of erroneous winds simulated by the atmospheric component in the eastern Pacific, errors grow fast though. Replacing the atmospheric model by a statistical one results in more reliable predictions over 1980–98. For lead times up to 1 yr, the model predicts well the observed anomalies between 1984 and 1993, including the sea level rises along the ITCZ after warm events and their subsequent equatorward migration. This success is attributed to the consistency between the observed anomalies over this period and the mechanisms involved in maintaining the oscillatory behavior of the model, including the off-equatorial meridional wind anomalies.

### 1. Introduction

Despite a long history in predicting short-range climate variations and the availability of a wide variety of ENSO forecast models today (for a review, see Latif et al. 1998), the use of oceanic data to initialize forecasts is fairly recent. The present study started in December 1991, a few months prior to the launch of the Ocean Topography Experiment Poseidon (TOPEX/Poseidon) satellite. The objective was then to prepare the procedures to use the very accurate altimetric data delivered by this satellite for improving ENSO predictions. Using sea level data to initialize a coupled model is indeed likely to have a long-lasting impact on forecasts, because these data monitor the vertically integrated density of the ocean, which happens to have a much longer memory than the atmosphere. A few studies have thus illustrated the importance of the oceanic data assimilation

on predictions (Kleeman et al. 1995; Ji and Leetma 1997; Rosati et al. 1997; Fischer et al. 1997; Ji et al. 1998; Schneider et al. 1998). In addition, coupled models are very sensitive to small changes in their initial conditions, so it is likely that the accuracy of the data used in the initialization process is a crucial issue.

This study uses the model presented in Zebiak and Cane (1987, hereafter CZ) and sea level estimates that are derived either from expendable bathy thermograph (XBT) since 1980 or from TOPEX/Poseidon since October 1992 (see appendix A). The CZ code, the climatological files, and the standard procedure to initialize the forecasts including the regular updates of the wind anomalies since 1964 are all provided by Dr. Zebiak from Lamont-Doherty Earth Observatory (LDEO). Our original approach consisted of improving the initial conditions of the model without modifying the code itself. We first developed and tested various methods for initializing the model with sea level data. In particular we decomposed the dynamic height anomalies into their Kelvin and Rossby components as in Perigaud and Dewitte (1996) and inserted them into the baroclinic ocean over 1980–95, or we applied the optimal Kalman filter

---

*Corresponding author address:* Claire Perigaud, Jet Propulsion Laboratory, MS 300/323, 4800 Oak Grove Drive, Pasadena, CA 91109.  
E-mail: cp@pacific.jpl.nasa.gov

developed for a baroclinic model similar to CZ (Fukumori 1995) to assimilate the TOPEX/Poseidon data over 1992–95. Various procedures to initialize the mixed layer and the atmosphere with the assimilated baroclinic fields, with the observed SST, and/or with the observed winds were also tested. These experiments are not described in detail here; the important message to retain for now is that none of them led to improved forecasts. One series with sea level initialization is reported here for illustration. For comparison, the corresponding plots are provided for two CZ series published in the literature, either the series initialized with the LDEO1 procedure described in Cane et al. (1986), or the one initialized with the LDEO2 procedure described in D. Chen et al. (1995). The sole difference between these two procedures is the wind used to initialize the system: each time the baroclinic model is integrated in an initialization step increment, LDEO1 forces it with the observed wind anomalies only, whereas LDEO2 nudges the latter with the model wind anomalies.

The experiments presented in this paper correspond to model simulations that we classify either as “forced” or “coupled.” Forced simulations refer to those where the model is integrated in time with some data input. They include all the experiments for which observations have been introduced in the model, whether observations are wind, or sea level, or SST, or a combination of oceanic and atmospheric data, and whether data are introduced into the model via direct insertion, nudging, or optimal assimilation. By contrast, coupled simulations refer to those where no data are ever introduced during the experiment, the model evolving freely after some forced initial conditions. Coupled experiments include multidecadal simulations and forecasts. Models behave very differently whether they are integrated in a forced or in a coupled context, and recent efforts have been devoted to make the forced initial conditions more consistent with the coupled model to reduce the shock at the transition. Thus, the LDEO2 procedure prepares forced initial conditions that are more consistent with the coupled mode than LDEO1. It significantly reduces the number of forecast cases for which the error grows (Xue et al. 1997). It delivers forecasts that are extremely good over the period 1980–92 (D. Chen et al. 1995, 1997). It fails to predict the 1997–98 El Niño event, but after assimilating tide gauge sea level data in the baroclinic model and nudging sea level and wind as in LDEO2 for the wind only, Chen et al. (1998) succeed in predicting the last event with forecasts initiated from March 1997. Part of the present paper investigates the reasons why differences in the initialization procedures lead the same CZ model to successes or failures in forecasting.

Although most of the sea level data used in our study do not have the accuracy of TOPEX/Poseidon, we are convinced that the preliminary results did not fail because of data errors. Neither did they because of inadequate methods of assimilation. One of the purposes

of this paper is to provide evidence that failure to forecast with CZ is due to inadequate parameterizations and not so much to ad hoc initialization nor missing physics. The data used in this study are presented either in Perigaud and Dewitte (1996) or in Perigaud et al. (2000) where they are compared with CZ forced and coupled simulations, respectively. Dewitte and Perigaud (1996, hereafter DP96) describe the new parameterization of the subsurface temperature in the mixed layer,  $T_{sub}$ , and the spatiotemporal relationship between SST and heating released by atmospheric convection,  $Conv$ , which can be introduced in the atmospheric model. The model with the revised parameterizations is named  $T_{sub}.Conv$ . DP96 also analyze the improved anomalies simulated by  $T_{sub}.Conv$ , in a forced context. Cassou and Perigaud (2000, hereafter CP00) analyze multidecadal coupled integrations and demonstrate that  $T_{sub}.Conv$  behaves very differently than CZ.  $T_{sub}.Conv$  simulates SST, sea level, and wind anomalies in much better agreement with data than CZ, and in order to further improve the agreement, the atmospheric component was replaced by a statistical one in a model referred to as  $T_{sub}.Astat$  (also presented in CP00). After analyzing the behaviors of these three intermediate coupled models (ICMs) (CZ,  $T_{sub}.Conv$ , and  $T_{sub}.Astat$ ) during forced or multidecadal coupled integrations, the present paper examines their forecasting skills, based on several series of 2-yr-long coupled experiments initialized every month from a variety of forced conditions. The paper is organized with section 2 for the CZ forecast series, section 3 for  $T_{sub}.Conv$ , section 4 for  $T_{sub}.Astat$ , and section 5 for a summary and perspectives.

## 2. Forecasts delivered by the CZ model with various initialization procedures

This section presents an analysis of the forecasts delivered by the CZ model after having applied three different initialization procedures (see Table 1a). The CZ model is first used to deliver forecasts initialized with the standard LDEO1 procedure, meaning that the ocean component started from rest in 1964 has been forced by the “detrended” Florida State University (FSU) wind stress anomalies and that the latter has provided the SST, which has then been used to drive the atmospheric model. The forecasts initialized with the standard LDEO1 procedure are referred to as CZ.STD. The second procedure uses sea level data to initialize the model as described in appendix A and delivers a series of forecasts referred to as CZ.SL. Finally, the LDEO2 procedure giving relatively more coupled initial conditions is applied to deliver a series of forecasts referred to as CZ.CPLIC. Most of the forecast series shown in this paper are provided every month since January 1980 because sea level data are not available before that date. Series of SST Niño-3 indices (averages of SST anomalies over 5°S–5°N and 150°–90°W) are presented with a zoom between January 1981 and January 1993 in Figs.

TABLE 1. Identification of the CZ, Tsub.Conv, and Tsub.Astat experiments.

(a)				
Name: CZ	Friction	Initialization procedure	Wind forcing for generating ICs	
CZ.STD	30	LDEO1	FSU	
CZ.SL	30	SL (method 1)	FSU	
CZ.CPLIC	30	LDEO2	FSU	
(b)				
Tsub.Conv	Friction	Convect weight (%)	Initialization procedure	Wind forcing for generating ICs
Tsub.Conv1	6	100	LDEO1	FSU
Tsub.Conv2	24	10	LDEO1	FSU
Tsub.Conv2.SL	24	10	SL (method 2)	FSU
Tsub.Conv3	6	100	LDEO1	Astat (first SVD)
T.Con3.CPLIC	6	100	LDEO2	Astat (first SVD)
(c)				
Tsub.Astat	Friction	Coupl. coeff.	Initialization procedure	Wind forcing for generating ICs
Tsub.Astat1	6	1.0	LDEO1	Astat (first SVD)
Tsub.Astat2	6	1.0	LDEO1	Astat (seven SVD)
Tsub.Astat3	6	1.2	LDEO1	Astat (seven first SVD)

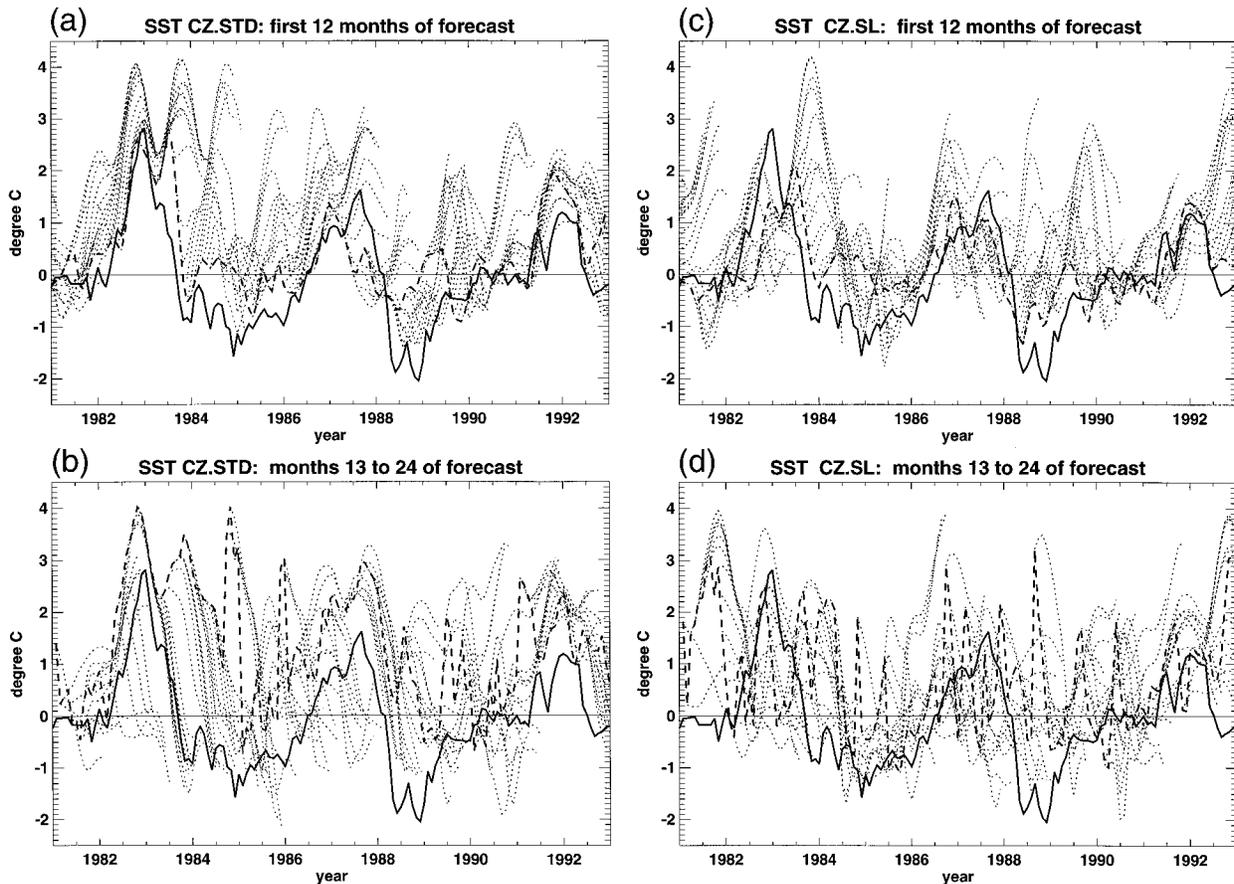


FIG. 1. Time series of Niño-3 SST index between Jan 1981 and 1993 from data (solid), or from CZ during 12-month-long forecasts (dotted). (top) The first year of forecasts (with the initial conditions in dashed); (bottom) the second year (with the 12-month lead time forecasts in dashed); (left) CZ.STD; (right) CZ.SL.

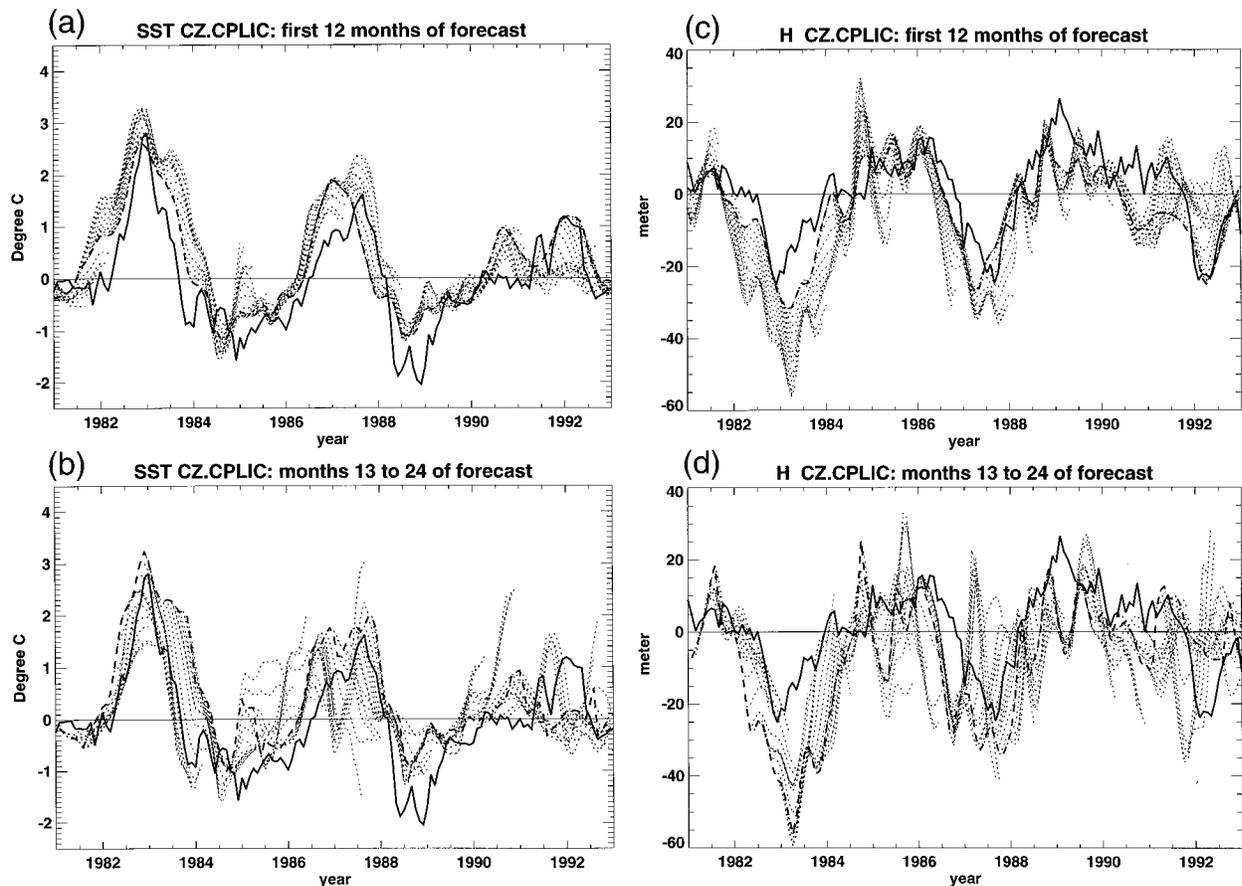


FIG. 2. Same as Fig. 1 but for (a), (b) SST Niño-3 and (c), (d) H Niño-W simulated by CZ.CPLIC.

1 and 2. For clarity, forecasts during only 12-month-long segments are plotted per diagram, the top diagrams corresponding to the first 12 months and the bottom ones to the following 12 months of prediction.

#### a. SST forecasts

The CZ model initialized with the LDEO1 procedure tends to systematically predict growing warm events with peaks up to  $3^{\circ}$  or  $4^{\circ}\text{C}$  after a year, which then decay back to normal or slightly cold conditions (Fig. 1). The reader should not be surprised by the large errors between the individual forecasts and the observations. Similar diagrams are found in the published literature (D. Chen et al. 1995, 1997). Such diagrams allow one to question the validity of the conventional statistics (correlations and rms errors as a function of lead time) used to determine the model predictive skill. For example, the correlation between the observed and the 12-month lead predicted time series over 1981–93 (represented by the dashed line in Fig. 1b) is equal to 0.5, well above the 95% level of confidence equal to 0.2. Such statistics make the reader think that the model is skillful, whereas it is obvious from Fig. 1 that the model

predictions do not agree with the observed SST. Indeed we choose to avoid the use of the conventional statistics for evaluating the model predictive skills as a function of lead time because these can be misleading for ENSO timescales analyzed with 1–2-yr-long forecasts (C. Wunsch and S. Philander 1999, personal communication). Figure 1 shows an overwhelming tendency of the CZ model to predict warm anomalies. The reasons why this model simulates a warm bias in a forced or coupled context have been analyzed in Perigaud and Dewitte (1996) and in Perigaud et al. (2000), respectively.

One of the series of forecasts performed by the CZ model after initialization with sea level data is presented in Figs. 1c,d. The model initialized with sea level data delivers SST forecasts that are neither better nor worse than the above (see Table 2). Nevertheless, sea level data certainly have a strong impact on the individual predictions. Individual CZ.SL forecasts differ from the standard ones by more than  $2^{\circ}\text{C}$  most of the time. By contrast, comparing the dashed lines in Figs. 1a and 1c indicates that the impact of sea level data on the SST simulated in a forced context is little. A similar type of behavior is reproduced in CZ.STD and CZ.SL. Most of the individual predictions undergo a strong oscillation

TABLE 2. Rms error between observed and predicted time series over 1981–93 for lead time equal to 0, 6, and 12 months. SST Niño-3 is the average of SST over 5°S–5°N, 150°W–90°W, TX Niño-4 is the average of the zonal wind stress over 5°S–5°N, 160°E–150°W, and H Niño-W is the average of the thermocline depth over 5°S–5°N, 130°E–170°E. For the wind, the observations are Astat winds (with FSU, the errors are even larger). The observed rms variability of the three signals are 1.0°C, 0.13 dyn cm<sup>-2</sup>, and 17 m, respectively. The reader should take these numbers with caution before concluding that one forecast is better than the other. Note that errors are close to the level of observed variability, that errors do not always increase with lead time, and that error in one field can decrease while the error in another field can increase. In addition, this list of three indexes is far from being complete, other indexes with the meridional wind or the off-equatorial signals are as important. Finally, indices can be misleading because they depend on the region where the average is taken and it can happen that the model badly locates the key features. The reader should also be aware that these numbers depend a lot on the duration of the period considered, extending the period in order to have more forecast cases does not give more reliable statistics though (Perigaud and Neelin 1999, unpublished manuscript).

Name	SST Niño-3 (°C)			TX Niño-4 (dyn cm <sup>-2</sup> )			H Niño-W (m)		
CZ	0.7	1.3	1.5	0.15	0.27	0.30	15	23	34
CZ.SL	0.7	1.4	1.5	0.14	0.28	0.24	4	33	35
CZ.CPLIC	0.7	0.8	0.8	0.11	0.15	0.17	10	17	22
Tsub.Conv1	0.7	1.0	1.0	0.10	0.12	0.11	9	10	14
Tsub.Conv2	0.7	1.0	1.4	0.10	0.11	0.27	9	11	23
Tsub.Conv2.SL	0.7	1.1	1.3	0.10	0.13	0.23	4	11	17
Tsub.Conv3	0.6	0.9	1.2	0.08	0.11	0.15	9	15	13
Tsub.Astat1	0.6	0.8	1.0	0.08	0.11	0.13	9	13	15
Tsub.Astat2	0.6	0.8	0.9	0.07	0.08	0.13	8	11	14
Tsub.Astat3	0.6	0.8	0.8	0.07	0.09	0.12	8	10	15

during the 2 yr of forecasting. The series predicted at a lead time equal to 12 months (dashed lines in Figs. 1b and 1d) are quite different from each other, but they both have the same large errors with data, with sharp jumps from month to month conversely to the initial conditions in SST or sea level. After 12 months of forecast, the model has completely lost the memory of the oceanic initialization. CZ.STD and CZ.SL suffer from severe deficiencies, in particular from strong shocks that take place at the transition between the forced and the coupled mode.

The LDEO2 procedure was developed to reduce such shocks. This procedure is quite efficient and delivers very good predictions between 1981 and 1993, even for lead times longer than 12 months (Figs. 2a,b; Table 2). The reader is invited to read appendix B to be informed of the detailed differences between the series analyzed here and the one published as LDEO2 in the literature. The CZ.CPLIC forecasts are by far the most realistic predictions of the SST Niño-3 index over 1981–93 among all the ones published so far. So one may argue that the procedure used to initialize the CZ.SL forecasts is inadequate, because it does not assimilate the sea level data in a context that is consistent with the physics of the coupled model. This is indeed a very valid concern and this is why we have tested various methods of sea level data assimilation in a more coupled context than LDEO1. However, all methods applied to CZ led to forecasts that are as poor as in CZ.STD or CZ.SL. Note that Chen et al. (1997) also found that the LDEO2 procedure does not work when the nudging is applied to the SST instead of the wind and that (Chen et al. 1998) correct the failure of the 1997–98 LDEO2 forecast by adding sea level data. These successes and failures are all the more intriguing as the LDEO2 initialization procedure does not correct errors that are intrinsic to the

model itself. This is also why the LDEO2 procedure will be tested with the other models presented in this paper (see section 3).

It can be expected that the LDEO2 procedure introduces features in the initial conditions that have the characteristics of the deficiencies found in the coupled simulations (see Perigaud et al. 2000). The latter are briefly recalled here: 1) The zonal wind stress anomaly along the equator has its maximum located almost always at the same place in the central Pacific, about 20° to the east of the date line, whereas the observed ones are mostly centered around the dateline. 2) The thermocline in the western Pacific is 2–3 times more upwelled than in reality during and after a warm event. 3) The equatorial zonal wind at the eastern boundary of the model has an opposite sign to the one in the central Pacific, with an amplitude at least twice as large as in the observations. 4) The zonal wind anomalies beyond 9° of latitude in the eastern Pacific are three times stronger than the strongest signal ever observed there. 5) The latter are associated with unrealistically large off-equatorial thermocline displacements, with signs that can be opposite to those observed. 6) The off-equatorial wind and thermocline anomalies play a key role in the oscillatory behavior of the coupled model. The reader can verify the first point with Fig. 1 of D. Chen et al. (1995). The other points are successively examined below.

### b. Thermocline in the equatorial wave guide

Figures 2c,d present thermocline depth anomalies averaged over Niño-W (5°S–5°N, 130°–170°E). The initial thermocline simulated by CZ.CPLIC is shallower than the observed one with differences as large as 20 m during most of 1983 and prior to the 1992 warming (Fig. 2c). Actually, it is always shallower than the observed

one. This is not the case in the standard initialization (not shown). The tendency for the thermocline to shoal in the western Pacific more than in the observations is characteristic of the coupled CZ behavior and explained in Perigaud et al. (2000). This cold reservoir in the western Pacific compensates the tendency to predict overly warm SST in the eastern Pacific. During the forecasts, the thermocline becomes shallower and shallower, differences with observations can reach 45 m (Fig. 2d). Note that when the SST index is very well predicted, as for example in April 1983 for lead times equal to 12 months (Fig. 2b), the thermocline in Niño-W is three times shallower than the observed one (Fig. 2d). These results question the validity of forecasts based on the SST signal only. Other fields like the thermocline must be examined as well. The baroclinic ocean undergoes higher fluctuations via the 9-month mobile mode (see Mantua and Battisti 1995), which are not found in the SST. Note that even in this case where the oscillatory behavior of the thermocline differs from the SST, statistics for the thermocline sorted out per lead times appear successful: the correlation between the observed and the predicted time series remains positive and significantly larger than the 95% level of confidence for all lead times. Analyzing lags between thermocline, wind, and SST signals as time increases in a given forecast experiment and analyzing their full amplitude relative to the prescribed background without correcting for model drifts is more informative. Observations often exhibit a lead of the wind (and thermocline) relative to the SST during the warming phase, whereas the CZ coupled model does not. Indeed, instead of a lead, the CZ model simulates the wind peaks 1–2 months after the SST warm peaks (see Perigaud et al. 2000). This lag and the important role of the off-equatorial coupling (see section 2d) explain why the shallowest peaks of the thermocline are reached several months after the observed ones. The rises of the thermocline and SST warm growths happen in 1982, 1986, and 1990–91 earlier by 6–7 months than in the observations or in the initial conditions driven by FSU winds. They are associated with the erroneous location of the coupled westerlies along the equator and the role of the off-equatorial coupling between the wind curl and the thermocline (see below).

Analyzing time series only is not sufficient either, because indexes can be misleading when box averages cover regions where anomalies are not homogenous. The choice of the box locations is usually based on where the observed variations are the largest, but the signals simulated by the coupled model can have very different spatial patterns. The observed variability of the thermocline displacements is compared in Fig. 3 to the initial conditions simulated by the CZ model for the various procedures. The conditions obtained with LDEO1 agree to some extent with observations (Figs. 3a,b). Not surprisingly, the variability is much better reproduced by CZ.SL (Fig. 3c). Indeed the difference

between the latter and data is smaller than 5 m everywhere within  $5^\circ$  of the equator over the whole period. By contrast, the initial conditions of CZ.CPLIC (Fig. 3d) exhibit a minimum in the central equatorial Pacific where observations present a maximum. This node of variability associated with the fixed location of the equatorial wind anomaly is one of the deficiencies of the model when it is integrated in a coupled context (Perigaud et al. 2000).

The agreement between the CZ.SL predictions and observations holds for a short while, as illustrated in Figs. 4a–d for a 1-month lead. This is not the case for the CZ.CPLIC forecasts. Prior to the 1982–83 (Fig. 4a) and 1991–92 events (Fig. 4d), the latter predict a thermocline that is severely tilted with a cold reservoir in the western Pacific, opposite to the observed warm. Actually during the warm events themselves (Fig. 4c) or for increasing lead times (Fig. 4e), the model thermocline keeps on shoaling in the western Pacific well beyond the observed level. After the 1983 event (Fig. 4b), it is shallower than observations by 10 m on both sides of the equator. The shoaling of the model thermocline all along the equator is systematically found after warm events during coupled simulations (Perigaud et al. 2000). Note that it is shallow in the eastern Pacific in 1983 while the SST is still warm. Such a situation has never been observed during warm events, but it systematically happens in CZ after warm peaks, because the model assumes very little change of the subsurface temperature in case of thermocline shoaling in the eastern Pacific, the SST being mostly influenced by the anomalous surface current convergence. By contrast, the observed thermocline is shallow in the central Pacific and downwelled in the western Pacific while the observed SST is already back to normal during this period. The unrealistically cold reservoir initiated by the LDEO2 procedure after 1983 certainly contributes to reduce the number of growing warm events predicted by CZ.CPLIC (Fig. 4f). For the CZ.SL forecasts, the growth mode of warm events dominates and rapidly eliminates the good initialization: after 6 months, the predicted thermocline does not look at all like the observed one, but is very typical of the period simulated after warm peaks, with shallow positions in the western Pacific, and a pronounced slope reversal in the far eastern Pacific where the wind blows eastward. The latter corresponds to erroneous zonal wind anomalies along the equator.

### c. Zonal wind along the equator

Conversely to a common belief that the atmosphere has little memory, its initialization plays a key role for the CZ model because one of the nonlinearities of the coupled system comes from the initial guess provided to the atmosphere component to solve the iterative moisture convergence feedback (see Zebiak 1986). For amplitudes of the SST Niño-3 index greater than 0.1, the

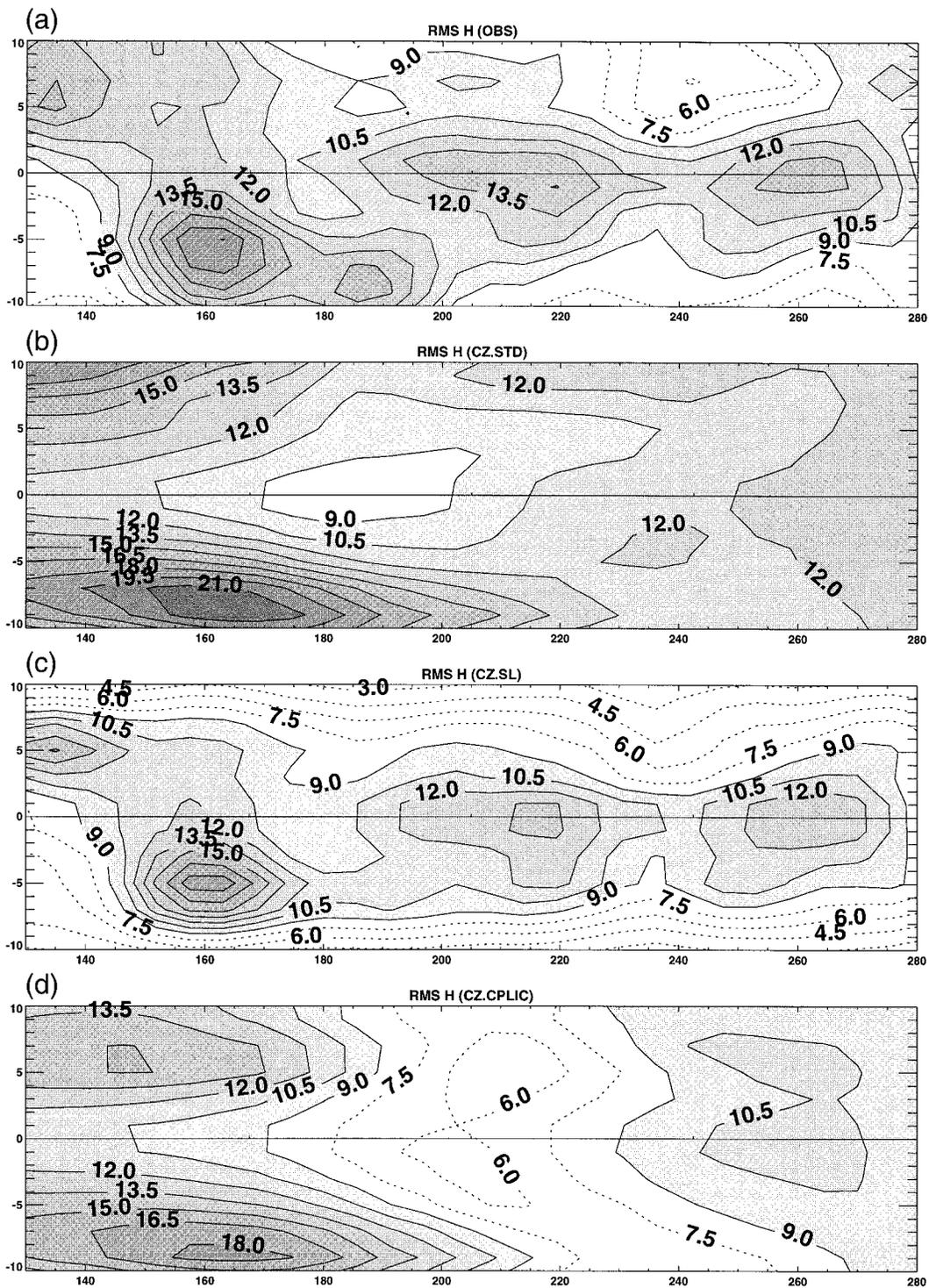


FIG. 3. Maps of rms variability of the thermocline depth in m over the period 1980–93: (a) derived from data or initialized by (b) the LDEO1 procedure (CZ.STD), (c) the data assimilation method 1 described in appendix A (CZ.SL), and (d) the LDEO2 procedure (CZ.CPLIC).

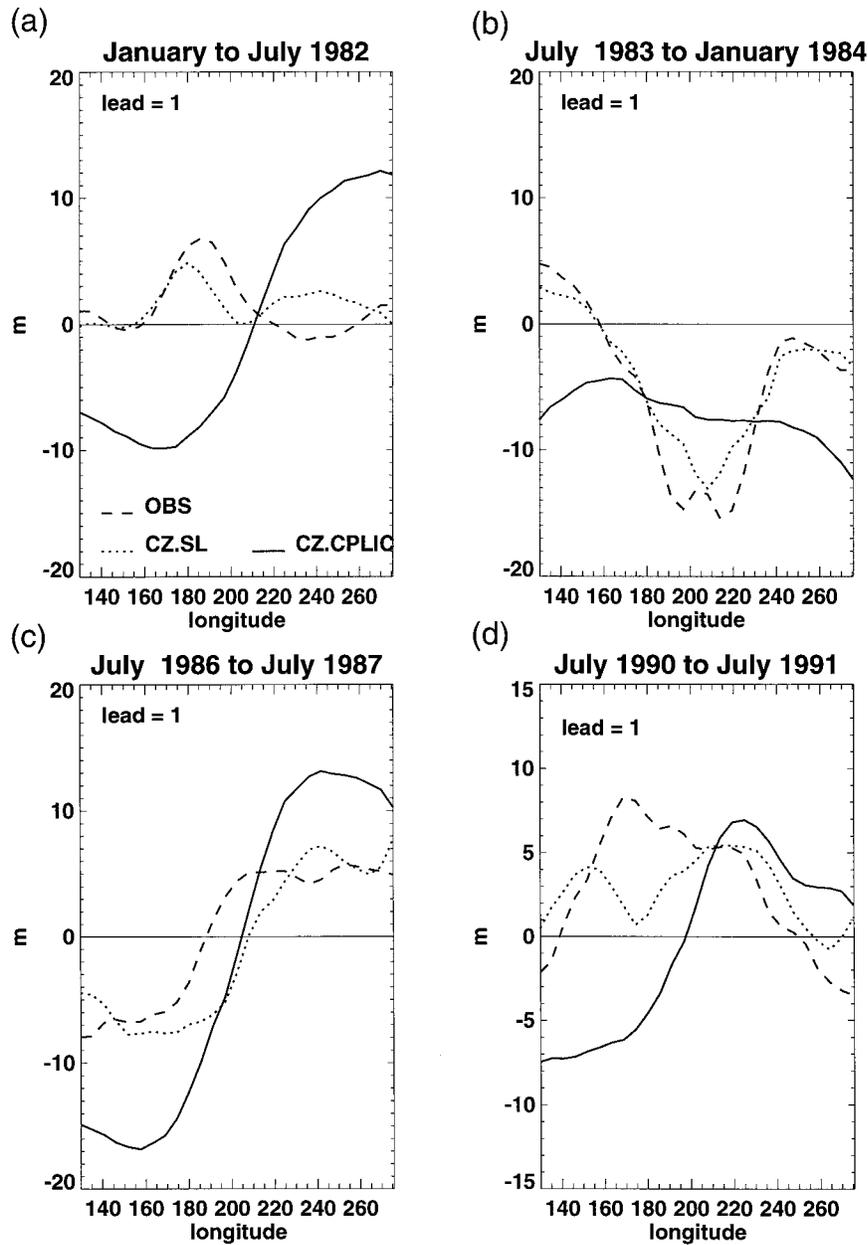


FIG. 4. Thermocline depth in m as a function of longitude on average between 1°S and 1°N derived from observations (dashed), CZ.SL (dotted), or CZ.CPLIC (solid). Plots in (a)–(d) correspond to a lead time equal to 1 month, plots in (e) and (f) to 6 month. Plots have been averaged in time over the period indicated in the title of each panel.

feedback term is initialized with the wind convergence of the previous time step, and, consequently, the CZ atmosphere is not a “slave” to the ocean state. This is a reason why the LDEO2 procedure can have a strong impact on the initial wind conditions, even though it induces small changes in the initial SST conditions. Figure 5 presents along the equator the zonal wind anomalies delivered at the initial conditions by the LDEO2 procedure together with two estimates derived from observations, namely FSU and Astat, the wind recon-

structed after the singular value decomposition (SVD) between SST and wind observations (see CP00). Astat winds will also be used in sections 3 and 4 of this paper. Using these two observed estimates is useful here as they contain different information, the Astat winds having less mesoscale high-frequency variability than the FSU winds, but also missing part of the low-frequency variability in the western Pacific. Wind anomalies presented in Fig. 5 are averaged over a minimum of 6 months in order to compare low-frequency signals only.

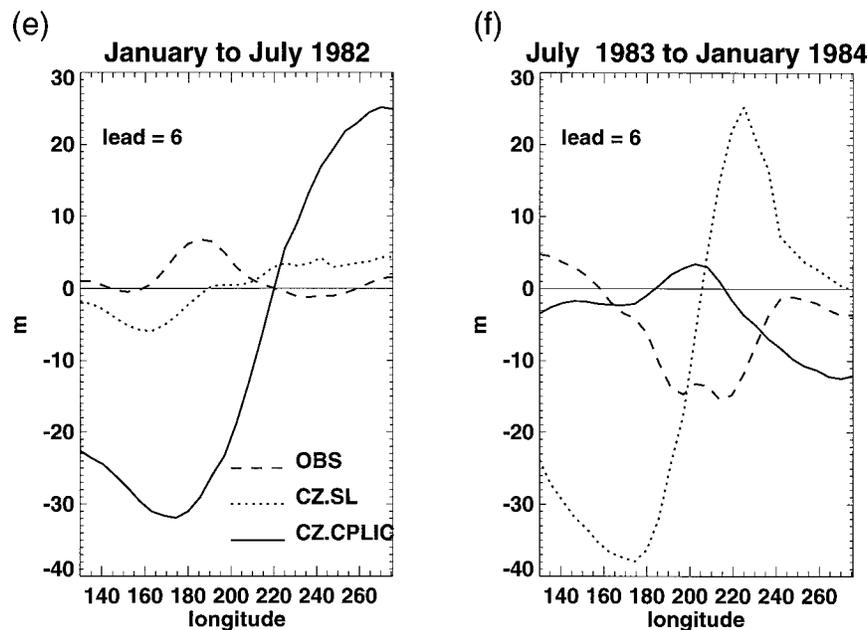


FIG. 4. (Continued)

Comparing the results indicates that the simulated wind is very different from observations, while the two observed estimates agree quite well. Even several months prior to the beginning of the observed warm event growth in 1982, the model wind coupled to the model SST exhibits strong westerlies in the central Pacific (Fig. 5a). Actually as early as January 1986 and October 1990 (not shown), the LDEO2 also simulates strong westerlies in the central Pacific, whereas the observed growth has not started yet. During warm events, the LDEO2 procedure simulates a wind that is blowing maximum to the east in the central Pacific (close to 160°W), but also strong easterlies in the far eastern Pacific (Figs. 5b,d,f). The latter systematically appear during the simulated warm events and play a key role in the coupled behavior of the model (Perigaud et al. 1997). Observations do not show such strong anomalies there. These model anomalies also explain the reversal of the thermocline tilt along the equator between the central and the eastern Pacific as in Fig. 4f. Systematically during cold events as in 1984 or 1988, the LDEO2 procedure simulates strong westerlies at the eastern boundary that are not realistic either (Figs. 5c,e). Note that the initial wind simulated by LDEO1 exhibits a similarly large mismatch with data (not shown). Nevertheless, it is worth analyzing the LDEO2 wind as in Fig. 5 rather than the LDEO1 wind, because the former is the wind consistent with the LDEO2 oceanic conditions and it partly explains the thermocline mismatch with data presented in Fig. 4. However, it is necessary to examine the off-equatorial anomalies to fully understand the mechanisms responsible for the successes and failures of the CZ forecasts.

*d. Off-equatorial wind and thermocline*

Because the off-equatorial anomalies were found to play a crucial role in the coupled model behavior (see Perigaud et al. 2000), the thermocline and wind initiated by LDEO2 are now examined away from the equator. More specifically, warm SST conditions are associated with westerlies along the equator and easterlies beyond 9° of latitude, the latter being more than three times stronger than the strongest anomaly ever observed there. As the model wind is more trusted along the equator than away from it, the LDEO2 procedure applies a nudging coefficient that gives less and less weight to the model wind as latitude increases. However, the coefficient is larger than 0.45 everywhere, and because the model wind is so strong outside the equatorial band, the nudged result is dominated by the model and not the observed signal there. Figures 6a,b presents the zonal wind stress anomalies averaged between 8° and 12°N during 5 months centered around the warm events. In 1983, the observed wind has weak westerlies in the eastern Pacific, whereas the LDEO2 wind is very strong with the opposite sign. At that time, the wind simulated by CZ.STD has the same wrong sign as LDEO2, with an intensity as large as 0.8 dyn cm<sup>-2</sup> (see the thin plain lines in Fig. 6a). The LDEO2 wind does not even have the observed sign then. Note that because the atmospheric model is not linear and depends on the model wind estimated at the previous time steps, the LDEO2 wind is not accurately obtained by simply applying the FSU and the CZ.STD winds with their respective nudging ratios at a given time. Consistently with the coupled behavior of the model, the same scenario is reproduced

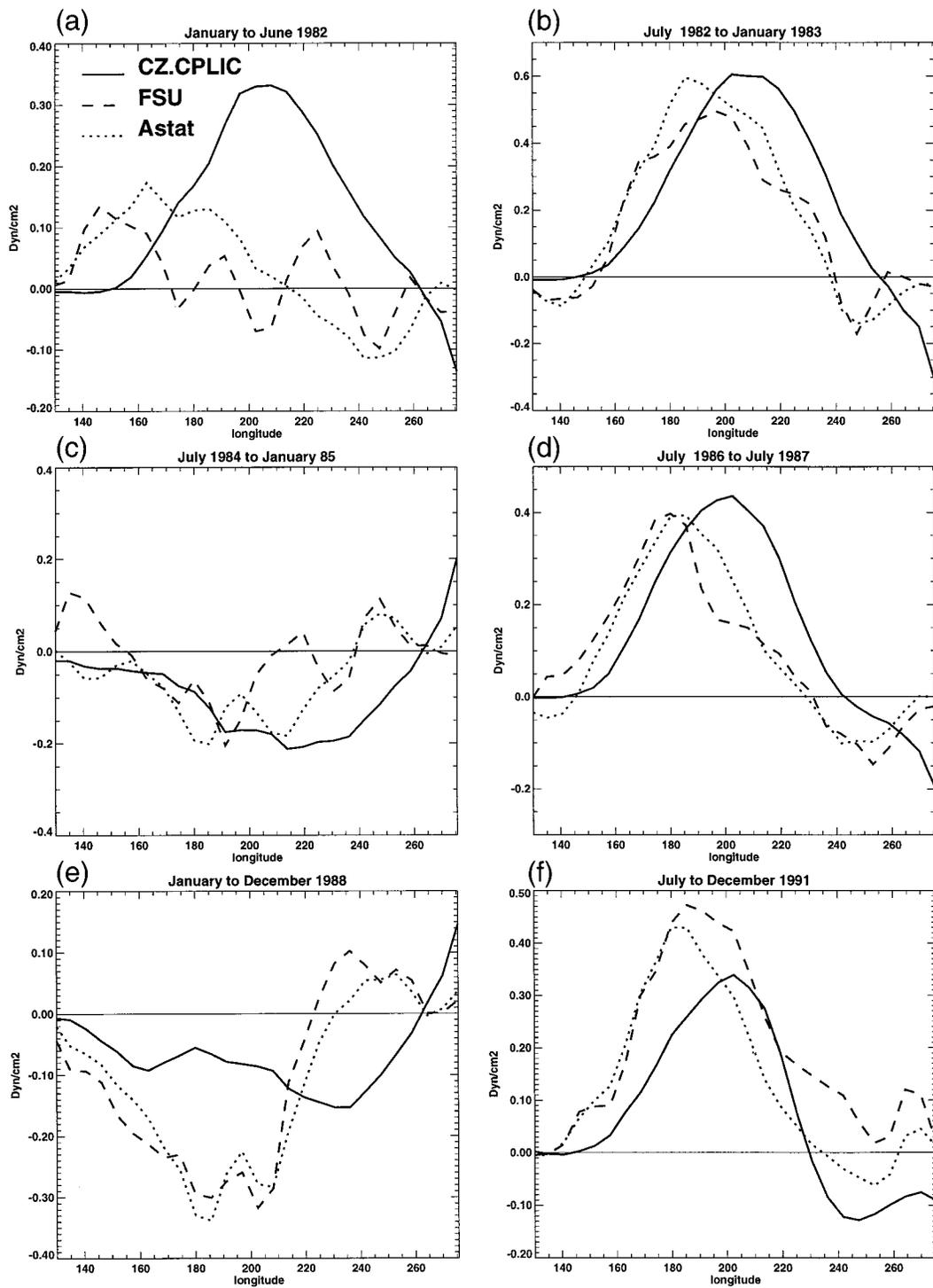


FIG. 5. Zonal wind stress anomalies observed by FSU data (dashed) or by Astat wind (dotted) or simulated by the LDEO2 procedure for initializing the forecasts (solid) as a function of longitude along the equator. Model and data are averaged in time over the months indicated in the title of each panel.

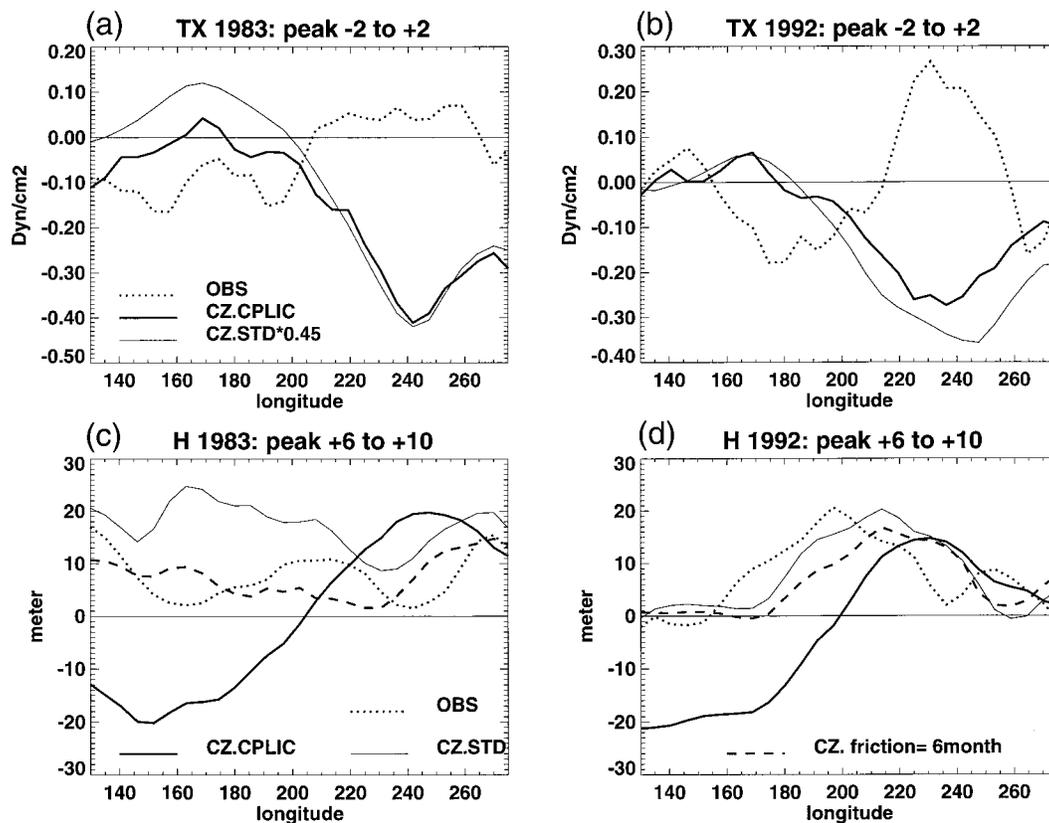


FIG. 6. Comparison between observations (OBS) and initial conditions obtained by the LDEO1 (CZ.STD) or LDEO2 procedures (CZ.CPLIC) as a function of longitude for off-equatorial bands. (a), (b) Zonal wind stress averaged between 8° and 12°N and during 5 months centered on Jan 1983 or 1992. The values for the CZ.STD solid thin curve have been multiplied by 0.45, which corresponds to the value of the nudging coefficient applied in LDEO2 at these latitudes. (c), (d) Thermocline depth averaged between 6° and 10°N and between Jul and Oct 1983 or 1992. The dashed line corresponds to the CZ model with a friction equal to 6 months instead of 30 months in CZ.STD or CZ.CPLIC.

for all warm events. Thus, in 1992, although the observed wind shows strong westerlies, the LDEO2 procedure simulates strong easterlies (Fig. 6b).

Associated with the strong westerlies blowing at that time along the equator, the off-equatorial easterlies create a very strong cyclonic curl (the meridional wind stress is relatively much weaker than the zonal one in CZ). By “quasi Sverdrup” balance, the pair of cyclonic curls on both sides of the equator is associated with shallow positions of the thermocline in the western Pacific. Figures 6c and 6d present the thermocline positions averaged over 6° and 10°N during 6 months after the observed warm peaks in 1983 and 1992, respectively. As in the multidecadal coupled experiments with the CZ model, the thermocline anomalies simulated by the LDEO2 procedure for both events is tilted to very shallow positions in the western Pacific. This is not specific to the north. Actually the shallowest of all the thermocline forecasts is found in the southwestern Pacific (see Perigaud et al. 2000). But in the northwestern Pacific, the observed and simulated thermocline do not even have the same sign. In fact the observed thermocline is deeper than normal over the whole domain

in the north at that time. Note that the thermocline simulated by LDEO1 is downwelled as in reality. The fact that the latter is deeper in 1983 than the observed one is due to the weak baroclinic friction, as demonstrated when the LDEO1 procedure is applied with a friction equal to 6 months instead of the standard 30 months (dashed lines in Figs. 6c,d). The sign of the CZ.CPLIC thermocline in the off-equatorial western Pacific does not even match the observed one. The LDEO2 procedure has considerably degraded the initial conditions for the off-equatorial wind and thermocline. It has built up a cold reservoir in the western Pacific that compensates for the warming tendency of the model during forecasts. One can now understand why nudging the SST instead of the wind is not successful in improving forecasts (Chen et al. 1997): off-equatorial model SST anomalies do not differ much from the observations whereas off-equatorial winds do.

*e. LDEO2 with sea level data: The LDEO3 procedure*

Recently the LDEO2 procedure has been completed with an assimilation of sea level data in a procedure

named LDEO3, which has delivered good statistics for Niño-3 indices over 1975–97 (Chen et al. 1998). In 1996–97, the sea level data have contributed to a reduction in the cold SST bias that LDEO2 is erroneously producing (see Fig. 2 of Chen et al. 1998). Prior to 1996, differences between the LDEO2 and LDEO3 initial SST are small. As the LDEO3 nudging for the wind is the same as in LDEO2, erroneous features in the wind similar to the ones described above are reproduced by LDEO3. This is also the case for the 1997–98 event, which reproduces the observed warm SST Niño-3 index. Then the LDEO3 conditions have strong off-equatorial easterlies and thermocline shoaling. The fact that these conditions with erroneous thermocline positions are obtained, although sea level data have been assimilated, is explained by the choice of the LDEO3 nudging done in the baroclinic model. It gives very little weight to sea level data away from the equator. The weight drastically decreases from 80% at the equator with an  $e$ -folding scale of  $2^\circ$  in latitude, meaning that at  $5^\circ\text{N}$  or  $5^\circ\text{S}$  and beyond, the LDEO3 procedure trusts the data to less than 1% and the model thermocline to more than 99%. By contrast, the nudging done for the wind gives more weight to data than model when distance from the equator increases, but the model wind is still retained up to 45% at  $7^\circ$  of latitude or beyond as in LDEO2. Therefore, sea level data in the LDEO3 procedure do not correct the off-equatorial thermocline anomalies that are erroneous, because they are associated with the erroneous model wind. One can expect that the thermocline initialized by LDEO3 in late 1997–98 is far from reality in the western off-equatorial Pacific. Although the LDEO2 and LDEO3 procedures have greatly reduced the shock at the initialization of forecasts, they have severely degraded the oceanic initial conditions in comparison with the CZ.SL or the LDEO1 ones. These results support the need for improving at first the model, before increasing the consistency of the initial conditions with the coupled model.

Kleeman et al. (1995) are able to improve forecasts by assimilating observed thermocline depth anomalies into a coupled model of similar complexity to CZ. Our results illustrate that conclusions relative to the use of data to improve forecasts strongly depend on the model. Indeed the thermocline in their model has a more important impact on the SST because the SST changes explicitly depend on the thermocline displacements, they do not depend on baroclinic nor Ekman nor climatological nor anomalous currents. The thermocline has a more important impact on the wind also because their atmosphere, conversely to CZ, is effectively a slave to the ocean state via SST. Bennett et al. (1998) assimilate thermocline depth and wind data into a model that has similar physics to CZ, the ocean being the same and the atmosphere being the CZ model without the iterative convergence feedback term. They point out some model data inconsistencies that suggest that this type of ICM misses some physics. It is certain that a model with only

one vertical baroclinic mode for the ocean and the atmosphere and a fixed mixed layer depth lacks a lot of the physics that play an important role in El Niño. Nevertheless, however simple a model is, very different behaviors can be simulated for a given set of equations and climatological background, depending on which coupled mode is retained among those allowed by the parameterizations (Neelin and Jin 1993).

In summary, two ways of using data to improve forecasts have been examined. One consists of using more data to improve the initial conditions while the other consists of reducing the impact of data in order to increase the consistency of the initial state with the model. Both lead to very limited success. Rather than using data for the initial conditions only, a third option consists in using them to improve the model itself. As proposed in DP96 and CP00, data provide information to revise the parameterizations and allow for reduction of the model deficiencies in a forced context as well as during multidecadal coupled simulations. Let us now examine the forecasts obtained with the revised parameterizations.

### 3. Forecasts delivered by Tsub.Conv

The Tsub.Conv model is the CZ model where the parameterization of the subsurface temperature is modified according to oceanic measurements and where the iterative convergence feedback is replaced by a parameterization of convective heating derived from observations (see DP96 and CP00). Conv is an atmosphere model that resembles the one developed in Kleeman (1991) where the wind stress anomaly is diagnostically derived from SST. As recommended in Perigaud and Dewitte (1996) and unless specified otherwise, the friction applied in the baroclinic model is of the order of 3–6 months so that the thermocline variability driven by observed winds beyond  $5^\circ$  of latitude is reduced to a level close to the observed one (see Figs. 6c,d). Tsub.Conv is first initialized as in LDEO1, starting with the CZ initial conditions of December 1979. The Tsub.Conv initial states obtained since 1980 are slightly improved as the 1988 cold event is then well reproduced (see DP96), and the thermocline and wind are the most improved signals (see Table 2). Various series of forecasts are generated to test the sensitivity to the parameters that have been investigated in CP00. The reader can refer to Table 1b for a list of the experiments presented here.

#### *a. Characteristics of forecasts initialized as in LDEO1*

Figure 7 compares Tsub.Conv1 time series of SST and thermocline at a 6-month lead with data and CZ.STD forecasts. Results show that the Tsub.Conv model is not biased toward warm SST in the east nor upwelled thermocline in the west as CZ. The reduction of this bias corresponds to one of the two major im-

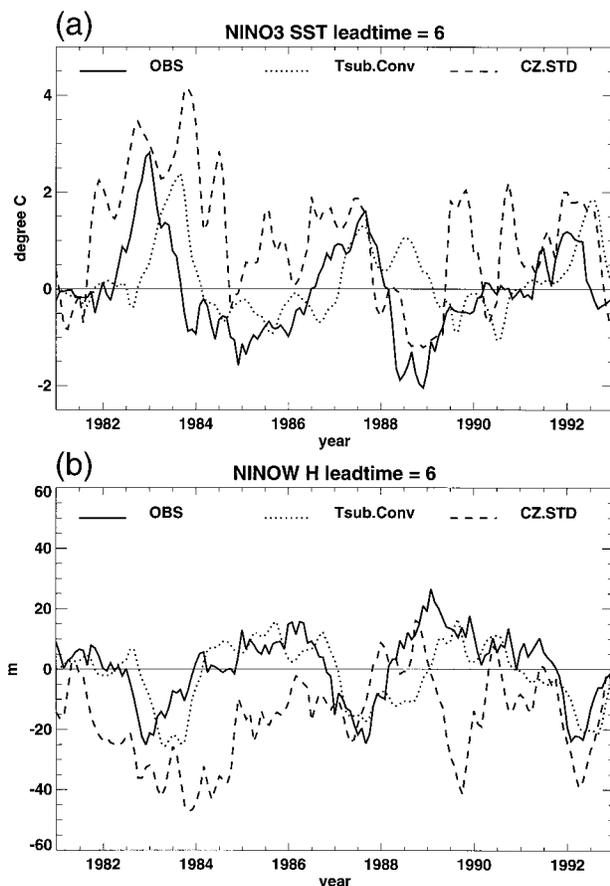


FIG. 7. Time series of the Niño-3 SST and the Niño-W thermocline between Jan 1981 and 1993 from data (solid), or from 6-month lead time forecasts simulated by CZ.STD (dashed) or by Tsub.Conv (dotted).

provements achieved by Tsub.Conv in comparison with CZ in simulating realistic coupled anomalies (CP00). The other expected improvement is relative to the wind whose maximum along the equator should be better located. The zonal wind variability over 1980–93 is presented for observations (Fig. 8a) or forecasts with a 6-month lead for CZ.STD (Fig. 8b), CZ.CPLIC (Fig. 8c), or Tsub.Conv (Fig. 8d). The observed reference used here is the Astat winds (Fig. 8a). It will be used again in this section and the reader can verify in Perigaud et al. (2000) that there is little difference with the corresponding map derived from the FSU winds. Tsub.Conv forecasts have the wind pattern that agrees the best with observations. The wind maximum is located to the west of the other forecasts, only  $10^\circ$  to the east of the date line, where it is observed. Note that it is possible to correct for the weak amplitude of Tsub.Conv by modifying the weight of the convective heating in the atmosphere (see CP00). This is addressed later in this section.

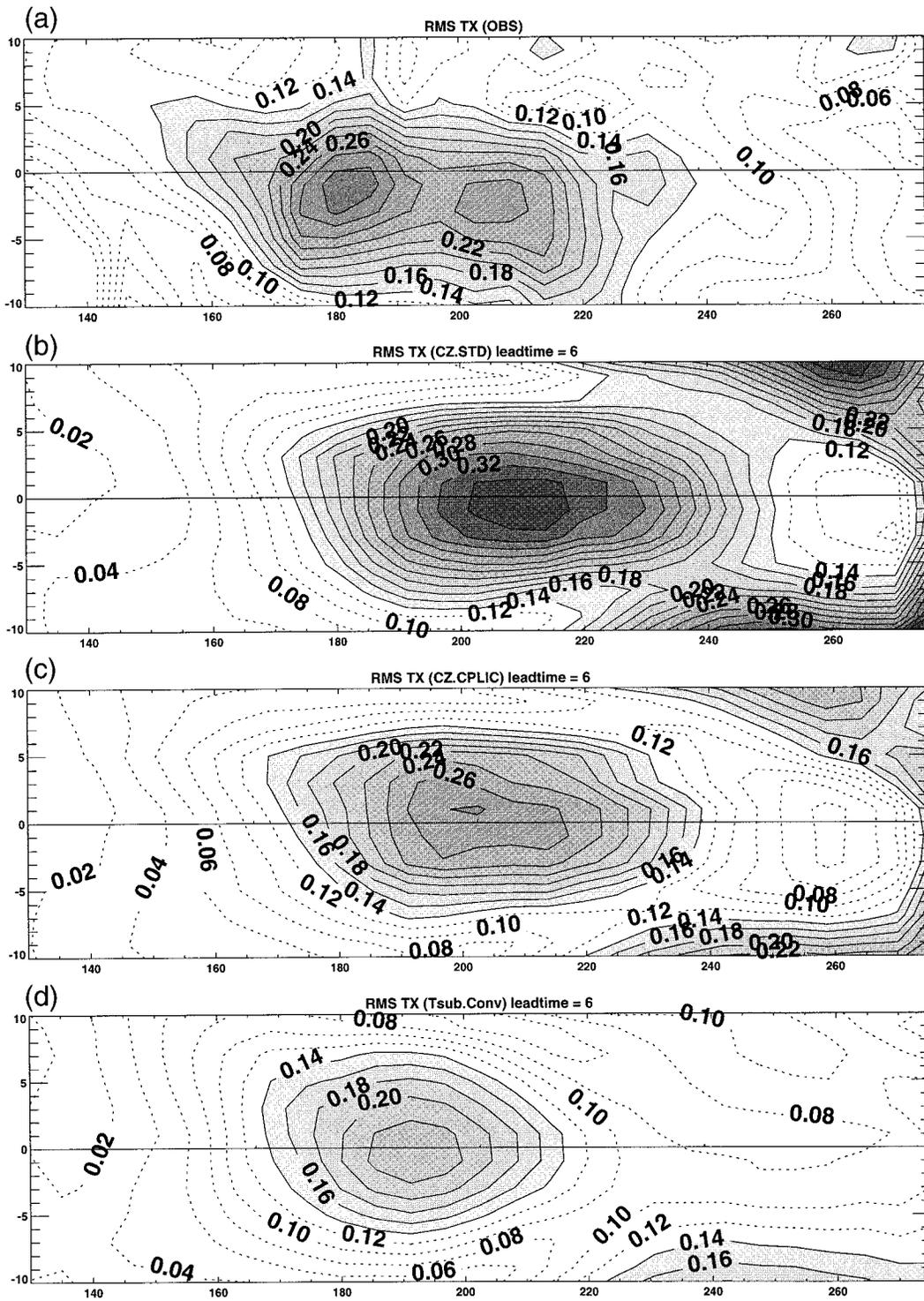
Although the errors are much smaller for Tsub.Conv than for CZ at 6-month lead, they are still large. The Tsub.Conv1 series correspond to forecasts that decay to

the “quasi normal” state when the lead time increases (Fig. 9a). The spaghetti diagrams illustrate that Tsub.Conv1 misses the growth of all warm and cold events and obviously has a very poor predictive skill. Note that the conventional statistics for this series (not shown) indicate an apparently good performance with errors much smaller than for CZ.STD.

Tests are then done with a weaker friction (i.e., larger time decay) in the baroclinic component so that the coupled model exhibits an oscillatory regime (this is the case for decay times larger than 12 months). Various weights given to the convective heating relative to the local heating in the atmosphere are also tested because the weaker the weight, the bigger the intrusion of westerlies in the eastern Pacific, and the stronger the warm events (see CP00). As anticipated, Tsub.Conv forecasts are found very sensitive to the choices of these coefficients. The Tsub.Conv2 series presented in Fig. 9b has a 24-month friction and a convective weight of 10%. Forecasts then do predict the 1982–83 and 1986–87 warmings. They also succeed in predicting the 1988–89 cooling and reversal. These apparently successful predictions hold up to about 1-yr lead time. But the reader can see that Tsub.Conv2 has a tendency to predict systematic warm growth like CZ.STD. Results from CP00 allow one to explain that it is for the wrong reasons that the forecast indices appear good in 1982–83 and 1986–89. Both CZ with a 30-month friction and Tsub.Conv with a 24- or even a 12-month friction develop large and erroneous off-equatorial anomalies of wind and thermocline that play a key role in the coupled behavior.

#### b. Forecasts initialized with sea level data

The fact that the prediction of the Niño-3 SST index depends on the off-equatorial ocean strengthens the need for a good ocean initialization away from the equator as much as in the equatorial band. Tsub.Conv2.SL is obtained with the same parameterization as Tsub.Conv2 but has been initialized with sea level data. The procedure corresponds to the second method described in appendix A that provides a better initialization of the baroclinic ocean beyond  $5^\circ$  of latitude in comparison with the first method used in CZ.SL. The impact of the sea level data on the initial SST anomalies is small (Fig. 9c), except in 1982 and late 1989. Overall, the use of sea level data does not significantly reduce the initial error for the SST nor the wind (see Table 2). Except for those initiated in 1982 or late 1989, Tsub.Conv2.SL predictions resemble a lot the previous ones. The sea level data have a much smaller impact on Tsub.Conv2 than on CZ (cf. CZ.STD and CZ.SL). This is because the initial conditions, in particular the wind initialized by Tsub.Conv2 or Tsub.Conv2.SL, resemble a lot more FSU winds than CZ.STD or CZ.SL. The most important step to achieve is to reduce the growth of the unstable coupled mode, possibly by reducing the discontinuity



between the forced and the coupled context, which appears so strong during interevents.

*c. Forecasts initialized with a more “coupled” procedure*

In order to reduce the initial shock undergone by forecasts, various initialization procedures are tested. The two presented now, Tsub.Conv3 and Tsub.Conv3.CPLIC have been obtained with a 6-month friction and with the atmospheric convection weight fixed to the value of Tsub.Conv1, because these parameters are the most consistent with observations.

Effort was put on increasing the consistency between the observed and the model wind as much as possible. Instead of using FSU data to force the ocean component, the wind comes from the Astat model in order to retain the variability associated with ENSO. In addition here, because the second SVD wind mode of the Astat model has its maximum amplitude in the eastern equatorial Pacific and favors the coupled mode growth (see CP00), only the first SVD mode is now retained to initialize the model. Tsub.Conv3 and Tsub.Conv3.CPLIC forecasts are initialized with Astat winds following the LDEO1 and the LDEO2 procedures, respectively.

The Tsub.Conv3 forecasts are presented in Fig. 9d. Forcing the model with the first SVD wind instead of FSU has a slight beneficial impact on the initial SST and wind (cf. with the dashed lines in Figs. 9a,b and see Table 2) by removing the erroneous cold feature in 1989 as well as correcting the overly warm second peak in 1983 (the 1989 feature is also corrected when all the SVD modes are retained to force the model; see section 4). It is verified that the initial wind of Tsub.Conv3 agrees very well with the Astat wind (see Table 2). However different the initial conditions are, it is striking that most of the Tsub.Conv3 forecasts resemble Tsub.Conv1 with big initial shocks and poor predictive skill. The model has a tendency to predict decaying warm events when it should predict growing warm events as in 1982, 1986, and 1992; growing warm events when it should predict the decays of warm events as in 1983 and 1992; or the growth of the cold event as in 1988. This suggests that the high-frequency finescale features of the FSU wind anomalies that could a priori be a source of error growth are not.

Finally, the LDEO2 procedure was applied to Tsub.Conv and this wind with the exact same nudging coefficients and interpolation schemes as in CZ.CPLIC. Figure 9e illustrates that the procedure does not affect much at all the initial conditions nor the forecasts. The initial shock is certainly less brutal when the LDEO2 procedure is applied (cf. with Fig. 9d), but the impact of the procedure is very small on the initial conditions and for all lead times. The statistics for Tsub.Conv3.CPLIC are undistinguishable from the ones given in Table 2 for Tsub.Conv3. It is quite remarkable that the impact of the LDEO2 procedure is so small on

Tsub.Conv, whereas it is so large on CZ. It is clear that the LDEO2 procedure can significantly affect the initial conditions only if the model wind is drastically different from reality, which is not the case for the Tsub.Conv model.

In summary, the use of data to revise the CZ model into the Tsub.Conv model has considerably reduced the errors in forecasting the SST Niño-3 index (cf. Fig. 1a with Figs. 9a or Fig. 9d). Compared to CZ, the initial conditions and the forecasts during the first 6 months are much closer to the observed thermocline, wind, and SST anomalies over all the Pacific within 15° of the equator. Tsub.Conv forecasts are not much sensitive to changes in the initial conditions because the coupled wind and ocean resemble much more the observed ones than those simulated by CZ. With increasing lead time, predictions decay to the normal state when the model has a friction of the order of 3–6 months. This is not the case with weaker frictions, but whatever its value or that of the atmospheric convective weight, Tsub.Conv skill is limited because of its erroneous mode of growth located in the eastern Pacific. Although the winds may not be the only model deficiency in the eastern Pacific, replacing the Conv by the Astat atmosphere is worth the investigation, because the coupled mechanisms developed by Tsub.Astat are quite different (see CP00).

#### 4. Forecasts delivered by Tsub.Astat

Because this study is focused on predictions, the data covering the period that is going to be predicted should not be taken into account in the statistical atmosphere. So, the Tsub.Astat model presented in CP00, which is defined with statistics computed over the period 1970–95, cannot be used here. This section first gives the characteristics of the Tsub.Astat model used here. It then reports the SST forecast series obtained by this model for different initialization procedures and parameters (see Table 1c). Finally it analyzes in detail the various predicted anomalies to examine the extent and limitations of the apparent success.

*a. Characteristics of Tsub.Astat model used for forecasting*

It was clearly identified in CP00 that the second SVD mode plays a key role in the oscillatory regime of Tsub.Astat. Decomposed over the 1970–95 period, the amplitude of the second SVD wind mode reaches its maximum at the 1982–83 event. When the SVD decomposition is performed over 1970–80 only, Tsub.Astat does not reproduce an oscillatory regime, whereas it does with the decomposition over 1970–83. So rather than updating the SVD modes as time advances before each new forecasting year, as done in Syu and Neelin (2000), it is decided for simplicity to fix them to those obtained over 1970–83.

We verified that the coupled behavior is then similar

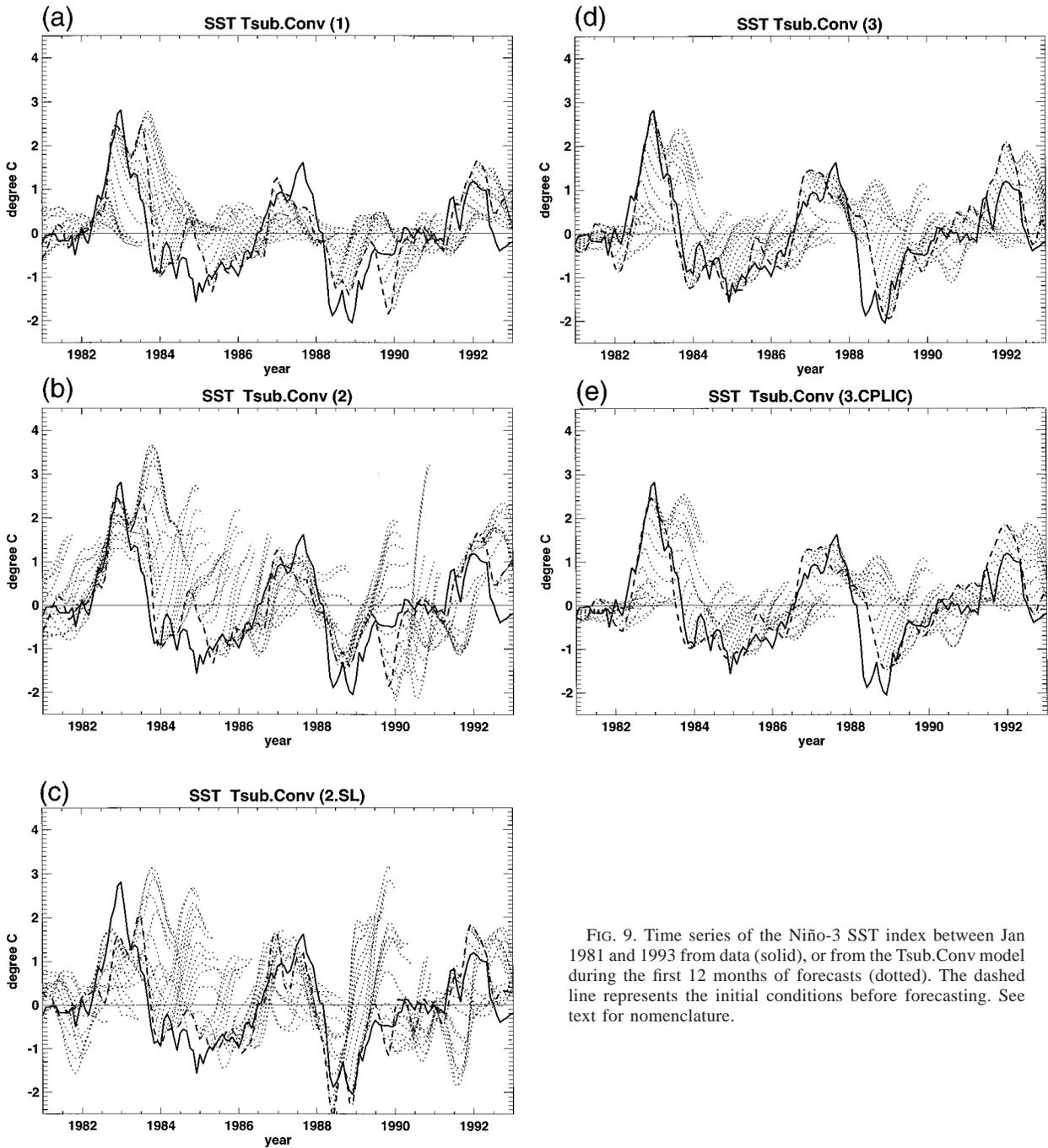


FIG. 9. Time series of the Niño-3 SST index between Jan 1981 and 1993 from data (solid), or from the Tsub.Conv model during the first 12 months of forecasts (dotted). The dashed line represents the initial conditions before forecasting. See text for nomenclature.

to the one presented in CP00. It is found in CP00 that the Tsub.Astat model is much less sensitive than Tsub.Conv to changes in parameter value for the friction or drag coefficients. It is also explained that the simulated wind stress and baroclinic anomalies in the north have signs opposite to the ones simulated by Tsub.Conv during and after warm peaks, in agreement with observations. In the rest of the paper, the friction is fixed to 6 months. Tsub.Astat then does not oscillate during de-

cade-long coupled experiments; it simulates a quasi-steady state around either warm, or cold or quasi-normal conditions after 3–4 yr of integration. This choice is nevertheless justified because it corresponds to the one that best fits observations with the forced simulations, and results from the last section indicate that failure in predicting is not necessarily due to the nonoscillatory behavior of the model. In addition, CP00 show that for a given friction, the fact that the model behavior is

oscillatory or not depends on the SVD itself, demonstrating that Tsub.Astat models can easily be switched from a stable to an oscillatory regime. Similarly, Penland et al. (2000) consider an ICM as a possible tool for ENSO forecasting even though the model resides in the stable domain, because ENSO-like behavior can be reproduced by adding stochastic forcing.

*b. SST forecast series delivered by Tsub.Astat*

In order to directly compare predictions that are initialized with the exact same oceanic conditions, the Tsub.Astat (1) forecasts are first initialized by forcing the ocean with the first SVD wind mode as in Tsub.Conv3. Tsub.Astat1 and Tsub.Conv3 forecasts are quite different (cf. Fig. 10a and Fig. 9d). Although there are still some cases during the decay of the 1983, 1987, and 1992 events that grow warm again, many Tsub.Astat forecasts predict the reversals and decays with some skill and the rms errors are smaller (see Table 2). In addition, Tsub.Astat1 does predict the warm growths past July 1982, past April 1986, and past May 1992; it also predicts the cooling and reversal past March 1988. All these forecasts are improved compared to Tsub.Conv3.

The choice is now made to present the series of forecasts Tsub.Astat2 obtained when the model is initialized by the wind reconstructed with all seven SVD modes used in Astat as in CP00. Several series were generated with the various parameters tested in CP00. Consistently with the changes explained in this reference, forecast series are slightly modified. Two of them are presented here. As the quality of these series is considerably improved compared to all the previous ones, forecasts are now presented up to December 1997 (Fig. 10b). The introduction of higher SVD modes, especially the second one, in the initial conditions contributes to a “boost” in the reversals and decays of events more clearly than in Tsub.Astat1. Similar success is also obtained with FSU winds (not shown). This is a big difference from Tsub.Conv, which simulates erroneous warm growth and overly warm initial conditions with FSU or the full SVD winds. Most of the individual Tsub.Astat2 forecasts agree with data to some extent. The decays of the warm events in 1983 and 1987, with the continuation of the latter into the cold event in 1988, are captured as well as the warming mild trends between 1989 and 1994. Note that Tsub.Astat produces erroneous warm anomalies during late 1989 and underestimates the growing warm events of 1983 and 1997. Similar errors are found in most of the other published forecasts (Rosati et al. 1997; Kirtman and Zebiak 1997; Syu and Neelin 2000).

Because it is possible with Tsub.Astat to simulate stronger warm events by slightly changing the value of parameters while keeping realistic amplitude patterns and behavior simulated in the tropical Pacific within 15° of the equator (see CP00), several experiments are per-

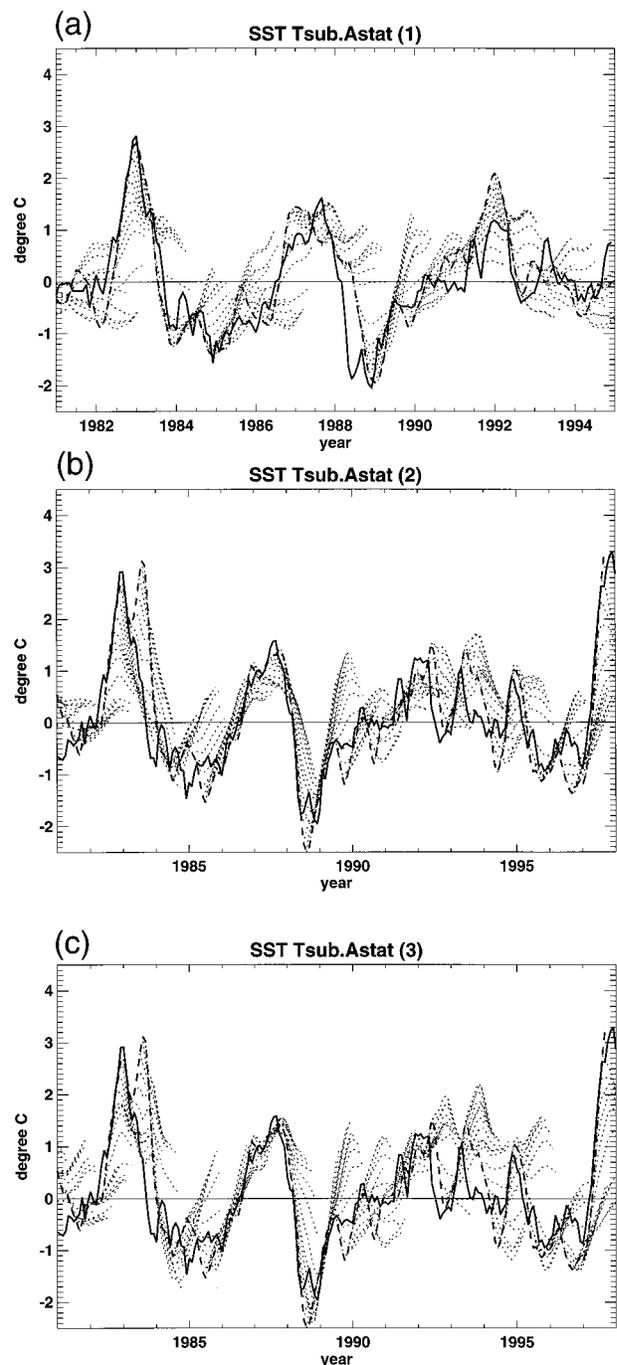


FIG. 10. Same as in Fig. 9 but for Tsub.Astat. Note that the two time series in (b) and (c) cover a longer period than (a), up to Dec 1997.

formed to test if the strength of the 1983 and 1997 events could be recovered without degrading the rest of the forecasting. Such a series with an increased coupling coefficient (Tsub.Astat3; see Table 2) exhibit similar behavior and skill to Tsub.Astat2 (Fig. 10c). The 6-month lead predicted growths of the 1982 and 1997

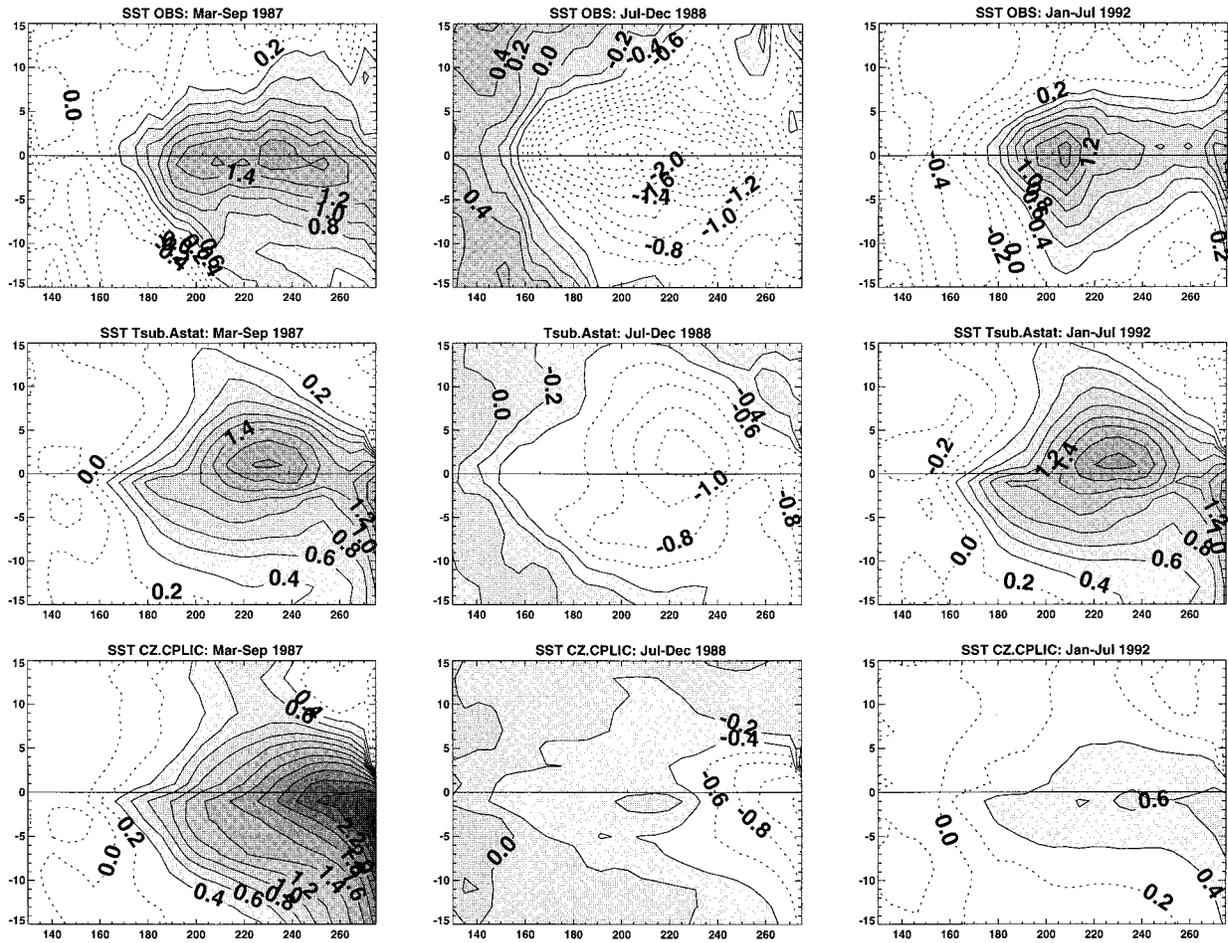


FIG. 11. Maps of SST anomalies in  $^{\circ}\text{C}$  derived from (top) data, (middle) Tsub.Astat, or (bottom) CZ.CPLIC forecasts with a 12-month lead time. Left maps are averaged anomalies between Mar and Sep 1987, middle between Jul and Dec 1988, and right between Jan and Jul 1992.

warm events agree better with observations, but they are still weaker and the model fails to predict these events in forecasts initialized prior to May 1982 and prior to March 1997, respectively. Tsub.Astat3 is no less skilled than the best published forecasts; the failure in predicting these events more than 6 months in advance deserves special attention (Perigaud and Cassou 2000). At this point, it is rather necessary to examine and explain the good performance of Tsub.Astat in forecasting the Niño-3 index for the 1984–93 period. Validation with data is performed below for the wind, the SST, and the thermocline anomalies over this period.

### c. One-year lead SST and wind forecasts

The 1-yr-lead Tsub.Astat3 forecasts are now compared with observations and with the 1-yr-lead CZ.CPLIC forecasts for the 1987 and 1992 warm events or for the 1988 cold event. The predicted SST fields are presented in Fig. 11. Tsub.Astat warm predictions in 1987 and in 1992 have an offshore and a coastal maxima

of about  $1.4^{\circ}\text{C}$ , whereas CZ.CPLIC predicts only the coastal maximum with an overly large amplitude close to  $2.6^{\circ}\text{C}$  in 1987. These patterns and amplitude correspond to the characteristics of the two models during long coupled simulations (see CP00). The observed SST anomalies do not have a coastal maximum in 1987, but they exhibit a two-peak pattern in 1992 in very good agreement with Tsub.Astat. The cold event in Tsub.Astat is too weak, but it is central as in observations and not coastal as in CZ.CPLIC. Note that these maps illustrate that indices such as the SST Niño-3 index can be misleading because of averaging effects.

The corresponding observed and predicted zonal wind stress anomalies are presented in Fig. 12. The westerlies in 1987 are predicted with a fairly good location and amplitude for Tsub.Astat, whereas CZ.CPLIC is much too strong and its maximum is located  $20^{\circ}$  to the east of the date line. Similar success in forecasting the amplitude and patterns of the wind is achieved by Tsub.Astat in 1992. More importantly, CZ.CPLIC predicts very strong easterlies in the eastern

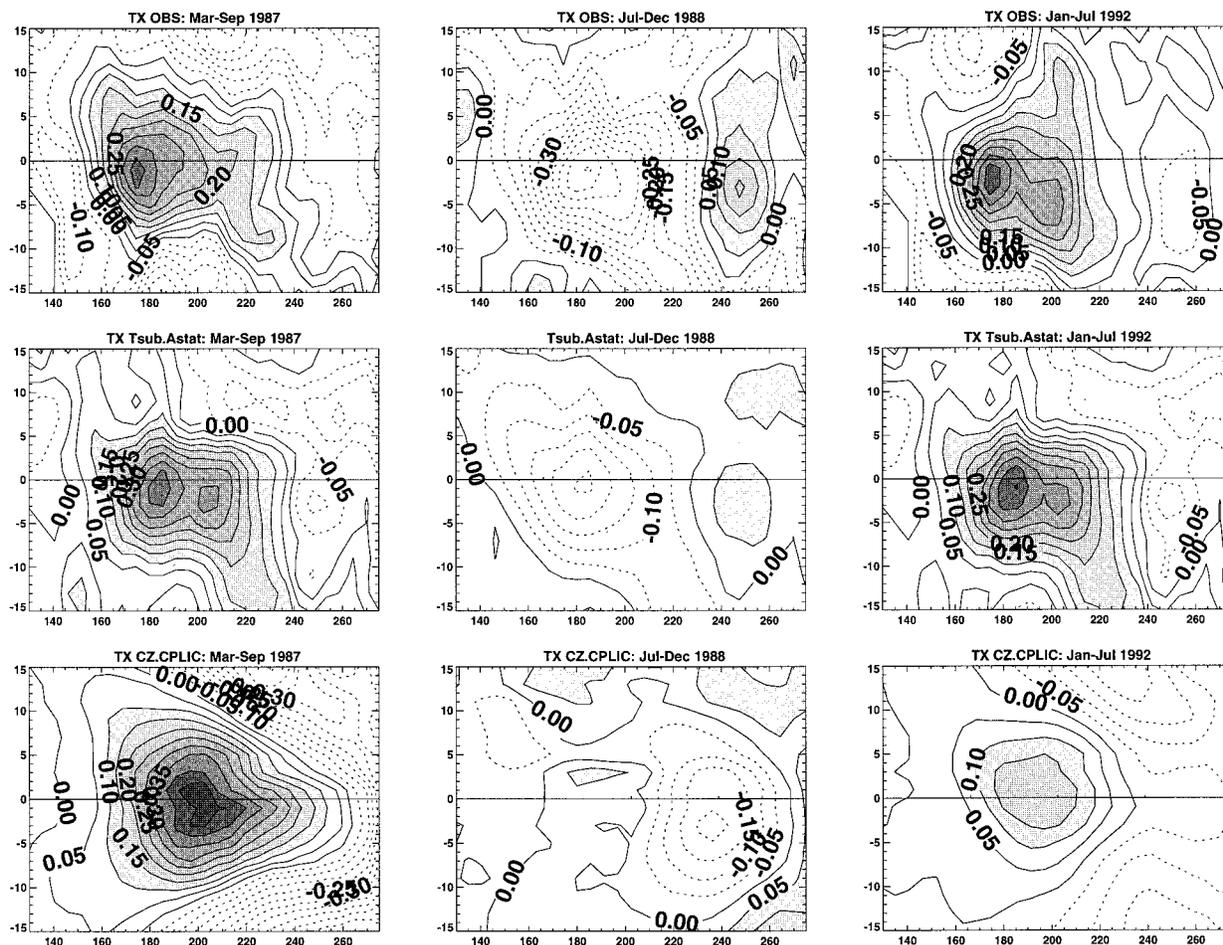


FIG. 12. Same as in Fig. 11 but for the zonal wind stress anomaly in  $\text{dyn cm}^{-2}$ .

Pacific, along the equator, as well as off-equator. This is not the case for observations. During the cold event, the wind pattern is very well predicted by Tsub.Astat. By contrast, the easterlies predicted by CZ.CPLIC are located  $60^\circ$  to the east of the date line and thus the predicted wind happens to have a sign opposite to the one observed there.

The meridional wind stress anomalies must also be validated as they play a strong role in the off-equatorial thermocline and coupled adjustments. The corresponding anomalies are presented in Fig. 13. It is striking that the meridional component predicted by CZ.CPLIC is weak, weaker than observed and much weaker than the zonal one. For Tsub.Astat, it reaches values that agree well with observations and that are larger than the zonal component in the north and central parts of the basin and in the southwest. Patterns and amplitudes related to the ITCZ and South Pacific convergence zone (SPCZ) displacements are remarkably well predicted by Tsub.Astat for both warm and cold events. By contrast, CZ.CPLIC winds are poorly located in the eastern Pacific.

*d. Meridional displacements of the predicted thermocline*

As explained in CP00, the oscillatory behaviors of both CZ and Tsub.Astat are sensitive to the charge and discharge of the oceanic heat content outside the equatorial band. This quantity can be monitored by the zonal average of the thermocline between the western and the eastern boundaries of the Pacific basin. It is analyzed for observations or forecasts as a function of latitude and time in Fig. 14. Because El Niño is a low-frequency signal, this quantity is coupled to the wind stress curl anomalies via the quasi-Sverdrup balance rather than via Ekman pumping. The sign of the curl in the  $5^\circ$ – $15^\circ\text{N}$  band happens to be given by the meridional wind anomalies for Tsub.Astat, and by the zonal ones for CZ. Thus warm peaks simulated in long coupled runs are followed by a charge (discharge) for Tsub.Astat (CZ) of the oceanic heat content in the north. Let us examine if forecasts reproduce these scenarios and how they compare with observations.

It is striking that the observed 1984–92 period is



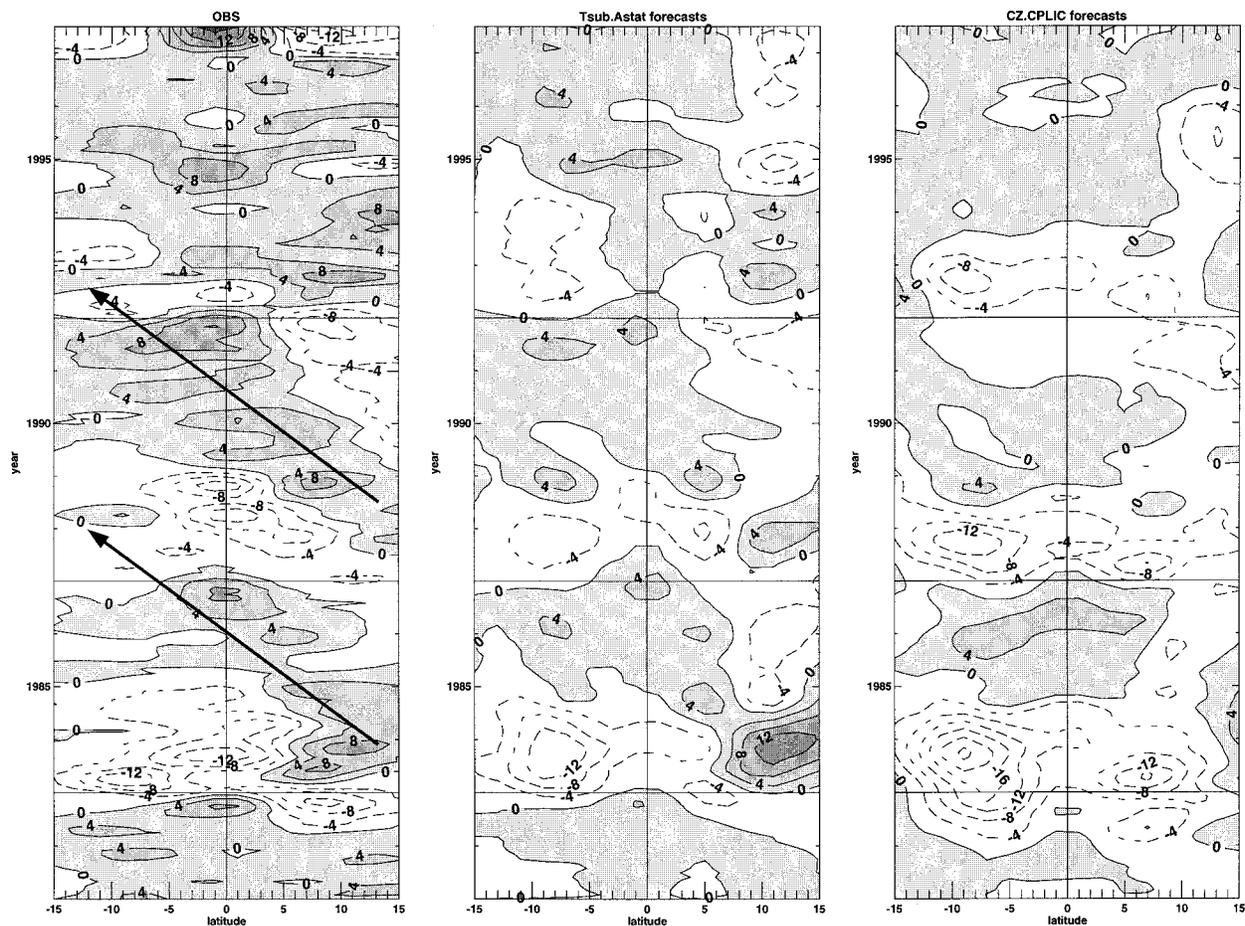


FIG. 14. Thermocline anomaly (in m) zonally averaged over the Pacific as a function of latitude and years between Jan 1981 and Dec 1997. Left panel corresponds to observations, middle to Tsub.Astat, and right to CZ.CPLIC ensemble forecasts for lead times between month 1 and 12.

mates of data uncertainty. More specifically, the wind anomalies beyond 5° of latitude are very strong with signs opposite to the observed ones, and are associated with very shallow thermoclines in the western Pacific that are not observed. The LDEO3 assimilation of sea level data applied in addition to the LDEO2 wind nudging does not correct the thermocline nor wind errors because the model thermocline is given a much stronger weight than the data everywhere poleward of 1° from the equator. We regard these as cases where the procedures or the model should be revised. Using sea level in addition to wind to improve the LDEO1 initial conditions does not work either. The CZ model has a tendency to systematically predict warm events, associated with very strong wind cyclonic curls in the off-equatorial eastern Pacific and very shallow thermocline in the western Pacific. Given a poor model, there is a limit of what initialization can do. The ultimate improvement in forecast skill is to improve the model.

The impact of reducing model deficiencies is illustrated here with ICMs without adding physics to the model. A number of caveats emerge from such simpli-

fied physics. For instance, the oceanic vertical structure simplified with only one baroclinic mode does not adequately represent both the equatorial and off-equatorial thermocline displacements and renders the choice of the friction more critical than in a multimode ocean. The mixed layer has a constant depth and surface currents as well as their convergence may contain large errors. Similarly to the large sensitivity of the coupled simulations explained in CP00, forecasts are very much dependent on the values of anomalous and climatological upwelling. Keeping a simple modeling context in mind, it is worth adding vertical modes in the baroclinic ocean as in Y. Chen et al. (1995) and Dewitte (2000). The Gill (1980) type of atmosphere, whether it is in the CZ or Conv version, has a serious flaw in generating spurious winds in the eastern Pacific. As long as this type of atmospheric dynamics is kept, the model will have this flaw. Refining the parameterizations or adding mechanisms to drive the winds farther away from the coast cannot fix the problem in the east. For coupling to simple ocean models, alternatives include intermediate atmosphere models (Wang and Li 1993; Neelin and Zeng

2000) or atmospheric general circulation models (AGCMs) (e.g., Kirtman and Zebiak 1997). In this study, the dynamical atmospheric model is replaced by a statistical one. This choice was done because coupling a given ocean model with a new atmosphere is a multiyear effort, even in the case of intermediate or statistical atmospheres. Another reason is that the coupling with Astat reproduces mechanisms outside the equatorial band that agree reasonably well with observations and that have signs opposite to the ones simulated with Gill's atmosphere. Whenever a coupled model has some statistics built into it, it is common to hear that the limited success in forecasting is due to the statistical component of the model. Results in this paper highlight that this is not the only source of limitations. The dynamic model itself needs to be improved. More complete physics is certainly the long-term future objective needed for climate predictions. Nevertheless, the simplified physics used here have led to results such as the role of the off-equatorial variability and the connection between the position of the equatorial easterlies and the strength of warm events that apply to other models (Dewitte 2000).

The Tsub.Conv model simulates initial conditions and forecasts during the first 6 months that are much more realistic than the standard CZ model. But it still has a tendency to predict warm conditions and erroneous winds in the eastern equatorial Pacific. The main difference between CZ and Tsub.Conv in terms of forecasting is that the Tsub.Conv initial conditions agree much better with observations, and changing the wind used in the initialization procedure, either by removing some of its energy in the eastern Pacific or by nudging it to the model wind following LDEO2, does not greatly affect individual forecasts.

The Tsub.Astat model is the most reliable prediction system among the ones presented here. Forecasts do not depend as much on friction or coupling coefficients. With a lead time up to 1 yr, the model performs well in predicting the observed anomalies of SST, sea level, and zonal and meridional winds, all over the Pacific between 15°S and 15°N for the period between 1983 and 1993. It predicts the oceanic recharge in the north coupled to the southward migration of the ITCZ that is observed after the 1983 and the 1987 warm events. Note that if the meridional wind stress anomaly is maintained at zero during a time integration of a forecast experiment, the model fails to predict the coming events. So it appears necessary to simulate realistic processes outside the equatorial band to have some skill in forecasting. The success of a model in predicting a particular El Niño event does not guarantee its success for other events. Tsub.Astat fails to predict the warm trend prior to May 1982 or March 1997, whereas it succeeds in predicting the 1986–87 and 1992 warm events, and the 1988 cold events. Sea level data can be used to improve the initial conditions. In particular, TOPEX/Poseidon data are very efficient at improving the 5°–15° bands of

latitude prior to March 1997, but the reader can anticipate that this will not be sufficient to improve the 1997–98 forecasts, since the mechanisms at work for these events do not follow the scenario reproduced by Tsub.Astat (see Perigaud and Cassou 2000). This study emphasizes the uniqueness of each event and illustrates variations taking place at decadal timescales. It is clear that in order to make progress in forecasting statistical analyses of predicted indices are not sufficient, one needs to understand the mechanisms that make a model successful or not for a particular forecast.

*Acknowledgments.* The research described in this paper was carried at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration, including Grants 960504 and 959852, with the University of California, Los Angeles. Authors thank Drs. S. Zebiak and M. Cane from LDEO for providing their code of the coupled ocean–atmosphere model, input files, initialization procedure, as well as their time and helpful discussions.

## APPENDIX A

### Model Initialization with Sea Level Data

Data are used either for validating or for initializing the baroclinic anomalies simulated by the ICMs. As the ocean models assume only one vertical mode, the model thermocline depth anomalies are equal to sea level anomalies divided by the model density ratio (0.0057 in CZ). Four datasets providing estimates of sea level or thermocline depth anomalies are used in this study, one comes from TOPEX/Poseidon (TP) altimetric data since October 1992 and three from XBT temperature profiles since 1980.

For each of these datasets, the monthly climatology is computed over as many complete years as possible (4 for TP and 15 for XBT). For TP, it is necessary to add the reference surface provided by the XBT average during the overlapping period with the satellite mission, so that the TP anomalies can be relative to the same 15 years. Data are initially on a 2° by 1° grid on a monthly basis, except for TP along-track data, which is every 10 days, which provides a resolution of 0.5° with a maximum intertrack distance of 3°. For model validation, the observed anomalies have all been estimated on the atmospheric model grid (2° in latitude and 5.625° in longitude). For model initialization, except in the Kalman filter experiment, the anomalies have been interpolated in space to match the baroclinic model grid (0.5° in latitude and 2° in longitude) and in time every 10 days to match the model time step.

Hydrographic data provide proxies of sea level or thermocline depth. Three different quantities were derived. The first two, described in Perigaud and Dewitte (1996), are the 20°C isotherm depth (D20) and the oce-

anic heat content in the upper 400 m (T400) over 1980–94 provided by Dr. Smith from Bureau of Meteorology Research Centre (BMRC). The third one is the dynamic height at the surface relative to 400 m ( $h_{dyn}$ ) derived from the vertical temperature profiles provided by the BMRC since 1980 with regular updates up to the present time (Smith 1995). D20 can be considered as an estimate of the thermocline, T400, and  $h_{dyn}$  as proxies of sea level variations. T400,  $h_{dyn}$ , and TP take into account variations of density over a thicker layer than the upper layer assumed in the baroclinic model (150 m) while D20 has a position closer to 150 m than 400 m. So D20 is directly used for comparison or insertion in the models, whereas the other three estimates are normalized to match the model physics. The normalization factors for the sea level datasets were obtained by matching their level of variability averaged over the Pacific between 15°S and 15°N to the one for D20.

In the text, the model thermocline is compared with the normalized  $h_{dyn}$  rather than D20 because the latter covers a shorter period, but validation has also been done with D20 over 1980–94 and leads to the same results concerning the deficiency of the CZ model in the western Pacific. On the other hand, initializing the models with one dataset rather than the other matters a lot. The CZ.SL and Tsub.Conv2.SL series presented in this paper are the ones initialized with T400. There are forecasts initialized with D20 that differ from these by more than 2°C in a few months. Overall CZ.SL performs no better no worse with D20 than with T400.

The procedure for initializing the CZ.SL forecasts consists of two steps. First, sea level data are decomposed into Kelvin and Rossby meridional modes up to the fifth one as explained in Perigaud and Dewitte (1996). Then, the CZ model is integrated in time with the FSU winds and the “observed” baroclinic fields inserted at each time step in order to determine the boundary conditions of the baroclinic model in the western Pacific, as well as the Ekman currents, the SST, and the atmospheric fields.

Actually for both steps, alternatives have been tested. For step 1, a second method was developed in order to recover more variability beyond the wave guide. It consists of projecting sea level data on the Kelvin component only and retaining the rest of the sea level as the “non-Kelvin” component, with the currents derived from the non-Kelvin sea level by geostrophy. This second method is the one applied for the series Tsub.Conv2.SL. Because the model domain covers 124°E–100°W and has no landmask, the data used in CZ.SL have been extrapolated from the ocean onto the land encountered over the model domain before applying the first step described above. Then the boundary conditions at the westernmost point (124°E) are obtained by verifying the conservation of the meridionally integrated zonal transport. Because extrapolating data over land is not an accurate solution, other options were tested. One consists in using the observed Kelvin and

Rossby fields east of 150°E only and running the baroclinic model forced by FSU winds to determine the values west of 150°E, but this approach needs a matching zone to avoid inconsistencies at the 150°E transition and is not accurate either. Two alternate options that are more reliable have been considered. One is to add a landmask in the model. The second is to apply an optimal Kalman filter and smoother to assimilate the along-track TP sea level variations as in Fukumori (1995) to initialize the baroclinic model forced by FSU wind anomalies between October 1992 and December 1993.

An alternative for step 2 consists of using the wind simulated by the model where sea level and FSU wind data have been inserted in a first initialization, to reforce the baroclinic model and reinitialize all the oceanic and atmospheric fields again. This second procedure significantly degrades the baroclinic fields compared to the first one in the CZ case. So rather than gaining in model consistency, the initial conditions retained in the text are the closest to the data, meaning that the CZ.SL series and the results presented below correspond to no more than one initialization of the atmosphere, Ekman layer, and mixed layer with the observed baroclinic ocean.

All these experiments support the fact that our failure to improve the CZ forecasts is not primarily due to data errors nor to the inadequacy of the methods applied for initializing the model. At the same time, they illustrate how sensitive the forecasts are to very small changes in the initial conditions and support the need for accurate data and optimal methods of data assimilation. They also suggest that, although the initial shock is certainly a source of concern, initialization in a coupled context as in the LDEO2 procedure should be applied to reduce this shock only if it does not degrade the initial conditions beyond the limits of the estimated a priori data errors.

## APPENDIX B

### Experiments Following the LDEO2 Procedure

Based on the LDEO2 procedure described in D. Chen et al. (1995), we performed several experiments to try to reproduce the published series. Our first experiment consisted of applying the nudging between the model wind and the nondetrended FSU wind starting from the CZ.STD forced conditions of January 1980. Our results were different from theirs and we then thought that it was either because the observed winds must be the detrended version of the FSU dataset, or because the nudging must be started from rest in January 1964. Next, we used only the detrended winds (as for all the series presented in this paper). Then, we found that CZ.CPLIC forecasts are surprisingly sensitive to the starting date of the LDEO2 procedure. Experiments with the procedure started either from rest in January 1964 (CPLIC.64), or from the LDEO1 conditions of January 1970 (CPLIC.70), or January 1974 (CPLIC.74), or Jan-

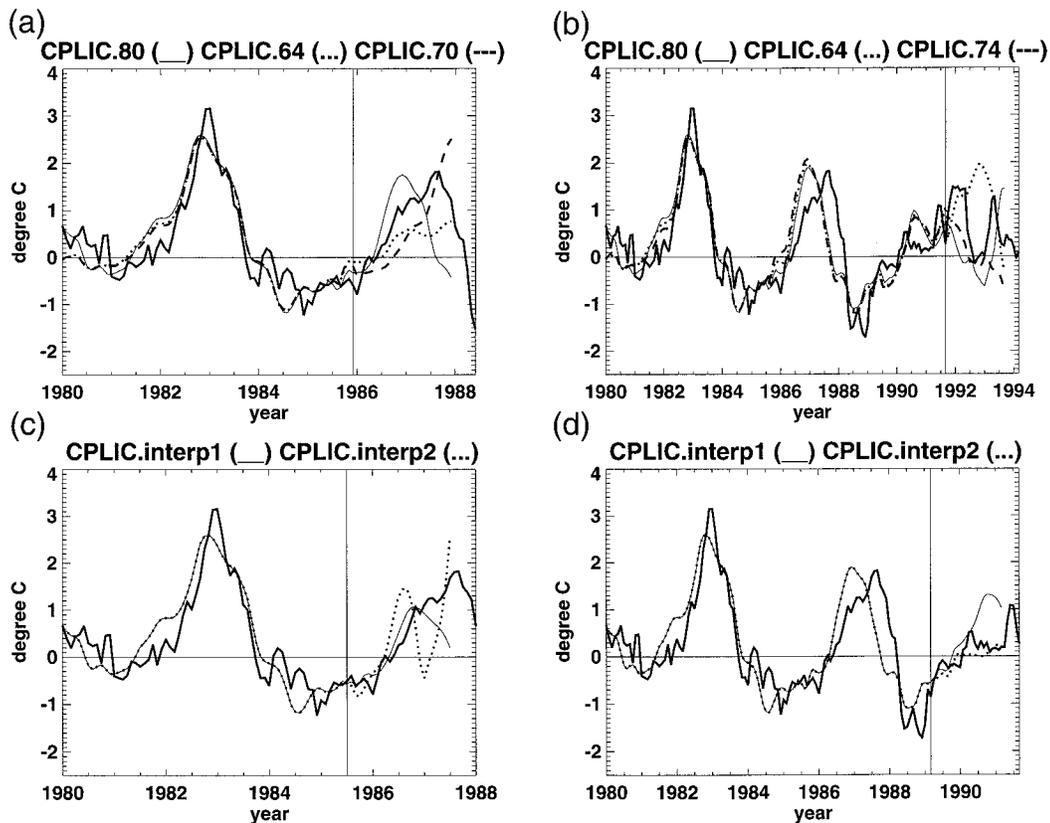


FIG. B1. Niño-3 SST indices as a function of time. In all panels, the observed index is represented by the thick solid line and the vertical bar separates the period of initialization from the period of forecasting. The curves correspond to the series simulated by the CZ model initialized with the LDEO2 procedure for the various tests described in appendix B.

uary 1980 (CPLIC.80) are presented in Figs. B1a and B1b. Results appear similar during the initialization step (before the vertical bar), but forecasts are very different (after the vertical bar). Two predictions arbitrarily chosen many years after the starting date can differ by more than  $2^{\circ}\text{C}$ . CPLIC.80 is the CZ.CPLIC series retained in this paper (it is the one that has the best statistics).

Although all the CZ.CPLIC series look good, none is identical to the series presented in Chen et al. (1995a, 1997). One possible explanation is the uncertainty introduced by the interpolation routines applied to match the various grids involved in nudging FSU wind with the model wind. Because the nudging is applied on the atmospheric grid (S. Zebiak 1995, personal communication), we estimated FSU data on the model grid prior to applying the LDEO2 procedure either by taking the average of the FSU points contained in the atmospheric model grid element weighted by a coefficient function of their distance to the center of the element (interp1), or by adding a smoothing on the 4 nearest adjacent points of the grid in latitude and longitude (interp2). Results of applying the LDEO2 procedure since January 1980 to interp1 or interp2 winds show that the initial

series are very similar, but individual forecasts can be significantly different (Figs. B1c and B1d).

The interpolation that delivers the forecasts retained in the paper is CPLIC.interp1 since 1980. As mentioned above, the CZ.CPLIC series is not identical to the one published by Chen et al. (1995, 1997), but it nevertheless appears to be as good (see Fig. 3). It was also verified that the CZ.CPLIC wind and thermocline anomalies agree well with the results presented in their papers. The CZ.STD forecasts are strictly identical to the ones referred to as LDEO1 in the literature.

Finally, we performed the LDEO2 (interp1 detrended since 1980) experiment where the temporal scheme applied in CZ is replaced by the one applied in Mantua and Battisti (1995). This means that instead of updating the surface current anomalies and upwelling fields after the new SST and the heating fields for the atmosphere have been calculated, the wind stress anomaly is first used to force the dynamic ocean model and then the surface currents prior to updating the SST and the atmosphere. We found that the initial indices are not significantly affected by the time integration procedure, but the forecasts are. In summary, these experiments em-

phasize that even for models as simple as CZ, the coupler that determines the characteristics of the communication between the ocean and the atmosphere in space and time plays a critical role in forecasting.

## REFERENCES

- Bennett, A. F., B. S. Chua, D. E. Harrison, and M. J. McPhaden, 1998: Generalized inversion of Tropical Atmosphere–Ocean (TAO) data and a couple model of the tropical Pacific. *J. Climate*, **11**, 1768–1792.
- Cane, M. A., S. E. Zebiak, and S. C. Dolan, 1986: Experimental forecasts of El Niño. *Nature*, **321**, 827–832.
- Cassou, C., and C. Perigaud, 2000: ENSO simulated by intermediate coupled models and evaluated with observations over 1970–96. Part II: Role of the off-equatorial ocean and meridional winds. *J. Climate*, **13**, 1095–1123.
- Chen, D., S. E. Zebiak, A. J. Busalacchi, and M. A. Cane, 1995: An improved procedure for El Niño forecasting. *Science*, **269**, 1699–1702.
- , —, M. A. Cane, and A. J. Busalacchi, 1997: Initialization and predictability of a coupled ENSO forecast model. *Mon. Wea. Rev.*, **125**, 773–788.
- , M. A. Cane, S. E. Zebiak, and A. Kaplan, 1998: The impact of sea-level data assimilation on the Lamont model prediction of the 1997/98 El Niño. *Geophys. Res. Lett.*, **25**, 2837–2840.
- Chen Y. Q., D. S. Battisti, and E. S. Sarachick, 1995: A new ocean model for studying the tropical oceanic aspects of ENSO. *J. Phys. Oceanogr.*, **25**, 2065–2089.
- Dewitte, B., 2000: Sensitivity of an intermediate ocean–atmosphere coupled model of the tropical Pacific to its oceanic vertical structure. *J. Climate*, **13**, 2363–2388.
- , and C. Perigaud, 1996: El Niño–La Niña events simulated with Cane and Zebiak’s model and observed with satellite and in situ data. Part II: Model forced with observations. *J. Climate*, **9**, 1188–1207.
- Fischer M., M. Latif, M. Flugel, and M. Ji, 1997: The impact of data assimilation on ENSO simulations and predictions. *Mon. Wea. Rev.*, **125**, 819–829.
- Fukumori, I., 1995: Assimilation of TOPEX sea-level measurements with a reduced gravity, shallow-water model of the tropical Pacific Ocean. *J. Geophys. Res.*, **100**, 25 027–25 039.
- Gill, A. E., 1980: Some simple solutions for heat induced tropical circulation. *Quart. J. Roy. Meteor. Soc.*, **106**, 447–462.
- Ji, M., and A. Leetma, 1997: Impact of data assimilation on ocean initialization and El Niño predictions. *Mon. Wea. Rev.*, **125**, 742–753.
- , D. W. Berhinger, and A. Leetmaa, 1998: An improved coupled model for ENSO prediction and implications for ocean initialization. Part II: The coupled model. *Mon. Wea. Rev.*, **126**, 1022–1034.
- Kirtman, B., and S. E. Zebiak, 1997: ENSO simulation and prediction with a hybrid coupled model. *Mon. Wea. Rev.*, **125**, 2620–2640.
- Kleeman, R., 1991: A simple model of the atmospheric response to ENSO SST anomalies. *J. Atmos. Sci.*, **48**, 3–18.
- , A. M. Moore, and N. R. Smith, 1995: Assimilation of subsurface thermal data into an intermediate tropical coupled ocean–atmosphere model. *Mon. Wea. Rev.*, **123**, 3103–3113.
- Latif, M., and Coauthors, 1998: A review of the predictability and prediction of ENSO. *J. Geophys. Res.*, **103**, 14 375–14 393.
- Mantua, N. J., and D. S. Battisti, 1995: Aperiodic variability in the Cane–Zebiak coupled ocean–atmosphere model: Ocean atmosphere interactions in the western equatorial Pacific. *J. Climate*, **8**, 2897–2927.
- Neelin, J. D., and F. F. Jin, 1993: Modes of interannual tropical ocean–atmosphere interaction—A unified view. Part II: Analytical results in the weak-coupling limit. *J. Atmos. Sci.*, **50**, 3504–3522.
- , and N. Zeng, 2000: A quasi-equilibrium tropical circulation model—Formulation. *J. Atmos. Sci.*, **57**, 1741–1766.
- Penland C., M. Flugel, and P. Chang, 2000: Identification of dynamical regimes in an intermediate coupled ocean–atmosphere model. *J. Climate*, **13**, 2105–2115.
- Perigaud, C., and B. Dewitte, 1996: El Niño–La Niña events simulated with Cane and Zebiak’s model and observed with satellite and in situ data. Part I: Model data comparison. *J. Climate*, **9**, 66–84.
- , and C. Cassou, 2000: Importance of oceanic decadal trends and westerly wind bursts for forecasting El Niño. *Geophys. Res. Lett.*, **27**, 389–392.
- , S. Zebiak, F. Melin, J. P. Boulanger, and B. Dewitte, 1997: On the role of meridional wind anomalies in a coupled model of ENSO. *J. Climate*, **10**, 761–773.
- , F. Melin, and C. Cassou, 2000: ENSO simulated by intermediate coupled models and evaluated with observations over 1970–96. Part I: Role of the off-equatorial variability. *J. Climate*, **13**, 1065–1094.
- Rosati, A., K. Miyakoda, and R. Gudgel, 1997: The impact of ocean initial conditions on ENSO forecasting with a coupled model. *Mon. Wea. Rev.*, **125**, 754–772.
- Schneider, E. K., B. Huang, Z. Zhu, D. G. DeWitt, J. L. Kinter III, B. Kirtman, and J. Shukla, 1998: Ocean data assimilation, initialization, and predictions of ENSO with a coupled GCM. *Mon. Wea. Rev.*, **127**, 1187–1207.
- Smith, N. R., 1995: An improved system for tropical ocean subsurface temperature analyses. *J. Atmos. Oceanic Technol.*, **12**, 850–870.
- Syu, H. H., and J. D. Neelin, 2000: ENSO in a hybrid coupled model. Part II: Prediction with piggyback data assimilation. *Climate Dyn.*, **16**, 19–34.
- Wang, B., and T. Li, 1993: A simple tropical atmosphere model of relevance to short-term climate variations. *J. Atmos. Sci.*, **50**, 260–284.
- Xue, Y., M. A. Cane, and S. E. Zebiak, 1997: Predictability of a coupled model of ENSO using singular vector analysis. Part II: Optimal growth and forecast skill. *Mon. Wea. Rev.*, **125**, 2074–2093.
- Zebiak, S. E., 1986: Atmospheric convergence feedback in a simple model for El Niño. *Mon. Wea. Rev.*, **114**, 1263–1271.
- , and M. A. Cane, 1987: A model El Niño–Southern Oscillation. *Mon. Wea. Rev.*, **115**, 2262–2278.