Evaluating Observing Requirements for ENSO Prediction: Experiments with an Intermediate Coupled Model

Rebecca E. Morss

National Center for Atmospheric Research,* Boulder, Colorado

David S. Battisti

Joint Institute for the Study of the Atmosphere and Ocean, University of Washington, Seattle, Washington

(Manuscript received 16 June 2003, in final form 10 February 2004)

ABSTRACT

The Tropical Atmosphere Ocean (TAO) array of moored buoys in the tropical Pacific Ocean is a major source of data for understanding and predicting the El Niño–Southern Oscillation (ENSO). Despite the importance of the TAO array, limited work has been performed to date on the number and locations of observations required to predict ENSO effectively. To address this issue, this study performs a series of observing system simulation experiments (OSSEs) with a linearized intermediate coupled ENSO model, stochastically forced. ENSO forecasts are simulated for a number of observing network configurations, and forecast skill averaged over 1000 years of simulated ENSO events is compared.

The experiments demonstrate that an OSSE framework can be used with a linear, stochastically forced ENSO model to provide useful information about requirements for ENSO prediction. To the extent that the simplified model dynamics represent ENSO dynamics accurately, the experiments also suggest which types of observations in which regions are most important for ENSO prediction. The results indicate that, using this model and experimental setup, subsurface ocean observations are relatively unimportant for ENSO prediction when good information about sea surface temperature (SST) is available; adding subsurface observations primarily improves forecasts initialized in late summer. For short lead-time (1–2 month) forecasts, observations within approximately 3° of the equator are most important for skillful forecasts, while for longer lead-time forecasts, forecast skill is increased by including information at higher latitudes. For forecasts longer than a few months, the most important region for observations is the eastern equatorial Pacific, south of the equator; a secondary region of importance is the western equatorial Pacific. These regions correspond to those where the leading singular vector for the ENSO model has a large amplitude. In a continuation of this study, these results will be used to develop efficient observing networks for forecasting ENSO in this system.

1. Introduction

The Tropical Atmosphere Ocean (TAO) array of moored buoys was deployed during the Tropical Ocean Global Atmosphere (TOGA) program to provide a consistent source of real-time in situ observations across the tropical Pacific Ocean. Since the early 1990s, the TAO array has regularly provided surface atmospheric wind, sea surface temperature (SST), and subsurface ocean data at approximately 70 locations between 8°N and 8°S and 95°W and 137°E. Along with satellites, ships, and other in situ observing platforms in the tropical Pacific, the TAO array is a major source of observations for understanding and predicting El Niño, La Niña, and the Southern Oscillation (ENSO).

When the TAO array was designed, ENSO variability and dynamics were much less well understood than they are today, and operational ENSO prediction was in its early stages. Given the recent advances in ENSO understanding and prediction—in large part due to the TAO array and the TOGA program—it seems appropriate to revisit the design of the ENSO observing network. This includes assessing how varying the number and locations of different types of observations affects the skill of ENSO forecasts.

A number of previous studies have addressed the importance of different observations for ENSO prediction. Such studies have employed a variety of ENSO prediction methods, including coupled ocean–atmosphere general circulation models (Ji and Leetmaa 1997; Rosati et al. 1997; Ji et al. 1998; Segschneider et al. 2001; Wang et al. 2002), hybrid models (dynamical ocean models coupled to statistical atmospheric models; Fi-

---

Corresponding author address: Dr. Rebecca E. Morss, National Center for Atmospheric Research, P.O. Box 3000, Boulder, CO 80307. E-mail: morss@ucar.edu

scher et al. 1997; Syu and Neelin 2000), intermediate coupled models (Kleeman et al. 1995; Chen et al. 1998; Ballabrera-Poy et al. 2001), and statistical models (Latif and Graham 1992; Smith et al. 1995; Johnson et al. 2000; Xue et al. 2000). They have also tested a variety of data sources and a variety of data insertion or assimilation methods. Most of these studies, however, focus primarily on one aspect of ENSO observing networks: the importance of ocean heat content information (including subsurface temperature, thermocline depth, and sea level) relative to that of ocean surface information (including SST and surface atmospheric winds). Thus, despite the TAO array's importance, current understanding of the type, number, and locations of observations required for skillful ENSO prediction is limited.

In addition, most previous studies have evaluated the importance of different observations for hindcasting ENSO, that is, for predicting past ENSO events. Hindcast studies are restricted by the length of the data record suitable for such experiments, generally shorter than 50 years for SST and wind data, and shorter than 30 years for ocean heat content data. Consequently, most studies evaluate the importance of different observations for fewer than 30 years of ENSO forecasts (e.g., Chen et al. 1998; Johnson et al. 2000; Xue et al. 2000), and many of the frequently cited studies base conclusions on fewer than 15 years of ENSO forecasts (e.g., Kleeman et al. 1995; Fischer et al. 1997; Ji and Leetmaa 1997; Rosati et al. 1997). Given that the characteristics of the ENSO cycle vary on decadal and longer time scales, it is unclear how well results from studies with short data records will apply to predictions of future ENSO events.

To allow an evaluation over a longer time period than is feasible for hindcasts, this study evaluates observing networks for ENSO prediction using observing system simulation experiments (OSSEs). Rather than use real observations, OSSEs simulate observations by periodically sampling from a numerical model run that simulates the evolution of the physical system being studied. Because the experiments use synthetic data, results can be compared for many more configurations of observations and simulated ENSO forecasts than is possible using real data.

This study simulates the real ENSO system and ENSO forecasts using an intermediate coupled model that is a linearized variant of the Zebiak–Cane (1987) ENSO model, with stochastic forcing representing atmospheric noise. Because the experiments use the same dynamical model to evolve the simulated real and forecast states, they are similar to the standard "identical twin" experiments that have been used to study other aspects of ENSO, such as predictability and data assimilation (e.g., Miller 1990; Goswami and Shukla 1991; Hao and Ghil 1994; Blanchet et al. 1997; Alves et al. 2001; Sun et al. 2002; and references therein). The experiments performed here are not truly identical twins, however, because different realizations of stochastic forcing are used to evolve the real and forecast states. Nevertheless, the experiments retain the standard limitation of OSSEs: the applicability of the results to the real world depends on how well the simulation represents the true system.

Section 2 describes the experimental design, including the ENSO prediction model, the simulated observations, the data assimilation system, and the setup of the OSSEs. Next, to demonstrate the general behavior of the OSSE system, section 3 presents results from several reference experiments, including forecasts created using perfect initial conditions, many SST and/or thermocline depth observations, no observations, and persistence. The next two sections evaluate how the average skill of ENSO forecasts is affected by varying the number and locations of platforms that observe both SST and thermocline depth. Section 4 tests modifying the meridional extent, meridional spacing, and longitudinal spacing of observations, and section 5 tests removing observations in different regions. Section 6 summarizes and compares the results with those from other studies.

A continuation of this study, Morss and Battisti (2004, hereafter MB2), builds upon the results presented here by developing observing networks that are efficient for ENSO prediction in this system, where efficient is defined as providing reasonably skillful forecasts for relatively few observations. Since a major difference between this study and many other studies of ENSO prediction is the number of ENSO events examined, MB2 also explores the dependence of the OSSE results on the duration of the simulated data record.

2. Experimental design

a. Numerical prediction model

The forecast model used is an intermediate coupled ENSO anomaly model, the T80 model described and studied in Thompson and Battisti (2000, 2001; hereafter TB2000, TB2001). Its characteristics are only summarized here; detailed discussion of the model physics and behavior can be found in Thompson (1998a), TB2000, and TB2001.

The T80 model is a linearized, moderately damped variant of the Zebiak–Cane (1987) ENSO model, stochastically forced. The T80 model was constructed by 1) linearizing the Battisti (1988) version of the Zebiak–Cane model (Thompson 1998a); 2) reducing the atmosphere–ocean coupling until the system was stable (TB2000); 3) modifying two parameters, the western boundary reflection efficiency and the ocean mechanical damping, to more realistic values [based on results obtained since Zebiak and Cane (1987)] and examining the effects on transient and modal growth (TB2000); and 4) using different values of these parameters, within a realistic range, to construct four candidate models, including the T80 model (TB2000). Each of the four

candidate models was then run with stochastic forcing; after comparing times series, power spectra, and other characteristics of the simulations with Comprehensive Ocean–Atmosphere Data Set (COADS) SST data, the T80 model was selected as the most realistic (TB2000, TB2001). The T80 model's leading mode is similar to the observed ENSO mode and peaks in boreal winter, with a period of 3.85 yr and an annual growth rate of 0.80 (TB2000).

The prognostic variables in the model are ocean dynamics, representing thermocline depth and ocean surface currents, and SST. The atmospheric surface winds are obtained diagnostically from the SST. Because the system is moderately damped, all disturbances are the result of transient growth initiated by the stochastic forcing, which represents the effects of atmospheric noise. The stochastic forcing is applied to the model prognostic variables; it is white in time, is uncorrelated in space, and has an amplitude that produces the observed level of variance in the standard indices of ENSO [e.g., Niño-3 (5°S–5°N, 90°–150°W) SST index, Southern Oscillation index, and Cold Tongue index]. Other forms of stochastic forcing, including structured SST noise (using higher-order EOFs derived from COADS data) and seasonal Kelvin wave perturbations in the western Pacific (simulating the effects of wind bursts related to the Madden–Julian oscillation), were tested in Thompson (1998b) and TB2001 and did not substantially change the model behavior.

In longitude, the model domain extends from 123°E to 81°W and is discretized at 15 equally spaced grid points. The zonal gridpoint spacing is approximately 11° (most longitudes discussed in the paper are rounded to the nearest whole degree). In latitude, the model is discretized in terms of Hermite polynomials for SST, and in terms of Kelvin and Rossby waves for the ocean dynamics. Eight Hermite polynomials are retained for SST, and the Kelvin wave and eight Rossby waves are retained for the ocean dynamics. To obtain meridional grid points, the Hermite polynomials and Kelvin/Rossby waves are projected onto a grid with approximately 0.77° latitudinal spacing (most latitudes discussed in the paper are rounded to the nearest tenth of a degree). Increasing the zonal resolution of the model or retaining additional meridional modes produces no substantial changes other than a poleward extension of the southeastern SST maximum in the leading singular vector (Thompson 1998a).

Because the OSSE methodology (described in section 2d) uses the numerical model to generate synthetic ENSO events, the applicability of the study's results to the real world depends on how well the model simulates the real ENSO system. Although the model is an imperfect representation of the full system, as noted above, its leading mode is similar to the observed ENSO mode (see Thompson 1998a and TB2000 for details). The model is also closely related to the Zebiak–Cane (1987) model, which is well established, has been used in numerous studies of ENSO, and has predictive skill similar to that of other ENSO prediction schemes (Latif et al. 1998). One major difference between this model and the Zebiak–Cane model is the linearization; however, as noted by Xue et al. (1994) and Chen et al. (1997), both the Zebiak–Cane and Battisti versions of the model are essentially linear to small perturbations out to approximately 9 months. For additional discussion of the relationship of this model and its variants to the Zebiak–Cane model and the real ENSO system, see Chen et al. (1997), Thompson (1998a), TB2000, and references therein.

The T80 version of the model has not been tested extensively in ENSO prediction. However, Thompson (1998b) used the T80 model to generate analog hindcasts of ENSO, using COADS SST data from 1950–90. At a lead time of 6 months, the analog hindcasts had a Nino-3 SST anomaly correlation of 0.5 and an rms SST error of 0.8 (see Thompson 1998b for details); this skill is not substantially different from other ENSO prediction schemes initialized without ocean subsurface information (Latif et al. 1998), particularly given that the forecast methodology was not tuned on a subset of the data and that the model output was not postprocessed. Overall, therefore, the model used here is sufficiently representative of ENSO dynamics to be useful for an OSSE study. Nevertheless, only a limited set of dynamics and types of ENSO events are represented in the model, and the results should be interpreted with this in mind.

b. Simulated observations

The simulated observations in this study are of SST and thermocline depth, taken once per month. Assimilating SST observations is, using this numerical model, similar to assimilating surface wind information; assimilating thermocline depth is similar to assimilating upper-ocean heat content or sea level data. All observing locations remain fixed throughout each experiment, simulating immobile observing platforms such as TAO moorings, island stations, or geostationary satellites. However, given the idealizations in the experimental setup, the results are not intended to identify locations for specific observing platforms; instead, they indicate general features of observing networks that are important for skillful ENSO prediction in this system, assuming that observing locations cannot be purposefully moved between observing times.

All observations in all experiments are assumed to be independent, in other words, to have uncorrelated errors. The observation error covariance matrix, used to both assimilate data (section 2c) and simulate observation errors (section 2d), is therefore diagonal. The standard deviations of the observation errors are set to 0.2°C for all SST observations and 2 m for all thermocline depth observations (errors represent an average over a 30-day period). Although more complex obser-

vation error statistics could easily be simulated, we chose not to include features such as spatial correlations in observation errors or spatial variation in observation representativeness errors in these experiments to facilitate interpreting the results.

c. Data assimilation system

The simulated observations are incorporated into the model simulation using a standard three-dimensional variational data assimilation system (3DVAR), which is similar to an optimal interpolation (OI) scheme. The 3DVAR implemented here is based on the data assimilation system described in Derber and Rosati (1989; hereafter DR89) and currently used operationally at the National Centers for Environmental Prediction (Ji et al. 1995), with one major difference: because the observations in this study are idealized, they do not require temporal smoothing and thus are not inserted continuously.

The data assimilation system generates an "analysis" state, used as the initial conditions for forecasts, by combining the observations with a "background" field, the 1-month forecast valid at the observation time. To determine how to weight the observations and the background field, 3DVAR uses statistics given to it in observation and background error covariance matrices. The observation error covariance matrix is described in section 2b. To develop the background error covariance matrix, we assumed that errors in the SST background field are uncorrelated with errors in the ocean dynamics background field, and that background errors in different projections in latitude are uncorrelated.[1] The only remaining correlations are longitudinal; these are defined using the parameterization given in DR89, with latitudinal dependence removed:

$$ae^{-r^2/b^2}, \tag{1}$$

where r = east–west distance between grid points, b = longitudinal correlation length scale, and a = background error variance.

The parameter b determines how far (east and west) each observation's influence extends. In DR89, b is set to 570 km, with the east–west distance used in the calculation of r decreased by a factor of 2.28 near the equator (in their nearly global data assimilation system). The effective b near the equator in DR89 is therefore 2.28×570 km ≈ 1300 km. For these experiments, we set b to 1300 km for both the SST and ocean dynamics fields in the background error covariance matrix at all longitudes and for all projections in latitude.

The parameter a determines how the data assimilation system weights each observation relative to the back-

ground field. Since DR89 (and similar real-world implementations of 3DVAR for ENSO prediction) insert data continuously over a 30-day window, their choice of a cannot be extended to this study. Instead, we have chosen a so that for a single observation near the equator, the observation and the background field are weighted approximately equally. Thus, a is set to 1.7×10^{-4} for SST and 5.4×10^{-4} for ocean dynamics, at all longitudes and for all projections in latitude. Although this choice of a is somewhat arbitrary, many of the experiments have been performed using different values of a, producing results similar to those shown.

Using this parameterization, the background error covariance matrix is calculated once and remains the same for all the experiments shown. Because the ENSO model has only 260 degrees of freedom, the background error covariance matrix and the matrix that must be inverted in 3DVAR to solve for the analysis increments are only 260×260. The full matrices can therefore be stored and inverted.

Because the matrices in the data assimilation system are relatively small, one could easily develop more sophisticated background error statistics for this 3DVAR. For example, the parameters a and b could be varied with longitude, latitude, variable, and observing network, or the full time-averaged background error covariances could be estimated for different observing networks using the OSSEs. One could also easily implement a more sophisticated data assimilation scheme, such as a four-dimensional variational algorithm (4DVAR) or a variant of a Kalman filter. For these experiments, however, we chose to use 3DVAR and simple error statistics both to keep the data assimilation system similar to those currently operational and to simplify interpreting the effects of changing the observing network. The relatively simple data assimilation scheme is also reasonable given the simplicity of the model and the experimental setup; a more sophisticated scheme would likely produce unrealistically optimistic results with few observations.

Figure 1 depicts the analysis increments produced by the data assimilation system, using the aforementioned background error covariance matrix, for a single SST observation and a single thermocline depth observation. Due to the assumption that SST and ocean dynamics background errors are uncorelated, the SST observation does not affect the ocean dynamics field, and the thermocline depth observation does not affect the SST field. Note that a thermocline depth observation near the equator influences the analysis on a broader meridional scale than an SST observation near the equator. This occurs because the data assimilation system determines how to spread observational information meridionally using the meridional projections in the ENSO model, and the meridional scale of these projections is larger for ocean dynamics than for SST.

Because the data assimilation system interpolates information between observations, some of the differences among results from different observing networks

---

[1] Although studies of the error covariances in this system indicate that these background errors are correlated, these assumptions were made for several reasons related to the goals of this study; see following discussion.
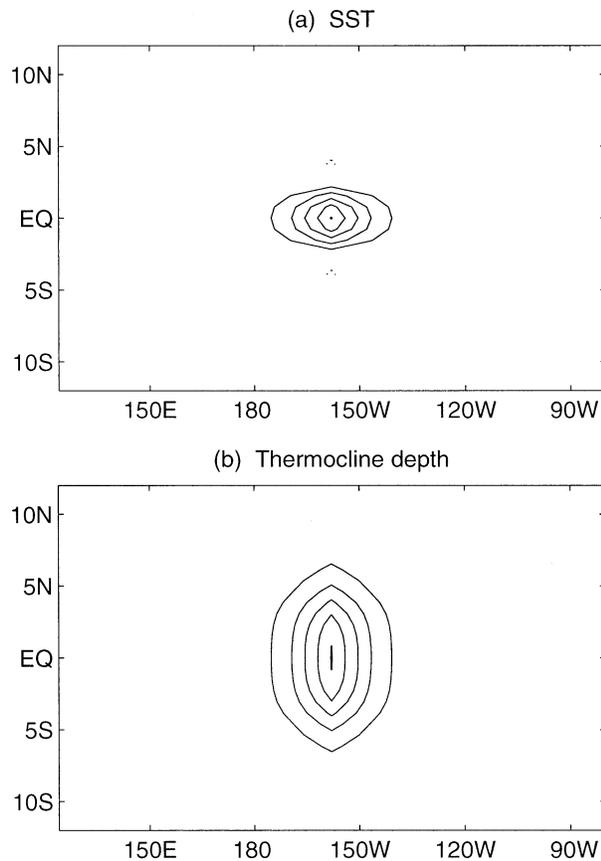
(a) SST

(b) Thermocline depth

FIG. 1. (a) SST analysis increment produced by the data assimilation system for one SST observation of 1°C at 0°, 159°W. (b) Thermocline depth analysis increment for one thermocline depth observation of 10 m at 0°, 159°W. The contour interval is 0.1°C in (a) and 1 m in (b); positive contours are solid and negative contours are dotted. The maximum in both (a) and (b) is approximately one-half the value of the observation, as designed by the choice of magnitude for the background error covariance matrix.

could be related to how well the data assimilation system fills in different missing information. With this relatively simple data assimilation system, however, removing observations of a variable in a region clearly increases the average analysis error in that variable and region—to the extent that one can generally identify which observations were taken where in an experiment simply by looking at the spatial distribution of time-averaged analysis errors. In addition, the spatial distributions of time-averaged analysis and forecast errors generally depict how the increased analysis errors in a region propagate into short lead-time forecast errors. Results from different observing networks can therefore be interpreted in terms of how important good information about initial conditions in different variables and regions is for forecasting ENSO.

d. Setup of OSSEs

To run the OSSEs, first the ENSO model is run from an arbitrary initial state to generate a ''true'' (''nature'')

run, designed to simulate the evolution of the real-world ENSO system. To simulate an ''analysis/forecast'' (''model'') run, the ENSO model is run from another arbitrary initial state, using the same dynamical equations but a different sequence of random numbers in the stochastic forcing. Once per month, a set of ''observations'' is created by sampling SST and/or thermocline depth from the true state at the appropriate locations and then adding random errors consistent with the statistics in the observation error covariance matrix. These simulated observations are assimilated into the analysis/forecast run using 3DVAR, which combines them with a 1-month ''forecast'' valid at the observation time. The resulting analysis becomes the new model state, used as the initial conditions for forecasting to the next observation time and out to 24 months. Each set of forecasts uses a different realization of the stochastic forcing.

The experiments are thus run as a continual cycle of analyses and forecasts, with a 1-month data assimilation interval. Each experiment starts with the same true state and uses the same sequence of stochastic forcing in its true run, so that each experiment is trying to forecast the same sequence of ENSO states. Prior to gathering skill statistics, each experiment is run for a 20-yr spinup period to equilibrate the analysis/forecast run to the observing network being tested. In subsequent sections, the forecast lead time is defined as the time lag between the initial and end time of a forecast.

The OSSEs run here are similar to standard identical twin experiments in the sense that the model used to generate the analysis/forecast run includes the same dynamics as the model used to generate the true run. However, the forecast and true models use different realizations of the stochastic forcing, and thus the experiments are not truly identical twins. Because this experimental setup simulates a nearly perfect forecast model, the techniques that are often employed in ENSO prediction experiments to compensate for model error, such as the 3-month averaging of forecasts, bias correction, or restarting the analysis/forecast run, are not required.

A major advantage of OSSEs compared to studies with real-world data is that the experiments one can perform are not limited by either the events experienced in the real world or the data taken as the events occurred. To capitalize on these advantages, this study tests idealized observing networks and, for all of the results shown in this paper, accumulates skill statistics over 1000 years of events. One thousand years is long enough to provide reasonably stable statistics for comparing observing networks; results for shorter periods of time are discussed in MB2.

Another major advantage of OSSEs is that perfect information is always available about the true state of the system being predicted. All of the results shown, therefore, evaluate analyses and forecasts with respect to the corresponding true state. Four error norms were used: anomaly correlation (AC) of the Niño-3 SST in-

TABLE 1. Description of the reference experiments discussed in section 3.

Experiment	Description
Perfect IC	Standard OSSE, except that at each assimilation time, the initial model state is set to the true state (i.e., OSSE with perfect initial conditions).
FullTh	Standard OSSE, with 375 simulated SST and thermocline depth observations at the locations shown in Fig. 2.
FullT	Standard OSSE, with 375 SST observations at the locations shown in Fig. 2 (no thermocline depth observations).
Fullh	Standard OSSE, with 375 thermocline depth observations at the locations shown in Fig. 2 (no SST observations).
NoObs	Standard OSSE, but with no observations.
Persist	Initial conditions generated as in Perfect IC experiment, with forecasts generated by persisting the initial anomaly (i.e., persistence forecasts).

dex, and root-mean-square error (rmse) in SST in the Niño-3 region, in the Niño-3.4 region ($5°N$–$5°S$, $170°$–$120°W$), and in equatorial SST ($10°S$–$10°N$). Unless otherwise mentioned, results for the four norms are similar; most results are shown for only the Niño-3 AC norm.

The major disadvantage of OSSEs is that the results are only applicable to the real world to the extent that the simulated system represents the real system. The observations simulated, for example, are idealized; real-world observations often do not directly measure model variables (e.g., thermocline depth), and they can have errors not represented in our simulation. More importantly, the model used to generate the true run is an imperfect representation of the true coupled atmosphere–ocean system. A major limitation of the results presented, therefore, is the simplified dynamics included in the T80 ENSO model, and the results should be interpreted with this in mind.

3. Reference experiments

a. Annually averaged forecast skill

To demonstrate the basic behavior of the experimental system and provide a framework for interpreting subsequent results, first we discuss results from the six experiments described in Table 1. Figure 3 shows errors
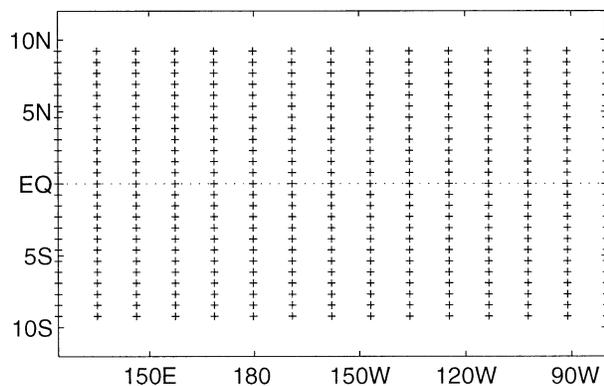

FIG. 2. Locations of the 375 observations in the FullTh, FullT, and Fullh networks: at 15 longitudes, spaced by $11°$, and at 25 latitudes, spaced by $0.77°$ from $9.2°S$–$9.2°N$.

in initial conditions and 1–24-month forecasts, averaged over a 1000-yr OSSE. Results are depicted for two error norms: AC of the Niño-3 SST index and rmse in Niño-3.4 SST. Results for rmse in equatorial Pacific SST and in Niño-3 SST (not shown) are similar to those depicted for Niño-3.4 SST.

As expected given the stochastic forcing in the ENSO model, even forecasts beginning with perfect initial conditions (PerfectIC experiment) have errors. For each forecast lead time, the PerfectIC results represent the maximum forecast skill attainable by improving the initial conditions, averaged over this 1000-yr period. On average, 15-month forecasts from perfect initial conditions have Niño-3 ACs less than 0.6, and 12-month PerfectIC forecasts have ACs less than 0.7.[2] Most results shown hereafter will be limited to forecast lead times of 12 months.

In the FullTh experiment, perfect information about the initial conditions is replaced by a large number of observations (at the locations depicted in Fig. 2) assimilated into the model state using 3DVAR. FullTh forecasts are nearly as skillful as PerfectIC forecasts, indicating that the data assimilation system performs well with many observations; it generates initial conditions that are not only accurate, but also accurate in the sense that matters for forecasts. Compared to the PerfectIC results, the FullTh results are a more realistic measure of maximum attainable forecast skill, that is, the skill when a large amount of imperfect information about the initial conditions is available. The remaining results are interpreted with respect to these two estimates of maximum attainable forecast skill in this simulated system; all plots with the same form as Fig. 3 throughout the paper show the same PerfectIC and FullTh results using the same line styles.

When many SST but no thermocline depth observations are available (FullT experiment), Fig. 3 shows that forecasts are on average somewhat less skillful than those in the PerfectIC and FullTh experiments, but not dramatically so. Differences in forecast skill can be con-

[2] Note that this is only a long-term average; in some subsets of this 1000-yr period, the average skill of PerfectIC forecasts is significantly higher or lower. This variability is discussed in MB2.
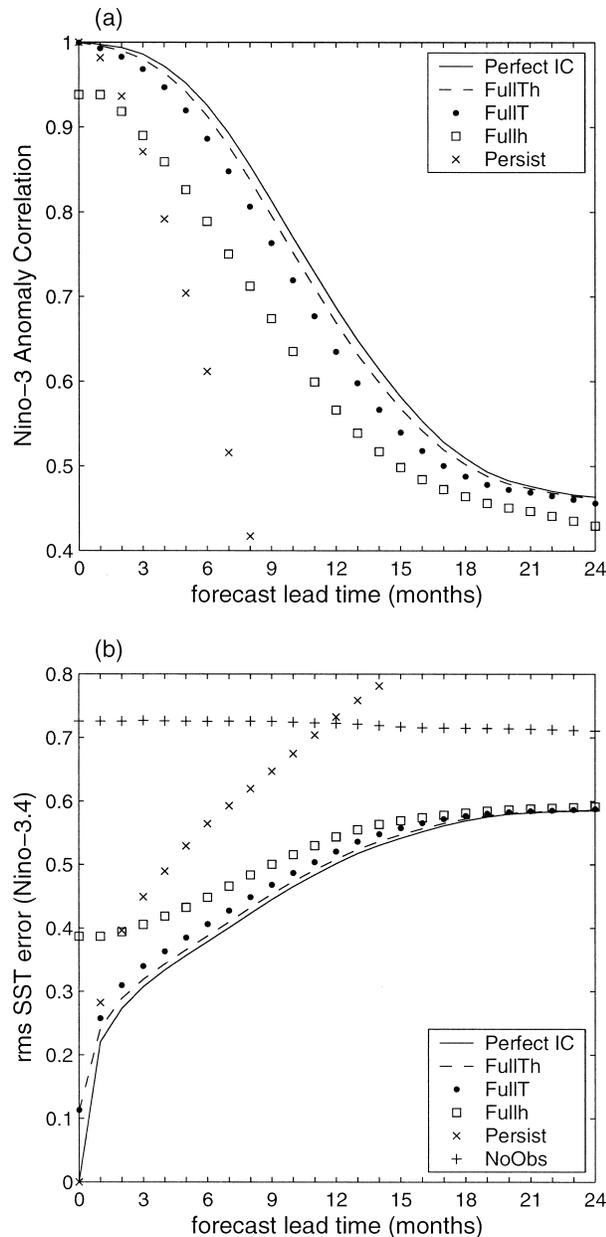
FIG. 3. (a) Anomaly correlation of the Niño-3 SST index for initial conditions (forecast lead time = 0 months) and forecasts with lead times ranging from 1 to 24 months for the six experiments described in Table 1. Results are averaged over a 1000-yr OSSE. The anomaly correlations for the NoObs experiment are all close to zero and thus do not appear. (b) As in (a), but for rms SST error in the Niño-3.4 region. The average skill of persistence forecasts reaches a Niño-3 AC minimum of −0.67 and an rms Niño-3.4 SST error maximum of 0.90 at about 23 months of lead time, thereafter asymptoting to an AC of zero and a saturated rms error.

sidered in two ways: 1) as differences in skill for forecasts at the same lead time or 2) as differences in the forecast lead time at which average skill drops below a certain threshold, that is, as a loss in lead time. In other words, one can say that FullT forecasts are slightly less

skillful than FullTh forecasts because 1) FullT Niño-3 ACs are approximately 1%–5% smaller than FullTh Niño-3 ACs or 2) 8-month FullT forecasts are approximately as accurate as 9-month FullTh forecasts (a 1-month loss in lead time). This indicates that good SST information is sufficient to produce ENSO forecasts nearly as skillful as possible, while adding subsurface ocean information on average benefits forecasts only slightly (using this ENSO model and experimental setup).

When the many SST observations in the FullT experiment are replaced with many thermocline depth observations (Fullh experiment), forecast skill decreases noticeably (a degradation of 5%–12% in Niño-3 AC, or a 2–3-month loss in lead time). The magnitude of the difference between the Fullh and FullT results is partly related to the magnitudes of the observation errors. However, decreasing the magnitude of the thermocline depth observation errors does not improve Fullh forecasts to the level of FullT. This suggests that information about the SST initial conditions is more important for forecasting ENSO skillfully in this system than information about the initial ocean heat content. This is corroborated by results for an upper-ocean equatorial heat content norm (not shown); at lead times greater than 2 months, the FullT array forecasts heat content more skillfully than the Fullh array. Note that this occurs despite the key role of ocean memory in this system and is likely related to the damped ENSO mode and stochastic forcing.

Nevertheless, Fig. 3 shows that the Fullh analyses and forecasts retain significant skill, and that they are much more skillful than when no observations are taken (NoObs experiment). Thus, even with no direct information about SST, the 3DVAR data assimilation system and ENSO model together are able to use good information about ocean heat content, inserted every month, to generate fairly accurate SST analyses and forecasts—at least in this experimental setup, in which the dynamics of the system being predicted are well represented by the forecast model.

The skill of persistence forecasts (Persist experiment), while initially high, drops rapidly as the forecast lead time increases. At lead times greater than 2 months, all of the experiments shown except NoObs produce forecasts that are on average more skillful than persistence forecasts. At lead times greater than 6 months, the average Niño-3 AC for persistence forecasts is less than 0.6; at lead times greater than 12 months, the average Niño-3 AC for persistence forecasts drops below 0. The persistence skill of upper-ocean equatorial heat content (not shown) is nearly identical to the persistence skill of Niño-3 SST.

Because the results for different observing networks are generally easier to differentiate using the Niño-3 AC norm than the rms SST norms, most of the remaining results are presented for only the Niño-3 AC norm. Un-
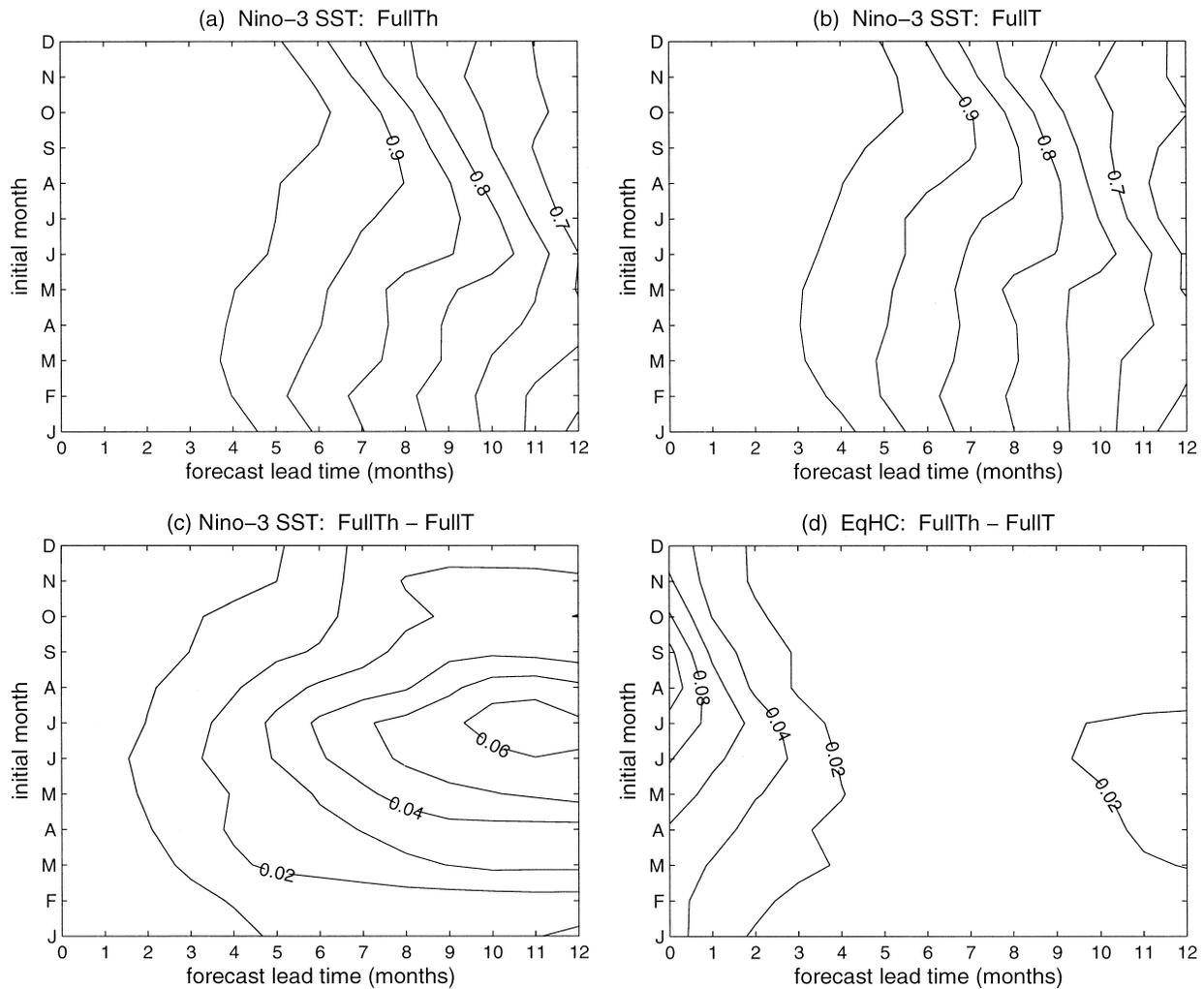
FIG. 4. (a) Anomaly correlation of Niño-3 SST forecasts as a function of start month (*y* axis) and forecast lead time (*x* axis) for the FullTh experiment. (b) As in (a) for the FullT experiment. (c) Difference between (a) and (b). (d) As in (c), but for anomaly correlation of upper-ocean equatorial heat content. Contour intervals are 0.05 in (a) and (b), 0.01 in (c), and 0.02 in (d).

less otherwise mentioned, all results discussed are similar for the other error norms tested.

b. Seasonal variability in forecast skill

Many ENSO prediction schemes exhibit seasonal variability in forecast skill, with SST forecast skill dropping off significantly as forecasts pass through boreal spring (e.g., Latif et al. 1998). Although the causes of this "spring prediction barrier" are not completely understood, it coincides with the season when tropical Pacific SST anomalies are least persistent, as indicated by a spring minimum in lag autocorrelations in SST observations (e.g., Wright 1985). As noted in TB2001, the T80 model used here also exhibits a seasonal SST persistence barrier, although it occurs 1–2 months later and is less pronounced than in observations. A similar late spring barrier occurs in numerical forecasts in this system, as depicted in Fig. 4a (for the FullTh experiment).

Autocorrelations of upper-ocean equatorial heat content anomalies exhibit a minimum in boreal winter, with heat content anomalies least persistent over the late fall and early winter and most persistent over the spring (e.g., Balmaseda et al. 1995; McPhaden 2003). This offset in the phasing of SST and heat content anomaly persistence barriers has led to suggestions that initializing ENSO forecasts with upper-ocean heat content information may improve predictions over the spring SST barrier (e.g., Balmaseda et al. 1995; McPhaden 2003). Supporting this suggestion, several studies with real data have found that improving the initialization of upper-ocean heat content reduces the prominence of the spring SST prediction barrier (e.g., Smith et al. 1995; Xue et al. 2000).

Like observations, this model exhibits an upper-ocean equatorial heat content persistence barrier out of phase with the SST persistence barrier, with heat content anomalies least persistent beginning in August (several

months earlier than observations) and most persistent during the spring. Does this offset mean that good information about upper-ocean heat content can reduce the prominence of the SST prediction barrier, carrying the forecasts through the spring period when SST anomalies are least persistent? Figure 4a depicts the seasonal variability in forecast skill for the FullTh experiment, when many SST and thermocline depth observations are available; Fig. 4b depicts the same results, but without thermocline depth observations (the FullT experiment). In comparing Figs. 4a and 4b, removing heat content observations does not appear to significantly modify the late spring barrier. Rather, as indicated by the difference between Figs. 4a and 4b depicted in Fig. 4c, adding heat content observations most improves SST forecasts that *begin* in boreal summer.

Figure 4d depicts the same results as Fig. 4c (FullTh − FullT results), but for anomaly correlation in upper-ocean equatorial heat content (defined as thermocline depth integrated across the Pacific from 5°S–5°N). In this system, heat content observations most improve heat content initial conditions in late summer; 3–12 months later, these improved initial conditions lead to improved SST forecasts (Fig. 4c). Note that, in this system 1) perturbations initialized at any month reach maximum growth when allowed to evolve over boreal summer and they peak at the end of the year (TB2000), and 2) the magnitude of heat content anomalies tends to be smallest in late summer (not shown). The first factor suggests that inserting heat content information may be most important in late summer because doing so prevents heat content errors from growing unrestrained over this period of rapid growth, leading up to the winter peak in ENSO events. Because the magnitude of 1-month forecast errors and analysis increments in heat content is similar year-round, the second factor suggests that inserting heat content information may be most important in late summer because at this time of year the corrections to heat content are a larger percentage of the anomalies. We hypothesize that some combination of these factors, both related to the annual cycle in the model, is responsible for the results in Fig. 4. Further study is required, however, to explore the reasons for these results and whether they extend to other ENSO prediction systems. Subsequent results are shown only for annual averages.

4. Effects of varying extent and spacing of observations

This section explores the effects, averaged over 1000 years of simulated ENSO events, of modifying the number and placement of SST and thermocline depth observations, assuming that both variables are sampled at the same locations. Results from the different observing networks are compared as in section 3, in terms of the average skill of analyses and of forecasts generated at different lead times. All of the observing networks test-
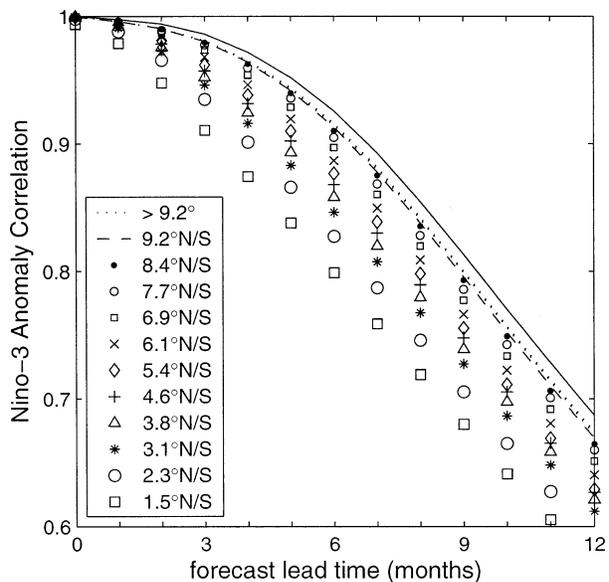

FIG. 5. Anomaly correlation of the Niño-3 SST index for initial conditions and 1–12-month forecasts, averaged over a 1000-yr OSSE, for observing networks with different meridional extents. All of the networks are evenly spaced grids that are symmetric about the equator, with 15 north–south transects spaced by 11° longitude, and observations within the transects spaced by 0.77° latitude. The networks' meridional extents vary from 1.5°N/S to >9.2°N/S and are indicated in the figure legend. The 1.5°N/S network, for example, has observations at 1.5°S, 0.77°S, 0°, 0.77°N, and 1.5°N, for a total of 15 × 5 = 75 observing locations; each time the network is extended north and south by 0.77°, 30 observing locations are added. The 9.2°N/S network is the same as the FullTh network shown in Fig. 2. Results for all networks with meridional extents greater than 9.2°N/S are similar and are depicted with dotted lines. As in Fig. 3, results from the PerfectIC experiment are depicted with a solid line.

ed here are modifications of evenly spaced grids that are symmetric about the equator, similar to the network shown in Fig. 2 or the TAO array.

Because large amounts of information about Pacific SST and surface winds are routinely available from sources besides the TAO array (e.g., satellites), SST tends to be better observed than ocean heat content. Simulating SST and thermocline depth observations at the same locations may therefore be unrealistic. To account for this, a set of experiments similar to those presented in sections 4–5, but assuming many SST observations and moving only thermocline observations, was also performed. However, as discussed in section 3, the difference in forecast skill between the FullTh and FullT experiments is small—sufficiently small that when SST is well observed, it is often difficult to differentiate among results for different thermocline-observing networks. Consequently, we defer to MB2 for a discussion of important aspects of thermocline-observing networks when SST is already well observed.

a. Varying the meridional extent of observations

Figure 5 depicts how varying the meridional extent of observations in this system affects average analysis
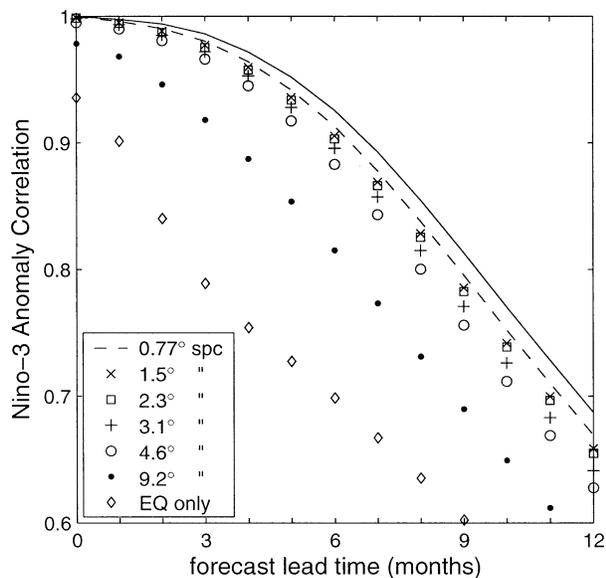
FIG. 6. As in Fig. 5, but for different meridional spacings of observations. All of the observing networks except EQ only are evenly spaced grids that are symmetric about the equator, with 15 north–south transects (spaced by 11° longitude) that observe from 9.2°S/N. The meridional spacings of observations vary from 0.77° to 9.2° and are indicated in the legend. The networks with 0.77°, 1.5°, 2.3°, 3.1°, 4.6°, and 9.2° spacing contain 375, 195, 135, 105, 75, and 45 observing locations, respectively; each time the spacing is doubled, the number of observing locations is approximately halved. The 0.77° spacing network is the same as the FullTh network shown in Fig. 2. The EQ only network has 15 observing locations at 15 longitudes along the equator.

and forecast skill. Each observing network tested has the same meridional spacing of observations (0.77°) and the same longitudinal distribution of observations (15 north–south transects, spaced by 11° longitude). All of the networks tested include observations at the equator and are symmetric about the equator; only the poleward extent of the network is changed.

Up to 9.2°N/S, extending the network poleward continues to increase forecast skill. Adding observations poleward of 9.2°, however, provides little benefit. Based on these results, 9.2° was chosen for the meridional extent of the FullTh network, and none of the remaining networks tested have observations poleward of 9.2°.

b. Varying the meridional spacing of observations

Next, we vary the meridional density of observations, keeping the same meridional extent (9.2°N/S) and longitudinal distribution of observations (15 north–south transects, spaced by 11°). To do so, we begin with the FullTh network shown in Fig. 2, which has observations every 0.77° in latitude, and gradually increase the meridional spacing until observations are located at only 9.2°S, 0°, and 9.2°N. The results are depicted in Fig. 6, along with the results when observations are taken only at the equator.

Figure 6 indicates that in this system, even a network with observations at only the equator can produce forecasts with some skill. Adding observations at other latitudes does benefit analyses and forecasts. However, a fairly large meridional spacing of observations, for example, 4.6°, is sufficient to produce analyses and forecasts that are reasonably accurate compared to what is possible. Decreasing the meridional spacing of observations from 4.6° to 3.1° or 2.3° improves forecasts somewhat; decreasing the spacing to less than 2.3° only increases average forecast skill by a small amount.

Decreasing the meridional spacing of observations provides additional information about the meridional structure of the errors in the initial conditions. It also adds observations in regions that are already observed; independent of added information about meridional structure, this can improve initial conditions by facilitating error checking among observations, decreasing the effective observation error. To evaluate how much of the forecast improvements shown in Fig. 6 are due to more information about meridional structure rather than simply more observations, we tested doubling the number of observations without changing the meridional spacing. For these "double observation" experiments, two simulated observations (with independently generated random errors) are taken at each observing location and time. The results (not shown) are then compared with those from the standard "single observation" experiments.

When double observations are taken for the 3.1° spacing network ($105 \times 2 = 210$ observations at each time), forecasts are at least as skillful as those from the single observation 1.5° spacing experiment (195 observations). And, when double observations are taken for the 2.3° spacing network (270 observations), forecasts are nearly as skillful as those from the single observation 0.77° spacing experiment (375 observations). This indicates that observations closer together than every 3.1° latitude improves forecasts primarily because it increases the number of observations rather than because it provides additional information about meridional structure. When double observations are taken for the 4.6° spacing network (150 observations), on the other hand, forecasts are somewhat more skillful than those from the single observation 4.6° spacing experiment, but noticeably less skillful than those from the single observation 2.3° spacing experiment (130 observations). This indicates that forecasts benefit from at least some information about initial condition errors on scales of less than 4.6° latitude.

c. Varying the longitudinal spacing of observations

The next set of experiments varies the longitudinal density of observations, keeping the meridional distribution (25 observations spaced equally from 9.2°S/N) the same. The FullTh network has observations spaced by 11° in longitude (the gridpoint spacing); Fig. 7 de-
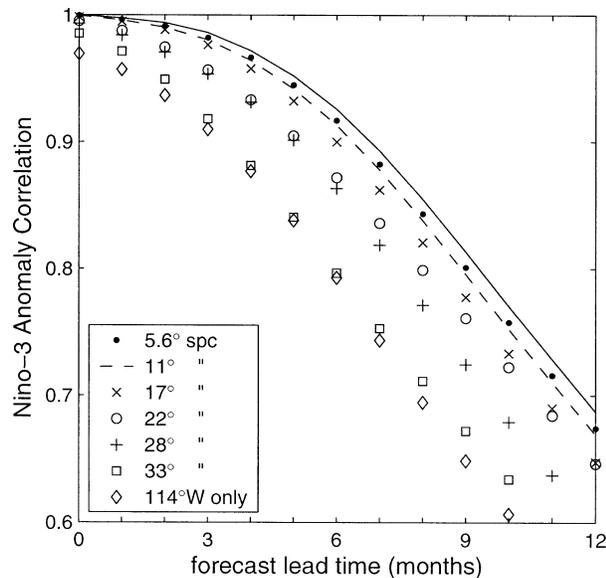
FIG. 7. As in Fig. 5, but for different longitudinal spacings of observations. All of the networks have observations at 25 latitudes spaced by 0.77° from 9.2°S/N. The networks with longitudinal spacings of 5.6°, 11°, 17°, 22°, 28°, and 33° contain 29, 15, 10, 8, 6, and 5 north–south transects, and 725, 375, 250, 200, 150, and 125 observing locations, respectively; each time the spacing is doubled, the number of observing locations is approximately halved. The 11° spacing network is the same as the FullTh network shown in Fig. 2. The 114°W only network has one north–south transect, observing at 25 latitudes along 114°W.

picts the results when the longitudinal spacing is decreased to 5.6°, and when it is gradually increased to 33°. Results are also depicted when observations are taken at only one longitude, 114°W.

As the 5.6° results in Fig. 7 indicate, decreasing the longitudinal spacing of observations to less than the gridpoint spacing improves forecasts very little. This is expected, given that the FullTh network generates nearly perfect initial conditions. Increasing the longitudinal spacing of observations to more than 11°, on the other hand, degrades forecasts. When two independent observations of each type are taken at each observing location (the double observation experiments described in section 4b), the results (not shown) indicate that observing every 11° in longitude benefits forecasts not only because it provides additional observations, but also because it provides additional information about the longitudinal structure of initial condition errors.

To evaluate the relative importance of the longitudinal and latitudinal spacing of observations, one can compare Fig. 7 with Fig. 6. For example, the network with 11° longitudinal and 2.3° latitudinal spacing in Fig. 6 produces forecasts that are approximately as skillful as the network with 17° longitudinal and 0.77° latitudinal spacing in Fig. 7, even though the first network contains fewer observing locations (135 versus 250). Such comparisons suggest that, in general, information about initial condition errors on longitudinal scales of 10°–20°

is as important for ENSO forecasts in this system as information on latitudinal scales of 2°–3°.

Note, however, that this comparison is only valid in general, since as the longitudinal density decreases, forecast skill depends increasingly on the specific longitudes at which observations are located. Take, for example, the results when observations are taken at only one longitude: forecasts can be as skillful as those shown in Fig. 7 for 114°W, or, for other longitudes (not shown), no more skillful than when no observations are taken. In addition, when observations are taken at only 114°W, forecasts are nearly as skillful as those from the 33° spacing network, which has observations at 5 longitudes (and 5 times as many observations). The 33° spacing network has observations at 126° and 92°W, but not at 114°W; this suggests that observations are more important at 114°W than at other longitudes. This dependence of forecast skill on the specific locations of observations is explored next.

5. Effects of removing observations in specific regions

To investigate where observations are most important for forecasting ENSO skillfully, we tested removing observations from the FullTh network (Fig. 2) in specific regions. Figure 8 depicts a sample set of results, when observations are removed at all longitudes for various 4.6°-wide bands of latitude. Note that unlike in Figs. 5–7, each of the networks tested in Fig. 8 contains the same number of observing locations; only the meridional distribution is different.

As Fig. 8a depicts, removing observations north of the equator, from 2.3°–6.9°N or from 4.6°–9.2°N, degrades forecasts very little. In contrast, removing observations in the same bands south of the equator degrades forecasts noticeably. Although such north–south asymmetry in the importance of observations could be caused by asymmetry in how the data are assimilated, recall from section 2c that the simple observation and background error covariances used in the data assimilation system are north–south symmetric, so that information from observations in the Northern and Southern Hemispheres is spread symmetrically. In addition, Fig. 8b, along with plots of the spatial distribution of time-averaged analysis and forecast errors in each experiment (not shown), indicates that removing observations north of the equator degrades analyses as much as removing observations south of the equator, but that the degraded analyses north of the equator have a much smaller influence on forecasts. In other words, in this system, accurate initial conditions south of the equator are more important for ENSO forecasting than accurate initial conditions north of the equator.

A closer look at Figs. 8a–b indicates that the latitudes where observations are most important depend on the forecast lead time; analyses and short lead-time forecasts (<3 months) are most degraded by removing ob-
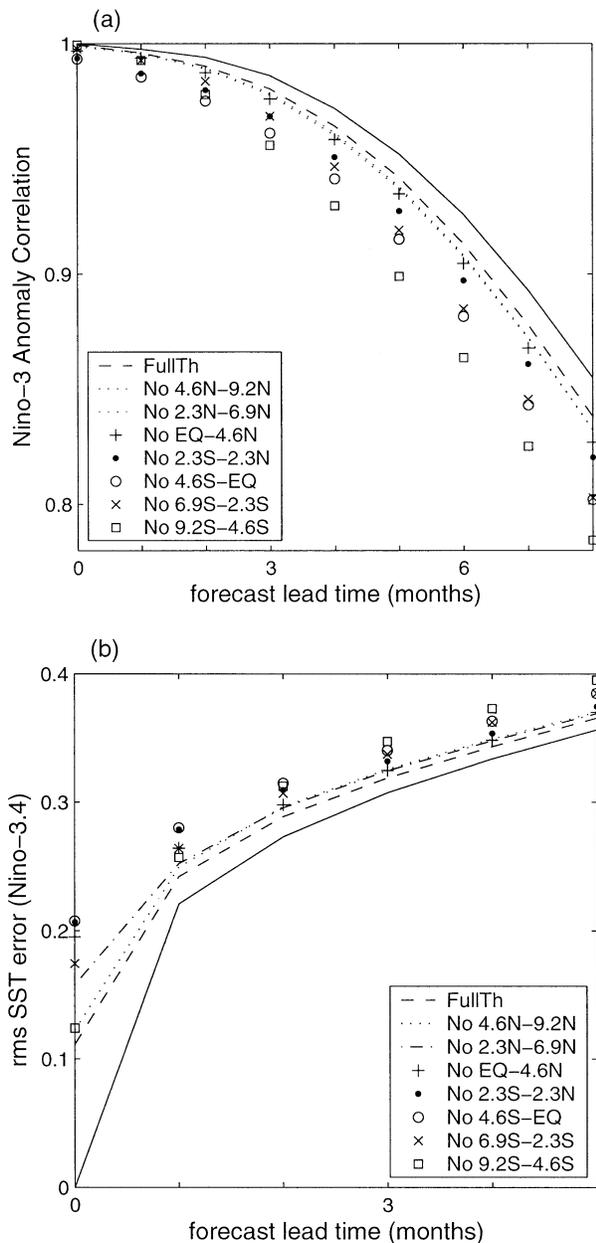
FIG. 8. (a) As in Fig. 5, but for the FullTh observing network (Fig. 2) with observations removed at all longitudes in different 4.6°-wide bands of latitude. The band in which observations have been removed is indicated in the legend; the No 2.3S–2.3N network, for example, has observations at 15 longitudes from 9.2°–3.1°S and from 3.1°–9.2°N. All of the networks contain 270 observing locations. (b) As in (a), but for rms SST error in the Niño-3.4 region. To facilitate viewing the results for short lead-time forecasts, results are shown only for lead times up to 9 months in (a) and 5 months in (b).

waves take only 2–3 months to travel east across the basin, and thus can affect forecasts quickly. Observations outside the equatorial waveguide, on the other hand, provide information about the Rossby waves in the oceanic initial conditions. Rossby waves not only travel more slowly than Kelvin waves, but also must travel westward and be reflected at the western boundary, then travel eastward as Kelvin waves before they can have a large effect on SST; the time scale for this adjustment is at least 6–9 months. This means that information about the initial conditions outside the equatorial waveguide tends to affect forecasts more at greater lead times. Consequently, as the lead time increases, the influence of observations farther south of the equator grows relative to that of observations closer to the equator.

Figure 9 depicts a sample set of results when observations are removed at all latitudes for different bands of longitude. Two of the networks tested, No 123E–167E and No 179E–137W, contain the same number of observing locations; each has no observations in a 44°-wide band in longitude. The third network, No 114W–92W, has more observing locations, with none only in a 22°-wide band in longitude.[3]

Figure 9a shows that removing observations from 114°–92°W degrades Niño-3 SST analyses and forecasts significantly more than removing observations from 123°–167°E or 179°E–137°W. The No 114W–92W network has fewer observations in the Niño-3 region than No 123E–167E and No 179E–137W; thus, the difference in Niño-3 *analysis* accuracy is expected. The difference in *forecast* skill in Fig. 9a, on the other hand, is not a consequence of the Niño-3 error norm. This is evident in Fig. 9b; when the same results are evaluated using a Niño-3.4 SST error norm, 2-month or longer lead-time forecasts are still degraded most by removing observations from 114°–92°W. In fact, even though removing observations from 179°E–137°W (in the Niño-3.4 region) significantly degrades Niño-3.4 SST analyses, it has little effect on Niño-3.4 SST forecasts longer than several months. Therefore, observations in the eastern third of the equatorial Pacific are most important for ENSO forecasting in this system. Observations in the western equatorial Pacific are of secondary importance, and observations in the central Pacific are of little importance.

Removing observations in different localized regions of the domain, that is, the FullTh network with a "hole" in a specific region, was also tested. The results (not shown) indicate that observations are most important at the intersection of the regions indicated by the results in Figs. 8 and 9—south of the equator, east of approximately 130°W.

servations near the equator, while 3-month or longer lead-time forecasts are most degraded by removing observations from 4.6°–9.2°S. Short lead-time forecasts are likely most affected by observations near the equator because these observations provide information about Kelvin waves in the oceanic initial conditions; Kelvin

[3] If observations are removed in a 44°-wide band in the eastern third of the domain (126°–81°W), the data assimilation system has difficulty interpolating information into the data void well enough to produce initial conditions that can be used in forecasting.
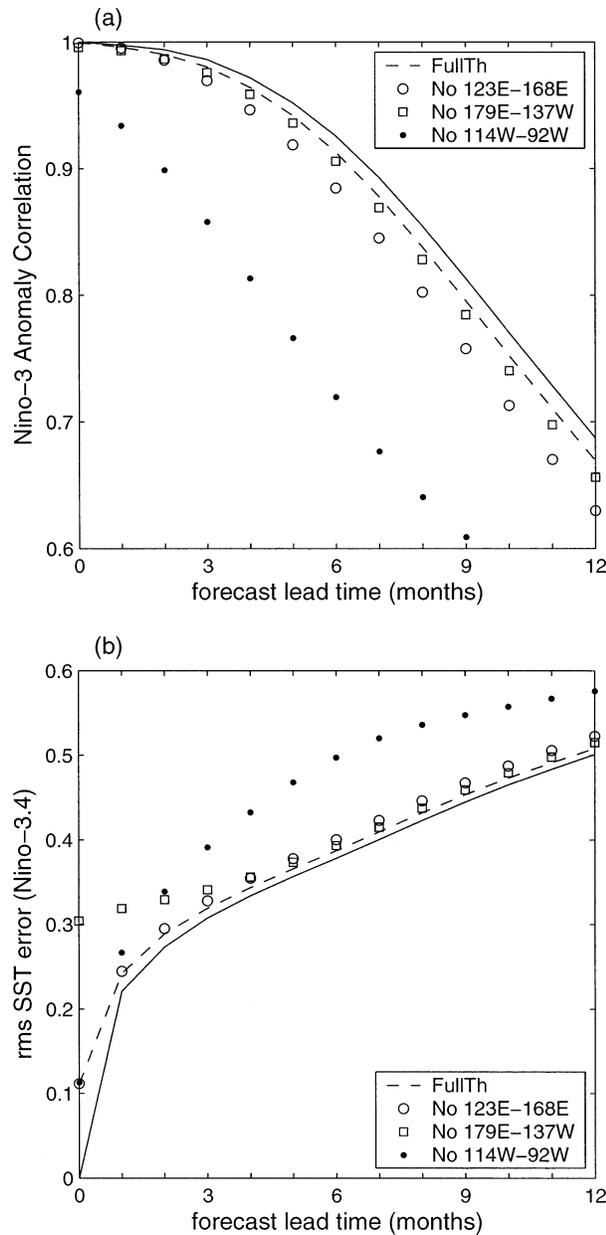
(a)

(b)

FIG. 9. (a) As in Fig. 5, but for the FullTh observing network (Fig. 2) with observations removed at all latitudes in different bands of longitude. The No 123E–167E and No 179E–137W networks have no observations in the western and central third of the domain, respectively, leaving 250 observing locations; the No 114W–92W network has no observations in a 22° band within the eastern third of the domain, leaving 300 observing locations. For clarity, results are not shown for networks with observations removed in other 22° bands within the eastern third of the domain; the No 103W–81W network produces slightly less skillful forecasts than the No 179E–137W network, and the No 126W–103W network produces forecasts that are less skillful than No 103W–81W but more skillful than No 114W–92W. (b) As in (a), but for rms SST error in the Niño-3.4 region.

(a) (5,9) SV1: SST

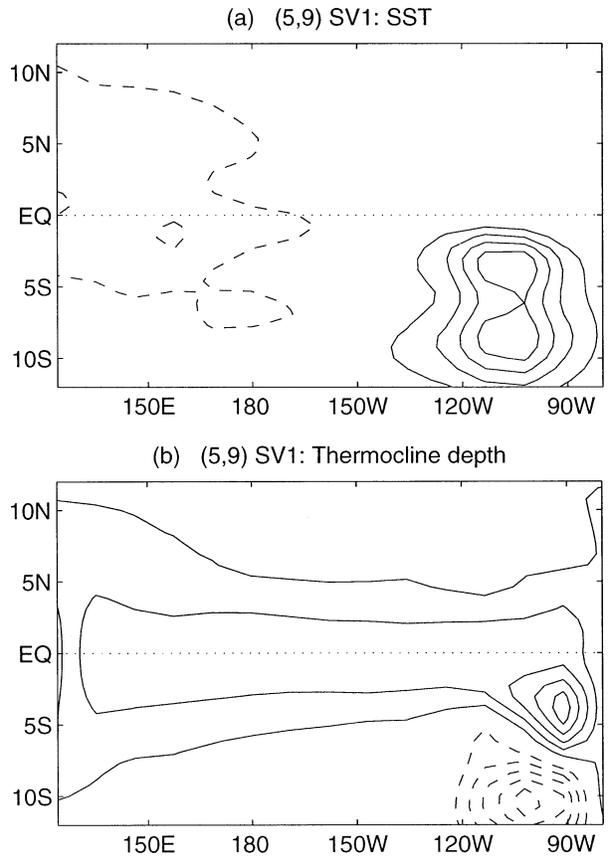(b) (5,9) SV1: Thermocline depth

FIG. 10. (a) SST and (b) thermocline depth fields of the leading singular vector (optimal perturbation) for a 1 May start date and a 9-month optimization period (5, 9). The contour intervals are 0.2°C in (a) and 1 m in (b). The singular vector is scaled so that the absolute value of the SST field has a maximum of 1°C. The sign of the singular vector is such that it leads to an El Niño event; the solid contours are positive and the dashed contours are negative. As discussed in Thompson (1998a), a 1-2-1 zonal filter is applied.

Error growth in this ENSO model is dominated by the leading singular vector (optimal perturbation), which develops into the ENSO mode (TB2000). The spatial structure of the leading singular vector therefore indicates which errors in initial conditions, if they exist, will grow most rapidly into errors in ENSO forecasts (in a specified norm). As in TB2000, the leading singular vector for this model has a similar structure for all start times and for all optimization times between 4 and 12 months. The leading singular vector for a 9-month optimization time starting in May, a local maximum in transient growth, is depicted in Fig. 10 (calculated using the model's adjoint as described in TB2000). Note that the leading singular vector has the largest amplitude in the southeastern tropical Pacific—in the same region where the OSSEs indicate that observations are most important. In other words, accurate initial conditions are most important in the regions where the ENSO model's singular vectors indicate that initial condition errors will grow rapidly into ENSO forecast errors.
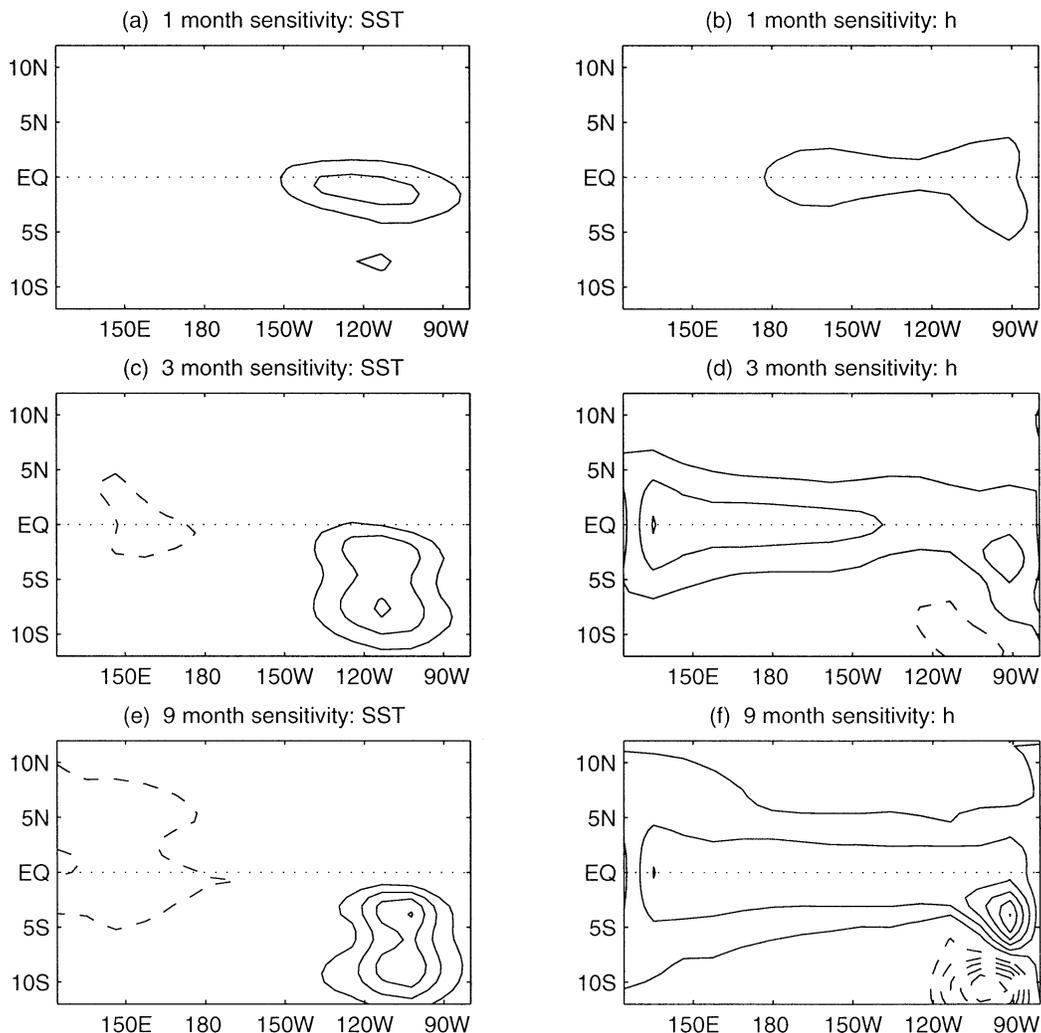
FIG. 11. Adjoint-derived sensitivities of (a), (b) 1-month, (c), (d) 3-month, and (e), (f) 9-month Niño-3 SST forecasts to 1 May perturbations, (a), (c), (e) in SST and (b), (d), (f) in thermocline depth initial conditions. To approximate the latitudinal extent for the Niño-3 SST norm, the first Hermite polynomial is used. The contour intervals are 0.5°, 0.75°, and 1.0°C in (a), (c), and (e), respectively, and 2 m, 3 m, and 4 m in (b), (d), and (f), respectively. Positive contours (indicating that initial perturbations lead to final perturbations of the same sign) are solid, and negative contours are dashed. As discussed in Thompson (1998a), a 1-2-1 zonal filter is applied.

Singular vectors identify the initial perturbations that will grow most rapidly, from which one can infer sensitivity of forecasts to initial perturbations. However, the sensitivity of a specific forecast norm to perturbations in initial conditions can also be calculated directly using the model's adjoint (for details see, e.g., van Oldenborgh et al. 1999 or Errico and Vukicevic 1992). Figure 11 depicts, for this ENSO model, the adjoint-derived sensitivities of Niño-3 SST forecasts to perturbations in initial conditions. Results are shown for 1-, 3-, and 9-month forecasts, all beginning in May. For forecast lead times of 6 months or greater, the adjoint-derived sensitivities are quite similar to the leading singular vector. At shorter lead times, however, the sensitivities show more clearly than the singular vectors how the most important regions for observations evolve with the lead

time: from the central-eastern equatorial Pacific for 1-month forecasts, to the western Pacific for 2–3 month forecasts, to farther off the equator in longer lead-time forecasts, as the model dynamics propagate information through the basin. This evolution of sensitive regions corroborates the results discussed earlier; the influence of off-equatorial information increases relative to the influence of near-equatorial information as the forecast lead time increases.

Chen et al. (1997) describe the physical reason for the southeast tropical Pacific SST maximum in this model's leading singular vector: it forces a northwesterly wind anomaly in the eastern equatorial Pacific, which relaxes the mean southeasterly wind stress in this region; this decreases upwelling of cold water into the surface layer, warming the eastern near-equatorial Pacific. As

discussed in Thompson (1998a), the thermocline depth singular vector in the southeast tropical Pacific affects eastern near-equatorial Pacific SST by deepening the thermocline, which affects the temperature of upwelled water. Thus, the primary mechanism for transferring information from observations in the southeastern tropical Pacific to the Niño-3 region is upwelling: (negative) anomalous upwelling acting on the background vertical temperature gradient and background upwelling of anomalously warm water. Information from observations in the western equatorial Pacific is transferred to the Niño-3 region primarily through eastward-propagating Kelvin waves. Further discussion of the dynamics of the leading singular vector can be found in Chen et al. (1997) and Thompson (1998a); a comparison with singular vectors derived in other studies is provided in MB2.

6. Summary and discussion

This study explores observing requirements for ENSO prediction by performing observing system simulation experiments (OSSEs) with a linearized intermediate coupled ENSO model, a variant of the Zebiak–Cane (1987) model, stochastically forced. Results are compared for forecasts generated with a variety of observation densities, spacings, and distributions, averaged over 1000 years of simulated ENSO events. The experiments show that it is possible to use this type of ENSO model in an OSSE framework to evaluate observing networks for ENSO prediction. Like all OSSE results, the results from this study are only applicable to the real world to the extent that the simulated system represents the true system. At a minimum, however, the results identify features of observing networks that are important for predicting ENSO in this system, suggesting consideration for future studies of ENSO observing networks (with OSSEs and with real data).

Because the ENSO model is linearized and, for the parameters used here, moderately damped, forecast skill is limited by the stochastic forcing, which represents atmospheric transients. The skill of forecasts from perfect initial conditions can therefore be used to estimate the potential predictability of ENSO. The results indicate that, in this system, ENSO forecasts with Niño-3 anomaly correlations greater than 0.6 are on average possible only for lead times shorter than 15 months. When many SST and thermocline depth observations are assimilated, forecasts are only slightly less skillful than those from perfect initial conditions, indicating that the data assimilation system performs well when many observations are available. The remaining results are interpreted with respect to these estimates of maximum attainable forecast skill.

When only thermocline depth observations are available, SST forecasts are reasonably skillful, indicating the importance of ocean memory for the ENSO system. However, nearly all of the ENSO forecast skill possible in this system can be achieved by observing and assimilating only SST—adding observations of thermocline depth on average improves forecasts only slightly. Recall that the data assimilation system contains no correlations between SST and thermocline depth information. Consequently, the only mechanism for transferring information from SST to thermocline depth (or vice versa) is the model dynamics.

The relative effects of ocean heat content and ocean surface information on ENSO forecasts have been investigated in several other studies, using real data to hindcast past ENSO events. Ocean surface information is generally incorporated by inserting SST data or by forcing with atmospheric winds; sources for ocean heat content information include subsurface temperature, thermocline depth, and sea level observations. Studies with coupled ocean–atmosphere general circulation models have found, using several different models and data assimilation systems, that assimilating subsurface ocean information improves ENSO forecasts for at least some data combinations, model configurations, forecast start and lead times, and forecast norms (Ji and Leetmaa 1997; Rosati et al. 1997; Ji et al. 1998; Segschneider et al. 2001; Wang et al. 2002). Using dynamical ocean models coupled to statistical atmospheric models (hybrid models), Fischer et al. (1997) and Syu and Neelin (2000) also found that inserting subsurface ocean information was beneficial, as did Kleeman et al. (1995) and Chen et al. (1998) using intermediate coupled models. With Markov models, Xue et al. (2000) found that subsurface data were very important for ENSO prediction—in fact, the most important predictor tested—while Johnson et al. (2000) and C. Penland (2001, personal communication) found that subsurface data were a relatively unimportant predictor compared to SST. Thus, many (but not all) previous studies suggest that initializing ENSO prediction schemes with ocean heat content information as well as SST and/or surface wind information increases ENSO forecast skill.

Results from this study about the relative importance of subsurface and surface ocean information could disagree with those from other studies for several reasons. First, the ENSO model used in this study may be missing aspects of the real ENSO system that cause subsurface information to be more important. For example, thermocline perturbations may be too closely linked to SST perturbations in the model compared to the real world. In addition, subsurface data benefit real-world numerical model hindcasts at least in part because they help compensate for errors in the model's formulation and climatology, such as a tendency to generate a thermocline that is too diffuse when forced with only SST and/or wind stress data (Rosati et al. 1997; Ji and Leetmaa 1997). Since this study uses the same dynamical ENSO model to forecast ENSO as it does to generate the synthetic ENSO events being forecasted, the forecast model has no errors in dynamics for which to compensate,

resulting, perhaps, in a much smaller benefit from incorporating subsurface observations.

It is sometimes argued that because the ocean contains most of the memory of the tropical Pacific coupled ocean–atmosphere system, intelligently incorporating subsurface ocean information into prediction models is an important component of skillful ENSO prediction (e.g., Kleeman 1993; Kleeman et al. 1995; Miller et al. 1995; Fischer et al. 1997; Ji and Leetmaa 1997; Chen et al. 1998; Schneider et al. 1999; Xue et al. 2000; Perigaud et al. 2000; Segschneider et al. 2000). However, for the ENSO mode that is supported in the model used in the present study, the ocean does contain most of the memory, yet nearly all of the forecast skill possible can be achieved with only SST observations. Hence, our results suggest that this inference may not be valid. Naturally, this result may be a consequence of the ENSO model used and the experimental setup, and thus requires further study.

Several studies have suggested that because the minimum in lag autocorrelation of upper-ocean heat content occurs at a different time of year than the minimum in lag autocorrelation of SST, accurate initialization of heat content may reduce the SST "spring prediction barrier" evident in many ENSO prediction schemes (e.g., Balmaseda et al. 1995; McPhaden 2003). To investigate this hypothesis, the seasonal variability in forecast skill from OSSEs with and without thermocline depth observations was also explored. In this system, adding ocean heat content observations has little effect on the spring prediction barrier. Instead, it improves SST forecasts that begin in boreal summer, at a range of lead times.

Experiments were also conducted to explore how moving SST and thermocline depth observing locations affects forecast errors. The results identify several important aspects of observing networks for ENSO prediction in this system. First, SST or thermocline depth observations poleward of approximately 10° contribute little to forecast skill. Note, however, that at this resolution of the T80 ENSO model, analysis increments have little amplitude poleward of 10°, even when observations are taken poleward of 10°. Furthermore, the ocean component of the model does not adequately represent the physics responsible for subtropical SST anomalies (poleward of approximately 10°). Hence, to the extent that subtropical SST anomalies affect the winds within the equatorial wave guide (see, e.g., Vimont et al. 2003), the present coupled model is insufficient to evaluate the importance of observations in the subtropics.

Second, predicting ENSO skillfully in this system on average requires observations (of SST and thermocline depth simultaneously) spaced longitudinally by approximately 10°–15° and meridionally by approximately 2.0°–2.5°. More complex, real-world forecasting systems may benefit from more closely spaced observations due to observational limitations, more complex dynam-

ics, or higher model resolution. However, when multiple observations are simulated independently at each observed location and time, the results suggest that for some of the observing networks tested in this study, forecast skill is limited not by the observational spacing, but rather by the observations' errors or their weighting by the data assimilation system. In particular, a meridional observation spacing of 3.0°–4.5° appears to be sufficient.

Third, in this system, observations in some regions are significantly more important for skillful ENSO forecasting than observations in other regions. For example, SST and thermocline depth observations between 5° and 10°S have a major effect on forecast skill, while observations north of 5°N have little effect. For shorter (2 months or less) lead-time forecasts, observations near the equator tend to have the largest influence, while for longer lead-time forecasts, off-equatorial observations tend to be more important. Observations in the eastern tropical Pacific have a major influence on ENSO forecasts, observations in the western tropical Pacific have less influence, and observations in the central Pacific (from approximately 140°W to the date line) have almost no influence, except on analyses and 1–3-month forecasts in the Niño 3.4 region. The most important region for observations in this system is south of the equator, east of approximately 130°W. This coincides with the region where the leading singular vector and adjoint-derived sensitivities for this ENSO model indicate that perturbations in initial conditions are most likely to affect ENSO forecasts. Previous studies of the dynamics of this model's leading singular vector indicate that information from the southeastern tropical Pacific is communicated to the Niño-3 region primarily through upwelling.

These results suggest that uniformly spaced observations may not be the most cost-effective configuration for ENSO-observing networks. Based on the results presented here, a continuation of this study, MB2, develops and discusses observing networks that are efficient for predicting ENSO in this system, that is, that produce reasonably skillful forecasts with relatively few observations. Two types of efficient observing networks are addressed: networks that observe SST and thermocline depth at the same locations, and networks of thermocline observations when many SST observations are available from another source (e.g., satellites). A major difference between this study and many other studies of ENSO prediction is the number of ENSO events examined; consequently, MB2 also examines the dependence of the OSSE results on the duration of the simulated data record.

postdoctoral fellowship at the National Center for Atmospheric Research.

REFERENCES

Alves, J. O. S., K. Haines, and D. L. T. Anderson, 2001: Sea level assimilation experiments in the tropical Pacific. *J. Phys. Oceanogr.,* **31,** 305–323.

Ballabrera-Poy, J., A. J. Busalacchi, and R. Murtugudde, 2001: Application of a reduced-order Kalman filter to initialize a coupled atmosphere–ocean model: Impact on the prediction of El Niño. *J. Climate,* **14,** 1720–1737.

Balmaseda, M. A., M. K. Davey, and D. L. T. Anderson, 1995: Decadal and seasonal dependence of ENSO prediction skill. *J. Climate,* **8,** 2705–2715.

Battisti, D. S., 1988: Dynamics and thermodynamics of a warming event in a coupled tropical atmosphere–ocean model. *J. Atmos. Sci.,* **45,** 2889–2919.

Blanchet, I., C. Frankignoul, and M. A. Cane, 1997: A comparison of adaptive Kalman filters for a tropical Pacific Ocean model. *Mon. Wea. Rev.,* **125,** 40–58.

Chen, D., 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., D. S. Battisti, T. N. Palmer, J. Barsugli, and E. S. Sarachik, 1997: A study of the predictability of tropical Pacific SST in a coupled atmosphere–ocean model using singular vector analysis: The role of the annual cycle and ENSO cycle. *Mon. Wea. Rev.,* **125,** 831–845.

Derber, J., and A. Rosati, 1989: Global oceanic data assimilation system. *J. Phys. Oceanogr.,* **19,** 1333–1348.

Errico, R. M., and T. Vukicevic, 1992: Sensitivity analysis using an adjoint of the PSU–NCAR Mesoscale Model. *Mon. Wea. Rev.,* **120,** 1644–1659.

Fischer, M., M. Latif, M. Flügel, and M. Ji, 1997: The impact of data assimilation on ENSO simulations and predictions. *Mon. Wea. Rev.,* **125,** 819–829.

Goswami, B. N., and J. Shukla, 1991: Predictability of a coupled ocean–atmosphere model. *J. Climate,* **4,** 3–22.

Hao, Z., and M. Ghil, 1994: Data assimilation in a simple tropical ocean model with wind stress errors. *J. Phys. Oceanogr.,* **24,** 2111–2128.

Ji, M., and A. Leetmaa, 1997: Impact of data assimilation on ocean initialization and El Niño prediction. *Mon. Wea. Rev.,* **125,** 742–753.

——, ——, and J. Derber, 1995: An ocean analysis system for seasonal to interannual climate studies. *Mon. Wea. Rev.,* **123,** 460–481.

——, D. W. Behringer, 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.

Johnson, S. D., D. S. Battisti, and E. S. Sarachik, 2000: Empirically derived Markov models and prediction of tropical Pacific sea surface temperature anomalies. *J. Climate,* **13,** 3–17.

Kleeman, R., 1993: On the dependence of hindcast skill on ocean thermodynamics in a coupled ocean–atmosphere model. *J. Climate,* **6,** 2012–2033.

——, A. M. Moore, and N. R. Smith, 1995: Assimilation of subsurface thermal data into a simple ocean model for the initialization of an intermediate tropical coupled ocean–atmosphere forecast model. *Mon. Wea. Rev.,* **123,** 3103–3113.

Latif, M., and N. E. Graham, 1992: How much predictive skill is contained in the thermal structure of an oceanic GCM? *J. Phys. Oceanogr.,* **22,** 951–962.

——, and Coauthors, 1998: A review of the predictability and prediction of ENSO. *J. Geophys. Res.,* **103,** 14 375–14 393.

McPhaden, M. J., 2003: Tropical Pacific Ocean heat content variations and ENSO persistence barriers. *Geophys. Res. Lett.,* **30,** 1480, doi:10.1029/2003GL016872.

Miller, R. N., 1990: Tropical data assimilation experiments with simulated data: The impact of the tropical ocean and global atmosphere thermal array for the ocean. *J. Geophys. Res.,* **95,** 11 461–11 482.

——, A. J. Busalacchi, and E. C. Hackert, 1995: Sea surface topography fields of the tropical Pacific from data assimilation. *J. Geophys. Res.,* **100,** 13 389–13 426.

Morss, R. E., and D. S. Battisti, 2004: Designing efficient observing networks for ENSO prediction. *J. Climate,* **17,** 3074–3089.

Perigaud, C. M., C. Cassou, B. Dewitte, L.-L. Fu, and J. D. Neelin, 2000: Using data and intermediate coupled models for seasonal-to-interannual forecasts. *Mon. Wea. Rev.,* **128,** 3025–3049.

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. P. Kirtman, and J. Shukla, 1999: Ocean data assimilation, initialization, and predictions of ENSO with a coupled GCM. *Mon. Wea. Rev.,* **127,** 1187–1207.

Segschneider, J., D. L. T. Anderson, and T. N. Stockdale, 2000: Toward the use of altimetry for operational seasonal forecasting. *J. Climate,* **13,** 3115–3138.

——, ——, J. Vialard, M. Balmaseda, T. N. Stockdale, A. Troccoli, and K. Haines, 2001: Initialization of seasonal forecasts assimilating sea level and temperature observations. *J. Climate,* **14,** 4292–4307.

Smith, T. M., A. G. Barnston, M. Ji, and M. Chelliah, 1995: The impact of Pacific Ocean subsurface data on operational prediction of tropical Pacific SST at the NCEP. *Wea. Forecasting,* **10,** 708–714.

Sun, C., Z. Hao, M. Ghil, and J. D. Neelin, 2002: Data assimilation for a coupled ocean–atmosphere model. Part I: Sequential state estimation. *Mon. Wea. Rev.,* **130,** 1073–1099.

Syu, H.-H., and J. D. Neelin, 2000: ENSO in a hybrid coupled model. Part II: Prediction with piggyback data assimilation. *Climate Dyn.,* **16,** 35–48.

Thompson, C. J., 1998a: Initial conditions for optimal growth in a coupled ocean–atmosphere model of ENSO. *J. Atmos. Sci.,* **55,** 537–557.

——, 1998b: A linear, stochastic, dynamical model of El Niño/Southern Oscillation. Ph.D. thesis, University of Washington, 235 pp.

——, and D. S. Battisti, 2000: A linear stochastic dynamical model of ENSO. Part I: Model development. *J. Climate,* **13,** 2818–2832.

——, and ——, 2001: A linear stochastic dynamical model of ENSO. Part II: Analysis. *J. Climate,* **14,** 445–466.

van Oldenborgh, G. J., G. Burgers, S. Venzke, C. Eckert, and R. Giering, 1999: Tracking down the ENSO delayed oscillator with an adjoint OGCM. *Mon. Wea. Rev.,* **127,** 1477–1495.

Vimont, D. J., J. M. Wallace, and D. S. Battisti, 2003: The seasonal footprinting mechanism in the Pacific: Implications for ENSO. *J. Climate,* **16,** 2668–2675.

Wang, G., R. Kleeman, N. Smith, and F. Tseitkin, 2002: The BMRC coupled general circulation model ENSO forecast system. *Mon. Wea. Rev.,* **130,** 975–991.

Wright, P. B., 1985: The Southern Oscillation: An ocean–atmosphere feedback system? *Bull. Amer. Meteor. Soc.,* **66,** 398–412.

Xue, Y., M. A. Cane, S. E. Zebiak, and M. B. Blumenthal, 1994: On the prediction of ENSO: A study with a low-order Markov model. *Tellus,* **46A,** 512–528.

——, A. Leetmaa, and M. Ji, 2000: ENSO prediction with Markov models: The impact of sea level. *J. Climate,* **13,** 849–871.

Zebiak, S. E., and M. A. Cane, 1987: A model El Niño–Southern Oscillation. *Mon. Wea. Rev.,* **115,** 2262–2278.