Assimilation of Subsurface Thermal Data into a Simple Ocean Model for the Initialization of an Intermediate Tropical Coupled Ocean–Atmosphere Forecast Model

RICHARD KLEEMAN, ANDREW M. MOORE, AND NEVILLE R. SMITH

Bureau of Meteorology Research Centre, Melbourne, Victoria, Australia

(Manuscript received 27 October 1994, in final form 12 April 1995)

ABSTRACT

An adjoint variational assimilation technique is used to assimilate observations of both the oceanic state and wind stress data into an intermediate coupled ENSO prediction model. This method of initialization is contrasted with the more usual method, which uses only wind stress data to establish the initial state of the ocean. It is shown that ocean temperature data has a positive impact on the prediction skill in such models. On the basis of hindcasts for the period 1982–91, it is shown that NINO3 SST anomaly correlations greater than 0.7 can be obtained for hindcasts of duration up to 13 months and greater than 0.6 up to 16 months. There are also clear indications of skill at two years.

1. Introduction

During the past decade a large number of coupled ocean–atmosphere models have been developed for the purpose of exploring interannual variability in the tropical Pacific. These have included the relatively simple or “intermediate” coupled models (e.g., Zebiak and Cane 1987; and Kleeman 1993); the so-called hybrid coupled models in which one component is simple (usually the atmosphere) and the other a general circulation model (e.g., Neelin 1990); and the fully coupled general circulation models (e.g., Philander et al. 1992; and Latif et al. 1993).

The first dynamical model to demonstrate an ability to predict the El Niño–Southern Oscillation (ENSO) was that of Zebiak and Cane (1987). Their prediction system involved establishing the initial state of the ocean by forcing the ocean-model component of the coupled model with the observed wind anomalies for several years prior to the forecast date. Forecasts were then made by running the coupled model forward in time from the established ocean initial conditions. Note that the atmospheric model does not need initialization since it is effectively a “slave” to the ocean state as expressed through the model sea surface temperature (SST). This particular forecast system has also been used with success by Latif et al. (1993) to initialize their coupled general circulation model.

Although there has clearly been a great deal of success in using winds to establish the ocean state, it seems intuitively reasonable to expect that direct observations of the ocean state should positively impact on model initialization. Due largely to the volunteer observing ship XBT (expendable bathythermograph) program (TWXPPC 1993) and the TOGA TAO array (Hayes et al. 1991), subsurface thermal observations now exist at a reasonable density and for a long enough period to test their impact on the predictive skill of coupled models.

Very recently, coupled general circulation models in which the ocean state is established with both winds and subsurface thermal data have been shown to have potential predictive capacity (Ji et al. 1994; Rosati and Miyakoda 1994, personal communication). The question therefore arises as to whether direct observations can be used in the initialization of an intermediate coupled model. The answer to this is not immediately obvious because of the highly parameterized nature of the ocean models used. The results of Kleeman (1993), however, provide a strong indication of how the observations might be used: He showed that good skill can be obtained in an intermediate coupled model if thermocline perturbations dominate the SST equation of the model. Providing such a dominance exists, the skill of the model is relatively insensitive to the particular way in which the SST parameterization is carried out. This central role for thermocline perturbations in forecasting has also been argued for from OGCM results by Latif and Graham (1992). Such a situation strongly suggests that if an adequate proxy for thermocline perturbations can be obtained from the ocean subsurface thermal data, then the latter data will be of use in the coupled model. In the present work we demonstrate that such a proxy exists; that the subsurface thermal data can be successfully assimilated into a shal-
low water equation model, and that the data can increase the skill of an intermediate coupled model.

The paper is organized as follows: section 2 contains a brief description of the coupled model used; section 3 provides a description of the thermal data used and also justifies the proxy relation used to assimilate the data into the coupled model; section 4 describes the assimilation method and the results of assimilating both wind and thermal data; section 5 shows the impact of the ocean assimilation on the hindcast skill of the coupled model; and section 6 contains a summary and discussion of the results obtained and suggestions for further work.

2. Coupled model description

The coupled model to be used here has been described in detail elsewhere (Kleeman 1991; and Kleeman 1993). We provide a brief description here and refer the reader to these publications for a more extensive discussion.

The atmospheric component of the model consists of a steady-state dynamical Gill (1980) model forced by a latent and a direct thermal heating. The former kind of heating dominates in general and is obtained by vertically integrating the steady state moisture equation with the assumption that all atmospheric quantities have a prescribed vertical structure.

The model tends to respond more strongly to SST anomalies in regions where the total SST is greatest. Its performance has been assessed by forcing it with a wide range of observed SST anomaly patterns and generally it positions wind and precipitation anomalies in good agreement with observations.

The ocean model consists of a dynamical component given by the longwave approximation of the shallow-water equations with the Kelvin mode and lowest six Rossby modes retained. Meridional boundaries are located at 124°E and 80°W. The SST is assumed to have a fixed meridional structure about the equator with an e-folding radius of 10°. The equatorial SST equation is given by

$$ T_r = \alpha(x) \theta(h) - \kappa T_e, $$

where $h$ is the equatorial thermocline perturbation (obtained from the dynamical model), while the functions $\alpha$ and $\theta$ are depicted in Fig. 1. The function $\alpha$ is intended to model the observed fact that thermocline perturbations are much more associated with SST anomalies in the eastern rather than the western Pacific. The function $\theta$ provides the simplest possible amplitude limiting nonlinearity. The final term in Eq. (1) represents the damping effects on equatorial SST anomalies of the mean equatorial upwelling and the heat flux response.

As was emphasized in the previous section, the skill of the coupled model is relatively insensitive to the details of $\alpha$, $\theta$, and $\kappa$. They (and the shallow-water speed of the ocean $c_0$) are chosen to roughly optimize the hindcast skill for the period 1972–86 obtained by initializing the ocean model with FSU winds. Relevant values are listed in Table 1.

The horizontal resolution of the atmospheric model and zonal resolution of the ocean SST grid are approximately 3°. The ocean dynamics is solved exactly by the method of characteristics and thus the time step was chosen so that each meridional mode traveled one grid point per step with the Kelvin mode grid being the same as that of SST (the Rossby mode grids are appropriately finer by factors of 3, 5, . . . , 11).

Coupling is achieved via a linearized stress law with SST and wind stress anomalies being exchanged between ocean and atmosphere once a month (there is little sensitivity if a shorter exchange period is used).

Table 1. Model parameter values.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Meaning</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha_0$</td>
<td>western thermocline</td>
<td>$6.8 \times 10^{-8}$ °C m$^{-1}$ s$^{-1}$</td>
</tr>
<tr>
<td>$\kappa$</td>
<td>relaxation coefficient</td>
<td>$2.72 \times 10^{-7}$ s$^{-1}$</td>
</tr>
<tr>
<td>$h_{max}$</td>
<td>limiting thermocline depth</td>
<td>22.5 m</td>
</tr>
<tr>
<td>$c_0$</td>
<td>shallow-water speed</td>
<td>2.3 m s$^{-1}$</td>
</tr>
</tbody>
</table>
Hindcast tests are performed by initializing the ocean model every three months for a specified time period. In the trials to be described below, the 10-yr period 1982–91 was chosen to test the hindcast skill of the model.

3. Thermal data

The subsurface information to be assimilated into the ocean model was derived from the statistical analysis scheme of Smith et al. (1991) with improvements presented in Smith (1995). We provide a brief overview of method and results and the interested reader is referred to the above papers for further details.

The analysis scheme uses the method of optimal interpolation to combine irregular observations of ocean temperature with a statistical forecast based on previous analyses to produce an estimate of the subsurface thermal structure on a regular space time grid (1° in latitude and longitude; 1 month in time). The observations are principally from the volunteer observing ship XBT program and from the TOGA TAO array. The statistical forecast is a method for extrapolating past information to the period of the analysis. For each month the analysis system produces a global estimate of the temperature field at specific levels as well as estimates of various derived fields, such as the 400-m depth-averaged temperature, together with an estimate of error variance for the fields. The error variance for the 400-m depth-averaged temperature anomaly, denoted here by \( \sigma_T^2 \), essentially represents the spatial and temporal sampling density of the effective, independent information contained in the database for the period in question: \( \sigma_T^2 \) becomes very large in data-sparse regions and small in well-sampled regions [a lower limit of \((0.1^\circ C)^2\) is imposed by the analysis system, which represents crudely the limit of instrumental accuracy].

The analysis scheme is able to capture in a reasonably smooth way the interannual variations in the upper-ocean heat content. This can be seen in Fig. 2, which shows an analysis of the 400-m depth-averaged equatorial temperature anomaly for the period 1982 to the present. The scheme is sufficiently robust that when a 10-day rather than the usual monthly analysis period is used, equatorial Kelvin waves propagating from the western to the eastern Pacific can clearly be discerned.

As discussed in the previous section, a proxy for modeled thermocline perturbations is required in order to assimilate the thermal data successfully. Two possibilities are the depth of the 20°C isotherm and the upper-ocean heat content (we use the 400-m averaged temperature to represent this). The physical basis of the first is that the 20°C isotherm is usually located in the middle of the thermocline in the climatology of the equatorial Pacific. The basis for the second is that it provides a good measure of the activity of the first two baroclinic modes since it is intimately related to sea-level variations and Cane (1984) has shown that observations of the latter can essentially be explained by a model involving the former. The assumption underlying the ocean model is that these low-order baroclinic modes provide an estimate of the thermocline perturbation. A little thought shows that the second proxy is more likely to be physically related to the model \( h \).

To test the ability of the ocean model to assimilate the proxies, EOF analyses of the proxies obtained using the thermal analysis above were computed for the period 1980–93. The same analysis was also carried out for the ocean model \( h \) computed by forcing the model with the FSU winds for the same period. From this it was concluded (as expected intuitively) that the second proxy was in better agreement with the model \( h \). Depicted in Figs. 3a and 3b are the spatial patterns for the first EOFs of model \( h \) and heat content from the statistical analysis. In Fig. 3c the time series associated with the EOFs are compared. Note that units have been rescaled to aid in comparison. As can be seen, there is
good agreement between the spatial patterns in Figs. 3a and 3b and excellent agreement between the time series. Also both EOFs explain a similar large amount of the variance (~45%). In view of this, this proxy was used in all assimilation experiments. The conversion coefficient between the two variables was taken from the rescaling constant used in the first EOF comparison and has the value 22.8 m °C⁻¹.

4. Assimilation method and results

The technique used to assimilate information into the ocean model is a full space–time variational technique and uses a standard adjoint technique. Discussion of this technique may be found in many places in the literature, for example, in Lewis and Derber (1985), Le Dimet and Talagrand (1986), and also Thacker.
We provide a brief overview of this technique here with particular emphasis on the distinguishing details of our methodology.

A cost function $J$ is introduced to measure the deviation of $h$ from the observed values of the same quantity over the assimilation domain and time interval. A term is also added to ensure that solutions are smooth in the zonal direction:

$$J = \int dxdydt \left[ \frac{(h_m - h_o)^2}{(\sigma_n)^2} + \gamma^2 \left( \frac{\partial^2 h_m}{\partial x^2} \right)^2 \right]. \quad (2)$$

The subscript $m$ refers to a model value, and $o$ refers to an "observed" value derived as outlined in the previous section. Here $\sigma_n$ is the estimated error in the observed value of $h$ and is determined from the Smith analysis estimate of $T'_{100}$ error $\sigma_T$, by multiplying by the proxy conversion coefficient discussed in the previous section. The constant $\gamma$ is set at such a value that zonal wavelengths of $15^\circ$ in $h_m$ with amplitude $\Delta h$ are penalized as much as deviations of $\Delta h$ in $h_m$ about $h_o$ in regions of excellent data coverage (i.e., where $\sigma_T$ reaches its instrumental minimum of $0.1^\circ$). The justification for the term is that model perturbations should be smooth on the same scales as the analyses (these have zonal decorrelation scales of $15^\circ$ based on observations in the equatorial waveguide by Meyers et al. 1991). A similar term was not required in the meridional direction because of the six Rossby mode truncation used in the model. Without the term, a small amount of spurious high zonal wavenumber noise is present in assimilation runs due to the very large spatial variations in $\sigma_T$ (see Thacker 1988 for further discussion on this point).

The basic idea of the technique used is to vary the initial conditions of the ocean model in such a way as to minimize $J$. To do this, the gradient of $J$ with respect to the initial conditions is required and we obtain this by the method of Lagrange multipliers: The dynamical equations governing the ocean model are solved using the method of characteristics for the ancillary variables $q$ (see Gill 1982, p. 435) which describe the amplitudes of the equatorial ocean waves and satisfy the equations

$$q_{n+1} = Aq_n + BZ_n, \quad (3)$$

where the subscripts refer to the time-step number and the vector notation (bold) indicates that the $q$ have a zonal and meridional mode dependence. The matrix $A$ gives the form of the dynamical evolution of the model, whereas the constant matrix $B$ operates on

$$Z_n = \begin{bmatrix} \tau_n(t_n) \\ \tau'_n(t_n) \end{bmatrix}$$

to give the forcing for each component of $q$. Note that the ocean-model forcing is kept constant for one month at a time.

The Lagrangian is obtained by appending the inner product of Lagrange multipliers $\lambda_n$ with the equations of motion [Eq. (3)] to the cost function:

$$L = J + \sum_{j=1}^{N} \lambda_j (q_j - Aq_{j-1} - BZ_{j-1}). \quad (4)$$

The gradient of the cost function with respect to the variables of interest is then obtained by taking the partial derivative of $L$ with respect to this variable. At the extrema of $L$ (and $J$), the gradient of $L$ with respect to the $\lambda_n$ and with respect to the $q_n$ (except $q_0$, the initial conditions) is zero. With this assumption the original equations of motion are obtained as well as the adjoint equations for the model:

$$\lambda_n - (A)^* \lambda_{n+1} + \frac{\partial J}{\partial q_n} = 0 \quad n = 1, \ldots, N, \quad (5)$$

where $(A)^*$ is the adjoint (or transpose) matrix to $A$.

The gradient of $J$ with respect to $q_0$, the initial conditions, is obtained from $\partial L/\partial q_0$

$$\nabla_{q_0} J = -(A)^* \lambda_1 = -\lambda_0, \quad (6)$$

where the equivalence in this equation means we are extending Eq. (5) to include $n = 0$. With the gradients in Eq. (6), the model $q_0$, which minimize $J$ can now be obtained as follows:

A first guess for $q_0$ is used to integrate the ocean model [Eq. (3)] over the assimilation period. This then defines a trajectory $q_n$. The adjoint model [Eq. (5)] is then integrated back in time from the final time to the initial time using the condition $\lambda_{N+1} = 0$. The trajectory $q$, determines the third term in Eq. (5). This integration defines an adjoint trajectory $\lambda_n$, which then enables the gradients from Eq. (6) to be evaluated. A conjugate gradient technique due to Polak and Ribiere (1969) (see also Moore 1991; Thacker 1990) is then used to compute the new initial conditions, $q_0$.

The procedure is continued iteratively with a decreasing $J$ at every step until convergence to the minimum of $J$ is obtained. In general, around 100–200 iterations are required depending on the length of the assimilation and the particular time period of data assimilation. For the experiments to be described below, 200 iterations were always used. An example of cost function convergence as a function of iteration number is displayed in Fig. 4. It is worth observing that the cost function for this problem can never, in general, be reduced to zero even if the third smoothing term in the cost function is ignored. This is because the initial conditions for the assimilation alone, which are defined for one time slice only, would have to be varied in such a way to ensure the observed thermocline displacements matched the model ones at $N_m$ time slices, where $N_m$ is the number of months in the assimilation. This then implies that the problem will in general be overdetermined for assimilation periods of more than 1 month.
and an exact fit solution for these initial conditions would not be possible. Of course, if the data density is poor so that \( (\sigma_k)^{-1} \) is close to zero in many areas, the residual cost function may be very small since the problem will then be overdetermined only to a very small degree (i.e., a solution for the initial conditions with vanishing cost function "almost" exists). In view of the fact that the cost function does not vanish, it is important to check that the convergence seen is real and not simply due to slowing of convergence in regions of parameter space that have small gradients. This was checked by commencing the iteration procedure from different starting points for the initial conditions. It was found that solutions obtained for thermocline displacements differed only marginally once convergence was apparent from a monitoring of the cost function.

We turn now to the results of the assimilation for particular cases. Two 12-month assimilations were performed for the years 1987 and 1988, which were, respectively, warm and cold ENSO events. Results to be discussed were found to be relatively insensitive to the period of assimilation used.

Depicted in Figs. 5a–c are Hovmöller diagrams of 1987 equatorial anomalies of \( T_{400} \) for the assimilation, the original Smith analysis and the ocean model forced with the observed FSU winds (note that a spinup period of 34 years was used prior to the period under study for this case). Figures 5d–f are the corresponding quantities for 1988.

Focusing first on the Smith analysis and the wind-forced ocean-model results (Figs. 5a and 5c for 1987 and 5d and 5f for 1988), we see broad agreement between the two results as one would expect from the EOF analysis of the previous section. Some difference of note are apparent, however: The 1987 warm event tends to peak earlier in the Smith analysis and is generally stronger in the wind-forced model particularly toward the end of 1987 where the Smith analysis calls for an earlier termination of the event. Also absent in the Smith analysis are the sizable vacillations in the date line region present in the wind-forced case. Large negative peaks are apparent in the central Pacific during mid and late 1988 in the Smith analysis but are absent in the wind-forced results.

The assimilation runs (Figs. 5b and 5e) show results that tend to be intermediate between the two runs above but probably are closer to the wind-forced ocean run. The subsurface information has, however, clearly influenced the assimilation: There is a more pronounced decline in eastern Pacific warm conditions in the assimilation toward the end of 1987 than in the wind-forced model; the vacillations around the date line in the wind-forced model have been strongly reduced in the assimilation consistent with the Smith analysis and finally the central Pacific cold peak at the end of 1988 present in the Smith analysis seems to have influenced the assimilation. Finally the assimilation seems to have generated features that are absent in the two other depictions: The strong far eastern Pacific warming during March–June 1987 is weakly present in the wind-forced case but absent in the Smith analysis. This assimilation generated feature represents a "dynamical" compromise between the subsurface observations and the observed surface forcing. We hope to investigate this interesting effect in greater detail in a separate study.

The assimilation systems behavior for a particular month was also examined. Displayed in Fig. 6 (in the same order as Fig. 5) are the x-y plots of \( T_{400} \) for April 1987. The assimilation results are taken from the 12-month 1987 run. Note that the Smith analysis has been truncated at the sixth Rossby mode to aid in comparison (this is also the form in which the data is assimilated). As for the equatorial section, it is apparent that the basic zonal structure of the Smith analysis (Fig. 6a) is qualitatively similar to the wind-forced model output (Fig. 6c). As previously, however, there are notable differences with perhaps the greatest being the much broader meridional structure in the eastern Pacific in the wind-forced results. The assimilation run (Fig. 6b) again shows intermediate behavior between the Smith analysis and the wind-forced results with the meridional structure in the central eastern Pacific being the most obvious example of this. Also evident is that the western Pacific negative anomalies in the assimilation tend to be a compromise between the marked asymmetry evident in the wind-forced case and the near symmetry observed in the Smith analysis.

In summary, the data assimilation system is able to successfully capture many features of the Smith (1995) analysis and also features of thermocline displacements forced into the ocean model by the FSU winds. It is able to do so in a way that is dynamically consistent with the simple ocean model used. There appear to be no obvious drawbacks in imposing this dynamical constraint since the assimilation products behave well over the entire space–time interval used in the variational technique.
Fig. 5. Hovmöller diagrams of equatorial $T_{200}$ from (a) the 1987 Smith analysis, (b) the 1987 assimilation run, (c) the 1987 wind-forced ocean model run, (d) the 1988 Smith analysis, (e) the 1988 assimilation run, and the (f) 1988 wind-forced ocean model run. Note that the assimilation and model values have been converted to temperature units by use of the proxy relation. The contour interval is 0.25°C in all cases.
5. Coupled model hindcast results

The ability of the assimilation system to produce initial conditions suitable for coupled model forecasts was also examined. The period chosen to test this capability was from January 1982 through to October 1991 with initial conditions for the coupled model being prepared every 3 months within this time interval. Hindcasts were made by running the coupled model for 2 years from these prepared initial conditions. A sample of 40 hindcasts were thus prepared with the final hindcast finishing in September 1993.

Initial conditions for the hindcasts were prepared in two different ways. First, the "traditional" method was used, which involved forcing the ocean model with the observed FSU winds from January 1979 to October 1991 and extracting the appropriate model variables at the appropriate times mentioned above. Second, the assimilation system was run for 12 months prior to the required initial condition dates.

The hindcast skill of the model was estimated by two methods: By calculating the correlation of the model NINO3 index (this is the SST anomaly for the region 5°N–5°S, 90°–150°W) with the observed index (derived from Reynolds 1988 analyses) and also by calculating the rms difference between the model and observed NINO3 index. This was done for hindcast lags up to 24 months. Figures 7a and 7b depict these estimated skills for the two methods of initialization together with the skill of persistence.

The estimated skill for the assimilation case is clearly greater than that of the "traditional" wind-forced initialization. This increase occurs at all hindcast lags and for both the correlation and rms error measures of skill. Judged by the numerical weather prediction yardstick of useful skill occurring when the anomaly correlation exceeds 0.6, there appears to be useful skill for this method of initialization out to around 16 months. Also this skill exceeds 0.7 (and hence explains more than 50% of the variance) for the first 13 months of the hindcasts. Notable is that the greatest increase in skill occurs between 8 and 16 months. Finally notice that there are now strong indications of skill at 24 months. The rms error of hindcasts shows a similar improvement. Note that the small decline in error over the first year for the assimilation hindcasts is due to the fact that the coupled model is marginally stable and so the variance of hindcasts tends to drop with time (see Kleeman 1991 for further discussion on this point).

To get a more concrete understanding of how improvements in hindcasts are possible as a result of assimilation, a specific case is examined. Depicted in Fig.
Fig. 7. (a) Anomaly correlation curves and (b) rms error curves of NINO3 as a function of hindcast lag for a model initialized with winds alone, a model initialized by using a 12-month data assimilation, and persistence. Note that the persistence curve anomaly correlation curve has been truncated at 12 months (values after this time are weakly negative).

8 are the two hindcasts for the initial conditions of January 1988. Also shown is the observed variation in NINO3 for the 2-yr period of the hindcasts. This period was characterized by an initial rapid cooling from the 1987 warm event into the 1988 cold event. As can be seen, the traditional method of initialization has signifi-

Fig. 8. Hindcasts commencing January 1988 of NINO3 using wind-forced and data assimilation initialization together with the verifying observations.
ificant problems in picking this cooling and even in late 1988 when a cool event was well established, the hindcast was calling for warm conditions. On the other hand, the assimilation run produces a significantly better hindcast although it still does not show as sharp a decline as observed. As was noted in the previous section (cf. Fig. 5), the Smith analysis showed significantly smaller positive thermocline perturbations in the latter part of 1987 compared to the wind-forced model (perturbations at the beginning of the year were comparable). These perturbations also became negative sooner in the former case. It appears that this difference in behavior is responsible for the more rapid (and hence more realistic) cooling in the assimilation hindcast.

One of the interesting features of the assimilation technique used here is the strong spatial variation in the coefficient $\sigma_0$ used in the cost function. This simply reflects the fact that subsurface data is collected in a very spatially inhomogenous fashion via ships of opportunity and fixed moorings. To test the usefulness of this information to the assimilation, the hindcast experiments above were repeated with $\sigma_0$ set to a uniform value. Compared to the results above there was a drop in skill at all lags of between 0.1 and 0.3 in the correlation coefficient with longer lags being more degraded than shorter ones. Such a result indicates the value of observational density information and suggests that a similar measure for the wind data may prove useful in generalizations of the present system.

Finally, two experiments were conducted to test the sensitivity of the above results to the assimilation period used. The experiments above were repeated with assimilation time intervals of 24 and 6 months. In both cases there was again an improvement in skill over the wind-forced case at all lags and for both anomaly correlation and rms error. The 24-month case showed skill intermediate between the 12-month case and the wind-forced case. This is probably due to the fact that the observed winds are playing a stronger role in constraining the assimilation in the former case (a 4-yr assimilation is not very different to a wind-forced run). The 6-month case showed less skill than the 12-month case for lags less than 12 months but somewhat greater skill for lags between 12 and 24 months.

6. Summary and discussion

Anomalies of the ocean temperature averaged over the upper 400 m were found to be a good proxy for shallow-water ocean model thermocline displacements. Use of this proxy relation allowed the successful assimilation of subsurface thermal data into a simple ocean model using an adjoint variational technique. Such a technique amounts to an optimal dynamical interpolation of the observational data since a trajectory through the data that is consistent with the dynamical ocean model and optimal with respect to the observations is found. The existence of a measure of analysis error allowed the large spatial variation in data density to be taken into account in the cost function.

Use of this assimilation technique to initialize an intermediate coupled model led to a marked increase in the estimated hindcast skill of the model for the period 1982–91. Useful hindcast skill based on exceeding a level of 0.6 for the NINO3 anomaly correlation was found for hindcast lag times up to and including 16 months. Previous estimates of skill based on the 1972–86 period had indicated useful skill for around 8 months (see Kleeman 1993). Interestingly, skill was indicated also at 24 months suggesting that useful forecasts in excess of 2 years may be possible with future more sophisticated coupled models. Of course some caution needs to be exercised in this regard since the skill estimates are based on a fairly short and possibly atypical time period (three quite strong warm events are included).

The conclusion of Kleeman (1993) that thermocline perturbations are crucial to coupled model skill has been strongly confirmed by the results presented here since the ocean model SST here is controlled by such perturbations alone and the only ocean observations assimilated amount to a simple proxy for this variable.

Since such a variable is demonstrably important it is probably worth improving on the fairly crude proxy used here to extract it from the data. Physically, such a variable represents a good approximation, a linear combination of the amplitudes of the significant oceanic baroclinic modes (these are excited primarily by surface momentum forcing). Since in general the subsurface thermal observations have quite good vertical resolution, decomposition into baroclinic modes would be straightforward. These amplitudes could then be assimilated into an obvious generalization of the ocean model used here. The reasonably large spectral gap in shallow-water speeds between the first few baroclinic modes in the tropical Pacific and the mild hindcast skill dependency on shallow water speed seen in Kleeman (1993), suggest that some improvement in model skill will ensue. In addition, the results of Cane (1984) indicate that perhaps only two vertical modes may be required. Such a model is currently under investigation.

Another generalization of the system described here would be to allow the wind stress data to become part of the cost function $J$ and hence subject to variation like the initial conditions. Such a system has been described by Thacker and Long (1988). For the successful implementation of such a system, estimates of wind stress anomaly error would probably be required since the density of surface data for such analyses varies strongly in space and time. In addition some indication of the crudeness of the thermocline proxy relation would be required to appropriately weight the two parts of the cost function. These issues are also currently under investigation.

It was interesting to note the moderate sensitivity of the results to the assimilation period used: The e-fold-
ing decay time within the linear model is around 2.5 years. Thus, as the assimilation period approaches this timescale, variations in the initial conditions will influence the final state of the model less and less relative to the state of the model early in the assimilation period. Therefore, most of the work of changing these initial conditions will go into changing the cost function contribution generated by the ocean early in the assimilation period and not that generated near the final state, which acts as the initial conditions for coupled forecasts. On this basis, we would expect the skill of the model to approach the wind-forced case as the assimilation period was lengthened, as indeed occurs. On the other hand as the assimilation period is shortened, the usefulness of the space–time interpolation might be expected to be reduced since information that was assimilated in one part of the basin cannot influence the final state at another part of the basin since it has insufficient time to propagate there during the assimilation period.

Since basin transit times of the Rossby modes in the ocean are between 8 and 19 months, we would expect that an assimilation period of this order would be required to fully utilize the dynamical space–time interpolation implicit in the variational technique used here. The reduced skill of the model in the first 12 months of hindcasts with a 6-month assimilation period seems to confirm this surmise although the mild increase in skill in the second year suggests that perhaps the situation is somewhat more complex than alluded to here. This issue is clearly of some importance and a more detailed study is planned.

Acknowledgments. One of the authors (RK) wishes to thank Benno Blumenthal of Lamont Doherty Geological Observatory for stimulating and useful discussions related to the work here.

REFERENCES


