Application of a reduced order Kalman filter to initialize a coupled atmosphere-ocean model: Impact on the prediction of El Niño.

Joaquim Ballabrera-Poy
Universities Space Research Association
NASA / GSFC

Antonio J. Busalacchi
Laboratory for Hydrospheric Processes
NASA / GSFC

Ragu Murtugudde
ESSIC, University of Maryland
NASA / GSFC

February 2000

Corresponding author:
Joaquim Ballabrera-Poy
NASA/GSFC
Mailcode 970
Greenbelt, MD 20771

Telf : (301) 614 5685
Fax: (301) 614 5666
E-mail: joaquim@neptune.gsfc.nasa.gov

Abstract

A reduced order Kalman Filter, based on a simplification of the Singular Evolutive Extended Kalman (SEEK) filter equations (Pham et al. 1998), is used to assimilate observed fields of the surface wind stress, sea surface temperature and sea level into the nonlinear coupled ocean-atmosphere model of Zebiak and Cane (1987). The SEEK filter projects the Kalman Filter equations (Kalman and Bucy 1960) onto a subspace defined by the eigenvalue decomposition of the error forecast matrix, allowing its application to high dimensional systems.

The Zebiak and Cane model couples a linear reduced gravity ocean model (Cane and Patton 1984) with a single vertical mode atmospheric model of Zebiak (1986). The compatibility between the simplified physics of the model and each observed variable is studied separately and together. The results show the ability of the model to represent the simultaneous value of the wind stress, SST and sea level, when the fields are limited to the latitude band 10°S - 10°N.

In this first application of the Kalman Filter to a coupled ocean-atmosphere prediction model, the sea level fields are assimilated in terms of the Kelvin and Rossby modes of the thermocline depth anomaly. An estimation of the error of these modes is derived from the projection of an estimation of the sea level error over such modes. This method gives a value of 12 for the error of the Kelvin amplitude, and 6 m of error for the Rossby component of the thermocline depth.

The ability of the method to reconstruct the state of the equatorial Pacific and predict its time evolution is demonstrated. The method is shown to be quite robust for predictions

up to six months, and able to predict the onset of the 1997 warm event fifteen months before its occurrence.

1. Introduction

The prediction of the El Niño - Southern Oscillation (ENSO) and its relation with global climate anomalies is one of the major research efforts in short-term climate forecasting. As the large-scale climatic variability of the ENSO can be described in terms of low frequency modes which are successfully modeled by simple linear models, the studies of ENSO have been supported by a large hierarchy of models, ranging from complex general circulation models (e.g. Philander et al. 1992; Latif et al. 1993; Kirtman et al. 1996) to one-layer, reduced-gravity models (eg: Philander et al. 1984; Gill 1985; Hirst 1986; Cane and Zebiak 1987; Battisti 1988; Battisti and Hirst 1989).

The success of the simplified models is due to a combination of two facts. First, the fact that the inertia ("memory") of the tropical, coupled ocean-atmosphere system lies in the upper ocean (for more details see Neelin et al. 1994), which has much longer time scales than the atmospheric component. Second, the fact that the equatorial ocean is characterized by a stable density structure consisting of a warm upper layer and a cool deep layer separated by a sharp, near-surface pycnocline. This situation can be well described by the reduced gravity models which assume that the ocean consists of a thin surface layer of density ρ overlaying an infinitely-deep, motionless lower layer of density $\rho + \Delta\rho$.

The Zebiak and Cane model (Zebiak and Cane 1987), hereafter named ZC, is a nonlinear model for the anomalies of the system about a basic state derived from monthly mean climatology of sea surface temperature (SST), surface winds and thermocline depths in the

tropical Pacific. Although this model was originally developed for process studies (Cane et al. 1986) it gave the first successful forecast of an ENSO event and it has since been used routinely in real-time ENSO forecasts (Latif et al. 1998).

As the memory of the coupled system lies in the ocean, the initialization of the several ENSO prediction systems has been focused mainly on the ocean component, in accord with the hypothesis that the coupled state evolves (in a deterministic way) from an initial state determined by the thermal state of the upper-ocean (Cane et al. 1986; Xue et al. 1994; Ji et al. 1995; Fischer et al. 1995).

In the case of the ZC model, Cane et al. (1986) use an initialization procedure in which the ocean component is forced by observed wind stresses starting in January 1964 and ending at the time when the forecast begins. Then, the computed SST anomalies are used to run the atmospheric component. The final states from the uncoupled component models are used as the initial conditions for a forecast with the model run in a purely coupled manner. Using this method, the initial conditions of the SST and the winds are assumed to be in accord with the thermocline depth but not necessarily with their current observed values (Cane et al. 1986).

An improved procedure was developed by Chen et al. (1995) allowing the complete interaction between both ocean and atmospheric components during the initialization process. In their procedure, the model wind stress is relaxed towards the observed values by introducing a nudging term in the wind stress equations of the atmospheric model during the initialization. This nudging approach reduces both the error of the initial condition, and some of the small-scale, high-frequency modes appearing when the ocean component is forced by ob-

served wind stress (Chen et al. 1995). As a result, the skill of the ENSO forecasts is greatly improved. This methodology has been used to perform operational ENSO forecasts. This method, however, failed to predict the 1997-98 ENSO event, showing the need of further improvements in the physics of the model (Kleeman 1993; Dewitte and Perigaud 1996), the assimilation method (Kleeman et al 1995), or the quality of the data as it only relied on wind observations (Chen et al. 1999).

The positive impact of assimilating subsurface thermal data in a complex GCM model was shown by Rosati et al. (1997). Kleeman et al. (1995) improved the initialization of the intermediate coupled model of Kleeman (1993) by using the adjoint approach to assimilate subsurface thermal data.

In the case of the ZC model, Chen et al. (1998,1999) study the influence of either new wind fields or tide-gauge observations to improve the thermocline structure at the initial time. Both approaches have been able to predict retrospectively the 1997-98 El Niño, with no need to change the physics of the coupled model.

The impact of SSH observations has been also studied by Xue et al. (2000, manuscript submitted to *J. Climate*) using a linear Markov model best fit to the ZC (Blumenthal 1991). Their results indicate that the prediction skill of the Markov model is better when the sea level information is used during the training period.

The work presented here follows a similar vein, but with emphasis on an advanced assimilation methodology, multivariate data sets, and a coupled initialization approach. This is accomplished by applying a reduced order Kalman filter to assimilate gridded fields of SST (Reynolds and Smith 1994), the Florida State University (FSU) wind stress analysis

(Stricherz et al. 1992) and the sea surface height (SSH) anomalies (Busalacchi et al. 1994) into the ZC model. In a reduced gravity model, the upper-layer depth D can be related to the SSH (Gill 1982) by

$$D(x,y,t) = \frac{\rho}{\Delta\rho}\mathrm{SSH}(x,y,t) \tag{1}$$

Then, the thermocline depth information is directly assimilated into the ZC model in terms of the Kelvin and Rossby modes. This method differs from the indirect assimilation used by Chen et al. (1998) who nudge the Kelvin and Rossby modes of the coupled model to the modes derived from a previous assimilation of the sea level in the stand-alone Cane and Patton (1984) ocean model.

This paper represents the first attempt of a direct assimilation of multiple fields in the fully nonlinear ZC model. Because of the simplified physics of the model, the model shows discrepancies with each one of the observed fields (Perigaud and Dewitte 1996). These discrepancies increase when these fields are considered simultaneously. Thus, a preliminary step must be the study of the compatibility between the observations and the model.

The data used in this study are presented in section 2. Section 3 presents a brief review of the ZC model and the multivariate EOF basis used to reduce the degrees of freedom of the assimilation method. The reduced order Kalman filter is presented in section 4. The compatibility between the data and the model is presented is section 5. The method used to determine the confidence of each field is presented in section 6. The assimilation experiments, the measure of the forecast skill and several sensitivity studies are described in section 7.

Finally, a summary and concluding discussion are given in section 8.

2. The data

a. FSU winds

The wind fields used here come from the monthly pseudo-stress analysis from merchant ship observations provided by the Florida State University (Stricherz et al. 1992). Hereafter, these data will be referred to as FSU winds. Research quality pseudo-stresses are available until 1997. Pseudo-stresses for 1998 are obtained from the quick-look product. Following Stricherz et al. (1992), a drag coefficient $C_D = 1.6 \cdot 10^{-3}$ is used to compute the wind stresses.

Monthly anomalies are computed from the four years running mean, consistent with the methodology employed by Cane et al (1986). That is, at each month, the reference state is obtained by the mean between the current month and the respective month for the three previous years. This methodology is used to remove the long term trend of the data (Zebiak 1989). Finally, field anomalies are filtered with a 1-2-1 filter in latitude, longitude and time.

b. Reynolds SST

Maps of the SST from November 1981 to December 1998 are derived from the optimum interpolated SST analysis of Reynolds and Smith (1994). Anomaly fields are constructed from the monthly mean field of the Reynolds fields during November 1961 and December 1996. A 1-2-1 filter in latitude, longitude and time is applied to the anomaly fields.

c. TOPEX/Poseidon altimetry

Monthly anomalies of the thermocline depth are derived from the TOPEX/POSEIDON (T/P) sea level fields (Busalacchi et al. 1994). The long term surface topography anomaly

maps are used to construct monthly mean maps from January 1993 to December 1996. Then, equation (1) is used to compute the thermocline depth.

Because the ZC model projects the dynamic equations onto the equatorial modes, we have decided to project the thermocline depth onto the equatorial Kelvin and several Rossby modes using the method described in Boulanger and Menkes (1995). Once the coefficients of the several Rossby modes are computed, the Rossby component of the thermocline depth is constructed by a combination of the Rossby modes. This methodology has the advantage of decomposing the sea level observations into prognostic variables of the model, diminishing the possible impact of the incompatibilities between the observations and the simplified physics of the model. In section 5, the compatibility between the physics of the model and the different Rossby modes will be examined.

In this first application of a reduced order Kalman Filter, the analysis period is limited to January 1993 - December 1998. This is the time period for which comprehensive and coincident fields of T/P sea level, wind stress and SST observations are available for the model domain.

The temporal evolution of the meridional average of the zonal wind stress, the Kelvin amplitude, the first Rossby mode, and the SST anomalies are given in Fig. 1. A positive value of the wind stress shows the presence of westerly anomalies (weakening the Walker circulation in the equatorial Pacific). A positive value of the thermocline depth indicates a downwelling, associated with an increase of SST in the east.

A cursory examination of plots in Fig. 1 reveals the basic characteristics of the equatorial Pacific dynamics, as the large-scale nature of the warm/cold oscillations. Also, the high

correlation between the SST and the thermocline depth justifies the parameterization of vertical mixing in the SST equation of the ZC model. Finally, the correlation between western wind stress anomalies with the eastern sea surface temperature anomalies suggests how local SST changes are related to remote forcing.

During the period Jan 93 - May 95, westerlies are present in the western Pacific extending periodically towards the central Pacific where the anomalous wind is easterly. These easterlies in the eastern side of the basin inhibit the propagation of the downwelling Kelvin waves generated by the central westerlies, limiting the warming of the eastern waters. After May 95, easterlies are present over the entire equatorial Pacific, indicating stronger trade winds. Under these conditions, an upwelling Kelvin wave excited in the central Pacific will be amplified during its eastward travel, causing a cooling of the equatorial SST. The maximum strength of the central trades is observed at the beginning of 1996. Then, the wind anomalies decrease and finally, reverse to westerlies. The amplitude of the westerlies increase rapidly and propagate to the eastern region, amplifying the downwelling Kelvin waves during all of 1997, thereby warming the central and eastern Pacific water. As the downwelling Kelvin waves propagate into the eastern equatorial Pacific, an upwelling wave is present in the central Pacific since October 1997 preceding the termination of the warm event on December 1997. The origin of this upwelling Kelvin wave (generated by the reflection of Rossby waves at the western boundary and/or excited by the winds) is not provided by the plots in Fig. 1.

3. Data error estimates

As the Kalman filter is a weighted combination of a first guess and observations, a measure of the confidence of the observations has to be specified *a priori*. A simple estimation of the observation errors is given here.

Following Shriver and O'Brien (1995) the mean wind speed over the tropical Pacific is approximately 7 m s^{-1}, or 49 m^2 s^{-2} in pseudostress units. The error in anemometer measurements of wind speed from moored buoys is approximately 10%. The error associated with estimates of wind speed and direction is larger than that. Therefore, an optimistic measure of the error of the wind speed can be given as 10% of the mean wind speed, i.e., 4.9 m^2 s^{-1} $\approx$ 0.094 dyn cm^{-2}.

McPhaden et al. (1998) have verified the accuracy of the optimum interpolation SST analysis of Reynolds and Smith (1994) using independent data. They compute the root mean square (rms) between the monthly analyzed fields and in situ TOGA-TAO moorings at the equator at three different locations (110°W, 140°W, and 165°E). The rms was computed from January 1982 to January 1993. The mean rms differences at each location are 0.38°C, 0.39°C, and 0.24°C respectively. The SST anomaly error used here is obtained as the mean of these three values, i.e. 0.34 °C.

Cross correlation between T/P altimeter data and dynamical topography fields from TOGA/TAO data show RMS differences of 4 cm in the equatorial region, but higher values off the equator (Busalacchi et al. 1994). Smaller error measures have been reported on monthly and longer time-scales, but a value of 4 cm will be used here, corresponding to an error of 7 m for the thermocline depth.

As the Kelvin and the Rossby components of the thermocline depth are assimilated

separately, a distinct error value has to be given to the assimilation algorithm. Here, we compute an equipartition of the thermocline depth error on two components by assuming that the confidence of the decomposition algorithm is similar for each wave. This is not really true because of the fact that high Rossby modes have their maximum off the equator, where the reduced gravity approximation is less accurate. Thus, our hypothesis is valid only as the reduced gravity approximation is valid. Furthermore, assuming that the errors of the Kelvin and Rossby modes have zero mean and are uncorrelated, the error variance can be described as

$$< \varepsilon_h^2 >=< a_k^2 > \psi_k^2(y) + \sum_n < a_n^2 > \psi_n^2(y) \tag{2}$$

because of the following relations

$$< a_k > = < a_n > = 0, \qquad n = 1, \ldots$$

$$< a_k a_n > = < a_n a_m > = 0 \qquad n \neq m$$

where $\psi_k(y)$ is the meridional structure of the Kelvin wave, $\psi_n(y)$ the meridional structure of the n-th Rossby wave, and a_k and a_n is the amplitude of these modes.

The algorithm of Boulanger and Menkes (1996) is used to solve equation (2). The left hand side term is now a constant for all the latitudes. A minimum of 8 Rossby modes has to be used for a reasonable reconstruction of such a constant function over the latitude band of [10°S-10°N]. The error associated with the Kelvin coefficient is obtained directly by this procedure. The error of the Rossby component is obtained by the longitudinal average of the second term on the right hand side. Figure 2 shows the errors of the Kelvin amplitude coefficient (solid line), and the error of the Rossby field (dashed line). The error

estimates converge as the number of Rossby modes increases. From the asymptotic limits, $\varepsilon_k = (< a_k^2 >)^{1/2} = 12$, and $\varepsilon_r = 6$ m. As the Kelvin and Rossby waves are prognostic variables of the ZC model (Zebiak and Cane 1987), the sensitivity to these values will be examined later.

4. The model

The standard configuration of the ZC model is used here. The ocean domain is a rectangular basin limited by 30°S-30°N and 124°E-80°W. Details of the model equations can be found in Zebiak and Cane (1987) and Battisti (1988), so only a brief description of the model is given here.

The dynamical model is a single-layer, reduced-gravity, anomaly model on the equatorial β-plane focusing only on the seasonal and interannual time scales. The ocean model has been simplified to determine only a subclass of all the possible motions: the low-frequency, long zonal-scale equatorial Kelvin and Rossby waves forced by an anomalous surface wind stress (Cane and Patton 1984). As the Ekman component of the surface velocity is not negligible the surface layer is divided into two sub-layers. The upper sub-layer has constant thickness and is directly acted upon by the wind. The SST is calculated separately through a nonlinear equation, including three-dimensional advection, but it does not have a direct feedback to the ocean dynamics. The atmospheric model is based on the steady-state, linear model of a thermally forced tropical atmosphere (limited to a single vertical mode) of Gill (1980). The forcing of the atmosphere is given by a local heating depending only on the SST and a second term introduced to simulate the fact that the heating due to the low-level

moisture convergence anomaly is not dependent on the SST alone, but also on the wind convergence (Zebiak 1986).

Due to the special role of SST, the different processes in the thermodynamic equation have been carefully parameterized. The heat flux anomaly is assumed to be proportional to the local SST anomaly, acting always to damp the temperature towards its climatological mean state (Zebiak and Cane 1987). The temperature is affected by vertical advection only in the presence of upwelling. The entrained anomalous temperature is defined by an analytical equation that is a function of the thermocline motions, adjusted to fit the observations (Cane and Zebiak 1985).

The time-step of the ocean model is 10 days. As the atmospheric model gives the steady-state solution associated with each SST field, the atmosphere may adjust too rapidly to the ocean changes. To avoid this, the ZC model allows a time-dependency of the moisture term of the atmospheric heating. The change of heating is computed every time step. Then, the assumed background convergence is the total convergence at the previous time-step, rather than the mean convergence as in the steady-state model. However, this procedure could allow the development of small scale anomalies that may persist and become completely unrelated to subsequent SST fields (Zebiak and Cane 1987). It has been shown that the method used to filter out these unrealistic anomalies can modify the behavior of the model. Zebiak and Cane (1986) recalculated the atmospheric anomalies when the temperature anomalies (given by Niño-3 SST) were small. This procedure favored a period of the warm events close to 4 years.

In spite of the success of the model to simulate and predict the evolution of the Niño-3

SST, the descriptive skill of the model can be seriously affected by its simplicity: the lack of any variability generated at the mid-latitudes, the absence of the 30-60 day waves (Cane et al. 1986), the coarse resolution of the model distorting the simulation of coastal upwelling processes in the eastern Pacific (Cane and Zebiak 1985), and the simplified relation between the atmospheric heating and SST (Zebiak 1986; Dewitte and Perigaud 1996). Therefore, dynamical incompatibilities between the model and the observations can be important, particularly in the off equatorial region. This has a direct effect on assimilation algorithms. For example, the nudging constant of Chen et al. (1995) has a meridional dependency to take account of the fact that the winds generated by the model are less accurate away from the equator. The nudging term gives more weight to the observations in the off-equatorial regions, allowing the observations to modify the physics of the atmospheric model where the model is less accurate.

a. The multivariate EOFs

The statistical properties of the coupled model are studied in terms of multivariate EOFs of a set of states obtained from a control run of the model. The coupled model is integrated from rest, with an initial wind stress anomaly of 0.5 dyn/cm^2 is imposed during the first four months in the region 163°E - 163°W and 5°S - 5°N. After this period of time, the model is integrated in a coupled manner for 200 years.

To compute a set of multivariate Empirical Orthogonal Functions (EOF), a multivariate state vector is defined to express the state of the system:

$$\mathbf{x} = (\mathbf{T}, \tau_x, \tau_y, \mathbf{u}, \mathbf{v}, \mathbf{k}, \mathbf{h}, \mathbf{u}_a, \mathbf{v}_a, \mathbf{y}^o, \mathbf{q}), \tag{3}$$

where $\mathbf{T}$ represents the sea surface temperature; τ_x and τ_y are the zonal and meridional components of the wind stress; $\mathbf{u}_a$ and $\mathbf{v}_a$ are the zonal and meridional components of surface wind; $\mathbf{k}$ is the amplitude of the equatorial Kelvin wave, $\mathbf{h}$ is the Rossby component of the upper layer depth, $\mathbf{u}$ and $\mathbf{v}$ are the zonal and meridional components of the depth integrated current anomalies between the thermocline and surface, $\mathbf{y}^o$ is the surface wind divergence and $\mathbf{q}$ is the atmospheric heating. All the variables are normalized by the mean standard deviation of each one during the 200 years of simulation. The number of components of the state vector (3) is 35457.

A set of 107 multivariate EOFs is constructed from the monthly mean states during the last 115 years. Figure 3 shows the eigenvalue spectrum of the covariance matrix. The 95% confidence levels, indicated by the dashed line have been calculated from a Monte Carlo simulation of 100 eigenvalue decomposition of individual covariance matrices containing randomly generated Gaussian variables of zero mean and unit variance (Overland and Preisendorfer 1982). Only the first 12 multivariate EOFs pass the confidence level. That is, this rule states that we cannot distinguish the multivariate EOFs of order higher than 12 from ones generated by a spatially and temporally uncorrelated random process. Such a small number is due to the large amount of variability explained by the first few EOFs: The first multivariate EOF explains 42% of the variability of the model during the 115 years considered. The second and third EOFs explain more than 15% and 9% respectively. More than 90% of the variability of the system is explained by the first 14 modes.

Figures 4-5 show the SST, wind stress, Kelvin amplitude and Rossby component associated with the first two multivariate EOFs. The first mode (when considered with positive

sign) corresponds to a warming of the eastern equatorial Pacific. The associated wind stress corresponds to a weakening of the trades, i.e. a weakening of the Walker circulation, while the Hadley circulation is increased in the eastern Pacific. Both effects tend to increase the wind convergence east of the dateline. This convergence is maximum near 110°W. In this region, the Kelvin amplitude is maximum, corresponding to an increase of the thermocline depth. In this first multivariate mode, both Kelvin and Rossby components of the thermocline act together to increase the depth of the thermocline in the eastern Pacific and to decrease the depth of the thermocline in the western Pacific. The Kelvin wave associated with this mode represents a downwelling Kelvin wave reaching the eastern boundary. The Rossby wave corresponds to an upwelling wave off the western boundary. This pattern of circulation is characteristic of the warm phase of ENSO, the El Niño phenomenon (Battisti 1988).

The SST of the second multivariate EOF also corresponds to a warming of the surface waters east of the dateline. The main differences with the SST pattern of the first EOF are the lower amplitude of the warming, and the weak zonal gradients near the equator, compared with the strong meridional gradients. This increase of the SST anomalies is not produced by an east/west tilting of the thermocline, but rather a meridional structure corresponding to an increase of the thermocline depth near the equator and a decrease at higher latitudes. The zonal structure of the field is weak and concentrated mainly east of the dateline.

5. Compatibility between data and model

The assimilation experiments presented below use multivariate EOFs of the model to

extrapolate the information from the available observations $\mathbf{y}^o$ to each component of the state vector $\mathbf{x}$. As the observations can be incompatible with the simplified dynamics of the model, an *a priori* study of the compatibility between the observations and the dynamics of the model is necessary. In this section, this is done in terms of the multivariate EOFs of the system.

The observations do not give information about the full state vector $\mathbf{x}$ but only about a small number of its components. Therefore, the classical principal component analysis cannot be carried out. On the other hand, as the observations can be physically incompatible with the model physics, a residual may be present when observations are expressed in terms of the EOFs of the model.

Let $\mathbf{x}^t$ be the true state of the ocean, $\overline{\mathbf{x}}$ be the mean state of the model and $\mathbf{S}$ be the matrix defined by the EOFs of the model. The true state can always be expressed as

$$\mathbf{x}^t = \overline{\mathbf{x}} + \mathbf{S}\mu + \mathbf{e}^r, \tag{4}$$

where μ is a vector with the amplitude of each element of the subspace defined by $\mathbf{S}$. $\mathbf{e}^r$ is the residual error. The only knowledge of the true state comes from the observations

$$\mathbf{y}^o = \mathbf{H}\mathbf{x}^t + \mathbf{e}^o, \tag{5}$$

where a linear relation is assumed between the observed value $\mathbf{y}^o$ and the components of the state vector. $\mathbf{e}^o$ is the observation error.The innovation vector around the mean state is defined by

$$\mathbf{d} = \mathbf{y}^o - \mathbf{H}\overline{\mathbf{x}}. \tag{6}$$

Then, the innovation vector can be expressed as

$$\mathbf{d} = \mathbf{H}\mathbf{S}\mu + \mathbf{H}\mathbf{e}^*, \tag{7}$$

where $\mathbf{H}\mathbf{e}^* = \mathbf{H}\mathbf{e}^r + \mathbf{e}^o$.

Equation (7) can be considered as a parametric definition of the error not accounted by the subspace defined by $\mathbf{S}$ at the observation locations, $\mathbf{H}\mathbf{e}^*$, in terms of the coefficients μ. If we require that the two terms of equation (7) have no common information, i.e.,

$$(\mathbf{H}\mathbf{S})^\top \mathbf{H}\mathbf{e}^* = 0, \tag{8}$$

then, we obtain

$$(\mathbf{H}\mathbf{S})^\top \mathbf{d} = \mathbf{A}\mu \tag{9}$$

$$\mathbf{H}\mathbf{e}^* = \left[\mathbf{I} - \mathbf{H}\mathbf{S}\mathbf{A}^{-1}(\mathbf{H}\mathbf{S})^\top\right]\mathbf{d} \tag{10}$$

where $\mathbf{A} = (\mathbf{H}\mathbf{S})^\top \mathbf{H}\mathbf{S}$ is a $r \times r$ symmetric matrix.

Equation (9) is formally equivalent to the equation used by Smith et al. (1996) for the least-square fitting of EOFs to observations. This is due to the fact that condition (8) minimizes the projection of the error onto the space defined by the EOFs. Therefore, this method has the same disadvantages as the method of Smith et al. 1996), i.e., the projection fails when the number of observations is not large enough (Kaplan et al. 1998).

Kaplan et al. (1998) discuss an alternative EOF fitting by the minimization of

$$\mathcal{J}[\mu] = (\mathbf{d} - \mathbf{H}\mathbf{S}\mu)^\top \mathbf{R}^{-1}(\mathbf{d} - \mathbf{H}\mathbf{S}\mu), \tag{11}$$

where $\mathbf{d}$ is given by equation (6). The set of principal components minimizing (11) is the

solution of

$$(\mathbf{H}\,\mathbf{S})^{\top}\,\mathbf{R}^{-1}\,\mathbf{d} = \mathbf{A}'\mu, \tag{12}$$

where $\mathbf{A}' = (\mathbf{H}\,\mathbf{S})^{\top}\,\mathbf{R}^{-1}\,\mathbf{H}\,\mathbf{S}$ is also a $r \times r$ symmetric matrix. In this case, the innovation vector is split in two orthogonal terms with the inner product $< \mathbf{a}, \mathbf{b} >= \mathbf{a}^{\top}\,\mathbf{R}^{-1}\,\mathbf{b}$. Comparison equations (9) and (12) shows that both methods only differ on the metric of the inner product. Finally, note that since the components of the state vector, equation (3), have been normalized by the standard deviation of the model solutions, the estimates of the obervational error variances (section 3) also have been normalized.

The subspace $\mathbf{S}$ is representative of observations if $\mathbf{H}\,\mathbf{S}\mu$ expresses the same information as $\mathbf{d}$. Therefore, we can define the representativity of $\mathbf{S}$ as the projection between the innovation vector and $\mathbf{H}\,\mathbf{S}\mu$:

$$Rep = \cos\left(\mathbf{d}, \mathbf{H}\,\mathbf{S}\mu\right), \tag{13}$$

where the cosine is computed with respect to the chosen inner product.

Because of its definition (13), a representativity value of $1/\sqrt{2} \approx 0.71$ means that the innovation vector is equally projected over the EOF subspace and over its complementary. That is, representativity values below 0.71 mean that the complementary space represents the observations better than the EOF basis.

Equations (9)-(13) are applied to the monthly averaged maps of each gridded field (Reynolds SST, FSU wind stress, and the Kelvin and Rossby modes of the thermocline depth), and the value of μ and R are computed using maps of FSU wind stresses from

January 1961 to December 1998.

a. FSU wind stress

Figure 6 shows the temporal evolution of the representativity of the EOFs to describe the spatial patterns of the FSU wind stress anomalies (from the 1961-92 climatology). In this figure, two curves are shown. The solid one is obtained when the data are limited to the band 10°S - 10°N, while the dotted line is obtained with 19°S - 19 °N data (i.e., the maximum latitude at which the ZC model computes SST anomalies). The representativity of the EOFs measures the ability of such a basis to fit the observations. Therefore, when the number of basis functions increases or the number of observations decreases, better the fit could be. Similarly, the representativity decreases when the data are extended to all the model grid. However, the representativity of both data sets is clearly larger than the critical value.

b. Reynolds SST

Figure 7 shows the representativity of the EOFs to describe the spatial patterns of the Reynolds SST fields. The results show the high skill of the model to reconstruct the spatial distribution of the SST, especially at low latitudes, where the mean representativity is .98.

If figures 6 and 7 are compared to the time evolution of the Niño-3 SST, it can be observed that the representativity is maximum at each culmination of warm and cold events and is minimum at each rapid transition from one event to the other, specially when high latitudes are considered (the representativity systematically diminishes before and after each event). This behavior can be explained by the inherent variation of the noise to signal ratio, and by

the fact that the parameters of the model have been adjusted to simulate the dynamics of large warm events of the ENSO cycle.

c. Kelvin and Rossby modes of the thermocline depth

Figure 8 shows the variation of the representativity depending on the number of Rossby modes used. Except for the last months, the maximum representativity is obtained when only the first Rossby mode is combined with the Kelvin wave. Oscillations of the representativity of the model are not directly related to the Niño-3 SST index, and a period of stable high representativity appears between middle 1995 to the end of 1996, a period quite calm in terms of Niño-3 variations.

As in the case of winds and the SST, the representativity decreases when information at the high latitudes is considered by the use of higher Rossby modes. The mean representativity decreases from .98 when 1 or 2 Rossby modes are used to 0.96 when 5 or 10 modes are used.

To ensure the compatibility between the physics of the model and the observations, only the first two Rossby modes will be retained. This does not represent too much loss of information in the 10°S - 10°N band. For example, figure 9 shows the thermocline depth in December 1998 derived from sea level (Fig. 9a), and reconstructed by the Kelvin and the two first Rossby modes (Fig. 9b). The first two modes are sufficient to describe most of the spatial structure of the thermocline depth. Similarly, the comparison between the Niño-3 and Niño-4 values of the thermocline depth reconstructed using the Kelvin and the first Rossby mode (not shown) and Fig. 1c are indistinguishable.

d. Combining observations

The degradation of the representativity of the model when multiple data fields are considered simultaneously is shown in Fig. 10. The best interdependency accounted by the ZC corresponds to the link between the thermocline depth and the SST (dashed line). The lesser skill of the model to account for the spatial patterns of the wind stress and the thermocline depth is clear as well (dotted line), specially during year 1994.

When all the observed fields are considered together, the mean representativity of the model decreases to 0.88, considerably smaller than the representativity of the model when each field is considered alone, but still higher than the critical value.

Finally, Fig. 11 provides an example of the compatibility between observations of SST and oceanic equatorial waves with the physics of the model. Compared are the observed value of the Niño-3 index with the index associated to the SST anomaly field derived from the EOF projection to observations of the Kelvin and first Rossby modes. The strong similarities between both curves suggests that the ocean model captures the basic relationship between thermocline depth and SST in the eastern equatorial Pacific. This result also shows the significance of the Kelvin and first Rossby wave to determine the current state of the thermocline depth over that region.

As a conclusion, in spite of the strong simplifications of the ZC model, and its well known limitations to locate correctly some spatial features of the physical fields (Zebiak 1986; Cane et al. 1986; Dewitte and Perigaud 1996), the ZC model is relatively adept at representing the covariability of the wind stress, the SST and the thermocline depth when observations are limited to low latitudes, and the thermocline depth is introduced as a function of the

first Rossby modes.

6. The assimilation method

a. The reduced order Kalman Filter

The reduced order Kalman Filter (ROKF) used in this work is based on SEEK (Singular Evolutive Extended Kalman Filter, Pham et al. 1998) equations. The SEEK filter is a sub-optimal version of the Kalman Filter (Kalman and Bucy 1960) based on two hypothesis: the reduced rank of the error covariance matrices, and the conservation of the rank in time. Although these hypothesis cannot be justified for a nonlinear system, the SEEK filter has been used successfully to assimilate T/P SSH into a primitive equation tropical Pacific ocean model (Verron et al. 1999), and to reconstruct the mesoscale circulation of the mid-latitude western Atlantic circulation (Brasseur et al. 1999).

Analysis-step equations of the SEEK filter are directly derived from the analysis equations of the Kalman filter by expressing the error covariance matrix, $\mathbf{P}^f$, at time t_k, as

$$\mathbf{P}_k^f = \mathbf{S}_k \Lambda_k \mathbf{S}_k^\top, \tag{14}$$

where subindex k refers to the time t_k. If n is the number of components of the state vector (3), then matrices $\mathbf{P}_k^f$, $\mathbf{S}_k$, and Λ_k, are of order n. However, if the rank of $\mathbf{P}_k^f$ is r, only r columns of matrices $\mathbf{S}_k$ and Λ_k should be retained. Thus, the numerical cost of the analysis step is reduced if the Kalman filter equations are expressed in terms of the r-dimensional space defined by the columns of $\mathbf{S}_k$.

Following the notation recommended by Ide et al. (1997), if $\mathbf{y}_k^o$ represents a set of observations at time t_k, and $\mathbf{x}_k^f$ is the current guess of the system state, the new estimate

of the state of the system, $\mathbf{x}_k^a$, is obtained using equations (see Pham et al. 1998):

$$\Lambda_{k+1} = \left[\Lambda_k^{-1} + (\mathbf{H}_k \mathbf{S}_k)^\top \mathbf{R}_k^{-1} \mathbf{H}_k \mathbf{S}_k\right]^{-1}, \tag{15}$$

$$\mathbf{K}_k = \mathbf{S}_k \Lambda_{k+1} (\mathbf{H}_k \mathbf{S}_k)^\top \mathbf{R}_k^{-1}, \tag{16}$$

$$\mathbf{x}_k^a = \mathbf{x}_k^f + \mathbf{K}_k \left[\mathbf{y}_k^o - \mathbf{H}_k \mathbf{x}_k^f\right], \tag{17}$$

$$\mathbf{P}_k^a = \mathbf{S}_k \Lambda_{k+1} \mathbf{S}_k^\top, \tag{18}$$

where $\mathbf{H}$ is the observation matrix, relating the components of the state vector (3) with the observations $\mathbf{y}^o$. The relative weight of the first guess and the observations is given by the gain matrix $\mathbf{K}_k$, which depends on error covariance matrices $\mathbf{P}_k^f$ (estimation of the error of $\mathbf{x}^f$) and $\mathbf{R}_k$ (estimation of the error of $\mathbf{y}^o$). Equations (15)-(18) are valid in the measure that equation (14) is valid.

Under the hypothesis that the model is linear and exact, the rank of the error covariance matrix is constant in time, and the the time evolution of $\mathbf{P}^f$ depends only on the time evolution of the columns of $\mathbf{S}_k$ (Pham et al. 1998). Then, the forecast-step of the Kalman Filter can be rewritten as:

$$\mathbf{x}_{k+1}^f = \mathbf{M} \mathbf{x}_k^a, \tag{19}$$

$$\mathbf{S}_{k+1} = \mathbf{M} \mathbf{S}_k, \tag{20}$$

$$\mathbf{P}_{k+1}^f = \mathbf{S}_{k+1} \Lambda_{k+1} \mathbf{S}_{k+1}^\top. \tag{21}$$

The generalization of equations (19)-(21) to a non-linear, non-perfect model is not straightforward because the rank of the covariance matrices may not be conserved in time, because of the dynamical coupling between the space defined by $\mathbf{S}_k$ and its null-space, i.e., equation

(21) is not longer valid. Moreover, because of the nonlinearity of the model, equation (20) cannot be solved explicitly, and only an approximation of the time evolution of the sub-space basis can be obtained.

These facts lead us to use an alternative method of equations (20)-(21). Let $\mathbf{d}_k$ be the innovation vector at time t_k, about the current forecast of the state of the system

$$\mathbf{d}_k = \mathbf{y}_k^o - \mathbf{H}\,\mathbf{x}_k^f. \tag{22}$$

It can be shown that

$$\mathrm{E}\left[\,\mathbf{d}_k\,\mathbf{d}_k^\top\,\right] = \mathbf{H}\,\mathbf{P}_k^f\,\mathbf{H}^\top + \mathbf{R}_k \tag{23}$$

$$= \mathbf{H}\,\mathbf{S}\Lambda_k\,(\mathbf{H}\,\mathbf{S})^\top + \mathbf{R}_k' \tag{24}$$

where $\mathbf{R}_k'$ contains both the observational error covariance and the error covariance on the null-space of $\mathbf{S}$ (Cane et al. 1996). Now, if equations (7)-(9) are used, we obtain

$$\mathbf{d}_k\,\mathbf{d}_k^\top = \mathbf{H}\,\mathbf{S}\mu_k\mu_k^\top(\mathbf{H}\,\mathbf{S})^\top + \mathbf{H}\,\mathbf{e}^*\,\mathbf{e}^{*\top}\,\mathbf{H}^\top \tag{25}$$

because of the orthogonality of the decomposition of the innovation vector. Note that the residual vector $\mathbf{e}^*$ is a function of the observational error and the error not accounted by the subspace defined by $\mathbf{S}$.

Because of the similarity between equations (24) and (25), we use $\mu_k\mu_k^\top$ as an *ad hoc* approximation of the error covariance on the analysis subspace, Λ_k. Two relationships between μ_k and Λ_k are used here.

$$\Lambda_k = \mathrm{diag}(\mu_k\mu_k^\top), \tag{26}$$

$$\text{or} \qquad \Lambda_k = \frac{1}{k}\sum_{j=1}^{k}\mu_j\mu_j^\top. \tag{27}$$

In any case, this methodology allows us to compute a non-stationary, non-diagonal estimate of the forecast error covariance matrix $\mathbf{S}\Lambda_k\mathbf{S}^\top$. By using equation (26) or (27), the parameterization of the dynamical coupling between the subspace defined by $\mathbf{S}$ and its complementary, and the error of the model, are not necessary because Λ_k comes directly from the current misfit between the observations and the current state of the system.

In summary, the ROKF algorithm uses equation (19) to obtain the current state of the system. Equations (22), (7)-(9) are used to compute the set of principal components $\mathbf{s}_k$ for the estimation of the current forecast error covariance by equations (26) or (27). Finally, equations (15)-(17) are used to correct the state of the system.

b. Nudging the surface wind stress

The performance of the reduced order KF will be compared with the performance of the nudging of surface wind stress in the equation for wind stresses. The method used here replicates the method proposed by Chen et al. (1995). That is, at each time step the anomalous model wind stress $\tau = (\tau_x, \tau_y)$ is substituted by $\alpha\tau_o + (1-\alpha)\tau$, where τ_o is the observed anomalous wind stress. Following Chen et al. (1995), α is a function of the latitude. In our experience of assimilation, α is a linear function of the latitude, with a value of 0.25 near the equator, and increasing by 0.1 per grid point up to the latitude where it attains the value 0.55, which is then held constant. That is, the model winds are given more weight in the equatorial region and less weight in the higher latitudes.

Neither nudging or the Kalman filter use the hypothesis of a perfect model. In the case of nudging, a new term is introduced into the dynamical equations in order to correct the output of the model. As this term is linear by respect the corrected field, the behavior of

the model dynamics is linearized when the data-model misfits are important. On the other hand, in the present application, the nudging of the wind stress does not modify the value of the prognostic fields (the Kelvin amplitude, and the Rossby component of the upper layer depth anomaly), but its forcing term. That is, nudging does not modify the current state of the system, but its ulterior time evolution. This approach strongly reduces the possible dynamical imbalances between the prognostic fields. On the contrary, algorithms based on the Kalman filter compute the dynamical evolution of the system with no relaxation term, allowing the full development of the nonlinearities of the dynamics. Account for the model error is done by computing an estimate about the error of the new system state. Therefore, misstatements in the determination of such an error are allowed to propagate in a nonlinear way, contaminating all the dynamical scales of the system. On the other hand, all the components of the system are updated by equation (17). This allows a potential way for a faster extrapolation of the information from observations to all the variables of the model, but also increases the risk of dynamical imbalances, i.e. initialization shocks may exist after each assimilation step.

7. Model initialization experiments

A set of assimilation experiments, using both nudging of the wind stress and the ROKF presented in the previous section, are performed in order to obtain the initial conditions to predict the time evolution of the Niño-3 SST index.

Topex/Poseidon data is available since October 1992. Then, prediction experiments are restricted to the period January 1993 - December 1998. Note that the goal of these

experiments is to examine the potential of a ROKF to assimilate, directly and simultaneously, different types of data into the ZC model, but not to give an exhaustive comparison between different assimilation methods (this would require longer time series, including several warm and cold events).

Nudging and ROKF also differ in the temporal distribution of the observations used to identify the state of the system. In the case of nudging, the coupled model is integrated from January 1961 until the initial time for the forecast. During this period, the model wind stress is relaxed towards the observations. Thus, the final state is used as initial conditions for the prediction of ENSO. This means that initial conditions for ENSO forecasts are a function of the state of the system from 1961. On the contrary, the ROKF is initialized on October 1992 with the long term mean anomaly of the model. Then, observations are used to give a first estimate of the state of the system at this time. Then, sequential assimilation cycles are done until reaching the initial time for the forecasts.

Initial states for 15-month ENSO forecasts are computed each month from January 1993 to December 1998. ENSO forecasts are compared with observations via the NIÑO 3 index. The correlation and the root mean square (RMS) of the error between the predicted values and the observations, are used to assess the performance of the initialization.

a. Nudging the wind stress

The first experiment presented here uses the nudging method to retrieve the initial conditions for ENSO forecasts, and thereby establish a baseline for the ROKF experiments. Wind observations are considered on the model grid, that is between 19°S - 19°N, and are the only data assimilated.

Figure 12 displays Niño-3 index forecasts with these experiments. The thick solid line represents the hindcast Niño-3 SST index, the thin solid lines show the 15 month forecasts, and the dashed-line indicates the observed value. The curves show low correlation between hindcasts and observations during the whole period under consideration, indicating that nudging of the wind stress is not able to reproduce the temperature of the eastern Pacific. Forecasts without assimilation (thin solid lines) show little dispersion by respect the time evolution with assimilation (thick solid line), indicating small impact of new observations onto the evolution of the system, i.e., a strong inertia of the model. This explains the delay to identify the onset and the climax of the 1997 warm episode.

Figure 13 compares the correlation and RMS error for different lead time predictions with the initialized model, and the Niño-3 persistence. Zero lead time represents the initialization time. The results show the small correlations at short lead times, errors greater than persistence for lead times shorter than 8 months, and better than persistence for longer lead times.

b. Reduced Order Kalman Filter

The ROKF is now used to assimilate simultaneously wind stress, SST and the Kelvin and the first Rossby mode of the thermocline depth into the ZC model. As the assimilation procedure starts in October 1992, the assimilation results might still be very sensitive to the initial conditions. As before, each analyzed field is saved to be used as initial conditions for a fifteen-month forecast.

Initialization experiments are performed with different values of r, i.e. with different number of multivariate EOFs. The correlation between the observations and the associated

lead time forecasts is displayed in Fig. 14(a). This plot shows the correlations for values of r ranging from one to seventy. The plot also displays the correlation obtained with the nudging initialization. The ROKF initialization is providing higher correlation at short time leads than nudging does. This is related to the fact that SST observations are being now assimilated. The most striking feature of these results is that the method is robust only for lead times up to 4 or 5 months. This fact may be related to the results of Goswami and Shukla (1991) who found that the growth of small initial errors in the ZC coupled model is governed by processes with two well-separated time scales. The fast time scale process induces error doubling scales of about 5 months. The slow time scales induces error doubling scales of about 15 months. Therefore, even if the ROKF seems to be able to identify the current state of the system (high correlation at the initial time), the method is unable to capture the fast-growing error components, and the correlation between the prediction states and observations diminishes for lead times longer than the doubling scales of such errors.

Different initialization experiments of the coupled model have been performed by using a non-diagonal Λ_k (equation (27)) instead of a diagonal matrix (equation (26)), or by using the error-metric, equation (12), instead of the identity metric, equation (9); or by assimilating SST observations distributed over all the model grid instead of SST observations limited within the band [10°S-10°N]. In all the cases, the robustness of the method is limited to 4 to 5 months. As an example, Fig. 14(b) displays the correlation between observations and forecast for one of these initialization experiments, for which equations (27), (12) have been used to assimilate observations with SST observations distributed over all the model grid.

The comparison of both panels of Fig. 14 shows several common features. Note that

high prediction skill of the month-to-month predictions is obtained by using a small number of modes (7 EOFs), and the use of higher order, non statistically significant modes does not improve the correlation for such short-term forecasts. The fact that seven modes are enough to reach a high correlation skill, is explained by the high amount of variance explained by the first modes, so the basic structure of the state of the system is recovered by using a small number of modes. However, the rapid decrease of the correlation between forecasts and observations indicates that these modes are not able to identify the time evolution of the system. Higher order modes are necessary to extend the lead times of the predictions, indicating that the lead order EOF modes capture the stationary variance of the fields, but not its time evolution.

Another common feature of all these initialization experiments is the fact that there always exists a range of optimal values for the number of EOFs providing high forecast-observation correlation values for lead times ranging from 9 to 15 months. In all the experiments, the range of optimal values is always located between 40-60 modes, consisting of more than one single value, which implies that such high correlations are not due to chance.

Figure 15 shows ENSO predictions obtained by initialization experiments using 20 EOFs (Fig. 15(a)), and 47 EOFs (Fig. 15(b)). In these experiments, equations (9), and (26) are used, and SST observations are extended until 19°S - 19°N. Hereafter these experiments are cited as r20 and r47. Comparing the thick solid line on Figs. 15(a) and 15(b), it can be noticed that the initial conditions of the prediction experiments become closer to the observed value as the number of modes increases.

The main shortcoming of the r20 initialization is the inability of the forecasts to identify

the termination of warm events, specially at the end of 1993, or at the end of 1998. To determine the reason for this, we compare the initial conditions given by these initialization experiments. Figure 16 shows the difference between r20 and r47. Fields with differences larger than two standard deviations are the zonal wind stress and the Kelvin wave amplitude. The impact of these differences has been examined by resetting to zero the wind stress components or the Kelvin amplitude of the initial conditions provided by the experiment r47. The correlation and the error RMS of these experiments are shown in Fig. 17. Note the small impact of resetting the wind stress to zero on the prediction skill of the solutions, and the larger impact of resetting the Kelvin amplitude. This behavior was expected because the oceanic Kelvin wave is one of the prognostic variables of the model, while the atmospheric model mainly depends on the current SST distribution. Examining the initial Kelvin and Rossby fields in experiment r20 (not shown), it was found that both fields were too weak near the western boundary (Fig 16 shows that the differences of the Rossby waves are generally large there). Then, both propagation of the Kelvin wave as well as the boundary reflection of Rossby waves, provide too weak Kelvin waves that rapidly contaminate the solution at the central and eastern equatorial Pacific. Note that the failure of the lead multivariate EOFs to reconstruct the required wave amplitudes is directly related to one of the well known properties of the model, i.e. the weak variability of the coupled model west of the dateline (Cane et al. 1986).

Initial conditions obtained with the r47 experiment are shown in Fig. 18. Note that the region, near the western boundary of the system, with no observations of Kelvin and Rossby waves (see Fig. 1), has now been filled (Fig. 18), according to the dynamics of the model, i.e.,

the delayed oscillator mechanism. All fields show a good agreement with the observations, even if, as discussed before, none of the reconstructed fields displays the variability of the observation near the western boundary. Forecasts from these initial conditions successfully predict the warm event of 1997-98 by identifying its onset, rapid increase of the temperature, amplitude of the event, and its duration. Forecasts since May 1996 predict a warm event for the 1997, indicating that preconditions of such an event may be present in the tropical Pacific at least fifteen months prior to the event, in agreement with the remarks of McPhaden and Yu (1999). These experiments also indicate that a nonlinear, coupled model based on the delayed oscillator may generate the correct amplitude of the warm event starting from conditions before November 1996, the time when wind anomalies associated with the Madden-Julian Oscillation appear. Such anomalies have been suggested as being important for the timing and amplitude of the event (McPhaden and Yu 1999).

1) Sensitivity to the wave errors

The error of the Kelvin and Rossby modes have been determined by assuming that the reduced gravity approach is strictly verified over the band [10°S-10]degN] (see section 3). The sensitivity of the results to these values has also been examined. This has been done with six additional experiments. In these new model initializations, the Kelvin and Rossby errors have been, separately or simultaneously, increased (and diminished) 20% of the value given in section 3.

The results, not shown, show small impact to changes on the error estimates for the Kelvin and Rossby modes. Results are almost insensitive to changes on Rossby error estimates because of the large number of observations, 2378, being assimilated. Results are slightly

more sensitive to changes in the Kelvin error estimate. When the Kelvin error estimate is increased 20%, the skill of the forecasts diminishes for 6 to 8 month lead time predictions, but remain unchanged for higher lead times. However, when the Kelvin error estimate is reduced 20%, we have noticed a more noticeable decrease of the skill of the predictions for lead-times larger than 12 months. Thus, the Kelvin error estimate is low enough to provide an estimate of the thermocline structure over the equator, but is large enough to avoid an overfitting that degrades the prediction for interanual time scales.

8. Summary

A simplified reduced order Kalman Filter has been applied to the coupled ocean-atmosphere nonlinear model of Zebiak and Cane. The multivariate EOFs of the coupled model are used to reduce the dimension of the problem. In the case of the state vector used in this work, the size of the covariance matrices $\mathbf{P}$ used by the classic Kalman filter is 5 Gb. When the ROKF is used with $r = 100$, the size of the matrices $\mathbf{S}$ and Λ is 14 Mb and 40 Kb respectively. The ROKF is used to assimilate fields of the SST, surface wind stresses and thermocline depth (derived from the T/P sea level) into the coupled model every month.

Projecting the Kalman Filter equations onto the multivariate EOFs of the model means that, although incomplete, the physics of the model is used to reduce the degrees of freedom of the KF algorithm. To ensure a relative compatibility with the model, observations are limited to the band 10°N - 10 °S. A representativity study shows that the ability of the EOFs to account for the covariability of the observed fields decreases at high latitudes.

The numerical cost of the filter is highly reduced by computing an estimate of the forecast

covariance matrix, on the subspace defined by the multivariate EOFs, from the available observations. On the other hand, as the time evolution of Λ comes from the observations instead of being computed by the time evolution of the error covariance matrix, the dynamical interaction between the EOF subspace and its nullspace has been avoided.

The prediction skill of the forecasts initialized from the outputs of the filter has been compared to persistence and the nudging of surface wind stress. The forecasts obtained from the filter show higher correlations with observations than the results coming from the nudging experiment for all lead times. However, the method is robust, to changes in the number of EOFs for forecasts shorter than 4 to 5 months, indicating that the method is not able to identify the error associated with the fast error modes of the model described by Goswami and Shukla (1991). The evolution of the forecast skill as a function of the number of modes is insensitive to changes in the metric used to fit the EOFs to the observations, and to how the current value of Λ is constructed. Therefore, such a limit may be related to the nature of the basis functions, i.e. the multivariate EOFs. These functions describe the stationary modes of variance of the fields rather than the time evolution of perturbations of the system. Therefore, a natural extension of the present work is the use of basis of functions accounting for the dynamic evolution of the fields, as the singular vectors of the transition operator of the ZC model, or the evolutive version of the SEEK filter, allowing the time evolution of the matrix $\mathbf{S}$.

Besides the inability of the EOF basis to capture the growing errors of the coupled model, the failure of the lead EOFs to provide initial states containing sufficient information to identify the actual evolution of the system is related to known shortcomings of the physical

model itself. Thus, a common feature of all the experiments performed by the authors is that the optimal number of EOFs is always larger than the number of modes passing the statistical significance test of Overland and Preisendorfer (1982).

Despite the fact that the filter has been shown to be robust for time scales shorter than 4-5 months, the filter has provided initial conditions of the tropical Pacific, predicting the onset of the 1997-98 warm event fifteen months before its occurrence. These predictions start from states prior to the first westerly wind burst associated with the Madden-Julian oscillation, often suggested as a triggering mechanism for the onset of warm events. Such a result indicates the possibility that the delayed oscillator mechanism may itself be the origin of the correct amplitude of the 1997-98 El Niño event.

ACKNOWLEDGEMENT. We acknowledge NASA RTOPs grants 621-75-02, 229-15-35, and 291-01-08 that supported this work. The authors would like to thank Drs. Cane, Kaplan, and Vintzileos for helpful discussions and criticisms. We also thank Eric Hackert for providing TOPEX/Poseidon data and his invaluable experience, as well as to Dr. Blumenthal of IRI/LDEO Climate Data Library for providing the oceanographic and atmospheric datasets used in this work.

References

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

Blumenthal, M.B. 1991: Predictability of a coupled ocean-atmosphere model, *J. Climate*, **4**, 766-784.

Boulanger, J.-P., and C. Menkes, 1995: Propagation and reflection of long equatorial waves in the Pacific ocean during 1992-1993 El Niño, *J. Geophys. Res.*, **100**, 25041-25059.

Brasseur, P., J. Ballabrera-Poy, and J. Verron, 1999: Assimilation of altimetric data in the mid-latitude oceans using the SEEK filter with an eddy-resolving, primitive equation model, *J. Mar. Sys.*, **22**, 269-294.

Busalacchi, A.J., 1997: Oceanic observations, *J. Meteorol. Soc. Japan*, **75**, 131-154.

Busalacchi, A.J., M.J. McPhaden, and J. Picaut, 1994: Variability in equatorial Pacific sea surface topography during the verification phase of the TOPEX/POSEIDON Mission, *J. Geophys. Res.*, *99*, 25725-25738.

Cane, M.A., and R.J. Patton, 1984: A numerical model for low-frequency equatorial dynamics, *J. Geophys. Res.*, **14**, 1853-1863.

Cane, M.A., and S.E. Zebiak, 1985: A theory for El Niño and the Southern Oscillation, *Science*, **228**, 1095-1087.

Cane, M. A., S.E. Zebiak, and S.C. Dolan, 1986: Experimental forecasts of El Niño, *Nature*, **321**, 827-832.

Cane, M. A., A. Kaplan, R. N. Miller, B. Tang, E. C. Hackert, and A. J. Busalacchi, 1996: Mapping tropical Pacific sea level, Data assimilation via a reduced state space Kalman Filter, *J. Geophys. Res.*, **101**, 22599-22617.

Chen, D., S.E. Zebiak, A.J. Busalacchi, and M.A. Cane, 1995: An improved procedure for El Niño forecasting: Implications for predictability, *Science*, **269**, 1699-1702.

Chen, D., S.E. Zebiak, M.A. Cane, and A.J. Busalacchi, 1997: Initialization and Predictability of a coupled ENSO forecast model, *Mon. Wea. Rev.*, **125**, 773-788, 1997.

Chen, D., M.A. Cane, and 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, D., M.A. Cane, and S.E. Zebiak, 1999: The impact of NSCAT winds on predicting the 1997/98 El Niño: A case study with the Lamont-Doherty Earth Observatory model, *J. Geophys. Res.*, **104**, 11321-11327.

Dewitte, B, and C. Perigaud, 1996: El Niño-La Niña events simulated with Cane and Zebiak's model and observed with satellite or in situ data. Part II: Model forced with observations, *J. Clim.*, **9**, 1188-1207.

Gill, A.E., 1982: *Atmosphere-Ocean dynamics*, Academic Press.

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

Ide, K., A.F. Bennett, P. Courtier, M. Ghil, and A.C. Lorenc, 1997: Unified notation for data assimilation: operational, sequential and variational. *J. Meteorol. Soc. Jpn.*, **75**, 181-189.

Kalman, R. E., and R. S. Bucy, 1960: A new approach to linear filtering and prediction problems, *J. of Basic Engineering (ASME)*, **82D**, 35-45.

Kaplan, A., M.A. Cane, Y. Kushnir, A.C. Clement, M.B. Blumenthal, and B. Rajagopalan, 1998: Analyses of global sea surface temperature 1856-1991, *J. Geophys. Res.*, **103**, 18567-18589.

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

Kleeman, R., 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 ocean-atmosphere forecast model, *Mon. Wea. Rev.*, **123**, 3103-3113.

M. Latif, T.P. Barnett, M.A. Cane, M. Flügel, N.E. Graham, H. von Storch, J.-S. Xu, and S.E. Zebiak, 1994: A review of ENSO prediction studies, *Climate Dynamics*, **9**, 167-179.

McPhaden, M.J., A.J. Busalacchi, R. Cheney, J.-R. Donguy, K.S. Gage, D. Halpern, M. Ji, P. Julian, G. Meyers, G.T. Mitchum, P.P. Niiler, J. Picaut, R.W. Reynolds, N. Smith,

and K. Takeuchi, 1998: The tropical ocean-global atmosphere observing system: A decade of progress, *J. Geophys. Res.*, **103**(C7), 14169-14240.

McPhaden, M.J., and X. Yu, 1999: Equatorial waves and the 1997-98 El Niño, *J. Geophys. Res. Lett.*, **26**, 2961-2964.

Miller, R.N, A.J. Busalacchi, and E.C. Hackert, 1995: Sea surface topography fields of the tropical Pacific from data assimilation, *J. Geophys. Res.*, **100**, 13389-13425.

Neelin, J.D., M. Latif, and F.-F. Jin, 1994: Dynamics of coupled ocean-atmosphere models: The tropical problem, *Annu. Rev. Fluid. Mech.*, **26**, 617-659, 1994.

Overland, J. E., and R. W. Preisendorfer, 1982: A significance test for principal components applied to a cyclone climatology, *Mon. Wea. Rev.*, **110**, 1-4.

Perigaud, C., and B. Dewitte, 1996: El Niño-La Niña events simulated with Cane and Zebiak's model and observed with satellite and In situ data. Part I: Model Data comparison, *J. Clim.*, **9**, 66-84.

Pham, D. T., J. Verron, and M.-C. Roubaud, 1998: Singular evolutive extended Kalman filter with EOF initialization for data assimilation in oceanography, *J. Mar. Syst.*, **16**, 323-340.

Reynolds, R.W., and T.M. Smith, 1994: Improved global sea surface temperature analyses, *J. Climate*, **7**, 929-948.

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.

Shriver, J.F., and J.J. O'Brien, 1995: Low-frequency variability of the equatorial Pacific using a new pseudostress dataset: 1930-1989, *J. Clim.*, **8**, 2762-2768.

Smith, T.M, R.W. Reynolds, R.E. Livezey, and D.C. Stokes, 1996: Reconstruction of historical sea surface temperatures using Empirical Orthogonal Functions, *J. Clim.*, **9**, 1403-1420.

Stricherz, J. N., J. J. O'Brien, and D. M. Legler, 1992: Atlas of Florida State University Tropical Pacific Winds for TOGA, Mesoscale Air-Sea Interaction Group technical report, Fla. State Univ., Tallahassee.

Verron, J., L. Gourdeau, D. T. Pham, R. Murtugudde, and A. J. Busalacchi, 1999: An extended Kalman filter to assimilate satellite altimeter data into a non-linear numerical model of the tropical Pacific: method and validation, *J. Geophys. Res.*, **104**, 5441-5458.

Zebiak, S. E., 1986: Atmospheric convergence feedback in a simple model for El Niño,*Mon. Wea. Rev.*, **114**, 1263-1271.

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

Zebiak, S. E., 1989: Oceanic heat content variability and El Niño cycles, *J. Phys. Oceanogr.*, **19**, 475-486.

Figure Captions

Figure 1: Longitude-time for observed fields of zonal wind stress, Kelvin amplitude, first Rossby mode, and sea surface temperature. Plots dispay averaged values over the band [5°S-5°N].

Figure 2: Convergence of the error of the Kelvin wave amplitude and the error of the Rossby field as the number of Rossby modes increases.

Figure 3: Eigenvalue spectrum of the multivariate EOFs of the coupled model of Zebiak and Cane. Also shown are the 95% confidence levels (dashed line) from a Monte Carlo simulation.

Figure 4: First multivariate EOF: (a) Sea Surface temperature anomaly. (b) Surface wind stress anomaly. (c) Zonal variability of the equatorial Kelvin wave. (d) Rossby component of the upper layer depth anomaly. Units are standard deviations.

Figure 5: Second multivariate EOF: (a) Sea Surface temperature anomaly. (b) Surface wind stress anomaly. (c) Zonal variability of the equatorial Kelvin wave. (d) Rossby component of the upper layer depth anomaly. Units are standard deviations.

Figure 6: Ability of the multivariate EOFs to represent the spatial structure of the FSU wind stress during the period Jan 1961 to Dec 1998. Solid line: data is limited to the band 10°S - 10°N. Dotted line: data is for the band 19°S - 19°N.

Figure 7: Ability of the multivariate EOFs to represent the spatial structure of the Reynolds SST during the period Nov 1981 to Dec 1998. Solid line: data is limited to the band 10°S - 10°N. Dotted line: data is on the band 19°S - 19°N.

Figure 8: Ability of the multivariate EOFs to represent the spatial structure of the thermocline depth derived from Topex/Poseidon sea surface topography during the period Oct 1992 to Dec 1998. The maps of thermocline depth are decomposed in terms of meridional equatorial modes. The different representativity with different number of different Rossby modes (in addition to the Kelvin mode) is shown.

Figure 9: Spatial structure of the thermocline depth monthly anomaly in December 1998. (a) Derived from Topex/Poseidon sea surface level. (b) Reconstructed from the wave decomposition of the previous field. The modes used are the Kelvin and the two first Rossby modes. Amplitude in cm.

Figure 10: Ability of the multivariate EOFs to represent the spatial structure of all the data combined during the period Oct 1992 to Dec 1998.

Figure 11: NIÑO 3 index derived from observations (red line), and derived by projecting the sea level (Kelvin and first Rossby mode) observations onto the SST via the multivariate EOFs of the model (black line).

Figure 12: 15-month forecast of the NIÑO3 SST index. The thick solid line is associated to the initial conditions. Dashed line is the observed value. Thin solid lines show the evolution of forecasts starting every month. The initialization experiment is done by nudging the surface wind stress.

Figure 13: Forecast skill measured by the correlation (a), and the RMS error (b), between predicted and observed NIÑO3 SST index from January 1993 through December 1998. The initialization experiment is done by nudging the surface wind stress.

Figure 14: Forecast skill measured by the correlation between predicted and observed NIÑO3 SST index from January 1993 through December 1998. The number of EOFs used with the ROKF ranges from one to seventy. The ordinate axis represents the lead time (in months). (a) SST observations are limited to the band [10°S-10°N], and the identity metric is used to fit the EOFs to the data. (b) SST observations are assimilated over all the model grid, and the metric of the data-EOF fitting is given by the inverse of the error covariance matrix.

Figure 15: 15-month forecast of the NIÑO3 SST index. The thick solid line is associated to the initial conditions. Dashed line is the observed value. Thin solid lines show the evolution of forecasts starting every month. The ROKF with 20 EOFs (a), and 47 EOFs (b) are used. The metric is the identity matrix, $\Lambda_k = \mathrm{diag}(\mathbf{s}_k \mathbf{s}_k^T)$, and SST is assimilated over all the model grid.

Figure 16: Longitude-time plots of the difference between r20 and r47, averaged over the band [5°S-5°N].

Figure 17: Forecast skill measured by the correlation (a), and the RMS error (b), between predicted and observed NIÑO3 SST index from January 1993 through December 1998. The solid line corresponds to the experiment r47. Dashed line is obtained when the initial state resulting from experiment r47 is modified by resetting to zero the surface wind stress. The dash-dot line is obtained as before but resetting to zero the Kelvin wave amplitude.

Figure 18: Longitude-time plots of the initial conditions given by experiment r47, averaged over the band [5°S-5°N].

Fig 1.

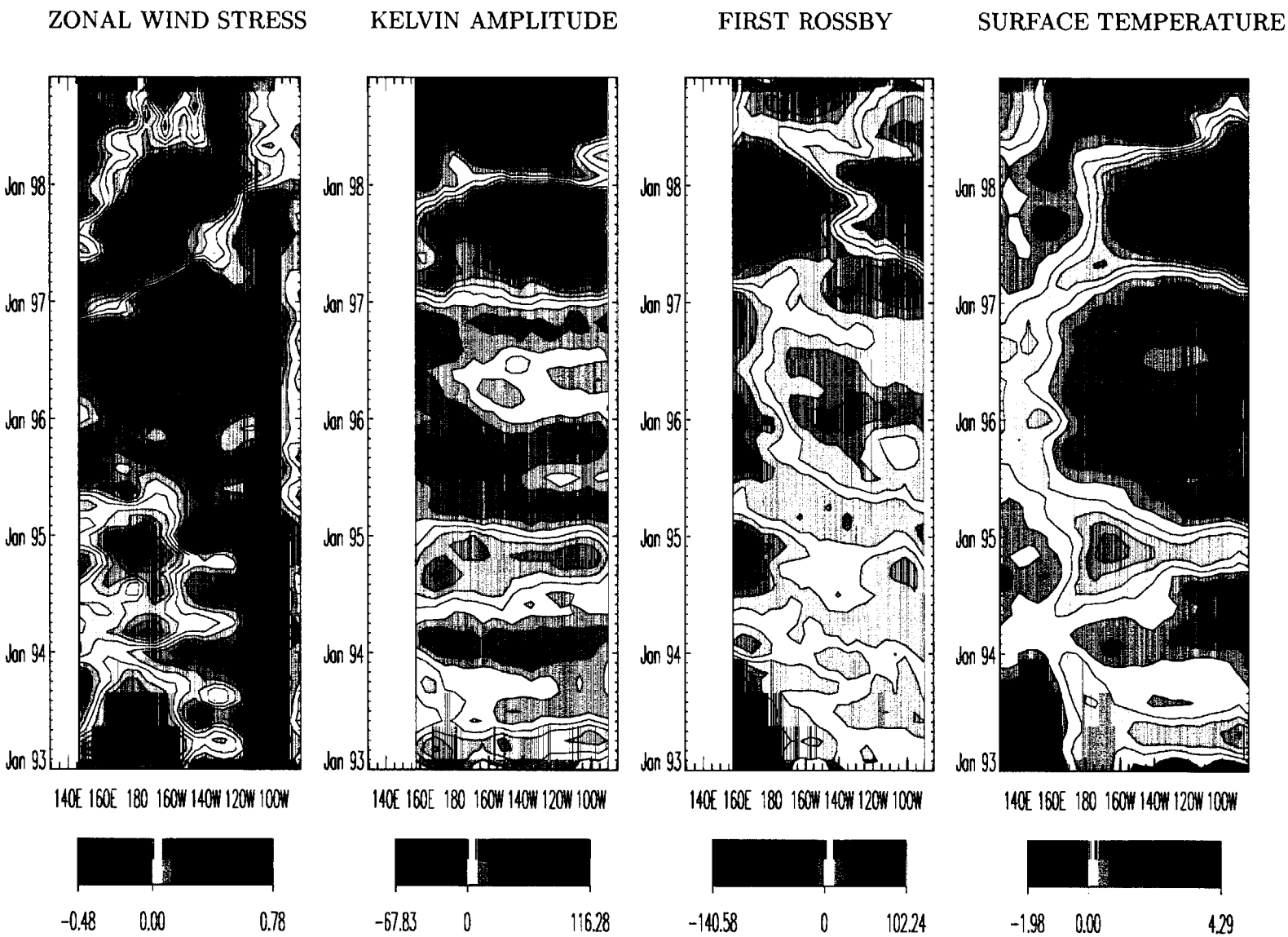

ZONAL WIND STRESS
KELVIN AMPLITUDE
FIRST ROSSBY
SURFACE TEMPERATURE
Jan 98
Jan 97
Jan 96
Jan 95
Jan 94
Jan 93
140E 160E 180 160W 140W 120W 100W
-0.48 0.00 0.78
-67.83 0 116.28
-140.58 0 102.24
-1.98 0.00 4.29

Fig 2

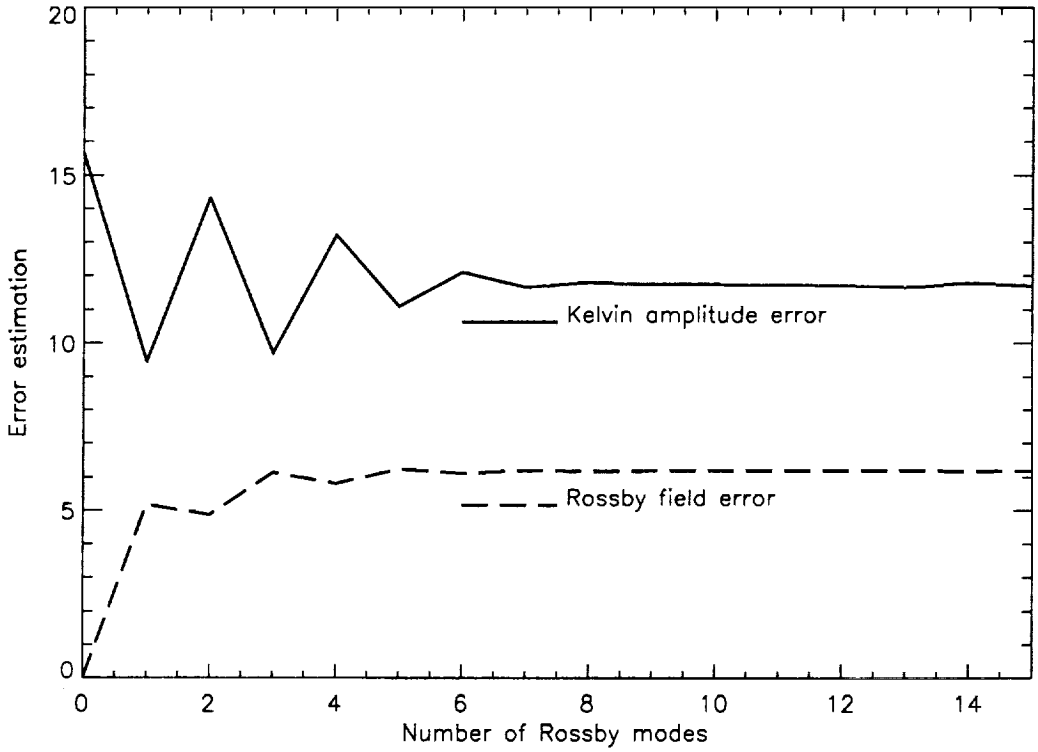

20
15
10
5
0
Error estimation
Kelvin amplitude error
Rossby field error
0
2
4
6
8
10
12
14
Number of Rossby modes

Fig 3.

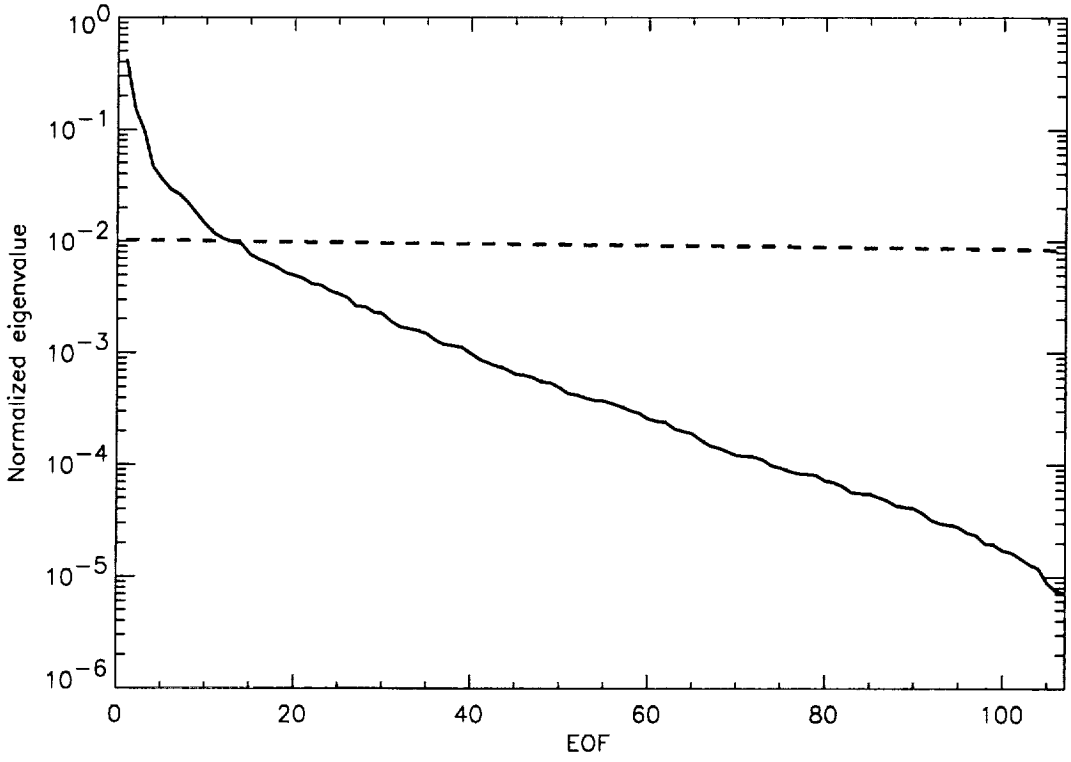

Normalized eigenvalue
10^0
10^{-1}
10^{-2}
10^{-3}
10^{-4}
10^{-5}
10^{-6}
0
20
40
60
80
100
EOF

Fig 4.

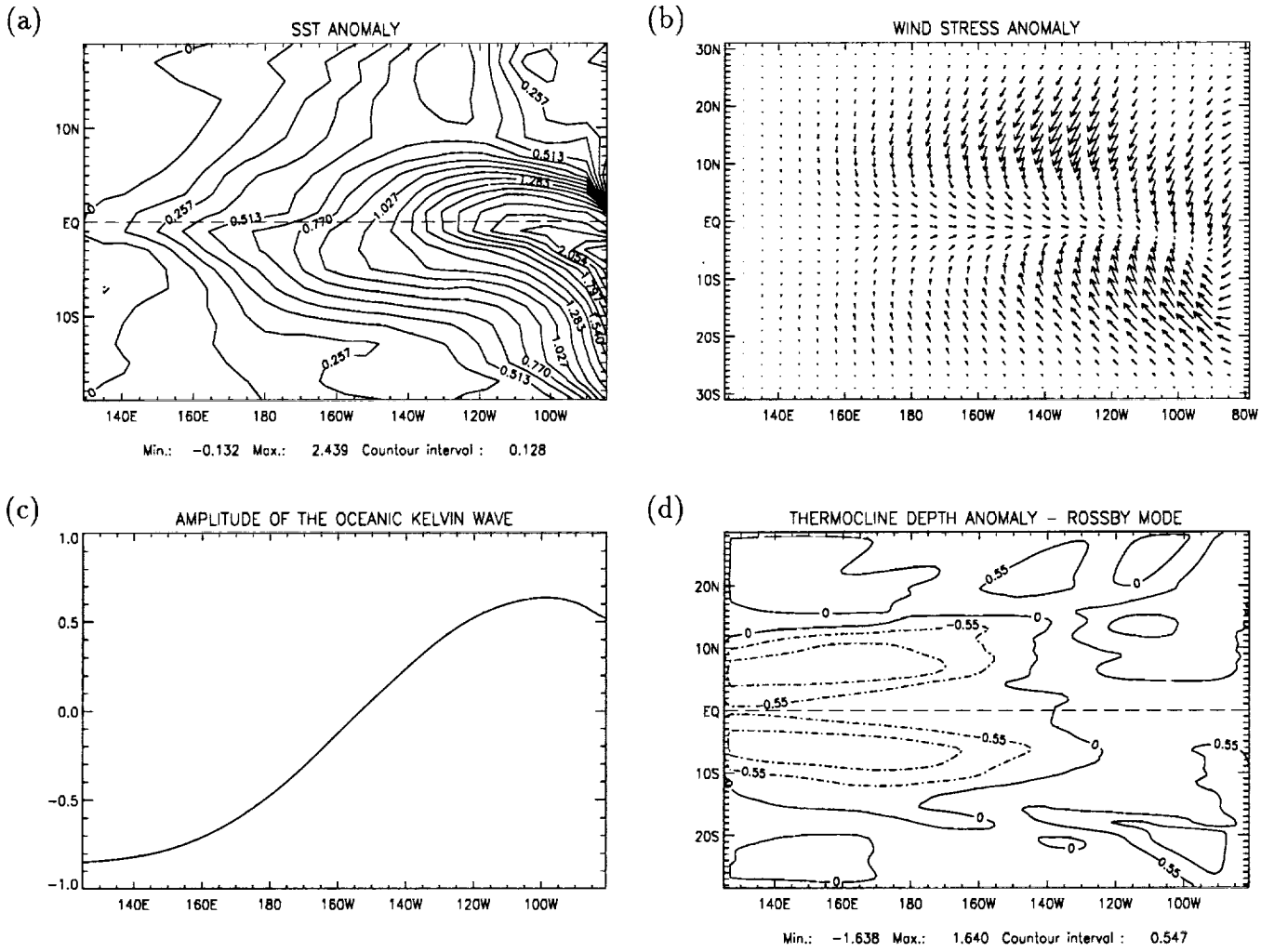

(a)
SST ANOMALY
10N
EQ
10S
140E
160E
180
160W
140W
120W
100W
Min.: −0.132 Max.: 2.439 Countour interval : 0.128
(b)
WIND STRESS ANOMALY
30N
20N
10N
EQ
10S
20S
30S
140E
160E
180
160W
140W
120W
100W
80W
(c)
AMPLITUDE OF THE OCEANIC KELVIN WAVE
1.0
0.5
0.0
−0.5
−1.0
140E
160E
180
160W
140W
120W
100W
(d)
THERMOCLINE DEPTH ANOMALY - ROSSBY MODE
20N
10N
EQ
10S
20S
140E
160E
180
160W
140W
120W
100W
Min.: −1.638 Max.: 1.640 Countour interval : 0.547

Fig 5.

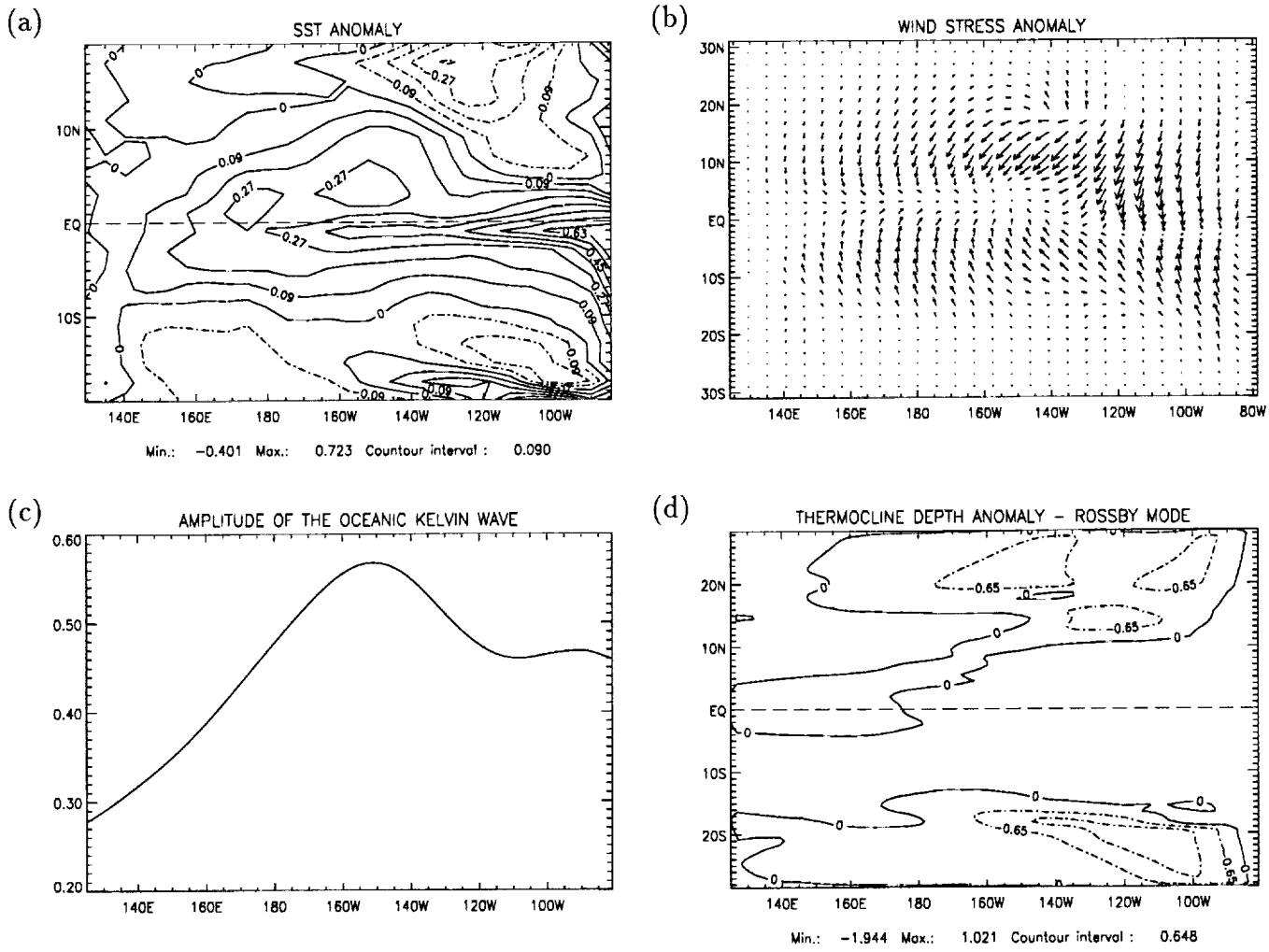

(a)
SST ANOMALY
10N
EQ
10S
140E
160E
180
160W
140W
120W
100W
Min.: −0.401 Max.: 0.723 Countour interval : 0.090
(b)
WIND STRESS ANOMALY
30N
20N
10N
EQ
10S
20S
30S
140E
160E
180
160W
140W
120W
100W
80W
(c)
AMPLITUDE OF THE OCEANIC KELVIN WAVE
0.60
0.50
0.40
0.30
0.20
140E
160E
180
160W
140W
120W
100W
(d)
THERMOCLINE DEPTH ANOMALY – ROSSBY MODE
20N
10N
EQ
10S
20S
140E
160E
180
160W
140W
120W
100W
Min.: −1.944 Max.: 1.021 Countour interval : 0.648

Fig 6.

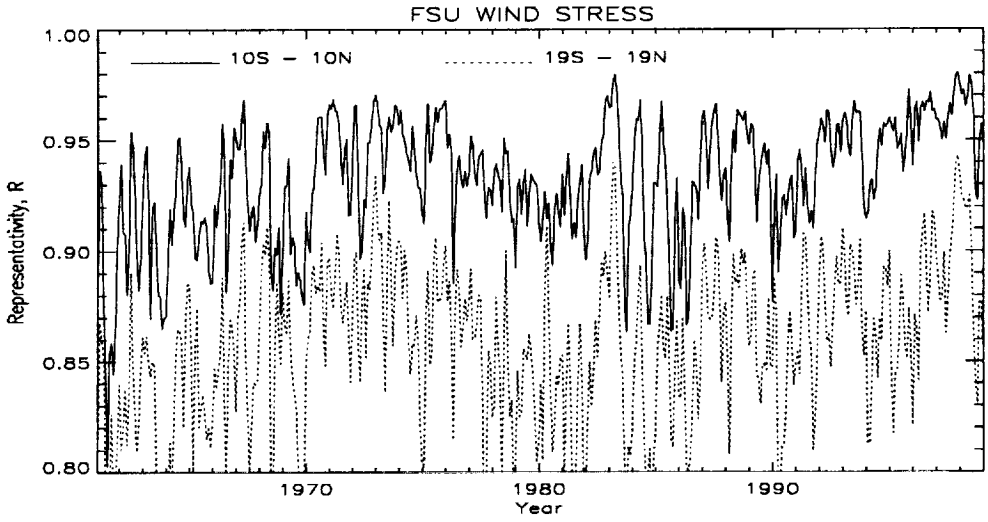

FSU WIND STRESS
10S – 10N
19S – 19N
1.00
0.95
0.90
0.85
0.80
Representativity, R
1970
1980
1990
Year

Fig 7.

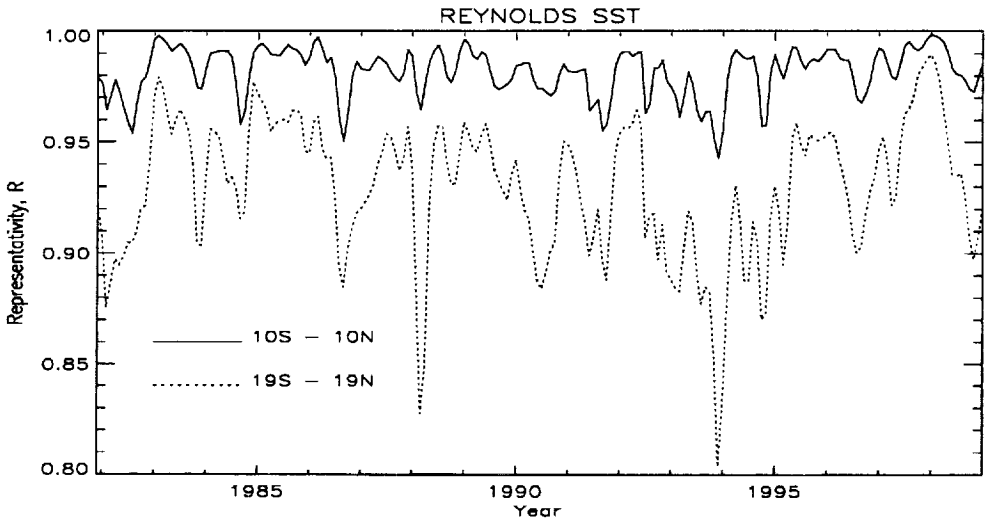

REYNOLDS SST
1.00
0.95
0.90
0.85
0.80
Representativity, R
10S – 10N
19S – 19N
1985
1990
1995
Year

Fig 8.

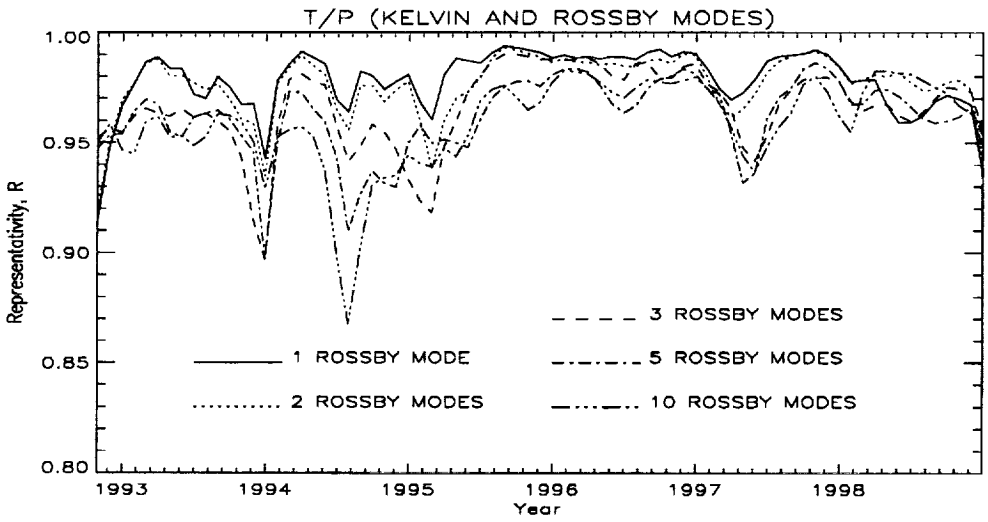
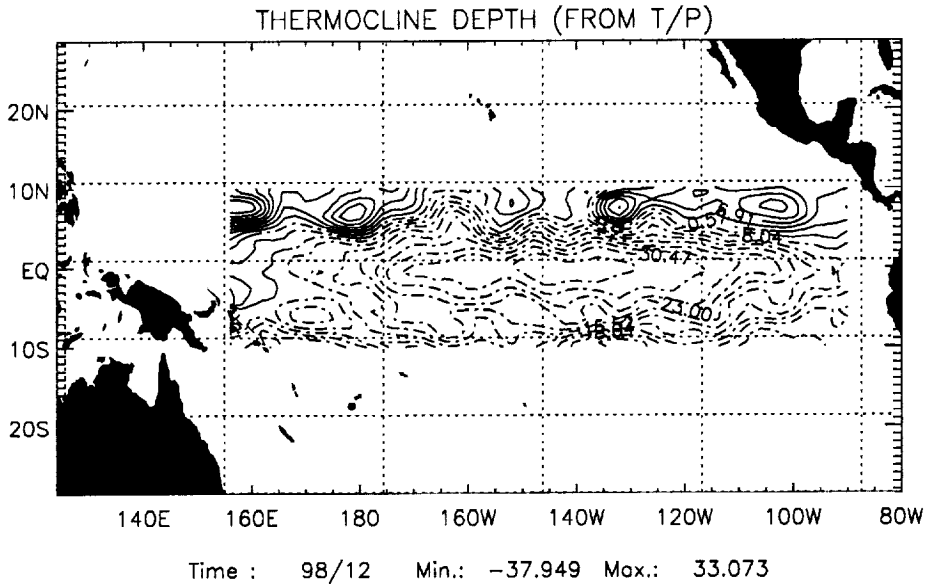

T/P (KELVIN AND ROSSBY MODES)
1.00
0.95
0.90
0.85
0.80
Representativity, R
1993
1994
1995
1996
1997
1998
Year
1 ROSSBY MODE
2 ROSSBY MODES
3 ROSSBY MODES
5 ROSSBY MODES
10 ROSSBY MODES

Fig 9.

(a)


THERMOCLINE DEPTH (FROM T/P)
20N
10N
EQ
10S
20S
140E
160E
180
160W
140W
120W
100W
80W
Time : 98/12 Min.: −37.949 Max.: 33.073

(b)

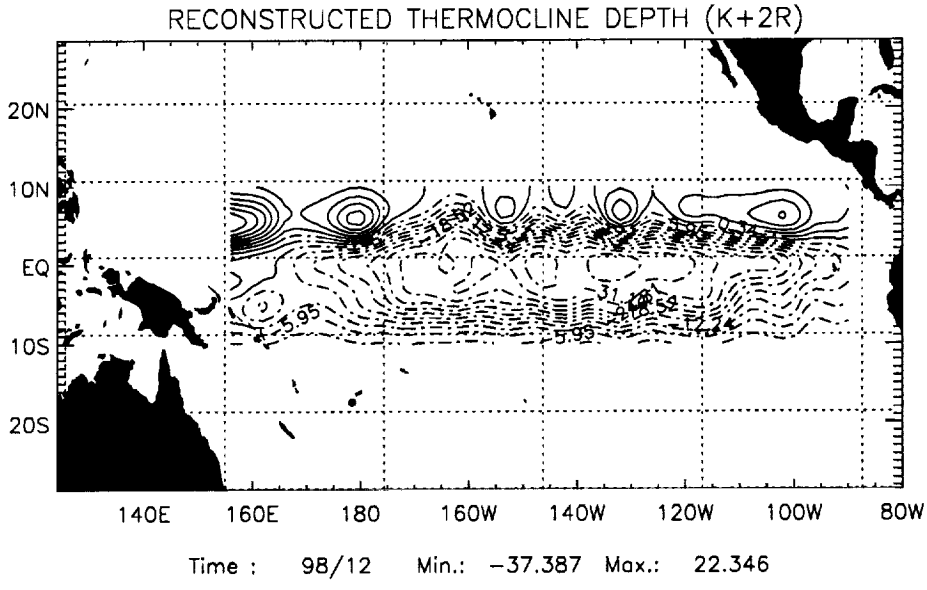

RECONSTRUCTED THERMOCLINE DEPTH (K+2R)
20N
10N
EQ
10S
20S
140E
160E
180
160W
140W
120W
100W
80W
Time : 98/12 Min.: −37.387 Max.: 22.346

Fig 10.

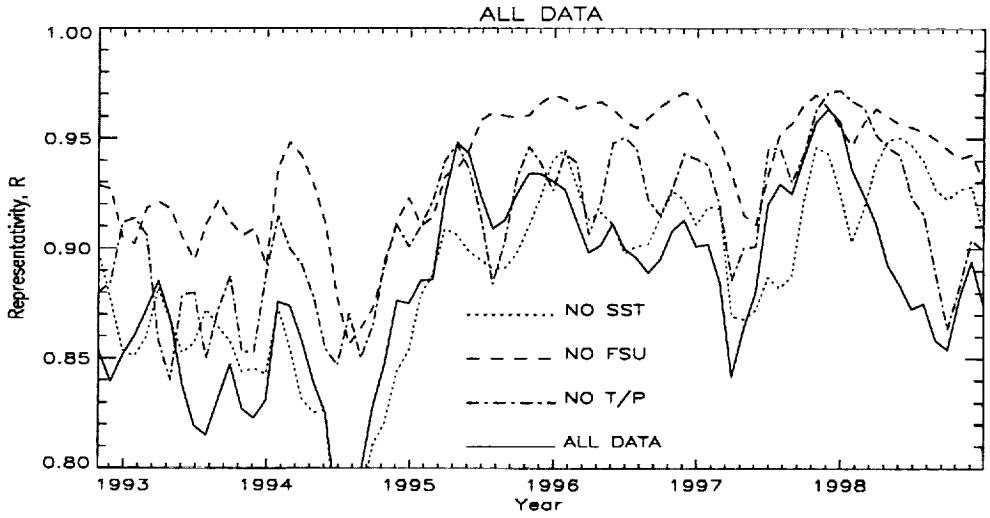

ALL DATA
1.00
0.95
0.90
0.85
0.80
Representativity, R
1993
1994
1995
1996
1997
1998
Year
NO SST
NO FSU
NO T/P
ALL DATA

Fig 11.

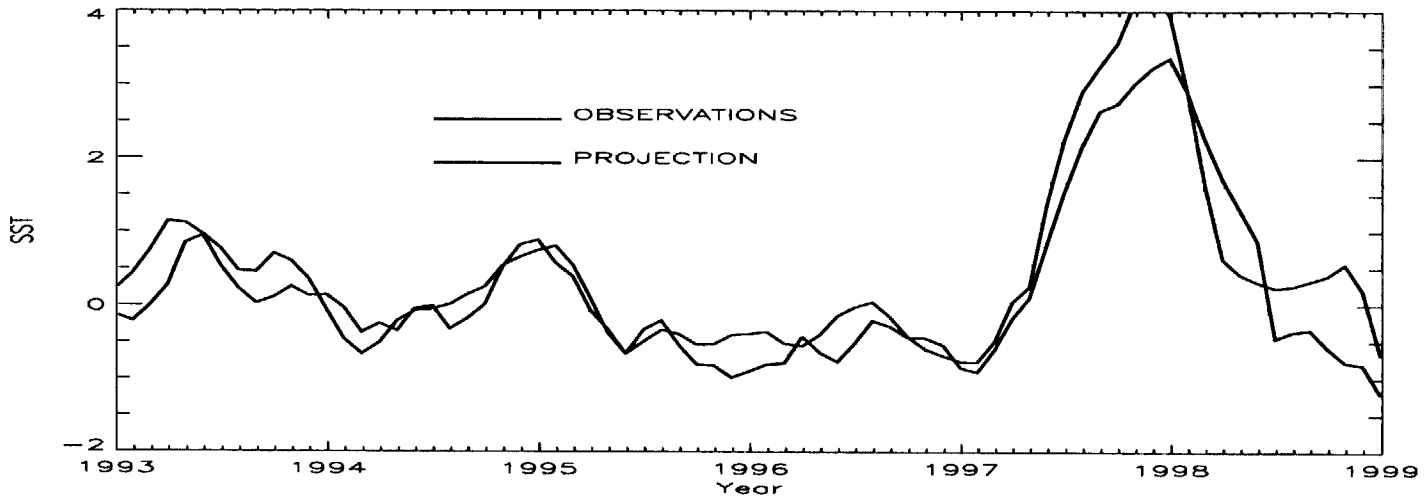

OBSERVATIONS
PROJECTION
SST
4
2
0
−2
1993
1994
1995
1996
1997
1998
1999
Year

Fig 12.

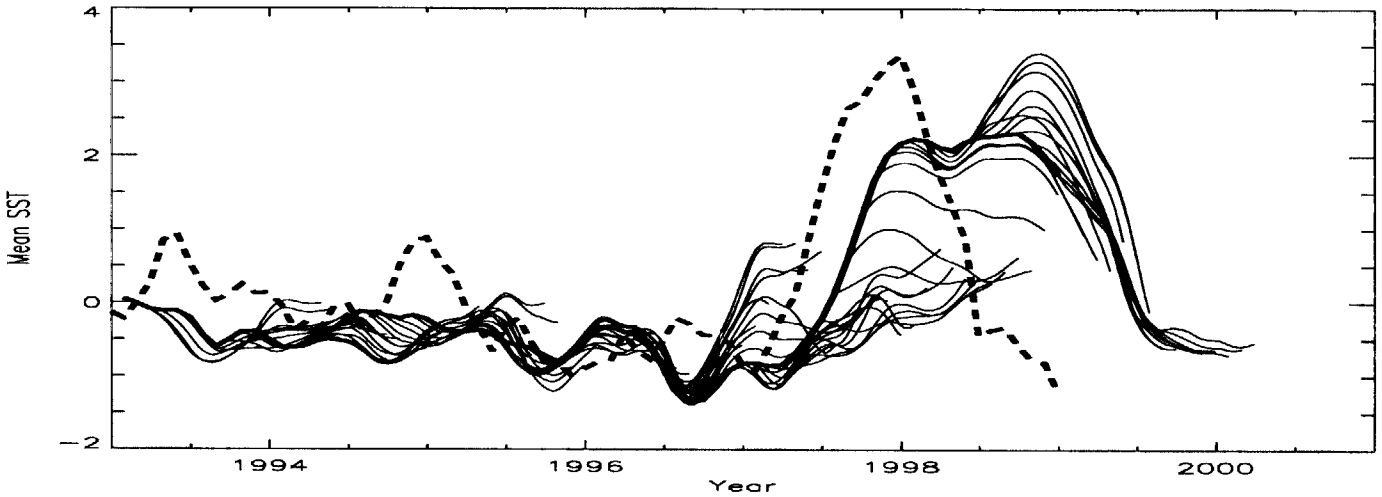


Fig 13.

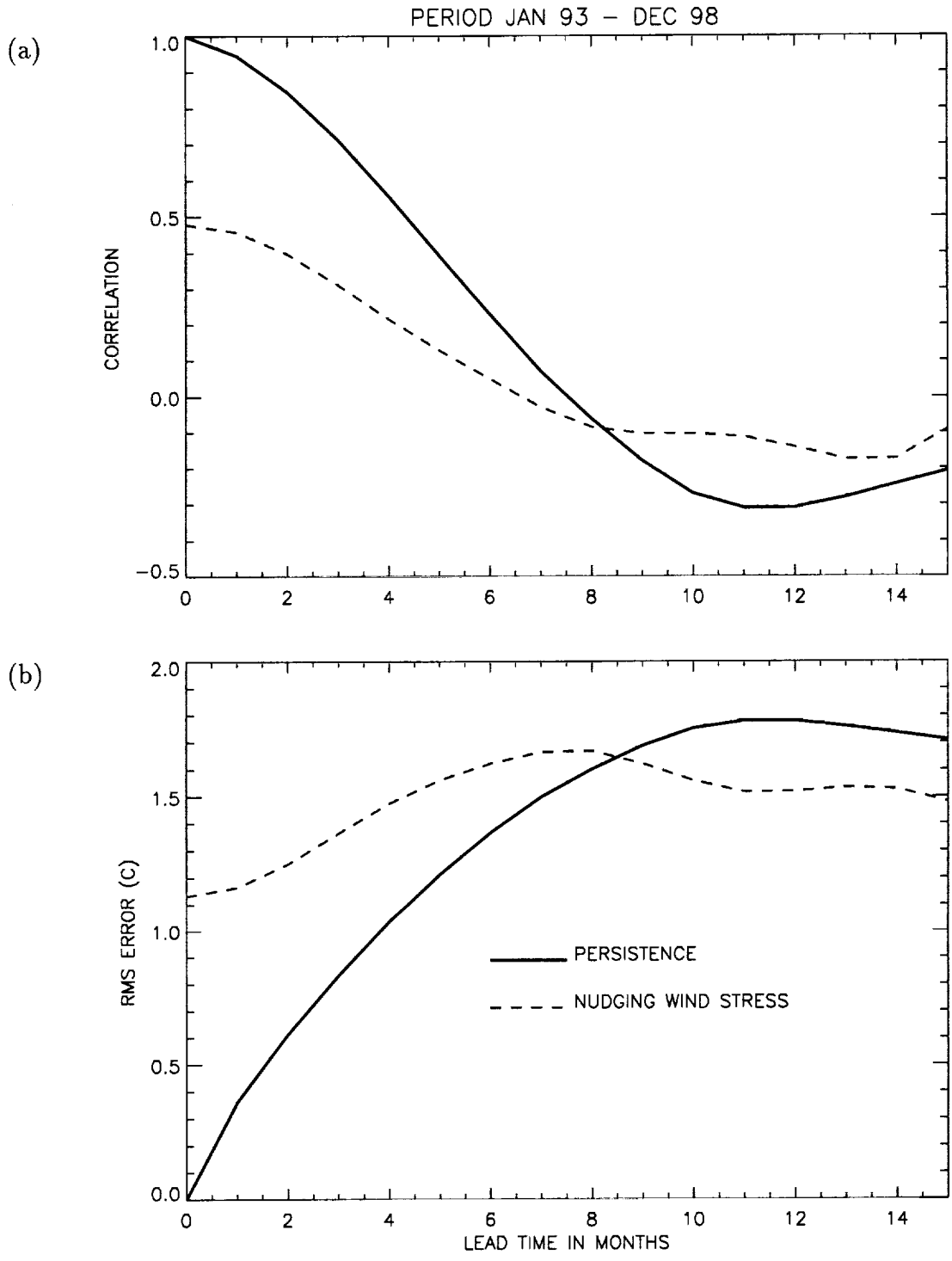
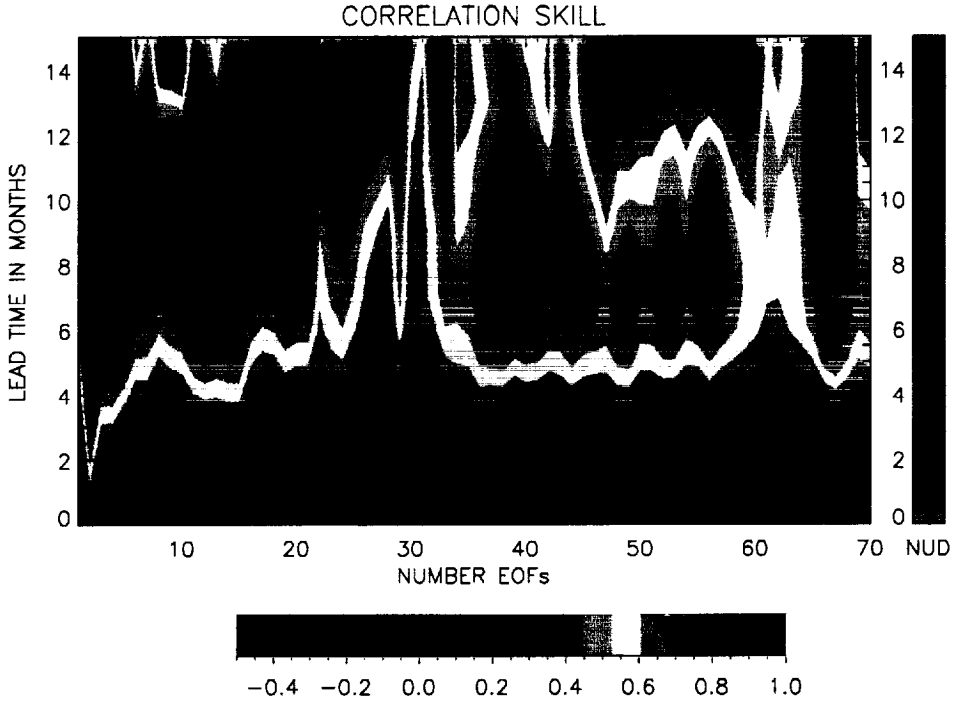

(a)
PERIOD JAN 93 – DEC 98
CORRELATION
1.0
0.5
0.0
−0.5
0
2
4
6
8
10
12
14
(b)
RMS ERROR (C)
2.0
1.5
1.0
0.5
0.0
PERSISTENCE
NUDGING WIND STRESS
LEAD TIME IN MONTHS

Fig 14.

(a)


CORRELATION SKILL
LEAD TIME IN MONTHS
0
2
4
6
8
10
12
14
10
20
30
40
50
60
70
NUD
NUMBER EOFs
−0.4
−0.2
0.0
0.2
0.4
0.6
0.8
1.0

(b)

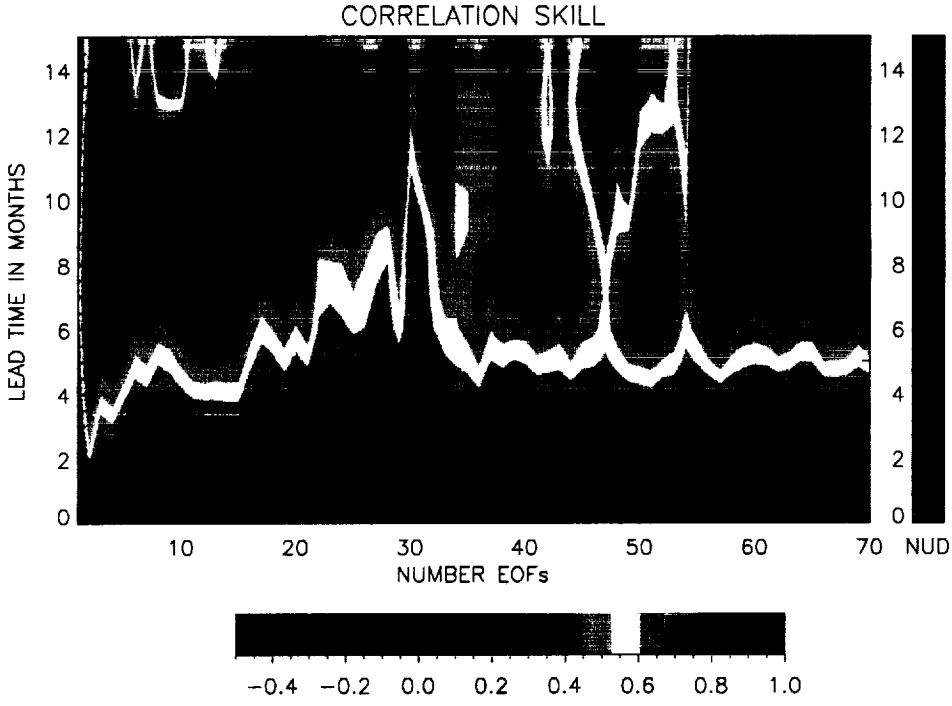
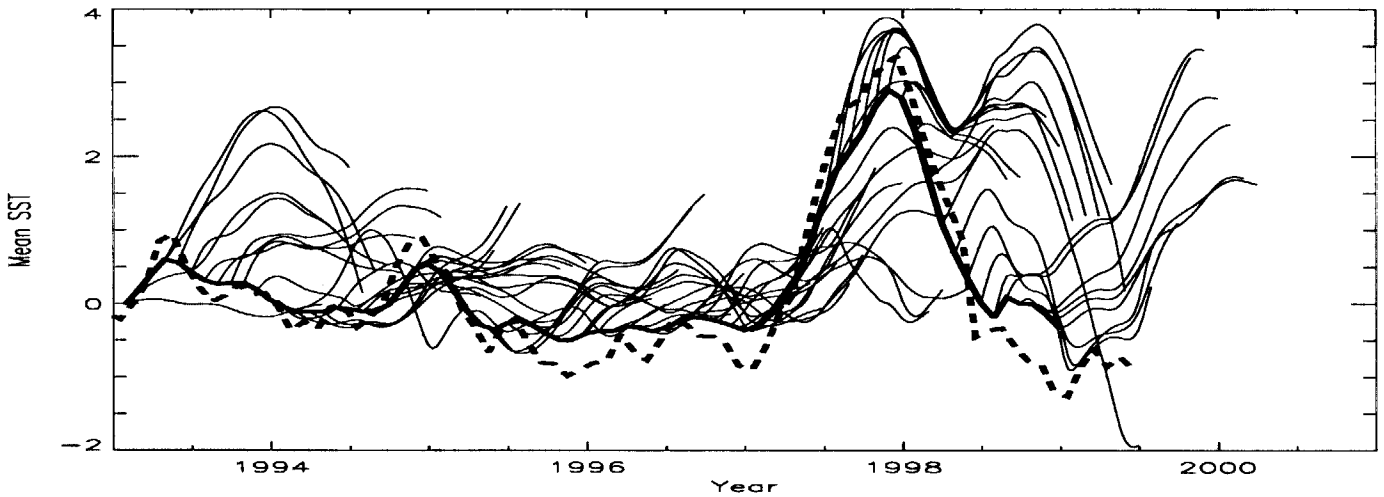

CORRELATION SKILL
LEAD TIME IN MONTHS
0
2
4
6
8
10
12
14
10
20
30
40
50
60
70
NUD
NUMBER EOFs
−0.4
−0.2
0.0
0.2
0.4
0.6
0.8
1.0

Fig 15.

(a)


Mean SST
4
2
0
−2
1994
1996
1998
2000
Year

(b)

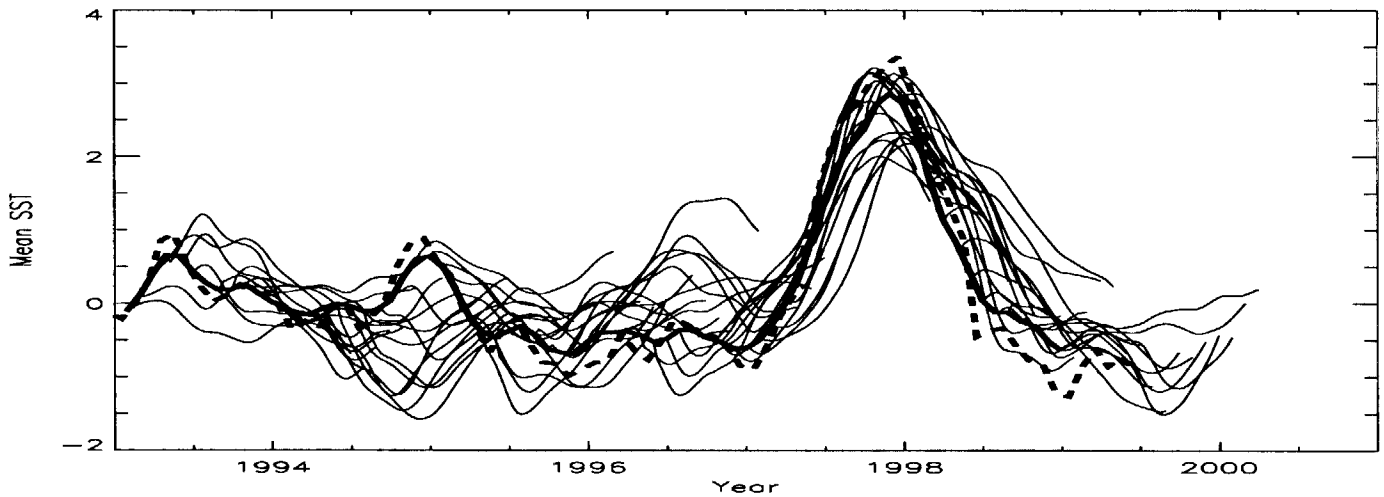

Mean SST
4
2
0
−2
1994
1996
1998
2000
Year

Fig 16.

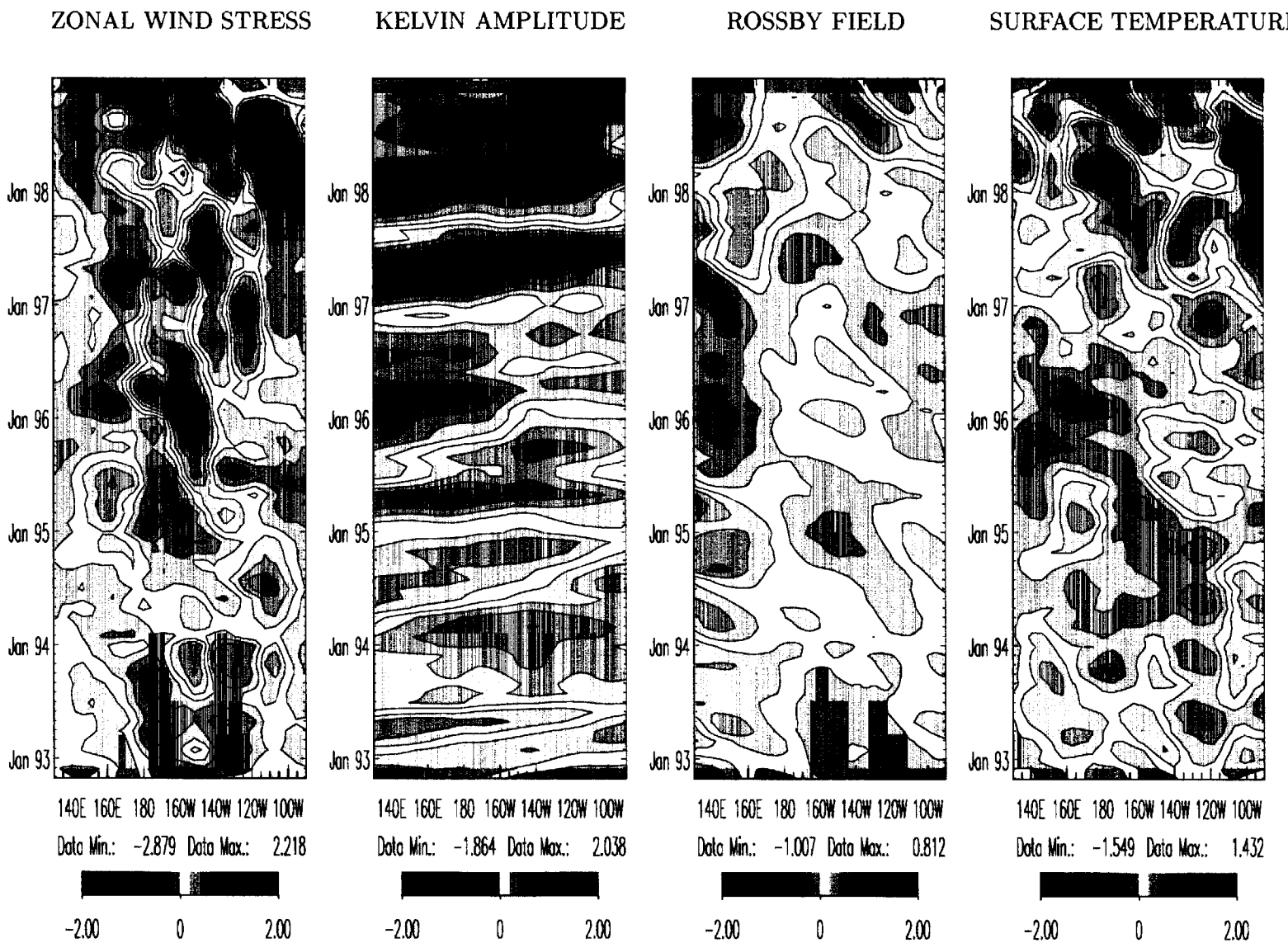
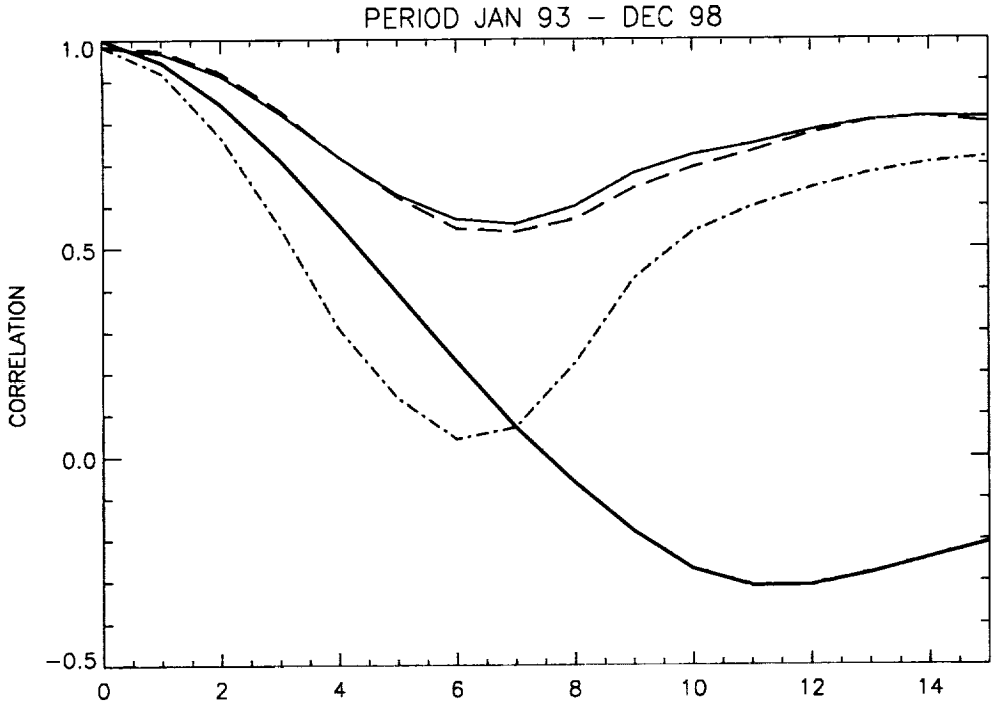

ZONAL WIND STRESS
KELVIN AMPLITUDE
ROSSBY FIELD
SURFACE TEMPERATURE
Jan 98
Jan 97
Jan 96
Jan 95
Jan 94
Jan 93
140E 160E 180 160W 140W 120W 100W
Data Min.: -2.879 Data Max.: 2.218
Data Min.: -1.864 Data Max.: 2.038
Data Min.: -1.007 Data Max.: 0.812
Data Min.: -1.549 Data Max.: 1.432
-2.00 0 2.00

Fig 17.

(a)


PERIOD JAN 93 – DEC 98
CORRELATION
1.0
0.5
0.0
−0.5
0
2
4
6
8
10
12
14

(b)

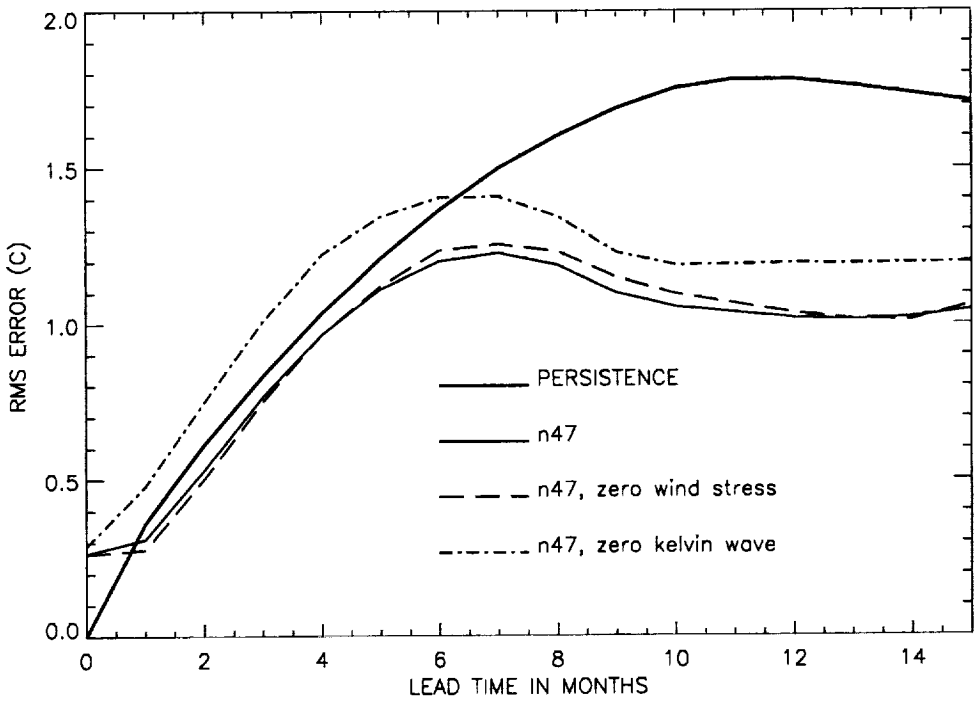
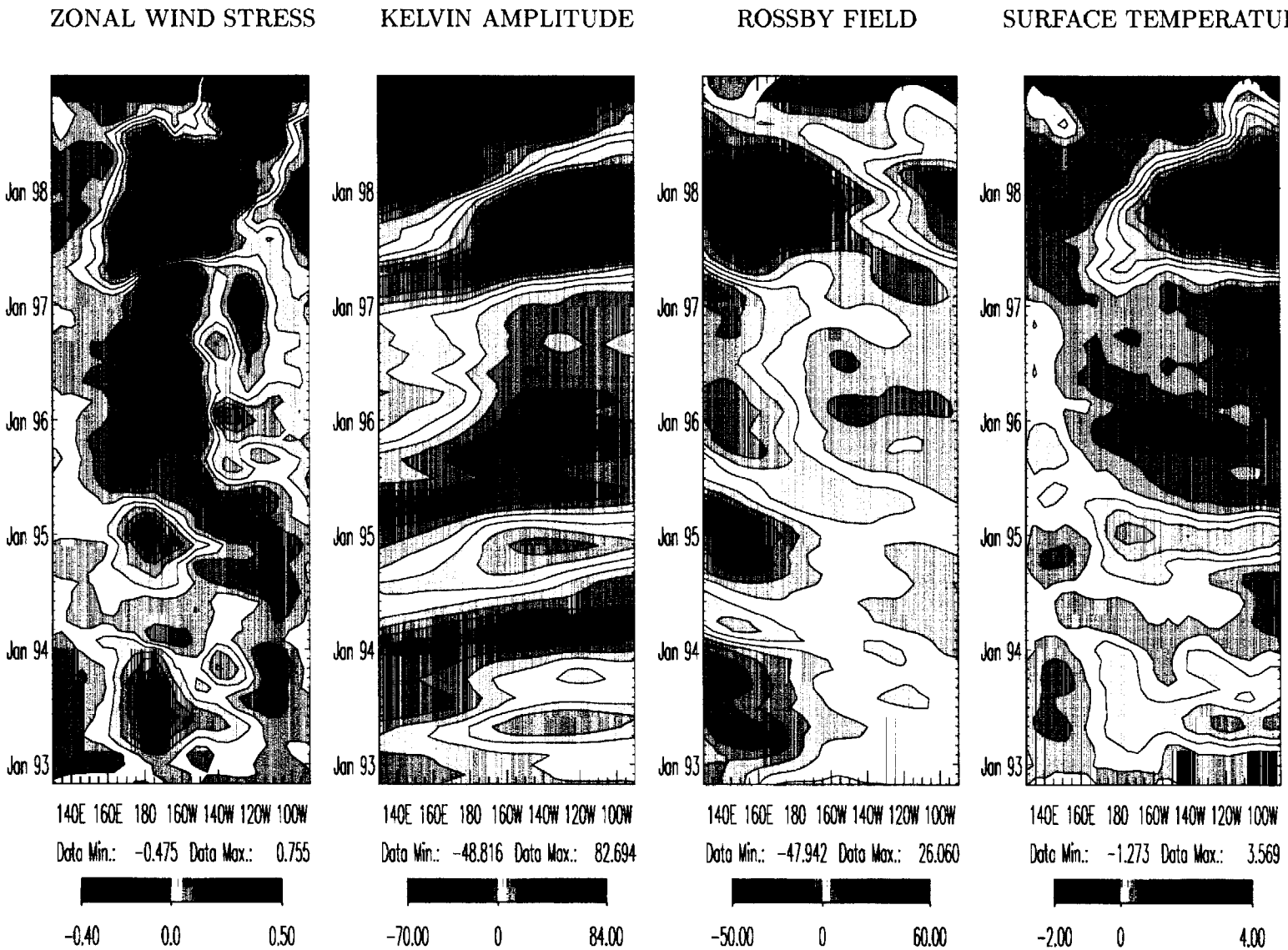

RMS ERROR (C)
2.0
1.5
1.0
0.5
0.0
0
2
4
6
8
10
12
14
LEAD TIME IN MONTHS
PERSISTENCE
n47
n47, zero wind stress
n47, zero kelvin wave

Fig 18.


ZONAL WIND STRESS
KELVIN AMPLITUDE
ROSSBY FIELD
SURFACE TEMPERATURE
Jan 98
Jan 97
Jan 96
Jan 95
Jan 94
Jan 93
140E 160E 180 160W 140W 120W 100W
Data Min.: -0.475 Data Max.: 0.755
-0.40 0.0 0.50
Data Min.: -48.816 Data Max.: 82.694
-70.00 0 84.00
Data Min.: -47.942 Data Max.: 26.060
-50.00 0 60.00
Data Min.: -1.273 Data Max.: 3.569
-2.00 0 4.00