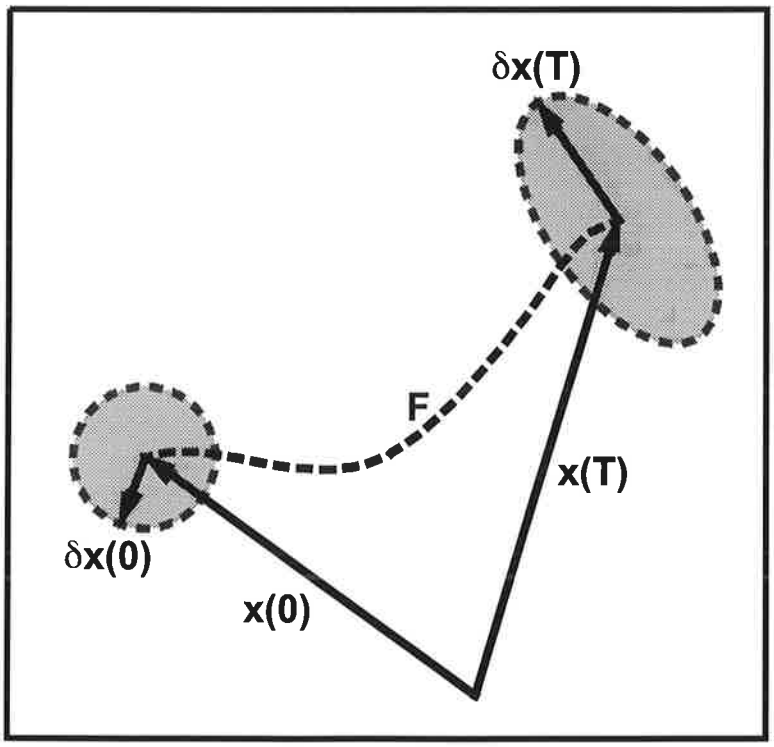
Max-Planck-Institut
für Meteorologie

EXAMENSARBEIT Nr. 55



ON PREDICTABILITY LIMITS OF ENSO

A Study Performed with a Simplified Model of the Tropical Pacific Ocean-Atmosphere System

von

Christian Eckert

HAMBURG, Juli 1998

Dissertation zur Erlangung des Doktorgrades

Autor:

Christian Eckert

Max-Planck-Institut
für Meteorologie

MAX-PLANCK-INSTITUT
FÜR METEOROLOGIE
BUNDESSTRASSE 55
D - 20146 HAMBURG
GERMANY

Tel.: +49-(0)40-4 11 73-0
Telefax: +49-(0)40-4 11 73-298
E-Mail: <name> @ dkrz.de

On Predictability Limits of ENSO

A Study Performed with a Simplified Model of the Tropical Pacific Ocean–Atmosphere System

Dissertation
zur Erlangung des Doktorgrades
der Naturwissenschaften im Fachbereich
Geowissenschaften
der Universität Hamburg

vorgelegt von

Christian Eckert
aus
Freiburg i.Br.

Hamburg

1998

ISSN 0938-5177

Als Dissertation angenommen vom Fachbereich Geowissenschaften der Universität Hamburg

aufgrund der Gutachten von Dr. M. Latif
und Dr. U. Luksch

Hamburg, den 6. Juli 1998

Prof. Dr. G. Miehlich
Dekan
des Fachbereichs Geowissenschaften

Contents

Abstract

Two processes limiting the predictability of the El Niño/Southern Oscillation (ENSO) phenomenon were investigated. First, the perpetual action of fluctuations in wind stress forcing that are not correlated to ENSO itself but that are an integral part of the tropical atmosphere-ocean system was included into a simplified coupled model of ENSO. The implications of this random element for the dynamics and predictability were studied. Second, the growth of small errors in the initial conditions of the coupled model was analysed. This was accomplished by computing its singular vectors and singular values, i.e. the spatial structures and the amplification rates of those initial state perturbations that grow most strongly over a given time interval.

The simplified coupled model of ENSO used consists of an ocean general circulation model coupled to a diagnostic atmosphere model. Following common terminology, such a model is called a Hybrid Coupled Model (HCM). The HCM was designed to simulate the interannual climate variability of the ocean–atmosphere system in the tropical Pacific region. The diagnostic atmosphere exploits the statistical correlation of anomalous sea surface temperature and wind stress present in observations. Via linear regression both quantities are related in a reduced state space of their leading Empirical Orthogonal Functions (EOFs).

To study the effect of random perturbations during the forecast, the coupled model was complemented by a stochastic anomalous wind stress field. This stochastic part was derived from high-pass filtered data. It mimics the observed statistics of the random perturbations. The singular vectors, on the other hand, were derived by generating the linearised numerical code of the HCM and the corresponding adjoint code. This was done with the help of an automatic differentiation tool.

Both above-mentioned processes are important in limiting the predictability of ENSO. They can be understood as parts of a stochastic dynamical system which suffers from the imperfect knowledge of its initial state and the unpredictable components during simulation. It is shown that each ENSO prediction faces a natural limit of predictability depending on the season and the phase of ENSO at its start. I conclude that the limit of ENSO predictability is substantially shorter than the typical ENSO cycle period.

I. Introduction

The El Niño/Southern Oscillation (ENSO) phenomenon is the dominant interannual natural climate fluctuation. It manifests itself, among other things, in a quasi-periodic large scale anomalous warming (El Niño) or cooling (La Niña) of surface waters in the eastern and central equatorial Pacific. Philander (1990) gives a thorough introduction and Neelin et al. (1994) review recent research activities.

Observed indices of ENSO, e.g. as derived from the UKMO Global Ice and Sea Surface Temperature (GISST) data set, Parker et al. (1994), show that ENSO extremes occur irregularly in time and with varying amplitudes. The typical recurrence time of El Niño is 2 to 7 years, and its amplitudes, measured in terms of sea surface temperature anomalies (SSTA), averaged over the eastern equatorial Pacific NINO3 index region (5°N – 5°S and 150°W – 90°W), typically range from 1°C to 2°C. Additionally, ENSO is more or less tightly phase locked to the annual cyle. The degree of phase locking shows marked interdecadal variations as observed by Balmaseda et al. (1995) who performed a statistical analysis of the GISST NINO3 time series. The atmospheric signature of ENSO, the Southern Oscillation, is characterised by large scale surface pressure anomalies (SLPA) of opposite sign in the eastern and western tropical Pacific. The Southern Oscillation Index (SOI) quantifies these deviations from the climatic mean. It is defined as the difference of sealevel pressure anomalies measured at Tahiti (French Polynesia) and Darwin (Australia). The SOI and the NINO3 SSTA time series are highly anti-correlated; Fig.1. To illustrate the observed atmospheric and oceanic characteristics of ENSO further, Fig.2 depicts the leading Empirical Orthogonal Functions (EOFs) of SSTA and of sea level pressure anomalies (SLPA), along with the corresponding prinicipal components, for the last two decades.

The above mentioned characteristics of ENSO indicate that ocean and atmosphere act in concert to maintain a coupled process. Such a theoretical explanation for the ENSO phenomenon was first given by Bjerknes (1966, 1969). According to his still reigning paradigm, ENSO owes its existence to air-sea interaction. The warming of the ocean's surface waters causes the trade winds to weaken which in turn leads to further oceanic adjustment to eventually produce the observed cyclic deviations from the climatic mean. There are different hypotheses about the details of the coupled dynamical process just mentioned; see also Neelin et al. (1994). Presently, the delayed action oscillator scenario, Battisiti and Hirst (1989), Suarez and Schopf (1988), is favoured by many scientists. Local instabilities, equatorial

ocean wave dynamics, and amplitude limiting nonlinear effects are assumed to produce the observed low frequency cycle. To illustrate the dynamics of this scenario, let us assume that the surface waters in the eastern equatorial Pacific are warming. Associated anomalous eastward winds excite two kinds of oceanic wave modes: on the one hand, equatorial Kelvin waves that propagate eastward and that enhance the warming trend, on the other hand, equatorial Rossby waves are exited that propagate to the western boundary of the Pacific. After reflection at the western boundary, these Rossby waves are transformed into Kelvin waves, now carrying an uppwelling signal to the east, i.e. they oppose the warming. Eventually, this delayed response dominates the local warming instability, stops the warming, and a cooling in the eartern parts begins. After this, a cold phase of the ENSO cycle will develop, with the signs of the wind and temperature signals reversed.

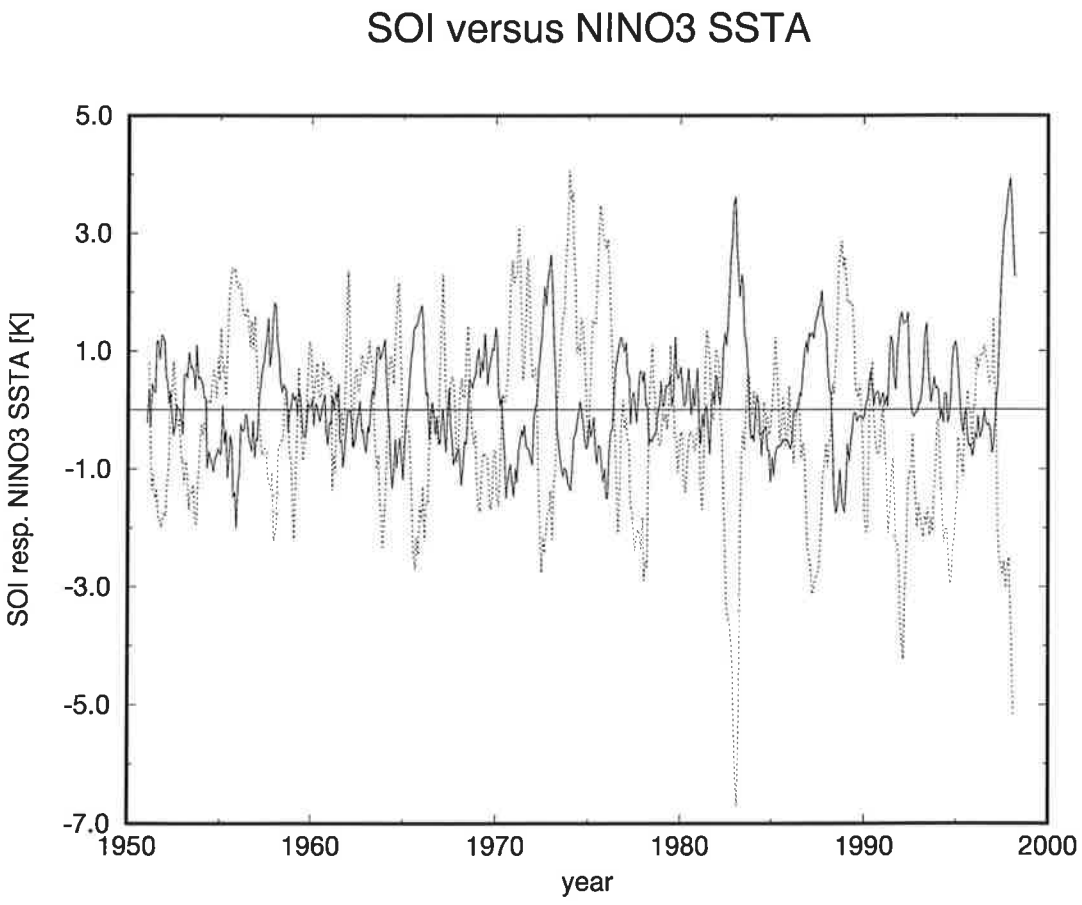


Fig.2: Time evolution of the Southern Oscillation Index (SOI) and of the NINO3 index (5°N – 5°S and 150°W – 90°W) starting in 1951 up to present based on monthly means. SOI: dotted line, NINO3: solid line. The SOI data were smoothed by a 3 month running mean. CPC (1998a).

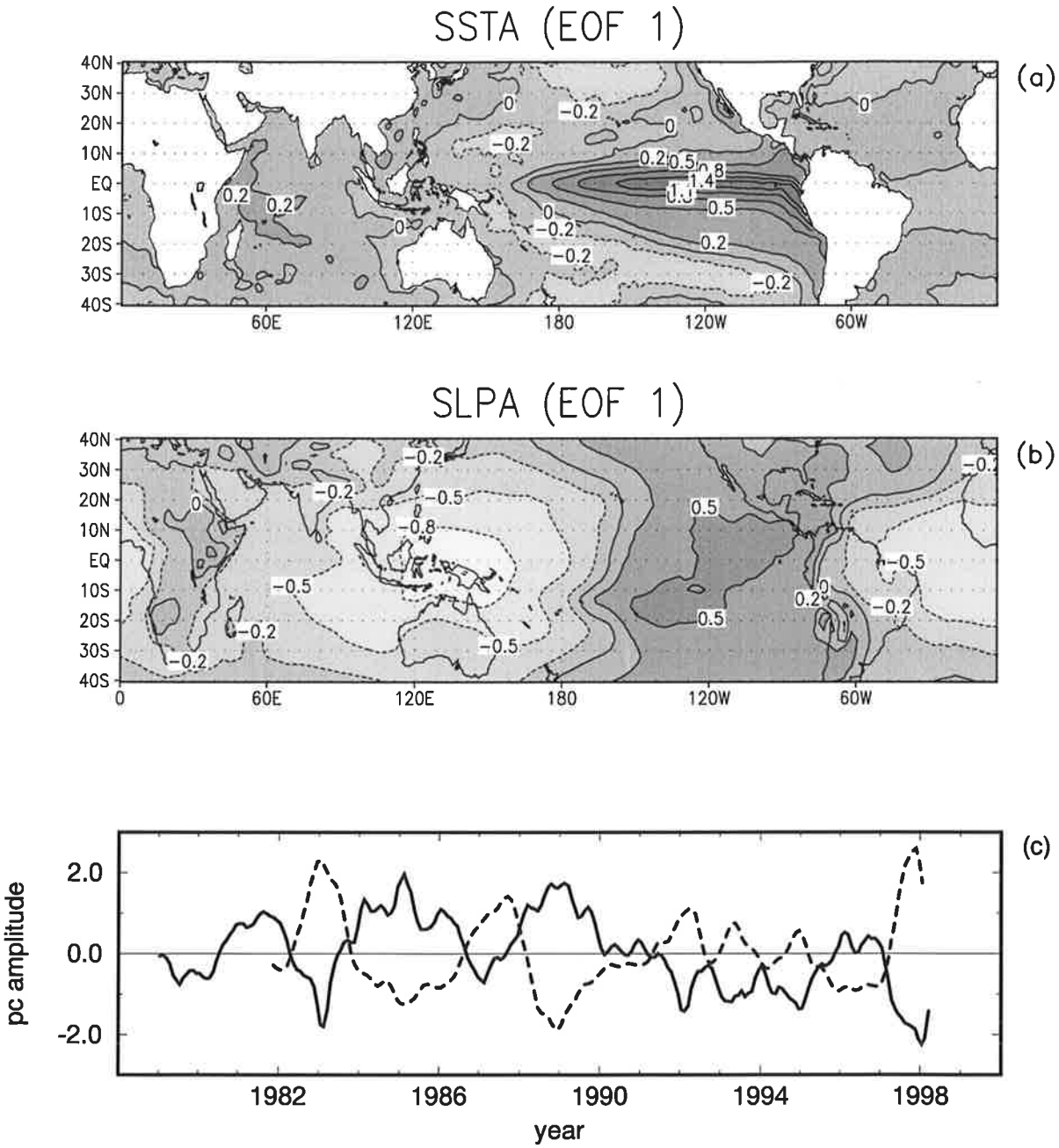


Fig.3: Leading EOFs of SSTA, (a), and SLPA, (b), on a monthly basis and the corresponding principal components, (c). They account for 24.2% (SSTA) and 16.9% (SLPA) of the variance contained in the data sets. The data were taken from the datasets of Reynolds, Reynolds and Smith (1994), for SSTA, and NCEP daily analyses, NCEP (1998), for SLPA. In (c) the thick dashed line corresponds to the SSTA data, the thick solid line to the SLPA data. The time series were filtered by a triangular 5-point moving average.

Among the hypotheses being able to explain the high degree of variability in the data records one may distinguish at least two. The first is linked to the characteristics exhibited by non-linear dynamical systems. Jin et al. (1994) and Tziperman et al. (1994), for instance, explain the irregularity of ENSO by non-linear interactions between the annual cycle and the fundamental ENSO mode. Varying parameters of their simulation models yielded ENSO frequencies that are rational fractions of the annual frequency. Irregularity in turn occurs by the overlapping of frequency locked regimes. Instead of varying internal parameters, Chang et al. (1994) studied the response of their coupled ocean-atmosphere model to different amplitudes of the seasonally varying part of heat flux forcing. Apart from chaotic regimes and transition regimes to chaos, they get three phase locked regimes with 1, 2, and 3 years period. Consequently, changes in the forcing amplitude with time may render ENSO irregular. A second candidate for rendering ENSO irregular, is stochastic forcing. Kleeman and Power (1994) studied the influence of random wind stress perturbations on a coupled model. Forecast ensembles initialised in the 1970s and 1980s showed considerable spread after the start of the individual predictions when random forcing was included. In this thesis, the influence of stochastic wind stress forcing on ENSO is studied as well. A stochastically forced ocean–atmosphere system which mimics the observed statistics of the random wind stress fluctuations was designed. This approach is related to Hasselmann (1976) who pointed out the importance of stochastic forcing for inducing climate variability. Blanke et al. (1997) performed a similar study. Although they use a different approach to get an estimate of the stochastic wind stress forcing, their results aggree well with those presented here and in Eckert and Latif (1997). In both models the sensitivity to the random perturbations introduced is quite large. The models's predictability is reduced to less than an average cycle length.

The research efforts in predicting the state of ENSO indicate that this climate fluctation is predictable for about one year ahead. Barnston et al. (1994) and Latif et al. (1998) give reviews on this topic. The forecast models' successes in predicting specific events are, however, very variable. This leads to the question, which processes limit the predictability of ENSO. Of course, systematic errors of the forecast models are an important source of ignorance. As our understanding improves this source of error should become negligible. Apart from this, there are intrinsic sources of inderterminableness a forecaster has to account for. These are intimately related to the above mentioned causes for the observed irregularity of ENSO.

As a working hypothesis, the assumption that the ENSO phenomenon is representable as a stochastically driven dynamical system is made in this thesis. In

general, such a system can be characterised by a probability density function (pdf). Making a forecast, one seeks to specify a very sharply peaked pdf at the start. The pdf evolves in time during the forecast. It may shrink in some directions and become broader in others. Eventually, i.e. for very long forecast intervals, it asymptotes into the climatological distribution of the dynamical process. It is the task of the forecaster to describe the time evolution of the pdf during the forecast interval. From this description one can infer the time span over which one is willing to favour the time dependent forecast. Afterwards, a forecast made on the basis of the stationary pdf, the climatology, will be as skillful as the time dependent one, with the advantage of being much cheaper. The time span during which the time dependent forecast is the better one, is defined here as the limit of predictability. This limit can be highly variable, depending on the state of the ocean–atmosphere system for the specific forecast time. The general limit of predictability can now be understood as a random variable associated with the dynamical system. At present, prediction models of different degrees of complexity, ranging from purely statistical to fully coupled ocean–atmosphere general circulation models, are used to forecast ENSO. Most of the forecasts issued, especially those derived from dynamical models, are not made in the way described above. Either a single forecast is issued or a small ensemble of forecasts is taken to get some indication of the likely spread (CPC 1998b). The skill of the forecast model is estimated by carrying out hindcast experiments and subsequent comparison with measurements.

The amplification of initial errors can be caused by different mechanisms such as unstable modes on which the error structure projects, Goswami & Shukla (1991), or the non-selfadjoint nature of the linearised system dynamics, Blumenthal (1991), Penland and Sardeshmukh (1995). The latter can yield rapid initial error growth even in asymptotically stable linear dynamical systems. The knowledge of the error structures that yield largest growth over the forecast time can be exploited in several ways. For example, one may use these error structures to define those geographical regions in which measurements should be improved to yield better forecasts. Perturbations of this kind can also improve the estimate of the likely spread of the forecast. Several authors investigated initial error growth in their ENSO prediction models. The methods to obtain an estimate of error growth differ and the relevant structures determined differ as well. The simplest possible approach that is only feasible for relatively low-dimensional models is to compute finite difference approximations to the tangent linear propagator. The tangent linear propagator is the linear mapping that maps an initial state perturbation of the system linearly onto the corresponding final state perturbation. This was done by Chen et al. (1997). Next, one may project the simulation model onto dominant

modes of variability of the model, e.g. the leading EOFs, and subsequently fit linear models in the reduced state space. Such studies were performed by Blumenthal (1991) and Xue et al. (1994). Another study of Xue (1997) estimated the tangent linear propagator of her forecast model by a method originally proposed by Lorenz (1965), involving a reduction of the model's state space by projection onto EOFs as well. Moore and Kleeman (1996) were the first to compute those initial state perturbations that grow most strongly over the forecast time, the singular vectors. To do this a singular vector analysis of the linearised forecast model has to be performed. The latter was obtained by linearsing the model equations which were subsequently coded. In a series of publications Moore and Kleeman discuss the dynamics of error growth in this model and various sensitivity studies; see Moore and Kleeman (1997a) and Moore and Kleeman (1997b).

In this study, a Hybrid Coupled Model (HCM) is used as a simulation model for ENSO. Such an HCM consists of a primitive equation ocean model coupled to a greatly simplified atmosphere component. The atmosphere is paramterised by a regression of SSTA and wind stress anomalies. The regression matrix was derived from observations. The influence of random atmospheric fluctuations is modeled by extending the HCM to a stochastically forced HCM. The observed statistics of wind stress fluctuations not correlated to ENSO is simulated. To study the influence of initial state perturbations, the HCM's linearisation is determined directly from the numerical code. This is accomplished by the use of an automatic differentiation tool developed by Giering (1997), the TAMC (Tangent Linear and Adjoint Model Compiler). The determination of the modes that exhibit strongest growth over the forecast interval is done with the help of an iterative algorithm which only requires the specification of the action of a symmetric linear operator on a given input vector. This property is well suited to study dynamical systems with a large number of independent variables, such as the HCM.

The contents of this thesis is organised as follows. In chapter II the Hybrid Coupled Model is described in some detail. The effect of random wind stress perturbations on the model's predictability is investigated in chapter III . Chapter IV deals with the time evolution of perturbations in the model's initial conditions and their effect on the outcome of a forecast. The most strongly growing initial perturbations are determined by singular vector analysis. Finally, chapter V contains a summary and a discussion together with some remarks on future work.

II. ENSO and the Hybrid Coupled Approach

1. Introduction to Hybrid Coupled Models

Seemingly, Neelin (1989, 1990) was the first to introduce the name "Hybrid Coupled Model" for the concept we are going to describe in this chapter, namely, to model the coupled ocean-atmosphere system in the tropics by using an elaborate ocean model coupled to a greatly simplified atmospheric component.

The basic assumption underlying this approach is that the atmosphere is in statistical equilibrium with the forcing sea surface temperature on the time scales considered, i.e. the adjustment of the atmosphere to a given ocean state is very much faster than the typical time scales of ocean dynamics.

The model used here was developed by M. Flügel (1994), who applied it to obtain forecasts of the state of ENSO from 1979 to 1988. The atmosphere consists of a linear regression of observed anomalies of SST and wind stress. To retain the spatial correlations of the respective anomalies, the regression was done in EOF space. The model's forecast skill is comparable to other state of the art forecast models. Since the model details are important later on, the setup is described at some length in the following sections.

Barnett et al. (1993) used the same ocean model component in an anologous HCM experiment. The diagnostic atmosphere, however, was obtained in a slightly different manner. The regression was done using the anomaly fields of individual months. Changes in the response characteristics of the atmosphere during the annual march can thereby be incorporated into the linear relationship. As a consequence the model variability is increased. This HCM visited different dynamical regimes during a 150yr integration. Starting with an irregular behaviour, around year 30 it got locked into a two year cycle and finally changed to a 3 year cycle between years 90 and 100. The transitions to different dynamical regimes were be ascribed to a slow warming trend in the oceanic temperature field by the investigators.

Recently, van Oldenborgh et al. (1997) used a Hybrid Coupled Model to study the oceanic dynamical features during the years 1987 and 1988. They focus on the sensitivity of the NINO3 SST index to itself at earlier times. This was done by using the adjoint of the ocean general circulation model. The adjoint of the statistical atmosphere model can be used to relate the NINO3 index to itself. The results obtained are consistent with the delayed action oscillator scenario described in the introduction. In chapter IV a description of adjoint models will be given.

2. The Diagnostic Atmosphere

As already mentioned in the previous section, the diagnostic atmosphere module consists of a regression of SSTA and anomalous wind stress over the domain of the ocean model. The regression is done in EOF space, retaining the five leading EOFs of the respective anomaly fields. This approach takes account of the nonlocal and large scale atmospheric response to the observed anomalous patterns of SSTA. The Florida State University (FSU) wind stress data, Goldenberg and O'Brien (1981) and the SST dataset of Reynolds (1988) were taken to derive the regression matrix.

Formally, the regresssion matrix is obtained as follows. After EOF decomposition of the observed fields of anomalous SST and wind stress, which we denote by $T_{obs}(x,t)$ and $\tau_{obs}(x,t)$, respectively, we may represent the latter by

$$T_{obs}(x,t) = \sum_n \alpha_n(t) e_n(x) \quad \text{and} \tag{1}$$

$$\tau_{obs}(x,t) = \sum_m \beta_m(t) f_m(x) \quad . \tag{2}$$

The numbers

$$C^{atm}_{mn} = \frac{\langle \beta_m \alpha_n \rangle}{\langle \alpha_n^2 \rangle} \tag{3}$$

constitute the entries of the regression matrix C^{atm}. Angle brackets indicate time averages. The matrix C^{atm} solves the minimisation problem

$$\langle \, || \vec{\beta} - C^{atm} \vec{\alpha} ||^2 \, \rangle = \min \quad . \tag{4}$$

Here, the norm $||.||$ is based on the euclidean scalar product. For a given pattern of SSTA the anomalous atmospheric response can now be estimated by

$$\tau_{est}(x,t) = \sum_m \beta_{est,m}(t) f_m(x) \tag{5}$$

with β_{est} being calculated as

$$\beta_{est,m}(t) = \sum_n C^{atm}_{mn} \alpha_n(t) \quad . \tag{6}$$

On the interannual time scale, the simulated wind stress anomalies agree well with those observed. Flügel (1994) correlated both quantities obtaining correlations of about 0.8 in the central equatorial Pacific. Before these correlations were

computed, the observed wind stress anomalies were slightly smoothed by applying a 5-month running mean.

3. The Ocean Model

The oceanic part of the HCM is an improved version of the model used by Latif (1987). It is based on the primitive equations with some simplifying approximations. The hydrostatic and Boussinesq approximations are applied and the model is formulated on the equatorial β-plane. The model domain extends from 130°E to 70°W zonally and from 30°S to 30°N meridionally. The northern and southern boundaries are implemented as solid walls, whereas the eastern and western boundaries mimic the coast lines. The domain is flat-bottomed at a depth of 4000m. Prognostic variables are the fields of horizontal velocity, temperature, and sea surface elevation. Numerically, the model is formulated on an Arakawa E-grid. The basic equations are split into a vertically averaged part, the barotropic system, and the remainder, the baroclinic system. The barotropic system is solved implicitly, whereas the baroclinic system is treated by an explicit scheme. This kind of decomposition allows a time step of 2.25 hours. The zonal resolution in the constituting rectangular subgrids of the staggered E-grid is 6 degrees of longitude, and the meridional resolution amounts to 50km right at the equator decreasing to higher latitudes where it is 400km at the northern- and southernmost points. In the vertical, 13 irregularly spaced levels are used, with 10 of them in the upper 300m. The uppermost layer has a thickness of 20m. Vertical mixing of momentum and heat is parametrised according to Pacanowski and Philander (1981). The horizontal eddy viscosity is assumed to be constant over the model domain, and explicit horizontal diffusion of heat is not taken into account. In the course of each model time step a convective adjustment is applied to avoid an unstable stratification. It consists of one sweep through the water column, mixing two adjacent boxes whenever unstable stratification occurrs. Salinity variations are not taken into account. Therefore, the equation of state density is linearly related to temperature. Thus, the governing equations for the ocean's interior are:

$$\frac{\partial}{\partial t}\vec{u} + (\vec{v}\cdot\nabla)\vec{u} = -\nabla p - \beta y\vec{k}\times\vec{u} + \frac{\partial}{\partial z}(A_v\frac{\partial}{\partial z}\vec{u}) + A_h\nabla^2\vec{u} \tag{7}$$

$$\frac{\partial}{\partial t}T + (\vec{v}\cdot\nabla)T = \frac{\partial}{\partial z}(k_v\frac{\partial}{\partial z}T) \tag{8}$$

$$\nabla\cdot\vec{v} = 0 \tag{9a}$$

$$\frac{\partial}{\partial t}\eta(t) = -\nabla_h \cdot \int_{-H}^{0} \vec{u} dz \tag{9b}$$

$$\rho(T) = \rho_o(1 - \alpha T) \tag{10}$$

Here, $\vec{u}$ and $\vec{v}$ denote the horizontal and total velocity, respectively. Vertical velocities are determined from the equation of continuity; Eq.(9a). The vector $\vec{k}$ is a unit vertical vector and y the meridional coordinate. A_v is the Richardson number dependent vertical eddy viscosity, A_h the horizontal eddy viscosity, and k_v the Richardson number dependent vertical diffusion coefficient for heat. Temperature and sea level are abbreviated by T and η, respectively. The reference density is signified by ρ_o, and the thermal expansion coefficient of sea water by α. ∇_h denotes the horizontal gradient operator $(\frac{\partial}{\partial x}, \frac{\partial}{\partial y})$.

The above equations are complemented by appropriate boundary conditions. At the surface, the model is forced by annual cycles of fluxes of momentum and heat. The thermal forcing enters as a relaxation to an equivalent climatological temperature with a time constant of about 30 days. This procedure is along the lines of Haney (1971). The flux of momentum due to the wind stress is paramterised by the drag law $\vec{\tau} \sim \vec{v}_{wind}|\vec{v}_{wind}|$. At the bounding walls the velocities vanish. The actual model configuration posesses 36694 prognostic variables.

Forcing the ocean model with observed anomalies of wind stress reveals systematic deviations of the simulated SSTA compared to the observed. For this reason the simulated SSTA is not directly used as input for the statistical atmosphere. Instead, a correction adjusts the systematic model errors. Once again, the leading 5 EOFs of the model simulation are related to those derived from the observed anomalies via linear regression. In Eq.(6) the coefficients α_n have to be replaced by their estimators, $\alpha_{int,n}$, calculated from this model output correcting interface:

$$\alpha_{int,n}(t) = \sum_l C_{nl}^{int} \alpha_l(t) \quad . \tag{11}$$

As a reference climatology for computing anomalies which serve as input for the model correcting interface and the statistical atmosphere an integration forced by the FSU wind stress data from 1967 to 1985 was taken.

4. Interannual Variability as Simulated by the HCM

The atmospheric model response was magnified by a factor of 1.4 to obtain a self-sustained oscillation of the coupled model. This additional scaling factor is termed the coupling strength. At least two reasons motivate the necessity of such a scaling. On the one hand, the regression ansatz reduces the variance of the atmospheric response, on the other hand the ocean model has a relatively weak dynamical response to an imposed wind stress forcing due to its rather diffuse thermocline. The specific value was chosen to run the model in a self-sustained regime. Coupling is done on a monthly basis with linear interpolation in between.

The coupled model exhibits a very regular cycle with a period of approximately 60 months. Fig.3 shows a 120yr control intergration of the HCM. Apart from the strictly periodic simulation of warmings and coolings in the NINO3 index region, the amplitudes of the model ENSO extremes are slightly underestimated compared to observations. The dynamics of the modeled El Niño and La Niña events, however, resembles the observed. This can be seen by applying a Principal Oscillation Pattern (POP) analysis to the combined fields of SSTA, sea level anomalies, and wind stress anomalies. Performing a POP analysis consists of fitting a multivariate first order Markov process to the given dataset, minimising the one time step prediction error. The normal modes of the process serve as an alternative basis to describe the data. They are characterised by an e-folding time and an oscillation period in case they correspond to a complex eigenvalue of the system matrix, or an e-folding time only if the eigenvalue is real. The technique was proposed by Hasselmann (1988). A review along with applications and possible refinements can be found in von Storch et al. (1995). The leading POP mode of the control integration has an oscillation period of 62 months, an e-folding time of 450 months, and accounts for 83% of the variance in the combined data set. This result indicates that most of the information contained in the dataset is well modeled by one oscillatory mode.

Fig.4 shows a longitude time section along the equator of the control integration projected onto the leading POP mode. The anomalous SSTs and wind stress anomalies reveal a standing pattern, whereas the sea level anomalies propagate slowly from west to east. This simulation is in accord with the delayed action oscillator scenario, described in the introduction. Performing the same analysis with observations of the respective quantities reveals similarities as well as systematic deficiences of the model. In Fig.5 a POP analysis of the obervations is shown. The National Centers for Environmental Prediction (NCEP) reanalysis data for SSTA

and sea level anomalies, Ji et al. (1995), were combined with the FSU dataset for wind stress anomalies for the period February 1982–December 1992. Once again SSTA and wind stress anomalies are characterised by standing patterns, whereas the sea level anomalies propagate eastward. The modeled SSTA is displaced to the west and the observed time dependence is much more variable. The leading POP mode of this dataset has an e-folding time of 16 months, an oscillation period of 52 months, and accounts for 15 % of the variance.

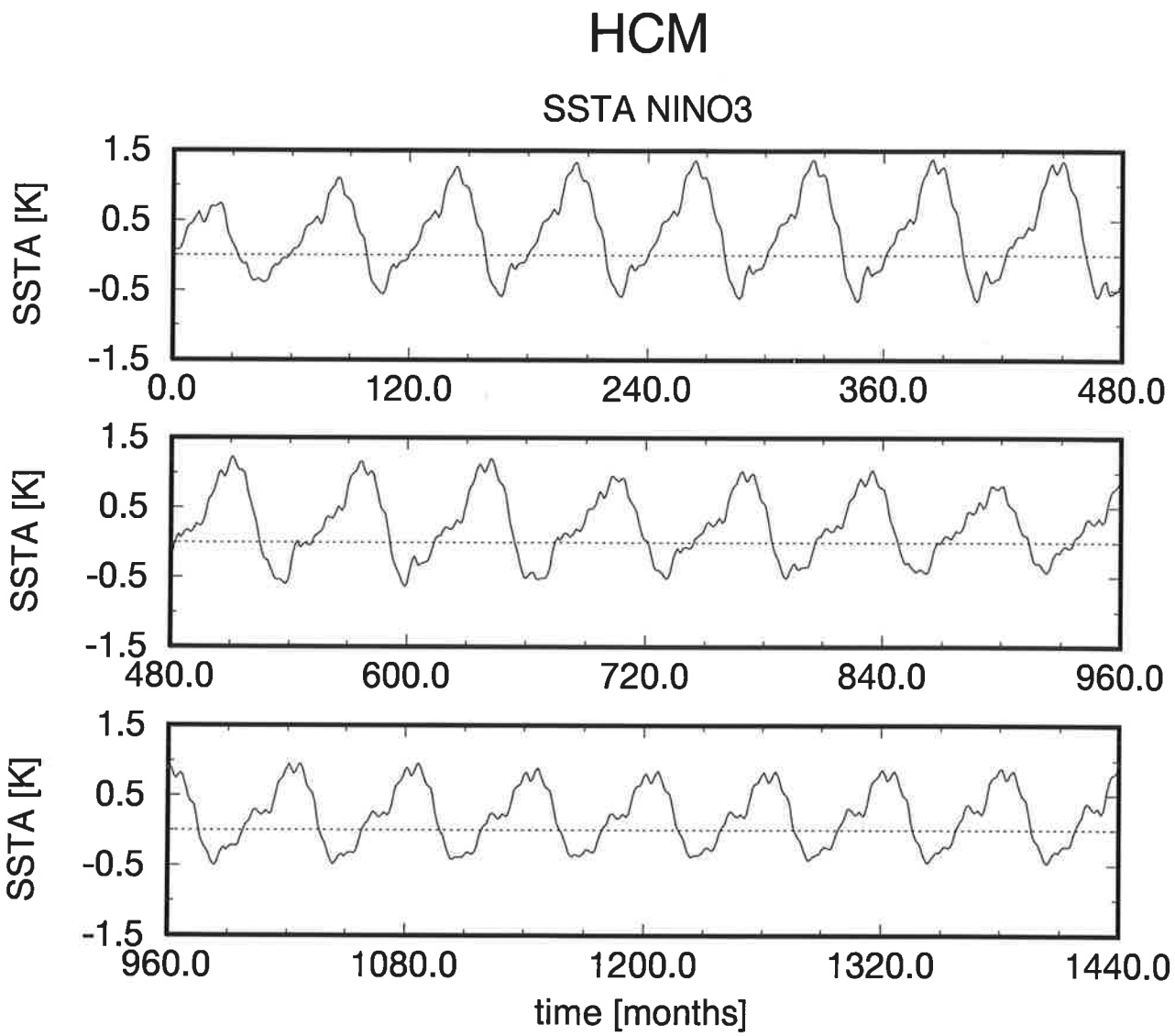


Fig.3: Time series of anomalous sea surface temperature (SSTA), averaged over the NINO3 index region (150° – 90°W, 5°S - 5°N), obtained by integrating the Hybrid Coupled Model (HCM) for 120 yr.

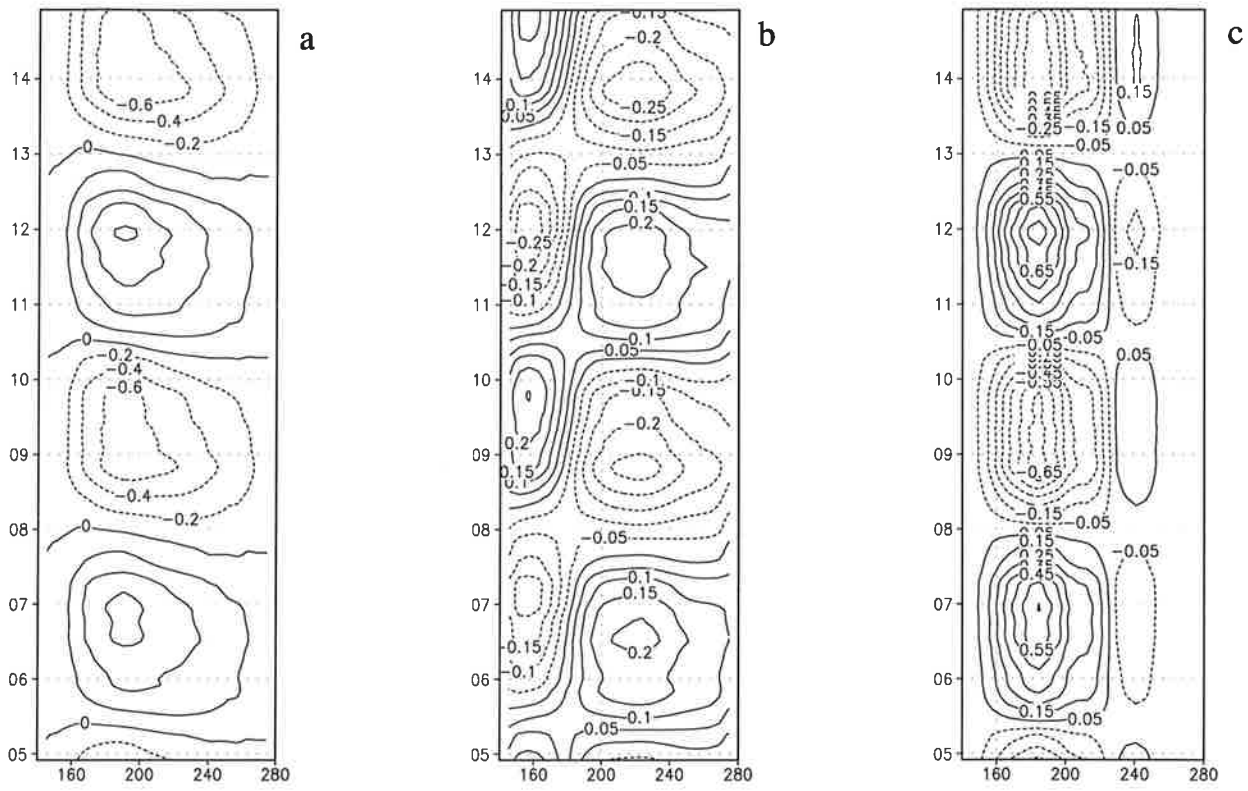


Fig.4: Longitude time section along the equator from year 5 to 15 during the 120-yr integration of the HCM. The combined scaled fields of SSTA, anomalous sea level, and anomalous zonal wind stress were subjected to a POP analysis. Shown are the projections of the individual fields onto the dominant POP mode. This leading POP mode has a period of 62 months, an e-folding time of 450 months, and accounts for 83% of the variance of the dataset. To get physical units, the depicted fields have to be multiplied by factors of (a) 3.46 K, (b) 36.6 cm, and (c) 7.3 $\times 10^{-2}$ Pa, respectively.

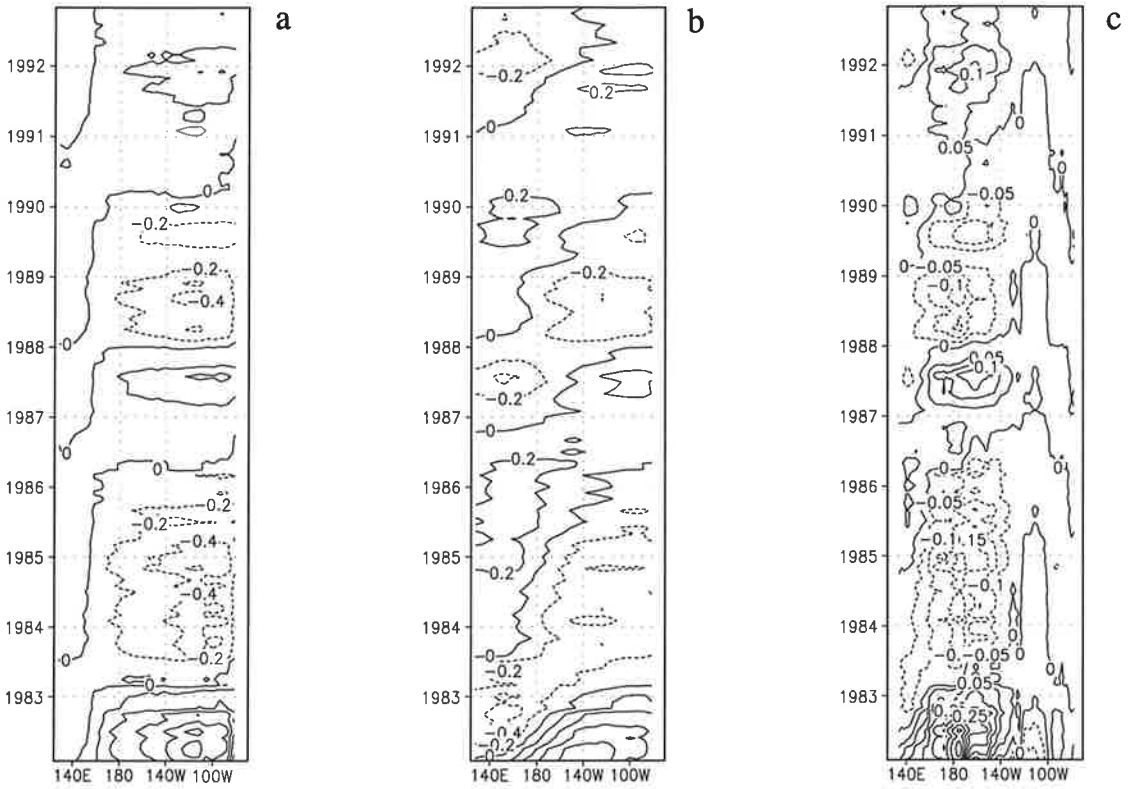


Fig.5: Longitude time section analogous to Fig. 4 . In this case, the NCEP reanalysis data for SSTA and anomalous sea level were combined with the zonal wind stress anomalies of the FSU dataset. The individual fields were scaled with the factors given below before they were pooled. The dominant POP mode has a period of 52 months, an e-folding time of 16 months, and accounts for 15% of the variance of the data set. The individual fields have to be multiplied by factors of (a) 7.68 K, (b) 51.5 cm, and (c) 2.5 $\times 10^{-1}$ Pa to get physical units.

III. Limitation of Predictability due to Random Forcing

1. The Stochastically Forced HCM

The FSU dataset, from which the data to derive the statistical atmosphere were taken, exhibits a high degree of variability not correlated to the ENSO phenomenon. To incorporate this kind of variability into the model, the wind stress anomaly data were high-pass filtered to remove the ENSO time scales. Variability with time scales of less than 12 months was retained fully and to reduce end effects, the filter weights were gradually reduced to zero at 16 months period. As already noted by other investigators, see Kleeman & Power (1994), such high-pass filtered wind stress data are consistent with a white noise forcing with respect to time. The leading EOFs of the high-pass filtered wind stress data exhibit large-scale structures which are likely to affect the interannual variability of the HCM, see Fig.6 which depicts the first and the second EOF. To retain these spatial structures, the following approach was used to model the random wind stress fluctuations. The leading 10 EOFs were linearly combined with the expansion coefficients taken from a white noise process having the appropriate variance. The remainder is simulated as being white temporally and spatially. Its variance amounts to that contribution which is not accounted for by the leading 10 EOFs. Each time the statistical atmosphere is called to produce a wind stress anomaly an additional random wind stress fluctuation is generated. Their sum constitutes the new wind stress anomaly by which the ocean model is forced.

This extension of the HCM will be called "Stochastically Forced HCM" further on. Fig.7 displays a schematic of the whole model. Another 120yr integration was performed. As expected, a much higher degree of variability is observed. The model ENSO extremes now occur irregularly with varying amplitudes, see Fig.8 which shows the NINO3 index during this 120yr integration of the extended model. Sometimes the underlying cyclic behaviour is damped, sometimes it is reinforced by the random signal. A further POP analysis of the combined fields of SSTA, wind stress anomalies, and sea level anomalies, shows that the basic ENSO dynamics is retained, the e-folding time of the leading mode is 22.6 months, its oscillation period 61.9 months, and it accounts for 32.4% of the variance in the data set. The stochastically forced model's simulation is in closer aggreement with observations. This can as well be seen by comparing the spectra of the NINO3 time series of the control integration and the integration with noise added to the GISST dataset which extends from 1900 to 1992. Fig.9 shows such a comparison of the maximum entropy spectra calculated for the three time series. The spectral peaks

of the HCM are broadened when the random wind stress component is added to the forcing. On decadal and larger time scales the variability of the stachastically forced HCM is still less than observed. This may be explained by the lack of interaction with the large-scale ocean circulation. Since the focus of this study are limitations of predictability on seasonal to interannual time scales this should not affect the results obtained here.

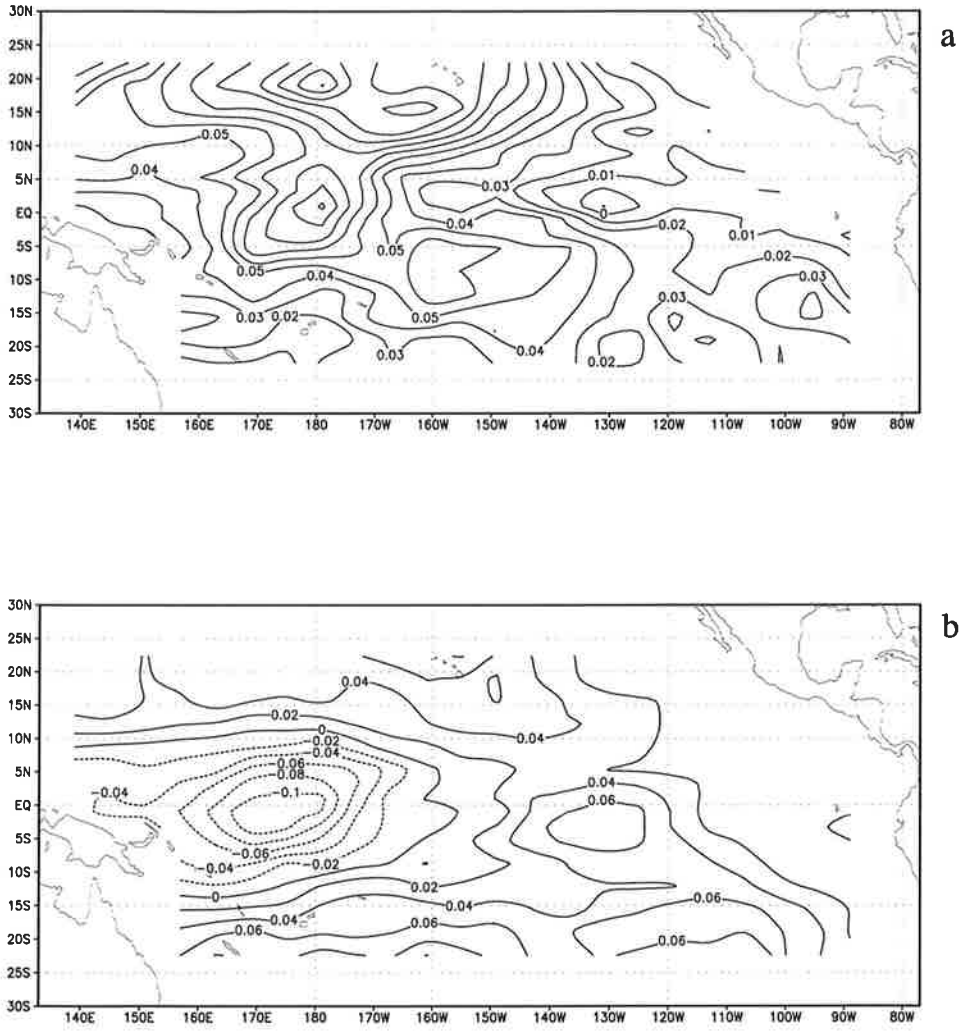


Fig.6: The first (a) and the second (b) EOF in units of Pa of the high-pass filtered wind stress anomalies of the FSU dataset. They explain 10.3% and 7% of the variance, respectively.

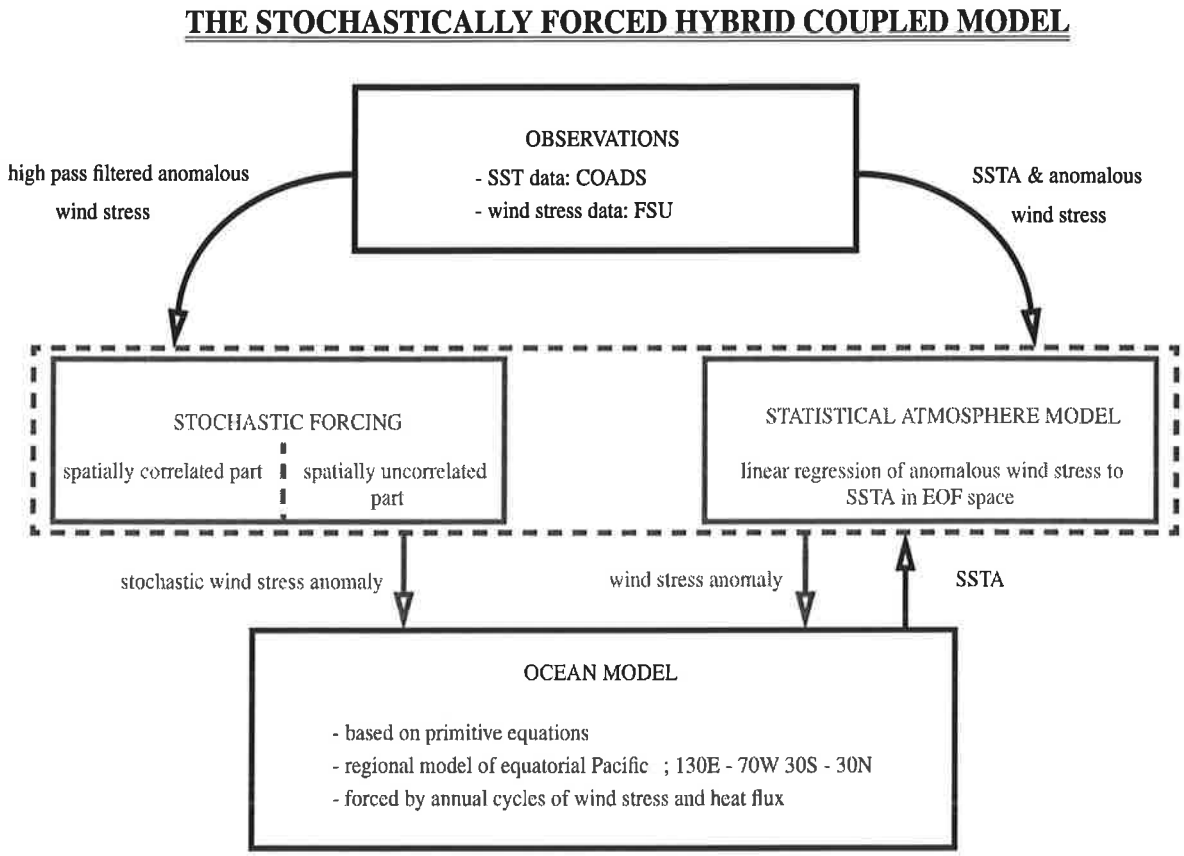


Fig.7: Schematic picture of the stochastically forced HCM that was used in this predictability study. The additional stochastic component in the wind stress forcing of the ocean model consists of two parts, a spatially correlated one and a spatially uncorrelated one. For further details see the text.

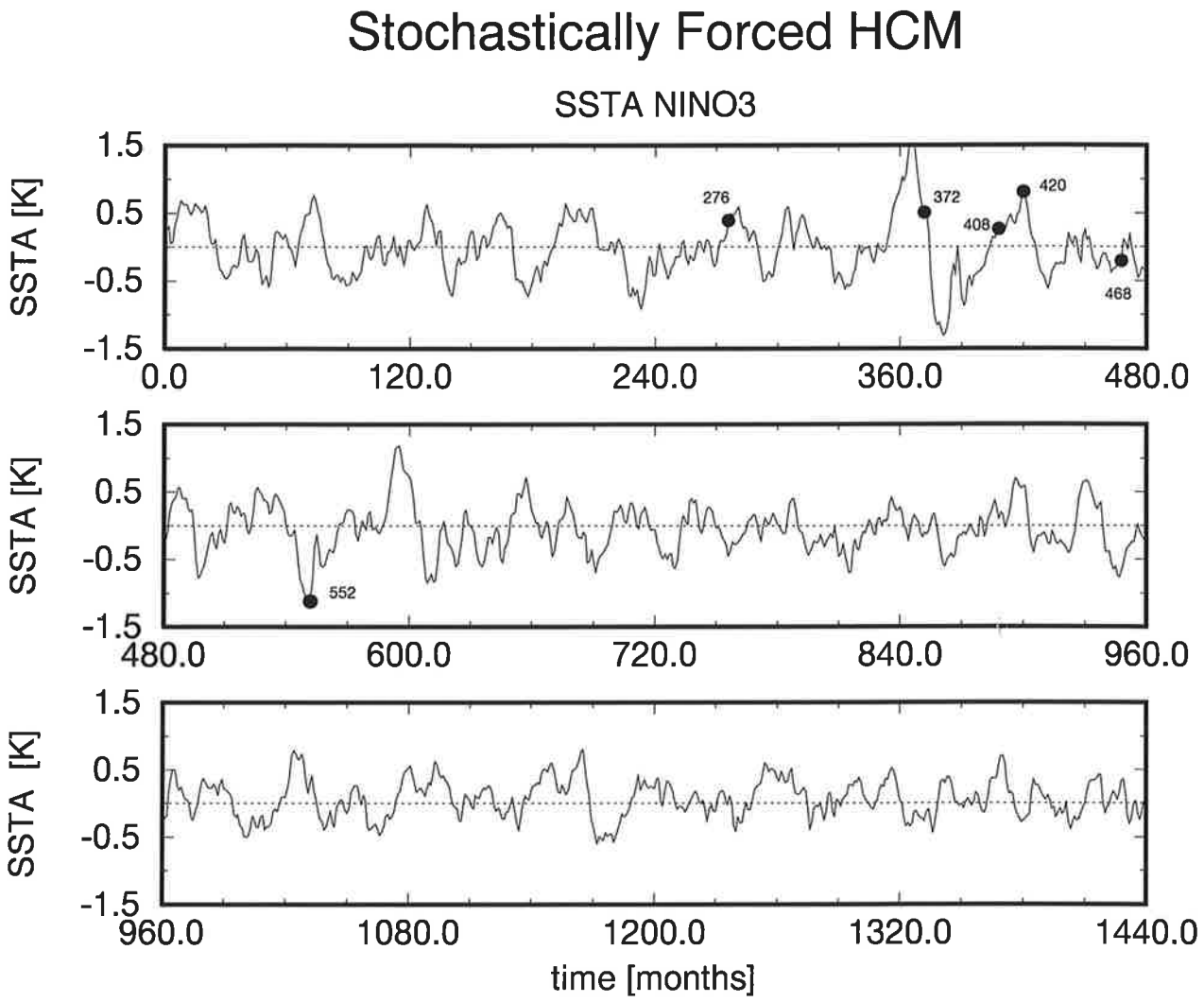


Fig.8: Time series of SSTA averaged over the NINO3 index region obtained by integrating the stochastically forced HCM for 120 years. The labeled dots mark the restart dates that were used to generate ensembles of trajectories with different realisations of the stochastic wind stress component. This is discussed in the following section.

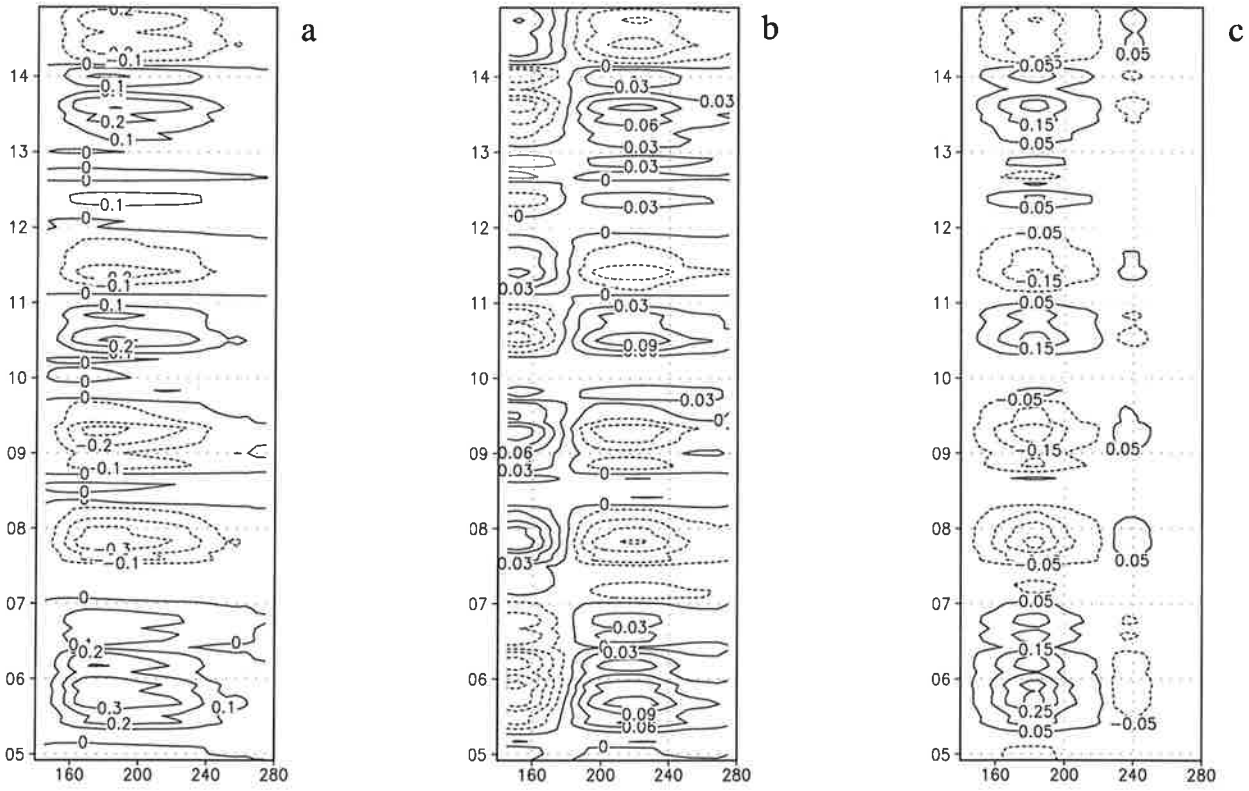


Fig.9: Longitude time section analogous to Fig.4 using the data of the 120yr integration of the stochastically forced HCM. The individual fields were scaled with the factors given below before they were pooled. The dominant POP mode has a perid of 61.9 months, an e-folding time of 22.6 months, and accounts for 32.4% of the variance of the dataset. The individual fields have to multiplied by factors of (a) 4.74 K, (b) 49.3 cm, and (c) 1.3×10^{-1} Pa to get physical units.

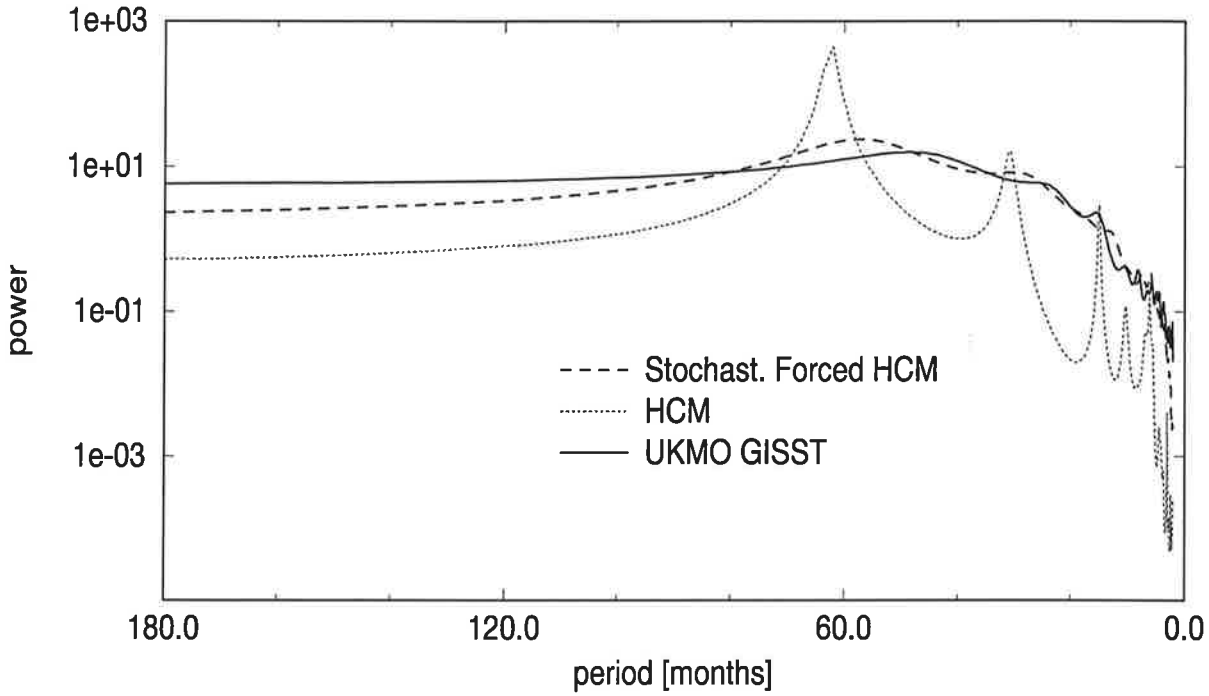


Fig.10: Maximum entropy spectra of the NINO3 time series of the UKMO GISST dataset lasting from 1900 to 1992, the Hybrid Coupled Model (using the data corresponding to Fig.3), the stochastically forced Hybrid Coupled Model (using the data corresponding to Fig.8). The spectra were obtained by using the SSA toolkit, Dettinger et al. (1995), which yielded stable spectral estimates for a range orders, typically 30-40, of the fitted autoregressive models.

2. Restart Ensemble Experiments

In a typical forecast experiment, the forecast model is initialised with the past history of forcing for a certain period of time. Then the forecast is started for the time span of interest, which in the case of ENSO ranges from one season up to about $1\frac{1}{2}$ years. Since the specific forceast is only one realisation of a stochastic dynamical process, ensembles of forecasts all starting with indentical initial conditions were created. Six restart ensembles were generated. Table 1 lists the restart dates. Additionally, the restarts are marked by black dots in Fig.8. Different phases of the model ENSO cycle were chosen. Each restart experiment comprised 216 members and was carried out for 72 months. Fig.11 and Fig.12 show such restart ensembles. Selected forecast trajectories as well as the time evolution of the ensemble mean and of the ensemble standard deviation are plotted. Fig.13 depicts the time evolution of the individual ensemble standard deviations seperately. The restart ensemble for month 276 shows a slower increase in ensemble variance than the one started at month 420. This can be explained in the following way. The ensemble of month 420 was started during the peak phase of a model El Niño event. Most of the ensemble members evolve into a La Niña event one year after this. But this is happening with slightly different phases which leads to a large spread among the ensemble members. The spread diminishes again after the cold event is reached by most forecasts.

Restart Month	Phase of Model ENSO Cycle
276	onset of El Niño
372	El Niño to La Niña
408	onset of El Niño
420	El Niño extreme
468	quiet phase
552	La Niña extreme

Table 1: List of the restart months corresponding to Fig.8 . The stochastically forced HCM was reinitialised at these dates and integrated forward for 72 months with different realisations of the stochastic part.

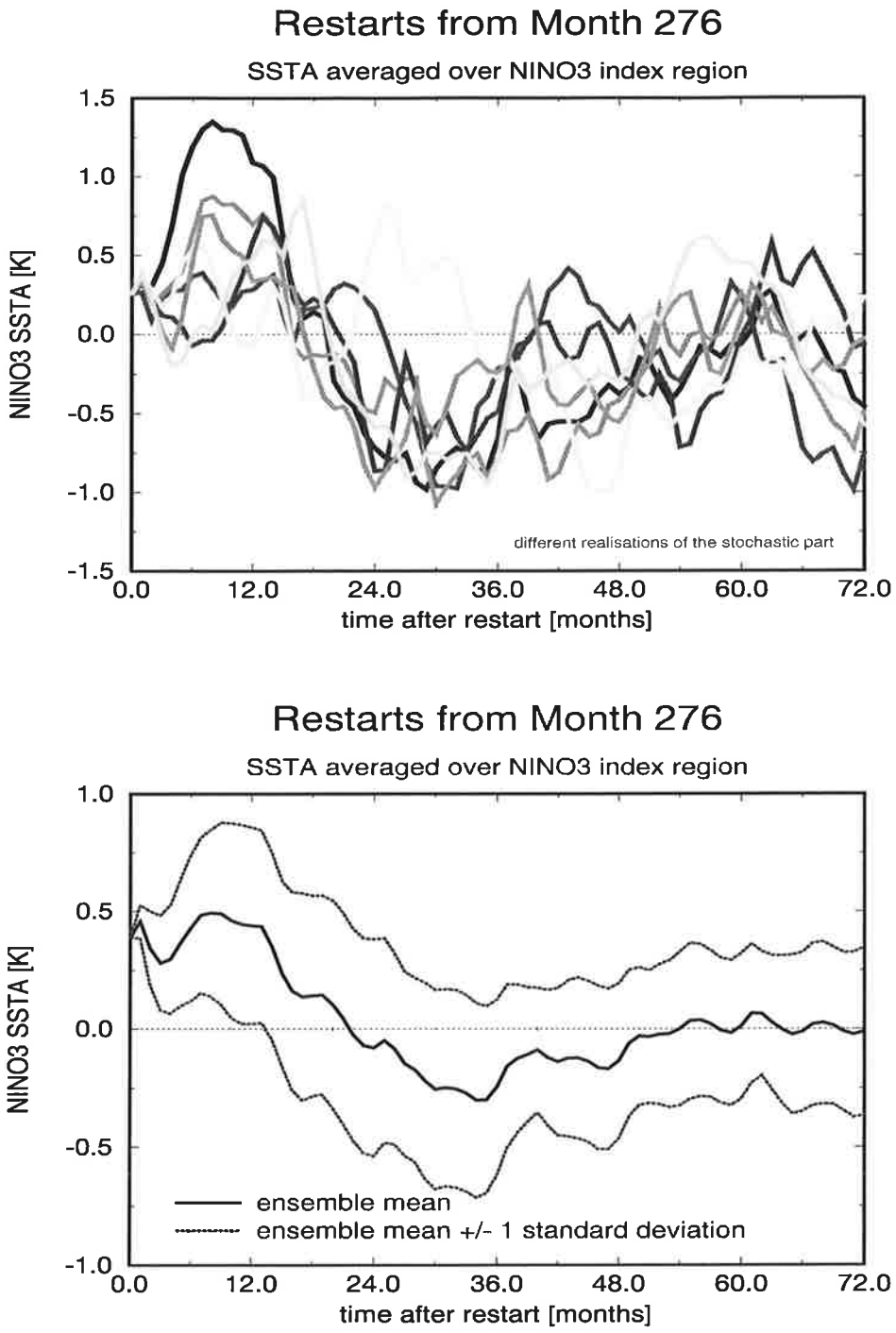


Fig.11: NINO3 time series of SSTA of the restart experiment initialised at month 276 of the integration of the stochastically forced HCM. The upper panel shows a few individual realisations, the lower one depicts the ensemble mean $\pm$ one standard deviation.

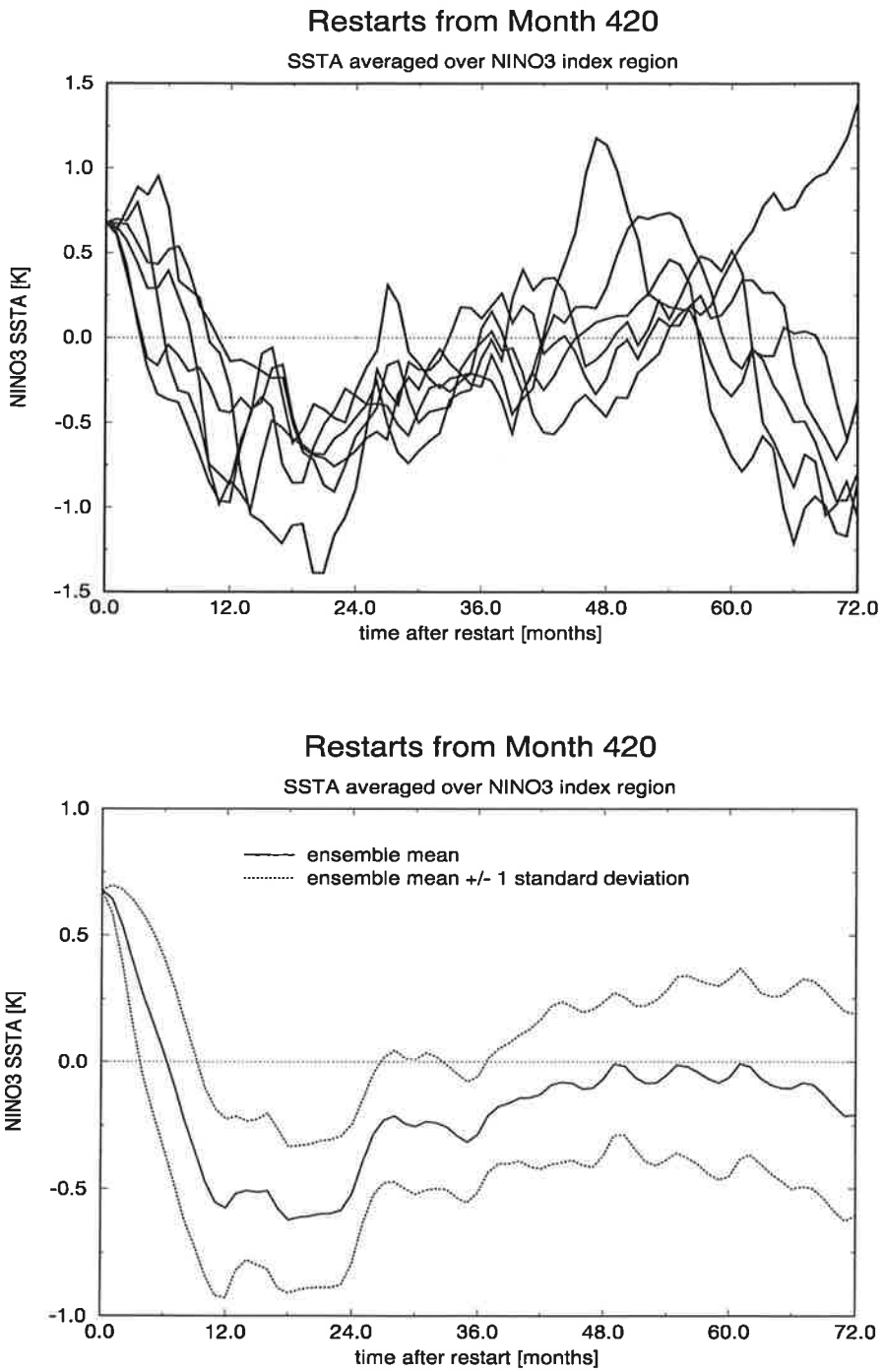


Fig.12: NINO3 time series analogous to Fig.11 for restart month 420: The upper panel shows some individual realisations, the lower one the time evolution of the ensemble mean $\pm$ one standard deviation.

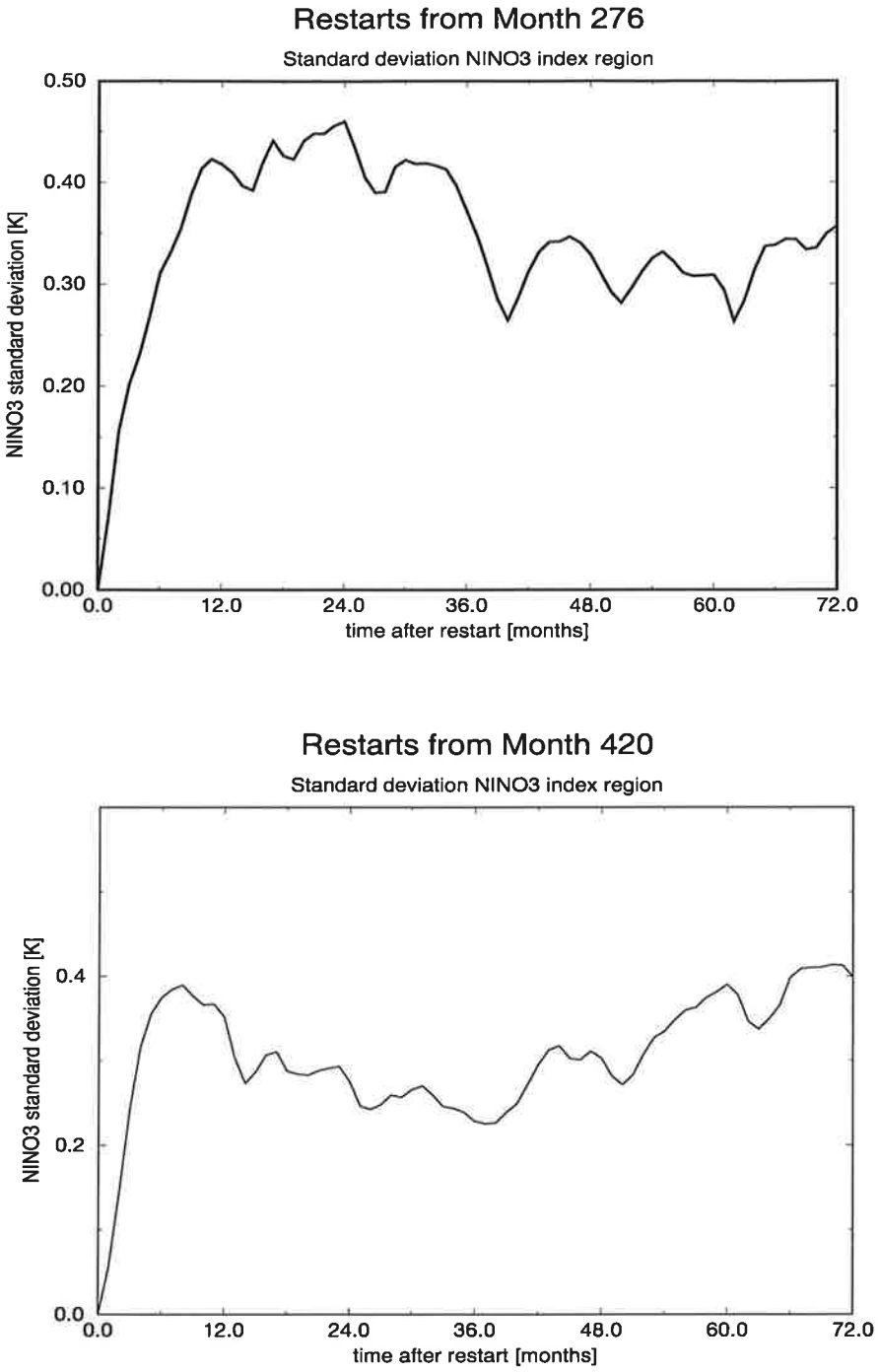


Fig.13: Time evolution of the ensemble standard deviations of the ensembles initialised at months 276 (upper panel) and at month 420 (lower panel). The ensemble started at month 420 shows a faster increase in ensemble variance resp. ensemble standard deviation than that started at month 276. For an explanation see text of this section.

3. Linear Stochastic Processes

In the following two sections the limitations to the modeled ENSO variability which occurr due to the random wind stress fluctuations are quantified. The six restart ensembles that were described in the previous section are taken as reference data. They serve as a substitute for observations. As a first step towards infering the model's predictability characteristics, the restart experiments were fitted to low dimensional linear stochastic processes. The success of the previous POP analyses in extracting the interannual variability contained in the observations as well as in the model data motivates such an approach.

Once again, the fields of SSTA, zonal wind stress anomalies and sea level anomalies were taken from the integration of the stochastically forced HCM were pooled and subjected to an EOF analysis. Next, the individual restarts were projected onto the leading EOFs. The time evolution of the expansion coefficients was then modeled by linear stochastic models.

EOF No.	Eigenvalue	Explained Variance	Cumulative Expl. Var.
1	8.26	35.91	35.91
2	1.21	5.25	41.16
3	1.08	4.68	45.84
4	0.78	3.38	49.22
5	0.62	2.68	51.90
6	0.52	2.27	54.17

Table 2: Result of the EOF analysis of the combined fields of SSTA, anomalous sealevel, and anomalous zonal wind stress for the 120yr integration of the stochastically forced HCM. The individual fields were scaled with the maximum absolute value they attained during the noise run before they were pooled.

A linear stochastic dynamical system is characterised by the stochastic differential equation

$$\dot{x}(t) = Ax(t) + B\eta(t) \quad \text{and} \quad x(0) = x_o \quad . \tag{12}$$

The vector $x(t)$ represents the model's state in the reduced state space, A and B denote the system matrix and the stochastic forcing, respectively, with $\eta(t)$ representing a Gaussian white noise process of unit variance and zero mean. The diffusion matrix of the random process is given by BB^T. Here, the superscript denotes the transpose of the matrix B. From this one can infer the time evolution of the mean and the covariance matrix, given appropriate initial values:

$$\frac{d}{dt}\langle x\rangle = A\langle x\rangle \quad ; \quad \langle x\rangle(0) = x_o \tag{13}$$

$$\frac{d}{dt}C = AC + CA^T + BB^T \quad ; \quad C(0) = C_o \tag{14}$$

Angle brackets denote ensemble means in this case. For a derivation of the above equations see e.g. van Kampen (1992) or Honerkamp (1990).

To obtain a fit that is as close as possible to the ensemble data over the forecast time interval $[0, T^*]$, a cost function $F(A, B)$ was defined that measures the misfit between the data and the linear stochastic model:

$$\begin{aligned} F(A,B) = \sum_{t=1}^{T^*} [&(\langle x\rangle - \langle x\rangle_{ens})^T M_{1,t} (\langle x\rangle - \langle x\rangle_{ens}) \\ &+ (C - C_{ens})^T M_{2,t} (C - C_{ens})] \quad . \end{aligned} \tag{15}$$

C denotes a column vector with its entries given by the main diagonal and the upper triangle of the covariance matrix. The subscript *ens* signifies the estimates of the ensemble mean and the ensemble covariance. The weighing matrices $M_{1,t}$ and $M_{2,t}$ were derived from the ensemble assuming normal variates

$$M_{1,t} = NC_{ens}^{-1} \tag{16}$$

$$\begin{aligned} M_{2,t} &= \mathrm{cov}(C_{ens,ij}, C_{ens,kl})^{-1} \\ &= \left[\frac{1}{N}(C_{ens,il}C_{ens,jk} + C_{ens,ik}C_{ens,jl})\right]^{-1} \end{aligned} \tag{17}$$

where N is equal to the ensemble size. For a derivation of the second matrix see Kendall et al. (1983). The cost function was minimised in parameter space by use

of an optimisation procedure. Among the many possible approaches a software tool developed by Giering and Kaminski (1996) was used. This tool generates the adjoint of the numerical evolution equations of Eq.(13) and Eq.(14). With the help of the adjoint of a given model the gradient of the cost function with respect to the entries of A and B can be determined, and this enters a minimisation algorithm which was taken from a standard software package (Numerical algorithms Group 1991). The generation and the use of adjoint models will be described in chapter IV.

Various dimensionalities of the reduced state space and time intervals $[0, T^*]$ were considered for carrying out the fit procedure outlined above. As dimension of the reduced state space 4, 5, and 6 were taken, the time spans ranged from 36 months up to 60 months. Fig.14 shows an example of the outcome of the fit procedure for the restart ensemble intialised at month 276 of the noise run. In Appendix A the results are summarised for all the fits performed. Two tables list the dominant interannual modes of the fitted processes and their simulated variances in comparison with the data.

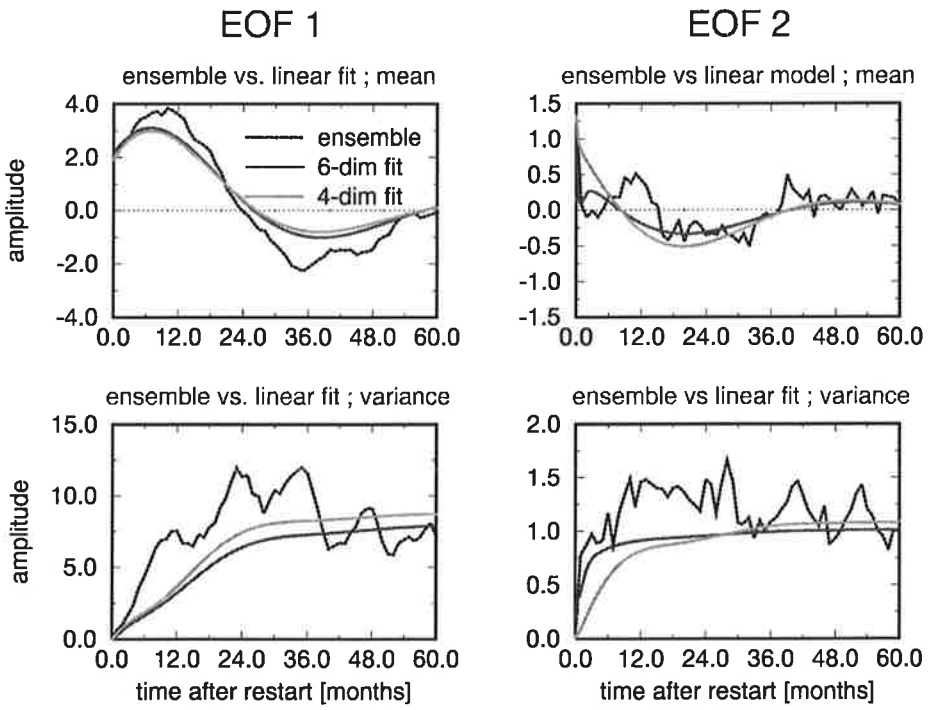


Fig.14: Ensemble data vs. the fitted linear stochastic processes for restart month 276. Depicted are the first and second coefficients of the ensemble mean and variance vs. the fitted the linear stochastic models of dimensionalities 4 and 6. The linear processes were fitted for a time interval of 60 months.

The fitted stochastic models are able to extract the information about the ensemble averages and the ensemble covariances on the interannual time scale contained in the restart ensembles. With these linear stochastic models it is now possible to quantify the restart ensembles's predictability characteristics.

4. Measure of Predictability

A motivation for the definiton of an ensemble's measure of predictability provides Fig.15. For some time t after the start it sketches the time-dependent probability density function (pdf) and the stationary pdf of a stochastic dynamical system.

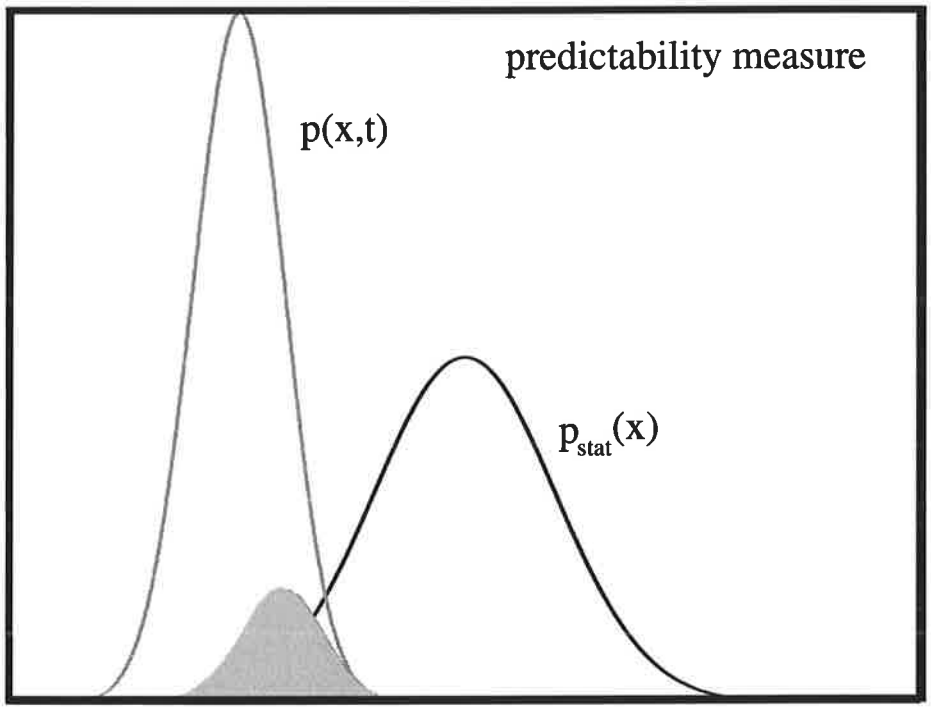


Fig.15: Illustration of the predictability measure explained in the text. The shaded area measures the overlap of the time-dependent and the stationary probability density functions.

As a measure of predictability, the overlap of the time dependent and the stationary pdf is used. At the start, the time dependent pdf is sharply peaked and well separable from the asymtotic one; the overlap of both pdfs is small. As time

goes by, the time dependent probability density widens, resulting in an increasing overlap. Eventually, both distributions are practically identical and the overlap is maximal. As long as the overlap is small the chances are good to disriminate between both pdfs. During the course of time the probability densities merge together. After this has happened, the climatological forecast will be the most adequate.

The above motivation can now be applied to linear stochastic processes. A one-dimensional linear stochastic process, the Ornstein-Uhlenbeck process

$$\dot{x} = -\alpha x + \sigma\eta, \quad and \quad x(0) = x_o \tag{18}$$

possesses as stationary and time-dependent pdfs, Honerkamp (1990),

$$p_{stat}(x) = \sqrt{\frac{\alpha}{\pi\sigma^2}} \exp\left(-\frac{\alpha x^2}{\sigma^2}\right) \quad \text{and} \tag{19}$$

$$p(x;t) = \sqrt{\frac{\alpha}{\pi\sigma^2(1-e^{-2\alpha t})}} \exp\left(-\frac{\alpha(x - x_o e^{-\alpha t})^2}{\sigma^2(1-e^{-2\alpha t})}\right) \quad , \tag{20}$$

respectively. The overlap of these pdfs increases during the forecast. Therefore, an appropriate measure of predictability is given by the following expression:

$$s_{OU}(t) = 1 - 2\sqrt{\pi\sigma\sigma(t)} \int dx p(x;t) p_{stat}(x) \quad . \tag{21}$$

Inserting (19) and (20) yields:

$$s_{OU}(t) = 1 - \sqrt{\frac{2(1-e^{-2\alpha t})}{2-e^{-2\alpha t}}} \exp\left[-\frac{\alpha x_o e^{-2\alpha t}}{\sigma^2(1-e^{-2\alpha t})}\right] \tag{22}$$

with $\sigma(t) = \sigma\sqrt{(1-e^{-2\alpha t})}$. This definition has the properties

$$s_{OU}(0) = 1 \quad \text{and} \quad s_{OU}(\infty) = 0 \quad . \tag{23}$$

The generalisation of the definition (21) to an n-dimensional linear stochastic process, characterised by the stochastic differential equation (12) is given by

$$s(t) = 1 - (2\sqrt{\pi})^n [\det(C(t)C_{stat})]^{1/4} \int dx p(x;t) p_{stat}(x) \quad , \tag{25}$$

with $C(t)$ and C_{stat} denoting the covariance matrix at time t and the stationary covariance matrix, respectively. The pdfs $p(x;t)$ and $p_{stat}(x)$ are multivariate normal distributions. Again, the normalising factors in front of the integral yield

an initial value of 1 and an asymptotic value of 0 for $s(t)$ provided all the ensemble members start with indentical initial conditions. This is the case for the fitted linear stochastic models described in the previous section. The behaviour of $s(t)$ between the two limiting cases $t = 0$ and $t = \infty$ depends on the specific process, i.e. the system dynamics, the noise forcing and on the initial state of the system.

Fig.16 summarises the time evolution of the predictability measure determined for the restart experiments. Before the predictability measure was calculated, the fitted linear stochastic models were projected onto the leading two EOF expansion coefficients. Such a projection does not alter the results substantially, since the third and higher principal components saturate quickly to their stationary values. Thus, they contain little information relevant to characterise the intermediate and asymptotic behaviour of $s(t)$. In all cases, the predictability is lost well before a typical cycle period has elapsed. Among the ensembles there are different characteristics observable. Especially the restarts initialised in the peak warm phases exhibit a peculiar behaviour which can be interpreted in the following way. Since all ensemble members tend to evolve into a cold state and the stochastic forcing induces phase differences between them the time dependent probability distribution widens relatively quickly, yielding a rapid decrease in $s(t)$. The subsequent cold state represents a relatively large deviation from the climatic mean. After the regions around zero anomaly are passed, therefore, the climatic and the time dependent probability distribution reduce their overlap and the predictability measure recovers. For longer forecast time intervals the time dependent probability distribution merges into the asymptotic one and $s(t)$ gradually vanishes as in the other ensembles.

In this chapter, the initial error distribution was not considered, all ensemble members start with indentical initial conditions. Additional initial errors would limit the system's predictability still further. They would add a term $C_{ini}(t) = \exp(At)C(0)\exp(At)^T$ to the time evolution of the ensemble covariance matrix, which can yield considerable initial error growth due to the non-self-adjointness of $\exp(A)$, the one time step propagator. This was discussed by Blumenthal (1991), using an estimate of the propagator obtained from the cyclostationary POP method, which takes into account the seasonality in the underlying model. Here, the system matrix and the driving noise were estimated by fitting a linear stochastic process locally – that is, over a short time interval – to a restart ensemble of the HCM. As such, it takes into account the model's actual state. The fits are optimised to monitor the time evolution of the underlying system for a finite time span. The results obtained suggest that in making actual ENSO predictions, ensemble forecasts which include noise forcing will provide valuable

estimates of the forecasts' uncertainty.

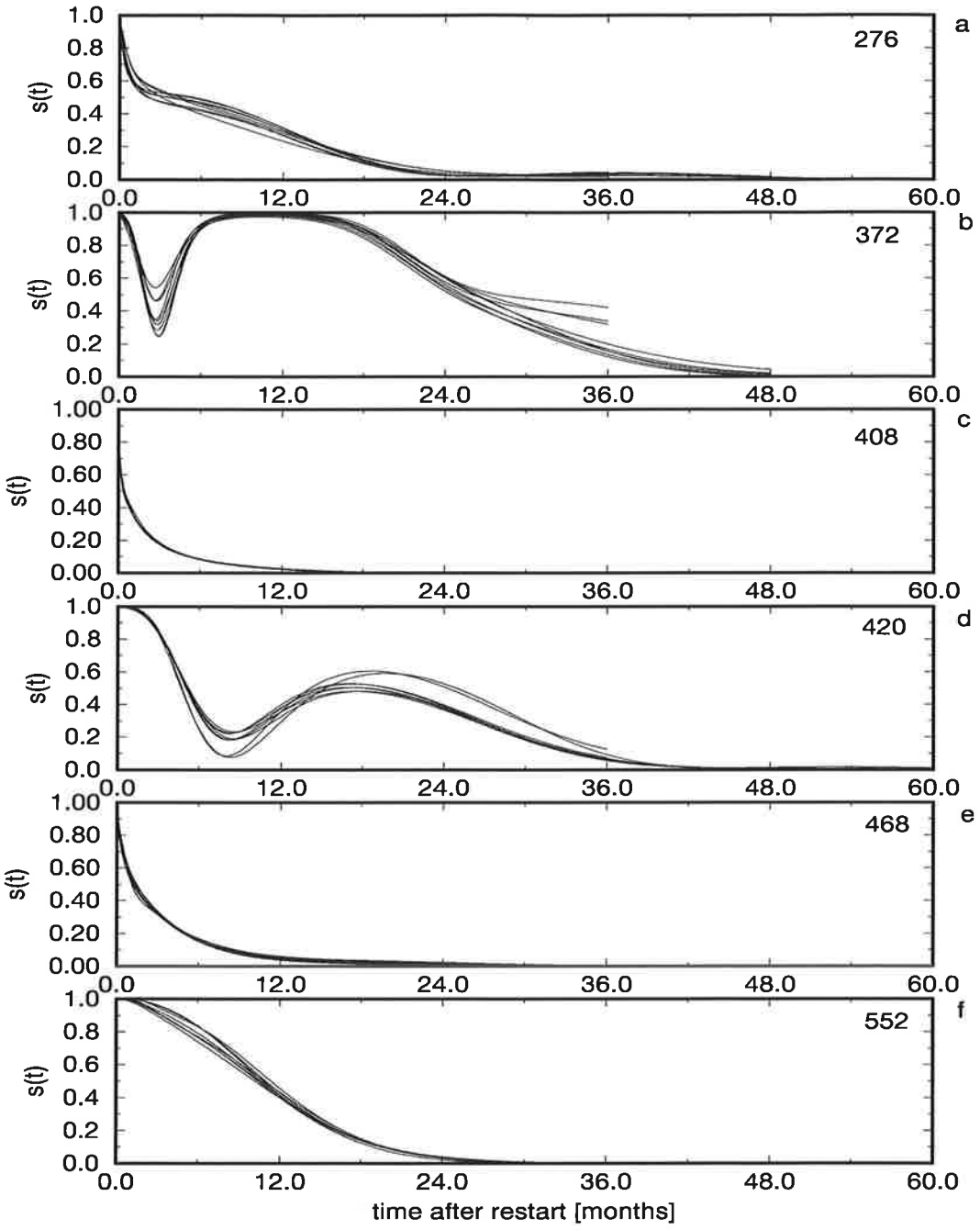


Fig.16: Predictability measures obtained for the restart experiments. The restart months are indicated in the upper right corner of each panel.

The time evolution of the predictability measure $s(t)$ indicates a clear dependence on the system's state and may therefore be used as an objective measure of the expected forecast quality.

There are, of course, some caveats to consider. The estimate of the stochastic forcing included into the HCM may suffer from errors in the underlying data set, and the way of constructing the random perturbation cannot be justified rigorously. This implies that the noise variance could have been over- or underestimated, which in turn would affect the coupled model's predictability. However, the fact that the results obtained here are consistent with those of Blanke et al. (1997) strengthens the reliability of the results. The dominant period of the model ENSO mode is longer than the observed one. Further, the approximating linear systems do not fit the stochastically forced HCM perfectly, since they are formulated in reduced state spaces, and misfits remain after optimisation. The sensitivity of the stochastically forced HCM to different noise amplitudes and coupling strengths was investigated (not shown here). The reduction of the noise amplitude induced a more regular behaviour, eventually reproducing the HCM's control run. Reducing the coupling strength at a constant noise level yielded an increasingly irregular behaviour. At zero coupling strength and non-vanishing noise forcing, the NINO3 time series did not exhibit marked interannual fluctuations any more. The choice of the coupling strength and the noise level, although admittedly subjective, was guided by the spatio-temporal characteristics of the model simulations.

Nevertheless, it is expected that the main results of this chapter remain valid. The stochastic forcing is important to model ENSO realistically, and the ENSO predictability is limited by it, leading to a loss of predictability before a typical cycle period has elapsed.

IV. Limitation of Predictability due to Uncertain Initial Conditions

1. The Growth of Initial Errors, Optimal Perturbations, and Singular Vector Analysis

In the foregoing chapter all forecasts were made with identical initial conditions. Here, we are going to concentrate on the issue of errors in the initial state and their subsequent evolution during the forecast. At first, the random perturbations occurring during the forecast will not be considered. Afterwards, the additional random forcing will be included.

In general, the coupled model's state cannot be specified accurately, since measurements are not complete, the measurements themselves have some errors, or the initialisation procedure is imperfect. Let us assume that the initial uncertainty of a forecast can be described by some multivariate probability distribution with the dimension of the probability space being given by number of prognostic variables of the forecast model. This multivariate probability density evolves in the course of time. At forecast time it has changed its original shape and volume. Looking in more detail into the time evolution, one can discern directions in the initial error field in which errors are magnified and others in which errors are damped during the forecast. This implies that one is mainly interested in determining those directions in which initial erros grow most strongly. If the initial error field is projected onto these directions, one can obtain an estimate of the likely spread of the forecast.

One approach to determine these directions in the model's state space, consists of the investigation of sufficiently small perturbations, so that linear perturbation theory can be applied. Let the forecast be denoted by

$$x(T^*) = M(x(0)) \quad . \tag{25}$$

The initial state $x(0)$ is mapped onto the final state $x(T^*)$ some time T^* later via the forecast model M. A small perturbation $\delta x(0)$ of the initial state results in a perturbed final state

$$x(T^*) + \delta x(T^*) = M(x(0) + \delta x(0)) \quad , \tag{26}$$

which in a linear approximation yields for the final state's change

$$\delta x(T^*) = L(T^*, 0)\delta x(0) \quad . \tag{27}$$

The linear mapping $L(T^*,0)$ is called the tangent linear propagator over the time interval $[0,T^*]$. It has to be determined using the forecast model M. Given norms in the spaces of initial and final state perturbations, $\|.\|_0$ and $\|.\|_{T^*}$, one can define lengths of the perturbations and therefore their growth with respect to these norms. The amplification rate of a perturbation is defined as:

$$a = \frac{\|\delta x(T^*)\|_{T^*}}{\|\delta x(0)\|_0} \quad . \tag{28}$$

Scalar products in the spaces of initial and final state perturbations induce such norms:

$$\|\delta x(0)\|_0^2 = \langle \delta x(0), \delta x(0) \rangle_0 = \delta x(0)^T M_0 \delta x(0) \quad \text{and} \tag{29a}$$

$$\|\delta x(T^*)\|_{T^*}^2 = \langle \delta x(T^*), \delta x(T^*) \rangle_{T^*} = \delta x(T^*)^T M_{T^*} \delta x(T^*) \quad . \tag{29b}$$

Depending on the problem to be studied, one may favour specific choices of the positive-definite metric matrices M_0 and M_{T^*}.

The optimal perturbations are by definition those initial perturbations, that yield largest amplification rates over the integration time T^*. Mathematically, these can be determined by solving a maximisation problem for the final state perturbation's length given a unit perturbation at initial time:

$$\langle \delta x(T^*), \delta x(T^*) \rangle_{T^*} + \lambda[1 - \langle \delta x(0), \delta x(0) \rangle_0] = \max \quad . \tag{30}$$

In the above equation λ denotes a Lagrange multiplier, introduced to incorporate the constraint $\|\delta x(0)\|_0 = 1$. If the evolved initial perturbation $\delta x(T^*)$ is represented by the action of the tangent linear propagator $L(T^*,0)$ on the initial perturbation, and use is made of the definition of the adjoint tangent linear propagator $L(T^*,0)^T$, namely

$$\langle \delta z, L(T^*,0)\delta x(0) \rangle_{T^*} = \langle L^T(T^*,0)\delta z, \delta x(0) \rangle_0 \quad , \tag{31}$$

then one obtains

$$\begin{aligned} \langle \delta x(T^*), \delta x(T^*) \rangle_{T^*} &= \langle L(T^*,0)\delta x(0), L(T^*,0)\delta x(0) \rangle_{T^*} \\ &= \langle L^T(T^*,0)L(T^*,0)\delta x(0), \delta x(0) \rangle_0 \quad . \end{aligned} \tag{32}$$

The perturbation δz is arbitrarly chosen in the space of final state perturbations.

Inserting the metric matrices and differentiating with respect to the initial state perturbation to determine the maximum, yields a symmetric eigenvalue problem:

$$M_0^{-\frac{1}{2}} L(T^*,0)^T M_{T^*} L(T^*,0) M_0^{-\frac{1}{2}} \delta y(0) = \sigma^2 \delta y(0) \quad . \tag{33}$$

To render the eigenvalue problem symmetric, the transformation

$$\delta y(0) = M_0^{\frac{1}{2}} \delta x(0) \quad , \tag{34}$$

which does not change the eigenvalue spectrum, has to be applied. Note that in (31) and in (32) the superscript T denotes the adjoint with respect to the scalar products chosen, whereas in (33) the adjoint refers to the usual euclidean scalar product. For the sake of simpler notation the two cases were not distinguished in these fromulae.

The set of eigenvectors determined from this maximisation problem forms a complete orthonormal basis of the space of initial state perturbations. An arbitrary initial perturbation can be expanded as a linear combination of these eigenvectors. Following common use the optimal perturbations are also called singular vectors, since there exists an intimate relationship to the well known singular value decomposition of matrices. The SVD of the tangent linear propagator reads:

$$L_M(T^*,0) := M_{T^*}^{\frac{1}{2}} L(T^*,0) M_0^{-\frac{1}{2}} = U \Sigma V^T \quad . \tag{35}$$

The matrices U and V are unitary, i.e. $U^T = U^{-1}$ and $V^T = V^{-1}$. Σ is a diagonal matrix, whose entries are called the singular values of L_M. The columns of V are called the right singular vectors and the columns of U the left singular vectors of the decomposed matrix, respectively. Inserting this into (33), we obtain

$$L_M^T L_M = V \Sigma^2 V^T \quad . \tag{36}$$

Therefore, the above determined optimal perturbations are exactly the right singular vectors of the propagator $L_M(t,0)$. They form a basis of the space of initial perturbations, whereas the left singular vectors span the space of final state perturbations:

$$L_M L_M^T = U \Sigma^2 U^T \quad . \tag{37}$$

An evolved right singular vector v_i is given by $\sigma_i u_i$.

The concept of singular vector analysis has been widely applied in atmospheric sciences, see e.g. Farrel and Ioannou (1996a and 1996b) and references therein.

In Numercal Weather Prediction singular vectors can be used as efficient perturbations to generate forecast ensembles, Buizza et al. (1993) and Palmer et al. (1998).

In the context of seasonal predictability several studies already exist as well. They differ in the way the symmetric linear operator was obtained, and in the way the matrices M_0 and M_t were specified. Xue et al. (1994) analysed error growth in the forecast model of Zebiak and Cane (1987). After projection onto the leading multivariate EOFs of the model a cyclostationary POP analysis using monthly transition matrices was performed, see Blumenthal (1991). This highly reduced linear model served as a reference to compute its singular vectors. Xue et al. identified the first two singular vectors contributing most to the initial error growth in the linear model. Initial error growth in such an asymtotically stable linear model is intimately related to the non-selfadjoint nature of the system matrix. For example, two POPs with very different decay times but similar spatial structure can induce rapid transient error growth after the start of a forecast. If that fraction of a model's initial state which erroneously leads to an El Niño condition, could be subtracted, this would lead to better forecasts. A technically much simpler approach to estimate the tangent linear propagator of a forecast model was chosen by Chen et al. (1997). Since their forecast model posesses a moderate number of prognostic variables the tangent linear propagator was estimated by perturbing each state variable slightly and carrying out a model integration for the time interval of interest. This finite difference approximation was then used to calculate the tangent linear propagator of the SST field. Singular vector analysis revealed one dominant structure which varied little in its spatial characteristcs, but which had markedly different amplification rates with respect to an ENSO cycle and with respect to the annual cycle as well. The norm used in this study was based on the euclidean scalar product. The onset phases of El Niño and the transitions from warm conditions to La Niña states were characterised by larger growth rates than the maxima of the model ENSO. During the course of a year, maximum growth rates were found in boreal summer, and the period from January to April showed the smallest amplification rates. The leading singular vector can be characterised as a dipolar pattern with respect to the meridian at 150°W. It evolves into a structure similar to a model El Niño after 3 months integration time. Moore and Kleeman (1996) determine the singular vectors of the coupled model described in Kleeman (1993) by a method that does not exploit the reduction of state space like the above- mentioned studies. Instead, they determined the model's linearisation from the model equations and coded it subsequently. The adjoint model was then determined from the discretised tangent linear model equations. The perturbation

energy was used as a norm in the calculations. The singular value spectrum was dominated by the leading singular vector for all experiments performed. Largest growth rates during the annual cycle were observed in boreal spring. During the onset of El Niño the growth rates were larger than during the onset of La Niña. This indicates better chances to predict cold anomalies skillfully than warm states. In Moore and Kleeman (1997a) and (1997b) a detailed analysis of the dynamics and energetics of the singular vectors of the coupled model is presented.

2. Determination of the Leading Optimal Perturbations of the HCM

In this study, an approach is persued that consists in the derivation of the tangent linear model and its adjoint directly from the numerical code of the Hybrid Coupled Model. The numerical representation of the Hybrid Coupled Model can be viewed as the composition of individual mappings. Each statement in the computer code and each structure therein can be viewed as such a mapping:

$$x(t) = \left[S_N \circ S_{N-1} \circ ... \circ S_2 \circ S_1\right] x(0) \quad . \tag{38}$$

Linearising this sequence of statements is straightforward:

$$\delta x(t) = \left[TL_N \circ TL_{N-1} \circ ... \circ TL_2 \circ TL_1\right] \delta x(0) \quad . \tag{39}$$

Here, TL_i signifies the linearised version of statement S_i in the numerical code which involves the Jacobian of the individual mappings. The bracketed composition of linear mappings represents the tangent linear propagator $L(T^*, 0)$ used in Eq.(33). The adjoint tangent linear propagator $L^T(T^*, 0)$ can formally be obtained by transposing (39), sucessively applying the rule $(AB)^T = B^T A^T$ for the composition of linear mappings A and B:

$$L^T(T^*, 0) = TL_1^T \circ TL_2^T \circ ... \circ TL_{N-1}^T \circ TL_N^T \quad . \tag{40}$$

The very simple example below is intended to exemplify the strategy for generating tangent linear statements and their corresponding adjoint statements from given statements. Such an example is instructive, since it shows some important features of the process of automatic differentiation. Let the following statement be contained in an algorithm to be differentiated:

$$A = A^2 + \sin(B) \quad . \tag{41}$$

The tangent linear statement is readily obtained:

$$\delta A = 2A\delta A + \cos(B)\delta B \quad . \tag{42}$$

Now, an easy way to generate the corresponding tangent linear code is to first evaluate the tangent linear statement (42), and after this to calculate statement (41), since statement (42) requires the value A attained before (41) is excuted. No, seperate storage of data is thus required to run this tangent linear code. The work space for the tangent linear code will be slightly less than twice the original work space of the nonlinear algorithm. Statement (42) can as well be represented by the linear mapping

$$\begin{pmatrix} \delta B \\ \delta A \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ \cos B & 2A \end{pmatrix} \begin{pmatrix} \delta B \\ \delta A \end{pmatrix} \quad , \tag{43}$$

which after transposition yields for the adjoint mapping

$$\begin{pmatrix} \delta^* B \\ \delta^* A \end{pmatrix} = \begin{pmatrix} 1 & \cos B \\ 0 & 2A \end{pmatrix} \begin{pmatrix} \delta^* B \\ \delta^* A \end{pmatrix} \quad . \tag{44}$$

Consequently, the adjoint statements that have to be coded read:

$$\begin{aligned} \delta^* B &= \delta^* B + \cos B \delta^* A \\ \delta^* A &= 2A\delta^* A \end{aligned} \tag{45}$$

Again, the values of A and B attained before (41) is excuted are required to evaluate the adjoint statements in (45). Since the order of computations in the adjoint code is reversed, see (40), they either have to be recomputed, or they have to be stored during the run of the nonlinear algorithm and supplied to (45) when needed. Thus, a considerable amount of book-keeping is needed to efficiently implement the adjoint model. In practice certain variables will be recomputed during the run of the adjoint model and others will be read in from a secondary storage device when needed. Due to these requirements, the adjoint code requires the largest fraction of computing time for the evaluation of the symmetric linear operator (33). To be specific, the run times of the adjoint tangent linear HCM, the tangent linear HCM, and the nonlinear HCM have the proportions $2.8 : 1.7 : 1$.

Differentiating an algorithm can be automatised. Each mapping of the numerical code has to be differentiated according to well known rules. To perform such a differentiation, the differentiation tool has to know the set of variables with respect to which a given statement has to be differentiated. Such variables are called active variables. Variables that are not active are called passive variables of the algorithm. The set of active and passive variables can be determined by

performing a dependence analysis prior to the differentiation. The differentiation tool must apply the rules of differentiation to simple statements, such as addition, multiplication etc., built in functions and procedures of the programming language, and it must be able to resolve programming structures, such as loops, conditional statements etc. . A software tool that performs exactly these tasks was developed by R. Giering, see Giering (1997) and Giering and Kaminski (1996). The tool is called Tangent Linear and Adjoint Model Compiler (TAMC).

All that is needed to generate the linearised code of the HCM by the TAMC are some modifications of the FORTRAN source code to match the programming conventions expected by the TAMC. The interested reader will find a good account on the techniques of automatic differentiation in Giering (1996) or in Griewank and Corliss (1991) and references therein. After specification of the input and output variables, the TAMC performs a dependence analysis to indentify the active and the passive variables contained in the numerical code. Then, the tangent linear and the adjoint tangent linear computer codes can be generated.

Tangent linear and adjoint models can be used in different contexts. Apart from the computation of optimally growing perturbations, adjoint models can be used to determine the sensitivities of a dynamical model to small perturbations at an earlier time, e.g. van Oldenborgh et al. (1997). Particularly interesting in this respect is the determination of the gradient of a scalar function with respect to e.g. internal model parameters or initial conditions. The linear stochastic processes discussed in the previous chapter were determined this way. Another example is the assimilation of data into circulation models of the atmosphere or of the ocean can be done with the help of the corresponding adjoint models. The advantage of this approach is that the assimilation procedure is consistent with the dynamics of the underlying general circulation model. Such applications along with an introduction to the generation of adjoint numerical models is excellently described in Talagrand (1991).

The TAMC generated computer codes were subsequently tested. The adjoint code was tested by introduction of a simple scalar function. Computation of its gradients and comparison with finite difference approximations yielded good agreement. The tangent linear code was tested by comparing scalar products according to (32). The results obtained aggree within computational accurracy.

Since the tangent linear HCM has a large number of input and output variables, an algorithm suited to handle these large dimensions for the solution of the eigenvalue problem (33) is needed. In this study a modified version of the Lanczos algorithm was chosen, Parlett and Scott (1979). The Lanczos algorithm

reduces the high dimensional problem to a low dimensional one, which approximates the largest or the smallest eigenvalues of the symmetric eigenvalue problem (33). Details about the Lanczos algorithm are given in Appendix B.

As a final step, norms at the initial and final time of the model integration have to be specified. In the experiments described in the next section, an energy norm was used. The perturbation energy equation for a stratified ocean reads, see Gill (1982):

$$\frac{\partial}{\partial t}\int \frac{1}{2}\rho_o(\delta u^2+\delta v^2+\delta w^2)dz+\frac{\partial}{\partial t}\Big\{\frac{1}{2}\rho_o(0)g\delta\eta^2+\int\frac{1}{2}\Big(\frac{g^2\delta\rho}{\rho_o N^2}\Big)dz\Big\}$$

$$+\frac{\partial}{\partial x}\int \delta p\delta u dz+\frac{\partial}{\partial y}\int \delta p\delta v dz=0 \quad . \tag{44}$$

For the definition of a norm, we retain those terms containing the prognostic variables of the ocean model, i.e. the kinetic energy term for the horizontal velocities, the term containing the density perturbation which is related to temperature perturbations by the equation of state used in the ocean model, and the term containing the sea surface elevation. Additionally, the associated entries in the metric matrices are multiplied by weights, which are smallest in the equatorial region and largest at the northern and at the southern boundaries. This implies that the equatorial regions may attain larger perturbation amplitudes relative to the off-equatorial parts of the basin. The motivation for this additional weighting reflects the assumption that the equatorial regions are most important for ENSO dynamics. The choice of norm has the property, that regions of large vertical temperature gradients can attain larger perturbation amplitudes, since there the metric weights are smaller compared to regions with weaker vertical temperature gradients. The perturbations in the thermocline region will thus be important for the resulting singular vectors. Since the equation of state of the ocean model relates density linearly to temperature the density perturbation $\delta\rho$ can easily be substituted in (44) by $\rho_o\alpha\delta T$, see (10).

3. Dynamics of Optimal Perturbation Growth

In this section, several sets of experiments are discussed. First, the leading 10 singular vectors and singular values of the HCM were determined during a model ENSO cycle without noise forcing present. During the 5 years of such a cycle every quarter of a year a new calculation was started. Forecast intervals of 3, 6, 9 and 12 months were considered. This first set comprises a total of 80 individual experiments. In a further set of singlar vector calculations the atmospheric feedback was switched off in the HCM and in its linearisation. One year during this annual cycle case was chosen, and again forecast intervals of 3, 6, 9, and 12 months were taken for the determination of the leading 10 singular vectors and values. This set comprises 16 experiments. In a third set of experiments, the leading singular vectors of the stochastically forced HCM were computed for a restart date already employed in chapter III. to generate model trajectories with different realisations of the stochastic wind stress component. In this case 4 different realisations of wind stress noise were considered.

In Fig.17 a typical singular value spectrum is shown. In most experiments carried out the first singular value dominates the spectrum. The amplification rates with respect to the chosen norm are considerable with average values of $a = 68.5$ in the 80 ENSO cycle experiments and $a = 63.9$ in the 16 annual cycle experiments. The individual leading singular values of the two sets of experiments are displayed in Fig.18 and Fig.19. The amplification rates exhibit marked dependences on the annual cycle and on the phase of the model ENSO cycle.

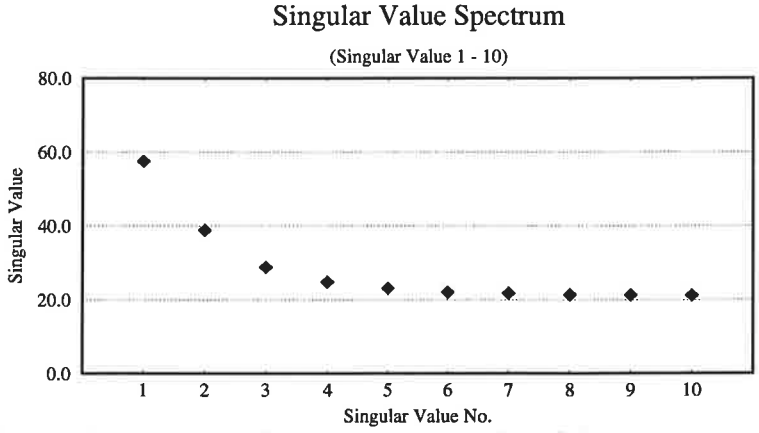


Fig.17: A typical singular value spectrum. Shown are the leading 10 singular values of a coupled experiment starting in January of the second year during the model ENSO cycle. The singular values were optimised for 3 months integration time.

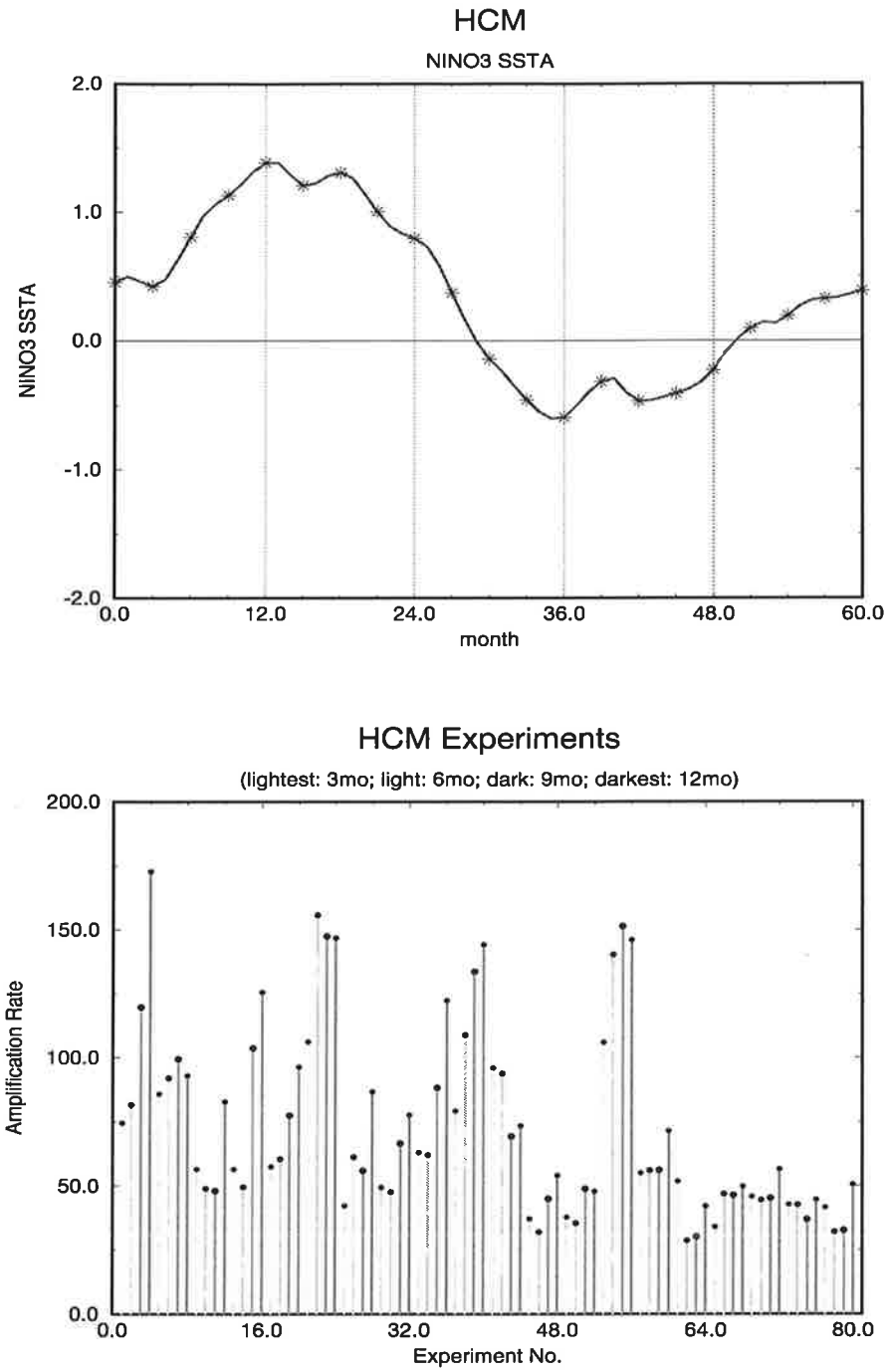


Fig.18: Amplification rates of the leading singular vectors of the experiments performed during a model ENSO cycle. The upper panel shows the restart dates, the lower panel the amplification rates obtained. Different shading indicates the optimisation time of 3, 6, 9, or 12 months.

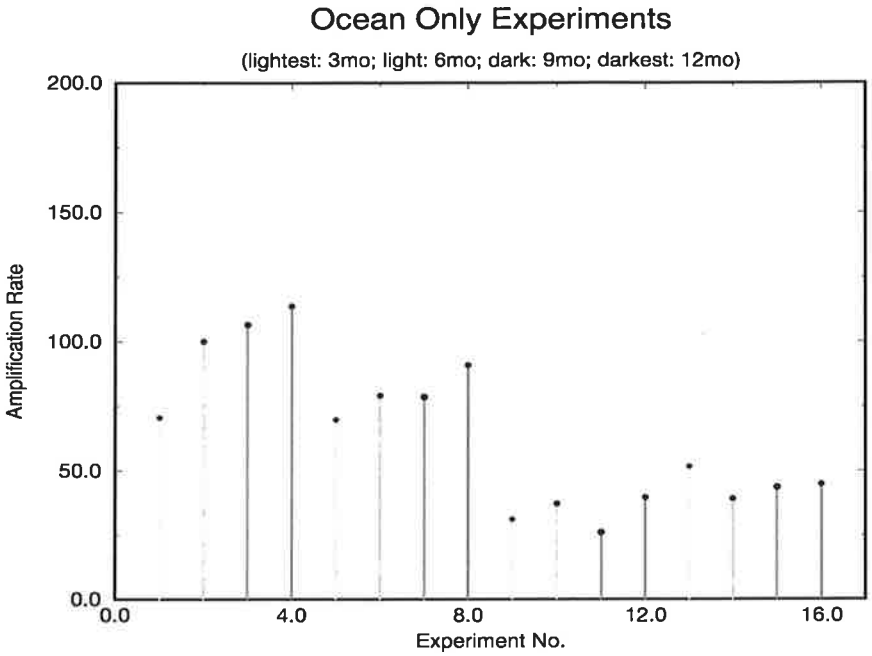
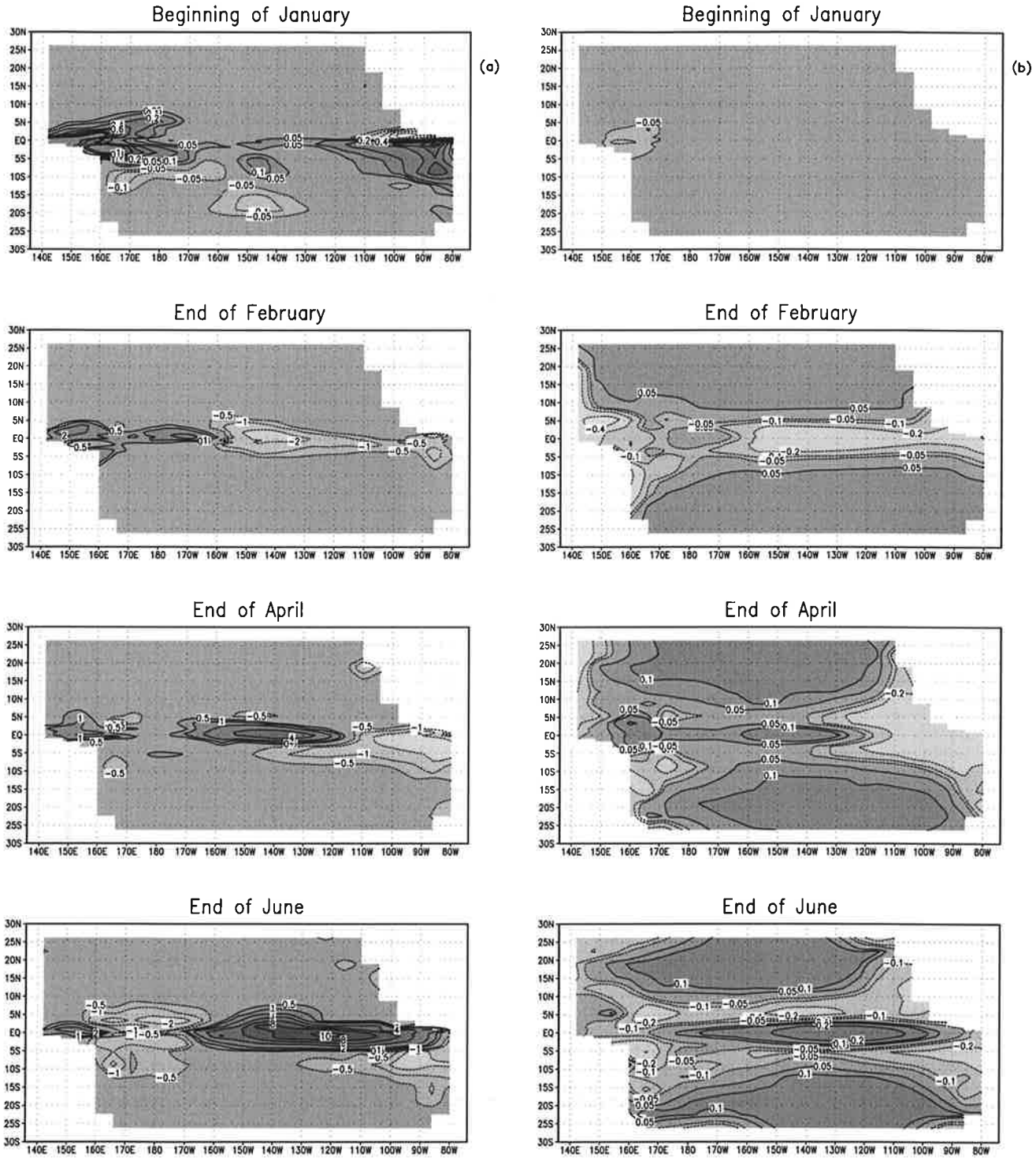


Fig.19: Amplification rates of the leading singular vectors of the experiments performed during an uncoupled model run. The 16 experiments are grouped according to their starting date: 1–4 January starts, 5–8 April starts, 9–12 July starts, and 13–16 October starts. 3,6,9, and 12 months optimisation times were chosen for each starting date.

As a common feature of all the experiments, integrations started in January and in April tend to exhibit much larger amplification rates than those initialised in July and in October. The dependence on the ENSO cycle is reflected in the reduced amplification rates during year 5 of the model ENSO cycle and in the sometimes enhanced amplification rates during the rest of the cycle.

So far, only the amplification rates of the singular vectors were presented. The spatial patterns associated with the singular values are depicted in Figs. 20 – 23 on the following pages. The time evolutions of two singular vectors, one obtained from the ENSO cycle experiments and one obtained from the uncoupled experiments, is shown.


Beginning of January
(a)
Beginning of January
(b)
End of February
End of February
End of April
End of April
End of June
End of June

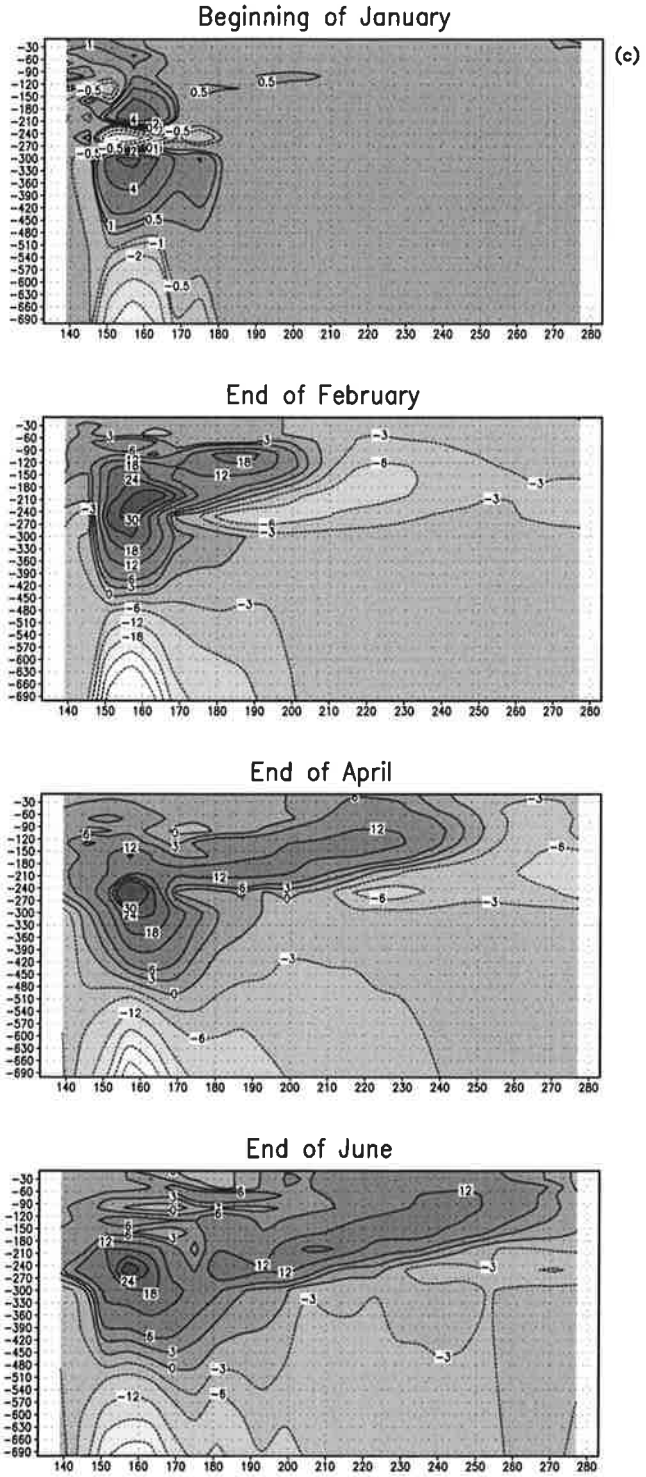


Fig.20: Leading singular vector of the experiment started in January of the second year during the coupled integration. In (a) the time evolution of the surface temperature perturbation is shown. (b) depicts the evolution of the sea level perturbation, and (c) contains a vertical section along the equator depicting the temperature perturbation. Starting in the western part of the basin, the perturbation gradually propagates to the East. The vertical section along the equator reveals that two baroclinic modes propagate to the eastern boundary. Biggest amplitudes in the deeper oceanic layers are located in the thermocline region, where large vertical temperature gradients in the reference state are present. The temperatures are in units of [K] and the sea level is in [m]. The fields may be scaled by an abritrary common factor since the calculations involved in the determination of singular vectors are linear.

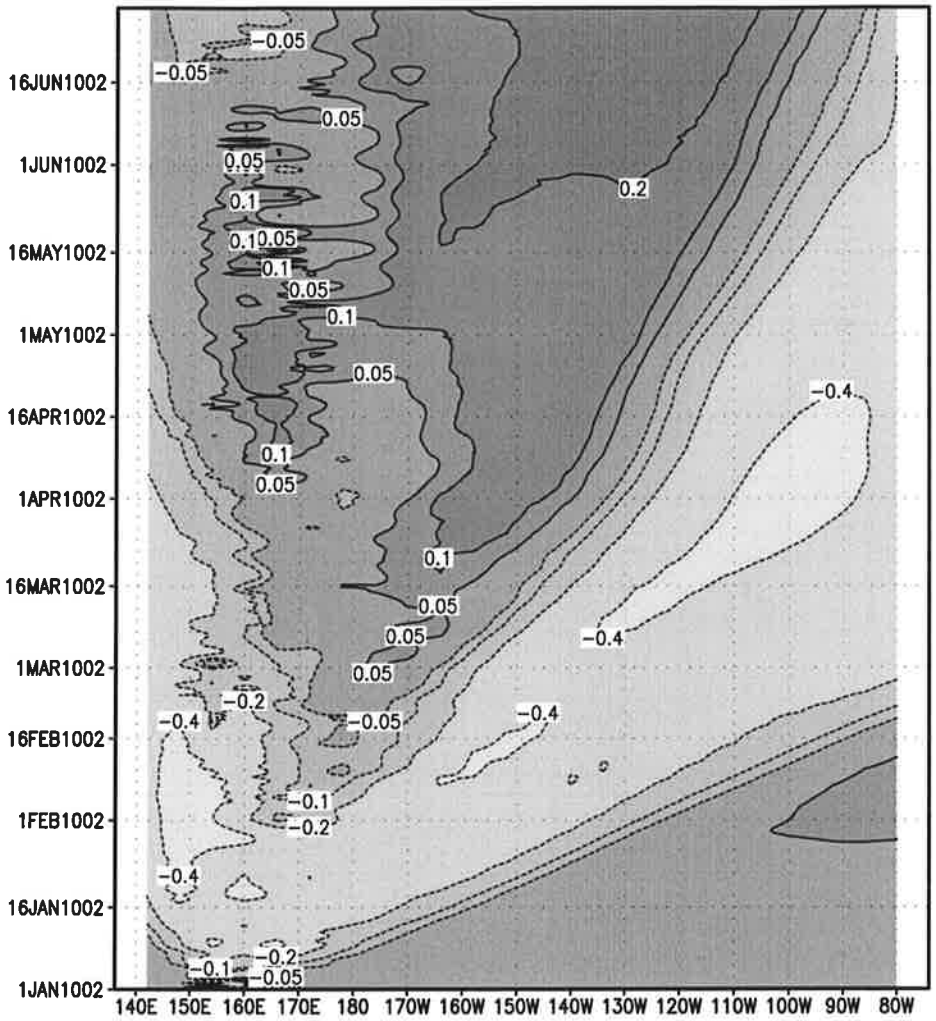


Fig.21: Longitude time plot along the equator showing the evolution of the sea level perturbation of Fig.20 (b). Clearly discernible are the two baroclinic modes propagating from West to East. The first mode traverses the basin in approximately 2-3 months, whereas the second mode needs about 4-5 months. The units of the sea level perturbation are in [m].

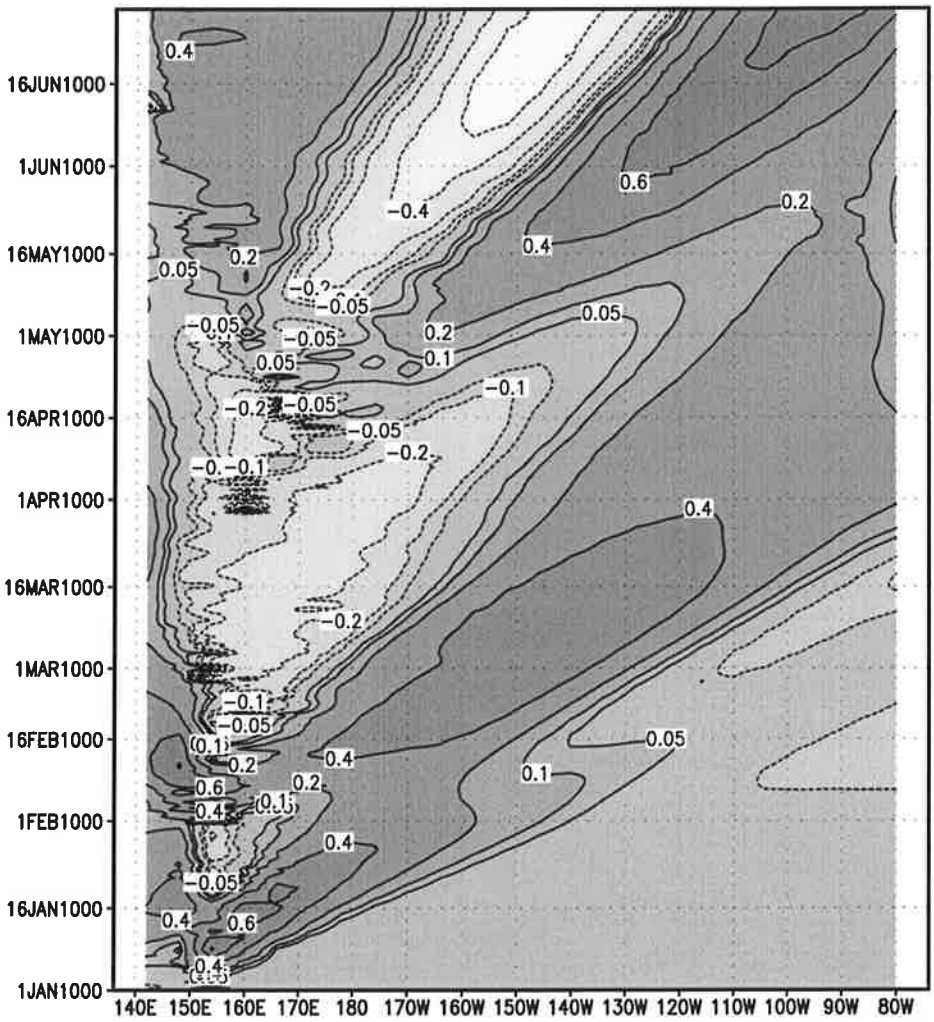
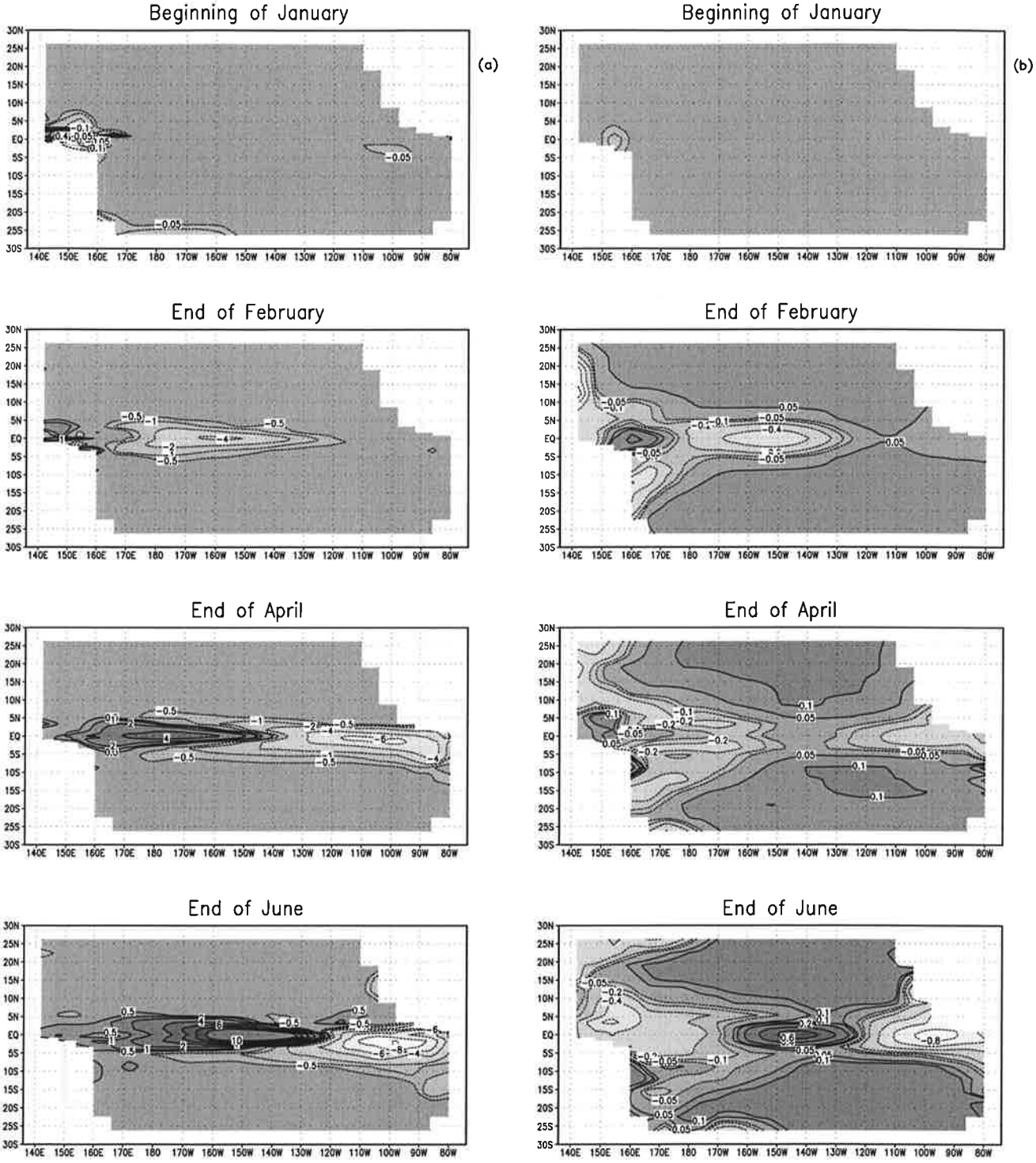


Fig.22: Longitude time plot along the equator analogous to Fig.21 showing the evolution of the sea level perturbation of an uncoupled singular vector optimised for 6 months (experiment no. 2 in Fig.19).


Beginning of January
(a)
Beginning of January
(b)
End of February
End of February
End of April
End of April
End of June
End of June

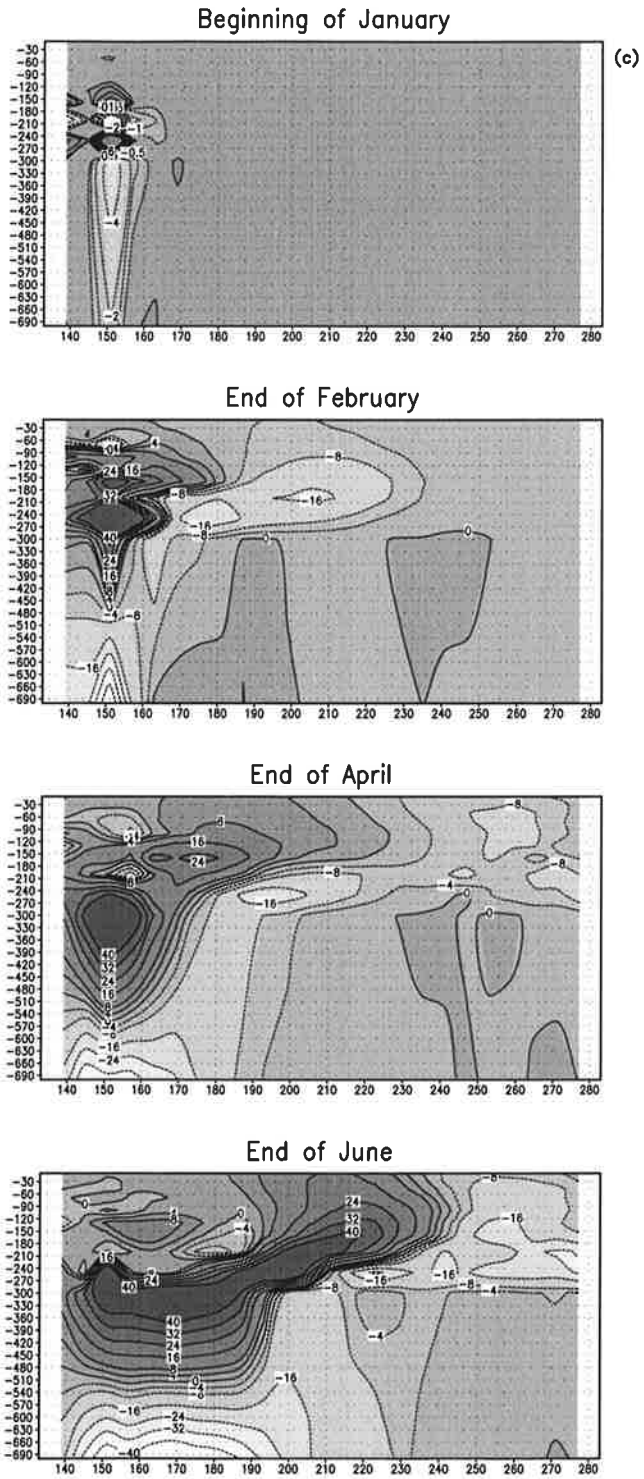


Fig.23: Leading singular vector of an uncoupled experiment started in January and optimised for 6 months. In (a) the time evolution of the surface temperature perturbation is shown, (b) depicts the evolution of the sea level perturbation, and (c) gives a vertical section along the equator of the temperature perturbation. The temperatures are in units of [K] and the sea level is in [m]. The remarks about scaling mentioned in Fig.20 apply here as well. This singular vector does not exhibit marked amplitudes in the eastern equatorial Pacific as was observed for the coupled case in Fig.20.

The time evolutions of the singular vectors have characteristic features. In both cases presented, the largest amplitudes are found initially in the western part of the ocean basin, the warm pool region. The perturbations gradually affect the central and eastern part of the model Pacifc during the singular vector's time evolution. This happens through the propagation of the first and the second baroclinic equatorial Kelvin wave modes propagate from West to East. The ellipsoidal spatial structure as well as the propagation speeds correspond well with those presented in Philander (1990). The propagation speeds inferred from Figs.21 and 22, agree well with the $2.4\frac{m}{s}$ and $1.4\frac{m}{s}$ for the first and the second baroclinic mode, respectively, given in the former reference. These propagation speeds correspond to a temperature profile representative for the central equatorial Pacific. The variable depth of the thermocline in the ocean general circulation model induces variable propagation speeds as seen in Figs.21 and 22. Having reached the eastern boundary, the equatorial Kelvin waves reflect into coastal Kelvin waves which propagate to the North and to the South along the boundaries. Some of the wave signal is reflected into Rossby waves which propagate westward. The coupled singular vector's SST signal is large in the eastern part of the ocean as well. Such a signal yields an anomalous westerly wind response in the central and eastern part of the Pacific region by virtue of the diagnostic atmosphere, Flügel (1994). It enhances the development of the warming simulated during the evolution of the coupled singular vector. This explains the much more organised structure of the coupled singular vector compared to the uncoupled run. In the latter case the lack of atmospheric feedback inhibits the amplification of a given signal. Therefore, the uncoupled singular vector does not show large SST amplitudes in the eastern Pacific. To summarise, the coupled singular vectors in most cases exhibit larger growth rates than the uncoupled ones, and their spatial structures are affected by the action of the diagnostic atmosphere.

The singular vectors are derived by linear calculations. If they are added to the HCM's initial state, they will evolve in a different way, depending on the degree of non-linearity of the HCM. Here, scaling of the singular vectors does matter. To compare the time evolution of a singular vector with its non-linear evolution after scaling and superposition onto the HCM's initial state, the experiment shown in Fig.24 was done. The singular vector shown in Fig.20 was scaled such that its largest amplitude in the temperature field was 1K. The SST of the scaled singular vector is shown in Fig.24 (a). The lower panel on the left depicts the final state of the linearly propagated signal, and the lower panel on the right shows the difference field of an integration performed with the scaled singular vector added to the initial state and the reference run. Even after 6 months integration time the

correspondence is remarkably good. Of course, the details visible in the final states differ, but the spatial sturctures closely resemble each other and the amplitudes attained at the end of the integration are in good aggreement as well.

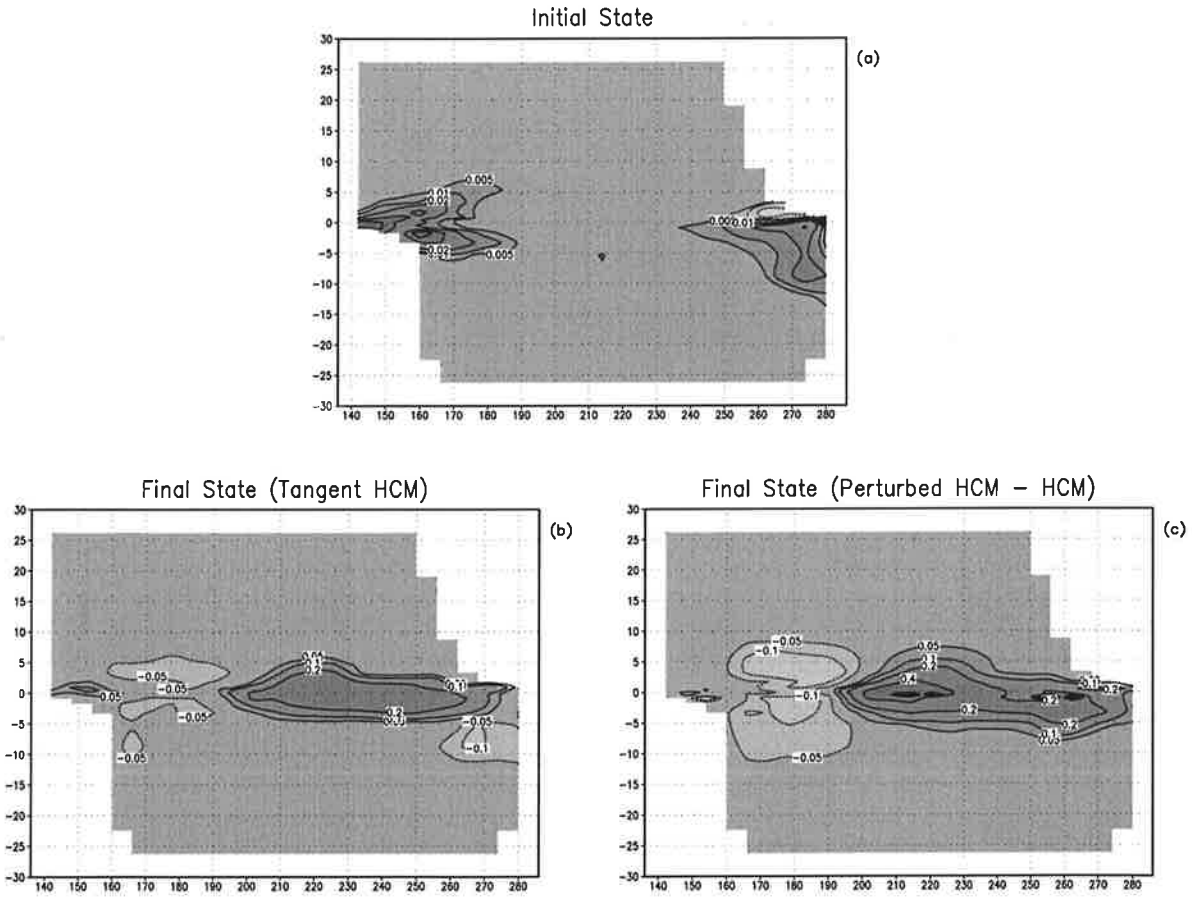


Fig.24: Comparison of non-linear and linear time evolution of the singular vector shown in Fig.20. The singular vector was scaled such that the maximum perturbation amplitude equals 1K; (a) the scaled singular vector, (b) the tangent linearly evolved structure, and (c) the non-linearly evolved structure obtained by subtrating the unperturbed intergration from the perturbed one.

The previous example of a singular vector's time evolution indicates that such a perturbation may well affect the outcome of an HCM ENSO forecast, since large portions of the eastern equatorial Pacific experience a warming. To investigate this further, several initial states during a model ENSO cycle were perturbed by their corresponding leading singular vectors. These were scaled to yield maximum perturbation amplitudes of -2K,-1K, 1K, and 2K in the oceanic temperature field. Fig.25 depicts the outcome of the experiment. The individual ensembles of perturbed model integrations show different spreads during integration time. The ensemble intialised during the peak phase of a model El Niño exhibits the largest spread while the trajectory starting during a peak La Niña phase is less affected. The transition phases from cold to warm ENSO states and from warm to cold ENSO states lie between the two extremes. However, the impacts on the forecasts

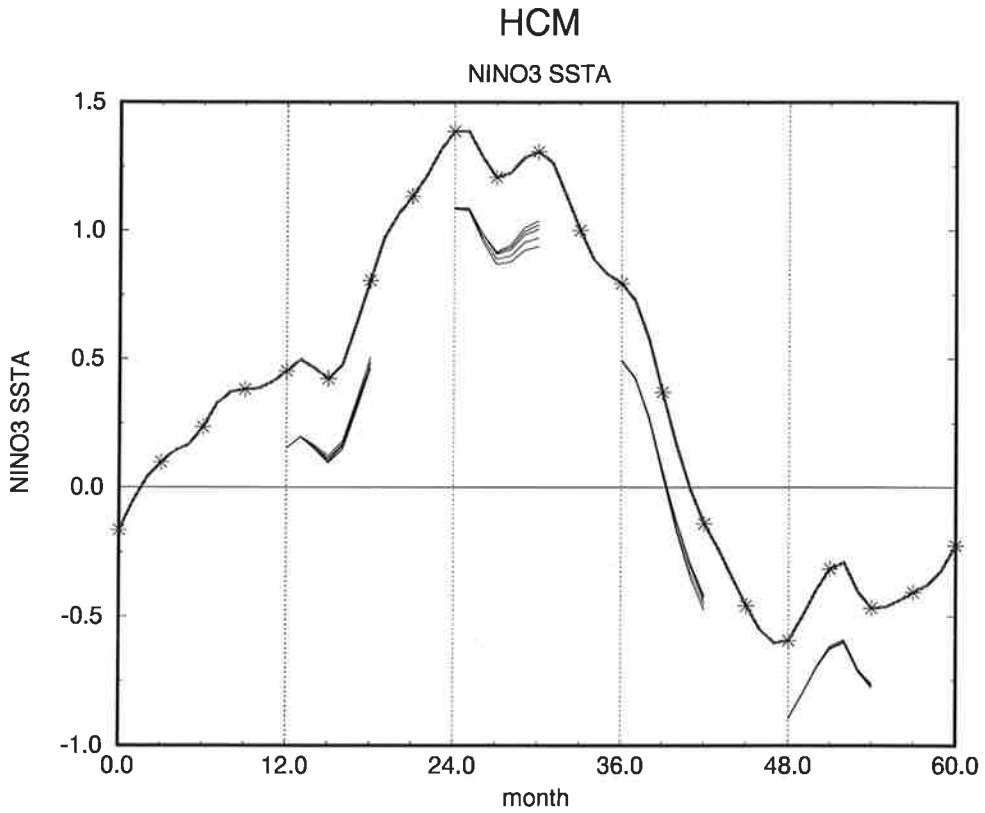


Fig.25: Effect of the leading singular vectors on the NINO3 index. Each ensemble of integrations shows the effect of adding the leading singular vector to the HCM. The maximum perturbation amplitudes in each ensemble were chosen to be -2K,-1K,1K, and 2K in the perturbation's temperature field. The ensemble of perturbed trajectories and the relevant piece of the reference run are displaced by -0.3K releative to the HCM NINO3 index.

are generally small, indicating that the model resides in a regime which exhibits relatively high predictability.

The extent to which an ENSO forecast will be affected by the addition of a singular vector for a specific forecast interval depends on several factors. In general, the perturbation amplitude will be chosen such that it reflects the forecasters uncertainty of the model's initial condition. This uncertainty can be included in the singular vector calculations by using the inverse of the initial error's covariance as a norm in the space of initial state perturbations. Further, the susceptibility to perturbations may vary, during the annual cycle, as well as during the ENSO cycle. Random perturbations like those considered in chapter III that perpetually disturb the ocean-atmosphere system induce a large variability among the oceanic initial states and alter the stability of the latter.

As a final set of experiments, singular vectors of the stochastically forced HCM (see chapter III) were computed. The stochastic forcing enters the HCM as an external forcing analogous to the annual cycle of wind stress. This implies that only minor modifications of the tangent linear and of the adjoint tangent linear HCM have to be made in order to include this additional element.

The restart month 276, Fig.8 p.22, was taken for a suite of singular vector calculations. Four different realisations of the stochastic wind stress component were chosen and the leading 10 singular vectors of each realisation were computed for an integration time of 6 months. Fig.26 shows the final states of the leading singular vectors belonging to the four different realisations of the stoachstic part of the wind stress field. The amplification rates as well as the final structures of SST differ considerably among the realisations. This leads to the conclusion that it might be a difficult task to determine an adequate set of perturbations for an ENSO forceast via singular vector analysis, since the singular vectors depend on the path taken by the coupled system during the forecast. They depend on the inevitably random nature of part of the atmospheric wind stress forcing. Singular vectors that yield significant perturbations may well exist, but their determination will require as many singular vector calculations as there are realistions of stochastic forcing in a forecast ensemble. As already mentioned in the concluding remarks to chapter III, the random wind stress fluctuations may have been over-estimated. If this was the case, chances are good that the sensitivity of the singular vectors to the random wind stress component will be less pronounced.

This study of initial error growth in an HCM and in a stochastically forced version of the latter by computing the leading singular vectors revealed different properties that are important for the quantification of initial error growth in the

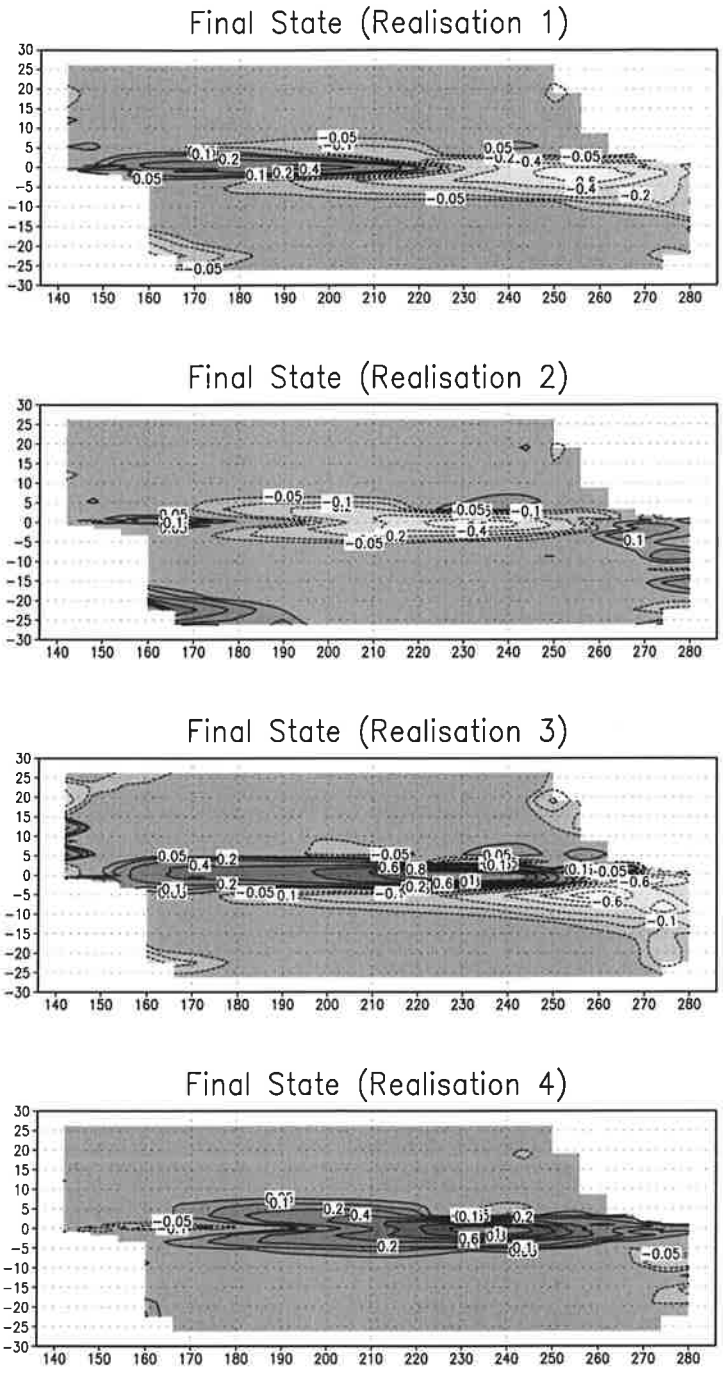


Fig.26: Final states of the leading singular vectors obtained for 4 different realisations of the stochastically forced HCM started at restart month 276. The amplification rates are, starting with the first realisation: 37.4, 33.5, 49.2, 34.4.

context of ENSO forecasting. The amplification rates of the singular vectors depended on the phases of the annual and ENSO cycles. The structures of the leading coupled singular vectors and their subsequent time evolution are affected by the atmospheric feedback. Singular vectors added to the non-linear HCM evolve in a similar way as the tangent linearly propagated ones. The effect on the outcome of a forceast is variable. It depends on the phase of the model ENSO cycle and, of course, on how well a given initial error field projects on the singular vectors.

The studies of Moore and Kleeman (1996), Chen et al. (1997), Xue et al. (1994), and Xue et al. (1997) show similarities as well as striking differences when compared to the results obtained here. The different degrees of complexity of the underlying forecast models as well as the diverse approaches to get estimates of the leading singular vectors and singular values, however, complicates an objective comparison between these studies. A common feature shared by all investigations is the dependence of results to the phases of the models' annual and ENSO cycles.

VI. Summary, Conclusions, and Outlook

Two aspects which are of relevance to the ENSO predictability limits were investigated studied in this thesis. Wind stress fluctuations not correlated to ENSO which perpetually interfere with the ENSO dynamics and the growth of initial uncertainties during a forecast were considered. Both aspects were studied with the help of a Hybrid Coupled Model (HCM) of ENSO and its extension to a stochastically forced HCM.

Random wind stress fluctuations not correlated to ENSO that are present in the tropical atmosphere are an important source of ENSO irregularity. Since the leading EOFs of these wind stress fluctuations exhibit large-scale structures that are likely to have an impact on the ENSO dynamcis, they were added to the HCM in order to mimic the observed random variability. This extension of the original HCM is termed stochastically forced HCM. As presented in chapter III, the fluctuations result in a considerable spread among forecast trajectories initialised with identical initial conditions. The variability simulated by the stochastically forced HCM is in qualitative aggreement with that observed. Spectra of the NINO3 index time series taken from observations and from the stochastically forced HCM show a high degree of similarity on interannual time scales. Restart ensembles during different phases of the model ENSO cycle were computed. A measure of predictability was defined by fitting linear stochastic models over time intervals of 36, 48, and 60 months in a reduced state space spanned by the leading EOFs of the pooled fields of SSTA, sea level anomalies, and anomalous zonal wind stress. This measure of predictability reflects the overlap of the time-dependent and of the stationary probability density functions (pdf) of the linear stochastic processes. In the early phases of the forecasts the overlap of both pdfs is small, leading to high predictability. As time goes by the overlap increases and the measure of predictability asymptotes to zero. In each case studied, predictability was lost before a typical cycle period was completed. For intermediate forecast times different evolutions of the predictability measure were obtained. In some cases a rather rapid decrease and a subsequent recovery is observed. After this local maximum in predictive skill is passed the predictability measure asymptotes to zero as in the other cases. Such a peculiar behaviour can readily be explained by the fact that the predictability measure incorporates the dynamics of the system. While the diffusion present in the stochastic process widens the time-dependent pdf, the dynamical part changes the relative position of the ensemble mean to the climatology. Therefore, large excursions of the ensemble mean relative to its stationary value can lead to increased predictability at intermediate forceast times.

The ensemble experiments performed with the stochastically forced HCM did not incorporate errors in the initial conditions. The latter is a further candidate for limited ENSO predictability. To get an estimate of the relevance of initial error growth, the strongest growing perturbations of the coupled model during a forecast were determined by singular vector analysis. The determination of the HCM's singular vectors requires the specification of the tangent linear HCM and the adjoint tangent linear HCM. Among the possible approaches a novel one was pursued: the tangent linear and the adjoint HCM required for the determination of the singular vectors and the singular values were generated by an automatic differentiation tool. Earlier studies either studied the singular vectors of low dimensional approximations to the forceast models or they coded the linearised model equations. The automatic differentiation tool is able to generate the linear codes after some minor modifications of the HCM's FORTRAN source code. Advantages of this approach consist in the high flexibility to generate new linear versions after changes in the HCM's code and the fact that the numerically exact tangent linear and adjoint tangent linear models are generated. The high dimensionality of the eigenvalue problem that has to be solved to obtain the singular vectors and values of the HCM precludes the determination of the complete eigenvalue spectrum. In the experiments presented in chapter IV the ten leading singular vectors were determined by employing the Lanczos algorithm. This algorithm iteratively aproximates the eigenvalues and eigenvectors of a symmetric linear operator. The linear operator only enters the algorithm by its action on a given input vector. The latter property can be exploited for the treatment of large sparse matrices and linear operators that are given as a computer program. A set of singular vector calculations was done during a model ENSO cycle for forecast intervals of 3, 6, 9, and 12 months. Additionally, the leading ten singular vectors of the uncoupled ocean model during an annual cycle for the same forecast intervals as in the coupled cases and for four starting dates during one year (January, April, July, and October) were determined. The amplification rates of the singular vectors showed a marked dependence on the ocean model's annual cycle as well as on the HCM's ENSO cycle. Largest amplification rates are obtained for forecasts started in January and in April, whereas the starts in July and in October showed lower amplification rates. Amplification rates were higher in the coupled cases than in the uncoupled cases. The spatial structures observed during the time evolution of the leading singular vectors correspond to equatorial wave dynamics. The perturbation's largest amplitudes are observed in the western part of the model ocean basin. During integration the perturbation gradually affects the central and the eastern parts of the ocean. Large perturbation amplitudes are observed in the deeper ocean layers

as well. The region where the model ocean's thermocline is located shows large temperature signals. This feature compares well with actual observations of the evolution of anomalous equatorial temperature signals by the TOGA/TAO array; NOAA (1998). A third set of singular vector calculations was carried out with the stochastically forced HCM. Four different realisations of the stochastic wind stress field were taken and the correpsonding leading ten singular vectors optimised for 6 months integration time were determined. The four realisations have different amplification rates and they evolve into remarkably different final states.

The tangent linear time evolution of the singular vectors compares well with a perturbed run of the non-linear HCM subtracted from the reference run in the SST field. Therefore, a superposition of the leading singular vectors onto an HCM's initial state will lead to similar perturbation patterns as in the linear integrations. To assess the effect such perturbations might have on the outcome of a forecast, the leading singular vectors optimised for 6 months were added to initial states of the HCM in different phases of the model ENSO cycle. Two peak phases, El Niño and La Niña, and two transition phases, from warm to cold ENSO states and from cold to warm ENSO states were taken. The singular vectors were scaled such that the maximum perturbation amplitude in the oceanic temperature field amounted to -2K,-1K,1K, or 2K. The effect on the NINO3 time series was variable. The perturbations were most effective during the El Niño phase and least effective during the La Niña phase. This experiment shows that singular vectors have the potential to affect the outcome of an ENSO forecast. The appropriate scaling of the singular vectors in the context of actual forecasts, however, will depend on the error attributed to the initial conditions of the forcast model. This error structure can be incorporated into the determination of the singular vectors by using as a norm in the space of initial state perturbations the inverse of the error covariance. To conclude, the singular vector calculations of chapter IV showed that it is feasible to determine the most strongly growing perturbations by using the numerically exact tangent linear and adjoint tangent linear versions of the forecast model. Since the adjoint of the forecast model can as well be used for data assimilation, an operational ENSO forecast model may use the leading singular vectors for the generation of perturbations of the model's initial state to assess the likely spread of the forecasts.

The singular vectors and their associated singular values are model dependent. It is advisable to be cautious in transferring the results obtained to actual dynamical processes in the real ocean–atmosphere system. Most importantly, singular vector analysis gives hints to the requirements of a given model to yield better forecasts, and it enables the foecaster to assess the effect initial errors will have on

the outcome of a forecast.

Comparing both processes studied in this thesis, the stochastic wind stress fluctuations are likely to be the dominant agent in limiting ENSO predictability. The moderate effect obtained by adding the leading singular vectors to initial states of the HCM supports this conjecture. Additionally, the spread among forecast trajectories when random wind stress forcing is included are sustantially larger than than those obtained by adding singular vectors to initial states of the HCM. However, the latter finding strongly depends on the actual amplitude of the singular vector added to the initial state of the HCM.

Three directions of possible future research should be mentioned. First, the singular vector studies begun here could be pursued further by e.g. analysing the singular vector dynamics in the ocean general circulation model in more detail. The processes enabling singular vector growth can be identified by analysing the time evolution of oceanic variables in the tangent linear HCM. Second, the HCM concept allows an efficient way to parameterise the atmospheric response to an anomalous oceanic circulation in the tropical regions. For the singular vector study the adjoint of the HCM was generated. An efficient way to generate an HCM that optimally predicts the state of ENSO can be obtained by optimising the entries of the atmospheric feedback matrix such that a maximal prediction skill is attained. The regression matrix may serve as a first guess for an optimisation of a cost function that measures the misfit between observations and model output over the forcast interval. The adjoint model enables the determination of the gradient of the cost function with respect to the entries of the atmospheric feedback matrix. Third, atmospheric random variability induces substantial irregularity in the recurrence of ENSO extremes and in the amplitudes of the latter. This random element was incorporated into the HCM by simply parameterising it by statistical means. The same holds for the determination of the atmospheric response to given SSTA. The question arises to which extent results will carry over to a dynamically based representation of the atmospheric dynamics. An atmospheric model of intermediate complexity coupled to an ocean general circulation model would allow to address questions of this kind. Especially the singular vector structure of such a coupled model would present an interesting area of research, since the nature of the modeled coupled dynamics would reflect itself in the singular vectors.

Appendices

1. Appendix to Chapter III

Restart Month	Dimension	Time Interval	Period	e-folding Time
276	4	36 48 60	60.9 60.7 62.6	25.5 26.6 24.8
276	5	36 48 60	59.8 62.1 63.0	33.4 30.2 27.2
276	6	36 48 60	60.3 62.6 63.5	37.1 31.8 27.4
372	4	36 48 60	27.3 23.8 23.0	8.0 6.4 7.1
372	5	36 48 60	27.4 23.6 24.1	7.4 7.3 7.6
372	6	36 48 60	27.3 23.8 24.4	6.6 6.8 6.8
408	4	36 48 60	47.1 50.8 52.1	11.4 12.1 11.5
408	5	36 48 60	45.8 50.3 50.2	10.0 11.8 10.5
408	6	36 48 60	46.6 56.6 51.4	10.8 12.7 11.2

Table A1: Periods and e-folding times of the interannual eigenmodes of the fitted system matrices belonging to the stochastically forced HCM's restart ensembles.

Restart Month	Dimension	Time Interval	Period	e-folding Time
420	4	36 48 60	51.4 57.0 66.1	11.9 12.4 17.4
420	5	36 48 60	59.7 61.7 63.9	13.7 13.1 16.1
420	6	36 48 60	59.2 61.5 64.7	13.1 12.9 16.0
468	4	36 48 60	112.0 66.8 58.6	11.3 14.1 16.6
468	5	36 48 60	124.3 69.1 60.5	11.9 15.4 19.0
468	6	36 48 60	121.3 67.2 61.2	12.9 18.0 21.0
552	4	36 48 60	302.1 99.7 102.5	7.1 13.8 15.1
552	5	36 48 60	233.6 107.6 112.3	7.2 15.5 16.6
552	6	36 48 60	102.5 99.1 103.9	9.3 15.7 16.9

– Table A1 continued –

Rest. Month	Dim.	Time Interv.	EOF1 (sim.)	EOF1 (dat.)	EOF2 (sim.)	EOF2 (dat.)	Dim. EOFs (sim.)	Dim. EOFs (dat.)
276	4	36	11.80	10.66	1.17	1.21	14.81	13.85
		48	9.24	7.93	1.11	1.18	12.16	11.05
		60	8.41	7.03	1.07	1.10	11.29	10.04
276	5	36	10.79	10.66	1.11	1.21	14.37	14.49
		48	8.47	7.93	1.04	1.18	11.94	11.65
		60	7.84	7.03	1.02	1.10	11.24	10.59
276	6	36	10.80	10.66	1.01	1.21	14.86	15.03
		48	8.44	7.93	1.03	1.18	12.39	12.18
		60	7.79	7.03	1.01	1.10	11.66	11.10
372	4	36	2.49	2.62	0.98	1.09	4.21	5.65
		48	3.27	4.75	0.77	1.00	5.26	7.45
		60	3.98	6.81	0.78	1.08	5.91	9.82
372	5	36	2.91	2.62	0.95	1.09	5.71	6.21
		48	4.00	4.75	0.92	1.00	6.82	8.22
		60	5.02	6.81	0.92	1.08	7.88	10.38
372	6	36	2.75	2.62	0.95	1.09	6.00	6.70
		48	3.76	4.75	0.91	1.00	7.04	8.76
		60	4.73	6.81	0.91	1.08	8.04	10.89
408	4	36	7.09	8.21	1.03	1.10	9.97	11.22
		48	7.06	7.32	1.05	1.16	9.95	10.42
		60	7.00	7.41	1.05	1.14	9.90	10.50
408	5	36	6.90	8.21	1.00	1.10	10.25	11.78
		48	6.91	7.32	1.03	1.16	10.30	10.99
		60	6.84	7.41	1.01	1.14	10.21	11.05
408	6	36	6.91	8.21	1.00	1.10	10.74	12.30
		48	6.91	7.32	1.03	1.16	10.77	11.48
		60	6.82	7.41	1.02	1.14	10.64	11.54

Table A2: Variances simulated by the fitted stochastic processes (sim.) and variances of the input data (dat.).

Rest. Month	Dim.	Time Interv.	EOF1 (sim.)	EOF1 (dat.)	EOF2 (sim.)	EOF2 (dat.)	Dim.EOFs (sim.)	Dim.EOFs (dat.)
420	4	36	3.86	3.50	1.04	1.06	5.64	6.46
		48	4.17	5.45	1.01	1.00	5.93	8.32
		60	4.37	7.25	1.00	1.08	6.12	10.29
420	5	36	4.93	3.50	0.97	1.06	8.02	7.02
		48	5.16	5.45	0.97	1.00	8.30	8.83
		60	5.21	7.25	0.96	1.08	8.37	10.87
420	6	36	4.81	3.50	0.98	1.06	8.37	7.52
		48	5.06	5.45	0.97	1.00	8.68	9.35
		60	5.11	7.25	0.95	1.08	8.71	11.36
468	4	36	8.46	9.55	0.99	1.19	11.25	12.66
		48	8.40	9.61	1.03	1.19	11.26	12.75
		60	8.11	7.96	1.06	1.20	11.01	11.13
468	5	36	8.57	9.55	0.97	1.19	11.84	13.24
		48	8.44	9.61	1.01	1.19	11.79	13.39
		60	8.12	7.96	1.04	1.20	11.50	11.69
468	6	36	8.57	9.55	0.96	1.19	12.31	14.05
		48	8.49	9.61	1.01	1.19	12.32	13.90
		60	8.13	7.96	1.04	1.20	11.98	12.22
552	4	36	3.13	5.88	0.92	0.99	5.86	8.83
		48	4.21	7.63	0.94	1.14	6.99	10.72
		60	4.82	9.18	0.96	1.13	7.62	12.24
552	5	36	3.10	5.88	0.90	0.99	6.31	9.41
		48	4.49	7.63	0.93	1.14	7.78	11.33
		60	5.23	9.18	0.96	1.13	8.56	12.87
552	6	36	3.16	5.88	0.89	0.99	6.81	9.89
		48	4.28	7.63	0.92	1.14	8.00	11.87
		60	4.93	9.18	0.94	1.13	8.69	13.38

– Table A2 continued –

2. Appendix to Chapter IV

2.1. The Lanczos Algorithm

Lanczos (1950) proposed an algorithm for computing a tridiagonal matrix T orthogonally congruent to a given symmetric (n,n)-matrix A, i.e. $T = Q^T A Q$ and $Q^T = Q^{-1}$.

Subsequent investigations revealed that the numerical treatment of the algorithm poses some problems, see Parlett and Scott (1979) or Stoer (1990). Software packages now available incorporate efficient means to circumvent the undesirable features of the algorithm and to exploit its strengths. In this thesis the *laso*-algorithm described by Parlett and Scott (1979) is used.

The mathematical foundation of the Lanczos algorithm consists in the following theorem:

Given a symmetric (n,n)-matrix A and some non-trivial n-vector q of unit length, the sequence of vectors $q_1, ..., q_i$ obtained by the recursion

$$q_1 = q \quad \gamma_1 q_0 := 0 \tag{A1}$$

$$A q_i = \gamma_i q_{i-1} + \delta_i q_i + \gamma_{i+1} q_{i+1} \qquad (i \geq 1) \tag{A2}$$

with the real scalars δ_i and γ_{i+1} and the vector q_{i+1} being given by

$$\delta_i = q_i^T A q_i \tag{A3}$$

$$\gamma_{i+1} = ||r_i|| \quad \text{and} \quad r_i = A q_i - \delta_i q_i - \gamma_i q_{i-1} \tag{A5}$$

$$q_{i+1} = r_i / \gamma_{i+1} \tag{A6}$$

forms an orthonormal basis of the vector space spanned by

$$q, Aq, A^2 q, ..., A^{i-1} q \quad . \tag{A7}$$

The above recursion terminates at some $m \leq n$; $\gamma_{m+1} = 0$. Written in matrix notation, one gets:

$$A Q_i = Q_i T_i + \gamma_{i+1} (0, \ldots, 0, q_{i+1}) \quad i = 1, ..., m \quad . \tag{A8}$$

The matrix Q_i is formed by $q_1, \ldots, q_i$ as column vectors, and T_i reads

$$T_i = \begin{pmatrix} \delta_1 & \gamma_2 & & 0 \\ \gamma_2 & \delta_2 & \ddots & \\ & \ddots & \ddots & \gamma_i \\ 0 & & \gamma_i & \delta_i \end{pmatrix}$$

Especially for $i = m$ this implies , since $\gamma_{m+1} = 0$,

$$AQ_m = Q_m T_m \quad . \tag{A9}$$

The proof of this theorem proceeds by induction over i and is given in e.g. Stoer (1990).

As an immediate consequence, this theorem provides a transformation of the eigenvalue problem

$$Ax = \lambda x \tag{A10}$$

into an eigenvalue problem for a tridiagonal matrix T. Further investigation reveals that the largest respectively smallest eigenvalues of A are well approximated by those of T_i for relatively small i, so that the above sequence of vectors $q_1, \ldots, q_i$ is not required to be computed until the index m is reached. The symmetric matrix A enters the algorithm only by its action on an input vector q_i at step i of the above recursion. Due to these facts, the Lanczos-algorithm is well suited for eigenvalue problems involving large sparse matrices that require special forms of internal representation on a computer or problems that involve linear operators that are not given as matrices but by a computer program. The eigenvalue problem that had to be solved in chapter IV belongs to the latter class.

Acknowledgements

I'd like to thank all those people who directly and/or indirectly assisted in the process of evolution of this thesis. Out of the long list I only mention a few, apologising that I did not select more. Starting with Klaus Hasselmann and Mojib Latif who planted the seed and enabled its growth in the exceptionally pleasant atmosphere of the Max-Planck-Institut für Meteorologie. Going on to Moritz Flügel who provided an early version of the Hybrid Coupled Model, Ralf Giering who gave advice during the use of the TAMC, and Matthias Münnich who often added fertilising suggestions and questions. Last but not least, I thank my wife Caroline for her permanent support.

I thank you all.

Christian Eckert

Bibliography

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

Barnett, T.P., M. Latif, N. Graham, M. Flügel, S. Pazan, and W. White, 1993, ENSO and ENSO-related Predictability. Part I: Prediction of Equatorial Pacific Sea Surface Temperature with a Hybrid Coupled Ocean–Atmosphere Model, *J. Climate*, **6**, 1545-1566.

Barnston, A.G., H.M. van den Dool, S.E. Zebiak, T.P. Barnett, M. Ji, D.R. Rodenhuis, M.A. Cane, A. Leetma, N.E. Graham, C.R. Ropelewski, V.E. Kousky, E.A. O'Lenic, and R.E. Livezey, 1994, Long-Lead Seasonal Forecasts – Where do we stand?, *Bull. Amer. Meteor. Soc.*, 2097-2114.

Battisti. D.S., A.C. Hirst, 1989, Interannual variability in a tropical atmosphere-ocean model: influence of the basic state, ocean geometry, and nonlinearity, *J. Atmos. Sci.*, **46**, 1687-1712.

Bjerknes, J., 1966, A possible response of the atmospheric Hadley circulation to equatorial anomalies of ocean temperature, *Tellus*, **18**, 820-828.

Bjerknes, J., 1969, Atmospheric teleconnections from the equatorial Pacific, *Mon. Weather. Rev.*, **97**, 163-172.

Blanke, B., J.D. Neelin, and D. Gutzler, 1997, Estimating the Effect of Stochastic Wind Stress Forcing on ENSO Irregularity, *J. Climate*, **10**, 1473-1486.

Blumenthal, B., 1991, Predictability of a Coupled Ocean–Atmosphere Model, *J. Climate*, **4**, 766-784.

Buizza, R., J. Tribbia, F. Molteni, and T. Palmer, 1993, Computation of optimal unstable structures for a numerical weather prediction model, *Tellus*, **45A**, 388-407.

Chang, P., B. Wang, T. Li, and L. Ji, 1994, Interactions between the Seasonal Cycle and the Southern Oscillation – Frequency Entrainment and Chaos in an Intermediate Coupled Ocean–Atmosphere Model, *Geophys. Res. Lett.*, **21**, 2817-2820.

Chen, Y.-Q., D.S. Battisti, T.N. Palmer, J. Barsugli, and E.S. Sarachik, 1997, A Study of the Predictability of Tropical Pacific SST in a Coupled Atmosphere/Ocean Model Using Singular Vector Analysis: The Role of the Annual Cycle and the ENSO Cycle, *Mon. Wea. Rev.*, **125**, 831-845.

CPC, 1998a, Climate Prediction Center, Data & Indices, available from: *http://nic.fb4.noaa.gov/data/cddb/cddb/sstoi.indices.*

CPC, 1998b, Climate Prediction Center, Experimental Long-Lead Forecast Bulletins, available from: *http://nic.fb4.noaa.gov:80/products/predictions/experimental/bulletin/.*

Dettinger, M.D., M. Ghil, C.M. Strong, W. Weibel, and P. Liou, 1995, Software Expedites Singular Spectrum Analysis of Noisy Time Series, *EOS*, **76**, No. 2, January 1995.

Eckert, Ch., M. Latif, 1997, Predictability of a Stochastically Forced Hybrid Coupled Model of El Niño, *J. Climate*, **10**, 1488-1504.

Farrel, B.F., P.J. Ioannou, 1996a, Generalized Stability Theory. Part I: Autonomous Operators, *J. Atmos. Sci.*, **53**, 2025-2040.

Farrel, B.F., P.J. Ioannou, 1996b, Generalized Stability Theory. Part II: Nonautonomous Operators, *J. Atmos. Sci.*, **53**, 2041-2053.

Flügel, M., 1994, Untersuchungen zu Mechanismus und Vorhersagbarkeit von ENSO mit einem vereinfachten gekoppelten Ozean–Atmosphäre–Modell, Ph.D. Thesis, Univ. Hamburg (Germany), 67 pp.

Fraedrich, K., Ch. Ziehmann-Schlumbohm, 1993, Predictability Experiments with Persistence Forecasts in a Red Noise Atmosphere, *Journal*, **Vol.**, pages.

Giering, R., 1996, *Erstellung eines Adjungierten Modells zur Assimilierung von Daten in ein Modell der globalen ozeanischen Zirkulation*, Ph.D. Thesis, University of Hamburg. [Also available as Examensarbeit Nr.44, ISSN: 0938-5177, Max-Planck-Institut für Meteorologie, Bundesstraße 55, D-20146 Hamburg.]

Giering, R., 1997, The TAMC Manual, [Available from R. Giering, Massachusetts Institute of Technology, Department of Earth and Planetary Sciences, Center for Global Change Science, Cambridge, MA 02139 USA]

Giering, R., and T. Kaminski, 1996, Recipes for Adjoint Code Construction, *Transactions on Mathematical Software*, in press. [Also available as Report No. 212, ISSN 0937-1060, Max-Planck-Institut für Meteorologie, Bundesstraße 55, D-20146 Hamburg, GERMANY.]

Gill, A.E., 1982, Atmosphere–Ocean Dynamics, *Academic Press Inc.*, pp. 662.

Goldenberg, S., and J. O'Brien, 1981, Time and Space Variability of the Tropical Pacific Wind Stress, *Mon. Wea. Rev.*, **109**, 1190-1207.

Goswami, B.N., and J. Shukla, 1991, Predictability of a Coupled Ocean–Atmosphere Model, *J. Climate*, **4**, 3-22.

Griewank, A., and G.F. Corliss (Eds.), 1991, *Automatic Differentiation of Algorithms: Theory, Implementation and Application*, SIAM, Philadelphia, Penn. .

Haney, R.H., 1971, Surface Thermal Boundary Conditions for Ocean Circulation Models, *J. Phys. Oceangr.*, **1**, 241-248.

Hasselmann, K., 1976, Stochastic Climate Models. Part I: Theory, *Tellus*, **28**, 289-305.

Hasselmann, K., 1988, PIPs and POPs: The Reduction of Complex Dynamical Systems Using Principal Interaction and Oscillation Patterns, *J. Geophys. Res.*, **93**, 11015-11021.

Honerkamp, J., 1990, *Stochastische Dynamische Systeme*, VCH WEinheim, ISBN: 3-527-27945-8, pp. 340.

Ji, M., A. Leetma, and J. Derber, 1995, An Ocean Analysis System for Seasonal to interannual climate studies, *Mon. Wea. Rev.*, **123**, 460-481.

Jin, F.-F., D. Neelin, and M. Ghil, 1994, El Niño on the Devil's Staircase: Annual Subharmonic Steps to Chaos, *Science*, **246**, 70-72.

Kendall, Sir M., A. Stuart, and J.K. Ord, 1983, *The Advanced Theory of Statistics*, Fourth Ed., Vol. 3, Charles Griffin & Company Limited, London & High Wycombe, p. 283 f. .

Kleeman, R., 1993, On the Dependence of Hindcast Skill on Ocean Thermodynamics in a Coupled Ocean-Atmosphere Model, *J. Climate*, **6**, 2012-2033.

Kleeman, R., and S.B. Power, 1994, Limits to Predictability in a Coupled Ocean–Atmosphere Model due to Atmospheric Noise, *Tellus*, **46A**, 529-540.

Kleeman, R., and A. Moore, 1997, A Theory for the Limitation of ENSO Predictability Due to Stochastic Atmospheric Transients, *J. Atmos. Sci.*, **54**, 753-767.

Lanczos, C., 1950, An Iteration Method for the Solution of the Eigenvalue Problem of Linear Differential and Integral Operators, *J. Res. National Bureau of Standards*, **45**, 255-282.

Latif, M., 1987, Tropical Ocean Circulation Experiments, *J. Phys. Oceanogr.*, **17**, 246-263.

Latif, M., D. Anderson, T. Barnett, M. Cane, R. Kleeman, A. Leetma, J. O'Brien, A. Rosati, and E. Schneider, 1998, TOGA Review Paper: "Predictability and Prediction", *J. Geophys. Res.*, in press.

Latif, M., 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, *Clim. Dyn.*, **9**, 167-179.

Lorenz, E.N., 1965, A Study of the Predictability of a 28-variable Atmospheric Model, *Tellus*, **17**, 321-333.

Moore, A., and R. Kleeman, 1996, The Dynamics of Error Growth in a Coupled Model of ENSO, *Quart. J. of the Royal Met. Soc.* , **122**, 1405-1446.

Moore, A., and R. Kleeman, 1997a, The Singular Vectors of a Coupled Ocean–Atmosphere Model of ENSO. I.: Thermodynamics, Energetics, and Error Growth, *Quart. J. of the Royal Met. Soc.* , **123**, 953-981.

Moore, A., and R. Kleeman, 1997b, The Singular Vectors of a Coupled Ocean–Atmosphere Model of ENSO. I.: Sensitivity Studies and Dynamical Interpretation, *Quart. J. of the Royal Met. Soc.* , **123**, 983-1006.

NAG Fortran Library Manual Mark 15, 1991, (Numerical Algorithms Group, Wilkinson House, Jordan Hill Road, Oxfod, UK).

NCEP, 1998, NCEP Daily Global Analyses, available from: *http://www.cdc.noaa.gov/cdc/data.nmc.html.*

Neelin, J.D., 1989, Interannual Oscillations in an ocean general circulation model coupled to a simple atmosphere model, *Trans. Roy. Phil. Soc.* , **329A**, 189-205.

Neelin, J.D., 1990, A hybrid coupled general circulation model for El Niño Studies, *J. Atmos. Sci.* , **47**,674-693.

Neelin, J.D., M. Latif, and F.-F. Jin, 1994, Dynamics of Coupled Ocean–Atmosphere Models: The Tropical Problem, *Ann. Rev. Fluid Mech.*, **26**, 617-659.

NOAA, 1998, The El Niño Theme Page, TOGA/TAO Data, available from: *http://www.pmel.noaa.gov/toga-tao/el-nino/home.html.*

Parker, D.E., C.K. Folland, A. Bevan, M.N. Ward, M. Jackson, and K. Maskell, 1994, *Marine Surface Data for Analysis of Climatic Fluctuations on Interannual to Century Time Scales*, National Research Council: Natural Climate Variability on Decade to Century Time Scales, Eds. D.G. Martinson, K. Bryan, M. Ghil, M.M. Hall, T.R. Karl, E.S. Sarachik, S. Soroostian, and L.F. Talley, National Academy Press, Washington.

Pacanowski, R.C., S.G.H. Philander, 1981, Parameterization of Vertical Mixing in Numerical Models of Tropical Oceans, *J. Phys. Oceanogr.*, **11**, 1143-1151.

Palmer, T.N., R. Gelaro, J. Barkmeijer, and R. Buizza, 1998, Singular Vectors, Metrics, and Adaptive Observations, *J. Atmos. Sci.*, **55**, 633-653.

Parlett, B.N., D.S. Scott, 1979, The Lanzos Algorithm with Selective Orthogonalization, *Mathematics of Computation*, **33**, 217-238.

Penland, C., P.D. Sardeshmukh, 1995, The optimal Growth of Tropical Sea Surface Temperature Anomalies, *J. Climate*, **8**, 1999-2024.

Philander, S.G.H., 1990, *El Niño, La Niña, and the Southern Oscillation*, Academic Press, San Diego, pp. 293.

Reynolds, R.W., 1988, A Real-Time Global Sea Surface Temperature Analysis, *J. Climate*, **1**, 75-86.

Reynolds, R.W., T.M. Smith, 1994, Improved Global Sea Surface Temperature Analyses Using Optimum Interpolation, *J. Climate*, **7**, 929-948.

Stoer, J., R. Bulirsch, 1990, *Numerische Mathematik*, Vol. 2, Springer, Heidelberg, pp. 341.

Suarez, M.J., P.S. Schopf, 1988, A delayed action oscillator for ENSO, *J. Atmos. Sci.*, **45**, 549-566.

Talagrand, O., 1991, The Use of Adjoint Equations in Numerical Modeling of the Atmospheric Circulation, in: A. Griewank and G. Corliss (Eds.), *Automatic Differentiation of Algorithms: Theory, Implementation and Application*, SIAM, Philadelphia, Penn., p. 169-180.

Tziperman, E., L. Stone, M.A. Cane, and H. Jarosh, 1994, El Niño Chaos: Overlapping of Resonances between the Seasonal Cycle and the Pacific Ocean–Atmosphere Oscillator, *Science*, **264**, 72-74.

van Kampen, N.G., 1992, *Stochastic Processes in Physics and Chemistry*, North-Holland, 210ff.

van Oldenborgh, G.J., G. Burgers, S. Venzke, C. Eckert, and R. Giering, 1997, Tracking Down the ENSO Delayed Oscillator with an Adjoint OGCM, *Mon. Wea. Rev.*, submitted.

von Storch, H., G. Bürger, R. Schnur, and J.-S. von Storch, 1995, Principal Oscillation Patterns: A Review, *J. Climate*, **3**, 377-400.

Xue, Y., M.A. Cane, S.E. Zebiak, M.B. Blumenthal, 1994, On the Prediction of ENSO: A Study with a Low Order Markov Model, *Tellus*, **46A**, 512-528.

Xue, Y., M.A. Cane, S.E. Zebiak, 1997, Predictability of a Coupled Model of ENSO using Singular Vector Analysis. Part I.: Optimal Growth in Seasonal Background and ENSO Cycles, *Mon. Wea. Rev.*, **125**, 2043-2056.

Zebiak, S.E., M.A. Cane, 1987, A Model El Niño/Southern Oscillation, *Mon. Wea. Rev.*, **115**, 1263-1271.