Construction of Low Dimensional Models of El Niño-Southern Oscillation Using Empirical Orthogonal Functions

Thesis by

Mark Stephen Roulston

In Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy

California Institute of Technology

Pasadena, California

United States of America

2000

(Submitted November 22, 1999)

Acknowledgements

First I must thank my *de jure* and *de facto* advisors, Yuk Yung and David Neelin respectively.

Thanks go to Andy Ingersoll for general discussions on the research as well as for introducing me to David Neelin and urging me to go to the GFD Summer School at Woods Hole in 1998, which was very useful and stimulating.

The research in this thesis was 99% computational so I would have been up a certain creek without a certain implement on many occasions if it weren't for Mike Black. Equally vital support came from Kay, Irma and Tammie.

No acknowledgements would be complete without thanking past and present office mates: Eric, Mark, Ian, Chris, Adam, Sarah and Matt. They provided many hours of occasionally useful but generally fruitless discussions. And extra thanks to Matt for bringing me up to speed on two centuries of linguistic vandalism. The afternoon coffee crowd are thanked for conversation almost as stimulating as the caffeine itself and the Friday Beer crowd for, well, being the Friday Beer crowd. And finally to my adopted "family" of Anthony, Mark, Sue, Albert and Ilona.

Abstract

The main theme of this thesis is the construction and analysis of low dimensional dynamical models of El Niño-Southern Oscillation (ENSO). Low dimensional models of ENSO have been constructed [Val86, WF96, Jin97b] but all the models were built on *ad hoc* assumptions and simplifications. The low dimensional models in this thesis were constructed from a higher dimensional intermediate coupled model (ICM) by Galerkin projection on truncated basis sets of empirical orthogonal functions (EOFs), a method that has previously been applied to atmospheric models and models of turbulence [Sel95, HLB96]. This method makes no simplications to the physics of the ICM but instead projects it into a lower dimensional subspace of its full state space that has been *empirically shown* to contain the maximum amount of variance of any subspace of the same dimension.

Analysis of the reduced models shows that the variability of the full model can be explained by three types of mode; I. self-sustaining, II. linearly stable but nonlinearly coupled to a self-sustaining mode, III. linearly stable and excited by noise. When driven by noise the stable modes can couple to the primary modes, and this provides an additional pathway for noise to perturb these leading modes.

The implication is that the behavior of the model without stochastic forcing can be well reproduced by the low dimensional models obtained by projection onto the leading EOFs but that many more EOFs are required to properly model the response to stochastic forcing.

Experiments with the full model also indicate that the model is not efficient at trans-

ferring power from high (intraseasonal) frequencies to low (interannual) frequencies. In particular the impact of a Madden-Julian type oscillation in the 30-60 day frequency range on the interannual ENSO variability is negligble compared to the impact of the interannual component of atmospheric variability.

Contents

List of Figures

List of Tables

Glossary

ANOMALY Difference between a particular quantity and its mean value *for that time of year.*

ATTRACTOR The set of points in state space visited by the trajectory of a dynamical system after transients have died out.

BETA-PLANE An approximation in which the $\sin\lambda$ variation of the Coriolis parameter is replaced with a linear dependence on latitude.

BOX COUNTING DIMENSION If the number of boxes of size ε required to cover an object goes as $n \sim \varepsilon^{-D}$ then the object's box counting dimension is D.

CLOSURE The problem of accounting for the effect of unresolved modes or scales on the evolution of resolved modes or scales.

CORRELATION DIMENSION A lower bound on the box counting dimension.

EMBEDDING DIMENSION The minimum dimension of the space required to embed an object.

EOF Empirical orthogonal function.

ENSO El Niño-Southern Oscillation.

FNN False nearest neighbors, points which appear close in a reconstructed state space due to projection effects.

GCM General circulation model, a model of the ocean or atmosphere that contains the full primitive equations.

HCM Hybrid coupled model, an ocean GCM coupled to a simple atmospheric model.

HOPF BIFURCATION A point in parameter space below which a stable point exists and above which the stable point becomes a limit cycle.

ICM Intermediate coupled model, a model of the ocean-atmosphere system that is simpler than a GCM but still retains ocean dynamics.

LIMIT CYCLE A periodic solution with no other periodic solutions close to it.

MJO Madden-Julian oscillation, an atmospheric oscillation in the tropics with a period of 30 to 60 days.

NINO 3 Region in the eastern tropical Pacific. (See fig.)

NINO 4 Region in the central/western tropical Pacific. (See fig.)

OAGCM Ocean-atmosphere GCM, an ocean GCM coupled to an atmospheric GCM, the most complex type of climate model.

PSEUDO-STATE SPACE A space topologically equivalent to the true state space.

PSEUDO-WINDSTRESS The vector $|\mathbf{V}|\mathbf{V}$ where $\mathbf{V}$ is the surface wind velocity vector.

QB Quasi-biennial, with a period of about 2 years.

QQ Quasi-quadrennial, with a period of about 4 years.

SLAVE ATMOSPHERE An atmospheric model that has no dynamics and is always in equilibrium with atmospheric heating.

SOI Southern oscillation index, the normalized anomaly in the difference in sea-level pressure between Darwin (Australia) and Tahiti.

SST(A) Sea surface temperature (anomaly).

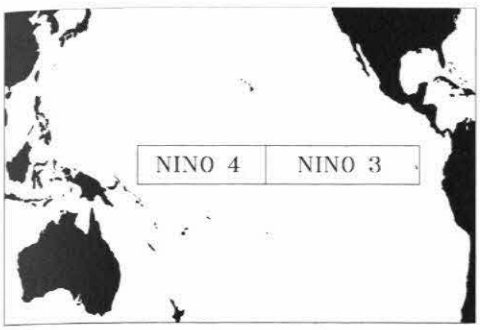
STATE SPACE The space which contains the state vectors of a system.

STATE VECTOR A vector consisting of all the variables that specify the state of a system.

TAO Tropical Atmosphere Ocean experiment. An array of buoys in the tropical Pacific maintained by NOAA.

THERMOCLINE The region in the upper ocean where there is a steep temperature

gradient.



NINO 3 and NINO 4

Chapter 1

Introduction

1.1 Historical Background

The El Niño-Southern Oscillation (ENSO) is the single largest perturbation to the Earth's climate on interannual timescales. The first clue to the existence of ENSO was the observation of an abnormally warm counter current off the coast of Peru once every few years. This was brought to the attention of science by Peruvian geographer Luis Carranza in 1891 [Phi89]. The atmospheric component of ENSO, the Southern Oscillation, was discovered by Sir Gilbert Walker earlier this century as a result of his work on correlations in global weather patterns [Wal23, Wal24, Wal28, WB32, Wal37]. It was not until the extensive observing campaign of International Geophysical Year in 1957, which coincided with an El Niño event, that the connection between the anomalous sea surface temperature (SST) patterns in the tropical Pacific and the the Southern Oscillation was established. In the late 1960s Jacob Bjerknes outlined a theory for ENSO which relied on positive feedback between the tropical Pacific ocean and atmosphere to maintain an El Niño state [Bje66, Bje69]. In brief, the zonal SST gradient across the Pacific is maintained by easterly trade winds which lead to a deep thermocline off Indonesia and a very shallow thermocline off Peru. Over the warm waters off Indonesia moist convection occurs. The

rising air travels east aloft before sinking over the colder waters off Peru. The air then returns west along the surface completing what Bjerknes named the *Walker Circulation* in honor of Sir Gilbert Walker. Hence the zonal SST gradient drives the zonal atmospheric circulation and the surface winds of this circulation maintain the SST gradient. Should the winds weaken then the SST gradient will decrease and this will lead to a further weakening of the winds. An El Niño state may then ensue, characterized by a nearly horizontal thermocline across the Pacific and a very weak, or even reversed Walker Circulation. Although this picture due to Bjerknes is still accepted, it fails to explain why the tropical Pacific should come out of such a state once it enters it. Clearly the system has some sort of "memory."

The devastating El Niño of 1982 stimulated a considerable amount of research activity both in the creation of observing systems and in modeling [Gla96]. Numerical models of the Pacific Ocean were created [CP84] and since a simple, linear model of the tropical atmosphere already existed [Gil80], the first coupled models of the ocean-atmosphere system were constructed [PYP84, AM85, Hir86]. These models were used to study the nature of unstable modes in the coupled system. They were not intended to be used for forecasting El Niño events. The first attempt to produce a dynamical model of the tropical Pacific ocean-atmosphere system capable of forecasting warm events was that of Cane and Zebiak [Zeb84, CZ85, CZD86, ZC87] (hereafter the CZ model). This model exhibited an oscillation with a timescale of about 3 years. It has also been used with some success to forecast warm events. Versions of this model are still used extensively both for making operational forecasts and to try to increase our fundamental understanding of ENSO [BZC95a, BZC95b, CZBC95, PD96, DP96, CZC97]. A similar model was created by Schopf and Suarez [SS88a]. This model also showed ENSO type vacillations. Through studying their model Suarez and Schopf proposed what became known as the *delayed oscillator model* of ENSO [SS88b]. A similar model was proposed by Battisti and Hirst [BH89]. In the delayed oscillator model, a weakening of the trade winds in the central

Pacific creates a westward propagating, upwelling Rossby wave and an eastward propagating, downwelling Kelvin wave. The arrival of the Kelvin wave in the eastern basin causes an increase in SST there which reinforces the weakening of the zonal wind and an El Niño state develops. Meanwhile, the Rossby wave undergoes a phase preserving reflection off the western boundary and returns east as an upwelling Kelvin wave which causes a decrease in SST and begins moving the system towards a La Niña phase. The delayed oscillator model has since been elaborated upon in an attempt to show that ENSO may be a chaotic system [TSCJ94]. Recently satellite measurements of sea surface height have provided observational evidence supporting the delayed oscillator model [BM99]. However, the delayed oscillator model assumes the coupling between the ocean and the atmosphere. More generally ENSO is thought to be a modified scattering mode of the ocean destabilized by coupled processes [JN93a, NJ93, JN93b].

While efforts have been underway to improve our understanding of ENSO physics, the skill of operational forecasts has also been improved, motivated by an increased appreciation of the socio-economic benefits of good ENSO predictions [Gla96]. For example, it has been estimated that even modest skill at forecasting the phase of ENSO could save U.S. agriculture $240 million a year [SAB+98].

1.2 Outstanding Issues

While much progress has been made in the past two decades in understanding ENSO, there are still several important issues that have not been resolved. Two of these issues, which are related to each other, are outlined below:

- Is the ENSO system above the bifurcation at which it becomes unstable and capable of self-sustained interannual oscillations or is it below the bifurcation and essentially a noise excited oscillator?

- Is the irregularity the result of low dimensional processes intrinsic to the tropical Pacific ocean-atmosphere system (chaotic hypothesis) or high dimensional processes some of which are extrinsic to the tropical Pacific (stochastic hypothesis)?

1.2.1 Oscillations: Limit Cycle or Noise?

It has been established that as the coupling between the ocean and atmosphere is increased in models of the tropical Pacific, the system passes through a Hopf bifurcation [Nee89, Nee90]. Below the bifurcation the mean state of the system is a stable fixed point, above the bifurcation the meanstate is unstable and there is a limit cycle. If the system is below the bifurcation, it can still produce interannual variability if it is excited by stochastic forcing. Whether the real tropical Pacific is above or below the bifurcation is difficult to determine. Some have suggested that ENSO can be modeled as a stable, linear system [PM93, CS95, Pen96, Bur99] which implies the system is below the bifurcation. However, most modeling efforts have assumed that the system is above the bifurcation and possesses an intrinsic amplitude determined by the coupled dynamics [ZC87, SS88a, JN93a]. Observational studies also imply that ENSO does not have statistics consistent with a linear process driven by Gaussian noise [BS99].

The question of whether the mean state is stable or whether the system has a limit cycle essentially determines what sets the amplitude of ENSO. If the system is noise excited, then the amplitude of the noise is a major influence on the amplitude. If the system has a limit cycle, then this provides an intrinsic amplitude to ENSO independent of stochastic forcing. Since the amplitude of most of the ENSO variability found in many global climate models is too small [ROB+96, KMG97], the question of what processes determine this amplitude is very significant.

1.2.2 Irregularity: Chaos or Noise?

The observed ENSO cycle is aperiodic. There is a broad peak in its power spectrum that extends over frequencies corresponding to periods of 3 to 7 years. One of the keys to understanding ENSO is knowing the source of this irregularity. One possibility is that ENSO is a low dimensional, *chaotic* system, *i.e.*, it can be described by a small number of deterministic equations but the sensitivity of these equations to initial conditions makes their behavior intrinsically unpredictable. This hypothesis was summed up by Zebiak and Cane after they successfully produced aperiodic behavior in their deterministic model: *"We need not appeal to random forcing of unknown origin in order to account for the aperiodicity of ENSO; it can result from strictly determinisitic processes"* [ZC87]. This hypothesis can be called *intrinsic irregularity.*

An alternative hypothesis is that the coupled ocean and atmosphere form a low dimensional *non-chaotic* system that in the absence of noise would oscillate periodically. "Noise," in the form of synoptic scale atmospheric transients, effectively provides a stochastic forcing to the oscillator causing it to oscillate irregularly. *"The chances of [ENSO] being in an inherently oscillatory, rather than an intrinsically chaotic regime, are substantial."* [JNG96]. This hypothesis can be described as *extrinsic irregularity.* Of course it is possible that the Pacific ocean-atmosphere system would behave chaotically in the absence of atmospheric transients, but in addition is forced by the effectively stochastic synoptic weather patterns. An important source of atmospheric variability in the Pacific is the 30 to 60 day Madden-Julian oscillation responsible for westerly wind bursts that have been implicated in starting El Niño events.

Both the intrinsic and extrinsic irregularity scenarios have been explored. With strong coupling the Tziperman model behaves chaotically, but it also behaves irregularly in the periodic region of parameter space when a stochastic forcing is added [SSHP98]. Another study explored the parameter space of an ICM and found that periodic regimes were

more common than chaotic ones [JNG96]. Studies of a stochastically forced ocean GCM have found that realistic atmospheric transients can produce a marked broadening of the ENSO spectral peak in a model which is periodic when such forcing is absent [Bla97].

1.3 The Model Hierarchy

Models of ENSO can be placed into four categories. Following loosely the classification system used by Neelin *et al.* [NBH+98], these categories are, in decreasing order of complexity, ocean-atmosphere general circulation models (OAGCMs), ocean general circulation models (OGCMs), intermediate coupled models (ICMs) and low dimensional models (LDMs).

1.3.1 Ocean-Atmosphere GCMs

These are the most complex of all ENSO models. They consist of a fully nonlinear ocean circulation model coupled to a fully nonlinear atmospheric circulation model. Their primary use is in studying global climate change over decadal timescales, but as part of these studies they have been used to try to reproduce ENSO type behavior in the tropical Pacific [SHGP87, PLPN89, Mee90, MB92]. The reason for trying to simulate an internally generated ENSO in these global coupled models is twofold. Firstly the presence of ENSO helps to validate the model. Any global climate model that fails to produce an interannual oscillation in the tropical Pacific is failing to produce one of the most important modes of internal climate variability. Secondly the models have been used to address the question of whether long term changes in global climate, particularly those induced by an increase in CO_2, will affect the frequency or amplitude of future El Niño events. OAGCMs are very high dimensional. To estimate the dimension of the state space of such models consider the lowest resolution model used in an ENSO study [PLPN89]. This consisted of an R15 atmospheric model [Lau85] which has 256 variables per field, per vertical level, and the

model had 9 vertical levels. Thus the atmospheric model has a state space with more than 9000 dimensions. This was then coupled to an ocean GCM with a resolution of approximately 4.5° in latitude and 3.8° in longitude and 12 vertical levels, giving over 10,000 dimensions.

1.3.2 Ocean GCMs

On the next rung down the ladder of complexity is the ocean GCM coupled to a simple atmosphere. These models consist of an ocean general circulation model similar to those of OAGCMs, although in some cases only the tropical Pacific is included. Coupled to these are simple atmospheric models. Because of the combination of a complex ocean with a simple atmosphere, these models are sometimes called *hybrid coupled models* (HCMs). Some of the atmospheric models are based on the Gill model [Gil80] while others use statistical atmospheric models created by finding empirical relationships between sea surface temperature and surface windstress. The atmospheric models are referred to as *slave atmospheres* or *steady-state atmospheres* because they contain no dynamics and respond instantaneously to SST forcing. This can be justified because the timescale for atmospheric adjustment is short compared to the typical timescales of the ocean dynamics. Even though they do not contain dynamical atmospheres, OGCMs are still very high dimensional. An example is that developed at UCLA [Nee89]. The ocean GCM component is a 10 level model of the tropical Pacific. The mesh is variable but there are on the order of 1000 grid-points per level, giving a state space of more than 40,000 dimensions. The steady-state atmospheric model does not contribute any degrees of freedom to the dynamical system. Nevertheless the ocean model alone makes such models very high dimensional.

1.3.3 Intermediate Coupled Models

The most famous example of an ICM is the previously mentioned Cane-Zebiak model. The ocean component of this model covers the tropical Pacific (124°E to 80°W and 30°S to 30°N) with 2° latitude and 5.625° longitude grid, giving approximately 400 grid points. The model has four fields on this grid; the two components of horizontal currents, thermocline depth and temperature. Thus even this model has a state space of well over 1000 dimensions. The atmosphere used in the CZ model is a steady-state atmosphere based on the Gill model although some ICMs have been run with empirical, steady-state atmospheric models.

1.3.4 Low Dimensional Models

In an attempt to understand ENSO as a dynamical system, some researchers have created very low dimensional models of ENSO. Typically these models consist of a handful of independent variables. One of the earliest examples was that of Vallis [Val86] who attempted to model ENSO using the well known Lorenz system of equations [Lor63]. This model did not contain any elements of the delayed oscillator model and did not produce the preferred frequency that is observed in ENSO. More recent attempts have been based on the delayed oscillator model [TSCJ94] while others have approached the problem by stripping down ICMs [WF96, Jin97b, Jin97a]. The dimensionality of these models is very low. The Vallis model consisted of three coupled ODEs. The Tziperman model consisted of just one delay-differential equation, the trajectory of which can be embedded in three dimensions. The Wang model reduces to two coupled nonlinear ODEs while the Jin ocean recharge model consists of two coupled linear ODEs.

1.4 Reduced Dynamical Models

As was discussed above intermediate models have had some success at modeling ENSO, even to the point of operational forecasting. The low dimensional models of ENSO which have been studied to date have been created by making *ad hoc* assumptions. It would be desirable to construct a low dimensional model of ENSO dynamics which is more rigorously related to intermediate level models. One method for reducing a spatially extended system to its essential degrees of freedom is the method of *empirical orthogonal functions* (EOFs). [1] This method has been applied to the general problem of turbulence [HLB96] as well as to atmospheric models [RK75, Sel95, Zhu96, Sel97a]. In the method a spectral dynamical model is reprojected onto a new set of orthogonal basis functions (EOFs) which are determined either from runs of the model in its conventional form or from observational data.

1.5 This Thesis

The central theme of this thesis is the construction of a low dimensional ENSO model from a high dimensional ICM by using Empirical Orthogonal Functions as a basis set and without making any *ad hoc* simplifications of the equations of the original ICM. In Chapter 2 the ICM which was ultimately reduced is described. The general form of the model will be recognizable to those familiar with the Cane-Zebiak or Jin-Neelin ICMs. It is to be hoped that by choosing such a well known and much used type of model as the starting point, insights obtained from the reduced model can be applied to understand a large number of model studies. In Chapter 3 the general behavior of the full ICM is described both in its uncoupled and coupled configurations. In Chapter 4 the response

[1] This method goes by a multitude of names. It is also known as Proper Orthogonal Decomposition, Karhunen Loève Decomposition, Principal Component Analysis, Singular Value Decomposition and Singular System Analysis.

of the full model to various forms of stochastic forcing is described. In this chapter it is shown that stochastic forcing at intraseasonal frequencies is not very effective at perturbing the interannual oscillations of the model. In Chapter 5 algorithms for estimating the dimensionality of the ENSO model are presented with their resulting estimates of dimension. Chapter 6 outlines the method by which the model was reduced by projecting it onto truncated basis sets of its EOFs. In Chapter 7 the resulting low dimensional models are analyzed. The processes which destabilize the dominant interannual mode and modify its frequency are elucidated. It is also demonstrated that secondary, linearly stable modes manifest themselves through nonlinear coupling to the dominant mode and that it is the effectiveness of that coupling rather than how close they are to instability that governs their strength. It is also shown that the coupling between modes provides a pathway for wind noise to perturb the leading modes.

Chapter 2

A Description of the Model

2.1 The Oceanic Component

The ocean model used was similar to the oceanic component of previous intermediate models, in particular, the CZ model [Zeb84, ZC87], and the ICM developed by Jin *et al.* [JN93a]. The primary difference between the model described in this chapter and CZ-type models is not in the equations themselves but in the method of their solution. The CZ model is a finite difference model whereas in this work a spectral approach is adopted. [1] Figure 2.1 shows a schematic of the model in its coupled configuration.

The equatorial ocean model consists of two layers. The upper layer is the *mixed layer* and has a thickness H_M. The temperature in the mixed layer is constant with depth and will be written as $T(x, y, t)$. Below the mixed layer is the *thermocline layer* with a mean thickness H_T. The thermocline itself forms the lower boundary of the thermocline layer. The mean depth of the thermocline is $H = H_M + H_T$. The temperature in the thermocline layer is a given function of distance from the thermocline interface. The perturbation in the thickness of the thermocline layer is $h(x, y, t)$. The ocean above the thermocline will

[1] The CZ model in common with many ICMs employs a long-wave approximation. Also an expansion in hermite functions is used in the meridional direction.

be referred to as the *upper ocean*. The depth averaged zonal and meridional currents in the upper ocean will be written as $u(x,y,t)$ and $v(x,y,t)$ respectively. The zonal and meridional currents in the mixed layer will be written as (u_M, v_M) and similarly the currents in the thermocline layer will be written as (u_T, v_T). From the above definitions it follows that

$$Hu = H_M u_M + H_T u_T \tag{2.1}$$

$$Hv = H_M v_M + H_T v_T \tag{2.2}$$

If the shear velocity is defined as $(u_S, v_S) = (u_M, v_M) - (u_T, v_T)$ then Eqs. 2.1 and 2.2 give

$$u_M = u + \frac{H_T}{H} u_S \tag{2.3}$$

$$v_M = v + \frac{H_T}{H} v_S \tag{2.4}$$

The dynamical part of the model consists of the linearized shallow-water equations on an equatorial beta-plane. These equations describe the evolution of $u(x,y,t)$, $v(x,y,t)$ and $h(x,y,t)$.

$$\frac{\partial u}{\partial t} = \beta y v - g' \frac{\partial h}{\partial x} + \frac{\tau_x}{\rho H} - ru \tag{2.5}$$

$$\frac{\partial v}{\partial t} = -\beta y u - g' \frac{\partial h}{\partial y} + \frac{\tau_y}{\rho H} - rv \tag{2.6}$$

$$\frac{\partial h}{\partial t} = -H \left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} \right) - rh \tag{2.7}$$

In the above equations β is the gradient of the Coriolis parameter, g' is the reduced gravity, ρ is the mean density of the upper ocean, r is the coefficient of damping by Rayleigh friction and (τ_x, τ_y) is the windstress vector at the surface. It is assumed that the shear currents are in equilibrium with the windstress and therefore the following steady state equations for the shear current may be written down

$$qu_S - \beta y v_S = \frac{\tau_x}{\rho H_M} \tag{2.8}$$

$$qv_S + \beta y u_S = \frac{\tau_y}{\rho H_M} \tag{2.9}$$

where q is the coefficient of friction between the mixed and thermocline layers. Equations 2.8 and 2.9 can be readily solved to give

$$\rho H_M(\beta^2 y^2 + q^2)u_S = q\tau_x + \beta y \tau_y \tag{2.10}$$

$$\rho H_M(\beta^2 y^2 + q^2)v_S = q\tau_y - \beta y \tau_x \tag{2.11}$$

The upwelling velocity, w, in the mixed layer is given by

$$w = w_T + w_S \tag{2.12}$$

where w_T and w_S are given by

$$w_T = H_M\left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y}\right) \tag{2.13}$$

$$w_S = H_M\frac{H_T}{H}\left(\frac{\partial u_S}{\partial x} + \frac{\partial v_S}{\partial y}\right) \tag{2.14}$$

The equation for the mixed layer temperature is

$$\frac{\partial T}{\partial t} = -u_M\frac{\partial T}{\partial x} - v_M\frac{\partial T}{\partial y} + \gamma\mathcal{H}(w)w\frac{(T_B(h) - T)}{H_M} - \alpha T \tag{2.15}$$

where $\mathcal{H}$ is the heaviside step function, T_B is the temperature at the base of the mixed layer and γ is the fraction of upwelling water entrained in the mixed layer. The first two terms in Eq. 2.15 describe the horizontal advection of temperature. The third term is a parameterization of the effect of upwelling on the temperature of the mixed layer. The heaviside step function means that this term is only non-zero if water is entering the mixed layer from below. This is because since the mixed layer temperature is constant with depth when water leaves the mixed layer (negative w) there is no effect on the temperature of the remaining water. However, when water enters the mixed layer from below (positive w), this water is at temperature T_B and thus when it is mixed throughout

the mixed layer the temperature will change at a rate proportional to $w(T_B - T)$. Note that the Eq. 2.15 contains all the nonlinear terms of the model. The ocean dynamics are linear.

The temperature at the base of the mixed layer, T_B, is parameterized as a function of h. It is assumed that the subsurface temperature structure below the mixed layer is fixed to the thermocline. The following function was used for T_B, in degrees Celsius.

$$T_B = T_0 \tanh\left(\frac{h - h_0}{H^*}\right) - T_0 \tag{2.16}$$

This parameterization function is shown in Fig. 2.2. Note that for convenience the zero in temperature has been moved. The temperature at the center of the thermocline is about 20°C.

2.2 The Atmospheric Component

The Gill model for tropical atmospheric dynamics has been used extensively in intermediate coupled models of the tropical ocean-atmosphere system [Gil80, Gil82]. It is a linear model but it models the response of the tropical atmosphere to SST forcing reasonably well and nonlinear models do not produce substantial improvements [AD93]. The model consists of the linearized shallow-water equations on a beta-plane.

$$-\beta y V + \frac{\partial \phi}{\partial x} + RU = 0 \tag{2.17}$$

$$\beta y U + \frac{\partial \phi}{\partial y} + RV = 0 \tag{2.18}$$

$$c_a^2 \left(\frac{\partial U}{\partial x} + \frac{\partial V}{\partial y}\right) + R\phi = -Q \tag{2.19}$$

In the above equations R is the atmospheric damping coefficient, c_a is the atmospheric Kelvin wave speed, U and V are the zonal and meridional components of lower tropospheric wind respectively and ϕ is the lower tropospheric geopotential height. The variable

Q is the middle tropospheric heating which is taken to be mainly due to latent heat and is thus a function of SST. The atmospheric model is a "slave" atmosphere. It adjusts instantaneously to changes in Q. Using such a steady-state atmospheric model is justified if Q depends only on SST which changes on a timescale associated with the ocean dynamics of weeks to months whereas the atmosphere adjusts on a timescale of days.

2.3 Coupling the Ocean and Atmosphere

The ocean and atmosphere are coupled in both directions. The heating term in the atmospheric model is a function of SST and the windstresses that force the ocean model are functions of the wind velocity obtained from the atmospheric model.

The actual functional dependence of the windstress on the wind velocity was assumed to be linear. The equations are as follows:

$$\tau_x = \rho_a C_D W \mathbf{U} = K_S W \mathbf{U} \tag{2.20}$$

$$\tau_y = \rho_a C_D W \mathbf{V} = K_S W \mathbf{V} \tag{2.21}$$

where ρ_a is the atmospheric density, C_D the coefficient of surface drag and W is a characteristic windspeed. The parameters ρ_a and C_D were combined into a single parameter, K_S.

The atmospheric heating term, Q, was assumed to be a linear function of SST.

$$Q = K_Q T \tag{2.22}$$

where K_Q has units of $m^2\, s^{-3}\, {}^\circ C^{-1}$.

2.4 Estimating the External Windstress

The windstress can be separated into two components, an *internal* windstress, which is the result of SST forcing within the tropical Pacific, and an *external* windstress which is

caused by processes outside the tropical Pacific. The coupled model should be able to generate the internal component itself but it must be forced with the external windstress. The observed windstress is the sum of the internal and external components. The external component can be estimated by using the observed SST and the atmospheric model to estimate the internal component and then subtracting this from the observed windstress. If the observed, external and internal windstresses are represented by τ_{obs}, τ_{ext} and τ_{int} respectively, then

$$\tau_{ext} = \tau_{obs} - \tau_{int} \tag{2.23}$$

If T_{obs} is the observed SST, then τ_{int} can be estimated by

$$\tau_{int} = K_S W \mathcal{A}(K_Q T_{obs}) \tag{2.24}$$

where $\mathcal{A}$ is the atmospheric model.

The windstress forcing used was the Florida State University (FSU) pseudo-windstress product covering the period from 1961 to 1994 [GO81]. This dataset consists of the monthly mean components of magnitude of the wind velocity at the surface on a 2° grid covering the area 124°E to 80°W and 30°S to 30°N. The dataset consists of the vectors

$$\mathbf{S} = \mathbf{V}|\mathbf{V}| \tag{2.25}$$

where $\mathbf{V}$ is the wind velocity. The vector $\mathbf{S}$ is converted into an actual surface stress using a parameterization similar to the one used for the internal windstress.

$$\tau_x = \rho_a C_D S_x \qquad \tau_y = \rho_a C_D S_y \tag{2.26}$$

The observed SST used was actually the SST database from 1950-1991 reconstructed from ship records and satellite observations for the period after 1981 [SRLS96]. The SST reconstructions were used with the atmospheric model described above to estimate τ_{int}. Since the atmospheric model is linear, the product of the two parameters in which there

is most uncertainty, K_Q and W, can be treated as single parameter. The value of $K_Q W$ was chosen so as to minimize the mean square error between the internal zonal windstress anomaly and the observed zonal windstress anomaly. The optimal value was found to be $K_Q = 0.013\, m^3\, s^{-4}\, ^\circ C^{-1}$ (see Fig. 2.3). With this value 58% of the observed windstress variance can be explained as the internal windstress correlated with tropical Pacific SST. Figure 2.4 shows the zonally averaged observed zonal windstress. The annual component of the the external windstress was then calculated by subtracting the estimated internal windstress from the observed windstress. The zonally averaged seasonal cycle in the estimated external zonal windstress is shown in Fig. 2.5.

Subtraction of the seasonal external windstress from the total external windstress leaves the anomalous external windstress. This component of the external windstress is sometimes referred to as *wind noise* since it has a broadband spectrum and is generated by processes outside of the model domain. This component of the windstress can be modeled stochastically. Figure 2.6 shows the amplitude spectrum of the residual external windstress averaged over the NINO 4 region (150°W-90°W, 5°S-5°N). The spectrum is red. This is probably largely due to the ocean acting as a low-pass filter to white atmospheric noise. The white component of the atmospheric noise was estimated by first projecting the anomalous external windstress onto its leading EOFs. The time series of the coefficients of these EOFs had red power spectra. The average power at frequencies above $4\, year^{-1}$ was used as an estimate of the "floor" of the atmospheric white noise. White noise time series with this power were constructed for each EOF coefficient. The resulting wind field was then projected back onto the original basis. The atmospheric noise was constructed in its EOF domain because the coefficients of the EOFs are linearly uncorrelated by construction. In the standard basis there are linear correlations between coefficients of the basis functions. These correlations would be destroyed if the time series of each coefficient were simply replaced by a white noise time series.

2.5 The Spectral Representation

Equations 2.5, 2.6, 2.7 and 2.15 are partial differential equations. They may be solved using finite difference methods. However, since the aim is ultimately to reduce the system to its essential degrees of freedom a *spectral* approach was adopted. In this method each of the fields, u, v, h and T was expressed as a weighted sum of basis functions. The basis functions vary in space the weighting coefficients vary in time. To be as general as possible, the ith basis function for the fields u, v, h and T will be written as F_i^u, F_i^v, F_i^h and F_i^T respectively. The weighting coefficients for the ith basis function of the fields will be written as u_i, v_i, h_i and T_i. Furthermore, let the number of basis functions for each of the fields be written as N_u, N_v, N_h and N_T. Thus u, v, h and T can be written as

$$u(x,y,t) = \sum_{i=1}^{N_u} u_i(t) F_i^u(x,y) \tag{2.27}$$

$$v(x,y,t) = \sum_{i=1}^{N_v} v_i(t) F_i^v(x,y) \tag{2.28}$$

$$h(x,y,t) = \sum_{i=1}^{N_h} h_i(t) F_i^h(x,y) \tag{2.29}$$

$$T(x,y,t) = \sum_{i=1}^{N_T} T_i(t) F_i^T(x,y) \tag{2.30}$$

With an infinite number of basis functions, the basis sets can be complete and any field can be synthesized. The Eqs. 2.5, 2.6 and 2.7 and 2.15 are PDEs and in principle have an infinite number of degrees of freedom. Substitution of the sums in Eqs. 2.27-2.30 when $N_u, N_v, N_h, N_T \to \infty$ into the model equations leads to a countable infinity of ODEs. The aim was to reduce the number of ODEs and thus the number of degrees of freedom to the minimum number required to accurately model the system. This was done by a judicious choice of the basis sets.

The model can be written in the spectral representation as

$$\frac{d}{dt}\begin{bmatrix}\mathbf{u}\\ \mathbf{v}\\ \mathbf{h}\\ \mathbf{T}\end{bmatrix} = \begin{bmatrix} -r\mathcal{I} & \beta\mathcal{Y} & -g'\mathcal{D}_x & \mu\mathcal{A}_x \\ -\beta\mathcal{Y} & -r\mathcal{I} & -g'\mathcal{D}_y & \mu\mathcal{A}_y \\ -H\mathcal{D}_x & -H\mathcal{D}_y & -r\mathcal{I} & 0 \\ 0 & 0 & 0 & -\alpha\mathcal{I}\end{bmatrix}\begin{bmatrix}\mathbf{u}\\ \mathbf{v}\\ \mathbf{h}\\ \mathbf{T}\end{bmatrix} + \mathcal{N}\begin{pmatrix}\mathbf{u}\\ \mathbf{v}\\ \mathbf{h}\\ \mathbf{T}\end{pmatrix} + \mathcal{F}(t) \quad (2.31)$$

where $\mu = K_S K_Q/\rho H$. In Eq. 2.31 $\mathcal{I}$ is the identity operator, $\mathcal{Y}$ represents multiplication by the y coordinate and $\mathcal{D}_x$ and $\mathcal{D}_y$ represent the spatial derivative operators. The operators $\mathcal{A}_x$ and $\mathcal{A}_y$ represent the atmospheric model. They are obtained by inverting the system of equations, Eq. 2.17-2.19. The atmospheric model can be written as

$$\begin{bmatrix} R\mathcal{I} & -\beta\mathcal{Y} & \mathcal{D}_x \\ \beta\mathcal{Y} & R\mathcal{I} & \mathcal{D}_y \\ c_a^2\mathcal{D}_x & c_a^2\mathcal{D}_y & R\mathcal{I}\end{bmatrix}\begin{bmatrix}\mathbf{U}\\ \mathbf{V}\\ \Phi\end{bmatrix} = -\begin{bmatrix}\mathbf{0}\\ \mathbf{0}\\ \mathbf{Q}\end{bmatrix} \quad (2.32)$$

Inversion of the matrix in Eq. 2.32 leads to

$$\begin{bmatrix} . & . & \mathcal{A}_x \\ . & . & \mathcal{A}_y \\ . & . & . \end{bmatrix}\begin{bmatrix}\mathbf{0}\\ \mathbf{0}\\ \mathbf{Q}\end{bmatrix} = \begin{bmatrix}\mathbf{U}\\ \mathbf{V}\\ \Phi\end{bmatrix} \quad (2.33)$$

The operators $\mathcal{A}_x$ and $\mathcal{A}_y$ can thus be extracted from the inverted matrix.

The nonlinear terms in Eq. 2.31 are represented by the function $\mathcal{N}$ and $\mathcal{F}(t)$ represents time dependent forcing.

A characteristic timescale for the model is given by $\sqrt{1/\beta c}$ where c is the Kelvin wave speed given by $c = \sqrt{g'H}$. The model state variables u, v, h and T can be nondimensionalized using the following characteristic scales.

$$\tilde{u} = \tilde{v} = c = \sqrt{g'H} \quad (2.34)$$

$$\tilde{h} = H \quad (2.35)$$

$$\tilde{T} = \frac{c\sqrt{\beta c}}{\mu ||\mathcal{A}_x||} \tag{2.36}$$

where $||\mathcal{A}_x||$ is the norm of the atmospheric model matrix. The windstress forcing can also be nondimensionalized using the characteristic windstress given by

$$\tilde{\tau} = \rho H c\sqrt{\beta c} \tag{2.37}$$

Let $\mathbf{x}$ be the nondimensionalized model state vector given by

$$\mathbf{x} = \left[\frac{\mathbf{u}}{c}, \frac{\mathbf{v}}{c}, \frac{\mathbf{h}}{H}, \frac{\mathbf{T}}{\tilde{T}}\right] \tag{2.38}$$

The model equations can now be written in the following nondimensional form

$$\frac{d\mathbf{x}}{dt} = \mathbf{M}\mathbf{x} + \mathbf{N}(\mathbf{x}) + \mathbf{F}(t) \tag{2.39}$$

where $\mathbf{M}$ is the nondimensionalized model matrix, $\mathbf{N}$ is the nondimensionalized nonlinear part of the model and $\mathbf{F}(t)$ is the nondimensionalized time dependent forcing.

2.6 The Standard Basis Sets

For the initial runs of the model, a basis set of conventional functions was used. The standard basis sets used for the fields u, v, h were products of fourier functions in the zonal direction and hermite functions in the meridional direction. These functions are solutions to the free dynamical equations, Eqs. 2.5, 2.6 and 2.7. The boundary condition is that $u = 0$ at $x = 0$ and at $x = L$ where L is the zonal extent of the ocean basin. The hermite functions decay to zero as $|y| \to \infty$. For the SST field fourier functions were used in both the zonal and the meridional directions. If the basis set contains the m lowest fourier functions and n lowest hermite functions, then the functions in Eqs. 2.27, 2.28,

2.29 and 2.30 are

$$
\begin{aligned}
F^u_{lm+k} &= \sin\left(\frac{k\pi x}{L}\right) D_l\left(\frac{y}{\sqrt{c/\beta}}\right) \quad (k=1,...,m \;\; l=0,...,n-1) \\
F^v_{lm+k} &= \cos\left(\frac{k\pi x}{L}\right) D_l\left(\frac{y}{\sqrt{c/\beta}}\right) \quad (k=0,...,m \;\; l=0,...,n-1) \\
F^h_{lm+k} &= \cos\left(\frac{k\pi x}{L}\right) D_l\left(\frac{y}{\sqrt{c/\beta}}\right) \quad (k=0,...,m \;\; l=1,...n-1) \\
F^T_{lm+k} &= \cos\left(\frac{k\pi x}{L}\right) \cos\left(\frac{l\pi y}{W}\right) \quad (k=0,...m-1 \;\; l=0,...,n-1)
\end{aligned}
$$

where D_l is the hermite function of order l. The number of functions in each basis set is

$$N_u = N_T = mn, \quad N_v = N_h = (m+1)n \tag{2.40}$$

The dimension of the model state space is thus $N_u + N_v + N_h + N_T = 4mn + 2n$.

2.7 Numerical Methods

The model described by Eq. 2.39 was integrated using a fourth order Runge Kutta scheme. The linear operators, $\mathcal{L}$, $\mathcal{D}_x$, $\mathcal{D}_y$, and $\mathcal{Y}$ were constructed numerically on a grid 1° in longitude and latitude covering the area from 124°E to 80°W and 15°S to 15°N. The nonlinear terms were evaluated pseudo-spectrally on a coarser grid with a resolution of 6° in longitude and 3° in latitude.

2.8 Parameter Values

The parameters β, ρ, ρ_a and L have unambiguous physical interpretations and the values used were $2 \times 10^{-11}\, m^{-1}s^{-1}$, $1000\, kg\, m^{-3}$, $1.3\, kg\, m^{-3}$ and $1.74 \times 10^7 m$ respectively. The value of L corresponds to 156° of longitude. The value of H was taken to be $150\, m$ as in the CZ model. Other parameters used were $r = (2.5\, years)^{-1}$, $H_M = 50\, m$, $\gamma = 0.5$

and $\alpha = (125\,days)^{-1}$. The reduced gravity was set at $g' = 0.0486\,ms^{-2}$ to give a Kelvin wave speed of $2.7\,ms^{-1}$. The value of α corresponds to a coupling sensitivity of $19\,Wm^{-2}K^{-1}$. The windstress coefficient, C_D, was set at 1.5×10^{-3} [GO81]. The parameters are summarized in Table 2.1.

Table 2.1: The parameters of the standard model.

m	zonal functions	8
n	meridional functions	7
β	planetary vorticity gradient	$2 \times 10^{-11}\, m^{-1} s^{-1}$
ρ	mean density of upper ocean	$1000\, kg\, m^{-3}$
ρ_a	surface density of atmosphere	$1.2\, kg\, m^{-3}$
L	zonal extent of basin	$1.74 \times 10^{7}\, m$
g'	reduced gravity	$0.0486\, ms^{-2}$
H	mean thermocline depth	$150\, m$
c	Kelvin wavespeed ($\sqrt{g'H}$)	$2.7\, ms^{-1}$
L_0	oceanic Rossby radius ($\sqrt{c/\beta}$)	$3.7 \times 10^{5}\, m$
r	upper ocean Rayleigh friction	$(2.5\, years)^{-1}$
C_D	windstress coefficient	1.5×10^{-3}
H_M	thickness of mixed layer	$50\, m$
q	mixed layer Rayleigh friction	$(2\, days)^{-1}$
α	oceanic Newtonian cooling coefficient	$(125\, days)^{-1}$
γ	entrainment fraction	0.50
T_0	temperature contrast across thermocline/2	$10°C$
h_0	temperature structure asymmetry parameter	$40\, m$
H^*	vertical scale of thermoclinemetry parameter	$50\, m$
R	atmospheric Rayleigh friction	$(2.3\, days)^{-1}$
c_a	atmospheric Kelvin wavespeed	$60\, ms^{-1}$

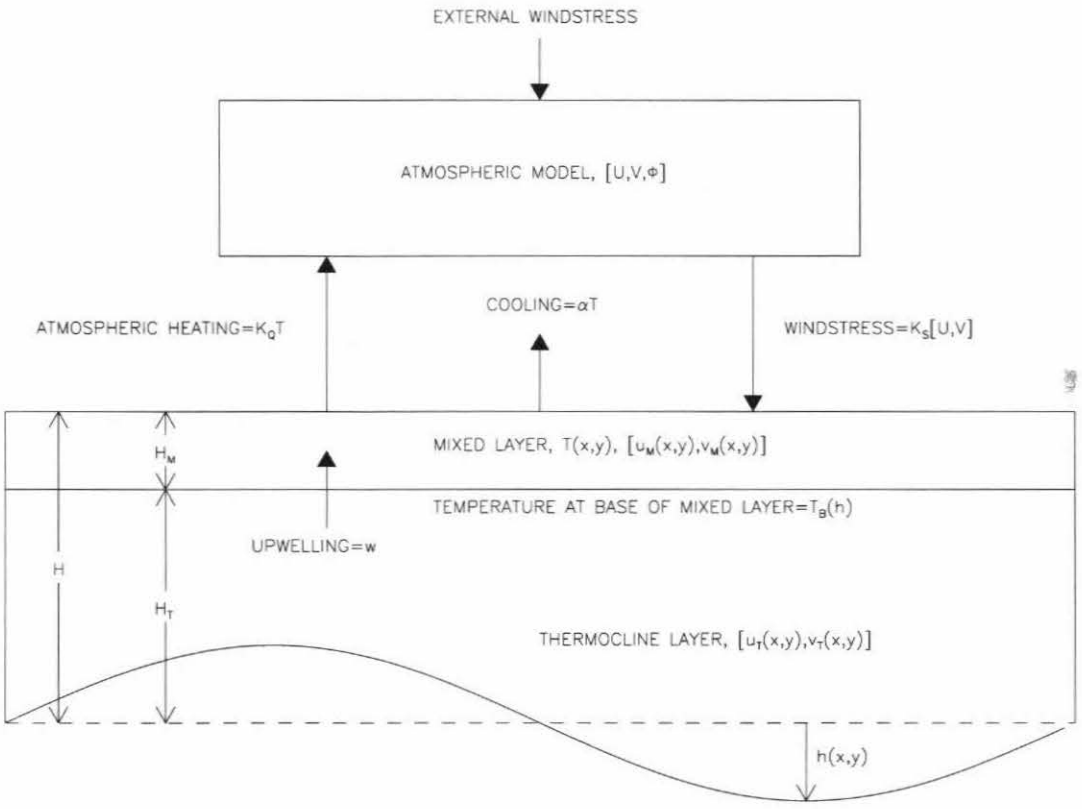


Figure 2.1: A schematic of the model in its coupled configuration.

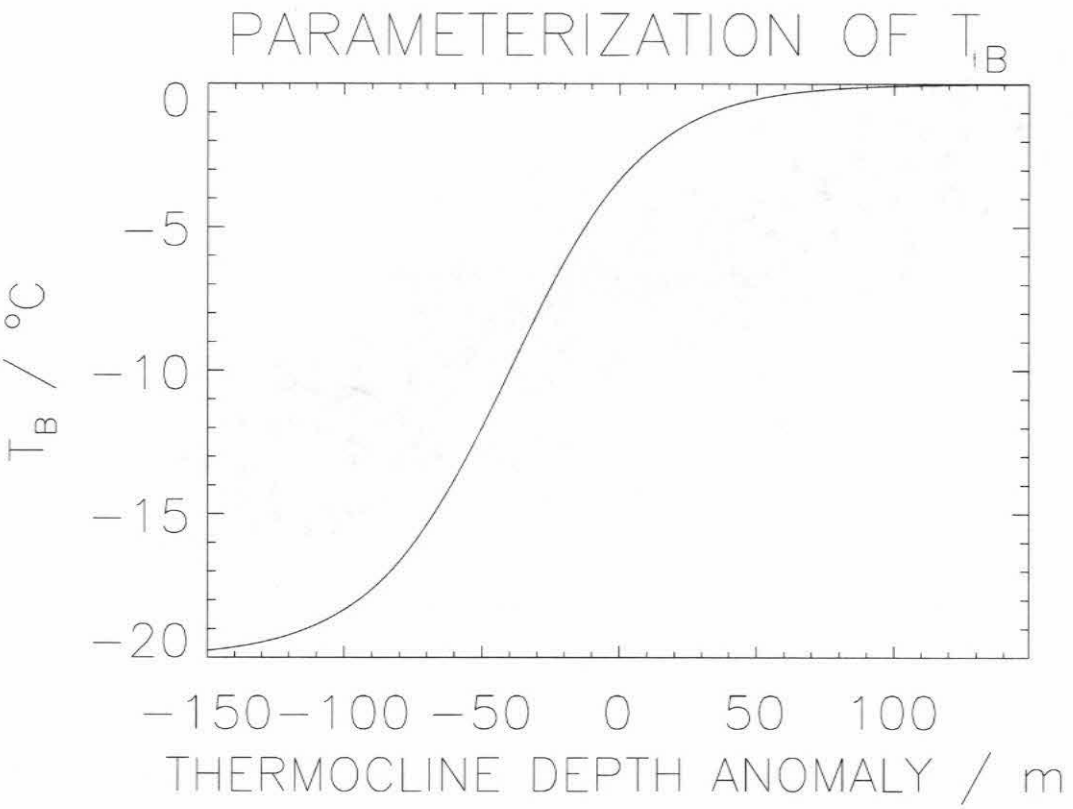


Figure 2.2: The temperature at the base of the thermocline, T_B as a function of thermocline depth anomaly, h. The relationship is that of Eq. 2.16 for the parameter values $T_0 = 10°C$, $h_0 = 40\,m$ and $H^* = 50\,m$.

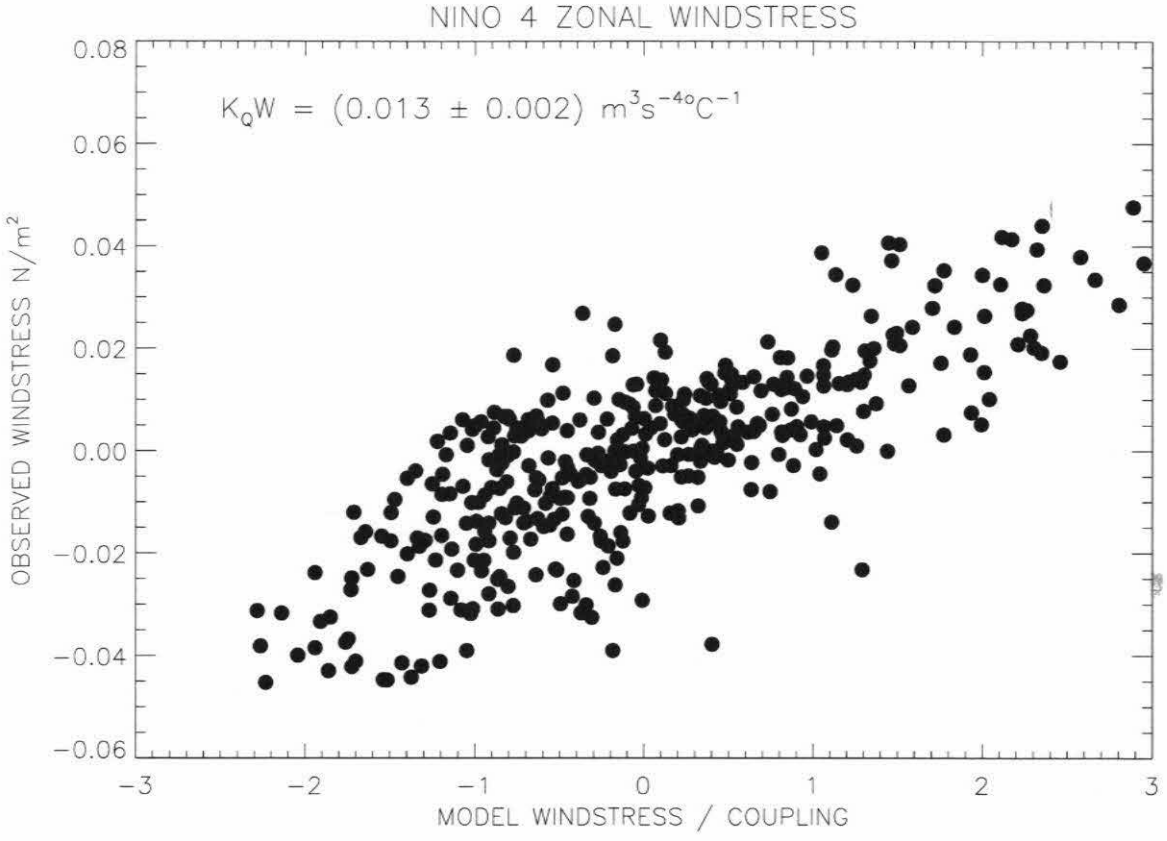


Figure 2.3: The observed NINO 4 zonal windstress anomaly plotted against the windstress produced by driving the atmospheric model with observed SST.

The implied value of K_QW is $(0.013 \pm 0.002)\, m^3\, s^{-4}\, {}^\circ C^{-1}$. To estimate the uncertainty the number of degrees of freedom must be known. Because both time series have high serial correlations, the number of degrees of freedom is less than the number of points. The number of degrees of freedom was estimated by calculating the autocorrelation functions of the series. The first zero was at 8 months for both series. The estimated number of degrees of freedom is thus 384 months/8 months = 48.

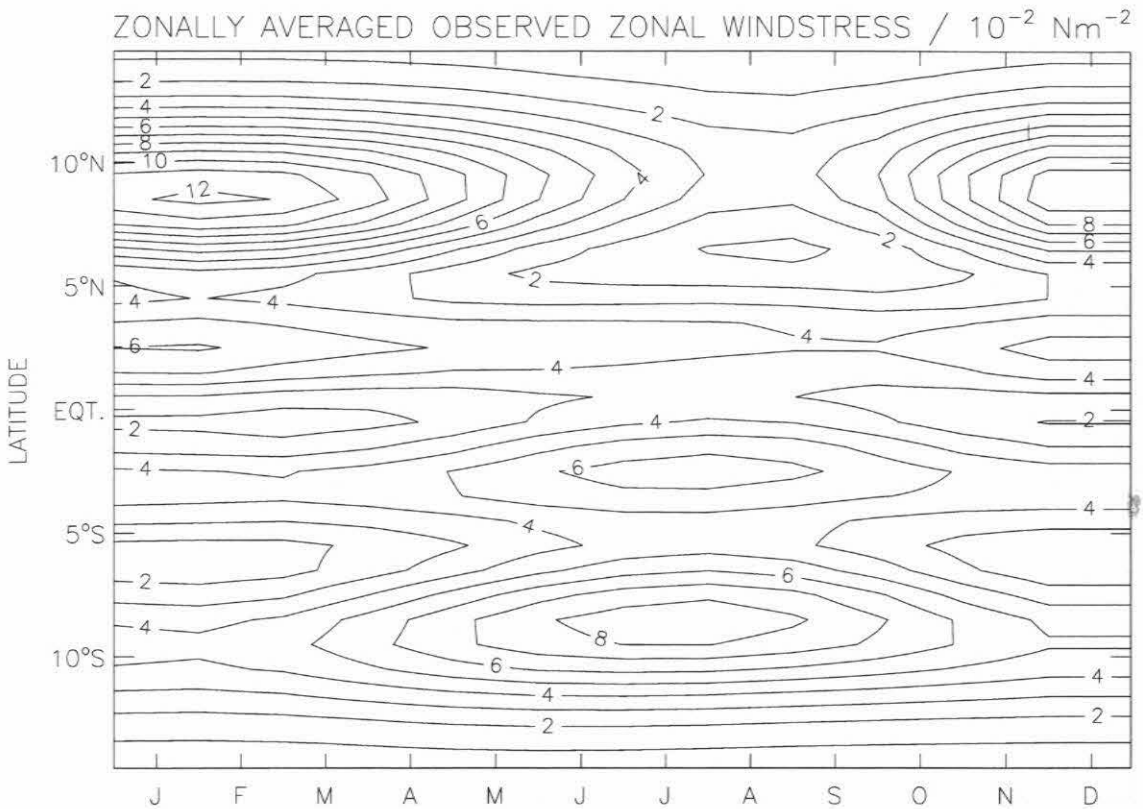


Figure 2.4: Observed windstress.

Zonally averaged seasonal cycle in the observed (FSU) zonal windstress. Units are $10^{-2}\,N\,m^{-2}$ and positive indicates an easterly windstress. A value of $K_S = 0.0019\,kg\,m^{-3}$ was assumed. This is the windstress dataset that was used to estimate the *external* component of the windstress.

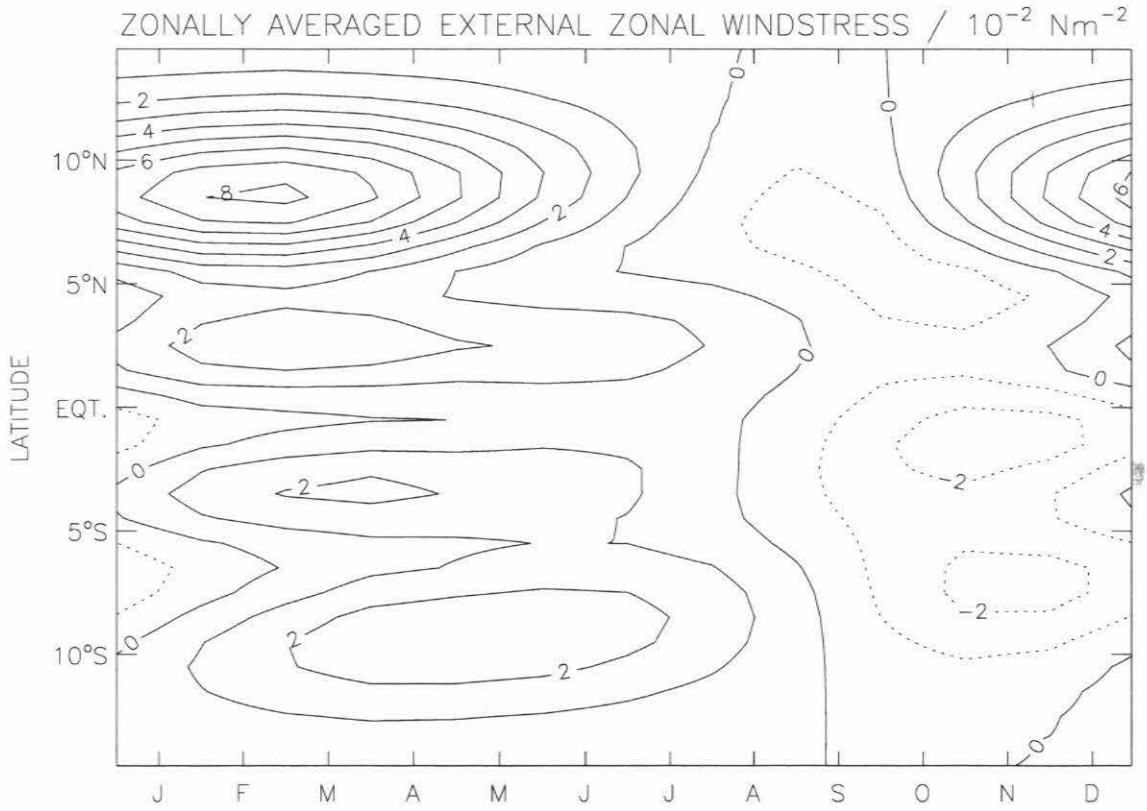


Figure 2.5: Zonally averaged seasonal cycle in the estimated external zonal windstress. Units are $10^{-2}\,N\,m^{-2}$ and positive indicates an easterly windstress. This is the seasonal component of the observed windstress (Fig. 2.4) that *cannot* be explained as a linear response to SST in the tropical Pacific.

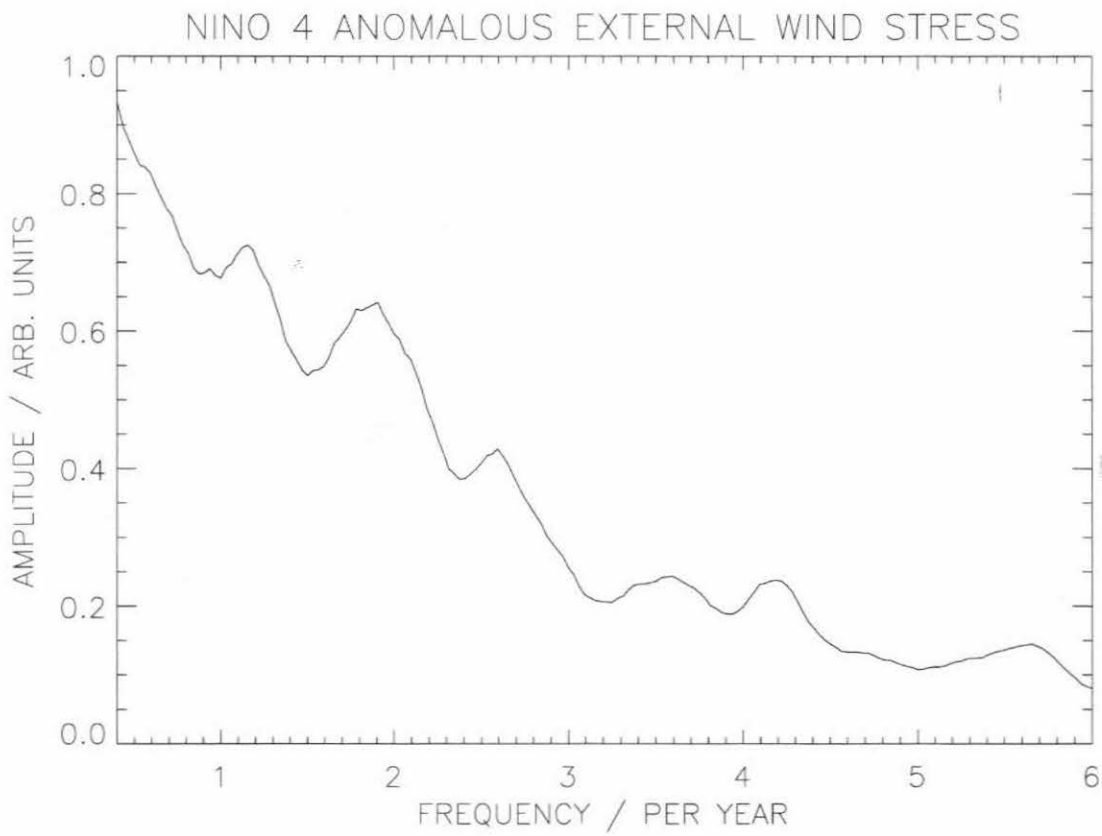


Figure 2.6: Amplitude spectrum of the residual component of the estimated external zonal windstress averaged over the NINO 4 region. This is the nonseasonal component of the observed windstress that *cannot* be explained as a linear response to SST in the tropical Pacific.

Chapter 3

Behavior of the Model

3.1 Uncoupled Model

Before any attempt was made to run the model in its coupled mode, it was run uncoupled, driven by the FSU pseudo-windstress for the period 1961 to 1994. The parameter values used were those listed in Table 2.1.

Figure 3.1ab shows the time series of SST anomaly in the NINO 3 and NINO 4 regions. The NINO 3 region is in the Eastern Pacfic, covering the region 150°W-90°W and 5°S-5°N. NINO 4 lies in the mid to western Pacific and covers 150°W-160°E and 5°S-5°N. The red line shows the model SST anomalies. The blue line shows observed SST anomalies which were obtained from the SST reconstructions [SRLS96].

Figure 3.1cd shows the model thermocline depth anomalies on the equator at 165°E and 120°W (red lines). The blue lines which begin in 1987 are objective analyses of the depth anomalies of the 20°C isotherm at 165°E and 120°W obtained using the TAO buoy array [KM95b, KM95a].

From Fig. 3.1 it can be seen that the forced model reproduces the equatorial thermocline displacements well over the period for which observations are available. Furthermore, the relative amplitudes of the SST anomalies in the NINO 4 and NINO 3 regions are re-

produced.

Figure 3.2 shows the mean and standard deviations in the equatorial profiles of SST, thermocline depth and upwelling produced by the forced model. The SST gradient along the equator is consistent with observations. The largest changes in SST are found in the Eastern Pacific, as is the case in reality. This is not because there is more thermocline depth variability there. The RMS variation in thermocline depth is about the same in the western and eastern parts of the basin. The reason for the greater SST variability in the east is partly due to the much larger upwelling there and also because the gradient in the parameterization of T_B is greater at $h = -50\,m$ than at $h = +10\,m$.

Figure 3.3 shows longitude-time plots of the observed and modeled SST anomaly along the equator and the amplitude spectra of NINO 3 SST. Note that the modeled SST field displays more high frequency variability than the observed field.

Figure 3.4 compares the seasonal cycle in SST of the forced model with the observed seasonal cycle. The agreement in the NINO 3 region is good although the coldest SST occurs a couple of months later than in the observations. In the NINO 4 region the model does a poor job of reproducing the observed SST. This disparity can be explained by recognizing that in the NINO 3 region upwelling is important in determining the SST whereas in NINO 4 surface heat flux is more important [SZC88]. The model includes upwelling and thus does reasonably well in NINO 3 but does not include seasonal changes in surface heat flux leading to poor agreement further west. Since the SST variability in the western Pacific is relatively small, failure to reproduce the annual cycle in these regions does not represent a major deficiency in the model.

The forced model run output was used to make another estimate of the coupling coefficient. Figure 3.5 shows a plot of the observed zonal windstress anomaly in the NINO 4 region versus the windstress produced by the model divided by the coupling coefficient, $K_Q W$. From this plot a value of $K_Q W = 0.012 \pm 0.002\,m^3\,s^{-4}\,{}^\circ C^{-1}$ was inferred. This is consistent with the value of $K_Q W$ inferred using the atmospheric model driven by the

observed SST.

To test the atmospheric model, the SST produced by the forced model was used to drive the atmospheric model. Figure 3.6 shows the resulting windstresses in the NINO 3 and NINO 4 regions and also the observed windstresses. [1] The model reproduces the observed windstresses reasonably well indicating that there should be self-consistency in the coupled model.

3.2 Coupled Model

3.2.1 Constant Windstress

Before the coupled model was forced with a seasonally varying windstress it was forced with a constant windstress. The spatial pattern of this windstress was the estimated external windstress for the month of April. April was chosen because the estimated external windstress is most strongly easterly during this month (see Fig. 2.5). A coupling of $K_Q W = 0.0097\, m^3\, s^{-4}\, ^\circ C^{-1}$ was chosen as this coupling placed the system just above the bifurcation at which sustained interannual oscillations appear. Figure 3.7 shows a plot of equatorial SST as a function of time. The NINO 3 SST oscillates with a period of about 4.5 years and a peak-to-peak amplitude of about 2°C. The longitude-time plot in Fig. 3.7b shows the spatial pattern of oscillation typical of ENSO with the SST variance confined to the Eastern Pacific although the NINO 3 region and the region closest to the eastern boundary are out of phase. This phase difference is not present in observations. The amplitude spectrum of NINO 3 SST shows a dominant peak corresponding to a period of about 4.5 years and a smaller peak corresponding to about 2.1 years. It is interesting to note that analysis of observed equatorial Pacific SST has found a quasi-quadrennial

[1] The "observed" windstresses are obtained from the FSU pseudo-windstresses by assuming that $K_S = 0.0019\, kg\, m^{-3}$.

(QQ) mode with a period of 53 months and a quasi-biennial (QB) mode with a period of 28 months [JNG95]. Analysis of the southern oscillation index also found a low frequency component with a period of 4-6 years and a high frequency component with a period of 2-3 years [KDG92]. In Chapter 6 it will be shown that these two modes of oscillation in the model can be separated by EOF analysis.

3.2.2 Seasonal Windstress

The model was run in its coupled configuration for a range of values of $K_Q W$. Figure 3.8 shows the NINO 3 SST anomaly for these runs. The model shows interannual variability for values of $K_Q W$ of $0.0136\, m^3\, s^{-4}\, ^\circ C^{-1}$ and higher. As $K_Q W$ is increased the behavior of the model becomes more irregular. There appears to be a bias towards strong La Niña events over El Niño events which is not seen in the observed time series of NINO 3 SST. This may be due to problems with the parameterization of subsurface temperature. The original CZ model uses a more detailed parameterization with spatially varying parameters.

The estimates of the coupling parameter calculated using Figs. 2.3 and 3.5 are both very close to the bifurcation above which the seasonally forced model is capable of self-sustained interannual oscillations. This highlights the difficulty in determining whether the system is strictly noise excited or whether it possesses a limit cycle and hence an intrinsic amplitude.

Figure 3.9 shows the behavior of the model with a coupling strength of $K_Q W = 0.0136\, m^3\, s^{-4}\, ^\circ C^{-1}$. The behavior is clearly more complicated than the case with no seasonal cycle, but there is still a sharp interannual peak. The period of the interannual oscillation is now 5.3 years. The peak-to-peak amplitude of the interannual oscillation is approximately 1°C. Figure 3.9b shows that the general spatial pattern of ENSO remains although there is more SST variability in the west than in the non-seasonal case. Also

visible in the amplitude spectrum is the annual cycle as well as several smaller peaks.

Figure 3.10 shows the behavior of the model with a coupling of $K_Q W = 0.0140\, m^3\, s^{-4}\, ^\circ C^{-1}$. The peak representing the annual cycle remains and has a similar amplitude to the case with lower coupling, but in the more strongly coupled run the interannual behavior becomes far more complicated with a marked broadening of the interannual peak at around a period of 5 years. The amplitude of the interannual oscillations becomes highly variable with peak-to-peak amplitudes ranging from less than 3°C to more than 8°C. However, the spatial pattern of the SST changes remains ENSO-like with warming and cooling events confined to the east of the basin.

Both the coupled runs shown in Figs. 3.9 and 3.10 capture many of the key aspects of ENSO. The primary interannual peak is at a lower frequency than the peak in the spectrum of observed NINO 3 SST which is closer to 3 years (Fig. 3.4), but the spatial pattern of SST variation is very similar to the pattern observed. In both coupled runs the amplitude of the interannual variability is higher relative to the annual cycle than in the observations.

Hereafter the case of $K_Q W = 0.0136\, m^3\, s^{-4}\, ^\circ C^{-1}$ will be called the *regular regime* and the case $K_Q W = 0.0140\, m^3\, s^{-4}\, ^\circ C^{-1}$ will be referred to as the *irregular regime*.

Figure 3.11 shows the equatorial profiles of mean SST, thermocline depth anomaly and upwelling along with the RMS variation of these quantities. It should be compared with Fig. 3.2 for the forced model case. The agreement is reasonably good. The temperature contrast across the basin is a couple of degrees higher in the forced case than the coupled case. This is due to a slightly deeper eastern thermocline in the coupled model. There is also more variation in thermocline depth in the middle of the basin in the forced case than in the coupled case. The main features of the equatorial profiles are reproduced.

The seasonal cycles in SST of the coupled models are shown in Figs. 3.12 and 3.13. As in the forced case the model reproduces the SST cycle considerably better in the NINO 3 region than in the NINO 4 region. In the NINO 3 region the amplitude is in agreement

with observations, but the warmest SST occurs three months prematurely in the model. In the NINO 4 region the amplitude agreement is good, but the model's seasonal cycle is 6 months out of phase.

3.3 Stochastically Forced Model

The model in the regular regime was forced with the stochastic wind forcing with a white spectrum described in Chapter 2. The result is shown in Fig. 3.14. It should be compared with Fig. 3.9. The ENSO peak has been substantially broadened as have the peaks at 0.8 and 1.2 $year^{-1}$. The pattern of SST change remains ENSO-like, with most variability in the east, but from Fig. 3.14a and 3.14b it can been seen to be more irregular than in the case without wind noise.

If the model lies below the bifurcation, it can still exhibit interannual variability if excited by noise. Figure 3.15 shows the behavior of the model with a coupling of $K_Q W = 0.0134\, m^3\, s^{-4}\, ^\circ C^{-1}$ which places it just below the bifurcation when forced with the seasonal external windstress and the stochastic forcing. Notice the similarities between Fig. 3.15 and Fig. 3.14. The interannual peak is somewhat sharper in the model above the bifurcation, but in general both the spatial and temporal behavior of the models is similar. This indicates the difficulty in determining whether the noise driven system would exhibit interannual variability in the absence of noise.

A more detailed analysis of the model's response to stochastic forcing can be found in the following chapter.

3.4 Phase Synchronization

An important feature of the ENSO cycle is its phase synchronization with the seasonal cycle. El Niño events appear to have a preference for peaking during Northern winter

(hence the name). Figure 3.16a shows the record of the Southern Oscillation Index (SOI) with the peaks of 30 El Niño events (characterized by negative SOI) marked with squares. Figure 3.16b and 3.16c show the annual distribution of the event peaks binned by month and season respectively. Note the tendency for event peaks to occur between September and March. The significance of this tendency was determined by splitting the year into the periods September-February and March-August containing 20 and 10 warm events respectively. The entropy of this distribution is 0.918 bits or 0.636 *nats*. [2] A test statistic for entropy can be defined as [Rou97]

$$Z_H = \sqrt{\frac{B}{2}} \left(\frac{2N(\ln B - H)}{B-1} - 1 \right) \tag{3.1}$$

where N is the number of events and B is the number of bins. For $H = 0.636\,nats$ the value of the statistic is $Z_H = 2.4$ implying a 99% significance to the deviation of this distribution from uniformity.

For comparison, Figs. 3.17, 3.18, 3.19 and 3.20 show the annual distribution of warm events in the coupled models in the regular and irregular regimes without wind noise, in the regular regime when wind noise is added and below the bifurcation when wind noise is added. All the distributions were constructed by identifying peaks in the NINO 3 SST anomaly in 200 year runs of the model. In all the cases the model has a strong tendency towards warm events in the JJA season, in contrast to the observations which suggest a preference for the SON season. This discrepancy might be explained by the approximate three month phase difference between the model's seasonal cycle in NINO 3 SST and the observed cycle (Figs. 3.12 and 3.13). Both the model results and observations suggest a preference for peaks in the NINO 3 SST anomaly during the part of the year when the seasonal SST in the region is below average.

[2] A *nat* is the unit of entropy when natural logarithms are used instead of base-2 logarithms. It is equal to $1/\ln 2$ bits.

3.5 Comparison of Spectra

Figure 3.21 and 3.22 show comparisons of the amplitude spectra of the observed NINO 3 SST between 1950 and 1999 and amplitude spectra of 49 year segments of runs of the coupled models. In Fig. 3.21 the models were not forced with stochastic wind forcing whereas in Fig. 3.22 wind noise with a white spectrum was added. From Fig. 3.21 it can be seen that ampitude of the interannual variability in both models is close to that observed and the amplitude of the seasonal variability in the seasonal model also matches observations well. However, the frequency of the interannual peak is too low in the seasonal model and neither model has the broad peak at interannual frequencies seen in the observations. Figure 3.22 indicates that the addition of stochastic wind forcing broadens the interannual peaks. The peak broadening seen in the nonseasonal model does resemble the observations but the dominant interannual frequency of the seasonal model is still too low. When looking at these comparisons it should be remembered that these spectra have been created using time series just 49 years long. To get an idea of the natural variability in the ENSO spectrum, the Southern Oscillation Index (SOI) was used since this time series extends back into the last century. Figure 3.23 shows the amplitude spectra of three 49 year segments of the SOI. Note that there is variability in the frequency of the interannual peak. During the period 1882 to 1930 the dominant period is close to 3.3 years whereas between 1915 and 1963 the peak is closer to 5 years.

3.6 Normal Modes of the Ocean Model

Figure 3.24 shows the eigenvalues associated with the normal modes of the uncoupled ocean model dynamics (which are linear). Figure 3.24a shows eigenvalues in the interannual frequency range. It has been nondimensionalized with the basin crossing time of a Kelvin wave for comparison to previous results [NJ93]. These interannual modes are

discretized versions of the continuous scattering modes found by Moore [Moo68]. Figure 3.24b shows the eigenvalues of all the normal modes of the uncoupled ocean model. Notice the very high frequency modes that arise because the long wave approximation was not made. The interannual modes have decay times between 2 and 2.5 years. This decay time is set by the coefficient of Rayleigh friction, r. It has been established that it is scattering modes with interannual frequencies that give rise to ENSO through destabilization and modification through atmospheric coupling [NJ93]. This is the general origin of ENSO. The question of which modes are selected and how they are modified as well as how nonlinear processes may couple the modified modes will be addressed in Chapter 7.

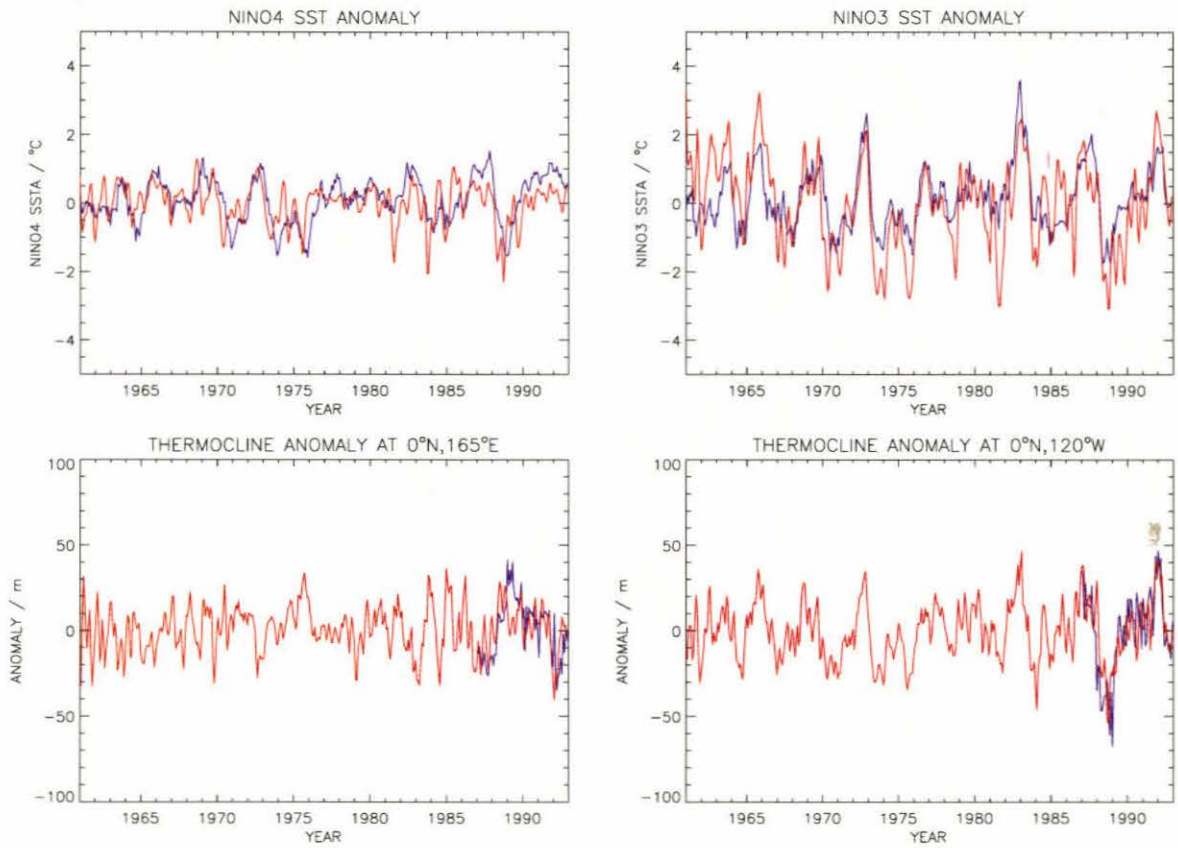


Figure 3.1: The results of the 34 year forced run in which the ocean model was forced with the observed FSU windstress. The red line is the model and the blue line is observations in each figure. The observed thermocline is taken to be the 20°C isotherm. (a) NINO 4 SST anomaly; (b) NINO 3 SST anomaly; (c) thermocline depth anomaly at (0°N,165°E); (d) thermocline depth anomaly at (0°N,120°W). Note that the thermocline depth observations are from the TAO array and begin in 1987.

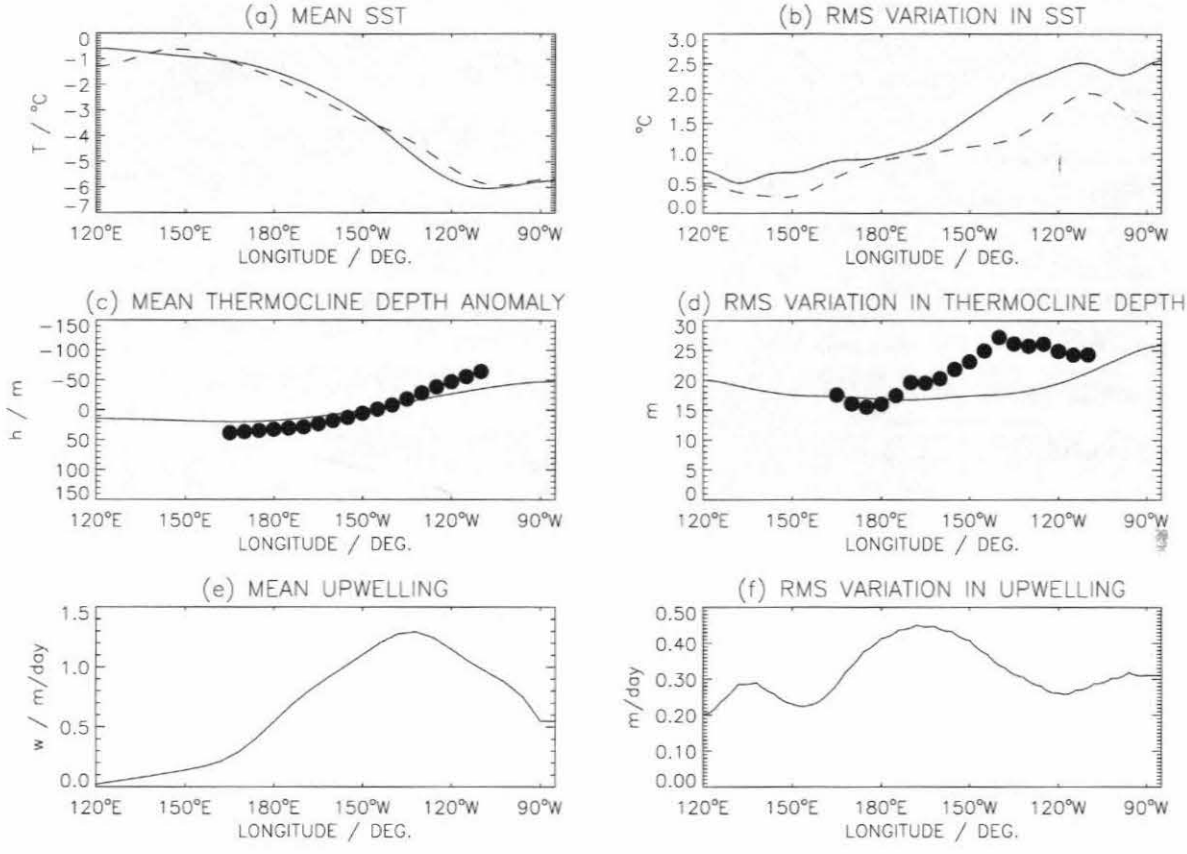


Figure 3.2: Equatorial profiles of variable mean and standard deviations produced by the forced model. (a) mean SST along the equator; (b) standard deviation in equatorial SST; (c) mean equatorial profile in thermocline depth anomaly; (d) standard deviation in thermocline depth; (e) mean equatorial profile of upwelling; (f) standard deviation in the equatorial upwelling. The dashed curves in (a) and (b) are for the observed SST. The circles in (c) and (d) are for the 20°C isotherm measured by the TAO array.

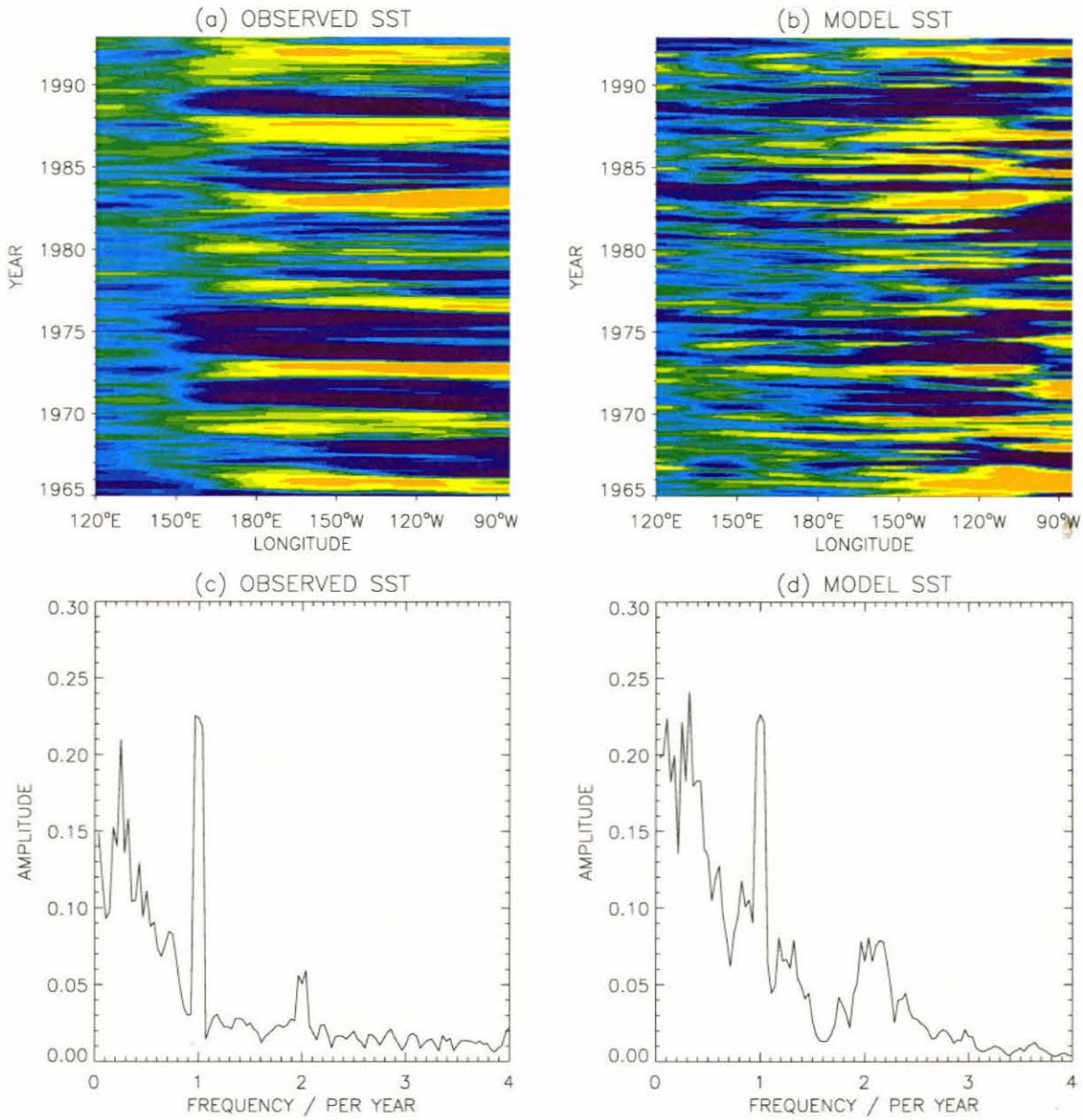


Figure 3.3: Observed SST and the result from the model forced with observed windstress.

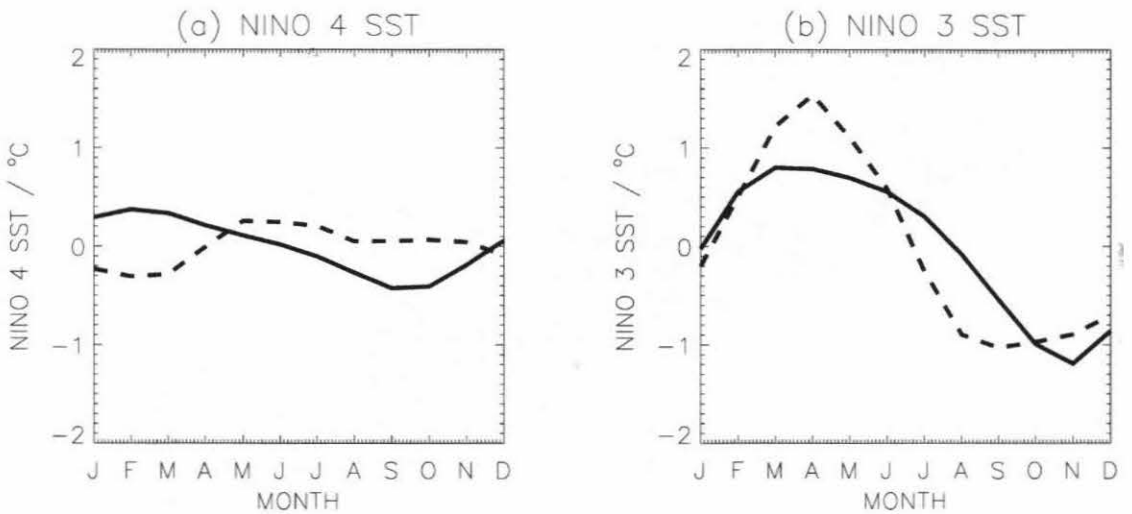


Figure 3.4: The observed SST seasonal cycle (*dashed line*) and the seasonal cycle produced by the forced model run (*solid line*).

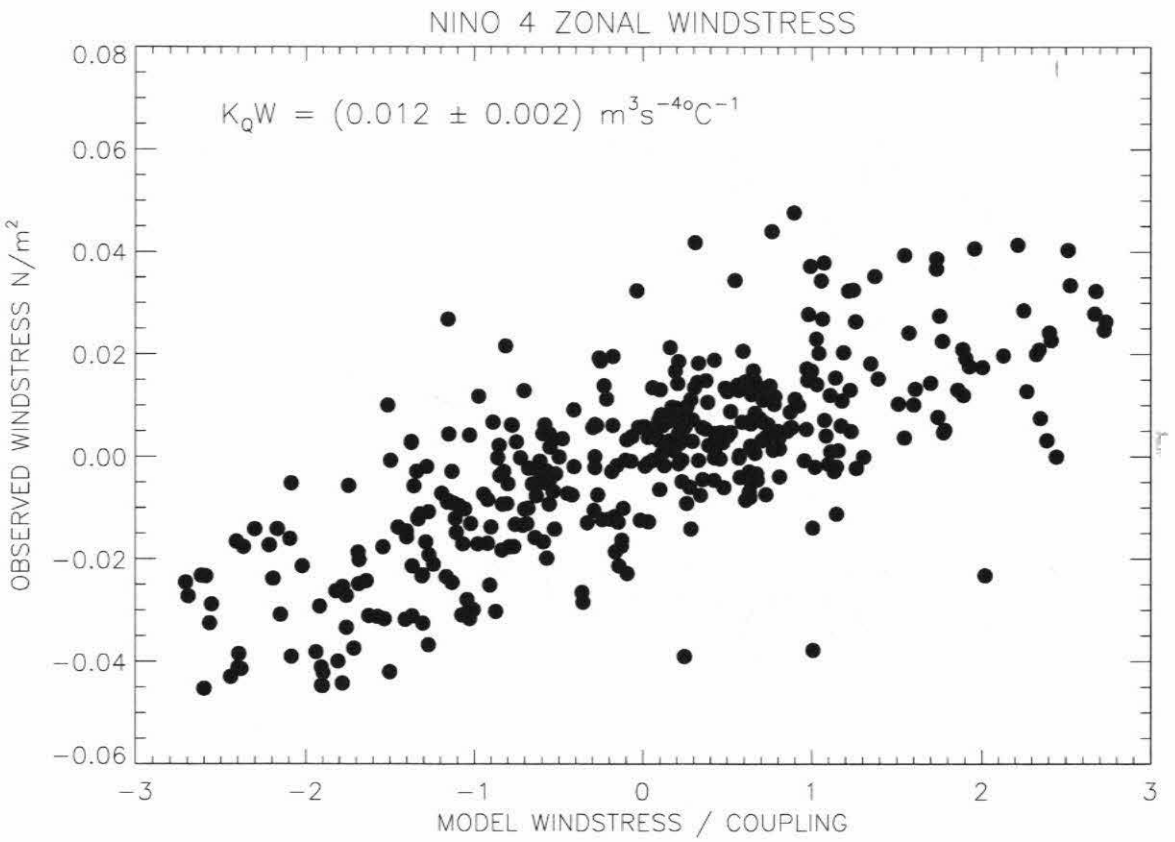


Figure 3.5: The observed NINO 4 zonal windstress anomaly plotted against the windstress produced by the atmospheric model driven by the SST produced by the uncoupled model. The implied value of K_QW is $(0.012 \pm 0.002)\, m^3\, s^{-4}\, {}^\circ C^{-1}$.

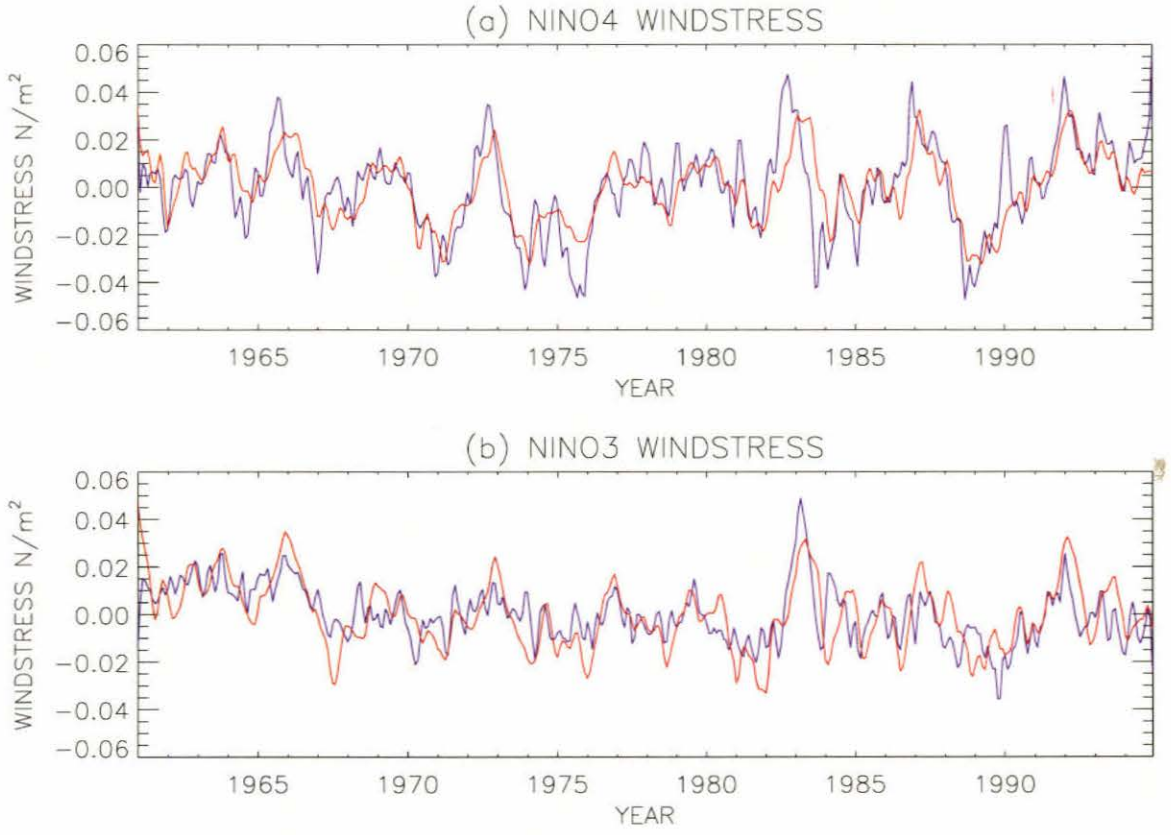


Figure 3.6: A comparison of the observed and modelled zonal windstresses in the NINO 4 and NINO 3 regions. The blue lines are the anomalies calculated from the FSU dataset while the red lines are the anomalous windstresses generated by the model in its uncoupled configuration using a value of $K_Q W = 0.012\, m^3\, s^{-4}\, ^\circ C^{-1}$.

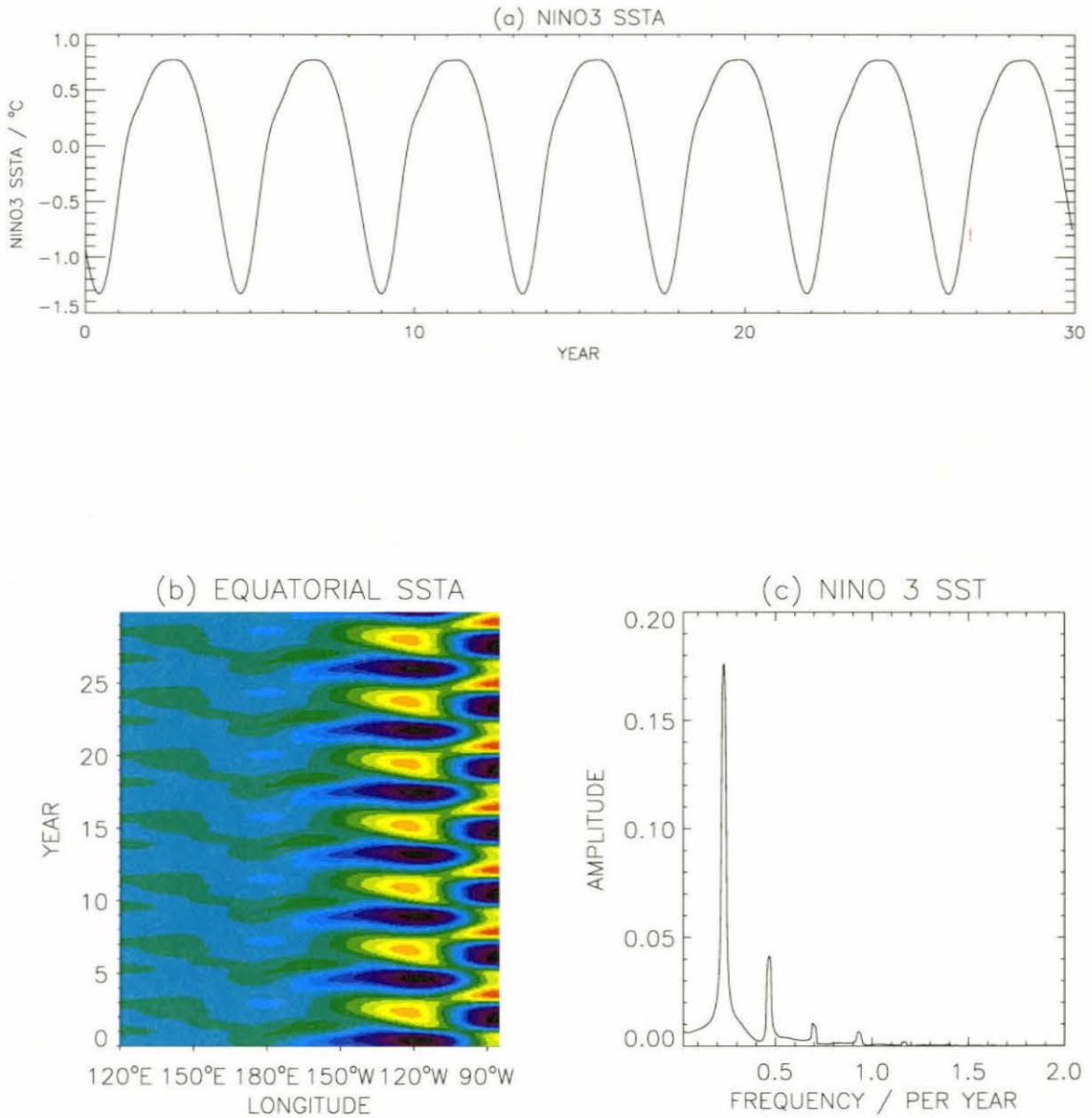


Figure 3.7: The results from the coupled model run forced by the estimated external windstress for April. The coupling was $K_Q W = 0.0097\, m^3\, s^{-4}\, ^\circ C^{-1}$. The spectrum in (c) was calculated from a 200 year run of the model.

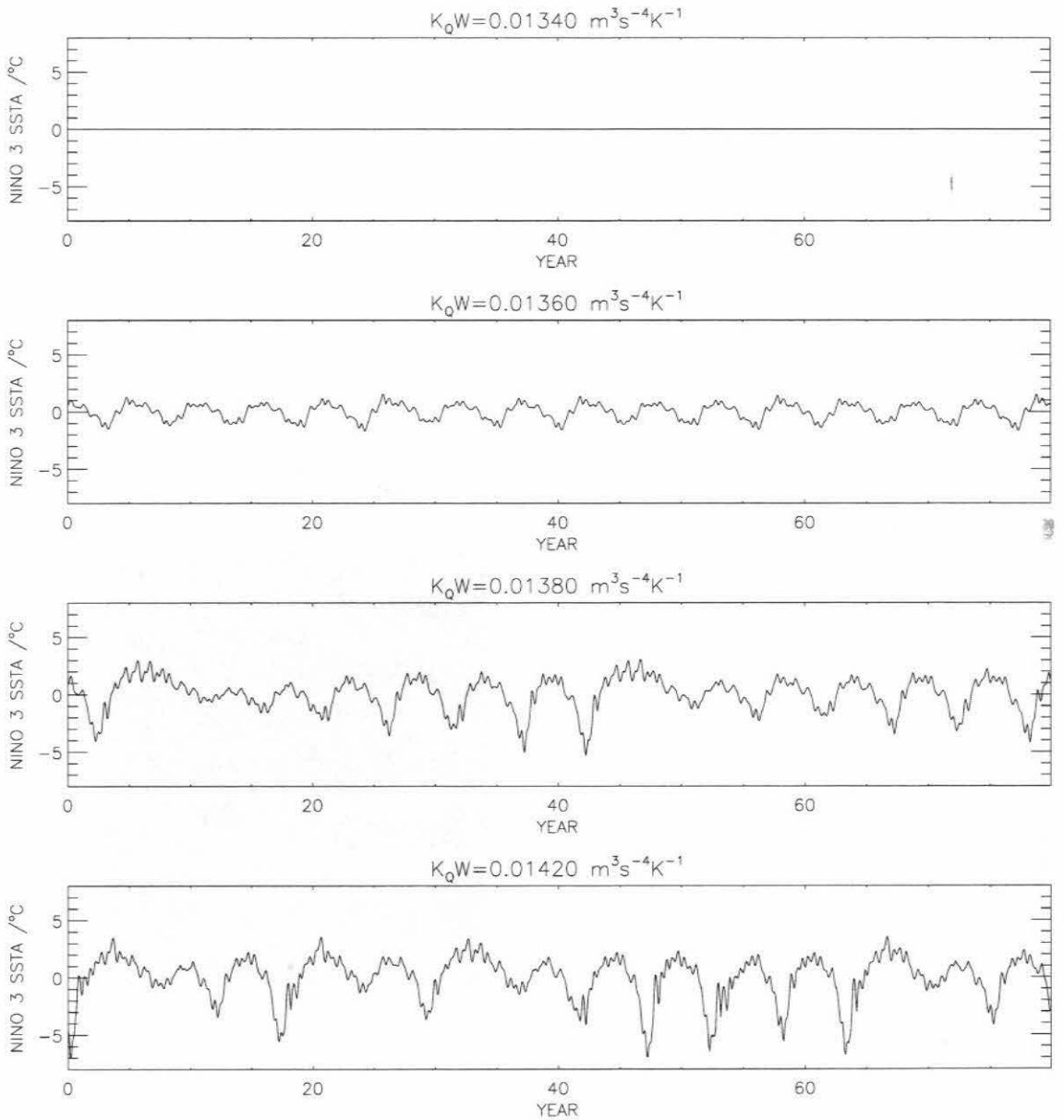


Figure 3.8: The NINO 3 SST anomaly from coupled runs of the model.

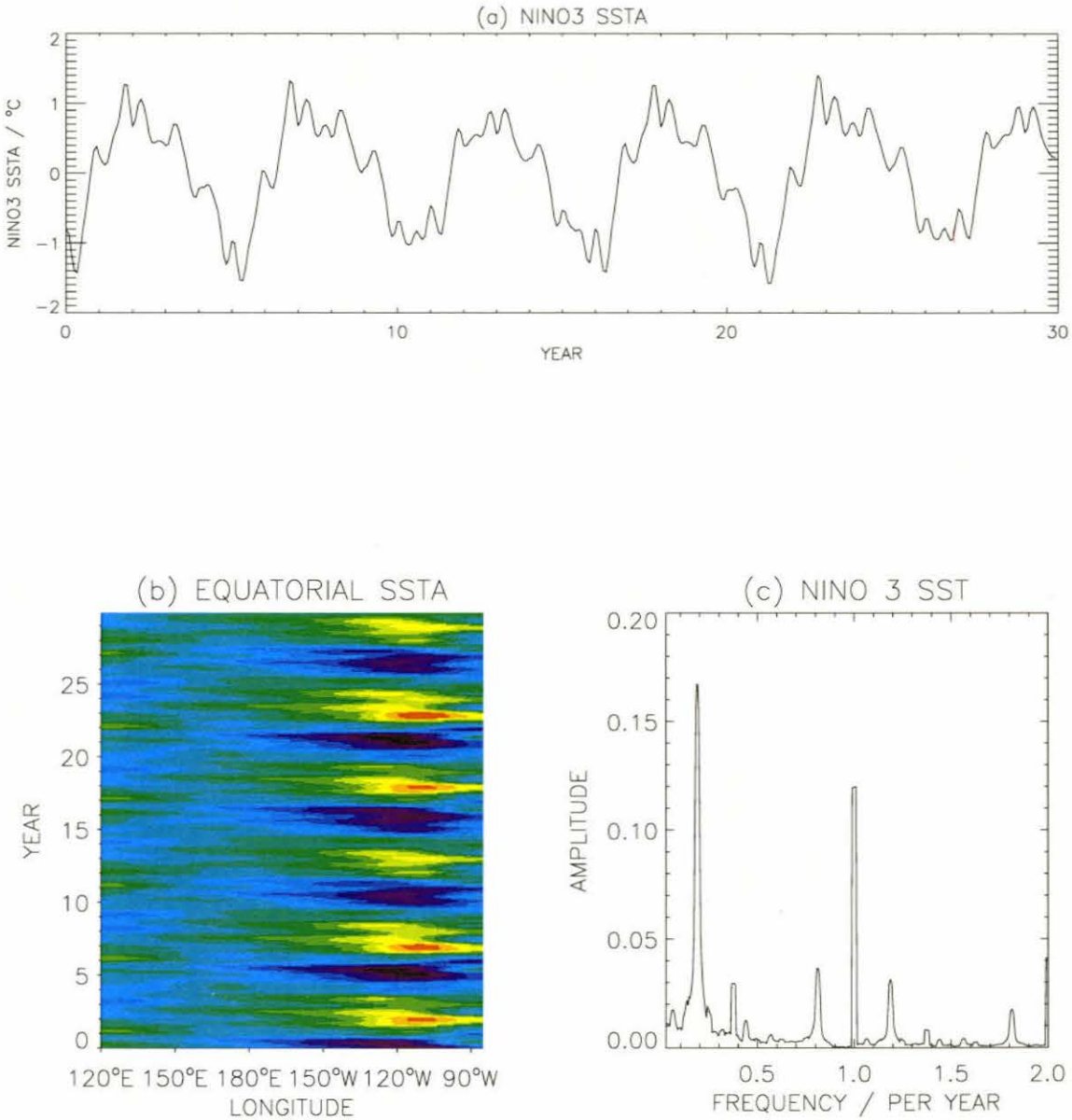


Figure 3.9: The results from a coupled model run forced with the estimated external seasonal windstress. The coupling was $K_Q W = 0.0136\, m^3\, s^{-4}\, ^\circ C^{-1}$. The spectrum in (c) was calculated from a 200 year run of the model.

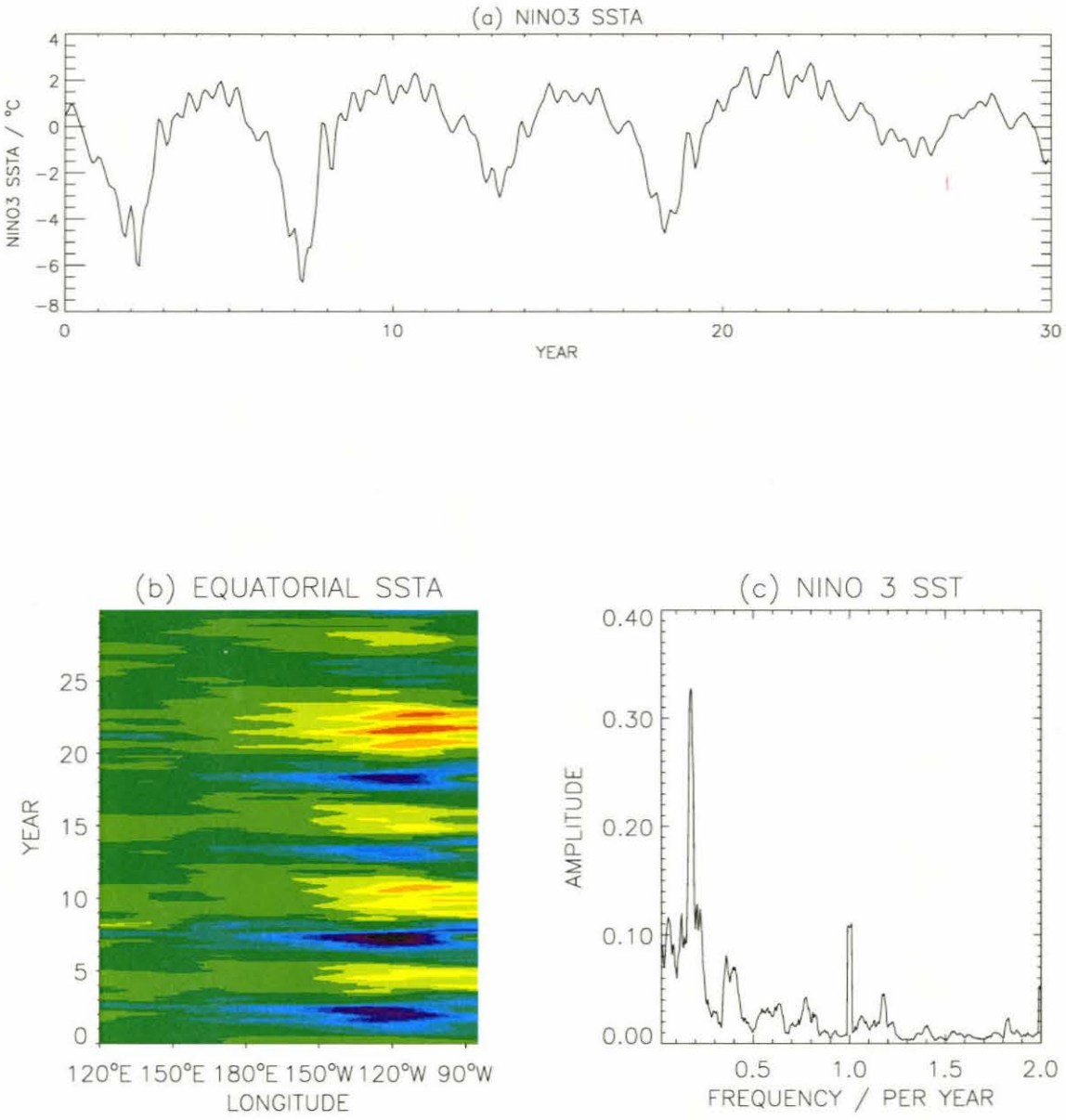


Figure 3.10: As Fig. 3.9 but with $K_Q W = 0.0140\, m^3\, s^{-4}\, ^\circ C^{-1}$.

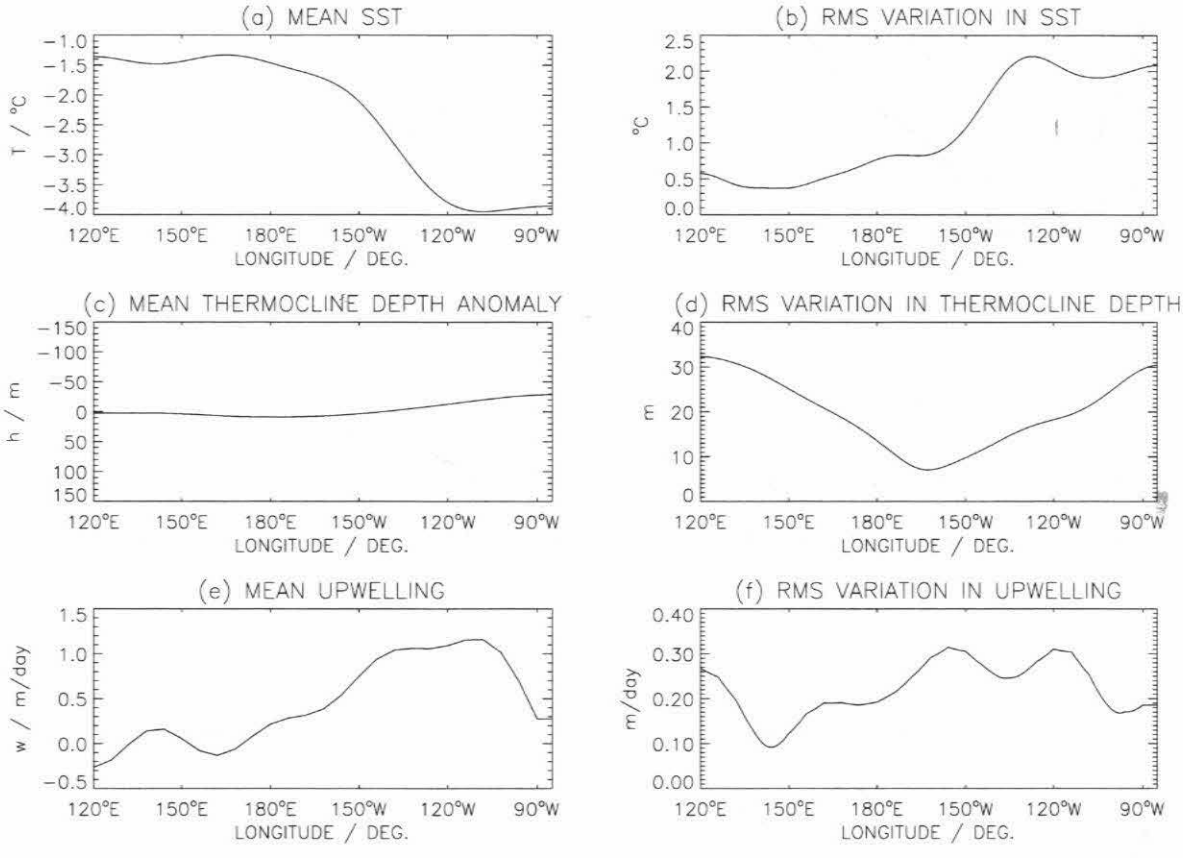


Figure 3.11: Equatorial profiles of variable means and standard deviations for the coupled case. The coupling was $K_Q W = 0.0136\, m^3\, s^{-4}\, ^\circ C^{-1}$. (a) mean SST along the equator; (b) standard deviation in equatorial SST; (c) mean equatorial profile in thermocline depth anomaly; (d) standard deviation in thermocline depth; (e) mean equatorial profile of upwelling; (f) standard deviation in the equatorial upwelling.

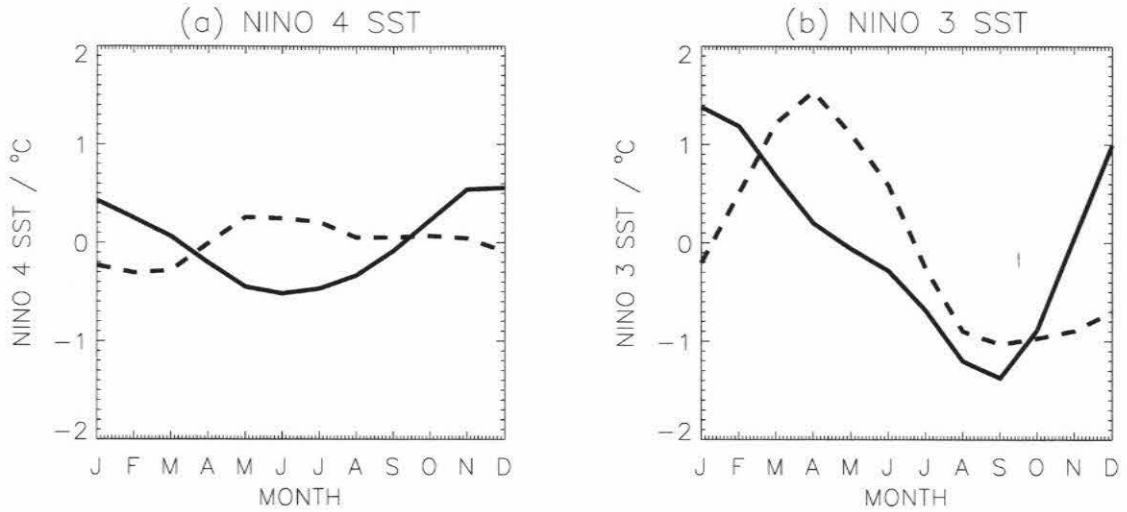


Figure 3.12: The observed seasonal cycle in SST (*dashed line*) and the seasonal cycle produced by the coupled model (*solid line*). The coupling was $K_Q W = 0.0136\, m^3\, s^{-4}\, ^\circ C^{-1}$.

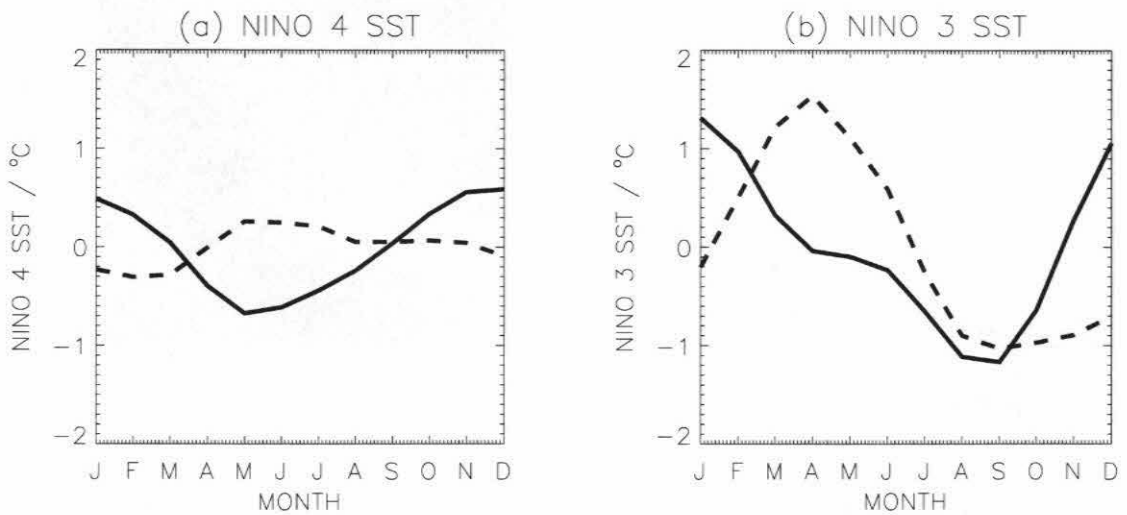


Figure 3.13: As Fig. 3.12 but with $K_Q W = 0.0140\, m^3\, s^{-4}\, ^\circ C^{-1}$.

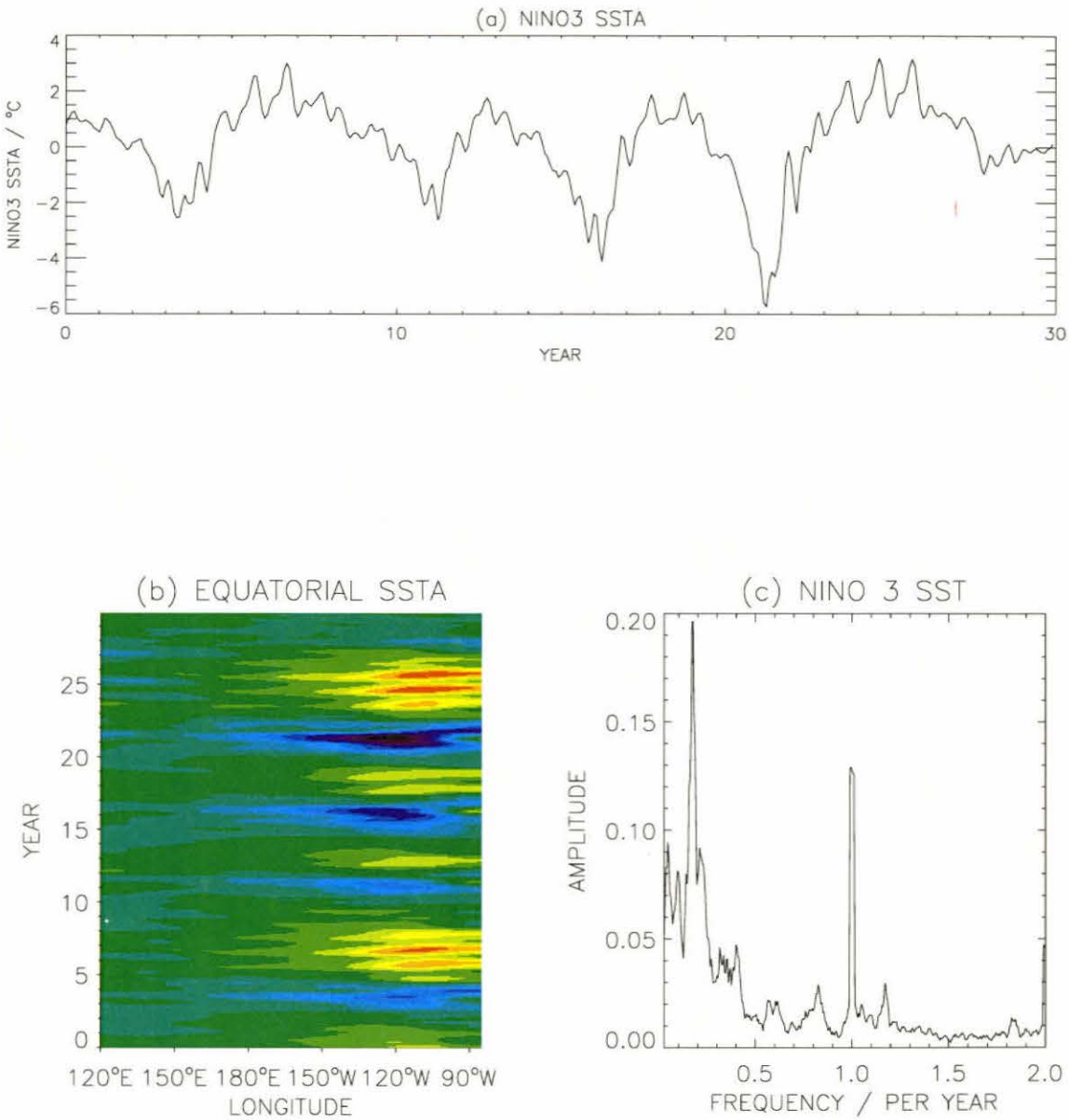


Figure 3.14: As Fig. 3.9 but with stochastic wind forcing. The model is above the bifurcation.

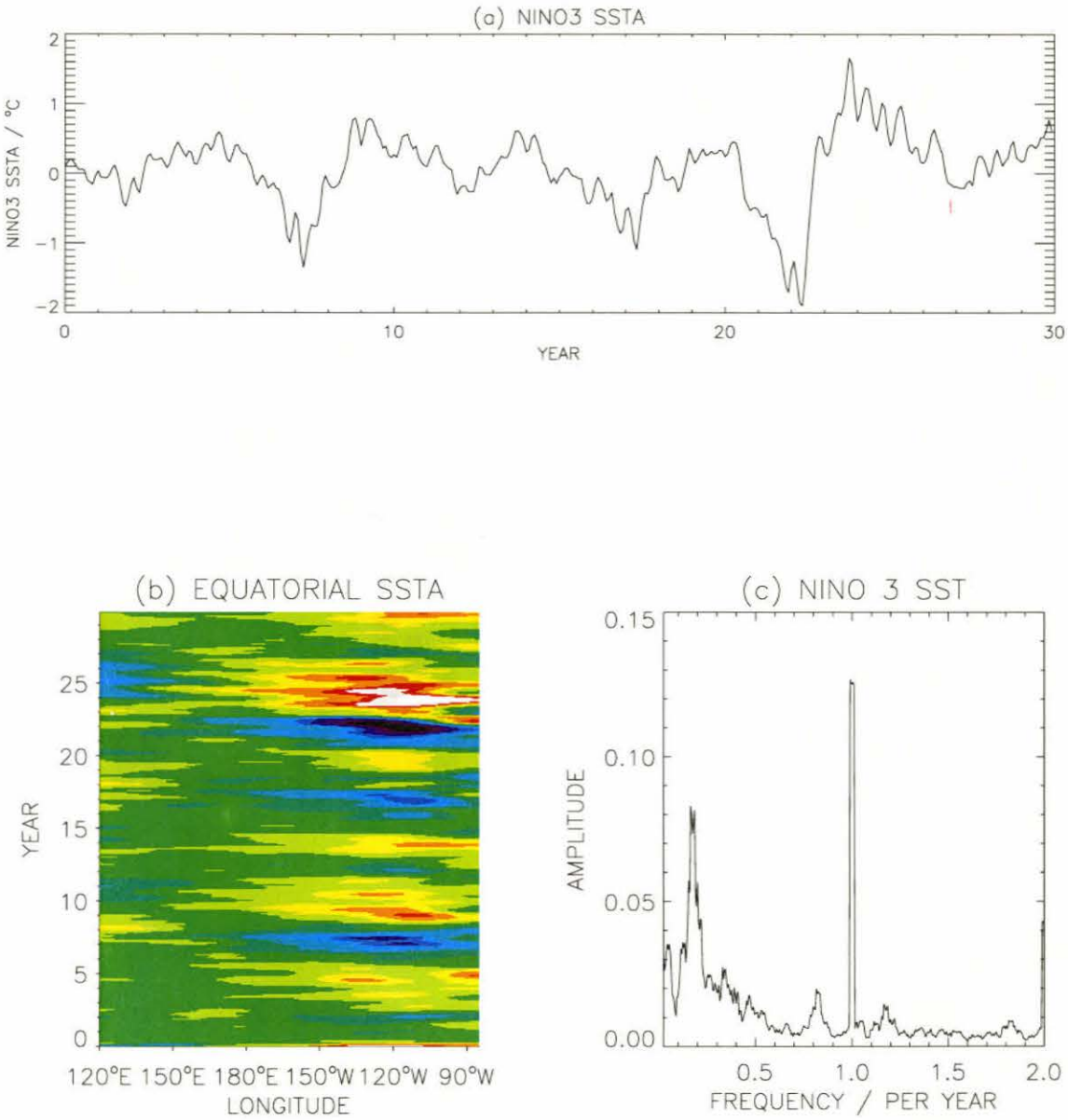


Figure 3.15: As Fig. 3.9 but with stochastic wind forcing and $K_Q W = 0.0134\, m^3\, s^{-4}\, ^\circ C^{-1}$. The model is below the bifurcation.

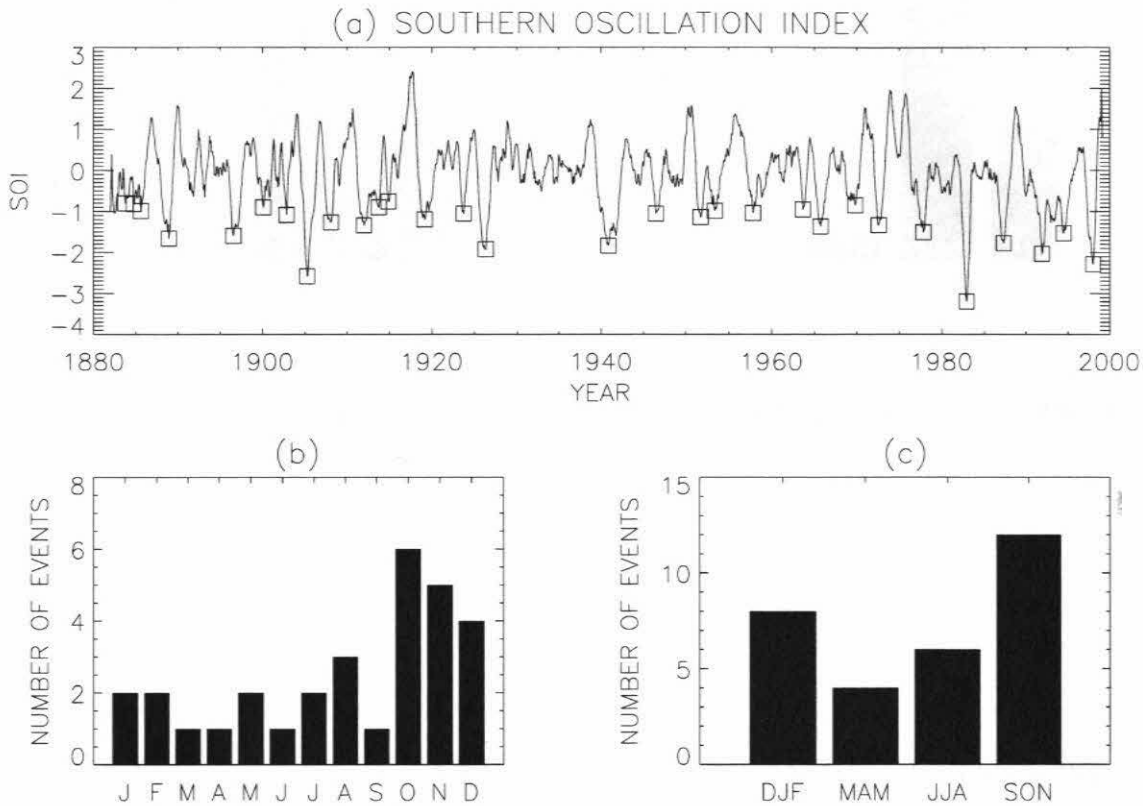


Figure 3.16: (a) The Southern Oscillation Index (normalized pressure difference between Tahiti and Darwin) since 1882. The index has been smoothed with a 6 month moving average filter. The peaks in 30 El Niño events occurring in this time period are marked with squares. (b) The monthly distribution of the event peaks. (c) The seasonal distribution of the event peaks.

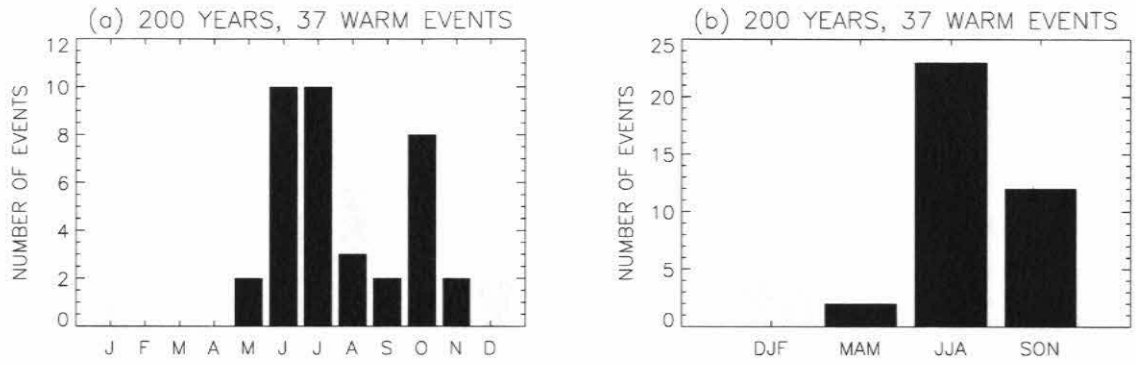


Figure 3.17: The distribution of El Niño events in the coupled model. The coupling was $K_Q W = 0.0136\, m^3\, s^{-4}\, ^\circ C^{-1}$.

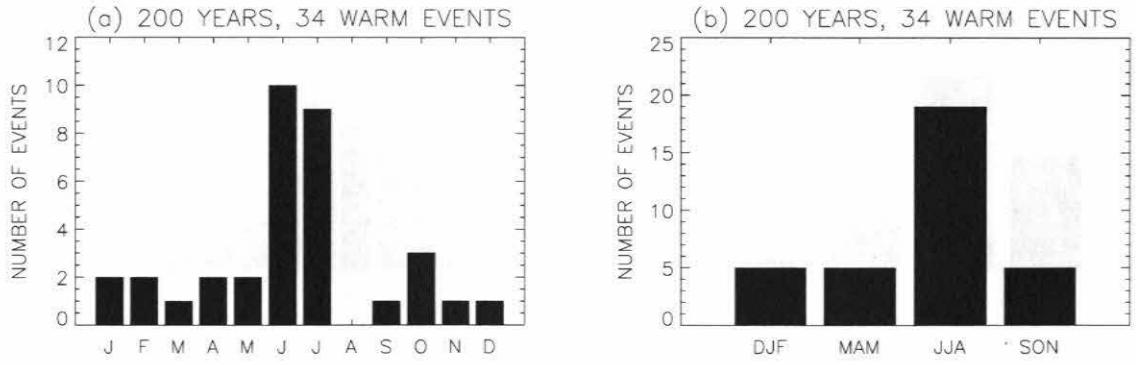


Figure 3.18: As Fig. 3.17 but with $K_Q W = 0.0140\, m^3\, s^{-4}\, ^\circ C^{-1}$.

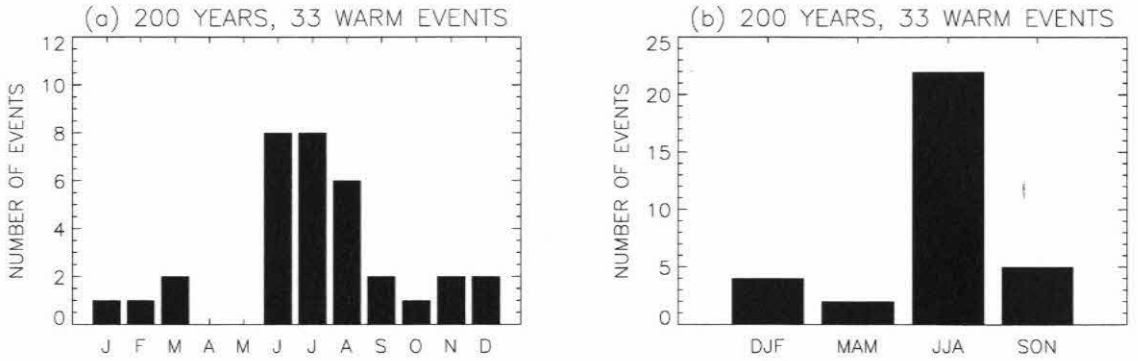


Figure 3.19: As Fig. 3.17 but with stochastic wind forcing.

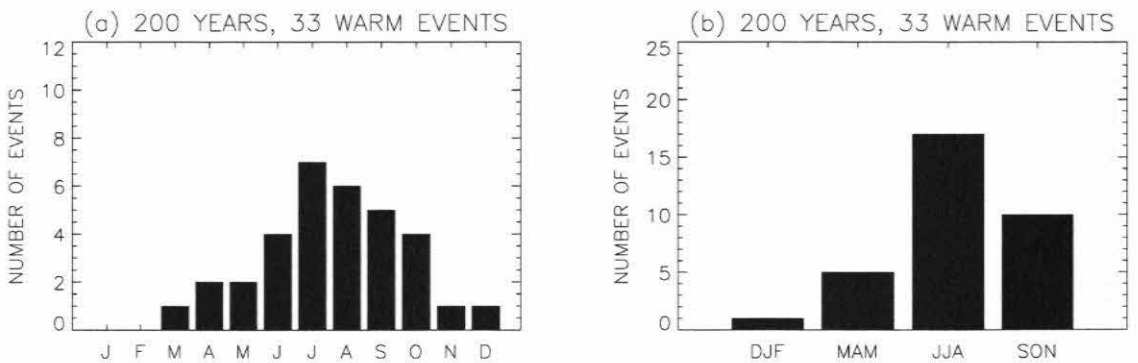


Figure 3.20: As Fig. 3.19 but with $K_Q W = 0.0134\, m^3\, s^{-4}\, {}^\circ C^{-1}$ and stochastic wind forcing.

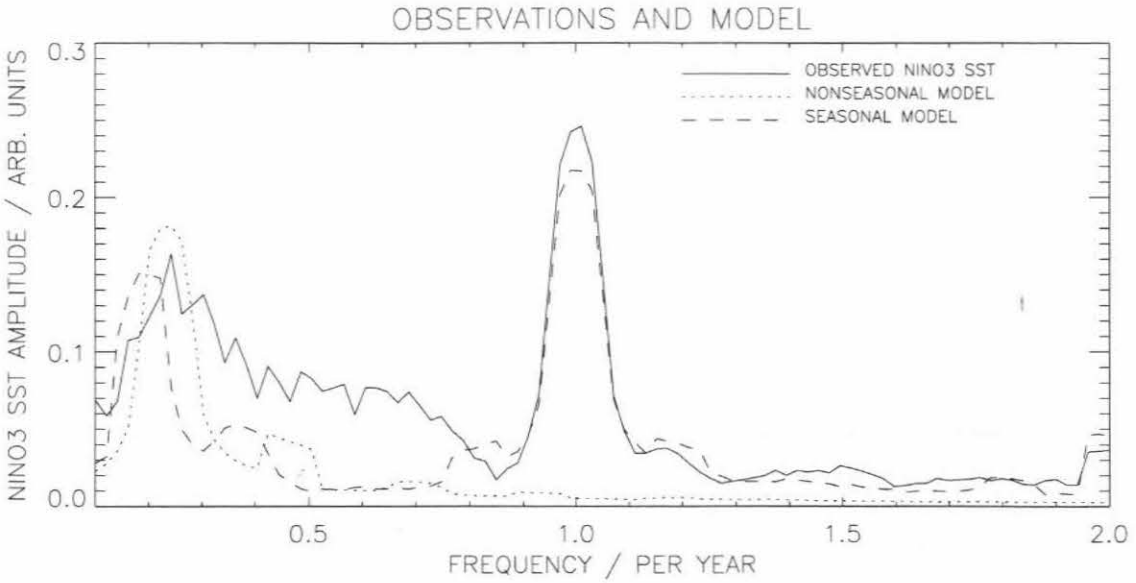


Figure 3.21: The amplitude spectra of observed NINO 3 SST compared with runs of the coupled model. The observed NINO 3 SST is from 1950 to 1999. The spectra of the coupled runs were calculated from 49 year segments.

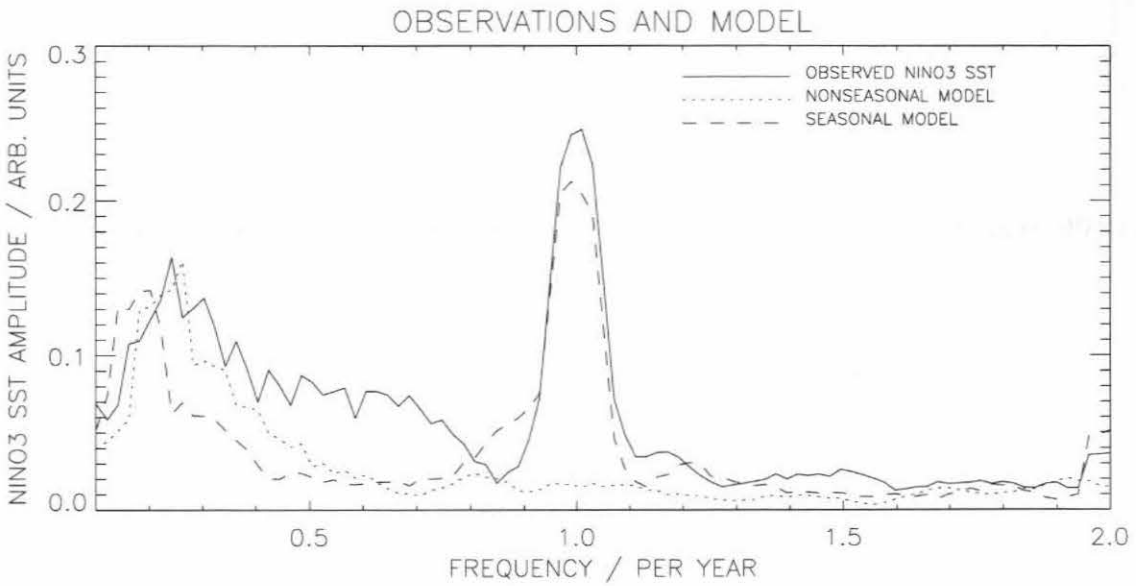


Figure 3.22: As Fig. 3.21 but the coupled models were forced with stochastic wind forcing.

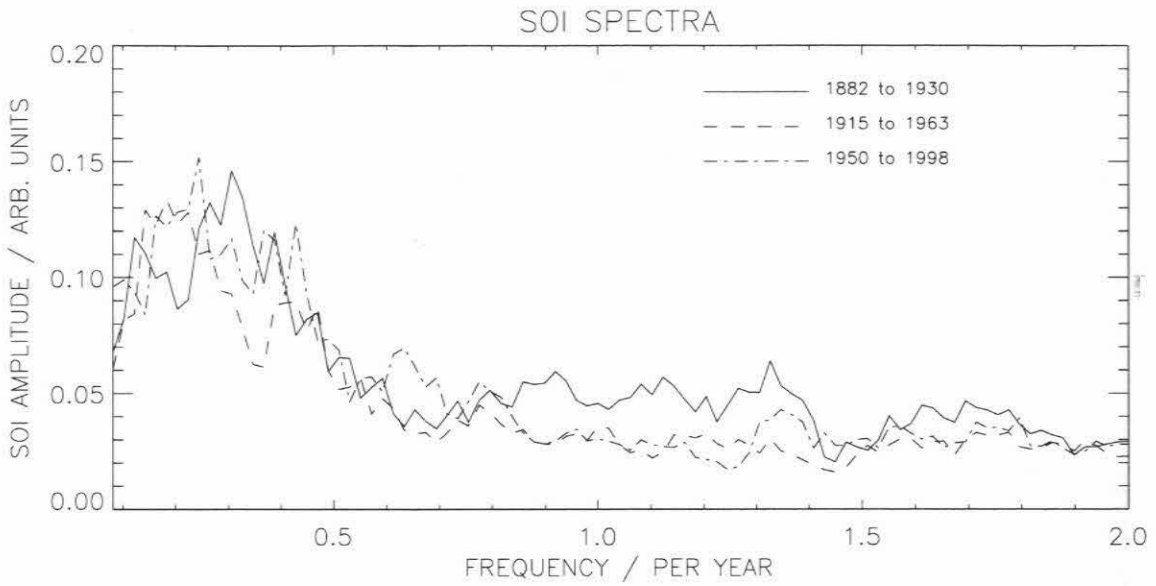


Figure 3.23: The amplitude spectra of the Southern Oscillation Index for three 49 year periods.

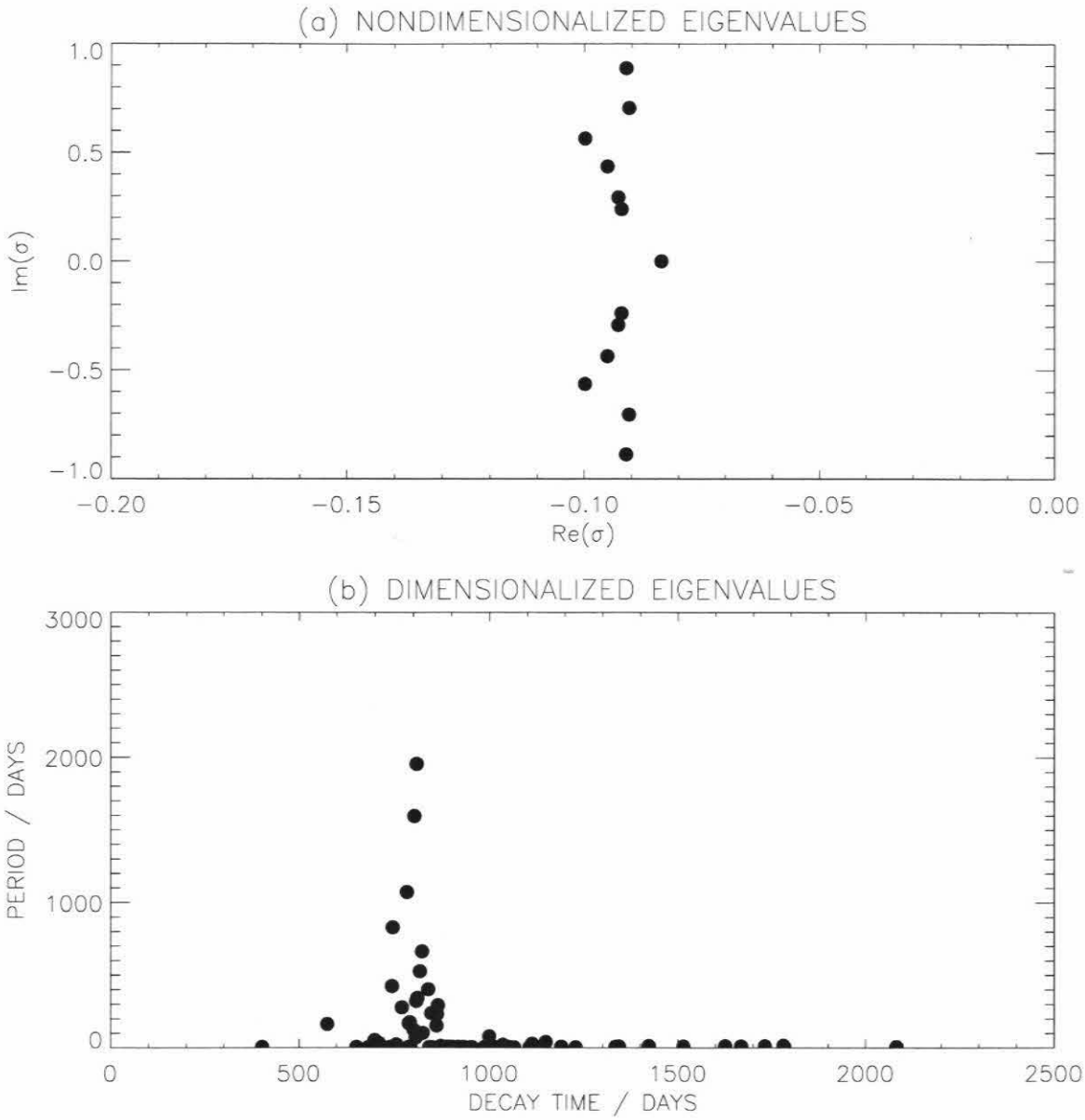


Figure 3.24: Eigenvalues of the uncoupled ocean model. (a) The lower frequency eigenvalues nondimensionalized with the Kelvin wave basin crossing time. (b) All the eigenvalues in the period-decay time plane.

Chapter 4

Response of the Model to Stochastic Forcing

4.1 Introduction

In the early stage of ENSO research the main goal was to understand the oscillatory nature of the phenomenon. Now that a basic understanding of the mechanism of the oscillation has been obtained [NBH+98], more attention is being paid to the effect that stochastic forcing has on the system [KP94, MK96, KM97, Bla97, SSHP98]. In particular, the effect of the Madden-Julian *intraseasonal* atmospheric oscillation (MJO) on the *interannual* ENSO system has been investigated [Zeb89, MK99]. In this chapter it will be argued that the ENSO system is not strongly nonlinear enough to effectively channel power from high frequency forcings to the low frequency ENSO mode. More specifically the 30-60 day component of the MJO has little effect on the dominant 3-7 year component of ENSO.

4.2 Method

The investigation was conducted using an intermediate coupled model (ICM) of the tropical Pacific ocean-atmosphere system based on previous ICMs [ZC87, JN93a]. The model is described in Chapters 2 and 3. The ocean model consisted of a linearized shallow-water model on a beta-plane and a nonlinear equation for temperature in a fixed depth mixed layer. Horizontal and meridional advection of SST was included and vertical entrainment into the mixed layer was parameterized. The atmospheric component of the model was a linear, steady-state Gill-type model [Gil80].

The stochastic component of the atmospheric noise was estimated as follows. The reconstructions of SST in the tropical Pacific for the period 1961-1994 [SRLS96] and the Florida State University pseudo-windstress data set for the same period [GO81] were used to estimate the coupling of the SST to the model's atmospheric component. That part of the windstress variance that could be explained by this linear, atmospheric model was then substracted to leave the residual windstress. Figure 4.1 shows the power spectrum of this residual windstress field averaged over the NINO 4 region (150°W-90°W, 5°S-5°N). The spectrum of the residual windstress is red. This reddening is probably due to the ocean acting as a lowpass filter to atmospheric noise which has a whiter spectrum. A stochastic model of this white atmospheric noise was estimated by projecting the residual noise onto its EOFs. The time series of each EOF's coefficient are uncorrelated by construction so each time series can be modeled independently. The power spectrum of each time series was calculated and the power in higher frequencies (4 to 6 $year^{-1}$) was used as an estimate of the "noise floor." New stochastic time series with white spectra defined by this noise floor were then constructed. An example of a noise floor estimate is shown in Fig. 4.1 by the dashed line.

Since the purpose of the study was to investigate the response of the model to stochastic forcing of different frequency ranges, the stochastic forcing was lowpass filtered to re-

move all frequency components with periods less than 6 months. The white noise was also highpass filtered to remove frequency components with periods greater than 6 months. The three noise products that resulted will be referred to as *white noise*, *low frequency noise* and *high frequency noise* respectively.

Three model regimes were studied. The subcritical regime has a coupling strength just below the value at which the model displays sustained interannual oscillations when forced with a constant easterly windstress. The supercritical regime has a coupling strength high enough for the model to oscillate without noise excitation. In the supercritical seasonal case the model displays sustained oscillations when forced with an annual cycle in windstress.

4.3 Results

Each regime of the model was forced with the three noise products. The results are shown in Figs. 4.2, 4.3 and 4.4. Figure 4.2 shows the subcritical case. Without noise excitation the model shows no variability (Fig. 4.2b). When forced with the white noise product, the model displays a spectral peak corresponding to a period of about 4 years although there is also power spread across the entire interannual range (Fig. 4.2f). The model's response to the low frequency noise is not that different from its response to the white noise (Fig. 4.2i) whereas when the model is forced with the high frequency noise the response is quite different (Fig. 4.2l). A peak at 4 years does appear but its amplitude is less than 20% of the peak's amplitude for the other two noise driven cases. There is some power at other interannual frequencies, but the amount is negligible compared to the response to white noise and low frequency noise. The small amount of power at interannual frequencies is evidence of nonlinear processes channeling power from the higher frequencies excited by the forcing; however, these processes do not seem to be very efficient leading only to a very low amplitude excitation at lower frequencies. Inspection of the time series of

NINO 3 SSTA in Fig. 4.2k shows that the high frequency forcing does produce a small high frequency response in the model. Such a response can be explained by purely linear processes.

Figure 4.3 shows the results obtained with the model in the supercritical regime. The unforced model displays a sustained oscillation with a period of about 4 years as well as a smaller peak with a period of just over 2 years (Fig. 4.3c). When forced by white noise, the 4 year peak is broadened and again power appears at all interannual frequencies (Fig. 4.3f). As in the subcritical case the response of the model to the low frequency noise is similar to its response to the white noise (Fig. 4.3i). However, when the supercritical model is forced with only the high frequency components of the stochastic forcing, there is a little difference between its behavior and its behavior when no forcing is present (Fig. 4.3l). The amplitude of the dominant peak is marginally smaller but that peak and the secondary peak are no broader than in the noise-free case.

The experiments described above were conducted in the absence of a seasonal cycle. In reality interaction between the seasonal cycle and the interannual ENSO cycle gives rise to interesting behavior not observed in the nonseasonal case [JNG94, JNG96]. To investigate how the seasonal cycle might affect the response of the model to stochastic forcing, the model was run with seasonal forcing in addition to stochastic forcing. The coupling was chosen to place the seasonally forced model just above the bifurcation where it is capable of self sustained interannual oscillation. The results of these experiments are shown in Fig. 4.4. The spectrum of the seasonal forcing is shown in Fig. 4.4a. The response of the model is shown in Fig. 4.4c. The dominant interannual mode has a period of slightly more than 5 years. The secondary interannual peak corresponds to a frequency of just over 2.5 years. The amplitude of the annual cycle in NINO 3 SST is comparable to the amplitude of the interannual oscillation. When forced with the white noise product the annual cycle remains unchanged, but the primary and secondary interannual peaks are broadened (Fig. 4.4f). As in the nonseasonal cases power is present at all interannual

frequencies. When only the low frequency components of the stochastic forcing are used to drive the model, its response is quite similar to its response to the white noise forcing (Fig. 4.4i). When only the high frequency components of the stochastic forcing are present, the response of the model is different to the white noise case. The dominant interannual peak is not as broad as it is in the white noise case although it does show more broadening than when the high frequency noise forces the model in the absence of a seasonal cycle. Also there is less power present at other interannual frequencies than in the case of white noise forcing.

One of the most important sources of intraseasonal variability in the tropical Pacific is the Madden-Julian oscillation which has a spectral peak corresponding to a period of 30-60 days [MJ71, MJ72]. The ICM was forced with the high frequency stochastic windstress supplemented with a more narrow band stochastic forcing with power in the 30-60 day range. The amplitude spectrum of this MJO-like forcing is shown in Fig. 4.5.

The result of forcing the supercritical model with this MJO-like forcing in the absence of a seasonal cycle is shown in Fig. 4.6. The amplitude of the primary interannual peak is reduced but there is little peak broadening although there is some transfer of a small amount of power to other interannual frequencies. The time series of the NINO 3 SST anomaly shows that the frequency of the dominant interannual oscillation is largely unchanged although its amplitude is modulated on a decadal timescale. Figure 4.7 shows the response of the seasonally forced model to the additional MJO-like forcing. Again there is little peak broadening of the interannual oscillation. Power appears at interdecadal frequencies and this has the effect of modulating the amplitude of the interannual oscillation. Comparison of Fig. 4.6 and Fig. 4.4k and 4.4l and Fig. 4.7 and Fig. 4.5k and 4.5l reveals that the additional power in the wind forcing in the 30-60 day range does not change the response of the model much from the case of the standard high frequency noise.

4.4 Discussion and Conclusions

The response of the subcritical model to the different stochastic products indicates that there are nonlinear processes operating that transport energy from intraseasonal frequencies to interannual frequencies. However, the low amplitude of the response to high frequency forcing (Fig. 4.2l) suggests that these processes are not particularly efficient. Even with stronger coupling (Fig 4.3) the effect of stochastic forcing at intraseasonal frequencies on the interannual variability of the model is not great. In both cases the similarity of the response of the model to the white noise product and the low frequency noise product is consistent with the type of response that would be expected from a largely linear system.

The presence of a seasonal cycle does appear to increase the effectiveness of high frequency noise at modifying the low frequency response of the model, but still it is the stochastic forcing at interannual frequencies that has by far the biggest impact on the interannual response of the model.

The addition of extra noise power at frequencies corresponding to periods of 30 to 60 days does not have a major impact on the response of the model. This implies that the intraseasonal variability associated with the Madden-Julian oscillation may not have a significant impact on the interannual variability associated with ENSO. This result is consistent with previous work [Zeb89] which found that forcing at intraseasonal frequencies had only a marginal effect on the statistics and the predictability of the Cane-Zebiak ICM. Based on the results in this work a more general conclusion can be drawn that only interannual forcing is effective at modifying the interannual response of the model.

The spatial patterns of variability characteristic of MJO may project onto the modes of the system with interannual frequencies, but it is the low frequency variability of these patterns which will have a significant impact on ENSO, not the variability in the 30-60 day range.

The results and hence the conclusions presented above are model dependent. A model with stronger nonlinear processes could be more efficient at channeling power from intraseasonal to interannual frequencies. However, some studies of the statistics of ENSO observations have concluded that ENSO can be described as a stable, linear, noise driven system [Pen96, Bur99]. While this modeling study does not address the issue of whether the real system is in the stable or unstable regime, it does suggest that the response of the system to noise is largely linear in both regimes.

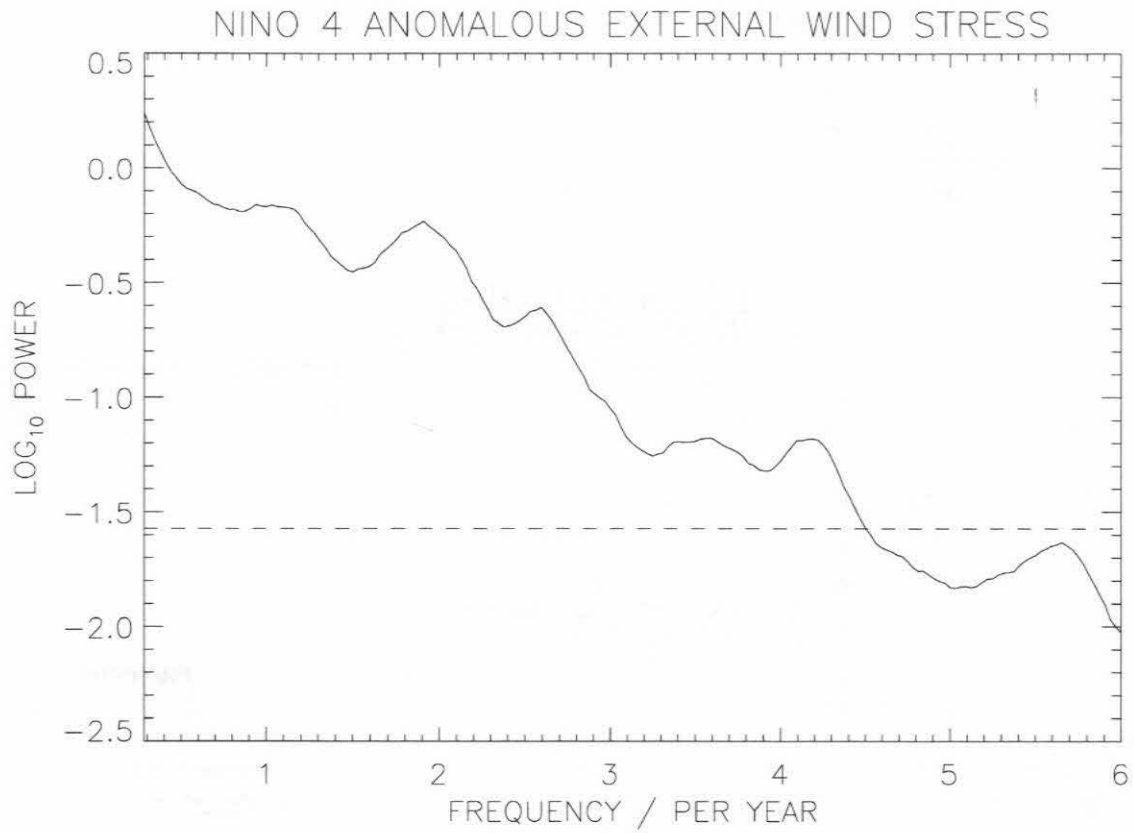


Figure 4.1: The power spectrum (*solid curve*) of the residual windstress that could not be explained by a linear model relating windstress to SST. The dashed line shows the estimated "noise floor" which was used to construct a stochastic time series with a white spectrum.

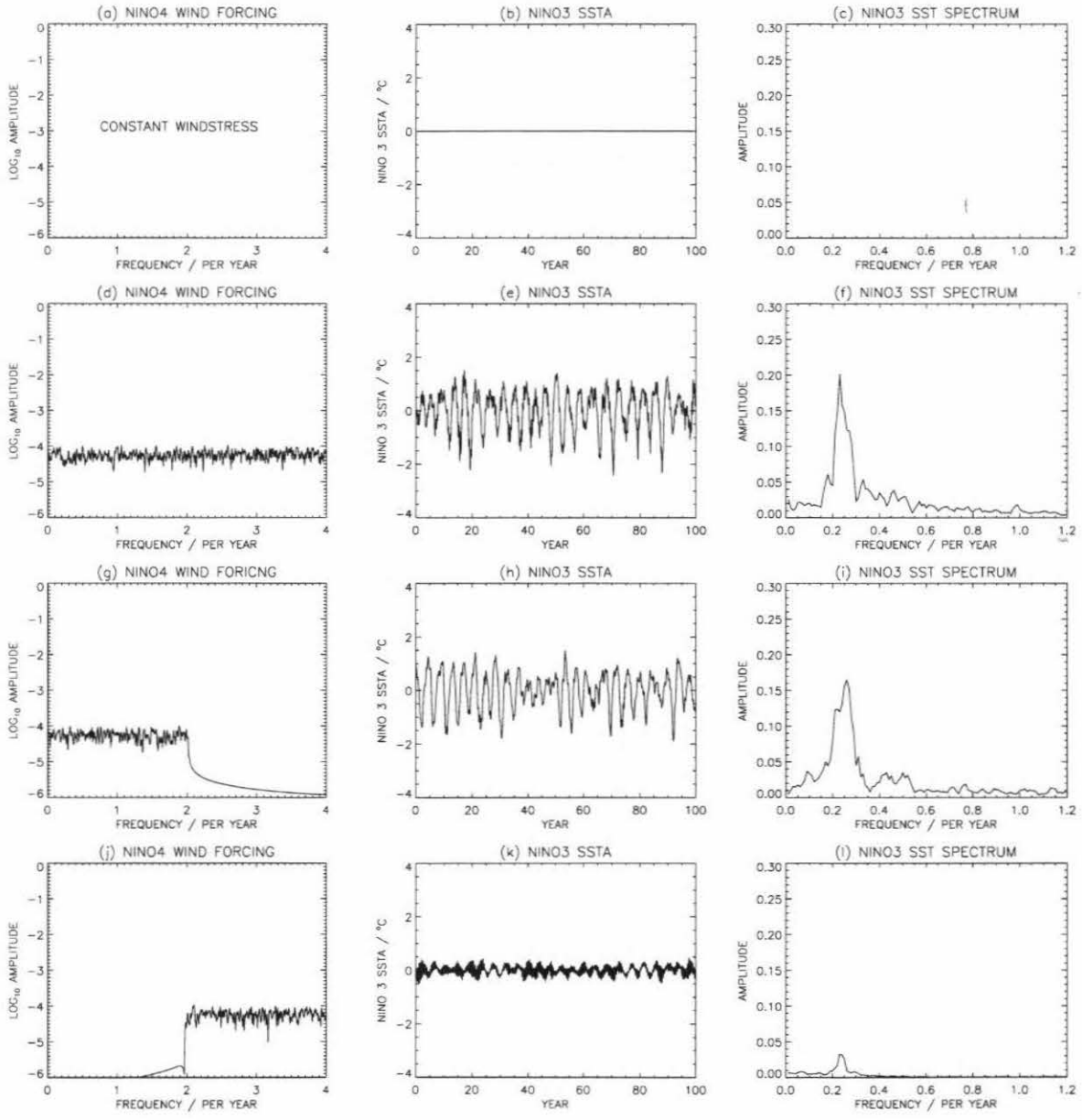


Figure 4.2: The response of the coupled model to stochastic forcing when the coupling strength puts it just below the bifurcation.

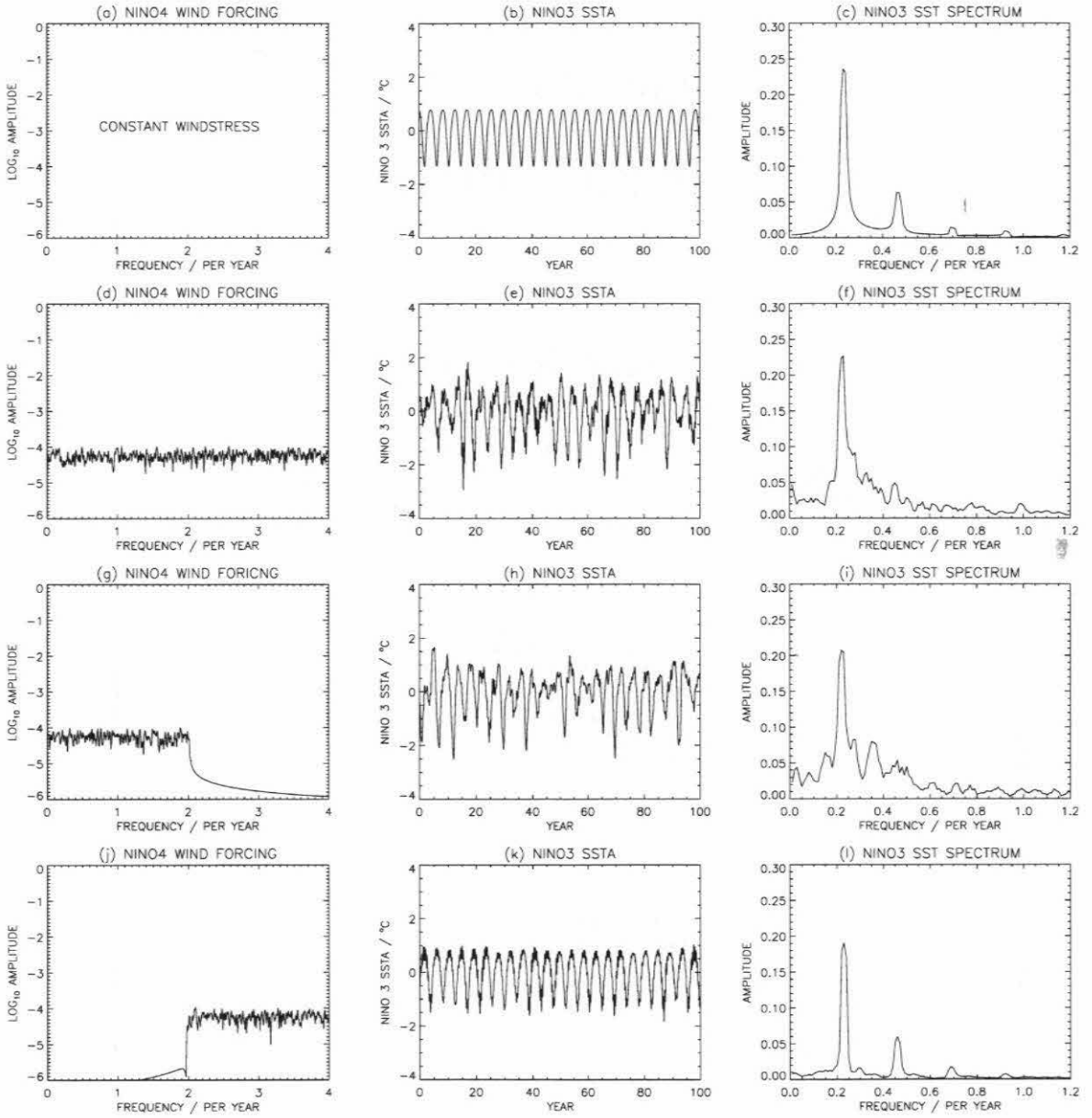


Figure 4.3: The response of the coupled model to stochastic forcing when the coupling strength puts it just above the bifurcation.

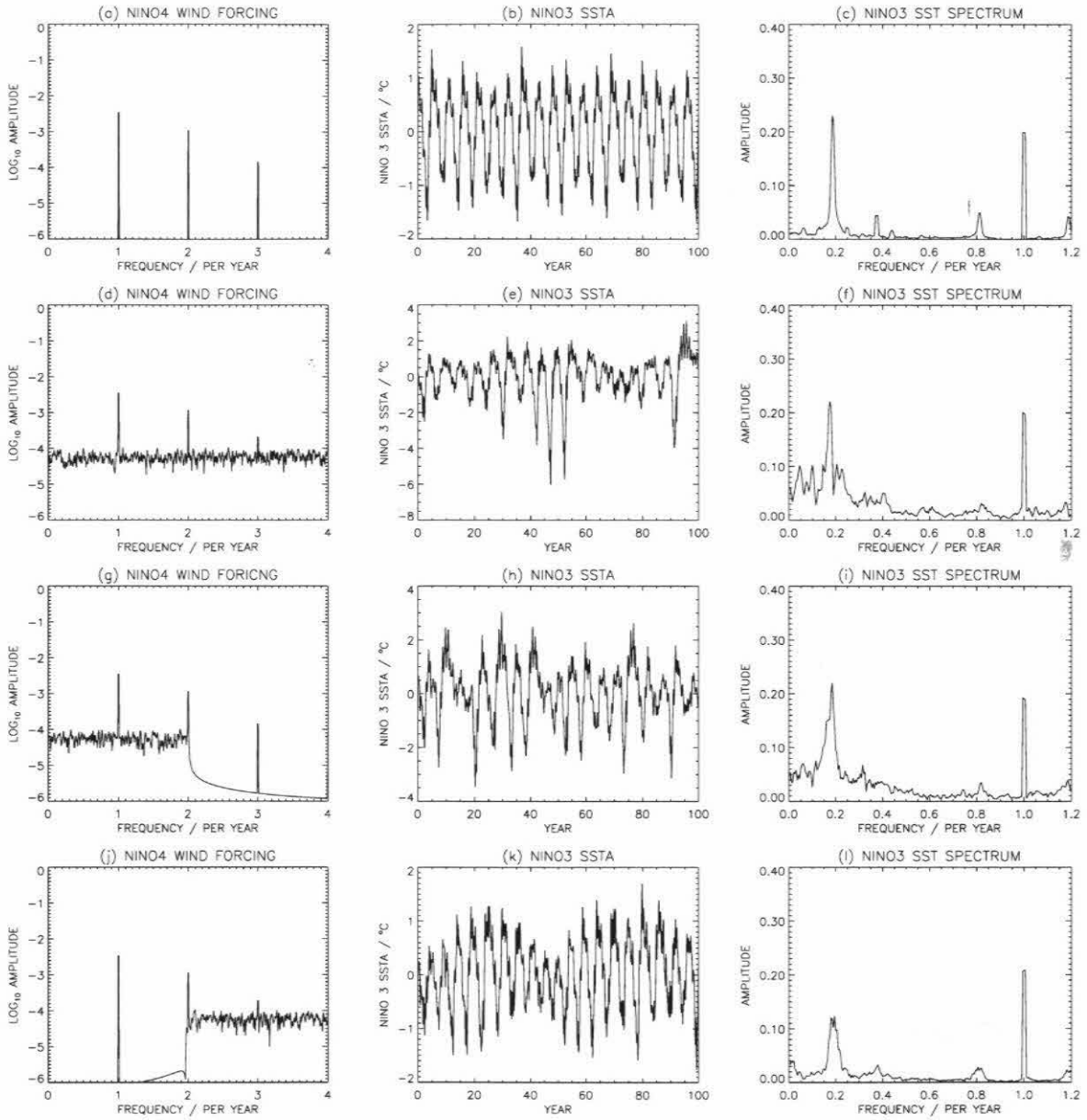


Figure 4.4: The response of the coupled model to stochastic forcing and seasonal forcing when the coupling strength puts it just above the bifurcation.

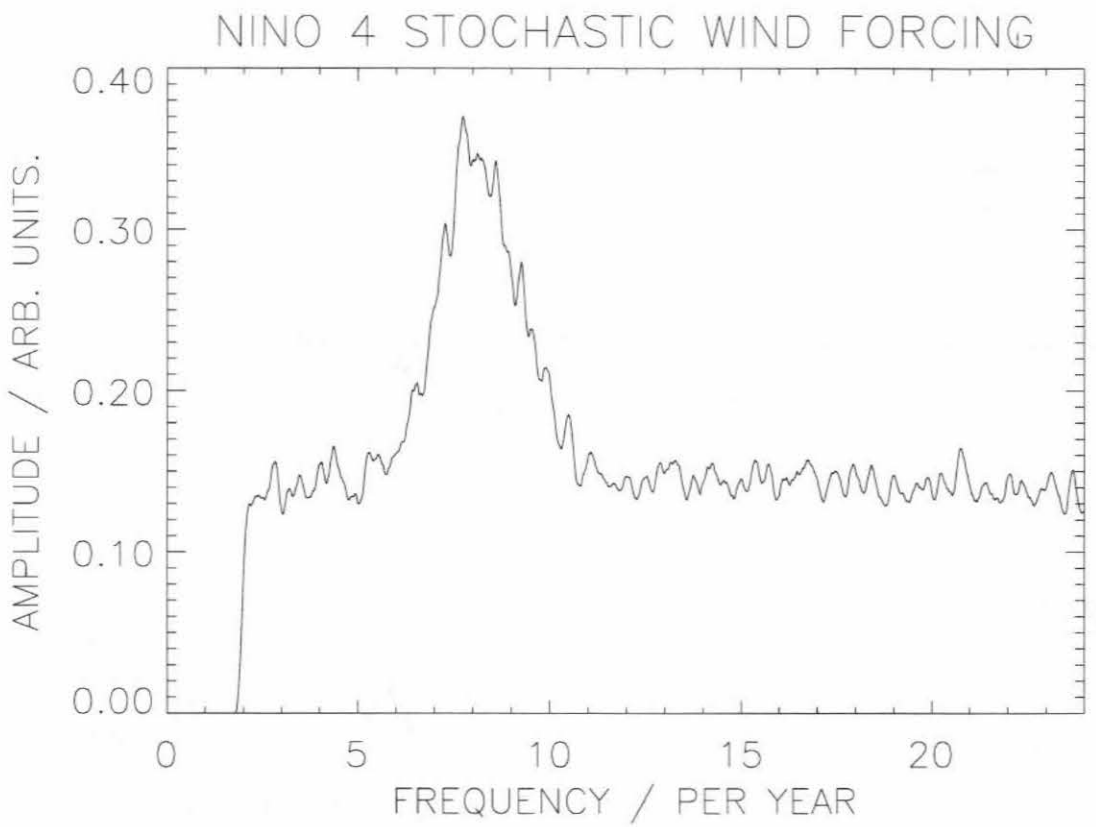


Figure 4.5: The amplitude spectrum of the "MJO" stochastic forcing with a spectral peak in the 30-60 day period range.

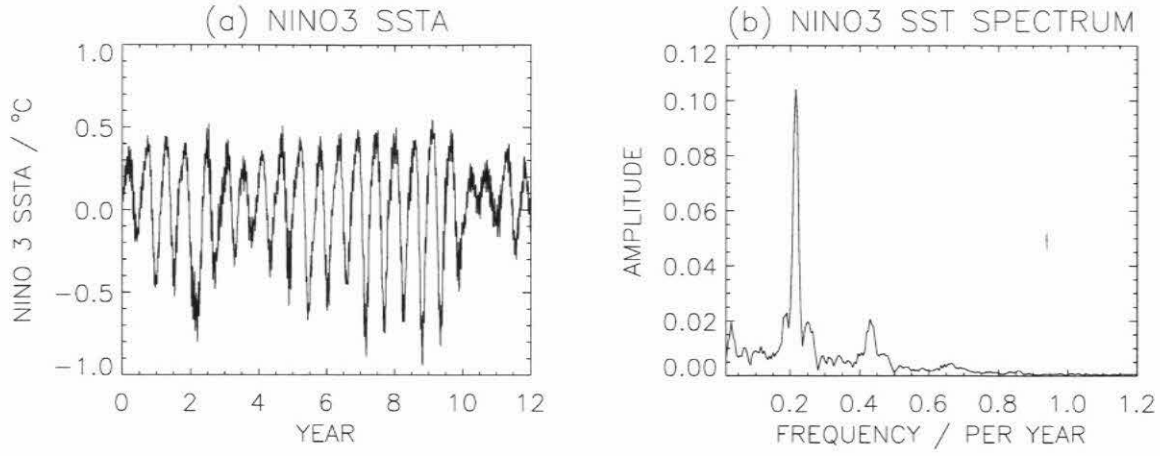


Figure 4.6: The NINO 3 SST anomaly and the amplitude spectrum of NINO 3 SST for the model forced with an MJO-like forcing without a seasonal cycle.

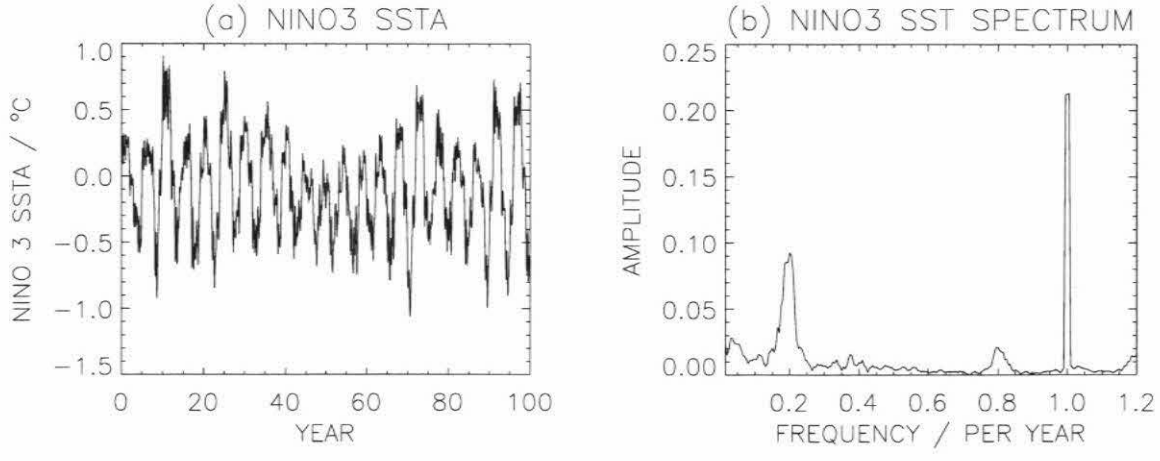


Figure 4.7: The NINO 3 SST anomaly and the amplitude spectrum of NINO 3 SST for the model forced with an MJO-like forcing with a seasonal cycle.

Chapter 5

Dimensionality of ENSO

5.1 Introduction

One of the outstanding questions concerning the dynamics of El Niño-Southern Oscillation (ENSO) is whether its irregularity is the result of chaotic dynamics or due to stochastic forcing in the form of atmospheric transients [NBH+98]. The hypothesis of chaotic dynamics implies that the unpredictable behavior of the system is actually due to *low dimensional* deterministic dynamics of the coupled ocean-atmosphere system. On the other hand, the hypothesis of stochastic forcing holds that the irregularity is due to the *high dimensional* stochastic forcing processes.

In principle the dimensionality of a dynamical system can be determined from a time series of a scalar variable [GP83, KBA92, Aba95]. In practice estimating the dimension of a system from observed data is fraught with difficulties.

In this chapter the ENSO model described in Chapters 2 and 3 will be used to study the feasibility of estimating the dimension of the ENSO system from observations.

5.2 The Datasets

The datasets used were the NINO 3 SST time series from two 1000 year runs of the intermediate coupled model. One run came from the model in its "irregular" regime in which it exhibits irregular behavior in the absence of stochastic forcing. The other run came from the stochastically forced model. The stochastically forced model was above the bifurcation at which it begins to display self-sustaining interannual oscillations but was also forced with random wind noise with a white spectrum. Figures 5.1 and 5.2 show 100 year samples of the NINO 3 SST anomaly time series and amplitude spectra of NINO 3 SST for the 1000 year runs.

Two methods for determining the dimensionality were used; the Grassberger-Procaccia algorithm for estimating the correlation dimension and the method of false nearest neighbors, for estimating the embedding dimension.

Both of these methods require a pseudo-state space reconstruction, that is the reconstruction of trajectories in a space which is topologically equivalent to the true state space.

5.3 The Method of Delays and Mutual Information

The method of delays provides a means to reconstruct the attractor of a system in a multidimensional state space using a time series of a scalar variable. The theoretical foundation for the method is the Takens-Mañé embedding theorem [Tak81, MnY81]. Suppose that s_t is a scalar time series generated by a multidimensional dynamical system. The topological structure of the multidimensional dynamics can be reconstructed by constructing data vectors, $\mathbf{y}_t$, from s. The d-dimensional data vectors are given by

$$\mathbf{y}_t = [s_t, s_{t+T}, s_{t+2T}, ..., s_{t+(d-1)T}] \tag{5.1}$$

where T is a time delay. Note that in terms of the information contained in the vector this is equivalent to constructing a data vector from time derivatives of s.

$$\mathbf{x}^*(t) = \left[s, \frac{ds}{dt}, \frac{d^2s}{dt^2}, ..., \frac{d^{d-1}s}{dt^{d-1}}\right]_t \tag{5.2}$$

where all the derivatives are evaluated at time t. In this context d will be called the *unfolding dimension*. The reconstruction obtained using Eq. 5.1 will obviously not be the same reconstruction obtained using Eq. 5.2 but they will be topologically equivalent. Furthermore, they will be topologically equivalent to the true state space for large enough d. The reconstructed state space is often called the *pseudo-state space*. Due to the problems associated with calculating derivatives numerically, Eq. 5.1 is usually used. This introduces the problem of how to choose T.

In the case of an infinite time series of noise free data, any value of T will lead to a faithful pseudo-state space reconstruction. In the absence of this luxury an optimal value must be chosen, and unfortunately there is no formal theoretical result that suggests one. Instead several approaches have been suggested based on a combination of heuristic reasoning and practical experience. The simplest of these approaches is to calculate the autocorrelation function of s_t and choose T to be the lag at the first zero of the autocorrelation. The validity of this approach has been questioned because autocorrelation is an inherently linear measure of dependency and pseudo-state space reconstructions are designed to be applied to nonlinear systems. Therefore, another approach which draws upon information theory has been suggested [FS86]. In this approach the first minimum of the mutual information is taken to be the appropriate value for T. Mutual information is an information theoretic functional that is related to information entropy. For a discrete variable, X, the entropy is defined as

$$H(X) = -\sum_{i=1}^{B_X} p_i \ln p_i \tag{5.3}$$

where the sum is over the B_X states that X can assume and p_i is the probability that X

will be in state i. The joint entropy of two discrete variables, X and Y, is defined as

$$H(X,Y) = -\sum_{i=1}^{B_X}\sum_{j=1}^{B_Y} p_{ij} \ln p_{ij} \tag{5.4}$$

where the sum is over the B_X states that X can assume and the B_Y states that Y can assume and p_{ij} is the probability that X is in state i *and* Y is in state j. Having defined entropy and joint entropy, the mutual information of two discrete variables, X and Y, can be defined as

$$I(X;Y) = H(X) + H(Y) - H(X,Y) \tag{5.5}$$

Mutual information can be thought of as a generalized correlation analogous to the linear correlation coefficient but sensitive to any relationship, not just linear dependence. The bias and errors on the mutual information can be calculated using standard error analysis (see Appendix A).

Once an appropriate delay has been chosen the pseudo-state space can be reconstructed and the dimension estimates can be made.

5.4 Correlation Dimension

The correlation dimension is defined as

$$D_2 = \lim_{\varepsilon \to 0} \frac{1}{\ln \varepsilon} \ln \sum_i p_i^2 \tag{5.6}$$

where p_i is the probability that a point on the attractor lies in the ith box which has side ε. Grassberger and Procaccia [GP83] presented an algorithm for determining D_2 from data. They introduced the correlation integral which they defined as

$$C(\varepsilon) = \lim_{N \to \infty} \frac{1}{N^2} \sum_{ij} \mathcal{H}(\varepsilon - |\mathbf{y}_i - \mathbf{y}_j|) \tag{5.7}$$

where $\mathbf{x}_i$, $i = 1, 2, ..., N$ are the vectors created from the time series using the method of delays and $\mathcal{H}$ is the heaviside step function. Notice that $C(\varepsilon)$ is nothing more than the probability that two points of the attractor lie within the same hyperdimensional box of sidelength ε. The value of D_2 can be estimated by measuring the gradient of $\ln C(\varepsilon)$ plotted against $\ln \varepsilon$. Since the necessary value of d is initially unknown, the procedure is repeated for increasing values of d until the estimate of D_2 saturates.

5.5 Embedding Dimension

The method of false nearest neighbors (FNN) was developed as a means to determine the minimum embedding dimension, M, of a dynamical system from a time series, s_t, where $t = 1, 2, ..., N$ [KBA92, Aba95]. The basic principle behind the method is that if points in a true, M-dimensional, state space are projected onto d dimensions, where $d < M$, then points which are far apart in the true, M-dimensional, space will be close together in d dimensions. These points are referred to as *false nearest neighbors*. The aim of the method is to determine the value of d which eliminates false nearest neighbors. To determine the number of false neighbors in dimension d, the first step is to project the univariate data onto d-dimensions using the method of delays. Let $\mathbf{y}_t$ be the data vectors.

$$\mathbf{y}_t = [s_t, s_{t+T}, ..., s_{t+(d-1)T}] \tag{5.8}$$

where T is the delay time. The next step is to locate the nearest neighbor of $\mathbf{y}_t$ in the d-dimensional space. This neighbor should not be a point on the same short segment of the state space trajectory, that is one that is close to $\mathbf{y}_t$ in time. This condition is best satisified by not considering points occurring in the time interval $[t - KT, t + KT]$ where K is an integer of order unity. The nearest neighbor of $\mathbf{y}_t$ will be designated $\mathbf{y}_k$. If $\mathbf{y}_k$ is a false neighbor of $\mathbf{y}_t$ then, when the data is projected onto $d + 1$ dimensions, it may move out of the neighborhood of $\mathbf{y}_t$. Let R_d be the Euclidean distance between $\mathbf{y}_t$ and

$\mathbf{y}_k$ in d dimensions and let R_{d+1} be the distance between $\mathbf{y}_t$ and $\mathbf{y}_k$ in $d+1$ dimensions. The criterion for deciding that a neighbor is false is that

$$\sqrt{\frac{R_{d+1}^2 - R_d^2}{R_d^2}} > \alpha_0 \tag{5.9}$$

where α_0 is some threshold value. The value of R_{d+1}^2 is given by

$$R_{d+1}^2 = \sum_{i=1}^{d+1} (s_{t+(i-1)T} - s_{k+(i-1)T})^2 \tag{5.10}$$

$$= \sum_{i=1}^{d} (s_{t+(i-1)T} - s_{k+(i-1)T})^2 + (s_{t+dT} - s_{k+dT})^2 \tag{5.11}$$

that is

$$R_{d+1}^2 = R_d^2 + (s_{t+dT} - s_{k+dT})^2 \tag{5.12}$$

Thus, the condition for a point to be classed as a false neighbor given by Eq. 5.9 can be written

$$\frac{|s_{t+dT} - s_{k+dT}|}{R_d} > \alpha_0 \tag{5.13}$$

The value of α_0 must be chosen such that false neighbors can be distinguished. Abarbanel [Aba95] has found that the ability of the criterion to identify false neighbors is insensitive to the value of α_0 but, on the basis of numerical experiments, recommends a value of about 15.

In addition to the condition described by Eq. 5.9 there is another criterion for identifying a false neighbor, namely that the additional distance introduced in raising the dimension by one is of the order of the diameter of the attractor. This condition is expressed as

$$\frac{|s_{t+dT} - s_{k+dT}|}{R_A} > \beta_0 \tag{5.14}$$

where R_A is the "radius" of the attractor, defined as the square root of the variance of the time series, and β_0 is on the order of 2.

5.6 Relationship Between Dimensions

The Fractal Whitney Embedding Prevalance Theorem gives a relationship between the box-counting dimension and the embedding dimension. The box-counting dimension is related to the correlation dimension but is intuitively easier to understand. The theorem states that a projection of the attractor into a space the dimension of which is more than twice the box-counting dimension of the attractor is almost certain to be an embedding. A nonrigorous heuristic derivation of this theorem is as follows [SYC91]. The number of boxes, n, of size ε required to cover an object of box-counting dimension D_0 goes, by definition, as

$$n \sim \varepsilon^{-D_0} \tag{5.15}$$

The expression in Eq. 5.15 is clearly true for a plane where $n \sim \varepsilon^{-2}$ and for a 3D solid where $n \sim \varepsilon^{-3}$ but D_0 does not need to be an integer and indeed it is not when the object is a *fractal*, unlike the embedding dimension which is an integer even for fractal objects. The number of boxes, N, of size ε required to fill the volume of the M-dimensional pseudo-state space that the object is being mapped into goes as

$$N \sim \varepsilon^{-M} \tag{5.16}$$

The fraction of boxes that the object will occupy in the pseudo-state space is thus n/N. The probability that part of the object will be mapped into a box occupied by another part of the object is $n \times n/N$. Using Eqs. (5.15) and (5.16) this probability goes as

$$n \times \frac{n}{N} \sim \frac{\varepsilon^{-2D_0}}{\varepsilon^{-M}} = \varepsilon^{M-2D_0} \tag{5.17}$$

This probability goes to zero with ε provided that $M > 2D_0$.

It should be noted that the box-counting dimension, D_0, is not the same as the correlation dimension, D_2. The correlation dimension is actually a lower bound on the box-counting dimension. While the Whitney Embedding theorem relates D_0 to the embedding dimension in practice, D_2 is more easily estimated from data.

5.7 Limits on Dimension Estimates

Soon after the methods outlined above were introduced there were several claims made for the discovery of low dimensionality in time series of weather and climate variables. None of these claims turned out to be valid. Most were flawed in the same way, having far too little data to be consistent with the dimensionality claimed. Smith showed that to determine the correlation dimension of an M-dimensional system to within a fractional accuracy of α requires at least N_{min} uncorrelated data points where

$$N_{min} \geq \left(\frac{R(1+\alpha)}{2\alpha} \right)^M \tag{5.18}$$

where R is the range of ε over which the gradient of $\ln C / \ln \varepsilon$ must be constant. Equation 5.18 shows that N_{min} increases exponentially with the dimension of the system [Smi88]. To estimate D_2 to within an accuracy of 5% N_{min} must be at least 42^M.

The maximum embedding dimension that can be determined using the method of false nearest neighbors can also be estimated. Consider a time series of N random numbers uniformly distributed on the interval $[0, 1]$. If the method of delays is used to construct a d dimensional pseudo-state space, then the approximate distance between a point in this space and its nearest neighbor, L_{min}, is given by

$$L_{min} \approx \frac{1}{2N^{1/d}} \tag{5.19}$$

The expectation value of $|y_{d+1} - x_{d+1}|$ if y_{d+1} and x_{d+1} are both random, uncorrelated and uniform on $[0, 1]$ is $1/3$. Thus the number of times that the first FNN criterion given by Eq. 5.9 is satisfied can be expected to fall close to zero when

$$\frac{1/3}{L_{min}} < \alpha_0 \tag{5.20}$$

Using Eq. 5.19 this condition becomes

$$\frac{2}{3} N^{1/d} < \alpha_0 \tag{5.21}$$

Thus for a given N the condition Eq. 5.9 fails to identify false nearest neighbors for

$$d \geq \frac{\log N}{\log(3\alpha_0/2)} \tag{5.22}$$

For $\alpha_0 \approx 15$ it is found that for the first FNN criterion to be meaningful, N must be at least about 20^d. In fact, for a "nearest neighbor" to be a meaningful concept, the distance between nearest neighbors must be smaller than some fraction, f, of the size of the attractor, that is

$$L_{min} < f R_A \tag{5.23}$$

Since for the case of a uniformly distributed random series $R_A \approx O(1)$, then

$$\frac{1}{2N^{1/d}} < f \tag{5.24}$$

thus

$$N > (2f)^{-d} \tag{5.25}$$

So again it is found that for a dimension estimate of d to be meaningful, the number of independent data points must exceed a^d where $a \approx O(10)$.

In an attempt to avoid spurious dimension estimates from time series, the method of surrogate data was introduced [TGL+91]. This method seeks to quantify the probability of a null hypothesis being able to explain the time series. One of the most common null hypotheses is *isospectral noise*. This is a random time series which has the same power spectrum as the time series being investigated. An isospectral surrogate can be constructed by fourier transforming the time series, picking new phases from a distribution uniform on $[0, 2\pi]$, and then performing an inverse fourier transform. An ensemble of such surrogates can be used to estimate whether the actual time series is significantly different from isospectral noise.

5.8 Results

The dimension estimating algorithms were applied to both the 1000 year time series of NINO 3 SST.

Figures 5.3 and 5.4 show the time lagged mutual information for both the test series. The first minimum in the lagged mutual information in both cases is 8 months. The peak at 12 months represents the seasonal cycle. The smaller, broader peaks around 30 and 40 months are due to the interannual variability. The lagged mutual information of the observed NINO 3 SST shows peaks at 6, 18, 24 and 36 months not present in either of the models. These peaks are the result of the seasonal cycle. Their absence in the models indicates that the seasonal cycle is relatively weaker in these models than in the observations.

Figure 5.5 shows the results of the correlation dimension estimates for the time series produced by the model without stochastic forcing along with similar estimates for its isospectral surrogate. The curves correspond to unfolding dimensions of 2 to 10. Saturation appears to be occurring for a correlation dimension of $D_2 \approx 1.5$. Note that no such saturation occurs for the isospectral surrogate time series.

Figure 5.6 shows the results of the correlation dimension estimates for the time series generated by the stochastically forced model. There is a hint that saturation might be occurring for $D_2 \approx 2.0$, but this is equivocal.

Figure 5.7 shows the results of the false nearest neighbors analysis performed on the time series produced by the noise free model. The FNN fraction falls close to zero for an embedding dimension of about 4. The fall off in FNN fraction is significantly faster than for any members of the ensemble of isospectral surrogates. The Whitney Embedding theorem implies that an embedding dimension of 4 is consistent with the correlation dimension estimate of 1.5. This result implies that in principle only 4 independent model variables are required to capture the dynamics of the model that produced the time series.

Figure 5.8 shows the results of the FNN analysis on the time series from the stochasically forced model. The FNN fraction falls faster than for the isospectral surrogates but it doesn't reach zero, even at an unfolding dimension of 10. Beyond an unfolding dimension of 6, the FNN fraction appears to level off at around 4%. This type of fall off to a residual level is seen in low dimensional systems contaminated by noise [Aba95].

All the methods employed to estimate the dimension suggest low dimensionality for the unforced coupled model. The number of independent data points used to make the estimates was $1000\,years/8\,months \approx 1500$. This number of data points puts the dimension estimates at the limit of credibility, but the comparison with surrogate data sets does bolster the case for low dimensionality.

The results in this chapter should be compared with previous efforts to estimate the dimensionality of ENSO models. The correlation dimension of the Cane-Zebiak model has been estimated to be $D_2 \approx 3.5$ based on a 1024 year time series of monthly averaged SST in the Eastern Pacific [TSCJ94]. Chang *et al.* used both the Grassberger-Procaccia algorithm and the method of false nearest neighbors to estimate the dimensionality of ENSO in an Intermediate Coupled Model, a Hybrid GCM (Ocean GCM with a steady-state atmosphere) and a full Ocean-Atmosphere GCM. They performed nonlinear time series analysis on the output of 1000 year runs of each of these models. They found evidence of low dimensionality for the ICM and HGCM ($D_2 \approx 2.5$, $M \approx 5$). No evidence for low dimensionality was found for the full GCM. This result implies that the atmosphere is the main source of high dimensionality in the coupled ocean-atmosphere system.

5.9 Extension to Real Data

All the results discussed above have been based on output generated by models. The obvious question is can these methods be applied to real time series that describe ENSO? The main obstacle to doing this has already been stated; these methods require copious

amounts of data. All the model studies were performed on time series at least 1000 years long and even then some of the results are on the threshold of credibility. The longest instrumental records describing ENSO are the sea level pressure measurements from Tahiti and Darwin which begin in 1882, far too short to make even a tentative detection of low dimensionality let alone a robust detection. There are historical records of ENSO based on indicators such as Nile flooding [Qui92], North American tree-ring records [LF85, Lou92] and ice cores [TMTA84, TMTB85]. However, these records are probably too qualitative or too displaced from the equatorial Pacific to be used for reconstructing a low dimensional "ENSO attractor," if such a thing exists. There is one source of paleoclimate data though that is ideally located (in the tropical Pacific), has high temporal resolution (monthly or even submonthly) and could provide records long enough to detect low dimensionality (millenia of data). This source is tropical corals. The abundance of the oxygen-18 isotope in corals is determined by the temperature of the water in which the coral grew as well as the local amount of precipitation [WW72, Dan64]. Temperature and precipitation reconstructions that overlap with the modern instrumental records indicate a high level of agreement with the observations [CSFM92]. Reconstructions predating the modern observations by several centuries have also been obtained [DWCG94] as well as reconstructions several decades long using 100,000 year old coral [KSJH99]. Although corals seem to offer the best hope for constructing time series long enough to put a reasonable lower bound on the dimensionality of ENSO, there are still several issues to be addressed. The biggest problem is probably the finite life of corals which typically live for only about 20 years. Continual time series longer than the lifetime of a single coral can be reconstructed by finding different corals that overlap, but constructing an unbroken coral record stretching over a millenium is a daunting task. However, it is possible that a continuous time series might not be absolutely necessary to apply the algorithms described in this chapter.

There exists a property, called *ergodicity*, that is possessed by some dynamical systems.

One of the most useful features of an ergodic system is the equivalence of time averages and ensemble averages. Usually this property is exploited by using long runs of a system as a proxy for an ensemble of shorter runs, but in this case the corollary may be useful. To demonstrate that strictly continuous time series are not necessary for the application of these algorithms, the 1000 year time series were cut into 16 segments each 50 years long and with random breaks between each segment ranging from zero up to 10 years. The method of delays was applied to each segment to construct a pseudo-state space and the false nearest neighbor algorithm was applied. The same segmentation and state space reconstruction procedure was applied to 10 isospectral surrogates of the original series.

5.10 Results for Discontinuous Time Series

Figures 5.9 and 5.10 show the result of the FNN algorithm for the discontinuous time series. Again the result for the irregular model indicates low dimensionality and that for the stochastic model suggests a low dimensional system contaminated by noise. In both cases the FNN fraction falls off faster than the ensembles although the significance level isn't as high as for the continual case. This difference can be explained by the smaller amount of data. In the disjoint case only 800 years of data is utilized. Furhermore, when using the method of delays, the number of independent pseudo-state space points that can be reconstructed from a time series of length τ with an unfolding dimension of d is $\tau/T - d + 1$. So if the data set consists of N time series of duration τ, the number of points lost is $N(d-1)$. So the more fragmented the time series, the fewer points can be reconstructed. Nevertheless, the FNN algorithm does not utilize any information concerning the temporal sequence of the points in pseudo-state space; this is why it can be applied to discontinuous data.

5.11 Discussion and Conclusions

The results presented above indicate that the ICM model can be represented by a low dimensional dynamical system. The correlation dimension estimate and the embedding dimension estimate both imply an embedding dimension of about 4. Interestingly, even when the model is forced by stochastic noise, it is distinguishable from colored noise. The stochastically driven model has an FNN behavior typical of a low dimensional system contaminated by noise.

It has also been demonstrated that estimates of the dimensionality of the system can be made using an ensemble of nonoverlapping, discontinuous time series. This can be done because although the method of delays does utilize the temporal order of data to reconstruct points in pseudo-state space, the dimension estimating algorithms do not make reference to how these points connect up in time. This fact may be of importance for the application of these methods to paleoclimate reconstructions where obtaining extended and unbroken time series can be very difficult.

While the methods described in this chapter indicate low dimensional dynamics for the forced model, they do not say much about how a low dimensional model can be constructed from the PDEs that describe the physics of the model given in Chapter 2. In the following chapters this question will be addressed using Empirical Orthogonal Functions.

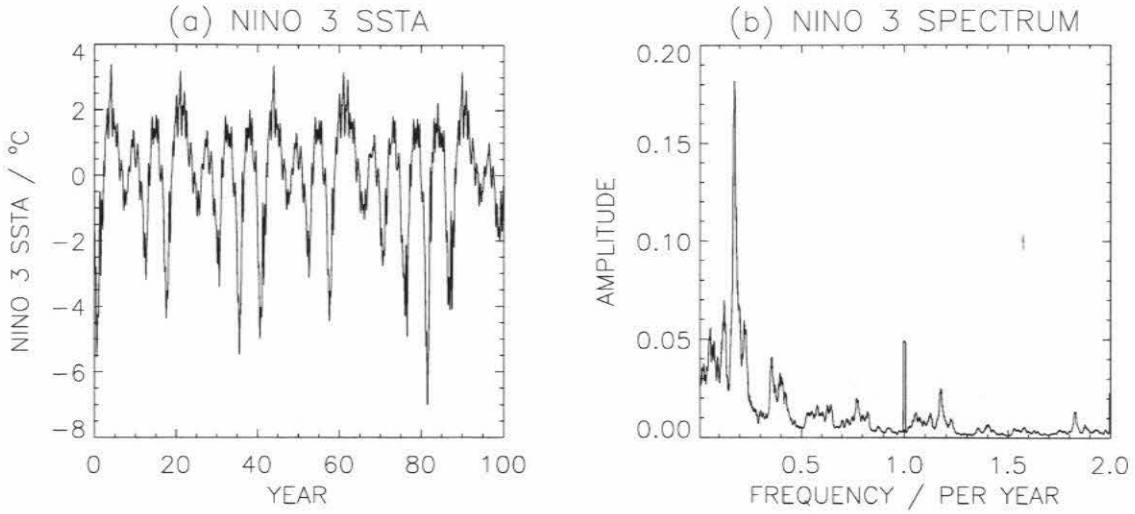


Figure 5.1: The behavior of NINO 3 SST in the noise free irregular regime.

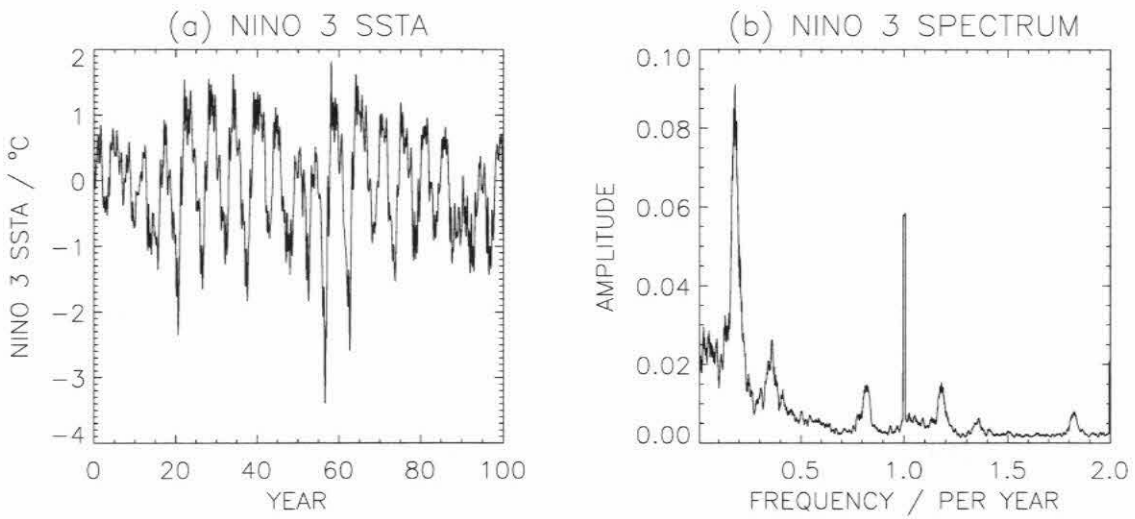


Figure 5.2: The behavior of NINO 3 SST in the stochastically forced model.

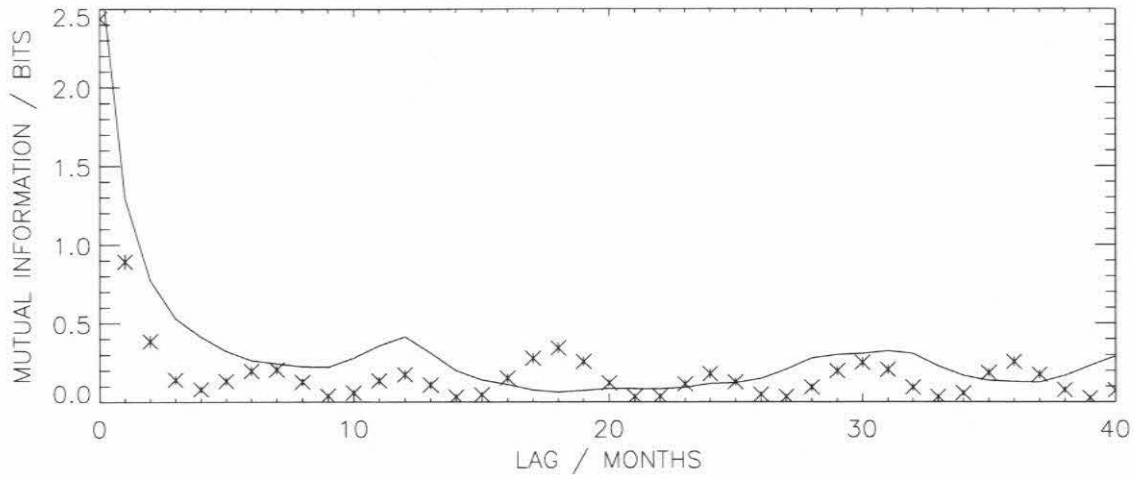


Figure 5.3: Plot of the lagged mutual information for the deterministic model (solid curve) compared with the mutual information of the observed NINO 3 SST series for 1950-1997.

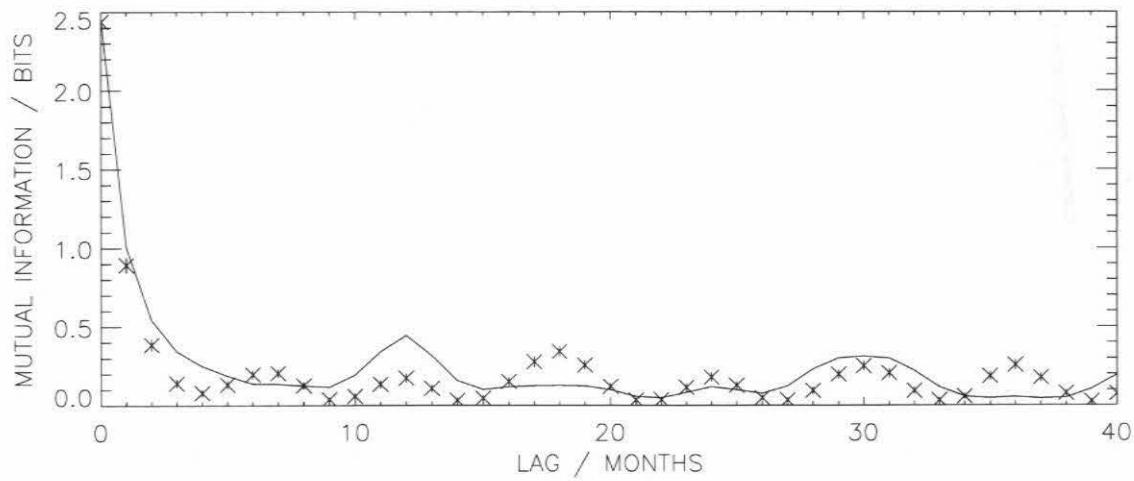


Figure 5.4: As Fig. 5.3 except for the stochastically forced model.

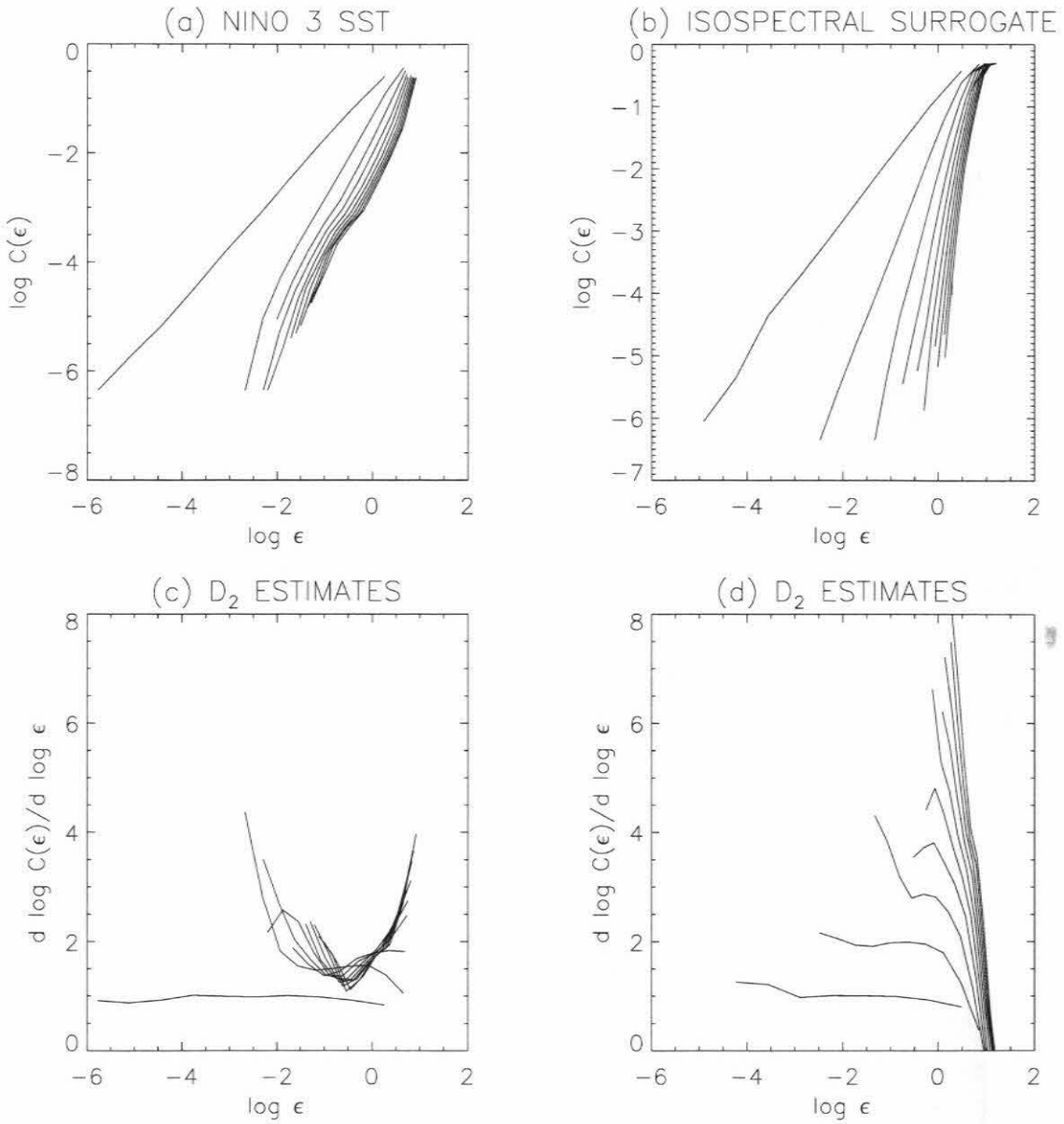


Figure 5.5: Correlation dimension estimate for the irregular regime without stochastic forcing. (a) Plot of $\log C(\varepsilon)$ v. $\log \varepsilon$ for the NINO 3 SST time series generated by the model, (b) as in (a) but for an isospectral surrogate of the model time series, (c) the gradient of the curves in (a), (d) the gradient of the curves in (b). The curves correspond to unfolding dimensions of 2 to 10.

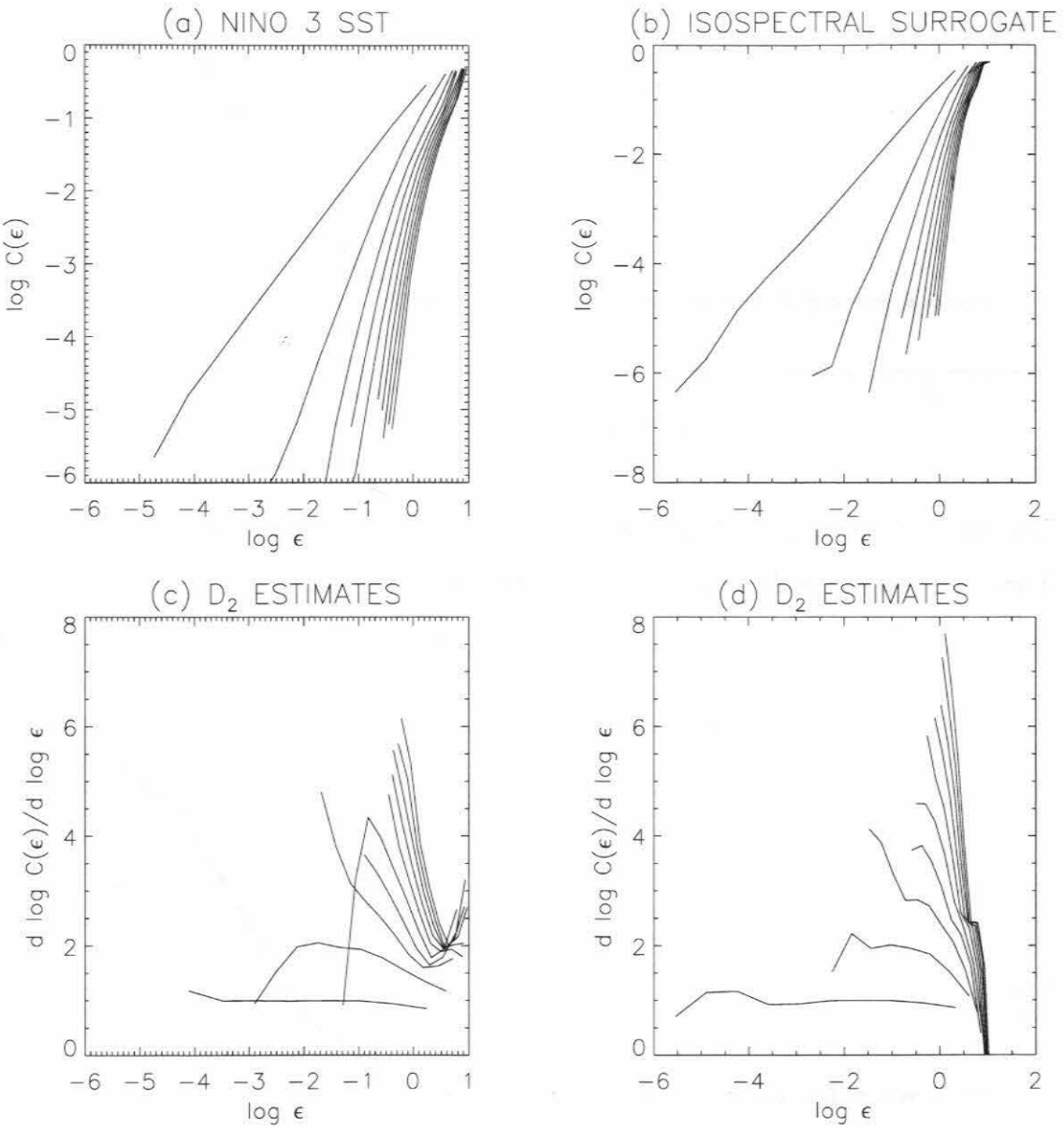


Figure 5.6: As Fig. 5.5 except for the stochastically forced model.

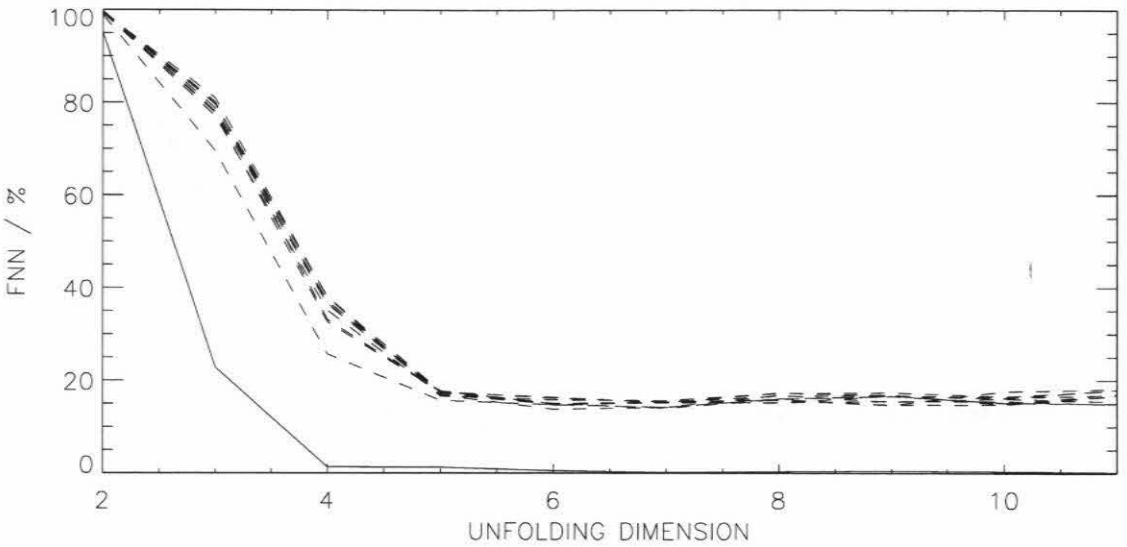


Figure 5.7: FNN estimate for the irregular regime without stochastic forcing. The solid line is for the NINO 3 SST time series generated by the model. The dashed lines are for an ensemble of 10 isospectral surrogates.

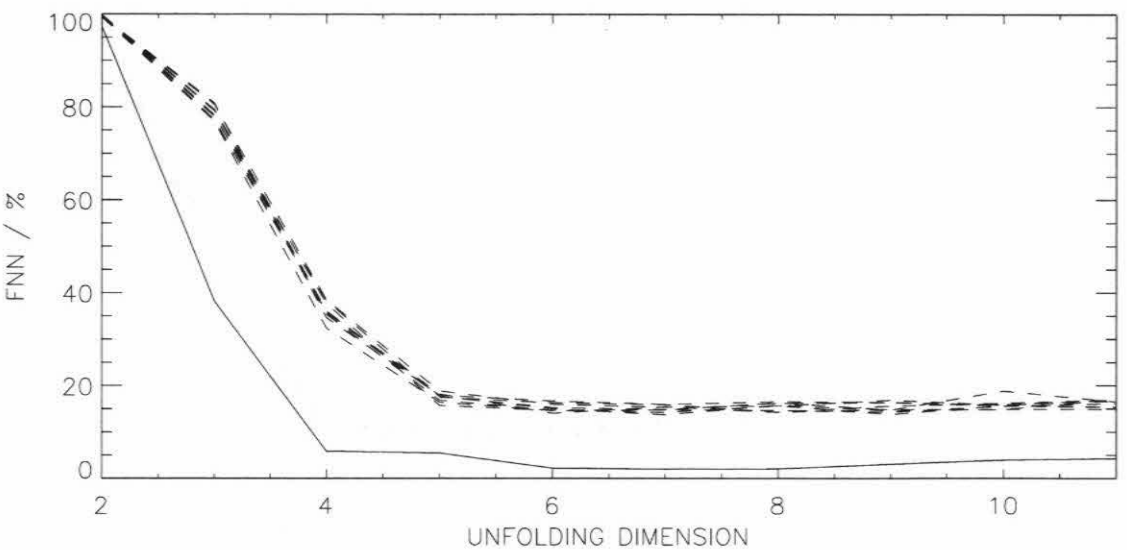


Figure 5.8: As for Fig. 5.7 except for the stochastically forced model.

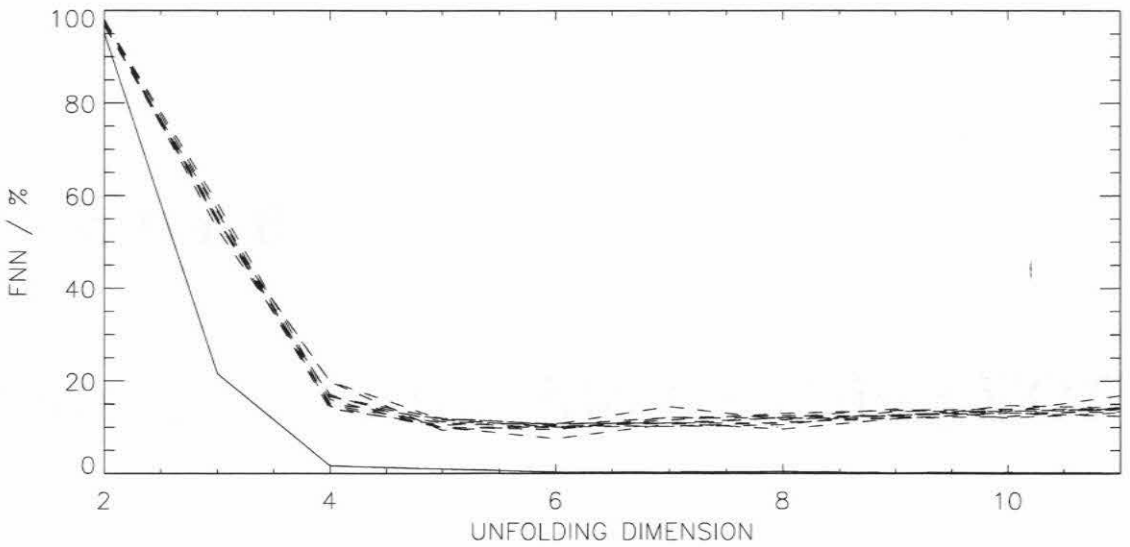


Figure 5.9: FNN estimate for the irregular regime without stochastic forcing determined from the dataset consisting of disjoint 50 year segments. The dashed lines are for an ensemble of 10 isospectral surrogates.

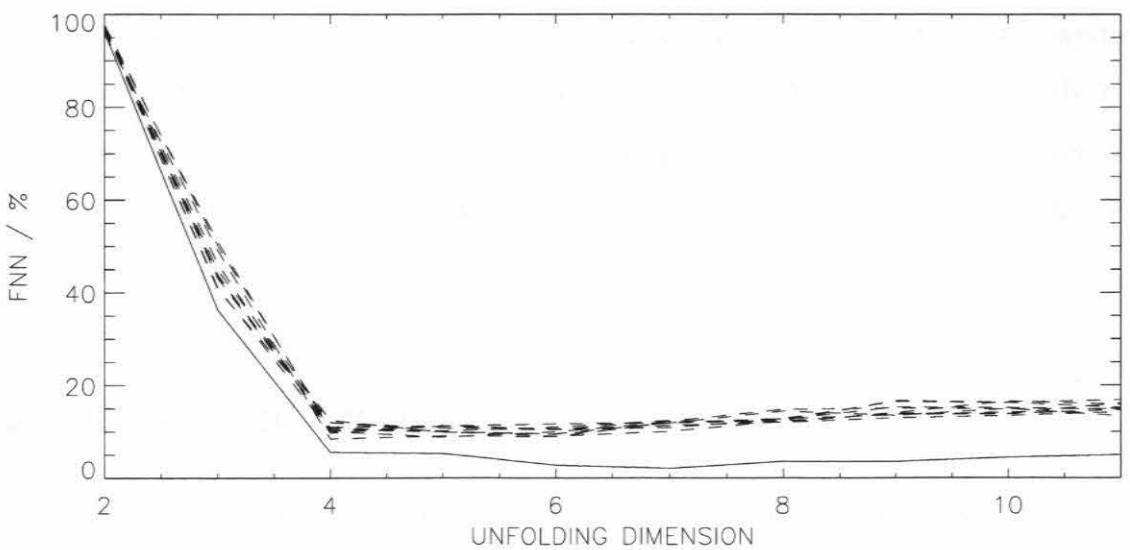


Figure 5.10: As for Fig. 5.9 except for the stochastically forced model.

Chapter 6

Reduction of the Model Using EOFs

6.1 Introduction

The system in question, the intermediate coupled model of the tropical Pacific ocean-atmosphere system, has *a priori* a 238 dimensional state space in the configuration which was used for all the numerical experiments described in Chapter 3. The results of the dimension estimates in Chapter 5 all suggest that the dimensionality of the system's attractor is much lower. In fact, these results suggest that the model is actually a low dimensional dynamical system. [1] A common approach to constructing a low dimensional dynamical model from a model with a high dimensional state space is to project the model onto a truncated set of EOFs.

6.2 Empirical Orthogonal Functions

The method of Empirical Orthogonal Function analysis is also known as Proper Orthogonal Decomposition (POD) or Karhunen Loẽve Decomposition after its developers

[1] "Low" in this sense usually implies $< O(10)$.

[Le45, Kar46]. It was first used in meteorology by Lorenz in 1956 [Lor56]. A decade later it was introduced into the field of turbulence modeling [Lum67]. Initially EOF analysis was used as an investigative tool to try to extract the salient features of the data set under investigation. More recently researchers have attempted to use EOFs as a basis set for Galerkin projections of models. This method was first used in a meteorological context for a vertical model [Ruk63], but later it was applied to horizontal barotropic models of the atmosphere [RK75, Sel93, Sel95]. In addition the method has been used to model convection in the atmospheric boundary layer [Zhu96]. The method has more recently been applied to a horizontal baroclinic model of the atmosphere [Sel97a].

Let $\mathbf{x}_t = [x_{1t}, x_{2t}, ..., x_{mt}]$ be a state vector with m components which describe the state of the system at a time t. Let $\langle \mathbf{x} \rangle$ be the mean state vector and let $\mathbf{x}'_t$ be the deviation of the state vector from the mean at time t. The problem is to find a new basis which is optimal in some sense. This can be done by maximizing the average projection of the data onto the new basis set. Let $\mathbf{e}_i$ $(i = 1, ..., m)$ be unit vectors defining the new basis. The aim is to maximize $\langle (\mathbf{e}_i^T \mathbf{x}')(\mathbf{e}_i^T \mathbf{x}')^T \rangle$ where $\langle . \rangle$ denotes averaging over the time index. The constraint is that $\mathbf{e}_i^T \mathbf{e}_i = 1$. Thus the function to be maximized is

$$J = \langle (\mathbf{e}_i^T \mathbf{x}')(\mathbf{e}_i^T \mathbf{x}')^T \rangle - \lambda_i (\mathbf{e}_i^T \mathbf{e}_i - 1) \tag{6.1}$$

where λ_i is the undetermined multiplier. Differentiation of Eq. 6.1 with respect to $\mathbf{e}_i$ leads to

$$\langle \mathbf{x}' \mathbf{x}'^T \rangle \mathbf{e}_i = \lambda \mathbf{e}_i \tag{6.2}$$

Thus the new basis vectors are the eigenvectors of the matrix $\langle \mathbf{x}' \mathbf{x}'^T \rangle$. Furthermore,

$$\langle (\mathbf{e}_i^T \mathbf{x}')(\mathbf{e}_i^T \mathbf{x}')^T \rangle = \lambda_i \tag{6.3}$$

So the data has the largest average projection onto the eigenvector associated with the largest eigenvalue of $\langle \mathbf{x}' \mathbf{x}'^T \rangle$. Note that $\langle \mathbf{x}' \mathbf{x}'^T \rangle$ is a real symmetric matrix and thus it has orthogonal eigenvectors which can be normalized to form a complete orthonormal basis.

Since the mean state vector has been subtracted from the data, λ_i is the variance associated with the ith basis vector. The essential point about this basis set is that for any value of D more variance is contained in the subspace spanned by the leading D basis vectors than any other D-dimensional subspace. These basis vectors are the empirical orthogonal functions of the data set.

The covariance matrix calculated from a finite data set will contain sampling errors and these errors will propagate to the eigenvalue estimates of the matrix. North *et al.* [NBT+82] estimated the sampling error in the ith eigenvalue to first order as

$$\delta\lambda_i \approx \lambda_i \sqrt{\frac{2}{N}} \tag{6.4}$$

where N is the number of realizations of the data vector. For the estimate of an eigenvalue and its associated eigenvector to be considered robust, the error in the eigenvalue should be significantly smaller than the separation between the eigenvalue and the eigenvalue closest to it. If $\Delta\lambda_i$ is the difference between λ_i and the eigenvalue that lies closest to λ_i, then the ratio $\Delta\lambda_i/\delta\lambda_i$ provides a rough measure of the robustness of the eigenvalue estimate. A value greater than unity indicates the eigenvalue is separated from other eigenvalues by more than the estimated sampling error a value less than unity means that the eigenvalue is not well separated. When eigenvalues are not well separated, eigenvector mixing can occur leading to a large amount of variablity in the EOFs calculated using different samples.

To apply EOF analysis to the output of the coupled model, state vectors were constructed. The state vectors used were the nondimensionalized state vectors given by Eq. 2.38.

6.3 Results of the EOF Analysis

EOF analysis was performed on the output of five runs of the coupled model. The results are shown in Figs. 6.1 to 6.10.

Figures 6.1 and 6.2 show the results from the case without a seasonal cycle or stochastic wind noise for a coupling strength that puts the model just above the bifurcation where it displays sustained interannual oscillations. The leading 2 EOFs contain 95% of the variance and the leading 4 EOFs contain 99% of the variance. Figure 6.1b indicates that the eigenvalues of the leading 32 EOFs are well separated.

Figures 6.3 and 6.4 show what happens when the stochastic wind forcing with a white spectrum is added. There is clearly more variance contained in the higher EOFs than in the case without stochastic forcing. The leading 4 EOFs only contain 62% of the variance and 42 EOFs are required to explain 99% of the variance. The spatial patterns of the first 2 EOFs in Fig. 6.2 are very similar to the leading 2 EOFs in Fig. 6.4. There is less similarity between the higher EOFs. This is not that suprising since in Fig. 6.1 the leading 2 EOFs do contain 95% of the variance. The stochastic forcing contributes to most of the variance in the higher EOFs in Fig. 6.3. Due to the slower falloff in the eigenvalue spectrum, only the leading 16 eigenvalues are well separated.

Figures 6.5 and 6.6 show the results for a lower value of coupling for which the model is below the bifurcation and does not display self-sustaining oscillations. In this case the stochastic noise is required to excite the interannual oscillations. The eigenvalue spectrum looks very similar to the noise forced case above the bifurcation and again over 40 EOFs are required to capture 99% of the variance. Figure 6.5b indicates that only the leading 8 eigenvalues are well separated. A comparison of the leading 8 EOFs in Fig. 6.6 with those in Fig. 6.4 shows that the leading 8 EOFs are very similar in both cases. (When making such comparisons it should be remembered that the signs of the EOF patterns are arbitrary.) The implication is that in these stochastically forced cases, which EOFs

contain the most variance does not depend greatly on whether the model is above or below the bifurcation where self-sustaining oscillations are possible. In the supercritical case the leading 2 EOFs contain almost all the variance associated with the self-sustaining oscillation the other EOFs are capturing the behavior of stable, noise-excited oscillators. In the subcritical case all the variance is due to stable, noise-excited oscillators but it appears that the distinction makes little difference to the EOFs of the noise forced cases.

Figures 6.7 and 6.8 show the results of the EOF analysis performed on the output of the coupled model with a seasonal cycle added, and the coupling strength puts the model in the *regular* regime. The leading 2 EOFs contain 91% of the variance, the leading 4 EOFs contain 96% of the variance and 99% can be explained with the first 8 EOFs. The plot of eigenvalue separation in Fig. 6.7b shows that the separation of about the leading 25 eigenvalues is larger than the error in these eigenvalues. Figures 6.9 and 6.10 show the results when the coupling is increased to put the model in the *irregular* regime. In this case the leading 4 EOFs only contain 90% of the variance and 16 EOFs are needed to capture 99% of the variability. In this case only about the first 12 eigenvalues are well separated.

The time series of the coefficients associated with the leading EOFs are shown in Figs. 6.11 to 6.16.

In Fig. 6.11 the leading 4 EOFs of the case with no seasonal cycle and no stochastic forcing is shown. The spectrum for this case is shown in Fig. 3.7. The dominant 4.5 year oscillation is captured by EOF 1 and EOF 2. The secondary oscillation, with a period of about 2.1 years, is contained in EOF 3 and EOF 4. An inspection of the spatial patterns of the h-field of the leading 4 EOFs in Fig. 6.2 shows that the 4.5 year oscillation is meridionally symmetric whereas the secondary 2.1 year oscillation is asymmetric in the north-south direction. The distinction between the SST fields of EOFs 1 and 2 and EOFs 3 and 4 is less clear. Jiang *et al.* found that spatially the SST variability associated with the QQ mode is similar to that of the QB mode [JNG95] and they did not have

subsurface observations of h. This work suggests that if a similar analysis was performed on the thermocline depth field, it would be found that QQ and QB variability in the thermocline depth are associated with meridionally symmetric and asymmetric patterns respectively. The patterns of EOF 1 and EOF 2 are similar to the characteristic pattern of ENSO variability with the thermocline sloping from east to west and SST varability confined to the eastern basin.

In Fig. 6.12 the time series are for the leading 4 EOFs of the case when stochastic wind forcing with a white spectrum is added. The leading 2 EOFs still contain the interannual oscillation with a period around 4.5 years although the time series are now affected by the noise. However, the time series of EOF 3 and EOF 4 do not bear any resemblence to the corresponding time series in Fig. 6.11. The dominant frequency of these two time series is subannual. The difference isn't suprising since the spatial forms of EOF 3 and EOF 4 in the two cases are quite different (Fig. 6.2 and Fig. 6.4). Projecting the output of the noise forced model onto the leading EOFs obtained from the model without noise forcing gives the result shown in Fig. 6.13. In this case EOF 3 and EOF 4 still contain an oscillation with a period of around 2.1 years, although like the 4.5 year oscillation it is affected by noise. Figure 6.14 shows the output of the model below the bifurcation driven with stochastic wind forcing onto the EOFs of the noise-free model above the bifurcation (Fig 6.2). The leading 2 EOFs still contain an oscillation with a period of approximately 4.5 years and EOF 3 and EOF 4 still contain the oscillation with a period of 2.1 years. In this case both of these oscillations must be noise excited because the model is incapable of self-sustaining interannual oscillations below the bifurcation.

The results of the EOF analyses of the noise driven models suggest that the main effect that stochastic forcing has on SST variability is by increasing the variance of higher EOFs rather than by increasing the variance of the leading EOFs that contain the self-sustaining oscillation. This question will be further investigated in Chapter 7.

Figure 6.15 shows the time series of the coefficients of the leading 4 EOFs for the

case with seasonal forcing but no stochastic forcing. Notice that EOF 1 and EOF 2 are dominated by the annual cycle while EOF 3 and EOF 4 are dominated by the interannual oscillation with a period of about 5 years. Inspection of the spatial patterns of these EOFs in Fig. 6.8 shows that EOF 3 and EOF 4 do resemble that characteristic pattern of ENSO variability, with zonally asymmetric thermocline and SST variability in the east. Note also that EOF 1 and EOF 2 also have similarities with this pattern of ENSO variability. This highlights the fact that ENSO variability and seasonal variability are not orthogonal. This can also be seen in the time series. The coefficients of EOF 1 and EOF 2 do contain interannual variability and EOF 3 and EOF 4 do contain seasonal variability.

Figure 6.16 shows the time series of the seasonal case with stronger coupling that puts the model in the irregular regime. As in Fig. 6.15 EOF 1 and EOF 2 are dominated by the annual cycle while EOF 3 and EOF 4 contain most of the interannual variability, although in this case this interannual variability is irregular.

6.4 Comparison with Observations

The amount of variance captured by the leading EOFs can be compared with the result of performing EOF analyses on the observed SST fields for the period 1950-1991 [PS95, Pen96]. It was found that the leading 15 EOFs contained 65% of the SST variance and the leading 20 EOFs contained 75% of the variance. The amount of variance contained in the leading EOFs of the SST components of the model state vectors is comparable to that contained in the leading EOFs of the full state vectors. Thus it appears that in reality there is more variance associated with the higher EOFs than in any of the model runs, including those with stochastic forcing. This might be explained by the exclusion of heat flux noise which is more difficult to estimate than the windstress forcing used [KM97].

A study using Multivariate Singular Systems Analysis (M-SSA) which identified the quasi-quadrennial (QQ) and quasi-biennial (QB) oscillations in observed SST found that

the QQ oscillation accounted for 39% of the variance and the QB oscillation for 19% of the variance [JNG95]. The variance contained in the QQ-like oscillation of EOF 1 and EOF 2 in Fig. 6.13 is 50% while the QB-like oscillation of EOF 3 and EOF 4 accounts for a further 7% of the variance. Thus if EOF 1 and EOF 2 are identified with the QQ oscillation and EOF 3 and EOF 4 with the QB oscillation, then it appears that the QQ:QB variance ratio of the model is higher than in reality.

6.5 Projection and Closure

Reduced versions of the full model were constructed by Galerkin projection of the model onto truncated sets of the EOFs which span a lower dimensional subspace of the statespace of the full model.

Let $\mathbf{P}$ be the matrix of vectors defining the EOFs. That is $[P_{i1}, P_{i2}, ...P_{iN}]$ is the ith EOF, where N is the dimension of the model state space. The nondimensionalized model is given by Eq. 2.39 which is repeated here.

$$\frac{d\mathbf{x}}{dt} = \mathbf{M}\mathbf{x} + \mathbf{N}(\mathbf{x}) + \mathbf{F}(t) \tag{6.5}$$

If $\mathbf{y}$ is the reprojected state vector, $\mathbf{y} = \mathbf{P}\mathbf{x}$, then the reprojected model is given by

$$\frac{d\mathbf{y}}{dt} = \mathbf{P}\mathbf{M}\mathbf{P}^T\mathbf{y} + \mathbf{P}\mathbf{N}(\mathbf{P}^T\mathbf{y}) + \mathbf{P}\mathbf{F}(t) \tag{6.6}$$

The model given by Eq. 6.6 is exactly equivalent to the model given by Eq. 6.5 since the basis set has not been truncated. Let the row vectors of $\mathbf{P}$ be ordered in decreasing order of their corresponding eigenvalues. Let N_r be the number of EOFs that are to be retained and let $N_d = N - N_r$ be the number to be discarded. The projection matrix can now be decomposed as follows.

$$\mathbf{P} = \begin{bmatrix} \mathbf{P_r} \\ \mathbf{P_d} \end{bmatrix} \tag{6.7}$$

where $\mathbf{P}_r$ is a $N_r \times N$ matrix containing the retained basis vectors and $\mathbf{P}_d$ is a $N_d \times N$ matrix containing the discarded basis vectors. The reprojected state vector, $\mathbf{y}$, can also be decomposed.

$$\mathbf{y} = [\mathbf{y}_r, \mathbf{y}_d] \tag{6.8}$$

where $\mathbf{y}_r$ is the vector of the retained components and $\mathbf{y}_d$ is the vector of the discarded components. The truncated, reprojected model is thus given by

$$\frac{d\mathbf{y}_r}{dt} = \mathbf{P}_r\mathbf{M}\mathbf{P}_r^T\mathbf{y}_r + \mathbf{P}_r\mathbf{N}(\mathbf{P}_r^T\mathbf{y}_r) + \mathbf{P}_r\mathbf{F}(t) \tag{6.9}$$

However, the model given by Eq. 6.9 can be improved by attemping to model the effects of the discarded EOFs on the dynamics of the retained EOFs. The model used is known as the *closure* scheme. Most attempts at constructing closure schemes for truncated models have been statistical-empirical. In most standard atmospheric and oceanic models, the effect of unresolved modes is parameterized by introducing an eddy diffusivity. Essentially the same approach has been used in low order EOF models in which extra linear dissipation is added to prevent model drift [Sel97a]. However, the effects of discarded modes isn't always dissipative. Selten attempted to model the effects of discarded modes by finding a linear combination and quadratic combinations of the retained coefficients that minimized the error in the tendency equations (time derivatives) of the reduced model. This approach was found to improve short range forecasts but led to the model being unstable over longer integration periods [Sel97b]. Later attempts at this kind of closure attempted to solve the problem of instability by imposing constraints on the closure scheme to bound the total energy of the model [AS97].

In this work a simple closure scheme was adopted. The coefficients of each of the discarded EOFs were specified as constants and equal to their mean value in the full model. Let $\bar{\mathbf{y}}_d$ be the vector of the mean coefficients of the discarded EOFs. The projected

model thus becomes

$$\frac{d\mathbf{y}_r}{dt} = \mathbf{P}_r\mathbf{M}\mathbf{P}_r^T\mathbf{y}_r + \mathbf{P}_r\mathbf{M}\mathbf{P}_d^T\bar{\mathbf{y}}_d + \mathbf{P}_r\mathbf{N}(\mathbf{P}_r^T\mathbf{y}_r + \mathbf{P}_d^T\bar{\mathbf{y}}_d) + \mathbf{P}_r\mathbf{F}(t) \tag{6.10}$$

where the terms involving $\bar{\mathbf{y}}_d$ are the closure terms.

The behavior of the truncated models is investigated in Chapter 7.

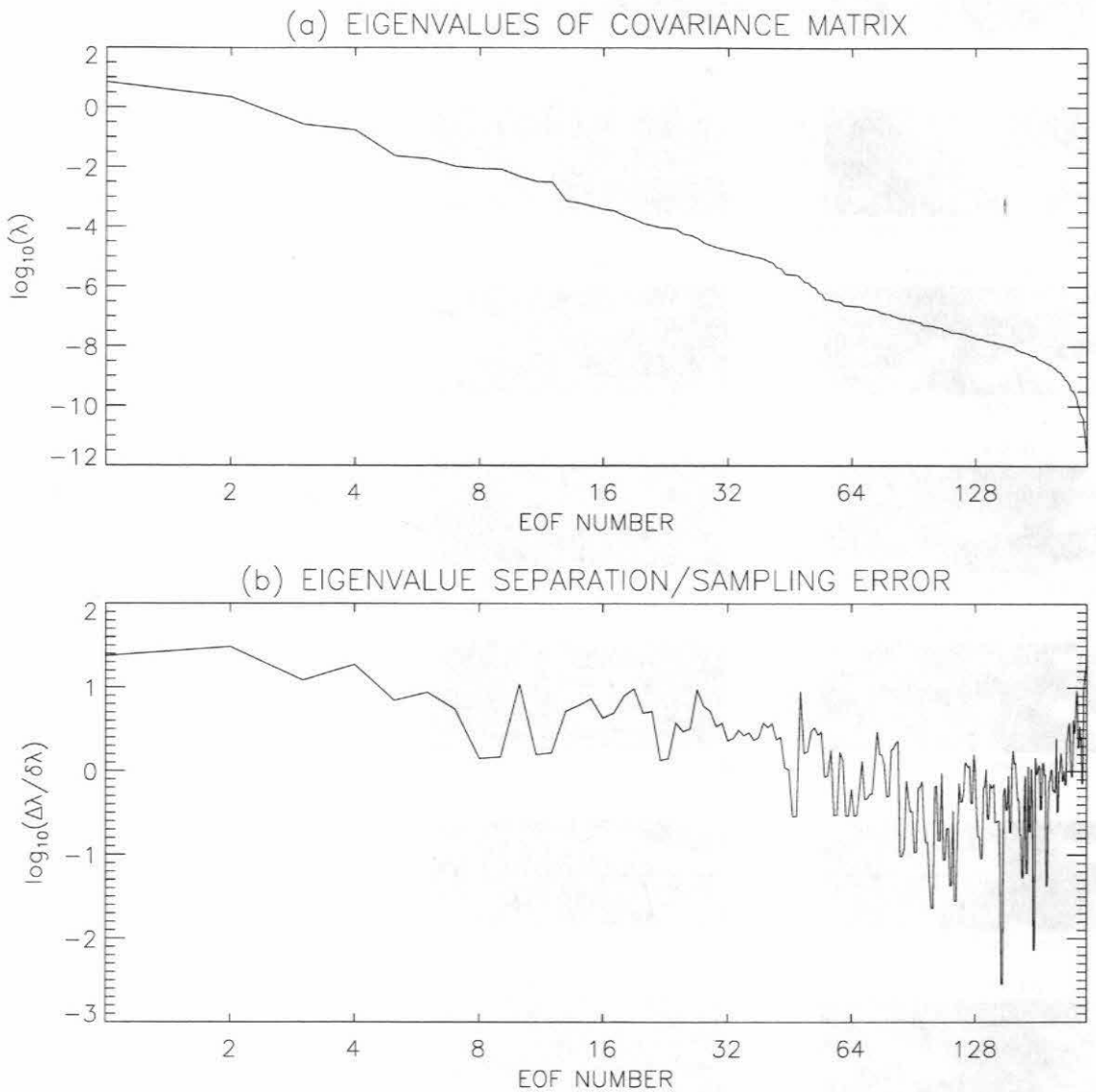


Figure 6.1: Eigenvalues of the covariance matrix of the coupled run with $K_Q W = 0.0097\, m^3\, s^{-4}\, ^\circ C^{-1}$ no seasonal cycle and no stochastic forcing. The model was above the bifurcation.

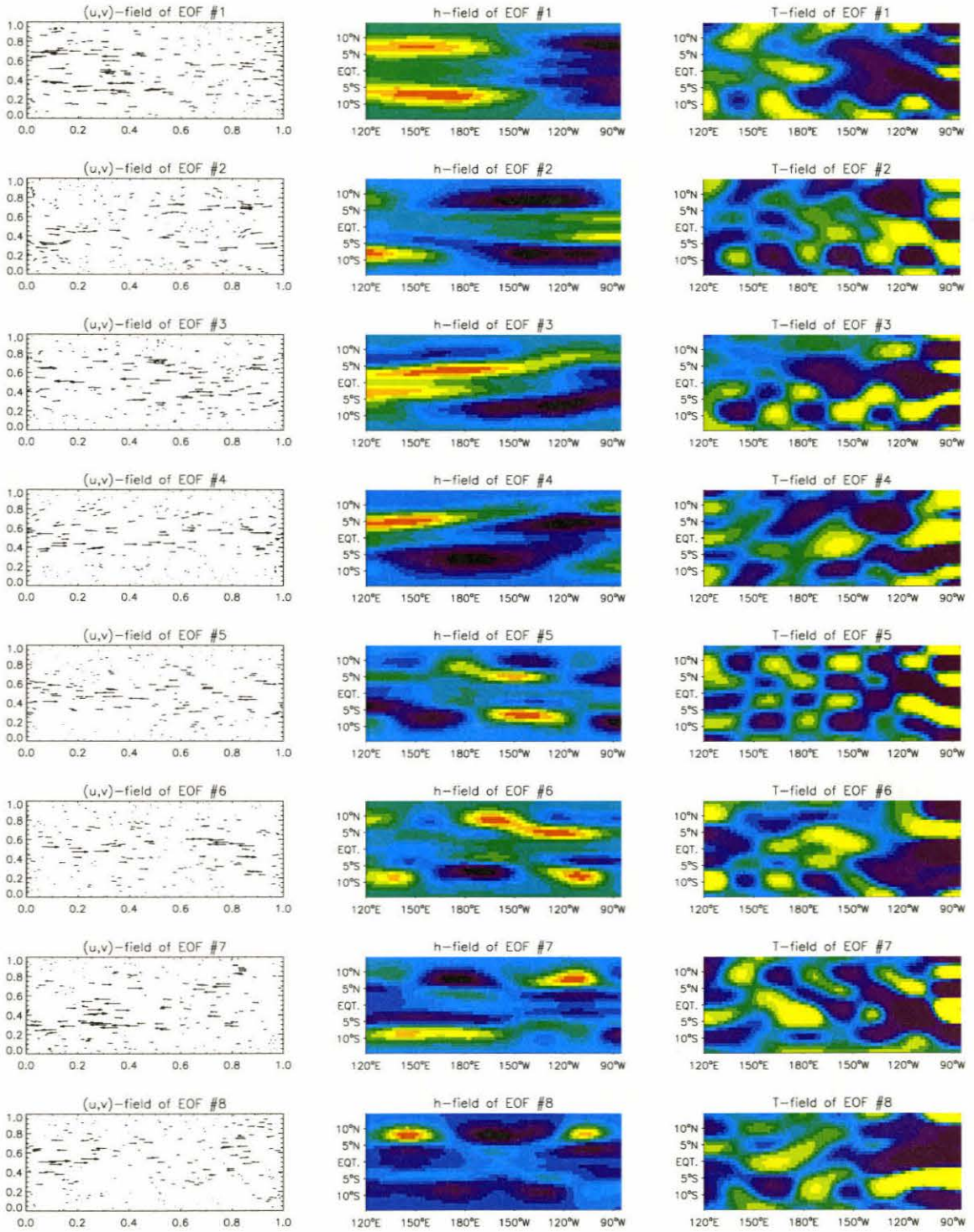


Figure 6.2: The leading EOF patterns associated with the eigenvalues in Fig. 6.1.

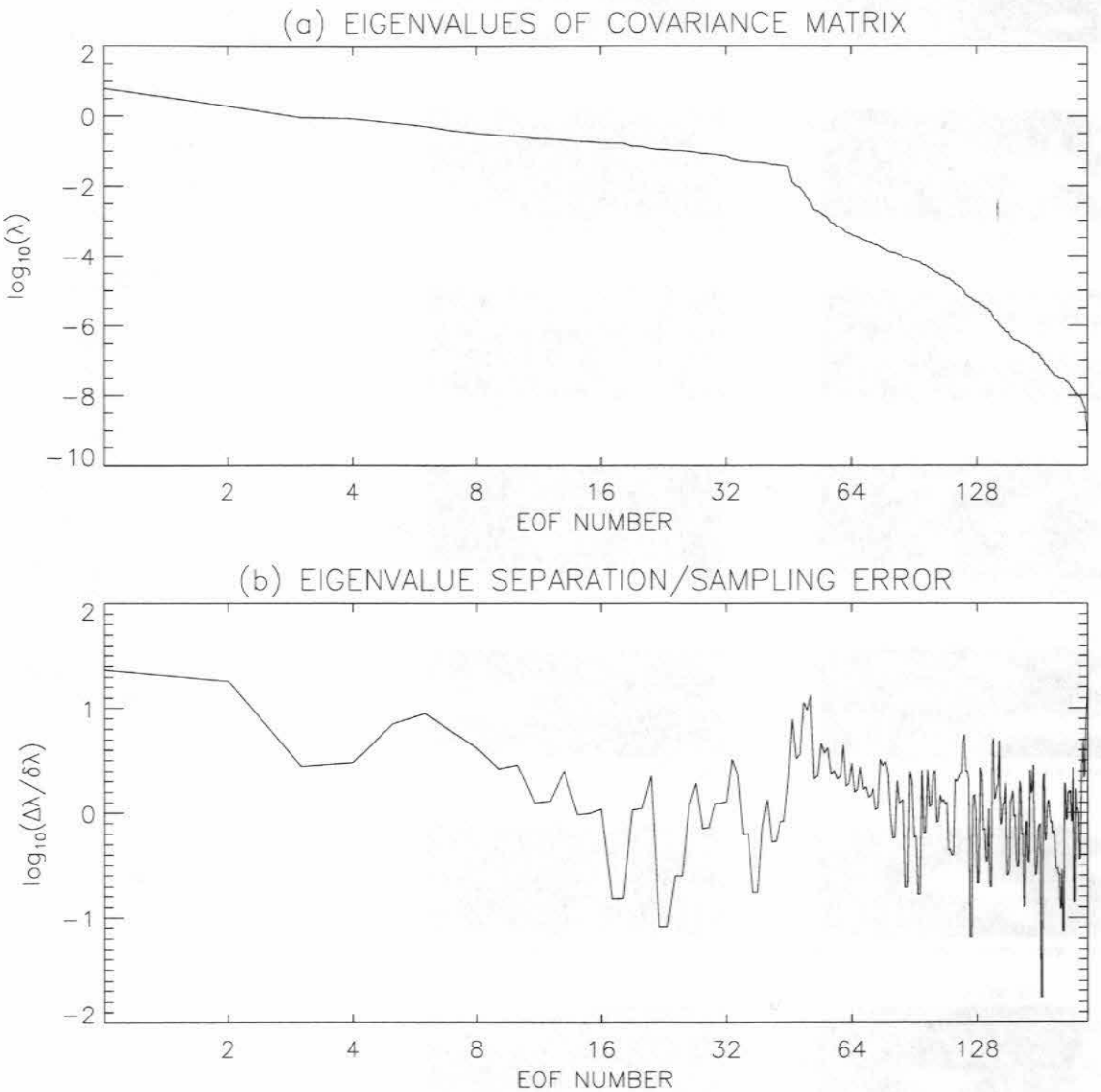


Figure 6.3: Eigenvalues of the covariance matrix of the coupled run with $K_Q W = 0.0097\, m^3\, s^{-4}\, {}^\circ C^{-1}$ with no seasonal cycle but with stochastic wind forcing with a white spectrum. The model was above the bifurcation.

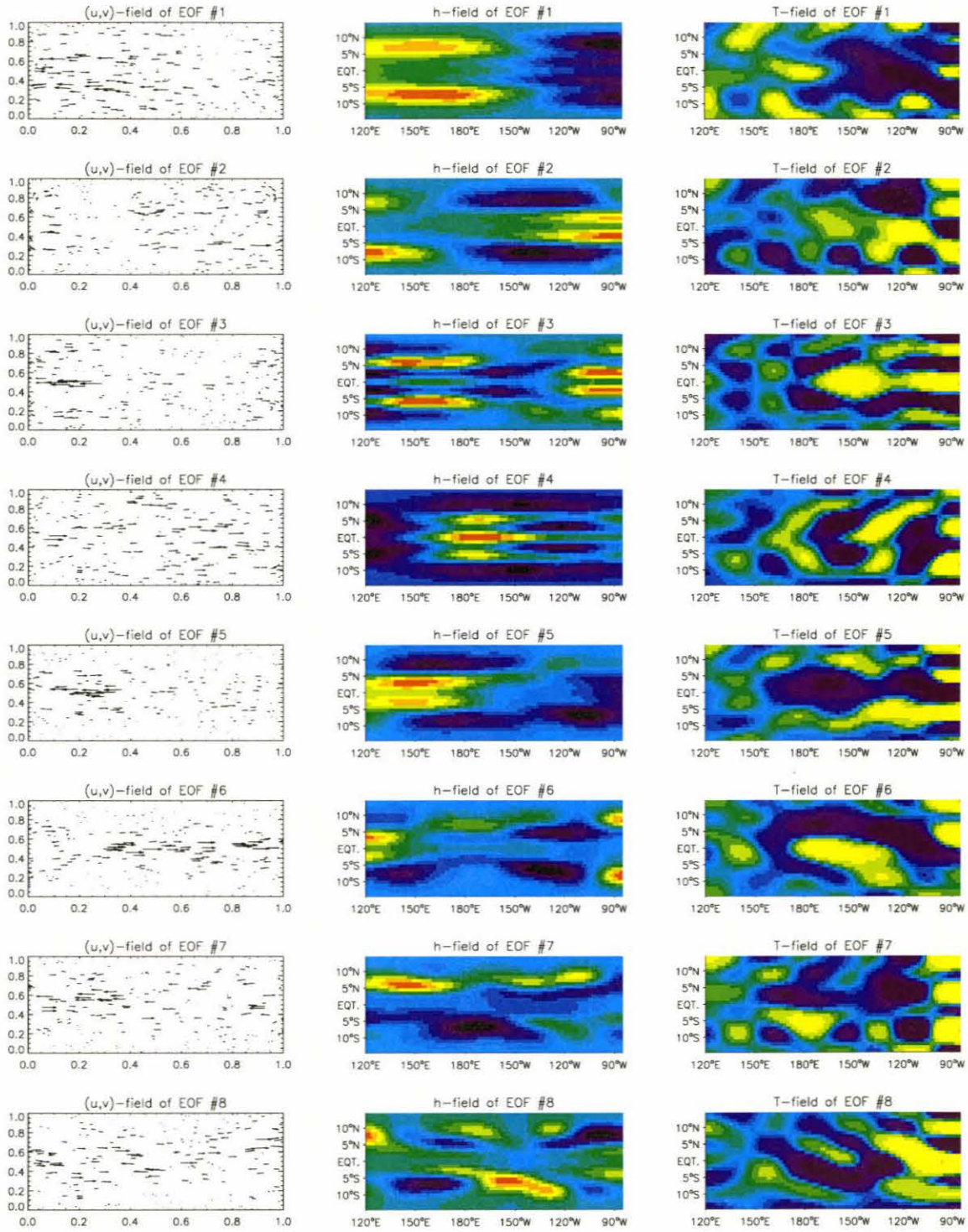


Figure 6.4: The leading EOF patterns associated with the eigenvalues in Fig. 6.3.

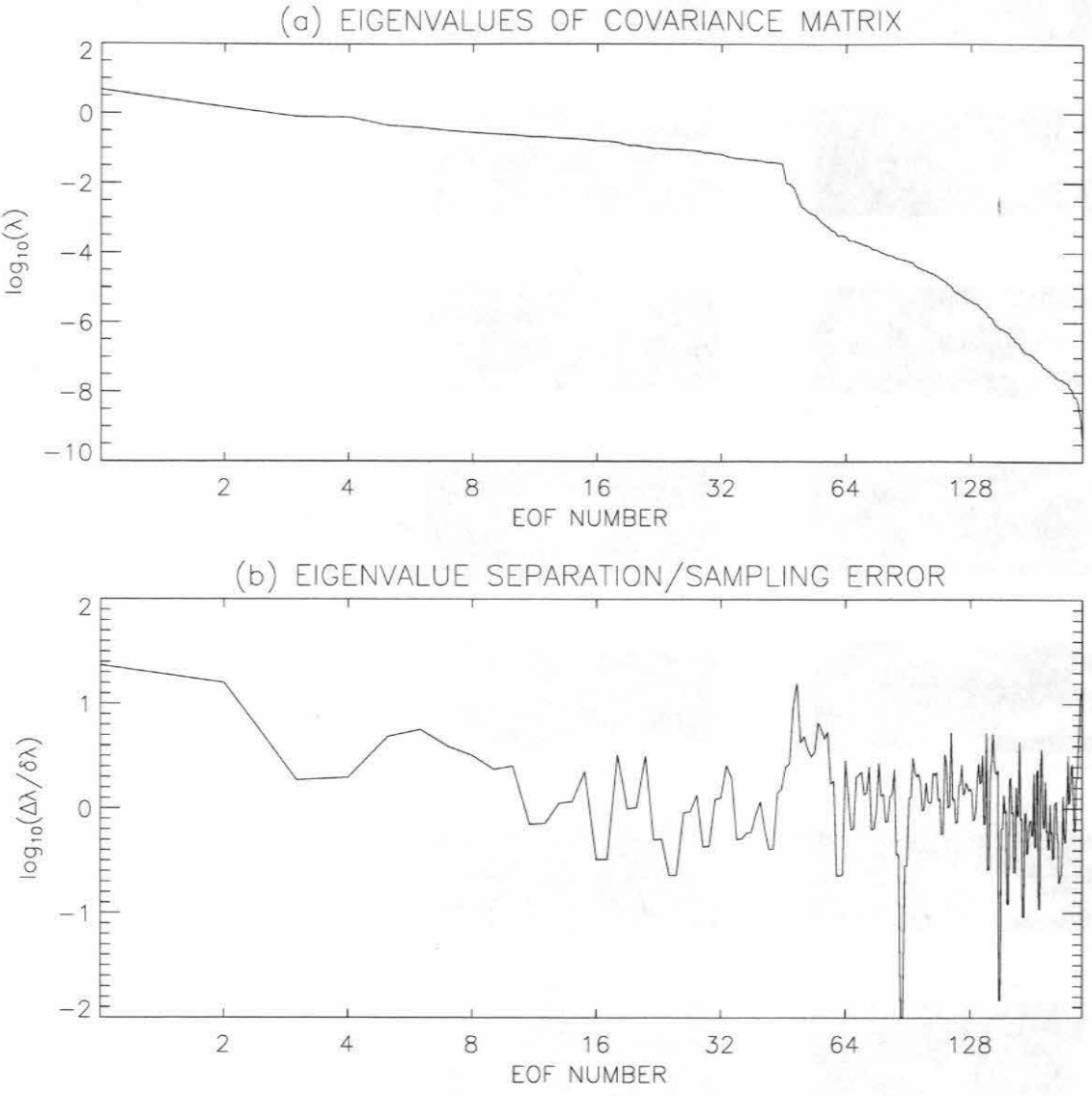


Figure 6.5: Eigenvalues of the covariance matrix of the coupled run with $K_Q W = 0.0090\, m^3\, s^{-4}\, {}^\circ C^{-1}$ with no seasonal cycle but with stochastic wind forcing with a white spectrum. The model was below the bifurcation.

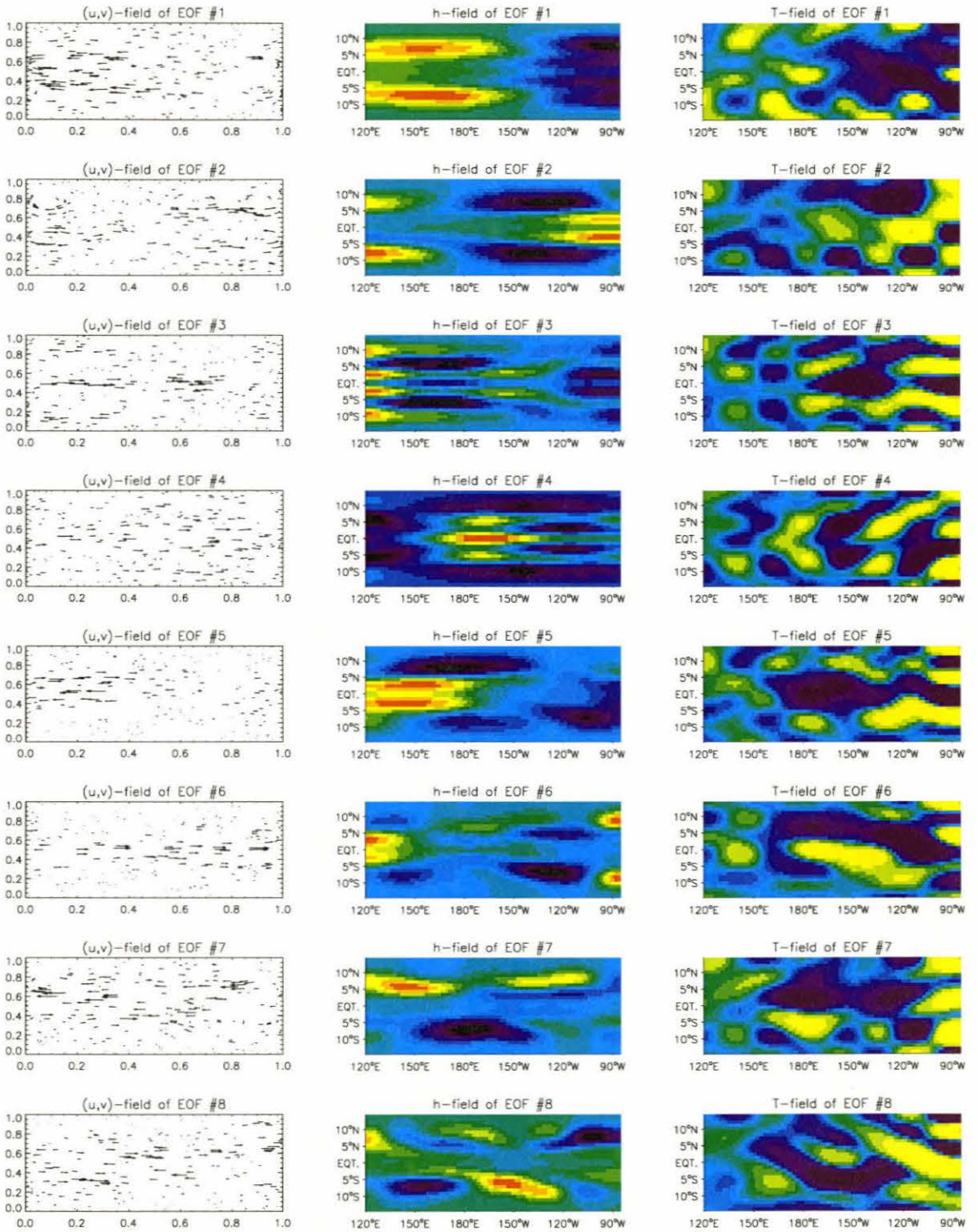


Figure 6.6: The leading EOF patterns associated with eigenvalues in Fig. 6.5.

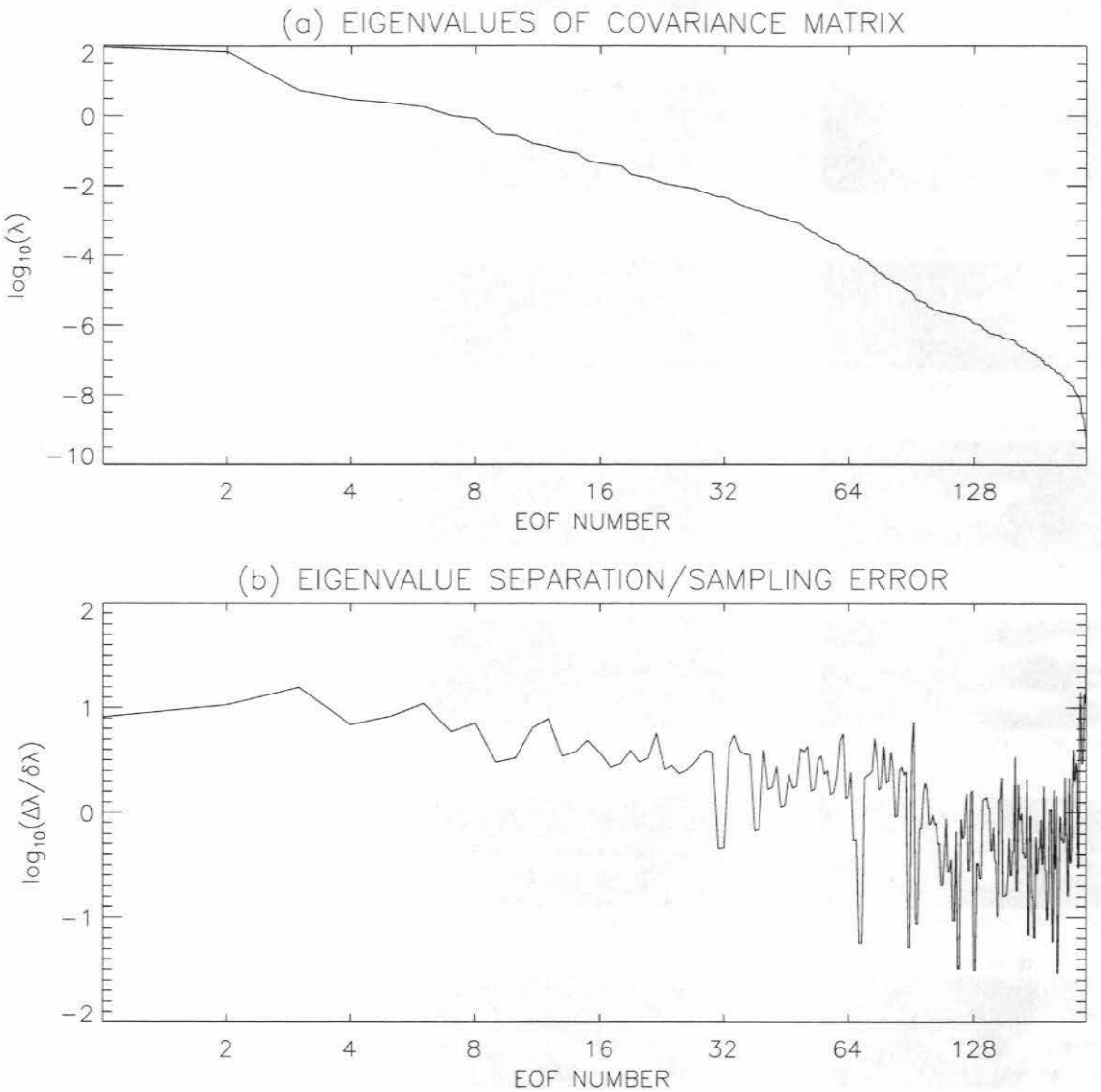


Figure 6.7: Eigenvalues of the covariance matrix of the coupled run with $K_Q W = 0.0136\, m^3\, s^{-4}\, {}^\circ C^{-1}$ with a seasonal cycle but no stochastic forcing.

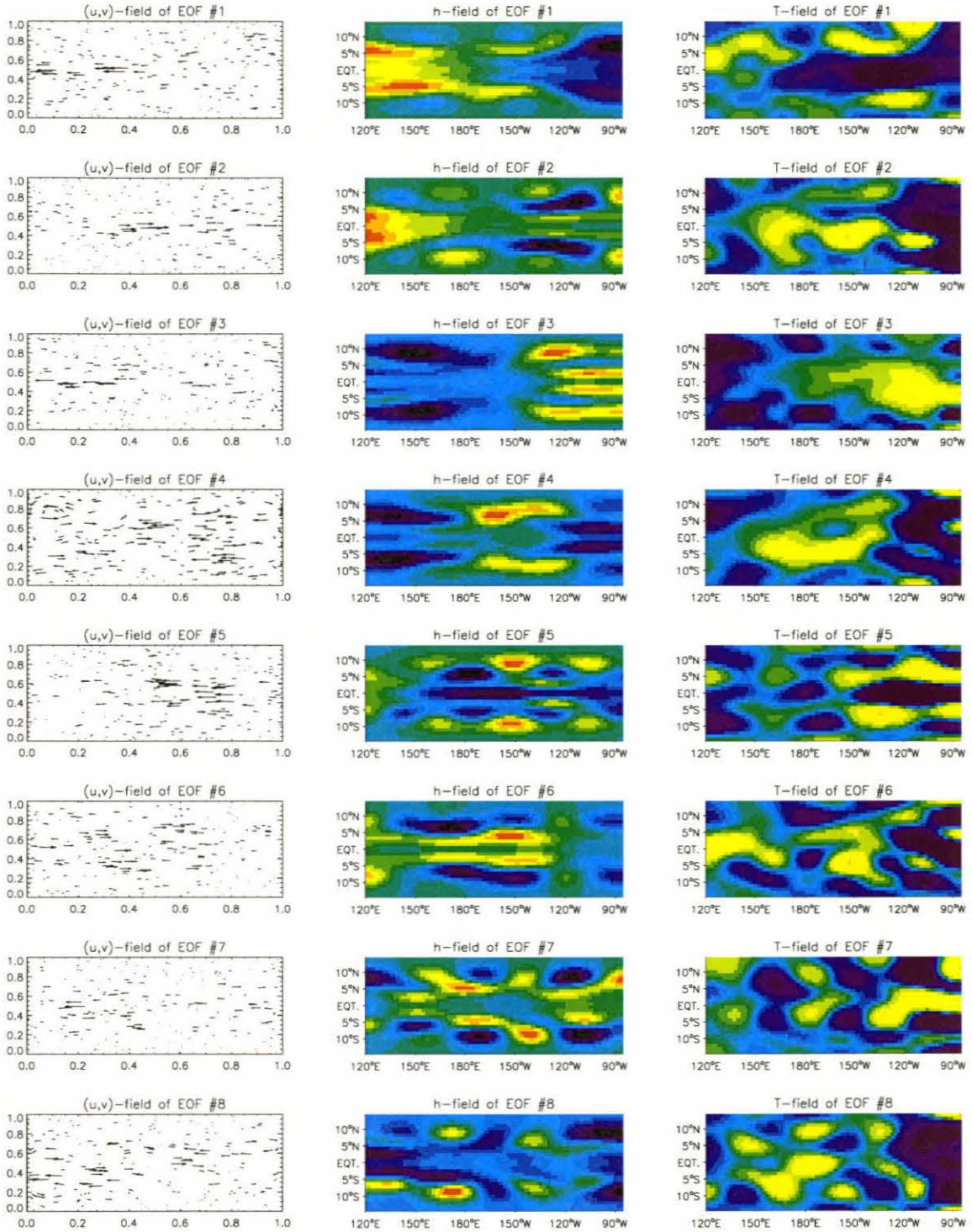


Figure 6.8: The leading EOF patterns associated with the eigenvalues in Fig. 6.7.

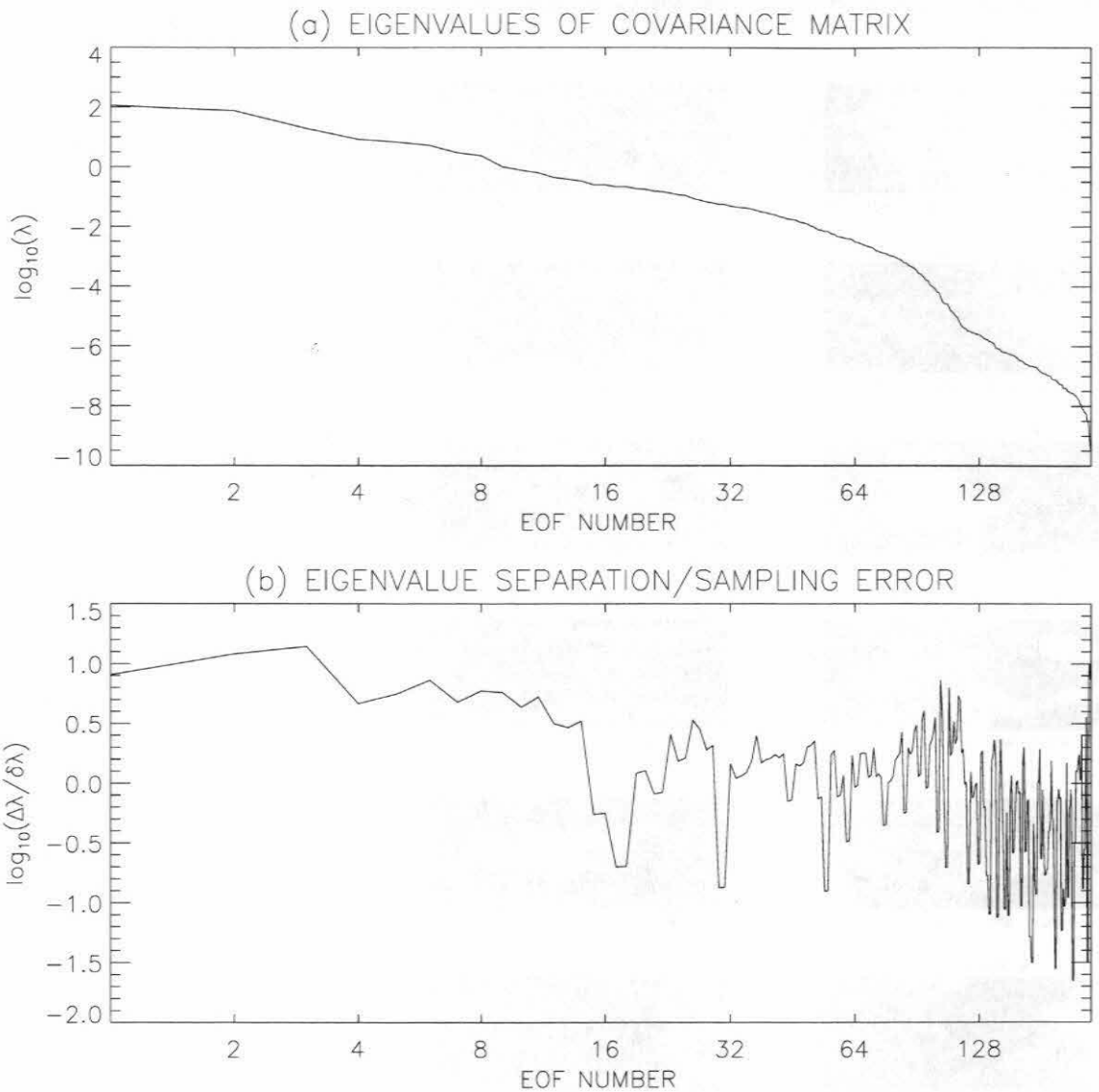


Figure 6.9: Eigenvalues of the covariance matrix of the coupled run with $K_Q W = 0.0140\, m^3\, s^{-4}\, {}^\circ C^{-1}$ with a seasonal cycle but no stochastic forcing.

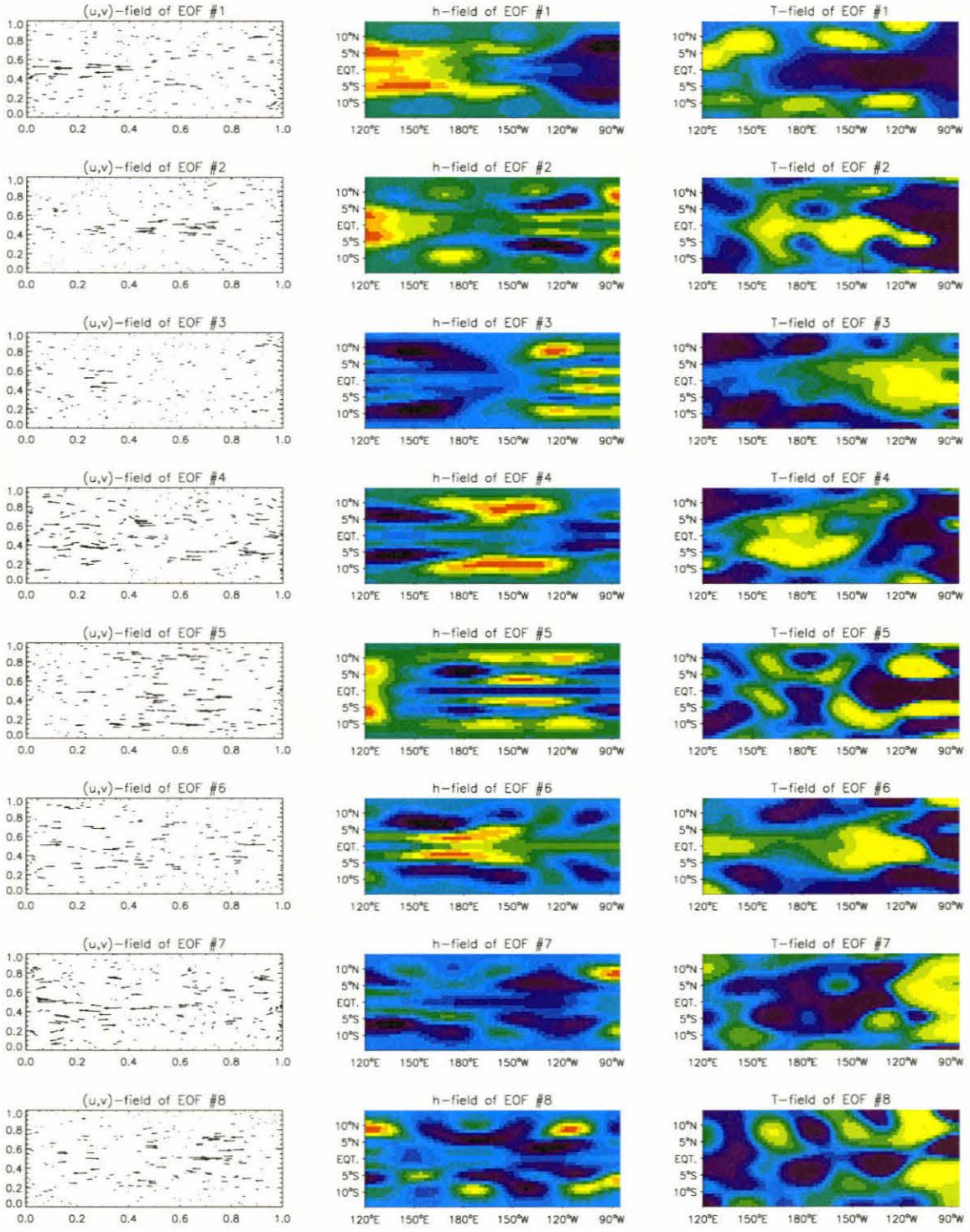


Figure 6.10: The leading EOF patterns associated with the eigenvalues in Fig. 6.9.

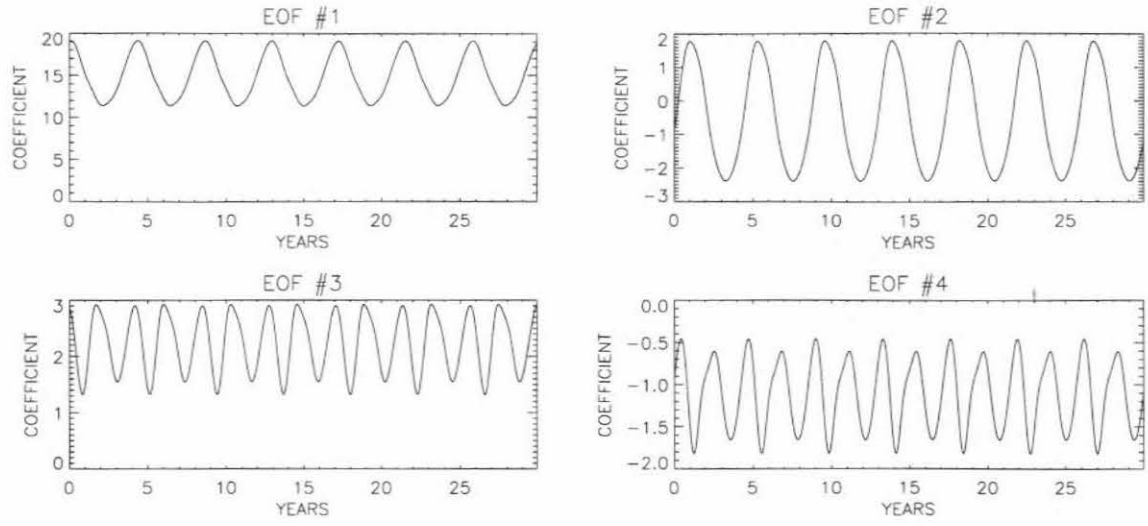


Figure 6.11: The time series of the coefficients of the leading 4 EOFs with $K_Q W = 0.0097\, m^3\, s^{-4}\, {}^\circ C^{-1}$, no seasonal cycle and no stochastic forcing.

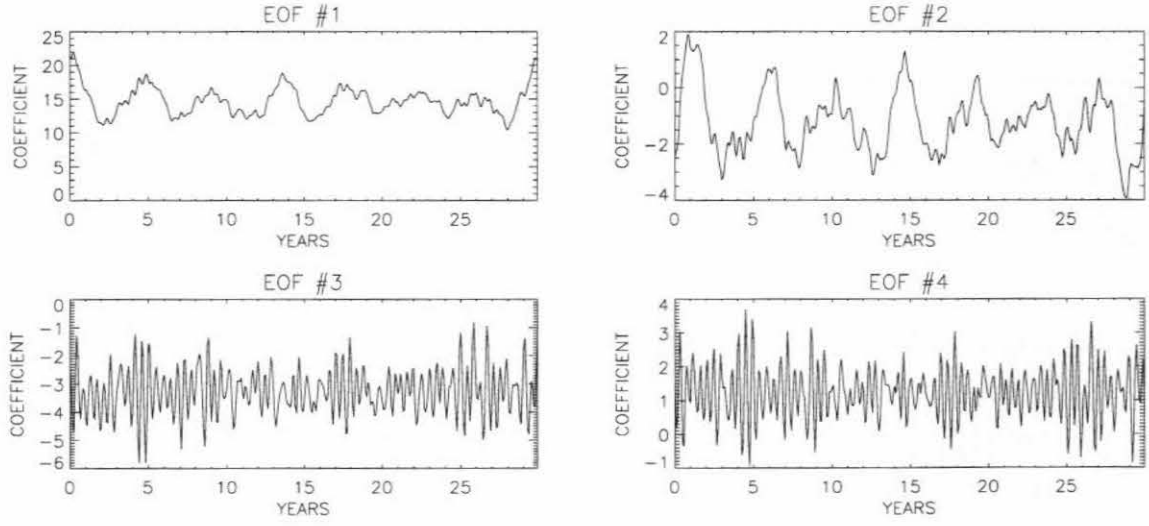


Figure 6.12: The time series of the coefficients of the leading 4 EOFs with $K_Q W = 0.0097\, m^3\, s^{-4}\, {}^\circ C^{-1}$ and no seasonal cycle but with stochastic wind forcing with a white spectrum.

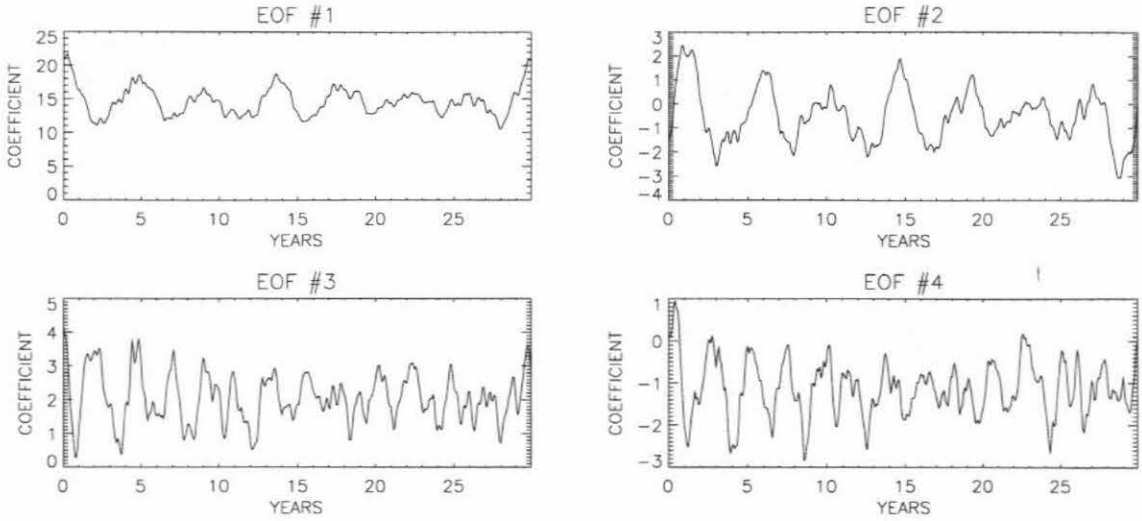


Figure 6.13: The case with $K_Q W = 0.0097\, m^3\, s^{-4}\, ^\circ C^{-1}$ and no seasonal cycle but with stochastic wind forcing projected onto the EOFs of the noise free case (Fig. 6.2).

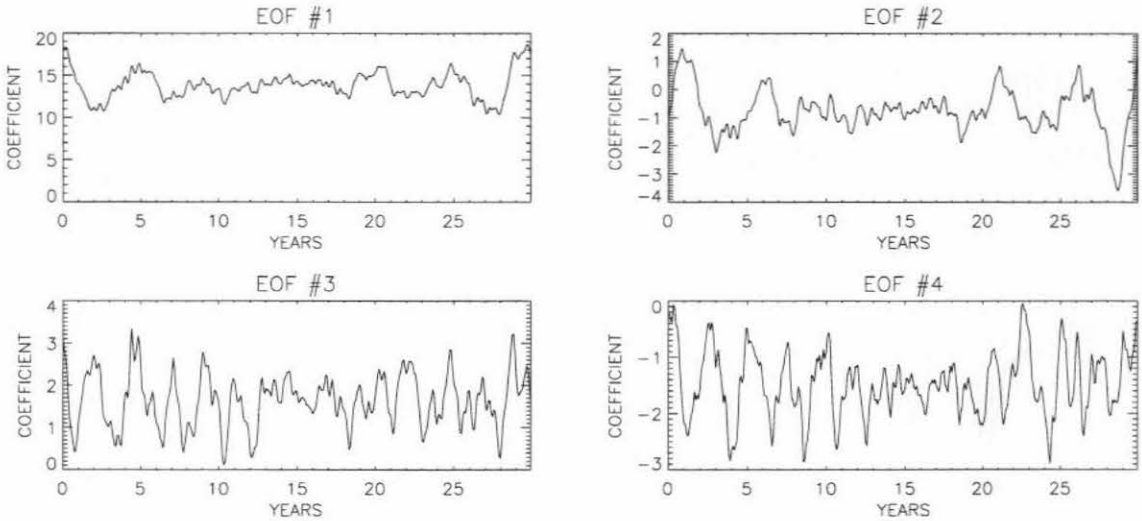


Figure 6.14: The case with $K_Q W = 0.0090\, m^3\, s^{-4}\, ^\circ C^{-1}$ and no seasonal cycle but with stochastic wind forcing projected onto the EOFs of the noise free case (Fig. 6.2).

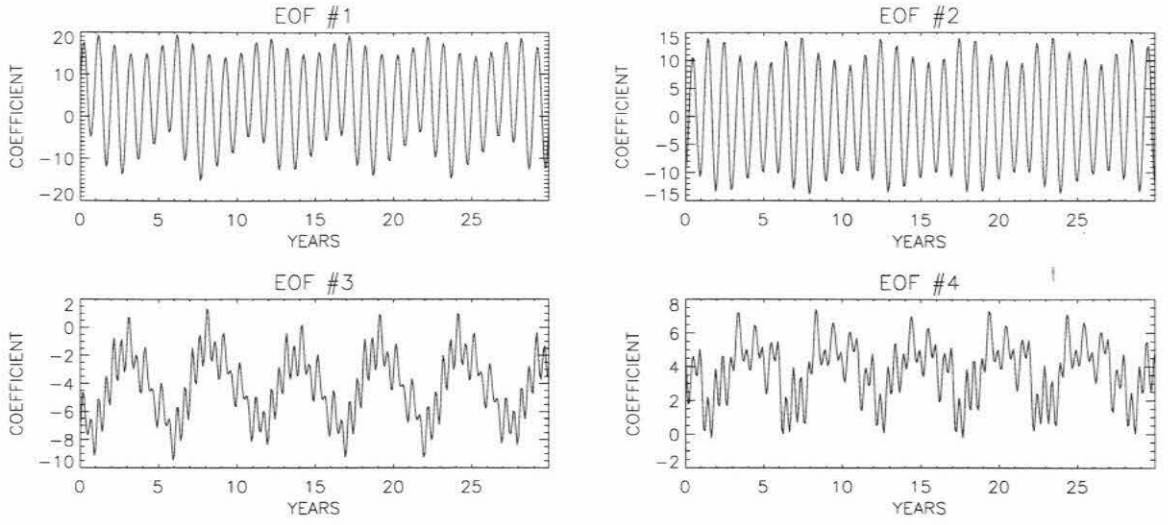


Figure 6.15: The time series of the coefficients of the leading 4 EOFs with $K_Q W = 0.0136\, m^3\, s^{-4}\, {}^\circ C^{-1}$ with a seasonal cycle but no stochastic forcing.

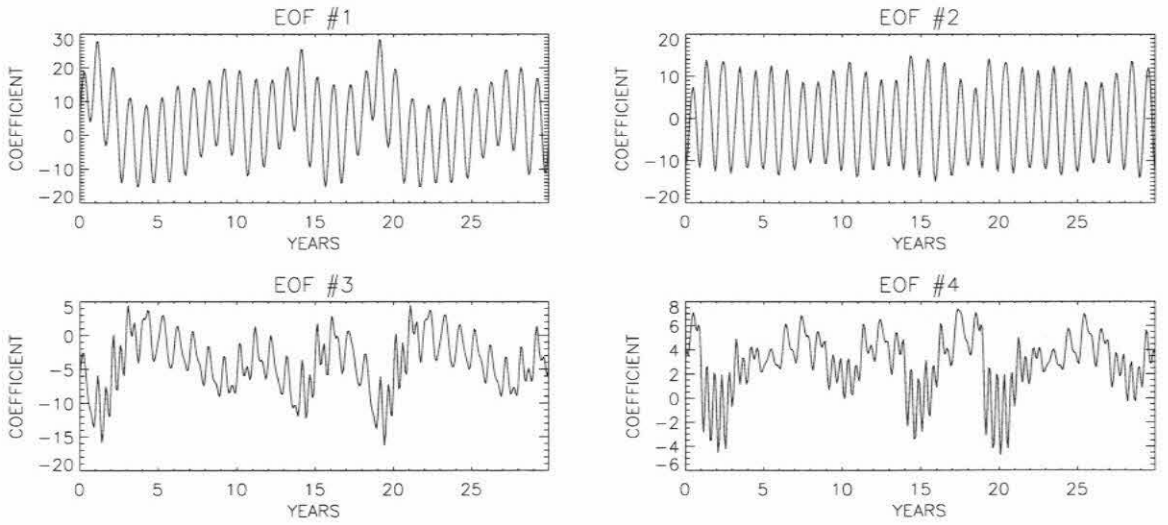


Figure 6.16: The time series of the coefficients of the leading 4 EOFs with $K_Q W = 0.0140\, m^3\, s^{-4}\, {}^\circ C^{-1}$ with a seasonal cycle but no stochastic forcing.

Chapter 7

The Reduced Models

7.1 Nonseasonal Model without Noise

Before any attempt at analysis of the reduced models, their performance will be compared to that of the full models described in Chapter 3.

The coupled model was projected onto its leading 2 EOFs (see Fig. 6.2) and forced with the constant wind forcing. The output of this reduced model is shown in Fig. 7.1. It should be compared with Fig. 3.7. The oscillation in NINO 3 SSTA (panel a in both figures) is similar in both cases. However, the peak-to-peak amplitude in the 2 EOF model is about 3°C in comparison with the 2°C of the full model. The oscillation in the reduced model is also more sinusoidal. The patterns of SST variability (panel b) are very similar. The most obvious difference between the full and 2 EOF models can be seen in their amplitude spectra (panel c). The 2.1 year oscillation of the full model is not present in the 2 EOF model. This absence is not suprising given the time series shown in Fig. 6.11 which showed that the leading 2 EOFs contain the 4.5 year "quasi-quadrennial" (QQ) oscillation and that the 2.1 year oscillation is contained in the third and fourth EOFs.

Inclusion of EOF 3 and EOF 4 is required to retain the secondary "quasi-biennial" peak. The output from the 4 EOF model is shown in Fig. 7.2. The 2.1 year oscillation is

retained in this model, the amplitude of the NINO 3 SSTA oscillation is closer to the full model and it becomes less sinusoidal due to presence of the second frequency component. Figure 7.3 shows the behavior of the reduced model when the leading 8 EOFs are included. The additional 4 EOFs have little effect on the performance of the model.

The first step in the analysis of the reduced model was to linearize the 4 EOF model about its mean state and to find its normal modes and their corresponding eigenvalues. These eigenvalues are plotted on the complex plain in Fig. 7.4. They were calculated for increasing values of coupling. The smallest symbol size represents the uncoupled model projected onto the EOFs of the coupled model while the largest symbol size represents the coupling strength used in the coupled model. From Fig. 7.4 it can be seen that for the standard coupling strength one pair of eigenvalues obtains positive real parts. The point when this pair crosses the imaginary axis represents the Hopf bifurcation above which the model exhibits sustained interannual oscillations. Notice that for zero coupling the real part of the eigenvalues is close to the coefficient of Rayleigh friction, r. This figure thus shows how coupled processes turn a somewhat obscure ocean mode into an unstable coupled mode. The oscillation period associated with this pair is about 4.5 years; this is the dominant peak in the spectrum of the model shown in Fig. 7.2c. Figure 7.4 also shows a second pair of eigenvalues which for the standard coupling correspond to a period of about 2.1 years but which have a negative real part indicating that in the linear model this mode of oscillation will be decaying. However, Fig. 7.2c does show a peak at this period which does represent a sustained oscillatory component. There must be a transfer of energy from the unstable QQ mode to the stable QB model due to nonlinear coupling.

To investigate what processes destabilize the QQ mode, the model was projected onto the leading 2 EOFs and linearized about its mean state.

Let X_1 and X_2 be the coefficients of EOF 1 and EOF 2 respectively. The equations

for their evolution are

$$\frac{dX_1}{dt} = (r_{11}+\alpha_{11}+\beta_{11}+g_{11}+H_{11}+K_{11}+A_{11}+a_{11}+W_{11}+w_{11})X_1 + (r_{12}+\alpha_{12}+\beta_{12}+g_{12}+H_{12}+K_{12}+A_{12}+a_{12}+W_{12}+w_{12})X_2 \quad (7.1)$$

$$\frac{dX_2}{dt} = (r_{21}+\alpha_{21}+\beta_{21}+g_{21}+H_{21}+K_{21}+A_{21}+a_{21}+W_{21}+w_{21})X_1 + (r_{22}+\alpha_{22}+\beta_{22}+g_{22}+H_{22}+K_{22}+A_{22}+a_{22}+W_{22}+w_{22})X_2 \quad (7.2)$$

In the equations the r, α, β, g and H terms have been normalized with respect to their standard values given in Table 2.1. The K terms are related to the feedback between SST and the dynamical equations. The A terms are the linearized horizontal advection due to the mean currents in the upper ocean while the a terms are the linearized advection terms due to the shear component of the currents. The W terms are the linearized entrainment terms due to the mean upwelling in the upper ocean while the w terms are due to the upwelling caused by the convergence of the shear currents.

Table 7.1 shows the coefficients of the model. The EOF coefficients, X_1 and X_2, are dimensionless. The coefficients in the table are all in units of $(year^{-1})$. The frequency of the oscillation is approximately given by

$$\omega \approx \sqrt{(g_{12}+H_{12}+K_{12})(g_{21}+H_{21}+K_{21})} \quad (7.3)$$

which gives an approximate frequency of $\omega \approx 1.5\, year^{-1}$ corresponding to a period of about 4 years. Eq. 7.3 implies a strong dependence of the frequency of the mode on the Kelvin wave speed given by $\sqrt{gH}$. An inspection of Table 7.1 indicates that the r, α and g terms tend to stabilize the system whereas the H and K are the major contributors to destabilization. Horizontal advection and vertical entrainment also contribute to destabilization. When interpreting the contribution of advection and entrainment, it should be remembered that by projecting the model onto a small set of EOFs that combine the u, v, h and T fields, part of the effects of advection and entrainment, which cause T to correlate

with the dynamics, has already been "hardwired" into the model. The contribution of K to destabilization is to be expected and is the essence of Bjerknes' hypothesis that positive feedback between the ocean and atmosphere gives rise to El Niño events [Bje66, Bje69]. The contributions of advection and entrainment to instability also make sense. Divergence of atmospheric winds causes divergence of surface water leading to entrainment of colder water and cooling of the ocean surface, promoting further subsidence and divergence of the winds. Westward horizontal advection enhances the SST gradient which strengthens the easterly winds which drive westward advection. Meridional advection also has a positive feedback effect that may be even stronger than the effect of zonal advection [PZM+97]. Thus Bjerknes' explanation for how the tropical Pacific ocean-atmosphere system becomes unstable certainly explains why the QQ mode becomes unstable in the reduced model. However, the 2 EOF reduced model was constructed with the knowledge that the QQ mode of oscillation dominates the variability. The reduced model provides some explanations for what processes destabilize the QQ mode and what processes modify its frequency, but there remains the question: why does the QQ mode become unstable rather than another mode? It has been shown that the QB mode is not linearly unstable and is only excited by being nonlinearly coupled to the QQ mode. What other modes of oscillation exist in the coupled system? Figure 7.5 shows the normal modes of the full coupled model projected onto its leading 8 EOFs (Fig. 6.2) and linearized about its mean state. There is a correspondence between four of the eigenvalues in Fig. 7.5 and the four in Fig. 7.4. but notice that the correspondence is not exact. This is because normal modes and EOFs are *not the same.* If the EOFs exactly coincided with the normal modes then addition of EOFs to the system would merely add normal modes without affecting the existing normal modes. The EOFs are not the same as normal modes. The EOFs are orthogonal by construction, and this is not necessarily true for the normal modes. An inspection of Fig. 7.5 and the output of the 8 EOF model shown in Fig. 7.3 reveals some interesting points. Clearly the dominant oscillation of the system is the QQ mode

with a period of about 4.5 years. This is to be expected since the mode is unstable when the model is fully coupled. However, the other mode of oscillation has a period of about 2.1 years which corresponds to a pair of eigenvalues with negative real values. As stated above this indicates nonlinear coupling between the QB mode and the QQ mode. But notice in Fig. 7.5 that there are two pairs of eigenvalues with less negative growth rates than the QB mode. These have periods of about 1.2 years and 2.9 years in the fully coupled regime. These modes of oscillations do not show up in the amplitude spectrum in Fig. 7.3c. The QB mode is more stable than either of these modes, yet it does manifest itself in the spectrum, presumably its proximity to a multiple of the frequency of the QQ mode means that it couples more effectively to the system's only self-sustaining oscillatory mode. The version of the 2 EOF model which includes the quadratic terms which provide the nonlinear coupling between modes is given in Appendix B.

7.2 Seasonal Model without Noise

From Fig 6.15 it can be seen that the annual cycle and the interannual oscillation are contained in the leading 4 EOFs of the seasonally forced model. The model was projected onto the leading 4 EOFs (shown in Fig. 6.8) and forced with the seasonal cycle. The resulting behavior is shown in Fig. 7.6. It was found that changing the reduced gravity in the reduced model from the $0.0486\, m\, s^{-2}$ to $0.0600\, m\, s^{-2}$ resulted in the reduced model having an interannual frequency closer to the full model. With the standard value the period of the interannual oscillation was about 6.3 years. Figure 7.6 should be compared with the corresponding full model run shown in Fig. 3.9. The dominant interannual oscillation and the seasonal cycle are reproduced, but the other peaks in the spectrum shown in Fig. 3.9c are not present in Fig. 7.6. This is not suprising since these other oscillations are contained in the higher EOFs. The effect of increasing the number of EOFs included in the model was investigated. The model was projected onto 8 EOFs

and 16 EOFs. The resulting behavior is shown in Figs. 7.7 and 7.8 respectively. The parameters of the 8 and 16 EOF models were identical to those of the full model.

The 8 EOF model exhibits an interannual oscillation and an annual oscillation. There is also an additional peak at a period of about 1.2 years. The amplitude of the interannual peak is much larger than in the full model case.

The 16 EOF model shows subsidiary peaks similar to the full model although their amplitudes are not exactly the same. The dominant interannual oscillation is at the same frequency, but its amplitude is more than twice that of the full model. The QB mode is also present although its amplitude is also too large. The secondary peaks on either side of the seasonal cycle corresponding to periods of about 1.2 years and 0.8 years are reproduced, but there is extra peak at 1.5 years which has negligible amplitude in the full model. Although the 16 EOF reduced model does not reproduce the behavior of the full model *exactly*, it does capture its general dynamics, including relatively minor spectral peaks.

To understand the changes in the spectra of the seasonal model as the number of retained EOFs increases, the normal modes of the linearized models were calculated. These are shown in Figs. 7.9, 7.10, 7.11 for the 4 EOF, 8 EOF and 16 EOF models respectively. Notice that in every case *there is only one unstable mode* and it has a period between 4 and 5 years. In the 4 EOF model the other mode has a period very close to 1 year. This mode is excited by the seasonal forcing. In the 8 EOF model one of the extra modes has a period of about 1.25 years; it is still stable but manifests itself because it couples to another mode. Since this mode does not manifest itself in the nonseasonal model, it probably couples to the seasonally driven, 1 year mode, rather than the interannual QQ mode. In the 16 EOF model there are several modes with periods of more than a year. One can be identified with the QB mode of the nonseasonal model. All the peaks in the spectrum in Fig. 7.8c can be identified with a normal model in Fig. 7.11 but all the normal modes remain stable except for the QQ mode.

7.3 Stochastically Forced Models

The response of the reduced models to stochastic windstress was investigated. The stochastic wind forcing with a white spectrum was used to force the reduced models. Table 7.2 shows how much of the variance of the white noise windstress projects onto the leading EOFs of the nonseasonal model shown in Fig 6.2. Notice that less than 6% of the noise variance projects onto the EOFs associated with QQ and QB modes. If these modes were only affected by the noise that directly projected onto their spatial patterns, then the noise forcing would not be very efficient at perturbing their evolution.

Figure 7.12 shows the amplitude spectra of the reduced models when driven with the white wind noise. Note that not until 32 EOFs are included does the reduced model even begin to reproduce the response of the full model. If fewer EOFs are included, the model spectrum is not broadband enough. It is found that the spectra of the time series of the coefficients of the leading EOFs follow a trend similar to that of the NINO 3 SSTA time series. This implies that in the models that include more EOFs, more noise energy flows into the leading modes and that this energy is not the wind noise that projects directly onto the those modes. The leading modes must be getting this energy by coupling to many other modes which are being excited by noise. Therefore, even if the aim is only to understand the impact of noise on the leading modes of the system, many other modes should be included since these provide a route for noise power to reach the leading modes. An alternative approach would be to parameterize the effect of coupling to noise driven modes on the leading modes, but naively projecting the noise onto the leading modes will greatly underestimate its impact on their evolution.

7.4 Discussion and Conclusions

In this chapter it has been shown that the primary interannual oscillatory behavior of the ICM can be captured by low dimensional models. It has been found that the overall interannual behavior of the full model is the product of the oscillations of a range of coupled modes with frequencies in the interannual range. The manifestation of a particular mode may be for one of the following three reasons.

I. The mode is linearly unstable, *e.g.*, the quasi-quadrennial mode in Fig. 7.1 and 7.2. Such modes manifest themselves without stochastic forcing and in the absence of other modes.

II. The mode is linearly stable but couples to another mode which is linearly unstable or to seasonal forcing, *e.g.*, the quasi-biennial mode in Fig. 7.2. Such modes manifest themselves in the absence of noise provided that the unstable mode they couple to is present.

III. The mode is linearly stable and does not couple particularly well to unstable modes but can be excited by stochastic forcing. There are many such modes in the full model and it is the superposition of such modes that leads to the broadband response of the stochastically forced model.

In the case of the model above the bifurcation at which the leading mode becomes unstable, all three types of modes can contribute to the SST variability if stochastic forcing is present. If the model is below the bifurcation, every mode is a type III mode. The flow of energy between the modes is summarized in Fig. 7.13.

The behavior of the model in the absence of stochastic forcing can be well reproduced by low dimensional models which reproduce the type I and type II modes. However the response of the model to stochastic forcing requires higher dimensional models that capture the dynamics of many modes because these modes contribute a significant amount of variability and they also couple to the leading modes thus providing an additional

pathway for noise power to influence the primary modes of oscillation.

The recognition of the existence of the different types of modes also provides a way to reconcile the forecasting success of models that treat ENSO as a stable, noise driven system [PM93, PS95, Pen96, Bur99] and those which assume a self-sustaining oscillation. Even when the stochastically forced model is above the bifurcation at which the primary mode becomes self-sustaining, a large amount of variance is still contained in stable, noise driven modes. For example, in the model run described by Fig. 6.13, the unstable mode contains 50% of the variance, the QB mode contains 6% and 44% of the variance is contained in stable, noise excited modes. This means that the even above the bifurcation the amplitude of the stochastic wind forcing has a significant impact on variability in the interannual frequency range that is identified with ENSO. The implication for models which attempt to model both ENSO and general atmospheric variability from basic principles is that if the amplitude of atmospheric variability is too low, the resulting ENSO related variability will also be too low even if the coupling between the ocean and the atmosphere is well modeled.

One final question is why the number of EOFs required to model the seasonal model without stochastic forcing is higher than the embedding dimension estimated in Chapter 5. In that chapter it was estimated that a four dimensional state space would suffice to model the noise free seasonal model whereas in this chapter it was found that 8 to 16 EOFs are required to capture anything beyond the leading interannual mode and the seasonal cycle. One explanation is that projecting the model onto EOFs is a linear transformation whereas the estimate of the embedding dimension does not assume that the coordinate transformation is linear. It is quite possible that some nonlinear transformation of the model variables would lead to an even lower dimensional model although the interpretation of the model may be inhibited by the complexity of the change in state space coordinates.

Table 7.1: Coefficients of the linearized 2 EOF model. All coefficients are in units of $(year^{-1})$.

r_{11}	-0.3926	r_{12}	-0.0087	r_{21}	-0.0087	r_{22}	-0.3922
α_{11}	-0.0847	α_{12}	0.0617	α_{21}	0.0617	α_{22}	-0.0869
β_{11}	0.0000	β_{12}	0.0952	β_{21}	-0.0952	β_{22}	0.0000
g_{11}	-0.3845	g_{12}	-0.3752	g_{21}	1.9738	g_{22}	-0.1424
H_{11}	0.3728	H_{12}	-2.1188	H_{21}	0.4480	H_{22}	0.0961
K_{11}	0.4214	K_{12}	-0.1169	K_{21}	-1.5214	K_{22}	0.3856
A_{11}	0.0258	A_{12}	-0.0803	A_{21}	-0.0271	A_{22}	0.1120
a_{11}	0.0344	a_{12}	-0.0278	a_{21}	-0.0191	a_{22}	0.0273
W_{11}	0.0061	W_{12}	-0.0333	W_{21}	-0.0026	W_{22}	0.0252
w_{11}	0.0058	w_{12}	-0.0102	w_{21}	0.0063	w_{22}	-0.0035

Table 7.2: Projection of noise onto the model EOFs.

NO. OF EOFS	CUMULATIVE NOISE VARIANCE
2	2.6%
4	5.8%
8	9.6%
16	29%
20	33%
32	51%

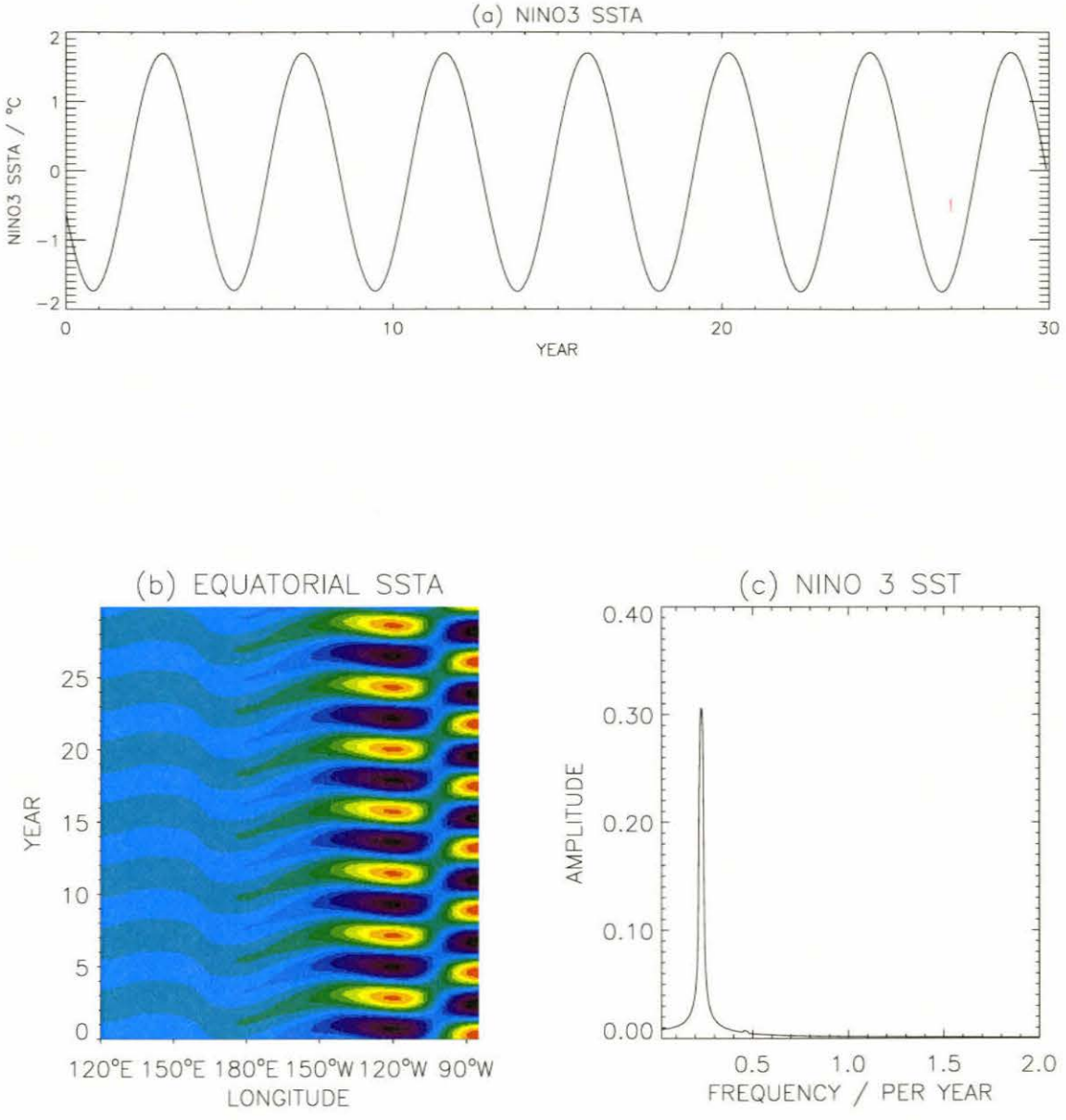


Figure 7.1: Output from the 2 EOF model using the leading 2 EOFs in Fig. 6.2 and a coupling of $K_Q W = 0.0097\, m^3\, s^{-4}\, ^\circ C^{-1}$.

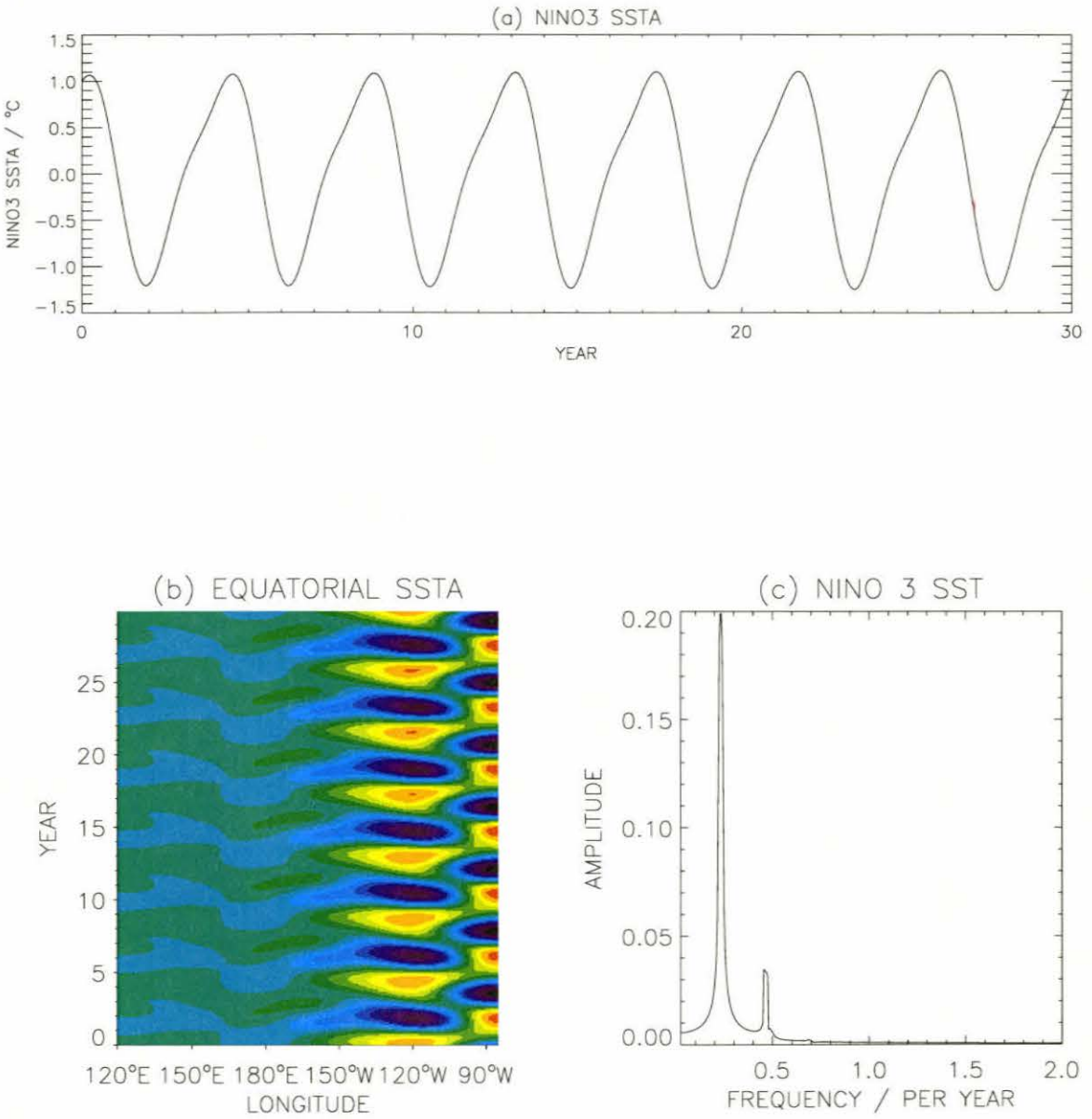


Figure 7.2: Output from the 4 EOF model using the leading 4 EOFs in Fig. 6.2 and a coupling of $K_Q W = 0.0097\, m^3\, s^{-4}\, ^\circ C^{-1}$.

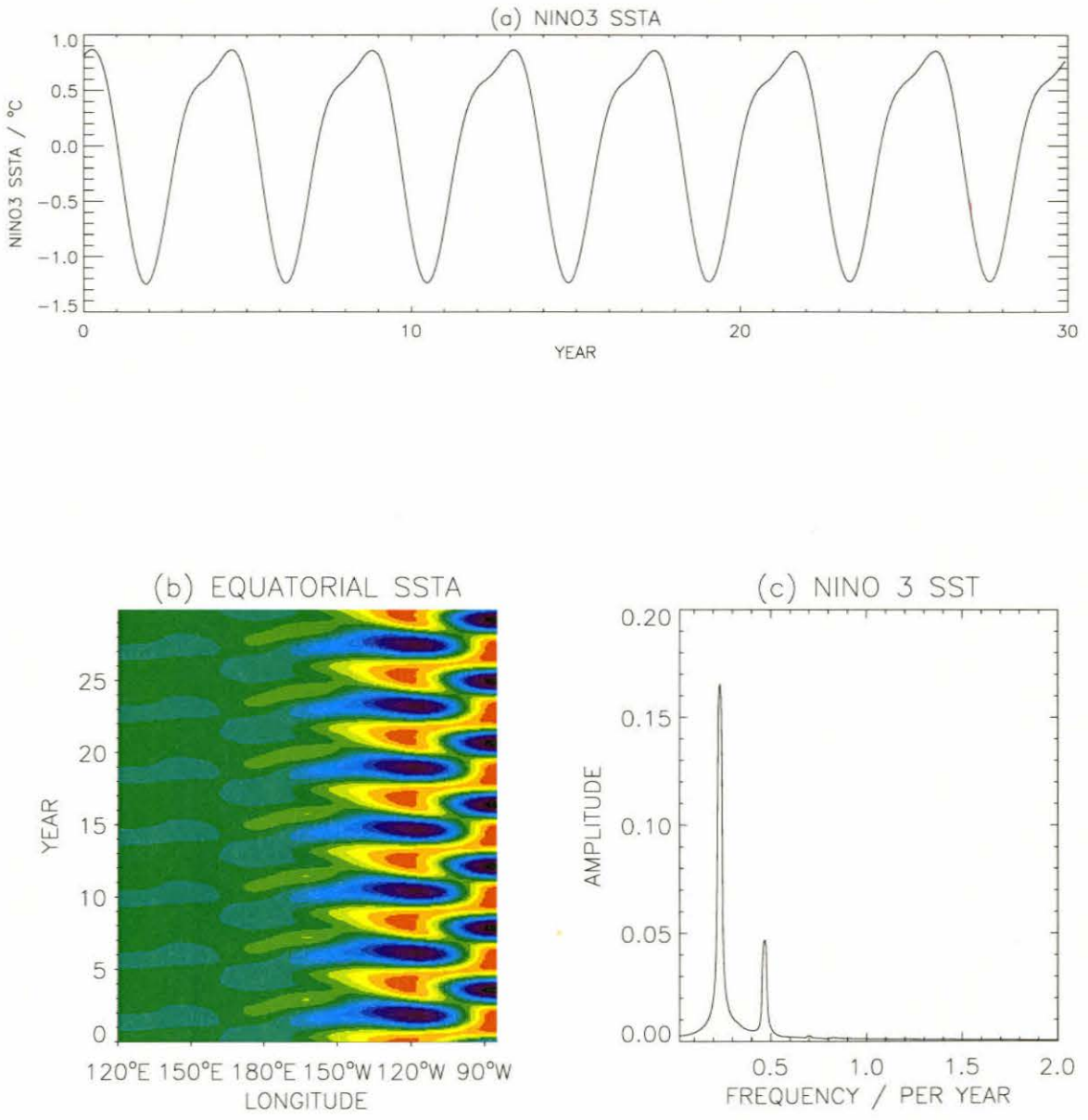


Figure 7.3: Output from the 8 EOF model using the leading 8 EOFs in Fig. 6.2 and a coupling of $K_Q W = 0.0097\, m^3\, s^{-4}\, ^\circ C^{-1}$.

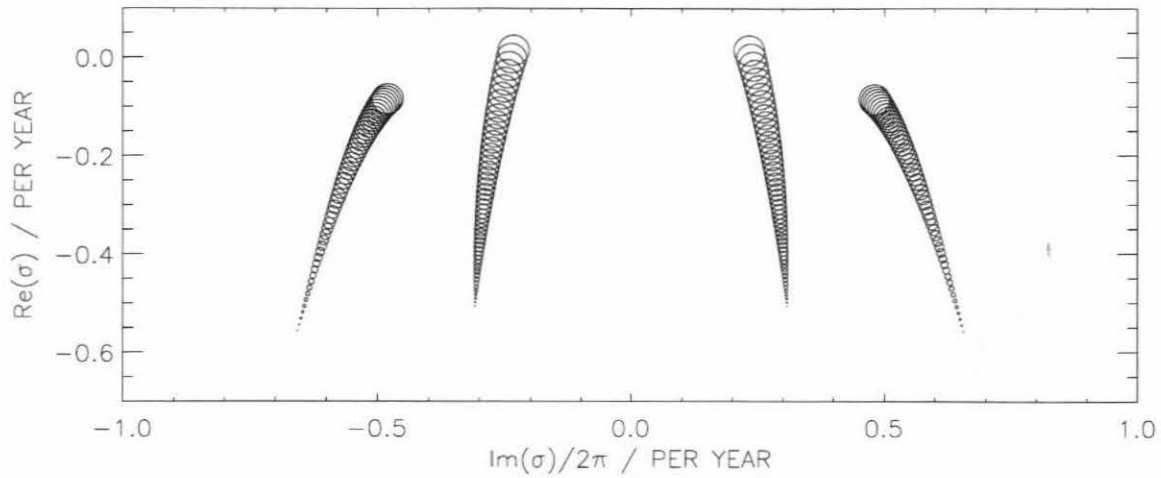


Figure 7.4: The eigenvalues of the normal modes of the 4 EOF reduced model linearized about its mean state. The size of the symbol represents the strength of the coupling, advection and vertical entrainment, ranging from zero to the actual value used in the model.

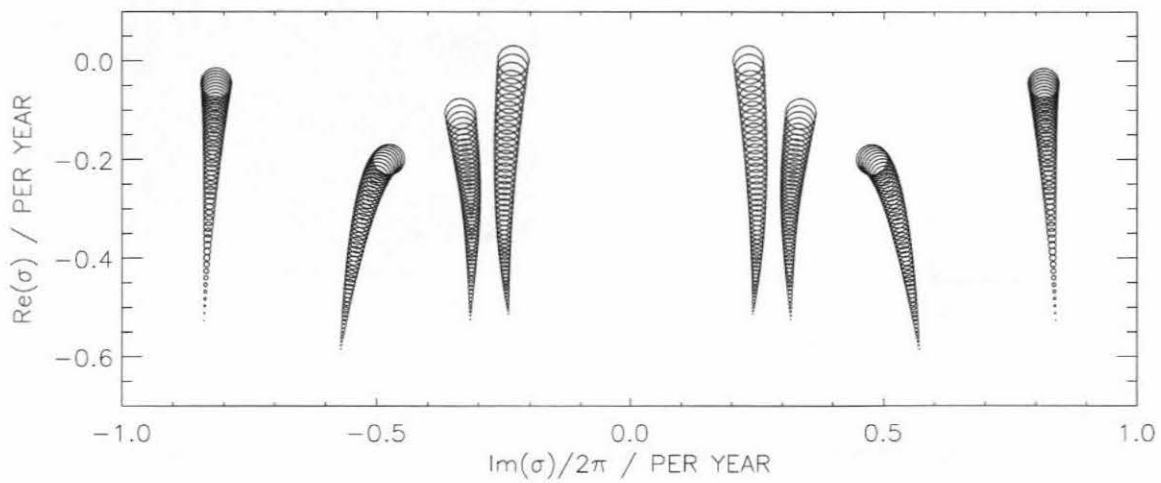


Figure 7.5: As Fig. 7.3 but for the 8 EOF reduced model.

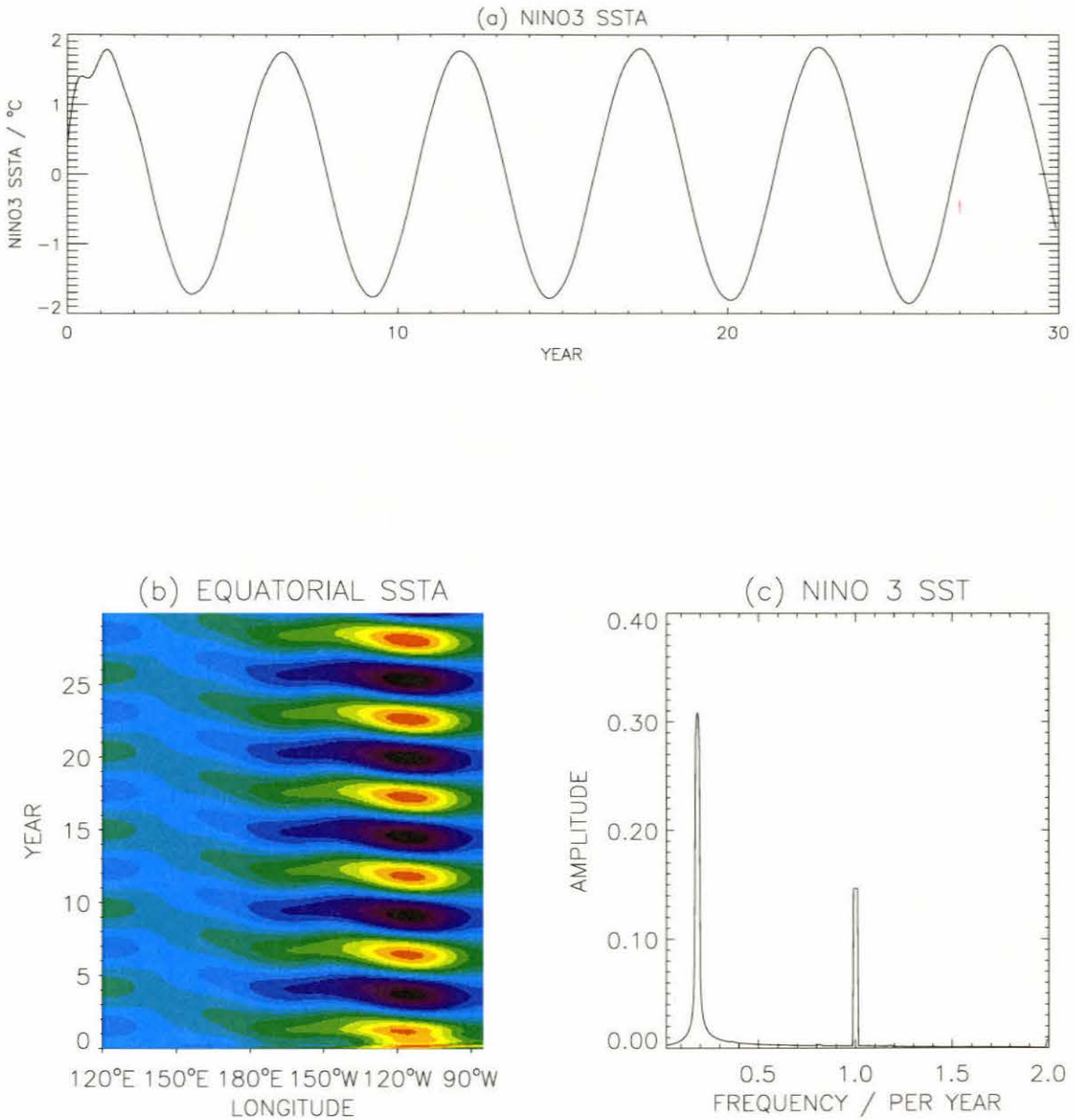


Figure 7.6: The behavior of the 4 EOF model constructed using the leading 4 EOFs in Fig. 6.8 and a coupling strength of $K_Q W = 0.0136\, m^3\, s^{-4}\, ^\circ C^{-1}$ with seasonal forcing.

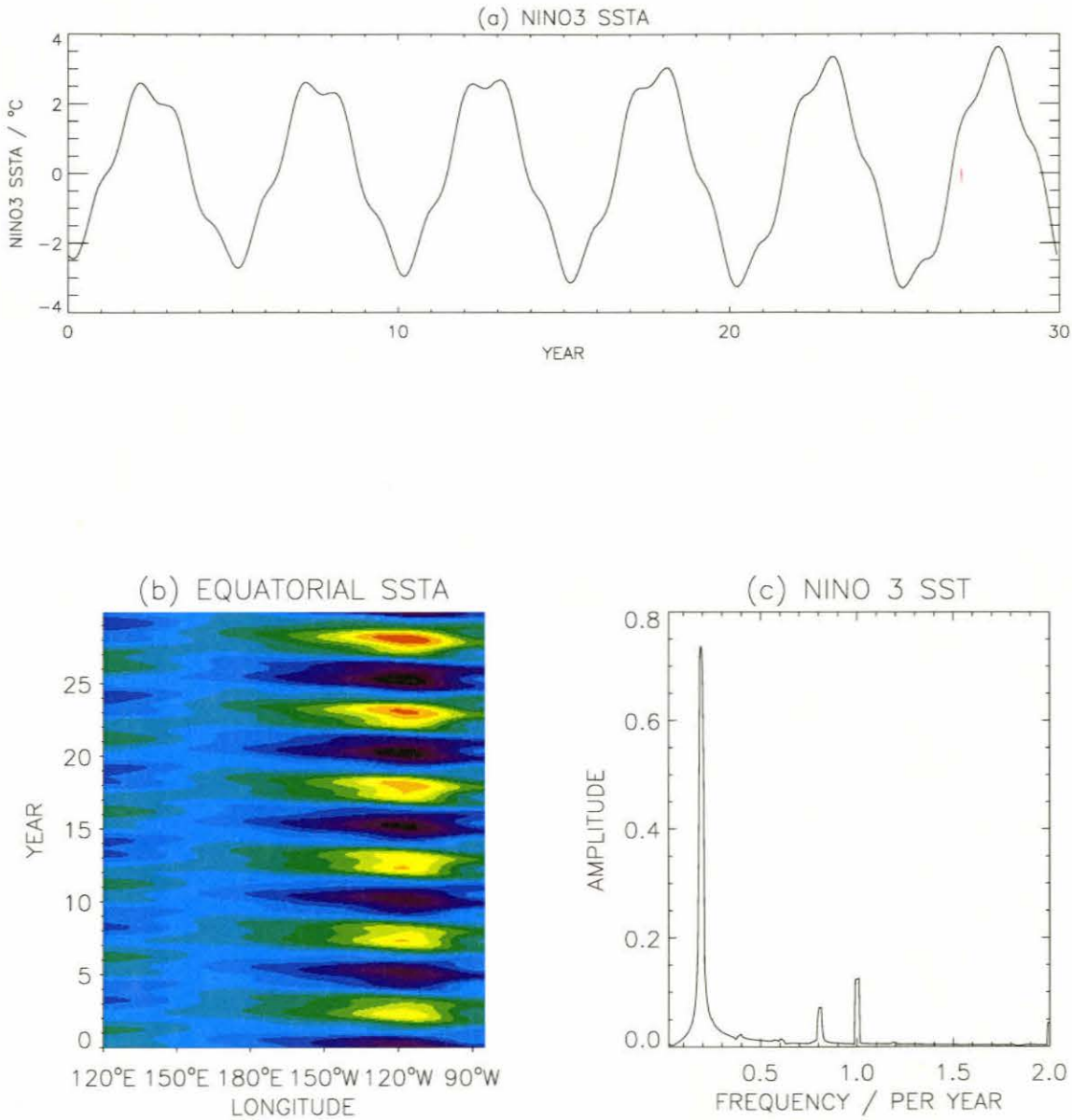


Figure 7.7: As Fig. 7.6 but with 8 EOFs.

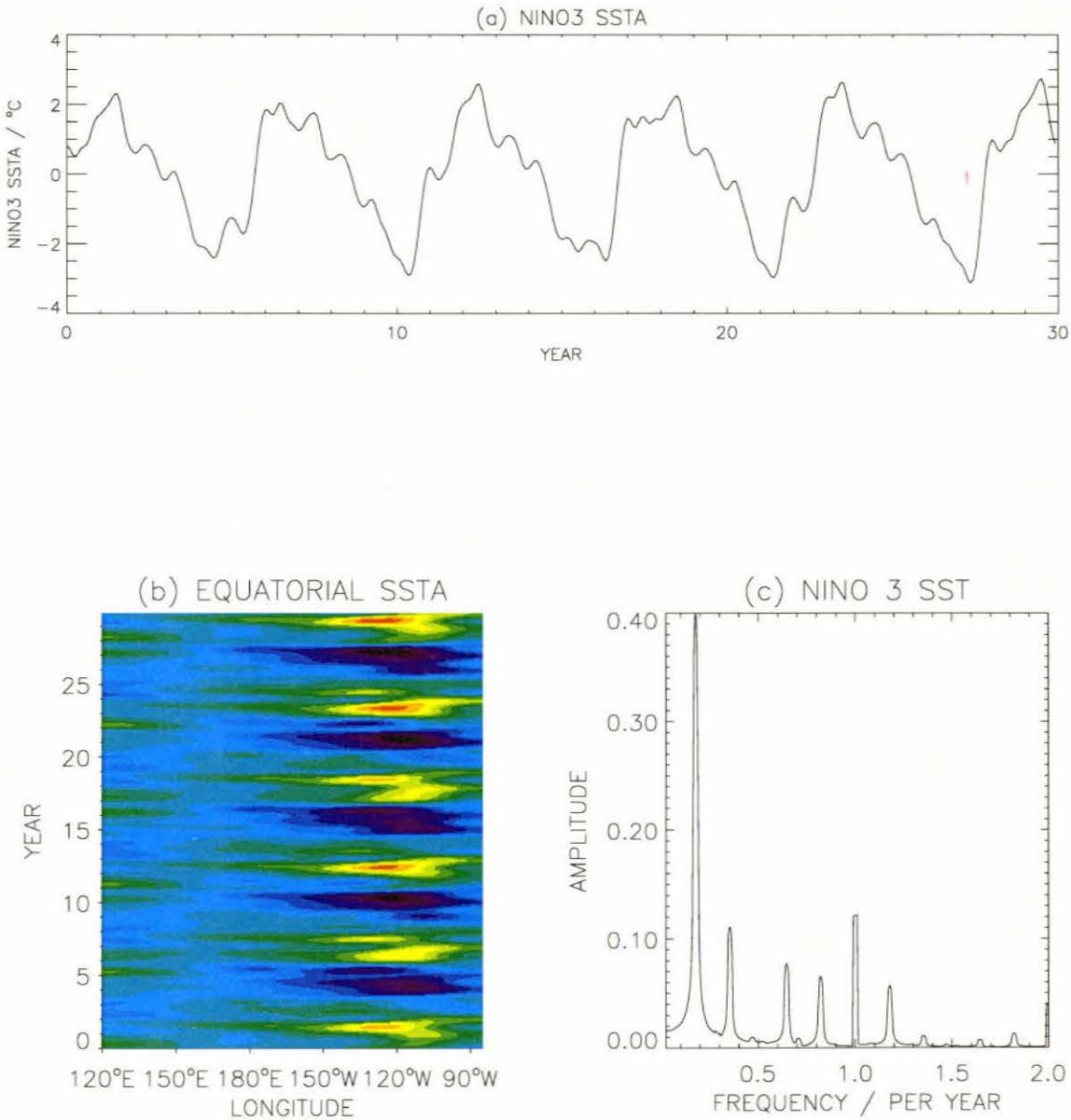


Figure 7.8: As Fig. 7.7 but with 16 EOFs.

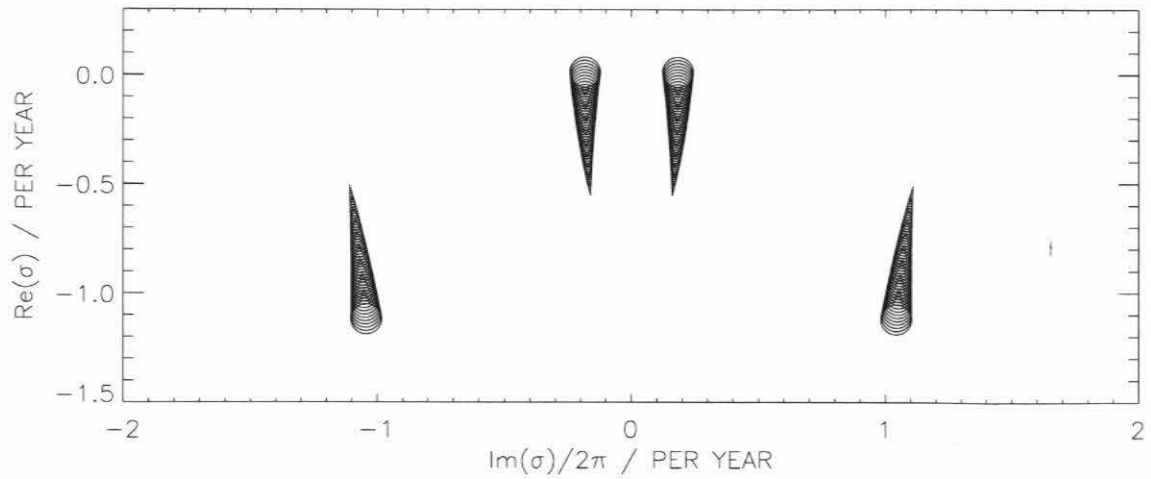


Figure 7.9: The eigenvalues of the normal modes of the 4 EOF reduced seasonally forced model, linearized about its mean state with a coupling strength of $K_Q W = 0.0136\, m^3\, s^{-4}\, {}^\circ C^{-1}$. The size of the symbols indicates the strength of the coupling, from zero to the standard value.

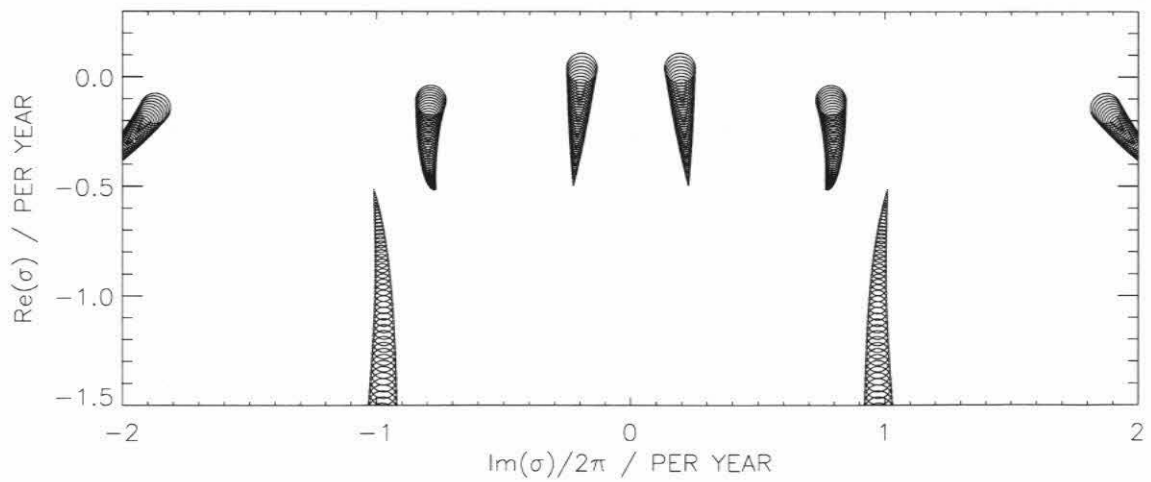


Figure 7.10: As Fig. 7.9 but for the 8 EOF seasonal model.

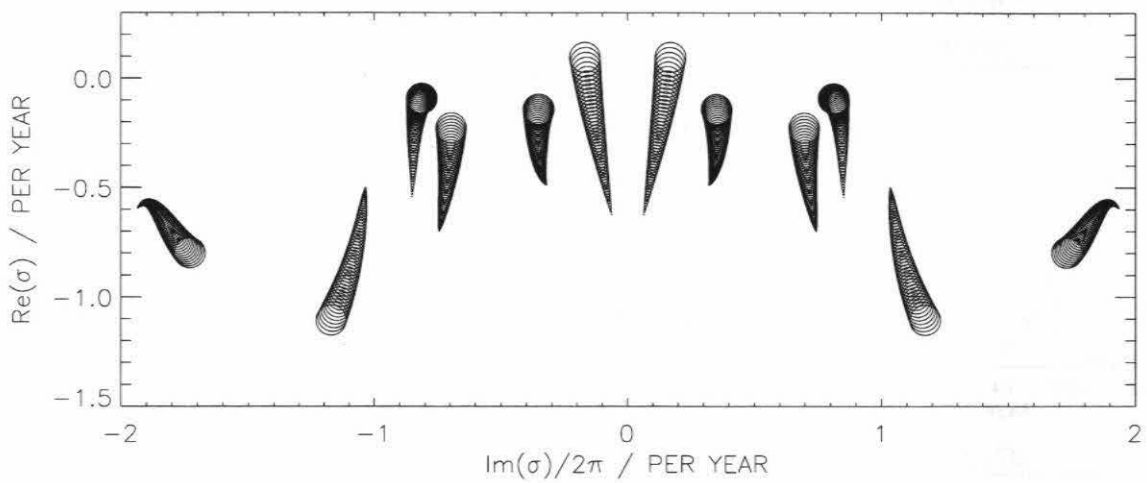


Figure 7.11: As Fig. 7.9 but for the 16 EOF seasonal model.

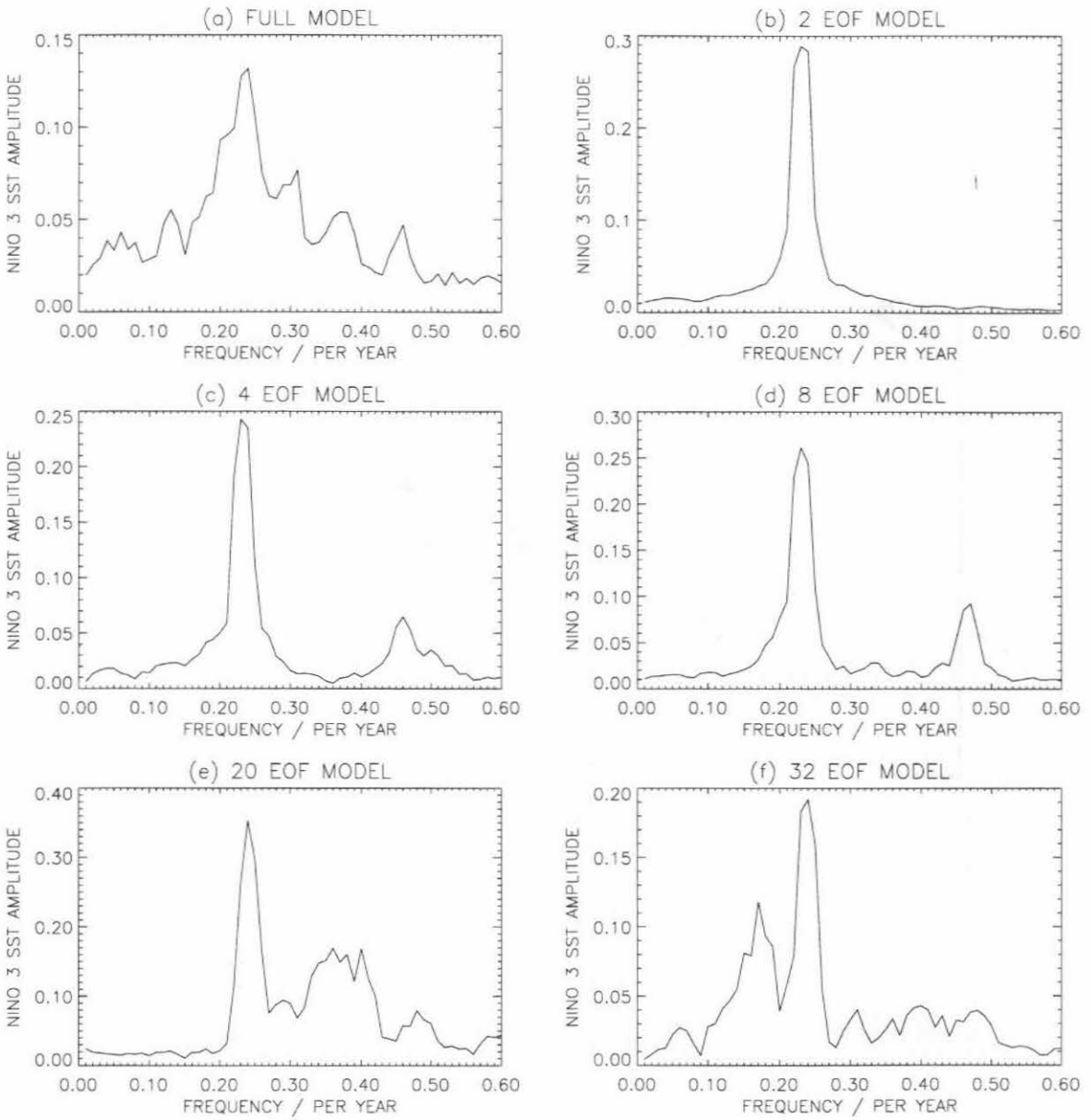


Figure 7.12: Amplitude spectra of the reduced models driven by stochastic wind forcing.

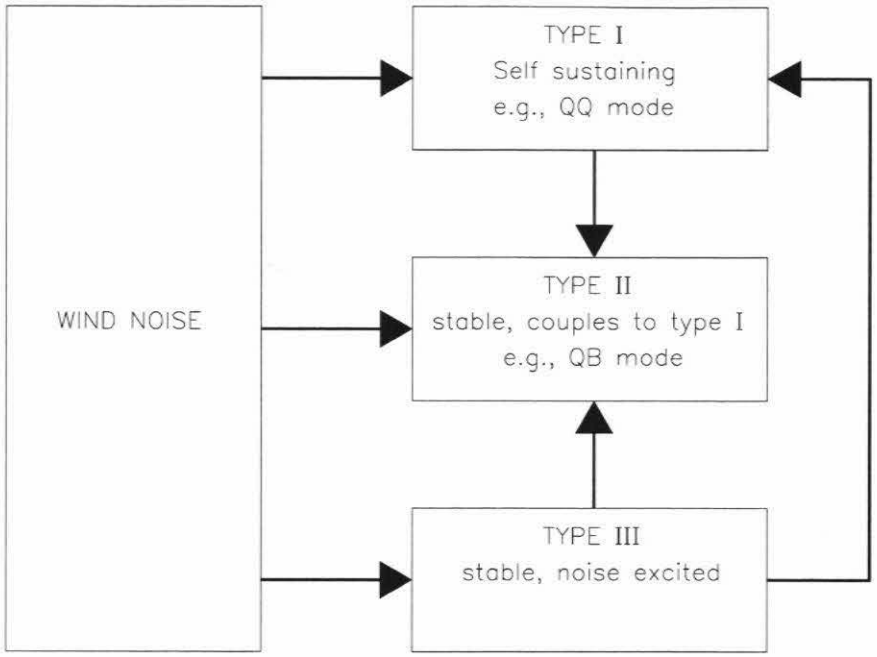


Figure 7.13: The flow of energy between modes in the full coupled model.

Appendix A: Error Estimates on Mutual Information

When the mutual information is determined from a finite amount of data the result will be biased with respect to the result that would be obtained if the true probability and joint probability distributions were known. The following expressions give estimates for what the "true" mutual information, I_∞, is based on the mutual information calculated from data, I_{obs} [Rou99].

$$I_\infty \approx I_{obs} + \frac{B_X^* + B_Y^* - B_{XY}^* - 1}{2N} \pm \sigma_I \tag{A.1}$$

where

$$\sigma_I = \sqrt{\frac{1}{N}\sum_{k=1}^{B_X}\sum_{l=1}^{B_Y}(\ln q_k^X + \ln q_l^Y - \ln q_{kl} + I_{obs})^2 q_{kl}(1 - q_{kl})} \tag{A.2}$$

where N is the number of data points, q_{ij} is the observed joint distribution of X and Y and q^X and q^Y are the observed distributions of X and Y respectively, that is

$$q_k^X = \sum_{j=1}^{B_Y} q_{kj} \qquad q_l^Y = \sum_{i=1}^{B_X} q_{il} \tag{A.3}$$

B_X^* is the number of bins for which $q_i^X \neq 0$, B_Y^* is the number of bins for which $q_i^Y \neq 0$ and B_{XY}^* is the number of bins for which $q_{ij} \neq 0$. For these formulae to give reasonable error estimates, the minimum number of data points in the most populated bin of the joint probability distribution should be greater than about 10.

Appendix B: Quadratic 2 EOF Model

To reduce the 2 EOF model to analytic form, the entrainment term in the SST equation was modified.

$$\frac{\gamma}{H_M}\mathcal{H}(w)w(T_B(h)-T) \rightarrow \frac{\gamma}{H_M}\mathcal{H}(\bar{w})w\left(\left(\frac{dT_B}{dh}\right)_{\bar{h}}(h-\bar{h})+T_B(\bar{h})-T\right) \tag{B.1}$$

where an overbar denotes the mean value. In Eq. A1 the linearization of the subsurface temperature profile and the relacement of the heaviside function with its mean value leads to an equation which is quadratic in the variables u, v, h and T. Note that with only terms up to second order there is no bounding nonlinearity, thus the model either decays or grows without limit depending on parameter values.

Let X_1 and X_2 be the coefficients of EOF 1 and EOF 2 respectively. The equations for their evolution are

$$\begin{aligned}\frac{dX_1}{dt} &= (r_{11}+\alpha_{11}+\beta_{11}+g_{11}+H_{11}+K_{11})X_1 \\ &+(r_{12}+\alpha_{12}+\beta_{12}+g_{12}+H_{12}+K_{12})X_2+c_1+F_1 \\ &+(a_{111}X_1^2+a_{122}X_2^2+a_{112}X_1X_2+a_{11}X_1+a_{12}X_2+a_1) \\ &+(w_{111}X_1^2+w_{122}X_2^2+w_{112}X_1X_2+w_{11}X_1+w_{12}X_2+w_1)\end{aligned} \tag{B.2}$$

$$\begin{aligned}\frac{dX_2}{dt} &= (r_{21}+\alpha_{21}+\beta_{21}+g_{21}+H_{21}+K_{21})X_1 \\ &+(r_{22}+\alpha_{22}+\beta_{22}+g_{22}+H_{22}+K_{22})X_2+c_2+F_2 \\ &+(a_{211}X_1^2+a_{222}X_2^2+a_{212}X_1X_2+a_{21}X_1+a_{22}X_2+a_2) \\ &+(w_{211}X_1^2+w_{222}X_2^2+w_{212}X_1X_2+w_{21}X_1+w_{22}X_2+w_2) \qquad \text{(B.3)}\end{aligned}$$

In the Eqs. B.2 and B.3 the r, α, β, g and H terms are proportional to the respective parameters in the full equations, *i.e.*, if r is increased by 10% from its standard value all the r terms get multiplied by 1.1. The K terms are related to the feedback between SST and the dynamical equations. The a terms are the horizontal advection terms and the w terms are the vertical entrainment terms.

The EOF coefficients, X_1 and X_2, are dimensionless. The coefficients in the tables are all in units of $(year^{-1})$.

Table B.1: Coefficients of the dynamical terms of the 2 EOF model.

r_{11}	-0.3926	r_{12}	-0.0087	r_{21}	-0.0087	r_{22}	-0.3922
α_{11}	-0.0847	α_{12}	0.0617	α_{21}	0.0617	α_{22}	-0.0869
β_{11}	0.0000	β_{12}	0.0952	β_{21}	-0.0952	β_{22}	0.0000
g_{11}	-0.3845	g_{12}	-0.3752	g_{21}	1.9738	g_{22}	-0.1424
H_{11}	0.3728	H_{12}	-2.1188	H_{21}	0.4480	H_{22}	0.0961
K_{11}	0.4214	K_{12}	-0.1169	K_{21}	-1.5214	K_{22}	0.3856

Table B.2: Constant terms of the 2 EOF model.

c_1	1.1442	c_2	1.0143
F_1	-2.2191	F_2	-13.4455

Table B.3: Coefficients of the advection terms of the 2 EOF model.

a_{111}	1.4×10^{-3}	a_{11}	0.0171	a_{211}	-2.4×10^{-3}	a_{21}	0.0312
a_{122}	1.7×10^{-3}	a_{12}	-0.0130	a_{222}	-6.6×10^{-3}	a_{22}	-0.0499
a_{112}	-6.3×10^{-3}	a_1	-0.008	a_{212}	1.2×10^{-2}	a_2	-0.1485

Table B.4: Coefficients of the entrainment terms of the 2 EOF model.

w_{111}	8.0×10^{-5}	w_{11}	0.0093	w_{211}	1.9×10^{-4}	w_{21}	-0.0027
w_{122}	-1.1×10^{-3}	w_{12}	-0.0400	w_{222}	2.5×10^{-3}	w_{22}	-0.0428
w_{112}	-2.8×10^{-4}	w_1	0.1463	w_{212}	-1.3×10^{-3}	w_2	-0.0947

Bibliography

[Aba95] H.D.I. Abarbanel. *Analysis of Observed Chaotic Data.* Springer-Verlag, New York, 1995.

[AD93] M.R. Allen and M.K. Davey. Empirical parameterization of tropical ocean-atmosphere coupling: The "inverse Gill problem". *J. Climate*, 6:509–530, 1993.

[AM85] D.L.T. Anderson and J.P. McCreary. Slowly propagating distrubances in a coupled ocean-atmosphere model. *J. Atmos. Sci.*, 42:615–629, 1985.

[AS97] U. Achatz and G. Schmitz. On the closure problem in the reduction of complex atmospheric models by PIPs and EOFs: A comparison for the case of a two-layer model with zonally symmetric forcing. *J. Atmos. Sci.*, 54:2542–2474, 1997.

[BH89] D.S. Battisti and A.C. Hirst. Interannual variability in the tropical atmosphere-ocean system: Influence of the basic state, ocean geometry, and nonlinearity. *J. Atmos. Sci.*, 46:1687–1712, 1989.

[Bje66] J. Bjerknes. A possible response of the atmospheric Hadley circulation to equatorial anomalies of ocean temperature. *Tellus*, 18:820–829, 1966.

[Bje69] J. Bjerknes. Atmospheric teleconnections from the equatorial Pacific. *Mon. Wea. Rev.*, 97:163–172, 1969.

[Bla97] B. Blanke. Estimating the effect of stochastic wind stress forcing on ENSO irregularity. *J. Climate*, 10:1473–1486, 1997.

[BM99] J.-P. Boulanger and C. Menkes. Long equatorial wave reflection in the Pacific Ocean from TOPEX/POSEIDON data during the 1992-1998 period. *Clim. Dyn.*, 15:205–225, 1999.

[BS99] G. Burgers and D.B. Stephenson. The "normality" of El Niño. *Geophys. Res. Lett.*, 26:1027–1030, 1999.

[Bur99] G. Burgers. The El Niño stochastic oscillator. *Clim. Dyn.*, 15:521–531, 1999.

[BZC95a] G. Burger, S.E. Zebiak, and M.A. Cane. Quasi-fixed points and periodic orbits in the Zebiak-Cane ENSO model with applications in Kalman filtering. Part I: Monthly quasi-fixed points. *Mon. Wea. Rev.*, 123:2802–2813, 1995.

[BZC95b] G. Burger, S.E. Zebiak, and M.A. Cane. Quasi-fixed points and periodic orbits in the Zebiak-Cane ENSO model with applications in Kalman filtering. Part II: Periodic orbits. *Mon. Wea. Rev.*, 123:2814–2824, 1995.

[CP84] M.A. Cane and R.J. Patton. A numerical model for low-frequency equatorial dynamics. *J. Phys. Oceanogr.*, 14:1853–1863, 1984.

[CS95] Penland C. and P.D. Sardeshmukh. The optimal growth of tropical sea surface temperature anomalies. *J. Climate*, 8:1999–2024, 1995.

[CSFM92] J.E. Cole, G.T. Shen, R.G. Fairbanks, and M. Moore. *Coral monitors of El Niño/Southern Oscillation dynamics across the equatorial Pacific*, pages 349–375. In Diaz and Markgraf [DM92], 1992.

[CZ85] M.A. Cane and S.E. Zebiak. A theory for El Niño and the Southern Oscillation. *Science*, 228:1084–1087, 1985.

[CZBC95] D. Chen, S.E. Zebiak, A.J. Busalacchi, and M.A. Cane. An improved procedure for El Nino forecasting: Implications for predictability. *Science*, 269:1699–1702, 1995.

[CZC97] D. Chen, S.E. Zebiak, and M.A. Cane. Initialization and predictability of a coupled ENSO forecast model. *Mon. Wea. Rev.*, 125:773–788, 1997.

[CZD86] M.A. Cane, S.E. Zebiak, and S.C. Dolan. Experimental forecasts of El Niño. *Nature*, 321:827–832, 1986.

[Dan64] W. Dansgaard. Stable isotopes in precipitation. *Tellus*, 16:436–468, 1964.

[DM92] H.F. Diaz and V. Markgraf, editors. *El Niño Historical and Paleoclimatic Aspects of the Southern Oscillation.* Cambridge Univ. Press, Cambridge, 1992.

[DP96] B. Dewitte and C. Perigaud. El Niño - La Niña events simulated with Cane and Zebiak's model and observed with satellite or in situ data. Part II: model forced with observations. *J. Climate*, 9:1188–1207, 1996.

[DWCG94] R.B. Dunbar, G.M. Wellington, M.W. Colgan, and P.W. Glynn. Eastern Pacific sea surface temperature since 1600 A.D.: The δ^{18}O record of climate variability in the Galápagos Islands. *Paleoceanography*, 9:291–315, 1994.

[FS86] A.M. Fraser and H.L. Swinney. Independent coordinates for strange attractors from mutual information. *Phys. Rev. A*, 33:1134–1140, 1986.

[Gil80] A.E. Gill. Some simple solutions for heat-induced tropical circulation. *Quart. J.R. Met. Soc.*, 106:447–462, 1980.

[Gil82] A.E. Gill. *Atmosphere-ocean dynamics.* Academic Press, San Diego, 1982.

[Gla96] M.H. Glantz. *Currents of change: El Niño's impact on climate and society.* Cambridge Univ. Press, Cambridge, 1996.

[GO81] S.B. Goldenberg and J.J. O'Brien. Time and space variability of tropical Pacific wind stress. *Mon. Wea. Rev.*, 109:1190–1207, 1981.

[GP83] P. Grassberger and I. Procaccia. Measuring the strangeness of strange attractors. *Physica D*, 9:189–208, 1983.

[Hir86] A.C. Hirst. Unstable and damped equatorial modes in simple couple ocean-atmosphere models. *J. Atmos. Sci.*, 43:606–630, 1986.

[HLB96] P. Holmes, J.L. Lumley, and G. Berkooz. *Turbulence, coherent structures, dynamicals systems and symmetry.* Cambridge Univ. Press, Cambridge, 1996.

[Jin97a] F.-F. Jin. An equatorial rechange paradigm for ENSO II: A stripped down coupled model. *J. Atmos. Sci.*, 54:830–845, 1997.

[Jin97b] F.-F. Jin. An equatorial recharge paradigm for ENSO I: Conceptual model. *J. Atmos. Sci.*, 54:811–829, 1997.

[JN93a] F.-F. Jin and J.D. Neelin. Modes of interannual tropical ocean-atmosphere interaction - a unified view. Part I: Numerical results. *J. Atmos. Sci.*, 50:3477–3503, 1993.

[JN93b] F.-F. Jin and J.D. Neelin. Modes of interannual tropical ocean-atmosphere interaction - a unified view. Part III: Analytical results in fully coupled cases. *J. Atmos. Sci.*, 50:3523–3540, 1993.

[JNG94] F.-F. Jin, J.D. Neelin, and M. Ghil. El Niño on the devil's staircase: annual subharmonic steps to chaos. *Science*, 264:70–72, 1994.

[JNG95] N. Jiang, J.D. Neelin, and M. Ghil. Quasi-quadrennial and quasi-biennial variability in the equatorial Pacific. *Clim. Dyn.*, 12:101–112, 1995.

[JNG96] F.-F. Jin, J.D. Neelin, and M. Ghil. El Niño/Southern Oscillation and the annual cycle subharmonic frequency-locking and aperiodicity. *Physica D*, 98:442–465, 1996.

[Kar46] K. Karhunen. Zur spektraltheorie stochastischer Prozesse. *Ann. Acad. Sci. Fennicae.*, A1 34, 1946.

[KBA92] M.B. Kennel, R. Brown, and H.D.I. Abarbanel. Determining minimum embedding dimension using a geometrical construction. *Phys. Rev. A*, 45:3403–3411, 1992.

[KDG92] C.L. Keppenne, M.D. Dettinger, and M. Ghil. Adaptive filtering and prediction of the southern oscillation index. *J. Geophys. Res.*, 97:20,449–20,454, 1992.

[KM95a] W.S. Kessler and M.J. McPhaden. The 1991-1993 El Niño in the central Pacific. *Deep-Sea Res. II*, 42:295–333, 1995.

[KM95b] W.S. Kessler and M.J. McPhaden. Oceanic equatorial waves and the 1991-93 El Niño. *J. Climate*, 8:1757–1774, 1995.

[KM97] R. Kleeman and A.M. Moore. A theory for the limitation of ENSO predictability due to stochastic atmospheric transients. *J. Atmos. Sci.*, 54:753–767, 1997.

[KMG97] T.R. Knutson, S. Manabe, and D.F. Gu. Simulated ENSO in a global coupled ocean-atmosphere model: Multidecadal amplitude modulation and CO_2 sensitivity. *J. Clim.*, 10:138–161, 1997.

[KP94] R. Kleeman and S.B. Power. Limits to predictability in a coupled ocean-atmosphere model due to atmospheric noise. *Tellus A*, 46:529–540, 1994.

[KSJH99] Hughen K.A., D.P. Schrag, S.B. Jacobsen, and W. Hantoro. El Niño during the last interglacial period recorded by a fossil coral from Indonesia. *Geophys. Res. Lett.*, 26:3129–3132, 1999.

[Lau85] N.C. Lau. Modeling the seasonal dependence of the atmospheric responses to observed El Niños 1962-1976. *Mon. Wea. Rev.*, 113:1970–1996, 1985.

[Le45] M. Loẽve. Functions aléatoire de second ordre. *Comptes Rendus Acad. Sci. Paris*, 220, 1945.

[LF85] J.M. Lough and H.C. Fritts. The Southern Oscillation and tree rings: 1600-1961. *J. of Climate and Appl. Meteorology*, 24:952–966, 1985.

[Lor56] E.N. Lorenz. *Empirical orthogonal functions and statistical weather prediction.* MIT Press, Cambridge, MA, 1956.

[Lor63] E.N. Lorenz. Deterministic nonperiodic flow. *J. Atmos. Sci.*, 20:130–141, 1963.

[Lou92] J.M. Lough. *An index of the Southern Oscillation reconstructed from western North American tree-ring chronologies*, pages 215–226. In Diaz and Markgraf [DM92], 1992.

[Lum67] J.L. Lumley. *The structure of inhomogeneous turbulence*, pages 166–178. Nauka, Moscow, 1967.

[MB92] G. Meehl and G.W. Branstator. *Coupled climate model simulation of El Niño/Southern Oscillation: implications for paleoclimate*, pages 69–91. In Diaz and Markgraf [DM92], 1992.

[Mee90] G.A. Meehl. Seasonal cycle forcing of El Niño-Southern Oscillation in a global coupled ocean-atmosphere GCM. *J. Climate*, 3:72–98, 1990.

[MJ71] R.A. Madden and P.R. Julian. Detection of a 40-50 days oscillation in the zonal wind in the tropical Pacific. *J. Atmos. Sci.*, 28:702–708, 1971.

[MJ72] R.A. Madden and P.R. Julian. Description of global-scale circulation cells in the tropics with a 40-50 day period. *J. Atmos. Sci.*, 29:1109–1123, 1972.

[MK96] A.M. Moore and R. Kleeman. The dynamics of error growth and predictabililty in a coupled model of ENSO. *Q.J.R. Meteorol. Soc.*, 122, 1996.

[MK99] A.M. Moore and R. Kleeman. Stochastic forcing of ENSO by the intraseasonal oscillation. *J. Climate*, 12:1199–1220, 1999.

[MnY81] Rand D. Mañé, R. and L.S. Young. Springer, Berlin, 1981.

[Moo68] D.W. Moore. *Planetary-gravity waves in an equatorial ocean.* PhD thesis, Harvard, 1968.

[NBH+98] J.D. Neelin, D.S. Battisti, A.C. Hirst, F.-F. Jin, Y. Wakata, T. Yamagata, and S.E. Zebiak. ENSO Theory. *J. Geophys. Res.*, 103:14,261–14,290, 1998.

[NBT+82] G.R. North, Bell, T.L., R.F. Cahalan, and F.J. Moeng. Sampling errors in the estimation of empirical orthogonal functions. *Mon. Wea. Rev.*, 110:699–706, 1982.

[Nee89] J.D. Neelin. Interannual oscillations in an ocean general circulation model coupled to a simple atmosphere model. *Phil. Trans. R. Soc. Lond.*, A329:189–205, 1989.

[Nee90] J.D. Neelin. A hybrid coupled general circulation model for El Niño studies. *J. Atmos. Sci.*, 47:674–693, 1990.

[NJ93] J.D. Neelin and F.-F. Jin. Modes of interannual tropical ocean-atmosphere interaction - a unified view. Part II: Analytical results in the weak-coupling limit. *J. Atmos. Sci.*, 50:3504–3522, 1993.

[PD96] C. Perigaud and B. Dewitte. El Niño - La Niña events simulated with Cane and Zebiak's model and observed with satellite or in situ data. Part I: model data comparisons. *J. Climate*, 9:66–84, 1996.

[Pen96] C. Penland. A stochastic model of IndoPacific sea surface temperature anomalies. *Physica D*, 98:534–558, 1996.

[Phi89] S.G.H. Philander. *El Niño, La Niña, and the Southern Oscillation.* Academic Press, San Diego, 1989.

[PLPN89] S.G.H. Philander, N.C. Lau, R.C. Pacanowski, and M.J. Nath. Two different simulations of Southern Oscillation and El Niño with coupled ocean-atmosphere general circulation models. *Phil. Trans. R. Soc. Lond.*, A329:167–178, 1989.

[PM93] C. Penland and T. Magorian. Prediction of Niño 3 sea surface temperature using linear inverse modeling. *J. Climate*, 6:1067–1076, 1993.

[PS95] C. Penland and P.D. Sardeshmukh. Error and sensitvity analysis of geophysical eigensystems. *J. Climate*, 8:1988–1998, 1995.

[PYP84] S.G.H. Philander, T. Yamagata, and R.C. Pacanowski. Unstable air-sea interactions in the tropics. *J. Atmos. Sci.*, 41:604–613, 1984.

[PZM+97] C. Perigaud, S.E. Zebiak, F. Melin, J.P. Boulanger, and B. Dewitte. On the role of meridional wind anomalies in a coupled model of ENSO. *J. Climate*, 10:761–773, 1997.

[Qui92] W.H Quinn. *A study of Southern Oscillation-related climatic activity for A.D. 622-1900 incorporating Nile River flood data.* In Diaz and Markgraf [DM92], 1992.

[RK75] J. Rinne and V. Karhila. A spectral barotropic model in horizontal empirical orthogonal functions. *Q.J.R. Met. Soc.*, 101:365–382, 1975.

[ROB+96] E. Roeckner, J.M. Oberhuber, A. Bacher, M. Christoph, and I. Kirchner. ENSO variability and atmospheric response in a global coupled atmosphere-ocean GCM. *Clim. Dynam.*, 12:737–754, 1996.

[Rou97] M.S. Roulston. Significance testing of information theoretic functionals. *Physica D*, 110:62–66, 1997.

[Rou99] M.S. Roulston. Estimating the errors on measured entropy and mutual information. *Physica D*, 125:285–294, 1999.

[Ruk63] L.V. Rukhovetz. The optimum representation of the vertical distribution of certain meteorological elements. *Izv. Acad. Sci. U.S.S.R, Geoph. Ser.*, 4:626–326, 1963.

[SAB+98] A.R. Solow, R.F. Adams, K.J. Bryant, D.M. Legler, J.J. O'Brien, B.A. McCarl, W. Nayda, and R. Weiher. The value of improved ENSO prediction to U.S. agriculture. *Climatic Change*, 39:47–60, 1998.

[Sel93] F.M. Selten. Toward an optimal description of atmospheric flow. *J. Atmos. Sci.*, 50:861–877, 1993.

[Sel95] F.M. Selten. An efficient description of the dynamics of barotropic flow. *J. Atmos. Sci.*, 52:915–936, 1995.

[Sel97a] F.M. Selten. Baroclinic empirical orthogonal functions as basis functions in an atmpospheric model. *J. Atmos. Sci.*, 54:2099–2114, 1997.

[Sel97b] F.M. Selten. A statistical closure of a low-order barotropic model. *J. Atmos. Sci.*, 54:1085–1093, 1997.

[SHGP87] K.R. Sperber, S. Hameed, W.L. Gates, and G.L. Potter. Southern Oscillation simulated in a global climate model. *Nature*, 329:140–142, 1987.

[Smi88] L.A. Smith. Intrinsic limits on dimension calculations. *Phys. Lett. A*, 133:283–288, 1988.

[SRLS96] T.M. Smith, R.W. Reynolds, R.E. Livezey, and D.C. Stokes. Reconstruction of historical sea surface temperatures using empirical orthogonal functions. *J. Climate*, 9:1403–1420, 1996.

[SS88a] P.S. Schopf and M.J. Suarez. Vacillation in a coupled ocean-atmosphere model. *J. Atmos. Sci.*, 45:549–566, 1988.

[SS88b] M.J. Suarez and P.S. Schopf. A delayed action oscillator for ENSO. *J. Atmos. Sci.*, 45:3283–3287, 1988.

[SSHP98] L. Stone, P.I. Saparin, A. Huppert, and C. Price. El Niño chaos: The role of noise and stochastic resonance on the ENSO cycle. *Geophys. Res. Lett.*, 25:175–178, 1998.

[SYC91] T. Sauer, J.A. Yorke, and M. Casdagli. Embedology. *J. Statistical Physics*, 65:579–616, 1991.

[SZC88] R. Seager, S.E. Zebiak, and M.A. Cane. A model of the tropical Pacific sea surface temperature climatology. *J. Geophys. Res.*, 93:1265–1280, 1988.

[Tak81] F. Takens. *Detecting strange attractors in turbulence, Warwick 1981.* Springer, Berlin, 1981.

[TGL+91] J. Theiler, B. Galdrikian, A. Longtin, S. Eubank, and J.D. Farmer. *Using surrogate data to detect nonlinearity in time series*, pages 163–188. Addison-Wesley, Redwood City, 1991.

[TMTA84] L.G. Thompson, E. Mosley-Thompson, and B.M. Arnao. El Niño-Southern Oscillation events recorded in the stratigraphy of the tropical Quelccaya ice cap, Peru. *Science*, 226:50–53, 1984.

[TMTB85] L.G. Thompson, E. Mosley-Thompson, and J.F. Bolzan. A 1500-year record of tropical precipitation in ice cores from the Quelccaya ice cap, Peru. *Science*, 229:971–973, 1985.

[TSCJ94] E. Tziperman, L. Stone, M.A. Cane, and H. Jarosh. El Niño chaos: overlapping resonances between the seasonal cycle and the Pacific Ocean-atmosphere oscillator. *Science*, 264:72–74, 1994.

[Val86] G.K. Vallis. El Niño: a chaotic dynamical system? *Science*, 232:243–245, 1986.

[Wal23] G.T. Walker. Correlation in seasonal variations of weather VIII. *Mem. Ind. Meteor. Dept.*, 24:75–131, 1923.

[Wal24] G.T. Walker. World weather IX. *Mem. Ind. Meteor. Dept.*, 24:275–332, 1924.

[Wal28] G.T. Walker. World weather III. *Mem. Roy. Metoer. Soc.*, 17:97–106, 1928.

[Wal37] G.T. Walker. World weather VI. *Mem. Roy. Meteor. Soc.*, 4:119–139, 1937.

[WB32] G.T. Walker and E.W. Bliss. World weather V. *Mem. Roy. Meteor. Soc.*, 4:53–84, 1932.

[WF96] B. Wang and Z. Fang. Chaotic oscillations of tropical climate: A dynamic system theory for ENSO. *J. Atmos. Sci.*, 53:2786–2802, 1996.

[WW72] J.N. Weber and P.M.J. Woodhead. Temperature dependence of oxygen-18 concentration in reef coral carbonates. *J. Geophys. Res.*, 77:463–473, 1972.

[ZC87] S.E. Zebiak and M.A. Cane. A Model El Niño-Southern oscillation. *Mon. Wea. Rev.*, 115:2262–2278, 1987.

[Zeb84] S.E. Zebiak. *Tropical atmosphere-ocean interaction and the El Niño/Southern Oscillation phenomenon.* PhD thesis, M.I.T., 1984.

[Zeb89] S.E. Zebiak. On the 30-60 day oscillation and the prediction of El Niño. *J. Clim.*, 2:1381–1387, 1989.

[Zhu96] Y. Zhuang. A low-order model of coherent structures in the convective atmospheric surface layer. *Q.J.R. Met. Soc.*, 122:1075–1094, 1996.