Stochastic Parameterisation Schemes Based on Rigorous Limit Theorems

by

Joel David Culina
B.Sc., McMaster University, 2001
M.A., York University, 2002

A Dissertation Submitted in Partial Fulfillment of the
Requirements for the Degree of

Doctor of Philosophy

in the School of Earth and Ocean Sciences

University of Victoria

Stochastic Parameterisation Schemes Based on Rigorous Limit Theorems

by

Joel David Culina
B.Sc., McMaster University, 2001
M.A., York University, 2002

Supervisory Committee

Dr. Adam H. Monahan, Supervisor (School of Earth and Ocean Sciences)

Dr. Stan E. Dosso, Departmental Member (School of Earth and Ocean Sciences)

Dr. Norm McFarlane, Additional Member (Canadian Centre for Climate Modelling and Analysis)

Dr. Boualem Khouider, Outside Member (Department of Mathematics and Statistics)

Supervisory Committee

Dr. Adam H. Monahan, Supervisor (School of Earth and Ocean Sciences)

Dr. Stan E. Dosso, Departmental Member (School of Earth and Ocean Sciences)

Dr. Norm McFarlane, Additional Member (Canadian Centre for Climate Modelling and Analysis)

Dr. Boualem Khouider, Outside Member (Department of Mathematics and Statistics)

Abstract

In this study, theorem-based, generally applicable stochastic parameterisation schemes are developed and applied to a quasi-geostrophic model of extratropical atmospheric low-frequency variability (LFV). Hasselmann's method is developed from limiting theorems for slow-fast systems of ordinary differential equations (ODEs) and applied to this high-dimensional model of intermediate complexity comprised of partial differential equations with complicated boundary conditions. Seamless, efficient algorithms for integrating the parameterised models are developed, which require only minimal changes to the full model algorithm. These algorithms may be readily adapted to a range of climate models of greater complexity in parameterising the effects of fast, sub-grid scale processes on the resolved scales. For comparison, the Majda-Timofeyev-Vanden-Eijnden (MTV) parameterisation method is applied to this model.

The seamless algorithms are first adapted to probe the multiple regime behaviour that characterises the full model LFV. In contrast to the conclusions of a previous study, it is found that the multiple regime behaviour is not the result of a nonlinear interaction between the leading two planetary-scale modes, but rather is the result of interactions among these

two modes and the leading synoptic-scale mode.

The low-dimensional Hasselmann stochastic models perform well in simulating the statistics of the planetary-scale modes. In particular, a model with only one resolved (planetary-scale) mode captures the multiple regime behaviour of the full model. Although a fast-evolving synoptic-scale mode is of primary importance to the multiple regime behaviour, deterministic averaged forcing and not multiplicative noise is responsible for the regime behaviour in this model. The MTV models generate non-Gaussian statistics, but generally do not perform as well in capturing the climate statistics.

Table of Contents

List of Tables

List of Figures

Acknowledgments

I would like to thank my supervisor, Dr. Adam Monahan, for providing the opportunity to work towards a Ph.D., for providing generous funding and resources and for fruitful discussions and skillful editing. I am deeply grateful to Dr. Sergey Kravtsov for his exceptional patience and selflessness in assisting me throughout the program (and for letting me stay at his house and meeting his wonderful family!). I would also like to thank committee member Dr. Boualem Khouider for supervising a directed reading course and for fruitful discussions. Thank you to committee members: Drs. Stan Dosso, Norm McFarlane and David Muraki.

Thank you to John Dorocicz and Ed Wiebe for the technical support.

I wish to thank all of my colleagues and friends, past and present, for their help and care during my Ph.D. study.

Special thanks go to my parents, sister and Emmy for their unwavering love and support.

Finally, thank God.

Chapter 1

Introduction

1.1 Motivation

The climate system is characterised by multiple scales in time and space. All processes comprising this multi-scale system are interdependent, at least to some degree. Indeed, the 'climate system' is used interchangeably with the 'climate-weather system', reflecting the fact that slowly-evolving processes (i.e., climate) may be intrinsically tied to fast-evolving processes (i.e., weather). A prototypical climate-weather system is the atmosphere-ocean system. The fast-evolving atmospheric weather processes can dominate the evolution of the slowly-evolving oceanic climate processes, as with wind-driven ocean currents. In turn, the ocean can impact atmospheric variability both on the weather timescale and the climate timescale. Another example of a climate-weather system, and the system considered in this dissertation, is atmospheric low-frequency variability (LFV), determined by the interaction between slowly-evolving, planetary-scale modes and fast-evolving, synoptic-scale modes. Using a model of LFV from Kravtsov et al. (2005) (hereafter KRG05), the model's timescale separation is exploited in this dissertation by applying rigorously-justified limit theorems to derive reduced equations in the planetary-scale modes alone.

A dimensionally reduced climate model derived from a climate-weather model may seem to be a step backwards, especially in terms of accuracy. However, (numerical) climate models do not explicitly resolve all fast-evolving weather processes, instead parameterising the effect of various unresolved processes (e.g., convection). The rigorous reduction methods of this dissertation can be applied to any slow-fast system that satisfies the assumptions

of the underlying theory, including any slow-fast subsystem within the greater climate system. Thus, these methods have the potential to serve as systematic, generally applicable parameterisation schemes in climate models. Parameterisations are necessary and abundant in climate models for a variety of reasons, including intractable complexity and limited resolution. The methods in this dissertation correct for limited resolution by parameterising the effect of sub-grid scale processes. Even in the highest resolution model, sub-grid scale processes cannot be altogether ignored without severely compromising the model's quality. However, sub-grid scale parameterisations are generally highly simplified representations of reality and are often limited to the modeling of specific processes within specific models; i.e., they are often ad hoc parameterisations. The methods developed in this dissertation are grounded in theory and are generally applicable in parameterising the effect of sub-grid scales on the resolved scales.

In addition to their potential as parameterisation schemes, the reduction methods serve in understanding the underlying climate dynamics of the climate-weather system. In the full climate-weather system, the explicitly resolved fast-evolving processes can obscure the dynamics at longer timescales of interest. Reduction methods simplify the *representation* of the interactions between these scales, by eliminating the degrees of freedom associated with the weather through rigorously-based closure schemes. Furthermore, these reduction methods can reveal which of the long timescale modes are important to the climate dynamics. For example, the dynamics underlying climate variability on decadal timescales is of interest in global warming studies. However, there is potentially a significant effect of monthly and annual-scale weather variability on decadal-scale climate variability. An equation in the decadal-scale modes alone allows for a comparison of the relative importance of the interaction among these modes and the forcings that arise through climate-weather interactions.

1.2 Stochastic climate modeling and atmospheric low-frequency variability (LFV)

Compelled by the fact that climate variability has a predominantly red noise signal (e.g., Pandolfo (1993)), Klaus Hasselmann introduced the concept of stochastic climate modeling

in Hasselmann (1976). Specifically, Hasselmann made the connection between the dominant red noise component of the climate spectrum and integrated white noise, as part of a stochastic differential equation (SDE) of a particular form. He reasoned that the slowly-evolving ocean (climate) integrates the effectively stochastic contributions from the weather, analogous to the classical case of the Brownian motion of a grain of pollen induced by random bombardments of neighbouring water molecules.

It was a profound realisation that a deterministic (although turbulent) climate system can potentially be modeled by a stochastic process. The concept quickly grew to prominence because it had wide implications and great potential, but the potential remains largely untapped. There have been valuable stochastic modeling studies of a range of climate systems. For example, the study of the $100,000$-year glacial cycle spawned the original mathematical idea of stochastic resonance. Originally proposed in Benzi et al. (1981) and Benzi et al. (1982), the marked periodicity in the glacial/interglacial cycle was attributed to the resonance between periodic, astronomical forcing (acting on long timescales) and intrinsic dynamics of the climate system (acting on shorter timescales). Although the evidence weighs against stochastic resonance in this cycle, stochastic resonance found application in many other disciplines and it remains an intriguing possibility behind such climate processes as millennial-scale changes in glacial period ocean circulation (e.g., Ganopolski and Rahmstorf (2002)). On shorter timescales, stochastic climate modeling has been successfully applied to atmosphere-ocean processes, such as El Nino-Southern Oscillation (ENSO) (e.g., Penland and Sardeshmukh (1995)) and the processes determining ocean sea-surface temperature (SST) (e.g., Sura et al. (2006) and Sura and Sardeshmukh (2008)). Another important application of stochastic modeling is in the study of the large-scale atmospheric circulation, particularly atmospheric low-frequency variability (LFV).

Atmospheric low-frequency variability (LFV) is a phenomenon manifest at the planetary scale, characterised by a maximum in power at around wavenumbers 2, 3 and 4 in the wintertime extratropical troposphere, corresponding to timescales of weeks to months (Pandolfo 1993). The concentration of power at these timescales is at first surprising, as frequent, significant changes are most obvious in the high-frequency/synoptic-scale weather

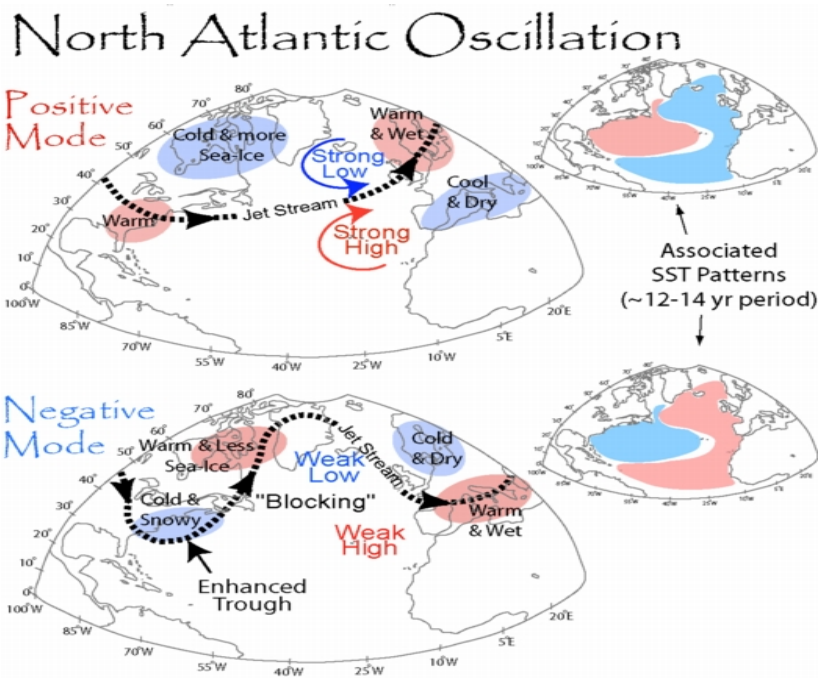

Figure 1.1: Schematic of the two phases of the North Atlantic Oscillation (from www.unh.edu).

systems in the midlatitudes in the wintertime. That there is in fact dominant variability among the lowest wavenumbers is suggested by the presence of a few persistent and recurrent large-scale circulation patterns. Qualitatively, these patterns can be divided into two categories: zonal and blocked flows. High-amplitude blocking is the most visually striking example, evident in large-scale circulation maps by its horseshoe shape (Figure 1.1). Persistence and recurrence of the planetary-scale circulation is illustrated in Figure 1.2, which shows several successive days of streamlines for two different wintertimes. The robustness of the flow pattern within a particular winter illustrates persistence of the flow, and that the same pattern should recur in multiple years illustrates recurrence. Although blocking patterns are manifest on regional scales, simple one-point correlation maps reveal larger structures at work in the dynamics of blocking (Haines 1994), a notable example of which is the North-Atlantic Oscillation (NAO) 'teleconnection' pattern (Wallace and Gutzler 1981).

A more revealing picture of planetary-scale, low-frequency activity is provided by Principal Component Analysis (PCA). PCA performed in Thompson and Wallace (2000) on the Northern Hemisphere sea-level pressure field reveals that the leading mode of variability is

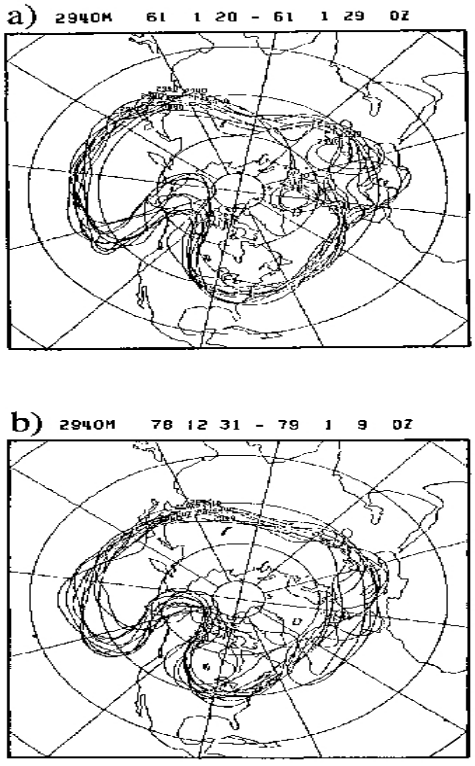

Figure 1.2: Image from Kimoto and Ghil (1993) with caption:"An example of recurrent flow patterns: limited contour maps of a)20-29 January 1961, and b) 31 December 1978-9 January 1979".

a planetary-scale annular mode, which is the apparent counterpart to the Southern Hemisphere annular mode. Called the Arctic Oscillation (AO) (also called the Northern annular mode (NAM)), the AO is deep, equivalent barotropic and largely zonally symmetric (Figure 1.3). It is characterised as well by a meridional dipole in zonal-wind perturbations, similar over the North Atlantic to the NAO, such that the NAO may be considered a regional manifestation of the hemispheric-wide AO (Thompson and Wallace 2001). The KRG05 model used in this dissertation can capture an AO-like mode, which is characterised by meridional shifts in the eddy-driven jet. In addition to the nearly zonally symmetric AO, a wave-train pattern similar to the Pacific-North American (PNA) teleconnection pattern emerges as the second mode of variability in PCA of real atmospheric data (Quadrelli and Wallace 2004). In this latter study, PCA was performed on a wide range of pressure fields and temporal scales, consistently revealing that AO and a PNA-like field dominate the description of long timescale variability.

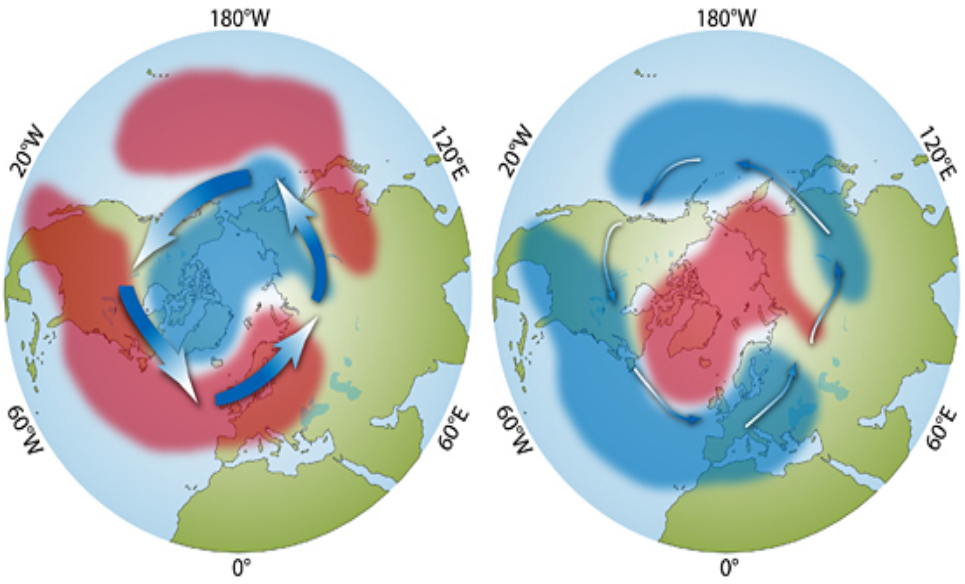

Figure 1.3: Schematic of the two phases of the Arctic Oscillation (from www.whoi.edu). The AO captures a meridional dipole in sea-level pressure; the high-index phase is depicted on the left, characterised by anomalously low sea-level pressure and low surface temperatures in the Arctic, and strong Westerlies.

AO and the PNA-like mode are statistical modes of the atmosphere, that have only relatively recently been identified. It is not surprising then that the first-order dynamics of LFV; i.e., the dynamics underlying the enhanced variability on timescales of weeks to seasons, is not fully understood. In particular, it is not yet determined whether LFV is an effectively linear or nonlinear phenomenon, nor if it is effectively a multi-scale system in the planetary and synoptic scale modes. LFV arises from nonlinear interactions among planetary-scale modes, from feedback between (single or multiple) planetary-scale modes and synoptic-scale modes, or from a single planetary-scale mode (i.e., background climate state) energised by synoptic-scale modes.

Deterministic atmospheric models linearised around a background state have long performed well in simulating certain features of circulation variability, but not the large-scale persistent and recurrent flow configurations. By adding stochastic forcing to represent the effect of fast-evolving processes, these patterns were better simulated. For example, in Whitaker and Sardeshmukh (1998), the statistics of the extratropical synoptic eddies were determined with reasonable accuracy based on the assumption that they are stochastically forced disturbances of a stable flow. The equations of motions are linearised about a mean

state,

$$\frac{d\tilde{x}}{dt} = L\tilde{x}, \tag{1.1}$$

white noise, dW/dt, is added to account for the (assumed) fast-evolving error in linearisation (that is, to account for processes that the linear dynamics cannot), and a damping term is added for stability:

$$\frac{d\tilde{x}}{dt} = (L + D)\tilde{x} + \frac{dW}{dt}. \tag{1.2}$$

If the asymptotically stable linear operator $L+D$ is non-normal (i.e., if it does not commute with its adjoint) and if its eigenvectors are non-orthogonal, then additive noise forcing may induce significant transient growth in the solution of (1.2), which may correspond to the persistent and recurrent streamfunctions states that characterise LFV (Farrell and Ioannou 1995, 1996). In this perspective, the primary role of the synoptic scales in driving LFV is seen to be as a source of energy and sink of enstrophy. There have been subsequent studies demonstrating the success of linear modeling with additive noise, including Winkler et al. (2001) and the rigorously-based study of Franzke et al. (2005) (hereafter FMV05).

Beginning with Egger and Schilling (1983), the first-order dynamical importance of the synoptic scales has been argued for using a range of models and real atmospheric data. Motivated by the observed organisation of the storm-track anomalies, Egger and Schilling (1983) constructed a linearised model driven by red noise fitted to a time-series of observed synoptic-scale flow. Based on the success of the stochastic reduced model, it was concluded that a large fraction of planetary-scale variability is induced by the nonlinear interaction of the synoptic-scale modes. In Branstator (1992) and Branstator (1995), using models linearised about specific low-frequency anomalies, it was found that storm-track organisation is governed by feedback between the anomalies (with respect to the linearised state) and the linearised state itself. This feedback dominates in both the formation and maintenance of organised storm-track anomalies, which in turn sustain their associated low-frequency anomalies.

The underlying assumption of linearised climate models is the existence of a single, dynamically-meaningful, background climate state. In Whitaker and Sardeshmukh (1998)

for example, the background state is the time-mean flow of real atmospheric data, which is predominantly zonal because of the dominance of zonal flow (Westerlies) in the midlatitudes. A compelling alternative paradigm was introduced in Charney and DeVore (1979) in which it was proposed that there are a multiplicity of 'background' climate states, or multiple regimes in the midlatitude flow. Based on the existence of physically-relevant multiple equilibria corresponding to zonal and blocked flows in a truncated barotropic model in the planetary scale modes alone, it was suggested that the large-scale circulation is analogous to a potential well system, with multiple local minima in potential corresponding to multiple quasi-stationary climate states. In this perspective, nonlinearity among planetary-scale modes determines the first-order dynamics of LFV, with the secondary synoptic-scale weather noise acting to (eventually) kick the system out of its quasi-stationary states. A single background state determined as the time average over a long time interval derived from a system with multiple wells may be dynamically meaningless as it does not necessarily correspond to any of the local potential minima.

Following Charney and DeVore (1979), there were many studies that were concerned with identifying these regimes, primarily through statistical analysis. A particularly popular statistical method is to examine (estimates of) probability density functions (PDFs) over a phase space spanned by statistically-determined planetary-scale modes (e.g., Kimoto and Ghil (1993)). Not surprisingly in light of the success of linear models, the statistical significance of the existence of multiple regimes is weak at best. As well, there are fundamental problems with the basis of the phase space, including the practical limitations on the number of dimensions and the potentially poor correspondence between statistical and dynamical modes (e.g., Monahan et al. (2003)).

There have been recent attempts at fitting climate data to nonlinear, multi-scale stochastic processes to determine the significance of nonlinearity at the planetary scale and the level of feedback between scales. For example, in Berner (2005), the probability density evolution equation (i.e., Fokker-Planck equation (FPE)) was fitted to mid-latitude circulation data to determine the nature of the drift and diffusion coefficients of the associated stochastic differential equation (SDE). Statistically significant nonlinearity was identified, as well as

the importance of state-dependent or multiplicative noise (i.e., feedback between scales), but the difficulty and ambiguity in inverting the FPE has been highlighted in Franzke and Majda (2006) and Crommelin and Vanden-Eijnden (2006). In Sura et al. (2005), the FPE is dealt with directly in analysing the non-Gaussian features of real atmospheric data. By assuming that large-scale atmospheric variability satisfies a stationary FPE corresponding to linear dynamics, the good fit of observations to the model led these authors to conclude that nonlinearity among planetary-scale does not play a significant role, leaving multiplicative noise as the cause of the departures from Gaussianity.

A novel approach to uncovering multiple regimes in the large-scale flow is the Hidden Markov Model (HMM) method of Franzke et al. (2008). The HMM method consists of fitting (e.g., by maximum likelihood estimates) climate data to a Markov chain, the states of which potentially correspond to atmospheric regimes. Persistence is the primary determinant of the existence of regimes, as they are identified if transitions among states are improbable. In Franzke et al. (2008), multiple regimes were identified in each model from a hierarchy of models, except in a GCM. A drawback of this method is that the number of states of the Markov chain is not predicted by the method.

By virtue of the uncertainty over the existence of multiple planetary-scale states, it is necessary to move beyond the assumption of a single climate state in analyzing LFV. Indeed, the dynamical picture can change considerably by allowing for the possibility of multiple climate states. For example, the synoptic-scale dynamics of blocking is understood, yet blocking patterns are poorly simulated in conceptual models and in GCMs. The feedback of smaller (synoptic) scales on blocking patterns is "strikingly visual" (Figure 1.1) and is now "universally accepted" (Haines 1994). In particular, a transient trough just upstream of the block is stretched meridionally by the block. Low potential vorticity is advected from the subtropics into the block, reinforcing its anticyclonic motion against dissipation. However, conceptual blocking models involving fast transients, beginning with Shutts (1983) and Vautard and Legras (1988), have not resolved whether these transients play a leading or secondary role in the formation and maintenance of the block. Their role becomes clearer if the contribution of planetary-scale processes to blocking is considered. Haines and Holland

(1998) deduced that different planetary-scale states, such as a blocked flow state, support different local anomalies, such as blocking. Moreover, their analysis suggests that feedback between the high-frequency transient eddies, and not the blocking-scale stationary waves, and the planetary-scale modes is essential to the existence of blocking. By comparison, Branstator (1992) and Branstator (1995) determined that the transients and the stationary waves both interact with the large-scale patterns in maintenance and formation of blocking-scale storm-track formations.

By virtue of its zonal symmetry and barotropic structure, AO is amenable to analysis by simplified zonally and vertically averaged nonlinear equations (e.g., Feldstein and Lee (2002), Lorenz and Hartmann (2001), Lorenz and Hartmann (2002)). In the latter studies, it was determined that the transient eddies, more so than the stationary waves, maintain the high and low index states of AO, which in turn provide conditions which are favourable to these maintaining eddies. The positive feedback between synoptic eddies and zonal wind was deduced from their correlation and the similarity of the redness of synoptic-scale forcing of the zonal wind and of the redness of the zonal wind itself. However, Kravtsov et al. (2003) argued that their results are also consistent with a passive steering of synoptic eddies by planetary-scale waves, without feedback onto these waves. In this study and in KRG05, regimes were determined to arise through nonlinear planetary-scale interactions, with energy provided by the synoptic-scale modes, similar to the mechanism proposed in Charney and DeVore (1979). In Crommelin (2004), multiple regime behaviour, characterised by transitions among zonal and wavy quasi-stationary states, is actually generated without synoptic-scale modes. In models with synoptic-scale modes, connections between different regimes arising from these planetary-scale mechanisms are shown to be close (in phase space) to the corresponding connections in the full, multi-scale model (Legras and Ghil (1985), Crommelin (2003)). The existence of multiple circulation states and the importance of synoptic eddies relative to nonlinear interactions among large-scale modes in generating LFV remain open questions.

1.3 Rigorously-based reduction methods

The problem of LFV is particularly suited to reduction methods (through which effective low-dimensional dynamics are obtained from high-dimensional systems) because there are relatively few dynamically important modes. The low-dimensional nature of LFV is demonstrated in D'Andrea and Vautard (2001), Franzke and Majda (2006) (hereafter FM06) and Kondrashov et al. (2006). All three studies are concerned with the reduction of the three-layer quasi-geostrophic (QG) model of Marshall and Molteni (1993) (the MM93 model), which is a leading conceptual model of LFV. Although the reduction methods differ, the results are consistent among the three studies. In D'Andrea and Vautard (2001), the fast-evolving processes are catalogued and randomized; a rigorously-based method is applied in FM06 (and in this dissertation); and in Kondrashov et al. (2006), the climate data is fitted to multi-level polynomial equations. In the former two studies, out of 1449 empirical orthogonal functions (EOFs) that comprise the full model phase space, the climate statistics were simulated reasonably well with an effective climate equation in the first 10 EOFs. In both cases, the optimal results were achieved with a state-dependent closure. In Kondrashov et al. (2006), a polynomial regression in the first 15 EOFs generated an excellent climate simulation, with a state-independent closure of spatially-coherent and white in time additive noise. Based on all three studies, it would appear that the dynamical modes that give rise to LFV are contained in the space of the first 15 EOFs, and that it is possible to represent the latter 5 of these EOFs in a reduced equation in the first 10 EOFs only, through stochastic and deterministic state-dependent corrections.

Rigorously-based reduction methods involve the application of rigorously justified mathematical theorems, which provide explicit formulae and conditions under which the reduction is always valid. Such guidance is crucial in understanding LFV because of the range of possible dynamics under which LFV may exist. While the rigorously-justified reduction theorems do not impose linearity assumptions or other such assumptions on fundamental aspects of the system, the reduced equations are strictly valid for large scale separation only. However, the (moderate) success of stochastic reduced models of the MM93 model, in which there does not exist a significant 'spectral gap', suggests that this assumption

potentially can be relaxed. Moreover, these methods are amenable to a variety of 'minimal' tuning schemes (such as the scheme of FM06 to be explained and applied in this dissertation). It is hoped that the limiting climate dynamics and the climate dynamics of the climate-weather model do not significantly differ, and that these tuning schemes serve to correct the differences that do exist.

Two rigorously-based methods will be applied in this dissertation: 'Hasselmann's method' and the 'Majda-Timofeyev-Vanden-Eijnden (MTV) method'. As noted above, stochastic climate modeling was introduced in Hasselmann (1976). Although it was a profound realisation in itself that the deterministic climate system could be represented by a stochastic process, it is maybe more remarkable that Hasselmann constructed an SDE primarily from physical considerations that was later mathematically justified as a reduced equation. This SDE is driven by deterministic averaging of the climate forcing over weather realisations and a state-dependent stochastic correction of the averaged term.

Deterministic averaging of a slow-fast system was applied on intuitive grounds in early celestial mechanics applications, but it was not rigorously justified as a limit theorem until in Bogolyubov and Mitroplskii (1961). Applied to a slow-fast stochastic system, Khasminskii (1966) justified the emergence of a stochastic correction to the deterministic averaged solution. By comparison, Hasselmann (1976) proposed a reduced equation of a fully deterministic, cross-coupled slow-fast system with a state-dependent or multiplicative stochastic correction to the averaged forcing. Only recently, in Kifer (2003), and from a less general slow-fast system has Hasselmann's SDE been rigorously justified, as an improvement over an averaged ODE. Yuri Kifer has been particularly active in this field; e.g., Kifer (2001, 2004a, 2005), although a reduced SDE with multiplicative noise has yet to be rigorously derived from the most general deterministic system (although there have been mathematically formal derivations of such a reduced multiplicative SDE in, e.g., Just et al. (2001)). In Arnold et al. (2003) (hereafter AI03), Hasselmann's method is formalised and expanded to include a hierarchy of deterministic and stochastic reduced equations, including the stochastic equation derived in Khasminskii (1966) and the original SDE from Hasselmann (1976). In AI03, Hasselmann's method is applied to a simple toy atmosphere-ocean model. In this

dissertation, Hasselmann's method is adapted to the KRG05 model of LFV by combining the methods and ideas of AI03 and Fatkullin and Vanden-Eijnden (2004).

The 'MTV method' is the second reduction method applied in this dissertation, based on the rigorously-justified theory in Kurtz (1973) and the premise of stochastic climate modeling developed in Hasselmann (1976). Both Hasselmann's SDEs and the MTV SDE in the slow variable alone are derived from a slow-fast system and have a similar form, but generally, their solutions differ. The MTV method, as developed in Majda et al. (2001), Majda et al. (2003) and Majda et al. (2006), consists of a formal means by which to derive a reduced SDE, and in these papers, the method is successfully applied to toy climate models. In FMV05 and FM06, the method is applied to QG models of LFV, and includes several simplifications that are necessary in practical application to complex climate models. In this dissertation, the MTV method is applied to the KRG05 model for comparison with the Hasselmann method.

1.4 Objectives and dissertation outline

The goals of this dissertation are:

- ⋆ to develop Hasselmann's method for application to complex climate models,
- ⋆ to compare the Hasselmann and MTV methods in terms of their ability to simulate the climate statistics of the KRG05 model,
- ⋆ and to better understand the bifurcation dynamics that give rise to jet bimodality in the KRG05 model.

The latter goal extends to contributing to a general understanding of the dynamics underlying LFV. It is also of value to better understand the KRG05 model itself, as it has been used in several subsequent studies, in combination with an ocean model, to study the mechanisms underlying decadal timescale variability (e.g., Kravtsov et al. (2006), Kravtsov et al. (2007), Kravtsov et al. (2008)).

This dissertation is organised as follows. Chapter 2 details the limiting theories underlying the MTV and Hasselmann reduction methods. Chapter 3 develops the numerical schemes necessary for application of Hasselmann's method to complex climate models.

Chapter 4 describes the quasi-geostrophic model of LFV (the KRG05 model) to which the reduction methods are applied, as well as the conclusions reached in KRG05 regarding the model dynamics. The reduced Hasselmann and MTV equations of the KRG05 model are formally derived in Chapter 5. In Chapter 6, the parameters of the reduced equations are computed and preliminary analysis of the dynamics of the unreduced KRG05 model is presented. A statistical comparison of the climate component of the unreduced KRG05 model and the MTV and Hasselmann climate approximations is detailed in Chapter 7 and Chapter 8, respectively. Conclusions are presented in Chapter 9.

Chapter 2

Limit theorems underlying the Hasselmann and MTV methods

This chapter begins with a description of the slow-fast systems to which the Hasselmann and MTV reduction methods apply. This is followed by derivations of the Hasselmann effective ODE and SDEs and the MTV SDE, with a focus on intuition-building (detailed derivations are given in the cited references).

The MTV and Hasselmann methods are systematic methods for reducing the climate-weather/slow-fast system of equations to a system in the climate/slow variable alone. Each method is based on mathematically justified theories describing how the slowly-evolving variable of the unreduced slow-fast system converges to the effective (slow) variable of the reduced model, in the limit of large timescale separation. The asymptotic convergence implies that the effective variable and the unreduced model slow variable are close for sufficiently large separation between slow and fast timescales.

The deterministic and stochastic approximations comprising Hasselmann's method and the stochastic approximation of the MTV method are similarly derived: each approximation involves expectations with respect to the distribution of the weather variable and each stochastic approximation is a diffusion process (i.e., the solution of a SDE). In fact, the only Hasselmann stochastic approximation valid in theory on long timescales is a generalisation of the MTV approximation. If the more stringent assumptions of the MTV method are met, then the MTV approximation is superior to this Hasselmann approximation as a parameterisation and as a tool for understanding the climate dynamics. Among the approximations

comprising Hasselmann's method, the deterministic averaged approximation is the least accurate in a (theoretical, asymptotic) sense, but its simplicity, efficiency, and effectiveness on long timescales in toy models of the atmosphere (Fatkullin and Vanden-Eijnden 2004), might make it a more attractive pararmeterisation than the stochastic reductions.

In applying both strategies, it is first assumed that the (vector-valued) state variable can be divided into two variables with strongly differing timescales (e.g., climate and weather). Following AI03, this is written as a coupled system of ODEs:

$$\begin{aligned} \frac{dx}{dt} &= f(x,y), \quad \text{(slow climate ODE)} \\ \frac{dy}{dt} &= \frac{1}{\epsilon}g(x,y), \quad \text{(fast weather ODE)} \end{aligned} \tag{2.1}$$

where $0 < \epsilon \ll 1$ is the ratio of the y to x timescales (the initial conditions are excluded in all differential equations, except where needed for clarity). The assumption of sufficient timescale separation is key to the accuracy of the deterministic and stochastic approximations.

The Hasselmann approximations in the slow variable alone are derived directly from the slow-fast system, (2.1). In applying the MTV method, it is additionally assumed that each of f and g have an $\mathcal{O}(\epsilon)$ component (such that ϵ now serves as the scale separation between the slow and fast variables x and y and as the scale separation parameter between the slow and fast tendency forcings). Thus, the timescale of interest is the coarse-grained time $\tau = t\epsilon$. For example, a rigorously justified reduction theorem in Kurtz (1973) applies to reduction of:

$$\begin{aligned} \frac{dx}{d\tau} &= \frac{1}{\epsilon}f_0(x,y) + f_1(x,x), \\ \frac{dy}{d\tau} &= \frac{1}{\epsilon}g_0(x,y) + \frac{1}{\epsilon^2}g_1(y,y), \end{aligned} \tag{2.2}$$

to an equation in the climate variable x alone on coarse-grained time. The KRG05 model that will be reduced in this dissertation is assumed to be of the form (2.1), to which Hasselmann's reduction methods are applied, and on coarse-grained time, it is assumed to be of the form (2.2), to which the MTV method is applied.

Although the state variable remains divided into slow (x) and fast (y), there are processes acting on three timescales in (2.2): slow-slow (climate-climate) interactions, slow-fast (climate-weather) interactions and fast-fast (weather-weather) interactions. The MTV decomposition of the tendency forcing is physically plausible because, for example, the self-interaction of the climate modes, f_1, is expected to act on a longer timescale than the climate-weather interaction, f_0. This former, $\mathcal{O}(1)$ forcing (in τ time), also called the bare truncation forcing, is unchanged in the effective climate equation. The weather-weather interaction, g_1, acts alone on the weather timescale, such that in the limit of large timescale separation the weather distribution is independent of the climate variable. However, the effective climate is dependent on the weather distribution through the $\mathcal{O}(\epsilon^{-1})$ forcings f_0 and g_0, which give rise to state-dependent (multiplicative) noise and noise-induced drift (to be discussed).

As a consequence of the climate-independence of the weather distribution, the MTV SDE can be explicitly determined, and for a low-dimensional climate variable can be rapidly integrated. By contrast, integration of the Hasselmann equations requires the sampling of the climate-dependent weather attractor concurrent with the advancement of the effective climate state, which adds considerably to the computational costs. The explicit determination of the effective SDE by the MTV method is also an advantage because the individual forcings can be analyzed to determine their significance to the overall climate dynamics. However, the MTV method may be inappropriate in systems in which there is significant feedback between the climate and weather. For example, the MTV SDE may poorly approximate the planetary-scale modes of the Haines and Holland (1998) model, in which the feedback between the fastest transients and these modes is crucial in the development and maintenance of the model LFV (cf. Chapter 1). The KRG05 model, studied in this dissertation, includes parameter ranges for which weak nonlinearity of planetary-scale dynamics dominates, and also ranges for which the faster-scale processes are dynamically involved in such a way that the distribution of y is not independent of x.

2.1 Hasselmann's method

The starting point of Hasselmann's method is the ODE system (2.1), with sufficiently large time-scale separation, $0 < \epsilon \ll 1$. On the climate timescale (i.e., the $\mathcal{O}(1)$ timescale by which (2.1) is scaled), only the distribution and not the particular path of the weather variable (y) is important to the evolution of the climate variable. Thus, the reduced equations in the climate variable alone are driven by statistical information about the weather. This information for all approximations is obtained through integration of the weather equation:

$$\dot{y}_s^x = g(x, y_s^x), \qquad y_0^x = y_0, \tag{2.3}$$

where the superscript x denotes that the climate variable is held fixed. This equation is the limiting equation of (2.1) on fine-grained time, $s = t/\epsilon$. The limiting dynamics approximates the weather dynamics on the weather timescale, over which the climate variable changes slowly. The distribution of the fast dynamics is determined from (2.3) by the assumption of ergodicity. Essentially, ergodicity of the fast weather dynamics means that y_s^x generates a limiting distribution on its phase space, such that the time and phase space averages are equal. In practice, then, climate averages and covariances with respect to the weather distribution are found as time-averages over iterates of the weather equation (2.3). Since the fast distribution is dependent on the slow variable x in general, a family of distributions μ_x, indexed by the slow variable, is obtained.

In AI03, Hasselmann's method was applied to a 6-variable model, comprised of three ODEs each for an idealised atmosphere and an idealised ocean. The fast weather equation is a 3-variable atmosphere model dependent on a single ocean parameter. In this 'simple' model, the fast equation for a fixed ocean parameter has very complex dynamics as a consequence of the existence of multiple weather attractors. Thus, at a particular state of the ocean, the atmosphere settles on a different attractor depending on its initialisation, and so there are multiple families of distributions $(\mu_x)_i, i = 1, ..., N$. A simple example of such a system is a system with two potential wells with an infinitely high potential barrier between wells, such that each well has its own ergodic dynamics. Depending on the initial

condition, the weather variable remains in a subspace of the full phase space, corresponding to one of the two wells.

For the much higher-dimensional fast dynamics of the KRG05 model, it appears from numerical experiments that the weather distribution is independent of the initial condition, implying the existence of a single attractor and unique ergodicity for each climate variable. Generally, the presence of many fast-evolving modes for fixed low-frequency modes results in a well-mixed turbulence, which is expected to inhibit the formation of multiple invariant distributions for the fast variable (for a given climate state). In any case, a "path-following method" (described in Chapter 3), as in AI03, is applied in this dissertation to guard against unphysical weather attractor hopping in the case of non-unique ergodicity.

The following discussion of the Hasselmann method follows that presented in AI03.

2.1.1 Deterministic Averaging (A)

Assuming ergodicity of the fast dynamics, the fast variable is averaged out of the climate tendency forcing:

$$\lim_{T\to\infty} \frac{1}{T} \int_0^T f(x, y_t^x)\, dt = \int f(x,y)\, \mu_x(dy) \equiv \overline{f}(x). \tag{2.4}$$

The averaged function $\overline{f}$ drives the effective climate ODE, denoted (A):

$$\frac{d\overline{x}}{dt} = \overline{f}(\overline{x}). \tag{2.5}$$

In the limit of infinite time-scale separation the climate variable from the unreduced system (2.1) is equal to the solution of the effective climate ODE; i.e.,

$$\lim_{\epsilon\to 0} x_t^\epsilon(x_0) = \overline{x}_t(x_0), \tag{2.6}$$

on the $\mathcal{O}(1)$ timescale for all initial x_0 and μ_x-almost all initial y_0 (the strongest form of pathwise convergence). The convergence (2.6) holds under very restrictive assumptions (Kifer 2004b). Under more general assumptions, convergence (still pathwise) is justified in L_1, or convergence in the average, over initial conditions (Kifer 2004a).

The (A) approximation will not hold in general on the $\mathcal{O}(\epsilon^{-1})$ timescale. Thus, for the physical case of a small but nonzero scale separation parameter, the climate approximation should track the climate variable from the full climate-weather system, but potentially for a limited period of time only. On long timescales on which the climate statistics are determined, it is not guaranteed that the (A) approximation will approximate these statistics (much less the climate paths). Thus, it may not be possible to determine these statistics from the (A) approximation, either from a single long time-integration or from an ensemble of (short time-integrated) (A) approximations over several initial climate conditions. In the latter case, an accurate integration may just be too short to capture long timescale features of the climate dynamics.

2.1.2 Diffusion approximations (L), (W), (N) and (N+)

As noted above, the (A) approximation can break down on timescales of interest for climate variability. Although the climate-weather system is deterministic, and the (A) approximation is deterministic as well, higher-order corrections to the averaged climate are stochastic. Specifically, the best approximations of the deterministic climate system with large scale separation are given by stochastic differential equations, and are therefore known as diffusion approximations.

The assumption of ergodicity of the fast dynamics is common to both the averaging and diffusion approximations. As a smoothly oscillating fast variable can have ergodic dynamics, it can be averaged out of the slow climate equation, but the difference between the full model climate and (A) approximation trajectories is smooth and cannot be modelled as a stochastic process. Ergodicity alone of the fast dynamics is evidently insufficient to induce the random and rapidly fluctuating dynamics of a noise process in the limit of large timescale separation. The additional necessary ingredient for convergence of the climate-weather system to a climate diffusion process is the sufficiently strong mixing of the fast dynamics. A strongly mixing process is one that rapidly decorrelates, such that the past and future are only weakly dependent (not just weakly correlated). At the limit lies Gaussian white noise, which is 'δ-correlated', and hence has independent values at each individual point in time (Arnold 1974). As $\epsilon \to 0$, the strongly mixing weather dynamics induces a

white noise term in the climate equation.

The role of a mixing condition in the emergence of white noise from a deterministic chaotic process can be illustrated in the context of a simple multi-scale system, to which the Central Limit Theorem (CLT) may be easily applied. Informally, the CLT asserts that the sum of a sufficiently large number of independent variables (with finite variances) has an approximately normal distribution (Arnold 1974). In its simplest form, the CLT states that for a sequence of *i.i.d.* (i.e., independent and identically distributed) random variables (X_i), with $E(X_i) = \mu$ and $Var(X_i) = \sigma^2$:

$$\lim_{N\to\infty} \frac{\frac{1}{N}\sum_{i=1}^{N} X_i - \mu}{\frac{1}{\sqrt{N}}} \sim \mathcal{N}(0, \sigma^2). \tag{2.7}$$

Even in its simplest form, the CLT is remarkable because convergence is to the Gaussian distribution from a sequence with *any* distribution of finite variance, as long as it is *i.i.d.*.

Now consider from Givon and Kupferman (2004):

$$\begin{aligned} \dot{x_t} &= \frac{f(y_t)}{\sqrt{\epsilon}} \\ \dot{y_t} &= \frac{g(y_t)}{\epsilon}, \end{aligned}$$

and the corresponding discretisation:

$$\begin{aligned} \frac{x_{n+1} - x_n}{\epsilon} &= \frac{f(y_n)}{\sqrt{\epsilon}} \\ \frac{y_{n+1} - y_n}{\epsilon} &= \frac{g(y_n)}{\epsilon}. \end{aligned} \tag{2.8}$$

Assume to start that y is a stochastic process, and that $\mathbb{E}f = 0$, where the expectation is with respect to the measure associated with y. The small parameter ϵ serves both as a time scaling and as the time step, such that $(x_n, y_n) = (x(n\epsilon), y(n\epsilon))$. Both processes are fast-evolving on the $\mathcal{O}(1)$ timescale, but it is still a slow-fast system because the timescale of x is $\mathcal{O}(\epsilon^{-1/2})$ slower than that of y. We are interested in the limit to a continuous-time system, as $\epsilon \to 0$.

System (2.8) can be summed to give:

$$x(M) = x(0) + \epsilon \sum_{k=1}^{\frac{M}{\epsilon}} \frac{f(y_k)}{\sqrt{\epsilon}}, \tag{2.9}$$

where M is a fixed, bounded value. The analogy between the above expression and the sum in (2.7) is clear: the limiting equation (2.9) becomes the limiting equation (2.7) by setting $\epsilon = 1/N \to 0$. Although the sequence of $f(y_n)$ is not *i.i.d.*, it does not have to be for the CLT to hold in a more general form (Arnold 1974). In particular, the assumption on the fast variables can be weakened to strongly mixing in justifying the CLT. Assume now that the function g is such that the fast process y_n in (2.8) is a strongly mixing stochastic process, such that successive iterates of y_n lose their dependence quickly. The infinite sum of iterates in (2.9) is unaffected by the weak dependence among finite successions of iterates. Thus, the limiting sum in (2.9) is also Gaussian distributed, and hence, the increments of the continuous-time limiting process are Gaussian distributed. Gaussian distributed increments are a characteristic of the Wiener process, or Brownian Motion. It is justified in Billingsley (1968) that indeed the limiting equation in x is driven by Gaussian white noise according to the SDE:

$$dX(t) = \sigma dW(t),$$

where the diffusion coefficient, σ, is:

$$\sigma^2 = 2 \int_0^\infty \mathbb{E}\left[f(y_t) f(y_0)\right] \, dt. \tag{2.10}$$

As the Wiener process is the integral of Gaussian white noise, the summand, $f_0(y_k)/\sqrt{\epsilon}$, plays the role of Gaussian white noise. An integrated lag-covariance of form similar to (2.10) determines the diffusion coefficients in the MTV and Hasselmann SDEs.

The assumption that y is a stochastic process is not needed to obtain the previous asymptotic result. However, it is difficult to understand how the strongly mixing assumption in particular can be satisfied if the fast dynamics is deterministic, and the initial conditions are precisely known. After all, unlike general random processes, the future of a deterministic

process is entirely predictable given these initial conditions. However, it is possible to formally establish an equivalence between a deterministic process and a stochastic process with a generating partition function, such as a Markov partition. A generating partition is a way of constructing an isomorphism between a deterministic process and a simple stationary stochastic process (Walters 1982). As they are isomorphic systems, to each (important) relation in the stochastic system, including the relations defining ergodicity and mixing, there is an analogous relation in the deterministic system (Walters 1982). If indeed the stochastic process is ergodic and mixing, then the isomorphic deterministic process is said to be ergodic and mixing (Givon and Kupferman 2004).

The implications of the CLT are profound: white noise can potentially be generated from a simple, *deterministic*, low-dimensional (highly nonlinear) climate-weather system. In particular, an infinite-dimensional 'heat bath' is not needed to derive white noise; rather, the large number of degrees of freedom that are needed are introduced by the mixing property and timescale separation.

We will denote by (L) the diffusion approximation obtained through application of the CLT. Through the rigorous derivation of (L) at the pathwise level, first by Khasminskii (1966) for a fast stochastic process uncoupled to the slow process, a clear interpretation emerges that (L) is a first order correction to the *solution* of the averaged ODE. Specifically, the (L) approximation is given by the (A) approximation with stochastic fluctuations driven by the averaged climate superposed as:

$$x_t \overset{D}{\approx} \overline{x}_t + \sqrt{\epsilon}\zeta_t^x, \qquad \zeta_0^x = 0, \tag{2.11}$$

where $\overset{D}{\longrightarrow}$ denotes convergence in distribution or weak convergence. The (L) approximation is determined by a system of two equations: the (A) equation and an equation for the stochastic correction as an Ornstein-Uhlenbeck (O-U) or Gauss-Markov process. An O-U process is the Gaussian-distributed, bounded solution of a linear SDE, describing a damped system driven by white noise fluctuations, used in many contexts such as a model of position and velocity of a particle subject to friction and molecular bombardment (Arnold 1974). In (L), the O-U process is driven by the averaged solution, but the averaged solution is

independent of the O-U process; specifically,

$$\dot{\overline{x}}_t = \overline{f}(\overline{x}_t) \tag{2.12}$$

$$d\zeta_t^x = \nabla_x \overline{f}(\overline{x}_t)\zeta_t^x dt + \sigma(\overline{x}_t)dW_t, \tag{2.13}$$

where W_t is the standard Wiener process, $\nabla_x(\overline{f})$ is the Jacobian of $\overline{f}$ with respect to the slow variable, x, and σ is the non-negative matrix square root of:

$$[\sigma(x)\sigma(x)^T] = \int_{-\infty}^{\infty} cov(f(x, y_t^x), f(x, y_0^x))\ dt. \tag{2.14}$$

Equation 2.13 and all other SDEs are interpreted in the Ito sense (Arnold 1974) unless otherwise noted. On the timescale on which the (A) approximation is valid, it resembles a low-pass filtered version of the (L) approximation.

In the interpretation in terms of the CLT, the O-U process emerges as the $\sqrt{\epsilon}$-scaled error in the (A) approximation of the full model climate variable:

$$\frac{x_t - \overline{x}_t}{\sqrt{\epsilon}} \xrightarrow{D} \zeta_t; \tag{2.15}$$

rearranging (2.15) yields the (L) approximation (2.11).

Although its rigorous derivation is complex, (2.15) is simply a (rigorously-justified) asymptotic expansion of the climate variable in the timescale separation parameter, ϵ. However, it is not a 'typical' asymptotic expansion because the correction is a diffusion process, and also because there are no additional expansion terms. In (2.1), with a periodic fast process y, for example, the slow equation has an infinite number of corrections in ϵ, but with an ergodic and sufficiently mixing chaotic or stochastic fast process, the first-order expansion term is the only one in the expansion (Khasminskii 1966). The linear stochastic (L) model is a rigorously-justified form of the linearised climate models with additive noise of, e.g., Farrell and Ioannou (1995), Penland and Sardeshmukh (1995) and Whitaker and Sardeshmukh (1998). In these models, the time-mean circulation is the climate variable x. As it is constant (by construction), the timescale separation between the weather and

climate modes is assured. While the mean climate is determined by the statistics of the eddies, the trajectory of the weather variable does not feed back on the climate modes. In the more general (L) approximation, the climate state evolves in time and the drift and diffusion terms of the weather respond accordingly (again without feeding back on the climate modes). If the full (L) model is assumed to represent the eddies in a zonally-averaged model, then the averaged flow might represent stationary waves which influence the fast-evolving transient eddies, but which are not in turn influenced by these transient eddies.

A more physically plausible approximation in which the feedback of the weather trajectory on the climate variable is permitted is denoted as (N). In particular, the (N) approximation is a SDE for the climate variable driven by a state-dependent or multiplicative noise term in addition to the averaged forcing:

$$dz_t^\epsilon = \overline{f}(z_t^\epsilon)dt + \sqrt{\epsilon}\sigma(z_t^\epsilon)dW_t, \tag{2.16}$$

where $\overline{f}$ and σ are defined above. In contrast to the (L) approximation, the averaged climate trajectory is no longer resolved itself, as the averaged climate and stochastic weather forcings are now mutually interacting within one SDE. The convergence of the climate variable from the full climate-weather system to the (N) climate diffusion is again in distribution.

A simpler approximation not considered in AI03 (nor in any study of Hasselmann's reduced equations) is one that is seemingly intermediate between the (N) and (L) models, in which the slow-fast feedback is simplified by suppressing state dependence of the diffusion matrix σ, yielding an additive noise SDE. This SDE, which will be called the (W) SDE, only allows for weather to influence climate through the averaging measure and an additional 'one-way feedback' of the fast onto the slow, through the effect of Gaussian white noise on the climate solution. Variability of the weather variable in this approximation is assumed to be independent of the climate state. By comparison, the (L) approximation also allows for feedback through the averaging measure, but the one-way feedback is in the opposite direction, from slow onto fast. Thus, it is expected that the (W) approximation is an improvement over the (L) approximation in modelling the climate because (W) accounts

for more of the effect of the weather processes on the climate. If the (W) approximation is effective in modelling the climate variable, then it would indicate that the corrective (two-way) feedback of multiplicative noise is insignificant.

In theory, the corrective two-way feedback of the (N) model is not an improvement over the Gaussian correction of the (L) model on short timescales. Neither of these models are valid in theory on long timescales, but the (N) model alone can potentially generate a non-Gaussian climate distribution in the case that the solution of the (A) approximation is Gaussian distributed. For example, as discussed in Just et al. (2001), the (A) approximation does not simulate well-hopping in a damped double-well system with small periodic forcing coupled to the (fast) Lorenz system (the reduced models of the double-well system are derived with respect to the distribution of the fast Lorenz system). The (L) approximation cannot induce jumps between wells because it determined as a Gaussian stochastic correction superposed on the existing (A) solution, which is confined to one well. On the other hand, the (N) approximation allows for the interaction of the slow averaged equation and the stochastic correction, permitting the possibility of well-hopping. Since (N) allows for such a feedback, it is not surprising that results from AI03 suggest that (N) outperforms (L) on long timescales. The (W) approximation also permits the possibility of well-hopping, but the exit timescale from a regime, for example, may be dependent on the interaction between high-frequency climate fluctuations represented by the noise term and the climate itself.

In principle, (N) is not strictly valid on the long $\mathcal{O}(\epsilon^{-1})$ timescale; an additional corrective term given by a (deterministic) noise-induced drift is needed. The noise-induced drift is the manifestation of finite temporal correlations that remain in the limit of large scale separation (Penland 2003). In Sura and Newman (2008), the emergence of noise-induced drift is heuristically illustrated with a simplified equation relating rapid-wind variability and air-sea thermal coupling:

$$\frac{d\delta T}{dt} = -\delta T(|\overline{\mathbf{U}}| + |\mathbf{U}|'). \tag{2.17}$$

Since the anomalous wind speed $|\mathbf{U}|' \equiv \left(|\mathbf{U}| - \overline{|\mathbf{U}|}\right) \in [-\overline{|\mathbf{U}|}, \infty)$ varies rapidly relative to

mean wind speed $\overline{|\mathbf{U}|}$, it is replaced by white noise in the equation for air-sea temperature difference, δT. However, white noise is $\delta-$correlated, whereas, in reality, there are small intervals over which $|\mathbf{U}|'$ is effectively constant. To determine the mean effect of this discrepancy, assume that, at a point in time at which $\delta T = \delta T_0$, the wind anomaly forcing has density $p(|\mathbf{U}|')$. The ensemble mean trajectory of δT for positive anomalous wind speed, $< \delta T_+ >$, and for negative anomalous wind speed, $< \delta T_- >$, are:

$$\begin{aligned} < \delta T_+ > &= \delta T_0 \int_0^\infty e^{-(\overline{|U|}+|U|')t} p \; d|U|' \\ < \delta T_- > &= \delta T_0 \int_{-\infty}^0 e^{-(\overline{|U|}+|U|')t} p \; d|U|' \\ &= \delta T_0 \int_0^\infty e^{-(\overline{|U|}-|U|')t} p \; d|U|'. \end{aligned} \tag{2.18}$$

Thus, the mean response to the anomalous wind forcing,

$$< \delta T >= \delta T_0 e^{-\overline{|U|}t} \int_0^\infty \cosh(-|U|'t) p \; d|U|', \tag{2.19}$$

is non-zero. The noise-induced drift arises because the exponential decay of the air-sea temperature difference is different depending on whether the anomalous wind speed forcing is positive or negative; that is, "the effects of equal but opposite wind anomalies do not cancel each other" (Sura and Newman 2008).

This enhancement of (N) that includes noise-induced drift, (N+), is formally derived for the fully-coupled system at the level of the Fokker-Planck equation (FPE) in Just et al. (2001):

$$dz_t^\epsilon = (\overline{f}(z_t^\epsilon) + \epsilon D(z_t^\epsilon))dt + \sqrt{\epsilon}\sigma(z_t^\epsilon)dW_t, \tag{2.20}$$

where the noise-induced drift is

$$D(x) \equiv \int_0^\infty \mathbb{E}(\nabla_x f(x, y_t^x) - \nabla_x \overline{f}(x))(f(x, y_0^x) - \overline{f}(x)) \; dt. \tag{2.21}$$

This term can be interpreted as corresponding to the correction term arising from interpreting (2.16) as a Stratonovich SDE, as is natural when white noise arises as an approximation

to stochastic processes with a finite autocorrelation time. The first Jacobian can be expanded using the chain rule: $\nabla_x f(x, y_t^x) = \nabla_x f(x, y_t) + \nabla_y f(x, y_t) \nabla_x y_t^x$. Unfortunately, the particular factor $\nabla_x y_t^x$ is very difficult to evaluate in complex models with many degrees of freedom (such as the KRG05 model). The time evolution of $\nabla_x y_t^x$ can be simply determined by applying the derivative ∇_x to the unresolved equations:

$$\dot{\nabla_x y_t^x} = \nabla_x g(x, y_t) + \nabla_y g(x, y_t) \nabla_x y_t^x. \tag{2.22}$$

For a suitable initial condition, $\nabla_x y_0^x$, (2.22) can be integrated to give the time evolution of $\nabla_x y_t^x$, as needed in calculation of D. However, the derivative of the fast modes at the initial time, $\nabla_x y_0^x$, must still be evaluated. As will be discussed in the next section, the factor $\nabla_x y_t^x$ vanishes in the MTV SDE because of the assumption that the unresolved weather is independent of the resolved climate.

This discussion has characterised a hierarchy of Hasselmann approximations, all of which are driven by the conditional distribution of the climate for fixed weather, $\mu_x(y)$. To zeroth order in the scale separation parameter ϵ, the weather variable averages out of the climate equation to yield an effective deterministic (A) equation in the climate variable alone. For a chaotic or stochastic fast mixing (e.g., exponentially mixing) weather variable, noise becomes important on long timescales, as ϵ-scaled corrections. The linear diffusion correction of the (L) approximation is the only true correction in the sense that the first-order (A) dynamics is unaffected by this linear diffusion. In particular, the (L) approximation is comprised of the superposition of the (A) approximation and a Gaussian distributed diffusion driven by the (A) approximation. By contrast, Hasselmann's nonlinear diffusion (N) approximation includes a state-dependent noise term. Although formally (N) and (L) are equally good approximations in the limit of infinite timescale separation, numerical studies, including AI03, show that (N) captures long timescale dynamics better than does (L) (and (A)). A natural (but not rigorously justified) simplification of (N) is to assume that the diffusion coefficient is state-independent, yielding the additive noise (W) approximation. The (W) approximation would be successful if the climate-weather feedback is primarily captured by the (conditional) averaged forcing. Finally, the (N+) ap-

proximation is the only approximation that is rigorously valid on the long $\mathcal{O}(\epsilon^{-1})$ timescale. The (N+) SDE includes the same stochastic correction as the (N) SDE, but has an additional noise-induced drift term. Analytic expressions exist for Hasselmann's models at all levels of approximation.

Approximation	Form
(A)	$\dot{\overline{x}} = \overline{f} = \int f(x,y)\, d\mu_x(y)$
(L)	$z = \overline{x} + \sqrt{\epsilon}\zeta(\overline{x})$
(W)	$\dot{z} = \overline{f}(z) + \sqrt{\epsilon}\sigma\dot{W}$
(N)	$\dot{z} = \overline{f}(z) + \sqrt{\epsilon}\sigma(z)\dot{W}$
(N+)	$\dot{z} = \overline{f}(z) + \epsilon D + \sqrt{\epsilon}\sigma(z)\dot{W}$

Table 2.1: Hasselmann's reduced equations

Additional expansion terms in ϵ in the (L) approximation are not possible, at least in theory (although the half-order correction in ϵ can be replaced by higher-order corrections, up to first order (Kifer 2001)). Just et al. (2001) speculate on the nature of additional expansion terms in the (N+) model. For $\epsilon < 0.1$, the (N+) approximation captures the statistics of the double-well dynamics they consider. For a timescale separation of less than an order of magnitude, it is determined that both higher-order lag-correlations of the fast dynamics and non-Gaussian noise are needed to correct the stochastic approximation, requiring a much more sophisticated model. The range $\epsilon > 0.1$ is the range of timescale separation in that of both the KRG05 model of LFV to be studied in this dissertation and in many other climate-weather systems. This is not to suggest that there is no hope for Hasselmann's stochastic models, as the exaggerated potential barrier between wells of the double-well model and low-dimensionality of the fast Lorenz dynamics are not characteristic of realistic geophysical models. Based on mathematical considerations, the (N+) approximation should be superior to all others. Ultimately, (N) may have more practical significance than (N+) because without the need to evaluate the Jacobian derivatives, (N) is simple to implement and more cost-effective than (N+). The (N+) SDE can also in principle be simulated using numerical methods for Stratonovich SDEs, but the derivatives associated with these methods are costly to compute in high-dimensional models.

2.1.3 Illustrative example of application of Hasselmann's method

To illustrate averaging and the diffusion approximations, a simple SDE is considered. First consider the ODE:

$$\begin{aligned} \dot{x} &= -y^3 + \sin(\pi t) + \cos(\sqrt{2}\pi t), \\ \dot{y} &= -\epsilon^{-1}(y - x). \end{aligned} \tag{2.23}$$

For $\epsilon \ll 1$, the solution of the fast equation adjusts rapidly to the current value of x, such that

$$y = x + \mathcal{O}(\epsilon); \tag{2.24}$$

that is, the fast variable is "slaved" to the slow variable (Gardiner 1996). The limiting equation as $\epsilon \to 0$ (Figure 2.1) is then:

$$\dot{x} = -x^3 + \sin(\pi t) + \cos(\sqrt{2}\pi t). \tag{2.25}$$

Now consider the SDE:

$$\begin{aligned} \dot{x} &= -y^3 + \sin(\pi t) + \cos(\sqrt{2}\pi t), \\ \dot{y} &= -\epsilon^{-1}(y - x) + \epsilon^{-1/2}\dot{W}. \end{aligned} \tag{2.26}$$

With the addition of white noise, the fast equation is now an O-U process for which the equilibrium conditional probability density is:

$$\rho_x(y) = \frac{e^{-\frac{(y-x)^2}{2}}}{\sqrt{2\pi}}. \tag{2.27}$$

Thus, instead of tending to a fixed point, y tends to a non-trivial attractor, on which it statistically equilibrates (for sufficiently large scale separation). The averaged equation is

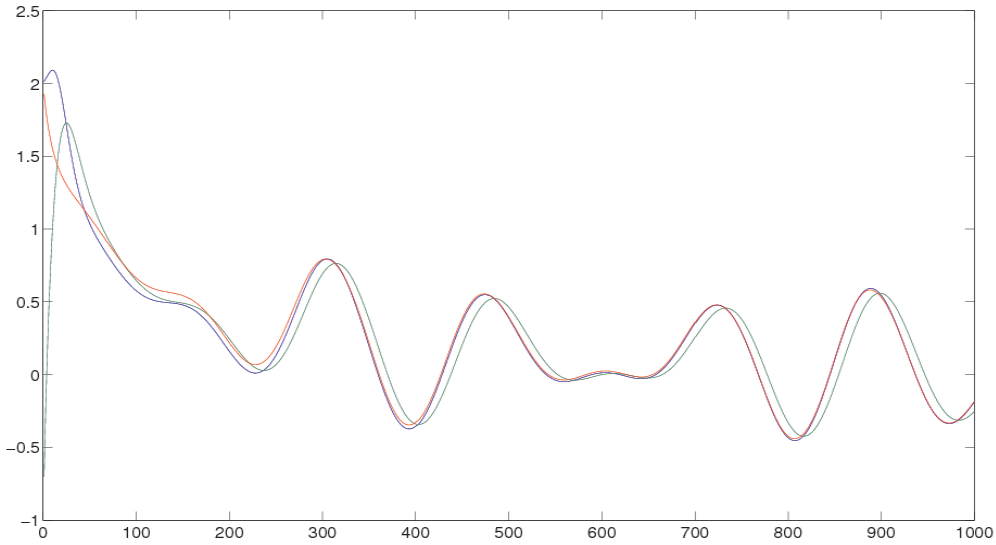

Figure 2.1: The solution of the deterministic slow-fast system, (2.23) (x blue and y green) with $\epsilon = 0.1$, and the solution of the averaged equation, (2.25) in red.

analytically obtained by integrating the slow tendency forcing with respect to this density:

$$\begin{aligned} \dot{\overline{x}} &= \int \left[-y^3 + \sin(\pi t) + \cos(\sqrt{2}\pi t)\right] \rho_x(y)\, dy \\ &= -\overline{x}^3 - \frac{3}{2}\overline{x} + \sin(\pi t) + \cos(\sqrt{2}\pi t). \end{aligned} \tag{2.28}$$

Notice the additional tendency forcing term, $-3\overline{x}/2$. In the limit of large timescale separation, the slow solution of (2.26) converges *pathwise* to the solution of the averaged equation (Figure 2.2). Note that pathwise convergence implies convergence in distribution.

The (L) approximation is given by a corrective Gauss-Markov process added to the solution of (2.28): $\overline{x_t} + \sqrt{\epsilon}\zeta_t$. A drift factor, $\partial\overline{f}/\partial x$, and a diffusion coefficient, σ, drive the linear SDE:

$$\dot{\zeta} = (-3\overline{x}^2 - \frac{3}{2})\zeta + \sigma(\overline{x})\dot{W},$$

whose solution is the Gauss-Markov process, ζ. The derivative can be simply calculated because $\overline{f}$ is known explicitly, but the diffusion coefficient is non-stationary. The diffusion

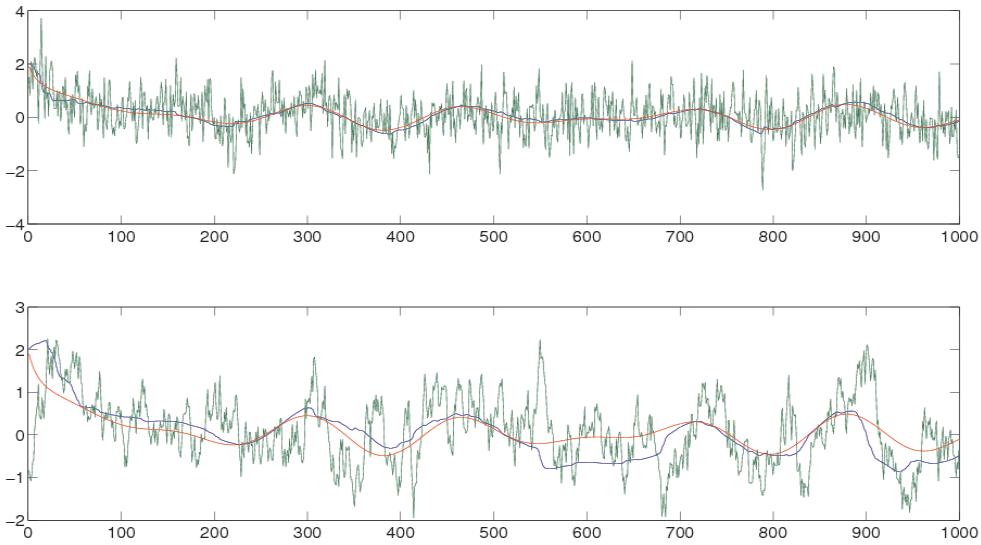

Figure 2.2: The solution of the stochastic slow-fast system, (2.2) (x blue and y green), with $\epsilon = 0.01$ (top) and $\epsilon = 0.1$ (bottom), and the solution of the averaged equation, (2.28) (red).

coefficient is the integrated lag-covariance of the full model slow tendency:

$$\begin{aligned}\sigma(x)^2 &= \lim_{T\to\infty} 2\int_0^T Cov(-(y_t^x)^3 + \sin(\pi t) + \cos(\sqrt{2}\pi t), -(y_0^x)^3 + \sin(\pi 0) + \cos(\sqrt{2}\pi 0))\, dt \\ &= \lim_{T\to\infty} 2\int_0^T Cov(-(y_t^x)^3 + \sin(\pi t) + \cos(\sqrt{2}\pi t), -(y_0^x)^3 + 1)\, dt\end{aligned}$$

where in calculating the covariance, the expectation is taken with respect to the joint distribution of y_t^x and y_0^x.

The (N) approximation is simpler to compute, as it is determined by just $\overline{f}$ and σ, and does not require calculation of the drift factor, $\partial\overline{f}/\partial x$. The diffusion coefficient σ has the same formula as above, except it is not dependent on the (A) solution, $\overline{x}$, but rather it feeds back onto the effective slow solution. The (N+) approximation includes a noise-induced drift, calculated from a similar equation to that for σ:

$$D(x) = \lim_{t\to\infty}\int_0^T \mathbb{E}(\frac{\partial}{\partial x}(-(y_t^x)^3 + sin(\pi t) + cos(\sqrt{2}\pi t)) - 3x^2 - \frac{3}{2}x)(-(y_t^x)^3 + 1 - 3x^2 - \frac{3}{2})\, dt.$$

The derivative $\partial f(y_t^x)/\partial x$ is nonzero because the unresolved fast variable y_t^x is dependent on the resolved variable x. A comparison of the full model with the Hasselmann (A) and (N) models, integrated with numerical schemes developed specifically for the Hasselmann reduced approximations, is presented in the next chapter.

2.2 The MTV method

The MTV method, like Hasselmann's method, applies to systems in which the weather statistically equilibrates on a timescale on which the climate variable does not change significantly. This method further assumes that the weather distribution is independent of climate; i.e., $\mu_x = \mu$. In fact, this assumption follows from the assumed form of the climate-weather system. In (2.2), g_1 alone acts on the weather timescale, such that the weather statistics are determined from the climate-independent equation:

$$\dot{c} = g_1(c, c), \tag{2.29}$$

where the variable c is used in place of y^x because the weather is no longer climate dependent (hence, there is no x to fix). As a consequence of the climate-independence of the weather distribution, the functional form of the MTV SDE can be explicitly determined, and for a low-dimensional climate can be rapidly integrated.

The MTV SDE has the same mathematical form as the (N+) SDE, with bare-truncation terms, noise-induced drifts and multiplicative noise. From a pathwise perspective, the emergence of multiplicative noise in (2.2) is intuitive from consideration of the CLT. Recall that the CLT was applied to (2.8) in justifying the emergence of noise in the limit of infinite timescale separation. As (2.8) is a simpler form of (2.2), (2.8) can also be used to predict the form of the effective climate equation of the latter climate-weather equation. With a slow tendency forcing given by $f/\epsilon = f(y)/\epsilon$, (2.8) yielded additive noise as $\epsilon \longrightarrow 0$. Since the $\mathcal{O}(\epsilon^{-1})$ slow tendency forcing of (2.2) is dependent on x as $f_0/\epsilon = f_0(x, y)/\epsilon$, it is expected that state-dependent multiplicative noise emerges. Considering the example system from Section 2.1 relating rapid wind variability and air-sea thermal coupling, it is expected that a noise-induced drift also emerges. Note that in the context of this example

system that the MTV method assumes that the Gaussian PDF, p, of the anomalous wind forcing is independent of the air-sea temperature difference, δT.

Although path-wise descriptions are useful for developing intuition, rigorous reduction of (2.2) to an effective equation in the climate variable alone has been carried out only at the level of the probability density evolution equation. Moreover, formal reduction of (2.2) appears to be possible only at this level (Gardiner 1996). However, this is not a shortcoming on long timescales on which path-wise predictability is not relevant. Moreover, formal perturbation analysis applied to the density evolution equation is a relatively simple means of deriving a reduced climate equation and some of the conditions under which the reduction holds. Of course, it is not necessary to apply perturbation methods to derive the reduced equation of (2.2) because the effective climate equation has been rigorously derived in Kurtz (1973). However, perturbation methods can be readily applied to a variety of systems for which there are no rigorous reduction theories (e.g., Majda et al. (2003)).

The 'MTV method' essentially consists of identifying a multi-scale system of a particular form (e.g., (2.2)) appropriate for modelling the climate system under consideration and applying perturbation methods to derive the reduced system and the conditions under which it holds. In earlier work on the MTV method, for example Majda et al. (2001, 2002, 2003), it was assumed that the limiting fast weather variable, c, is a simple stationary stochastic process. This assumption was relaxed in Majda et al. (2006), in which it was formally justified that a fully deterministic system can be reduced to an effective SDE. For illustration, the MTV method is applied below to a simpler system in which y is a diffusion process independent of x (thus, $y = c$ in this example). The MTV method is applied to a system of the form (2.2) in FMV05, FM06, and in this dissertation.

As an illustration of how the MTV results are obtained, consider the special case of the system from Givon et al. (2004):

$$\begin{aligned} \frac{dx}{d\tau} &= \frac{1}{\epsilon} f_0(x,y) + f_1(x,x) \\ \frac{dy}{d\tau} &= \frac{1}{\epsilon^2} g_0(y) + \frac{1}{\epsilon} \beta_0(y) \frac{dW}{d\tau}, \end{aligned} \tag{2.30}$$

where $dW/d\tau$ is white noise. In transforming from (2.1) to (2.30), time is rescaled as $\tau = t\epsilon$

(i.e., time is coarse-grained) and f is decomposed into two forcing functions with strongly differing timescales. The associated FPE for the density $\rho = \rho(x, y, \tau, \epsilon)$ is

$$\frac{\partial \rho}{\partial \tau} = \frac{1}{\epsilon^2} L_1 \rho + \frac{1}{\epsilon} L_2 \rho + L_3 \rho, \tag{2.31}$$

where

$$\begin{aligned} L_1 \rho &= -\frac{\partial}{\partial y} \cdot (g_0 \rho) + \frac{1}{2} \frac{\partial}{\partial y} \cdot \left[\frac{\partial}{\partial y} \cdot (\beta_0 \beta_0^T \rho) \right] && (2.32) \\ L_2 \rho &= -\frac{\partial}{\partial x} \cdot (f_0 \rho) && (2.33) \\ L_3 \rho &= -\frac{\partial}{\partial x} \cdot (f_1 \rho). && (2.34) \end{aligned}$$

The goal of reduction is to find the effective FPE in the slow variable only,

$$\frac{\partial \overline{\rho}}{\partial \tau} = L \overline{\rho}, \quad \overline{\rho} = \overline{\rho}(x, \tau), \tag{2.35}$$

from which can be derived an effective SDE.

Equation (2.35) is formally derived through a perturbation expansion of ρ:

$$\rho(x, y, \tau, \epsilon) = \rho_0(x, y, \tau) + \epsilon \rho_1(x, y, \tau) + \epsilon^2 \rho_2(x, y, \tau) + ... \tag{2.36}$$

The expansion is inserted into (2.31) and like terms in ϵ are collected, which gives a series of equations:

$$\begin{aligned} L_1 \rho_0 &= 0, && (2.37) \\ L_1 \rho_1 &= -L_2 \rho_0, && (2.38) \\ L_1 \rho_2 &= -\frac{\partial \rho_0}{\partial \tau} - L_3 \rho_0 - L_2 \rho_1, && (2.39) \\ &\vdots \end{aligned}$$

The effective equation in the climate variable alone and some of the conditions under which it holds are obtained by solving (2.37)-(2.39).

First, the assumption of the existence of a limiting invariant weather density $\rho_\infty(y)$ implies that $L_1\rho_\infty(y) = 0$. Thus, it follows that $L_1\rho_\infty(y)\overline{\rho}(x,\tau) = 0$ and hence a solution ρ_0 of (2.37) is:

$$\rho_0(x, y, \tau) = \rho_\infty(y)\overline{\rho}(x, \tau), \tag{2.40}$$

where $\overline{\rho}$ is the (as yet undetermined) density of the effective equation. Obviously, the effective climate SDE cannot be derived from this stationary FPE, $\overline{\rho}$, is derived, by examining the higher-order equations in ϵ.

It suffices to solve the second equation (2.38) that:

$$< f_0 > \equiv 0, \tag{2.41}$$

where $< \cdot >$ is the expectation with respect to the limiting invariant weather density, $\rho_\infty(y)$. That is, the leading-order x-dynamics averages to zero. This condition does not imply that there can be no drift in the climate equations, as there are bare-truncation and noise-induced drifts . Rather, (2.41) ensures that there are no $\mathcal{O}\left(\epsilon^{-1}\right)$ terms in the effective climate equation, or "fast-wave effects" (Majda et al. (2001)). The condition (2.41) was relaxed in Majda et al. (2001), leading to a significantly more complex effective climate equation. The climate equations of FMV05 and FM06 are derived with the condition (2.41).

The two leading order equations alone do not yield a prognostic equation. However, applying the previous conditions to (2.39), the effective FPE is derived for the leading-order density:

$$\frac{\partial\overline{\rho}}{\partial\tau} = - < L_2 L_1^{-1} L_2\overline{\rho} > + < L_3\overline{\rho} > . \tag{2.42}$$

Convergence is in distribution as the reduction is at the level of the density evolution equation. The MTV model is obtained as the SDE for which 2.42 is the FPE.

Perturbation expansion is the simplest formal justification that unambiguously yields the effective SDE, and some of the conditions of its application. However, it can yield inconsistent conditions at higher orders in ϵ (Just et al. 2001). Standard projection techniques for SDEs from statistical mechanics are less simple than perturbation expansions,

but they are formally exact when reduction is possible. A formally exact reduction of (2.30) to (2.42) is possible by *defining* the projection operator:

$$\mathcal{P}\rho(x,y,\tau) \equiv \rho_\infty(y)\overline{\rho}(x,\tau). \tag{2.43}$$

The projection is applied to the associated FPE, and the equation is formally reduced to one in $\overline{\rho}$ only.

There is a limited class of climate-weather equations that yield an analytic expression for the effective climate FPEs (2.42). The weather dynamics must have particularly 'nice' properties to ensure that analytic expressions can be obtained for both the inverse of its Fokker-Planck operator, L_1^{-1}, and the invariant weather distribution. An example of a such a weather process is an Gauss-Markov process with $g_0 = y$ and β_0 constant (motivating its selection in deriving effective equations in Majda et al. (2001), Majda et al. (2002) and Majda et al. (2003)). However, an analytic expression for the FPE is not needed to derive an explicit effective SDE. Instead, the assumption of ergodicity of the fast dynamics can be exploited to numerically derive from (2.42) the drift and diffusion operators of the associated SDE. In particular, the drift $F(x) = F_0(x) + F_1(x)$ is:

$$F_0(x) = \int_0^\infty \left[\lim_{T\to\infty}\frac{1}{T}\int_0^T \nabla_x f_0(x,y_s) f_0(x,y_{s+t})\, ds\right]\, dt, \tag{2.44}$$

$$F_1(x) = \lim_{T\to\infty}\frac{1}{T}\int_0^T f_1(x,x)\, dt = f_1(x,x), \tag{2.45}$$

and the diffusion σ is determined from

$$\frac{1}{2}[\sigma(x)\sigma(x)^T] = \int_0^\infty \left[\lim_{T\to\infty}\frac{1}{T}\int_0^T f_0(x,y_s) f_0^T(x,y_{s+t})\, ds\right]\, dt. \tag{2.46}$$

Note the similarity to the drift and diffusion formulae of Hasselmann's (N+) SDE. Due to the independence of x and y, the troublesome term $\nabla_x y^x$ in the (N+) SDE is eliminated in the MTV SDE.

These expressions for drift and diffusion are derived from a simplified form of (2.2). The

only difference in generalisation is to the noise-induced drift that becomes:

$$F_0(x) = \int_0^\infty \left\{ \lim_{T\to\infty} \frac{1}{T} \int_0^T \nabla_{x,c} f_0(x, c_{t+s}) \left(f_0(x, c_t), g_0(x, c_t) \right) \, dt \right\} \, ds, \tag{2.47}$$

where c is the solution of the general limiting weather equation (2.29). The tendency forcing g_0 and derivatives with respect to the fast modes arise because g_0 has the same $\mathcal{O}(\epsilon^{-1})$ timescale as the slow mode tendency forcing f_0. Thus, in the general case, the weather influences the climate through its invariant distribution and through g_0, despite the independence of the climate and weather processes.

Although the formulae for the (N+) and MTV SDEs are similar, only for the latter SDE can explicit functional forms be determined for complex systems. The key to the tractability of the MTV formulae is the independence of the climate and weather. In practice, additional assumptions beyond those required by theory are needed in order to derive the explicit equation. A particularly strong assumption is required in FMV05 and FM06 that the climate-weather interaction functions f_1 and g_0 can be written as tensor functions of state variables; i.e., that f_1 and g_0 are multilinear functions. As quadratic nonlinearities, through advection, are the dominant (intrinsic) nonlinearities at large scales in the mid-latitude atmosphere, this effectively bilinear system may be reducable into an explicit SDE in the low-frequency modes alone. Indeed, the Marshall and Molteni model (1993), to which the MTV method is applied in FM06, is determined by a bilinear system. However, as we shall see, the KRG05 model includes a non-quadratic nonlinearity (of secondary importance) which must be approximated by a multilinear function in order to derive an explicit effective climate equation. The implementation of the MTV method in the KRG05 model is detailed in Chapter 5.

2.2.1 Illustrative examples of application of MTV method

Consider the example introduced in Chapter 1 of the symbiotic relationship between low-frequency and storm-track anomalies. If this system is assumed to be of the form (2.2), then on coarse-grained time the storm-track anomalies are forced by fast, $\mathcal{O}(\epsilon^{-1})$ interactions between the low-frequency modes and storm-track anomalies, and the fastest, $\mathcal{O}(\epsilon^{-2})$ self-

interactions among storm-track anomalies, the latter of which determines the invariant distribution of the storm-track anomalies. The effective climate feedback, which replaces the degrees of freedom associated with the synoptic-scale modes, is determined by the first two moments of the $\mathcal{O}(\epsilon^{-1})$ interactions with respect to the invariant weather distribution. In the effective equation, the low-frequency modes may be maintained by positive feedback between slow and fast processes through the multiplicative noise term, simulating the re-reinforcement of these modes by the storm-track anomalies. Nevertheless, the assumption of an invariant distribution is clearly not representative of the situation in which the statistics of storm-track anomalies change significantly depending on the low-frequency anomaly.

Now consider the example from Givon et al. (2004):

$$\begin{aligned} \dot{x} &= x - x^3 + \frac{4}{90\epsilon} y_2, \\ \dot{y_1} &= \frac{10}{\epsilon^2}(y_2 - y_1), \\ \dot{y_2} &= \frac{1}{\epsilon^2}(28y_1 - y_2 - y_1 y_3), \\ \dot{y_3} &= \frac{1}{\epsilon^2}\left(y_1 y_2 - \frac{8}{3} y_3\right). \end{aligned} \tag{2.48}$$

The slow dynamics is characterised by a double well potential, and the fast vector y solves the Lorenz equations for parameter settings at which the dynamics is chaotic.

The reduced equation in x only is found by applying the MTV method. Perturbation analysis can be applied to (2.48) to formally derive the effective SDE. However, the Lorenz-driven system has the same form as (2.30), from which the general reduced equation has already been derived. The fast dynamics is here deterministic, but this does not change the effective equation (Majda et al. 2006), as the Lorenz dynamics at the given parameter settings is chaotic and well-mixing. The reduced equation is particularly simple because the operators are all just scalars. The drift and diffusion formulae are given by (2.45) and (2.46), with $f_1 = f_1(x)$ and $f_0 = f_0(y)$. The bare truncation term remains as is in the reduced equation and the noise-induced drift is zero because $\nabla_x f_0(y) = 0$. In (2.48), the diffusion coefficient is a constant, but it cannot be analytically solved because of the

complex measure associated with the Lorenz equations. The reduced equation reads:

$$\frac{dX}{dt} = X - X^3 + \sigma \frac{dW}{dt} \tag{2.49}$$

with the diffusion constant solving $\sigma^2 = 2(4/90)^2 \int_0^\infty < y_2(t) y_2(0) > dt$. For $\epsilon^2 = 0.001$, the bimodal distribution of the full model is almost exactly reproduced by the reduced equation (Givon et al. 2004).

As the assumptions which give rise to the MTV SDE are relaxed in deriving the (N+) SDE, the same effective equation (2.49) is derived in the Hasselmann framework. On the timescale of (2.1), from which the Hasselmann reduced equations are derived, it is assumed that the forcing $x - x^3$ is $\mathcal{O}(\epsilon)$. On the long $\mathcal{O}(\epsilon^{-1})$ timescale on which only the (N+) equation (among Hasselmann's equations) is strictly valid, the (N+) equation reduces to the additive noise (W) equation (2.49).

Chapter 3

Numerical schemes for Hasselmann's method

If the (A), (W), (L), (N) and (N+) equations are to be applied to realistic geophysical models exactly as justified in theory, then they must be integrated with a degree of brute computational force. In particular, the distribution of the (high-dimensional) weather component conditioned on the climate cannot be a priori (or offline) determined, as in the highly simplified model considered in AI03. In this dissertation, a 'superparameterisation' (or online) scheme is implemented, which will be called the 'cumulative averaging method'. The use of the term superparameterisation is in analogy with its standard use in the climate literature, in which it was recently developed (Grabowski 2001) to model directly mesoscale and smaller-scale processes, particularly cloud processes (that dynamically influence processes on the planetary scale). Instead of a traditional cloud parameterisation scheme, smaller-scale processes are sampled from a cloud-resolving model and this information is coupled to a GCM. Similarly, the cumulative averaging method entails sampling of the weather attractor after each time step of the climate equation, to be used towards determination of the conditional weather distribution. If the timescale separation is sufficiently large, then only a sparse sampling is needed at each time step to derive an efficient and accurate climate approximation. The 'seamless' version of this method is also outlined, which is needed for the practical application of this method to the high-dimensional KRG05 model.

The cumulative averaging method is described below in the context of the representative climate-weather model, (2.1). Only one timescale, the climate timescale, with climate variable t is considered. A time step of the effective climate equation, or a macrotime step, is denoted Δt, and a time step of the weather equation, or a microtime step, is denoted δt.

For each macrotime step, the number of integration microtime steps is denoted N. Of these N microtime steps, N_1 are treated as 'spin-up' to bring the weather trajectory down onto its instantaneous (x-dependent) weather attractor. An ensemble of R weather realisations is initialised at the beginning of the climate integration by running a long weather integration, for fixed initial climate variable, and choosing from this integration R well separated points to use as initial points.

3.1 Cumulative averaging method

This superparameterisation method was first developed in Fatkullin and Vanden-Eijnden (2004) using consistency results as a formal justification. In Weinan et al. (2005), the cumulative averaging schemes were applied to integration of the (A) and (N+) differential equations. Convergence to these reduced equations was rigorously justified, and the bounds on the rates of convergence were established. The intuition behind the method is developed below in application to the deterministic and stochastic Hasselmann approximations.

3.1.1 (A) approximation

The cumulative averaging method applies to computation of all expectations, including those that determine the noise-induced drifts and multiplicative noise. The simplest expectation that determines the (A) approximation (and is common to all of Hasselmann's approximations) is the effective climate tendency forcing: the conditional average of the unreduced model climate forcing, or equivalently, the time-average of the climate forcing over realisations of the weather for fixed climate. A standard approximation of this mean is:

$$\overline{f}(x) = \lim_{T\to\infty} \frac{1}{T} \int_0^T f(x, y_t^x)\, dt \approx \frac{1}{(N-N_1)R} \sum_{r=1}^{R} \sum_{n=N_1+1}^{N} f(x, y_{r,n}^x). \tag{3.1}$$

After the effective climate variable is advanced a macrotime step, the weather ODE must be initialised to find the average with respect to a new conditional distribution. An efficient initialisation is:

$$y_{r,1}^{x(t)} = y_{r,N}^{x(t-\Delta(t))}; \tag{3.2}$$

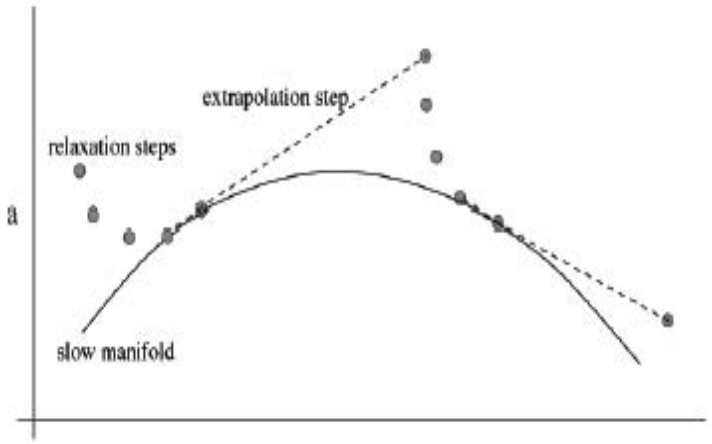

Figure 3.1: Figure from www.math.princeton.edu/ weinan illustrating the extrapolation integration scheme. The slow equation is advanced by a large extrapolation time step (or macrotime step) and relaxation steps (or microtime steps) bring the solution back onto the slow manifold.

i.e., the last weather iterate from the previous weather integration is the initial weather iterate in the next weather integration. This is an efficient approach because the 'distance' between successive attractors, on which the weather solutions settle, will be small because x changes by a small amount over the time step Δt. Since the last weather iterate is on its attractor, the new initial weather iterate will be close to the new attractor. Figure 3.1 is a graphical representation of an extrapolation integration scheme, which is a simpler version of the cumulative averaging method. In this simple case, the fast weather solution is "slaved" to a slow manifold, such that it relaxes to a single point on the manifold; in the general case, the weather solution tends to a non-degenerate distribution on a (non-trivial) attractor (Gardiner 1996).

The initialisation (3.2) is used in AI03 in finding a library of weather distributions prior to integration of the effective climate equation. The model in AI03 has a 3-variable weather ODE coupled to one of the three climate variables, which modulates the conditional weather distribution. The motivation behind using the path-following approach is to ensure that, past a bifurcation point of the climate parameter, a particular branch of the weather solution is followed. Each branch has its own distinct family of 3-variable weather distributions occupying a subspace of the entire space of the weather variable. Different initialisations of the weather ODE can lead its solution to different attractors. Thus, past the bifurcation

point, random initialisations of the weather ODE could result in a wildly fluctuating solution of the climate equation. There appears to be unique ergodicity of the weather dynamics for fixed climate in the KRG05 model, but initialisation (3.2) nevertheless guards against the possibility of wandering between different weather attractors.

Although the relaxation to the weather attractor may be fast, an average must be taken, for which a large sampling of the attractor, and hence, long integration time, may be required to find a converging averaged solution. However, it is a remarkable fact that for sufficiently large scale separation and small macrotime step, setting $N = R = 1$ and $N_1 = 0$ (no spin-up) in calculation of the average (3.1) with initialisation (3.2) yields a good approximation of the statistics of $\overline{x}$. That is, the averaged statistics can be well approximated without averaging at each macrotime step. With an $\mathcal{O}(1)$ macrotime step, there is an $\mathcal{O}(\epsilon^{-1})$ efficiency gain with these settings. These remarkable efficiency gains, justified and quantified in Weinan et al. (2005), are possible because averaging is in fact occurring cumulatively over many macrotime steps. Since successive attractors are 'close', the average over successive attractors is close to the average over one attractor. A simple argument illustrates this idea (for $R = 1$ and $N_1 = 0$):

$$\begin{aligned}
x_{m+1} &\approx x_0 + ... + \overline{f}(x_{m-1})\Delta t + \overline{f}(x_m)\Delta t \\
&\approx x_0 + ... + \Delta t \left[\frac{1}{N} \sum_{i=1}^{N} f(x_{m-1}, y_{1,i}) + \frac{1}{N} \sum_{i=1}^{N} f(x_m, y_{1,i}) \right] \\
&\approx x_0 + ... + 2\Delta t \frac{1}{2N} \sum_{i=1}^{2N} f(x_{m-1}, y_{1,i}) \\
&\vdots \\
&\approx x_0 + ... + L\Delta t \frac{1}{LN} \sum_{i=1}^{LN} f(x_{m-(L-1)}, y_{1,i})
\end{aligned}$$

where L is sufficiently small that x_m does not change significantly from $x_{m-(L-1)}$. Effectively, the average is found over a macrotime step of size $L\Delta t$, but with an actual, stable macrotime step of Δt. Over LN summands, the approximation of the continuous-time average can be good even if $N = 1$.

Notice that the summations can be grouped together because of the initialisation scheme

and the small size of the macrotime step (and thus the small change in the slow variable). The initialisation (3.2) is not necessary for the argument to hold, but it does connect successive integrations of the fast equation such that, effectively, there is a continuous integration of the fast equation as the climate equation is integrated. In fact, as discussed in Fatkullin and Vanden-Eijnden (2004), the averaging scheme with $N-N_1 = 1$ is consistent with:

$$\begin{aligned} \dot{x} &= \frac{1}{R}\sum_{r=1}^{R} f(x, y_{r,N}^x) \\ \dot{y_r^x} &= \frac{1}{\epsilon\delta} g(x, y_{r,N}^x) \end{aligned} \tag{3.3}$$

in the limits $\Delta t \to 0$, $\delta t \to 0$ and $\frac{\Delta t}{N\delta t} \to \delta$. That is, the fast variables are effectively slowed down by a factor of δ. In turn, this differential equation converges to the averaged differential equation (A), in the limit $\epsilon\delta \to 0$. However, for small R, the averaging scheme in fact yields a solution that is somewhere in between the solutions of the unreduced model with the fast dynamics slowed down by a factor of δ and the averaged model; i.e., it is a hybrid unreduced-averaged solution. There is an efficiency gain over the unreduced KRG05 model if $\delta > 1$, or if the fast modes are slowed down. The gain can be dramatic for large scale separation: in Fatkullin and Vanden-Eijnden (2004) the averaged Lorenz-96 toy model of the atmosphere with $N = R = 1$ and $N_1 = 0$ (i.e., no spin-up and no averaging at each macrotime step) is a successful model of the low-frequency modes which is more efficient by a factor of 32 than the parent Lorenz-96 model with the setting of $\epsilon = 2^{-7}$. The setting $\delta > 1$ causes the 'stretching out' of the fast dynamics over successive macrotime steps, thereby increasing its timescale. Slowing down the weather dynamics will not affect the climate dynamics if there is sufficient scale separation (i.e., if $\epsilon \ll 0$) because the climate component of the climate-weather system remains sufficiently close to the limiting climate. Although the hybrid model is not an exact integration of the original model, it may be close enough to justify its use as a numerically efficient parameterisation scheme.

With the settings $R = 1$ and $N - N_1 = 1$, the hybrid (A)-unreduced approximation may be much closer to the solution of the unreduced model than to the solution of the (A)

model. Thus, the hybrid model with no averaging at each macrotime step and with fast modes sufficiently *sped up* (i.e., by setting $\delta < 1$ to be sufficiently small) may give a good approximation of the limiting dynamics represented by the determinisitic and stochastic reduced equations. A correspondence between these models would validate the Hasselmann reduced models as limiting models of the climate dynamics. Certainly, if the model with fast modes sped up is unstable, as it may be in a neighbourhood of a bifurcation point, then the validity of the reduced models would be cast into doubt. The model considered in this dissertation (the KRG05 model) does have a bifurcation point, but the reduced models are found at points sufficiently removed from the bifurcation point.

3.1.2 (N) approximation

All of the conditional expectations (conditional on the climate value), including lag-covariance, for the reduced averaged and stochastic Hasselmann models can be found in a cumulative sense, over many macrotime steps of the climate variable. However, not all of the calculations in determining the noise-induced drift and diffusion terms are derived in a cumulative fashion. For example, the diffusion term is the integral of the lag-covariance of the slow tendency forcing (2.14), numerically approximated as:

$$\sigma\sigma^T = \frac{2\delta t}{R} \sum_{j=1}^{R} \sum_{m=N_1+1}^{N} (f(x, y^x_{j,m}) - \overline{f})(f(x, y^x_{j,N_1+1}) - \overline{f}). \tag{3.4}$$

For a particular lag, $m_0 - (N_1 + 1)$, the lag-covariance is determined in a cumulative fashion, but integration over multiple time-lags is not cumulative. Thus, the weather integrations at *each* climate value should be carried out for a sufficiently long time to approximate the integral of the lag-covariance up to infinite time lag. In particular, the uppermost lag should be of the order of the decorrelation timescale of the fast modes. This timescale is not expected to be too large when there is good scale separation (and the mixing assumption is satisfied). Since there is at most an order of magnitude scale separation between climate and weather modes in the KRG05 model, calculation of the time integral up to its decorrelation timescale would require a timescale in excess of the macrotime step of the effective climate equation, rendering the scheme inefficient. This necessitates a simplification of the

calculation of the diffusion term.

In deriving the Hasselmann (N) approximation of the KRG05 model, a simple extrapolation scheme is applied in calculation of the diffusion term σ. Prior to integrating this stochastic equation, the integrated lag-covariances are calculated up to short and long time-lags for a range of climate values derived from the unreduced KRG05 model. An extrapolation factor relating these integrated lag-covariances is determined, to be used as the factor by which the integrated lag-covariance up to a short time-lag is upscaled in approximating the integrated lag-covariance up to a long time-lag. This procedure ensures that σ has the correct scaling, while permitting an efficient integration of the effective climate equation. This scheme is illustrated and the various scales are derived in Chapter 6.

The noise-induced drift of the (N+) approximation has a similar formula to that for the diffusion, σ, but the derivative $\nabla_x y_t^x$ in the formula for this drift is very inefficiently approximated at each climate value. Recall from Chapter 2 that a differential equation (2.22) for $\nabla_x y_t^x$ was derived which only required the evaluation of the derivative of the fast modes at an initial time and at each macrotime step. However, it is not clear how to initialise this equation such that the initial value is close to the attractor of $\nabla_x y_t^x$. In Weinan et al. (2005), efficient schemes are developed that circumvent this problem, but these schemes apply to very simple models only, which rules out their application in deriving a (N+) approximation of the complex KRG05 model. In this dissertation, (N+) models will not be found.

As an illustrative example, consider the example SDE (2.26) with the timescale separation of $\epsilon = 0.1$. The algorithm for integration of the (N) SDE, the solution of which approximates x in (2.26), is outlined below.

Algorithm for integrating (N) equation

Initial time: $t = t_0$

Integration of weather equation for ensemble member $R = r$

y_0^X

$y_{\delta t}^X = y_0^X + \left[-(y_0^X - X) + \dot{W}\right] \delta t$

$\vdots$

(spin-up fast equation N_1 microtime steps)

$y_{(N_1+1)\delta t}^X = y_{N_1\delta t}^X + \left[-(y_{N_1\delta t}^X - X) + \dot{W}\right] \delta t$

$f(1, r) = -(y_{(N_1+1)\delta t}^X)^3 + sin(\pi t_0) + cos(\sqrt{2}\pi t_0)$

(save the full model slow tendency forcing, driven by y^X)

$\vdots$

$y_{N\delta t}^X = y_{(N-1)\delta t}^X + \left[-(y_{(N-1)\delta t}^X - X) + \dot{W}\right] \delta t$

$f(N - N_1, r) = -(y_{N\delta t}^X)^3 + sin(\pi t_0) + cos(\sqrt{2}\pi t_0)$

Integration of effective climate equation

$\overline{f} = \frac{\sum_{j=1}^{R} f(N-N_1,j)}{R}$

(effective slow tendency forcing)

$\sigma^2 = 2\delta t \sum_{i=1}^{N-N_1} \frac{\sum_{j=1}^{R} (f(i,j)-\overline{f})(f(1,j)-\overline{f})}{R}$

(diffusion coefficient, upper time-lag of $(N - N_1)\delta t$)

$X(t + \Delta t) = X(t) + \overline{f}\Delta t + \sqrt{0.1}\sigma\sqrt{\Delta t}\xi$

(advance effective slow mode ($\xi \sim \mathcal{N}(0, 1)$))

Initialisation for next integration of weather equation

$y_0^X = y_{N\delta t}^X$

For this example, the (N) and (A) simulations (the latter plotted in Figure 2.2) are virtually identical. This is not surprising as both the variance of the fast variable (2.27) and the unreduced model tendency forcing $f = f(y, t)$ are independent of the slow variable, x.

3.2 Seamless cumulative averaging

In practice, the 'Algorithm for integrating (N) equation' cannot be applied to the high-dimensional KRG05 model (to be detailed next chapter) because that would entail resolving almost $10,000$ fast equations. However, the same integration can be carried out without explicitly resolving the fast modes. In particular, the reduced Hasselmann equations can be computed using only the original system of PDEs and the projection of the original system onto the slow modes. The fast weather subsystem is obtained as the difference between the full system and the slow subsystem. The seamless integration algorithms for Hasselmann's reduced equations are developed in Chapter 5, which require only minor modifications to the original KRG05 algorithm.

Chapter 4

The Kravtsov et al. (2005) model of LFV

4.1 The KRG05 model equations

The Kravtsov et al. (2005) (KRG05) model is a channel model of the midlatitude troposphere, governed by the two-layer quasi-geostrophic (QG) equations on the β-plane. In the midlatitudes, the QG equations are the equations of motion of the planetary and synoptic scale waves alone, with smaller-scale waves (e.g., gravity and sound waves) filtered out. The continuous QG equations are an excellent representation of the large-scale dynamics in the mid-latitudes because of the large scale separation between the slowly-evolving and fast-evolving processes, illustrating that such a scale separation between balanced and unbalanced dynamics does indeed exist in the real climate system, and can be successfully exploited to derive a useful equation in the balanced modes alone. Note that the reduction methods considered in this dissertation - in which both slow and fast modes are contained within the balanced dynamics - are distinct from the asymptotic analysis used to derive QG dynamics from the primitive equations. The two-layer KRG05 model has poor vertical resolution, but as LFV in the real atmosphere is predominantly equivalent barotropic, the model is successful in qualitatively capturing important large-scale features.

A model schematic is depicted in Figure 4.1. The model lower surface is a Northern Hemisphere-like, zonally-asymmetric surface in midlatitudes, with a body of water representing the North Atlantic basin. To the north of the basin, the strip of land represents sea-ice cover, but it has the same properties, including albedo, as the land surface. The entire surface has flat topography.

The model dynamics are given explicitly in terms of the barotropic and baroclinic streamfunctions, respectively ψ and λ:

$$\frac{\partial q_\psi}{\partial t} + J(\psi, q_\psi) = -k\nabla^2\psi - \sum_{n=1}^{3} k^{(n)}\nabla^2\psi^{(n)} + A_H\nabla^6\psi - h_1h_2\left[J(\lambda, q_\lambda) + \frac{k}{h_1}\nabla^2\lambda\right] \quad (4.1)$$

$$\frac{\partial q_\lambda}{\partial t} + (h_2 - h_1)J(\lambda, q_\lambda) = \frac{f_0}{H_a}\frac{1}{h_1h_2}F(x, y; \lambda, \psi) - \frac{h_2}{h_1}k\nabla^2\lambda - \sum_{n=1}^{3} k^{(n)}\nabla^2\lambda^{(n)} + A_H\nabla^6\lambda$$
$$- \left[J(\lambda, q_\lambda) + J(\psi, q_\lambda) + \frac{k}{h_1}\nabla^2\psi\right], \quad (4.2)$$

where the barotropic and baroclinic components of the potential vorticity are respectively,

$$q_\psi = \quad \nabla^2\psi + \beta y, \quad (4.3)$$

$$q_\lambda = \quad \nabla^2\lambda - \tfrac{1}{R_d^2}\lambda \equiv \quad \mathcal{L}\lambda. \quad (4.4)$$

In terms of physical processes, these equations read:

total tendency (barotropic potential vorticity) =

bottom Ekman drag + scale-selective Ekman drag + supervisc.+ interaction with bc. mode

total tendency (baroclinic potential vorticity) =

thermal forcing + Ekman drags + supervisc.+ interaction with bt. mode.

The equations are expressed in non-dimensional terms. The nondimensional heights of the thinner lower layer and thicker upper layer are $h_1 = 0.3$ and $h_2 = 0.7$, respectively, R_d is the Rossby radius of deformation, f_0 is the Coriolis parameter, β is the gradient of the Coriolis parameter at $45°N$, k^{-1} is the spin-down timescale (governed by the bottom Ekman drag), A_H is the superviscosity coefficient, and $k^{(1)}$, $k^{(2)}$ and $k^{(3)}$ are enhanced Ekman damping constants applied to the lowest frequency waves (corresponding to $\psi^{(1)}$, $\psi^{(2)}$ and $\psi^{(3)}$, and $\lambda^{(1)}$, $\lambda^{(2)}$ and $\lambda^{(3)}$). This latter drag is necessary in rigid-lid models because the lack of meridional and vertical dispersion results in accumulation of vorticity at the lowest frequencies (Kravtsov et al. 2003). The superviscosity acts on the opposite end of the spectrum, on the high frequencies, parameterising unresolved turbulent fluxes of

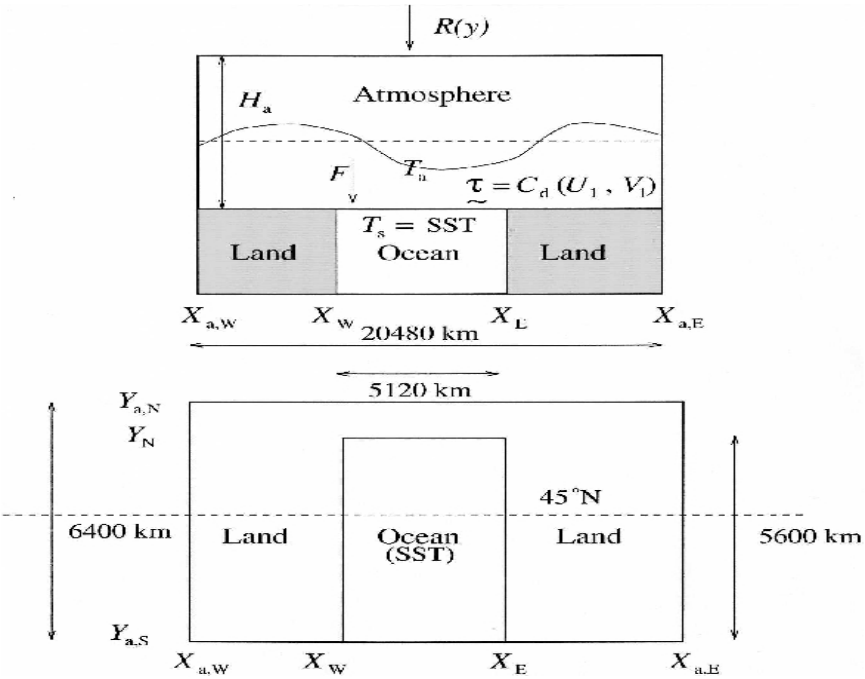

Figure 4.1: Schematic from Kravtsov et al. (2005) of the KRG05 model of LFV over a NH-latitude-like surface. Top is a cross-section of the model and bottom is a plan view. The surface boundary corresponds to the North Atlantic basin and surrounding land and sea-ice.

enstrophy to small scales. The equations are discretised onto an evenly-spaced horizontal grid with 41 longitudinal levels and 128 latitudinal levels. Combined with the two vertical levels, there are $128 \times 41 \times 2$ total gridpoints in the system.

The model is forced by a temporally constant but meridionally varying insolation, R, determined according to a standard observationally-based formula. The atmosphere is transparent to the shortwave forcing, and emits long-wave radiation both upward and downward. This long-wave flux, B, is linear in atmospheric temperature, which itself is linear in the baroclinic streamfunction. The surface heat fluxes, F, are parameterised in terms of the atmospheric temperature and, over water, in terms of the sea surface temperature (SST) as well. The free atmospheric forcing is parameterised as a relaxation to an equilibrium temperature.

The land, including sea-ice, is modelled as a blackbody with an insulating condition imposed that it emits only long-wave radiation and no sensible or latent heat fluxes. Thus,

the surface atmospheric forcing, F, over land is

$$F = \frac{1}{\rho_a c_{P,a} \Delta\theta_s}(R - B),$$

where ρ_a is the representative atmospheric density, $c_{P,a}$ is the atmospheric heat capacity and $\Delta\theta_s$ is the difference in potential temperature between the layers.

Zonal thermal contrasts arise from differences between air-sea and air-land surface heat fluxes. The ocean is also modelled as a blackbody, but there is no insulating condition, allowing for sensible and latent heat fluxes between the ocean and atmosphere. These fluxes are linear in ψ and λ, except for the wind-stress component of the heat flux. As this is a dry model (i.e., water is not explicitly modelled) the sensible and latent heat fluxes are parameterised together in a single forcing term, H_{SL}, determined by a standard bulk formula dependent on SST, humidity differences between the ocean and air, and the wind speed at the ocean surface. In particular,

$$H_{SL} = \rho_a U(C_h c_{P,a}(T_s - T_a) + C_e L[q(T_s) - q(T_a)]), \tag{4.5}$$

where L is the latent heat of vaporization of water, ρ_a is the representative atmospheric density, $c_{P,a}$ is the atmospheric heat capacity, and C_h and C_e are empirically determined. The wind speed is derived from wind stress (τ) according to:

$$U = \left[\frac{(\tau_{x-dir}^2 + \tau_{y-dir}^2)^{\frac{1}{2}}}{\rho_a C_d}\right]^{\frac{1}{2}}, \tag{4.6}$$

where C_d is an empirically determined drag coefficient. (Wind stress τ is a linear function of the atmospheric lower-layer velocities.) As detailed in the next chapter, the non-quadratic nonlinear wind speed must be linearised in order to apply the MTV method within the framework established in FMV05 and FM06.

Coupled with the SST-dependent oceanic long-wave radiative flux (O), the atmospheric

surface forcing relationship over the ocean is:

$$F = \frac{1}{\rho_a c_{P,a} \Delta\theta_s}(O + H_{SL} - 2B)$$

The forcing over the ocean acts to strongly dampen atmospheric variability.

The boundary conditions to close the QG equations include periodic flow at the zonal boundaries, and the Neumann conditions of no parallel flow ('no slip') nor normal flow ('no flow') at the north and south boundaries. The no flow condition in particular implies that each of the streamfunctions must not vary with longitude along each of the northern and southern walls. In addition to these standard conditions, integral constraints of momentum and heat are imposed to ensure that the QG equations satisfy all conservation laws associated with the parent primitive equations. In particular, the momentum constraint implies that the circulation around the boundary is internally forced and the heat constraint ensures that the system is thermally forced by the boundary heat exchanges only, parameterised by F. The constraint derivations are outlined below, following McWilliams (1977).

In a two-layer incompressible fluid with the rigid-lid approximation and no flow at the boundaries, the heat content of the system is proportional to the integral interface displacement. For example, if the system is forced by a negative heat flux, it cools by a thickening of the lower, 'colder' (in the sense of potential temperature) layer and a corresponding thinning of the upper, warmer layer. Equivalently, there is a simultaneous cooling of the atmosphere and an entrainment of mass into the lower layer from the upper layer. The interface displacement is itself proportional to the baroclinic mode. The tendency equation relating the baroclinic mode to the thermal forcing is:

$$-\frac{1}{R_d^2}\frac{\partial}{\partial t}\iint_\Omega \lambda \, d\Omega = \frac{f_0}{H_a}\frac{1}{h_1 h_2}\iint_\Omega F \, d\Omega. \tag{4.7}$$

This statement of thermal energy conservation is in fact equivalent to the statement of mass conservation. This same equation (4.7) may be derived from consideration of the continuity equation, starting with the assumption that there is no integrated flow through

solid boundaries.

The barotropic integral momentum constraint is derived by integrating the PDE for the barotropic streamfunction over the channel domain. The area integral is simplified by applying the divergence theorem and the periodic boundary conditions on the west and east walls. Additionally, the streamfunction on the southern wall is set to $\psi \equiv 0$ since the barotropic mode is defined only up to an arbitrary constant. The resulting momentum constraint is:

$$\frac{\partial}{\partial t}\int_N \psi_y \, dx = \int_N (A_H \psi_{y^{(5)}} + ...) \, dx, \tag{4.8}$$

where the integrals are over the northern wall of the domain and $\psi_{y^{(5)}}$ is the fifth partial derivative of ψ with respect to y.

Integrating the baroclinic equation over the channel domain and applying the periodic boundary conditions and divergence theorem gives:

$$\frac{\partial}{\partial t}\left[\int_N \lambda_y \, dx - \int_S \lambda_y \, dx - \frac{1}{R_d^2}\iint_\Omega \lambda \, d\Omega\right] =$$
$$\int_N (A_H \lambda_{y^{(5)}} + ...) \, dx - \int_S (A_H \lambda_{y^{(5)}} + ...) \, dx + \iint_\Omega \frac{f_0}{H_a}\frac{1}{h_1 h_2} F \, d\Omega.$$

The equation is further simplified by subtracting from it the thermal constraint equation to eliminate the double integrals from both sides, leaving only integrals on the boundaries. Finally, the equation is split into two by assuming that the integral of momentum on each of the walls is forced only at the wall; i.e.,

$$\begin{aligned}\frac{\partial}{\partial t}\int_N \lambda_y \, dx &= \int_N (A_H \lambda_{y^{(5)}} + ...) \, dx \\ \frac{\partial}{\partial t}\int_S \lambda_y \, dx &= \int_S (A_H \lambda_{y^{(5)}} + ...) \, dx.\end{aligned} \tag{4.9}$$

The PDEs and the integral constraints, (4.8) and (4.9), together close the system. In discretising the system, the barotropic mode is fully determined over $128 \times 39 + 1$ grid points, because it satisfies $\psi(x, y, t) \equiv 0$ along the entire southern wall and $\psi(x, y, t) = \psi(t)$ along the northern wall. Since the baroclinic mode is fully determined over $128 \times 39 + 2$ grid points, the full system of QG PDEs is fully determined over $128 \times 39 \times 2 + 3$ grid points.

In the KRG05 algorithm, the barotropic and baroclinic streamfunctions are advanced in time in two stages: first, stepping forward according to the PDEs (4.2) and second, applying the momentum constraints. An Euler discretisation of the first stage solution is:

$$\begin{aligned} \Delta^{-1}\left[q_\psi(x,y,t) + Dq_\psi(x,y,t)\delta t - \beta y\right] &\equiv \psi_{00}(x,y,t+\delta t) \\ \mathcal{L}^{-1}\left[q_\lambda(x,y,t) + Dq_\lambda(x,y,t)\delta t\right] &\equiv \lambda_{00}(x,y,t+\delta t) \end{aligned} \tag{4.10}$$

where Δ^{-1} and $\mathcal{L}^{-1}$ are (linear) inverse operators (cf. (4.3, 4.4)) that satisfy $\psi\big|_S \equiv 0$, $\psi\big|_N \equiv 0$, $\lambda\big|_S \equiv 0$, and $\lambda\big|_N \equiv 0$.

The general solution consistent with the boundary conditions is then:

$$\begin{aligned} \psi(x,y,t+\delta t) &= \psi_{00}(x,y,t+\delta t) + C_1(t)\psi_{01}(y) \\ \lambda(x,y,t+\delta t) &= \lambda_{00}(x,y,t+\delta t) + C_2(t)\lambda_{01}(y) + C_3(t)\lambda_{10}(y). \end{aligned} \tag{4.11}$$

The function ψ_{01} is the solution of (4.3) with $q_\psi = 0$ (i.e., homogeneous Laplace equation) and $\psi\big|_S \equiv 0$ and $\psi\big|_N \equiv 1$. The functions λ_{01} and λ_{10} are the solutions of (4.4) with $q_\lambda = 0$ (i.e., homogeneous Helmholtz equation) and, respectively, $\psi\big|_S \equiv 0$ and $\psi\big|_N \equiv 1$, and $\psi\big|_S \equiv 1$ and $\psi\big|_N \equiv 0$. The coefficients, C_1, C_2 and C_3 are found by substituting (4.11) into the three integral momentum equations, which incorporate the heat/mass constraint.

The PDEs and integral constraints can be combined into one autonomous system, with a unified tendency forcing. An explicit tendency forcing is necessary in application of the MTV method, following FMV05 and FM06.

In the next section, the conclusions of KRG05 are presented, for comparison with the results obtained in this dissertation.

4.2 Model results from KRG05

The KRG05 model was designed to study low-frequency variability in the extratropical atmosphere. A natural control variable for the model LFV is the Ekman dissipation spin-down timescale, k^{-1}, which acts as a bifurcation parameter (from hereon k^{-1} represents the dimensional spin-down timescale). A bifurcation occurs at the realistic spin-down timescale

of $k^{-1} = 5.6$ days, characterised by the emergence of two metastable jet states from one jet state (i.e., 'jet bimodality' emerges, though multiple regime behaviour in the KRG05 model is mostly manifest as a long-tailed distribution as opposed to a bimodal distribution). For $k^{-1} > 5.6$ days, the sole model jet shifts irregularly between high and low-latitude states, as depicted in Figure 4.2. As the spin-down timescale increases (or equivalently, the bottom drag decreases), the persistence of the jet in each state increases, as does the (latitudinal) distance between the states. At very high values of spin-down timescale, the timescale between regime transitions becomes very large. The range of observed spin-down timescales lies largely in the range corresponding to jet bimodality (Lorenz and Hartmann 2001). As well, the leading empirical orthogonal function (EOF) of low-pass filtered model data, that captures much of the meridional displacement of the jet, is similar to its counterpart in the real atmosphere, the Arctic Oscillation (AO). However, large meridional shifts of the jet axis are not seen in the real atmosphere. Another leading mode of real atmospheric data corresponds to the strengthening/weakening of the jet, but the corresponding mode in the KRG05 model is not dynamically important (except at very low spin-down timescales of $k^{-1} < 2.0$ days).

For $k^{-1} > 2.0$ days, the leading EOFs of unfiltered model data are a propagating wavenumber-four pair of EOFs, in spatial and temporal quadrature, and a wavenumber-four modulated stationary EOF, which is also the leading EOF of zonally-averaged data (Figure 4.3). Wave-4 is defined to consist of the first two of these EOFs (and the associated PCs) and the stationary mode consists of the one EOF (and its associated PC). Similar to the AO pattern, the stationary mode EOF has a significant zonally symmetric component that becomes more dominant as the spin-down timescale increases, such that the stationary mode EOF better captures the meridional displacements of the jet as the bottom drag decreases. KRG05 claimed that the dynamical modes responsible for meridional displacements between jet regimes are contained in the space spanned by these three leading EOFs. In particular, the multiplicity of jet states was interpreted as resulting from weakly nonlinear interactions between the stationary and wave-4 modes. This determination was based primarily on three points (KRG05): "First, mode 1 and wave 4 are the only two

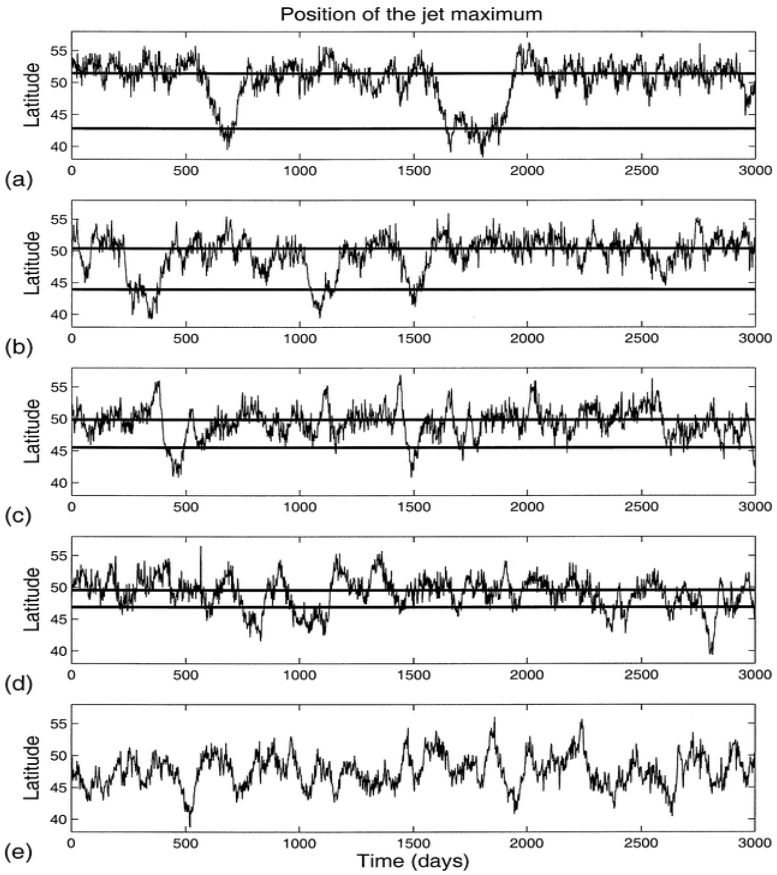

Figure 4.2: Figure from KRG05. "Time series of the jet-center position: a)k^{-1} = 7.7 days; b)k^{-1} = 7.1 days; c)k^{-1} = 6.7 days; d)k^{-1} = 6.3 days; and e)k^{-1} = 5.9 days. Heavy solid lines mark the climatological location of the jet in the high-latitude and low-latitude states."

modes whose variances change significantly when crossing the bifurcation point...Second, the power spectra of wave 4 show a persistent decrease of frequency as k is increasing to its critical value, where the dominant frequency of wave 4 is very low (Figure 4.4); hence, the interaction of wave 4 with [the stationary mode], whose temporal behaviour resembles red noise, is likely to become increasingly important there. Last, the largely zonal [stationary mode] possesses a wave-4 modulation".

The dynamics in the region of jet bimodality was further analysed in KRG05 based on the linearised vorticity equation and the climatological vorticity budget. It was found that there is a correspondence between the leading statistical modes (the stationary mode and the propagating wave-4, wave-5 and wave-6 modes) and the dynamical modes of the linearised equation. As such, it was concluded that these modes are Rossby waves. It was also determined from the linearised equation that the advection of the stationary mode's

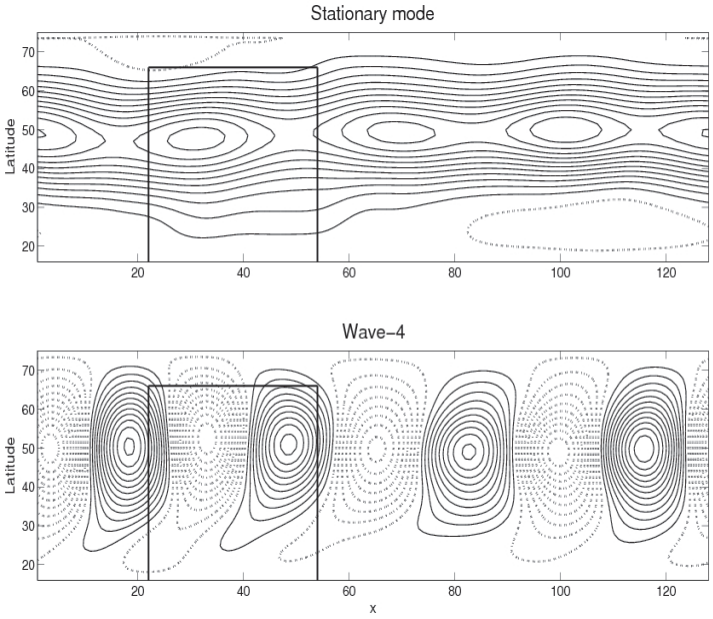

Figure 4.3: The stationary (top) and wave-4 (bottom) EOFs at $k^{-1} = 6.7$ days of the bulk (barotropic & baroclinic) streamfunction fields at $k^{-1} = 6.7$ days (contour interval=0.0025).

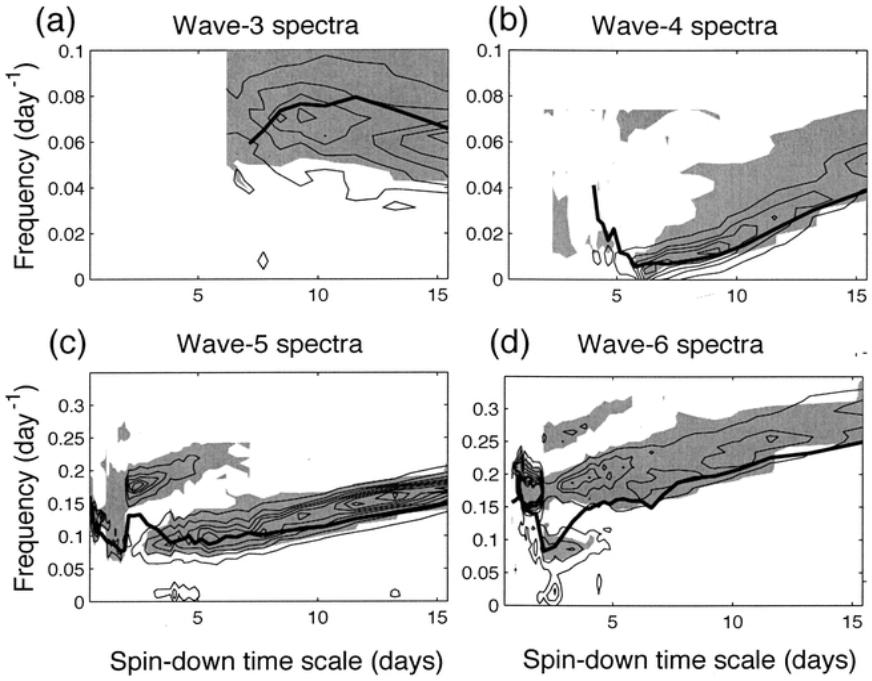

Figure 4.4: Figure from KRG05. Spectra of leading propagating modes over range of k^{-1}; "(a) wave 3 (CI= 0.03); (b) wave 4 (CI= 0.4); (c) wave 5 (CI= 0.05); and (d) wave 6 (CI= 0.02). Contours are power spectral density ($0.78{\times}10^1 6 m^4 s^{-2} day^{-1}$): shaded areas are statistically significant at the 95% level against a null hypothesis of red noise. Heavy solid lines depict the fundamental frequency of the mode as obtained by linear stability analysis."

vorticity by the climatological flow and the advection of the climatological vorticity by the stationary mode nearly cancel, which is a characteristic property of stationary Rossby waves. Since the stationary mode is associated with negligible forcing of the climatological vorticity equation in both regimes, it was determined to be close to a free mode of the system. KRG05 claim that based on the lack of exchange of energy with the climatological flow, the free stationary mode is not as damped by nonlinear effects at low frequencies as are the propagating modes, resulting in its dominant variability at low frequencies. At shorter timescales, the stationary and propagating modes do interact, with wave-4 emerging as the dominant mode in interaction with the stationary mode on these timescales.

The leading synoptic eddies, the propagating wave-5 and wave-6 modes in particular, were determined in KRG05 to be second-order drivers of the bimodal dynamics. A primary effect of these eddies is in the modulation of the persistence and frequency of the jet in the high and low latitude regimes. Wave-5 in particular, which acts to reduce the meridional temperature gradient, is less energetic in the low-latitude regime, resulting in a stronger jet in this regime. Wave-4 feeds off the greater jet energy, resulting in greater disturbances of the low-latitude regime relative to the high-latitude regime. Consequently, the high-latitude regime is preferred and more persistent. However, it is the interaction between the leading two planetary-scale modes rather than the interaction between the planetary and synoptic modes which was interpreted as being responsible for the jet bimodality. In Kravtsov et al. (2003), in which essentially the same model is analyzed, the forcing of the low-frequency modes by the synoptic eddies, defined as the high-pass filtered streamfunction at timescales of 10 days and less, was replaced by its temporally-white, structurally-maintained stochastic analogue without significantly affecting the statistics of the low-frequency modes.

In the next chapter, the Hasselmann and MTV methods are applied to the KRG05 model (slow-fast QG PDE) to obtain closed equations for the low-frequency modes (ODEs and SDEs for the slow modes alone). The subsequent analysis will allow for a critical assessment of the conclusions of KRG05 in terms of systematically derived low-dimensional models.

Chapter 5

Application of the MTV and Hasselmann methods to the KRG05 model

As mentioned in the previous chapter, the PDEs and integral constraints of the KRG05 model can be combined into an autonomous PDE:

$$\begin{aligned}\frac{\partial \psi}{\partial t} &= f(\psi, \lambda) \\ \frac{\partial \lambda}{\partial t} &= g(\psi, \lambda).\end{aligned}$$

To illustrate the implementation of the reduction methods, the equation for the 'bulk streamfunction', symbolically denoted $p = (\psi, \lambda)$, is considered:

$$\begin{aligned}\frac{\partial p}{\partial t} &= F(p) \\ &= (f(p), g(p)).\end{aligned}$$

The above system can be recast in proper matrix notation; subsequent manipulations are carried out on the symbolic system.

There are two primary steps in systematic mode reduction: first, the equations are projected onto a suitable basis, which (hopefully) yields a slow-fast system; second, reduction theories are applied to derive a system in the slow variable only. The projection must yield a system of ODEs, as the reduction theories apply to systems of ODEs, not PDEs. Potentially, there are several bases, such as EOFs and PIPs (Principal Interaction Patterns) (Kwasniok 1996), by which a system may be efficiently partitioned. For example, a few

leading EOFs may capture a large fraction of the variance of the field, but these leading EOFs do not necessarily correspond to the leading dynamical modes. In some fluid dynamical applications involving 'bursting' or rapid transition events which are dynamically significant but carry little variance, a large set of leading EOFs fails to capture the relevant dynamics (Crommelin and Majda 2004). It is also demonstrated in Crommelin and Majda (2004) that the leading EOFs are insufficient to describe the dynamics of a simple model of LFV characterised by irregular transitions between atmospheric regimes. This earlier study found that a truncated model determined by projection onto PIPs best captures the dynamics. However, there are several sensitivities in deriving PIPs that might limit their utility. Moreover, EOFs are common in systematic reduction studies, including D'Andrea and Vautard (2001), FMV05 and FM06. Since the primary objective of this study is to extend Hasselmann's method to an atmosphere model of intermediate complexity, and to compare with the MTV method, EOFs here serve as the basis for reduction.

The set of normalised eigenvectors $(e_i(x,y))$ of the covariance function of the streamfunction $p(x,y,t)$, denoted the EOFs, satisfies:

$$p(x,y,t) = \overline{p}(x,y) + \sum \alpha_i(t) e_i(x,y),$$

where $\overline{p}$ denotes the time-mean and

$$\alpha_i(t) = (p, e_i),$$

where $(\cdot,\cdot)$ is an inner product. The inner product associated with the L_2 norm of the streamfunction field, $\|\psi\|_0^2 = \int \psi^2 \, dA$, where dA is an infinitesimal area of the channel, is used in reduction of the KRG05 model over other possibly more physical norms such as the total energy norm. The total energy norm, applied towards reduction in FM06, is an integral norm that ensures that total energy associated with a conservative equation remains conserved by projection onto any set of leading EOFs (Selten 1995, Ehrendorfer 2000). In the system under consideration, the total energy norm is not necessarily superior to other norms at conservation of total energy because of the presence of non-conservative

terms driving the forced-dissipative dynamics in the QG equations.

As the grid cells of the KRG05 model have a unit area, except at the boundaries at which they have a half-unit area, the inner product in the interior of the channel associated with the L_2 norm is the dot product. For clarity, all inner products are written as dot products.

Applying the dot product, the projection of the governing equation onto the set of EOFs (e_i) yields a system of interdependent ODEs for the PC time-series $\alpha_i(t)$:

$$\frac{d\alpha_i}{dt} = e_i \cdot \left[F\left(\overline{p} + \sum \alpha_j e_j\right)\right]. \tag{5.1}$$

The EOFs-PCs are divided into slowly-evolving and fast-evolving modes based on their autocorrelation timescales, such that: $\sum \alpha_i e_i = \sum_{i \in S_a} a_i e_i + \sum_{j \in S_b} b_j e_j$; $S_a \cap S_b = \emptyset$, where S_a and S_b are the index sets of the slow and fast modes, respectively (Majda et al. 2001, Berner 2005).

The ODE (5.1) can be re-written to reflect the partitioning into slow (a) and fast modes (b) as:

$$\frac{da_i}{dt} = e_i \cdot \left(F(\overline{p} + \sum a_k e_k + \sum b_k e_k)\right) \tag{5.2}$$

$$\frac{db_j}{dt} = e_j \cdot \left(F(\overline{p} + \sum a_k e_k + \sum b_k e_k)\right). \tag{5.3}$$

Except for the absence of the explicit scaling parameter ϵ, this system of ODEs is now in the form (2.1). Its absence does not affect the dynamics of the averaged ODE, but does potentially affect the scaling of the diffusion term in the Hasselmann SDEs, as described below.

Recall from last chapter that the system is fully determined over $128 \times 39 \times 2 + 3$ grid points. Thus, the rank of the covariance matrix from which the EOFs are derived (and thus the number of EOFs that determine the system) is $128 \times 39 \times 2 + 3$. The QG PDEs are projected onto these EOFs to give a system of $128 \times 39 \times 2 + 3$ ODEs. As there are only a few slowly-evolving PCs (a_i) in the KRG05 model, there are almost $10,000$ fast PCs.

For a broad range of values of the bottom drag parameter, a larger reduction is achieved

with EOFs computed on the bulk streamfunction, p, relative to EOFs computed on the barotropic and baroclinic streamfunctions individually. In particular, at high spin-down timescales at which there are large meridional displacements of the jet, the dynamics is strongly equivalent barotropic. Thus, the number of slowly-evolving bulk EOFs is half the number of slowly-evolving barotropic and baroclinic EOFs.

We now turn to the details of the application of the reduction methods to the KRG05 model.

5.1 Application of the Hasselmann method to the KRG05 model

5.1.1 (A) approximation

The averaged equation for the slow mode only can be written succinctly as:

$$\frac{da_i}{dt} = \int e_i \cdot \left[F\left(\overline{p} + \sum a_k e_k + \sum b_k e_k \right) \right] \, d\mu_a(b). \tag{5.4}$$

As discussed in Chapter 2, the measure μ_a cannot be analytically determined as a function of the slow mode for all but the simplest systems (of which the KRG05 model is not one). Although an explicit ODE cannot be found, by ergodicity of the fast dynamics:

$$\frac{db_j^a}{ds} = e_j \cdot \left[F\left(\overline{p} + \sum a_k e_k + \sum b_k^a(s) e_k \right) \right], \tag{5.5}$$

the effective tendency forcing for a particular value of the slow mode can be computed as the time-average over a realisation of the fast equation:

$$\overline{f}_i = \lim_{T \to \infty} \frac{1}{T} \int_0^T e_i \cdot \left[F\left(\overline{p} + \sum a_k e_k + \sum b_k^a(s) e_k \right) \right] \, ds. \tag{5.6}$$

The averaged equation becomes:

$$\frac{da_i}{dt} = \overline{f}_i(a). \tag{5.7}$$

Thus, the numerical integration of the effective system is divided into an integration of the effective slow equation and an integration of the fast equation. As discussed in Chapter 3, for sufficiently large scale separation and small macrotime step, a good approximation of

the averaged solution is derived with a weather integration interval that is less than the macrotime step; that is, the averaged ODE is more efficiently integrated than is the full climate-weather model.

It might seem necessary to resolve all the fast modes in order to carry out the numerical integration, given that the fast values are to be averaged out of the slow equation. Since there are almost $10,000$ fast PCs in the KRG05 model, the cost of resolving the fast modes would be enormous. However, the averaged slow tendency forcing can be found without explicitly resolving the fast modes, using only the original full model PDE and the slow modes ODE. To see this, note that the total model tendency forcing function F and its argument are the same in (5.6) and (5.5). The function $F(\overline{p} + \sum a_k e_k + \sum b_k^a(s) e_k)$ is simply determined from integration of the full model with the slow modes (a_i) fixed, without explicitly resolving the fast modes. Iterates of F, not the fast modes, are saved, to be projected onto the slow modes and time-averaged.

The alterations to the original model code to implement such a scheme are minimal. The scheme can be made more versatile because the *explicit* full model tendency forcing F is not needed. Rather, $\partial p/\partial t$ is evaluated directly, as shown in the algorithm outline below. This versatility is of practical significance, as F is not explicitly defined in the KRG05 code, nor in models of similar or greater complexity for which it may be impossible to explicitly derive F. By comparison, application of the MTV method requires that F is explicitly derived.

Seamless algorithm for integrating averaged (A) equation

Initial time: $t = t_0$, slow mode: $a_i(t_0)$, streamfunction: $p(0)$

$e_i \cdot (p(0) - \overline{p}) = a_i(t_0)$

Integration of weather equation

$p(0)$

$p^*(\delta t)$

(advance streamfunction one microtime step, according to full model)

$p(\delta t) = p^*(\delta t) - \sum e_i \cdot [p^*(\delta t) - p(0)]$

(subtract out slow and add in original slow to hold slow variables fixed)

$p^*(2\delta t)$

$p(2\delta t) = p^*(2\delta t) - \sum e_i \cdot [p^*(2\delta t) - p(0)]$

$\vdots$

$p(N\delta t) = p^*(N\delta t) - \sum e_i \cdot [p^*(N\delta t) - p(0)]$

Integration of effective climate equation

$a_i(t_0 + \Delta t) = a_i(t_0) + e_i \cdot \left[\frac{1}{N - N_1} \sum_{j=N_1+1}^{N} \frac{p^*(j\delta t) - p((j-1)\delta t)}{\delta t}\right] \Delta t$

Next integration of weather equation

$p(0) = p(N\delta t) - \sum a_i(t_0) e_i + \sum a_i(t_0 + \Delta t) e_i$

5.1.2 (N) approximation

There is very little that is added to the averaging algorithm to integrate the (N) equation. The saved tendency forcings, $e_i \cdot [(p^*(j\delta t) - p((j-1)\delta t))/\delta t]$, that are used to calculate the averaged tendency are also used to calculate the diffusion coefficient σ, according to formula (3.4). An accurate approximation of σ may require that the fast integration is carried out for large N, which would necessitate using a simplified scheme for this calculation. An extrapolation scheme is used here in computation of σ, as detailed in Chapter 6.

Due to the absence of an explicit ϵ in (5.3), there is uncertainty in the scaling of the

noise term. To see this, assume that the climate forcing scales as ϵ, such that:

$$\begin{aligned} \frac{da_i}{ds} &= e_i \cdot \left[F(\overline{p} + \sum a_k e_k + \sum b_k e_k) \right] \equiv \epsilon \mathcal{F}(a, b, \epsilon) \\ \frac{db_j}{ds} &= e_j \cdot \left[F(\overline{p} + \sum a_k e_k + \sum b_k e_k) \right] \equiv \mathcal{G}(a, b, \epsilon), \end{aligned} \tag{5.8}$$

where $\mathcal{F}$ and $\mathcal{G}$ are $\mathcal{O}(1)$. In theory, the effective climate equation is derived on coarse-grained time $t = s\epsilon$ (e.g., on the timescale of days from a timescale of hours) by substituting $\mathcal{F}$ into the formulae for drift (2.4) and diffusion (2.14). In practice, in which $\mathcal{F}$ is not derived and the scaling in ϵ is unknown, $e_i \cdot F = \epsilon \mathcal{F}$ is used in place of $\mathcal{F}$. Substituting $e_i \cdot F$ into (2.4) and (2.14), both the averaged drift $\overline{f}$ and the diffusion σ scale as ϵ and thus the noise term is underestimated by a factor of $\sqrt{\epsilon}$. There are additional uncertainties in the scaling of the noise because of uncertainties in the computation of ϵ itself (cf. Chapter 6). In Chapter 8, the Hasselmann stochastic equations are derived by assuming that the KRG05 model is scaled correctly in ϵ and that ϵ is correctly determined, but ϵ is also varied with these uncertainties in mind.

5.2 Application of the MTV method to the KRG05 model

As applied in FMV05 and FM06, the MTV method requires that the governing QG equations be of a sufficiently simple form; specifically, each operator must be written as a tensor. All operators in these two previous studies can be written as tensors because they are either linear ($\hat{L}$ and L), bilinear ($\hat{B}$ and B) or constant (G) operators. In symbolic form:

$$\begin{aligned} \frac{\partial p}{\partial t} &= F(p) \\ &= \hat{L}p + B(p, p) + \hat{G}. \end{aligned} \tag{5.9}$$

New operators for anomalies around the climatological base state are defined as follows: $L\cdot = \hat{L} \cdot + B(\cdot, \overline{p}) + B(\overline{p}, \cdot)$ and $G = \hat{G} + L(\overline{p}) + B(\overline{p}, \overline{p})$. Expanding p in the EOF basis, applying the new operators and projecting onto the EOFs yields a system in the slow (a)

and fast (b) variables:

$$\begin{aligned}\frac{da_i}{dt} &= e_i \cdot \left[L\left(\sum a_k e_k + \sum b_k e_k\right) + B\left(\sum a_k e_k + \sum b_k e_k, \sum a_k e_k + \sum b_k e_k\right) + G\right] \\ \frac{db_j}{dt} &= e_j \cdot \left[L\left(\sum a_k e_k + \sum b_k e_k\right) + B\left(\sum a_k e_k + \sum b_k e_k, \sum a_k e_k + \sum b_k e_k\right) + G\right]\end{aligned} \tag{5.10}$$

The KRG05 model cannot be written entirely in the form (5.10) because its forcing includes a non-quadratic nonlinear expression for the wind speed function (4.6), which cannot be written as a tensor. The limitation of the MTV method to multilinear systems in its present form presents a practical limitation of its applicability. For example, many subgrid scale processes in high-resolution GCMs are parameterised by non-quadratic nonlinear equations (e.g., parameterisations with thresholds). Tensor fields can be used to approximate such functions, but the computational requirements can be prohibitive in calculating individual effective tensors, much less a field of effective tensors. In the KRG05 model, this forcing term has been replaced by its linear best fit of a constant function, without significant ramifications to the model dynamics. The primary influence of this approximation is a slight shift of the dynamics towards higher spin-down timescales. The multilinear approximation of the KRG05 model is used in application of both the MTV and Hasselmann's methods for the sake of direct comparison; however, there is no requirement that Hasselmann's method be applied to a multilinear system.

With the approximation of constant wind-stress, the now bilinear KRG05 model equations can be written in the form (FM06):

$$\begin{aligned}\dot{a}_i(\tau) &= H_i^a + \sum_j L_{ij}^{aa} a_j(\tau) + \sum_{jk} B_{ijk}^{aaa} a_j(\tau) a_k(\tau) \\ &\quad + \sum_j L_{ij}^{ab} b_j(\tau) + 2\sum_{jk} B_{ijk}^{aab} a_j(\tau) b_k(\tau) + \sum_{jk} B_{ijk}^{abb} b_j(\tau) b_k(\tau) \\ \dot{b}_i(\tau) &= H_i^b \\ &\quad + \sum_j L_{ij}^{ba} a_j(\tau) + \sum_j L_{ij}^{bb} b_j(\tau) + \sum_{jk} B_{ijk}^{baa} a_j(\tau) a_k(\tau) + 2\sum_{jk} B_{ijk}^{bab} a_j(\tau) b_k(\tau) \\ &\quad + \sum_{jk} B_{ijk}^{bbb} b_j(\tau) b_k(\tau).\end{aligned} \tag{5.11}$$

The first superscript in the coefficient factors refers to the variable on the LHS of the equation, and the subsequent superscripts refer to the variables on which the operator acts. The forcings are assumed to act on three different timescales in ϵ: the tensor representing nonlinear self-interaction of fast modes, B^{bbb} acts on the $\mathcal{O}\left(\epsilon^2\right)$ timescale, the linear and bilinear terms involving both slow and fast modes act on the $\mathcal{O}\left(\epsilon\right)$ timescale and the terms involving only slow modes and the forcing terms act on the $\mathcal{O}\left(1\right)$ timescale. Another timescale is introduced in FMV05, in which it is assumed that the time-independent forcing H acts on the longest $\mathcal{O}(\epsilon^{-1})$ timescale.

The operator tensor coefficients are derived by isolating the associated spatial terms from the time-dependent PCs. For example, the ij^{th} tensor coefficient of the superviscosity (L^{SV}) operator is determined from:

$$e_i \cdot \begin{bmatrix} \Delta^{-1}(A_H \nabla^6(a_j e_j)) \\ \mathcal{L}^{-1}(A_H \nabla^6(a_j e_j)) \end{bmatrix} = e_i \cdot \begin{bmatrix} \Delta^{-1}(A_H \nabla^6(e_j)) \\ \mathcal{L}^{-1}(A_H \nabla^6(e_j)) \end{bmatrix} a_j \equiv L_{ij}^{(SV)} a_j. \tag{5.12}$$

The bilinear operator tensor derived as above is symmetrised; i.e., $B_{ijk} = B_{ikj}$, without loss of generality. To see this, let $\hat{B}$ denote the (generally asymmetric) bilinear operator tensor derived as above and let φ be the vector of PCs. Then $\varphi^T \hat{B}_i \varphi$ is the tendency forcing of the i^{th} PC by the bilinear term. Since $\varphi^T \hat{B}_i \varphi$ is a scalar, it equals its transpose; i.e., $\varphi^T \hat{B}_i \varphi = \varphi^T \hat{B}_i^T \varphi$. Thus, $\varphi^T \hat{B}_i \varphi = \varphi^T \left[(\hat{B}_i^T + \hat{B}_i)/2 \right] \varphi \equiv \varphi^T B_i \varphi$. Note that in each of the appendices of FMV05 and FM06, only $\hat{B}$ is defined and it is incorrectly identified as B.

The effective climate equation for a model of the form (5.11) is derived by substituting the appropriate terms from (5.11) into the equations for the effective drift (2.47) and diffusion (2.46). As all forcings are tensors, the climate variable a can be isolated in the drift and diffusion equations, leaving time-integrals and time-averages of (multilinear) functions of the weather variables. These latter expressions comprise the effective climate tensors.

For the bilinear system, the first four moments of the weather variable are needed in determining the effective climate tensors. Note that the fourth moment is the highest moment because a product of bilinear tensors (e.g., a product of polynomials of degree 2) yields at most a product of four weather variables. The distribution of the weather variable

is determined from the limiting weather equation:

$$\dot{c}_i = \sum_{jk} B^{bbb}_{ijk} c_j c_k. \tag{5.13}$$

Simplifications are needed in practice in computation of these moments. In particular, the limiting weather statistics are approximated in FMV05 and FM06 by the statistics of b of the unreduced model. Additionally, b is assumed to be Gaussian, thus requiring computation of only the first two weather moments. In the KRG05 model, this latter assumption is a reasonable assumption for the synoptic-scale modes. Even with these simplifications, the computation of the climate tensors is enormously costly in higher dimensional models, including the intermediate-level complexity KRG05 model. However, this computation is done only once offline (i.e., before integration of the MTV SDE).

Under these assumptions, the formulae for the effective climate tensors of the climate-weather equation (5.11) are derived in FMV05, such that the effective KRG05 climate equation is found by plugging into these expressions the appropriate information from the KRG05 model. In its general form, the effective climate SDE is given by:

$$\begin{aligned} da_i(t) &= H^a_i dt + \sum_j L^{aa}_{ij} a_j(t) dt + \sum_{jk} B^{aaa}_{ijk} a_j(t) a_k(t) dt && (5.14)\\ &\quad + \tilde{H}_i dt + \sum_j \tilde{L}_{ij} a_j(t) dt && (5.15)\\ &\quad + \sum_{jk} \tilde{B}_{ijk} a_j(t) a_k(t) dt && (5.16)\\ &\quad + \sum_{jkl} \tilde{M}_{ijkl} a_j(t) a_k(t) a_l(t) dt && (5.17)\\ &\quad + \sqrt{2} \sum_j \sigma^{(1)}_{ij}(a(t)) dW^{(1)}_j + \sqrt{2} \sum_j \sigma^{(2)}_{ij} dW^{(2)}_j. && (5.18) \end{aligned}$$

Thus, in the bilinear system, climate-weather interactions induce time-independent, linear, bilinear and cubic noise-induced drifts and additive and multiplicative noises. The bare truncation forcing and noise-induced drift together comprise the effective drift. Assuming $H = \epsilon H$ in the parent equations (5.11) implies that H^a disappears; that is, the bare truncation time-independent forcing acts on such a long timescale that it does not impact

the $\mathcal{O}(1)$ climate dynamics. In application to the KRG05 model, the MTV approximations are found with and without H.

Because of the potentially stringent assumptions of the theory underlying the MTV method, including the requirement of a large scale separation, the effective slow dynamics will generally differ from the slow dynamics of the unreduced system. To improve the accuracy of the MTV reduction, a fitting scheme was developed and applied in FM06 (called a 'minimal regression fitting' in FM06) by which tendency forcing terms are scaled according to their physical origin:

$$
\begin{aligned}
da_i(t) &= \lambda_B \left[H_i^a dt + \sum_j L_{ij}^{aa} a_j(t) dt + \sum_{jk} B_{ijk}^{aaa} a_j(t) a_k(t) dt \right] && (5.19)\\
&+\lambda_A^2 \sum_j \tilde{L}_{ij}^{(2)} a_j(t) dt + \lambda_A \sqrt{2} \sum_j \sigma_{ij}^{(2)} dW_j^{(2)} && (5.20)\\
&+\lambda_M^2 \left[\sum_j \tilde{L}_{ij}^{(3)} a_j(t) dt + \sum_{jkl} \tilde{M}_{ijkl} a_j(t) a_k(t) a_l(t) dt \right] && (5.21)\\
&+\lambda_L^2 \left[\sum_j \tilde{L}_{ij}^{(1)} a_j(t) dt \right] && (5.22)\\
&+\lambda_M \lambda_L \left[\tilde{H}_i^{(1)} dt + \sum_{jk} \tilde{B}_{ijk} a_j(t) a_k(t) dt \right] && (5.23)\\
&+\lambda_A \lambda_F \tilde{H}_i^{(2)} dt && (5.24)\\
&+\sqrt{2} \sum_j \sigma_{ij}^{(1)}(a(t)) dW_j^{(1)}, && (5.25)
\end{aligned}
$$

where the multiplicative noise matrix $\sigma^{(1)}$ satisfies

$$
\lambda_L^2 Q_{ij}^{(1)} + \lambda_L \lambda_M \sum_k U_{ijk} \alpha_k(t) + \lambda_M^2 \sum_{kl} V_{ijkl} \alpha_k(t) \alpha_l(t) = \sum_k \sigma_{ik}^{(1)} \left[\alpha(t)\right] \sigma_{jk}^{(1)} \left[\alpha(t)\right].
$$

Scaling parameters are assigned to unreduced model tensors based on the interactions which give rise to the effective climate tensors. Naturally, the bare truncation tensors are grouped together, and are each assigned the scaling parameter λ_B. There are two "triad interactions" which induce noise-induced drift and noise terms in the effective climate equation. B^{abb} and B^{bab}, which interact in generating additive noise and a linear correction, are each

assigned the scaling parameter λ_A. The scaling parameter λ_M is assigned to B^{aab} and to B^{baa}, the interaction between which gives rise to a cubic noise-induced drift and which contributes to multiplicative noise. The linear tensors L^{ab} and L^{ba} couple to induce a linear noise-induced drift and contribute to multiplicative noise, and are each assigned the scaling parameter λ_L. The scaling parameter λ_F is assigned to the remaining tensor, L^{bb}. There are additional interactions between triads and linear tensors, which give rise to time-independent forcings and quadratic non-linearities, and contribute to multiplicative noise. Although these interactions are distinct from the above-described interactions, no new scaling parameters are assigned. Thus, the scaling parameters are assigned based on only a subset of the interactions responsible for the effective climate equation.

However, the MTV scaling parameters can be set to correct for the scaling errors arising from the absence of an explicitly-resolved ϵ in the KRG05 model. In finding the effective Hasselmann stochastic equations, this problem can be resolved by an appropriate time rescaling of the system. In finding the MTV equations, this problem is not resolved by a time rescaling of the system, as ϵ serves not only as the scale separation parameter between climate and weather variables, but also as the scale separation parameter between climate forcings and as the scale separation parameter between weather forcings. On the coarse-grained time τ, the MTV ansatz for the KRG05 model is represented by (cf. (5.11)):

$$\begin{aligned}
\dot{a}_i(\tau) &\equiv \mathcal{F}_1(a,a,\epsilon) \\
&\quad + \frac{1}{\epsilon}\mathcal{F}_0(a,b,\epsilon) \\
\dot{b}_i(\tau) &\equiv H_i^b \\
&\quad + \frac{1}{\epsilon}\mathcal{G}_0(a,b,\epsilon) \\
&\quad + \frac{1}{\epsilon^2}\mathcal{G}_1(b,b,\epsilon),
\end{aligned} \tag{5.26}$$

where $\mathcal{F}_0$, $\mathcal{F}_1$, $\mathcal{G}_0$ and $\mathcal{G}_1$ are *assumed* to be $\mathcal{O}(1)$. The effective climate feedback (i.e., all effective forcings except the bare truncation forcing) is found by substituting $\mathcal{F}_0$ and $\mathcal{G}_0$ into the formulae for drift (2.47) and diffusion (2.46). Since these $\mathcal{O}(1)$ forcings are not derived in practice, the $\mathcal{O}(\epsilon^{-1})$ climate-weather interaction tensors from (5.11) are used

in their place. However, the effective climate feedback can be correctly scaled by setting $\lambda_A = \lambda_L = \lambda_M = \lambda_F = \epsilon^{-2}$. Indeed, it will be demonstrated in Chapter 7 that a tuning of either the bare truncation forcing or the effective climate feedback often yields the optimal fitting.

Chapter 6

Preliminary tests and analyses

In Section 6.1, the components and parameters needed for application of the reduction methods to the KRG05 model are derived. All parameters are in fact used only in integration of the Hasselmann stochastic equations. In Section 6.2, the 'Seamless algorithm for integrating averaged (A) equation' developed in Chapter 5 is used to run a modified form of the unreduced KRG05 model, with the goal of understanding the dynamics underlying jet bimodality.

We first establish some terminology that will be used in this and successive chapters. The climate and the weather timescales are the timescales of the climate and weather variables in (2.1). Also, recall from Chapter 4 that the propagating wave-4 mode is defined to consist of two EOF-PC pairs and the stationary mode consists of one EOF-PC pair. If these two modes alone determine the slow dynamics, then the slow dynamics is called the '3-variable slow dynamics' because it is determined by the evolution of three PCs associated with the two modes. As well, the 'stationary mode' without qualification refers to the mode which is most associated with the meridional shifts of the jet. Another stationary mode is the jet-intensity mode, but it is only important at the lowest spin-down timescales, $k^{-1} < 2.0$ days.

6.1 Components and parameters of the reduced equations

The reduction theorems are based on a number of assumptions that guarantee the convergence of the statistics of the full model slow variable to the statistics of the effective model (slow) variable, in the limit of infinite timescale separation. These assumptions include the

existence of limiting slow dynamics and ergodicity and mixing of the fast dynamics. Due to the complexity of geophysical models, these assumptions cannot be rigorously verified in such models, nor can they even all be thoroughly verified numerically. For example, ergodicity of the weather dynamics at a particular climate state can only be tested numerically, and then only for a small sample of fixed climate states.

If these assumptions are satisfied, then the limiting solution will be a good approximation of the climate solution if there is sufficient timescale separation between the slow and fast modes of the system. Thus, a (testable) assumption underlying the practical application of these theorems is the existence of such a timescale separation. In fact, the 'sufficient' quality of the timescale separation must be evaluated subjectively, and then, based on the timescales of the modes comprising the system, the division can be made into slow and fast modes and the timescale separation parameter ϵ can be computed. The parameter ϵ appears as an explicit scaling parameter in the Hasselmann stochastic equations only. Other parameters are needed in practice for the efficient computation of the Hasselmann stochastic equations. These parameters are related to the mixing property of the weather dynamics.

The primary feature of interest of the KRG05 model is the dynamics associated with the bifurcation to a bimodal jet distribution. The various parameters are determined at values of the spin-down timescale above and below the bifurcation point at $k^{-1} = 5.6$ days: at $k^{-1} = 2.3, 4.7, 6.7$, and 12.0 days. Multiple regime behaviour is pronounced at $k^{-1} = 6.7$ days, and this spin-down timescale is sufficiently removed from the bifurcation point at $k^{-1} = 5.6$ days to avoid errors associated with numerical approximation in a small neighbourhood of the bifurcation point. Aside from understanding the dynamics, simulation of the bimodal distribution in particular is a strict test of reduction efficacy.

6.1.1 Timescale separation

Autocorrelation timescale of y and value of the scale separation parameter ϵ

The definition of the timescale separation parameter, ϵ, should be consistent with the premise underlying these mode reduction methods that there is a timescale on which the weather statistically equilibrates and the climate does not significantly change. A standard

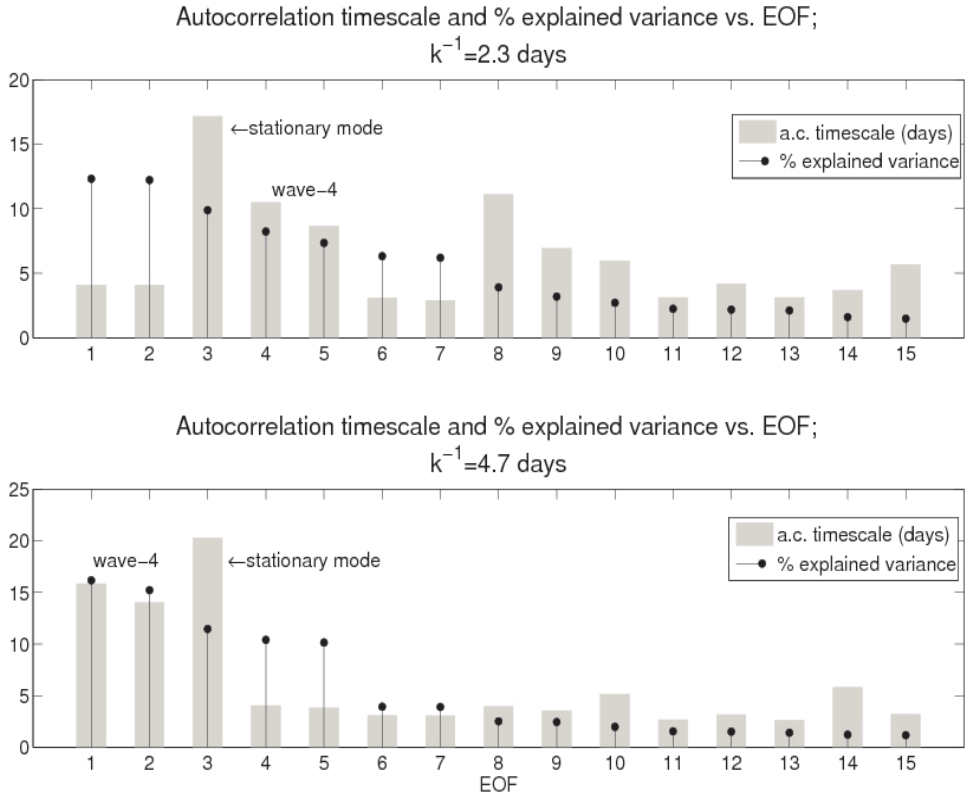

Figure 6.1: Autocorrelation timescale (days) and percentage of explained variance at values of the spin-down timescale at which there is no jet bimodality.

operational measure, from e.g., Majda et al. (2001) and Berner (2005), is the ratio of autocorrelation timescales of weather/fast modes to climate/slow modes. In principle, a very small ϵ is consistent with the situation in which the weather decorrelates much faster than does the climate. As in Berner (2005), the autocorrelation timescale of a principal component (PC), α_i, is defined as:

$$\tau_{\alpha_i} = \frac{1}{<(\alpha_i)^2>} \int_0^\infty |<\alpha_i(s)\alpha_i(0)>| \, ds.$$

The autocorrelation timescales of the first few hundred PCs are examined in the KRG05 model because of the possibility of slowly-evolving EOFs that explain little of the variance.

Plots of autocorrelation timescale and explained variance at values of bottom drag relevant to the bifurcation dynamics, Figure 6.1 and Figure 6.2, demonstrate a seemingly clear division into slow and fast modes. At the three highest values of spin-down timescale,

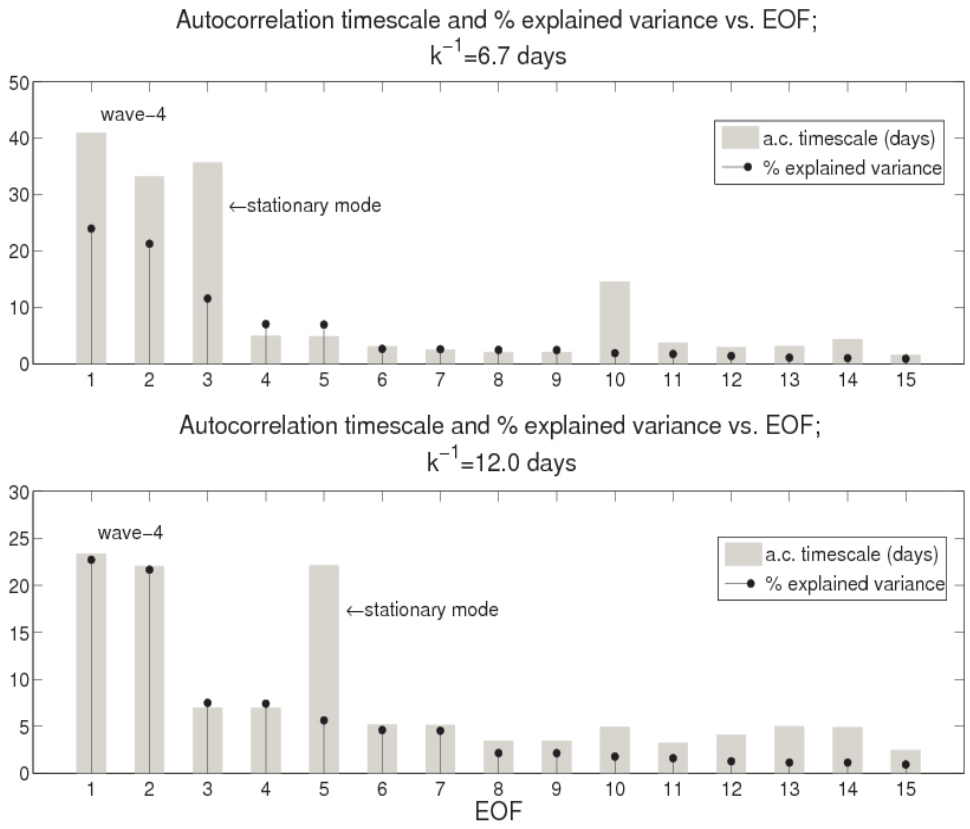

Figure 6.2: Autocorrelation timescale (days) and percentage of explained variance at values of the spin-down timescale at which there is jet bimodality

the propagating wave-4 mode, consisting of two EOFs in spatial and temporal quadrature, and the stationary mode, primarily describing the meridional shifts of the jet, appear to comprise the set of slow modes and the rest of the modes comprise the fast modes. In KRG05, these two modes alone were determined to be of first-order importance to the bifurcation dynamics, generating bimodality through nonlinear interaction between the modes on long timescales. At $k^{-1} = 6.7$ days, in particular, there is an order of magnitude difference between the autocorrelation timescales of these two slow modes and the fast modes (except EOF-10). Thus, by defining these two slow modes alone as the climate modes, a reasonable choice for the scale separation parameter is $\epsilon = 0.1$, by which the noise-induced drift and noise terms in the Hasselmann approximations are scaled. The value of ϵ is larger at $k^{-1} = 4.7$ days and at $k^{-1} = 12.0$ days. At $k^{-1} = 2.3$ days, another stationary EOF (EOF-8) that is primarily associated with the jet intensity, is also slowly evolving. In

Spin-down timescale	Values of ϵ (rounded)	
	1-variable	3-variable
$k^{-1} = 2.3$ days	0.66	0.33
$k^{-1} = 4.7$ days	0.75	0.25
$k^{-1} = 6.7$ days	1	0.1
$k^{-1} = 12.0$ days	1	0.25

Table 6.1: Values of scale separation parameter ϵ

FMV05 and FM06, the timescale separation ratio is defined as the ratio of the slowest fast mode to the fastest slow mode. If EOF-8 is included in the set of weather modes, then using this definition the scale separation parameter is unity. However, since this EOF may be dynamically unimportant, as it is at higher spin-down timescales (KRG05), it may not have an influence in setting the autocorrelation timescale of the weather dynamics as a whole. If this is the case, then a better estimate of the scale separation parameter is $\epsilon = 0.33$, between the slow stationary and wave-4 modes and the leading two fast propagating waves, wave-5 (EOF-1 and EOF-2) and wave-6 (EOF-6 and EOF-7). Although the definition from the MTV studies eliminates ambiguity in determining the value of ϵ, it tends to overestimate its value, leading to overestimation of noise in the Hasselmann approximations. However, since ϵ does not appear in the MTV approximation, this overestimation is unimportant in calculation of the MTV SDE.

As will be demonstrated in Chapter 7 and Chapter 8, there are circumstances in which it is sensible to restrict the climate variables to the stationary mode only, and to include wave-4 among the weather modes. This separation is motivated not by a timescale separation, but rather by the (non-)existence of limiting dynamics (i.e., the 1-variable reduced models are considered because they yield bounded solutions). Table 6.1 gives the estimated scale separation values in the case that only the stationary mode (1-variable) is defined as a climate mode and in the case that the stationary and wave-4 modes (3-variable) are defined as climate modes. In the 1-variable case, the timescale separation parameter is defined as the ratio of timescales between the wave-4 and stationary modes (consistent with the definition in FMV05 and FM06).

Autocorrelation timescale of y^x

In Fatkullin and Vanden-Eijnden (2004), it is stated that "the unambiguous test to determine if the [fast modes] are fast is to compute their ACFs [autocorrelation functions] at fixed $X = x$ (that is, the autocorrelation function of y^x)". The model considered in this earlier study includes an explicitly defined timescale separation parameter, ϵ. In one particular test, $\epsilon \ll 1$ and yet it was found that the fast and slow ACFs are nearly identical. Due to the significant modulation of the fast variable by the slow variable, the fast variable evolves on both fast and slow timescales. The ACF determined from a long time-integration captures the statistics on the slow timescale only.

Since the explicit dependence of y^x on x is unknown in the much more complex KRG05 model, it is only practical to examine snapshots of the autocorrelation timescales of y^x for a few values of x. Nevertheless, it is informative to apply this test to the KRG05 model, particularly at $k^{-1} = 6.7$ days. At this spin-down timescale, the autocorrelation timescale of y^x, with *x=(stationary mode, wave-4 mode)*, is roughly similar to that of y (for which the slow variable is not fixed). With *x=(stationary mode)*, the autocorrelation timescale of the wave-4 mode, which no longer is fixed in y^x, is of particular interest. Again, the autocorrelation timescales of wave-4 at y^x and at y are similar. It thus appears that at $k^{-1} = 6.7$ days the wave-4 mode is indeed slow and not a "hidden" fast variable that appears slow because of its modulation by the slowly-evolving stationary mode (Fatkullin and Vanden-Eijnden (2004)). This is surprising because, as will be shown, the 1-variable reduced model in the stationary mode alone yields the jet bimodality of the unreduced model despite the fact that the slowly-evolving wave-4 mode is assigned to the set of fast modes in derivation of the 1-variable reduced equation.

In fact, at each of $k^{-1} = 2.3$, 4.7, 6.7 and 12.0 days, y and y^x have similar autocorrelation timescales with *x=(stationary mode)*, with *x=(wave-4)* and with *x=(stationary mode, wave-4 mode)*. Thus, the stationary and wave-4 modes are slow modes of the system from very low to very high values of bottom drag. Since there is good scale separation between these modes and the other modes in the system, it is likely that the effective climate modes (which determine the reduced climate equations) are contained in the space defined by these two

modes.

6.1.2 Mixing of weather dynamics

Ergodicity and mixing of the fast, weather dynamics are key assumptions in the reduction to a stochastic equation in the climate modes alone (mixing is not needed in reduction to Hasselmann's deterministic averaged (A) equation). These two properties are extremely difficult to evaluate in simple systems, much less in climate models. Nevertheless, there are parameters related to the mixing of the weather dynamics that must be determined prior to integrating Hasselmann's stochastic equations. Specifically, the scaling of the diffusion term σ (i.e., the integrated lag-covariance up to a long time-lag) is needed in the additive noise (W) model and the scaling factor separating the integrated lag-covariances up to short and long time-lags is needed for an efficient integration of the (N) model.

The noise term of the (N) model is given by $\sqrt{\epsilon}\sigma dW/dt$, where ϵ is the scale separation parameter and σ is the state-dependent (i.e., climate-dependent) diffusion matrix. By comparison, the noise term of the (W) model is given by $\sqrt{\epsilon}\gamma I dW/dt$, where γ is a scaling factor and I is the identity matrix. It is natural then to define γ as the matrix norm of σ.

The diffusion matrix σ has a theoretically-justified formula, which for practical implementation must be simplified. Recall from Chapter 3 that the lag-covariance is calculated in a cumulative sense, but not the time-integral component of the time-integrated lag-covariance. With a rapidly mixing weather component, σ can be efficiently computed without simplification because the lag-covariance tends to zero sufficiently fast. In the KRG05 model, the weather modes may mix sufficiently rapidly to ensure the validity of the stochastic approximations, but their decorrelation timescale is on the order of days, which would exceed the greatest possible macrotime step in an (accurate) integration of the effective equation. To efficiently integrate the effective system, the weather equation at each climate realisation is integrated up to a time-lag that is less than the macrotime step and the integrated lag-covariance up to this time-lag is multiplied by an a priori determined extrapolation factor to determine the covariance integral.

Consider the KRG05 model at $k^{-1} = 6.7$ days, with only the stationary mode defined as a climate mode. For a particular upper lag, a time series of the integrated lag covariance is

Spin-down timescale	γ		Extrapolation factor	
	1-variable	3-variable	1-variable	3-variable
$k^{-1} = 2.3$ days	20	50	6	6
$k^{-1} = 4.7$ days	25	75	6	6
$k^{-1} = 6.7$ days	20	50	6	5
$k^{-1} = 12.0$ days	20	50	6	5

Table 6.2: Values of the norm of σ (γ) and the extrapolation factor, determined as the ratio of the norm of the lag-covariance integrated up to two days to that integrated up to 1 hour.

found using the stationary mode (i.e., climate) values from an integration of the unreduced model. Generally, the norm of the integrated lag-covariance is computed for each climate realisation, but here the integrated lag-covariance is just a (positive) scalar. Several such time series are found, yielding a series of time series of integrated lag covariances indexed by the upper time-lag. As depicted at top in (Figure 6.3), this series appears to converge at a lag timescale of approximately two days. The average of this convergent series gives the scaling factor γ in the 1-variable (W) model at $k^{-1} = 6.7$ days. In the 1-variable (N) model, it would be prohibitively costly to integrate for two days the weather equation at each climate realisation, as the macrotime step of the effective climate SDE is on the order of hours (and multiple ensembles are needed in practice to sufficiently sample the weather attractor). However, the upscaled (by a factor of 6) time-series of lag-covariances integrated up to only 1 hour coincides with the series integrated up to 2 days, including the spike in these values (bottom plot in Figure 6.3). Thus, in deriving σ in the (N) model it is only necessary to integrate the weather equation at each climate realisation up to a lag of one hour.

This same procedure is carried out at each spin-down timescale for a climate defined to consist of the stationary mode only (1-variable case) and defined to consist of the stationary and wave-4 modes (3-variable case). In all cases, the extrapolation factor is determined as the ratio of the integrated lag-covariance for a time-lag of 2 days to that for a time-lag of 6 hours.

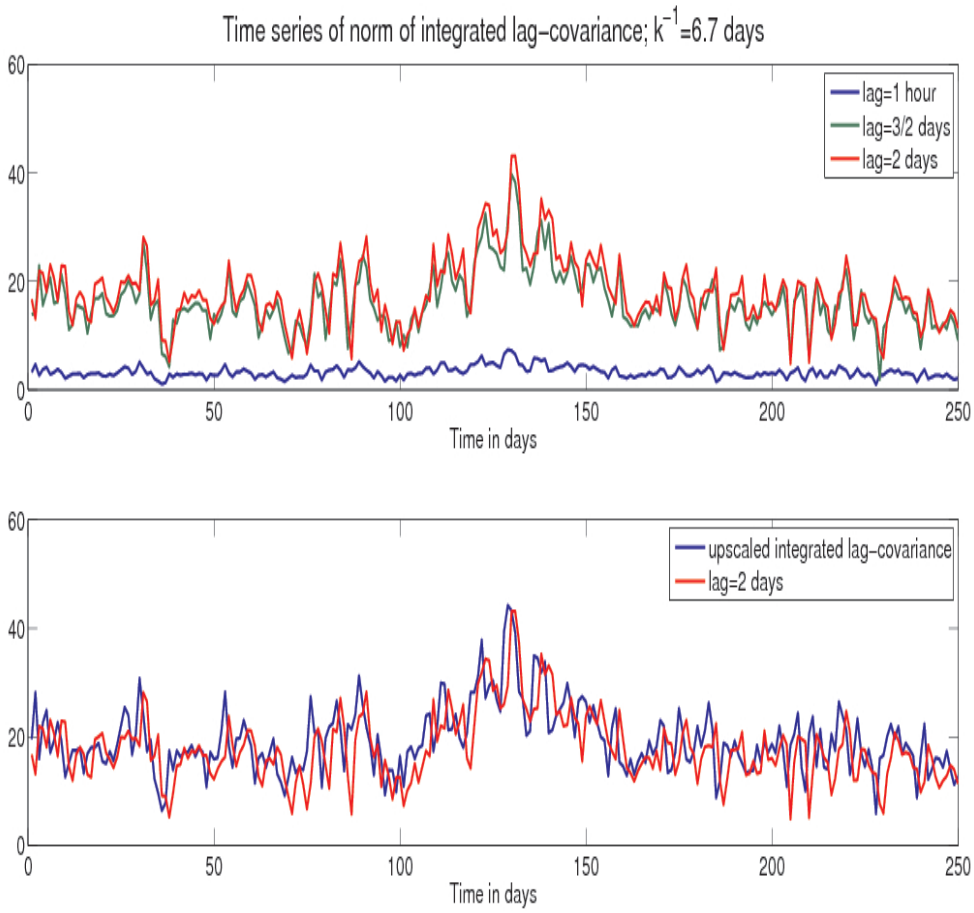

Figure 6.3: Time series of norm of integrated lag-covariance with only the stationary mode defined as the climate mode. Bottom: there is agreement between the lag-covariance integrated up to 1 hour and upscaled by a factor of 6 and the lag-covariance integrated up to two days.

6.2 Modified unreduced KRG05 models

In this section, the 'Seamless algorithm for integrating averaged (A) equation' from Chapter 5 is modified to isolate individual processes in the *unreduced* KRG05 model (and not the reduced (A) model) through the elimination of particular modes and interactions among modes to determine their effect on jet bimodality. A different understanding of the dynamics underlying jet bimodality than that in KRG05 is obtained.

6.2.1 The effect of particular interactions on jet bimodality

Using a seamless algorithm, it is a simple matter to directly test the effect of particular interactions on the model dynamics. In particular, the unreduced KRG05 model is derived

from the 'Seamless algorithm for integrating averaged (A) equation' by simply setting $R = N = 1$, $N_1 = 0$ and $\Delta t = \delta t$, and individual interactions between modes of interest are suppressed by subtracting out the tendency forcings that determine these interactions. Using the individual tendency forcing operators of the KRG05 model that have already been computed towards finding the MTV approximations, the primary modification to the existing averaging algorithm is to subtract out from the set of 'slow' ODEs the tendency forcings of interest. It is convenient but not necessary to use these explicitly determined tendency forcing operators.

For example, the stationary mode-wave-4 nonlinear interactions are removed from the systems as follows:

Unreduced model modified to suppress stationary mode-wave-4 nonlinear interactions

Initial time: $t = t_0$, slow mode: $a_i(t_0)$, streamfunction: $p(0)$

$e_i \cdot (p(0) - \overline{p}) = a_i(t_0)$

a_1 is the stationary mode PC and a_2 and a_3 are the wave-4 PCs

Fast integration $p(0)$

$p^*(\delta t)$

(advance streamfunction one microtime step, according to full model)

$p(\delta t) = p^*(\delta t) - \sum e_i \cdot [p^*(\delta t) - p(0)]$

(subtract out slow and add in original slow to hold slow variables fixed)

Effective slow integration

$a_i(t_0 + \Delta t) = a_i(t_0) + e_i \cdot \left[\frac{p^*(\delta t) - p(0)}{\delta t}\right] \Delta t - \sum_{i=1}^{25} \sum_{j=1}^{3} \sum_{k=1}^{3} B_{ijk} a_j(t_0) a_k(t_0) \Delta t$

(Subtract out the stationary-wave-4 Jacobian forcings from unreduced model

(approximated by projection of Jacobian forcings onto first 25 *EOFs))*

Next fast integration

$p(0) = p(\delta t) - \sum a_i(t_0) e_i + \sum a_i(t_0 + \Delta t) e_i$

Since it is desired to remove the stationary-wave-4 nonlinear interactions from the full KRG05 model (instead of just removing its projection onto leading modes), the ODE of 'slow' modes is expanded to include as many modes as is computationally feasible. In particular, these interactions are subtracted out from the first 25 EOFs, which explain over 90% of the variance, to approximate their removal from the full system. Note that the stationary mode-wave-4 nonlinear interactions are here defined to consist of the stationary mode self-interaction, the wave-4 self-interaction and the (nonlinear) advection involving both the stationary and wave-4 modes. The stationary-wave-4-wave-5 nonlinear interactions considered below are defined to consist of the self-interactions of all three modes and the advections involving the stationary and wave-4 modes, the stationary and wave-5 modes

and the wave-4 and wave-5 modes.

If jet bimodality persists without a particular set of nonlinear interactions, then these interactions are not necessary for its existence. Only nonlinear interactions are removed because bimodality is a nonlinear phenomenon (among slowly-evolving modes and/or between slowly-evolving and fast-evolving modes). It is concluded in KRG05 that the jet bimodality results from a "weak" nonlinear interaction (i.e., interaction on long timescales) between the two slowest-evolving modes, the stationary and wave-4 modes. However, the jet bimodality is fundamentally unchanged by subtracting out the stationary-wave-4 nonlinear interactions (Figure 6.4). Thus, the nonlinear interactions of these two modes are not responsible for jet bimodality. This determination is based on the similarity of the statistics, including bimodality of the stationary mode distribution, of the unreduced model with and without these interactions. The similarity is particularly evident at $k^{-1} = 7.7$ days, at which the jet bimodality is more easily discerned in the unreduced model because of the greater separation between regimes (and hence, for clarity, the results are examined at this spin-down timescale instead of at $k^{-1} = 6.7$ days (Figure 6.4)). Without these interactions, the low latitude regime (approximately correpsonding to PC values in the range $[-400, -200]$) becomes more favoured than in the unreduced model, but there clearly exist two regimes. As jet bimodality is robust to suppression of the nonlinear interactions of these two modes, these interactions are not responsible for the bimodality (in contrast to the conclusions of KRG05).

To further understand the secondary effect of the nonlinear interactions of the stationary and wave-4 modes on jet bimodality, the individual tendency forcings that comprise these interactions are examined. Suppressing only the advection forcings involving both the stationary mode and the wave-4 mode, the statistics of this modified model are virtually unchanged from those of the unmodified model. Removing the stationary mode self-interaction also has no impact on jet bimodality, but removing the wave-4 self-interaction results in a nearly symmetric distribution between the high and low latitude regimes (Figure 6.4). It is deduced in KRG05 that the wave-4 self-interaction broadens the climatological jet and facilitates meridional shifts by counteracting the stabilizing effect of the leading synoptic ed-

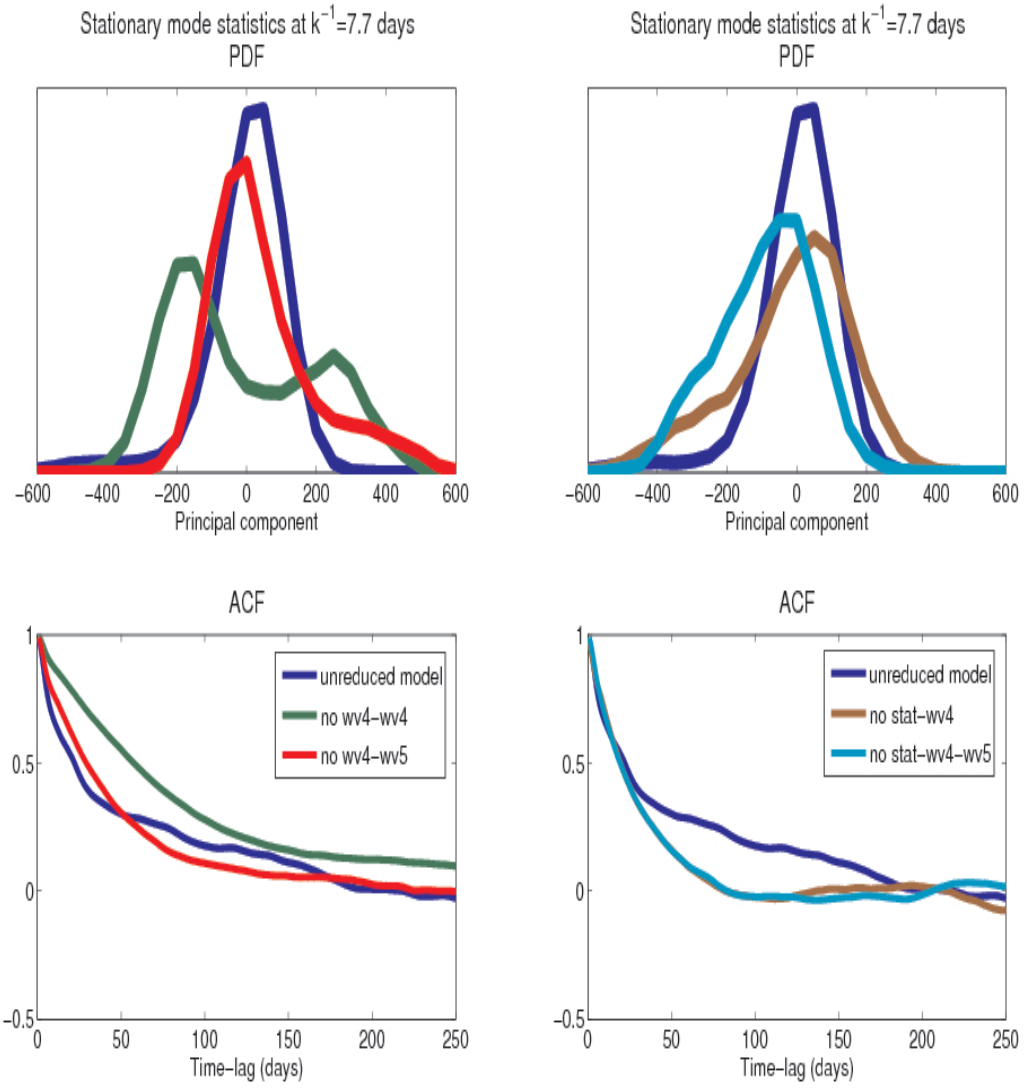

Figure 6.4: Statistics of the unreduced model at $k^{-1} = 7.7$ days with nonlinear interactions suppressed as indicated. In contrast to the conclusions of the KRG05 model, jet bimodality exists with the nonlinear interaction of the stationary and wave-4 modes suppressed (labeled: 'no stat-wv4').

dies. The results here suggest the opposite effect, that *removing* the wave-4 self-interaction counteracts the (supposed) stabilizing effect of these eddies on the high-latitude regime.

Of course, the secondary effect of the nonlinear interactions of the stationary and the wave-4 modes does not rule out the possibility that higher-order interactions involving these two modes are responsible for generating jet bimodality. Indeed, the regime behaviour is destroyed by eliminating the nonlinear interactions of three modes: the stationary mode, the wave-4 mode and the wave-5 mode, as evidenced by the quasi-Gaussian distribution of the stationary mode of the model without these interactions. Among the individual interactions that comprise the stationary mode-wave-4-wave-5 nonlinear interactions, it has already been established that the wave-4 self-interaction has an important secondary effect on the jet dynamics, although if it alone is removed, then the jet bimodality persists. Removing the stationary mode-wave-5 nonlinear interactions have an insignificant effect, but as a result of removing the wave-4-wave-5 nonlinear interactions, the low-latitude regime becomes favoured over the high-latitude regime, more so than just by removing the wave-4 self-interaction (Figure 6.4). Thus, interactions of the wave-4 and wave-5 modes act to stabilize the high-latitude regime, such that the high-latitude regime is favoured in the unreduced model.

Removing another synoptic eddy mode of potential importance (as determined in KRG05) - the wave-6 mode - does not impact the jet bimodality significantly. In fact, the nonlinear interactions of the stationary, wave-4 and wave-5 modes comprise the smallest subset of interactions that cannot be removed from the system without destroying the jet bimodality. Thus, it appears that jet bimodality is a result of the nonlinear interaction of three modes: the stationary, wave-4 and wave-5 modes.

6.2.2 The effect of particular modes on jet bimodality

It is even a simpler matter to remove a particular mode from the system. A particular mode is removed by simply suppressing it at each microtime step and each macrotime step, through simple manipulations of the seamless algorithm. However, in contrast to removing a particular nonlinear interaction, removing a leading mode significantly affects the mean and variance (and other moments) of the system. Moreover, it is not clear how

to renormalise the system to isolate the dynamical changes related to the removal of a particular mode. For example, the unreduced model without wave-4 does not generate jet bimodality, but the stationary mode distribution has a large, negative mean value and its variance is significantly decreased (not shown).

However, the low-frequency modulation of the stationary mode in the region without jet bimodality can be diagnosed with this method. At $k^{-1} = 4.7$ days, the stationary mode ACF of the KRG05 model without wave-5 decays exponentially and without wave-4 decays with oscillation, similar to the ACF of the unmodified model (Figure 6.5). The ACF is not significantly affected by the removal of wave-4 despite the fact that wave-4 carries the most variance in the system. It thus appears that wave-5 is primarily responsible for the low-frequency modulation of the stationary mode.

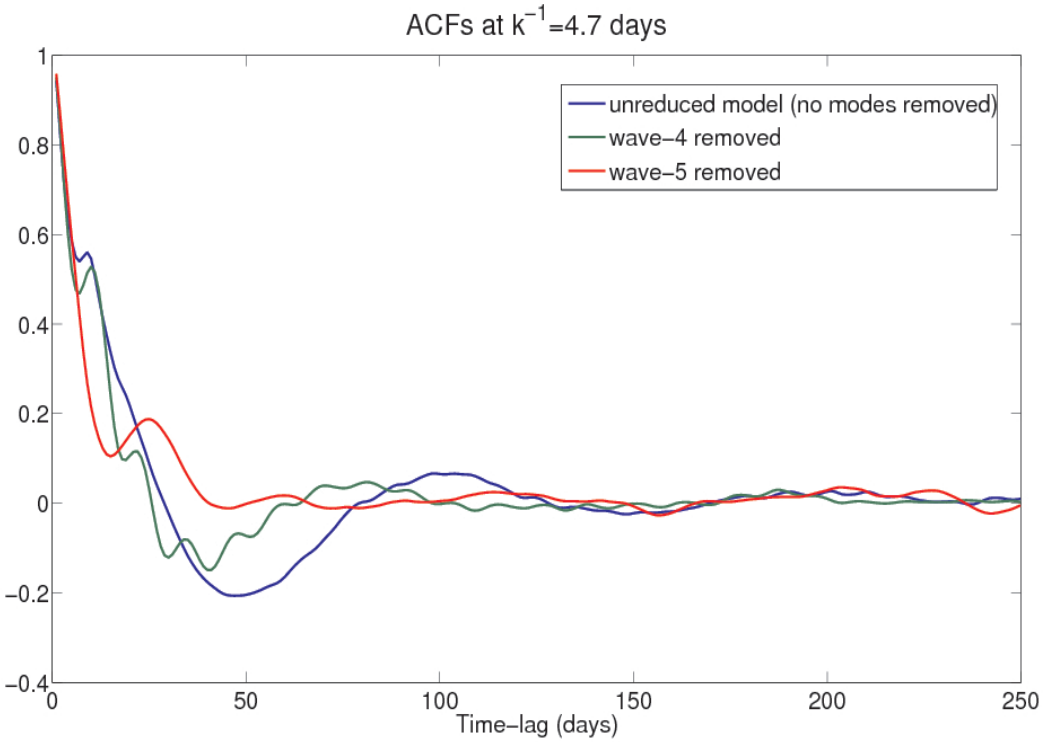

Figure 6.5: Stationary mode ACF of the KRG05 model with the indicated modes removed.

Chapter 7

Results of MTV method applied to KRG05 model

In this chapter, MTV approximations of the KRG05 model climate are computed, following the procedure detailed in Chapter 5. These approximations are found over a range of spin-down timescales (k^{-1} = 2.3, 4.7, 6.7 and 12.0 days), with the goal of better understanding the bifurcation dynamics that gives rise to jet bimodality. Jet bimodality is of particular interest because the meridional shifts of the zonal mean zonal flow by which it is characterised are roughly similar to the observed oscillations of the Arctic Oscillation (although regime behaviour is not obvious from observations). In KRG05, it was concluded that a long timescale, nonlinear interaction between the stationary and wave-4 modes is responsible for the jet bimodality. Since there is good scale separation between these two modes and the weather modes, it might be expected that the effective climate equation in these two modes will generate jet bimodality.

The MTV approximations are derived from rigorously justified limiting theorems, but because of departures from the underlying assumptions, a fitting scheme was developed in FM06 (called a 'minimal regression fitting' scheme in FM06) (cf. Chapter 5). Scaling parameters are a priori assigned to the climate-climate and the climate-weather interactions which determine the effective climate equation, and these parameters may be adjusted (away from a value of 1) in order to better simulate the climate statistics of the unreduced model. A number of corrections are possible with this fitting scheme, including the correction of climate drift (i.e., the potential non-zero mean of the climate PCs). In FMV05, the climate drift was eliminated by assuming that the time-independent bare truncation forcing H acts on the $\mathcal{O}(\epsilon^{-1})$ timescale, which eliminates it from the effective equation. In

FM06 (using a different model - the Marshall and Molteni model (1993) - than in FMV05), H was assumed to act on the $\mathcal{O}(1)$ timescale and the climate drift was eliminated through application of the MTV fitting scheme. In this chapter, the MTV models with and without nonzero H are considered and the fitting scheme is also applied.

In both FMV05 and FM06, the approximations that were derived under the assumption that the cross-correlations of the weather modes are zero:

$$\mathcal{B}_{ij}(s) = \lim_{T\to\infty} \frac{1}{T} \int_0^T b_i(t+s) b_j(t) dt \equiv 0 \tag{7.1}$$

for all $i \neq j$ do not significantly differ from the approximations without this assumption. The MTV approximations of the KRG05 model with assumption (7.1) are found in the region of jet bimodality (Section 7.1) and in the region without jet bimodality (Section 7.2). Within each of these sections, the untuned MTV approximations, derived with and without nonzero H, are presented first. The MTV tuning scheme is then applied to find the best-fit MTV approximations. In Section 7.3, the results without the assumption (7.1) on the weather statistics are discussed. The results of this chapter are summarised and discussed in Section 7.4.

As will be demonstrated, for certain spin-down timescales the 3-variable MTV model is a poor approximation to the climate variables of the full system. For this reason, it is interesting to consider other fast-slow separations: in particular, treating the stationary mode alone as the climate variable and including the wave-4 mode in the weather variables. In the following these reduced models will be referred to as 1-variable.

7.1 Spin-down timescales of $k^{-1} = 6.7$ and 12.0 days

7.1.1 Untuned MTV approximations

In the parameter range with jet bimodality, the untuned 1-variable MTV models in the stationary mode alone do not yield good approximations of the stationary mode statistics of the unreduced model. At $k^{-1} = 6.7$ days, the jet bimodality of the unreduced model is not captured, either with or without nonzero H. Each approximation is characterised by a strong climate drift, as evidenced by the shift in its PDF towards high latitudes (i.e.,

towards large postive PC values) (Figure 7.1). H primarily affects the mean, with a greater shift in the PDF occurring with $H = 0$. The variance is reasonably captured by these approximations, but their PDFs are quasi-Gaussian and their ACFs decay too rapidly. By contrast, the unreduced model stationary mode statistics are characterised by a strong skewness and a slowly-decaying ACF, which reflect excursions of the jet to the low-latitude regime.

At $k^{-1} = 12.0$ days, the stationary mode of the unreduced model has a Gaussian distribution with significantly less variance than at lower spin-down timescales, reflecting that the jet remains at high latitudes (on millennial timescales). The 1-variable MTV approximations have the correct variance and a Gaussian distribution, but have the same deficiencies as at $k^{-1} = 6.7$ days. In particular, the MTV approximations are characterised by a climate drift towards higher latitudes, with a greater climate drift occurring under the assumption that $H = 0$, and their ACFs decay too rapidly.

In the 1-variable case, there is little scale separation between the climate and weather modes, particularly between the stationary mode and the wave-4 weather mode, which may account for the poor quality of the approximations. In the 3-variable case, there is good scale separation of the stationary and wave-4 modes from all other modes. However, the 3-variable MTV models yield reasonably good approximations only of the stationary mode statistics. At $k^{-1} = 6.7$ days, the MTV approximation with nonzero H captures the first two moments of the stationary mode of the unreduced model, but not its skewness (Figure 7.2). With $H = 0$, the distribution of the MTV model stationary mode is skewed in the direction of low latitudes, but lacks the strong skewness which reflects multiple regime behaviour; as well, this mode has a climate drift towards higher latitudes. Both models (with and without nonzero H) give reasonable approximations of the stationary mode ACF despite their inability to capture the pronounced jet bimodality of the unreduced model. It is maybe not surprising that the jet bimodality is not entirely captured considering that wave-4 is so poorly approximated. In particular, the strong low-frequency oscillation of wave-4 of the unreduced model is not captured (with or without nonzero H), as evidenced by the significantly decreased variance and lag-correlation timescale (Figure 7.2).

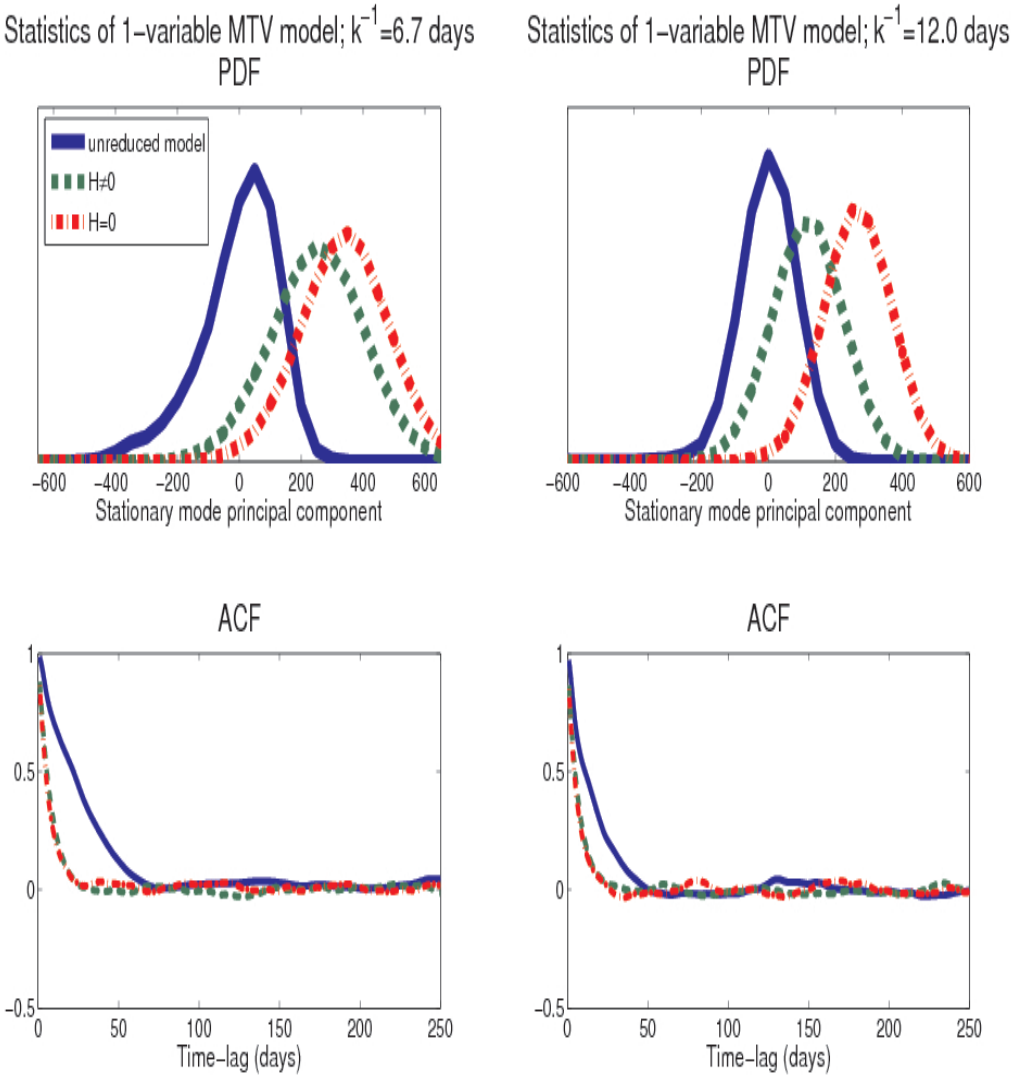

Figure 7.1: The statistics of the untuned 1-variable MTV models at spin-down timescales in the region of jet bimodality. Each simulation is characterised by a significant climate drift.

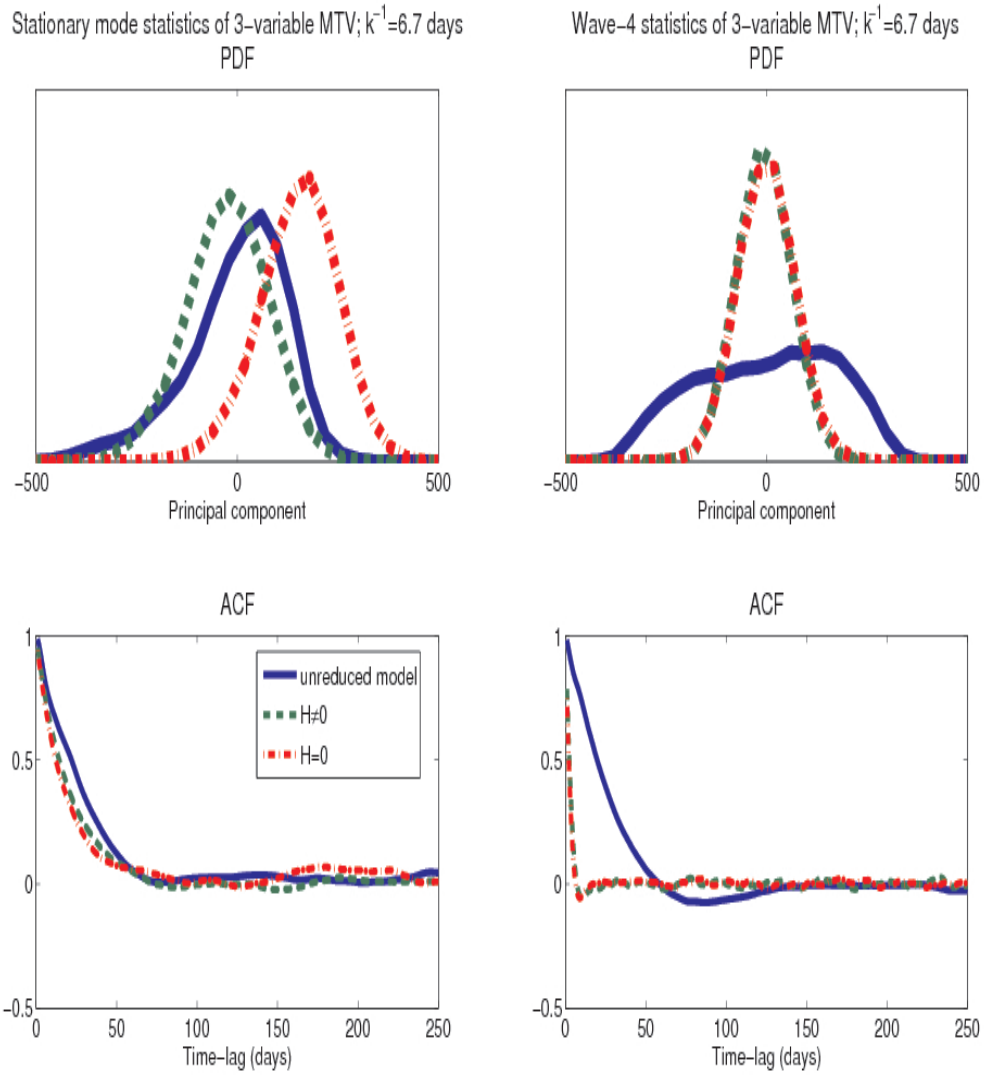

Figure 7.2: The statistics of the untuned 3-variable MTV model at $k^{-1} = 6.7$ days. The stationary mode statistics are reasonably well simulated, but the pronounced jet bimodality of the unreduced model is not captured. The dominant low-frequency oscillations of wave-4 are not captured.

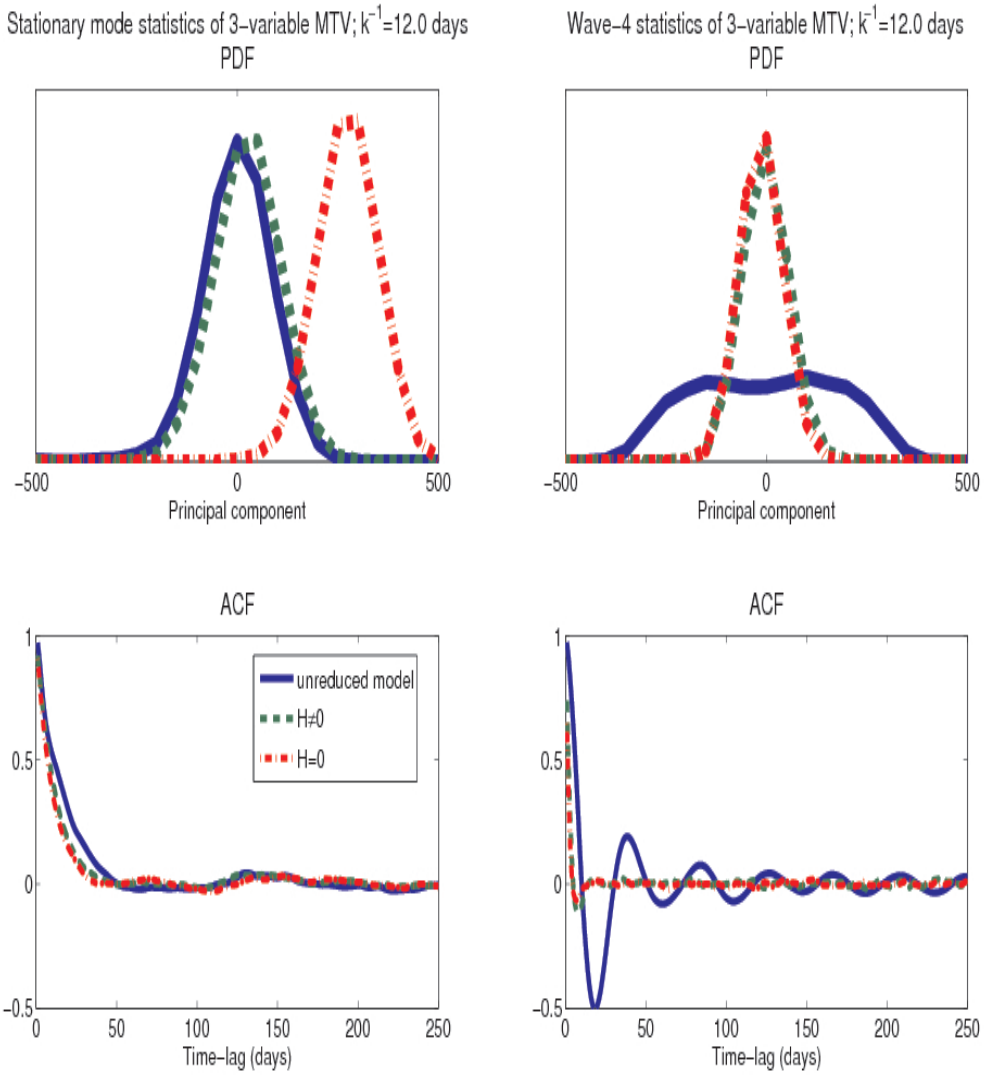

Figure 7.3: The statistics of the untuned 3-variable MTV model at $k^{-1} = 12.0$ days. The stationary mode statistics are very well simulated, but the dominant low-frequency oscillations of wave-4 are not captured.

As at $k^{-1} = 6.7$ days, the 3-variable MTV model with nonzero H at $k^{-1} = 12.0$ days captures the first two moments of the unreduced model stationary mode (Figure 7.3). Since the stationary mode of the unreduced model has a Gaussian distribution at $k^{-1} = 12.0$ days, the MTV model with nonzero H performs very well in simulating the stationary mode statistics of the unreduced model. With $H = 0$, there is a significant climate drift towards higher latitudes. Wave-4 of the unreduced model is dominated by a low-frequency oscillation, as reflected in its ACF. However, this is not captured by the MTV approximation of wave-4, which is instead similar to the wave-4 approximation at $k^{-1} = 6.7$ days.

7.1.2 MTV approximations with MTV fitting procedure

In FM06, a 'minimal' fitting procedure was developed (as detailed in Chapter 5), which introduced scaling factors into the effective climate equation (5.25). There are five such scaling factors, each of which is a priori assigned to multiple unreduced model forcings based on their interactions in generating the effective climate forcings. In fact, scaling parameters are assigned based on only a subset of the interactions responsible for the effective climate equations. Such a fitting procedure, even if (somewhat) physically based, represents a departure from the goal of this dissertation of applying rigorously-based limiting methods in modeling the climate component of the climate-weather system.

However, the number of scaling factors can be effectively reduced, and hence the degrees of freedom of the fitting procedure can be reduced, by equating scaling factors. In particular, the number of scaling parameters can be reduced to two by setting $\lambda_A = \lambda_L = \lambda_M = \lambda_F$, which controls the scaling of the effective climate feedback (i.e., all the noise and noise-induced drift terms). By equating these scaling factors, the fitting procedure consists of adjusting the timescale of the bare truncation forcing relative to that of the effective climate feedback. As discussed in Chapter 5, this particular fitting procedure can correct the scaling problems that arise from the absence of an explicitly resolved scale separation parameter in the KRG05 model equations (5.11). Generally, the MTV fitting procedure corrects for the departures of the KRG05 model from the assumptions underlying the MTV method, including the absence of a large scale separation between the slow and fast modes.

The approximations to follow are derived under the assumption that H acts on the $\mathcal{O}(1)$ timescale. The optimal results are not significantly improved under the assumption that H vanishes from the effective equation. Unless otherwise noted, scaling parameters are set to the default value of 1.

In the region of jet bimodality, the untuned 1-variable MTV models reasonably approximate the variance, but not the mean and ACF of the stationary mode from the full model. With a downscaled effective climate feedback, these latter two statistics are better simulated. At $k^{-1} = 12.0$ days, the 1-variable model with the setting $\lambda_A = \lambda_L = \lambda_M = \lambda_F = 0.75$ (i.e., with the effective climate feedback downscaled by a factor of 0.75^2) yields an excellent approximation of the unreduced model stationary mode statistics (Figure 7.4).

At $k^{-1} = 6.7$ days, the first two moments are well simulated with a downscaled effective climate feedback, but jet bimodality is not captured. With the settings $\lambda_A = \lambda_L = \lambda_M = \lambda_F = 0.66$, there is a slight skewness in the direction of low latitudes, but the presence of jet bimodality is not clear as it is in the unreduced model (Figure 7.4). Without the pronounced multiple regime behaviour, the MTV model ACF decays too rapidly.

In FM06, the bare truncation forcing was downscaled, with the effect of eliminating the climate drift. Here as well the mean is adjusted by changing the relative timescales of the bare truncation forcing and the effective climate feedback. However, in application to the KRG05 model, the climate drift is eliminated with a downscaled effective climate feedback. The choice was made to downscale the climate feedback instead of to upscale the bare truncation forcing to better match the lag-correlation timescales of the reduced and unreduced models. Since the variance is already well generated by the untuned MTV approximation, the first two moments and the timescale are reasonably well simulated by adjusting only the bare truncation forcing relative to the effective climate feedback. However, the simulation of higher-order moments does not improve as a result of these improved statistics.

The bare truncation MTV models (not shown) are found by setting $\lambda_B = 1$ and $\lambda_A = \lambda_L = \lambda_M = \lambda_F = 0$. At each of $k^{-1} = 6.7$ days and $k^{-1} = 12.0$ days, the solution of the 3-variable bare truncation model is a limit cycle, the oscillations of which have a frequency

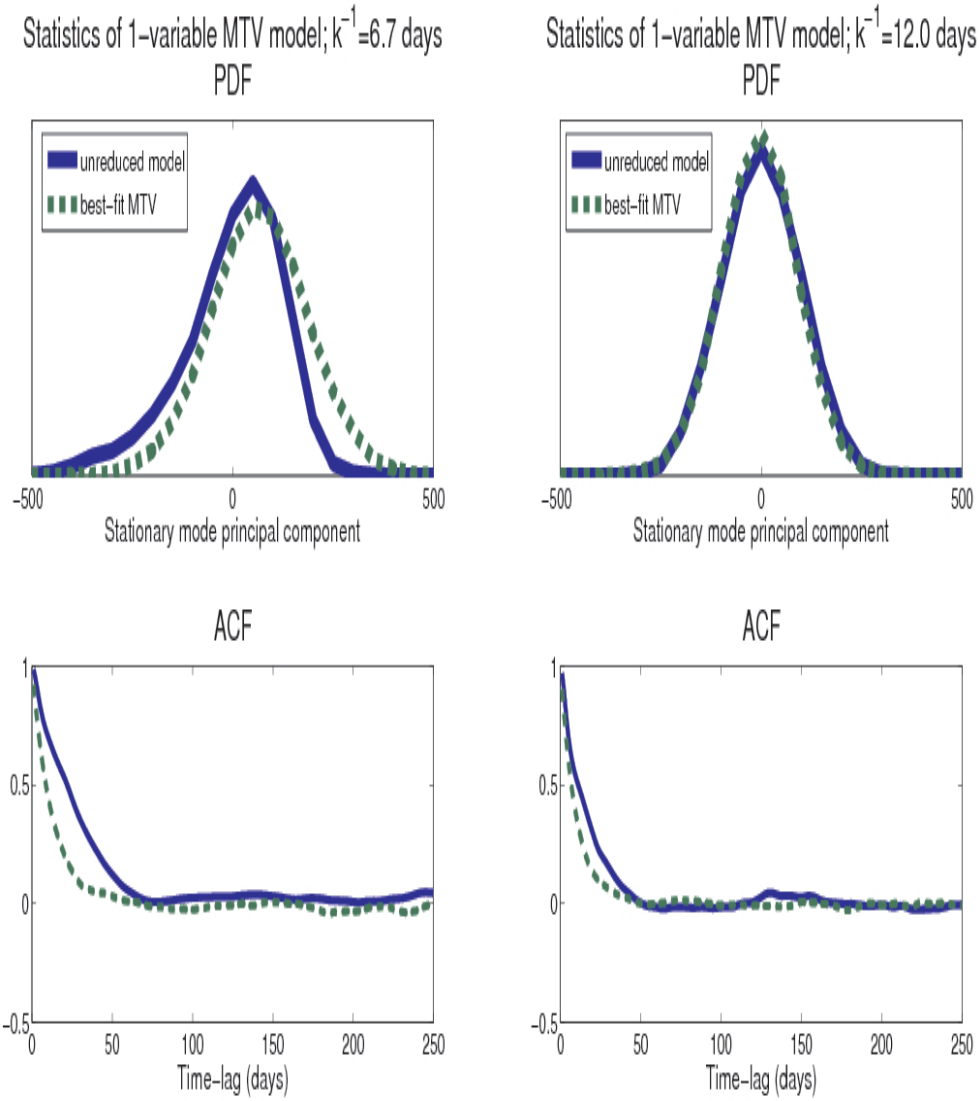

Figure 7.4: The statistics of the tuned 1-variable MTV models in the region of jet bimodality. Jet bimodality cannot be generated by adjusting the MTV tuning parameters.

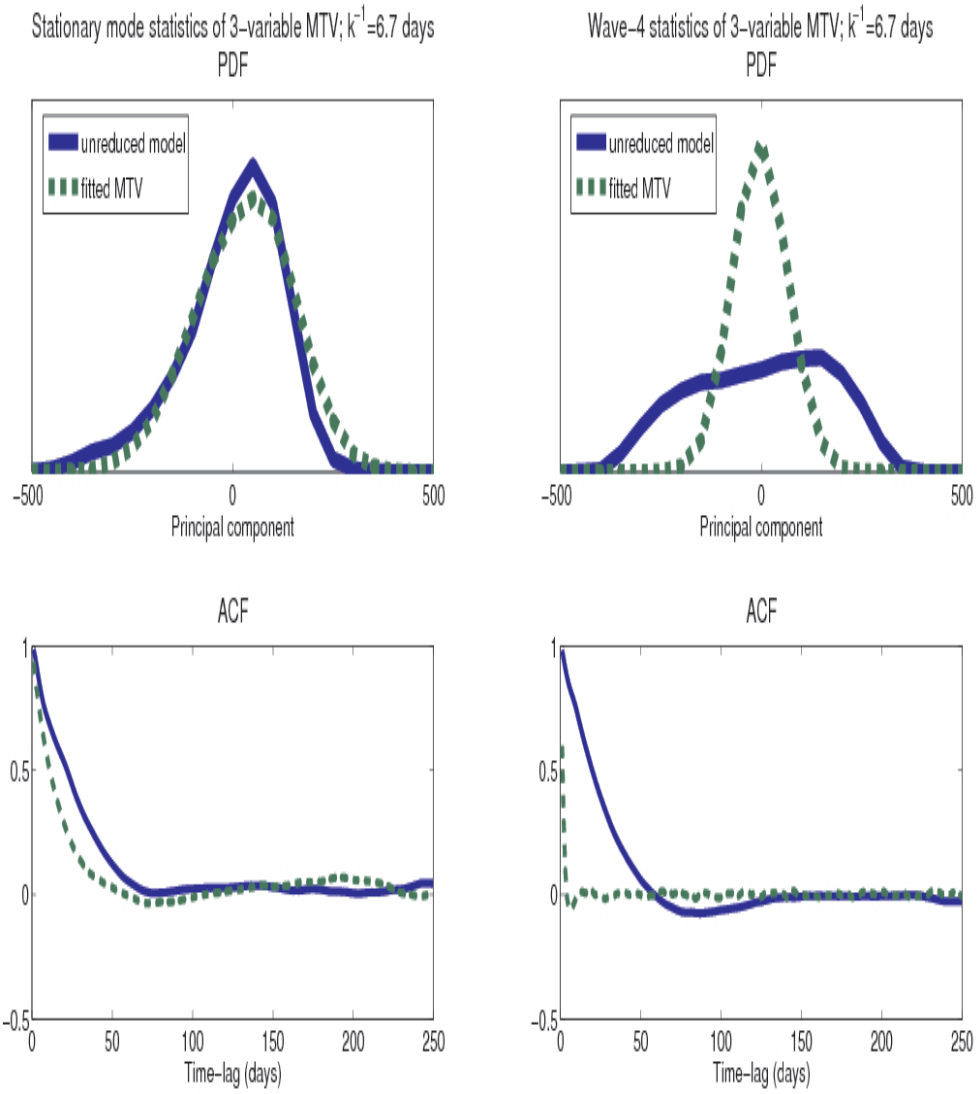

Figure 7.5: The statistics of the 3-variable MTV model with upscaled climate feedback at $k^{-1} = 6.7$ days. The stationary mode statistics are better simulated than in the untuned model, despite the lack of improvement in the wave-4 simulation.

of approximately 1/20 days^{-1}. At $k^{-1} = 6.7$ days, this frequency is an order of magnitude greater than the fundamental frequency of wave-4 of the unreduced KRG05 model (Figure 4.4), but at $k^{-1} = 12.0$ days there is approximate agreement between the two frequencies. As wave-4 is dominated by an oscillation at the latter spin-down timescale, wave-4 of the bare truncation model and of the unreduced model do not significantly differ. At each spin-down timescale, the bare truncation model has a large climate drift, particularly manifest in the stationary mode.

In the 3-variable case, different parameter settings optimise the simulation of the stationary mode and the simulation of wave-4. At $k^{-1} = 6.7$ days, the 3-variable untuned MTV model yields a reasonable approximation of the stationary mode. With the effective climate feedback upscaled with the setting $\lambda_A = \lambda_L = \lambda_M = \lambda_F = 1.66$, the stationary mode PDF has a moderate skewness in the direction of low latitudes (Figure 7.5), but the distinctive multiple regime behaviour of the unreduced model is not generated. With an upscaled effective climate feedback, there are no significant changes in the wave-4 statistics. Rather, oscillations are introduced into wave-4 with a downscaled effective climate feedback. With the setting $\lambda_A = \lambda_L = \lambda_M = \lambda_F = 0.33$, the wave-4 PDF is well simulated, but the oscillations in fact dominate, as evidenced by the large oscillation in the ACF (Figure 7.6). Interestingly, the stationary mode does appear to have jet bimodality, but with a dominant low-latitude regime.

At $k^{-1} = 12.0$ days, the bare truncation model yields a reasonable approximation of wave-4. It is thus not surprising that the 3-variable MTV model with a downscaled effective climate feedback, with the same setting as before of $\lambda_A = \lambda_L = \lambda_M = \lambda_F = 0.33$, yields a good approximation of the wave-4 statistics (Figure 7.7). However, the stationary mode has a strong climate drift towards low-latitudes with this setting; simulation of this mode is optimised in the untuned 3-variable MTV model. As is the case at $k^{-1} = 6.7$ days, other scaling combinations for which the statistics of *both* modes are captured were not found.

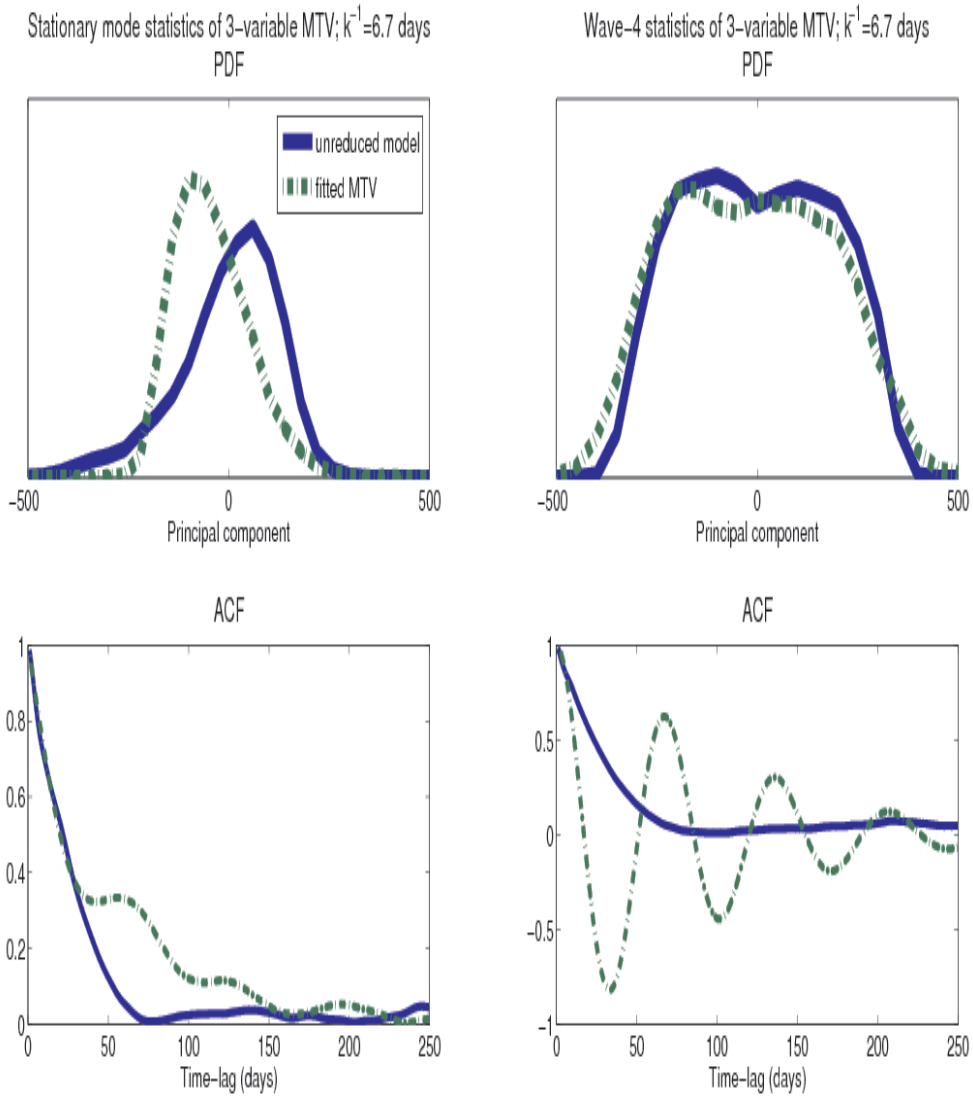

Figure 7.6: The statistics of the 3-variable MTV model with downscaled effective climate feedback at $k^{-1} = 6.7$ days. With this tuning, the PDF of wave-4 is best simulated. Wave-4 of the unreduced model has a concentration of power at low frequencies, but not to the extent as in this tuned model.

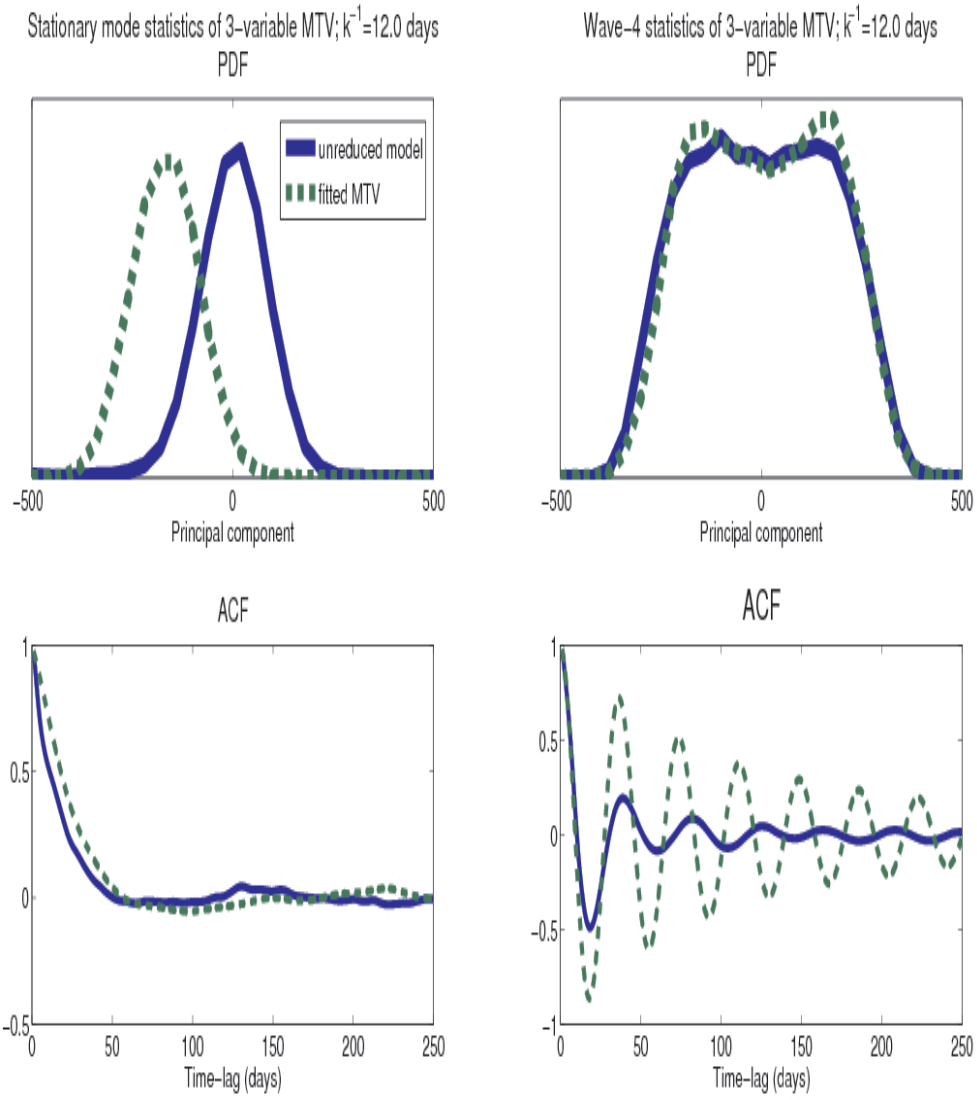

Figure 7.7: The statistics of the 3-variable MTV model with downscaled effective climate feedback. The wave-4 statistics are reasonably well simulated with this tuning, but a climate drift is evident in the stationary mode.

7.1.3 Jet bimodality induced by additional fitting

As discussed, the scaling parameters of the MTV fitting scheme are assigned to unreduced model tensors based on the interactions which give rise to the effective climate tensors. The unreduced model tensors which give rise to the effective climate tensors may be divided into three groups: the bare truncation tensors, the climate-weather interactions which project onto the climate modes and the climate-weather interactions which project onto the weather modes. A new parameter, λ_C, is assigned to each climate-weather interaction which projects onto the climate modes, yielding the effective climate equation:

$$
\begin{aligned}
da_i(t) &= \lambda_B \left[H_i^a dt + \sum_j L_{ij}^{aa} a_j(t) dt + \sum_{jk} B_{ijk}^{aaa} a_j(t) a_k(t) dt \right] && (7.2)\\
&+\lambda_A^2 \lambda_{\mathbf{C}} \sum_j \tilde{L}_{ij}^{(2)} a_j(t) dt + \lambda_A \lambda_{\mathbf{C}} \sqrt{2} \sum_j \sigma_{ij}^{(2)} dW_j^{(2)} && (7.3)\\
&+\lambda_M^2 \left[\lambda_{\mathbf{C}}^{\mathbf{2}} \sum_j \tilde{L}_{ij}^{(3)} a_j(t) dt + \lambda_{\mathbf{C}} \sum_{jkl} \tilde{M}_{ijkl} a_j(t) a_k(t) a_l(t) dt \right] && (7.4)\\
&+\lambda_L^2 \left[\lambda_{\mathbf{C}} \sum_j \tilde{L}_{ij}^{(1)} a_j(t) dt \right] && (7.5)\\
&+\lambda_M \lambda_L \left[\lambda_{\mathbf{C}}^{\mathbf{2}} \tilde{H}_i^{(1)} dt + \lambda_{\mathbf{C}} \sum_{jk} \tilde{B}_{ijk} a_j(t) a_k(t) dt \right] && (7.6)\\
&+\lambda_A \lambda_F \lambda_{\mathbf{C}} \tilde{H}_i^{(2)} dt && (7.7)\\
&+\sqrt{2} \sum_j \sigma_{ij}^{(1)}(a(t)) dW_j^{(1)}, && (7.8)
\end{aligned}
$$

where the multiplicative noise matrix $\sigma^{(1)}$ satisfies

$$
\lambda_{\mathbf{C}}^{\mathbf{2}} \left[\lambda_L^2 Q_{ij}^{(1)} + \lambda_L \lambda_M \sum_k U_{ijk} \alpha_k(t) + \lambda_M^2 \sum_{kl} V_{ijkl} \alpha_k(t) \alpha_l(t) \right] = \sum_k \sigma_{ik}^{(1)} \left[\alpha(t) \right] \sigma_{jk}^{(1)} \left[\alpha(t) \right].
$$

With the addition of λ_C, the scaling of the climate-weather interaction projected onto the climate modes may be adjusted, such that the relative scalings of the three groups of forcings which give rise to the effective climate equation may be adjusted.

With appropriate tuning, jet bimodality is induced in the MTV models at $k^{-1} = 6.7$

days. The other key parameter is $\theta \equiv \lambda_M = \lambda_L$, which controls the scaling of the nonlinear component of the effective climate feedback as well as the scaling of the multiplicative noise. In the 1-variable case, a reasonable approximation of the unreduced model jet bimodality is obtained with $H = 0$ and with the settings $\lambda_C = 0.166$ and $\theta = 9.5$ (Figure 7.8). The time independent bare truncation forcing need not be eliminated to obtain jet bimodality, but the results are optimised with $H = 0$. The results are sensitive to θ, with significant changes occurring in persistence and recurrence within each of the regimes with a change in this parameter. In particular, the greater is θ, the more the probability density is shifted to the low-latitude regime, such that the low-latitude regime becomes favoured with sufficiently increased θ. In the limit of large θ, excursions to the high-latitude regime cease altogether.

Plots of the drift potential (i.e., the potential of the bare-truncation and noise-induced drifts) illustrate the importance of the drift as a function of θ to jet bimodality. As depicted in Figure 7.8, the potential at high latitudes remains approximately constant with increasing θ, but at low latitudes the potential decreases. With $\theta = 9.5$, the potential is at a minimum at a high latitude and increases slowly as the low latitudes are approached, which is consistent with the skewness in the stationary mode PDF in the direction of low latitudes. With $\theta = 11.0$, potential is double-welled with a (slightly) deeper well at low latitudes, which is again consistent with the behaviour of the corresponding MTV approximation. The linear noise-induced drift and the bilinear and cubic noise-induced drifts are upscaled the most with these settings, with multiplicative noise upscaled significantly less. It thus appears that the aspect of multiplicative noise most important for jet bimodality in this model is the noise-induced drift.

With very similar parameter adjustments, jet bimodality is induced in the 3-variable MTV model as well. As in the 1-variable case, $\lambda_C = 0.166$ and $H = 0$, but θ must be made larger to generate a good approximation of the stationary mode statistics (Figure 7.9). In fact, the stationary mode is better simulated in the 3-variable case than in the 1-variable case. On the other hand, wave-4 is very poorly approximated, with a lag-correlation timescale of less than a day, which is less than (very short) lag-correlation timescale of wave-4 the untuned model (Figure 7.9).

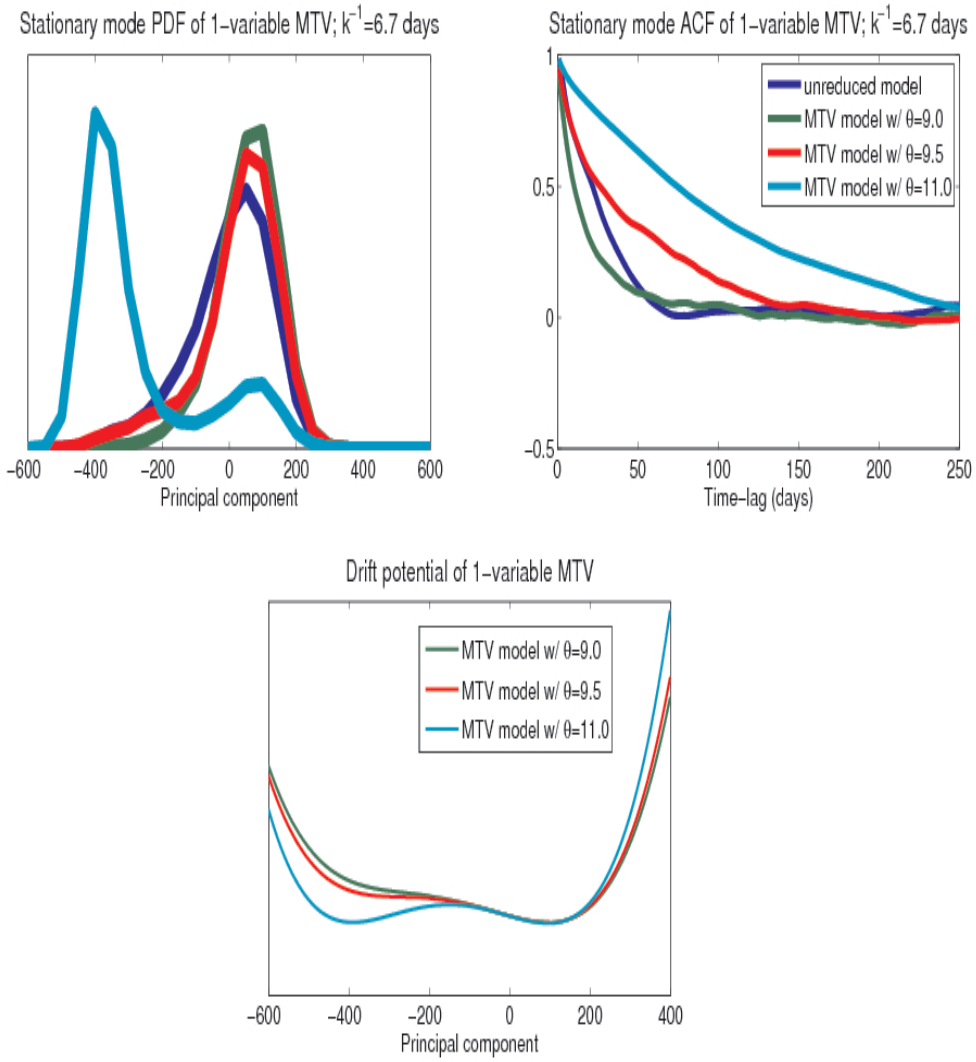

Figure 7.8: The statistics (top) and the potential (bottom) of the tuned 1-variable MTV model, with one MTV scaling parameter and one additional parameter adjusted. Each of these models is derived with $\lambda_C = 0.166$ and with $H = 0$, and the value of θ is varied as indicated. Jet bimodality is induced with these settings, although a particularly good match with the stationary mode statistics of the unreduced model is not obtained.

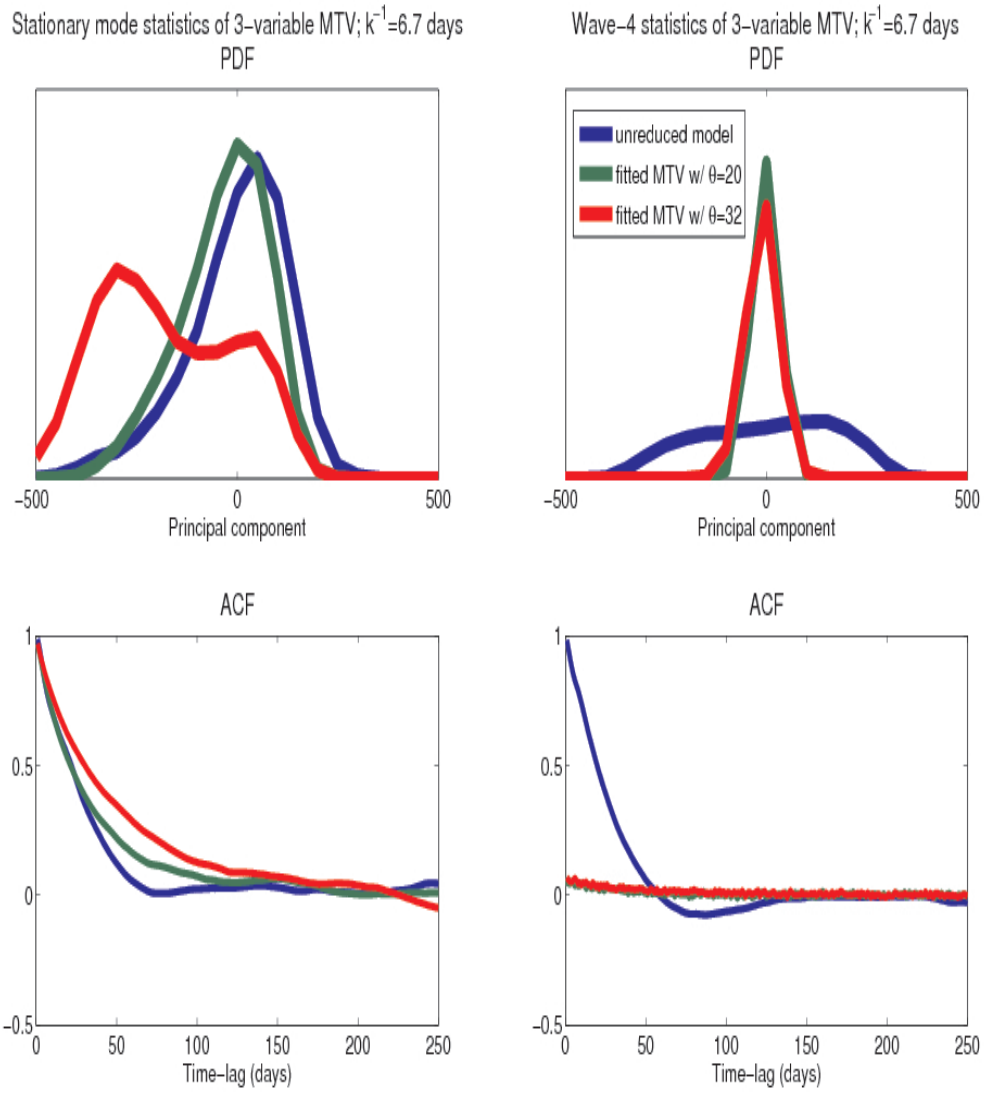

Figure 7.9: As in Figure 7.8, but for the tuned 3-variable model. The stationary mode statistics are better captured with the appropriate tuning, but wave-4 is very poorly is simulated, with a lag-correlation timescale of less than 1 day.

7.2 Spin-down timescales of $k^{-1} = 2.3$ and 4.7 days

7.2.1 Untuned MTV approximations

For values of the spin-down timescale below the bifurcation point (that is, without jet bimodality), each of the untuned 1-variable MTV models is able to capture the stationary mode PDF of the unreduced model, but does not perform as well in simulating the temporal characteristics. The time-independent bare truncation forcing H does not impact the model approximations. At $k^{-1} = 4.7$ days, the MTV model generates a Gaussian distribution in good agreement with the stationary mode distribution of the unreduced model (Figure 7.10). However, the simulated lag-correlation timescale is too short and the ACF decays exponentially, whereas it decays with oscillation in the unreduced model.

At $k^{-1} = 2.3$ days, the 1-variable MTV model yields an excellent simulation of the stationary mode PDF of the unreduced model, but as at $k^{-1} = 4.7$ days, the MTV model fails to simulate the decay with oscillation in the stationary mode ACF (Figure 7.10). However, the lag-correlation timescale is better simulated than at $k^{-1} = 4.7$ days.

In the 3-variable case, the MTV models generate reasonably good approximations of the stationary mode statistics only. As in the 1-variable case, the 3-variable MTV models with and without $H = 0$ do not significantly differ. At $k^{-1} = 4.7$ days, the stationary mode PDF and its lag-correlation timescale are well simulated, but the damped oscillation in the ACF is not captured (Figure 7.11). As the timescale is not well simulated in the 1-variable case, the 3-variable approximation of the stationary mode statistics is an improvement over the 1-variable approximation. Despite the nearly identical statistics of the two wave-4 components of the unreduced model, the MTV approximation of these two components differ. Their PDFs are similar, but the lag-correlation timescale of wave-4 component (II) is larger than that of component (I), although less than that of component (II) of the unreduced model.

The difference between the wave-4 components of the 3-variable MTV approximation at $k^{-1} = 2.3$ days is more pronounced. As at $k^{-1} = 4.7$ days, the statistics of the wave-4 components of the unreduced model are nearly identical. The MTV approximation of wave-4 component (I) is similar to the wave-4 approximations on the side with jet bimodality. By

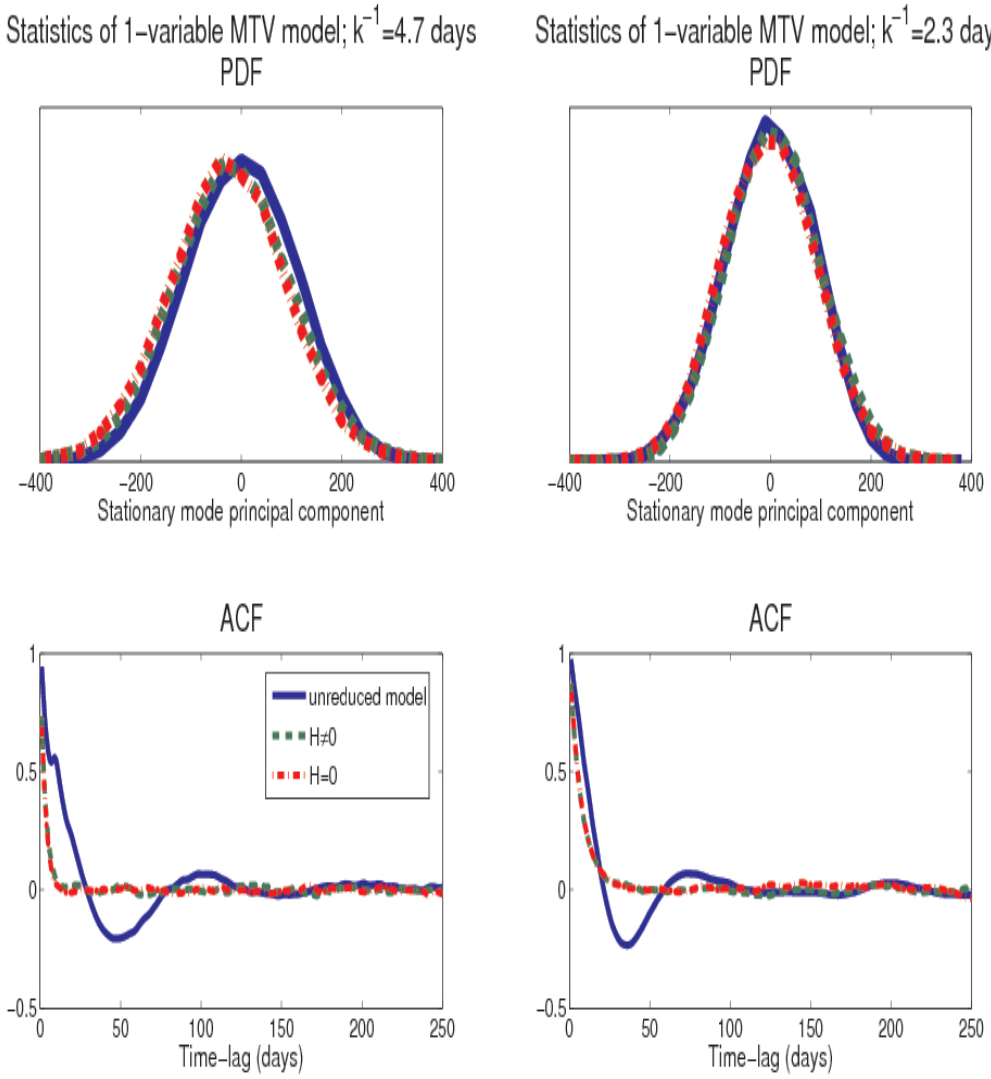

Figure 7.10: The statistics of the untuned 1-variable MTV models at spin-down timescales in the region with a single jet state.

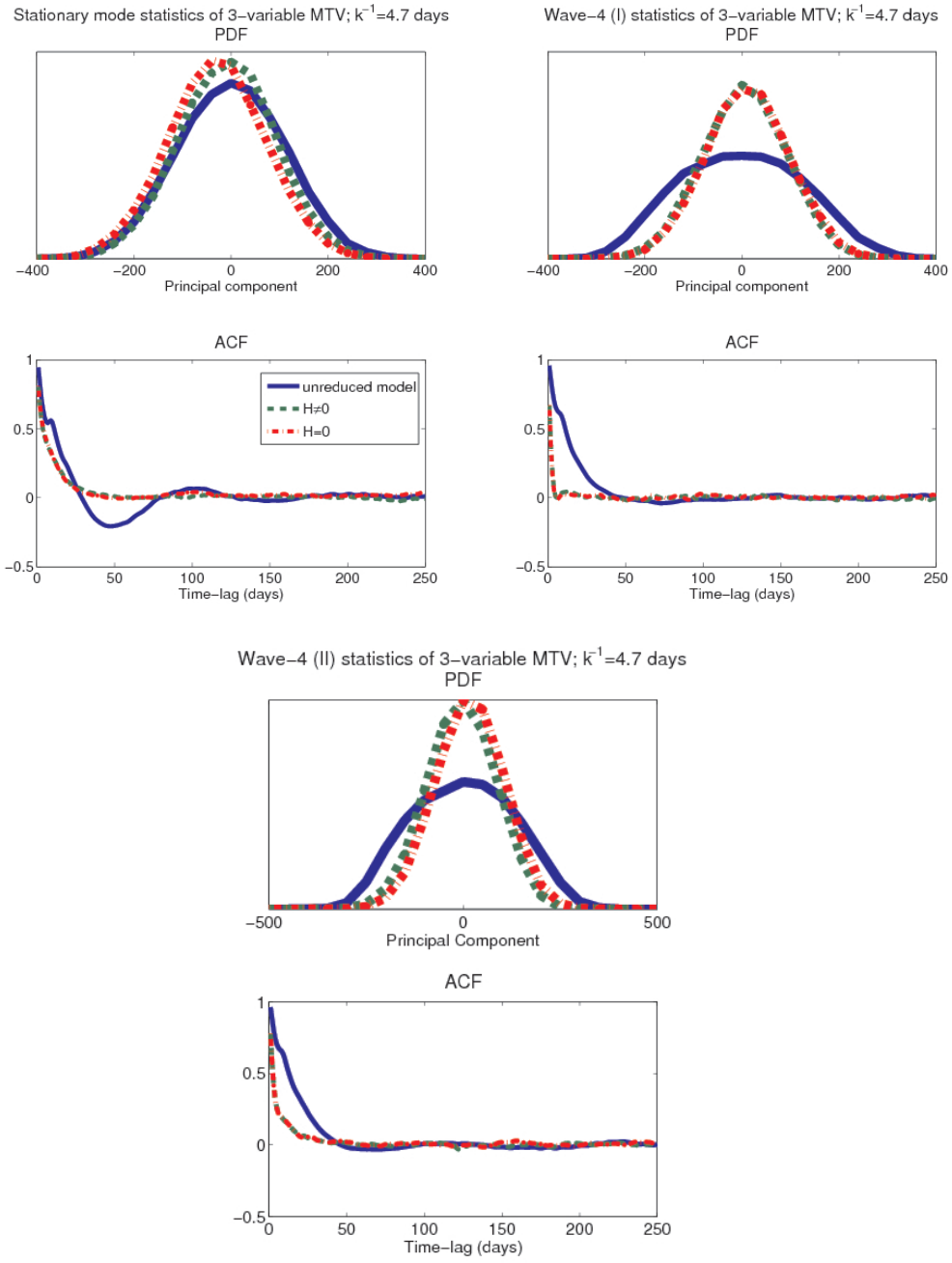

Figure 7.11: The statistics of the untuned 3-variable MTV model at $k^{-1} = 4.7$ days. Both wave-4 components of the MTV model are depicted as there are differences in their statistics (but the statistics of the wave-4 components of the unreduced model are nearly identical).

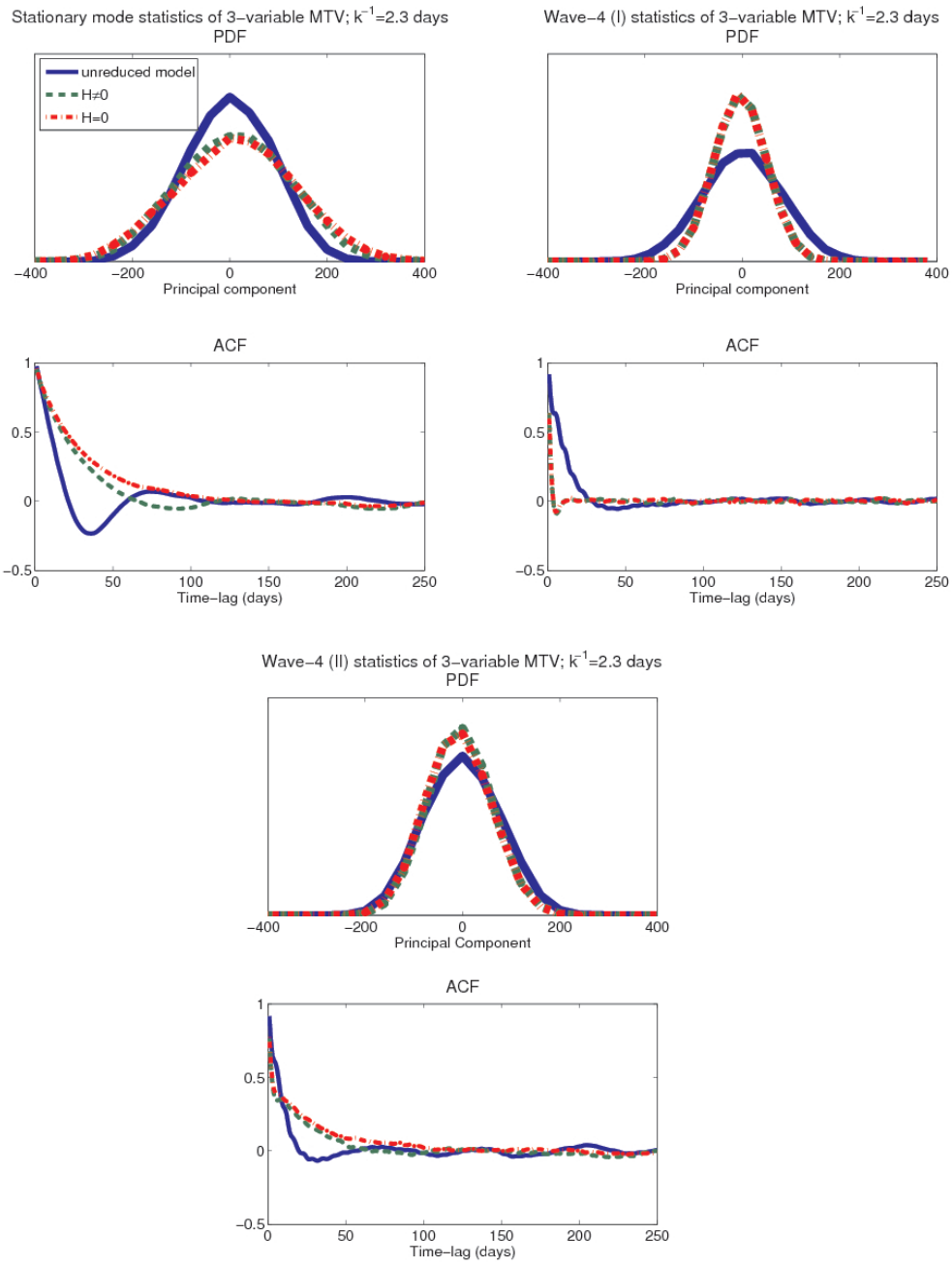

Figure 7.12: As in Figure 7.11 but at $k^{-1} = 2.3$ days. The lag-correlation timescale of one of the MTV wave-4 components is significantly longer than that of the other.

contrast, the ACF of the MTV approximation of wave-4 component (II) decays very slowly on time-lags of weeks (Figure 7.12). The MTV model yields a reasonable approximation of the stationary mode statistics, although the low-frequency oscillation in the stationary mode is not captured.

On the side of the bifurcation point without jet bimodality, the stationary mode EOFs (not shown) have a greater wave-4 modulation than on the side with jet bimodality. It is thus likely that a significant portion of the *dynamical* stationary mode is captured by at least one of the *statistically-determined* wave-4 mode components. Indeed, this is made evident by the appropriate tuning of the MTV model.

7.2.2 MTV approximations with MTV fitting procedure

At each of $k^{-1} = 2.3$ and 4.7 days, the untuned 1-variable model does not capture the oscillation with lag in the ACF. However, this feature cannot be derived by tuning the model, as an oscillation in the reduced model requires at least two degrees of freedom. At $k^{-1} = 2.3$ days, the statistics are otherwise captured by the untuned model. At $k^{-1} = 4.7$ days, the ACF of the 1-variable untuned model also decays too rapidly, which can be rectified simply by fine-graining the effective equation; i.e., by setting $\lambda_B = \lambda_A = \lambda_L = \lambda_M = \lambda_F < 1$ (not shown).

The bare truncation MTV models (not shown) are found by setting $\lambda_B = 1$ and $\lambda_A = \lambda_L = \lambda_M = \lambda_F = 0$. As on the side of the bifurcation point with jet bimodality, the solution of the 3-variable bare truncation model at $k^{-1} = 4.7$ days is a limit cycle. However, the frequency of the oscillations of approximately 1/40 days^{-1} is half of that on the side with jet bimodality. At $k^{-1} = 2.3$ days, the 3-variable bare truncation model has a fixed point attractor.

At $k^{-1} = 2.3$ days, the statistics of the 3-variable MTV model cannot be significantly improved by adjusting the MTV tuning parameters. However, it can be made clear that the statistical wave-4 component (II) captures a significant portion of the dynamical stationary mode. By scaling down the effective climate feedback by setting $\lambda_A = \lambda_L = \lambda_M = \lambda_F = 0.66$, the stationary mode variance decreases significantly and the ACFs of the two wave-4 components are made similar (Figure 7.13). By instead increasing the scaling of the

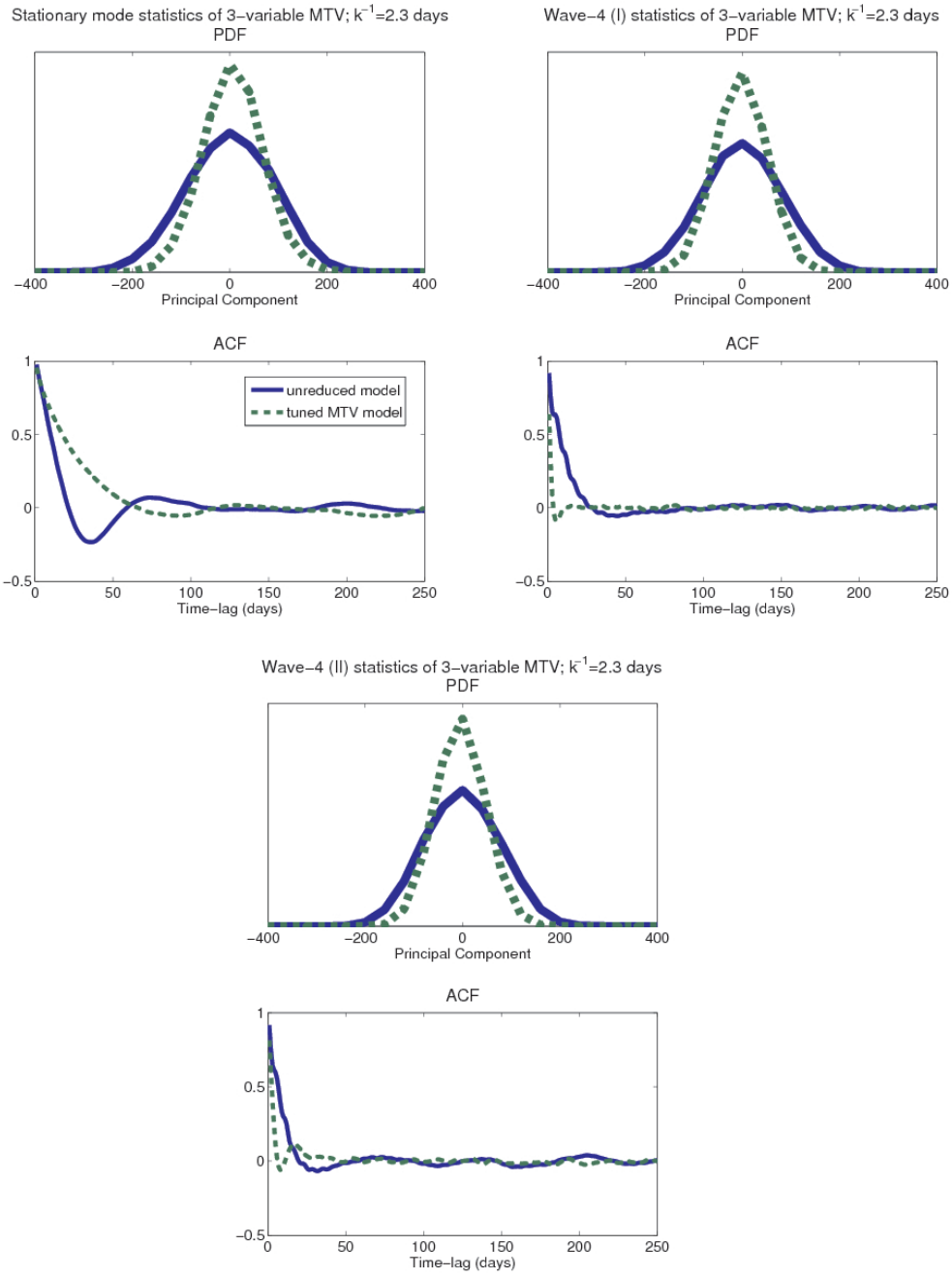

Figure 7.13: 3-variable MTV model at $k^{-1} = 2.3$ days with downscaled effective climate feedback. The statistics of the wave-4 components do not significantly differ.

effective climate feedback, jet bimodality is induced in the reduced system, as is manifest in the bimodal distributions of the stationary mode *and* wave-4 component (II). Wave-4 component (I) does not change significantly with this tuning, despite the significant changes which occur in the other two variables.

Whereas at $k^{-1} = 2.3$ days jet bimodality is induced with a downscaled bare truncation forcing, at $k^{-1} = 4.7$ days, the stationary mode shifts towards higher latitudes with this tuning. This is corrected by also scaling down the linear terms scaled by λ_A and λ_F with a setting $\lambda_B = \lambda_A = \lambda_F < 1$ (equivalently, with a setting $\theta > 1$). In fact, the simulation is not particularly sensitive to changes in this scaling for $\lambda_B = \lambda_A = \lambda_F < 0.5$. As depicted in Figure 7.14, the tuned 3-variable model with the bare truncation forcing, additive noise and some linear noise-induced drift terms eliminated, derived by setting $\lambda_B = \lambda_A = \lambda_F = 0$, generates a reasonable approximation of the lag-correlation timescale and PDF of the stationary mode and of wave-4 component (II). However, the statistics of wave-4 component (I) do not improve significantly with this tuning. Since a portion of the dynamical stationary mode likely projects onto the statistical wave-4 component (II), the improvement in component (II) is just a reflection of the improvement in the simulation of the stationary mode.

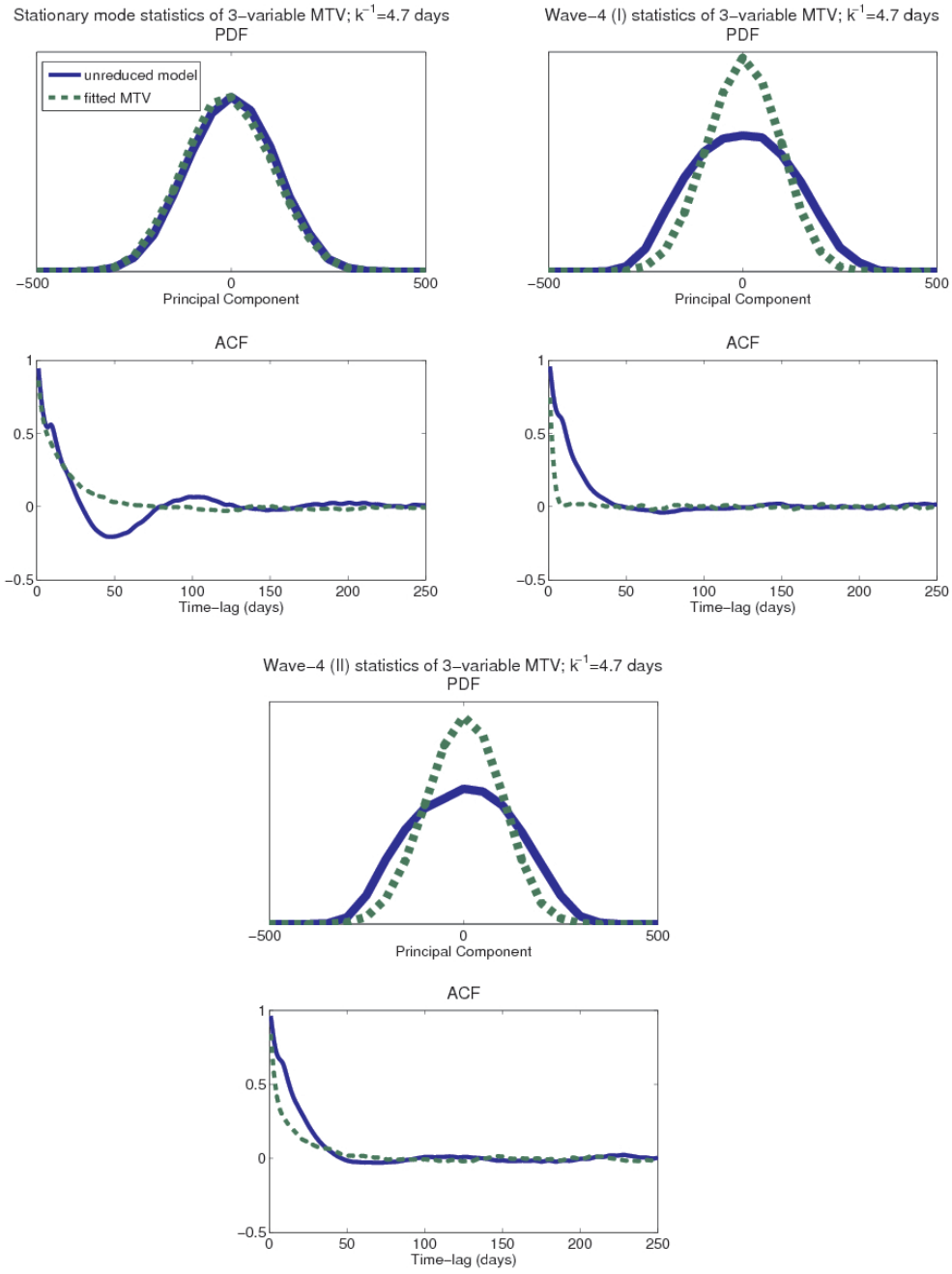

Figure 7.14: 3-variable MTV model at $k^{-1} = 4.7$ days with the bare truncation forcing, linear noise-induced drifts and additive noise removed.

7.3 MTV models with full cross-correlations

The above results were derived by assuming that the lagged cross-correlations of the weather modes are zero. This assumption greatly reduces the amount of computation in deriving the MTV coefficients, but is not generally valid. In FMV05 and FM06, however, no difference was found between the models derived with and without cross-correlations, for a number of reduced models. Applied to the KRG05 model, if 9 of the 12 noise and noise-induced drift terms are computed with cross-correlations and the other 3 without cross-correlations, then the results are the same as those obtained with all 12 terms computed without cross-correlations. However, by computing all 12 terms with cross-correlations, the results significantly differ from the results without cross-correlations. In particular, the reduced models derived with full cross-correlations yield very poor approximations of the KRG05 climate.

These three forcing tensors - two linear noise-induced drift tensors and the additive noise tensor ($H^{(2)}$, $L^{(2)}$ and $Q^{(2)}$ in (5.25)) - are computed with the largest number of summands: using the first $1,000$ EOFs and an upper lag of 200 days with a sampling resolution of $1/2$ of a day (and a sample size of $365,000$ days), there are $10^{14} - 10^{15}$ summands in their computation, which is three orders of magnitude greater than the number of summands in the next largest computation. (In FMV05 and FM06, the integrated lag-correlations were computed from a simulation of size $1,000,000$ days sampled every $1/2$ day.) The results are not sensitive to a change in the number of EOFs, to the sampling size nor to the sampling resolution. It is clear in application to the KRG05 model that error builds up in computation of these three forcings to the point of overwhelming the reduced dynamics.

7.4 Summary and discussion

The MTV method was applied to the KRG05 model at spin-down timescales on both sides of the bifurcation point to obtain low-dimensional climate models. Since at each of these spin-down timescales the stationary and wave-4 modes are the dominant slowly-evolving modes, 3-variable reduced models in these two modes were derived, as well as 1-variable reduced models in the stationary mode alone. Without tuning, the 1-variable model performed reasonably well on the side without jet bimodality, but in the region of jet bimodality, only the variance was captured. The climate drift was eliminated and the lag-correlation

timescale was improved by scaling down the effective climate feedback, but jet bimodality at $k^{-1} = 6.7$ days could not be induced by adjusting the MTV tuning parameters.

At each spin-down timescale, the untuned 3-variable model gave a reasonably good approximation of the stationary mode statistics and a poor approximation of the wave-4 statistics. The 3-variable bare truncation models yielded either limit cycles or fixed point solutions, neither of which resemble the full MTV approximations. Thus, the a priori determined effective climate feedback (i.e., noise and noise-induced drift) generated (most of) the necessary corrections in simulating the stationary mode statistics. At each spin-down timescale, the simulated wave-4 did not have the strong low-frequency oscillation of wave-4 of the unreduced model, which was manifest in its significantly reduced variance and shorter lag-correlation timescale. Only at the very high spin-down timescale of $k^{-1} = 12.0$ days with a downscaled effective climate feedback was a good approximation of both the wave-4 PDF and ACF obtained, but this scaling caused a climate drift in the stationary mode.

It was concluded in KRG05 that jet bimodality is the result of a long timescale interaction between the stationary and wave-4 modes, which in the 3-variable case would satisfy the assumption of the MTV theory of strong scale separation between the bare truncation forcing and the climate-weather interactions. However, a different conclusion was reached in Chapter 6 through that fact that jet bimodality persisted in the unreduced model with the nonlinear interaction between these modes removed; rather, the nonlinear interaction among the stationary, wave-4 and wave-5 modes was implicated as the cause of jet bimodality. This underlying dynamics may be behind the failure of the untuned 3-variable MTV model to capture jet bimodality, but it still remains the case that there is good scale separation between these modes and the weather modes. Jet bimodality can be induced by tuning, but the dynamical picture becomes muddled. In particular, jet bimodality was induced in *both* the 1-variable and 3-variable models at $k^{-1} = 6.7$ days by scaling up the nonlinear noise-induced drift. In the 3-variable case, the oscillation in wave-4 remained absent with this rescaling. These results suggest that jet bimodality occurs independently of wave-4. However, given the ad hoc nature of these tunings, the induced jet bimodality

may just be an artifact of turning up the nonlinearity in the system. Indeed, jet bimodality was induced by scaling up the nonlinearity in the reduced model at a spin-down timescale (of $k^{-1} = 2.3$ days) far removed from the region of jet bimodality.

The inability of the untuned 3-variable model at $k^{-1} = 6.7$ days to generate jet bimodality may be a symptom of the general inability (at all spin-down timescales) of the untuned 3-variable model to simulate the low-frequency oscillation of the propagating wave-4 mode. With downscaled effective climate feedback, oscillations were induced in wave-4 at $k^{-1} = 4.7$, 6.7 and 12.0 days and the PDFs well captured (and the ACF as well at the latter spin-down timescale). At $k^{-1} = 12.0$ days in particular, it is not surprising that the wave-4 statistics were generated with a significantly downscaled effective climate feedback as wave-4 of the unreduced model appears to be to first order decoupled from the weather modes. Generally, it may be the case that the wave-4 oscillation is too strongly damped by noise in the untuned model. It is also possible that the problem arises from the deterministic forcings of the reduced model. It was determined in KRG05 that the stationary and wave-4 modes are close to eigenmodes of the full system linearised about the climatology (cf. Chapter 4). In the untuned reduced system, this quality of the wave-4 mode in particular may be lost. A tuning with the constraint that the eigenvalues of the linearised operator of the full model are preserved might better capture the statistics of both modes.

Based on the (very) limited number of applications of the MTV method to realistic geophysical models, the propagating modes appear to be most problematic in general. In Strounine et al. (submitted), in which the MTV method was applied to the Marshall and Molteni (1993) model, it was also found that the statistics of the stationary modes are reasonably well approximated, but that there are shortcomings in the simulation of the propagating modes. In fact, the MTV method was applied with Kalman filter estimation of the tuning parameters. With optimal tuning, the opposite problem arose of spurious low-frequency oscillations in the wave modes.

In Strounine et al. (submitted), the number of tuning parameters was increased from five to ten independent parameters to facilitate application of the Kalman filter estimation of these parameters. Recall that the MTV scaling parameters are assigned to unreduced

model tensors based on the interactions which give rise to the effective climate tensors. This basis for assigning parameters was largely disregarded in Strounine et al. (submitted), and instead the bare truncation forcing terms and individual effective climate feedback terms were linearly scaled by newly defined parameters. However, they did not achieve significantly more success in simulating the climate statistics than in FM06, in which the MTV method with the (trial-and-error) MTV scaling scheme was applied to the same model. In particular, the moderate non-Gaussianity which characterises LFV was not captured by the reduced model in either of these studies. In this dissertation, typically only one parameter, either the scaling of the bare truncation forcing or effective climate feedback, was adjusted. As in FM06, the climate drift was corrected by such a rescaling, and in the 3-variable model at $k^{-1} = 6.7$ days, the stationary mode PDF developed a stronger skewness in better agreement with that of the unreduced model. Certainly, with ten parameters, it is more likely that the problems in simulating wave-4 would be rectified, but the resulting model is even further away from the original reduced model.

The sequential estimation of the ten parameters via Kalman filtering required a minimal input of data to estimate the MTV coefficients, which were then corrected by this data assimilation method. In particular, the coefficients were determined from a sample size of $30,000$ days sampled once daily, and the unreduced model data were assimilated using a $5,000$ day sample. In addition to providing a systematic means of determining the optimal tuning coefficients and reducing the enormous computational costs in estimating the MTV coefficients directly, Kalman filtering may have prevented the error build-up that overwhelmed the reduced KRG05 model computed with full cross-correlations. However, the tuning is not 'minimal', which obscures the connection of the reduced model to the rigorously justified asymptotic theory. Nevertheless, it does provide a systematic method of deriving a (fitted) stochastic equation with multiplicative noise.

Since individual forcings are resolved in the MTV model, a detailed budget analysis is possible. In FMV05 and FM06, the importance of five groups of forcings, scaled by the five tuning parameters, were considered in a budget analysis. In FM06, λ_M and λ_L, which scale the nonlinear terms and multiplicative noise of the effective climate feedback, were

determined to be important. However, since the MTV approximations did not capture the non-Gaussianity of the climate modes, this analysis does not answer the important questions concerning the dynamics of LFV. Moreover, the connection between the effective and unreduced model forcings through the MTV tuning parameters is ambiguous. In Strounine et al. (submitted), there was no untuned model to analyse, but the models with and without multiplicative noise were compared. Whereas the potential importance of multiplicative noise was argued for in FM06, it was shown in Strounine et al. (submitted) that this noise was not important. A budget analysis of the MTV models of the KRG05 model was not considered because the primary feature of interest - pronounced jet bimodality - was not simulated by the untuned models.

Overall, the MTV method as applied in FM06 did not perform particularly well in reduction of the KRG05 model, especially in the 3-variable case. The failure of the untuned 3-variable models to generate the wave-4 statistics illustrates that timescale separation alone is insufficient to obtain a good reduced stochastic model. Indeed, the theory underlying the MTV method includes the stringent assumption that the scale separation between climate and weather modes also serves as the scale separation between the climate tendency forcings and as the scale separation between the weather tendency forcings. Furthermore, it is assumed that the average of the climate-weather interactions projected onto the climate modes is zero (cf. Chapter 2). These assumptions are relaxed in deriving Hasselmann's equations, which are examined in the next chapter.

Chapter 8

Results of Hasselmann's method applied to KRG05 model

There are multiple Hasselmann models ((A), (L), (W), (N) and (N+)), which form a hierarchy of models based on their different abilities to capture the unreduced model dynamics (cf. Table 2.1). It is theoretically justified that the (L) and (N) models are superior to the (A) model on bounded timescales and that only the (N+) model is valid on the long, $\mathcal{O}(\epsilon^{-1})$ timescale. The (L) approximation in particular is given by the *solution* of the (A) model plus a correction given by a Gauss-Markov process with mean zero (2.11). Special cases of the (L) model have been used to successfully simulate LFV in, e.g., Farrell and Ioannou (1995), Penland and Sardeshmukh (1995) and Whitaker and Sardeshmukh (1998). However, in application to the KRG05 model, it can be immediately deduced from the relevant (A) approximations that the (L) approximations are unable to capture the (pronounced) jet bimodality. Thus, the (L) approximations are not determined in this study. For computational reasons, the (N+) approximations are also not found. Expressed in Itô form, the (N+) SDE involves derivatives of the weather modes, which are unmanageable because of the high-dimensional dynamics of the KRG05 model.

Nevertheless, the (A), (W) and (N) Hasselmann models are informative and the stochastic models in particular prove capable of simulating the climate statistics of the unreduced model. These Hasselmann models are integrated using the seamless algorithm developed in Chapter 5. This algorithm requires minimal changes to the original KRG05 model algorithm, as only the climate modes need be resolved. The (W) and (N) model parameters

were determined in Chapter 6; particularly, the values of the timescale separation parameter ϵ (cf. Table 6.1) and the scaling values associated with the diffusion matrix σ (i.e., γ and the extrapolation factor; cf. Table 6.2). The scaling of σ, denoted by γ, is needed to properly scale the noise term in the (W) model. In the (N) model, an extrapolation factor is required to approximate the integrated lag-covariance for large lag by the upscaled (by the extrapolation factor) integrated lag-covariance for small lag. The scale separation ϵ explicitly appears in the (N) model noise term given by $\sqrt{\epsilon}\sigma dW/dt$ (and is defined to appear in the (W) model noise term $\sqrt{\epsilon}\gamma I dW/dt$). In theory, in comparing the statistics of the unreduced and (N) models, ϵ in the (N) model should not be treated as a free parameter as it represents the fixed scale separation between the climate and weather modes of the unreduced model. However, since ϵ is not explicitly defined in the unreduced model equations, there is uncertainty in the scaling of these equations in ϵ and thus uncertainty in the scaling of the (N) model noise term (cf. Chapter 6). To account for this uncertainty, the noise scaling is tuned by adjusting the value of ϵ. Furthermore, in the absence of arbitrarily large scale separation, the Hasselmann method cannot be expected to be perfectly accurate; ϵ serves as a convenient tuning parameter to bring the reduced and unreduced models into better agreement.

The Hasselmann reduced model approximations are found in the region of jet bimodality at $k^{-1} = 6.7$ and 12.0 days (Section 8.1) and in the region with a single jet state at $k^{-1} = 2.3$ and 4.7 days (Section 8.2). Within each of these sections, the (A), (W) and (N) Hasselmann approximations are found with the a priori determined scaling parameters, and also with ϵ adjusted towards finding best fit approximations. The ϵ-tuned Hasselmann approximations are important as well for comparison with the similarly tuned MTV approximations. In Section 8.3, computationally efficient hybrid Hasselmann-unreduced models are derived and integrated. The results are summarised and discussed in Section 8.4.

8.1 Spin-down timescales of $k^{-1} = 6.7$ and 12.0 days

8.1.1 Deterministic Averaging (A)

Deterministic averaging is the first method considered among Hasselmann's reduction methods as it is simple to implement and it serves as the basis of the stochastic corrections. The

1-variable averaged (A) ODE in the stationary mode alone has a fixed-point attractor, but as shown below, its forcing function $\overline{f}$ captures the dynamics underlying jet bimodality. The 3-variable averaged solutions are non-trivial, but they are only valid (in theory) on the short, $\mathcal{O}(1)$ timescale, and on the longer timescale of jet bimodality, noise becomes important. Even on short timescales, the stochastic models are expected to outperform the deterministic averaged models.

Indeed, the 3-variable averaged (A) dynamics differs from the slow dynamics of the unreduced model in the region of jet bimodality. At $k^{-1} = 6.7$ days, the most obvious difference is the absence of noise in the averaged model wave-4 mode, which is dominated by an oscillation with a period on the order of 200 days (Figure 8.1). As well, the stationary mode distribution has significantly decreased variance and is no longer bimodal, with the jet residing only at high latitudes. For this averaged solution, the macrotime step is $\Delta t = 3$ hours, the microtime step is $\delta t = 10$ minutes, the spin-up is $N_1 = 12\delta t$, the number of ensemble members is $R = 18$, and the number of averaging iterates per ensemble member is $N = 1$. The number of weather ensemble members over which the average is taken per macrotime step is large but needed to sufficiently reduce the noise error associated with the cumulative averaging scheme. In particular, spurious excursions to the low-latitude regime not characteristic of the true averaged dynamics are altogether eliminated using a sufficiently large R. By contrast, wave-4 reaches its asymptotic state with significantly less averaging.

A similar 3-variable averaged dynamics exists at $k^{-1} = 12.0$ days, characterised by an oscillating wave-4 mode (Figure 8.1). Given its long timescale, it is possible that the wave-4 oscillation is an artifact of averaging (which is valid in theory on short timescales only). However, at each of $k^{-1} = 6.7$ and 12.0 days, the oscillation of the wave-4 mode of (A) is similar to that of the unreduced model. In particular, the frequency of the oscillation of the averaged wave-4 mode agrees with the "fundamental frequency" of the wave-4 mode of the unreduced model, as evident from Figure 8.1. This agreement is particularly obvious at $k^{-1} = 12.0$ days because wave-4 of the unreduced model is dominated by an oscillation (before averaging). Deterministic averaging thus reveals that a common wave-4 dynamics

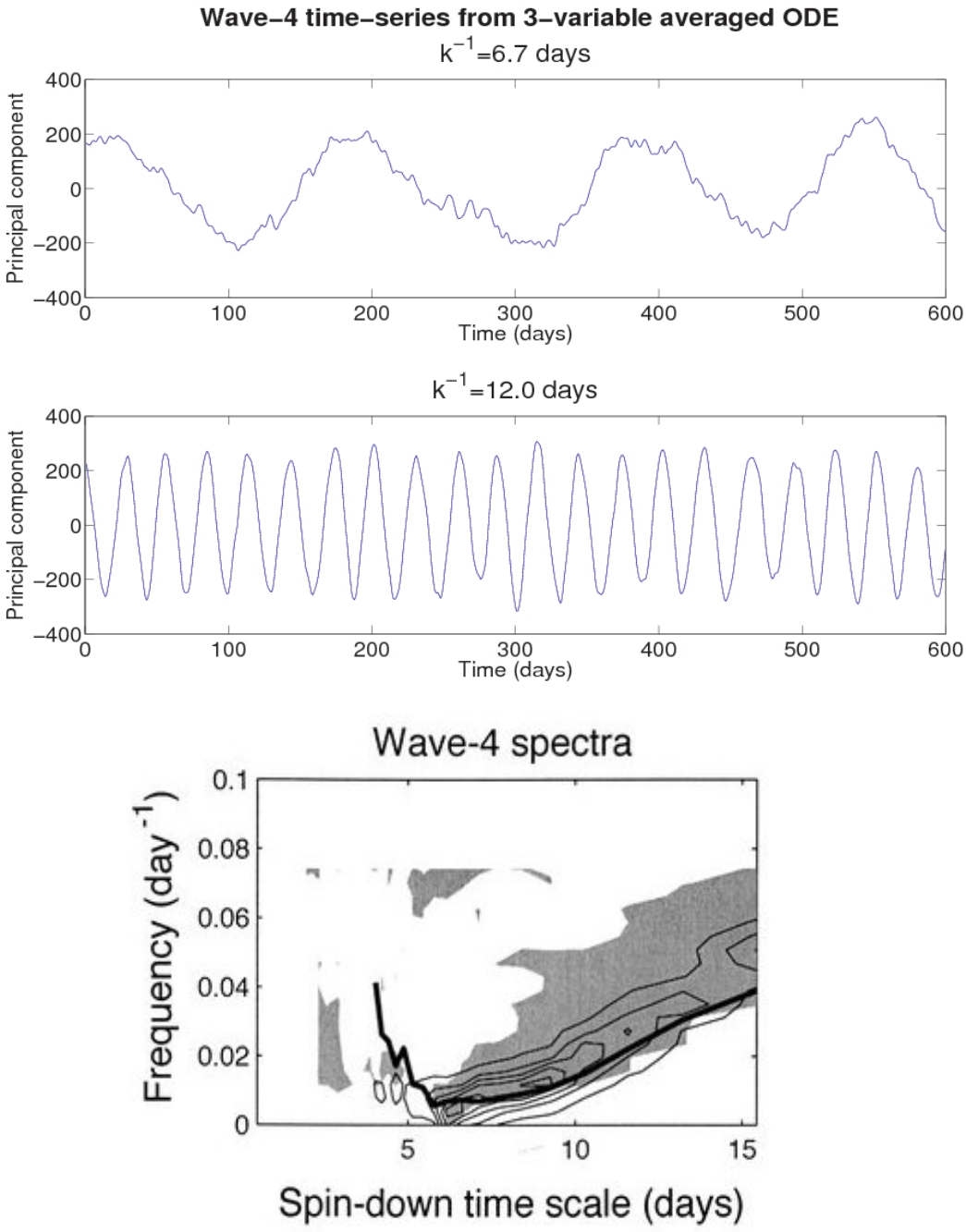

Figure 8.1: Top: typical time series of the wave-4 component of the 3-variable averaged models in the region of jet bimodality. Wave-4 of the averaged model is a simple oscillation in the region of jet bimodality. At each spin-down timescale, the frequency of this oscillation is in approximate agreement with the fundamental frequency of wave-4 of the **unreduced model** (as seen by comparison with bottom plot from KRG05).

exists in the whole region of jet bimodality, which is captured by the averaged dynamics. Averaging filters out the secondary noisy effects, leaving only the dominant oscillation.

Although the averaged (A) model captures the dominant low-frequency oscillation of wave-4, it fails to capture the (very) low-frequency feature of jet bimodality that is most manifest in the stationary mode. Rather, on long timescales there is decoupling of the stationary and wave-4 modes of the (A) model. Indeed, the significantly decreased variance of the stationary mode and the simple oscillation of the wave-4 mode are broadly consistent with the numerical solutions of a decoupled 1-variable ODE (which has a fixed point attractor) and 2-variable ODE (the solution of which may be an oscillation).

With the addition of noise and noise-induced drift to the averaged forcing, the resultant equation is valid on long timescales. It is thus possible that important long timescale features are captured by the averaged forcing function. Indeed, although the 1-variable averaged ODE has a fixed-point attractor, its potential function, given by $F(a) = -\int \overline{f}(a)\, da$, reveals the importance of the averaged function to the dynamics of jet bimodality. At $k^{-1} = 6.7$ days in particular, there is a plateau in the potential approximately corresponding to the location of the low-latitude regime in the unreduced model, as depicted in Figure 8.2. A trough approximately corresponds to the location of the high-latitude regime of the unreduced model at which the jet mostly resides. The potential is not double-welled, as is expected from the fact that the stationary mode PDF of the unreduced model is skewed but not bimodal. Rather, the shape of the potential reflects that the meridional shifting of the jet slows down, or reaches a 'sticking point', in a low-latitude band. At $k^{-1} = 12.0$ days, there is a more pronounced plateau in the potential of the 1-variable averaged drift, but it is situated at much lower latitudes and is separated from the high-latitude regime trough by a larger potential barrier than is the one at $k^{-1} = 6.7$ days. Consistent with this potential, the jet of the unreduced model at $k^{-1} = 12.0$ days remains at high latitudes on millennial timescales.

The Hasselmann stochastic approximations are found below. Based on the potential of the 1-variable averaged (A) equations, it is expected that jet bimodality can be generated in the 1-variable stochastic models without state-dependent noise.

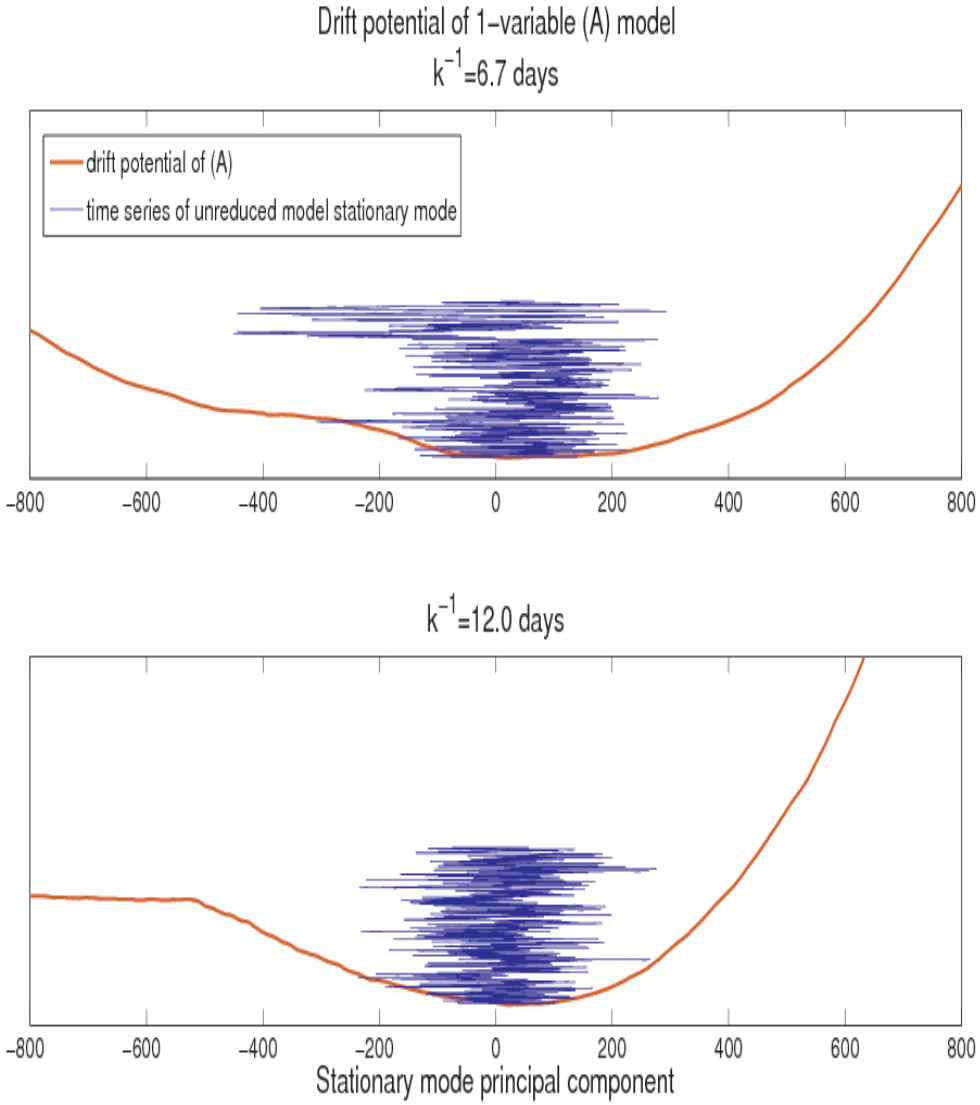

Figure 8.2: Potential of the 1-variable averaged function in the region of jet bimodality. A time-series of the stationary mode of the unreduced model is superposed onto each plot of the potential (such that the y-axis indicates both the potential and the time in days). There are plateaus in the potential at both spin-down timescales, corresponding to the locations of the low-latitude regimes.

8.1.2 Diffusion approximations (W) and (N)

Indeed, at $k^{-1} = 6.7$ days, a 1-variable (W) SDE, consisting of averaged forcing and additive or state-independent Gaussian white noise, suffices to generate the correct jet bimodality statistics. The reduced model PDF is shifted slightly relative to the stationary mode PDF of the unreduced model, but, as evidenced by the pronounced tail of the PDF, the jet bimodality is captured (Figure 8.3). Indeed, the jet bimodality is obvious from a time series of the (W) approximation (not shown), which is characterised by infrequent and persistent visits of the jet to the low-latitude regime. The statistics of the 1-variable multiplicative noise (N) SDE differ insignificantly from those of the additive noise SDE. Note that a 1-variable (L) stochastic approximation cannot yield the regime behaviour because it is a Gaussian-distributed solution, determined as the sum of the fixed-point averaged solution and a Gaussian-distributed correction.

The noise terms of the (W) and (N) approximations are scaled according to the results of Chapter 6. As it was found that there is no scale separation between the stationary and wave-4 modes, the scale separation parameter is given by $\epsilon = 1$. It was also determined that the diffusion coefficient σ has a magnitude of approximately $\gamma = 20$. Accordingly, Gaussian white noise driven by a standard Brownian motion/Wiener process satisfying $W(\Delta t) \sim N(0, \Delta t)$ is scaled by a factor of $\sqrt{\epsilon}\gamma = 20$ in the 1-variable (W) model that generates jet bimodality. At each macrotime step of the (N) approximation, σ is determined by multiplying by a factor of 6 the integrated lag-correlation determined up to a lag of one hour. With an upper-lag of one-half of a day and a correspondingly smaller extrapolation factor, the results of the (N) approximation do not change significantly.

It is suggested by the nature of the potential of the 1-variable averaged equation at $k^{-1} = 6.7$ days that nonlinearity in the deterministic averaged dynamics has a dominant role in driving jet bimodality, which is further supported by the quality of the (W) approximation in simulating jet bimodality. Figure 8.4 demonstrates that the climate variable/stationary mode is effectively independent of the noise term, which is further supported by the similarity of the (N) and (W) approximations. Thus, jet bimodality is governed by deterministic, averaged forcing of the stationary mode, with all other modes averaged out.

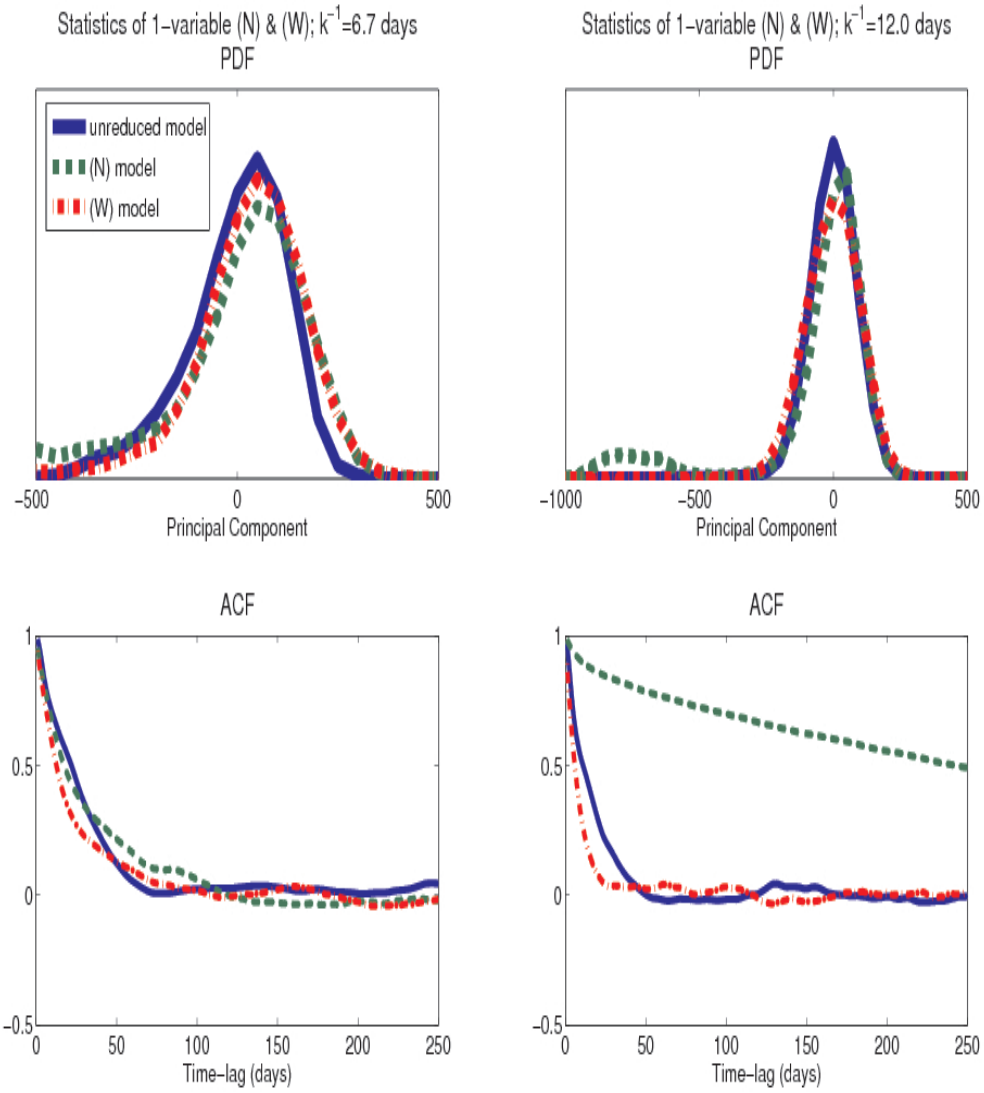

Figure 8.3: Statistics of 1-variable (W) and (N) stochastic Hasselmann models in the region of jet bimodality. At $k^{-1} = 6.7$ days, the (W) and (N) models perform equally well in capturing the jet bimodality of the unreduced model. At $k^{-1} = 12.0$ days, the (N) model generates pronounced multiple regime behaviour, whereas the (W) model generates the single jet state statistics of the unreduced model (on millennial timescales).

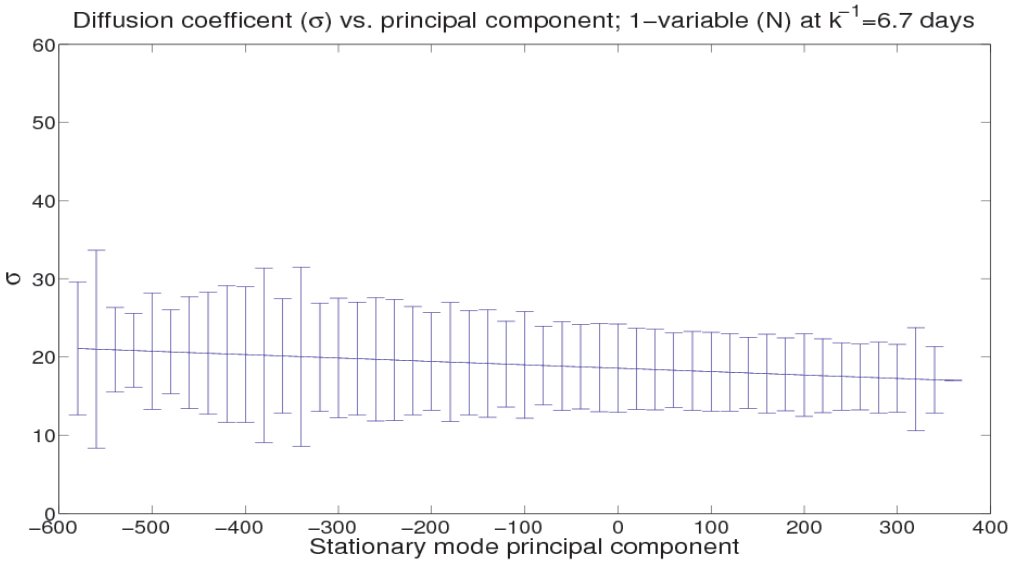

Figure 8.4: Plot of the diffusion coefficient σ against the stationary mode principal component. The error bars indicate the standard deviation of the error in computation of σ, which arises from approximation of σ in a high-dimensional system. It is clear that σ is effectively independent of the climate variable.

The averaged forcing captures the important climate-weather interactions, particularly, the nonlinear interactions among the stationary, wave-4 and wave-5 modes, that give rise to jet bimodality. Additive noise is the source of variance or energy exciting the reduced system to sample the full range of these interactions. Note that the success of the 1-variable averaged equation does not contradict the results of Chapter 6, in which it was determined that the stationary mode self-interaction, defined as the self-advection of the stationary mode, is unimportant. Rather, averaging induces other interactions in the stationary mode besides the stationary mode self-interaction that account for the important effects of the weather modes.

At the very low bottom drag/high spin-down timescale of $k^{-1} = 12.0$ days, the 1-variable (N) and (W) models display significant differences. With the a priori determined parameters of $\epsilon = 1$ and $\gamma = 20$, the 1-variable (W) model generates the stationary mode statistics of the unreduced model (Figure 8.3). In contrast, the (N) approximation displays marked bimodality. The regime positions of the (N) approximation are consistent with the potential of the averaged forcing (Figure 8.2), and for the particular (a priori determined) noise scaling, the low-latitude regime has a persistence timescale of years and a recurrence

timescale of decades. The recurrence timescale is particularly sensitive to the strength of the noise, as a decrease in the (N) model noise strength by only 15% increases to centuries the timescale between visits to the low-latitude regime. In this case, the (N) and (W) approximations both resemble the stationary mode of the unreduced model on timescales of $O(1000)$ years.

On longer timescales, the (N) model with downscaled noise captures the regime behaviour of the unreduced model. Although excursions to the low-latitude regime did not occur in a 1000-year integration of the unreduced model at $k^{-1} = 12.0$ days (from which the statistics of the unreduced model were derived), one such excursion did occur in a long integration at the nearby spin-down timescale of $k^{-1} = 13.0$ days (not shown). The persistence within the low-latitude regime was of $O(10)$ years, in approximate agreement with the persistence timescale within the low-latitude regime of the (N) model with downscaled noise. Excursions to the low-latitude regime would have occurred eventually in the untuned (W) model, possibly with persistence and recurrence timescales in agreement with those of the unreduced model. However, multiplicative noise does act to enhance the persistence timescale within the low-latitude regime. Within each regime, the means of σ differ, with a smaller noise magnitude in the low-latitude regime (Figure 8.5). It thus appears that the jet is less energetic in the low-latitude regime as a result of climate-weather interactions, which allows it to persist on longer timescales within this regime than it would persist in the (W) model. As well spikes in the time series of σ (not shown) coincide with transitions to the low-latitude regime, but it is not clear which variable (between σ and the stationary mode principal component) is in the lead.

There was ambiguity in determining the value of ϵ in the 1-variable case in the region of jet bimodality because of the disparity in the autocorrelation timescales among weather modes; particularly between the wave-4 mode and the rest of the weather modes. With wave-4 defined as a climate mode, there is a clearly defined scale separation between the climate and weather modes. In particular, the noise scaling in the 3-variable stochastic models at $k^{-1} = 6.7$ days is $\sqrt{\epsilon}||\sigma|| = \sqrt{0.1}||\sigma||$ and the scaling at $k^{-1} = 12.0$ days is $\sqrt{\epsilon}||\sigma|| = \sqrt{0.25}||\sigma||$. However, since there is uncertainty in the scaling of the system, the

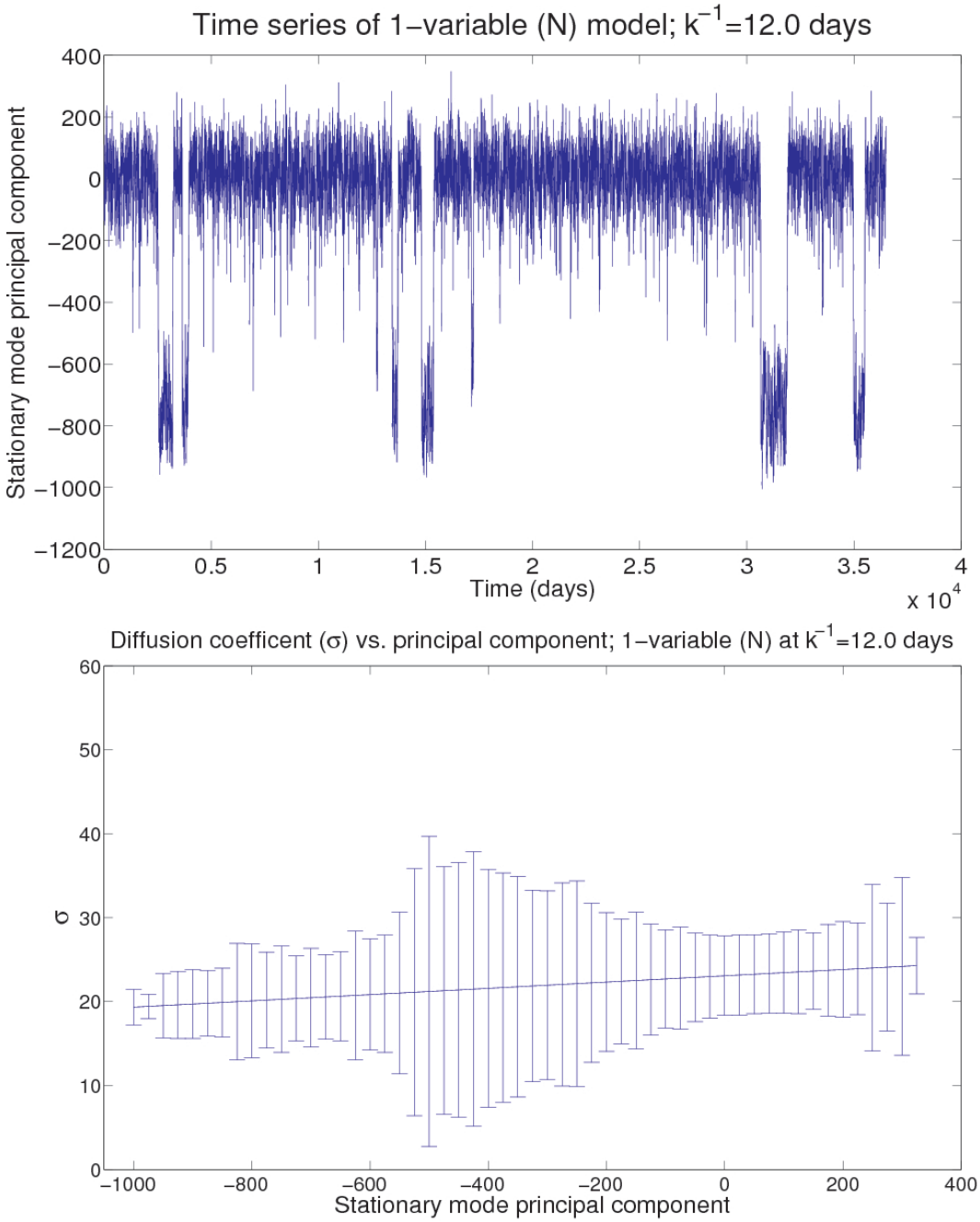

Figure 8.5: Top: time series of 1-variable (N) model at $k^{-1} = 12.0$ days with a priori determined values of the scaling parameters. Bottom: corresponding plot of σ against the stationary mode PC, as in Figure 8.4. There is greater error in computation of σ at points in between the regimes because of the smaller sample size and the spikes in σ which coincide with transitions to the low-latitude regime.

solutions with ϵ varied will be examined.

At $k^{-1} = 6.7$ days, the 3-variable (N) model with $\epsilon = 0.1$ yields a reasonably good approximation of the climate statistics of the unreduced model. Jet bimodality is generated, but the variance and skewness of the stationary mode distribution are decreased relative to that of the unreduced model (Figure 8.6). The decay in the ACF of both modes is captured, but the wave-4 ACF decays with a long timescale oscillation, similar to that of the averaged (A) model. In contrast to the 1-variable case, there are significant differences between the (N) and (W) models, as the similarly scaled 3-variable (W) model (with $\epsilon = 0.1$ and $\gamma = 50$) has a nearly Gaussian distribution with significantly greater variance. In simulating the climate statistics of the unreduced model, the (N) model is superior to the (W) model. In fact, it seems that the unreduced model statistics would be captured by increasing the variance of the (N) model.

However, the stationary mode PDF of the (N) model shifts towards higher latitudes with an increase in the noise strength. As depicted in (Figure 8.7), the variance and skewness are reasonably well captured with the noise scaling increased with $\epsilon = 0.5$, but there is a strong climate drift. With a further increase in the noise scaling, the 3-variable solution becomes unstable.

At $k^{-1} = 12.0$ days, the 3-variable Hasselmann stochastic models with $\epsilon = 0.25$ have unbounded solutions. With significantly downscaled noise, the solutions are bounded, but the jet bimodality generated by the 1-variable (N) model is not generated in the 3-variable case. With the noise scaling decreased to $\epsilon = 0.1$, the 3-variable (N) and (W) models capture the climate statistics of the unreduced model (Figure 8.8). Moreover, there is little difference between the (N) and (W) approximations. With the tuning of a single parameter (the timescale separation) it is possible to bring the full and reduced models into excellent agreement.

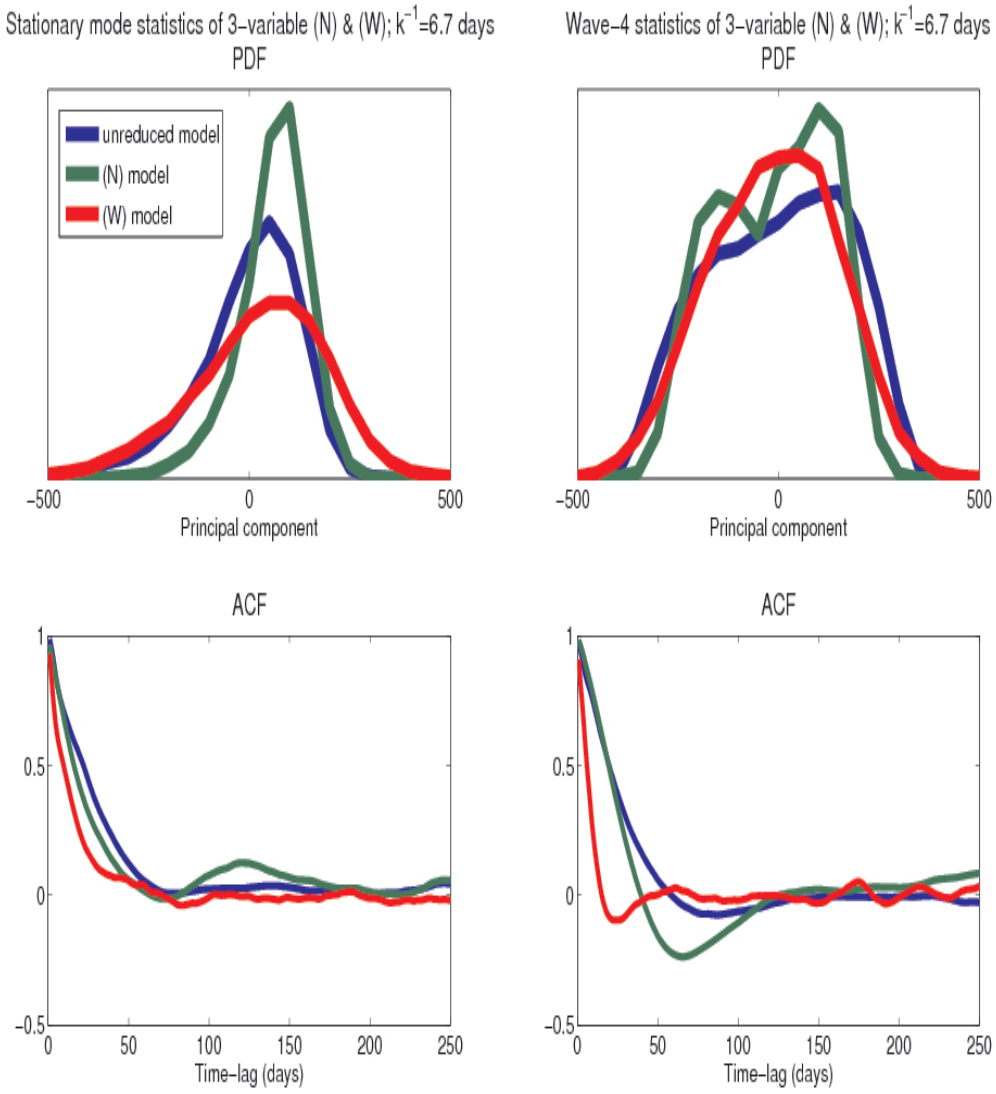

Figure 8.6: Statistics of the 3-variable (N) and (W) models at $k^{-1} = 6.7$ days with a priori determined parameter values.

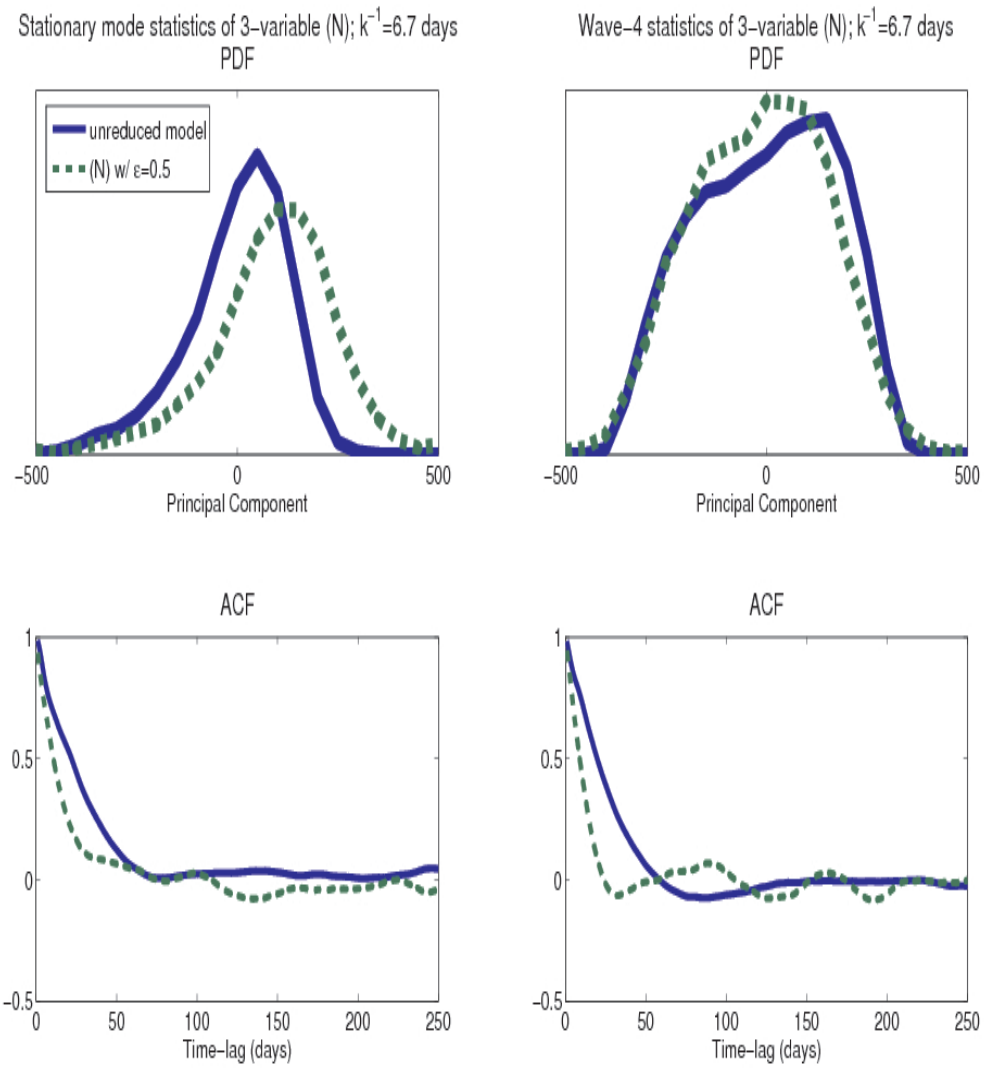

Figure 8.7: Statistics of 3-variable (N) model at $k^{-1} = 6.7$ days with noise upscaled with $\epsilon = 0.5$. The variance and skewness are better approximated with this upscaled noise, but there is a significant climate drift.

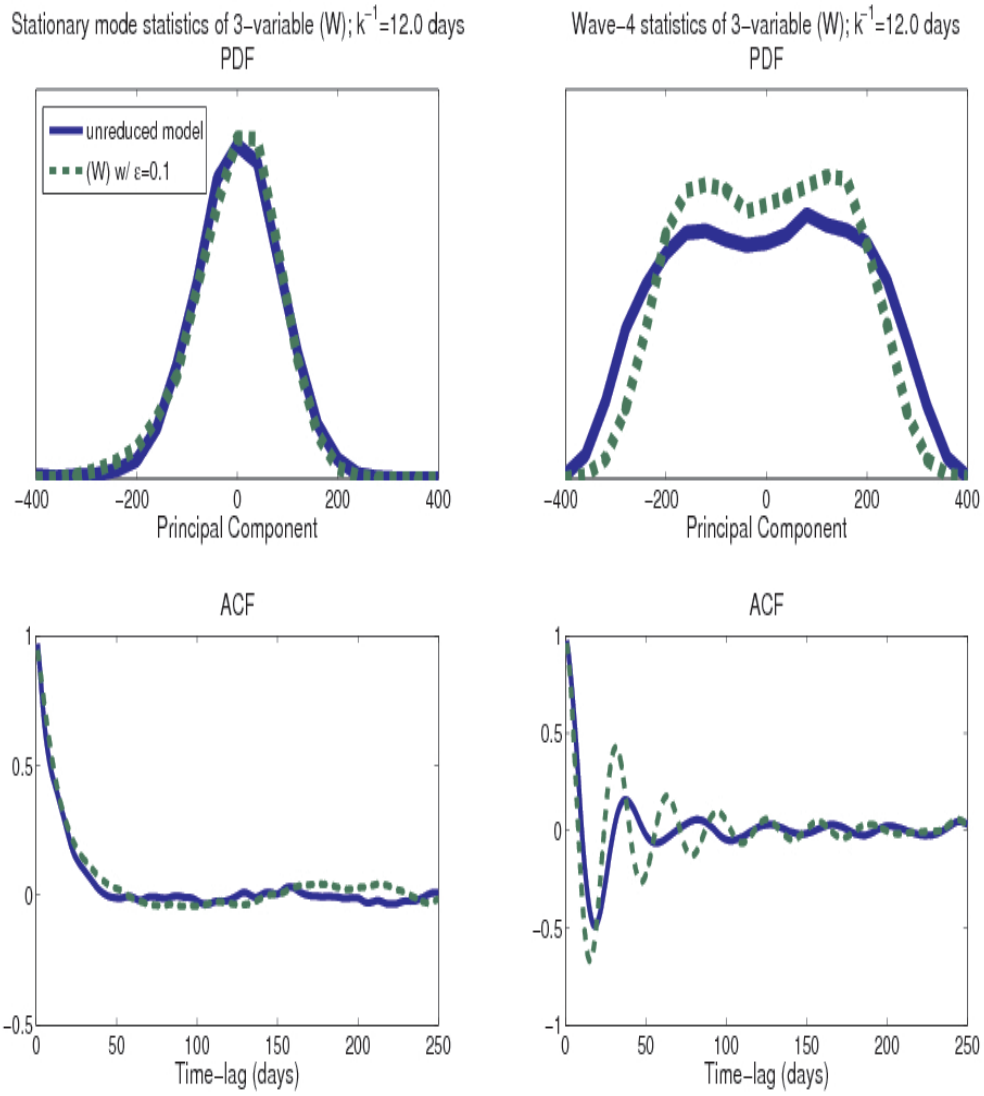

Figure 8.8: Statistics of the 3-variable (W) model at $k^{-1} = 12.0$ days with the noise scaling parameter decreased from the a priori determined value of $\epsilon = 0.25$ to $\epsilon = 0.1$. With this tuning, there is good correspondence between the (N) and (W) statistics ((N) not shown) and the climate statistics of the unreduced model.

8.1.3 Comparison of Hasselmann and MTV approximations

In the 1-variable case, jet bimodality was generated by the untuned Hasselmann (N) models, but not the untuned MTV models. At $k^{-1} = 6.7$ days, the stationary mode statistics were captured by the 1-variable (W) and (N) models without tuning, whereas a tuning of the MTV tuning parameters only improved the first two moments and the temporal characteristics of the simulation. Multiple regime behaviour was induced by a tuning of multiple (artificially) free parameters. It was found in both the untuned Hasselmann models and this tuned MTV model that deterministic dynamics underlies jet bimodality, as evidenced by their similar drift potentials (Figure 8.2 and Figure 7.8). The MTV model was tuned to yield jet bimodality by giving greater weighting to the nonlinear noise-induced drift terms. By comparison, the untuned Hasselmann stochastic models, in which noise-induced drift terms are absent, generates jet bimodality through the averaged forcing. However, the mechanisms for generating jet bimodality do not necessarily differ between the reduced models, as will be discussed in Section 8.4.

At $k^{-1} = 12.0$ days as well, the 1-variable averaged forcing is responsible for the existence of jet bimodality. The (N) model significantly underestimated the recurrence timescale between visits to the low-latitude regime, but this was improved with only a slight decrease in the noise scaling. The noise, in addition to the noise-induced drift, was also downscaled in the 1-variable MTV model to generate a good approximation, but the effect of the downscaled effective climate feedback was to eliminate climate drift. In fact, jet bimodality could not be simulated with the MTV tuning procedure.

In the 3-variable case, the Hasselmann models outperform the MTV models in simulating the statistics of both modes. At $k^{-1} = 6.7$ days, the MTV model performed well only in simulating the stationary mode statistics, which were best captured with an upscaled effective climate feedback. The statistics of the Hasselmann (N) and MTV models appear to be converging as the noise and noise-induced drift are upscaled. As depicted in Figure 8.9, the jet shifts towards higher latitudes in both the (N) and MTV models with these terms upscaled. Wave-4 of the MTV model is insensitive to these rescalings, but the variance and lag-correlation timescale of wave-4 of the (N) model are reduced as the noise is increased.

However, the Hasselmann model is unbounded with sufficiently increased noise.

At $k^{-1} = 12.0$ days, the Hasselmann (N) and (W) models are bounded only with significantly downscaled noise, but they did generate the statistics of both modes with such a noise term. The untuned MTV model generated a good approximation of the stationary mode statistics, and the wave-4 statistics were well simulated with a significantly downscaled effective climate feedback. However, the downscaled effective feedback caused a climate drift in the stationary mode.

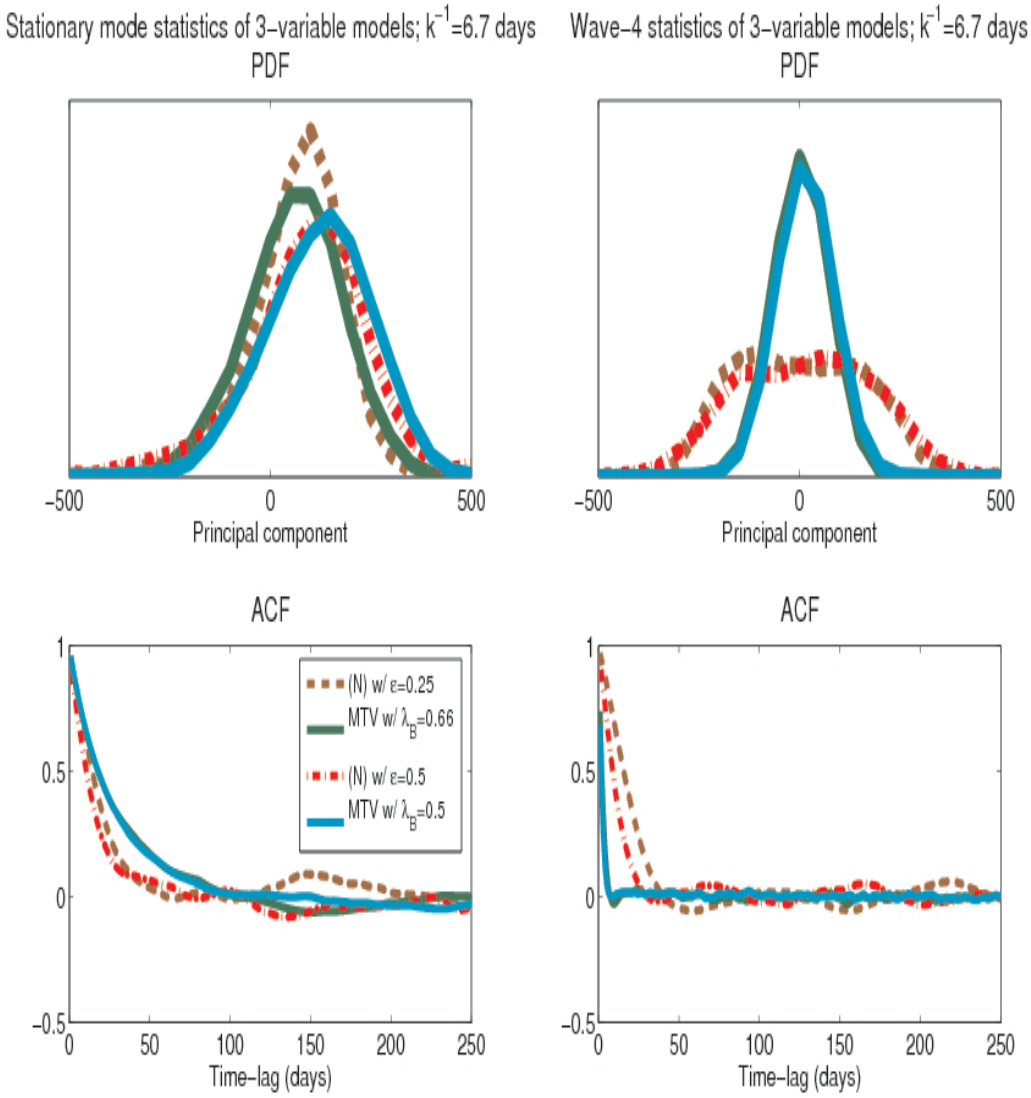

Figure 8.9: The tuned (N) model with upscaled noise and the MTV model with downscaled bare-truncation forcing (equivalently, upscaled noise and noise-induced drift).

8.2 Spin-down timescales of $k^{-1} = 2.3$ and 4.7 days

8.2.1 Deterministic averaging (A)

At $k^{-1} = 2.3$ days and at $k^{-1} = 4.7$ days, the stationary and wave-4 modes are the dominant slowly-evolving modes (at least in terms of variability) and there exists a single jet state. However, there are significant changes in the climate dynamics as the bifurcation point of $k^{-1} = 5.6$ days is approached, which are particularly manifest in the wave-4 spectrum. In particular, the power in the wave-4 spectrum becomes concentrated at progressively lower frequencies as $k^{-1} \to 5.6$ days (Figure 8.1). At $k^{-1} = 4.7$ days, the 'fundamental frequency' of wave-4 is on the order of $1/100$ days^{-1}.

The 3-variable averaged models at $k^{-1} = 2.3$ and 4.7 days capture these temporal changes as the bifurcation point is approached. At $k^{-1} = 2.3$ days, the ACF of the stationary mode of the averaged model decays with oscillation in excellent agreement with that of the unreduced model (Figure 8.10). The wave-4 ACFs of the averaged and unreduced models are in good agreement, with only a slight oscillation with lag in the averaged model ACF. At $k^{-1} = 4.7$ days, the ACFs of both the stationary and wave-4 modes decay with oscillation, but only the stationary mode ACF of the unreduced model behaves in this fashion (Figure 8.11). However, the low-frequency oscillation in the wave-4 ACF reflects the dominance of low-frequencies of the order of $1/100$ days^{-1} in the wave-4 spectrum.

At each spin-down timescale, the variance of the averaged model approximations is significantly less than the climate variance of the unreduced model. This is generally true, on both sides of the bifurcation point, as the noise variance arising from slow-fast interactions is eliminated through deterministic averaging. Note that at $k^{-1} = 4.7$ days alone, the stationary mode PDF of the averaged model is significantly shifted towards low latitudes.

It was demonstrated in Chapter 6 that this oscillation in the stationary mode ACF of the unreduced model is primarily due to the action of the propagating, synoptic-scale wave-5 mode, and not the wave-4 (climate) mode. Thus, the averaged effect of the synoptic-scale eddies on the climate forcing is sufficient to capture this low-frequency modulation of the stationary mode. Generally, it is expected that such deterministic, non-chaotic features in

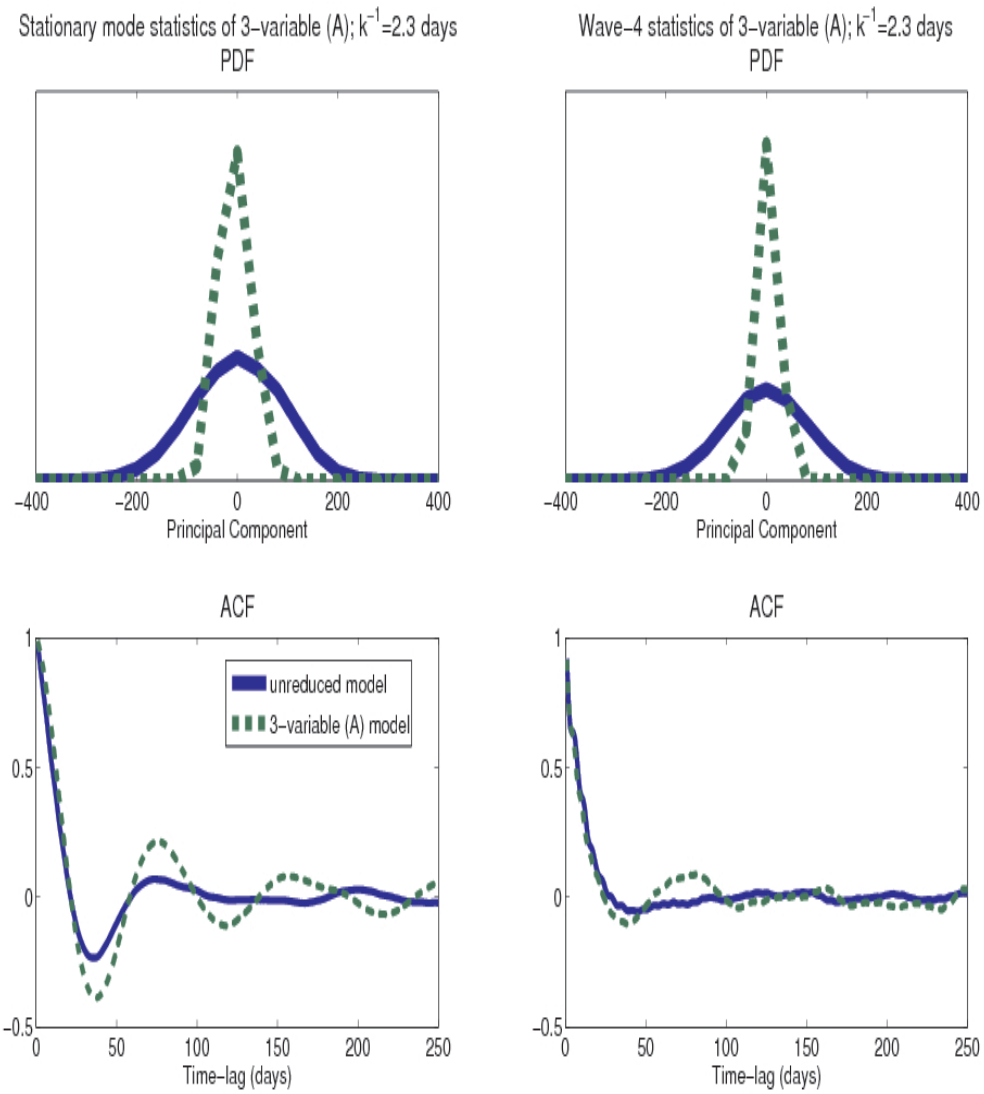

Figure 8.10: Statistics of 3-variable (A) model at $k^{-1} = 2.3$ days.

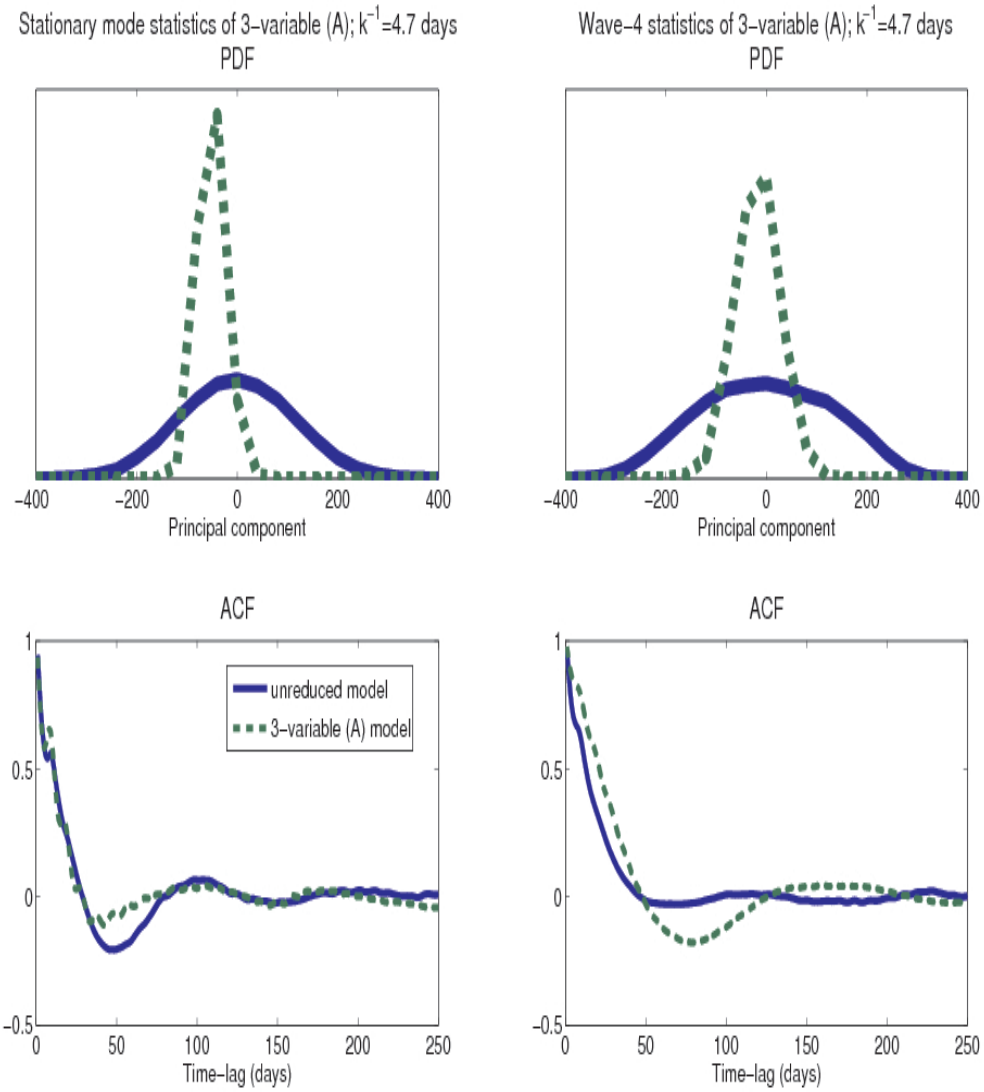

Figure 8.11: Statistics of 3-variable (A) model at $k^{-1} = 4.7$ days.

the climate dynamics, even if arising from climate-weather interactions, should be captured by deterministic averaging (assuming that there is sufficient timescale separation between climate and weather modes).

Despite the (deterministic) oscillatory components of the stationary and wave-4 modes in the reduced model, both modes have noisy time series. The stochastic nature of wave-4 in particular contrasts with the dominant oscillation in wave-4 on the side of the bifurcation point with jet bimodality. In fact, only the wave-4 component, and not the stationary mode component, of the 3-variable averaged solution undergoes significant changes in crossing the bifurcation point. The noisiness of the wave-4 and stationary modes suggests that there is coupling between these modes, in contrast to the apparent decoupling of these modes on the side of the bifurcation point with jet bimodality. At $k^{-1} = 6.7$ and 12.0 days, the oscillation of wave-4 of the 3-variable averaged dynamics is characteristic of a 2-dimensional ODE that is independent of the stationary mode. At $k^{-1} = 2.3$ and 4.7 days, the noisiness of both the stationary and wave-4 modes implies that the 3-variable averaged dynamics is derived from a 3-dimensional ODE with mutually interacting variables.

The potential functions of the 1-variable averaged models at $k^{-1} = 2.3$ and 4.7 days are concave up (Figure 8.12), unlike the potential at $k^{-1} = 6.7$ days, which has a large plateau in potential corresponding to the location of the low-latitude regime (Figure 8.2). At $k^{-1} = 5.3$ days, at a point very close to the bifurcation point, the potential remains concave up (not shown). These potentials are consistent with the existence of a single jet state regime for values of the spin-down time below the bifurcation point.

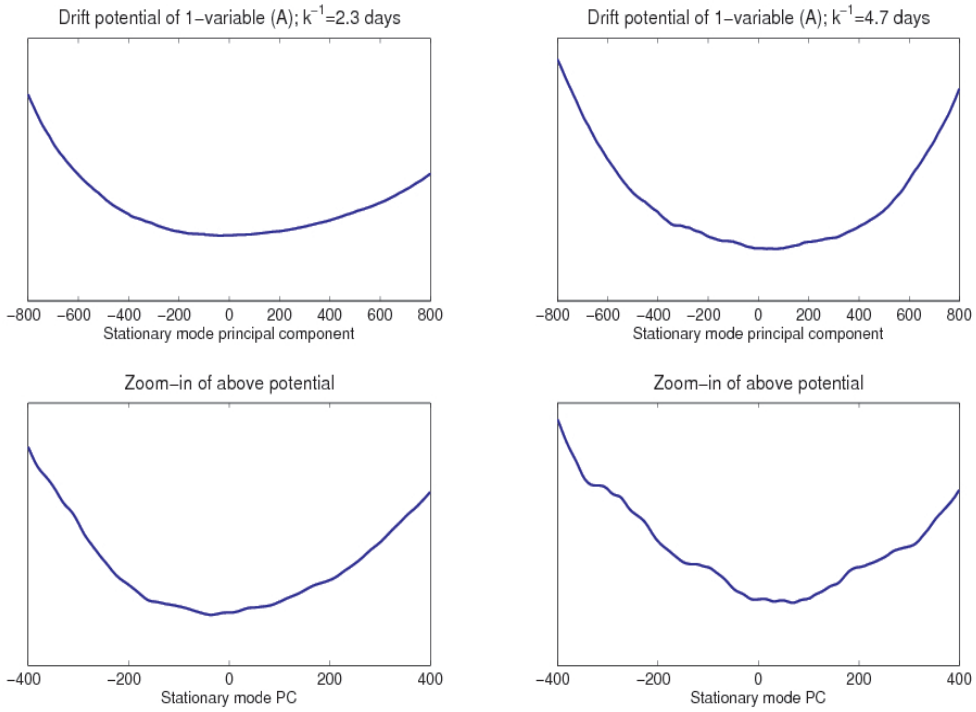

Figure 8.12: Potential of the 1-variable averaged functions in the region without jet bimodality. It is suggested by these plots that the corresponding 1-variable (W) approximations will not generate jet bimodality.

8.2.2 Diffusion approximations (W) and (N)

At $k^{-1} = 4.7$ days, the 1-variable (W) and (N) models with a priori determined values of ϵ and γ generate similar statistics which reasonably approximate the stationary mode statistics of the unreduced model (Figure 8.13). The stochastic models generate the Gaussian distribution of the unreduced model stationary mode, but with a slightly increased variance. The ACFs of the stochastic and unreduced models decay similarly, but the stochastic models fail to capture the low-frequency oscillation with lag.

At the very low spin-down timescale of $k^{-1} = 2.3$ days, the 1-variable (N) and (W) models yield almost identical results. The stochastic models reasonably approximate the stationary mode statistics of the unreduced model, but the PDF has slightly decreased variance and the ACF does not decay with oscillation. The 1-variable models are not expected to generate this oscillation as an oscillation requires at least two degrees of freedom.

Using the a priori determined scaling parameters, the 3-variable stochastic approxima-

tions are bounded at $k^{-1} = 2.3$ days but not at $k^{-1} = 4.7$ days. At the latter spin-down timescale, the Hasselmann (N) and (W) models with sufficiently decreased noise scaling are bounded. As depicted in Figure 8.14, the (N) model with noise scaling decreased from the a priori determined value of $\epsilon = 0.25$ to $\epsilon = 0.166$ approximately captures the variance of both modes, but there is a significant climate drift towards higher latitudes evident in the shifted stationary mode PDF. As ϵ is decreased, the stationary mode shifts towards lower latitudes from its position with the setting $\epsilon = 0.166$, but the variance also decreases. The (N) model with $\epsilon = 0.166$ does capture the decay of the ACFs of both modes, but not the low-frequency modulation of the stationary mode.

At $k^{-1} = 2.3$ days, the 3-variable (N) model approximation with a priori determined parameters is bounded and performs reasonably well in simulating the climate statistics (Figure 8.15). The two wave-4 components of the (N) model do not significantly differ from each other (and are each very well simulated). The stationary mode of the (N) model has too little variance, but the decay of the ACF is well captured. Despite the ability of the 3-variable (A) model to capture the low-frequency modulation of the stationary mode, this feature is lost in the 3-variable (N) model.

In fact, the 3-variable (N) approximation at $k^{-1} = 2.3$ days is bounded only by taking the lag-covariance up to a lag of at least 2 hours (and adjusting the extrapolation factor appropriately) in computation of σ. By comparison, the 3-variable (N) model at $k^{-1} = 4.7$ days is insensitive to such changes; in particular, the 3-variable (N) model with $\epsilon = 0.25$ remains unbounded with a larger upper time lag.

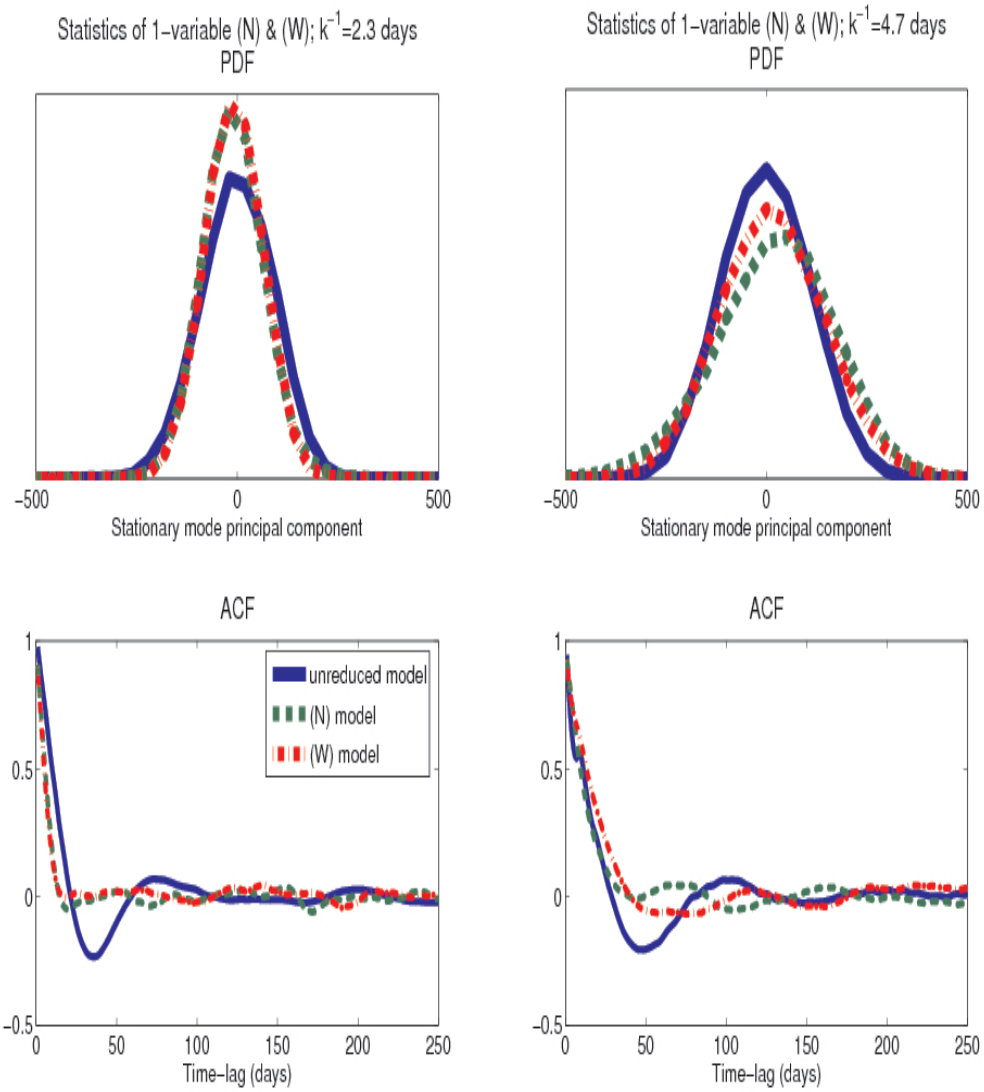

Figure 8.13: Statistics of the 1-variable (N) and (W) models in the region without jet bimodality.

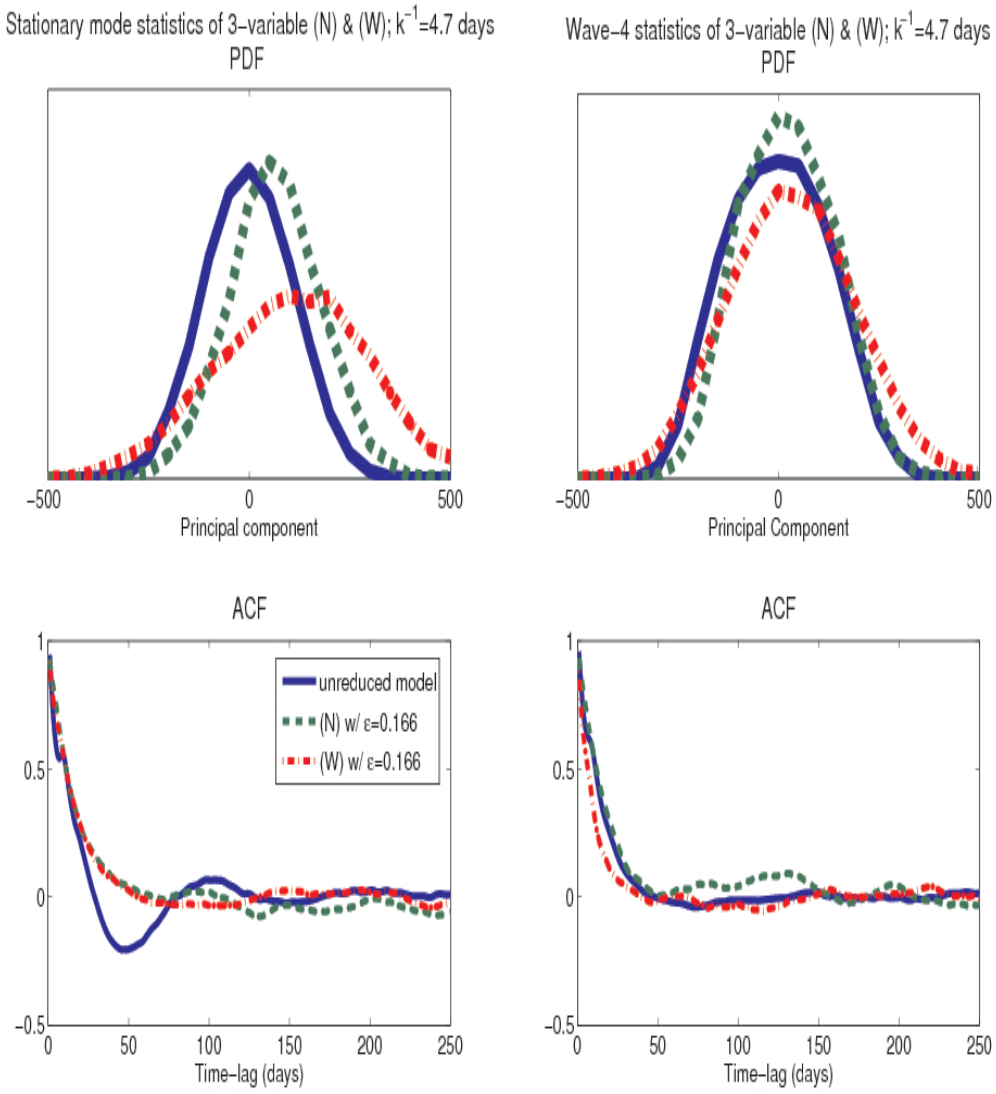

Figure 8.14: Statistics of the 3-variable (N) model with the noise scaling parameter decreased to $\epsilon = 0.166$.

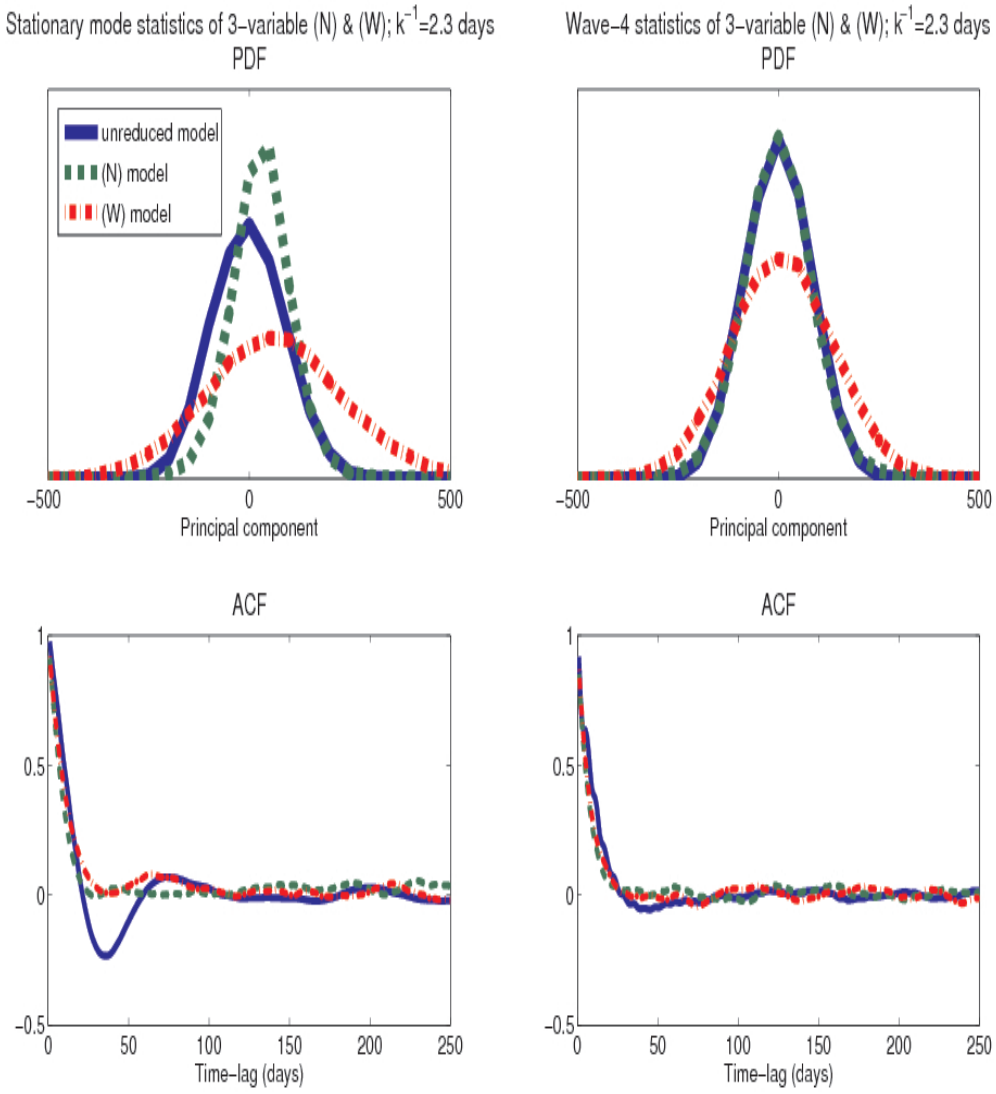

Figure 8.15: Statistics of the 3-variable (N) and (W) models at $k^{-1} = 2.3$ days with a priori determined values of the scaling parameters.

8.2.3 Comparison of Hasselmann and MTV approximations

On the side of the bifurcation point without jet bimodality, both the 1-variable Hasselmann and MTV models perform reasonably well in capturing the stationary mode statistics of the unreduced model.

At $k^{-1} = 4.7$ days, the untuned 3-variable MTV model outperforms the tuned 3-variable Hasselmann (N) models in simulating the stationary mode statistics of the unreduced model. The untuned (N) model has a large climate drift, and is ultimately (exponentially) unstable, whereas the MTV model has no drift. As at $k^{-1} = 6.7$ days, both types of models respond similarly to a rescaling of the noise and noise-induced drift terms. As depicted in Figure 8.16, there is excellent correspondence between the (N) model with $\epsilon = 0.166$ and the MTV model with the bare truncation forcing downscaled by setting (coincidentally) $\lambda_B = 0.166$. As the noise scaling in the 3-variable (N) model is decreased and the scaling of the bare truncation forcing in the MTV model is increased, the jet shifts to lower latitudes in both cases. However, the untuned MTV model (with $\lambda_B = 1$) yields a good approximation of the stationary mode statistics whereas the stationary mode of the tuned (N) model with a correctly centred jet has too little variance.

At $k^{-1} = 2.3$ days, the untuned 3-variable Hasselmann (N) outperforms the untuned 3-variable MTV model. There is little difference between the statistics of the two wave-4 components of the (N) model whereas the lag-correlation timescales of the two components of the MTV model significantly differ. With a downscaled bare truncation forcing, jet bimodality was induced in the MTV model but not in the (N) model.

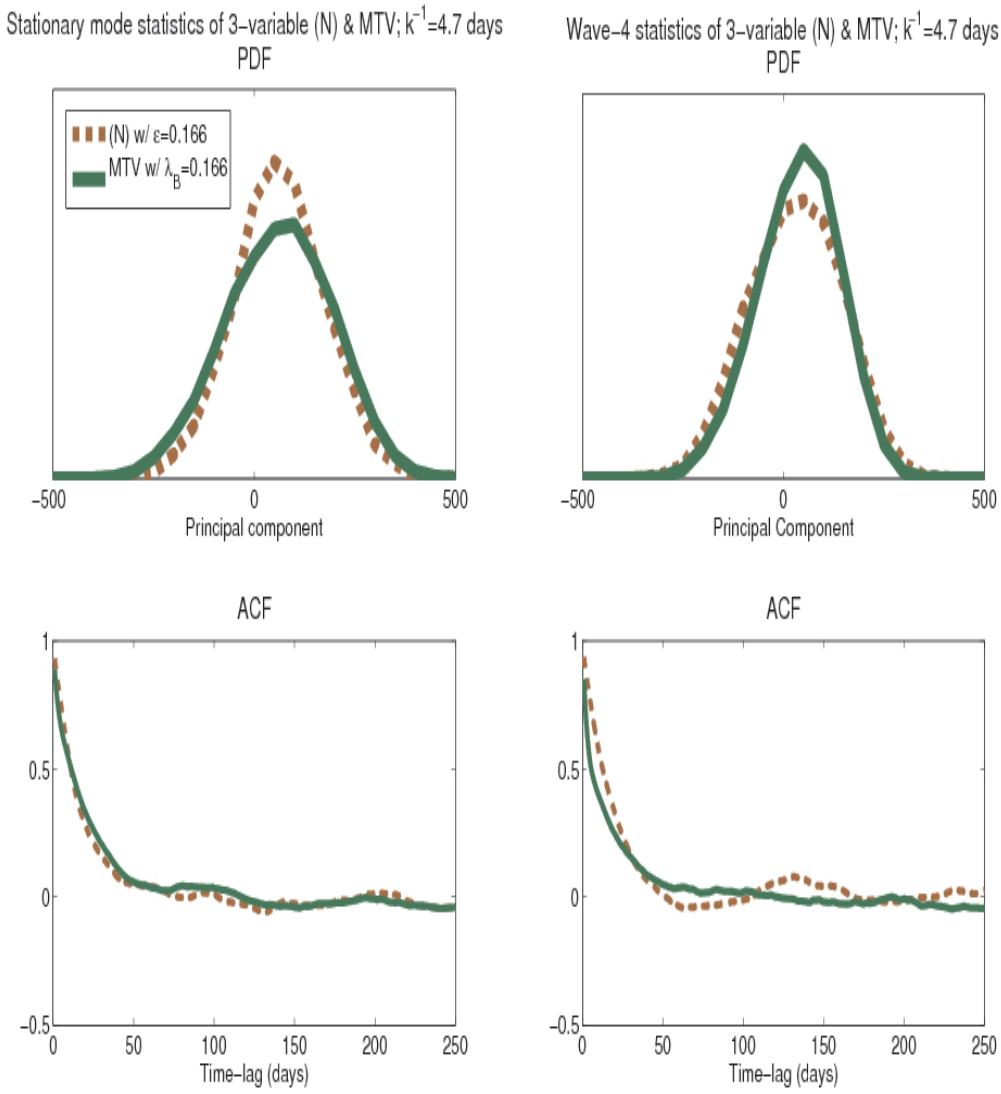

Figure 8.16: Statistics of the tuned 3-variable (N) and MTV models at $k^{-1} = 4.7$ days. They respond similarly to an increase in the importance of noise (and noise-induced drift).

8.3 Hybrid approximations

The above deterministic and stochastic approximations were found without consideration for computational efficiency. In particular, long spin-up times and large numbers of ensemble members were used in integration of the (A), (W) and (N) equations. In this section, computationally efficient, accurate approximations of the unreduced model statistics are sought. To this end, the same (cumulative averaging) scheme is applied as in computation of Hasselmann's approximations, but with parameter settings that yield efficient integrations. Generally, these efficiently derived solutions lie somewhere in between the solutions of a Hasselmann model and the unreduced model. Accordingly, these approximations are called 'hybrid approximations' (cf. Chapter 3). Depending on which Hasselmann model is used, the solution may be an (A), (W) or (N)-unreduced hybrid approximation. Since hybrid models are computationally efficient relative to the unreduced model and simple to construct, these models are potentially useful as (super)parameterisations in complex models. However, a hybrid approximation is not useful for dynamical interpretation, as the degree to which it resembles a Hasselmann approximation is unknown. It is thus necessary to base dynamical interpretation on Hasselmann's approximations as derived in the previous sections.

In this first part below, the emphasis is placed on maximizing the real time speed of integration. By any measure of efficiency, it is necessary that the macrotime step is greater than the integration time of the weather equations at each climate realisation; i.e., $\delta \equiv \Delta t/(N\delta t) > 1$, such that the weather modes are effectively slowed down (cf. Chapter 3). By applying the re-initialisation rule (3.2) to minimize spin-up time, it is possible to derive an approximation with $\Delta t \gg N\delta t$ and without averaging (i.e., $N - N_1 = 1$ and $R = 1$) in systems with sufficient timescale separation and with a rapidly mixing weather. Applied to the KRG05 model, these settings do not yield a good approximation (see below), which necessitates averaging over multiple weather iterates at each macrotime step. Rather than finding the average over a single weather realisation, multiple ensembles are used for this purpose (i.e., $N - N_1 = 1$ and $R > 1$). Integration of these hybrid models can be made (much) faster than integration of the unreduced models in real time through parallel

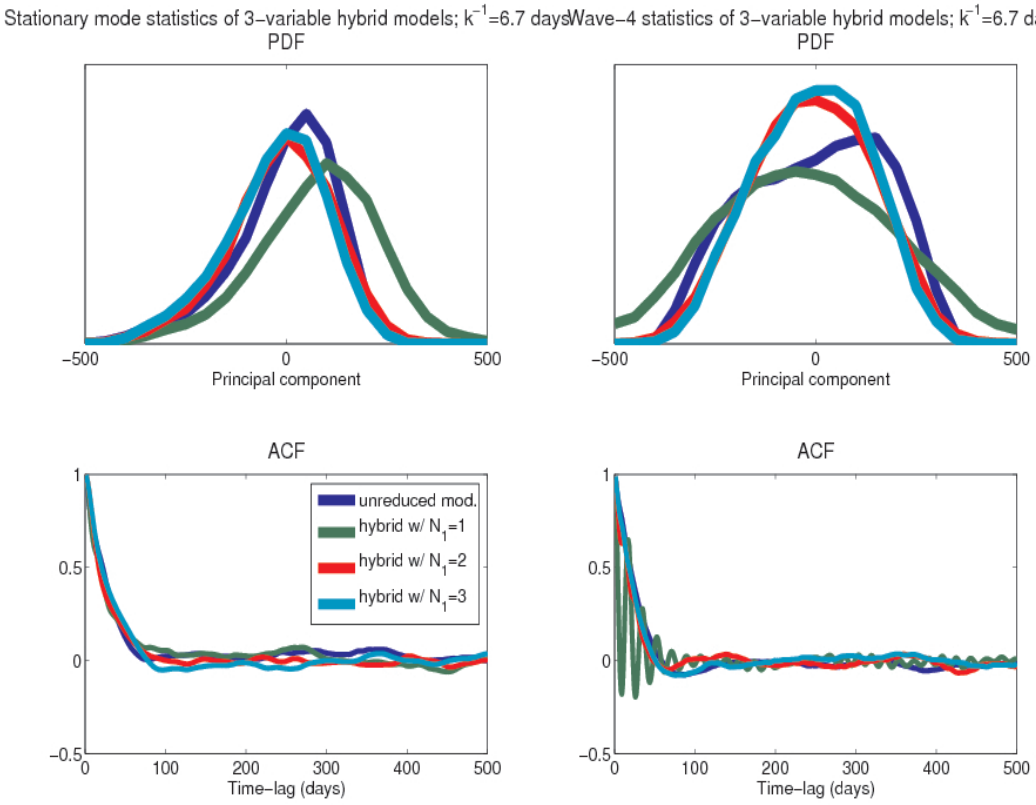

Figure 8.17: Hybrid (A)-unreduced model with the number of spin-up steps N_1 varied as indicated.

computing.

The potential efficiency gain of the hybrid models is explored below by considering the 3-variable case at $k^{-1} = 6.7$ days. These hybrid models are integrated with $\Delta t = 12\delta t = 2$ hours and $N - N_1 = 1$, and R and N_1 are varied to determine the settings that minimize computation time and yield good approximations of the climate statistics of the unreduced model.

The 3-variable (A)-unreduced hybrid models are first considered. In Figure 8.17, the model statistics are shown for $R = 4$ ensemble members and with the number of spin-up steps varied as $N_1 = 1, 2$ and 3, for an efficiency gain of a factor $6, 4$ and 3, respectively. With $N_1 = 1$, the ACF of wave-4 oscillates rapidly, suggesting that this amount of spin-up is insufficient to reach the weather attractor. With one more spin-up time step, the wave-4 ACF no longer oscillates and the hybrid model statistics are in reasonable agreement with those of the unreduced model. With an increase to $N_1 = 3$, the statistics do not

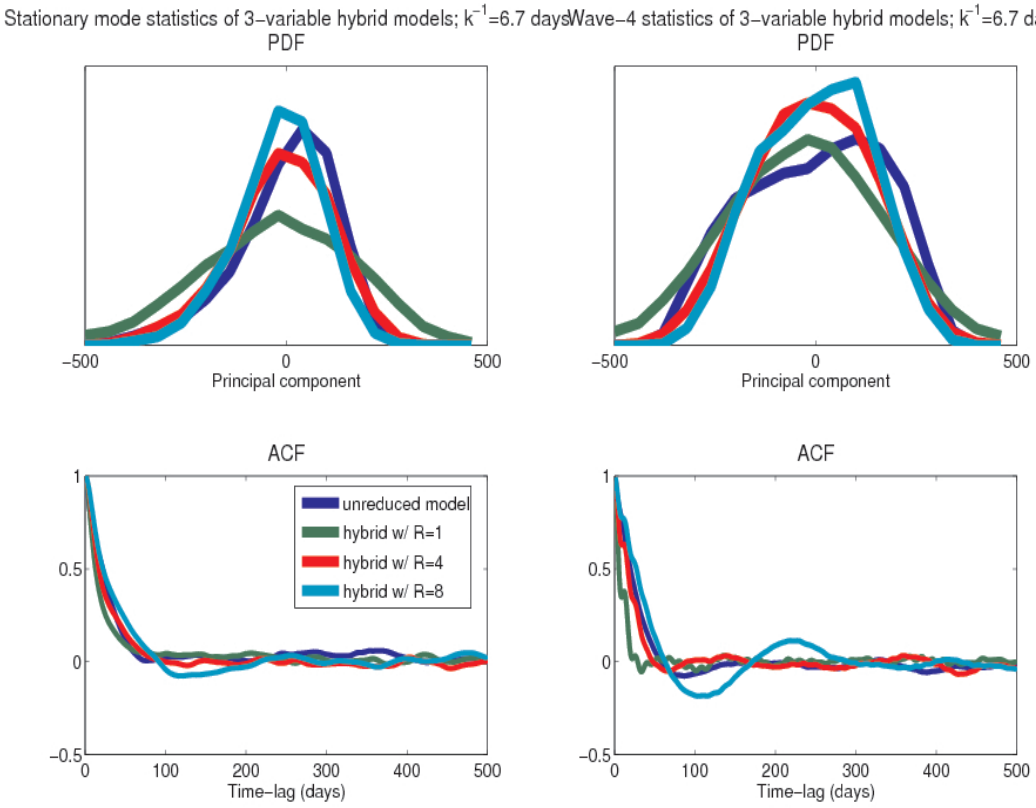

Figure 8.18: Hybrid (A)-unreduced model with number of averaging ensemble members R varied as indicated.

change significantly, suggesting that a spin-up of only 20 minutes (and altogether, a weather integration time of 30 minutes) is needed when the climate macrotime step is 2 hours.

In Figure 8.18, the number of spin-up steps is fixed at $N_1 = 2$ and the number of ensemble members R, over which the average is taken at each macrotime step, is varied. With $R = 1$, the approximation is poor as the system has too much variance and there is no evidence of jet bimodality. With $R = 4$, the hybrid model statistics are a good fit to those of the unreduced model, including the pronounced skewness in the stationary mode PDF that is the manifestation of multiple regime behaviour. Not surprisingly, the 3-variable hybrid approximation with $R = 8$ (averaging) ensemble members best resembles the 3-variable (A) approximation, as is most apparent from the oscillation with lag in the hybrid model ACF. This feature of the ACF is the manifestation of the dominant oscillation in the wave-4 dynamics of the averaged model.

The statistics of the hybrid (W)-unreduced and (N)-unreduced models with a priori determined scaling parameters do not significantly differ from those of the (A)-unreduced model. This was also the case in Fatkullin and Vanden-Eijnden (2004), in which the computationally efficient averaged models were successful in simulating the climate statistics of a model (the Lorenz-96 (L96) model) with very large scale separation between climate and weather variables. They claimed that the L96 model with the appropriate parameter settings is "intrinsically stochastic on the $\mathcal{O}(1)$ timescale, and the small corrections arising on the $\mathcal{O}(\epsilon^{-1})$ timescale have a very small effect on the long-time statistical properties of the system". This may be true of the L96 model, but here it is evident that the efficient hybrid (A)-unreduced models significantly differ from the 3-variable (A) model at $k^{-1} = 6.7$ days, which is characterised by a significantly decreased variance and a dominant wave-4 oscillation.

Although the macrotime step of $\Delta t = 2$ hours is greater than the the KRG05 model time step (and weather time step) of $\delta t = 10$ minutes by an order of magnitude, it is a small step compared to the $\mathcal{O}(month)$ timescale of the stationary and wave-4 modes. Indeed, the macrotime step can be increased by an order of magnitude and again R and N_1 can be varied, but the efficiency gain in generating a good approximation of the climate statistics is no greater than with the smaller macrotime step. However, there appears to be significant potential for efficiency gain, which can be realised with an appropriate (but as yet undetermined) re-intialisation scheme.

To illustrate this potential, consider that a direct numerical simulation (DNS) of the stationary and wave-4 modes is obtained with the settings $R = N - N_1 = 1$, $\Delta t = 24$ hours and $N\delta t = 24$ hours (such that $\delta \equiv \Delta t/(N\delta t) = 1$). In the 1-variable case in which only the stationary mode is defined as a climate mode, it is possible to increase the macrotime step and spin-up time to 48 hours. It is thus sufficient in simulating the climate statistics to (appropriately) sample the weather variable only once per day.

With these same settings but with increased spin-up (i.e., with $\delta < 1$), intriguing results are obtained. In the 3-variable case, the *deterministic* (A)-unreduced hybrid model with $\Delta t = 24$ hours and small δ yields a good approximation of the *stochastic multiplicative*

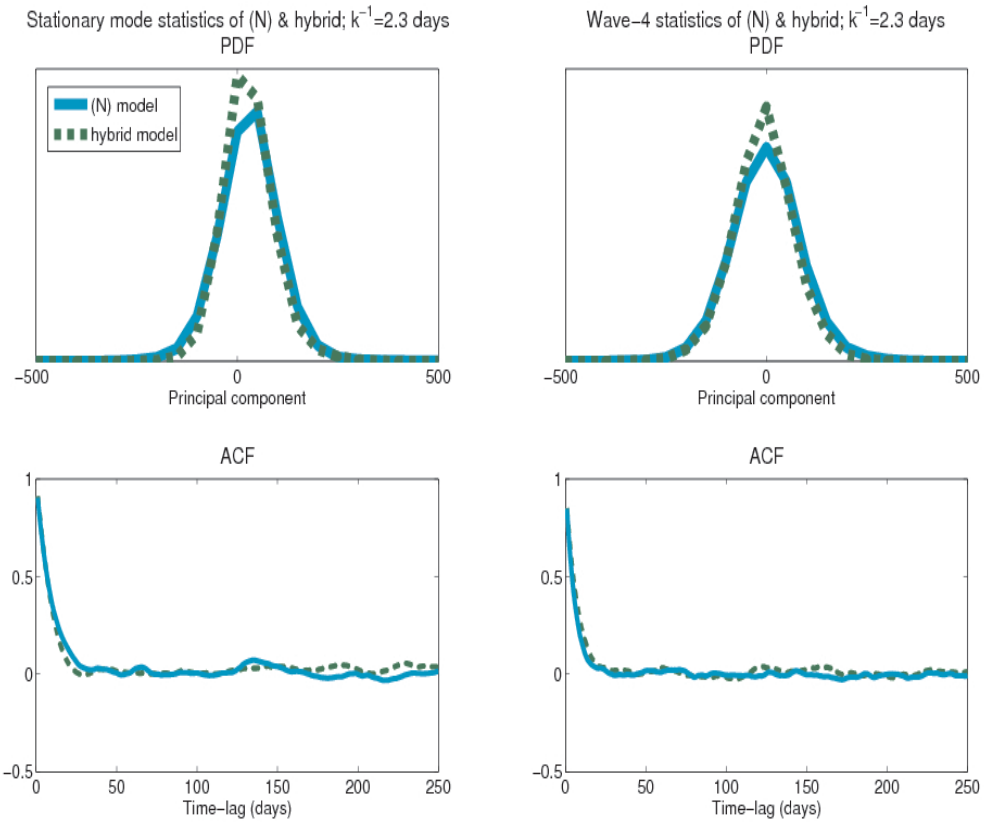

Figure 8.19: The 3-variable (N) model and the hybrid (A)-unreduced model with $\delta = 0.1$ and $\Delta t = 24$ hours.

(N) model approximation. At $k^{-1} = 2.3$ days, the hybrid model with a 3-variable climate yields statistics which are invariant for $\delta < 0.1$. As depicted in (Figure 8.19), there is excellent agreement between these statistics and those of the 3-variable (N) model (which were determined in Section 8.2). At $k^{-1} = 4.7$ days and at $k^{-1} = 6.7$ days, the stationary mode PDF of the hybrid model shifts towards higher latitudes with decreasing δ. For $\delta < 0.1$, the climate component of the hybrid model integration is unstable. Similarly, the 3-variable (N) model at each of these spin-down timescales is characterised by a shifting stationary mode PDF with an increase in the noise scaling and an unbounded solution for sufficiently increased noise scaling (Figure 8.20).

In the 1-variable case, the hybrid model solutions at all spin-down timescales are invariant for $\delta < 0.1$. On the side of the bifurcation point with jet bimodality, the hybrid models with a time step of $\Delta t = 48$ hours capture the jet bimodality generated by the

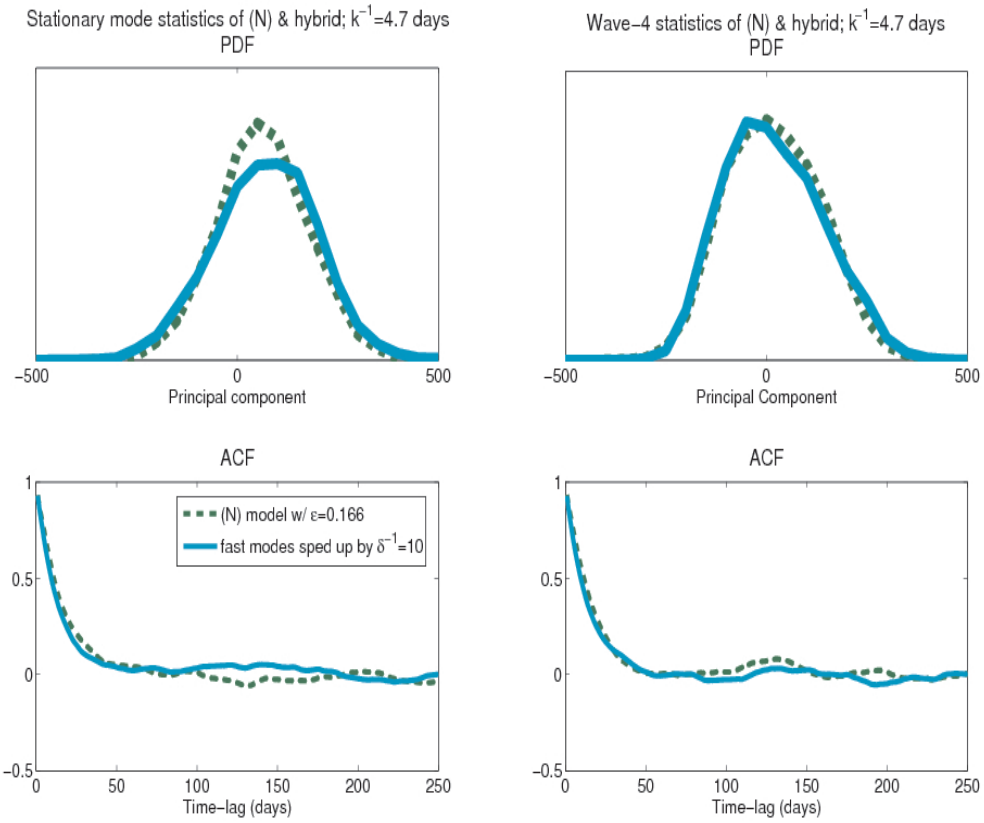

Figure 8.20: As in 8.19. For $\delta < 0.1$, the hybrid model solution is exponentially unbounded.

1-variable (N) models (integrated with an $\mathcal{O}(1)$ hour time step). The ability of this model to capture the state-dependent noise of the (N) model is exemplified at the high spin-down timescale of $k^{-1} = 12.0$ days. As depicted in Figure 8.21, the hybrid model captures the pronounced multiple regime behaviour particular to the 1-variable (N) model, including the years timescale persistence within the low-latitude regime and the decades timescale between excursions to the low-latitude regime.

With $\delta < 1$ the fast modes are sped up, and with no averaging at each climate time step, the hybrid (A)-unreduced model is closest to the unreduced model. Assuming that there exists limiting dynamics, the slow component of the unreduced model with fast modes sped up approaches the solution of the averaged (A) equation as $\delta \to 0$. Indeed, with these settings and with a sufficiently small Δt, the (A) solution, characterised by the dominant wave-4 oscillation is evident (not shown). With a large Δt on the order of days, a (state-dependent) noise error is introduced into the solution. In application to the KRG05 model,

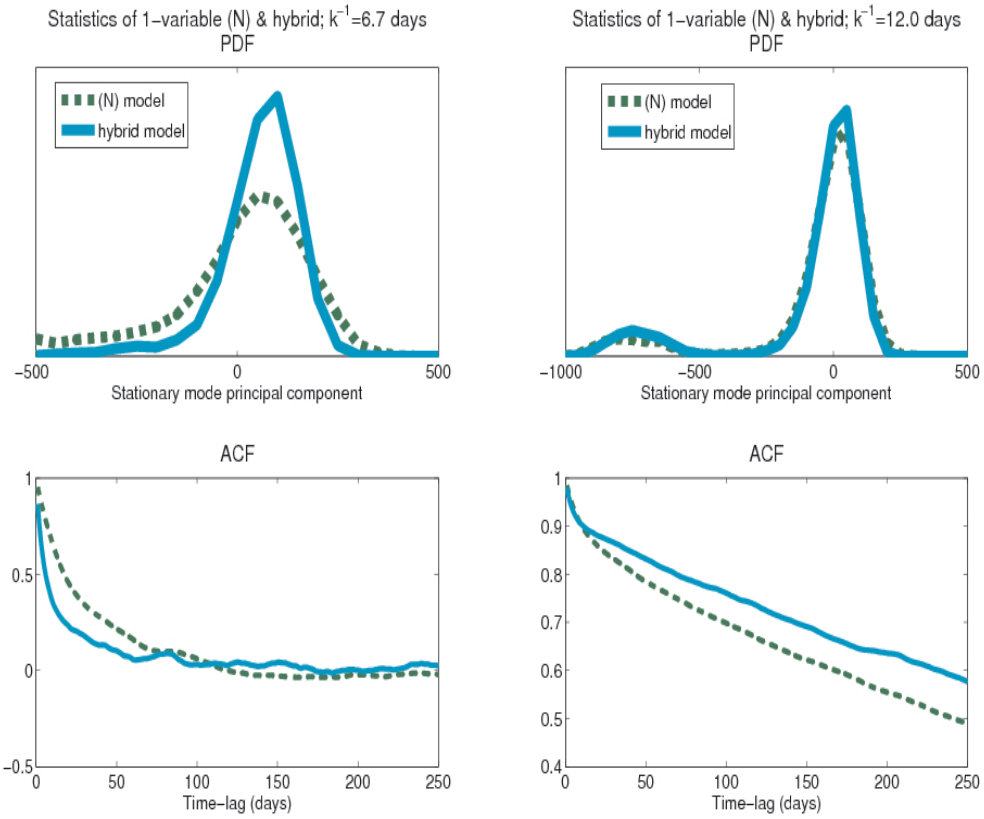

Figure 8.21: The 1-variable (N) models and hybrid (A)-unreduced model with $\delta = 0.1$ and $\Delta t = 48$ hours.

it appears that for an appropriately sized climate time step that this state-dependent noise error (that arises in approximation of the limiting/averaged dynamics) in combination with the (coarsely) approximated averaged forcing generates the same statistics as the (N) model.

8.4 Summary and discussion

Deterministic averaging of the climate forcing proved to be a powerful reduction method. In particular, the 1-variable averaged forcing in the stationary mode alone captures the dynamics underlying jet bimodality at $k^{-1} = 6.7$ days, particularly, the nonlinear interaction among the stationary, wave-4 and wave-5 modes. At this spin-down timescale, the 1-variable averaged system had to be only energized by additive noise to generate the statistics of the stationary mode, in which the jet bimodality is most manifest. In fact, the 1-variable (W) additive noise model performed at least as well as the 1-variable (N) multiplicative noise model in simulating the stationary mode statistics of the unreduced model at all spin-down timescales considered. However, at the high spin-down timescale of $k^{-1} = 12.0$ days, multiplicative noise had an impact on very long timescales in determining the nature of the (very infrequent) visits to the low-latitude regime.

The 3-variable averaged ODEs generated long timescale features such as the low-frequency oscillation in wave-4 and the low-frequency modulation of the stationary mode. With multiplicative noise, only a weak jet bimodality was generated at $k^{-1} = 6.7$ days, which was made stronger with increased noise scaling, but with the effect of increasing the climate drift. In fact, in a neighbourhood of the bifurcation point (including $k^{-1} = 4.7$ and 6.7 days), the 3-variable (N) SDEs generated strong climate drifts, manifest in their shifted stationary mode PDFs. At the very high spin-down timescale of $k^{-1} = 12.0$ days, the solutions of the stochastic models were unbounded with the a priori determined values of the scaling parameters, but with a significantly downscaled noise, the (N) and (W) models generated very good approximations of the statistics of the stationary and wave-4 modes. At the low-spin down timescale of $k^{-1} = 2.3$ days, the climate statistics were reasonably well approximated by the (N) model with the a priori determined parameters.

In construction of the 3-variable (W) models, it was assumed that the diffusion matrix has the form $\sqrt{\epsilon}\gamma I$, where ϵ and γ are scaling factors and I is the identity matrix. A diffusion matrix of this simple form was used successfully in Whitaker and Sardeshmukh (1998). Through the steady-state fluctuation-dissipation relation (FDR), the geographical distribution of synoptic eddy covariance (in a linearised model) was linked solely to the

spatial structure of the synoptic eddy drift by assuming an additive noise forcing with a diffusion matrix of this form. In the Hasselmann framework, the diffusion matrix $\sqrt{\epsilon}\gamma I$ is valid if with respect to the joint distribution of the lagged weather variables the cross-correlations of the three climate forcings are zero (simply, the crossed climate forcings are not statistically linked via the weather) and the auto-correlations are equal to each other. This appears to be the case only at the high spin-down timescale of $k^{-1} = 12.0$ days, at which wave-4, dominated by (deterministic) oscillations, is to first-order decoupled from the weather modes (and is thus not statistically linked to the stationary mode via the weather). As in Whitaker and Sardeshmukh (1998), the simple diagonal diffusion matrix is valid only if the important interactions occur among the resolved climate modes.

As the (W) model is not rigorously defined, alternative formulations of the diffusion matrix can be considered. One possibility is to find the average $< \cdot >$ over the range of climate realisations as $\sigma_{W*}\sigma_{W*}^T =< \sigma(a)\sigma(a)^T >$, where σ_{W*} is the diffusion matrix of the re-defined additive noise (W*) model. For example, σ_{W*} at $k^{-1} = 6.7$ days may be derived from a simulation of the 3-variable (N) model at $k^{-1} = 6.7$ days with the a priori determined scaling parameter of $\epsilon = 0.1$. Recall that the stationary mode variance was underestimated in this (N) model. As depicted in Figure 8.22, the 3-variable (W*) model captures the jet bimodality manifest in the stationary mode, but its simulation of the wave-4 statistics does not significantly differ from that of the 3-variable (W) model. In particular, the lag-correlation timescale of wave-4 is significantly underestimated by the additive noise models, but not by the (N) model. As the (W*) model is not a theoretically justified model, it cannot be expected to outperform the (N) model in general. However, it is an interesting direction for future work to see under which circumstances the (W*) model or a similar such additive noise model outperforms the theoretically-justified multiplicative noise (N) model.

The (W*) model is an improvement over the (N) model with upscaled noise because of the climate drift in the latter model. In fact, in a neighbourhood of the bifurcation point, the (N) approximation has a climate drift that is strongly dependent on the scale separation parameter ϵ. This result begs the question of whether the (N+) model, with

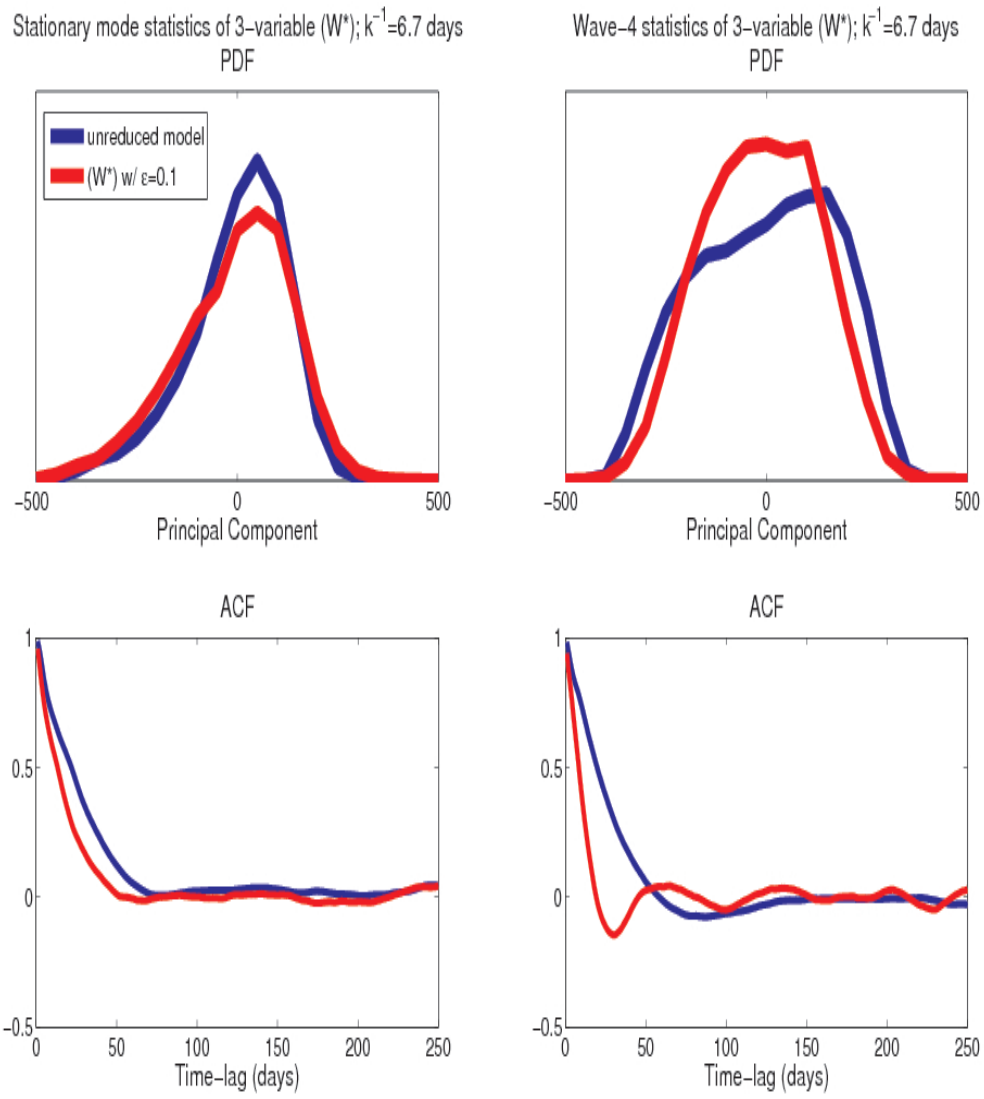

Figure 8.22: Statistics of the 3-variable (W*) model (see text).

noise-induced drift correction, would prevent such a climate drift, and in general would better capture the climate statistics. The (N+) model is the only Hasselmann model that is strictly valid on the long $\mathcal{O}(\epsilon^{-1})$ timescale. The (N) model performed very well simulating the ocean statistics in the idealised 6-variable atmosphere-ocean model considered in AI03. However, the scale separation in this model was set to extremely small values of $\mathcal{O}(10^{-5})$, such that the $\mathcal{O}(1)$ timescale on which the (A) and (N) models are valid may be longer than the timescales of interest in the ocean model. Moreover, on the $\mathcal{O}(1)$ timescale, the error in approximation is expected to be very small because of the large scale separation parameter. With the significantly smaller timescale separation and long timescale features of the KRG05 model, it is not guaranteed that these features would be well captured by the (N) model.

In the 1-variable model at $k^{-1} = 6.7$ days, multiplicative noise was not required to produce jet bimodality. The potential generated by averaging has an extended shoulder corresponding to the low-latitude regime. At least in this model, the long timescale correction of noise-induced drift was not needed to capture the stationary mode statistics. It does not follow that the long timescale correction will also be unnecessary in the 3-variable case. Nor does it follow that the success of the reduction method with wave-4 included among the weather modes precludes the success with wave-4 included among the climate modes. Rather, the success of the 1-variable model simply implies that the detailed weather dynamics are not important to the dynamics underlying jet bimodality. It would be expected that with an additional resolved mode - the wave-4 mode - and with the relatively large scale separation between the stationary and wave-4 modes and the weather modes that the 3-variable model would perform at least as well as the 1-variable model in generating jet bimodality. Of course, the best way to determine the effect of the noise-induced drift is to integrate the (N+) model itself.

To better understand the success of the 1-variable Hasselmann stochastic model in generating the correct jet bimodality, it is informative to compare this model with the corresponding MTV model. As mentioned, deterministic averaged forcing, and not state-dependent noise nor noise-induced drift, is responsible for generating jet bimodality in the

Hasselmann models. By comparison, jet bimodality was induced in the 1-variable MTV model by scaling up the nonlinear noise-induced drift. However, it is possible that the mechanisms by which jet bimodality was induced in the Hasselmann and MTV models are the same. The average in the MTV model is with respect to the climate-independent distribution of weather variables, whereas the weather distribution in the Hasselmann model is derived from the full climate-dependent weather equation. The climate-dependent component of the weather equation factors into the effective MTV equation through the noise-induced drift. The explicit dependence of the noise-induced drift on this component does not derive from $\nabla_x y$ of the (N+) SDE because of the independence of the climate and weather processes in the MTV framework. Thus, the deterministic averaged forcing of the 1-variable model at $k^{-1} = 6.7$ days may include the effect of the noise-induced drift terms responsible for jet bimodality in the MTV model. However, since parameter tuning of the MTV model leads to ambiguities in dynamical interpretation, jet bimodality cannot unambiguously be ascribed to the quadratic and cubic nonlinearities of the MTV model. Furthermore, there is no contradiction between the results presented in this chapter and the conclusion in Chapter 6 that jet bimodality arises due to a complicated interaction between the stationary, wave-4, and wave-5 modes. The averaging procedure accounts for these interactions in averaged form, parameterising the effects of the unresolved modes on the dynamics of the resolved modes.

As a parameterisation scheme, the Hasselmann method as implemented does not have the computational efficiency of the MTV method. It is instead analogous (in a general sense; cf. Chapter 3) to the superparameterisation method of Grabowski (2001), the goal of which was to improve the more efficient but less accurate parameterisation schemes of cloud processes. Here as well, the Hasselmann models provided a superior approximation of the climate dynamics compared to the MTV models, especially in the 3-variable case, by more accurately accounting for the effects of the fast dynamics through online sampling of the conditional weather distribution. The Hasselmann integrations were made efficient compared to those of the unreduced model through the 'path-following' re-initialisation and cumulative averaging schemes. Importantly, these schemes are simple to implement

in complex climate models, requiring minimal changes to the original unreduced model algorithms. At the very least, the Hasselmann method can be (readily) implemented to determine whether the MTV model, as a special case of the (N+) model in particular, can be expected to produce good results. If the Hasselmann models generate poor results, then it is expected that the untuned MTV model will not generate better results.

Chapter 9

Conclusion

In this dissertation, rigorously-based reduction methods were developed and applied to a quasi-geostrophic model of mid-latitude low-frequency variability (the KRG05 model) to derive deterministic and stochastic models in the slowly-evolving (planetary scale) modes alone. Hasselmann's method was developed from limiting theories for slow-fast systems of ODEs and applied to the KRG05 model of intermediate complexity comprised of PDEs with complicated boundary conditions. Seamless (i.e., 'equation-free') superparameterisation schemes for integrating the (deterministic and stochastic) reduced models were developed, which required minimal changes to the KRG05 model. In particular, the reduced approximations were derived without the need to resolve the fast modes and the tendency forcing of the slow modes of the unreduced model. For comparison, the MTV method as implemented in Franzke et al. (2005) (FMV05) and Franzke and Majda (2006) (FM06) was applied to the KRG05 model. The MTV model is a stochastic model derived under stricter assumptions than in derivation of the Hasselmann models.

Prior to deriving the reduced approximations, parameters of the Hasselmann stochastic equations were computed. It was found that in computation of the diffusion term σ that the lag-covariance integrated (in principle) over an infinite time-interval can be well approximated by the lag-covariance integrated over a short time-interval and extrapolated by an a priori determined factor. This (accurate) simplification is necessary to efficiently integrate the Hasselmann stochastic equations. Tests of the underlying dynamics were also developed to aid in interpretation of the reduced models. In particular, the seamless reduction schemes were adapted to integrate the (unreduced) KRG05 model having suppressed

particular nonlinear interactions of potential importance to LFV (jet bimodality in particular). Suppressing the nonlinear interaction between the two planetary-scale modes (the stationary and wave-4 modes) did not significantly alter the model jet bimodality, implying that this interaction is not responsible for jet bimodality, in contrast to the conclusions of KRG05. Rather, it appeared that jet bimodality resulted from the nonlinear interaction among these two modes and the leading synoptic-scale mode (wave-5).

As at all spin-down timescales considered there is a significant timescale separation between the slowly-evolving stationary and wave-4 modes and the leading synoptic-scale modes, 3-variable reduced models (from the nearly $10,000$-variable KRG05 model) in these two modes were derived. Since the 3-variable bare truncation models (with the effects of the unresolved modes simply ignored) generate either fixed point solutions (at $k^{-1} = 2.3$ days) or limit cycles, parameterising the climate-weather interactions is necessary in deriving accurate reduced models. One-variable models in the stationary mode alone were also examined.

With theoretically-justified corrections (under the MTV assumptions, which include arbitrarily large scale separation between climate and weather modes) to the bare truncation forcing, the untuned 3-variable MTV models generated reasonably good approximations of the stationary mode statistics and poor approximations of the wave-4 statistics. In particular, the strong, low-frequency oscillation in wave-4 was not captured by the untuned MTV models. To improve the simulations, the MTV tuning procedure was applied, and in most cases only one parameter - either the scaling of the bare truncation forcing or that of the effective climate feedback - was adjusted. However, it was found that different parameter settings optimised the simulation of the stationary mode and the simulation of wave-4. For example, the 3-variable model at $k^{-1} = 6.7$ days with an upscaled effective climate feedback generated a stronger skewness in the stationary mode reflecting a jet bimodality (although a weaker jet bimodality than in the unreduced model), but only with a downscaled effective climate feedback were the wave-4 statistics improved (and oscillations generated). In the 1-variable case, the statistics were reasonably well approximated by the untuned MTV models at spin-down timescales below the bifurcation point, but in the region

of jet bimodality, only the stationary mode variance was well captured. Again adjusting only one parameter, the first two moments and the ACF were well approximated; however, jet bimodality was not generated with this tuning. With an additional tuning parameter, the jet bimodality at $k^{-1} = 6.7$ days was induced in both the 1-variable and 3-variable models, but in the latter case, the simulation of wave-4 did not improve.

Rescaling the bare truncation forcing or effective climate feedback partly corrects for the absence of a large scale separation between fast and slow variables. Although the results were often optimised with this one-parameter tuning (optimised from among several multi-parameter tuning combinations), the tuned 3-variable models in particular could not capture the statistics of both modes. Generally, a first attempt at tuning the MTV equations would involve adjusting this one parameter, rescaling the nonlinearity in the noise-induced drift and multiplicative noise through the parameter $\lambda_M = \lambda_L \equiv \theta$ and adjusting the parameter defined in this dissertation (λ_C). The former two parameters were adjusted with some success here and in FMV05 and FM06. If these tunings fail, then the next step would be to apply a systematic fitting procedure such as Kalman filtering. As applied in Strounine et al. (submitted), Kalman filtering served not only as a systematic method of tuning but also as an efficient means by which to compute the MTV coefficients. However, this tuning may significantly alter the reduced model dynamics.

The MTV approximation makes a number of stringent assumptions regarding the dynamics, not all of which will generally be true in climate-weather system. In particular, the scaling assumed by MTV implies that the invariant distribution of the weather modes is (asymptotically) independent of the climate modes. While this scaling may be appropriate for some problems, it is not for others. As an example, in the coupled atmosphere/ocean boundary layer dynamics considered by Monahan and Culina (in preparation), the (fast) surface wind speed PDF is strongly influenced by the atmospheric boundary layer thickness, which itself is a function of the (slow) surface stratification. One possible reason for the superior performance of the Hasselmann approximations to the MTV approximations in the present study is a potential dependence of the invariant measure of the fast modes on the slow modes. Unfortunately, this dependence is difficult to establish numerically (in

particular given the very high dimensionality of the weather modes).

Jet bimodality was generated by the untuned 1-variable Hasselmann additive noise (W) model at $k^{-1} = 6.7$ days. Based on the success of this additive noise model and the nature of the potential function of the 1-variable deterministic averaged forcing, it is evident that the dynamics underlying jet bimodality is generated by the averaged forcing; i.e., by the stationary mode tendency forcing (of the unreduced model) averaged with respect to the distribution of all other modes conditioned on the state of the stationary mode. The effective independence of the diffusion function σ on the climate variable does *not* imply that the climate-weather feedback is unimportant to jet bimodality and hence that the stationary mode self-interaction is responsible for jet bimodality. Rather, deterministic averaging induces nonlinear functions in the stationary mode besides the stationary mode self-interaction, which represent the nonlinear relationship between the stationary, wave-4 and wave-5 modes that is responsible for jet bimodality. Additive noise is the source of variance or energy exciting the reduced system to sample the full range of these interactions. In contrast to the paradigm put forth in Sura et al. (2005), deterministic nonlinearity and not multiplicative noise is responsible for the non-Gaussianity of the KRG05 model. As the present model is a channel model with idealised dynamics and Sura et al. (2005) considered real atmospheric data, the results here do not refute those in Sura et al. (2005); rather, the present results suggest that those of the earlier study are not as general as had been suggested.

It might be expected that with an additional resolved mode - wave-4 - and with a greater scale separation between the stationary and wave-4 modes and the weather modes that the 3-variable model would outperform the 1-variable model in capturing jet bimodality. The untuned 3-variable model performed very well in simulating the wave-4 statistics, and it did in fact generate jet bimodality, but the regime behaviour was weaker than that produced by the unreduced and 1-variable models. The (W*) model, with a diffusion matrix determined from the average over climate realisations of the integrated lag-covariance associated with the untuned (N) model, yielded an excellent approximation of the stationary mode statistics, including the pronounced skewness of the PDF. However, the wave-4

lag-correlation timescale was significantly less than that of the unreduced model. The (N) model with upscaled noise generated the correct variance and skewness, but with a climate drift in addition to an underestimation of the wave-4 lag-correlation timescale. Similar to the MTV models, only one of the two modes could be simulated with a high degree of accuracy. However, the 3-variable Hasselmann models do clearly outperform their MTV counterparts in simulating *both* modes, at all spin-down timescales considered.

The scale separation parameter ϵ appears explicitly in the (N) model as a theoretically justified scaling of the noise term. Due to uncertainty in its calculation and the absence of a large scale separation between climate and weather modes, it is natural to adjust this parameter. In the KRG05 model, this tuning can also be crucial to the accuracy of the approximations. For the 1-variable (N) model at $k^{-1} = 12.0$ days, the timescale between excursions to the low-latitude regime decreased from decades to centuries by decreasing the strength of the noise term by only 15%. In the 3-variable case at $k^{-1} = 4.7$, 6.7 and 12.0 days, seemingly accurate estimates of ϵ were obtained, as at these spin-down timescales there is a better-defined scale separation between the slow and fast modes. Nevertheless, at $k^{-1} = 4.7$ and 12.0 days, the solutions were unbounded with the a priori determined values of ϵ. With noise downscaled by decreasing ϵ, the solutions were bounded, and at the latter spin-down timescale in particular, a very good approximation of both modes was obtained. A tuning of the noise term could not bridge the differences between the reduced and unreduced models in all cases, as in the 3-variable case at $k^{-1} = 6.7$ days. It is not obvious how to introduce additional tuning parameters, as the (N) model in particular is comprised of only two forcings. By comparison, the MTV model is amenable to multi-parameter tuning as the individual forcings comprising the reduced model may each be tuned (as in Strounine et al. (submitted)). Although not beneficial towards understanding the underlying physics, this feature of the MTV model is useful towards accurate parameterisation of the unresolved weather features.

However, the hybrid Hasselmann-unreduced models can also be tuned by adjusting the number of spin-up steps N_1 and averaging ensemble members R. As the solution of a hybrid model lies somewhere in between the solutions of the Hasselmann and unreduced

models, a decrease in R brings this solution closer to the unreduced model solution. Despite the significant differences between the solutions of the (A) and unreduced models, a good approximation of the statistics of the stationary and wave-4 modes of the unreduced model was (efficiently) generated by the hybrid model with appropriately chosen values of N_1 and R. It is very encouraging (for the potential of the hybrid model to serve as an efficient parameterisation scheme) that a good approximation of the climate component of the KRG05 model was efficiently obtained with a relatively small scale separation and with significant differences between the solutions of the (A) and unreduced models.

This dissertation presents the first application of Hasselmann's method to a deterministic (but turbulent) climate model of high dimensionality. The primary goals of this study were to study the separation of such a model into fast and slow variables, to develop efficient algorithms for carrying out the averaging, and to study the hierarchy of reduced models. A more detailed analysis would entail a thorough test of the assumptions underlying the limiting theories which gives rise to the reduced equations. As discussed, the assumptions of mixing and ergodicity of the weather dynamics are not practical to evaluate in complex climate models. It is also an assumption of the limiting theories that the limit exists for arbitrarily large scale separation, and that as the limit is approached, the slow variables become arbitrarily close to the solutions of the averaged and stochastic equations. In fact, this limit may not exist in the neighbourhood of a bifurcation point.

Naturally, a more thorough examination of the reduced dynamics would entail integration of the full suite of reduced models - the (A), (L), (W), (N) and (N+) models. These models form a hierarchy from the least accurate (A) model to the (in principle) most accurate (N+) model, which best accounts for the effect of the weather modes on the climate. As the (W) model is not rigorously defined, alternative formulations of the diffusion matrix can be considered. Given the success of the 3-variable (W*) model at $k^{-1} = 6.7$ days in capturing jet bimodality, it is an interesting direction for future work to see under which circumstances the (W*) model or a similar such additive noise model outperforms the theoretically-justified multiplicative noise (N) model.

The (N+) model is particularly important. Although the stochastic Hasselmann models

generated reasonably good approximations of the KRG05 model climate, it is possible that the noise-induced drift correction would improve the results. Expressed in Itô form, the (N+) SDE involves derivatives of the weather modes, which are unmanageable in high-dimensional systems. However, the time evolution of this derivative can be determined, so in fact only the initial condition of this quantity needs to be specified (cf. Chapter 2; 2.22). With an appropriate re-initialisation scheme, this derivative need only be evaluated once at the beginning of the effective climate integration. Alternatively, it may be possible to integrate the (N+) SDE using numerical methods from Stratonovich calculus (e.g., Ewald and Penland (2009)).

Klaus Hasselmann in Hasselmann (1976) was the first to show that the slowly-evolving components of the climate system under the influence of fast-evolving turbulent processes can be modelled as stochastic processes in a systematic fashion. This program was not carried out at the time because of the huge computational requirements. Subsequent stochastic climate modelling studies utilized SDEs of a simpler form, that were derived either empirically (e.g., the linear inverse modelling of Penland and Sardeshmukh (1995)) or in an ad hoc fashion. Modern computational resources have allowed us to directly apply rigorously justified limiting theories to model high-dimensional climate systems (but requiring the development of new algorithms). This dissertation is one of the first studies to do this, and is the first using the limiting theories of Hasselmann's method. The success in this dissertation in applying this method and the potential for application to models of greater complexity hold promise for systematic approaches to stochastic climate modelling in future studies.

Bibliography

Arnold, L., 1974: *Stochastic Differential Equations: Theory and Applications*. Wiley, New York, 228 pp.

Arnold, L., P. Imkeller, and Y. Wu, 2003: Reduction of deterministic coupled atmosphere-ocean models to stochastic ocean models: a numerical case study of the Lorenz-Maas system. *Dynamical Systems*, **18**, 295–350.

Benzi, R., G. Parisi, A. Sutera, and A. Vulpiani, 1982: Stochastic resonance in climatic change. *Tellus*, **24**, 10–16.

Benzi, R., A. Sutera, and A. Vulpiani, 1981: The mechanism of stochastic resonance. *J. Phys. A: Math. Gen.*, **14**, L453–L457.

Berner, J., 2005: Linking nonlinearity and non-gaussianity of planetary wave behaviour by the fokker-planck equation. *J. Atmos. Sci.*, **62**, 2098–2117.

Billingsley, P., 1968: *Convergence of probability measures*. Wiley, New York, 253 pp.

Bogolyubov, N. and Y. Mitroplskii, 1961: *Asymptotic methods in the theory of nonlinear oscillations*. Gordon & Breach, New York.

Branstator, G., 1992: The maintenance of low-frequency atmospheric anomalies. *J. Atmos. Sci.*, **49**, 1924–1945.

— 1995: Organization of storm track anomalies by recurring low-frequency circulation anomalies. *J. Atmos. Sci.*, **52**, 207–226.

Charney, J. G. and J. G. DeVore, 1979: Multiple flow equilibria in the atmosphere and blocking. *J. Atmos. Sci.*, **36**, 1205–1216.

Crommelin, D., 2003: Regime transitions and heteroclinic connections in a barotropic atmospher. *J. Atmos. Sci.*, **60**, 229–246.

— 2004: Observed nondiffusive dynamics in the large-scale atmospheric flow. *J. Atmos. Sci.*, **61**, 2384–2396.

Crommelin, D. and A. Majda, 2004: Strategies for model reduction: Comparing different optimal bases. *J. Atmos. Sci.*, **61**, 2206–2217.

Crommelin, D. and E. Vanden-Eijnden, 2006: Reconstruction of diffusions using spectral data from time-series. *Comm. Math. Sci.*, **4**, 651–668.

D'Andrea, F. and R. Vautard, 2001: Extracting macroscopic dynamics: model problems and algorithms. *Quart. J. Roy. Met. Soc.*, **127**, 1357–1374.

Egger, J. and H. D. Schilling, 1983: On the theory of the long-term variability of the atmosphere. *J. Atmos. Sci.*, **40**, 1073–1085.

Ehrendorfer, M., 2000: The total energy norm in a quasigeostrophic model. *J. Atmos. Sci.*, **57**, 3443–3451.

Ewald, B. and C. Penland, 2009: Numerical generation of stochastic differential equations in climate models. *Computational methods for the Atmosphere and the Oceans*, North-Holland/Elsevier.

Farrell, B. F. and P. J. Ioannou, 1995: Stochastic dynamics of the mid-latitude atmospheric jet. *J. Atmos. Sci.*, **52**, 1642–1656.

— 1996: Generalized stability theory. part I: autonomous operators. *J. Atmos. Sci.*, **53**, 2025–2040.

Fatkullin, I. and E. Vanden-Eijnden, 2004: A computational strategy for multiscale systems with applications to lorenz 96 model. *J. Comp. Phys.*, 605–638.

Feldstein, S. and S. Lee, 2002: Is the atmospheric zonal index driven by an eddy feedback? *J. Atmos. Sci.*, **16**, 1212–1227.

Franzke, C., D. Crommelin, A. Fisher, and A. Majda, 2008: Hidden markov model perspective on regimes and metastability. *J. Clim.*, **21**, 1740–1757.

Franzke, C. and A. Majda, 2006: Low-order stochastic mode reduction for a prototype atmospheric gcm. *J. Atm. Sci.*, **63**, 457–479.

Franzke, C., A. Majda, and E. Vanden-Eijnden, 2005: Low-order stochastic mode reduction for a realistic barotropic model climate. *J. Atm. Sci.*, **62**, 1722–1745.

Ganopolski, A. and S. Rahmstorf, 2002: Abrupt glacial climate changes due to stochastic resonance. *Phys. Rev. Letters*, **88**, 153–158.

Gardiner, C., 1996: *Handbook of Stochastic Methods*. Springer, Berlin, 442 pp.

Givon, D. and R. Kupferman, 2004: White noise limits for discrete dynamical systems driven by fast deterministic. *Phys. A - Stat. Mech. and its Apps.*, **335**, 385–412.

Givon, D., R. Kupferman, and A. Stuart, 2004: Extracting macroscopic dynamics: model problems and algorithms. *Nonlinearity*, **17**, R55–R127.

Grabowski, W., 2001: Coupling cloud processes with the large-scale dynamics using the cloud-resolving convection parameterization. *J. Atmos. Sci.*, **58**, 978–97.

Haines, K., 1994: Low-frequency variability in atmospheric middle latitudes. *Surveys in Geophys.*, **15**, 1–61.

Haines, K. and A. Holland, 1998: Vacillation cycles and blocking in a channel. *Q. J. R. Meteorol. Soc.*, **124**, 873–895.

Hasselmann, K., 1976: Stochastic climate models. *Tellus*, **28**, 473–484.

Just, W., H. Kantz, C. Rödenbeck, and M. Helm, 2001: Stochastic modelling: replacing fast degrees of freedom by noise. *J. Phys. A: Math. Gen.*, **34**, 3199–3213.

Khasminskii, R. Z., 1966: On stochastic processes defined by differential equations with a small parameter. *Theory Probab. Appl.*, **11**, 211–28.

Kifer, Y., 2001: Averaging and climate models. *Stochastic Climate Models*, Birkhäuser.

— 2003: l^2 diffusion approximation for slow motion in averaging. *Stoch. and Dyn.*, **3**, 213–246.

— 2004a: Averaging principle for fully coupled dynamical systems and large deviations. *Erg. Theory and Dyn. Sys.*, **24**, 841–871.

— 2004b: Some recent advances in averaging. *Modern Dynamical Systems and Applications*, Cambridge University Press.

— 2005: Another proof of the averaging principle for fully coupled dynamical systems with hyperbolic fast motions. *Disc. and Cont. Dynam. Systems*, **13**, 1187–1201.

Kimoto, M. and M. Ghil, 1993: Multiple flow regimes in the northern hemisphere winter. part I: Methodology and hemispheric regimes. *J. Atmos. Sci.*, **50**, 2625–2644.

Kondrashov, D., S. Kravtsov, and M. Ghil, 2006: Empirical mode reduction in a model of extratropical low-frequency variability. *J. Atmos. Sci.*, **63**, 1859–1877.

Kravtsov, S., P. Berloff, W. K. Dewar, J. C. McWilliams, and M. Ghil, 2006: Dynamical origin of low-frequency variability in a highly nonlinear mid-latitude coupled model. *J. Clim.*, **19**, 6391–6408.

Kravtsov, S., W. K. Dewar, P. Berloff, J. C. McWilliams, and M. Ghil, 2007: A highly nonlinear coupled model of decadal variability in a mid-latitude ocean-atmosphere model. *Dyn. Atmos. Oceans*, **19**, 6391–6408.

Kravtsov, S., W. K. Dewar, M. Ghil, J. C. McWilliams, and P. Berloff, 2008: A highly nonlinear coupled model of decadal variability in a mid-latitude ocean-atmosphere model. *Physica D*, **237**, 584–599.

Kravtsov, S., A. W. Robertson, and M. Ghil, 2003: Low-frequency variability in a baroclinic β channel model with land-sea contrast. *J. Atmos. Sci.*, **60**, 2267–2293.

— 2005: Bimodal behaviour in the zonal mean flow of a baroclinic β-channel model. *J. Atmos. Sci.*, **62**, 1746–1769.

Kurtz, T., 1973: A limit theorem for perturbed operator semigroups with applications to random evolutions. *J. Funct. Anal.*, **12**, 55–67.

Kwasniok, F., 1996: The reduction of complex dynamical systems using principal interaction patterns. *Physica D*, **92**, 28–60.

Legras, B. and M. Ghil, 1985: Persistent anomalies, blocking and variations in atmospheric predictability. *J. Atmos. Sci.*, **42**, 433–471.

Lorenz, D. J. and D. L. Hartmann, 2001: Eddy-zonal flow feedback in the southern hemisphere. *J. Atmos. Sci.*, **26**, 3312–3327.

— 2002: Eddy-zonal flow feedback in the northern hemisphere winter. *J. Atmos. Sci.*, **16**, 1212–1227.

Majda, A. J., I. Timofeyev, and E. Vanden-Eijnden, 2001: A mathematical framework for stochastic climate models. *Comm. Pure Appl. Math.*, **54**, 891–974.

— 2002: A priori tests of stochastic mode reduction strategy. *Physica D*, **170**, 206–252.

— 2003: Systematic strategies for stochastic mode reduction in climate. *J. Atmos. Sci.*, **60**, 1705–1722.

— 2006: Stochastic models for selected slow variables in large deterministic systems. *Nonlinearity*, **19**, 769–794.

Marshall, J. and F. Molteni, 1993: Toward a dynamical understanding of planetary-scale flow regimes. *J. Atmos. Sci.*, **50**, 1792–1818.

McWilliams, J., 1977: A note on a consistent quasi-geostrophic model in a multiply connected domain. *Dynamics of Atmospheres and Oceans*, **1**, 422–441.

Monahan, A. H. and J. Culina, in preparation: Stochastic averaging of simple climate models.

Monahan, A. H., J. C. Fyfe, and L. Pandolfo, 2003: The vertical structure of wintertime climate regimes of the northern hemisphere extratropical atmosphere. *J. Climate*, **16**, 2005–2021.

Pandolfo, L., 1993: Observational aspects of the low-frequency intraseasonal variability of the atmosphere in middle latitudes. *Advances in Geophysics*, Academic Press, volume 34, 93–174.

Penland, C., 2003: A stochastic approach to nonlinear dynamics: A review. *Bull. Am. Met. Soc.*, **84**, ES43–ES52.

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

Quadrelli, R. and J. M. Wallace, 2004: A simplified linear framework for interpreting patterns of northern hemisphere wintertime climate variablity. *J. Clim.*, **17**, 3728–3744.

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

Shutts, G. J., 1983: The propogation of eddies in diffluent jet streams: Eddy vorticity forcing of blocking flow fields. *Q. J. R. Meteor. Soc*, **109**, 737–761.

Strounine, K., S. Kravtsov, D. Kondrashov, and M. Ghil, submitted: Reduced models of atmospheric low-frequency variability: Parameter estimation and comparative performance. *Physica D*.

Sura, P. and M. Newman, 2008: The impact of rapid wind variability upon air-sea thermal coupling. *J. Phys. Oceanogr.*, **38**, 639–647.

Sura, P., M. Newman, and M. Alexander, 2006: Daily to decadal sea surface temperature variability driven by state-dependent stochastic heat fluxes. *J. Phys. Oceanogr.*, **36**, 1940–1958.

Sura, P., M. Newman, C. Penland, and P. Sardeshmukh, 2005: Multiplicative noise and non-gaussianity: A paradigm for atmospheric regimes? *J. Atm. Sci.*, **62**, 1391–1409.

Sura, P. and P. Sardeshmukh, 2008: A global view of non-gaussian sst variability. *J. Phys. Oceanogr.*, **38**, 639–647.

Thompson, D. and J. Wallace, 2001: Regional climate impacts of the northern hemisphere annular mode. *Science*, **293**, 85–89.

Thompson, D. W. and J. M. Wallace, 2000: Annular modes in the extratropical circulation. part I: Month-to-month variability. *J. Climate*, **13**, 1000–1016.

Vautard, R. and B. Legras, 1988: On the source of mid-latitude low-frequency variability. part II: Nonlinear baroclinic equilibration of weather regimes. *J. Atmos. Sci.*, **45**, 2845–2867.

Wallace, J. M. and D. S. Gutzler, 1981: Teleconnections in the geopotential height field during the northern hemisphere winter. *Mon. Weath. Rev.*, **109**, 784–812.

Walters, P., 1982: *An Introduction to Ergodic Theory*. Springer-Verlag, New York, 250 pp.

Weinan, E., D. Liu, and E. Vanden-Eijnden, 2005: Analysis of multiscale methods for stochastic differential equations. *Comm. on Pure and App. Math.*, **58**, 1544–1585.

Whitaker, J. S. and P. D. Sardeshmukh, 1998: A linear theory of extratropical synoptic eddy statistics. *J. Atmos. Sci.*, **55**, 237–258.

Winkler, C. R., M. Newman, and P. D. Sardeshmukh, 2001: A linear of wintertime low-frequency variablity. part I: Formulation and forecast skill. *J. Clim.*, **14**, 4474–4494.