

APPLICATION OF EMPIRICAL AND DYNAMICAL CLOSURE
METHODS TO SIMPLE CLIMATE MODELS

Lauren Elizabeth Padilla

A DISSERTATION
PRESENTED TO THE FACULTY
OF PRINCETON UNIVERSITY
IN CANDIDACY FOR THE DEGREE
OF DOCTOR OF PHILOSOPHY

RECOMMENDED FOR ACCEPTANCE
BY THE DEPARTMENT OF
MECHANICAL AND AEROSPACE ENGINEERING
Advisers: Geoffrey K. Vallis and Clarence W. Rowley

January 2013

© Copyright by Lauren Elizabeth Padilla, 2012. All Rights Reserved.

Abstract

This dissertation applies empirically- and physically-based methods for closure of uncertain parameters and processes to three model systems that lie on the simple end of climate model complexity. Each model isolates one of three sources of closure uncertainty: uncertain observational data, large dimension, and wide ranging length scales. They serve as efficient test systems toward extension of the methods to more realistic climate models.

The empirical approach uses the Unscented Kalman Filter (UKF) to estimate the transient climate sensitivity (TCS) parameter in a globally-averaged energy balance model. Uncertainty in climate forcing and historical temperature make TCS difficult to determine. A range of probabilistic estimates of TCS computed for various assumptions about past forcing and natural variability corroborate ranges reported in the IPCC AR4 found by different means. Also computed are estimates of how quickly uncertainty in TCS may be expected to diminish in the future as additional observations become available.

For higher system dimensions the UKF approach may become prohibitively expensive. A modified UKF algorithm is developed in which the error covariance is represented by a reduced-rank approximation, substantially reducing the number of model evaluations required to provide probability densities for unknown parameters. The method estimates the state and parameters of an abstract atmospheric model, known as Lorenz 96, with accuracy close to that of a full-order UKF for 30-60% rank reduction.

The physical approach to closure uses the Multiscale Modeling Framework (MMF) to demonstrate closure of small-scale, nonlinear processes that would not be resolved directly in climate models. A one-dimensional, abstract test model with a broad spatial spectrum is developed. The test model couples the Kuramoto-Sivashinsky equation to a transport equation that includes cloud formation and precipitation-like processes.

In the test model, three main sources of MMF error are evaluated independently. Loss of nonlinear multi-scale interactions and periodic boundary conditions in closure models were dominant sources of error. Using a reduced order modeling approach to maximize energy content allowed reduction of the closure model dimension up to 75% without loss in accuracy. MMF and a comparable alternative model performed equally well compared to direct numerical simulation.

Acknowledgements

This dissertation would not have been possible without the support of many to whom I am indebted and grateful. I thank my advisors Geoffrey Vallis and Clarence Rowley for their wisdom, support, and open-mindedness throughout my time at Princeton. They encouraged me to pursue my interests at the intersection of engineering and climate science by guiding me through their respective fields which resulted in a fruitful graduate experience for me.

Many thanks go out to Professor Luigi Martinelli for his mentorship and helpful scientific conversation at almost a dozen quarterly meetings of the Joint University Program. I am also grateful to Rob Socolow for encouraging my interdisciplinary interests through involvement in the Princeton Energy and Climate Scholars. I appreciate all of the many excellent teachers I have learned from including my engineering professors and the scientists at the Geophysical Fluid Dynamics Laboratory, especially Isaac Held and Michael Winton for many conversations with respect to the work on climate sensitivity.

Thank you to the readers of this dissertation, Yannis Kevrekidis and Geoffrey Vallis, for their helpful comments. And thank you to my final public oral examiners, Isaac Held and Marcus Hultmark.

I am grateful to have been supported by the dedicated staff of several great Princeton programs. In the Program in Atmospheric and Oceanic Sciences, I especially thank Anna Valerio, Laura Rossi, Joanne Curcio, and Cindy Kandell. In the Department of Mechanical and Aerospace Engineering I was fortunate to have the support of Jessica O’Leary, Jill Ray, Deborah Brown, Maureen Hickey, Marcia Kuonen, Kathy Opitz, Candy Reed, and Louis Riehl. Thank you to Judy Swan from the Writing Center for whipping my scientific writing skills into shape. Thank you to Amy Pszczolkowski from career services for helping me think about life after the dissertation.

I appreciate the comraderie of my officemates, Amanda O’Rourke, Sam Potter, Steve Brunton, Brandt Belson, Jonathan Tu, Kevin Chen, Imène Goumiri, Carla Bahri, Anthony DeGennaro, Scott Dawson and Mark Luchtenburg. I also thank the members past and present of the Princeton Energy and Climate Scholars for helping me think about my research in a broad context. To my friends and fellow students, especially, Elena Krieger, Darren Pais, Arijeet Pal, Jitendra Kanodia, David Turnbull, Jessica Shang, Tristen Hohman, Andy Stewart, Megan Leftwich, and many others,

it was a pleasure spending time with you in and out of academic settings.

Thank you to my family for their loving support, especially to both sets of my parents for their encouragement from the first and third coasts. To my mom and dad, I am here today thanks to your unflagging emphasis on education over three decades. With all my heart I thank my husband Todd, who has steadfastly walked every step of this endeavor with me, for your constant assurance, friendship, and love.

A final word of thanks goes to all of those who helped me in any capacity while I completed this work.

This work was supported by grants from the Federal Aviation Administration Joint University Program and the Cooperative Institute for Climate Science.

This dissertation carries the number T-3251 in the records of the Department of Mechanical and Aerospace Engineering.

To Todd,
love and home.

Contents

Abstract	iii
Acknowledgements	iv
List of Tables	x
List of Figures	xi
List of Algorithms	xiv
1 Introduction	1
1.1 Motivation and goals	3
1.2 Approach	4
1.3 Overview and contributions	5
2 Empirical closure for a globally-averaged energy balance model	9
2.1 Energy balance models of the climate	10
2.1.1 Two-box model	10
2.1.2 Mixed-layer model	11
2.2 Equilibrium climate sensitivity, transient climate sensitivity, and transient climate response	12
2.2.1 Probabilistic estimates of climate change	15
2.3 Parameter estimation method	16
2.3.1 Unscented Kalman filter approach	17
2.3.2 Scaled-unscented sigma points	19
2.4 Validation of the method with data from comprehensive climate model	22
2.4.1 Comprehensive climate model data	22

2.4.2	Parameter estimates	23
2.4.3	Effect of assumptions on forcing uncertainty	25
2.5	Concluding remarks	27
3	Closure estimates from climate observations	28
3.1	Observational data	29
3.2	Parameter estimates	31
3.3	Sensitivity to temperature and forcing	33
3.3.1	Sensitivity to magnitude of uncertainty in forcing and temperature	34
3.3.2	Sensitivity to natural temperature fluctuations	37
3.3.3	Sensitivity to forcing trajectories	38
3.4	Sensitivity to mixed layer depth	39
3.5	Sensitivity to historical time-period	41
3.6	Sensitivity to prior uncertainty	43
3.7	Sensitivity to skew	45
3.8	Concluding remarks	47
4	Empirical closure for higher order climate systems	48
4.1	Adaptive-Covariance-Rank Unscented Kalman Filter	50
4.1.1	Adaptive covariance rank	51
4.1.2	Measurement update in factored form	52
4.2	An abstract atmospheric test model	56
4.3	State estimates	56
4.4	Parameter estimates	60
4.5	Limitations	62
4.6	Concluding remarks	62
5	Dynamical closure of a momentum-moisture-like model	64
5.1	Multiscale modeling framework overview	65
5.1.1	MMF equations and approximations	67
5.1.2	Criteria for successful test model	71

5.2	MMF test models	73
5.2.1	Kuramoto-Sivashinsky equation	73
5.2.2	KS-moisture system	76
5.2.3	Filtered moisture equation	80
5.2.4	Moisture closure equation	81
5.2.5	Momentum equation	82
5.2.6	Numerical simulation	83
5.3	Concluding remarks	84
6	Performance evaluation of dynamical closure	85
6.1	Neglecting filter-closure interactions	85
6.2	Periodic boundary error	89
6.2.1	Advection	89
6.2.2	Downscaling	90
6.2.3	Upscaling	94
6.3	Closure model dimension reduction	95
6.4	Coupling error	98
6.5	Concluding remarks	100
7	Conclusion	102
7.1	Future research directions	104
A	Filter Review	106

List of Tables

3.1	Natural variability and scale factor uncertainty for four experiments	34
6.1	Comparison of q' variance from DNS and two closure models	91
6.2	Upscaling simulation statistics	95
6.3	Comparison of q' spatial variance with decreasing resolution	97
6.4	Coupled MMF simulation statistics	100

List of Figures

1.1	Diagram of thesis organization	5
2.1	Comparison of EBM and GFDL CM2.1 model responses	12
2.2	Comparison of TCR and TCS	14
2.3	Map between probability densities in λ and TCS	15
2.4	GFDL's CM2.1 historical temperature realizations	22
2.5	GFDL's CM2.1 historical forcing estimates	23
2.6	Sensitivity parameter estimates from CM2.1 model data	24
2.7	Comparison of CM2.1 and mixed-layer model temperature response	24
2.8	Estimates of individual components of CM2.1 forcing	26
2.9	Comparison of parameter estimates for two different approaches to forcing uncertainty	27
3.1	Observed historical temperature record	29
3.2	Estimates of historical forcing	30
3.3	Estimates of transient sensitivity parameter and aerosol scale factor	32
3.4	Evolution of TCS probability density	33
3.5	Range of TCS probability densities	34
3.6	Variation of transient sensitivity parameter estimates with natural variability and forcing uncertainty for the period 1900 to 2008	35
3.7	Variation of transient sensitivity parameter estimates with natural variability and forcing uncertainty for the period 1970 to 2030	36
3.8	Experiments with alternative temperature and forcing time series	37
3.9	Sensitivity of parameter estimates to individual forcing trajectories	38

3.10	Sensitivity of parameter estimates to mixed layer depth	40
3.11	Variation of surface temperature response with mixed layer depth	40
3.12	Variation of TCS probability density with mixed layer depth	41
3.13	Estimates of transient sensitivity parameter and aerosol scale factor from 1970 onward	42
3.14	Sensitivity of climate sensitivity parameter to prior uncertainty	44
3.15	Sensitivity of aerosol scale factor to prior uncertainty	44
3.16	Impact of skew on parameter estimates	46
3.17	Impact of skew on TCS probability density over time	46
4.1	True and observed L96 fields	57
4.2	Comparison of full and adaptive UKF state estimates	58
4.3	Comparison of full and adaptive UKF percent error	59
4.4	Comparison of full and adaptive covariance rank	59
4.5	Comparison of full and adaptive UKF state estimate at latitude 20	60
4.6	Comparison of full and adaptive UKF parameter estimates	61
5.1	Coupling of multiscale domains	69
5.2	Eigenvalues of the linearized Kuramoto-Sivashinsky equation	75
5.3	Energy spectra for Kuramoto-Sivashinsky equation	76
5.4	Scale factor in moisture source term	78
5.5	Illustration of moisture source strength	80
6.1	Effect of advection approximations on \bar{q} realizations	86
6.2	Effect of advection approximations on spectrum of q	87
6.3	Effect of advection approximations on spectrum of $(uq)_x$	87
6.4	Downscaling realizations of q'	91
6.5	Comparison of energy spectra of q' in downscaling experiments	92
6.6	Comparison of spatial variance of q' over time in downscaling experiments	93
6.7	Spatial variance of \bar{q} and M over time in upscaling experiments	95
6.8	Comparison of q' realizations with decreasing resolution	96
6.9	Comparison of energy spectra of q' with decreasing resolution	97

6.10 Comparison of \bar{q} realizations in coupled MMF experiments	98
6.11 Comparison of spatial variance of \bar{q} in coupled MMF experiments	99

List of Algorithms

1	Unscented Kalman filter	21
2	Adaptive-covariance-rank unscented Kalman filter	55
3	Time stepping routine in MMF filtered-closure model simulations.	84

Chapter 1

Introduction

Climate models are systems of equations describing the behavior of Earth's land, ocean, and atmosphere and their interactions with each other as well as outer space, based on physical laws. Although they are deeply rooted in physical first principles, climate models include unknown parameters and processes that need closure through parameterization or other means. They will continue to need closure for the foreseeable future because of inherent sources of uncertainty and complexity in the climate system.

Physical intricacies of climate processes are often left out of global models for at least two main reasons. They are either on scales too small to be resolved by the model given available computational resources, for example as in boundary layer turbulence and cumulus convection. Or equations describing their behavior are poorly understood as is the case for aspects of ice dynamics, vegetation, and ecosystems. In these cases closures are necessary and often empirical closure models known as parameterizations fill the void.

Even for processes that are modeled explicitly, closure may be necessary because it may be impossible to know the value of pertinent parameters. For example, the surface albedo is an important parameter in determining the reflectivity of solar radiation however its precise value through history and across the globe is continually in flux due to natural or human-made shifts in geology and land use. In some climate models the strength of feedbacks are considered parameters and are similarly unknown, such as the sensitivity of the climate's temperature response to changes in radiative forcing.

Climate model closure is not a straightforward calibration or modeling procedure. Calibration

or estimation of parameters from observations is made difficult by uncertainty about the timing and magnitude of climate forcings like solar radiation, volcanic eruptions, and emissions of greenhouse gases. Additionally, past observations may be sparse and error-prone. For example, climate records with high spatial coverage tend to be short and the longer records tend to include relatively few spatial locations with little vertical resolution. Proxy records like paleo-climate data and satellite observations do not measure climate variables directly so they are subject to calibration error of their own. The size of the system and the wide range of scales involved limits our ability to use what we know physically to provide closure models. In the case of cloud micro-physics, bulk equations (e.g. Grabowski and Smolarkiewicz (1996)) are used in regional cloud resolving models to provide physical insight into small-scale convective processes. However macroscopic parameterizations used in global climate models are developed based on cloud resolving simulations at only a fraction of possible atmospheric conditions.

Despite these difficulties, climate models have been closed with varying degrees of success using approaches that may be lumped into two quite general categories. We call them the expert model approach and the systematic model approach.

Expert models use heuristic techniques to tune model parameters during model development to achieve output that is consistent with the observed world (e.g., Mauritsen et al. (2012)). The structure of a parameterization is based on an expert’s physical intuition and empirical experience with converting detailed physical equations into simpler macroscopic ones. Some concerns about this approach include that the collection of expert models around the world give quite different simulation results, especially with respect to climate change. (Randall et al. (2007); Kiehl (2007); Meinshausen et al. (2008)) In general, there is no clear systematic way to assess the uncertainty in the results. (Allen and Stainforth (2002); Lopez et al. (2006)) Furthermore, the models tend to maximize the detailed physical processes included, making them computationally expensive and limiting results to a few simulations of each model.

A more systematic approach involves applying empirical and physical closures in a mathematically consistent framework. Observations are used to identify a collection of best performing models out of a large ensemble of models with different parameter settings that adequately sample a range of feasible parameter values (e.g. Stainforth et al. (2005); Annan et al. (2005b)). A major benefit is that systematic evaluation of uncertainty may be conducted using the ensemble results.

Physically-based closure models are developed systematically using scale separation in such a way as to avoid introducing new parameters (e.g. Grabowski and Smolarkiewicz (1999); Majda (2007)). An advantage of parameterizations designed in this manner is that they alleviate some of the need for model tuning.

Systematic approaches to closure may help reduce biases and minimize the contribution closures make to the overall error in climate model output. They can also help improve understanding of the range of these errors. However, the use of systematic techniques has been limited mainly due to their greater computational burden. When they are applied, compared to expert models, some physical realism may need to be sacrificed to afford the many simulations needed for uncertainty analysis.

The complexity of systematic techniques for climate modeling makes an assessment of their benefits difficult or ambiguous, which further inhibits their widespread use. Therefore, to improve our understanding of the benefits and limitations, this thesis investigates the effectiveness of systematic closure techniques by applying them to the simplest possible climate models for which the success or failure of the technique is clear.

The main contribution of this thesis is the first application of three systematic closure techniques: the unscented Kalman filter, the adaptive-covariance-rank unscented Kalman filter, and the multiscale modeling framework to three simple climate models: a mixed-layer energy balance, the Lorenz 1996 abstract atmosphere, and a coupling of the Kuramoto-Sivashinsky equation to a scalar moisture-like transport equation. The main benefit of these applications is learning about the skill, appropriateness, and benefit of the methods for climate models in an experimental setting where sources of uncertainty and complexity may be controlled and restricted. An additional outcome is that through closure of the energy balance climate model, insight is achieved into predictions of future climate including how certain those predictions may be.

1.1 Motivation and goals

This work is motivated by the fact that closure parameters and closure models (which may also be known as parameterizations) are dominant sources of error in climate modeling that propagate through the system equations leading to uncertainty in their output. Systematic closure techniques

using concepts adapted from mathematical methods in engineering have not been widely applied to climate models primarily due to computational expense.

The goal of this thesis is to apply closure techniques that use mathematical tools from system identification and model reduction to systematically assess and reduce closure errors. The methods we use address, one at a time, three distinct issues that arise in climate models making closure difficult: unknown or limited observational data, high model dimensions, and extreme scope in time and length scales. In addition to the applications, we also seek to improve the methods and to develop and implement our revised approach.

1.2 Approach

There are several key aspects to our approach. First, the types of closures investigated occupy extreme ends of the modeling spectrum, they are either entirely observationally-based or entirely physically-based closures and we evaluate them independently. By observational or empirical closure we mean calibration or estimation of parameters based on past observations of the system. On the other hand, physical closure involves the development of a closure model based on known physics whose output supplies a system with its unknown terms. Although they are treated separately in this thesis, they are actually complementary techniques, rather than competing approaches to be compared, that could be used to treat different closure problems in the same model. We apply them to separate models for clearer understanding of their benefits and appropriateness to specific kinds of closure problems.

The mathematical methods employed in the techniques are drawn from the engineering fields of system identification and model reduction. Namely, this includes the use of Kalman filtering in the observationally-based closure methods and the use of modal decomposition, Galerkin projection, and dimension reduction in the physically-based closure methods.

We restrict applications to simple models that contain just a few unknown parameters with dimensions that easily fit within the computational constraints of an ordinary laptop meaning that many of the physical features of the climate system are neglected and spatial resolution is limited. The advantages of using simple models are multifold. For example, with fewer parameters to constrain, simple models have fewer sources of error and simpler closure problems for which it is

easier to implement closure methods. Simple climate models require less input data and produce less output data making pre- and post- processing less time consuming. Also their discretization for numerical integration is less complex and they require less computation time per simulation.

Finally, the approach of this thesis is interdisciplinary. The application of closure method to climate model is revealing from both the engineering perspective and the climate modeling perspective. The relevance of the simple models for climate is important, we are not interested in applying the methods to arbitrary dynamical systems. The simple climate applications highlight the strengths and weaknesses of the closure methods for addressing the key issues of uncertain observations, high dimension and wide ranging scales that are pertinent to more comprehensive climate models.

1.3 Overview and contributions

	Model/Method	Results	Closure Issue
Empirical Closure	Chapter 2 (EBM/UKF)	Chapter 3	Uncertain data
	Chapter 4 (L96/AUKF)	Chapter 4	High dimension
Physical Closure	Chapter 5 (KS-q/MMF)	Chapter 6	Wide range of scales

Figure 1.1: Diagram of thesis organization. Chapters in the left column describe theory behind the two kinds of closure methods, introduce the simple climate model test beds and provide other preliminaries. The chapters listed in the center column present the main results of applying the methods to the models. The right column is the focus of the application.

This thesis is organized according to the diagram of figure 1.1. The first half of this work, chapters 2–4, is dedicated to observationally-based or empirical closure methods. The second half, chapters 5–6, is focused on a physically-based method. The general progression is to introduce a model and method for its closure followed by simulation results demonstrating the performance. Each model and method demonstrates one of three main issues in climate closure.

The specific contributions of this thesis are the first application of three distinct closure methods to simple climate models, the development of a new method for efficient parameter estimation, and

the development of a new model as a better test bed for a multiscale closure technique. We begin by applying a filtering technique to determine the single unknown in a minimal climate system. The main result of this is a probabilistic estimate of the sensitivity parameter governing climatic change in the next few decades and a proof of concept that the filtering equations work well for a simple climate problem in which there is great uncertainty in the constraining data. However, the exact filter equations used in this case would be prohibitively expensive if paired with a more complex model. So the next area of work provides a simplification of the nonlinear filter equations for use with a higher-dimensional model. The approach is successful in our test model, however for application to comprehensive climate systems it requires some heuristic modifications. Up against the computational limit of systematic approaches leveraging observational data, the focus turns to a physically-based closure technique that has been used in comprehensive climate models to address the issue of resolving processes at extreme ends of the time and length spectrum, albeit with ambiguous success. This method of closure is evaluated on a test model designed specifically to alleviate much of this ambiguity.

Chapter 2 introduces an estimation technique based on the Kalman filter to close an energy balance climate model (EBM) with an unknown parameter pertaining to the transient climate sensitivity. The model itself is the subject of section 2.1. Key climate concepts are defined in section 2.2 including the notion of a transient sensitivity, which is a fairly new idea in the climate literature. The unscented Kalman filter (UKF) technique for parameter estimation is explained in section 2.3.1. The method is validated using data from a comprehensive climate model in section 2.4. Section 2.4.3 outlines a unique feature of the method that allows forcing estimates to be revised as observed data accumulates.

Chapter 3 provides detailed results of the parameter estimation. A key contribution of this chapter is the first use of the UKF to estimate a range of probability densities for the transient climate sensitivity (TCS). Section 3.1 outlines the observed climate records input to the UKF. The main probabilistic estimate of the TCS and its evolution over time is made in section 3.2. The results in this section also include the first computation of the potential to learn TCS given expected future data. The remainder of the chapter is focused on probing the sensitivity of the estimates to assumptions. In section 3.3, the filter is used to novelly compute a range of probabilistic estimates of the TCS based on direct variations of the model’s natural variability and forcing uncertainty. Most

of the work in chapters 2–3 has appeared in Padilla et al. (2011), a paper coauthored with Geoffrey Vallis and Clarence Rowley. This work was also presented at the 11th International Meeting on Statistical Climatology.

Chapter 4 presents a novel adaptation to the UKF that improves its efficiency for large systems and then applies the new filter to state and parameter estimation in an abstract atmospheric model (L96). Section 4.1 outlines the modified theory called adaptive-UKF (AUKF) which relies on the techniques of singular value decomposition and dimension reduction. Section 4.2 describes the L96 test model. The performance of the AUKF is demonstrated for state estimation in section 4.3 and for joint state-parameter estimation in section 4.4. Section 4.5 discusses some of the limitations of the approach for extremely large systems. The theory and the state estimation example have appeared in Padilla and Rowley (2010), coauthored with Clarence Rowley and with input from Geoffrey Vallis. The paper was presented at the 49th IEEE Conference on Decision and Control.

Chapter 5 reviews a closure technique known as the multiscale modeling framework (MMF) and introduces its application to a novel test system that couples the Kuramoto-Sivashinsky equation to an evolution equation for moisture-like processes. The main contribution here is the development of a simplified model for testing the MMF method in which sources of error may be controlled and investigated one at a time. Also, a flexible code for numerical integration of the model has been written that facilitates testing modified MMF implementations. An overview of MMF and criteria for successfully testing it are given in section 5.1. The test model and how it meets the criteria are described in section 5.2. The details of the closure problem in the large-scale test model are given in section 5.2.3. The closure model, whose output is provided to the large-scale model, is derived in section 5.2.4.

Chapter 6 provides detailed results of the MMF closure in the test model and evaluates the errors introduced with each successive MMF approximation. Section 6.1 covers the error due to neglect of certain non-linear interaction terms. Section 6.2 highlights the dominant source of error out of the entire approach, the imposition of periodic boundary conditions on the closure model. And section 6.3 shows how closure model performance degrades as its resolution decreases. Here the use of reduced order models via Galerkin projection and dimension reduction to improve the efficiency of the closure models is a novel contribution. Finally, the errors are coupled together in simulations of the complete MMF test system in section 6.4. An important contribution here is the comparison

of the performance of the model using MMF to a model of comparable computational expense. MMF adds considerable cost to simulation, however previous studies have only compared MMF results to those generated by models of lower cost. The test model and the results of its closure using MMF have been presented at several quarterly meetings of the Joint University Program of the Federal Aviation Administration.

Chapter 2

Empirical closure for a globally-averaged energy balance model

Analysis of empirical closure techniques begins on one of the simplest climate models to study; a linear representation of Earth's global energy exchange with space buffered by energy storage in its oceans. This coarse model of the global energy budget has several free parameters that need closure which pertain to rates of radiative transfer. The main reason to use a simple climate model in this case is to permit parameter closure using an entirely observationally-based method.

In this chapter, we introduce a canonical two-component energy balance model and explain several simplifying modifications made. We proceed to define the parameters of the model that need closure and explain their relevance to short-term prediction of climate warming under specific future emissions scenarios. Next we provide an overview of the theory behind an observationally-based closure technique based on the Kalman filter and give the specifics of implementation for our climate application. Finally, we validate this method using data obtained by simulations of a physically realistic general circulation model of the climate.

2.1 Energy balance models of the climate

Global energy balance models (EBMs) of the climate system first appeared in the literature in the work of Budyko (1969) and Sellers (1969). Since then there have been many incarnations of varying complexity. A zero-dimensional EBM suitable for our purpose of parameter estimation is similar to that used, for example, by Raper et al. (2002) and Held et al. (2010). In contrast to these models, however, we include stochastic forcing to represent the effects of natural variability. Our model, which characterizes the energy balance of the ocean’s mixed layer, is derived from a two-box (deep ocean, and mixed layer) model. We neglect the effects of heat storage in the deep ocean to reduce the number of unknown parameters. A consequence of this simplification is that parameter values and therefore future predictions of warming are valid only out to several decades.

2.1.1 Two-box model

Consider the following two-component model, equation (2.1), that may be the minimal system appropriate for studying future climate warming on century-long timescales. It contains two independent variables representing perturbation surface temperature (T) and deep ocean perturbation temperature (T_o), namely

$$\begin{aligned} C \frac{dT}{dt} &= F_{net} - \gamma T - \beta(T - T_o) \\ C_o \frac{dT_o}{dt} &= \beta(T - T_o) \end{aligned} \tag{2.1}$$

where γ and β are positive parameters, C and C_o are heat capacities of the mixed layer and deep ocean, respectively, with $C_o \gg C$, and F is the net perturbation to the climate forcing (including both natural and anthropogenic factors). In final equilibrium, $T = T_o$, and the temperature response to a specified forcing, say F_{2CO_2} , is given by $T_{ECS} = F_{2CO_2}/\gamma$.

When $F_{2CO_2} = 3.71 \text{ W/m}^2$, the change in radiative forcing under a doubling of CO_2 concentration over pre-industrial levels, then T_{ECS} is defined to be the Equilibrium Climate Sensitivity (Randall et al., 2007).

2.1.2 Mixed-layer model

On decadal timescales, the response of the deep ocean in the two-box model is small, so we approximate $T_o \approx 0$. The system then reduces to

$$C \frac{dT}{dt} = F_{net} - \lambda T, \quad (2.2)$$

an evolution equation for the mixed-layer response to external forcing, where $\lambda = \gamma + \beta$. Since the separate values of γ and β , and thus T_{ECS} , would be poorly constrained by observations, we limit our focus to determination of λ and transient climate responses.

We emphasize that the parameter λ determines the transient climate sensitivity, not the equilibrium climate sensitivity, because it includes the rate of heat uptake by the deep ocean as well as the outgoing infra-red radiation. Our λ is the same as the quantity ρ , termed the climate resistance by Gregory and Forster (2008), although our methods of finding it differ: we account for the time delay due to mixed layer heat capacity and therefore may make use of volcanic effects in constraining λ . The combination of γ and β is also similar to the sum of positive and negative feedbacks discussed in Baker and Roe (2009).

The parameter C represents the heat capacity of the system on decadal timescales, and we take its value to be that corresponding to a mixed layer of 60 m deep. Our results are fairly insensitive to this, and indeed a value of $C = 0$ does not give significantly different results.

We make one further modification, adding a stochastic forcing S , so that the model equation is

$$C \frac{dT}{dt} = F_{net} - \lambda T + S. \quad (2.3)$$

The term S parameterizes internally forced temperature variability. It satisfies the Ornstein–Uhlenbeck process (Majda et al., 2001; Vallis et al., 2004),

$$\frac{dS}{dt} = -\frac{S}{\tau} + \frac{\sigma_S \sqrt{2}}{\sqrt{\tau}} w_S \quad (2.4)$$

where τ determines the temporal correlation of the variability, σ_S is the standard deviation of the variability, and w_S is a white noise process.

Several realizations of the system (2.3)–(2.4) illustrate that it satisfactorily emulates the global-average response of a comprehensive climate model when given corresponding forcing and a nominal value of λ . These are plotted in figure 2.1. Each realization is initialized with a different random seed for w_S .

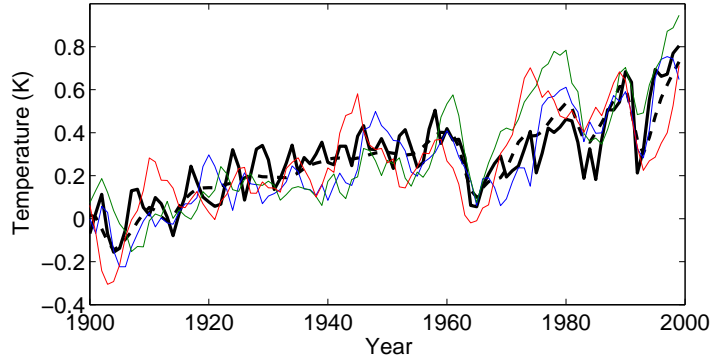


Figure 2.1: Three realizations of the stochastic mixed-layer model response (equations (2.3)–(2.4), thin color) compared to the mean of GFDL's CM2.1 model 20th century temperature response (solid). Also shown is the deterministic response of the simple model where the random variable S is fixed at zero (dashed).

2.2 Equilibrium climate sensitivity, transient climate sensitivity, and transient climate response

The model derived in the previous section is relevant to short term climate responses, on the order of several decades. Whereas much of previous work has focused on the steady-state response, which, given the heat capacity of the oceans, may take centuries to realize. This section provides definitions of established measures of climate responses, like the equilibrium response, and introduces a novel measure called the transient climate sensitivity (TCS). This section explains the relevance of TCS for near-term decision-making and how the sensitivity parameter of the model (2.3)–(2.4) may be used to obtain a concise mathematical expression for the TCS.

The equilibrium response of global-mean, near-surface temperature to an increase in greenhouse gas concentrations (e.g., a doubling of CO_2 levels) is given, definitionally, by the equilibrium climate sensitivity (ECS), which is considered an unambiguous and convenient measure of the sensitivity of the climate system to external forcing. However, given the long timescales involved in bringing the ocean to equilibrium the ECS may only be realized on a timescale of many centuries or more

and so its relevance to policy makers, and indeed to present society, has been debated. Of more relevance to the short and medium term — that is, timescales of a few years to about a century — is the transient climate response (TCR, Hegerl et al., 2007), which is the global and annual mean surface temperature response after about 70 years given a 1% CO₂ doubling rate. (Sometimes an average may be taken from 60 to 80 years or similar to ameliorate natural variability.)

Although the detailed response of the atmosphere to a doubling in CO₂ will depend on the rate at which CO₂ is added to the atmosphere, recent work with comprehensive models suggests that surface temperatures respond quite quickly to a change in radiative forcing, reaching a quasi-equilibrium on the timescale of a few years (in part determined by the mixed-layer depth) prior to a much slower evolution to the true equilibrium (e.g., Held et al., 2010). In the quasi-equilibrium state, the rate of change of surface temperature is a small fraction of its initial increase, and the response following a doubling of CO₂ may be denoted the transient climate sensitivity (TCS). The TCS may be expected to be very similar to the TCR, but its definition does not depend so strictly on there being a particular rate of increase of greenhouse gases. As long as the CO₂ doubles over a time period short enough for deep ocean temperature to remain far from equilibrium (less than 100 years, for example), the response to that doubling will likely be nearly independent of the emissions path.

Because the response is relatively rapid, we may use the transient sensitivity parameter in (2.2) to define the transient climate sensitivity in the following way,

$$T_{\text{TCS}} \equiv \frac{F_{2\text{CO}_2}}{\lambda}. \quad (2.5)$$

Here $F_{2\text{CO}_2}$ is the forcing corresponding to doubled CO₂, and this is approximately equal to the quasi-equilibrium response to a forcing change in the time dependent system (2.1).

Since the TCS is defined in terms of a single parameter, λ , it may be easier to determine from observations than the ECS. By summing the atmospheric feedback strength and the rate of ocean heat uptake [also an uncertain quantity (Hegerl et al., 2007; Forest et al., 2002)], these quantities do not need to be constrained separately. The overall response uncertainty, however, may still be dominated more by uncertainty in atmospheric feedbacks than the uptake of heat by the ocean (Knutti and Tomassini, 2008; Baker and Roe, 2009).

To illustrate the small differences between TCR and TCS, we simulate the response of the two-time-constant model (equation (2.1)) for three different emissions pathways to double CO_2 concentration: an instantaneous doubling, increases at a rate of 1% per year, and increases at a rate of 0.7% per year which ensures doubling is achieved in 100 years. The surface and deep ocean temperature responses to each of the forcing scenarios are plotted in figure 2.2. Key parameter values of the model are as follows: $\text{ECS} = 3.5 \text{ K}$, $\lambda = 2 \text{ W m}^{-2} \text{ K}^{-1}$, deep ocean depth of 5000 m, and mixed layer depth of 60 m.

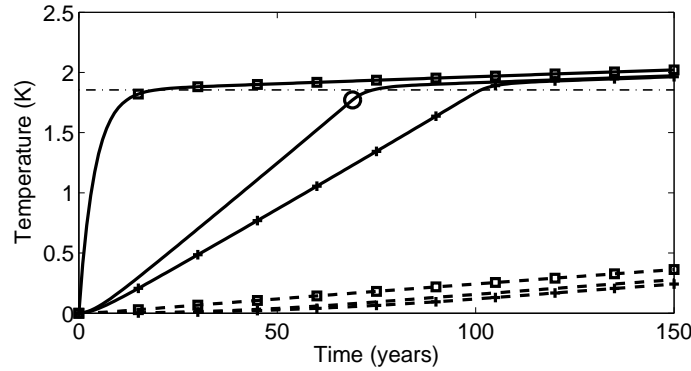


Figure 2.2: Surface (solid) and deep ocean (dashed) temperature responses to three different emissions pathways to double CO_2 : instantaneous (squares), 1% increases per year compounded continuously (no marker, with TCR marked as open circle), and 0.7% increases per year such that doubling is achieved in exactly 100 years (crosses). TCS as computed by equation (2.5) is the horizontal line (dash-dot).

In the first 150 years of the responses shown, the deep ocean temperature changes remain quite small, increasing to 10% of the ECS or less. The surface temperature reaches quasi-equilibrium, after instantaneous double forcing, in just 20 years. At $T = 1.86 \text{ K}$, which is the TCS computed according to (2.5), the surface temperature, is about 50% of the way to ECS. It will not reach even 75% of equilibrium for another 800 years. For each of the three emissions pathways, the surface temperature change at the end of the fast transient is almost identical due to the slow deep ocean temperature response. This would be the case for any other emissions pathway, as long as CO_2 doubling is achieved well before the ocean reaches equilibrium.

The TCR in this model is 1.77 K and is indicated by an open circle in figure 2.2 at the time of CO_2 doubling in the 1% per year scenario. The TCR is about 4% smaller than TCS because TCS is evaluated at the end of the fast transient while there may be some committed yet unrealized warming at the time TCR is measured. If, in reality, there is much less separation between the

timescale of the mixed-layer response and that of the whole-ocean then the TCS and TCR will differ more.

2.2.1 Probabilistic estimates of climate change

The previous example highlighted the mean TCS response given a mean value of λ , however, we are in fact more interested in how the entire probability density of λ influences the distribution of TCS. Similar to the transformation used by Roe and Baker (2007), when we take equation (2.5) as the map between λ and TCS, the following relationship describes the probability distribution for TCS,

$$\Pr(T_{\text{TCS}}) = \Pr(\lambda) \cdot \left| \frac{d\lambda}{dT} \right| = \frac{1}{\sqrt{2\pi}\sigma_\lambda} \exp \left[-\frac{1}{2\sigma_\lambda^2} \left(\frac{F_{2\text{CO}_2}}{T_{\text{TCS}}} - \bar{\lambda} \right)^2 \right] \cdot \frac{F_{2\text{CO}_2}}{T_{\text{TCS}}^2}. \quad (2.6)$$

The TCS distribution is computed assuming a normal probability density for λ with mean, $\bar{\lambda}$ and standard deviation, σ_λ . We illustrate the relationship in figure 2.3. Notice that the maximum likelihood TCS is not equal to the mean TCS because the transformation of the Gaussian λ leads to a skewed TCS distribution. The skew in this case is a mathematical artifact of the nonlinear map (2.5). As long as the standard deviation of λ is fairly well constrained by observations, the skew in TCS will remain small. Since the decision to make λ Gaussian was somewhat arbitrary, we will also examine how the results differ if we replaced λ everywhere with $1/\gamma$ where γ is Gaussian in the following chapter (3.6).

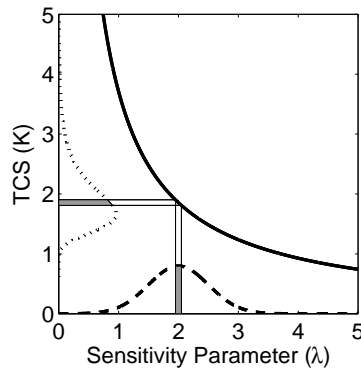


Figure 2.3: The map between λ and TCS defined by equation (2.5) (solid) and representative probability distributions for λ (dashed along horizontal axis) and TCS (dotted along vertical axis). The shaded areas underneath the curves illustrate that the map is area preserving.

2.3 Parameter estimation method

Various observationally-based estimates have been made of both ECS and TCS (or TCR) using a variety of statistical techniques and a range of model complexity; Knutti et al. (2008) provide a useful review.

Giorgi and Mearns (2002), Tebaldi et al. (2005) and Greene et al. (2006), for example, employ ensembles of comprehensive climate models, such as are described in the IPCC reports (e.g., Randall et al., 2007). These models try to represent the physical processes of the climate system, including processes determining aerosol forcing, in as explicit a way as possible. Although the physical parameterizations are tuned to simulate climate consistent with that observed, the ECS and TCS are not directly tuned by fitting to past climates; rather, they are obtained by integration of the model into the future under specific emissions scenarios. Still, model agreement with 20th century climates seems to depend in part on the trade off between historical aerosol level and climate sensitivity, and so some implicit tuning of climate sensitivity may occur (Kiehl, 2007; Knutti, 2008; Huybers, 2010).

Results from a collection of models developed in this manner may be combined to give a distribution of model sensitivities, but the distributions are effectively distributions of opportunity, rather than being properly controlled. The IPCC models only sparsely sample the space of all model formulations and may be compromised by the repeated use of observations for model development and verification (Sanderson and Knutti (2012)). To better fill out model-space, Ghil et al. (2008) have proposed the use of stochastic parameterizations. The idea is that, for a given level of noise, results from a diverse set of models will form a large diffuse cluster of probable output.

In other work, the uncertainty space of a single comprehensive model is explicitly explored. For example, the parameters of atmospheric and coupled models are sampled extensively resulting in the perturbed-physics ensembles of the climateprediction.net project. (Stainforth et al., 2005; Rowlands et al., 2012) Relatedly, using the technique of optimal fingerprinting (Stott and Tett, 1998), realizations of a single comprehensive model under different forcing scenarios are given probability weights based on the agreement of their spatial response patterns with past observations (e.g., Hasselmann, 1997; Allen et al., 2000; Stott and Kettleborough, 2002; Stott et al., 2006, and others).

At the other end of the model-complexity spectrum are methods that compute sensitivities based on linear regression of past forcing and observed climate such as Gregory and Forster (2008) and Murphy (2010). Advantages of these methods are their more direct use of observed data and independence of model biases and tuning.

The way that we shall proceed is to construct a simple but physically based model and then to try to constrain the parameters that determine the model’s transient climate sensitivity by a direct comparison with observations. In terms of model complexity, our methodology is closer to simple regression calculations than to the use of GCMs, but differs notably in that we seek to obtain time-dependent, probabilistic information. Specifically, we will constrain the output of a simple energy balance model by observations of the 20th century surface temperature record, using a particular nonlinear form of the Kalman filter as a way of estimating parameters over time. This sequential approach allows us to explicitly examine the way in which probability distributions depend on the underlying uncertainty assumptions and length of the observed record. Similar to the idea of Ghil et al. (2008), we expand the range of our simple model’s results by stochastically parameterizing natural variability and considering a broad range of uncertainty in historical forcing. Set against this, compared to the general circulation models, is the less comprehensive nature and the lack of detail of the predictions made.

2.3.1 Unscented Kalman filter approach

In order to estimate the parameter λ from past observations of temperature and forcing we use an adaptation of the Kalman filter applicable to nonlinear systems called the unscented Kalman filter based on Julier (2002) and van der Merwe (2004). (The term λT in (2.2) is formally nonlinear because both λ and T are regarded as state variables. Physically λ is a constant parameter, but the Kalman filter adjusts its value to find the best fit.) Although there are many methods by which to find probability distributions for unknown parameters, we use the nonlinear filter because it is a simple method to implement and provides well-founded probability estimates, though they are restricted to remain Gaussian. The recursive method has the additional advantage that in computing the posterior distribution given a time series of observations from t_1 to t_N , the posterior at every intermediate time, t_i , is automatically calculated. This feature thus enables one to study the evolution of uncertainty over time with the addition of observed temperatures. The filter

also accounts for model dynamics and time delays. A simple, static regression of the temperature against the forcing would, given sufficient data, give similar values for the TCS but less probabilistic information, with less ability to determine the effects of forcing uncertainty and natural variability separately.

The unscented Kalman filter (UKF) resembles the classical Kalman filter for linear systems and Gaussian random variables in that it is an approximate recursive Bayesian method. In general, all recursive Bayesian methods make use of prior and observed distributions of a model’s state to form an updated or posterior state based on the linear combination of the prior and observed states that minimizes the posterior error covariance. Each iteration, the forecast of the posterior at t_i becomes the prior at t_{i+1} . This process is outlined for a generic filter in appendix A which reviews key concepts and terms in sequential filtering.

For nonlinear systems, the minimization of the posterior error covariance cannot be solved exactly. Many different filters have been developed to address this problem. In the UKF, the error covariance and mean state are approximately computed from the statistics of an ensemble of state estimates. The UKF follows the general filter steps from appendix A but with the additional step of computing this ensemble. Ensemble member states are selected as ± 1 standard deviation perturbations about the mean, thus they are called sigma points. More details regarding sigma points are provided in section 2.3.2, including their mathematical definition (equation (2.7)). The complete filter process with update equations from appendix A written specifically in terms of the UKF implementation is outlined in algorithm 1.

The UKF may be thought of as a particular type of ensemble Kalman filter (Evensen, 1994, 2007) frequently used in data assimilation and sometimes applied to parameter estimation (Annan et al., 2005a, Annan et al., 2005b). For small systems, the UKF requires far fewer ensemble members and has equal or better accuracy than the standard implementations of the ensemble Kalman filter. Also, the sigma points are deterministically recomputed, enhancing accuracy and aiding in avoiding collapse of the ensemble. Nonetheless, had a conventional ensemble Kalman filter been used it would likely have given similar results.

Since λ is regarded as a variable with respect to the filter, the state to be estimated is the 3-dimensional vector $[T \quad \lambda \quad S]^T$. The dynamic update for λ is the persistence model, $\lambda_{k+1} = \lambda_k$. Known as state-space augmentation (Gelb et al. (1974)), the approach is a common technique

for including parameters in filter equations (e.g., Anderson (2001); Annan and Hargreaves (2004); Ambadan and Tang (2009); Yang and Delsole (2009)).

In addition to the state variables, the filter samples two sources of uncertainty in the system, input and output noise. The input uncertainties include Gaussian forcing uncertainty, w_F , and noise driving the natural variability of the system, w_S , with error-covariance R^w . The output uncertainty is additive white noise, v , representing measurement error in the historical temperature record, and has error-covariance R^v . The mean state is initialized with a best guess and the initial mean noise is zero as outlined in the first step of algorithm 1.

Each uncertainty source is perturbed, along with the state variables, to form the sigma-point ensemble according to the scaled-unscented equations of Julier (2002) as in algorithm 1 step 2a. The ensemble for our model consists of perturbations in 6 dimensions yielding a total of 13 sigma points, 12 symmetric perturbations plus 1 to include the mean. When considering alternative models for forcing uncertainty in section 2.4.3, the state dimension increases by one with the introduction of an aerosol forcing scale factor, α , and the noise dimension decreases by one with the elimination of w_F .

Each sigma point is forecast according to the nonlinear model (2.3) as in step 2(b)i of algorithm 1. The forecast mean state and error covariance are computed from the statistics of the forecast points at step 2(b)ii. (This is in contrast to the extended Kalman filter which loses some accuracy because it relies on linearized system equations and a single state to propagate the error covariance.) The forecast sigma points are mapped onto the space of observed variables by the measurement equation for our system, which happens to be linear in this case although that is not a requirement of the filter, $h_k(x_k^{(i)f}) = T_k^{(i)f} + v^{(i)}$ (step 2(c)i). The mean and covariance of the forecast sigma points are updated (or corrected) with the weighted difference between a real observation of the global average surface air temperature and the expected value (i.e. mean) of the modeled observations, $T_{obs} - E[T_k^{(i)f}]$ (steps 2(c)ii–2(c)iii). Finally, the covariance and mean state are augmented with the noise variables to prepare for the next iteration (step 2(c)iv).

2.3.2 Scaled-unscented sigma points

Following Julier (2002), we apply the scaled-unscented transformation when computing sigma points. This has the advantage of preventing loss of positive-definiteness in the error-covariance

over time while retaining up to fourth order accuracy. The scaled-unscented sigma points are calculated according to the following equations.

$$\begin{aligned}\chi_k^{(0)} &= x_{k-1}^{aug} \\ \chi_k^{(i)} &= x_{k-1}^{aug} + \sqrt{(L + \lambda)} \sqrt{P_{k-1}^{aug}}^{(i)}, \quad i = 1, \dots, L \\ \chi_k^{(i+L)} &= x_{k-1}^{aug} - \sqrt{(L + \lambda)} \sqrt{P_{k-1}^{aug}}^{(i)}, \quad i = 1, \dots, L\end{aligned}\tag{2.7}$$

where $\sqrt{P_{k-1}^{aug}}^{(i)}$ is the i -th column of a square root or Cholesky factorization of the covariance matrix after the analysis step of the previous iteration. Additionally, $\chi_k^{(i)}$ denotes the i -th column of the $L \times (2L + 1)$ sigma point matrix χ_k . $L = L_x + L_w + L_v$ is the dimension of the augmented state vector, the sum of the state, input, and output dimensions.

The corresponding covariance weights are

$$\begin{aligned}\omega_{c,k}^{(0)} &= \frac{\lambda}{L + \lambda} + 1 - \alpha^2 + \beta \\ \omega_{c,k}^{(i)} &= \frac{1}{2(L + \lambda)} \quad i = 1, \dots, 2L,\end{aligned}\tag{2.8}$$

where $\lambda = \alpha^2(L + \kappa) - L$. Here κ is the original parameter from the unscaled transformation affecting the accuracy of the higher order moments of χ_k . Setting $\kappa = 3 - L$ leads to a mean and covariance with up to fourth order accuracy. For systems of dimension greater than 3, a negative κ may cause loss of positive-definiteness in the covariance so Julier (2002) introduced the scaled-unscented transformation. As the scaling parameter $\alpha \in (0, 1]$ approaches zero, the original sigma points are drawn into a ball of decreasing radius thus minimizing the effects of error in higher order moments for large-dimensional systems. The parameter $\beta > 0$, in the zero-th covariance weight, provides an additional degree of freedom by which 4th-order error may be minimized. For the most concise notation, we write the weights as the $(2L + 1) \times (2L + 1)$ diagonal matrix $\Omega_{c,k}$ with the weights $\omega_{c,k}^{(i)}$ as the diagonal entries. Mean weights, $\omega_{m,k}^{(i)}$, are identical to the covariance weights of (2.8) except that the zero-weight is $\omega_{m,k}^{(0)} = \lambda/(L + \lambda)$.

Algorithm 1 Unscented Kalman filter

1. Initialize augmented state $x_0^{aug} = [x_0^a \quad 0 \quad 0]^T$ and covariance,

$$P_0^{aug} = \begin{bmatrix} P_0^{x^a} & 0 & 0 \\ 0 & R_0^w & 0 \\ 0 & 0 & R_0^v \end{bmatrix}. \quad (2.9)$$

2. For $k = 1$ to ∞

- (a) Calculate sigma points, χ_k , according to equations (2.7)–(2.8).

- (b) *Forecast or Dynamic Update*

- i. Forecast each sigma point by $\chi_k^f = f(\chi_k)$.
 - ii. Compute the statistical moments of the forecast sigma points: the mean state and error-covariance,

$$x_k^f = E[\chi_k^f] = \sum_{i=0}^{2L} \omega_{m,k}^{(i)} \chi_k^{(i)f} \quad (2.10)$$

$$P_k^{x^f} = A_k \Omega_{c,k} A_k^T \quad (2.11)$$

where $A_k = \chi_k^f - x_k^f$ is a matrix of perturbations about the mean forecast.

- (c) *Analysis or Measurement Update*

- i. Map the forecast sigma points to measurement space by $h_k(\chi_k^f, \chi_k)$ and form $Y_k = h_k(\chi_k^f, \chi_k) - E[h_k(\chi_k^f, \chi_k)]$, the perturbations about the mean measurement.
 - ii. Compute the cross and measurement covariances and weighting matrix,

$$P_k^{xy} = A_k \Omega_{c,k} Y_k^T \quad (2.12)$$

$$P_k^{yy} = Y_k \Omega_{c,k} Y_k^T \quad (2.13)$$

$$K_k = P_k^{xy} (P_k^{yy})^{-1}. \quad (2.14)$$

- iii. Compute the analysis mean and covariance,

$$x_k^a = x_k^f + K_k (y_k^{obs} - E[h_k(\chi_k^f, \chi_k)]) \quad (2.15)$$

$$P_k^{x^a} = P_k^{x^f} - K_k P_k^{yy} K_k^T. \quad (2.16)$$

- iv. Form augmented state, x_k^{aug} , and block-diagonal covariance, P_k^{aug} .
-

2.4 Validation of the method with data from comprehensive climate model

In this section we show that the use of the method (i.e., the energy-balance model in conjunction with the Kalman filter) is able to emulate the evolution over the 20th century of a comprehensive climate model (GFDL's CM2.1, Delworth et al., 2006), and furthermore that the method can predict the TCS of the GCM using only its forcing and temperature record of the 20th century. Since only a single realization of the real temperature record exists, we examine the extent to which single realizations of the GCM can be used to constrain transient climate sensitivity, rather than the average over an ensemble of integrations.

2.4.1 Comprehensive climate model data

To do this we consider separately as constraining data the five CM2.1 AR4 20th century integrations and their mean shown in figure 2.4. (The individual runs are smoothed with a 3 year moving average to reduce some of the unforced variability in the time series since this will be accounted for in the magnitude of σ_S .)

We model the perturbation forcing as the sum of a mean forcing record and white noise in each year, $F = \bar{F} + w_F$. The mean forcing, \bar{F} , for these experiments is the mean of ten forcing runs computed in Held et al. (2010), shown in figure 2.5. The uncertainty in the forcing, w_F , has standard deviation $\sigma_F = 1 \text{ W m}^{-2} \text{ K}^{-1}$.

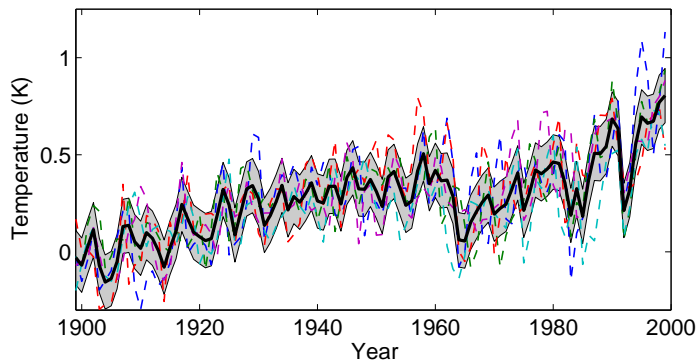


Figure 2.4: GFDL's CM2.1 historical temperature realizations. Five individual realizations (dashed color) and their mean (solid) with $\pm 0.14 \text{ K}$ about the mean (shaded).

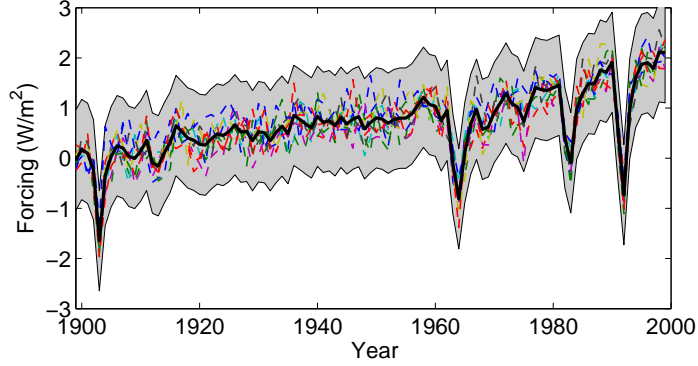


Figure 2.5: GFDL's CM2.1 historical forcing estimates. Data provided by Isaac Held. Ten individual realizations (dashed color) and their mean (solid) with $\pm 1 \text{ W/m}^2$ about the mean (shaded).

2.4.2 Parameter estimates

The estimate of the transient climate sensitivity parameter (λ) and its standard deviation, σ_λ , as determined by the nonlinear Kalman filter, are shown in figure 2.6. By the year 2000, the mean estimate of $\lambda = 2.6 \text{ W m}^{-2} \text{ K}^{-1}$ corresponds to a most likely value for $T_{TCS} = 1.4 \text{ K}$ which agrees well with the known TCS of CM2.1 of 1.5 K. The value of λ from the individual runs is less constrained but the estimates remain within a standard deviation of the mean with the exception of run 4 (cyan in figure 2.6) which remains within the 90% confidence interval. The uncertainty range throughout the 100 year time period is a little greater for the individual runs than the mean because they have greater natural variability. For run 4 the estimate deviates the farthest from the true λ as variability on longer time-scales in the temperature obscures forced features that help determine sensitivity. In the experiments run with the real temperature record, we partially offset this difficulty by removing the ENSO signal.

With the value of λ fixed at the mean estimate found by the end of the observational record, $\lambda = 2.6 \text{ W m}^{-2} \text{ K}^{-1}$, the EBM reproduces the temperature response of the comprehensive climate model, as shown in figure 2.7. Here the stochastic component of the EBM has been set to zero to more easily compare forced features in the responses. The EBM without any natural variability agrees well with the main features of the CM2.1 response such as to volcanic eruptions in the years 1902, 1963, 1982, and 1991 as well as the overall trend.

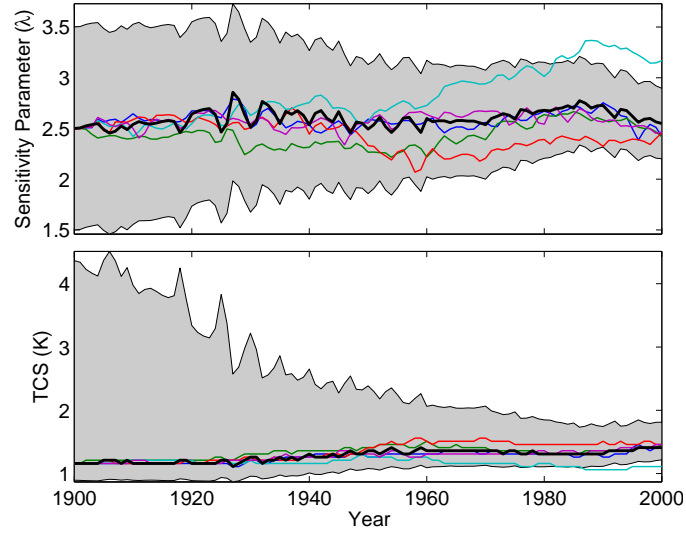


Figure 2.6: (top) Estimates of the sensitivity parameter λ for constraining temperature data from the individual CM2.1 realizations of figure 2.4 (thin color) and the 5-realization mean (solid) with its 90% confidence interval (shaded). (bottom) The maximum likelihood transient climate sensitivity corresponding to each of the distributions for λ in top panel with the 90% confidence interval about the mean TCS (shaded).

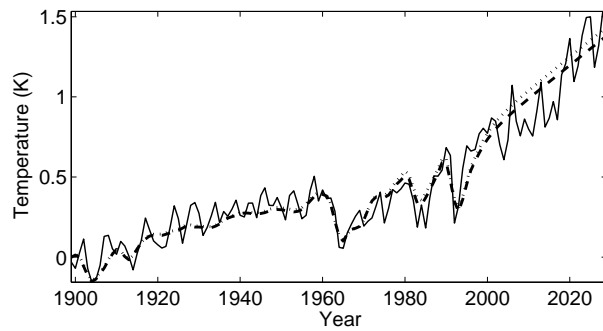


Figure 2.7: Comparison of temperature responses for the CM2.1 model (solid) and the mixed-layer model of section 2.1.2 with zero-mean Gaussian forcing uncertainty (dashed) and scaled forcing uncertainty (dotted) described in section 2.4.3. The CM2.1 data is the 5-realization mean appended with the first 30 years of the 1%/year to CO_2 doubling run. The stochastic component of the mixed-layer model has been set to zero.

2.4.3 Effect of assumptions on forcing uncertainty

In this section we further explore the effects of our assumptions regarding the uncertainty in the radiative forcing, the uncertainty that is often regarded as the biggest single impediment to calculating the equilibrium climate sensitivity from the past record. In the previous calculations corresponding to figure 2.6, this was modeled as a white noise, which is a good assumption for the forcing uncertainty in CM2.1, as may be observed in the forcing variations shown in figure 2.5. However, this is not a good assumption for the uncertainty in actual historical forcing, so in this section we introduce a more realistic model for forcing uncertainty and present results for CM2.1 data, as a precursor to doing the same for observed data.

The IPCC attributes the greatest source of forcing uncertainty to anthropogenic aerosols, reporting a 90% confidence range of -0.5 to -2.2 W m⁻² in 2005 (Forster et al., 2007). Although other sources of uncertainty are not insignificant, for simplicity we restrict uncertainty in our new forcing model to anthropogenic aerosols, and from here forward, aerosols (without a qualifying adjective) means those of anthropogenic origin. We separate the total historical forcing into aerosol and all other components,

$$F_{net}(t) = F_{other}(t) + \alpha F_{aero}(t). \quad (2.17)$$

We suppose aerosols are known only within a multiplicative scale factor α , which is a unity-mean, normally distributed random variable. Scaling the magnitude of aerosol forcing is an approach that has been adopted previously by Harvey and Kaufmann (2002); Forest et al. (2006) and others. The variance of α and the variance of F_{net} are related by,

$$\sigma_{F_{net}}^2 = F_{aero}^2 \sigma_{\alpha}^2, \quad (2.18)$$

which is consistent with the idea that the greater the magnitude of the aerosol forcing, the greater is the uncertainty about it. According to this model, $F_{other}(t)$ is known exactly, and is defined for both the GCM and real world applications as the sum of greenhouse gas, solar, and volcanic contributions estimated by Gregory and Forster (2008). (CM2.1 forcing data separated into individual components was not available.) The nominal aerosol estimate, F_{aero} , varies depending on the application. In this application to a GCM, we are concerned with CM2.1 forcing and temperature,

so we estimate CM2.1 aerosol forcing as the smoothed difference $F_{aero}^{CM2} = F_{net}^{CM2} - F_{other}$. The net CM2.1 forcing along with components: *aerosol* and *other* are plotted in figure 2.8. Also shown in the figure is the 90% confidence interval about the aerosols. Notice that since the variance of the forcing scales with the magnitude of the aerosols, when F_{aero}^{CM2} is near zero, uncertainty is quite small. Here the prior variance of the scale factor, σ_α^2 , is chosen such that the forcing variance $\sigma_{F_{net}}^2$ in 2005 is consistent with the IPCC confidence interval. By fitting a Gaussian and rounding up, we approximate the IPCC variance as $\sigma_{F_{net}}^2(2005) = 0.36 \text{ W}^2 \text{ m}^{-4}$. Then by equation (2.18), with $F_{aero}^{CM2}(2005) = -0.5 \text{ W/m}^2$, the variance of α at the start of assimilation is $\sigma_\alpha^2 = 1.4$.

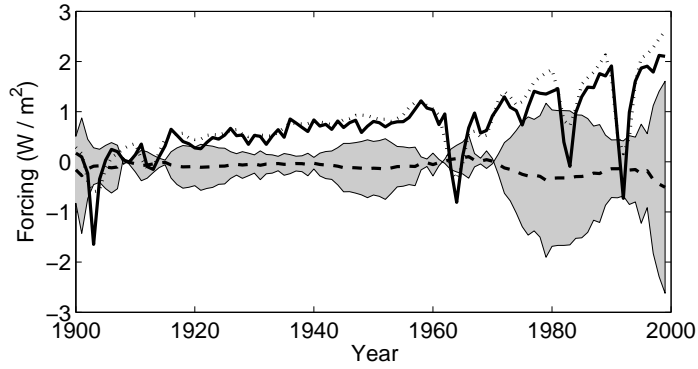


Figure 2.8: GFDL's CM2.1 net historical forcing (solid, same as the mean forcing in figure 2.5) and estimates of the individual forcing components: *other* (dotted) and *aerosol* (dashed) with 90% confidence interval (shaded).

We now allow the nonlinear Kalman filter to simultaneously constrain the parameters α and λ with the CM2.1 temperature data. In figure 2.9, we compare the results of a calculation with the new scaled forcing model to the results from the previous calculation with additive white forcing uncertainty. The spread of the λ -density increases, as expected, since errors in the longer term forcing trend are now taken into account. The mean estimate of λ remains very similar, with most likely TCS 1.4K closely matching the TCS of CM2.1, which is really 1.5 K. The mean scale factor estimate remains between about 0.5 and 1 ending the assimilation at about 1. The posterior uncertainty about α narrows slightly, indicating that temperature observations do marginally constrain α . In allowing the filter to estimate part of the forcing trend, there are now two mechanisms by which the model may be corrected to emulate the increasing observed temperature record: decreased aerosol forcing, achieved when $\alpha < 1$, and decreased λ (more sensitive). This makes it difficult to uniquely determine either parameter from the temperature time series alone, thus it

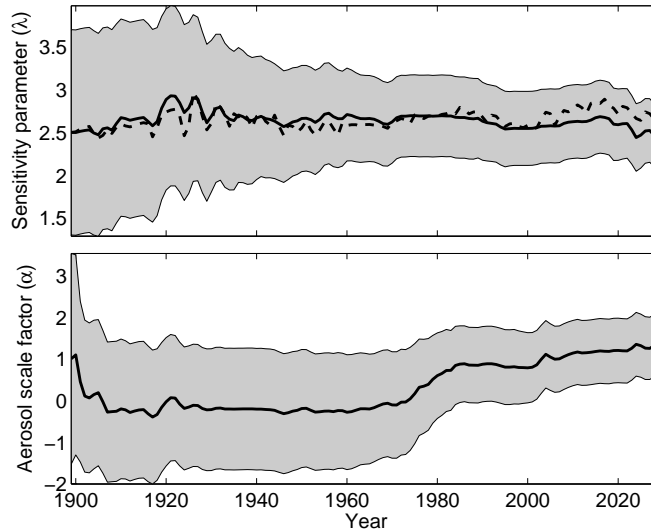


Figure 2.9: (top) The estimate of sensitivity parameter λ assuming scaled forcing uncertainty (solid) with the original estimate from white noise assumptions (dashed) shown again for comparison. The shaded band is $\pm 1\sigma$ about the mean. (bottom) The estimate of aerosol scale factor α with $\pm 1\sigma$ (shaded)

is important to consider entire probability distributions rather than focusing solely on mean estimates. Nevertheless, and as shown in figure 2.7, ignoring stochastic variability, the 20th century GCM response closely resembles the mean evolution of the EBM with estimates of sensitivity parameter, $\lambda = 2.55 \text{ W/m}^2$, and forcing scale factor, $\alpha = 0.8$, conditioned on observations through the year 2000.

2.5 Concluding remarks

The main conclusion to be drawn from the above exercise is that the methodology of using the EBM in conjunction with an unscented Kalman filter, when applied to the 20th century record of globally averaged surface temperature and forcing taken from a comprehensive climate model, is able to estimate, within reasonable error bounds, the transient climate sensitivity of the comprehensive model of about 1.5K for a doubling of CO_2 . Using this value of TCS along with estimates of natural variability, the EBM is able to produce plausible trajectories of 20th century warming that are visually indistinguishable from trajectories of the GCM. There is, however, some sensitivity to the nature of the assumed uncertainty in the forcing. Nevertheless, these results give us confidence to proceed with applying the method to real data.

Chapter 3

Closure estimates from climate observations

Given the success of the Kalman filter method using data from a comprehensive climate model in the preceding chapter, in this chapter we apply the same method for closing the unknown climate sensitivity parameter (λ) using real observations of the temperature and climate forcing recorded over the 20th century. The use of historical data introduces additional challenges to estimation of λ . The actual trajectory of globally-averaged temperatures over the past century has been somewhat less linear than the temperature response of the GFDL CM2 general circulation model (GCM). For example, the real temperature response to past volcanic eruptions was weaker than predicted, periods of unforced variability were present with autocorrelation as long as decades, and relatively little temperature change occurred in the 2000 decade. Due to greater uncertainty about the fit of the linear model (2.3)–(2.4), rather than providing a single probability density for λ , the contribution of this chapter is a range of distributions that sample the uncertainties associated with real forcing and temperature data. Also presented is the sensitivity of results to other sources of uncertainty such as prior distributions, length of the data record, and value of the mixed layer depth.

3.1 Observational data

The observed temperature, figure 3.1, is derived from Thompson et al. (2009), which itself is derived from the HadCRUT data (Brohan et al., 2006). We have annually averaged the residual after subtracting the ENSO signal which has the effect of making volcanic responses more pronounced. In the figure we also show some temperatures after 2008; these have been extrapolated with synthetic data generated by the mixed layer model (2.2) introduced in the previous chapter, where the climate sensitivity parameter has been set to $\lambda = 2.0 \text{ W m}^{-2} \text{ K}^{-1}$, which is approximately the mean estimate we obtain in 2008, and a forcing corresponding to a 1% increase per year in CO_2 with no change in aerosol forcing. That is, we essentially create a climate realization after 2008 using the simple model that we can then analyze with the Kalman filter. (The results are not especially sensitive to the slope of the extrapolated data on the time scales considered and yearly variability in the record is not a major factor in the results, as we show in section 3.3.2.) The standard deviation of measurement errors in the observed temperature record is taken to be $\sigma_v = 0.06 \text{ K}$, as estimated by Brohan et al. (2006).

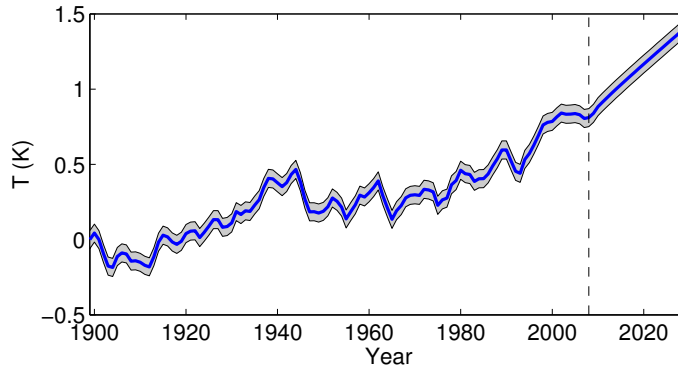


Figure 3.1: Observed global temperature change after removal of ENSO signal. Measurement error of $\pm 0.06 \text{ K}$ is shown shaded. Vertical dash line marks the end of real data and beginning of use of synthetic data in the future. Data beyond 2008 was generated by linear extrapolation.

To account for natural variability we assume for our base case that the natural variability [S in equation (2.4)] has a standard deviation of $\sigma_S = 0.07 \text{ W m}^{-2}$, which gives $\sigma_T \approx 0.13 \text{ K}$, this being the standard deviation of the observed detrended 20th century temperature record. Since these parameters are only approximate, we also calculate results for a range of values of σ_S and thus σ_T . These are discussed in section 3.3.1.

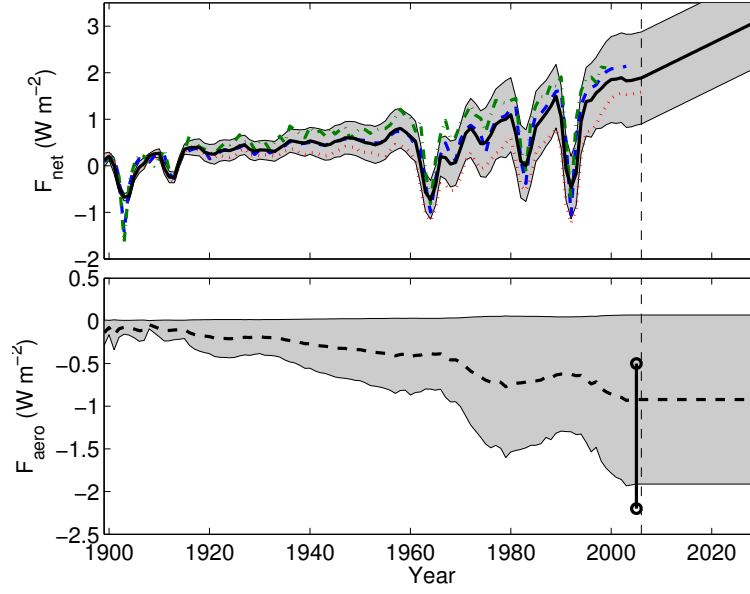


Figure 3.2: (top) Mean of three 20th century forcing estimates (solid black) from the following sources, which are also plotted individually for comparison: GFDL (green dash-dot), GF08 (red dotted), and GISS (blue dashed), plus 90% confidence interval (shaded). Data after 2006 assumes CO_2 forcing increases 1% per year. (bottom) Estimate of anthropogenic aerosol forcing with levels held constant after 2006 (black dashed) plus prior uncertainty at the 90% level (shaded) and the corresponding IPCC interval (black vertical).

The forcing is derived from three sources which we denote as GISS (Hansen et al., 2007), GFDL (Held et al., 2010) and GF08 (Gregory and Forster, 2008). These forcings, shown in the top plot of figure 3.2, are obtained by slightly different techniques and represent slightly different levels of atmospheric adjustment. For our base case we take F_{net} to be the mean of the three, although due to the differing lengths of the records after 2000 the mean forcing is the average of GISS and GF08, and from 2005 to 2006 it is GF08 alone. After 2006 we extrapolate the forcing by assuming 1% per year increases in CO_2 with no change to other components. Differences in the time series become apparent in the middle of the 20th century and continue to grow throughout the time period due mainly to unknown aerosol forcing. Therefore, we model this growing uncertainty as an unknown scale factor as we did for GCM data in section 2.4.3 by equation (2.17). Had we just used one of the forcing records, instead of the mean, there would have been small quantitative differences that are within the uncertainty bounds that we also calculate. Further discussion of the consistency of results with respect to forcing assumptions is provided in section 3.3.3.

Consistent with the method developed in section 2.4.3, F_{net} is separated into F_{other} , the sum

of Gregory and Forster (2008) solar, volcanic and greenhouse gas contributions, and anthropogenic aerosol forcing, F_{aero} , the smoothed difference between F_{net} and F_{other} . We plot F_{aero} in the bottom plot of figure 3.2 and infer the prior uncertainty about F_{aero} , also shown, from the IPCC aerosol uncertainty range in 2005, assuming that the uncertainty is proportional to the aerosol level itself. Approximating the IPCC 90% confidence interval of -0.5 to -2.2 W m^{-2} as Gaussian as we did previously for the GCM experiments, by equation (2.18) the prior variance of the scale factor at initialization of the filter should be $\sigma_\alpha^2 = 0.42$. This yields the prior 90% confidence interval shown as the shaded regions in figure 3.2 for F_{net} (top) and F_{aero} (bottom).

The IPCC interval estimated in the year 2005 is also shown for comparison in the bottom plot. Although the IPCC range extends to slightly higher aerosol forcing than our Gaussian fit, we are not concerned about this small level of disagreement. Both ranges likely overestimate the aerosol uncertainty because they are prior guesses, unconstrained by data. The observational data we assimilate will revise and constrain the forcing uncertainty. Recall from section 2.4.3 that the variance of the forcing is proportional to both the variance of the scale factor and the magnitude of the aerosol forcing. Since we let the Kalman filter determine the most likely trajectory of the scale factor and its uncertainty, the posterior confidence interval of the forcing will narrow slightly as uncertainty about the scale factor narrows.

3.2 Parameter estimates

With the forcing and temperature data described above, beginning in 1900, we employ the nonlinear filter to sequentially update the probability density for λ and α as more observations are included as constraints. The time-varying mean and standard deviation of the λ -density and α -density are shown to steadily narrow over time in figure 3.3 in the top and bottom panels, respectively.

The uncertainty in the distribution declines throughout the period of observation as a result of more data points unveiling the temperature trend. One may notice that the year 2000 estimates of mean λ are slightly more sensitive compared to 2008. This is a result of the flat to decreasing temperature trend in the last decade while forcing continued to increase. Beyond 2008, there is (as expected) little change in the mean estimate because we have fixed the sensitivity of the synthetic data close to the 2008 sensitivity.

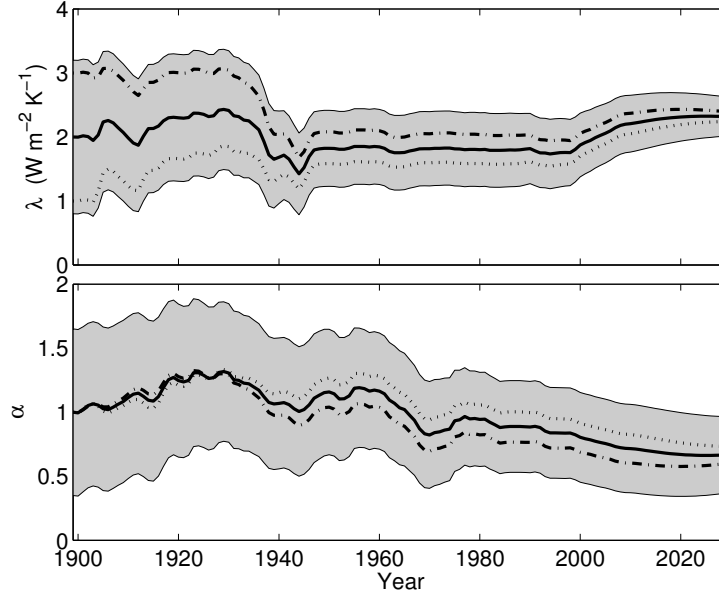


Figure 3.3: (top) Mean estimate of transient sensitivity parameter, λ , for three prior means: $\lambda(1900) = 2.0$ (solid), 1.0 (dotted), and $3.0 \text{ W m}^{-2} \text{ K}^{-1}$ (dash-dot) and $\pm 1\sigma$ (gray shaded). (bottom) Mean estimate of scale factor, α , for the three prior mean λ values and $\pm 1\sigma$ uncertainty (gray shaded).

Also shown are the results of shifting the prior λ mean to high and low values of 1 and $3 \text{ W m}^{-2} \text{ K}^{-1}$. Despite the quite different starting points in 1900, the mean trajectories converge steadily over time. More of the effects of varying initial uncertainty are discussed in sections 3.3.1 and 3.6.

Mapping the time evolution of the Gaussian λ -density of figure 3.3 to the transient climate sensitivity via equation (2.5) yields the skewed TCS probability density whose 90% confidence interval as a function of time is shown as the shaded region in figure 3.4. The peak in the distribution is plotted as the dashed line. The prior distribution in year 1900 is noticeably skewed, with the 95th percentile including temperatures in excess of 10 K. As more data becomes available, the posterior TCS distribution continues to narrow until around 1940, when a large unforced temperature perturbation causes the sensitivity estimate to decrease. Even though the overall spread of the λ -density narrows, the spread in TCS actually increases. This is a feature of the nonlinear relationship between λ and TCS in equation (2.5). For small λ , large improvements in our understanding of λ translate into only modest improvements in the confidence bounds of the TCS. Similarly, the decline in uncertainty after 2000 may be attributed to the increase in λ . For

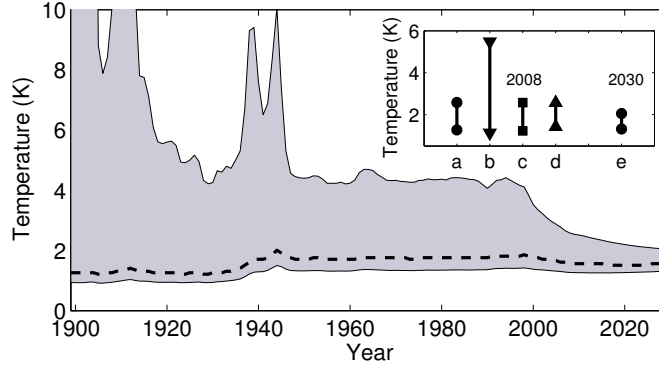


Figure 3.4: Evolution of TCS probability as more data from the past is used as a constraint. Most likely TCS indicated by the dashed line. Shaded region is 90% confidence interval. Inset: 90% confidence intervals for the uncertainty scenarios outlined in table 3.1: most probable (a), high past natural variability (b), forcing uncertainty is high (c), and forcing uncertainty is low (d), all in 2008, and most probable in 2030 (e).

further discussion of the implications of (2.5) for skewing the TCS distribution see section 3.7.

In later years, the distribution of λ has narrowed sufficiently that the skewness of the TCS distribution is no longer a prominent feature. In 2008, the 90% TCS confidence interval is 1.3–2.6 K and that range is reduced by 45% by 2030. (To avoid the effects of high aerosol uncertainty, in section 3.5 we describe the effects of only using data from 1970 on; in fact these lead to similar estimates of λ and TCS.)

3.3 Sensitivity to temperature and forcing

To understand the robustness of the parameter estimates to assumptions about past temperature change and forcing, we vary these assumptions and present the subsequent changes to the parameter probability distributions. In the first set of experiments, described in section 3.3.1, while holding the mean forcing and temperature trajectories constant, the magnitude of the uncertainty about each trajectory is varied. In section 3.3.2, we present the results of changing the forcing and temperature trajectories, for example by supposing there were no volcanic eruptions or that the trends were linear. Finally, in section 3.3.3, the trajectory of the mean historical forcing is varied among the estimates from three different studies.

3.3.1 Sensitivity to magnitude of uncertainty in forcing and temperature

For comparison with the parameter estimates obtained in section 3.2 using assumptions about uncertainty detailed in section 3.1 that we consider most plausibly supported by observations, we also consider three limiting cases for past uncertainty: forcing uncertainty 50% larger, forcing uncertainty 50% smaller - both with our plausible estimate of unforced variability, and plausible forcing uncertainty with larger natural variability in the temperature record. These uncertainty scenarios are summarized in table 3.1 along with the resulting 90% confidence intervals after assimilation of surface temperature data up to 2008 and 2030. These same confidence intervals are also shown in the inset plot of figure 3.4 and the complete probability densities of TCS are plotted for selected scenarios in figure 3.5.

Experiment	σ_T (K)	σ_α	90% C.I. 2008 (K)	90% C.I. 2030 (K)
Most plausible	0.13	0.65	1.3–2.6	1.3–2.0
Large F uncertainty	0.13	0.98	1.2–2.6	1.3–2.0
Small F uncertainty	0.13	0.33	1.4–2.6	1.4–2.1
Large natural variability	0.5	0.65	1.1–5.5	1.2–3.5

Table 3.1: Summary of natural variability and scale factor uncertainty for each of four different experiments. The resulting TCS confidence intervals given each set of assumptions are also included.

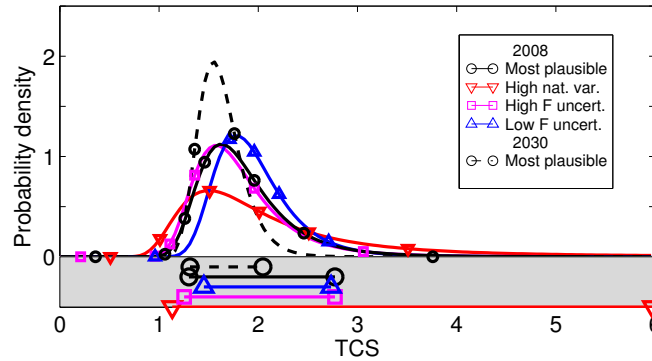


Figure 3.5: The range of PDFs and 90% confidence intervals for the transient climate sensitivity for the various sets of assumptions made. With data up to 2008: estimates based on most plausible forcing uncertainty and natural variability, forcing uncertainty underestimated, forcing uncertainty overestimated, and natural variability overestimated. With data up to 2030: most plausible forcing uncertainty and natural variability.

The combined range of confidence intervals is an indication of the sensitivity of estimates to assumptions about uncertainty magnitude. As expected, the larger the forcing uncertainty and

natural variability, the broader becomes the spread in the estimated λ . Since the mean value of λ stays about the same in all cases, the TCS distributions also broaden with the development of a fat tail most easily seen in figure 3.5. All of the estimates by the year 2030 have narrowed substantially from their 2008 ranges. The additional two decades of observations provide a strong constraint on the tail of the distribution, curtailing the highest TCS values, while having very little effect at the low end. The broadest ranges in both 2008 and 2030 were obtained for the large natural variability scenario, in which we more than tripled the standard deviation of natural variations, $\sigma_T = 0.5$ K. At this level of variability, the temperature observations do little to constrain the distribution due to the very low signal-to-noise ratio.

Besides the above extreme cases, we also experimented with values of σ_T varying between 0 and 0.5 K and the initial uncertainty in the scale factor, σ_α , varying between 0.3 and 1.0. The parameter estimates from a full factorial set of experiments are shown as contours plotted against $(\sigma_\alpha, \sigma_T)$ pairs in figures 3.6 and 3.7. In both figures, the top plot is the mean estimate of λ and the bottom is the standard deviation, σ_λ . The values of $(\sigma_\alpha, \sigma_T) = (0.65, 0.13)$ from the most plausible scenario above are drawn for reference as dashed lines in figure 3.6.

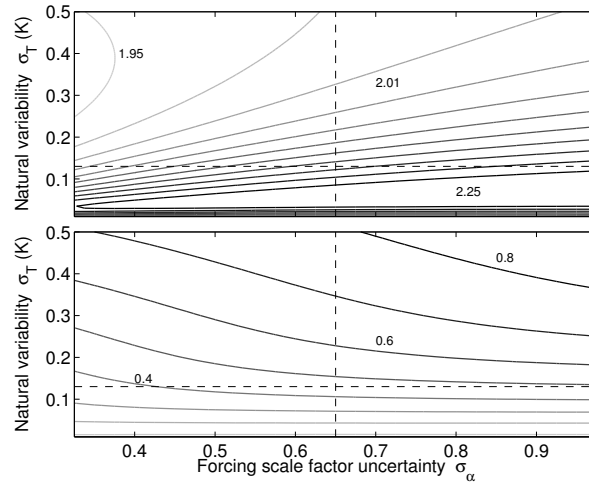


Figure 3.6: (top) Contours of mean transient sensitivity parameter, λ , where the interval between contours is $0.03 \text{ W m}^{-2} \text{ K}^{-1}$. (bottom) Uncertainty, σ_λ , with contour interval $0.1 \text{ W m}^{-2} \text{ K}^{-1}$. λ mean and uncertainty estimates at each value of forcing uncertainty and natural variability are from 2008, after assimilation historical observations between 1900 and 2008. Dashed lines indicate most plausible scenario for forcing uncertainty and natural variability.

We observe that the mean transient sensitivity parameter is not very sensitive to either natural variability or scale factor uncertainty and that of the two, it is the least sensitive to scale factor

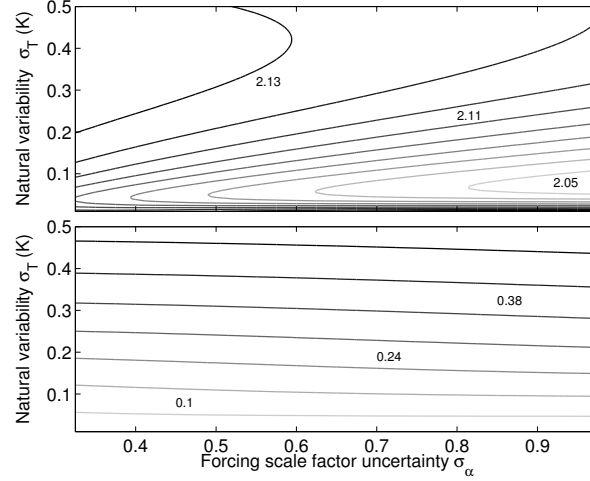


Figure 3.7: (top) Same as for figure 3.6 with contour interval $0.01 \text{ W m}^{-2} \text{ K}^{-1}$ and (bottom) same as for figure 3.6 with contour interval $0.07 \text{ W m}^{-2} \text{ K}^{-1}$ but where estimates are shown from the year 2030 after assimilation of historical observations between 1970 and 2030.

uncertainty, as evidenced by the low sloping contours in top of figure 3.6. Regarding the broadness of the λ -distribution, in the bottom of figure 3.6, we see that above $\sigma_T = 0.13 \text{ K}$, increasing both sources of uncertainty leads to broader distributions. However, the flatness of the contours below about $\sigma_T = 0.13 \text{ K}$ indicates that in this region, even large changes in the scale factor uncertainty have minor to no affect on the width of the λ -distribution. Here, uncertainty about λ is primarily limited by the magnitude of the natural variability. The same observation can be made from the similarity of the confidence intervals a, c, and d in the figure 3.4 inset.

Unlike, measurement uncertainty (for example) which may be reduced in the future with improved technology, one cannot reduce the natural variability of the climate system. According to our method, further uncertainty reduction in predictions of future climate change can only occur when the signal emerges more clearly from the noise in longer observational time series, such as shown in figure 3.7. This limitation, however, may be alleviated by recent work from Chekroun et al. (2011) in which past natural variability is used to predict and constrain the path of future variations. In future work, this approach could be used with our method to restrict the ensemble of model forecasts made in the prediction step of the Kalman filter.

3.3.2 Sensitivity to natural temperature fluctuations

Given that we have only one realization of the temperature record, and with forcing estimates that are partly dependent on GCM results, we now show that our parameter estimates are not overly sensitive to fluctuations in each trajectory. In figure 3.8, we repeat experiments, replacing the actual temperature and forcing data with alternate time series shown in figure 3.8 (a) and (b).

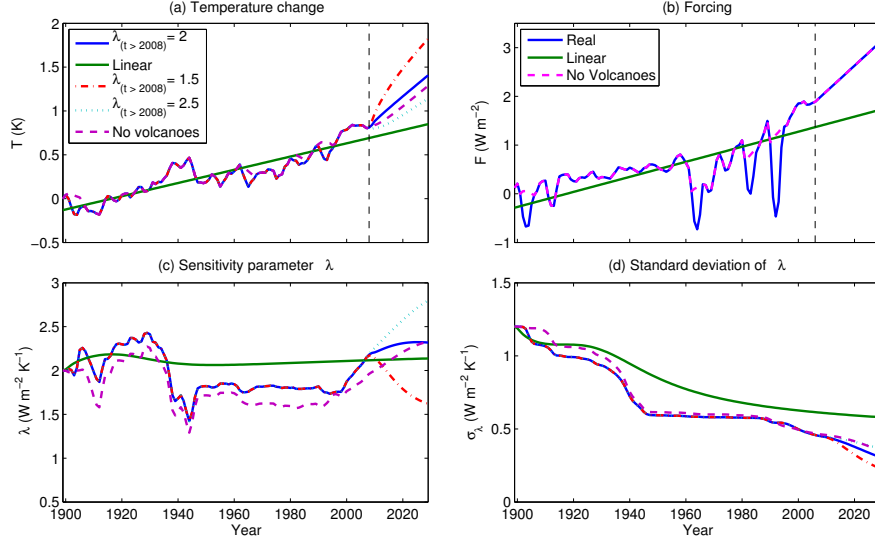


Figure 3.8: Alternative temperature (a) and forcing (b) time series for experiments in which: the data has been linearly fit (green solid), the temperatures after 2008 (denoted with a vertical dashed line) vary in slope according to prescribed sensitivities of $1.5 \text{ W m}^{-2} \text{ K}^{-1}$ (red dash-dot) and $2.5 \text{ W m}^{-2} \text{ K}^{-1}$ (cyan dotted), and volcanic effects have been interpolated out (magenta dashed). Also, the evolution of the mean (c) and standard deviation (d) of the probability density for λ when constrained by the alternative data.

First we apply straight line fits to the data from 1900 to 2008, effectively removing all unforced variability as well as some forced features such as due to volcanoes. For the linear data, the results are not significantly different, with the most noticeable change being a slightly higher uncertainty. The estimates vary smoothly over time due to the lack of any fluctuations in the trajectories.

In another experiment, we remove volcanic features through linear interpolation between the data in years bounding an eruption. Estimates of λ without volcanoes are slightly lower, indicating a higher TCS, or more sensitive climate than predicted with volcanoes included in the record — see also figure 3.10. In general, the observed response to volcanoes is, arguably, somewhat weaker than might be expected, presuming the forcing to be correct. It is possible that, because volcanoes cool, the effective oceanic heat capacity is larger in the period after an eruption, but further study

of this seems warranted. However, overall, estimates with and without volcanoes are similar.

Experiments so far have examined how uncertainty changes in the future as more observations are collected using just one realization of future temperatures derived from the simple model (2.2) with a nominal choice of $\lambda = 2.0 \text{ W/m}^2$, which was near the mean estimate in 2008. Here we also look at the effects of varying the assumed future data modeled with $\lambda = 1.5$ and $\lambda = 2.5 \text{ W m}^{-2} \text{ K}^{-1}$ which effectively modifies the slope of the temperature record after 2008. The λ estimates diverge immediately from their 2008 value toward the prescribed values of 1.5 and 2.5 $\text{W m}^{-2} \text{ K}^{-1}$ but do not deviate more than the standard deviation of λ in the 20 year period to 2030. Also, the reduction in uncertainty about λ is very similar for both records, so our estimate of how much uncertainty may be reduced by 2030 is valid even though we may be assuming incorrect future temperatures.

3.3.3 Sensitivity to forcing trajectories

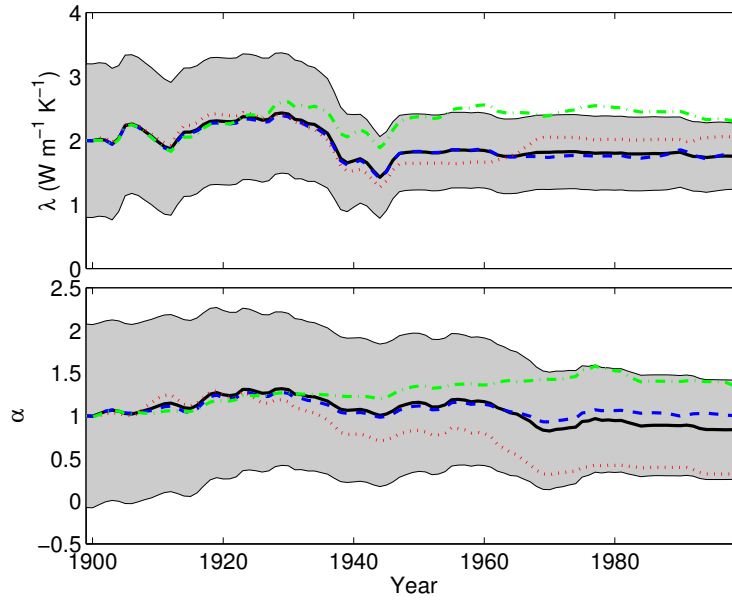


Figure 3.9: (top) The evolution of λ when the forcing is taken to be GISS alone (green dashed), GF08 alone (red dash-dot), GFDL alone (magenta dotted), and the three forcing mean (blue solid). The gray band is one standard deviation about the mean estimate. (bottom) As for top panel but for α . Here the gray band is the 90% confidence range.

In this section we revisit the issue of the forcing time-series. As is evident from figure 3.6 (bottom), the uncertainty in forcing may play a major role in limiting the skill in our observational

estimates of TCS. We have modeled the forcing uncertainty as an unknown scaling of the contribution of anthropogenic aerosols because we note that the three forcings, GISS, GFDL, and GF08, that we have used, diverge from each other with time (figure 3.2 (top)) rather than oscillating round each other, as in the GCM forcing error of figure 2.5.

Given this, we may see if our estimates of the sensitivity parameter are particularly sensitive to the specific forcing time series. Thus, for example, suppose that we only had available one of the three forcings we have used — would the computed parameter, λ , differ noticeably? The results of doing just this are shown in figure 3.9. The estimates of λ remain within about one standard deviation of the mean estimate as indicated by the gray shaded region regardless of which forcing trajectory is assumed: GISS and GF08 forcing yield λ well within the envelope and GFDL forcing leads to λ slightly outside. A white noise uncertainty model would underestimate the uncertainty in λ and cause the estimates using the three forcings individually to fall well outside the one standard deviation envelope. The estimates of the scale factor, which vary more widely due to the quite different forcing time series, exceed the standard deviation but remain within the 90% confidence range about the mean.

3.4 Sensitivity to mixed layer depth

All calculations up to this point have been made assuming an effective heat capacity, C , in the mixed layer energy balance (2.2) corresponding to a mixed layer 60 m deep. We now explore the effects of varying the mixed layer depth, H , on estimates of TCS and the sensitivity parameter, λ .

We find that the mean λ , constrained by surface temperatures through 2008 that include volcanic effects, is largely unaffected by depths ranging from $H = 40$ m to $H = 200$ m, in figure 3.10. However, for $H < 40$ m, the mean estimate of λ increases rather sharply (meaning the temperature response becomes less sensitive to forcing). This occurs because low-heat-capacity models have a large temperature response to volcanic forcing whereas the observed temperature, in fact, responded very little, as is evident in figure 3.11 where the simulated temperatures with $H = 1$ m overshoot observation in volcano years. To correct for this, Kalman filter nudges the model’s sensitivity so that the response to volcanoes is not overly exaggerated, in better agreement with observations. Thus a model with too low heat capacity would falsely predict a very insensitive climate.

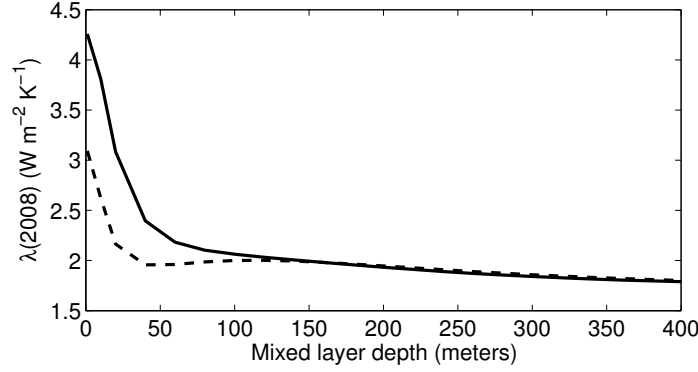


Figure 3.10: Estimates of mean transient climate sensitivity parameter by 2008 as a function of mixed layer depth in the simple model (2.2), with (solid) and without (dashed) volcanic response included in the constraining temperature data.

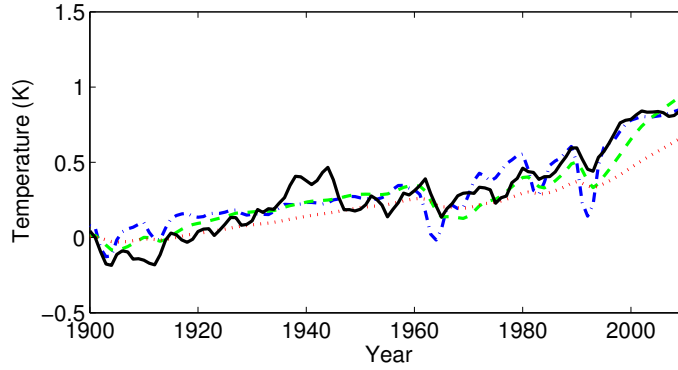


Figure 3.11: Surface temperature response of the simple model, equation (2.2), with λ and α fixed at 2008 means, for mixed layer depths of 1 m (blue dash-dot), 120 m (green dashed), and 400 m (red dotted) compared to historical temperatures (black solid).

On the other hand, increasing the heat capacity smooths and dampens the model's temperature response. For large depths, $H > 200$ m, the model is quite insensitive to forcing fluctuations, so mean λ is estimated as a bit more sensitive in order to better simulate the historical record. For example, see the simulation with $H = 400$ m in figure 3.11, the volcanic response is too smooth and the fit becomes increasingly poor beyond 1970. A model with too high heat capacity would falsely predict a very sensitive climate.

When we remove volcanic effects by interpolation, λ remains quite constant for all mixed layer depths as small as 10 m. Therefore, simulating volcanoes helps to constrain the mixed layer depth. Volcanoes indicate that our model's mixed layer depth should be larger than 40 m and smaller than 200 m. Therefore, our choice of $H = 60$ m is appropriate, although at the lower end of this range, sensitivity estimates would err toward the insensitive.

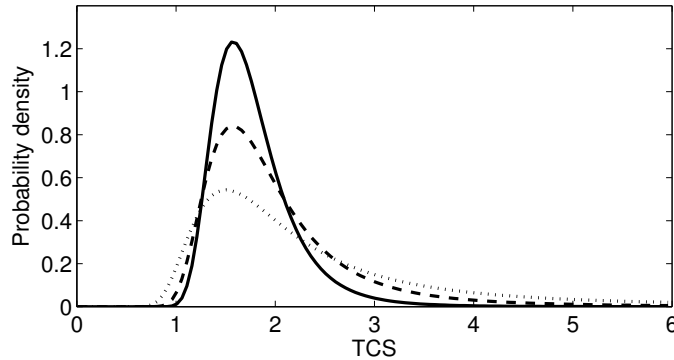


Figure 3.12: Probability density distributions for the transient climate sensitivity in 2008 with mixed layer depths of 60 m (solid), 120 m (dashed), and 400 m (dotted).

Also, as mixed layer depth increases, uncertainty surrounding estimates of the parameters λ and α steadily increases. This is expected because the greater the thermal inertia of the model, the longer the time series of temperature observations needed to increase the signal to noise ratio and thus reduce uncertainty in the nonlinear Kalman filter. Overestimating the mixed layer depth, therefore, may lead to slow-to-converge parameter estimates that ignore much of the observed signal. On the other hand, underestimating the depth could result in overconfident probability densities. Increasing uncertainty with depth is evident in the lengthening of the tails of the distributions of figure 3.12, where TCS probability densities are compared for mixed layer depths of $H = 60, 120$, and 400 m.

3.5 Sensitivity to historical time-period

The uncertainty in aerosol forcing is generally believed to be greatest in the middle part of the last century and less in the last third of the century and in the early part of the 21st century. Therefore, one may be interested in estimates of TCS using only the data from 1970 on, and this route was taken by Gregory and Forster (2008). The disadvantage is that the shorter time period means that the uncertainty due to natural variability can be expected to be larger. In figure 3.13 we show results using only data from 1970 and using otherwise the same uncertainty assumptions as used previously. The change in the aerosols and thus the three forcing estimates (GISS, GFDL and GF08) in this period are now very similar to each other, and so our estimate of the standard deviation of the forcing scale factor is rather conservative (i.e., probably too large).

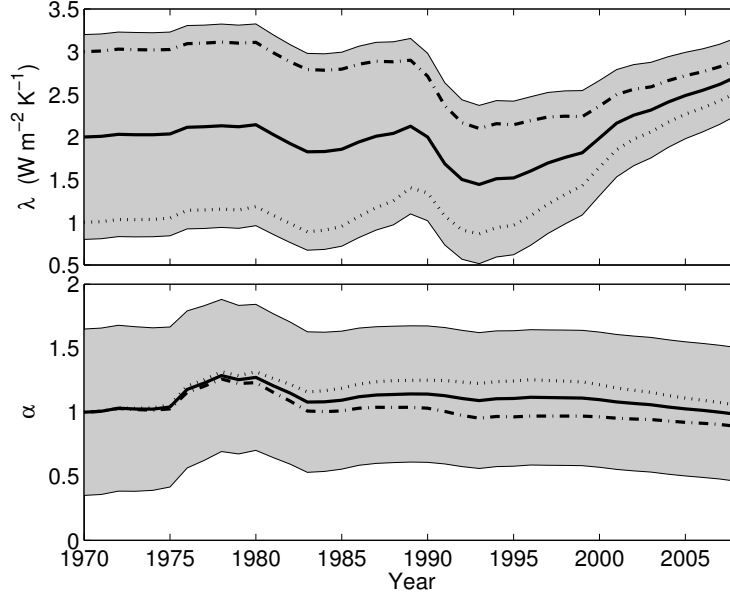


Figure 3.13: As for figure 3.3 but using data only from 1970.

Rather encouragingly, estimates of λ and so of TCS calculated in this manner are similar to the estimates constrained by data all the way back to 1900, although the relatively steady temperatures after 2000 dominate the shorter time window and lead to slightly less sensitive estimates of TCS, with a 90% confidence interval of 1.1–1.9 K by 2008. The spread of this interval is slightly narrower, even with the smaller data set, than the estimates from the longer record, and this may appear somewhat paradoxical. One reason for the reduced spread may be that the simple EBM better simulates the more recent period and so the model learns faster. A second reason is that the generally larger λ leads to a TCS distribution with less skew and a reduced likelihood of a very large temperature response: given two lambda distributions with the same spread but centered about different mean values, the distribution with the less sensitive mean will have a smaller TCS uncertainty range. The results using the shorter record are also a little more sensitive to the priors that are chosen, and so the results arising from the use of temperature record over the entire century probably reflect better the true uncertainty, but as noted the differences are small.

Previously, in section 3.3.1, we also examined how uncertainty estimates change should only data from 1970 projected to 2030 constrain the parameters for the many $(\sigma_\alpha, \sigma_T)$ combinations. Figure 3.7 (top) shows that λ is largely unaffected by variations in this shorter time period. The bottom plot shows that the standard deviation of λ decreases as the natural variability and forcing

uncertainty decrease, although λ uncertainty is much more sensitive to changes in natural variability over this shorter time period.

3.6 Sensitivity to prior uncertainty

The prior in any Bayesian analysis is often difficult to assess and may require some subjective analysis. Our aim is to show that any subjectivity in prior λ distribution does not compromise the objectivity of the data and posterior. We vary the spread of the Gaussian prior to get a sense of how strongly it influences the posterior results (the Kalman filter requires all distributions to be Gaussian so a completely uninformative prior is not an option). Figures 3.14 and 3.15 show the effect the prior uncertainty of λ has on the posterior distributions for the parameters λ and α by the end of the 130 year data assimilation period.

As evidenced by figure 3.14, about 50 years of data contain enough information to effectively forget the prior; the posteriors have converged to the same distribution. The narrowest prior used, $\sigma_\lambda(1900) = 0.6 \text{ W m}^{-2} \text{ K}^{-1}$, is probably too small since its posterior standard deviation remains slightly smaller than the others until 2030 (figure 3.14 bottom) and the data hardly reduces its value. Halving the prior λ uncertainty had a minimal effect on the mean and uncertainty after 2000. Increasing prior uncertainty caused greater variance in λ before 1940 with quick convergence to the distributions shown in figure 3.3.

In figure 3.15, it is clear that the distribution of the parameter α is not affected by prior assumptions in λ . Note that varying the prior mean value of λ is addressed in section 3.2. These experiments give us confidence in our choice of prior uncertainty, $\sigma_\lambda = 1.2 \text{ W m}^{-2} \text{ K}^{-1}$, and the objectivity of the final posterior distributions.

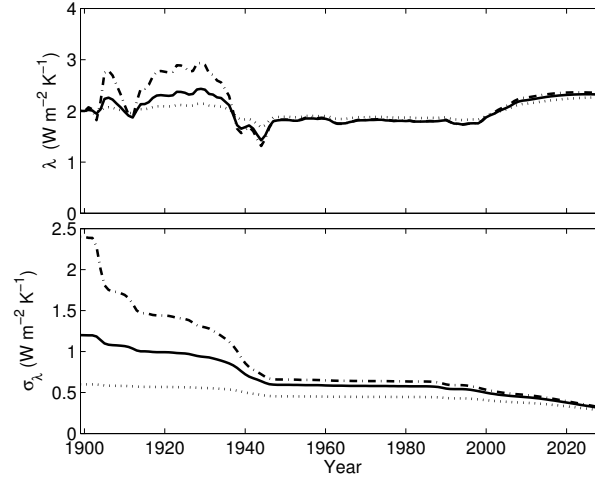


Figure 3.14: Estimates of the mean (top) and standard deviation (bottom) of the transient climate sensitivity parameter, λ , for three different prior assumptions on uncertainty in λ : $\sigma_\lambda(1900) = 0.6$ (dotted), 1.2 (solid), and 2.4 (dash-dot) $\text{W m}^{-2} \text{K}^{-1}$.

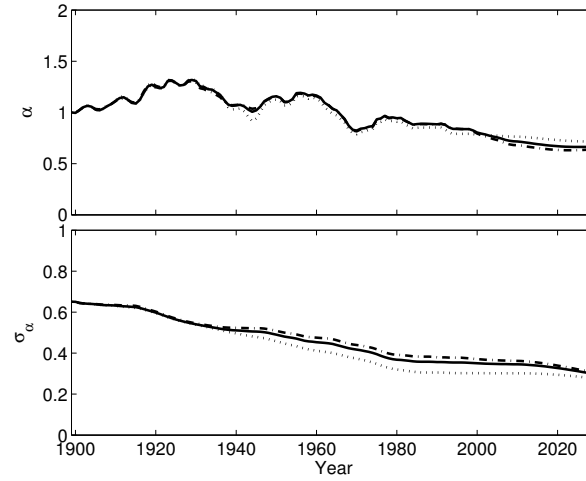


Figure 3.15: Estimates of the mean (top) and standard deviation (bottom) of the forcing scale factor, α , for the same prior assumptions on uncertainty in λ as in figure 3.14.

3.7 Sensitivity to skew

In addition to testing the sensitivity to the spread of the prior, one may also question whether it is appropriate to restrict the nature of the prior distribution to Gaussian, which is a limitation of the Kalman filter. Although Roe and Armour (2011) make compelling arguments justifying the use of Gaussian feedback parameter distributions, we nevertheless explore the consequences for our own approach. As a preliminary step in understanding the influence of Gaussianity, we apply our method to the simple climate model where the sensitivity parameter is redefined in terms of an inverse, $\lambda = 1/\gamma$. We let the inverse sensitivity parameter, γ , have a Gaussian prior uncertainty, and estimate the evolution of this probability distribution keeping all other assumptions constant. This results in the following relationships between probability distributions, similar to what was derived in chapter 2 equation (2.6).

$$\begin{aligned}\Pr(T_{\text{TCS}}) &= \frac{1}{\sqrt{2\pi}\sigma_\gamma} \exp \left[-\frac{1}{2\sigma_\gamma^2} \left(\frac{T_{\text{TCS}}}{F_{2\text{CO}_2}} - \bar{\gamma} \right)^2 \right] \cdot \frac{1}{F_{2\text{CO}_2}} \\ \Pr(\lambda) &= \frac{1}{\sqrt{2\pi}\sigma_\gamma} \exp \left[-\frac{1}{2\sigma_\gamma^2} \left(\frac{1}{\lambda} - \bar{\gamma} \right)^2 \right] \cdot \frac{1}{\lambda^2}.\end{aligned}\tag{3.1}$$

With Gaussian γ -distributions, the corresponding λ -distributions will be skewed and the corresponding TCS distributions will no longer have skew, which is evident in figure 3.16 where the left panel shows λ and the right panel shows TCS. Notice that the TCS distributions do not have tails when they are computed from Gaussian γ .

The prior distributions of both λ and TCS using Gaussian λ versus Gaussian γ are starkly different. However, as observed data is assimilated, the distributions converge so that by 2008, they are quite similar, as shown by the improved overlap in the solid curves in figure 3.16 compared to the lack of overlap in the prior dashed curves. Also the overlap in TCS curves by 2008 is the best near the peak and toward higher temperature change. This is encouraging because arguably policy makers are most concerned about the upper bounds of climate change.

The evolution of TCS over time (figure 3.17) is computed from the revised sensitivity parameter just as we did in section 3.2 from the Gaussian λ . The figures are plotted on the same axes for easier comparison. The sharp peaks or spikes in the tail of the density in figure 3.4, especially

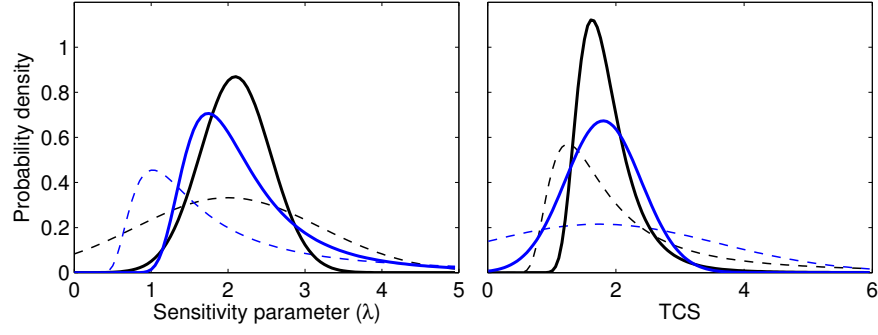


Figure 3.16: (left) Probability density distributions of sensitivity parameter λ . Distributions assuming Gaussian λ are black. Distributions assuming Gaussian inverse λ are blue. Prior distributions as prescribed in 1900 are dashed. Posterior distributions computed in 2008 are solid. (right) Same as left panel but for TCS.

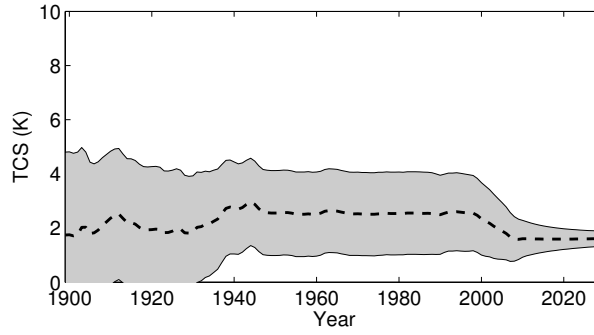


Figure 3.17: The same plot of TCS probability evolution over time as in figure 3.4 but where an inverse sensitivity parameter γ is the Gaussian unknown estimated by the Kalman filter.

around 1935-1945, are not present in figure 3.17. This confirms that the spikes are in fact an artifact of the earlier method for which λ was Gaussian. The maximum likelihood value (dashed line) still shows a perturbation around 1935-1945 because of a fluctuation in the observations then with comparatively small change in forcing, however, because the relationship between TCS and γ is linear, this does not lead to a fattening (or spiking) of the tail.

Also of note, the amount of ‘learning’ or reduction in the 90% confidence range of TCS that occurs between 2008 and 2030 is greater for the revised method, about 70% reduction compared to 45%, so the skew does impede uncertainty reduction. However, as the TCS distributions computed by either method grow narrower with additional observations, capturing the skewness becomes less important to characterizing the upper range of temperature change. With sufficient data, the second-order Kalman filter approach seems justified, therefore we do not see an immediate need to upgrade to higher order methods, such as particle filters.

3.8 Concluding remarks

The use of the nonlinear Kalman filter in conjunction with a semi-empirical model allows us to estimate the distribution of transient climate sensitivity, how the distribution explicitly depends on forcing uncertainty and natural variability, and how the distribution may change in the future as more data becomes available. The nature of the forcing uncertainty, whether scaled or not, as well as its magnitude, also affects our resulting probability distributions.

Although our estimates are certainly sensitive to these uncertainties and to natural variability, they may be sufficiently narrow as to still be useful. For uncertainties ranging from very large forcing uncertainty to very small forcing uncertainty, our confidence intervals for TCS range from 1.2–2.6 K to 1.4–2.6 K. With a much larger portion of the observed temperature change attributed to natural variability, our TCS interval increases to 1.1–5.5 K. Our probabilistic estimate of the range of TCS that we believe to be best justified by data, namely 1.3–2.6 K with a most probable estimate of 1.6 K, is broadly consistent with the TCR range of IPCC AR4 climate models whose median and mean are 1.6 K and 1.8 K, with 90% confidence interval of 1.2–2.4 K (Randall et al., 2007; Meehl et al., 2007).

To obtain much tighter estimates using methodology similar to ours would require both longer time-series and significantly reduced uncertainties in the forcing. Without that, attributing the observed temperature rise definitively to climate feedbacks or erroneous forcing is rather difficult. It may be possible to by-pass this difficulty by looking at the temperature increase of the two hemispheres separately or by looking at still more regional changes, as in Harvey and Kaufmann (2002) or Tomassini et al. (2009). The idea would be to increase the amount of constraining data while introducing as few new underdetermined parameters as possible.

The results we obtain are largely independent of those from comprehensive climate models, although we verify that our methodology works in part by comparison with a model (as well as the ability to reproduce the observed 20th century record), and the forcings shown in figure 3.2 do to some degree involve model calculations.

Chapter 4

Empirical closure for higher order climate systems

To regain some of the physical details neglected by the globally-averaged climate model of the previous chapters, we are interested in closing the parameters of more complex climate models. However, the unscented Kalman filter approach of the previous chapter does not scale well. Computational effort increases linearly with the size of the state, parameter, and observation spaces as more sigma points will be needed to adequately sample high dimensional error spaces. Generally, the more complex the model becomes, the more unknown parameters arise necessitating larger constraining observational datasets. For comprehensive GCMs that may have upwards of $O(10^5)$ dimensional state-spaces, the number of sigma points required makes the UKF method unfeasible.

However, more efficient nonlinear filter methods do not include many of the features of the UKF that make it attractive to estimation for climate systems. For example, the widely used Extended Kalman Filter (EKF) (Gelb et al., 1974) approximates the dynamic and measurement equations with a linearized system model so that the traditional Kalman Filter (Kalman, 1960) equations can be applied. The EKF presents two main difficulties: the tangent linear models may be difficult to derive and the first order approximation of the errors between the true and estimated states may lead to poor filter performance and even divergence of the filter.

The UKF, on the other hand, is a member of the growing family of alternative filters for nonlinear systems that we refer to as derivativeless, ensemble methods. The error covariances for

these methods are approximated by an evolving ensemble of states propagated forward with full nonlinear equations rather than a single state and linearized system equations. Derivativeless, ensemble filters are generally easier to implement than the EKF, as one does not require a tangent to the nonlinear map describing the dynamics and measurements. The UKF is particularly useful for strongly nonlinear systems as it has provable second-order accuracy in approximating the error covariances.

Several solutions to the derivativeless, ensemble filtering problem for large systems have been proposed and these generally fall into two categories: reduced models, for example Farrell and Ioannou (2001); Morelande and Ristic (2006); Lu et al. (2007), and reduced filters, such as Heemink et al. (2001); Uzunoglu et al. (2007); Lermusiaux and Robinson (1999); Ambadan and Tang (2009). The goal of reduced models is to approximate the true system equations with a lower dimensional model that can be filtered by the usual techniques. On the other hand, the goal of reduced-filter methods is to span almost all of the error space with fewer vectors in the most uncertain directions while preserving the dynamics of the full dimensional model.

Of the reduced-filter approaches, none of the methods have been developed generally enough to handle systems in which the dynamic and measurement equations are both nonlinear and process and measurement noise are not purely linearly additive. Some of the methods have additional shortcomings compared to the UKF. For example, the Ensemble Kalman Filter (EnKF) (Evensen, 2003) is a derivativeless ensemble method in which ensemble members are selected by Monte Carlo sampling. For the basic EnKF, the ensemble is not recomputed every filter iteration as in the UKF, so the rank of the error covariance is prone to collapse for undersized ensembles. In the UKF, it is possible for the number of dimensions spanned by the sigma point ensemble to grow at every iteration of the filter which aids in the prevention of diminishing error covariance rank.

EnKF methods using very small ensembles have had some success for state and parameter estimation in atmosphere general-circulation models (Annan et al., 2005b) when used in conjunction with the techniques of *inflation* and *covariance localisation*, designed to reduce the risk of filter divergence due to sampling errors. Inflation rescales the ensemble of points by increasing their distance from the mean (Evensen, 2007), though there is no clear method for determining the optimal scale factor. Localisation increases the effective rank of the error covariance by discarding unrealistic cross-covariances between distant grid points in the assimilation (Hamill, 2006). Both

techniques increase the complexity of the filter. Additionally, EnKF methods are not as easily implemented for systems in which both the state and measurement noise are nonlinearly related through the measurement operator.

Thus it would be beneficial to develop a computationally tractable filter based on the original UKF that preserves many of its desirable features. Luo and Moroz (2009) started in this direction by incorporating the unscented transform into the EnKF and using a truncated singular value decomposition to reduce the number of sigma points. However their approach still requires the storage of large-dimensional covariance matrices. The error subspace statistical estimation (ESSE) technique of Lermusiaux and Robinson (1999) avoids large matrix computations by using covariance update equations in factored form. Recent work by Ambadan and Tang (2009) applied ESSE to the UKF for linear measurement operators. Our work presented here may be viewed as extending the ESSE approach for UKF to include nonlinear measurements.

In this chapter, we introduce an efficient, easily implemented, deterministic filter technique for large systems. We take advantage of strengths of the UKF: high accuracy, simple application to nonlinear noise and nonlinear operators, derivative-free implementation, and we modify the traditional equations to allow for adaptive rank in the error covariance. Through this reduction, our goal is to make the computational expense of the UKF comparable to that of the EnKF without losing the advantages of the UKF. The chapter is organized as follows. Section 4.1 outlines the development of the modified filter. Sections 4.2-4.4 demonstrate the approach for both state and parameter estimation on a one-dimensional atmospheric model known as the Lorenz 96 model.

4.1 Adaptive-Covariance-Rank Unscented Kalman Filter

The goal of modifying the full-rank UKF (presented in chapter 2 as algorithm 1) is to improve the computational efficiency of the algorithm with minimal loss in accuracy. For high-dimensional systems, the dominant costs come from the $2(L_x + L_w) + 1$ unique forward model integrations of step 2(b)i in the original algorithm and the storage in memory and computation of the $L_x \times L_x$ covariance matrix of step 2(b)ii and 2(c)iii. Adapting the linear method of Lermusiaux and Robinson (1999) to general nonlinear systems, our approach reduces the cost of computation in two ways. First, we decrease the number of sigma points, and therefore the covariance rank, by keeping only those

points that span the dominant eigenvectors of the error space. Second, we directly update a factored form of P_k^x avoiding manipulations of the full sized matrix. We call the approach *adaptive* rather than *reduced* because there is an opportunity for the covariance rank to grow again when the sigma point ensemble is reformed by symmetrically perturbing the mean sigma point and forecasting all points by the nonlinear dynamics. The rank grows and decays with the number of dominant error directions found in the forecast sigma points. The mathematical details of the approach follow.

4.1.1 Adaptive covariance rank

In the original formulation of the UKF, the $L_x \times L_x$ error covariance matrix scales with the dimension of the system, which in the case of the climate system may be greater than $O(10^5)$, however, not all of the error dimensions may be crucial for accurate state and parameter estimation. Recall from algorithm 1 equation (2.11) that the UKF error covariance of the forecast state is defined,

$$P_k^{x^f} = A_k \Omega_{c,k} A_k^T,$$

where A_k is the matrix defining how the sigma points are distributed about the mean state estimate,

$$A_k = \chi_k^f - x_k^f,$$

(see algorithm 1 step 2(b)i). A factorization of the error covariance into its eigenvalues and vectors, $P_k^{x^f} = U_k \Sigma_k^2 U_k^T$, reveals the error dimensions that have the greatest variability (those corresponding to the largest eigenvalues) down to the least variable (smallest eigenvalues). For many systems, the majority of the variance is contained in a fraction of the error dimensions. Those dimensions without much variance may be ignored without substantially altering $P_k^{x^f}$. The following is a procedure for systematically eliminating extraneous error dimensions, thereby reducing the rank of $P_k^{x^f}$.

The singular value decomposition of the perturbation matrix is formed, $A_k = U_k \Sigma_k V_k^T$. A reduced-rank approximation of $P_k^{x^f}$ is obtained by retaining the first p_k singular values, σ_j , such that the ratio of singular values,

$$\rho = \frac{\sum_{j=1}^{p_k} \sigma_j}{\sum_{j=1}^{p_{k-1}} \sigma_j}, \quad (4.1)$$

is less than one, representing a reduction in the total variance of $P_k^{x^f}$. The choice of ρ should be tailored to each application, however a recommended starting point is $\rho = 0.99$. The dimensions of U_k , Σ_k , and V_k are truncated to form the reduced matrices U_{rk} , Σ_{rk} , and V_{rk} which have dimension $L_x \times p_k$, $p_k \times p_k$, and $2(p_k + L_w + L_v) + 1 \times p_k$, respectively.

The perturbation matrix is now approximated by

$$A_k \approx U_{rk} \Sigma_{rk} V_{rk}^T, \quad (4.2)$$

so that with substitution into (2.11), the reduced-rank approximation of the state error covariance is $P_k^{x^f} \approx U_{rk} \Sigma_{rk} V_{rk}^T \Omega_{c,k} V_{rk} \Sigma_{rk} U_{rk}^T$.

The rank of R_k^w and R_k^v are also reduced separately by similarly truncating their respective singular values at p_k^w and p_k^v . If these matrices are constant this may be computed offline as part of the initialization routine. We then use U_{rk}^w , Σ_{rk}^w , and U_{rk}^v , Σ_{rk}^v when forming the sigma point sets.

The rank reduction described above is one piece of the adaptation of the error covariance. The rank also has potential to increase at every iteration in nonlinear systems because the sigma points (which number twice the rank) may be forecast into a distribution spanning more dimensions than at the previous iteration.

4.1.2 Measurement update in factored form

To find U_{rk} and Σ_{rk} after analysis, we begin by rewriting the covariance analysis update equation (2.16) in terms of the other known covariances,

$$P_k^{x^a} = P_k^{x^f} - P_k^{xy} (P_k^{yy})^{-1} (P_k^{xy})^T. \quad (4.3)$$

Being symmetric, $P_k^{x^a}$ has an eigenvalue decomposition $P_k^{x^a} = U_{rk}^a (\Sigma_{rk}^a)^2 (U_{rk}^a)^T$. Substituting the UKF definition of P_k^{xy} and P_k^{yy} from algorithm 1 (equations (2.12)–(2.13)) and then inserting the

approximation (4.2) gives

$$\begin{aligned}
U_{rk}^a (\Sigma_{rk}^a)^2 (U_{rk}^a)^T &= A_k \Omega_{c,k} A_k^T - A_k \Omega_{c,k} Y_k^T (P_k^{yy})^{-1} Y_k \Omega_{c,k} A_k^T \\
&\approx U_{rk}^f \Sigma_{rk}^f (V_{rk}^f)^T \\
&\quad \cdot [\Omega_{c,k} - \Omega_{c,k} Y_k^T (P_k^{yy})^{-1} Y_k \Omega_{c,k}] \\
&\quad \cdot V_{rk}^f \Sigma_{rk}^f (U_{rk}^f)^T \\
&= U_{rk}^f \Pi_k (U_{rk}^f)^T \\
&= U_{rk}^f H_k \Lambda_k H_k^T (U_{rk}^f)^T,
\end{aligned} \tag{4.4}$$

where in the final step we have replaced Π_k with its eigenvalue decomposition since it is symmetric as in Lermusiaux and Robinson (1999). Thus we arrive at the update equations

$$U_{rk}^a = U_{rk}^f H_k \tag{4.5}$$

$$\Sigma_{rk}^a = \sqrt{\Lambda_k}, \tag{4.6}$$

where according to equation 4.4, H_k and Λ_k are computed by eigenvalue decomposition of the $p_k \times p_k$ matrix $\Pi = \Sigma_{rk}^f (V_{rk}^f)^T [\Omega_{c,k} - \Omega_{c,k} Y_k^T (P_k^{yy})^{-1} Y_k \Omega_{c,k}] V_{rk}^f \Sigma_{rk}^f$.

On the next iteration of the algorithm, there will be $2L_{rk} + 1$ sigma points, where $L_{rk} = p_k + p_k^w + p_k^v$, given by

$$\begin{aligned}
\chi_k^{(0)} &= x_{k-1}^{aug} \\
\chi_k^{(i)} &= x_{k-1}^{aug} + \sqrt{(L_{rk} + \lambda)} S_k^{(i)}, \quad i = 1, \dots, L_{rk} \\
\chi_k^{(i+L_{rk})} &= x_{k-1}^{aug} - \sqrt{(L_{rk} + \lambda)} S_k^{(i)}, \quad i = 1, \dots, L_{rk}
\end{aligned} \tag{4.7}$$

with corresponding weights defined

$$\begin{aligned}
\omega_{c,k}^{(0)} &= \frac{\lambda}{L_{rk} + \lambda} + 1 - \alpha^2 + \beta \\
\omega_{c,k}^{(i)} &= \frac{1}{2(L_{rk} + \lambda)} \quad i = 1, \dots, 2L_{rk}.
\end{aligned} \tag{4.8}$$

Mean weights, $\omega_{m,k}^{(i)}$, are identical to the covariance weights except that the zero-weight is $\omega_{m,k}^{(0)} = \lambda/(L_{rk} + \lambda)$. The weights are the same as those for the UKF defined in (2.8) but with L replaced

by L_{rk} , see section 2.3.2 for more information. The $S_k^{(i)}$ are the i columns of the block-diagonal $L \times L_{rk}$ dimensional augmented covariance matrix square root

$$S_k = \begin{bmatrix} U_{rk}^a \Sigma_{rk}^a & 0 & 0 \\ 0 & U_{rk}^w \Sigma_{rk}^w & 0 \\ 0 & 0 & U_{rk}^v \Sigma_{rk}^v \end{bmatrix} \quad (4.9)$$

Note that we need only keep track of the U_{rk} and Σ_{rk} in creating the next set of sigma points, and in practice we never need to reconstruct the full sized P_k^x . Algorithm 2 summarizes the steps for the adaptive-rank UKF.

Algorithm 2 Adaptive-covariance-rank unscented Kalman filter

1. Initialize P_0^{aug} and $x_0^{aug} = [x_0^a \ 0 \ 0]^T$
2. For $k = 1$ to ∞
 - (a) Calculate sigma points and weights as in (4.7)-(4.8).
 - (b) *Forecast or Dynamic Update*
 - i. Forecast each sigma point by $\chi_k^f = f(\chi_k)$. (Note that because it is assumed the process noise does not have dynamics, χ_k^f is $L_x \times 2L_{rk} + 1$, whereas χ_k , which includes perturbations to process and measurement noise, is $L \times 2L_{rk} + 1$.)
 - ii. Compute mean forecast,

$$x_k^f = E[\chi_k^f] = \sum_{i=0}^{2L_{rk}} \omega_{m,k}^{(i)} \chi_k^{(i)f}$$

as in (2.10) and form $A_k = \chi_k^f - x_k^f$.

- iii. Compute the reduced-rank approximation of A_k , R_k^w , and R_k^v as in (4.1)-(4.2) to find U_{rk}^f , Σ_{rk}^f , V_{rk}^f , and the reduced noise matrices.
- (c) *Analysis or Measurement Update*
 - i. Map the forecast sigma points in measurement space by $h_k(\chi_k^f, \chi_k)$ and form $Y_k = h_k(\chi_k^f, \chi_k) - E[h_k(\chi_k^f, \chi_k)]$.
 - ii. Compute the cross and measurement covariances and weighting matrix,

$$P_k^{xy} = A_k \Omega_{c,k} Y_k^T \tag{4.10}$$

$$P_k^{yy} = Y_k \Omega_{c,k} Y_k^T \tag{4.11}$$

$$K_k = P_k^{xy} (P_k^{yy})^{-1}. \tag{4.12}$$

- iii. Compute the analysis $x_k^a = x_k^f + K_k(y_k^{obs} - E[h_k(\chi_k^f, \chi_k)])$ and factors U_{rk}^a and Σ_{rk}^a according to (4.5), and (4.6).
 - iv. Form the block-diagonal augmented covariance root S_k^a by equation (4.9) and augmented state $x_k^{aug} = [x_k^a \ 0 \ 0]^T$ for the next iteration.
-

4.2 An abstract atmospheric test model

We illustrate the performance of the adaptive algorithm on a one-dimensional abstract model of the atmosphere derived by Lorenz (2006), from here on referred to as the L96 system. We chose this system of J coupled ODEs because the dimension is easily scaled to the higher dimensions for which the benefit of using an adaptive filter becomes evident. Additionally, the system is loosely representative of the physical behavior found in realistic atmospheric models in that it evolves chaotically, conserves total energy, and has dissipating effects, external forcing, and nonlinear interaction terms. The evolution equations are below.

$$\frac{dx_j}{dt} = (x_{j+1} - x_{j-2})x_{j-1} - \lambda x_j + F$$

The system is periodic at the endpoints such that $x_{j+J} = x_{j-J} = x_j$. The x_j can be thought of as a scalar atmospheric quantity, like temperature, at discrete longitudes, j , dispersed about a single latitude ring with periodic boundary conditions. The L96 system is discretized in time with a time step of 0.05 normalized time units using a fourth-order Runge-Kutta integration scheme. The same discretization is used in the filter's forecast simulations and in a truth run (shown in figure 4.1 (top)) used to generate the observed data. For both the filter simulations and the truth run, we set the forcing to $F = 8$, a value for which the system behaves chaotically. The transfer coefficient has been fixed at $\lambda = 1$. Process noise is introduced in the truth run by adding zero mean Gaussian noise w with standard deviation 0.01 to all J state equations at each integration time step.

4.3 State estimates

We begin by evaluating the adaptive filter performance for state estimation. To demonstrate the filter performance when the observations are a nonlinear function of both the state and the measurement noise, we add Gaussian noise with 0.1 standard deviation to the truth run and then square the noisy data. We then sample this field at M random grid points every 0.1 time units (twice the integration time step) to simulate the sparse, time-varying availability of climate observations. The measurement operation thin is

$$y_m = (x_m + v_m)^2,$$

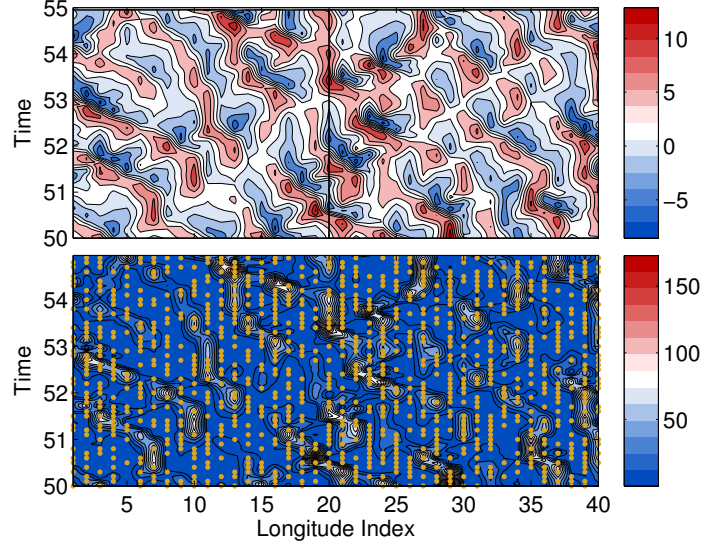


Figure 4.1: (top) True state of the L96 system at 40 discrete longitudes for a random initial condition. (bottom) Observation of the L96 system. Contours are the square of the state shown at top. Dots indicate the location of nonlinear observations made in space and time, drawn at random each time step.

for $m = 1 \dots M$ where $1 \leq M \leq J$. Moving sensors are a common feature of climate-observing systems, which rely heavily on adaptive sensor networks. The filter algorithm easily accounts for the variable measurement operator. Figure 4.1 (bottom) shows a segment of the true system evolution as well as the observed noisy squared field. Here the locations of each point in space and time at which a measurement is sampled are shown as dots.

In the simple example presented here, $L_x = L_w = J = 40$, but may be increased in future experiments, and $L_v = M = 20$. Thus the overall dimension of the system is $L = J + J + M = 100$, for a total of 201 sigma points. While the standard deviations of the true process and measurement noises are 0.01 and 0.1, respectively, it is unlikely these values would be known precisely in reality so the estimates provided to the filter in R^w and R^v are overestimated by double as 0.02 and 0.2. The initial state covariance has a variance of $4.0\text{e-}4$ along the diagonal.

We run the assimilation for 1000 iterations, or equivalently, 100 time units. For clarity in the figures, output is shown for a smaller representative time window. Figure 4.2 shows the time evolution of the full-rank UKF and adaptive-rank UKF state estimates in the top and bottom panels respectively. The two estimates preserve all of the qualitative features of the true state and deviate from it with similar errors as evidenced by comparison of percent errors in figure 4.3. The

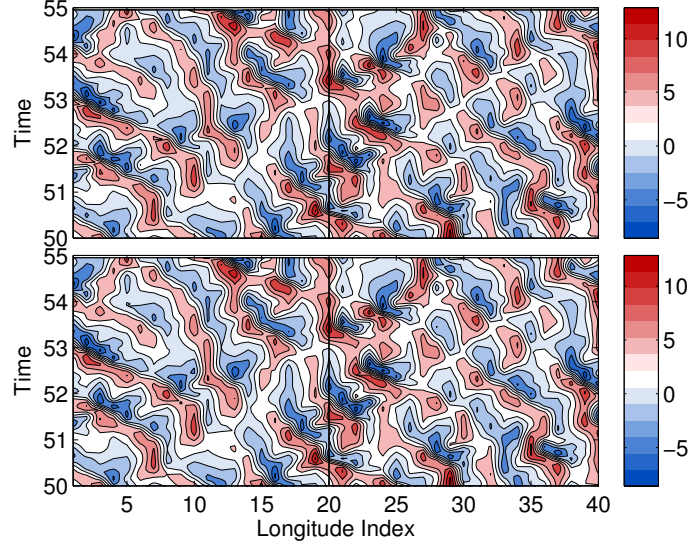


Figure 4.2: (top) The state estimated with a full-rank UKF. (bottom) The state estimated by an adaptive-rank UKF. The vertical line corresponds to the latitude of the output shown in figure 4.5

greater average error for the adaptive algorithm is explained by the differences in the estimates of the error covariance matrix. After an initial transient, the adaptive filter underestimates the magnitude of the singular values of the error covariance by about 50%, as evidenced by the fraction of total singular values captured using the adaptive covariance plotted in figure 4.4 (bottom). The algorithm is designed to keep only the variance contribution of the leading singular values at every iteration. If the forecast sigma points do not significantly expand the error dimensions in the future iterations, the fraction of singular values captured could be driven to zero. For this reason, we set a lower limit on the state-covariance rank at $p_{min} = 16$ to prevent rank collapse from which the covariance would be unrecoverable. In this experiment, we set the truncation threshold in equation 4.1 at 99.9% for the state variance, 80% for the process variance, and 99.9% for the measurement variance.

Figure 4.4 (top) shows the evolution of the rank of the augmented error covariance. Initialized with full rank, $L = 100$, the rank quickly diminishes to values oscillating around $L_r = 70$. The rank is seen to saturate at $L_r = p + p^w + p^v = 68$ which is consistent with a lower limit $p_{min} = 16$ because for the constant process and measurement covariances truncated at the 80th and 99.9th percentiles, their respective ranks will always be $p^w = 32$ and $p^v = 20$.

Evidently in this case, L_r is dominated by the ranks of the process and measurement noise

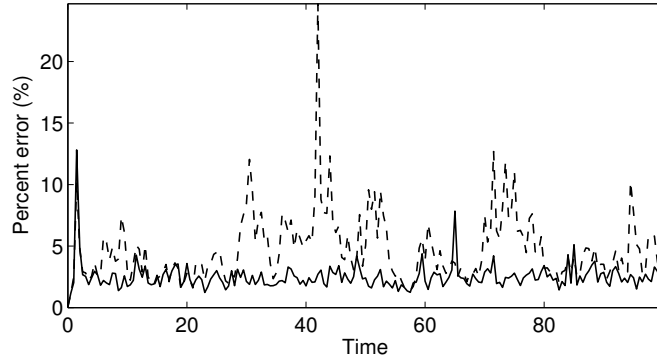


Figure 4.3: Comparison of the percent error averaged over all longitudes for the full- (solid) and adaptive-rank (dash) algorithms.

covariances, p^w and p^v . Most of the improvement in the overall rank is due to the decrease in p . This low effective dimension of the error subspace corresponding to the state variables may indicate potential to further reduce the rank of the error subspace corresponding to the process noise variables. Designing alternative strategies to minimize the noise subspaces is an interesting area for future work.

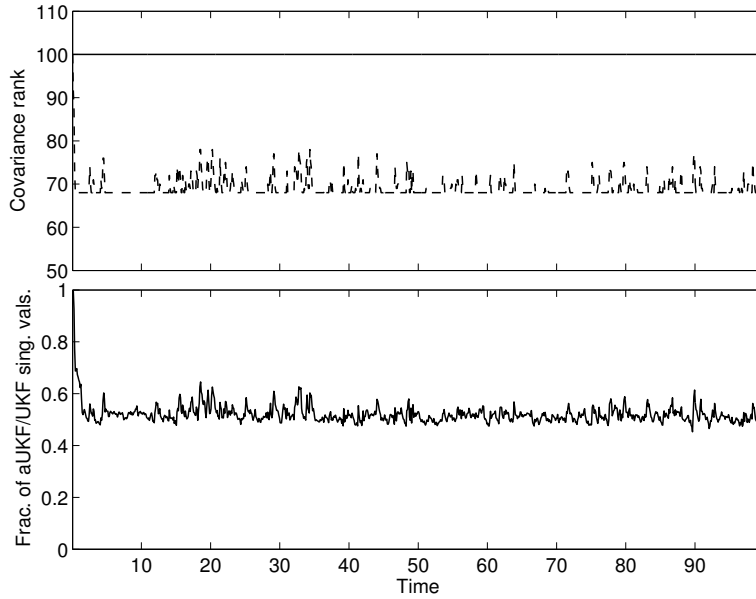


Figure 4.4: (top) The augmented covariance rank in the full (solid) and adaptive-rank (dash) UKF. (bottom) Comparison of the fraction of singular values from the UKF covariance that are present in the covariance of the adaptive algorithm.

For an alternate representation of the estimate error, the cross section at the 20th latitude index is shown in figure 4.5 for the true, full, and adaptive estimates along with the square root of

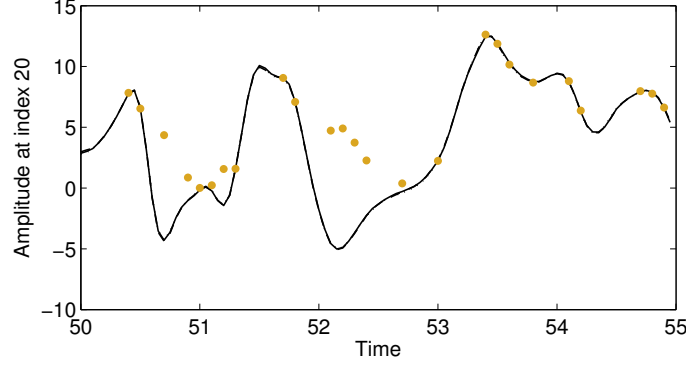


Figure 4.5: Cross-section of the true state at latitude 20 as well as estimates by both the full and adaptive filters (solid). True, full and adaptive curves lie on top of each other. (bottom) Square-root of the observations made at this latitude (dot).

the sparse nonlinear observations indicated by dots. The curves lie nearly on top of one another. Note that with the square measurement function, there are no negative observations, yet the state estimate remains accurate for the 100 time unit duration. For the marginal sacrifice in accuracy, the adaptive filter achieves runtime improvements of about 40% over the full UKF because it consistently forecasts the system at about 40% fewer unique sigma points.

4.4 Parameter estimates

The adaptive filter is similarly successful at joint state-parameter estimation and thus may be used to estimate closure parameters, including their uncertainty bounds, as was done in chapters 2 and 3 using the more computationally intensive full-rank UKF. For consistency with the experiments of these chapters, the closure parameter estimated for the L96 system is also a transfer coefficient λ like the climate sensitivity parameter of the energy balance model in equations (2.3)–(2.4) via state-space augmentation. The dynamic update for the parameter is likewise the persistence model. The experiment is conducted using the same observations and noise statistics applied for the state estimates. The simulations are initialized as the state at time 50 and a guess of $\lambda = 3$, with prior uncertainty, $\sigma_\lambda = \sqrt{10}$.

Comparison of the full and adaptive filters in figure 4.6 shows that the two approaches are almost identically successful at predicting the value of λ . The convergence of both algorithms onto the correct λ with very high certainty after just two assimilation steps is due to the lack of

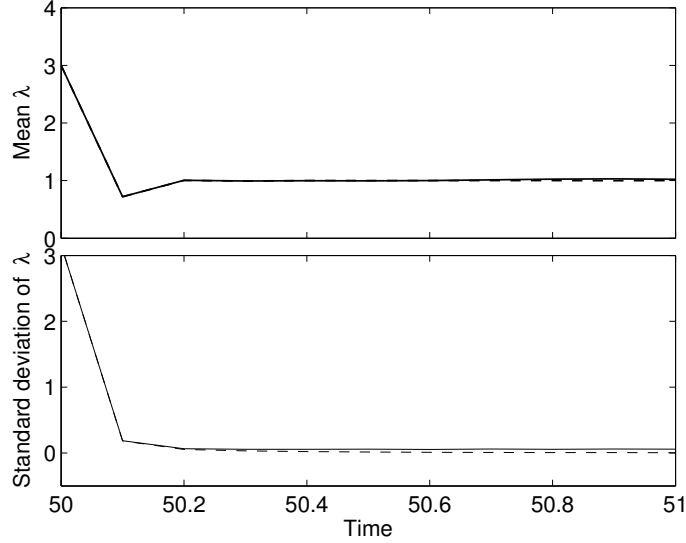


Figure 4.6: (top) Estimate of the parameter λ as noisy, nonlinear observations are assimilated over a 5 unit time period using the full-rank (solid) and the adaptive-rank (dash) UKF. (bottom) Corresponding uncertainty bounds (standard deviation) of λ for full-rank (solid) and adaptive-rank (dash) UKF.

other sources of uncertainty that were present in the climate model problem of chapters 2 and 3. This example has comparatively smaller unforced variability and no uncertainty in the external forcing F . Additionally, the L96 model has more states and the filter uses more observations of these states than the single temperature observation of the globally-averaged climate model. Although incorporating these additional uncertainties would be expected to slow the convergence and increase uncertainty about the parameter estimate, the performance of both filters would be degraded equally since they both capture the leading uncertainties equally well.

There is no evidence of the destabilization of the system from multiplicative parameter estimation that Yang and Delsole (2009) encountered with the L96 system using an ensemble square root filter (Whitaker and Hamill, 2002), although their application included more unknown parameters. For more complex estimation problems, should members of the sigma-point ensemble become divergent, the smoothed parameter update from Yang and Delsole (2009) could directly replace the persistence model without any modification to our adaptive algorithm needed.

4.5 Limitations

The computational savings using the adaptive method was about 40% in this general example of a nonlinear system where the state and process noise and observations and measurement noise were nonlinearly related. For special cases, for example with linearly additive process noise, additional efficiency may be achieved because forecasts of the sigma points corresponding to process noise perturbations may be neglected. In our example, this would have resulted in $2p+1$ unique forecasts compared to the $2(p+p_w)+1$ unique forecast simulations that were actually computed, for a savings of about 60%.

Although these savings are substantial, they do not achieve the order of magnitude reductions needed to filter $O(10^5)$ comprehensive global climate models. For these complex systems, even a 90% reduction in the number of sigma point forecasts may not be enough. Even a 99.9% reduction would still require 100 simulations of a GCM, a formidable task for climate-relevant time periods. This extreme loss of covariance information would lead to rank collapse in the adaptive UKF. To stabilize the filter, additional approximations may be applied, namely the techniques of inflation and covariance localization, originally developed for the EnKF. However given the heuristic nature of these techniques and the severe reduction in the sigma point ensemble size, there is no reason to expect the adaptive UKF under these conditions to have any advantages over the random sampling of the ensemble Kalman filter.

4.6 Concluding remarks

We have presented an efficient algorithm for nonlinear state estimation of high dimensional systems using a adaptive-rank unscented Kalman filter. The cost savings of the algorithm are two-fold, in that the number of sigma points is adapted to capture only the dominant state and noise variances, and the covariance information is propagated in a factored form.

We demonstrated the concept of the algorithm with successful performance on the L96 system, an abstract, chaotic model of the atmosphere. The computational savings of the adaptive method were significant and the method represents a reasonable approach to parameter estimation in climate models that incorporate more realistic physical processes and greater spatial resolution than the globally-averaged climate system introduced previously in chapter 2.

However, it may be that the advantages are specific to the system in question. For example, in a system where the errors do not collapse onto a significantly smaller subspace, to achieve significant improvement in computational cost, more than an acceptable amount of accuracy would be sacrificed. Also, for very large systems requiring severe rank reduction, the use of the adaptive filter becomes less obviously advantageous over previously developed methods like the EnKF. Another difficulty arises in obtaining enough observations of the climate system to constrain very large state and parameter spaces.

Limitations of observationally-based filter techniques for very large systems motivate us to explore the opposite end of the spectrum of closure techniques. Leaving observational methods for now, the remainder of this thesis evaluates the complementary (rather than competing) physically-based closure technique known as multiscale modeling.

Chapter 5

Dynamical closure of a momentum-moisture-like model

Purely empirical methods for closure of parameterized climate models are limited by the computational expense of the large number of simulations required for systematic uncertainty evaluation. They are also limited by the quality and availability of constraining data, especially as the number of unknowns increases with the complexity of the physical detail in the model. In contrast to the empirical approaches examined thus far, in this chapter, we analyze an emerging approach to parameterization that leverages physical understanding of the climate system. This approach, initially known as Cloud-Resolving Convection Parameterization (Grabowski and Smolarkiewicz, 1999) and Superparameterization (Khairoutdinov and Randall, 2001), as well as more recently the Multiscale Modeling Framework (MMF) (Jung and Arakawa, 2005), has been applied to the parameterization of cloud processes in global climate models. MMF provides closure to climate models by simulating closure models that couple to a filtered or large-scale system. The closure model consists of physically-based evolution equations for sub-filter or small-scale processes that otherwise would be neglected in the filtered system.

Previous studies using global climate models are ambiguous as to the benefit of MMF, showing improved simulation of certain convective processes while actually making others worse. The ambiguity arises out of the complexity of applying the approach to high-dimensional, physically intricate climate systems (results are highly sensitive to the many design options (Blossey et al.,

2009; Marchand and Ackerman, 2010; DeMott et al., 2011)) and out of the difficulty of validating against a modeled or observed truth. It is important to clarify what can be gained from MMF because it comes at the high price of increased computation time by several orders of magnitude.

Our work applies MMF to a simple system that we have developed expressly for the purpose of evaluating the costs and benefits of MMF unencumbered by computational expense while preserving dynamical relevance to the climate system. We have developed a chaotic, deterministic abstract model of the coupling of momentum and moisture in the atmosphere that uses the Kuramoto-Sivashinsky equation as a source of turbulent-like momentum coupled to a moisture-like equation that includes cloud water formation and precipitation-like processes. Our system allows us to evaluate three of the main MMF approximations: (1) elimination of nonlinear interactions between filter- and sub-filter- scale variables in advection processes, (2) imposition of periodic boundary conditions on the closure model, and (3) reduction in the state-dimension of the closure model. Most importantly, the system can be used to evaluate these sources of MMF error independently. We also use the system to illustrate how error sources interact through feedback in the full coupling of filtered and closure-scale equations.

We begin by providing background information on MMF and an overview of equations and approximations. This is followed by detailed criteria that a simple model must meet to rigorously test the MMF approach, all in section 5.1. Next we describe the equations and key features of the test system that has been developed and how MMF has been applied to it, in section 5.2. Evaluation of the MMF error sources follows in the next chapter.

5.1 Multiscale modeling framework overview

The multiscale modeling framework (MMF) can be thought of succinctly as a reduced-dimension solution to closure problems that arise in nonlinear systems when they are filtered in space and/or time. Filtered systems are common in climate modeling where some of the small-scale processes may be too computationally burdensome to simulate directly. Low-pass filtering reduces the order of the equations, retaining the large-scale fields of interest. Filtered equations may then be solved numerically on a coarse grid to save cost. However, due to nonlinearities, sub-filter-scale processes may have an important effect on the evolution of the filter-scale fields. Leaving these effects un-

closed would lead to errors in the scales of interest.

In conventional parameterizations, sub-filter-scale effects have typically been modeled by empirical relationships to the filter-scale variables. For convection parameterizations, this process may involve distilling observations to a few idealized scenarios for boundary conditions and forcing. These scenarios are then modeled in detail over a small geographic extent using large eddy simulations, also called cloud resolving models (CRMs) (often employing the moist, anelastic, non-hydrostatic equations). Based on the CRM results, parametric relationships between filter-scale variables and sub-filter-scale effects are formed. However these relationships may not be good approximations outside the idealized scenarios small regions (Neggers et al. (2012)). Additionally, sub-filter-scale effects diagnosed from filtered fields have no memory or transient development independent of the filtered solution. Such closures only capture the homogeneous statistical effects of sub-filter-scale processes. When imposing scale separation at a filter level for which sub-filter-scale processes are not truly homogeneous, conventional parameterization may lead to biased filter-scale solutions.

In climate models, sub-filter-scale convective processes may have significant heterogeneous development and important interactions with other sub-filter-scale processes. One approach to improving simulation of convection in climate has been to resolve clouds more directly. The need for coarse cloud parameterization may be eliminated by uniformly increasing model resolution to a horizontal spacing of less than 10km, effectively extending the domain of a CRM to cover the entire globe. Brief simulations (on the order of weeks) resolving clouds throughout a global domain have been achieved at this resolution (Satoh et al., 2008), but remain too expensive to compute for the longer time periods relevant to climate change. In addition, even global cloud resolving models may have additional unresolved processes that need parameterization, therefore they do not preclude the need for improved closure techniques.

In contrast to conventional methods, MMF solves reduced-dimension versions of the physical equations that describe sub-filter-scale effects. The sub-filter equations of motion used in MMF efficiently permit development of heterogeneous, transient and local dynamics. In climate applications of MMF, instead of conventional parameterizations, CRMs are employed directly; one is nested in every grid-column of a coarse (order 100km horizontal spacing) climate model. Computational savings over a global CRM are accomplished by reducing the size of the nested model domains so

that they occupy only a fraction of the horizontal extent of the outer grid-column.

Despite difficulties evaluating MMF in complex systems, the trend in MMF experiments has been to apply the framework to increasingly computationally intensive and physically realistic systems. In the framework initially conceived by Grabowski et al. (1996), one CRM was driven by a vertical profile of large-scale conditions observed a priori in the GARP Atlantic Tropical Experiment (GATE). The coupling was designed using relaxation of momenta to ensure the small-scale dynamics would develop yet would not deviate severely from the known GATE observations. This technique was more like dynamical downscaling approaches since it lacked feedback to the large-scale system. Later, large-scale feedback was completed using an outer model covering a 400km square area coupled to 100 CRMs. (Grabowski and Smolarkiewicz, 1999; Grabowski, 2001). The first application of MMF to an atmospheric model with global coverage, the atmospheric component of the NCAR Community Climate System Model (CCSM), used over 8000 CRMs. (Khairoutdinov and Randall, 2001) Simulation of one model year cost two weeks of computer time using 64 374MHz processors. Tao et al. (2009) applied MMF to the NASA Goddard finite-volume atmospheric general circulation model - its over 12,000 cloud-resolving models require 365 hours on 384 processors per year of model simulation. Most recently, Stan et al. (2010) applied MMF to the fully coupled atmosphere-ocean CCSM.

In sharp contrast, our work explores the framework at the opposite end of the model-complexity spectrum. There has been progress in the development of simple systems for understanding MMF using stochastic closure models by Kerstein (1988) and Majda and Grote (2009). The contribution of this thesis is the first application of MMF to a simple, deterministic test model.

5.1.1 MMF equations and approximations

We now outline general equations for the framework and highlight three of its important approximations. We conclude by defining criteria for a model that would be successful in testing the significance of the errors in these approximations. The following derivation illustrates how MMF may be used to solve a closure problem in general nonlinear equations. It is based on the revised formulation of Grabowski (2004), the asymptotic derivation of Majda (2007) and the most recent clarification from Randall and Khairoutdinov (2010), all of which have sought to clarify the method and give it more sound mathematical footing over the original framework in Grabowski

and Smolarkiewicz (1999).

Consider the system of partial-differential equations (5.1) whose state is $v(x, t)$ and is comprised of advection terms, B , and other linear and nonlinear sources/sinks, S ,

$$v_t = B(v) + S(v). \quad (5.1)$$

The subscript t denotes the partial derivative with respect to time.

We define a filter operator $\overline{(\cdot)}$, which throughout this work is usually a box filter, although other filters like the sharp spectral filter could be applied – the main requirement is that the filter commutes with differentiation, such that

$$\begin{aligned} v(x, t) &= \bar{v}(X, t) + v'(x, t) \\ \overline{v'(x, t)} &= 0. \end{aligned} \quad (5.2)$$

Then applying the filter to equation (5.1), and substituting (5.2), we arrive at the filtered system (5.3) that depends on both \bar{v} and v' ,

$$\bar{v}_t = \overline{B(\bar{v} + v')} + \overline{S(\bar{v} + v')}. \quad (5.3)$$

The system must be closed with an expression for v' . Conventional closure parameterizations would ignore any small-scale effects in \bar{B} , letting $\overline{B(v)} \approx B(\bar{v})$, and would define an approximate function for sources and sinks, $\overline{S(v)} \approx \tilde{S}(\bar{v})$. The approximate system, $\bar{v}_t = B(\bar{v}) + \tilde{S}(\bar{v})$, now a function of only the filtered state, is solved for the large-scale behavior.

In contrast, MMF evaluates sub-filter-scale or closure equations – partial-differential equations with resolution fine enough to resolve many of the length scales in v' . The results of the closure equations are supplied as input to the filtered system. The sub-filter-scale or closure equation, without approximation, is,

$$v'_t = B(\bar{v} + v') + S(\bar{v} + v') - \bar{v}_t. \quad (5.4)$$

For computational tractability, key approximations are made to both the filtered equations and the closure equations. In both equations, it is common to neglect the advection terms related to

interactions between large and small scales as in, $B(v) \approx B(\bar{v}) + B(v')$. Interacting terms $B(\bar{v}, v')$ are considered small by an argument for scale separation, though whether an actual separation exists for the climate system has been debated. Omission of interacting terms simplifies the coupling between the filter and closure equations although it is not strictly necessary and as we will see in section 6.1 of the next chapter, it leads to substantial error in simulations of the test system.

For the closure equation, the main simplification is to model the sub-filter-scale behavior on periodic subdomains, one for each grid point in the discretized domain of the filtered system. Periodic boundary conditions are advantageous because they are computationally simple to implement, they allow for greatly parallelized processing, existing cloud-resolving models already use them, and the actual boundary conditions may not be known. The constraint of periodic boundary conditions means that all inputs to the subdomain models must also be periodic, thus the filter-scale variables are commonly approximated as constants across each subdomain.

The periodic boundary approximation, as will be shown in section 6.2, is a major source of error in MMF simulations. Periodic boundaries create a barrier between neighboring subdomains such that communication is only possible through interaction with filter-scale variables. Consequently, simulations will not converge to direct numerical simulation (DNS) with increasing resolution.

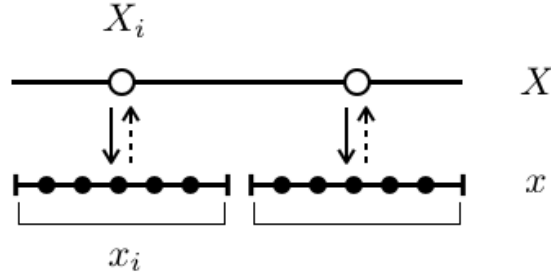


Figure 5.1: A segment of a one-dimensional MMF computational domain illustrating the coupling of filter- and closure-scale grids. Closed circles are closure-scale grid points and open circles denote filter-scale grid points. One closure-scale subdomain with periodic lateral boundary conditions communicates with each filter-scale grid point. Arrows point in direction of information flow, dashed indicates average, solid indicates actual values passed.

The coupling between the filter-scale domain and the closure-scale subdomains is illustrated in one spatial dimension in the schematic of figure 5.1. The filter-scale variables are discretized at the open-circle points on the large-scale domain, X , whereas the closure-scale variables are discretized at the closed-circle points on x . Each filter-scale grid point, X_i , is in communication

with the corresponding set of closure-scale grid points, x_i , that make up its subdomain. Within a subdomain information is shared between all points, however, outside the subdomain, only box-filtered (subdomain-average) values are passed to the filter-scale model and no information is passed to neighboring subdomains. Thus the closure-scale solution, $v'(x, t)$, is a piece-wise discontinuous function. In the example domain shown discretized with two filter-scale grid points, the filter-scale sources/sinks and advection would be given respectively by,

$$\begin{aligned}\overline{S(X_i, t)} &= \overline{S(\bar{v}(X_i, t) + v'(x_i, t))} \\ \overline{B(X_i, t)} &= \overline{B(\bar{v}(X_i, t))} + \overline{B(v'(x_i, t))}, \quad \text{for } i = 1, 2.\end{aligned}\tag{5.5}$$

Note that because net horizontal advective transport on a horizontally periodic domain will be zero, there will be no horizontal component of the closure-scale advection effects, $\overline{B(v'(x_i, t))}$. In climate modeling MMF applications, $\overline{B(v'(x_i, t))}$ would have a non-trivial vertical component. For our one-dimensional test system with a purely horizontal domain, $\overline{B(v'(x_i, t))} = 0$. In other words, for our system, MMF will not lead to any improvement in modeling advection; all of the benefit of MMF will be attributed to modeling sources and sinks, \bar{S} . For further discussion related to this see section 6.2.1.

The key approximation to achieving computational savings with MMF is the reduction in the dimension of the closure equation. This is typically done by reducing the dimension of the subdomains to something less than the dimension required to span an entire filter-scale grid-box. In comprehensive climate models, the subdomain dimensions have been reduced in a variety of ways, most commonly by eliminating one horizontal dimension to achieve planar (horizontal-vertical) subdomains (Grabowski, 2004). Others have experimented with three-dimensional subdomains that occupy just a fraction of the horizontal extent of a filter-scale grid column (Khairoutdinov et al., 2005; Xing et al., 2009). Another configuration used two planar subdomains at right angles per grid column (Randall et al., 2003). Our simple test model will allow us to investigate alternative approaches to dimension reduction besides shortening the spatial domain.

In summary, the three main approximations applied to MMF systems are as follows,

1. Terms involving filter-closure scale interactions are neglected in expressions for advective transport.

2. The boundaries of the closure models are periodic, therefore the filtered variables are constant in time and space over the closure model domain and information is not advected from subdomain to subdomain.
3. The dimension of the closure model is reduced in order to reduce computation time. Note that this is the most important step. The MMF system is not more efficient than solving the system at high resolution unless the closure model is made efficient by eliminating some dimensions.

The effects of these approximations may conveniently be studied in the test system that we have developed and that we introduce in section 5.2. The test system was designed specifically to address several of the difficulties in evaluating MMF effectiveness for more complex systems.

5.1.2 Criteria for successful test model

Perhaps the greatest difficulty in evaluating MMF is the lack of a comparable observational truth. Observational data is sparse; available data relevant to clouds and convection spans fairly short periods of time with respect to climate change. Additionally, observations may not map directly to climate variables. For example, most satellite retrievals assume that clouds occur in a single layer, whereas climate models have vertical resolution capable of simulating multiple cloud layers. To cope with this, MMF model output has been translated into comparable satellite retrievals by instrument simulators (Marchand and Ackerman, 2010). In lieu of satellite measurements, other studies have made use of reanalysis products or output from high resolution weather models that use data assimilation (DeMott et al., 2011; Pritchard et al., 2011). However these observations are not completely independent of model results.

Likewise, there does not exist a comparable simulated truth. Direct numerical simulation of cloud processes would require grid resolution down to micro-scales. Global CRMs are emerging, however they are too expensive to simulate for climate time scales. Most studies have compared their MMF results to climate simulations using conventional convection closure which are much less expensive simulations. Increasing the resolution of conventional models so the computation time is comparable to MMF is not easily done because the parameterizations do not take the grid size as an input. Existing parameterizations would need to be re-tuned or overhauled to accommodate

the change.

Additional features of global climate models make them difficult to use for unambiguous evaluation of MMF. Besides convection, many other processes are parameterized, also contributing to the overall error in climate models. Since parameterizations may be tuned to compensate each other, an improvement in convection parameterization due to MMF may lead to poorer performance in the system overall, making the overall benefit of MMF ambiguous. In physically detailed systems, MMF simulations are expensive, so the length and number of validation runs is limited, prohibiting complete assessment of the many MMF configurations. Also, results have been shown to be sensitive to the physics included in the CRMs, which may vary greatly (Tao et al., 2009).

In light of these difficulties, the simplified system most useful to studying MMF will have several key dynamical, physical, mathematical, and computational characteristics. First and foremost, with respect to dynamics, filtering the system equations should result in a closure problem that can be addressed by MMF that is similar to the closure problems found in climate models that are addressed by parameterizations. In other words, filtering should yield filter-scale equations that can be solved on a coarse grid to achieve computational savings over DNS. Since these filter-scale equations will have unclosed terms dependent on sub-filter-scale behavior, we should be able to derive physically consistent evolution equations to simulate sub-filter-scale realizations that will close the filtered system.

Physically, the system will behave similarly to just a single process found in climate models, for example the transport of moisture. This will allow us to attribute errors directly to one closure term instead of the many found in climate models.

Mathematically, the system should be easy to understand and manipulate analytically. Toward this end, the domain should occupy just one spatial dimension with a minimal number of component equations. It should be possible to compute eigenvalues analytically, at least for a linearized version of the system. Analysis of the energetics should be straightforward using Fourier transforms.

Finally, the computational burden of numerical integrations of the system should be minimized, permitting ensembles of MMF simulations for a range of run times. High-resolution direct numerical simulations should be quickly computed to provide the unequivocal truth simulations critically need for validation.

5.2 MMF test models

The first test model we considered that could meet the above criteria is known as the Kuramoto-Sivashinsky (KS) equation. However, for reasons explained below, the KS model alone turns out not to be a suitable test bed and motivates the development of our eventual model, the KS-moisture system, which includes the KS equation as a source of turbulent motion.

5.2.1 Kuramoto-Sivashinsky equation

The KS equation was an appealing first model because it shares several key properties with turbulent momentum in the atmosphere yet covers a simple one-dimensional geometry. Its solutions span a broad energy spectrum with energy continuously transferred between high and low wavenumbers so that accurate simulation of the system with small spatial scales filtered out would strongly depend on the evolution of sub-filter scales, requiring a closure model. Also, its dynamics can be easily understood in spectral space where critical wavenumbers and eigenvalues of the linear system are written algebraically. However, we found that although the K-S equation is physically relevant and computationally tractable for the study of MMF, it is not a system for which MMF can be applied independently from other closures necessary to prevent solutions to the equation from growing unbounded. This section describes the KS system and explains how its dynamical features make MMF insufficient at modeling the effect of small-scale behavior on the filtered equation.

The KS system was originally derived to model the propagation of waves in reaction-diffusion chemical systems (Kuramoto and Tsuzuki (1976)). It has also been used to describe instabilities of flame fronts in combustible mixtures (Sivashinsky (1977)) and thickness variations of a liquid film flowing down a vertical plane (Sivashinsky and Michelson (1980)). For our purposes, we think of (5.9) as a fluid dynamic equation and its solutions, $u(x, t)$, as zonal velocity or height perturbations of a thin fluid on a ring of latitude, x .

The system,

$$u_t + uu_x + u_{xx} + \nu u_{xxxx} = 0 \tag{5.6}$$

is solved on the periodic domain, $0 \leq x < 2\pi$. Subscripts denote partial differentiation.

The single parameter, ν , is a damping coefficient like viscosity or inverse Reynolds number (Re). The key processes of advection, energy production, and dissipation are expressed by the uu_x ,

u_{xx} , and νu_{xxxx} terms, respectively. The mixing provided by advection transfers energy from large length scales to small scales where it is dissipated. The length scale of dissipation is determined by the Reynolds number. In the absence of mixing, all solutions to the KS equation for $Re > 1$ would diverge to infinity.

The critical length scale separating energy production and dissipation may best be understood by examining the eigenvalues of just the linear part of the system in Fourier space,

$$\hat{u}(k, t)_t = (k^2 - \nu k^4) \hat{u}(k, t). \quad (5.7)$$

Here \hat{u} is the Fourier transform of u and a function of wavenumber, k , and time, t . From (5.7) it is evident we now have a diagonal system in which the Fourier modes are completely uncoupled and each mode's eigenvalue, $e_k = k^2 - \nu k^4$, indicates whether that mode makes an energy producing or dissipating contribution to the system as a whole. The all real eigenvalues of the linear system are plotted in figure 5.2 at different Reynolds numbers. We see that at low wavenumbers, the k^2 term dominates the νk^4 term leading to positive, unstable, energy-producing eigenvalues. As the wavenumber increases, the dissipative term, νk^4 becomes stronger, eventually overwhelming k^2 , making clear the reason for fourth order diffusion. At the critical wavenumber, $k_{crit} = 1/\sqrt{\nu}$, the production and dissipation are in perfect balance corresponding to a zero eigenvalue. Dissipation dominates for wavenumbers beyond k_{crit} with all of the eigenvalues being negative and hence stable. As Re increases, k_{crit} shifts toward the higher end of the spectrum, increasing the number of waves that produce energy.

The Fourier transformed system also helps us understand the shape of the energy spectrum. The eigenvalues are maximized at a peak wavenumber, $k_{peak} = 1/\sqrt{2\nu}$ which corresponds to the peak in the energy spectrum. The rapid decay of the eigenvalues beyond the critical wavenumber creates steep attenuation of energy in the spectrum. We illustrate the spectrum features for our particular choice of $Re = 100$ in figure 5.3. This value was selected to achieve the desired complexity of spatial structures.

Applying a spatial filter to the KS equation results in the filtered equation below with a clear closure problem,

$$\bar{u}_t + \overline{uu_x} + \bar{u}_{xx} + \nu \bar{u}_{xxxx} = 0. \quad (5.8)$$

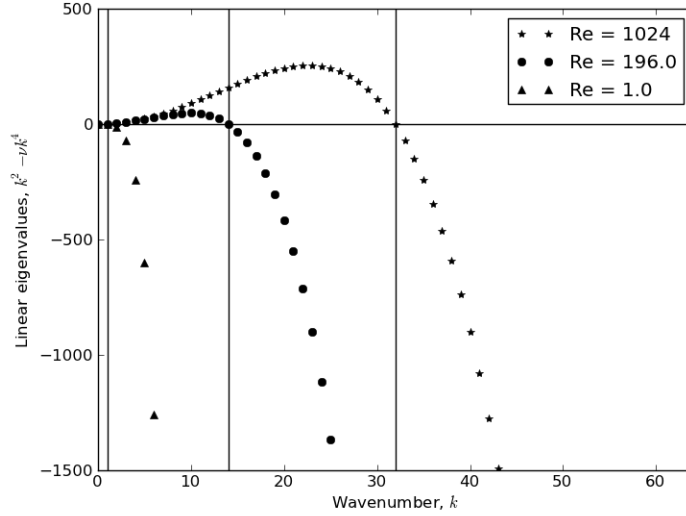


Figure 5.2: Eigenvalues of the linear system (5.7) for decreasing Reynolds numbers as indicated by the legend. The vertical lines indicate the critical wavenumber, $k_{crit} = \sqrt{Re}$, at which the eigenvalues cross over from unstable to stable.

Here the nonlinear term requiring closure provides the critical energy transfer needed to keep solutions bounded. Model this closure term incorrectly and energy production grows unbounded as it would in the unstable linearized system (5.7) causing solutions to diverge to infinity. Even though MMF can approximate the closure term with good statistical accuracy, the filtered system is extremely sensitive to its accuracy and subtle errors lead to unbounded solutions. In simulations where the closure term was computed ahead of time from DNS, the solution of the filtered equation grew to infinity after a short integration time due to rounding errors in the pre-computed term.

MMF does not come close to approximating the closure term to within rounding error for several reasons. First, sub-filter-scale advection is grossly limited by the use of periodic subdomains for the closure simulations. Nothing is advected between subdomains because the subdomains do not share boundaries. Next, were modifications made to MMF to alleviate the periodic subdomain limitation, closure errors would arise due to the omission of cross terms in the quadratic advection. Finally, were cross terms included, the simulations of sub-filter-scales would still remain approximate due to the reduced spatial extent of the subdomains to achieve cost savings.

In summary, the closure terms that MMF is capable of generating are too approximate to provide the energy transfer necessary to keep solutions of the filtered KS equation from diverging to infinity. The need for a more forgiving model motivates the slightly more complex but better

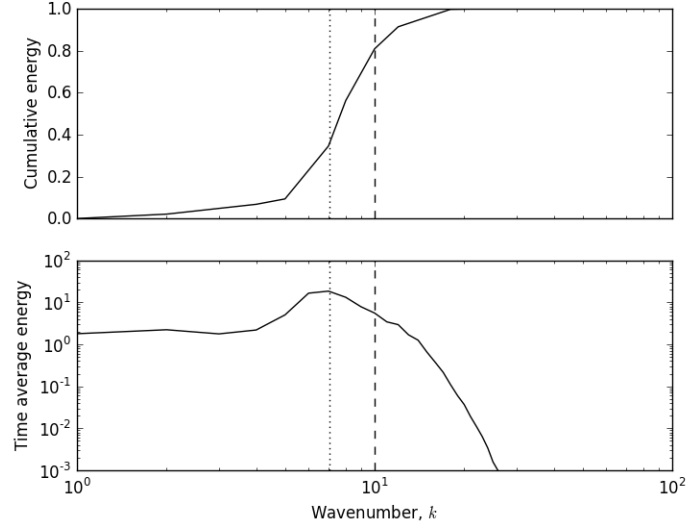


Figure 5.3: Time-average cumulative energy and energy distribution as a function of wavenumber for $Re = 100$. Vertical lines are k_{peak} (dot) and k_{crit} (dash).

test model that we now describe. The key requirement of the improved model is that the term(s) needing closure not be the determining factor in whether solutions remain bounded.

5.2.2 KS-moisture system

Given what we learned about MMF and the KS system, we designed the system described in this section to improve on the shortfalls of KS and meet the criteria for a successful test model outlined in section 5.1.2. Although the system is slightly more complex, the guiding design principle remains to seek a model of minimal complexity while retaining dynamics relevant for studying the consequences of MMF approximations. Our system now has two components, representing momentum and moisture, discretized in just one spatial-dimension to keep simulation costs to a minimum,

$$u_t + uu_x + u_{xx} + \nu u_{xxxx} = -\gamma q \quad (5.9)$$

$$q_t + (uq)_x + \mu q_{xxxx} = K(q)(\beta u), \quad (5.10)$$

on the periodic domain, $0 \leq x < 2\pi$. Subscripts denote partial differentiation.

We can think of q , as a cloud water mixing ratio and u as a horizontal velocity. The mois-

ture dynamics (5.10) include internal dissipation and two non-linearities that will lead to closure problems: advective transport and a moisture source with a variable time-scale.

The turbulent momentum equation (5.9), the KS equation with some feedback from the moisture, drives the moisture equation providing a source of high-wavenumber energy. Solutions to the KS equation are chaotic and have energy at a broad range of length scales with this energy peaking at a critical length and then attenuating as scales become finer as discussed in 5.2.1. The broader the energy spectra of solutions to KS, the broader will be the moisture spectra and the greater the need for closure of sub-filter-scale moisture dynamics.

Coupling KS to the moisture equation (5.10) ensures that a wide range of wavenumbers in q will be excited. As stated earlier in section 5.1.2, critical to the success of the test model is that accurate characterization of its low-wavenumber dynamics depends strongly on modeling the high-wavenumber behavior. A moisture response without any high-wavenumber energy would mean that the system could be accurately approximated at very coarse resolution, making small-scale closure unnecessary. We refer to the coupled test system (5.9)-(5.10) as KS-q from this point forward.

Also important to achieving the desired energy spectrum for the moisture dynamics is the strength of the diffusion coefficient μ . The stronger the diffusion, the greater the damping of higher order waves. We must choose a diffusion strength that will maximize the wavenumbers containing significant energy yet minimize the development of discontinuities due to wavebreaking. We want to maximize the width of the distribution so the need for closure is obvious.

The main closure problem to be addressed by MMF arises from the nonlinear moisture source term, the details of which are now explained. Depending on the relative strength of moisture and momentum in a given location, the term acts as either a source of moisture when $\beta u > 0$, which we liken to the formation of clouds, or as a moisture sink when $\beta u < 0$, which we give the physical significance of cloud depletion via moisture precipitating out of the atmosphere. Cloud formation may occur given any present amount of cloud moisture when momentum is positive, which we can think of as correlated to the strength of updrafts due to surface heating or to the strength of vertical wind shear. The opposite sign momentum results in cloud loss or rain. Conditions leading to the strongest cloud loss are high levels of existing clouds (positive q) combined with negative or downdrafting or shear free momentum.

To introduce nonlinearity, the rates of cloud formation and depletion are made asymmetric

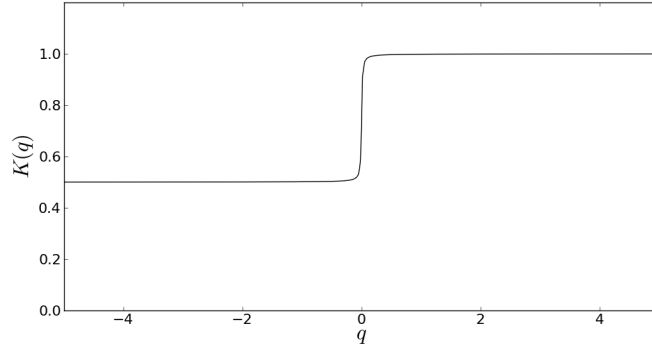


Figure 5.4: The scale factor, $K = \arctan(\alpha q) \times (K_p - K_e)/\pi + (K_p + K_e)/2$, is a function of the relative strengths of cloud moisture and momentum.

through the scale factor $K(q)$ which varies smoothly between two bounds, $K_e = 0.5$ and $K_p = 1.0$, depending on the magnitude and sign of q as shown in figure 5.4. This has the effect of attenuating the magnitude of the source or sink wherever q is negative. The parameter α is a scale factor that determines the sharpness of the transition between the two limits of K . We chose a rather high value, $\alpha = 100$ to increase the relative importance of q' compared to \bar{q} when q is close to the transition point at 0. With this value of α , very subtle fluctuations in q' become important to the accuracy of the solution; this creates a scenario where any benefit from MMF will be more obvious.

The moisture dynamics feed back to influence momentum through the $-\gamma q$ term in (5.9). This is consistent with the idea that cloud formation is associated with a vertical mixing of the atmosphere that will have a tendency to act as a drag force on horizontal momentum and dissipate vertical wind shear. The moisture source and the drag force terms of (5.9)-(5.10) lead to solutions in which u and q oscillate with respect to one another aperiodically in time.

Assuming for the moment that q and u don't vary in space, the KS- q system reduces to the classical coupled harmonic oscillator with complex-conjugate eigenvalues,

$$\lambda = \pm \sqrt{K\gamma\beta}i,$$

that vary in time with K . The nonlinear scaling of K causes the eigenvalue pair to oscillate between two limits on the imaginary axis, a high-frequency pair resonating at undamped natural frequency, $\omega_0^p = \sqrt{K_p\gamma\beta}$, when clouds are abundant and a low frequency pair corresponding to a smaller resonant frequency, $\omega_0^e = \sqrt{K_e\gamma\beta} = \sqrt{2}/2\omega_0^p$, in the absence of clouds. A possible physical

interpretation of these dynamics is that a lack of clouds indicates a more stable atmosphere in which the time scale of oscillations is slower. The coupling parameters, $\beta = 10$ and $\gamma = 5$, are chosen so that q and u are roughly the same order of magnitude and to achieve a desired $\omega_0^p \approx 2\pi$.

The oscillator dynamics are important for balancing time the moisture source spends with its mean above and below zero to prevent a steady drift of mean moisture. As strong precipitation causes q to grow negative, u increases, shifting the mean of $K(q)\beta u$ above zero and tipping the balance back toward cloud formation. Perhaps a more physically relevant model would replace u with u_x in the moisture source term where u_x makes a better proxy for the surface divergence that causes the updrafts and downdrafts leading to cloud formation and precipitation. However, it would destroy the coupling between u and q and additional terms in both momentum and moisture equations would be needed to maintain a balanced system in which q does not diverge to infinity. For simplicity, we opt to keep the original formulation.

Now we consider the effects of the nonlinearity for different degrees of spatial variation in q and u . The source term is designed so that its average value depends strongly on the spatial resolution of q (which will be modeled according to MMF). Spatial perturbations in βu are attenuated asymmetrically by the scale factor $K(q)$ and this has a significant effect on the magnitude of the domain-averaged moisture source, $(K(q)p)^*$, as is illustrated by figure 5.5. Here we have supposed that $p(x) = \beta u(x)$ is a linear combination of simple sine and cosine waves. In the left-hand plot, $q = \cos(x)$ has comparatively little spatial variance. It includes just one cosine mode at wavenumber one that could be modeled with coarse resolution as it would in a large-scale simulation without closure using MMF. On the other hand, in the right plot, $q = \cos(x) + \cos(3x)$ includes a higher order mode effect from $\cos(3x)$ that could be captured through MMF closure simulations. The average contribution of the moisture source term in the right plot is about half that of the left plot. This shows that the resolution of small-scale spatial fluctuations in q is necessary to accurately capture the amplitude of the moisture source. In general, we would expect the variance in \bar{q} to be too large without the attenuating effects of small-scale q .

A final feature of the KS-q system is the option to resolve the moisture and momentum processes at arbitrary length-scales independently, giving us the option to filter the moisture equation, which has been the main process targeted for MMF in climate, while retaining all scales in the KS equation which we showed in the previous section cannot be closed using MMF. Thus we have a

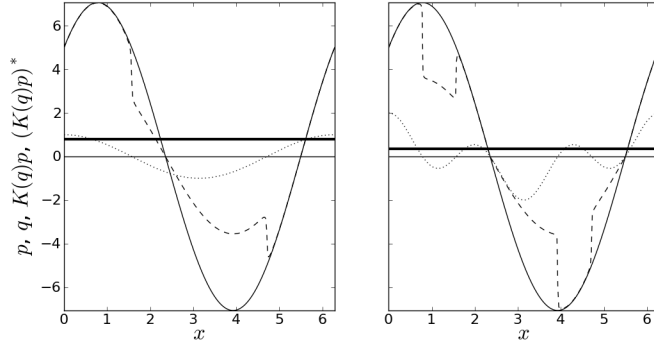


Figure 5.5: (left) The moisture source, $K(q)p$ (dash) and the mean moisture source, $(K(q)p)^*$ (bold) given that $p = \beta(.5 \sin x - .5 \cos x)$ (solid), and $q = \cos x$ (dot). (right) Same as for left but with $q = \cos x + \cos 3x$.

system where we can apply MMF to one process and not another to eliminate confounding sources of error. Such a feature would be complicated to implement in a full physics climate model. This process is outlined below.

5.2.3 Filtered moisture equation

We obtain the large-scale moisture equation (5.12) by applying the same assumptions and box filter from section 5.1.1 to equation (5.10), beginning by separating momentum and moisture into filtered and closure scales,

$$\begin{aligned} u(x, t) &= \bar{u}(X, t) + u'(x, t) \\ q(x, t) &= \bar{q}(X, t) + q'(x, t). \end{aligned} \tag{5.11}$$

Without approximation, the filtered moisture equation corresponding to equation (5.3) from section 5.1.1 is,

$$\bar{q}_t + \overline{(\bar{u}\bar{q})}_X + \mu \bar{q}_{XXX} = \overline{K(q)\beta u} - \overline{(\bar{u}q')}_X - \overline{(u'\bar{q})}_X - \overline{(u'q')}_X, \tag{5.12}$$

on the domain $0 \leq X < 2\pi$. All terms on the right hand side are functions of sub-filter-scale variables and therefore need closure. After neglecting filter-closure interaction terms in advection, $\overline{(\bar{u}q')}_X$ and $\overline{(u'\bar{q})}_X$ would be omitted from (5.12). This would leave $\overline{(\bar{u}\bar{q})}_X + \overline{(u'q')}_X$ which corresponds to the approximation, $\overline{B(\bar{v})} + \overline{B(v')}$, made in section 5.1.1.

The specific expressions needed to close the remaining terms on the right-hand side of (5.12) that correspond to the general equations (5.5) are,

$$\begin{aligned}\overline{K\beta u}(X_i, t) &= \overline{K(\bar{q}(X_i, t) + q'(x_i, t))\beta(\bar{u}(X_i, t) + u'(x_i, t))}, \\ \overline{(u'q')_X}(X_i, t) &= \overline{(u'(x_i, t)q'(x_i, t))_x}, \quad \text{for } i = 1, \dots, N_F,\end{aligned}\tag{5.13}$$

however, as explained earlier and discussed again in 6.2.1, $\overline{(u'q')_X}(X_i, t) = 0$ due to the periodicity of the closure-scale subdomains, x_i . In the above equation, N_F is the number of large-scale grid points.

Recall that solutions to the filtered equation may be obtained using fewer computational resources than solving the original equations at high resolution provided that the closure models that simulate the unknown terms are solved efficiently by a model reduction, usually a reduction in the extent of the spatial domain. Details regarding the closure model follow.

5.2.4 Moisture closure equation

The closure model is obtained by substitution of (5.11) into (5.10),

$$q'_t + (u'q')_x + \mu q'_{xxxx} = K(q)\beta u - \bar{q}_t - \mu \bar{q}_{XXXX} - (\bar{u}\bar{q})_x - (\bar{u}q')_x - (u'\bar{q})_x.\tag{5.14}$$

The system is solved on the periodic subdomains $0 \leq x_i < 2\pi/N_F$ where N_F is the number of large-scale grid points. Terms on the right hand side of the equation involve inputs from the filtered model. Due to the periodicity imposed at the boundaries of each subdomain, filter-scale inputs to the closure model, \bar{u} , \bar{q} are considered constants across the extent of each subdomain and their derivatives, \bar{q}_t , \bar{q}_X , \bar{q}_{XXXX} , are zero. As was done for the filtered moisture equation, the filter-closure interaction terms in advection are also neglected in the moisture closure equation.

With the simplifications described above, rearrangement of the remaining terms in (5.14) is consistent with the closure model of Randall and Khairoutdinov (2010),

$$q_t = B_C + B_G + S_C,$$

which is their equation (6). In their notation, B_G , the adiabatic tendency of q due to filter-scale

effects, corresponds to $-(\bar{u}\bar{q})_X$ in our equation (5.14) or $B(\bar{v})$ from section 5.1.1. Their adiabatic tendency of q due to advection on the closure scale, B_C , corresponds to our term $-(u'q')_x$ or $B(v')$. The sources/sinks of q due to closure-scale physics, S_C , corresponds to our source term plus dissipative losses, $K(q)\beta u - \mu q'_{xxxx}$ or $S(v)$. One can also verify that by filtering equation (5.14) we recover (5.12) and that this remains consistent with Randall and Khairoutdinov (2010) equation (8).

While equation (5.14) provides q' leading to a satisfactory closure of equation (5.12), there is nothing inherent in the equation to ensure the closure is economical. The key to computational savings with MMF is to reduce the cost of simulating q' in the closure model through dimension reduction. Our simplified model allows us to systematically study the effect of reducing the dimension of the closure-scale models in several ways, albeit with fewer configuration options than those available to the higher-dimensional climate systems where typically an entire spatial degree of freedom, like longitudinal variation, is omitted. In this one-dimensional system, the subdomain dimension may be reduced such that it spans less than the spacing between filter-scale grid points. Alternately, a spectral truncation could be applied to the model equations eliminating negligible wave dynamics, and this is discussed further in the next chapter in 6.3.

5.2.5 Momentum equation

Ordinarily, the momentum equation (5.9) would be divided into filtered and closure parts as was done for moisture. The coupled system of filter-scale momentum and moisture equations together would comprise a large-scale system solved numerically on a coarse grid. This system would interact with an array of closure-scale systems made up of the small-scale momentum and moisture equations solved on fine grids.

However, the main goal of this work is to illustrate unambiguously the strengths and weaknesses of MMF in controlled experiments without other sources of error. As we saw from section 5.2.1, MMF is not sufficiently accurate for application to the KS equation alone because the boundedness of solutions is highly sensitive to the term needing closure, $\overline{uu_x}$. Although we could develop an alternative expression to close this term, it would not have anything to do with MMF and would introduce unrelated errors.

Instead, the way in which we proceed is to always solve the momentum equation (5.9) on

one continuous high-resolution grid over the full 2π -length domain (as opposed to on the periodic subdomains). Where $\bar{u}(X_i, t)$ is an input into the moisture equations, the high-resolution solution $u(x, t)$ is filtered and discretized at the large-scale grid points. Similarly, $u'(x_i)$ is computed as $u(x_i, t) - \bar{u}(X_i, t)$.

One complication is that the momentum equation is forced by $-\gamma q(x)$. An approximate high-resolution forcing is computed by concatenation of the piece-wise moisture solutions, $(\bar{q}(X_i) + q'(x_i))$, and interpolation onto the continuous fine grid. There may be some errors introduced at the discontinuous boundaries of the piece-wise solutions, however these are minor compared to errors introduced by MMF approximations.

5.2.6 Numerical simulation

For direct numerical integration of the KS-q system, all terms are discretized in space using spectral-Galerkin approximation except for the moisture source which is computed by spectral-collocation. The spatial resolution is $N = 64$ which includes wavenumbers up to $k = \pm 32$. The time-stepper is implicit trapezoidal for all linear terms and explicit second-order Adams-Bashforth for the nonlinear terms. The time step used for all simulations is $\Delta t = 0.001$.

The filtered system uses the same integration scheme as KS-q except that the spatial resolution varies between $N_F = 2$ and $N_F = 8$ with the resolved wavenumbers varying accordingly. The moisture closure models are integrated in physical space on 36 or 16 grid points using the same time-stepper as KS-q. The first- and fourth- order derivatives are computed using centered and five-point differences, respectively, both with second-order accuracy. Outputs to the filtered equation are transformed to Fourier-space by collocation. The filtered and closure models use the same time step as KS-q DNS.

We also use a spectral-Galerkin numerical scheme to solve the closure model equations to investigate model reduction in the following chapter in section 6.3. It has similar performance to the finite-difference scheme and is easier to implement different spectral truncations for lowering the cost of simulations. Resolutions include $k = \pm 16, 8$, and 4 wavenumbers. It uses the same time-stepper and time step as the finite-difference closure model. To facilitate using a variety of closure models, a flexible Python code manages inputs and outputs at the interface between closure and filtered numerical models.

MMF simulations iterate according to algorithm 3 below. The closure model simulations may easily be computed in parallel.

Algorithm 3 Time stepping routine in MMF filtered-closure model simulations.

1. Initialize $\bar{q}(X)^0, q(x)^0, u(x)^0, q'(x_j)^0$ for $j = 1$ to N_F
 2. For $t = 0$ to T
 - (a) Time step equation (5.9) and (5.12) to compute $u(x)^{t+1}$ and $\bar{q}(X)^{t+1}$
 - (b) For every filter-scale grid point, X_j , where $j = 1$ to N_F
 - i. Pass $u(x_j)^t, \bar{q}(X_j)^t$ as inputs to the j^{th} closure model
 - ii. Time step the closure model (5.14) to compute $q'(x_j)^{t+1}$
 - iii. Compute (5.13) from the closure model solution
 - (c) Supply the vectors $\overline{K(q)\beta u}(X)^{t+1}$ and $-\gamma q(x)^{t+1}$, aggregated from the closure model results, as input to (5.9) and (5.12) on the next time step.
-

5.3 Concluding remarks

This chapter reviewed the closure technique of multiscale modeling and introduced a novel test system that couples the KS equation to an evolution equation for moisture-like processes. An earlier test model using just the KS equation was rejected because the terms that needed the closure were the very terms that determined whether solutions grew unbounded. The moisture source term in the KS-q system, which does not have this property, is now the focus of our multiscale modeling efforts. The KS-q model is an appropriate test bed for the MMF method because it permits investigation of error sources one at a time. The results of this are the topic of the next chapter.

Chapter 6

Performance evaluation of dynamical closure

In this chapter, we provide an analysis of the errors due to each of the MMF approximations outlined in chapter 5. Our approach applies just one approximation at a time. First we consider the effects of neglecting filter-closure scale interactions in the quadratic terms of the filter-scale equation (5.12) in section 6.1. Next in section 6.2, we examine errors introduced by periodic boundary conditions applied to the closure-scale equation (5.14). In this section the closure model output and filtered model output are considered separately, without coupling to each other. This helps us dissect the error contribution of each equation without the possibility of canceling or compounding errors over time from coupling. In section 6.3, we evaluate the effect of reducing the dimension of the closure model. The coupled filtered-closure system is addressed in section 6.4.

6.1 Neglecting filter-closure interactions

A common approximation in MMF applications has been to neglect cross terms involving filter-scale and closure-scale variables that arise out of separating scales in the quadratic advection terms. We assess the validity of this approximation by simulating the KS-q system (5.9)–(5.10) where the advection term, A , is provided as forcing, computed ahead of time from DNS in three different ways: (1) exact advection, $(uq)_x$, as in DNS, (2) cross-terms omitted, retaining only $(\bar{u}\bar{q})_x$ and $(u'q')_x$ terms, and (3) all terms involving closure-scales omitted, leaving just the filter-scale advection term,

$(\bar{u}\bar{q})_X$.

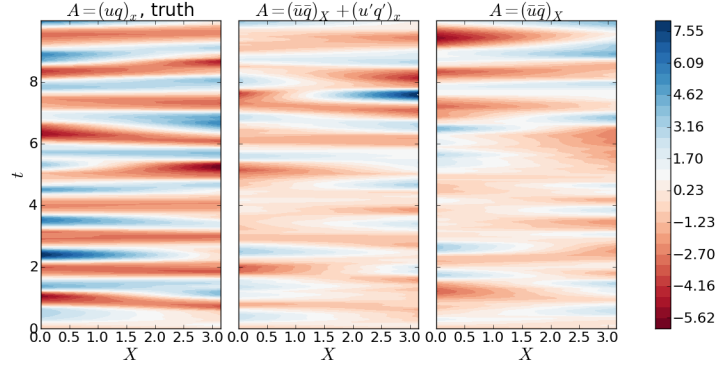


Figure 6.1: A realization of the filtered moisture field, \bar{q} , satisfying $q_t + A + \nu q_{xxxx} = K\beta u$, where (left) all scales are included in advection, (center) filtered-closure scale-interactions are omitted from advection, and (right) all advection terms involving closure-scale variables are omitted.

The resulting moisture solutions, q , are coarsely box-filtered with box length $\Delta = \pi$ to illustrate how well each advection closure captures filter-scale behavior. A comparison of \bar{q} for the three advection scenarios, given a particular initial condition, is plotted in figure 6.1. We see that beyond about one time unit, the individual eddies (local minima or maxima) become less clearly defined as the advection term is increasingly made approximate. The simulation neglecting only filter-closure interactions looks to have statistical behavior closer to the DNS simulation than the simulation neglecting all closure-scale terms. This is evident by comparison of the strength and size of eddies: the temporal duration of the eddies is generally longest and the amplitude weakest in the simulation neglecting all closure-scale terms (far right plot). For a more rigorous analysis of the eddy energy at various length scales, we turn to a spectral analysis.

The time-averaged spectra of q and $(uq)_x$ are plotted in figures 6.2 and 6.3, respectively, for each of the advection scenarios. In figure 6.2, it is evident that the cross-terms are responsible for a large amount of energy in the q spectrum at all wavenumbers. The omission of all closure terms introduces significant error in addition to this for wavenumbers greater than the cutoff of $k = 1$. (Although our MMF application uses a box filter, the filtering effects are similar to applying a sharp spectral filter with cutoff wavenumber $k_{cut} = \pi/\Delta = 1$.) Beyond $k = 2$, the performance of the no-closure-term scenario is as bad as if we had neglected advection altogether. Thus, when all closure-scale terms are neglected, energy found at high wavenumbers is due to the nonlinear moisture source term rather than any mixing due to advection.

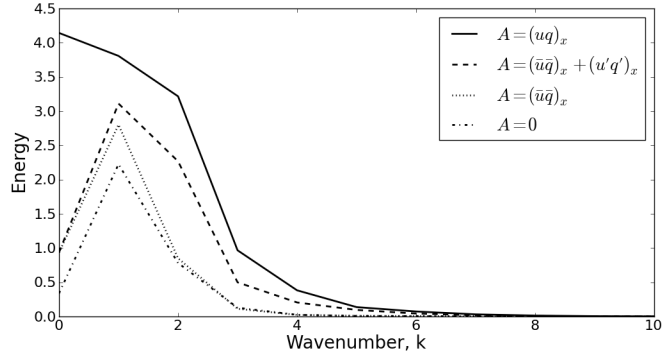


Figure 6.2: Comparison of the time-averaged energy spectra for q corresponding to the realizations of figure 6.1: all scales included in advection (solid), filter-closure scale interactions omitted from advection (dash), and all advection terms involving closure-scale variables omitted (dot). In addition, the spectrum without any advection of moisture at all is shown for reference (dash-dot).

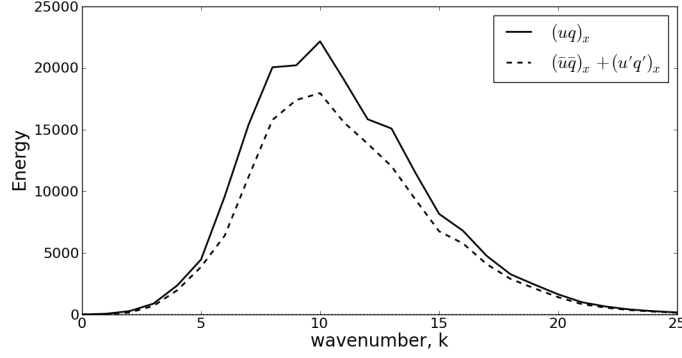


Figure 6.3: Comparison of the time-averaged energy spectra of $(uq)_x$ including (solid) and neglecting (dash) cross-terms.

The greatest discrepancy in the q spectra occurs at the zeroth wavenumber, but as we observe in figure 6.3, there is no error in the spectrum of $(uq)_x$ from cross-term omission at $k = 0$. The error in advection spectra does not become significant until $k \geq 3$ and increases steadily from there to a peak around $k = 7$. The energy discrepancy for q at $k = 0$ is actually due to indirect effects from the system's coupled-oscillator-like behavior. High wavenumber energy in the q spectrum that is lost due to omission of terms feeds back to the momentum equation whose higher modes will likewise be less excited. Less energy in the spectrum of u feeds back to the energy of q at $k = 0$ by way of diminished amplitude in the mean mode of the moisture source $K(q)\beta u$.

We explain the behavior observed in figures 6.2 and 6.3 in greater detail by considering advection expressed in Fourier space and by defining cross-terms and closure-terms with respect to the Fourier

modes they include.

The derivation of advection in Fourier space and term definitions are as follows. Momentum, moisture, and advection may be written, respectively, as the following series of Fourier mode products,

$$u(x, t) = \sum_k \hat{u}_k(t) \phi_k(x) \quad (6.1)$$

$$q(x, t) = \sum_k \hat{q}_k(t) \phi_k(x) \quad (6.2)$$

$$(uq)_x(x, t) = \sum_k \hat{a}_k(t) \phi_k(x), \quad (6.3)$$

with ϕ_k the Fourier mode at the k^{th} wavenumber. Substituting (6.1) and (6.2) into (6.3) and applying standard Galerkin projection, the coefficient for the k^{th} Fourier mode in advection is a series of products of u and q Fourier coefficients,

$$\hat{a}_k = \sum_{j=-N/2}^{N/2} ik \hat{u}_j \hat{q}_{k-j}. \quad (6.4)$$

Here, N is the total number of modes included in the series and $i = \sqrt{-1}$. Given this, a cross-term may be defined as any product of Fourier coefficients in (6.4), $\hat{u}_j \hat{q}_{k-j}$ where $|j|$ and $|k-j|$ lie on opposite sides of the filter cutoff wavenumber. A closure-scale term is one in which both $|j|, |k-j| > k_{cut}$.

With these definitions, it is evident that mean advection is zero or $\hat{a}_0 = 0$, which is consistent with the fact that on a periodic domain, the net transport due to advection must be zero. Consequently, both spectra shown in figure 6.3 intersect origin and neglecting cross terms in \hat{a}_0 is not a source of error in the energy of \hat{q}_0 . The importance of cross-terms increases with increasing wavenumber to reach a maximum near $k = 7$ and then diminishes. The increase is due to two factors. First, the number of cross terms in Fourier series (6.4) increases with wavenumber such that \hat{a}_1 has 2 cross terms: $\hat{u}_{-1} \hat{q}_2, \hat{u}_2 \hat{q}_{-1}$; \hat{a}_2 has 4 cross terms: $\hat{u}_0 \hat{q}_2, \hat{u}_{-1} \hat{q}_3, \hat{u}_2 \hat{q}_0, \hat{u}_3 \hat{q}_{-1}$; and \hat{a}_k where $k \geq 3$ have all 6 cross terms: $\hat{u}_0 \hat{q}_k, \hat{u}_1 \hat{q}_{k-1}, \hat{u}_{-1} \hat{q}_{k+1}, \hat{u}_k \hat{q}_0, \hat{u}_{k-1} \hat{q}_1, \hat{u}_{k+1} \hat{q}_{-1}$, the maximum number of interactions possible between coefficients on opposite sides of the cutoff. Second, the peak wavenumber in the u spectrum occurs near $k = 7$. After that, the error diminishes because

the energy in u diminishes according to fourth order dissipation.

When all closure-scale terms are neglected, the error in the advection energy spectrum is infinite for $k \geq 3$. This is because there are no components in (6.4) that involve only large-scale terms, i.e. no terms have both $|j|, |k - j| \leq k_{cut}$.

The above analysis tells us that cross-terms and closure-scale terms in advection are of first order importance. Neglecting them will be a major source of error in MMF simulations of our system.

6.2 Periodic boundary error

In this section the periodic boundary assumption is evaluated. We begin with a discussion of what that means for the ability of MMF to close closure-scale advection terms in section 6.2.1. Next we evaluate the assumption in each of the MMF equations separately. Initially we evaluate the closure equation by dynamical downscaling in which (5.14) is forced by filter-scale inputs computed ahead of time from DNS. The resulting realizations of q' and their statistical properties and energy spectra are compared to the correct DNS values in section 6.2.2. Then the q' realizations are used to compute the terms needed to close the filter-scale equation (5.12) in section 6.2.3.

6.2.1 Advection

As was made evident by the analysis of section 6.1, small-scale advective transport has leading order effects on the energy spectra of the large-scale KS-q model. Unfortunately, due to its periodic subdomain configuration MMF cannot be used as a closure method for these effects. According to MMF, the unknown advection terms would be computed on each subdomain with subdomain-average values passed to the filtered equation. With periodic boundary conditions, the net advection out of a subdomain is artificially constrained to be zero. This undesirable feature of MMF closure models prohibits modeling advection transport other than at the filter scale between large-scale grid points.

There has been some work in the literature to remedy the problem by Jung and Arakawa (2005). Their scheme prescribes Dirichlet boundary conditions by interpolating the large-scale solution at the borders of each closure model domain. However, alternatives to periodic boundaries remain

an area for additional research. One issue with prescribing boundaries is that they cannot evolve with the closure model; they are determined by the large-scale and fixed for the duration of each large-scale timestep. Foreseeable solutions may have more in common with finite element methods like discontinuous Galerkin (Cockburn (2004); Xu and Shu (2006)) where elements pass fluxes at their borders rather than merely sharing a common prescribed boundary.

In three-dimensional climate applications the problem may be less severe than in our simple test system because it only affects horizontal advection. The cloud-resolving closure models occupy the full vertical extent of a filter-scale grid column which has aperiodic vertical boundary conditions so vertical transports can be modeled.

Because advection will be a major source of error for MMF given our model, we will proceed by removing advective transport from the moisture equation for the rest of the simulations in this chapter. This will allow us to focus on evaluating MMF as closure for the moisture source term alone, which is the main process targeted for improvement by MMF in more comprehensive climate model applications. However, even without the complications of advection, there are still other limitations we will encounter due to the assumption of periodicity. This is the topic of the remainder of section 6.2.

6.2.2 Downscaling

The following results are from simulations of the closure model (5.14) alone where time series of filter-scale inputs are computed a priori by filtering KS-q DNS. The closure model provides realizations of q' which we simulate for two different box filter length scales: box length $\Delta = \pi$ and $\pi/2$. This is equivalent to discretizing the domain such that the number of filter-scale grid points, N_F , is equal to 2 and 4 points (or up to wavenumbers 1 and 2), respectively. This is also the number of subdomains on which we run the closure model.

Representative simulation results are shown in figure 6.4 for half of the full length 2π domain. The other half is computed, of course, but is statistically very similar and so not shown here. Comparing the results from $N_F = 4$ to $N_F = 2$, we see that closure simulations using four shorter subdomains are less accurate than simulations on just two long subdomains. Temporal frequency of fluctuations seem to agree between the true and model estimated simulations as expected because we did not filter the KS-q system in the time domain. However, the modeled simulations with $N_F = 4$

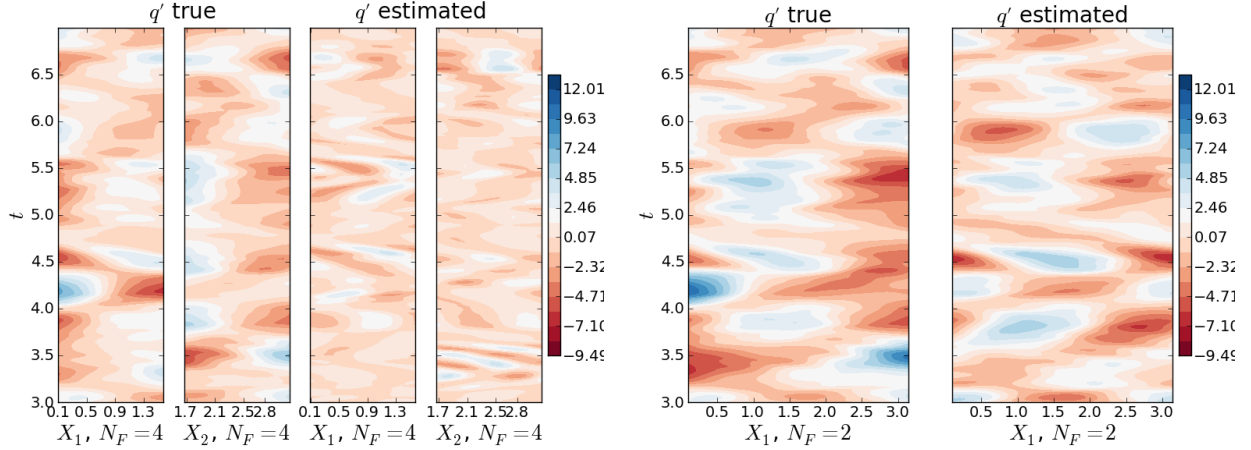


Figure 6.4: Comparison of q' realizations from DNS and closure models with two filter length-scales: $N_F = 4$ (left group) and $N_F = 2$ (right group). Within each grouping, the left plots are from DNS and the right plots are from closure models.

do not fluctuate with enough amplitude compared to the true simulations. The perturbations, q' , are too weak and this is a consequence of imposing periodic boundaries on short subdomains.

Periodicity constrains the values and gradients of solutions at the boundaries to agree with each other. This may inhibit wave development and lead to damping of the overall spatial variability if the subdomain is too short. This turns out to be the case for $N_F = 4$; the subdomain length is about equal to the length of the largest eddy we try to capture with the closure model so the constrained boundaries have a profound effect on the magnitude of eddies at this scale in the interior of the subdomain.

To preserve variability in the interior in the presence of damped boundaries, it is necessary to increase the subdomain length to twice the eddy length or longer. The simulations with $N_F = 2$ satisfy this requirement, and so we see marked improvement in the accuracy of spatial variations. The statistics of the realizations provided in table 6.1 and the energy spectra in figure 6.5 confirm the finding that the spatial variance is too weak for $N_F = 4$ but acceptable for $N_F = 2$.

N_F	$\text{Var}(q')$	
	DNS	Closure Model
4	2.19	0.99
2	4.64	3.98

Table 6.1: Spatial variance, $\overline{|q'(x)|^2}$, in simulations of the closure model averaged over the simulation time and over all subdomains from DNS and two closure models.

Close inspection of the spectra for $N_F = 2$ in figure 6.5 reveals yet another limitation of periodic

boundaries. Although the overall variance and peak energy are quite accurate, substantial energy is missing at wavenumbers two and higher. High energy modes are needed to capture the aperiodic boundaries of the DNS, a phenomenon well understood by Gibbs. In the periodic closure models such discontinuities at the boundaries are quickly diffused by 4th order dissipation.

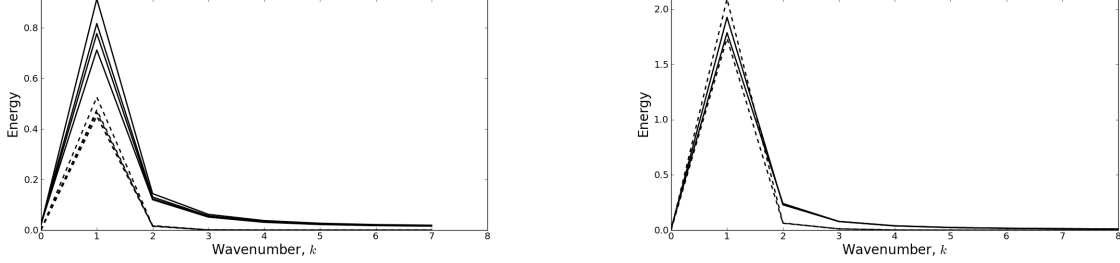


Figure 6.5: Comparison of energy spectra of q' from each closure model realization at the two filter scales: $N_F = 4$ (left) and $N_F = 2$ (right). DNS spectra shown as solid lines and closure model spectra are dashed.

Although the variance is underestimated for $N_F = 4$, both discretizations capture an important feature of the closure-scale behavior: heterogeneity in the development of eddies, or intermittent periods of high and low eddy activity. Intermittent behavior is evident in the time-series of spatial variance shown in figure 6.6 for both $N_F = 4$ and $N_F = 2$. This is the main benefit of using a dynamical closure model derived from first principles; one achieves heterogenous solutions in agreement with the true dynamics of the system without any need for empirical tuning.

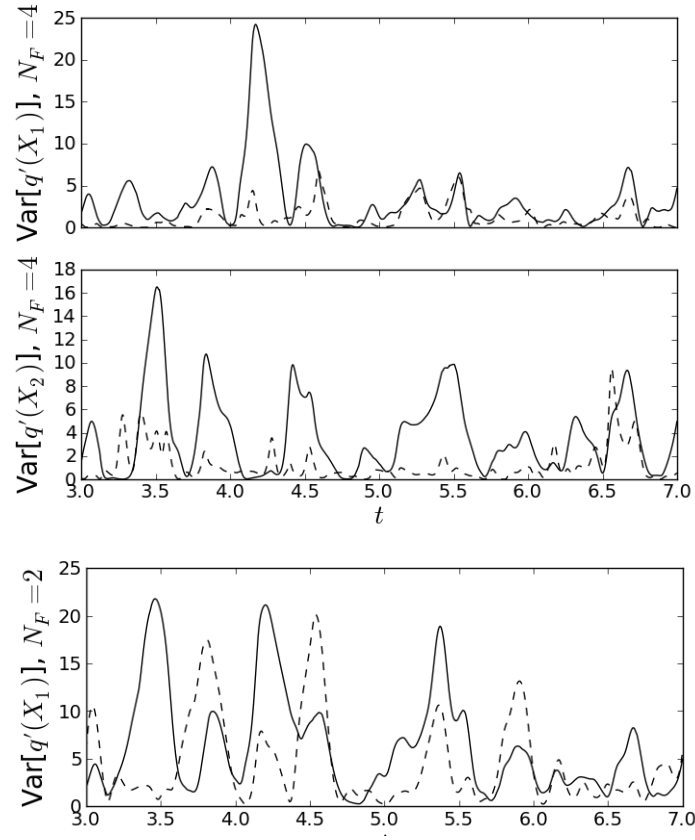


Figure 6.6: Comparison of spatial variance of q' over time from DNS (solid) and closure model solutions (dash). The top two plots are results from two representative subdomains where $N_F = 4$. The bottom plot is from one representative subdomain where $N_F = 2$.

6.2.3 Upscaling

We now use the realizations of q' computed via downscaling to compute the closure terms and pass them back to the filtered equation (5.12) as external forcing. The q' from the i^{th} closure model are supplied in the filter-scale moisture source term defined as

$$\overline{M(u, q)}(X_i) = \overline{K(\bar{q}(X_i) + q'(x_i))\beta(\bar{u}(X_i) + u'(x_i))}, \quad (6.5)$$

where $i = 0, 1$. As we learned in the previous section that closure simulations are more accurate for $N_F \leq 2$, we use $N_F = 2$ as the filter-scale resolution for all of the upscaled simulations.

We compare solutions for \bar{q} where the unknown \overline{M} is computed from DNS, from MMF closure models, and from large-scale closure alone. Large-scale closure is computed without any knowledge of sub-filter-scale perturbations, $\overline{M} = K(\bar{q})\beta\bar{u}$.

The MMF simulation substantially outperforms the large-scale closure as evident in table 6.2 when comparing the spatial variances of \bar{q} averaged over the simulation time period. The small discrepancy between the MMF closure and DNS is due to the missing high wavenumber energy in q' (see figure 6.5) that is responsible for additional attenuation of \bar{q} . We confirmed this by adapting the closure model code to accept Dirichlet/Neumann boundary conditions extracted from DNS. With near perfect simulation at the now aperiodic boundaries, the MMF performance was excellent. Unfortunately, this is not a viable solution because in actual coupled simulations, the boundary values would not be known a priori, they evolve with the simulation.

The large-scale-closure solution has several long periods of high variance that skew its result as shown in the top plot of figure 6.7. Each one is preceded by a spike in the variance of \overline{M} shown in the bottom plot. Large-scale-closure \overline{M} has higher spatial variance for longer time periods than for DNS or MMF because it is missing the attenuating effects that come with greater resolution of q as was explained in section 5.2.2. Also, the large \overline{M} fluctuations persist for longer because \bar{u} and \bar{q} vary less rapidly in time than u and q . This is evident by the bottom plot of figure 6.7 but also from the integral of $|\overline{M}|$ over the simulation period provided in table 6.2. The slow variations in the large-scale source term make its overall contribution to the evolution of \bar{q} , $\int |\overline{M}| dt$, almost 1.5 times greater than for DNS and MMF closure.

	$\text{Var}(\bar{q})$	$\text{Var}(q')$	$\int \bar{M} dt$
DNS	1.89	4.64	187
MMF closure	3.48	3.98	185
Large-scale closure	16.8	-	268

Table 6.2: Comparison of simulation statistics time-averaged over the simulation period: the spatial variance of \bar{q} , the spatial variance of q' and the integral of the magnitude of the moisture source \bar{M} (6.5) of the KS-q system using different closure models.

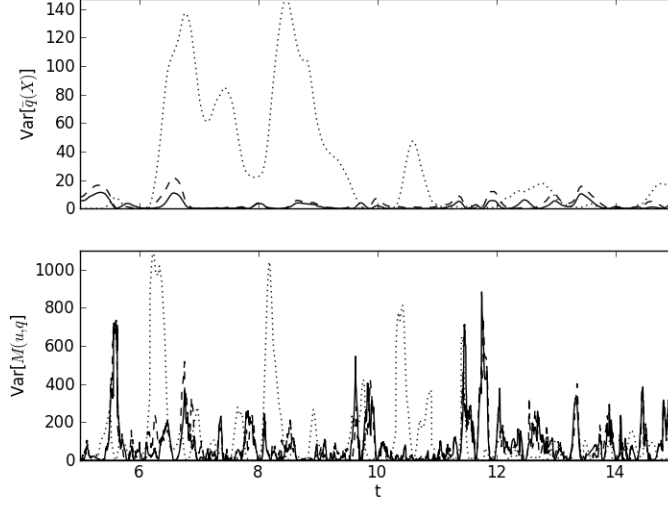


Figure 6.7: (top) Spatial variance of \bar{q} over time from DNS (solid), MMF closure (dash), and large-scale closure (dot). (bottom) Same as top but for the moisture source \bar{M} (6.5).

6.3 Closure model dimension reduction

The final approximation to investigate is the reduction of the state dimension of the closure model to achieve savings in computational cost. For CRMs this is typically accomplished by shortening the horizontal extent of the closure model subdomain. Given what we have learned about the effect of periodic boundary conditions on very short subdomains, shortening the subdomain length further is not possible without greatly increasing error. Instead, we propose reducing the number of Fourier modes used as a basis for a spectral version of the closure model.

Evident from the energy spectra plotted in figure 6.5, most of the energy and hence the variability in the closure-scale moisture solutions is contained in the first 4 Fourier modes. Therefore, a reasonable reduced order model should be obtained eliminating as much as mode 5 and higher, which would still retain about 94% of the energy in the spectrum.

We run downscaling simulations of the closure model decreasing the number of complex modes

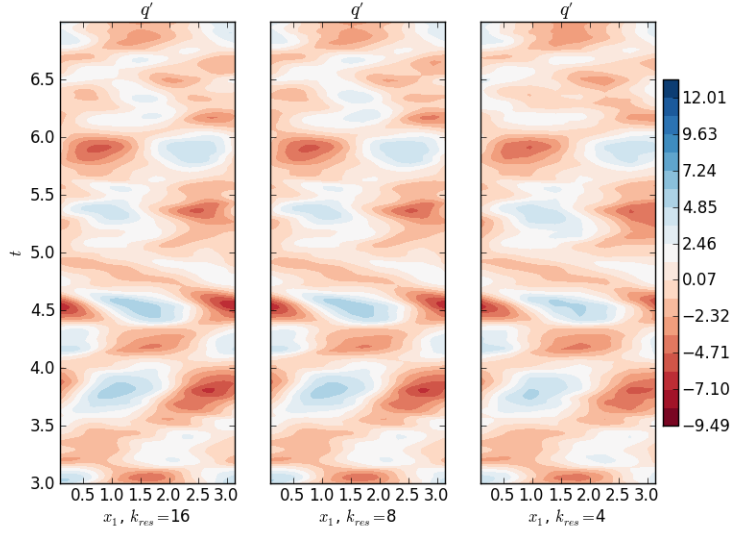


Figure 6.8: Realizations of q' generated by spectral closure models with the number of Fourier modes resolved decreasing from left to right, $k_{res} = 16$, 8, and 4, respectively.

resolved from $k_{res} = 16$ down to the coarsest simulation at $k_{res} = 4$. Note that because solutions are all real, we only need to numerically integrate equations for Fourier coefficients corresponding to positive wavenumbers, the negative coefficients are their complex conjugates.

The resulting realizations of q' in figure 6.8 are shown for the same initial condition on the same closure model subdomain (x_1) as in figure 6.4. The left plot with $N_F = 2$ in figure 6.4 is the truth from DNS to which these runs should be compared. The reduced-order runs perform very well, capturing the distribution of major spatial features correctly with only minor loss of some of the finer features as resolution decreases.

The effect of resolution on the variance of q' is summarized in table 6.3. The variances among the three resolutions are almost identical though they underestimate the true variance slightly due to the periodic boundary effects explained in section 6.2.2. The energy spectra in figure 6.9 shows the same missing energy in modes two and higher that was observed in figure 6.5 using the finite-difference closure model.

Cost savings is achieved when the resolution of each of the two closure models is less than 16 modes. Direct numerical simulations of the entire 2π domain requires 32 Fourier modes. MMF simulations reduce cost by about 50% when the closure models resolve 8 modes and by 75% when closure models resolve 4 modes.

k_{res} (% Energy)	$\text{Var}(q')$	Percent Error
DNS	4.64	-
16 (100%)	3.95	14.9%
8 (97.7%)	3.96	14.7%
4 (93.7%)	3.97	14.5%

Table 6.3: Spatial variance, $\overline{|q'(x)|^2}$, and percent error in simulations of the closure model averaged over the simulation time and over all subdomains, as the dimension of the model, k_{res} , is reduced by truncating the number of Fourier modes that are resolved.

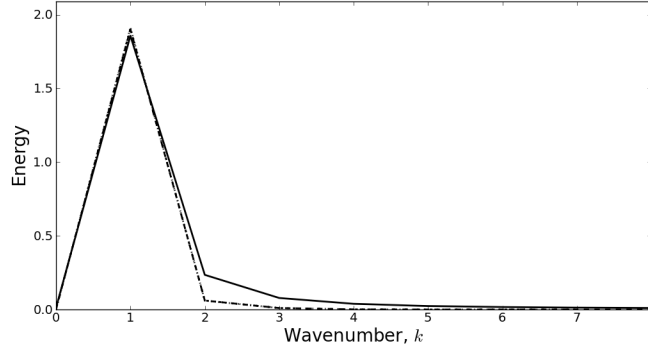


Figure 6.9: Comparison of energy spectra of q' averaged over all subdomains for each of the resolutions in table 6.3. DNS spectrum is shown as solid line and closure model spectra are dashed and lie on top of one another.

Our approach to dimension reduction in which we project the closure model dynamics onto a Fourier basis and select which modes to retain based on their energy content is an old approach in the area of low-dimensional modeling (Lumley (1970); Holmes et al. (1996)) but is a new contribution to MMF. Previous applications achieve reduction by eliminating spatial dimensions which does not take energy content into account and would not have had as successful low-order results in our test model. Our example is only the beginning of how dynamical considerations, like energy, can be used to build low-dimensional closure models. The area of reduced order modeling is rich with techniques for systems with more complex geometries than our one-dimensional periodic example. A great advantage of MMF is that it breaks a system into distinct, parallel, regional models with clear large-scale inputs and outputs. This makes MMF very compatible with advanced techniques from reduced order modeling and system identification that may be used to build highly efficient input-output maps to approximate the regional model dynamics.

6.4 Coupling error

In this final section of experiments, we present the results of simulations where the filtered equation is fully coupled to the closure models. In fully coupled simulations all of the previously examined sources of error (besides advection which is still ignored here) come into play simultaneously. Plus additional error is introduced due to the fact that respective inputs to the equations are no longer prescribed from DNS computed ahead of time as they were for the downscaling and upscaling experiments. For the first time in this chapter, the momentum equation responds dynamically to the MMF estimates of q through the forcing $-\gamma q$. Error in either field leads to changes in the evolution of the other. Because the simulations are not anchored to prescribed inputs, our goal with MMF is to match the statistics of the true model rather than specific realizations.

We compare MMF runs to DNS and to simulations of the filtered equation with only large-scale closure where $N_F = 2$. We also compare the MMF results to simulations of the filtered equation with higher resolution, $N_F = 16$, which is comparable in expense to the MMF runs that use two closure models resolving 4 Fourier modes each. This comparison is rarely made for more comprehensive climate models because their parameterizations are not generally designed to scale with resolution. Such a comparison would require substantial new model development to fit parameterizations to the higher resolution model. Our system has the advantage of easy adaptation to various resolutions.

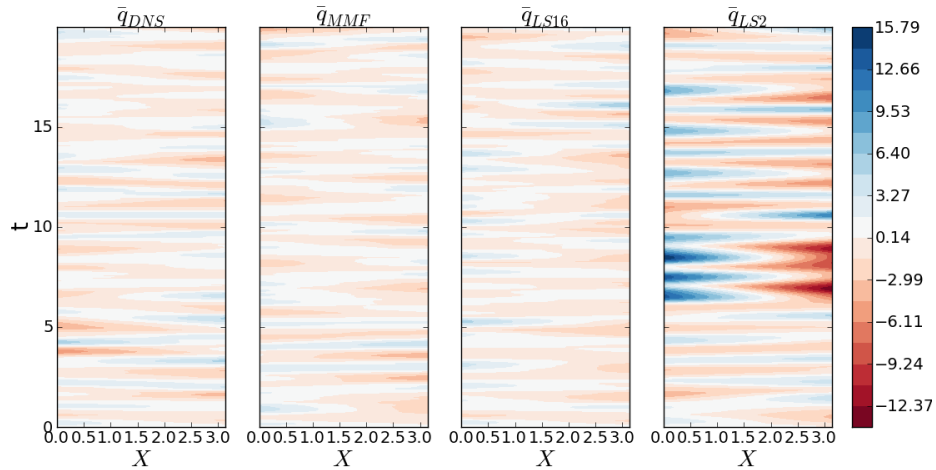


Figure 6.10: Coupled realizations of the filtered moisture field, \bar{q} for four different model configurations: DNS, MMF, LS2, and LS16 from left to right, respectively.

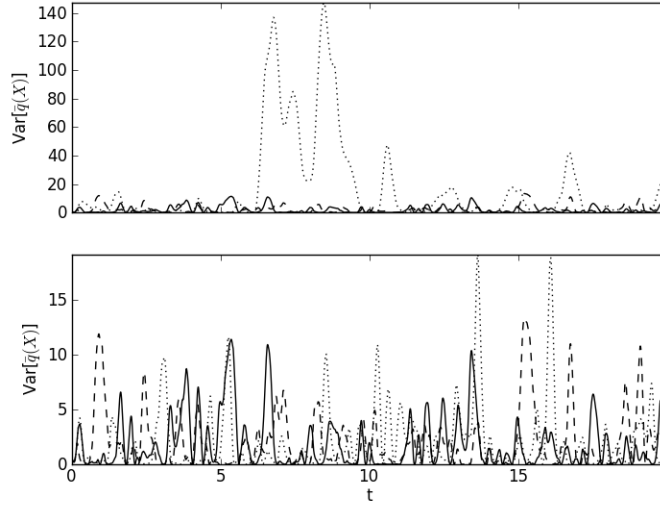


Figure 6.11: (top) Comparison of spatial variance of \bar{q} over time from DNS (solid), MMF (dash), and LS2 (dot). (bottom) Comparison of spatial variance of \bar{q} over time from DNS (solid), MMF (dash), and LS16 (dot).

Realizations of \bar{q} are shown in figure 6.10 for qualitative comparison of four model configurations: DNS, MMF with $N_F = 2$, $k_{res} = 4$, large-scale closure with $N_F = 2$ (called LS2) and large-scale closure with $N_F = 16$ (called LS16). The fluctuations in the LS2 simulation are obviously too strong over several time periods. LS16 performs much better and is quite similar to MMF and DNS solutions. The frequency and magnitude of oscillations are in agreement on average. The LS16 simulation does have two peaks in spatial variance at about $t = 14$ and $t = 16$ that are larger than any individual fluctuation found in the DNS or MMF runs. This is apparent in the bottom plot of figure 6.11 which shows the spatial variance of LS16, DNS, and MMF over time. Generally, MMF and LS16 both capture equally well the average variance and the intermittency of periods of high variance interspersed with less active periods. The top plot shows the same comparison but with LS2. All four simulations could not be shown on the same axes because the LS2 simulation has three very large peaks that far exceed peaks in all of the other runs.

The statistics of the runs are summarized in table 6.4. Perhaps surprisingly, compared to the upscaling statistics from table 6.2, the fully coupled MMF performance is actually much improved. In spite of the missing energy in the spectrum of q' that caused too much variance in \bar{q} in the upscaling experiments, the agreement of \bar{q} from DNS and MMF in the coupled experiments is excellent. This suggests that even in this simple system, there exists cancellation of errors. The

Resolution	Var(\bar{q})	Var(q')
DNS, $N = 64$	1.89	4.64
MMF, $N_F = 2, k_{res} = 4$	2.05	4.15
LS16, $N_F = 16$	2.14	-
LS2, $N_F = 2$	16.8	-

Table 6.4: Spatial variance of \bar{q} and where applicable of q' for four different model configurations averaged over the simulation time and over all subdomains.

cancellation occurs by the following mechanism. The oscillator behavior allows u to respond to errors in the magnitude of q . The magnitude of the moisture source depends on u and q oppositely: missing energy in u causes decreased variance in \bar{q} which counteracts the variance increasing effects of missing energy in q . We verify that indeed this cancellation is occurring by measuring a slight drop in the spatial variance of u .

6.5 Concluding remarks

In summary, this chapter detailed a series of simulation results with the KS-q test model to highlight the strengths and weaknesses of MMF. We showed that the approximations of neglecting cross-terms in the filtered equation and applying periodic boundaries to the closure models are the most dominant sources of error. In spite of these difficulties, we showed that MMF may be a good closure technique in cases where the terms needing closure depend primarily on local effects rather than nonlocal processes occurring outside of the closure model. The moisture source term in our system is an example of the former, whereas sub-filter-scale advective transport proved to fall in the latter category.

We also introduced a new way of thinking about dimension reduction in the closure models based on energy content. The Galerkin projection of model dynamics onto Fourier modes and subsequent truncation of low-energy modes is a common approach to dimension reduction in the area of reduced order modeling but it has not previously been used with MMF. This has most likely been due to the availability of existing cloud resolving models that have already reduced physical dimensions from three- to two-dimensional domains. Application of reduced order modeling techniques is a promising future direction for attaining better low cost closure models.

Finally, we compared results from the MMF system to both DNS and a large-scale model with

expense comparable to the MMF. Both of these comparisons would be difficult if not impossible for comprehensive global climate models. We found that, statistically, both MMF and the comparable large-scale model are almost equally good approximations to DNS. This raises the question of where to invest computational resources as they become increasingly available. Assuming that the use of global cloud resolving models for climate applications remains prohibitively expensive for another decade, should present model development focus on higher resolution MMF closure models or increasing global model resolution and updating parameterizations to match?

Neither of these approaches converges to a global cloud resolving model with increasing resolution. However, MMF may be a preferred approach for a few reasons not directly addressed in our work with the KS-q system. First, MMF provides a framework for coupling existing large-scale general circulation models of the climate with existing cloud-resolving models which minimizes the amount of new model development. Conventional parameterizations, on the other hand, do not generally take grid-resolution as an input parameter so higher-resolution parameterizations would need to be developed anew. Second, if the climate has greater scale separation than our test system, MMF may prove more economical. An increase in resolution would extend the spectrum of resolved scales in a continuous manner. Supposing processes act at large and small scales with some spectral band gap between them, MMF would enable modeling just the relevant scales discretely. Finally, MMF closure models may capture the interaction of multiple sub-filter-scale processes. This kind of interaction is difficult to model by conventional parameterizations which do not interact in a dynamically consistent way. In the words of Jung and Arakawa (2005), MMF may be thought of as a useful ‘physics-coupler’ providing a viable intermediate approach to modeling cloud processes until the computational power is available to resolve them globally.

Chapter 7

Conclusion

The main contributions of this thesis are: three applications of systematic closure methods to simple climate models, adaptations to the formulation of two of the methods, development of one new model, and an advancement in the scientific understanding of uncertainty in short range climate change. The systematic closure methods are the unscented Kalman filter, the adaptive-covariance-rank unscented Kalman filter, and the multiscale modeling framework. They are applied for the first time to a zero-dimensional energy balance, the Lorenz 1996 abstract atmospheric system, and the coupled Kuramoto-Sivashinsky and scalar transport equations, respectively. Each model embodies one of the major difficulties in closure of the climate system: uncertain and limited data, high dimension, and vast scale. Each application demonstrates how one of these difficulties may be addressed systematically.

In the energy balance example, uncertain data was the primary concern. This included unknown historical forcing and natural variability in the temperature record. These factors make it difficult to precisely estimate the transient climate sensitivity. The unscented Kalman filter was shown to efficiently account for uncertainty in data and provide probabilistic estimates of TCS rather than focusing strictly on a single value for it. The UKF was an advantageous approach for its efficiency, second-order error bounds, and built in uncertainty analysis through the Kalman update of both mean and covariance. Also, because UKF is a sequential Bayesian method, it allowed us to experiment with the effects of initial uncertainty assumptions and to easily make future predictions of uncertainty reductions given additional data.

Our successful application of the UKF method led to a scientific contribution in short range

climate prediction. We produced a set of TCS probability densities mostly independent of global climate model results that were consistent with findings of the IPCC AR4 and other studies. We also found that the uncertainty in the predicted TCS range has the potential to narrow by about 45% given two additional decades of future data. Giving credibility to our results, we validated the method on output from the comprehensive global climate model CM2.1 developed at GFDL by showing that it estimated the TCS of CM2.1 to within predicted error bounds.

In the Lorenz 1996 system we addressed the issue of large dimension. We designed an adaptation to the UKF to minimize the number of model evaluations required to make probabilistic state and parameter estimates. The new algorithm, called adaptive-covariance-rank UKF, reduced computational expense by modal decomposition of the error covariance matrix and truncation of its higher order modes. This led to a smaller error space for sampling by sigma points. We found it achieved cost savings of 30-60% over the original UKF formulation without the need to introduce any additional heuristic tools.

Work on a scalar transport equation for cloud-like processes coupled to the Kuramoto-Sivashinsky equation highlighted the problem of wide ranging length scales. We designed a model that could be divided into filter-scale and sub-filter-scale dynamics and applied the multiscale modeling framework to efficiently simulate sub-filter-scale dynamics. The model was developed especially to address some of the complications in understanding the performance of MMF with more complex climate models. Our model made possible direct comparison of MMF performance with DNS and an alternative model of comparable expense. We found that the MMF and comparable models had similar accuracy but that the MMF may be advantageous for coupling multiple small-scale processes and in systems with greater scale separation.

Our application of MMF modified the usual means of achieving dimension reduction in the physical domain of the sub-filter-scale models to instead use reduced order models truncated to maximize energy content. Using a reduced order modeling approach allowed reduction of the closure model dimension up to 75% while retaining the same level of accuracy. This would not have been the case if we had made the same reduction by cropping the physical domain length.

The MMF application also illustrated some of the strengths of MMF as well as some weaknesses. Specifically, we explained that MMF is not an appropriate closure method for the Kuramoto-Sivashinsky equation because the boundedness of its solutions was highly sensitive to the accuracy

of the closure. Also, with sub-filter-scale models that do not communicate at their borders, MMF is not effective at closing terms like advection that depend on transport between sub-filter-scale domains. The MMF is better suited to closing terms dependent on more localized dynamics like the source term in our moisture transport equation. Finally, the periodic boundaries of the closure models limited the minimum length of subdomains to two eddy lengths. Anything shorter resulted in artificial attenuation of variability on the interior of the subdomain.

In summary, this thesis applied methods adapted from estimation techniques in the fields of system identification and model reduction to models adapted from the fields of climate science and dynamical systems. Recommendations for continued work at the intersection of these fields are made below.

7.1 Future research directions

Uncertain climate data. Future work with respect to the problem of uncertain or limited observational and forcing data could address the use of datasets with more spatial resolution than the globally-averaged surface temperature and forcing records used in this work. For example, a first step could be to resolve the data for each hemisphere and modify the energy balance appropriately. Similarly, the model could be extended to include upwelling-diffusion processes in a horizontally averaged global ocean. This would allow for use of temperature data at depths below the surface. It would also make possible UKF estimates of the equilibrium climate sensitivity. One challenge in this direction will be that with increased resolution, additional unknown parameters will arise in the model. It would be useful to understand whether the increased number of observational constraints overcomes the increased uncertainty in the model. Also, experiments could be repeated with longer and improved forcing estimates such as those recently made by Friend (2011) dating back to AD 1.

High dimension. Further work in the area of parameter estimation for very large systems using Kalman filters will involve additional strategies to minimize the number of times a model must be evaluated to generate accurate enough probability densities. One specific area of research would address the process and measurement noise spaces. If these uncertainties can be approximated as constants or as linearly additive then explicit model runs to sample these errors may not be needed.

Breadth of scales. A promising area for future work with multiscale modeling is in the design of alternative sub-filter-scale modeling techniques. Already there has been some work examining ways to improve the communication between neighboring closure models. (Jung and Arakawa (2005); Majda (2007)) These efforts attempt to alleviate the errors arising from periodic boundary conditions. Further investigation would be useful including understanding how finite element methods like discontinuous Galerkin (Cockburn (2004); Xu and Shu (2006)) could help. Another line of work could use reduced order modeling techniques on the cloud-resolving closure models used in more realistic multiscale applications. Toward this end, Bailon-Cuba and Schumacher (2011) have recently built a low-dimensional model for turbulent Rayleigh-Benard convection using Galerkin projection of the Boussinesq equations onto a modal basis obtained using Proper Orthogonal Decomposition (Lumley (1970); Holmes et al. (1996)) also known in the weather and climate field as Empirical Orthogonal Function evaluation (Lorenz (1956)). This model could be tested as a closure model in the multiscale framework. Similar models but including cloud equations could be developed.

Combined methods. Finally, this work focused on the application of one method at a time to models that isolate a single closure issue. However, no physically realistic climate model exhibits only one of these issues. Conversely, no single method addresses every issue. For realistic systems, one or more methods may be applied. Future work could involve developing criteria for prioritizing the closure issues that should and could be addressed for more complex models. New methods could be designed that merge aspects of the filtering and multiscale approaches. For example, adding sigma points to the UKF that correspond to MMF design parameters would allow one to sample some of the structural uncertainties in the model, like sub-filter-scale domain length, closure model resolution, or cloud microphysical formulation.

Appendix A

Filter Review

This appendix provides an overview of commonly used filter equations for nonlinear systems. We review these concepts in order to establish general concepts and define variables that are referenced in the specific filter implementations in the main text. This baseline will also aid in the comparison of the original UKF presented in 2.3.1 with our modified adaptive-rank version in 4.1.

In general, filter algorithms follow a common two-step format where, after initialization, the system state estimate and error covariance are first forecast in time according to a dynamic model, and second, corrected by a set of observations to the extent possible given measurement errors.

Consider the following discrete-time, nonlinear system,

$$\begin{aligned}x_k &= f(x_{k-1}, u_k, w_k) \\ y_k &= h_k(x_k, u_k, v_k)\end{aligned}\tag{A.1}$$

where f is a forward model operator and h_k is a measurement operator that may vary every k -th filter iteration, as in, for example, a moving sensor network. The system state, x_k is L_x -dimensional with covariance

$$P_k^x = E[(x_k - E(x_k))(x_k - E(x_k))^T].$$

The system has known input u_k . The system also has unknown input w_k , which is L_w -dimensional zero-mean Gaussian process noise with covariance $R_k^w = E[w_k w_k^T]$. The system outputs y_k may be a function of L_v -dimensional zero-mean Gaussian measurement noise v_k , with covariance $R_k^v = E[v_k v_k^T]$. For implementation of the UKF, one typically considers an augmented system of dimen-

sion $L = L_x + L_w + L_v$, in which the noise variables are appended to the state vector:

$$x_k^{aug} = \begin{bmatrix} x_k \\ w_k \\ v_k \end{bmatrix}.$$

Assuming there is no cross-correlation between the state and noise variables, we form the augmented block-diagonal covariance matrix

$$P_k = \begin{bmatrix} P_k^x & 0 & 0 \\ 0 & R_k^w & 0 \\ 0 & 0 & R_k^v \end{bmatrix}. \quad (\text{A.2})$$

In the *forecast* step, filters provide estimates, x_k^f and $P_k^{x^f}$, of the state, x_k and its covariance at a future time according to the following,

$$x_k^f = E[f(x_{k-1}, u_k, w_k)]$$

$$P_k^{x^f} = E[(x_k - x_k^f)(x_k - x_k^f)^T]$$

In the measurement update, the following equations are used to provide the corrected quantities, often called the state and covariance after *analysis*:

$$x_k^a = x_k^f + K_k(y_k - E(y_k)) \quad (\text{A.3})$$

$$P_k^{x^a} = P_k^{x^f} - K_k P_k^{yy} K_k^T. \quad (\text{A.4})$$

Here, the gain

$$K_k = P_k^{xy} (P_k^{yy})^{-1} \quad (\text{A.5})$$

weights the correction based on the relative contribution to the overall uncertainty from dynamic and measurement factors. The matrices P_k^{xy} and P_k^{yy} are the state-measurement error cross-

covariance and measurement error covariance, and are given by

$$P_k^{xy} = E[(x_k - x_k^f)(y_k - E(y_k))^T]$$

$$P_k^{yy} = E[(y_k - E(y_k))(y_k - E(y_k))^T].$$

For detailed derivations of these equations, see, for instance, Simon (2006). Specific filter implementations vary in their assumptions about the system (for example, linearity), and in their approximations of the expected values and error covariances in the above expressions.

Bibliography

- Allen, M. R. and D. A. Stainforth, 2002: Towards objective probabilistic climate forecasting. *Nature*, **419**, 228.
- Allen, M. R., P. Stott, J. F. B. Mitchell, J. F. B. Schnur, and T. L. Delworth, 2000: Quantifying the uncertainty in forecasts of anthropogenic climate change. *Nature*, **407**, 617–620.
- Ambadan, J. T. and Y. Tang, 2009: Sigma-point Kalman filter data assimilation methods for strongly nonlinear systems. *Journal of the Atmospheric Sciences*, **66** (2), 261–285.
- Anderson, J., 2001: An ensemble adjustment Kalman filter for data assimilation. *Mon. Wea. Rev.*, **129**, 2884–2903.
- Annan, J. D. and J. C. Hargreaves, 2004: Efficient parameter estimation for a highly chaotic system. *Tellus*, **56A**, 520–526.
- Annan, J. D., J. C. Hargreaves, R. Ohgaito, A. Abe-Ouchi, and S. Emori, 2005a: Efficiently constraining climate sensitivity with ensembles of paleoclimate simulations. *SOLA*, **1**, 181–184.
- Annan, J. D., D. J. Lunt, J. C. Hargreaves, and P. J. Valdes, 2005b: Parameter estimation in an atmospheric GCM using the ensemble Kalman filter. *Nonlin. Processes Geophys.*, **12** (3), 363–371.
- Bailon-Cuba, J. and J. Schumacher, 2011: Low-dimensional model of turbulent rayleigh-benard convection in a cartesian cell with square domain. *Phys. Fluids*, **23**, 1–16.
- Baker, M. and G. Roe, 2009: The shape of things to come: why is climate change so predictable? *Journal of Climate*, **22** (17), 4574–4589.
- Blossey, P., C. Bretherton, and M. Wyant, 2009: Subtropical low cloud response to a warmer climate in a superparameterized climate model. part ii: Column modeling with a cloud resolving model. *J. Adv. Model. Earth Syst.*, **1**, 1–14.
- Brohan, P., J. J. Kennedy, I. Harris, S. F. B. Tett, and P. D. Jones, 2006: Uncertainty estimates in regional and global observed temperature changes: a new dataset from 1850. *J. Geophys. Res.*, **111**, 1–21.
- Budyko, M., 1969: The effect of solar radiation variations on the climate of the earth. *Tellus*, **21** (5), 611–619.
- Chekroun, M. D., D. Kondrashov, and M. Ghil, 2011: Predicting stochastic systems by noise sampling, and application to the el nino-southern oscillation. *PNAS*, 1–6.

- Cockburn, B., 2004: *Discontinuous Galerkin Methods for computational fluid dynamics*, Vol. 3, chap. 4, 91–127. John Wiley and Sons.
- Delworth, T. L., et al., 2006: GFDL’s CM2 global coupled climate models. Part I: Formulations and simulation characteristics. *J. Climate*, **19**, 643–674.
- DeMott, C., C. Stan, D. Randall, J. Kinter, and M. Khairoutdinov, 2011: The asian monsoon in the super-parameterized CCSM and its relationship to tropical wave activity. *J. Climate*, **24**, 5134–5156.
- Evensen, G., 1994: Sequential data assimilation with a nonlinear quasi-geostrophic model using monte carlo methods to forecast error statistics. *J. Geophys. Res.*, **99**, 10 143–10 162.
- Evensen, G., 2003: The Ensemble kalman filter: theoretical formulation and practical implementation. *Ocean Dynamics*, **53** (4), 343–367.
- Evensen, G., 2007: *Data Assimilation The ensemble Kalman filter*. Springer.
- Farrell, B. and P. Ioannou, 2001: State estimation using a reduced-order Kalman filter. *Journal of the Atmospheric Sciences*, **58** (23), 3666–3680.
- Forest, C. E., P. H. Stone, and A. P. Sokolov, 2006: Estimated pdfs of climate system properties including natural and anthropogenic forcings. *Geophys. Res. Lett.*, **33**, 1–4.
- Forest, C. E., P. H. Stone, A. P. Sokolov, M. R. Allen, and M. D. Webster, 2002: Quantifying uncertainties in climate system properties with the use of recent climate observations. *Science*, **295** (5552), 113–117.
- Forster, P., et al., 2007: Changes in atmospheric constituents and in radiative forcing. *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, S. Solomon, D. Qin, M. Manning, Z. Chen, M. Marquis, K. Averyt, M. Tignor, and H. Miller, Eds., Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.
- Friend, A., 2011: Response of earth’s surface temperature to radiative forcing over a.d. 1-2009. *J. Geophys. Res.*, **116** (1–15).
- Gelb, A., J. Kasper, R. Nash, C. Price, and A. Sutherland, 1974: *Applied Optimal Estimation*. M.I.T. Press.
- Ghil, M., M. D. Chekroun, and E. Simonnet, 2008: Climate dynamics and fluid mechanics: Natural variability and related uncertainties. *Physica D: Nonlinear Phenomena*, **237** (14–17), 2111–2126.
- Giorgi, F. and L. O. Mearns, 2002: Calculation of average, uncertainty range, and reliability of regional climate changes from AOGCM simulations via the “reliability ensemble averaging” (REA) method. *Journal of Climate*, **15** (10), 1141–1158.
- Grabowski, W., 2001: Coupling cloud processes with the large-scale dynamics using the cloud-resolving convection parameterization (CRCP). *J. Atmos. Sci.*, **58**, 978–997.
- Grabowski, W., 2004: An improved framework for superparameterization. *J. Atmos. Sci.*, **61**, 1940–1952.

- Grabowski, W. and P. Smolarkiewicz, 1996: Two-time-level semi-lagrangian modeling of precipitating clouds. *Mon. Wea. Rev.*, **124**, 487–497.
- Grabowski, W. and P. Smolarkiewicz, 1999: CRCP: a cloud resolving convection parameterization for modeling the tropical convecting atmosphere. *Physica D: Nonlinear Phenomena*, **133**, 171–178.
- Grabowski, W., X. Wu, and M. Monkrieff, 1996: Cloud-resolving modeling of tropical cloud systems during phase III of GATE. Part i: Two-dimensional experiments. *J. Atmos. Sci.*, **53** (24), 3684–3709.
- Greene, A. M., L. Goddard, and U. Lall, 2006: Probabilistic multimodel regional temperature change projections. *Journal of Climate*, **19** (17), 4326–4343.
- Gregory, J. M. and P. M. Forster, 2008: Transient climate response estimated from radiative forcing and observed temperature change. *J. Geophys. Res.*, **113**, 1–15.
- Hamill, T., 2006: Ensemble-based atmospheric data assimilation. *Predictability of Weather and Climate*, 124–156.
- Hansen, J., et al., 2007: Dangerous human-made interference with climate: a giss modele study. *Atmos. Chem. Phys.*, **7** (9), 2287–2312.
- Harvey, L. D. D. and R. Kaufmann, 2002: Simultaneously constraining climate sensitivity and aerosol radiative forcing. *J. Climate*, **15** (20), 2837–2861.
- Hasselmann, K., 1997: Multi-pattern fingerprint method for detection and attribution of climate change. *Climate Dynamics*, **13** (9), 601–611.
- Heemink, A., M. Verlaan, and A. Segers, 2001: Variance reduced ensemble Kalman filtering. *Monthly Weather Review*, **129** (7), 1718–1728.
- Hegerl, G., et al., 2007: Understanding and attributing climate change. *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, S. Solomon, D. Qin, M. Manning, Z. Chen, M. Marquis, K. Averyt, M. Tignor, and H. Miller, Eds., Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.
- Held, I., M. Winton, K. Takahashi, T. L. Delworth, F. Zeng, and G. K. Vallis, 2010: Probing the fast and slow components of global warming by returning abruptly to pre-industrial forcing. *J. Climate*, **23**, 2418–2427.
- Holmes, P., J. Lumley, and G. Berkooz, 1996: *Turbulence, coherent structures, dynamical systems and symmetry*. Cambridge University Press.
- Huybers, P., 2010: Compensation between model feedbacks and curtailment of climate sensitivity. *J. Climate*, **23**, 3009–3018.
- Julier, S., 2002: The scaled unscented transformation. *Proceedings of the American Control Conference*, IEEE, Vol. 6, 4555–4559.
- Jung, J. and A. Arakawa, 2005: Preliminary tests of multiscale modeling with a two-dimensional framework: sensitivity to coupling methods. *Mon. Wea. Rev.*, **133**, 649–662.

- Kalman, R., 1960: A new approach to linear filtering and prediction problems. *Transactions of the ASME—Journal of Basic Engineering*, **82 (Series D)**, 34–45.
- Kerstein, A., 1988: A linear- eddy model of turbulent scalar transport and mixing. *Combustion Science and Technology*, **60**, 391–421.
- Khairoutdinov, M. and D. Randall, 2001: A cloud resolving model as a cloud parameterization in the near community climate system model: preliminary results. *Geophys. Res. Lett.*, **28 (18)**, 3617–3620.
- Khairoutdinov, M., D. Randall, and C. DeMott, 2005: Simulations of the atmospheric general circulation using a cloud-resolving model as a superparameterization of physical processes. *J. Atmos. Sci.*, **62**, 2136–2154.
- Kiehl, J. T., 2007: Twentieth century climate model response and climate sensitivity. *Geophys. Res. Lett.*, **34**, 1–4.
- Knutti, R., 2008: Why are climate models reproducing the observed global surface warming so well? *Geophys. Res. Lett.*, **35**, 1–5.
- Knutti, R. and L. Tomassini, 2008: Constraints on the transient climate response from observed global temperature and ocean heat uptake. *Geophys. Res. Lett.*, **35 (9)**, 1–5.
- Knutti, R., et al., 2008: A review of uncertainties in global temperature projections over the twenty-first century. *Journal of Climate*, **21 (11)**, 2651–2663.
- Kuramoto, Y. and T. Tsuzuki, 1976: Persistent propagation of concentration waves in dissipative media far from thermal equilibrium. *Prog. of Theor. Phys.*, **55 (2)**, 356–369.
- Lermusiaux, P. F. J. and A. R. Robinson, 1999: Data assimilation via error subspace statistical estimation. Part I: Theory and schemes. *Monthly Weather Review*, **127 (7)**, 1385–1407.
- Lopez, A., C. Tebaldi, M. New, D. Stainforth, M. Allen, and J. Kettleborough, 2006: Two approaches to quantifying uncertainty in global temperature changes. *Journal of Climate*, **19 (19)**, 4785–4796.
- Lorenz, E., 1956: *Empirical orthogonal functions and statistical weather prediction*. MIT, Dept. of Meteorology.
- Lorenz, E., 2006: Predictability—A problem partly solved. *Predictability of weather and climate*, T. Palmer and R. Hagedorn, Eds., Cambridge University Press, 40–58.
- Lu, Z., T. K. Leen, R. van der Merwe, S. Frolov, and A. M. Baptista, 2007: Sequential data assimilation with sigma-point Kalman filter on low-dimensional manifold. Technical report TR-07-001, NSF-STC for Coastal Margin Observation and Prediction.
- Lumley, J., 1970: *Stochastic tools in turbulence*. Academic Press.
- Luo, X. and I. Moroz, 2009: Ensemble Kalman filter with the unscented transform. *Physica D: Nonlinear Phenomena*, **238**, 549–562.
- Majda, A., 2007: Multiscale models with moisture and systematic strategies for superparameterization. *J. Atmos. Sci.*, **64**, 2726–2734.

- Majda, A. and M. Grote, 2009: Mathematical test models for superparameterization in anisotropic turbulence. *PNAS*, **106**, 5470–5474.
- Majda, A., I. Timofeyev, and E. Eijnden, 2001: A mathematical framework for stochastic climate models. *Communications on Pure and Applied Mathematics*, **LIV**, 0891–0974.
- Marchand, R. and T. Ackerman, 2010: An analysis of cloud cover in multiscale modeling framework global climate model simulations using 4 and 1 km horizontal grids. *J. Geophys. Res.*, **115**, 1–19.
- Mauritsen, T., et al., 2012: Tuning the climate of a global model. *J. Adv. Model. Earth Syst.*, **4**, 1–18.
- Meehl, G., et al., 2007: *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, chap. Global Climate Projections, 747–845. Cambridge University Press.
- Meinshausen, M., S. C. B. Raper, and T. M. L. Wigley, 2008: Emulating IPCC AR4 atmosphere-ocean and carbon cycle models for projecting global-mean, hemispheric and land/ocean temperatures: MAGICC 6.0. *Atmospheric Chemistry and Physics Discussions*, **8** (2), 6153–6272.
- Morelande, M. and B. Ristic, 2006: Reduced sigma point filtering for partially linear models. *2006 IEEE International Conference on Acoustics, Speech and Signal Processing*, Vol. 3.
- Murphy, D. M., 2010: Constraining climate sensitivity with linear fits to outgoing radiation. *Geophys. Res. Lett.*, **37** (9), 1–5.
- Neggers, R., A. Siebesma, and T. Heus, 2012: Continuous single-column model evaluation at a permanent meteorological supersite. *BAMS*, **93** (9), 1389–1400.
- Padilla, L. and C. Rowley, 2010: An adaptive-covariance-rank algorithm for the unscented Kalman filter. *Decision and Control (CDC), 2010 49th IEEE Conference on*, 1324–1329.
- Padilla, L., G. K. Vallis, and C. Rowley, 2011: Probabilistic estimates of transient climate sensitivity subject to uncertainty in forcing and natural variability. *J. Climate*, **24**, 5521–5537.
- Pritchard, M., M. Moncrieff, and R. Somerville, 2011: Orographic propagating precipitation systems over the united states in a global climate model with embedded explicit convection. *J. Atmos. Sci.*, **68**, 1821–1840.
- Randall, D. and M. Khairoutdinov, 2010: *SP-CAM coupling*.
- Randall, D., M. Khairoutdinov, A. Arakawa, and W. Grabowski, 2003: Breaking the cloud parameterization deadlock. *BAMS*, 1547–1564.
- Randall, D., et al., 2007: *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, chap. Climate Models and Their Evaluation. Cambridge University Press, United Kingdom and New York, NY, USA.
- Raper, S. C. B., J. M. Gregory, and R. J. Stouffer, 2002: The role of climate sensitivity and ocean heat uptake on aogcm transient temperature response. *Journal of Climate*, **15** (1), 124–130.
- Roe, G. H. and K. Armour, 2011: How sensitive is climate sensitivity? *Geophys. Res. Lett.*, **38**, 1–5.

- Roe, G. H. and M. B. Baker, 2007: Why is climate sensitivity so unpredictable? *Science*, **318** (5850), 629–632.
- Rowlands, D., et al., 2012: Broad range of 2050 warming from an observationally constrained large climate model ensemble. *Nature Geosci.*, 1–5.
- Sanderson, B. and R. Knutti, 2012: On the interpretation of constrained climate model ensembles. *Geophys. Res. Lett.*, **39**, 1–6.
- Satoh, M., T. Matsuno, H. Tomita, H. Miura, T. Nasuno, and S. Iga, 2008: Nonhydrostatic icosahedral atmospheric model (nicam) for global cloud resolving simulations. *J. Comp. Phys.*, **227**, 3486–3514.
- Sellers, W., 1969: A global climatic model based on the energy balance of the earth-atmosphere system. *J. Appl. Meteor.*, **8**, 392–400.
- Simon, D., 2006: *Optimal state estimation: Kalman, H infinity and nonlinear approaches*. John Wiley and Sons.
- Sivashinsky, G., 1977: Nonlinear analysis of hydrodynamic instability in laminar flames—i. derivation of basic equations. *Acta Astronautica*, **4**, 1177–1206.
- Sivashinsky, G. and D. Michelson, 1980: On irregular wavy flow of a liquid film down a vertical plane. *Prog. of Theor. Phys.*, **63** (6), 2112–2114.
- Stainforth, D. A., et al., 2005: Uncertainty in predictions of the climate response to rising levels of greenhouse gases. *Nature*, **433** (7024), 403–406.
- Stan, C., M. Khairoutdinov, C. DeMott, T. Krishnamurti, D. Straus, D. Randall, J. Kinter, and J. Shukla, 2010: An ocean-atmosphere climate simulation with an embedded cloud resolving model. *Geophys. Res. Lett.*, **37**, 1–6.
- Stott, P. A. and J. A. Kettleborough, 2002: Origins and estimates of uncertainty in predictions of twenty-first century temperature rise. *Nature*, **416**, 723–726.
- Stott, P. A., J. F. B. Mitchell, M. R. Allen, T. L. Delworth, J. M. Gregory, G. A. Meehl, and B. D. Santer, 2006: Observational constraints on past attributable warming and predictions of future global warming. *Journal of Climate*, **19** (13), 3055–3069.
- Stott, P. A. and S. F. B. Tett, 1998: Scale-dependent detection of climate change. *J. Climate*, **11**, 3282–3294.
- Tao, W., et al., 2009: A multiscale modeling system; developments, applications, and critical issues. *BAMS*, 515–534.
- Tebaldi, C., R. L. Smith, D. Nychka, and L. O. Mearns, 2005: Quantifying uncertainty in projections of regional climate change: A bayesian approach to the analysis of multimodel ensembles. *Journal of Climate*, **18** (10), 1524–1540.
- Thompson, D. W. J., J. M. Wallace, P. D. Jones, and J. J. Kennedy, 2009: Identifying signatures of natural climate variability in time series of global-mean surface temperature: Methodology and insights. *Journal of Climate*, **22** (22), 6120–6141.

- Tomassini, L., P. Reichert, H. Kunsch, C. Buser, R. Knutti, and M. Borsuk, 2009: A smoothing algorithm for estimating stochastic, continuous time model parameters and its application to a simple climate model. *J. Royal Stat. Soc.*, **58** (5), 679–704.
- Uzunoglu, B., S. Fletcher, M. Zupanski, and I. Navon, 2007: Adaptive ensemble reduction and inflation. *Quarterly Journal of the Royal Meteorological Society*, **133** (626), 1281–1294.
- Vallis, G. K., E. P. Gerber, P. J. Kushner, and B. A. Cash, 2004: A mechanism and simple dynamical model of the north atlantic oscillation and annular modes. *Journal of the Atmospheric Sciences*, **61** (3), 264–280.
- van der Merwe, R., 2004: Sigma point Kalman filters for probabilistic inference in dynamic state space models. Ph.D. thesis, Oregon Health and Science University.
- Whitaker, J. and T. M. Hamill, 2002: Ensemble data assimilation without perturbed observations. *Mon. Wea. Rev.*, **130**, 1913–1924.
- Xing, Y., A. Majda, and W. Grabowski, 2009: New efficient sparse space-time algorithms for superparameterization on mesoscales. *Mon. Wea. Rev.*, **137**, 4307–4324.
- Xu, Y. and C.-W. Shu, 2006: Local discontinuous Galerkin methods for the Kuramoto-Sivashinsky equations and the Ito-type coupled KdV equations. *Computer methods in applied mechanics and engineering*, **195**, 3430–3447.
- Yang, X. and T. Delsole, 2009: Using the ensemble Kalman filter to estimate multiplicative model parameters. *Tellus A*, **61**, 601–609.