

ABSTRACT

Title of dissertation: SYNCHRONIZATION IN
CHAOTIC SYSTEMS:
COUPLING OF CHAOTIC MAPS,
DATA ASSIMILATION, AND
WEATHER FORECASTING

Seung-Jong Baek, Doctor of Philosophy, 2007

Dissertation directed by: Distinguished University Professor Edward Ott
Department of Electrical and Computer Engineering

The theme of this thesis is the synchronization of coupled chaotic systems. Background and introductory material are presented in Chapter 1.

In Chapter 2, we study the transition to coherence of ensembles of globally coupled chaotic maps allowing for ensembles of non-identical maps and for noise. The transition coupling strength is determined from a transfer function of the perturbation evolution. Analytical results are presented and tested using numerical experiments. One of our examples suggests that the validity of the perturbation theory approach can be problematic for an ensemble of noiseless identical ‘nonhyperbolic’ maps, but can be restored by noise and/or parameter spread.

The problem of estimating the state of a large evolving spatiotemporally chaotic system from noisy observations and a model of the system dynamics is studied in Chapters 3 – 5. This problem, referred to as ‘data assimilation’, can be

thought of as a synchronization problem where one attempts to synchronize the model state to the system state by using incoming data to correct synchronization error.

In Chapter 3, using a simple data assimilation technique, we show the possible occurrence of temporally and spatially localized bursts in the estimation error. We discuss the similarity of these bursts to those occurring at the ‘bubbling transition’ in the synchronization of low dimensional chaotic systems.

In general, the model used for state estimation is imperfect and does not exactly represent the system dynamics. In Chapter 4 we modify an ensemble Kalman filter scheme to incorporate the effect of model bias for large chaotic systems based on augmentation of the system state by the bias estimates, and we consider different ways of parameterizing the model bias. The experimental results highlight the critical role played by the selection of a good parameterization model for representing the form of the possible bias in the model.

In Chapter 5 we further test the method developed in Chapter 4 via numerical experiments employing previously developed codes for global weather forecasting. The results suggest that our method can be effective for obtaining improved forecasting results when using an ensemble Kalman filter scheme.

SYNCHRONIZATION IN CHAOTIC SYSTEMS:
COUPLING OF CHAOTIC MAPS, DATA ASSIMILATION, AND
WEATHER FORECASTING

by

Seung-Jong Baek

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2007

Advisory Committee:

Professor Edward Ott, Chair/Advisor
Professor Eyad H. Abed
Professor Thomas M. Antonsen
Professor Brian R. Hunt
Professor Rajarshi Roy
Doctor Istvan Szunyogh

This dissertation is dedicated to my mother and father, and to my wife.

ACKNOWLEDGMENTS

I would like to acknowledge many people for helping me during my graduate study. My deepest gratitude is to my advisor, Prof. Edward Ott. His patience and support helped me to finish my doctoral research. He encouraged me to develop analytical thinking and research skills, and greatly assisted me with scientific writing. I am also very grateful for working with Profs. Brian Hunt and Istvan Szunyogh. Along with my advisor, they have patiently reviewed draft after draft of every paper that I wrote up, and provided me with deep insight on research projects.

I would like to thank my committee members, Profs. Thomas Antonsen, Eyad Abed, and Rajarshi Roy, for their valuable comments and suggestions. I also wish to thank Drs. Gyorgyi Gyarati, Aleksy Zimin, Takemasa Miyoshi, John Harlim, Elana Fertig, and Hong Li for helping me to write computer programs to perform various numerical experiments. They also generously gave me their time to discuss problems on my programs and work, and encouraged me to keep pursuing my doctoral research.

Finally, I thank my parents and my sister for their encouragement and endless support. I am also extremely grateful to Yoon Namkung, my wife and best friend, for her patience and for helping me keep my life in proper perspective and balance. They always have faith in me despite their incomprehension of my work.

TABLE OF CONTENTS

List of Figures	vi
1 Introduction	1
1.1 Synchronization of a large number of heterogeneous globally coupled chaotic maps	1
1.2 Error bursts in data assimilation for estimating the state of a spatiotemporal chaotic model	3
1.3 The application of state augmentation for correcting model errors in ensemble Kalman filters	4
1.4 Application of the error correction scheme to a real weather forecasting environment	5
2 Onset of Synchronization in Systems of Globally Coupled Chaotic Maps	7
2.1 Introduction	7
2.2 System Model and Analysis	10
2.2.1 System Model	10
2.2.2 Stability Analysis	12
2.3 Examples: One-Dimensional Maps	16
2.3.1 An Ensemble of Shifted Bernoulli Maps	16
2.3.2 A Heterogeneous Ensemble of Modified Bernoulli Maps	20
2.3.3 An Ensemble of Logistic Maps	23
2.3.4 An Ensemble of Logistic Maps with Noise	26
2.3.5 A Heterogeneous Ensemble of Logistic Maps	28
2.4 Examples: Two-Dimensional Maps	29
2.4.1 Multidimensional Extension of the Theory	29
2.4.2 An Ensemble of Cat Maps	32
2.5 Conclusion	33
3 Localized Error Bursts in Estimating the State of Spatiotemporal Chaos	35
3.1 Introduction	35
3.2 Experimental Design	37
3.3 Numerical Experiments	41
3.3.1 Experiment 1	41
3.3.2 Experiment 2	46
3.4 Conclusion and Discussion	50
4 Local Ensemble Kalman Filtering in the Presence of Model Bias	52
4.1 Introduction	52
4.2 Bias Modeling and Data Assimilation	55
4.2.1 Bias Model I	56
4.2.2 Bias Model II	59
4.2.3 Bias Model III	62
4.2.4 Augmented Local Ensemble Kalman Filter	63

4.3	Numerical Experiments	65
4.3.1	Experimental Setup	65
4.3.2	Perfect Forecast Model	70
4.3.3	Data Assimilation with Type A Truth Bias	72
4.3.4	Data Assimilation with Type B Truth Bias	77
4.3.5	Data Assimilation with Type C Truth Bias	79
4.3.6	Settling Time	79
4.3.7	A Simple State Dependent Model Error	83
4.4	Conclusions and Discussion	85
5	Surface pressure bias correction	88
5.1	Introduction	88
5.2	Data assimilation with bias correction	91
5.2.1	Bias modeling	91
5.2.2	Bias correction with LETKF	94
5.3	Numerical experiments	98
5.3.1	Experimental setup	98
5.3.2	Data assimilation without bias correction	103
5.3.3	Bias correction for surface pressure	108
5.4	Conclusions and discussion	116
A	Decay of Q_m with Increasing m	119
B	Orbit Densities for Ensembles of Logistic Maps	121
C	Background Error Covariance Matrix	123

LIST OF FIGURES

2.1	Results for an ensemble of $N = 10^5$ Bernoulli maps with and without noise: (a) shows $\langle a^2 \rangle$ versus k . The time average $\langle a^2 \rangle$ is computed using 1000 iterations. The power spectral densities of a_n at the positive critical values are shown in (b) for the cases without noise ($k = k_{c1}^{(0)}$, solid graph) and with noise ($k = k_{c1}$, dash-dot graph). The predicted values of critical coupling strengths and frequencies of oscillation agree with the experimental results.	17
2.2	Linear scale plot of $\langle a^2 \rangle$ and orbits of a_n for an ensemble of 10^5 noisy Bernoulli maps: (a) shows $\langle a^2 \rangle$ versus k (dots) and a quadratic curve fit (solid line) to the data in the range $1.5 \leq k \leq 1.8$. (b) shows orbits of a_n for $k = 1.53 > k_{c1} \approx 1.49$ (points in halo about origin) and $k = 1.45 < k_{c1}$ (points clustered near origin). $\langle a^2 \rangle$ approaches zero linearly and the orbits encircle the origin consistent with a Hopf bifurcation.	18
2.3	Results for ensembles of $N = 10^4, 10^5, 10^6, 10^7$ Bernoulli maps: (a) shows $\langle a^2 \rangle$ versus k . (b) shows $\overline{\langle a^2 \rangle}$ versus N . Within the incoherent state $\langle a^2 \rangle$ varies as N^{-1}	19
2.4	$Q(e^{i\omega})$ for an ensemble of $N = 10^6$ modified Bernoulli maps with uniformly distributed parameters: (a) shows six numerical approximations to Q_m . The six approximations are near zero at $m = 4, 5$, but then diverge from each other due to the combined effect of chaotic dynamics and finite ensemble size. (b) shows $Q(e^{i\omega})$ evaluated by averaging the six results from (a) for $m = 0 - 4$ and taking $Q_m = 0$ for $m \geq 5$. The imaginary part of $Q(e^{i\omega})$ has four zero-crossing points and we label three of them and corresponding real parts of these three points. The real parts corresponding to ω_{c11} and ω_{c12} are close to each other.	22
2.5	Results for ensembles of $N = 10^5, 10^6$ modified Bernoulli maps with uniformly distributed parameters: (a) shows $\langle a^2 \rangle$ versus k , (b) is the power spectral density of a_n at $k = 2.48 \approx k_{c1}$	23
2.6	Q_m for an ensemble of $N = 10^9$ logistic maps with parameter $\mu = 3.9$: (a) shows Q_m up to $m = 17$ and (b) shows Q_m plotted up to $m = 35$. We see that our approximations do not converge to small values before diverging from each other.	24

2.7	$\langle a^2 \rangle$ versus k for ensembles of logistic maps: Data for ensembles of sizes $N = 10^4, 10^5, 10^6$ are shown in (a). We plot $\langle a^2 \rangle$ for ensembles of $N = 10^5, 10^6, 10^7$ logistic maps versus k in a narrower range of k in (b). We see that the N dependence is confined to a very small region near $k = 0$ and that, as N increases, the confined region becomes narrower.	25
2.8	$Q(e^{i\omega})$ for an ensemble of $N = 10^6$ noisy logistic maps with parameter $\mu = 3.9$ and normally distributed noise: (a) shows three numerical approximations to Q_m . The three approximations are near zero at $m = 8, 9$, but then diverge from each other. (b) shows $Q(e^{i\omega})$ evaluated from (a) assuming that $Q_m = 0$ after $m = 8$	26
2.9	Experimental results for an ensemble of $N = 10^4, 10^5$ noisy logistic maps with parameter $\mu = 3.9$ and normally distributed noise: (a) shows a semilogarithmic plot of $\langle a^2 \rangle$ versus k , and (b) shows $\langle a^2 \rangle$ (dots) and a quadratic curve fit to the numerical data in the range $0.29 \leq k \leq 0.44$ (solid line) for $N = 10^5$	28
2.10	Power spectral density and orbits of a_n for $N = 10^4$: The frequency of oscillation and the power spectral density of a_n at $k = k_{c1}$ are plotted in (a). (b) shows orbits of a_n for $k = 0.30 > k_{c1} \approx 0.29$ (points in halo about origin) and $k = 0.28 < k_{c1}$ (points clustered near origin).	29
2.11	$Q(e^{i\omega})$ for a noiseless ensemble of $N = 10^9$ logistic maps with parameter μ uniformly distributed in the interval (3.88, 3.96): (a) shows three numerical approximations to Q_m . The three approximations are near zero at $m = 12$, but then diverge from each other. (b) shows $Q(e^{i\omega})$ evaluated from (a) assuming that $Q_m = 0$ after $m = 12$	30
2.12	Experimental results for a noiseless ensemble of $N = 10^4, 10^5, 10^6$ logistic maps with parameter μ uniformly distributed in the interval (3.88, 3.96): (a) shows $\langle a^2 \rangle$ versus k , and (b) shows the power spectral density of a_n at $k = 0.24 \approx k_{c1}$ for $N = 10^6$	31
2.13	Experimental results for an ensemble of $N = 10^5$ cat maps with and without noise: (a) shows k versus $\langle a^2 \rangle$ for the cases with noise and without noise. (b) show the power spectral densities of a_n at the positive critical values for the cases without noise ($k = k_{c1}^{(0)}$, solid graph) and with noise ($k = k_{c1}$, dash-dot graph).	33
3.1	Illustration of the Lorenz-96 model.	38
3.2	Time evolution of the “true state” for a randomly selected short time interval.	42

3.3	Time evolution of the rms analysis error. Also shown is the time mean analysis error (white dashed line), $\langle\langle\mathbf{e}^a\rangle\rangle = 0.0093$	43
3.4	Spatial distribution of the background error (a) and the analysis error (b) at a given time during a burst event. Note that the difference between the scales of the y-axes.	43
3.5	$\langle\rho(w, \mathbb{S}_1)\rangle$ (solid) and $\langle\rho(w, \mathbb{S}_2)\rangle$ (dashed).	44
3.6	The spatial distribution of the burst events.	45
3.7	The 19-th ($K_{19,k}$) and 20-th ($K_{20,k}$) row of the gain matrix \mathbf{K} for Experiment 2.	47
3.8	Time evolution of the rms analysis error. Also shown is the time mean analysis error (white dashed line), $\langle\langle\mathbf{e}^a\rangle\rangle = 0.0103$	47
3.9	The spatial distribution of the burst events. (b) has a different vertical scale from (a) in order to see the burst events at other locations. . . .	48
3.10	Spatial distribution of the background error (a) and the analysis error (b) at a given time during a burst event.	49
3.11	$\langle\rho(w, \mathbb{S}_1)\rangle$ (solid) and $\langle\rho(w, \mathbb{S}_2)\rangle$ (dashed).	49
4.1	Illustration of Bias Model I: \mathbf{x}_n^t is the true state evolving according to \mathbf{F}^t from the previous true state, and \mathbf{x}_n^m is the forecast model state evolving according to \mathbf{F}^m from the previous true state.	57
4.2	Illustration of data assimilation with Bias Model I: Data assimilation produces an unbiased analysis for the true state, \mathbf{x}_n^a , and an analysis for the bias correction term, \mathbf{b}_n^a	58
4.3	Illustration of Bias Model II: \mathbf{x}_n^t and \mathbf{x}_n^m evolve according to their own dynamics but the behavior of the forecast model is similar to the behavior of the truth.	60
4.4	Illustration of data assimilation with Bias Model II: Data assimilation produces an analysis of the best forecast model state, $\mathbf{x}_n^m = \mathbf{x}_n^t - \mathbf{c}_n^t$, and an analysis for the correction term, \mathbf{c}_n^t	62
4.5	Illustration of a local state centered about location m	64

4.6	Time averaged rms analysis error, $\langle\langle\mathbf{e}^a\rangle\rangle$, versus variance inflation coefficient, μ , for the perfect forecast model experiment: With a perfect forecast model, any attempt to estimate and correct for a bias results in slightly higher analysis error.	71
4.7	Time averaged rms analysis error, $\langle\langle\mathbf{e}^a\rangle\rangle$, versus μ for the case of Type A truth bias: Note that (b) shows the same results as (a) for Bias Model I and Bias Model III but for a different vertical scale. . .	74
4.8	The average bias estimate of location i is shown as \circ . The approximate true bias $\beta_i\Delta t$ is shown as the solid curve.	75
4.9	Time average of the analysis error as a function of location: The Type A truth bias is corrected best when we perform the assimilation using Bias Model III.	75
4.10	Figure (a) depicts the case $\mathbf{c} = \mathbf{0}$, corresponding Bias Model I, while figure (b) depicts the case $\mathbf{c} \approx -\beta/2$ in Bias Model III.	76
4.11	The average bias estimate of location i is shown as $+$. The the value of $-\beta_i\Delta t/2$ is shown as the solid curve.	77
4.12	Time averaged rms analysis error, $\langle\langle\mathbf{e}^a\rangle\rangle$, versus μ for the case of Type B truth bias. (b) has a different vertical scale from (a) for transparency.	78
4.13	Bias estimate for the Bias Model II is shown as \circ . The true bias (4.49) is shown as the solid curve.	78
4.14	Time averaged rms analysis error, $\langle\langle\mathbf{e}^a\rangle\rangle$, versus μ for the case in which the truth has Type C truth bias: In order to correct for the biases, the augmented state used in the assimilations must contain both the \mathbf{b} and \mathbf{c} bias estimates, Here, we again use a different vertical scale in (b) from (a) for transparency.	80
4.15	Time averaged innovation for the case that no bias estimation is performed with various biases in the truth.	81
4.16	Time averaged rms analysis error, $\langle\langle\mathbf{e}^a\rangle\rangle$, versus μ with various α_b and $\alpha_c = 0$ for the case of Type C truth bias corrected using Bias Model III: Small diffusion improves the performance of the assimilation up to $\alpha_b = 0.01$	83
4.17	Time averaged rms analysis error, $\langle\langle\mathbf{e}^a\rangle\rangle$, versus μ with the simple state dependent model error: With Bias Model I and diffusion process, we can improve the performance in terms of the rms analysis error.	84

5.1	Illustration of local region: Dots represent grid points and the grey box represent a local region. (a) is the top view of the local region and (b) is the side view of it. The local region in the figure is a $3 \times 3 \times 1$ box. The black dot in the local region, \mathbf{r} , is the grid point whose states are estimated in the assimilation time using the states and the observation within this local region.	95
5.2	6-hour forecast bias of the surface pressure without bias correction: In many regions, the amount of the bias is proportional to the magnitude of the orography difference but sometimes they are negatively correlated. There are also regions that shows small difference in orography but have moderate amount of the biases.	99
5.3	Orography difference between SPEEDY and NCEP models: Surface geopotential height of NCEP model is subtracted from that of SPEEDY model. Discrepancies are mostly on mountain areas. The scales are different in negative and positive directions.	99
5.4	Correlation between orography difference and surface pressure bias.	99
5.5	Local observation areas at latitude 31.5° between longitude 96° and 120° : The lowest line with circle \circ represents the surface pressure (PS) and above it, four sigma level lines are depicted. Along the sigma levels, we put symbols to indicate the grid point where the model stats are estimated. Symbol X's indicate the available observations which are at levels 925hPa, 850hPa, 700hPa, 500hPa, and 300hPa above the surface. Two $3 \times 3 \times 1$ local regions are shown as a dashed line and a dash-dotted line. We see that they are different from the region in Fig. 5.1(b) and each region contains different number of observations due to the different coordinate space between model space and observation space.	101
5.6	Surface pressure analysis and 6-hour forecast rms errors and biases versus variance inflation without bias correction: The analysis error of surface pressure is getting smaller as the variance inflation is getting larger. However, the 6-hour forecast error remains almost the same.	104
5.7	Zonal wind analysis rms errors (a), 6-hour forecast rms errors (b), analysis biases (a), and 6-hour forecast biases (b) without bias correction: Similar to the case of the surface pressure, the analysis errors become smaller as the covariance inflation coefficient increases but the forecast errors stay the same after $\gamma = 1.0$	106

5.8 Temperature analysis rms errors (a), and 6-hour forecast rms errors (b), analysis biases (c), and 6-hour forecast biases (d) without bias correction: The forecast rms error, the analysis bias and the forecast bias of the temperature become smaller as the covariance inflation coefficient increases. For the inflation coefficient greater than 1.0, the analysis rms error of the temperature is greater than the forecast rms error. 107

5.9 6-hour forecast bias of the surface pressure with type II surface pressure bias correction: Most of the large biases existed in the experiments in Sec. 5.3.2 are mitigated to the magnitude less than 1 hPa except for the polar areas. 111

5.10 Time averages of analysis and 6-hour forecast rms errors of the zonal wind (a) and rms values of the analysis and 6-hour forecast bias of the zonal wind (b) with and without surface pressure bias correction: Solid lines represent the analysis rms errors and biases, and dashed lines represent the forecast rms errors and biases. The accuracy of the forecast is more improved than the accuracy of the analysis. At the highest level, the analysis error is larger than the case that no bias correction is performed. 112

5.11 Time averages of analysis and 6-hour forecast rms errors of the temperature (a) and rms values of the analysis and 6-hour forecast bias of the temperature (b) with and without surface pressure bias correction: Solid lines represent the analysis rms errors and biases, and dashed lines represent the forecast rms errors and biases. The accuracy of the analysis is better than the accuracy of the forecast. At the highest level, the biases of the analysis and the forecast are both degraded. 113

5.12 Rms, global spread and global average of the bias estimate: In a few time steps, the bias estimate captures the global structure of the bias in the surface pressure and slowly push the surface pressure to a high air mass state where the SPEEDY model can provide a dynamics similar to that of the NCEP model. 115

B.1 The histogram approximated orbit densities in the interval $[0.35, 0.55]$ for ensembles of 10^7 identical logistic maps with parameter $\mu = 3.9$ with the initial perturbation $\delta x_0 = 10^{-3}$ (grey) and without perturbation (black): We set time t to 0 at which the perturbation is applied. (a) and (b) show the densities at $t = 4$ and $t = 10$ respectively. . . . 122

C.1	Convergence of the Frobenius norms of P_k^b and $P_k^b - P_{k-1}^b$ for Experiment 1.	123
C.2	P^b after 40 iteration steps for Experiment 1.	124
C.3	P^b after 40 iteration steps for Experiment 2.	124

Chapter 1

Introduction

The overall theme of this thesis is synchronization of coupled dynamical systems, especially where the elements of the systems are chaotic. Four issues that involve this general theme will be addressed.

- (i) Synchronization of a large number of heterogeneous globally coupled chaotic maps.
- (ii) Error bursts in data assimilation for estimating the state of a spatiotemporal chaotic model.
- (iii) The application of state augmentation for correcting model errors in ensemble Kalman filters for estimating the state of spatiotemporal chaotic systems.
- (iv) Application of the results in part (iii) to a real weather forecasting environment.

1.1 Synchronization of a large number of heterogeneous globally coupled chaotic maps

First, we consider systems of large ensembles of globally coupled dynamic systems [6]. There are abundant examples of synchronous behavior of dynamical systems observed in nature including the sinoatrial node in hearts [62], networks of neurons [77, 101, 78, 43, 99], suspensions of yeast cells [22], flashing fireflies [12, 42, 13], chirping crickets [93, 83], as well as observations in laboratories including those on arrays of electrochemical oscillators [94, 52], semiconductor laser arrays [53], and Josephson junction arrays [97]. It is known that these large ensembles of globally

coupled dynamic units experience a transition to a coherent state at some critical coupling strength. In particular, for low coupling the individual units essentially evolve independently (the “incoherent state”), but, as the coupling passes a critical value, the average over all units of an appropriate function of the unit state begins to take on a macroscopic value, signaling the onset of coherent, synchronous behavior of the system. The transition from incoherence to synchronization with increase of the coupling has been very extensively studied in the case where the uncoupled dynamics is *periodic*.

Recently, attempts to study the transition from incoherence to synchrony in the case where the individual units are *chaotic* systems of ordinary differential equation have been made [75, 79, 69]. Pikovsky et al. [75] numerically show that the transition to coherent behavior in a system of globally coupled non-identical Rössler systems is due to the synchronization of phases of the individual units. They concentrate on the case that the individual systems have phase-coherent attractors and think of the transition as the synchronization transition in a system of coupled noisy limit-cycle oscillators. However, they do not develop a theory for the transition to synchrony. Sakaguchi [79] analyzes a large ensemble of globally coupled identical Rössler systems by assuming that, at the onset of coherence, the average motion of the ensemble is sinusoidal in time. Ott et al. [69] investigated the stability of the incoherent state for ensembles of globally coupled continuous-time dynamical systems using a perturbation method and analytic continuation. They numerically applied their theory to ensembles of globally coupled heterogeneous Lorenz systems with the parameters uniformly distributed in periodic, chaotic and mixed parameter regions. Refs. [75, 79, 69] all treated ensembles of *noise-free* chaotic systems.

Concerning large systems of globally coupled chaotic *maps*, starting with the work of Kaneko [50], it has been known that, depending on parameters, these systems can exhibit either incoherent behavior, or system-wide coherent oscillatory

(i.e., synchronized) behavior, and work has been done investigating the nature of these behaviors [50, 72, 74, 81, 82, 16, 89]. However, only one paper (Topaj et al. [89]) has so far given an analytical treatment of the transition from incoherence to coherence for a large globally coupled map system. This paper, however, treats only the very special case of identical $2x \bmod 2\pi$ maps. In Chapter 2, we establish a general method to determine the critical coupling strength and the frequency of oscillation at the onset of coherence for general chaotic maps.

1.2 Error bursts in data assimilation for estimating the state of a spatiotemporal chaotic model

Next, we consider the problem of estimating the state of an evolving spatiotemporally chaotic system from noisy observations of the system state and a model of the system dynamics [7, 8]. Spatiotemporal chaos refers to systems that exhibit complex behavior both in time and space. There are many examples of such systems, including the Earth's atmosphere [21, 49], laboratory fluids [19, 64], chemical systems [28], oceans [76], etc. We view this estimation problem as a process in which one attempts to synchronize the state of the model system to the state of the observed system.

Estimating the current state of the system is a key first step in both control and prediction of a system's behavior: One can predict the future of the system by running a model of the system with the current estimated system's state as the model's initial condition, or one can attempt to control the system by introducing active feedback perturbations whose strength and configuration are appropriately formulated based upon knowledge of the system state. Thus, the performance of the prediction or the control is heavily dependent on the accuracy of estimates of the current state of the system. Such estimations are particularly nontrivial for high-

dimensional spatiotemporal systems, for which the entire state cannot be measured directly (e.g., in weather forecasting). Data assimilation is the process by which an estimate of the system is obtained through the observations of the system and a model for the system [21, 49]. Through data assimilation one endeavors to accurately synchronize the state of the model system to the state of the observed system.

Typically, the observations have errors and the model is imperfect. As a result, the model system may not always be in accurate synchrony with the observed system, and occasionally large discrepancies from the observed system may occur. In the synchronism of low dimensional chaotic systems, analogous desynchronization events are known to occur as a result of the so-called ‘bubbling transition’ [4, 102] in which intermittent bursts of desynchronization are observed. We find that similar events are possible in high dimensional spatiotemporally chaotic systems, and that the resulting intermittent bursts are localized in small spatial regions. In Chapter 3, we show these phenomena through numerical experiments.

1.3 The application of state augmentation for correcting model errors in ensemble Kalman filters

Considering the problem of estimating the current state of a spatiotemporally chaotic system, one typically has to use an imperfect model. We refer to *model error* as the difference between the dynamics of a real process and the dynamics of the best available model of that process. Model error is thought to be a key issue in weather forecasting.

Kalman filters have been considered for estimating atmospheric states to be used as initial conditions in forecast models [34]. In Kalman filter methodology, model error has been treated for a general setting [33] and for the weather prediction [24]. Recently, the ensemble technique has been proposed as a computationally

feasible means of applying Kalman filtering to the very high dimensional states inherent in global atmospheric models [30]. One of our goals in this work is to investigate the incorporation of model error correction in an ensemble Kalman filter. For this purpose we will restrict our considerations to the example of one particular ensemble Kalman filter, the Local Ensemble Kalman Filter (LEKF) proposed by Ott et al. [70].

The technique we propose belongs to the family of schemes usually called *state space augmentation* methods. In these techniques the state vector is augmented with the uncertain model parameters, and the augmented state is estimated using data assimilation. The augmented state space approach may provide an accurate estimate of the parameters even for a highly chaotic system, as recently demonstrated on a simple model by Annan and Hargreaves [3]. A practical approach for weather prediction, first suggested by Derber [26], is to assume that the errors (bias) can be approximately represented in the form of a limited number of bulk error terms. Then, the task is to estimate the parameters of the bulk error terms. We present the results of the numerical experiments which show that our scheme can estimate biases in Chapter 4.

1.4 Application of the error correction scheme to a real weather forecasting environment

Model error is thought to be a key issue in Numerical Weather Prediction (NWP) in that the presence of the model error can lead to cause large discrepancies between the forecasts and the true atmospheric states. The difference between the model evolution and the true atmospheric evolution stems from such causes as the forecast model state being a finite dimensional representation of the infinite dimensional atmospheric fields and the governing atmospheric equations and associated

parameters not being fully known. Additional sources include orography differences and other boundary conditions that may not be known exactly.

Using our approach discussed in Sec. 1.3, we investigate the potential benefit of bias correction for a case in which there are biases in the surface pressure by numerical experiments simulating the situation of the numerical weather prediction. We apply our model bias correction technique to the data assimilation scheme of Ott et al. [70] as formulated by Hunt et al. [47] called the Local Ensemble Transform Kalman Filter. It is found that by using the new bias model for the surface pressure biases, the accuracies of the forecasts of the surface pressure, temperature, meridional wind and zonal wind are all improved. Detailed results are presents in Chapter 5.

Chapter 2

Onset of Synchronization in Systems of Globally Coupled Chaotic Maps

2.1 Introduction

The onset of synchronization in large ensembles of globally coupled dynamical units is of interest in many fields [98, 65, 86, 54]. Systems of this type have been examined in biology, where it is thought that rhythms are essential for maintaining life. In these biological situations rhythms are typically generated by a large number of cells or groups of cells, each one of which has a tendency to oscillate when isolated [35, 36]. For example, heartbeats are stimulated by the sinoatrial (SA) node located on the right atrium, and the node consists of thousands of coupled pacemaker cells [35, 36, 62]. Each cell in the SA node has slightly different intrinsic frequency and, through electrical coupling, achieves a consensus as to when to fire [62]. Similar mechanisms are observed in networks of neurons [37, 77, 101, 78, 43], coupled neurons in the suprachiasmatic nucleus (the circadian center) [99], and suspensions of yeast cells [22]. Insects also exhibit synchronized behavior [12, 42, 13, 93, 83]. For example, it has been observed that a large number of certain types of fireflies flash on and off in unison. They apparently watch each other and adjust their flashing according to their neighbors. A small group of fireflies starts to flash synchronously, and the number of synchronized participants grows so that the whole swarm finally flashes in unison [12, 42, 13]. Swarms of crickets and grasshoppers also chirp in unison through a similar process [93, 83]. Nonbiological examples of synchrony in large systems of globally coupled dynamical units occur in arrays of globally cou-

pled chaotic electrochemical oscillators [94, 52], semiconductor laser arrays [53], and Josephson junction arrays [97]. It is known that these systems experience a transition to a coherent state at some critical coupling strength. In particular, for low coupling the individual units essentially evolve independently (the “incoherent state”), but, as the coupling passes a critical value, the average over all units of an appropriate function of the unit state begins to take on a macroscopic value, signaling the onset of coherent, synchronous behavior of the system. The transition from incoherence to synchronization with increase of the coupling has been very extensively studied in the case where the uncoupled dynamics is *periodic* [98, 65, 86, 54]. In particular, the well-studied ‘Kuramoto model’ [54] considers many periodic oscillators whose uncoupled dynamics is described by the simple phase evolution equation, $d\theta^{(i)}(t)/dt = \omega^{(i)}$, for the phase $\theta^{(i)}$ of oscillator i , which has natural frequency $\omega^{(i)}$.

Very recently attempts to study the transition from incoherence to synchrony in the case where the individual units are *chaotic* systems of ordinary differential equation have been made [75, 79, 69]. Pikovsky et al. [75] numerically show that the transition to coherent behavior in a system of globally coupled non-identical Rössler systems is due to the synchronization of phases of the individual units. They concentrate on the case that the individual systems have phase-coherent attractors and think of the transition as the synchronization transition in a system of coupled noisy limit-cycle oscillators. However, they do not develop a theory for the transition to synchrony. Sakaguchi [79] analyzes a large ensemble of globally coupled identical Rössler systems by assuming that, at the onset of coherence, the average motion of the ensemble is sinusoidal in time. Ott et al. [69] investigated the stability of the incoherent state for ensembles of globally coupled continuous-time dynamical systems using a perturbation method and analytic continuation. They numerically applied their theory to ensembles of globally coupled heterogeneous Lorenz systems with the parameters uniformly distributed in periodic, chaotic and mixed parameter

regions. Refs. [75, 79, 69] all treated ensembles of *noise-free* chaotic systems.

Concerning large systems of globally coupled chaotic *maps*, starting with the work of Kaneko [50], it has been known that, depending on parameters, these systems can exhibit either incoherent behavior, or coherent oscillatory (i.e. synchronized) behavior, and work has been done investigating the nature of these behaviors [50, 72, 74, 81, 82, 16]. However, only one paper (Topaj et al. [89]) has so far given an analytical treatment of the transition from incoherence to coherence for a large globally coupled map system. This paper, however, treats only the special case of identical $2x \bmod 2\pi$ maps.

In the present paper, we present a general method to determine the critical coupling strength and the frequency of oscillation at the onset of coherence in a system of globally coupled chaotic maps. Our analysis allows for arbitrary map functions, the inclusion of noise, and the treatment of ensembles that are heterogeneous (i.e., it is not required that all maps in the ensemble are the same). Also, we investigate techniques for numerical application of the analysis. Our analysis adapts to *discrete-time* (maps) the perturbation method which Ott et al. [69] applied to systems of globally coupled *continuous-time* dynamical systems. The goal is to relate the evolution of a perturbation of the individual uncoupled elements to the evolution of a perturbation of the globally coupled system. Thus, we represent the behavior of coupled systems in terms of the behavior of uncoupled elements.

The system model and the analysis of the system is presented in Section 2.2. One-dimensional map examples employing and testing the analysis are given in Section 2.3. One result of the numerical experiments in Section 2.3.3 is that there is an apparent failure of the perturbation theory in a case of an ensemble of identical noiseless nonhyperbolic maps (in particular, the logistic map). However, the introduction of noise or of parameter spread appears to have a regularizing effect which seems to restore the validity of the perturbation theory approach. Section 2.4

gives an example for an ensemble of two dimensional maps. The examples in Section 2.3 and 2.4 contain ensembles of shifted Bernoulli maps, ensembles of modified Bernoulli maps in which a parameter of the map is uniformly distributed throughout an interval, ensembles of identical logistic maps, ensembles of identical noisy logistic maps, ensembles of noiseless logistic maps with a uniform distribution of the map parameter in an interval, and ensembles of cat maps.

2.2 System Model and Analysis

2.2.1 System Model

In this section we study systems of globally coupled, chaotic, one dimensional maps similar to the system studied in Ref. [89]. (The generalization to higher dimensional maps is made in Sec. 2.4.) In our theory, we assume that there is a mixing chaotic attractor so that almost every orbit of the uncoupled system yields the same long-time statistical behavior. The general form of the system we consider is

$$\begin{aligned} x_{n+1}^{(i)} &= f(x_n^{(i)}, \mu_i) + w_n^{(i)} + kg(x_n^{(i)})(\langle q(x_n) \rangle - \langle q(x) \rangle_*), \quad i = 1, \dots, N \\ \langle q(x_n) \rangle &= \frac{1}{N} \sum_{i=1}^N q(x_n^{(i)}). \end{aligned} \quad (2.1)$$

The quantity $w_n^{(i)}$ in Eqs. (2.1) is a random noise where $E[w_n^{(i)}] = 0$, $E[w_n^{(i)} w_k^{(j)}] = \sigma^2 \delta_{nk} \delta_{ij}$, $E[\cdot]$ denotes the expectation value, σ^2 is the variance of the noise, and for each n and i the $w_n^{(i)}$ are identically distributed. Moreover, we assume that x_n and w_n are independent. We note that additive noise makes the time averaged orbit probability distribution function of an uncoupled map smooth so that the distribution has no singularities [80, 67, 68], and noise may also eliminate small periodic windows within the chaotic parameter region. The maps $f(x, \mu_i)$ are assumed to

arise from a one parameter family with parameter μ_i , and we assume that the μ_i are distributed in some specified manner. For example, in our examples (Sec. 2.3) we will consider the case where the μ_i are distributed uniformly in an interval, and also the case of identical maps where all the μ_i are the same ($\mu_i = \mu$). Also in Eqs. (2.1), N is the ensemble size, k is the coupling coefficient, and the function $g(x)$ and $q(x)$ are assumed to be smooth and bounded. The symbol $\langle \cdot \rangle$ denotes the average over the ensemble (the average over i) at a fixed time n , and thus depends on n . The symbol $\langle \cdot \rangle_*$ denotes the infinite time mean for a typical orbit x_n of a noisy uncoupled map ($k = 0$ in Eqs. (2.1), giving $x_{n+1}^{(i)} = f(x_n^{(i)}, \mu_i) + w_n^{(i)}$) averaged over μ_i ; thus $\langle \cdot \rangle_*$ is time independent. In the limit $N \rightarrow \infty$, a possible solution of Eqs. (2.1) is $\langle q(x_n) \rangle = \langle q(x) \rangle_*$, in which case the coupling has no effect. We refer to this solution as the “incoherent state”. For large finite N , it is expected and numerically observed that for parameter values in the predicted $N \rightarrow \infty$ incoherent state $\langle q(x_n) \rangle$ executes small fluctuations of order $N^{-1/2}$ about $\langle q(x) \rangle_*$. As $|k|$ is increased, the incoherent state becomes unstable, and the mean field $a_n = \langle q(x_n) \rangle - \langle q(x) \rangle_*$ begins to have a macroscopic [i.e., $\mathcal{O}(N^0)$ rather than $\mathcal{O}(N^{-1/2})$] value. This transition typically occurs at some critical nonzero coupling coefficient k_c [69, 89]. In fact there can be two critical k_c values, a positive one, at which coherence arises as k increases through the critical value, and a negative one at which coherence arises as k decreases through the critical value. It was shown for globally coupled, noiseless, continuous-time systems [69] that the transition from the incoherent state to the coherent state can be analyzed by a perturbation method. Here, we develop this perturbation method to analyze globally coupled, noisy, discrete-time systems (i.e., maps). For convenience we will drop the superscript (i) if there is no confusion for all following equations.

2.2.2 Stability Analysis

We carry out the stability analysis for the system in the $N \rightarrow \infty$ limit. Thus, for the purposes of this analysis, given a function h , the symbol $\langle h(x) \rangle$ is now understood to mean

$$\langle h(x_n) \rangle = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N h(x_n^{(i)}).$$

We note that, while $N \rightarrow \infty$ in the analysis, our numerical experiments necessarily have finite N , and, as we will see in Secs. 2.3 and 2.4, the finiteness of N will have a profound effect for implementation of our theory in numerical examples, even though N will be very large in these examples. For compactness of notation, where it is unlikely to create confusion, we will henceforth drop the subscripts and superscripts i denoting the individual maps.

To perform the $N \rightarrow \infty$ stability analysis, we assume a small perturbation from the incoherent state and investigate its stability. Let δ_n be a perturbation to x_n when there is no coupling, and let δx_n be the perturbation with coupling,

$$\delta_{n+1} = f'(x_n, \mu) \delta_n, \quad \delta_0 = 1 \tag{2.2}$$

$$\delta x_{n+1} = f'(x_n, \mu) \delta x_n + kg(x_n) \langle q'(x_n) \delta x_n \rangle. \tag{2.3}$$

(Note that, while the noise does not appear explicitly in Eqns. (2.2) and (2.3), its presence is still important since it influences the orbit x_n : $x_{n+1}^{(i)} = f(x_n^{(i)}, \mu_i) + w_n^{(i)}$.) We are interested in how the mean field perturbation $\langle q'(x_n) \delta x_n \rangle$ develops from an initial perturbation $\langle q'(x_0) \delta x_0 \rangle$. Setting

$$\delta x_n = z_n \delta_n, \tag{2.4}$$

in Eq. (2.3), and employing Eq. (2.2), we obtain

$$z_{n+1} - z_n = k \frac{g(x_n)}{\delta_{n+1}} \langle q'(x_n) \delta x_n \rangle,$$

recursive application of which yields

$$z_{n+1} = k \sum_{p=0}^n \frac{g(x_p)}{\delta_{p+1}} \langle q'(x_p) \delta x_p \rangle + z_0, \quad (2.5)$$

where $z_0 = \delta x_0$. Using the relation (2.4), we can obtain $\langle q'(x_n) \delta x_n \rangle$ from Eq. (2.5),

$$\langle q'(x_{n+1}) \delta x_{n+1} \rangle = k \left\langle \sum_{p=0}^n \frac{\delta_{n+1} q'(x_{n+1}) g(x_p)}{\delta_{p+1}} \langle q'(x_p) \delta x_p \rangle \right\rangle + \delta x_0 \langle q'(x_{n+1}) \delta_{n+1} \rangle. \quad (2.6)$$

Now, we assume exponential instability of the incoherent state so that the mean of the perturbation grows exponentially with n , $\langle q'(x_n) \delta x_n \rangle = \lambda^n \langle q'(x_0) \delta x_0 \rangle$ with $|\lambda| > 1$. Then, we can rewrite Eq. (2.6) as

$$\lambda^{n+1} = k \left\langle \sum_{p=0}^n \frac{\delta_{n+1} q'(x_{n+1}) g(x_p)}{\delta_{p+1}} \lambda^p \right\rangle + \frac{\delta x_0 \langle q'(x_{n+1}) \delta_{n+1} \rangle}{\langle q'(x_0) \delta x_0 \rangle}.$$

Letting $n \rightarrow \infty$, assuming convergence of the summation, and setting $m = n - p$, we have

$$1 = \frac{k}{\lambda} \left\langle \sum_{m=0}^{\infty} \frac{P_m}{\lambda^m} \right\rangle \equiv kQ(\lambda), \quad (2.7)$$

where

$$P_m = \frac{\delta_{n+1} q'(x_{n+1}) g(x_{n-m})}{\delta_{n-m+1}}.$$

The ratio $\delta_{n+1}/\delta_{n-m+1}$ is

$$\frac{\delta_{n+1}}{\delta_{n-m+1}} = f'(x_n, \mu) f'(x_{n-1}, \mu) \cdots f'(x_{n-m+1}, \mu). \quad (2.8)$$

For large m , this quantity increases exponentially with m as $\xi^m(x_0, \mu)$ where $\xi(x_0, \mu)$ is the Lyapunov number of the map $f(x, \mu)$ for the initial condition x_0 . [For almost all x_0 , $\xi(x_0, \mu)$ is the same number (i.e., what is usually referred as “the” Lyapunov number of the chaotic attractor), but there are special choices of x_0 (e.g., x_0 on an unstable periodic orbit) for which $\xi(x_0, \mu)$ takes on a different value.] For Eq. (2.7) to have meaning we require that the summation over m converges. This will be so if

$$|\lambda| > \xi_{\max} \equiv \sup_{x_0, \mu} \xi(x_0, \mu), \quad (2.9)$$

where x_0 is in the attractor basin and μ is in the range of parameter values used in the ensemble. If $|\lambda| < \xi(x_0, \mu)$, the m -th term in the sum increases exponentially with m as $[\xi(x_0, \mu)/\lambda]^m$, and Eq. (2.7) is meaningless. For the case where λ satisfies Eq. (2.9), the exponential convergence of the summation implies that we can interchange the order of the average and the summation. Thus we obtain

$$1 = \frac{k}{\lambda} \sum_{m=0}^{\infty} \frac{Q_m}{\lambda^m} \equiv kQ(\lambda), \quad (2.10)$$

where $Q_m = \langle P_m \rangle$ or

$$Q_m = \left\langle \frac{\delta_{n+1}}{\delta_{n-m+1}} q'(x_{n+1}) g(x_{n-m}) \right\rangle. \quad (2.11)$$

Because we are dealing with chaotic situations (i.e., Lyapunov number greater than one), Eq. (2.9) implies that $Q(\lambda)$, given by Eq. (2.10), can, so far, only be used for λ sufficiently larger than one. We now argue that Eq. (2.10) can be expected to apply for $|\lambda| < 1$. This is crucial, since it is required in order to use the theory for studying the onset of coherence. We heuristically argue in Appendix A that the quantity Q_m can be expected to decrease exponentially with m for typical chaotic maps (a similar argument is presented in Ref. [69] for the case of continuous time

systems). Assuming this to be the case, we have

$$|Q_m| < K\xi_*^{-m}, \quad (2.12)$$

where K and ξ_* are positive constants and $\xi_* < 1$. Thus the sum in Eq. (2.10) now converges for all $|\lambda| > \xi_*$. Hence, while Eq. (2.10) was derived for $|\lambda| > \xi_{\max}$, Eq. (2.12) implies that we can analytically continue $Q(\lambda)$ from the region $|\lambda| > \xi_{\max}$ to the region $\xi_{\max} \geq |\lambda| > \xi_* < 1$.

At the transition of the incoherent state to the coherent state (i.e., at $k = k_c$) the system is marginally stable so that $|\lambda| = 1$ or $\lambda = e^{i\omega}$ with ω real. Thus, $k_c Q(e^{i\omega}) = 1$. Taking the imaginary part of this equation, we obtain an equation for the frequency of oscillation ω at the transition

$$\text{Im}\{Q(e^{i\omega})\} = 0. \quad (2.13)$$

After solving this equation for ω , we obtain the critical coupling strength,

$$k_c = Q(e^{i\omega})^{-1}. \quad (2.14)$$

In addition, expanding Eq. (2.10) about $k = k_c$ and $\lambda = e^{i\omega}$, we obtain the following result for λ near the transition

$$\lambda = e^{i\omega} + \frac{k_c - k}{k_c^2 Q'(e^{i\omega})} + \mathcal{O}\{(k_c - k)^2\},$$

where $Q'(\lambda) = dQ(\lambda)/d\lambda$.

2.3 Examples: One-Dimensional Maps

2.3.1 An Ensemble of Shifted Bernoulli Maps

Our first example is an ensemble of Bernoulli maps,

$$\begin{aligned}
 f(x_n^{(i)}, \mu_i) &= (2x_n^{(i)} + \mu_i) \bmod 2\pi, & i = 1, \dots, N \\
 g(x) &= \sin 2x + \sin 4x, \\
 q(x) &= \cos x,
 \end{aligned}
 \tag{2.15}$$

where μ_i is a shift which is in general different for each map i . Because of the simplicity of this example, $Q(\lambda)$ can be obtained analytically. After plugging Eqs. (2.15) into Eq. (2.11), and taking the noise w to be symmetrically distributed around $w = 0$, we obtain

$$\begin{aligned}
 Q_n(\lambda) &= -\frac{1}{\lambda} \sum_{p=0}^n \langle \sin x_{n+1} (\sin 2x_p + \sin 4x_p) \rangle \left(\frac{2}{\lambda}\right)^{n-p} \\
 &= -\frac{1}{2\lambda} \left(\langle \cos w_n \rangle \langle \cos \mu \rangle + \frac{2}{\lambda} \langle \cos 2w_{n-1} \rangle \langle \cos w_{n-2} \rangle \langle \cos 3\mu \rangle \right).
 \end{aligned}$$

The second equality results from noting that, since $x_p \bmod 2\pi$ has a uniform density in $(0, 2\pi)$, all terms in the summation are zero except for $p = n$ and $p = n - 1$. In the above, if the ensemble of shift parameters $\{\mu_i\}$ is generated from a probability density $\rho(\mu)$, then

$$\langle \cos M\mu \rangle \equiv \int \rho(\mu) \cos(M\mu) d\mu.$$

Taking the noise w to be normally distributed, we obtain

$$Q(e^{i\omega}) = -\frac{1}{2} \left(\frac{e^{-\sigma^2/2}}{e^{i\omega}} \langle \cos \mu \rangle + \frac{2e^{-5\sigma^2/2}}{e^{2i\omega}} \langle \cos 3\mu \rangle \right). \tag{2.16}$$

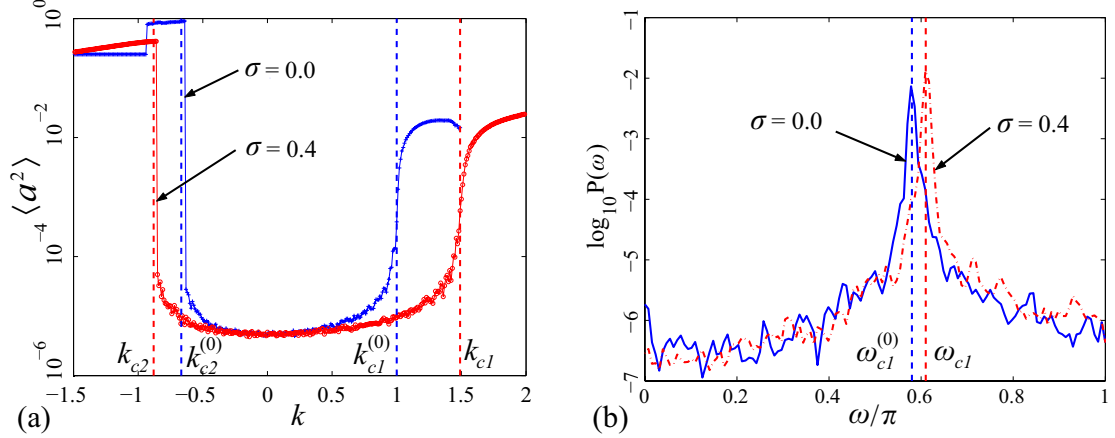


Figure 2.1: Results for an ensemble of $N = 10^5$ Bernoulli maps with and without noise: (a) shows $\langle a^2 \rangle$ versus k . The time average $\langle a^2 \rangle$ is computed using 1000 iterations. The power spectral densities of a_n at the positive critical values are shown in (b) for the cases without noise ($k = k_{c1}^{(0)}$, solid graph) and with noise ($k = k_{c1}$, dash-dot graph). The predicted values of critical coupling strengths and frequencies of oscillation agree with the experimental results.

For simplicity, in what follows we take $\rho(\mu) = \delta(\mu)$, so that $\langle \cos \mu \rangle = \langle \cos 3u \rangle = 1$, which with Eq. (2.13) and Eq. (2.14), yields critical values

$$\begin{aligned} k_{c1} &= e^{5\sigma^2/2}, & \omega_{c1} &= \arccos\left(-\frac{e^{2\sigma^2}}{4}\right), \\ k_{c2} &= -2(e^{-\sigma^2/2} + 2e^{-5\sigma^2/2})^{-1}, & \omega_{c2} &= 0. \end{aligned} \quad (2.17)$$

For $\sigma = 0$, Eqs. (2.17) agree with the result of Ref. [89].

We present results of numerical experiments in Figure 2.1 for the case that the system has no noise and the case that the system has normally distributed noise with $\sigma^2 = 0.16$. Without noise (i.e., $\sigma^2 = 0$), Eqs. (2.17) yield $k_{c1}^{(0)} = 1$ at $\omega_{c1}^{(0)} = \arccos(-\frac{1}{4}) \approx 0.58\pi$ and $k_{c2}^{(0)} = -2/3$ at $\omega_{c2}^{(0)} = 0$. With $\sigma^2 = 0.16$, Eqs. (2.17) yield the critical values $k_{c1} \approx 1.49$ at $\omega_{c1} \approx 0.61\pi$ and $k_{c2} \approx -0.88$ at $\omega_{c2} = 0$. In Figure 2.1(a), $\langle a^2 \rangle$ is the time average of the square of the mean field $a_n = \langle q(x_n) \rangle - \langle q(x) \rangle_*$ and this average is computed over 1000 iterations in time. We see that the mean field starts to have macroscopic values near the predicted critical values of the coupling strengths (the vertical dashed lines). Also, the power

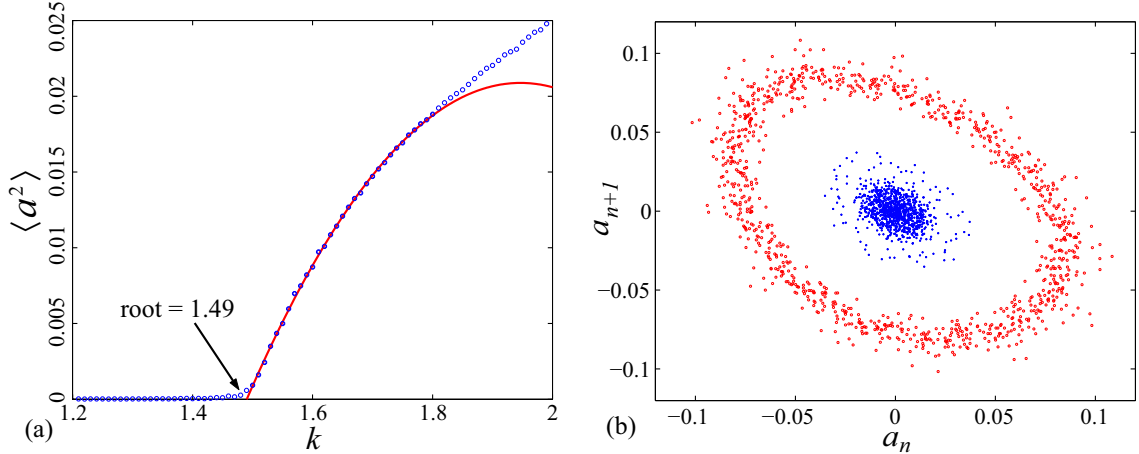


Figure 2.2: Linear scale plot of $\langle a^2 \rangle$ and orbits of a_n for an ensemble of 10^5 noisy Bernoulli maps: (a) shows $\langle a^2 \rangle$ versus k (dots) and a quadratic curve fit (solid line) to the data in the range $1.5 \leq k \leq 1.8$. (b) shows orbits of a_n for $k = 1.53 > k_{c1} \approx 1.49$ (points in halo about origin) and $k = 1.45 < k_{c1}$ (points clustered near origin). $\langle a^2 \rangle$ approaches zero linearly and the orbits encircle the origin consistent with a Hopf bifurcation.

spectral density of the sequence of a_n , Figure 2.1(b), shows that the frequencies of oscillation at $k = k_{c1}^{(0)}$ and $k = k_{c1}$ coincide with the predicted values $\omega_{c1}^{(0)}$ and ω_{c1} for each case. Here, and in the following examples, the power spectral densities are estimated using Welch's method [85].

In Figure 2.2(a), we replot $\langle a^2 \rangle$ versus k for the above described ensemble of noisy Bernoulli maps using a linear scale (dots). Also, a quadratic curve fit to the numerical data in the range $1.5 \leq k \leq 1.8$ is shown as a solid line in the same figure. We see that the fitted line agrees well with the experimental results for $k \leq 1.8$. Consistent with the expectation that the transition is a Hopf bifurcation (since the frequency of oscillation at transition (Figure 2.1) is nonzero), $\langle a^2 \rangle$ approaches zero linearly with $(k - k_{c1})$. (Close examination of the numerical results in Figure 2.2(a) very near k_c shows a slight rounding of the, otherwise sharp, transition due to finite N .) Figure 2.2(b) shows a_{n+1} versus a_n for two values of k , one slightly past k_{c1} (coherent) and one slightly before k_{c1} (incoherent). For $k = 1.45 < k_{c1} = 1.49$, we see that the orbit points appear as a cloud centered at the origin as expected for the

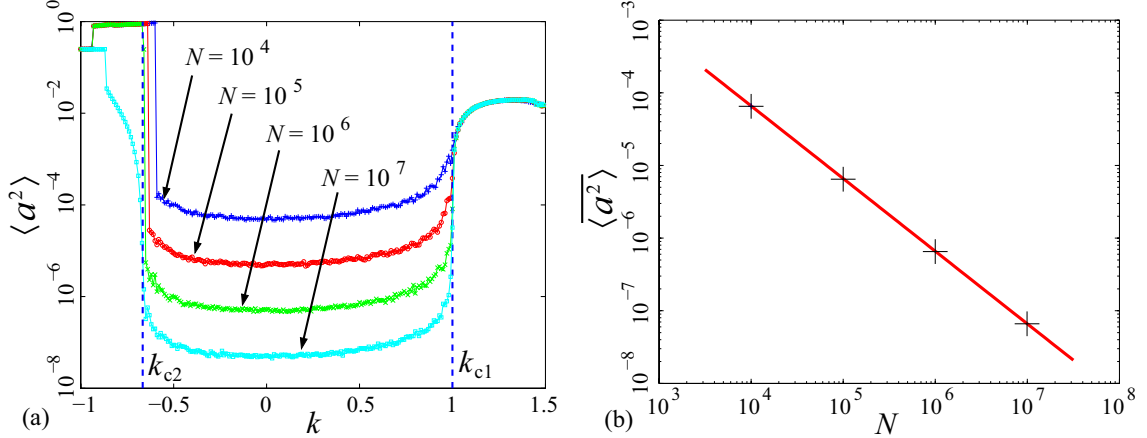


Figure 2.3: Results for ensembles of $N = 10^4, 10^5, 10^6, 10^7$ Bernoulli maps: (a) shows $\langle a^2 \rangle$ versus k . (b) shows $\overline{\langle a^2 \rangle}$ versus N . Within the incoherent state $\langle a^2 \rangle$ varies as N^{-1} .

incoherent state. For $k = 1.53 > k_{c1}$, the orbit points appear as a loop encircling the origin. As expected, at each step n points in the ring, on average, advance in angle by almost ω_{c1} radians. (A plot similar to that in Figure 2.2(b) appears in Ref. [89] for the noiseless case, $\sigma = 0$.)

Figure 2.3 demonstrates the effect of varying the ensemble size N . Figure 2.3(a) shows results for our coupled noiseless Bernoulli map example for $N = 10^4, 10^5, 10^6$, and 10^7 . We note that these graphs differ appreciably only in the range $k_{c2}^{(0)} < k < k_{c1}^{(0)}$ corresponding to the incoherent state. Figure 2.3(b) shows the values of $\langle a^2 \rangle$ averaged over the range, $(k_{c2}^{(0)} + 0.2) \leq k \leq (k_{c1}^{(0)} - 0.2)$, that is within the incoherent region (we denote this average $\overline{\langle a^2 \rangle}$), versus N on a log-log plot. Also shown in Figure 2.3(b) is a straight line of slope -1 . We see that, similar to what is expected for a sum of random variables, a scaling of $\overline{\langle a^2 \rangle}$ as N^{-1} is consistent with the data. The behavior seen in Figure 2.3 is also seen for all our other examples (except for that in Sec. 2.3.3). Regarding Figure 2.1(b), we also note that, as the ensemble size is increased, the spectral power not at ω_c decreases, and the spectral peak at ω_c becomes sharper, consistent with the spectrum approaching a delta function at $\omega = \omega_c$ as $N \rightarrow \infty$.

Remark: In the supercritical Hopf bifurcation of a discrete time system, the resulting stable orbit lies on a closed curve bifurcating off the basic periodic orbit that was stable before the bifurcation (in our case, the incoherent state, which has period one). On this curve, the orbit can be either periodic (consisting of a finite number of discrete points) or quasiperiodic (filling out the curve, as in Figure 2.2(b)). Generically, $\omega_c/2\pi$ will be irrational, and, for most (in the Lebesgue sense) $(k - k_c)$ values near zero, the orbit will be quasiperiodic, although there is an open dense set of values of $(k - k_c)$ for which there is an attracting periodic orbit. We note that, in our case, due to finite N , the bifurcation is noisy, and this can wash out small windows of periodic behavior.

2.3.2 A Heterogeneous Ensemble of Modified Bernoulli Maps

The preceding example, coupled Bernoulli maps, is useful because it allows an analytic solution for $Q(\lambda)$ (preceding section and Ref. [89]). In more typical cases, analytical solution for $Q(\lambda)$ is not possible, and numerical techniques for calculating $Q(\lambda)$ must be formulated. Furthermore, the maps in the ensemble may not all be identical. In order to illustrate these points, our second example is an ensemble of noiseless modified Bernoulli maps depending upon a map parameter μ that is uniformly distributed in the interval $(1, 2)$,

$$f_i(x) = 2x \bmod 2\pi + \mu_i \sin x, \quad \mu_i \sim U(1, 2), \quad i = 1, \dots, N$$

$$g(x) = \sin 4x, \quad q(x) = \cos x,$$

where $U(1, 2)$ is the uniform distribution over the interval $(1, 2)$.

Since we do not have a closed form expression for the natural invariant density in this case, we evaluate $Q(e^{i\omega})$ numerically in the following way. First, we produce $N = 10^6$ points, x_0 , uniformly distributed in the interval $(0, 2\pi)$ with a random num-

ber generator, and we also produce the same number of randomly chosen parameters uniformly distributed in the interval $(1, 2)$. Then we evolve the uncoupled ($k = 0$) system forward in time for 3000 steps saving the values of $x^{(i)}$ for the last 31 iterations. Using these values we construct histogram approximation to the invariant density using bins of width $2\pi \times 10^{-3}$ in x . We let x_{3000} be x_{n+1} in Eq. (2.11). Using the saved data, we obtain $\delta_{n+1}^{(i)}/\delta_{n-m+1}^{(i)}$ (from Eq. (2.8)) and $g(x_{n-m}^{(i)})$, and employ Eq. (2.11) to obtain an approximation to Q_m for $m = 0, \dots, 30$. Note that, in the incoherent state, $x^{(i)}$ for the system (2.1) has an invariant density resulting from the uncoupled individual maps. Thus, if the ensemble has an infinite number of maps ($N \rightarrow \infty$), and each orbit in the ensemble is given an initial perturbation δx_0 (as in our analysis in Sec. 2.2), then the uncoupled ensemble will eventually settle down to the invariant density after a sufficiently large number of iterations. It is, therefore, expected that for $N \rightarrow \infty$ the mean field perturbation $\langle q'(x_n)\delta x_n \rangle$ converges to zero as n increases. Hence Q_m converges to zero with increasing m in the large ensemble limit $N \rightarrow \infty$. However, due to the finite ensemble size ($N = 10^6$), our computation of Q_m does not converge to zero. What happens is that as m increases Q_m eventually becomes small; say it assumes a small value at $m = m_c$. However, as m becomes larger, our computed approximations to Q_m become inaccurate. Since $\delta_{n+1}^{(i)}/\delta_{n-m+1}^{(i)}$ on average increases exponentially with m (chaos), the individual terms in the average (2.11) becomes larger and larger as m increases. On the other hand, for $N \rightarrow \infty$, the average decreases with m . Thus as m increases cancellation between terms in the average must become more and more precise. Hence to obtain good statistics for Q_m demands exponentially larger and larger ensemble size N as m increases. Thus for any finite N we expect our numerical computation of Q_m to breakdown as m increases.

We plot six numerical approximations to Q_m with different randomly chosen initial conditions in Figure 2.4(a). We see that our approximations to Q_m become

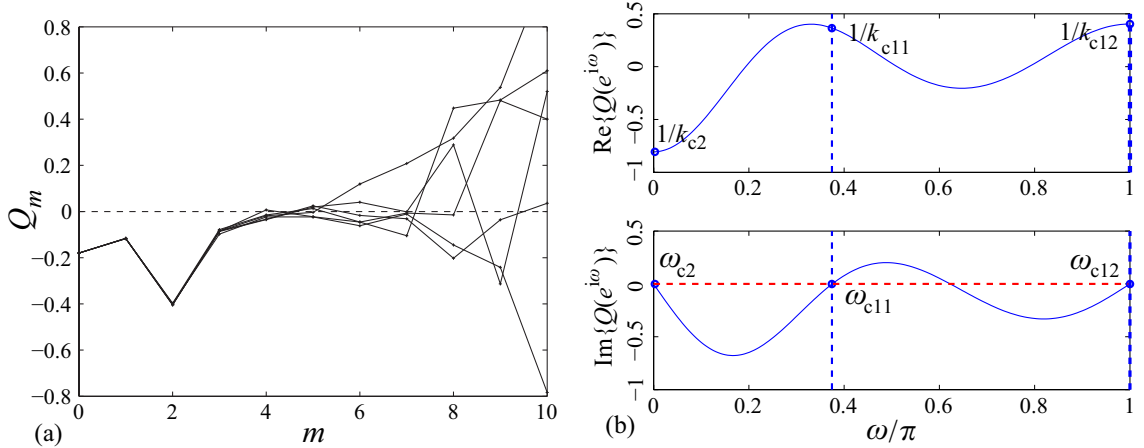


Figure 2.4: $Q(e^{i\omega})$ for an ensemble of $N = 10^6$ modified Bernoulli maps with uniformly distributed parameters: (a) shows six numerical approximations to Q_m . The six approximations are near zero at $m = 4, 5$, but then diverge from each other due to the combined effect of chaotic dynamics and finite ensemble size. (b) shows $Q(e^{i\omega})$ evaluated by averaging the six results from (a) for $m = 0 - 4$ and taking $Q_m = 0$ for $m \geq 5$. The imaginary part of $Q(e^{i\omega})$ has four zero-crossing points and we label three of them and corresponding real parts of these three points. The real parts corresponding to ω_{c11} and ω_{c12} are close to each other.

small at around $m = 4$ or 5 , but increase after that and clearly become unequal. To obtain $Q(e^{i\omega})$, we set Q_m to be zero for $m \geq 5$ and take the average over our six approximations. The real and imaginary parts of the resulting approximation to $Q(e^{i\omega})$ are shown in Figure 2.4(b). (When the imaginary part of $Q(e^{i\omega})$ crosses zero, the real part has a maximum or minimum near these crossing points.) The greatest positive $Q(e^{i\omega})$ at a crossing point and the smallest negative $Q(e^{i\omega})$ at a crossing point are the reciprocals of the positive and negative critical coupling strengths respectively. In this example, the imaginary part of $Q(e^{i\omega})$ crosses zero four times in the plotted range. We label three of these zeros ω_{c2} , ω_{c11} and ω_{c12} . The positive values of $Q(e^{i\omega})$ at $\omega_{c11} \approx 0.37\pi$ and $\omega_{c12} \approx \pi$ are close to each other, although the value at ω_{c12} is larger. From the real part of $Q(e^{i\omega})$, we obtain critical coupling strengths, $k_{c2} \approx -1.24$ and $k_{c12} \approx 2.48$ corresponding to ω_{c2} and ω_{c12} .

Results from coupled ensembles of 10^5 and 10^6 noiseless modified Bernoulli maps are shown in Figure 2.5(a) along with the critical values (vertical dashed

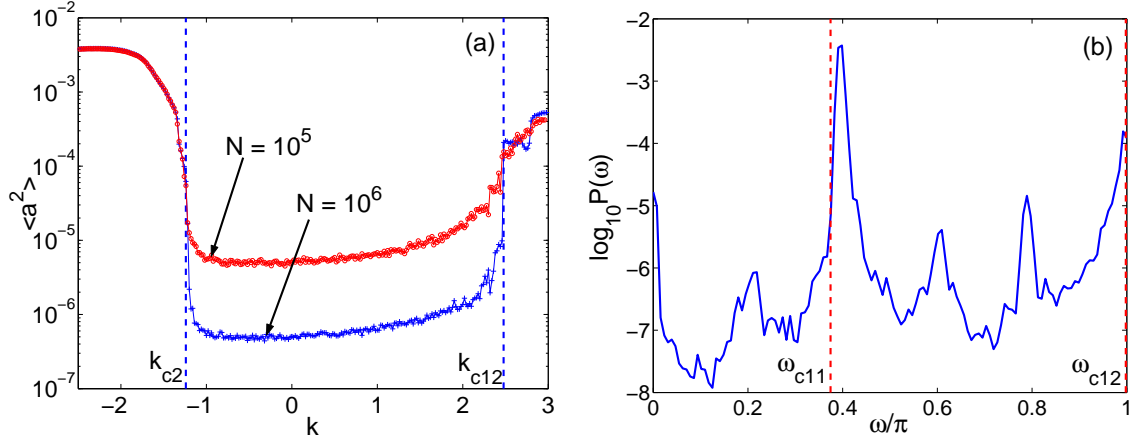


Figure 2.5: Results for ensembles of $N = 10^5, 10^6$ modified Bernoulli maps with uniformly distributed parameters: (a) shows $\langle a^2 \rangle$ versus k , (b) is the power spectral density of a_n at $k = 2.48 \approx k_{c1}$.

lines) which we obtained from our numerical approximation to $Q(e^{i\omega})$. We see that k_{c12} and k_{c2} closely agree with the experimental results. Figure 2.5(b) shows the power spectral density of a_n for $k = k_{c12} = 2.48$; we note that peaks are present both at ω_{c11} (vertical dashed line) and at ω_{c12} , and that the peak at ω_{c11} is, in fact, larger, even though k_{c12} is less than $1/Q(e^{i\omega_{c11}})$.

2.3.3 An Ensemble of Logistic Maps

Our third example is a noiseless ensemble of logistic maps with $g(x) = 1$ and $q(x) = x$,

$$\begin{aligned} x_{n+1}^{(i)} &= f(x_n^{(i)}) + k(\langle x_n \rangle - \langle x \rangle_*), \quad i = 1, \dots, N, \\ f(x_n^{(i)}) &= \mu x_n^{(i)}(1 - x_n^{(i)}), \end{aligned} \quad (2.18)$$

where all maps have identical parameters ($\mu_i = \mu = 3.9, i = 1, \dots, N$). In this case, we were not able to obtain useful results by use of Eq. (2.10). We include this example mainly to illustrate that numerical implementation of Eq. (2.10) can sometimes be problematic, and to speculate on why that might be the case.

In this example, we again do not have a closed form expression for the invariant density of an uncoupled map. Hence, we attempted to evaluate Q_m numerically

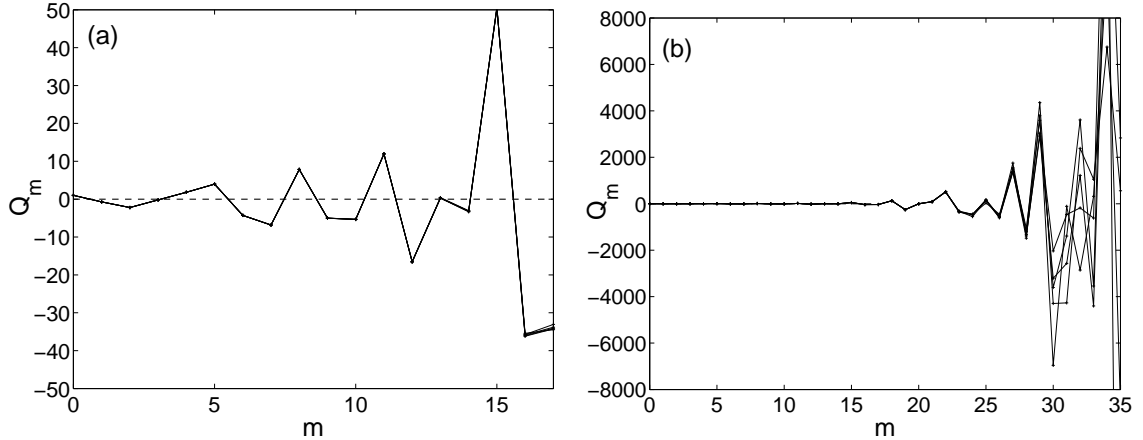


Figure 2.6: Q_m for an ensemble of $N = 10^9$ logistic maps with parameter $\mu = 3.9$: (a) shows Q_m up to $m = 17$ and (b) shows Q_m plotted up to $m = 35$. We see that our approximations do not converge to small values before diverging from each other.

using Eqs. (2.8) and (2.11) which in this case is simply

$$Q_m = \left\langle \frac{\delta_{n+1}}{\delta_{n-m+1}} \right\rangle. \quad (2.19)$$

In Figure 2.6, using $N = 10^9$, we plot five approximations to Q_m up to $m = 35$ obtained using different random initial conditions (as in Sec. 2.3.2). We see that the five approximations stay close to each other up to $m = 28$ without converging to zero. Past $m = 28$, they diverge from each other. Our numerical approximations to Q_m do not converge to zero before diverging from each other, and we thus cannot predict the critical coupling strength from the theory. Since it is impractical for us to increase N further we cannot proceed further. Indeed, since Figure 2.6 indicates growing oscillations of Q_m with increasing m , it is questionable that increase of N would solve the problem.

Note that the logistic map has dense periodic windows in the chaotic parameter range and that the natural invariant density $\rho(x)$ of the logistic map for typical chaotic parameter values has a dense countable set of x -values at which $\rho(x)$ is infinite [66]. We speculate that this could be the root of our problem in applying

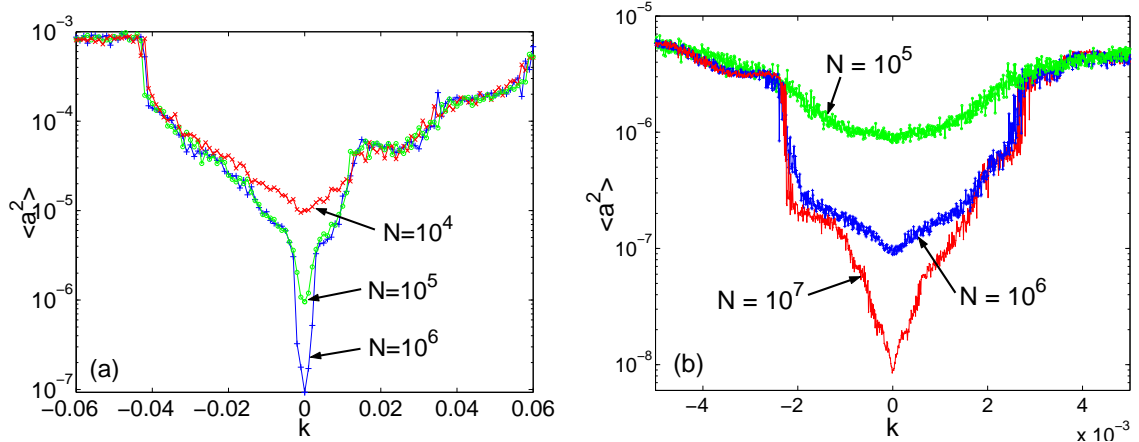


Figure 2.7: $\langle a^2 \rangle$ versus k for ensembles of logistic maps: Data for ensembles of sizes $N = 10^4, 10^5, 10^6$ are shown in (a). We plot $\langle a^2 \rangle$ for ensembles of $N = 10^5, 10^6, 10^7$ logistic maps versus k in a narrower range of k in (b). We see that the N dependence is confined to a very small region near $k = 0$ and that, as N increases, the confined region becomes narrower.

Eq. (2.10) (see the Appendix B). In particular, both of these features call the application of the perturbation theory used in Sec. 2.2 into question.

In addition, we find that the behavior found for this example is qualitatively different from the behavior found for the examples in Secs. 2.3.1 and 2.3.2. In particular, Figure 2.7 shows $\langle a^2 \rangle$ versus k for $N = 10^4, 10^5, 10^6$, and 10^7 . We see that, unlike Figure 2.3(a), the N dependence is confined to a very small region near $k = 0$, and this confinement becomes narrower as N increases. Thus, if there are critical values $k_{c1} > 0$ and $k_{c2} < 0$, bounding an incoherent state in $k_{c2} < k < k_{c1}$, these values have very small magnitude. Another possibility (which we suspect might be the case) is that there may be no incoherent state, except at $k = 0$, and that, as soon as k is nonzero, coherent behavior arises discontinuously. Such a situation would be outside the scope of our perturbation theory.

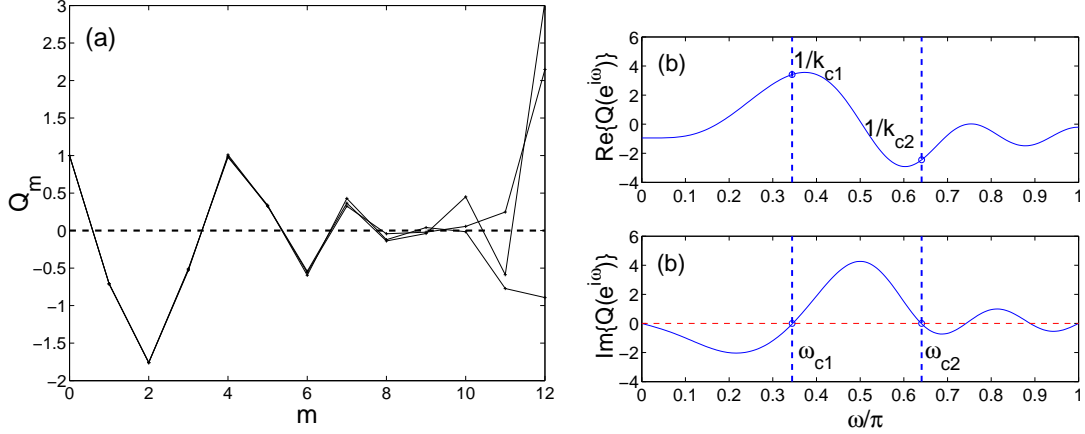


Figure 2.8: $Q(e^{i\omega})$ for an ensemble of $N = 10^6$ noisy logistic maps with parameter $\mu = 3.9$ and normally distributed noise: (a) shows three numerical approximations to Q_m . The three approximations are near zero at $m = 8, 9$, but then diverge from each other. (b) shows $Q(e^{i\omega})$ evaluated from (a) assuming that $Q_m = 0$ after $m = 8$.

2.3.4 An Ensemble of Logistic Maps with Noise

In our fourth example, we consider the case studied in the previous section, but with noise added. As mentioned in Sec. 2.2, adding noise makes the orbit density smooth and may eliminate small periodic windows [80, 67, 68]. Hence, we can expect that the confinement of N dependence of $\langle a^2 \rangle$ shown in Figure 2.7 will be widened and we confirmed this dependence by numerical experiments. Also, we find that the noise promotes convergence of Q_m , and that application of Eq. (2.10) now yields accurate and useful results. We consider Eqs. (2.1) with $f(x) = 3.9x(1-x)$, $g(x) = 1$, $q(x) = x$, and $w_n^{(i)}$ normally distributed with variance $\sigma^2 = 10^{-4}$ (see Sec. 2.2). Note that $x_{n+1}^{(i)}$ in Eqs. (2.1) could fall outside the basin of attraction of the map ($0 < x < 1$) because of the noise $w_{n+1}^{(i)}$ (the coupling term is on the order of 10^{-4} near the incoherent state with $N = 10^5$ noisy logistic maps). To prevent any variable from escaping the basin, we replace $x_{n+1}^{(i)}$ by $x_{n+1}^{(i)} \bmod 1$, if it falls outside the basin.

Again we do not have a closed form expression for the orbit density, and hence we rely on a numerical approximation to Q_m to obtain $Q(e^{i\omega})$. Three approximations

to Q_m for different random number seeds are shown in Figure 2.8(a). These three plots show good agreement with each other up to $m = 9$, where they assume small values. $Q(e^{i\omega})$ derived from one of these approximations with Q_m set to zero for $m \geq 9$ is shown in Figure 2.8(b). Using the data in Figure 2.8(b), we predict that the critical values of the coupling coefficient will be $k_{c1} \approx 0.29$ and $k_{c2} \approx -0.39$, and that the corresponding frequencies of oscillation at the onset of coherence will be $\omega_{c1} \approx 0.34\pi$ and $\omega_{c2} \approx 0.64\pi$.

In Figure 2.9(a) we plot the time average of the square of the mean field a_n for ensembles of 10^4 and 10^5 noisy logistic maps, $\langle a^2 \rangle$, and the predicted values of the critical coupling strengths (vertical dashed lines), k_{c1} and k_{c2} . We replot $\langle a^2 \rangle$ using a linear scale in Figure 2.9(b) (dots). Also shown in Figure 2.9(b) as a solid line is a quadratic curve fit to the numerical data in the range $0.29 \leq k \leq 0.44$ which agrees well with the experimental results for $k \leq 0.51$. Extrapolating the fitted quadratic curve to $\langle a^2 \rangle = 0$, we obtain an accurate estimate of the critical value of the coupling strength, $k \approx 0.29$, confirming the theoretical prediction. From Figure 2.9(b) we see that $\langle a^2 \rangle$ approaches zero linearly with $(k - k_{c1})$ consistent with a Hopf bifurcation ($\omega_{c1} \neq 0$).

The frequency of oscillation at $k = k_{c1}$ obtained from $Q(e^{i\omega})$ is shown in Figure 2.10(a) as a vertical dashed line, along with the power spectral density of a_n . The dominant frequency of the spectrum agrees with the frequency predicted by our analysis. Figure 2.10(b) shows a_{n+1} versus a_n for two values of k , one just past k_{c1} (coherent) and one just before k_{c1} (incoherent). For $k = 0.28 < k_{c1} = 0.29$, we see that the orbit points appear as a cloud centered at the origin. For $k = 0.30 > k_{c1}$, the orbit points appear as a loop encircling the origin (at the frequency ω_{c1}).

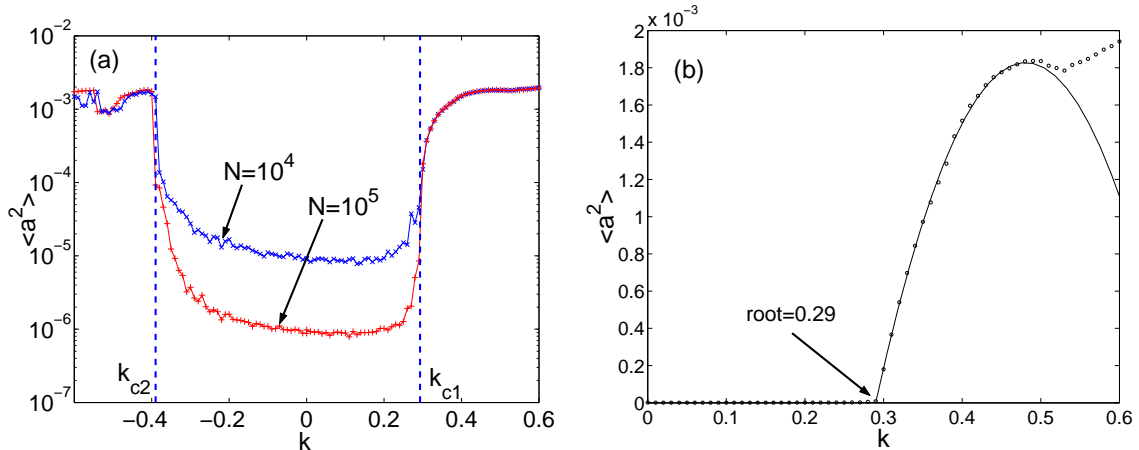


Figure 2.9: Experimental results for an ensemble of $N = 10^4, 10^5$ noisy logistic maps with parameter $\mu = 3.9$ and normally distributed noise: (a) shows a semilogarithmic plot of $\langle a^2 \rangle$ versus k , and (b) shows $\langle a^2 \rangle$ (dots) and a quadratic curve fit to the numerical data in the range $0.29 \leq k \leq 0.44$ (solid line) for $N = 10^5$.

2.3.5 A Heterogeneous Ensemble of Logistic Maps

In our fifth example, as in Sec. 2.3.3, we again consider a noiseless ensemble of logistic maps (2.18), but now with the map parameter, μ , uniformly distributed in the interval (3.88, 3.96). We find that the introduction of parameter spread appears to have a regularizing effect and, for this noiseless case, we now obtain results in agreement with our perturbation theory. A similar regularizing effect has been observed by Shibata and Kaneko [81].

Figure 2.11(a) shows six approximations to Q_m for different random number seeds. They agree well with each other up to $m = 12$, where they assume small values. $Q(e^{i\omega})$ shown in Figure 2.11(b) is derived from one of these approximations with Q_m set to zero for $m \geq 12$. From Figure 2.11(b) we predict that the critical coupling strengths will be $k_{c1} \approx 0.24$ and $k_{c2} \approx -0.13$, and that the corresponding frequencies of oscillation at the onset of coherence will be $\omega_{c1} \approx 0.37\pi$ and $\omega_{c2} \approx 0.60\pi$.

In Figure 2.12(a) we plot the time average of the square of the mean field for noiseless ensembles of $10^4, 10^5$, and 10^6 logistic maps with the parameter, μ ,

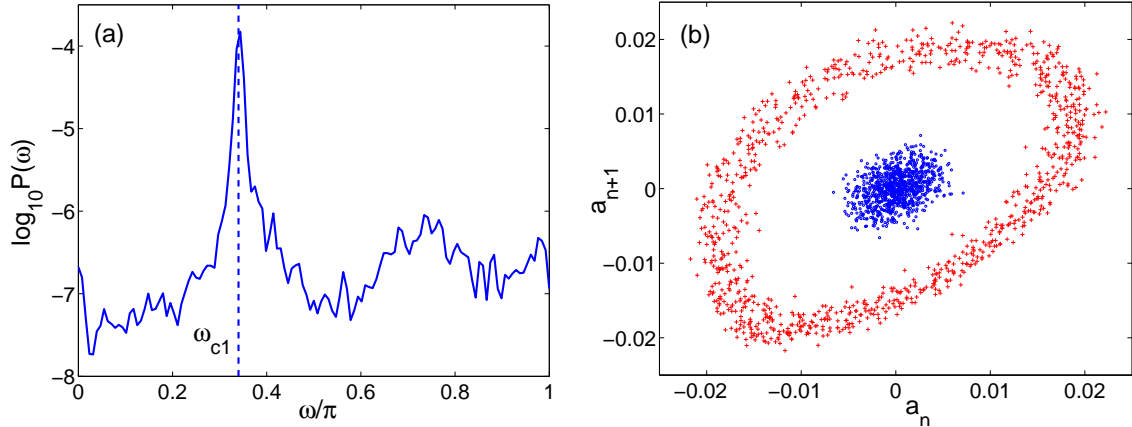


Figure 2.10: Power spectral density and orbits of a_n for $N = 10^4$: The frequency of oscillation and the power spectral density of a_n at $k = k_{c1}$ are plotted in (a). (b) shows orbits of a_n for $k = 0.30 > k_{c1} \approx 0.29$ (points in halo about origin) and $k = 0.28 < k_{c1}$ (points clustered near origin).

uniformly distributed in the interval (3.88, 3.96). The power spectral density of a_n at $k \approx k_{c1}$ is shown in Figure 2.12(b). We see that the predicted values agree well with the numerical experiments. Note that the peak around $\omega = 0.74\pi$ is also expected from the data in Figure 2.11(b). The real part of $Q(e^{i\omega})$ at $\omega \approx 0.74\pi$, where the imaginary part crosses zero, has a value comparable with that at $k = k_{c1}$.

2.4 Examples: Two-Dimensional Maps

2.4.1 Multidimensional Extension of the Theory

In this section we examine globally coupled multi-dimensional systems for which, analogous to the system (2.1), the considered system is

$$\mathbf{x}_{n+1}^{(i)} = \mathbf{f}(\mathbf{x}_n^{(i)}) + \mathbf{w}_n^{(i)} + \mathbf{G}(\mathbf{x}_n) \mathbf{K} (\langle \mathbf{q}(\mathbf{x}_n) \rangle - \langle \mathbf{q}(\mathbf{x}) \rangle_*), \quad i = 1, \dots, N,$$

where $\mathbf{x}_n = [x_{1,n} \ x_{2,n} \ \dots \ x_{r,n}]^T$, r is the dimension of a map, \mathbf{K} is the coupling matrix, $\mathbf{G}(\mathbf{x}_n)$ is a matrix function, and $\mathbf{q}(\mathbf{x}_n)$ is a vector function of \mathbf{x}_n . Here, $\mathbf{w}_n^{(i)}$

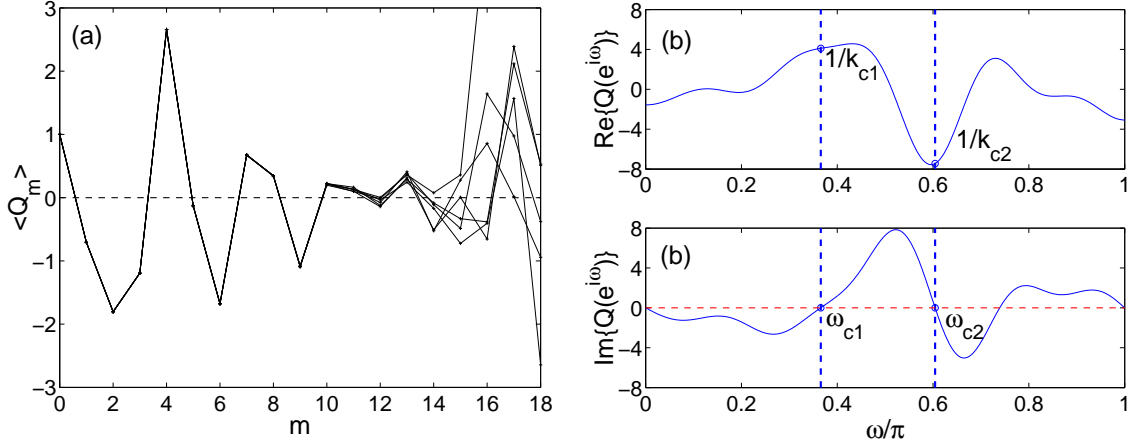


Figure 2.11: $Q(e^{i\omega})$ for a noiseless ensemble of $N = 10^9$ logistic maps with parameter μ uniformly distributed in the interval $(3.88, 3.96)$: (a) shows three numerical approximations to Q_m . The three approximations are near zero at $m = 12$, but then diverge from each other. (b) shows $Q(e^{i\omega})$ evaluated from (a) assuming that $Q_m = 0$ after $m = 12$.

is random noise where $E[\mathbf{w}_n^{(i)}] = \mathbf{0}$, $E[\mathbf{w}_n^{(i)} \mathbf{w}_k^{(j)\top}] = \Sigma \delta_{ij} \delta_{nk}$, Σ is the covariance matrix, and we assume that the noise at each iterate is identically distributed and that \mathbf{x}_n and \mathbf{w}_n are independent.

Let $\mathbf{Df}(\mathbf{x}) = \partial \mathbf{f}(\mathbf{x}) / \partial \mathbf{x}$ and

$$\mathbf{M}(n, p) = \begin{cases} \mathbf{Df}(\mathbf{x}_{n-1}) \mathbf{Df}(\mathbf{x}_{n-2}) \cdots \mathbf{Df}(\mathbf{x}_p), & n \geq p + 1 \\ \mathbb{I}, & n = p \end{cases}$$

where \mathbb{I} is the identity matrix. Then proceeding as in Sec. 2.2, we assume that the mean of the perturbation $\delta \mathbf{x}$ grows exponentially with n , $\langle \mathbf{Dq}(\mathbf{x}_n) \delta \mathbf{x}_n \rangle = \mathbf{v} \lambda^n$, where $\mathbf{v} = \langle \mathbf{Dq}(\mathbf{x}_0) \delta \mathbf{x}_0 \rangle$. Letting $n \rightarrow \infty$, assuming convergence of the summation, and setting $m = n - p$, we obtain

$$(\mathbb{I} - \mathbf{Q}(\lambda) \mathbf{K}) \mathbf{v} = \mathbf{0}.$$

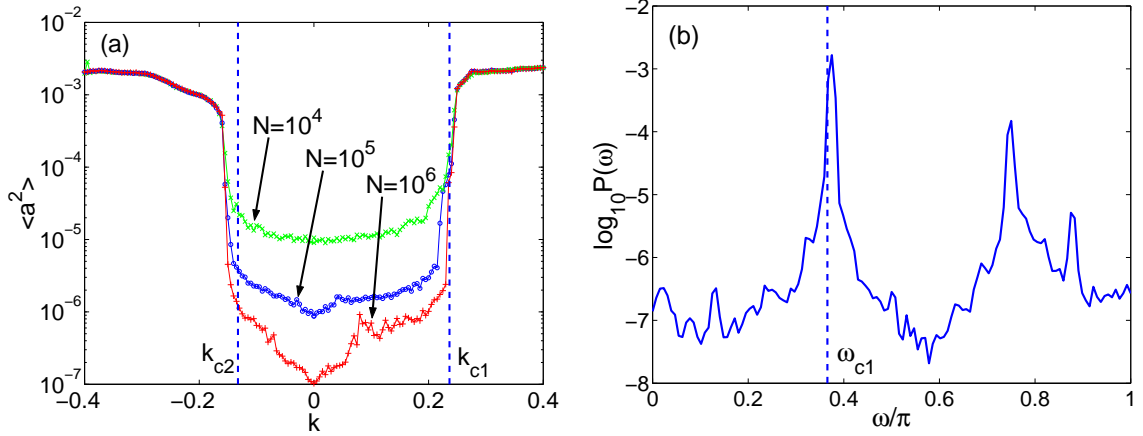


Figure 2.12: Experimental results for a noiseless ensemble of $N = 10^4, 10^5, 10^6$ logistic maps with parameter μ uniformly distributed in the interval $(3.88, 3.96)$: (a) shows $\langle a^2 \rangle$ versus k , and (b) shows the power spectral density of a_n at $k = 0.24 \approx k_{c1}$ for $N = 10^6$.

where

$$Q(\lambda) = \sum_{m=1}^{\infty} Q_m / \lambda^m, \quad (2.20)$$

$$Q_m = \langle \mathbf{Dq}(\mathbf{x}_n) \mathbf{M}(n, n-m+1) \mathbf{G}(\mathbf{x}_{n-m}) \rangle,$$

which yields

$$\det(\mathbb{I} - Q(\lambda)K) = 0. \quad (2.21)$$

By setting $\lambda = e^{i\omega}$ in Eq. (2.21), we can determine the critical coupling strength and the frequency of oscillation (see Sec. 2.2).

2.4.2 An Ensemble of Cat Maps

We now illustrate Eq. (2.21) by application to an ensemble of globally coupled two dimensional maps. In particular, we take $\mathbf{f}(\mathbf{x})$ to be the cat map,

$$\begin{aligned} \mathbf{f}(\mathbf{x}) &= \begin{bmatrix} x + y \\ x + 2y \end{bmatrix} \bmod 2\pi, & \mathbf{q}(\mathbf{x}) &= \begin{bmatrix} \cos x \\ 0 \end{bmatrix}, \\ \mathbf{G}(\mathbf{x}) &= \begin{bmatrix} \sin(2x + 3y) & 0 \\ 0 & 0 \end{bmatrix}, & \mathbf{K} &= \begin{bmatrix} k & 0 \\ 0 & 0 \end{bmatrix}, \end{aligned} \quad (2.22)$$

where $\mathbf{x}_n = [x_n \ y_n]^T$. For the noise, we choose

$$E[\mathbf{w}_n^{(i)}] = \mathbf{0}, \quad E[\mathbf{w}_n^{(i)} \mathbf{w}_k^{(j)T}] = \begin{bmatrix} \sigma^2 & 0 \\ 0 & \sigma^2 \end{bmatrix} \delta_{ij} \delta_{nk},$$

where $\mathbf{w}_n^{(i)} = [w_{x,n}^{(i)} \ w_{y,n}^{(i)}]^T$. We denote the element at the k -th row and the l -th column of a matrix \mathbf{A} by $[\mathbf{A}]_{kl}$. Then, after plugging Eqs. (2.22) into Eqs. (2.20), we obtain

$$\begin{aligned} [\mathbf{Q}_n(\lambda)]_{11} &= - \sum_{p=0}^{n-1} \langle \sin x_n \sin(2x_p + 3y_p) \rangle \frac{[\mathbf{M}(n, p + 1)]_{11}}{\lambda^{n-p}} \\ &= - \frac{1}{2\lambda^2} \langle \cos(w_{x,n} + w_{x,n-1} + w_{y,n-1}) \rangle. \end{aligned}$$

The second equality results from noting that, since the measure generated by orbits of the uncoupled noisy cat maps is uniform in $0 \leq x \leq 2\pi$, $0 \leq y \leq 2\pi$, all terms in the summation are zero except for $p = n - 2$. For normally distributed noise, $[\mathbf{Q}_n(\lambda)]_{11} = -(2\lambda^2)^{-1} \exp(-3\sigma^2/2)$. From the condition (2.21), we obtain

$$\frac{1}{k} = - \frac{e^{-3\sigma^2/2}}{2e^{2i\omega}},$$

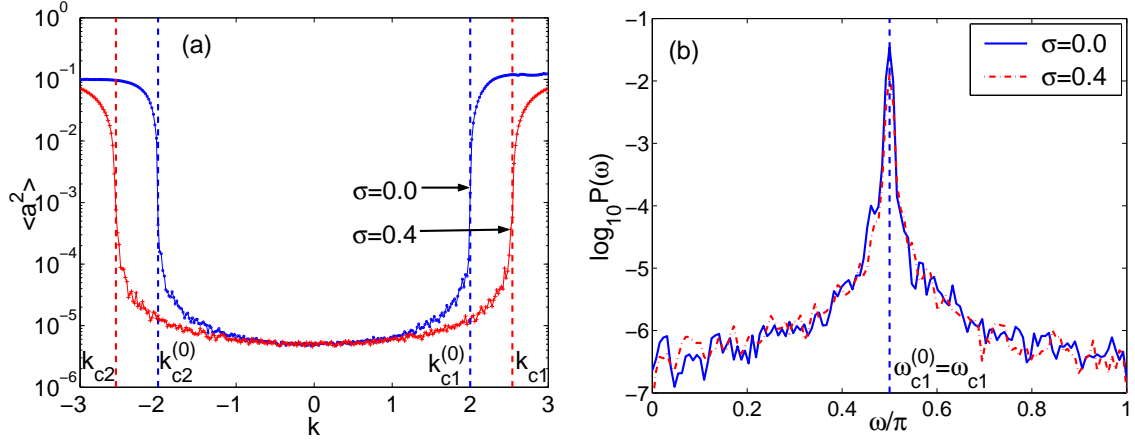


Figure 2.13: Experimental results for an ensemble of $N = 10^5$ cat maps with and without noise: (a) shows k versus $\langle a^2 \rangle$ for the cases with noise and without noise. (b) show the power spectral densities of a_n at the positive critical values for the cases without noise ($k = k_{c1}^{(0)}$, solid graph) and with noise ($k = k_{c1}$, dash-dot graph).

which yields the critical values $k_{c1} = 2e^{3\sigma^2/2}$ at $\omega_{c1} = \pi/2$ and $k_{c2} = -2e^{3\sigma^2/2}$ at $\omega_{c2} = 0$.

Figure 2.13 shows results of numerical experiments for this system without noise and with noise ($\sigma^2 = 0.16$). Without noise, the critical values are $k_{c1}^{(0)} = 2$ at $\omega_{c1}^{(0)} = \pi/2$ and $k_{c2}^{(0)} = -2$ at $\omega_{c2}^{(0)} = 0$. With noise, the critical values are $k_{c1} \approx 2.54$ at $\omega_{c1} = \pi/2$ and $k_{c2} \approx -2.54$ at $\omega_{c2} = 0$. In Figure 2.13(a), we see that the transition occurs near the predicted critical values for each case. Also, Figure 2.13(b) shows that the predicted frequencies of oscillation at $k = k_{c1}^{(0)}$ and $k = k_{c1}$ (which is $\pi/2$ in both cases) match the peaks of the power spectral densities of a_n for each k .

2.5 Conclusion

A large class of globally coupled systems of chaotic maps experience a transition from incoherence to coherence at critical values of a coupling coefficient. We have shown that these critical values can be determined from a perturbation method, and we apply our method to ensembles of homogeneous chaotic maps, ensembles of chaotic maps with distributed parameters, and ensembles of chaotic maps with noise.

We have shown that numerical approximations to $Q(e^{i\omega})$ can be sufficiently accurate to yield good predictions for the transition, provided that a large enough number of elements is used in obtaining the approximations.

In our numerical experiments we obtained good agreement with our theory for ensembles of a large number of noisy Bernoulli maps, for a noiseless system of modified Bernoulli maps with distributed parameters, for noisy logistic maps, and for a noisy system of cat maps (a two dimensional example). However, we did not obtain useful results from our analysis when we attempted to apply it to an ensemble of identical noise-free logistic maps. We speculate that this may be due to the facts that the natural invariant density $\rho(x)$ of the logistic map, in common with other generic nonhyperbolic maps, has a dense countable set of x -values at which $\rho(x)$ is infinite and that the map is structurally unstable (it has a dense set of periodic windows in its chaotic parameter range). Structural instability, for example, implies that a small perturbation can result in totally different dynamics, and hence application of a perturbation method may be questionable. On the other hand, we have found that either noise or parameter spread appears to restore the validity of the perturbation theory approach. It would be worthwhile to further investigate noiseless ensembles of smooth maps that have a dense set of periodic windows.

Chapter 3

Localized Error Bursts in Estimating the State of Spatiotemporal Chaos

3.1 Introduction

Spatiotemporal chaos refers to systems that exhibit complex behavior both in time and space. There are many examples of such systems, including the Earth's atmosphere [21, 49], laboratory fluids [19, 64], chemical systems [28], oceans [76], etc. A common problem for systems of this type is that of estimating its current state. For example, the problem of weather forecasting consists in first estimating the current state of the atmosphere, and then producing a forecast by integration of an atmospheric model using the estimated current state as the initial conditions [21, 49]. In addition to forecasting, state estimation is also a key step if one wanted to *control* a spatiotemporally chaotic system: based on the current state estimate, active perturbations to the system would be intelligently applied so as to promote meeting desired goals for the future evolution of the system.

Even though the system is chaotic, if we had *perfect* knowledge of the current state and the dynamics of a system, in principle (i.e. with an infinite computer), we would be able to predict the future of the system for all time. In practice, however, observations of the system have errors and the model for the system is imperfect. The standard procedure in such circumstances is to maintain a good estimate of the system state by running the system model and periodically applying corrections to the model state based on observations. This process is called *data assimilation* [21, 49]. Its purpose is essentially that of accurately synchronizing the state of the

model system to the state of the observed system. Thus this problem might be thought of as being related to recent work on the synchronism of low dimensional chaotic dynamical systems¹ [84, 66, 73], although with the obvious difference that we are here dealing with a high dimensional spatiotemporally chaotic system. In the case of synchronization of low dimensional chaotic dynamical systems, it was found that there exists a threshold in parameter space past which a desynchronizing transition, called the bubbling transition, takes place [4, 5, 91, 92, 102]. One of the characteristic possible manifestations of the bubbling transition is the presence of intermittent bursts of desynchronization. In particular, slightly above the bubbling transition there are relatively short, irregularly occurring, epochs of large desynchronization, between which the systems are well synchronized. As the bubbling transition is approached from above, these desynchronization bursts become rarer, essentially disappearing below the transition [91, 92, 102].

The main finding of the present paper is that for data assimilation in spatiotemporally chaotic systems, a phenomenon similar to bubbling is possible. In particular, we perform numerical data assimilation experiments on a simple spatiotemporally chaotic model, and we observe bursts in the error of the state estimate (essentially desynchronization bursts). As compared to the previously mentioned work, the novel feature of these bursts is that they are spatially localized. Furthermore, in the case of a homogeneous observing network (see Section 3.3.1) the bursts are intermittent in *both space and time*. That is, individual burst events occur in a random-like manner in small localized spatial regions and time intervals. In what follows we will describe and characterize these space/time error bursts.

In Section 3.2, we briefly describe the Lorenz-96 model [56] used in our numerical experiments, as well as the data assimilation scheme that we use to estimate the system state. Also, we outline how we conduct our numerical experiments. Our

¹For example, So et al. [84] treat the problem of determining the state of a chaotic system by making use of the phenomenon of synchronization of chaotic process.

numerical results are presented and discussed in Section 3.3. Conclusions are given in Section 3.4.

3.2 Experimental Design

In this paper we consider an illustrative case of spatio-temporal chaos originally proposed by Lorenz [56] as a simplified testing ground for data assimilation techniques for atmospheric dynamics. The model has features similar to many other spatio-temporally chaotic systems, including local interactions, propagating nonlinear wave-like disturbances, high fractal dimension of the attractor, and many positive Lyapunov exponents. As such, we expect that our general findings for the behavior of data assimilation applied to this model will be similar to those of many other systems in nature (e.g., oceans, planetary cores, stellar atmospheres, etc.) and in the laboratory (e.g., experiments on large aspect ratio Rayleigh-Benard convection [M. Schatz, private communication]).

The Lorenz-96 model [56] is an N -dimensional spatiotemporally chaotic system defined by the system of differential equation

$$\frac{dx_i}{dt} = (x_{i+1} - x_{i-2})x_{i-1} - x_i + F, \quad i = 1, \dots, N, \quad (3.1)$$

where $x_{-1} = x_{N-1}$, $x_0 = x_N$, $x_{N+1} = x_1$ and F is a constant. The variables of this model can be thought of as values of a space and time dependent scalar variable at discrete locations around a circle (Fig. 3.1). Furthermore, interactions between the variables are allowed only at close locations. In our numerical experiments, we choose $N = 40$ and $F = 8$, for which the leading Lyapunov exponent corresponds to a doubling time of 0.42 dimensionless time unit and the fractal dimension of the attractor is 27.1 [56].

We generate a time series of true system state by integrating (3.1) with a

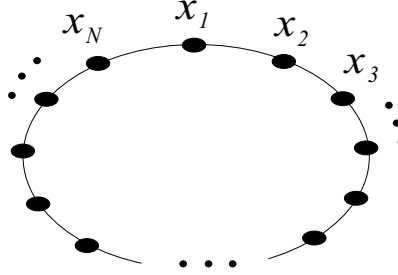


Figure 3.1: Illustration of the Lorenz-96 model.

fourth order Runge-Kutta scheme with time step $t_s = 0.05$ dimensionless time unit. To refer to the true state, we introduce the vector $\mathbf{x}^t = [x_1^t \ x_2^t \ \dots \ x_N^t]^T$ (component x_i^t is the true state at location i). The initial state $\mathbf{x}^t(0)$ is taken as the combination of the steady state solution, $x_i^t = F$ at each location i , and a small perturbation ($\sim 10^{-3}$) at a randomly selected location. For this initial condition, the behavior of the system appears to reach a statistically steady chaotic state after a short transient time, t_0 . We define $\mathbf{x}^t(n)$ by evaluating the true state n time steps after the transition ended, i.e., $\mathbf{x}^t(n) \equiv \mathbf{x}^t(t_0 + nt_s)$, where n is a positive integer.

We assume that the true state is observed at P of the N locations ($P \leq N$). A vector of simulated noisy observations $\mathbf{y} = [y_1, y_2, \dots, y_P]^T$ is generated by adding random noise to the true state at the observational locations, i.e.,

$$\mathbf{y} = \mathbf{H}\mathbf{x}^t + \mathbf{v}, \quad (3.2)$$

where $\mathbf{v} = [v_1, v_2, \dots, v_P]^T$ is a prescribed observation error and \mathbf{H} is a $P \times N$ matrix such that

$$H_{ij} = \begin{cases} 1, & \text{if observation } i \text{ is taken at location } j, \\ 0, & \text{otherwise.} \end{cases}$$

Our simple scheme for state estimation assumes that the observational errors are normally distributed with zero mean and error covariance matrix $\mathbf{R} = E[\mathbf{v}\mathbf{v}^T] = \sigma_v^2 \mathbb{I}_P$,

where $E[\cdot]$ denotes the expectation and \mathbb{I}_P is the $P \times P$ unity matrix. This particular choice of \mathbf{R} is equivalent to assuming that the observational error variance is σ_v^2 at all locations and the observational errors at the different locations are uncorrelated. The observational noise we generate, however, is uniformly distributed in the interval $(-\sqrt{3}\sigma_v, \sqrt{3}\sigma_v)$. The discrepancy between the assumed and true observational error statistics mimics common practice. For example, estimation schemes used in atmospheric science [21, 49] typically assume a normal error distribution for the sake of mathematical convenience, although this formulation allows for the occurrence of large errors on some rare occasion. Real observational error, on the other hand, are typically bounded.

By using the numerical solution of equation (3.1) to simulate the system dynamics, we assume that we have perfect knowledge of the behavior of the system. Our goal is to use this model and the noisy observations to obtain an estimate of the true state. This estimate, called the *analysis* and denoted by $\mathbf{x}^a(n)$, is chosen as the state \mathbf{x} that minimizes the quadratic cost function [1]

$$J(\mathbf{x}) = (\mathbf{x} - \mathbf{x}^b)^T (\mathbf{P}^b)^{-1} (\mathbf{x} - \mathbf{x}^b) + (\mathbf{y} - \mathbf{H}\mathbf{x})^T \mathbf{R}^{-1} (\mathbf{y} - \mathbf{H}\mathbf{x}). \quad (3.3)$$

Here $\mathbf{x}^b(n)$ is a short term forecast (called the *background*) obtained by integrating the model from the preceding analysis $\mathbf{x}^a(n-1)$. \mathbf{P}^b is the assumed covariance matrix of the background error, i.e., \mathbf{P}^b is an estimate of $E[\mathbf{e}^b(\mathbf{e}^b)^T]$, where $\mathbf{e}^b = \mathbf{x}^b - \mathbf{x}^t$. We note that $\mathbf{x}^a(n)$ is a maximum likelihood estimate of the true state if the observational and background errors are normally distributed, have zero means, and \mathbf{R} and \mathbf{P}^b are known [1]. Our choice for \mathbf{R} and its relation to the “true” observational error statistics have already been explained. After a short detour, in which we describe how the state estimate is updated, we will return to explain the algorithm to obtain \mathbf{P}^b .

The minimizer of $J(\mathbf{x})$ (Eq. (3.3)) is [1]

$$\mathbf{x}^a = \mathbf{x}^b + \mathbf{K}(\mathbf{y} - \mathbf{H}\mathbf{x}^b), \quad (3.4)$$

where

$$\mathbf{K} = ((\mathbf{P}^b)^{-1} + \mathbf{H}^T \mathbf{R}^{-1} \mathbf{H})^{-1} \mathbf{H}^T \mathbf{R}^{-1}, \quad (3.5)$$

is the *gain matrix*, representing the gain in the knowledge of the state after comparing the observations to the background (our best estimate of the state prior to collecting the observations). Based on numerical experiments, we choose the transient time t_0 to be 5,000 time steps. Once t_0 is passed, observation of the true state and data assimilation are performed once every 10 time steps.

Now we explain our iterative process (also described in Ref. [70]) to obtain \mathbf{P}^b . The algorithm starts with a reasonable, but more or less ad hoc estimate of \mathbf{P}^b . Then a time series of the “true” background error vectors $\mathbf{e}^b(n) = \mathbf{x}^t(n) - \mathbf{x}^b(n)$, $n = 1, 2, \dots, M$, is generated and the updated \mathbf{P}^b is chosen to be $(M - 1)^{-1} \sum_{n=1}^M \mathbf{e}^b(n)(\mathbf{e}^b(n))^T$. In the next iteration step, this updated \mathbf{P}^b is used in (3.4) and (3.5) to generate new samples of the “true” background errors. This step is repeated until the distance between consecutive estimates of \mathbf{P}^b becomes small, where we measure distance by the Frobenius matrix norm, i.e., the square root of the sum of the squares of all the matrix elements. (For a detailed description of the resulting \mathbf{P}^b see Appendix C.) We emphasize that our estimate of \mathbf{P}^b is not necessarily optimal in the sense of providing the smallest possible analysis error among all matrices, but it has the desirable feature that the true and assumed covariances are consistent ².

²We note that the \mathbf{P}^b we use is time independent, and thus, the resulting analysis may not be as good as could be obtained by use of a Kalman filter type procedure. However, it should be remarked that, due to its computational speed, a constant \mathbf{P}^b is used in important practical situations (e.g., weather forecasting). This is not a significant issue for our present paper, in that, the phenomenon we study, error bursting, is conjectured by us to occur independent of details in how \mathbf{P}^b determined. See also Sec. 3.4

The assimilations are done once every ten steps at times $n = 10m$. The analysis error for the m -th assimilation is defined as

$$\mathbf{e}^a(m) = \mathbf{x}^a(m) - \mathbf{x}^t(m). \quad (3.6)$$

We use the root-mean-square (rms)

$$\langle \mathbf{e}^a(m) \rangle = \sqrt{\frac{1}{N} \sum_{k=1}^N (e_k^a(m))^2} \quad (3.7)$$

of \mathbf{e}^a to assess the quality of the state estimate at a given time. Moreover, we are using the time mean of the rms error,

$$\langle \langle \mathbf{e}^a \rangle \rangle = \frac{1}{T} \sum_{m=1}^T \langle \mathbf{e}^a(m) \rangle, \quad T \gg 1, \quad (3.8)$$

to measure the overall performance of the estimation scheme.

3.3 Numerical Experiments

3.3.1 Experiment 1

In the first numerical experiment, we observe all model variables; i.e., $P = N = 40$ and \mathbf{H} is the $N \times N$ identity matrix, \mathbb{I}_N . The chaotic behavior of the model for our chosen parameters is illustrated by showing the evolution of the true state for a randomly selected short time interval in Fig. 3.2. We see that the spatiotemporal patterns are wave-like, but the amplitude and the period are not regular in time or space.

In Fig. 3.3, for the case $\sigma_v = 10^{-4}$, we plot the rms analysis error (3.7) along with the time mean rms error (3.8) for the period $1 \leq m \leq 10^5$. We can see that the rms error at a given time is typically similar to the time mean error, but there are a

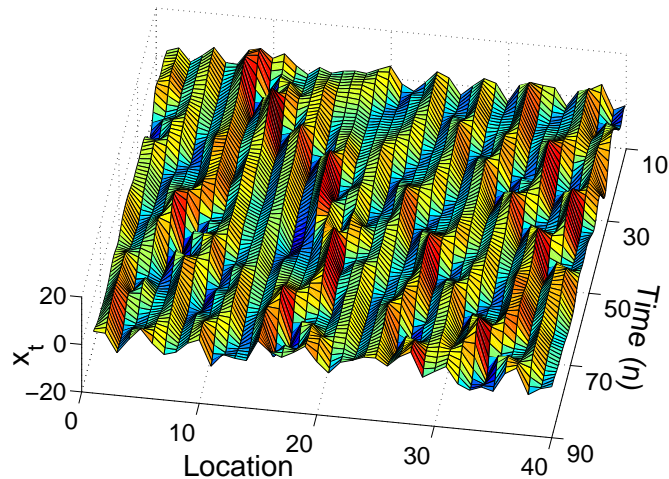


Figure 3.2: Time evolution of the “true state” for a randomly selected short time interval.

few occasions on which the error far exceeds its time mean. We call these occasions *bursts*. More precisely, we define bursts by choosing a threshold, $2\langle\langle\mathbf{e}^a\rangle\rangle$, i.e., a burst occurs at time m if $\langle\mathbf{e}^a(m)\rangle \geq 2\langle\langle\mathbf{e}^a\rangle\rangle$. For this threshold, a burst occurs once in 4300 assimilation steps on average. Also, it is important to note that bursts do not exist longer than one or two assimilation steps.

In what follows, we further investigate the spatial distribution of the analysis errors during burst events. In Fig. 3.4, we plot the spatial distribution of the background (forecast) error, $\mathbf{e}^b = \mathbf{x}^b - \mathbf{x}^t$ and the analysis error $\mathbf{e}^a = \mathbf{x}^a - \mathbf{x}^t$ for the burst that occurred at an arbitrarily chosen time, $m = 82585$. We see that the large analysis errors are confined to a narrow region of locations. Another important feature is the strong similarity between the distribution of the analysis errors and background errors. This indicates that the large localized analysis errors are due to large localized background errors at the same locations. Bursts at other times we examined were qualitatively similar.

To quantify the spatial extent of localization in the error patterns, we define

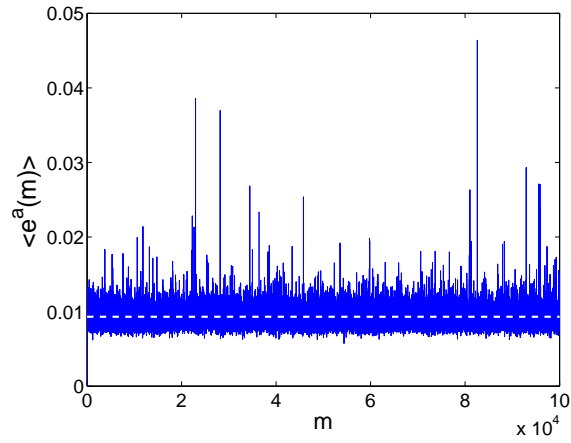


Figure 3.3: Time evolution of the rms analysis error. Also shown is the time mean analysis error (white dashed line), $\langle e^a \rangle = 0.0093$.

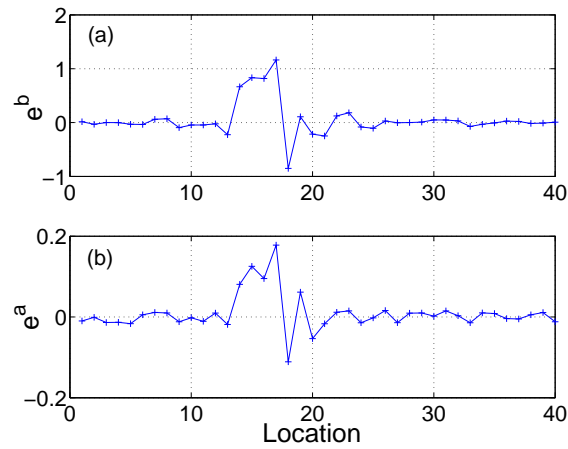


Figure 3.4: Spatial distribution of the background error (a) and the analysis error (b) at a given time during a burst event. Note that the difference between the scales of the y-axes.

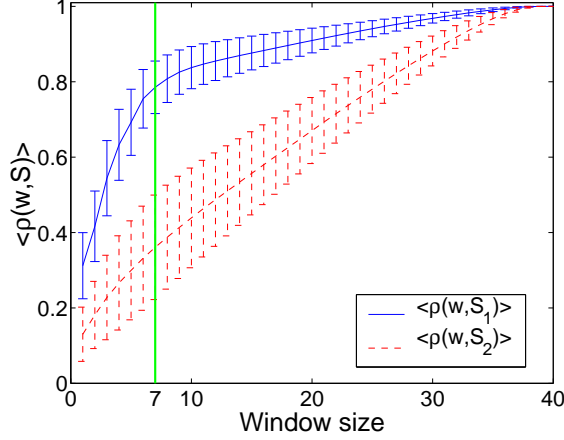


Figure 3.5: $\langle \rho(w, \mathbb{S}_1) \rangle$ (solid) and $\langle \rho(w, \mathbb{S}_2) \rangle$ (dashed).

a ratio,

$$\rho(w, m) = \max_j \frac{\sum_{k=j}^{j+w-1} (e_k^a(m))^2}{\sum_{k=1}^N (e_k^a(m))^2}. \quad (3.9)$$

This ratio tells us the largest portion of the total rms error that is confined to a local window of length w at a given time. We are primarily interested in the average of $\rho(w, m)$ over given sets (\mathbb{S}) of the time indices,

$$\langle \rho(w, \mathbb{S}) \rangle = \frac{1}{|\mathbb{S}|} \sum_{m \in \mathbb{S}} \rho(w, m), \quad (3.10)$$

where $|\mathbb{S}|$ is the size of \mathbb{S} . We evaluate the average (3.10) over all burst events, i.e., $\mathbb{S}_1 = \{m : \langle \mathbf{e}_a(m) \rangle \geq 2 \langle \langle \mathbf{e}_a \rangle \rangle\}$, and also over a randomly selected set of times, \mathbb{S}_2 , at which bursts are not observed, i.e., $\mathbb{S}_2 = \{m : m \notin \mathbb{S}_1 \text{ and } |\mathbb{S}_2| = |\mathbb{S}_1|\}$.

In Fig. 3.5, we plot the average ratio, $\langle \rho(w, \mathbb{S}_1) \rangle$ with $|\mathbb{S}_1| = 3077$ (solid line), and $\langle \rho(w, \mathbb{S}_2) \rangle$ with $|\mathbb{S}_2| = 3077$ (dashed line). The largest difference between the ratios $\langle \rho(w, \mathbb{S}_1) \rangle$ and $\langle \rho(w, \mathbb{S}_2) \rangle$ is at window size $w = 7$ (gray vertical line in Fig. 3.5) indicating that bursts are typically confined to a window of size 7.

In Fig. 3.6, we plot the number of bursts with window size 7 at each location. That is, we evaluate $\rho(w, m)$ in Eq. (3.9) for $w = 7$ and $m \in \mathbb{S}_1$, and we record the j which yields $\rho(7, m)$. We find that the $|\mathbb{S}_1| = 3077$ bursts are fairly evenly

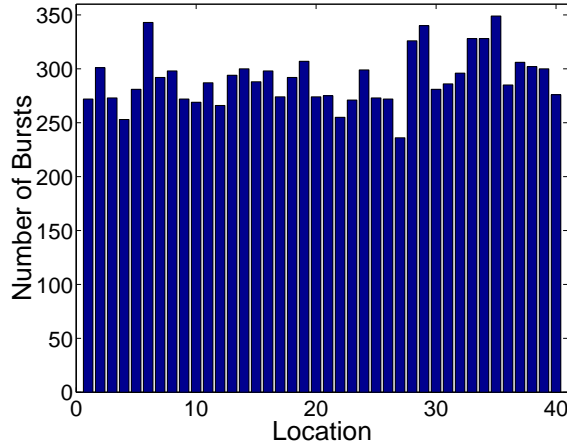


Figure 3.6: The spatial distribution of the burst events.

distributed over all locations. This result is not unexpected since neither the model dynamics, nor the observational errors change with location.

A similar kind of bursting phenomenon, known as a bubbling transition, was observed in two coupled chaotic oscillators [4, 5, 91, 92, 102]. For a certain coupling strength, in the absence of noise, the system of two identical coupled chaotic oscillators has a stable manifold on which the two oscillators are synchronized. Additionally, the manifold may possess transversely unstable invariant sets. In the presence of parameter mismatch or noise, the synchronized trajectories are near to, but not exactly on, the stable manifold. When the trajectories come close to the unstable invariant sets, they are repelled in the transverse direction (the two oscillators are desynchronized), and, if there are no other attractors, the trajectories subsequently return to being close to the stable manifold. These desynchronized bursts occur as a system parameter passes the bubbling transition [102]. In our case, a one way coupling is established once every 10 time steps through the assimilation process. In other words, the data assimilation attempts to synchronize the model system to the true system based on the observations. However, due to the noise in the observations, the two systems cannot be exactly synchronized, hence bursts may occur. We note that bursts are not seen when the observational noise

variance σ_v^2 is reduced and that bursts become more frequent with increasing σ_v^2 . Although we have not attempted an analysis of our high dimensional system establishing the burst mechanism in terms of unstable invariant sets (as previously done for low dimensional systems [91, 92, 102]), the similarity of the burst phenomenon we observe to that for low dimensional synchronized chaotic systems leads us to conjecture that the mechanisms involved might be also similar.

3.3.2 Experiment 2

Our second numerical experiment is identical to the first one except that observations are never taken at a fixed location (location 20). This seemingly slight change leads to important changes in the spatial distribution of the analysis error. To better understand these changes, we recall that the analysis at the j -th location is obtained by

$$x_j^a = x_j^b + \sum_{k=1}^P K_{jk} [y_k - (\mathbf{H}\mathbf{x}^b)_k]. \quad (3.11)$$

The (j, k) element of \mathbf{K} , K_{jk} , controls how an observation taken at location k affects the state estimate at location j . [$(\mathbf{H}\mathbf{x}^b)_k$ is the k -th element of the vector $\mathbf{H}\mathbf{x}^b$.] Since \mathbf{R} is always diagonal in our experiments, and \mathbf{P}^b is near diagonal when observations are taken at each location, \mathbf{K} is also near diagonal. Thus the state estimate at a given location is to a rough approximation a linear combination of the background and the observation at that same location. The situation is substantially different at a location, where observations are not taken. This is illustrated by Fig. 3.7, in which two rows of \mathbf{K} are plotted; one of them ($K_{19,k}$) is associated with a typical location (location 19), while the other one ($K_{20,k}$) is associated with the missing observation. At location 20, the elements of the gain matrix \mathbf{K} are smaller, indicating smaller corrections in the state estimate at that location based on the observed information. Also, the corrections are influenced by a number of observations at nearby locations.

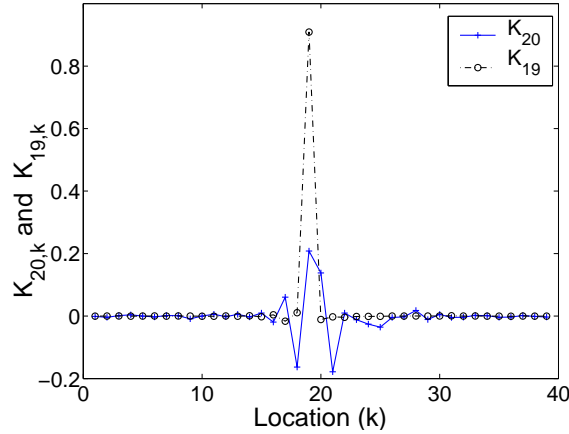


Figure 3.7: The 19-th ($K_{19,k}$) and 20-th ($K_{20,k}$) row of the gain matrix K for Experiment 2.

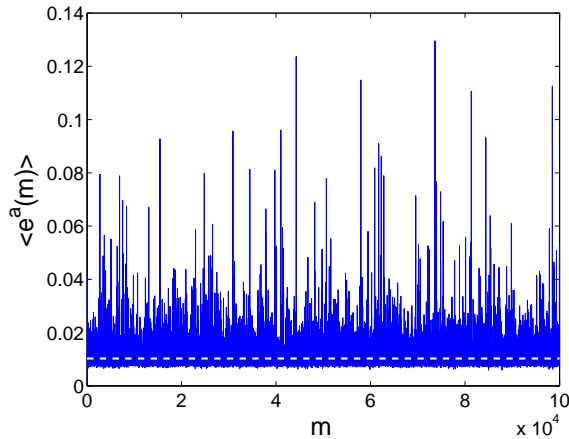


Figure 3.8: Time evolution of the rms analysis error. Also shown is the time mean analysis error (white dashed line), $\langle\langle e^a \rangle\rangle = 0.0103$.

The rms error for $1 \leq m \leq 10^5$ is plotted in Fig. 3.8. By comparing Figs. 3.3 and 3.8, we can see that bursts now occur much more frequently than in the case in which all locations are observed. The time mean rms error is now $\langle\langle e^a \rangle\rangle \approx 0.0103$ which is one order of magnitude higher than in the previous experiment (we increase the threshold value ($2\langle\langle e^a \rangle\rangle$) accordingly). The average frequency of bursts is about one per 1069 assimilation steps.

The large analysis errors during burst events are localized, but as an important

new feature, they are almost always confined to the location where observations are not taken (Figs. 3.9 and 3.10(a)). The portion of bursts occurring at location other than 20 is less than 1% (the vertical scale in Fig. 3.9(a) goes up to 250,000 while that in Fig. 3.9(b) goes up to 40). Fig. 3.10(a) suggests that bursts occur due to insufficient correction of the state estimate at location 20. This is in contrast to the behavior of bursts, in which the background and analysis errors have the same spatial distribution. Another interesting feature, seen in Fig. 3.9, is the very efficient reduction of background errors by the data assimilation step between locations 21 and 26. This is due to the specific structure of P^b , that correctly expects an elevated error level in this region, thus forcing the analysis to give larger weights to the observations (see Appendix).

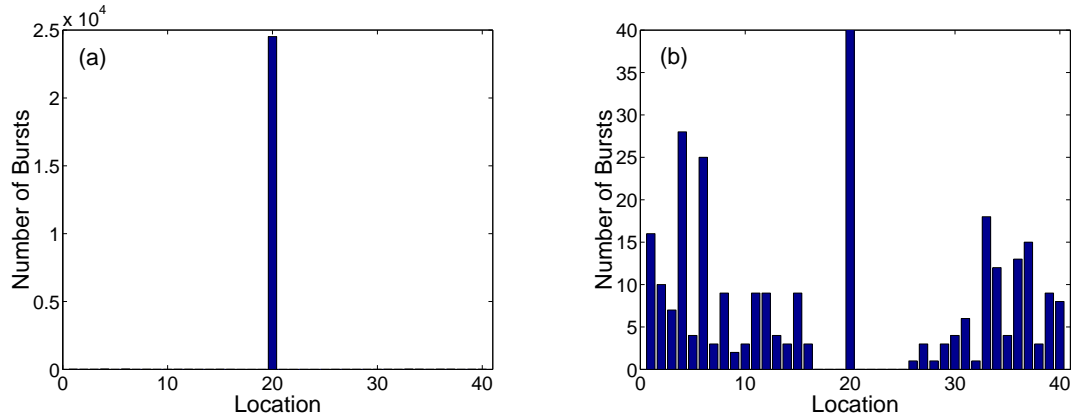


Figure 3.9: The spatial distribution of the burst events. (b) has a different vertical scale from (a) in order to see the burst events at other locations.

Fig. 3.11 indicates that there is also an important change in the spatial extent of the burst. The largest difference between $\langle \rho(w, \mathbb{S}_1) \rangle$ and $\langle \rho(w, \mathbb{S}_2) \rangle$ is at window size $w = 1$, which is very different from $w = 7$ observed in Experiment 1.

Finally, we note that we also conducted experiment in which observations were not taken at two far apart locations. The results were very similar to those shown in Experiment 2, except that the bursts were approximately equally distributed

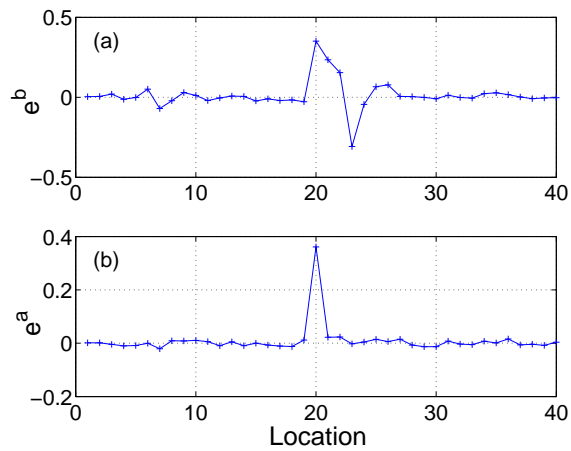


Figure 3.10: Spatial distribution of the background error (a) and the analysis error (b) at a given time during a burst event.

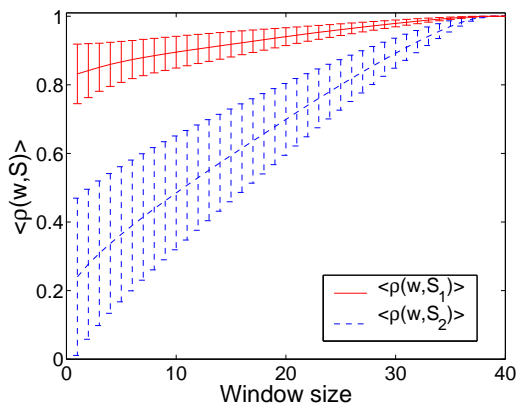


Figure 3.11: $\langle \rho(w, \mathbb{S}_1) \rangle$ (solid) and $\langle \rho(w, \mathbb{S}_2) \rangle$ (dashed).

between the two locations. However, when two neighboring locations were not observed, our iterative procedure for the estimation of \mathbf{P}^b failed. We decided not to pursue this experiment.

3.4 Conclusion and Discussion

In this paper, we have demonstrated that large, temporally intermittent and spatially localized error bursts can occur in data assimilation estimates of the state of a spatiotemporally chaotic system. We considered two scenarios; in one of them, the system was fully observed, while in the other, one of the model variables was never observed. Error bursts occurred in both cases, though some characteristics of the burst events were distinctly different in the two experiments. When all variables were observed, bursts occurred with equal likelihood at the different locations, and they were typically confined to a 7-location wide local region. On the other hand, when a selected variable was never observed, the burst events were almost always confined to the unobserved location and were much more localized.

In our estimation scheme \mathbf{P}^b is constant, which is the formulation most operational weather prediction center (e.g., the National Weather Services) has implemented. There exists a more advanced family of schemes, the Kalman filters, in which \mathbf{P}^b is dynamically evolved (e.g., Ref [71] and references therein). It was shown [70] that bursts can occur even in a Kalman filter scheme and that the occurrence of bursts can be prevented by inflating (artificially increasing) the background error covariance \mathbf{P}^b . Inflating \mathbf{P}^b can be viewed as increasing the coupling between the observed and model systems at estimation time (see Eqs. (3.4) and (3.5)). In this sense, the bursting observed in this paper is similar to the phenomenon of bubbling transition in the synchronization of two low dimensional chaotic systems: a weaker coupling parameter leads to intermittent bursts of desynchronization.

In general, the error burst phenomenon we find is undesirable in practice. For

example, this phenomenon may explain some of the occasional large errors known to occur in operational numerical weather prediction systems (e.g., see Fig. 5.6.1 in Ref [49]). Thus, when designing data assimilation systems, the avoidance of rare large error may be an essential goal in addition to the more obvious goal of maintaining a low *time averaged* error.

Chapter 4

Local Ensemble Kalman Filtering in the Presence of Model Bias

4.1 Introduction

In many situations the dynamics of a real process may differ from those of the best available model of that process. We refer to this difference as *model error*. Model error is thought to be a key issue in weather forecasting in that the presence of model error can lead to large discrepancies between the forecasts and the true atmospheric states. In this connection, we note (i) that Kalman filters have been considered for estimating atmospheric states to be used as initial conditions in forecast models [34], and, (ii) that the general Kalman filter methodology has long been adapted to account for model error [33]. Recently, the ensemble technique has been proposed as a computationally feasible means of applying Kalman filtering to the very high dimensional states inherent in global atmospheric models [31, 45, 44, 2, 11, 40, 95, 70]. One of our goals in this paper is to investigate the incorporation of model error correction in an ensemble Kalman filter. Here, we consider only the case where the evolution of the model error is governed by a deterministic equation, i.e., the model error has no random component. We refer to this type of error as model bias, since when the state of the model is described by a probabilistic variable, as is the case in data assimilation, such errors become equal to the expected error in the model forecast.

We will restrict our considerations to the example of one particular ensemble Kalman filter, the Local Ensemble Kalman Filter (LEKF) proposed by Ott et al. [70]. (This scheme has been successfully tested on a reduced resolution version of the operational Global Forecast System (GFS) of the National Centers for Envi-

ronmental Prediction (NCEP) for the perfect forecast model scenario [87].) We believe that our results in the present paper, using the LEKF example, may also be more generally applicable, providing an indication of what to expect if model bias correction is attempted using other related ensemble Kalman filter methods. In addition, some of our ideas may also be useful in designing weakly constrained 4DVAR schemes [e.g., 26, 103], which also allow for model errors.

The technique we propose belongs to the family of schemes usually called *state space augmentation* methods [18]. In these techniques the state vector is augmented with the uncertain model parameters, and the augmented state is estimated using the forecast model in conjunction with observations. In principle, the uncertain parameters can occur in otherwise completely known forecast model equations. In such a case, the augmented state space approach may provide an accurate estimate of the parameters even for a highly chaotic system, as recently demonstrated on a simple model by Annan and Hargreaves [3]. In reality, the equations governing the motion of the atmospheric flow are not known exactly, thus uncertainties also arise due to our limited knowledge of the dynamics. Also, estimating all parameters of the forecast model equations would not be computationally feasible due to the large number of the parameters. A practical approach, first suggested by Derber [26], is to assume that the uncertainties in the forecast model can be approximately represented in the form of a limited number of bulk error terms. Then the task is to estimate the parameters of the bulk error terms. We recall that since the error terms are modeled as random vectors, the parameters to be estimated are the mean errors (model biases).

The information encapsulated in the bias can be used either to modify the forecast model equations or to modify the analysis scheme. Here we follow the second approach. That is, we treat the forecast model as a “black box”, that does not yield the true time evolution of the atmosphere, and we attempt to use this

black box in conjunction with observations to account for model bias in the state estimation.

The two key components of the aforementioned strategy are the selection of a bias model that efficiently represents the bias and the design of a computational strategy that can efficiently estimate the parameters of the bias model. The most common assumption is that the bias is constant or has a simple evolution in time. It is also frequently assumed that the uncertainties in the forecast model state and the bias are uncorrelated. These assumptions were used to derive the bias estimation schemes of Dee and Da Silva [24], Dee and Todling [25], Carton et al. [15], Martin et al. [61], and Bell et al. [10].

The scheme we propose allows for correlations between the uncertainties of the forecast model state and the bias. This additional flexibility is necessitated by the structure of our technique (see Section 4.2.4), and is affordable due to the high computational efficiency of the LEKF approach. In Section 4.2, we introduce three different bias models. Bias Model I is a simple additive correction to the model forecast. Bias Model II is motivated by envisioning a situation in which the forecast model evolution takes place on an attractor that is shifted from the attractor for the true system evolution. Bias Model III is essentially a combination of Bias Models I and II. Section 4.3 presents the results of numerical experiments with the Lorenz-96 model [e.g., 56] for several cases of the difference between the forecast model evolution and the evolution of the true state. Conclusions and discussion follow in Section 4.4.

A main result is the importance of selecting a bias model that effectively parameterizes the form of possible bias in the forecast model. In particular, if the bias model can parameterize the possible bias of the forecast model, then our results suggest that substantial improvement in forecasts may result. On the other hand, if the parameterization of the model bias through the bias model does not sufficiently

capture the form of the true biases in the forecast model, then substantial forecast improvements were not obtained in our numerical experiments.

4.2 Bias Modeling and Data Assimilation

The discrepancy between the forecast model evolution and the evolution of the real atmosphere has two sources: (i) due to numerical solution on a grid of a finite number of points, the forecast model state is a finite dimensional representation of the infinite dimensional atmospheric fields, and (ii) the equations that govern the true evolution of the atmosphere are not known exactly. These two sources of forecast model errors are not independent, since the errors in the forecast model formulation are mainly associated with the inherent problems of considering only a limited number of interactions between the finite number of components of the state vector and the imperfect representation of the effects of the subgrid processes on the motions at the resolved scales.

Denoting the true atmospheric state at t_n by \mathbf{x}_n^t , the true atmospheric evolution is denoted

$$\mathbf{x}_{n+1}^t = \mathbf{F}^t(\mathbf{x}_n^t), \quad (4.1)$$

where \mathbf{F}^t is the (unknown) true atmospheric evolution operator and \mathbf{x}_{n+1}^t is the true atmospheric state at time $t_{n+1} = t_n + \Delta t$. Denoting the forecast model state at time t_n by \mathbf{x}_n^m , the black-box produces a forecast model state at time t_{n+1} ,

$$\mathbf{x}_{n+1}^m = \mathbf{F}^m(\mathbf{x}_n^m), \quad (4.2)$$

where \mathbf{F}^m is the forecast model evolution operator that mimics \mathbf{F}^t . Note that the dimensions of \mathbf{x}^t and \mathbf{x}^m are, in general, different; for example, in the case of real weather forecasting the true state is infinite dimensional and the forecast model state is finite dimensional. In what follows we will treat a scenario in which \mathbf{x}^m and \mathbf{x}^t

have the same (finite) dimensionality, and we henceforth assume this circumstance. With respect to the situation of atmospheric weather forecasting, this assumption restricts the character of the errors and their means (biases) that can be addressed by our bias models. In particular, we regard our treatment to follow as addressing only those types of forecast model biases that can be represented as dynamics in the state space of the forecast model variables. Thus, we ignore dynamics that occur at smaller scales than the forecast model resolves. On the other hand, if we think of the dynamics at the unresolved scales as random perturbations to the forecast model dynamics, our methods may be able to correct for the mean bias due to such perturbations. Meanwhile, the uncertainty in the small-scale fluctuations is modeled as representativeness error in the observation error statistics.

In this section we define three ways of modeling the bias that can arise due to forecast model error. We refer to these as Bias Model I, Bias Model II, and Bias Model III.

4.2.1 Bias Model I

In general, it is desired to have the forecast model state as close to the true state as possible so that, assuming that the forecast model and the true evolution operators are the same, the forecast model state stays near the true state after its evolution. In practice, however, the forecast model evolution operator differs from the true evolution operator. As a result, even if we evolve the forecast model state from an initial condition corresponding to the true state at the initial time (e.g., $\mathbf{x}_{n-1}^m = \mathbf{x}_{n-1}^t$), it is likely that the forecast model state departs from the true state as it evolves. In Bias Model I, we attempt to estimate

$$\mathbf{b}_n^t = \mathbf{F}^t(\mathbf{x}_{n-1}^t) - \mathbf{F}^m(\mathbf{x}_{n-1}^t), \quad (4.3)$$

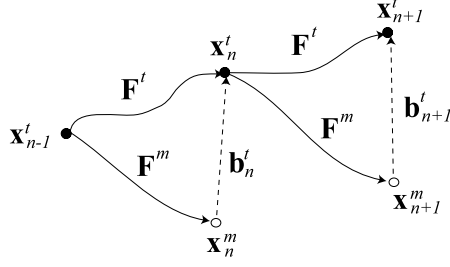


Figure 4.1: Illustration of Bias Model I: \mathbf{x}_n^t is the true state evolving according to \mathbf{F}^t from the previous true state, and \mathbf{x}_n^m is the forecast model state evolving according to \mathbf{F}^m from the previous true state.

i.e., the departure of the forecast model state from the true state as illustrated in Fig. 4.1.

In order to estimate \mathbf{b}_n^t , we must use some model of how it is related to past values of the bias \mathbf{b}_i^t ($i < n$). For example, we can assume that the bias is constant in time. If so, the true system evolution can be written in terms of the model evolution as follows:

$$\mathbf{x}_n^t = \mathbf{F}^m(\mathbf{x}_{n-1}^t) + \mathbf{b}_n^t, \quad (4.4)$$

$$\mathbf{b}_n^t = \mathbf{b}_{n-1}^t. \quad (4.5)$$

Though we could write this system more concisely as $\mathbf{x}_n^t = \mathbf{F}^m(\mathbf{x}_{n-1}^t) + \mathbf{b}$, where \mathbf{b} is an unknown parameter vector, we write the system in terms of the augmented state vector $(\mathbf{x}_n^t, \mathbf{b}_n^t)$ in order to facilitate the iterative estimation of both \mathbf{x}_n^t and \mathbf{b}_n^t by our data assimilation procedure. More generally, we can replace eq. (4.5) with another model for the bias of the form,

$$\mathbf{b}_n^t = \mathbf{G}^b(\mathbf{b}_{n-1}^t, \mathbf{x}_{n-1}^t), \quad (4.6)$$

where \mathbf{G}^b is the evolution operator for the bias correction term. Another alternative is to assume that the model error evolution is a Markov process. In that case \mathbf{b}_n^t is

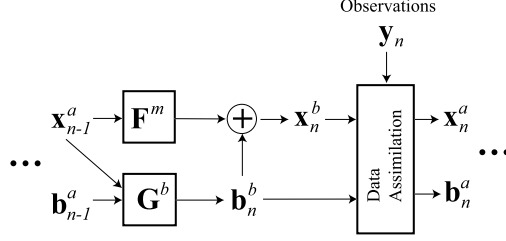


Figure 4.2: Illustration of data assimilation with Bias Model I: Data assimilation produces an unbiased analysis for the true state, \mathbf{x}_n^a , and an analysis for the bias correction term, \mathbf{b}_n^a .

represented as \mathbf{b}_{n-1}^t multiplied by a matrix that describes the temporal covariance between the model errors at different spatial location, and the right-hand side of eq. (4.6) also includes an additive random term [e.g., 48, 20, 103].

Then given an estimate $(\mathbf{x}_{n-1}^a, \mathbf{b}_{n-1}^a)$ of the augmented state vector at time t_{n-1} (the “analysis” from the previous data assimilation), we take the forecast (or “background”) of this vector $(\mathbf{x}_n^b, \mathbf{b}_n^b)$ at time t_n to be

$$\mathbf{x}_n^b = \mathbf{F}^m(\mathbf{x}_{n-1}^a) + \mathbf{b}_n^b, \quad (4.7)$$

$$\mathbf{b}_n^b = \mathbf{G}^b(\mathbf{b}_{n-1}^a, \mathbf{x}_{n-1}^a), \quad (4.8)$$

where we have assumed bias evolution by eq. (4.6). We then perform data assimilation using $(\mathbf{x}_n^b, \mathbf{b}_n^b)$, and the observations at time t_n to obtain the analysis $(\mathbf{x}_n^a, \mathbf{b}_n^a)$.

This way of taking forecast model error into account is illustrated in Fig. 4.2, and is the general scheme used in several previous methods appearing in the literature [e.g., 24, 25, 15, 61, 10]. The vector \mathbf{y}_n in Fig 4.2 is the observation of the true state at time t_n , which we assume to obey a model equation of the form,

$$\mathbf{y}_n = \mathbf{H}(\mathbf{x}_n^t) + \epsilon_n, \quad (4.9)$$

where \mathbf{H} is the observation operator mapping the true states to the observations. and ϵ_n is the observational noise.

Basically, in Bias Model I it is supposed that the best forecast is produced when the input to the forecast model evolution is as close to the truth as possible. One can imagine problems with this. For example, say that atmospheric balance for the forecast model is not the same as that for the true atmosphere. Then, if a very good estimate of the true state at time t_n is inserted into the forecast model, the forecast model state at time t_n could often be out of balance, and spurious gravity wave activity might be excited. In the practice of numerical weather prediction, such spurious gravity wave activity is prevented by a filtering process, called initialization, applied to the fields provided by the data assimilation process [e.g., 59, 9, 58, 57]. The general wisdom is that a well designed Kalman filter might eliminate the need for initialization process. This consideration motivates Bias Model II.

4.2.2 Bias Model II

A consequence of the imperfect model is that the forecast model system has a different attractor from the true system. In some cases, it might be desirable to let the forecast model state follow its own attractor, since plugging a very good approximation of the true state into the forecast model system can result in completely different dynamics (like gravity wave excitation). In addition, one can envision a situation in which forecast model dynamics and true dynamics become more similar through an (a priori unknown) coordinate transformation. For instance, such transformations were rigorously derived to correct for truncation errors in numerical solution of the two-dimensional Navier-Stokes equations [60]. Having found a similar transformation for the weather prediction model, we may obtain a better estimate of the true trajectory by applying this transformation to an appropriate forecast model trajectory after it has been computed than by forcing the forecast model state to be close to the truth and then computing its trajectory. For simplicity, we assume the transformation is just a shift of the forecast model state to the

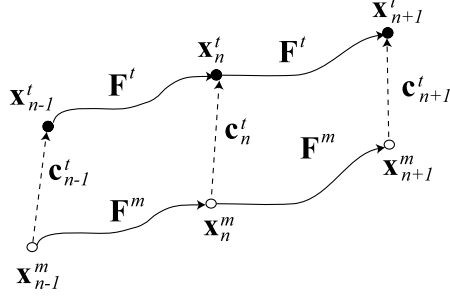


Figure 4.3: Illustration of Bias Model II: \mathbf{x}_n^t and \mathbf{x}_n^m evolve according to their own dynamics but the behavior of the forecast model is similar to the behavior of the truth.

true state, and we define the bias \mathbf{c}_n^t at time t_n by

$$\begin{aligned} \mathbf{c}_n^t &= \mathbf{F}^t(\mathbf{x}_{n-1}^t) - \mathbf{F}^m(\mathbf{x}_{n-1}^m) \\ &= \mathbf{F}^t(\mathbf{x}_{n-1}^t) - \mathbf{F}^m(\mathbf{x}_{n-1}^t - \mathbf{c}_{n-1}^t). \end{aligned} \quad (4.10)$$

A schematic illustration of this bias model is shown in Fig. 4.3. The forecast model state is not pushed to the true state. Instead, it mimics the true dynamics in a shifted location of the state space.

Unlike the bias in Bias Model I, the bias in Bias Model II at time t_n depends not only on \mathbf{F}^t , \mathbf{F}^m , and \mathbf{x}_{n-1}^t , but also on the previous bias \mathbf{c}_{n-1}^t . Nonetheless, we may assume that for some choice of \mathbf{c}_{n-1}^t , the correction term \mathbf{c}_n^t approximately obeys a simplified evolution model such as $\mathbf{c}_n^t = \mathbf{c}_{n-1}^t$, or more generally $\mathbf{c}_n^t = \tilde{\mathbf{G}}^c(\mathbf{c}_{n-1}^t, \mathbf{x}_{n-1}^t)$. In terms of this model, we approximate the true system evolution by the augmented model system,

$$\mathbf{x}_n^t = \mathbf{F}^m(\mathbf{x}_{n-1}^t - \mathbf{c}_{n-1}^t) + \mathbf{c}_n^t, \quad (4.11)$$

$$\mathbf{c}_n^t = \tilde{\mathbf{G}}^c(\mathbf{c}_{n-1}^t, \mathbf{x}_{n-1}^t). \quad (4.12)$$

For this bias model (and for Bias Model III to follow), our goal is not that the analysis state vector \mathbf{x}_n^a closely approximates the true state \mathbf{x}_n^t , but rather that

it approximates the best forecast model state $\mathbf{x}_n^m = \mathbf{x}_n^t - \mathbf{c}_n^t$ from which to make future forecasts. Thus, we rewrite eqs. (4.11) and (4.12) as

$$\mathbf{x}_n^m = \mathbf{F}^m(\mathbf{x}_{n-1}^m), \quad (4.13)$$

$$\mathbf{c}_n^t = \mathbf{G}^c(\mathbf{c}_{n-1}^t, \mathbf{x}_{n-1}^m), \quad (4.14)$$

where $\mathbf{G}(\mathbf{c}, \mathbf{x}) = \tilde{\mathbf{G}}^c(\mathbf{c}, \mathbf{x} + \mathbf{c})$. We can then write the background augmented state vector $(\mathbf{x}_n^b, \mathbf{c}_n^b)$ in terms of the previous analysis $(\mathbf{x}_{n-1}^a, \mathbf{c}_{n-1}^a)$ as follows:

$$\mathbf{x}_n^b = \mathbf{F}^m(\mathbf{x}_{n-1}^a), \quad (4.15)$$

$$\mathbf{c}_n^b = \mathbf{G}^c(\mathbf{c}_{n-1}^a, \mathbf{x}_{n-1}^a), \quad (4.16)$$

In taking this approach, one must keep in mind that the bias should be added to the forecast model state vector whenever making comparisons to observations. Thus instead of eq. (4.9), we use the observation model,

$$\mathbf{y}_n = \mathbf{H}(\mathbf{x}_n^m + \mathbf{c}_n^t) + \boldsymbol{\epsilon}_n \quad (4.17)$$

when performing data assimilation. The analysis $(\mathbf{x}_n^a, \mathbf{c}_n^a)$ represents an approximation to the augmented state vector $(\mathbf{x}_n^m, \mathbf{c}_n^t)$, and thus forecasts made using \mathbf{x}_n^a as the initial condition should also be corrected by the approximated bias in order to better predict the true system. Data assimilation with Bias Model II is illustrated in Fig. 4.4.

To the best of our knowledge, Bias Model II is a novel approach to the effects of model errors on the accuracy of the state estimates. Hansen [41] also argued for the model attractor, but he suggested the use of a multimodel approach as opposed to the state augmentation we propose.

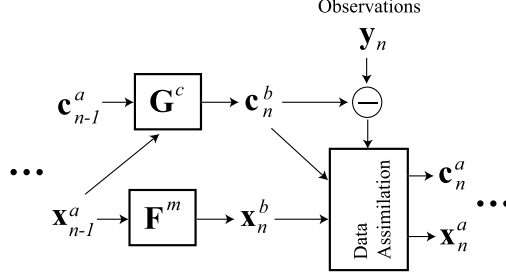


Figure 4.4: Illustration of data assimilation with Bias Model II: Data assimilation produces an analysis of the best forecast model state, $\mathbf{x}_n^m = \mathbf{x}_n^t - \mathbf{c}_n^t$, and an analysis for the correction term, \mathbf{c}_n^t .

4.2.3 Bias Model III

In Bias Model III, we combine Bias Model I and Bias Model II. Formally, we combine the equations describing the previous two bias models in the following manner. We evolve the analysis augmented state vector $(\mathbf{x}_{n-1}^a, \mathbf{b}_{n-1}^a, \mathbf{c}_{n-1}^a)$ to the background at the next step using the model

$$\mathbf{x}_n^b = \mathbf{F}^m(\mathbf{x}_{n-1}^a) + \mathbf{b}_n^b \quad (4.18)$$

$$\mathbf{b}_n^b = \mathbf{G}^b(\mathbf{x}_{n-1}^a, \mathbf{b}_{n-1}^a, \mathbf{c}_{n-1}^a) \quad (4.19)$$

$$\mathbf{c}_n^b = \mathbf{G}^c(\mathbf{x}_{n-1}^a, \mathbf{b}_{n-1}^a, \mathbf{c}_{n-1}^a), \quad (4.20)$$

and we compare the background state with observations according to eq. (4.17). Since \mathbf{x}_n^b and \mathbf{c}_n^b represent the best available approximations to \mathbf{x}_n^m and \mathbf{c}_n^t prior to the data assimilation at time t_n , the observation increment we use is

$$\mathbf{y}_n - \mathbf{H}(\mathbf{x}_n^b + \mathbf{c}_n^b). \quad (4.21)$$

Notice that if $\mathbf{G}^b(\mathbf{x}, \mathbf{b}, \mathbf{c}) = \mathbf{0}$, then this model reduces to Bias Model II, while if $\mathbf{G}^c(\mathbf{x}, \mathbf{b}, \mathbf{c}) = \mathbf{0}$, this model reduces to Bias Model I. In its simplest form, our model uses $\mathbf{G}^b(\mathbf{x}, \mathbf{b}, \mathbf{c}) = \mathbf{b}$ and $\mathbf{G}^c(\mathbf{x}, \mathbf{b}, \mathbf{c}) = \mathbf{c}$. However, we find a slightly different bias evolution function to be advantageous in some situations (see Sec. 4.3.6).

4.2.4 Augmented Local Ensemble Kalman Filter

For the purposes of all subsequent discussion we henceforth take the system state at time t_n to be a scalar variable $x_{n,i}$ defined on a discrete one dimensional spatial domain, $i = 1, 2, \dots, N$. Thus we represent the system state as a vector $\mathbf{x}_n = [x_{n,1}, x_{n,2}, \dots, x_{n,N}]^T$, where the superscript T denotes the transpose.

Once a suitable model for the bias is chosen, it can be incorporated into the formulation of the Kalman filter. For example, in the case of Bias Model III, the new equations can be obtained by replacing the state \mathbf{x}_n , by the augmented state, $\mathbf{v}_n = [\mathbf{x}_n, \mathbf{b}_n, \mathbf{c}_n]^T$, in the Kalman filter equations. Here, the correction terms, $\mathbf{b}_n = [b_{n,1}, b_{n,2}, \dots, b_{n,N}]^T$, and $\mathbf{c}_n = [c_{n,1}, c_{n,2}, \dots, c_{n,N}]^T$, have the same dimension, N , which is typically equal to the number of grid point variables in a numerical weather prediction model. By inserting the augmented state into the Kalman filter equations, we assume that $\psi(\mathbf{v}_n^b)$, the background probability distribution of the augmented state, is Gaussian; that is,

$$\psi(\mathbf{v}_n^b) \sim \exp \left[-\frac{1}{2}(\mathbf{v}_n^b - \bar{\mathbf{v}}_n^b)^T (\mathbf{P}_v^b)^{-1} (\mathbf{v}_n^b - \bar{\mathbf{v}}_n^b) \right], \quad (4.22)$$

where $\bar{\mathbf{v}}_n^b$ is the background mean of the augmented state, and \mathbf{P}_v^b is the background error covariance matrix for the augmented state.

The main computational challenge in designing an augmented Kalman filter is to find a computationally efficient approach to estimate \mathbf{P}_v^b , whose dimension increases by N when a new parameter is added to the state. One frequently applied approach to reduce the computational burden associated with the estimation of \mathbf{P}_v^b is to assume that many entries of the matrix are zero, e.g., by assuming that the (non-augmented) state and the bias parameters are uncorrelated. We propose a different approach, which involves estimating the background mean and the background error covariance matrix by an ensemble, and solving the ensemble Kalman

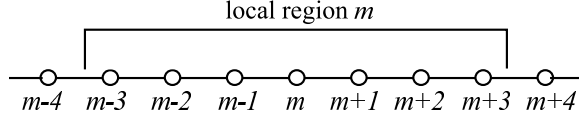


Figure 4.5: Illustration of a local state centered about location m .

filter equations locally in grid space applying the Local Ensemble Kalman Filter [70] to the augmented state. The LEKF scheme estimates ‘local states’ as illustrated in Fig. 4.5. In particular, considering the LEKF procedure without model bias correction [i.e., as in 70], for each point m on the spatial grid, we consider a neighborhood consisting of the $2l + 1$ points centered at m ; these points have locations $m - l, m - l + 1, \dots, m, \dots, m + l - 1, m + l$ (e.g., $l = 3$ in Fig. 4.5). At time $t = t_n$, the LEKF does data assimilation on local regions centered at each grid point using the local state,

$$\mathbf{x}_n(m) = \begin{bmatrix} x_{n,m-l} \\ x_{n,m-l+1} \\ \vdots \\ x_{n,m} \\ \vdots \\ x_{n,m+l-1} \\ x_{n,m+l} \end{bmatrix}. \quad (4.23)$$

The global analysis state (i.e., the analysis state at $t = t_n$ at each grid point over the entire grid) is then taken to be the state at the center of each local region (see Ott et al. [70] for further discussion).

In order to adapt the LEKF to correct for model bias, we augment each local state to include the bias estimate of the bias model employed. For example, for Bias Model III, we form an augmented local state, $\mathbf{v}_n(m) = [\mathbf{x}_n(m), \mathbf{b}_n(m), \mathbf{c}_n(m)]^T$, for the data assimilation at location m . Similarly, for Bias Model I, $\mathbf{v}_n(m) =$

$[\mathbf{x}_n(m), \mathbf{b}_n(m)]^T$, and, for Bias Model II, $\mathbf{v}_n(m) = [\mathbf{x}_n(m), \mathbf{c}_n(m)]^T$. Since the augmented local state is derived from the global state \mathbf{v}_n , it can be also assumed to have a Gaussian distribution,

$$\begin{aligned} \psi_m(\mathbf{v}_n^b(m)) &\sim \exp \left\{ -\frac{1}{2} [\mathbf{v}_n^b(m) - \bar{\mathbf{v}}_n^b(m)]^T \right. \\ &\quad \left. \times [\mathbf{P}_n^b(m)]^{-1} [\mathbf{v}_n^b(m) - \bar{\mathbf{v}}_n^b(m)] \right\}, \end{aligned} \quad (4.24)$$

where $\bar{\mathbf{v}}_n^b(m)$ is the background mean of the augmented local state, and $\mathbf{P}_n^b(m)$ is the background error covariance matrix for the augmented local state. In this way, the dimension of the space for data assimilation is reduced to $2(2l + 1)$ for Bias Model I or II and to $3(2l + 1)$ for Bias Model III. An important property of this scheme is that it allows for (and also requires) the estimation of cross-correlations between uncertainties in the state estimates and uncertainties in the estimation of the model bias terms.

4.3 Numerical Experiments

4.3.1 Experimental Setup

For testing our assimilation scheme, we consider the Lorenz-96 model [56] which is defined by the system of differential equations,

$$\frac{dx_i}{dt} = (x_{i+1} - x_{i-2})x_{i-1} - x_i + \Theta, \quad i = 1, \dots, N, \quad (4.25)$$

where $x_{-1} = x_{N-1}$, $x_0 = x_N$, $x_{N+1} = x_1$ and Θ is a constant. The variables form a cyclic chain and may be thought of as roughly analogous to the values of some unspecified scalar meteorological quantity at N equally spaced sites along a latitude

circle. For compactness of notation, we will also represent eq. (4.25) as

$$\frac{d\mathbf{x}}{dt} = \mathbf{L}(\mathbf{x}), \quad (4.26)$$

where $\mathbf{x} = [x_1, x_2, \dots, x_N]^T$. We solve eq. (4.25) with a fourth-order Runge-Kutta method using a time step of 0.05 dimensionless units for which the system is computationally stable. Lorenz and Emanuel [56] consider this time step as roughly corresponding to 6 hours of real atmospheric evolution. With $\Theta = 8.0$ and $N = 40$, Lorenz and Emanuel demonstrate that the system (4.25) results in a westward (i.e., in the direction of low index of locations) progression of individual maxima and minima and an eastward progression of the center of activity with a dominant wavenumber of 8. In addition, they also find that the system is chaotic with 13 positive Lyapunov exponents and a Lyapunov dimension of 27.1. Throughout our numerical experiments we use $\Theta = 8.0$ and $N = 40$.

In what follows we will assume that our forecast model dynamics is given by eq. (4.25) but that the true dynamics of the system whose state we are concerned with obeys dynamics that may differ from those of our forecast model. We will consider situations in which the true dynamics differ from the forecast model in three ways, which we refer to as Type A truth bias, Type B truth bias, and Type C truth bias. The dynamical behaviors of the true system in these three cases are as follows:

$$\frac{d\mathbf{x}}{dt} = \mathbf{L}(\mathbf{x}) + \boldsymbol{\beta} \quad (\text{Type A}), \quad (4.27)$$

$$\frac{d\mathbf{x}}{dt} = \mathbf{L}(\mathbf{x} + \boldsymbol{\zeta}) \quad (\text{Type B}), \quad (4.28)$$

$$\frac{d\mathbf{x}}{dt} = \mathbf{L}(\mathbf{x} + \boldsymbol{\zeta}) + \boldsymbol{\beta} \quad (\text{Type C}), \quad (4.29)$$

where $\boldsymbol{\beta} = [\beta_1, \beta_2, \dots, \beta_N]^T$ and $\boldsymbol{\zeta} = [\zeta_1, \zeta_2, \dots, \zeta_N]^T$. When the true dynamics is

described by the same equation [eq. (4.25)] as the forecast model, we say that the forecast model is ‘perfect’. Note that Bias Model I would be a natural choice if it were known that the deviation of the true dynamics from the model dynamics (4.26) was such that the true dynamics belonged to a family of systems of the form given by (4.27) (Type A truth bias). Similar statements apply with regard to the relation between Bias Model II and Type B truth bias and between Bias Model III and Type C truth bias.

With small values of β and ζ , we conjecture that the systems (4.27)–(4.29) exhibit behaviors similar to those of system (4.25). In our numerical experiments, the elements of β and ζ vary in space (i) and have the forms

$$\beta_i = A \sin\left(2\pi\frac{i-1}{N}\right), \quad (4.30)$$

$$\zeta_i = B \sin\left(2\pi\frac{i-1}{N}\right), \quad i = 1, \dots, N, \quad (4.31)$$

where A and B are scalar constants.

The true states are generated by integrating one of the three equations (4.27)–(4.29), while the forecast model states are generated by integrating eq. (4.26). The evolution operators, \mathbf{F}^t and \mathbf{F}^m , are the integrations of the above equations (4.26)–(4.29) from some time t to $t + \Delta t$ where $\Delta t = 0.05$ and the states are available at every discrete time $t_n = t_0 + n\Delta t$, where t_0 is the time at which an experiment begins and n is a positive integer.

We assume that the observations are available at every time t_n for $n \geq 0$ and the state variables themselves are directly observed. Thus the observation operator in eqs. (4.9) and (4.17) is the identity operator [i.e., $\mathbf{H}(\mathbf{x}) = \mathbf{x}$]. We also assume that the observational noise ϵ_n has zero expected value and is uncorrelated, white and Gaussian with variance σ^2 . Thus the local observation error covariance matrix is a diagonal matrix whose components are σ^2 . Correspondingly, we generate our

simulated ‘observations’ (4.9) by adding uncorrelated Gaussian random numbers with variance σ^2 to the true state variables x_i^t and form a local observation $\mathbf{y}_n(m) = [y_{n,m-l}, \dots, y_{n,m+l}]^T$. Throughout our numerical experiments, we take $\sigma^2 = 0.09$.

Data assimilations are done at every integration time t_n . The analysis error is defined as

$$\mathbf{e}_n^a = \bar{\mathbf{x}}_n^a - \mathbf{x}_n^t, \quad (4.32)$$

for Bias Model I, and

$$\mathbf{e}_n^a = \bar{\mathbf{x}}_n^a + \bar{\mathbf{c}}_n^a - \mathbf{x}_n^t, \quad (4.33)$$

for Bias Model II and Bias Model III, where $\bar{\mathbf{x}}_n^a$ is the ensemble mean of the analysis and where $\bar{\mathbf{c}}_n^a$ is the ensemble mean of the estimate of the Type II bias. We use the root-mean-square (rms) of the analysis error,

$$\text{rms}\{\mathbf{e}_n^a\} = \sqrt{\frac{1}{N} \sum_{i=1}^N (e_{n,i}^a)^2}, \quad (4.34)$$

to assess the quality of the analysis at a given time, and the time mean of the rms error over a long time interval T ,

$$\langle\langle \mathbf{e}^a \rangle\rangle = \frac{1}{T} \sum_{n=n_0+1}^{n_0+T} \text{rms}\{\mathbf{e}_n^a\}, \quad T \gg 1, \quad (4.35)$$

to measure the overall performance of the assimilation scheme. Here, n_0 is the time we allow for the analysis to converge to the true state.

To improve the analyses in our numerical experiments, we employ variance inflation,

$$\hat{\mathbf{P}}_n^a(m) \rightarrow \hat{\mathbf{P}}_n^a(m) + \frac{\mu\Lambda}{k} \mathbf{I}_k, \quad (4.36)$$

where $\hat{\mathbf{P}}_n^a(m)$ is the local analysis error covariance matrix defined in the ‘internal’ coordinate system [70] whose basis is the set of eigenvectors of the local background

error covariance matrix $\mathbf{P}_n^b(m)$, μ is an inflation coefficient, and $\Lambda = \text{Trace}\{\hat{\mathbf{P}}_n^a(m)\}$. This particular form of variance inflation was proposed in Ott et al. [70] where it is referred to as ‘enhanced variance inflation’. Enhanced variance inflation has the effect of enhancing the estimated probability of error in directions that formally show only very small error probability. [This modification of $\hat{\mathbf{P}}_n^a(m)$ also modifies the ensemble perturbations through the square root filter; see Ott et al. [70].] The general purpose of employing a variance inflation is to correct for the loss of variance in the ensemble due to nonlinearities and sampling errors. Most importantly, variance inflation can also stabilize the Kalman filter in the presence of model errors, as it was shown in Ott et al. [71] for the Lorenz-96 model and in Whitaker et al. [96] for the NCEP GFS model preparing a historical reanalysis data set. Since variance inflation schemes are computationally less expensive than the state augmentation method, we hope to see that the technique we propose here lead to larger improvements in the accuracy of the state estimates than what can be achieved by simply tuning the variance inflation coefficient μ .

For the dimension of local states used in the LEKF, we select 13 (i.e., $l = 6$) which is known to be a good choice for the Lorenz model (4.25) with $\Theta = 8$ and $N = 40$ [70]. Hence, the augmented local states have 26 dimensions for the states used in Bias Model I or Bias Model II, and 39 dimensions for the states used in Bias Model III. In our numerical experiments we choose the number of ensemble members to be the same as the dimension of the augmented local states so that the local background error covariance matrix has full rank. This choice means that the ensemble size is 13 when bias is not estimated in the assimilation, 26 when Bias Model I or II is used, and 39 when Bias Model III is used. Thus we take into account the added dimensionality of the augmented local states, anticipating that this increased dimensionality necessitates correspondingly increased ensemble size in order to properly represent it. This increased ensemble size is part of the added

computational cost that is paid in order to correct for model bias. In practice, for given computer resources, the need for a large ensemble may thus necessitate consideration of benefit trade-offs amongst ensemble size, local domain size, model resolution, etc.

Finally, for bias evolution (eqs. (4.8), (4.16), (4.19), and (4.20)) we use $\mathbf{G}^b(\mathbf{x}, \mathbf{b}, \mathbf{c}) = \mathbf{b}$ and $\mathbf{G}^c(\mathbf{x}, \mathbf{b}, \mathbf{c}) = \mathbf{c}$, until Sec. 4.3.6 where we consider different evolution.

4.3.2 Perfect Forecast Model

We first test our bias models for the case of a perfect forecast model, i.e., for the case when the true values of β and ζ in eqs. (4.27)–(4.29) are $\mathbf{0}$. In this cast, the evolution operators for the true state and the forecast model state are identical, $\mathbf{F}^t = \mathbf{F}^m$.

In order to generate the true states for this run, we first integrate eq. (4.25) for 10^4 time steps from a random initial condition, allowing the system to approach its attractor. After this, we perform data assimilation at every time step. The initial ensemble members for the first data assimilation are generated by adding independent, zero mean, normally distributed random numbers of variance 1.3 to the true state at every spatial point i . Before obtaining the rms time mean of the analysis error (4.35), we run 2×10^4 data assimilations to allow convergence. Past this time (denoted n_0), it is found that the rms analysis error reaches a statistically steady state in which it fluctuates about a temporally constant mean value which we denote $\langle\langle \mathbf{e}^a \rangle\rangle$.

The data plotted in Fig. 4.6 show the time averaged rms analysis error (4.35) as a function of the inflation coefficient μ (4.36). Here, the rms analysis error is averaged over $T = 3 \times 10^4$ time steps. For the case in which no state augmentation is employed in the assimilation (data plotted as * symbols) the best performance is obtained near $\mu = 0.005$ for which $\langle\langle \mathbf{e}^a \rangle\rangle = 0.057$. The other three curves in Fig. 4.6

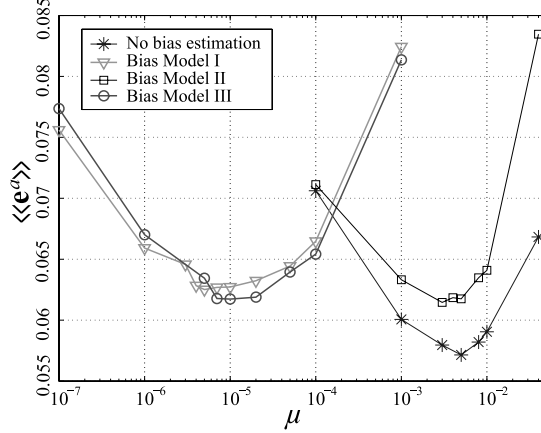


Figure 4.6: Time averaged rms analysis error, $\langle\langle \mathbf{e}^a \rangle\rangle$, versus variance inflation coefficient, μ , for the perfect forecast model experiment: With a perfect forecast model, any attempt to estimate and correct for a bias results in slightly higher analysis error.

show the analysis errors for the cases when the same observations are assimilated with using the three different bias models of Sec. 4.2.4 in the state estimation. In these cases, we try to estimate a bias that is zero in reality. The estimated bias terms tend to fluctuate about zero, resulting in a slight ($\sim 8\%$) increase of the error. The above results from examining this case provide a standard against which we can compare results that we will subsequently obtain for situations with error in the forecast model.

In Fig. 4.6, we see that the minima of $\langle\langle \mathbf{e}^a \rangle\rangle$ appear at a lower inflation coefficient when the states are augmented by an estimate of the bias, \mathbf{b}^b (i.e., for Bias Models I and III). In order to see why this occurs we consider the local perturbations for the augmented local background,

$$\delta \mathbf{v}_n^{b(j)}(m) = \mathbf{v}_n^{b(j)}(m) - \bar{\mathbf{v}}_n^b(m), \quad (4.37)$$

where $\{\mathbf{v}_n^{b(j)}(m)\}$ are the ensemble members of the augmented local background. These perturbations are used to estimate the local background error covariance

matrix [70]

$$\mathbf{P}_n^b(m) = \mathbf{V}_n^b(m) [\mathbf{V}_n^b(m)]^T, \quad (4.38)$$

where

$$\mathbf{V}_n^b(m) = k^{-\frac{1}{2}} [\delta \mathbf{v}_n^{b(1)}(m) | \delta \mathbf{v}_n^{b(2)}(m) | \dots | \delta \mathbf{v}_n^{b(k+1)}(m)] \quad (4.39)$$

and $k + 1$ is the number of the local ensemble members. We rewrite eq. (4.37) using eq. (4.7) as

$$\delta \mathbf{v}_n^{b(j)}(m) = \begin{bmatrix} \delta \mathbf{x}_n^{b(j)}(m) \\ \delta \mathbf{b}_n^{b(j)}(m) \end{bmatrix} = \begin{bmatrix} \delta \tilde{\mathbf{x}}_n^{b(j)}(m) + \delta \mathbf{b}_n^{b(j)}(m) \\ \delta \mathbf{b}_n^{b(j)}(m) \end{bmatrix}, \quad (4.40)$$

where $\{\delta \mathbf{x}_n^{b(j)}(m)\}$ are perturbations for the local background, $\tilde{\mathbf{x}}_n^b = \mathbf{F}^m(\mathbf{x}_{n-1}^a)$, and $\{\delta \mathbf{b}_n^{b(j)}(m)\}$ are perturbations for the local prediction for Bias Model I. In our experiments we observe that $\{\delta \mathbf{b}_n^{b(j)}(m)\}$ are only weakly correlated with each other and almost uncorrelated with $\{\delta \tilde{\mathbf{x}}_n^{b(j)}(m)\}$. Effectively, therefore, uncorrelated random vectors are added to the state perturbations $\{\delta \tilde{\mathbf{x}}_n^{b(j)}(m)\}$ in eq. (4.40). In consequence, $\mathbf{P}_n^b(m)$ (4.38) is effectively inflated mostly on the diagonal components by the amount of the variance of $\{\delta \mathbf{b}_n^{b(j)}(m)\}$. (This effective inflation created by using eq. (4.7) to obtain the background will also be observed when the forecast model is not perfect as shown in the following subsections.)

4.3.3 Data Assimilation with Type A Truth Bias

In this experiment, we perform data assimilation using the three augmented local states as described in Section 4.2.4 and an unaugmented state when the true state is evolved using eq. (4.27) with $A = 0.2\Theta = 1.6$ in eq. (4.30) corresponding to Type A truth bias. Again the forecast model state is evolved using eq. (4.25). We can approximate the bias \mathbf{b}_n^t given by eq. (4.3) as follows. Recall that \mathbf{F}^t and \mathbf{F}^m are the time Δt maps of the true dynamics (eq. (4.27)) and the forecast model

(eq. (4.26)) and that eq. (4.3) is based on the assumption that $\mathbf{x}^m(t_{n-1}) = \mathbf{x}^t(t_{n-1})$. Taking the difference between the true equation (4.27) and the model equation (4.26),

$$\frac{d}{dt}(\mathbf{x}^t - \mathbf{x}^m) = \mathbf{L}(\mathbf{x}^t) + \boldsymbol{\beta} - \mathbf{L}(\mathbf{x}^m) \approx \boldsymbol{\beta}, \quad (4.41)$$

then, integrating eq. (4.41) for the time interval $t_{n-1} \leq t \leq t_n$ with the initial condition $\mathbf{x}^t(t_{n-1}) = \mathbf{x}^m(t_{n-1}) = \mathbf{x}_{n-1}^t$, to obtain $\mathbf{b}_n^t = \mathbf{x}^t(t_n) - \mathbf{x}^m(t_n)$, yields

$$\mathbf{x}^t - \mathbf{x}^m \approx \int_{t_{n-1}}^{t_n} \boldsymbol{\beta} dt = \boldsymbol{\beta} \Delta t. \quad (4.42)$$

Using eq. (4.3) and (4.42) we obtain $\mathbf{b}_n^t \approx \boldsymbol{\beta} \Delta t$. For the situation in this section, $\Delta t = 0.05$, and we have taken $\boldsymbol{\beta}$ to be constant in time,

$$\beta_i = A \sin\left(2\pi \frac{i-1}{N}\right) = 1.6 \sin\left(2\pi \frac{i-1}{N}\right). \quad (4.43)$$

Time averaged rms analysis errors for each case are shown in Fig. 4.7. In the case where the bias is not estimated, the error is around 0.167 at $\mu \approx 0.7$, which is still lower than the rms error of the noisy observations. If we, however, augment the state using Bias Model I the error is reduced dramatically, and slightly more if we augment the state using Bias Model III, yielding $\langle\langle \mathbf{e}^a \rangle\rangle = 0.068$ and 0.061, respectively, at $\mu \approx 10^{-5}$ [Fig. 4.7(b)]. If, however, we augment the state using Bias Model II, then the rms error is around 0.225 at $\mu \approx 0.3$, worse than what is obtained when no bias estimation is employed. Here, we see again that using the unbiased background (4.7) for data assimilation effectively inflates the local background error covariance matrix, and a smaller variance inflation yields the lowest analysis error, $\langle\langle \mathbf{e}^a \rangle\rangle$. The good results obtained when the state is augmented using either Bias Model I or Bias Model III might reasonably be ascribed to the fact that estimation of \mathbf{b}_n^t can be regarded as correcting for precisely the form of truth bias that is present

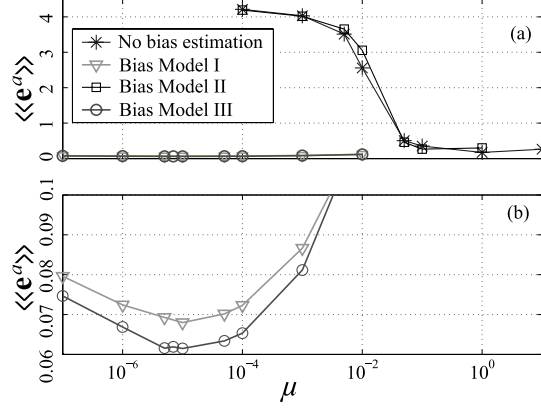


Figure 4.7: Time averaged rms analysis error, $\langle\langle \mathbf{e}^a \rangle\rangle$, versus μ for the case of Type A truth bias: Note that (b) shows the same results as (a) for Bias Model I and Bias Model III but for a different vertical scale.

when the truth evolves by eq. (4.27).

In Fig. 4.8, we plot $\langle \bar{\mathbf{b}}^a \rangle$, the time average of the ensemble mean of the bias estimate,

$$\langle \bar{\mathbf{b}}^a \rangle = \frac{1}{T} \sum_{n=n_0+1}^{n_0+T} \bar{\mathbf{b}}_n^a, \quad (4.44)$$

$$\bar{\mathbf{b}}_n^a = \frac{1}{k+1} \sum_{i=1}^{k+1} \mathbf{b}_n^{a(j)}, \quad (4.45)$$

where $T = 2000$ and $n_0 = 15000$, for the experiment with the state augmented using Bias Model I at $\mu = 10^{-5}$ (where the rms error is minimum in Fig. 4.7). We see that $\langle \bar{\mathbf{b}}^a \rangle$ agrees well with the approximation to \mathbf{b}_n^t given by (4.42) and (4.43) (shown as the solid curve). Also, although the shape is not shown here, for the case that the state is augmented using Bias Model III, $\langle \bar{\mathbf{b}}^a \rangle$ again agrees well with (4.42) and (4.43).

We now examine the analysis errors using Bias Model I and III. In Fig. 4.9, we plot the analysis error (4.32) averaged over 5000 time steps,

$$\langle \mathbf{e}^a \rangle = \frac{1}{T} \sum_{n=n_0+1}^{n_0+T} \mathbf{e}_n^a, \quad (4.46)$$

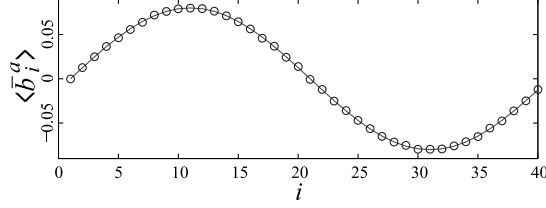


Figure 4.8: The average bias estimate of location i is shown as \circ . The approximate true bias $\beta_i \Delta t$ is shown as the solid curve.

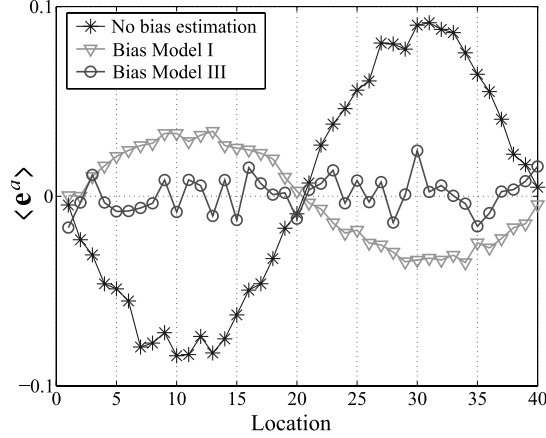


Figure 4.9: Time average of the analysis error as a function of location: The Type A truth bias is corrected best when we perform the assimilation using Bias Model III.

(here, $T = 5000$ and $n_0 = 15000$) for the case that no bias estimation is performed in data assimilation ($*$), the case that the estimation is performed using Bias Model I in the assimilation (∇), and the case that the bias estimation is performed using Bias Model III (\circ). The variance inflation coefficients are $\mu = 1.0, 10^{-5}$, and 10^{-5} , respectively, at which the errors, $\langle \langle e^a \rangle \rangle$, are minimum for each case.

In order to understand why Bias Model III does better than Bias Model I for this case, recall that both models determine the background state \mathbf{x}_n^b at time t_n by $\mathbf{x}_n^b = \mathbf{F}^m(\mathbf{x}_{n-1}^a) + \mathbf{b}_n^b$ where \mathbf{b}_n^b is approximately constant in time (eqs. (4.7) and (4.18)). However, Bias Model I tries to make \mathbf{x}_n^b close to \mathbf{x}_n^t , whereas Bias Model III tries to make $\mathbf{x}_n^b + \mathbf{c}_n^b$ close to \mathbf{x}_n^t , where \mathbf{c}_n^b is also approximately constant in time. Suppose that \mathbf{b}_n^b has converged to the time-independent approximation $\beta \Delta t$ of \mathbf{b}_n^t (see eq. (4.42)), that \mathbf{c}_n^b has converged to a constant vector \mathbf{c} , and that the

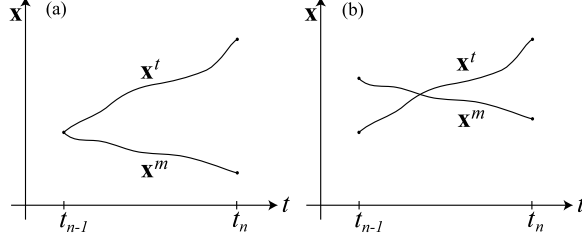


Figure 4.10: Figure (a) depicts the case $\mathbf{c} = \mathbf{0}$, corresponding Bias Model I, while figure (b) depicts the case $\mathbf{c} \approx -\beta/2$ in Bias Model III.

analysis at time t_{n-1} is perfect: $\mathbf{x}_{n-1}^a + \mathbf{c} = \mathbf{x}_{n-1}^t$. Consider the model trajectory $\mathbf{x}^m(t)$ of eq. (4.26) with $\mathbf{x}^m(t_{n-1}) = \mathbf{x}_{n-1}^a$ and the true trajectory $\mathbf{x}^t(t)$ of eq. (4.27) with $\mathbf{x}^t(t_{n-1}) = \mathbf{x}_{n-1}^t$. Then $\mathbf{x}_n^t = \mathbf{x}^t(t_n)$ and $\mathbf{x}_n^b = \mathbf{x}^m(t_n) + \beta\Delta t$, and the background is most accurate if $\mathbf{x}_n^b + \mathbf{c} = \mathbf{x}_n^t$. As in eq. (4.41),

$$\begin{aligned}
\mathbf{x}_n^t - (\mathbf{x}_n^b + \mathbf{c}) &= \mathbf{x}^t(t_n) - (\mathbf{x}^m(t_n) + \beta\Delta t + \mathbf{c}) \\
&= \mathbf{x}^t(t_{n-1}) - (\mathbf{x}^m(t_{n-1}) + \beta\Delta t + \mathbf{c}) \\
&\quad + \int_{t_{n-1}}^{t_n} [\mathbf{L}(\mathbf{x}^t(t)) + \beta - \mathbf{L}(\mathbf{x}^m(t))] dt \\
&= \mathbf{x}_{n-1}^t - (\mathbf{x}_{n-1}^a + \beta\Delta t + \mathbf{c}) \\
&\quad + \int_{t_{n-1}}^{t_n} [\mathbf{L}(\mathbf{x}^t(t)) - \mathbf{L}(\mathbf{x}^m(t))] dt + \beta\Delta t \\
&= \int_{t_{n-1}}^{t_n} [\mathbf{L}(\mathbf{x}^t(t)) - \mathbf{L}(\mathbf{x}^m(t))] dt.
\end{aligned} \tag{4.47}$$

Thus, we desire that the average value of $\mathbf{L}(\mathbf{x}^t(t)) - \mathbf{L}(\mathbf{x}^m(t))$ to be as small as possible over the interval $t_{n-1} \leq t \leq t_n$. Using $\mathbf{c} = -\beta\Delta t/2$ makes this average zero to first order, and is thus superior to using $\mathbf{c} = \mathbf{0}$, which corresponds to Bias Model I. (see Fig. 4.10.) This is confirmed in Fig. 4.11, from which one observes $\langle \bar{\mathbf{c}}^a \rangle \approx -\beta\Delta t/2$ in the experiment with Bias Model III.

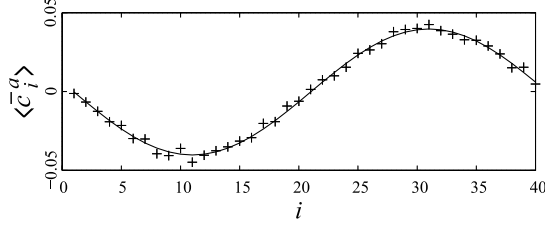


Figure 4.11: The average bias estimate of location i is shown as $+$. The the value of $-\beta_i \Delta t / 2$ is shown as the solid curve.

4.3.4 Data Assimilation with Type B Truth Bias

In this experiment, we simulate a bias by evolving the true state with eq. (4.28) with temporally constant ζ and estimate it with three different augmentation methods as done in Sec. 4.3.3. To obtain the true value of \mathbf{c}^t , defined by (4.10), we use

$$\begin{aligned}
 \frac{d\mathbf{c}^t}{dt} &= \frac{d\mathbf{x}^t}{dt} - \frac{d\mathbf{x}^m}{dt} \\
 &= \mathbf{L}(\mathbf{x}^t + \zeta) - \mathbf{L}(\mathbf{x}^m) \\
 &= \mathbf{L}(\mathbf{x}^t + \zeta) - \mathbf{L}(\mathbf{x}^t - \mathbf{c}^t).
 \end{aligned} \tag{4.48}$$

A trivial solution to eq. (4.48) is $\mathbf{c}^t = -\zeta$ whose i -th element is given by

$$c_{n,i}^t = -B \sin\left(2\pi \frac{i-1}{N}\right) = -1.6 \sin\left(2\pi \frac{i-1}{N}\right), \tag{4.49}$$

where we take $B = 0.2\Theta = 1.6$ in this experiment.

The time averaged rms analysis errors for each case are shown in Fig. 4.12. We see that, when we augment the state using Bias Model II in the assimilation, we can correct for the bias. The minimum rms error for this assimilation is about 0.061 and occurs near $\mu = 0.003$. The minimum rms error for the assimilation with the augmented state using the Bias Model III is about 0.062 and occurs at a lower μ value (as expected) of about $\mu = 2 \times 10^{-5}$. Without bias correction, the minimum rms error is 0.263 and occurs near $\mu = 4.0$, similar to what is obtained using Bias

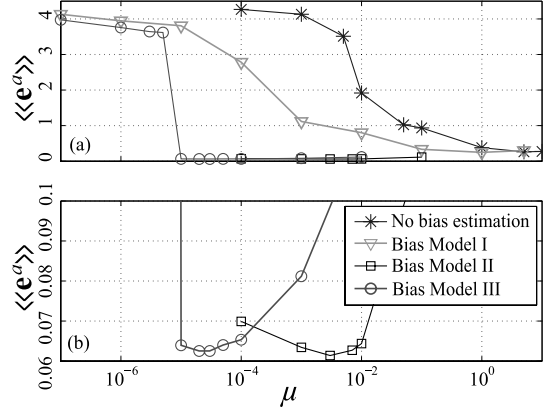


Figure 4.12: Time averaged rms analysis error, $\langle\langle \mathbf{e}^a \rangle\rangle$, versus μ for the case of Type B truth bias. (b) has a different vertical scale from (a) for transparency.

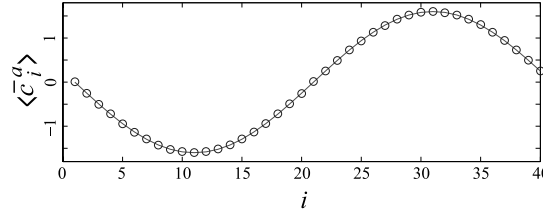


Figure 4.13: Bias estimate for the Bias Model II is shown as \circ . The true bias (4.49) is shown as the solid curve.

Model I ($\langle\langle \mathbf{e}^a \rangle\rangle \approx 0.234$ near $\mu = 0.5$). Also, due to the variance inflation effect of Bias Model I, the best assimilation result with the augmented state using the Bias Model III occurs at a lower value of variance inflation than the assimilation with the augmented state using Bias Model II. For the same reason, the rms error, $\langle\langle \mathbf{e}^a \rangle\rangle$, for the case of assimilation with the augmented state using Bias Model I has lower values in the region of variance inflation, $10^{-5} \leq \mu \leq 1.0$, than for the case of assimilation with no state augmentation.

We show $\langle \bar{\mathbf{c}}^a \rangle$ in Fig. 4.13 for the case in which the state is augmented using Bias Model II at $\mu = 0.003$. It is seen that the result (plotted as \circ) agrees very well with eq. (4.49) (plotted as the solid line). We obtain the same result for $\langle \bar{\mathbf{c}}^a \rangle$ when the state is augmented using Bias Model III at $\mu = 2 \times 10^{-5}$.

4.3.5 Data Assimilation with Type C Truth Bias

Now, we combine the two biases in the truth [see eq. (4.29)], and estimate them with three augmented states as done in previous sections. In this case, we can regard the Type A truth bias as added to a system that already has Type B truth bias. Hence, a differential equation for \mathbf{b}^t can be written as

$$\frac{d\mathbf{b}^t}{dt} = \mathbf{L}(\mathbf{x}^t + \boldsymbol{\zeta}) + \boldsymbol{\beta} - \mathbf{L}(\mathbf{x}^t - \mathbf{c}^t), \quad (4.50)$$

where $\boldsymbol{\zeta}$ and $\boldsymbol{\beta}$ are constant in time in the present experiment. A solution to eq. (4.50) is

$$\frac{d\mathbf{b}^t}{dt} = \boldsymbol{\beta} \quad \text{if} \quad \mathbf{c}^t = -\boldsymbol{\zeta}; \quad (4.51)$$

that is, the individual true bias is the same as if the system has only one bias. The quantities $\boldsymbol{\beta}$, $\boldsymbol{\zeta}$, are thus again given by eq. (4.43) and (4.49).

Figure 4.14 shows the resulting time averaged rms analysis error, $\langle\langle \mathbf{e}^a \rangle\rangle$, for each estimation method. The best result is obtained using Bias Model III. This might be anticipated since augmentation by one bias estimate alone cannot satisfy the solution (4.51). The minimum rms error is around 0.062 and occurs near 2×10^{-5} ; this is the same as in the previous experiments. The other assimilation methods yields $\langle\langle \mathbf{e}^a \rangle\rangle \approx 0.261$ at $\mu \approx 4.0$ without state augmentation, $\langle\langle \mathbf{e}^a \rangle\rangle \approx 0.236$ at $\mu \approx 0.5$ using Bias Model I, and $\langle\langle \mathbf{e}^a \rangle\rangle \approx 0.224$ at $\mu \approx 0.3$ using Bias Model II.

4.3.6 Settling Time

Even though the state augmented LEKF can correct for various biases in the true state, for the truth biases considered here it requires longer settling time for the forecast model state to converge toward the true state as compared to the LEKF without state augmentation. We regard the time it takes for the rms analysis

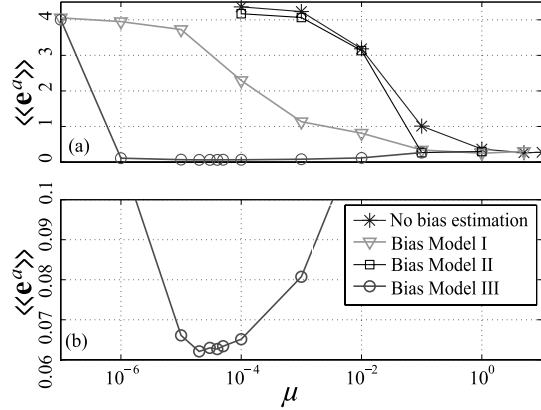


Figure 4.14: Time averaged rms analysis error, $\langle\langle \mathbf{e}^a \rangle\rangle$, versus μ for the case in which the truth has Type C truth bias: In order to correct for the biases, the augmented state used in the assimilations must contain both the \mathbf{b} and \mathbf{c} bias estimates, Here, we again use a different vertical scale in (b) from (a) for transparency.

error (4.34) to settle near its time averaged rms analysis error (4.35) as the settling time. With a perfect forecast model as in Sec. 4.3.2, the settling time is around 50 time steps with the $k + 1 = 13$ ensemble members we are using for the regular (i.e., without state augmentation) LEKF assimilation scheme. The settling time, however, becomes between 100 and 200 time steps when either Bias Model I or Bias Model II are used in the assimilation to respectively correct for Type A or B truth bias. The longest settling time, which is near 15000 time steps, appears when assimilations are done using Bias Model III to correct for Type B and Type C truth biases (when using Bias Model III to correct for Type A truth bias, we found settling times that were generally below 500).

However, it turns out that we can easily correct this problem by using a priori information on the bias which could be obtained by looking at the innovation, the difference between the forecast and the observation,

$$\bar{\mathbf{d}}_n = \mathbf{H}(\bar{\mathbf{x}}_n^b) - \mathbf{y}_n, \quad (4.52)$$

for the case when no bias estimation is performed. We plot the time average of

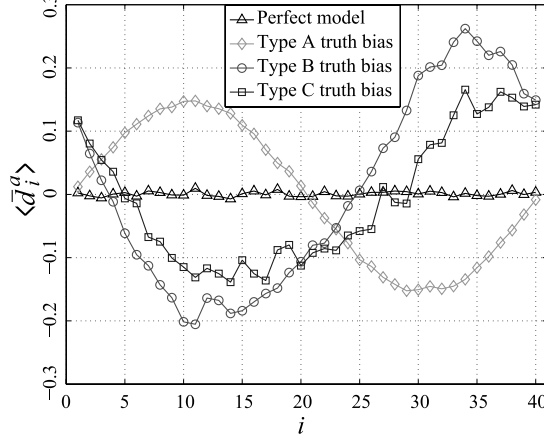


Figure 4.15: Time averaged innovation for the case that no bias estimation is performed with various biases in the truth.

the innovation in Fig. 4.15 for the case that no bias estimations are performed even though Type A, B or C truth bias is present in the truth. Since we take the observations to be unbiased, we can think of the time averaged innovation as the forecast bias. We see that these averages are large and vary slowly in space.

In the previous experiments, the initial ensemble variance for the bias estimate is 0.1. By increasing the initial variance to 1.0 (somewhat larger than the spread in the time averaged innovations), we can dramatically decrease the longest settling time from 15000 time steps to 800 time steps, while for the case that the settling time is already small (between 100 and 200 time steps) no significant change in the settling time is observed. We can decrease the settling time further if we exploit the fact that the biases vary slowly in space. To incorporate this added knowledge into our data assimilation scheme we now use a diffusion process for the time evolution, eqs. (4.8) and (4.16), of the biases,

$$b_{n+1,i}^b = (1 - 2\alpha_b)b_{n,i}^a + \alpha_b b_{n,i-1}^a + \alpha_b b_{n,i+1}^a, \quad (4.53)$$

$$c_{n+1,i}^b = (1 - 2\alpha_c)c_{n,i}^a + \alpha_c c_{n,i-1}^a + \alpha_c c_{n,i+1}^a, \quad (4.54)$$

where α_b and α_c are diffusion coefficients. By introducing diffusion in this way,

rapid spatial variation of the bias estimates is damped, leading to smooth spatial variation consistent with the actual case, eqs. (4.30) and (4.31), and the evidence of Fig. 4.15.

Our experiments show that there is a modest improvement in the settling time in the case that there is one type of bias in the truth (truth bias A or B) and that it is corrected using the corresponding bias model (Bias Model I or II, respectively); the settling time is decreased from between 100 and 200 time steps to between 80 and 130 time steps when α_b (or α_c) is increased from zero to $\alpha_b = 0.01$ (or $\alpha_c = 0.01$). When we consider the case of Type C truth bias and augment the state using Bias Model III, nonzero α_b diffusion (with $\alpha_c = 0$) can achieve a large decrease in the settling time, from 800 time steps to 300 time steps.

All of the decreases in settling time we have described come without significant increase in the time averaged rms analysis error $\langle\langle \mathbf{e}^a \rangle\rangle$. We find that using diffusive evolution on bias estimates actually decreases $\langle\langle \mathbf{e}^a \rangle\rangle$ in some cases. Figure 4.16 shows the time asymptotic analysis errors as a function of α_b for the case of Type C truth bias and Bias Model III assimilation. We see that a small amount of diffusion, in addition to shortening the settling time, also improves the time asymptotic performance of the assimilation. This improvement is not seen in the experiments with Bias Model I and II.

Finally, we note that if diffusion (eqs. (4.53) and (4.54)) is added in the perfect forecast model case (Sec. 4.3.2), all three cases of augmentation have the same values of rms analysis error as that of the unaugmented case. That is, curves corresponding to the four cases in Fig. 4.6 have the same minimum values with appropriate amounts of diffusion. Evidently diffused evolution of the bias estimate allows the estimate to converge to the truth faster, and also reduces the rms analysis error of the state augmented estimates.

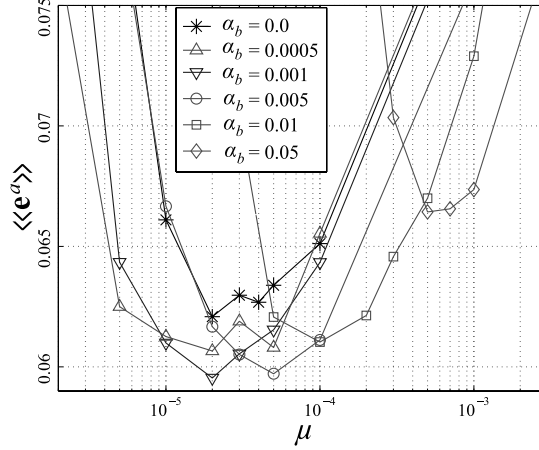


Figure 4.16: Time averaged rms analysis error, $\langle\langle \mathbf{e}^a \rangle\rangle$, versus μ with various α_b and $\alpha_c = 0$ for the case of Type C truth bias corrected using Bias Model III: Small diffusion improves the performance of the assimilation up to $\alpha_b = 0.01$.

4.3.7 A Simple State Dependent Model Error

In this section, we introduce a simple model error, γx_i^2 , which is proportional to the square of the state variable. That is, the dynamic equation for the true system is as follows:

$$\frac{dx_i}{dt} = (x_{i+1} - x_{i-2})x_{i-1} - x_i - \gamma x_i^2 + \Theta, \quad i = 1, \dots, N, \quad (4.55)$$

where $\gamma = 0.05$ and $\Theta = 10.0$. Here, Θ is increased to maintain chaotic behavior of the true dynamics (introducing γx_i^2 without changing Θ from its previous value of $\Theta = 8.0$ results in dominance of time periodic behavior of the true dynamics). For the forecast model, we use eq. (4.25) with $\Theta = 10.0$. Through the numerical experiment, we obtain the time average $\langle -\gamma x_i^2 \rangle \approx -0.95$ at each point i while $-\gamma x_i^2$ itself has large temporal fluctuations ranging from 0 to around -10 . The task we undertake here is to successfully estimate the time mean effect of the model bias, $\langle -\gamma x_i^2 \rangle$, with our bias estimation schemes. Using analysis similar to that in

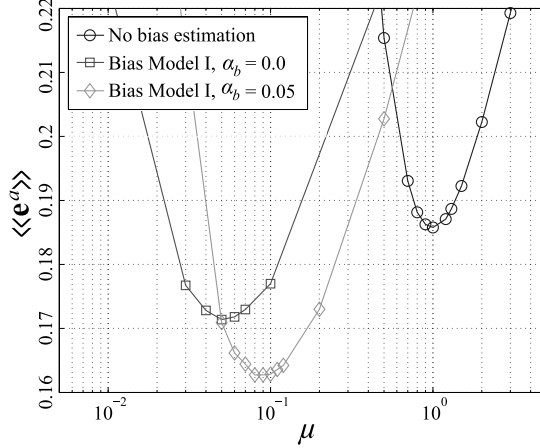


Figure 4.17: Time averaged rms analysis error, $\langle\langle \mathbf{e}^a \rangle\rangle$, versus μ with the simple state dependent model error: With Bias Model I and diffusion process, we can improve the performance in terms of the rms analysis error.

eqs. (4.41) and (4.42), we obtain

$$\langle b_i^t \rangle \approx \left\langle -\gamma \int_{t_{n-1}}^{t_n} (x_i^t)^2 dt \right\rangle \approx -\gamma \langle (x_i^t)^2 \rangle \Delta t, \quad (4.56)$$

and hence $\langle b_i^t \rangle \approx -0.048$ for each location i .

In Fig. 4.17, we plot the rms analysis errors of the numerical experiments. Without bias estimation the minimum rms analysis error $\langle\langle \mathbf{e}^a \rangle\rangle \approx 0.186$ is obtained at $\mu \approx 1.0$. Among the three bias models, Bias Model I produces the best result, $\langle\langle \mathbf{e}^a \rangle\rangle \approx 0.171$ at $\mu \approx 0.05$. In terms of the difference with the error in the perfect model case ($\langle\langle \mathbf{e}^a \rangle\rangle \approx 0.057$), this represents a 12% improvement toward the perfect model performance. If we employ diffusive evolution (with $\alpha_b = 0.05$), we can further decrease the rms error to obtain $\langle\langle \mathbf{e}^a \rangle\rangle \approx 0.163$ at $\mu \approx 0.09$ and 18% improvement toward the perfect model performance. In both cases (with and without diffusion), we obtain $\langle \bar{b}_i^a \rangle \approx -0.05$, which is a good estimate of $\langle b_i^t \rangle$, with smaller spatial variations when we use diffusion.

We also find that the performance of the assimilation with Bias Model I is not sensitive to the selection of Θ in the model equation while the performance of

the assimilation without bias estimation is sensitive to the selection of Θ . We can also achieve the same performance with Bias Model III with appropriate diffusion ($\alpha_b = 0.05$, $\alpha_c = 0.2$), but we cannot achieve it with Bias Model II. We conjecture that the reason is because the form of the bias in eq. (4.55) is closer to Type A truth bias than Type B.

4.4 Conclusions and Discussion

In this paper, we considered three bias models for use in state space augmentation strategies to mitigate the effects of model biases on forecasts:

- Bias Model I is based on the assumption that the best background information is obtained when the initial condition of the short term forecast that provides the background (the analysis at the previous assimilation time) is as close to the truth as possible.
- Bias Model II is based on the assumption that there exists a transformation from orbits on the attractor of the forecast model to orbits on the attractor of the true system.
- Bias Model III combines Bias Model I and Bias Model II.

While Bias Model I was considered by others in earlier papers for schemes other than the LEKF, Bias Model II and Bias Model III are (to the best of our knowledge) first introduced in the present paper.

To evaluate the performance of the proposed bias models for use in augmented ensemble Kalman filtering, we carried out experiments with the Lorenz-96 model. While we used the original model equation of Lorenz and Emanuel [56] to evolve the model state, we employed altered versions of Lorenz and Emanuel's equation to generate sets of time series of the true states. Each alteration of the equation

corresponded to distinctly different types of model biases. The main results of these numerical experiments are the following:

- The effectiveness of the different bias models strongly depends on the actual form of the true model bias. In our numerical experiments it was found that when the bias model was suited to the bias of the forecast model in modeling the true dynamics, then good results were obtained. However, when this was not the case, the results were not improved by the model bias correction scheme. This suggests that serious consideration of the choice of the bias model may be crucial in obtaining a successful scheme for model bias mitigation.
- For the bias models we considered, Bias Model III performed as well or better than the other bias models in terms of average analysis error, at the expense of requiring a larger ensemble and in some cases increasing the settling time. In most cases, the inclusion of parameters that were not present in the model bias did not yield improved performance. However, in the case of Type A truth bias, Bias Model III did outperform Bias Model I, due to the fact that the model bias was added to a continuous time forecast model, while the bias correction was applied at discrete times (see Sec. 4.3.3).
- We found (Sec. 4.3.6) that the model bias correction scheme took many more iteration steps to converge than in the case in which no model biases are present. The settling time strongly depends on the actual model bias, and on the bias model employed. As a result of the possibility of long settling times, one might anticipate that use of these model correction schemes may become problematic. We can, however, dramatically reduce the settling time by increase of the initial ensemble spread of the bias estimates. Moreover, when the model bias is slowly varying in space, we demonstrated that choosing a diffusive evolution of the model bias can also reduce the settling time.

- In a case with state-dependent, and thus time-varying, additive model error (Sec. 4.3.7), Bias Model I estimated a model bias that was close to the time average of the model error. In this sense it found the best estimate of the model error within our constant-in-time parameterization. The improvement in performance compared to no bias estimation was modest but significant.

Finally we note, that state space augmentation is not the only way to account for the effect of model errors in the state estimation process. As we mentioned earlier, variance inflation (both additive and multiplicative) can improve the resilience of Kalman filter schemes to the effects of model errors. Promising results were achieved by employing additive variance inflation schemes [46, 39] and by using hybrid Ensemble/3DVAR schemes [29] first proposed by Hamill and Snyder [38]. Nevertheless, we believe that optimal parameterization of model errors in the data assimilation process is a very promising means of mitigating model error and that much remains to be done along these lines. We view our paper as a contribution only to the initial phase of the quest for efficient model error parameterization algorithms.

Chapter 5

Correction of surface pressure model bias with an ensemble-based Kalman filter

5.1 Introduction

The dynamics of a numerical weather prediction model and the dynamics of the real atmosphere are different. In general, we view errors as fluctuating in space and time, and we think of them as the sum of a component that is smoothly varying on the space-time discretization scale of the model and a component that has the characteristics of random noise. In this chapter, our focus is on the slowly varying, component, which we call model bias.

The contribution of model bias to the discrepancy between a background forecast and the true atmospheric state can be comparable, or even larger, than the contribution of the initial condition errors and the chaotic model dynamics. Model errors have many sources, such as the finite resolution discretization of the continuous atmospheric fields, limited knowledge and imperfect representation of the sub-grid physical processes and imperfect specification of the boundary conditions. Although some of these sources are completely independent, it is not feasible to identify and parameterize each of them independently. One way to account for model bias, first suggested by Derber [26], is to assume that the total effect of all sources of the bias in the forecast model can be represented in the form of a limited number of bulk error terms added to the model forecast. Here, the amplitude of the bulk error terms is specified by parameters, which evolve smoothly with respect to the model discretization. Then, the approach is to estimate these parameters as

part of the data assimilation process.

The general problem of model bias estimation in Kalman filtering was first studied by Friedland [33], who suggested a scheme in which the state was augmented by a bias component. In Friedland's work the dynamics was taken to be linear and the bias was decoupled and estimated separately from the model state. This approach was introduced to the data assimilation literature by Dee and Da Silva [24], who achieved separation of the estimates of the state and the bias by assuming that the uncertainties in the two estimates were uncorrelated. Since then, this approach has been applied with some success to the assimilation of atmospheric observations [25, 15, 61, 10, 55, 17, 51, 27] More recently, Baek et al. [8] and Zupanski and Zupanski [104] suggested incorporating the method of state augmentation into the formulation of the ensemble-based Kalman filter data assimilations schemes.

In Baek et al. [8], we have also shown that the traditional approach to bias correction (henceforth referred to as Bias Model I), in which the background is first corrected with the estimated bias from the previous analysis cycle and then the state is estimated by updating the bias corrected background based on the latest observations, is often inefficient in improving the accuracy of the state estimate. To address this problem, we also proposed a new approach to incorporate the effects of model bias into the data assimilation process in Baek et al. [8]. This approach, called Bias Model II, is motivated by envisioning a situation in which the forecast model evolution takes place on an attractor shifted from the attractor of the true dynamics. In such a situation, making a correction to the background states that moves the background state estimate from the model attractor to the the true system attractor, as done in Bias Model I, may trigger an adjustment process during the next model integration step. The effects of such an adjustment on the accuracy of the state estimate are unpredictable and often negative. To avoid triggering a strong adjustment process , in Bias Model II we search for a state estimate that

best represents the true state on the the model attractor. This involves finding the parameters of a transformation between the model attractor and the true attractor. Bias Model I and II are defined more precisely in Section 2. [We note that other approaches based on assuming a mismatch between the model attractor and the true attractor were also proposed recently by Drécourt et al. [27] and Toth and Peña [90]].

In this chapter, we investigate the potential benefits of correcting the bias in the surface pressure state variable with Bias Model II. To simulate the situation faced in numerical weather prediction, we use two forecast models at different resolutions and with different level of sophistication in the physical parameterization packages to generate the “true” atmospheric state and to analyze and forecast the “true” states. In particular, the true atmospheric states are generated by the 2004 version of Global Forecast System (GFS) model of the National Centers for Environmental Prediction (NCEP) at resolution T62L28 (the global atmosphere is represented by $144 \times 73 \times 28$ grid points), while the analyses and forecasts are obtained with the Simplified Parameterization Primitive-Equation Dynamics (SPEEDY) model at resolution T30L7 (the global atmosphere is represented by $96 \times 48 \times 7$ grid points). The physical parameterization schemes and the surface boundary condition in the SPEEDY model are strongly simplified compared to those in the NCEP GFS. Since the vertical coordinate of the SPEEDY model is sigma, defined by the ratio between pressure and the surface pressure, correcting the bias in the surface pressure affects the assimilation of all other variables. Also, because the systematic difference between the surface pressure in the SPEEDY and NCEP GFS models is mainly due to the difference between the orographies of the two models, it is especially appropriate to use Bias Model II: it would not make sense to bias correct the background surface pressure in the SPEEDY model to match the surface pressure of the NCEP GFS model, which is associated with a much higher resolution orography.

For data assimilation we use the latest, computationally most efficient, version of the the scheme first proposed in Ott et al. [70]. This version was formulated by Hunt et al. [47] and is called the Local Ensemble Transform Kalman Filter (LETKF). The most important feature of this scheme is that it obtains the analysis independently for each grid point, assimilating all observations influencing the analysis at a given grid point simultaneously. This allows for a computationally efficient parallel implementation. Although our experiments employ the LETKF, we believe that the results we obtain regarding the quality of the bias estimates would hold for any other ensemble-based data assimilation schemes [e.g., for those proposed by 14, 45, 44, 2, 11, 40, 95, 105].

We present the bias model and the data assimilation scheme in Section 2. Section 3 presents the results of our numerical experiments, while conclusions and discussion are offered in Section 4.

5.2 Data assimilation with bias correction

5.2.1 Bias modeling

In the numerical weather prediction, the real atmospheric states are in infinite dimensional space and forecast model states that are used to predict the atmospheric states are in finite dimensional space (or a model space). We denote a finite dimensional representation (or a model space representation) of the infinite dimensional atmospheric state at time t_n by \mathbf{x}_n^t and call it ‘true state’. Then, the true atmospheric evolution in the model space can be written by

$$\mathbf{x}_{n+1}^t = \mathbf{f}^t(\mathbf{x}_n^t), \quad (5.1)$$

where \mathbf{x}_{n+1}^t is the true state at time $t_{n+1} = t_n + \Delta t$ and \mathbf{f}^t is the evolution operator for the true state. In general, one wants to devise a forecast model that mimics the model space representation (5.1) in order to predict the true atmospheric state into the future. Denoting the forecast model state at time t_n by \mathbf{x}_n^m , the forecast model evolution can be written by

$$\mathbf{x}_{n+1}^m = \mathbf{f}^m(\mathbf{x}_n^m), \quad (5.2)$$

where \mathbf{f}^m is the forecast model evolution operator.

Traditionally, it is assumed that the best estimate of the current true state can generate the best forecast for the true state. In Bias Model I, we follow this approach. This is similar to assuming that the evolution of the true state (5.1) and the evolution of the forecast model state (5.2) are the same so that the forecast model state stays near the true state after the evolution. In practice, however, the two evolution operators differ from each other. As a result, even if the forecast model state is evolved from an initial condition corresponding to the true state at the initial time (e.g. $\mathbf{x}_{n-1}^m = \mathbf{x}_{n-1}^t$), it is likely that the forecast model state departs from the true state as it evolves. Denoting this departure at time t_n by \mathbf{b}_n^t , or

$$\mathbf{b}_n^t = \mathbf{f}^m(\mathbf{x}_{n-1}^t) - \mathbf{f}^t(\mathbf{x}_{n-1}^t), \quad (5.3)$$

we can write the evolution of the true state (5.1) in terms of the forecast model evolution as follows:

$$\mathbf{x}_n^t = \mathbf{f}^m(\mathbf{x}_{n-1}^t) - \mathbf{b}_n^t, \quad (5.4)$$

$$\mathbf{b}_n^t = \mathbf{g}^b(\mathbf{b}_{n-1}^t, \mathbf{x}_{n-1}^t), \quad (5.5)$$

where \mathbf{g}^b is the evolution operator for the bias correction term. Then given an estimate (or ‘analysis’) of the augmented state $(\mathbf{x}_{n-1}^a, \mathbf{b}_{n-1}^a)$ at time t_{n-1} , we take

the forecast (or ‘background’) of the augmented state $(\mathbf{x}_n^b, \mathbf{b}_n^b)$ at time t_n to be

$$\mathbf{x}_n^b = \mathbf{f}^m(\mathbf{x}_{n-1}^a) - \mathbf{b}_n^b, \quad (5.6)$$

$$\mathbf{b}_n^b = \mathbf{g}^b(\mathbf{b}_{n-1}^a, \mathbf{x}_{n-1}^a). \quad (5.7)$$

A consequence of the discrepancies between the true evolution and the forecast model evolution is that the true system and the forecast model system may have different attractors to each other. In Bias Model II, we try to keep the actual forecast model state within or near its own attractor and to find a transformation between the two attractors rather than try to move the model state to the true state at the assimilation time. For simplicity, we assume that the transformation is just a shift of the forecast model state to the true state. Denoting the shift at time t_n by \mathbf{c}_n^t , or

$$\mathbf{c}_n^t = \mathbf{f}^m(\mathbf{x}_{n-1}^m) - \mathbf{f}^t(\mathbf{x}_{n-1}^t) = \mathbf{f}^m(\mathbf{x}_{n-1}^t + \mathbf{c}_{n-1}^t) - \mathbf{f}^t(\mathbf{x}_{n-1}^t), \quad (5.8)$$

we can rewrite the true evolution (5.1) as

$$\mathbf{x}_n^t = \mathbf{f}^m(\mathbf{x}_{n-1}^t + \mathbf{c}_{n-1}^t) - \mathbf{c}_n^t, \quad (5.9)$$

$$\mathbf{c}_n^t = \mathbf{g}^c(\mathbf{c}_{n-1}^t, \mathbf{x}_{n-1}^t + \mathbf{c}_{n-1}^t), \quad (5.10)$$

where \mathbf{g}^c is the evolution operator of the transformation. Since our goal in this model is that the analysis state \mathbf{x}_n^a closely approximates the model state $\mathbf{x}_n^m = \mathbf{x}_n^t + \mathbf{c}_n^t$ from which we can make future forecasts, we rewrite eqs. (5.9) and (5.10) as

$$\mathbf{x}_n^m = \mathbf{f}^m(\mathbf{x}_{n-1}^m), \quad (5.11)$$

$$\mathbf{c}_n^t = \mathbf{g}^c(\mathbf{c}_{n-1}^t, \mathbf{x}_{n-1}^m). \quad (5.12)$$

Then we can write the background state $(\mathbf{x}_n^b, \mathbf{b}_n^b)$ in terms of the previous analysis

state $(\mathbf{x}_{n-1}^a, \mathbf{c}_{n-1}^a)$ as follows:

$$\mathbf{x}_n^b = \mathbf{f}^m(\mathbf{x}_{n-1}^a), \quad (5.13)$$

$$\mathbf{c}_n^b = \mathbf{g}^c(\mathbf{c}_{n-1}^a, \mathbf{x}_{n-1}^a). \quad (5.14)$$

We note that, since \mathbf{x}_n^a and \mathbf{x}_n^b are the analysis and the forecast of \mathbf{x}_n^m , the analysis and the background for the true state \mathbf{x}_n^t are $\mathbf{x}_n^a + \mathbf{c}_n^a$ and $\mathbf{x}_n^b + \mathbf{c}_n^b$.

5.2.2 Bias correction with LETKF

Once a suitable model for the bias is chosen, it can be incorporated into the formulation of the local ensemble transform Kalman filter [LETKF, 70, 47] using state space augmentation [33, 24]. The LETKF scheme estimates local states in a local region. As illustrated in Fig. 5.1 for instance, the states at grid point $\mathbf{r} = (\text{longitude}, \text{latitude}, \text{height})$ can be estimated using the states and the observations within the local region which is colored grey and contains 9 grid points. To perform data assimilation for this local region at time t_n , we establish a local state vector, $\mathbf{x}_{\mathbf{r}}$ (we drop time index n in this section for simplicity), which may, for example, contain temperature, meridional wind, zonal wind and surface pressure for each grid point so that the dimension of the local state vector in Fig. 5.1 is $4(\text{states}) \times 9(\text{grids}) = 36$. We also assume that we have local bias estimates $\mathbf{b}_{\mathbf{r}}$ or $\mathbf{c}_{\mathbf{r}}$ depending on the bias model that we have chosen. Then, we form an augmented local state vector,

$$\mathbf{v}_{\mathbf{r}} = \begin{bmatrix} \mathbf{x}_{\mathbf{r}} \\ \mathbf{b}_{\mathbf{r}} \\ \mathbf{c}_{\mathbf{r}} \end{bmatrix}, \quad (5.15)$$

and, by applying Kalman filter equations, we obtain a local analysis. The global analysis is then taken to be the state at the center of each local region.

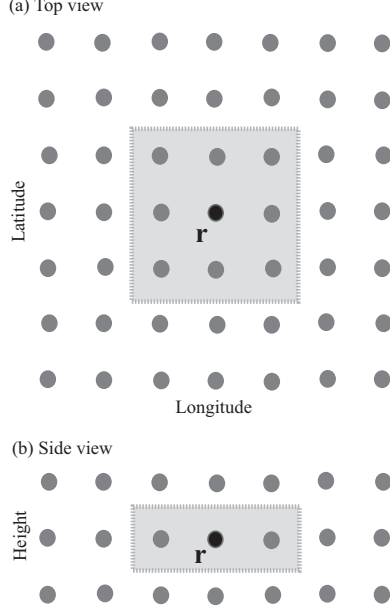


Figure 5.1: Illustration of local region: Dots represent grid points and the grey box represent a local region. (a) is the top view of the local region and (b) is the side view of it. The local region in the figure is a $3 \times 3 \times 1$ box. The black dot in the local region, \mathbf{r} , is the grid point whose states are estimated in the assimilation time using the states and the observation within this local region.

By inserting the local augmented state vector into the Kalman filter equations, we assume that $\psi(\mathbf{v}_{\mathbf{r}}^b)$, the background probability distribution of the local augmented state, is Gaussian; that is,

$$\psi(\mathbf{v}_{\mathbf{r}}^b) \sim \exp \left\{ -\frac{1}{2} [\mathbf{v}_{\mathbf{r}}^b - \bar{\mathbf{v}}_{\mathbf{r}}^b]^\top [\mathbf{P}_{\mathbf{r}}^b]^{-1} [\mathbf{v}_{\mathbf{r}}^b - \bar{\mathbf{v}}_{\mathbf{r}}^b] \right\}, \quad (5.16)$$

where $\bar{\mathbf{v}}_{\mathbf{r}}^b$ is the background mean of the augmented local state, and $\mathbf{P}_{\mathbf{r}}^b$ is the background error covariance matrix for the augmented local state. In LETKF, $\bar{\mathbf{v}}_{\mathbf{r}}^b$ and $\mathbf{P}_{\mathbf{r}}^b$ are estimated using an ensemble of augmented local state vectors $\mathbf{v}_{\mathbf{r}}^{b(i)}$. That is, $\bar{\mathbf{v}}_{\mathbf{r}}^b$ can be estimated as a sample mean,

$$\bar{\mathbf{v}}_{\mathbf{r}}^b = \frac{1}{k} \sum_{i=1}^k \mathbf{v}_{\mathbf{r}}^{b(i)}, \quad (5.17)$$

and \mathbf{P}_r^b as an unbiased sample variance,

$$\mathbf{P}_r^b = (k-1)^{-1} \mathbf{V}_r^b [\mathbf{V}_r^b]^\top, \quad (5.18)$$

where k is the number of ensemble members, $\mathbf{v}_r^{b(i)}$ is the i -th ensemble member ($i = 1, \dots, k$), \mathbf{V}_r^b is the matrix of background ensemble perturbations,

$$\mathbf{V}_r^b = [\delta \mathbf{v}_r^{b(1)} | \delta \mathbf{v}_r^{b(2)} | \dots | \delta \mathbf{v}_r^{b(k)}], \quad (5.19)$$

and

$$\delta \mathbf{v}_r^{b(i)} = \mathbf{v}_r^{b(i)} - \bar{\mathbf{v}}_r^b. \quad (5.20)$$

In this way, our scheme allows for the estimation of cross-correlations between uncertainties in the state estimates and uncertainties in the estimation of the model bias and simultaneously estimates the atmospheric state and the model bias.

In what follows, we briefly introduce LETKF formulation for the augmented state vectors (see Hunt et al. [47] for further discussion). Although, a similar formulation could be applied to other ensemble Kalman filters [e.g., 8], we employ LETKF since it has been found to be effective and efficient for operational weather models [88]. By applying the observation operator, $\hat{\mathbf{h}}(\cdot)$ to an ensemble of the global augmented background, we obtain an ensemble of the global background observation,

$$\mathbf{y}^{b(i)} = \hat{\mathbf{h}}(\mathbf{v}^{b(i)}), \quad i = 1, \dots, k. \quad (5.21)$$

From these ensemble members (5.21), we can obtain an ensemble of local background observation, $\{\mathbf{y}_r^{b(i)}\}$, the local mean, $\bar{\mathbf{y}}_r^b$, and the local perturbation matrix, \mathbf{Y}_r^b , similar to eqs. (5.17), (5.19) and (5.20). Then, the analysis equations for the

augmented local state become

$$\bar{\mathbf{v}}_{\mathbf{r}}^a = \bar{\mathbf{v}}_{\mathbf{r}}^b + \mathbf{V}_{\mathbf{r}}^b \tilde{\mathbf{P}}_{\mathbf{r}}^a (\mathbf{Y}_{\mathbf{r}}^b)^\top \mathbf{R}_{\mathbf{r}}^{-1} (\mathbf{y}_{\mathbf{r}}^o - \bar{\mathbf{y}}_{\mathbf{r}}^b), \quad (5.22)$$

$$\tilde{\mathbf{P}}_{\mathbf{r}}^a = \left[(k-1)\mathbf{I} + (\mathbf{Y}_{\mathbf{r}}^b)^\top \mathbf{R}_{\mathbf{r}}^{-1} \mathbf{Y}_{\mathbf{r}}^b \right]^{-1}, \quad (5.23)$$

$$\mathbf{P}_{\mathbf{r}}^a = \mathbf{V}_{\mathbf{r}}^b \tilde{\mathbf{P}}_{\mathbf{r}}^a (\mathbf{V}_{\mathbf{r}}^b)^\top, \quad (5.24)$$

where $\bar{\mathbf{v}}_{\mathbf{r}}^a$ is the mean analysis for the augmented local state, $\mathbf{P}_{\mathbf{r}}^a$ is the local analysis error covariance matrix, $\mathbf{y}_{\mathbf{r}}^o$ is the local observation vector which consists of the observations within the local region, and $\mathbf{R}_{\mathbf{r}}$ is the local observation error covariance matrix. LETKF determines the matrix of local analysis ensemble perturbation, $\mathbf{V}_{\mathbf{r}}^a$, using a square root of the local analysis error covariance matrix:

$$\mathbf{V}_{\mathbf{r}}^a = (k-1)^{1/2} \mathbf{V}_{\mathbf{r}}^b \left[\tilde{\mathbf{P}}_{\mathbf{r}}^a \right]^{1/2}. \quad (5.25)$$

An ensemble of the local augmented analysis, $\{\mathbf{v}_{\mathbf{r}}^{a(i)}\}$, can then be obtained by adding $\mathbf{v}_{\mathbf{r}}^a$ to each of the columns of $\mathbf{V}_{\mathbf{r}}^a$:

$$\mathbf{v}_{\mathbf{r}}^{a(i)} = \bar{\mathbf{v}}_{\mathbf{r}}^a + \mathbf{V}_{\mathbf{r}}^{a(i)}, \quad (5.26)$$

where $\mathbf{V}_{\mathbf{r}}^{a(i)}$ is the i -th column of $\mathbf{V}_{\mathbf{r}}^a$.

In practice, we modify this approach using multiplicative covariance inflation on the analysis, in effect multiplying $\mathbf{P}_{\mathbf{r}}^a$ by a factor greater than 1. This prevents the filter covariances from decaying toward $\mathbf{0}$ over time. Notice that, since the ensemble of bias estimates does not change during the forecast step, its covariance remains constant during each forecast, and (like the state variable covariance) decreases during each analysis. Thus, if no covariance inflation is applied, the filter's uncertainty about its bias estimates will decrease to zero over time; this is undesirable unless the bias is truly time-independent. We find that it is often advantageous to apply

more inflation to the bias estimates than to the state variables, and to do so we inflate \mathbf{V}_r^a before applying eq. (5.26); see also Sections 5.3.2 and 5.3.3.

5.3 Numerical experiments

5.3.1 Experimental setup

The true states are generated by running the NCEP model for 3 months (91 days), starting from the operational NCEP analysis, which was truncated to T62L28 resolution, at 0000 UTC on 1 January 2004. Observations are generated with six-hour frequency, at 0000 UTC, 0600 UTC, 1200 UTC and 1800 UTC, at seven pressure levels, at 925, 850, 700, 500, 300, 200 and 100 hPa. The horizontal locations of the observations are chosen to coincide with the horizontal locations of the model grid points in the SPEEDY model. When some of seven observed pressure levels are below the surface of the SPEEDY model at a horizontal location, no observation is created at those pressure levels at the given horizontal location. The observations are generated by adding random “observation noise” to the “true” state at the observational locations. The observation noise has zero mean for all observed quantities and standard deviation of 1 hPa for the surface pressure, 1 K for the temperature and 1 m/s for the wind.

Our goal is to account for the bias in the 6-hour background forecasts with the SPPEEDY model. This bias for the 3-month period is shown in Figure 5.2. As can be expected, this surface pressure bias is dominantly due to the difference between the “true” orography of the NCEP model and the “model” orography of the SPEEDY model (Figures 5.3 and 5.4).

If we removed the bias from the surface pressure background, as done in Bias Model I, the surface pressure analysis would have a value consistent with the higher resolution “true” orography. This would induce an adjustment process of the surface

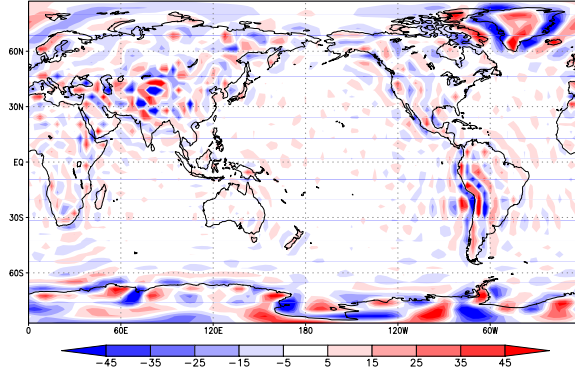


Figure 5.2: 6-hour forecast bias of the surface pressure without bias correction: In many regions, the amount of the bias is proportional to the magnitude of the orography difference but sometimes they are negatively correlated. There are also regions that shows small difference in orography but have moderate amount of the biases.

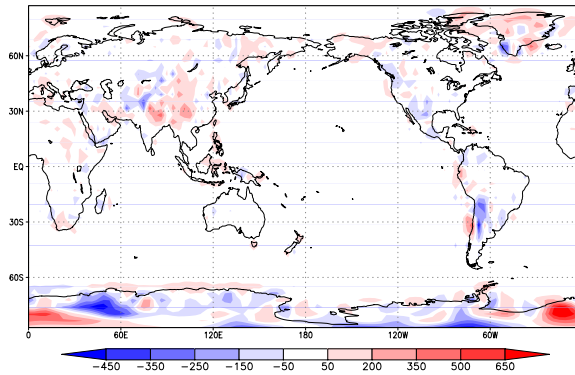


Figure 5.3: Orography difference between SPEEDY and NCEP models: Surface geopotential height of NCEP model is subtracted from that of SPEEDY model. Discrepancies are mostly on mountain areas. The scales are different in negative and positive directions.

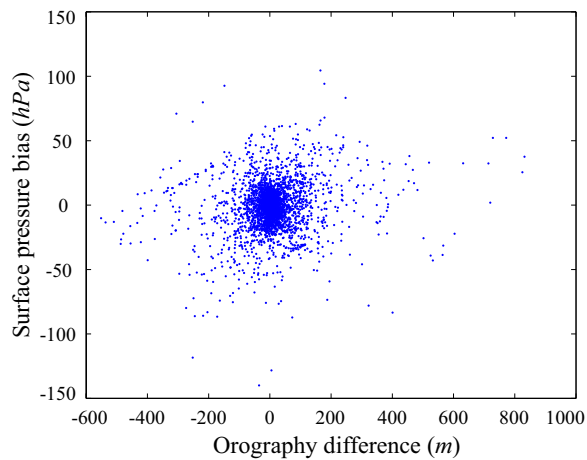


Figure 5.4: Correlation between orography difference and surface pressure bias.

pressure, and of the other state variables through their dynamical relation to the surface pressure, to the SPEEDY orography in the forecast phase of the next analysis cycle. Thus the surface pressure bias of our experimental design, and of the real-life situation, is an ideal candidate to be treated with Bias Model II.

The ensemble of 6-hour model forecast at time t_n is generated by integrating the SPEEDY model from ensemble of the analysis at previous time t_{n-1} . The initial ensemble states at time t_0 , $\{\mathbf{x}_0^{b(i)}\}$, are taken from a free run of the SPEEDY model. The initial condition for the free run is obtained from the NCEP true state at 1800 UTC on 31 March 2004. The initial ensemble estimates for the bias are obtained from Gaussian random numbers with certain means and variances which will be described later. Then the augmented initial ensemble, $\{\mathbf{v}_0^{b(i)}\}$, is formed by augmenting the initial states by the initial bias estimates. We start our data assimilation cycle at time t_0 (= 0000 UTC on 1 January 2004) by assimilating the observation, \mathbf{y}_0^o , into the initial ensemble, $\{\mathbf{v}_0^{b(i)}\}$, to generate the analysis ensemble, $\{\mathbf{v}_0^{a(i)}\}$.

Throughout our experiments, we employ $3 \times 3 \times 1$ grid points for the local regions since we found that increasing the dimension of the local regions does not significantly improve the accuracy of the analyses and forecasts but requires more ensemble members and longer computational time. The local dimension is, therefore, $9 \times 4 = 36$ for the states without bias correction and can be increased up to 72 if we correct for the biases in all state variables with one type of bias model. In all of our experiments, we use 60 ensemble members which must be enough to account for the dimension of the local states augmented by various combinations of bias estimates.

Since the local region is defined in the model space, the number of observations in the local regions for each grid points can be different. In some areas, the sigma levels are dense and therefore the local regions in these areas only can contain very few observations or no observation. For example, we depict two local regions at

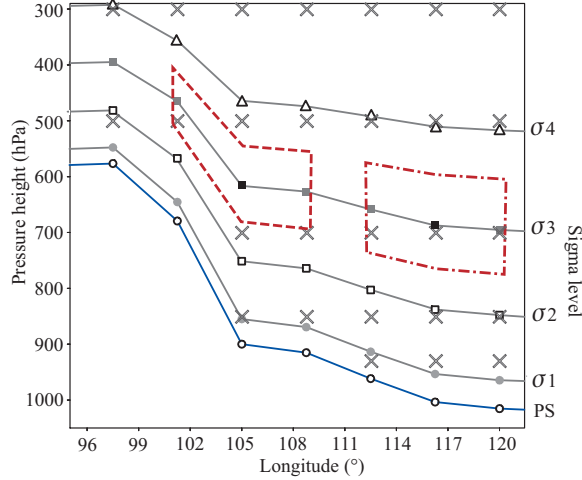


Figure 5.5: Local observation areas at latitude 31.5° between longitude 96° and 120° : The lowest line with circle \circ represents the surface pressure (PS) and above it, four sigma level lines are depicted. Along the sigma levels, we put symbols to indicate the grid point where the model stats are estimated. Symbol X's indicate the available observations which are at levels 925hPa, 850hPa, 700hPa, 500hPa, and 300hPa above the surface. Two $3 \times 3 \times 1$ local regions are shown as a dashed line and a dash-dotted line. We see that they are different from the region in Fig. 5.1(b) and each region contains different number of observations due to the different coordinate space between model space and observation space.

latitude 31.5° and at longitude 96° through 120° along the third sigma level in Fig. 5.5. It is essentially the same with Fig. 5.1 except that it is on the pressure coordinates and contains observations. The observations of winds and temperatures are available at 7 pressure heights above the surface (in the figure we can see only 4 levels). In the Fig. 5.5, the local region enclosed with dashed line has one grid point for the observation of the wind and the temperature that can be used to estimate the states of three grid point enclosed by the line, while the local region enclosed by dash-dotted line has three grid points. It happens near the mountain areas and therefore the accuracy of the analysis and forecast may be worse in these areas.

To assess the quality of the analysis and the forecast, the errors are evaluated in the observation space with the known “true” state at the observation location, \mathbf{y}_n^t , which are generated by running the NCEP model. We define 6-hour forecast

error at time t_n , \mathbf{e}_n^b , as

$$\mathbf{e}_n^b = \bar{\mathbf{y}}_n^b - \mathbf{y}_n^t, \quad (5.27)$$

where $\bar{\mathbf{y}}_n^b$ is the ensemble mean of the background observation (5.21), and similarly the analysis error, \mathbf{e}_n^a , can be defined by

$$\mathbf{e}_n^a = \bar{\mathbf{y}}_n^a - \mathbf{y}_n^t, \quad (5.28)$$

where $\bar{\mathbf{y}}_n^a$ is the ensemble mean of the analysis observation. Then, for each vertical level, we use the root-mean-square (rms) of the analysis and forecast errors over the horizontal grids. For example, the rms of the temperature analysis error is given by

$$\text{rms}(\mathbf{e}_{\tau,n}^a) = \sqrt{\frac{1}{N \times M} \sum_{i=1}^N \sum_{j=1}^M [\mathbf{e}_{\tau,n}^a(i,j)]^2}, \quad (5.29)$$

where $N \times M$ is the number of horizontal grid points at each vertical level and $\mathbf{e}_{\tau,n}^a(i,j)$ is the temperature analysis error at grid (longitude, latitude)=(i,j) at time t_n . (Here, we use subscript τ to represent temperature error. For other states, we use subscript v for meridional wind, u for zonal wind and p_s for surface pressure.) The overall performance of the assimilation scheme is measured by the time average of the rms error over a long time interval T ,

$$\langle\langle \mathbf{e}_{\tau}^a \rangle\rangle = \frac{1}{T} \sum_{n=n_0+1}^{n_0+T} \text{rms}(\mathbf{e}_{\tau,n}^a), \quad (5.30)$$

where n_0 indicates the time we allow for the analysis to converge near the true state. In our experiments, we allow $n_0 = 10$ days to converge and hence the rms error is averaged over next $T = 81$ days to obtain $\langle\langle \mathbf{e}_{\tau}^a \rangle\rangle$. Similarly, in order to assess the ability to correct for the model bias, we examine the analysis bias, $\langle \mathbf{e}_n^a \rangle$, and the forecast bias, $\langle \mathbf{e}_n^b \rangle$, which are defined by the time average of \mathbf{e}_n^a and \mathbf{e}_n^b over the time

interval T . For example, the analysis bias for the temperature can be obtained by

$$\langle \mathbf{e}_T^a \rangle = \frac{1}{T} \sum_{n=n_0+1}^{n_0+T} \mathbf{e}_{T,n}^a. \quad (5.31)$$

We also examine the rms of the bias over the horizontal grids to assess overall bias correction performance for each vertical level. For the temperature bias, we obtain its rms by

$$\text{rms}(\langle \mathbf{e}_T^a \rangle) = \sqrt{\frac{1}{N \times M} \sum_{i=1}^N \sum_{j=1}^M \langle \mathbf{e}_T^a(i, j) \rangle^2}. \quad (5.32)$$

5.3.2 Data assimilation without bias correction

In our first numerical experiment, we try to keep the analysis observation, $\bar{\mathbf{y}}_n^a$, and the background observation, $\bar{\mathbf{y}}_n^b$, close to the true state at the observation location, \mathbf{y}_n^t , for all time without explicitly accounting for the model error. That is, we just use a covariance inflation scheme to achieve our goal. General purpose of employing a covariance inflation is to correct for the loss of covariance in the ensemble due to nonlinearities and sampling errors. Covariance inflation can also stabilize the Kalman filter in the presence of model errors [70, 96]. In our numerical experiments, we employ multiplicative covariance inflation on the analysis error covariance,

$$\mathbf{P}_r^a \rightarrow (1 + \gamma) \mathbf{P}_r^a, \quad (5.33)$$

where γ is an inflation coefficient. From eq. (5.25), this is equivalent to multiplying the analysis perturbations by $\sqrt{1 + \gamma}$ and since the inflation does not depend on the location in our experiments, we can inflate the perturbation globally, such that

$$\mathbf{V}^a \rightarrow \sqrt{1 + \gamma} \mathbf{V}^a. \quad (5.34)$$

Figure 5.6 (a) shows the rms values of the analysis bias and forecast bias

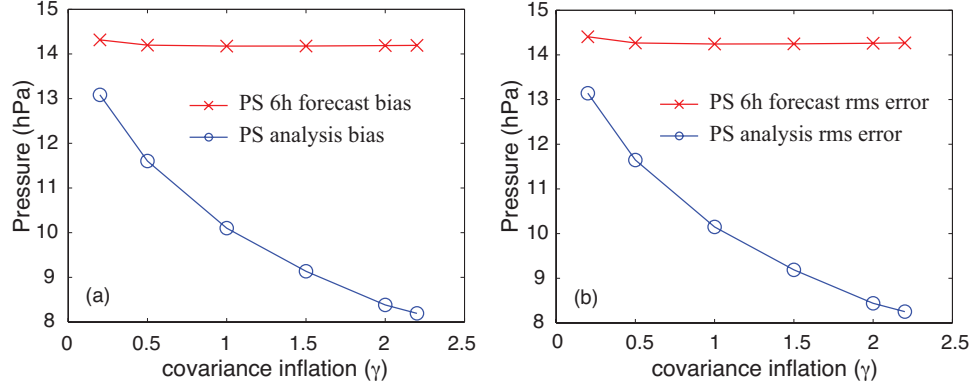


Figure 5.6: Surface pressure analysis and 6-hour forecast rms errors and biases versus variance inflation without bias correction: The analysis error of surface pressure is getting smaller as the variance inflation is getting larger. However, the 6-hour forecast error remains almost the same.

of the surface pressure, $\text{rms}(\langle \mathbf{e}_{\text{PS}}^a \rangle)$ and $\text{rms}(\langle \mathbf{e}_{\text{PS}}^b \rangle)$, and Fig. 5.6 (b) shows the time averaged rms errors of the analysis and forecast, $\langle \langle \mathbf{e}_{\text{PS}}^a \rangle \rangle$ and $\langle \langle \mathbf{e}_{\text{PS}}^b \rangle \rangle$, versus covariance inflation coefficient, γ . Since more weight is put on the observation by increasing the inflation coefficient, the analysis is getting closer to the observation and the rms analysis error is decreasing as the inflation coefficient is increasing. Meanwhile, the forecast rms error does not have significant improvement as the inflation coefficient increases and has its minimum at around $\gamma = 1.0$. By comparing Fig. 5.6 (a) and (b), we can see that large portion of the error in the surface pressure comes from its bias and a better analysis of the surface pressure does not produce a better forecast.

The same characteristics are seen in the zonal wind as shown in Fig. 5.7 where analysis rms errors, 6-hour forecast rms errors, analysis biases, and 6-hour forecast biases of the zonal wind for a range of the covariance inflation coefficients are depicted. It is apparent that a better analysis does not generate a better forecast beyond $\gamma = 1.0$. We also see that the analysis rms error and bias keep getting better as the coefficient is increasing until the model states become out of accepted range for the SPEEDY model at $\gamma = 2.2$ like the surface pressure. It is also seen in

the meridional wind but we do not show this case here.

The temperature error, however, has a little different characteristics. 6-hour forecast rms errors and biases of the temperature also shows the same characteristics as the winds and the surface pressure for which the accuracy of them is not improved beyond certain covariance inflation as shown in Fig. 5.8. Meanwhile, the analysis rms error of the temperature is not always getting better as the covariance inflation is increasing. Each height has different optimal values of the inflation coefficient and even for some coefficients, the analysis rms errors are worse than the forecast rms errors. For example, we see that the analysis rms error at 700hPa height with $\gamma = 2.2$ is worse than the forecast rms error at the same height with the same coefficient.

It has been shown that, in local ensemble Kalman filtering, increasing the inflation coefficient beyond a certain value does not improve the analysis and further increasing may degrades the performance [70, 8]. In our experiments, however, since the model error is too large, we see the improvements for large values of the inflation coefficient. Furthermore, too much increase of the coefficient makes the spread of analysis ensemble so large that the states of some ensemble members are out of balance and rejected by the forecast model so that the forecasts corresponding to those ensemble members are not available. In this experiment, it happens around $\gamma = 2.2$. Even though we could employ a filtering process to balance out the states to prevent such an event [for example, initialization process can be employed to prevent spurious gravity waves [59, 58]], we decided not to employ such a process to increase the inflation coefficient since the 6-hour forecast errors are already saturated around $\gamma = 1.0$ and further increase of γ does not improve the quality of the forecasts.

The forecast bias of the surface pressure, $\langle \mathbf{x}_{PS}^b \rangle$, with covariance inflation $\gamma = 1.5$ is shown in Fig. 5.2. One may notice that the regions with large biases are at and near the regions with large orography differences in Fig. 5.3, even though we

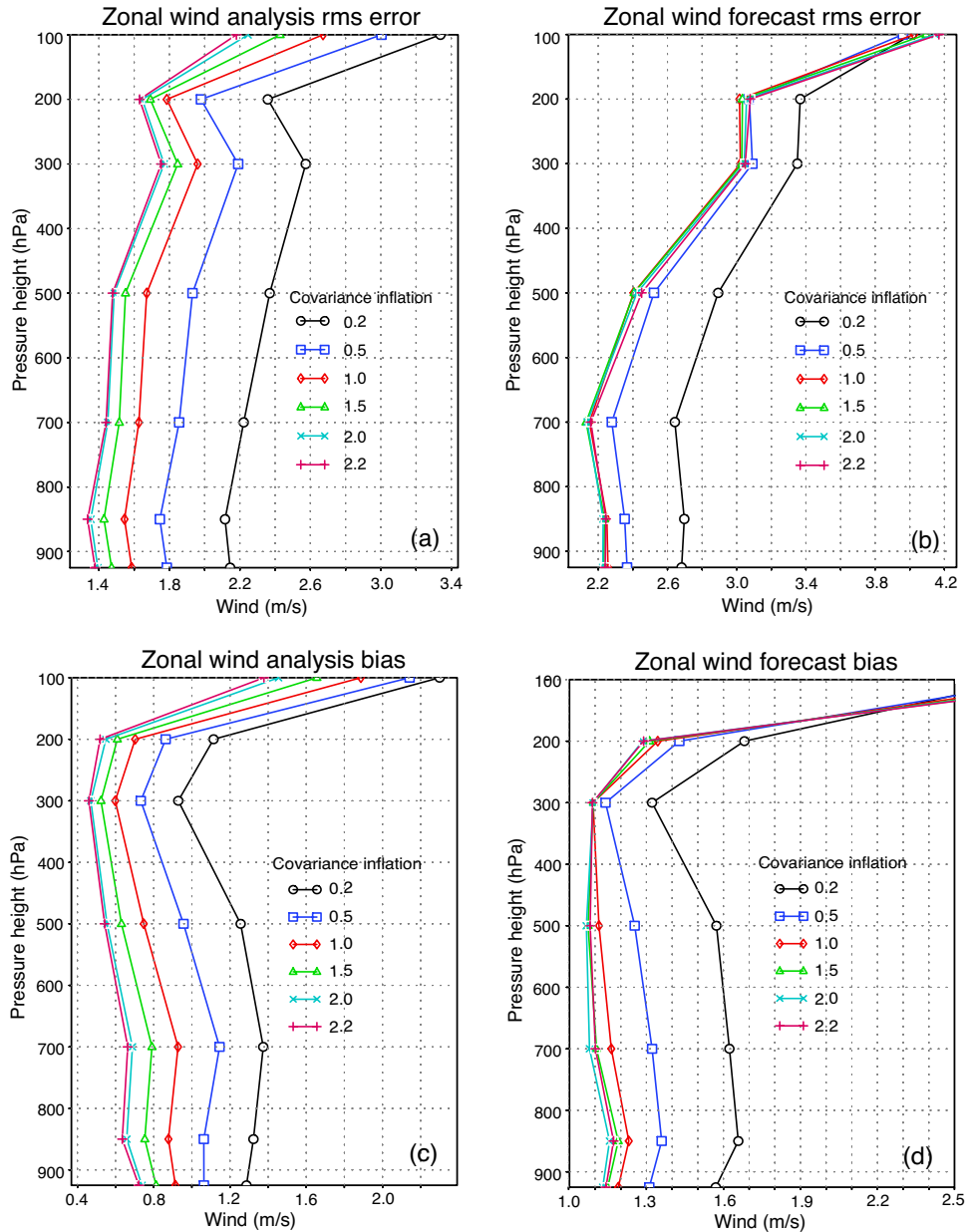


Figure 5.7: Zonal wind analysis rms errors (a), 6-hour forecast rms errors (b), analysis biases (a), and 6-hour forecast biases (b) without bias correction: Similar to the case of the surface pressure, the analysis errors become smaller as the covariance inflation coefficient increases but the forecast errors stay the same after $\gamma = 1.0$.

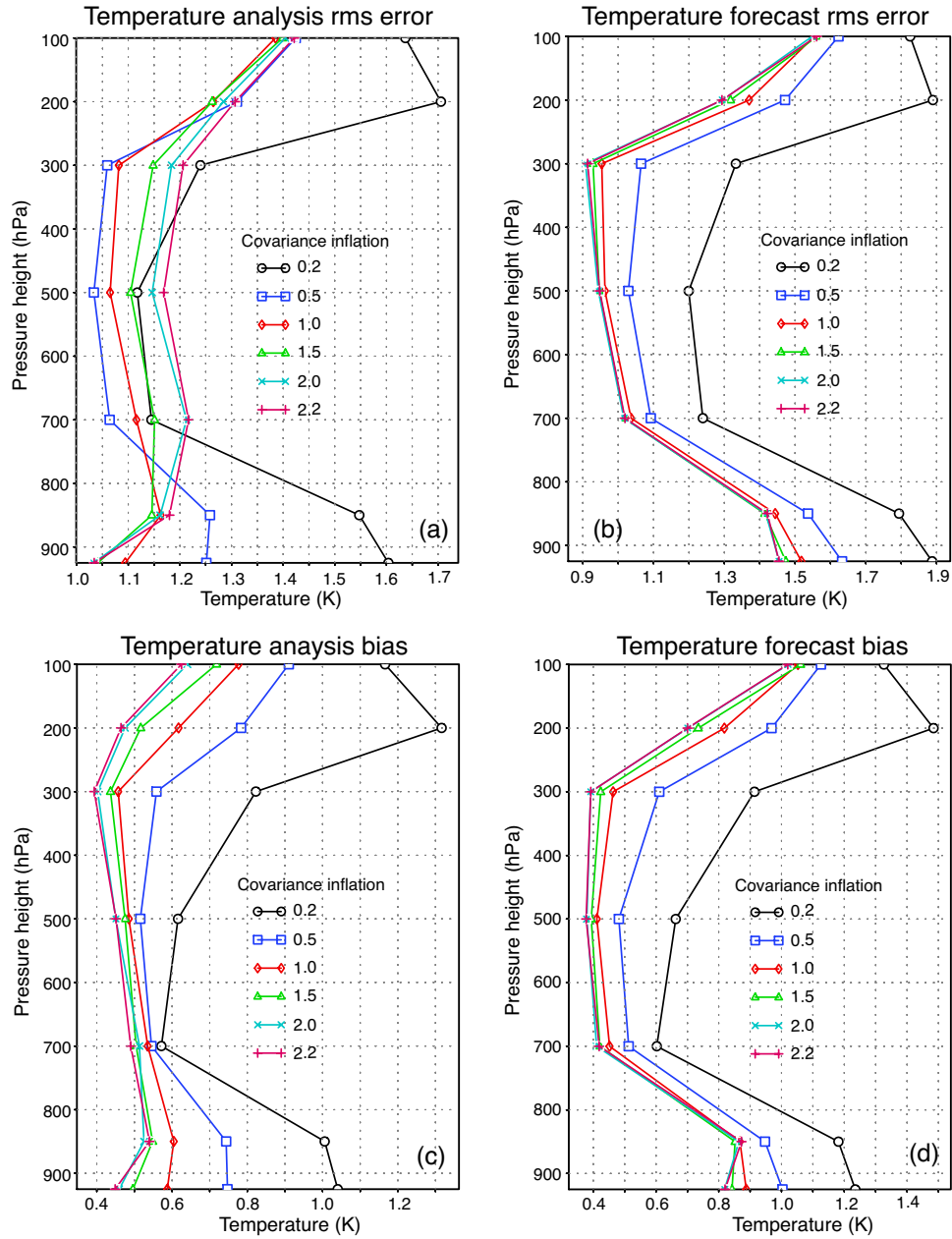


Figure 5.8: Temperature analysis rms errors (a), and 6-hour forecast rms errors (b), analysis biases (c), and 6-hour forecast biases (d) without bias correction: The forecast rms error, the analysis bias and the forecast bias of the temperature become smaller as the covariance inflation coefficient increases. For the inflation coefficient greater than 1.0, the analysis rms error of the temperature is greater than the forecast rms error.

adjust the surface pressure to compensate the orography difference. We employ a simple formula to perform the adjustment. Denoting the adjusted surface pressure for the orography difference Δz by $x_{\text{PS}}^b(\Delta z)$

$$x_{\text{PS}}^b(\Delta z) = x_{\text{PS}}^b \cdot \exp\{g\Delta z/R\bar{T}\}, \quad (5.35)$$

$$\bar{T} = T_s + 0.5\Gamma\Delta z, \quad (5.36)$$

where x_{PS}^b is the surface pressure forecast obtained from the SPEEDY model, $g = 9.8 \text{ m/s}^2$ is the gravitational acceleration, $R = 287 \text{ J kg}^{-1}\text{K}^{-1}$ is the gas constant for dry air, T_s is the surface temperature and $\Gamma = -6.5 \text{ K km}^{-1}$ is the lapse rate of temperature. Although the adjustment formula (5.35) is very simple, the resulting accuracy of the forecasts are better than that of the forecasts in case that we do not adjust orography differences, especially in the regions with large orography differences in our experiments. However, the biases in that region are still large and the amplitude of the bias is proportional to the amplitude of the orography difference but some are negatively correlated. The forecast biases of the surface pressure is also in the region where the orography differences are so small that the adjustments by the above formula are negligible. As a result, the overall forecast bias of the surface pressure is still large even after the orography adjustment.

5.3.3 Bias correction for surface pressure

We correct for the bias in the surface pressure since the error in the surface pressure mostly comes from the bias in the surface pressure unlike other state variables and hence the correction for the surface pressure bias may result in a significant improvement in the accuracy of the surface pressure analysis and forecast. Furthermore, the surface pressure variables are included in all the local state vectors even for the higher levels and, therefore, if the surface pressure becomes more accurate,

the accuracy of the other state variables in all the levels can also be improved. In this section, we mainly present the results from the case that we correct the surface pressure bias with Bias Model II since this scheme is most effective in improving overall accuracies of the analysis and the forecast.

The observation operator for surface pressure bias correction, $\hat{\mathbf{h}}$, consists of a sequence of two operations, the vertical interpolation to obtain background observation in pressure coordinate and subtraction of the bias estimate from the surface pressure to correct for the surface pressure bias. That is, we write

$$\hat{\mathbf{h}}(\mathbf{v}) = \mathcal{H}(\mathbf{x}) - \begin{bmatrix} \mathbf{0} \\ \mathbf{c} \end{bmatrix}, \quad (5.37)$$

where \mathcal{H} is the vertical interpolation operator and \mathbf{c} is the estimate of the surface pressure bias. Note that the interpolation precedes the subtraction. As shown in Fig. 5.5, an observation at a certain pressure height is assimilated to estimate a model state at a sigma level whose pressure height is closer to the pressure height of the observation. The pressure height of this sigma level is obtained by a simple function of the surface pressure of the model state. Therefore, in order to obtain the background observation or the analysis observation, we must use the surface pressure of the model state to obtain the pressure heights of the sigma levels and interpolate the state variables based on these pressure heights to obtain the background observation or the analysis observation.

In correcting for the surface pressure bias, we modify the inflation scheme (5.34) so that each state perturbations and bias perturbations are multiplied by

different amounts of covariance inflation,

$$\mathbf{V}^a \rightarrow \begin{bmatrix} \sqrt{1 + \gamma_x} \mathbf{X}^a \\ \sqrt{1 + \gamma_{\text{PS}}} \mathbf{X}_{\text{PS}}^a \\ \sqrt{1 + \gamma_c} \mathbf{C}^a \end{bmatrix}, \quad (5.38)$$

where γ_{PS} is the inflation coefficient for the surface pressure state variables, γ_x is the coefficient for the state variables other than surface pressure, γ_c is the coefficient for the bias estimates, and subscript PS represents surface pressure. Here, \mathbf{X}^a , \mathbf{X}_{PS}^a and \mathbf{C}^a are perturbation matrices for state variables and bias estimates as defined in eqs. (5.19) and (5.20). We keep the ratio of the three coefficient in a sufficient range in order to prevent some ensemble members from being out of balance for the forecast model. In general, if we inflate each state variables with different ratio, it is possible to have one or more ensemble members whose state variables are not reasonable for the atmospheric states to the forecast model. In our experiments, we have to keep the ratio of the three coefficients between 1/3 and 3.

In Fig. 5.9, we show the 6-hour forecast bias after we correct for the surface pressure bias using Bias Model II with the covariance inflation coefficients $\gamma_x = 1.0$, $\gamma_{\text{PS}} = 0.4$ and $\gamma_c = 0.8$. The initial bias estimate for this experiment has zero mean and ensemble spread of 10 hPa. Comparing with Fig. 5.2, we can see that the bias correction scheme significantly reduces the bias in the surface pressure. (Note that the color-coded pressure scale in Fig. 5.9 is 1/20 of that in Fig. 5.2.) Hence, the time averages of the analysis and the forecast rms errors for the surface pressure, $\langle\langle \mathbf{e}_{\text{PS}}^a \rangle\rangle$ and $\langle\langle \mathbf{e}_{\text{PS}}^b \rangle\rangle$, are significantly improved as shown in table 5.1 such that both the analysis and forecast rms errors are more than 8 times smaller than before.

The improved accuracy in the surface pressure analysis and forecast leads to improvement in accuracies in the analysis and the forecast of other state variables as shown in Figs. 5.10 and 5.11. The time averaged rms analysis error (solid lines)

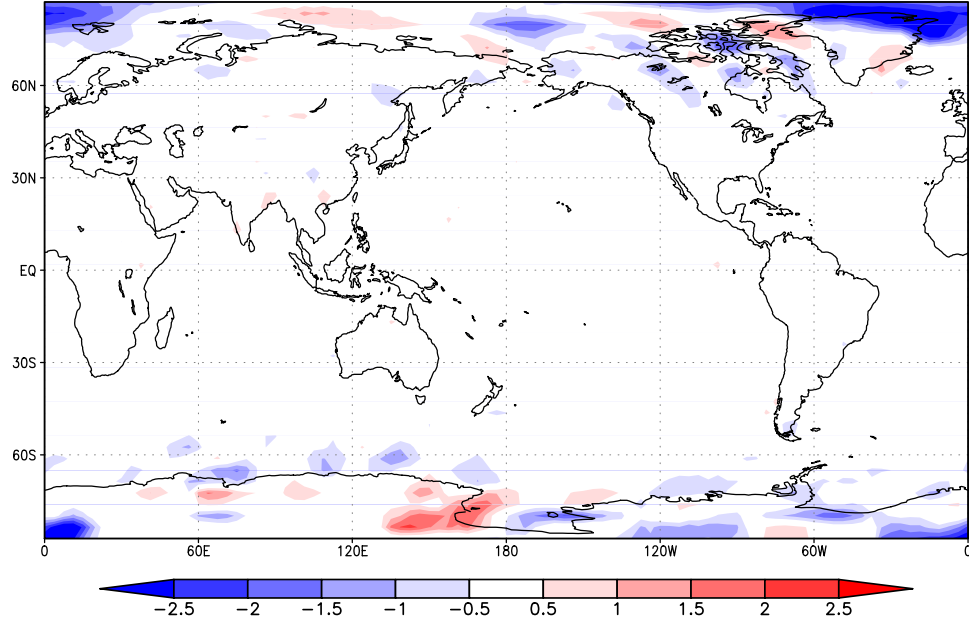


Figure 5.9: 6-hour forecast bias of the surface pressure with type II surface pressure bias correction: Most of the large biases existed in the experiments in Sec. 5.3.2 are mitigated to the magnitude less than 1 hPa except for the polar areas.

Table 5.1: Time average of the surface pressure analysis and forecast rms error: Bias correction for the surface pressure with Bias Model II results in a significant improvement in the analysis and the forecast.

	No correction	Correction with Bias Model II
$\langle\langle e_{PS}^a \rangle\rangle$	7.95 hPa	0.90 hPa
$\langle\langle e_{PS}^b \rangle\rangle$	14.25 hPa	1.75 hPa

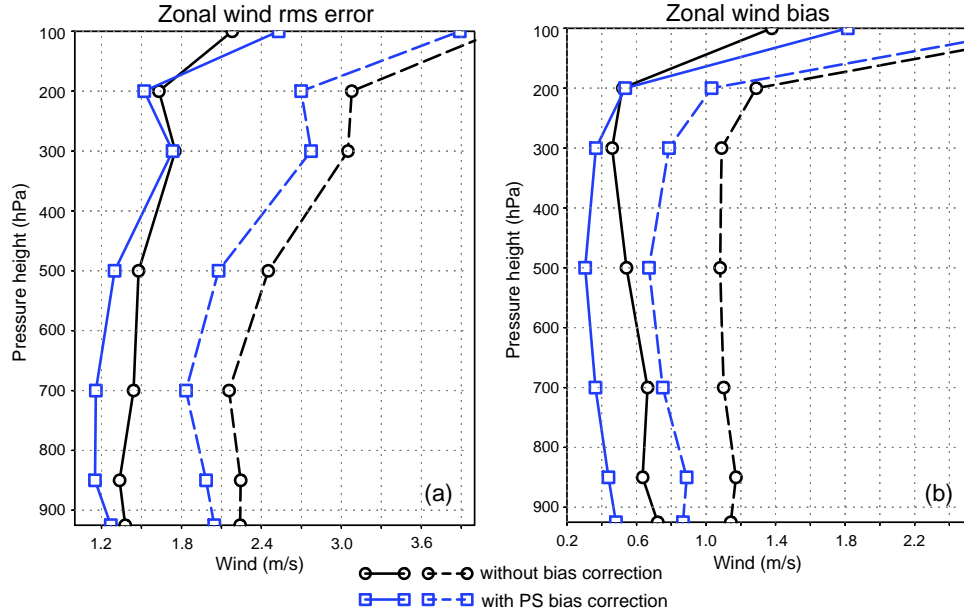


Figure 5.10: Time averages of analysis and 6-hour forecast rms errors of the zonal wind (a) and rms values of the analysis and 6-hour forecast bias of the zonal wind (b) with and without surface pressure bias correction: Solid lines represent the analysis rms errors and biases, and dashed lines represent the forecast rms errors and biases. The accuracy of the forecast is more improved than the accuracy of the analysis. At the highest level, the analysis error is larger than the case that no bias correction is performed.

and forecast error (dashed lines) of zonal wind are plotted in Fig. 5.10(a), and rms values of the analysis bias (solid lines) and forecast bias (dashed lines) are shown in Fig. 5.10(b) with and without surface pressure bias correction. For the comparison, we plot the same quantities of the zonal wind using the line with symbol circle \circ in the case that the surface pressure bias is not corrected. In this case, covariance inflation coefficient $\gamma = 2.2$ is used. The forecast and the analysis is improved at all pressure levels except the analysis at pressure level 100hPa. The same pattern is observed in the analysis and forecast errors of the meridional wind variable (not shown here). In both wind state variables, the forecasts are more improved than the analyses by correcting for the surface pressure bias.

In Fig. 5.11, we plot the time averaged rms error of the analysis (solid lines) and forecast (dashed lines) and rms values of the analysis (solid lines) and forecast

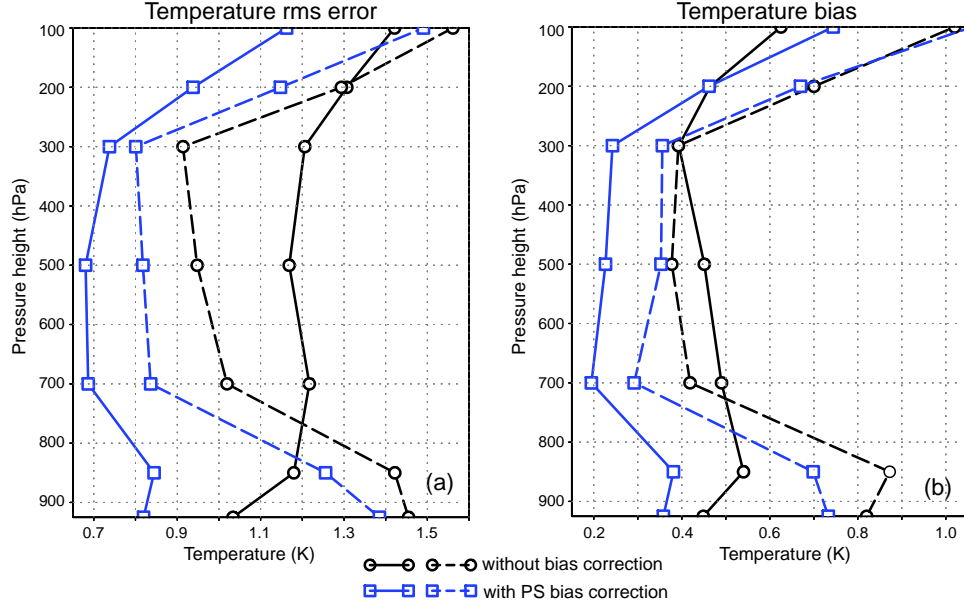


Figure 5.11: Time averages of analysis and 6-hour forecast rms errors of the temperature (a) and rms values of the analysis and 6-hour forecast bias of the temperature (b) with and without surface pressure bias correction: Solid lines represent the analysis rms errors and biases, and dashed lines represent the forecast rms errors and biases. The accuracy of the analysis is better than the accuracy of the forecast. At the highest level, the biases of the analysis and the forecast are both degraded.

(dashed lines) biases for the temperature. As done in Fig. 5.10, we also plot the same quantities from the experiments where no bias correction performed. Unlike the wind variables, for the temperature, the accuracy of the analysis is more improved than the accuracy of the forecast. Also, both of the analysis bias and the forecast bias are degraded at the highest level.

In the SPEEDY model, the top and th bottom layer provide a bulk representation of the stratosphere and the planetary boundary layer [63] while, in the NCEP model, several model layers represent these layers. From the experiments in Sec. 5.3.2, we can see that the dynamics of the bottom layer of the SPEEDY model represents the planetary boundary layer of the NCEP model than the top layer represents the stratosphere of the NCEP model and it is still true in case that the surface pressure bias is corrected.

In fig. 5.12, we plot the rms value (solid line), $\text{rms}(\mathbf{c}_n^b)$, the global spread

(dashed line), $\text{spr}(\mathbf{c}_n^b)$, and the global mean (dash-dotted line), $\text{ave}(\mathbf{c}_n^b)$, of the bias estimate for the surface pressure. From eq. (5.29), the rms value of the bias estimate at time t_n is given by

$$\text{rms}(\mathbf{c}_n^b) = \sqrt{\frac{1}{N \times M} \sum_{i=1}^N \sum_{j=1}^M [c_n^b(i, j)]^2}, \quad (5.39)$$

where \mathbf{c}_n^b is the bias estimate of the surface pressure forecast at time t_n . The global mean is an averaged value of the bias estimate over the globe at time t_n ,

$$\text{ave}(\mathbf{c}_n^b) = \frac{1}{N \times M} \sum_{i=1}^N \sum_{j=1}^M c_n^b(i, j), \quad (5.40)$$

and the global spread is the rms value of the bias estimate whose global mean is eliminated,

$$\text{spr}(\mathbf{c}_n^b) = \sqrt{\frac{1}{N \times M} \sum_{i=1}^N \sum_{j=1}^M [c_n^b(i, j) - \text{ave}(\mathbf{c}_n^b)]^2}. \quad (5.41)$$

The bias estimate captures the global shape of the surface pressure bias within few days and increases slowly after that due to the seasonal variation of the surface pressure in the NCEP model. This seasonal variation is also observed in the state variable of the surface pressure, \mathbf{x}_{PS} . That is, the seasonal variation is split to the state variable, \mathbf{x}_{PS} , and the bias estimate, \mathbf{c} .

The global mean of the bias estimate increases throughout the experimental period but slows down at the end of the interval. As the bias estimate increases, the model state for the surface pressure also increases. It indicates that the SPEEDY model with the states that has an air mass higher than normal provides a dynamics close to that of the NCEP model. In our experiments, the bias estimate finds the global shape of the surface pressure bias, but it slowly finds the difference of the surface pressure between the two models that leads to higher air mass in the

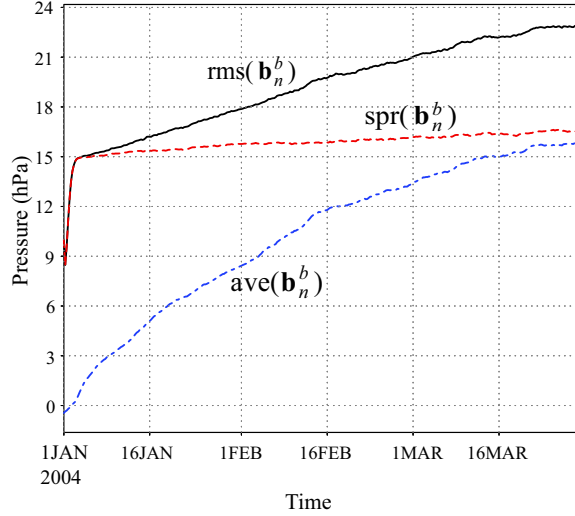


Figure 5.12: Rms, global spread and global average of the bias estimate: In a few time steps, the bias estimate captures the global structure of the bias in the surface pressure and slowly push the surface pressure to a high air mass state where the SPEEDY model can provide a dynamics similar to that of the NCEP model.

SPEEDY model. Finding the global shape of the surface pressure bias does not require the other state variables to be changed since the surface pressure is corrected for the bias after the interpolation and it does not change the balance between state variables. However, finding the optimal air mass requires the other state variables to be changed to have a proper balance with increased surface pressure and it happens very slowly in our experiments. In order to further examine the increase of the mean of the bias estimate, we reuse the ensemble of the final forecast (i.e., the forecast at 0000 UTC on 1 April 2004) as the initial ensemble states at time t_0 (i.e., 0000 UTC on 1 January 2004) and run another 3 months data assimilation. We keep reusing the final forecast again until the mean of the bias estimate does not increase during 3 months experimental period. After 3 reuse, the mean of the bias estimate keeps the value around 29.5 hPa. During this process, the accuracy of the forecast and the analysis remains the same indicating that the process is a normal phenomenon.

5.4 Conclusions and discussion

In this chapter, we presented a modified Ensemble Transform Kalman Filter (LETKF) scheme to correct for model bias and investigated the potential benefit of correction for the bias in the surface pressure.

We simulated the situation of numerical weather prediction by using Global Forecast System of the National Centers for Environmental Prediction with high resolution to generate the true atmospheric states while using the SPEEDY (Simplified Parametrization, primitivE-Equation DYNamics) model to perform forecasts with low resolution. And hence, in the simulated situation, the true atmospheric states and the forecast model states have different resolution, different boundary conditions and different physical parameterizations. We also assumed that the observation takes place in pressure coordinates while the model state evolves in σ -coordinates.

We introduced three different bias models for use in the state space augmentation method. Bias Model I is based on the assumption that the best forecast is obtained when the analysis at the previous assimilation time is as close to the true state as possible. In Bias Model II, it is assumed that there exists a transformation from trajectories of the model state to the trajectories of the true atmospheric state. Bias Model III combines Bias Model I and Bias Model II. We found that, among the above bias models, correcting the bias in the surface pressure using Bias Model II is most effective in such a way that it improves the accuracies of the analyses and the forecasts of all the atmospheric states considered here, surface pressure, zonal wind, meridional wind and temperature.

In our experiments, it seems that the SPEEDY model with high air mass better follows the dynamics of the NCEP model than that with normal air mass does. Our bias correction scheme successfully estimates this difference as well as the spatial distribution of the surface pressure bias. It takes only few days to correct for the bias in the surface pressure while it takes long time to reach the state with

optimal air mass. However, it is shown that the slow movement to the optimal state is normal and the performance is not degraded along the process.

We also found that it is important to make the covariance inflation coefficients different for each state estimates and the bias estimates to have the bias correction scheme work properly. However, the difference should not be too large otherwise the states of some ensemble members of the analysis from the inflated covariance tends to be out of balance for the forecast model.

Bias Model II can also be considered as a model for the observation bias with a different situation and is applicable to the correction for the observation bias. In reality, both the forecast model and the observation can be biased depending on the nature of the observation and the model states and it may be reasonable to consider the forecast model bias and the observation bias at the same time. It is, however, not clear that the model bias and the observation bias can always be correctly and simultaneously identified in the analysis [23] although it has to be considered in the future to improve the quality of the forecasts.

Finally, we note that it is possible to further improve the accuracy of other variables, for example, the temperature, by estimating the bias in the temperature in addition to estimating the bias in the surface pressure. It may increase the dimension of the augmented local state vectors and hence require more ensemble members to properly estimate the covariance matrices. One can reduce the dimension of the augmented state by employing a different parameterization for the bias. For example, it can be assumed that the bias in the surface pressure is proportional to the surface temperature as done by Fertig et al. [32] in correcting for the bias in the satellite observation. Effective parameterization of the bias lead to smaller number of ensemble members or inclusion of many bias estimates that can mitigate the effect of model bias in many atmospheric states. It might even be possible to differentiate the bias in the forecast model and the bias in the observation model

and hence they can be corrected at the same time.

Appendix A

Decay of Q_m with Increasing m

In this Appendix we give a heuristic argument suggesting that, in typical cases, it is reasonable to hypothesize that, for large m , Q_m decays exponentially with increasing m . For definiteness, we consider the case of a map $x_{n+1} = f(x_n)$ of the real line, $-\infty \leq x \leq +\infty$, which has a single chaotic attractor in some bounded region of x , and we also assume that this attractor has a natural invariant measure [also called a SRB (Sinai-Ruelle-Bowen) measure]. By definition the natural measure is the unique invariant measure, ν , such that, for any smooth function $s(x)$, the time average of $s(x_n)$ over an orbit is $\int s(x)d\nu$ for orbits generated by Lebesgue almost every initial condition x_0 in the basin of attractor. Recalling that $g(x)$ and $q(x)$ are smooth bounded functions, we anticipate that the decay of Q_m does not depend critically on details of these functions. Thus we consider the illustrative example of Eq. (2.19). Using Eq. (2.8) we express Eq. (2.19) as

$$Q_m = \langle f'(x_n)f'(x_{n-1}) \cdots f'(x_{n-m+1}) \rangle.$$

The average, $\langle \cdots \rangle$, is over an infinite number of initial conditions $x_0^{(i)}$ which are distributed on the attractor according to the natural invariant measure. Since the natural measure is invariant,

$$\begin{aligned} Q_m &= \langle f'(x_{m-1})f'(x_{m-2}) \cdots f'(x_0) \rangle \\ &= \langle \delta_m \rangle. \end{aligned}$$

The quantity $\langle \delta_m \rangle$ has a simple geometric interpretation. Imagine that, at time $m = 0$, we displace all the initial condition by the same amount dx_0 . That is,

we rigidly translate the natural invariant measure by an amount dx_0 . Thus, at any subsequent time m , $\langle \delta_m \rangle dx_0$ is the displacement of the evolved orbits averaged over all orbits. In other words, $\langle \delta_m \rangle dx_0$ is the displacement of the centroid of the evolved measure from the centroid of the natural invariant measure. Since, by definition, the natural invariant measure is generated by the time average of Lebesgue almost any initial condition in the basin of the attractor, $\langle \delta_m \rangle$ should relax to zero as m increases. Thus, for the example (2.19), our hypothesis that Q_m decays exponentially, is equivalent to the hypothesis that the displaced centroid of a cloud of orbits relaxes exponentially to its equilibrium value. This is rigorously true for the case of hyperbolic attractors [100], and we also adopt it as a useful working hypothesis for the general case. We caution, however, that this hypothesis may not always be valid (see Sec. 2.3.3 and Appendix B).

Appendix B

Orbit Densities for Ensembles of Logistic Maps

In this Appendix, we attempt to gain understanding concerning the observed lack of convergence found for the example in Sec. 2.3.3. To do this we numerically examine how the orbit density evolves after a small perturbation from the natural time asymptotic invariant density. We evolve a large number of orbits (10^7 with $\mu = 3.9$), initially uniformly distributed, forward in time for many iterates, (to approximately reach the invariant orbit density), and, by duplicating, two identical orbit distributions are created. Then, one of the orbit distributions is perturbed by adding the same small perturbation $\delta x_0 = 10^{-3}$ to every orbit points. We then evolve both sets of orbits forward in time and observe the orbit densities to see how the perturbed density relates to the unperturbed invariant density. We divide the interval $(0, 1)$ into 1000 subintervals, count the number of orbits in each subinterval, normalize the numbers, and plot these numbers for each subinterval. This histogram procedure yields an approximation to the density with resolution 10^{-3} .

In Figure B.1, we plot our histogram approximation of the orbit densities for the perturbed case (grey) and for the unperturbed case (black) in a small interval $0.35 \leq x \leq 0.55$. The perturbation $\delta x_0 = 10^{-3}$ rigidly shifts the original invariant density slightly to the right. Setting t to 0 at this moment, Figure B.1(a) and (b) show the orbit densities at $t = 4$ and $t = 10$ respectively. We find that the perturbed density is distorted significantly by around $t = 8$ so that the outstanding peaks in the perturbed density do not match those in the unperturbed (i.e. invariant) density. Thus the small perturbation δx_0 in the orbit location points leads to large perturbations in the absolute value of the histogram approximation of orbit density

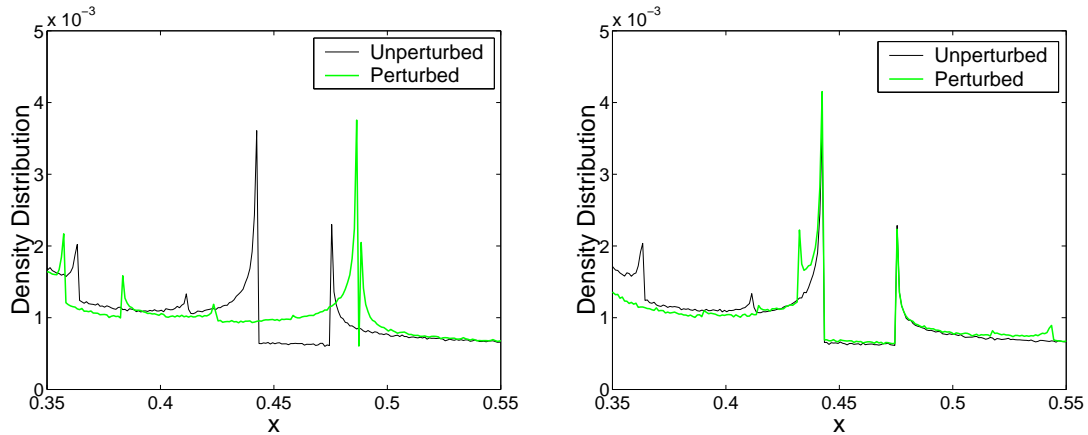


Figure B.1: The histogram approximated orbit densities in the interval $[0.35, 0.55]$ for ensembles of 10^7 identical logistic maps with parameter $\mu = 3.9$ with the initial perturbation $\delta x_0 = 10^{-3}$ (grey) and without perturbation (black): We set time t to 0 at which the perturbation is applied. (a) and (b) show the densities at $t = 4$ and $t = 10$ respectively.

near points where the histogram approximated density has strong narrow peaks. After $t = 8$, the histogram approximated perturbed density becomes closer to the histogram approximated invariant density. However, it is to be expected that as the resolution of the histogram is increased, large differences in the approximated densities would be observed out to later and later times. This is a reflection of the singular nature of the density and suggests that the perturbation theory approach is not valid.

Appendix C

Background Error Covariance Matrix

In Fig. C.1, we plot the Frobenius norms of \mathbf{P}_k^b and $\mathbf{P}_k^b - \mathbf{P}_{k-1}^b$ for each iteration step k of our algorithm for determining \mathbf{P}^b . Here, we use 5×10^4 samples of background error vectors to estimate \mathbf{P}^b for each iteration step assuming that the observational error variance is $\sigma_v^2 = 10^{-4}$ (the value used in Section 3.3). We see that, after about 10 iteration steps, $\|\mathbf{P}_k^b\|$ and $\|\mathbf{P}_{k-1}^b\|$ changes only very little, and $\|\mathbf{P}_k^b - \mathbf{P}_{k-1}^b\|$ is at least one order of magnitude smaller than $\|\mathbf{P}_k^b\|$.

In Fig. C.2, we show the estimate of the background error covariance at iteration step $k = 40$. (We note that beyond $k = 5$, the estimate maintains a similar shape throughout the process.) The background errors at neighboring locations are correlated, but the spatial correlation length is rather small.

We also obtain an estimate of the background error covariance matrix in the case when observations are not taken at location 20. The evolution of the Frobenius norms of \mathbf{P}_k^b and $\mathbf{P}_k^b - \mathbf{P}_{k-1}^b$ (not shown) is very similar to that shown in Fig. C.1. On the other hand, important changes can be observed in the shape of the \mathbf{P}^b (Fig. C.3).

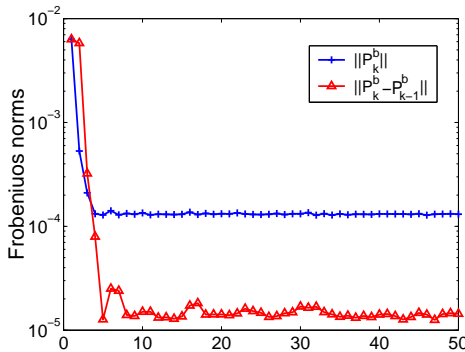


Figure C.1: Convergence of the Frobenius norms of \mathbf{P}_k^b and $\mathbf{P}_k^b - \mathbf{P}_{k-1}^b$ for Experiment 1.

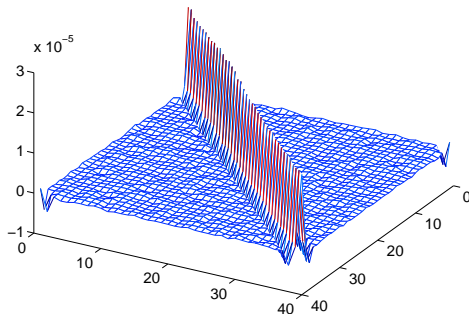


Figure C.2: \mathbf{P}^b after 40 iteration steps for Experiment 1.

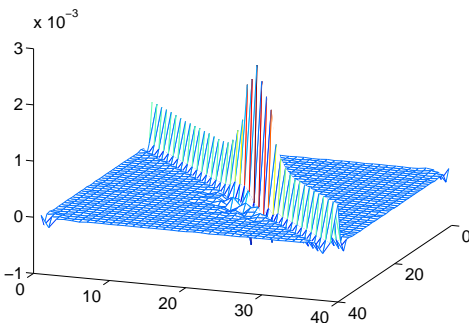


Figure C.3: \mathbf{P}^b after 40 iteration steps for Experiment 2.

First, the diagonal elements of \mathbf{P}_k^b are not identical any more; a new narrow region of elevated background errors emerges in the neighborhood of the missing observation. Interestingly, this region is centered at location 23 and not at location 20 where observations are not taken. This is due to the wave-like propagation (Fig. 3.2) of initial uncertainty toward the higher indices in the Lorenz-96 model. Since \mathbf{P}^b describes errors in 10 time-step model integrations, the error propagation leads to a shift in the location of the largest errors toward the higher indices. Secondly, the off-diagonal elements of \mathbf{P}^b describing covariances between locations near to location 20 become somewhat larger.

REFERENCES

- [1] B. D. O. Anderson and J. B. Moore. *Optimal Filtering*. Prentice-Hall, Englewood Cliffe, 1979.
- [2] J. L. Anderson. An Ensemble Adjustment Kalman Filter for Data Assimilation. *Mon. Wea. Rev.*, 129:2884–2903, 2001.
- [3] J. D. Annan and J. C. Hargreaves. Efficient parameter estimation for a highly chaotic system. *Tellus*, 56A:520–526, 2004.
- [4] P. Ashwin, J. Buescu, and I. N. Stewart. Bubbling of attractors and synchronization of chaotic oscillators. *Phys. Lett. A*, 193:126–139, 1994.
- [5] P. Ashwin, J. Buescu, and I. N. Stewart. From attractor to chaotic saddle: A tale of transverse instability. *Nonlinearity*, 9:703–737, 1996.
- [6] S. J. Baek and E. Ott. Onset of synchronization in systems of globally coupled chaotic maps. *Phys. Rev. E*, 69:066210, 2004.
- [7] S. J. Baek, B. R. Hunt, I. Szunyogh, A. Zimin, and E. Ott. Localized error bursts in estimating the state of spatiotemporal chaos. *Chaos*, 14:1042–1049, 2004.
- [8] S. J. Baek, B. R. Hunt, E. Kalnay, E. Ott, and I. Zunyogh. Local ensemble Kalman filtering in the presence of model bias. *Tellus*, 58A:293–306, 2006.
- [9] F. Baer and J. J. Tribbia. On Complete Filtering of Gravity Modes Through Nonlinear Initialization. *Mon. Wea. Rev.*, 105:1536–1539, 1977.
- [10] M. J. Bell, M. J. Martin, and N. K. Nichols. Assimilation of data into an ocean model with systematic errors near the equator. *Quart. J. Roy. Meteor. Soc.*, 130:873–893, 2004.
- [11] C. H. Bishop, B. J. Etherton, and S. J. Majumdar. Adaptive Sampling with the Ensemble Transform Kalman Filter. Part I: Theoretical Aspects. *Mon. Wea. Rev.*, 129:420–436, 2001.
- [12] J. Buck. Synchronous rhythmic flashing of fireflies. II. *Q. Rev. Biol.*, 63:265–289, 1988.

- [13] J. Buck and E. Buck. Mechanism of Rhythmic Synchronous Flashing of Fireflies. *Science*, 159:1319–1327, 1968.
- [14] G. Burgers, P. J. van Leeuwen, and G. Evensen. Analysis scheme in the ensemble kalman filter. *Mon. Wea. Rev.*, 126:1719–1724, 1998.
- [15] J. A. Carton, G. Chepurin, and X. Cao. A simple ocean data assimilation analysis of the global upper ocean 1950-95. Part I: Methodology. *J. Phys. Oceanogr.*, 30:294–309, 2000.
- [16] M. Cencini, M. Falcioni, D. Verngi, and A. Vulpiani. Macroscopic chaos in globally coupled maps. *Physica D*, 130:58–72, 1999.
- [17] G. A. Chepurin, J. A. Carton, and D. Dee. Forecast model bias correction in ocean data assimilation. *Mon. Wea. Rev.*, 133:1328–1342, 2005.
- [18] S. E. Cohn. An Introduction to Estimation Theory. *J. Meteor. Soc. Japan*, 75:257–288, 1997.
- [19] M. C. Cross and P. C. Hohenberg. Pattern formation outside of equilibrium. *Rev. Mod. Phys.*, 65:851–1112, 1993.
- [20] R. Daley. The Effect of Serially Correlated Observation and Model Error on Atmospheric Data Assimilation. *Mon. Wea. Rev.*, 120:164–177, 1992.
- [21] R. Daley. *Atmospheric Data Analysis*. Cambridge University Press, Cambridge, 1991.
- [22] S. Dano, P. G. Sorensen, and F. Hynne. Sustained oscillations in living cells. *Nature*, 402:320–322, 1999.
- [23] D. P. Dee. Bias and data assimilation. *Quart. J. Roy. Meteor. Soc.*, 131:3323–3343, 2005.
- [24] D. P. Dee and A. M. Da Silva. Data assimilation in the presence of forecast bias. *Quart. J. Roy. Meteor. Soc.*, 124:269–295, 1998.
- [25] D. P. Dee and R. Todling. Data assimilation in the presence of forecast bias: the GEOS moisture analysis. *Mon. Wea. Rev.*, 128:3268–3282, 2000.
- [26] J. C. Derber. A variational continuous assimilation technique. *Mon. Wea. Rev.*, 117:2437–2446, 1989.

- [27] J. Drécourt, H. Madsen, and D. Rosbjerg. Bias aware kalman filters: comparison and improvements. *Adv. Water Resour.*, 29:707–718, 2006.
- [28] I. R. Epstein and J. A. Pojman. *An Introduction to Nonlinear Chemical Dynamics*. Oxford University Press, Oxford, 1998.
- [29] B. J. Etherton and C. H. Bishop. Resilience of Hybrid Ensemble/3DVAR Analysis Schemes to Model Error and Ensemble Covariance Error. *Mon. Wea. Rev.*, 132:1065–1080, 2004.
- [30] G. Evensen. The Ensemble Kalman Filter: theoretical formulation and practical implementation. *Ocean Dynam.*, 53:343–367, 2003.
- [31] G. Evensen. Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *J. Geophys. Res.*, 99C5:10143–10162, 1994.
- [32] E. J. Fertig, S. J. Baek, B. R. Hunt, E. Ott, I. Szunyogh, J. A. Aravéquia, E. Kalnay, H. Li, and J. Liu. Observation bias correction with an ensemble kalman filter. *Tellus*, Submitted, 2007.
- [33] B. Friedland. Treatment of bias in recursive filtering. *IEEE Trans. Autom. Contr.*, 14:359–367, 1969.
- [34] M. Ghil, S. Cohn, Tavantzis J., K. Bube, and E. Isaacson. *Dynamic Meteorology: Data Assimilation Methods*, chapter Applications of estimation theory to numerical weather prediction, pages 139–224. Springer-Verlag, New York, 1981.
- [35] L. Glass. Synchronization and rhythmic processes in physiology. *Nature (London)*, 410:277–284, 2001.
- [36] L. Glass and M. C. Mackey. *From Clocks to Chaos*. Princeton University Press, Princeton, 1988.
- [37] C. M. Gray. Synchronous oscillations in neuronal systems: Mechanisms and functions. *J. Computat. Neurosci.*, 1:11–38, 1994.
- [38] T. M. Hamill and C. Snyder. A Hybrid Ensemble Kalman Filter-3D Variational Analysis Scheme. *Men. Wea. Rev.*, 128:2905–2919, 2000.

- [39] T. M. Hamill and J. S. Whitaker. Accounting for the Error due to Unresolved Scales in Ensemble Data Assimilation: A Comparison of Different Approaches. *Mon. Wea. Rev.*, 133:3132–3147, 2005.
- [40] T. M. Hamill, J. S. Whitaker, and C. Snyder. Distance-Dependent Filtering of Background Error Covariance Estimates in an Ensemble Kalman Filter. *Mon. Wea. Rev.*, 129:2776–2790, 2001.
- [41] J. A. Hansen. Accounting for Model Error in Ensemble-Based State Estimation and Forecasting. *Mon. Wea. Rev.*, 130:2373–2391, 2002.
- [42] F. E. Hanson, J. F. Case, E. Buck, and J. Buck. Synchrony and Flash Entrainment in a New Guinea Firefly. *Science*, 174:161–164, 1971.
- [43] J. J. Hopfield and C. D. Brody. What is a moment? Transient synchrony as a collective mechanism for spatiotemporal integration. *Proc. Natl. Acad. Sci.*, 98:1282–1287, 2001.
- [44] P. L. Houtekamer and H. L. Mitchell. A Sequential Ensemble Kalman Filter for Atmospheric Data Assimilation. *Mon. Wea. Rev.*, 129:123–137, 2001.
- [45] P. L. Houtekamer and H. L. Mitchell. Data Assimilation Using an Ensemble Kalman Filter Technique. *Mon. Wea. Rev.*, 126:796–811, 1998.
- [46] P. L. Houtekamer, H. L. Mitchell, G. Pellerin, M. Buehner, M. Charron, L. Spacek, and B. Hansen. Atmospheric Data Assimilation with an Ensemble Kalman Filter: Results with Real Observations. *Mon. Wea. Rev.*, 133:604–620, 2005.
- [47] B. R. Hunt, E. J. Kostelich, and I. Szunyogh. Efficient data assimilation for spatiotemporal chaos: A local ensemble transform kalman filter. *Physica D*, 230:112–126, 2007.
- [48] A. H. Jazwinski. *Stochastic processes and filtering theory*. Academic Press, San Diego, 1970.
- [49] E. Kalnay. *Atmospheric Modeling, Data Assimilation and Predictability*. Cambridge University Press, Cambridge, 2003.
- [50] K. Kaneko. Globally Coupled Chaos Violates the Law of Large Numbers but Not the Central-Limit-Theorem. *Phys. Rev. Lett.*, 65:1391–1394, 1990.

- [51] C. L. Keppenne, M. M. Rienecker, N. P. Kurkowski, and D. A. Adamec. Ensemble Kalman filter assimilation of temperature and altimeter data with bias correction and application to seasonal prediction. *Nonlinear Proc. Geoph.*, 12:491–503, 2005.
- [52] I. Z. Kiss, Y. Zhai, and J. L. Hudson. Emerging coherence in a population of chemical oscillators. *Science*, 296:1676–1678, 2002.
- [53] G. Kozyreff, A. G. Vladimirov, and P. Mandel. Global Coupling with Time Delay in an Array of Semiconductor Lasers. *Phys. Rev. Lett.*, 85:3809–3812, 2000.
- [54] Y. Kuramoto. *Chemical Oscillations, Waves and Turbulences*. Springer-Verlag, Berlin, 1984.
- [55] J. F. Lamarque, B. Khattatov, V. Yudin, D. P. Edwards, J. C. Gille, L. K. Emmons, M. N. Deeter, J. Warner, D. C. Ziskin, G. L. Francis, S. Ho, D. Mao, and J. Chen. Application of a bias estimator for the improved assimilation of measurements of pollution in the troposphere (MOPITT) carbon monoxide retrievals. *J. Geophys. Res.*, 109:D16304, 2004.
- [56] E. N. Lorenz and K. A. Emanuel. Optimal sites for supplementary weather observations: Simulation with a small model. *J. Atmos. Sci.*, 55:399–414, 1998.
- [57] P. Lynch. The Dolph-Chebyshev Window: A Simple Optimal Filter. *Mon. Wea. Rev.*, 125:655–660, 1997.
- [58] P. Lynch and X. Y. Huang. Initialization of the HIRLAM model using a digital filter. *Mon. Wea. Rev.*, 120:1019–1034, 1992.
- [59] B. Machenhauer. On the dynamics of gravity oscillations in a shallow water model with applications to normal mode initialization. *Contrib. Atmos. Phys.*, 50:253–271, 1977.
- [60] L. G. Margolin, E. S. Titi, and S. Wynne. The postprocessing Galerkin and nonlinear Galerkin methods : A truncation analysis point of view. *SIAM J. Numer. Anal.*, 41:695–714, 2003.
- [61] M. J. Martin, M. J. Bell, and N. K. Nichols. Estimation of systematic error in an equatorial ocean model using data assimilation. *Int. J. Numer. Meth. Fluids*, 40:435–444, 2002.

- [62] D. C. Michaels, E. P. Matyas, and J. Jalife. Mechanisms of sinoatrial pacemaker synchronization - a new hypothesis. *Circ. Res.*, 61:704–714, 1987.
- [63] F. Molteni. Atmospheric simulations using a GCM with simplified physical parametrizations. i: model climatology and variability in multi-decadal experiments. *Clim. Dynam.*, 20:175–191, 2003.
- [64] S. W. Morris, E. Bodenschatz, D. S. Cannell, and G. Ahlers. Spiral defect chaos in large aspect ratio Rayleigh-Bénard convection. *Phys. Rev. Lett.*, 71:2026–2029, 1993.
- [65] S. Nadis. Complex systems: All together now. *Nature (London)*, 421:780–782, 2003.
- [66] E. Ott. *Chaos in Dynamical Systems*, chapter 2, 10. Cambridge University Press, Cambridge, second edition, 2002.
- [67] E. Ott and J. D. Hanson. The effect of noise on the structure of strange attractors. *Phys. Lett. A*, 85:20–22, 1981.
- [68] E. Ott, W. D. Withers, and J. A. Yorke. Is the dimension of chaotic attractors invariant under coordinate changes? *J. Stat. Phys.*, 36:687–697, 1984.
- [69] E. Ott, P. So, E. Barreto, and T. Antonsen. The onset of synchronization in systems of globally coupled chaotic and periodic oscillators. *Physica D*, 173:29–51, 2002.
- [70] E. Ott, B. R. Hunt, I. Szunyogh, A. V. Zimin, E. J. Kostelich, M. Corazza, E. Kalnay, D. J. Patil, and J. A. Yorke. A local ensemble Kalman filter for atmospheric data assimilation. *Tellus*, 56A:415–428, 2004.
- [71] E. Ott, B. R. Hunt, I. Szunyogh, A. V. Zimin, E. J. Kostelich, M. Corazza, E. Kalnay, D. J. Patil, and J. A. Yorke. Estimating the state of large spatio-temporally chaotic systems. *Phys. Lett. A*, 330:365–370, 2004.
- [72] G. Perez and H. A. Cerdeira. Instabilities and nonstatistical behavior in globally coupled systems. *Phys. Rev. A*, 46:7492–7497, 1992.
- [73] A. Pikovsky, M. Rosenblum, and J. Kurths. *Synchronization: A Universal Concept in Nonlinear Sciences*. Cambridge University Press, Cambridge, 2001.

- [74] A. S. Pikovsky and J. Kurths. Do Globally Coupled Maps Really Violate the Law of Large Numbers? *Phys. Rev. Lett.*, 72:1644–1646, 1994.
- [75] A. S. Pikovsky, M. G. Rosenblum, and J. Kurths. Synchronization in a population of globally coupled chaotic oscillators. *Europhys. Lett.*, 34:165–170, 1996.
- [76] N. Pinardi and J. Woods, editors. *Ocean Forecasting: Conceptual Basis and Applications*. Springer-Verlag, Berlin, 2002.
- [77] M. I. Rabinovich and H. D. I. Abarbanel. The role of chaos in neural systems. *Neuroscience*, 87:5–14, 1998.
- [78] N. F. Rulkov. Regularization of synchronized chaotic bursts. *Phys. Rev. Lett.*, 86:183–186, 2001.
- [79] H. Sakaguchi. Phase transition in globally coupled Rössler oscillators. *Phys. Rev. E*, 61:7212–7214, 2000.
- [80] S. Seshadri, V. Balakrishnan, and S. Lakshmibala. Noise-amplitude dependence of the invariant density for noisy, fully chaotic one-dimensional maps. *Phys. Rev. E*, 60:386–390, 1999.
- [81] T. Shibata and K. Kaneko. Heterogeneity-induced order in globally coupled chaotic systems. *Europhys. Lett.*, 38:417–, 1997.
- [82] T. Shibata and K. Kaneko. Collective Chaos. *Phys. Rev. Lett.*, 81:4116–4119, 1998.
- [83] E. Sismondo. Synchronous, Alternating, and Phase-Locked Stridulation by a Tropical Katydid. *Science*, 249:55–58, 1990.
- [84] P. So, E. Ott, and W. P. Dayawansa. Observing chaos: Reducing and tracking the state of a chaotic system from limited observation. *Phys. Lett. A*, 176:421–427, 1993.
- [85] P. Stoica and R. Moses. *Introduction to Spectral Analysis*, chapter 2. Prentice-Hall Inc., Upper Saddle River, 1997.
- [86] S. Strogatz. *Sync: The Emerging Science of Spontaneous Order*. Hyperion, New York, 2003.

- [87] I. Szunyogh, E. J. Kostelich, G. Gyarmati, D. J. Patil, and B. R. Hunt. Assessing a local ensemble kalman filter: Perfect model experiments with the NCEP global model. *Tellus*, 57A:528–545, 2005.
- [88] I. Szunyogh, E. J. Kostelich, G. Gyarmati, E. Kalnay, B. R. Hunt, E. Ott, and J. A. Yorke. A local ensemble transform kalman filter data assimilation system for the NCEP global model. *Tellus*, Submitted, 2007.
- [89] D. Topaj, W. H. Kye, and A. Pikovsky. Transition to Coherence in Populations of Coupled Chaotic Oscillators: A Linear Response Approach. *Phys. Rev. Lett.*, 87:074101, 2001.
- [90] Z. Toth and M. Peña. Data assimilation and numerical forecasting with imperfect models: the mapping paradigm. *Physica D*, 230:146–158, 2007.
- [91] S. C. Venkataramani, B. R. Hunt, and E. Ott. Bubbling transition. *Phys. Rev. E*, 54:1346–1360, 1996.
- [92] S. C. Venkataramani, B. R. Hunt, E. Ott, D. J. Gauthier, and J. C. Bienfang. Transitions to Bubbling of Chaotic Systems. *Phys. Rev. Lett.*, 77:5361–5364, 1996.
- [93] T. J. Walker. Acoustic Synchrony: Two Mechanisms in Snowy Tree Cricket. *Science*, 166:891–894, 1969.
- [94] W. Wang, I. Z. Kiss, and J. L. Hudson. Experiments on arrays of globally coupled chaotic electrochemical oscillators: Synchronization and clustering. *Chaos*, 10:248–256, 2000.
- [95] J. S. Whitaker and T. M. Hamill. Ensemble Data Assimilation without Perturbed Observations. *Mon. Wea. Rev.*, 130:1913–1924, 2002.
- [96] J. S. Whitaker, G. P. Compo, X. Wei, and T. M. Hamill. Reanalysis without radiosondes using ensemble data assimilation. *Mon. Wea. Rev.*, 132:1190–1200, 2004.
- [97] K. Wiesenfeld, P. Colet, and S. H. Strogatz. Synchronization Transition in a Disordered Josephson Series Array. *Phys. Rev. Lett.*, 76:404–407, 1996.
- [98] A. T. Winfree. On Emerging Coherence. *Science*, 298:2336–2337, 2002.

- [99] S. Yamaguchi, H. Isejima, T. Matsuo, R. Okura, K. Yagita, M. Kobayachi, and H. Okamura. Synchronization of Cellular Clocks in the Suprachiasmatic Nucleus. *Science*, 302:1408–1412, 2003.
- [100] L. S. Young. Statistical properties of dynamical systems with some hyperbolicity. *Ann. Math.*, 147:585–650, 1998.
- [101] D. H. Zanette and A. S. Mikhailov. Mutual synchronization in ensembles of globally coupled neural networks. *Phys. Rev. E*, 58:872–875, 1998.
- [102] A. V. Zimin, B. R. Bunt, and E. Ott. Bifurcation scenarios for bubbling transition. *Phys. Rev. E*, 67:016204, 2003.
- [103] D. Zupanski. A General Weak Constraint Applicable to Operational 4DVAR Data Assimilation Systems. *Mon. Wea. Rev.*, 125:2274–2292, 1997.
- [104] D. Zupanski and M. Zupanski. Model error estimation employing an ensemble data assimilation approach. *Mon. Wea. Rev.*, 134:1337–1354, 2006.
- [105] M. Zupanski. Maximum likelihood ensemble filter: theoretical aspects. *Mon. Wea. Rev.*, 133:1710–1726, 2005.