REDUCED-RANK SIGMA-POINT KALMAN FILTER FOR GEOPHYSICAL DATA ASSIMILATION

by

Manoj K Kizhakkeniyil

B.Sc. (Physics, Mathematics, Chemistry), Mangalore University, 2002
M.Sc. (Physics), Goa University, 2006

DISSERTATION SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY
IN
NATURAL RESOURCES AND ENVIRONMENTAL STUDIES

UNIVERSITY OF NORTHERN BRITISH COLUMBIA

June 2015

Abstract

The main goal of my research was to develop a practical scheme for the sigma-point Kalman filter (SPKF) for its application in a realistic climate model. Large computational expense has been an obstacle to applying the SPKF to a high-dimensional system. I addressed this issue by developing an advanced SPKF data-assimilation system. My work also addressed several other factors related to the practical implementation of SPKF. The main objectives of this research were to: (i) investigate two methods to construct a reduced-rank sigma-point unscented Kalman filter (RRSPUKF); (ii) propose a localization scheme for the SPKF; and (iii) implement RRSPUKF in a realistic climate model.

I present two methods to approximate the error covariance by a reduced-rank approximation. In the first method, truncated singular-value decomposition (TSVD) is applied on the error-covariance matrix calculated in the data space (RRSPUKF(D)) while in the second method TSVD is applied on the error-covariance matrix calculated in the ensemble space (RRSPUKF(E)). The new algorithms are first tested on the Lorenz-96 model, a one-dimensional atmospheric "toy" model. The performance of both rank-reduction methods are close to that of the full-rank SPKF.

I propose a localization method for RRSPUKF(E). The results from numerical experiments on the Lorenz-96 model showed that when the localization and inflation were implemented, the optimal estimate was achieved with a finite number of sigma points.

The realistic model I used in this study was the Zebiak-Cane (ZC) model, an intermediate complexity coupled El Niño Southern Oscillation (ENSO) prediction model. The RRSPUKFs are implemented for the ZC model with the assimilation of sea surface temperature anomalies. The results showed that both RRSPUKF(D and E) were able to correctly analyze the phase and intensity of all major ENSO events during the study period with relatively similar estimation accuracy. Furthermore, the RRSPUKF was compared against ensemble square-root filter (EnSRF), showing that the overall analysis skill of RRSPUKF and EnSRF are comparable to each other but the former is more robust than the latter.

Contents

Abstract i

List of Tables vi

List of Figures vii

Glossary x

Acknowledgements xii

1 Introduction 1

1.1 Compact Overview 1

1.2 Background 2

1.3 Data assimilation 7

1.4 Motivation 12

1.5 Objectives 16

1.6 Outline 18

2 Assimilation methods and Models used in this study **21**

2.1 Introduction 21

2.2 Theoretical review 22

2.2.1 The framework of EnKF 25

2.2.2 Formulation of EnSRF 27

2.2.3 Sigma point unscented Kalman filter (SPUKF) 28
2.3 Models 31
2.3.1 The Lorenz-96 model 32
2.3.2 Zebiak-Cane model 33

3 RRSPUKFs: Derivation and application on the Lorenz-96 model . 35
3.1 Introduction 35
3.2 Reduced-rank sigma-point Kalman filter (RRSPUKF) 36
3.2.1 RRSPUKF(D): TSVD in the data space 37
3.2.2 RRSPUKF(E): TSVD in the Ensemble space 39
3.3 Application on the Lorenz-96 model 41
3.4 Experimental setup 42
3.4.1 Generation of truth and observation 43
3.4.2 Assimilation methods and procedure 43
3.4.3 Analysis error 44
3.5 Results and discussion 44
3.5.1 State estimation experiments with RRSPUKF(D) and RRSPUKF(E) 44
3.5.2 Sensitivity experiments with the RRSPUKF(E) method 49
3.6 Summary 53

4 Localization in Sigma Point Kalman Filter 57
4.1 Introduction 57
4.2 Inbreeding, filter divergence and spurious correlation 58
4.3 Covariance localization 59
4.4 Covariance inflation 60
4.5 Localization in RRSPUKF(E) 61
4.6 Numerical experiments 64

4.6.1 Experimental setup 64
4.6.2 Results and discussion 66
4.7 Summary 73

5 Application of RRSPUKF in an ENSO Model 76
5.1 Introduction 76
5.2 Model and methods 77
5.2.1 Zebiak-Cane Model LDEO5 Version 77
5.2.2 Methods and experimental setup 78
5.3 Results and discussion 83
5.3.1 Sensitivity experiments with RRSPUKF(D) 83
5.3.2 Comparison of RRSPUKF(D) and SPUKF 85
5.3.3 Forecast skill 91
5.3.4 Comparison between the RRSPUKF(D) and RRSPUKF(E) approaches 92
5.3.5 Comparison between the RRSPUKF(D) and EnSRF 96
5.4 Summary 100

6 Conclusions and Future directions 103
6.1 Introduction 103
6.2 General overview 104
6.3 Concluding summary 106
6.4 Future directions 111
6.4.1 Implementation of RRSPUKF(E) for a GCM 112
6.4.2 A hybrid assimilation by coupling 3D-Var/4D-Var + SPKF . . . 113
6.4.3 Non-Gaussian statistics in SPKF 114

Bibliography 116

List of Tables

3.1 Computation time in seconds for various data assimilation methods on a Linux PC with a 2.0GHz Intel Pentium Dual Core processor. 49

4.1 Localization experiment details. T4.2 and T4.3 are experiment numbers for the experiments to generate Table 4.2 and Table 4.3 respectively. 66

4.2 The dependence of time mean RMSE on the localization radius and number of sigma points. The inflation coefficient is 0.03. 72

4.3 The dependence of time mean RMSE on the inflation coefficient and number of sigma points. The localization radius is 6. 74

5.1 ENSO state estimation: Experiment summary 82

List of Figures

1.1 The dependence on the initial condition for the Lorenz-63 model. Here four series of X-coordinate of the Lorenz-63 model diverge markedly over time from small differences in the initial condition. 6

3.1 The time averaged relative RMSE as a function of number of sigma points (ensemble members) used for the Lorenz-96 model 46

3.2 Comparison between the true value and analysis for variable X1 with time step for (a) the full-rank SPUKF, (b) RRSPUKF(D) with 31 sigma points (ensemble members) and (c) RRSPUKF(E) with 31 sigma points. 47

3.3 The variation of the root mean-squared error over the 40 variables with time step for (a) the full-rank SPUKF, (b) RRSPUKF(D) with 31 sigma points (ensemble members), and (c) RRSPUKF(E) with 31 sigma points. 48

3.4 The time mean of the RMSE as a function of truncated modes (ensemble size) for RRSPUKF(E) method. 50

3.5 Comparison between the true value and analysis of RRSPUKF(E) for the variable X1. The analysis from the truncated modes of 3, 5, and 7 are shown in Figures 3.5a–3.5c, indicating a poor estimation skill in all three cases. 51

3.6 Comparison between the true value and analysis of RRSPUKF(E) for the vaiable X1. The analysis from the full-rank SPUKF and truncated modes of 41, 31, 21 and 11 are shown in Figures 3.6a–3.6e. 52

3.7 The variation of the RMSE over the 40 variables with time step for the different truncated modes (i.e., ensemble size) for the RRSPUKF(E) method. 53

3.8 The variance explained by the truncated mode (ensembles) with respect to the total variance for the RRSPUKF(E) method. 54

4.1 The time mean RMS error of the RRSPUKF(E) scheme as a function of number of sigma points (ensemble members). The results are shown for different size (N_x) system. 67

4.2 The time mean RMS error of the RRSPUKF(E) scheme as a function of number of sigma points (ensemble members) when localization is used. The results are shown for different size (N_x) system. 68

4.3 The time mean RMS error of the RRSPUKF(E) scheme as a function of number of sigma points (ensemble members) when both localization and inflation are used. The results are shown for different size (N_x) system. 70

4.4 The time mean RMS error of the RRSPUKF(E) scheme as a function of localization radius. The inflation parameter is 0.03. 71

4.5 The time mean RMS error of the RRSPUKF(E) scheme as a function of inflation coefficient. The localization radius is 6. 73

5.1 Graphical depiction of the Niño regions (from National Centers for Environmental Prediction (NCEP)). 82

5.2 The root mean-squared error (RMSE) of SSTA (5°S-5°N, 120-170°W) between analysis and truth as a function of number of ensembles used. 84

5.3 The RMSE of SSTA between the analysis and truth averaged over the period 1971–2000 for the RRSPUKF(D) when different numbers of truncated modes (l) are used. The number of ensembles is equal to ($2l$+1). 86

5.4 The correlation coefficient of SSTA between the analysis and truth averaged over the period 1971–2000 for the RRSPUKF(D) when different numbers of truncated modes (l) are used. The number of ensembles is equal to ($2l$+1). 87

5.5 The RMSE of the Niño-3.4 index of the analysis against the observed counterpart. The panel (a) is for RRSPUKF(D) when 41 ensembles are used and panel (b) is for the full-rank SPUKF. 88

5.6 The RMSE and correlation coefficient of SSTA between the analysis and truth for full-rank SPUKF and RRSPUKF(D) with 41 ensembles for the period 1971–2000 (a-d). Figure 5.6(e) and (f) are the difference in SSTA RMSE and correlation coefficient of SPUKF and RRSPUKF(D) where the statistically significant areas (at 95% confidence level) are shaded. 89

5.7 One-time step forecast for the Niño-3.4 region (5°S-5°N, 120-170°W). + indicates the Niño-3.4 index of the observation and the solid line indicates the Niño-3.4 index of a one month lead forecast of RRSPUKF(D) with 41 ensembles. 90

5.8 The Niño-3.4 forecast RMSE and correlation with 12-month lead time for RRSPUKF(D) for the cases with different number of ensemble members. 91

5.9 The root mean-squared error (RMSE) of SSTA (5°S-5°N, 120-170°W) between analysis and truth as a function of the number of ensembles used for the RRSPUKF(D) and RRSPUKF(E) methods. 93

5.10 The RMSE and correlation coefficient of SSTA between the analysis and truth for RRSPUKF(E) and RRSPUKF(D) with 41 ensembles (a-d). (e) and (f) are the difference in RMSE and in correlation coefficient of SSTA between RRSPUKF(E) and RRSPUKF(D). The statistically significant areas (at 95% confidence level) are shaded. 94

5.11 The explained variance (%) for the RRSPUKF(D and E) methods as a function of number of ensembles used. 95

5.12 The RMS difference between the Niño-3.4 index of the analysis against the observed counterpart for (a) RRSPUKF(D) and (b) EnSRF when 41 ensembles are used. 97

5.13 SSTA RMSE and correlation coefficient for EnSRF and RRSPUKF(D) method for the period 1971–2000 when 41 ensembles are used (a-d). . (e) and (f) are the difference in RMSE and correlation coefficient of SSTA of EnSRF and RRSPUKF(D) where the statistically significant areas (at 95% confidence level) are shaded. 98

5.14 Niño-3.4 index forecast RMSE and correlation with 12 month lead time for RRSPUKF(E), RRSPUKF(D) and EnSRF. The vertical error bars are the sampling errors at 95% confidence interval obtained using bootstrap experiments at each lead time. 99

Glossary

3D-Var	3 Dimensional Variational
4D-Var	4 Dimensional Variational
AGCM	Atmospheric General Circulation Model
BLUE	Best Linear Unbiased Estimate
EAKF	Ensemble Adjustment Kalman Filter
ENSO	El Niño Southern Oscillation
EnKF	Ensemble Kalman Filter
EnSRF	Ensemble Square-Root Filter
EKF	Extended Kalman Filter
ESSE	Error Subspace Statistical Estimation
ETKF	Ensemble Transform Kalman Filter
GCM	Global Circulation Model
LDEO5	Lamont-Doherty Earth Observatory, model version 5
LETKF	Local Ensemble Transform Kalman Filter

OGCM	Oceanic General Circulation Model
OSSE	Observation System Simulation Experiment
NCEP	National Centers for Environmental Prediction
NWP	Numerical Weather Prediction
RRSPUKF	Reduced-Rank Sigma-point Unscented Kalman Filter
RRSPUKF(D)	Reduced-Rank Sigma-point Unscented Kalman Filter - Data
RRSPUKF(E)	Reduced-Rank Sigma-point Unscented Kalman Filter - Ensemble
SIR	Sequential Important Resampling
SPCDKF	Sigma-point Central Difference Kalman Filter
SPKF	Sigma-point Kalman Filter
SPPF	Sigma-point Particle Filter
SPUKF	Sigma-point Unscented Kalman Filter
SSTA	Sea Surface Temperature anomaly
TLM	Tangent Linear Model
TSVD	Truncated Singular-Value Decomposition
SVD	Singular-Value Decomposition
TPCA	Truncated Principal-Component Analysis
ZC	Zebiak-Cane

Acknowledgements

First and foremost, I would like to express my deepest gratitude to my supervisor Dr. Youmin Tang and co-supervisor Dr. Peter Jackson for their guidance, mentoring, patience, constant encouragement, insight and knowledge imparted to me during my study period at UNBC. Their keen and vigorous academic observation enlightens me not only in this thesis but also give motivation for my future study.

I would also like to take this opportunity to acknowledge and express my gratitude to my graduate supervisory committee members Dr. Stephen Déry, Dr. Michael Gillingham and Dr. Liang Chen for their input and suggestions along the way, as well as for taking time to review this work.

This thesis greatly benefits from Youmin's courses at UNBC, from which I learned many helpful statistical methods and tools. Thanks to Dr. Peter Jackson, Dr. Ken Otter and Dr. Phil Burton for they opened my mind with interesting scientific topics of physical, biological and social sciences during my course study. I have also benefited from the courses of Dr. Peter Jackson and Dr. Stephen Déry on dynamical meteorology.

Special thanks to Dr. Ziwang Deng for his helpful comments, suggestions, and discussions in my manuscripts. Many thanks to Dr. Yanjie Cheng and Dr. Dake Chen for providing me the latest Zebiak-Cane model.

I am grateful for the numerical computation support of Dr. Jean Wang at the HPC Lab of UNBC. A special word of gratitude to all the members of my research group, Dr. Ziwang Deng, Dr. Yanjie Cheng, Jaison Ambadan, Waqar Younas, Siraj Ul Islam, Xiaoqin Yan and Dejian Yang for their friendships and discussions on research.

My deepest thanks go to my wonderful family and marvelous friends (of whom there are simply too many to mention by name) for their monumental, unwavering support and encouragement.

Finally, it would be remiss of me not to say thank you to UNBC and members of its staff for their assistance in numerous administrative and financial matters. Special thanks to UNBC Physics department for hiring me as a Teaching Assistant. I really enjoyed my teaching duties.

This work is supported by the Natural Sciences and Engineering Research Council of Canada (NSERC) Discovery program and Pacific Century Graduate Scholarship by British Columbia's Ministry of Advanced Education.

Chapter 1

Introduction

1.1 Compact Overview

The geophysical system is not a controlled laboratory experiment. Therefore models are an important tool to understand the mechanisms in geophysical systems. They describe how the components of the earth system changes using mathematical equations representing the laws of physics and data collected from observations. Advanced models, such as coupled general circulation models help to undestand many processes and mechanism of the earth's climate system. They are used to simulate the past and present climate as well as for future projections. Modern weather forecasting is based on numerical weather prediction (NWP) models. The models require current state of the atmosphere ocean called as the initial conditions to interpolate in to future. The initial conditions are obtained from observations. The weather system is chaotic and higly sensitive to initial conditions. Data assimilation helps to pro

duce high quality initial condition, optimally combining the model output with the observations. There are various schemes for data assimilation. Sigma-point Kalman filter (SPKF) is a derivative-less, deterministic Kalman filter optimal for nonlinear state-space estimation. The SPKF has many advantages over other popular data-assimilation methods such as ensemble Kalman filter (EnKF) and Ensemble square-root filter (EnSRF).

Large computational expense has been an obstacle while applying SPKF to a high-dimensional system such as the oceanic general circulation model (OGCM) or atmospheric general circulation model (AGCM). The overall goal of this dissertation is to develop advanced SPKF data-assimilation method for its application in a realistic climate model. The presentation of this chapter is organized as follows. Section 1.2 describes the broad problem domain addressed by the body of work presented in this dissertation. Section 1.3 reviews the general concepts and goals of data assimilation. This section also contain a compact literature overview of different data-assimilation methods. After this, the motivation for this study is given in Section 1.4. Objectives of my study is given in Section 1.5. Finally, an outline of the dissertation is provided in Section 1.6.

1.2 Background

The realistic model I used in this study is the LDEO5 version of the ZC model (Zebiak and Cane, 1987; Chen et al., 2004) for El Niño Southern Oscillation (ENSO) prediction. El Niño Southern Oscillation is the strongest signal in the variability of the global climate system on the interannual time-scales. It occurs in the tropical Pacific Ocean irregularly at intervals of 2–7 years and influences the global climate

through teleconnections (Wang et al., 1999). The ENSO event in 1997–1998 was the strongest in 50 years of data recording. El Niño-related impacts included heavy rains and flooding along the coast of Central America leading to extensive damage to people's lives, assets, crops and infrastructure. Droughts in many Asian countries were also associated with El Niño; it had an impact on the weather conditions around the globe through teleconnections and ocean currents. There were persistent forest fires, above normal temperature and below normal precipitation reported in Canada influenced by ENSO (Hsieh et al., 1999). The accurate prediction of extreme weather events and seasonal climatic phenomena is very important for society. The timely and more precise prediction of weather extremes will significantly minimize the loss of lives and economic impacts.

Predicting the future state of a physical-system requires a model to propagate the state of the system through time as well as knowledge of the current state of the system (Hunt et al., 2007). All models are, in general, a simplification of the real-world systems. Models are used to refer to a large number of things, from a set of hypotheses to a scaled replica of the physical-system. The main functions of models include understanding and explaining the mechanism and processes of the concerned system, predicting the future state of the system even before the observation is made and derivation of new principles. Collected observations help to improve the model structure and fine tune the model (Shenk and Franklin, 2001). Generally, models can be classified into either theoretical or statistical models depending on how they are made and their intended use. Theoretical models are usually built either before or after the data are collected, although they are mainly used as ad hoc models. The main purpose of a theoretical model is explanatory, to understand the mechanism and processes of the concerned system and also for predicting the future state of the system. Statistical models are post hoc models; they are used for the interpretation

of the data, are constructed during the analysis of the data collected and they are typically without a mechanism (Shenk and Franklin, 2001). Regression models are examples of statistical models. Theoretical models can apply to a wide range of systems and often give in-depth knowledge about the system compared to statistical models. Most of the theoretical models are expressed in the form of mathematical equations.

Earth-system models are generally theoretical models that contain mathematical equations representing the physics of the atmosphere and ocean. More often they do not have analytic solutions and are solved numerically (Jöckel, 2012). The main advantage of using a theoretical model over a statistical model is that, the underlying mechanism that the former model is based on does not change from place to place or, in other words, it is "intellectually transportable" (Shenk and Franklin, 2001). Statistical models, however, are based on empirical relationships derived from the past time-series data and the statistical relation may not hold well because the environment is so variable. For example, the El Niño event of 1982–83 could not be predicted well because at that time statistical models were used to predict the seasonal climate. Zebiak and Cane (1987) developed a theoretical model for ENSO, applying the basic equations of physics such as Newton's laws of motion, the laws of thermodynamics and the continuity equation to demonstrate the air-sea interactions in the tropical Pacific. The model demonstrated the prospects of making a long-term seasonal forecast (Webster and Palmer, 1997).

Modern weather forecasts use numerical weather prediction (NWP) models, which are built upon the general principles of fluid dynamics and thermodynamics. The earth-system models contain partial differential equations that govern the geophysical fluid systems. These partial differential equations do not have analytic solutions and they

are solved numerically (Jöckel, 2012) to obtain future states starting from an initial condition describing the current state of the system (Wang et al., 2000). Before we obtain the solution, this initial condition must be provided, which, along with the equations in the models, controls the evolution of the solution trajectory in space and time.

The sensitive dependence on the initial condition (Lorenz, 1963) can be seen in Figure 1.1 in the model solution of the Lorenz-63 model. When the Lorenz-63 model (Lorenz, 1963) is initiated from arbitrarily close alternative initial conditions (X, Y, Z), the model solutions spread out from each other after a certain number of steps. If we consider the model solution started from the initial condition (5.0, 5.0, 5.0) as the true state, then the closer the other initial conditions are to (5.0, 5.0, 5.0), the longer it takes for their solutions to spread away from the true state. This indicates the importance of preparing the initial conditions with high quality and accuracy to predict the future state of a chaotic system such as weather. Data assimilation helps to prepare high quality initial conditions, merging the information from the model as well as the observations (Wang et al., 2000; Kalnay, 2003).

Earth-system models simulate climate in detail, but at the same time they have limitations too. The uncertainties in the model prediction can arise from the inherent nonlinearity of the atmospheric system that makes the system chaotic (Lorenz, 1963, 2006), incomplete physics (e.g., cloud physics), errors in specifying the boundary condition or due to the errors in the initial condition (Stott and Kettleborough, 2002).

Both the models and observations contain errors. Some of the model limitations are errors in modelling the dynamics of the earth-system, linearization, errors in numerical approximations and errors in the parameterization of the sub-grid scale physical

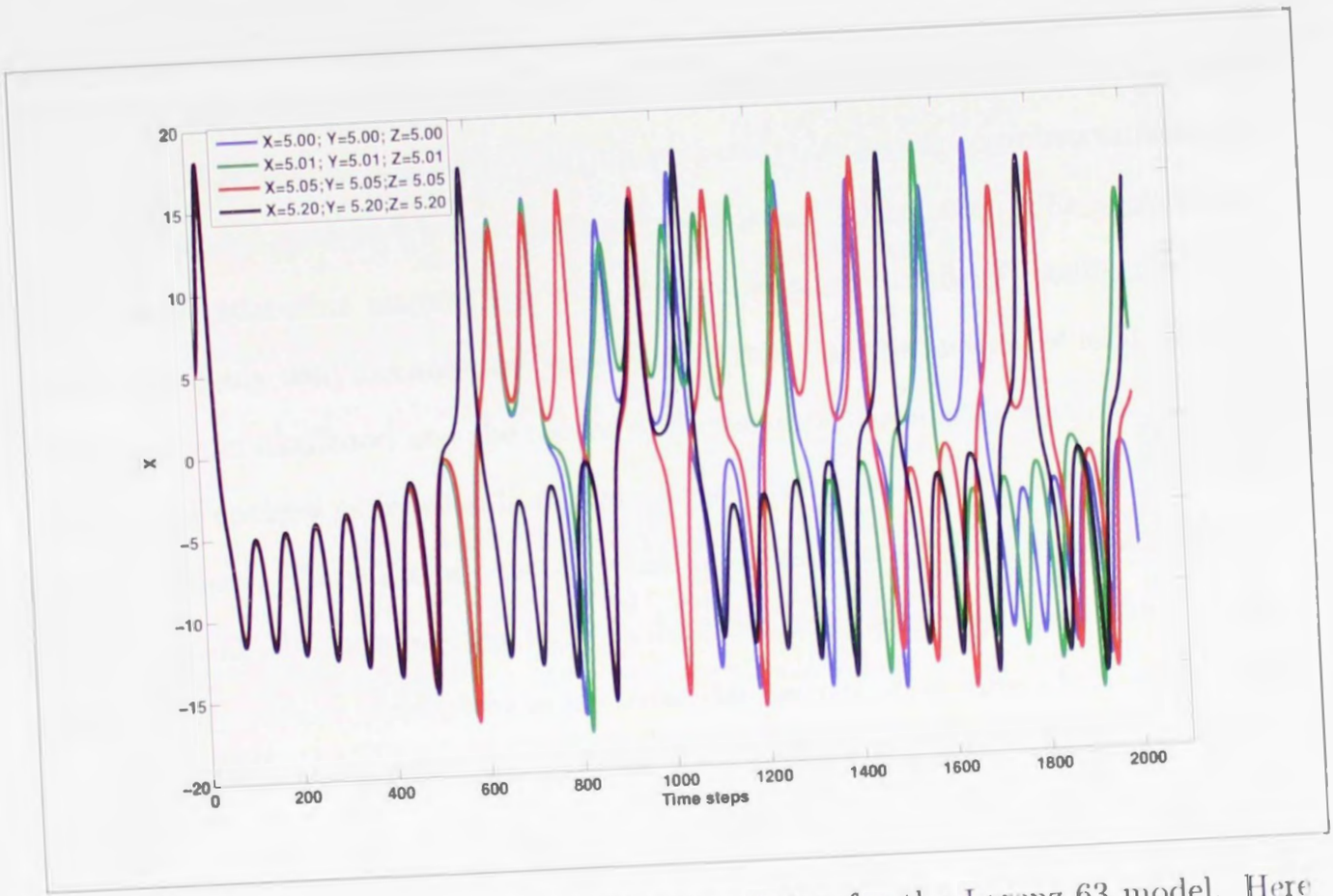

Figure 1.1: The dependence on the initial condition for the Lorenz-63 model. Here four series of X-coordinate of the Lorenz-63 model diverge markedly over time from small differences in the initial condition.

processes. For applications such as weather forecasting and climate prediction, it is not possible to have observations of all the variables at all the model grid points. The observations are limited in space and time. Moreover, the state has to be estimated from the available observations. Sources of initial-condition errors are because of fewer observations than the number of model grid points in both spatial density and temporal frequency for all the variables, errors in the data (instrumental errors and representativeness errors), errors in quality control, errors in objective analysis and missing variables (Teixeira et al., 2007). The limitations in the accurate depiction of the atmospheric state can be attributed to imperfect models, imperfect and inaccurate observations and the inherent chaotic nature of the system.

In short, because of the errors in the dynamical model and observations there is a need

to balance the information from them properly. Optimal estimation is the process of estimating the state of a system with minimal errors by combining observations and a mathematical model with their respective statistical uncertainty. The applications of optimal estimation ranges from weather forecasting to artificial intelligence. The most commonly used methods for optimal estimation are the method of least squares, the maximum likelihood and the recursive Bayesian estimation. In earth science, the purpose of optimal estimation is to provide best estimates of the state of the ocean or atmosphere. These provide better estimates than can be obtained by using solely observational data or numerical models. In earth science, optimal estimation of the state of the ocean or atmosphere is known as data assimilation. One of the main goals of data assimilation in NWP is to provide the initial conditions for weather prediction models.

1.3 Data assimilation

Data assimilation is used to optimally estimate the state of a physical-system using imperfect numerical models and noisy observations. In general, a data-assimilation system consists of three parts: a mathematical model, a set of observations and a data-assimilation method (Robinson and Lermusiaux, 2002). The assimilation of the observation into the numerical model helps to adjust the data dynamically, improving the quality of the estimated state. The mathematical state of the dynamical system are described by a set of variables known as the state variables. The dynamics link the observations with other state variables and influence them. The estimated state or analysis can be used as the initial conditions for the forecast. The main applications of data assimilation include: (i) estimating the current state of the atmospheric and

oceanic variables, which can be used as the initial condition for the future forecast; and (ii) obtaining a good trajectory of the past climate by use of all the observations available.

Numerical weather prediction is an initial-value problem. Data assimilation produces an analysis, which is a maximum-likelihood estimate of atmospheric and oceanic states and serves as optimal initial conditions to initialize the numerical models. Data assimilation can also be used for parameter estimation (O'Neill, 2004). In the early days of NWP, the computational cost to be paid for analysis was irrelevant in contrast to the cost of a 24-hour forecast. However, with advanced computational facilities, highly developed data-assimilation methods, and a vast network of observations, the cost of the computations required by the assimilation is typically the cost of a 24-hour forecast (Talagrand, 1997; Wang et al., 2000). In oceanography, data assimilation has become an important tool for studying the ocean circulation. Compared to the atmosphere we have much fewer observations in the ocean, and also the direct measurements can be difficult and expensive in the ocean. Data assimilation combines ocean models with observations, with the results, called the analysis, providing more complete and accurate information about the ocean compared to what we get from model simulations or observations alone (Talagrand, 1997; Fukumori, 2012).

Data assimilation is a mathematically rigorous process in which the optimal state of the system is estimated as a weighted combination of model forecasts and observations. The weights are determined according to the estimated uncertainties of both the model forecast and observation data. In general, the methodologies used for data assimilation can be grouped into two types: (i) the variational method (Dimet and Talagrand, 1986; Courtier et al., 1994, 1998); and (ii) the sequential method (Kalman, 1960; Kalman and Bucy, 1961; Kalnay, 2003; Auroux and Blum, 2008).

The variational method is based on optimal-control theory. The most popular method in this approach is the adjoint method. Its main objective is to minimize a cost function that measures the misfit between the temporal sequence of model states and observations (Dimet and Talagrand, 1986; Talagrand, 2012) available over a given assimilation window. The popular variational methods used in NWP are three- and four-dimensional variational assimilation, usually denoted as 3D-Var and 4D-Var, respectively. In the variational assimilation method, the information contained in the data are propagated forward in time through the nonlinear model (forecast model) and propagated backward in time though the tangent linear model (TLM) (Talagrand, 2012), which is computationally very expensive.

The 3D-Var method has been used in many operational weather-prediction centres. In 3D-Var, the forecast and observation errors are treated as time-invariant and Gaussian, described by stationary error-covariances (Talagrand, 1997). The main limitation of the 3D-Var is that it does not consider the dynamically dependent errors of the nonlinear dynamics and it assumes that all observations are made at the analysis time. The 4D-Var, considered as an advanced version of the 3D-Var, seeks the initial condition that leads to a forecast fitting best with the observations in a given assimilation window. In 4D-Var, the background state is compared to the observation at the exact time of the observation in the assimilation window and the background-covariance matrix is implicitly evolved within the given assimilation window. The minimization of the 4D-Var cost function, which measures the misfit between the temporal sequence of model states and observations, requires finding the gradient using the adjoint model (Yang et al., 2007). The main disadvantage of 4D-Var in meteorological and oceanographical applications is the complexity in the computations of the adjoint of the assimilating model and various observation operators, their validations and maintenance (Talagrand, 2012).

The sequential method is based on estimation theory, for example, the Kalman-based filters (Kalman, 1960; Kalman and Bucy, 1961; Kalnay, 2003; Auroux and Blum, 2008). Sequential-assimilation methods provide the estimate of the state of the system sequentially using a probabilistic framework, and they also provide the estimate of the uncertainty associated with the state update. Sequential-assimilation methods are easy to implement because there is no need to derive the adjoint model. These methods estimate the state recursively in two steps. The first step is the forecast step to get the prior-error statistics, and the second step is the update step that incorporates the observation information and update the prior distribution and provide the posterior-error statistics (Bertino et al., 2003). One of the main feature of the recursive estimators such as the Kalman filter is that they satisfy the Markov property, it implies that future state of the system is solely based on the present state and are independent of the past states (Ambadan, 2013).

Optimal interpolation (OI) is a fairly simple sequential data-assimilation method commonly used for operational NWP (Daley, 1991; Kalnay, 2003). Optimal interpolation is a minimum variance estimator. In OI, similar to 3D-Var, the background covariance used is constant in time and is obtained from a long-term statistical mean. One of the main disadvantages of OI is that the stationary background-covariance should represent the variability in the whole assimilation domain at all time (Bertino et al., 2003).

The Kalman filter is one of the main sequential data-assimilation methods. In Kalman filters, the goal is to find the best linear unbiased estimate (BLUE), which is assumed as the linear combinations of prior estimate and observations. When the model is linear and errors are assumed to be Gaussian, Kalman filters give the optimal solution (Bertino et al., 2003). For a weakly-nonlinear system, Kalman filters can perform

quite well but for a highly nonlinear model the solution of the Kalman Filter is not optimal (Miyoshi, 2005). The advanced derivatives of Kalman filters have been developed for their application in nonlinear estimation such as the ensemble Kalman filter (EnKF) (Evensen, 1994), which uses Monte Carlo sampling to estimate the prior-error statistics. More information about the Kalman filter and its advanced variants are given in Section 1.4 and Chapter 2.

An alternate method is using particle filters (van Leeuwen and Evensen, 1996; Anderson and Anderson, 1999; Miller et al., 1999; Kivman, 2003; van Leeuwen, 2009). The main advantage of particle filters is that they are suitable even when the dynamics are nonlinear and the error statistics are non-Gaussian. In particle filters there is no need to linearize the nonlinear model. In particle filters, the model probability-density function is represented by a discrete set of ensembles called "particles". The probability-density function is estimated by propagating the particles forward in time through the nonlinear model. When the observations are available the particles are re-sampled incorporating the observation information (van Leeuwen, 2009). In a particle filter, a finite number of particles will lead to filter degeneracy (the effective particle size reduces to a very small number of particles) and the filter fails. However, a large number of particles is computationally very expensive and this limits the application of particle filters in geophysical-systems (van Leeuwen, 2012).

Sequential data-assimilation methods based on ensemble methods are becoming more popular over the variational data-assimilation methods because of their flow-dependent (dynamically changing instead of a stationary) structure, no adjoint nor TLMs are required, their higher potential for parallelization and they are easy to code (Hamill, 2006). In the sequential algorithms the state can be updated whenever the observations are available and not required to restart the assimilation cycle.

1.4 Motivation

The Kalman filter and its derivatives have been widely applied in atmospheric and oceanic sciences in recent years (Evensen, 2003; Nerger, 2003; Robert et al., 2006; Deng and Tang, 2009; Evensen, 2009). The standard Kalman Filter (Kalman, 1960) is a linear analysis system based on Bayes' theorem and gives the BLUE making use of past observations and dynamics of the model. The application of the standard Kalman filter is not practical for NWP models because of its large number of degrees of freedom and nonlinearity (Miyoshi, 2005; Ambadan and Tang, 2009).

The extended Kalman filter (EKF) is an expansion of the Kalman filter, in which linearization is employed to approximate nonlinear systems. In the EKF, the traditional Kalman filter equations are applied after linearizing all the nonlinear models. The main drawback of EKF is the requirement of a Jacobian or the TLM for the linearization of the nonlinear models, which is extremely difficult to implement in a large-dimensional system (Julier and Uhlmann, 1997; Ambadan and Tang, 2009). In addition, the EKF retains only the first-order statistical moments in evaluating error statistics, often leading to underestimation in error covariance and subsequently to sub-optimal filtering (Evensen, 1992; Gauthier et al., 1993; Miller et al., 1994).

Instead, the EnKF, proposed by Evensen (1994), estimates the error statistics using ensembles generated by integrating the nonlinear models (Evensen, 1994; Houtekamer and Mitchell, 1998). Unlike EKF, EnKF neither requires linearization of the dynamical model nor neglects the contribution from higher-order terms to error statistics (Evensen, 1997). EnKF and its derivatives have become powerful and widely used methods within the data-assimilation community due to their algorithmic simplicity,

the relative ease of implementation as well as affordable computational cost (Anderson, 2001; Whitaker and Hamill, 2002; Evensen, 2003; Chen and Snyder, 2007; Li and Xiu, 2008). An extensive review of the literature on EnKF can be found in Evensen (2003, 2009).

To avoid the under-estimation of forecast-error covariance due to the decrease in ensemble spread, the observations are often perturbed in the standard EnKF (Burgers et al., 1998; Houtekamer and Mitchell, 1998). However, Whitaker and Hamill (2002) found that the perturbed observations can act as an additional source of sampling error, making the estimate of analysis-error covariance less accurate. They developed the ensemble square-root filter (EnSRF) to avoid the perturbation of observations, which uses the traditional Kalman gain to update the ensemble mean and the reformulated Kalman gain to update the deviation from the mean. For the same ensemble size, EnSRF can yield better results than the EnKF without any additional computational expense (Whitaker and Hamill, 2002; Tippett et al., 2003). There are different formulations of the square-root filter that have been proposed including an ensemble adjustment Kalman filter (EAKF) (Anderson, 2001), an ensemble transform Kalman filter (ETKF) (Bishop et al., 2001) apart from the EnSRF (Whitaker and Hamill, 2002). It is not clear if there is any particular formulation of the square-root filter that is better than the others (Tippett et al., 2003).

A main challenge in EnKF and EnSRF is choosing an appropriate ensemble size. A small ensemble size often results in inaccurate representation of the error-covariance matrix whereas a large ensemble size is not computationally feasible for high-dimensional systems. Therefore, some specific questions are likely to arise such as how can we generate, with a greater degree of accuracy, the finite samples for the optimal estimate of prediction-error covariance? Furthermore, what is the percentage

of the estimated error variance related to the true counterpart for a given ensemble size? The methods currently used for generating ensemble members in the standard EnKF are unable to answer these questions (Manoj et al., 2014; Tang et al., 2014a). Another concern in the traditional EnKF is its treatment of nonlinear measurement functions. As argued in Tang and Ambadan (2009) as well as Tang et al. (2014a), when the measurement function is nonlinear, the current EnKF algorithm contains an implicit assumption: the forecast of the measurement function is unbiased or the mean of the forecast equals to the forecast of the mean. The implicit assumption could cause errors in estimating the Kalman gain (Tang and Ambadan, 2009; Tang et al., 2014a).

Recently, Ambadan and Tang (2009) as well as Luo and Moroz (2009) introduced a new ensemble-based filter to the field of earth sciences using the simple Lorenz system, which is called the sigma-point Kalman filter (SPKF). The SPKFs are a derivative-less Kalman filters optimal for nonlinear state-space estimation (Julier et al., 1995; Julier and Uhlmann, 2002, 2004; Van der Merwe et al., 2004). In SPKF the error statistics and model state are analyzed using the transformation of a deterministically chosen minimal set of weighted sample points called sigma points, which captures the true mean and covariance of the prior random variable completely and the posterior mean and covariance to the second order and up to third order for Gaussian inputs (Julier et al., 1995; Julier, 2002; Van der Merwe, 2004). The SPKF family includes sigma-point unscented Kalman filter (SPUKF) based on scaled unscented transformation (Julier et al., 1995; Julier, 2002), sigma-point central difference Kalman filter (SPCDKF) based on Stirlings interpolation formulas (Nørgaard et al., 1998, 2000; Ito and Xiong, 2000) and their square-root variants (Van der Merwe and Wan, 2001; Van der Merwe, 2004). This study is only confined to the SPUKF.

A challenge in implementing SPKF for a realistic high-dimensional system is the large computational expense (Fan and You, 2009). For a high-dimensional system such as an ocean or atmosphere global circulation model (GCM), it is not computationally feasible to satisfy the requirement that the number of sigma points should be greater than twice the number of system states (Chandrasekar et al., 2008; Han and Li, 2008; Ambadan and Tang, 2009; Luo and Moroz, 2009). For implementing SPKF, the state vector is redefined through state augmentation by concatenating the model states, process noise and observation noise. This will further increase the number of sigma points and make SPKF unsuitable for realistic climate models. When the process noise is additive, the computational complexity can be reduced by using the non-augmented state vector (Van der Merwe and Wan, 2001; Fan and You, 2009). There are some reduced-order sigma-point methods using simplex sigma points (Julier, 2002, 2003), which can decrease sigma points to $L+1$ for an L dimensional system. This number is still computationally infeasible.

Recently, Ambadan and Tang (2009) suggested a subspace-projection approach for the dimensional compression by the application of truncated principal-component analysis (TPCA) on the multidimensional sigma-point space. In this method the sigma points in the principal-component space, which will retain the main features of the original sigma-point space, are used to approximate the error propagation. Luo and Moroz (2009) suggested that a subspace approach based on singular-value decomposition (SVD) can be applied on the error-covariance matrix to reduce the sigma points. Another method for rank-reduction is the Cholesky based decomposition of the error-covariance matrix and subsequent rank-reduction of the square-root matrix (Stewart, 1998; Chandrasekar et al., 2008).

The ensemble formulation of the Kalman filter and its variants has been an active area

of research in data-assimilation problems in geoscience because of its flow-dependent error structure and ease of implementation. The main disadvantage of advanced Kalman filters like EnKF and EnSRF is that selection of the initial ensemble size is not deterministic and the requirement of a greater size of ensembles for the accurate approximation of the mean and variance of the prior and posterior distribution. The SPKF is a deterministic derivative-less Kalman filter optimal for nonlinear state-space estimation problems. SPKF has been a great success in the field of optimal estimation in many areas of engineering. SPKF has a great potential in the field of oceanography and meteorology. But the large computational expense has been a main challenge in applying the SPKF to a high-dimensional system. The requirement that the number of sigma points should be greater than twice the number of system states is a disadvantage for its application in realistic climate models. This necessiates the requirement of developing a practically implementable approximation for SPKF.

1.5 Objectives

Given the computational constraints of applying a full-rank SPUKF for data-assimilation problems in oceanography and meteorology, the data-assimilation community is quite interested in the development of a reduced-rank SPUKF method for data assimilation. I address this important issue by developing an advanced SPKF data-assimilation method. Specifically, I:

1. investigate two methods to construct a reduced-rank, sigma-point unscented Kalman filter (RRSPUKF);

2. develop a localization scheme and study its impacts on the performance of

RRSPUKF; and

3. implement the RRSPUKF in a realistic climate model, which is, to my knowledge, the first trial in the data-assimilation community.

To investigate the first objective I propose two methods of rank reductions. Both methods of rank reduction employ the truncated singular-value decomposition (TSVD) to factorize the covariance matrix and reduce its rank through truncation. In the first method, SVD is applied on the error-covariance matrix calculated in the data space (RRSPUKF(D)) while in the second method the SVD is applied on the error-covariance matrix calculated in the ensemble space (RRSPUKF(E)). The reduced-rank, square-root matrix is used to select the most important sigma points that can retain the main statistical features of the original sigma points. Observation simulation experiments (or perfect model experiments) are performed using the Lorenz-96 model.

To achieve the second objective, I propose a hybrid-localization scheme for the RRSPUKF and test using the Lorenz-96 model (Lorenz and Emanuel, 1998). The experiments are conducted in the perfect model setup. Many sensitivity experiments are conducted with varying parameters to understand how and why changes in the filter parameters (ensemble size, covariance localization, and covariance inflation) affect the quality of the analysis.

The third objective is realized by implementing the RRSPUKFs to a realistic climate model. The realistic climate model employed in this study is the LDEO5 (Lamont-Doherty Earth Observatory, version 5) version of the Zebiak-Cane (ZC) model, which is an intermediate-complexity, coupled ENSO prediction model. Some important properties and the estimation accuracy of the RRSPUKF are explored. The properties

of the RRSPUKF are also compared with a full-rank SPUKF and EnSRF.

1.6 Outline

This dissertation addresses a practical application of SPKF, developing two reduced-rank approximations of SPKF and introducing a localization scheme for the SPKF. The reduced-rank SPKFs are applied to a realistic climate model. The organization of this dissertation is as follows.

Chapter 2 contains a brief theoretical review of different data-assimilation methods and the introduction to the models used in this study. In Section 2.2 the formulations of EnKF, EnSRF and SPUKF are discussed in a self contained manner. In EnKF and EnSRF, the method of generation of ensemble size is not deterministic in a rigorous statistical sense. In contrast, SPUKF presents a deterministic way of ensemble generation. Section 2.3 contains a brief introduction to the Lorenz-96 (Lorenz, 1996; Lorenz and Emanuel, 1998) model as well as the ZC model (Zebiak and Cane, 1987).

Chapter 3 addresses the first objective and discusses the formulation of the reduced-rank sigma-point unscented Kalman filters and a reduced-rank sigma-point unscented Kalman filter based on the data space (RRSPUKF(D)) is proposed and derived. Rank reduction employs the TSVD to factorize the covariance matrix and reduce its rank through truncation. Section 3.3 describes the theoretical formulation of reduced-rank sigma-point unscented Kalman filter based on the ensemble space (RRSPUKF(E)), TSVD is employed for rank reduction here as well. Before going to the application of RRSPUKFs for a realistic climate model, Section 3.4 describes the data-assimilation experiments with the Lorenz-96 model in the perfect model assumption to get a basic

idea of the performance of both the RRSPUKFs.

In Chapter 4, I propose and discuss the hybrid-localization scheme for RRSPUKF. This chapter addresses such issues as inbreeding, filter divergence and spurious correlation and how to improve the performance of the filter by employing localization and inflation schemes. The implementation details of covariance localization and covariance inflation is described in Sections 4.3 and 4.4, respectively. Section 4.5 proposes and describes the theoretical formulation of the hybrid-localization scheme for RRSPUKF(E). Section 4.6 contains the numerical experiments with the Lorenz-96 model in the perfect model assumption. After descriptions of implementation, the results of multiple sensitivity experiments are shown and compared.

Chapter 5 addresses my final objective, i.e, the application of RRSPUKFs in a realistic climate model. The realistic model used in this study is the ZC model for ENSO prediction. This chapter contains implementation details of RRSPUKF(D and E) methods in the ZC model for ENSO prediction by assimilating sea surface temperature anomaly (SSTA) data. The ZC model computes anomalies of atmospheric and oceanic fields, relative to a specified monthly mean climatology. There is detailed information about the model, experimental setup and implementation details. Four data-assimilation methods (RRSPUKF(D), RRSPUKF(E), SPUKF and EnSRF) are implemented for assimilating SSTA into ZC model. Multiple sensitivity experiments are conducted by varying the number of sigma points. Estimation accuracy and many important properties of different methods are discussed.

In Chapter 6, I summarize my results with a detailed discussion and give the conclusion of the present research. This chapter also provides some suggestions for future work. Throughout Chapters 2 through 5, I use the 1st person plural to acknowledge

the contributions of others[1] to this work

[1] I use the 1st person plural to acknowledge the contributions of my supervisor and co-supervisor to this work

Chapter 2

Assimilation methods and Models used in this study

2.1 Introduction

The purpose of this Chapter is to give a theoretical review of data assimilation methods in this dissertation as well as to give a brief introduction to the models used. The theory and derivation of a Kalman filter, EnKF, EnSRF and SPUKF are given in this chapter, and the advantages of SPUKF over EnSRF and EnKF are also discussed. The theory and derivation of EnSRF presented here is mainly based on the works by Whitaker and Hamill (2002) and Tippett et al. (2003) whereas the theory of SPUKF is mainly based on the works by Julier et al. (1995), Julier and Uhlmann (1997), Julier (2002), Van der Merwe (2004) and Van der Merwe et al. (2004c). The

information about the models used in this study is given in the last section of this Chapter.

2.2 Theoretical review

The data-assimilation problem is an estimation problem; its objective is to estimate the state of oceanic/atmospheric fields by fusing the observations with a dynamical model. Suppose X_t^b is the forecast state of a dynamical system of dimension L at time t, and Y_t is the observation at time t. We can represent the state-space equations of the dynamical system as below (wherever possible, notations are based on Ide et al. (1997)):

$$X_t^b = f(X_{t-1}^a, q_{t-1}) \tag{2.1}$$

$$Y_t = h(X_t^b, r_t) \tag{2.2}$$

where $f()$ is the nonlinear model that takes state X_{t-1}^a to X_t^b and q_{t-1} is the random model error following a Gaussian distribution with zero mean and covariance Q_{t-1}. Further $h()$ is the nonlinear observation or measurement function, which maps the model variable to observational space. Usually the dimension of X_t^b is greater than the dimension of Y_t because the observation is sparse and irregular in space and time. Furthermore, r_t is the observation noise, which is also Gaussian with mean zero and covariance R_t.

If we assume that X_t^{true} is the true state of the system then the forecast error and

forecast-error covariance can be defined as:

$$\varepsilon_t^b = X_t^{true} - X_t^b \tag{2.3}$$

$$P_t^b = \left\langle \varepsilon_t^b (\varepsilon_t^b)^T \right\rangle \tag{2.4}$$

where the $\left\langle . \right\rangle$ is the statistical expectation operator. The observation error and observation-error covariance can be denoted by:

$$r_t = Y_t - h(X_t^{true}) \tag{2.5}$$

$$R_t = \left\langle r_t (r_t)^T \right\rangle \tag{2.6}$$

Our objective is to find the updated state (called analysis) X_t^a from the forecast state X_t^b and observation Y_t. We can write X_t^a as the linear weighted mean of X_t^b and Y_t:

$$X_t^a = X_t^b + K_t(Y_t - HX_t^b) \tag{2.7}$$

where H is the linearized measurement operator and K is the weight. Analysis error and analysis-error covariance can be represented as:

$$\varepsilon_t^a = X_t^{true} - X_t^a \tag{2.8}$$

$$P_t^a = \left\langle \varepsilon_t^a (\varepsilon_t^a)^T \right\rangle \tag{2.9}$$

On solving equations (2.3), (2.4), (2.7) and (2.8), the equations for analysis error and

analysis-error covariance can be rewritten as:

$$\varepsilon_t^a \approx \varepsilon_t^b + K(r_t - H\varepsilon_t^b) \tag{2.10}$$

$$P_t^a = (I - K_t H)P_t^b(I - K_t H)^T + K_t R K_t^T \tag{2.11}$$

where I is an identity matrix. The optimal solution is obtained when the trace of P_t^a is a minimum and the optimal weight K called the Kalman gain and is defined as:

$$K_t = P_t^b H^T (H P_t^b H^T + R)^{-1} \tag{2.12}$$

On substituting the value of K in equation (2.11) we get:

$$P_t^a = (I - K_t H)P_t^b \tag{2.13}$$

In the traditional Kalman filter the forecast-error covariance is calculated using the equation:

$$P_t^b = M P_{t-1}^a M^T + Q \tag{2.14}$$

where M is the TLM of the nonlinear model. The large number of degrees of freedom and nonlinearity restrict the direct implementation of Kalman filters for GCMs. The ensemble formulation of a Kalman filter known as ensemble Kalman filter (EnKF) proposed by Evensen (1994) is a practical implementation of Kalman filter for the nonlinear system with large dimension.

2.2.1 The framework of EnKF

The formulation of EnKF is based on the idea that if the forecast model is interpreted as a stochastic differential equation, the forecast-error statistics can be approximated using ensemble integrations (Evensen, 1994, 1997). With the EnKF introduced by Evensen (1994) and Houtekamer and Mitchell (1998), ensemble members are integrated forward in time and updated (updated ensembles are known as analysis ensembles) whenever new observations are available, namely:

$$X_t^a = X_t^b + K_t(Y_t - HX_t^b) \tag{2.15}$$

$$K_t = P_t^b H^T (H P_t^b H^T + R)^{-1} \tag{2.16}$$

$$P_t^b = E[(X_t^b - \overline{X_t^b})(X_t^b - \overline{X_t^b})^T] \tag{2.17}$$

$$P_t^a = (I - K_t H) P_t^b \tag{2.18}$$

where X_t^a is the analyzed state and H is the linearized measurement operator. The superscript b indicates the model-forecast state, K_t is the Kalman gain, P_t^b is the forecast-error covariance, E[.] represents the expectation value, the overbar denotes the ensemble mean and P_t^a is the analysis-error covariance. I is an identity matrix.

Ensemble Kalman filter has been widely used in atmospheric and oceanic sciences because of its algorithmic simplicity and it is relatively easy to implement (Deng et al., 2012; Tang et al., 2014b). There are some disadvantages also for an EnKF data-assimilation system. The size of the ensemble used is a decisive factor in the performance of the EnKF. A finite ensemble size reduces the accuracy of the estimate of the statistical moments whereas a large ensemble size is computationally not afford-

able. The limited ensemble size causes the underestimation of the error covariance and noisy correlation (Houtekamer and Mitchell, 1998, 2001). To solve this problem, covariance inflation and localization is used in advanced EnKFs. More discussion about the covariance inflation (Anderson and Anderson, 1999) and localization (Houtekamer and Mitchell, 2001) methods can be found in Chapter 4.

Another disadvantage of EnKF is the assumption of the linear measurement operator (Ambadan and Tang, 2009; Tang et al., 2014a). The current EnKF algorithm contains an implicit assumption: the forecast of the measurement function is unbiased or the mean of the forecast equals to the forecast of the mean. The direct application of the nonlinear measurement function in EnKF and EnSRF imposes an implicit linearization process using ensemble members, which is approximately hold true only if:

$$\overline{h(X_{i,t}^b)} = h(\overline{X_t^b}) \tag{2.19}$$

Detailed derivation can be seen in Tang et al. (2014a) and the references there in. In many real-world situations the measurement function is nonlinear. For example the observation data are satellite radiances, but the variable required for assimilation is temperature. Linearizing the measurement function will result in estimation errors and the improvement due to the assimilation of corresponding observations may be small (Houtekamer and Mitchell, 2001). SPKF addresses this issue by using a different approach for calculating the Kalman gain and error covariances (Ambadan and Tang, 2009; Tang et al., 2014a).

2.2.2 Formulation of EnSRF

In the classic EnKF schemes, randomly perturbed observations are usually used to avoid the underestimation of the analysis-error covariance (Burgers et al., 1998; Houtekamer and Mitchell, 1998). Perturbing the observation, however, will act as another source of sampling error and decrease the accuracy of the analysis-error covariance matrix (Whitaker and Hamill, 2002). Thus an optimum approach that makes use of unperturbed observations, but avoids systematic underestimation of analysis-error covariance, was developed by Whitaker and Hamill (2002), called the ensemble square-root filter (EnSRF).

For convenience, analysis state equation (2.15) can be rewritten as the sum of mean analysis state $\overline{X_t^a}$ and the deviation from the mean state $X_t^{a'}$:

$$\overline{X_t^a} = \overline{X_t^b} + K_t(\overline{Y_t} - H\overline{X_t^b}) \tag{2.20}$$

$$X_t^{a'} = X_t^{b'} + \widetilde{K_t}(Y_t^{'} - HX_t^{b'}) \tag{2.21}$$

where K_t is the traditional Kalman gain and $\widetilde{K_t}$ is the gain used to update the deviation from the mean. For the traditional EnKF, $\widetilde{K_t} = K_t$. Here $Y_t^{'}$ is the randomly perturbed observation errors (Whitaker and Hamill, 2002; Burgers et al., 1998). In EnSRF, equation (2.20) will be written as:

$$X_t^{a'} = X_t^{b'} - \widetilde{K_t}HX_t^{b'} = (I - \widetilde{K_t}H)X_t^{b'} \tag{2.22}$$

$\widetilde{K_t}$ is chosen such that its definition will satisfy the analysis-error covariance equation

(2.18). Then the solution will be equal to:

$$\widetilde{K_t} = P_t^b H^T [\sqrt{(HP_t^b H^T + R)}^{-1}]^T \times [\sqrt{(HP_t^b H^T + R)} + \sqrt{R}]^{-1} \tag{2.23}$$

In EnSRF, observations can also be processed one at a time: in that case $HP_t^b H^T$ and R reduce to scalars. Rewriting equation (2.22), if $\widetilde{K_t} = \alpha K_t$, where α is a constant, on solving for α the solution obtained is (Whitaker and Hamill, 2002):

$$\alpha = \left(1 + \sqrt{\frac{R}{HP_t^b H^T + R}}\right)^{-1} \tag{2.24}$$

2.2.3 Sigma point unscented Kalman filter (SPUKF)

In EnKF, the ensemble members are usually generated randomly using some well-designed random perturbation schemes (Evensen, 2003). In EnKF and EnSRF, the ensemble size is usually determined considering the limitations dictated by the complexity of the assimilation system, or calculation capacity, or in some cases by sensitivity studies (Mitchell et al., 2002; Lorenc, 2003; Liang, 2007; Curn et al., 2012), which is not deterministic in a rigorous statistical sense. In contrast, SPUKF uses a deterministically chosen scheme to generate ensemble members (i.e., sigma points) by the scaled unscented transformation formula (Julier, 2002):

$$\left.\begin{aligned} \Psi_{t,0} &= \overline{X_t^a} \\ \Psi_{t,i}^{+} &= \overline{X_t^a} + [\sqrt{(L+\lambda)P_t^a}]_i \text{ , for } i = 1, 2.....L \\ \Psi_{t,i}^{-} &= \overline{X_t^a} - [\sqrt{(L+\lambda)P_t^a}]_{i-L} \text{ , for } i = L+1, ...2L \end{aligned}\right\} \tag{2.25}$$

There is a set of weights associated with these sigma points:

$$\left.\begin{aligned} w_0^m &= \frac{\lambda}{L+\lambda} \text{ , for } i = 0 \text{ and} \\ w_i^m &= \frac{1}{2(L+\lambda)} \text{ , for } i = 1, ...2L \end{aligned}\right\} \tag{2.26}$$

where the superscripts m correspond to the mean. Here $2L + 1$ ensemble members are generated according to equation (2.24). L is the dimension of augmented states concatenating the model states, process noise and observation noise. $[\sqrt{(L+\lambda)P_t^a}]_i$ denotes the i^{th} column of the weighted matrix square-root of the analysis-error covariance matrix. $\lambda = \alpha^2(L+k) - L$ is a scaling parameter. α is a small positive value between 0 and 1 and decides the spread of the sigma points around the mean state $\overline{X_t^a}$. It should ideally be small to minimize the higher-order effects when the nonlinearities are strong. k is a control parameter that guarantees positive semi-definiteness of the covariance matrix and is an important issue in implementing SPUKF. The value of k also depends on the value of α. k ensures that λ is positive, otherwise the weight of the zeroth sigma point, $w_0^m < 0$ and the non-positive weights can make the covariance non-positive semidefinite (Julier, 2002; Van der Merwe, 2004). β is a non-negative weighting term introduced to reduce the approximation error. If the state distribution is Gaussian, $\beta = 2$ is an optimal choice (Van der Merwe, 2004; Julier, 2002). In this thesis the values of $\alpha = 1$, $k = 0$ and $\beta = 2$ are used, whose rationales and the detailed derivations can be found in Van der Merwe (2004) and references therein.

SPUKF makes use of the following reformulated equation to obtain the Kalman gain K_t and analysis-error covariance matrix P_t^a so that it does not require a TLM or

linearized measurement operator:

$$K_t = P_t^{xy}(P_t^{yy})^{-1} \tag{2.27}$$

$$P_t^a = P_t^b - K_t P_t^{yy}(K_t)^T \tag{2.28}$$

P_t^{xy} is the cross-covariance matrix of the state and observation prediction error. P_t^{yy} is the covariance matrix of the observation prediction error. For a detailed derivation see Van der Merwe (2004) and Ambadan and Tang (2009) and references therein.

When the new observation arrives the analysis-error covariances are updated using equation (2.27) and states are updated using the following equation:

$$\overline{X_t^a} = \overline{X_t^b} + K_t(\overline{Y_t} - H(\overline{X_t^b})) \tag{2.29}$$

At this step one can generate new sigma points by:

$$\Psi_t = [\Psi_{t,0}, \Psi_{t,i}^+, \Psi_{t,i}^-] \text{ , for } i = 1, 2.....L \tag{2.30}$$

The sigma-point vector is propagated through equations (2.1) and (2.2). The mean and covariance are calculated using the following formulas:

$$\overline{X_t^b} = \sum_{i=0}^{i=2L} w_i^m X_{t,i}^b \tag{2.31}$$

$$\overline{Y_t^b} = \sum_{i=0}^{i=2L} w_i^m Y_{t,i}^b \tag{2.32}$$

$$P_t^b = \sum_{i=0}^{i=2L} w_i^c (X_{t,i}^b - \overline{X_t^b})(X_{t,i}^b - \overline{X_t^b})^T \tag{2.33}$$

The cross covariance P_t^{xy} and projection covariance P_t^{yy} are evaluated by the following equations:

$$P_t^{xy} = \sum_{i=0}^{i=2L} w_i^c (X_{t,i}^b - \overline{X_t^b})(Y_{t,i}^b - \overline{Y_t^b})^T \tag{2.34}$$

$$P_t^{yy} = \sum_{i=0}^{i=2L} w_i^c (Y_{t,i}^b - \overline{Y_t^b})(Y_{t,i}^b - \overline{Y_t^b})^T \tag{2.35}$$

The weights associated with the covariance calculation is:

$$\begin{aligned} w_0^c &= \frac{\lambda}{L+\lambda} + (1 - \alpha^2 + \beta) \text{ , for } i = 0 \text{ and} \\ w_i^c &= \frac{1}{2(L+\lambda)} \text{ , for } i = 1, ...2L \end{aligned} \tag{2.36}$$

Using the SPUKF, the second order accuracy can be preserved in the mean and covariance. Higher-order information of the system in the covariance can be included by using the scaling parameters α and β (Julier, 2002; Julier and Uhlmann, 2004).

2.3 Models

In this study, we use the Lorenz-96 model (Lorenz, 1996; Lorenz and Emanuel, 1998) and the Zebiak-Cane model (Chen et al., 2004), hereafter ZC model). A brief introduction to the models is given in the Subsections 2.3.1 and 2.3.2. A more detailed description about the Lorenz-96 model can be found in Lorenz (1996) and Lorenz and Emanuel (1998); more detail about the ZC model can be found in Zebiak and Cane (1987).

2.3.1 The Lorenz-96 model

The Lorenz-96 model (Lorenz, 1996; Lorenz and Emanuel, 1998) is a highly-nonlinear model often used as a test-bed for new data-assimilation algorithms. This model has many common aspects with atmospheric models even though it is not meteorologically very realistic. One important property is that it behaves chaotically.

The model equation is defined as:

$$\frac{dx_i}{dt} = x_{i-1}(x_{i+1} - x_{i-2}) - x_i + F \tag{2.37}$$

where $i = 1, 2, ...Nx$, represents closed cyclic boundary conditions. The x_i's could be loosely treated as an atmospheric variable at discrete longitudes around a constant latitude circle such that $x_0 = x_{Nx}, x_{-1} = x_{Nx-1}, x_{Nx+1} = x_1$. This sytem is used because the dimension is small and could be easily scaled and also because of its similarities in chaotic behaviour and timescale of predictability as found in realistic weather models. The term F is an external forcing and the model behaves chaotically when $F = 8$. The Runge-Kutta fourth-order scheme is used to solve the model equation with time step equal to 0.05. Performing data assimilation every time step corresponds approximately to 6-hours in a global weather model (Lorenz and Emanuel, 1998). A more detailed discussion of the model and its characteristics can be found in Lorenz (1996) and Lorenz and Emanuel (1998). The experiments are conducted in the perfect model set up: that is, a long-term integration from an arbitrary initial condition is used as the true state and synthetic observation is generated by adding normal random errors to the true state.

2.3.2 Zebiak-Cane model

The LDEO5 version of the ZC model (Zebiak and Cane, 1987; Chen et al., 2004) is used as the realistic climate model in this study. The ZC model is a coupled ocean-atmosphere model and has been widely used for predicting the timing, phase and intensity of ENSO events for both experimental and operational purposes since the late 1980s (Webster and Palmer, 1997; Karspeck and Anderson, 2007). This model has been used by Chen et al. (2004) for the successful long-term retrospective prediction of ENSO for 148 years (1856–2003). This was the first dynamical coupled model developed to predict ENSO. This model has components of both discharge-recharge and delayed oscillator physics of ENSO.

The ZC model is an anomaly model that computes anomalies of atmospheric and oceanic fields, relative to a specified monthly mean climatology. The atmospheric dynamics follow Gill (1980), with a steady-state, linear, shallow-water equations on an equatorial beta plane. The circulation is forced by a heating anomaly that depends on the SST anomaly and moisture convergence parameterized in terms of surface wind convergence. The model domain is confined to the tropical Pacific Ocean (101.25°E–73.125°W, 29°S–29°N). The grid resolutions for ocean dynamics is 2° longitude by 0.5° latitude and for SST physics and the atmosphere the model grid is 5.625° longitude by 2° latitude. The ocean dynamics use the reduced-gravity model. The surface currents are generated by spinning-up the model with monthly mean climatological winds and is then used for the current anomaly calculations. The thermodynamics describe the SST anomaly and heat-flux change. The governing equation of thermodynamics contain three-dimensional temperature advection by the calculated anomalous current and the specified mean current (Cane et al., 1986; Zebiak and Cane, 1987). The

governing equation of thermodynamics in the ZC model is given by:

$$\frac{\partial T}{\partial t} = -\overline{U}.\nabla T - U.\nabla(\overline{T}+T) - [M(\overline{W}+W) - M(\overline{W})]\frac{\partial \overline{T}}{\partial Z} - M(\overline{W}+W)\frac{\partial T}{\partial Z} - \alpha T \tag{2.38}$$

where the barred and unbarred quantities represent mean and anomalies respectively. U and W represent horizontal surface currents and upwelling velocity respectively and α is a thermal damping coefficient. The horizontal advection terms are represented by the first two terms on the right hand side of the above equation. The effects of anomalous upwelling in the presence of the mean vertical temperature gradient, $\frac{\partial \overline{T}}{\partial Z}$, and the total upwelling in the presence of the anomalous vertical temperature gradient, $\frac{\partial T}{\partial Z}$, are represented in the third and fourth terms respectively. The last term is a linear damping term that represents the change of SSTA due to the heat exchange between the ocean and atmosphere (Zebiak and Cane, 1987; Battisti, 1988; Cheng et al., 2010). The function $M(x)$ is defined by:

$$M(x) = \begin{cases} 0, & x \leq 0 \\ x, & x > 0 \end{cases} \tag{2.39}$$

The function $M(x)$ ensures that surface temperature is affected by vertical advection only in the presence of upwelling (Zebiak and Cane, 1987).

The surface wind-stress anomaly is generated from the background mean winds specified by the observation and surface wind anomalies produced by the atmospheric model. The ocean dynamics time step is 10 days. Further details about the ZC model can be found in Cane et al. (1986) and Zebiak and Cane (1987).

Chapter 3

RRSPUKFs: Derivation and application on the Lorenz-96 model

3.1 Introduction

This Chapter constitutes some of the main theoretical contributions of this dissertation. The advantages of the the SPUKF algorithm compared to EnKF and EnSRF in estimating the true mean and covariance have already been reviewed in the previous chapters. In SPUKF, the sigma points are generated deterministically by the scaled unscented transformation formula (Julier, 2002). The main limitation to apply the SPUKF to a higher-dimensional system such as NWP is the requirement of more than twice the number of sigma points than the system states. If the system dimension is L, there is a requirement of $2L + 1$ sigma points for estimating the mean and co-

variance, and that is computationally unfeasible. One of the possible solutions is to investigate if the number of sigma points can be lowered by taking the reduced-rank approximation of the full SPUKF, which can retain most of the main features of the original sigma points. The possibility of constructing a reduced-rank approximation is investigated in the Section 3.2.

3.2 Reduced-rank sigma-point Kalman filter (RRSPUKF)

The reduced-rank approximation is mainly based on the concept that most large-scale geophysical phenomena can be approximated by a finite number of degrees of freedom and their dominant variability can be explained by a limited number of modes (Temam, 1991; Lermusiaux, 1997; Lermusiaux and Robinson, 1999; Ambadan and Tang, 2009). Ideally these modes should be able to describe the evolving dynamics of the system. Common techniques used to reduce the number of modes to a limited number that is still able to represent the dominant features of the system are dynamical singular vectors (Palmer and Zanna, 2013), empirical orthogonal functions (Lorenz, 1965), principal oscillation and interaction patterns (Penland, 1989) and radial functions and wavelets (Lermusiaux, 1997; Lermusiaux and Robinson, 1999). Lermusiaux and Robinson (1999) proposed the concept of error-subspace statistical estimation (ESSE) to find the dominant error subspace, described by error-subspace singular vectors and values. In the following section a similar concept is applied in the case of SPUKF, assuming that most important errors of the original sigma-point space can be estimated using a fewer number of the dominant sigma points (Manoj

et al., 2014).

As can be seen in the equation below, in SPUKF, for the complete description of all the errors we need to use the number of sigma points determined by the dimension of the error-covariance matrix, P_t^a:

$$\left.\begin{aligned} \Psi_{t,0} &= \overline{X_t^a} \\ \Psi_{t,i}^{+} &= \overline{X_t^a} + [\sqrt{(L+\lambda)P_t^a}]_i \text{ , for } i = 1, 2.....L \\ \Psi_{t,i}^{-} &= \overline{X_t^a} - [\sqrt{(L+\lambda)P_t^a}]_{i-L} \text{ , for } i = L+1, ...2L \end{aligned}\right\} \tag{3.1}$$

by finding the reduced-rank approximation of P_t^a, the number of sigma points can be reduced. Two methods are proposed for the rank reduction and both are based on the TSVD. The formulation of both methods is described in Section 3.2.1.

3.2.1 RRSPUKF(D): TSVD in the data space

In the RRSPUKF(D) method, TSVD is applied on the covariance matrix constructed in the data space, so the filter will be denoted as reduced-rank, sigma-point unscented Kalman filter (Data) (RRSPUKF(D)). As can be seen in equation (3.1), the number of sigma points is determined by the dimension of the error-covariance matrix. Thus, the central point of RRSPUK(D) is to effectively reduce the rank of the error-covariance matrix, which will be achieved by the TSVD method. The analysis-error covariance P_t^a in (2.15) is symmetric and can be decomposed as:

$$P_t^a = E_{t,L}^a D_{t,L}^a (E_{t,L}^a)^T \tag{3.2}$$

where $D_t^a = diag(\sigma_{t,1}^2 ... \sigma_{t,L}^2)$ is a diagonal matrix consisting of the eigenvalues $\sigma_{t,i}^2$, sorted in descending order, i.e., $\sigma_{t,i}^2 > \sigma_{t,j}^2 \geq 0$ for $i > j$; $E_t^a = [e_{t,1}, ...e_{t,L}]$ is the matrix consisting of the corresponding eigenvectors $e_{t,i}$.

To reduce the number of sigma points, a reduced-rank, error-covariance matrix, denoted by $\widetilde{P_t^a}$, can be approximately obtained by the truncation of equation (3.1), namely:

$$\widetilde{P_t^a} = E_{t,l}^a D_{t,l}^a (E_{t,l}^a)^T \tag{3.3}$$

where l is the truncation number. The new sigma points are generated as follows:

$$\Psi_{t,0} = \overline{X_t^a} \tag{3.4}$$

$$\Psi_{t,i}^+ = \overline{X_t^a} + [\sqrt{(l+\lambda)}]\sigma_{t,i} e_{t,i} \text{ , for } i = 1, 2...l \tag{3.5}$$

$$\Psi_{t,i}^- = \overline{X_t^a} - [\sqrt{(l+\lambda)}]\sigma_{t,i-l} e_{t,i-l} \text{ , for } i = l+1, ...2l \tag{3.6}$$

where σ is the square root of the eigen value and e is the eigen vector.

The new sigma point vector now becomes:

$$\Psi_{t+1} = [\Psi_{t+1,0}, ..., \Psi_{t+1,2l}] \tag{3.7}$$

The number of sigma-points generated by equations (3.4 to 3.6) are $2l + 1$.

3.2.2 RRSPUKF(E): TSVD in the Ensemble space

In the RRSPUKF(E) method, TSVD is applied on the covariance matrix constructed in the ensemble subspace (Ensemble), denoted as RRSPUKF(E). The RRSPUKF(D) can effectively save computational time through reducing the sigma points from $2L+1$ to $2l+1$. The saving is especially significant when the error covariance P_t^a can be approximated well by a few leading modes so that a small l can have a good truncation accuracy. However, a large challenge in RRSPUKF(D) is the computation of P_t^a itself. When the state dimension is large, it is computationally expensive to compute such a big matrix and it is even more difficult to store.

Another approach to seek the most important sigma points, which can avoid the aforementioned challenge, is through TSVD in the ensemble space (Manoj et al., 2014; Tang et al., 2014b). When the state dimension L_M is much greater than the ensemble dimension n, i.e $L_M \gg n$, it is possible to compute the SVD on the nXn covariance matrix (von Storch and Hannoschock, 1984; Wilks, 2011), which is computationally affordable for a system with large dimension. The computational complexity of the algorithm to find the covariance matrix in the data space is $O(L_M^2 \times n)$; however, the mathematical complexity of the algorithm to find the covariance matrix in the ensemble subspace is $O(n^2 \times L_M)$. The difference in computational complexity between the two algorithms is significant when $L_M \gg n$, that is usually the case for a realistic NWP model. One of the main advantage of RRSPUKF(E) over RRSPUKF(D) is that in RRSPUKF(E) we can calculate the analysis-error covariance explicitly, which is a requirement for generating the sigma points, even for a large-dimensional model.

In the RRSPUKF(E) method all the ensembles are updated according to equation (2.3) and then from these members find the analysis-error covariance matrix P_t^{aE} in

the ensemble subspace according to:

$$P_t^{aE} = E[(X_t^a - \overline{X_t^a})^T (X_t^a - \overline{X_t^a})] \tag{3.8}$$

where $\overline{X_t^a}$ is the analysis mean. Then SVD is performed on the nXn analysis-error covariance matrix:

$$P_t^{aE} = F_t^a G_t^a (F_t^a)^T \tag{3.9}$$

where F_t^a is the eigenvectors and G_t^a is a diagonal matrix of the eigenvalues.

The eigenvectors of P_t^{aE} and P_t^a are different, but the leading n eigenvectors of P_t^a can be computed from the eigenvectors F_t^a of P_t^{aE} (von Storch and Hannoschock, 1984; Wilks, 2011) using:

$$E_{t,l}^a = \frac{(X_t^a - \overline{X_t^a})^T * F_{t,l}^a}{\|(X_t^a - \overline{X_t^a})^T * F_{t,l}^a\|} \tag{3.10}$$

where $l = 1..n$. The role of the denominator is to make sure that the resulting E_t^a have unit length. The maximum value of l is $l = (n-1)/2$, so that the total number of ensemble members is n for every time step. The new sigma points are generated in the same way as in equation (3.4 to 3.6), i.e. the mean and $(n-1)/2$ of positive and negative pairs of perturbation to the mean.

3.3 Application on the Lorenz-96 model

In this section, the practicality of using the reduced-rank sigma-point Kalman filters algorithm as an effective data-assimilation method for highly nonlinear models is explored. The highly nonlinear model used in this study is the Lorenz (1996) model. The model equation is given by:

$$\frac{dx_i}{dt} = x_{i-1}(x_{i+1} - x_{i-2}) - x_i + F \tag{3.11}$$

where $i = 1, 2, ...N_x$. This model simulates the time evolution of an unspecified scalar atmospheric variable, x, at N_x equidistant grid points around a constant latitude circle. When the value of external forcing $F = 8$, the system behaves chaotically (Lorenz and Emanuel, 1998). The fourth-order Runge-Kutta scheme is used for the time integration of the model with time step = 0.05.

In the field of data assimilation, this model, along with the Lorenz (1963) three-variable model, is often used as a test-bed for examining the properties of various data-assimilation methods (Miller et al., 1994; Ambadan and Tang, 2009; Kalnay et al., 2012) not only because of its smaller degrees of freedom, allowing various experiments, but also because it includes many dynamic features of realistic weather systems (Lorenz, 2006). For example, it has error growth characteristics similar to those of full NWP models. A more detailed description about the model and its characteristics can be found in Lorenz and Emanuel (1998).

One of the reasons for using the Lorenz-96 model for data assimilation experiments is that, if new assimilation algorithms give unfavourable results for a highly idealized

model such as the Lorenz-96 model, it is highly unlikely for the same method to get reasonable results for a high-dimensional realistic weather model. A full-scale experiment using a high-dimensional model is a time consuming and vast undertaking. The experiment results from simple models can give some directions on general trends and possible behaviours.

3.4 Experimental setup

We perform the data assimilation experiments under the observing system simulation experiments (OSSE), also known as the perfect model experiments. OSSEs are generally designed to test the performance of different data assimilation algorithms. They may also be used to assess the potential impact of an observation array to be deployed using data-assimilation methods (Masutani et al., 2010; Houtekamer, 2012). In an OSSE, the model trajectory is created by a long integration of the model forward in time from a known state. The model trajectory is known as the "truth". Synthetic observations are created by sampling the model trajectory at a specified location and interval using an observation forward operator (for a real model the frequency and locations of sampled observations should be identical to that of an actual observing system). In the next step, these synthetic observations are assimilated into the model using the state of the art data-assimilation method and the resultant product is known as the analysis. The assimilation performance can be evaluated comparing the analysis with the "truth".

3.4.1 Generation of truth and observation

The experimental setup is similar to that of Lorenz (2006), where N_x=40 and F, the magnitude of the forcing, is set to 8, in which the system is chaotic. The nature or observation simulation run is created by integrating the system over 4000 time steps by using the fourth-order Runge-Kutta scheme. The integration step is set to 0.05 (i.e., 6-hours). The data from the nature run in the time steps 1000–2000 are taken as the truth. The observational data sets are generated by adding normally distributed noise $N(0, \sqrt{2})$ to the truth.

3.4.2 Assimilation methods and procedure

We use three data-assimilation methods (SPUKF, RRSPUKF(D), RRSPUKF(E)). To implement the SPUKF method, the state vector is redefined as the concatenation of the model states, model errors, and measurement errors as discussed in Section 2.2.3. For a full-rank, sigma-point space, there is a total of 241 sigma points or 241 ensemble members. We assume that the model and observation errors are uncorrelated in both space and time and that the observations of all model states are observed and assimilated every five time steps. Because there is no general method of setting the model error, the magnitude of model error used in the Kalman filter is often determined experimentally by trial and error or by statistical methods such as the Monte Carlo method (Miller et al., 1994). In our experiments, the model errors are intentionally designed in such a way that the model would not drift too far from the true state.

3.4.3 Analysis error

The quality of the analysis is measured in terms of the root mean-square error (RMSE) and time averaged relative RMS error (RRMSE) between the analysis ensemble mean and the true state:

$$RMSE = \left[\frac{1}{N_x}\sum_{n=1}^{N_x}(X_t^a - X_t^{tr})^2\right]^{\frac{1}{2}} \tag{3.12}$$

Here N_x is number of the model dimension (N_x=40 in this case):

$$RRMSE = \frac{1}{T_{max}}\sum_{t=1}^{T_{max}}\frac{\|X_t^a - X_t^{tr}\|_2}{\|X_t^{tr}\|} \tag{3.13}$$

where T_{max} is the number of assimilation cycle and $\|.\|_2$ is the L2 norm.

3.5 Results and discussion

3.5.1 State estimation experiments with RRSPUKF(D) and RRSPUKF(E)

One important element in the application of the RRSPUKF method is the size of the truncated modes of TSVD, or effective sigma points. A small size fails to characterize important error statistics and leads to poor assimilation analysis, whereas a large size is likely computationally unaffordable. A good strategy for determining the size of the

truncated modes is sensitivity experiments, which are often used in EnKF to examine the analysis error as a function of the ensemble size (Mitchell et al., 2002; Lorenc, 2003; Liang, 2007; Curn et al., 2012). Experiments are conducted using different truncated modes after the SVD using both the RRSPUKF(D) and RRSPUKF(E) methods and the results are compared with that of a full-rank SPUKF.

Figure 3.1 shows the variation of RRMSE as a function of the number of sigma points used. As can be seen in Figure 3.1, below 21 sigma points the estimation skill is very poor for both schemes but as the number of sigma points increases the estimation error decreases. When there are 31 sigma points or more, further improvement is minimal suggesting that 31 sigma points could be used for assimilation, balancing the accuracy and computational cost.

Figure 3.2(a) shows the assimilation solution for the variablr X1 of the Lorenz-96 model for the full-rank SPUKF compared with the truth with time step t . In the second case, the RRSPUKF(D) is used to reduce the sigma points. In this case, 31 sigma points are selected that account for more than 83% of the total variance (Figure 3.2(b)). The model states can be fairly well estimated by the RRSPUKF(D), although its RMSE is not as small as the full SPKF (Figure 3.3). In the third case, the RRSPUKF(E) algorithm is used to reduce the number of sigma points. As in the second case, 31 sigma points are chosen. Shown in Figure 3.2(c) is the simulation of variable X1 from the RRSPUKF(E) algorithm compared to the truth. Figure 3.3 compares the estimation error (RMSE) in the cases of (a) Full-rank SPUKF, (b) RRSPUKF(D) with 31 sigma points, and (c) RRSPUKF(E) with 31 sigma points. In the case of both RRSPUKF(D) and RRSPUKF(E), the RMSE is very close to that of a full-rank SPUKF. As the full-rank SPUKF contain more number of sigma points it takes longer time to reach the equilibrium value. Comparison between Figures 3.3(c)

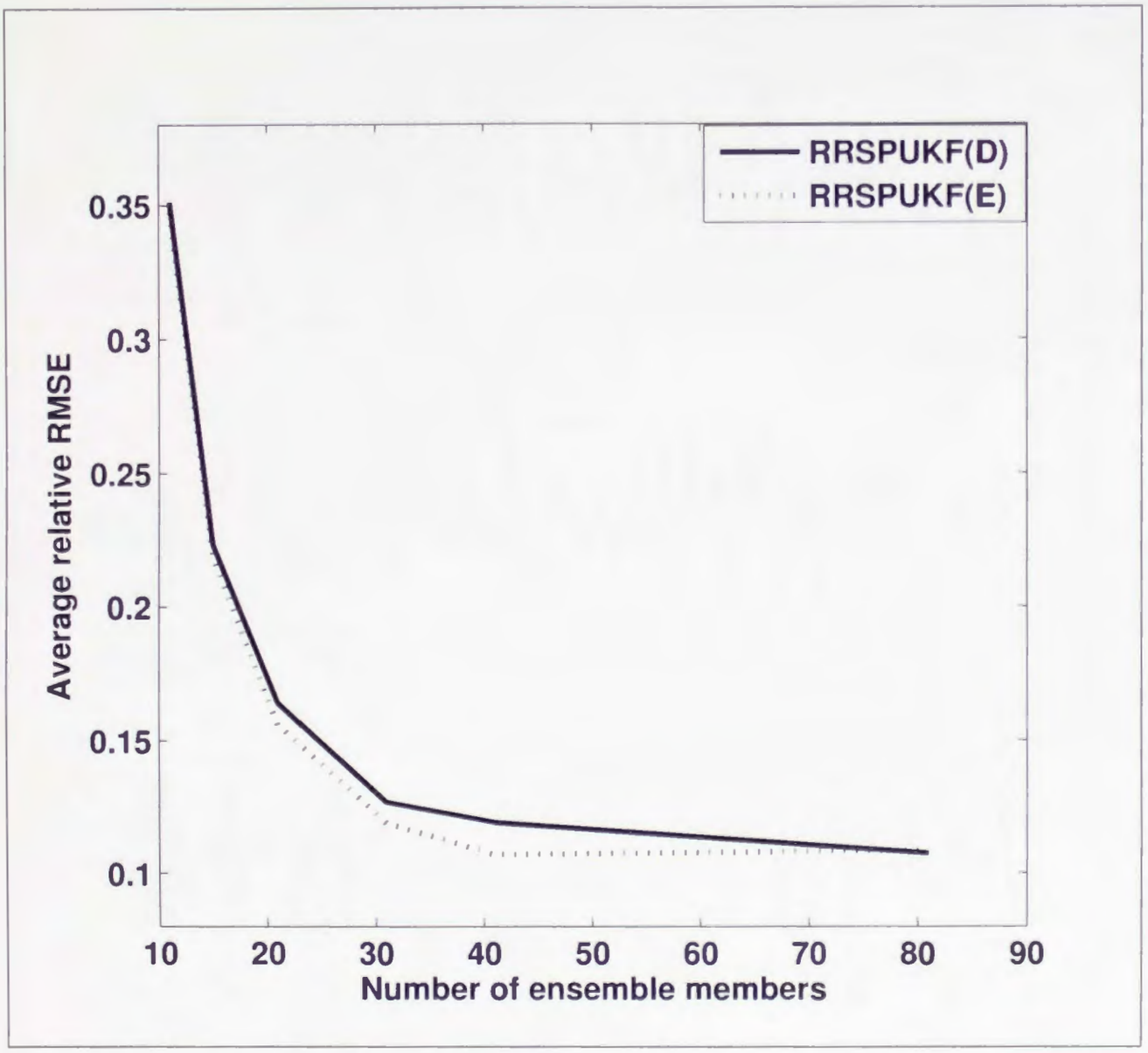

Figure 3.1: The time averaged relative RMSE as a function of number of sigma points (ensemble members) used for the Lorenz-96 model.

and 3.3(b) reveals that the RRSPUKF(E) is comparable with the RRSPUKF(D) with the same approximate error magnitude, suggesting a possible solution to applying SPUKF in high-dimensional systems. More detailed discussion about the RRSPUKF(E) is given in Section 3.5.2.

Computation time for various methods using a Linux PC with a 2.0GHz Intel Pentium Dual Core processor is shown in Table 3.1. The computation time is for 1000 cycles of analysis-forecast steps on the Lorenz-96 model. This computation timing results from

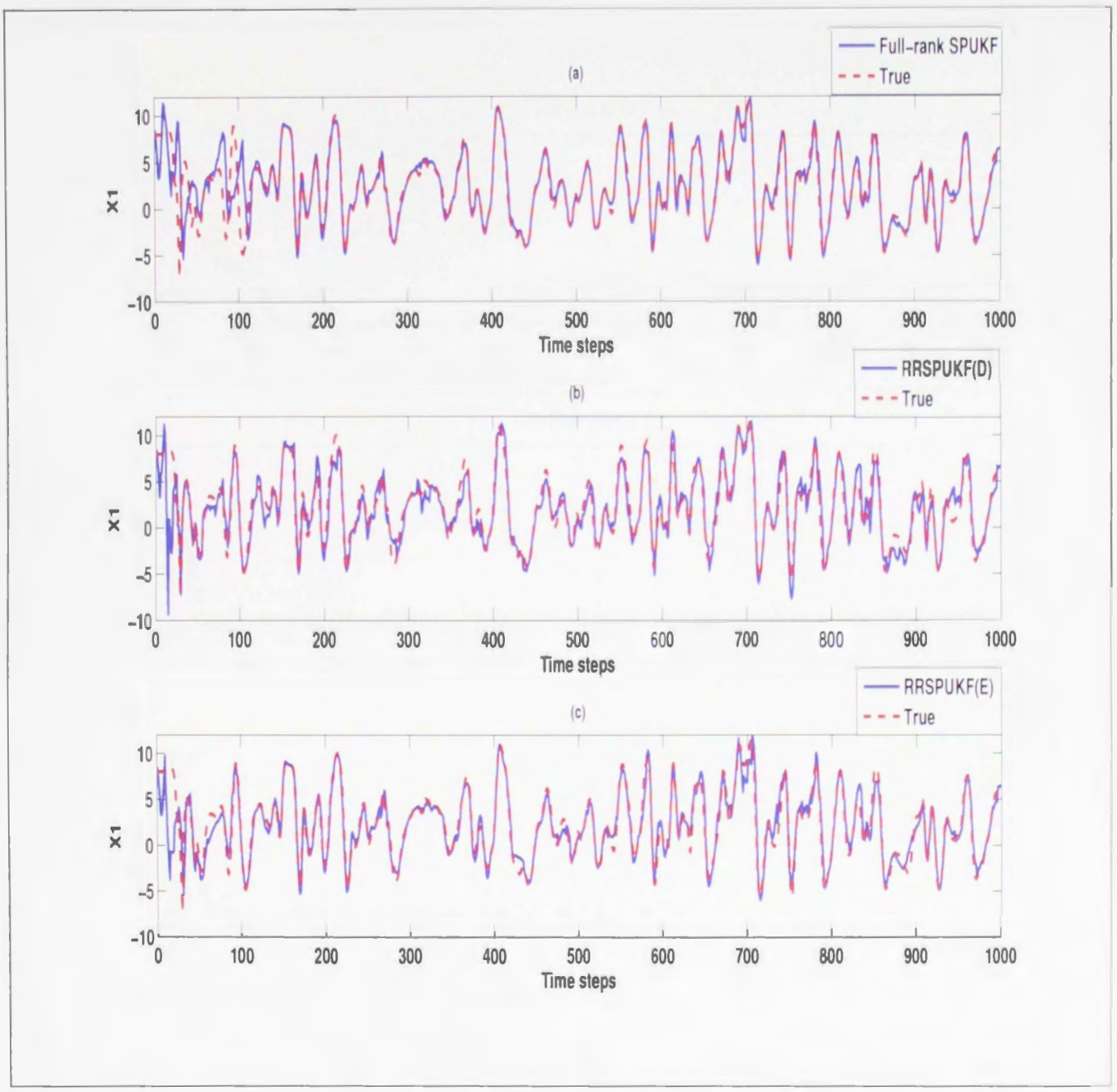

Figure 3.2: Comparison between the true value and analysis for variable X1 with time step for (a) the full-rank SPUKF, (b) RRSPUKF(D) with 31 sigma points (ensemble members) and (c) RRSPUKF(E) with 31 sigma points.

the experiments on Lorenz-96 model give a general trend of how RRSPUKF methods perform compared to the full-rank SPUKF. RRSPUKF(E) and (D) are much faster than the Full-Rank SPUKF. The timing of RRSPUKF(E) is slightly better than the RRSPUKF(D). The timing results from both the RRSPUKFs are comparable because of the low-dimensionality of the Loren-96 model. For a realistic atmospheric or oceanic

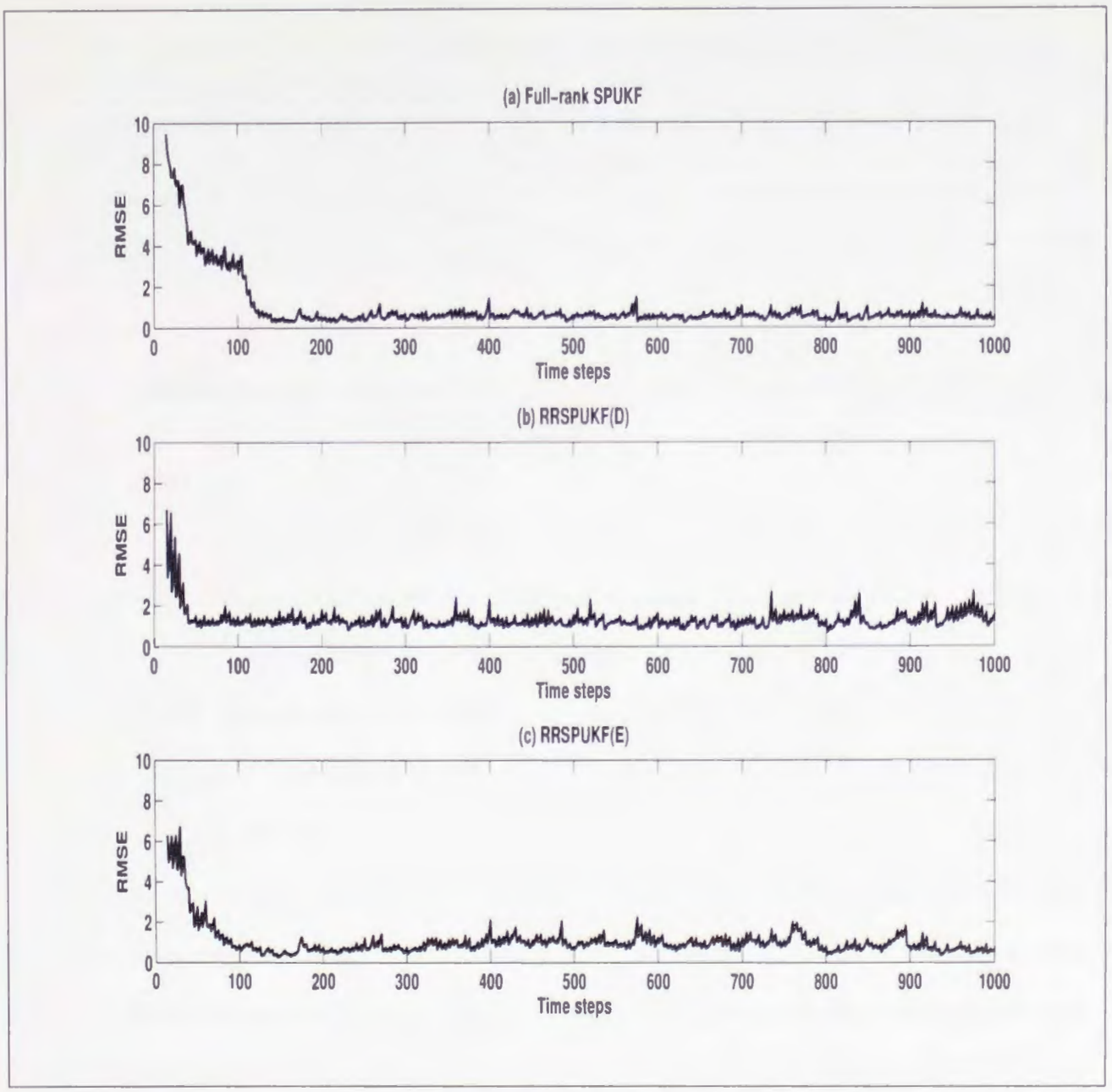

Figure 3.3: The variation of the root mean-squared error over the 40 variables with time step for (a) the full-rank SPUKF, (b) RRSPUKF(D) with 31 sigma points (ensemble members), and (c) RRSPUKF(E) with 31 sigma points.

system with high-dimensionality and large number of observations we can afford only RRSPUKF(E) method.

Table 3.1: Computation time in seconds for various data assimilation methods on a Linux PC with a 2.0GHz Intel Pentium Dual Core processor.

Assimilation method	Computation time (in seconds)
Full-rank SPUKF	1398.72
RRSPUKF(D)(31-sigma points)	152.7
RRSPUKF(E)(31-sigma points)	146.7

3.5.2 Sensitivity experiments with the RRSPUKF(E) method

In the sensitivity experiments we use different truncated modes after the SVD. Figure 3.4 displays the sensitivity experiment results of RRSPUKF(E) method, and the variation of the time mean of the RMSE with the different number of sigma points (i.e., ensemble size). As shown in Figure 3.4, when the number of sigma points increases from 21 to 31, the average error quickly decreases, whereas when the number of sigma points changes from 31 to 41, the improvement in the analysis is not very significant. This suggests that 31 sigma points are most likely sufficient to characterize the error statistics in this case. Figures 3.5 and 3.6 shows the assimilation solution for a different number of sigma points and full SPUKF in comparison to the truth. When the number of sigma points are 3, 5 and 7, the assimilation solution is far away from the truth but as the number of sigma points increases the assimilation solutions are closer to the truth (Figures 3.5 and 3.6). When the number of sigma points are 31 or above the assimilation solution is similar to that of a full-rank sigma-point Kalman filter (Figure 3.6).

Figure 3.7(e) shows the assimilation solution for the variation of the RMSE over the 40 variables with time step t for the full-rank SPKF. Full-rank SPKF can result in a

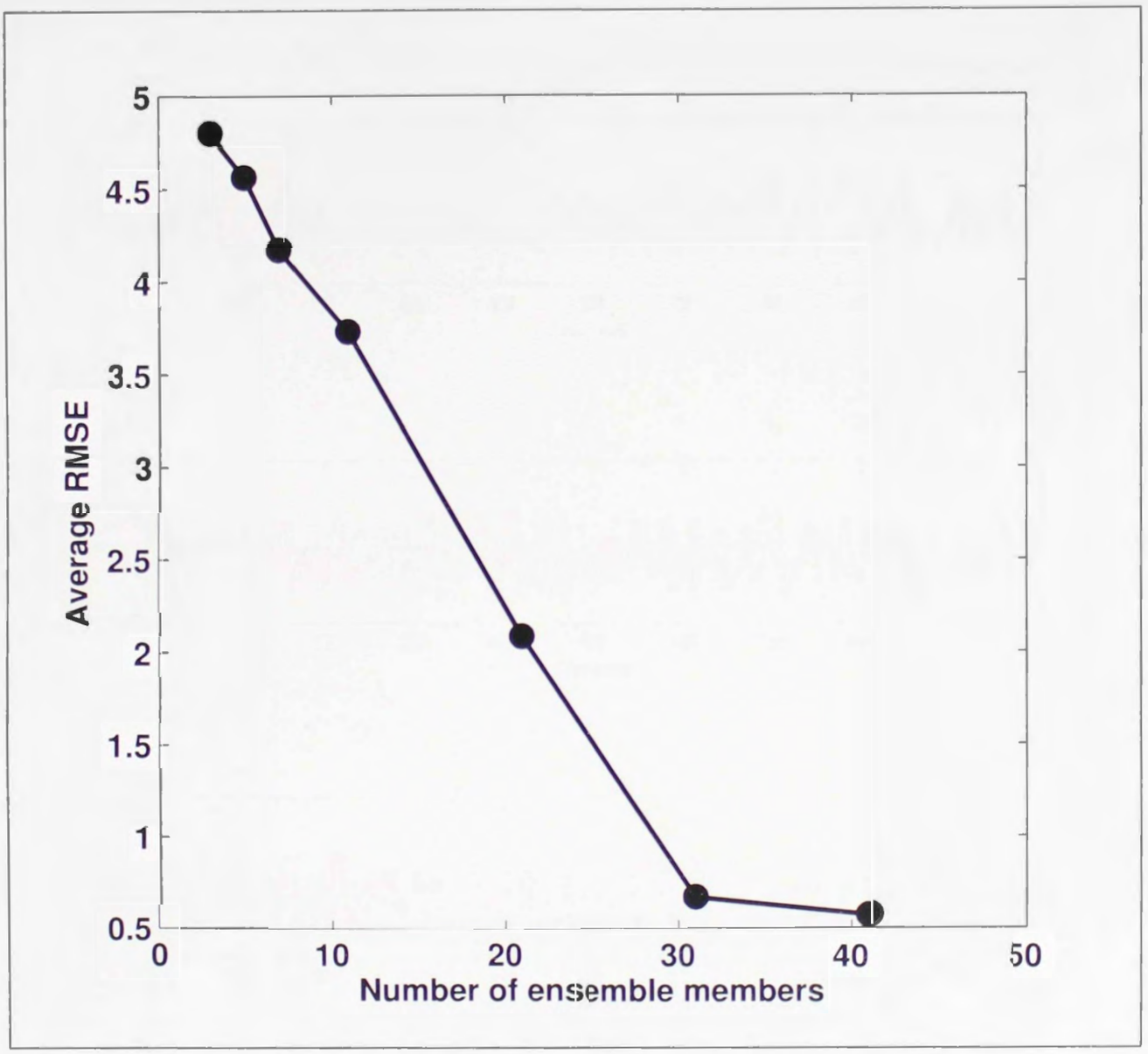

Figure 3.4: The time mean of the RMSE as a function of truncated modes (ensemble size) for RRSPUKF(E) method.

very small RMSE for the estimation of the model states after approximately 120 time steps. The RMSE of the analyses compared to the truth for (b) 41 sigma points, (c) 31 sigma points, (d) 21 sigma points, (e) 11 sigma points are also shown as a function of time step, further confirming that the ensemble size of 31 is sufficient and that the improvement with additional points is only subtle.

Another method for the determination of the truncated modes is based on the variance explained by the truncated modes with respect to the total variance of a full-rank

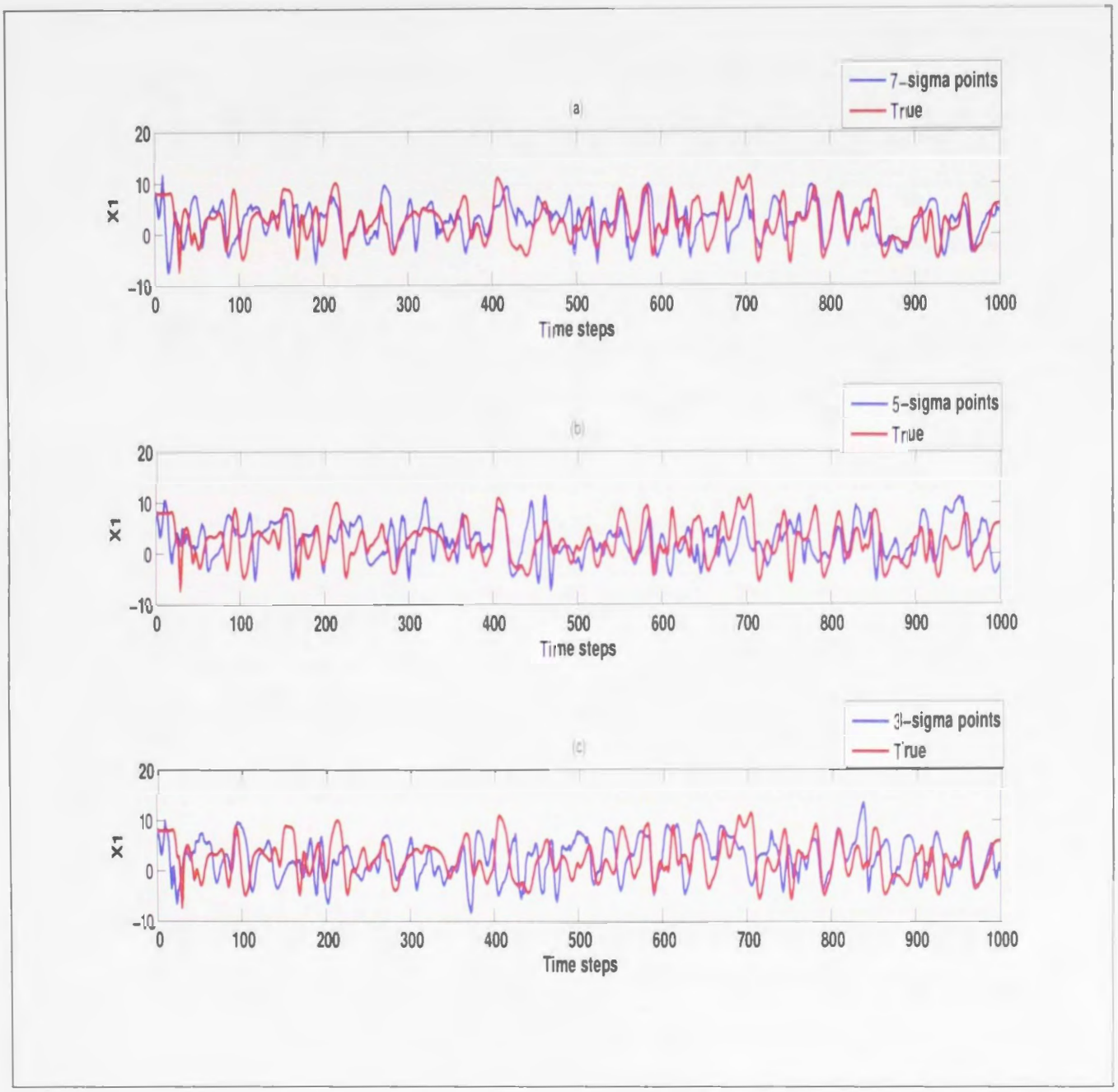

Figure 3.5: Comparison between the true value and analysis of RRSPUKF(E) for the variable X1. The analysis from the truncated modes of 3, 5, and 7 are shown in Figures 3.5a–3.5c, indicating a poor estimation skill in all three cases.

covariance matrix P_t^a. For RRSPUKF(E), however, the total variance P_t^a is unknown because RRSPUKF(E) only calculates the covariance P_t^{aE} of the ensemble space, which is related to the number of initial perturbations. Therefore, one challenge in using RRSPUKF(E) is determining the truncated modes. An approximate solution is to explore the explained variance with respect to the total variance of P_t^{aE} (Figure 3.8).

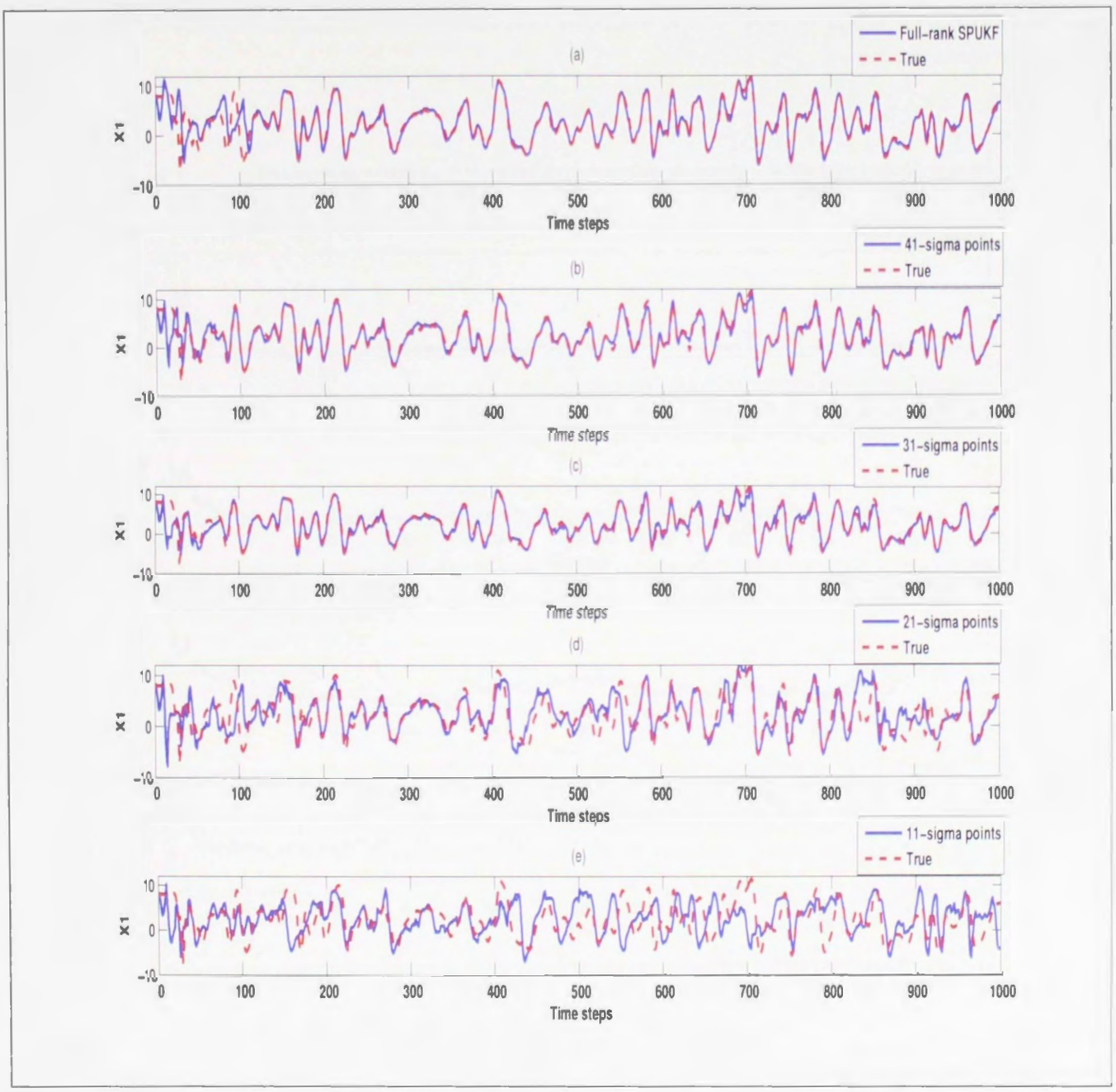

Figure 3.6: Comparison between the true value and analysis of RRSPUKF(E) for the vaiable X1. The analysis from the full-rank SPUKF and truncated modes of 41, 31, 21 and 11 are shown in Figures 3.6a–3.6e.

The variance explained increases with the ensemble size, but the rate is not much higher, especially beyond the size of 11. However, the assimilation experiment with 11 ensemble members is poor, indicating that such a strategy may not be useful (Figure 3.4).

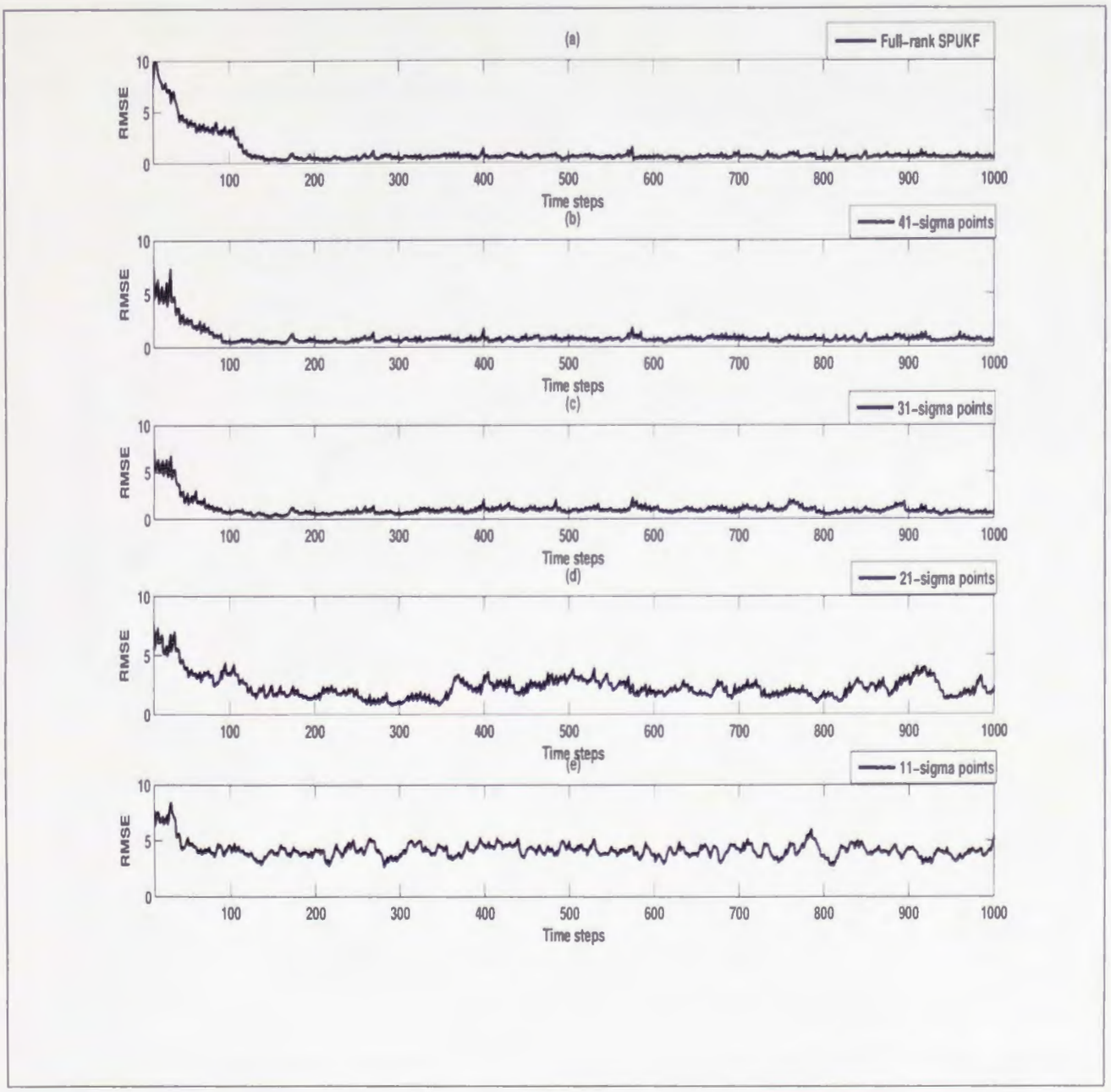

Figure 3.7: The variation of the RMSE over the 40 variables with time step for the different truncated modes (i.e., ensemble size) for the RRSPUKF(E) method.

3.6 Summary

In this chapter, we discussed the formulation of reduced-rank approximations of SPUKF. The reduced-rank approximation was mainly based on the concept that most large-scale geophysical processes can be approximated by a limited number of

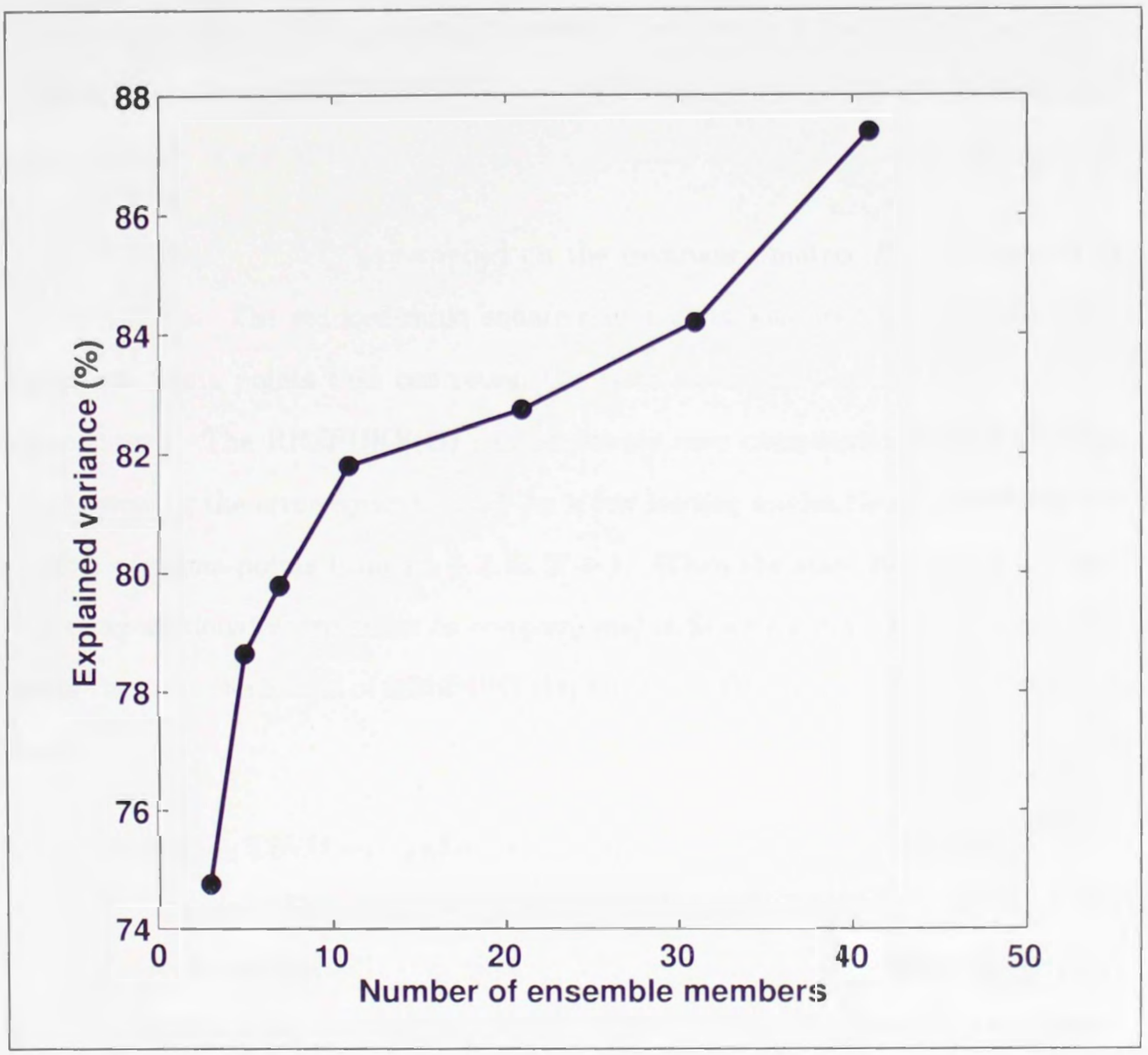

Figure 3.8: The variance explained by the truncated mode (ensembles) with respect to the total variance for the RRSPUKF(E) method.

degrees of freedom and their dominant variability patterns can be described by a finite number of modes (Temam, 1991; Lermusiaux, 1997; Lermusiaux and Robinson, 1999; Ambadan and Tang, 2009). We applied a similar concept as in error-subspace statistical estimation (Lermusiaux and Robinson, 1999) for the reduced-rank approximation of the SPUKF, assuming that the most important statistical features of the original sigma-point space can be estimated using a fewer number of the dominant sigma points (Manoj et al., 2014). Two methods: (i) reduced-rank sigma-point unscented

Kalman filter (Data) (RRSPUKF(D)); and (ii) reduced-rank sigma-point unscented Kalman filter (Ensemble) (RRSPUKF(E)) were investigated to find the reduced-rank approximation of the SPUKF.

In RRSPUKF(D), TSVD was applied on the covariance matrix P_t^a, constructed in the data space. The reduced-rank, square-root matrix was used to select the most important sigma points that can retain the main statistical features of the original sigma points. The RRSPUKF(D) can effectively save computational time through approximating the error covariance P_t^a by a few leading modes thereby reducing the number of sigma-points from $2L + 1$ to $2l + 1$. When the state dimension is large, it is computationally expensive to compute and difficult to store P_t^a itself. That is one of the large challenges of RRSPUKF(D) for the application in a large-dimensional model.

In RRSPUKF(E), TSVD was applied on the covariance matrix P_t^{aE}, constructed in the ensemble subspace. The computational complexity of the RRSPUKF(E) algorithm ($O(n^2 \times L_M)$) is much smaller than that of the RRSPUKF(D) algorithm ($O(L_M^2 \times n)$). The difference in computational complexity between the two algorithms is significant when the model dimension is much greater than the ensemble dimension, that is usually the case for a realistic NWP model.

Data-assimilation experiments were conducted using the Lorenz-96 model. Three different data-assimilation methods (SPUKF, RRSPUKF(D), RRSPUKF(E)) were applied. The results of both the RRSPUKFs were compared to each other and with that of a full-rank SPUKF. Sensitivity experiments were conducted using different truncated modes after the singular-value decomposition. A small size of sigma points failed to capture the main features of the important error statistics and lead to poor estimate, whereas a large number of sigma points were computationally not feasible.

The results showed that the performance or estimation skill of both the RRSPUKF methods were poor when the number of sigma points were 21 or below, but when the number of sigma points increased the estimation accuracy also increased. When the number of sigma points was 31 or more, further improvement was subtle suggesting that 31 sigma points could be used for assimilation, balancing the accuracy and computational cost. Comparison between the results from RRSPUKF(E) and RRSPUKF(D) showed that RRSPUKF(E) was comparable with the RRSPUKF(D) with the same approximate error magnitude. The results of RRSPUKFs were compared with that of full-rank SPUKF, showing that, both the reduced-rank schemes have analysis skill close to that of a full-rank sigma-point unscented Kalman filter even with minimum number of sigma points suggesting a possible solution to applying SPUKF in high-dimensional systems.

Chapter 4

Localization in Sigma Point Kalman Filter

4.1 Introduction

In this chapter the potential problems within the RRSPUKF(E) are investigated along with their possible solutions so that RRSPUKF(E) can be implemented for realistic atmospheric and oceanic models. As we have already seen in Chapter 3, the computational cost limits the size of sigma points we can afford. On the other hand, too small a number of sigma points may cause problems such as spurious correlation, inbreeding and filter divergence (Hamill et al., 2001; Whitaker and Hamill, 2002; Lorenc, 2003; Petrie, 2008; Deng et al., 2010, 2011, 2012; Tang et al., 2014a). Thus, it is interesting to explore the possibility of implementing a localization strategy for RRSPUKF(E).

In Section 4.5 we are proposing a hybrid-localization scheme for RRSPUKF(E) so that the problems with inbreeding, filter divergence and spurious correlation will be minimized. Numerical experiments are conducted using Lorenz-96 model with three different number of state variables.

4.2 Inbreeding, filter divergence and spurious correlation

When the background-error covariance is underestimated there is greater certainty in the forecast estimate and the filter will give more weight to the forecast compared to observations. On the other hand, if the background covariance is too large, the filter will give less weight to the forecast. Undersampling will lead to a phenomenon known as inbreeding, where the forecast-error covariance is underestimated (Lorenc, 2003; Petrie, 2008). If the background covariance is too small, the filter starts giving smaller weights to the observational data. Ignoring the observation information by the ensemble members in the successive assimilation cycles will lead to a process known as filter divergence (Houtekamer and Mitchell, 1998; Hamill et al., 2001; Ehrendorfer, 2007).

The background covariance generated using a limited number of ensemble members often produces spuriously large correlations between grid points, which are at great distances apart (Houtekamer and Mitchell, 1998; Hamill et al., 2001). The spurious correlation between independent, greatly distant grid points that are not physically related can exaggerate the forecast-error covariance, may introduce error in the analysis, and can cause filter divergence (Houtekamer and Mitchell, 1998; Hu et al., 2010).

It has been also observed by Houtekamer and Mitchell (1998) and Hamill et al. (2001) that avoiding the observations that are at remote locations from the analysis grid can also improve the analysis.

The errors in the covariance estimated from a finite number of ensembles can often cause a biased mean and insufficient variance in the analysis. In the following forecast steps, the chaotic dynamics of the model may take the ensemble further away from the true state of the system. In the successive assimilation cycles, the forecast-error covariance is further underestimated, and comparatively even more disregard the observation information, resulting in degraded ensemble of forecasts (Hamill et al., 2001). The methods to combat the problems of undersampling include covariance localization, covariance inflation and localization as in the local ensemble transform Kalman filter (LETKF) (Ott et al., 2004; Miyoshi, 2005; Szunyogh et al., 2005).

4.3 Covariance localization

There are various methods developed to overcome or lessen the undersampling problems in EnKF. One good method is covariance localization proposed by Gaspari and Cohn (1999) and by Houtekamer and Mitchell (2001). In this method, the background-covariance matrix P_t^b is localized by applying a Schur product (an element by element multiplication) with a distance dependent correlation function ρ. The value of the correlation function varies from a value of unity at the observation location to 0.0 at some predetermined distances:

$$(\rho \circ P_t^b)_{i,j} = \rho_{i,j} P_{t_{i,j}}^b \tag{4.1}$$

The resultant product is also a covariance matrix (Horn, 1990) and the refined covariance matrix is free of, or has minimal errors due to, spurious noise. Studies on EnKF conducted by Hamill et al. (2001) and Houtekamer and Mitchell (2001) suggested that analyses were smoother and closer to the truth when the influences of observations beyond a specified distance were omitted .

4.4 Covariance inflation

One method for addressing the underestimation of the background error-covariance matrix is covariance inflation (Anderson and Anderson, 1999; Li et al., 2009; Deng et al., 2010). There are various methods to do that. One method adds a positive semi-definite matrix to the covariance matrix. There is no unique way, however, of choosing the semi-definite inflation matrix to add to the covariance matrix. This process, of finding the suitable positive semi-definite matrix, is treated as a part of model tuning. Adding a positive semi-definite matrix to the background-covariance matrix is also one of the ways to account for the model errors in EnKF (Hunt et al., 2007). Several different model error schemes have been proposed for the covariance inflation (Li et al., 2009). The most common schemes are additive inflation (Corazza et al., 2007), multiplicative inflation (Anderson and Anderson, 1999) and hybrid methods (Li et al., 2009). In the additive-inflation method an identity matrix scaled by a small real random vector is added to the raw covariance matrix $P^b_{t,e}$. Adding a random vector can act as an additional source of noise and destroy the correlation structure of the covariance matrix.

In the multiplicative-inflation scheme the raw covariance matrix is multiplied by a small number greater than one. This direct inflation of background-error covariance

is equivalent to the inflation of model states if the model errors are assumed to be proportional to the background error (Li et al., 2009). In this method, the raw background-error covariance P_t^b is inflated by a factor $(1+\phi)$, where ϕ is the inflation factor also called as the tunable parameter, so the new background covariance will be:

$$P_t^b \leftarrow (1+\phi)P_{t,e}^b \tag{4.2}$$

The size of the parameter ϕ may change according to the number of ensemble members (Whitaker and Hamill, 2002). We use this scheme in Section 4.6 and in Chapter 5. Multiplicative inflation is flow-dependent and can avoid the divergence and inbreeding problem of the filter of finite ensemble size (Anderson, 2001; Hamill and Whitaker, 2005; Furrer and Bengtsson, 2007; Karspeck and Anderson, 2007; Oke et al., 2007).

4.5 Localization in RRSPUKF(E)

The localization scheme used in the EnKFs addresses the computational efficiency

The main challenge for implementing the localization in the SPKF is that the sigma points are generated based on the global analysis-error covariance. As we have already discussed the global analysis without localization for a large-dimensional model would cause the following problems: (i) it would be computationally expensive to update equation (3.7); and (ii) the analysis will be degraded because of the spurious correlation between the remote grid points. In this Section(4.5) we propose a hybrid-localization scheme for the RRSPUKF(E), in which model states are analyzed locally at each model grid but the sigma points are generated globally.

The atmospheric or oceanic model state vector at time t can be denoted by $X_{s,t}$, where s is the location of the grid in the numerical models and is usually fixed and discrete. In general, $s(lat, lon, z)$ has three components, the geographical latitude (lat), longitude (lon) and the altitude (z). For simplicity we discuss the localization in the case of a one-dimensional model that can be easily extended to a two or three dimensional model. Therefore, in our case s has only one coordinate, say $s(lon)$. To perform the analysis locally, first the local vector containing the information from the localization region is defined (Patil et al., 2001; Ott et al., 2004). If the analysis grid points are centred at the grid point v with the location coordinate $s(lon)$ and the localization radius is d, then the grid points inside the local region are $X_{s_{lon-d},t}$ to $X_{s_{lon+d},t}$. There is a total $2d+1$ grid points centred at v. The equation for the local analysis at grid point v is:

$$X^a_{loc,t} = X^b_{loc,t} + K_{loc,t}\left(Y_{loc,t} - H_{loc}(X^b_{loc,t})\right) \tag{4.3}$$

where the subscript loc in the equation above denotes the local subspace. $X^b_{loc,t}$ are the local vectors from the model forecast. $K_{loc,t}$ is the Kalman gain calculated in the

local subspace:

$$K_{loc,t} = P^{xy}_{loc,t}(P^{yy}_{loc,t})^{-1} \tag{4.4}$$

where $P^{xy}_{loc,t}$ and $P^{yy}_{loc,t}$ are the cross covariance and projection covariance in the local subspace respectively, and calculated according to the equations given in Section 2.2.3 of Chapter 2.

A similar operation is performed on each model grid v using only the observations local to v and the analysis $X^a_{v,t}$ is retained. The global-analysis vector X^a_t is the combination of all the $X^a_{v,t}$.

The global analysis-covariance matrix in the ensemble space is constructed using the equation:

$$P^{aE}_t = E[(X^a_t - \overline{X^a_t})^T(X^a_t - \overline{X^a_t})] \tag{4.5}$$

On performing SVD on the P^{aE}_t:

$$P^{aE}_t = F^a_t G^a_t (F^a_t)^T \tag{4.6}$$

Then the leading eigenvectors are found using the equation:

$$E^a_{t,l} = \frac{(X^a_t - \overline{X^a_t})^T * F^a_{t,l}}{\|(X^a_t - \overline{X^a_t})^T * F^a_{t,l}\|} \tag{4.7}$$

The global sigma points are generated according to the equations given in Section 3.2.2 of Chapter 3.

4.6 Numerical experiments

Numerical experiments are conducted to test the performance of RRSPUKF(E) with a hybrid-localization scheme on the Lorenz-96 model (Lorenz, 1996; Lorenz and Emanuel, 1998)) using an OSSE scheme. The sensitivity experiments in this section are motivated by Ott et al. (2004).

The model equation is given by:

$$\frac{dx_i}{dt} = x_{i-1}(x_{i+1} - x_{i-2}) - x_i + F \tag{4.8}$$

where $i = 1, 2, ...N_x$.

4.6.1 Experimental setup

In our experiments, the value of the external forcing $F = 8$ is used and the fourth-order Runge-Kutta scheme is used for the time integration of the model with time step=0.05. The experiments are conducted for several settings of the model by varying the number of spatial coordinates ($N_x = 40, 80$ and 120) as in Ott et al. (2004). The nature or observational run is created by integrating the models over 3800 time steps by using the fourth-order Runge-Kutta scheme. The integration step is set to 0.05 (i.e., 6-hours). The data from the model trajectory in the time steps 0–1000 for the initial spin-up are discarded and the data from the time step 1000–3800 are taken as the truth. The observation vector at each time step is created by adding random normal error to the truth, where the observation error-covariance matrix R is an identity matrix. The RRSPUKF(E) method is used to assimilate the observations

every ten time steps and all the variables are observed.

As the Lorenz-96 model has only one spatial dimension, the localization is carried out as follows. If the analysis grid is denoted by v and the localization radius by d , the grid points considered in the local analysis are $v-d,...v-1, v, v+1, ..., v+d$ at an increment of one grid point in either direction.

In the first set of experiments, different dimensional size of the system ($N_x = 40, 80$ and 120) are considered and assimilate the synthetic observations using RRSPUKF(E). Here no localization is applied and the analysis is performed on the global domain using a varying number of sigma points. The new sigma points are generated following equations (3.4) to (3.10). There is no covariance inflation used (i.e., $\phi = 0$). In the second set of experiments the localization radius is fixed as $d = 6$, which is the default value used for the Lorenz-96 system (Ott et al., 2004). The inflation parameter is fixed to zero ($\phi = 0$) and how many ensembles are required to reach the minimum RMSE is checked in the cases of systems with different dimension ($N_x = 40, 80$ and 120). In the third set of experiments, the inflation parameter is set to $\phi = 3\%$ and the localization radius $d = 6$. In the fourth scheme, the impact of the localization radius (d) on the assimilation analysis is explored. For this, the following parameter values are used. Namely, the number of model variables ($N_x = 40$), number of ensemble members ($k = 7$), and inflation parameter ($\phi = 3\%$). As the Lorenz-96 model has only one spatial dimension, the local region is denoted by a single spatial index (d). For a localization radius (d), the observations from $2d + 1$ grid points centred at the analysis point are used. Varying d from 2 to 7 is used in this experiment. In the fifth set of experiments, the values of $d = 6$, and $k = 7$ are kept constant but the value of the inflation parameter ϕ varies from 0 to 0.07.

The quality of the analysis is measured in terms of RMSE of the analysis-ensemble

mean against the true state. The details of the specific experiments are summarized in Table 4.1.

Table 4.1: Localization experiment details. T4.2 and T4.3 are experiment numbers for the experiments to generate Table 4.2 and Table 4.3 respectively.

Expt. number	Assimilation method	Model variables (N_x)	Localization radius (d)	Inflation parameter (ϕ)	Number of ensembles (k)
1	RRSPUKF(E)	40, 80 and 120	0	0	3 to 91
2	RRSPUKF(E)	40, 80 and 120	6	0	3 to 31
3	RRSPUKF(E)	40, 80 and 120	6	0.03	3 to 31
4	RRSPUKF(E)	40	2 to 7	0.03	7
5	RRSPUKF(E)	40	6	0 to 0.07	7
T4.2	RRSPUKF(E)	40	2 to 7	0.03	3 to 31
T4.3	RRSPUKF(E)	40	6	0 to 0.07	3 to 31

4.6.2 Results and discussion

In this Section the performance of the hybrid-localization scheme is assessed in the presence of varying parameters. The influence of different parameters such as the number of ensemble members (k), inflation factor (ϕ), localization radius (d) and the number of model variables (N_x) is checked on the quality of the analysis.

The time mean RMS error of the RRSPUKF (E) scheme as a function of number of sigma points for different size (N_x) systems is shown in Figure 4.1. If the model has more degrees of freedom it requires a greater number of ensembles to describe the evolution of the mean and covariance. From the results of the first experiment (Figure 4.1) it can be seen how the RMSE varies as a function of ensemble members for different dimensional sizes of the system ($N_x = 40, 80$ and 120). It can also be seen that the RMSE converges to approximately 0.90 irrespective of the dimension of the system. Systems of different sizes need different numbers of ensemble members to

reach the minimum value of RMSE. When the system dimension is small ($N_x = 40$, the solid line), a smaller number of sigma points is needed to reach the level off value of RMSE. As the system size increases, more sigma points are required to reach the level off value of RMSE as evidenced in the case of $N_x = 80$ and $N_x = 120$.

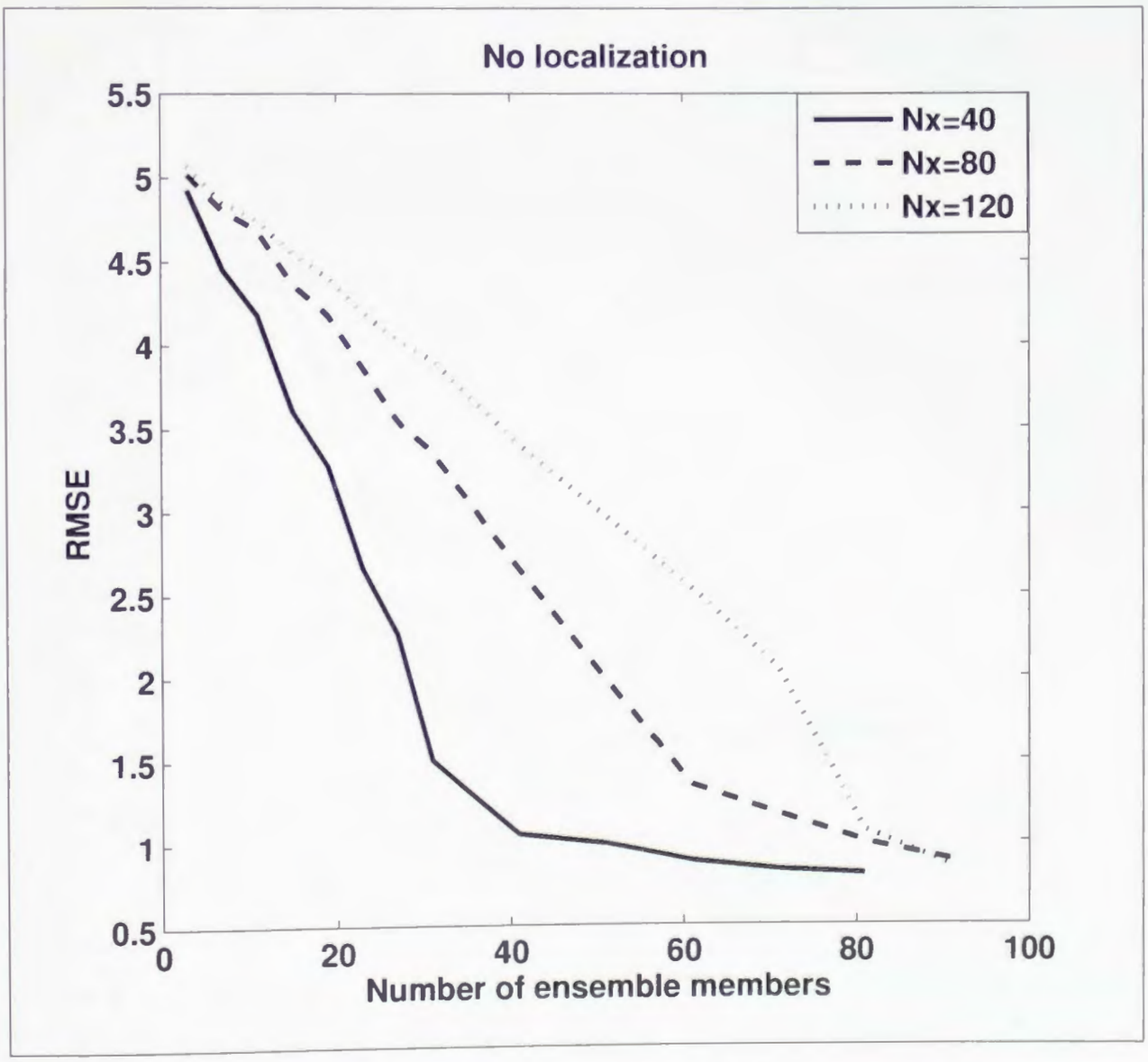

Figure 4.1: The time mean RMS error of the RRSPUKF(E) scheme as a function of number of sigma points (ensemble members). The results are shown for different size (N_x) system.

In the second set of experiments the localization is introduced. The time mean RMS error of the RRSPUKF(E) scheme as a function of number of sigma points (ensemble members) when localization is used for different size (N_x) systems is shown in Fig-

ure 4.2. Here the covariances are described by the the local vectors. Figure 4.2 shows

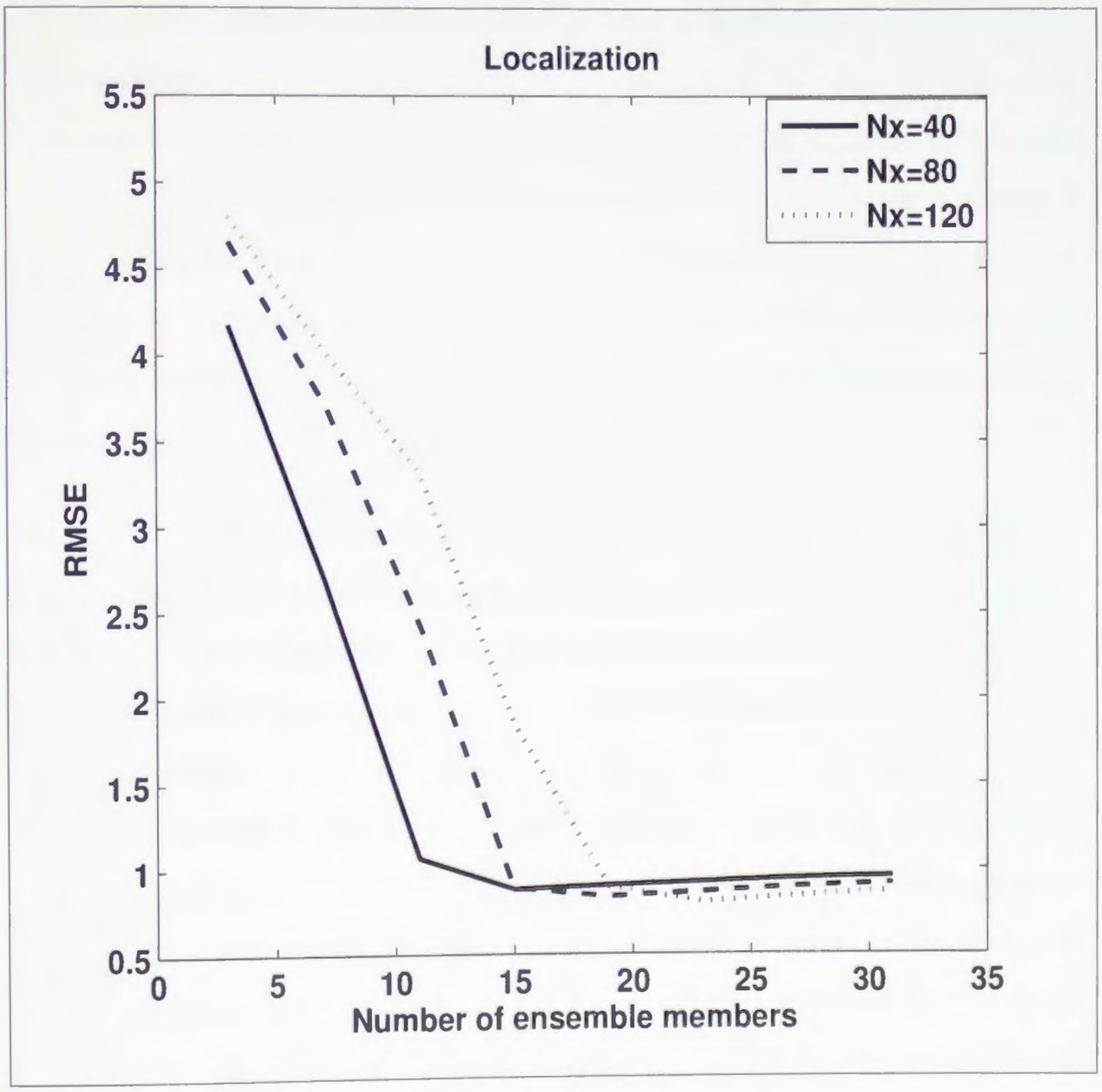

Figure 4.2: The time mean RMS error of the RRSPUKF(E) scheme as a function of number of sigma points (ensemble members) when localization is used. The results are shown for different size (N_x) system.

that introducing the localization significantly reduces the number of sigma points required to achieve the minimum estimation error. When the system size $N_x = 40$ (solid line), the number of sigma points required to achieve the minimum RMSE is around 11 and it does not change with further increases in the number of sigma points.

When the system sizes are $N_x = 80$ and $N_x = 120$, the minimum number of sigma points required to reach the minimum analysis error is 15 and 21, respectively. When comparing Figure 4.2 with Figure 4.1, it can be seen that as the result of localization, the number of sigma points required to achieve the minimum RMSE is significantly reduced. The results indicate that the background uncertainty is low-dimensional and much smaller than the number of degrees of freedom of the system. Another advantage of this localization scheme for application in a GCM is that analysis can be done in parallel for each local region independent of the system dimension, making the method suitable for parallel computation.

The time mean RMS error of the RRSPUKF(E) scheme as a function of the number of sigma points when both localization and inflation are used for different size (N_x) systems is shown in Figure 4.3, where both localization and inflation are introduced. The RMSE reached the minimum value at a sigma points size of seven for all the three cases with different system dimension ($N_x = 40, 80$ and 120) and does not change further with increase in the number of sigma points used (Figure 4.3). This result has important implications when we use the SPKF on a high-dimensional system, because it addresses the importance of localization and covariance inflation. In the absence of parallel computation, the time required for the local analysis increases linearly but if we can use parallel algorithms the local analysis can be done simultaneously. One of the very important highlights is the time used for local analysis is independent of the dimension of the system N_x and the RMSE is independent of system dimension for a simple model such as the Lorenz-96 model. Therefore, $N_x = 40$-variable system is used for further experiments.

The results of the fourth set of experiments (Figure 4.4) show how RMSE changes with localization radius when the number of ensemble members ($k = 7$), and the inflation

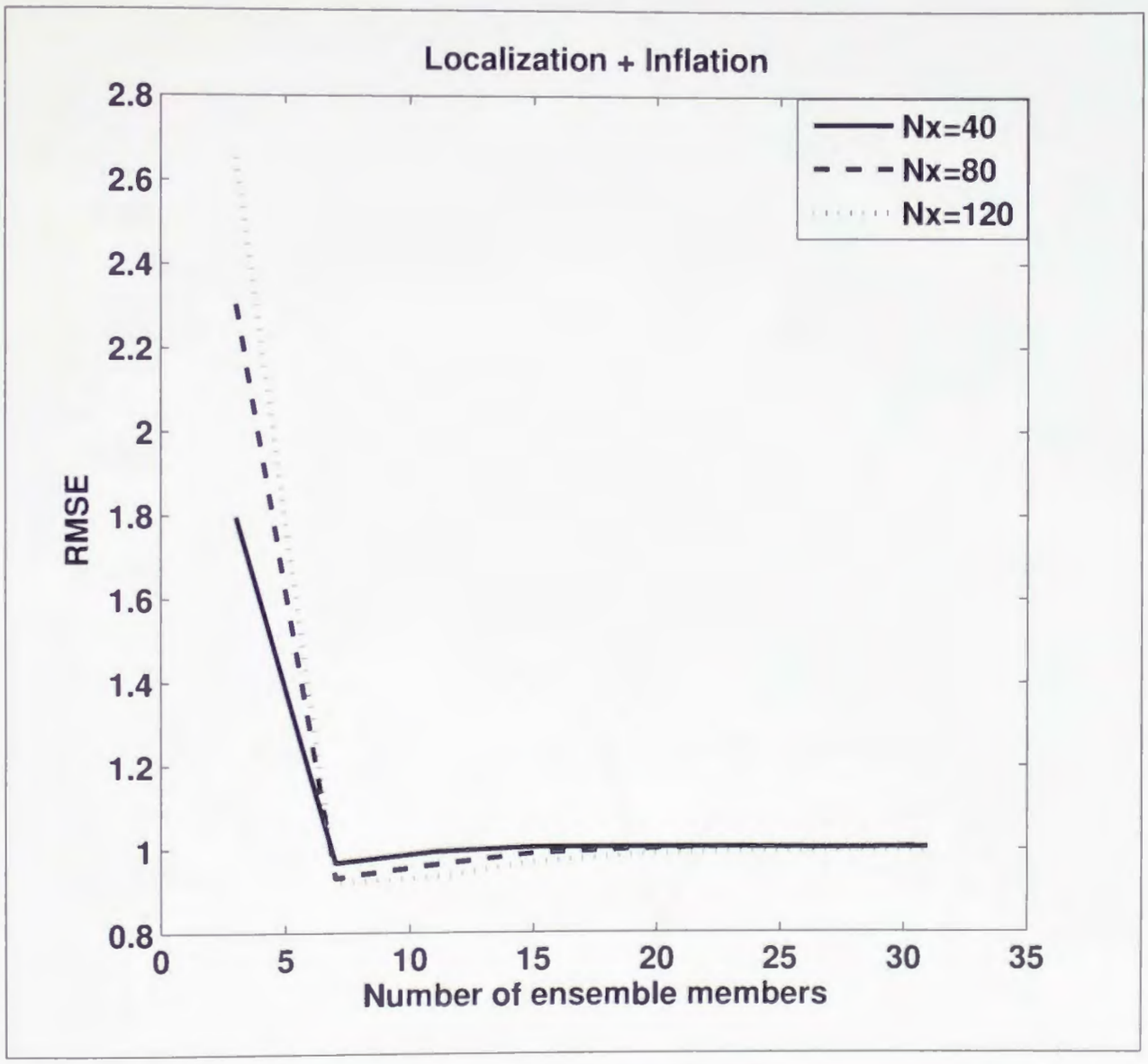

Figure 4.3: The time mean RMS error of the RRSPUKF(E) scheme as a function of number of sigma points (ensemble members) when both localization and inflation are used. The results are shown for different size (N_x) system.

parameter ($\phi = 3\%$), and the number of model variables ($N_x = 40$) are fixed. The minimum RMSE is obtained when the localization radius is $d = 6$, suggesting that it may be a good choice for the further data-assimilation experiments with Lorenz-96 system (Figure 4.4). This result indicates that the background uncertainty can be well approximated in a low-dimensional ($d = 6$) local space provided that the values used for external forcing (F), number of ensemble members (k) and the inflation parameter (ϕ) are 8, 7, and 3, respectively. We have conducted more sensitivity experiments

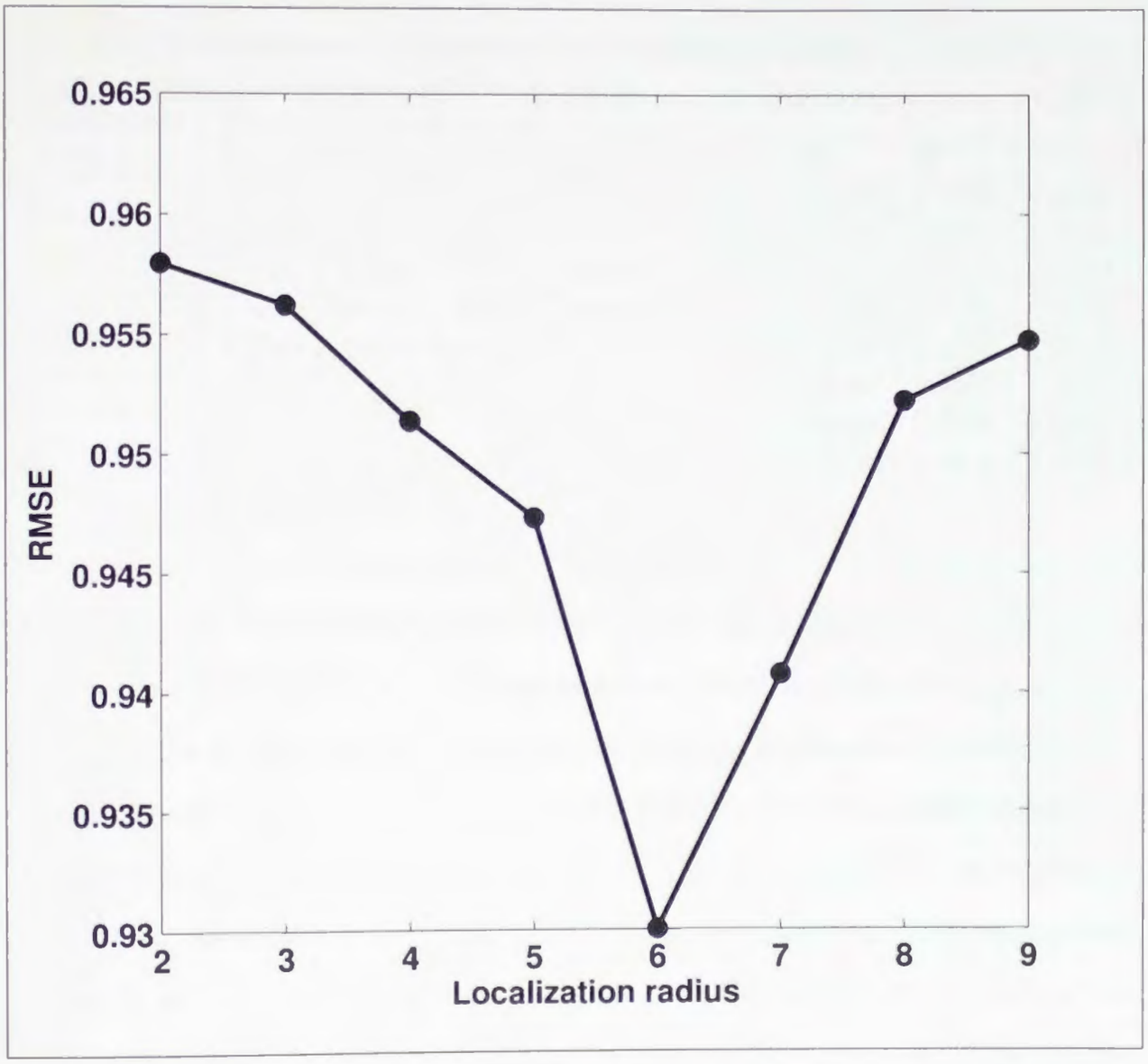

Figure 4.4: The time mean RMS error of the RRSPUKF(E) scheme as a function of localization radius. The inflation parameter is 0.03.

where the value of variance inflation parameter ϕ is fixed as 0.03, but the localization radius (d) varied from 2–9 and the number of ensemble members (k) from 3–31 and the results are summarized in the Table 4.2. In all cases, when the sizes of sigma points are < 7 resulted in large RMSE irrespective of the localization radius. This suggests that the state in the respective local space is higher-dimensional and requires more sigma points to approximate the background-covariance matrix. The minimum RMSE is obtained for an ensemble size of 7 and localization radius of 6.

Table 4.2: The dependence of time mean RMSE on the localization radius and number of sigma points. The inflation coefficient is 0.03.

Localization radius ↓	Number of sigma points							
	3	7	11	15	19	23	27	31
2	2.5940	0.9579	0.9769	0.9904	0.9973	0.9996	1.0008	1.0007
3	2.3366	0.9562	0.9662	0.9912	1.0003	1.0016	0.9992	1.0048
4	2.3400	0.9514	0.9759	0.9925	0.9967	1.0011	1.0045	0.9999
5	2.4076	0.9473	0.9649	0.9879	0.9952	1.0020	1.0058	1.0001
6	2.3283	0.9300	0.9771	0.9892	0.9980	1.0008	1.0014	1.0003
7	2.3923	0.9408	0.9708	0.9853	0.9960	1.0007	1.0006	1.0013
8	2.3220	0.9522	0.9665	0.9902	0.9977	0.9996	0.9998	1.0002
9	2.3779	0.9547	0.9722	0.9903	0.9996	1.0000	1.0008	1.0033

In the fifth set of experiments when the values of $d = 6$, and $k = 7$ are kept constant (Figure 4.5) and the inflation parameter varies, the minimum RMSE is obtained at the value of ϕ is 0.03 or 3%. It can be also seen from Figure 4.5 that when the inflation parameter $\phi = 0$, RMSE is large indicating the necessity of inflation. More sensitivity experiments are conducted to check the influence of inflation when a different number of sigma points are used and the results are summarized in Table 4.3. In this set of experiments, the localization radius $d = 6$ is kept as a constant and other parameters (ϕ and k) are varied. The inflation parameter ϕ is varied from 0.0 to 0.07 at an increment of 0.01, and number of ensemble members (k) from 3–31. The results are summarized in Table 4.3. When the inflation parameter $\phi = 0$ results in large RMSE, which shows the importance of inflation, to cope with the inbreeding and filter divergence. It can be also seen from the Table 4.3 that as the number of ensembles increases, the value of the inflation coefficient needed to reach the minimum RMSE decreases. The optimum inflation coefficient for different ensemble sizes may not be the same (Table 4.3). For example, when the ensemble size is 7 the minimum RMSE obtained when the inflation parameter is 0.03, whereas when the number of sigma points is 11 the minimum RMSE is obtained when the inflation parameter is 0.01.

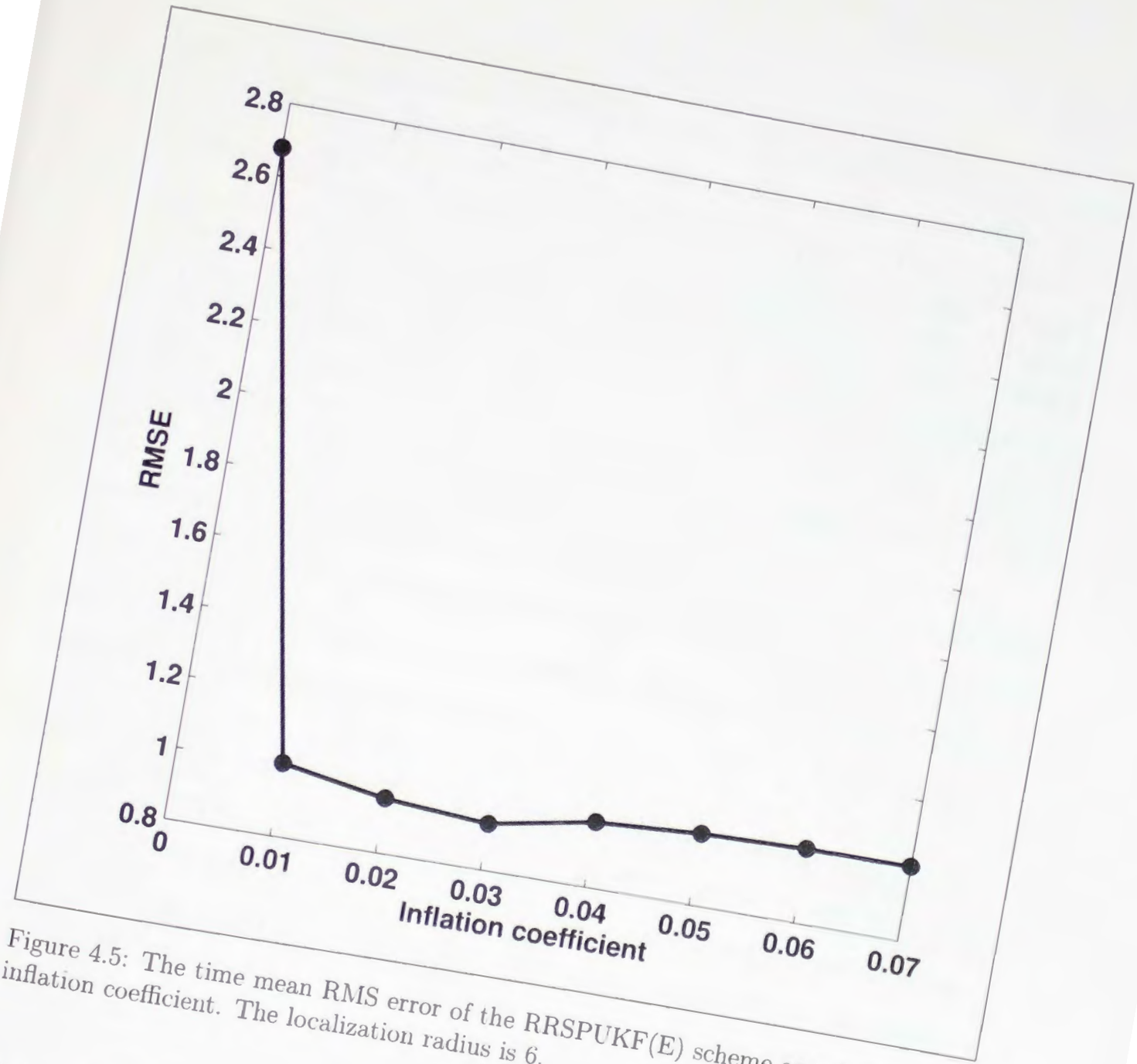

Figure 4.5: The time mean RMS error of the RRSPUKF(E) scheme as a function of inflation coefficient. The localization radius is 6.

4.7 Summary

In this chapter, a hybrid-localization scheme was introduced for the RRSPUKF(E) data-assimilation method, combining local analysis and global generation of sigma points. The localization scheme was based on the assumption that, for a global model, the state vectors of forecast uncertainties lie in a local subspace of much smaller

Table 4.3: The dependence of time mean RMSE on the inflation coefficient and number of sigma points. The localization radius is 6.

Inflation	Number of sigma points							
coefficient ↓	3	7	11	15	19	23	27	31
0	4.3628	2.6755	1.1446	0.9831	0.9922	0.9412	0.9466	0.9543
0.01	3.1660	1.0043	0.9266	0.9592	0.9732	0.9823	0.9908	0.9915
0.02	2.3481	0.9539	0.9699	0.9950	0.9990	1.0000	0.9991	1.0018
0.03	1.7910	0.9300	0.9915	1.0051	1.0072	1.0055	1.0061	1.0059
0.04	1.5470	0.9846	1.0097	1.0082	1.0103	1.0106	1.0089	1.0073
0.05	1.4610	1.0028	1.0130	1.0170	1.0127	1.0101	1.0090	1.0084
0.06	1.3105	1.0116	1.0220	1.0201	1.0133	1.0111	1.0081	1.0039
0.07	1.2880	1.0132	1.0224	1.0233	1.0136	1.0100	1.0084	1.0056

dimension than that of the full state vector in the global domain (Ott et al., 2004). The localization was important because it makes the filter more computationally feasible and may make the analysis estimate optimal by reducing the problems of spurious correlation, inbreeding and filter divergence.

The localization was performed by updating the analysis at each model grid point using the observation within a predefined local subspace centred at that grid point. Combining the local analysis at each grid point the global-analysis vector is constructed. The global analysis-covariance matrix was constructed in the ensemble space and the global sigma points were generated as in RRSPUKF(E). Covariance inflation was also applied for the RRSPUKF(E) to overcome the underestimation of covariance matrix because of the too small number of sigma points.

Numerical experiments of our methods utilized the Lorenz-96 model. The performance of the hybrid-localization scheme was assessed in the presence of varying parameters such as the number of sigma points (k), inflation factor (ϕ), localization radius (d) and the number of model variables (N_x). When the localization was implemented, the number of sigma points required to achieve the minimum RMSE was significantly reduced compared to a case where no localization was used for three different cases

of model variables (N_x). When both the localization and inflation schemes were introduced, the optimal estimate obtained was much less than the case where only localization was introduced. Another interesting finding from the same experiment was that the accurate analysis obtained was independent of the system dimension of N_x in the case of a simple system like Lorenz-96. If the results can be generalized, this will be a positive step towards the potential application of RRSPUKF(E) for an oceanic or atmospheric GCM. The sensitivity experiments also provide suitable parameter choices for the assimilation experiment using the Lorenz-96 model. For the Lorenz $N_x = 40$-variable system, a choice of $k = 7$, $\phi = 0.03$ and $d = 6$ gave the minimum RMSE.

The numerical experimental results indicated that localization and inflation in RRSPUKF(E) was very effective in improving its performance. Key practical aspects of the localization scheme in RRSPUKF(E) data-assimilation method were: (i) a high dimensional estimate of the background covariance was calculated in a local subspace much smaller than the global domain using a small number of ensemble; (ii) the analysis at each local subspace can be done independently making it suitable for parallel computation; (iii) only low-level matrix operations were required; and (iv) there was potential to obtain the optimal estimate at a very modest cost compared to the case where there was no localization used. One of the limitation of this study was that, the Lorenz-96 model, used as the assimilation platform, was a "toy" model; a full atmospheric or oceanic GCM is much more complex, and the results from Lorenz-96 model can, at best, only indicate some general trends and possible behaviours.

Chapter 5

Application of RRSPUKF in an ENSO Model

5.1 Introduction

This Chapter focuses on the investigation of the possibility of applying a reduced rank SPUKF to a realistic climate model. The important properties and the estimation accuracy of the reduced rank SPUKF are explored. Emphasis is placed on the implementation of the reduced rank SPUKF and on its comparison against a full rank SPUKF using a realistic climate model which has not been reported before. The model used is the LDEO5 (Lamont Doherty Earth Observatory version 5 version of the Zebiak Cane El Nino prediction model which is an intermediate complexity model. The RRSPUKF results are also compared with the EnSRF results.

5.2 Model and methods

5.2.1 Zebiak-Cane Model LDEO5 Version

The realistic model we use in this study is the LDEO5 version of the ZC model (Chen et al., 2004), which is an ENSO prediction model (Zebiak and Cane, 1987). ENSO is the strongest signal in the variability of the global climate system on the interannual time-scales. It occurs in the tropical Pacific Ocean irregularly at intervals of 2–7 years and influences the global climate through teleconnections (Wang et al., 1999). In Canada, the largest interannual variation in winter temperature is influenced by ENSO. During the warm phase of ENSO, most of Canada experiences above normal winter temperatures and below normal summer precipitation (Hsieh et al., 1999). Déry and Wood (2005) found a correlation between large-scale teleconnections such as ENSO to the total annual freshwater discharge in 64 rivers in northern Canada.

Studies conducted by Chen et al. (2004) suggest that model-based ENSO prediction depends more on the initial conditions than on unpredictable atmospheric noise. In this study we use the intermediate-complex ZC model as a representative of a realistic climate model to assess the performance of SPKF algorithms. The ZC model has been widely used for predicting the timing, phase and intensity of ENSO events for both experimental and operational purposes since the late 1980s (Webster and Palmer, 1997; Karspeck and Anderson, 2007). Another rationale for using the ZC model is that we can afford the full sigma points required by the full-rank SPUKF method for assimilation and compare their performance with RRSPUKF method.

The ZC is an anomaly model, which computes anomalies of atmospheric and oceanic

fields, relative to a prescribed monthly mean climatological background. The dynamics of the atmospheric component of ZC model are similar to that of Gill (1980), using a steady-state, linear, shallow-water equation on an equatorial beta plane. The atmospheric circulation is forced by a heating anomaly that depends on the heat-flux due to SST anomaly and moisture convergence parameterized in terms of surface wind convergence. The atmospheric model generates the wind fields, which are then converted to stress anomalies that force the ocean model. The dynamics of the ocean component of ZC use the reduced-gravity model. The dynamics are modelled as a single baroclinic mode of the shallow-water equations acting beneath a shallow surface mixed layer (Karspeck and Anderson, 2007). The surface currents are generated by spinning-up the model with monthly mean climatological wind. The thermodynamic equations describe the SST anomaly and heat-flux change. The model time step is 10 days. The model domain is confined to the tropical Pacific Ocean (101.25°E-73.125°W, 29°S-29°N). The grid for ocean dynamics is 2° longitude by 0.5° latitude and for SST physics and the atmosphere the model grid is 5.625° longitude by 2° latitude. Further details about the ZC model can be found in Section 2.3.2.

5.2.2 Methods and experimental setup

We apply the RRSPUKF(D and E) data-assimilation methods to assimilate SST anomalies into the ZC model. The detailed derivation of both the data-assimilation methods is explained in Chapter 3 (RRSPUKF(D) in Section 3.2.1 and RRSPUKF(E) in Section 3.2.2). The estimation skill of reduced-rank methods have been compared to a full-rank SPUKF (reviewed in Section 2.2.3) and EnSRF (see Section 2.2.2).

with random vectors may act as a source of additional noise and destroy the correlation balance. The inflated covariance matrix is:

$$P_t^b \leftarrow (1+\phi)P_{t,e}^b \tag{5.1}$$

where $P_{t,e}^b$ is the raw background-error covariance; ϕ is the inflation coefficient usually obtained by tuning experiments (also called tunable parameter). In this study, the value ϕ is set to 0.15, which was found to result in the best analysis based on a trial and error method and the literature (Karspeck and Anderson, 2007).

5.2.2.3 Covariance localization

To avoid the possible impact of spurious correlation on the analysis, we apply covariance localization in this study (Hamill et al., 2001; Houtekamer and Mitchell, 2001). The background error-covariance matrix P_t^b is multiplied (the Schur product) by a distance dependent localization correlation matrix ρ, and their product will also be a covariance matrix (Horn, 1990; Gaspari and Cohn, 1999). The modified covariance matrix has been defined in Section 4.3 as $\rho \circ P_t^b$:

$$(\rho \circ P_t^b)_{i,j} = \rho_{i,j} P_{t_{i,j}}^b \tag{5.2}$$

with the members of the correlation matrix:

$$\rho_{i,j} = exp(-\frac{r_{i,j}^2}{d^2}) \tag{5.3}$$

where ρ is a localization function that has a value of unity at the analysis grid point and its value decreases as the distance from the analysis grid point increases, d is the decorrelation length and r is the distance between the grid points i and j. The localization radius used in this study is 45° becasue it was found to be a good value based on the sensitivity tests and earlier studies (Jiang et al., 2012).

5.2.2.4 SSTA assimilation

In our experiments we assume that the observation errors and model errors are uncorrelated in space and time. The total number of spatial grid points within the assimilation domain is 540. When SPUKF is implemented, the state vector is redefined as the concatenation of the model states, model errors and measurement errors, thus the augmented state dimension will become 3×540. Consequently, the number of a full SPUKF sigma points is $(2 \times (3 \times 540)) + 1$. In our study, to decrease the number of sigma points, rank reductions are applied. In RRSPUKF(D) the TSVD is applied to the analysis error-covariance matrix constructed in the data space, P_t^a, and in RRSPUKF(E) TSVD is applied in the analysis error-covariance matrix constructed in the ensemble space, P_t^{aE}. The total number of sigma points $2L+1$ is thus reduced to $2l+1$. The estimate accuracy of the truncated covariance matrix depends on the choice of l. The selection of l should be chosen cautiously in such a way that it should not be too small nor too large. If l is too small, some of the information from the P_t^a will be lost and a large l will create a computational burden. The Niño-3.4 index, defined as the average SST anomaly in the equatorial Pacific (5°S-5°N, 120-170°W), is used as an indicator to track the phase and intensity of the ENSO events. The Niño-3.4 region (Figure 5.1) has large variability on ENSO time scales. The experiment details are summarized in Table 5.1.

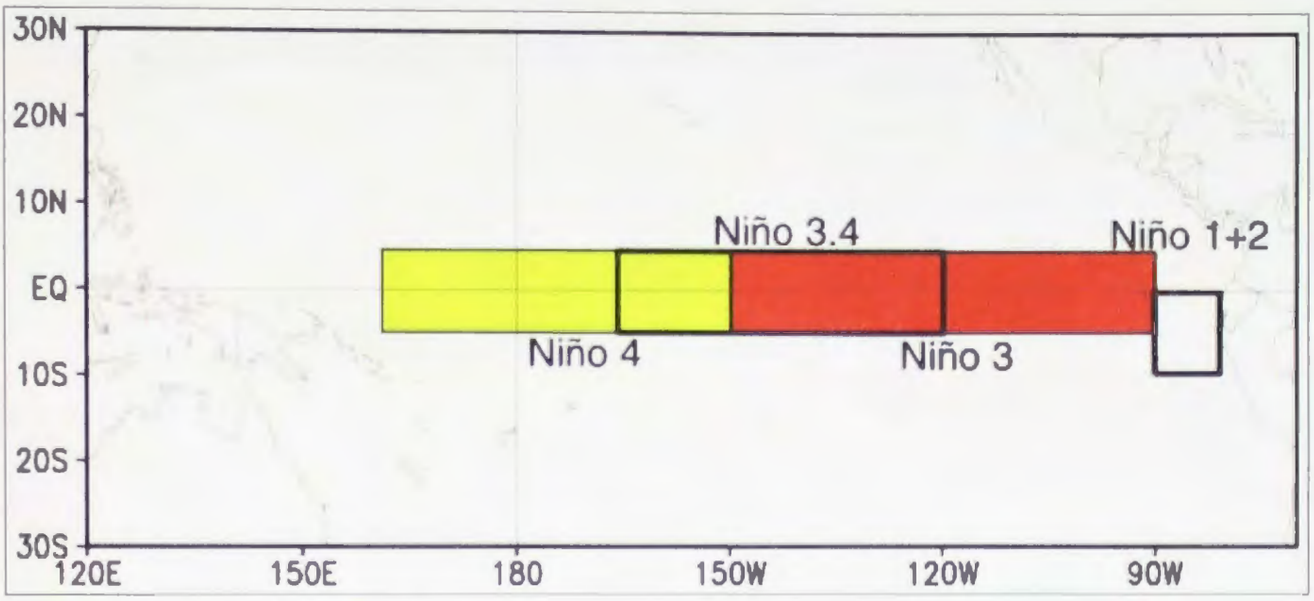

Figure 5.1: Graphical depiction of the Niño regions (from National Centers for Environmental Prediction (NCEP)).

Table 5.1: ENSO state estimation: Experiment summary

Assimilation method	Time period	Observational data	Analysis update	Model error	Localization method
SPUKF	1971-2000	Kaplan et al. (1998) SSTA data	monthly	Multiplicative inflation	Covariance localization
RRSPUKF(D)	1971-2000	Kaplan et al. (1998) SSTA data	monthly	Multiplicative inflation	Covariance localization
RRSPUKF(E)	1971-2000	Kaplan et al. (1998) SSTA data	monthly	Multiplicative inflation	Covariance localization
EnSRF	1971-2000	Kaplan et al. (1998) SSTA data	monthly	Multiplicative inflation	Covariance localization

For the sensitivity experiments with RRSPUKF(D) different numbers of sigma points ranging from 5 to 101 are tested and the results are explained in Section 5.3.1. To further explore the impact of the RRSPUKF(D) assimilation method with different ensemble size on ENSO prediction, the SSTA hindcasts of the duration of 12 months, initialized from the first day of each calendar month, are performed between the period 1971–2000. The Kaplan SSTA data set is used to verify the model predictions. The results of the hindcast experiments are discussed in the Section 5.3.3

Calculation of explained variance: The variance explained by truncated modes is

calculated as below, with each variable having the same notation as that in Chapter 3:

$$P_t^a = E_{t,L}^a D_{t,L}^a (E_{t,L}^a)^T \tag{5.4}$$

$$Variance_{exp} = \left(\Sigma_{i=1}^{i=l} D_{i,i}^a / _{trace(D_{t,L}^a)} \right) \times 100\% \tag{5.5}$$

5.2.2.5 Statistical significance test

The bootstrap method is used to perform a statistical significance test for RMSE and correlation skill, as follows: 1) For a given grid point, the analysis and observed sample of the same time are combined to form a pair sample over 1971–2000; 2) 95% of the pair samples are randomly chosen to calculate the correlation coefficient and RMSE between analysis and observation; 3) process (2) is repeated 10000 times to produce 10000 correlations and RMSEs, from which the standard deviation is obtained; and 4) the 95% confidence interval is used as the threshold value of sampling errors.

5.3 Results and discussion

5.3.1 Sensitivity experiments with RRSPUKF(D)

In this section the performance of RRSPUKF(D), as a function of the number of truncated modes of the SVD, is examined. A comparison among different truncated modes is shown in terms of root mean-squared error (RMSE) between the analysis values and the observed values (Figure 5.2). When the number of sigma points in-

creases from 5 to 101, the RMSE of Niño-3.4 SSTA analysis compared to observations decreases. However, when the number of sigma points used is beyond 40 sigma points (i.e., $l = 20$), there is little improvement in the analysis. When the number of truncated modes is less than approximately 10 there is considerable increase in the RMSE. One of the reasons for higher RMSE when the ensemble number is small is due to underestimation of the covariance matrix by the finite truncated modes.

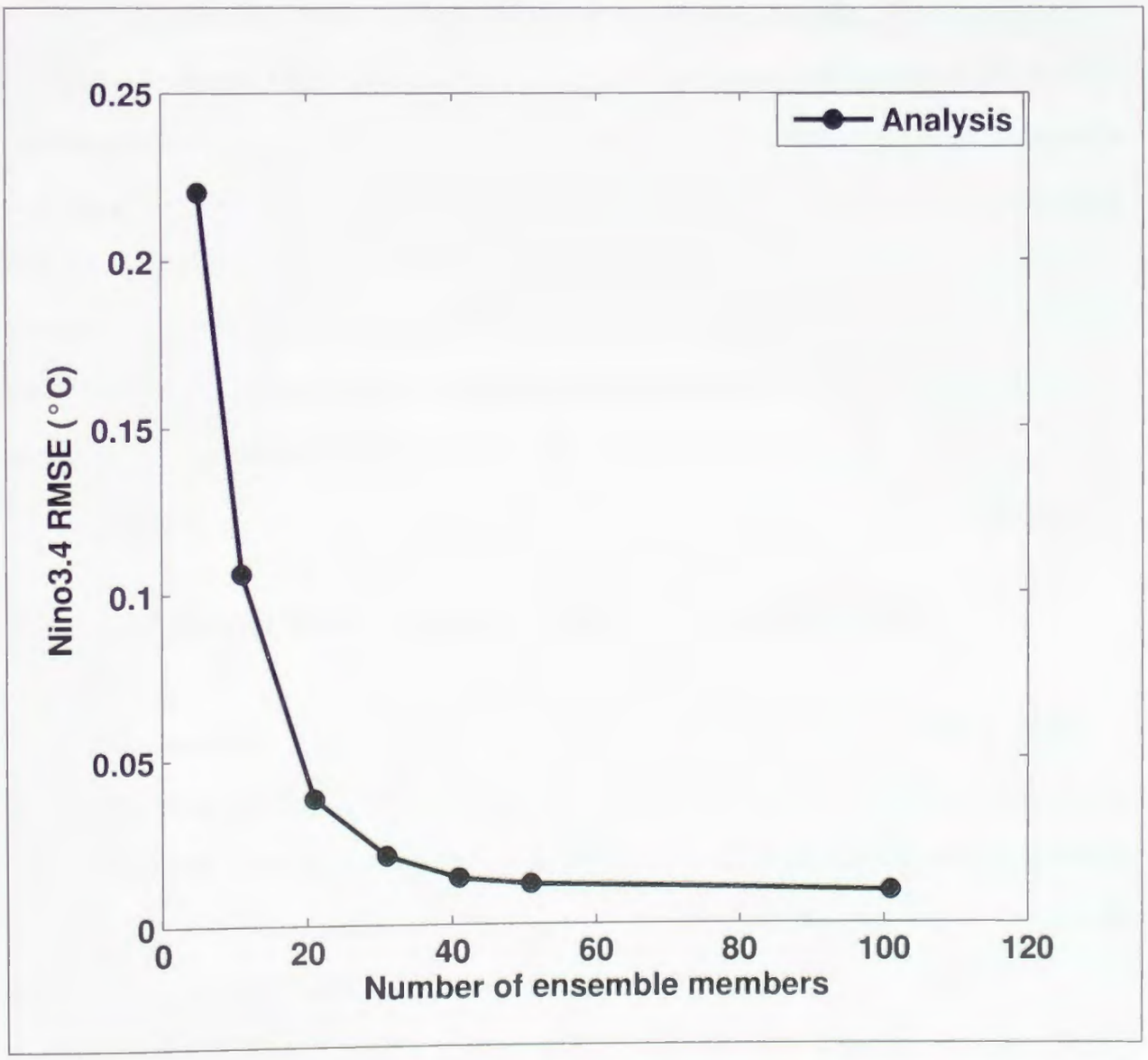

Figure 5.2: The root mean-squared error (RMSE) of SSTA (5°S-5°N, 120-170°W) between analysis and truth as a function of number of ensembles used.

The RMSE and correlation skill for the period 1971–2000 over the entire basin are

shown in Figure 5.3 and Figure 5.4. As in Figure 5.2, the analysis has a large error when the ensemble size is small, as indicated by the higher RMSE and lower correlation coefficient. This is a case of filter divergence, which usually occurs when the forecast covariance is small compared to the measurement uncertainties and the filter assigns small weight to the measurement (Furrer and Bengtsson, 2007; Karspeck and Anderson, 2007). When the ensemble number increases, the RMSE decreases; but beyond 40 ensembles, there is little improvement in the analysis. This suggests that an ensemble size of 40 is sufficient to approximate the mean and spread of the forecast distribution as well as to characterize the coherent structure between the measurement and the prior state. Different ensemble sizes from 31–51 can lead to similar behaviour in both RMSE and correlation, but below 21 ensembles the skill is very poor (Figures. 5.3 and 5.4). This indicates that if the basic characteristics can be captured by the minimum truncated modes, the sensitivity of the assimilation performance to further increasing the truncated modes is not high.

5.3.2 Comparison of RRSPUKF(D) and SPUKF

Based on sensitivity experiments of the analysis error to the truncation modes l, it appears that 41 sigma points (i.e., $l = 20$) are sufficient to estimate the mean and covariance. The resulting analysis is very good and comparable with a full-rank SPUKF as shown in Figure 5.5 and Figure 5.6. The variance explained by the 41 sigma points is more than 85% of the total variance.

Figures 5.5 (a) and (b) show the RMSE error of the Niño-3.4 index (the average SST anomaly in 5°S-5°N, 120-170°W) between the analysis and the observation for both the full-rank SPUKF and RRSPUKF(D) with 41 sigma points, respectively. As can

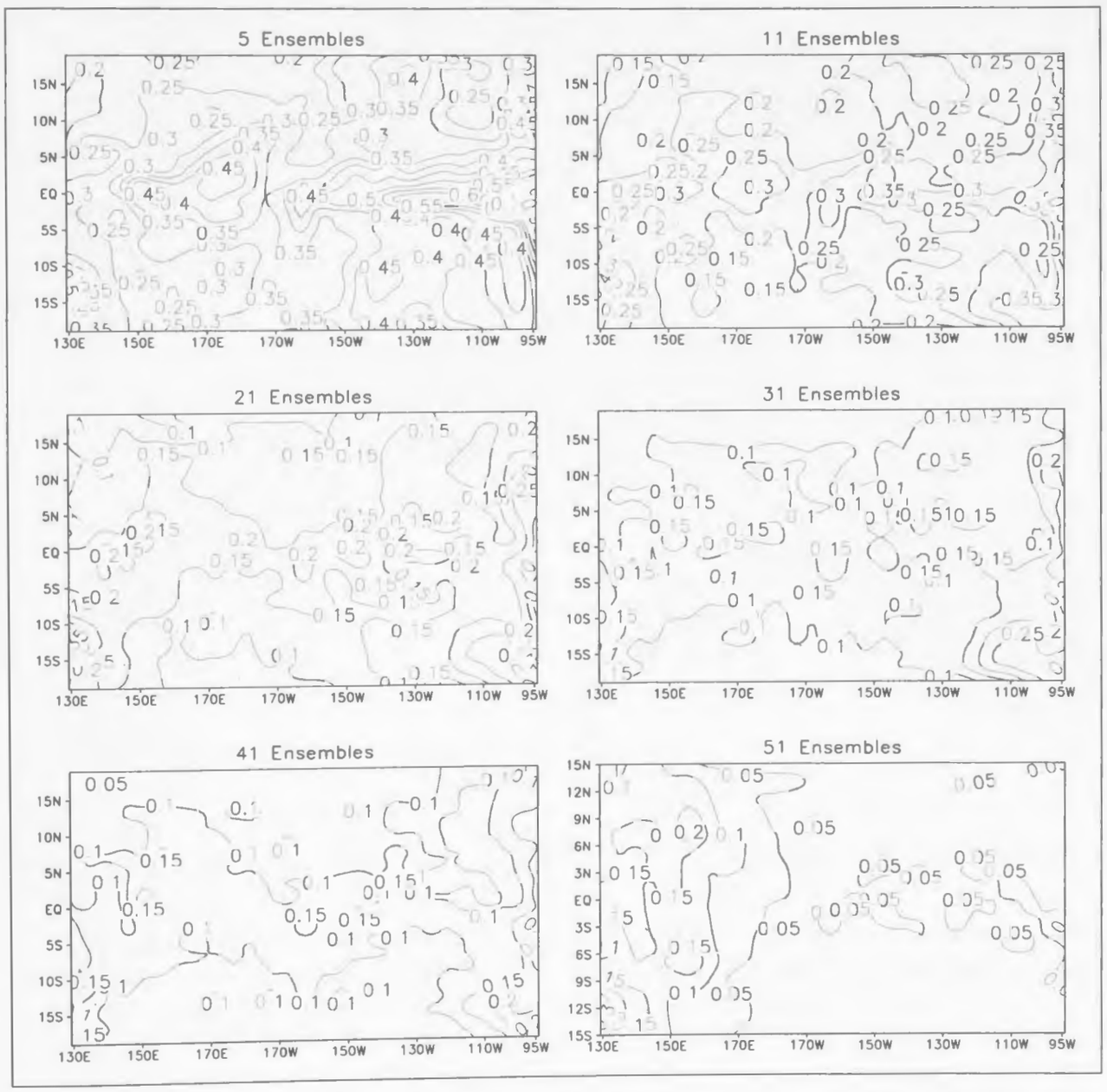

Figure 5.3: The RMSE of SSTA between the analysis and truth averaged over the period 1971–2000 for the RRSPUKF(D) when different numbers of truncated modes (l) are used. The number of ensembles is equal to $(2l+1)$.

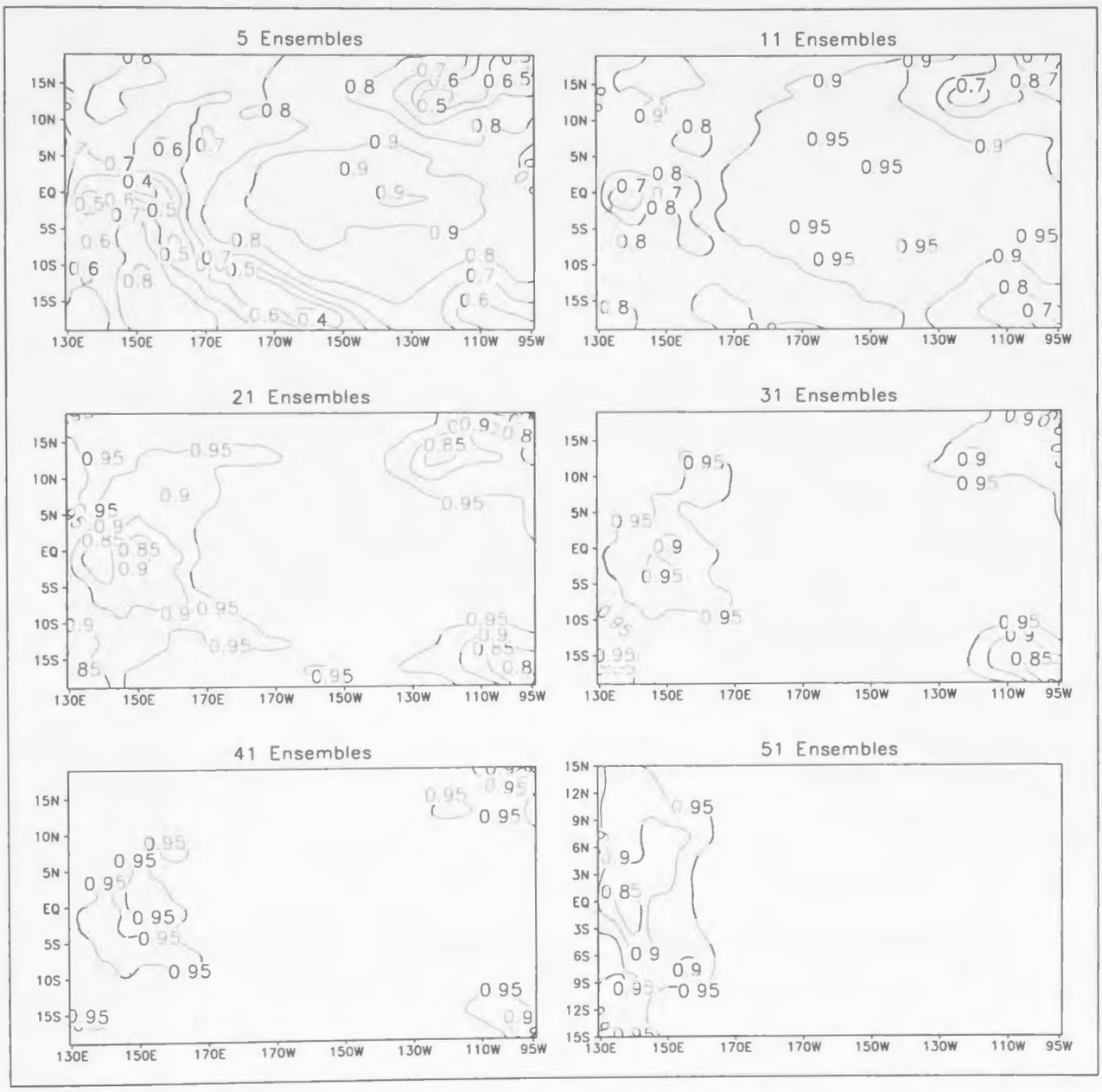

Figure 5.4: The correlation coefficient of SSTA between the analysis and truth averaged over the period 1971–2000 for the RRSPUKF(D) when different numbers of truncated modes (l) are used. The number of ensembles is equal to $(2l+1)$.

be seen in Figure 5.5, the RMSE of RRSPUKF(D) is similar to that of the full-rank SPUKF in most of the assimilation window. This can also be seen in Figure 5.6 where the RMSE and correlation between the analysis and observation for both methods are compared in the entire basin. The bootstrap experiments are conducted to perform the statistical significance test for RMSE and correlation skill. Figures 5.5e and 5.5f show

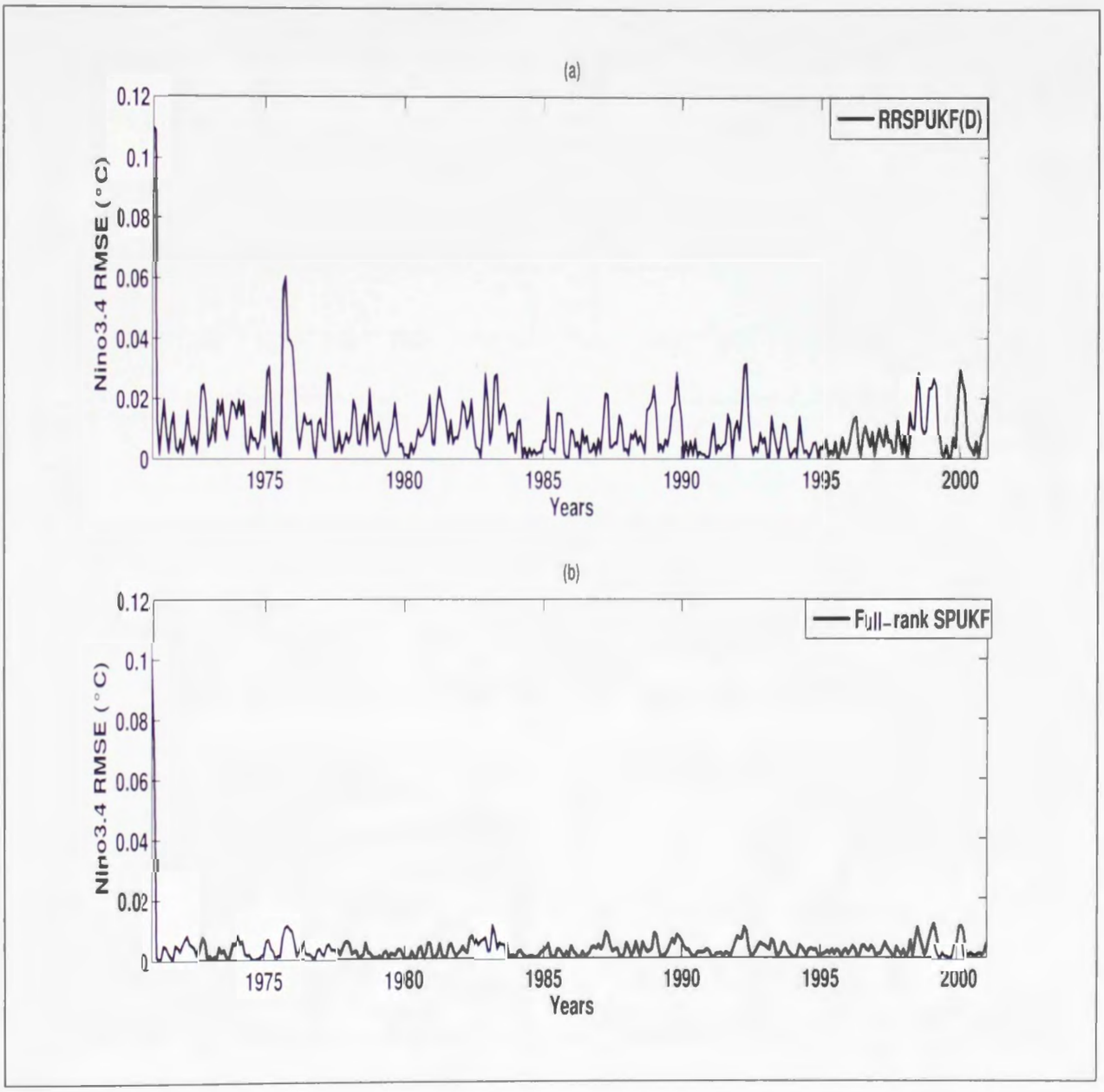

Figure 5.5: The RMSE of the Niño-3.4 index of the analysis against the observed counterpart. The panel (a) is for RRSPUKF(D) when 41 ensembles are used and panel (b) is for the full-rank SPUKF.

the difference of RMSE and correlation between SPUKF and RRSPUKF(D), with the statistical significance area shaded at the confidence level of 95%. The correlation skill of RRSPUKF(D) has no significant difference from that of a full-rank SPUKF for most of the domain of the assimilation (Figure 5.6f). The difference of RMSE between RRSPUKF(D) and SPUKF, however, is significant for almost the entire assimilation

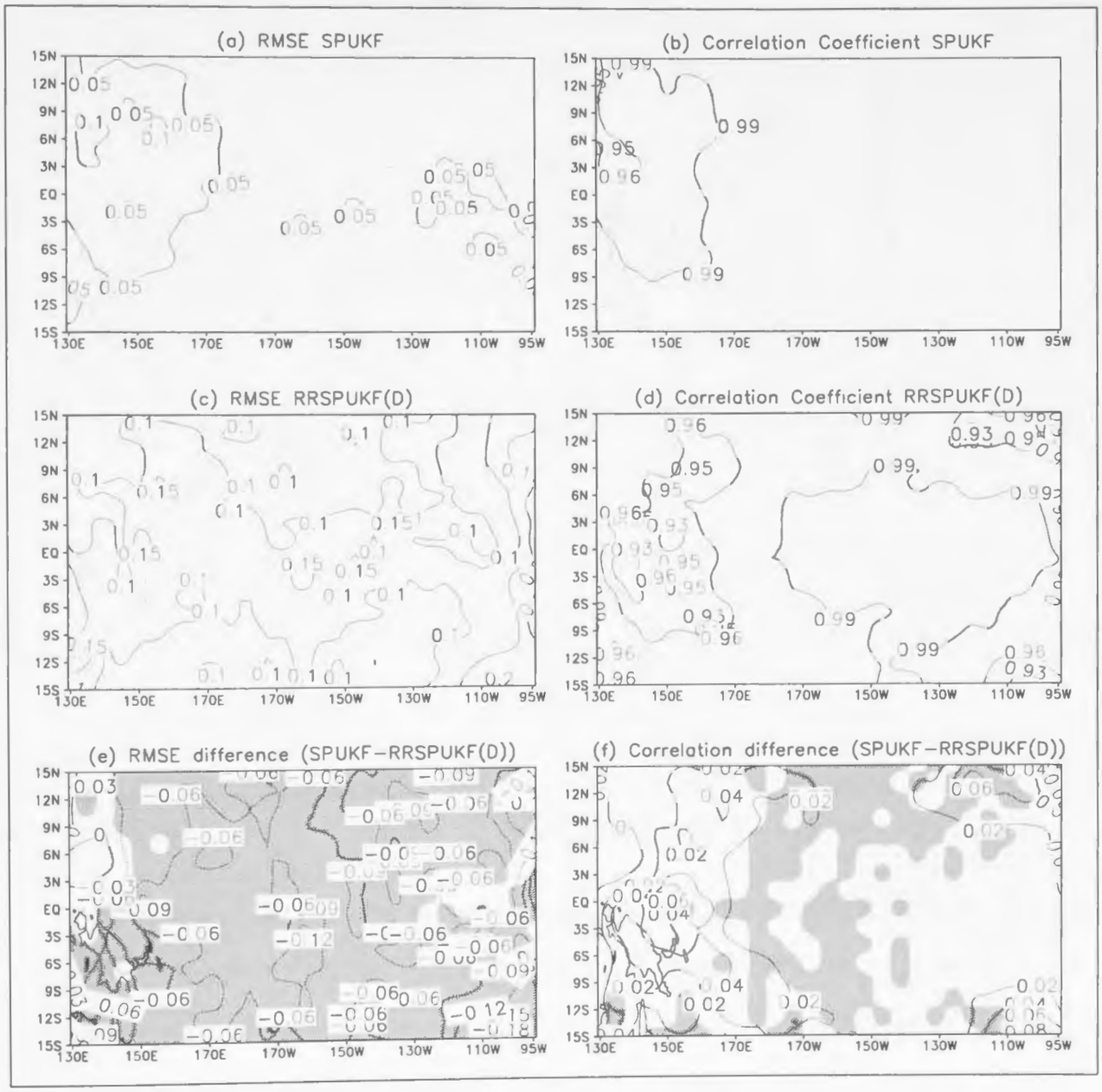

Figure 5.6: The RMSE and correlation coefficient of SSTA between the analysis and truth for full-rank SPUKF and RRSPUKF(D) with 41 ensembles for the period 1971–2000 (a-d). Figure 5.6(e) and (f) are the difference in SSTA RMSE and correlation coefficient of SPUKF and RRSPUKF(D) where the statistically significant areas (at 95% confidence level) are shaded.

domain as shown in Figure 5.6e. This is mainly because the RMSE itself is very small in both methods, leading to very small criteria values. Thus, it can be argued that the RRSPUKF could fairly well simulate the performance of the full-rank SPUKF in terms of the measure of correlation skill, and also can generate small RMSE.

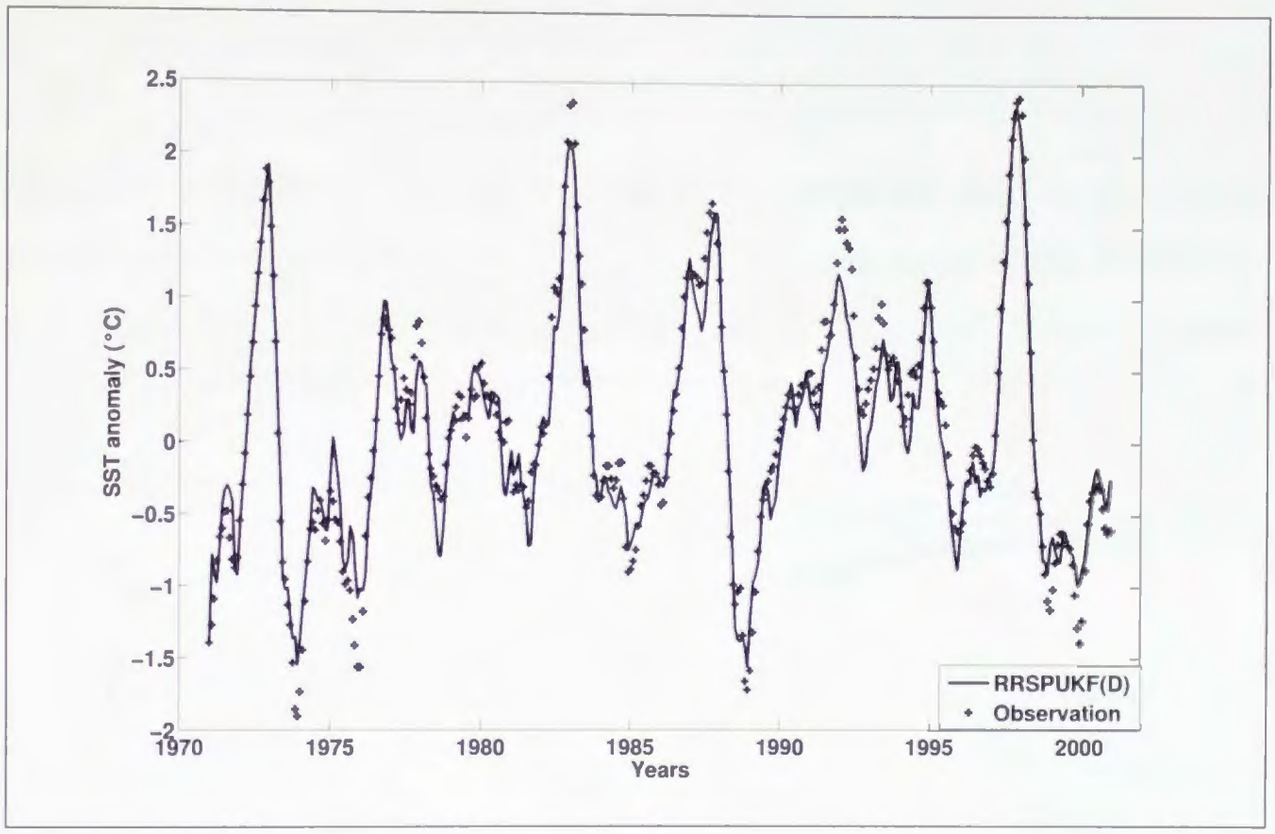

Figure 5.7: One-time step forecast for the Niño-3.4 region (5°S-5°N, 120-170°W). + indicates the Niño-3.4 index of the observation and the solid line indicates the Niño-3.4 index of a one month lead forecast of RRSPUKF(D) with 41 ensembles.

Figure 5.7 shows the one step forecast (one-month lead forecast) of Niño-3.4 SSTA index from the RRSPUKF(D) with 41 sigma points during the time period 1971–2000. The solid line is the model state before assimilation (i.e., the prediction initialized from the analysis of last step analysis) and + denotes the observations. The RRSPUKF(D) filter has a good capability for estimating the phase and intensity of all major ENSO events during the entire period (Figure 5.7). The large errors occurs at a few time steps when there is high frequency variability in SSTA, which is reasonable given that the observation used for assimilation can also contain uncertainties.

5.3.3 Forecast skill

The quality of the ENSO hindcast of 12-month lead initialized from an ensemble of analysis obtained using the RRSPUKF(D) is evaluated in terms of the RMSE and correlation skill of predicted SSTA against the observed counterpart. Figure 5.8 shows

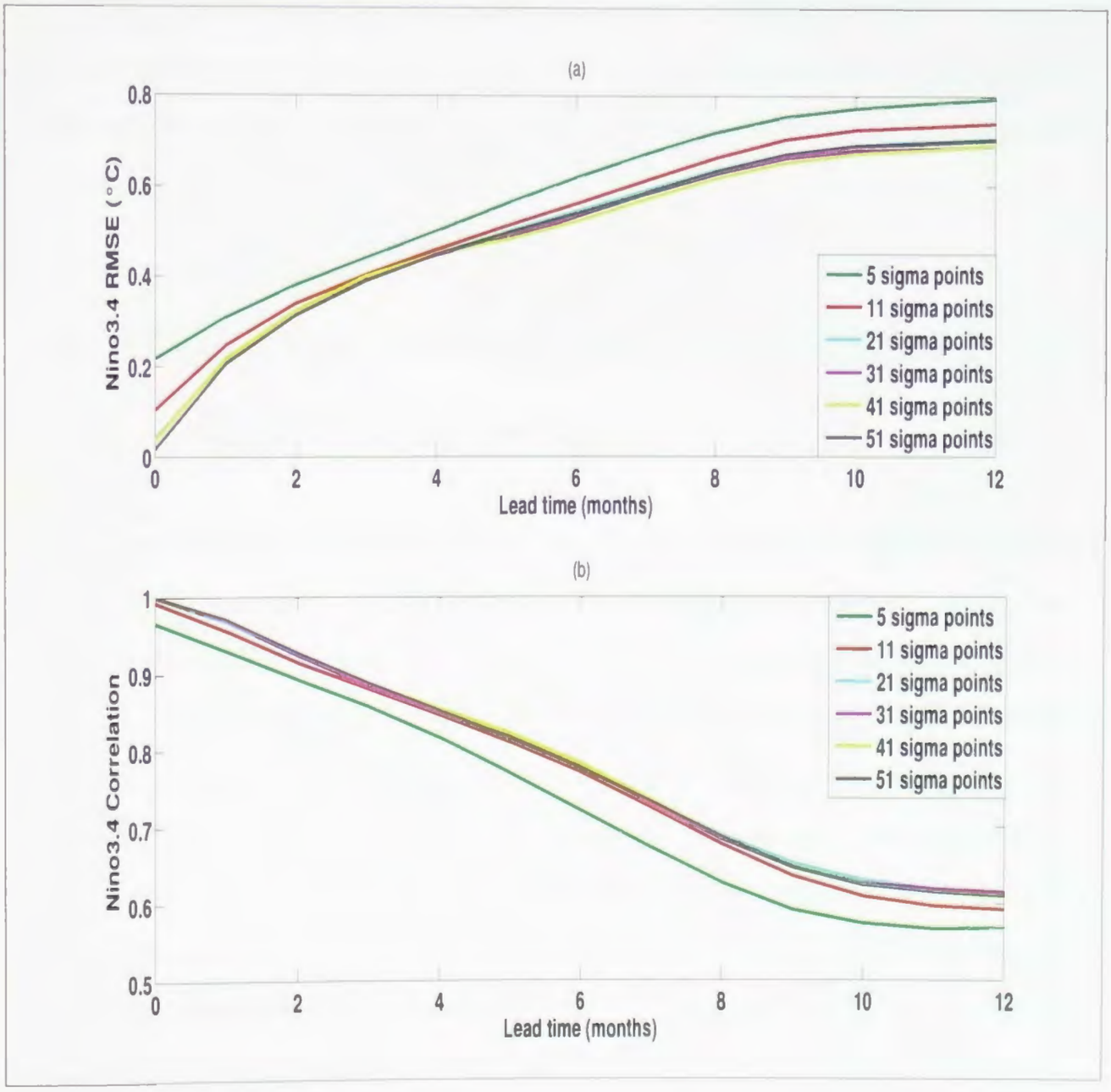

Figure 5.8: The Niño-3.4 forecast RMSE and correlation with 12-month lead time for RRSPUKF(D) for the cases with different number of ensemble members.

the RMSE and correlation skill of Niño-3.4 SSTA, as a function of lead time for differ-

ent ensemble sizes. The correlation skill is 0.5 or higher for predictions at lead times of up to 12 months irrespective of the ensemble size. Both the RMSE and correlation skills decrease with lead time, as characterized in a nonlinear or stochastic dynamical system such as ENSO (Jin et al., 1994; Penland and Sardeshmukh, 1995; Kirtman and Schopf, 1998; Chen et al., 2004). On the other hand, the RMSE decreases and correlation skill increases as the ensemble size increases. The prediction skill starts to level off and have little improvement when the number of ensembles is approximately above 20. This result is consistent with those obtained in the sensitivity experiments in Section 5.3.1.

5.3.4 Comparison between the RRSPUKF(D) and RRSPUKF(E) approaches

An alternate method for finding the dominant sigma points is through TSVD in the ensemble subspace as introduced in Chapter 3. This method is computationally feasible even for a high-dimensional model. Figure 5.9 compares the RMSE skill of analyzed Niño-3.4 SSTA index obtained by the RRSPUKF(D) and RRSPUKF(E) methods, respectively, for the period from 1971–2000, against the observations. The RRSPUKF(D) is better than RRSPUKF(E) when the ensemble size is small, indicating that the latter is more inefficient in using small ensembles to capture the main features of the full covariance matrix. This is because the RRSPUKF(D) directly decomposes the full covariance whereas the RRSPUKF(E) uses the ensemble to approximate the covariance matrix. As the ensemble size increases, however, the difference between the two methods becomes very small. Figure 5.10 shows the RMSE (a and c) and the correlation (b and d) skill of analyzed SSTA using the two methods

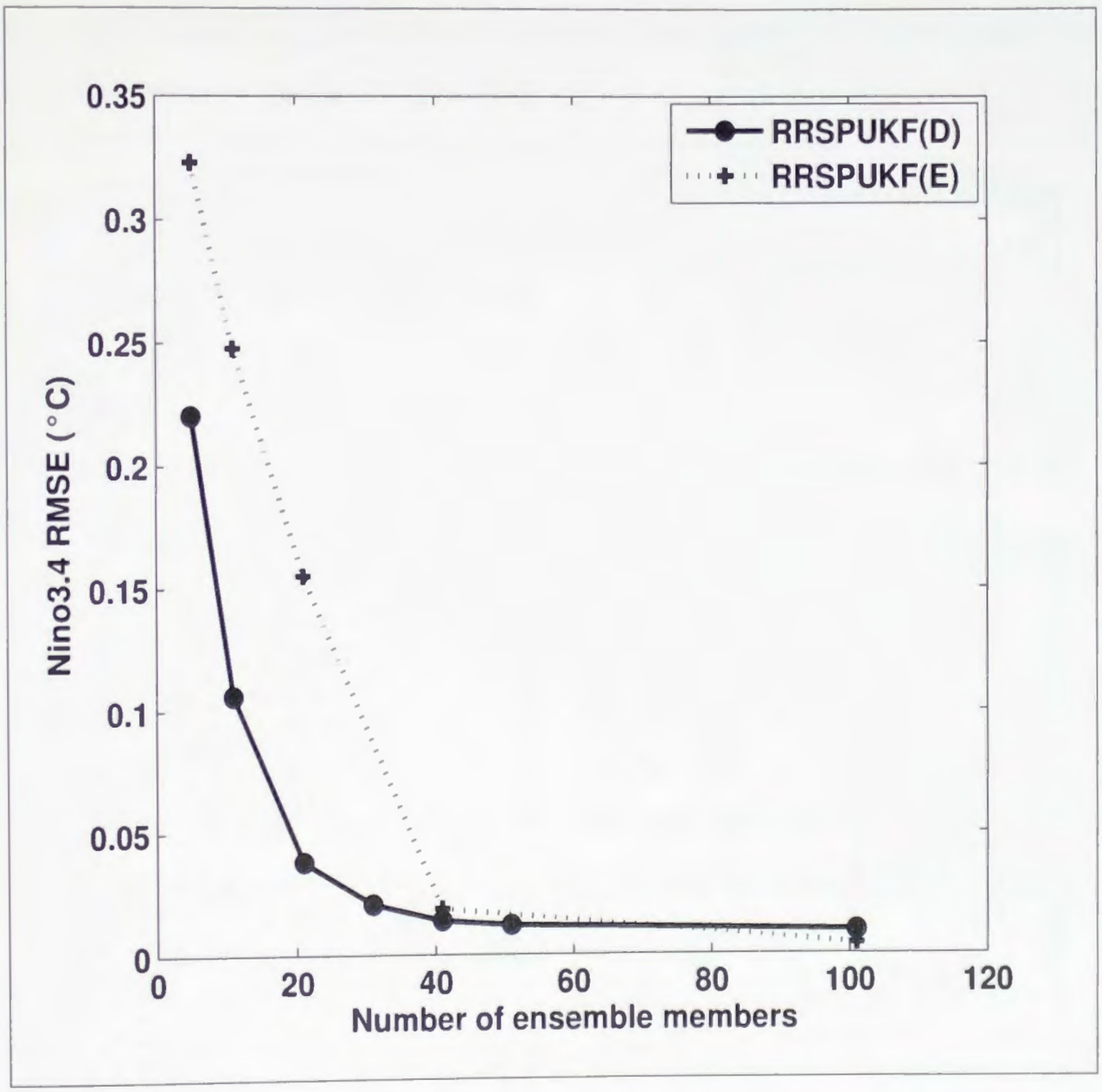

Figure 5.9: The root mean-squared error (RMSE) of SSTA (5°S-5°N, 120-170°W) between analysis and truth as a function of the number of ensembles used for the RRSPUKF(D) and RRSPUKF(E) methods.

for the period 1971–2000 when the ensemble size is 41. Figures 5.10e and 5.10f show the differences in RMSE and in correlation values between the two methods. Similar to Figure 5.6 and as discussed in Section 5.3, the bootstrap method is used to perform the statistical significance test to a 95% confidence level. There is no significant difference of correlation and RMSE skill between the two methods for almost

the entire domain of the assimilation, indicating that the skill of RRSPUKE(E) is entirely comparable to the RRSPUKF(D).

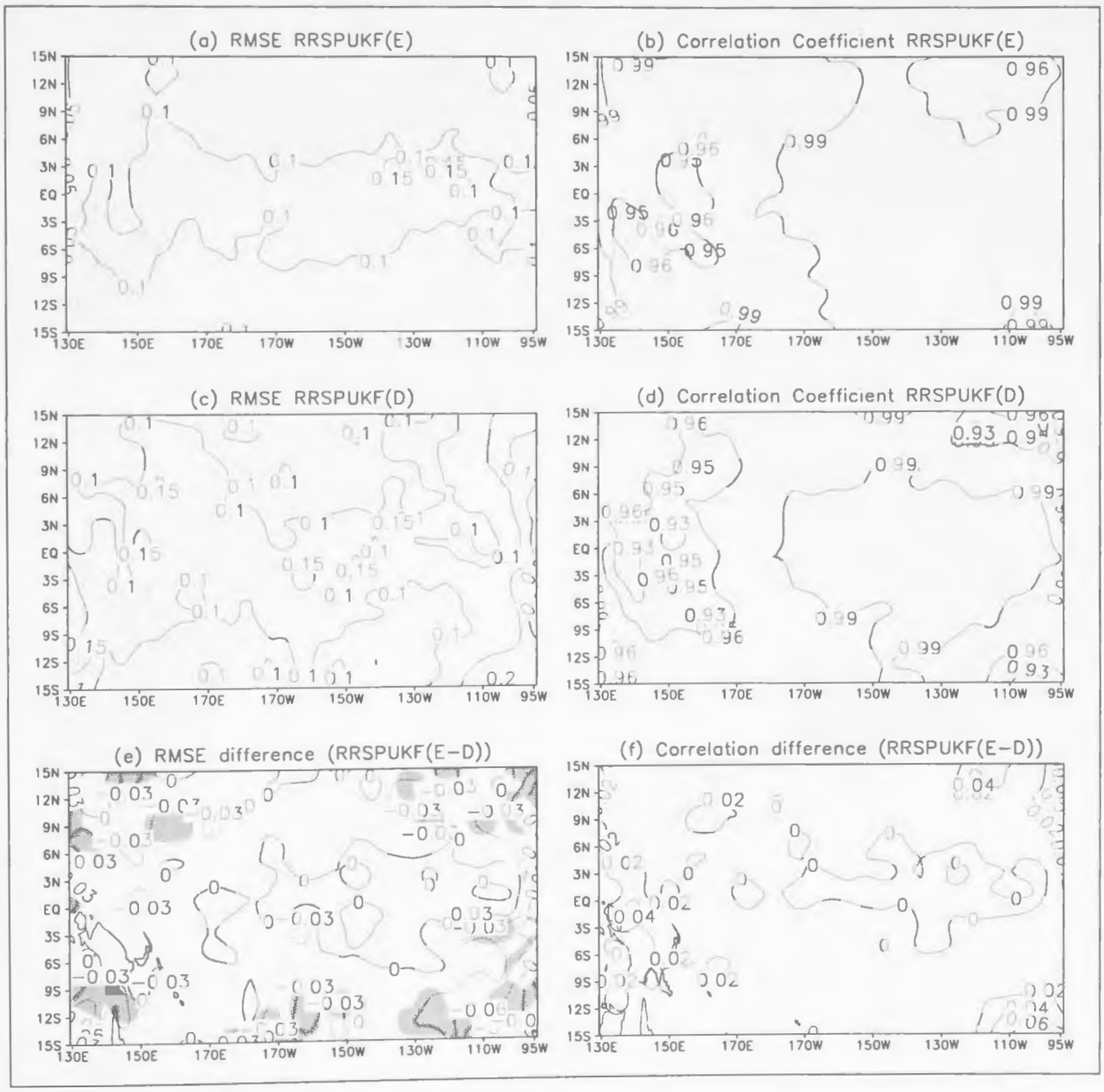

Figure 5.10: The RMSE and correlation coefficient of SSTA between the analysis and truth for RRSPUKF(E) and RRSPUKF(D) with 41 ensembles (a-d). e and f are the difference in RMSE and in correlation coefficient of SSTA between RRSPUKF(E) and RRSPUKF(D). The statistically significant areas (at 95% confidence level) are shaded.

Figure 5.11 shows the explained variance as a function of ensemble size for both RRSPUKF(E) and the RRSPUKF(D). The explained variance is larger in

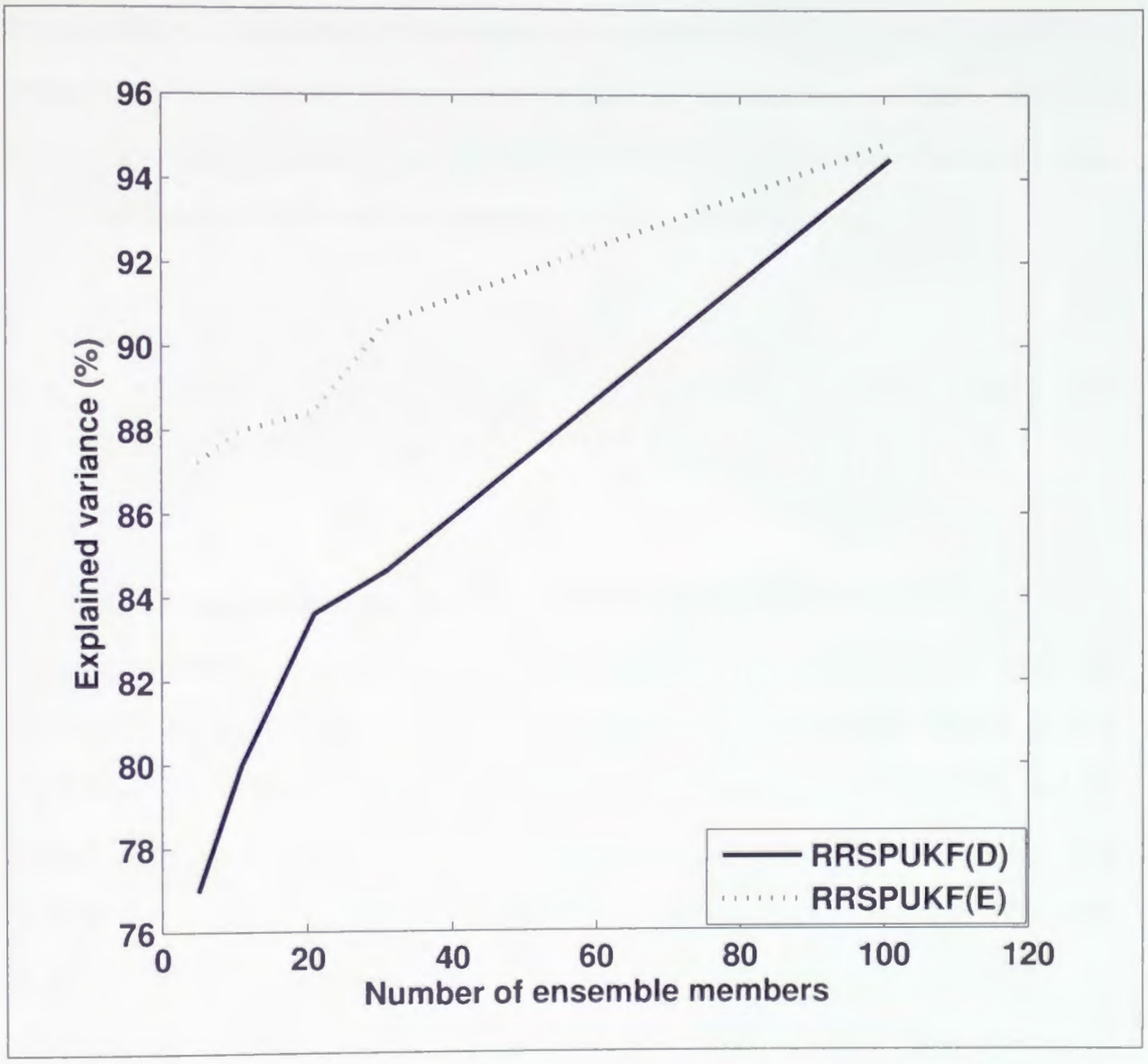

Figure 5.11: The explained variance (%) for the RRSPUKF(D and E) methods as a function of number of ensembles used.

RRSPUKF(E) than in the RRSPUKF(D) (Figure 5.11). This is interesting because the RRSPUKF(D) usually has better performance than the RRSPUKF(E) for the same ensemble size (Figure 5.9). One possible reason is that both methods have different total variances (denominators), making them incomparable. It should be noted that the RRSPUKF(E) only uses half of the ensemble (i.e., throwing out the other half) to calculate the covariance matrix to keep the same ensemble size of $2l+1$, given a prescribed number of l. Furthermore, because the ensemble size is not large

enough, there is inaccurate representation or underestimation for the prediction covariance. The SVD that is performed on such an inaccurate covariance matrix can lead to sub-optimal sigma points, suggesting that explained variance cannot be taken as the single entity to decide the number of sigma points to be used.

5.3.5 Comparison between the RRSPUKF(D) and EnSRF

Assimilation experiments are also performed using the EnSRF algorithm with the same experimental setup as those in RRSPUKF(D). In all experiments both the RRSPUKF(D) and EnSRF start from the same initial conditions. Shown in Figure 5.12(a) is the RMSE of the Niño-3.4 index analysis by RRSPUKF(D) and (b) EnSRF, both with ensemble size of 41 and against the observation counterpart. The analysis from two different data-assimilation methods can diverge from each other at certain assimilation steps along the assimilation track. It can be based on the ability of the assimilation algorithm in capturing the observation information and mix with the model in the transitioning states. If one particular data-assimilation method can approximate the error covariance better that method can better estimate the state at phase transition steps. Thus Figure 5.12 suggests that RRSPUKF(D) is probably better than EnSRF in the assimilation of some transition states using noisy observations. The RMSE and correlation over the entire basin is also compared in Figure 5.13, showing that the RRSPUKF(D) results in the smaller RMSE and slightly higher correlation skill than EnSRF, although their differences are not statistically significant for most of the domain (Figures 5.13e and 5.13f). The better performance of RRSPUKF over the EnSRF is probably because of its superior mathematical prop-

times less than 4-5 months. As the lead time increases the RMSE of RRSPUKF(D) and RRSPUKF(E) methods are significantly lower than that of EnSRF method but correlation skill are still comparable for all three methods, indicating the RMSE skill is more sensitive to the assimilation method.

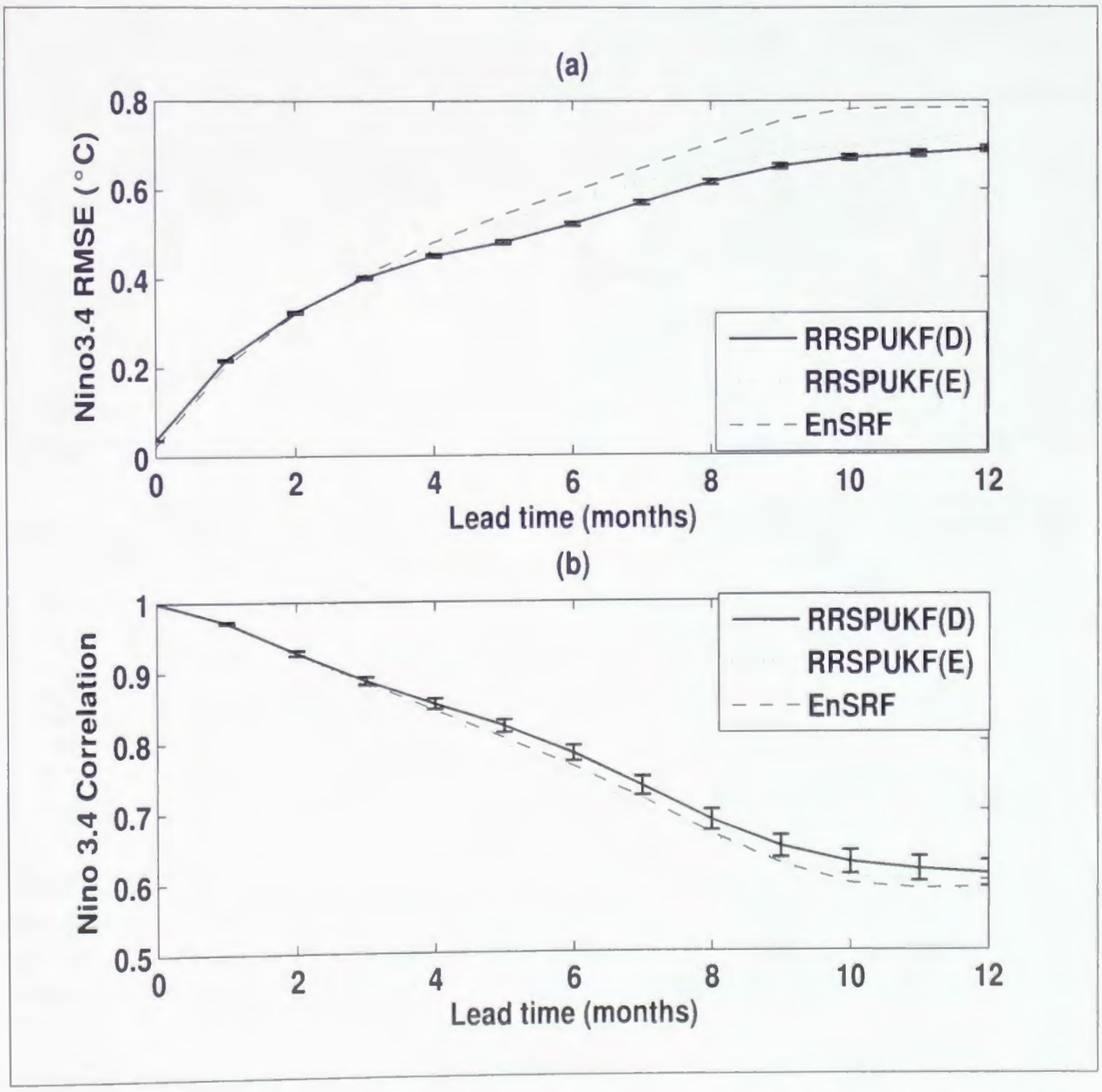

Figure 5.14: Niño-3.4 index forecast RMSE and correlation with 12 month lead time for RRSPUKF(E), RRSPUKF(D) and EnSRF. The vertical error bars are the sampling errors at 95% confidence interval obtained using bootstrap experiments at each lead time.

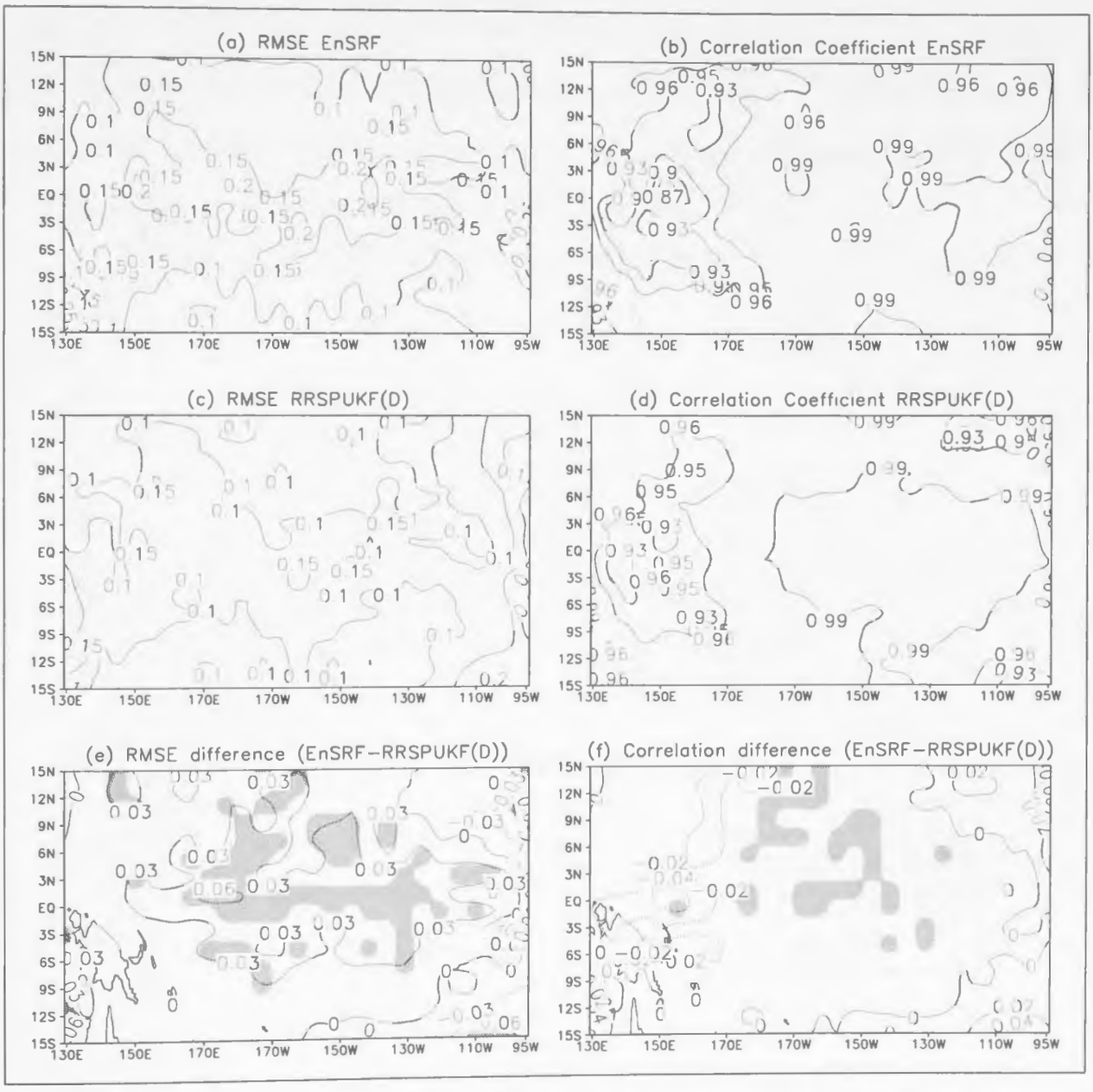

Figure 5.13: SSTA RMSE and correlation coefficient for EnSRF and RRSPUKF(D) method for the period 1971 2000 when 41 ensembles are used (a-d). e and f are the difference in RMSE and correlation coefficient of SSTA of EnSRF and RRSPUKF(D) where the statistically significant areas (at 95% confidence level) are shaded.

As in Figure 5.8, Figure 5.14 shows the hindcast experiments of Niño-3.4 SSTA initialized by the RRSPUKF(D), EnSRF and RRSPUKF(E). The hindcast settings are the same as those in Figure 5.8. Bootstrap experiments are performed to test the statistical significance for both RMSE and correlation skill. The experiment setting is similar to that described in Section 5.3. All three methods have similar skill at shorter lead

times less than 4-5 months. As the lead time increases the RMSE of RRSPUKF(D) and RRSPUKF(E) methods are significantly lower than that of EnSRF method but correlation skill are still comparable for all three methods, indicating the RMSE skill is more sensitive to the assimilation method.

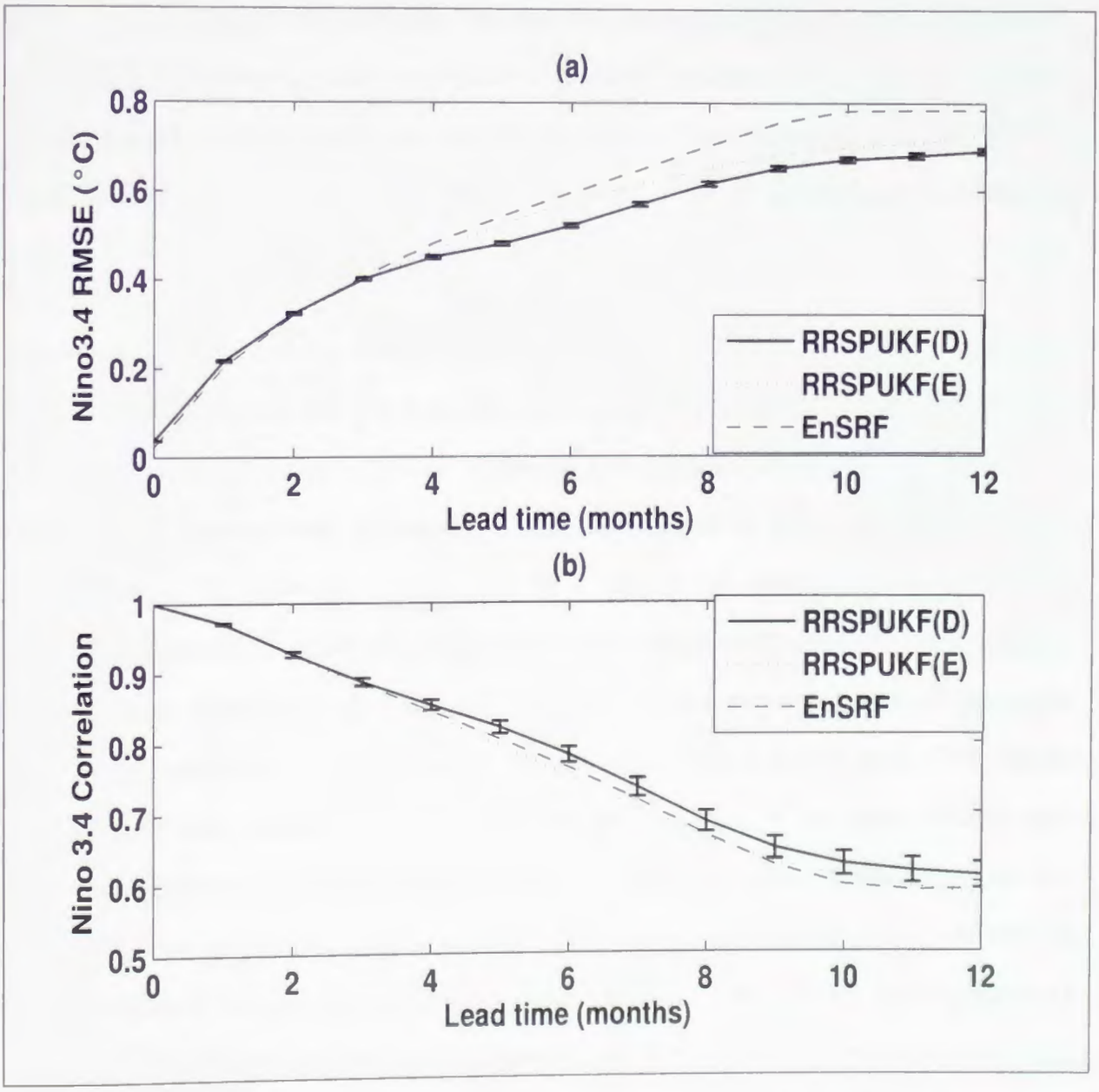

Figure 5.14: Niño-3.4 index forecast RMSE and correlation with 12 month lead time for RRSPUKF(E), RRSPUKF(D) and EnSRF. The vertical error bars are the sampling errors at 95% confidence interval obtained using bootstrap experiments at each lead time.

5.4 Summary

In this Chapter, both reduced-rank sigma-point unscented Kalman filter (Data) (RRSPUKF(D)) and reduced-rank sigma-point unscented Kalman filter (Ensemble) (RRSPUKF(E)) were applied on a realistic El Niño Southern Oscillation (ENSO) prediction model (LDEO5) and assimilated SST data. Four methods (RRSPUKF(D), RRSPUKF(E), SPUKF and EnSRF) were implemented and tested on the ZC model.

Sensitivity experiment were conducted using RRSPUKF(D) with varying number of truncated modes after the SVD on the covariance matrix. The results showed that when the number of sigma points increases the RMSE of Niño-3.4 SSTA analysis compared to observations decreases. When the number of sigma points were above 40, however, the estimation skill levels off. One of the reasons for higher RMSE when the ensemble number was small was due to underestimation of the covariance matrix by the finite truncated modes. The RMSE and correlation skill of the entire basin were compared as a function of the number of sigma points used. The results showed that the estimate had more error when the number of sigma points were small as evidenced from the higher RMSE and lower correlation coefficient. This was a case of filter divergence which usually occurs when the background covariance is underestimated because of the small number of ensembles. When the background covariance is underestimated there is greater certainty in the forecast estimate and the filter will give more weight to the forecast compared to observations.

The estimation skills of RRSPUKF(D) and RRSPUKF(E) were compared to each other. When the truncation reaches a trade-off between cost expense and estimation

accuracy, RRSPUKFs were able to analyze the phase and intensity of all major ENSO events from 1971–2000 with comparable estimation accuracy. For smaller ensemble size the RRSPUKF(D) performed better than RRSPUKF(E), suggesting that the latter was more inefficient in using small ensembles to capture the main features of the full covariance matrix. This was because the RRSPUKF(D) directly decomposes the full covariance matrix whereas the RRSPUKF(E) uses the ensemble to approximate the covariance matrix. As the number of sigma points increased, however, the difference between the two methods became minimal. When 41 sigma points were used there was no significant difference of correlation and RMSE skill between the two methods for almost the entire domain of the assimilation, indicating that the skill of RRSPUKE(E) was entirely comparable to the RRSPUKF(D). The main advantage of RRSPUKF(E) method over RRSPUKF(D) method is that the former method significantly reduces the computational cost and memory storage and makes the SPUKF method feasible for a high-dimensional system.

The estimation skills of RRSPUKF with 41 sigma points were very close to that of a full-rank sigma-point unscented Kalman filter (SPUKF). The RRSPUKF could fairly well simulate the performance of the full-rank SPUKF in terms of the measure of correlation skill, and also can generate small RMSE. The results showed that both the RRSPUKFs were more computationally efficient than the full-rank SPUKF, in spite of losing some estimation accuracy.

The estimation skill of ensemble square-root filter (EnSRF) had also compared to RRSPUKF, and the results showed that RRSPUKF was more robust than the latter. The better performance of RRSPUKF over the EnSRF was probably because of the superior mathematical properties of the SPKF algorithm as explained in Chapters 1 and 2. Namely, the RRSPUKF can better estimate the first- and second-order sta-

tistical moments using the SVD technique to approximate the known true statistical moments in contrast to an estimate of unknown true moments in EnSRF. As the RRSPUKF algorithm can better estimate the error covariance compared to EnSRF, it has better ability in capturing the observation information and mix with the model, thereby providing better state estimate. Furthermore, in RRSPUKF the truncation error can be quantified but in the case of EnSRF the selection of the number of ensembles is arbitrary and true prediction covariance is usually unknown (Manoj et al., 2014).

Chapter 6

Conclusions and Future directions

6.1 Introduction

This dissertation primarily addressed the issue of developing a practical scheme of sigma-point Kalman filter (SPKF) for its application in a realistic climate model. The SPKF assimilation system with its deterministic and robust mathematical properties is expected to produce a more reliable estimate of the state of the ocean-atmosphere system. To the best of my knowledge, this was the first trial in the assimilation community to apply SPKF in a realistic climate model. This chapter provides a concise synopsis of the work presented in this dissertation. Important contributions of this study are also discussed. There is also a section on possible future directions included at the end.

6.2 General overview

In Chapter 2, I discussed the theoretical formulation of the Kalman filter, ensemble Kalman filter (EnKF), ensemble square-root filter (EnSRF) and sigma-point unscented Kalman filter (SPUKF). The ensemble-based Kalman filters such as EnSRF have successfully demonstrated their ability in oceanic and atmospheric data assimilation (Deng et al., 2010, 2011). The main disadvantages of these methods are the uncertainty in constructing the ensembles and the requirement of the greater size of ensembles for the accurate approximation of the mean and variance of the prior and posterior distribution of model states (Kim et al., 2007; Chandrasekar et al., 2008; Ambadan and Tang, 2009). The SPUKF is relatively new in atmospheric and oceanographic data assimilation, which deterministically chooses ensemble members and thereby reduces uncertainty in generating ensembles (Julier, 2003). Also, the SPUKF makes use of a derivative-less optimal estimation technique that uses a minimal number of statistically weighted samples called sigma points to calculate the error statistics. When transformed through the nonlinear model they capture the true mean and covariance accurately to the second order in the Taylor series expansion (Julier, 2002). The SPUKF uses reformulated equations to find the Kalman gain and error covariance so that there is no need to calculate the Jacobian or tangent linear model (TLM) as in the case of extended Kalman filter (EKF); and there is no need to linearize the nonlinear measurement function as in the EnKF-based filters (Tang and Ambadan, 2009; Tang et al., 2014a).

In Chapter 3, I described the derivation of two reduced-rank methods, reduced-rank sigma-point unscented Kalman filter (Data) (RRSPUKF(D)) and reduced-rank sigma-point unscented Kalman filter (Ensemble) (RRSPUKF(E)). Despite all the merits of

SPUKF, the need for twice the number of sigma points than the system states remains a major computational concern for the application of SPUKF into realistic ocean or atmosphere models. Given the computational constraints of full-order error-covariance estimation, the data-assimilation community is very interested in the development of a reduced order scheme for the state estimation. The main objective of this thesis was to develop a reduced-rank sigma-point unscented Kalman filter, and to apply the reduced-rank filter to a realistic climate model. In the reduced-rank methods, the challenge of the computational constraint for SPUKF is successfully overcome by employing a truncated singular-value decomposition (TSVD) technique for dimensional compression or rank reduction of the error-covariance matrix. I applied both methods on the Lorenz-96 model (Lorenz and Emanuel, 1998) in the perfect model scenario. The results showed that the reduced-rank methods significantly decrease the number of sigma points and their estimation accuracy is close to that of a full-rank SPUKF.

In Chapter 4, I discussed the common problems in ensemble-based filters such as inbreeding, filter divergence and spurious correlation. I implemented a hybrid-localization scheme for RRSPUKF(E). In this work, I placed more emphasis on RRSPUKF(E) because with that method the analysis error-covariance matrix is estimated in the ensemble space, which avoids the estimation of analysis error-covariance matrix of model state space globally. In the localization scheme, the analysis was performed in a local subspace whose dimension may be greatly lower than the system dimension but the sigma points are generated based on the global domain. The main advantage of this scheme is that it not only makes the filter more computationally efficient, but also makes the analysis estimate more accurate by reducing the spurious correlation between remote grid points. In my experiments I used different system dimensions (N_x=40, 80 and 120) and the results showed that, to achieve the optimal

estimate (in terms of minimum RMSE), the number of sigma points needed to increase as the system dimension increases in the absence of localization. But when the localization was introduced, the number of sigma points required to achieve the minimum error mainly depends on the local dimension and is independent of the dimension of the global domain. The results of numerical experiments provide a good test-bed for the application of the hybrid-localization scheme in an actual atmospheric or oceanic model.

In Chapter 5, I applied both the RRSPUKF(D) and RRSPUKF(E) schemes to a realistic climate model. I used the LDEO5 of the Zebiak-Cane (ZC) model and assimilated sea surface temperature anomaly (SSTA) data. The performance of reduced-rank schemes was compared with that of the full-rank SPUKF. The sensitivity experiments results showed that both the RRSPUKF (D and E) schemes, even with a minimum number of sigma points as low as 41, were sufficient to get an analysis near to the full-rank SPUKF analysis. The ability of RRSPUKF (D and E) and EnSRF in terms of RMSE and correlation coefficients was compared. When the same number of ensembles were used, the performance of RRSPUKF was slightly better than the EnSRF, which suggests that the RRSPUKF scheme can obtain more accurate estimates because of its better algorithmic properties. My results showed that when the ensemble size is as small as 41, the performance of both schemes RRSPUKF (D and E) is comparable.

6.3 Concluding summary

The aim of data assimilation is to optimally estimate the state of the geophysical system from all available information, which are mainly the dynamical models and

observations (Talagrand, 1997). The output of the data-assimilation process known as the analysis is used as the initial condition for the future forecast. It can be also used to get more information about the area where the observation is sparse such as the deep layers of the ocean. Data assimilation is an essential component of the numerical weather prediction (NWP) process, where reasoning under uncertainty is a requirement. Methods based on recursive Bayesian estimation provides a valuable tool for the sequential state-optimization problems. Kalman filters give the best linear unbiased estimate (BLUE) for linear systems; however, its direct implementation is not practical for NWP problems because of its large dimensionality and non-linearity.

The ensemble formulation of the Kalman filter, EnKF and its derivatives are widely used in numerous geophysical data-assimilation problems because of their algorithmic simplicity and they are easy to implement. The main problem with EnKF and its advanced variants such as EnSRF is that there is no well designed technique to choose the optimal ensemble size. They also cannot provide information about the percentage of the estimated error variance related to the true counterpart for a given ensemble size. The SPUKF is a deterministic, derivative-less algorithm optimal for nonlinear estimation problems. SPUKF makes use of a novel deterministic sampling strategy based on the unscented transformation. The new ensembles called the sigma points can accurately describe the statistical moments of the nonlinear model. In spite of all the benefits of SPUKF, the requirement of more than twice the number of sigma points than the system states makes it impractical for their application in estimation problems in meteorology or oceanography. This necessitates the requirement of developing a practically implementable approximation for SPUKF. This dissertation has mainly focused on that issue and investigated two methods to approximate the error covariance by reduced-rank approximation and also proposed a localization scheme so

that SPKF can be applied to a realistic high-dimensional climate model. The main contribution of this dissertation is described in the following paragraphs.

I introduced and investigated two computationally efficient algorithms for sigma-point Kalman filters, making them viable candidates for application in realistic atmospheric and oceanic models. The reduced-rank approximation is mainly based on the concept that most large scale geophysical phenomena can be approximated by a finite number of degrees of freedom and their dominant variability can be explained by a limited number of modes. Assuming that most important errors of the original sigma-point space can be estimated using a fewer number of the dominant sigma points, two reduced-rank approximation methods are proposed in this work. The schemes are reduced-rank sigma-point unscented Kalman filter (Data) (RRSPUKF(D)) and reduced-rank sigma-point unscented Kalman filter (Ensemble) (RRSPUKF(E)). Both methods are based on the truncated SVD to factorize the covariance matrix and reduce its rank through truncation. In the first technique, the SVD is applied on the error covariance matrix calculated in the data space (RRSPUKF(D)) whereas in the second technique the SVD is applied on the error covariance matrix calculated in the ensemble space (RRSPUKF(E)). The reduced-rank, square-root matrix is used to select the most important sigma points that can retain the main statistical features of the original sigma points.

The RRSPUKF(D) can effectively save computational time through reducing the number of sigma points. The saving is especially significant when the error covariance in the data space, P_t^a can be approximated well by a few leading modes so that the fewer number of sigma points can approximate the statistical moments with a good truncation accuracy. However, a large challenge in RRSPUKF(D) is the computation of P_t^a itself. When the state dimension is large, it is computationally expensive to

estimate the error covariance in the data space, even difficult to compute and store such a big matrix. When the state dimension L_M is much greater than the ensemble dimension n, i.e $L_M \gg n$, it is possible to compute the SVD on the nXn covariance matrix constructed in the ensemble subspace (von Storch and Hannoschock, 1984; Wilks, 2011), which is computationally affordable for a system with large dimension. The computational complexity of the algorithm to find the covariance matrix in the data space is $O(L_M^2 \times n)$; however, the mathematical complexity of the algorithm to find the covariance matrix in the ensemble subspace is $O(n^2 \times L_M)$. The difference in computational complexity between the two algorithms is significant when $L_M \gg n$, that is usually the case for a realistic NWP model. In the RRSPUKF(E) method TSVD is applied on the analysis error-covariance matrix P_t^{aE} in the ensemble subspace. Both the new algorithms are first tested using the Lorenz 96 (Lorenz and Emanuel, 1998) model, which is a one-dimensional atmospheric toy model, as test-bed. The estimation skills of both the rank-reduction methods were near to that of full-rank SPUKF.

I proposed and tested a hybrid-localization scheme for RRSPUKF(E) method. The localization is very important in the case of RRSPUKFs for its practical application in a global circulation model (GCM) because we are approximating the full-rank covariance with its reduced-rank approximation. Localization makes the filter more computationally feasible and makes the analysis estimate optimal by reducing the problems of spurious correlation, inbreeding and filter divergence because of smaller ensemble size. The main challenge for implementing the localization in the SPKF is that the sigma points are generated based on the global analysis-error covariance. In this dissertation I developed a localization scheme in that the analysis is performed locally and the sigma-points are generated globally. The numerical experiments results from the Lorenz-96 model showed that when the localization and inflation are

introduced, the optimal estimate obtained is independent of the system dimension. The sensitivity experiment results suggested possible ways to improve the filter performance. The localization scheme gave encouraging results making the RRSPUKF(E) method a suitable candidate for assimilation in a high-dimensional system.

I explored the application of SPUKF and RRSPUKF(D and E) in a realistic climate model of intermediate complexity. To the best of my knowledge, this was the first time SPKF was applied to a realistic model in the atmospheric/oceanic sciences. The realistic model used in this study was the LDEO5 version of the ZC model for ENSO prediction. Emphasis was placed on the implementation of RRSPUKF for the LDEO5 with the assimilation of SSTA. Both the RRSPUKFs could analyze well the phase and intensity of all major ENSO events during the study period with comparable estimation accuracy. Furthermore, the RRSPUKF is compared against the EnSRF, showing that the overall analysis skill of RRSPUKF and EnSRF are comparable to each other but the former is more robust than the latter. The sensitivity experiment result showed that RRSPUKF with 41 ensembles can have estimation accuracy close to that of a full-rank SPUKF, making it a potential candidate for data assimilation in large-dimensional atmospheric or ocean GCMs.

The advantage of RRSPUKF(E) over RRSPUKF(D) is in its affordable computational cost and lower memory usage. The main challenge of RRSPUKF(D) is to represent the full analysis-covariance matrix explicitly, which is still computationally unaffordable for high dimensional models. In the case of RRSPUKF(E) the ensembles are used to approximate the covariance matrix, calculated in the reduced dimension (ensemble space). The localization scheme makes RRSPUKF(E) more computationally efficient because it may reduce the analysis errors from spurious correlation, inbreeding and filter divergence, making it an ideal candidate for implementation on a real weather

model.

My goal was to develop a practical scheme for the SPKF for their application in high-dimensional systems. The studies conducted by Tang et al. (2014a) have shown that the Kalman filter equations used in SPKFs have better estimation accuracy when the measurement function is nonlinear compared to the one used in EnKF and En-SRFs. The direct application of the nonlinear measurement function in EnKF and EnSRF imposes an implicit linearization process using ensemble members. The performance of RRSPUKFs were more robust than EnSRF when applied for assimilating SSTA into ZC model (Manoj et al., 2014). In RRSPUKFs the ensembles are generated deterministically (Manoj et al., 2014; Tang et al., 2014b) compared to the arbitrary nature of initial ensemble generation in EnKFs (Liang, 2007; Curn et al., 2012). RRSPUKFs can generate finite number of sigma points for the optimal estimate of prediction-error covariance(Manoj et al., 2014; Tang et al., 2014b). The implementation of localization and inflation schemes in RRSPUKF(E) resulted in its better performance. The RMSE was independent of system dimension for a simple model such as the Lorenz-96 model similar to the findings of Ott et al. (2004). Similar to the LETKF data-assimilation system (Miyoshi, 2005; Li et al., 2009) RRSPUKF(E) has an advantage for parallel computation making it the best candidate for application in systems of high-dimensionality.

6.4 Future directions

A couple of key challenges related to the SPKF data-assimilation methods have been identified. The first challenge is the potential application of RRSPUKF(E) in an atmospheric or oceanic GCM. The second issue concerns the inability of ensemble

based filters such as SPKF to assimilate asynchronous observations. Another issue is the assumption of Gaussian distribution for the observation and background errors in SPKF. In the following subsections I briefly discuss these issues and also some possible ways that they may be addressed.

6.4.1 Implementation of RRSPUKF(E) for a GCM

In this dissertation I explored practical algorithms for SPKF to make it suitable for application in an atmospheric or oceanic GCM. The RRSPUKF(E) method with the hybrid-localization scheme proposed in this dissertation showed estimation accuracy close to that of a full-rank SPUKF making it an ideal candidate for implementation into a GCM. In this study RRSPUKF(E) is applied on a number of models with varying complexity and the results indicated that the method is very effective in assimilating data. However, the method has some limitations as well for application into a GCM in an operational setting. When we develop an optimal realistic RRSPUKF(E) data-assimilation system for an ocean or atmosphere GCM, there are many important issues that need to be addressed. The model resolution, number of variables and the observations to be assimilated will be much larger for a GCM than what we have encountered so far. Designing the model error, observation error and initial perturbation will be much harder in the case of a GCM. Another important factor is the efficiency of computation. Unlike in simple models we cannot afford to conduct so many sensitivity experiments with varying parameters to choose the optimal parameter settings in the case of a GCM. The results from the experiments with simple models can be used as a guideline for the application of RRSPUKF(E) in a complex GCM.

6.4.2 A hybrid assimilation by coupling 3D-Var/4D-Var + SPKF

The main advantage of advanced ensemble based methods such as SPKF over variational methods is that, in the former the estimate of the background-covariance matrix is flow-dependent or it estimates the "errors of the day". One of the primary advantage of 4D-Var assimilation method is its ability to assimilate asynchronous observations. The coupling of sequential evolution of covariance matrices with variational data-assimilation methods is an active area of research in the data-assimilation community (Zhang and Zhang, 2012; Zhang et al., 2013). Recent studies have shown that the data-assimilation system combining EnKF with 4D-Var performed better than uncoupled 4D-Var or EnKF (Zhang and Zhang, 2012).

Most of the studies on coupling ensemble-based data-assimilation methods with variational methods use EnKF as a representative of the ensemble-based data-assimilation method. As discussed in this dissertation, SPKF has superior mathematical properties over EnKF and it will be interesting to explore a hybrid data-assimilation system coupling the advanced SPKF such as RRSPUKF(E) with the 4D-Var method. Following Zhang and Zhang (2012), the basic idea of coupling is that, the explicit covariance information from the sigma-point forecasts are used in the minimization of the cost function of 4D-Var. The 4D-Var analysis trajectory is used as the mean of the SPKF analysis. Positive and negative pair of analysis perturbations are generated according to the SPKF scheme for the next cycle of forecasts. In the hybrid approach, both the SPKF and 4D-Var will run in parallel while benefiting from the exchange of information with each other at discrete time intervals. The main challenge of coupling SPKF with 4D-Var will be the computational cost.

6.4.3 Non-Gaussian statistics in SPKF

Kalman filter data-assimilation methods give the optimal estimate, BLUE, when the models as well as the measurement operator that relates model states to observations are linear and the associated background and observation errors are Gaussian (Ambadan and Tang, 2011; van Leeuwen, 2012; Ambadan, 2013). When we move away from the assumption of linearity and Gaussianity, that is often the case with the real-world problems, the optimality of the estimate obtained from the ensemble becomes questionable (van Leeuwen, 2012). The assumption of Gaussianity does not hold for short term (daily or weekly) averages of many atmospheric variables (Sura and Hannachi, 2014). Recent studies have shown that multiplicative (state-dependent) noise models can be used to account for the internal (dynamical and numerical) and external (observation) errors and their associated deviations from Gaussian probability distributions (Ambadan and Tang, 2011; Sura and Hannachi, 2014). The challenges of multiplicative noise can be solved with a hybrid data-assimilation method of a particle filter and SPKF, known as sigma-point particle filter (SPPF) (Van der Merwe, 2004; Ambadan and Tang, 2011).

SPPF makes use of the important properties of both the particle filter and SPKF. Unlike the Kalman filters, the particle filter solves the complete nonlinear data-assimilation problem and the probability-density function is fully propagated in time. Furthermore, a particle filter does not need artificial techniques like covariance inflation and localization to obtain optimal estimate for a high-dimensional system. The main disadvantage of the particle filter is filter degeneracy; this can be solved by introducing a good resampling strategy (van Leeuwen, 2012). In SPPF, SPKF technology is used for sequential important resampling (SIR). For application in a

high dimensional system. SPKF resampling technology is not computationally feasible. The data assimilation community is very interested in developing a hybrid particle filter, which can use the RRSPUKF Ev scheme for the SIR.

Bibliography

Ambadan, J. T. (2013). *Evaluation of Stochastic Kinetic Energy Backscatter*. PhD thesis, International Max Planck Research School on Earth System Modelling, Hamburg, Germany.

Ambadan, J. T. and Tang, Y. (2009). Sigma-point kalman filter data assimilation methods for strongly nonlinear systems. *Journal of the Atmospheric Sciences*, 66:261–285.

Ambadan, J. T. and Tang, Y. (2011). Sigma-point particle filter for parameter estimation in a multiplicative noise environment. *Journal of Advances in Modeling Earth Systems*, 3:1–16.

Anderson, J. L. (2001). An ensemble adjustment kalman filter for data assimilation. *Monthly Weather Review*, 129:2884–2903.

Anderson, J. L. and Anderson, S. L. (1999). A monte carlo implementation of the nonlinear filtering problem to produce ensemble assimilations and forecasts. *Monthly Weather Review*, 127:2741–2758.

Auroux, D. and Blum, J. (2008). A nudging-based data assimilation method: the back and forth nudging (bfn) algorithm. *Nonlinear Processes in Geophysics*, 15:305–319.

Battisti, D. S. (1988). Dynamics and thermodynamics of a warming event in a coupled tropical atmosphere–ocean model. *Journal of the Atmospheric Sciences*, 45:2889–2919.

Bertino, L., Evensen, G., and Wackernagel, H. (2003). Sequential data assimilation techniques in oceanography. *International Statistical Review*, 71:223–241.

Bishop, C. H., Etherton, B. J., and Majumdar, S. J. (2001). Adaptive sampling with the ensemble transform kalman filter. Part 1: Theoretical aspects. *Monthly Weather*

Review, 129:420–436.

Burgers, G., van Leeuwen, P. J., and Evensen, G. (1998). Analysis scheme in the ensemble kalman filter. *Monthly Weather Review*, 126:1719–1724.

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

Chandrasekar, J., Kim, I. S., Bernstein, D. S., and Ridley, A. J. (2008). Reduced-rank unscented kalman filtering using cholesky-based decomposition. *International Journal of Control*, 81:1779–1792.

Chen, D., Cane, M. A., Kaplan, A., Zebiak, S. E., and Huang, D. (2004). Predictability of El Niño over the past 148 years. *Nature*, 428:733–736.

Chen, Y. and Snyder, C. (2007). Assimilating vortex position with an ensemble kalman filter. *Monthly Weather Review*, 135:1828–1845.

Cheng, Y., Tang, Y., Zhou, X., Jackson, P., and Chen, D. (2010). Further analysis of singular vector and ENSO predictability in the lamont model - Part 1: singular vector and the control factors. *Climate Dynamics*, 35:807–826.

Corazza, M., Kalnay, E., and Yang, S. C. (2007). An implementation of the local ensemble kalman filter in a quasi geostrophic model and comparison with 3d-var. *Nonlinear Processes in Geophysics*, 14(1):89–101.

Courtier, P., Andersson, E., Heckley, W., Vasiljevic, D., Hamrud, M., Hollingsworth, A., Rabier, F., Fisher, M., and Pailleux, J. (1998). The ECMWF implementation of three-dimensional variational assimilation (3d-var). i: Formulation. *Quarterly Journal of the Royal Meteorological Society*, 124:1783–1807.

Courtier, P., Thépaut, J.-N., and Hollingsworth, A. (1994). A strategy for operational implementation of 4d-var, using an incremental approach. *Quarterly Journal of the Royal Meteorological Society*, 120:1367–1387.

Curn, J., Marinescu, D., Lacey, G., and Cahill, V. (2012). Estimation with non-white gaussian observation noise using a generalised ensemble kalman filter. In *Robotic and Sensors Environments (ROSE), 2012 IEEE International Symposium*, pages 85–90, Magdeburg, Germany.

Daley, R. (1991). *Atmospheric Data Analysis*. Cambridge University Press, Cambridge, United Kingdom.

Deng, Z. and Tang, Y. (2009). Reconstructing the past wind stresses over the tropical

Pacific Ocean from 1875 to 1947. *Journal of Applied Meteorology and Climatology*, 48:1181–1198.

Deng, Z., Tang, Y., Chen, D., and Wang, G. (2012). A time-averaged covariance method in the EnKF for Argo data assimilation. *Atmosphere-Ocean*, 50:129–145.

Deng, Z., Tang, Y., and Freeland, H. J. (2011). Evaluation of several model error schemes in the EnKF assimilation: Applied to Argo profiles in the Pacific Ocean. *Journal of Geophysical Research: Oceans*, 116:doi:10.1029/2011JC006942.

Deng, Z., Tang, Y., and Wang, G. (2010). Assimilation of Argo temperature and salinity profiles using a bias-aware localized EnKF system for the Pacific Ocean. *Ocean Modelling*, 35:187 – 205.

Déry, S. J. and Wood, E. F. (2005). Decreasing river discharge in northern Canada. *Geophysical Research Letters*, 32:doi:10.1029/2005GL022845.

Dimet, F.-X. L. and Talagrand, O. (1986). Variational algorithms for analysis and assimilation of meteorological observations: theoretical aspects. *Tellus A*, 38A:97–110.

Ehrendorfer, M. (2007). A review of issues in ensemble-based kalman filtering. *Meteorologische Zeitschrift*, 16:795–818.

Evensen, G. (1992). Using the extended kalman filter with a multilayer quasi-geostrophic ocean model. *Journal of Geophysical Research: Oceans*, 97:17905–17924.

Evensen, G. (1994). Sequential data assimilation with a nonlinear quasi-geostrophic model using monte carlo methods to forecast error statistics. *Journal of Geophysical Research: Oceans*, 99:10143–10162.

Evensen, G. (1997). Advanced data assimilation for strongly nonlinear dynamics. *Monthly Weather Review*, 125:1342–1354.

Evensen, G. (2003). The ensemble kalman filter: theoretical formulation and practical implementation. *Ocean Dynamics*, 53:343–367.

Evensen, G. (2009). *Data assimilation, The Ensemble Kalman Filter, 2nd ed.* Springer-Verlag, Berlin, Germany.

Fan, C. and You, Z. (2009). Highly efficient sigma point filter for spacecraft attitude and rate estimation. *Mathematical Problems in Engineering*, 2009:doi:10.1155/2009/507370.

Fukumori, I. (2012). Combining models and data in large-scale oceanography: examples from the consortium for estimating the circulation and climate of the ocean (ecco). In Blayo, E., Bocquet, M., Cosme, E., and Cugliandolo, L. F., editors, *Advanced Data Assimilation for Geoscience*, pages 535–551, Oxford, United Kingdom. Oxford University Press.

Furrer, R. and Bengtsson, T. (2007). Estimation of high-dimensional prior and posterior covariance matrices in kalman filter variants. *Journal of Multivariate Analysis*, 98:227 – 255.

Gaspari, G. and Cohn, S. E. (1999). Construction of correlation functions in two and three dimensions. *Quarterly Journal of the Royal Meteorological Society*, 125:723–757.

Gauthier, P., Courtier, P., and Moll, P. (1993). Assimilation of simulated wind lidar data with a kalman filter. *Monthly Weather Review*, 121:1803–1820.

Gill, A. E. (1980). Some simple solutions for heat-induced tropical circulation. *Quarterly Journal of the Royal Meteorological Society*, 106:447–462.

Hamill, T. (2006). Ensemble based atmospheric data assimilation. In *Predictability of Weather and Climate*, pages 124–156. Cambridge University Press, Cambridge, United Kingdom.

Hamill, T. M. and Whitaker, J. S. (2005). Accounting for the error due to unresolved scales in ensemble data assimilation: A comparison of different approaches. *Monthly Weather Review*, 133:3132–3147.

Hamill, T. M., Whitaker, J. S., and Snyder, C. (2001). Distance-dependent filtering of background error covariance estimates in an ensemble kalman filter. *Monthly Weather Review*, 129:2776–2790.

Han, X. and Li, X. (2008). An evaluation of the nonlinear/non-gaussian filters for the sequential data assimilation. *Remote Sensing of Environment*, 112:1434–1449.

Horn, R. (1990). The hadamard product. Matrix theory and applications. In *American Mathematical society, Proceedings of Symposia in Applied Mathematics*, pages 87–169, Providence, United States.

Houtekamer, P. (2012). Short-range error statistics in an ensemble kalman filter. In Blayo, E., Bocquet, M., Cosme, E., and Cugliandolo, L. F., editors, *Advanced Data Assimilation for Geoscience*, pages 267–278, Oxford, United Kingdom. Oxford University Press.

Houtekamer, P. L. and Mitchell, H. L. (1998). Data assimilation using an ensemble kalman filter technique. *Monthly Weather Review*, 126:796–811.

Houtekamer, P. L. and Mitchell, H. L. (2001). A sequential ensemble kalman filter for atmospheric data assimilation. *Monthly Weather Review*, 129:123–137.

Hsieh, W. W., Tang, B., and Garnett, E. (1999). Teleconnections between Pacific sea surface temperatures and Canadian prairie wheat yield. *Agricultural and Forest Meteorology*, 96:209–217.

Hu, X.-M., Zhang, F., and Nielsen-Gammon, J. W. (2010). Ensemble-based simultaneous state and parameter estimation for treatment of mesoscale model error: A real-data study. *Geophysical Research Letters*, 37:doi:10.1029/2010GL043017.

Hunt, B. R., Kostelich, E. J., and Szunyogh, I. (2007). Efficient data assimilation for spatiotemporal chaos: A local ensemble transform kalman filter. *Physica D: Nonlinear Phenomena*, 230:112–126.

Ide, K., Courtier, P., Ghil, M., and Lorenc, A. C. (1997). Unified notation for data assimilation : Operational, sequential and variational. *Journal of the Meteorological Society of Japan. Ser. II*, 75:181–189.

Ito, K. and Xiong, K. (2000). Gaussian filters for nonlinear filtering problems. *Automatic Control, IEEE Transactions*, 45:910–927.

Jiang, C., Gille, S. T., Sprintall, J., Yoshimura, K., and Kanamitsu, M. (2012). Spatial variation in turbulent heat fluxes in Drake passage. *Journal of Climate*, 25:1470–1488.

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

Jöckel, P. (2012). Earth system modeling. In Schumann, U., editor, *Atmospheric Physics*, Research Topics in Aerospace, pages 577–590. Springer Berlin Heidelberg, Berlin, Germany.

Julier, S. (2002). The scaled unscented transformation. *Proceedings of the American Control Conference*, 6:4555–4559.

Julier, S. (2003). The spherical simplex unscented transformation. *Proceedings of the American Control Conference*, 3:2430–2434.

Julier, S. and Uhlmann, J. (2002). Reduced sigma point filters for the propagation of means and covariances through nonlinear transformations. *Proceedings of the*

American Control Conference, 2:887–892.

Julier, S. and Uhlmann, J. (2004). Unscented filtering and nonlinear estimation. *Proceedings of the IEEE*, 92:401–422.

Julier, S., Uhlmann, J., and Durrant-Whyte, H. (1995). A new approach for filtering nonlinear systems. *Proceedings of the American Control Conference*, 3:1628–1632.

Julier, S. J. and Uhlmann, J. K. (1997). New extension of the kalman filter to nonlinear systems. *Proc. SPIE 3068, Signal Processing, Sensor Fusion, and Target Recognition VI*, 3068:182–193.

Kalman, R. E. (1960). A new approach to linear filtering and prediction problems. *Trans. American Society of Mechanical Engineers (ASME), Journal of Basic Engineering*, 82:35–45.

Kalman, R. E. and Bucy, R. S. (1961). New results in linear filtering and prediction theory. *Trans. American Society of Mechanical Engineers (ASME), Journal of Basic Engineering*, 83:95–107.

Kalnay, E. (2003). *Atmospheric modelling, data assimilation and predictability*. Cambridge University Press, Cambridge, United Kingdom.

Kalnay, E., Ota, Y., Miyoshi, T., and Liu, J. (2012). A simpler formulation of forecast sensitivity to observations: application to ensemble kalman filters. *Tellus A*, 64:doi:10.3402/tellusa.v64i0.18462.

Kaplan, A., Cane, M. A., Kushnir, Y., Clement, A. C., Blumenthal, M. B., and Rajagopalan, B. (1998). Analyses of global sea surface temperature 1856–1991. *Journal of Geophysical Research: Oceans*, 103:18567–18589.

Karspeck, A. R. and Anderson, J. L. (2007). Experimental implementation of an ensemble adjustment filter for an intermediate ENSO model. *Journal of Climate*, 20:4638–4658.

Kim, I., Chandrasekar, J., Palanthandalam-Madapusi, H., Ridley, A., and Bernstein, D. (2007). State estimation for large-scale systems based on reduced-order error-covariance propagation. *American Control Conference, 2007.*, 2007:5700–5705.

Kirtman, B. P. and Schopf, P. S. (1998). Decadal variability in ENSO predictability and prediction. *Journal of Climate*, 11:2804–2822.

Kivman, G. A. (2003). Sequential parameter estimation for stochastic systems. *Nonlinear Processes in Geophysics*, 10:253–259.

Lermusiaux, P. F. J. (1997). *Error subspace data assimilation methods for ocean field estimation: Theory validation and applications.* PhD thesis, Harvard University, Cambridge, United States.

Lermusiaux, P. F. J. and Robinson, A. R. (1999). Data assimilation via error subspace statistical estimation. Part I: Theory and schemes. *Monthly Weather Review*, 127:1385–1407.

Li, H., Kalnay, E., Miyoshi, T., and Danforth, C. M. (2009). Accounting for model errors in ensemble data assimilation. *Monthly Weather Review*, 137:3407–3419.

Li, J. and Xiu, D. (2008). On numerical properties of the ensemble kalman filter for data assimilation. *Computer Methods in Applied Mechanics and Engineering*, 197:3574–3583. Stochastic Modeling of Multiscale and Multiphysics Problems.

Liang, B. (2007). *An ensemble Kalman filter module for automatic history matching.* PhD thesis, The University of Texas at Austin.

Lorenc, A. C. (2003). The potential of the ensemble kalman filter for NWP-a comparison with 4D-var. *Quarterly Journal of the Royal Meteorological Society*, 129:3183–3203.

Lorenz, E. N. (1963). Deterministic nonperiodic flow. *Journal of the Atmospheric Sciences*, 20:130–141.

Lorenz, E. N. (1965). A study of the predictability of a 28-variable atmospheric model. *Tellus*, 17:321–333.

Lorenz, E. N. (1996). Predictability: a problem partly solved. *in Proc. Seminar on Predictability (ECMWF, Reading, Berkshire, UK)*, 1:1–18.

Lorenz, E. N. (2006). Predictability: a problem partly solved. In *Predictability of Weather and Climate*, pages 40–58. Cambridge University Press, Cambridge, United Kingdom.

Lorenz, E. N. and Emanuel, K. A. (1998). Optimal sites for supplementary weather observations: Simulation with a small model. *Journal of the Atmospheric Sciences*, 55:399–414.

Luo, X. and Moroz, I. (2009). Ensemble kalman filter with the unscented transform. *Physica D: Nonlinear Phenomena*, 238:549 – 562.

Manoj, K. K., Tang, Y., Deng, Z., Chen, D., and Cheng, Y. (2014). Reduced-rank sigma-point kalman filter and its application in ENSO model. *Journal of Atmospheric*

and Oceanic Technology, 31:2350–2366.

Masutani, M., Woollen, J. S., Lord, S. J., Emmitt, G. D., Kleespies, T. J., Wood, S. A., Greco, S., Sun, H., Terry, J., Kapoor, V., Treadon, R., and Campana, K. A. (2010). Observing system simulation experiments at the national centers for environmental prediction. *Journal of Geophysical Research: Atmospheres*, 115:doi:10.1029/2009JD012528. D07101.

Miller, R. N., Carter, E. F., and Blue, S. T. (1999). Data assimilation into nonlinear stochastic models. *Tellus A*, 51:167–194.

Miller, R. N., Ghil, M., and Gauthiez, F. (1994). Advanced data assimilation in strongly nonlinear dynamical systems. *Journal of the Atmospheric Sciences*, 51:1037–1056.

Mitchell, H. L., Houtekamer, P. L., and Pellerin, G. (2002). Ensemble size, balance, and model-error representation in an ensemble kalman filter. *Monthly Weather Review*, 130:2791–2808.

Miyoshi, T. (2005). *Ensemble Kalman filter experiments with a primitive-equation global model.* PhD thesis, University of Maryland, College Park, United States.

Nerger, L. (2003). *Parallel Filter Algorithms for Data Assimilation in Oceanography.* PhD thesis, University of Bremen, Bremen, Germany.

Nørgaard, M., Poulsen, N. K., and Ravn, O. (1998). Advances in derivative-free state estimation for nonlinear systems. In *IMM-Technical Report-1998-15*. Informatics and Mathematical Modelling, Technical University of Denmark, Lyngby, Denmark.

Nørgaard, M., Poulsen, N. K., and Ravn, O. (2000). New developments in state estimation for nonlinear systems. *Automatica*, 36:1627–1638.

Oke, P., Sakov, P., and Corney, S. (2007). Impacts of localisation in the EnKF and EnOI: experiments with a small model. *Ocean Dynamics*, 57:32–45.

Oke, P. R., Brassington, G. B., Griffin, D. A., and Schiller, A. (2008). The bluelink ocean data assimilation system (bodas). *Ocean Modelling*, 21:46–70.

Oke, P. R., Schiller, A., Griffin, D. A., and Brassington, G. B. (2005). Ensemble data assimilation for an eddy-resolving ocean model of the Australian region. *Quarterly Journal of the Royal Meteorological Society*, 131:3301–3311.

O'Neill, A. (2004). Making the most of earth observation with data assimilation. *ESA bulletin*, 118:33–38.

Ott, E., Hunt, B., Szunyogh, I., Zimin, A., Kostelich, E., Corazza, M., Kalnay, E., Patil, D., and Yorke, J. (2004). A local ensemble kalman filter for atmospheric data assimilation. *Tellus A*, 56:415–428.

Palmer, T. and Zanna, L. (2013). Singular vectors, predictability and ensemble forecasting for weather and climate. *Journal of Physics A: Mathematical and Theoretical*, 46:doi:10.1088/1751–8113/46/25/254018.

Patil, D., Hunt, B., Kalnay, E., Yorke, J., and Ott, E. (2001). Local low dimensionality of atmospheric dynamics. *Physical Review Letters*, 86:5878–5881.

Penland, C. (1989). Random forcing and forecasting using principal oscillation pattern analysis. *Monthly Weather Review*, 117:2165–2185.

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

Petrie, R. E. (2008). Localization in the ensemble kalman filter. Master's thesis, University of Reading, Reading, United Kingdom.

Robert, C., Blayo, E., and Verron, J. (2006). Comparison of reduced-order, sequential and variational data assimilation methods in the tropical Pacific Ocean. *Ocean Dynamics*, 56:624–633.

Robinson, A. R. and Lermusiaux, P. (2002). Data assimilation for modeling and predicting coupled physical - biological interactions in the sea. In Robinson, A., McCarthy, J., and Rothschild, B., editors, *The Sea*, volume 12, pages 475–536, New York, United States. John Wiley and Sons, Inc.

Shenk, T. M. and Franklin, A. B. (2001). *Modeling in natural resource management; development, interpretation and application.* Island Press, Washington, United States.

Stewart, G. (1998). *Matrix Algorithms: Vol 1, Basic Decompositions.* Society for Industrial and Applied Mathematics, Philadelphia, United States.

Stott, P. A. and Kettleborough, J. A. (2002). Origins and estimates of uncertainty in predictions of twenty-first century temperature rise. *Nature*, 416:723–726.

Sura, P. and Hannachi, A. (2014). Perspectives of non-gaussianity in atmospheric synoptic and low-frequency variability. *Journal of Climate*, Ahead of Print:doi: 10.1175/JCLI–D–14–00572.1.

Szunyogh, I., Kostelich, E. J., Gyarmati, G., Patil, D. J., Hunt, B. R., Kalnay, E.,

Ott, E., and Yorke, J. A. (2005). Assessing a local ensemble kalman filter: perfect model experiments with the National Centers for Environmental Prediction global model. *Tellus A*, 57:528–545.

Talagrand, O. (1997). Assimilation of observations, an introduction. *Journal of the Meteorological Society of Japan*, 75:191–209.

Talagrand, O. (2012). Four-dimensional variational assimilation (4d-var). In Blayo, E., Bocquet, M., Cosme, E., and Cugliandolo, L. F., editors, *Advanced Data Assimilation for Geoscience*, pages 3–26, Oxford, United Kingdom. Oxford University Press.

Tang, Y. and Ambadan, J. T. (2009). Reply. *Journal of the Atmospheric Sciences*, 66:3501–3503.

Tang, Y., Ambadan, J. T., and Chen, D. (2014a). Nonlinear measurement function in the ensemble kalman filter. *Advances in Atmospheric Sciences*, 31:551–558.

Tang, Y., Deng, Z., Manoj, K. K., and Chen, D. (2014b). A practical scheme of the sigma-point kalman filter for high-dimensional systems. *Journal of Advances in Modeling Earth Systems*, 6:21–37.

Tang, Y., Deng, Z., Zhou, X., Cheng, Y., and Chen, D. (2008). Interdecadal variation of ENSO predictability in multiple models. *Journal of Climate*, 21:4811–4833.

Teixeira, J., Reynolds, C. A., and Judd, K. (2007). Time step sensitivity of nonlinear atmospheric models: Numerical convergence, truncation error growth, and ensemble design. *Journal of the Atmospheric Sciences*, 64:175–189.

Temam, R. (1991). Approximation of attractors, large eddy simulations and multiscale methods. *Proceedings: Mathematical and Physical Sciences*, 434:23–39.

Tippett, M. K., Anderson, J. L., Bishop, C. H., Hamill, T. M., and Whitaker, J. S. (2003). Ensemble square root filters. *Monthly Weather Review*, 131:1485–1490.

Van der Merwe, R. (2004). *Sigma-point Kalman filters for probabilistic inference in dynamic state-space models*. PhD thesis, Oregon Health and Science University, Beaverton, United States.

Van der Merwe, R. and Wan, E. (2001). The square-root unscented kalman filter for state and parameter-estimation. *Acoustics, Speech, and Signal Processing. Proceedings. (ICASSP '01). 2001 IEEE International Conference*, 6:3461–3464.

Van der Merwe, R., Wan, E., and Julier, S. I. (2004). Sigma-point kalman filters

for nonlinear estimation and sensor fusion: Applications to integrated navigation. *AIAA Guidance, Navigation and Control Conf. Providence, RI, American Institute of Aeronautics and Astronautics*, 2004:5120–5122.

van Leeuwen, P. J. (2009). Particle filtering in geophysical systems. *Monthly Weather Review*, 137:4089–4114.

van Leeuwen, P. J. (2012). Particle filters for the geosciences. In Blayo, E., Bocquet, M., Cosme, E., and Cugliandolo, L. F., editors, *Advanced Data Assimilation for Geoscience*, pages 291–318, Oxford, United Kingdom. Oxford University Press.

van Leeuwen, P. J. and Evensen, G. (1996). Data assimilation and inverse methods in terms of a probabilistic formulation. *Monthly Weather Review*, 124:2898–2913.

von Storch, H. and Hannoschock, G. (1984). Comments on empirical orthogonal function- analysis of wind vectors over the tropical Pacific region. *Bulletin of the American Meteorological Society*, 65:162.

Wang, B., Zou, X., and Zhu, J. (2000). Data assimilation and its applications. *Proceedings of the National Academy of Sciences*, 97:11143–11144.

Wang, H.-J., Zhang, R.-H., Cole, J., and Chavez, F. (1999). El Niño and the related phenomenon Southern Oscillation (enso): The largest signal in interannual climate variation. *Proc. of the National Academy of Sciences U S A*, 96:11071–11072. 10500128[pmid].

Webster, P. J. and Palmer, T. N. (1997). The past and the future of El Niño. *Nature*, 390:562–564.

Whitaker, J. S. and Hamill, T. M. (2002). Ensemble data assimilation without perturbed observations. *Monthly Weather Review*, 130:1913–1924.

Wilks, D. (2011). Principal component (EOF) analysis. In Wilks, D. S., editor, *Statistical Methods in the Atmospheric Sciences*, volume 100 of *International Geophysics*, pages 519–562. Academic Press, San Diego, United States.

Yang, S.-C., Corazza, M., Carrassi, A., Kalnay, E., and Miyoshi, T. (2007). Comparison of ensemble-based and variational-based data assimilation schemes in a quasi-geostrophic model. In *AMS 10th Symposium on Integrated Observing and Assimilation Systems for the Atmosphere, Oceans, and Land Surface.*, Atlanta, United States.

Zebiak, S. E. and Cane, M. A. (1987). A model El Niño–Southern Oscillation. *Monthly Weather Review*, 115:2262–2278.

Zhang, F., Zhang, M., and Poterjoy, J. (2013). E3DVar: Coupling an ensemble kalman filter with three-dimensional variational data assimilation in a limited-area weather prediction model and comparison to E4DVar. *Monthly Weather Review*, 141:900–917.

Zhang, M. and Zhang, F. (2012). E4DVar: Coupling an ensemble kalman filter with four dimensional variational data assimilation in a limited area weather prediction model. *Monthly Weather Review*, 140:587–600.