

## ABSTRACT

Title of dissertation: ASSIMILATING SATELLITE OBSERVATIONS  
WITH A LOCAL ENSEMBLE KALMAN FILTER

Elana Judith Fertig  
Doctor of Philosophy, 2007

Dissertation directed by: Professor Brian R. Hunt  
Department of Mathematics

Numerical weather prediction relies on data assimilation to estimate the current state of the atmosphere. Generally speaking, data assimilation methods combine information from observations and from a prior forecast state, taking into account their respective uncertainties. Ensemble-based data assimilation schemes estimate the forecast uncertainty with the sample covariance from an ensemble of forecasts. While these schemes have been shown to successfully assimilate conventional observations of model state variables, they have only recently begun to assimilate satellite observations. This dissertation explores some of the complications that arise when ensemble-based schemes assimilate satellite observations.

Although ensemble data assimilation schemes often assume that observations are taken at the time of assimilation, satellite observations are available almost continuously between consecutive assimilation times. In Chapter 2, we formulate a “four-dimensional” extension to ensemble-based schemes that is analogous to the operationally used scheme 4D-VAR. Using perfect model experiments with the Lorenz-96 model, we find that the four-dimensional ensemble scheme can per-

form comparably to 4D-VAR.

Many ensemble data assimilation schemes utilize spatial localization so that a small ensemble can capture the unstable degrees of freedom in the model state. These local ensemble-based schemes typically allow the analysis at a given location to depend only on observations near that location. Meanwhile, the location of satellite observations cannot be pinpointed in the same manner as conventional observations. In Chapter 3, we propose a technique to update the state at a given location by assimilating satellite radiance observations that are strongly correlated to the model state there. For satellite retrievals, we propose incorporating the observation error covariance matrix and selecting the retrievals that have errors correlated to observations near the location to be updated. Our selection techniques improve the analysis obtained when assimilating simulated satellite observations with a seven-layer primitive equation model, the SPEEDY model.

Finally, satellite radiance observations are subject to state-dependent, systematic errors due to errors in the radiative transfer model used as the observation operator. In Chapter 4 we propose applying state-space augmentation to ensemble based assimilation schemes to estimate satellite radiance biases during the data assimilation procedure. Our approach successfully corrects such systematic errors in simulated biased satellite observations with the SPEEDY model.

ASSIMILATING SATELLITE OBSERVATIONS  
WITH A LOCAL ENSEMBLE KALMAN FILTER

by

Elana Judith Fertig

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

Advisory Committee:  
Professor Brian R. Hunt, Chair/Advisor  
Dr. Ricardo Todling, NASA Advisor  
Professor Edward Ott  
Professor Istvan Szunyogh  
Professor Harry Tamvakis

© Copyright by  
Elana Judith Fertig  
2007

## Acknowledgments

I owe thanks to many people for making this dissertation possible. I would like to first thank my entire dissertation committee. Thanks to Brian Hunt, Edward Ott, and Istvan Szunyogh who helped me develop and pursue relevant research projects. I have grown immensely from their advice and consideration. I also owe thanks to Ricardo Todling for the countless hours he dedicated to bring a scaled down version of the NASA fvGCM operational model to the University of Maryland.

Thanks to the entire Chaos Weather group for giving me the opportunity to work on real world systems from early in my graduate school career and countless public speaking opportunities. Eugenia Kalnay has been optimistic and supportive in all the projects I have pursued. Thanks to Eric Kostelich for his continual coding efforts in developing and improving the LETKF system. I owe both Ricardo Todling and Eric Kostelich further thanks for their patience in teaching myself, Hong Li, and Junjie Liu their systems. I am also grateful for the time I spent working with José Aravèquia, Hong Li, and Junjie Liu. Thanks to their patience, advice, teaching, and friendship during our countless hours coding together, I learned and achieved things I never would have been able to do on my own.

A special thank you to Gyorgyi Gyarmati, Alfredo Nava-Tudela, and Aleksey Zimin for fighting to make sure that the computers always ran smoothly. Thanks to Alverda McCoy for assisting me in jump through all the graduate school hoops. I would also like to thank David Levermore, the ESSIC staff, and the NASA Goddard staff for generously arranging and providing the funding for my time at the University of Maryland.

I would like to thank Arlindo DaSilva, Jim Jeung, Jim Purser, and Jeff Whitaker for their helpful suggestions for my research. Thanks, too, to all my fellow graduate students, who were always there as a support. I would like to especially acknowledge Seung-Jong Baek, Chris Danforth, Amy Finkbiner, Fernando Galaz-Garcia, JT Halbert, Russ Halper, John Harlim, Matt Hoffman, David Kuhl, Aaron Lott, Dagmar Merkova, Elizabeth Satterfield, Suzanne Sindi, Poorani Submaranian, and James White for their suggestions, advice, and conversations - helpful both in science and beyond. I owe a special thanks to Janet Gusdorff for always listening to and believing in me.

Last, but certainly not least, I would like to thank my entire family. Thanks for support from Ari Davidow, Ari Fertig, Arnie Fertig, Gail Fertig, and Judith Pinnolis. I also thank my grandparents, Leo Goldstein, Sybil Goldstein, Arlene Kamin, Robert Kamin, Gina Klein, Myron Klein, Eileen Pinnolis, Deborah Shain, Murray Silverman, Bella Weiner, and Samuel Weiner, for their continual kvelling about me. Thanks to my parents, Merom and Louise Klein, for helping me keep my priorities in order and goals in mind. Thanks, too, to my parents, Harriet and Louis Weiner, for their support and scientific advice. Thanks to my brothers, Ken and David Weiner, for helping me to keep things in perspective, making me laugh, and always being interested in my research. Finally, I would like to thank my husband, Ben Fertig, for his love, humor, and kindness. His continual support of me and conviction in my abilities from qualifying exams to dissertation has been integral to my completing my PhD.

# Table of Contents

List of Tables	vi
List of Figures	vii
1 Introduction	1
1.1 Motivation . . . . .	1
1.2 Data Assimilation . . . . .	3
1.3 Assimilating Satellite Observations . . . . .	8
1.4 Outline of Thesis . . . . .	10
1.5 Figures . . . . .	13
2 A Comparative Study of 4D-VAR and a 4D Ensemble Kalman Filter	15
2.1 Introduction . . . . .	15
2.2 Formulation . . . . .	17
2.2.1 4D-VAR . . . . .	17
2.2.2 4D-LETKF . . . . .	18
2.3 Results . . . . .	21
2.4 Summary . . . . .	25
2.5 Figures . . . . .	26
3 Assimilating Nonlocal Observations with a Local Ensemble Kalman Filter	28
3.1 Introduction . . . . .	28
3.2 Background . . . . .	31
3.2.1 Ensemble Kalman Filter . . . . .	31
3.2.2 Localization . . . . .	35
3.2.3 Assimilating Nonlocal Observations . . . . .	37
3.2.4 SPEEDY Model . . . . .	39
3.3 Experimental Design . . . . .	40
3.4 Results . . . . .	44
3.4.1 Grid Point Observations . . . . .	44
3.4.2 Radiance Observations . . . . .	44
3.4.3 Retrieval Observations . . . . .	46
3.5 Conclusions . . . . .	48
3.6 Appendix: Covariance Localization . . . . .	53
3.7 Appendix: Diagonalization . . . . .	55
3.8 Figures . . . . .	58
3.9 Tables . . . . .	62
4 Satellite Observation Bias Correction with an Ensemble Kalman Filter	63
4.1 Introduction . . . . .	63
4.2 Satellite Radiance Biases . . . . .	65
4.2.1 Online Bias Correction . . . . .	69
4.2.2 Online Bias Correction with an Ensemble Kalman Filter . . . . .	70

4.3	Experimental Design . . . . .	74
4.3.1	SPEEDY Model . . . . .	74
4.3.2	Community Radiative Transfer Model . . . . .	75
4.3.3	Truth and Observations . . . . .	77
4.3.4	LETKF and Bias Correction . . . . .	79
4.4	Results . . . . .	81
4.5	Conclusions . . . . .	87
4.6	Figures . . . . .	89
4.7	Tables . . . . .	99

## List of Tables

3.1	The observation error covariance for the simulated retrieval observations at a fixed horizontal location. . . . .	62
3.2	The correlation between the errors of observations at a single horizontal point and model levels $k$ and $l$ is given by the entry in this table corresponding to these model levels. . . . .	62
4.1	The global, February average of the temperature analysis RMS error (K) at each model level for selected combinations of predictors. Here, 6% variance inflation is applied to the analysis variables and 6% variance inflation is applied to the bias parameters. . . . .	99
4.2	As for Table 4.2. Here, 6% variance inflation is applied to the analysis variables and 12% variance inflation is applied to the bias parameters. . . . .	99
4.3	The global, February average of the observation RMS error (K) for representative channels when applying the bias correction parameters estimated using 6% variance inflation for the analysis variables and 12% for the bias parameters. . . . .	100

## List of Figures

1.1	A schematic analysis cycle used in data assimilation schemes, based on Figure 1.4.2 in Kalnay (2002). . . . .	13
1.2	The horizontal observational distribution within $\pm 3$ hours of a typical analysis step, from Figure 1.4.1 in Kalnay (2002). . . . .	14
2.1	The mean analysis error as a function of the analysis window for 4D-VAR and as a function of time between analyses for 4D-LETKF. . . . .	27
2.2	The mean forecast errors as a function of forecast time for 4D-VAR and 4D-LETKF. . . . .	27
3.1	A schematic for the observation selection schemes used for nonlocal, radiance-like observations. . . . .	58
3.2	Locations of rawinsonde observations used for the SPEEDY model. . . . .	59
3.3	Average analysis RMS error from assimilating rawinsondes with different amounts of vertical localization. . . . .	59
3.4	Average analysis RMS error from assimilating simulated rawinsonde observations and satellite radiances. . . . .	60
3.5	Average analysis RMS error from assimilating simulated rawinsonde observations and satellite retrievals. . . . .	60
3.6	Comparing the analysis RMS error from assimilating simulated rawinsondes and radiances with that from assimilating simulated rawinsondes and satellite retrievals. . . . .	61
4.1	Weighting function of representative channels. . . . .	89
4.2	The true bias for two representative channels. . . . .	90
4.3	The correlation between the true bias and each of the predictors for each satellite channel. . . . .	91
4.4	The 500 hPa temperature analysis rms error for each analysis time. . . . .	92
4.5	The spread in the background bias parameters with predictors $p_1$ , $p_3$ , and $p_4$ . . . . .	93

4.6	As for Figure 4.4. Here, the analysis variables are inflated by 6% while the bias parameters are inflated by 12%. . . . .	94
4.7	As for Figure 4.5. Here, the analysis variables are inflated by 6% while the bias parameters are inflated by 12%. . . . .	95
4.8	Temperature analysis rms error at each model level. . . . .	96
4.9	Observation rms error for each satellite channel. . . . .	97
4.10	The results from using $p_1$ to correct only channels above $701.618 \text{ cm}^{-1}$ . . . . .	98

# Chapter 1

## Introduction

### 1.1 Motivation

Weather models forecast the future state of the atmosphere from an estimate of the current state of the atmosphere. Even if the model perfectly describes the atmospheric dynamics, for sufficiently long forecast lead times the chaotic nature of the model can turn small errors in the initial conditions into dramatic errors in forecasts. To minimize these errors, data assimilation schemes seek an accurate estimate for the current state of the atmosphere using both current and past observations. Current observations are incorporated into the analysis when the data assimilation is performed. Past observations are incorporated by using a forecast of the initial condition estimated by the data assimilation scheme as the prior forecast the next time that data assimilation is performed (Figure 1.1, based on Figure 1.4.2 in Kalnay, 2002).

Historically, in order to reduce computational cost, several data assimilation schemes, including 3D-VAR and PSAS, assumed that the forecast uncertainty is constant. While such schemes have been successful for operational data assimilation, their estimates are subject to “errors of the day” (Kalnay, 2002; Liu *et al.*, 2007). The next generation of data assimilation schemes avoid these errors by using flow dependent uncertainties, implicitly in 4D-VAR and explicitly in ensemble Kalman

filter schemes. Several studies have shown the potential of the developing ensemble Kalman filter schemes for operational use (e.g., Houtekamer *et al.*, 2005; Szunyogh *et al.*, 2005; Whitaker *et al.*, 2004; Houtekamer and Mitchell, 2005; Whitaker *et al.*, 2007; Liu *et al.*, 2007; Szunyogh *et al.*, 2007).

Many studies with ensemble Kalman filters have focused on assimilating conventional observations, which are direct measurements of model state variables available at irregular locations (Figure 1.2 based on Figure 1.4.1 in Kalnay, 2002). Although satellite observations do not measure model variables directly, they cover more of the globe than do conventional observations (Figure 1.2, middle right panel). Satellite observations are known to improve the estimates for the current state of the atmosphere in regions like the Southern Hemisphere and oceans, which lack conventional data (e.g., Kalnay, 2002). Therefore, studies with ensemble Kalman filters are starting to incorporate satellite observations (e.g., Houtekamer *et al.*, 2005; Houtekamer and Mitchell, 2005). Ensemble schemes must account for complications that arise because satellite observations are available continuously between analysis times, measure weighted vertical averages of the atmospheric state at a single horizontal location, and have significant errors relative to the observation operator in use. This dissertation examines these aspects of satellite data assimilation with an ensemble Kalman filter.

The data assimilation methods described in this section are discussed in more detail in Section 1.2. Then, Section 1.3 describes satellite data assimilation and Section 1.4 outlines the complications associated with assimilating satellite observations that are addressed in the following chapters of this dissertation.

## 1.2 Data Assimilation

Data assimilation estimates the current atmospheric state by combining information from a prior forecast of this state, called the background ( $\mathbf{x}^b$ ), with atmospheric observations ( $\mathbf{y}$ ). How these two states are weighted in doing this combination depends on their respective uncertainties as given by their error covariance matrices ( $\mathbf{P}^b$  for  $\mathbf{x}^b$  and  $\mathbf{R}$  for  $\mathbf{y}$ ). This estimate, or analysis state ( $\mathbf{x}^a$ ) and possibly its error covariance matrix ( $\mathbf{P}^a$ ) are then used as an initial condition for future forecasts and the analysis cycle continues (Figure 1.1).

The observations used in data assimilation schemes are assumed to be related to the atmospheric state,  $\mathbf{y} = \mathbf{h}(\mathbf{x}^t) + \boldsymbol{\epsilon}$ , where  $\mathbf{x}^t$  represents the unknown true state of the atmosphere. The error term,  $\boldsymbol{\epsilon}$ , has covariance  $\mathbf{R} = \langle \boldsymbol{\epsilon}\boldsymbol{\epsilon}^\top \rangle$ , where  $\top$  denotes the transpose. The observation operator  $\mathbf{h}$  transforms atmospheric model states into observation space. The  $\mathbf{h}$  operator primarily performs spatial interpolations for most conventional observations, but may be a more complex, nonlinear operator for observations like satellite radiances.

Assuming that the errors of the background and observations are Gaussian distributed and unbiased, a least-squares approach can combine these measures to obtain the analysis state. This state can be found by maximizing a likelihood function, or equivalently, minimizing the following cost function

$$2J(\mathbf{x}) = (\mathbf{x} - \mathbf{x}^b)^\top \mathbf{P}^{b-1} (\mathbf{x} - \mathbf{x}^b) + (\mathbf{y} - \mathbf{h}(\mathbf{x}))^\top \mathbf{R}^{-1} (\mathbf{y} - \mathbf{h}(\mathbf{x})) \quad (1.1)$$

(e.g., Kalnay, 2002). Some operational data assimilation schemes, including 3D-VAR, find the analysis state by directly minimizing this cost function using a con-

stant background error covariance  $\mathbf{P}^b$ . This latter assumption is made for computational efficiency, but may degrade the quality of the analysis and introduce time-varying “errors of the day” (Kalnay, 2002; Liu *et al.*, 2007). 4D-VAR schemes (the “strong constraint” formulation is described in more detail in Chapter 2) overcome these limitations by modifying the cost function in Equation (1.1) to seek the model trajectory which best fits the all available observations within some time window. In this way, 4D-VAR can implicitly account for time dependent errors (e.g., Thépaut *et al.*, 1993).

While the data assimilation techniques described above accurately estimate the initial state of the atmosphere, they do not directly incorporate dynamic estimates of the forecast uncertainty. By using approximations to the Kalman filter equations, ensemble-based schemes can efficiently estimate and utilize the forecast uncertainty when obtaining the analysis. We formulate the Kalman filter equations and related ensemble-based schemes in the remainder of this section.

In addition to the assumptions made about the observations, for the purpose of presenting the classical Kalman filter technique, we also assume that the analysis state and its error covariance are propagated by a linear forecast model  $\mathbf{M}$ . Then, if at analysis time  $t_{n-1}$  the analysis state is  $\mathbf{x}_{n-1}^a$  and its uncertainty is given by the covariance matrix  $\mathbf{P}_{n-1}^a$ , the background state and covariance at time  $t_n$  are given by

$$\mathbf{x}_n^b = \mathbf{M}\mathbf{x}_{n-1}^a, \quad (1.2)$$

$$\mathbf{P}_n^b = \mathbf{M}\mathbf{P}_{n-1}^a\mathbf{M}^\top. \quad (1.3)$$

In this way, an estimated analysis state with Gaussian errors is forecasted into a background state with Gaussian errors, and the Kalman filter (Kalman, 1960; Kalman and Bucy, 1961) can be applied to estimate the analysis state and its covariance at time  $t_n$ .<sup>1</sup> Specifically, the analysis equations are given by

$$\mathbf{x}^a = \mathbf{x}^b + \mathbf{K} (\mathbf{y} - \mathbf{h}(\mathbf{x}^b)), \quad (1.4)$$

$$\mathbf{P}^a = (\mathbf{I} - \mathbf{K}\mathbf{H}) \mathbf{P}^b, \quad (1.5)$$

$$\mathbf{K} = \mathbf{P}^b \mathbf{H}^\top (\mathbf{H} \mathbf{P}^b \mathbf{H}^\top + \mathbf{R})^{-1}, \quad (1.6)$$

where  $\mathbf{H}$  is the linearization of the observation operator  $\mathbf{h}$  around the background  $\mathbf{H} = \frac{\partial \mathbf{h}}{\partial \mathbf{x}}(\mathbf{x}^b)$ . If  $\mathbf{h}$  is linear, then the analysis state given by Equations (1.4) and (1.6) is the minimum of the cost function in Equation (1.1). Furthermore, in this case the Kalman filter provides the minimum variance unbiased estimate for the analysis state (Kalman, 1960; Kalman and Bucy, 1961), and, if the dynamics were linear, then the Kalman filter is equivalent to 4D-VAR in the limit of long time windows.

In the formulation of the Kalman filter in the previous paragraph, we assumed that the forecast model is linear. For a nonlinear forecast model, the extended Kalman filter provides an approach to replace the linear model in Equation (1.2) with the nonlinear forecast model, and linearizes this model about the forecast trajectory to obtain the matrix  $\mathbf{M}$  in Equation (1.3).

If the model state has dimension  $m$ , then  $\mathbf{P}^a$ ,  $\mathbf{P}^b$ , and  $\mathbf{M}$  are  $m \times m$  matrices. For operational weather models,  $m$  is more than  $10^7 - 10^8$ . Therefore, propagating

---

<sup>1</sup>In the remainder of this section we consider all variables at time  $t_n$  only, so we drop the time index.

the error covariance as proposed in Equation (1.3) is prohibitively expensive to implement. Ensemble Kalman filter techniques reduce this cost by approximating Equation (1.3) with the sample covariance of an ensemble of forecasts made from a set of initial conditions whose perturbations from the analysis state are chosen according to the analysis covariance in Equation (1.5). If there are enough ensemble members and forecast model is roughly linear over the forecast time, this approach will ideally come close to the optimality of the Kalman filter analysis (Evensen, 1994).

Specifically, ensemble Kalman filters forecast a set of perturbed initial conditions, creating an ensemble of  $k$  forecasts, each denoted by  $\mathbf{x}^{b(i)}$ . The background state  $\mathbf{x}^b$  is the ensemble mean of these states

$$\mathbf{x}^b = k^{-1} \sum_{i=1}^k \mathbf{x}^{b(i)}. \quad (1.7)$$

The sample covariance of the ensemble of forecasts yields a time dependent background error covariance,

$$\mathbf{P}^b = \frac{\sum_{i=1}^k (\mathbf{x}^{b(i)} - \mathbf{x}^b) (\mathbf{x}^{b(i)} - \mathbf{x}^b)^\top}{k - 1}. \quad (1.8)$$

The analysis state and its error covariance are obtained by substituting Equations (1.7) and (1.8) into Equations (1.4) - (1.6).

The ensemble Kalman filter also determines an ensemble of analysis states based on the analysis state and its error covariance. The forecast model is then applied to evolve each analysis ensemble member to obtain the background ensemble at the next analysis step. The means of generating this ensemble differentiates

ensemble-based schemes. In many ensemble schemes, the analysis for the  $i^{\text{th}}$  ensemble member is given by

$$\mathbf{x}^{a(i)} = \mathbf{x}^a + \mathbf{X}^{a(i)}. \quad (1.9)$$

Here  $\mathbf{X}^{a(i)}$  represents the  $i^{\text{th}}$  column of an  $m \times k$  matrix “square root” matrix  $\mathbf{X}^a$ , for which  $\mathbf{X}^a (\mathbf{X}^a)^\top = (k - 1) \mathbf{P}^a$ . The choice of the particular square root further distinguishes several ensemble Kalman filter schemes (Tippett *et al.*, 2003).

Unlike the standard modifications of the Kalman filter, such as the extended Kalman filter, an ensemble Kalman filter does not require explicitly linearizing the forecast model; the ensemble forecast implicitly gives information about the linearization. Furthermore,  $\mathbf{P}^b$  has only  $k - 1$  linearly independent perturbations, where we assume  $k$  to be small compared to  $m$ . Presuming that this limited number of ensemble members captures the model dynamics, the number of degrees of freedom in the Kalman filter equations is reduced. Several ensemble schemes, including the Ensemble Transform Kalman Filter (ETKF, Bishop *et al.*, 2001), the Local Ensemble Kalman Filter (LEKF, Ott *et al.*, 2004), and the Local Ensemble Transform Kalman Filter (LETKF, Hunt *et al.*, 2007), exploit this low-dimensionality to solve Equations (1.4) - (1.6) in the space spanned by the ensemble, reducing the cost associated with solving these equations.

In most operational settings, ensemble Kalman filter schemes are only computationally feasible for fewer than 100 ensemble members. These few ensemble members are unable to capture all of the unstable degrees of freedom in the global model state. Furthermore, globally using so few ensemble members may introduce

spurious correlations in the background error covariance estimated in Equation (1.8). However, as shown by Patil *et al.* (2001), atmospheric dynamics tend to be low dimensional within local regions. Therefore, many ensemble Kalman filter schemes employ some form of localization. Ideally, the limited ensemble size will be sufficient to represent the state uncertainty within these local regions (Ott *et al.*, 2004; Szunyogh *et al.*, 2005).

### 1.3 Assimilating Satellite Observations

Satellites observe the radiance emitted from the atmosphere at particular window frequencies continually during their orbit. Accordingly, these instruments can provide data assimilation systems a wealth of information about traditionally data poor regions, such as the Southern Hemisphere and oceans (Figure 1.2). However, this information is less straightforward to assimilate than conventional observations, which directly measure model state variables.

Because satellites observe continuously, they provide information at greater frequency than the analysis is performed. However, in many of the data assimilation schemes presented in the previous section (but not all, e.g., 4D-VAR), the observations are assumed to be taken at the analysis time. Therefore, the data assimilation schemes must account for the model dynamics between analysis times when assimilating satellite observations. 4D-VAR schemes naturally handle this because they consider model trajectories over an analysis time window. Evensen and van Leeuwen (2000), Anderson (2001), and Hunt *et al.* (2004) have made similar

modifications to the ensemble Kalman filter schemes to assimilate asynchronous observations. These “four-dimensional” ensemble-based schemes are beginning to be applied to operational models and data (Houtekamer and Mitchell, 2005; Szunyogh *et al.*, 2007; Whitaker *et al.*, 2007).

Further challenges arise because, unlike conventional observations, satellite radiance observations only indirectly represent the atmospheric variables used in the forecast models, such as temperature and humidity. For example, the case of radiance observations, the relationship between the observations and the relevant model state variables (i.e., temperature and humidity) is given by a radiative transfer model. Accordingly, the observation operator for each radiance observation is given by

$$h(\mathbf{x}) = \tau_\nu(0)\kappa_\nu(0)B(T(0)) + \int_0^\infty B_\nu(T(z))\frac{\partial\tau_\nu(z)}{\partial z}dz, \quad (1.10)$$

where  $\nu$  represents the observing frequency,  $z = 0$  represents the surface,  $\kappa_\nu(0)$  is the emissivity of the surface,  $B_\nu(T(z))$  is the Planck function, and  $\frac{\partial\tau}{\partial z}$  is called the *weighting function* because it indicates amount which a radiance observation depends on the Planck function at a particular height (e.g., Liou, 2002). Because Equation (1.10) involves an integral over all model levels, the radiance observations depend on the atmospheric state in a fundamentally nonlocal manner. This nonlocal nature presents a problem for previous ensemble Kalman filter schemes that rely on localization to assimilate radiances (see Chapter 2). Furthermore, there are often systematic errors in the radiative transfer model, which often arise from errors in calibrating the spectroscopic data associated with the satellite (Eyre, 1992). These

errors can introduce a state dependent bias to the satellite observations, contradicting the assumption of unbiased observation errors made in all the data assimilation systems and harming the resulting analysis.

## 1.4 Outline of Thesis

In this dissertation, we explore some of the complications arising from assimilating satellite observations that are described in the previous section. Each of the following chapters has been prepared for publication in *Tellus A*. Therefore, each chapter has been written to stand alone and may have slight differences in notation.

In Chapter 2, we formulate a four dimensional Ensemble Kalman Filter (the 4D-Local Ensemble Transform Kalman Filter, 4D-LETKF) based on Hunt *et al.* (2004) to assimilate observations available between analysis times, such as satellite observations. This scheme minimizes a cost function similar to that in a 4D-VAR method described in Section 1.2 in the space spanned by the ensemble trajectories. Using perfect model experiments with the Lorenz-96 model, we compare assimilation of simulated asynchronous observations with 4D-VAR and 4D-LETKF. We find that both schemes have comparable error when 4D-LETKF is performed sufficiently frequently and when 4D-VAR is performed over a sufficiently long analysis time window. We explore how the error depends on the time between analyses for 4D-LETKF and the analysis time window for 4D-VAR. This chapter has appeared in *Tellus A* as Fertig *et al.* (2007b).

In Chapter 3, we explore the nonlocality associated with satellite observa-

tions and the challenges it presents for local ensemble Kalman filter schemes. We propose a technique to assimilate these nonlocal observations in which the observation operator is applied to the global model state and then appropriate observations are selected to estimate the atmospheric state for each model grid point. The issue of how to choose appropriate observations is investigated with numerical experiments using the ensemble scheme LETKF on a seven layer primitive equation model, the the Simplified Parameterizations, primitivE-Equation DYnamics (SPEEDY) model. We run numerical experiments that assimilate simulated point and nonlocal radiance-like observations and numerical experiments that assimilate simulated point and retrieval-like observations. For radiances, the best analysis results are obtained from a scheme that updates the state at a given location by assimilating all those radiance observations that depend strongly on the model state near that location. For retrievals, we assume a known non-diagonal error covariance matrix. We use a scheme that updates the state at a given location by assimilating retrieval observations near that location (“local retrievals”) and also retrieval observations with errors correlated to those of the local retrievals. The analysis that results from assimilating these retrievals has comparable accuracy to the analysis obtained from assimilating radiances.

Finally, in Chapter 4 we consider the use of an ensemble Kalman filter to correct satellite radiances for their state dependent biases relative to the observation operator in use. Our approach is to use state-space augmentation to estimate satellite biases during the ensemble data assimilation procedure. We illustrate our approach by applying it to LETKF to assimilate simulated biased AIRS brightness

temperature observations on the SPEEDY model. The bias parameters estimated by LETKF successfully reduce the analysis error and observation bias.

1.5 Figures

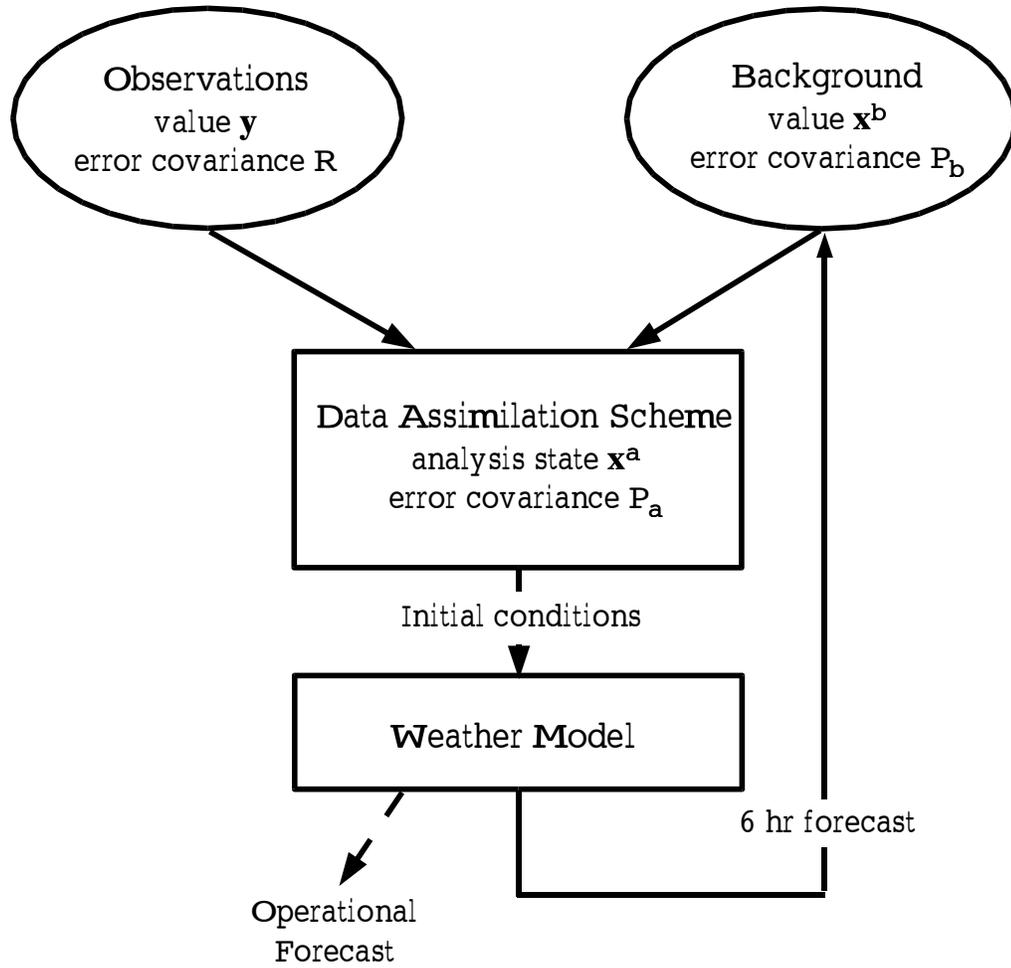


Figure 1.1: A schematic analysis cycle used in data assimilation schemes, based on Figure 1.4.2 in Kalnay (2002).

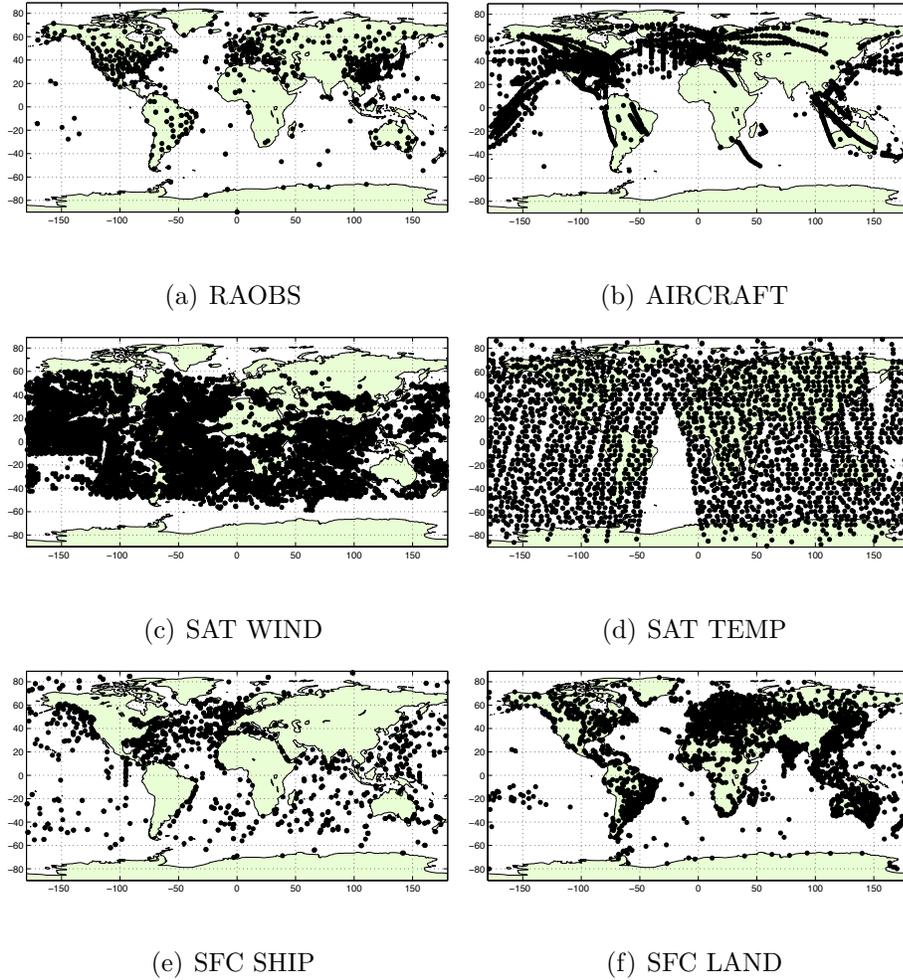


Figure 1.2: The horizontal observational distribution within  $\pm 3$  hours of a typical analysis step, based on Figure 1.4.1 in Kalnay (2002). Here, rawinsonde (RAOBS), aircraft, satellite derived wind (SAT WIND), surface ship (SFC SHIP), and surface land (SFC LAND) are considered to be conventional observations. As their name suggests the surface observations are only available to measure surface variables. SAT TEMP represents the typical horizontal distribution of satellite radiance observations.

## Chapter 2

### A Comparative Study of 4D-VAR and a 4D Ensemble Kalman

#### Filter: Perfect Model Simulations with Lorenz-96

##### 2.1 Introduction

Operational data assimilation schemes traditionally have assimilated available observations as though they, or their innovations (observation minus forecast, in observation space), occurred at the analysis time. With a growing number of observations from instruments such as satellites, many observations are available between analysis times. However, employing three dimensional data assimilation schemes to match the frequency of these asynchronous observations would be prohibitively expensive and could introduce imbalances in the resulting analysis states. New generations of data assimilation schemes, most notably the 4D-VAR and Ensemble Kalman Filter (EnKF) techniques, are able to more accurately take into account the timing of asynchronous observations.

One approach to assimilate observations at various time is 4D-VAR (see Le Dimet and Talagrand, 1986; Courtier *et al.*, 1994; Rabier *et al.*, 1998, 2000). This assimilation system is currently the operational data assimilation scheme at the European Centre for Medium-Range Weather Forecasts and Meteo-France, and is being developed for operation at centers including the Canadian Meteorological

Centre and the Japan Meteorological Agency. “Strong constraint” 4D-VAR seeks a model trajectory that best fits the available observations during a specified time window before the analysis time.

Another approach to data assimilation is the EnKF technique (Evensen, 1994; Burgers *et al.*, 1998; Houtekamer and Mitchell, 1998; Anderson, 2001; Bishop *et al.*, 2001; Whitaker and Hamill, 2002; Ott *et al.*, 2004; Zupanski, 2005). These schemes evolve an ensemble of model trajectories to estimate background uncertainty. Though not yet operational, experiments, such as those of Houtekamer *et al.* (2005), Szunyogh *et al.* (2005), and Whitaker *et al.* (2004), have shown the potential of EnKF for operational data assimilation. In these implementations of EnKF on operational models, observations are still assimilated as though they were taken at the analysis time. As Lorenc (2003) points out, when assimilating asynchronous observations with EnKF, one should use the time sequence of ensemble states between analysis times to account for model state correlations in time as well as space. In this way, Evensen and van Leeuwen (2000), Anderson (2001), and Hunt *et al.* (2004) extend EnKF to accurately assimilate asynchronous observations at the correct time. Their methods have been successfully applied to operational models by Houtekamer and Mitchell (2005), Szunyogh and Kostelich (Personal Communication, 2006), and Whitaker *et al.* (2007).

In this paper, we present 4D-LETKF, a simplified version of the four-dimensional Ensemble Kalman Filter in Hunt *et al.* (2004), and compare it to 4D-VAR in a perfect model scenario using the Lorenz-96 system (Lorenz, 1996). In Section 2.2, we describe our implementation of 4D-VAR, and then derive 4D-LETKF using a mod-

ification of the 4D-VAR cost function. In Section 2.3, we describe our experimental design and present our numerical results. We find that the two methods yield similar results when the time between analysis is short enough for 4D-LETKF and when the analysis time window is long enough for 4D-VAR, and we discuss our results further in Section 2.4.

## 2.2 Formulation

### 2.2.1 4D-VAR

Since we consider a perfect model scenario in this paper, we formulate here a strong constraint 4D-VAR scheme. It seeks an initial condition close to the background state (determined by the prior analysis) from which the resulting exact model trajectory remains closest to the observations over the analysis time window,  $[t_0, t_n]$ . More precisely, it minimizes a cost function:

$$\begin{aligned}
 J(\mathbf{x}(t_0)) &= \frac{1}{2} (\mathbf{x}(t_0) - \mathbf{x}^b)^\top \mathbf{B}^{-1} (\mathbf{x}(t_0) - \mathbf{x}^b) \\
 &\quad + \frac{1}{2} \sum_{l=0}^n (\mathbf{y}_l^o - H_l(\mathbf{x}(t_l)))^\top \mathbf{R}_l^{-1} (\mathbf{y}_l^o - H_l(\mathbf{x}(t_l)))
 \end{aligned} \tag{2.1}$$

where  $\top$  denotes the transpose,  $\mathbf{x}(t_0)$  is an  $m$ -dimensional the model state at the start of the analysis window. The model state at each observation time is obtained by integrating the nonlinear model from  $\mathbf{x}(t_0)$ .  $\mathbf{x}^b$  is the the  $m$ -dimensional background forecast at the same time and  $\mathbf{B}$  is an  $m \times m$  background error covariance matrix, which is typically constant, homogeneous, and isotropic. The observation state at time  $t_l$  is given by the  $s_l$ -dimensional vector  $\mathbf{y}_l^o$ .  $\mathbf{R}_l$  is the associated  $s_l \times s_l$

observation error covariance matrix, for  $l = 0, \dots, n$ . The observation operator,  $H_l$ , maps the model state  $\mathbf{x}(t_l)$  to the observation space at time  $t_l$ . For this formulation, observations taken at different times are assumed to have uncorrelated errors.

The cost  $J$  is only a function of the initial model state,  $\mathbf{x}(t_0)$ . Therefore, once we determine  $\mathbf{x}(t_0)$  that minimizes the total cost  $J$ , the integrated state  $\mathbf{x}(t_n)$  is the 4D-VAR analysis at time  $t_n$ . For this study, we obtain the minimum using a BFGS algorithm adapted from Numerical Recipes in Fortran (Press *et al.*, 1992, pg. 418). This algorithm requires the gradient of the cost function, which we compute using the adjoint technique presented in Lawson *et al.* (1995). For large systems, Hessian preconditioning is often employed to ensure that the minimization algorithm converges quickly on an accurate state. In this study, we judged preconditioning to be unnecessary because here the BFGS algorithm converges to the minimum state quickly (generally after about 20 iterations).

### 2.2.2 4D-LETKF

EnKF replaces the time-independent background error covariance matrix  $\mathbf{B}$  with a time-dependent sample covariance matrix  $\mathbf{P}^b = (k-1)^{-1}\mathbf{X}^b(\mathbf{X}^b)^\top$ , where  $\mathbf{X}^b$  is an  $m \times k$  matrix of background ensemble perturbations. That is,

$$\mathbf{X}^b = [\mathbf{x}^{b(1)} - \bar{\mathbf{x}}^b \mid \mathbf{x}^{b(2)} - \bar{\mathbf{x}}^b \mid \dots \mid \mathbf{x}^{b(k)} - \bar{\mathbf{x}}^b], \quad (2.2)$$

where  $\mathbf{x}^{b(i)}$  denotes the  $i$ th ensemble member and  $\bar{\mathbf{x}}^b$  denotes the ensemble mean. In 4D-LETKF, the deviation of a model solution  $\mathbf{x}(t)$  from the background mean state  $\bar{\mathbf{x}}^b(t)$  is approximated between analysis times by a linear combination of the

background ensemble perturbations by:

$$\mathbf{x}(t) \approx \bar{\mathbf{x}}^b(t) + \mathbf{X}^b(t)\mathbf{w} \quad (2.3)$$

where  $\mathbf{w} \in \mathbb{R}^k$  is a time-independent weight vector. As in Hunt *et al.* (2004), the analysis determines which weight vector makes this linear combination “best fit” the observations over the analysis time window, in the sense of minimizing a cost function like (2.1).

The projection of  $\mathbf{x}(t)$  to the observation space at time  $t_l$  is approximated by:

$$H_l(\mathbf{x}(t_l)) \approx H_l(\bar{\mathbf{x}}^b(t_l) + \mathbf{X}^b(t_l)\mathbf{w}) \approx H_l(\bar{\mathbf{x}}^b(t_l)) + \mathbf{Y}_l^b\mathbf{w}. \quad (2.4)$$

The  $i$ th column vector of the  $s_l \times k$  matrix  $\mathbf{Y}_l^b$  is defined to be  $H_l(\mathbf{x}^{b(i)}(t_l)) - \overline{H_l(\mathbf{x}^{b(i)}(t_l))}$ , where  $\overline{H_l(\mathbf{x}^{b(i)}(t_l))}$  represents the ensemble mean of the projection of the background state on observation space.

Replacing  $\mathbf{B}$  by  $\mathbf{P}^b(t_0)$  and  $\mathbf{x}^b$  by  $\bar{\mathbf{x}}^b(t_0)$  and substituting approximations (2.3) and (2.4) into the cost function (2.1) yields the modified cost function:

$$\begin{aligned} \tilde{J}(\mathbf{w}) = & \frac{1}{2}(k-1)\mathbf{w}^\top\mathbf{w} + \frac{1}{2}\sum_{l=1}^n[\mathbf{y}_l^o - H_l(\bar{\mathbf{x}}^b(t_l)) - \mathbf{Y}_l^b\mathbf{w}]^\top \\ & \times \mathbf{R}_l^{-1}[\mathbf{y}_l^o - H_l(\bar{\mathbf{x}}^b(t_l)) - \mathbf{Y}_l^b\mathbf{w}]. \end{aligned} \quad (2.5)$$

Here, the  $m$ -dimensional minimization problem is reduced to a  $k$ -dimensional problem, reducing the cost of implementation if the ensemble size  $k$  is less than the number of model variables  $m$ .

The minimum of (2.5) occurs at

$$\mathbf{w}^a = \tilde{\mathbf{P}}^a \left( \sum_{l=1}^n \mathbf{Y}_l^{b\top} \mathbf{R}_l^{-1} (\mathbf{y}_l^o - H_l(\bar{\mathbf{x}}^b(t_l))) \right), \quad (2.6)$$

where  $\tilde{\mathbf{P}}^a$  is the  $k \times k$  matrix given by:

$$\tilde{\mathbf{P}}^a = \left( (k-1)\mathbf{I} + \sum_{i=1}^n \mathbf{Y}_i^{b\top} \mathbf{R}_i^{-1} \mathbf{Y}_i^b \right)^{-1}. \quad (2.7)$$

These equations correspond to the Kalman filter analysis mean and covariance equations; here the background mean and covariance for  $\mathbf{w}$  are  $\mathbf{0}$  and  $(k-1)^{-1}\mathbf{I}$ , respectively. Now, the model state corresponding to (2.6) is the mean analysis state:

$$\bar{\mathbf{x}}^a = \bar{\mathbf{x}}^b + \mathbf{X}^b \mathbf{w}^a, \quad (2.8)$$

The analysis ensemble is generated as follows:

$$\mathbf{x}^{a(i)} = \bar{\mathbf{x}}^a + \mathbf{X}^b \mathbf{W}^{(i)}. \quad (2.9)$$

where  $\mathbf{W}^{(i)}$  is the  $i$ th column of the matrix  $\mathbf{W} = \left[ (k-1)\tilde{\mathbf{P}}^a \right]^{1/2}$ . The forecasts from these analysis ensemble states are then used for the next analysis as the background ensemble states.

We remark that a cost function similar to (2.5), but without the linear approximation (2.4) to the observation operator, is used in the Maximum Likelihood Ensemble Filter (Zupanski, 2005). The analysis (2.6) - (2.9) is equivalent to the Ensemble Transform Kalman Filter (Bishop *et al.*, 2001) with the Centered Spherical Simplex Ensemble (Wang *et al.*, 2004). It is also equivalent, though formally less similar, to the analysis described in Hunt *et al.* (2004) and Ott *et al.* (2004). In this paper, the analysis is performed locally like in Ott *et al.* (2004), as described in the next section.

## 2.3 Results

We test both 4D-VAR and 4D-LETKF on Lorenz-96, a toy model with variable  $\mathbf{x}$  in  $m$ -equally spaced points around a circle of constant latitude. The  $j$ th component is propagated in time following differential equation:

$$\frac{d\mathbf{x}_j}{dt} = \frac{1}{120} [(\mathbf{x}_{j+1} - \mathbf{x}_{j-2})\mathbf{x}_{j-1} - \mathbf{x}_j + F] \quad (2.10)$$

where  $j = 1, \dots, m$  represents the spatial coordinate (“longitude”).

Following Lorenz (1996), like the study of Ott *et al.* (2004), we choose the external forcing to be  $F = 8$  and the number of spatial elements to be  $m = 40$ . We have inserted the factor  $\frac{1}{120}$  so that, according to Lorenz’s estimate, the time scale of the dynamics roughly matches that of a global weather model, with  $t$  measured in hours. We use a fourth-order Runge-Kutta scheme for time integration of (2.10) with timestep  $\Delta t = 1.5$  hours.

We perform all simulations by assuming a perfect model scenario. That is, a long integration from an arbitrary initial condition is assumed to be the “true” state. We create the observations,  $\mathbf{y}^o$ , by adding uncorrelated random noise with standard Gaussian distribution (mean 0, variance 1) to the true state. To simulate the asynchronous observations, we make 10 uniformly distributed observations at every timestep (1.5 hours). We rotate the observation locations so that for every 6 hour period, we make one observation available at each model grid point.

To ensure consistency between the 4D-VAR and 4D-LETKF experiments, the assimilation experiments use the same truth and observations. At each analysis time we compute the analysis error as the Root-Mean-Square (RMS) difference between

the true state and the analysis ensemble mean for 4D-LETKF (or simply the analysis state for 4D-VAR). Furthermore, we also compute the RMS difference between the mean forecast from the analysis ensemble for 4D-LETKF (or the forecast from 4D-VAR) and the true state at 6 hour intervals up to a 5 day forecast. We then average the analysis errors and the forecast errors over time by taking the RMS over  $T/1.5n$  analysis cycles, where the analysis is done every  $n$  timesteps ( $1.5n$  hours) and  $T$  is the total length of the simulation. In our simulations, we choose  $T = 120,000$  hours (approximately 13.5 years). We vary the value of  $n$  to examine how the analysis error depends on the time between analyses for 4D-LETKF and on the analysis time window for 4D-VAR.

For the 4D-VAR experiments in this perfect model scenario, we obtain the constant background error covariance matrix  $\mathbf{B}$  for each analysis window iteratively. We initially run 4D-VAR for  $T/1.5n$  analysis cycles using an arbitrary background covariance matrix  $\mathbf{B}_0$  and compute the covariance  $\mathbf{B}_1$  of the differences between the true and analysis states at all of the analysis times. Next, we run 4D-VAR using  $\mathbf{B}_1$  as the background error covariance matrix and again compute the covariance  $\mathbf{B}_2$  of the differences between the truth and background. We repeat this process until the average analysis error does not change significantly. To ensure optimality, we then replace the error covariance matrix found by the iterative algorithm,  $\mathbf{B}$ , with  $\alpha\mathbf{B}$  and tune  $\alpha$  to empirically minimize the analysis RMS error. For all analysis windows in this study,  $\alpha$  was found to be close to one. This covariance matrix is similar to that used for the constant error covariance for OI by Wang *et al.* (2006). In this scenario, we found that the analysis errors obtained from this method were

similar to the analysis errors obtained when using a covariance matrix generated with the NMC method (Parrish and Derber, 1992).

For 4D-LETKF, we obtain the analysis ensemble at each grid point by computing the 4D-LETKF equations using only the observations within a local region. That is, for each grid point we make a separate computation of (2.6)-(2.9) using only the rows and columns of  $\mathbf{Y}_l$  and  $\mathbf{R}_l$  corresponding to the observations in its region. Following Ott *et al.* (2004), the local region for a grid point is centered at that grid point and contains a total of 13 grid points. We present results produced from an ensemble of 15 members. Similar to Ott *et al.* (2004), we found that additional ensemble members did not significantly benefit the analysis when using this localization. We apply multiplicative variance inflation, as in (Whitaker and Hamill, 2002), to compensate for the effects of model nonlinearity and limited ensemble size. Multiplicative inflation replaces the background covariance matrix  $\mathbf{P}^b$  with  $(1+r)\mathbf{P}^b$  for some  $r > 0$ ; we do this in LETKF by replacing  $(k-1)\mathbf{I}$  with  $(k-1)\mathbf{I}/(1+r)$  in (2.7). We tune  $r$  to minimize the analysis RMS error. For these studies, the optimal inflation factor  $r$  increased with the length of the analysis window.

We show the average analysis error as a function of the analysis time window in Figure 2.1 for 4D-LETKF (solid) and 4D-VAR (dashed). For 4D-LETKF, which should not use the same observation in more than one analysis, the analysis window must correspond to the time between analyses, but for 4D-VAR the time between analyses can be chosen independently of the analysis window. We see that the average analysis error of 4D-LETKF grows with the time between analyses, and the average analysis error of 4D-VAR decreases with the length of the analysis window.

The 4D-VAR scheme remains stable for a larger range of analysis windows than 4D-LETKF. The mean analysis RMS error of 4D-LETKF appears to saturate at approximately 0.23 for analysis windows between 0.25 and 1 days, while for 4D-VAR the mean analysis error appears to approach a similar value for analysis windows between 4 and 4.5 days.

For the model we consider here it is computationally feasible to use a large enough ensemble that no localization is necessary. In Figure 2.1 we also display results for 4D-LETKF (dot-dashed) with 50 ensemble members and no localization; that is, the analysis at each grid point uses all of the observations. In this case, the average analysis error is 5–10% better than for the smaller (15 member) ensemble with localization.

To further compare the analysis of 4D-VAR and 4D-LETKF, we compare their average forecast errors as a function of forecast time in Figure 2.2. The average forecast error at initial time 0 is the average analysis error. For 4D-VAR we forecast from the 4 day window analysis (dashed), while for 4D-LETKF we run the forecast from the 1 day window analysis. For each method, the analysis window we use yields near optimal results (for that method) according to Figure 2.1. Because 4D-LETKF provides initial conditions for an ensemble forecast, consider both the mean of an ensemble forecast (solid) and a forecast from the mean analysis state (dot-dashed) to compare with the 4D-VAR forecast. We observe that the 4D-VAR and 4D-LETKF forecasts from the analysis mean have comparable mean RMS errors. The advantage of the ensemble forecast becomes apparent after two and a half days. (We expect that an ensemble forecast initialized with appropriate perturbations

from the 4D-VAR analysis would yield a similar advantage.)

## 2.4 Summary

In assimilating the asynchronous observations considered here, 4D-VAR and 4D-LETKF yield comparable average analysis and forecast errors when 4D-VAR is performed with a long enough analysis time window and when 4D-LETKF is performed sufficiently frequently. In an operational setting, the time between analyses (and thus the time window for 4D-LETKF) may not be adjustable, while the 4D-VAR analysis window can be chosen as large as computational constraints allow. For the scenario of this paper, our results indicate that if the desired time between analyses is 1 day or less, the mean analysis from 4D-LETKF with 15 ensemble members is of similar quality to the 4D-VAR analysis with a 4 day time window. Our results for a 50 member ensemble, shown in Figure 2.1, indicate that the loss of accuracy in 4D-LETKF due to a limited ensemble size does not depend significantly on the time between analyses. Since this large ensemble allows for full rank covariances and a global analysis, we presume that the loss of accuracy as the time between analyses increases is due primarily to model nonlinearity in the following sense: the longer the time between analyses, the greater the background ensemble spread becomes during the analysis time window, and, therefore, the less accurate it is to approximate model trajectories as linear combinations of ensemble trajectories. While 4D-VAR is not strongly affected by nonlinearity, in practice model error will affect both methods. Its effect will increase the errors as the analysis time window

grows.

The introduction of model error could further distinguish the two 4D data assimilation schemes. 4D-VAR generally accounts for model error by using the “weak constraint formulation”, which adds additional terms of the cost function. In an Ensemble Kalman Filter, additional amounts or different types of variance inflation are generally used to counteract model error. While this is the simplest approach to take in 4D-LETKF, it is also possible to minimize a modified cost function in the space spanned by the ensemble trajectories.

As implemented in this paper, both 4D-VAR and 4D-LETKF assimilate the asynchronous observations at comparable computational cost. However, the implementation of 4D-LETKF can be dramatically sped up by computing the analysis for each grid point in parallel (Szunyogh *et al.*, 2005). Furthermore, like other Ensemble Kalman Filters, the 4D-LETKF data assimilation scheme is model independent and thus can easily be adapted to new and evolving models without the human cost involved in determining the adjoint of the model for 4D-VAR.

## 2.5 Figures

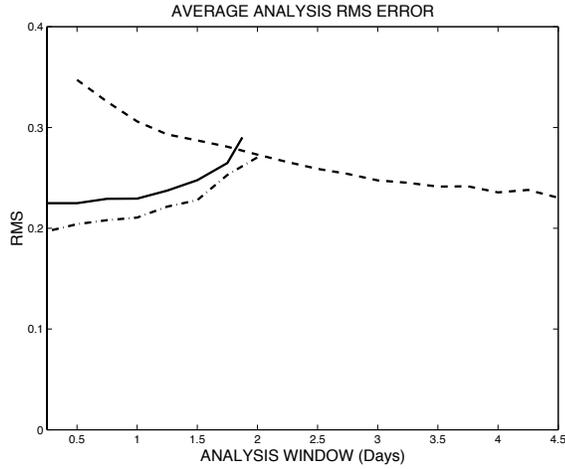


Figure 2.1: The mean analysis error as a function of analysis window for 4D-VAR (dashed). The mean analysis error is shown as a function for time between analyses for 4D-LETKF with 15 ensemble members and localization (solid) and 50 ensemble members and no localization (dot-dashed). The 15 (50) ensemble member 4D-LETKF experiments used inflation factors of 10% (8%) for the 0.5 day,  $n = 8$ , analysis window, 23% (14%) for the 1 day,  $n = 16$ , analysis window, and 65% (24%) for the 1.75 day,  $n = 28$ , analysis window. The 4D-LETKF analysis diverged for all inflation factors below 100% for analysis windows longer than 1.75 days and 2 days in the 15 and 50 ensemble member experiments, respectively.

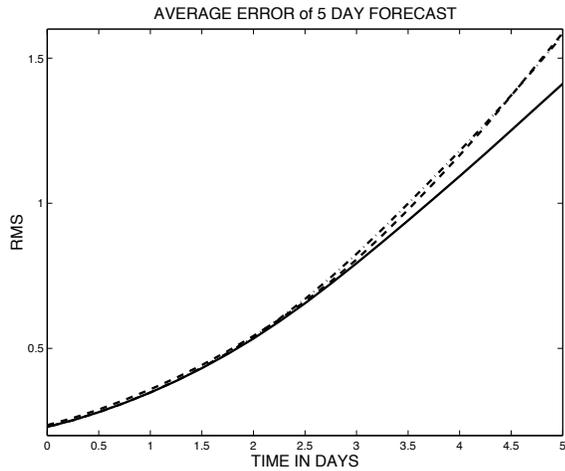


Figure 2.2: Mean forecast errors as function of forecast time for 4D-VAR with 4 day (dashed) and for a 4D-LETKF with 1 day analysis window using an ensemble forecast (solid) and forecast from the mean analysis state (dot-dashed). The mean analysis error of each scheme is at time 0.

## Chapter 3

# Assimilating Nonlocal Observations with a Local Ensemble Kalman Filter

### 3.1 Introduction

Localization is essential to many ensemble-based Kalman filter data assimilation schemes (Houtekamer and Mitchell, 1998; Keppenne, 1999; Anderson, 2001; Hamill *et al.*, 2001; Houtekamer and Mitchell, 2001; Whitaker and Hamill, 2002; Ott *et al.*, 2004). When implementing an ensemble-based scheme globally, spurious correlations in the background ensemble typically arise between distant atmospheric events. These correlations are due to the practical necessity of using a limited number of ensemble members. Local implementations of these ensemble-based schemes eliminate such spurious long-range correlations. Essentially, localization reduces the required ensemble size because it reduces the dimensionality of the estimated state (Szunyogh *et al.*, 2005, 2007).

Local ensemble-based methods typically allow the analysis at a given grid point to depend only on observations within a surrounding local region. The observations in this local region are assumed to be strongly correlated with the model state at the analysis location. Assimilation experiments suggest that choosing a region containing only a subset of the horizontal model grid points and vertical model

levels benefits the ensemble-based Kalman filter analysis (e.g., Hamill *et al.*, 2001; Szunyogh *et al.*, 2005; Houtekamer *et al.*, 2005; Whitaker *et al.*, 2007). However, a challenge arises when assimilating satellite observations in this local framework. Many satellite channels provide radiance observations sensitive to the dynamics over a broad layer of the atmosphere (e.g., Liou, 2002). Consequently, resulting observations are correlated to the atmospheric state over several model levels and selecting which satellite observations to use to perform a given local analysis is not straightforward.

Houtekamer *et al.* (2005) and Houtekamer and Mitchell (2005) address this problem for Advanced Microwave Sounding Unit-A (AMSU-A) radiance observations by treating each observation as if it were a local observation made at the pressure level for which the corresponding satellite channel is most sensitive. With this vertical coordinate, they then assimilate those observations within the local region. While Houtekamer and Mitchell (2005) demonstrated that with this selection strategy the AMSU-A observations have benefited the analysis, it is unknown to what extent the localization prevents the analysis from extracting information from these observations. In particular, an observation assigned to a particular model level may still depend strongly on the dynamics at more distant levels. Therefore, the local analysis at those distant levels may benefit from assimilating that observation. However, that observation may not be selected to update that level because its assigned model level is too distant.

Another potential approach is to assimilate retrievals instead of the radiance observations. We note, however, that satellite retrievals suffer from a complication

similar to that described for radiances. Retrievals invert sets of nonlocal satellite observations (e.g., radiances) to obtain local observations of the analysis (model) variables. The errors of the resulting point observations become correlated over broad layers of the atmosphere. That is, some of the off-diagonal elements of the observational error covariance matrix become nonzero. Thus selecting only those retrievals located within the local region may neglect correlations between observations.

This study examines methods for assimilating observations sensitive to broad layers of the atmosphere with ensemble-based schemes. In particular, we advocate and test the following methodology. For remotely sensed observations, which are related to the dynamics at multiple levels of the atmosphere, we apply the observation operator globally before performing localization. Because the observation operator is applied to the global model state before any localization is performed, there is flexibility in choosing which observations are assimilated in each local analysis. For radiance observations, we can perform the analysis at a given grid point by assimilating any observations related to the dynamics in the associated local region. For retrieval observations, we incorporate a non-diagonal observation error covariance matrix in the assimilation and select those observations located in the local region and any other retrieval observations with errors significantly correlated to those in the local region.

We test the above technique for a mixture of simulated point, nonlocal radiance-like, and correlated retrieval-like observations with the seven level SPEEDY primitive equation model of Molteni (2003). In this paper, we assimilate these obser-

vations with a specific ensemble Kalman filter scheme. In particular, we use the localization set-up proposed in Ott *et al.* (2004) with the algorithmic improvements of Hunt *et al.* (2007). The latter is referred to as the Local Ensemble Transform Kalman Filter (LETKF, Hunt *et al.*, 2007).

A brief introduction to ensemble-based Kalman filter data assimilation techniques and the LETKF scheme are presented in Section 3.2. This section explains the methods of localization employed and then develops a means to select nonlocal observations corresponding to a local region in model space. Section 3.2 also introduces the SPEEDY model, which we use in our perfect model experiments presented in Section 3.3. The results and a summary of conclusions are presented in Sections 3.4 and 3.5, respectively. These sections suggest updating the state at a location by assimilating all nonlocal observations correlated to the model state near that location. These results are obtained for simplified simulated satellite observations using a particular ensemble Kalman filter scheme. However, we expect the results will hold more generally when assimilating real satellite observations with any data assimilation scheme requiring localization.

## 3.2 Background

### 3.2.1 Ensemble Kalman Filter

Kalman filter based data assimilation schemes obtain a best guess (or “analysis”) for the state of the atmosphere by combining a forecast started in the past (called the “first guess” or “background”) and observations. The basic formulation

is as follows (Kalman, 1960; Kalman and Bucy, 1961),

$$\mathbf{P}^a = (\mathbf{I} - \mathbf{K}\mathbf{H})\mathbf{P}^b, \quad (3.1)$$

$$\mathbf{x}^a = \mathbf{x}^b + \mathbf{K}(\mathbf{y} - \mathbf{h}(\mathbf{x}^b)), \quad (3.2)$$

$$\mathbf{K} = \mathbf{P}^b\mathbf{H}^\top (\mathbf{H}\mathbf{P}^b\mathbf{H}^\top + \mathbf{R})^{-1}. \quad (3.3)$$

Here,  $\mathbf{x}^a$  and  $\mathbf{x}^b$  are column vectors containing the analysis and background states, respectively. These states have corresponding error covariance matrices  $\mathbf{P}^a$  and  $\mathbf{P}^b$ . The observations,  $\mathbf{y}$ , are assumed to be given by  $\mathbf{y} = \mathbf{h}(\mathbf{x}^t) + \boldsymbol{\eta}$ , where  $\mathbf{x}^t$  denotes the unknown true state and  $\boldsymbol{\eta}$  represents the measurement error. The corresponding observation error covariance matrix is given by  $\mathbf{R} = \langle \boldsymbol{\eta}\boldsymbol{\eta}^\top \rangle$ . The angle brackets denote an ensemble average over realizations of the random noise vector  $\boldsymbol{\eta}$ . The superscript  $\top$  denotes the matrix transpose. Note that the observation operator  $\mathbf{h}$  converts from model to observation space. This observation operator is linearized around the background mean to obtain a matrix  $\mathbf{H}$  for use in the Kalman filter equations. If the observation operator and model equations are linear, and if  $\boldsymbol{\eta}$  is a zero mean Gaussian random vector, then this Kalman filter analysis is the best linear unbiased estimate of the system state and its error covariance (Kalman, 1960; Kalman and Bucy, 1961)).

After each analysis, one must propagate the analysis state and its covariance forward in time to obtain the background state and covariance for the next analysis. Due to limitations of computational power, the size of the covariance matrices make it essentially impossible to directly evolve these covariances in the context of operational data assimilation for numerical weather prediction. To address this problem,

ensemble based Kalman filter schemes evolve the atmospheric model state starting from a set of  $k$  perturbed initial conditions, creating an ensemble of  $k$  forecasts. Each forecast is denoted by  $\mathbf{x}^{b(i)}$ , where  $i = 1, 2, \dots, k$ . The background state is given by  $\mathbf{x}^b = \overline{\mathbf{x}^{b(i)}}$ , where the over-bar represents the ensemble mean. The covariance is now estimated using the sample covariance

$$\mathbf{P}^b = (k - 1)^{-1} \mathbf{X}^b \mathbf{X}^{b\top} = \frac{\sum_{i=1}^k (\mathbf{x}^{b(i)} - \mathbf{x}^b) (\mathbf{x}^{b(i)} - \mathbf{x}^b)^\top}{k - 1}, \quad (3.4)$$

where the  $i$ th column of the matrix  $\mathbf{X}^b$  is given by  $\mathbf{x}^{b(i)} - \mathbf{x}^b$ . In contrast to some other approaches, this formulation does not require linearization of the model. Furthermore,  $\mathbf{P}^b$  has rank at most  $k - 1$ . Thus, if  $k$  is chosen small enough, the number of degrees of freedom in the Kalman Filter equations can be greatly reduced. However, while this truncation reduces the size of the matrix, it can also introduce severe sampling fluctuations for realistically affordable ensemble sizes  $k$ . As discussed in Section 3.2.2, this latter problem can be efficiently addressed by “localization”.

Equation (3.4) is substituted into the Kalman Filter equations (3.1)–(3.3) to give:

$$\mathbf{P}^a = (k - 1)^{-1} (\mathbf{X}^b - \mathbf{K} \mathbf{Y}^b) \mathbf{X}^{b\top}, \quad (3.5)$$

$$\mathbf{x}^a = \mathbf{x}^b + \mathbf{K} (\mathbf{y} - \mathbf{h}(\mathbf{x}^b)), \quad (3.6)$$

$$\mathbf{K} = (k - 1)^{-1} \mathbf{X}^b \mathbf{Y}^{b\top} [(k - 1)^{-1} \mathbf{Y}^b \mathbf{Y}^{b\top} + \mathbf{R}]^{-1}, \quad (3.7)$$

where  $\mathbf{Y}^b$  is the matrix of ensemble perturbations in observation space. That is,  $\mathbf{Y}^b$  takes the place of  $\mathbf{H} \mathbf{X}^b$  by letting the  $i$ th column of  $\mathbf{Y}^b$  be

$$\mathbf{Y}^{b(i)} = \mathbf{h}(\mathbf{x}^{b(i)}) - \overline{\mathbf{h}(\mathbf{x}^{b(i)})}, \quad (3.8)$$

where  $\overline{\mathbf{h}(\mathbf{x}^{b(i)})}$  represents the ensemble mean of the background ensemble in observation space. This substitution avoids the cost involved in finding the linearized observation operator required by  $\mathbf{H}\mathbf{X}^b$ .

Equations (3.5)–(3.7) form the basis for many ensemble-based Kalman filter algorithms (Evensen, 1994; Burgers *et al.*, 1998; Houtekamer and Mitchell, 1998; Anderson, 2001; Bishop *et al.*, 2001; Whitaker and Hamill, 2002; Ott *et al.*, 2004). However, as posed here these equations are costly to solve because the matrix inverted in Equation (3.7) has as many dimensions as the number of observations, which is typically as large as  $O(10^6)$ . Several methods (i.e., Houtekamer and Mitchell, 2001; Anderson, 2003; Whitaker and Hamill, 2002) reduce the cost of an ensemble-based scheme by assimilating observations sequentially, so that each evaluation of the right hand side of Equation (3.7) involves only the inverse of a scalar. Alternately, several schemes are made efficient by choosing a basis in which they need only invert a matrix of size roughly  $k \times k$  (i.e., Bishop *et al.*, 2001; Ott *et al.*, 2004; Hunt *et al.*, 2007). In this study, we use an efficient modification of the scheme of Ott *et al.* (2004), formulated in Hunt *et al.* (2007) which has been called the Local Ensemble Transform Kalman Filter (LETKF). Following the notation of the latter paper, Equations (3.5)–(3.7) become

$$\tilde{\mathbf{P}}^a = [(k-1)\mathbf{I} + \mathbf{Y}^{b\top}\mathbf{R}^{-1}\mathbf{Y}^b]^{-1}, \quad (3.9)$$

$$\mathbf{P}^a = \mathbf{X}^b\tilde{\mathbf{P}}^a\mathbf{X}^{b\top}, \quad (3.10)$$

$$\mathbf{x}^a = \mathbf{x}^b + \mathbf{X}^b\tilde{\mathbf{P}}^a\mathbf{Y}^{b\top}\mathbf{R}^{-1}(\mathbf{y} - \mathbf{h}(\mathbf{x}^b)). \quad (3.11)$$

Here,  $\tilde{\mathbf{P}}^a$  is the representation of  $\mathbf{P}^a$  in the space spanned by the ensemble pertur-

bations.

All ensemble-based Kalman filter methods must also obtain an analysis ensemble. The LETKF algorithm obtains this analysis for each ensemble member by letting

$$\mathbf{x}^{a(i)} = \mathbf{x}^a + \mathbf{X}^{a(i)}, \quad (3.12)$$

where  $\mathbf{X}^{a(i)}$  is the  $i$ th column of the matrix

$$\mathbf{X}^a = \mathbf{X}^b \left( (k-1) \tilde{\mathbf{P}}^a \right)^{1/2}, \quad (3.13)$$

where by  $\mathbf{M}^{1/2}$  we mean the symmetric square root of a symmetric matrix  $\mathbf{M}$ . Notice that by Equation (3.10),

$$\mathbf{P}^a = (k-1)^{-1} \mathbf{X}^a \mathbf{X}^{a\top}. \quad (3.14)$$

### 3.2.2 Localization

In a local region, the ensemble size may be sufficient to represent a large portion of the state uncertainty (Szunyogh *et al.*, 2005). Therefore, ensemble-based Kalman filter schemes typically apply some form of localization, whose effect is to make the analysis at a given location depend only on the observations within an appropriate region in model space. A feature differentiating ensemble-based data assimilation schemes proposed in the literature is their method of localization. Many ensemble-based Kalman filter techniques assimilate observations sequentially, and perform localization by limiting the influence of an observation on the analysis to a surrounding local region (Anderson, 2001; Hamill *et al.*, 2001; Houtekamer and Mitchell, 2001; Whitaker and Hamill, 2002). In practice, this is accomplished by ta-

pering the entries in the background error covariance matrix to zero outside the local region of influence. We refer to this localization as covariance localization and describe its implementation in the Appendix (Section 3.6). The sequential assimilation of observations neglects correlations between observational errors which, if present, are represented by nonzero off-diagonal entries in the error covariance matrix  $\mathbf{R}$ . If these off-diagonal entries are significant, they can be taken into account by simultaneously assimilating observations in correlated batches (Houtekamer and Mitchell, 2001). Alternatively,  $\mathbf{R}$  can be diagonalized by a coordinate transformation in observation space, and the transformed observations can be assimilated sequentially (Parrish and Cohn, 1985; Anderson, 2003). As we discuss in the Appendix (Section 3.7), diagonalizing  $\mathbf{R}$  can be problematic when observation error correlations are nonlocal, because the transformed observations are not associated with a specific location and thus their influence is hard to locate.

Alternately, Ott *et al.* (2004) employ localization in model grid space. In this scheme, a local region is defined for each model grid point. The local region is centered at that grid point and can, for example, be a box with length  $2l + 1$  model grid points in both horizontal coordinates and depth  $2v + 1$  model levels. The parameters  $l$  and  $v$  should be tuned in a given application to optimize performance. The model state at the center grid point is updated using all the observations within the local box. With this model grid localization, the analysis at each grid point is independent and observations are assimilated simultaneously. These properties enable parallel implementations of grid point localization to be realized naturally (Szunyogh *et al.*, 2005, 2007). Because of this computational advantage, we use grid

point localization in this paper’s numerical experiments. However, we expect the results obtained to hold regardless of which localization method is employed.

### 3.2.3 Assimilating Nonlocal Observations

In this section, we discuss options for selecting which nonlocal observations to use in the local model grid analysis for a given grid point. Because we may select observations that depend on model variables outside the local region defined for uncorrelated point observations, we apply the observation operator to the global model state for each ensemble member. We include only those rows of  $\mathbf{Y}^b$  and  $\mathbf{h}(\mathbf{x}^b)$  corresponding to the selected observations, truncate the observation error covariance matrix accordingly, and solve the resulting ensemble-based Equations (3.9)–(3.11) exactly.

Selecting which satellite radiances to assimilate is complicated by the fact that they do not have a single well-defined vertical location. We note that the weighting function at a particular model point indicates the sensitivity of that observation to the state at that model grid point (Liou, 2002). Houtekamer *et al.* (2005) and Houtekamer and Mitchell (2005) choose to assign radiance observations to the model level for which the magnitude of the weighting function is largest. They use this location to select observations within the local region in model space. Throughout this paper, we refer to this method for selecting radiance observations as “maximum-based selection”. We represent this scheme graphically in Figure 3.1(a).

An alternative to the maximum-based selection is to examine the weighting

function for each nonlocal observation, and choose to assimilate it if a “significant” weight is assigned to any model state vector component within the local region; otherwise, this observation is not used in the assimilation. We refer to this observation selection as “cutoff-based selection” and represent it graphically in Figure 3.1(b). Maximum-based selection may be regarded as a particular case of cutoff-based selection in which the only significant weight is the largest weight. The Appendix (Section 3.6) applies the cutoff-based selection to covariance localization schemes. In Section 3.3, we run a series of numerical experiments to compare the maximum-based and cutoff-based observation selection schemes. We note that we avoid the additional complication of the nonlinearity of the satellite observation operator in this study by letting the observation operator be a linear weighting function.

As mentioned earlier, satellite retrievals are nonlocal in that their errors can be strongly correlated across significant distances (e.g., Kalnay, 2002). These correlations are manifested in nonzero values in the off-diagonal terms of the observation error covariance matrix,  $\mathbf{R}$ . The magnitude of the correlation ( $\mathbf{R}_{ij}/\sqrt{\mathbf{R}_{ii}\mathbf{R}_{jj}}$ ) between the errors of two observations,  $i$  and  $j$ , can be used to select which correlated observations to assimilate. That is, we assimilate all the retrieval observations located within the local region and all other retrieval observations for which the magnitude of the correlation of their errors with the errors of any observation in the local region is above a certain threshold. We retain the entries in the observation error covariance matrix corresponding to the selected retrievals and use the resulting error covariance matrix as  $\mathbf{R}$  in the analysis (Equations (3.9)–(3.11)).

In the numerical experiments in Section 3.4, we tune the threshold value used

to select retrieval observations. For comparison, we also assimilate only observations inside the local region, neglecting the correlations with more distant observations. In this case, we scale the entries of the observation error covariance by a constant value and tune that constant to obtain the analysis with the smallest error.

### 3.2.4 SPEEDY Model

The primitive-equations-based SPEEDY (short for Simplified Parameterizations, primitivE-Equation DYnamics) model is a global atmospheric circulation model (Molteni, 2003). The model inputs are zonal wind,  $u$ , meridional wind,  $v$ , absolute temperature,  $T$ , specific humidity,  $q$ , and surface pressure,  $p_s$ , at a resolution of 96 zonal grid points, 48 meridional grid points, and seven sigma levels (0.950, 0.835, 0.685, 0.510, 0.340, 0.200, 0.080). To solve the equations, the input variables are converted to the vertical component of vorticity  $\zeta$ , horizontal divergence  $\nabla$ ,  $T$ ,  $q$ , and  $\log(p_s)$  at seven sigma levels. The primitive equations are formulated in terms of these variables and integrated in spectral space. Triangular truncation is performed at wavenumber 30 (T30). The resulting forecasts are converted back into the model input variables.

As detailed in Molteni (2003), the SPEEDY model uses the 1981-90 reanalysis from the European Centre for Medium-Range Weather Forecasts (ECMWF) to obtain its boundary conditions at the bottom of the atmosphere for fields including topographical height, land-sea mask, sea ice fraction, sea surface temperature (SST), snow depth, bare-surface albedo, surface temperature and moisture in the top soil

layer, and land-surface covered by vegetation. The SPEEDY model utilizes simplifications of the basic physical parameterization schemes present in many general circulation models (GCMs) for processes such as convection, large-scale condensation, clouds, short-wave radiation, long-wave radiation, surface fluxes, and vertical diffusion. We employ the SPEEDY model for the numerical experiments in this study because it retains many of the characteristics of operational GCMs, but at orders of magnitude lower computational cost.

### 3.3 Experimental Design

We explore the methods proposed in Section 3.2.3 for selecting nonlocal observations carrying out perfect model experiments with the SPEEDY model. The known “true” state of the atmosphere is a 180 time step (roughly one and a half month) integration of the SPEEDY model. Following Miyoshi (2005), the initial condition for the truth results from a year long run of the SPEEDY model from its default rest state on January 01, 1981. For consistency, we assimilate the same observations of this truth for all of the numerical experiments. To determine the quality of the assimilations, we convert the analysis and true states to pressure coordinates, at pressure levels (925, 850, 700, 500, 300, 200, 100 hPa). We then compute the global root mean square (RMS) error between the true and analysis states at each pressure level. The average RMS error refers to the average of the RMS error over the last month of the assimilation cycle.

Every six hours, we create simulated local and nonlocal remotely-sensed ob-

servations. We obtain local “point” observations by adding zero mean Gaussian-distributed noise to the modified truth. The noise has standard deviation 1 m/s for  $u$  and  $v$ , 1 K for  $T$ , 0.0001 kg/kg for  $q$ , and 1 hPa for  $p_s$ . Following Miyoshi (2005), we simulate a rawinsonde network by retaining only those point observations which are closest to the real rawinsonde locations at the 12Z synoptic time. For example, the corresponding observation distribution for  $p_s$  is shown in Figure 3.2.

We obtain simulated remotely sensed radiances by computing weighted averages of  $T$ . For given horizontal location  $(i, j)$ , the radiance-like observations are of the form  $\mathbf{y}_{(i,j)} = \mathbf{H}\mathbf{x}_{(i,j)}^t + \boldsymbol{\eta}$ . The vector  $\mathbf{x}_{(i,j)}^t$  contains the “true” temperature for all levels at the horizontal grid point  $(i, j)$ . The noise,  $\boldsymbol{\eta}$ , is a zero mean Gaussian distributed random vector with zero mean and standard deviation 2 K. We choose a larger standard deviation for satellite observations than for gridpoint observations to simulate the relative difficulties associated with assimilating real satellite observations. The observation resulting from applying a single row of  $\mathbf{H}$  to the model state  $\mathbf{x}_{(i,j)}^t$  represents a single satellite “channel”. That is, for channel  $n$ , the simulated radiance observation is

$$y_{(i,j,n)} = \sum_{m=1}^7 \mathbf{H}_{nm} x_{(i,j,m)}^t + \eta, \quad (3.15)$$

where  $x_{(i,j,m)}^t$  is the “true” temperature at the horizontal location  $(i, j)$  and model level  $m$  and  $y_{(i,j,n)}$  is the corresponding radiance observation for channel  $n$ . For this study, we have as many radiance observations as model levels (i.e.,  $n$  runs from 1 to

7). We take  $\mathbf{H}_{nm} = 2^{-|n-m|-1}$  if  $|n - m| \leq 3$ , and 0 otherwise; so that

$$\mathbf{H} = \begin{bmatrix} 1/2 & 1/4 & 1/8 & 1/16 & 0 & 0 & 0 \\ 1/4 & 1/2 & 1/4 & 1/8 & 1/16 & 0 & 0 \\ 1/8 & 1/4 & 1/2 & 1/4 & 1/8 & 1/16 & 0 \\ 1/16 & 1/8 & 1/4 & 1/2 & 1/4 & 1/8 & 1/16 \\ 0 & 1/16 & 1/8 & 1/4 & 1/2 & 1/4 & 1/8 \\ 0 & 0 & 1/16 & 1/8 & 1/4 & 1/2 & 1/4 \\ 0 & 0 & 0 & 1/16 & 1/8 & 1/4 & 1/2 \end{bmatrix}. \quad (3.16)$$

We retain the simulated radiances at every horizontal model grid point  $(i, j)$  to simulate the relative density of satellite observations.

In this preliminary study, we wish to explore the methods proposed in Section 3.2.3 in the case of retrieval observations with known non-diagonal error covariance matrices. Therefore, we consider idealized retrievals for which we can quantify the full observation error covariance matrix, without the complications that typically arise from horizontal correlations in retrievals or correlations between the retrieval and background errors. Because our observation operator  $\mathbf{H}$  is invertible, we do not need to use any background information in our retrieval scheme, so that the resulting retrievals can be assimilated correctly with an appropriate non-diagonal error covariance matrix (Joiner and da Silva, 1998). Accordingly, we create idealized retrieval observations  $\mathbf{y}_{(i,j)}^{\text{ret}}$  from our simulated radiance observations  $\mathbf{y}_{(i,j)}$ , which we obtain by solving at each horizontal grid point the equation  $\mathbf{y}_{(i,j)} = \mathbf{H}\mathbf{y}_{(i,j)}^{\text{ret}}$ . Thus the observation error covariance corresponding to the resulting retrievals,  $\mathbf{R}^r$ , becomes  $\mathbf{R}^r = \mathbf{H}^{-1}\mathbf{R}\mathbf{H}^{-1\top}$  (e.g., Joiner and da Silva, 1998). Here,  $\mathbf{R}$  is the observation error

covariance matrix of the radiance observations, which is four times the identity in this case. The resulting observation error covariance matrix for the retrievals at a fixed horizontal point is shown in Table 3.1. Retrieval observation errors at two different horizontal grid points are uncorrelated. However, those at the same horizontal location but at different model levels are correlated. The magnitude of the correlation between retrieval observation errors at a single horizontal point is shown in Table 3.2.

The significant anti-correlation between errors at adjacent levels can be reduced by blending the observational information with a background state, based on the climatology of the model, similarly to operational practice. However, as we discussed above, this introduced other error correlations that are harder to account for in data assimilation. Indeed, we find that “regularizing” the retrievals in this manner increased the analysis errors, despite the reduced vertical correlations. We discuss these results further in Section 3.5.

The ensemble-based Kalman filter experiments on the SPEEDY model adapt the LETKF code from Miyoshi (2005) and Harlim and Hunt (2007) to assimilate nonlocal observations. Based on Miyoshi (2005), all the experiments presented use 20 ensemble members and a local region containing 7 meridional and 7 zonal grid points. We determine the optimal vertical depth of the local regions from assimilation experiments using the local “point” observations alone. In these studies, we fix the depth of the local region and assimilate all point observations within that region. We then use the local region resulting in the smallest analysis error as the basis for applying the observation selection schemes when assimilating the simulated

nonlocal radiance and retrieval observations.

## 3.4 Results

### 3.4.1 Grid Point Observations

We first assimilate the local “rawinsonde” observations to determine the optimal depth of the local regions. We consider a variety of depths for the local regions and assimilate all the local observations within these regions. In Figure 3.3, we compare the analysis RMS error of zonal wind (panel (a)) and temperature (panel (b)) for local regions one level deep (solid) and a scheme which uses local regions one level deep at the surface and top of the model and local regions three levels deep elsewhere (dashed). The analysis RMS error is smallest when the local regions are only one level deep. Using local regions containing more than three levels further degraded the analysis (results not shown). For the ensemble size of 20 and horizontal localization of  $7 \times 7$  grid points, we find that these results hold over all analysis times and for all other model variables (results not shown).

### 3.4.2 Radiance Observations

Figure 3.4 shows the time averaged RMS error as a function of vertical level for a mixed observing network of rawinsonde and radiance observations and an observing network of rawinsonde observations alone (thick solid). We compare the cutoff-based and maximum-based selection schemes for the radiance observations. Specifically, in the cutoff-based selection scheme, we assimilate those channels of the

radiance observations for which the observation operator has a nonzero weight (thin solid line), a weight of at least 0.125 (dash-dotted), and a weight of at least 0.25 (dashed). In the maximal-based selection scheme, we assimilate those channels for which the maximal weight 0.5 (dotted) within the local region.

Regardless of the observation selection scheme used to assimilate the radiances, adding them to the observing network reduces the analysis RMS error for both the zonal wind and temperature. The RMS error for the temperature, the analysis variable most directly related to the radiances, does not vary dramatically with the number of observations selected for assimilation. However, the RMS error of the zonal wind is larger for the maximum-based selection scheme than for the cutoff-based selection schemes. For the cutoff-based selection scheme, the zonal wind RMS error generally decreases when additional radiances with smaller magnitudes of the weighting function in the local region are added to the assimilated data set. However, there is a slight disadvantage to assimilating radiance observations from all channels with nonzero weight in the local region. The analysis RMS error of the remaining variables behaves similarly to that of the zonal wind (results not shown). Furthermore, it is possible that the maximum-based observation selection scheme would perform better if it were assimilating more satellite radiances. However, we find that extending the depth of the local regions and applying the maximum-based selection to the extended local region degrades the analysis (results not shown).

### 3.4.3 Retrieval Observations

We also update 1-level deep local regions by assimilating simulated rawinsonde observations and retrievals. In this case, we vary the number of retrieval observations assimilated in any local region and truncate the observation error covariance matrix,  $\mathbf{R}$ , to reflect the choice of observations. Specifically, we assimilate retrieval observations only within the local region (dotted), observations correlated to those in the local region with a correlation coefficient of magnitude at least 0.25 (dashed), 0.15 (dot-dashed), and 0.05 (thin solid line). The resulting average analysis RMS error at each level is shown in Figure 3.5. The thick solid line shows the RMS error from assimilating the rawinsondes alone. Assimilating those retrievals located in the local region has a positive impact on the analysis. However, we find that increasing the number of retrieval observation assimilated and incorporating their error covariance matrix further improves the analysis for temperature, zonal wind, and the other model variables.

For these simulated retrievals, Table 3.2 shows that the correlation coefficient between neighboring observation errors is roughly  $-0.6$ , indicating that they are strongly anti-correlated. By assimilating observations only within the local region, we are neglecting these strong (anti-)correlations between observation errors and, therefore, not we do not obtain an optimal analysis. On the other hand, assimilating retrievals whose errors are correlated and incorporating a non-diagonal error covariance matrix can account for these strong (anti-)correlations accurately.

When assimilating real retrieval observations, elements of the matrix  $\mathbf{R}$  are

unknown. Often,  $\mathbf{R}$  is assumed to be a diagonal matrix, whose entries are tuned to optimize the analysis. Because our local region is only one level deep, when we assimilate only those retrievals located strictly within the local region, we in fact use only the diagonal entries of  $\mathbf{R}$  when assimilating only those retrievals located strictly within the local region. Therefore, we tried scaling these diagonal entries by factors both larger and smaller than one, but found that the analysis was not significantly improved (results not shown).

On the other hand, when we assimilate correlated retrievals from outside the local region, the results in Figure 3.5 do use nonzero off-diagonal error covariances. We repeated these experiments setting the off-diagonal entries in  $\mathbf{R}$  to zero, and found that the analysis errors got larger but were significantly better than the errors obtained using only retrievals in the local region (results not shown). Thus, even when the correlations between retrieval errors cannot be quantified, including retrievals whose errors are correlated to those in the local region may improve the analysis.

In our experiments, the observation operator and corresponding error covariance are known perfectly for both the simulated radiances and retrievals. Therefore, the simulated radiance and retrievals contain the same information and should provide similar analyses when using the same amount of that information. To verify whether this is the case for our data assimilation system, we compare the analysis RMS error for the best results obtained from assimilating the simulated rawinsonde and radiance observations (using cutoff-based selection to assimilate those radiances with a weight of at least 0.125 in the local region, solid line) to those obtained from

assimilating the simulated rawinsonde and retrieval observations (selecting observations in the local region and those retrieval observations with a coefficient of magnitude at least 0.15, dashed-line). Figure 3.6 shows the RMS error as a function of vertical level. In both fields, the RMS error is similar at all levels. Similar results hold for the other analysis variables (results not shown). The comparable accuracy of the analyses from assimilating radiances and retrievals suggests that in each case we are using the observation information consistently when we extend the local region for the satellite observations and also incorporate the observation error covariance matrix for retrievals.

### 3.5 Conclusions

In this paper, we compare different vertical localization approaches when using an ensemble Kalman filter to assimilate a combination of simulated rawinsonde and satellite observations for the SPEEDY model. In Section 3.4 we compare the RMS analysis error for each approach. We have also computed but do not include here forecast errors, which yield a similar comparison.

Using a 20 member ensemble, we find that the analysis error associated with assimilating simulated rawinsonde observations is smallest when using 1-level deep local regions. The dramatic increase in the analysis error from increasing the depth of local regions reflects the coarse vertical resolution of the seven level SPEEDY model. Increasing the depth of the local regions here by even one grid point in depth increases their extent by hundreds of hPa. In more realistic weather models,

increasing the depth of the local regions does not cause the local regions to span such a large extent of the atmosphere. Therefore, for more realistic weather models, slightly increasing the depth of the local regions has only a slight impact on the quality of the analysis. However, dramatically increasing the depth of these local regions degrades the analysis in perfect model experiments with the NCEP GFS model (Szunyogh *et al.*, 2005). Houtekamer *et al.* (2005); Whitaker *et al.* (2007); Szunyogh *et al.* (2007) also find that significant vertical localization is necessary to get the best analysis from rawinsonde observations.

The numerical experiments described in Section 3.4 suggest that when using vertical localization, a properly chosen cutoff-based selection for which radiance-like observations to assimilate is better than either maximum-based selection. With cutoff-based selection, we update the model state at a given location by assimilating radiances for which the value of the weighting function is "sufficiently" large in the local region from which rawinsonde observations are assimilated, rather than using only the radiances for which the maximum of the weighting function is in the local region as in maximum-based selection (see Section 3.2.3). In our experiments with simulated rawinsonde and radiance observations, we obtain the least analysis error using cutoff-based selection and a 1-level deep local region. We find that maximum-based selection results in greater analysis error, whether we retain the 1-level deep local region or enlarge it. The advantage of cutoff-based selection is most significant for analysis variables other than the temperature. These variables depend on the observed radiances only through their correlations with the temperature.

We infer that for purposes of vertical localization, radiance-like observations

should be considered wherever they depend significantly on the model state rather than only where they depend maximally on the model state. The numerical experiments here represent a simplified case, in which the observation operator is a known linear weighting function. Radiance observation operators are typically complicated nonlinear functions, of which the weighting function is only a component. Nonetheless, the weighting function can still be extracted and indicates the sensitivity of that observation to the model variables at a given location. One therefore can use cutoff-based selection to choose appropriate radiances for a given local region when using real data.

We note that maximum based selection is equivalent to cutoff based selection with a large cutoff value. Therefore, we expect the analysis from the two schemes to be most different when the cutoff value is small enough and the weighting function is broad enough to incorporate significantly more observations in the analysis than would maximum based selection (Figure 3.1). Accordingly, we expect cutoff based selection to benefit the analysis for radiances with broad weighting functions. In particular, we expect the benefit of using cutoff-based selection to the analysis to be greatest when the vertical distance over which the weighting functions exceeds the cutoff is comparable to or larger than the vertical localization distance. On the other hand, if the weighting function is sharply peaked, both the maximum and cutoff selection schemes will select similar observations for assimilation and, therefore, provide an analysis of comparable quality, regardless of the depth of the local regions.

Our numerical experiments suggest that one also should assimilate retrievals

from a region that is larger than the local region used for the assimilation of point observations. We find that the number of observations selected for assimilation influences the analysis of the model variables more strongly for our simulated retrievals than for our simulated radiances. Using a non-diagonal error covariance matrix and assimilating observations with errors strongly correlated to those in the local region improves the analysis in all variables. The analyses obtained by assimilating the retrieval observations in this way have similar average RMS errors to those obtained with the assimilation of the radiances at the optimal selection of the analysis parameters.

As we discussed in Section 3.3, the simple retrieval algorithm used in this study exaggerates the correlations between errors of the retrieval observations over what they would be when using a typical retrieval algorithm. Nonetheless, the errors of the resulting retrievals would still have significant correlations. Accordingly, we expect that the analysis will benefit from incorporating a non-diagonal observation error covariance matrix and including more retrievals than those in the local region for point observations, although the benefit may be less significant than for the simplified retrievals used in the numerical experiments here.

A further complication from retrieval observations arises because the correlations between their errors are often unknown. In practice, they are generally assumed to be uncorrelated and the diagonal entries of the observation error covariance matrix  $\mathbf{R}$  are tuned to provide the optimal analysis. However, for the simplified retrieval observations used here, we observe that simply scaling the diagonal entries of  $\mathbf{R}$  only slightly improves the analysis in some of the model variables. We gain the

most significant advantage from assimilating those observations with strongly correlated errors simultaneously, using a non-diagonal  $\mathbf{R}$ . Therefore, when retrieval error correlations are unknown, including additional retrieval observations suspected of having errors correlated errors to those in the local region and incorporating these correlations in the estimate for the observation error covariance matrix may still improve the analysis. Similar techniques could be applied to account for known horizontal correlations in retrieval errors.

As we discussed in Section 3.3, many retrieval algorithms introduce correlations between the retrieval observation errors and errors in the background state used in the data assimilation. Therefore, directly assimilating these retrievals may not provide an optimal analysis, as we observed for simulated regularized retrievals. Several techniques for reducing the unwanted observation-background error correlations are described in the literature (e.g., Joiner and da Silva, 1998; Rodgers, 2000) either by modifying the retrieval observations themselves or by introducing a non-local observation operator. In the latter case, a cutoff-based selection similar to the one proposed for radiance observations should be applied. Future studies should consider these techniques for more realistic retrieval observations than the idealized retrievals used here.

In summary, spatial localization is essential to many ensemble-based data assimilation schemes. Typically the local regions used for ensemble-based Kalman filter schemes on operational models contain more levels than those used for the SPEEDY model. However, these regions are still small enough that relevant satellite observations can depend on the model variables outside the local region. The

numerical experiments in Section 3.4 and the discussion of Section 3.2 suggest that one should localize the impact of observations differently depending on the observation type. Nonlocal observations do not necessarily need to be used in the analysis for all model variables on which they depend, but they need to be used in the analyses at the location of all model variables they are related to sufficiently strongly, either through the observation operator (as for satellite radiances) or through correlated errors (as for retrievals). The distance over which nonlocal observations influence the analysis is a parameter that can be tuned independently of the localization distance for point observations. In this paper, we have expressed this parameter as a cutoff for the radiance weighting function or the correlation between retrieval errors, but if these quantities are not known, one can tune the localization distance directly.

### 3.6 Appendix: Covariance Localization

An alternative to the grid point localization scheme in this paper is covariance localization (Anderson, 2003; Bishop *et al.*, 2001; Hamill *et al.*, 2001; Houtekamer and Mitchell, 2001; Whitaker and Hamill, 2002). This localization scheme limits the distance over which each observation affects the analysis by tapering the entries of the background matrix to zero beyond a fixed horizontal distance  $L$  and vertical distance  $V$ . Each entry of the background covariance matrix  $\mathbf{P}^b$  relates two model variables, and here distance is measured between the locations of these model variables. In a sequential assimilation scheme, the term  $\mathbf{P}^b$  in Equation (3.3) then causes

entries in the Kalman gain matrix  $\mathbf{K}$  to be zero beyond the horizontal distance  $L$  and vertical distance  $V$  from the location of the observation(s) being assimilated.

Houtekamer *et al.* (2005) implement this approach by modifying Equation (3.7) such that

$$\mathbf{K} = (k - 1)^{-1} [\rho_V \circ \rho_H \circ (\mathbf{X}^b \mathbf{Y}^{b\top})] [(k - 1)^{-1} \rho_V \circ \rho_H \circ (\mathbf{Y}^b \mathbf{Y}^{b\top}) + \mathbf{R}]^{-1}, \quad (3.17)$$

Here,  $\rho_H$  acts on a covariance matrix by multiplying each entry by a function that depends on horizontal distance, decreasing from one at a distance of 0 to zero at a distance of  $L$  or greater (This function is often chosen to be approximately Gaussian, as in Gaspari and Cohn (1999)). Similarly,  $\rho_V$  multiplies by a function of vertical distance that is zero beyond distance  $V$ . In this implementation, distances are between model variable locations and observation locations in  $\mathbf{X}^b \mathbf{Y}^{b\top}$ , and between observation locations in  $\mathbf{Y}^b \mathbf{Y}^{b\top}$ .

With covariance localization, one can allow nonlocal observations to have a larger region of influence than point observations as follows. To use cutoff-based selection, as described in Section 3.2.3, for vertical localization, one should first determine an appropriate vertical localization distance  $V$  for point observations. For nonlocal observations, increase  $V$  to  $\tilde{V}$ , where  $\tilde{V} - V$  is the distance over which the cutoff (for the weighting function in the case of satellite radiances or for correlation in the case of retrievals) is exceeded. Then, define a modified vertical localization operator  $\tilde{\rho}_V$  that tapers covariances to zero beyond the vertical distance  $\tilde{V}$  by scaling vertical distances by a factor of  $V/\tilde{V}$ . Each nonlocal observation must also be assigned a nominal location; for satellite radiances, one can use the maximum of the

weighting function. Finally, replace  $\rho_V$  by  $\tilde{\rho}_V$  in Equation (3.17) when assimilating nonlocal observations. This will be simplest in a sequential scheme, assimilating nonlocal observations separately from point observations.

### 3.7 Appendix: Diagonalization

In the context of a sequential ensemble-based scheme, Anderson (2003) proposes to account for spatial correlations of observations errors by diagonalizing the observation error covariance matrix. Because the observation error covariance matrix  $\mathbf{R}$  is symmetric, it can be diagonalized by an orthogonal transformation  $\mathbf{R} = \mathbf{S}\mathbf{\Lambda}\mathbf{S}^\top$ . The columns of  $\mathbf{S}$  contain the orthonormal eigenvectors of  $\mathbf{R}$  and  $\mathbf{\Lambda}$  has the corresponding eigenvalues along its diagonal. Recall that observations are given by  $\mathbf{y} = \mathbf{h}(\mathbf{x}^t) + \eta$ , where  $\langle \eta\eta^\top \rangle = \mathbf{R}$ . Instead of assimilating the original retrieval, Anderson (2003) creates a new observation  $\mathbf{y}^* = \mathbf{S}^\top \mathbf{y} = \mathbf{S}^\top \mathbf{h}(\mathbf{x}^t) + \mathbf{S}^\top \eta$ . For this observation,  $\mathbf{R}^* = \langle \mathbf{S}^\top \eta \eta^\top \mathbf{S} \rangle = \mathbf{S}^\top \mathbf{R} \mathbf{S} = \mathbf{\Lambda}$ , which is diagonal. The resulting observations have uncorrelated errors. Their observation operator can be taken to be  $\mathbf{h}^* = \mathbf{S}^\top \mathbf{h}$ . This operator contains all the sensitivity to nonlocal behavior. Therefore, we could select which observations to assimilate to update a given local region by examining the magnitude of the linearized observation operator. We could use this magnitude as the basis of maximum-based or cutoff-based observation selection.

However, applying the eigenvalue decomposition makes it problematic to distinguish observations that are strongly correlated to the dynamics within a local

region. To better understand this we recall that for observations which have only slightly correlated errors, the corresponding off diagonal entries of the observation error covariance matrix,  $\mathbf{R}$ , are small. For example, in the numerical experiments in Section 3.4, the  $\mathbf{R}$  matrix for retrieval observations at a given horizontal grid point is given in Table 1. Its structure suggests that an accurate analysis would result from assimilating only those observations with errors that are most strongly correlated to the errors of those observations located within the local region. However, the corresponding eigenvectors typically do not maintain this local structure. We have found that most entries of these eigenvectors have approximately the same magnitude. For example, for the experiments in Section 3.3, the matrix whose columns correspond to the eigenvectors of  $\mathbf{R}$  is

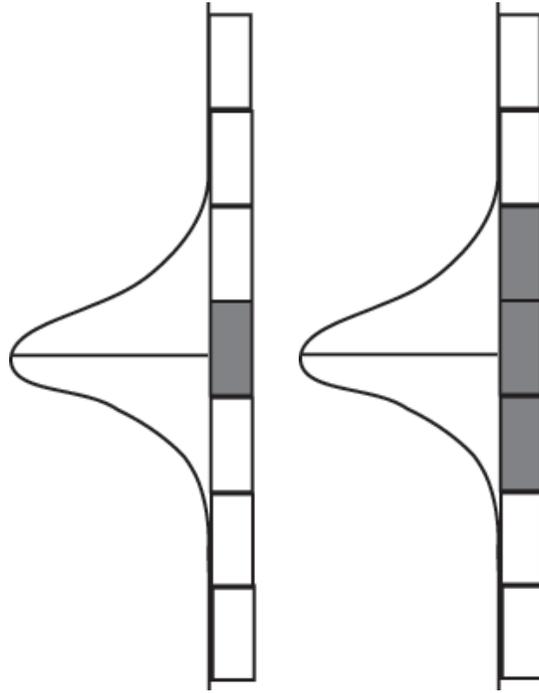
$$\mathbf{S} = \begin{pmatrix} 0.24 & 0.44 & 0.53 & 0.48 & 0.33 & -0.28 & 0.21 \\ 0.36 & 0.47 & 0.24 & -0.12 & -0.39 & 0.52 & -0.40 \\ 0.44 & 0.30 & -0.25 & -0.51 & -0.20 & -0.39 & 0.45 \\ 0.48 & 0.00 & -0.44 & 0.00 & 0.63 & 0.00 & -0.43 \\ 0.44 & -0.30 & -0.25 & 0.51 & -0.20 & 0.39 & 0.45 \\ 0.36 & -0.47 & 0.24 & 0.12 & -0.39 & -0.52 & -0.40 \\ 0.24 & -0.44 & 0.53 & -0.48 & 0.33 & 0.28 & 0.21 \end{pmatrix}. \quad (3.18)$$

Because the magnitude of all these terms is comparable, there would be no means to distinguish which of the observations  $\mathbf{y}^*$  one should assimilate.

Nonetheless, some ensemble-based Kalman filter algorithms require diagonalization to assimilate observations with correlated errors. In these cases, observations can still be selected based on their correlation coefficients, as described in Section

3.2.3. The  $\mathbf{R}$  matrix could then be truncated to reflect this choice of observations. Applying the diagonalization to the truncated  $\mathbf{R}$  matrix would ensure that only those observations with errors most strongly correlated those of observations in the local region are assimilated to update the state in that region.

### 3.8 Figures



(a) Maximum-Based Selection (b) Cutoff-Based Selection

Figure 3.1: A schematic for the observation selection schemes used for nonlocal, radiance-like observations. The curve represents the weighting function at each model level. The rectangles represent several local regions around these model levels. In practice, these regions overlap, though not depicted as such in this figure. The shaded rectangles represent local regions that assimilate this observation.

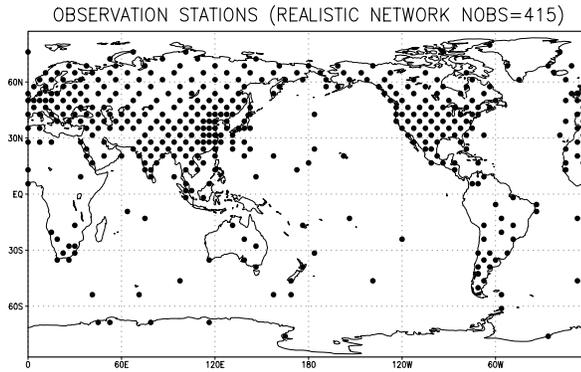
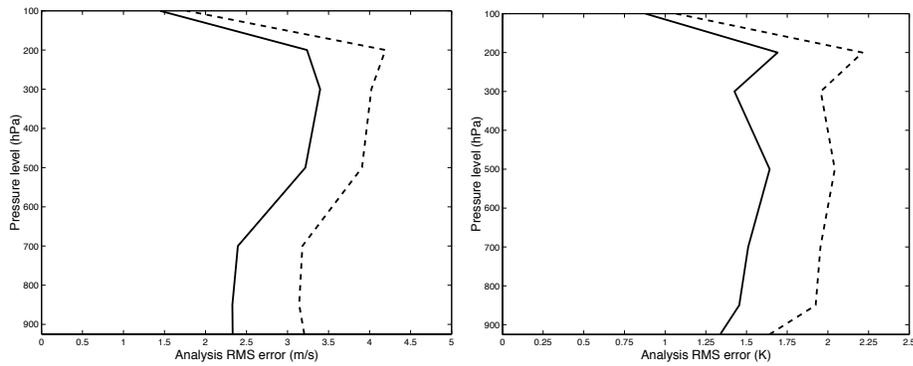


Figure 3.2: For the SPEEDY model experiments in this paper, we simulate rawinsonde observations at the model grid points closest to real rawinsonde locations at the 12Z synoptic time developed by Miyoshi (2005). Here, we show the locations of the model grid points of the rawinsonde stations for the  $p_s$  observations. Similar locations are used for the other variables, and the observation density varies with height as for real rawinsonde observations.



(a) Zonal Wind ( $u$ )

(b) Temperature ( $T$ )

Figure 3.3: The average analysis RMS error as a function of vertical model level from assimilating simulated rawinsonde observations. The depth of the vertical local region contains at most 1 level (solid) and 3 levels except at the model surface and top where local regions are 1 level deep.

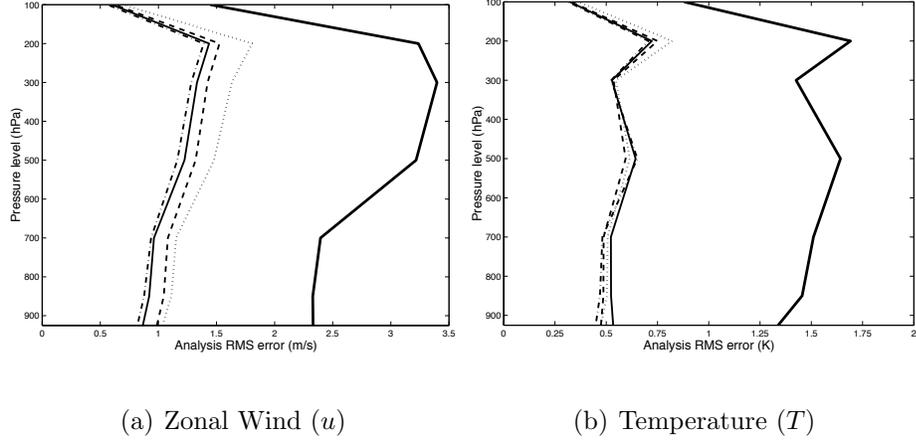


Figure 3.4: The average analysis RMS error as a function of vertical model level in from assimilating simulated rawinsonde observations and satellite radiances. We select rawinsonde observations within a local region 1 level deep (thick solid) and we also assimilate the radiance observations if the value of the  $\mathbf{h}$  operator there is greater or equal to 0.5 (dotted), 0.25 (dashed), 0.125 (dash-dotted), and 0.0625 (thin solid).

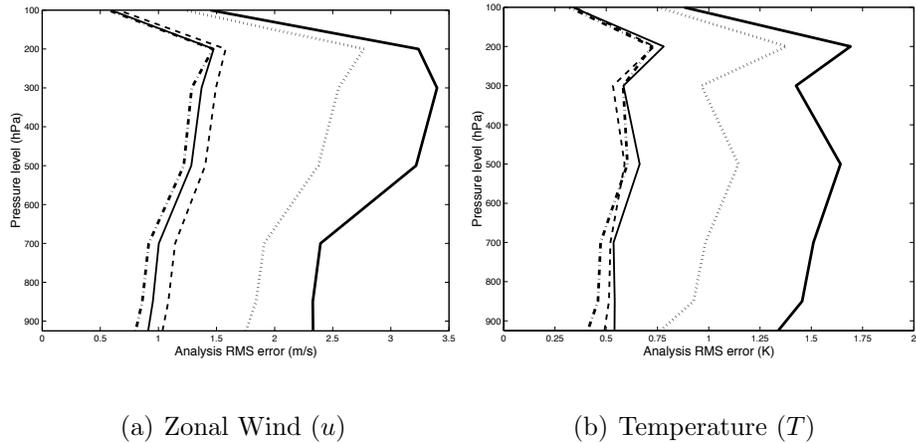
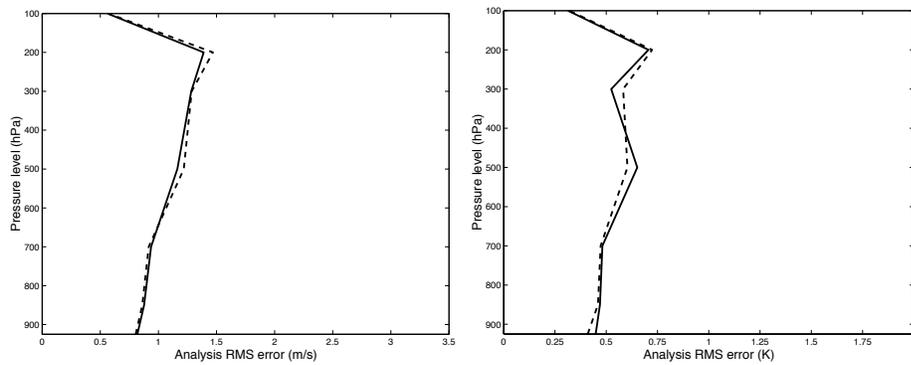


Figure 3.5: The average analysis RMS error as a function of vertical model level from assimilating simulated rawinsonde observations alone (thick solid) and rawinsondes with satellite retrievals. We assimilate retrievals within the local region (dotted) and also those retrieval observations outside the local region with a correlation coefficient of at least 0.25 (dashed), 0.15 (dash-dotted), and 0.05 (thin solid).



(a) Zonal Wind ( $u$ )

(b) Temperature ( $T$ )

Figure 3.6: We compare the best result from assimilating the simulated rawinsonde and satellite radiance observations (solid) and the simulated rawinsonde and satellite retrieval observations (dashed) by plotting the average analysis RMS error as a function of vertical model level.

### 3.9 Tables

Model Level	1	2	3	4	5	6	7
1	36.7	-33.3	8.2	-4.6	9.4	-6.1	1.3
2	-33.3	61.0	-37.5	11.0	-10.8	13.1	-6.1
3	8.2	-37.5	60.5	-36.9	11.5	-10.8	9.4
4	-4.6	11.0	-36.9	59.3	-36.9	11.0	-4.6
5	9.4	-10.8	11.5	-36.9	60.5	-37.5	8.2
6	-6.1	13.1	-10.8	11.0	-37.5	61.0	-33.3
7	1.3	-6.1	9.4	-4.6	8.2	-33.3	36.7

Table 3.1: The observation error covariance for the simulated retrieval observations at a fixed horizontal location.

Model Level	1	2	3	4	5	6	7
1	1.00	-0.70	0.17	-0.10	0.20	-0.13	0.04
2	-0.70	1.00	-0.62	0.18	-0.18	0.21	-0.13
3	0.17	-0.62	1.00	-0.62	0.19	-0.18	0.20
4	-0.10	0.18	-0.62	1.00	-0.62	0.18	-0.10
5	0.20	-0.18	0.19	-0.62	1.00	-0.62	0.17
6	-0.13	0.21	-0.18	0.18	-0.62	1.00	-0.70
7	0.04	-0.13	0.20	-0.10	0.17	-0.70	1.00

Table 3.2: The correlation between the errors of observations at a single horizontal point and model levels  $k$  and  $l$  is given by the entry in this table corresponding to these model levels.

## Chapter 4

# Satellite Observation Bias Correction with an Ensemble Kalman Filter

### 4.1 Introduction

Satellite observations provide information essential to numerical weather prediction, particularly in traditionally data poor regions like the Southern Hemisphere (e.g., Kalnay, 2002). However, satellite radiance observations are subject to biases due to errors in the instrument calibration and in the observation operator that maps the model variables to observation variables (e.g., Eyre, 1992). These biases must be corrected in order to best utilize the information from radiance observations. In this study, we focus on correcting biases from errors in the observation operator.

When assimilating radiance observations directly, many data assimilation schemes rely on offline corrections of these radiance observations (Eyre, 1992; Harris and Kelly, 2001). Offline bias correction schemes, however, have to have some important limitations. For example, their parameters have to be tuned manually and they must be re-tuned to account for new instruments, new forward operators, and biases which evolve as the instrument ages (Dee, 2005). Therefore, some variational data assimilation schemes correct for radiance observation biases as part of the assimi-

lation procedure (Derber and Wu, 1998; Dee, 2005) using *state space augmentation* (Friedland, 1969). These techniques append bias parameters to the model state vector and update these parameters as part of the data assimilation process. Typically the uncertainty in the bias parameters is assumed to be uncorrelated to the uncertainty of the forecast model state (Derber and Wu, 1998; Dee, 2005).

While ensemble schemes have been found to assimilate radiance observations efficiently (Houtekamer *et al.*, 2005; Houtekamer and Mitchell, 2005), their ability to likewise account for radiance biases online has not been demonstrated. Baek *et al.* (2006), Keppenne *et al.* (2005), and Zupanski and Zupanski (2006) apply state space augmentation to estimate low dimensional model biases with an ensemble Kalman filter. The approaches of Baek *et al.* (2006) and Zupanski and Zupanski (2006) also takes into account the correlation between uncertainties in the model bias and uncertainties in the forecast due to uncertainties in the initial conditions and the chaotic nature of the forecast models.

In this paper, we use state-space augmentation for ensemble-based assimilation schemes to estimate radiance observation biases online. We apply this set-up to assimilate simulated unbiased rawinsonde observations and simulated biased AIRS radiances employing, as an example, a specific ensemble Kalman filter data assimilation scheme and a specific atmospheric model. In particular, for our data assimilation scheme we employ the algorithmically efficient Local Ensemble Transform Kalman Filter (LETKF) of Hunt *et al.* (2007) which is based on the localization scheme of Ott *et al.* (2004), and for our atmospheric model we use the “Simplified Parameterizations, primitivE-Equation DYnamics” model (SPEEDY) of Molteni

(2003). To represent the fact that the true form of the satellite radiance biases is unknown in reality, we create the simulated AIRS radiances with one form of bias and assume that the bias has another form during the assimilation. In both cases, we use the prototype Community Radiative Transfer Model (pCRTM) as the observation operator for the AIRS radiances. We simulate biased AIRS radiances by modifying the pCRTM in a nonlinear manner as suggested in Watts and McNally (2004), while our online bias correction estimates coefficients of a linear combination of predictive quantities proposed by Cameron (2003) and Harris *et al.* (2004). With this bias correction scheme, we significantly reduce the observation and analysis RMS errors associated with assimilating these simulated biased AIRS observations.

In Section 2, we describe the typical forms of the bias assumed for satellite radiances and how they are usually corrected. We further describe the LETKF data assimilation scheme and the state-space augmentation bias correction technique for LETKF in this section. We describe the setup for the perfect model experiments we perform with the SPEEDY model in Section 3. The results of these experiments are described in Section 4 and conclusions are provided in Section 5.

## 4.2 Satellite Radiance Biases

Satellite radiance biases arise from measurement biases and errors in the radiative transfer model, which transforms model states into radiance space. Therefore, a radiance observation,  $y$ , is modeled to be of the form

$$y = \tilde{h}(\mathbf{x}^t, \boldsymbol{\beta}) + \eta. \quad (4.1)$$

Here,  $\mathbf{x}^t$  is the true state of the atmosphere,  $\eta$  is the mean zero, random component of the radiance error, and  $\tilde{h}$  is a radiative transfer operator, modified to incorporate bias parameters  $\boldsymbol{\beta}$ . Because it includes the bias parameters,  $\tilde{h}$  will ideally correct for the biases associated with errors in the satellite calibration or the unmodified radiative transfer model,  $h$ . The bias parameters,  $\boldsymbol{\beta}$ , differ for each channel of each instrument (Eyre, 1992; Derber and Wu, 1998; Harris and Kelly, 2001).

Typically, the radiance bias is assumed to have two sources, scan-angle bias and air-mass bias. The scan angle bias depends on latitude and scan angle, while the air mass bias depends on the atmospheric state. We only consider the air-mass bias in this paper, but note that the techniques described in this paper could be applied to estimate the scan-angle bias as well.

Several operational centers, including the National Centers for Environmental Prediction (NCEP) and the European Centre for Medium-Range Weather Forecasts (ECMWF), assume that the air-mass bias is a linear combination of a small set of model state variables, called predictors. That is,

$$\tilde{h}(\mathbf{x}^t, \boldsymbol{\beta}) = h(\mathbf{x}) + \sum_{i=1}^N \beta_i p_i(\mathbf{x}), \quad (4.2)$$

where  $h(\mathbf{x})$  is the radiative transfer model derived from physical principles and each  $p_i$  is one of  $N$  predictors (Eyre, 1992; Derber and Wu, 1998; Harris and Kelly, 2001). Each coefficient  $\beta_i$  is assumed to be a global constant that depends on the satellite instrument and channel. In this study, we use a subset of the predictors found to correct AIRS biases (Cameron, 2003; Harris *et al.*, 2004). Specifically, we choose the following predictors:

- $p_1$ : Constant term (i.e.,  $p_1 = 1$ ).
- $p_2$ : 200 - 50 hPa thickness.
- $p_3$ : 850 - 300 hPa thickness.
- $p_4$ : Surface skin temperature (i.e., model surface temperature).
- $p_5$ : The modeled radiance itself (i.e.,  $p_5 = h(\mathbf{x})$ ).

Throughout the rest of this paper we will refer to each predictor by its index. In the numerical experiments with the SPEEDY model presented in this paper, the thicknesses used for  $p_2$  and  $p_3$  are found by computing the difference between the geopotential height at the appropriate levels. The SPEEDY model obtains the surface skin temperature used for  $p_4$  from the sea-surface temperature over ocean and by modifying the land-surface temperature over land to incorporate the forecasted surface pressure (Molteni, 2003).

Watts and McNally (2004) propose an alternate form for the radiance biases. They choose a physical modification of the radiative transfer model. Normally, the radiative transfer model is given by

$$h(\mathbf{x}) = \tau(0)\kappa(0)B(T(0)) + \int_0^\infty B(T(z))\frac{\partial\tau(z)}{\partial z}dz, \quad (4.3)$$

where  $z = 0$  is at the surface,  $\kappa(0)$  is the emissivity of the surface,  $T(z)$  is the temperature profile associated with the state  $\mathbf{x}$  at the latitude and longitude of the satellite observation, and  $B(T(z))$  is the Planck function. The term  $\frac{\partial\tau}{\partial z}$  is called the *weighting function* because it characterizes the contribution to the radiance from a

given height when combined with the Planck function evaluated there. The function  $\tau(z)$  represents the transmittance from height  $z$  and is given by

$$\tau(z) = \exp\left(-\int_0^\infty \kappa(z)\rho(z)dz\right). \quad (4.4)$$

Here,  $\kappa(z)$  is the absorption coefficient from height  $z$  and  $\rho(z)$  is the density at that height. Watts and McNally (2004) assume that the greatest source of error in the radiative transfer model arises from poor knowledge of  $\kappa(z)$ , and they assume that  $\kappa(z)$  has a constant fractional error,  $\gamma$ . That is, they make the following replacement

$$\kappa \rightarrow \gamma\kappa, \quad (4.5)$$

so that

$$\tau(z) \rightarrow \tau(z)^\gamma. \quad (4.6)$$

To correct biases from other sources, they add a global constant,  $\delta$ , to the modified radiative transfer model. Again,  $\gamma$  and  $\delta$  are assumed to be global constants which depend only on the satellite and the channel. For most AIRS long-wave temperature channels,  $\gamma \approx 1.05$  (Watts and McNally, 2004).

We use the bias form proposed by Watts and McNally (2004) with  $\gamma = 1.05$  and  $\delta = 0$  to simulate biased radiance observations. We use the more standard bias correction form in Equations (4.1) and (4.2) to estimate these biases. The goal of the present paper is not to address which functional form of the bias estimate should be used to correct real radiance biases, but rather to show that we can accurately estimate bias correction parameters with an ensemble Kalman filter.

### 4.2.1 Online Bias Correction

Bias parameters for radiance observations can be estimated online as part of the data assimilation by using state space augmentation (e.g., Derber and Wu, 1998; Dee, 2005). In the data assimilation schemes using such an approach, the analysis updates an augmented state vector,

$$\mathbf{z} = \begin{bmatrix} \mathbf{x} \\ \boldsymbol{\beta} \end{bmatrix}, \quad (4.7)$$

based on the observed information, where  $\mathbf{x}$  is the model state and  $\boldsymbol{\beta}$  is the vector of bias parameters for all satellite instruments and channels. That is, the analysis seeks the augmented state vector,  $\mathbf{z}^a$ , which best fits its background state (or *first-guess*,  $\mathbf{z}^b$ ) and the available observations,  $\mathbf{y}$ .

Because systematic errors in satellites change slowly in time, the bias parameters are typically assumed to remain constant between successive analysis times. That is, the evolution of the bias vectors is modeled by persistence, so that

$$\boldsymbol{\beta}_t^b = \boldsymbol{\beta}_{t-1}^a, \quad (4.8)$$

where  $\boldsymbol{\beta}_{t-1}^a$  is the updated bias vector at the previous analysis time and  $\boldsymbol{\beta}_t^b$  is the background bias vector at the current analysis time.

The state space augmentation approach also requires an estimate for the background error covariance of the augmented state vector,  $\mathbf{z}$ . For simplicity, many 3D-VAR and 4D-VAR implementations assume that the error of each bias parameter is uncorrelated to the errors of the other bias parameters, to the errors of the bias parameters for other satellite channels and instruments, and to the errors of the

model state (Derber and Wu, 1998; Dee, 2005). Such an assumption is not made in the ensemble-based Kalman filter scheme we describe in the next section.

#### 4.2.2 Online Bias Correction with an Ensemble Kalman Filter

Our ensemble Kalman filter bias correction scheme uses an ensemble of augmented state vectors to estimate a flow-dependent background error covariance,  $\mathbf{P}^b$  (Baek *et al.*, 2006). In the following formulation, considering that based on Equation (4.7) each augmented state vector  $\mathbf{z}$  contains the model state  $\mathbf{x}$  and bias parameters  $\boldsymbol{\beta}$ , the modified observation operator is given by  $\tilde{h}(\mathbf{x}, \boldsymbol{\beta}) = \tilde{h}(\mathbf{z})$ . An ensemble of  $k$  forecasts of these augmented state vectors are obtained from model runs starting from a set of  $k$  initial conditions, where  $\mathbf{z}^{b(i)}$  represents the forecast corresponding to the  $i^{\text{th}}$  ensemble member ( $i = 1, \dots, k$ ). Here, forecasts of the atmospheric state portion  $\mathbf{x}$  of  $\mathbf{z}$  are obtained by integrating the atmospheric model, while forecasts of the bias vector portion  $\boldsymbol{\beta}$  of  $\mathbf{z}$  are obtained from persistence (Equation (4.8)). The background error covariance can be obtained from the sample covariance of the ensemble members,

$$\mathbf{P}^b = (k - 1)^{-1} \mathbf{Z}^b (\mathbf{Z}^b)^\top. \quad (4.9)$$

Here,  $\top$  denotes the matrix transpose.  $\mathbf{Z}^b$  is the matrix of ensemble perturbations for the augmented state,

$$\mathbf{Z}^b = \begin{bmatrix} \mathbf{z}^{b(1)} - \mathbf{z}^b & \mathbf{z}^{b(2)} - \mathbf{z}^b & \dots & \mathbf{z}^{b(k)} - \mathbf{z}^b \end{bmatrix}, \quad (4.10)$$

where the “background state”  $\mathbf{z}^b$  is the ensemble mean,

$$\mathbf{z}^b = k^{-1} \sum_{i=1}^k \mathbf{z}^{b(i)}. \quad (4.11)$$

Ensemble schemes generally use the Kalman filter equations to seek the linear combination of the ensemble members which best fits the available observations. Applying this methodology to the ensemble of augmented state vectors, ensemble schemes can simultaneously update atmospheric state and bias parameters. In what follows, we formulate the resulting analysis equations using LETKF (Hunt *et al.*, 2007), but note that a similar formulation could be applied to any ensemble Kalman filter scheme. We choose LETKF in this paper because, like its predecessor LEKF (Ott *et al.*, 2004), it has been found to be an accurate and efficient data assimilation scheme for operational weather models (Szunyogh *et al.*, 2005, 2007; Whitaker *et al.*, 2007). Applying LETKF to the ensemble of augmented state vectors, the analysis equations become

$$\mathbf{z}^a = \mathbf{z}^b + \mathbf{Z}^b \tilde{\mathbf{P}}^a (\mathbf{Y}^b)^\top \mathbf{R}^{-1} \left( \mathbf{y} - \overline{\tilde{\mathbf{h}}(\mathbf{z}^{b(i)})} \right), \quad (4.12)$$

$$\mathbf{P}^a = \mathbf{Z}^b \tilde{\mathbf{P}}^a (\mathbf{Z}^b)^\top, \quad (4.13)$$

$$\tilde{\mathbf{P}}^a = \left[ (k-1) \mathbf{I} + (\mathbf{Y}^b)^\top \mathbf{R}^{-1} \mathbf{Y}^b \right]^{-1}. \quad (4.14)$$

Here,  $\tilde{\mathbf{h}}$  maps the background state into observation space for all the observations to be assimilated. The  $j^{\text{th}}$  row of the vector  $\overline{\tilde{\mathbf{h}}(\mathbf{z}^{b(i)})}$  is the ensemble mean of the background ensemble in observation space for the  $j^{\text{th}}$  observation. If the  $j^{\text{th}}$  observation is believed to be biased, then the corresponding row of the operator  $\tilde{\mathbf{h}}$  is a function of the model state and the bias parameters. Otherwise, the corresponding row  $\tilde{\mathbf{h}}$  is a function of the model state alone. This operator is also used to define the matrix of ensemble perturbations in observation space,  $\mathbf{Y}^b$ . Its  $i^{\text{th}}$  column is given

by

$$\mathbf{Y}_i^b = \tilde{\mathbf{h}}(\mathbf{z}^{b(i)}) - \overline{\tilde{\mathbf{h}}(\mathbf{z}^{b(i)})}. \quad (4.15)$$

The information provided by the analysis error covariance matrix,  $\mathbf{P}^a$ , is used to generate the analysis ensemble consistent with the resulting analysis state. LETKF uses Equation (4.13) to determine the analysis of the augmented state for the  $i^{\text{th}}$  ensemble by

$$\mathbf{z}^{a(i)} = \mathbf{z}^{a(i)} + \mathbf{Z}_i^a, \quad (4.16)$$

where  $\mathbf{Z}_i^a$  is the  $i^{\text{th}}$  column of  $\mathbf{Z}^a = (k-1)^{-1/2} \mathbf{Z}^b \left( \tilde{\mathbf{P}}^a \right)^{1/2}$ .

Computational restrictions often limit the number of ensemble members which can be used to obtain the analysis in Equations (4.12) – (4.14). Therefore, the resulting ensemble of augmented state vectors may not be sufficient to capture the global dynamics. However, a limited number of ensemble members is presumed sufficient to capture the dynamics within local regions (Patil *et al.*, 2001; Ott *et al.*, 2004; Oczkowski *et al.*, 2005; Szunyogh *et al.*, 2005; Kuhl *et al.*, 2007; Kalnay *et al.*, 2006).

In order to use a limited number of ensemble members, many ensemble schemes, including LETKF, use localization (Anderson, 2001; Hamill *et al.*, 2001; Houtekamer and Mitchell, 2001; Whitaker and Hamill, 2002; Ott *et al.*, 2004; Hunt *et al.*, 2007). As in the Local Ensemble Kalman Filter (LEKF Ott *et al.*, 2004), in LETKF the analysis for each grid point is obtained by solving Equations (4.12) – (4.14) exactly using only those observations in a local region centered at that grid point (Ott *et al.*, 2004; Hunt *et al.*, 2007). Here, these regions contain  $2l + 1$  gridpoints in the two

horizontal directions and  $2v + 1$  gridpoints in the vertical direction, for some choice of  $l$  and  $v$ . This type of localization technique (Ott *et al.*, 2004) allows fast parallel implementation (Szunyogh *et al.*, 2005, 2007). Because satellite observations depend on deep layers of the atmosphere, they cannot be selected for assimilation in the local regions this way. Based on Fertig *et al.* (2007a), we use a cutoff based localization for the satellite observations. Specifically, we use a satellite observation in the analysis at a grid point if its weighting function is above a cutoff value,  $\chi$ , somewhere within the local region around that grid point.

Although the LETKF analysis of model variables utilizes localization, we assume that the bias parameters are global constants that depend on the instrument, channel, and ensemble member. We take the following steps to obtain a global analysis for these parameters while still obtaining a local analysis for the model states. For each grid point, we obtain a local augmented state vector by appending those bias parameters corresponding to biased observations that are within the local region to the model state vector. We then update the local augmented state vector for that grid point as described above. Finally, we replace the estimated bias parameter at every grid point by a weighted global average of the analysis estimate for that bias parameter. The weighted average that we use is a surface area average weighting the grid points by the inverse of the corresponding analysis bias variance. That is, the global analysis for the  $i^{\text{th}}$  ensemble member of a component  $\beta$  of the bias state vector is given by

$$\beta^{a(i)} = \frac{\sum_j \frac{\cos(\phi_j)\beta_j^{a(i)}}{\sigma_j^2}}{\sum_j \frac{\cos(\phi_j)}{\sigma_j^2}}, \quad (4.17)$$

where  $j$  indexes all the model grid points,  $\phi_j$  is the latitude at grid point  $j$ , and  $\beta_j^{a(i)}$  is the analysis for the  $i^{\text{th}}$  ensemble member at grid point  $j$ . The variance of the analysis bias ensemble at grid point  $j$  is represented by  $\sigma_j^2$  and is given by

$$\sigma_j^2 = (k - 1)^{-1} \sum_{l=1}^k \left( \beta_j^{a(l)} - \beta_j^a \right)^2, \quad (4.18)$$

where  $\beta_j^a$  is the ensemble mean of the bias analysis at grid point  $j$ .<sup>1</sup> If instead the parameter varies zonally, as for scan-angle bias, this average could be taken over all model grid points at a single latitude or range of latitudes. Because this averaging is employed, the dimensionality of the augmented system is not greatly increased by including the bias parameters. Therefore, we do not require many additional ensemble members to estimate the observation biases, in contrast to Baek *et al.* (2006) where a pointwise estimate of the model bias term required a relatively large ensemble.

## 4.3 Experimental Design

### 4.3.1 SPEEDY Model

The numerical experiments in this paper are performed using the Simplified Parameterizations, primitive-Equation Dynamics (SPEEDY) model (Molteni, 2003). Because the SPEEDY model evolves the primitive equations, it reflects the dynamics of a full general circulation model. However, its computational cost is

---

<sup>1</sup>In the numerical experiments in this paper we do not observe a significant difference in the analysis using simpler means of weighting the average (e.g., by the number of observations in each local region, results not shown).

significantly lower than that of operational models and it has been widely used for climate simulations (Molteni, 2003; Miyoshi, 2005). This computational efficiency is obtained by performing triangular truncation at wavenumber 30 (T30) and using simplified parameterization for several physical processes, including convection, large-scale condensation, cloud physics, radiation, surface fluxes, and vertical diffusion. The boundary conditions at the bottom of the atmosphere for variables including the sea-surface temperature and the land-surface temperature are obtained from the 1981-1990 ECMWF reanalysis (Molteni, 2003).

The SPEEDY model obtains its forecasts by solving the primitive equations for vorticity,  $\zeta$ , divergence,  $\nabla$ , absolute temperature  $T$ , specific humidity,  $q$ , and the logarithm of the surface pressure,  $\log(p_s)$ . The resulting values are converted to output variables of zonal wind,  $u$ , meridional wind,  $v$ , absolute temperature,  $T$ , specific humidity,  $q$ , and surface pressure,  $p_s$ , at 96 zonal grid points, 48 meridional grid points, and seven sigma levels (0.950, 0.835, 0.685, 0.510, 0.340, 0.200, 0.080) or seven pressure levels (925, 850, 700, 500, 300, 200, 100 hPa). Miyoshi (2005) modified the SPEEDY model code for the purpose of data assimilation, using values of  $u$ ,  $v$ ,  $T$ ,  $q$ , and  $p_s$  at seven sigma levels as initial conditions for each six hour forecast of the SPEEDY model.

### 4.3.2 Community Radiative Transfer Model

In this paper, we use the Prototype Community Radiative Transfer Model (pCRTM, Han *et al.*, 2005) to generate the simulated satellite observations and for

the corresponding observation operator. This model was developed by the Joint Center for Satellite Data Assimilation (JCSDA) to provide the community with a unified code to simulate satellite radiances. As part of this mission, the framework of the pCRTM and its more advanced counterpart, the CRTM, are both general enough to be easily adapted to simulate radiances from most satellite instruments and from most atmospheric model states. Furthermore, the more advanced CRTM is currently used by the operational NCEP Global Forecast System (GFS) and the Gridpoint Statistical Interpolation System (GSI) assimilation schemes. In this paper, we use the pCRTM to simulate satellite radiance to take advantage of its relative simplicity.

The pCRTM determines the radiance and brightness temperature from an input profile of pressure, temperature, water vapor, and ozone. Here, the profile values for pressure, temperature, and water vapor are obtained from the SPEEDY model. For simplicity, in this study we assume that the ozone profile has a constant value of 0.06720 ppmv. Furthermore, the numerical experiments in this paper are not concerned with scan angle bias since we assume that the satellite observes the atmosphere directly below the location of the satellite.

The pCRTM determines the radiance by solving a discrete form of Equation (4.3). This equation is given by

$$h(\mathbf{x}) = R_{\text{surface}} + \sum_{i=1}^{\text{top}} B(T_i) \Delta\tau_i, \quad (4.19)$$

where  $R_{\text{surface}}$  is the contribution to the radiance from the surface and  $T_i$  is the temperature at model level  $i$ . The values of  $\tau$  at each level of the profile are derived

from information about the spectral and transmittance coefficients for the desired satellite observation.  $\Delta\tau_i$  is the difference between the  $\tau$  values at levels  $i$  and  $i + 1$  in the profile, for all levels below the top of the profile.  $\Delta\tau$  is determined at the top of the profile by enforcing  $\sum_{i=1}^{\text{top}} \Delta\tau_i = 1$ . The pCRTM finds the brightness temperature by seeking the temperature that would provide a black body radiation of the satellite radiance.

In the discrete form of the radiative transfer model (Equation (4.19)),  $\Delta\tau_i$  indicates the influence of the radiance from the  $i^{\text{th}}$  layer of the profile. Therefore, we use this value of  $\Delta\tau_i$  to select observations using the cutoff based selection described in Section 4.2.2. In this study, we find that selecting observations with  $\Delta\tau_i$  of at least

$$\chi = 0.4 \times \left( \max_{i=1, \dots, \text{top}} \Delta\tau_i \right) \quad (4.20)$$

somewhere within the local region provides the smallest analysis rms error. Accordingly, we use this value in all the numerical experiments in this paper.

### 4.3.3 Truth and Observations

We perform the numerical experiments in this paper for the perfect model scenario with the SPEEDY model. Following Miyoshi (2005), the “true” state of the atmosphere is obtained by running the SPEEDY model for a year from its default rest state on January 01, 1981. We then obtain the true atmospheric trajectory by integrating the SPEEDY model from the resulting atmospheric state at January 01, 1982 for 240 timesteps (roughly two months). The quality of the analyses is

assessed by comparing the resulting analysis to this true state. Specifically, we compute the global root mean square (RMS) error between the true and analysis states in pressure coordinates. The average RMS error is the average of these errors over the last month of the assimilation cycle.

We create two sets of simulated observations from the known “true” state of the atmosphere: a set of rawinsonde observations and a set of radiance observations. The simulated rawinsonde observations are created by adding zero mean, Gaussian noise to the true state of the atmosphere at those points of the SPEEDY model that are closest to the real rawinsonde observations at the 12Z synoptic time (Miyoshi, 2005). We use a standard deviation of 1 m/s for  $u$  and  $v$ , 1 K for  $T$ , 0.0001 kg/kg for  $q$ , and 1 hPa for  $p_s$ .

Similarly, we simulate 15 long-wave temperature radiance channels from the AIRS 281 subset (wave numbers: 680.491, 691.391, 694.4, 696.052, 697.99, 699.102, 700.218, 701.618, 703.87, 706.991, 709.566, 712.739, 715.941, 723.029, 724.824  $\text{cm}^{-1}$ ) at each horizontal model grid point. We note that these channels represent a third of the long-wave temperature soundings from the M-11 hardware module in the 324 subset of AIRS channels for Numerical Weather Prediction. Figure 4.1 shows the weighting function (i.e., the values of  $\Delta\tau$  for each model layer) for representative channels.

These radiance observations are generated by applying the pCRTM to obtain brightness temperature observations from the true atmospheric state. A zero mean, Gaussian noise with standard deviation 0.5 K is added to the resulting brightness temperature observation. The bias in these observations is simulated by modifying

$\kappa(z)$  in the pCRTM as suggested by Watts and McNally (2004) and described in Section 4.2. Because the bias is presumed unknown by the analysis scheme, the unmodified pCRTM is used for the observation operator during the data assimilation.

However, for our simulated radiances the true bias in the observations can be calculated by applying both the modified pCRTM and unmodified pCRTM to the true state. The difference between these two values of the brightness temperature represents the true satellite bias. The extent to which the estimated bias parameters reduce the known bias provides an additional means to verify the data assimilation scheme. Furthermore, the correlation between the true bias and the predictors indicates which predictors can be expected to make the largest correction of the bias and, therefore, provides an additional measure to assess the performance of the bias correction scheme.

#### 4.3.4 LETKF and Bias Correction

We adapt the LETKF code developed originally by Miyoshi (2005) and Harlim and Hunt (2007) for the assimilation of rawinsonde observations to assimilate the rawinsonde and biased radiance observations on the SPEEDY model. In all these numerical experiments, we use 40 ensemble members. We find that local regions one level deep ( $v = 0$ ) with a diameter of 7 horizontal grid points ( $l = 3$ ) provide the best analysis for rawinsonde observations alone. We find that smaller local regions of 3 grid points in diameter ( $l = 1$ ) and one level deep provide the best analysis when we assimilate both the rawinsonde and the simulated satellite

observations. To avoid having the spread in the ensemble members collapse, the numerical experiments perform multiplicative variance inflation on the background ensemble members (both for the dynamical and the bias variables (e.g., Anderson, 2001)). Furthermore, the inflation serves as a forgetting factor so that the bias parameters can account for slowly evolving biases even though persistence is used for their forecast model (Hunt *et al.*, 2007). As will be further described in the next section (Section 4.4), the dynamical variables are always inflated by 6%, while the bias variables are inflated by either 6% or 12%.

We verify the results of the LETKF analysis by computing the root mean square (rms) distance between the analysis and truth (rms error). Similarly, we verify the bias correction parameters by applying the ensemble mean of the estimated bias parameters to the true values of the predictors to correct the set of biased satellite observations with no noise. We refer to the rms error of these corrected satellite observations relative to the unbiased satellite observations as the observation rms error. The time averaged results are shown for the last month of the assimilation experiments (February, 1982) to exclude the spin-up period.

In this study, we subtract the mean over time and space from each of the non-constant predictors before using them for the online bias correction. Thus, the bias estimated from the constant predictor ( $p_1$ ) describes the overall mean bias and the bias estimated from the other predictors represents variations from that mean. We create the initial bias ensemble by randomly sampling from a Gaussian distribution with mean zero and a standard deviation based on the desired initial spread of the ensemble. Over the two month period, we found that the analysis

and estimated bias parameters can be sensitive to the initial spread in the bias parameters. Specifically, we found that the bias parameters took longer to converge when their initial spread was large. The convergence was slower for experiments using three or more predictors (results not shown). However, the spread must start out large enough for the ensemble to capture the bias uncertainty. We therefore tune the initial spread in the bias parameters to minimize the analysis errors and observation biases over the entire two month period. Based on this tuning, we choose the initial standard deviation of the bias parameters to be 1 K for  $\beta_1$ , standard deviation 0.02 K/m for  $\beta_2$  and  $\beta_3$ , and standard deviation 0.004 K/K for  $\beta_4$  and  $\beta_5$ . We expect that if reasonable initial spreads are used, the impact of this choice will decrease to become negligible for longer analysis times.

#### 4.4 Results

To verify the significance of the satellite observation bias created in the form of Watts and McNally (2004), we plot the true bias (as defined in Section 3.3) of the simulated observations at the first analysis time (06Z January 01, 1982) in Figure 4.2 for the channel corresponding to  $696.052 \text{ cm}^{-1}$  (panel (a)) and the channel corresponding to  $723.029 \text{ cm}^{-1}$  (panel (b)). We observe that the bias is significant and location dependent. In particular, the bias has the smallest magnitude near the poles and its magnitude increases near the tropics to almost 1 K. Though Figure 4.2 shows the bias for a particular time and particular channels, it qualitatively represents the typical true bias at all times and for all simulated channels.

The true satellite bias can be used to infer which predictors we can expect to correct for the bias most efficiently. Though the correlations between the predictors and the true bias cannot be computed in reality, they enable us to further evaluate the performance of the bias correction scheme. We plot the correlation between each of the predictors described in Section 4.2 and the true bias of the simulated satellite observations in Figure 4.3. We observe that the true bias is most strongly correlated to predictors  $p_2$  and  $p_3$  for all satellite channels at a wave number below  $712.739 \text{ cm}^{-1}$ . On the other hand, predictors  $p_4$  and  $p_5$  are more strongly correlated to the bias for wave numbers above  $715.941 \text{ cm}^{-1}$ . However, none of the predictors are well correlated to the bias for channels at wave numbers  $700.218$  and  $701.618 \text{ cm}^{-1}$ . This poor correlation is observed because the bias of these channels is substantially smaller than the others (a fact which we will explore later in this section). Overall, predictors  $p_3$  and  $p_4$  show the strongest correlations to the true bias. We also find these predictors to be more useful than  $p_2$  and  $p_5$  in reducing the analysis error. For completeness, we will show results using all five predictors at the end of this section. Here, we discuss results using only predictors  $p_1$  (constant),  $p_3$  (850 - 300 hPa thickness), and  $p_4$  (skin temperature).

In Figure 4.4, we plot the global average of the 500 hPa temperature analysis rms error over time when inflating all background variables by 6%. For reference, we plot the analysis rms error when assimilating rawinsonde observations alone (cyan), rawinsonde and biased satellite observations using the unmodified pCRTM (gray), and rawinsonde and biased satellite observations using the modified pCRTM (black). From these reference runs, it is apparent that the biased satellite observations have

a significant negative impact on the analysis. Nonetheless, were these observations unbiased they could dramatically improve the analysis over that obtained from assimilating the rawinsonde observations alone.

The analysis rms error for the bias correction is shown in Figure 4.4 for the following combination of predictors:  $p_1$  (green);  $p_1$  and  $p_3$  (blue);  $p_1$  and  $p_4$  (red); and  $p_1$ ,  $p_3$ , and  $p_4$  (purple). All these bias correction experiments improve the analysis rms error significantly. The skin temperature,  $p_4$ , has the most significant impact on the 500 hPa analysis. When using  $p_1$ ,  $p_3$ , and  $p_4$ , the 500 hPa temperature analysis rms error more slowly reduces to the value obtained from using  $p_1$  and  $p_4$  alone. The results are similar for the surface pressure (results not shown). The bias correction clearly has an impact on the 500 hPa zonal wind, meridional wind, and humidity analyses (results not shown). However, there appears to be a less significant difference between the different predictors for these variables. These results are further summarized for all model levels in Table 4.1. This table shows that predictors  $p_1$  and  $p_4$  improve the analysis rms error in most lower levels, while  $p_1$  and  $p_3$  improve the analysis most in the higher model levels.

In order for the online bias correction to continually update the bias estimate, the spread (standard deviation) of the ensemble of bias parameters must remain significantly larger than zero. Therefore, Figure 4.5 plots the background spread for each of the predictors for a representative subset of the channels corrected in the run with  $p_1$ ,  $p_3$ , and  $p_4$  (panels (a), (b), and (c), respectively) over time. The spread in the bias parameters for channel 723.029  $\text{cm}^{-1}$  remains nearly constant after the spin-up time. However, for the majority of the other channels the spread of the bias

parameters decreases dramatically as the analysis cycle proceeds. A similar decrease in the bias spread is observed for bias correction experiments with all combinations of the predictors (results not shown). For channels like  $723.029\text{ cm}^{-1}$ , where the bias remains nearly constant, we observe that the weighting function is above the cutoff value for several model levels (results not shown). Therefore, the analysis of the bias parameters corresponding to these channels combines information from more local regions, keeping the spread relatively large.

Multiplicative variance inflation often prevents the background ensemble from collapsing and serves as a forgetting factor of past observations and dynamics. However, increasing the variance inflation of all background variables degrades the quality of the analysis and does not prevent the spread of the bias parameters from collapsing (results not shown). Instead, we compensate for the decreasing spread in the ensemble of bias parameters by inflating these variables by 12% and retaining the 6% variance inflation for the other analysis variables. In this case, the analysis rms error over time is shown at 500 hPa in Figure 4.6 and is similar to that obtained using a 6% variance inflation for all variables (Figure 4.4). Here, the analysis error associated with using  $p_1$ ,  $p_3$ , and  $p_4$  reaches its optimal value more quickly. The similarity between the analysis error when using different inflation factors holds at all model levels (Table 4.1 for 6% inflation of all variables and Table 4.2 for 12% inflation of bias variables). Furthermore, Figure 4.7 reveals that the spread of the bias correction run with  $p_1$ ,  $p_3$ ,  $p_4$  (panels (a), (b), and (c), respectively) plateaus to a more reasonable value for all channels. The results are similar for the other bias correction experiments (results not shown).

We also explore the impact of the bias correction on the global, time averaged temperature analysis RMS error for at all seven SPEEDY model levels in Figure 4.8. The color scheme in this figure is the same as that used in Figure 4.4. In this figure, we see that the  $p_4$  term provides the greatest improvement to the analysis at low levels, below 400 hPa. On the other hand,  $p_3$  provides the greatest improvement to the analysis above 300 hPa. The combination of predictors  $p_1$ ,  $p_3$ , and  $p_4$  provides the analysis with the smallest error at almost all levels. Below 500 hPa, the analysis error almost as low as that obtained from assimilating the satellite observations with the true observation operator. We further summarize the global, time averaged temperature rms error for all the predictors proposed in Section 4.2 in Table 4.2. As we mentioned earlier, we find that using  $p_2$  or  $p_5$  also decreases the analysis rms error, but not as significantly as using  $p_3$  or  $p_4$ . Also, including  $p_2$  and  $p_5$  in the bias correction with  $p_1$ ,  $p_3$ , and  $p_4$  does not improve the analysis rms error.

Ideally, the bias parameters predicted by the data assimilation improve the analysis rms errors because they reduce the bias in the observations. Figure 4.9 plots the global, time averaged observation rms error for each of the choices of predictors for each satellite channel. We observe that  $p_3$  improves the observation bias most for low wave numbers, while  $p_4$  improves the observation bias for higher wave numbers. This result is consistent with the structure of the correlation between the true observation bias and these predictors shown in Figure 4.3. Furthermore, using predictors  $p_1$ ,  $p_3$ , and  $p_4$  together significantly reduces the observation rms for most channels. We observe that the bias correction scheme is unable to correct the already small biases for channels at wave numbers of 700.218 and 701.618  $\text{cm}^{-1}$ .

Closer inspection reveals that the weighting function for these channels is such that they rarely meet the localization criterion necessary for them to be used in the analysis. Therefore, these observations have little impact on the analysis of the atmospheric variables or the bias parameters. Accordingly, the estimated bias parameters do not improve the relatively small biases for these channels. The global, time averaged observation rms error is summarized for selected channels using all choices of predictors in Table 4.3. As for the analysis errors, we see that none of the results using  $p_2$  or  $p_5$  improve significantly compared to the results using  $p_3$  and/or  $p_4$ . We also see that for each of the channels the predictors that are best correlated with the true bias (see Figure 4.3) yield the greatest improvement in the observation rms error.

We observe in Figure 4.9 that for the lower wave number channels, using certain combinations of the predictors increases the rms error of the observations above their value with no bias correction. However, the bias correction still significantly improves the analysis rms error throughout the atmosphere (Figures 4.6 and 4.8). For example, while using  $p_1$  (constant) improves the analysis at all but the highest model level, this predictor only improves the observation rms error for channels with wave numbers above  $701.618 \text{ cm}^{-1}$  (Figure 4.9). To explore the impact of the increased observation rms error further, we tried using  $p_1$  to correct only the bias for channels above  $701.618 \text{ cm}^{-1}$ . We plot the resulting observation and analysis rms errors in panels (a) and (b) of Figure 4.10, respectively. The observation rms error from using  $p_1$  to correct the bias only for channels above wave number  $701.618 \text{ cm}^{-1}$  (dotted) mirrors that obtained from using  $p_1$  for all channels (dashed) above  $701.618$

$\text{cm}^{-1}$  and the true observation bias (solid in panel (a)) for the rest of the channels. In spite of this apparent improvement in the overall observation rms error, using  $p_1$  to correct the bias only for channels above  $701.618 \text{ cm}^{-1}$  degrades the temperature analysis rms error for the six lower model levels.

## 4.5 Conclusions

Our results show that ensemble data assimilation schemes, such as LETKF, can estimate several observation bias parameters as part of the assimilation procedure. If correlated to the true observation bias, the estimated bias parameters significantly reduce the analysis error and observation bias. In some cases, the observation biases may be increased in order to reduce the analysis error. Furthermore, making the inflation parameter larger for the bias parameters than for the model variables prevents the spread in the bias correction from collapsing. In this way, the online bias estimation can respond to slow changes in the bias parameters.

Computing a global average of the local estimates for the bias parameters ensures accuracy, efficiency, and flexibility. For one, obtaining a global estimate of the bias parameters ensures that they do not capture the random errors from a small set of observations in place of the bias. Furthermore, global averaging effectively reduces the dimensionality of the bias problem so that no additional ensemble members are required to estimate the bias parameters. Finally, averaging provides flexibility because it could be applied to estimate the bias parameters over any sufficiently large region. In this way, the bias correction scheme could obtain

bias estimates over latitude strips as easily as it obtains them globally in this paper.

Though we applied the observation bias correction scheme to a particular formulation of the ensemble Kalman filter in this paper, we expect the results would hold for any ensemble scheme. Furthermore, they need not be limited to estimating bias parameters which combine with predictors to determine the bias. For example, this correction scheme could as easily be applied to estimate the bias parameters of Watts and McNally (2004). As for other online observation bias schemes, bias correction in ensemble Kalman filter schemes require that there be only a limited number of bias parameters to avoid overfitting. In this way, there is great flexibility to estimating the observation biases with an ensemble Kalman filter.

## 4.6 Figures

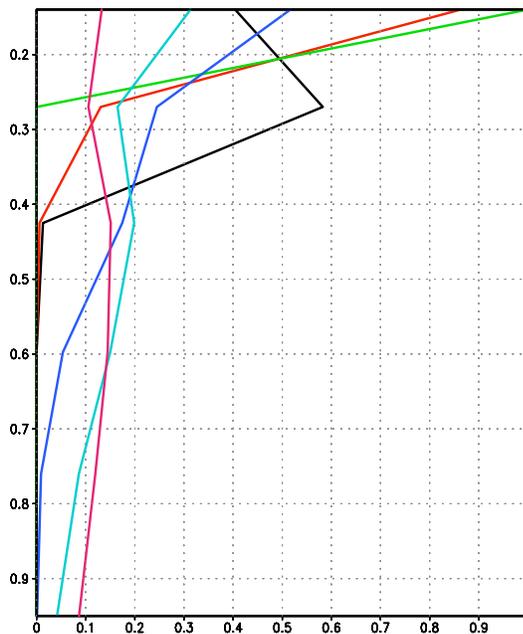


Figure 4.1: The weighting function (profile of  $\Delta\tau$  at 180 W and 35 N) for representative channels:  $680.491 \text{ cm}^{-1}$  (black),  $697.99 \text{ cm}^{-1}$  (red, similar shape to channels  $691.391 \text{ cm}^{-1}$  -  $699.102 \text{ cm}^{-1}$ ),  $700.218 \text{ cm}^{-1}$  (green, similar to channel  $701.618 \text{ cm}^{-1}$ ),  $706.991 \text{ cm}^{-1}$  (dark blue, similar to channels  $703.87 \text{ cm}^{-1}$  -  $709.991 \text{ cm}^{-1}$ ),  $712.739 \text{ cm}^{-1}$  (light blue), and  $723.029 \text{ cm}^{-1}$  (magenta, similar to channels  $715.941 \text{ cm}^{-1}$  -  $724.824 \text{ cm}^{-1}$ ).



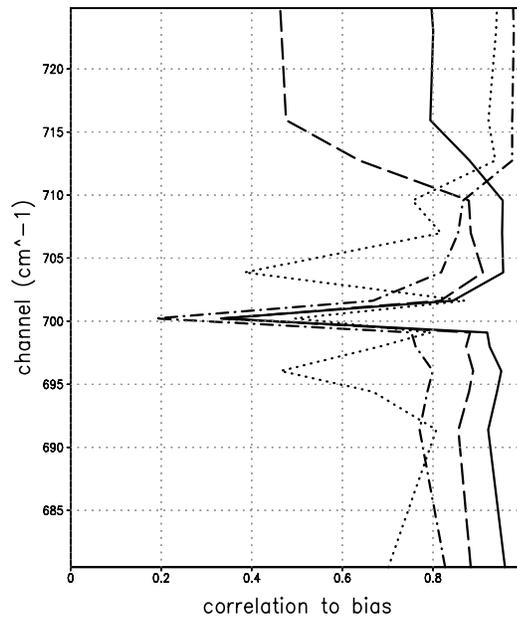


Figure 4.3: The correlation between the true bias and each of the predictors for each satellite channel. ( $p_2$  is dashed;  $p_3$  is solid;  $p_4$  is dash-dotted; and  $p_5$  is dotted.)

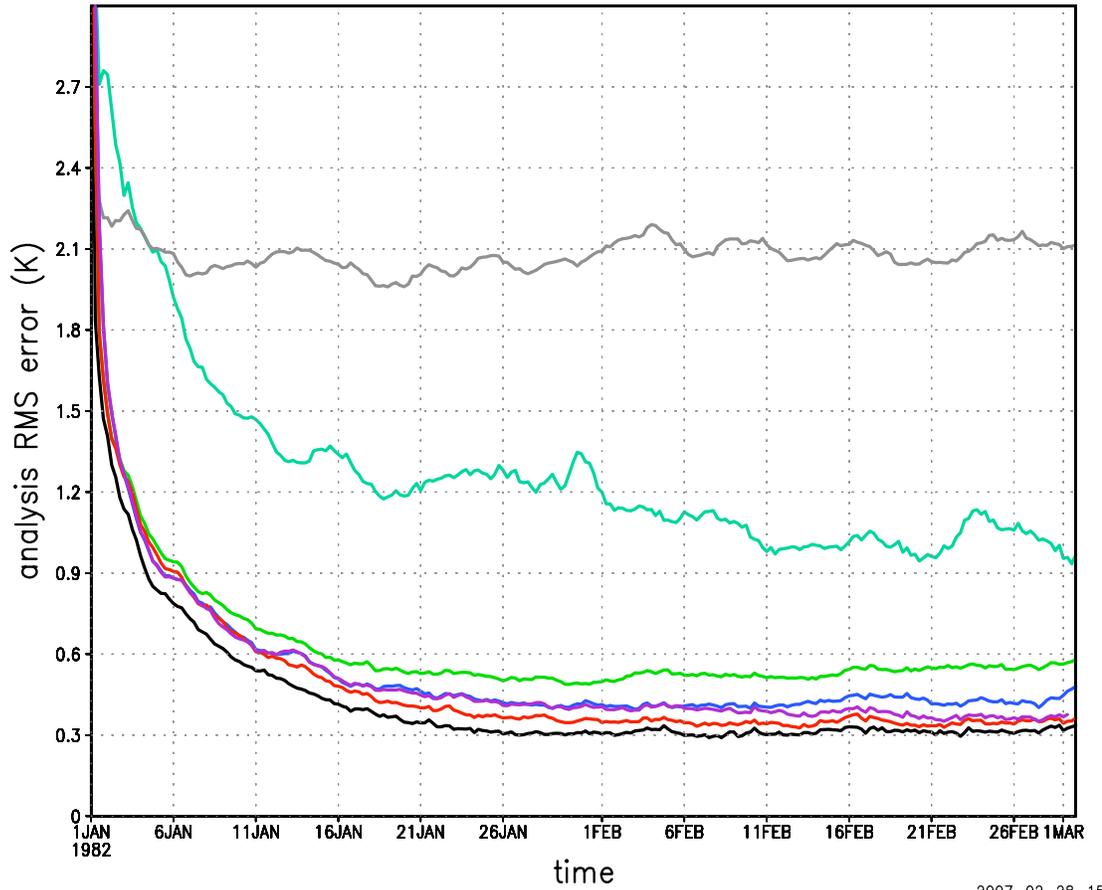
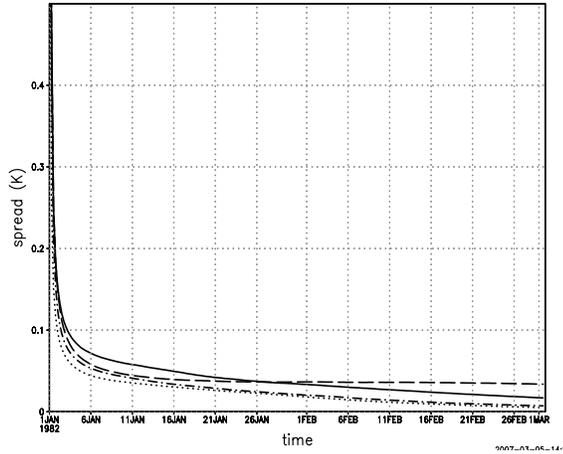
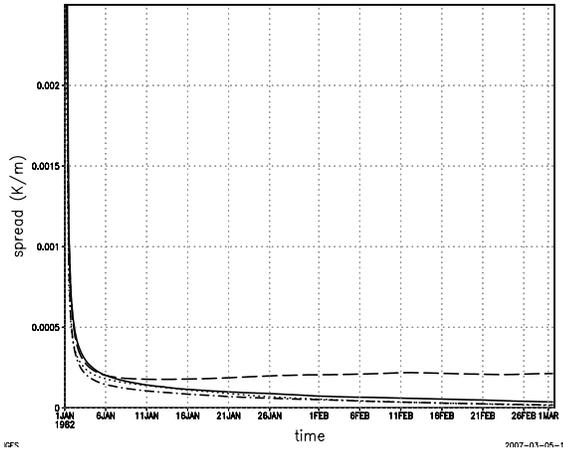


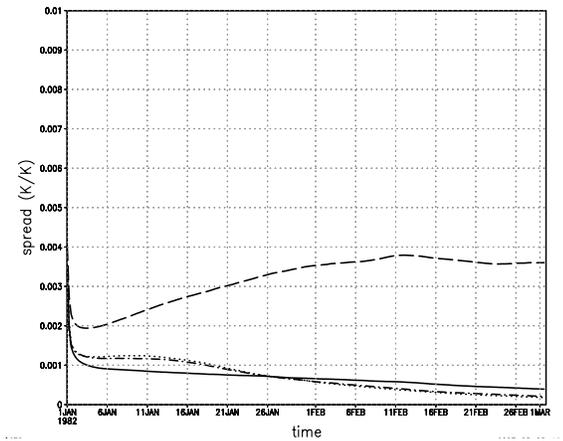
Figure 4.4: The 500 hPa temperature analysis rms error for each of the analysis times. The following color scheme is used in this figure: the results for rawinsonde observations alone are cyan; for rawinsonde and biased satellite observations using the unmodified pCRTM are gray; for rawinsonde and biased satellite observations using the modified pCRTM are black; for predictor  $p_1$  are green; for predictors  $p_1$  and  $p_3$  are blue; and for predictors  $p_1$  and  $p_4$  are red; and  $p_1$ ,  $p_3$ , and  $p_4$  are purple. Here, both the analysis and bias variables are inflated by 6%.



(a) Spread of  $p_1$



(b) Spread of  $p_3$



(c) Spread of  $p_4$

Figure 4.5: The spread in the background bias parameters for the run with predictors  $p_1$ ,  $p_3$ , and  $p_4$ . The spread of these predictors is in panels (a) to (c), respectively. Each line represents a different satellite channel, where the solid line is for channel  $680.491 \text{ cm}^{-1}$ , the dotted line is for channel  $696.052 \text{ cm}^{-1}$ , the dash-dotted line is for channel  $709.566 \text{ cm}^{-1}$ , and the dashed line is for channel  $723.029 \text{ cm}^{-1}$ .

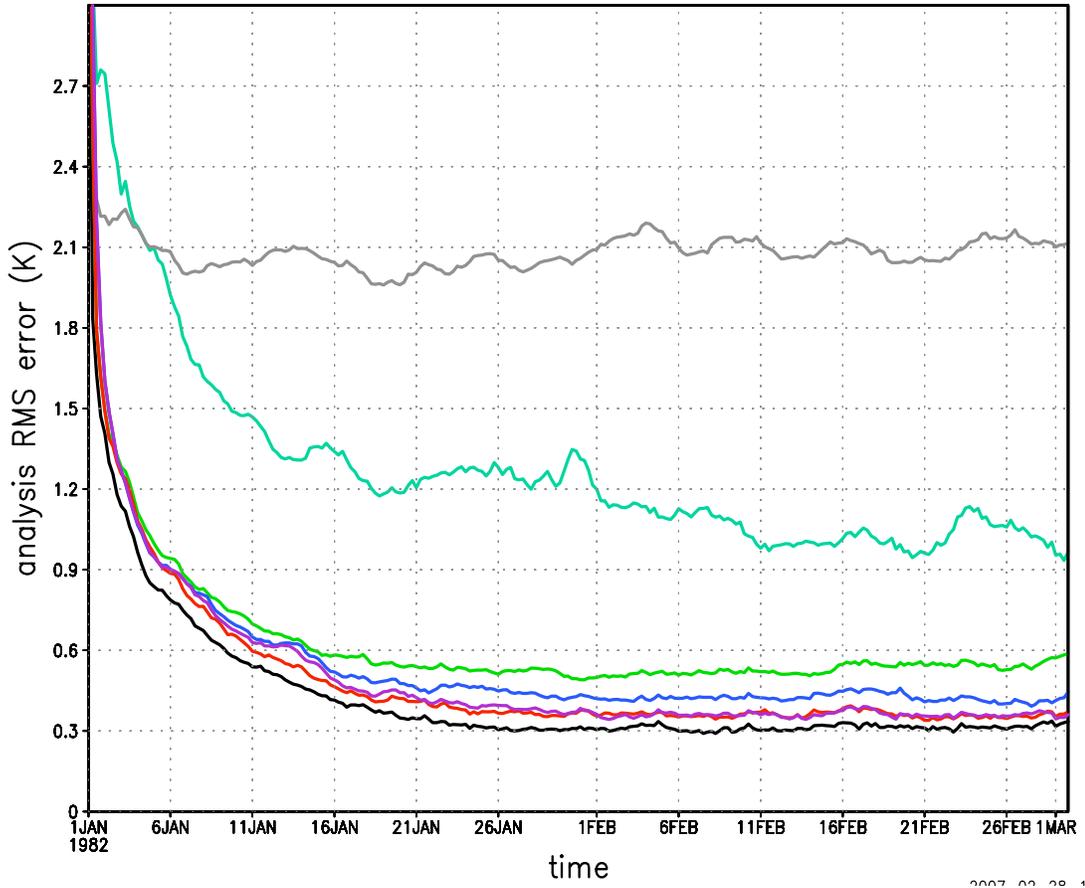
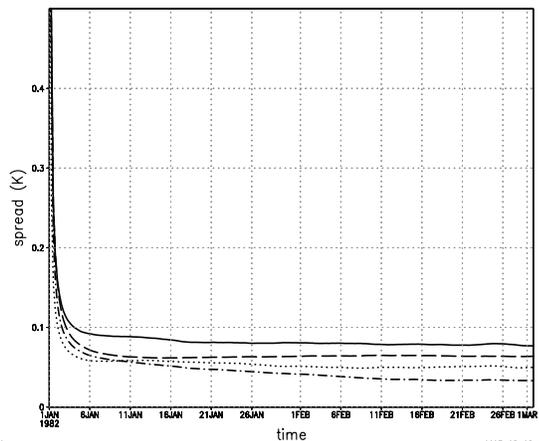
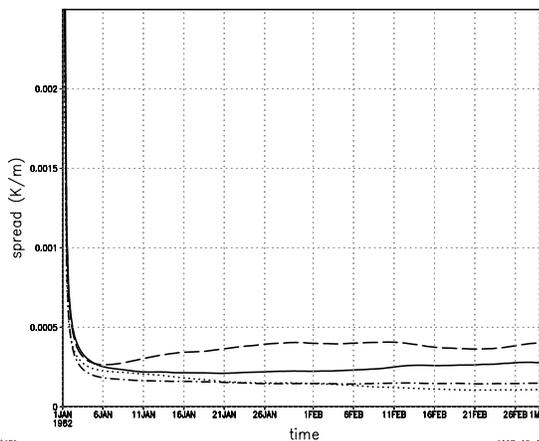


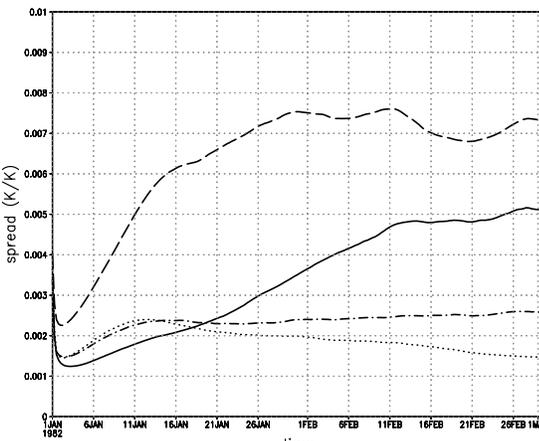
Figure 4.6: As for Figure 4.4. Here, the analysis variables are inflated by 6% while the bias parameters are inflated by 12%.



(a) Spread of  $p_1$



(b) Spread of  $p_3$



(c) Spread of  $p_4$

Figure 4.7: As for Figure 4.5. Here, the analysis variables are inflated by 6% while the bias parameters are inflated by 12%.

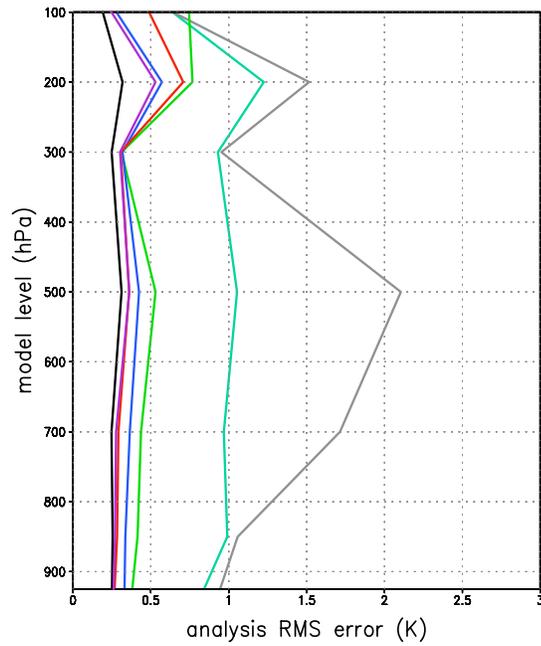


Figure 4.8: The global, February average of the temperature analysis rms error at each model level. The color scheme for this plot is the same as that used in Figure 4.4. Here, the analysis variables are inflated by 6% while the bias parameters are inflated by 12%.

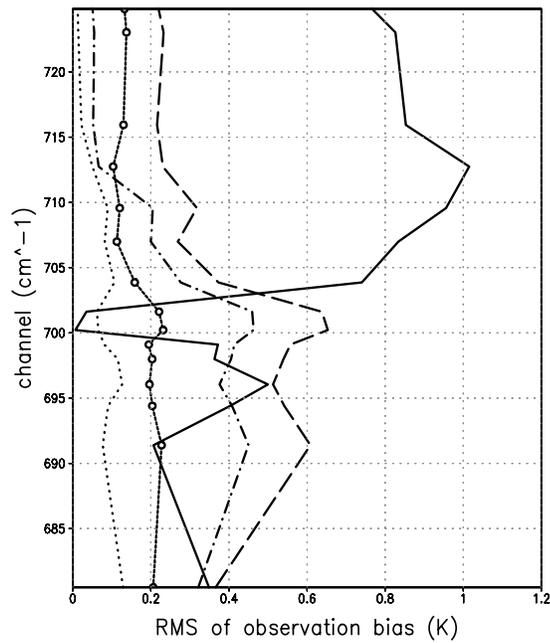
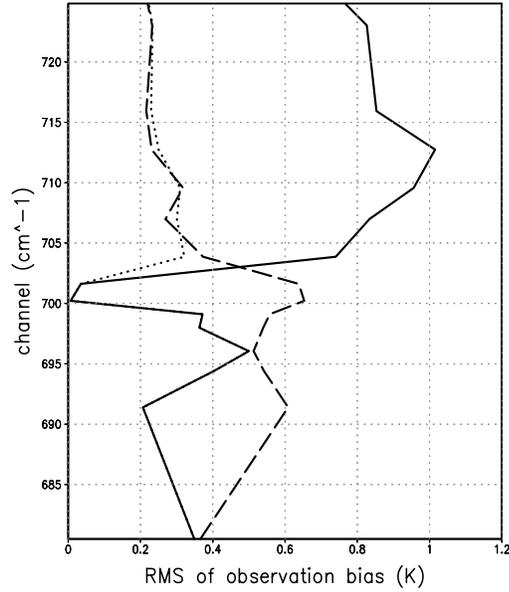
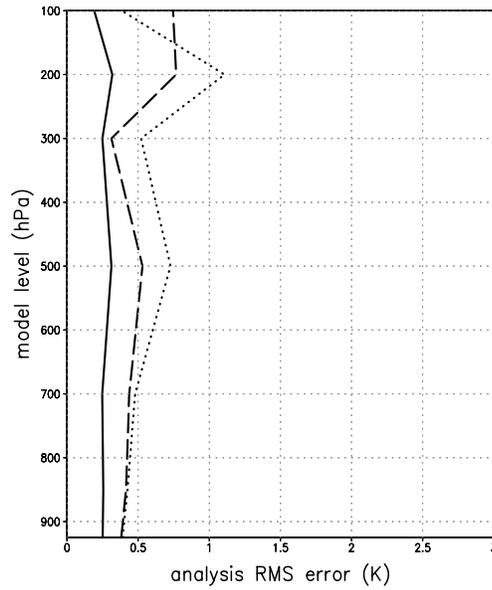


Figure 4.9: The global, February average of the observation rms error for each satellite channel. The observation rms error obtained when using no predictors is solid; for  $p_1$  is dashed; for  $p_1$  and  $p_3$  is thickly dotted with open circles; for  $p_1$  and  $p_4$  is dash-dotted; and for  $p_1$ ,  $p_3$ , and  $p_4$  is lightly dotted.



(a) Observation rms error



(b) Temperature analysis rms error

Figure 4.10: The results from using  $p_1$  to correct only channels above  $701.618 \text{ cm}^{-1}$  (dotted) as compared to using  $p_1$  at all levels (dashed). In panel (a) showing the RMS error of the observations, the true bias is plotted (solid) for reference. In panel (b) showing the temperature analysis, we plot the results obtained from using the true modified observation operator (solid).

<b>Experiment</b>	<b>Model Level</b>						
	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>	<b>6</b>	<b>7</b>
modified pCRTM (unbiased)	0.25	0.25	0.25	0.31	0.25	0.32	0.19
$p_1$	0.40	0.44	0.46	0.53	0.30	0.79	0.79
$p_1, p_3$	0.33	0.34	0.37	0.42	0.28	0.46	0.25
$p_1, p_4$	0.27	0.28	0.28	0.35	0.29	0.72	0.58
$p_1, p_3, p_4$	0.29	0.29	0.31	0.38	0.30	0.50	0.28

Table 4.1: The global, February average of the temperature analysis RMS error (K) at each model level for selected combinations of predictors. Here, 6% variance inflation is applied to the analysis variables and 6% variance inflation is applied to the bias parameters.

<b>Experiment</b>	<b>Model Level</b>						
	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>	<b>6</b>	<b>7</b>
modified pCRTM (unbiased)	0.25	0.25	0.25	0.31	0.25	0.32	0.19
$p_1$	0.38	0.41	0.44	0.53	0.31	0.77	0.75
$p_1, p_2$	0.37	0.40	0.44	0.51	0.34	0.60	0.33
$p_1, p_3$	0.33	0.34	0.37	0.42	0.32	0.57	0.28
$p_1, p_4$	0.27	0.28	0.29	0.36	0.31	0.71	0.49
$p_1, p_5$	0.29	0.30	0.32	0.38	0.32	0.77	0.60
$p_1, p_3, p_4$	0.26	0.27	0.28	0.36	0.30	0.53	0.25
$p_1, p_2, p_3, p_4, p_5$	0.26	0.27	0.28	0.39	0.33	0.56	0.34

Table 4.2: As for Table 4.2. Here, 6% variance inflation is applied to the analysis variables and 12% variance inflation is applied to the bias parameters.

## 4.7 Tables

Experiment	Channel ( $\text{cm}^{-1}$ )			
	680.491	696.052	709.566	723.029
unmodified pCRTM (biased)	0.35	0.50	0.96	0.83
$p_1$	0.37	0.51	0.32	0.23
$p_1, p_2$	0.20	0.21	0.16	0.21
$p_1, p_3$	0.21	0.20	0.12	0.14
$p_1, p_4$	0.32	0.38	0.20	0.05
$p_1, p_5$	0.37	0.47	0.27	0.09
$p_1, p_3, p_4$	0.13	0.13	0.09	0.01
$p_1, p_2, p_3, p_4, p_5$	0.19	0.12	0.14	0.02

Table 4.3: The global, February average of the observation RMS error (K) for representative channels when applying the bias correction parameters estimated using 6% variance inflation for the analysis variables and 12% for the bias parameters.

## Bibliography

- Anderson, J. 2001. An ensemble adjustment Kalman filter for data assimilation. *Monthly Weather Review*, **129**, 2884–2903.
- Anderson, J. 2003. A local least squares framework for ensemble filtering. *Monthly Weather Review*, **131**, 634–642.
- Baek, S.-J., Hunt, B., Kalnay, E., Ott, E., and Szunyogh, I. 2006. Local ensemble Kalman filtering in the presence of model bias. *Tellus*, **58A**, 293–306.
- Bishop, C., Etherton, B., and Majumdar, S. 2001. Adaptive sampling with the ensemble transform Kalman filter. Part I: Theoretical aspects. *Monthly Weather Review*, **129**, 420–436.
- Burgers, G., van Leeuwen, P., and Evensen, G. 1998. Analysis scheme in the ensemble Kalman filter. *Monthly Weather Review*, **126**, 1719–1724.
- Cameron, J. 2003. The effectiveness of the AIRS bias correction for various air-mass predictor combinations. Technical report, Met Office NWP.
- Courtier, P., Thépaut, J.-N., and Hollingsworth, A. 1994. A strategy for operational implementation of 4D-VAR, using an incremental approach. *Quarterly Journal of Royal Meteorological Society*, **120**, 1367–1387.
- Dee, D. 2005. Bias and data assimilation. *Quarterly Journal of the Royal Meteorological Society*, **131**, 3323–3343.
- Derber, J. and Wu, W.-S. 1998. The use of TOVS cloud-cleared radiances in the NCEP SSI analysis system. *Monthly Weather Review*, **126**, 2287–2299.
- Evensen, G. 1994. Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo Methods to forecast error statistics. *Journal of Geophysical Research*, **99**, 10143–10162.
- Evensen, G. and van Leeuwen, P. 2000. An ensemble Kalman smoother for nonlinear dynamics. *Monthly Weather Review*, **128**, 1852–1867.
- Eyre, J. 1992. A bias correction scheme for simulated TOVS brightness temperatures. Technical Memorandum 186, ECMWF, Reading, UK.
- Fertig, E., Hunt, B., Ott, E., and Szunyogh, I. 2007a. Assimilating nonlocal observations with a local ensemble Kalman filter. *Tellus Ser. A*, (Submitted).
- Fertig, E., Harlim, J., and Hunt, B. 2007b. A comparative study of 4D-VAR and 4D ensemble Kalman filter: Perfect model simulations with Lorenz-96. *Tellus A*, **59**, 96–101.
- Friedland, B. 1969. Treatment of bias in recursive filtering. *IEEE Trans. Autom. Contr.*, **14**, 359–367.

- Gaspari, G. and Cohn, S. 1999. Construction of correlation functions in two and three dimensions. *Quarterly Journal of Royal Meteorological Society*, **125**, 723–757.
- Hamill, T., Whitaker, J., and Snyder, C. 2001. Distance-dependent filtering of background covariance estimates in an ensemble Kalman filter. *Monthly Weather Review*, **129**, 2776–2790.
- Han, Y., van Delst, P., Liu, Q., Weng, F., Yan, B., and Derber, J. 2005. *User’s Guide to the JCSDA Community Radiative Transfer Model (Beta Version)*. Joint Center for Satellite Data Assimilation, Camp Springs, Maryland.
- Harlim, J. and Hunt, B. 2007. A non-gaussian ensemble filter for assimilating infrequent noisy observations. *Tellus*, **59A**, 225–237.
- Harris, B. and Kelly, G. 2001. A satellite radiance-bias correction scheme for data assimilation. *Quarterly Journal of Royal Meteorological Society*, **127**, 1453–1468.
- Harris, B., Cameron, J., Collard, A., and Saunders, R. 2004. Effect of air-mass predictor choice on the AIRS bias correction at the Met Office. Technical report, Thirteenth International TOVS Study.
- Houtekamer, P. and Mitchell, H. 1998. Data assimilation using an ensemble Kalman filter technique. *Monthly Weather Review*, **126**, 796–811.
- Houtekamer, P. and Mitchell, H. 2001. A sequential ensemble Kalman filter for atmospheric data assimilation. *Monthly Weather Review*, **129**, 123–137.
- Houtekamer, P. and Mitchell, H. 2005. Ensemble Kalman filtering. *Quarterly Journal of the Royal Meteorological Society*, **131**, 3269–3289.
- Houtekamer, P., Mitchell, H., Pellerin, G., Buehner, M., and M. Charron, e. a. 2005. Atmospheric data assimilation with the ensemble Kalman filter: Results with real observations. *Monthly Weather Review*, **133**, 604–620.
- Hunt, B., Kalnay, E., Kostelich, E., Ott, E., Patil, D., Sauer, T., Szunyogh, I., Yorke, J., and Zimin, A. 2004. Four-dimensional ensemble Kalman filtering. *Tellus*, **56A**, 273–277.
- Hunt, B., Kostelich, E., and Szunyogh, I. 2007. Efficient data assimilation for spatiotemporal chaos: a local ensemble transform Kalman filter. *Physica D.*, (In Press).
- Joiner, J. and da Silva, A. 1998. Efficient methods to assimilate remotely sensed data based on information content. *Quarterly Journal of the Royal Meteorological*, **124**(549), 1669–1694.
- Kalman, R. 1960. A new approach to linear filtering and predictions problems. *Journal of Basic Engineering*, **82 D**, 35–45.

- Kalman, R. and Bucy, R. 1961. New results in linear filtering and prediction theory. *Journal of Basic Engineering*, **83 D**, 95–107.
- Kalnay, E. 2002. *Atmospheric Modeling, Data Assimilation, and Predictability*. Cambridge University Press.
- Kalnay, E., Hunt, B., Ott, E., and Szunyogh, I. 2006. Ensemble forecasting and data assimilation: two problems with the same solution? In T. Palmer and R. Hagedorn, editors, *Predictability of Weather and Climate*. Cambridge University Press.
- Keppenne, C. 1999. Data assimilation into a primitive-equation model with a parallel ensemble Kalman filter. *Monthly Weather Review*, **128**, 1971–1981.
- Keppenne, C., Rienecker, M., Kurowski, N., and Adamec, D. 2005. Ensemble Kalman filter assimilation of temperature and altimeter data with bias correction and application to seasonal prediction. *Nonlinear Processes in Geophysics*, **12**, 491–503.
- Kuhl, D., Szunyogh, I., Kostelich, E. J., Patil, D. J., Gyarmati, G., Oczkowski, M., Hunt, B. R., Kalnay, E., Ott, E., and Yorke, J. A. 2007. Assessing predictability with a local ensemble kalman filter. *Submitted to Journal of the Atmospheric Sciences*, (In Press).
- Lawson, L., Spitz, Y., Hofmann, E., and Long, R. 1995. A data assimilation technique applied to a predator-prey model. *Bulletin of Mathematical Biology*, **57**(4), 593–617.
- Le Dimet, F.-X. and Talagrand, O. 1986. Variational algorithm for analysis and assimilation of meteorological observations: Theoretical aspects. *Tellus*, **38A**, 97–110.
- Liou, K. 2002. *An Introduction to Atmospheric Radiation*. Academic Press, New York, second edition.
- Liu, J., Fertig, E., Kalnay, E., Hunt, B., Kostelich, E., Szunyogh, I., and Todling, R. 2007.
- Lorenc, A. 2003. The potential of the ensemble Kalman filter for NWP—a comparison with 4D-Var. *Quarterly Journal of Royal Meteorological Society*, **129**, 3183–3203.
- Lorenz, E. 1996. Predictability - a problem partly solved. In *Proceedings on predictability, held at ECMWF on 4-8 September 1995*.
- Miyoshi, T. 2005. *Ensemble Kalman filter experiments with a primitive-equation global model*. Ph.D. thesis, University of Maryland.
- Molteni, F. 2003. Atmospheric simulations using a GCM with simplified physics parameterizations. i: Model climatology and variability in multi-decadal experiments. *Clim. Dyn.*, **20**, 175–191.

- Oczkowski, M., Szunyogh, I., and Patil, D. J. 2005. Mechanisms for the development of locally low dimensional atmospheric dynamics. *Journal of the Atmospheric Sciences*, **62**, 1135–1156.
- Ott, E., Hunt, B., Szunyogh, I., Zimin, A., Kostelich, E., Corazza, M., Kalnay, E., and Yorke, J. 2004. A local ensemble Kalman filter for atmospheric data assimilation. *Tellus*, **56A**, 415–428.
- Parrish, D. and Cohn, S. 1985. A Kalman filter for a two-dimensional shallow-water model: Formulation and preconditioning experiments. Office Note 304, National Meteorological Center, Washington, DC.
- Parrish, D. and Derber, J. 1992. The National Meteorological Center's spectral statistical-interpolation analysis system. *Monthly Weather Review*, **120**(8), 1747–1763.
- Patil, D., Hunt, B. R., Kalnay, E., Yorke, J. A., and Ott, E. 2001. Local low dimensionality at atmospheric dynamics. *Physical Review Letters*, **86**(26), 5878–5881.
- Press, W., B.P.Flannery, S.A.Teukolsky, and Vetterling, W. 1992. *Numerical recipes in Fortran*. Cambridge University Press.
- Rabier, F., Thépaut, J.-N., and Courtier, P. 1998. Extended assimilation and forecast experiments with a four-dimensional variational assimilation system. *Quarterly Journal of Royal Meteorological Society*, **124**, 1861–1887.
- Rabier, F., Järvinen, H., Mahfouf, J.-F., and Simmons, A. 2000. The ECMWF operational implementation of four-dimensional variational assimilation: Experimental results with simplified physics. *Quarterly Journal of Royal Meteorological Society*, **126**, 1148–1170.
- Rodgers, C. 2000. *Inverse Methods for Atmospheric Sounding: Theory and Practice*. World Scientific Publishing.
- Szunyogh, I., Kostelich, E., Gyarmati, G., Patil, D., Hunt, B., Kalnay, E., Ott, E., and Yorke, J. 2005. Assessing a local ensemble Kalman filter: Perfect model experiments with the National Centers for Environmental Prediction global model. *Tellus*, **57A**, 528–545.
- Szunyogh, I., Kostelich, E., Gyarmati, G., Kalnay, E., Hunt, B., Ott, E., and Yorke, J. A. 2007. A local ensemble transform Kalman filter data assimilation for the NCEP global model. *Tellus*, (Submitted).
- Thépaut, J.-N., Hoffman, R., and Courtier, P. 1993. Interactions of dynamics and observations in a four-dimensional variational assimilation. *Monthly Weather Review*, **121**, 3393–3414.

- Tippett, M., Anderson, J., Bishop, C., Hamill, T., and Whitaker, J. 2003. Ensemble square-root filters. *Monthly Weather Review*, **131**, 1485–1490.
- Wang, X., Bishop, C., and Julier, S. 2004. Which is better, an ensemble of positive-negative pairs or a centered spherical simplex ensemble? *Monthly Weather Review*, **132**(7), 1590–1605.
- Wang, X., Hamill, T., Whitaker, J., and Bishop, C. 2006. A comparison of hybrid ensemble transform Kalman filter-OI and ensemble square-root filter analysis schemes. *Monthly Weather Review*.
- Watts, P. and McNally, A. 2004. Identification and correction of radiative transfer modelling errors for atmospheric sounders: AIRS and AMSU-A. In *ECMWF Workshop on Assimilation of high spectral resolution sounders in NWP*.
- Whitaker, J. and Hamill, T. 2002. Ensemble data assimilation without perturbed observations. *Monthly Weather Review*, **130**, 1913–1924.
- Whitaker, J., Compo, G., Wei, X., and Hamill, T. 2004. Reanalysis without radiosondes using ensemble data assimilation. *Monthly Weather Review*, **132**, 1190–1200.
- Whitaker, J., Hamill, T., and Wei, X. 2007. Ensemble data assimilation with the ncep global forecast system. *Monthly Weather Review*, (Submitted).
- Zupanski, D. and Zupanski, M. 2006. Model error estimation employing an ensemble data assimilation approach. *Monthly Weather Review*, **134**, 1337–1354.
- Zupanski, M. 2005. Maximum likelihood ensemble filter: Theoretical aspects. *Monthly Weather Review*, **133**, 1710–1726.