University of Reading

Nonlinear bias correction for numerical weather prediction models

PhD in Mathematics

Department of Mathematics

Jason Arvid Otkin

March 2020

Declaration

I confirm that this is my own research and that the use of all material from other sources has been properly and fully acknowledged.

- Jason Arvid Otkin

Publications

The results presented in Chapters 3, 4, and 5 of this thesis have been, or will be, included in the following peer-reviewed publications:

Otkin, J. A., R. Potthast, and A. Lawless, 2018: Nonlinear bias correction for satellite data assimilation using Taylor series polynomials. *Mon. Wea. Rev.*, **146**, 263-285.

Otkin, J. A., and R. Potthast, 2019: Assimilation of all-sky SEVIRI infrared brightness temperatures with nonlinear bias corrections in a regional-scale ensemble data assimilation system. *Mon. Wea. Rev.*, **147**, 4481-4509.

Otkin, J. A., R. W. E. Potthast, and A. S. Lawless 2020: Nonlinear conditional bias estimation for data assimilation. Conditionally accepted for publication in *SIAM Journal on Applied Dynamical Systems*.

All research results presented in these publications were produced by Jason Otkin, with co-authors providing guidance and review.

Abstract

Data assimilation is an inverse problem that seeks to optimally combine information from a set of observations with a first guess analysis to generate the best estimate of the current state of a dynamic system. It is an essential part of numerical weather prediction because the accuracy of a model forecast is closely tied to the accuracy of the initial conditions. Thus, the goal of this thesis is to enhance our ability to assimilate satellite brightness temperatures through development of bias correction (BC) methods to remove systematic errors from the observations and model background.

In the first part of the thesis, we introduce an innovative BC method that uses a Taylor series polynomial expansion of the observation-minus-background (OMB) departures to remove linear and nonlinear conditional biases from all-sky satellite infrared brightness temperatures. Passive monitoring experiments reveal that variables sensitive to the cloud top height are the most effective BC predictors and that higher-order Taylor series terms are necessary to account for complex nonlinear biases in the OMB departures. Active data assimilation experiments using the nonlinear BC method show that the model background is most improved when higher-order cloud-sensitive predictors are employed.

Following this work, we use the Lorenz-63 model to develop a model bias estimation method based on an asymptotic expansion of the model dynamics for small time scales and small perturbations in one of its parameters. The model bias estimators are subsequently used to improve the model background error covariance matrix used during the data assimilation step. It is shown that the combination of a static matrix with a dynamic matrix that varies with time leads to more accurate model analyses and forecasts. Together, results from this thesis demonstrate that bias predictors derived from polynomial expansions of modeled and observed variables can improve the performance of data assimilation systems.

Acknowledgements

Earning a PhD is a long and arduous process that requires the support of many people to be successful. I have been fortunate to have received encouragement and support from various people and organizations on both sides of the Atlantic Ocean as I worked part-time on my PhD while also having a full-time job and being a parent.

First, I would like to thank my wife and daughter for allowing me the time and space to work on my PhD. This was a big time commitment for all of us, but somehow we found enough time to get it done. I am very thankful for your willingness to pick up the slack while I worked on my PhD during evenings and weekends, and especially during my longer absences to meet with my advisors in Reading. I would not have been able to do this without your support during the past four and a half years. Our daughter has also provided additional motivation along the way. She is such a delight to be around. Her smiles and hugs always made things easier and the prospect of having more time to play with her pushed me to work harder each day.

I also owe a tremendous debt of gratitude to my supervisors Dr. Roland Potthast and Dr. Amos Lawless for their excellent support, guidance, and encouragement throughout my PhD. I am especially thankful for all the times that Roland met with me in Reading. These trips took him away from his family in Germany so I would like to send a special thank you to them for letting me borrow some of his time. Both of my advisors are exceptional people and I feel very blessed to have had them guide me down the road toward my PhD.

I want to thank my monitoring committee members Dr. Peter Chamberlain, Dr. Sarah Dance, and Dr. Andrea Moiola for their input and encouragement. I also extend my appreciation to colleagues at the University of Wisconsin-Madison for dealing with my long absences as I worked on my PhD. Finally, I acknowledge financial support from the University of Reading via an International Research Studentship award.

Last, but not least, I am also grateful for my parents for instilling a strong work ethic in me and for providing a loving and caring environment that encouraged education.

Contents

Chapter 1

Introduction

1.1 Motivation

Numerical weather prediction (NWP) models simulate the evolution of the atmosphere and other aspects of the earth system using a set of partial differential equations and various parameterization schemes that together account for processes occurring within and at the boundaries between the atmosphere, land, and ocean. The dynamic and thermodynamic characteristics of the atmosphere at a given location, and how they evolve with time, are typically described using Newton's law of motion for fluid air and the laws of thermodynamics. These continuous nonlinear partial differential equations are generally impossible to solve exactly; therefore, numerical approximations are used to transform them into discrete difference equations suitable for use in NWP models (Strikwerda 2004; Mengaldo et al. 2019). Because many environmental processes are either too complex to adequately represent using these equations or occur at spatial and temporal scales that are too small to be fully resolved on the model grid, parameterization schemes are used to estimate their statistical impact on the resolved variables (Stensrud 2007). Parameterization schemes account for a wide range of unresolved processes such as turbulent mixing in the planetary boundary layer, the evolution of cloud hydrometeors as cloud systems grow and decay, and the transfer of electromagnetic radiation through the atmosphere and its interaction with clouds, aerosols, and the earth's surface.

Despite the numerous challenges associated with simulating complex and chaotic processes, such as those that occur in the atmosphere, the accuracy of NWP model forecasts has substantially increased during the past several decades. For example, Magnusson and Kallen (2013) showed

that the average accuracy of a 3-day forecast generated using the 1980 version of the European Centre for Medium-range Weather Forecasts (ECMWF) operational modeling system could be obtained in a 6-day forecast using the 2010 version of the system. This extension of the same forecast skill to longer lead times greatly enhances the utility of the model forecasts by providing stakeholders and the general public additional time to prepare for adverse weather conditions. A recent study by Hoffman et al. (2018) has shown that forecast errors have continued to decrease during the past 10 years at all of the major operational weather forecasting centers, and that the errors are much smaller than they were 40 years ago.

The notable increase in forecast skill can be attributed to advancements in many areas, such as improvements made to the forecast models through development of more sophisticated parameterization schemes and to increases in spatial resolution, both of which permit a more accurate representation of physical processes. The development of advanced data assimilation capabilities has proven to be equally important because the more effective assimilation of various observations has greatly improved the accuracy of the datasets that are used to initialize NWP model forecasts (Magnusson and Kallen 2013). All other things being equal, the accuracy of a model forecast is closely tied to the quality of the initial and lateral boundary conditions that it is provided. Thus, efforts to further improve the accuracy of these datasets through the more effective assimilation of existing observations are critical for gaining additional increases in forecast skill. This is especially true in cloudy regions where model errors remain large due to the tendency for nonlinear error growth and the difficulties associated with assimilating satellite brightness temperatures affected by clouds (McNally 2002; Geer and Bauer 2011). Indeed, most observations impacted by clouds and precipitation are currently not being assimilated at global NWP centers (Yang et al. 2016). Neglecting observations impacted by clouds is problematic however because they tend to be located in dynamically active regions where more accurate initial conditions could help constrain potentially rapid error growth during the forecasts. Therefore, one of the primary objectives of this thesis is to improve the accuracy of model initialization datasets through the more effective use of cloud-sensitive satellite infrared brightness temperatures in data assimilation systems. This goal will be accomplished through development of a novel all-sky bias correction method that is able to remove complex biases from these important observations prior to their assimilation.

1.2 Data Assimilation

Data assimilation refers to a mathematical technique that seeks to optimally combine imperfect observations with a model background or "first guess" that also contains errors in order to determine the best estimate of the atmospheric state and lower boundary conditions associated with the land and ocean surfaces. Output from a short-range model forecast is typically used as the first guess because it provides a more complete representation of the atmospheric conditions and is necessary to make the inverse problem well-posed given that the number of observations is typically much smaller than the dimension of the NWP model. Various data assimilation methods have been developed during the past 60 years ranging from simple methods such as optimal interpolation (Gandin 1962) to more complex methods such as variational data assimilation (Lewis and Derber 1985; Le Dimet and Talagrand 1986; Klinker et al. 2000) and the ensemble Kalman filter (EnKF; Evensen 1994; Houtekamer and Mitchell 1998). More recently, hybrid data assimilation methods that combine aspects of the variational and ensemble approaches have also been developed (e.g., Buehner et al. 2010; Clayton et al. 2013). All of these latter methods include covariance information that can be used to update unobserved variables, thereby allowing remotely sensed observations such as satellite brightness temperatures to be assimilated. This is important because satellite observations are collectively the most important observation type assimilated in most operational NWP modeling systems (Bauer et al. 2010).

Though various mathematical perspectives can be used to discuss theoretical aspects of data assimilation, the most common approach is to view it as a Bayesian estimation problem. Variational data assimilation methods seek the solution that maximizes the *a posteriori* probability $p(\mathbf{x}|\mathbf{y})$ through minimization of a quadratic cost function $J(\mathbf{x})$. For 3-dimensional variational (3DVAR) assimilation, where all of the observations valid during a given assimilation window are assimilated at the same analysis time, this leads to the following cost function:

$$J(\mathbf{x}) = \frac{1}{2}(\mathbf{x}-\mathbf{x}_b)^{\mathrm{T}}\mathbf{B}^{-1}(\mathbf{x}-\mathbf{x}_b) + \frac{1}{2}(H(\mathbf{x})-\mathbf{y})^{\mathrm{T}}\mathbf{R}^{-1}(H(\mathbf{x})-\mathbf{y}), \tag{1.1}$$

where $\mathbf{x} \in \mathbb{R}^n$ and $\mathbf{x}_b \in \mathbb{R}^n$ are the model analysis and background vectors, $\mathbf{B} \in \mathbb{R}^{n\times n}$ and $\mathbf{R} \in \mathbb{R}^{m\times m}$ are the model background and observation error covariance matrices, H is the observation operator that maps from the model space to the observation space, $\mathbf{y} \in \mathbb{R}^m$ is the observation vector, $n \in \mathbb{N}$ is the dimension of the model space, and $m \in \mathbb{N}$ denotes the number of observations. Because the

NWP model and observations contain errors, the model background $\mathbf{x}_b$ and observations $\mathbf{y}$ only approximate the true system state $\mathbf{x}^{true}$:

$$\mathbf{x}_b = \mathbf{x}^{true} + \boldsymbol{\epsilon}_b + \boldsymbol{\beta}_b \tag{1.2}$$

$$\mathbf{y} = H(\mathbf{x}^{true}) + \boldsymbol{\epsilon}_o + \boldsymbol{\beta}_o, \tag{1.3}$$

where $\boldsymbol{\epsilon}_b \in \mathbb{R}^n$ and $\boldsymbol{\epsilon}_o \in \mathbb{R}^m$ represent random errors in the model background and observations, respectively, and $\boldsymbol{\beta}_b \in \mathbb{R}^n$ and $\boldsymbol{\beta}_o \in \mathbb{R}^n$ refer to their biases. In most data assimilation systems, the model background and observations are assumed to be unbiased or that any biases in the datasets can be removed using suitable bias correction methods. The model background and observation error covariance matrices can then be defined as $\mathbf{B} = \mathbb{E}[\boldsymbol{\epsilon}_b(\boldsymbol{\epsilon}_b)^T]$ and $\mathbf{R} = \mathbb{E}[\boldsymbol{\epsilon}_o(\boldsymbol{\epsilon}_o)^T]$, where $\mathbb{E}$ is the expectation operator.

If the above framework is extended to four dimensions (4DVAR) so that the observations can be incorporated at their exact time of measurement, the cost function becomes:

$$J(\mathbf{x}_0) = \frac{1}{2}(\mathbf{x}_0 - \mathbf{x}_b)^{\mathrm{T}}\mathbf{B}^{-1}(\mathbf{x}_0 - \mathbf{x}_b) + \frac{1}{2}\sum_{i=1}^{N}(H_i(\mathbf{x}_i) - \mathbf{y}_i)^{\mathrm{T}}\mathbf{R}_i^{-1}(H_i(\mathbf{x}_i) - \mathbf{y}_i) \tag{1.4}$$

where $\mathbf{x}_0 \in \mathbb{R}^n$ is the model state at the beginning of the assimilation window, $\mathbf{x}_i = M(\mathbf{x}_0) \in \mathbb{R}^n$ is the model state at each observation time t_i obtained through integration of the numerical model M from the initial state $\mathbf{x}_0$ at time t_0, H_i and $\mathbf{R}_i \in \mathbb{R}^{m\times m}$ are the observation operator and error covariance matrix for the ith observation $\mathbf{y}_i$, and N is the total number of observations. This is the standard formulation for strong-constraint 4DVAR where the numerical model M is assumed to be perfect. The first term on the righthand sides of (1.1) and (1.4) measures the departure of the analysis vector from the model background, weighted by the inverse of the background error covariances. Likewise, the second term in each equation measures the departure of the observations from their corresponding model equivalents, weighted by the inverse of the observation error covariances. Iterative approaches such as the conjugate gradient method (Hestenes and Stiefel 1952) are typically used to find the solution that minimizes the cost function $J(\mathbf{x})$.

The $\mathbf{B}$ matrix in variational data assimilation systems has traditionally been used to represent average conditions over a long period of time; however, this means that it cannot properly capture the true background uncertainty at a given time if it deviates strongly from climatology, such as often occurs in dynamically active regions. Because of this limitation, some operational weather

forecasting centers have developed new methods to generate $\mathbf{B}$ that allow it to better capture the true uncertainty during a given assimilation cycle. One way this is accomplished is by combining information about current atmospheric conditions from a reduced resolution ensemble with a static (e.g., climatological) higher-resolution $\mathbf{B}$ matrix (Isaksen et al. 2010). An alternative approach is to use an EnKF data assimilation system where $\mathbf{B}$ is updated during each assimilation cycle using an ensemble of full-resolution model forecasts integrated from the previous analysis time. By using an ensemble to approximate the forecast uncertainty, an EnKF is a Monte Carlo implementation of the Bayesian update problem, which can be expressed as:

$$p^{(a)}(\mathbf{x}|\mathbf{y}) = \frac{p^{(b)}(\mathbf{x})p^{(o)}(\mathbf{y}|\mathbf{x})}{p(\mathbf{y})}, \tag{1.5}$$

where $p^{(a)}(\mathbf{x}|\mathbf{y})$ is the posterior probability distribution conditioned on the observations $\mathbf{y}$, $p^{(b)}(\mathbf{x})$ is the prior probability distribution, $p^{(o)}(\mathbf{y}|\mathbf{x})$ is the observation error distribution conditioned on $\mathbf{x}$, and $p(\mathbf{y})$ is the probability distribution of the observations. Because the observations are assumed to be independent of the current model state $\mathbf{x}$, $p(\mathbf{y})$ can be set to a constant value. We then want to find the model analysis $\mathbf{x}$ that maximizes the posterior probability $p^{(a)}(\mathbf{x}|\mathbf{y})$. This is equivalent to finding the posterior state that best represents the new observations $\mathbf{y}$ and the prior state. The probability distributions are usually assumed to be Gaussian, which leads to:

$$p^{(b)}(\mathbf{x}) = \frac{1}{\sqrt{2\pi|\mathbf{B}|^n}}\exp\Big(-\frac{1}{2}(\mathbf{x}-\mathbf{x}_b)^T\mathbf{B}^{-1}(\mathbf{x}-\mathbf{x}_b)\Big), \tag{1.6}$$

and

$$p^{(o)}(\mathbf{y}|\mathbf{x}) = \frac{1}{\sqrt{2\pi|\mathbf{R}|^m}}\exp\Big(-\frac{1}{2}(\mathbf{y}-h(\mathbf{x}))^T\mathbf{R}^{-1}(\mathbf{y}-h(\mathbf{x}))\Big), \tag{1.7}$$

Then, from Bayes theorem, the posterior probability distribution becomes:

$$p^{(a)}(\mathbf{x}|\mathbf{y}) \propto \exp\Big(-\frac{1}{2}(\mathbf{x}-\mathbf{x}_b)^T\mathbf{B}^{-1}(\mathbf{x}-\mathbf{x}_b)) - \frac{1}{2}(\mathbf{y}-h(\mathbf{x}))^T\mathbf{R}^{-1}(\mathbf{y}-h(\mathbf{x}))\Big), \tag{1.8}$$

where we ignore the constants multiplying the exponents in (1.6) and (1.7) because they are independent of $\mathbf{x}$. The maximum likelihood of the posterior analysis is the $\mathbf{x}$ that maximizes the terms in the exponent of (1.8).

In an EnKF system, the ensemble mean analysis $\overline{\mathbf{x}}^a \in \mathbb{R}^n$ during a given assimilation cycle is then solved using:

$$\overline{\mathbf{x}}^a = \overline{\mathbf{x}}^f + \mathbf{K}(\mathbf{y} - H(\overline{\mathbf{x}}^f)), \tag{1.9}$$

where $\bar{\mathbf{x}}^f \in \mathbb{R}^n$ is the forecast ensemble mean and $\mathbf{K}$ is the Kalman gain matrix given by:

$$\mathbf{K} = \mathbf{P}^f \mathbf{H}^{\mathrm{T}} (\mathbf{H}\mathbf{P}^f \mathbf{H}^{\mathrm{T}} + \mathbf{R})^{-1} \tag{1.10}$$

Here, $\mathbf{P}^f \in \mathbb{R}^{n \times n}$ denotes the forecast error covariance matrix, which is equivalent to the background error covariance matrix $\mathbf{B}$ used in variational data assimilation methods except that it has the important distinction of varying with time, thereby allowing it to better capture the true uncertainty in the model background. An EnKF is a sequential data assimilation method that is a Monte Carlo approximation of the extended Kalman filter (Kalman 1960) where the true covariance matrix is replaced by the sample covariance matrix computed from an ensemble. This approximation allows it to be implemented very efficiently and to scale well to the very large dimensions used by modern geophysical models (Houtekamer and Zhang 2016). Covariance localization methods are typically used to make the ensemble less rank deficient and to limit the impact of spurious correlations occurring at long distances from a given grid point (Houtekamer and Mitchell 1998; Hamill et al. 2001; Mitchell et al. 2002). Deficient ensemble spread resulting from a finite ensemble size may also lead to filter divergence where the observations are no longer able to pull the model trajectory back toward the observations. This problem can be mitigated through use of covariance inflation methods that act to increase the variance explained by the ensemble and therefore allow the observations to have a larger impact on the analysis (Anderson and Anderson 1999; Zhang et al. 2004; Anderson 2009). The EnKF methodology has become widely used during the past 15 years due to its ease of implementation. It has become particularly popular in high-resolution regional-scale data assimilation systems for which flow-dependent background error covariances are especially important (Bonavita et al. 2010; Caron et al. 2015; Schwartz et al. 2015).

Though the mathematics underlying 4DVAR and EnKF data assimilation systems differ in important ways, it is important to note that in the theoretical context of a linear NWP model, linear observation operator, and Gaussian statistics, that the EnKF and 4DVAR solutions will be identical at the end of the assimilation window provided that the EnKF employs a sufficiently large ensemble (Lorenc 1986). For both methods, the observation error covariance matrix $\mathbf{R}$ includes contributions due to measurement error and representativeness error. It is often assumed to be diagonal, which means that it is unable to account for correlated observation errors that may occur both in space and between different channels on a given satellite sensor. This assumption is problematic for all-sky satellite observations because correlated observation errors can arise due

to limitations in the forward observation operator that maps from the model to observation space and in the inability of the NWP model to properly resolve the atmospheric and surface features seen in the observations. The impact of correlated observation errors has traditionally been mitigated to some extent through use of variance inflation and observation thinning procedures (e.g., Courtier et al. 1998; Hilton et al. 2009; Bormann et al. 2016; Janjic et al. 2018).

Desroziers et al. (2005) developed a diagnostic tool that can be used to estimate **R** by considering relationships between the observation-minus-background (OMB) and observation-minus-analysis statistics. Though their diagnostic relies on linearity and assumes that the covariance matrices used in the data assimilation system are close to correct, it has been used successfully in various settings (Bormann et al. 2010a,b, 2011; Miyoshi et al. 2013; Stewart et al. 2014; Waller et al. 2016a). A recent study by Waller et al. (2016b) used this method to investigate spatial and inter-channel correlations in satellite all-sky infrared brightness temperatures. They found that strong inter-channel correlations occur between the water vapor channels, as well as between the various surface-sensitive channels, and that the correlations vary spatially across the model domain. Another recent study by Michel (2018) investigated horizontal spatial correlations in all-sky satellite observations. Building upon the work of Fisher (2005), they showed that it is possible to model the structure of the spatial correlations by applying a sequence of linear operators. Their method uses the Lanczos algorithm to provide a low-rank eigenvalue decomposition of the correlation matrix that is then regularized and explicitly inverted.

The matrices **B** and **R** are important because the *a posteriori* model analysis will be pulled more strongly toward (away from) the observations when there is greater (less) uncertainty in the model background or when the observation error uncertainty is smaller (larger). It is therefore necessary to correctly estimate these covariance matrices in order to make full use of the information content provided by the model background and the observations. Despite its importance, it is impossible to specify the exact form of the **B** matrix in real-world applications because the true state of the system $\mathbf{x}^{true}$ is unknown. For ensemble data assimilation systems, the ensemble mean from a short-term forecast $\overline{\mathbf{x}}^{f}$ serves as the best estimate of the true state when computing the background error covariance matrix; however, in variational data assimilation systems, it is common to estimate the **B** matrix using the so-called NMC method (Parrish and Derber 1992). With this method, the background error covariances are estimated by taking differences between pairs of forecasts valid at the same time but initialized at different times. The two forecasts, in this case represented by 24 and 48 h forecasts valid at the same time, can be written as a linear combination of the true state

and their associated errors:

$$\mathbf{x}^{48} = \mathbf{x}^{true} + \boldsymbol{\epsilon}^{48} + \boldsymbol{\beta}^{48} \tag{1.11}$$

$$\mathbf{x}^{24} = \mathbf{x}^{true} + \boldsymbol{\epsilon}^{24} + \boldsymbol{\beta}^{24}, \tag{1.12}$$

where $\boldsymbol{\epsilon}^{48}$ and $\boldsymbol{\epsilon}^{24}$ are the random errors and $\boldsymbol{\beta}^{48}$ and $\boldsymbol{\beta}^{24}$ are the biases for each of the forecasts. The true state will then cancel out when taking the difference between the forecasts, and assuming that there is no bias or that the bias is constant with time, the difference becomes:

$$\mathbf{x}^{diff} = \mathbf{x}^{48} - \mathbf{x}^{24} = \boldsymbol{\epsilon}^{48} - \boldsymbol{\epsilon}^{24}. \tag{1.13}$$

The background error covariance matrix can then be written as:

$$\mathbf{B} = \overline{(\mathbf{x}^{48} - \mathbf{x}^{24})(\mathbf{x}^{48} - \mathbf{x}^{24})^T} = \overline{(\boldsymbol{\epsilon}^{48} - \boldsymbol{\epsilon}^{24})(\boldsymbol{\epsilon}^{48} - \boldsymbol{\epsilon}^{24})^T}, \tag{1.14}$$

where the overbar denotes an average over space or time.

The assumption that the model forecasts are unbiased or that the biases are the same for both forecasts is questionable because the model accuracy can strongly depend upon the prevailing atmospheric and surface conditions. As such, it is reasonable to assume that the model forecasts may exhibit large and potentially nonlinear biases in some situations that are conditioned on one or more predictors. This problem also arises in ensemble data assimilation systems because the use of an imperfect NWP model may lead to a forecast ensemble that has too much or too little spread for some model variables in certain situations. This means that the impact of a model bias on the background error covariance matrix $\mathbf{P}^f$ could vary with space and time such that the use of covariance inflation methods that are applied uniformly to all grid points will not be able to account for the model bias in a realistic manner. Thus, a second objective of this thesis is to develop a mathematical method that can be used to identify model biases using the model state variables and then use these estimates to improve the background error covariance matrix. The model bias estimation method discussed in Chapter 5 is based on an asymptotic expansion of the model variables and in its most general form does not require direct knowledge of the relationship between the model bias and the underlying model dynamics. We choose to use these model bias estimators to improve the quality of the $\mathbf{B}$ matrix in a variational data assimilation system rather than to explicitly develop a new model error covariance matrix $\mathbf{Q} \in \mathbb{R}^{n \times n}$. This choice is justified because the $\mathbf{B}$ matrix is computed using imperfect forecasts that include a model error component.

Our model bias estimation method is able to account for at least some portion of the model error in the **B** matrix without having to go through the computational expense of estimating and inverting an independent model error covariance matrix **Q**.

1.3 Thesis Goals

A key assumption in data assimilation is that the model first guess and observations are characterized only by random zero-mean errors when they are combined to create the *a posteriori* analysis. In reality, however, biases are often present both in the observations and in the model first guess. Systematic errors in the assimilation system can manifest themselves through regularly recurring spatial patterns in the analysis increments or through the presence of persistently large increments in certain regions or during specific atmospheric flow regimes (Dee 1995). They can also be diagnosed in the OMB departures during the data assimilation step (Eyre 1992, 2016). OMB biases can be especially large, and also display complex nonlinear patterns, when cloud-affected satellite observations are assimilated. For example, deficiencies in the forward radiative transfer model used to compute the model-equivalent brightness temperatures and the inability of the parameterization schemes in the NWP model to properly represent some cloud properties could lead to large OMB departures. Though substantial progress has been made in our ability to accurately simulate clouds and other aspects of the earth system, large errors remain when model forecasts are compared to observations (Fabry and Sun 2010; Geer and Bauer 2011). This result points toward the continued need to develop mathematical methods to remove biases from the observations and model background in order to promote the more effective use of observations in modern data assimilation systems.

In this thesis, the primary goal is to enhance our ability to assimilate cloud-sensitive satellite observations through development of new methods to identify and remove systematic errors from the OMB departures and model background. This will be accomplished by 1) developing a new bias correction (BC) method to remove complex nonlinear biases from all-sky infrared brightness temperatures via a Taylor series polynomial expansion of the OMB departures, 2) demonstrating that its use leads to more accurate short-range forecasts in a high-resolution regional-scale ensemble data assimilation system, and 3) improving the background error covariance matrix **B** through development of dynamic model bias estimators based on an asymptotic expansion of the model dynamics in an idealized model of atmospheric convection.

1.4 Thesis Outline

The remaining sections of the thesis are organized as follows. Note that comprehensive literature reviews are included at the beginning of Chapters 3-5 consistent with the topics discussed in each of the three journal articles that comprise this thesis.

- **Chapter 2** provides background material describing the primary mathematical methods employed during this thesis.

- **Chapter 3** introduces a new BC method that uses a Taylor series polynomial expansion of the OMB departures to remove linear and nonlinear conditional biases from all-sky satellite infrared brightness temperatures. The OMB departures were accumulated over a 5-day period in which the satellite observations were passively monitored and conventional observations were actively assimilated. The results showed that quantities sensitive to the cloud top height were the most effective BC predictors and that higher-order Taylor series terms were necessary to account for complex nonlinear conditional biases in the OMB departures. The work presented in this chapter also appears in the January 2018 issue of *Monthly Weather Review* (Otkin et al. 2018).

- **Chapter 4** examines the impact of using the nonlinear BC method developed in Chapter 3 during cycled data assimilation experiments in which all-sky satellite infrared brightness temperatures were actively assimilated during a 3-day period. Comparisons to radiosonde observations showed that the model background was less accurate when the satellite observations were assimilated without first removing their biases; however, the results were substantially improved when the nonlinear BC method was used to remove the linear and nonlinear conditional biases from the observations prior to their assimilation. Experiments employing BC predictors sensitive to the cloud top height had the smallest forecast errors because they were most effective at removing large conditional biases associated with a deficiency in upper-level clouds in the model background. The work presented in this chapter also appears in the December 2019 issue of *Monthly Weather Review* (Otkin and Potthast 2019).

- **Chapter 5** presents a new method to generate dynamic model bias estimators and then uses them to improve the performance of a 3DVAR data assimilation system. The Lorenz (1963)

model was chosen for this part of the thesis so that all aspects of the system and bias estimators could be studied in a way that would not be possible if a full physics NWP model were used. In particular, the asymptotics of the Lorenz (1963) model are derived for small changes in one of its parameters and then statistics from cycled data assimilation experiments are used to demonstrate that the asymptotics accurately represent the behavior of the model and that the coefficients of the nonlinear asymptotical expansion can be reasonably estimated by solving a least squares minimization problem. It is then shown that the combination of a constant **B** matrix with a dynamic **B** matrix that varies with time based on the model bias estimators leads to more accurate model analyses. The work presented in this chapter has been conditionally accepted for publication in the *SIAM Journal on Applied Dynamical Systems* (Otkin et al. 2020).

- **Chapter 6** summarizes the results of the thesis and discusses opportunities for future work.

Chapter 2

Mathematical Background

An overarching theme of this thesis is its use of polynomial equations and regularized least square regression problems to identify a set of bias correction coefficients that can then be used to remove linear and nonlinear conditional biases from the observation departures or the model background. A polynomial equation is a linear combination of one or more algebraic terms built using addition, subtraction, multiplication, and exponentiation. Each term in the equation consists of the product of a constant and a finite number of variables (or indeterminates) raised to a non-negative integer power. For a single variable case, a polynomial function $f(x)$ has the form:

$$f(x) = a_n x^n + a_{n-1} x^{n-1} + ... + a_2 x^2 + a_1 x + a_0, \tag{2.1}$$

where $a_0, ..., a_n$ are the constants, and the degree of the function refers to the largest exponent of any one term with a non-zero constant. For a polynomial function containing more than one variable, the degree of a given term is the sum of the exponents for each of the variables in that term. Polynomials with smaller degrees are given specific names such as constant, linear, quadratic, and cubic for polynomials of degree 0, 1, 2, and 3, respectively. Polynomial functions are smooth and continuous and can be used to evaluate all values of x in the domain of f. Unless otherwise specified, polynomial functions can have complex constants, variables, and values (Barbeau 2003).

As will be shown in subsequent chapters, it is important to account for nonlinear error characteristics when attempting to remove biases from the all-sky observation departures and model background. One way that this can be accomplished is by using a Taylor series polynomial expansion that includes higher order terms that are able to capture nonlinear error dependencies if they

exist in the data. Taylor's theorem is one of the basic tools used in mathematics and serves as the starting point for more advanced asymptotic analysis (Apostol 1974). It provides an approximation of a function $f(x)$, whose form can be known or unknown, in a neighborhood surrounding a given point c, and can be thought of as the extension of simple linear approximations to higher order polynomials. The most important type of series is the power series:

$$f(x) = \sum_{n=0}^{\infty} a_n (x-c)^n, \tag{2.2}$$

where a_n are the coefficients and the domain of $f(x)$ is the set of x surrounding c where the series converges. Though convergence is not guaranteed for all functions, in practice, most functions of interest have a convergent Taylor series. To determine which power series provides the best fit to a given function $f(x)$, the power series and function must agree at c, and $f(x)$ should have derivatives of all orders in an open interval I containing c. For each positive integer n and for each $x \in I$, this leads to the Taylor series for $f(x)$:

$$f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(c)}{n!} (x-c)^n, \tag{2.3}$$

where $f^{(n)}(c)$ is the nth derivative of f evaluated at the point c and $n!$ is the factorial of n. Partial sums are subsequently used to approximate the function $f(x)$ to the nth order and to estimate the size of the remainder term that represents the error of the Taylor series:

$$f(x) = T_n(x) + R_n(x) \tag{2.4}$$

where

$$T_n(x) = \sum_{n=0}^{n} \frac{f^{(n)}(c)}{n!} (x-c)^n \tag{2.5}$$

is the nth-order Taylor series polynomial expansion of $f(x)$ evaluated at the point c, and the remainder term $R_n(x)$ is defined as:

$$R_n(x) = \frac{f^{(n+1)}(z_n)}{(n+1)!} (x-c)^{n+1} \tag{2.6}$$

for some z_n between c and x (Abramowitz and Stegun 1972). If $R_n(x) \to 0$ as $n \to \infty$ for all $x \in I$, then the Taylor series generated by $f(x)$ at $x = c$ converges to $f(x)$ on I. Note that the remainder term is not part of the Taylor series.

For illustrative purposes, if we assume that the bias in a given observation departure dataset **dy** can be described by a real or complex function $f(z)$ of a single variable (e.g., predictor) that is infinitely differentiable around a real or complex number c, then this function can be represented as an N order Taylor series polynomial expansion:

$$\mathbf{dy} = \left(f(c) + \frac{f'(c)(z^{(i)} - c)}{1!} + \frac{f''(c)(z^{(i)} - c)^2}{2!} + \frac{f'''(c)(z^{(i)} - c)^3}{3!} + ... + \frac{f^{(n)}(c)(z^{(i)} - c)^n}{n!} \right)_{i=1,...,m} \tag{2.7}$$

where m is the number of observations, **dy** is the m x 1 observation departure vector, $f^{(n)}(c)$ is the nth derivative of f evaluated at the point c, and $z^{(i)}$ is the value of the variable used to predict the bias for the ith observation. The $i = 1,...,m$ notation outside the parentheses indicates that the Taylor series terms are computed separately for each element of the **dy** vector using the equation shown in the parentheses. The variable used as the bias predictor is chosen based on its ability to capture linear or nonlinear characteristics of the error distribution. The single variable case shown above can subsequently be generalized to be a function of more than one predictor.

If we continue with the above example using a single variable, but truncate the Taylor series expansion to only include terms up to the third order and define the bias correction coefficients as $b_n = \frac{f^{(n)}(a)}{n!}$, (2.7) can then be rewritten as:

$$\mathbf{dy} = \left(b_0 + b_1(z^{(i)} - c) + b_2(z^{(i)} - c)^2 + b_3(z^{(i)} - c)^3 \right)_{i=1,...,m} \tag{2.8}$$

or, more compactly, in matrix notation as:

$$\mathbf{dy} = \mathbf{Ab}, \tag{2.9}$$

where **A** is an m x n matrix containing the n Taylor series terms for each observation and **b** is an n x 1 vector containing the bias correction coefficients that we want to solve for. The first column of **A** contains ones, with the remaining columns containing the linear and higher order Taylor series terms. It is readily apparent in (2.8) that the higher order terms represent nonlinear bias predictors because the exponents are ≥ 2, with the $(z^{(i)} - c)^2$ and $(z^{(i)} - c)^3$ polynomials representing the quadratic and cubic terms, respectively.

In real-world situations, (2.9) will typically represent an overdetermined system of m linear equations in n unknown coefficients because $m > n$. This kind of system generally will not have

an analytic solution; therefore, least squares regression methods are commonly used to obtain an approximate solution that best fits the set of equations. The goal of regression analysis is to model the expected value of a dependent variable y based on an independent variable x. With least squares regression methods, the solution is obtained by minimizing the sum of the square of the residuals, where residual refers to the difference between the observed value of the dependent variable and the fitted value provided by the regression model. Unlike linear least squares regression that finds the best linear fit to the data, polynomial regression is more flexible because it uses an nth degree polynomial of x to model the potentially nonlinear relationship between the independent variable x and biases in the dependent variable y. As such, it provides a useful means with which to diagnose and remove both linear and nonlinear conditional biases from a dataset. Though it fits a nonlinear model to the data, polynomial regression is considered to be a special case of multiple linear regression because the estimation problem remains linear being that the regression function is linear in the unknown parameters estimated from the data. A polynomial regression model takes the same form as that shown in (2.9). The coefficients that best fit the linear system of equations can be found by solving the quadratic minimization problem $\hat{b} = \min_b S(b)$, where the objective function S is given by:

$$S(b) = \sum_{i=1}^{m} |dy_i - \sum_{j=1}^{n} A_{ij} b_j|^2 = \|\mathbf{dy} - \mathbf{Ab}\|^2 \tag{2.10}$$

and $\|\cdot\|$ is the Euclidean norm. The least squares solution can then be found by differentiating S with respect to b, and equating to zero, which leads to:

$$\mathbf{b} = (\mathbf{A}^{\mathrm{T}}\mathbf{A})^{-1}\mathbf{A}^{\mathrm{T}}\mathbf{dy}, \tag{2.11}$$

where $(\mathbf{A}^{\mathrm{T}}\mathbf{A})$ is a symmetric, n x n square matrix, thereby making it easy to compute its inverse.

It is well known in inverse modeling that some form of regularization may be necessary to make the problem well-posed (Nakamura and Potthast 2015). Regularization refers to the situation where the original inverse problem is replaced by a different solution that roughly matches the desired solution (low bias), is less sensitive to changes in the data (low variance), and has a parameter that permits a reasonable balance between the low bias and low variance considerations. The aim is to reduce the complexity of the model in order to prevent overfitting of the solution due to the use of a large number of parameters. Various regularization methods have been developed; however, the most common approach is to use Tikhonov regularization given its simplicity.

With this method, the conditioning of the inverse problem is improved by adding a term to the least squares estimator shown in (2.11) that takes the form of the identity matrix multiplied by a regularization parameter α_{reg} (Tikhonov 1963). With this change, the least squares solution can be found using:

$$\mathbf{b} = (\alpha_{reg} I + \mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T \mathbf{dy}. \tag{2.12}$$

By adding a multiple of the identity matrix to $\mathbf{A}^T \mathbf{A}$, this acts as a positive constant shift to the main diagonal, thereby decreasing the condition number of the matrix. The effects of the regularization disappear as $\alpha_{reg} \to 0$. It is important to use a reasonable value for α_{reg} because a regularization that is too small will leave the problem ill-posed whereas a regularization that is too large will change the nature of the problem and thus lead to larger bias. A heuristic approach can sometimes be used to identify a suitable value for α_{reg}; however, automated approaches such as ridge trace method (Hoerl and Kennard 2000), quasi-optimal criterion (Kindermann and Neubauer 2008), generalized cross validation (Jia et al. 2015), and Morozov discrepancy principle (Bonesky 2009) can also be used to obtain the optimal value for this parameter.

Chapter 3

Nonlinear Bias Correction – Passive Monitoring Experiments

3.1 Abstract

Output from a high-resolution ensemble data assimilation system is used to assess the ability of an innovative nonlinear bias correction (BC) method that uses a Taylor series polynomial expansion of the observation-minus-background departures to remove linear and nonlinear conditional biases from all-sky satellite infrared brightness temperatures. Univariate and multivariate experiments were performed in which the satellite zenith angle and variables sensitive to clouds and water vapor were used as the BC predictors. The results showed that even though the bias of the entire observation departure distribution is equal to zero regardless of the order of the Taylor series expansion, there are often large conditional biases that vary as a nonlinear function of the BC predictor. The linear 1st order term had the largest impact on the entire distribution as measured by reductions in variance; however, large conditional biases often remained in the distribution when plotted as a function of the predictor. These conditional biases were typically reduced to near zero when the nonlinear 2nd and 3rd order terms were used. The univariate results showed that variables sensitive to the cloud top height are effective BC predictors especially when higher order Taylor series terms are used. Comparison of the statistics for clear-sky and cloudy-sky observations revealed that nonlinear departures are more important for cloudy-sky observations as signified by the much larger impact of the 2nd and 3rd order terms on the conditional biases. Together, these results indicate that the nonlinear BC method is able to effectively remove the bias

from all-sky infrared observation departures.

3.2 Introduction

The ability to generate accurate cloud and water vapor (WV) analyses suitable for numerical weather prediction (NWP) models is perhaps the most challenging aspect of modern data assimilation (DA) systems because they typically assume Gaussian error statistics and that linear relationships exist between the observations and model state variables. Cloud processes, however, are inherently nonlinear with complex interactions occurring between different cloud hydrometeor species and the local thermodynamic environment at spatial and temporal scales that are typically much smaller than those represented by NWP models. Likewise, WV content can change rapidly in space and time and can influence the evolution of the cloud field in nonlinear ways. These and other factors can make it very challenging to effectively assimilate information from cloud and WV sensitive observations.

Remotely sensed observations obtained using geostationary and polar-orbiting satellites provide the only reliable source of high-resolution cloud and WV information covering large geographic domains. Sophisticated visible, infrared, and microwave sensors onboard various satellite platforms provide information about the spatial distribution and characteristics of the cloud and WV fields. For regional-scale NWP, observations from geostationary satellites are especially useful because their continuous viewing of the same area with high temporal and spatial resolution allow them to more easily constrain the evolution of rapidly changing weather features (Vukicevic et al. 2006; Errico et al. 2007). Satellite observations, however, often exhibit biases when compared to their model equivalents computed using the NWP model background; therefore, bias correction (BC) methods are typically required to assimilate these observations (Eyre 2016).

Observation-minus-background (OMB) biases can occur for a variety of reasons and can differ for clear and cloudy observations. For example, biases can arise from calibration errors in a satellite sensor or to instrument "drift" as a sensor ages. Biases can also be introduced by deficiencies in the forward radiative transfer models used to compute the model equivalent brightness temperatures. For clear-sky observations, biases may result from errors in the specification of surface emissivity, simplifications in the radiative transfer model equations, inadequate vertical resolution or a low model top in the NWP model, or the misspecification or absence of atmospheric

constituents (such as aerosols) observed by some satellite bands. In the context of clear-sky DA, biases can also be introduced by incomplete cloud screening procedures that allow some cloud-affected observations to pass quality control and thereby incorrectly enter the DA system. Indeed, most existing quality control methods were originally designed to remove all cloud-affected observations; however, these constraints are being relaxed as operational modeling centers move toward all-sky DA (e.g., Okamoto et al. 2014; Zhu et al. 2016). Exclusion of cloud-affected brightness temperatures has the undesirable consequence of removing observations that could have been used to improve the model initialization in cloudy areas of the model domain.

Additional uncertainties regarding the specification of cloud properties arise when assimilating cloud-affected infrared brightness temperatures. Though forward radiative transfer modeling for cloudy scenes has become more accurate in recent years, deficiencies remain, especially for ice clouds. Simulation of absorption and scattering properties for liquid clouds is relatively straightforward because the droplets are assumed to be spherical. However, there are larger uncertainties with ice cloud bulk optical properties because there is some dependence in the infrared on the shape of the ice particles (Yang et al. 2013). For example, an ice particle may take the form of a hexagonal plate, solid or hollow column, bullet rosette, or an aggregate of some form, and impact the bulk microphysical and optical properties that result from integration of the individual particle properties over the assumed size and habit distributions (Baum et al. 2014). In addition, the ice water path is related to both the cloud optical thickness and the cloud particle effective diameter. When computing simulated brightness temperatures, these diameters should be computed using the particle size distribution and cloud property assumptions made for each cloud species by a given microphysics scheme (e.g. Otkin et al. 2009; Cintineo et al. 2014; Thompson et al. 2016).

Biases in the OMB departures can also be caused by systematic errors in the NWP model forecasts that result from deficiencies in the parameterization schemes or other characteristics of the NWP model. It is well known that model forecasts containing large biases influence the behavior of BC methods and can degrade the performance of DA systems (Dee 2005; Dee and Uppala 2009; Eyre 2016). Biases can be especially large for model variables for which few observations are available to constrain their evolution, such as root zone soil temperature and moisture (Mahfouf 2010), or variables such as clouds and water vapor that are strongly influenced by parameterization schemes accounting for sub-grid scale processes. For example, uncertainties in microphysical parameters controlling cloud generation and decay processes can lead to systematic errors in the spatial extent, optical thickness, and height of the clouds, which in turn impacts the simulated

satellite brightness temperatures (Otkin and Greenwald 2008; Cintineo et al. 2014; Eikenberg et al. 2015). Ideally, a BC method would not remove the bias in the OMB departures associated with deficiencies in the NWP model because the observations should be used to correct such systematic errors. In the absence of a perfect reference analysis, however, it can be very difficult to determine whether a bias originates in the observations or forward radiative transfer model, both of which should be corrected, or in the model background (Dee 2005). Because of this uncertainty in bias attribution, all BC methods functionally act to correct the bias in the "observations" regardless of the true sources of the bias (Dee and Uppala 2009). Though this outcome is not desirable because it will limit the ability of the observations to reduce systematic errors in the analysis, it does satisfy the requirement by most DA methods that the observations are unbiased. In addition, the bias corrected observations can still be used to reduce random errors in the analysis. Eyre (2016) noted that the impact of model bias on the analysis accuracy depends on the rate at which the NWP model state relaxes back toward its own climatology after the assimilation update. If an NWP model quickly returns to its preferred state, then the analysis errors will continue to be large even if the model bias can be removed prior to computing the BC coefficients. This points toward the need to fix the bias at its source within the NWP model. The impact of model bias on a BC method can be reduced when high quality "anchor" observations with little or no bias are available; however, it is not apparent that such observations exist for water vapor and clouds.

BC methods can be broadly categorized into two types (Eyre 2016). The first type uses departures between the observations and their model equivalents accumulated over long time periods outside of the DA system to estimate and remove the bias from the observations prior to their assimilation. These so-called "static" BC methods typically use the satellite scan angle along with several atmospheric variables, such as the geopotential thickness over some layer, as the BC predictors. The BC coefficients for each satellite sensor and band are then computed using linear least squares regressions between the predictors and the observations. In practice, however, these "static" BC coefficients are regularly updated to account for changes in the model background due to changes in the NWP model or DA system, the addition of new observations, and upgrades to the forward radiative transfer model. Frequent retuning of a static BC method can be beneficial because it makes it more adaptable to changes in the models and observations. More detailed descriptions of static BC methods can be found in Eyre (1992), Harris and Kelly (2001), and Hilton et al. (2009).

With the second type of BC method, known as variational BC (VarBC), the BC coefficients are

updated simultaneously with the control vector during each DA cycle using the same set of observations and an augmented control vector (Derber et al. 1991; Parrish and Derber 1992; Derber and Wu 1998; Dee 2005; Auligne et al. 2007; Dee and Uppala 2009; Zhu et al. 2014). Like static BC methods, VarBC typically uses the satellite scan angle and several variables describing the atmospheric state as the predictors, with the total BC treated as a linear combination of all predictors. The BC coefficient for each predictor is computed during the minimization of the variational cost function. With an incremental DA approach with multiple outer loops, the BC coefficient increments evolve during each iteration of the inner loop and are updated at the end of each outer loop, which allows the coefficients to adjust with time and capture changes in observation quality. The state space augmentation approach used by VarBC also requires an estimate of the background covariances of the augmented state vector. For simplicity, most schemes assume that the error for a given BC parameter is uncorrelated with errors in other parameters for other satellite sensors and bands and with errors in the model background (Derber and Wu 1998; Dee 2005).

Most BC methods have been developed for use in variational or hybrid DA systems; however, several studies have also explored BC in ensemble DA systems. Fertig et al. (2009) developed a BC method for ensemble DA that is similar to VarBC in that it uses state augmentation to estimate the biases during the assimilation step. They showed that their method was able to reduce both the observation bias and the analysis error in perfect model experiments. Similar methods have also been used successfully in real data experiments assimilating microwave brightness temperatures (Szunyogh et al. 2008; Aravequia et al. 2011; Miyoshi et al. 2011). In high-resolution observing system simulation experiments assimilating infrared brightness temperatures, Cintineo et al. (2016) found that the analysis and forecast accuracy was improved when a simple fixed-value BC was applied to the clear-sky observations similar to that used by Stengel et al. (2009, 2013) in a variational DA system. Cintineo et al. (2016), however, did not bias-correct the cloudy observations prior to their assimilation because their bias was too complex to properly handle using a simple fixed-value BC applied uniformly to all cloudy observations. Zhu et al. (2016) handled biases in all-sky microwave observations by computing the BC coefficients using only cases where both the model background and the observations were either clear or cloudy. By doing this, they were able to reduce errors associated with mismatched cloud fields, while still preserving cloud-dependent information in the matched observations. Together, these results provide evidence that more sophisticated BC methods that can account for changes in cloud properties are necessary to effectively remove biases in the OMB departures.

In this study, we present a new BC method that can be used to diagnose and remove biases in all-sky infrared brightness temperatures using a Taylor series polynomial expansion of the OMB departures. This approach can diagnose both linear and nonlinear bias components through use of higher order Taylor series terms and a set of BC predictors. For example, with a 3rd order approximation, the 0th and 1st order terms represent the constant and linear bias components, whereas the 2nd (quadratic) and 3rd (cubic) order terms represent nonlinear bias components. We use this nonlinear BC (NBC) method to remove the bias from Scanning Enhanced Visible and Infrared Imager (SEVIRI) infrared brightness temperatures that were passively monitored during high-resolution ensemble DA experiments. The paper is organized as follows. The DA framework is described in Section 3.3, with a mathematical description of the NBC method presented in Section 3.4. Statistics obtained using the NBC method are shown in Section 3.5, with conclusions and a discussion presented in Section 3.6.

3.3 Experimental Design

3.3.1 SEVIRI Satellite Datasets

The SEVIRI sensor onboard the Meteosat Second Generation satellite provides accurate top-of-atmosphere radiance measurements across 12 visible and infrared spectral bands with a nadir resolution of 3 km for all infrared bands (Schmetz et al. 2002). The utility of the NBC method was evaluated using brightness temperatures from the 6.2 μm and 7.3 μm bands sensitive to WV over broad layers of the upper and middle troposphere, respectively, when skies are clear, while also being sensitive to clouds when they are present. Under clear conditions, the weighting functions that depict how much radiation from a given atmospheric height reaches the top of the atmosphere peak near 350 hPa (500 hPa) for the 6.2 μm (7.3 μm) bands, and then decrease to zero in the lower troposphere. When clouds are present, however, the weighting functions are truncated near the cloud top, which means that a larger portion of the top-of-atmosphere radiation originates at higher (e.g. colder) altitudes than would occur under clear-sky conditions. Their dual sensitivity to clouds and WV means that observations from these bands provide valuable information about the atmospheric state that is typically not available with conventional observations. Another motivation for using these bands is the expectation that their OMB departure statistics will be more Gaussian than would occur with infrared "window" bands because there will be a

smoother transition between the brightness temperatures in adjacent clear and cloudy areas.

Cloud top height retrievals made using SEVIRI observations were also obtained using software provided by the EUMETSAT Nowcasting Satellite Applications Facility and will be used as one of the BC predictors. The cloud top height for each satellite pixel was estimated by computing simulated clear-sky 10.8 μm brightness temperatures using the RTTOV radiative transfer model (Saunders et al. 1999) and temperature and humidity profiles from the global GME model (Majewski et al. 2002), and then inserting a cloud at successively higher levels until a best fit is obtained between the observed and simulated brightness temperatures (Derrien and Le Gleau 2005; Le Gleau 2016). To reduce the data volume and minimize the impact of spatially correlated errors in the observation departures, the cloud top height retrievals and SEVIRI brightness temperatures were horizontally thinned by a factor of 5 in the zonal and meridional directions. This reduces their horizontal resolution to ∼20-25 km across the model domain, and is ∼8 times coarser than the NWP model resolution. The cloud top height retrievals have a vertical resolution of 200 m; however, their uncertainty is larger, especially for semi-transparent clouds (Le Gleau 2016).

3.3.2 KENDA Data Assimilation System

Ensemble DA experiments in which conventional observations were actively assimilated and SEVIRI brightness temperatures were passively monitored were performed using the Kilometer-scale Ensemble Data Assimilation (KENDA) system (Schraff et al. 2016) developed by the Deutscher Wetterdienst (DWD). The KENDA system is based on the local ensemble transform Kalman filter method described by Hunt et al. (2007) and uses the Consortium for Small-scale Modeling (COSMO) model (Baldauf et al. 2011) as the NWP model. During this study, radiosonde, surface, wind profiler, and aircraft observations, were actively assimilated using a 1-h assimilation window, whereas SEVIRI 6.2 μm and 7.3 μm brightness temperatures were passively monitored. With KENDA, 4-D assimilation capabilities are obtained through inclusion of the observation operators within the COSMO model so that the model equivalents can be computed at the exact observation times during the forward integration of the ensemble. Temporally and spatially varying covariance inflation values are obtained at each grid point through a combination of multiplicative covariance inflation based on Anderson and Anderson (1999) and the relaxation to prior perturbations approach described by Zhang et al. (2004). Covariance localization is performed by updating the analysis at each grid point using only those observations located within a spec-

ified distance of the grid point. The vertical localization scale is fixed, but increases with height, whereas the horizontal scale is determined adaptively. For more detailed information about the KENDA system, the reader is referred to Schraff et al. (2016).

This study uses output from ensemble DA experiments that were performed on the COSMO-DE domain covering all of Germany and parts of surrounding countries with 2.8 km horizontal grid spacing. Lateral boundary conditions were obtained at hourly intervals from the 7-km resolution COSMO-EU domain run at the DWD, which in turn is driven by boundary conditions provided by the Icosahedral non-hydrostatic (ICON) model (Zangl et al. 2015). The COSMO-DE domain covers approximately 1200 x 1200 km and contains 50 vertical levels that are terrain-following in the lower troposphere and become horizontally flat in the upper troposphere and stratosphere. The model top is located at 22 km (i.e. about 40 hPa). The DA experiments employed 40 ensemble members along with a deterministic run that is initialized by applying the Kalman gain matrix from the assimilation update to the deterministic model background. The ensemble and deterministic runs were initialized at 00 UTC on 16 May 2014 and then updated at hourly intervals during a 5-day period ending at 00 UTC on 21 May 2014.

Atmospheric prognostic variables in the COSMO model include the horizontal and meridional wind components, temperature, pressure, and the mixing ratios for water vapor, cloud water, rainwater, pristine ice, snow, and graupel. Cloud microphysical processes, such as autoconversion, accretion, and self-collection, are represented using a simplified version of the Seifert and Beheng (2001) double-moment microphysics scheme that was reduced to a single-moment scheme for computational efficiency. Cloud formation and decay processes are parameterized based on the work of Lin et al. (1983). Heating rates due to radiative effects are updated at 15-min intervals using the δ-2-stream method developed by Ritter and Geleyn (1992). Deep convection is explicitly resolved whereas shallow convection is parameterized using a simplified version of the Tiedtke (1989) mass-flux scheme. A 2.5 order turbulent kinetic energy scheme developed by Raschendorfer (2001) is used to predict turbulence.

After an initial 12-h spin-up period, simulated SEVIRI brightness temperatures were generated for each ensemble member and the deterministic run at hourly intervals during a 4.5-day period from 13 UTC 16 May 2014 to 00 UTC 21 May 2014 using first-guess model output from 1-h COSMO-DE forecasts. The model profiles were interpolated to the thinned SEVIRI observation locations, and then simulated 6.2 μm and 7.3 μm brightness temperatures were computed using

version 10.2 of the RTTOV radiative transfer model (Saunders et al. 1999). RTTOV includes an enhanced cloud-scattering module that enables the use of cloud profiles located on the NWP model vertical grid (Matricardi 2005; Hocking et al. 2011). When computing cloudy brightness temperatures, RTTOV requires vertical profiles of liquid water content, ice water content, and fractional cloud cover. These quantities were computed using the COSMO model output and empirical relationships developed by Kostka et al. (2014). The default maximum-random cloud overlap scheme in RTTOV based on Raisanen (1998) was used during this study. RTTOV also includes several options to diagnose the ice particle effective diameters from the forecast ice water content based on relationships developed by Wyser (1998), Ou and Liou (1995), and McFarquhar et al. (2003) along with two ice crystal shape options (aggregates and randomly-oriented hexagonal crystals) that together are used to compute the ice radiative properties. For this study, we assume hexagonal ice crystals and compute the particle diameters using the McFarquhar et al. (2003) method. These settings were chosen because they provided the smallest overall bias during the 108-h study period based on six sensitivity experiments using the various ice crystal diameter and shape options. The mean brightness temperature for ice clouds between the best and worst options differed by approximately 1 K for the 6.2 μm band and 2.5 K for the 7.3 μm band during the entire study period (not shown), which illustrates the large uncertainty associated with the ice cloud property lookup tables in RTTOV.

3.4 Nonlinear Bias Correction (NBC) Method

Traditional BC methods remove biases between a given set of observed and model-equivalent satellite brightness temperatures through use of a set of BC predictors that describe the atmospheric state or characteristics of the satellite data. Both static and VarBC methods typically assume that a linear relationship exists between the departure bias and a given set of predictors or that a global constant can be added to the observations. This linear BC approach has been shown to work well for clear-sky observations possessing Gaussian error characteristics for which a set of constant and linear BC coefficients are sufficient to remove the bias; however, their use will be sub-optimal if the observation bias varies as a nonlinear function of some predictor. For satellite observations, nonlinear error dependencies are more likely to occur when cloudy observations are assimilated given the prevalence of nonlinear processes in clouds that could lead to complex errors in the forecast cloud field and the possibility that nonlinear error sources could be intro-

duced by the forward radiative transfer model used to compute the model-equivalent brightness temperatures. For example, with infrared brightness temperatures, it is possible that increased uncertainty simulating ice radiative properties in forward radiative transfer models could lead to biases that are a nonlinear function of some cloud property, such as cloud top height. Thus, given the increased interest in all-sky DA, it is desirable to develop BC methods that can remove both linear and nonlinear bias components from the innovations.

One method that can be used to account for nonlinear error dependencies in a set of observations is a Taylor series polynomial expansion that includes higher order terms that can capture nonlinear features of the error distribution if they exist. For a given set of observed and model-equivalent brightness temperatures corresponding to a specific satellite sensor and band, the observation departure vector is defined as:

$$\mathbf{dy} = \mathbf{y} - H(\mathbf{x}), \tag{3.1}$$

where $\mathbf{y}$ is the observation vector, $\mathbf{x}$ is the NWP model state vector, and $H(\mathbf{x})$ is the observation operator that is used to compute the model equivalent brightness temperatures. If we assume that the bias in the observation departures can be described by a real function $f(z)$ of a single variable (e.g., predictor) that is infinitely differentiable around a real number c, Eqn. 3.1 can be decomposed into an N order Taylor series expansion:

$$\mathbf{dy} = \left(f(c) + \frac{f'(c)(z^{(i)} - c)}{1!} + \frac{f''(c)(z^{(i)} - c)^2}{2!} + \frac{f'''(c)(z^{(i)} - c)^3}{3!} + ... + \frac{f^{(n)}(c)(z^{(i)} - c)^n}{n!} \right)_{i=1,...,m} \tag{3.2}$$

where $\mathbf{dy}$ is the m x 1 observation departure vector and m is the number of observations, $f^{(n)}(c)$ is the nth derivative of f evaluated at the point c, and $z^{(i)}$ is the predictor value for the ith observation. The $i = 1, ..., m$ notation outside the parentheses indicates that the Taylor series approximation is computed separately for each element of the $\mathbf{dy}$ vector using the equation within the parentheses. The variable used as the predictor is chosen based on its ability to capture some aspect of the observation departure bias, whereas the value $z^{(i)}$ of that variable for a given observation can be obtained from a variety of sources, such as the model background or a satellite retrieval. The constant c can be set to any value because $c + \delta c$ simply moves c to another constant value; therefore,

for convenience, we define c to be the mean of the predictor values:

$$c = \frac{\sum_{i=1}^{m} z^{(i)}}{m} \tag{3.3}$$

It is readily apparent from Eqn. 3.2 that the higher order terms represent nonlinear components because the exponents are ≥ 2, with the $(z-c)^2$ and $(z-c)^3$ polynomials representing the quadratic and cubic terms, respectively.

The single variable case shown in Eqn. 3.2 can subsequently be generalized to be a function of more than one predictor:

$$\begin{aligned}\mathbf{dy} = \Bigg(& f(a_1,...,a_d) + \sum_{j=1}^{d} \frac{\partial f(a_1,...,a_d)}{\partial x_j}(x_j^{(i)} - a_j) \\ & + \frac{1}{2!}\sum_{j=1}^{d}\sum_{k=1}^{d} \frac{\partial^2 f(a_1,...,a_d)}{\partial x_j \partial x_k}(x_j^{(i)} - a_j)(x_k^{(i)} - a_k) \\ + \frac{1}{3!}\sum_{j=1}^{d}\sum_{k=1}^{d}\sum_{l=1}^{d} & \frac{\partial^3 f(a_1,\cdots,a_d)}{\partial x_j \partial x_k \partial x_l}(x_j^{(i)} - a_j)(x_k^{(i)} - a_k)(x_l^{(i)} - a_l) + ... \Bigg)_{i=1,...,m}\end{aligned} \tag{3.4}$$

which can be written more compactly as:

$$\mathbf{dy} = \left(\sum_{n_1=0}^{d} \cdots \sum_{n_d=0}^{d} \left(\frac{\partial^{(n_1+\cdots+n_d)} f}{\partial x_1^{n_1} \cdots \partial x_d^{n_d}} \right)(a_1,\cdots,a_d) \frac{(x_1^{(i)} - a_1)^{n_1} \cdots (x_d^{(i)} - a_d)^{n_d}}{n_1! \cdots n_d!} \right)_{i=1,...,m}, \tag{3.5}$$

where d is the number of predictors, $f^{(n_d)}(a_d)$ denotes the nth partial derivative of f evaluated at the point a_d, and $x_d^{(i)}$ is the ith value for a given predictor x_d.

For illustrative purposes, if we assume a single variable, third order Taylor series expansion for a single satellite sensor and band, and define the BC coefficients such that $b_n = \frac{f^{(n)}(a)}{n!}$, Eqn. 3.2 can be written as:

$$\mathbf{dy} = \left(b_0 + b_1(z^{(i)} - c) + b_2(z^{(i)} - c)^2 + b_3(z^{(i)} - c)^3 \right)_{i=1,...,m} \tag{3.6}$$

or alternatively in matrix notation as:

$$\mathbf{dy} = \mathbf{Ab} \tag{3.7}$$

where **dy** is the m x 1 observation departure vector, **A** is an m x n matrix containing the n Taylor series terms $(z^{(i)} - c)^l$ for each ith observation, where $l = 0, ..., n-1$, and **b** is an n x 1 vector containing the BC coefficients. This is an overdetermined system of m linear equations in n unknown coeffi-

cients because $m > n$. The first column of $\mathbf{A}$ contains ones, with the remaining columns containing the linear and higher order Taylor series terms. Because this kind of system typically does not have an analytic solution, we instead want to find the coefficients $\mathbf{b}$ that best fit the equations by solving the quadratic minimization problem $\hat{b} = \min_b S(b)$, where the objective function S is given by:

$$S(b) = \sum_{i=1}^{m} |dy_i - \sum_{j=1}^{n} A_{ij} b_j|^2 = \|\mathbf{dy} - \mathbf{Ab}\|^2 \tag{3.8}$$

and $\|\cdot\|$ is the Euclidean norm. Because most real-world phenomena act as a low pass filter in the forward direction where $\mathbf{A}$ maps $\mathbf{b}$ to $\mathbf{dy}$, the inverse mapping will operate as a high-pass filter that amplifies noise and can therefore lead to a poorly conditioned problem. Preference, however, can be given to smaller norms by adding a Tikhonov regularization term, $\|\Gamma\mathbf{b}\|^2$, to Eqn. 3.8, which is a standard approach when solving inverse problems (Nakamura and Potthast, 2015). For simplicity, we choose a matrix that is a multiple of the identity matrix ($\Gamma = \alpha I$), such that:

$$\hat{S}(b) = \|\mathbf{dy} - \mathbf{Ab}\|^2 + \alpha \|I\mathbf{b}\|^2 \tag{3.9}$$

Sensitivity tests showed that α could be set to a very small value (10^{-9}) when one variable was used in the regression; however, a slightly larger value (10^{-6}) was found to work better for the multivariate regressions. These values were used for the univariate and multivariate experiments presented in Section 3.5. The least squares solution can then be found by differentiating $\hat{S}$ with respect to b, and equating to 0, such that:

$$\frac{\partial \hat{S}}{\partial b} = \mathbf{A}^{\mathrm{T}}\mathbf{dy} - (\alpha I + \mathbf{A}^{\mathrm{T}}\mathbf{A})\mathbf{b} = 0, \tag{3.10}$$

or alternatively, after rearranging and multiplying both sides of Eqn. 3.10 by $(\alpha I + \mathbf{A}^{\mathrm{T}}\mathbf{A})^{-1}$, we can solve for the b vector containing the BC coefficients using:

$$\mathbf{b} = (\alpha I + \mathbf{A}^{\mathrm{T}}\mathbf{A})^{-1}\mathbf{A}^{\mathrm{T}}\mathbf{dy} \tag{3.11}$$

where $(\alpha I + \mathbf{A}^{\mathrm{T}}\mathbf{A})$ is a symmetric, square matrix with dimensions n x n. The small dimensions of this matrix make it easy to compute its inverse, thereby making it feasible to include higher order Taylor series terms, additional predictors, and a large OMB departure dataset when computing the BC coefficients. After solving for $\mathbf{b}$, which is done separately for each satellite band and sensor, the BC coefficients can then be applied to $\mathbf{dy}$ to remove the linear and nonlinear conditional bias

components from the observations.

3.5 Results

In this section, the ability of the NBC method to remove biases from all-sky satellite infrared brightness temperatures is assessed using OMB departure statistics accumulated at hourly intervals during a 4.5 day period in which conventional observations were actively assimilated and SEVIRI observations were passively monitored. Figure 3.1 shows the evolution of the observed SEVIRI 6.2 μm brightness temperatures during this time period. At the start of the period on 16 May (Fig. 3.1a), an area of cold upper level clouds associated with a band of precipitation was located across the eastern half of the domain. This weather feature slowly weakened over Germany during the next two days (Fig. 3.1b, c), with the brightness temperatures becoming warmer as the convective clouds were replaced by cirrus and mid-level clouds. Generally clear skies characterized by warm brightness temperatures were also present across parts of the domain during this time period, with clear skies prevailing across most of the region on 19 May (Fig. 3.1d). A large area of convection with very cold upper-level clouds then moved into the western half of the domain on 20 May (Fig. 3.1e). Overall, it is evident that the study period contains a wide range of atmospheric conditions and cloud types that supports a realistic assessment of the NBC method during the warm season.

3.5.1 Univariate Bias Correction Results

To explore the ability of individual predictors to remove the bias from all-sky infrared observations, univariate NBC experiments were performed using the satellite zenith angle and various predictors sensitive to clouds and WV, such as the brightness temperature, cloud top height, and integrated water content over some vertical layer. This section presents results from a subset of these experiments that remove the bias from all-sky SEVIRI 6.2 μm observations. The impact of each predictor is assessed using OMB departure distributions normalized by the standard deviation in a given sample and with 2-D probability distributions of the departures plotted as a function of a given predictor. The results are evaluated separately for the original departure distribution and for distributions for which the bias has been removed using either a 0th (constant), 1st (linear), 2nd (quadratic), or 3rd (cubic) order Taylor series polynomial expansion.

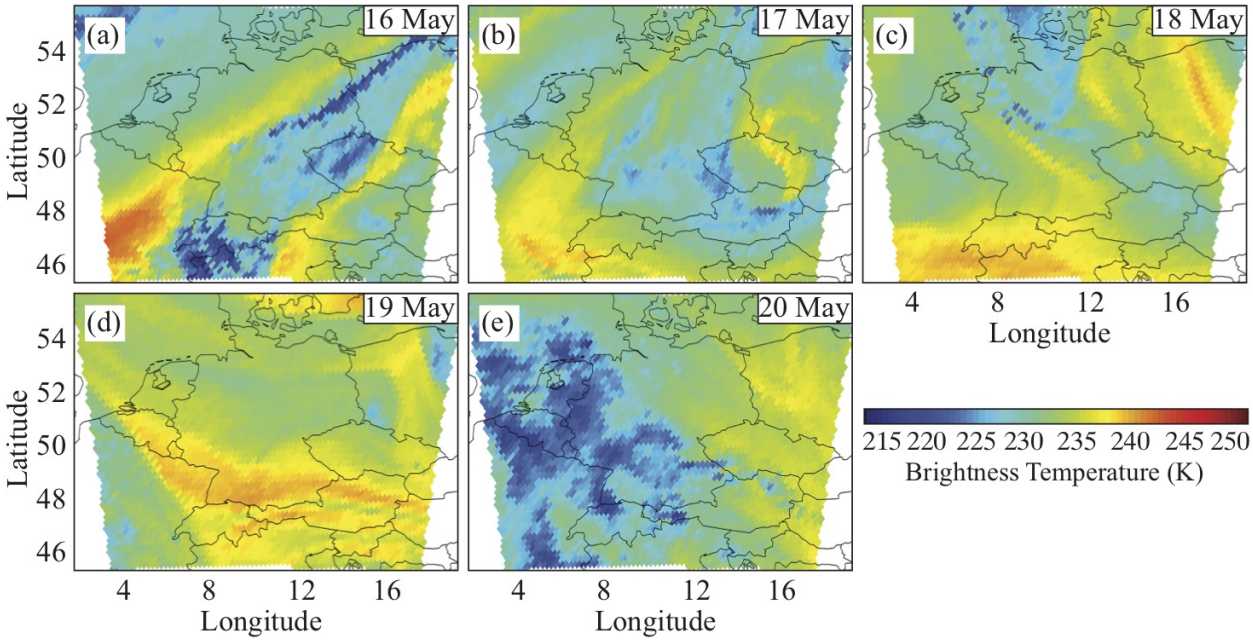

Figure 3.1: Observed SEVIRI 6.2 μm brightness temperatures (K) valid at 18 UTC on (a) 16 May, (b) 17 May, (c) 18 May, (d) 19 May, and (e) 20 May 2014.

3.5.1.1 Observed Brightness Temperature Predictor

As shown by the probability distributions in Fig. 3.2, the observed 6.2 μm brightness temperatures are an excellent predictor of their own bias, especially when higher order Taylor series terms are used. The horizontal magenta line in each panel depicts the mean bias of the entire distribution, whereas the shorter horizontal black lines depict the conditional bias in each column and will be used to assess how the bias varies as a function of the predictor value. This terminology is being used to differentiate biases conditioned on the predictor value from the bias of the overall distribution. For example, though each distribution except for the original distribution will have zero overall bias, this obscures the fact that the conditional bias could potentially vary as a function of the predictor value. Inspection of Fig. 3.2a reveals a nonlinear pattern in the conditional biases, with a tendency for the simulated brightness temperatures to be too warm (cold) when the observed brightness temperatures are colder (warmer) than 235 K. Though the mean bias of the distribution is relatively small (-0.83 K), the nonlinear pattern in the conditional biases means that constant and linear BC terms alone will be unable to remove all of the bias. For example, even though the constant BC term removes the mean bias from the distribution (Fig. 3.2b), its shape remains the same and therefore large conditional biases remain throughout the distribution. Likewise, the 1st order BC term removes the linear departure component by raising

(lowering) the cold (warm) end of the distribution, which reduces the conditional biases for the coldest brightness temperatures, but turns a positive bias into a negative bias for the warmest brightness temperatures (Fig. 3.2c). Removal of the constant and linear bias components exposes an asymmetric arch shape in the conditional biases that is largely removed when the 2nd order quadratic term is used (Fig. 3.2d), except for nonzero biases that remain at the cold and warm ends of the distribution. Finally, when the 3rd order cubic term is used, the general shape of the distribution is unchanged; however, it is evident that subtle improvements were made to it given that most of the conditional biases are now close to zero. Together, these results show that even though each BC distribution has zero mean bias, that the conditional biases in the distribution are much smaller when the higher order, nonlinear BC terms are applied to the observations.

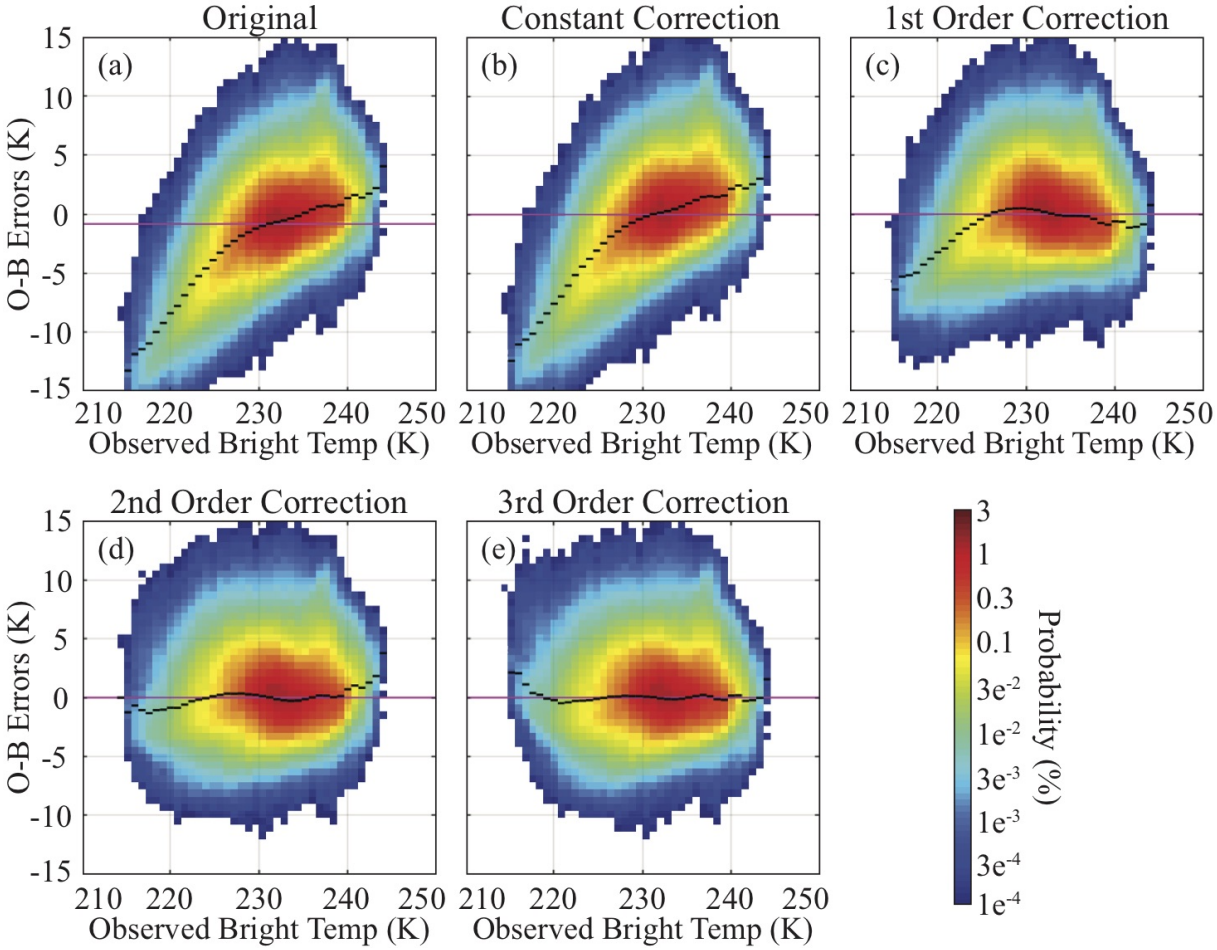

Figure 3.2: Probability distributions of 6.2 μm observation-minus-background departures plotted as a function of the observed 6.2 μm brightness temperatures (K) for the (a) original data, and the (b) constant, (c) 1st order, (d) 2nd order, and (e) 3rd order bias corrected observations when the observed 6.2 μm brightness temperature is used as the predictor. The horizontal black line segments represent the conditional bias in each column, whereas the horizontal magenta line depicts the mean bias of the entire distribution. Data were accumulated at hourly intervals during a 108-h period from 13 UTC on 16 May 2014 to 00 UTC on 20 May 2014.

Normalized OMB departure histograms computed using the original observations and the constant, 1st, 2nd, and 3rd order BC observations are shown in Fig. 3.3a-e. Each histogram is normalized based on its variance, with the curved red line on each panel representing a Gaussian distribution with zero mean and a variance equal to that of the sample. Overall, the variance and root mean square error (RMSE) are greatly reduced when the 1st order BC coefficients are applied to the observations (Fig. 3.3c), which is primarily due to the smaller departures for the colder brightness temperatures (e.g. Fig. 3.2c). The variance was further reduced when the 2nd order BC was used, with only minimal changes occurring when this was expanded to a 3rd order BC (Figs. 3.3d, e). The fact that the higher order terms only had a small impact on these statistics while simultaneously having a large positive impact on the conditional biases in Fig. 3.2 illustrates that more detailed analysis methods such as 2-D probability distributions can provide additional insight into the characteristics of the OMB departure distributions. Comparison of the histograms also shows that the negative skewness in the original distribution (Fig. 3.3a) changes to positive skewness after the BC terms are applied. This behavior primarily results from a conditional positive skewness for brightness temperatures < 230 K that is evident in Fig. 3.2a by the tendency for the conditional bias in each column to be located above the bin with the maximum probability. Because the same BC is applied to a given brightness temperature regardless of its OMB departure, the positive skewness in the conditional distributions is preserved as they are shifted upward, thereby leading to a positive skewness in the full BC distributions.

3.5.1.2 Cloud Top Height Predictor

Because infrared observations are very sensitive to the vertical distribution of clouds, an experiment was performed using the NWC SAF cloud top height retrievals as the BC predictor to better isolate the impact of clouds. To provide complete domain coverage, the clear-sky observations were assigned a height equal to the model terrain elevation. Overall, the conditional biases in the original distribution (Fig. 3.4a) are close to zero for cloud top heights < 7 km; however, the biases increase for clouds above this level and peak near -6 K for cloud top heights > 10 km. This is a complex error pattern that a constant BC scheme is unable to fix (Fig. 3.4b). Indeed, the upward shift of the distribution to remove the mean bias actually worsens the conditional biases for cloud top heights < 7 km, while leading to only minor improvements for the upper-level clouds. The linear correction (Fig. 3.4c) slightly improves the conditional biases for lower and upper-level

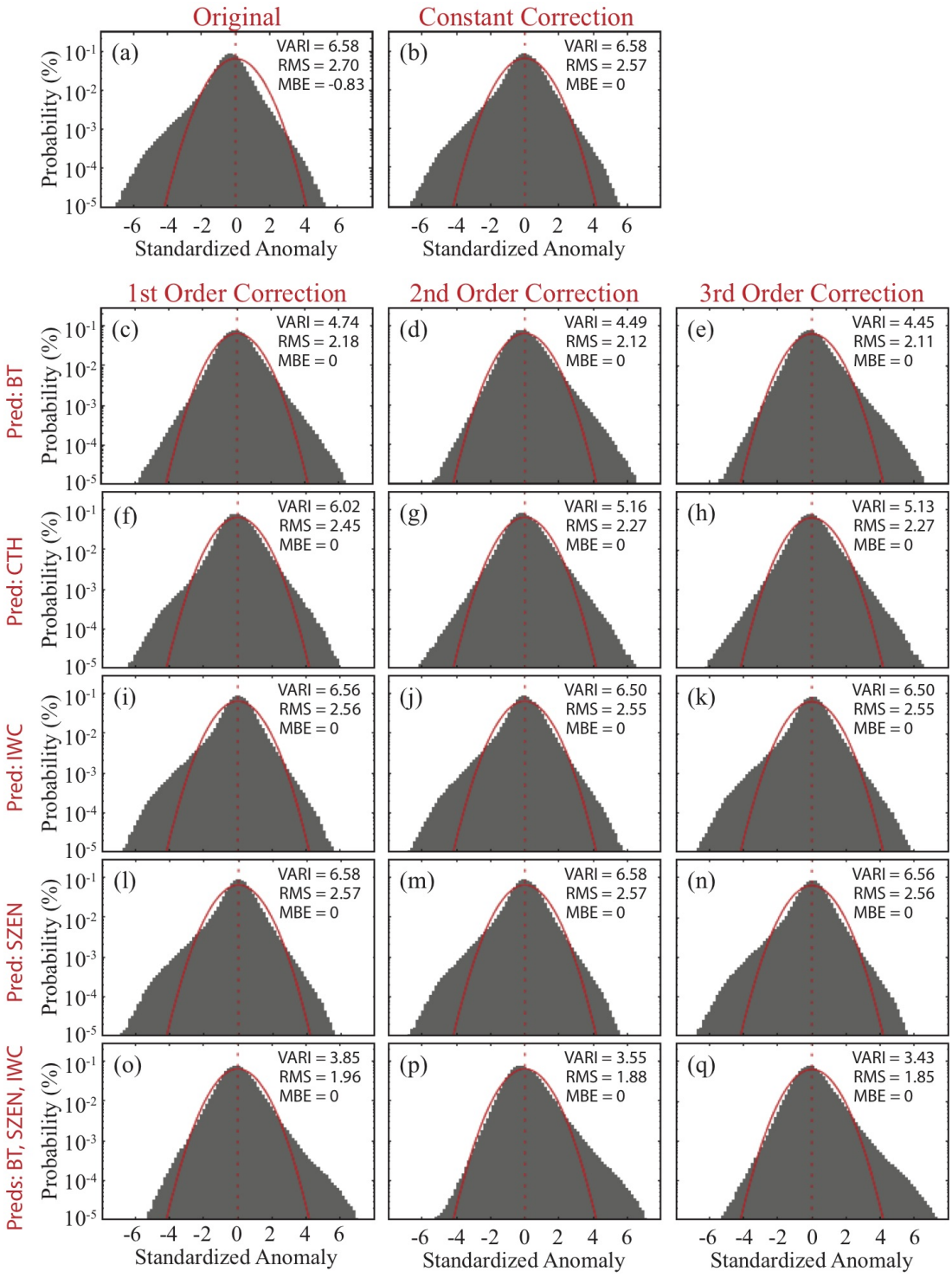

Figure 3.3: Probability density function of normalized 6.2 μm observation-minus-background departures for the (a) original and (b) constant bias correction distributions. The corresponding 1st, 2nd, and 3rd order bias correction error distributions when the (c-e) observed 6.2 μm brightness temperatures, (f-h), NWC SAF cloud top heights, (i-k) model-simulated total integrated water content (IWC) in the 100-700 hPa layer, (l-n) satellite zenith angle, or (o-q) observed 6.2 μm brightness temperatures, satellite zenith angle, and IWC are used as the predictors are also shown. Data were accumulated at hourly intervals during a 108-h period from 13 UTC on 16 May 2014 to 00 UTC on 20 May 2014.

clouds, but worsens the bias for mid-level clouds, which together slightly reduces the variance in the overall distribution (Fig. 3.3f). Use of the 2nd order quadratic term substantially improves the distribution by removing the arch in the conditional bias pattern by decreasing the magnitude of the positive (negative) OMB departures for cloud tops located in the middle (upper) troposphere (Fig. 3.4d). These changes resulted in a much smaller variance in the histogram (Fig. 3.3g). As was the case in the previous section, the 3rd order BC led to slightly smaller conditional biases across most of the distribution (Fig. 3.4e), but had minimal impact on the statistics of the overall distribution (Fig. 3.3h). Though the cloud top height predictor was unable to reduce the variance of the full distribution as much as the brightness temperature predictor did, the NBC method was still able to greatly improve the distribution by decreasing the conditional biases. Its use also led to a more symmetric OMB departure distribution (Fig. 3.3h). These results show that cloud top height information can be used to remove the bias from all-sky infrared observations if higher order Taylor series terms are used.

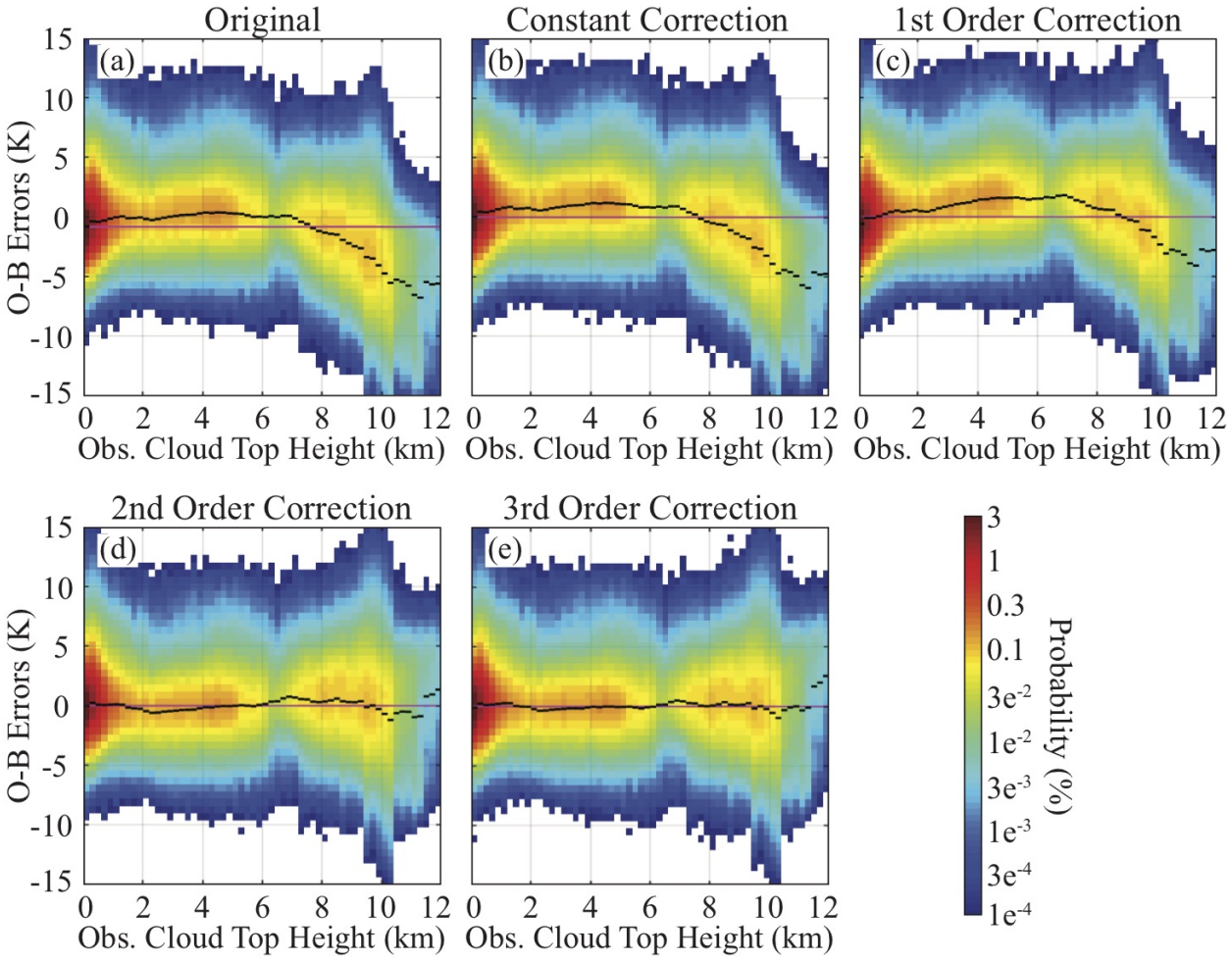

Figure 3.4: Same as Fig. 3.2 except for showing probability distributions plotted as a function of the NWC SAF cloud top height retrieval (km) when this quantity is also used as the BC predictor.

3.5.1.3 Vertically-Integrated Water Content Predictor

In this section, the impact of using a BC predictor that depicts the total water content over a vertical layer is assessed. Numerous experiments were performed using different vertical layers; however, for brevity, results are only shown for the predictor that encapsulates the total water content between 100 and 700 hPa because that is the portion of the atmosphere where 6.2 μm brightness temperatures are most sensitive. Unlike the previous predictors, this predictor is computed using model output. The total water content is calculated for each ensemble member by converting the WV and all cloud hydrometeor mixing ratios in each model layer into mm and then integrating over the 100-700 hPa layer. Inspection of Fig. 3.5a shows that this predictor has a less complex OMB departure pattern than occurred when the cloud top height and brightness temperatures were used as the predictors. There are however slightly larger biases on both ends of the distribution, with a small upward slope in the maximum probabilities as the total water content increases. This linear error trend is removed by the linear bias correction term (Fig. 3.5c), which reduces the conditional biases when the total water content is < 7 mm, but increases it elsewhere. The subtle arch in the conditional biases is subsequently removed after applying the 2nd order quadratic term (Fig. 3.5d), with only minor changes occurring after the 3rd order term is used (Fig. 3.5e). Comparison of the histograms (Figs. 3.3i-k) shows that the total water predictor had only a small impact on the variance of the full distribution; however, the scatterplots showed that it still improved the conditional bias across most of the distribution. Even so, this predictor still had a much smaller impact than the previous predictors that were directly sensitive to the cloud top height, which indicates that the location of the cloud top rather than the vertically integrated cloud and WV content is a more effective BC predictor for all-sky infrared brightness temperatures.

3.5.1.4 Satellite Zenith Angle Predictor

Given that the satellite zenith angle is widely used in operational BC methods, an additional experiment was performed using it as the BC predictor. After adjusting for the mean bias in the original distribution, the conditional biases are close to zero across the entire distribution, with only a slight downward trend in the bias for zenith angles $> 48^\circ$ (3.6b). Application of the 1st to 3rd order BC terms (Figs. 3.6c-e) eliminated most of these conditional biases; however, the impact of this predictor on the statistics of the entire distribution was negligible according to the his-

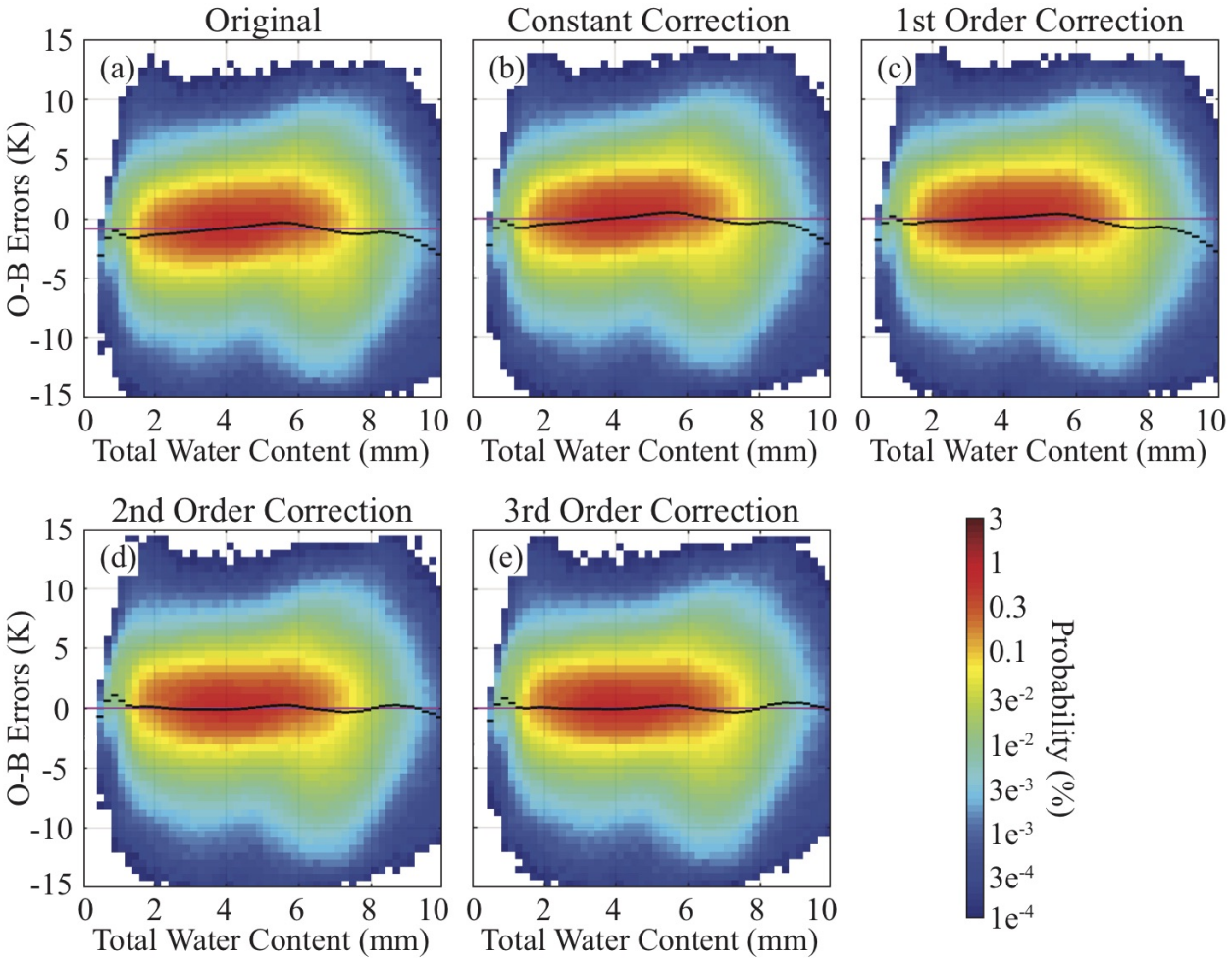

Figure 3.5: Same as Fig. 3.2 except for showing probability distributions plotted as a function of the vertically-integrated total water content (mm) over the 100-700 hPa layer when this quantity is also used as the BC predictor.

tograms (Figs. 3.3l-n). These results indicate that the bias in the observations is only very weakly related to the satellite zenith angle; however, the small improvements made to the conditional biases by the 2nd to 3rd order terms also show that there is a small nonlinear bias component that can be removed when using this predictor.

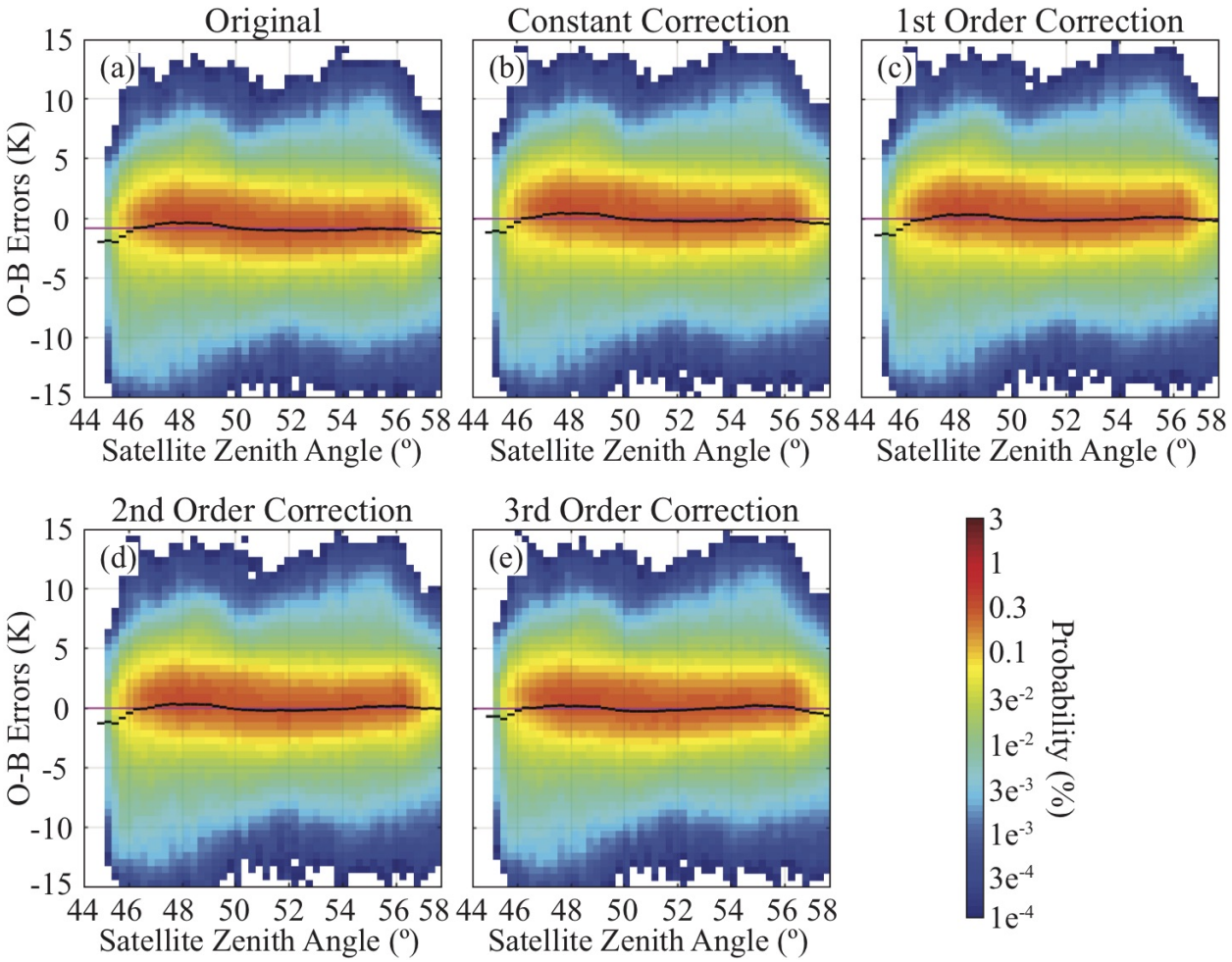

Figure 3.6: Same as Fig. 3.2 except for showing probability distributions plotted as a function of the satellite zenith angle (o) when this quantity is also used as the BC predictor.

3.5.2 Clear and Cloudy Sky Error Evaluation

Next, the relative impact of the linear and nonlinear BC terms on the clear and cloudy-sky observations is examined more closely using a subset of the 6.2 μm brightness temperatures for which both the model background and a given observation were identified as being clear or cloudy. Each observation was classified as clear or cloudy based on the NWC SAF cloud mask dataset whereas each model grid point was deemed to be clear (cloudy) if the sum of all cloud hydrometeor mixing ratios over the entire vertical profile was less (greater) than 10^{-6} kg kg^{-1}. The 2-D probability

distributions for the clear-sky matched observations are shown in Fig. 3.7, with the corresponding histograms shown in Fig. 3.8. The observed 6.2 μm brightness temperatures were used as the BC predictor. Inspection of Fig. 3.7a reveals that the original distribution contains both a systematic bias and a large linear trend where mostly negative OMB departures for the colder brightness temperatures transition into mostly positive departures for the warmer brightness temperatures. The linear trend indicates that the WV field in the model background is more uniform than observed such that the model tends to be too wet (dry) in regions where the observations indicate less (more) WV. Overall, most of the bias is removed from the clear-sky observation departures using only the constant and 1st order terms, with little or no impact due to the higher order terms (Figs. 3.7b-e). This behavior is consistent with existing BC schemes that use constant and linear corrections to remove the bias from clear-sky observation departures.

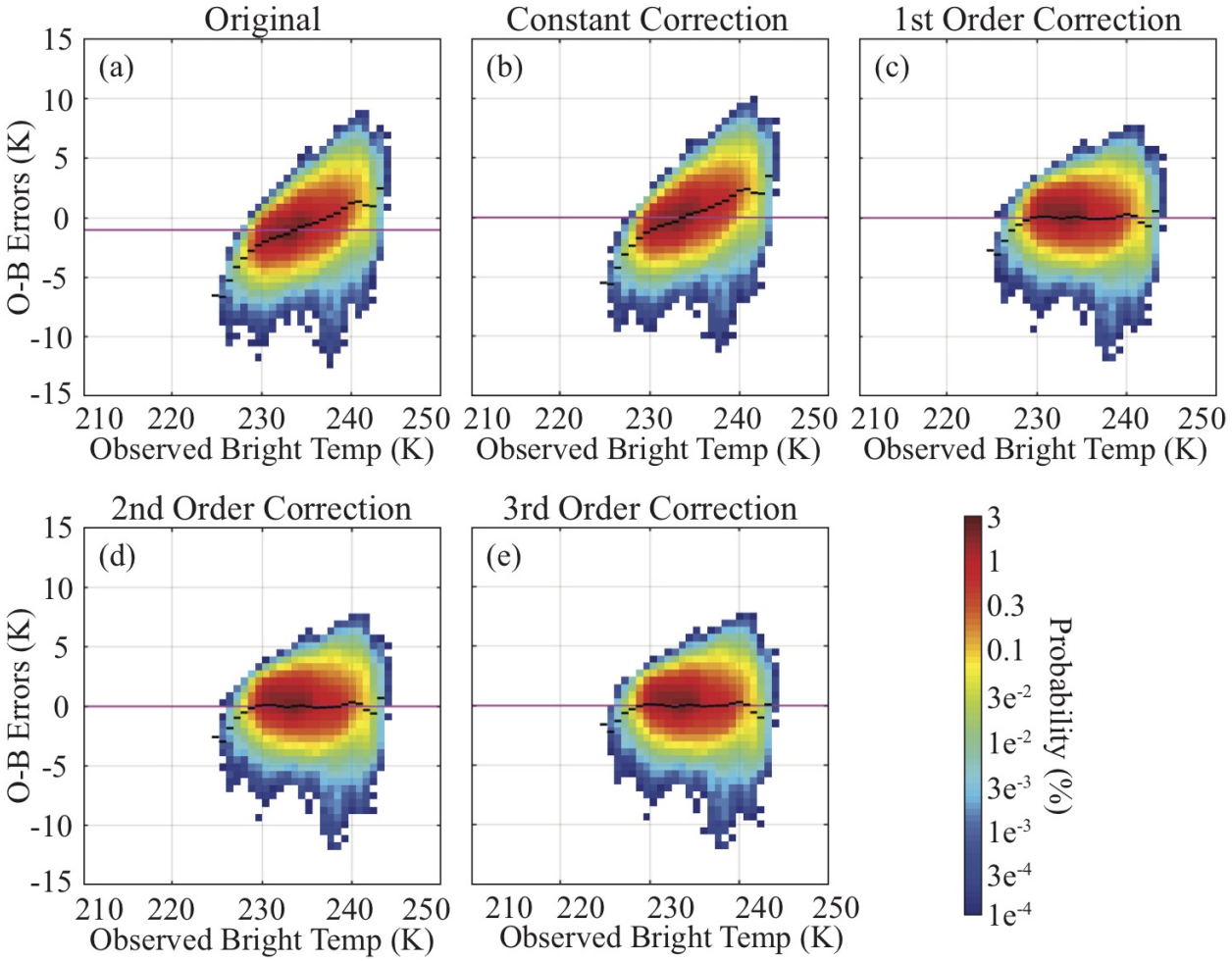

Figure 3.7: Same as Fig. 3.2 except for showing probability distributions for clear-sky matched observations plotted as a function of the observed brightness temperature (K) when this quantity is also used as the BC predictor.

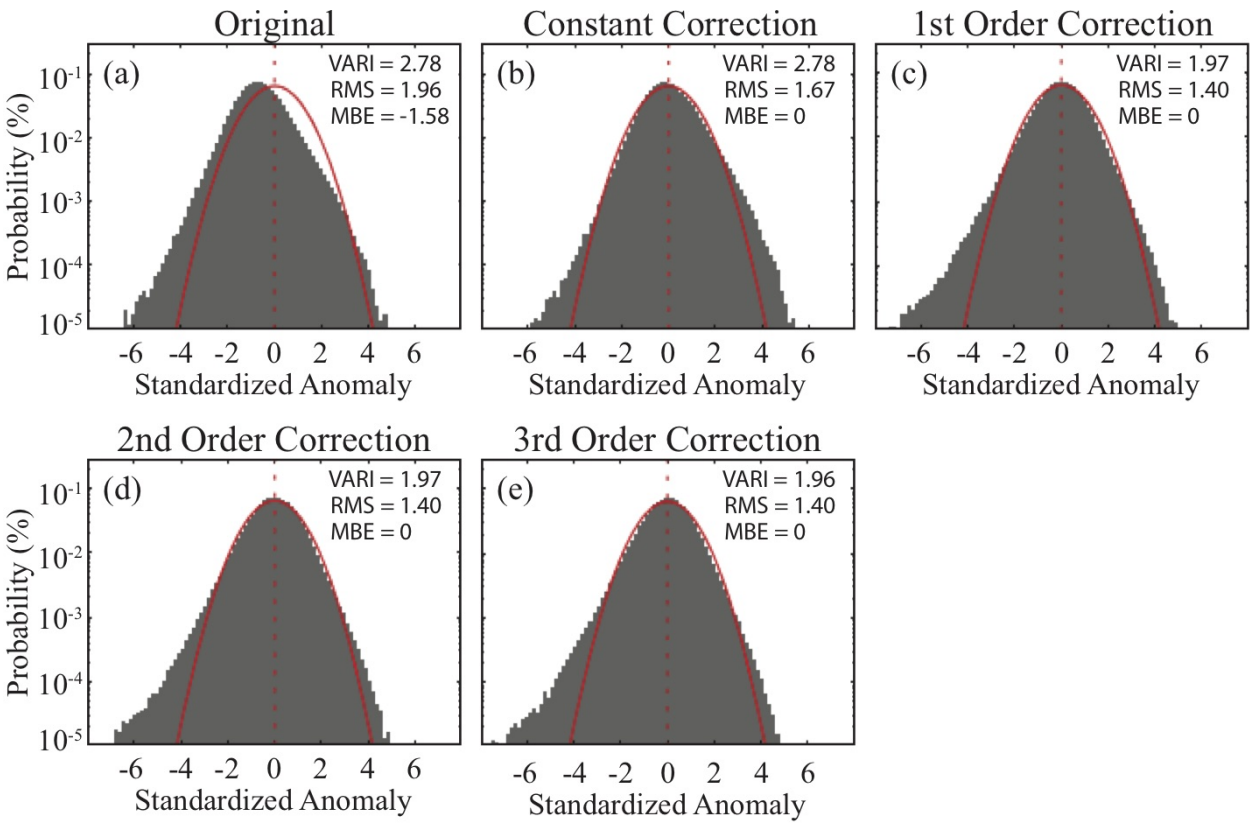

Figure 3.8: Probability density function of normalized clear-sky matched 6.2 μm observation-minus-background departures for the (a) original data, and the (b) constant, (c) 1st order, (d) 2nd order, and (e) 3rd order bias corrected observations when the observed 6.2 μm brightness temperature is used as the predictor. Data were accumulated at hourly intervals during a 108-h period from 13 UTC on 16 May 2014 to 00 UTC on 20 May 2014.

For the cloud-matched observations shown in Figs. 3.9 and 3.10, the NWC SAF cloud top height retrievals were used as the predictor. The OMB departure pattern and conditional biases for these observations are very similar to that shown in Fig. 3.4 when both clear and cloudy-sky observations were included in the regression. This includes the generally positive departures for mid-level clouds and the transition to large negative departures for the upper-level clouds (Fig. 3.9a). Large departures remained in the distribution for all cloud top heights after the constant and linear BC terms were applied to the observations (Fig. 3.9c). It is only when the 2nd and 3rd order terms are used that the conditional biases become close to zero throughout the entire distribution (Figs. 3.9d, e). The histograms in Fig. 3.10 also reveal that the quadratic and cubic terms had a much larger impact on the overall statistics than occurred for the clear-sky matched observations. These results provide further evidence that the nonlinear conditional biases evident in the all-sky scatterplots in Section 3.5.1 primarily result from biases associated with the cloudy observations. It also shows that the NBC method is an effective method to remove both linear

and nonlinear biases from all-sky infrared brightness temperature departures if a suitable cloud-sensitive variable is used as the predictor.

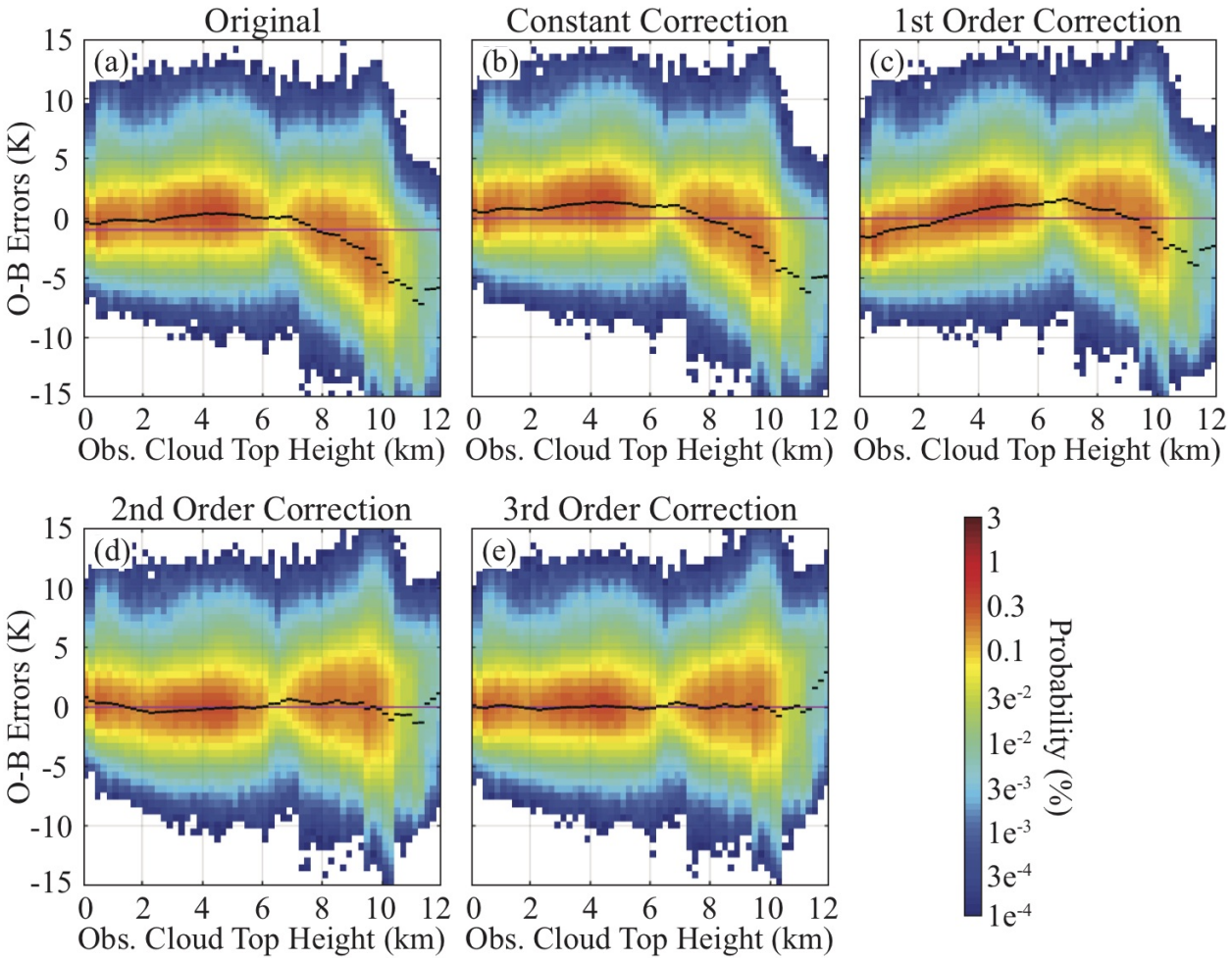

Figure 3.9: Same as Fig. 3.2 except for showing probability distributions for cloudy-sky matched observations plotted as a function of the NWC SAF cloud top height retrieval (km) when this quantity is also used as the BC predictor.

3.5.3 Multivariate Bias Correction Results

In addition to the univariate NBC experiments discussed in previous sections, multivariate experiments were performed to assess the impact of using more than one predictor to remove the observation bias. For a 3rd order polynomial expansion using two variables, it is necessary to solve for seven coefficients in Eqn. 3.11, whereas 22 coefficients are computed when three predictors are used. Because a direct approach is used to simultaneously estimate all of the BC coefficients, it is not possible to determine the individual contribution of each predictor on the OMB departures; however, the total contribution of all of the predictors within a given Taylor series order (e.g., 1st,

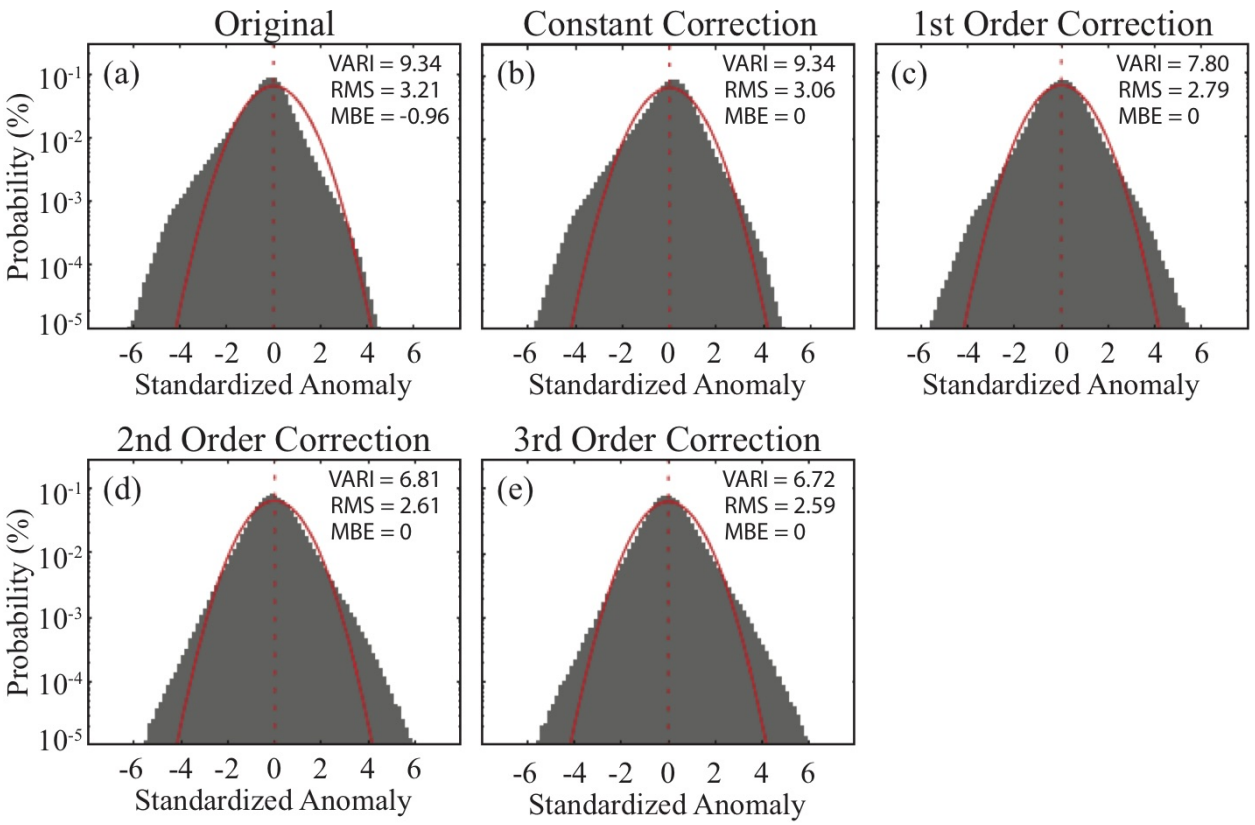

Figure 3.10: Probability density function of normalized cloudy-sky matched 6.2 μm observation-minus-background departures for the (a) original data, and the (b) constant, (c) 1st order, (d) 2nd order, and (e) 3rd order bias corrected observations when the NWC SAF cloud top height retrieval is used as the predictor. Data were accumulated at hourly intervals during a 108-h period from 13 UTC on 16 May 2014 to 00 UTC on 20 May 2014.

2nd, and 3rd) can still be inferred through comparison of the results obtained using different order expansions. Though using more than one variable greatly increases the size of the **A** matrix, it is still computationally efficient to solve for the inverse of $\mathbf{A}^T\mathbf{A}$ given its small dimensions.

Numerous experiments using different predictor combinations and a 2nd or 3rd order polynomial expansion were performed; however, for brevity, this section only includes results from the combination that had the largest impact on the OMB departure distributions. This particular configuration employed a 3rd order expansion with the satellite zenith angle, 100-700 hPa total water content, and observed brightness temperatures for a given satellite band used as the BC predictors for that band. A separate multi-variate experiment (not shown) that employed the cloud top height rather than the brightness temperature as the third predictor revealed that it had a smaller impact, similar to what occurred with the univariate experiments shown earlier. There may be some overlap between the brightness temperature and satellite zenith angle predictors; however, this should be minimal because the zenith angle predictor primarily accounts for potential biases

in the radiative transfer model associated with the path length through the atmosphere, whereas the brightness temperature predictor is being used as a proxy for the cloud top height given its strong sensitivity to the cloud top. Unlike the previous sections that focused exclusively on the 6.2 μm band, this section presents results from experiments that removed the bias from both of the SEVIRI WV-sensitive bands (e.g., 6.2 μm and 7.3 μm). All observations, both clear and cloudy-sky, were used during these experiments.

3.5.3.1 SEVIRI 6.2 μm Example

Figure 11 shows the OMB departure distributions for the 6.2 μm multivariate NBC experiment, with the corresponding normalized histograms shown in Figs. 3.3o-q. Comparison to Fig. 3.2 shows that the departure distributions for the multivariate case are similar to those from the univariate case employing only the observed brightness temperature as the BC predictor. This is not surprising given that the experiments employing the satellite zenith angle and total water content predictors both had a much smaller impact on the distributions (Figs. 3.5, 3.6). Overall, the shape of the distribution is improved after the linear term is used; however, there are still large conditional biases at both ends of the distribution (Fig. 3.11c). The arch pattern in the conditional bias was subsequently removed after the quadratic term was applied (Fig. 3.11d), with slightly smaller (larger) biases occurring at the warm (cold) end of the distribution after using the 3rd order cubic term (Fig. 3.11e). Though the distributions are similar to those shown in Fig. 3.2, it is evident that the width of the conditional distribution is less for all predictor values. This is encouraging because it shows that even though the impact of the satellite zenith angle and total water content predictors was relatively small when used individually, they still provided new information that further reduced the OMB departures when used in combination with the observed brightness temperature predictor. Inspection of the histograms (Figs. 3.3o-q) shows that the variance was greatly reduced compared to the univariate experiments; however, each of the distributions had a large positive skewness similar to that seen in Figs. 3.3c-e when the brightness temperature was used as the BC predictor. It is important to note however that quality control measures could potentially be used to reduce the skewness in the distribution after the BC terms are applied. This topic will be explored in a future study.

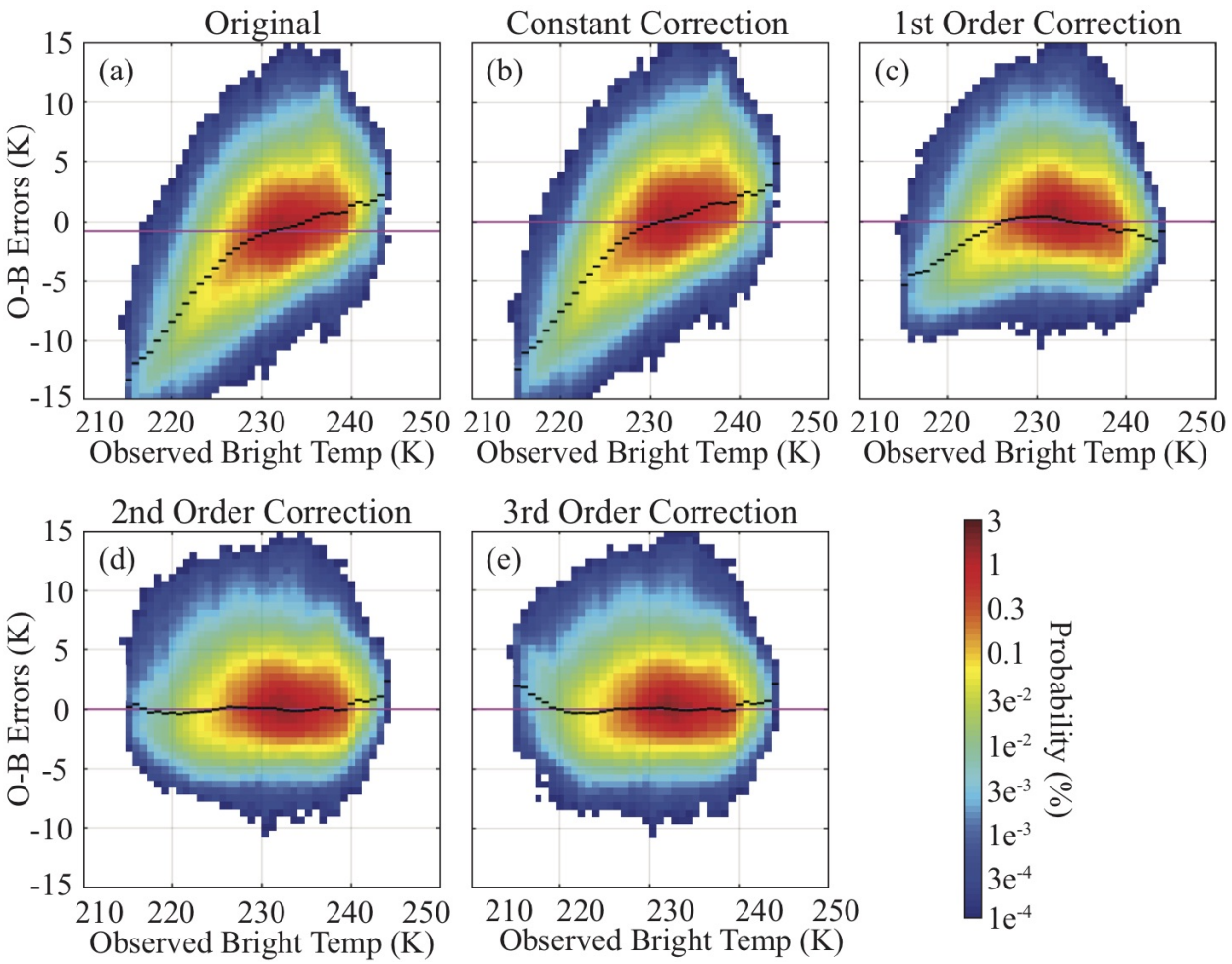

Figure 3.11: Same as Fig. 3.2 except for showing probability distributions plotted as a function of the observed 6.2 μm brightness temperatures when the observed 6.2 μm brightness temperature, satellite zenith angle, and vertically-integrated total water content from 100-700 hPa are used as the BC predictors.

3.5.3.2 SEVIRI 7.3 μm Example

In this section, we assess the ability of the multivariate NBC method to improve the observation error characteristics of the 7.3 μm band. As discussed in Section 3.3.1, observations from this band are sensitive to WV and clouds in the middle and upper troposphere, with a weighting function that peaks near 500 hPa in clear sky scenes. Overall, each of the OMB departure distributions (Fig. 3.12) have shapes that are similar to the corresponding 6.2 μm distributions (Fig. 3.11); however, their error range is larger because the weighting function for this band peaks at a lower level in the troposphere, thereby leading to potentially larger departures due to mismatched clouds in the observations and model background. Though the linear BC term substantially improves the distribution by making the departures less negative for colder brightness temperatures, non-zero

conditional biases remain across most of the distribution, with negative (positive) biases occurring for brightness temperatures colder (warmer) than 230 K (Fig. 3.12c). As occurred in the previous experiments, the conditional biases are almost eliminated after the 2nd order BC term is used, with minimal changes occurring due to the 3rd order term (Figs. 3.12d, e). The negative skewness present in the original histogram (Fig. 3.13a) switches to a large positive skewness after the linear BC term is used (Fig. 3.13c). Inspection of the OMB departure distributions shows that the positive skewness developed in response to the large upward shift in the conditional distributions for the colder brightness temperatures (Fig. 3.12a) that exposed the conditional positive skewness in the original distribution for warmer brightness temperatures that was being masked in the overall histogram by the large negative OMB departures. Another notable feature of the histograms is that their peaks are higher and narrower than the 6.2 μm histograms (Figs. 3.3o-q). This strongly non-Gaussian behavior was already present in the original histogram and is likely due to the large percentage of clear-sky observations containing small departures combined with fatter tails due to cloud displacement errors. Even so, these results show that the NBC method improved the distribution such that the variance was much lower and the conditional biases were reduced to near zero across most of the distribution. Also, as was the case with the 6.2 μm band, the linear BC term had the largest impact on the overall statistics; however, the variance was also reduced when using the higher order nonlinear BC terms.

3.6 Discussion and Conclusions

In this study, output from a high-resolution, regional-scale ensemble DA system was used to explore the ability of an innovative method to remove the bias associated with all-sky satellite infrared brightness temperatures using a Taylor series polynomial expansion of the OMB departures. This so-called NBC method uses OMB statistics accumulated over some period of time to remove linear and nonlinear conditional biases in a distribution through use of higher order Taylor series terms and a set of BC predictors. Nonlinear conditional biases can be identified using 2nd (quadratic) and 3rd (cubic) order terms (and even higher order terms if desired), whereas the constant and linear bias components can be diagnosed using the 0th and 1st order terms, respectively.

The ability of the NBC method to effectively remove the bias associated with all-sky SEVIRI infrared brightness temperatures was assessed using output from high-resolution ensemble DA

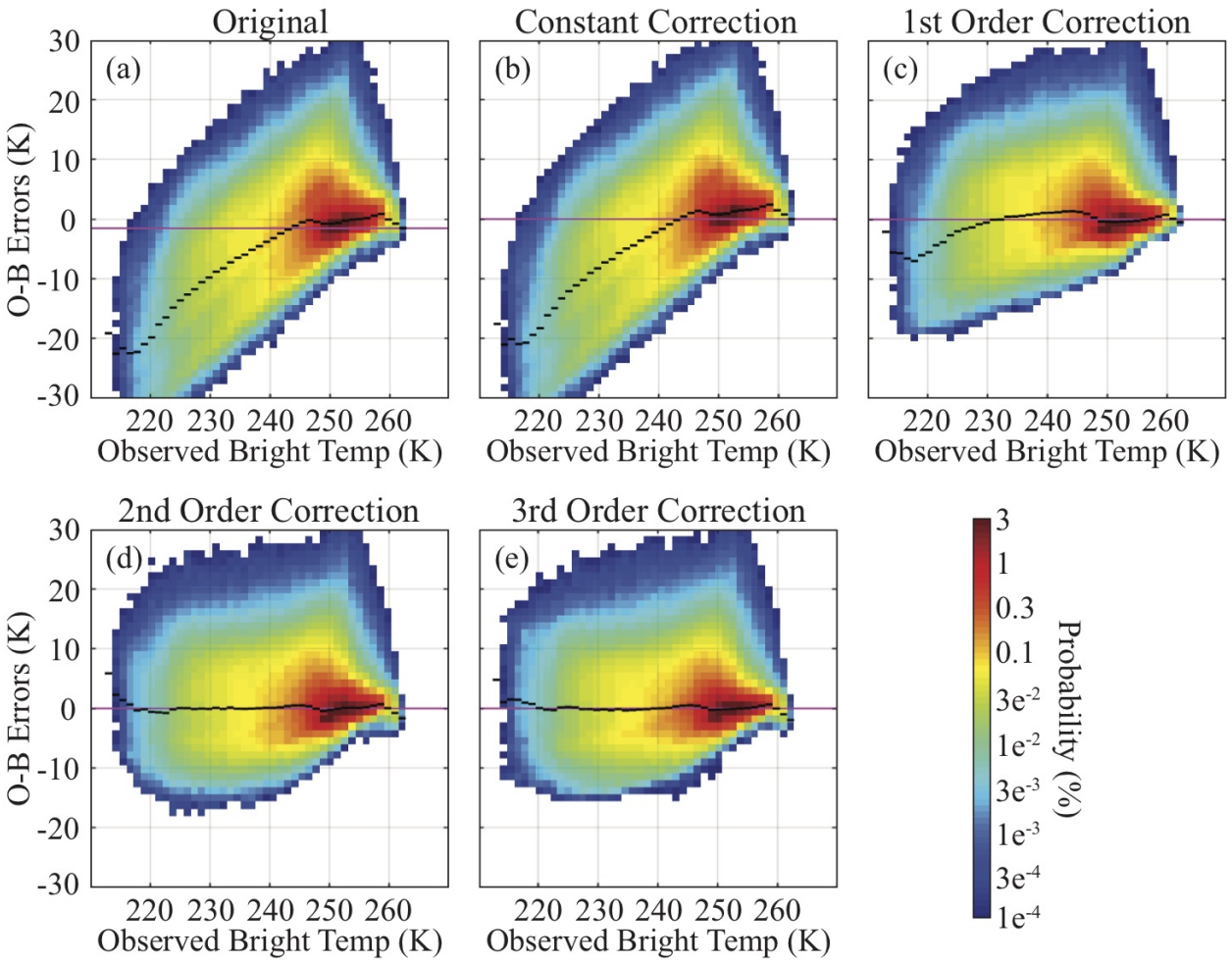

Figure 3.12: Probability distributions of 7.3 μm observation-minus-background departures plotted as a function of the observed 7.3 μm brightness temperatures (K) for the (a) original data, and the (b) constant, (c) 1st order, (d) 2nd order, and (e) 3rd order bias corrected observations when the observed 7.3 μm brightness temperature, satellite zenith angle, and model-integrated total water content from 100-700 hPa are used as the predictors. Data were accumulated at hourly intervals during a 108-h period from 13 UTC on 16 May 2014 to 00 UTC on 20 May 2014.

experiments performed using the KENDA system. OMB departure statistics for the 6.2 and 7.3 μm bands sensitive to clouds and WV in the upper and middle troposphere, respectively, were accumulated at hourly intervals during a 108-h period from 16-21 May 2014 using output from the COSMO-DE domain that covers Germany and surrounding areas with 2.8-km horizontal grid spacing. Conventional observations were actively assimilated, whereas the SEVIRI observations were passively monitored and therefore did not affect the analyses during the hourly assimilation cycles. Model-equivalent brightness temperatures were computed for each observation and ensemble member using the RTTOV radiative transfer model. The study period contained both clear-sky areas and a wide range of cloud types that together promoted a realistic assessment of the NBC method during the warm season.

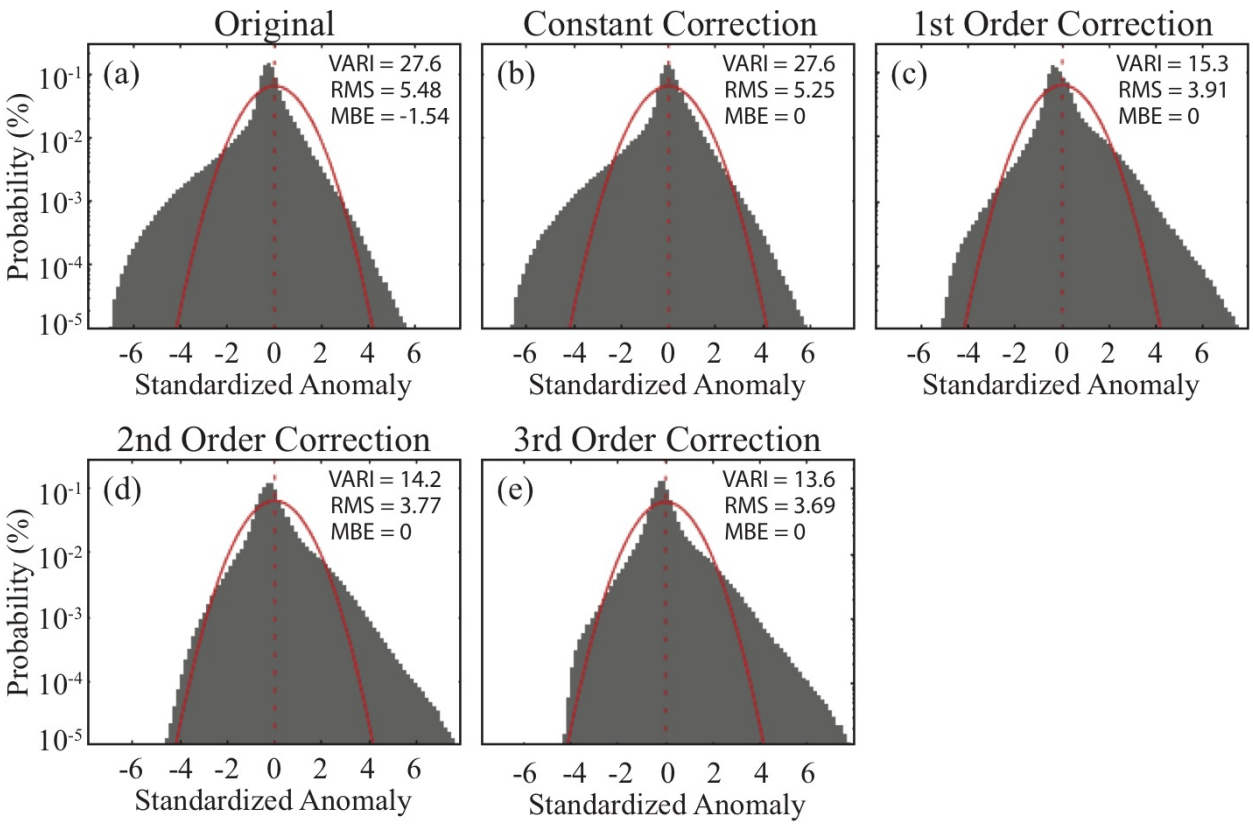

Figure 3.13: Probability density function of normalized 7.3 μm observation-minus-background departures for the (a) original data, and the (b) constant, (c) 1st order, (d) 2nd order, and (e) 3rd order bias corrected observations when the observed 7.3 μm brightness temperatures are used as the predictor. Data were accumulated at hourly intervals during a 108-h period from 13 UTC on 16 May 2014 to 00 UTC on 20 May 2014.

Univariate and multi-variate NBC experiments were performed using the satellite zenith angle and other predictors sensitive to clouds and WV, with their impact on the conditional bias and other aspects of the OMB departure distributions assessed using normalized histograms and probability distributions plotted as a function of the predictor. Overall, the results revealed that there are often strongly nonlinear conditional bias patterns in the OMB probability distributions that cannot be removed using only constant and linear BC terms. Though the overall bias of each distribution is equal to zero regardless of the order of the Taylor series expansion, there are often large conditional biases that vary as a function of the BC predictor. Because each SEVIRI band had a relatively small systematic bias, the constant BC term only had a small impact on the distributions. The linear 1st order term generally had the largest impact on the statistics of the entire distribution as measured by reductions in the variance; however, conditional biases often remained across much of the distribution. These conditional biases were typically reduced to near zero across the entire distribution only after the nonlinear 2nd and 3rd order terms were applied

to the OMB departures. Indeed, the conditional bias patterns often exhibited an arch shape for which the 2nd order quadratic term is ideally suited to remove. The tendency for the nonlinear terms to have a small impact on the variance of the entire distribution while simultaneously having a large positive impact on the conditional biases also illustrates that detailed analysis methods such as 2-D probability distributions provide valuable insight into the behavior of the BC method that is not possible using traditional 1-D error histograms.

Inspection of the univariate NBC results showed that the variance of the BC distributions was smallest when the brightness temperature observations were used as the BC predictor. The variance was also substantially reduced when the NWC SAF cloud top height retrievals were used as the predictor. Both of these predictors were able to diagnose and remove nonlinear biases associated with the cloudy observations. For example, large positive conditional biases for mid-level clouds transitioned into large negative conditional biases for upper-level clouds. Though not examined during this study, the different signs of the conditional biases for these clouds could be related to the ability of the COSMO model and RTTOV to properly simulate ice and mixed-phase cloud properties. The experiments using the satellite zenith angle or vertically-integrated water content showed that these BC predictors had a much smaller impact on the variance of the overall distribution. This behavior indicates that variables sensitive to the cloud top height are more effective BC predictors for all-sky infrared brightness temperatures, especially when higher order Taylor series terms are included. Even so, the multivariate experiments showed that though the zenith angle and total water content predictors only had a relatively small impact on the departure histograms when used individually, they still provided new information that greatly reduced the variance of the distribution when used in combination with the observed brightness temperature predictor.

Additional univariate NBC experiments were performed to examine the influence of linear and nonlinear components on the OMB departure distributions for clear- and cloudy-sky observations using a subset of the 6.2 μm brightness temperatures for which both a given observation and the corresponding model grid point were identified as being clear or cloudy. Overall, comparisons of the statistics for the clear-sky and cloudy-sky matched observations revealed that nonlinear error sources are much more important for cloudy sky observations as signified by the much larger impact of the 2nd and 3rd order Taylor series terms on the variance and the conditional biases in the distributions. For the clear-sky observations, the conditional biases could be effectively removed using only the 0th and 1st order terms, which is consistent with existing oper-

ational BC methods that typically remove the bias from clear-sky satellite observations using a set of constant and linear BC coefficients. These results show that the nonlinear conditional bias patterns evident in the all-sky OMB departure distributions primarily resulted from nonlinear biases in the cloudy-sky infrared brightness temperatures. They also show that the NBC method can effectively remove both linear and nonlinear conditional biases from all-sky infrared brightness temperatures provided that a suitable cloud-sensitive variable is used as one of the predictors.

Future work includes running cycled DA experiments using the KENDA system to assess the impact of the NBC method on the forecast accuracy when assimilating clear- and cloudy-sky infrared brightness temperatures. Additional experiments will be necessary to explore the ability of the method to remove biases from the OMB departures when the simulated brightness temperatures and cloud top heights are used as the BC predictors rather than their observed counterparts. Preliminary results indicate that predictors derived from the NWP model cloud field rather than the observations have a smaller impact on the overall statistics as measured by reductions in variance; however, they were still able to effectively remove the conditional biases across most of the distribution when higher order Taylor series terms were used. These results also indicate that it may be necessary to use up to a 4th order polynomial to remove the bias if the NWP-derived quantities are used rather than their observed counterparts. A more detailed assessment of this sensitivity is currently underway. Additional experiments will also be necessary to explore the ability of the NBC method to remove biases from infrared bands that are sensitive to the land surface or other atmospheric constituents such as ozone, as well as for all-sky microwave and visible radiances.

Though the NBC method used in this paper was implemented as a static, off-line method, it could also be incorporated into online methods such as VarBC through inclusion of additional nonlinear predictors. For example, the VarBC system at the Met Office uses Legendre polynomial predictors to remove residual scan biases and Fourier predictors to correct complex orbital biases in some satellite sensors (Cameron and Bell, 2016). Higher order predictors, such as the quadratic form of the temperature lapse rate and 4th order polynomial of the satellite angle bias, are also widely used in operational VarBC systems. Zhu et al. (2015) recently showed that inclusion of a quadratic aircraft ascent/descent term reduced the bias when assimilating aircraft temperature observations. Results from the current study could be used to help inform the development of operational DA systems as they continue to expand into all-sky satellite DA. Finally, many of the all-sky OMB departure distributions exhibited narrow peaks and fat tails that could potentially be

better represented using a Huber norm (Huber 1972) representation, which has been shown to lead to improved quality control and more observations being assimilated (Tavolato and Isaken 2015). Further research is necessary to determine if using a Huber norm in combination with the NBC method can improve existing quality control methods by identifying erroneous observations after the nonlinear conditional biases have been removed from the distribution. This approach could potentially preserve more cloud-affected observations where nonlinear biases are more prevalent, thereby leading to additional observations being assimilated in sensitive areas of the domain.

3.7 Acknowledgments

We thank each reviewer for their prompt reviews and detailed feedback that improved the manuscript. We gratefully acknowledge Jesse Stroik from the University of Wisconsin-Madison and Hendrik Reich, Andreas Rhodin, Robin Faulwetter, and Axel Hutt from the German DWD for their assistance porting and installing the KENDA system and basic cycling (BACY) scripts to the NOAA/NESDIS/STAR "S4" supercomputer located at the University of Wisconsin-Madison. The S4 supercomputer was used to perform all of the cycled DA experiments. The lead author was partially supported by the NOAA Joint Polar Satellite System (JPSS) program via CIMSS Cooperative Agreement NA15NES4320001 and by a University of Reading International Research Studentship.

Chapter 4

Nonlinear Bias Correction – Active Assimilation Experiments

4.1 Abstract

Ensemble data assimilation experiments were performed to assess the ability of satellite all-sky infrared brightness temperatures and different bias correction (BC) predictors to improve the accuracy of short-range forecasts used as the model background during each assimilation cycle. Satellite observations sensitive to clouds and water vapor in the upper troposphere were assimilated at hourly intervals during a 3-day period. Linear and nonlinear conditional biases were removed from the infrared observations using a Taylor series polynomial expansion of the observation-minus-background departures and BC predictors sensitive to clouds and water vapor or to variations in the satellite zenith angle. Assimilating the all-sky infrared brightness temperatures without BC degraded the forecast accuracy based on comparisons to radiosonde observations. Removal of the linear and nonlinear conditional biases from the satellite observations substantially improved the results, with predictors sensitive to the location of the cloud top having the largest impact, especially when higher order nonlinear BC terms were used. Overall, experiments employing the observed cloud top height or observed brightness temperature as the bias predictor had the smallest water vapor, cloud, and wind speed errors, while also having less degradation to temperatures than occurred when using other predictors. The forecast errors were smaller during these experiments because the cloud-height-sensitive BC predictors were able to more effectively remove the large conditional biases for lower brightness temperatures associated with a deficiency

in upper-level clouds in the model background.

4.2 Introduction

Indirect observations of the atmosphere, ocean, and land surface conditions obtained using sophisticated satellite remote sensing instruments are an indispensable component of the global observing system. For numerical weather prediction (NWP) applications, satellite radiances from visible, infrared, and microwave bands provide important information about atmospheric variables, such as temperature, winds, water vapor, and clouds, as well as lower boundary variables such as soil moisture, vegetation biomass, and sea surface temperatures. Satellite observations can also be used to detect the presence of aerosols and trace gases that are important for health and air quality models. Recent enhancements to the global satellite observing system through deployment of more accurate sensors onboard geostationary and polar-orbiting satellite platforms has made it possible to routinely monitor environmental conditions with high spatial and temporal resolution across the entire globe (Klaes et al. 2007; Strow 2013; Bessho et al. 2016; Schmit et al. 2017).

As satellite remote sensing capabilities have expanded and improved during the past several decades, substantial progress has also been made in our ability to extract more information from these important observations through development of advanced data assimilation (DA) methods and more accurate NWP models. Despite using only a small percentage of all available observations, satellite brightness temperatures and derived products such as atmospheric motion vectors still constitute more than 90% of the observations that are actively assimilated in most operational global NWP models (Bauer et al. 2010). Satellite observations are especially important in data sparse regions or for model state variables such as clouds and water vapor for which conventional in situ observations with high spatial and temporal resolution are not available.

Until the past decade, however, almost all efforts within the operational and research DA communities were directed toward optimizing the use of clear-sky brightness temperatures. This point of emphasis was not made because cloud-impacted observations were deemed unimportant, but rather, was due to the difficulty of using them in existing DA systems (Errico et al. 2007). Indeed, until the recent development of all-sky DA methods, the need to exclude observations impacted by clouds and precipitation meant that only a small percentage of available satellite ob-

servations were actively assimilated at global NWP centers (Yang et al. 2016). This limitation is even more severe for regional-scale NWP models where the entire domain may be covered by clouds (Lin et al. 2017). Though more effective assimilation of clear-sky satellite brightness temperatures has contributed to a steady increase in forecast skill, neglecting observations impacted by clouds is problematic because they tend to be located in dynamically active regions where the generation of more accurate initialization datasets through better use of these observations could help constrain potentially rapid error growth in NWP models (McNally 2002).

Observations sensitive to clouds and precipitation are challenging to use for a variety of reasons, as discussed by Errico et al. (2007). For example, though observation-minus-background (OMB) departure statistics are generally close to Gaussian for clear-sky observations, they can have substantial non-Gaussian error characteristics in the presence of clouds and precipitation (Bocquet et al. 2010; Okamoto et al. 2014; Harnisch et al. 2016; Otkin et al. 2018). Short-range model forecasts used as the first guess often exhibit large errors in the placement and characteristics of clouds and precipitation. Limited predictability of small-scale features and the difficulty of accurately modeling moist processes means that it is common for the model first guess to have much larger errors for clouds and precipitation than it does for dynamical variables such as temperature and geopotential height (Fabry and Sun 2010). Though representativeness errors can usually be ignored when assimilating clear-sky observations primarily sensitive to temperature, they become important for cloud-affected observations because they can lead to very large OMB departures that hinder their assimilation (Geer and Bauer 2011; Geer et al. 2012; Okamoto 2013). It is also more difficult to quantify the observation and model background errors because it can be challenging to separate signals associated with the individual atmospheric and land surface variables that contribute to the sensitivity of a given satellite observation (Bauer et al. 2011). Another prominent problem is the difficulty of modeling complex cloud properties in the radiative transfer models used to compute the model-equivalent brightness temperatures. Nonlinear error statistics due to deficiencies in the radiative transfer and NWP models could lead to erroneous analysis increments in the model state variables that in turn could impact balance and stability during the first few hours of the forecast (Errico et al. 2007). Last, it is also important to account for correlated observation errors because they can become very large in the presence of clouds and precipitation (Bormann et al. 2011, 2016; Campbell et al. 2017).

Despite these and other issues that make it more challenging to assimilate cloud-sensitive observations, substantial progress has still been made during the past decade (Geer et al. 2017,

2018). Successful efforts to assimilate all-sky satellite observations have occurred in tandem with improvements in the representation of water vapor and cloud features in NWP models and advances in the ability of radiative transfer models to accurately model radiative fluxes in clouds. These efforts have also been aided through the widespread adoption of four-dimensional variational (4DVAR) and ensemble DA methods that can more easily extract information about dynamical variables from cloud- and moisture-sensitive observations (Geer et al. 2014; Lien et al. 2016; Zhu et al. 2016). For example, Peubey and McNally (2009) demonstrated that four-dimensional variational methods could extract useful information about the wind field from moisture-sensitive satellite observations through the "tracer-advection" mechanism. Likewise, ensemble DA systems can infer the temperature, water vapor, and wind fields through ensemble covariances that link the model state variables to the simulated observations (Zhang et al. 2011; Houtekamer and Zhang 2016). Compared to DA methods that only assimilate clear-sky satellite observations, an important benefit of an all-sky DA approach is that it provides a unified treatment of cloud-free and cloud-impacted observations that negates the need to perform potentially unreliable and expensive cloud detection procedures (Bauer et al. 2010). An all-sky DA approach also promotes a more balanced use of satellite observations in clear and cloudy areas that helps overcome the tendency for operational DA systems to assimilate substantially more observations in regions that are not affected by clouds or precipitation (Geer et al. 2017).

Early efforts to assimilate all-sky satellite observations focused on microwave sounding channels that are sensitive to water vapor and non-precipitating cloud particles (Bauer et al. 2010). These channels were initially chosen because they have more Gaussian error characteristics than cloud-sensitive infrared and visible channels, thereby making them a logical starting point to explore the assimilation of all-sky observations. Whereas it was once thought that it may prove too difficult to assimilate water vapor and cloud-sensitive satellite observations (e.g., Bengtsson and Hodges, 2005), their impact has increased greatly in recent years (Geer et al. 2018). The direct assimilation of all-sky microwave observations was first accomplished in an operational DA system in 2009 at the European Centre for Medium-range Weather Forecasting (ECMWF) (Bauer et al. 2010). Since then, the impact of these observations has risen to nearly 20% (Geer et al. 2017), as measured using the forecast sensitivity observation impact metric (Langland and Baker 2004). This rapid increase in their impact means that all-sky microwave observations have become one of the most important sources of data in the ECMWF model, with an impact comparable to clear-sky satellite radiances and conventional observations. More recently, the National Centers for

Environmental Prediction has also started to assimilate all-sky microwave observations in their operational global forecasting system (Zhu et al. 2016). Numerous studies have documented the benefits of assimilating all-sky microwave observations in global and regional modeling systems (e.g., Aonashi and Eito 2011; Geer et al. 2014; Yang et al. 2016; Kazumori et al. 2016; Baordo and Geer 2016; Zhang and Guan 2017; Lawrence et al. 2018; Wu et al. 2019).

In contrast to the extensive resources that have been directed by the operational DA community toward the assimilation of all-sky microwave observations, much less attention has been given to increasing the use of cloud-sensitive infrared brightness temperatures. Indeed, until the past few years, most studies that explored the assimilation of all-sky infrared observations have done so using research models or within the context of observing system simulation experiments (OSSEs). Early studies by Vukicevic et al. (2004, 2006) assimilated cloudy-sky infrared brightness temperatures from the 10.7- and 12.0-μm bands on the Geostationary Operational Environmental Satellite (GOES) Imager using a 4DVAR assimilation system. Observations from these atmospheric window bands were shown to improve the depiction of upper-level ice clouds; however, they had less impact on liquid clouds occurring lower in the troposphere. Subsequent studies by Stengel et al. (2009, 2013) found that assimilation of cloud-impacted infrared observations from the 6.2- and 7.3-μm water vapor channels on the Spinning Enhanced Visible and Infrared Imager (SEVIRI) sensor led to more accurate analyses and forecasts in a high-resolution regional-scale model. Other investigators proposed several methods that could be used to assimilate information from cloud-impacted observations from hyperspectral sounders onboard polar-orbiting satellite platforms (Heillette and Garand 2007; Pavelin et al. 2008; McNally 2009; Pangaud et al. 2009; Guidard et al. 2011; Lupu and McNally 2012). All of these methods were designed to estimate the cloud top pressure or effective cloud amount, with these parameters then fed to the DA system. This process enabled the assimilation of some cloud information from these observations.

The direct assimilation of cloud and water vapor sensitive infrared brightness temperatures has also been investigated using regional-scale OSSEs. Most of these studies employed ensemble DA systems and were used to examine the potential impact of assimilating observations from the Advanced Baseline Imager (ABI) onboard the GOES-R satellite (currently GOES-16 and GOES-17). In studies assimilating both clear- and cloudy-sky brightness temperatures from the ABI 8.5 μm band, Otkin (2010, 2012a) showed that their assimilation improved the cloud field and that it was necessary to use a short horizontal localization radius to account for small-scale cloud features in the infrared observations. A subsequent study by Otkin (2012b) revealed that assimilation of

all-sky observations from the three water vapor sensitive bands on the ABI sensor had a large positive impact on 6-h precipitation forecasts during a high-impact winter storm. Jones et al. (2013a, 2014) examined the impact of simultaneously assimilating all-sky ABI brightness temperatures and Doppler radar reflectivity observations for an extratropical cyclone, where it was found that the most accurate analyses and forecasts were obtained when both observation types were assimilated because they are sensitive to different portions of the cloud field. The radar observations had a large positive impact on the cloud and wind fields in the lower troposphere, whereas the satellite observations provided additional improvements in the cloud and moisture fields in the upper troposphere. Other OSSE studies have shown similar positive results for various weather features, such as mesoscale convective systems and tropical cyclones (Zupanski et al. 2011; Cintineo et al. 2016; Zhang et al. 2016; Minamide and Zhang 2017, 2018; Pan et al. 2018).

Results from the various OSSE studies have been used to inform ongoing efforts by various groups to assimilate real all-sky infrared brightness temperatures and satellite-derived products. Most of these studies have focused on optimizing methods to assimilate data from geostationary satellite sensors in regional-scale ensemble DA systems. Geostationary satellite observations are very useful for these models because they are the only source of cloud and water vapor information with high spatial resolution. Moreover, unlike polar-orbiting satellites, geostationary sensors are also able to provide frequent observation updates that cover most, if not all, of the model domain. Some recent studies have shown positive results when assimilating satellite-derived products such as cloud water path or layer precipitable water (Jones et al. 2013b, 2015, 2016, 2018; Schomburg et al. 2015; Jones and Stensrud 2015; Kerr et al. 2015; Wang et al. 2018), whereas other studies have explored the direct assimilation of all-sky infrared brightness temperatures. Regardless, there is great potential in assimilating all-sky geostationary satellite observations in regional-scale models because clouds are the first observable aspect of convective systems (Gustafsson et al. 2018; Kurzrock et al. 2019).

Okamoto (2013) showed a slightly positive impact on temperature and wind analyses and 6-h forecasts when a subset of infrared brightness temperatures depicting spatially homogeneous clouds in the middle and upper troposphere were assimilated. Subsequent studies by Okamoto et al. (2014) and Harnisch et al. (2016) developed cloud-dependent all-sky observation error models where the error is allowed to vary as a function of a diagnosed cloud impact parameter. Similar in construct to the "symmetric" observation error model developed by Geer and Bauer (2011) for all-sky microwave observations, both models assign the largest errors to the most strongly cloud-

impacted observations given greater uncertainties in both the NWP and radiative transfer models in cloudy scenes. Minamide and Zhang (2017) have proposed an alternative method, known as adaptive observation error inflation, that scales the observation errors as a function of the first guess departure, with the largest errors given to observations with the largest departures. Application of these dynamical observation error models to all-sky infrared brightness temperatures generally leads to more Gaussian departure statistics.

Other studies have shown that assimilation of all-sky infrared observations from geostationary satellite sensors can improve forecasts for tropical cyclones, floods, and severe thunderstorms (Zhang et al. 2016, 2018; Honda et al. 2018a,b; Minamide and Zhang 2018). In particular, these case studies revealed that assimilation of all-sky observations improved the prediction of the mid-level mesocyclone during a tornadic thunderstorm and the structure of the inner core and outer rainband regions for several tropical cyclones. More accurate precipitation forecasts were also shown to lead to more skillful flood forecasts from a river discharge model (Honda et al. 2018b). Though the direct assimilation of all-sky infrared brightness temperatures is currently not included in any operational DA system, Geer et al. (2019) present promising early results from a semi-operational implementation of the ECMWF model. Their study assimilated all-sky observations from seven water vapor sensitive bands on the Infrared Atmospheric Sounding Interferometer sensor onboard the polar-orbiting Metop-A and Metop-B satellites. It was shown that the newly-developed all-sky DA approach gave results that were as good or better than the existing clear-sky-only approach, with the largest benefits found in the tropics where short-range forecasts were improved throughout the troposphere and stratosphere.

In this study, we advance efforts to assimilate all-sky infrared brightness temperatures from the cloud and water vapor sensitive 6.2-μm band on the SEVIRI sensor using a pre-operational version of the Kilometer-scale Ensemble Data Assimilation (KENDA) system run at the German Deutscher Wetterdienst (DWD). Experiments are run in which the nonlinear bias correction (NBC) method developed by Otkin et al. (2018) is used to remove systematic biases from the all-sky observations prior to their assimilation. Given the proven utility of clear-sky satellite BC methods (Eyre 2016), it is necessary to develop cloud-dependent BC methods for all-sky infrared brightness temperatures to make full use of these observations within modern DA systems. Cloud-dependent biases can occur for a variety of reasons. For example, deficiencies in the forward radiative transfer model used to compute the model-equivalent brightness temperatures, or the inability of the parameterization schemes in the NWP model to accurately represent the spatial

extent, thickness, and optical properties of clouds, can introduce systematic errors that vary as a nonlinear function of some cloud property, such as cloud top height (Dee 2005; Dee and Uppala 2009; Mahfouf 2010; Otkin and Greenwald 2008; Cintineo et al. 2014; Eikenberg et al. 2015). Though the accuracy of radiative transfer models has improved greatly in recent years, there are still large uncertainties regarding the specification of cloud properties, especially for ice clouds (Yang et al. 2013; Baum et al. 2014; Yi et al. 2016).

Most BC methods use a set of predictors describing aspects of the atmospheric state or characteristics of the satellite data to remove biases from the OMB departures (Eyre 2016). So-called "static" BC methods use a set of departures accumulated over long periods of time outside of the DA system to estimate and remove biases from the observations (Eyre 1992; Harris and Kelly 2001; Hilton et al. 2009). In contrast to the non-time-varying BC coefficients derived using static methods, variational BC (VarBC) methods update the BC coefficients during each DA cycle using an augmented control vector (Derber et al. 1991; Parrish and Derber 1992; Derber and Wu 1998; Dee 2005; Auligne et al. 2007; Dee and Uppala 2009; Zhu et al. 2014, 2016). Recently, Zhu et al. (2016) expanded an existing operational VarBC method so that it could be used to remove biases from all-sky microwave observations. To reduce errors associated with mismatched cloud fields, the BC coefficients with this method were computed using only situations where both the observed and model-equivalent brightness temperatures were diagnosed as clear or cloudy. Though most studies have focused on variational or hybrid DA systems, several studies have also explored their use in ensemble DA systems (Szunyogh et al. 2008; Fertig et al. 2009; Stengel et al. 2009, 2013; Miyoshi et al. 2010; Aravequia et al. 2011; Cintineo et al. 2016).

BC methods typically assume that a linear relationship exists between the OMB departure bias and a given set of predictors. Though previous studies have shown that linear BC methods are able to effectively remove biases from clear-sky satellite observations, these methods are suboptimal if the observation bias varies as a nonlinear function of some predictor. Otkin et al. (2018) showed that nonlinear conditional biases are more likely to occur for cloudy observations, which necessitates development of BC methods that can more easily capture complex error patterns when assimilating all-sky observations. Their study also showed that cloud-sensitive predictors, such as cloud top height (CTH) or the brightness temperatures themselves, are most effective at removing biases from all-sky infrared observations. In this study, we build upon the work of Otkin et al. (2018) by assessing the ability of linear and nonlinear BC predictors in the context of all-sky infrared brightness temperature assimilation to improve short-range (1-h) forecasts in

an ensemble DA system. The paper is organized as follows. The DA framework is described in section 4.3, with assimilation results using different linear and nonlinear BC predictors presented in section 4.4. Conclusions and a discussion are presented in section 4.5.

4.3 Experimental Design

4.3.1 SEVIRI Satellite Datasets

The DA experiments performed during this study employed all-sky infrared brightness temperatures from the SEVIRI sensor onboard the Meteosat Second Generation satellite, along with CTH retrievals provided by the EUMETSAT Nowcasting Satellite Applications Facility. The SEVIRI sensor observes the top-of-atmosphere radiances across 12 visible and infrared spectral bands, with a nadir resolution of 3 km for all infrared bands (Schmetz et al. 2002). This study focuses on the assimilation of clear and cloudy-sky brightness temperatures from the 6.2 μm band sensitive to clouds and water vapor in the upper troposphere. Under clear-sky conditions, the weighting function for this band peaks near 350 hPa for a standard mid-latitude atmosphere; however, it will shift upward and become truncated near the cloud top when clouds are present due to increased scattering. It will also peak at a higher (lower) atmospheric level if more (less) water vapor is present in the middle and upper troposphere. The dual sensitivity of this band to clouds and water vapor is advantageous for DA applications because increasing moisture and increasing cloud optical thickness influence the infrared brightness temperatures in a similar way. The resultant smoother dependence between water in its vapor and condensed (cloud) states will generally lead to more Gaussian statistics than would occur with an infrared atmospheric window band that has little or no sensitivity to water vapor.

As will be discussed in Section 4.4, CTH retrievals derived from SEVIRI observations were used as one of the BC predictors during the DA experiments. With this dataset, the CTH is estimated for each satellite pixel by first computing a simulated clear-sky 10.8 μm brightness temperature using the Radiative Transfer for TOVS (RTTOV) radiative transfer model (Saunders et al. 1993) and temperature and water vapor profiles from the global NWP model run at the DWD (Majewski et al. 2002). An opaque cloud is then inserted in the atmospheric profile at successively higher levels until the difference between the observed and simulated brightness temperatures is minimized (Derrien and Le Gleau 2005). The CTH retrievals have a nominal vertical resolution

of 200 m; however, their uncertainty is larger for semi-transparent clouds (Le Gleau 2016). To minimize the impact associated with spatially correlated errors, the CTH retrievals and SEVIRI brightness temperatures were horizontally thinned by a factor of five in the meridional and zonal directions. This reduces their horizontal resolution to ∼20-25 km across the model domain, which is ∼8 times coarser than the resolution of the NWP model employed during this study.

4.3.2 KENDA Data Assimilation System

Ensemble DA experiments were performed using a research version of the regional-scale KENDA system (Schraff et al. 2016) used at the DWD. A major development focus of KENDA in recent years has been the inclusion of cloud- and precipitation-sensitive observations that can be used to constrain the cloud and thermodynamic fields in convection-resolving models. KENDA employs a local ensemble transform Kalman filter (Hunt et al. 2007) during the analysis step and the Consortium for Small-scale Modeling (COSMO) NWP model (Baldauf et al. 2011) during the forecast step. All of the DA experiments were run on the COSMO-DE domain covering Germany and parts of surrounding countries with 2.8 km horizontal resolution. With this version of KENDA, the lateral boundary conditions were obtained at hourly intervals from the COSMO-EU domain run at the DWD, which in turn was driven by lateral boundary conditions from the global Icosahedral non-hydrostatic (ICON) model (Zangl et al. 2015). The COSMO-DE domain contains 50 terrain-following vertical layers, with the model top located near 22 km (about 40 hPa).

The COSMO model includes prognostic variables for atmospheric temperature, pressure, horizontal and meridional wind components, and the mixing ratios for water vapor, cloud water, rainwater, ice, snow, and graupel. Cloud microphysical processes are handled using a simplified version of the double-moment Seifert and Beheng (2001) microphysics scheme that was reduced to a single-moment scheme for computational purposes, whereas the parameterization of cloud formation and decay processes is based on Lin et al. (1983). Though deep convection is explicitly resolved on the COSMO-DE domain, a simplified version of the Tiedtke (1989) mass-flux scheme is used to parameterize shallow convection. Atmospheric turbulence is predicted using the 2.5 order turbulent kinetic energy scheme developed by Raschendorfer (2001). A δ-2 stream radiative transfer method is used to update atmospheric heating rates due to radiative effects at 15-min intervals (Ritter and Geleyn 1992).

The DA experiments employed a 40-member ensemble, along with a deterministic run that is

initialized by applying the Kalman gain matrix from the assimilation update to the deterministic model background. The ensemble and deterministic runs were initialized at 00 UTC on 28 May 2014 and then updated at hourly intervals during a 3-day period. Model-equivalent brightness temperatures for the SEVIRI 6.2 μm band were computed using version 10.2 of the RTTOV radiative transfer model that includes an enhanced cloud-scattering module that enables the use of cloud hydrometeor profiles located on the NWP model vertical grid (Matricardi 2005; Hocking et al. 2011). Vertical profiles of fractional cloud cover and ice and liquid water contents used to compute the cloudy-sky brightness temperatures were obtained using COSMO model output and empirical relationships developed by Kostka et al. (2014). The maximum-random cloud overlap scheme (Raisanen 1998) was used, with the ice crystals assumed to have a hexagonal shape and the effective particle diameters computed using the McFarquhar et al. (2003) method.

SEVIRI 6.2 μm brightness temperatures, along with radiosonde, surface, wind profiler, and aircraft observations, were actively assimilated at hourly intervals during each DA experiment. The corresponding model equivalents were computed at the exact observation times through inclusion of the various observation operators within the COSMO model. Covariance inflation values were computed at each grid point using a combination of the relaxation to prior perturbations approach described by Zhang et al. (2004) and multiplicative inflation based on Anderson and Anderson (1999). Covariance localization was performed by using only those observations located within a specified horizontal radius of a given analysis point. An adaptive horizontal localization radius was used for the conventional observations (Perianez et al. 2014); however, it was set to 35 km for the all-sky SEVIRI brightness temperatures given their uniform data coverage. The vertical localization scale was set to 0.7 in logarithm of pressure for the brightness temperatures, with the localization height determined using the peak of the satellite weighting function for the simulated brightness temperature from the deterministic run. The observation error was set to 4 K for the all-sky brightness temperatures, similar to that used in Otkin (2012b) and Cintineo et al. (2016). Though it may have been advantageous to use a cloud-dependent observation error model, that is beyond the scope of the current study.

4.3.3 Nonlinear Bias Correction Method

Systematic biases were removed from the satellite observations using the NBC method developed by Otkin et al. (2018). This method uses a Taylor series polynomial expansion of the OMB

departures for a given satellite band to remove linear and nonlinear conditional biases from the observations prior to their assimilation. A brief overview of the NBC method is provided here, with the reader referred to Otkin et al. (2018) for a more detailed description. To begin, the OMB departure vector is defined as:

$$\mathbf{dy} = \mathbf{y} - H(\mathbf{x}), \tag{4.1}$$

where $\mathbf{y}$ and $H(\mathbf{x})$ are vectors containing the observed and model-equivalent brightness temperatures, respectively, and H is the observation operator that is used to convert the NWP model first guess fields into simulated brightness temperatures. If we assume that any biases present in the OMB departures can be described by a real function $f(z)$ that is infinitely differentiable around a real number c, Eqn. 4.1 can be decomposed into an N order Taylor series polynomial expansion. A representative example in which a single predictor is used to identify biases in a given set of observations using a 3rd order expansion is shown in Eqn. 4.2:

$$\mathbf{dy} = \left(b_0 + b_1(z^{(i)} - c) + b_2(z^{(i)} - c)^2 + b_3(z^{(i)} - c)^3 \right)_{i=1,\ldots,m} \tag{4.2}$$

where m is the number of observations, $z^{(i)}$ is the predictor value for the ith observation, b_n are the 0...nth BC coefficients, and c is a constant that can be set to any value. The $(i = 1, \ldots, m)$ notation outside the parentheses indicates that the Taylor series terms are computed separately for each element of the observation departure vector. In this example, the first two terms on the right hand side represent the constant and linear bias components, whereas the last two terms represent the nonlinear 2nd order (quadratic) and 3rd order (cubic) components.

Eqn. 4.2 can be rewritten in matrix notation as $\mathbf{dy} = \mathbf{Ab}$, where $\mathbf{A}$ is an m x n matrix containing the n Taylor series terms for each observation and $\mathbf{b}$ is an n x 1 vector containing the BC coefficients. This is an overdetermined system of m linear equations in n unknown coefficients because $m > n$. The BC coefficients that best fit the set of equations can be identified by solving the quadratic minimization problem, which, after adding a Tikhonov regularization term (αI) to improve its conditioning, leads to:

$$\mathbf{b} = (\alpha I + \mathbf{A}^{\mathrm{T}}\mathbf{A})^{-1}\mathbf{A}^{\mathrm{T}}\mathbf{dy}, \tag{4.3}$$

where $(\alpha I + \mathbf{A}^{\mathrm{T}}\mathbf{A})$ is a symmetric, n x n square matrix, thereby making it easy to compute its inverse. The Tikhonov regularization term is defined to be a multiple of the identity matrix, which is a standard approach when solving inverse problems (Nakamura and Potthast 2015). The constant

α was set to a very small value (10^{-9}) following the results of Otkin et al. (2018).

For this study, the BC coefficients for the SEVIRI 6.2 μm band were updated during each assimilation cycle using only the observation departure statistics accumulated during the previous hour. This approach was used rather than accumulating statistics over a longer time period because it allows the BC coefficients to quickly adapt to changes in the cloud field, such as those associated with the diurnal cycle of convection and its impact on cloud properties in the upper troposphere. All of the observation departures for a given assimilation cycle were used to compute the BC coefficients, thereby providing a larger sample size and negating the need to identify cloud-matched observations when determining the coefficients. After calculating the BC coefficients, they were then applied separately to each observation and ensemble member.

4.4 Results

In this section, we assess the ability of all-sky infrared brightness temperatures from the SEVIRI 6.2 μm band to improve short-range forecasts when assimilated in an ensemble DA system after using various BC predictors to remove biases from the observations. Figure 4.1 shows the evolution of the upper-level cloud and water vapor fields during the 3-day assimilation period, as depicted by the observed SEVIRI 6.2 μm brightness temperatures. At the start of the period, an extensive area of cold, upper-level clouds associated with widespread precipitation extended from northwest-to-southeast across the domain (Fig. 4.1a). As this weather feature slowly moved southward and weakened during the next two and a half days, the lower brightness temperatures indicative of optically thick clouds were steadily replaced by higher brightness temperatures as the clouds became optically thinner and their spatial extent lessened. A small area of clear skies across the southwestern part of the domain was shunted southward during this time period, with a much larger area of clear skies developing behind the departing weather feature (Fig. 4.1e). Within these clear-sky areas, the highest brightness temperatures are associated with the driest conditions in the upper troposphere. Overall, this synopsis shows that there were a wide range of cloud and water vapor conditions in the upper troposphere that together support a realistic assessment of the impacts of the infrared brightness temperatures and bias predictors in the assimilation system.

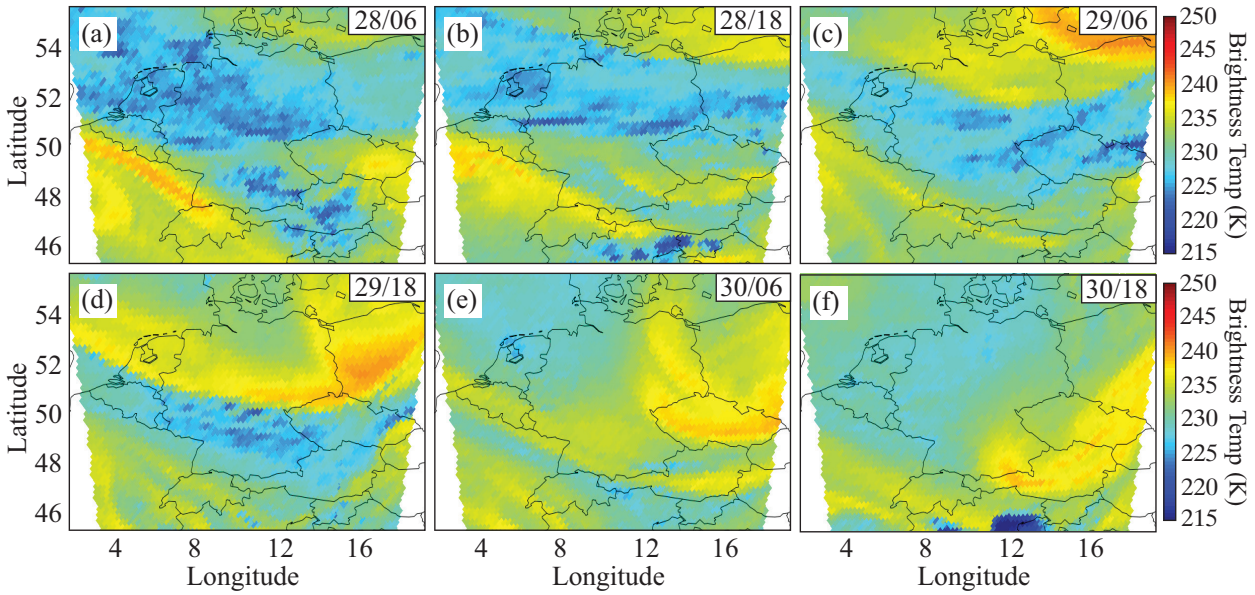

Figure 4.1: Observed SEVIRI 6.2 μm brightness temperatures (K) valid at (a) 06 UTC on 28 May, (b) 18 UTC on 28 May, (c) 06 UTC on 29 May, (d) 18 UTC on 29 May, (e) 06 UTC on 30 May, and (f) 18 UTC on 30 May 2014.

4.4.1 Assessing the Impact of Nonlinear Bias Corrections

Prior work by Otkin et al. (2018) found that it was necessary to use nonlinear BC predictors to remove cloud-dependent biases from passively monitored all-sky infrared brightness temperatures. Here, we extend their results by examining the impact of nonlinear BC predictors in cycled DA experiments where all-sky 6.2 μm brightness temperatures are actively assimilated. In particular, experiments are performed where the observation bias is removed using a 0th (constant), 1st (linear), 2nd (quadratic), or 3rd (cubic) order Taylor series polynomial expansion of the OMB departures when the observed cloud top height is used as the bias predictor. To provide complete domain coverage, satellite pixels identified as clear in the EUMETSAT CTH product were assigned a height equal to the model terrain elevation. These four experiments are hereafter referred to as OBSCTH-0TH, OBSCTH-1ST, OBSCTH-2ND, and OBSCTH-3RD, respectively. Results from these experiments are then compared to two baseline experiments in which the all-sky infrared observations are either not assimilated (No-Assim), or are actively assimilated, but without using BC (No-BC). The impact of the BC predictors is assessed using OMB departure statistics from the prior ensemble mean analyses accumulated at hourly intervals during the 72-h assimilation period. The prior analyses are used here to provide a measure of the observation impact on short-range (1-h) forecasts.

4.4.1.1 Brightness Temperature Bias Correction Statistics

To assess how the BC changes in relation to use of linear and nonlinear predictors, Fig. 4.2 shows the 2-D probability distribution of OMB departures for the 6.2 μm brightness temperatures from the No-Assim experiment (Fig. 4.2a), along with the corresponding BC distributions for each DA experiment. All of the distributions are plotted as a function of the observed 6.2 μm brightness temperatures. The magenta line in each panel denotes the mean of the entire distribution, whereas the shorter black lines depict the conditional mean in each column. Inspection of Fig. 4.2a reveals that, though the mean bias during the No-Assim experiment is relatively small (-0.76 K), the conditional biases exhibit an asymmetrical arch-shaped pattern that is a nonlinear function of the observed brightness temperatures. The conditional biases are close to zero for brightness temperatures near 235 K, and remain small for brightness temperatures $>$ 230 K; however, they become progressively more negative for lower brightness temperatures. The large negative biases for the lowest brightness temperatures indicate that the COSMO model forecasts are deficient in upper-level clouds or that there are biases in the RTTOV model used to compute the model-equivalent brightness temperatures. Regardless, assimilation of observations with such large biases could degrade the performance of the DA system. The simplest option is to exclude these observations, however, that is not ideal because they still contain valuable information about random errors in the cloud field if the biases can be removed.

Inspection of the corrections applied to the infrared observations during the active DA experiments (Figs. 4.2b-e) reveals that the mean BC is similar for all experiments despite the 2-D distributions having very different shapes. This occurs because the mean BC is most strongly influenced by the mean bias in the full set of OMB departures (Fig. 4.2a) and by the tendency for larger corrections for the lower brightness temperatures to be offset by smaller corrections for the higher brightness temperatures. Because the single bias predictor in the OBSCTH-0TH experiment (Fig. 4.2b) is only able to remove the mean bias during a given assimilation cycle, it has a narrower BC distribution than the other experiments. There is still some spread in the corrections during this experiment because the BC varies for each assimilation cycle due to changes in the prevailing atmospheric conditions. The constant corrections, however, are not optimal because they are unable to account for the large variations in the conditional biases across the OMB distribution (Fig. 4.2a). In contrast, more accurate corrections are obtained through use of the linear bias predictor during the OBSCTH-1ST experiment (Fig. 4.2c), as evidenced by the smaller (larger) BC for bright-

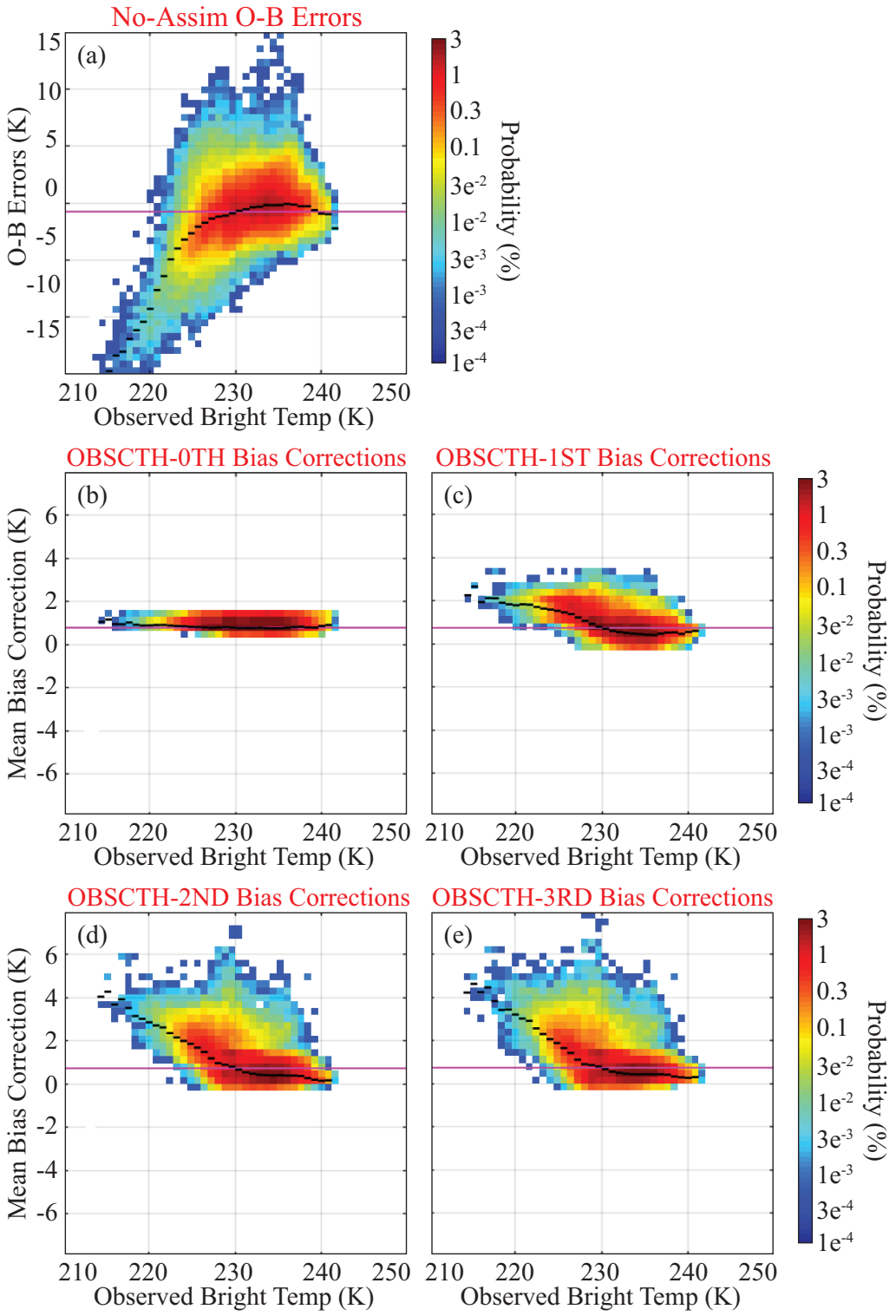

Figure 4.2: (a) Probability distribution of SEVIRI 6.2 μm observation-minus-background (O-B) brightness temperature departures (K) for the No-Assim experiment plotted as a function of the observed 6.2 μm brightness temperatures (K). (b-e) Probability distributions of SEVIRI 6.2 μm ensemble mean brightness temperature bias corrections (K) for the OBSCTH-0TH, OBSCTH-1ST, OBSCTH-2ND, and OBSCTH-3RD experiments plotted as a function of the observed 6.2 μm brightness temperatures (K). Data were accumulated at hourly intervals during a 72-h period from 00 UTC on 28 May 2014 to 00 UTC on 31 May 2014. The horizontal purple lines shows the mean O-B departure (panel a) or the mean bias correction (panels b-e), whereas the black line segments depict the conditional O-B bias (panel a) or the mean bias correction (panels b-e) in each column.

ness temperatures greater (less) than 230 K. The corrections for the lower brightness temperatures are even larger during the OBSCTH-2ND and OBSCTH-3RD experiments (Fig. 4.2d, e) because the additional nonlinear predictors are able to remove more of the conditional biases at those temperatures (Fig. 4.2a). Overall, these results indicate that the OBSCTH-2ND and OBSCTH-3RD experiments provide more accurate BC in the presence of complex nonlinear bias patterns.

4.4.1.2 Brightness Temperature Error Time Series

The evolution of the 6.2 μm brightness temperature root mean square error (RMSE) and bias during the 3-day assimilation period is shown in Fig. 4.3. The error statistics were computed using the ensemble mean brightness temperatures from the prior analyses for each assimilation cycle. Note that the bias is nonzero for all of the experiments because the statistics were computed using output from 1-h forecasts and prior to bias-correcting the satellite observations. Overall, there is a large diurnal cycle in the error statistics, with the largest RMSE and negative biases occurring during the daytime (09-18 UTC), followed by smaller errors at night. This error pattern is consistent with a lack of lower brightness temperatures during the afternoon when the deficiency in upper-level clouds associated with deep convection is most prominent (not shown). The large diurnal differences in the bias also illustrate why it is advantageous to compute the BC coefficients using observations from a single assimilation cycle because accumulation of OMB departures over longer time periods would obscure these important differences and therefore make the BC method less effective.

Inspection of the error time series reveals that the bias and RMSE are smallest during the No-BC experiment, which indicates that larger improvements are realized in the forecast cloud field when BC is not applied to the all-sky brightness temperatures. As will be shown in the next section, however, the improved fits to the satellite observations during the No-BC experiment do not translate into smaller errors for conventional observations that are not sensitive to clouds. Compared to the No-Assim experiment, the four experiments in which bias-corrected satellite observations were assimilated had similar biases, but much smaller RMSE, with values approaching those obtained during the No-BC experiment. The simultaneous large reductions in RMSE and small changes in bias demonstrate that even though the bias-corrected observations are unable to substantially reduce the bias, it is still possible to use them to fix random errors in the cloud and water vapor fields. Moreover, though there is a trend toward lower RMSE in all of the ex-

periments during the 3-day assimilation period due to a decrease in upper-level clouds (Fig. 4.1), this decrease in RMSE is larger for the experiments where infrared observations are assimilated. This result provides further evidence that the all-sky infrared brightness temperatures are able to improve the cloud field in the 1-h forecasts regardless of whether or not BC is applied to them prior to their assimilation.

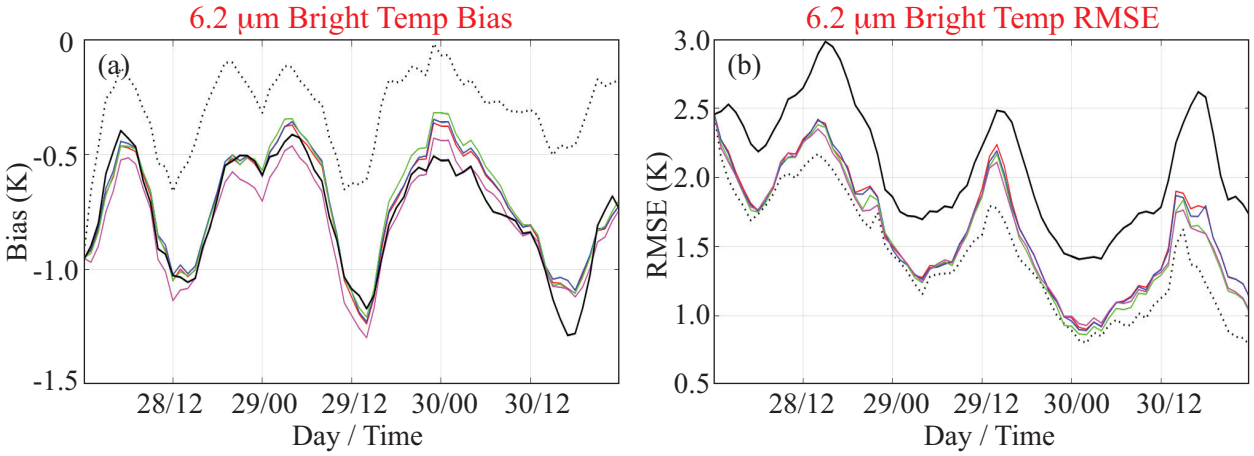

Figure 4.3: Time series showing the evolution of the SEVIRI 6.2 μm brightness temperature (a) bias (K) and (b) root mean square error (RMSE; K) computed using the ensemble mean prior analysis at hourly intervals from 00 UTC on 28 May 2014 to 00 UTC on 31 May 2014. Results are shown for the No-BC (dashed black line), OBSCTH-3RD (red line), OBSCTH-2ND (blue line), OBSCTH-1ST (green line), OBSCTH-0TH (magenta line), and No-Assim (solid black line) experiments.

4.4.1.3 Conventional Observation Error Analysis

To assess the impact of the nonlinear bias predictors on the thermodynamic and kinematic fields, Fig. 4.4 shows vertical profiles of RMSE for air temperature, relative humidity, and the zonal and meridional wind components computed using radiosonde observations accumulated over the 3-day assimilation period and binned into 100 hPa layers. For each variable, RMSE profiles are shown for the two baseline experiments (No-Assim and No-BC), followed by vertical profiles showing the percentage changes in RMSE for the remaining experiments computed with respect to each of the baseline experiments. This approach was used to make it easier to assess the impact of the bias predictors, while still being able to show the baseline error profiles. Negative (positive) changes mean that assimilation of the all-sky infrared observations decreased (increased) the errors relative to a given baseline experiment and therefore improved (degraded) the prior analysis fits to the radiosonde observations. Figure 4.5 shows the corresponding profiles of observation bias for each experiment. Only raw error profiles are shown for this metric because small biases

in the baseline experiments make the percentage changes difficult to evaluate.

Comparison of the temperature RMSE profiles for the baseline experiments reveals that the errors are up to 2% smaller (larger) in the upper (lower) troposphere when the all-sky observations are assimilated during the No-BC experiment (Fig. 4.4b). The RMSE and bias for the radiosonde temperatures were smaller below 400 hPa when the brightness temperature biases were removed during the OBSCTH experiments; however, the errors increased by several percent above this level (Fig. 4.4c, 4.5a). Because the largest BC is generally applied to lower brightness temperatures associated with clouds in the upper troposphere (e.g., Fig. 4.2), the larger errors near and above the tropopause indicate that removal of the brightness temperature bias may actually lead to some degradation in the fits to the radiosonde temperatures. The larger temperature errors occur during all of the OBSCTH experiments, however, which suggests that they may be related to removal of the mean brightness temperature bias rather than to removal of the conditional biases. It is also possible that some of the cloud and water vapor information from the all-sky satellite observations is being incorrectly aliased onto the temperature field. Further work is necessary to identify the cause of the larger temperature errors between 300 and 100 hPa.

For the relative humidity observations, the RMSE from the baseline experiments is relatively small near the surface, but increases rapidly to over 20% by 800 hPa. It then remains large in the middle troposphere before slowly decreasing with height in the upper troposphere (Fig. 4.4d). The bias profiles from the baseline experiments likewise indicate that the model background is too dry below 800 hPa, but too moist above this level (Fig. 4.5b). When all-sky brightness temperatures are assimilated during the No-BC experiment, the RMSE increases throughout most of the vertical profile (Fig. 4.4e), and the negative biases become even larger in the upper troposphere (Fig. 4.5b). Indeed, the relative humidity errors are larger in the No-BC experiment than they are in the No-Assim experiment despite the fact that the infrared observations are strongly sensitive to water vapor in the upper troposphere. As discussed previously, the negative conditional biases for brightness temperatures < 230 K indicate that the model background is deficient in upper level clouds (Fig. 4.2a). Thus, it appears that trying to add clouds more forcefully through assimilation of the non-bias-corrected observations leads to an incorrect aliasing of cloud information onto the water vapor field. Instead of increasing the cloud condensate in response to the negative OMB departures, the assimilation instead adds more water vapor to the model analyses. In contrast, both the RMSE and bias are greatly reduced when BC is applied to the infrared observations during the OBSCTH experiments (Figs. 4.4f, 4.5b). When combined with the brightness temperature statistics

shown in Fig. 4.3, this demonstrate that bias-correcting the all-sky infrared observations retains some cloud information during the assimilation while also improving the water vapor field.

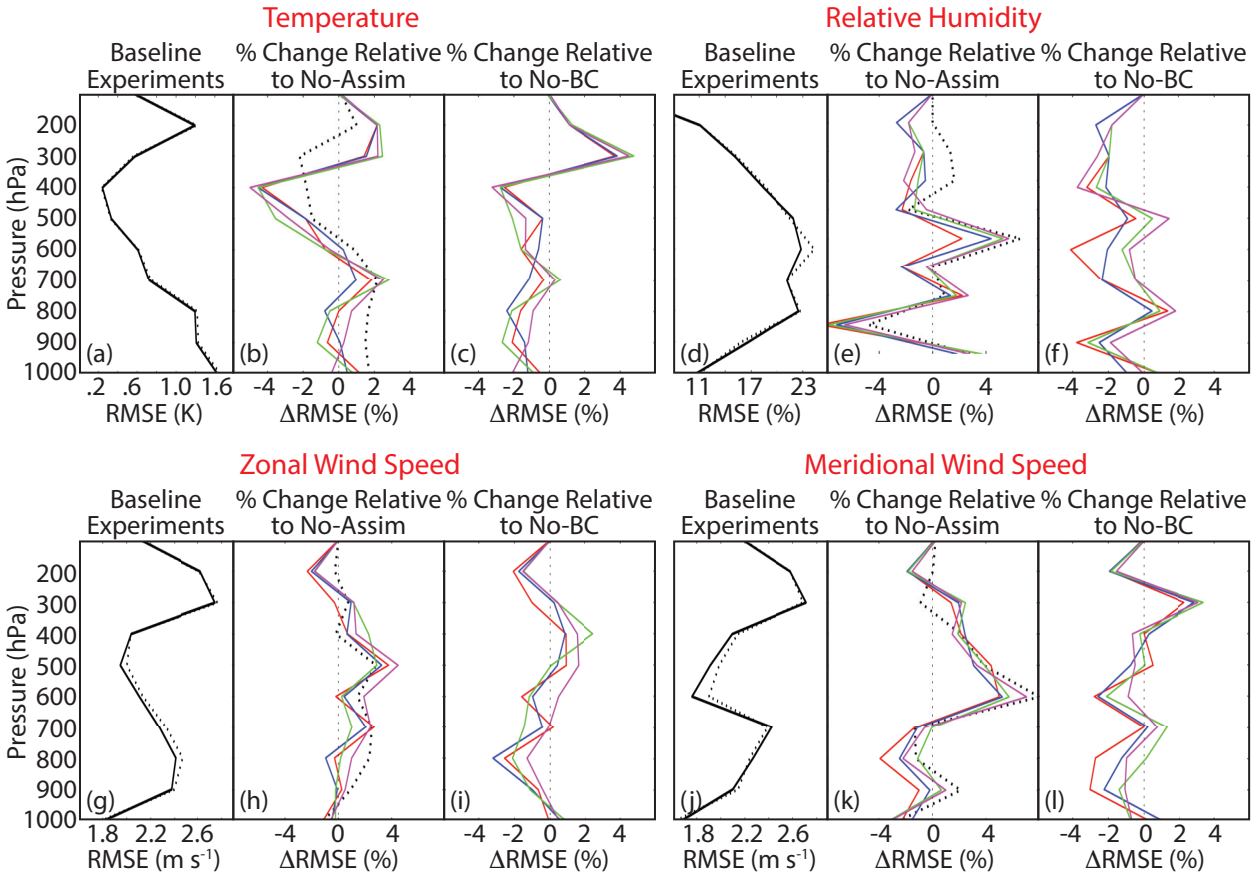

Figure 4.4: (a) Vertical profiles of temperature root mean square error (RMSE; K) from the No-Assim (black) and No-BC experiments (dashed black), with percentage changes in RMSE for the OBSCTH-3RD (red), OBSCTH-2ND (blue), OBSCTH-1ST (green), and OBSCTH-0TH (magenta) experiments relative to the No-Assim and No-BC experiments shown in panels (b) and (c). (d-f) Same as (a-c) except for showing vertical profiles of relative humidity RMSE (%). (g-i) Same as (a-c) except for showing vertical profiles of zonal wind speed RMSE (m s^{-1}). (j-l) Same as (a-c) except for showing vertical profiles of meridional wind speed RMSE (m s^{-1}). The error profiles were computed using data from the prior analyses over a 3-day period from 00 UTC on 28 May 2014 to 00 UTC on 31 May 2014.

For the zonal and meridional wind observations, the RMSE profiles from the baseline experiments have a sinusoidal appearance characterized by the largest errors in the lower and upper troposphere and smaller errors in the mid-troposphere (Fig. 4.4g, k). The biases in the baseline experiments are generally < 0.2 m s^{-1}, with the largest biases occurring near 600 and 700 hPa for the zonal and meridional wind components, respectively (Fig. 4.5c, d). The RMSE generally increases, especially for the meridional wind component, when the satellite observations are assimilated during the No-BC experiment (Fig. 4.4h, k). The wind errors are slightly reduced, however, when BC is applied to the infrared brightness temperatures during the OBSCTH exper-

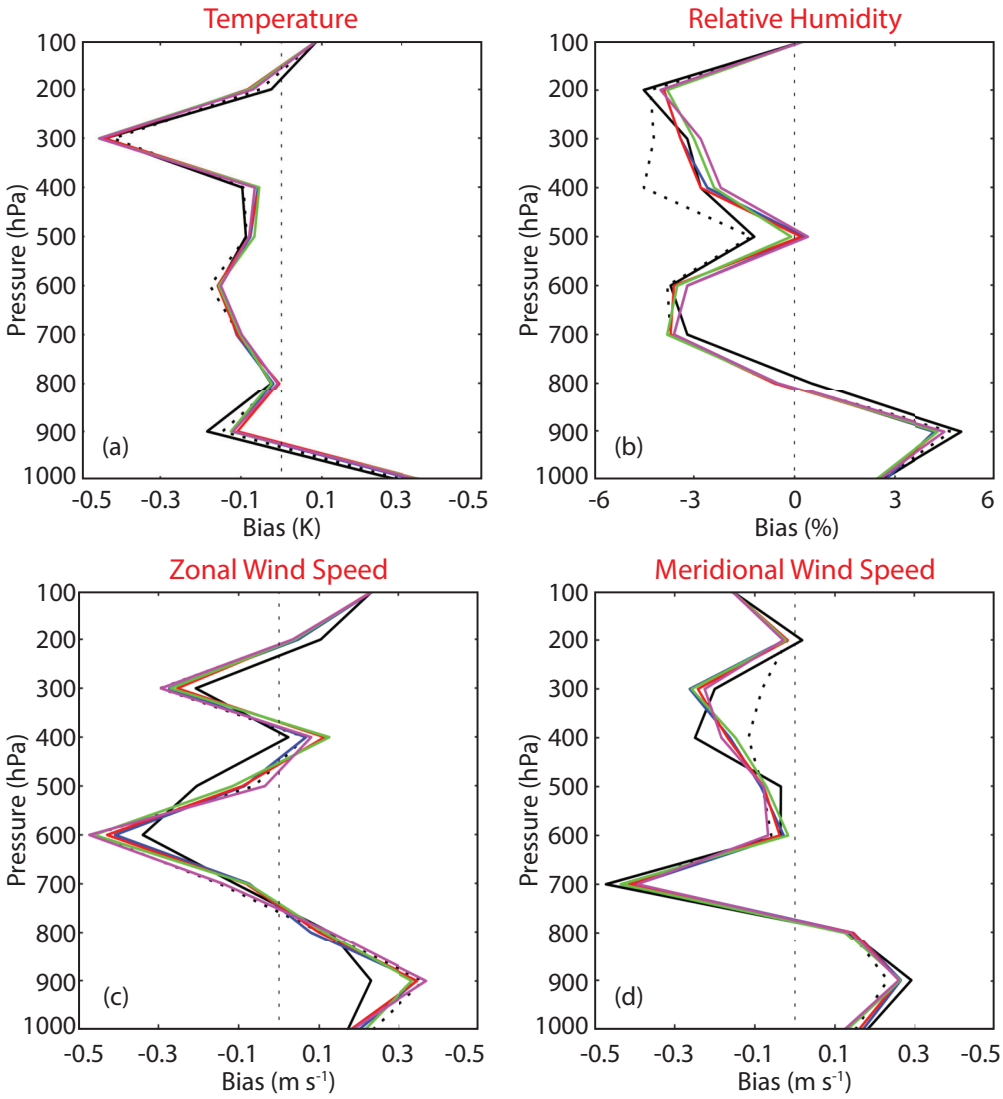

Figure 4.5: Vertical profiles of (a) temperature bias (K), (b) relative humidity bias (%), (c) zonal wind speed bias (m s^{-1}), and (d) meridional wind speed bias (m s^{-1}) for the No-Assim (solid black), No-BC (dashed black), OBSCTH-3RD (red), OBSCTH-2ND (blue), OBSCTH-1ST (green), and OBSCTH-0TH (magenta) experiments. The error profiles were computed using data from the prior analyses over a 3-day period from 01 UTC on 28 May 2014 to 00 UTC on 31 May 2014.

iments (Fig. 4.4i, l). Even so, it is evident that assimilation of the all-sky observations leads to a slightly negative impact on the mid-tropospheric winds and only a neutral to slightly positive impact in the lower troposphere and near the tropopause.

To more clearly assess the impact of the nonlinear BC predictors on each variable, summary statistics were computed using all of the radiosonde observations during the 72-h assimilation period. Table 4.1 shows the percentage changes in RMSE and bias for each OBSCTH experiment relative to the No-BC experiment. Overall, it is evident that bias-correcting the infrared brightness temperatures improves the quality of the model background fields. The largest improvements

(negative values) occur for the relative humidity field, with the bias reduced by at least 25% during each experiment. Smaller improvements occurred for the other variables. Comparison of the OBSCTH experiments reveals that there is a distinct advantage to using higher order BC terms to remove the bias from the all-sky brightness temperatures. For example, the RMSE for the relative humidity and wind observations steadily decrease as the BC predictor increases from the 0th (OBSCTH-0TH) to 3rd (OBSCTH-3RD) order. The impact of the higher order BC terms is less consistent for temperature and for the relative humidity bias; however, the errors are still smaller than occurred during the No-BC experiment. Together, the results presented in this section have shown that it is necessary to bias correct the infrared observations prior to their assimilation and that it is generally beneficial to include nonlinear BC predictors. This was demonstrated by the tendency for the higher order predictors to have a neutral-to-positive impact on the temperature and wind fields, while also improving the cloud and water vapor fields.

Table 4.1: Percentage changes in root mean square error (RMSE) and bias for the zonal and meridional wind speed, temperature, and relative humidity for the OBSCTH-0TH, OBSCTH-1ST, OBSCTH-2ND, and OBSCTH-3RD experiments relative to the No-BC experiment. The statistics were computed using all of the radiosonde observations and output from the prior ensemble mean analyses during the 72-h assimilation period.

	U	V	T		RH	
Experiment	RMSE	RMSE	RMSE	BIAS	RMSE	BIAS
OBSCTH-0TH - No-BC	-0.2%	-0.2%	-0.1%	-4.7%	-0.6%	-36.2%
OBSCTH-1ST - No-BC	-0.7%	-0.1%	-0.3%	-3.1%	-0.9%	-29.1%
OBSCTH-2ND - No-BC	-0.9%	-0.5%	-0.3%	-5.0%	-1.5%	-25.6%
OBSCTH-3RD - No-BC	-1.0%	-0.8%	-0.2%	-1.3%	-1.8%	-30.2%

4.4.2 Assessing the Impact of Different Bias Predictor Variables

In this section, we assess the ability of individual bias predictor variables sensitive to clouds and water vapor, or that depict variations in the satellite zenith angle, to improve the assimilation of all-sky infrared brightness temperatures during cycled DA experiments. Based on results from the previous section, all of the experiments employed a 3rd order polynomial expansion of the OMB departures to remove biases from the satellite brightness temperatures prior to their assimilation. In addition to the OBSCTH-3RD experiment presented in Section 4.4a (hereafter referred to as BC-OBSCTH), experiments were performed in which the observed SEVIRI 6.2 μm brightness temperatures (BC-OBSBT), satellite zenith angle (BC-SATZEN), or 100-700 hPa integrated water content (BC-IWC) were used as the bias predictors. The integrated water content predictor

was calculated by converting the water vapor and all cloud hydrometeor mixing ratios in each model layer into millimeters and then integrating over the 100-700 hPa layer. Together, these four predictors were chosen because they were also used during the passive monitoring experiments presented in Otkin et al. (2018). Here, we extend the results of that study by assessing the performance of these bias predictor variables when they are used during active DA experiments.

4.4.2.1 Observation Space Diagnostics

Figure 4.6 shows the evolution of the SEVIRI 6.2 μm brightness temperature bias, RMSE, ensemble spread, and consistency ratio (CR) for each experiment during the 3-day assimilation period. The statistics were computed for each assimilation cycle using brightness temperatures from the prior ensemble analyses. The ensemble spread is defined as:

$$\text{Spread} = \sqrt{\left\langle \frac{1}{N-1} \sum_{n=1}^{N} \left[H(\mathbf{x}_n) - \overline{H(\mathbf{x}_n)} \right]^2 \right\rangle}, \tag{4.4}$$

where N is the ensemble size, n is the index of a given ensemble member, and H is the observation operator (e.g., RTTOV) used to compute the model-equivalent brightness temperatures. The total ensemble spread is the combination of the observation error (σ_{obs}, set to 4 K) and ensemble spread, such that:

$$\text{Total Spread} = \sqrt{\sigma_{obs}^2 + \left\langle \frac{1}{N-1} \sum_{n=1}^{N} \left[H(\mathbf{x}_n) - \overline{H(\mathbf{x}_n)} \right]^2 \right\rangle}, \tag{4.5}$$

Finally, the RMSE and total spread are used to calculate the CR, which provides another diagnostic measure of the performance of the assimilation system:

$$\text{CR} = (\text{Total Spread})^2 / (\text{RMSE})^2 \tag{4.6}$$

With the CR, a value of 1 is desired because, in an ideal situation, the total spread should equal the RMSE for each observation type being assimilated. Values greater (less) than 1 indicate that there is too little (too much) ensemble spread and/or that the observation error is larger (smaller) than necessary (Dowell et al. 2004; Aksoy et al. 2009).

Inspection of the time series shows that the smallest RMSE and bias (Fig. 4.6a,b) occurred during the No-BC experiment, which is not surprising because assimilating non-bias-corrected observations should lead to the largest impact when assessed against themselves. Comparison

of the BC experiments reveals that the BC-SATZEN and BC-IWC experiments have larger biases and RMSEs than the BC-OBSBT and BC-OBSCTH experiments. The larger positive impact of the OBSBT and OBSCTH predictors on these two metrics is consistent with Otkin et al. (2018), who showed that variables sensitive to the cloud top height are more effective at identifying biases in all-sky infrared brightness temperatures. The results shown here indicate that using these predictors in active DA experiments also leads to smaller errors in the cloud and water vapor fields in the prior ensemble analyses when assessed using satellite observations.

The ensemble spread (Fig. 4.6c) generally decreases during the assimilation period due to a transition toward clearer skies and the cumulative impact of the all-sky brightness temperatures on the cloud and water vapor fields. The decrease in ensemble spread is accompanied by a corresponding increase in the CR (Fig. 4.6d), which peaks each morning when the RMSE reaches its diurnal minimum. Because the RMSE is smallest during the No-BC, BC-OBSCTH, and BC-OBSBT experiments (Fig. 4.6b), they also have the largest CRs. The large CR values during all of the active DA experiments reveal that it was sub-optimal to employ the same observation error variance for both clear and cloudy-sky observations during the entire assimilation period. Thus, combining an adaptive all-sky observation error model with the BC method would be beneficial; however, that is left for future work. In addition, inspection of rank histograms for each experiment (not shown) revealed that the ensemble spread is too small. This result points toward the need to also develop methods that increase the ensemble spread in cloud hydrometeors because they have the largest impact on the spread in the all-sky infrared brightness temperatures. One potential option would be to use the stochastic parameter perturbations method (Berner et al. 2017) to add perturbations to cloud source/sink terms to account for some of the uncertainty in cloud microphysics schemes. This has been shown to increase the spread in cloudy regions (Griffin et al. 2020).

4.4.2.2 Brightness Temperature Bias Correction Statistics

To further assess the behavior of each bias predictor, 2-D probability distributions of the ensemble mean BCs accumulated at hourly intervals during the 72-h assimilation period are shown for each experiment in Fig. 4.7. Overall, the BC-OBSBT and BC-OBSCTH experiments have similar distributions characterized by relatively small mean BCs for brightness temperatures > 230 K and then a strong upward trend in the mean BC for lower brightness temperatures (Fig. 4.7a,b). Even so, there are notable differences between these experiments, such as the larger BC for the lowest

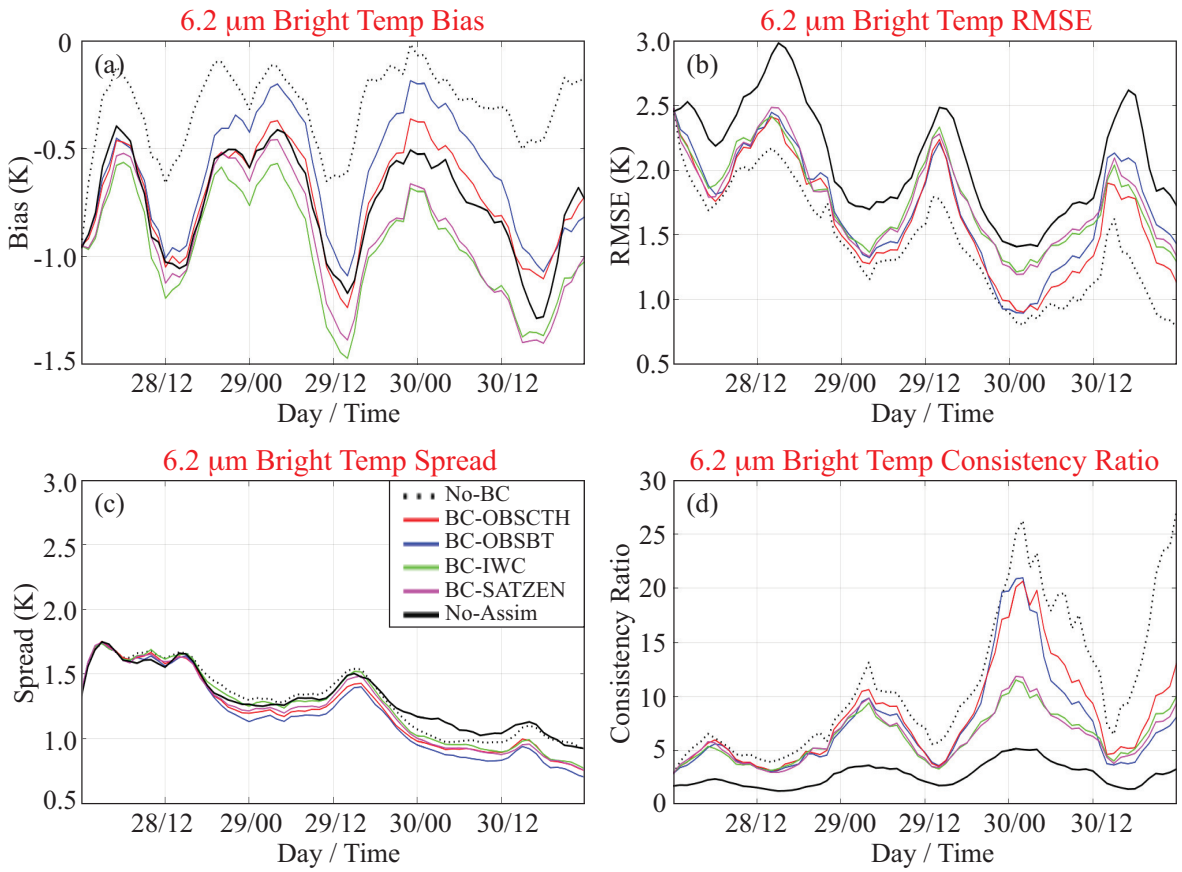

Figure 4.6: Time series showing the evolution of the SEVIRI 6.2 μm brightness temperature (a) bias (K), (b) root mean square error (RMSE; K), (c) spread (K), and (d) consistency ratio computed using the ensemble mean prior analysis at hourly intervals from 00 UTC on 28 May 2014 to 00 UTC on 31 May 2014. Results are shown for the No-BC (dashed black line), BC-OBSCTH (red line), BC-OBSBT (blue line), BC-IWC (green line), BC-SATZEN (magenta line), and No-Assim (solid black line) experiments.

brightness temperatures in the BC-OBSBT experiment and the wider vertical distribution for most brightness temperatures in the BC-OBSCTH experiment. The BC patterns for both experiments are flipped compared to the OMB departure distribution from the No-Assim experiment (Fig. 4.2a), which is good because that means that the OBSBT and OBSCTH predictors are able to account for the nonlinear, cloud-dependent conditional biases in that distribution. In contrast, the BC-IWC and BC-SATZEN experiments have much smaller BCs for the lowest brightness temperatures that then become larger for higher brightness temperatures. The mean BC is also larger during these experiments, which indicates that the IWC and SATZEN predictors did not have the same positive impact on the cloud field as the OBSBT and OBSCTH predictors. This behavior is consistent with the brightness temperature bias time series shown in Fig. 4.6a, and provides further evidence that it is necessary to use BC predictors sensitive to the cloud top height when assimilating all-sky infrared brightness temperatures.

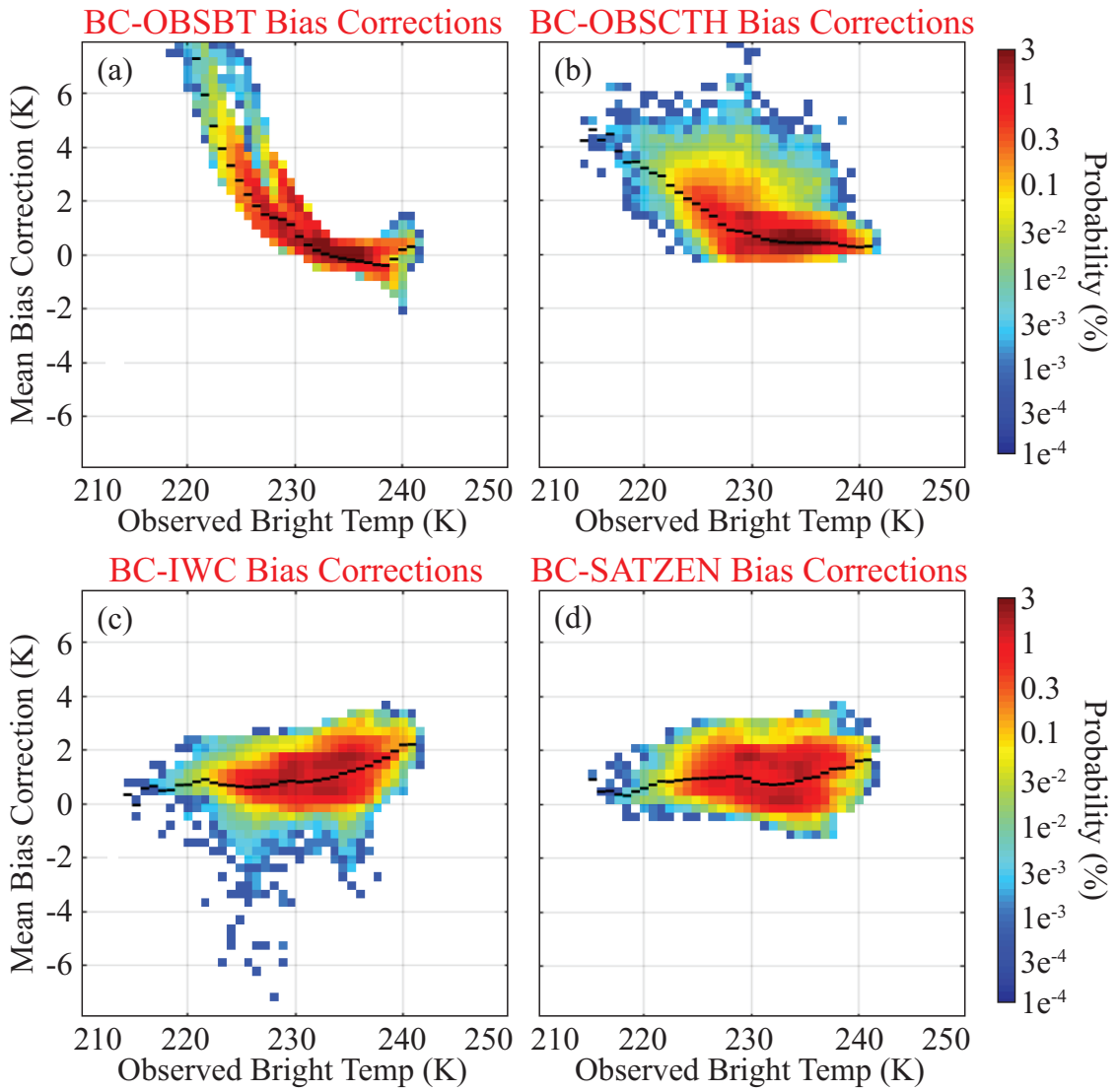

Figure 4.7: Probability distribution of SEVIRI 6.2 μm ensemble mean brightness temperature corrections (K) from the (a) BC-OBSBT, (b) BC-OBSCTH, (c) BC-IWC, and (d) BC-SATZEN experiments plotted as a function of the observed 6.2 μm brightness temperatures. Data were accumulated at hourly intervals during a 72-h period from 01 UTC on 28 May 2014 to 00 UTC on 31 May 2014. The horizontal black line segments represent the conditional bias in each column.

4.4.2.3 Brightness Temperature Innovations

Next, we examine the 6.2 μm brightness temperature innovations during each experiment using the 2-D probability distributions shown in Fig. 4.8. These distributions were constructed using the ensemble mean innovations accumulated at hourly intervals during the 72-h assimilation period. Inspection of Fig. 4.8a shows that the conditional mean innovations are close to zero across the entire distribution during the No-Assim experiment. This indicates that the conventional in-situ observations by themselves do not have a systematic impact on the cloud and water vapor fields in the upper troposphere. During the No-BC experiment (Fig. 4.8b), the innovation pattern

is very similar to the OMB departure distribution in the No-Assim experiment (Fig. 4.2a), with large (small) innovations occurring for lower (higher) brightness temperatures. This shows that the large conditional biases for the lower brightness temperatures are strongly corrected during this experiment, which is not surprising because BC was not applied to the brightness temperatures prior to their assimilation. A similar pattern emerges during the BC-IWC and BC-SATZEN experiments (Fig. 4.8e, f) because their smaller BCs for lower brightness temperatures (Fig. 4.7c, d) meant that large innovations were still possible during each assimilation cycle. In contrast, the mean innovations are very small across most of the distribution during the BC-OBSBT experiment (Fig. 4.8c) because the larger BCs for lower brightness temperatures (Fig. 4.7a) reduces the size of the resultant innovations. The distribution for the BC-OBSCTH experiment (Fig. 4.8d) has some larger negative innovations for the lower brightness temperatures, but is otherwise similar to the BC-OBSBT experiment. The smaller innovations during the BC-OBSBT and BC-OBSCTH experiments were likely beneficial because they limited potential imbalances in the model due to large analysis increments, while still leading to large reductions in the RMSE and bias (Fig. 4.6a, b).

4.4.2.4 Conventional Observation Error Analysis

Finally, we examine the impact of the infrared brightness temperatures and BC predictors on the accuracy of the prior ensemble mean analyses using OMB departure statistics accumulated during the 72-h assimilation period for the radiosonde temperature, relative humidity, and zonal and meridional wind observations. Figure 4.9 shows vertical profiles of RMSE for the No-Assim and No-BC experiments, along with percentage changes in RMSE for each BC experiment, whereas Fig. 4.10 shows the corresponding bias profiles. Summary statistics showing the percentage changes in RMSE and bias during each BC experiment relative to the No-Assim and No-BC experiments are shown in Tables 4.2 and 4.3, respectively.

Compared to the No-Assim experiment, the zonal and meridional wind speed errors in aggregate were slightly smaller during the BC-OBSBT and BC-OBSCTH experiments, but increased by $0.5-0.8\%$ during the BC-SATZEN and BC-IWC experiments (Table 4.2). Inspection of the zonal wind profiles (Fig. 4.9h, i) shows that the smaller RMSE during the BC-OBSBT and BC-OBSCTH experiments were primarily due to larger improvements in the upper and lower troposphere, with some degradation evident in the mid-troposphere. Both of these experiments also had the smallest meridional wind speed errors for most of the vertical layers (Fig. 4.9k, l). Indeed, the RMSE

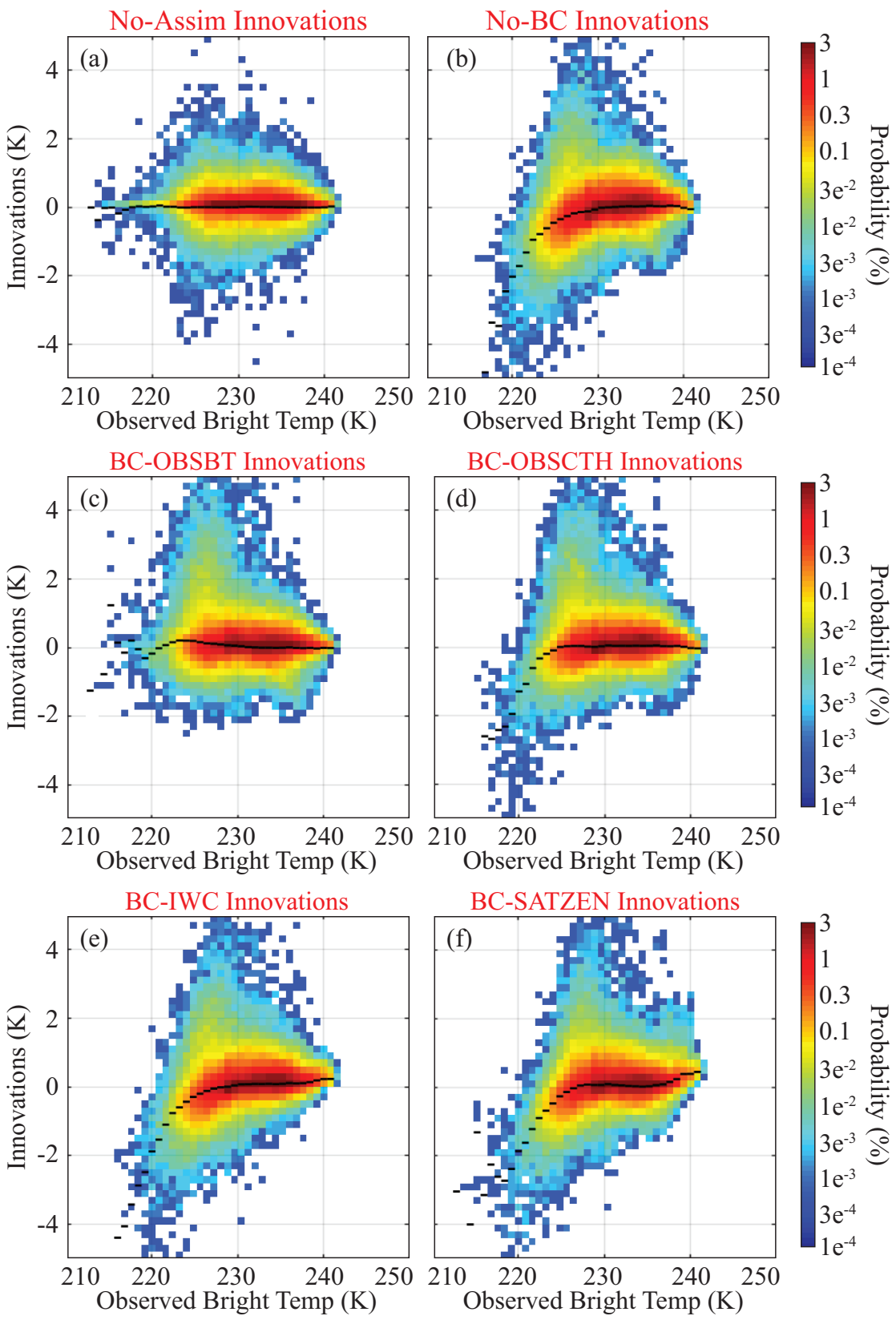

Figure 4.8: Probability distributions of SEVIRI 6.2 μm brightness temperature innovations (K) for the (a) No-Assim, (b) No-BC, (c) BC-OBSBT, (d) BC-OBSCTH, (e) BC-IWC, and (f) BC-SATZEN experiments plotted as a function of the observed 6.2 μm brightness temperatures (K). Data were accumulated at hourly intervals from 00 UTC on 28 May 2014 to 00 UTC on 31 May 2014. The black line segments depict the mean innovation in each column.

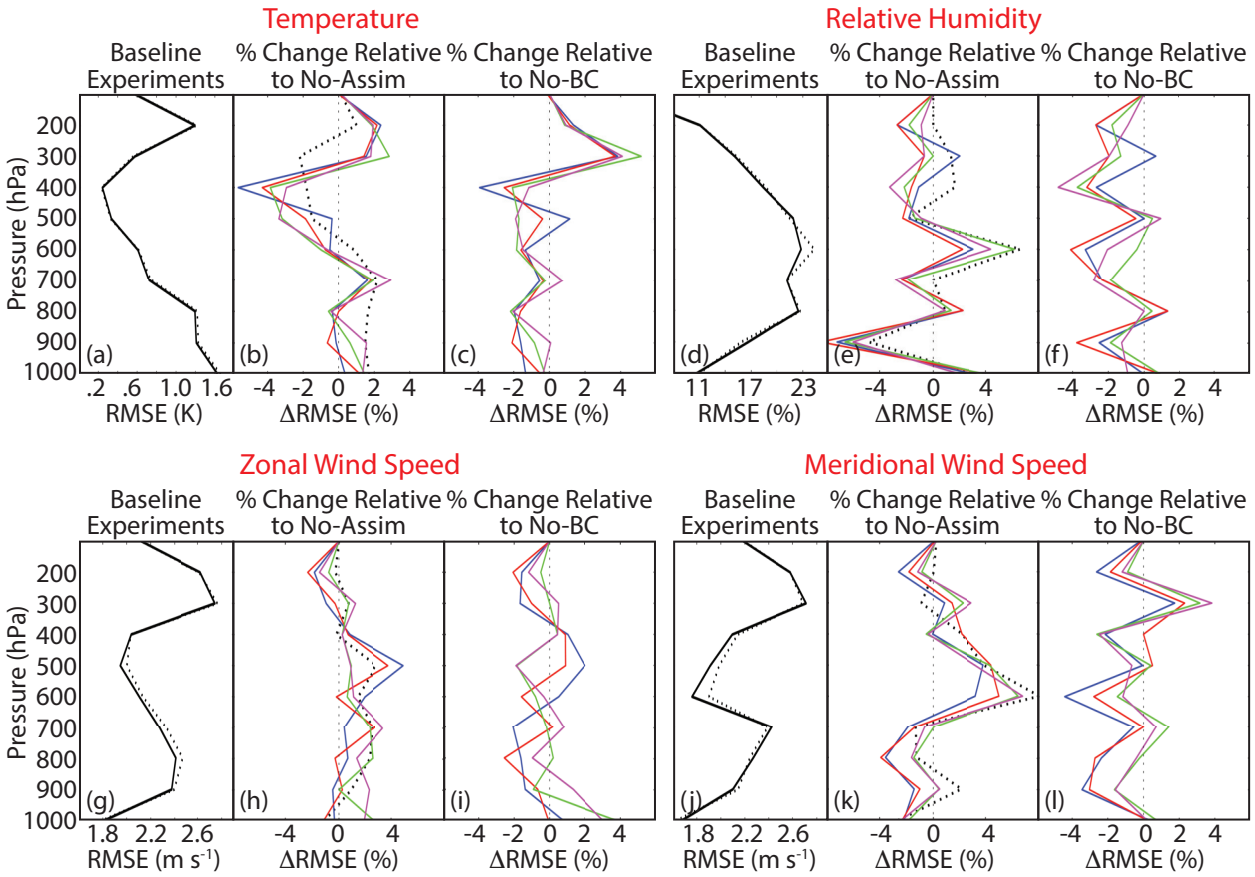

Figure 4.9: (a) Vertical profiles of temperature root mean square error (RMSE; K) from the No-Assim (solid black) and No-BC experiments (dashed black), with percentage changes in RMSE for the BC-OBSBT (blue), BC-OBSCTH (red), BC-IWC (green), and BC-SATZEN (magenta) experiments relative to the No-Assim and No-BC experiments shown in panels (b) and (c). (d-f) Same as (a-c) except for showing vertical profiles of relative humidity RMSE (%). (g-i) Same as (a-c) except for showing vertical profiles of zonal wind speed RMSE (m s^{-1}). (j-l) Same as (a-c) except for showing vertical profiles of meridional wind speed RMSE (m s^{-1}). The error profiles were computed using data from the ensemble mean prior analyses at hourly intervals over a 3-day period from 00 UTC on 28 May 2014 to 00 UTC on 31 May 2014.

for the meridional wind speed observations was 1.4% and 0.8% smaller during the BC-OBSBT and BC-OBSCTH experiments, respectively, compared to a neutral impact when the IWC and SATZEN predictors were used (Table 4.3).

Assimilation of the infrared brightness temperatures led to very different impacts on the RMSE and bias for the radiosonde temperature observations. For example, though the RMSE in each experiment increased by $0.8 - 1.0\%$ relative to the No-Assim experiment, the bias was substantially reduced, with decreases ranging from -1.7% during the No-BC experiment to -6.1% for the BC-SATZEN experiment (Table 4.2). Overall, the smallest biases were obtained during the various BC experiments, with all but BC-SATZEN also having slightly smaller RMSEs than

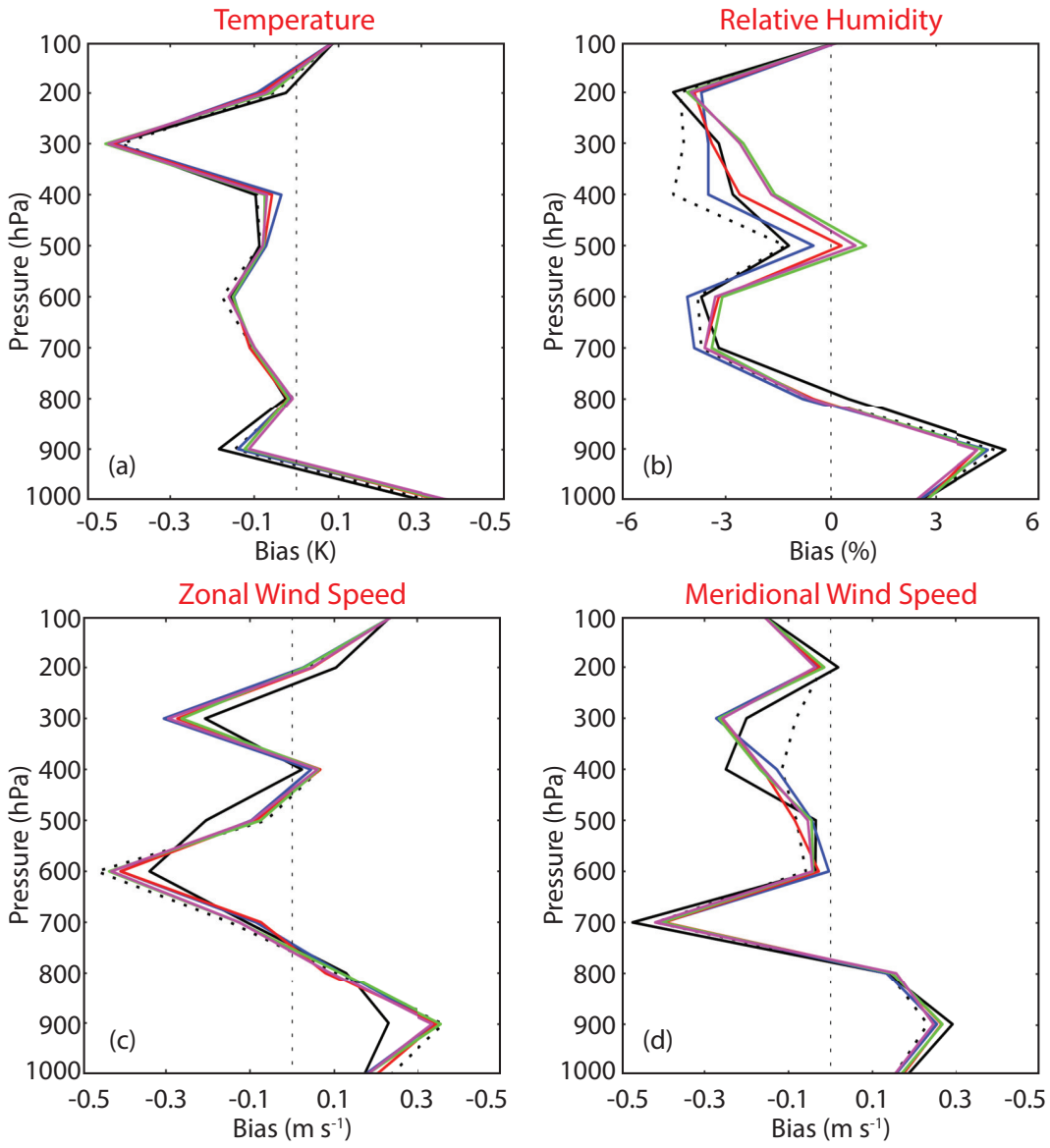

Figure 4.10: Vertical profiles of (a) temperature bias (K), (b) relative humidity bias (%), (c) zonal wind speed bias (m s^{-1}), and (d) meridional wind speed bias (m s^{-1}) for the No-Assim (solid black), No-BC (dashed black), BC-OBSBT (blue), BC-OBSCTH (red), BC-IWC (green), and BC-SATZEN (magenta) experiments. The error profiles were computed using data from the prior analyses over a 3-day period from 01 UTC on 28 May 2014 to 00 UTC on 31 May 2014.

the No-BC experiment (Table 4.3). Comparison of the vertical profiles shows that the temperature RMSEs were smaller within most of the troposphere during the BC experiments (Fig. 4.9c); however, the presence of much larger errors near the tropopause led to only a neutral to slightly positive impact when all of the temperature observations are considered (Table 4.3).

For relative humidity, assimilating the infrared brightness temperatures without BC led to sharply higher bias (30.1%) and RMSE (0.8%) during the No-BC experiment (Table 4.2). In contrast, the overall RMSE and bias were much smaller during the other experiments regardless of

which BC predictor was used (Table 4.3). Compared to the No-BC experiment, the largest RMSE reductions occurred during the BC-OBSCTH (-1.8%) and BC-SATZEN (-1.4%) experiments, with the largest bias reductions occurring during the BC-IWC (-45.2%), BC-SATZEN (-38.2%), and BC-OBSCTH (-30.2%) experiments. The error profiles in Fig. 4.9f show that, though there are some differences between the BC experiments, that the RMSEs are smaller in most of the troposphere relative to the No-BC experiment. The biases are also greatly reduced in the middle and upper troposphere (Fig. 4.10b).

In summary, the results presented in this section show that assimilation of infrared brightness temperatures that are not bias-corrected leads to larger errors for all metrics, except for the temperature bias, relative to the No-Assim experiment. Removal of the brightness temperature biases prior to their assimilation, however, greatly improves the impact of the satellite observations, with the largest percentage decreases in the errors realized for the relative humidity observations. Overall, the OBSCTH and OBSBT predictors were the most useful because not only did their use lead to more accurate cloud and water vapor fields, but they also produced the smallest RMSEs for the wind and temperature fields.

Table 4.2: Percentage changes in root mean square error (RMSE) and bias for the zonal and meridional wind speed, temperature, and relative humidity for the BC-OBSBT, BC-OBSCTH, BC-IWC, and BC-SATZEN experiments relative to the No-Assim experiment. The statistics were computed using all of the radiosonde observations and output from the prior ensemble mean analyses during the 72-hr assimilation period.

	U	V	T		RH	
Experiment	RMSE	RMSE	RMSE	BIAS	RMSE	BIAS
No-BC - No-Assim	0.9%	0.6%	1.0%	-1.7%	0.8%	30.1%
BC-OBSBT - No-Assim	0.0%	-0.8%	0.8%	-4.7%	-0.4%	9.8%
BC-OBSCTH - No-Assim	-0.1%	-0.2%	0.8%	-3.0%	-1.0%	-9.2%
BC-IWC - No-Assim	0.7%	0.6%	0.9%	-4.8%	-0.1%	-28.8%
BC-SATZEN - No-Assim	0.8%	0.5%	1.0%	-6.1%	-0.6%	-19.6%

Table 4.3: Percentage changes in root mean square error (RMSE) and bias for the zonal and meridional wind speed, temperature, and relative humidity for the BC-OBSBT, BC-OBSCTH, BC-IWC, and BC-SATZEN experiments relative to the No-BC experiment. The statistics were computed using all of the radiosonde observations and output from the prior ensemble mean analyses during the 72-hr assimilation period.

	U	V	T		RH	
Experiment	RMSE	RMSE	RMSE	BIAS	RMSE	BIAS
BC-OBSBT - No-BC	-0.9%	-1.4%	-0.2%	-3.1%	-1.2%	-15.6%
BC-OBSCTH - No-BC	-1.0%	-0.8%	-0.2%	-1.3%	-1.8%	-30.2%
BC-IWC - No-BC	-0.2%	0.0%	-0.1%	-3.2%	-0.9%	-45.2%
BC-SATZEN - No-BC	-0.1%	-0.1%	0.1%	-4.5%	-1.4%	-38.2%

4.4.3 Symmetric Bias Correction Predictors

In this section, we assess the impact of using "symmetric" predictors to remove the bias from all-sky infrared brightness temperatures. As discussed in the introduction, symmetric predictors that represent the average of an observed quantity and its corresponding model equivalent have been extensively used when developing all-sky observation error models. First introduced by Geer and Bauer (2011), symmetric predictors have been shown in various studies to lead to more Gaussian OMB departure statistics when a suitable cloud impact parameter is used to dynamically assign the error variance to each observation. This symmetric observation error approach is now widely used in operational DA systems that assimilate all-sky microwave radiances because it leads to more accurate forecasts through better utilization of the satellite observations.

Despite their widespread use in all-sky observation error models, it is not clear if symmetric variables can also serve as effective bias predictors, especially in the presence of complex nonlinear bias patterns. To explore their potential utility, two additional sets of experiments were run where the cloud top height or the 6.2 μm brightness temperatures were used as the bias predictor. These variables were chosen because they are either a direct measure of, or are sensitive to, the cloud height, which is an excellent measure of cloud impact in all-sky infrared brightness temperatures. Experiments were performed where observed (BC-OBSBT, BC-OBSCTH), simulated (BC-SIMBT, BC-SIMCTH), or symmetric (BC-SYMBT, BC-SYMCTH) quantities for each BC predictor variable were used to remove the bias from the infrared brightness temperatures prior to their assimilation. For the simulated cloud top height predictor, the cloud top was identified as the first model level looking downward from the model top in which the vertically-integrated cloud hydrometeor mixing ratio was $> 10^{-4}$ kg kg^{-1}. All of the cloud hydrometeor species predicted by the microphysics parameterization scheme were used when computing this quantity. The modeled land/ocean surface elevation was used as the predictor value when the accumulated cloud mixing ratio threshold was not surpassed. The same approach was used for the observed cloud top height retrievals where grid points identified as clear were also set to the model surface elevation.

Summary statistics showing the percentage changes relative to the No-BC experiment for the radiosonde temperature, relative humidity, and zonal and meridional wind speed observations are shown in Tables 4.4 and 4.5, respectively, for experiments using the various 6.2 μm brightness temperature or cloud top height quantities as the bias predictor. These statistics were computed using output from the prior ensemble mean analyses. Overall, the results show that using sym-

metric bias predictors does not lead to a more accurate model background. For experiments using the 6.2 μm brightness temperature predictors (Table 4.4), the error reduction for each radiosonde observation type is smaller during the BC-SYMBT experiment than it is during the BC-OBSBT experiment. Likewise, when the cloud top height quantities are used as the bias predictors (Table 4.5), the most accurate analyses are obtained when the observed quantity is used during the BC-OBSCTH experiment. The error reductions during the BC-SYMCTH experiment are either in between those obtained during the BC-OBSCTH and BC-SIMCTH experiments, or are smaller than both of them. A possible reason for the relatively poor performance during both of the symmetric bias predictor experiments is that, with the exception of relative humidity, the error reductions are consistently smaller when the simulated predictors are used to remove the bias from the all-sky infrared observations. Thus, inclusion of the model-simulated predictor value when computing the symmetric bias predictor is not beneficial. Instead, it is more effective to simply use the observed quantity as the bias predictor.

Table 4.4: Percentage changes in root mean square error (RMSE) and bias for the zonal and meridional wind speed, temperature, and relative humidity for the BC-OBSBT, BC-SYMBT, and BC-SIMBT experiments relative to the No-BC experiment. The statistics were computed using all of the radiosonde observations and output from the prior ensemble mean analyses during the 72-h assimilation period.

	U	V	T		RH	
Experiment	RMSE	RMSE	RMSE	BIAS	RMSE	BIAS
BC-OBSBT - No-BC	-0.9%	-1.4%	-0.2%	-3.1%	-1.2%	-15.6%
BC-SYMBT - No-BC	-0.1%	0.0%	-0.1%	-2.0%	-1.0%	-29.6%
BC-SIMBT - No-BC	1.0%	1.3%	0.6%	-1.1%	-0.8%	-55.8%

Table 4.5: Percentage changes in root mean square error (RMSE) and bias for the zonal and meridional wind speed, temperature, and relative humidity for the BC-OBSCTH, BC-SYMCTH, and BC-SIMCTH experiments relative to the No-BC experiment. The statistics were computed using all of the radiosonde observations and output from the prior ensemble mean analyses during the 72-h assimilation period.

	U	V	T		RH	
Experiment	RMSE	RMSE	RMSE	BIAS	RMSE	BIAS
BC-OBSCTH - No-BC	-1.0%	-0.8%	-0.2%	-1.3%	-1.8%	-30.2%
BC-SYMCTH - No-BC	-0.4%	-0.5%	0.0%	-3.1%	-1.2%	-27.1%
BC-SIMCTH - No-BC	-0.2%	0.5%	0.0%	-1.2%	-1.5%	-43.2%

To examine this behavior more closely, Fig. 4.11 shows 2-D probability distributions for the ensemble mean 6.2 μm brightness temperature BCs and innovations when the simulated, observed, and symmetric cloud top height bias predictors are used. Similar results are obtained for experiments employing the 6.2 μm brightness temperature predictors (not shown). Comparison

of the BC distributions reveals a relatively flat pattern during the BC-SIMCTH experiment (Fig. 4.11a), which shows that the model-derived cloud top height predictor is unable to account for the large negative conditional biases for brightness temperatures < 230 K (Fig. 4.2a). The smaller BCs for the lower brightness temperatures during this experiment stand in sharp contrast to the much larger BCs during the BC-OBSCTH experiment (Fig. 4.11e). Because the symmetric predictor is simply the mean of the observed and simulated quantities, the BC distribution during the BC-SYMCTH experiment (Fig. 4.11c) is a hybrid of the BC-OBSCTH and BC-SIMCTH distributions. As such, the smaller BCs for the lower brightness temperatures due to the impact of the model-simulated quantity leads to larger innovations than occurred during the BC-OBSCTH experiment (Fig. 4.11d, f). As was shown in the previous section, experiments containing larger innovations for the lower brightness temperatures associated with optically thick upper-level clouds were generally less accurate when assessed using radiosonde observations. This result suggests that, though symmetric predictors have been shown to improve the performance of all-sky observation error models, they may not work as well for all-sky BC. Further studies using other satellite bands and models are necessary to explore this in more detail.

4.5 Discussion and Conclusions

In this study, ensemble DA experiments were performed using the regional-scale KENDA system to evaluate the ability of all-sky infrared brightness temperatures to improve the accuracy of the ensemble prior analyses used during each assimilation cycle. Observations from the 6.2 μm band on the SEVIRI sensor were assimilated at hourly intervals over a 3-day period in May 2014. This infrared band is primarily sensitive to clouds and water vapor in the upper troposphere. Various experiments were performed in which different BC predictors were used to remove biases from the all-sky brightness temperatures prior to their assimilation. Results from these BC experiments were compared to baseline experiments in which the brightness temperatures were either not assimilated (No-Assim) or were assimilated without first removing their biases (No-BC). This study builds upon the passive monitoring experiments described in Otkin et al. (2018) by exploring the impact of linear and nonlinear BC predictors during experiments in which all-sky infrared brightness temperatures are actively assimilated.

Overall, inspection of the 6.2 μm brightness temperature OMB departure distribution from the No-Assim experiment revealed that the conditional biases exhibited a nonlinear pattern char-

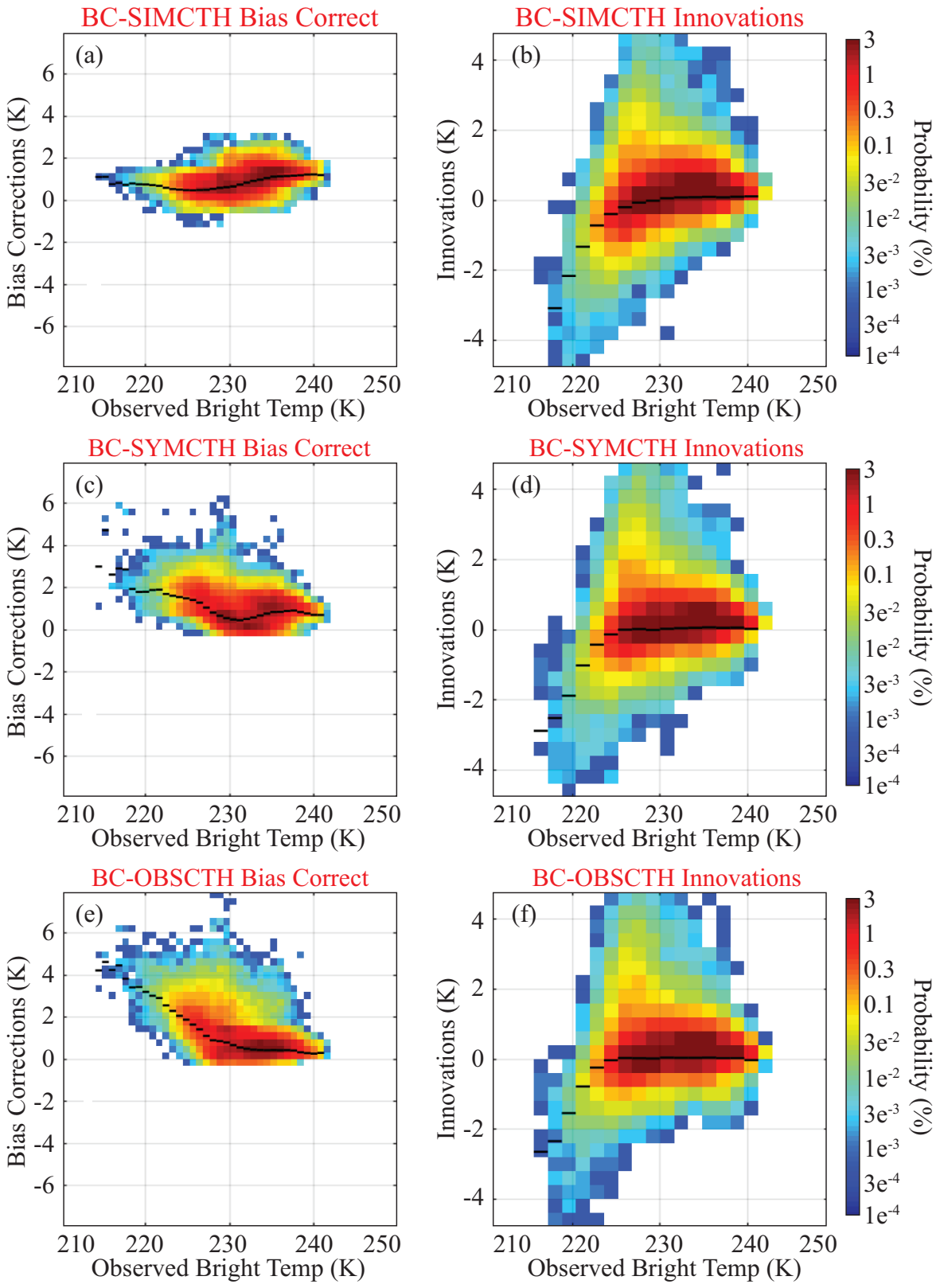

Figure 4.11: Probability distributions for the SEVIRI 6.2 μm brightness temperature (a) bias corrections and (b) innovations from the BC-SIMCTH experiment plotted as a function of the observed 6.2 μm brightness temperatures (K). (c-d) Same as (a-b), except for the BC-SYMCTH experiment. (e-f) Same as (a-b), except for the BC-OBSCTH experiment. Data were accumulated at hourly intervals from 00 UTC on 28 May 2014 to 00 UTC on 31 May 2014. The black line segments depict the mean bias correction or innovation in each column.

acterized by small biases for higher brightness temperatures and increasingly large negative biases for lower brightness temperatures. Though the negative conditional biases are likely at least partially due to inaccuracies in the forward observation operator, they also indicate that the model analyses do not contain enough cloud condensate in the upper troposphere. This deficiency, whether due to insufficient spatial coverage or cloud optical depth, represents a systematic bias in the NWP model depiction of the cloud field. Thus, trying to add these upper-level clouds during an assimilation cycle could be problematic because of aliasing of the cloud information onto other model state variables and the tendency for the model to revert back to its preferred state during the subsequent forecast period.

Evaluation of the No-BC experiment showed that assimilation of the infrared brightness temperatures without first removing their biases almost always degraded the accuracy of the ensemble prior analyses based on larger OMB departures for the radiosonde observations. In particular, the summary statistics showed that the relative humidity bias and RMSE were much larger during this experiment than they were during the No-Assim experiment. Despite having strong sensitivity to water vapor in the upper troposphere, assimilating infrared brightness temperatures without BC actually increased the relative humidity RMSE, primarily because of a large increase in the moist bias already present in the No-Assim experiment. The No-BC experiment was also characterized by smaller 6.2 μm brightness temperature OMB departures, which suggests that instead of adding clouds to the analysis, the DA system instead added more water vapor. An alternative explanation is that a portion of the cloud condensate added to the ensemble posterior analyses during a given assimilation cycle evaporated during the subsequent model integration period, thereby increasing the moist bias. Regardless, this result suggests that the analyses were being too strongly constrained by the all-sky infrared brightness temperatures during the No-BC experiment in situations where the model was unable to properly handle the additional cloud information.

The subsequent removal of linear and nonlinear conditional biases from the all-sky brightness temperatures through use of a 3rd order polynomial expansion of the OMB departures and various BC predictors led to smaller errors for all of the radiosonde observation types when compared to the No-BC experiment. The largest improvements occurred for the relative humidity observations where the moist bias in the upper troposphere was greatly reduced. Notable improvements also occurred in the temperature bias and in the RMSE for the zonal and meridional wind speed components during the BC-OBSBT and BC-OBSCTH experiments. The temperature

RMSE was also smaller in most of the troposphere; however, a spike of larger errors near and above the tropopause led to a neutral impact when all temperature observations were considered.

Comparison of the various predictors showed that those sensitive to the location of the cloud top had the largest positive impact on the model background based on improved fits to the radiosonde observations. The observed cloud top height and observed 6.2 μm brightness temperature predictors were the best overall because their use not only led to the smallest relative humidity errors, but also led to the largest error reductions for the zonal and meridional wind speed observations and the smallest degradation for the temperature RMSE. Both of these predictors were also characterized by smaller brightness temperature RMSE and bias. The larger improvements during the BC-OBSBT and BC-OBSCTH experiments were primarily due to the ability of the cloud-sensitive predictors to more effectively remove the large negative biases from brightness temperatures < 230 K. The larger BCs for these clouds then led to smaller brightness temperature innovations and presumably fewer model spin-up problems during the subsequent 1-h forecasts. Additional experiments using the OBSCTH predictor revealed that it was beneficial to use higher order nonlinear BC terms to remove the bias from the all-sky infrared brightness temperatures. For example, the radiosonde OMB departure errors generally decreased as the order of the polynomial expansion increased from the 0th order to the 3rd order. Finally, an additional set of experiments showed that symmetric bias predictors do not improve the model analyses as effectively as the observed predictors do by themselves. This suggests that, though symmetric predictors have proven utility for all-sky observation error models, they may not be as useful when developing all-sky BC methods.

This study has shown that assimilation of all-sky infrared brightness temperatures substantially improves the accuracy of the cloud and water vapor fields in the prior ensemble analyses when cloud-sensitive predictors and higher order BC terms are used to remove linear and nonlinear conditional biases from the observations prior to their assimilation. Though encouraging, additional studies are necessary to evaluate the ability of the NBC method and the all-sky infrared brightness temperatures to improve the model analyses during other seasons containing different cloud regimes potentially characterized by different conditional bias patterns. It will also be necessary to perform ensemble forecasts to evaluate how long the improved cloud and water vapor fields persist during the forecast period. It is important to note that the experiments performed during this study are only an initial step toward inclusion of the all-sky infrared observations in the KENDA system and that additional developments have the potential to substantially increase

their impact. For example, there is great promise in pairing the BC method to a dynamic all-sky observation error model because that could lead to more effective use of the clear- and cloudy-sky brightness temperatures. It would also be helpful to explore the benefits of more frequent assimilation updates and in assimilating brightness temperatures from more than one infrared band, though that would require development of a correlated observation error model. The results also suggest that some attention should be given to developing methods that can increase the ensemble spread in the cloud hydrometeor variables. These topics are all left to future work.

4.6 Acknowledgements

We gratefully acknowledge Jesse Stroik from the University of Wisconsin-Madison and Hendrik Reich, Andreas Rhodin, Robin Faulwetter, and Axel Hutt from the German DWD for their assistance porting and installing the KENDA system and basic cycling (BACY) scripts to the NOAA/NESDIS/STAR "S4" supercomputer located at the University of Wisconsin-Madison. The S4 supercomputer (Boukabara et al. 2016) was used to perform all of the cycled DA experiments. The lead author was partially supported by a University of Reading International Research Studentship.

Chapter 5

Model Bias Estimation – Lorenz-63 Model Experiments

5.1 Abstract

In this study, we develop dynamic model bias estimators based on an asymptotic expansion of the model dynamics for small time scales and small perturbations in a model parameter, and then use the estimators to improve the performance of a data assimilation system. We employ the well-known Lorenz (1963) model so that we can study all aspects of the dynamical system and model bias estimators in a detailed way that would not be possible with a full physics numerical weather prediction model. In particular, we first work out the asymptotics of the Lorenz model for small changes in one of its parameters and then use statistics from cycled data assimilation experiments to demonstrate that the asymptotics accurately represent the behavior of the model and that the coefficients of the nonlinear asymptotical expansion can be reasonably estimated by solving a least squares minimization problem.

In data assimilation, the background error covariance matrix usually estimates the uncertainty of the model background, which is then used along with the observation error covariance matrix to produce an updated analysis. If the uncertainty of the model background is strongly influenced by time-dependent model errors, then the development of nonlinear bias estimators that also vary with time could improve the performance of the assimilation system and the accuracy of the updated analysis. We demonstrate this improvement through the combination of a

constant background error covariance matrix with a dynamically-varying matrix computed using the model bias estimators. Numerical tests using the Lorenz (1963) model illustrate the feasibility of the approach and show that it leads to clear improvements in the analysis and forecast accuracy.

5.2 Introduction

Partial differential equations are widely used in scientific and technological fields to simulate the evolution of natural phenomena. For initial boundary condition problems such as those that are commonly encountered in atmospheric science, an accurate prediction of the spatial and temporal characteristics of various weather and climate features depends not only on the ability of a numerical model to realistically simulate the physical processes controlling their evolution, but also on the ability of a data assimilation system to provide the forecast model an accurate estimate of the initial conditions. Atmospheric data assimilation systems typically combine information from a short-range model forecast, or ensemble of forecasts, with a set of observations gathered over a specified time period to produce an analysis of the current state of the dynamical system that then serves as the initial conditions for the next model forecast. Commonly used data assimilation methods include variational assimilation that determines the analysis through minimization of a cost function, ensemble methods that use an ensemble of forecasts to dynamically estimate the sample covariances between different state components when determining how new observations impact the ensemble analysis, and so-called hybrid methods that combine aspects of variational and ensemble data assimilation methods. A wide range of literature is known today introducing and studying different data assimilation methods, see for example Lorenc et al. (2000), Kalnay (2003), Evensen (2009), Anderson and Moore (2012), van Leeuwen et al. (2015), Reich and Cotter (2015), Kleist et al. (2009), Nakamura and Potthast (2015), Houtekamer and Zhang (2016), and Bannister (2017).

Regardless of which assimilation methodology is employed, generation of the best possible analysis state $x^{(a)}$ through combination of the model first guess or background state $x^{(b)}$ with the available observations requires knowledge of the observation error and the underlying uncertainty in the model background $x^{(b)}$. The observation error uncertainty is usually determined by the covariance matrix $R \in \mathbb{R}^{m\times m}$ of the observation vector $y \in \mathbb{R}^m$ in observation space $\mathbb{R}^m$, where $m \in \mathbb{N}$ denotes the number of scalar observations. The uncertainty of the model background state $x^{(b)}$ is measured by the covariance matrix $B \in \mathbb{R}^{n\times n}$, where $n \in \mathbb{N}$ is the dimension of the model

space $\mathbb{R}^n$ and $x^{(b)} \in \mathbb{R}^n$. Variational data assimilation methods calculate the analysis state $x^{(a)} \in \mathbb{R}^n$ by minimizing the functional [1]

$$J(x) := \left\| x - x^{(b)} \right\|^2_{B^{-1}} + \|y - H(x)\|^2_{R^{-1}}, \;\; x \in \mathbb{R}^n, y \in \mathbb{R}^m \tag{5.1}$$

where $H : \mathbb{R}^n \to \mathbb{R}^m$ is the forward observation operator that maps the model state x into the simulated observation $H(x) \in \mathbb{R}^m$. For linear observation operators, it is well-known (c.f. Nakamura and Potthast (2015), Chapter 5) that the minimization of (5.1) is given by

$$x^{(a)} = x^{(b)} + BH^T(R + HBH^T)^{-1}(y - H(x^{(b)})). \tag{5.2}$$

Because the model background and observations are not perfect, accurate knowledge of the covariance matrices B and R is very important for data assimilation as they determine the weights that are applied to the model background and observations, respectively, when generating the updated analysis $x^{(a)}$. In addition, the matrix B spreads information spatially within a region surrounding the observation location and can also be used to add balance constraints between analysis variables based on physical principles (Bannister 2008a, b).

Despite its importance, an exact form for B cannot be determined for real-world applications because the true state of the dynamical system cannot be completely known due to a limited number of observations and the presence of errors in the observations that are available. For variational assimilation systems, the model background error covariances are often computed using the so-called National Meteorological Center (NMC) method that was first described by Parrish and Derber (1992). This method estimates B using differences between forecasts of different lengths valid at the same time. For example, forecast errors could be assessed by examining differences between 24 and 48 hour forecasts from model integrations initialized one day apart. These difference fields are usually obtained from a large collection of model forecasts covering time periods of a month or longer. As such, the NMC method generates a climatological estimate of B that may not properly represent the true model errors on any given day due to changes in the atmospheric conditions. Because of this, some operational weather forecasting centers have developed new methods to generate B. One approach is to use an ensemble of data assimilations (EDA) where an ensemble of reduced-resolution data assimilation cycles is performed in which the observations and model are

[1] Please note that the matrices and vectors in this chapter are denoted using italics font because that is what was used in the submitted journal article that is the basis of this chapter.

perturbed in some manner. A theoretical analysis by Isaksen et al. (2010) has shown that if the perturbations are drawn from the true distributions of observation and model errors, that the spread in the resultant EDA analyses about the unperturbed control analysis will be representative of the background error. This approach has the advantage of introducing some flow-dependency to the B matrix, thereby allowing it to better capture the errors of the day (Buehner et al. 2005, Isaksen et al. 2010, Raynaud et al. 2011).

Ensemble data assimilation systems such as the ensemble Kalman filter (EnKF) (e.g. Evensen 1994, 2009; Houtekamaer and Mitchell 1998, 2001, 2005; Evensen and van Leeuwen 2000; Anderson 2001; Whitaker and Hamill 2002; Snyder and Zhang 2003; Houtekamer et al. 2005) on the other hand re-compute B for each assimilation step using an estimator based on output from an ensemble $x^{(b,\ell)}$ of forecasts valid at the current analysis time, where $\ell = 1, ..., L$, and L is the size of the ensemble. For most applications, the standard stochastic estimator

$$B := \frac{1}{L-1} \sum_{\ell=1}^{L} (x^{(b,\ell)} - \bar{x}^{(b)})(x^{(b,\ell)} - \bar{x}^{(b)})^T, \qquad \bar{x}^{(b)} := \frac{1}{L} \sum_{\ell=1}^{L} x^{(b,\ell)}, \tag{5.3}$$

is used to compute the first guess ensemble mean $\bar{x}^{(b)}$ and the background error covariance matrix B. The stochastic estimator includes the uncertainty of the previous model analysis propagated to the current analysis time. Because the forecast model is an approximation of the real dynamical system, the distribution of the first guess ensemble could be sub-optimal due to the impact of systematic errors on the ensemble mean and ensemble spread. Similar problems can arise when using the NMC method because in situations where the model error varies with time, the forecast differences used to compute the covariances in B will include the dynamically-varying model bias. This could result in incorrect statistical relationships between the model variables. Both of these outcomes are not desirable because the inclusion of systematic model errors when generating B can degrade the accuracy of the posterior analysis $x^{(a)}$ obtained during the data assimilation step.

Various studies have focused on improving methods to estimate the background error covariances used by modern data assimilation systems; however, accounting for model error is challenging because of the large size of geophysical models (Dee 1995). One approach is to add perturbations to a subset of the model variables, such as temperature, at the initialization time, whereas another technique adds random perturbations to specific parameters in the parameterization schemes used to simulate sub-grid scale processes during each model time step. The goal with both approaches is to increase the range of possible forecast solutions to realistically address

the impact of systematic model errors and the underlying uncertainties in the parameterization schemes. Substantial research has been directed toward development of these methods, which have the potential to greatly improve the performance of assimilation systems (Buizza et al. 1999; Weisheimer et al. 2014; Romine et al. 2014; Ha et al. 2015; Shutts 2015; Berner et al. 2017). As a corollary to the above approaches, other studies have shown that the detrimental impact of systematic model errors in ensemble data assimilation systems can be reduced by using different parameterization schemes in each ensemble member (Meng and Zhang 2007; Fujita et al. 2007).

Another approach widely used in ensemble data assimilation systems to increase the ensemble spread is to apply additive or multiplicative covariance inflation during the assimilation step. Some amount of covariance inflation is often necessary because the rank deficiency of the system can lead to an underestimation of the ensemble variance and because systematic model errors can cause the model background $x^{(b)}$ to deviate greatly from reality. This in turn can lead to so-called filter divergence where the model analyses can no longer be pulled toward the observations during the data assimilation step (Houtekamer and Zhang 2016). In the case of additive covariance inflation, the impact of the unknown model error is treated by drawing random perturbations from some distribution and then adding them to either the model background $x^{(b)}$ or to the model analysis $x^{(a)}$. With multiplicative covariance inflation, the ensemble spread for selected model variables is multiplied by a real number to achieve the desired ensemble spread. Both methods have some adaptivity because observation-minus-background (OMB) statistics are used to estimate how much inflation is necessary. There is a very active community working on these approaches, see for example Hamill and Whitaker (2005), Anderson (2007, 2009), Houtekamer et al. (2009); Li et al. (2009a, b), Miyoshi (2011), and Whitaker and Hamill (2012).

Model error has often been ignored in variational data assimilation systems because it is difficult to quantify and has been viewed as having a minor impact compared to random errors in the initial conditions and systematic errors in the observations (Carrassi and Vannitsem 2010). Unlike ensemble assimilation systems where the background error covariance matrix B is dynamically estimated for each assimilation cycle using the ensemble output, additional statistical or dynamical assumptions are generally necessary when creating these estimates for variational systems. Studies by Derber (1989), Zupanski (1997), Vidard et al. (2004), and Tremolet (2006, 2007) have shown that treating the model error as part of the state estimation problem substantially improves the accuracy of the state estimates. Theoretical model error frameworks were developed by Griffith and Nichols (2000), Nicolis (2003, 2004) and Nicolis et al. (2009) based on the behavior of model errors

in deterministic models. These frameworks were then used by Carrissi and Vannitsem (2010) to derive evolution equations for the model error covariances and correlations that address errors due to parameterization schemes.

The desire to properly account for model error also underpins recent efforts to move from "strong-constraint" 4-dimensional variational systems that assume the forecast model is perfect to "weak-constraint" systems that include some estimate of the model error. This concept was introduced 50 years ago by Sasaki (1970), however, it was not implemented in a full-physics forecast model for several decades because of the lack of information with which to define and solve the problem and the computational burden associated with inverting the model error covariance matrix along with the other matrices already included in the strong-constraint formulation (Ngodock et al. 2017). The basic premise behind the weak-constraint approach is that it is sufficient to only approximately satisfy the model equations because they are not exact anyway due to incomplete knowledge of the physical processes being modeled or the need to simplify the governing equations due to computational limitations. Despite the challenges associated with implementing weak-constraint systems, their use has generally led to more accurate model analyses and forecasts when compared to strong-constraint systems due to the higher number of degrees of freedom. As such, they are becoming more widely used in variational data assimilation systems (Tremolet 2007; Lindskog et al. 2009; Ngodock et al. 2017). A recent study by Howes et al. (2017) has also shown that model errors can be accounted for in strong-constraint systems by allowing errors in both the model and the observations when considering the statistics of the innovation vector. They demonstrate that a more accurate estimate of the model state can be obtained when the combined model and observation error statistics are used instead of the standard observation-only error statistics.

In this paper, we seek to extend our understanding of how to identify and treat model bias in modern data assimilation systems. Key tasks of this research include: 1) studying the behavior of model errors in a nonlinear dynamical system, 2) developing nonlinear conditional model bias estimators using the observations and the model first guess and analysis states, and 3) employing these estimators during variational data assimilation experiments to assess their ability to improve the performance of the system. Numerical experiments are performed using the Lorenz-63 (L63) model (Lorenz 1963), which is a well-known and popular study object within the data assimilation and dynamical systems communities.

We begin by carrying out an asymptotic analysis of the L63 model when one of its parameters, in this case, the normalized Rayleigh number ρ, varies with time. In the L63 model, the ρ parameter is usually set to a constant value; however, we allow it to vary with time in order to introduce a model bias. This is accomplished through use of a coupled version of the L63 model where a background or "hidden" system $S2$ is used to control how the $\rho = \rho(t)$ parameter changes with time in the "primary" system $S1$ that is used to represent the truth. Though we chose to focus on variations in the ρ parameter during this study, the approach works in the same way for the other L63 model parameters. We then develop a nonlinear model bias estimator method based on the initial ideas discussed in Otkin et al. (2018) where the bias estimator is formulated as a polynomial expansion of the model variables and the coefficients of this expansion are determined by solving a least-squares minimization problem. The ability of this method to dynamically estimate the model error contribution to the matrix B and to improve the resultant OMB statistics is demonstrated by carrying out an experiment where B is represented as the sum of static and dynamically-varying components. Finally, we demonstrate the feasibility and potential utility of the asymptotic expansion and nonlinear bias estimation method by running numerical experiments using a 3-dimensional variational (3DVAR) data assimilation system and a coupled version of the L63 model.

A description of the coupled L63 modeling system and derivation of the model asymptotics are provided in Section 5.3. The utility of dynamically estimating the model background error covariance matrix B is discussed in Section 5.4, along with development of nonlinear conditional model error estimators. We then perform various numerical experiments using the L63 model in Section 5.5, first demonstrating the validity of the asymptotic expansion of the nonlinear model error estimators in Section 5.5.1. This is followed by a study of the optimality of the fixed and dynamic components of the B matrix used during data assimilation and then a study of the estimation of the nonlinear model error dynamics based on the first guess minus analysis statistics. Results from these sections will demonstrate the feasibility of using methods developed during this study to estimate nonlinear model errors without any prior knowledge or assumptions regarding the form of the model dynamics. Conclusions are presented in Section 5.6.

5.3 Estimating System Bias

5.3.1 Coupled Lorenz 1963 Model

We want to use a relatively simple atmospheric model to assess the behavior of nonlinear model biases and to develop ways to take into account those biases in a way that is complex enough to represent nonlinear atmospheric processes while being simple enough to provide insight into the nonlinear behavior of the system. To accomplish this goal, we have chosen to employ the L63 model (Lorenz 1963), which is widely used within the atmospheric data assimilation community (see for example Evensen 1997; Verlaan and Heemink 2001; Vukicevic and Posselt 2008; Pu and Hacker 2009; Ambadan and Tang 2009; Carrassi and Vannitsem 2010; Le and Bickel 2011; Hodyss 2011; Lei et al. 2012; Sakov et al. 2012; Yang et al. 2012; Zhang et al. 2012; Marzban 2013; and Goodliff et al. 2015) because it is less complex than a full physics numerical weather prediction model while maintaining strong nonlinearity representative of many atmospheric processes. The L63 model consists of a set of three coupled ordinary differential equations that provide a simplified description of dry convection. The model equations can be written as:

$$\tau \frac{dx_1}{dt} = \sigma(x_2 - x_1) \tag{5.4}$$

$$\tau \frac{dx_2}{dt} = \rho x_1 - x_2 - x_1 x_3 \tag{5.5}$$

$$\tau \frac{dx_3}{dt} = x_1 x_2 - \beta x_3 \tag{5.6}$$

where $x_1(t)$, $x_2(t)$, and $x_3(t)$ are the dependent variables, τ is a temporal scaling factor and σ, ρ, and β are the parameters of the model. For some parameter values, the system shows chaotic behavior because very small perturbations in the initial conditions can grow very rapidly into completely different solutions. The model was designed to simulate atmospheric dry cellular convection following the work of Saltzman (1962). The model simulates the evolution of a forced dissipative hydrodynamic system that possesses non-periodic and unstable solutions. The x_1 variable measures the intensity of convective motion, the x_2 variable measures the temperature difference between the ascending and descending currents, and the x_3 variable measures the distortion of the vertical temperature profile from linearity. The model parameters represent the Prandtl number (σ), a normalized Rayleigh number (ρ), and a non-dimensional wave number (β). The critical Rayleigh number for the system is 24.74; however, ρ is typically set to the slightly supercritical value of 28

following the work of Lorenz (1963). The σ and β parameters are set to 10 and $8/3$, respectively. Together, the values for these three parameters sustain the chaotic nature of the model. The behavior of the L63 model and the relationship between homoclinic bifurcations and the appearance of chaotic behaviors when changing the parameter values have been discussed by various authors such as Sparrow (1982), Afraimovich et al. (1977), Kaplan et al. (1979), and Barrio et al. (2012).

In this study, we investigate the sensitivity of the L63 model to perturbations in the ρ parameter and identify suitable predictors that can be used to estimate conditional biases in the state variables (x_1, x_2, x_3) due to these perturbations. We first generate a nature or "truth" simulation that tracks the evolution of the state variables over a certain period of time. The truth simulation is generated using a particular model for the behavior of ρ over time. Here, we choose to use a coupled version of the L63 model where each system $(S1, S2)$ is run at a different speed and one-way coupling occurs through the influence of $S2$ on the ρ parameter in $S1$, as is illustrated in Fig. 5.1. After some experimentation, we decided to set $\tau_{S1} = 1$ and $\tau_{S2} = 5$, which means that the hidden system $S2$ is integrated forward at one-fifth the speed of $S1$.

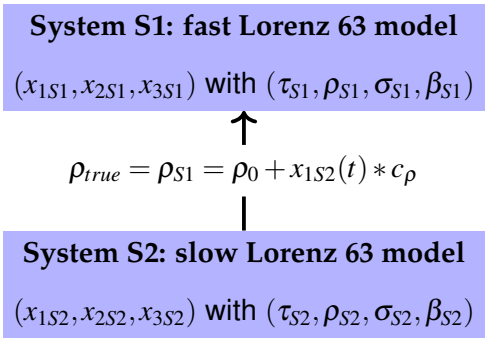

Figure 5.1: Coupled version of the Lorenz-63 model, with the fast system $S1$ dependent on the slow system $S2$. $S1$ is used to generate the nature simulation.

The state location x_{S2} obtained from the hidden system is then scaled by a factor of $c_\rho = 0.2$, with the scaled value subsequently used to perturb ρ_0, such that

$$\rho_{S1} = \rho_0 + x_{1S2}(t) \cdot c_\rho, \;\; t \in \mathbb{R}, \tag{5.7}$$

where $\rho_0 = \rho_{S2} = 28$, $x_{1S2}(t) \cdot c_\rho$ is the $\delta\rho$ perturbation obtained from $S2$, and ρ_{S1} is the resultant value used when integrating $S1$ during the next time step. The scaling of x_{1S2} by $c_p = 0.2$ means that $\delta\rho$ varies between approximately -4 and +4, which is reasonable because this represents departures up to 15% from ρ_0. The slowly varying autocorrelated $\delta\rho$ perturbations could be thought of as representing changes in the original L63 model equations due to the influence of the seasonal cycle on daily forecasts or the diurnal cycle on hourly forecasts. For example, parameters in cloud microphysics parameterization schemes are often assigned constant values even though some of them are known to vary, sometimes by up to several orders of magnitude, depending upon the stage of the cloud's life cycle. A similar approach was used by Zhang et al. (2012), where they attached an ocean slab model to the L63

model equations in order to represent the interaction between the slowly-varying ocean and the rapidly-changing atmosphere. Note that the parameters σ_{S1} and σ_{S2} were set to 10, and β_{S1} and β_{S2} were set to 8/3, as is typically done in the L63 model.

II. After generating the "truth" simulation using $S1$ in which the ρ_{S1} parameter varied with time, observations were generated for each state variable $(x_{1S1}, x_{2S1}, x_{3S1})$ at each model time step and then used in cycled data assimilation experiments employing a 3DVAR data assimilation system. The truth simulation and data assimilation experiments were started with the same initial conditions $(x_{1S1}, x_{2S1}, x_{3S1}) = (2, 3, 11)$; however, in the absence of data assimilation, they will follow different trajectories thereafter due to differences in the ρ parameter. The L63 model is integrated to the next time step using a 4th order Runge-Kutta time integration scheme. Various tests were performed using different observation error magnitudes and time step lengths, as will be shown in Section 5.5. Figure 5.2 shows the trajectory of the model state variables and evolution of the ρ_{S1} parameter during the truth simulation. It can be seen in Fig. 5.2b that the parameter ρ_{S1} may occasionally fall below the critical Rayleigh number for the system (24.74), which is also referred to as a sub-critical Hopf bifurcation point. Though this could temporarily lead to strange behavior in the model, we do not anticipate it having an impact on the experiments performed in Section 5.5 because ρ_{S1} will not remain at this value for long because it is ultimately controlled by variations in the state of the hidden system $S2$ as is shown in (5.7).

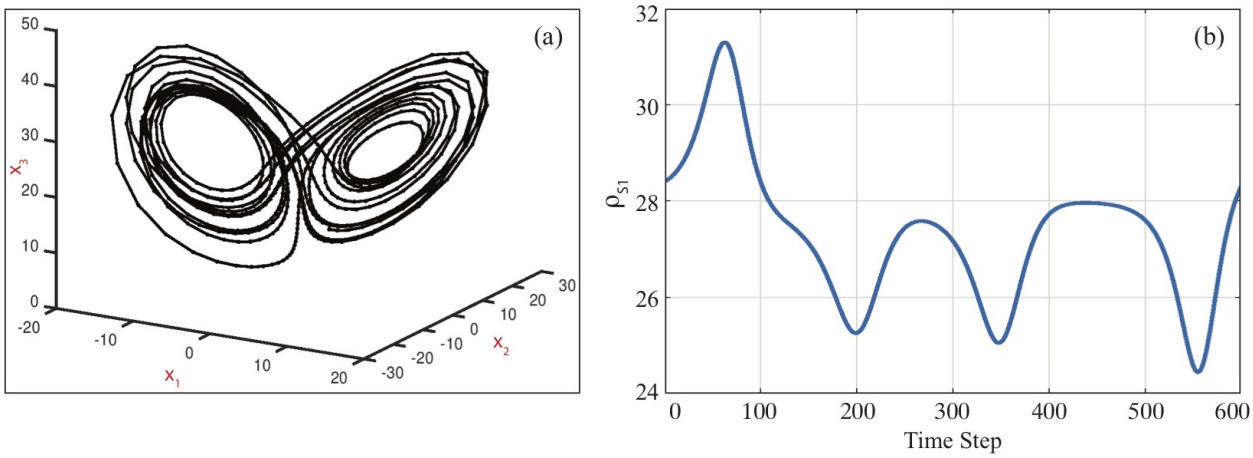

Figure 5.2: (a) Butterfly diagram showing the model trajectory during 600 time steps of the truth simulation using the coupled L63 system described in Section 5.3.1. (b) Time series showing the evolution of the ρ_{S1} parameter during the truth simulation, where ρ_{S1} for each model time step is set using (5.7).

The data assimilation experiments employed the typical L63 model equations, including $\rho = 28$; however, for these experiments, the equations represent an imperfect model because we know

that ρ is not constant during the truth simulation. Let us assume that we know that ρ varies with time, but that we only know its mean value ($\bar{\rho} = 28$) and not how it changes with time. The instantaneous difference between ρ in the data assimilation experiment and ρ in the truth simulation represents a model error; however, these differences correspond to conditional model biases when assessed over long time periods because $\delta\rho$ is a function of $S2$. Because errors in ρ directly impact the evolution of all three of the state variables in nonlinear ways, the instantaneous errors will potentially result in biases in the model state variables that are a nonlinear function of one or more predictors when assessed over long time periods. For example, a numerical weather prediction model may have the tendency to produce convection that is too strong during the day or too weak during the night, both of which will impact the sign and magnitude of the model biases in nonlinear ways during different parts of the diurnal cycle.

5.3.2 Asymptotics for Model Bias of the L63 System

Here, we first evaluate how the model variables (x_1, x_2, x_3) change in dependence on the model parameter ρ. In particular, we aim to develop an asymptotic estimator for the error in (x_1, x_2, x_3) depending on ρ and time t. The asymptotic analysis is performed using a Taylor series expansion with an explicit integral form of the error term. This approach is necessary because some of the constants will be zero in the higher order terms; therefore, we need to take sufficiently many terms into account to get the correct higher order terms.

Theorem 1. *The leading terms of the asymptotic analysis of the L63 system with respect to variations of $\rho = \rho_0 + \delta\rho$, where we use $t = t_0 + \delta t$ and $O(s)$ denotes a function bounded by $c|s|$ with some constant c in a neighborhood of $s = 0$, are given by*

$$x_1(\rho,t) - x_1(\rho_0,t) = \frac{1}{2}\sigma x_1(\rho_0,t_0)\cdot\delta\rho\cdot(\delta t)^2 + O(\delta\rho\cdot\delta t^3) \tag{5.8}$$

$$x_2(\rho,t) - x_2(\rho_0,t) = x_1(\rho_0,t_0)\cdot\delta\rho\cdot\delta t + O(\delta\rho\cdot\delta t^2) \tag{5.9}$$

$$x_3(\rho,t) - x_3(\rho_0,t) = x_1^2(\rho_0,t_0)\cdot\delta\rho\cdot(\delta t)^2 + O(\delta\rho\cdot\delta t^3) \tag{5.10}$$

Proof. We work out the proof in four steps, starting with some general setup and then considering the variables x_1, x_2, and x_3 in three steps.

Step 1. We begin by differentiating equations (5.4) - (5.6) with respect to ρ using the product rule,

where

$$x_1' = \frac{dx_1}{d\rho},\ x_2' = \frac{dx_2}{d\rho},\ x_3' = \frac{dx_3}{d\rho}$$

are the derivatives of the state variables with respect to ρ. Because the differentiation with respect to t and to ρ can be exchanged in the case of continuously differentiable functions, we obtain:

$$\frac{dx_1'}{dt} = \sigma x_2' - \sigma x_1' \tag{5.11}$$

$$\frac{dx_2'}{dt} = x_1'\rho + x_1 - x_2' - x_1' x_3 - x_1 x_3' \tag{5.12}$$

$$\frac{dx_3'}{dt} = x_1' x_2 + x_1 x_2' - \beta x_3'. \tag{5.13}$$

Note that all of the variables depend on time t and the parameter $\rho = \rho(t)$, and that the τ terms in equations (5.4) - (5.6) have been set to 1 to represent the original L63 model equations as described in Lorenz (1963).

To assess the sensitivity of the L63 model equations to variations in ρ at times t close to some initial time, t_0, we begin by looking at the scenario where the initial values for (x_1, x_2, x_3) are prescribed and identical for all ρ under consideration, such that at $t = t_0$:

$$x_1(\rho, t_0) = x_{1,0} \tag{5.14}$$

$$x_2(\rho, t_0) = x_{2,0} \tag{5.15}$$

$$x_3(\rho, t_0) = x_{3,0}. \tag{5.16}$$

This is an initial value problem where the derivatives of each equation with respect to ρ, (x_1', x_2', x_3'), are equal to zero at $t = t_0$, i.e.

$$x_1'(\rho, t_0) = 0,\ x_2'(\rho, t_0) = 0,\ x_3'(\rho, t_0) = 0. \tag{5.17}$$

This assumption is justified because in the case of perfect observations and full observability of the model state, you will obtain the true model state during each data assimilation step. With a perfect analysis, the first guess error at the next assimilation time will then be due to the model error that is predicted by the error estimators in Theorem 1 and will be independent of the value of

ρ at any particular point in time. After inserting these initial values into (5.11) - (5.13), we obtain:

$$\begin{aligned}
\frac{dx_1'}{dt}(\rho,t_0) &= 0 && (5.18)\\
\frac{dx_2'}{dt}(\rho,t_0) &= x_1(\rho,t_0) && (5.19)\\
\frac{dx_3'}{dt}(\rho,t_0) &= 0 && (5.20)
\end{aligned}$$

Step 2. Equation (5.19) reveals that the time rate of change of the sensitivity of x_2 with respect to ρ (i.e., x_2') is a function of its location along the x_1-axis. We now carry out an asymptotic analysis by an expansion of the functions with respect to variations in time $t = t_0 + \delta t$ and the parameter $\rho = \rho_0 + \delta\rho$. To assess the sensitivity of x_2 with respect to small variations in ρ, we employ (5.17) and (5.19) as follows. We estimate

$$\begin{aligned}
x_2(\rho,t) - x_2(\rho_0,t) &= \int_{\rho_0}^{\rho} x_2'(\tilde{\rho},t)\, d\tilde{\rho}\\
&= \int_{\rho_0}^{\rho} \Big(\underbrace{x_2'(\tilde{\rho},t_0)}_{=0} + \int_{t_0}^{t} \frac{dx_2'(\tilde{\rho},\tilde{t})}{d\tilde{t}}\, d\tilde{t} \Big)\, d\tilde{\rho}\\
&= \int_{\rho_0}^{\rho} \int_{t_0}^{t} \frac{dx_2'(\tilde{\rho},\tilde{t})}{d\tilde{t}}\, d\tilde{t}\, d\tilde{\rho}\\
&= \int_{\rho_0}^{\rho} \int_{t_0}^{t} \Big(\underbrace{\frac{dx_2'(\tilde{\rho},\tilde{t})}{d\tilde{t}}|_{t_0}}_{=x_1(\tilde{\rho},t_0)} + \int_{t_0}^{\tilde{t}} \frac{d^2x_2'(\tilde{\rho},s)}{ds^2} ds \Big)\, d\tilde{t}\, d\tilde{\rho}. && (5.21)
\end{aligned}$$

We estimate both terms in (5.21) separately. For the first term T_1, by (5.14) we obtain

$$\begin{aligned}
T_1 &= \int_{\rho_0}^{\rho} \int_{t_0}^{t} x_1(\tilde{\rho},t_0)\, d\tilde{t}\, d\tilde{\rho}\\
&= \int_{\rho_0}^{\rho} \int_{t_0}^{t} x_1(\rho_0,t_0)\, d\tilde{t}\, d\tilde{\rho}\\
&= x_1(\rho_0,t_0) \cdot \delta\rho \cdot \delta t, && (5.22)
\end{aligned}$$

where $x_1(\tilde{\rho},t_0)$ is replaced by $x_1(\rho_0,t_0)$ because the derivative of x_1 with respect to ρ is zero at t_0 following (5.17). The $\delta\rho$ and δt terms are obtained by solving the definite integrals, with $\delta\rho$ denoting the interval $[\rho_0,\rho]$ and δt denoting the interval $[t_0,t]$. The second term is estimated in a similar way by

$$\begin{aligned}
T_2 &= \int_{\rho_0}^{\rho} \int_{t_0}^{t} \int_{t_0}^{\tilde{t}} \frac{d^2x_2'(\tilde{\rho},s)}{ds^2}\, ds\, d\tilde{t}\, d\tilde{\rho}\\
&= O(\delta\rho \cdot \delta t^2). && (5.23)
\end{aligned}$$

Combining the estimates (5.22) and (5.23) then leads to

$$x_2(\rho,t) - x_2(\rho_0,t) \quad = \quad x_1(\rho_0,t_0)\cdot\delta\rho\cdot\delta t + O(\delta\rho\cdot\delta t^2). \tag{5.24}$$

This proves equation (5.9) in Theorem 1.

Step 3. To obtain an estimate for $x_1(\rho,t)$, we proceed as in equation (5.21) and, for a twice continuously differentiable function $x_1(\rho,t)$, estimate

$$\begin{aligned} x_1(\rho,t) &= x_1(\rho_0,t) + \int_{\rho_0}^{\rho} x_1'(\tilde{\rho},t)\, d\tilde{\rho} && (5.25)\\ &= x_1(\rho_0,t) + \int_{\rho_0}^{\rho} \left(x_1'(\rho_0,t) + \int_{\rho_0}^{\tilde{\rho}} x_1''(\tilde{\tilde{\rho}},t)\, d\tilde{\tilde{\rho}} \right) d\tilde{\rho} && (5.26) \end{aligned}$$

We note that by taking the derivative of (5.11) with respect to time and inserting (5.19) into the resultant equation, we obtain

$$\begin{aligned} \frac{d^2 x_1'(\rho,t_0)}{dt^2} &= \sigma\frac{dx_2'(\rho,t_0)}{dt} - \sigma\frac{dx_1'(\rho,t_0)}{dt} \\ &= \sigma x_1(\rho,t_0) && (5.27) \end{aligned}$$

and thus the derivative of (5.27) with respect to time gives

$$\frac{d^2 x_1''(\rho,t_0)}{dt^2} \quad = \quad \sigma x_1'(\rho,t_0) = 0. \tag{5.28}$$

Performing a third order expansion around t_0 then leads to an estimate for $x_1''(\rho,t)$:

$$x_1''(\rho,t) \quad = \quad O(\delta t^3). \tag{5.29}$$

After inserting (5.29) into (5.26) and then solving the definite integrals, we obtain:

$$x_1(\rho,t) \quad = \quad x_1(\rho_0,t) + x_1'(\rho_0,t)\cdot\delta\rho + O(\delta\rho^2\cdot\delta t^3) \tag{5.30}$$

To estimate $x_1'(\rho,t)$, with the help of (5.17) and (5.18), we derive:

$$\begin{aligned} x_1'(\rho,t) &= \underbrace{x_1'(\rho,t_0)}_{=0} + \int_{t_0}^{t} \frac{dx_1'(\rho,\tilde{t})}{d\tilde{t}}\, d\tilde{t} \\ &= \int_{t_0}^{t} \Big(\underbrace{\frac{dx_1'(\rho,\tilde{t})}{d\tilde{t}}\Big|_{t_0}}_{=0} + \int_{t_0}^{\tilde{t}} \frac{d^2x_1'(\rho,s)}{ds^2}\, ds \Big)\, d\tilde{t}. \end{aligned} \tag{5.31}$$

The second derivative of $x_1'(\rho,t)$ with respect to time t can be estimated by differentiating (5.11) with respect to t, and using (5.12) and (5.19), which yields:

$$\begin{aligned} \frac{d^2x_1'(\rho,t)}{dt^2} &= \frac{d}{dt}\Big(\frac{dx_1'(\rho,t)}{dt}\Big) \\ &= \frac{d}{dt}\Big(\sigma x_2'(\rho,t) - \sigma x_1'(\rho,t)\Big) \\ &= \sigma\frac{dx_2'}{dt}(\rho,t) - \sigma\frac{dx_1'}{dt}(\rho,t) \\ &= \sigma\frac{dx_2'}{dt}(\rho,t_0) - \sigma\underbrace{\frac{dx_1'}{dt}(\rho,t_0)}_{=0} + O(\delta t) \\ &= \sigma x_1(\rho,t_0) + O(\delta t). \end{aligned} \tag{5.32}$$

We insert this into (5.31) to conclude with

$$\begin{aligned} x_1'(\rho,t) &= \sigma x_1(\rho,t_0)\cdot\int_{t_0}^{t}\int_{t_0}^{\tilde{t}} 1\, ds\, d\tilde{t} + O(\delta t^3) \\ &= \sigma x_1(\rho,t_0)\cdot\int_{t_0}^{t}(\tilde{t}-t_0)\, d\tilde{t} + O(\delta t^3) \\ &= \sigma x_1(\rho,t_0)\cdot\frac{1}{2}(\delta t)^2 + O(\delta t^3). \\ &= \sigma x_1(\rho_0,t_0)\cdot\frac{1}{2}(\delta t)^2 + O(\delta t^3). \end{aligned} \tag{5.33}$$

Finally, we insert (5.33) into (5.30) with the help of (5.17) to obtain (5.8) in Theorem 1.

Step 4. In our final step, we estimate the behavior of $x_3(\rho,t)$. We note that similarly to $x_1'(\rho,t)$ given by (5.33) as in (5.21) we obtain:

$$\begin{aligned} x_2'(\rho,t) &= \underbrace{x_2'(\rho,t_0)}_{=0} + \int_{t_0}^{t}\frac{dx_2'(\rho,\tilde{t})}{d\tilde{t}}\, d\tilde{t} \\ &= x_1(\rho,t_0)\cdot\delta t + O(\delta t^2). \end{aligned} \tag{5.34}$$

Also, based on (5.20) we calculate

$$\begin{aligned} x_3'(\rho,t) &= \underbrace{x_3'(\rho,t_0)}_{=0} + \int_{t_0}^{t} \frac{dx_3'(\rho,\tilde{t})}{d\tilde{t}}\, d\tilde{t} \\ &= \int_{t_0}^{t} \Big(\underbrace{\frac{dx_3'(\rho,\tilde{t})}{d\tilde{t}}|_{t_0}}_{=0} + \int_{t_0}^{\tilde{t}} \frac{d^2x_3'(\rho,s)}{ds^2}\, ds \Big)\, d\tilde{t} \\ &= O(\delta t^2). \end{aligned} \tag{5.35}$$

Now, we follow the above lines to estimate

$$\begin{aligned} x_3(\rho,t) - x_3(\rho_0,t) &= \int_{\rho_0}^{\rho} x_3'(\tilde{\rho},t)\, d\tilde{\rho} \\ &= \int_{\rho_0}^{\rho} \Big(\underbrace{x_3'(\tilde{\rho},t_0)}_{=0} + \int_{t_0}^{t} \frac{dx_3'(\tilde{\rho},\tilde{t})}{d\tilde{t}}\, d\tilde{t} \Big)\, d\tilde{\rho}. \end{aligned} \tag{5.36}$$

Here, to obtain a sharper estimate than (5.35) and to evaluate the constant explicitly, we insert (5.13) into (5.36), which yields:

$$x_3(\rho,t) - x_3(\rho_0,t) = \int_{\rho_0}^{\rho} \int_{t_0}^{t} \Big(x_1'(\tilde{\rho},\tilde{t}) x_2(\tilde{\rho},\tilde{t}) + x_1(\tilde{\rho},\tilde{t}) x_2'(\tilde{\rho},\tilde{t}) - \beta x_3'(\tilde{\rho},\tilde{t}) \Big)\, d\tilde{t}\, d\tilde{\rho}. \tag{5.37}$$

Because $(x_1', x_2', x_3') = 0$ at t_0, we need to estimate the leading order term by its temporal change at t_0 as given in (5.18) - (5.20). We insert the asymptotics for $x_1'(\rho,t)$, $x_2'(\rho,t)$, and $x_3'(\rho,t)$ given by (5.33), (5.34), and (5.35) into (5.37) to estimate:

$$\begin{aligned} x_3(\rho,t) - x_3(\rho_0,t) &= \int_{\rho_0}^{\rho} \int_{t_0}^{t} \Big(x_1^2(\rho,t_0)\delta t + O(\delta t^2) \Big)\, d\tilde{t}\, d\tilde{\rho} \\ &= x_1^2(\rho,t_0) \cdot \delta t^2 \cdot \delta\rho + O(\delta\rho \cdot \delta t^3), \end{aligned} \tag{5.38}$$

where $x_1^2(\rho,t_0)\delta t$ is the leading order term, and all other terms have been absorbed into the order $O(\delta t^2)$ term. Thus, we have derived (5.10) in Theorem 1 and the proof is complete. □

Remark. In Step 3 of the proof, we could have performed the estimate slightly differently.

Using an approach similar to Steps 2 and 4, we obtain:

$$\begin{aligned} x_1(\rho,t) - x_1(\rho_0,t) &= \int_{\rho_0}^{\rho} x_1'(\tilde{\rho},t)\, d\tilde{\rho} \\ &= \int_{\rho_0}^{\rho} \Big(\underbrace{x_1'(\tilde{\rho},t_0)}_{=0} + \int_{t_0}^{t} \frac{dx_1'(\tilde{\rho},\tilde{t})}{d\tilde{t}}\, d\tilde{t} \Big)\, d\tilde{\rho} \\ &= \int_{\rho_0}^{\rho} \int_{t_0}^{t} \Big(\underbrace{\frac{dx_1'(\tilde{\rho},\tilde{t})}{d\tilde{t}}|_{t_0}}_{=0} + \int_{t_0}^{\tilde{t}} \frac{d^2x_1'(\tilde{\rho},s)}{ds}\, ds \Big)\, d\tilde{t}\, d\tilde{\rho} \end{aligned} \tag{5.39}$$

and then proceed as in (5.32) and (5.33) to obtain (5.8) as above. □

5.4 Improving Data Assimilation using Bias Estimators

Being able to accurately estimate errors in the model background $x^{(b)}$ is important for any practical implementation of a data assimilation algorithm. In this section, we first discuss the model error and model bias terminology and then study a simple Bayesian example to illustrate the importance of correctly estimating the model background error covariance matrix B. We then develop a generalized model error estimation method that is subsequently applied to the L63 model discussed in Section 5.3.2 to demonstrate the feasibility of dynamically estimating the model errors using nonlinear estimators based on the model variables. In Section 5.4.4, we show how the bias correction coefficient vector obtained through solving a least squares minimization problem can be used to estimate the unknown parameter using the analysis increments from the data assimilation system.

5.4.1 Nonlinear Model Bias and Error Terminology

In this section, we sharpen the terminology for model error, model bias, and conditional model bias, and compare the concepts. For a particular location, the model error is the instantaneous difference between the background state $x^{(b)}$ and the true state $x^{(true)}$ of the system. Model bias is then defined as the $x^{(b)} - x^{(true)}$ differences averaged over some period of time or region:

$$b_b := \mathbb{E}\{x^{(b)} - x^{(true)}\}, \tag{5.40}$$

where the bias is computed separately for different model quantities such as temperature, humidity, or cloud water path. If we then assume that the analysis state $x^{(a)}$ obtained during each assimilation cycle is the best estimate of the true system state, we can use the resultant $x^{(b)} - x^{(a)}$ differences as an approximation to the true model bias, with appropriate summation over particular regions or periods of time:

$$b_{b-a} := \mathbb{E}\{x^{(b)} - x^{(a)}\}. \tag{5.41}$$

The conditional model bias can then be defined as the mean deviation of the dependent variable from the true system state when the bias is a function of some other parameter or variable p referred to as the predictor. The conditional model bias can be estimated using:

$$b_{b-a}(p) = \mathbb{E}\{x^{(b)}(p) - x^{(a)}(p)\}. \tag{5.42}$$

For this study, we are interested in the situation where the bias predictor is a component of the model state.

If $b_{b-a}(p)$ varies in a nonlinear manner, then this behavior represents a nonlinear conditional bias and we will need to use nonlinear bias correction methods to remove the bias from the model variables. In this case, let us assume that the function $b_{b-a}(p)$ can be written as a superposition

$$b_{b-a}(p) = \sum_{\xi=1}^{N} \psi_\xi(p)\alpha_\xi \tag{5.43}$$

of nonlinear basis functions ψ_ξ with N unknown coefficients α_ξ. The solution of (5.43) can be understood as a generalized bias estimation equation because it structures the set of differences according to the predictor p and searches for a functional estimation of its behavior. We can then employ nonlinear bias correction methods such as that described in Otkin et al. (2018) to determine the bias correction coefficients based on a set of $b_{b-a}(p)$ differences. To do this effectively, we will need to obtain a large sample of differences covering a diverse range of system states.

It should also be noted that the estimation of the coefficients α_ξ in (5.43) using $x^{(b)} - x^{(a)}$ differences accumulated over multiple assimilation cycles subsequently leads to the capability to predict the instantaneous model error when those coefficients are applied to the current state during an individual assimilation cycle. This demonstrates that conditional model bias estimation and model error estimation are strongly related and show significant overlap. As discussed in Section 5.4.3, the forecast error in general can be represented as a combination of state estimation

error associated with the propagation of errors in the prior analysis to the current time and a second component that represents the true model error arising from the use of an imperfect model. The instantaneous model errors can therefore be viewed as conditional model biases because their characteristics likely depend on the state of the system.

The conditional model error estimators can be used for various purposes, including a) model bias correction where the model background is corrected prior to its use in the data assimilation system, b) model uncertainty estimation where the model error estimates are used to improve the background error covariance matrix B, and c) model development efforts where the error statistics are used to improve the accuracy of the numerical model. In this paper, we focus on application b) because we seek to employ knowledge regarding the behavior of the model errors to improve estimates of the model background uncertainty.

5.4.2 Study of a Simple Bayesian Example

A Bayesian data assimilation step employs Bayes formula

$$p^{(a)}(x) = cp^{(b)}(x)p(y|x), \;\; x \in \mathbb{R}^n \tag{5.44}$$

for estimating the posterior probability distribution $p^{(a)}(x)$ based on the prior probability distribution $p^{(b)}(x)$ and the observation error distribution $p(y|x)$. The prior distribution is usually assumed to be Gaussian in data assimilation systems, such that:

$$p^{(b)}(x) := \tilde{c}e^{-\frac{1}{2}(x-x^{(b)})^T B^{-1}(x-x^{(b)})}, \;\; x \in \mathbb{R}^n, \tag{5.45}$$

where $\tilde{c}$ is a constant and the background error covariance matrix B is estimated climatologically in classical variational assimilation systems or based on an ensemble of model states in an EnKF.

Here, we discuss and demonstrate the role of the correct estimate of B on the quality of the analysis mean and analysis distribution. For an EnKF system, the ensemble spread is used to estimate B, however, this estimate only contains part of the error when a numerical model is used because it does not include the difference between the model and the true state of the system. Variational data assimilation systems, such as 3DVAR, are also unable to consider these differences because B is chosen as fixed for a particular time period due to the way in which it is constructed. This means that the model bias and how it changes with time is not taken into account by either

assimilation methodology, which can substantially degrade their performance. For the remainder of this work, we restrict our attention to 3DVAR because that is what we used during the numerical experiments discussed in Section 5.5. We note however that similar arguments apply for ensemble data assimilation systems.

As a starting point, we derive the error representation explicitly for a one-dimensional Gaussian case with observation operator $H = I$. In one dimension, the best estimate of the current state (or analysis) during an assimilation step is given by:

$$x^{(a)} = x^{(b)} + \frac{q^{od}}{r + q^{od}}(y - x^{(b)}), \tag{5.46}$$

where y is the observation, r is the observation error uncertainty, $x^{(b)}$ is the first guess or background, and q^{od} represents the estimated variance of the error in the variable x. Now, let us assume that q_0^{od} is the true background error variance that includes model error, such that the correct analysis $x_0^{(a)}$ is represented as:

$$x_0^{(a)} = x^{(b)} + \frac{q_0^{od}}{r + q_0^{od}}(y - x^{(b)}) \tag{5.47}$$

The error between the analysis based on some uncertainty or variance q^{od} and the correct uncertainty or variance q_0^{od} is then given by:

$$\begin{aligned} |x^{(a)} - x_0^{(a)}| &= \left| \frac{q^{od}}{r + q^{od}} - \frac{q_0^{od}}{r + q_0^{od}} \right| \cdot |y - x^{(b)}| \\ &= \left| \frac{r(q^{od} - q_0^{od})}{(r + q^{od}) \cdot (r + q_0^{od})} \right| \cdot |y - x^{(b)}|. \end{aligned} \tag{5.48}$$

This result shows that the analysis error for each assimilation step is proportional to the observation departure $|y - x^{(b)}|$ and to the accuracy of the background error variance estimate $|q^{od} - q_0^{od}|$. Thus, development of new methods that can be used to generate a more accurate estimate of q^{od} will directly improve the quality of the analysis and performance of the assimilation system.

5.4.3 Dynamical Error and Bias Estimators

In this section, we develop a generalized method to diagnose model biases using the model variables. First, let us assume that the the forecast error $x_k^{(b)} - x_k^{(true)}$ at a given time k can be represented as the difference between the dynamical states that are obtained when the prior analysis $x_{k-1}^{(a)}$ is propagated by an imperfect model M and the true prior state $x_{k-1}^{(true)}$ is propagated by the perfect

model M^{true}:

$$x_k^{(b)} - x_k^{(true)} = M(x_{k-1}^{(a)}) - M^{true}(x_{k-1}^{true}) \tag{5.49}$$

The forecast error can then be decomposed into one part that is due to the propagation of the uncertainty error associated with the prior analysis state $M(x_{k-1}^{(a)}) - M(x_{k-1}^{true})$, and a second part that represents the true model error $M(x_{k-1}^{true}) - M^{true}(x_{k-1}^{true})$ during the propagation from the prior time:

$$x_k^{(b)} - x_k^{(true)} = \Big(M(x_{k-1}^{(a)}) - M(x_{k-1}^{true})\Big) + \Big(M(x_{k-1}^{true}) - M^{true}(x_{k-1}^{true})\Big). \tag{5.50}$$

Taking the variance on both sides of (5.50), and using

$$q^{state} := \mathrm{Var}(M(x_{k-1}^{(a)}) - M(x_{k-1}^{true})) \tag{5.51}$$

and

$$q^{model} = \mathrm{Var}(M(x_{k-1}^{true}) - M^{true}(x_{k-1}^{true})), \tag{5.52}$$

we obtain the total variance of the forecast error:

$$\begin{aligned} q^{total} &:= \mathrm{Var}\Big(x_k^{(b)} - x_k^{(true)}\Big) && (5.53)\\ &= q^{state} + q^{model} + 2 \cdot \mathrm{Cov}\Big(M(x_{k-1}^{(a)}) - M(x_{k-1}^{true}),\ M(x_{k-1}^{true}) - M^{true}(x_{k-1}^{true})\Big). && (5.54) \end{aligned}$$

It is a standard approach in data assimilation to assume that the initial condition uncertainty and model error are uncorrelated (Mitchell and Carrassi 2015), which means that the covariance term on the righthand side of (5.54) will equal zero and therefore the total variance of the forecast error can be given by

$$q^{total} = q^{state} + q^{model}, \tag{5.55}$$

where q^{state} reflects the influence of the variance of the estimate of the prior analysis propagated to the current analysis time using the model equations, and q^{model} is the variance in the model error E due to the use of an imperfect numerical model.

If some error estimators such as those shown in Theorem 1 are available, we can employ (5.55) to estimate q^{total} and then use it to improve the estimate of the analysis during a given data assimilation step. Though we typically will not know q^{state} in a complex real-world system, the development of a method that can be used to estimate the time-varying model error E, and thus

the variance q^{model}, allows us to employ a lower fixed q^{state} in our approach. This outcome is better than having to use a larger fixed q^{state}, which would otherwise be the case, because that would lead to an overestimate of the total error variance. In general, it will not be possible to carry out a full assessment of the model error due to incomplete knowledge of the governing equations; however, Theorem 1 shows that the model error asymptotically depends on the model variables, here in particular, $x_1(\rho_0, t_0)$. We can therefore employ nonlinear model error estimators to diagnose such dependencies as follows.

II. We begin with a general example where we study the estimation of an error that depends on the model state x and time t. We model the dependence on the states using basis functions $\varphi_\ell(x)$, $x \in \mathbb{R}^n$, with $\ell = 1, ..., N_\ell$. The dependence on time is modeled using basis functions ψ_k, $k = 1, ..., N_k$. Let us assume an ansatz of the form

$$E_j(x,t) = \sum_{\ell=1}^{N_\ell} \sum_{k=1}^{N_k} \beta_{\ell,k}^{(j)} \, \varphi_\ell(x) \psi_k(t), \;\; x \in \mathbb{R}^n, \, t \in \mathbb{R}, \tag{5.56}$$

for the model error E_j. For illustrative purposes, the functions $\psi_k(t)$ could be represented by $\sin(t)$ and $\cos(t)$ or by higher order trigonometric functions, whereas the functions $\varphi_\ell(x)$ could be represented by the polynomial terms in Theorem 1. In this situation, the terms would correspond to $\varphi_\ell(x) = x_1^{\xi_1} x_2^{\xi_2} x_3^{\xi_3}$, with ξ_1, ξ_2, ξ_3 counted by $\ell = 1, ..., N_\ell$, $\psi_1(t) \equiv 1$, and $\psi_k(t) = 0$ for $k > 1$. The coefficients $\beta_{\ell,k}^{(j)}$ are the unknown coefficients linking the true dynamics with the numerical model.

If we then observe the model error $E_j(x,t)$ for a selection of states $(x[\eta], t[\eta])$, $\eta = 1, ..., N_\eta$ such that the linear independence of φ_ℓ on $x[\eta]$ is satisfied and a set $t[\eta] \in [0,T]$ such that the linear independence of ψ_k is satisfied on this set, we know that the linear system

$$E_j(x[\eta]) = \sum_{\ell=1}^{N_\ell} \sum_{k=1}^{N_k} \beta_{\ell,k}^{(j)} \, \varphi_\ell(x[\eta]) \; \psi_k(t[\eta]), \tag{5.57}$$

$\eta = 1, ..., N_\eta$, has at most one solution for each $j = 1, ..., n$. It may be overdetermined if $N_\eta > N_\ell \cdot N_k$, and if the data is inconsistent would have no exact solution. In that case, we can use least squares methods to calculate approximate solutions.

Let us also discuss the case of non-uniqueness for the calculation of the bias correction coefficients. This situation can easily arise if two or more variables in the dynamical system under consideration are correlated. For example, the x_1 and x_2 variables in the L63 system display strong correlations in parts of the trajectory. Though the non-unique solution will not affect the quality

of the bias estimate for the time interval used to calculate the coefficients, it could potentially lead to large errors if these coefficients are used outside of the training period. Thus, we note that: 1) for time-local estimation of model biases, the consequences of non-uniqueness should be small, and 2) when the bias estimation tool is employed for longer time periods or for forecasting, it is important to have training periods that include conditions representative of the full climatology of the dynamical model.

III. Here, we illustrate the utility of the generalized framework developed in the previous section by applying it to the L63 model. First, let us assume that the true evolution of a hypothetical dynamical system, represented by M^{true}, depends on a particular parameter that varies with time, but that limitations in our understanding of the physical system means that it is assigned a constant value in the numerical model M used to represent the true dynamical system. An example is the dependence of the parameter ρ in the coupled L63 model described in Section 5.3.1, for which we have worked out the behavior of the model error for small time intervals δt and small changes $\delta\rho$ of ρ in Section 5.3.2. For this particular system, we observe the dependence of the error

$$E(\delta\rho) := \|x[\rho] - x[\rho_0]\|^2 \tag{5.58}$$

on the model state $x = (x_1, x_2, x_3)$ in Theorem 1, where ρ_0 is the true value at a given time t_0 in M^{true} and ρ is the constant value used by the imperfect model M. This dependence leads to the error estimate for the coupled L63 system:

$$E(\delta\rho) = x_1^2(\rho_0, t_0) \cdot \delta\rho^2 \cdot \delta t^2 + O(\delta\rho^2 \cdot \delta t^4), \tag{5.59}$$

where we added the squares of (5.8), (5.9), and (5.10), and then absorbed the higher order terms into the $O(\delta\rho^2 \cdot \delta t^4)$ term. It can be seen in (5.59) that the leading error term is proportional to x_1^2, which means that the expected model error is largest when the system state is located near the tips of the butterfly wings.

For this work, we use the analysis $x^{(a)}$ from each assimilation step as an approximation of the true state $x^{(true)}$ because the true state is unknown in a real-world system. Note that this approximation means that we will be unable to recover the full model error; however, because $x^{(a)}$ will be pulled toward the observations, we will still be able to estimate the model error under the assumption that the observations have small errors. We restrict our attention here to the L63

system with three variables in order to simplify the notation. The current model error E_j of the component x_j of the state $x \in \mathbb{R}^n$ is approximated by:

$$E_j := |x_j^{(a)} - x_j^{(b)}|, \tag{5.60}$$

where $j = 1,2,3$ corresponds to the three variables in the L63 system. Let us assume that knowledge of those parts of the system leading to model error at a specific time is such that after some manipulation the model error can be rewritten in the form of a triple sum:

$$E_j = \sum_{\xi_1,\xi_2,\xi_3=0}^{N_{coef}} \alpha^{(j)}_{\xi_1,\xi_2,\xi_3} x_1^{\xi_1} x_2^{\xi_2} x_3^{\xi_3}, \tag{5.61}$$

with coefficients $\alpha^{(j)}_{\xi_1,\xi_2,\xi_3}$, $\xi_1,\xi_2,\xi_3 = 0,...,N_{coef}$, where N_{coef} is the total number of coefficients determined by the maximum order of the polynomial and the number of model variables under consideration. For the L63 system containing three variables, $N_{coef} = 10$ for a 2nd order polynomial. The model error can be expressed as in (5.61) if we know that a hidden model exists but that we do not know the dependence of the true system because we cannot derive the asymptotics of the model equations. The ansatz (5.61) assumes some polynomial dependence of this relationship on the model variables $x \in \mathbb{R}^n$, as we have shown to be the case for the coupled L63 system. We also assume that the model errors do not have a temporal dependence such that the basis functions $\psi_k(t)$ in (5.56) can be set to 1.

Next, given a sequence of states $x[\eta]$ and their corresponding model errors $E_j[\eta]$ for $\eta = 1,...,N_{states}$ over some period of time, the above estimate leads to a linear system of equations:

$$A\alpha^{(j)} = q \tag{5.62}$$

for the $N_{coef} \times 1$ coefficient vector $\alpha^{(j)} = (\alpha^{(j)}_{0,0,0}, \alpha^{(j)}_{1,0,0}, \alpha^{(j)}_{0,1,0}, \alpha^{(j)}_{0,0,1}, \alpha^{(j)}_{1,1,0}, ...)^T$, where the sub-indices correspond to the polynomial order for the predictors (x_1, x_2, x_3) and the superscript denotes the model variable x_j. For example, the zeroth order coefficient for the x_1 variable is denoted as $\alpha^{(1)}_{0,0,0}$, whereas the second order coefficient for the $x_1 \cdot x_2$ mixed term is denoted as $\alpha^{(1)}_{1,1,0}$. Then, A is an $N_{states} \times N_{coef}$ matrix containing the N_{coef} polynomial terms for each observation:

$$A = A^{(j)} := \left(x_1^{\xi_1}[\eta] x_2^{\xi_2}[\eta] x_3^{\xi_3}[\eta] \right)_{\eta=1,...,N_{states};\ \xi_1,\xi_2,\xi_3=0,...,N_{coef}} \tag{5.63}$$

where η counts the rows and ξ_1, ξ_2, ξ_3 are subsequently ordered as column indices consistent with the ordering of the components of α, and

$$q = q^{(j)} := \Big(E_j[\eta]\Big)_{\eta=1,\ldots,N_{states}} \tag{5.64}$$

is the $N_{states} \times 1$ vector containing the model errors, with row index η. Finally, we can find the coefficients α that best fit the system of equations by solving the quadratic minimization problem, which leads to:

$$\alpha = (A^T A)^{-1} A^T q. \tag{5.65}$$

5.4.4 Parameter Estimation

We begin this section by noting that the asymptotics for the coupled L63 model shown in Theorem 1 reveal that the error, E_j, for each model variable $j = 1,2,3$ is proportional to the size of the hidden parameter $\delta\rho$, which means that the diagnosed conditional model bias should also be proportional to this parameter. In practice, however, this is not an easy relationship to capture because their proportionality depends in a very dynamic way on the current state of a modeling system characterized by chaotic behavior. Thus, without explicit knowledge of the model variables and the relationship between them and $\delta\rho$, it is impossible to draw conclusions about the size of $\delta\rho$.

However, based on the nonlinear model error estimators given by (5.8) - (5.10), we expect that the coefficient vector α in (5.61) will also be proportional to the size of the model bias. This vector depends on the average size of the analysis increment $x^{(a)} - x^{(b)}$ during the data assimilation steps rather than on the model state. The explicit dependence, unknown in general, is part of the estimation of the coefficients. Thus, we obtain a tool that can be used to dynamically diagnose the average size of the unknown parameter $\delta\rho$ by computing the mean of the coefficient vector α for each model variable $x_j = 1,2,3$. This leads to the following estimates for $\delta\rho$:

$$\delta\rho^{(1)}_{diag}(t) \approx c_1 \alpha^{(1)}_{1,0,0}(t) \quad \text{or} \quad \delta\rho^{(2)}_{diag}(t) \approx c_2 \alpha^{(2)}_{1,0,0}(t) \quad \text{or} \quad \delta\rho^{(3)}_{diag}(t) \approx c_3 \alpha^{(3)}_{2,0,0}(t) \tag{5.66}$$

where $c_1 = 2/\sigma(\delta t)^2$, $c_2 = 1/\delta t$, and $c_3 = 1/(\delta t)^2$, and we now need to carry out the bias estimation over time intervals $[t - \Delta t, t + \Delta t]$ with some $\Delta t > 0$ for which $\delta\rho$ can be considered a constant. We will demonstrate in Section 5.5.4 that this simple tool provides a reasonable approach to parameter

estimation for the dynamical system under consideration.

5.5 Numerical Results using the L63 Model

The purpose of this section is to use the L63 model to perform numerical experiments that demonstrate the validity of the model error identification and correction methods developed in the previous sections and their use within a data assimilation system. We begin by showing in Section 5.5.1 that the error asymptotics developed in Theorem 1 accurately represent the behavior of the L63 model and that they are able to capture the rapid evolution of the model error in each of the state variables. We then demonstrate in Section 5.5.2 that the model error asymptotics can be used to improve the model background error covariance matrix B through inclusion of a dynamic component that captures the current model errors. It is then shown in Section 5.5.3 that the coefficients of the nonlinear asymptotical expansion can be reasonably estimated by solving a regularized least squares minimization problem without explicit a priori knowledge of the error behavior. This is accomplished through use of a polynomial expansion of the model variables. Finally, we show in Section 5.5.4 that the ρ parameter can be reconstructed using the bias correction coefficient vector. Moreover, it is shown that it is possible to reconstruct this parameter using the analysis increments that are readily available in all data assimilation systems.

5.5.1 Analysis of the Asymptotic Error Estimators for the L63 Model

In this section, we assess the ability of the asymptotics derived in Theorem 1 to accurately capture the rapid evolution of model errors in the coupled L63 system during a cycled data assimilation experiment covering $N_t = 600$ assimilation cycles with an assimilation frequency $\delta t_{assim} = 0.06$. Though the true ρ parameter in the coupled L63 system varies with time following (5.7), it was set to a constant value ($\rho = 28$) during the data assimilation experiment to represent a dynamic and unknown model bias. Output from the truth simulation employing the time-varying ρ parameter was used to generate observations with zero measurement error (e.g., ε, the standard deviation of the errors drawn from an unbiased Gaussian distribution, was set to zero) for (x_1, x_2, x_3), which were then assimilated using a 3DVAR system. The analysis $x^{(a)}$ during a given assimilation cycle

was determined by minimizing the cost function:

$$J(x) := \left\| x - x^{(b)} \right\|^2_{B^{-1}} + \|y - H(x)\|^2_{R^{-1}}, \tag{5.67}$$

which, for a linear observation operator, is given by:

$$x^{(a)} = x^{(b)} + BH^T(R + HBH^T)^{-1}(y - H(x^{(b)})), \tag{5.68}$$

where $H = I$, the observation error covariance matrix R was given the form of the identity matrix scaled by the factor r,

$$R = r \cdot I, \tag{5.69}$$

and the background error covariance matrix B was given the form:

$$B = \begin{pmatrix} (x_1^{(b)} - x_1^{(true)})^2 & 0 & 0 \\ 0 & (x_2^{(b)} - x_2^{(true)})^2 & 0 \\ 0 & 0 & (x_3^{(b)} - x_3^{(true)})^2 \end{pmatrix}, \tag{5.70}$$

with $x^{(b)}$ being the background state, $x^{(true)}$ being the true dynamical state obtained from the truth simulation, and the diagonal elements of B containing the model error variances. We chose to use a diagonal matrix here because it is a reasonable place to start, and as is shown in this section, still has a positive impact on the assimilation performance. Given the strong correlations between the errors in the x_1 and x_2 variables (see Fig. 5.3), it is likely that including the off-diagonal elements would have led to even better results; however, their inclusion in the B matrix is left for future work. Note that the scaling factor r was set to 0.016 during this experiment.

Figure 5.3 shows the evolution of the true ρ parameter and the model errors $x_1^{(b)} - x_1^{(true)}$, $x_2^{(b)} - x_2^{(true)}$, and $x_3^{(b)} - x_3^{(true)}$ during the assimilation experiment. The true error for each model variable is shown in blue, whereas the model errors estimated using the asymptotic error estimators in (5.8) - (5.10) are depicted by the red dashed lines. For the asymptotic model error estimates, $x_1(\rho_0, t_0)$ is taken to be its instantaneous value at each assimilation time. Inspection of the error time series (Figs. 5.3a-c) reveals that the asymptotic error estimators are able to accurately capture the magnitude of the true errors in the model background, as well as their rapid changes with time, when all other errors in the system are eliminated. The model errors display more rapid variations than the ρ parameter (Fig. 5.3d) because the time step used by the coupled model is

five times faster than that used in the hidden model $S2$ to perturb ρ. The true ρ parameter oscillates in a quasi-periodic manner for an extended period of time either below or above $\rho = 28$, with occasional transitions between values less than or greater than this threshold as the hidden model driving the changes in ρ_{true} propagates from one wing of the butterfly to the other (see Fig. 5.2a). These quasi-periodic oscillations could be thought of as representing biases associated with the diurnal or seasonal cycles in atmospheric models.

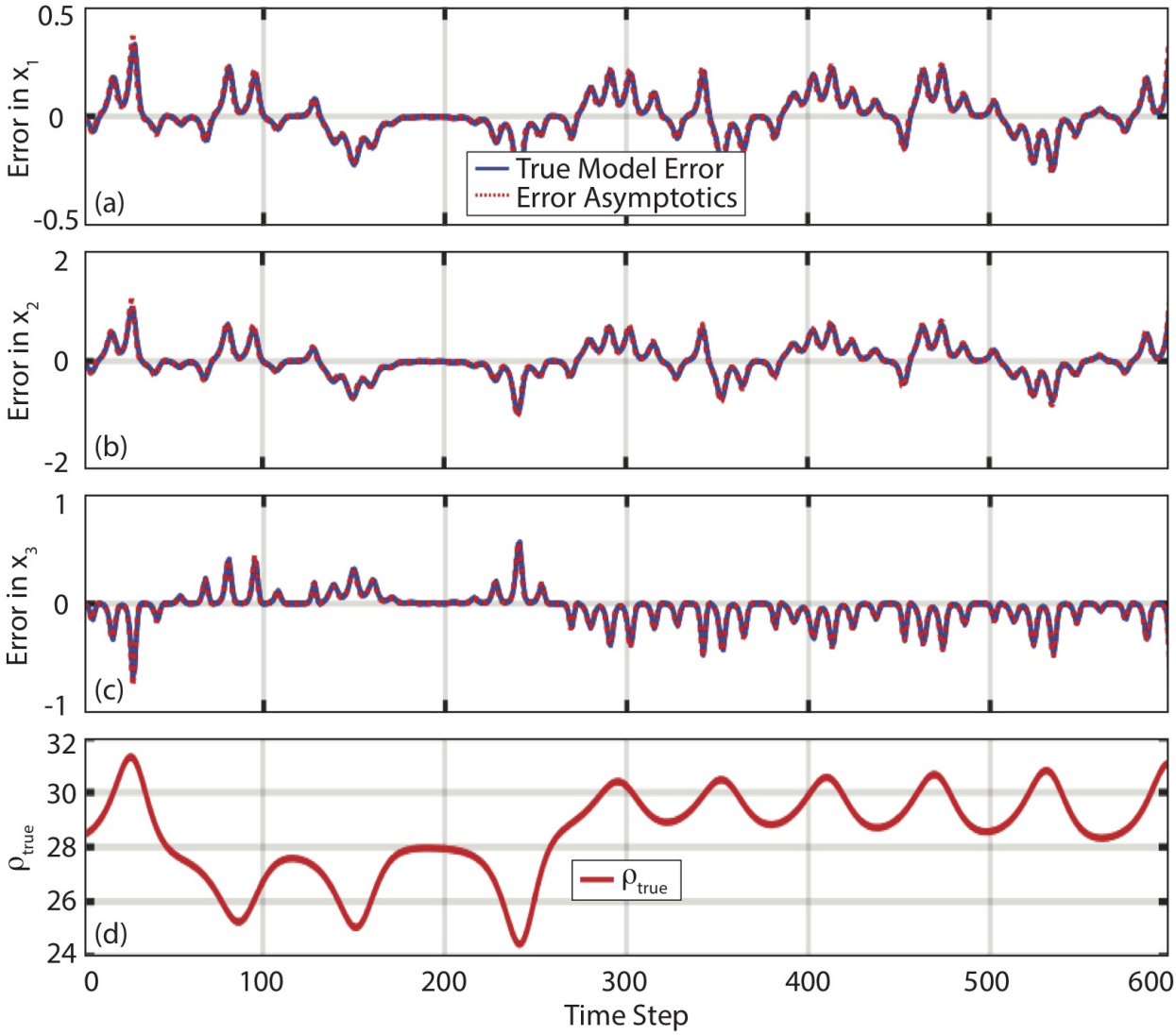

Figure 5.3: Time series showing the evolution of the true model error (blue lines) and asymptotic error estimations (red dashed lines) for the (a) x_1, (b) x_2, and (c) x_3 model state variables and for the (d) ρ_{true} parameter (red line) for an experiment lasting $N_t = 600$ assimilation cycles with $\delta t_{assim} = 0.06$ and the measurement error ε set to zero.

5.5.2 Using Bias Estimators to Improve Assimilation Performance

The development of methods to accurately estimate the model background error covariance matrix B is important for all data assimilation algorithms. In this section, we demonstrate that the assimilation quality, as measured using OMB statistics, can be improved through inclusion of appropriate model error estimators during the data assimilation step. We also examine the optimality of using either a fixed or dynamically varying B matrix and assess the influence of the observation error on these estimates.

For this exercise, we performed cycled 3DVAR data assimilation experiments using two versions of the L63 model where we chose to use a constant $\delta\rho = 1$ in the truth simulation or where we allowed $\delta\rho$ to vary with time based on the influence of the hidden system $S2$ described in Section 5.3.1. The first version is used to represent a situation where a given parameter that does not vary in the real world is assigned the wrong constant value in the numerical model. Here, we assume that we know the asymptotics describing the sensitivity of the model to small perturbations in ρ, but that we do not know the correct scaling factor c for $\delta\rho$. In other words, we know the true value of $\delta\rho$ only up to a constant $c \in \mathbb{R}$, which includes the case of a constant but unknown $\delta\rho$. For brevity, this section only includes results for the scenario in which $\delta\rho$ is allowed to vary with time. Note that even though the errors in the asymptotic estimates will be larger in this situation because the maximum size of $\delta\rho$ is larger, the conclusions regarding the importance of using the dynamically varying b matrix are the same for the experiments using the constant and time-varying $\delta\rho$ perturbations.

To assess the sensitivity to the matrix B, we initially performed an experiment where a constant covariance matrix of the form $B = b \cdot I \in \mathbb{R}^{3\times 3}$ was used during each assimilation cycle, where b is used to scale the identity matrix. We then searched for the constant b that produced the smallest OMB errors averaged over $N_t = 600$ assimilation cycles. Finally, we repeated the search using a dynamical B matrix, which as in (5.55), is the sum of a constant matrix as in (5.51) and a dynamical part as given by the term (5.52) that is computed using the model error estimators described in Theorem 1. The form of $B = B_k$ at time t_k, with the index $k = 1, 2, ..., N_t$ of analysis steps, is chosen as:

$$B_k = b \cdot \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} + \begin{pmatrix} error_{1,k}^2 & 0 & 0 \\ 0 & error_{2,k}^2 & 0 \\ 0 & 0 & error_{3,k}^2 \end{pmatrix}, \tag{5.71}$$

where the diagonal elements in the second part of (5.71) are defined as:

$$error_{1,k} = c \cdot 0.5 \cdot \sigma \cdot x_1(\rho_0, t_k) \cdot (\delta t)^2 \cdot \delta\rho_k \tag{5.72}$$

$$error_{2,k} = c \cdot x_1(\rho_0, t_k) \cdot \delta t \cdot \delta\rho_k \tag{5.73}$$

$$error_{3,k} = c \cdot x_1^2(\rho_0, t_k) \cdot (\delta t)^2 \cdot \delta\rho_k \tag{5.74}$$

Equations (5.72) - (5.74) correspond to the model first guess errors for x_1, x_2, and x_3, respectively, for each assimilation time t_k. The numerical experiments evaluated in this section were carried out using $c = 1$. We chose to use the model error estimators to improve the quality of the B_k matrix rather than to develop a new model error covariance matrix $Q \in \mathbb{R}^{n \times n}$ because the B matrix is estimated using imperfect forecasts that include a model error component as was discussed in Section 5.4.3. Our model error estimation method is able to account for at least some portion of the model error in the B matrix without having to go through the computational expense of estimating and inverting an independent model error covariance matrix Q during the data assimilation step.

Two examples illustrating the relationship between the size of b and the average model first guess errors when using either the constant or dynamic estimates for B during the assimilation experiments are shown in Fig. 5.4. The first example (Fig. 5.4a) has relatively frequent assimilation cycles ($\delta t_{assim} = 0.02$) and small random observation errors ($\varepsilon = 0.2$), whereas the observation errors are larger ($\varepsilon = 0.5$) and the observations are assimilated less frequently ($\delta t_{assim} = 0.04$) during the second example (Fig. 5.4b). Random errors added to each observation were drawn from a Gaussian distribution scaled by the value of ε chosen for each case.

In both examples, the behavior of the relationship shown in Fig. 5.4 is well-known in the field of inverse problems where a regularization that is too small increases the influence of the observation errors and a regularization that is too large will not be able to fully exploit the new information provided by the observations. The optimal B, which varies depending upon the observation and model errors present during a given assimilation cycle, will lead to the smallest first guess errors. Of importance for this discussion is that the smallest first guess errors for both examples occur when the dynamic B matrix is used. It is also evident that the optimal size of b decreases when the dynamical error estimators are used to scale B because they are better able to capture the actual errors in the model background during each assimilation cycle. Together, these examples demonstrate that it is highly desirable to employ dynamical estimators of the model first guess error in data assimilation algorithms.

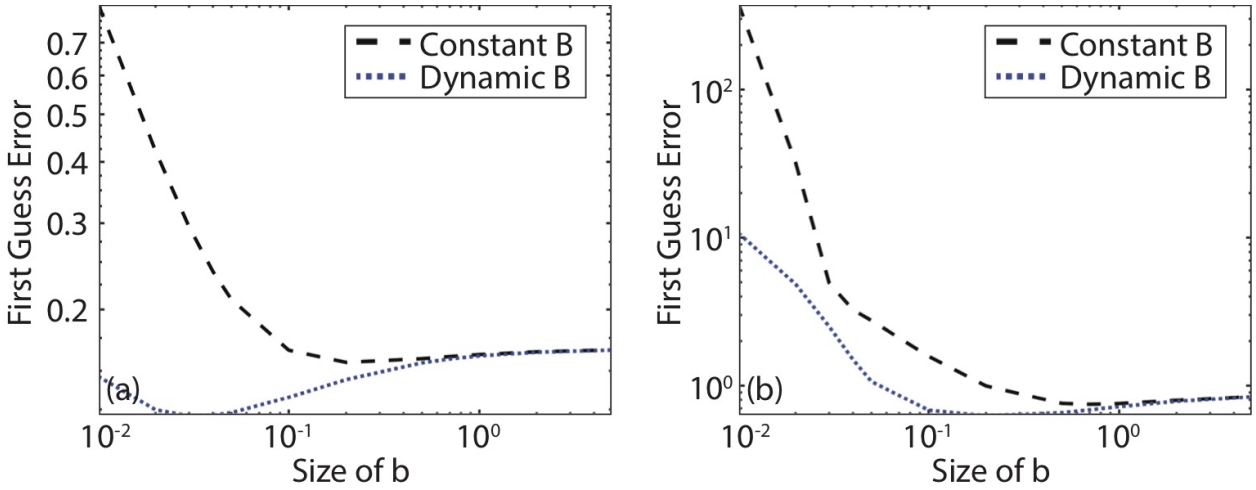

Figure 5.4: Scan of the average model first guess errors plotted as a function of the size of b when the background error covariance matrix B is a multiple of the identity matrix ($B = b \cdot I$) (black dashed line) or when it is obtained using the dynamic B estimator presented in (5.71) (blue dotted line). Panels (a) and (b) show results from experiments using assimilation update intervals δt_{assim} and random observation errors ε set to ($\delta t_{assim} = 0.02, \varepsilon = 0.2$) and ($\delta t_{assim} = 0.04, \varepsilon = 0.5$), respectively. The first guess error statistics were computed using output from 600 time steps.

5.5.3 Numerical Estimation of the Bias Estimator Polynomial Coefficients

In this section, we investigate the determination of the model bias estimator coefficients α using output from cycled 3DVAR experiments employing different assimilation intervals and observation error magnitudes. For these experiments, we employ the dynamical background error covariance matrix B shown in (5.71) during each data assimilation cycle, with the dynamic model errors for (x_1, x_2, x_3) computed using the asymptotic error estimators in (5.72) - (5.74) with the scaling factor c set to 1. Sensitivity tests revealed that the model error coefficients were stable over a broad range of values for the scaling factor b; therefore, for convenience, it was set to 0.1 during the experiments discussed in this section. This behavior and the chosen value for b are consistent with the results shown in Fig. 5.4.

Experimentation also revealed that the matrix A used to determine the bias correction coefficients α in (5.65) is ill-posed with singular values smaller than 10^{-4} and a condition number larger than 10^4. Therefore, to improve its conditioning, Tikhonov regularization was used by replacing the least squares estimator $A^{\dagger} = (A^T A)^{-1} A^T$ in (5.65) with the Tikhonov inverse:

$$Q := (\alpha_{reg} I + A^T A)^{-1} A^T \tag{5.75}$$

where α_{reg} is the Tikhonov regularization parameter. Sensitivity tests showed that setting α_{reg}

to a small value (10^{-5}) provided the most accurate results. This means that the bias correction coefficients for a given model variable can be determined using:

$$\alpha = (\alpha_{reg} I + A^T A)^{-1} A^T q \tag{5.76}$$

Table 5.1 shows results computed using truth-minus-background statistics accumulated over $N_t = 600$ assimilation cycles for two experiments, including one where perfect observations ($\varepsilon = 0$) were assimilated at $\delta t_{assim} = 0.01$ time intervals (left columns) and a second experiment where random errors were added to the observations ($\varepsilon = 0.01$) and the assimilation interval was increased to $\delta t_{assim} = 0.02$. The scaling factor r for the observation error covariance matrix in (5.69) was set to 10^{-5} and 10^{-4}, respectively, for each of these experiments, with $\delta\rho$ for a given time step obtained from the hidden system $S2$ described in Section 5.3.1. The coefficients of the polynomial expansion of the model bias are computed separately for each model variable (x_1, x_2, x_3). Here, we have used all polynomial terms up to the 2nd order when computing the dynamic B matrix in (5.71) because of the presence of the x_1^2 term in the asymptotics shown in (5.59). To ease interpretation of the results, we have included $\delta\rho$ and the constant 0.5, σ, δt, and (δt^2) terms as they appear in (5.72), (5.73), and (5.74) such that the estimation outcomes shown in Table 5.1 should be either 0 or 1 depending upon whether or not a given term is in the polynomial expansion. This means that the reconstructed bias correction coefficient $\alpha_{recon}(1,0,0)$ should equal one for x_1 and x_2, $\alpha_{recon}(2,0,0)$ should equal one for x_3, and all of the other α_{recon} values should be zero.

Inspection of Table 5.1 shows that the maximum error for each state variable (x_1, x_2, x_3) is 8% (e.g., $\alpha_{recon} = 0.92$) for the experiment in which perfect observations were assimilated, and that the errors for most of the remaining α_{recon} terms are very small. This demonstrates that the bias correction coefficients can be accurately estimated in this situation such that the only remaining sources of error are likely associated with numerical discretization errors or the exclusion of higher order polynomial terms from the asymptotical expansion (e.g., higher than the 2nd order). The error in each α_{recon} term increases during the second experiment where measurement errors were added to the observations prior to their assimilation. Even so, the results show that the method is still able to identify the dominant terms and that it is possible to obtain reasonable estimates for the bias correction coefficients in the presence of observation error. Finally, other experiments were performed where the size of the observation error and the length of the assimilation cycling interval were varied, with all of the experiments showing similar effects to those demonstrated in

Table 5.1 if reasonable observation errors and cycling intervals were used.

	Exp 1			Exp 2		
	for x_1	for x_2	for x_3	for x_1	for x_2	for x_3
$\alpha_{recon}(0,0,0)$	4.94E-02	-5.30E-03	-4.06E-03	1.16E-01	-7.26E-03	-1.28E-01
$\alpha_{recon}(1,0,0)$	0.92	1.06	2.37E-03	0.76	1.16	-5.95E-01
$\alpha_{recon}(2,0,0)$	-2.57E-04	4.08E-05	1.03	9.79E-03	1.52E-03	0.94
$\alpha_{recon}(0,1,0)$	2.38E-02	-4.92E-02	1.31E-02	2.65E-02	-1.13E-01	-2.45E-01
$\alpha_{recon}(0,2,0)$	3.13E-04	-2.35E-05	-3.94E-03	2.32E-03	-1.84E-04	5.25E-02
$\alpha_{recon}(0,0,1)$	-1.11E-02	6.09E-04	-4.77E-02	1.88E-04	5.97E-03	-1.30E-01
$\alpha_{recon}(0,0,2)$	2.13E-04	-1.85E-05	2.86E-03	-8.31E-04	-3.04E-04	3.10E-03
$\alpha_{recon}(1,1,0)$	-1.88E-04	-3.84E-06	-4.02E-02	-9.54E-03	-9.32E-04	-2.04E-01
$\alpha_{recon}(1,0,1)$	3.54E-03	-4.07E-04	4.04E-04	5.71E-03	-2.03E-03	2.42E-03
$\alpha_{recon}(0,1,1)$	-2.21E-03	-4.42E-05	-1.08E-04	-4.80E-03	6.95E-05	3.10E-02

Table 5.1: Reconstructed bias correction coefficients (α_{recon}) for each model variable (x_1, x_2, x_3) determined using (5.62) and truth-minus-background statistics accumulated over 600 assimilation cycles for two experiments employing different observation errors and assimilation update intervals. The 0th to 2nd order terms are shown in each row. Columns 2-4 and 5-7 show the results for experiments employing ($\delta t_{assim} = 0.01; \varepsilon = 0$) and ($\delta t_{assim} = 0.02; \varepsilon = 0.01$), respectively. The Tikhonov regularization parameter a_{reg} was set to 10^{-5} for both experiments.

5.5.4 Reconstruction of the ρ Parameter

In this section, we explore the effectiveness of using the bias correction coefficient vector α to reconstruct the ρ parameter within the data assimilation system. The truth simulation for this particular exercise was performed using the coupled L63 model described in Section 5.3.1. A cycled data assimilation experiment covering $N_t = 600$ assimilation cycles with $\delta t_{assim} = 0.04$ was then performed using observations from the truth simulation. Given that the true state of a real-world system is unknown, here we choose to use the analysis-minus-background difference as a proxy for the model error q in (5.65) because the model background $x^{(b)}$ and model analysis $x^{(a)}$ are both readily available from data assimilation systems.

Because ρ varies with time in the coupled L63 system used to perform the truth simulation, it is not advantageous to use assimilation statistics accumulated over a long time period to estimate the value of this parameter for a specific assimilation cycle. Instead, we compute the coefficient vector α using output from 10 consecutive assimilation cycles rather than from the full assimilation period. This length was chosen as a balance between the desire to acquire a large enough sample to robustly estimate $\delta\rho$ and the need to use a short enough time period to ensure that the instantaneous $\delta\rho$ values during a given time interval do not deviate strongly from the mean $\delta\rho$ over that interval. To ease comparison to the reconstructed mean $\delta\rho$, the average of the individual

$\delta\rho$ estimates obtained using the simple diagnostic tools shown in (5.66) are used to represent the true mean $\delta\rho$ over each time period. Together, these choices are consistent with the constraints that would be encountered in a real-world data assimilation system.

Figure 5.5 shows the evolution of the instantaneous model errors $x_1^{(b)} - x_1^{(a)}$, $x_2^{(b)} - x_2^{(a)}$, and $x_3^{(b)} - x_3^{(a)}$, along with the actual and reconstructed values for $\delta\rho$ for three experiments employing different observation errors. The images on the left show the true error for each model variable in blue, whereas the dashed red lines show the model errors estimated using the asymptotic error estimators in Theorem 1. For the images on the right, the black and blue lines denote the true instantaneous and true mean $\delta\rho$ values, respectively, whereas the red lines depict the corresponding mean $\delta\rho$ estimates reconstructed using the α vector. Results are shown for three experiments assimilating observations with measurement errors $\varepsilon = \{0, 0.02, \text{and } 0.04\}$ and scaling factors $r = \{0.0004, 0.0004, \text{and } 0.0016\}$ for the observation error covariance matrices.

Inspection of the time series in Fig. 5.5 reveals that the mean $\delta\rho$ values reconstructed from the coefficient vector α accurately capture the magnitude and evolution of the true $\delta\rho$ for the case where the assimilated observations have zero measurement error (Fig. 5.5b). The asymptotic error estimators also do an excellent job representing the true model errors during this experiment (Fig. 5.5a). As the observation error increases, however, the model error time series become more noisy (Fig. 5.5c, e) and the accuracy of the $\delta\rho$ reconstruction decreases due to the increased noise (Fig. 5.5d, f). The errors in the $\delta\rho$ reconstruction are largest for time periods when the true $\delta\rho$ reaches a local minimum or maximum because the rapid variation with time during those situations makes it more difficult to properly reconstruct $\delta\rho$. Regardless, these results show that it is possible to use the coefficient vector α to obtain useful information about the trajectory of $\delta\rho$ during the truth simulation. Because the true state was not used during this exercise, these results also demonstrate that reasonable parameter and model bias estimates can be obtained using differences between the model analysis and background states. This is important because whereas the true state of a real-world system is generally unknown, the model analysis and background states are both readily available from data assimilation systems.

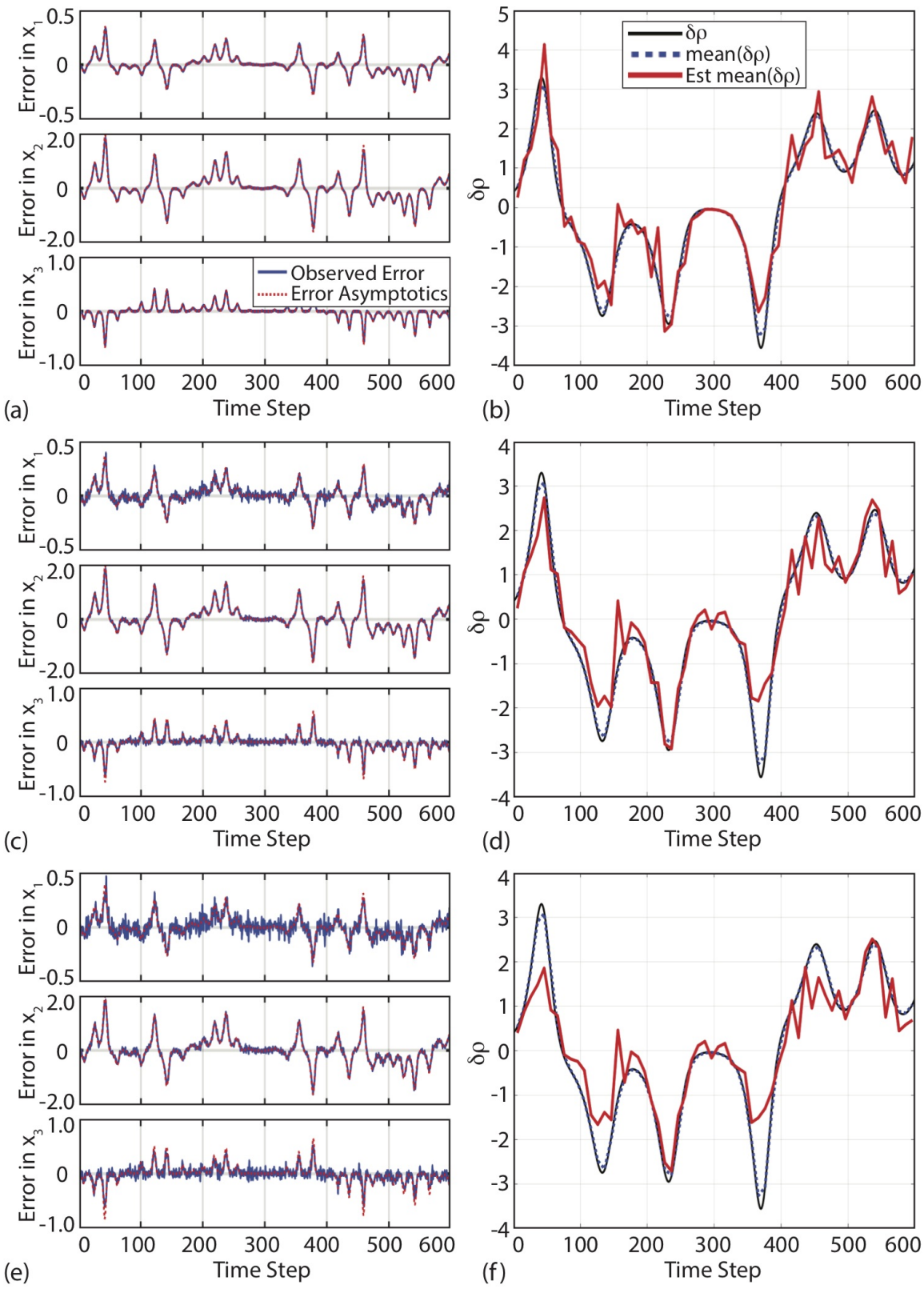

Figure 5.5: (a) Time series showing the evolution of the model error given by the first guess minus analysis (blue line) for $x_1, x_2,$ and x_3, and their estimation computed using the error asymptotics (dashed red line). Here, $\delta t = 0.04$ and $\varepsilon = 0$. (b) Time series showing the evolution of the true $\delta\rho$ (black line). The mean $\delta\rho$ parameter computed over intervals of 10 assimilation cycles is shown by the dashed blue line, with the corresponding dynamic estimation computed using the mean bias correction coefficients shown by the red lines. (c-d) Same as (a-b), except for the case where the assimilation experiment was performed using $\delta t = 0.04$ and $\varepsilon = 0.02$. (e-f) Same as (a-b), except for the case where the assimilation experiment was carried out using $\delta t = 0.04$ and $\varepsilon = 0.04$.

5.6 Conclusions

In this study, we have examined the behavior of dynamic model errors and their influence on the quality of the model analysis and first guess during cycled data assimilation experiments using the L63 model and a 3DVAR data assimilation system. We showed that conditional model biases due to errors in the specification of a model parameter can be represented as a polynomial function that can be estimated using the model background-minus-truth or background-minus-analysis statistics for the realistic situation where the modeling system consists of polynomial forcing terms. We have also suggested a regularized least squares regression method to estimate the model biases and then described how these model error estimators could be used in the data assimilation system to improve the accuracy of the model analysis and first guess.

We have carried out all derivations, estimations, and numerical experiments using the well-known L63 model to demonstrate the validity and feasibility of the ideas developed during this study. The L63 model allows us to study all parts of the system, bias estimators, and tools in a detailed way that would not be possible if we had used a full physics numerical model while still being able to represent the chaotic nonlinear characteristics of the real atmosphere. The results showed that the asymptotics are indeed a valid method to estimate an important part of the model first guess error, and that their use in data assimilation has the potential to improve the accuracy of the model background and analysis. We showed that model error estimators computed using the difference between the model background and analysis, which are readily available from all assimilation systems, are an effective way to estimate model error. In this framework, the model analysis serves as an approximation of the true state, which is unknown in a real-world system. Reasonable results can be achieved even when relatively large errors are present in the observations if Tikhonov regularization is employed during the estimation of the polynomial model error coefficients. Finally, we also show that the polynomial model bias coefficient vector can be used to reconstruct $\delta\rho$ during the assimilation experiments.

In the current work, we have restricted our attention to a small-scale system containing three state variables. Real-world NWP models and data assimilation systems have much deeper complexity and their dimensions are much larger than the system used here. Thus, future work is necessary to investigate the validity of the above ideas in high-dimensional models and to determine if the methods developed during this study can improve the representation of the background error covariance matrix B used by such systems. For the experiments presented in this paper, all of

the state variables were observed during each assimilation cycle, which of course is not possible in a real data assimilation system. It will be important to evaluate the utility of the method when the observation uncertainty is higher or the measurements do not observe the full state of the model. It is reasonable to expect that it will be more difficult to estimate the model errors in such situations. It is also possible that the size of the initial condition uncertainty relative to the model error could impact the performance of this method. For example, the model error contribution to the forecast uncertainty will typically increase relative to the initial condition uncertainty over longer time periods. This would suggest that the model error estimation method may be especially useful for longer assimilation windows or when the observations are assimilated less frequently. A final point to consider is that we already knew which model parameter was incorrectly specified in the L63 model during the data assimilation experiments, which made it possible for us to target its reconstruction using the bias correction coefficient vector. Though this knowledge made the problem easier to solve, it is still consistent with many real-world situations where it is known a priori that a certain parameter varies with time but has been assigned a constant value in the NWP model due to computational constraints or incomplete knowledge on how to predict its evolution. With this knowledge, it should be possible to use the general polynomial expansion of the model variables method developed in Section 5.5.3 to determine if there are relationships between any of the polynomial terms and a chosen parameter and then use that information to reconstruct the value of the parameter.

The dynamic B method developed during this study could be interpreted as providing dynamic additive covariance inflation capturing systematic model errors that are not represented by the static B used by variational systems nor by the dynamic B used by hybrid and EnKF assimilation methods. Inclusion of the dynamic model bias estimates in the B matrix could therefore make it possible to reduce the amount of covariance inflation that is used during the data assimilation step in EnKF systems. This is potentially advantageous because the dynamic B is computed based on the current conditions rather than using random perturbations drawn from a climatology as is typically done with additive covariance inflation methods. It may also provide a complementary approach to weak-constraint 4DVAR where instead of providing the model an additional degree of freedom through introduction of a model error forcing term, we instead enhance the quality of the B matrix through inclusion of the model bias estimates before it is used by the assimilation algorithm. More detailed investigations of these and other topics are left for future work.

5.7 Acknowledgements

The lead author was partially supported by a University of Reading International Research Studentship.

Chapter 6

Conclusions

Data assimilation is a type of inverse problem that seeks to optimally combine information from a set of observations with a spatially continuous first guess analysis in order to generate the best estimate of the 3-dimensional state of a dynamical system at a given time. The first guess analysis is typically provided by a previous model forecast, referred to as the model background, whereas observations are obtained from a variety of sources. Data assimilation is an important part of operational NWP systems because the skill of a model forecast is closely tied to the accuracy of the initial boundary conditions that are provided to the NWP model at the start of a forecast. Though tremendous advancements have been made in our ability to assimilate clear-sky satellite brightness temperatures, progress has been much slower for observations sensitive to clouds. Cloud-impacted observations are challenging to assimilate for a variety of reasons, one of which is that they are often characterized by complex biases that vary with time. These biases can occur due to the inability of the NWP model to properly simulate some cloud processes, inaccuracies in the forward radiative transfer model used to compute the model-equivalent brightness temperatures, and calibration errors in the satellite observations. Regardless of the precise reasons for the biases, their presence will lead to suboptimal results if they are not properly accounted for in the data assimilation system.

Therefore, to aid ongoing efforts within the research and operational NWP modeling communities to more effectively assimilate information from these important observations, this thesis has developed new methods that can be used to identify and remove biases from the observations and model background. This was accomplished through 1) development of an all-sky BC method that uses a Taylor series polynomial expansion of the OMB departures to remove linear

and nonlinear conditional biases from the satellite observations prior to their assimilation, and 2) development of a theoretical model bias estimation method that uses an asymptotic expansion of the model dynamics to improve the model background error covariance matrix used during the data assimilation step. Results and new insights gained from each of these complementary areas of research are described in greater detail in the following sections.

6.1 Nonlinear bias corrections for all-sky satellite data assimilation

For the first research topic, we used output from a high-resolution ensemble data assimilation system to assess the ability of an innovative BC method to remove linear and nonlinear conditional biases from all-sky satellite infrared brightness temperatures. This method removes biases from the satellite observations using a Taylor series polynomial expansion of the OMB departures, with one or more predictors used to capture the error characteristics in the departure distribution. In Chapter 3, the ability of the NBC method to remove the observation biases was assessed using output from cycled data assimilation experiments in which the all-sky infrared brightness temperatures were passively monitored during a 5-day period. After determining that the NBC method was able to effectively remove complex error patterns in the OMB departures, cycled data assimilation experiments were then performed for an independent 3-day period in which all-sky infrared brightness temperatures were actively assimilated at hourly intervals. Results from these active data assimilation experiments were presented in Chapter 4. The main conclusions from this portion of the thesis include:

- The passive monitoring experiments showed that all-sky infrared brightness temperatures are characterized by complex conditional bias patterns that cannot be properly accounted for using traditional linear BC methods designed to assimilate clear-sky observations.
- The linear 1st order term had the largest impact on the entire OMB departure distribution, as measured using reductions in variance; however, it was necessary to use the nonlinear 2nd and 3rd order terms to remove conditional biases that remained in the distribution after the constant and linear bias components were removed.
- Experiments showed that variables sensitive to the cloud top height are the most effective bias predictors for all-sky infrared brightness temperatures, especially when higher order Taylor series terms are used.

- Comparison of the statistics for clear-sky and cloudy-sky matched observations revealed that constant and linear terms are sufficient to remove the bias from clear-sky observations; however, these terms were unable to remove the more complex bias characteristics present in the cloudy-sky observations. This points toward the need to use nonlinear BC predictors when assimilating all-sky satellite brightness temperatures.

- Results from the active data assimilation experiments showed that assimilating the all-sky brightness temperatures without first removing their biases degraded the forecast accuracy based on comparisons to radiosonde observations.

- Removal of the linear and nonlinear conditional biases from the all-sky brightness temperatures prior to their assimilation substantially improved the results compared to the baseline experiment in which no BC was used. A set of experiments showed that errors in 1-h forecasts generally decreased as the order of the BC applied to the observations increased from the 0th order to the 3rd order.

- Cycled data assimilation experiments employing BC predictors sensitive to the cloud top height generally had the smallest errors because they were able to more effectively remove large conditional biases associated with deficiencies in the simulation of upper-level clouds.

- Sensitivity experiments showed that symmetric bias predictors representing the average of the observed and simulated quantities did not improve the model analyses as effectively as the observed predictors did by themselves. This behavior occurred because the relative lack of upper-level clouds in the model analyses prevented the simulated and symmetric quantities from serving as adequate predictors of the observation departure bias. This indicates that, though symmetric predictors have proven utility for all-sky observation error models, they may not be as useful when developing all-sky BC methods.

6.2 Dynamic model bias estimators

For the second research topic, we used the L63 model to develop a method to generate dynamic model error estimators that were then used to improve the model background error covariance matrix **B** employed during the data assimilation step. Unlike the first topic, the goal here was to directly identify model biases rather than overall biases in the OMB departure statistics. The L63 model consists of a set of three coupled ordinary differential equations that together provide a

simplified description of dry convection. It was chosen for this part of the thesis because it is less complex than a full physics model while still maintaining strong nonlinearity representative of many atmospheric processes. It also provides a fully accessible framework with complete knowledge of the model dynamics that promotes a deeper understanding of model biases and how to diagnose them. A model error was introduced to the system by adding small perturbations to the ρ parameter. An asymptotic analysis was performed to assess the sensitivity of the L63 model state variables (x_1, x_2, x_3) to changes in this parameter. A generalized method that uses the model variables to diagnose the model error was then developed, with the dynamic model error estimators subsequently used to modify the **B** matrix. Cycled data assimilation experiments were performed to demonstrate the utility of the model bias estimation method. The main conclusions from this part of the thesis include:

- The asymptotic analysis revealed that the model error asymptotically depends on the model variables, in particular $x_1(p_0, t_0)$, when small perturbations are added to the ρ parameter. The magnitude of the error for each model state variable is proportional to the size of the $\delta\rho$ perturbation.

- Cycled data assimilation experiments demonstrated the feasibility of dynamically estimating the model errors using nonlinear error estimators derived from a polynomial expansion of the model variables. It was shown that the asymptotics accurately represent the behavior of the L63 model and that the coefficients of the asymptotic expansion can be reasonably estimated by solving a regularized least squares regression problem when polynomial terms up to the second order are used. The asymptotic error estimators were also able to accurately capture the magnitude of the true errors in the model background and their rapid changes with time.

- Additional experiments examined the optimality of using either a fixed or dynamically varying **B** matrix during the data assimilation step. It was demonstrated that the combination of a constant **B** matrix with a dynamic **B** matrix that varies with time based on the model bias estimators leads to a more accurate model analysis. The smallest observation-minus-background departures occurred when the observations were assimilated more frequently and had smaller errors.

- A final set of experiments revealed that the bias correction coefficient vector computed using the generalized model bias estimation framework could be used to reconstruct the trajectory

of the ρ parameter. It was shown that the coefficient vector does not depend on the location of the model, but rather on the size of the analysis increment $\mathbf{x}^{(a)} - \mathbf{x}^{(b)}$. This means that it can be used to diagnose the size of the unknown $\delta\rho$ perturbation, without knowing its explicit dependence on the state variables, using only the model background and analysis information that is readily available in all data assimilation systems.

- Together, these results demonstrate that it is highly desirable to employ dynamic estimators of the model first guess error in data assimilation systems.

6.3 Future work

Additional studies are necessary to thoroughly assess the ability of the NBC method to improve the assimilation of all-sky satellite brightness temperatures when using different data assimilation methods, and in other regional and global NWP modeling systems. Because the results presented in Chapters 2 and 4 focused on the SEVIRI 6.2 μm band sensitive to clouds and water vapor in the upper troposphere, it would be beneficial to perform similar experiments using brightness temperatures from other infrared bands. Examples include bands that are sensitive to water vapor in other layers of the troposphere, as well as atmospheric window bands that provide information about surface properties when skies are clear and cloud top properties when clouds are present. Forecasts should also be performed to evaluate how long the improved cloud and water vapor fields in the initialization datasets last during the forecast period, as well as their impact on other aspects of the forecast such as precipitation and surface radiation. Such experiments performed over longer time periods and for different seasons would promote deeper insight into the behavior of the NBC method and its ability to improve the accuracy of the model analyses and forecasts.

It would also be useful to explore the ability of other types of basis functions, such as trigonometric functions, to diagnose nonlinear bias characteristics and how they vary with time. A related question would be to explore the relative merit of employing online bias corrections as were implemented during these experiments, or offline calculations that use error statistics accumulated over many weeks or longer. The latter approach would greatly increase the sample size and stability of the bias correction coefficients; however, this would come at the cost of no longer being able to respond quickly to changes in the error characteristics due to changes in the prevailing cloud regimes. It would also require re-computing the bias correction coefficients any time changes were

made to the NWP or radiative transfer models, or to other aspects of the data assimilation system.

The experiments performed during this thesis are an initial step toward the inclusion of the NBC method in operational data assimilation systems. There are many additional developments that have the potential to substantially increase its impact. For example, further research is necessary to determine if using a Huber norm in combination with the NBC method could improve existing quality control methods by identifying erroneous observations after the nonlinear conditional biases have been removed from the observation departures. There is also great promise in pairing the NBC method to a dynamic all-sky observation error model because that could potentially lead to the more effective use of the information provided by clear-sky and cloudy-sky brightness temperatures by assigning them different observation errors. Incorporation of both of these approaches could potentially lead to more cloud-affected observations being assimilated in sensitive areas of the domain where nonlinear error growth is more likely to occur. Development of an all-sky correlated observation error model is another outstanding issue that requires substantial research efforts. The results also suggest that attention should be given to developing new methods that can increase the ensemble spread in the cloud hydrometeor variables that are very important when assimilating cloud-sensitive observations.

Finally, for the model bias estimation research presented in Chapter 5, we restricted our attention to a small-scale idealized modeling system that consists of three coupled ordinary differential equations. Though this approach was advantageous because it allowed us to have complete knowledge of the model state while developing the model bias estimators, real-world models and data assimilation systems have much deeper complexity and their dimensions are not comparable to the L63 system that was employed here. Thus, it is of high interest to examine the ability of the generalized model bias estimator method to improve the assimilation performance within a high-dimensional, real-world environment. For example, inclusion of the dynamic model bias estimates in the background error covariance matrix $\mathbf{B}$ could make it possible to reduce the amount of covariance inflation used by ensemble data assimilation systems. This could be useful because whereas additive covariance inflation methods typically draw random perturbations from a climatology, the method developed in this thesis computes a dynamic $\mathbf{B}$ matrix using the current conditions. Likewise, incorporation of the dynamic model bias estimation method in weak-constraint 4DVAR systems could directly enhance the quality of the background error covariance matrix $\mathbf{B}$ or be used to estimate the model error covariance matrix $\mathbf{Q}$ that is often difficult to determine in practice. More detailed investigations of these and other topics are left for future work.

Chapter 7

Bibliography

Abramowitz, M., and I. A. Stegun, I. A.: 1972 *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, 9th printing. New York, Dover.

Afraimovich, V., V. V., Bykov, and L. P. Shilnikov, 1977: The origin and structure of the Lorenz attractor. *Sov. Phys. Dokl.*, **22**, 253–255.

Aksoy, A., D. C. Dowell, and C. Snyder, 2009: A multicase comparative assessment of the ensemble Kalman filter for assimilation of radar observations. Part I: Storm-scale analysis. *Mon. Wea. Rev.*, **137**, 1805–1824.

Ambadan, J. T., and Y. Tang, 2009: Sigma-point Kalman filter data assimilation methods for strongly nonlinear systems. *J. Atmos. Sci.*, **66**, 261–285.

Anderson, B. D., and J. B. Moore, 2012: Optimal Filtering. Dover Books on Electrical Engineering Series. Dover Publications, Incorporated.

Anderson J., and S. Anderson, 1999: A Monte Carlo implementation of the nonlinear filtering problem to produce ensemble assimilations and forecasts. *Mon. Wea. Rev.*, **127**, 2741–2758.

Anderson, J. L., 2001: An ensemble adjustment Kalman filter for data assimilation. *Mon. Wea. Rev.*, **129**, 2884–2903.

Anderson, J. L., 2007: An adaptive covariance inflation error correction algorithm for ensemble filters. *Tellus A*, **59**, 210–224.

Anderson, J. L., 2009. Spatially and temporally varying adaptive covariance inflation for ensemble

filters. *Tellus A*, **61**, 72–83.

Aonashi, K. and H. Eito, 2011: Displaced ensemble variations assimilation method to incorporate microwave imager brightness temperatures into a cloud resolving model. *J. Meteor. Soc. Japan*, **89**, 175–194, DOI:10.2151/jmsj.2011-301.

Apostol, T. M., 1974: Mathematical analysis (2nd ed.). Reading, MA, Addison-Wesley.

Aravequia J.A., I. Szunyogh, E. J. Fertig, E. Kalnay, D. Kuhl, and E. J. Kostelich, 2011: Evaluation of a strategy for the assimilation of satellite radiance observations with the local ensemble transform Kalman filter. *Mon. Wea. Rev.*, **139**, 1932–1951, doi: 10.1175/2010MWR3515.1.

Auligne T., A. P. McNally, and D. P. Dee, 2007: Adaptive bias correction for satellite data in a numerical weather prediction system. *Q. J. R. Meteorol. Soc.*, **133**, 631–642.

Baldauf M., A. Seifert, J. Forstner, D. Majewski, M. Raschendorfer, and T. Reinhardt, 2011: Operational convective-scale numerical weather prediction with the COSMO Model: Description and sensitivities. *Mon. Wea. Rev.*, **139**, 3887–3905.

Bannister, R. N., 2008: A review of forecast error covariance statistics in atmospheric variational data assimilation. I: Characteristics and measurements of forecast error covariances. *Q. J. R. Meteorol. Soc.*, **134**, 1951–1970.

Bannister, R. N., 2008: A review of forecast error covariance statistics in atmospheric variational data assimilation. II: Modelling the forecast error covariance statistics. *Q. J. R. Meteorol. Soc.*, **134**, 1971–1996.

Bannister, R. N., 2017: A review of operational methods of variational and ensemble variational data assimilation. *Q. J. R. Meteorol. Soc.*, **143**, 607–633.

Baordo F., and A. J. Geer, 2016: Assimilation of SSMIS humidity-sounding channels in all-sky conditions over land using a dynamic emissivity retrieval. *Q. J. R. Meteorol. Soc.*, **142**, 2854–2866, https://doi.org/10.1002/qj.2873.

Barbeau, E.J., 2003: Polynomials. Springer Science and Business Media, 455 pp.

Barrio, R., A. Shlnikov, and L. Shilnikov, 2012: Kneadings, symbolic dynamics and painting Lorenz chaos. *Int. J. Bifurcation and Chaos*, **22** 1230016, doi:10.1142/S0218127412300169.

Bauer P., A. J. Geer, P. Lopez, and D. Salmond, 2010: Direct 4D-Var assimilation of all-sky radi-

ances: Part I. Implementation. *Q. J. R. Meteorol. Soc.*, **136**, 1868–1885.

Bauer P., and Coauthors, 2011: Satellite cloud and precipitation assimilation at operational NWP centres. *Q. J. R. Meteorol. Soc.*, **137**, 1934–1951.

Baum, B. A., P. Yang, A. J. Heymsfield, A. Bansemer, A. Merrelli, C. Schmitt, and C. Wang, 2014: Ice cloud bulk single-scattering property models with the full phase matrix at wavelengths from 0.2 to 100 μm. *J. Quant. Spectrosc. Radiat. Transfer*, **146**, 123–139, doi:10.1016/j.jqsrt.2014.02.029.

Bengtsson L, and K. Hodges, 2005: On the impact of humidity observations in numerical weather prediction. *Tellus*, **57A**, 701–708.

Berner, J., and Coauthors, 2017: Stochastic parameterization: Toward a new view of weather and climate models. *Bull. Amer. Meteor. Soc.*, **98**, 565–588.

Bessho, K., and Coauthors, 2016: An introduction to Himawari-8/9—Japan's new-generation geostationary meteorological satellites. *J. Meteor. Soc. Japan*, **94**, 151–183, https://doi.org/10.2151/jmsj.2016-009.

Bocquet M., C. A. Pires, and L. Wu, 2010: Beyond Gaussian statistical modelling in geophysical data assimilation. *Mon. Wea. Rev.*, **138**, 2997–3023.

Bonavita, M., L. Torrisi, and F. Marcucci, 2010: Ensemble data assimilation with the CNMCA regional forecasting system. *Q. J. R. Meteorol. Soc.*, **136**, 132–145.

Bonesky, T., 2009: Morozov's discrepancy principle and Tikhonov-type functionals. *Inverse Problems*, **25**, Article ID 015015.

Bormann, N., and P. Bauer, 2010a: Estimates of spatial and interchannel observation-error characteristics for current sounder radiances for numerical weather prediction. I: Methods and application to ATOVS data. *Q. J. R. Meteorol. Soc.*, **136**, 1036-1050.

Bormann, N., A. Collard, and P. Bauer, 2010b: Estimates of spatial and interchannel observation-error characteristics for current sounder radiances for numerical weather prediction. II: Application to AIRS and IASI data. *Q. J. R. Meteorol. Soc.*, **136**, 1051-1063.

Bormann, N., A. J. Geer, and P. Bauer, 2011: Estimates of observation-error characteristics in clear and cloudy regions for microwave imager radiances from NWP. *Q. J. R. Meteorol. Soc.*, **137**, 2014–2023.

Bormann, N., M. Bonavita, R. Dragani, R. Eresmaa, M. Matricardi, and A. McNally, 2016: Enhancing the impact of IASI observations through an updated observation-error covariance matrix. *Q. J. R. Meteorol. Soc.*, **142**, 1767–178, https://doi.org/10.1002/qj.2774.

Boukabara, S. A., and Coauthors, 2016: S4: An O2R/R2O infrastructure for optimizing satellite data utilization in NOAA numerical modeling systems, a significant step toward bridging the valley of death. *Bull. Am. Meteorol. Soc.*, **97**, 2379–2394.

Buehner, M., P. Gauthier, and Z. Liu, 2005: Evaluation of new estimates of background- and observation-error covariances for variational assimilation. *Q. J. R. Meteorol. Soc.*, **131**, 3373-3383.

Buehner, M., P. L. Houtekamer, C. Charette, H. L. Mitchell, and B. He, 2010: Intercomparison of variational data assimilation and then ensemble Kalman filter for global deterministic NWP. Part II: One-month experiments with real observations. *Mon. Wea. Rev.*, **138**, 1567–1586.

Buizza, R., M. Miller, and T. N. Palmer, 1999: Stochastic representation of model uncertainties in the ecmwf ensemble prediction system. *Q. J. R. Meteorol. Soc.*, **125**, 2887–2908.

Cameron, J., and W. Bell, 2016: The testing and planned implementation of variational bias correction (VarBC) at the Met Office. *20th International TOVS study conference*, Madison, WI.

Campbell, W. F., E. A. Satterfield, B. Ruston, and N. L. Baker, 2017: Accounting for correlated observation error in a dual-formulation 4D variational data assimilation system. *Mon. Wea. Rev.*, **145**, 1019–1032.

Caron, J.-F., T. Milewski, M. Buehner, L. Fillion, M. Reszka, S. Macpherson, and J. St-James, 2015: Implementation of deterministic weather forecasting systems based on ensemble-variational data assimilation at Environment Canada. Part II: The regional system. *Mon. Wea. Rev.*, **143**, 2560–2580.

Carrassi, C., and S. Vannitsem, 2010: Accounting for model error in variational data assimilation: A deterministic formulation. *Mon. Wea. Rev.*, **138**, 3369–3386.

Cintineo, R., J. A. Otkin, M. Xue, and F. Kong, 2014. Evaluating the performance of planetary boundary layer and cloud microphysical parameterization schemes in convection permitting ensemble forecasts using synthetic GOES-13 satellite observations. *Mon. Wea. Rev.*, **142**, 163–182.

Cintineo, R., J. A. Otkin, T. Jones, S. Koch, and D. J. Stensrud, 2016: Assimilation of synthetic GOES-R ABI infrared brightness temperatures and WSR-88D radar observations in a high-resolution OSSE. *Mon. Wea. Rev.*, **144**, 3159–3180.

Clayton, A. M., A. C. Lorenc, and D. M. Barker, 2013: Operational implementation of a hybrid ensemble 4D-Var global data assimilation system at the Met Office. *Q. J. R. Meteorol. Soc.*, **139**, 1445-1461.

Courtier P., W. Heckley, J. Pailleux, D. Vasiljevic, M. Hamrud, A. Holingsworth, and F. Rabier, 1998: The ECMWF implementation of three-dimensional variational assimilation (3D-Var). Part 1: formulation. *Q. J. R. Meteorol. Soc.*, **124**, 1783–1807.

Dee, D. P., 2005: Bias and data assimilation. *Q. J. R. Meteorol. Soc.*, **131**, 3323–3343.

Dee, D. P, 1995: On-line estimation of error covariance parameters for atmospheric data assimilation. *Mon. Wea. Rev.*, **123**, 1128–1145.

Dee D. P., and S. Uppala, 2009: Variational bias correction of satellite radiance data in the ERA-Interim reanalysis. *Q. J. R. Meteorol. Soc.*, **135**, 1830–1841.

Derber, J. C., 1989: A variational continuous assimilation technique. *Mon. Wea. Rev.*, **117**, 2437–2446.

Derber J. C., D. F. Parrish, and S. J. Lord, 1991: The new global operational analysis system at the National Meteorological Center. *Wea. Forecasting*, **6**, 538–547.

Derber, J. C., and W.-S. Wu, 1998: The use of TOVS cloud-cleared radiances in the NCEP SSI analysis system. *Mon. Wea. Rev.*, **126**, 2287–2299.

Derrien M., and H. Le Gleau, 2005: MSG/SEVIRI cloud mask and type from SAF NWC. *Int. J. Remote Sens.*, **26**, 4707–4732.

Desroziers, G., L. Berre, B. Chapnik, and P. Poli, 2005: Diagnosis of observation, background and analysis-error statistics in observation space. *Q. J. R. Meteorol. Soc.*, **131**, 3385-3396.

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

Dowell, D. C., F. Zhang, L. Wicker, C. Snyder, and N. A. Crook, 2004: Wind and temperature retrievals in the 17 May 1981 Arcadia Oklahoma, supercell: Ensemble Kalman filter experiments.

Mon. Wea. Rev., **132**, 1982–2005.

Eikenberg, S., C. Kohler, A. Siefert, and S. Crewell, 2015: How microphysical choices affect simulated infrared brightness temperatures. *Atmos. Research*, **156**, 67–79.

Errico, R., P. Bauer, and J.-F. Mahfouf, 2007: Issues regarding the assimilation of cloud and precipitation data. *J. Atmos. Sci.*, **64**, 3685–3798.

Evensen, G., 1994: Sequential data assimilation with a nonlinear quasi-geostrophic model using monte carlo methods to forecast error statistics. *J. Geophys. Res. Oceans*, **99**, 10143–10162.

Evensen, G., 1997: Advanced data assimilation for strongly nonlinear dynamics. *Mon. Wea. Rev.*, **125**, 1342–1354.

Evensen, G., 2009: Data Assimilation: The Ensemble Kalman Filter. Earth and Environmental Science. Springer.

Eyre J. R., 1992: A bias correction scheme for simulated TOVS brightness temperatures. *Technical Memorandum 176*, European Centre for Medium-range Weather Forecasts, Reading, UK.

Eyre, J. R., 2016: Observation bias correction schemes in data assimilation systems: A theoretical study of some of their properties. *Q. J. R. Meteorol. Soc.*, **142**, 2284–2291.

Fabry, F., and J. Sun, 2010: For how long should what data be assimilated for the mesoscale forecasting of convection and why? Part I: On the propagation of initial condition errors and their implications for data assimilation. *Mon. Wea. Rev.*, **138**, 242–255.

Fertig E.J., S.-J. Baek, B. R. Hunt, E. Ott, I. Szunyogh, J. A. Aravequia, E. Kalnay, H. Li, and J. Liu, 2009: Observation bias correction with an ensemble Kalman filter. *Tellus*, **61A**, 210–226, doi: 10.1111/j.1600-0870.2008.00378.x.

Fisher, M., 2005: Accounting for correlated observation error in the ECMWF analysis. *Technical Memorandum R48.3*, European Centre for Medium-range Weather Forecasts, Reading, UK.

Fujita, T., D. J. Stensrud, and D. C. Dowell, 2007: Surface data assimilation using an ensemble Kalman filter approach with initial condition and model physics uncertainties. *Mon. Wea. Rev.*, **135**, 1846–1868.

Gandin, L. S., 1963: "Objective analysis of meteorological fields." Leningrad: Gridromet (in Russian), English translation by the Israel program for scientific translations, 1965, 242pp.

Geer A. J., and P. Bauer, 2011: Observation errors in all-sky data assimilation. *Q. J. R. Meteorol. Soc.*, **137**, 2024–2037.

Geer A. J., P. Bauer, and S. J. English, 2012: Assimilating AMSU-A temperature sounding channels in the presence of cloud and precipitation. *Technical Memorandum 670*, European Centre for Medium-range Weather Forecasts (ECMWF), Reading, UK.

Geer A. J., F. Baordo, N. Bormann, and S. English, 2014: All-sky assimilation of microwave humidity sounders. *Technical Memorandum 741*, European Centre for Medium-range Weather Forecasts (ECMWF), Reading, UK.

Geer, A. J., and Coauthors, 2017: The growing impact of satellite observations sensitive to humidity, cloud and precipitation. *Q. J. R. Meteorol. Soc.*, **143**, 3189–3206, https://doi.org/10.1002/qj.3172.

Geer, A. J., and Coauthors, 2018: All-sky satellite data assimilation at operational weather forecasting centre. *Q. J. R. Meteorol. Soc.*, **144**, 1191–1217, https://doi.org/10.1002/qj.3202.

Geer, A. J., S. Migliorini, and M. Matricardi, 2019: All-sky assimilation of infrared radiances sensitive to mid- and upper-tropospheric moisture and cloud. *Atmos. Meas. Tech.*, under review.

Le Gleau H., 2016: Algorithm theoretical basis document for the cloud products processors of the NWC/GEO. http://www.nwcsaf.org (accessed 29 March 2017).

Goodliff, M., J. Amezcua, and P. J. van Leeuwen, 2015: Comparing hybrid data assimilation methods on the Lorenz 1963 model with increasing non-linearity. *Tellus A*, **67**, 26928.

Griffin, S. M., J. A. Otkin, G. Thompson, M. Frediani, and J. Berner, 2020: Assessing the impact of stochastic cloud microphysics using GOES-16 infrared brightness temperatures. Conditionally accepted for publication in *Mon. Wea. Rev*.

Griffith, A. K., and N. K. Nichols, 2000: Adjoint methods in data assimilation for estimating model error. *Flow, Turbulence and Combustion*, **65**, 469-488.

Guidard V., N. Fourrie, P. Brousseau, and F. Rabier, 2011: Impact of IASI assimilation at global and convective scales and challenges for the assimilation of cloudy scenes. *Q. J. R. Meteorol. Soc.*, **137**, 1975–1987.

Gustafsson, N., and Coauthors, 2018: Survey of data assimilation methods for convective-scale numerical weather prediction at operational centres. *Q. J. R. Meteorol. Soc.*, **144**, 1218–1256,

https://doi.org/10.1002/qj.3179.

Ha, S., J. Berner, and C. Snyder, 2015: A comparison of model error representations in mesoscale ensemble data assimilation. *Mon. Wea. Rev.*, **143**, 3893–3911.

Hamill, T. M., J. S. Whitaker, and C. Snyder, 2001: Distance-dependent filtering of background error covariance estimates in an ensemble Kalman filter. *Mon. Wea. Rev.*, **129**, 2776–2790.

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

Harnisch F., M. Weissmann, and A. Perianez, 2016: Error model for the assimilation of cloud-affected infrared satellite observations in an ensemble data assimilation system. *Q. J. R. Meteorol. Soc.*, **142**, 1797–1808.

Harris, B. A, and G. Kelly, 2001: A satellite radiance-bias correction scheme for data assimilation. *Q. J. R. Meteorol. Soc.*, **127**, 1453–1468.

Heillette, S., and L. Garand, 2007: A practical approach for the assimilation of cloudy infrared radiances and its evaluation using AIRS simulated observations. *Atmos. Ocean*, **45**, 211–225.

Hestenes M. R., and E. Stiefel, 1952: Methods of conjugate gradients for solving linear systems. *J. Res. National Bureau Standards*, **49**, 409–436.

Hilton F., N. C. Atkinson, S. J. English, and J. R. Eyre, 2009: Assimilation of IASI at the Met Office and assessment of its impact through observing system experiments. *Q. J. R. Meteorol. Soc.*, **135**, 495–505.

Hocking J., P. Rayer, R. Saunders, M. Matricardi, A. Geer, P. Brunel, 2011: RTTOV v10 Users Guide, NWC SAF report. EUMETSAT: Darmstadt, Germany, 92 pp.

Hodyss, D., 2011: Ensemble state estimation for nonlinear systems using polynomial expansions in the innovation. *Mon. Wea. Rev.*, **139**, 3571–3588.

Hoerl, A. E., and R. W. Kennard, 2000: Ridge regression: biased estimation for nonorthogonal problems. *Technometrics*, **42**, 80–86.

Hoffman, R. N., V. K. Kumar, S.-A. Boukabara, K. Ide, F. Yang, and R. Atlas, 2018: Progress in forecast skill at three leading global oprational NWP centers during 2015-2017 as seen in summary assessment metrics (SAMs). *Wea. Forecasting*, **33**, 1661–1679.

Honda, T., and Coauthors, 2018a: Assimilating all-sky Himawari-8 satellite infrared radiances: A case of Typhoon Soudelor (2015). *Mon. Wea. Rev.*, **146**, 213–229, https://doi.org/10.1175/MWR-D-16-0357.1.

Honda, T., S. Kotsuki, G.-Y. Lien, Y. Maejima, K. Okamoto, and Y. Miyoshi, 2018b: Assimilation of Himawari-8 all-sky radiances every 10 minutes: Impact on precipitation and flood risk prediction. *J. Geophys. Res.*, **123**, 965–976, 10.1002/2017JD027096.

Houtekamer, P. L., and H. L. Mitchell, 1998: Data assimilation using an ensemble Kalman filter technique. *Mon. Wea. Rev.*, **126**, 796–811.

Houtekamer, P. L., and H. L. Mitchell, 2001: A sequential ensemble Kalman filter for atmospheric data assimilation. *Mon. Wea. Rev.*, **129**, 123–137.

Houtekamer, P. L, and H. L. Mitchell, 2005: Ensemble Kalman filtering. *Q. J. R. Meteorol. Soc.*, **131**, 3269–3289.

Houtekamer, P. L., H. L. Mitchell, and X. Deng, 2009: Model error representation in an operational ensemble Kalman filter. *Mon. Wea. Rev.*, **137**, 2126–2143.

Houtekamer, P. L, H. L. Mitchell, G. Pellerin, M. Buehner, M. Charron, L. Spacek, and B. Hansen, 2005: Atmospheric data assimilation with an ensemble Kalman filter: Results with real observations. *Mon. Wea. Rev.*, **133**, 604–620.

Houtekamer, P. L., and F. Zhang, 2016: Review of the ensemble Kalman filter for atmospheric data assimilation. *Mon. Wea. Rev.*, **144**, 4489–4532.

Howes, K. E., A. M. Fowler, and A. S. Lawless, 2017: Accounting for model error in strong-constraint 4D-Var data assimilation. *Q. J. R. Meteorol. Soc.*, **143**, 1227-1240.

Huber P. J., 1972: Robust statistics: A review. *Ann. Math. Stat.*, **43**, 1041–1067.

Hunt B. R., E. J. Kostelich, and I. Szunyogh, 2007: Efficient data assimilation for spatiotemporal chaos: A local ensemble transform Kalman filter. *Physica D*, **230**, 112–126, doi:10.1016/j.physd.2006.11.008.

Isaksen, L., M. Bonavita, R. Buizza, M. Fisher, J. Haseler, M. Leutbecher, and L. Raynaud, 2010: Ensemble of data assimilations at ECMWF. *Technical Memorandum 636*, European Centre for Medium-range Weather Forecasting (ECMWF), Reading, UK.

Janjic, T., N. Bormann, M. Bocquet, J. A. Carton, S. E. Cohn, S. L. Dance, S. N. Losa, N. K. Nichols, R. Potthast, J. A. Waller, and P. Weston, 2018: On the representation error in data assimilation. *Q. J. R. Meteorol. Soc.*, **144**, 1257–1278.

Jia, Y., Z. C. Yang, and Q. Z. Song, 2005: Experimental study of random dynamic loads identification based on weighted regularization method. *Journal of Sound and Vibration*, **342**, 113–123.

Jones, T. A., J. A. Otkin, D. J. Stensrud, and K. Knopfmeier, 2013a: Assimilation of satellite infrared radiances and Doppler radar observations during a cool season observing system simulation experiment. *Mon. Wea. Rev.*, **141**, 3273–3299, doi:10.1175/MWR-D-12-00267.1.

Jones, T. A, D. J. Stensrud, P. Minnis, and R. Palikonda, 2013b: Evaluation of a forward operator to assimilate cloud water path into WRF-DART. *Mon. Wea. Rev.*, **141**, 2272–2289, doi:10.1175/MWR-D-12-00238.1.

Jones, T. A., J. A. Otkin, D. J. Stensrud, and K. Knopfmeier, 2014: Forecast evaluation of an Observing System Simulation Experiment assimilating both radar and satellite data. *Mon. Wea. Rev.*, **142**, 107–124, doi:10.1175/MWR-D-13-00151.1.

Jones, T. A., and D. J. Stensrud, 2015: Assimilating cloud water path as a function of model cloud microphysics in an idealized simulation. *Mon. Wea. Rev.*, **143**, 2052–2081, doi:10.1175/MWR-D-14-00266.1.

Jones, T. A., D. J. Stensrud, L. Wicker, P. Minnis, and R. Palikonda, 2015: Simultaneous radar and satellite data storm-scale assimilation using an ensemble Kalman filter approach for 24 May 2011. *Mon. Wea. Rev.*, **143**, 165–194, doi:10.1175/ MWR-D-14-00180.1.

Jones, T. A., K. Knopfmeier, D. Wheatley, G. Creager, P. Minnis, and R. Palikondo, 2016: The NSSL Multiscale Ensemble. Part II: Combined radar and satellite assimilation. *Wea. Forecasting*, **31**, 297–327, doi:10.1175/WAF-D-15-0107.1.

Jones, T. A., X. G. Wang, P. Skinner, A. Johnson, and Y. M. Wang, 2018: Assimilation of GOES-13 Imager Clear-Sky Water Vapor (6.5 μm) Radiances into a Warn-on-Forecast System. *Mon. Wea. Rev.*, **146**, 1077–1107, doi:10.1175/MWR-D-17-0280.1.

Kalman, R. E., 1960: A new approach to linear filtering and prediction problems. *J. Basic Eng.*, **82**, 35–45.

Kalnay, E., 2003: Atmospheric Modeling, Data Assimilation and Predictability. Cambridge Uni-

versity Press.

Kaplan, J. L., and J. A. Yorke, 1979: Preturbulence: A regime observed in a fluid flow model of Lorenz. *Commun. Math. Phys.*, **67**, 93–108.

Kazumori, M., A. J. Geer, and S. J. English, 2016: Effects of all-sky assimilation of GCOM-W/AMSR2 radiances in the ECMWF numerical weather prediction system. *Q. J. R. Meteorol. Soc.*, **142**, 721–737, https://doi.org/10.1002/qj.2669.

Kerr, C. A., D. J. Stensrud, and X. Wang, 2015: Assimilation of cloud-top temperature and radar observations of an idealized splitting supercell using an observing system simulation experiment. *Mon. Wea. Rev.*, **143**, 1018–1034, https://doi.org/10.1175/MWR-D-14-00146.1.

Kindermann, S., and A. Neubauer, 2008: On the convergence of the quasioptimality criterion for (iterated) Tikhonov regularization. *Inverse Problems and Imaging*, **2**, 291–299.

Klaes, K. D., and Coauthors, 2007: An introduction to the EUMETSAT Polar System. *Bull. Am. Meteor. Soc.*, **88**, 1085–1096.

Kleist, D. T., D. F. Parrish, J. C. Derber, R. Treadon, W.-S. Wu, and S Lord, 2009: Introduction of the GSI into the NCEP global data assimilation system. *Wea. Forecasting*, **24**, 1691–1705.

Klinker, E., F. Rabier, G. Kelly, and J-F. Mahfouf, 2000: The ECMWF operational implementation of four-dimensional variational assimilation. III: Experimental results and diagnostics with operational configuration. *Q. J. R. Meteorol. Soc.*, **126**, 1191–1215.

Kostka P. M., M. Weissmann, R. Buras, B. Mayer, and O. Stiller, 2014: Observation operator for visible and near-infrared satellite reflectances. *J. Atmos. Oceanic Technol.*, **31**, 1216–1233.

Kurzrock, F., S. Cros, F. Chane-Ming, J. A. Otkin, L. Linguet, A. Hutt, G. Lajoie, and R. Potthast, 2018: The assimilation of geostationary meteorological satellite observations in regional NWP for cloudiness forecasting. *Meteorologische Zeitschrift*, **27**, 277–298.

Langland, R. H., and N. L. Baker, 2004: Estimation of observation impact using the NRL atmospheric variational data assimilation adjoint system. *Tellus A*, **56**, 189–201.

Lawrence, H., N. Bormann, A. J. Geer, Q. Lu, and S. J. English, 2018: Evaluation and Assimilation of the Microwave Sounder MWHS-2 Onboard FY-3C in the ECMWF Numerical Weather Prediction System. *IEEE Transactions on Geoscience and Remote Sensing*, **56**, 3333–3349, DOI:10.1109/TGRS.2018.2798292.

van Leeuwen, P. J., Y. Cheng, and S. Reich, 2015: Nonlinear Data Assimilation. Frontiers in Applied Dynamical Systems: Reviews and Tutorials. Springer.

Lei, J., and P. Bickel, 2011: A moment matching ensemble filter for nonlinear non Gaussian data assimilation. *Mon. Wea. Rev.*, **139**, 3964–3973.

Lei, L., D. R. Stauffer, S. E. Haupt, and G. S. Young, 2012: A hybrid nudging ensemble Kalman filter approach to data assimilation. Part I: Application in the Lorenz system. *Tellus A*, **64**, 18484.

Lewis, J. M., and J. C. Derber, 1985: The use of the adjoint equation to solve a variational adjustment problem with advective constraints. *Tellus A*, **37**, 309–322.

Lien, G.-Y., T. Miyoshi, and E. Kalnay, 2016: Assimilation of TRMM Multisatellite Precipitation Analysis with a Low-Resolution NCEP Global Forecast System. *Mon. Wea. Rev.*, **144**, 643–661.

Li, H., E. Kalnay, T. Miyoshi, and C. M. Danforth, 2009: Accounting for model errors in ensemble data assimilation. *Mon. Wea. Rev.*, **137**, 3407–3419.

Li, H., E. Kalnay, and T. Miyoshi, 2009: Simultaneous estimation of covariance inflation and observation errors within an ensemble Kalman filter. *Q. J. R. Meteorol. Soc.*, **135**, 523–533.

Lin, H., S. Weygandt, S. Benjamin, and M. Hu, 2017: Satellite radiance data assimilation with the hourly updated rapid refresh. *Wea. Forecasting*, **32**, 1273–1287, https://doi.org/10.1175/WAF-D-16-0215.1.

Lin Y. L., R. Farley, and H. Orville, 1983: Bulk parameterization of the snow field in a cloud model. *J. Climate Appl. Meteor.*, **22**, 1065–1092.

Lindskog, M., D. Dee, Y. Trémolet, E. Andersson, G. Radnóti, and M. Fisher, 2009: A weak-constraint four-dimensional variational analysis system in the stratosphere. *Q. J. R. Meteorol. Soc.*, **135**, 695–706.

Lorenc, A. C., 1986: Analysis methods for numerical weather pre- diction. *Q. J. R. Meteorol. Soc.*, **112**, 1177–1194.

Lorenc, A. C., and Coauthors, 2000: The Met Office global three-dimensional variational data assimilation scheme. *Q. J. R. Meteorol. Soc.*, **126**, 2991–3012.

Lorenz, E. N., 1963: Deterministic nonperiodic flow. *J. Atmos. Sci.*, **20**, 130–141.

Lupu C., and A. P. McNally, 2012: Assimilation of cloud-affected radiances from Meteosat-9 at ECMWF. *EUMETSAT/ECMWF Fellowship Programme Research Report No. 25*. European Centre for Medium-range Weather Forecasting (ECMWF), Reading, UK.

Magnusson, L., and E. Kallen, 2013: Factors influencing skill improvements in the ECMWF forecasting system. *Mon. Wea. Rev.*, **141**, 3142–3153.

Mahfouf, J.-F., 2010: Assimilation of satellite-derived soil moisture from ASCAT in a limited-area NWP model. *Q. J. R. Meteorol. Soc.*, **136**, 784–798, DOI:10.1002/qj.602.

Majewski, D., and Coauthors, 2002: The Operational Global Icosahedral-Hexagonal Gridpoint Model GME: Description and High-Resolution Tests. *Mon. Wea. Rev.*, **130**, 319–338.

Mallick, S., D. Dutta, and K.-H. Min, 2017: Quality assessment and forecast sensitivity of global remote sensing observations. *Adv. Atmos. Sci.*, **34**, 371–382.

Marzban, C., 2013: Variance-based sensitivity analysis: An illustration on the Lorenz'63 model. *Mon. Wea. Rev.*, **141**, 4069–4079.

Matricardi M., 2005: The inclusion of aerosols and clouds in RTIASI, the ECMWF Fast Radiative Transfer Model for the Infrared Atmospheric Sounding Interferometer. *Technical Memorandum 474*, European Centre for Medium-range Weather Forecasting (ECMWF), Reading, UK.

McFarquhar, G. M., S. Iacobellis, and R. C. J., Somerville, 2003: SCM simulations of tropical ice clouds using observationally based parameterizations of microphysics. *J. Clim.*, **16**, 1643–1664.

McNally A. P., 2002: A note on the occurrence of cloud in meteorologically sensitive areas and the implications for advanced infrared sounders. *Q. J. R. Meteorol. Soc.*, **128**, 2551–2556.

McNally, A. P., 2009: The direct assimilation of cloud-affected satellite infrared radiances in the ECMWF 4D-Var. *Q. J. R. Meteorol. Soc.*, **135**, 1214–1229.

Meng, Z., and F. Zhang, 2007: Tests of an ensemble Kalman filter for mesoscale and regional-scale data assimilation. Part II: Imperfect model experiments. *Mon. Wea. Rev.*, **135**, 403–1423.

Mengaldo, G., A. Wyszogrodzki, M. Diamantakis, S.-J. Lock, F. X. Giraldo, and N. P. Wedi, 2019: Current and emerging time-integration strategies in global numerical weather and climate prediction. *Arch. Computat. Methods Eng.*, **26**, 663–684.

Michel, Y., 2018: Revisiting Fisher's approach to the handling of horizontal spatial correlations of

observation errors in a variational framework. *Q. J. R. Meteorol. Soc.*, **144**, 2011-2025.

Minamide, M., and F. Zhang, 2017: Adaptive observation error inflation for assimilating all-sky satellite radiance. *Mon. Wea. Rev.*, **145**, 1063–1081, https://doi.org/10.1175/MWR-D-16-0257.1.

Minamide, M., and F. Zhang, 2018: Assimilation of all-sky infrared radiances from Himawari-8 and impacts of moisture and hydrometer initialization on convection-permitting tropical cyclone prediction. *Mon. Wea. Rev.*, **146**, 3241–3258.

Mitchell, H. L., P. L. Houtekamer, and G. Pellerin, 2002: Ensemble size, balance, and model-error representation in an ensemble Kalman filter. *Mon. Wea. Rev.*, **130**, 2791–2808.

Mitchell, H., and A. Carrassi, 2015: Accounting for model error due to unresolved scales within ensemble kalman filtering. *Q. J. R. Meteorol. Soc.*, **141**, 1417–1428.

Miyoshi T., Y. Sato, and T. Kadowaki, 2010: Ensemble Kalman filter and 4D-Var intercomparison with the Japanese operational global analysis and prediction system. *Mon. Wea. Rev.*, **138**, 2846–2866, doi:10.1175/2010MWR3209.1.

Miyoshi, T., 2011: The Gaussian approach to adaptive covariance inflation and its implementation with the local ensemble transform kalman filter. *Mon. Wea. Rev.*, **139**, 1519–1535.

Miyoshi, T., E. Kalnay, and H. Li, 2013: Estimating and including observation-error correlations in data assimilation. *Inverse Probl. Sci. Eng.*, **21**, 387-398.

Nakamura, G., and R. Potthast, 2015: Inverse Modeling: An introduction to the theory and methods of inverse problems and data assimilation. IOP Publishing, doi:10.1088/978-0-7503-1218-9.

Ngodock, H., M. Carrier, S. Smith, and I. Souopgui, 2017: Weak and strong constraints variational data assimilation with the NCOM-4dvar in the Agulhas region using the representer method. *Mon. Wea. Rev.*, **145**, 1755–1764.

Nicolis, C., 2003: Dynamics of model error: Some generic features. *J. Atmos. Sci.*, **60**, 2208–2218.

Nicolis, C., 2004: Dynamics of model error: The role of unresolved scales revisited. *J. Atmos. Sci.*, **61**, 1740–1753.

Nicolis, C., R. A. P. Perdigao, and S. Vannitsem, 2009: Dynamics of prediction errors under the

combined effect of initial condition and model errors. *J. Atmos. Sci.*, **66**, 766–778.

Okamoto K., 2013: Assimilation of overcast cloudy infrared radiances of the geostationary MTSAT-1R imager. *Q. J. R. Meteorol. Soc.*, **139**, 715–730.

Okamoto, K., A. P. McNally, and W. Bell, 2014: Progress towards the assimilation of all-sky infrared radiances: An evaluation of cloud effects. *Q. J. R. Meteorol. Soc.*, **140**, 1603–1614, doi:10.1002/qj.2242.

Okamoto, K., 2017: Evaluation of IR radiance simulation for all-sky assimilation of Himawari-8/AHI in a mesoscale NWP system. *Q. J. R. Meteorol. Soc.*, **143**, 1517–1527.

Okamoto, K. Y. Sawada, and M. Kunii, 2019: Comparison of assimilating all-sky and clear-sky infrared radiances from Himawari-8 in a mesoscale system. *Q. J. R. Meteorol. Soc.*, **145**, 745–766.

Otkin, J. A., and T. J. Greenwald, 2008: Comparison of WRF model-simulated and MODIS-derived cloud data. *Mon. Wea. Rev.*, **136**, 1957–1970.

Otkin, J. A., T. J. Greenwald, J. Sieglaff, and H.-L. Huang, 2009: Validation of a large-scale simulated brightness temperature dataset using SEVIRI satellite observations. *J. Appl. Meteorol. Climatol.*, **48**, 1613–1626, doi:10.1175/2009JAMC2142.1.

Otkin, J. A., 2010: Clear and cloudy sky infrared brightness temperature assimilation using an ensemble Kalman filter. *J. Geophys. Res.*, **115, D19207**, doi:10.1029/2009JD013759.

Otkin, J. A., 2012a: Assessing the impact of the covariance localization radius when assimilating infrared brightness temperature observations using an ensemble Kalman filter. *Mon. Wea. Rev.*, **140**, 543–561, doi:10.1175/MWR-D-11-00084.1.

Otkin, J. A., 2012b: Assimilation of water vapor sensitive infrared brightness temperature observations during a high impact weather event. *J. Geophys. Res.*, **117, D19203**, doi:10.1029/2012JD017568.

Otkin, J. A., R. Potthast, and A. Lawless, 2018: Nonlinear bias correction for satellite data assimilation using Taylor series polynomials. *Mon. Wea. Rev.*, **146**, 263–285.

Otkin, J. A., and R. Potthast, 2019: Assimilation of all-sky SEVIRI infrared brightness temperatures with nonlinear bias corrections in a regional-scale ensemble data assimilation system. *Mon. Wea. Rev.*, **147**, 4481-4509.

Otkin, J. A., R. E. Potthast, and A. S. Lawless 2020: Nonlinear conditional bias estimation for data assimilation. Conditionally accepted for publication in *SIAM Journal on Applied Dynamical Systems*.

Ou, S., and K.-N. Liou, 1995: Ice microphysics and climatic temperature feedback. *Atmos. Res.*, **35**, 127–138.

Pan, S. J., J. D. Gao, D. J. Stensrud, X. G. Wang, and T. A. Jones, 2018: Assimilation of radar radial velocity and reflectivity, satellite cloud water path, and total precipitable water for convective-scale NWP in OSSEs. *J. Atmos. Oceanic Technol.*, **35**, 67–89.

Pangaud, T., N. Fourrie, V. Guidard, M. Dahoui, F. and Fabier, 2009: Assimilation of AIRS radiances affected by mid to low level clouds. *Mon. Wea. Rev.*, **137**, 4276–4292. https://doi.org/10.1175/2009MWR3020.1.

Parrish D. F., and J. C. Derber, 1992: The National Meteorological Center's spectral statistical interpolation analysis system. *Mon. Wea. Rev.*, **120**, 1747–1763.

Pavelin, E.G., S. J. English, and J. R. Eyre, 2008: The assimilation of cloud-affected infrared satellite radiances for numerical weather prediction. *Q. J. R. Meteorol. Soc.*, **13**, 737–749.

Perianez A., H. Reich, and R. Potthast, 2014: Optimal localization for ensemble Kalman filter systems. *J. Meteorol. Soc. Japan*, **92**, 585–597. https://doi.org/10.2151/jmsj.2014-605.

Peubey, C. and A. P. McNally, 2009: Characterization of the impact of geostationary clear-sky radiances on wind analyses in a 4D-Var context. *Q. J. R. Meteorol. Soc.*, **135**, 1863–1876.

Pu, X. and J. Hacker, 2009: Ensemble-based Kalman filters in strongly nonlinear dynamics. *Adv. Atmos. Sci.*, **26**, 373–380.

Raisanen, P., 1998: Effective longwave cloud fraction and maximum-random overlap of clouds: A problem and a solution. *Mon. Wea. Rev.*, **126**, 3336–3340.

Raschendorfer M., 2001: The new turbulence parameterisation of LM. COSMO Newsl. 1, 89–97.

Reich S., and C. Cotter, 2015: Probabilistic Forecasting and Bayesian Data Assimilation. Cambridge University Press.

Raynaud, L., L. Berre, and G. Desroziers, 2011: An extended specification of flow-dependent background error variances in the Meteo-France global 4D-Var system. *Q. J. R. Meteorol. Soc.*, **137**,

607-619.

Ritter B., and J. F. Geleyn, 1992: A comprehensive radiation scheme for numerical weather prediction models with potential applications in climate simulations. *Mon. Wea. Rev.*, **120**, 303–325

Romine, G. S., and Coauthors, 2014: Representing forecast error in a convection permitting ensemble system. *Mon. Wea. Rev.*, **142**, 4519–4541.

Sakov, P., D. S. Oliver, and L. Bertino, 2012: An iterative EnKF for strongly nonlinear systems. *Mon. Wea. Rev.*, **140**, 1988–2004.

Saltzman, B., 1962: Finite amplitude free convection as an initial value problem. *J. Atmos. Sci.*, **19**, 329–341.

Sasaki, Y., 1970: Some basic formalisms in numerical variational analysis. *Mon. Wea. Rev.*, **98**, 875–883.

Saunders R., M. Matricardi, and P. Brunel, 1999: An improved fast radiative transfer model for assimilation of satellite radiance observations. *Q. J. R. Meteorol. Soc.*, **125**, 1407–1425.

Schmetz, J., P. Pili, S. Tjemkes, D. Just, J. Lerkmann, S. Rota, and A. Ratier, 2002: An introduction to Meteosat Second Generation (MSG). *Bull. Am. Meteor. Soc.*, **83**, 977–992.

Schmit, T. J., P. Griffith, M. M. Gunshor, J. M. Daniels, S. J. Goodman, and W. J. Lebair, 2017: A closer look at the ABI on the GOES-R series. *Bull. Am. Meteor. Soc.*, **98**, 681–698, https://doi.org/10.1175/BAMS-D-15-00230.1.

Schomburg A., C. Schraff, and R. Potthast, 2015: A concept for the assimilation of satellite cloud information in an Ensemble Kalman Filter: Single-observation experiments. *Q. J. R. Meteorol. Soc.*, **141**, 893–908, doi: 10.1002/qj.2407.

Schraff, C., H. Reich, A. Rhodin, A. Schomburg, K. Stephan, A. Perianez, and R. Potthast, 2016: Kilometer-Scale ensemble data assimilation for the COSMO model (KENDA). *Q. J. R. Meteorol. Soc.*, **142**, 1453–1472.

Schwartz, C. S., G. S. Romine, R. A. Sobash, K. R. Fossell, and M. L. Weisman, 2015: NCAR's experimental real-time convection-allowing ensemble prediction system. *Wea. Forecasting*, **30**, 1645–1654.

Seifert A, and K. Beheng, 2001: A double-moment parameterization for simulating auto-

conversion, accretion and selfcollection. *Atmos. Res.*, **59-60**, 265–281, doi:10.1016/S0169-8095(01)00126-0.

Shutts, G., 2015: A stochastic convective backscatter scheme for use in ensemble prediction systems. *Q. J. R. Meteorol. Soc.*, **141**, 2602– 2616.

Snyder, C., and F. Zhang, 2003: Assimilation of simulated doppler radar observations with an ensemble Kalman filter. *Mon. Wea. Rev.*, **131**, 1663–1677.

Sparrow, C., 1982: The Lorenz Equations: Bifurcations, Chaos, and Strange Attractors. Applied Mathematical Sciences, Vol. 41, Springer-Verlag, NY.

Stengel, M., P. Unden, M. Lindskog, P. Dahlgren, N. Gustafsson, and R. Bennartz, 2009: Assimilation of SEVIRI infrared radiances with HIRLAM 4D-Var. Quart. *Q. J. R. Meteorol. Soc.*, **135**, 2100–2109.

Stengel, M., M. Lindskog, P. Unden, and N. Gustafsson, 2013: The impact of cloud-affected IR radiances on forecast accuracy of a limited-area NWP model. *Q. J. R. Meteorol. Soc.*, **139**, 2081–2096.

Stensrud, D. J., 2007: Parameterization schemes: Keys to understanding numerical weather prediction models. Cambridge University Press. 480pp.

Stewart, L. M., S. L. Dancer, N. K. Nichols, J. R. Eyre, and J. Cameron, 2014: Estimating interchannel observation-error correlations for IASI radiance data in the Met Office system. *Q. J. R. Meteorol. Soc.*, **140**, 1236-1244.

Strow, L. L., and Coauthors, 2013: Spectral calibration and validation of the Cross-track Infrared Sounder on the Suomi NPP satellite. *J. Geophys. Res. Atmos.*, **118**, 12,486–12,496, https://doi.org/10.1002/2013JD020480.

Szunyogh, I., E. J. Kostelich, G. Gyarmati, E. Kalnay, and B. R. Hunt, 2008: A local ensemble transform Kalman filter data assimilation system for the NCEP global model. *Tellus*, **60A**, 113–130.

Tavolato, C., and L. Isaksen, 2015: On the use of a Huber norm for observation quality control in the ECMWF 4D-Var. *Q. J. R. Meteorol. Soc.*, **141**, 1514–1527.

Thompson, G., M. Tewari, K. Ikeda, S. Tessendorf, C. Weeks, J. A. Otkin, and F. Kong, 2016: Explicitly-coupled cloud physics and radiation parameterizations and subsequent evaluation

in WRF high-resolution convective forecasts. *Atmos. Res.*, **168**, 92–104.

Tiedtke M., 1989: A comprehensive mass flux scheme for cumulus parameterisation in large-scale models. *Mon. Wea. Rev.*, **117**, 1779–1799.

Tikhonov, A. N., 1963: Solution of incorrectly formulated problems and the regularization method. *Soviet Mathematics—Doklady*, **4**, 1035–1038.

Tremolet, Y., 2006: Accounting for an imperfect model in 4d-var. *Q. J. R. Meteorol. Soc.*, **132**, 2483–2504.

Trémolet, T., 2007: Model-error estimation in 4d-var. *Q. J. R. Meteorol. Soc.*, **133**, 1267–1280.

Verlaan M., and A. W. Heemink, 2001: Nonlinearity in data assimilation applications: A practical method for analysis. *Mon. Wea. Rev.*, **129**, 1578–1589.

Vidard, P. A., A. Piacentini, and F.-X. le Dimet, 2004: Variational data analysis with control of the forecast bias. *Tellus A*, **56**, 177–188.

Vukicevic, T., T. Greenwald, M. Zupanski, D. Zupanski, T. Vonder Haar, and A. S. Jones, 2004: Mesoscale cloud state estimation from visible and infrared satellite radiances. *Mon. Wea. Rev.*, **132**, 3066–3077, doi:10.1175/MWR2837.1.

Vukicevic, T., M. Sengupta, A. S. Jones, and T. Vonder Haar, 2006: Cloud-resolving satellite data assimilation: Information content of IR window observations and uncertainties in estimation. *J. Atmos. Sci.*, **63**, 901–919.

Vukicevic T., and D. Posselt, 2008: Analysis of the impact of model nonlinearities in inverse problem solving. *J. Atmos. Sci.*, **65**, 2803–2823.

Waller, J. A., D. Simonin, S. L. Dance, N. K. Nichols, and S. P. Ballard, 2016a: Diagnosing observation-error correlations for Doppler radar radial winds in the Met Office UKV model using observation-minus-background and observation-minus-analysis statistics. *Mon. Wea. Rev.*, **144**, 3533-3551.

Waller, J. A., S. P. Ballard, S. L. Dancer, G. Kelly, N. K. Nichols, and D. Simonin, 2016b: Diagnosing horizontal and interchannel observation-error correlations for SEVIRI observations using observation-minus-background and observation-minus-analysis statistics. *Rem. Sensing*, **8, 581**, https://doi.org/10.3390/rs8070581.

Wang, P., J. Li, B. Lu, T. J. Schmit, J. Lu, Y.-K. Lee, J. Li, and Z. Liu, 2018: Impact of moisture information from Advanced Himawari Imager measurements on heavy precipitation forecasts in a regional NWP model. *J. Geophys. Res. Atmos.*, **123**, 6022–6038. https://doi.org/10.1029/2017JD028012.

Weisheimer, A., S. Corti, T. N. Palmer, and F. Vitart, 2014: Addressing model error through atmospheric stochastic physical parametrizations: Impact on the coupled ECMWF seasonal forecasting system. *Philos. Trans. Roy. Soc. London*, **A372**, 2013029, doi:https://doi.org/10.1098/rsta.2013.0290.

Whitaker, J. S., and T. M. Hamill, 2002: Ensemble data assimilation without perturbed observations. *Mon. Wea. Rev.*, **130**, 1913–1924.

Whitaker, J. S., and T. M. Hamill, 2012: Evaluating methods to account for system errors in ensemble data assimilation. *Mon. Wea. Rev.*, **140**, 3078–3089.

Wu, T.-C., M. Zupanski, L. D. Grasso, C. D. Kummerow, and S.-A. Boukabara, 2019: All-sky radiance assimilation of ATMS in HWRF: A demonstration study. *Mon. Wea. Rev.*, **147**, 85–106.

Wyser, K. 1998: The effective radius in ice clouds. *J. Climate*, **7**, 1793–1802.

Yang, C., Z. Liu, J. Bresch, S. R. Rizvi, X. Y. Huang, and J. Min, 2016: AMSR2 all-sky radiance assimilation and its impact on the analysis and forecast of hurricane Sandy with a limited-area data assimilation system. *Tellus A*, **68**, 30917, http://dx.doi.org/10.3402/tellusa.v68.30917.

Yang, P., L. Bi, B. A. Baum, K.-N. Liou, G. Kattawar, M. Mishchenko, and B. Cole, 2013: Spectrally consistent scattering, absorption, and polarization properties of atmospheric ice crystals at wavelengths from 0.2 μm to 100 μm. *J. Atmos. Sci.*, **70**, 330–347.

Yang, S.-C., E. Kalnay, and B. R. Hunt, 2012 Handling nonlinearity in an ensemble Kalman filter: Experiments with the three-variable Lorenz model. *Mon. Wea. Rev.*, **140**, 2628–2646.

Yi, B., P. Yang, Q. Liu, P. van Delst, S.-A. Boukabara, and F. Weng, 2016: Improvements on the ice cloud modeling capabilities of the Community Radiative Transfer Model. *J. Geophys. Res. Atmos.*, **121**, 13577–13590, doi:10.1002/2016JD025207.

Zangl G., D. Reinert, P. Ripodas, and M. Baldauf, 2015: The ICON (ICOsahedral Non-hydrostatic) modelling framework of DWD and MPI-M: Description of the non-hydrostatic dynamical core. *Q. J. R. Meteorol. Soc.*, **141**, 563–579, doi:10.1002/qj.2378.

Zhang F., C. Snyder, and J. Sun, 2004: Impacts of initial estimate and observation availability on convective-scale data assimilation with an ensemble Kalman filter. *Mon. Wea. Rev.*, **132**, 1238–1253.

Zhang, F., M. Minamide, and E. Clothiaux, 2016: Potential impacts of assimilating all-sky infrared satellite radiances from GOES-R on convection-permitting analysis and prediction of tropical cyclones. *Geophys. Res. Lett.*, **43**, 2954–2963, https://doi.org/10.1002/2016GL068468.

Zhang, Y. J., F. Zhang, and D. J. Stensrud, 2018: Assimilating All-Sky Infrared Radiances from GOES-16 ABI Using an Ensemble Kalman Filter for Convection-Allowing Severe Thunderstorms Prediction. *Mon. Wea. Rev.*, **146**, 3363–3381.

Zhang, M., F. Zhang, X.-Y. Huang, and X. Zhang, 2011: Intercomparison of an ensemble Kalman filter with three- and four-dimensional variational data assimilation methods in a limited-area model over the month of June 2003. *Mon. Wea. Rev.*, **139**, 566–572, https://doi.org/10.1175/2010MWR3610.1.

Zhang, S., Z. Liu, A. Rosati, and T. Delworth, 2012: A study of enhancive parameter correction with coupled data assimilation for climate estimation and prediction using a simple coupled model. *Tellus A*, **64**, 10963.

Zhang, S., and L. Guan, 2017: Preliminary study on direct assimilation of cloud-affected satellite brightness temperatures. *Adv. Atmos. Sci.*, **34**, 199–208.

Zhu Y., J. Derber, A. Collard, D. Dee, R. Treadon, G. Gayno, and J. A. Jung, 2014: Enhanced radiance bias correction in the National Centers for Environmental Prediction's Gridpoint Statistical Interpolation data assimilation system. *Q. J. R. Meteorol. Soc.*, **140**, 1479–1492, doi:10.1002/qj.2233.

Zhu, Y., J. C. Derber, R. J. Purser, B. A. Ballish, and J. Whiting, 2015: Variational correction of aircraft temperature bias in the NCEP's GSI analysis system. *Mon. Wea. Rev.,* **143**, 3774–3803.

Zhu, Y., and Coauthors, 2016: All-sky microwave radiance assimilation in NCEP's GSI analysis system. *Mon. Wea. Rev.*, **144**, 4709–4735.

Zupanski, D., 1997: A general weak constraint applicable to operational 4dvar data assimilation systems. *Mon. Wea. Rev.*, **125**, 2274–2292.

Zupanski, D., M. Zupanski, L. D. Grasso, R. Brummer, I. Jankov, D. Lindsey, M. Sengupta, and

M. Demaria, 2011: Assimilating synthetic GOES-R radiances in cloudy conditions using an ensemble-based method. *Int. J. Remote Sens.*, **32**, 9637–9659.