Ensemble Kalman Filter Methods for High-Resolution Data Assimilation

February 2014

Keiichi KONDO

Ensemble Kalman Filter Methods

for High-Resolution Data Assimilation

A Dissertation Submitted to
the Graduate School of Life and Environmental Sciences,
the University of Tsukuba
in Partial Fulfillment of the Requirements
for the Degree of Doctor of Philosophy in Science
(Doctoral Program in Geoenvironmental Sciences)

Keiichi KONDO

ABSTRACT

In this study, a new method of 4 dimensional data assimilation technique has been developed in order to apply the ensemble Kalman filter (EnKF) to high-resolution models. First, to investigate the basic performance of EnKF, we have confirmed whether the EnKF converges to the extended Kalman filter (EKF) by increasing the ensemble members, using a barotropic general circulation model, called barotropic S-model, under the perfect model scenario in the spectral space. The barotropic S-model is based on the primitive equations and predicts the vertical mean state of the atmosphere. Although it has the predictability comparable to the operational prediction models, the direct computation of the EKF is possible. Therefore, we can assess the accuracy of the EnKF as a function of the ensemble members.

According to the result of the root mean square error (RMSE), the EnKF converges to the full-rank EKF when the ensemble member is increased to 100. It is demonstrated that the 20 ensemble members are insufficient with respect to the convergence. An empirical orthogonal function (EOF) analysis is conducted using the analysis error covariance matrices for both filters. The structure of the first EOF (EOF-1) indicates the characteristics of the baroclinic instability waves in mid-latitudes in both filters, showing the same geographical distributions when it has converged. It is concluded from the comparison of the RMSE that about 100 ensemble members are required for the non-localized EnKF to converge to the full-rank EKF for the practical assimilation in the spectral space under the perfect model scenario

of the barotropic general circulation model of the atmosphere.

Next, based on the confirmed convergence of the EnKF, a local ensemble transform Kalman filter (LETKF) is developed for the Nonhydrostatic Icosahedral Atmospheric Model (NICAM). We refer to this new assimilation system as the NICAM-LETKF in this study. In addition, an algorithm to adaptively estimate the inflation parameter and the observational errors is introduced to the LETKF. The feasibility and stability of the NICAM-LETKF are investigated under the perfect model scenario.

According to the results, it is confirmed that the converged analysis errors of the NICAM-LETKF are smaller than the observational errors, and the magnitude and distribution of the RMSEs are comparable to those of the ensemble spreads. In our experiments, it is confirmed that the inflation parameter is optimally tuned and the observational errors are close to the truth. It is concluded that the NICAM-LETKF works appropriately and stably under the perfect model scenario even if the inflation parameter and the observational errors are adaptively estimated within the LETKF.

Finally, a new method of multi-scale localization technique is developed in order to apply the LETKF to the high-resolution model. In a very high-resolution nonhydrostatic general circulation model such as NICAM, the excessive localization may limit the region of the available observations within a narrower area around the analysis point for a limited ensemble size. In this situation, the larger-scale structures than the localization would be excluded from the assimilation. To retain the larger-scale structures, a dual-localization method has been proposed, that consid-

ers two separated localization scales simultaneously. The method is examined with a low-resolution atmospheric model called SPEEDY. The dual-localization method analyzes small-scale analysis increment and large-scale analysis increment separately using spatial smoothing. The small and large localization scales and the smoothing function are the control parameters of the dual-localization method. In this study, we have investigated the parameter sensitivities in observation system simulation experiments. Also, the sensitivity to the two localization scales is carefully investigated. The results show that the dual-localization method outperforms the traditional single localization for relatively wide 400 km ranges of the two localization scales. This suggests that we can avoid fine tuning of the two localization scales. This new method can be easily extended to triple- or multi-scale localization methods.

In this study, we have confirmed that the EnKF converges to the EKF with low-resolution atmospheric model. We have concluded that the application of the multi-scale localization methods to the EnKF may construct the scientific basis of the future data assimilation for the very high-resolution models.

Key words: Data assimilation, EKF, EnKF, LETKF, Barotropic S-model, NICAM, dual-localization approach

CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION

Data assimilation generates accurate initial conditions by extracting the most information from both model forecasts and observations. Numerical weather prediction (NWP) is generally sensitive to the initial conditions, and it is essential to obtain better initial conditions using more advanced data assimilation methods. Many operational centers use 3D-Var (3-dimensional variational data assimilation; Parrish and Derber 1992), which is an economical and accurate statistical interpolation scheme, but 3D-Var cannot treat flow-dependent information. In some operational centers, 4D-Var (4-dimensional variational data assimilation; e.g., H. Liu and X. Zou 2001, and Pierre and Jean-Noël 2001) has been used recently, and the accuracy is much higher than the 3D-Var. The 4D-Var is one of advanced data assimilation methods which can the treat flow-dependent information. In addition to its high accuracy, the 4D-Var allows the assimilation of many observations, for example, asynchronous observations such as satellite radiances at their correct observation time.

Kalman filter (Kalman 1960, hereafter KF) is the optimal filter for linear systems with Gaussian error distribution. Extended Kalman filter may be proposed by Jazwinski (1970) to apply the KF to nonlinear systems. However, it is known that the KF may not be always stable in the case of being applied to nonlinear systems because the KF is based on the linear theory (e.g., Pham 2001). In the case of applying

the EKF to the general circulation models (GCMs), we need to consider their large degrees of freedom, and their orders are at least $O(10^6)$. So, we cannot implement the EKF with the GCMs because it is not feasible to compute a time evolution of error covariance matrix.

Then, Evensen (1994) approximated the covariance matrix of the KF using ensemble predictions, which was called an ensemble Kalman filter (EnKF). The EnKF is also one of advanced data assimilation methods. There are a number of studies about the comparison of the 3D-Var, 4D-Var or EnKF with various numerical models. For example, Whitaker et al. (2008) showed that the EnKF outperformed 3D-Var with the NCEP global forecast system, Fertig et al. (2007) compared LETKF and 4D-Var with the Lorenz-96 system (Lorenz 1996), and Kalnay et al. (2007) discussed the performances of EnKF and 4D-Var. On the other hand, the comparison of the EKF and the EnKF is hardly studied with the atmospheric models because it is not feasible to implement the EKF. Then Zang and Malanotte-Rizzoli (2003) compared the reduced-rank EKF and EnKF with the ocean model. They confirmed that for the strongly nonlinear case, the EnKF outperforms the reduced-rank EKF, and for the weak nonlinear case the performances of the reduced-rank EKF and EnKF are similar. Their reduced-rank EKF assumes a steady state filter. So the reduced-rank EKF cannot treat the rapidly changing atmospheric states, and it is necessary to compare EKF and EnKF with an atmospheric model.

Although the EnKF has not been hardly compared with the EKF yet in the atmospheric models, a number of research articles on EnKF have been published so

far. Various EnKF approaches have been suggested, such as the Ensemble Adjustment Kalman Filter (EAKF, Anderson 2001), Serial Ensemble Square Root Filter (Whitaker and Hamill 2002), Ensemble Transform Kalman Filter (ETKF; Bishop et al. 2001), and Local Ensemble Kalman Filter (LEKF; Ott et al. 2004). Hunt et al. (2007) suggested the Local Ensemble Transform Kalman Filter (LETKF) by applying the ETKF algorithm to the LEKF, and the LETKF has an important advantage in assimilating the observations in each "local patch", where the local patch is defined as an area around the analysis grid point and is also made up by the surrounding grid points. Due to this advantage, the LETKF has a higher performance for the implementation in parallel computers. Miyoshi and Yamane (2007) applied the LETKF to the Earth Simulator global model known as the AFES (Atmospheric General Circulation Model for the Earth Simulator; Ohfuchi et al. 2004) and investigated its performance in detail. Miyoshi et al. (2007) removed the local patch applied the new LETKF to AFES with a T159L48 resolution, and they investigated the stability of the LETKF without the local patches. Kalnay and Yang (2010) proposed and tested the RIP (Running-In-Place) approach with the LETKF to accelerate the spin-up time. And the EnKF are also applied to regional models. Zhang et al. (2006) investigated the EnKF with a nonhydrostatic regional model and gave details on the dependence of EnKF performance on error growth rate and scales. Seko et al. (2011) investigated an intense rainfall event by applying the LETKF to the JMANHM (Saito et al. 2006). Miyoshi and Kunii (2012) applied the LETKF to the WRF (Weather Research and Forecasting; Skamarock et al. 2005) model.

In high-resolution data assimilation, however, localization (Houtekamer and Mitchell 1998; Hamill et al. 2001) is a problem. When we apply EnKF to realistic meteorological problems, localization plays an essential role in dealing with spurious sampling errors due to a generally limited ensemble size. EnKF is generally unstable without localization mainly because of sampling errors between distant locations. Localization limits the impact of observations within a certain distance defined by a prescribed localization function. The localization function is usually a distance-dependent function that drops to zero at a certain distance, and the Gaspari and Cohn (1999)'s fifth-order polynomial function is a typical choice. The optimal localization scale depends on several factors including the model resolution and ensemble size; the localization scale becomes generally smaller for higher-resolution models. For example, Miyoshi et al. (2010) applied the 400-km localization scale for a T319/L40 (horizontal resolution up to 319 wavenumbers and 40 vertical levels) global model, and Miyoshi et al. (2007) applied the 500-km localization scale for the T159/L48 AFES. By contrast, for the T30/L7 SPEEDY model (Molteni 2003) the optimal localization scale is found to be 700 km (Kang, personal communication). In mesoscale model, Zhang et al. (2009) applied about 340-km, 110-km and 40-km localization scale for the WRF model whose corresponding horizontal resolutions are 40.5 km, 13.5 km and 4.5 km, respectively. Yussouf and Stensrud (2010) applied about a 1-km localization scale (the localization radius is 4 km) with 1-km horizontal resolution model for radar data assimilation. In such cases, observations are used only in such a limited distance, although at least some observations such as rawinsondes are expected to observe synoptic-scale weather with at least an O(100)-km scale.

In this study, in view of future data assimilation in NWP, EnKF methods are investigated. First, to investigate the performances of EnKF, the EKF and EnKF are compared. The comparison of the EKF and the EnKF has been hardly studied with the atmospheric models. In this study, Kondo (2007) and Kondo and Tanaka (2009a), the EnKF is compared with the EKF under a perfect model scenario using a barotropic S-model (Tanaka 2003), which is based on the primitive equations and has only 410 degrees of freedom. So, it is possible to implement the EKF at full rank. The full-rank EKF can treat the rapidly changing atmospheric states completely. We investigate the accuracy of the EnKF in comparison to the EKF by the root mean square error (RMSE) of both filters. In addition to comparing the analysis accuracy, we investigate how much the eigenvalues and the eigenvectors of the analysis covariance matrices coincide with each other for EnKF and EKF.

Next, the LETKF is applied to the NICAM (Nonhydrostatic Icosahedral Atmospheric Model; Satoh et al. 2008) and the NICAM-LEKTF are constructed for the future cloud-resolving data assimilation. The NICAM is a new type of ultra-high resolution atmospheric general circulation model, and is designed to perform cloud-resolving simulations by directly calculating deep convection and meso-scale circulation, which plays a key role not only in the tropical circulations but also in the entire general circulation of the atmosphere. As the model adopts the nonhydrostatic equations and icosahedral grid structure. However, the assimilation system for the NICAM has not been developed and the optimum initial condition for the NICAM does not yet exist. Kondo (2009) applied the non-local patch version of the LETKF

to the NICAM (referred to as NICAM-LETKF), but did not investigate the adaptive estimation of the inflation parameter and observational errors (Miyoshi 2005, Li et al. 2009a). Therefore, this study and Kondo and Tanaka (2009b) investigate the feasibility and stability of the NICAM-LETKF with adaptive estimation of both inflation parameter and observational errors under the perfect model scenario. This is the first test of the LETKF with a global non-hydrostatic model, although the model should still behave hydrostatically with the horizontal resolution of 224 km.

Finally, a new method of multi-scale localization is developed and is investigated for the high-resolution data assimilation in order to tackle the localization problem. If we have a very high-resolution mesoscale model with a 1-km or even finer grid spacing, the localization scale would likely be limited up to O(10)-km. Therefore, if we apply narrow localization up to O(10)-km for high resolution models, EnKF may not account for larger-scale errors even though the observations shall provide useful information at an O(100)-km scale. Multi-scale consideration is essential to tackle these situations and to use precious observations more effectively. In Zhang et al. (2009), three different localization scales are applied to three different resolutions models in a two-way nesting system to explore the uses of Doppler radar observations for cloud-resolving hurricane analysis, initialization and prediction. Buehner (2012) proposed a spatial and spectral covariance localization approach, which considers separate scales simultaneously. In this way, we can account for larger-scale error covariances separately from smaller-scale error covariances. Larger-scale error covariances are expected to have larger-scale structure, which usually requires large-

scale localization. This enables longer-distance impact from observations even if the model resolution is high. The separate consideration of higher-resolution error covariances allows simultaneous analysis of fine structure near the observations. Inspired by Buehner (2012), Miyoshi and Kondo (2013) and Kondo et al. (2013) proposed the dual-localization approach that considers two localization scales simultaneously. In the dual-localization method, longer-range analysis increments are analyzed by applying a spatial smoothing to the ensemble perturbations and using a larger localization scale, while simultaneously, shorter-range analysis increments are analyzed by using a smaller localization scale. Here, the smoothing function and two localization scales are the control parameters of the dual-localization approach.

To make this paper self-contained, in chapter II the EKF and EnKF are described. Chapter III, IV and V show the comparison of the EKF and EnKF, the NICAM-LETKF and the multi-localization approach. Finally, concluding summary is provided in chapter VI.

CHAPTER II

Kalman Filter and Ensemble Kalman Filter

2.1 Kalman Filter

Kalman (1960) suggested the KF algorithm which is optimal when evolution of error is linear and the probability distribution of error is the normal distribution. The EKF is suggested to the KF apply to nonlinear systems (Jazwinski 1970). By applying a linear model M to a state variable $\mathbf{x}_{t-1}^a$, a forecast at time t is written by

$$\mathbf{x}_t^f = \mathbf{M}\mathbf{x}_{t-1}^a, \tag{2.1}$$

where $\mathbf{x}_t$ denotes an n-dimensional state variable at time t. The superscripts f, a represent forecast and analysis, respectively. In the same way, the true evolution at t is written as

$$\mathbf{x}_t^t = \mathbf{M}^t\mathbf{x}_{t-1}^t, \tag{2.2}$$

where the superscript t represents the truth. However, since the model M is not perfect, so $\mathbf{x}_t^t$ is written as

$$\mathbf{x}_t^t = \mathbf{M}\mathbf{x}_{t-1}^t - \mathbf{q}, \tag{2.3}$$

where $\mathbf{q}$ denotes the model error, and $\langle \mathbf{q} \rangle = 0$ if the model is unbiased. $\langle\rangle$ represents an average over many cases or statistical expectation. The analysis/forecast errors $\delta \mathbf{x}^a$, $\delta \mathbf{x}^f$ are defined by

$$\delta \mathbf{x}^{a,f} = \mathbf{x}^{a,f} - \mathbf{x}^t. \tag{2.4}$$

By using Eqs. (2.3) and (2.4), at time t a forecast error covariance, i.e., background error covariance $\mathbf{P}_t^f$ is obtained by

$$\begin{aligned}
\mathbf{P}_t^f &= \left\langle \delta \mathbf{x}_t^f \left(\delta \mathbf{x}_t^f \right)^\top \right\rangle, \\
&= \left\langle \left(\mathbf{x}_t^f - \mathbf{x}_t^t \right) \left(\mathbf{x}_t^f - \mathbf{x}_t^t \right)^\top \right\rangle, \\
&= \left\langle \left(\mathbf{M}\mathbf{x}_{t-1}^a - \left(\mathbf{M}\mathbf{x}_{t-1}^t - \mathbf{q} \right) \right) \left(\mathbf{M}\mathbf{x}_{t-1}^a - \left(\mathbf{M}\mathbf{x}_{t-1}^t - \mathbf{q} \right) \right)^\top \right\rangle, \\
&= \left\langle \left(\mathbf{M} \left(\mathbf{x}_{t-1}^a - \mathbf{x}_{t-1}^t \right) + \mathbf{q} \right) \right) \left(\mathbf{M} \left(\mathbf{x}_{t-1}^a - \mathbf{x}_{t-1}^t \right) + \mathbf{q} \right) \right)^\top \Big\rangle, \\
&= \left\langle \left(\mathbf{M}\delta \mathbf{x}_{t-1}^a + \mathbf{q} \right) \left(\mathbf{M}\delta \mathbf{x}_{t-1}^a + \mathbf{q} \right)^\top \right\rangle, \\
&= \left\langle \left(\mathbf{M}\delta \mathbf{x}_{t-1}^a \right) \left(\mathbf{M}\delta \mathbf{x}_{t-1}^a \right)^\top + \left(\mathbf{M}\delta \mathbf{x}_{t-1}^a \right) \mathbf{q}^\top + \mathbf{q} \left(\mathbf{M}\delta \mathbf{x}_{t-1}^a \right)^\top + \mathbf{q}\mathbf{q}^\top \right\rangle, \\
&= \left\langle \mathbf{M}\delta \mathbf{x}_{t-1}^a \left(\delta \mathbf{x}_{t-1}^a \right)^\top \mathbf{M}^\top + \mathbf{q}\mathbf{q}^\top \right\rangle, \\
&= \mathbf{M}\mathbf{P}_{t-1}^a \mathbf{M}^\top + \mathbf{Q},
\end{aligned} \tag{2.5}$$

where $\mathbf{P}_{t-1}^a = \left\langle \delta \mathbf{x}_{t-1}^a \left(\delta \mathbf{x}_{t-1}^a \right)^\top \right\rangle$ represents an analysis error covariance at $t-1$. Some cross terms such as $\left\langle \left(\mathbf{M}\delta \mathbf{x}_{t-1}^a \right) \mathbf{q}^\top \right\rangle$ are zeroed out because we assume that there is no correlation between the analysis error $\delta \mathbf{x}_{t-1}^a$ and the model error $\mathbf{q}$.

In the EKF, a nonlinear model M is expressed using a linear tangent model $\mathbf{M}$ around $\mathbf{x}_t$.

$$\mathbf{M} = \left. \frac{\partial M}{\partial \mathbf{x}} \right|_{\mathbf{x}_t}. \tag{2.6}$$

Moreover, Eq. (2.1) is replaced by

$$\mathbf{x}_t^f = M(\mathbf{x}_{t-1}^a). \tag{2.7}$$

In the nonlinear scenario the dynamical system is integrated by the forecast step from time $t-1$ to t followed by an analysis step at time t to estimate the most likely system state. In the forecast step the EKF propagates the model solution from time $t-1$ to t, and in the analysis step the EKF combines the information provided by the observations at that time with the propagated information from the prior observations. In the EKF, Eqs. (2.7) and (2.5) describe how the model propagates the state variable and the covariance to the future.

In the analysis step, an analysis $\mathbf{x}^a$ is a weighted average of forecast $\mathbf{x}^f$ and observation $\mathbf{y}^o$. The analysis $\mathbf{x}^a$ is obtained by

$$\begin{aligned} \mathbf{x}_t^a &= (\mathbf{I} - \mathbf{K}_t\mathbf{H})\,\mathbf{x}_t^f + \mathbf{K}_t\mathbf{y}^o, \\ &= \mathbf{x}_t^f + \mathbf{K}_t\left(\mathbf{y}^o - \mathbf{H}\mathbf{x}_t^f\right), \end{aligned} \tag{2.8}$$

where $\mathbf{K}_t$ is the Kalman gain matrix at time t, which is a weight matrix to the observation. $\mathbf{H}$ denotes the linear tangent matrix of the observation operator. The observation error $\delta\mathbf{y}_t$ is obtained by

$$\delta\mathbf{y}_t = \mathbf{y}_t^o - \mathbf{H}\mathbf{x}_t^t, \tag{2.9}$$

and $\mathbf{y}_t^o - \mathbf{H}\mathbf{x}_t^f$ is written as

$$\begin{aligned} \mathbf{y}_t^o - \mathbf{H}\mathbf{x}_t^f &= \mathbf{y}_t^o - \mathbf{H}\mathbf{x}_t^t + \mathbf{H}\mathbf{x}_t^t - \mathbf{H}\mathbf{x}_t^f, \\ &= \delta\mathbf{y}_t^o - \mathbf{H}\delta\mathbf{x}_t^f. \end{aligned} \tag{2.10}$$

From Eq. (2.8) the analysis error $\delta\mathbf{x}_t^a$ is obtained by

$$\begin{aligned}
\delta\mathbf{x}_t^a &= \mathbf{x}_t^a - \mathbf{x}_t^t, \\
&= (\mathbf{I} - \mathbf{K}_t\mathbf{H})\,\mathbf{x}_t^f + \mathbf{K}_t\mathbf{y}^o - \mathbf{x}_t^t, \\
&= (\mathbf{I} - \mathbf{K}_t\mathbf{H})\,\delta\mathbf{x}_t^f + \mathbf{K}_t\delta\mathbf{y}^o. \qquad (2.11)
\end{aligned}$$

Therefore, the analysis error covariance $\mathbf{P}_t^a$ is obtained by

$$\begin{aligned}
\mathbf{P}_t^a &= \left\langle \delta\mathbf{x}_t^a \left(\delta\mathbf{x}_t^a\right)^\top \right\rangle, \\
&= \left\langle \left((\mathbf{I} - \mathbf{K}_t\mathbf{H})\,\delta\mathbf{x}_t^f + \mathbf{K}_t\delta\mathbf{y}^o\right)\left((\mathbf{I} - \mathbf{K}_t\mathbf{H})\,\delta\mathbf{x}_t^f + \mathbf{K}_t\delta\mathbf{y}^o\right)^\top \right\rangle, \\
&= \left\langle (\mathbf{I} - \mathbf{K}_t\mathbf{H})\,\delta\mathbf{x}_t^f \left(\delta\mathbf{x}_t^f\right)(\mathbf{I} - \mathbf{K}_t\mathbf{H})^\top \right\rangle + \left\langle (\mathbf{I} - \mathbf{K}_t\mathbf{H})\,\delta\mathbf{x}_t^f \left(\mathbf{K}_t\delta\mathbf{y}^o\right)^\top \right\rangle, \\
&\quad + \left\langle \mathbf{K}_t\delta\mathbf{y}^o \left((\mathbf{I} - \mathbf{K}_t\mathbf{H})\,\delta\mathbf{x}_t^f\right)^\top \right\rangle + \left\langle \mathbf{K}\delta\mathbf{y}^o \left(\mathbf{K}\delta\mathbf{y}^o\right)^\top \right\rangle, \\
&= (\mathbf{I} - \mathbf{K}_t\mathbf{H})\,\mathbf{P}_t^f\,(\mathbf{I} - \mathbf{K}_t\mathbf{H})^\top + \mathbf{K}_t\mathbf{R}\mathbf{K}^\top. \qquad (2.12)
\end{aligned}$$

The cross term $\delta\mathbf{x}_t^f \left(\mathbf{K}_t\delta\mathbf{y}^o\right)^\top$ is equal to zero because we assume that there is no correlation between $\delta\mathbf{x}_t^a$ and $\delta\mathbf{x}_t^f$. The variance of the analysis error is represent $trace(\mathbf{P}^a)$, and when $\dfrac{\partial}{\partial\mathbf{K}}(trace\mathbf{P}^a)$ is equal to zero, the Kalman gain $\mathbf{K}$ makes $trace(\mathbf{P}^a)$ the smallest. In order to obtain the Kalman gain $\mathbf{K}$, we should solve $\dfrac{\partial}{\partial\mathbf{K}}(trace\mathbf{P}^a) = 0$.

$$\begin{aligned}
\frac{\partial}{\partial\mathbf{K}}(trace(\mathbf{P}^a)) &= \frac{\partial}{\partial\mathbf{K}}\left[trace\left((\mathbf{I} - \mathbf{K}\mathbf{H})\,\mathbf{P}^f\,(\mathbf{I} - \mathbf{K}\mathbf{H})^\top + \mathbf{K}\mathbf{R}\mathbf{K}^\top\right)\right], \\
&= \frac{\partial}{\partial\mathbf{K}}\left[trace\left(\mathbf{P}^f - \mathbf{P}^f\mathbf{H}^\top\mathbf{K}^\top - \mathbf{K}\mathbf{H}\mathbf{P}^f + \mathbf{K}\mathbf{H}\mathbf{P}^f\mathbf{H}^\top\mathbf{K}^\top\right) + \mathbf{K}\mathbf{R}\mathbf{K}^\top\right].
\end{aligned}$$

And because of the forecast error covariance $\mathbf{P}^f$ is symmetric matrix,

$$\begin{aligned}
trace\left(\mathbf{P}^f\mathbf{H}^\top\mathbf{K}^\top\right) &= trace\left(\mathbf{K}\mathbf{H}\mathbf{P}^f\right)^\top, \\
&= trace\left(\mathbf{K}\mathbf{H}\mathbf{P}^f\right).
\end{aligned}$$

Therefore,

$$\begin{aligned}\frac{\partial}{\partial \mathbf{K}}(trace(\mathbf{P}^a)) &= \frac{\partial}{\partial \mathbf{K}}\left[trace(\mathbf{P}^f) - 2trace(\mathbf{KHP}^f)\right., \\ &\quad \left. + trace(\mathbf{KHP}^f\mathbf{H}^\top\mathbf{K}^\top) + trace(\mathbf{KRK}^\top)\right]. \end{aligned} \tag{2.13}$$

Here, the formulas of matrix is as follows:

$$\frac{\partial}{\partial \mathbf{A}}\left(trace(\mathbf{ABA}^\top)\right) = \mathbf{A}(\mathbf{B}+\mathbf{B}^\top), \tag{2.14}$$

$$\frac{\partial}{\partial \mathbf{A}}\left(trace(\mathbf{AB})\right) = \mathbf{B}^\top. \tag{2.15}$$

Applying the formulas (Gelb et al. 1974; eqs. (2.1-72) and (2.1-73)),

$$\begin{aligned}\frac{\partial}{\partial \mathbf{K}}(trace(\mathbf{P}^a)) &= -2\mathbf{P}^f\mathbf{H}^\top + 2\mathbf{KHP}^f\mathbf{H}^\top + 2\mathbf{KR}, \\ &= -2\left(\mathbf{I}-\mathbf{KH}\right)\mathbf{P}^f\mathbf{H}^\top + 2\mathbf{KR} = 0, \end{aligned} \tag{2.16}$$

where the observation error covariance matrix $\mathbf{R}$ is a symmetric matrix. Therefore,

$$\mathbf{K} = \mathbf{P}^f\mathbf{H}^\top\left(\mathbf{HP}^f\mathbf{H}^\top + \mathbf{R}\right)^{-1}. \tag{2.17}$$

By substituting Eq. (2.17) into Eq. (2.12), the analysis error covariance matrix $\mathbf{P}^a$ is obtained by $(\mathbf{HP}^f\mathbf{H}^\top + \mathbf{R} = \mathbf{A})$

$$\begin{aligned}\mathbf{P}^a &= \left(\mathbf{I}-\mathbf{KH}\right)\mathbf{P}^f\left(\mathbf{I}-\mathbf{KH}\right)^\top + \mathbf{KRK}^\top, \\ &= \left(\mathbf{I}-\mathbf{P}^f\mathbf{H}^\top\mathbf{A}^{-1}\mathbf{H}\right)\mathbf{P}^f\left(\mathbf{I}-\mathbf{P}^f\mathbf{H}^\top\mathbf{A}^{-1}\mathbf{H}\right)^\top + \mathbf{P}^f\mathbf{H}^\top\mathbf{A}^{-1}\mathbf{R}\left(\mathbf{P}^f\mathbf{H}^\top\mathbf{A}^{-1}\right)^\top, \\ &= \mathbf{P}^f - 2\mathbf{P}^f\mathbf{H}^\top\mathbf{A}^{-1}\mathbf{HP}^f + \mathbf{P}^f\mathbf{H}^\top\mathbf{A}^{-1}\mathbf{HP}^f\mathbf{H}^\top\mathbf{A}^{-1}\mathbf{HP}^f + \mathbf{P}^f\mathbf{H}^\top\mathbf{A}^{-1}\mathbf{RA}^{-1}\mathbf{HP}^f, \\ &= \mathbf{P}^f - 2\mathbf{P}^f\mathbf{H}^\top\mathbf{A}^{-1}\mathbf{HP}^f + \mathbf{P}^f\mathbf{H}^\top\mathbf{A}^{-1}\left(\mathbf{HP}^f\mathbf{H}^\top + \mathbf{R}\right)^{-1}\mathbf{A}^{-1}\mathbf{HP}^f, \\ &= \mathbf{P}^f - 2\mathbf{P}^f\mathbf{H}^\top\mathbf{A}^{-1}\mathbf{HP}^f + \mathbf{P}^f\mathbf{H}^\top\mathbf{A}^{-1}\mathbf{AA}^{-1}\mathbf{HP}^f, \\ &= \mathbf{P}^f - \mathbf{P}^f\mathbf{H}^\top\left(\mathbf{HP}^f\mathbf{H}^\top + \mathbf{R}\right)^{-1}\mathbf{HP}^f, \\ &= \left(\mathbf{I}-\mathbf{KH}\right)\mathbf{P}^f. \end{aligned} \tag{2.18}$$

In summary, the algorithm of the EKF is as follows: In the forecast step, the state variable $\mathbf{x}^f$ and the forecast error covariance $\mathbf{P}^f$ at t are written as

$$\mathbf{x}_t^f = M\mathbf{x}_{t-1}^a, \tag{2.19}$$

$$\mathbf{P}_t^f = \mathbf{M}\mathbf{P}_{t-1}^a\mathbf{M}^\top + \mathbf{Q}. \tag{2.5}$$

In the analysis step, the Kalman gain weight matrix $\mathbf{K}$ is written as

$$\mathbf{K}_t = \mathbf{P}_t^f\mathbf{H}^\top\left(\mathbf{H}\mathbf{P}_t^f\mathbf{H}^\top + \mathbf{R}\right)^{-1}. \tag{2.17}$$

The analysis $\mathbf{x}^a$ and the analysis error covariance $\mathbf{P}^a$ are written as

$$\mathbf{x}_t^a = \mathbf{x}_t^f + \mathbf{K}_t\left(\mathbf{y}^o - \mathbf{H}\mathbf{x}_t^f\right), \tag{2.8}$$

$$\mathbf{P}_t^a = \left(\mathbf{I} - \mathbf{K}_t\mathbf{H}\right)\mathbf{P}_t^f. \tag{2.18}$$

EKF can treat the flow-dependent covariance explicitly. In the EKF algorithm we must compute $\mathbf{M}\mathbf{P}_{t-1}^a\mathbf{M}^\top$, which size is n by n. n represents degree of freedom of the model M. In the atmospheric models, n is generally over $O(10^6)$, so it is not feasible to compute $\mathbf{M}\mathbf{P}_{t-1}^a\mathbf{M}^\top$ in the EKF.

2.2 Ensemble Kalman Filter

It is noteworthy that $\mathbf{P}_t^f$ in Eq. (2.5) requires the evolution of $\mathbf{P}_t^a$, which is impossible for the general AGCMs, since the degree of freedom n of the AGCMs is

at least $O(10^6)$. The implementation of EKF costs much computational resources to compute the covariance matrices of $n \times n$. In addition, the EKF uses the linearized tangent model $\mathbf{M}$ and its adjoint model $\mathbf{M}^\top$ to compute the covariance matrix, which is not easy and is so difficult to compute for the AGCMs, and if the nonlinear model M is updated, it is necessary that $\mathbf{M}$ and $\mathbf{M}^\top$ are updated.

In contrast, the EnKF does not require the computation of $\mathbf{M}$ and $\mathbf{M}^\top$ of $n \times n$ matrices. Moreover, the EnKF can use the nonlinear model M and nonlinear operator H instead of $\mathbf{M}$, $\mathbf{M}^\top$ and $\mathbf{H}$. The EKF is approximated by the EnKF using the technique of ensemble predictions. The main idea of the EnKF is as follows. First, we set an ensemble of the initial conditions at time $t-1$ which spread around $\mathbf{x}^a_{t-1}$. The analysis error covariance $\mathbf{P}^a_{t-1}$ can be estimated from the ensemble perturbations. Next, the nonlinear model propagates each ensemble member to time t. Then we can estimate $\mathbf{P}^f_t$ based on perturbations of the propagated ensemble members at time t.

There are many possible choices of the analysis ensemble. A variety of EnKFs have been proposed, and one of the main differences among them is how the analysis ensemble is chosen. In this section, we choose the LETKF approach. In the LETKF we compute the analysis step in the space spanned by the ensemble perturbations.

Let $\mathbf{X}$ denotes an $n \times m$ matrix composed of m ensemble members of n-dimensional variables in the local patch. The local patch is defined as a $l \times l$ area of grid points. $\mathbf{X}$ is written as

$$\mathbf{X} = \left[\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \cdots, \mathbf{x}^{(m)} \right], \tag{2.20}$$

where $\mathbf{x}^{(i)}$ denotes the i-th ensemble member. Predictions are performed for each ensemble member according to the nonlinear model,

$$\mathbf{X}_t^f = M(\mathbf{X}_{t-1}^a). \tag{2.21}$$

In order to estimate the forecast error covariance, we evaluate the ensemble of the forecast errors:

$$\begin{aligned} \delta\mathbf{X}_t^f &= \left[\, \mathbf{x}^{f(1)} - \bar{\mathbf{x}}^f, \mathbf{x}^{f(2)} - \bar{\mathbf{x}}^f, \cdots, \mathbf{x}^{f(m)} - \bar{\mathbf{x}}^f \,\right], \\ &= \mathbf{X}_t^f - \bar{\mathbf{X}}_t^f, \end{aligned} \tag{2.22}$$

where $\bar{\mathbf{x}}_t^f$ and $\bar{\mathbf{X}}_t^f$ denote the ensemble mean vector and matrix of the forecast, respectively. The forecast error covariance $\mathbf{P}_t^f$ may be estimated as,

$$\mathbf{P}_t^f = \frac{\delta\mathbf{X}_t^f \left(\delta\mathbf{X}_t^f\right)^\top}{m-1} = \delta\mathbf{X}_t^f \, \tilde{\mathbf{P}}_t^f \left(\delta\mathbf{X}_t^f\right)^\top, \tag{2.23}$$

where, $\tilde{\mathbf{P}}_t^f = \dfrac{\mathbf{I}}{m-1}$. The $\tilde{\mathbf{P}}_t^f$ is a covariance matrix in the space spanned by the ensemble forecast errors in the m-dimensional space. Similarly, the analysis error covariance matrix $\mathbf{P}_t^a$ may be computed by

$$\mathbf{P}_t^a = \delta\mathbf{X}_t^f \, \tilde{\mathbf{P}}_t^a \left(\delta\mathbf{X}_t^f\right)^\top, \tag{2.24}$$

where $\tilde{\mathbf{P}}^a$ is given by

$$\begin{aligned} \tilde{\mathbf{P}}_t^a &= \left[(m-1)\mathbf{I}/\rho + \left(\mathbf{H}\delta\mathbf{X}_t^f\right)^\top \mathbf{R}^{-1} \left(\mathbf{H}\delta\mathbf{X}_t^f\right)\right]^{-1}, \\ &= \left[\mathbf{V}\mathbf{\Lambda}\mathbf{V}^\top\right]^{-1}, \\ &= \mathbf{V}\mathbf{\Lambda}^{-1}\mathbf{V}^\top, \end{aligned} \tag{2.25}$$

where ρ is a covariance inflation parameter, which inflates the analysis error covariance matrix to avoid underestimating the error covariance and filter divergence. ρ is slightly

larger than 1, and the covariance inflation is expressed by multiplying the square root of ρ to the $\delta\mathbf{X}^f$. The analysis error covariance $\tilde{\mathbf{P}}_t^a$ can be calculated by solving the eigenvalue problem of $(\tilde{\mathbf{P}}_t^a)^{-1}$, where $\mathbf{\Lambda}$ is a diagonal matrix containing the eigenvalues, and $\mathbf{V}$ is a matrix containing eigenvectors with the orthonormal conditions $\mathbf{V}^\top\mathbf{V} = \mathbf{I}$. We assume that the observational errors are uncorrelated. Therefore, $\mathbf{R}$ is diagonal, and the inverse is trivial. In the model space, the analysis ensemble mean is evaluated as

$$\bar{\mathbf{x}}_t^a = \bar{\mathbf{x}}_t^f + \delta\mathbf{X}_t^f\tilde{\mathbf{P}}_t^a\left(\mathbf{H}\delta\mathbf{X}_t^f\right)^\top\mathbf{R}^{-1}\left(\mathbf{y}_t^o - \mathbf{H}\bar{\mathbf{x}}_t^f\right), \tag{2.26}$$

and the analysis perturbation may be obtained as,

$$\begin{aligned}\delta\mathbf{X}_t^a &= \delta\mathbf{X}_t^f\left[(m-1)\tilde{\mathbf{P}}_t^a\right]^{1/2}, \\ &= \delta\mathbf{X}_t^f\sqrt{m-1}\mathbf{V}\mathbf{\Lambda}^{-1/2}\mathbf{V}^\top.\end{aligned} \tag{2.27}$$

Finally, the analysis ensemble $\mathbf{X}_t^a$ is prepared using Eq. (2.27) as,

$$\mathbf{X}_t^a = \bar{\mathbf{X}}_t^a + \delta\mathbf{X}_t^a, \tag{2.28}$$

where $\bar{\mathbf{X}}_t^a$ is the ensemble mean matrix of the analysis. Furthermore, we can use the nonlinear observational operator H in place of the linear observational operator $\mathbf{H}$ as

$$\mathbf{H}\delta\mathbf{X}^f \simeq H(\mathbf{X}^f) - H(\bar{\mathbf{X}}^f), \tag{2.29}$$

$$\mathbf{H}\bar{\mathbf{x}}^f \simeq H(\bar{\mathbf{x}}^f). \tag{2.30}$$

Hence, LETKF has been completed without using the Kalman gain matrix and the linear operators, explicitly.

CHAPTER III

Comparison of the Extended Kalman Filter and the Ensemble Kalman Filter using the Barotropic General Circulation Model

3.1 Barotropic S-Model

The model equations used in this chapter are the 3D spectral primitive equations on a sphere (Tanaka 2003). A system of primitive equations with a spherical coordinate of longitude λ, latitude θ, pressure p, and time t may be reduced to three prognostic equations of horizontal motions and thermodynamics for three dependent variables (u, v, ϕ). Here, u and v are the horizontal velocity. The variable ϕ is a departure of the local isobaric geopotential from the global mean reference state. By expanding the state variables in 3D normal mode functions, we obtain a system of 3D spectral primitive equations in terms of the spectral expansion coefficients:

$$\frac{dw_i}{d\tau} = -i\sigma_i w_i - i\sum_{jk} r_{ijk} w_j w_k + f_i, \quad i = 1, 2, 3, \ldots, \tag{3.1}$$

where τ is a dimensionless time, σ_i is the eigen-frequency of the Laplace's tidal equation, w_i is the expansion coefficient of the state variables, f_i is the expansion coefficient of the external forcing of viscosity and diabatic heating rate, and r_{ijk} is the interaction coefficient for nonlinear wave-wave interactions.

In the 3D spectral representation, the vertical expansion basis functions may be divided into barotropic ($m = 0$) and baroclinic ($m \neq 0$) components. In this chapter, we use only the barotropic components ($m = 0$) of Rossby modes, by truncating all the baroclinic modes and high-frequency gravity modes (Kasahara 1977). Such a model is equivalent to predicting the vertical average of meteorological variables. The zonal and meridional wave truncation of the present model is equivalent to rhomboidal 20 with an equatorial wall. Because w_i is a complex-valued state variable, only the positive (and zero) zonal wave numbers are considered. Therefore, the degree of freedom of the system is reduced enormously to only N=410:

$$\frac{dw_i}{d\tau} = -i\sigma_i w_i - i\sum_{jk} r_{ijk} w_j w_k + f_i + g_i, \quad i = 1, 2, 3, \ldots. \tag{3.2}$$

The spectral equation for such a barotropic model Eq. (3.2) has the same form as for the baroclinic model Eq. (3.1), except for the fact that the barotropic-baroclinic interaction g_i appears on the right-hand side. Henceforth, we combine $f_i + g_i$ and designate it as the external forcing s_i of the barotropic model.

The parameterization of the external forcing $s_i = f_i + g_i$ in Eq. (3.2) is presented by Tanaka (1991, 1998) considering the dynamical processes of the baroclinic instability, the topographic forcing, the biharmonic diffusion, the zonal surface stress,

and the Ekman pumping. Since those parameterizations are fundamentally linear approximations to the true forcing and will never be perfect for the real atmosphere, we have attempted in this chapter to obtain the best linear fit of the forcing s_i based on the 56 years of the NCEP/NCAR reanalysis from 1950 to 2005 as in Tanaka and Nohara (2001):

$$s_i = \tilde{s}_i + s_i^{'} = \tilde{s}_i + \mathbf{A}_{ij} w_j + \mathbf{B}_{ij} w_j^* + \varepsilon_i, \tag{3.3}$$

where $\tilde{s}_i$ and $s_i^{'}$ are the climate and anomaly of s_i, respectively. The asterisk for w_i represents a complex conjugate of w_i. The linear matrices $\mathbf{A}_{ij}$ and $\mathbf{B}_{ij}$ can be obtained by the standard method of the least square fitting for $s_i^{'}$ to minimize the regression residual ε_i. However, the system matrices $\mathbf{A}_{ij}$ and $\mathbf{B}_{ij}$ mostly represent topographic forcing (TF). For this reason, we have added the energy source in terms of the parameterized baroclinic instability (BC) described by Tanaka (1998). The increased energy source at the synoptic scale requires an additional energy sink, so the biharmonic diffusion (DF), the zonal surface stress (DZ), and the Ekman pumping (DE) are also included in reference to the resulting energy spectrum of the model climate. The final form of the external forcing s_i is expressed as

$$s_i = \tilde{s}_i + \mathbf{A}_{ij} w_j + \mathbf{B}_{ij} w_j^* + (BC)_{ij} w_j + (DF)_{ij} w_j + (DZ)_{ij} w_j + (DE)_{ij} w_j. \tag{3.4}$$

This model with such a forcing is called "barotropic S-model" since a part of forcing is obtained statistically from observation, solving the inverse problem.

3.2 Experimental Settings

The EKF and the LETKF are implemented to the abovementioned barotropic S-model under the perfect model scenario.

First, the EKF and LETKF assimilate the observations every 6 hours in the spectral space. In the physical space with the small ensemble size, the spurious covariance among distant points due to sampling errors can be reduced by covariance localization (Houtekamer and Michell 2001; Hamill et al. 2001). According to Szunyogh et al. (2005), there is no clear difference between the ensemble size 40 and 80 in the case of optimum localization. However, in the spectral space, the error covariance of the spectral expansion coefficient w_i cannot be appropriately localized, because interactions between different coefficients are highly nonlocal. So, the LETKF in this study does not employ the localization when the assimilation is conducted in the spectral space. The LETKF experiments employ 20, 50, 100, 410 and 1000 ensemble members. The period of the experimental assimilation is from 00Z 1 January 1990 to 00Z 31 January 1990. The covariance inflation parameter $\rho^m_{nonlocal}(m = 20, 50, 100, 410, 1000)$ is optimized for each experiment to minimize the estimated errors ranging from 1.005 to 1.10. Specifically, $\rho^{20}_{nonlocal} = 1.10,\ \rho^{50}_{nonlocal} = 1.03,\ \rho^{100}_{nonlocal} = 1.01,\ \rho^{410}_{nonlocal} = 1.005,\ \rho^{1000}_{nonlocal} = 1.005.$ The truth of w^t_i is the model output from the control run, and the observation w^o_i is made by adding random vectors to the truth. The random vectors are generated from the normal distribution having the observational error statistics at the gravity

wave level, which is a white noise of 50 J/m^2 in its energy level (see the observational analysis by Tanaka 1985). The noise in its energy level is equivalent to about 5.9 m in barotropic height and 1.0 m/s in wind speed. For simplicity, we assume that all model variables are observed, i.e., the dimension of the observation w_i^o is the same as the dimension of the truth w_i^t. So, H and its tangent linear operator $\mathbf{H}$ are equal to the unit matrix $\mathbf{I}$. We assume that the model error covariance $\mathbf{Q}$ is 0 because the perfect model has been assumed. At each analysis time, we evaluate the analysis error using the RMSE between the true state and the analysis ensemble mean and compute the ensemble spread. The ensemble spread is measured using the trace of the analysis covariance matrix in Eq. (2.24).

Second, the LETKF experiment is conducted in the physical space for the confirmation purpose in order to investigate the impacts of the localization and in order to investigate the influence of assimilation in the physical space and in the spectral space. This LETKF is the non-local patch version (Miyoshi et al. 2007). The ensemble member is chosen to be 20, 50, 100 and 410 and the observations are assimilated every 6 hours. To assimilate the same observations of the experiments as in the spectral space, the observations in spectral space are transformed to the observations in physical space, and the observational covariance matrix is computed using the observations in physical space. The observational elements are barotropic height and horizontal wind components. The physical space for the barotropic model is constructed on the 2-dimensional space, with 72 grid points in longitudes and 30 grid points in Gaussian latitudes in the Northern Hemisphere. The observations cover

about 17 % grid points of the entire 2-dimensional grid space. It is assumed that the thinned observational errors are not correlated with each other. So the observational error covariance matrix is diagonal. In this study, Gaussian-like fifth order polynomial function (Gaspari and Cohn 1999) is adopted for the localization, by multiplying its inverse to the diagonal elements of the observational covariance matrix. The localization scale is defined by the one standard deviation. The horizontal localization scales for 20, 50, 100, and 410 members are 900 km, 900 km, 1200 km and 1200 km, respectively. The Gaussian-like function drops to zero at about 3300 km to 4400 km. The covariance inflation parameter $\rho^m_{local}(m = 20, 50, 100, 410)$ is changed for each experiment to minimize the estimated errors ranging from 1.07 to 1.08. Specifically, $\rho^{20}_{local} = 1.08,\ \rho^{50}_{local} = 1.07,\ \rho^{100}_{local} = 1.08,\ \rho^{410}_{local} = 1.08$. The other settings are the same with the non-localized LETKF experiments.

In this study, the non-localized LETKF is called "EnKF" which assimilates the observations in spectral space, and the localized LETKF is called "LETKF" which assimilates the observations in the physical space.

The initial ensemble for the first step of the assimilation experiments is prepared by the method of lagged average forecast (LAF) ensemble. First, we conduct many forecasts starting from different observational times. Let $\mathbf{x}_t$ be the initial value at the valid time t. If we write the forecast at time t starting from one day before as $\mathbf{x}_t^{f(1)}$, then the forecast with the initial value at k days before may be denoted as $\mathbf{x}_t^{f(k)}$. The forecast errors $\delta\mathbf{x}_t^{f(i)}$ up to $i = 1$ to k are then orthogonalized by analyzing the principal components. In this experiment, the norms of ensemble perturbations are

set to about 4 times of the observational error to have the magnitude of the forecast error at 7-day prediction because the error growth of the barotropic S-model is known to be relatively weak as demonstrated by Tanaka and Nohara (2001). By cycling the data assimilation, the norm of the ensemble perturbation becomes smaller. The predictions up to k days provide $m = 2k$ ensemble members of the initial perturbation by the pair of positive and negative signs of the vector. Thus, we can compute the initial perturbations by this method up to $m = 2N = 820$. When we need more initial perturbations, we repeat the same procedure using different years of the winter data, although those are not orthogonal. However, there is no problem because the EnKF updates these ensemble members to reflect the analysis error covariance at every analysis cycle.

The initial $\mathbf{P}^a$ of the EKF is set as 2 times of the observational covariance matrix $\mathbf{R}$. This assumption does not have any problem because the $\mathbf{P}^a$ is updated by the EKF at every analysis cycle.

3.3 Results

Figure 3.1 shows the eigenvalues of the analysis covariance matrices for the EKF and the EnKFs on 00Z 31 January 1990, after the filters have converged to the truth. There are only 19, 49, and 99 eigenvalues for the EnKF of 20, 50, and 100 ensemble members, respectively. The eigenvalues for the analysis covariance show the variance spectrum of the analysis error. The variance spectra of the EnKF for 20 ensemble

members differ from that of the EKF, but the spectra are approximately the same as that of the EKF for the 50 ensemble members. Interestingly, the magnitudes of eigenvalues are smaller than that of the EKF for the ensemble members 100 and more. The difference may come from the assumption of the linearization in the EKF, which is not assumed in the EnKF. It is found from the result that the EnKF has almost converged when the ensemble member has increased to 100. The magnitude of the 100th eigenvalue for the EKF is about $1/10^5$ of that of the first eigenvalue, suggesting that the matrix almost degenerates.

Figure 3.2 illustrates the first eigenvectors of the analysis covariance matrices for the EKF and the EnKFs on the same day as in Fig. 3.1, describing the most dominant pattern of the analysis error. We can see the waves of baroclinic instability over the east Pacific Ocean, the Northern American continent, and the Atlantic Ocean in the cases of the EKF and the EnKFs of 50, 100, 410 and 1000 members. On the other hand, these waves are absent in the cases of EnKFs of 20 members. In all cases there is another peak of error over the Arctic. We can confirm that both of the EKF and EnKF have converged to the same error pattern for 50, 100, 410 and 1000 members. The second eigenvectors (not shown) are not the same in the EKF and EnKF probably due to the assumption of the model linearization in the EKF.

Figure 3.3 shows the time series of the RMSEs against the truth of the state variables w_i and the analysis ensemble spreads for the EKF and the EnKFs during one month. The observational error is about 2.25×10^{-4}. The analysis errors for both EKF and EnKFs are reduced by assimilating the observations iteratively. In the case

of the EnKFs with 50, 100, 410 and 1000 members, the RMSEs are comparable to the ensemble spread, respectively. In the case of the EnKF of 20 members, however, the RMSE is larger than the spread. This shows that the EnKF of 20 members may underestimate the forecast error covariance and does not assimilate the observations. The RMSEs show daily variations, but the ensemble spreads show less daily variations. The RMSEs generally have the same order as the ensemble spreads except for the EnKF of 20 members in this study. The degree of the ensemble spreads can be adjusted by the covariance inflation, which is a constant but can be tuned so that the RMSEs is minimized and becomes comparable to the ensemble spreads.

Figure 3.4 illustrates the analysis fields of barotropic height (contours) and the analysis error against the truth on 00Z 31 January 1990. The analysis error of the EnKF for 20 members is larger than that of the EKF. That of the EnKF of 50 members is almost the same as that of the EKF. Moreover, the analysis error of the EnKF for 410 members is very similar to that of the EnKF for 1000 members. The error patterns of the EKF and EnKF of more than 50 members are similar, and are characterized by the baroclinic instability. These results indicate that the accuracy of the EnKF of 50 members and above is almost the same as that of the EKF.

In order to increase the statistical confidence of the result, we repeated the same experiments for 10 cases in different seasons and different years. Figure 3.5 shows the time series of the mean RMSEs against the truth of the state variables w_i for the average of the 10 cases. The analysis errors decrease for both the EKF and EnKFs by assimilating the observations iteratively. The analysis RMSE of EnKF

for 20 members is the largest, and that for 410 and 1000 members are the smallest. Interestingly, the analysis RMSEs for 100, 410 and 1000 members are smaller than that of the EKF, probably due to the error involved by the model linearization of the EKF. The magnitude of the analysis RMSE for the EKF is almost the same as the EnKF for sufficient number of the members. We see that the EnKF has converged to the EKF for 100 members.

Finally, Fig. 3.6 shows the time series of the RMSEs against the truth of the state variables w_i for the LETKF evaluated in the physical space with localization. The results are compared with the EnKF of 50 ensemble members during one month. The observational error is about 2.25×10^{-4} as before. The analysis RMSE of each LETKF decreases with time, and the RMSE becomes smaller than the observational error in each experiment. Moreover, the analysis ensemble spreads are comparable to the RMSEs in each experiment (not shown). It seems that the analysis RMSEs of the LETKF have converged for 20 members and above. However, the RMSE in the physical space is much larger than that in the spectral space. Although we attempted to expand the localization scale up to 1500 km, the RMSEs are not improved. It is found by the present study that the RMSE of the LETKF in the physical space is not converging to the EnKF in the spectral space where the localization is not assumed. It is considered that there are some reasons in this result, so this result is discussed in chapter VI.

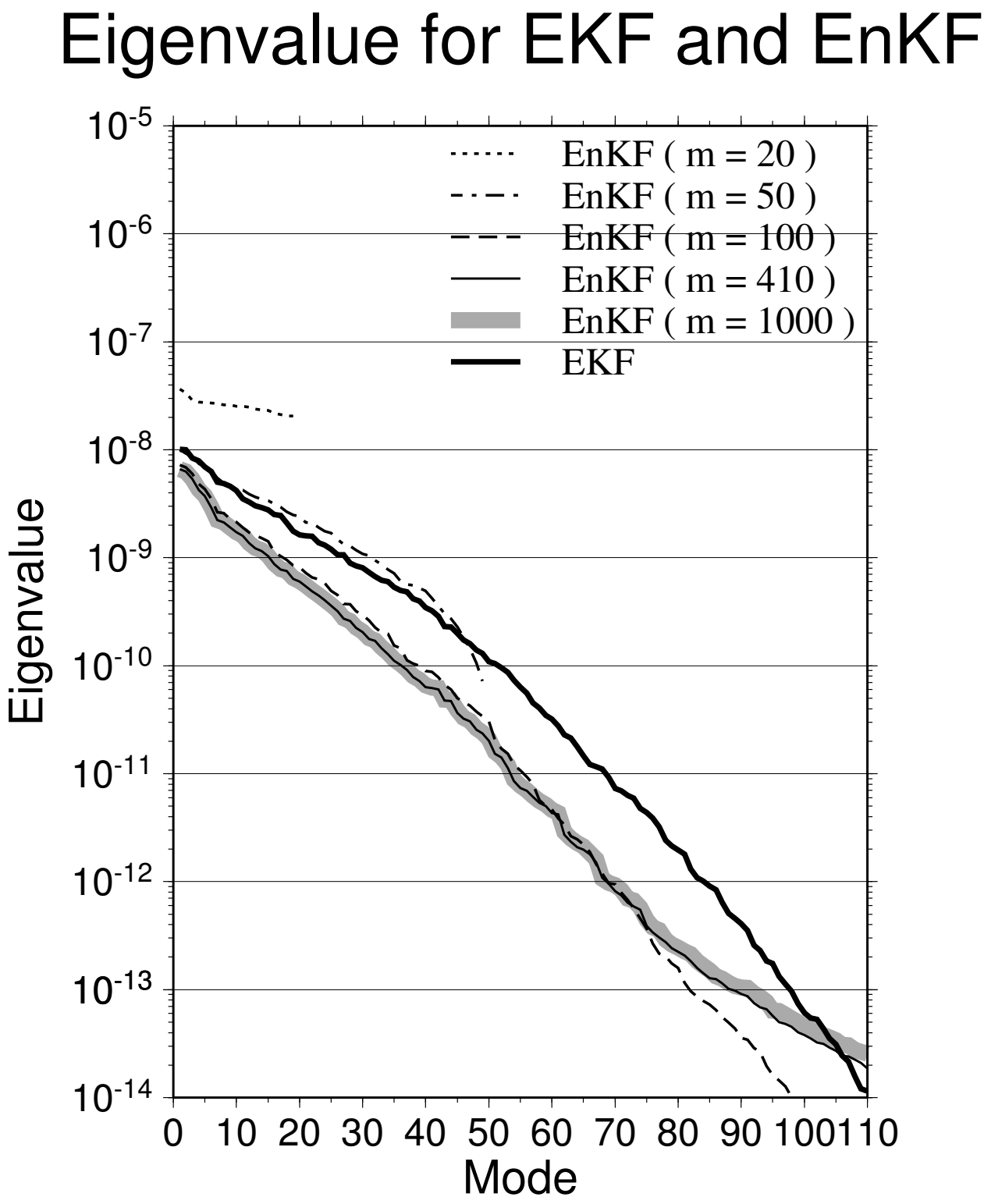


Figure 3.1. Eigenvalues of the analysis covariance matrix for EKF and EnKFs (ensemble member: 20, 50, 100, 410 and 1000) on 00Z 31 January 1990.

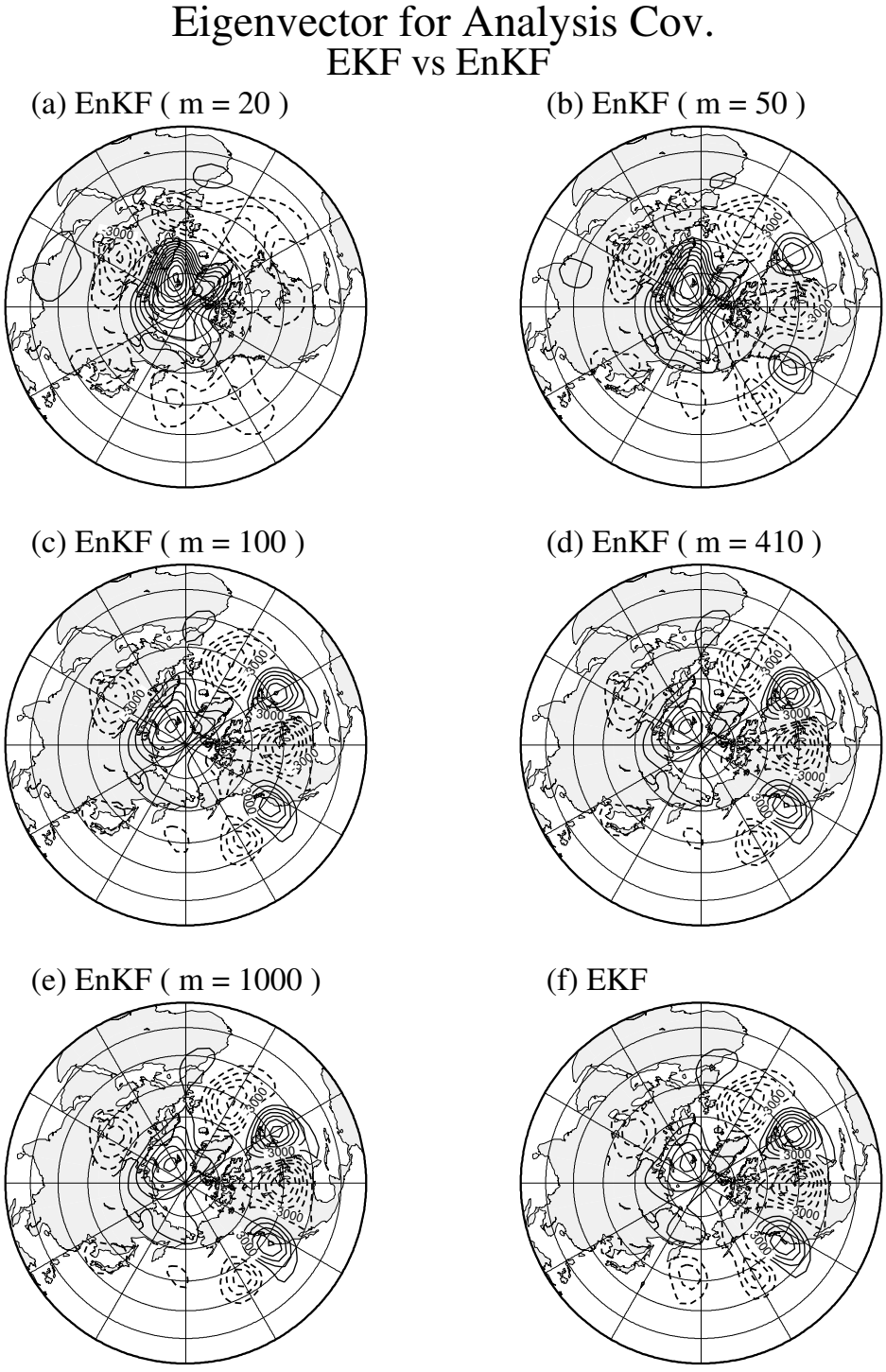


Figure 3.2. First eigenvectors of the analysis covariance matrix for EKF and EnKFs on 00Z 31 January 1990. (a) EnKF with 20 ensemble members, (b) EnKF with 50 ensemble members, (c) EnKF with 100 ensemble members, (d) EnKF with 410 ensemble members, (e) EnKF with 1000 ensemble members and (f) EKF.

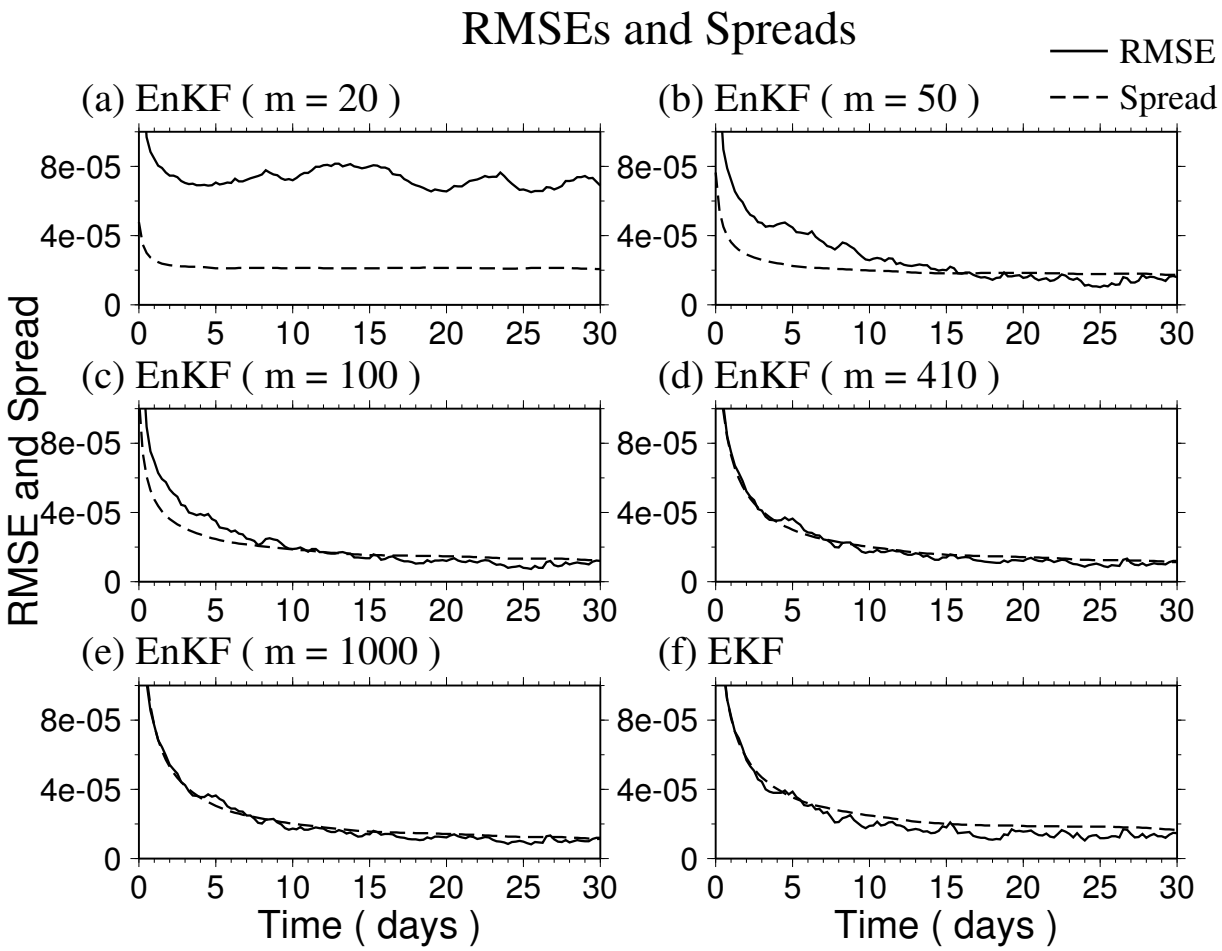


Figure 3.3. Time series of RMSEs (solid lines) against truth of the barotropic expansion coefficient w_i of the state variables (u, v, ϕ) and analysis ensemble spreads (dashed lines) for EKF and EnKF for 1 month from 00Z 1 January 1990 to 00Z 31 January 1990 (0 of the horizontal axes corresponds to 00Z 1 January 1990). (a) EnKF with 20 ensemble members, (b) EnKF with 50 ensemble members, (c) EnKF with 100 ensemble members, (d) EnKF with 410 ensemble members, (e) EnKF with 1000 ensemble members and (f) EKF. The observational error is about 2.25×10^{-4}.

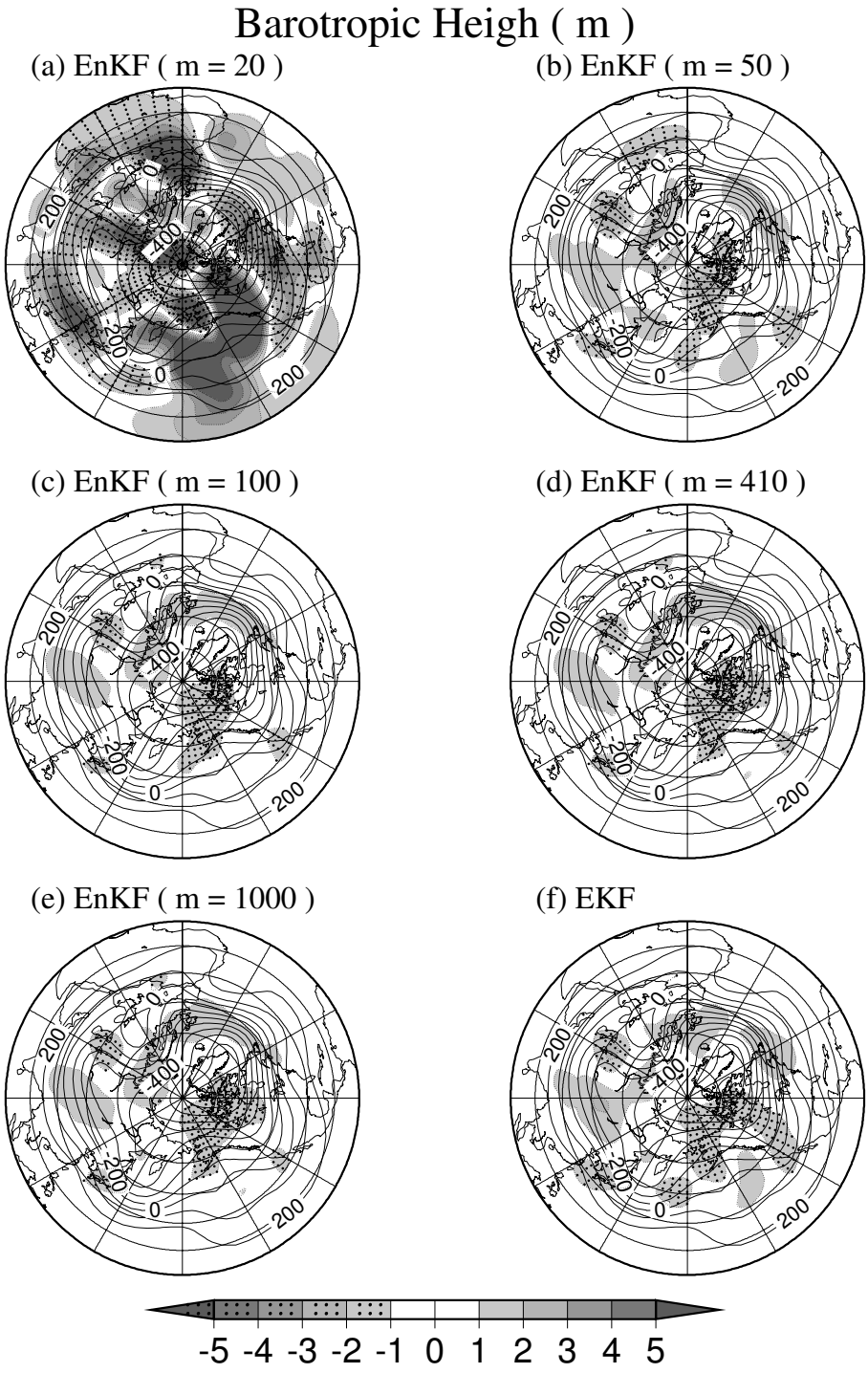


Figure 3.4. Analysis fields (contour) and the differences against truth (shaded) of the barotropic height (m) on 00Z 31 January 1990. (a) EnKF with 20 ensemble members, (b) EnKF with 50 ensemble members, (c) EnKF with 100 ensemble members, (d) EnKF with 410 ensemble members, (e) EnKF with 1000 ensemble members and (f) EKF.

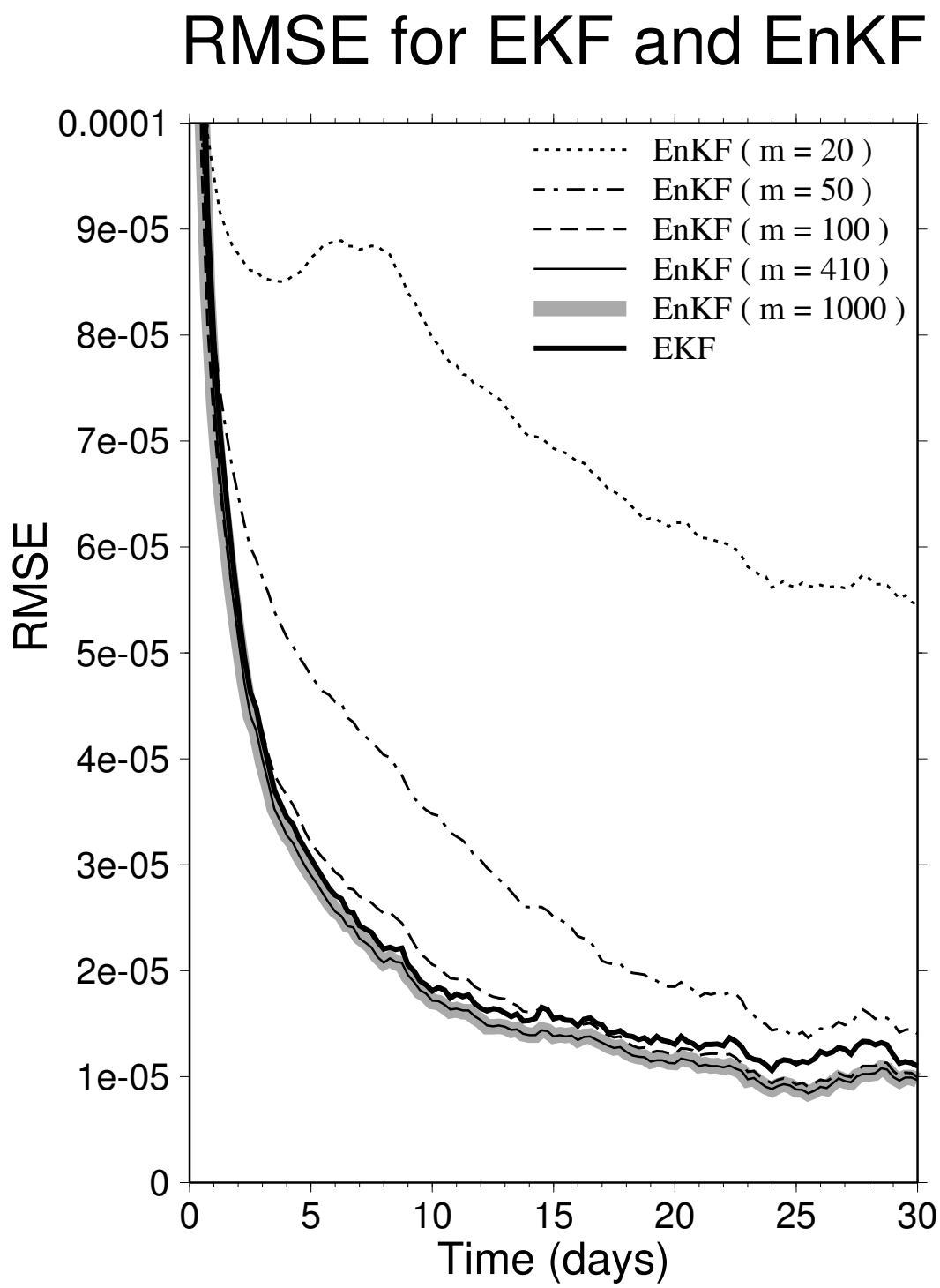


Figure 3.5. Average of RMSEs against truth of the barotropic expansion coefficient w_i of the state variables (u, v, ϕ) for EKF and EnKFs (ensemble member: 20, 50, 100, 410 and 1000). A number of the sample is 10, respectively. The observational error is about 2.25×10^{-4}.

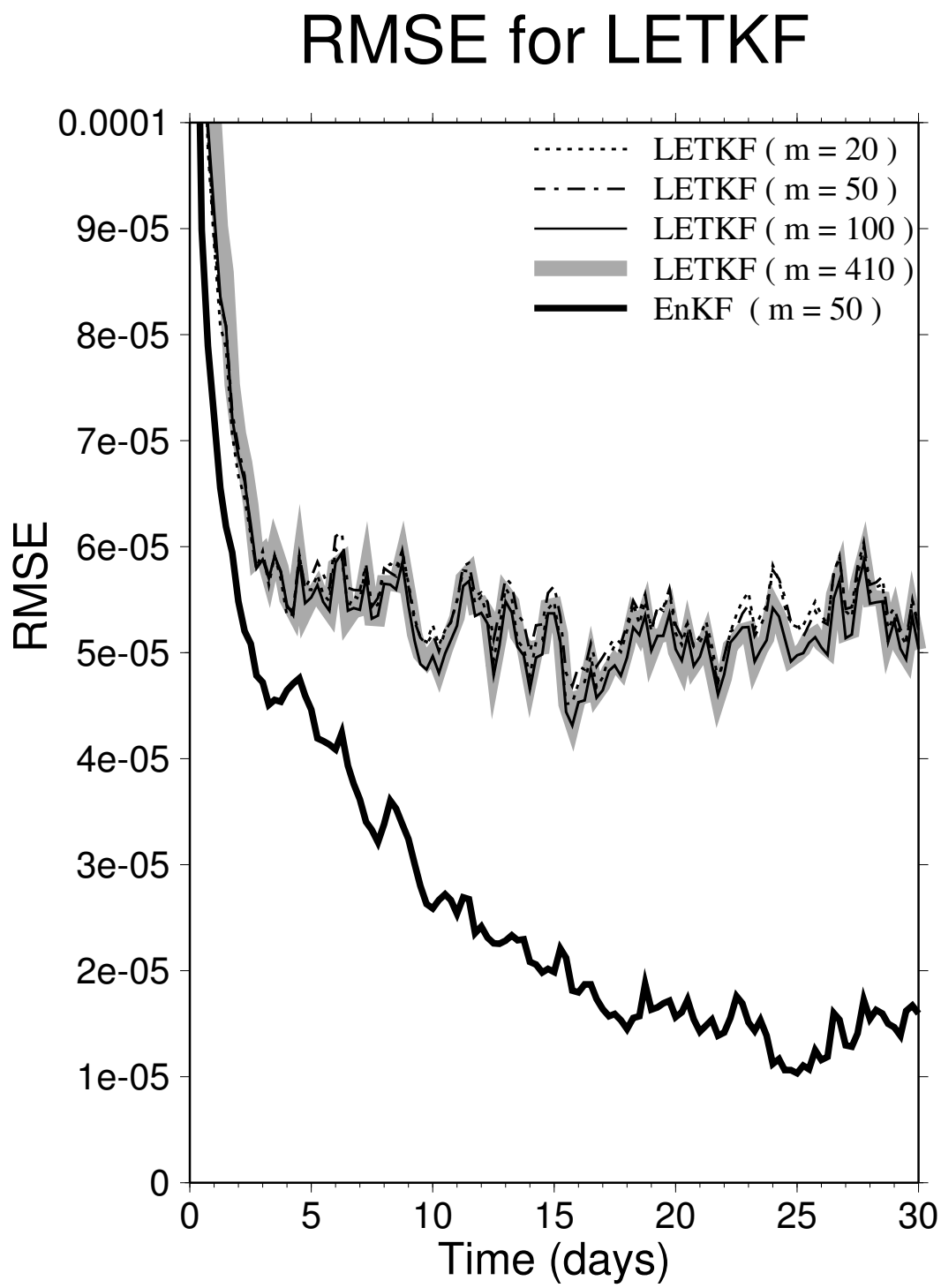


Figure 3.6. Time series of RMSEs against truth of the barotropic expansion coefficient w_i of the state variables (u, v, ϕ) for LETKF and EnKF (member=50) for 1 month from 00Z 1 January 1990 to 00Z 31 January 1990 (0 of the horizontal axes corresponds to 00Z 1 January 1990). The observational error is about 2.25×10^{-4}.

CHAPTER IV

Applying the Local Ensemble Transform Kalman Filter to the Nonhydrostatic Icosahedral Atmospheric Model (NICAM)

4.1 Methodology of the NICAM-LETKF

In this study, data assimilation experiments with the NICAM-LETKF are performed under the perfect model scenario. The forecast model used is the NICAM mentioned in introduction. Its horizontal resolution is 224 km (Glevel-5) and the number of vertical layers is 40. The prognostic variables are pressure, temperature, horizontal and vertical wind components and mixing ratios of water vapor, cloud water and rain water. In the NICAM, the horizontal wind is decomposed to three elements. So, the number of the prognostic variables is 9. The model physics used in this study are Louis's surface layer (Louis 1979), Mellor and Yamada Level 2 (Mellor and Yamada 1974), Arakawa and Schubert's cumulus parameterization (Arakawa and Schubert 1974, Arakawa 2004) and third-order Runge-Kutta method for time integration.

The control run for the truth is generated with an initial data by the JMA/GSM operational analysis on 12Z 30 December 2006, and integrated until 12Z 22 January 2007. The synthetic observations are generated by adding prescribed observational error to the truth. The error standard deviations are 1.0 hPa (pressure), 1.0 K (temperature), 1.0 m/s (horizontal wind) and 0.5 g/kg (mixing ratio of water vapor). The observations cover 10 % grid points of the entire horizontal 2-dimensional grid space and cover about 3.3 % grid points of the entire 3-dimensional grid space uniformly.

In this study, a Gaussian-like fifth order polynomial function (Gaspari and Cohn 1999) is adopted for the horizontal and vertical localization, by multiplying the function to the diagonal elements of the inverse of the observational covariance matrix. The localization scale is defined by the one standard deviation of the normal distribution. The horizontal localization scale is 500 km, and the Gaussian-like function drops to zero at about 1800 km from the analysis point. The vertical localization scale in 30°N–90°N and 30°S–90°S is 4.0 grid points, and that in 20°N–20°S is 3.0 grid points. In the other region the vertical localization scale changes smoothly by a linear interpolation. The Gaussian-like function drops to zero at about 10 grid points to 15 grid points in the vertical. The NICAM-LETKF assimilation cycle is every 6 hours, and the period of the experimental assimilation is from 12Z 1 January 2007 to 12Z 22 January 2007. The ensemble size is fixed to 40, and the initial ensemble members are computed from the JMA operational analysis integrated for 2 days by the NICAM. The dates of the JMA operational analysis are chosen at random to avoid the similar values.

In this study, four experiments are implemented, and in all experiments the multiplicative spread inflation is employed. The details of each experiment are summarized in Table 1. In the first experiment (Ex. 1), the spread inflation parameter is fixed temporally. In detail, the spread inflation is 1% in 30°N–90°N and 30°S–90°S, and is 3% in 20°N–20°S. In the other region the spread inflation changes smoothly by a linear interpolation. In the second experiment (Ex. 2), the spread inflation parameter is estimated adaptively, following the method shown in Miyoshi (2005) and Miyoshi (2011). In Miyoshi (2005) and Miyshi (2011), an estimate of the covariance inflation parameter Δ_o can be obtained from

$$\Delta_o = \frac{trace(\mathbf{d}_{o-f}\mathbf{d}_{o-f}^{\top}) \circ \mathbf{R}^{-1}}{trace(\mathbf{H}\mathbf{P}^f\mathbf{H}^{\top} \circ \mathbf{R}^{-1}) + trace(\mathbf{R} \circ \mathbf{R}^{-1})} - 1, \tag{4.1}$$

where $\mathbf{d}_{o-f}$ is the difference between the observation and forecast ensemble mean, and $\mathbf{P}^f$, $\mathbf{R}$ and $\mathbf{H}$ denote the forecast error covariance, the observational error covariance and the linear observational operator, respectively. The superscript $\top$, the subscript o and f denote the matrix transpose, observation and forecast, respectively. The spread inflation parameter is obtained from the covariance inflation parameter. In the third and fourth experiments (Ex. 3 and Ex. 4), not only the inflation parameter but also the observational errors are estimated adaptively. The method of estimating the observational error is shown in Desroziers et al. (2005), and the method for the covariance inflation parameter and the observational error is shown in Li et al. (2009a). In Desroziers et al. (2005) and Li et al. (2009a), an estimate of the observational error variance σ_o^2 can be obtained from

$$\sigma_o^2 = trace(\mathbf{d}_{o-a}\mathbf{d}_{o-f}^{\top})/p, \tag{4.2}$$

where $\mathbf{d}_{o-a}$ is the difference between the observation and analysis ensemble mean, and p and subscript a denote the number of observations and analysis, respectively. In this study, the estimation method is implemented for each observation variable. Li et al. (2009a) estimated the inflation parameter and the observational error variance adaptively at each analysis time step. However, if the number of observations is not large enough, a large sampling error is introduced. Therefore, Miyoshi (2005) and Li et al. (2009a) assumed that Δ_o and σ_o^2 were observed, respectively, to avoid this problem. They used a simple scalar KF approach, which usually uses the postprocess model output. By using this KF approach, the past information is accumulated, and the inflation parameter and observational error variance gradually converge to the optimum values while still allowing for time variations. The KF estimation is often ruined by an unrealistically large sampling error. So, to avoid this problem we impose reasonably wide upper and lower limits in the observed inflation Δ_o, e.g., 0.0 to 0.2, before applying the KF approach. In Ex. 3 and Ex. 4, the initial specifications of the observational errors are 3.0 and 0.1 times the true value, respectively, for each variable. At each analysis time, we evaluate the analysis error using the RMSE between the true state and the analysis ensemble mean and compute the ensemble spread.

Table 4.1. Configuration of each experiment.

	inf. param.	obs. error (init.)
Ex. 1	Constant (1–3 %)	True (–)
Ex. 2	Adaptive	True (–)
Ex. 3	Adaptive	Adaptive (3.0 times)
Ex. 4	Adaptive	Adaptive (0.1 times)

4.2 Results

Figure 4.1 shows the time series of the analysis RMSEs and ensemble spreads of 500 hPa geopotential height and 850 hPa temperature for Ex. 2. In the early period the RMSEs decrease with time. The RMSEs are larger than the ensemble spreads. After about three days spin-up, the RMSEs become comparable to the ensemble spreads. The RMSE and ensemble spread in the Tropics are larger than the other regions. This result shows that the uncertainty in the Tropics is large in spite of the perfect model experiments because the cumulus convection is active. In the temperature field, the RMSE and ensemble spread in the Southern Hemisphere are larger than that in the Northern Hemisphere. Such a tendency dominates in the lower troposphere, and there is no clear difference between the Southern Hemisphere and the Northern Hemisphere in the middle troposphere and above. This result implies that the cause is in the land-ocean distribution because the RMSE and ensemble spread are larger over the ocean, particularly in the temperature field. Moreover, for the other elements, e.g., SLP (Sea Level Pressure), wind components and water

vapor, the RMSEs are apparently smaller than the observational errors. For example, the analysis RMSEs of SLP in the Northern Hemisphere, Southern Hemisphere, and Tropics are about 0.2 hPa, 0.2 hPa and 0.3 hPa, respectively.

To see the horizontal distribution of the analysis errors, Fig. 4.2 illustrates the temporally averaged analysis RMSE and ensemble spread of 500 hPa geopotential height for Ex. 2, in which the inflation parameter is estimated and the observational error is perfectly known. The shaded areas show the analysis RMSE or ensemble spread, and the contours show 500 hPa geopotential height. As shown in both fields in Fig. 4.2, the analysis error distribution is comparable to that of the ensemble spread very much. Particularly the peaks of the RMSE correspond to those of the ensemble spread. The pattern represents the area which indicates large uncertainty. For example, in the Tropics the RMSE is very large along the ITCZ (The Intertropical Convergence Zone) because of the active cumulus convection. Therefore, the ensemble spread becomes large by the chaotic nature which originates by the uncertainty. The result is consistent with that in Fig. 4.1. On the other hand, over the east part of the North Pacific Ocean the westerly jet is meandering, and there is a ridge along the West Coast. In the upstream of the ridge, the RMSE is large, where an extratropical cyclone is developing in the surface. Moreover, in the area with large RMSE, the ensemble spread is also large as the RMSE. Therefore, it is confirmed that the NICAM-LETKF captures the characteristics of the regional analysis errors.

Figure 4.3 shows the time series of the analysis RMSEs of the 500 hPa zonal wind in Exs. 1, 2, 3, and 4. In the early stage the error level of Ex. 2 is the smallest.

Ten days after the beginning of the assimilation, however, a clear separation of the analysis RMSEs among the experiments is not seen. In the Ex. 4 the long adjusting period is necessary for the LETKF to converge. About ten days are needed so that the Ex. 4 is adjusted by the NICAM-LETKF, but eventually the RMSE in the Ex.4 becomes comparable to other experiments. In the other variables, similar results are obtained (not shown).

Figure 4.4 shows the adaptively estimated observational errors of pressure in the lowest layer of the NICAM in Exs. 3 and 4. The experiments start from the wrong observational error with 3.0 (Ex. 3) and 0.1 (Ex. 4) times the true value. In Ex. 3 the estimated observational error gradually decreases by assimilating data iteratively, and then the error converges to the true value. In Ex. 4 the estimated error becomes larger than the initial value in the early period. Then, the error converges to the true value such as in the case of the Ex. 3. About 18 days after the beginning of the assimilation, both of the estimated observational errors already become very close to the true value though the observational errors are not known. The estimated observational errors are slightly larger than the truth even though the initial observational error is very small in the case of the Ex. 4. Hence, it is confirmed that the algorithm which estimates the observational error adaptively works very well. In the other variables, similar results are obtained (not shown). As a result, we can get a good analysis by estimating the observational errors adaptively though the errors are not known.

According to Figs. 4.3 and 4.4, it seems that the NICAM-LETKF has small sensitivity to the observational errors because in the early assimilation period the

RMSE of the Ex. 3 is comparable to the RMSE of the Ex. 2 although the estimated observational errors are overestimated. This result is discussed in chapter VI.

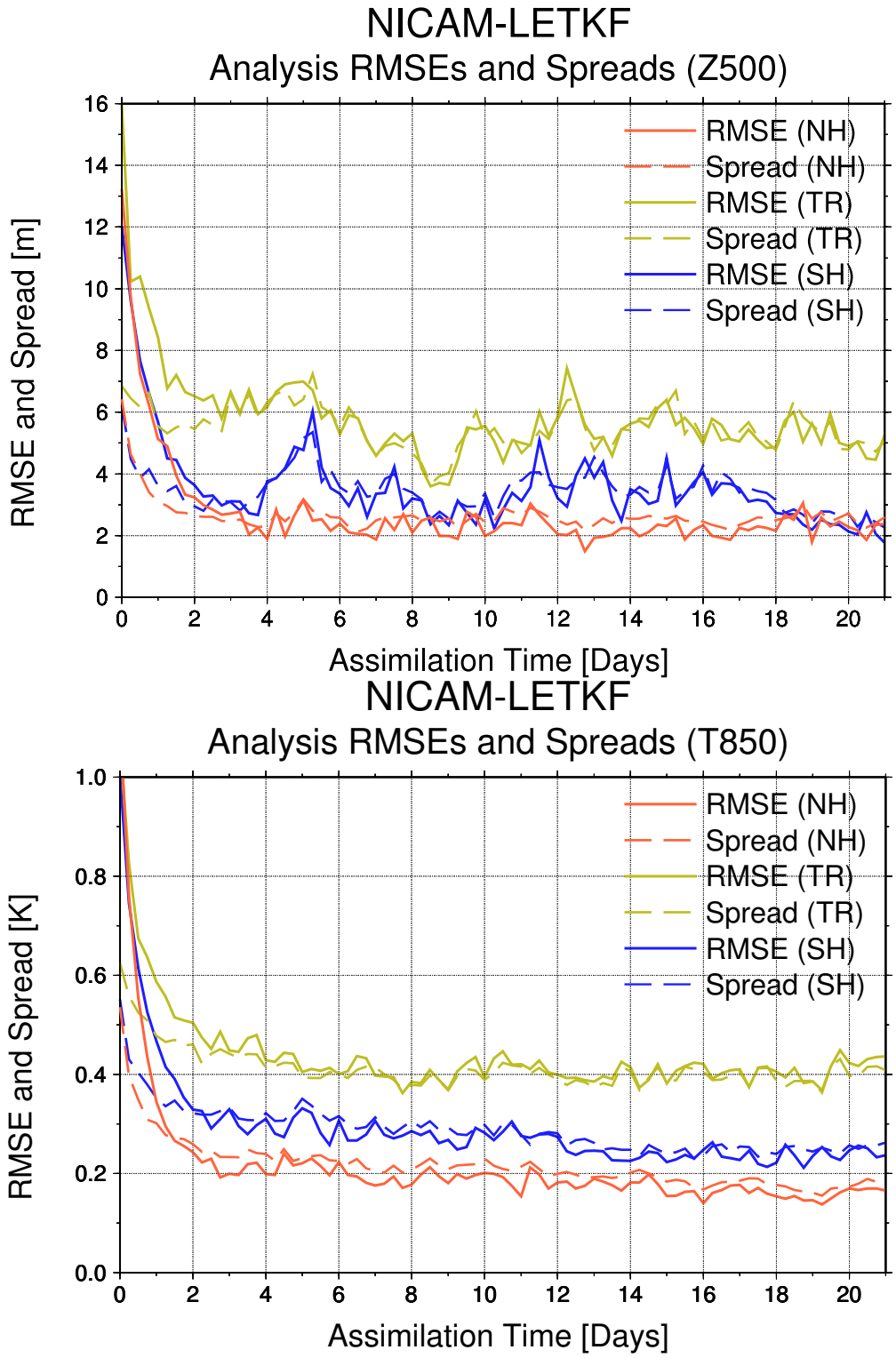


Figure 4.1. Time series of the analysis RMSEs and ensemble spreads of 500 hPa geopotential height (m) (top panel) and 850 hPa temperature (K) (bottom panel) for the Ex. 2. Initial time is 12Z 1 Jan 2007. The red, blue, and yellow lines are for the Northern Hemisphere (NH; 20°N–90°N), the Southern Hemisphere (SH; 20°S–90°S), and Tropics (TR; 20°N–20°S), respectively.

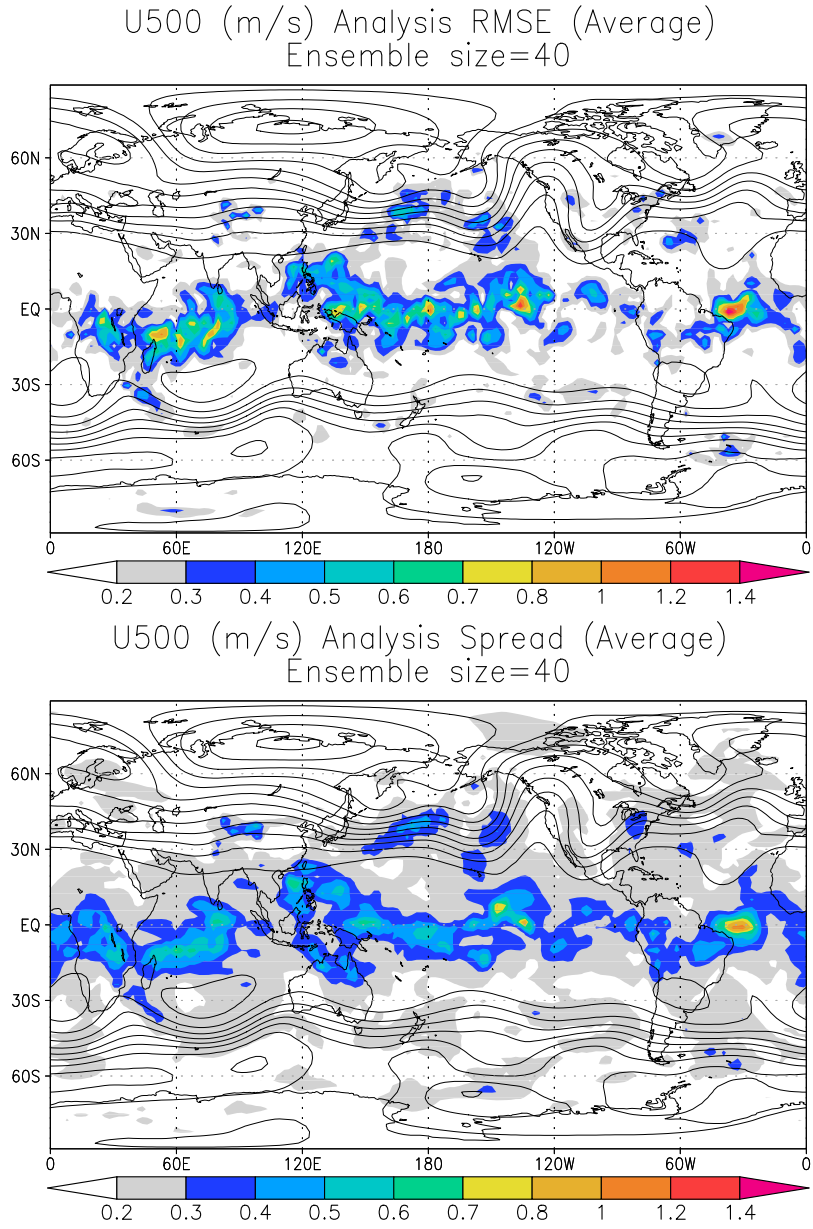


Figure 4.2. Spatial distributions of the analysis RMSEs (top panel) and ensemble spreads (bottom panel) of 500 hPa zonal wind and 500 hPa geopotential height (contour) for the Ex. 2, temporally averaged for 1 day from 00Z 21 January 2007 to 18Z 21 January 2007.

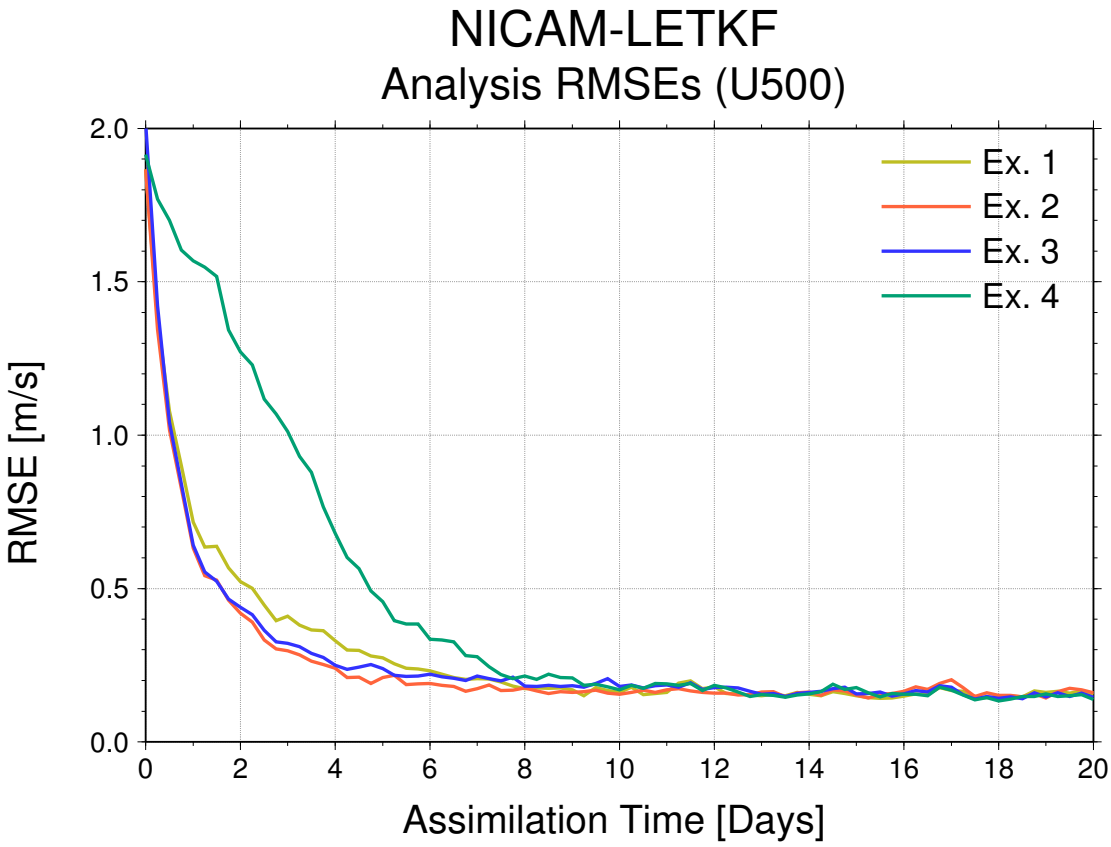


Figure 4.3. Time series of the analysis RMSEs of 500 hPa zonal wind (m/s) for the Exs. 1, 2, 3, and 4. The yellow, red, blue, and green lines are for the Exs. 1, 2, 3, and 4, respectively. Initial time is 12Z 1 Jan 2007.

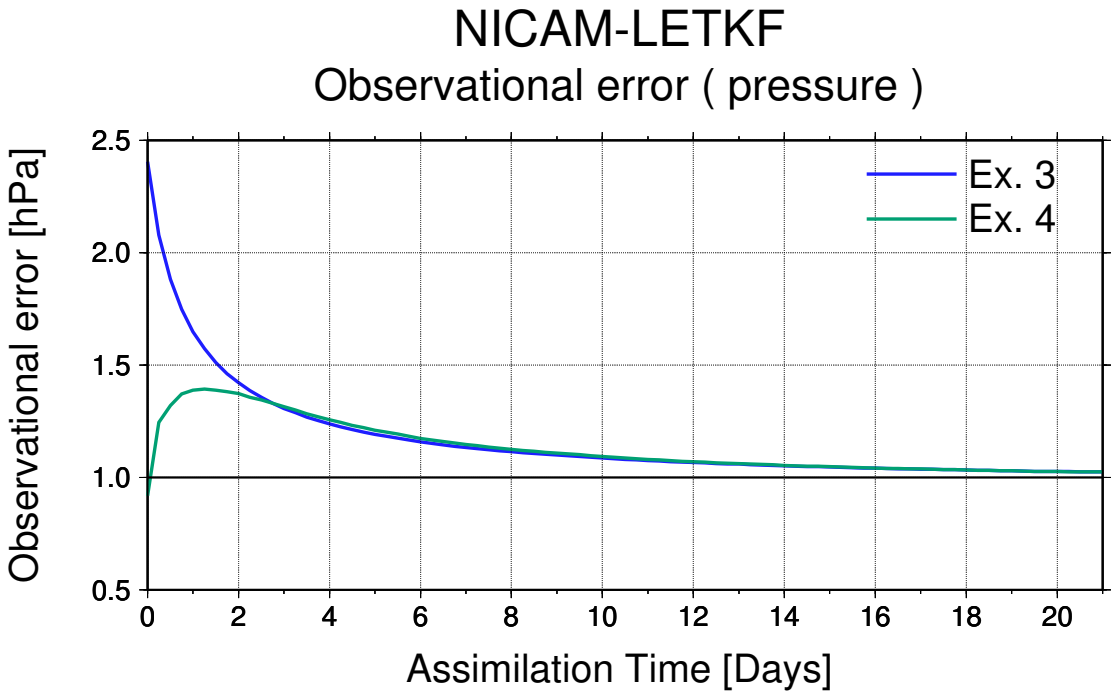


Figure 4.4. Time series of adaptively estimated observational error of the pressure at the lowest layer of the NICAM. The blue and green lines are for the Exs.3 and 4, respectively. The true observational error is 1.0 hPa. Initial time is 12Z 1 Jan 2007.

CHAPTER V

Multi-Scale Localization Approach

5.1 Dual-Localization Approach

We may limit the influence of observations within a narrower region than the range that the observations should impact, particularly when the model resolution is very high, since larger-scale structures than the localization scale are removed due to tight localization for the high-resolution model. The optimal localization scale depends on several factors including the model resolution and ensemble size; the localization scale becomes generally smaller for higher-resolution models. Then, Miyoshi and Kondo (2013) and Kondo et al. (2013) introduced the multi-localization approach, and they mainly investigated the dual-localization approach. The essence of the dual-localization approach is to find the analysis increments as a sum of the small-scale and large-scale components. The full-resolution ensemble perturbations and the reduced-resolution ensemble perturbations obtained by smoothing the full-resolution perturbations are combined with small-scale and large-scale localizations, respectively.

Figure 5.1 illustrates the raw increments computed from 20 ensemble members using the LETKF (a) with T30 SPEEDY model and (b) with T21 SPEEDY model,

assimilating the single-point observation at the star point. Figure 5.1(a) shows the full-resolution increments and 5.1(b) shows the reduced-resolution increments. There are some strong signals in the distance from the star point, which means that the long-range error correlations from the star point at the center. The long-range error correlations are likely to be sampling noise. Therefore, a distance-dependent localization function is usually applied, and then the localized results are illustrated in Fig. 5.2 with a 1000-km localization scale. These analysis increments of (a) and (b) are computed from the single-point observation only in the surrounding area. It is considered that they are more realistic analysis increments than Fig. 5.1. Figure 5.2(a) is noisier structure, however, in Fig. 5.2(b) the high-frequency sampling error is filtered out, and Fig. 5.2(b) will allow using broader localization without contamination.

Figure 5.3 illustrates the same analysis increments as Fig. 5.2, but with a 500-km localization scale. This includes only shorter-range covariances. Figure 5.3(a) has the finer structure near the observed location with higher-frequency components, but Fig. 5.3(b) does not keep the finer structure.

This dual-localization approach is to merge the fine short-range structure (Fig. 5.3(a)) and broad long-range structure (Fig. 5.2(b)). The weighted mean of Figs. 5.3(a) and Fig. 5.2(b) will reduce the fine structure and the long-range signals. In this study, we consider an alternative approach which uses the short-range structure of 5.3(a) and the only long-range component of 5.2(b). Then the short-range component is taken out from 5.2(b). This component is obtained by subtracting the increment with the 500-km localization with the smoothing perturbations (Fig. 5.3(b)) from

the increment with the 1000-km localization using the smoothing perturbations (Fig. 5.2(b)), which gives the only long-range components (Fig. 5.4). Finally, by adding short-range component (Fig. 5.3(a)) and the long-range component (Fig. 5.4), the merged dual-scale increment (Fig. 5.5) is obtained. This contains the fine structure near the observation and has the longer-range structure without high-frequency noise.

In this paper this dual-localization approach is derived based on the LETKF, but the same idea may be applied to other EnKF implementations in a straightforward manner. And this dual-localization approach formula is described as follow.

The smaller-scale, high-resolution component $\delta\mathbf{X}_h$ is obtained by the full-resolution forecast ensemble perturbations with a short-range localization scale (500-km localization, Fig. 5.3(a)). The subscript h represents using high-resolution perturbations. In the LETKF, $\delta\mathbf{X}_h$ is given by

$$\delta\mathbf{X}_h = \delta\mathbf{X}^f\tilde{\mathbf{P}}^a\left(\mathbf{H}\delta\mathbf{X}^f\right)^\top\left(\tilde{\boldsymbol{\rho}}_{short}\circ\mathbf{R}^{-1}\right)\left(\mathbf{y}^o-\mathbf{H}\bar{\mathbf{x}}^f\right)+\delta\mathbf{X}^f\left[(N-1)\tilde{\mathbf{P}}^a\right]^{1/2}, \quad (5.1)$$

where $\circ$ denotes the element-wise multiplication. Each column of the background ensemble perturbation matrix $\delta\mathbf{X}^f$ is composed of the difference between each ensemble forecast and the ensemble mean $\bar{\mathbf{x}}^f$, and the ensemble size is N. $\tilde{\mathbf{P}}^a$, $\mathbf{H}$, $\mathbf{R}$ and $\mathbf{y}^o$ denote the analysis error covariance in ensemble space, the linear tangent matrix of the observation operator, observation error covariance matrix (assumed to be diagonal) and observation vector, respectively. In the LETKF we apply observation localization $\tilde{\rho}_{short}$ to $\mathbf{R}^{-1}$ and weigh the observation error variances depending on the distance from the analyzed grid point (Hunt et al. 2007; Miyoshi and Yamane 2007;

Greybush et al. 2011).

The larger-scale, low-resolution component $\delta\mathbf{X}_l$ is computed from reduced-resolution forecast ensemble perturbations with a longer-range localization scale. $\delta\mathbf{X}_l$ is composed of $\delta\mathbf{X}_{l-long}$ and $\delta\mathbf{X}_{l-short}$. The subscript l represents low-resolution perturbations, followed by *long* (*short*) denoting a long-range (short-range) localization scale. $\delta\mathbf{X}_{l-long}$ (Fig. 5.2(b)) is obtained from $\delta\mathbf{X}_l^f$ with longer-range localization:

$$\delta\mathbf{X}_{l-long} = \delta\mathbf{X}_l^f \tilde{\mathbf{P}}_{l-long}^a \left(\mathbf{H}\delta\mathbf{X}_l^f\right)^\top \left(\tilde{\boldsymbol{\rho}}_{long} \circ \mathbf{R}^{-1}\right) \left(\mathbf{y}^o - \mathbf{H}\bar{\mathbf{x}}^f\right) + \delta\mathbf{X}_l^f \left[(N-1)\tilde{\mathbf{P}}_{l-long}^a\right]^{1/2}, \tag{5.2}$$

where the localization scale of $\tilde{\boldsymbol{\rho}}_{long}$ is longer than that of $\tilde{\boldsymbol{\rho}}_{short}$. Although $\delta\mathbf{X}_l^f$ is obtained by applying a spatial low-pass filter (i.e., spatial smoothing) to $\delta\mathbf{X}^f$, $\bar{\mathbf{x}}^f$ is not smoothed and is the same as that in Eq. (5.1). $\delta\mathbf{X}_{l-long}$ includes a short-range structure $\delta\mathbf{X}_{l-short}$, which needs to be removed. $\delta\mathbf{X}_{l-short}$ (Fig. 5.3(b)) is obtained from $\delta\mathbf{X}_l^f$ with smaller localization scale:

$$\delta\mathbf{X}_{l-short} = \delta\mathbf{X}_l^f \tilde{\mathbf{P}}_{l-short}^a \left(\mathbf{H}\delta\mathbf{X}_l^f\right)^\top \left(\tilde{\boldsymbol{\rho}}_{short} \circ \mathbf{R}^{-1}\right) \left(\mathbf{y}^o - \mathbf{H}\bar{\mathbf{x}}^f\right) + \delta\mathbf{X}_l^f \left[(N-1)\tilde{\mathbf{P}}_{l-short}^a\right]^{1/2}, \tag{5.3}$$

where $\tilde{\boldsymbol{\rho}}_{short}$ is the same as that in Eq. (5.1).

The resulting multi-scale increment $\delta\mathbf{X}$ is obtained from Eqs. (5.1), (5.2) and (5.3) and is shown in (Fig. 5.5). $\delta\mathbf{X}$ is computed by

$$\delta\mathbf{X} = \delta\mathbf{X}_h + \delta\mathbf{X}_{l-long} - \delta\mathbf{X}_{l-short}. \tag{5.4}$$

$\delta\mathbf{X}_h$ (Fig. 5.3(a)) provides the short-range, small-scale structure, and $\delta\mathbf{X}_l = \delta\mathbf{X}_{l-long} -$

$\delta\mathbf{X}_{l-short}$ (Fig. 5.4) provides the long-range, large-scale structure. If the long and short localization scales are the same, the dual-localization approach reduces to the single localization approach. Although the dual-localization approach requires solving the LETKF analysis updates three times, usually that is not a large load since the major part of LETKF computations is usually in the ensemble forecasting part, which remains the same.

Miyoshi and Kondo (2013) obtained the smoothed ensemble perturbations $\delta\mathbf{X}_l^f$ by applying the spectral truncation in the spherical harmonics. In this study, we apply the Lanczos filter (Lanczos 1956; Duchon 1979). Lanczos weight is given by

$$\text{Lanczos weight}(d) = \begin{cases} 1 & (d = 1) \\ \dfrac{\sin(f_c d)}{\pi d} \dfrac{\sin(\pi d/n)}{\pi d/n} & (0 < |d| < n) \\ 0 & (otherwise), \end{cases} \tag{5.5}$$

where d describes the distance from the center grid point. The Lanczos filter has two parameters: resolution parameter n and critical frequency f_c. In this study, n is chosen to be 10, and f_c is selected from 1/5, 1/8 and 1/11. The Lanczos filter is applied to the longitude and latitude directions. Figure 5.6 shows the response functions of the Lanczos filter and Fourier transform. The curve of Fourier transform is an ideal response function with cut-off wave number 21. The three curves of Lanczos filters are the smoothed response function with transition bands. The smaller f_c corresponds to more smoothing. The traditional single localization approach and the dual-localization approach are called "CTRL" and "DLOC", respectively, and sometimes subscript F or L is added, corresponding to the Fourier transform (F) and

Lanczos filter (L), respectively. Figure 5.7 illustrates similar analysis increments as Fig. 5.5, but with the Lanczos filter. If we choose a stronger smoothing parameter, $\delta\mathbf{X}_l$ becomes smaller and smaller by cancelling out positive and negative values, and the analysis increments become similar to those with single 500-km localization $\delta\mathbf{X}_h$ (Figs. 5.3(a) and 5.7(c)). This study compares the analysis accuracy among different smoothing functions.

5.2 Experimental Settings

In this study, twin experiments with a T30/L7 (horizontal resolution up to 30 wave numbers and 7 vertical levels) atmospheric general circulation model (AGCM) known as the SPEEDY model (Molteni 2003) are performed under the perfect model scenario. This type of experiments is also known as the OSSE (observing system simulation experiment). The horizontal resolution of the SPEEDY model is about 420 km (3.75 degrees in longitude), and the prognostic variables are zonal and meridional wind components (U, V), temperature (T), specific humidity (Q) and surface pressure (Ps). The experimental settings follow Miyoshi and Kondo (2013) and Kondo et al. (2013). Namely, the nature run starts at 0000 UTC 1 January 1982, and the observations are generated by adding uncorrelated white random numbers to the nature run. The observation error standard deviations are 1.0 m s^{-1} for U and V, 1.0 K for T, 1.0 $\times 10^{-3}$ kg kg^{-1} for Q and 1.0 hPa for Ps. The observations are taken every 6 hours at all 7 vertical levels at given horizontal stations of a radiosonde-like network, but

the observations of specific humidity and surface pressure are taken from the bottom to the 4th level and only at the surface, respectively. The assimilation cycle is every 6 hour, and the period of the experiments is from 0000 UTC 1 January 1982 to 0000 UTC 1 February 1983. The ensemble size is fixed at 20, and the initial ensemble members are chosen from the nature run in January 1984. All experiments are performed with adaptive covariance inflation (Miyoshi 2011). The CTRL experiment employs the traditional LETKF with 700-km horizontal and 0.1 ln p vertical localization parameters. The other experiments employ the dual-localization approach (DLOC) with different choices of smoothing functions and localization parameters. DLOC_F uses the spectral truncation at 21 wavenumbers, and DLOC_L5 uses Lanczos filter with fc = 1/5, and similarly to DLOC_L8 and DLOC_L11. The localization parameters are chosen by 100-km increment from 300 km to 900 km for the short localization, and from 600 km to 1300 km for the long localization. By definition, the long localization parameter is always greater than the short localization parameter. In order to investigate a statistical significance, four parallel experiments are implemented using different observational noises with the same observation error standard deviations.

5.3 Results

Figure 5.8 illustrates the results of the localization parameter survey for different choices of the smoothing functions. Figure 5.8 indicates RMSE (Root Mean Square

Error) of surface pressure, and the shaded area corresponds to the advantage of DLOC over CTRL. DLOC_F and DLOC_L5 are very similar (Figs. 5.8 (a) and (b)), and the smallest RMSE values are almost identical. Figures 5.8 (b), (c) and (d) show the impact of the smoothing strength. As the degree of smoothing becomes stronger, the optimal localization parameters shifts to larger scales, and the shaded area becomes smaller. The smallest RMSE in Fig. 5.8 (c) is almost the same as that of Fig. 5.8 (b), but Fig. 5.8 (d) shows larger RMSE. Hereafter, DLOC_L5 is further investigated and is denoted simply as DLOC_L.

Figure 5.9 is similar to Fig. 5.8, but for zonal wind, temperature and specific humidity. DLOC_L outperforms CTRL with relatively wide choices of the two localization parameters by about 400-km range, that is, 400-900 km for the short localization parameter and 700-1100 km for the long localization parameter. Using the longer localization parameter greater than 1200 km causes filter divergence. U and Ps (Fig. 5.8 (b)) show a little more sensitivity to the localization parameters than T and Q. The shapes of shaded area are similar except for Q. The positions of minimum RMSE differ among different variables. Figures 5.8 and 5.9 suggest that the best combination of short and long localization parameters be 600 km and 900 km. A similar localization parameter survey is performed for DLOC_F, and the optimal localization parameters are found to be 500 km and 900 km (not shown).

Figure 5.10 shows time series of analysis RMSEs. It is confirmed that overall advantage of dual localization (DLOC) over the manually-tuned single localization (CTRL). After about 4-month run (May 1982 and later), DLOC shows very clear

advantage, particularly for humidity. Figure 5.11 shows the spatial distribution of analysis RMSEs and spreads of CTRL and DLOC for the surface pressure (Figs. 5.11 (a) and (b)) and specific humidity (Figs. 5.11 (c) and (d)). The analysis RMSE patterns correspond to the spreads, and the RMSE of DLOC is smaller than that of CTRL in the most areas, so it shows that the DLOC is worked appropriately. In Fig. 5.11 (c; CTRL), the analysis spread is smaller than the analysis RMSE in the some areas, especially over the Northern America. On the other hands, in Fig. 5.11 (d; DLOC), the analaysis spread corresponds to the analysis RMSE. So, the analysis spread of CTRL is underestimated, and it is indecated that DLOC can estimate the analysis spread and that the analysis spread of DLOC corresponds to the analysis RMSE in specific humidity.

Figure 5.12 shows the horizontal power spectrums of the analysis error of zonal wind at the 4th model level ($\sim$500 hPa). Here, the optimal localization parameters are chosen. The analysis errors are computed from the difference between the nature run and analysis, and have the most power around the horizontal wave numbers 10 to 20. By applying the dual-localization approach, the analysis errors become much smaller in all wavenumbers, and the improvement was about 20 %. There is no significant difference between DLOC_F and DLOC_L.

Figures 5.13 and 5.14 illustrate the improvements for horizontal and vertical structure, respectively, and they greys out non-significant areas in which any one of the four parallel experiments disagrees. In general, DLOC has advantages not only over the ocean but also over the land. Over the ocean, it is probably considered

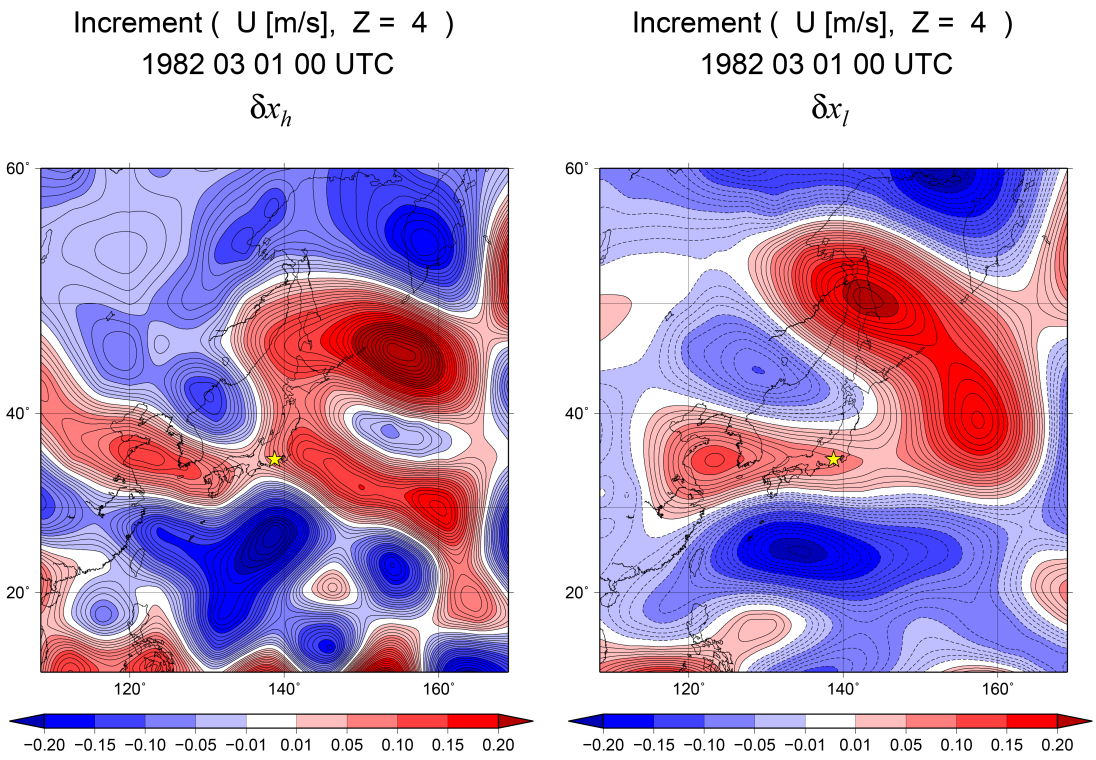
that the DLOC can catch some more distant signals of observations than CTRL. In fact, 900-km single localization gives more accurate analysis over the ocean than CTRL, however, the 900-km single localization is still worse than DLOC. On the other hands, over the land, 900-km localization gives less accurate analysis than CTRL, so the overall analysis accuracy of 900-km localization is not better than CTRL.

To investigate the covariance structures of the dual-localization approach, a 3200 ensemble member experiment is also implemented. Figure 5.15 illustrates the horizontal background error covariance maps of T at the 1st level. In the case of 20 ensemble members, the covariance has a lot of sampling errors even if using a 1000-km localization (Fig. 5.15(a)). By increasing the ensemble members up to 3200, the sampling error is much smaller than 20 ensemble members (Figs. Fig. 5.15(b)). In dual-localization approach with 20 ensemble members and 500-km and 1000-km localizations, the covariance structure is more similar to 3200 ensemble members than 20 ensemble members with 1000-km localization (Fig. 5.15(c)).

Figure 5.16 illustrates the correlation maps of the background error for Q at the 1st level, using 20, 50, 100, 200, 400, 800, 1600 and 3200 ensemble members. The sampling errors dominate and it is impossible to distinguish some important signals and the sampling errors with 20 ensemble members. However, by increasing the ensemble members up to 3200, the sampling errors are approximately deleted, and it is confirmed to capture some important signals. The signals have wavy patterns and extend in a westward direction. Figure 5.17 is similar to Fig. 5.16, but for Q using 20 and 3200 ensemble members. The signals do not have a wavy pattern, and

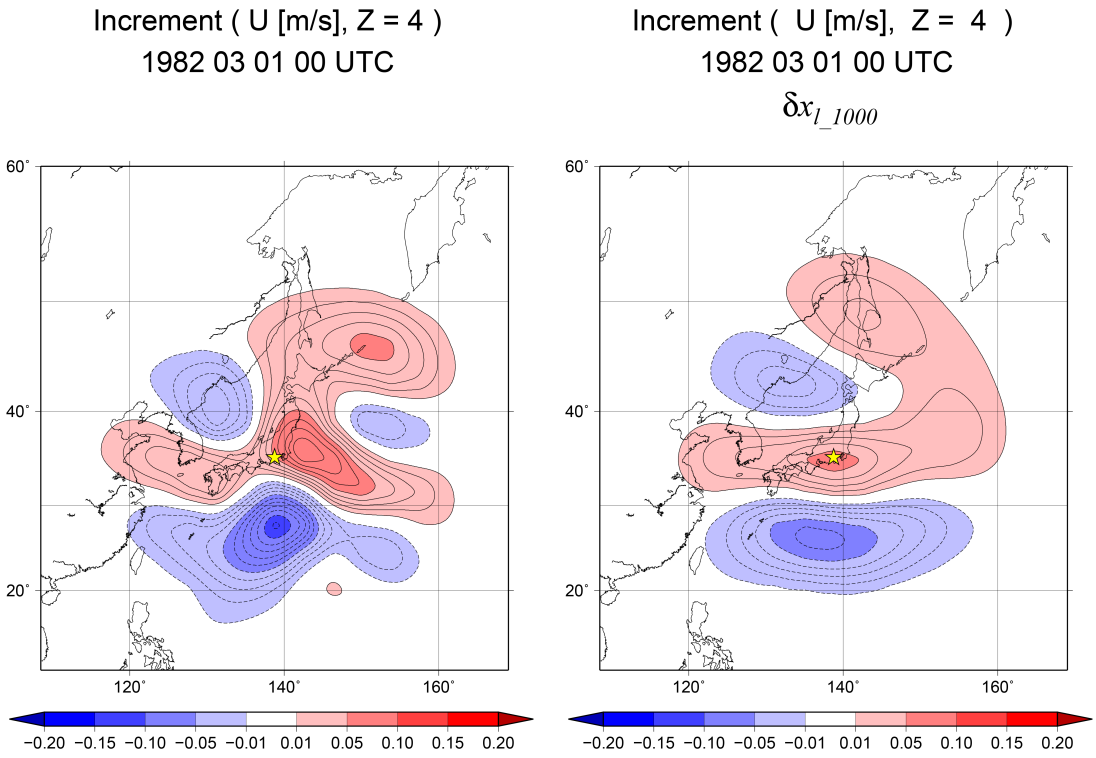
the direction is eastward. And the correlation structure is similar to PNA pattern (Pacific/North American pattern; Wallace and Gutzler, 1981). Moreover, the scales of signals are much larger than 1000-km localization scale.

Finally, the parallelization and computational time are mentioned. These techniques are implemented in K computer. The acceleration rations of LETKF are plotted in Fig. 5.18. If the parallelization ratio is 100 %, it is indicated a perfect acceleration (black solid line). The 400 member LETKF is more efficient than the 99.80 % parallelization ratio, and 1600 and 3200 members LETKFs outperform the 99.99 % parallelization ratio. Tables 5.1, 5.2, 5.3 and 5.4 show the computation times of LETKF by 12 nodes, 24 nodes, 48 nodes and 96 nodes, respectively. The assimilation occupies most of the total computational time, and the eigenvalue decomposition of analysis error covariance matrix (Eq. 2.25) occupies from 65 % to 85 % of the computational time of assimilation (not shown). The computational time of assimilation is proportional to about 7 times of the ensemble members and is proportional to 2 times of the number of nodes.



(a) High-resolution (T30) (b) smoothed-resolution (T21)

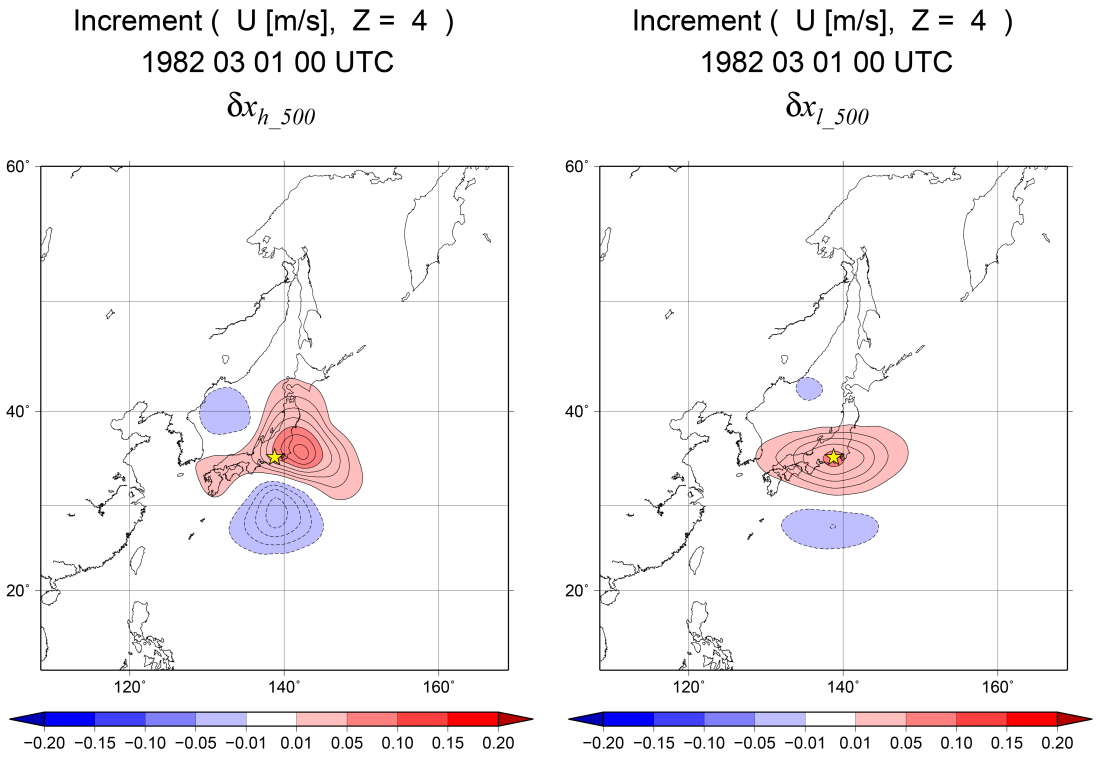
Figure 5.1. Analysis increments of zonal wind (m s^{-1}) at the 4th model level ($\sim$ 500 hPa) from a single profile observation at the star point using the SPEEDY model using (a) full-resolution (T30) ensemble perturbations without localization, (b) the same as (a) but smoothed resolution (T21). (a) is adopted from Miyoshi and Kondo (2013).



(a) L = 1000 km with high-resolution (T30)

(b) L = 1000 km with smoothed-resolution (T21)

Figure 5.2. Analysis increments of zonal wind (m s^{-1}) at the 4th model level (∼500 hPa) from a single profile observation at the star point using the SPEEDY model using (a) full-resolution (T30) ensemble perturbations with 1000-km localization scale, (b) the same as (a) but smoothed resolution (T21). These are adopted from Miyoshi and Kondo (2013).



(a) L = 500 km with high-resolution (T30)

(b) L = 500 km with smoothed-resolution (T21)

Figure 5.3. Analysis increments of zonal wind (m s^{-1}) at the 4th model level ($\sim$ 500 hPa) from a single profile observation at the star point using the SPEEDY model using (a) full-resolution (T30) ensemble perturbations with 500-km localization scale, (b) the same as (a) but smoothed resolution (T21). These are adopted from Miyoshi and Kondo (2013).

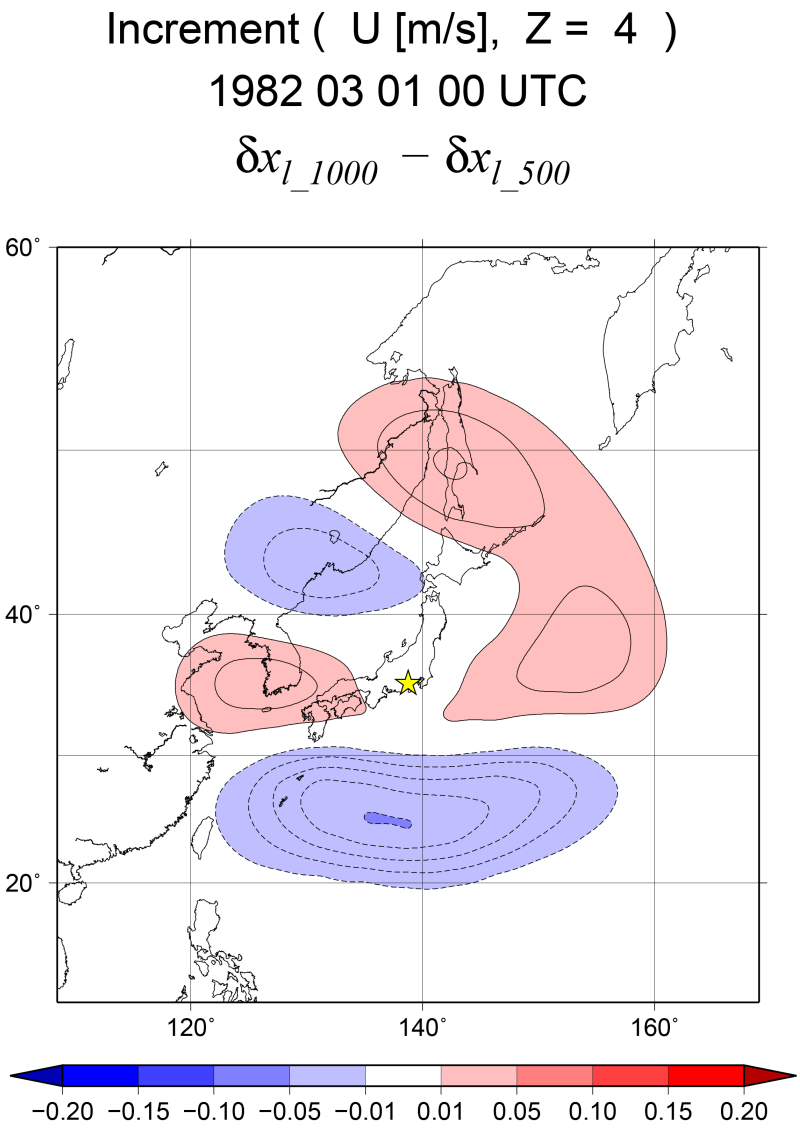


Figure 5.4. Difference of increment between Fig. 5.2(b) (large localization scale) and Fig. 5.3(b) (short localization scale). This is adopted from Miyoshi and Kondo (2013).

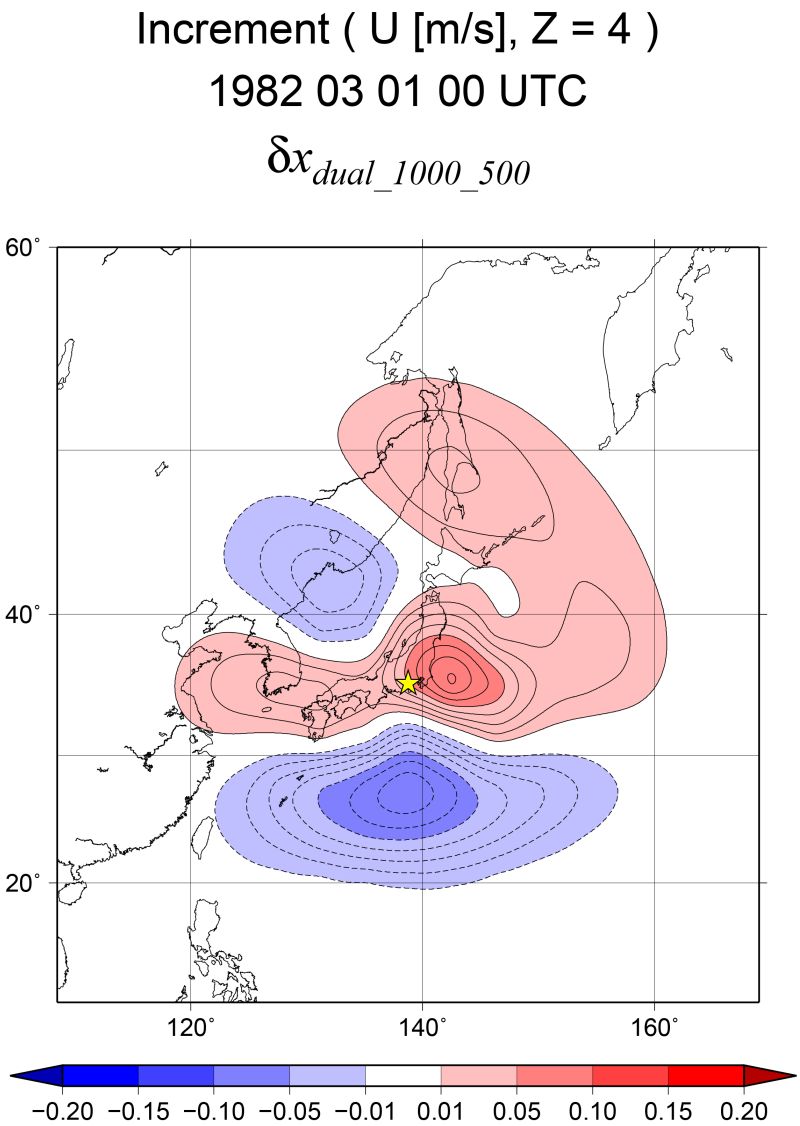


Figure 5.5. Dual-localization analysis increments of zonal wind (m s^{-1}) at the 4th model level (∼500 hPa) from a single profile observation at the star point using the SPEEDY model using . This is adopted from Miyoshi and Kondo (2013).

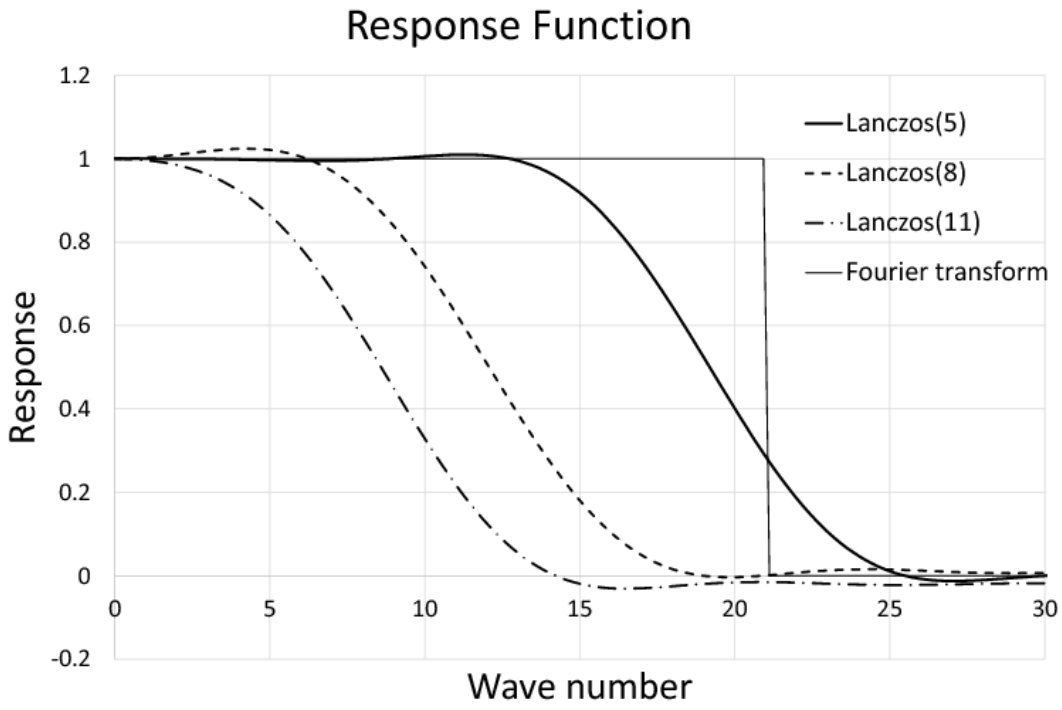
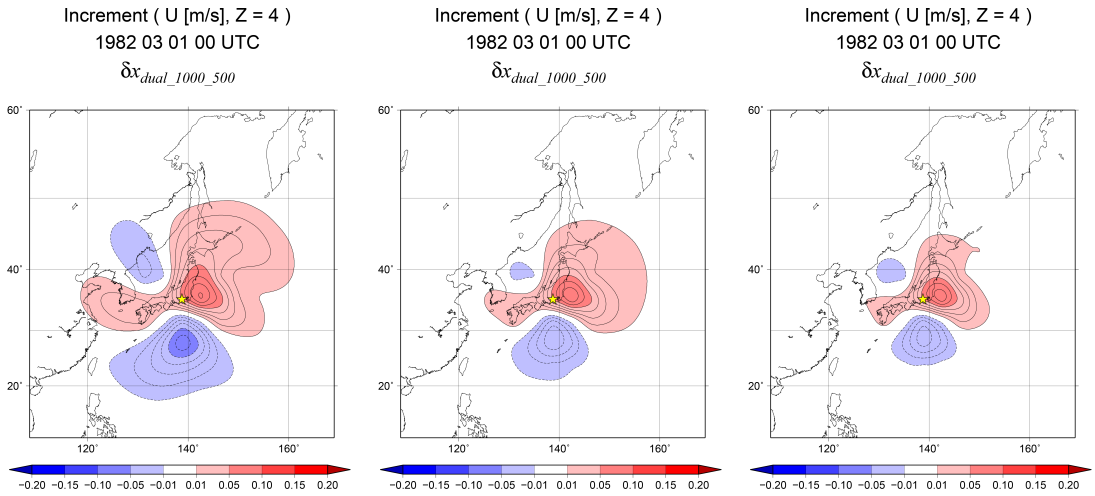
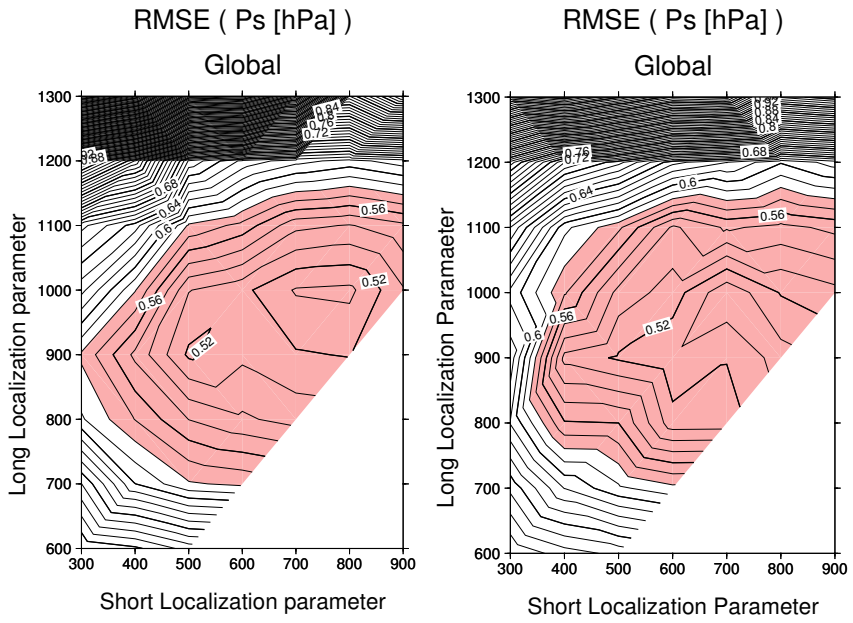


Figure 5.6. Response functions of the Fourier transform (thin solid) and the Lanczos filters with $f_c = 1/5$ (thick solid), 1/8 (dashed) and 1/11 (dash-dotted).



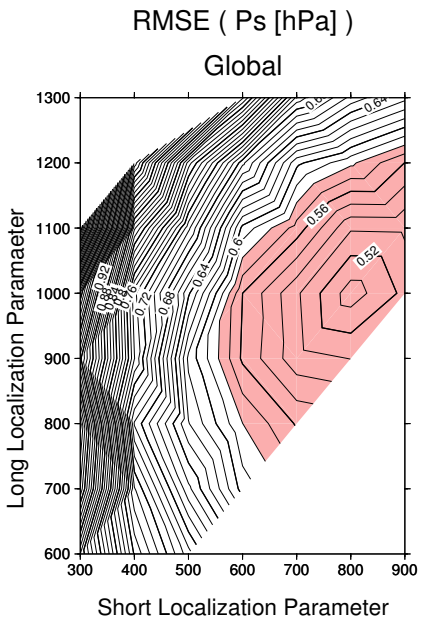
(a) $f_c = 1/5$ (b) $f_c = 1/8$ (c) $f_c = 1/11$

Figure 5.7. Similar to Fig. 5.5, but using the Lanczos filter with $f_c =$ (a) 1/5, (b) 1/8 and (c) 1/11.

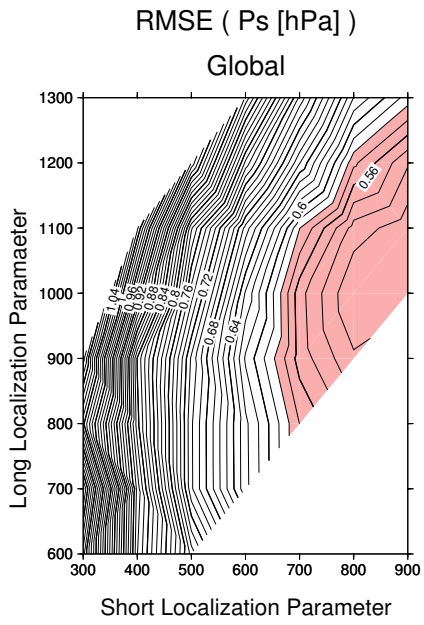


(a) DLOC_F

(b) DLOC_L5

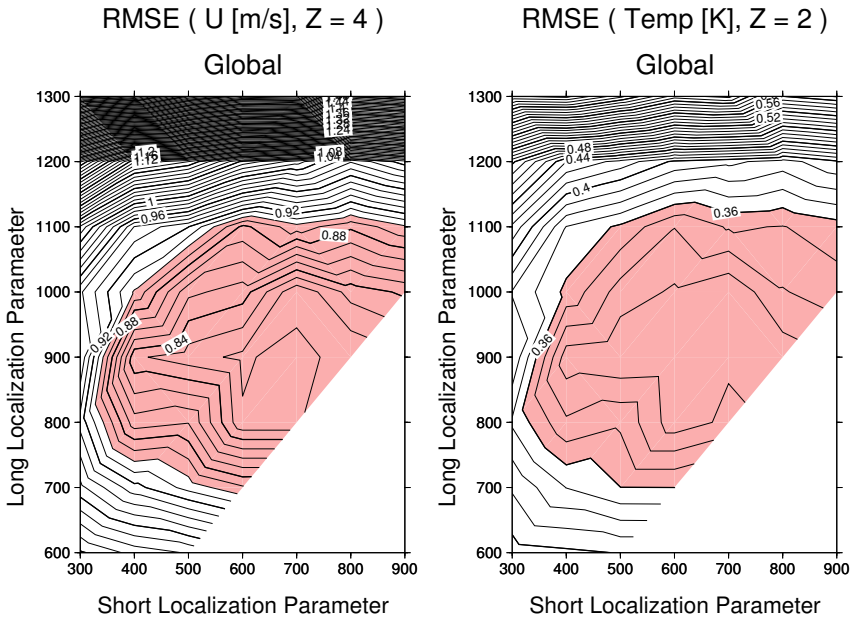


(c) DLOC_L8



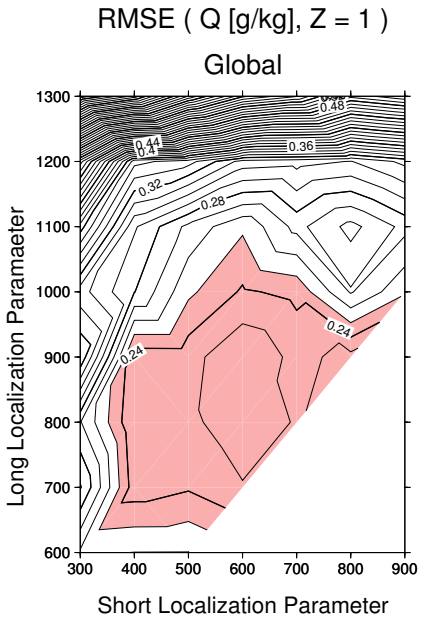
(d) DLOC_L11

Figure 5.8. Analysis RMSEs of surface pressure (hPa) of dual localization with various localization scale parameters, averaged for a year from 0000 UTC 1 February 1982 to 0000 UTC 1 February 1983, and using (a) T21 spectral truncation, (b) Lanczos filter with $f_c = 1/5$, (c) $f_c = 1/8$ and (d) $f_c = 1/11$. The shaded areas indicate improvements of DLOC over CTRL whose RMSE is 0.586 hPa.



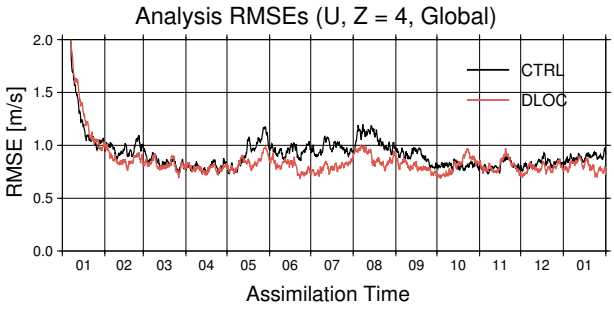
(a) Zonal wind at the 4th model level

(b) Temperature at the 2th model level

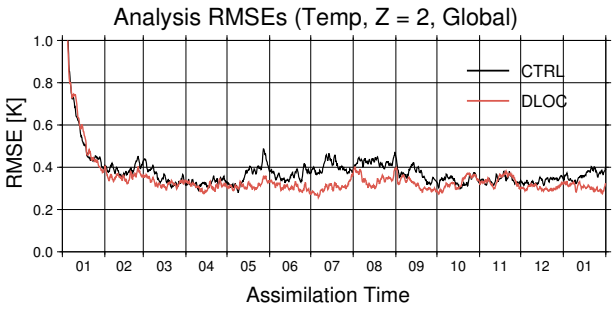


(c) Specific humidity at the lowest model level

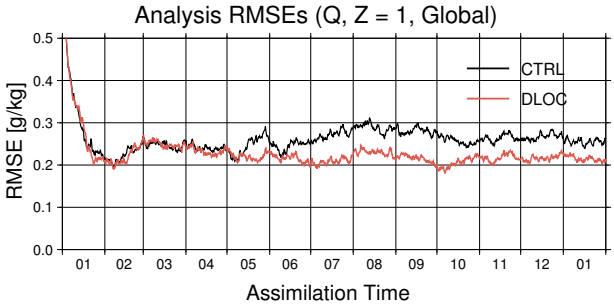
Figure 5.9. Similar to Fig. 5.8(b), but for (a) zonal wind (m s^{-1}) at the 4th model level, (b) temperature (K) at the second model level and (c) specific humidity (g kg^{-1}) at the lowest model level. The RMSEs of CTRL are 0.900 m s^{-1}, 0.366 K and 0.258 g kg^{-1}, respectively.



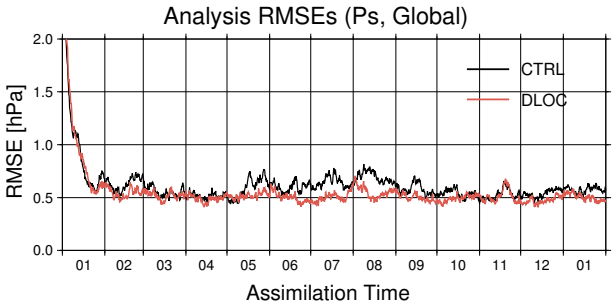
(a) Zonal wind (m s-1) at the 4th model level



(b) Temperature (K) at the 2nd model level

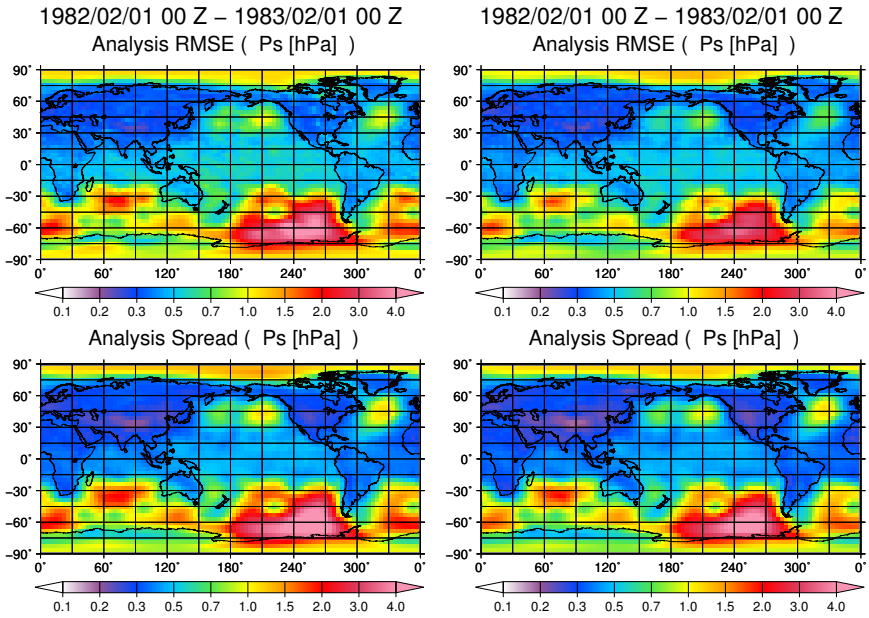


(c) Specific humidity (g kg-1) at the lowest level



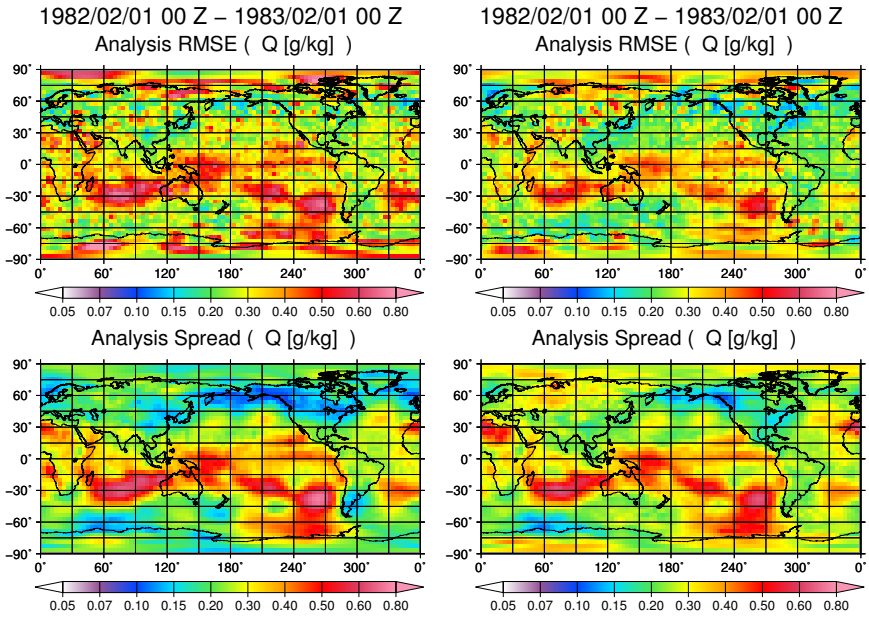
(d) Surface pressure (hPa)

Figure 5.10. Time series of analysis RMSE using Lanczos filter ($f_c = 1/5$) as a smoothing function for (a) zonal wind (m s^{-1}) at the 4th model level ($\sim$500 hPa), (b) temperature (K) at the 2nd model level ($\sim$850 hPa), (c) specific humidity (g kg^{-1}) at the lowest level ($\sim$925 hPa), and (d) surface pressure (hPa). Black and red curves indicate CTRL and DLOC experiments, respectively. Abscissa indicates the month.



(a) Surface pressure (hPa) of CTRL

(b) Surface pressure (hPa) of DLOC



(c) Specific humidity (g kg-1) at the lowest level of CTRL

(d) Specific humidity (g kg-1) at the lowest level of CTRL

Figure 5.11. Spatial distributions of 1-year-average analysis RMSE (top panel) and Spread (bottom panel) for (a) surface pressure of CTRL, (b) surface pressure of DLOC, (c) specific humidity of CTRL at the lowest level (∼925 hPa) and (d) specific humidity of DLOC at the lowest level (∼925 hPa).

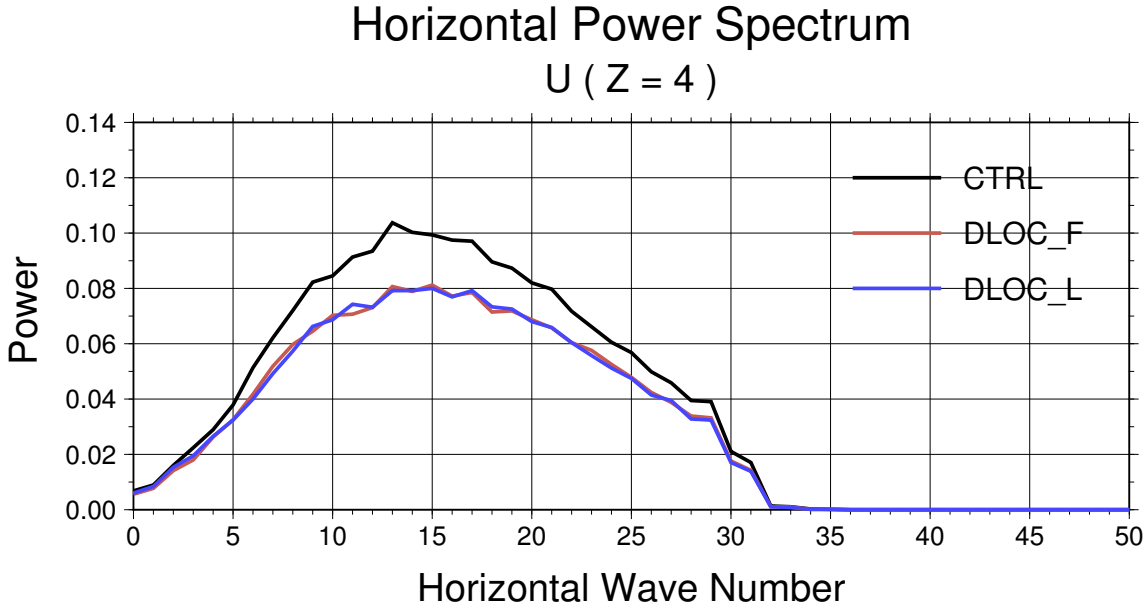
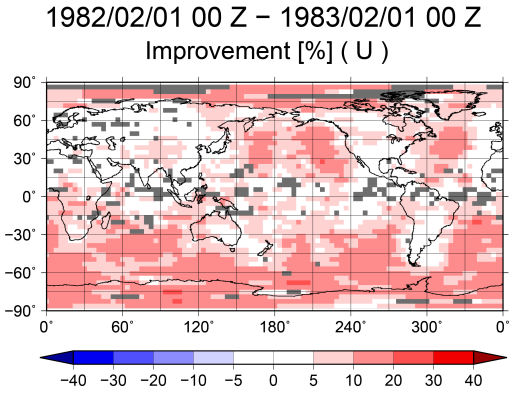
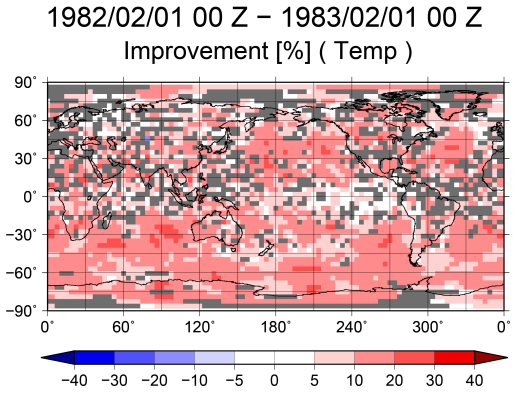


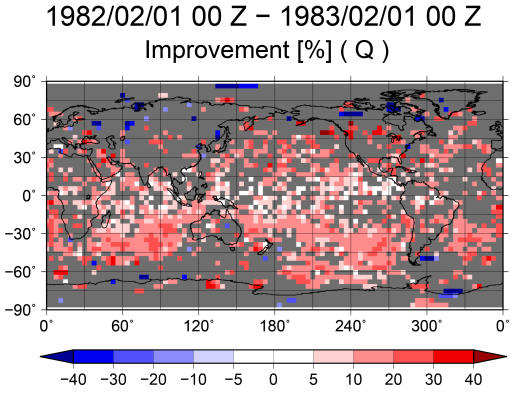
Figure 5.12. Horizontal power spectrum of the zonal wind analysis errors at the 4th model level for CTRL (black), DLOC_F (red) and DLOC_L (blue).



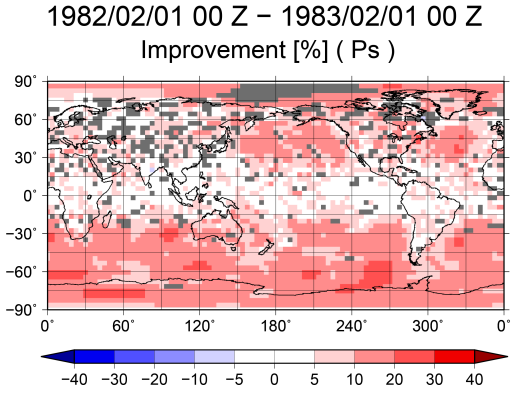
(a) Zonal wind at the 4th model level.



(b) Temperature at the 2th model level.

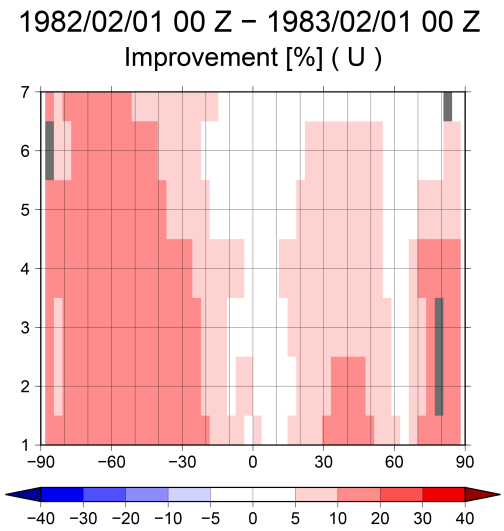


(c) Specific humidity at the 1st model level.

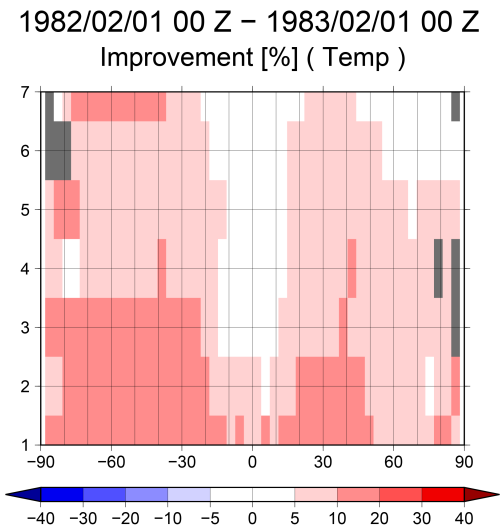


(d) Surface pressure.

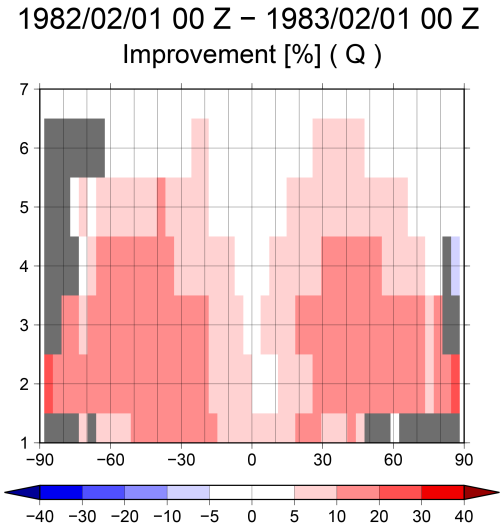
Figure 5.13. Improvements (%) of 1-year-average analysis RMSE of DLOC over CTRL for (a) zonal wind at the 4th model level (∼500 hPa), (b) temperature (K) at the 2nd model level (∼850 hPa), (c) specific humidity at the lowest level (∼925 hPa), and (d) surface pressure. Red (blue) indicates advantage (disadvantage) of DLOC.



(a) Zonal wind

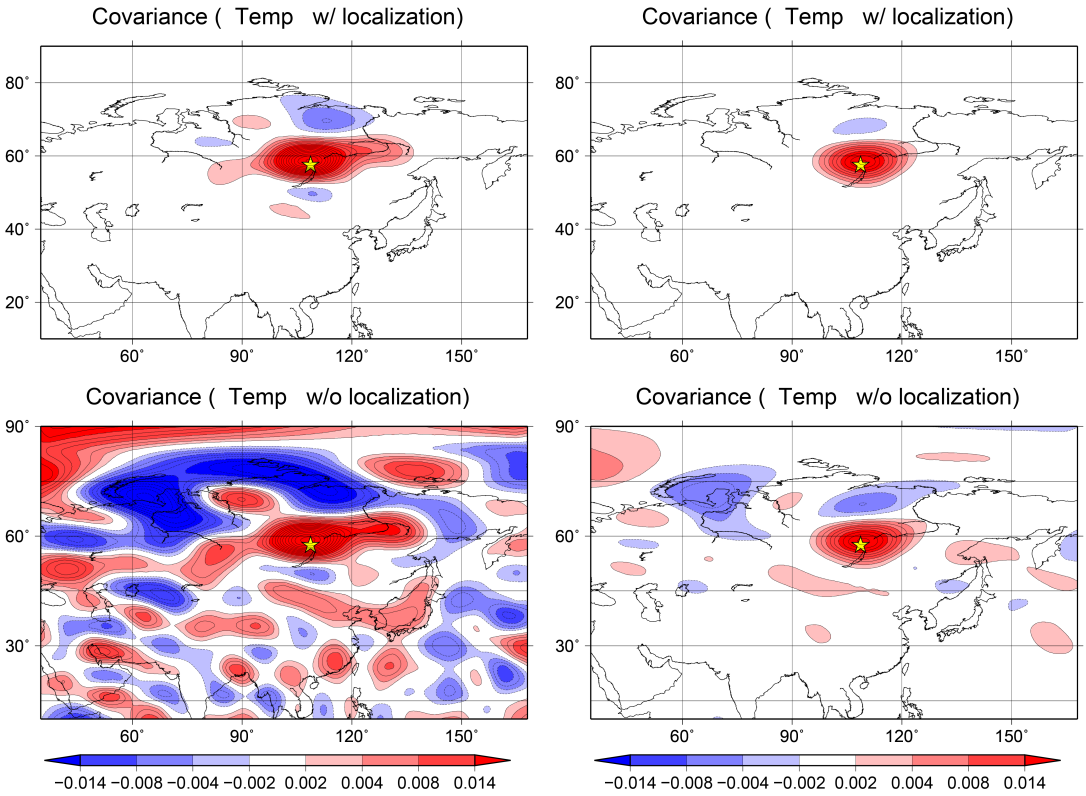


(b) Temperature



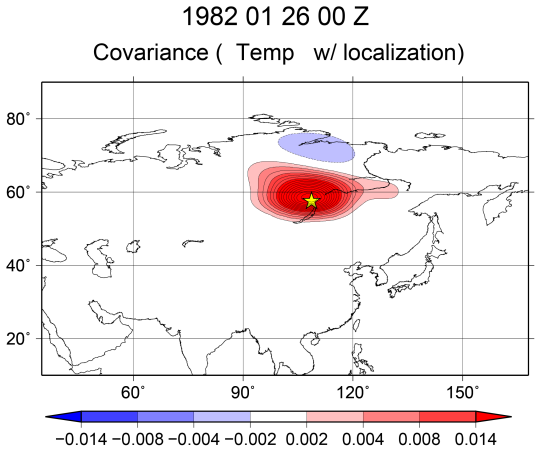
(c) Specific humidity

Figure 5.14. Improvements (%) of 1-year-average analysis RMSE of DLOC over CTRL for zonal-average vertical structure of (a) zonal wind, (b) temperature (K) and (c) specific humidity. Red (blue) indicates advantage (disadvantage) of DLOC.



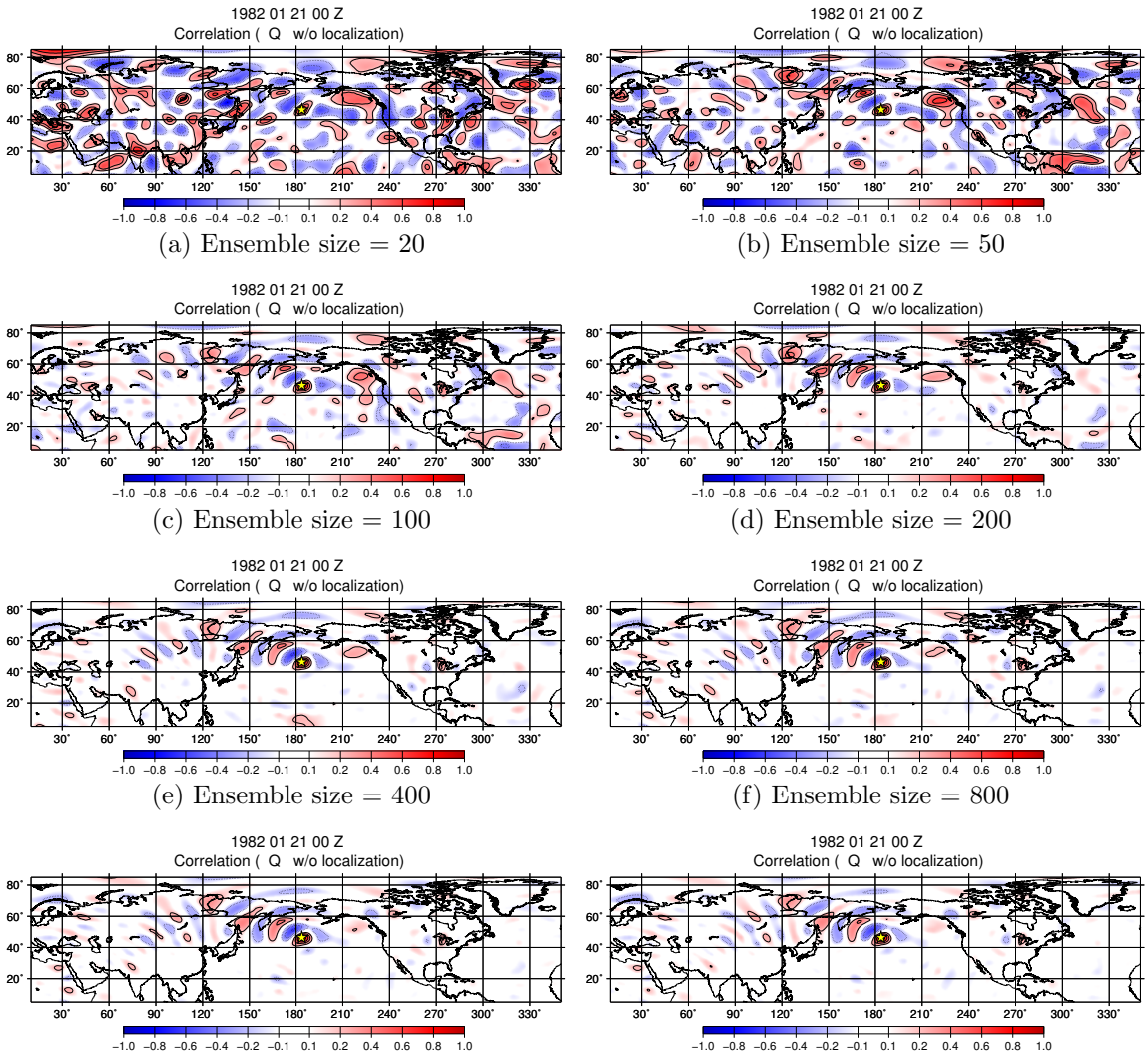
(a) Ensemble size = 20 with a single localization

(b) Ensemble size = 3200 with a single localization



(c) Ensemble size = 20 with a dual localization (500 km and 1000 km)

Figure 5.15. Horizontal background error covariance (K^2) maps of T at 1st level, (a) with 1000-km localization scale (top) and without localization (bottom) using 20 members, (b) with 1000-km localization scale (top) and without localization (bottom) using 3200 members and (c) with 500-km and 1000-km dual localization scales using 20 members.



(a) Ensemble size = 20

(b) Ensemble size = 50

(c) Ensemble size = 100

(d) Ensemble size = 200

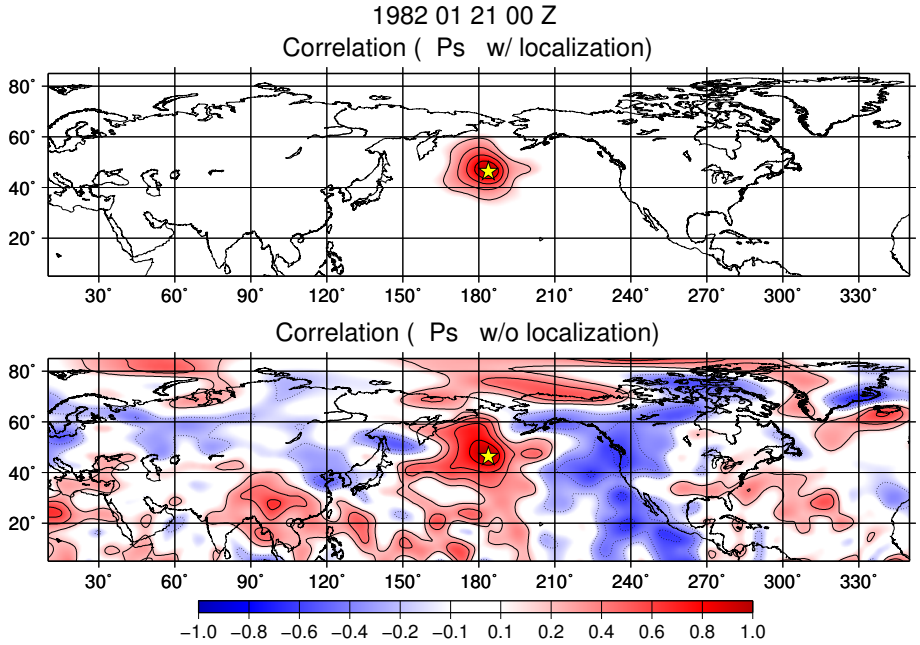
(e) Ensemble size = 400

(f) Ensemble size = 800

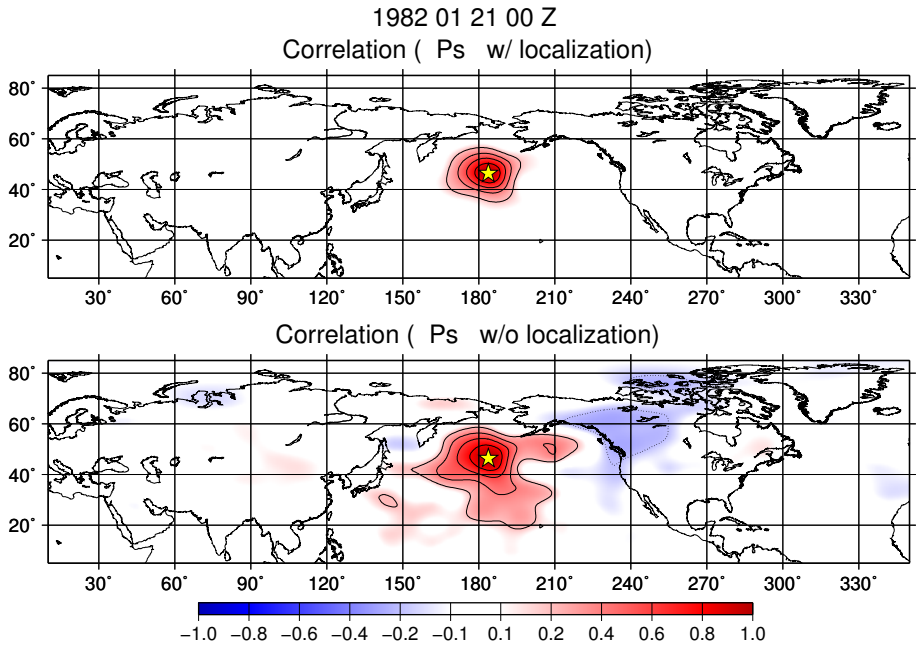
(g) Ensemble size = 1600

(h) Ensemble size = 3200

Figure 5.16. Horizontal background error correlation maps of Q at the 4th model level without localization, and using (a) ensemble size = 20, (b) ensemble size = 50, (c) ensemble size = 100, (d) ensemble size = 200, (e) ensemble size = 400, (f) ensemble size = 800, (g) ensemble size = 1600 and (h) ensemble size = 3200.



(a) Ensemble size = 20



(b) Ensemble size = 3200

Figure 5.17. Horizontal background error correlation maps of Ps (a) with 1000-km localization scale (top) and without localization (bottom) using 20 members, and (b) with 1000-km localization scale (top) and without localization (bottom) using 3200 members.

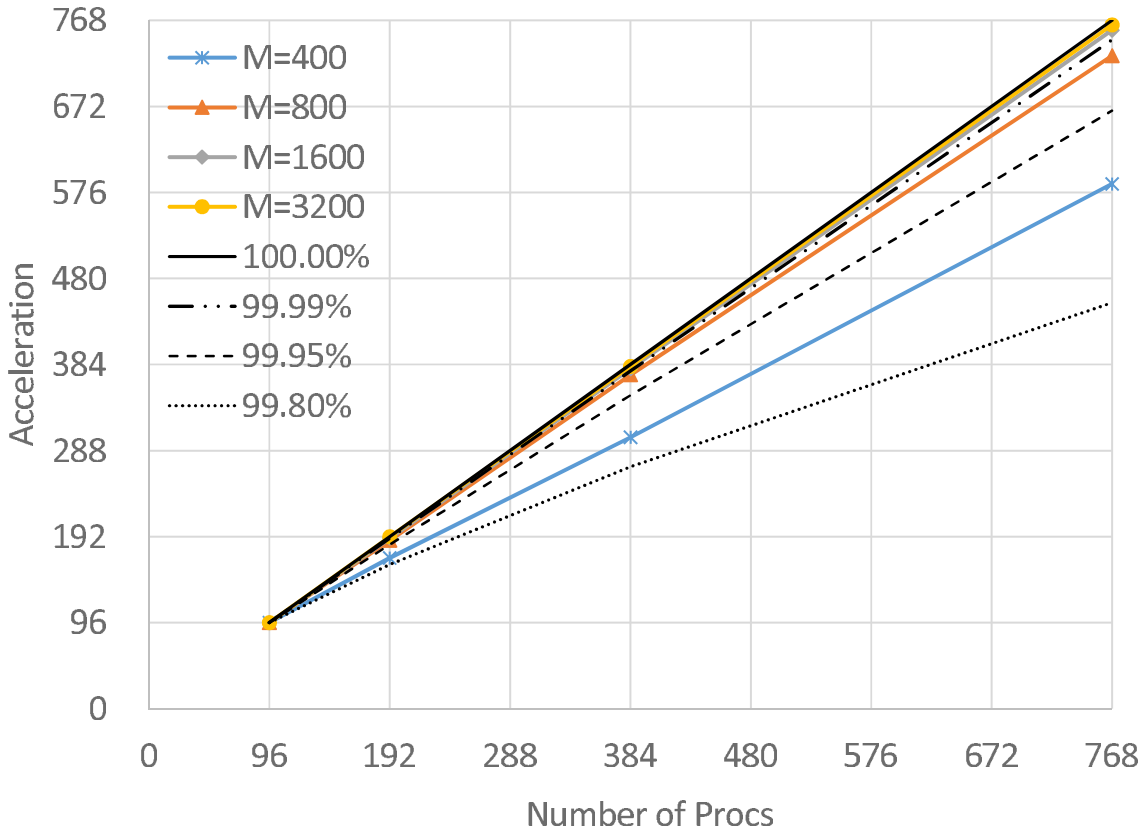


Figure 5.18. Acceleration ratios of LETKF with 400, 800, 1600 and 3200 ensemble members (color solid lines). Theoretically estimated acceleration ratios for parallelization are added (black lines).

Members	Computational time (sec)					
	Initialize	Read Obs.	Read Guess	Assimilation	Write Anal.	Total
100	2.00	0.19	0.65	2.95 (-)	0.47	6.40
200	1.04	0.36	1.17	14.76 (5.00)	0.69	18.15
400	0.07	0.68	2.19	90.10 (6.10)	1.63	94.80
800	1.06	1.19	4.16	666.20 (7.39)	3.27	675.99
1600	1.23	2.59	9.13	4,452.88 (6.68)	7.15	4,473.73
3200	0.97	9.37	16.26	31,645.46 (7.11)	13.38	31,711.64

Table 5.1. Computational time of LETKF by 12 nodes and 96 processors. () indicates the magnification of computational time when the ensemble members increases twice.

Members	Computational time (sec)					
	Initialize	Read Obs.	Read Guess	Assimilation	Write Anal.	Total
200	0.10	0.31	1.18	7.43 (-)	0.74	10.58
400	0.55	0.58	3.40	46.28 (6.23)	2.41	53.33
800	2.24	1.18	7.56	331.23 (7.16)	3.15	345.50
1600	2.05	2.57	13.31	2,232.83 (6.74)	6.47	2,257.40
3200	1.04	9.25	22.14	15,806.53 (7.08)	14.04	15,853.32

Table 5.2. Computational time of LETKF by 24 nodes and 192 processors. () indicates the magnification of computational time when the ensemble members increases twice.

Members	Computational time (sec)					
	Initialize	Read Obs.	Read Guess	Assimilation	Write Anal.	Total
400	2.19	0.62	6.45	23.12 (-)	6.24	38.74
800	2.15	1.28	10.51	167.01 (7.22)	3.01	184.07
1600	1.13	2.53	18.79	1,129.69 (6.76)	6.11	1,158.43
3200	0.35	8.27	48.02	7,991.79 (7.07)	12.96	8,061.66

Table 5.3. Computational time of LETKF by 48 nodes and 384 processors. () indicates the magnification of computational time when the ensemble members increases twice.

Members	Computational time (sec)					
	Initialize	Read Obs.	Read Guess	Assimilation	Write Anal.	Total
400	1.73	0.60	15.42	11.81 (-)	1.86	31.60
800	0.33	1.23	25.50	84.30 (7.14)	3.41	114.90
1600	2.30	2.32	39.50	563.44 (6.68)	6.81	614.20
3200	1.68	8.39	66.03	4,049.02 (7.19)	14.28	4,139.70

Table 5.4. Computational time of LETKF by 96 nodes and 768 processors. () indicates the magnification of computational time when the ensemble members increases twice.

CHAPTER VI

Concluding Summary

In this study, a new type of advanced data assimilation method is developed based on the EnKF method for very high-resolution models in view of future data assimilation in numerical weather predictions. First, in order to investigate the performances of EnKF, the EKF and EnKF are compared. Next, the LETKF is applied to the NICAM and the NICAM-LEKTF is constructed for the future cloud-resolving data assimilation. In the high resolution data assimilation, however, the region of available observations may be limited since larger-scale structures than localization scales are removed due to tight localization for the high-resolution model. For this reason, a new method of multi-scale localization technique is developed in this study in order to apply the LETKF to the high-resolution models. In this chapter, the results of this study are summarized as follows:

Comparison of the EKF and EnKF

In this study the accuracy of the full-rank EKF and EnKF is investigated under the perfect model scenario with the barotropic S-model. The accuracies of the EKF and EnKF are evaluated using the analysis error covariance matrix, RMSE and the ensemble spread.

First, the analysis error covariance matrices of the EKF an EnKF are compared in the spectral space by the eigenvalue decomposition of the variance of the analysis error. The magnitude of the 100th eigenvalue for the EKF is about $1/10^5$ of that of the first eigenvalue, suggesting that the covariance matrix is almost degenerate. It is found that the eigenvalue spectrum of the EnKF is smaller than that of the EKF, probably due to the assumption of the model linearization in the EKF. However, the most dominant error pattern represented by the EOF-1 of the analysis covariance matrix of EnKF has converged to the EKF when the ensemble member is increased to 50 or 100 and above. The spatial pattern represents the characteristics of the parameterized baroclinic instability which yields the largest error growth. Another rapid error growth is detected over the Arctic, and the result is probably the nonlinear interactions, but it should be investigated as a future work.

Next, the RMSEs and the ensemble spreads are compared for the EKF and EnKF. The RMSE is comparable to the ensemble spread for most of the experiments except for 20 members of the EnKF, although the experiment for the EnKF with 20 members does not result in filter divergence. The EnKF requires 100 ensemble members to accurately estimate the analysis error. The daily variations of the ensemble spreads are smaller than that of the RMSE. The former represents the characteristics of the model which does not strongly depend on the observation, whereas the latter is influenced by the introduced random noise in the time series. The accuracy of the EKF and the EnKFs is examined by the analysis error fields. The error in the analysis fields of the EnKFs is very similar to that of the EKF. The more the number of

members increases, the smaller the differences between the EKF and EnKF become. The results are confirmed from the comparison of the RMSEs for the EKF and the EnKF.

In order to investigate the statistical confidence of the result, we have repeated the same experiments for 10 cases in different seasons and different years. The magnitude of the analysis RMSE of the EKF is smaller than that of the EnKF with 50 ensemble members and is larger than that of the EnKF with 100 ensemble members. The RMSEs of the EnKF with 100, 410 and 1000 members are also smaller than that of the EKF, and the RMSE of the EnKF with 410 members is almost the same as that of 1000 members.

Here, the performance of the EKF and EnKF is discussed. According to the results of eigenvalue and RMSE of EKF and EnKF, the EnKF outperforms the EKF. In Zang and Malanotte-Rizzoli (2003), in the strongly nonlinear case the EnKF outperforms the reduced-rank EKF, and in the weakly nonlinear case the performances of the reduced-rank EKF and EnKF are similar. And in the barotropic S-model the error growth is relatively small because the dynamical instability of the model is weak (Tanaka and Nohara 2001). In this study, the difference of settings between the EKF and EnKF is whether to require the model linearization or not. Namely, the EnKF can directly treat the non-linear model although the EKF requires the model linearization, and the forecast error covariance of the EKF has some errors involved by the model linearization. So, the result is likely considered as the linearization of non-linear model. However, it is non-trivial that the EnKF outperforms the EKF.

The non-linear model may also have the negative effect on the performance of EnKF. In addition, we should also consider that the observation network influences to the performance. Moreover, it should be noted that the barotropic S-model may have a problem for the assimilation of the real observations due to the small internal instability. The small spread would not capture the true trajectory by the significant model errors.

Finally, the influence of the localization of the EnKF in the physical space is investigated for the barotropic S-model. In the physical space, the analysis errors of the EnKF become smaller than the observational errors. So the EnKF works appropriately. It seems that the analysis RMSEs of the EnKF have converged for the member size more than 20. This result agrees with that by Szunyogh et al. (2005). However, the comparison of the EnKF in the physical space and EnKF in the spectral space indicates that the RMSEs are larger in the physical space than that in the spectral space.

As the first reason, it is considered that the localization influences to the analysis field. Because the degree of freedom of the barotropic S-model is very small in the spectral space, the EnKF can assimilate all the observations even for the small ensemble size. On the other hand, the degree of the freedom is much larger in the physical space than in the spectral space for the barotropic S-model. For the physical space, it is necessary to apply the localization in order to reduce the spurious covariance among distant points due to the sampling errors, so that the filter will not diverge. In this study the influence of the localization is examined by applying the

EnKF to the physical space of the barotropic S-model. It is found that the EnKF can capture the analysis errors of regional scale such as synoptic disturbances. However, it cannot capture the analysis errors of the global scale such as teleconnection or zonally symmetric errors.

As the second reason, it is considered that the observations in the physical space have the local information and involve the retrieved errors. It is not guaranteed that the observational errors in physical space are equivalent to those in spectral space. On the other hand, in the spectral space the EnKF assimilates the observations whose elements are model variables. It may be important to note that the information of observations is very different between the spectral space and the physical space with the localization. In the spectral space, all state variables are assimilated as observations, whereas in the physical space, only few observations around the analysis point are assimilated. For this reason, the analysis results contain different number of observations and also different information about observations assuming no correlation between observational variables. Therefore, a careful comparison may be needed to understand the different analysis results. It is necessary to investigate the influence of the localization, the discrepancy of the assimilated observations and error correlation scale of the barotropic S-model in detail.

As is demonstrated in this study, it is found that the EnKF has a good performance compared to the full-rank EKF with the barotropic S-model in which the dynamic instability is weak. Therefore it is concluded that the EnKF converge to the EKF for the practical assimilation in the spectral space under the perfect model

scenario with the barotropic S-model, using 100 ensemble members.

NICAM-LETKF

In this study the LETKF is applied to the nonhydrostatic and realistic global atmospheric model called NICAM, and the data assimilation system is called the NICAM-LETKF. In addition, an algorithm which estimates not only the inflation parameter but also the observational errors adaptively is introduced to the NICAM-LETKF. We conducted three kinds of experiments to investigate the feasibility and stability of the NICAM-LETKF under the perfect model scenario: 1) the inflation parameter is fixed temporally (Ex. 1), 2) the inflation parameter is adaptively estimated, and the observational errors levels are perfectly specified (Ex. 2), and 3) the inflation parameter and the observational errors are adaptively estimated at a time (Exs. 3 and 4). It is confirmed that the LETKF works appropriately for the nonhydrostatic global model, although the model with the horizontal resolution of 224 km behaves hydrostatically.

First, it is confirmed that the NICAM-LETKF works stably without diverging, and the analysis errors become smaller than the observational errors in all variables. In addition, the magnitude and distribution of the analysis RMSEs are temporally and spatially comparable to those of the analysis ensemble spreads. The RMSEs are large in the area with large uncertainty such as the ITCZ or the developing extratropical cyclone. In such areas, each ensemble member spreads rapidly due to

the strong chaotic nature, resulting in the large ensemble spread. According to the above result, the inflation parameter is tuned optimally in the each grid point, so it is not necessary to tune the inflation parameter manually. These results indicate that the NICAM-LETKF combined with the adaptive estimation of the inflation parameter works appropriately and stably, and the NICAM-LETKF can capture the true analysis errors.

Second, in the case where the observational errors are perfectly known, and the inflation parameter is adaptively estimated (Ex. 2), the NICAM-LETKF converges fastest. On the other hand, in the case where the inflation parameter and observational errors are both estimated adaptively and the initial observational errors are 0.1 times the true value (Ex. 4), the NICAM-LETKF converges the most slowly. Since the observational errors are generally not perfectly known, the results are quite reasonable. In the early assimilation period, the accuracy of the analysis mainly depends on the inflation parameter and observational errors because the number of assimilated observations is insufficient. It may be important to note that the observational errors are overestimated by the present formulation, regardless of the magnitude of the initial observational errors. The observational error variance is obtained from Eq. 4.2. In the Eq. 4.2, it is assume that there is no correlation between the analysis error $\delta\mathbf{x}^a$ and the forecast error $\delta\mathbf{x}^f$. Actually, however, it is considered that there is a little correlation between the analysis error and the forecast error because the analysis fields and forecast field is similar in the area which a number of observations are a few or the ensemble spread is small. So, it is considered that the estimated observational

errors are overestimated. In the latter half of the analysis period, there is no clear difference among the analysis RMSEs of the 4 experiments. It is demonstrated that the observational errors sufficiently converge to the true values, and the inflation parameter is tuned optimally because the NICAM-LETKF assimilates sufficient number of observations.

We discuss about the sensitivity of the NICAM-LETKF to the observational error settings. The accuracies of the Exs. 2 and 3 are the same level in early period though the observational errors are overestimated in the Ex. 3. Therefore, it seems that the NICAM-LETKF has a small sensitivity to them. The adaptive estimation algorithm might have some advantages in the case such as assimilating real observations. The inflation parameter can be estimated appropriately only when the observational errors are adequately known. In addition, when the observation instrument is renewed and its observational error is changed, the LETKF can estimate its error by the adaptive estimation algorithm. Moreover, Li et al. (2009a) mentioned the possibility of the estimation of the error cross-correlations in such as satellite data.

In the assimilation with the real observations, the method which estimates the inflation parameter does not work properly because the observational errors are not perfectly known. Li et al. (2009a) reported the algorithm to adaptively estimate both the inflation parameter and the observational errors using the simple Lorenz-96 system and the SPEEDY model, which is based on the primitive equations. In this study, we demonstrated that the algorithm works appropriately even for the realis-

tic nonhydrostatic global model of NICAM. However, it is not clarified whether the NICAM-LETKF and the adaptive estimation algorithm works stably in the experiments with the real observations. It is suggested that this approach will suffer more in the real observations because of the model errors and the different sources of observations of similar variables. In Zhang et al. (2006), they tested the variance relaxation algorithm and reported that it is very effective and prevents filter divergence, preserving the structure of the local ensemble perturbations due to the error growth. These algorithms are useful to draw information from the observations as much as possible even for the experiments with the real observations. If the adaptive estimations of inflation and observation error are applied in the assimilation with the real observations, the model bias should be considered. The model bias has a bad influence upon the estimation algorithm of observation error. Li et al. (2009b) reported the model bias estimation algorithm (Dee and da Silva 1998) with the SPEEDY model, in which the model bias and state are separately estimated. In Dee and da Silva (1998), however, they assumed that the model error covariance has the same structure with the forecast error covariance because the model error covariance is unknown. This assumption is very strong, and it may be possible to relax the assumption using the adaptively estimated inflation parameter, but it should be the subject of the future work.

Multi-Scale Localization Approach

The dual-localization approach is proposed in Miyoshi and Kondo (2013), Kondo et al. (2013) and this study. The dual-localization approach analyzes the high-resolution and low-resolution components of analysis increments separately, using two localization scales and spatial smoothing. The high-resolution component is computed from the full-resolution perturbations and has a detailed structure of the error covariance by using a narrow localization. The low-resolution component is computed from the reduced-resolution perturbations and has a structure of the smoothed error covariance by using a wider localization. The error covariance computed from the low-resolution component has less sampling noise than the traditional single localization, and the influence of observations can reach to a longer distance. Therefore, the dual-localization approach is able to assimilate more observations efficiently through a larger localization and has a good performance more than the traditional single localization approach. These results are confirmed by comparing the covariance structures with 20 and 3200 ensemble members. The covariance structure of dual-localization approach is more similar to the covariance structure with 3200 ensemble members than the traditional single localization approach, using only 20 ensemble members. So, it is concluded that the dual-localization approach is able to capture larger scale structure removing the sampling errors. Therefore, this dual-localization approach (or multi-localization approach) is an important technique for the high-resolution models, in particular for the cloud-resolving data assimilation. On the other hands, it is confirmed that the scale of background error correlation are larger than the lo-

calization scale by increasing the ensemble member up to 3200. Some of the error correlations have wavy pattern from Pacific to Eurasia or PNA pattern, and the direction of strong correlation is very different in each variables. If we are able to treat the localization well, it is possible that the analysis errors are more improved.

This study also aims to investigate parameter sensitivities of the dual-localization approach. The two localization scale parameters and the choice of the smoothing function are the tuning parameters of the dual-localization method. The results showed robust performance of the dual-localization approach. Namely, dual localization outperformed traditional single localization with relatively wide choices of the two localization scale parameters by about 400-km range. In addition, two different smoothing functions are compared. Also, two smoothing functions, the spectral truncation as in Miyoshi and Kondo (2013) and the Lanczos filter with different smoothing strengths are investigated. The results showed no significant difference between the spectral truncation and Lanczos filter, and they gave essentially identical analysis performance if their smoothing strengths are similar. These findings are important in the sense that we may avoid a fine tuning of the parameters of dual localization. Since the Lanczos filter is simpler to implement and faster to compute, the Lanczos filter may be advantageous in implementing with higher-resolution regional models. Results also showed that the performance of the dual-localization approach strongly depends on the smoothing strength. With stronger smoothing, larger localization parameters give better results, although the minimum RMSE becomes slightly larger with stronger smoothing. Therefore, it is confirmed that DLOC outperformed CTRL

by about 20 % at all scales with the best combination of localization parameters and smoothing strength.

Furthermore, the dual-localization approach is extended to the triple-localization approach (not shown). As the result, the performance of triple-localization approach is almost the same as the dual-localization approach, although the choice of the largest localization scale is larger. And the triple-localization approach requires solving the LETKF analysis equations five times. So, it is concluded that the triple-localization approach is ineffective in the case of SPEEDY model because the horizontal resolution of the SPEEDY model is coarse. However, if the other higher-resolution models are used, it would be necessary to apply the triple-localization or multi-localization approach. If the observation stations are reduced in an effort to cut costs and are replaced by remote sensing data, it is difficult to assimilate larger scale information than the localization scale. Then, it is also considered that the multi-localization approach is one of the techniques to solve the problem in the high-resolution data assimilation.

Finally, the computational performance of LETKF is mentioned. The parallelization ratio of LETKF is more efficient than 99.99 % by increasing the ensemble size up to 1600 and 3200. To accelerate the computing speed, it is necessary to improve the eigenvalue decomposition of analysis error covariance matrix in ensemble space because the eigenvalue decomposition occupies most of the computational time. This result almost agrees with that by Miyoshi and Yamane (2007).

Future Subjects

In this study, we have confirmed that the EnKF converges to the EKF with low-resolution atmospheric model. We have concluded that the application of the multi-scale localization methods to the EnKF may construct the scientific basis of the future data assimilation for the very high-resolution models. It is desired to develop a new data assimilation method for the very high-resolution cloud-resolving global model such as NICAM. The multi-scale localization methods developed in this study can be applied to such high-resolution global and regional models in a future, using a high-performance parallel computing system.

ACKNOWLEDGMENTS

First of all, I would like to express special appreciation to Prof. Hiroshi L. Tanaka, Center for Computational Sciences, University of Tsukuba, for his variable comments and encouragements. I am grateful to Prof. Takemasa Miyoshi, RIKEN Advanced Institute for Computational Science, for variable comments and encouragements. I am also thankful to Profs. H. Ueda, K. Ueno, H. Kusaka and Y. Wakazuki for their variable comments and suggestions. I express my appreciation to Prof. T. Y. Tanaka of Meteorological Research Institute for accepting the committee of my doctoral thesis despite his crowded schedule. I am most grateful to Prof. Kalnay in University of Maryland, Dr. Matsueda in Oxford University, Dr. Terasaki in RIKEN Advanced Institute for Computational Science and Dr. S. Otsuka in RIKEN Advanced Institute for Computational Science for their various advice on my study. I am grateful to all other students and staff of the Climatology and Meteorology Group, the University of Tsukuba, for their comments and supports. Finally, I am most thankful to my family for their support and understanding.

REFERENCES

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

Arakawa, A., 2004: The cumulus parameterization problem: Past present and future. *J. Clim.*, **17**, 2493–2525.

Arakawa, A., and W. H. Schubert, 1974: Interaction of cumulus cloud ensemble with the large-scale environment. Part I. *J. Atmos. Sci.*, **31**, 671–701.

Bishop, C. H., B. J. Etherton, and S. J. Majumdar, 2001: Adaptive sampling with the ensemble transform Kalman filter. Part I: Theoretical aspects. *Mon. Wea. Rev.*, **129**, 420–436.

Buehner, M., 2012: Evaluation of a spatial/spectral covariance localization approach for atmospheric data assimilation. *Mon. Wea. Rev.*, **140**, 617–636.

Dee, D. P., and A. M. da Silva, 1998: Data assimilation in the presence of forecast bias. *Quart. J. Roy. Meteor. Soc.*, **124**, 1128–1145.

Desroziers G., L. Berre, B. Chapnik, and P. Poli 2005: Diagnosis of observation, backgraound and analysis error statistics in observation space. *Quart. J. Roy. Meteor. Soc.*, **131**, 3385–3396.

Duchon, C. E., 1979: Lanczos filtering in one and two dimensions. *J. Appl. Meteor.*, **18**, 1016–1022.

Evensen, G., 1994: Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *J. Geophys. Res.*, **99C5**, 10143–10162.

Fertig, E., J. Harlim, and B. R. Hunt, 2007: A comparative study of 4D-Var and a 4D ensemble Kalman filter: Perfect model simulations with Lorenz-96. *Tellus*, **59A**, 96–100.

Gaspari, G., and S. E. Cohn, 1999: Construction of correlation functions in two and three dimensions. *Quart. J. Roy. Meteor. Soc.*, **125**, 723–757.

Greybush, Steven J., E. Kalnay, T. Miyoshi, K. Ide and B. R. Hunt, 2011: Balance and ensemble Kalman filter localization techniques. *Mon. Wea. Rev.*, **139**, 511–522.

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

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.

Hunt, B. R., E. Kostelich, and I. Syzunogh, 2007: Efficient data assimilation for spatiotemporal chaos: A local ensemble transform Kalman filter. *Physica D*, **230**, 112–126.

Jazwinski, A. H., 1970: *Stochastic Processes and Filtering Theory.* Academic Press, 376pp.

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

Kalnay, E. and Shu-Chih Yang, 2010: Accelerating the spin-up of ensemble Kalman filtering. *Quart. J. Roy. Meteor. Soc.*, **136**, 1644–1651.

Kalnay, E., H. Li, T. Miyoshi, S. Yang, and J. Ballabrera-poy, 2007: 4-D-Var or ensemble Kalman filter? *Tellus*, **59A**, 758–773.

Kasahara, A., 1977: Numerical integration of the global barotropic primitive equations with Hough harmonic expansions. *J. Atmos. Sci.*, **34**, 687–701.

Kondo, K., 2007: A verification experiments of ensemble Kalman filter using a barotropic general circulation model. Bachelor Scholarly Paper. College of Natural Sciences, University of Tsukuba, 115pp.

Kondo, K., 2009: Applying a local ensemble transform Kalman filter to the nonhydrostatic icosahedral atmospheric model. Masters Scholarly Paper. Department of Geoscience, University of Tsukuba, 155pp.

Kondo, K. and H. L. Tanaka, 2009: Comparison of the extended Kalman filter and the ensemble Kalman filter using the barotropic general circulation model. *J. Meteor. Soc. Japan*, **87**, 347–359.

Kondo, K. and H. L. Tanaka, 2009: Applying the local ensemble transform Kalman filter to the nonhydrostatic icosahedral atmospheric model (NICAM). *SOLA*, **5**, 121–124.

Kondo, K. and T. Miyoshi and H. L. Tanaka, 2013: Parameter sensitivities of the dual-localization approach in the local ensemble transform Kalman filter. *SOLA*, **9**, 174–178.

Lanczos, C., 1956: *Applied Analysis*. Prentice-Hall, 539 pp.

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

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

Liu, H. and X. Zou, 2001: The impact of NORPEX targeted dropsondes on the analysis and 2-3-day forecasts of a landfalling Pacific winter storm using NCEP 3D-Var and 4D-Var systems. *Mon. Wea. Rev.*, **129**, 1987–2004.

Lorenz, E. N. 1996: Predictability: A problem partly solved. *In Proc. Predictability, ECMWF, 4-8 September 1995*.

Louis, J., 1979: A parametric model of vertical eddy fluxes in the atmosphere. *Bound.-Layer Meteor.*, **17**, 187–202.

Mellor, G. L., and T. Yamada, 1974: A hierarchy of turbulence closure models for planetary boundary layers. *J. Atmos. Sci.*, **31**, 1791–1806.

Miyoshi, T., 2005: Ensemble Kalman filter experiments with a primitive-equation global model. Ph.D. dissertation, University of Maryland, 197 pp.

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. doi: 10.1175/2010MWR3570.1.

Miyoshi, T., and S. Yamane 2007: Local ensemble transform Kalman filtering with an AGCM at a T159/L48 resolution. *Mon. Wea. Rev.*, **135**, 3841–3861.

Miyoshi, T., and M. Kunii, 2012: The local ensemble transform Kalman filter with the Weather Research and Forecasting model: Experiments with real observations. *Mon. Wea. Rev.*, **169**, 321–333.

Miyoshi, T. and K. Kondo, 2013: A multi-scale localization approach to an ensemble Kalman filter. *SOLA*, **9**, 170–173.

Miyoshi, T., S. Yamane, and T. Enomoto, 2007: Localizing the error covariance by physical distances within a local ensemble transform Kalman filter (LETKF). *SOLA*, **3**, 89–92.

Miyoshi T., Yoshiaki Sato, and Takashi Kadowaki, 2010: Ensemble Kalman filter and 4D-Var intercomparison with the japanese operational global analysis and prediction system. *Mon. Wea. Rev.*, **138**, 2846–2866.

Molteni, F., 2003: Atmospheric simulations using a GCM with simplified physical parametrizations. I: model climatology and variability in multi-decadal experiments. *Clim. Dyn.*, **20**, 175–191.

Ohfuchi, W., H. Nakamura, M. K. Yoshioka, T. Enomoto, K. Takaya, X. Peng, S. Yamane, T. Nishimura, Y. Kurihara, and K. Ninomiya, 2004: 10-km mesh meso-scale resolving simulations of the global atmosphere on the Earth Simulator: Preliminary outcomes of AFES (AGCM for the Earth Simulator). *J. Earth Simulator*, **1**, 8–34.

Ott, E., B. R. Hunt, I. Szunyogh, A. V. Zimin, E. J. Kostelich, and coauthors., 2004: A local ensemble Kalman filter for atmospheric data assimilation. *Tellus*, **56A**,

415–428.

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

Pham, D. T., 2001: Stocastic methods for sequentail data assimilation in strongly nonlinear systems. *Mon. Wea. Rev.*, **129**, 1194–1207.

Pierre G. and T. Jean-Noël, 2001: Impact of the digital filter as a weak constraint in the preoperational 4D-Var assimilation system of Meteo-France. *Amer. Meteor. Soc.*, **129**, 2089–2102.

Saito, K., T. Fujita, Y. Yamada, J. Ishida, Y. Kumagai, and coauthors., 2006: The operational JMA nonhydrostatic mesoscale model. *Mon. Wea. Rev.*, **134**, 1266–1298.

Satoh, M., T. Matsuno,T., H. Tomita, H. Miura, T. Nasuno, and S. Iga, 2008: Nonhydrostatic Icosahedral Atmospheric Model (NICAM) for global cloud resolving simulations. *Journal of Computational Physics*, the special issue on Predicting Weather, Climate and Extreme events. **227**, 3486–3514, doi: 10.1016/j.jcp.2007.02.006.

Seko, H., T. Miyoshi, Y. Shoji, and K. Saito, 2011: Data assimilation experiments of precipitable water vapour using the LETKF system: intense rainfall event over Japan 28 July 2008. *Tellus*, **63A**, 402–414.

Skamarock, W. C., J. B. Klemp, J. Dudhia, D. O. Gill, D. M. Barker, W. Wang, and J. G. Powers, 2005: *A description of the Advanced Research WRF* version 2. NCAR Tech. Note TN- 468 + STR, 88 pp.

Szunyogh, I., E. J. Kostelich, G. Gyarmati, D. J. Patil, B. R. Hunt, E. Kalnay, E. Ott, and J. A. Yorke, 2005: Assessing a local ensemble Kalman filter: Perfect model experiments with the National Centers for Environmental Prediction global model. *Tellus*, **57A**, 528–545.

Tanaka, H. L., 1985: Global energetics analysis by expansion into three dimensional normal mode functions during the FGGE winter. *J. Meteor. Soc. Japan*, **63**,

180–200.

Tanaka, H. L., 1991: A numerical simulation of amplification of low-frequency planetary waves and blocking formations by the up-scale energy cascade. *Mon. Wea. Rev.*, **119**, 2919–2935.

Tanaka, H. L., 1998: Numerical simulation of a life-cycle of atmospheric blocking and the analysis of potential verticity using a simple barotropic model. *J. Meteor. Soc. Japan*, **76**, 983–1008.

Tanaka, H. L., and D. Nohara, 2001: A study of deterministic predictability for the barotropic component of the atmosphere. *Science Rep.*, **22A**, Institute Geoscience, University of Tsukuba, 45 pp.

Tanaka, H. L., 2003: Analysis and modeling the Arctic Oscillation using a simple barotoropic model with baroclinic eddy forcing. *J. Atmos. Sci.*, **60**, 1359–1379.

Wallace, J. M., and D. S. Gutzler, 1981: Teleconnections in the geopotential height field during the northern hemisphere winter. *Mon. Wea. Rev.*, *109*, 784–812.

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

Whitaker, J. S., T. M. Hamill, X. Wei, Y. Song, and Z. Toth, 2008: Ensemble data assimilation with the NCEP global forecast system. *Mon. Wea. Rev.*, **136**. 463–482.

Yussouf, N. and D. J. Stensrud, 2010: Impact of phased-array radar observations over a short assimilation period: Observing system simulation experiments using an ensemble Kalman filter. *Mon. Wea. Rev.*, **138**, 517–538.

Zang, X., and P. Malanotte-Rizzoli, 2003: A comparison of assimilation results from the ensemble Kalman filter and the reduced-rank extended Kalman filter. *Nonlinear Processes in Geophysics*, **10**, 477–491.

Zhang F, Meng Z, Aksoy A. 2006. Tests of an ensemble Kalman filter for mesoscale and regional-scale data assimilation. Part I: Perfect model experiments. *Mon. Wea. Rev.*, **134**, 722–736.

Zhang, F., Y. Weng, J. A. Sippel, Z. Meng, and C. H. Bishop, 2009: Cloud-Resolving hurricane initialization and prediction through assimilation of doppler radar observations with an ensemble Kalman filter. *Mon. Wea. Rev.*, **137**, 2105–2125.