Tellus **000**, 000–000 (0000) Printed 16 February 2007 (Tellus LaTeX style file v2.2)

# Assimilating Nonlocal Observations using a Local Ensemble Kalman Filter

Elana J. Fertig[1]*, Brian R. Hunt[1], Edward Ott[2], and Istvan Szunyogh[3]

[1] *Institute for Physical Science and Technology and Department of Mathematics, University of Maryland, College Park, MD, USA;* [2] *Institute for Research in Electronics and Applied Physics, Department of Electrical and Computer Engineering and Department of Physics, University of Maryland, College Park, MD, USA;* [3] *Department of Meteorology and Institute for Physical Science and Technology, University of Maryland, College Park, MD, USA*

(Manuscript received 20 October 2006; Revised 16 February 2007)

## ABSTRACT

Many ensemble Kalman filter data assimilation schemes benefit from spatial localization, often in both the horizontal and vertical coordinates. On the other hand, satellite observations are often sensitive to the dynamics over a broad layer of the atmosphere; that is, the observation operator that maps the model state to the observed satellite radiances is a nonlocal function of the state. Similarly, errors in satellite retrieval observations can be correlated over significant distances. This nonlocality can present problems for assimilating satellite observations with local ensemble Kalman filter schemes. In this paper, we propose a technique in which the observation operator is applied to the global model state and then appropriate observations are selected to estimate the atmospheric state for each model grid point. The issue of how to choose appropriate observations is investigated with numerical experiments on a seven layer primitive equation model, the SPEEDY model. We assimilate both simulated point observations and either nonlocal radiance-like or retrieval-like observations with a particular ensemble Kalman filter, LETKF. The best analysis results are obtained from a scheme that updates the state at a given location by assimilating all those observations that are strongly correlated to the model state near that location.

* Corresponding author.
email:ejfertig@math.umd.edu

## 1 Introduction

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

Local ensemble-based methods typically allow the analysis at a given grid point to depend only on observations within a surrounding local region. The model dynamics in this local region is assumed to be strongly correlated with the model state at the analysis location. Assimilation experiments suggest that choosing a region containing only a subset of the horizontal model grid points and vertical model levels benefits the ensemble-based Kalman filter analysis (e.g., Hamill et al., 2001; Szunyogh et al., 2005; Houtekamer et al., 2005; Whitaker et al., 2006). However, a challenge arises when assimilating satellite observations in this local framework. Many satellite channels provide radiance observations sensitive to the dynamics over a broad layer of the atmosphere (e.g., Liou, 2002). Consequently, resulting observations are correlated to the atmospheric state over several model levels and selecting which satellite observations to use to perform a given local analysis is not straightforward.

Houtekamer et al. (2005) and Houtekamer and Mitchell (2006) address this problem for Advanced Microwave Sounding Unit-A (AMSU-A) radiance observations by treating each observation as if it were a local observation made at the pressure level for which the corresponding satellite channel is most sensitive. With this vertical coordinate, they then assimilate those observations within the local region. While Houtekamer and Mitchell (2006) demonstrated that with this selection strategy the AMSU-A observations have benefited the analysis, it is unknown to what extent the localization prevents the analysis from extracting information from these observations. In particular, an observation assigned to a particular model level may still depend strongly on the dynamics at more distant levels. Therefore, the local analysis at those distant levels may benefit from assimilating that observation.

However, that observation may not be selected to update that level because its assigned model level is too distant.

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

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

We test the above technique for a mixture of simulated point, nonlocal radiance-like, and correlated retrieval-like observations with the seven level SPEEDY primitive equation model of Molteni (2003). In this paper, we assimilate these observations with a specific ensemble Kalman filter scheme. In particular, we use the localization set-up proposed in Ott et al. (2004) with the algorithmic improvements of Hunt et al. (2006). The latter is referred to as the Local Ensemble Transform Kalman Filter (LETKF, Hunt et al., 2006).

A brief introduction to ensemble-based Kalman filter data assimilation techniques and the LETKF scheme are presented in Section 2. This section explains the methods of localization employed and then develops a means to select nonlocal observations corresponding to a local region in model space. Section 2 also introduces the SPEEDY model, which we use in our perfect model experiments presented in Section 3. The results and a summary of conclusions are presented in Sections 4 and 5, respectively. These sections suggest updat-

ing the state at a location by assimilating all nonlocal observations correlated to the model state near that location. These results are obtained for simplified simulated satellite observations using a particular ensemble Kalman filter scheme. However, we expect the results will hold more generally when assimilating real satellite observations with any data assimilation scheme requiring localization.

## 2 Background

### *2.1 Ensemble Kalman Filter*

Kalman filter based data assimilation schemes obtain a best guess (or "analysis") for the state of the atmosphere by combining a forecast started in the past (called the "first guess" or "background") and observations. The basic formulation is as follows (Kalman, 1960; Kalman and Bucy, 1961),

$$\mathbf{P}^a = (\mathbf{I} - \mathbf{K}\mathbf{H})\mathbf{P}^b, \tag{1}$$

$$\mathbf{x}^a = \mathbf{x}^b + \mathbf{K}\left(\mathbf{y} - \mathbf{h}\left(\mathbf{x}^b\right)\right), \tag{2}$$

$$\mathbf{K} = \mathbf{P}^b\mathbf{H}^\top\left(\mathbf{H}\mathbf{P}^b\mathbf{H}^\top + \mathbf{R}\right)^{-1}. \tag{3}$$

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

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

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

$$\mathbf{P}^b = (k-1)^{-1}\,\mathbf{X}^b\mathbf{X}^{b\top} = \frac{\sum_{i=1}^{k}\left(\mathbf{x}^{f(i)} - \mathbf{x}^b\right)\left(\mathbf{x}^{f(i)} - \mathbf{x}^b\right)^\top}{k-1}, \tag{4}$$

where the $i$th column of the matrix $\mathbf{X}^b$ is given by $\mathbf{x}^{f(i)} - \mathbf{x}^b$. In contrast to some other approaches, this formulation does not require linearization of the model. Furthermore, $\mathbf{P}^b$ has rank at most $k-1$. Thus, if $k$ is chosen small enough, the number of degrees of freedom in the Kalman Filter equations can be greatly reduced. However, while this truncation reduces the size of the matrix, it can also introduce severe sampling fluctuations for realistically affordable ensemble sizes $k$. As discussed in Section 2.2, this latter problem can be efficiently addressed by "localization".

Equation (4) is substituted into the Kalman Filter equations (1)–(3) to give:

$$\mathbf{P}^a = (k-1)^{-1}\left(\mathbf{X}^b - \mathbf{K}\mathbf{Y}^b\right)\mathbf{X}^{b\top}, \tag{5}$$

$$\mathbf{x}^a = \mathbf{x}^b + \mathbf{K}\left(\mathbf{y} - \mathbf{h}\left(\mathbf{x}^b\right)\right), \tag{6}$$

$$\mathbf{K} = (k-1)^{-1}\,\mathbf{X}^b\mathbf{Y}^{b\top}\left[(k-1)^{-1}\,\mathbf{Y}^b\mathbf{Y}^{b\top} + \mathbf{R}\right]^{-1}, \tag{7}$$

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

$$\mathbf{Y}^{b(i)} = \mathbf{h}\left(\mathbf{x}^{f(i)}\right) - \overline{\mathbf{h}\left(\mathbf{x}^{f(i)}\right)}, \tag{8}$$

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

Equations (5)–(7) form the basis for many ensemble-based Kalman filter algorithms (Evensen, 1994; Burgers et al., 1998; Houtekamer and Mitchell, 1998; Anderson, 2001; Bishop et al., 2001; Whitaker and Hamill, 2002; Ott et al., 2004). However, as posed here these equations are costly to solve because the matrix inverted in Equation (7) has as many dimensions as the number of observations, which is typically as large as $O(10^6)$. Several methods (i.e., Houtekamer and Mitchell, 2001; Anderson, 2003; Whitaker and Hamill, 2002) reduce the cost of an ensemble-based scheme by assimilating observations sequentially, so that each evaluation of the right hand side of Equation (7) involves only the inverse of a

scalar. Alternately, several schemes are made efficient by choosing a basis in which they need only invert a matrix of size roughly $k \times k$ (i.e., Bishop et al., 2001; Ott et al., 2004; Hunt et al., 2006). In this study, we use an efficient modification of the scheme of Ott et al. (2004), formulated in Hunt et al. (2006) which has been called the Local Ensemble Transform Kalman Filter (LETKF). Following the notation of the latter paper, Equations (5)–(7) become

$$\tilde{\mathbf{P}}^a = \left[(k-1)\,\mathbf{I} + \mathbf{Y}^{b\top}\mathbf{R}^{-1}\mathbf{Y}^b\right]^{-1}, \tag{9}$$

$$\mathbf{P}^a = \mathbf{X}^b\tilde{\mathbf{P}}^a\mathbf{X}^{b\top}, \tag{10}$$

$$\mathbf{x}^a = \mathbf{x}^b + \mathbf{X}^b\tilde{\mathbf{P}}^a\mathbf{Y}^{b\top}\mathbf{R}^{-1}\left(\mathbf{y} - \mathbf{h}\left(\mathbf{x}^b\right)\right). \tag{11}$$

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

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

$$\mathbf{x}^{a(i)} = \mathbf{x}^a + \mathbf{X}^{a(i)}, \tag{12}$$

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

$$\mathbf{X}^a = \mathbf{X}^b\left((k-1)\tilde{\mathbf{P}}^a\right)^{1/2}, \tag{13}$$

where by $\mathbf{M}^{1/2}$ we mean the symmetric square root of a symmetric matrix $\mathbf{M}$. Notice that by Equation (10),

$$\mathbf{P}^a = (k-1)^{-1}\,\mathbf{X}^a\mathbf{X}^{a\top}. \tag{14}$$

### *2.2 Localization*

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

Appendix A. The sequential assimilation of observations neglects correlations between observational errors which, if present, are represented by nonzero off-diagonal entries in the error covariance matrix $\mathbf{R}$. If these off-diagonal entries are significant, they can be taken into account by simultaneously assimilating observations in correlated batches (Houtekamer and Mitchell, 2001). Alternatively, $\mathbf{R}$ can be diagonalized by a coordinate transformation in observation space, and the transformed observations can be assimilated sequentially (Anderson, 2003). As we discuss in Appendix B, diagonalizing $\mathbf{R}$ can be problematic when observation error correlations are nonlocal, because the transformed observations are not associated with a specific location and thus their influence is hard to locate.

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

### *2.3 Assimilating Nonlocal Observations*

In this section, we discuss options for selecting which nonlocal observations to use in the local model grid analysis for a given grid point. Because we may select observations that depend on model variables outside the local region defined for uncorrelated point observations, we apply the observation operator to the global model state for each ensemble member. We include only those rows of $\mathbf{Y}^b$ and $\mathbf{h}(\mathbf{x}^b)$ corresponding to the selected observations, truncate the observation error covariance matrix accordingly, and solve the resulting ensemble-based Equations (9)–(11) exactly.

Selecting which satellite radiances to assimilate is complicated by the fact that they do not have a single well-defined vertical location. We note that the weighting function at a particular model point indicates the sensitivity of that observation to the state at

that model grid point (Liou, 2002). Houtekamer et al. (2005) and Houtekamer and Mitchell (2006) choose to assign radiance observations to the model level for which the magnitude of the weighting function is largest. They use this location to select observations within the local region in model space. Throughout this paper, we refer to this method for selecting radiance observations as "maximum-based selection". We represent this scheme graphically in Figure 1(a).

An alternative to the maximum-based selection is to examine the weighting function for each nonlocal observation, and choose to assimilate it if a "significant" weight is assigned to any model state vector component within the local region; otherwise, this observation is not used in the assimilation. We refer to this observation selection as "cutoff-based selection" and represent it graphically in Figure 1(b). Maximum-based selection may be regarded as a particular case of cutoff-based selection in which the only significant weight is the largest weight. Appendix A applies the cutoff-based selection to covariance localization schemes. In Section 3, we run a series of numerical experiments to compare the maximum-based and cutoff-based observation selection schemes. We note that we avoid the additional complication of the nonlinearity of the satellite observation operator in this study by letting the observation operator be a linear weighting function.

As mentioned earlier, satellite retrievals are nonlocal in that their errors can be strongly correlated across significant distances (e.g., Kalnay, 2003). These correlations are manifested in nonzero values in the off diagonal terms of the observation error covariance matrix, $\mathbf{R}$. The magnitude of the correlation ($\mathbf{R}_{ij}/\sqrt{\mathbf{R}_{ii}\mathbf{R}_{jj}}$) between the errors of two observations, $i$ and $j$, can be used to select which correlated observations to assimilate. That is, we assimilate all the retrieval observations located within the local region and all other retrieval observations for which the magnitude of the correlation of their errors with the errors of any observation in the local region is above a certain threshold. We tune this threshold value in the numerical experiments in Section 4. For comparison, we also assimilate only observations inside the local region, neglecting the correlations with more distant observations. In this case, we scale the entries of the observation error covariance by a constant value and tune that constant to obtain the analysis with the smallest error.

### *2.4 SPEEDY Model*

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

As detailed in Molteni (2003), the SPEEDY model uses the 1981-90 reanalysis from the European Centre for Medium-Range Weather Forecasts (ECMWF) to obtain its boundary conditions at the bottom of the atmosphere for fields including topographical height, land-sea mask, sea ice fraction, sea surface temperature (SST), snow depth, bare-surface albedo, surface temperature and moisture in the top soil layer, and land-surface covered by vegetation. The SPEEDY model utilizes simplifications of the basic physical parameterization schemes present in many general circulation models (GCMs) for processes such as convection, large-scale condensation, clouds, short-wave radiation, long-wave radiation, surface fluxes, and vertical diffusion. We employ the SPEEDY model for the numerical experiments in this study because it retains many of the characteristics of operational GCMs, but at an order of magnitude lower computational cost.

## 3 Experimental Design

We explore the methods proposed in Section 2.3 for selecting nonlocal observations carrying out perfect model experiments with the SPEEDY model. The known "true" state of the atmosphere is a 180 time step (roughly one and a half month) integration of the SPEEDY model. Following Miyoshi (2005), the initial condition for the truth results from a year long run of the SPEEDY model from its default rest state on January 01, 1981. For consistency, we assimilate the same observations of this truth for all of the numerical experiments. To determine the quality of the assimilations, we convert the analysis and true states to pressure coordinates, at pressure levels (925, 850, 700, 500, 300, 200, 100 hPa). We then compute the

global root mean square (RMS) error between the true and analysis states at each pressure level. The average RMS error refers to the average of the RMS error over the last month of the assimilation cycle.

Every six hours, we create simulated local and nonlocal remotely sensed observations. We obtain local "point" observations by adding zero mean Gaussian-distributed noise to the modified truth. The noise has standard deviation 1 m/s for $u$ and $v$, 1 K for $T$, 0.0001 kg/kg for $q$, and 1 hPa for $p_s$. Following Miyoshi (2005), we simulate a rawinsonde network by retaining only those point observations which are closest to the real rawinsonde locations at the 12Z synoptic time. For example, the corresponding observation distribution for $p_s$ is shown in Figure 2.

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

$$y_{(i,j,n)} = \sum_{m=1}^{7} \mathbf{H}_{nm} x^t_{(i,j,m)} + \boldsymbol{\eta}, \tag{15}$$

where $x^t_{(i,j,m)}$ is the "true" temperature at the horizontal location $(i, j)$ and model level $m$. For this study, we have as many radiance observations as model levels (i.e., $n$ runs from 1 to 7). We take $\mathbf{H}_{nm} = 2^{-|n-m|-1}$ if $|n - m| \leqslant 3$, and 0 otherwise; so that

$$\mathbf{H} = \begin{bmatrix} 1/2 & 1/4 & 1/8 & 1/16 & 0 & 0 & 0 \\ 1/4 & 1/2 & 1/4 & 1/8 & 1/16 & 0 & 0 \\ 1/8 & 1/4 & 1/2 & 1/4 & 1/8 & 1/16 & 0 \\ 1/16 & 1/8 & 1/4 & 1/2 & 1/4 & 1/8 & 1/16 \\ 0 & 1/16 & 1/8 & 1/4 & 1/2 & 1/4 & 1/8 \\ 0 & 0 & 1/16 & 1/8 & 1/4 & 1/2 & 1/4 \\ 0 & 0 & 0 & 1/16 & 1/8 & 1/4 & 1/2 \end{bmatrix}. \tag{16}$$

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

We also used the simulated radiance observations to create simulated retrieval observations $\mathbf{y}_{\text{ret}}$, which we obtain by solving the equation $\mathbf{y} = \mathbf{H}\mathbf{y}_{\text{ret}}$. Because, in this case, the $\mathbf{H}$ operator is linear and has full rank, its inverse can be applied to solve the equation. Thus the observation error covariance corresponding to the resulting retrievals, $\mathbf{R}^r$, becomes $\mathbf{R}^r = \mathbf{H}^{-1}\mathbf{R}\mathbf{H}^{-1\top}$ (e.g., Joiner and da Silva, 1998). Here, $\mathbf{R}$ is the observation error covariance matrix of the radiance observations, which is four times the identity in this case. The resulting observation error covariance matrix for the retrievals at a fixed horizontal point is shown in Table 1. Retrieval observation errors at two different horizontal grid points are uncorrelated. However, those at the same horizontal location but at different model levels are correlated. The magnitude of the correlation between retrieval observation errors at a single horizontal point is shown in Table 2. The significant anti-correlation between errors at adjacent levels can be reduced by blending the observational information with a background state, based on the climatology of the model, similarly to the operational practice. However, when we tried to "regularize" the retrievals in this manner, despite the reduced correlations we were unable to obtain analysis errors as low as those we report in Section 4.3.

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

## 4 Results

### *4.1 Grid Point Observations*

We first assimilate the local "rawinsonde" observations to determine the optimal depth of the local regions. We consider a variety of depths for the local regions and assimilate all the local observations within these regions. In Figure 3, each vertical localization scheme is named by the maximum number of levels a local region could contain: 1 (dotted), 3 (dashed), 5 (dash-dotted), and 7 (solid). The depth of the local region for each model level and for each

localization scheme is given in Table 3. To determine which one of these schemes provides the best analysis, we plot the time averaged analysis rms error as a function of vertical level for zonal wind, panel (a), and temperature, panel (b). For both of these analysis variables, the analysis rms error is smallest when each local region is only one level deep. The analysis RMS error grows with the vertical depth of each local region. For the ensemble size of 20 and horizontal localization of $7 \times 7$ grid points, we find that these results hold over all analysis times and for all other model variables (results not shown).

### *4.2 Radiance Observations*

Figure 4 shows the time averaged rms error as a function of vertical level for a mixed observing network of rawinsonde and radiance observations and an observing network of rawinsonde observations alone (thick solid). We compare the cutoff-based and maximum-based selection schemes for the radiance observations. Specifically, in the cutoff-based selection scheme, we assimilate those channels of the radiance observations for which the observation operator has a nonzero weight (thin solid line), a weight of at least 0.125 (dash-dotted), and a weight of at least 0.25 (dashed). In the maximal-based selection scheme, we assimilate those channels for which the maximal weight 0.5 (dotted) within the local region.

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

### *4.3 Retrieval Observations*

We also update 1-level deep local regions by assimilating simulated rawinsonde observations and retrievals. In this case, we vary the number of retrieval observations assimilated in any local region and truncate the observation error covariance matrix, $\mathbf{R}$, to reflect the choice of observations. Specifically, we assimilate retrieval observations only within the local region (dotted), observations correlated to those in the local region with a correlation coefficient of magnitude at least 0.25 (dashed), 0.15 (dot-dashed), and 0.05 (thin solid line). The resulting average analysis rms error at each level is shown in Figure 5. The thick solid line shows the rms error from assimilating the rawinsondes alone. Assimilating those retrievals located in the local region has a positive impact on the analysis. However, we find that increasing the number of retrieval observation assimilated improves the analysis for temperature, zonal wind, and the other model variables. For these simulated retrievals, Table 2 indicates that the correlation coefficient between neighboring observation errors is roughly $-0.6$, suggesting that they are strongly anti-correlated. By assimilating observations only within the local region, we are neglecting these strong (anti-) correlations between observation errors and therefore not we do not obtain an optimal analysis.

When assimilating real retrieval observations, elements of the matrix $\mathbf{R}$ are unknown. Often, $\mathbf{R}$ is assumed to be a diagonal matrix, whose entries are tuned to optimize the analysis. Because our local region is only one level deep, when we assimilate only those retrievals located strictly within the local region, we in fact use only the diagonal entries of $\mathbf{R}$ when assimilating only those retrievals located strictly within the local region. Therefore, we tried scaling these diagonal entries by factors both larger and smaller than one, but found that the analysis was not significantly improved (results not shown). On the other hand, when we assimilate correlated retrievals from outside the local region, the results in Figure 5 do use nonzero off-diagonal error covariances. We repeated these experiments setting the off-diagonal entries in $\mathbf{R}$ to zero, and found that the analysis errors got larger but were significantly better than the errors obtained using only retrievals in the local region (results not shown). Thus, even when the correlations between retrieval errors cannot be quantified, including retrievals whose errors are correlated to those in the local region may improve the analysis.

In our experiments, the observation operator and corresponding error covariance are known perfectly for both the simulated radiances and retrievals. Therefore, the simulated

radiance and retrievals contain the same information and should provide similar analyses when using the same amount of that information. To verify whether this is the case for our data assimilation system, we compare the analysis rms error for the best results obtained from assimilating the simulated rawinsonde and radiance observations (using cutoff-based selection to assimilate those radiances with a weight of at least 0.125 in the local region, solid line) to those obtained from assimilating the simulated rawinsonde and retrieval observations (selecting observations in the local region and those retrieval observations with a coefficient of magnitude at least 0.15, dashed-line). Figure 6 shows the rms error as a function of vertical level. In both fields, the rms error is similar at all levels. Similar results hold for the other analysis variables (results not shown). The comparable accuracy of the analyses from assimilating radiances and retrievals suggests that in each case we are using the observation information consistently when we extend the local region for the satellite observations.

## 5 Conclusions

In this paper, we compare different vertical localization approaches when using an ensemble Kalman filter to assimilate a combination of simulated rawinsonde and satellite observations for the SPEEDY model. In Section 4 we compare the rms analysis error for each approach. We have also computed but do not include here forecast errors, which yield a similar comparison.

Using a 20 member ensemble, we find that the analysis error associated with assimilating simulated rawinsonde observations is smallest when using 1-level deep local regions. The dramatic increase in the analysis error from increasing the depth of local regions reflects the coarse vertical resolution of the seven level SPEEDY model. Increasing the depth of the local regions here by even one grid point in depth increases their extent by hundreds of hPa. In more realistic weather models, increasing the depth of the local regions does not cause the local regions to span such a large extent of the atmosphere. Therefore, for more realistic weather models, slightly increasing the depth of the local regions has only a slight impact on the quality of the analysis. However, dramatically increasing the depth of these local regions degrades the analysis in perfect model experiments with the NCEP GFS model (Szunyogh et al., 2005). Houtekamer et al. (2005); Whitaker et al. (2006); Szunyogh et al. (2007) also find that significant vertical localization is necessary to get the best analysis from rawinsonde observations.

The numerical experiments described in Section 4 suggest that when using vertical localization, a properly chosen cutoff-based selection for which radiance-like observations to assimilate is better than either maximum-based selection. With cutoff-based selection, we update the model state at a given location by assimilating radiances for which the value of the weighting function is "sufficiently" large in the local region from which rawindsonde observations are assimilated, rather than using only the radiances for which the maximum of the weighting function is in the local region (see Section 2.3). In our experiments with simulated rawindsonde and radiance observations, we obtain the least analysis error using cutoff-based selection and a 1-level deep local region. We find that maximum-based selection results in greater analysis error, whether we retain the 1-level deep local region or enlarge it. The advantage of cutoff-based selection is most significant for analysis variables other than the temperature. These variables depend on the observed radiances only through their correlations with the temperature.

We infer that for purporses of vertical localization, radiance-like observations should be considered wherever they depend significantly on the model state rather then only where they depend maximally on the model state. The numerical experiments here represent a simplified case, in which the observation operator is a known linear weighting function. Radiance observation operators are typically complicated nonlinear functions, of which the weighting function is only a component. Nonetheless, the weighting function can still be extracted and indicates the sensitivity of that observation to the model variables at a given location. One therefore can use cutoff-based selection to choose appropriate radiances for a given local region when using real data. We expect the benefit to the analysis to be greatest when the vertical distance over which the weighting functions are significant is comparable to or larger than the vertical localization distance.

Our numerical experiments suggest that one also should assimilate retrievals from a region that is larger than the local region used for the assimilation of point observations. We find that the number of observations selected for assimilation influences the analysis of the model variables more strongly for our simulated retrievals than for our simulated radiances. Assimilating observations with errors strongly correlated to those in the local region improves the analysis in all variables. The analyses obtained by assimilating the retrieval observations in this way have similar average rms errors to those obtained with the assimilation of the radiances at the optimal selection of the analysis parameters.

As we discussed in Section 3, the simple retrieval algorithm used in this study exaggerates

the correlations between errors of the retrieval observations over what they would be when using a typical retrieval algorithm. Nonetheless, the errors of the resulting retrievals would still have significant correlations. Accordingly, we expect that the analysis will benefit from including more retrievals than those in the local region for point observations, although the benefit may be less significant than for the simplified retrievals used in the numerical experiments here.

A further complication from retrieval observations arises because the correlations between their errors are often unknown. In practice, they are generally assumed to be uncorrelated and the diagonal entries of the observation error covariance matrix $\mathbf{R}$ are tuned to provide the optimal analysis. However, for the simplified retrieval observations used here, we observe that simply scaling the diagonal entries of $\mathbf{R}$ only slightly improves the analysis in some of the model variables. We gain the most significant advantage from assimilating those observations with strongly correlated errors simultaneously, even if we use a diagonal $\mathbf{R}$. Therefore, when retrieval error correlations are unknown, including additional retrieval observations suspected of having errors correlated errors to those in the local region may still improve the analysis.

In summary, spatial localization is essential to many ensemble-based data assimilation schemes. Typically the local regions used for ensemble-based Kalman filter schemes on operational models contain more levels than those used for the SPEEDY model. However, these regions are still small enough that relevant satellite observations can depend on the model variables outside the local region. The numerical experiments in Section 4 and the discussion of Section 2 suggest that one should localize the impact of observations differently depending on the observation type. Nonlocal observations do not necessarily need to be used in the analysis for all model variables on which they depend, but they need to be used in the analyses at the location of all model variables they are related to sufficiently strongly, either through the observation operator (as for satellite radiances) or through correlated errors (as for retrievals). The distance over which nonlocal observations influence the analysis is a parameter that can be tuned independently of the localization distance for point observations. In this paper, we have expressed this parameter as a cutoff for the radiance weighting function or the correlation between retrieval errors, but if these quantities are not known, one can tune the localization distance directly.

## 6 Acknowledgments

The authors thank Takemasa Miyoshi for bringing the SPEEDY model to the University of Maryland and adapting it to use for data assimilation. We also thank Takemasa Miyoshi and John Harlim for making their implementation of LETKF on the SPEEDY model available for the numerical experiments in this paper. We thank Chris Danforth and Junjie Liu for their assistance with the SPEEDY model. We also thank J.T. Halbert, Eugenia Kalnay, Hong Li, Ricardo Todling, Jeffrey Whitaker, and the two anonymous reviewers for their helpful suggestions. This work was supported by the Office of Naval Research (Physics), the National Science Foundation, Grant ATM034225, a National Oceanic and Atmospheric Administration THORPEX grant, and the NPOESS Integrated Program Office (IPO). The first author would also like to acknowledge the support of University of Maryland's Earth System Science Interdisciplinary Center and NASA Goddard Space Flight Center.

## Appendix A: Covariance Localization

An alternative to the grid point localization scheme in this paper is covariance localization (Anderson, 2003; Bishop et al., 2001; Hamill et al., 2001; Houtekamer and Mitchell, 2001; Whitaker and Hamill, 2002). This localization scheme limits the distance over which each observation affects the analysis by tapering the entries of the background matrix to zero beyond a fixed horizontal distance $L$ and vertical distance $V$. Each entry of the background covariance matrix $\mathbf{P}^b$ relates two model variables, and here distance is measured between the locations of these model variables. In a sequential assimilation scheme, the term $\mathbf{P}^b$ in Equation (3) then causes entries in the Kalman gain matrix $\mathbf{K}$ to be zero beyond the horizontal distance $L$ and vertical distance $V$ from the location of the observation(s) being assimilated.

Houtekamer et al. (2005) implement this approach by modifying Equation (7) such that

$$\mathbf{K} = (k-1)^{-1} \left[\rho_V \circ \rho_H \circ \left(\mathbf{X}^b \mathbf{Y}^{b\top}\right)\right] \left[(k-1)^{-1} \rho_V \circ \rho_H \circ \left(\mathbf{Y}^b \mathbf{Y}^{b\top}\right) + \mathbf{R}\right]^{-1}, \qquad (17)$$

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

With covariance localization, one can allow nonlocal observations to have a larger region of influence than point observations as follows. To use cutoff-based selection, as described in Section 2.3, for vertical localization, one should first determine an appropriate vertical localization distance $V$ for point observations. For nonlocal observations, increase $V$ to $\tilde{V}$, where $\tilde{V} - V$ is the distance over which the cutoff (for the weighting function in the case of satellite radiances or for correlation in the case of retrievals) is exceeded. Then, define a modified vertical localization operator $\tilde{\rho}_V$ that tapers covariances to zero beyond the vertical distance $\tilde{V}$ by scaling vertical distances by a factor of $V/\tilde{V}$. Each nonlocal observation must also be assigned a nominal location; for satellite radiances, one can use the maximum of the weighting function. Finally, replace $\rho_V$ by $\tilde{\rho}_V$ in Equation (17) when assimilating nonlocal observations. This will be simplest in a sequential scheme, assimilating nonlocal observations separately from point observations.

## Appendix B: Diagonalization

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

However, applying the eigenvalue decomposition makes it problematic to distinguish observations that are strongly correlated to the dynamics within a local region. To better understand this we recall that for observations which have only slightly correlated errors, the corresponding off diagonal entries of the observation error covariance matrix, $\mathbf{R}$, are small. For example, in the numerical experiments in Section 4, the $\mathbf{R}$ matrix for retrieval

observations at a given horizontal grid point is given in Table 1. Its structure suggests that an accurate analysis would result from assimilating only those observations with errors that are most strongly correlated to the errors of those observations located within the local region. However, the corresponding eigenvectors typically do not maintain this local structure. We have found that most entries of these eigenvectors have approximately the same magnitude. For example, for the experiments in Section 3, the matrix whose columns correspond to the eigenvectors of $\mathbf{R}$ is

$$\mathbf{S} = \begin{pmatrix} 0.24 & 0.44 & 0.53 & 0.48 & 0.33 & -0.28 & 0.21 \\ 0.36 & 0.47 & 0.24 & -0.12 & -0.39 & 0.52 & -0.40 \\ 0.44 & 0.30 & -0.25 & -0.51 & -0.20 & -0.39 & 0.45 \\ 0.48 & 0.00 & -0.44 & 0.00 & 0.63 & 0.00 & -0.43 \\ 0.44 & -0.30 & -0.25 & 0.51 & -0.20 & 0.39 & 0.45 \\ 0.36 & -0.47 & 0.24 & 0.12 & -0.39 & -0.52 & -0.40 \\ 0.24 & -0.44 & 0.53 & -0.48 & 0.33 & 0.28 & 0.21 \end{pmatrix}. \tag{18}$$

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

Nonetheless, some ensemble-based Kalman filter algorithms require diagonalization to assimilate observations with correlated errors. In these cases, observations can still be selected based on their correlation coefficients, as described in Section 2.3. The $\mathbf{R}$ matrix could then be truncated to reflect this choice of observations. Applying the diagonalization to the truncated $\mathbf{R}$ matrix would ensure that only those observations with errors most strongly correlated those of observations in the local region are assimilated to update the state in that region.

## REFERENCES

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

Anderson, J.L. 2003. A local least squares framework for ensemble filtering. *Monthly Weather Review* **131**, 634–642.

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

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

Evensen, G. 1994. Sequential data assimilation with a nonlinear quasi-geostropic model using Monte Carlo Methods to forecast error statistics. *Journal of Geophysical Research* **99**, 10143–10162.

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

Hamill, T., Whitaker, J., and Snyder, C. 2001. Distant-dependent filtering of background covariance estimates in an ensemble Kalman filter. *Monthly Weather Review* **129**, 2776–2790.

Harlim, J. and Hunt, B.R. 2007 A non-Gaussian ensemble filter for assimilating infrequent noisy observations. *Tellus A.* **59A**, In Press.

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

Houtekamer, P.L. and Mitchell, H.L. 2001. A sequential ensemble Kalman filter technique. *Monthly Weather Review* **129**, 123–137.

Houtekamer, P.L., Mitchell, H.L., Pellerin, G., Buehner, M., Charron, M, and co-authors 2005. Atmospheric data assimilation with the ensemble Kalman filter: Results with real observations. *Monthly Weather Review* **133**, 604–620.

Houtekamer, P.L. and Mitchell, H.L. 2006. Ensemble Kalman Filtering. *Quarterly Journal of the Royal Meteorological Society*, In Press.

Hunt, B.R., Kostelich, E.J., and Szunyogh, I. 2006. Efficient data assimilation for spatiotemporal chaos: a local ensemble transform Kalman filter. *Physica D.*, Submitted.

Joiner, J. and da Silva, A.M. 1998. Efficient methods to assimilate remotely sensed data based on information content. *Quarterly Journal of the Royal Meteorological Society* **124**, 1669–1694.

Kalman, R.E. 1960. A new approach to linear filtering and prediction problems. *Journal of Basic Engineering* **82D**, 35–45.

Kalman, R.E. and Bucy, R.S. 1961. New results in linear filtering and prediction theory. *Journal of Basic Engineering* **83D**, 95–107.

Kalnay, E. 2003. *Atmospheric Modelling, Data Assimilation, and Predictability*. Cambridge University Press, New York.

Keppenne, C. 1999. Data assimilation into a primitive-equation model with a parallel ensemble Kalman filter. *Monthly Weather Review* **128**, 1971–1981.

Kuhl, D., Szunyogh, I., Kostelich, E.J., Patil, D.J., Gyarmati, G., Oczkowski, M., Hunt, B.R., Kalnay, E., Ott, E., and Yorke, J.A. 2006. Assessing predictability with a local ensemble Kalman filter. *Journal of the Atmospheric Sciences*, Submitted.

Liou, K. 2002. *An Introduction to Atmospheric Radiation*. Academic Press, New York, Second Edition.

Miyoshi, T. 2005. *Ensemble Kalman filter experiments with a primitive-equation global model*. Ph.D. thesis, University of Maryland.

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

Oczkowski, M., Szunyogh, I., and Patil, D.J. 2005. Mechanisms for the development of locally low dimensional atmospheric dynamics. *Journal of the Atmospheric Sciences* **62**, 1135–1156.

Ott, E., Hunt, B.R., Szunyogh, I., Zimin, A., Kostelich, E.J., Corrazza, M., Kalnay, E., and Yorke, J.A. 2004. A local ensemble Kalman filter for atmospheric data assimilation. *Tellus* **56A**, 415–428.

Patil, D.J, Hunt, B.R., Kalnay, E., Yorke, J.A., and Ott, E. 2005. Local low dimensionality of atmospheric dynamics. *Physical Review Letters* **86**, 5878–5881.

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

Szunyogh, I., Kostelich, E.J., Gyarmati, G., Kalnay, E., Hunt, B.R., Ott, E., Satterfield, E. and Yorke, J.A. 2007. A local ensemble transform Kalman filter data assimilation system for the NCEP global model. *Tellus*, Submitted.

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

Whitaker, J.S., Hamill, T.M., Wei, X., Song, Y. and Toth, Z. 2006. Ensemble data assimilation with the NCEP global forecast system. *Monthly Weather Review*, Submitted.

## FIGURE CAPTIONS

**Figure 1.** This figure depicts the observation selection schemes used for nonlocal, radiance-like observations. The curve represents the weighting function at each model level. The rectangles represent several local regions around these model levels. In practice, these regions overlap, though not depicted as such in this figure. The shaded rectangles represent local regions that assimilate this observation.

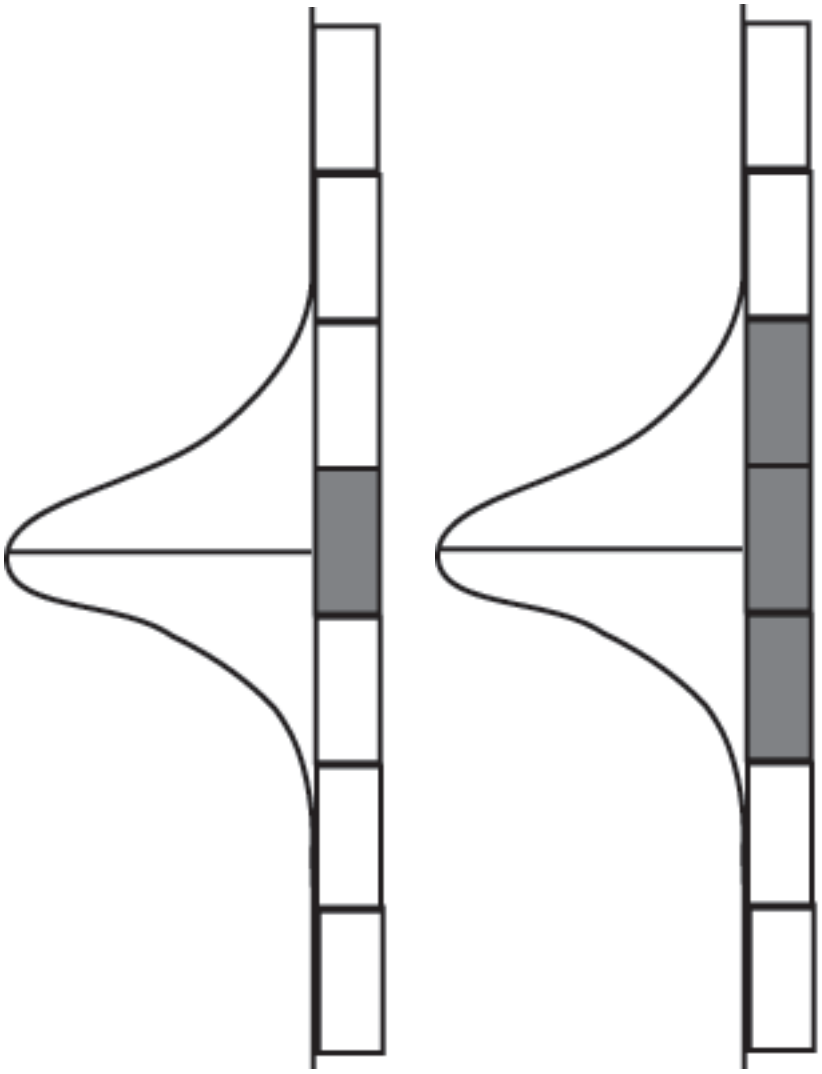
**Figure 2.** For the SPEEDY model experiments in this paper, we simulate rawinsonde observations at the model grid points closest to real rawinsonde locations at the 12Z synoptic time developed by Miyoshi (2005). Here, we show the locations of the model grid points of the rawinsonde stations for the $p_s$ observations. Similar locations are used for the other variables, and the observation density varies with height as for real rawinsonde observations.

**Figure 3.** We plot the average analysis RMS error as a function of vertical model level from assimilating simulated rawinsonde observations. The depth of the vertical local region contains at most 1 level (dotted), 3 levels (dashed), 5 levels (dash-dotted), and 7 levels (solid).

**Figure 4.** We plot the average analysis RMS error as a function of vertical model level in from assimilating simulated rawinsonde observations and satellite radiances. We select rawinsonde observations within a local region 1 level deep (thick solid) and we also assimilate the radiance observations if the value of the $\mathbf{h}$ operator there is greater or equal to 0.5 (dotted), 0.25 (dashed), 0.125 (dash-dotted), and 0.0625 (thin solid).

**Figure 5.** We plot the average analysis RMS error as a function of vertical model level from assimilating simulated rawinsonde observations alone (thick solid) and rawinsondes with satellite retrievals. We assimilate retrievals within the local region (dotted) and also those retrieval observations outside the local region with a correlation coefficient of at least 0.25 (dashed), 0.15 (dash-dotted), and 0.05 (thin solid).

**Figure 6.** We compare the best result from assimilating the simulated rawinsonde and satellite radiance observations (solid) and the simulated rawinsonde and satellite retrieval observations (dashed) by plotting the average analysis RMS error as a function of vertical model level



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

**Figure 1.** This figure depicts the observation selection schemes used for nonlocal, radiance-like observations. The curve represents the weighting function at each model level. The rectangles represent several local regions around these model levels. In practice, these regions overlap, though not depicted as such in this figure. The shaded rectangles represent local regions that assimilate this observation.

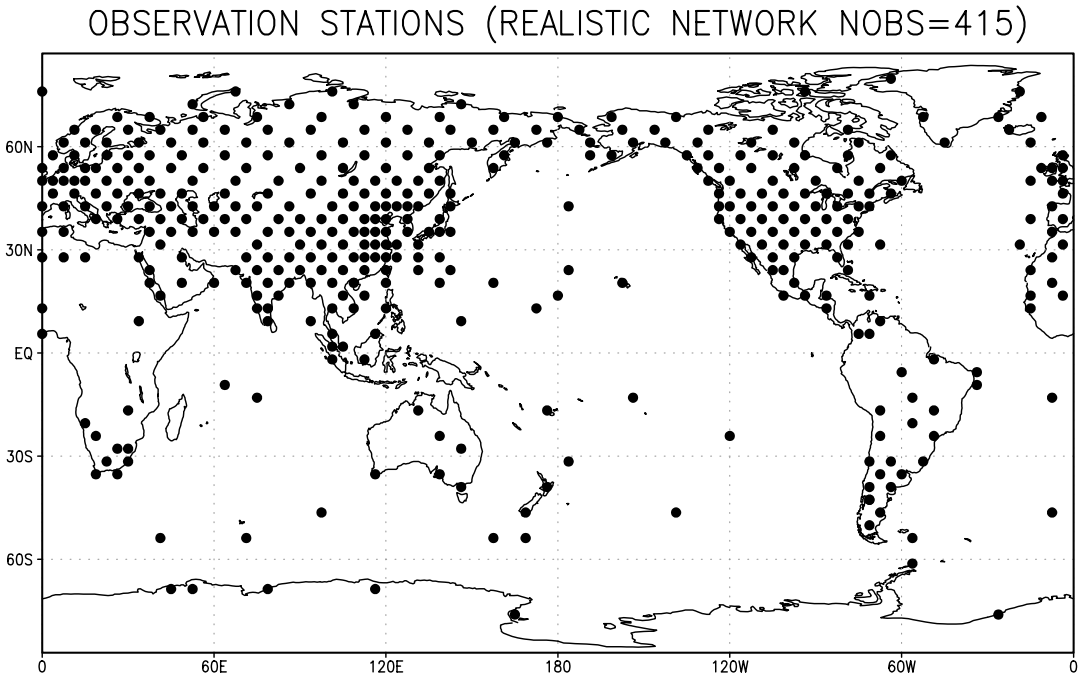
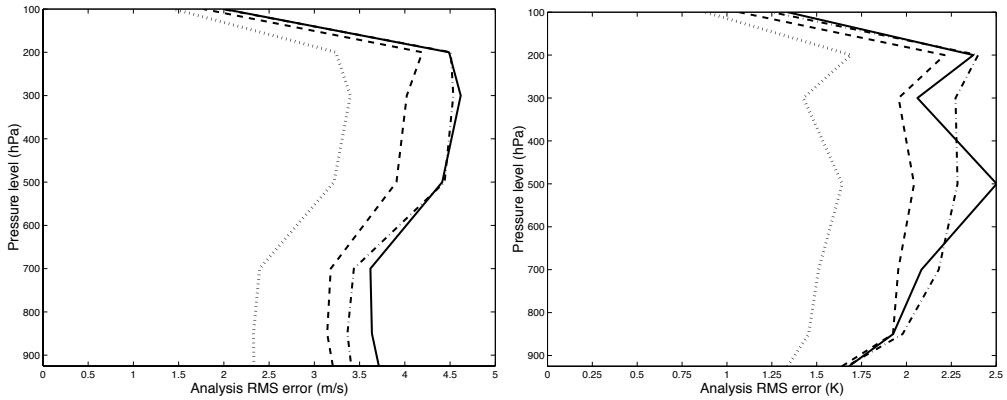
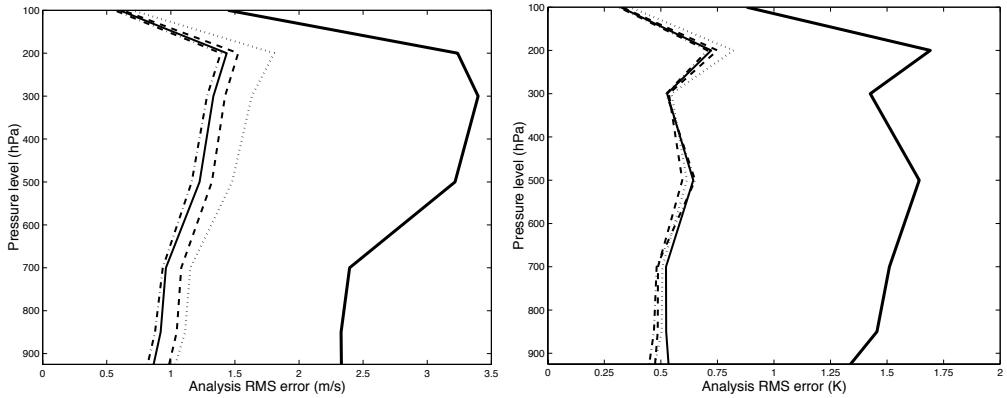



**Figure 2.** For the SPEEDY model experiments in this paper, we simulate rawinsonde observations at the model grid points closest to real rawinsonde locations at the 12Z synoptic time developed by Miyoshi (2005). Here, we show the locations of the model grid points of the rawinsonde stations for the $p_s$ observations. Similar locations are used for the other variables, and the observation density varies with height as for real rawinsonde observations.



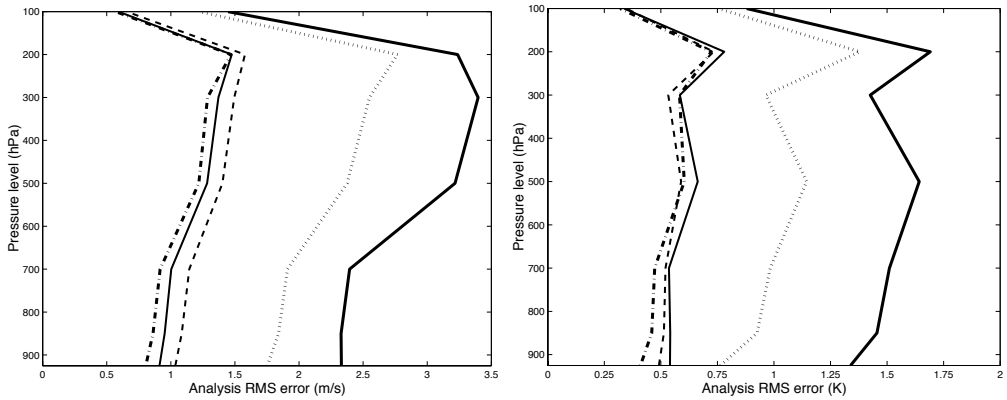

(a) Zonal Wind ($u$) (b) Temperature ($T$)

**Figure 3.** We plot the average analysis RMS error as a function of vertical model level from assimilating simulated rawinsonde observations. The depth of the vertical local region contains at most 1 level (dotted), 3 levels (dashed), 5 levels (dash-dotted), and 7 levels (solid).



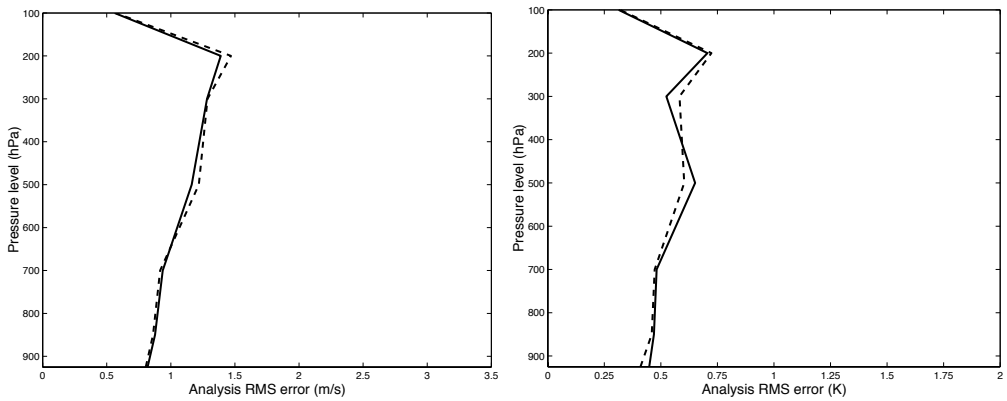

(a) Zonal Wind ($u$) (b) Temperature ($T$)

**Figure 4.** We plot the average analysis RMS error as a function of vertical model level in from assimilating simulated rawinsonde observations and satellite radiances. We select rawinsonde observations within a local region 1 level deep (thick solid) and we also assimilate the radiance observations if the value of the **h** operator there is greater or equal to 0.5 (dotted), 0.25 (dashed), 0.125 (dash-dotted), and 0.0625 (thin solid).




(a) Zonal Wind ($u$) (b) Temperature ($T$)

**Figure 5.** We plot the average analysis RMS error as a function of vertical model level from assimilating simulated rawinsonde observations alone (thick solid) and rawinsondes with satellite retrievals. We assimilate retrievals within the local region (dotted) and also those retrieval observations outside the local region with a correlation coefficient of at least 0.25 (dashed), 0.15 (dash-dotted), and 0.05 (thin solid).




(a) Zonal Wind ($u$) (b) Temperature ($T$)

**Figure 6.** We compare the best result from assimilating the simulated rawinsonde and satellite radiance observations (solid) and the simulated rawinsonde and satellite retrieval observations (dashed) by plotting the average analysis RMS error as a function of vertical model level.

## TABLE CAPTIONS

**Table 1.** This table gives the observation error covariance for the simulated retrieval observations at a fixed horizontal location.

**Table 2.** The correlation between the errors of observations at a single horizontal point and model levels $k$ and $l$ is given by the entry in this table corresponding to these model levels.

**Table 3.** We show the depth of the local regions at each model level for the vertical localization schemes 1, 3, 5, and 7. Each scheme is numbered based on the maximum number of levels a local region could contain.

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

**Table 1.** This table gives the observation error covariance for the simulated retrieval observations at a fixed horizontal location.

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

**Table 2.** The correlation between the errors of observations at a single horizontal point and model levels $k$ and $l$ is given by the entry in this table corresponding to these model levels.

| | **Vertical Localization** | | | |
|---|---|---|---|---|
| **Model Level** | **Scheme 1** | **Scheme 3** | **Scheme 5** | **Scheme 7** |
| | (depth in number of model levels) | | | |
| 1 | 1 | 1 | 1 | 1 |
| 2 | 1 | 3 | 3 | 3 |
| 3 | 1 | 3 | 5 | 5 |
| 4 | 1 | 3 | 5 | 7 |
| 5 | 1 | 3 | 5 | 5 |
| 6 | 1 | 3 | 3 | 3 |
| 7 | 1 | 1 | 1 | 1 |

**Table 3.** We show the depth of the local regions at each model level for the vertical localization schemes 1, 3, 5, and 7. Each scheme is numbered based on the maximum number of levels a local region could contain.