

Qin Xu<sup>1\*</sup>, Li Wei<sup>2</sup>, Jidong Gao<sup>1</sup>, Qingyun Zhao<sup>3</sup>, Kang Nai<sup>2</sup> and Shun Liu<sup>4</sup><sup>1</sup>NOAA/National Severe Storms Laboratory, Norman, Oklahoma<sup>2</sup>Cooperative Institute for Mesoscale Meteorological Studies, University of Oklahoma<sup>3</sup>Marine Meteorology Division, Naval Research Laboratory, Monterey, California<sup>4</sup>National Centers of Environmental Prediction and I. M. Systems Group, Rockville, Maryland

## 1. Introduction

It has been well recognized that using a Gaussian function with a synoptic-scale de-correlation length to model the background error covariance in data assimilation can inadvertently hamper the ability of the analysis to assimilate mesoscale structures. As a remedy to this problem, a superposition of Gaussians has been used for variational data assimilation at NCEP with increased computational cost (Purser et al. 2003), but the mesoscale features are still overly smoothed and inadequately resolved in the analyzed incremental fields even in areas covered by remotely sensed high-resolution observations, such as those from operational weather radars (Liu et al. 2005). This raises an important question on how to optimally assimilate high-resolution observations, such as those remotely sensed from radars and satellites, on the mesoscale and storm scale. Ideally and theoretically, if the background error covariance is exactly known and perfectly modeled in data assimilation, then all different types of observations can be optimally analyzed in a single batch at a single step. However, since the background error covariance is usually mostly unknown and often crudely modeled, a multi-step approach could be more effective and efficient than the single-step approach for assimilating various types of observations (including remotely sensed high-resolution observations) into a high-resolution mesoscale model.

Previously, Xie et al. (2013) proposed a sequential multi-step variational analysis approach for a multiscale analysis system with observations reused in each step in a fashion similar to Barnes successive correction scheme. These authors noted that the background error covariance should change with different steps to incorporate scale-dependent information (like the Barnes successive correction scheme) but left this issue to future studies for further improvements. Gao et al. (2013) adopted a real-time variational data assimilation system with a multi-step approach to analyze observations of different spatial resolutions sequentially from large to small scale with the background error covariance tuned/specified empirically in each step, so the issue concerning how to objectively estimate the background error covariance in each step was still unaddressed. By decomposing the cost function to allow the background error covariance estimated separately for two different spatial scales, Li et al. (2015) formulated a multi-scale variational scheme (which is not necessarily sequential) with the background error variance assumed to be known for each scale [see their (47) and (48)] and with the background error decorrelation length set simply to the minimum wavelength

resolved in each scale [see their (52) and (53)], but the issue on how to estimate the background error covariance was again left unaddressed. We believe that this issue is very important for a multi-step approach, as the analysis is largely determined by the background error covariance in each step.

For the traditional single-step variational analysis, the background error covariance can be estimated from time series (used as an ensemble) of innovation (that is, observation minus background in the observation space) by using the innovation method (Hollingsworth and Lonnberg 1986; Hollingsworth and Lonnberg 1986; Xu et al. 2001; Xu and Wei 2001, 2002) or from time series of difference between two model forecasts verifying at the same time by using the NMC method (Parrish and Derber 1992; Derber and Bouttier 1999). The background error covariance estimated by the above method can be readily used for the variational analysis in the first step of a multi-step approach. In each subsequent step, however, the background error covariance should be re-estimated [or re-computed as shown in (1b)] for the updated background, that is, the analysis from the previous step. The innovation method can be modified and used for the re-estimation if the observations used in current step are not previously used and thus are new and independent of the new background and if the following two conditions are also satisfied (as required by the innovation method). (i) The time series of new innovation (that is, new observation minus new background) still satisfy the ergodicity assumption (and thus can be used as an ensemble). (ii) The statistic structures of the new innovations remain to be horizontally homogeneous and isotropic. These two conditions often cannot be satisfied, as they require that the distribution of the observations used in each step is not only horizontally homogeneous (or nearly so) but also largely fixed in the time series. Thus, the innovation method must be simplified with reduced conditions to re-estimate the new background error variance only. In particular, as shown in (3.5) of Xu et al. (2015), by using the local spatial mean (instead of temporal mean) as the ensemble mean, the background error variance can be estimated efficiently as a smooth function of space from the new innovation field (rather than an ensemble collected from a time series) in each step of a multi-step approach. The background error de-correlation length, however, was still not objectively estimated but specified empirically in each step of the multi-step radar wind analysis system of Xu et al. (2015).

This study aims to explore a new multi-step approach with the background error covariance estimated efficiently and updated optimally (or nearly so) in each subsequent step. In particular, a two-step variational method is developed for idealized one-dimensional cases in section 2 and illustrated by numerical examples in section 3. Extensions of this approach for real-data applications are discussed in section 4, followed by conclusions in section 5.

---

\* Corresponding author address: Qin Xu, National Severe Storms Laboratory, 120 David L. Boren Blvd., Norman, OK 73072-7326; E-mail: Qin.Xu@noaa.gov

## 2. Update background error covariance

When the variational analysis is formulated optimally base on the Bayesian estimation theory (see chapter 7 of Jazwinski, 1970), the background state vector  $\mathbf{b}$  is updated to the analysis state vector  $\mathbf{a}$  by

$$\mathbf{a} = \mathbf{b} + \mathbf{B}\mathbf{H}^T(\mathbf{H}\mathbf{B}\mathbf{H}^T + \mathbf{R})^{-1}\mathbf{d}, \quad (1a)$$

and the background error covariance matrix  $\mathbf{B}$  is updated to the analysis error covariance matrix  $\mathbf{A}$  by

$$\mathbf{A} = \mathbf{B} - \mathbf{B}\mathbf{H}^T(\mathbf{H}\mathbf{B}\mathbf{H}^T + \mathbf{R})^{-1}\mathbf{H}\mathbf{B}, \quad (1b)$$

where  $\mathbf{R}$  is the observation error covariance matrix,  $\mathbf{H}$  is the (linearized) observation operator,  $\mathbf{d} = \mathbf{y} - \mathbf{H}(\mathbf{b})$  is the innovation vector,  $\mathbf{y}$  is the observation vector,  $\mathbf{H}(\cdot)$  denotes the observation operator and  $\mathbf{H}$  is the linearized  $\mathbf{H}(\cdot)$ . Here, (1) provides the precise formulation for updating  $\mathbf{B}$  to  $\mathbf{A}$  in each step of a multi-step variational analysis, but the required computation is impractically expensive for operational applications. Thus, the issue encountered here is how to simplify (1b) so that  $\mathbf{A}$  can be estimated efficiently with much reduced computational cost. This issue will be attacked in this section by formulating a two-step variational method for idealized one-dimensional cases. A related issue is whether this efficiently estimated  $\mathbf{A}$  can make the two-step analysis more accurate than the single-step analysis of observations of different resolutions if the number of iterations is not sufficiently large and thus the analysis is not truly optimal (which is often the case in operational variational data assimilation due to the computational constraints even if the background error covariance is assumed to be accurately modeled by a superposition of Gaussians). This issue will be examined in the next section with numerical examples.

For simplicity, we consider the following two types of observations: (i) coarse-resolution observations uniformly or quasi-uniformly distributed over the one-dimensional analysis domain of length  $D$  along the  $x$ -coordinate, and (ii) high-resolution observations over a fraction of the analysis domain. The coarse-resolution observations are analyzed in the first step over the entire domain to update the background state  $\mathbf{b}$  to  $\mathbf{a}$  and error covariance  $\mathbf{B}$  to  $\mathbf{A}$ . The high-resolution observations are analyzed in the second step over a nested domain using the above updated  $\mathbf{a}$  and  $\mathbf{A}$  as the new background state and error covariance. To complete such a two-step variational analysis, the key issue is to estimate  $\mathbf{A}$  efficiently. This is facilitated by the following properties:

I. For a single observation, say, at  $x = x_m$ , the inverse matrix  $(\mathbf{H}\mathbf{B}\mathbf{H}^T + \mathbf{R})^{-1}$  in (1b) reduces to  $(\sigma_b^2 + \sigma_o^2)^{-1}$ , so the  $ij^{\text{th}}$  element of  $\mathbf{A}$  is simply given by

$$A_{ij} = \sigma_b^2 [C_b(x_i - x_j) - \gamma_b C_b(x_i - x_m) C_b(x_m - x_j)], \quad (2)$$

where  $\gamma_b = \sigma_o^2 / (\sigma_b^2 + \sigma_o^2)$ ,  $\sigma_b^2$  (or  $\sigma_o^2$ ) is the background (or observation) error variance,  $C_b(x)$  is the background error correlation function,  $x_i$  (or  $x_j$ ) denotes the  $i^{\text{th}}$  (or  $j^{\text{th}}$ ) point in the analysis space  $R^N$ , and  $N$  is the number of analysis grid points.

II. For  $M$  uniformly distributed coarse-resolution observations over  $D$  in  $x$ , (1b) can be transformed into the following spectral form in the wavenumber space:

$$\mathbf{S}_a = \mathbf{S} - \mathbf{S}\mathbf{P}_{MN}^T(\mathbf{P}_{MN}\mathbf{S}\mathbf{P}_{MN}^T + \nu\mathbf{C})^{-1}\mathbf{P}_{MN}\mathbf{S}, \quad (3)$$

where  $\mathbf{S}_a \equiv \mathbf{F}_N\mathbf{A}\mathbf{F}_N^H$ ,  $\mathbf{S} \equiv \mathbf{F}_N\mathbf{B}\mathbf{F}_N^H$  (or  $\mathbf{C} \equiv \mathbf{F}_M\mathbf{R}\mathbf{F}_M^H$ ) is a diagonal matrix in  $R^N$  (or  $R^M$ ),  $\mathbf{F}_N$  (or  $\mathbf{F}_M$ ) is the normalized discrete Fourier transformation (DFT) matrix in  $R^N$  (or  $R^M$ ),  $\nu \equiv N/M (> 1)$ , and  $\mathbf{P}_{MN}$  is a  $M \times N$  matrix. When  $\nu$  is an odd integer,  $\mathbf{P}_{MN}$  is simply given by  $(\mathbf{I}, \dots, \mathbf{I})$  where  $\mathbf{I}$  is the unit matrix in  $R^M$ . When  $\nu$  is not an odd integer,  $\mathbf{P}_{MN}$  is still largely composed of  $\mathbf{I}$  and the detailed formulation can be found in (21) of Xu (2011). Since  $M < N$ ,  $\mathbf{S}_a$  is not diagonal but its nonzero off-diagonal elements are sparse and negligibly small. Using (3),  $\mathbf{S}_a$  can be easily computed from  $\mathbf{S}$  and  $\mathbf{C}$ . The diagonal part of  $\mathbf{S}_a$  can be then transformed efficiently by the inverse DFT (or simply the inverse discrete cosine transformation) back to the physical space in the form of  $\sigma_e^2 C_a(x_i - x_j)$  to estimate the  $ij^{\text{th}}$  element of  $\mathbf{A}$ , where  $\sigma_e^2$  (= constant) and  $C_a(x)$  denote the estimated analysis error variance and correlation function, respectively. When the coarse-resolution observations are not exactly uniform but quasi-uniform over the analysis domain,  $\mathbf{A}$  still can be estimated from  $\sigma_e^2 C_a(x_i - x_j)$  approximately.

III. Setting  $x_i = x_j$  in (2) gives to  $A_{ii} = \sigma_b^2 - \Delta\sigma_b^2(x_i)$ , where

$$\Delta\sigma_b^2(x_i - x_m) \equiv \gamma_b \sigma_b^2 C_b^2(x_i - x_m) \quad (4)$$

is the error variance reduction produced by analyzing a single observation at  $x = x_m$ . Note that this error variance reduction becomes negligibly small as  $|x_i - x_m|$  increases and renders  $C_b^2(x_i - x_m) \ll 1$ . Thus, if the spacing  $\Delta x_{co}$  of the  $M$  coarse-resolution observations is sufficiently large to render  $C_b^2(\Delta x_{co}) \ll 1$ , then the error variance reduction produced by analyzing the  $M$  observations can be estimated by

$$\Delta\sigma^2(x_i) \leq \Delta\sigma_{bm}^2(x_i) \equiv \sum_m \Delta\sigma_b^2(x_i - x_m), \quad (5)$$

where  $\sum_m$  denotes the summation over  $m$  for the  $M$  observations. When  $\Delta x_{co}$  is not sufficiently large and thus  $C_b^2(\Delta x_{co})$  is not very small,  $\Delta\sigma^2(x_i)$  decreases below the upper bound  $\Delta\sigma_{bm}^2(x_i)$  on the right-hand side of (5). Now consider that  $\sigma_b^2 C_b(x)$  is updated to  $\sigma_e^2 C_a(x)$  by analyzing the  $M$  coarse-resolution observations according to the above property 2, the error variance reduction produced by analyzing an additional single observation, say, at  $x = x_m$  can be estimated by

$$\Delta\sigma_e^2(x_i - x_m) \equiv \gamma_e \sigma_e^2 C_a^2(x_i - x_m), \quad (6)$$

where  $\gamma_e = \sigma_e^2 / (\sigma_e^2 + \sigma_o^2)$ . Here, (6) is derived similarly to (4), but it provides a rough estimate of the error variance reduction produced by analyzing the last observation if the  $M$  observations are sequentially analyzed. A weighted average of (4) and (6) can be formulated by

$$\Delta\sigma_{bc}^2(x_i - x_m) \equiv (1 - w_e)\Delta\sigma_b^2(x_i) + w_e\Delta\sigma_e^2(x_i), \quad (7)$$

where  $w_e = C_b^2(\Delta x_{co})$ . When  $\Delta x_{co}$  is sufficiently large to render  $C_b^2(\Delta x_{co}) \ll 1$ , (7) recovers (4). Using (7), the error variance reduction produced by analyzing the  $M$  coarse-resolution observations can be formulated as a function of  $x_i$  by

$$\Delta\sigma^2(x_i) \approx \sum_m \Delta\sigma_{bc}^2(x_i - x_m) - \sigma_b^2 + \sigma_e^2 - c, \quad (8)$$

where  $c$  is a constant to be determined. Substituting this into  $\sigma_a^2(x_i) \equiv \sigma_e^2 - \Delta\sigma^2(x_i)$  gives the following estimated form of analysis error variance:

$$\sigma_a^2(x_i) = \sigma_e^2 - \sum_m \Delta\sigma_{eb}^2(x_i - x_m) + c. \quad (9)$$

Note that  $\sum_i \sigma_a^2(x_i)/N$  is the spatially averaged error variance computed from (9) and it should match the value of  $\sigma_e^2$  estimated from the above property II. This gives

$$\begin{aligned} c &= \sum_m \sum_i \Delta\sigma_{eb}^2(x_i - x_m)/N \\ &= (1 - w_e)\gamma_b\sigma_b^2 \sum_m \sum_i C_b^2(x_i - x_m)/N \\ &\quad + w_e\gamma_e\sigma_e^2 \sum_m \sum_i C_a^2(x_i - x_m)/N \\ &= [(1 - w_e)\gamma_b\sigma_b^2 \sum_i C_a^2(x_i) + w_e\gamma_e\sigma_e^2 \sum_i C_a^2(x_i)]M/N, \end{aligned}$$

where  $\sum_i$  denotes the summation over  $i$  for all the  $N$  analysis grad points. The analysis error covariance function is then refined from  $\sigma_e^2 C_a(x_i - x_j)$  to

$$\sigma_a(x_i)\sigma_a(x_j)C_a(x_i - x_j), \quad (10)$$

where  $\sigma_a^2(x_i)$  is estimated in (9).

### 3. Numerical examples

In this section, we use the same observational data source (that is, the radial velocities scanned by the NSSL phased array radar for the Oklahoma squall line on 2 June 2004) and the same model-produced background field as those described in section 5.2 of Xu (2007) but with the following treatments: (i) The analysis domain length is set to  $D = N\Delta x = 110.16$  km with  $N = 9 \times 51 = 459$  and  $\Delta x = 0.24$  km, where  $\Delta x$  is the analysis grid resolution and is set to be the same as the original radar radial-velocity observation resolution. (ii) The  $M (= 9)$  coarse-resolution observations are generated by uniformly or quasi-uniformly thinning the original 459 observations over the entire domain of length  $D$  (as shown by the purple + signs in Fig. 1 for the  $M$  uniformly thinned coarse-resolution observations), while the  $M' (= 76)$  high-resolution observations are taken from the remaining original observations in a nested domain of length  $D/6$  (as shown by the cyan x signs connected by cyan dashed lines in Fig. 1 for uniform coarse-resolution observations). (iii) A linear spatial variation is estimated from the background field and is subtracted from both the observations and background field. This last treatment can facilitate the periodic extension of the analysis but is unnecessary for a real-data application, because the innovation and analysis increment are not affected by this treatment and they are treated as spatially homogeneous random fields in the variational analysis and thus already suitable for periodic extension.

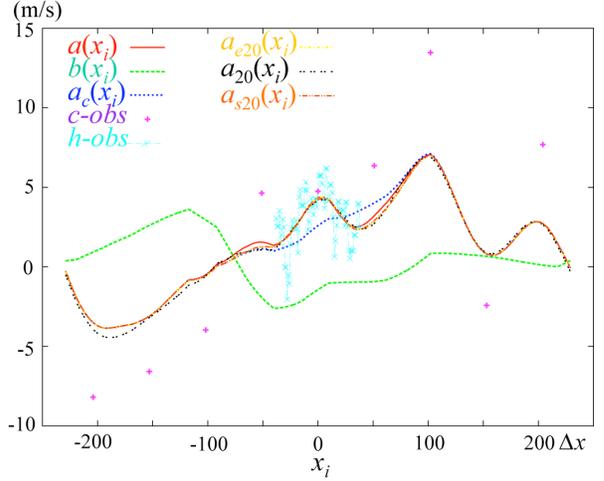


Fig. 1. Coarse-resolution observations (c-obs) plotted by purple + signs, high-resolution observations (h-obs) plotted by cyan x signs connected by dashed cyan lines, background field  $b(x_i)$  plotted by dotted green curve, optimally analyzed benchmark field  $a(x_i)$  plotted by solid red curve, single-step analyzed field  $a_{20}(x_i)$  with 20 iterations plotted by dashed black curve, and first-step analyzed field  $a_e(x_i)$  plotted by dotted blue curve. The dash-dotted yellow (or brown) curve plots the two-step analyzed field  $a_{20}(x_i)$  with 20 iterations and with  $\mathbf{B}$  updated by  $\sigma_e^2 C_a(x_i - x_j)$  [or  $\sigma_a(x_i)\sigma_a(x_j)C_a(x_i - x_j)$  in (10)].

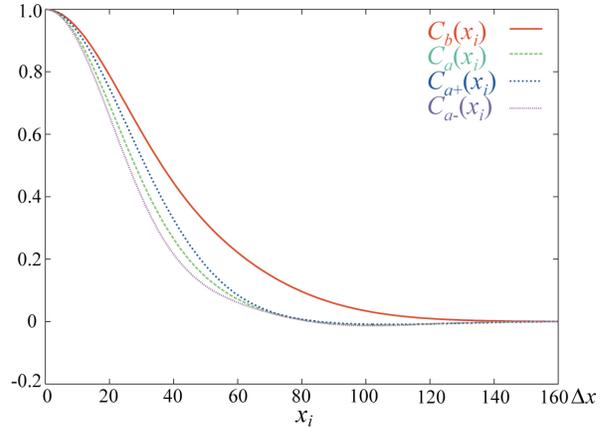


Fig. 2.  $C_b(x_i)$  plotted by solid red curve and  $C_a(x_i)$  plotted by dotted green curve. The dotted blue (or purple) curve plots  $C_{a+}(x_i)$  [or  $C_{a-}(x_i)$ ] – the correlation structure intercepted across the point marked by the + (or -) sign in Fig. 5a.

The background error variance is set to be the same as the observation error variance, that is,  $\sigma_b^2 = \sigma_o^2 (= 2.5^2 \text{ m}^2\text{s}^{-2})$ . The background error correlation function is modeled by a periodic extension of the following double Gaussians:

$$C_b(x) = \sum_i 0.6 \exp[-x^2/2L^2] + 0.4 \exp[-2x^2/L^2] \quad (11)$$

with  $L = 42\Delta x$ , and the periodic extension is formulated in the same way as in (1b) of Xu and Wei (2011). The structure of  $C_b(x)$  is plotted by the solid red curve in Fig. 2. The

background error power spectrum  $S(k_i)$  can be easily computed from  $\sigma_b^2 C_b(x)$  by the discrete cosine transformation [see (12)-(13) of Xu 2011], where  $k_i = i\Delta k$  denotes the  $i^{\text{th}}$  discrete wavenumber and  $\Delta k \equiv 2\pi/D$  is the minimum resolvable wavenumber. The computed power spectrum (that forms the diagonal matrix  $\mathbf{S}$ ) is plotted by the solid red curve in Fig. 3.

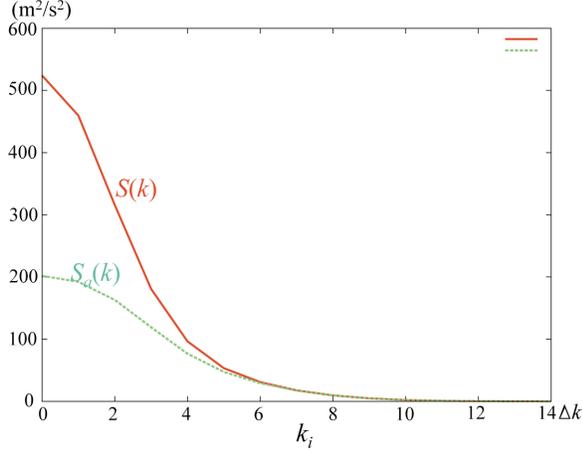


Fig. 3.  $S(k_i)$  plotted by solid red curve and  $S_b(k_i)$  plotted by dotted green curve.

There are two sets of observations. Both sets consist of  $M$  coarse-resolution observations and  $M'$  high-resolution observations, but the  $M$  coarse-resolution observations in the first (or second) set are uniformly (or quasi-uniformly) distributed. The observation errors are assumed to be spatially uncorrelated, so  $\mathbf{R} = \sigma_o^2 \mathbf{I}$  in  $R^{M+M'}$  for all the observations in each set. Since  $M + M'$  ( $= 9 + 76$ ) is not large, the inverse matrix  $(\mathbf{H}\mathbf{B}\mathbf{H}^T + \mathbf{R})^{-1}$  can be easily and accurately computed. Thus, the optimal analysis can be obtained by applying (1a) to all the  $M + M'$  observations together and then used as the benchmark to evaluate the accuracies of the analyses obtained from the control and test experiments for each set.

The control experiments also analyze all the  $M + M'$  observations together, but the analyses are performed by applying the standard conjugate-gradient descent algorithm with limited numbers of iterations (to mimic operational applications) to minimize the following cost function:

$$J = \mathbf{c}^T \mathbf{B} \mathbf{c} + |\mathbf{H}\mathbf{B}\mathbf{c} - \mathbf{d}|^T / \sigma_o^2, \quad (12)$$

where  $\Delta \mathbf{a} \equiv \mathbf{a} - \mathbf{b}$  is the analysis increment and is related to the transformed control vector  $\mathbf{c}$  by  $\Delta \mathbf{a} = \mathbf{B}\mathbf{c}$ . The test experiments analyze the  $M$  coarse-resolution observations in the first step, and then the  $M'$  high-resolution observations in the second step. In the first (or second) step, the analysis is performed by applying the standard conjugate-gradient descent algorithm with limited number of iterations to minimize the same form of cost function as in (12) but formulated for the  $M$  coarse-resolution (or  $M'$  high-resolution) observations with  $\mathbf{B}$  constructed by  $\sigma_b^2 C_b(x)$  (or updated to  $\mathbf{A}$ ).

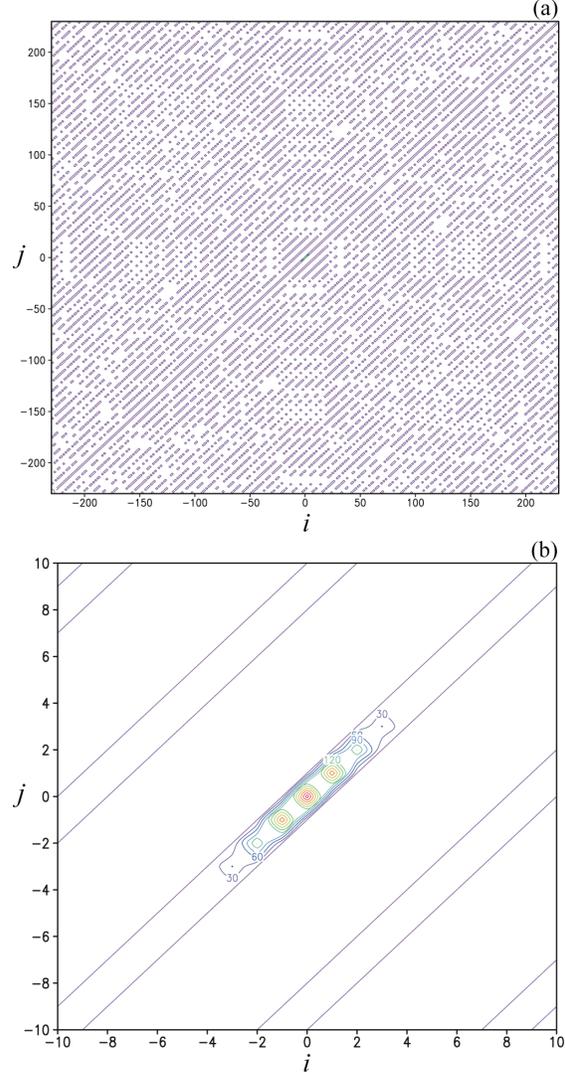
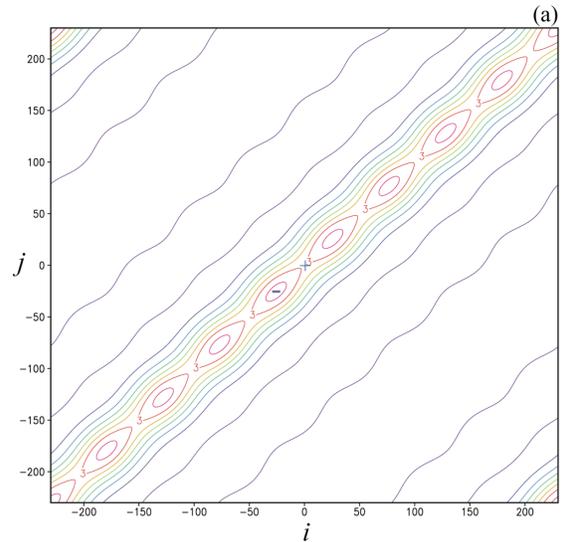


Fig. 4. (a) Full-matrix structure of  $\mathbf{S}_a$ . (b) Zoom-in structure of  $\mathbf{S}_a$ . The color contours plot the element value in  $\text{m}^2\text{s}^{-2}$ .



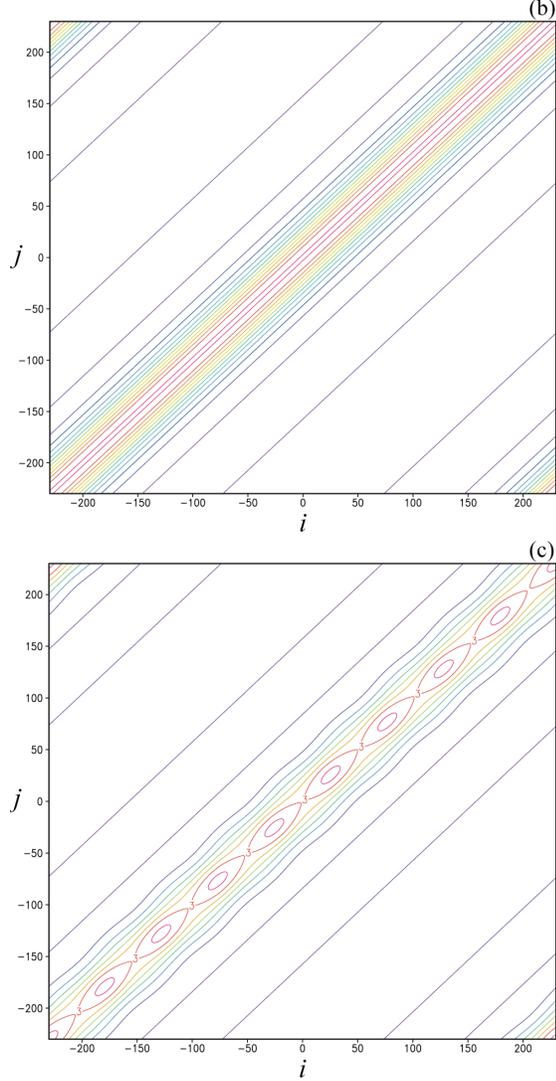


Fig. 5. (a) Structure of the benchmark  $\mathbf{A}$ . (b) Structure of  $\mathbf{A}$  estimated by  $A_{ij} = \sigma_e^2 C_a(x_i - x_j)$ . (c) Structure of  $\mathbf{A}$  estimated by  $\sigma_a(x_i) \sigma_a(x_j) C_a(x_i - x_j)$  in (10). The color contours plot the value of  $A_{ij}$  in  $\text{m}^2 \text{s}^{-2}$ .

After the  $M$  coarse-resolution observations are analyzed in the first step,  $\mathbf{S}$  is updated to  $\mathbf{S}_a$  according to (3). As shown by the full-matrix structure of  $\mathbf{S}_a$  in Fig. 4a,  $\mathbf{S}_a$  is not diagonal (due to the fact of  $M < N$ ) but its nonzero off-diagonal elements are sparse and negligibly small. The diagonal elements of  $\mathbf{S}_a$  are also negligibly small outside the center diagonal segment, as shown by the zoom-in structure of  $\mathbf{S}_a$  in Fig. 4b. Using (3), the diagonal part of  $\mathbf{S}_a$  can be easily computed from  $\mathbf{S}$  and  $\mathbf{C}$ . The analysis error power spectrum  $S_a(k_i)$  estimated by the diagonal part of  $\mathbf{S}_a$  is plotted by the dotted green curve in Fig. 3. In comparison with the solid red curve plotted for  $S(k_i)$  in Fig. 3, this dotted green curve shows that the background error reduction produced by the first-step analysis is largest for  $k_i = 0$ , decreases rapidly as  $k_i$  increases and becomes nearly zero as  $k_i > k_5$ . The inverse discrete cosine transformation of  $S_a(k_i)$  gives  $\sigma_e^2 C_a(x_i)$ . As shown by the dotted green curve for  $C_a(x_i)$

versus the solid red curve for  $C_b(x_i)$  in Fig. 2, the error correlation function is narrowed and thus the de-correlation length is reduced as  $C_b(x_i)$  is updated by  $C_a(x_i)$ . This is simply because the background errors are reduced by the first-step analysis mostly in long-wave structures as shown by the change of error power spectrum from  $S(k_i)$  to  $S_a(k_i)$  in Fig. 3.

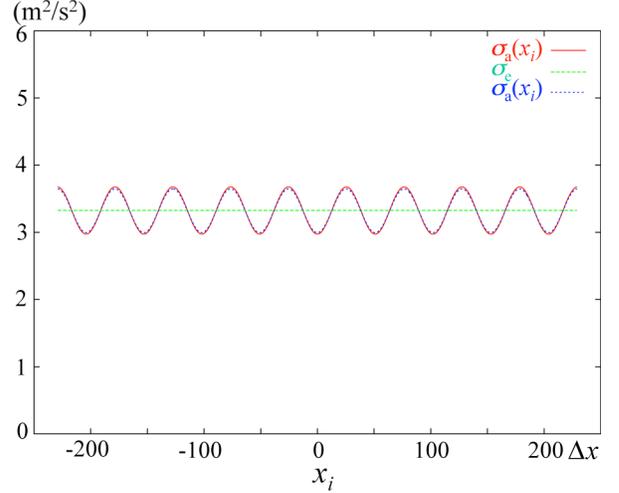


Fig. 6. True analysis error variance  $\sigma_a^2(x_i)$  (that is, the contour value along the diagonal line in Fig. 5a) plotted by solid red curve, and analysis error variance  $\sigma_a^2(x_i)$  estimated by (9) (that is, the contour value along the diagonal line in Fig. 5c) plotted by dotted blue curve. The dotted green line plots the constant value of  $\sigma_e^2$ .

Fig. 5a shows the structure of the benchmark  $\mathbf{A}$  that is precisely computed by the inverse DFT of  $\mathbf{S}_a$ . Fig. 5b shows the structure of  $\mathbf{A}$  estimated by  $A_{ij} = \sigma_e^2 C_a(x_i - x_j)$  from the property II in section 2. This estimated  $\mathbf{A}$  has the same coarse-grain structure as the benchmark  $\mathbf{A}$  in Fig. 5a but it cannot capture the detailed variations along the diagonal line and adjacent lines. Fig. 5c shows the structure of  $\mathbf{A}$  estimated by  $A_{ij} = \sigma_a(x_i) \sigma_a(x_j) C_a(x_i - x_j)$  in (10) from the property III in section 2. This estimated  $\mathbf{A}$  not only has the same coarse-grain structure as the benchmark  $\mathbf{A}$  but also captures the dominant variations along the diagonal line and adjacent lines. In particular, as shown in Fig. 6, the analysis error variance  $\sigma_a^2(x_i)$  estimated by (9) (that is, the contour value along the diagonal line in Fig. 5c) is very close to the true analysis error variance (that is, the contour value along the diagonal line in Fig. 5a). The dotted blue (or purple) curve in Fig. 2 shows the correlation structure intercepted horizontally across the point marked by the + (or -) sign in Fig. 5a. Note that the + (or -) sign in Fig. 5a is collocated with a coarse-resolution observation (or is between two adjacent coarse-resolution observations). This explains why the dotted blue (or purple) curve is slightly wider (or narrower) than the dotted green curve. When  $\mathbf{A}$  is estimated by (10), the error correlation structure is still uniform modeled by  $C_a(x)$ . The error correlation structure in the benchmark  $\mathbf{A}$  is not but nearly uniform and its variations are bounded between the dotted blue and purple curves in Fig. 2.

The analysis produced by the control experiment applied to the first set (9 uniform coarse-resolution observations plus 76 high-resolution observations) with 20 iterations is shown by the dashed black curve in Fig. 1. This curve is quite close to the benchmark optimal analysis shown by the solid red curve in Fig. 1, and the evaluated error is shown by the solid red curve in Fig. 7. The analysis produced in the first step by the test experiment applied to the 9 uniform coarse-resolutions observations (in the first data set) with 20 iterations is shown by the dotted blue curve in Fig. 1. This analysis is almost identical to the benchmark optimal analysis obtained by applying (1a) directly and only to the 9 uniform coarse-resolution observations. The dash-dotted yellow (or brown) curve in Fig. 1 shows the analysis produced by the test experiment applied to the remaining high-resolution observations in the second step with 20 iterations and with  $\mathbf{B}$  updated by  $\sigma_e^2 C_a(x_i - x_j)$  in the property II [or  $\sigma_a(x_i)\sigma_a(x_j)C_a(x_i - x_j)$  in (10)]. These two curves are almost identical to each other and they are very close to the solid red curve for the benchmark optimal analysis. The error evaluated by the difference of the dash-dotted yellow (or brown) curve with respect to the benchmark solid red curve in Fig. 1 is plotted by the dotted green (or blue) curves in Fig. 7. As shown, the analysis error is reduced slightly in the test experiment if  $\mathbf{B}$  is updated by  $\sigma_a(x_i)\sigma_a(x_j)C_a(x_i - x_j)$  in (10) instead of  $\sigma_e^2 C_a(x_i - x_j)$  in the property II.

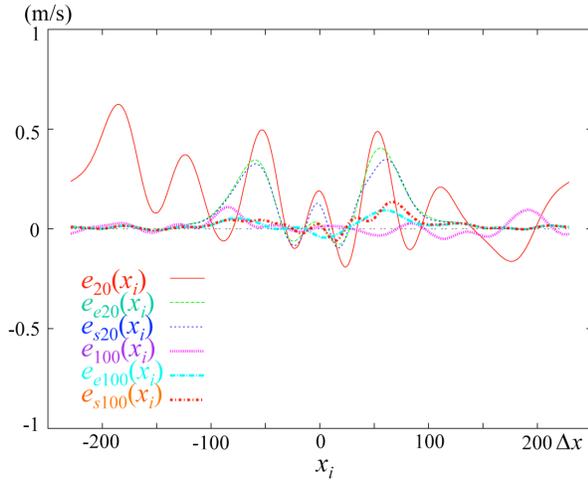


Fig. 7. Error  $e_{20}(x_i)$  of single-step analysis with 20 iterations plotted by solid red curve and error  $e_{100}(x_i)$  of single-step analysis with 100 iterations plotted by dotted purple curve. error  $e_{20}(x_i)$  of two-step analysis with 20 iterations plotted by The dotted green (or blue) curve plots the error of the two-step analysis  $e_{e20}(x_i)$  [or  $e_{s20}(x_i)$ ] with 20 iterations and with  $\mathbf{B}$  updated by  $\sigma_e^2 C_a(x_i - x_j)$  [or  $\sigma_a(x_i)\sigma_a(x_j)C_a(x_i - x_j)$  in (10)]. The dotted cyan (or brown) curve plots the error of the two-step analysis  $e_{e100}(x_i)$  [or  $e_{s100}(x_i)$ ] with 100 iterations and with  $\mathbf{B}$  updated by  $\sigma_e^2 C_a(x_i - x_j)$  [or  $\sigma_a(x_i)\sigma_a(x_j)C_a(x_i - x_j)$  in (10)].

When the number of iterations is increased from 20 to 100 in the control and test experiments applied to the first set of observations, the analysis errors are reduced significantly in all the tree experiments, as shown by the dotted purple curve (or dotted cyan and brown curves) for the control experiment

(or the two test experiments). In this case, the error reduction is more significant in the control experiment than in the two test experiments, but the analysis produced by the control experiment is still slightly less accurate than the analyses produced by the test experiments. When the number of iterations is further increased to 200, all the analyses become nearly identical to the benchmark optimal analysis, and the analysis produced by the control experiment becomes more accurate than the analyses produced by the test experiments.

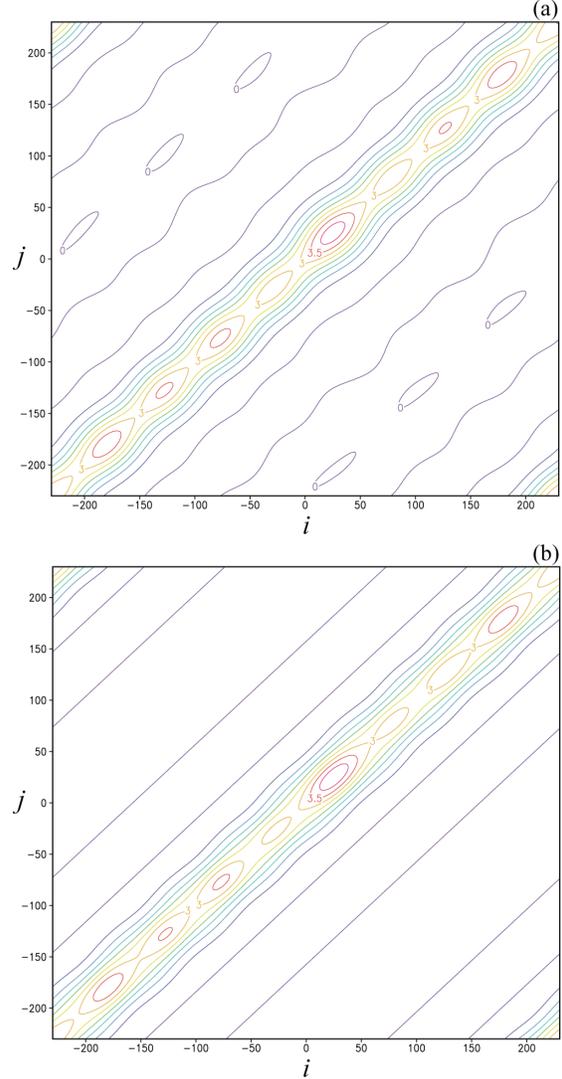


Fig. 8. As in Fig. 5a and Fig. 5c but for non-uniform coarse-resolution observations in the second set.

The above results remain qualitatively the same as the control and test experiments are applied to the second set of observations. In this case, the structure of the benchmark  $\mathbf{A}$  is computed precisely from (1b) and is shown in Fig. 8a, while Fig. 8b shows the matrix structure of  $\mathbf{A}$  estimated by  $A_{ij} = \sigma_a(x_i)\sigma_a(x_j)C_a(x_i - x_j)$  in (10) from the property III in section 2. This estimated  $\mathbf{A}$  not only has the same coarse-grain structure as the benchmark  $\mathbf{A}$  but also captures the dominant variations along the diagonal line and adjacent

lines. This feature is the same as seen from Fig. 5 for the results obtained from the uniform coarse-resolution observations in the first set. Again, as shown in Fig. 9, the analysis error variance  $\sigma_a^2(x_i)$  estimated by (9) (that is, the contour value along the diagonal line in Fig. 8b) is very close to the true analysis error variance (that is, the contour value along the diagonal line in Fig. 8a). The analysis errors are evaluated for the control experiment and test experiments in the same way as those for the first set of observations. As shown in Fig. 10, these errors have the same features as those shown and compared in Fig. 7.

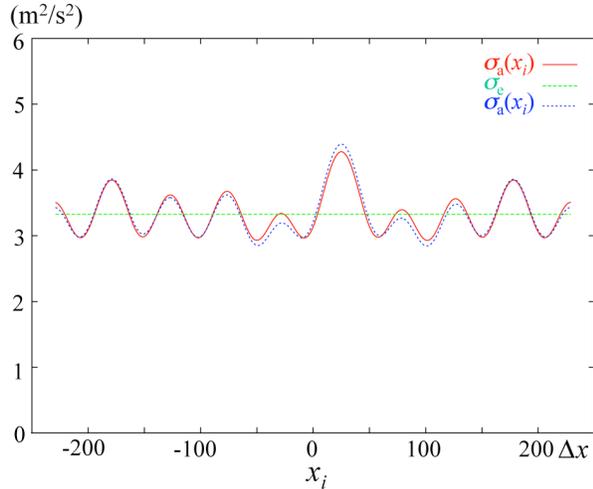


Fig. 9. As in Fig. 6 but for the second set observations.

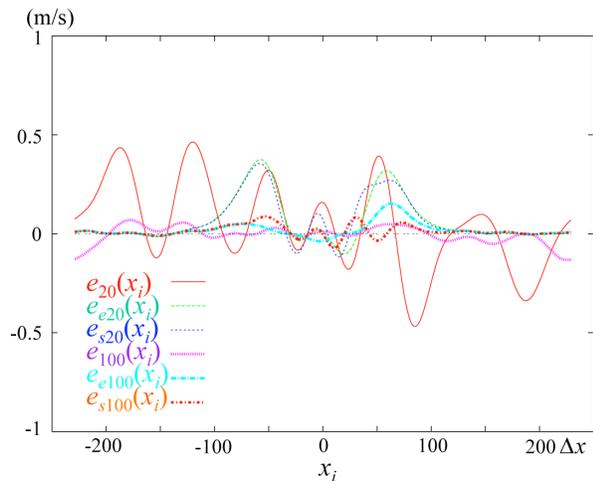


Fig. 10. As in Fig. 7 but for the second set of observations.

#### 4. Real-data applications

The property I in (2) and property III formulated by (9)–(10) can be readily extended and applied to a single observation in three-dimensional space. The property II formulated in (3) in the one-dimensional wavenumber space can be extended to the two-dimensional horizontal wavenumber space similarly to those derived in section 2.3 of Xu (2011) and demonstrated in section 3 of Xu and Wei

(2011), while the error correlation structure in the vertical direction may be kept the same and not updated after the coarse-resolution observations are analyzed in the first step. With the above extensions, the two-step approach developed in section 2 can be applied to observations of different horizontal resolutions in the three-dimensional space to improve the existing multi-scale variational data assimilation methods (Xie et al. 2013; Gao et al. 2013; Xu et al. 2015).

#### 5. Conclusions

In this paper, a two-step variational method is developed for analyzing observations of different spatial resolutions. The effectiveness of this approach is demonstrated for one-dimensional cases. In particular, the results obtained in sections 2 and 3 show, at least for the one-dimensional cases, that the analysis error covariance  $\mathbf{A}$  can be estimated efficiently and used to update the background error covariance  $\mathbf{B}$  after the coarse-resolution observations are analyzed in the first step. This can facilitate the two-step approach and make the two-step analysis more accurate than the single-step analysis of observations of different resolutions when the number of iterations is not sufficiently large and thus the single-step analyses is not truly optimal. Constrained by computational resources and capabilities, limited numbers of iterations are often used in variational data assimilation for operational applications. The analyses are thus not truly optimal and often overly smooth mesoscale features, especially when observations of different resolutions are analyzed in a single step, even if the background error covariance is accurately modeled. To improve this, the two-step method developed in this paper can and will be extended for real data assimilation applications. Specifically, as explained in section 4, the two-step method can be extended to analyze observations of different horizontal resolutions in the three-dimensional space more effectively and efficiently than the conventional single-step approach. The detailed formulations for these extensions are currently being developed and will be tested in our continued future studies.

#### Acknowledgments

The research work was supported by the ONR Grant N000141410281 and NSF grant AGS-1341878 to the University of Oklahoma (OU). Funding was also provided to CIMMS by NOAA/Office of Oceanic and Atmospheric Research under NOAA-OU Cooperative Agreement #NA17RJ1227, U.S. Department of Commerce.

#### REFERENCES

- Gao, J., T. M. Smith, D. J. Stensrud, C. Fu, K. Calhoun, K. L. Mnaross, J. Brogden, V. Lakshmanan, Y. Wang, K. W. Thomas, K. Brewster, and M. Xue, 2013: A realtime weather-adaptive 3DVAR analysis system for severe weather detections and warnings. *Wea. Forecasting*, **28**, 727–745.
- Hollingsworth, A., and P. Lonnberg, 1986: The statistical structure of short-range forecast errors as determined from radiosonde data. Part I: The wind field. *Tellus*, **38A**, 111–136.
- Jazwinski, A. H. 1970. *Stochastic Processes and Filtering Theory*. Academic Press, San Diego, 376 pp.

- Li, Z., J. C. McWilliams, K. Ide, and J. D. Farrara, 2015: A multiscale variational data assimilation scheme: formulation and illustration. *Mon. Wea. Rev.*, (in press)
- Liu, S., M. Xue, J. Gao, and D. Parrish, 2005: Analysis and impact of super-obbed Doppler radial velocity in the NCEP grid-point statistical interpolation (GSI) analysis system. Extended abstract, *17<sup>th</sup> Conf. Num. Wea. Pred.*, Washington DC, 13A-4.
- Lonnberg, P., and A. Hollingsworth, 1986: The statistical structure of short-range forecast errors as determined from radiosonde data. Part II: The covariance of height and wind errors. *Tellus*, **38A**, 137–161.
- Purser, R. J., W.-S. Wu, D. F. Parrish, and N. M. Roberts, 2003: Numerical aspects of the application of recursive filters to variational statistical analysis. part II: spatially inhomogeneous and anisotropic general covariances. *Mon. Wea. Rev.*, **131**, 1536–1548.
- Xie, Y., S. Koch, J. McGinley, S. Albers, P. E. Bieringer, M. Wolfson, and M. Chan, 2011: A space–time multiscale analysis system: a sequential variational analysis approach. *Mon. Wea. Rev.*, **139**, 1224–1240.
- Xu, Q., L. Wei, K. Nai, S. Liu, R. M. Rabin, and Q. Zhao, 2015a: A radar wind analysis system for nowcast applications. *Advances in Meteorology*. vol. **2015**, Article ID 264515, 13 pages.
- Xu, Q., L. Wei and K. Nai, 2015b: Analyzing vortex winds in radar observed tornadic mesocyclones for nowcast applications. *Wea. and Forecasting*, (in press).
- Xu, Q., and L. Wei, 2001: Estimation of three-dimensional error covariances. Part II: Analysis of wind innovation vectors. *Mon. Wea. Rev.*, **129**, 2939–2954.
- Xu, Q., and L. Wei, 2002: Estimation of three-dimensional error covariances. Part III: Height-wind error correlation and related geostrophy. *Mon. Wea. Rev.*, **130**, 1052–1062.
- Xu, Q., L. Wei, A. Van Tuyl, and E. H. Barker, 2001: Estimation of three-dimensional error covariances. Part I: Analysis of height innovation vectors. *Mon. Wea. Rev.*, **129**, 2126–2135.