Trade-Offs between Measurement Accuracy and Resolutions in Configuring Phased-Array Radar Velocity Scans for Ensemble-Based Storm-Scale Data Assimilation

HUIJUAN LU
Research Center for Numerical Prediction, State Key Laboratory of Severe Weather, Chinese Academy of Meteorological Sciences, Beijing, China, and Cooperative Institute for Mesoscale Meteorological Studies, University of Oklahoma, Norman, Oklahoma

QIN XU
NOAA/National Severe Storms Laboratory, Norman, Oklahoma

(Manuscript received 7 April 2008, in final form 30 September 2008)

ABSTRACT

Assimilation experiments are carried out with simulated radar radial-velocity observations to examine the impacts of observation accuracy and resolutions on storm-scale wind assimilation with an ensemble square root filter (EnSRF) on a storm-resolving grid (\( \Delta x = 2 \text{ km} \)). In this EnSRF, the background covariance is estimated from an ensemble of 40 imperfect-model predictions. The observation error includes both measurement error and representativeness error, and the error variance is estimated from the simulated observations against the simulated "truth." The results show that the analysis is not significantly improved when the measurement error is overly reduced (from 4 to 1 m s\(^{-1}\)) and becomes smaller than the representativeness error. The analysis can be improved by properly coarsening the observation resolution (to 2 km in the radial direction) with an increase in measurement accuracy and further improved by properly enhancing the temporal resolution of radar volume scans (from every 5 to 2 or 1 min) with a decrease in measurement accuracy. There can be an optimal balance or trade-off between measurement accuracy and resolutions (in space and time) for configuring radar scans, especially phased-array radar scans, to improve storm-scale radar wind analysis and assimilation.

1. Introduction

A phased-array weather radar has been constructed at the National Severe Storms Laboratory in Norman, Oklahoma. This establishes the first National Weather radar test bed equipped with the solid-state phased-array antenna, which is also the only phased-array radar of this kind at present time dedicated to weather radar research (Forsyth et al. 2005; Zrnic et al. 2007). One of the primary research goals is to explore the full advantages of the phased-array electronically controlled agile beam scan capabilities and make the best possible uses of these capabilities to improve numerical analysis and prediction of severe storms and other hazardous weather conditions. To this end, there is an important and yet very challenging issue concerning how to optimally design phased-array scan strategies for the intended meteorological applications. This is the motivation of this study.

One of the major advantages of the phased-array radar is the capability to adaptively scan weather phenomena at higher temporal resolution (1 min or less per volume scan) than is possible by the conventional weather radars [such as the Weather Surveillance Radar-1988 Doppler (WSR-88D)] that use mechanically steered reflector antennas (5 min or more per volume scan). With this capability, the enhanced temporal resolution can benefit storm-scale radar wind analysis and data assimilation (Qiu and Xu 1996; Shapiro et al. 2003; Xue et al. 2006). As the phased-array radar controls the beam direction electronically, it can transmit microwave pulses swiftly to many sufficiently different beam directions and thus acquire uncorrelated samples more efficiently than the WSR-88D to estimate the spectral moments (i.e., reflectivity, radial velocity, and spectrum width in level-II data) in each resolution volume and to
reduce the uncertainties (caused by random meteorological scatterers) of the estimated spectral moments. This can yield a gain in (level II) data accuracy or resolution (in space or time). However, regardless of this and other possible gains, there is always a limit on the number of independent samples acquirable per unit time for a given phased-array radar (or any other radar). This imposes an overall constraint on radar measurement accuracy and resolutions. Under this constraint, an increase in accuracy (or resolution) will cause a decrease in resolution (or accuracy) and, by the same token, a resolution increase in one spatial or temporal dimension will cause a resolution decrease in another dimension (without compromising the accuracy). Because of this, there should be a proper balance or trade-off between measurement accuracy and resolutions when radar scans are configured to improve storm-scale radar wind analyses and data assimilation. This problem will be studied in this paper. In particular, we will perform assimilation experiments with model-simulated radar radial-velocity observations of a supercell storm to examine the impacts of observation accuracy and resolutions on the assimilated storm winds by the ensemble square root filter (EnSRF) (Whitaker and Hamill 2002). As two important variants of the ensemble Kalman filter (EnKF) (Evensen 1994; Tippett et al. 2003), the EnSRF and EnKF with perturbed observations (Burgers et al. 1998) have been used to assimilate model-simulated and real radar observations with encouraging results [Snyder and Zhang 2003; Zhang et al. 2004; Dowell et al. 2004; Tong and Xue 2005 (hereinafter TX05), 2008; Xue et al. 2006; Lu 2007; Xu et al. 2008a; Gao and Xue 2008].

When the EnSRF is applied to simulated radar observations to address the issues concerned in this study, we need to consider both radar measurement error and representativeness error in the simulated observations. The radar measurement error will be mimicked by random noise. The representativeness error is an error due to the limited model resolution in the calculation of the observation operator. This is an error in the observation innovation but can be attributed to its observation part as a representativeness error. The total observation error is the sum of the measurement error and representativeness error, and the total error variance will be estimated from the simulated observations against the simulated “true” model state.

The theoretical analysis and numerical examples presented in Xu (2007) suggest that there can be a significant degree of information redundancy in radar observations if the observations are too dense to be resolved by an optimal analysis with a given background covariance. As the background covariance is parameterized by a smooth function on the model grid in a three- or four-dimensional variational data assimilation system, the covariance is often too smooth or too coarse to allow the analysis to absorb fine structures in radar observations. For storm-scale radar data assimilation with an EnKF such as the EnSRF considered in this study, the background covariance is computed from an ensemble of model-predicted fields on a storm-resolving model grid and thus contains much finer structures than a parameterized smooth covariance. However, as the covariance is estimated from a small ensemble of imperfect-model predictions (40 members in this study because of computational constraints), the estimated covariance may contain significant errors and spurious noisy structures. For the EnSRF analysis with such a covariance, the conventional (level II) radar data may still have excessive accuracy (or excessive spatial resolution), so the accuracy (or spatial resolution) may need to be reduced in favor of enhancing the temporal resolution. Such a trade-off can be beneficial to storm-scale radar wind assimilation, and this will be demonstrated in this paper.

As the phased-array radar is able to scan adaptively, trade-offs between measurement accuracy and resolutions can also be made adaptively in space and time according to the weather phenomena being observed. This complexity, however, is beyond the consideration of this study. Only simple trade-offs will be considered in this study. Except for those in favor of increasing the temporal resolution beyond 5 min (see section 4c), most of the trade-offs are applicable to both the phased-array and conventional weather radars. Since the objective considered in this study is to improve storm wind analyses, only radar radial-velocity observations will be assimilated by the EnSRF and the quality of the assimilation will be evaluated by the root-mean-square (RMS) error of the analyzed wind field.

The paper is organized as follows: the next section describes the model-simulated radar observations and assimilation system. Section 3 performs assimilation experiments with observations of different measurement accuracies and examines the impact of measurement accuracy on the EnSRF analysis. Section 4 performs assimilation experiments with observations of different resolutions and examines how the EnSRF analysis is affected by trade-offs between measurement accuracy and resolutions. Conclusions follow in section 5.

2. Simulated observations and assimilation system

a. Prediction model and simulated storm

The Advanced Regional Prediction System (ARPS; Xue et al. 2001) is used for both simulation and analysis,
and the classic 20 May 1977 Del City, Oklahoma, supercell storm case is simulated as in TX05. The ARPS is a fully compressible and nonhydrostatic atmospheric prediction system. For the version used in this study, it contains 12 prognostic state variables, including wind components \(u\), \(v\), and \(w\), potential temperature \(\theta\), pressure \(p\), the mixing ratios for water vapor \(q_u\), cloud water \(q_c\), rainwater \(q_r\), cloud ice \(q_i\), snow \(q_s\), and hail \(q_h\) plus the turbulent kinetic energy (TKE) used by the 1.5-order subgrid-scale turbulent closure scheme. All of these prognostic state variables except for TKE are subject to subgrid-scale turbulent closure scheme. All of these transport state variables are subject to the turbulent kinetic energy (TKE) used by the 1.5-order subgrid-scale turbulent closure scheme. All of these transport state variables are parameterized using the three-category ice scheme of Lin et al. (1983).

The physical domain is \(64 \times 64 \times 16\) km\(^3\) for all experiments. The model grid for the fine-resolution true state comprises \(131 \times 131 \times 35\) grid points (with horizontal grid spacing \(\Delta x = \Delta y = 0.5\) km and vertical grid spacing \(\Delta z = 0.5\) km), while the grid is coarsened to \(35 \times 35 \times 35\) grid points (with \(\Delta x = \Delta y = 2\) km and \(\Delta z = 0.5\) km) for all the assimilation experiments. Thus, the forecast model is imperfect. The fine-resolution true state (with \(\Delta x = \Delta y = 0.5\) km) is generated by a model simulation initialized from a modified Del City sounding as in TX05. The CAPE of the sounding is about 3300 J kg\(^{-1}\), and a 4-K ellipsoidal thermal bubble centered at \(x = 48, y = 16, \) and \(z = 1.5\) km, with a radius of 10 km in horizontal directions and 1.5 km in vertical, is used to initiate the storm. The fine-resolution true state is then averaged in each coarsened grid box (\(\Delta x = \Delta y = 2\) km) to obtain the coarsened-resolution true model state.

\subsection*{b. Simulated radar observations}

The temporal resolution of the operational WSR-88D radar scans in a storm scan mode is approximately 5 min per volume, and this is used as the benchmark temporal resolution in this study. All the simulated observations are in the radar polar coordinates, and each volume scan contains 14 tilts at the elevation angles of 0.48°, 1.45°, 2.4°, 3.3°, 4.3°, 5.2°, 6.2°, 7.5°, 8.7°, 10.0°, 12°, 14.0°, 16.7°, and 19.5°. On each tilt, the benchmark horizontal resolutions are 1° in the azimuthal direction and 0.5 km in the radial direction. Here, the range gate resolution is coarser than that of the real WSR-88D radial velocity (0.25 km) but it is consistent with the resolution of the fine-resolution true state from which the simulated observations are generated. The radial-velocity data are assumed to be available only in the precipitation region, where the reflectivity exceeds 10 dBZ. The reflectivity is calculated from the hydrometeor mixing ratios from the truth simulation as in TX05. The ground-based radar is located at the southwest corner of the computational domain, that is, at the coordinate origin, for all experiments.

Each radial velocity observation, \(V_r\), is simulated by

\[ V_r = u \cos \alpha \sin \beta + v \cos \alpha \cos \beta + w \sin \alpha + \varepsilon, \]  

where \(\alpha\) is the elevation angle, \(\beta\) is the azimuth angle of radar beam, and \((u, v, \) and \(w)\) are the three components of the fine-resolution true velocity at the observation point interpolated by trilinear interpolation. The \(\varepsilon\) represents random measurement errors that have a normal distribution with zero mean and a specified standard deviation \((\sigma)\). Strictly speaking, the simulated \(V_r\) observation in Eq. (1) should be an averaged value (weighted by radar antenna gain function and reflectivity) over the radar resolution volume. This complication, however, is not considered here (because it requires that the true reflectivity distribution be simulated accurately within each resolution volume and this is beyond the capability of the microphysics in the ARPS). Nevertheless, since the fine-resolution “truth” (with \(\Delta x = \Delta y = 0.5\) km and \(\Delta z = 0.5\) km) has about the same volume resolution as the radar resolution volume (with \(\Delta x \approx 0.25\) km, \(\Delta y = \Delta z = 1°\)) at the radial range \((r \approx 30-40\) km) where the simulated storm is “observed” from the radar, the representativeness error with respect to the fine-resolution truth should be much smaller than that with respect to the coarsened-resolution truth and thus can be neglected when the simulated \(V_r\) observation is generated from the fine-resolution truth in Eq. (1). Note that the radial velocity contribution from precipitation is assumed to be accurately estimated from radar observed reflectivity and thus excluded from the simulated radial-velocity observation in Eq. (1).

In the real radar data assimilation, the radial-velocity contribution from precipitation is often estimated empirically (Dowell et al. 2004), and the estimation error could be significantly non-Gaussian and biased. This complication is not considered in Eq. (1) for the simulated radial-velocity observations.

It is well known that radar velocity measurements can sometimes be contaminated by range folding, velocity aliasing, ground or sea surface clutters from sidelobes and/or ducted anomaly propagation, and clutter from nonmeteorological flying objects. In this paper, these gross errors are assumed to be removed by data quality controls (e.g., Gong et al. 2003; Zhang et al. 2005; Liu et al. 2005). In this case, random errors in radar radial-velocity measurements (level-II data) are caused mainly by inhomogeneities of velocity and reflectivity within each sampling volume, and the measurement error standard deviation may have a lower bound of about \(\sigma_m = 1\) m s\(^{-1}\) in a fair-weather turbulent boundary layer (see section 9.1.1.2 of Doviak and Zrnic 1993). Recently, by reformulating the conventional innovation
method, Xu et al. (2007) estimated radar radial-velocity observation error and background velocity error covariance from time series of radar radial-velocity innovation (observation minus independent background), and the estimated observation error standard deviation is \( \sigma = 4 \) m s\(^{-1}\). This estimate includes both the radar measurement error and representativeness error. As long as these two errors are uncorrelated, the estimated observation error variance \((\sigma^2 = 16 \text{ m}^2 \text{s}^{-2})\) is the sum of the measurement error variance and representativeness error variance, that is, \(\sigma^2 = \sigma_m^2 + \sigma_r^2\), where the subscripts \(m\) and \(r\) denote the measurement and representativeness part of the observation error, respectively. Since the above error variance was estimated in Xu et al. (2007) from radar observations collected during calm weather, the representativeness error variance \((\text{with } \Delta x = \Delta y = 3\text{-km grid resolution})\) could be as small as \(\sigma_r^2 \approx 0.1 \text{ m}^2 \text{s}^{-2}\) (see the remark at the end of section 3a). Thus, the estimated \(\sigma = 4 \) m s\(^{-1}\) can be considered as an upper bound for \(\sigma_m\). This upper bound will be used with the above lower bound \((\sigma_m = 1 \) m s\(^{-1}\)) in this paper to examine how the EnKF analysis is impacted by a change in observation accuracy.

c. Data assimilation procedure

The observation errors are assumed to be uncorrelated (or decorrelated), so the serial algorithm of EnSRF—a variant of the EnKF (Whitaker and Hamill 2002; Tippett et al. 2003)—can be used conveniently. A compactly supported smooth correlation function (Gaspari and Cohn 1999) is used to localize the background covariance computed from the ensemble. With the EnSRF, the model-state estimates are updated as follows:

\[
\mathbf{x}' = \mathbf{x} + W \cdot (y' - H\mathbf{x}) \quad \text{and} \quad (2)
\]

\[
\mathbf{x}_n' = \mathbf{x}_n' + W \cdot (y_{n}' - H\mathbf{x}_n) \quad \text{and} \quad (3)
\]

where \(\mathbf{x}\) denotes the model-state vector, the overbar denotes the ensemble average, the prime denotes the deviation from the mean, superscripts \(a\) and \(b\) denote the analysis and background, respectively, \(n\) is the index that identifies a particular ensemble member, \(y'\) is the observation, \(H\) is the observation operator, \(\circ\) denotes the Schur (elementwise) product, \(W\) is the localization weight computed from Eq. (4.10) of Gaspari and Cohn (1999), and \(y_{n}' = H\mathbf{x}_n' - H\mathbf{x}\) is the deviation of ensemble member \(n\) from the ensemble mean of the background projected to the observation space by observation operator \(H\). A simple trilinear interpolation is used in \(H\).

The Kalman gain \(\mathbf{K}\) (reduced to a column vector for single-observation analysis) and the \(\mathbf{\beta}\) factor are computed as follows:

\[
\mathbf{K} = (N - 1)^{-1} \sum_{n=1}^{N} y_n^{rb} y_n^{rb} \left[ \sigma^2 + \sum_{n=1}^{N} (y_n^{rb} y_n^{rb}) / (N - 1) \right]^{-1}
\]

and

\[
\mathbf{\beta} = \left[ 1 + \sigma^2 + \sum_{n=1}^{N} (y_n^{rb} y_n^{rb}) / (N - 1) \right]^{-1}
\]

where \(N\) is the ensemble size and \(\sigma^2\) is the observation error variance.

The ensemble has 40 members \((N = 40)\). As in TX05, the initial ensemble forecasts are started at 20 min of model time by adding random noises to a prior background state that is horizontally homogeneous and described by the environmental sounding in the vertical. The random noises are sampled from normal distributions with zero mean and standard deviation of 3 m s\(^{-1}\) for \(u, v,\) and \(w\), and 3 K for potential temperature. The pressure, moisture, and microphysical variables are not perturbed initially. The observations are assimilated every \(\tau = 5\) min in the control experiment. The first analysis is performed at \(t = 25\) min. After 15 more assimilation cycles, the last analysis is performed at the end \((t = 100\) min\) of the last (16th) cycle. The covariance localization radius is set to 8 km, which is the same as in TX05, and the covariance inflation factor is optimally tuned in each experiment.

3. Experiments with different observation accuracies

a. Experiment design

In this section, two EnSRF assimilation experiments (one control and one test experiment) are designed and performed with simulated high-resolution radial-velocity data \((0.5 \text{ km in radial range and } 1^\circ \text{ in azimuth})\) of different accuracies. As explained in section 2, the prediction model is imperfect as it has a relatively coarse horizontal resolution \((\Delta x = \Delta y = 2\) km\) with respect to the fine-resolution true state \((\Delta x = \Delta y = 0.5\) km\). The model error is not estimated or treated explicitly in the assimilation experiments. The conventional covariance inflation technique is used with the inflation factor optimally tuned in each experiment to partially account for the enhanced forecast error due to the imperfect model. Radial-velocity observations are generated from the fine-resolution true state by adding random measurement errors. In particular, by setting the error standard deviations to (i) \(\sigma_m = 1\) m s\(^{-1}\) and (ii) \(\sigma_m = 4\) m s\(^{-1}\), two sets
of observations are generated. These two sets are denoted by Om1 and Om4, respectively (see the first two rows of Table 1). As explained in section 2b, $\sigma_m = 1 \text{ m s}^{-1}$ is an estimated lower bound of $\sigma_m$, while $\sigma_m = 4 \text{ m s}^{-1}$ can be an upper bound for $\sigma_m$. The Om4 will be used in the control experiment, denoted by Em4. The Om1 will be used in the test experiment, denoted by Em1.

As explained in the introduction and section 2b, the total observation error variance is the sum of the measurement error variance and representativeness error variance, that is, $\sigma^2 = \sigma^2_m + \sigma^2_r$. The total observation error is estimated by the difference between the simulated radial-velocity observation and the coarsened-resolution true model state interpolated at each observation point (in the way described in section 2b for the simulated observations). The estimated differences are squared and then averaged over all the observation points (in the precipitation region where the reflectivity exceeds 10 dBZ) and over all the observation time levels to yield an estimate of $\sigma^2$ for each dataset. The representativeness error is also estimated independently by the difference between the two (fine resolution and coarsened resolution) true model states interpolated at each observation point and the estimated differences are squared and then averaged to yield an estimate of $\sigma^2_r$ for each dataset. The estimated representativeness error variances are found to be almost the same for all the datasets (listed in Table 1) and they all fall into the narrow range $3.00 \leq \sigma^2_r \leq 3.08 \text{ m}^2 \text{s}^{-2}$. The sum of the estimated $\sigma^2_r$ and specified $\sigma^2_m$ recovers the above estimated total observation variance $\sigma^2$ quite accurately (within $\pm 0.03 \text{ m}^2 \text{s}^{-2}$) for each of the six datasets. This indicates that the measurement error and representativeness error are uncorrelated for the simulated observations. Therefore, we can add the above estimated $\sigma^2_r = 3 \text{ m}^2 \text{s}^{-2}$ with the specified $\sigma^2_m$ to estimate $\sigma^2$ for each dataset, as listed in Table 1.

As a remark, it is necessary to mention that the representativeness errors can have considerable variations in time and space, although their spatially and temporally averaged values are almost the same for all the datasets. The spatially averaged representativeness error variances (over the observation points where the reflectivity exceeds 10 dBZ) are plotted by thick curves as functions of time in Fig. 1 for the six datasets. The datasets Om1 and Om4 have the same representativeness error, so their error variances averaged in the precipitation (or nonprecipitation) region are plotted by the same thick (or thin) solid curve.

![Fig. 1. Locally averaged representativeness error variances plotted as functions of time for the six datasets. The thick curves are the averages in the precipitation region (where the reflectivity is >10 dBZ). The thin curves are the averages in the non-precipitation region (where the reflectivity is <10 dBZ). The datasets Om1 and Om4 have the same representativeness error, so their error variances averaged in the precipitation (or nonprecipitation) region are plotted by the same thick (or thin) solid curve.](image-url)
other between different datasets than those (thick curves) averaged in the precipitation region. These error variances change very slowly with time, and their temporal averages do not exceed 0.12 m² s⁻² for the six datasets. The drastic decrease of the locally averaged representativeness error variance from the precipitation to nonprecipitation regions indicates that the significance of the representativeness error depends very much on the intensity of subgrid turbulence (not resolvable by the model grid but sampled by the observations).

b. Evaluation parameters

To evaluate the performances of the test experiment in comparison with that of the control experiment Em1, we need to introduce and define the following evaluation parameters:

(i) The analysis quality is evaluated by the RMS error of the ensemble mean velocity field. For the ith model variable, denoted by \( x_i \), the analysis RMS error is defined by

\[
\sigma_i = \left\{ \frac{1}{p} \sum_{t=1}^{p} (x_i^t - \bar{x}_i)^2 \right\}^{1/2},
\]

where \( \bar{x}_i \) denotes the ensemble mean background forecast (or analysis) of \( x_i \) before (or immediately after) each analysis time, \( x_i^t \) denotes the coarsened-resolution true value of \( x_i \) (on the \( \Delta x = \Delta y = 2 \)-km grid), and \( \frac{1}{p} \) denotes the average over the grid points covered by the observations (where reflectivity >10 dBZ).

(ii) The error variance (i.e., squared RMS error) averaged for the three velocity components at the end \( t = T = 100 \) min of the assimilation is denoted and defined by

\[
\sigma_v^2 = \left\{ \frac{1}{3 \cdot p} \sum_{t=1}^{p} (v_i^t - \bar{v}_i)^2 \right\}^{1/2},
\]

where \( \bar{v}_i \) denotes the average over \( i = 1, 2, 3 \) for the three velocity components \( (u, v, w) \).

(iii) To quantify the accuracy of the ensemble mean in the test experiment relative to that in the control experiment, it is convenient to introduce and define the relative RMS error (RRE) as follows:

\[
\text{RRE} = \sigma_v / \sigma_{v, \text{Em1}},
\]

where \( \sigma_{v, \text{Em1}} \) denotes the values of \( \sigma_v \) in the control experiment Em1.

(iv) The spatially averaged ensemble spread for \( x_i \) is defined by

\[
s_i = \left\{ \frac{1}{p} \sum_{t=1}^{p} (x_i^t - \bar{x}_i)^2 \right\}^{1/2},
\]

where the overbar denotes the ensemble average and \( \frac{1}{p} \) is the spatial average as in Eq. (6). Clearly, \( s_i^2 \) is the spatially averaged ensemble variance in \( x_i \) over the region covered by the observations. The average spread for the three velocity components at \( t = T \) can be then denoted and defined by \( s_v = \left[ s_i^2 \right]_{3 \cdot p} \) in the same way as \( \sigma_v \) in (7).

(v) Ideally, the ensemble spread in each variable should satisfy a consistency relationship with the ensemble mean RMS error. This consistency relation requires the ratio between the ensemble spread and the ensemble mean RMS error to be statistically equal to \( [N/(N + 1)]^{1/2} \) (Murphy 1988). For the ith variable \( x_i \), this ratio is defined and denoted by

\[
r_i = s_i / \sigma_i.
\]

We may call it the consistency ratio (the ith variable). For the ensemble size \( N = 40 \) used in this study, the ideal value for the consistency ratio is \( [N/(N + 1)]^{1/2} = 0.988 \).

(vi) The averaged consistency ratio for the three velocity components at the end \( t = T \) of the assimilation is defined and denoted by

\[
r_v = s_v / \sigma_v.
\]

If the uncertainty of the ensemble mean is well quantified by the ensemble spread, then \( r_i \) should be statistically equal to \( [N/(N + 1)]^{1/2} \) according to the consistency relationship. As mentioned in section 2c, the initial ensemble members are generated in the same way at the beginning of the assimilation in all the experiments. The initial consistency ratio is roughly in the same range of magnitude for each variable in all the experiments. After the initial time, however, the ratio will undergo different variations over the assimilation period in different experiments.

c. Results and discussion

The computed values of \( \sigma_v \), RRE, and \( r_v \) from the test experiment Em1 are listed in the first row of Table 2 in comparison with those from the control experiment Em4 (in the second row). The covariance inflation factor is optimally tuned (at the resolution of every 0.025 or 0.05) in each experiment. As listed, when the measurement error is reduced from \( \sigma_m = 4 \) m s⁻¹ (in Em4) to 1 m s⁻¹ (in Em1) and consequently the total observation error variance \( \sigma^2 \) is reduced from 19 to 4 m² s⁻², the optimal inflation factor is reduced from 1.15 to 1.05, the RRE is decreased by merely 2%, and the averaged consistency ratio \( r_v \) is decreased from 0.48 to 0.16. Since the analysis error reduction is very small, the decreased \( r_v \) is mainly a result of the reduced spread (from \( s_v = 1.56 \) to 0.51 m s⁻¹).
The ensemble mean RMS errors $\sigma_i$ and consistency ratios $r_i$ are plotted as functions of time over the 16 assimilation cycles in Figs. 2a–c and 2d–f, respectively, for $(u, v, w)$ from the above two experiments. As shown in Figs. 2a–c, Em1 (dotted lines) performs only slightly better than Em4 (solid lines) although the dataset Om1 is more accurate than Om4. Figures 2d–f show that the consistency ratios in Em1 are initially about the same as in Em4 but decrease in the first few assimilation cycles and then are maintained far below the ideal value (0.988) during the subsequent assimilation cycles. Clearly, the ensemble spread in Em1 is too small (0.51 m s$^{-1}$) to quantify the uncertainty of the ensemble mean.

The coarsened-resolution true fields of vertical velocity and horizontal perturbation wind at $z = 6$ km are plotted in a time series in the first row of Fig. 3 against the ensemble mean analyses from Em1 (second row) and Em4 (third row). The associated analysis error fields (computed by subtracting the coarsened-resolution true fields from their respective mean analysis fields) are plotted in Fig. 4 for Em1 (first row) and Em4 (second row). As shown in Fig. 3, the gross pattern of the true wind field is reasonably well assimilated (after five cycles) in these two experiments. The assimilated updrafts in Em1 and Em4 have about the same intensity as the true updrafts, and the major updraft patterns capture the true pattern almost equally well in the two experiments. As shown in Fig. 4, the Em4 analyses have slightly smaller errors in the vertical velocity but larger errors in the horizontal velocity than the Em1 analyses. This feature can be quantified by the RMS errors.
averaged in the precipitation region at $z = 6$ km. In particular, for the analyses produced from Em1 and Em4 at the 16th cycle ($t = 6000$ s), the horizontal velocity errors are 2.96 and 3.17 m s$^{-1}$, the vertical velocity errors are 2.69 and 2.39 m s$^{-1}$, and the vector velocity errors are 2.87 and 2.93 m s$^{-2}$, respectively. The detailed results in Fig. 4 indicate again that the analysis is not significantly improved as the observation error variance is reduced from $\sigma^2 = 19$ m$^2$ s$^{-2}$ (in Em4) to 4 m$^2$ s$^{-2}$ (in Em1).

Although the spread is apparently insufficient in Em1, the situation cannot be improved by further increasing the inflation (from the already optimally tuned value). On the contrary, as we noted during Em1 inflation tuning, increasing the inflation factor, say, from 1.05 to 1.1 can only worsen the analysis (with $\sigma_v$ increased from 3.21 to 3.43 m s$^{-1}$). Thus, there is a limit for improving the analysis by increasing the measurement accuracy especially when the total observation error is already sufficiently small. The existence of such a limit can be
understood as follows. First, with an imperfect model and limited ensemble size (of only 40 members), the ensemble-estimated background covariance may contain significant errors and spurious noisy structures even after localization. Second, if the observation error becomes smaller than the error of the ensemble-estimated background covariance, then the latter error will contribute dominantly to the analysis error and, therefore, further decreasing the observation error cannot effectively improve the analysis. Based on the above results and understanding, we will consider Om4 only and use it as a benchmark dataset to study trade-offs between observation accuracy and resolutions in the next section.

4. Experiments with trade-offs between measurement accuracy and resolutions

a. A simple rule for trade-offs

As explained in the introduction, a large number of uncorrelated samples need to be acquired to estimate

![Fig. 4. As in Fig. 3, but for the analysis error fields (computed by subtracting the coarsened-resolution “true” fields from their respective mean analysis fields) from (first row) Em1, (second row) Em4, (third row) E2km5min, and (last row) E2km1min. The vector scale for the horizontal wind is shown by the arrow (2.5 m s^{-1}) at the lower-left corner.](image)
the (level II) radial velocity in each resolution volume (typically 0.25 km × 1° × 1°) and to reduce the uncertainty (caused by random meteorological scatterers) of the estimated velocity. The total number of samples for all the radar radial-velocity estimates (level-II data) of a volume scan is the number of the estimates multiplying the number of samples needed for each estimate (in each resolution volume). The number of acquirable samples per unit time, however, always has a limit for a given phased-array radar (or any other radar), especially if the radar is designed for multifunction tasks. Constrained by this limit, if the observation resolution is increased, say, by n times in one spatial or temporal dimension, then the measurement error variance will have to increase also by n times (because the error variance of an estimate is inversely proportional to the number of independent samples used to yield the estimate according to the well-known central limit theorem in statistics). Therefore, as a simple rule for trade-offs between measurement accuracy and resolutions, the measurement error variance \( \sigma^2_m \) will decrease (or increase) by n times if the resolution is decreased (or increased) by n times in any spatial or temporal dimension.

Using the above rule, radar scans can be configured with different trade-offs between accuracy and resolutions from the benchmark dataset Om4. As we have seen in the previous section, Om4 has the same excessive spatial resolutions as Om1 (in comparison with the coarsened model resolution) but much lower measurement accuracy than Om1. By taking Om4 as a benchmark dataset, four sets of observations are generated by reducing the spatial resolutions in favor of increasing the measurement accuracy and temporal resolution. These four sets are denoted by O2km5min, O4km5min, O2km2min, and O2km1min, respectively. Their measurement error variances, observation error variances, and resolutions are listed in the last four rows of Table 1.

b. Trade-offs between accuracy and spatial resolutions

In this subsection, two test assimilation experiments are performed with O2km5min and O4km5min, and these experiments are denoted by E2km5min and E4km5min, respectively. The control experiment is Em4 as in section 3. The parameter values of the optimally tuned inflation factor \( c \), averaged RMS error \( \sigma_v \), relative RMS error, and averaged consistency ratio \( r_c \) from E2km5min and E4km5min are listed in the third and fourth rows of Table 2, respectively, in comparison with those from Em4 (in the second row). In addition, the CPU time ratios (with respect to the CPU time used by Em4) are listed in the last column. As listed in Tables 1 and 2, when the radial resolution is coarsened from 0.5 to 2 km and thus \( \sigma^2 \) is reduced from 19 m s\(^{-1}\) (in Em4) to 7 m\(^2\) s\(^{-2}\) (in E2km5min), the optimal inflation factor is increased from 1.15 to 1.2, the RRE is decreased by 4% (from 1 to 0.96), the averaged consistency ratio \( r_c \) is increased from 0.48 to 0.76, and the CPU time ratio is reduced from 1 to 0.24. However, when the radial resolution is further coarsened to 4 km, and the azimuthal resolution is coarsened from 1° to 2° in E4km5min and thus \( \sigma^2 \) is further and excessively reduced to 4 m\(^2\) s\(^{-2}\) (as in Em1), the analysis becomes worse and the RRE is increased by 3% (from 1 to 1.03) although the CPU time ratio is further reduced to 0.06. The optimal inflation factor remains to be 1.2, but the averaged consistency ratio \( r_c \) is increased to 105%. The worsened analysis accuracy in E4km5min is caused by the ineffectiveness of excessively increased observation accuracy (as in Em1) at the expense of excessively coarsened spatial resolutions. Judging from the analysis accuracy, E2km5min outperforms Em4 and E4km5min, and the observation set O2km5min provides the optimal trade-off (between the measurement accuracy and spatial resolutions) among the three (Om4, O2km5min, and O4km5min).

The ensemble mean RMS errors \( \sigma_i \) and consistency ratios \( r_i \) are plotted as functions of time over the 16 assimilation cycles in Figs. 5a–c and 5d–f, respectively, for \((u, v, w)\) from the two test experiments E2km5min and E4km5min versus those from the control experiment Em4. As shown in Figs. 5a–c, E2km5min (dotted lines) performs slightly better than Em4 (solid lines) and E4km5min (gray solid lines) performs slightly worse than E2km5min. Figures 5d–f show that the averaged consistency ratios in E2km5min and E4km5min are initially the same as in Em4 but become higher than that in Em4 during the subsequent assimilation cycles. The consistency ratio in E2km5min becomes quite close to the ideal value (=0.988) in the late stage of assimilation, and this indicates that the ensemble spread is adequately improved in E2km5min. The consistency ratio in E4km5min, however, becomes larger than the ideal value, so the ensemble spread is excessively increased in E4km5min. These results are consistent with those listed in Table 2.

The ensemble mean fields of vertical velocity and horizontal perturbation wind from E2km5min are plotted at \( z = 6 \) km in a time series in the last row of Fig. 3 in comparison with the true fields (in the first row) and from Em4 (the third row). The associated analysis error fields (analysis minus truth) from E2km5min are plotted in the third row in Fig. 4 in comparison with those from Em4 (the second row). As shown, the E2km5min analyses have slightly larger errors in the vertical velocity but smaller errors in the horizontal velocity than
the Em4 analyses. This feature is quantified by the RMS errors averaged in the precipitation region at \( z = 6 \) km. For example, for the analyses produced from E2km5min and Em4 at the 16th cycle (\( t = 6000 \) s), the horizontal velocity errors are 2.90 and 3.17 m s\(^{-1}\), the vertical velocity errors are 2.76 and 2.39 m s\(^{-1}\), and the vector velocity errors are 2.86 and 2.93 m s\(^{-1}\), respectively. These comparisons are consistent with the comparison between the domain-averaged RMS errors listed for E2km5min and Em4 in Table 2; and they all indicate that the analysis can be improved by coarsening the radial resolution from 0.5 km (in Om4) to 2 km (in O2km5min) in favor of increasing the measurement accuracy.

As we have seen, by assimilating O2km5min instead of the original Om4 that has quadruply higher radial resolution than O2km5min, the computational cost can be reduced quadruply. This trade-off also improves the analysis accuracy in addition to the computational efficiency. According to the formulations derived in Xu (2007) and Xu et al. (2008b) for measuring information content from observations for three- and four-dimensional data assimilations, the observation resolution should not be finer than that resolved by the background error covariance. The background error covariance estimated by an EnKF can have fine structures at the same resolution as the analysis grid, so the observation resolution can be as fine as but not finer than the EnKF analysis grid. This is also what we have seen from the results obtained in this section. After the spatial resolutions are optimally selected, it is necessary to consider the trade-off between the measurement accuracy and temporal resolution, and a proper trade-off between the accuracy and temporal resolution can further improve the EnSRF analysis. This problem is explored in the next subsection.

c. Trade-offs between accuracy and temporal resolution

In this subsection, two test assimilation experiments are performed with O2km2min and O2km1min, and these experiments are denoted by E2km2min and E2km1min, respectively. Because O2km2min and O2km1min have the same spatial resolutions as O2km5min, comparisons will be made with respect to E2km5min instead of Em4 [although the RRE is still computed relative to Em4 according to the definition in Eq. (8)]. The parameter values of \( c \), \( \alpha_r \), RRE, \( r_v \), and CPU time ratio from E2km2min and E2km1min are listed in the last two rows.
of Table 2 in comparison with those from E2km5min. As listed in Tables 1 and 2, when the temporal resolution is refined from 5 to 2 min and thus $\sigma^2$ is increased from 7 to 13 $m^2 s^{-2}$ in E2km2min, the optimal inflation factor is decreased from 1.2 to 1.05, the RRE is decreased by 11% (from 0.96 to 0.85), and the averaged consistency ratio $r_v$ is reduced to 0.53, but the CPU time ratio is increased from 0.24 to 0.66. When the temporal resolution is further refined to 1 min and thus $\sigma^2$ is increased to 23 $m^2 s^{-2}$ in E2km1min, the optimal inflation factor is reduced by 0.025 (from 1.05 to 1.025) and the CPU time ratio is further increased to 1.12. The analysis, however, is not much improved and the RRE is reduced by merely 1% (from 0.85 to 0.84), while the averaged consistency ratio $r_v$ is increased only slightly (from 0.53 to 0.54).

The ensemble mean RMS errors $\sigma_i$ and consistency ratios $r_i$ are plotted as functions of time over the 16 assimilation cycles in Figs. 6a–c and 6d–f, respectively, for $(u, v, w)$ from E2km2min and E2km1min versus those from E2km5min. As shown in Figs. 6a–c, E2km2min (dotted lines) outperforms E2km5min (solid lines), and E2km1min (dashed lines) slightly outperforms E2km5min. Figures 6d–f show that the averaged consistency ratios in E2km2min and E2km1min are initially the same as in E2km5min but not increased as fully as in E2km5min during the subsequent assimilation cycles. Clearly, the spreads are insufficient in E2km2min and E2km1min, but the analyses are not improved by increasing the inflation (from the already optimally tuned values). For example, the averaged RMS error $\sigma_v$ will increase from 2.76 (or 2.74) to 2.89 (or 3.18) m s$^{-1}$ if the inflation factor is increased from the optimal value 1.05 (or 1.025) to 1.1 in E2km2min (or E2km1min). Note that the inflation is applied, as in TX05, not to all the grid points but only a part of the grid points within the precipitation region (where the observed reflectivity exceeds 10 dBZ) or their 8-km vicinities and thus will be influenced by the analysis update in each cycle. This partial-space inflation is used here, as proposed in TX05, to avoid spurious convective activities outside the observed precipitation region, but it causes spatial discontinuities in the ensemble perturbation fields harmful to the analysis and ensemble integration, especially when the ensemble is updated very frequently, as in E2km2min and E2km1min. This explains why the analyses are not improved by further increasing the inflation factors in E2km2min and E2km1min although
the spreads are still apparently insufficient with the optimally tuned inflations in these two experiments.

The analysis error fields at $z = 6$ km from E2km1min are plotted in the last row in Fig. 4 in comparison with those from E2km5min (the third row) and Em4 (the second row). As shown, after the fifth analysis cycle, the analysis errors in E2km1min become notably smaller than those in E2km5min and Em4, especially in the main updraft regions. For the analyses produced at the 16th cycle ($t = 6000$ s) from E2km1min in comparison with those from E2km5min, the averaged RMS errors in the precipitation region at $z = 6$ km are reduced from 2.90 to 2.73 m s$^{-1}$ in the horizontal velocity, from 2.76 to 2.22 m s$^{-1}$ in the vertical velocity, and from 2.86 to 2.57 m s$^{-1}$ in the vector velocity. These results are consistent with those in Fig. 6 and Table 2. Judging from the analysis accuracy, E2km1min slightly outperforms E2km2min. Judging from the CPU time ratio, E2km2min is much more efficient than E2km1min. In either case, a proper trade-off between the accuracy and temporal resolution can indeed further improve the EnSRF analysis, as exemplified by O2km2min or O2km1min in comparison with O2km5min.

5. Conclusions

In this paper, assimilation experiments are designed and carried out with simulated radar velocity observations to examine the impacts of measurement accuracy and resolutions on storm-scale wind assimilation with the ensemble square root filter (EnSRF) on a storm-resolving grid ($\Delta x = 2$ km). In recognizing that the number of acquirable samples per unit time always has a limit for a given radar and that this limit imposes an overall constraint on radar measurement accuracy and resolutions, proper trade-offs between measurement accuracy and resolutions are considered in designing the assimilation experiments (see section 4a). The main results of the experiments are summarized as follows:

1) When the observations have excessive accuracy while the background covariance is not accurately estimated (from an ensemble of merely 40 imperfect-model predictions in this study), further increasing the measurement accuracy cannot effectively increase the EnSRF analysis accuracy. Because of errors (which are largely unknown and hard to quantify) in the ensemble-estimated background covariance, there is a limit for improving the EnSRF analysis by increasing the measurement accuracy (as suggested by the results of Em1 versus Em4 in section 3c).

2) When the observations have no excessive accuracy, the analysis can be improved by properly reducing the observation resolution (to 2 km in the radial direction) in favor of increasing the measurement accuracy, and there is an optimal trade-off between the accuracy and spatial resolutions.

3) When the spatial resolutions are optimally selected, the analysis can be further improved by properly enhancing the temporal resolution of radar volume scans (from 5 to 2 or 1 min) and compromising the measurement accuracy, and there is an optimal trade-off between measurement accuracy and temporal resolution.

The trade-offs between measurement accuracy and resolutions considered in this study can be important for optimally configuring radar scans, especially phased-array radar scans (Forsyth et al. 2005, Zrnic et al. 2007; Heinselman et al. 2008) to improve storm-scale radar wind assimilation. The quantitative aspects of the results, however, will change if the model grid resolutions become significantly different from those used in this study (see section 2a). In particular, if the horizontal grid spacing is reduced further (say, from $\Delta x = 2$ to 0.2 km for the purpose of resolving and analyzing tornadic vortices within a storm), then the optimal trade-off will likely be balanced at higher spatial and temporal resolutions but lower measurement accuracy than shown in this study, provided that the trade-off is made with respect to the same benchmark observations (i.e., Om4 to the full capabilities of the radar) as considered in section 4 of this paper. In this case, the EnSRF may behave quite differently than shown in this study. This speculation needs to be verified by future studies.

The formulations derived in Xu (2007) for measuring information content from observations for three-dimensional data assimilation were recently extended and used in Xu et al. (2008b) to examine the trade-offs between measurement accuracy and resolutions in configuring radar velocity scans for four-dimensional data assimilation. According to these previous studies, the information content will reduce if the observations become too dense to be resolved by the background error covariance or, equivalently, if the background error covariance is too smooth to resolve the observations. The background error covariance estimated by an EnKF can have fine structures at the same resolution as the analysis grid, so the observation resolution can be as fine as but not finer than the EnKF analysis grid. This point is supported by the results in this study. However, there is difficulty in optimizing radar azimuthal resolution over a broad radial range, because the azimuthal resolution (measured in arc length) is a linear function of the radial range distance. The difficulty can be avoided by optimizing the azimuthal resolution in the far radial range.
and then superobbing the observations along the azimuthal in the near-radar radial ranges. This approach is not examined in this study (since the simulated observations are not over a broad radial range) but should be considered in future real radar assimilations.

Finally, it is necessary to point out that the simple rule for trade-offs between measurement accuracy and resolution used in this study (see section 4a) is an idealized situation. In reality, radar-sampled (level I) velocities are not truly independent, and the statistical properties of random meteorological scatterers are not precisely stationary in each resolution volume. Consequently, when the observation resolution increases (or decreases) in a single spatial or temporal dimension, the trade-off will yield a smaller increase (or decrease) in the true measurement error variance than estimated by the simple rule. Furthermore, also in reality, radar radial-velocity observation errors are correlated between neighboring range gates and between neighboring beams (Xu et al. 2007). Although the simulated radar observations are thinned (with the range gate spacing increased from 0.25 to 0.5 km in Om1 and to 2 km in O2km5min and O2km2min) in this study to reduce or avoid error correlation in the radial direction, it remains necessary to deconvolute correlated observations between neighboring beams, because the azimuthal resolution is only 1° for the simulated datasets (except for O4km5min) in this study (see Table 1). In this case, the observation error variance will be increased because of the deconvolution. The above factors should be considered when the results obtained in this paper are used to assess the impacts of observation accuracy and resolutions on the EnSRF assimilation of real radar observations of storm winds. This problem deserves further investigation.

Acknowledgments. We are thankful to the anonymous reviewers for their insightful comments and suggestions that called our attention to the representativeness error (not considered in the original submission) and improved the quality and presentation of the work. We are also thankful to Mingjing Tong and Xue Ming for their help in providing the original ARPS EnKF code, and to Mingming Yao for her help in performing numerical experiments for the revision. The computational resource was provided by the OU Supercomputing Center for Education & Research at the University of Oklahoma. The research work was supported by the FAA Contract IA DTFAO3-01-X-9007 to NSSL and by ONR Grant N000140410312 to CIMMS, the University of Oklahoma. Funding was also provided to CIMMS by NOAA/Office of Oceanic and Atmospheric Research under NOAA/University of Oklahoma Cooperative Agreement NA17RJ1227, U.S. Department of Commerce. Part of this work was also funded by the Key Project of the National Natural Science Foundation of China Grant 2004GB418306.

REFERENCES


——, L. Wei, and S. Healy, 2008b: Measuring information content from observations for data assimilations: connection between different measures and application to radar scan design. Tellus, 61A, 144–153.