

Time-Expanded Sampling for Ensemble Kalman Filter: Assimilation Experiments with Simulated Radar Observations

QIN XU

NOAA/National Severe Storms Laboratory, Norman, Oklahoma

HUIJUAN LU

Cooperative Institute for Mesoscale Meteorological Studies, University of Oklahoma, Norman, Oklahoma, and State Key Laboratory of Severe Weather, Chinese Academy of Meteorological Sciences, Beijing, China

SHOUTING GAO

Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing, China

MING XUE

Center for Analysis and Prediction of Storms, and School of Meteorology, University of Oklahoma, Norman, Oklahoma

MINGJING TONG

Center for Analysis and Prediction of Storms, University of Oklahoma, Norman, Oklahoma

(Manuscript received 26 February 2007, in final form 26 October 2007)

ABSTRACT

A time-expanded sampling approach is proposed for the ensemble Kalman filter (EnKF). This approach samples a series of perturbed state vectors from each prediction run not only at the analysis time (as the conventional approach does) but also at other time levels in the vicinity of the analysis time. Since all the sampled state vectors are used to construct the ensemble, the number of required prediction runs can be much smaller than the ensemble size and this can reduce the computational cost. Since the sampling time interval can be adjusted to optimize the ensemble spread and enrich the ensemble structures, the proposed approach can improve the EnKF performance even though the number of prediction runs is greatly reduced. The potential merits of the time-expanded sampling approach are demonstrated by assimilation experiments with simulated radar observations for a supercell storm case.

1. Introduction

Since the ensemble Kalman filter (EnKF) was proposed by Evensen (1994) as an alternative data assimilation strategy, the EnKF technique has been refined in many aspects (see, e.g., Burgers et al. 1998; Anderson 2001; Bishop et al. 2001; Houtekamer and Mitchell 2001; Mitchell et al. 2002; Whitaker and Hamill 2002;

Tippett et al. 2003; Evensen 2003) and applied to data assimilation with not only larger-scale models but also storm-scale models (Snyder and Zhang 2003; Zhang et al. 2004; Dowell et al. 2004; Tong and Xue 2005; Xue et al. 2006, hereafter XTD06). The traditional (extended) Kalman filter (KF) propagates and updates not only the mean state vector but also the covariance matrix for a complete description of the assumed normal probability density function (pdf) of the model state (Jazwinski 1970; Cohn 1997). As the prognostic equation for the covariance matrix involves tangent linearization, it can be very complex to code and very expensive to integrate if the model state vector is of high dimensions,

Corresponding author address: Dr. Qin Xu, National Severe Storms Laboratory, 120 David L. Boren Blvd., Norman, OK 73072.

E-mail: qin.xu@noaa.gov

and the KF can become problematic when the model physics contain parameterized discontinuities (Xu 1996, 1997). Unlike the traditional KF, the EnKF propagates and updates an ensemble of state vectors based on the Monte Carlo method. As a prior estimate of the covariance matrix is directly computed from the ensemble and used for the analysis, the algorithm is simple and easy to code. Computationally, however, the EnKF is still very expensive because of large ensemble sizes required by the Monte Carlo method. Thus, how to reduce the computation cost of the EnKF is a major problem for its operational applications. This has motivated this study.

Another very challenging problem encountered by the EnKF concerns how to deal with model errors, especially model biases. The problem is notorious for all advanced assimilation techniques. Model errors are often closely linked to nonlinearities of the true processes, and they are usually unknown and difficult to estimate. The presence of unknown model errors can cause the filter (KF or EnKF) to diverge, that is, to cause the ensemble mean analysis to drift away from the observations and thus away from the true state as the assimilation proceeds through a large number of cycles. To prevent the EnKF from filter divergence as a result of unknown model errors as well as other reasons, some empirical treatments, such as the covariance inflation (Anderson 2001) and model error parameterization (Mitchell and Houtekamer 2000), have been used. The original model error problem, however, remains largely unsolved. It is thus desirable to explore other possible approaches to address or alleviate this problem perhaps more effectively from different perspectives. This is another motivation of this study, even though the attempted effort does not directly deal with model errors.

For an ensemble of size N , the EnKF updates N ensemble members at each analysis step to initialize N prediction runs and then propagates the ensemble to the next analysis time. For each analysis step, N predicted state vectors are sampled at the analysis time to compute prior mean and covariance. The computed covariance is then localized to eliminate spurious long-range correlation caused by limited ensemble size (Houtekamer and Mitchell 2001). In this study, we propose and test a new approach in which a series of perturbed state vectors are sampled from each prediction run not only at the analysis time but also at other time levels properly selected in the vicinity of the analysis time. As all the above sampled state vectors are used to compute the background covariance for the analysis, the ensemble size is increased without increasing the number of prediction runs. Thus, for a given ensemble

size, the number of prediction runs is reduced, and this can reduce the computational cost significantly.

The above time-expanded sampling is motivated by the consideration of timing errors in model predictions. This type of error occurs when the model predicted system develops or propagates faster (or slower) than the true system. Timing errors can be significant or even pervasive with considerable uncertainties in model predictions as suggested by many model verification studies (Manobianco and Nutter 1999; Colle et al. 2001; Mass et al. 2002; Colle and Charles 2007; Gregory and Grumm 2007). According to S. Weiss 2007, personal communication), Science and Operations Officer at the National Oceanic and Atmospheric Administration (NOAA)/National Weather Service (NWS) Storm Prediction Center, timing errors and timing-related spatial displacement errors were often distinct in operational model predictions, including those produced by the experimental convection-allowing Weather Research and Forecasting (WRF) models (Weiss et al. 2007; Xue et al. 2007), especially when considering the movement/location of precipitation systems. In the presence of timing error, the predicted field at a time before (or after) the analysis time may represent the true field better than the predicted field at the analysis time, at least, over the area covered by the weather system. Without a prior estimation, however, the timing error may be assumed to be statistically random with a zero mean. In this case, the localized forecast error covariance may be better estimated by the above proposed time-expanded sampling, that is, by sampling ensemble members not only at the analysis time but also at properly selected times before and after the analysis time.

Although the proposed time-expanded sampling is motivated by the consideration of timing errors, its applicability is not necessarily limited by this consideration. As will be demonstrated by the assimilation experiments performed in this paper, the proposed approach can reduce the computational cost and improve the filter performance even though the ensemble predictions are initialized with no systematic timing error. The potential merits of the proposed approach will be explored in this paper.

The paper is organized as follows: The next section describes the model-simulated radar observations and assimilation system. Section 3 designs assimilation experiments with the proposed time-expanded sampling. Assimilation experiments are performed with perfect model and imperfect model settings in sections 4 and 5, respectively, to explore and demonstrate the merits of the proposed approach. Conclusions follow in section 6.

2. Simulated observations and assimilation system

a. Prediction model and simulated storm

Tong and Xue (2005, hereafter TX05) demonstrated the ability of the EnKF method in retrieving not only the wind and thermodynamic variables but also multiple microphysical species associated with a multiclass ice microphysics scheme by using the Advanced Regional Prediction System (ARPS) model (Xue et al. 2001). They also discussed the relative role of radial velocity and reflectivity data as well as their spatial coverage in recovering the full flow and cloud fields. The current study uses the same ARPS configurations for the same storm case in TX05, that is, the 20 May 1977 Del City, Oklahoma, supercell (Ray et al. 1981). The model configuration contains 12 prognostic state variables, including wind components u , v , and w ; perturbation potential temperature θ' ; perturbation pressure p' ; turbulence kinetic energy (TKE); and the mixing ratios for water vapor q_v , cloud water q_c , rainwater q_r , cloud ice q_i , snow q_s , and hail q_h . As in TX05, all of these prognostic state variables except for TKE are subject to updating at each analysis step of assimilation.

The “true” state for our observing system simulation experiments (OSSEs) 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 $z = 1.5 \text{ km}$, with a radius of 10 km in horizontal directions and 1.5 km in vertical directions, is used to initiate the storm. As in TX05, for all experiments, the physical domain is $64 \times 64 \times 16 \text{ km}^3$ and is covered by a $35 \times 35 \times 35$ grid with horizontal grid spacings $\Delta x = \Delta y = 2 \text{ km}$ and vertical grid spacing $\Delta z = 0.5 \text{ km}$. The true state, however, is generated in two ways: (i) by using the same $35 \times 35 \times 35$ grid for the perfect model assimilation experiments (in section 4) and (ii) by using a $131 \times 131 \times 35$ grid (with the horizontal resolution enhanced to $\Delta x = \Delta y = 0.5 \text{ km}$) for the imperfect model assimilation experiments (in section 5).

b. Simulated radar observations

The radar volumetric observations are assumed to be available every 5 min, and this observation time interval, denoted by $T = 5 \text{ min}$, is consistent with the storm scan mode of the operational Weather Surveillance Radar-1988 Doppler (WSR-88D). Each radial velocity observation, V_r , is simulated by

$$V_r = u \cos\alpha \sin\beta + v \cos\alpha \cos\beta + w \sin\alpha + \varepsilon, \quad (1)$$

where α is the elevation angle and β is the azimuth angle of the radar beam, and (u, v, w) are the three components of the true velocity at the observation point obtained via trilinear interpolation from the true state grid. The ε represents random errors that are drawn from a normal distribution with zero mean and standard deviation of 1 m s^{-1} .

In this study, we will assimilate radial velocity data only, as in experiment VrP of TX05 (see their Table 1). The radial velocity data are assumed to be available only in precipitation regions where the reflectivity exceeds 10 dBZ and the reflectivity is calculated from the hydrometeor mixing ratios from the true simulation. As in TX05, the ground-based radar is located at the southwest corner of the computational domain, that is, at the x - y coordinate origin, for all experiments.

For the perfect model assimilation experiments, all observations are given on the scalar points of the staggered ARPS grid as in TX05. For the imperfect model assimilation experiments, the observations are given, via trilinear interpolation from the true state grid, on the polar coordinates according to the WSR-88D radar storm-mode scan configuration, which contains 14 tilts at the elevation angles of $0.48^\circ, 1.45^\circ, 2.4^\circ, 3.3^\circ, 4.3^\circ, 5.2^\circ, 6.2^\circ, 7.5^\circ, 8.7^\circ, 10.0^\circ, 12^\circ, 14.0^\circ, 16.7^\circ$, and 19.5° . On each tilt, the horizontal resolutions are 1° in the azimuthal direction and 2 km in the radial range direction for the radial velocity. Here, the range gate resolution is coarser than the real WSR-88D velocity resolution of 0.25 km. According to our additional experiments (not shown), making the range gate resolution for the (simulated) radar observations higher than the model resolution does not improve the EnKF performance and subsequent model predictions. Because of this, the range gate resolution is simply set to be the same as the model grid resolution.

c. Data assimilation procedure

The observation errors are assumed to be uncorrelated, so the serial algorithm of ensemble square root filter (EnSRF; Whitaker and Hamill 2002) can be conveniently used in this study. A compactly supported smooth correlation function (Gaspari and Cohn 1999) is used to localize the background covariance computed from the ensemble as in TX05 and XTD06.

With the EnSRF, the model state estimates are updated as follows:

$$\bar{\mathbf{x}}^a = \bar{\mathbf{x}}^b + \mathbf{W} \circ \mathbf{K}(\mathbf{y}^o - \overline{H\mathbf{x}}^b) \quad \text{and} \quad (2)$$

$$\mathbf{x}_n'^a = \mathbf{x}_n'^b - \mathbf{W} \circ \mathbf{K}\beta\mathbf{y}_n'^b, \quad (3)$$

where \mathbf{x} denotes the model state vector, the overbar denotes the ensemble mean, the prime denotes the deviation from the mean, superscripts a and b denote the analysis and background (or prior estimate), respectively, n is the index that identifies a particular ensemble member, \mathbf{y}^o is the observation vector, H is the observation operator, \circ denotes the Schur (element-wise) product, and \mathbf{W} is the weight computed from the Eq. (4.10) of Gaspari and Cohn (1999), which depends on the distance from the observation to the model grid points. In Eq. (3),

$$\mathbf{y}_n'^b \equiv H\mathbf{x}_n^b - \overline{H\mathbf{x}^b} \quad (4)$$

is the deviation of ensemble member n from the ensemble mean of the background projected to the observation space by observation operator H .

For the serial algorithm used here that assimilates one observation at a time, the Kalman gain \mathbf{K} , which reduces to a column vector, and the β factor are computed as follows:

$$\mathbf{K} = \langle \mathbf{x}_n'^b \mathbf{y}_n'^b \rangle (\sigma^2 + \langle \mathbf{y}_n'^b \mathbf{y}_n'^b \rangle)^{-1} \quad \text{and} \quad (5)$$

$$\beta = [1 + \sigma(\sigma^2 + \langle \mathbf{y}_n'^b \mathbf{y}_n'^b \rangle)^{-1/2}]^{-1}, \quad (6)$$

where

$$\langle (\cdot) \rangle \equiv \frac{1}{N-1} \sum_{n=1}^N (\cdot),$$

N is the ensemble size, and σ^2 is the observation error variance.

The ensemble size is constrained by the available computational capabilities and resources. Mitchell et al. (2002) used 64 members in their perturbed observation EnKF assimilation experiments and the analyses converged reasonably well. Snyder and Zhang (2003) and Dowell et al. (2004) used 50 members, and Zhang et al. (2004) even tried as few as 20 members in their EnSRF assimilation experiments. TX05 obtained very good results with 100 members and the perturbed observation EnKF algorithm while in XTD06, 40 members were used with EnSRF and the ARPS model. In this study, by applying the time-expanded sampling to the EnSRF, we seek to further reduce the ensemble sample size to 30 and take the samples out of only 10 prediction runs.

3. Experiment design and evaluation parameters

a. Experiment design and related considerations

With the time-expanded sampling proposed in this paper, the ensemble consists of the standard members and time-expanded members. The standard members are sampled at the analysis time from N_s perturbed runs in the same way as in TX05. The time-expanded members are sampled at equally separated time levels before and after the analysis time. We denote $t = t_j$ as the

j th analysis time (for $j = 1, 2, \dots, 16$ in each experiment). If we denote τ as the sampling time interval, then $2M\tau$ is the sampling time window, and the sampling time levels centered at the j th analysis time are $t = t_j + m\tau$ for $m = 0, \pm 1, \pm 2, \dots, \pm M$. Clearly, with the time-expanded sampling, the ensemble size is increased by a factor of $2M + 1$, that is, from $N = N_s$ to $N = N_s(2M + 1)$.

Assimilation experiments can be designed with different settings of M and τ , while the sampling time window is confined between the previous and next analysis times, that is, $2M\tau \leq 2T$, where $T = 5$ min is the assimilation cycle time length, which is the same as the observation time interval. In principle, the sampling time window can be extended beyond $2T$, but the sampling time must go back beyond $t = t_j - T$ to sample either the forecast initialized at $t = t_{j-2}$ from the previous cycle (without using observations at $t = t_j - T$) or the analysis obtained by a Kalman smoother (that uses the observations at $t = t_j - T$). These complications as well as increased cost are undesirable and therefore not considered in this study. In terms of the selection of M and τ , the experimental designs for the perfect model and imperfect model cases are similar. In each case, we have two control experiments, C30 and C10, designed with $M = 0$, and with $N = N_s = 30$ and 10, respectively. Clearly, since $M = 0$, these control experiments reduce to use of the conventional approach without time expansion.

For the perfect model case, the two control experiments, C30 and C10, are similar to the VrP experiment in TX05, except that the EnSRF is used and the ensemble sizes are reduced from 100 to 30 and 10, respectively. As in TX05, the initial ensemble forecasts are started at the 20 min of model time by adding random noise to a prior background state that is horizontally homogeneous and described by the environmental sounding in the vertical. The random noise is sampled from normal distribution 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. The observations are assimilated every $T = 5$ min. The first analysis is performed at $t = 25$ min. After a total of 16 assimilation cycles, the model prediction is launched at $t = 100$ min. For C30 with either the perfect or imperfect model, the covariance localization radius and inflation factor are set to 8 km and 1.07, respectively, which are the same as those used in TX05. For C10 with the perfect model, the covariance localization radius is tuned in 1-km increments to obtain an optimal value of 5 km and the inflation factor is tuned in 0.05 increments to arrive at an optimal value of 1.10. These parameter values are

also used as nearly optimal settings in C10 with the imperfect model. For all the test experiments with the time-expanded sampling, the covariance localization radius and inflation factor are simply set to 8 km and 1.07, respectively, that is, to be the same as in C30 with no further tuning.

To examine the effectiveness of the time-expanded sampling in comparisons with the conventional approach used in C10 and C30, test experiments are designed by setting $N_s = 10$ (which is the same as in C10) but with $M > 0$. As explained earlier, the sampling time window is constrained by $2M\tau \leq 2T$ in this study. The sampling time interval is thus confined by $\Delta t \leq \tau \leq T/M$, where Δt ($=6$ s) is the time step of ARPS integration and T ($=300$ s) is the length of each assimilation cycle, which is the same as the radar volume scan time interval here. Since T is short, large M will give rise to small τ so that the time expanded samples may be too similar or strongly correlated to provide much benefit. As we will see later, when M is increased from 0 to 1 with a fixed N_s ($=10$), the analyses and subsequent predictions are improved significantly (compared with those obtained with $M = 0$ in C10 without using the time-expanded sampling). However, when M is increased further (from 1 to 2 and beyond), it is found that the analyses and subsequent predictions are barely improved (see section 4d). Besides, large M will increase the computational cost. We therefore will only consider small values of, for example, $M = 1$ and 2, in this study. For the test experiments with $N_s = 10$ and $M = 1$, the initial 10 standard ensemble members are generated by adding random perturbations in the same way as in C10, but the total ensemble size for each analysis is expanded to $N = N_s(2M + 1) = 30$ by time-expanded sampling. In this case (with $N_s = 10$ and $N = 30$), the test experiment can be denoted by E10 \times 3 τ 90 (or E10 \times 3 τ 120, . . .) for $\tau = 90$ (or 120 s, . . .).

b. Evaluation parameters

To evaluate the performances of the test experiments in comparison with that of C30, we need to introduce and precisely define the following terms:

- The first term is the rms error of the expanded-ensemble mean (immediately before or after each analysis time) averaged over regions covered by the V_r observations (where reflectivity exceeds 10 dBZ). For the i th variable, denoted by x_i , this term is defined by

$$\varepsilon_i(t_{j\pm}) \equiv \left\{ \left[\overline{x_{ik}(t_{j\pm} + m\tau)} - x_i'(t_j) \right]^2 \right\}_p^{1/2}, \quad (7)$$

where t_{j-} (or t_{j+}) denotes the time immediately before (or after) the j th analysis, $x_{ik}(t_{j-} + m\tau)$ denotes the prior forecast ensemble member of x_i sample at

$t_{j-} + m\tau$ from the k th forecast run, $x_{ik}(t_{j+} + m\tau)$ denotes the posterior analysis ensemble member obtained from $x_{ik}(t_{j-} + m\tau)$ at the end of the j th assimilation cycle, the overbar labeled with mk denotes the double average of $()$ over m ($=0, \pm 1, \dots, \pm M$ for the $2M + 1$ sampling time levels) and over k ($=1, 2, \dots, N_s$ for N_s runs), x_i' denotes the true value of x_i , and $\{ \}_p$ denotes the spatial average over the grid points covered by observations (where the reflectivity > 10 dBZ).

- The second term is the rms error of the ensemble mean forecast (beyond the assimilation window, i.e., $t > t_{16} = 100$ min in our case) averaged over the entire domain. For the i th variable x_i , this term is defined by

$$\varepsilon_i^f(t) \equiv \left\{ \left[\overline{x_{ik}(t)} - x_i'(t) \right]^2 \right\}^{1/2} \quad \text{for } t > t_{16} = 100 \text{ min}, \quad (8)$$

where $x_{ik}(t)$ denotes the ensemble member of x_i sampled at t from the k th forecast run for $t > t_{16}$ beyond the assimilation window, the overbar labeled with k denotes the average of $()$ over k ($=1, 2, \dots, N_s$ for the N_s runs), and $\{ \}$ denotes the average over all the grid points. Note that only standard members are sampled in the forecast period beyond the assimilation window, so the forecast ensemble mean here has to be over the standard members only.

- The third term is the spatially averaged spread of the expanded ensemble at $t = t_{j\pm}$. For the i th variable, this term is defined by

$$s_i(t_{j\pm}) \equiv \left\{ \left[\overline{x_{ik}(t_{j\pm} + m\tau) - \overline{x_{ik}(t_{j\pm} + m\tau)}^{mk}} \right]^2 \right\}^{1/2}. \quad (9)$$

- The fourth term is the spatially averaged spread of the time-expanded ensemble members sampled (say, for the j th analysis) from each run and averaged over all the N_s runs. For the i th variable, this term is defined by

$$s_i^e(t_{j\pm}) \equiv \left\{ \left[\overline{x_{ik}(t_{j\pm} + m\tau) - \overline{x_{ik}(t_{j\pm} + m\tau)}^m} \right]^2 \right\}^{1/2}, \quad (10)$$

where the overbar labeled with m denotes the average of $()$ over m ($=0, \pm 1, \dots, \pm M$ for the $2M + 1$ sampling time levels).

- The fifth term is the spread of the standard members sampled (for the j th analysis) from all the N_s runs. For the i th variable, this parameter is defined by

$$s_i^s(t_{j\pm}) \equiv \left\{ \left[\overline{x_{ik}(t_{j\pm}) - \overline{x_{ik}(t_{j\pm})}^k} \right]^2 \right\}^{1/2}. \quad (11)$$

The above definitions of rms errors and spread are consistent with those used in TX05 and XTD06, with

the main difference being the introduction of time-expanded members here. Using the above terms, we introduce the following four parameters:

- The first parameter is the ratio between $\varepsilon_i(t_{j+})$ and $\varepsilon_i(t_{j+})_{C30}$ averaged over the 16 analysis time levels and 11 model variables, where $\varepsilon_i(t_{j+})_{C30}$ is defined in the same way as $\varepsilon_i(t_{j+})$ in Eq. (7) but for C30 specifically and only. This parameter is called the analysis relative rms error (RRE) for short and is defined by

$$\text{analysis RRE} \equiv [\varepsilon_i(t_{j+})/\varepsilon_i(t_{j+})_{C30}]_{j+}, \quad (12)$$

where $[\]_{j+}$ denotes the average over the 16 analysis time levels and $[\]_i$ denotes the average of $(\)$ over i for the 11 model variables.

- The second parameter is the averaged ratio between $\varepsilon_i^f(t)$ and $\varepsilon_i^f(t)_{C30}$, where $\varepsilon_i^f(t)_{C30}$ is defined in the same way as $\varepsilon_i^f(t)$ in Eq. (8) but specifically for C30 only. This parameter is called the forecast RRE for short and is defined by

$$\text{forecast RRE} \equiv [[\varepsilon_i^f(t)/\varepsilon_i^f(t)_{C30}]_t]_i, \quad (13)$$

where $[\]_t$ denotes the time average over the forecast period (from $t = 100$ to $t = 160$ min after the assimilation).

- The third parameter is the averaged ratio between $s_i(t_{1-})$ and $s_i(t_{1-})_{C30}$, where $s_i(t_{1-})$ is defined in Eq. (9) for the forecast ensemble at $t = t_{j-} = t_{1-}$ (immediately before the first analysis), and $s_i(t_{1-})_{C30}$ is the same as $s_i(t_{1-})$ but for C30 specifically and only. This parameter is called the initial relative spread (RS) for short and is defined by

$$\text{initial RS} \equiv [s_i(t_{1-})/s_i(t_{1-})_{C30}]_i. \quad (14)$$

- The fourth parameter is the averaged ratio between $s_i^e(t_{1-})$ and $s_i^s(t_{1-})$, where $s_i^e(t_{1-})$ is defined in Eq. (10) with $t_{j-} = t_{1-}$ and $s_i^s(t_{1-})$ is defined in Eq. (11) with $t_{j-} = t_{1-}$. This parameter is called the initial spread ratio (SR) for short and is defined by

$$\text{initial SR} \equiv [s_i^e(t_{1-})/s_i^s(t_{1-})]_i. \quad (15)$$

The above first two parameters are introduced to facilitate the comparison between the proposed time-expanded sampling with the conventional approach. The last parameter will be used to preselect the sampling time interval.

In Eqs. (14) and (15), the initial RS and SR are defined at $t = t_{1-}$. These definitions can be extended to any $t = t_{j\pm}$, and one can verify that

$$\begin{aligned} & \overline{[x_{ik}(t_{j\pm} + m\tau) - \overline{x_{ik}(t_{j\pm} + m\tau)}]^{2mk}} \\ &= \overline{[x_{ik}(t_{j\pm} + m\tau) - \overline{x_{ik}(t_{j\pm} + m\tau)}]^m + \overline{x_{ik}(t_{j\pm} + m\tau)}^m - \overline{x_{ik}(t_{j\pm} + m\tau)}^{mk}]^{2mk}} \\ &= \overline{[x_{ik}(t_{j\pm} + m\tau) - \overline{x_{ik}(t_{j\pm} + m\tau)}]^m]^{2mk}} + \overline{[\overline{x_{ik}(t_{j\pm} + m\tau)}^m - \overline{x_{ik}(t_{j\pm} + m\tau)}^{mk}]^{2mk}} \\ &= \overline{[x_{ik}(t_{j\pm} + m\tau) - \overline{x_{ik}(t_{j\pm} + m\tau)}]^m]^{2mk}} + \overline{[x_{ik}(t_{j\pm}) - \overline{x_{ik}(t_{j\pm})}]^k]^{2k}}, \end{aligned} \quad (16)$$

where $\overline{(\)}^{mkm} = \overline{(\)}^{mk}$ and

$$\overline{[x_{ik}(t_{j\pm} + m\tau) - \overline{x_{ik}(t_{j\pm} + m\tau)}]^m [\overline{x_{ik}(t_{j\pm} + m\tau)}^m - \overline{x_{ik}(t_{j\pm} + m\tau)}^{mk}]^{mk}} = 0$$

are used in the third step, and $\overline{x_{ik}(t_{j\pm} + m\tau)}^m \approx x_{jk}(t_{j\pm})$ is used in the last step. Applying $\{ \}$ to the two sides of Eq. (16) gives

$$s_i(t_{j\pm})^2 \approx s_i^e(t_{j\pm})^2 + s_i^s(t_{j\pm})^2. \quad (17)$$

This result indicates that the variance (i.e., the squared rms spread) of the total ensemble is approximately equal to the sum of the variance of the time-expanded members and the variance of the standard members. Note that $\overline{x_{ik}(t_{j\pm} + m\tau)}^m \approx x_{jk}(t_{j\pm})$ is the only approximation used in deriving Eqs. (16) and (17), so the ac-

curacy or validity of Eq. (17) depends solely on this approximation.

Note that the ensemble spread can be easily computed for the first analysis at $t = t_{1-}$ before the assimilation starts, but computing the spread ratio for a subsequent analysis, say, at $t = t_{j+}$ (for $j \geq 1$) requires the ensemble to be updated through j assimilation cycles. Because we intend to use the initial SR to guide the design of assimilation before the assimilation is actually carried out, we have simply set $t = t_{1-}$ in Eqs. (14) and (15). The initial RS and SR will be computed at $t = t_{1-}$ (before the assimilation starts) in all the experiments presented in this paper.

4. Perfect model assimilation experiments

In this section, a total of 13 perfect model assimilation experiments are performed and analyzed in details. These include two control experiments (C30 and C10) and six test experiments with $N_s = 10$ and $M = 1$ (thus $N = 30$), and five more experiments with $N_s = 10$ and $M = 2$ ($N = 50$). We first examine the six experiments with $M = 1$ and compare them against the two controls. In the first eight rows of Table 1 are the values of the four parameters calculated for these experiments, according to Eqs. (12)–(15). As shown in the table, measured by the smallness of the analysis RRE and forecast RRE, there is an optimal sampling time interval within the sampling time window (constrained by $2M\tau \leq 2T = 10$ min) for $M = 1$. This feature and its implications are examined in the following section.

a. Optimal time interval for time-expanded sampling

As we can see from the analysis RREs and forecast RREs listed in Table 1, the six test experiments perform significantly better than C10, and they all have the same number of prediction runs, that is, $N_s = 10$. Since their analysis RREs and forecast RREs are around 1, the six test experiments have about the same level of performance as C30. Among the six test experiments, two experiments, that is, E10×3τ168 and E10×3τ186, perform better than C30, and the best is E10×3τ168. Measured by the analysis RRE and forecast RRE, the optimal sampling time interval is clearly given by $\tau = 168$ s among the six selected values of τ for the above test experiments. The above results demonstrate that (i) the time-expanded sampling can improve the filter performance, and (ii) the improvement can be optimized by properly selecting the sampling time interval. Clearly, to improve the ensemble representation by using the time-expanded sampling, τ should be sufficiently large to reduce the similarity and correlation between the time-expanded members but not too large to cause the deterioration of the ensemble representation of the pdf for the forecast background field at the analysis time.

As shown by the results of E10×3τ168 and E10×3τ186 in Table 1, when the sampling time interval τ is optimal or nearly optimal, the initial RS is slightly larger than 1.5 and the initial SR is around 1. According to the definition in Eq. (14), the initial RS measures the initial ensemble spread (immediately before the first analysis) in the concerned experiment with respect to that in C30. As will be shown later (see Fig. 5), the initial ensemble spread in C30 is insufficient as it is significantly smaller than the rms error of the ensemble

TABLE 1. Evaluation parameter values for perfect model assimilation experiments. See Eqs. (12)–(15) for the definitions of the four evaluation parameters.

Expt	Analysis RRE	Forecast RRE	Initial RS	Initial SR
C30	1.00	1.00	1.00	
C10	1.49	1.26	0.83	
E10×3τ90	1.18	1.01	1.03	0.58
E10×3τ120	1.11	0.99	1.19	0.73
E10×3τ168	0.95	0.91	1.55	1.00
E10×3τ186	0.97	0.93	1.72	1.11
E10×3τ210	1.02	0.98	1.99	1.28
E10×3τ240	1.10	1.08	2.46	1.57
E10×5τ72	1.11	1.02	1.25	0.74
E10×5τ84	0.96	0.94	1.41	0.86
E10×5τ90	0.93	0.93	1.49	0.93
E10×5τ120	0.99	1.02	2.14	1.36
E10×5τ144	1.21	1.05	3.37	2.23

mean. In E10×3τ168, the initial ensemble spread is enhanced by a factor of 1.55 (as indicated by the initial RS). This enhanced spread partially explains the improved performance of E10×3τ168 relative to that of C30.

As shown in Table 1, the initial SR is 1.0 for E10×3τ168. This implies that $s_i^e(t_{1-}) \approx s_i^s(t_{1-})$ according to Eq. (15), but this approximation is crude because of the variability of $s_i^e(t_{1-})/s_i^s(t_{1-})$ over the 11 variables, especially over the hydrometeor variables ($q_c, q_r, q_i, q_s,$ and q_h). With this approximation, we have $s_i(t_{1-}) \approx \sqrt{2}s_i^s(t_{1-})$ according to Eq. (17), so the total ensemble spread is about $\sqrt{2}$ times the standard member spread. Note that the standard members are generated initially by adding random perturbations only to the (u, v, w, θ) component fields in E10×3τ168 in the same way as in C10, so their initial spread should be the same as that in C10 but the initial spreads are reflected mainly in the above four component fields. For these four component fields, the initial spreads $s_i(t_{1-})$ in C10 are very close to those in C30 and $s_i(t_{1-})/s_i(t_{1-})_{C30}$ are between 0.96 and 0.99 (larger than the variable-averaged value listed for C10 in Table 1). This implies that $s_i^s(t_{1-})$ in E10×3τ168 are very close to $s_i(t_{1-})$ in C30 for these four component fields, while $s_i^e(t_{1-}) \approx s_i^s(t_{1-})$ in E10×3τ168 as explained above. This indicates again that the initial ensemble spread, at least for the above four component fields, is enhanced by about a factor of $\sqrt{2}$ by the inclusion of time-expanded members. This suggests that the sampling time interval τ may be pre-selected by tuning the initial SR to 1, and the detailed procedure will be discussed at the end of section 5.

In the above test experiments, as N_s is 10 instead of 30, the time-expanded sampling saves more than a half CPU cost in integrating and updating the ensemble in

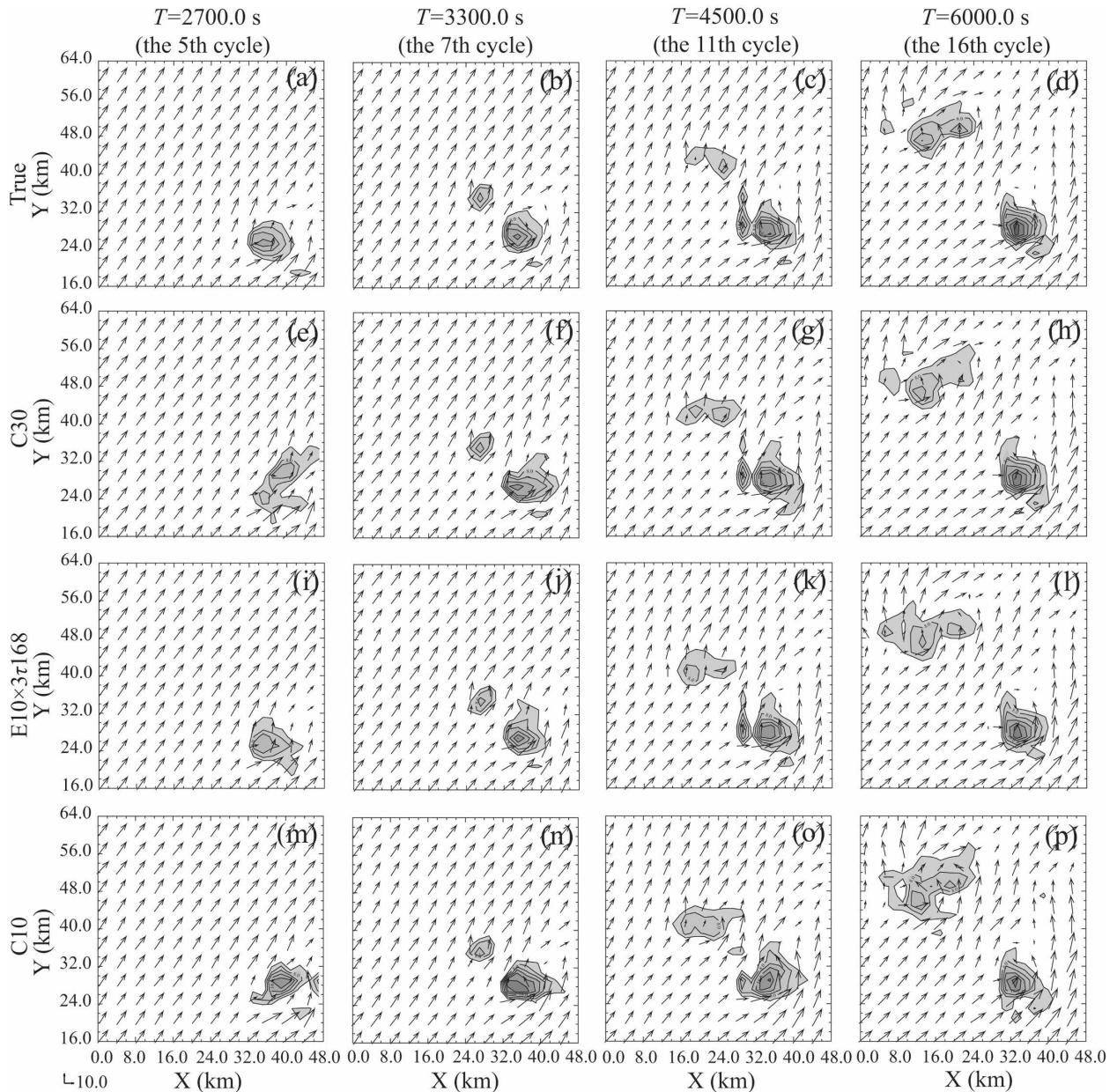


FIG. 1. Vertical velocity (contours and shading at intervals of 4 and $\geq 4 \text{ m s}^{-1}$) and horizontal wind (vectors plotted every other grid point) at level $z = 6 \text{ km}$: (a)–(d) true simulation and analyses from (e)–(h) C30, (i)–(l) $E10 \times 3\tau168$, and (m)–(p) C10, at the 5th, 7th, 11th, and 16th cycles during the assimilation period. The vector scale for the horizontal wind is shown by the arrow (10 m s^{-1}) at the lower-left corner.

comparison with the conventional approach in C30. When the sampling time interval τ is properly selected to make the initial RS close to 1 (for $M = 1$ with a perfect model setting), the time-expanded sampling performs better than the conventional approach in C30. Note that C30 performs similar to experiment VrP in TX05 (see their Fig. 4). The performances of the test experiments will be examined in details and compared

with the performances of C10 and C30 in the subsequent sections.

b. Ensemble mean analyses and forecasts

The true fields of vertical velocity and horizontal wind at $z = 6 \text{ km}$ are plotted in a time series in the first row of Fig. 1 against the ensemble mean analyses in C30 (second row), $E10 \times 3\tau168$ (third row), and C10 (last

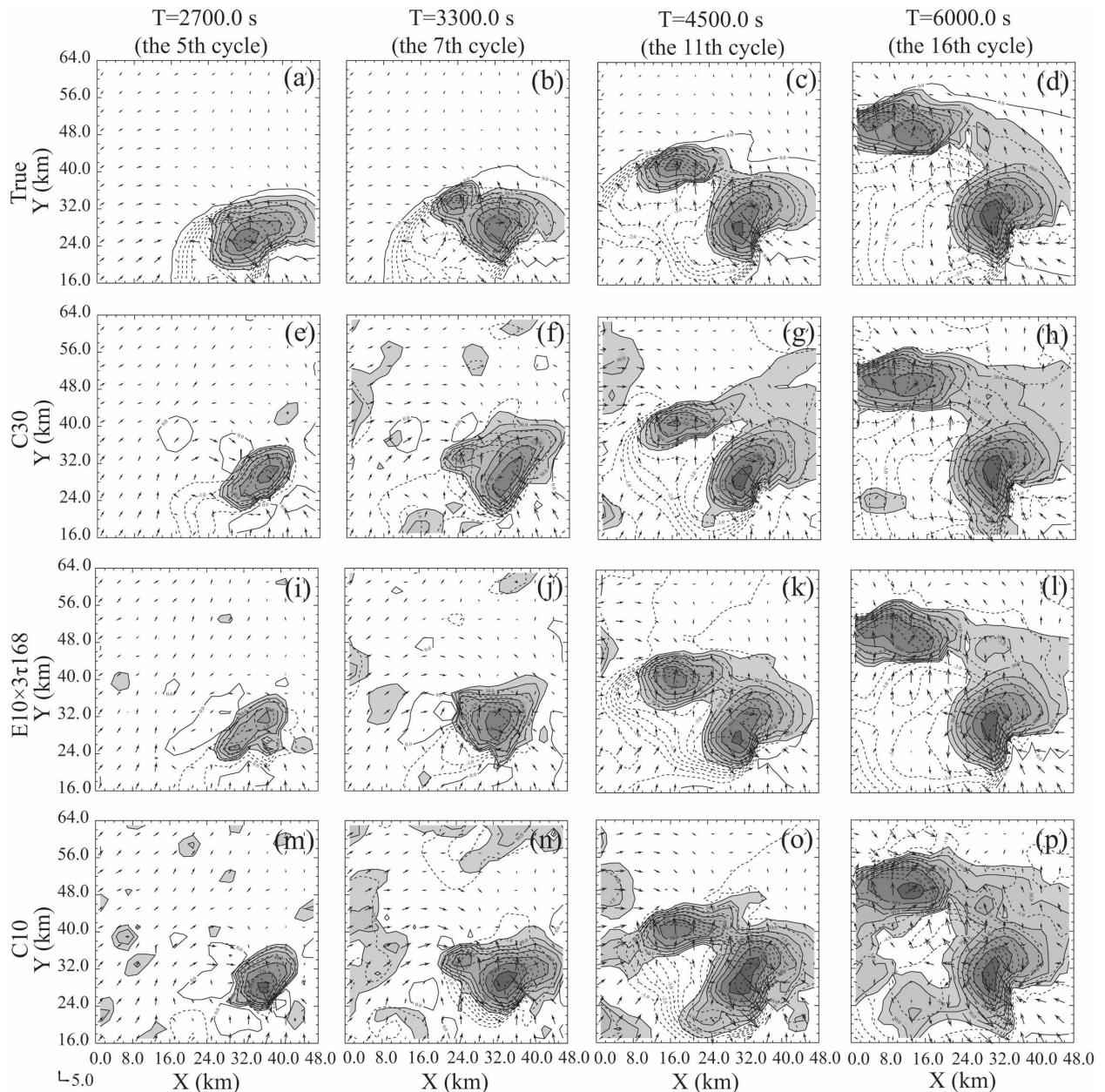


FIG. 2. Horizontal perturbation wind (vectors), perturbation potential temperature (thick black lines for 0 K and thin dashed contours for negative at 0.5-K intervals), and computed reflectivity (thin solid contours and shading at intervals of 5 dBZ, starting from 15 dBZ) at $z = 250$ m: (a)–(d) true simulation, and analyses from (e)–(h) C30, (i)–(l) $E10 \times 3r168$, and (m)–(p) C10. The arrow (10 m s^{-1}) at the lower-left corner shows the vector scale for the horizontal wind.

row). As shown, after five analysis cycles, the basic structures of updraft and horizontal flow are reasonably retrieved. It is evident that $E10 \times 3r168$ performs better than C10 in terms of their closeness to the truth fields. Here $E10 \times 3r168$ performs even slightly better than C30 for most of the 11 model variables over the first nine analysis cycles (see Fig. 3). By the fifth cycle, the low-level cold pool and the associated divergent flow in

C10, C30, and $E10 \times 3r168$ are all weaker than the true fields, especially in $E10 \times 3r168$ (Fig. 2i), and the cold pool centers are dislocated slightly to the northeast from the true center. In the subsequent cycles, hydrometeor retrievals are improved significantly and the analyses continue to improve, especially in $E10 \times 3r168$. In this case, as shown in Fig. 2, the magnitude and location of the perturbation potential temperature be-

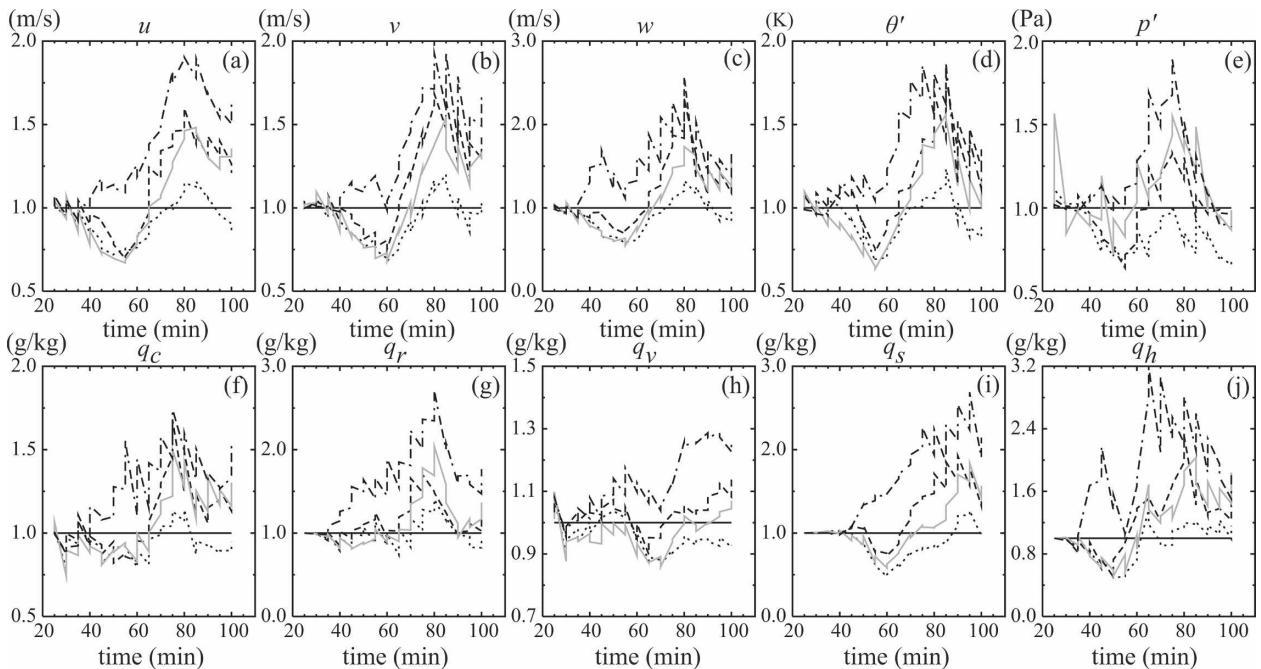


FIG. 3. Ensemble mean rms error ratios, $\varepsilon_i(t_{j\pm})/\varepsilon_i(t_{j\pm})_{C30}$, plotted as functions of assimilation time for (a) u , (b) v , (c) w , (d) perturbation potential temperature θ' , (e) perturbation pressure p' , (f) water vapor specific humidity q_w , (g) mixing ratios for cloud water q_c , (h) rainwater q_r , (i) snow q_s , and (j) hail q_h produced in C30 (straight thick solid lines), C10 (dashed-dotted lines), E10 \times 3 τ 90 (dashed lines), E10 \times 3 τ 168 (dotted lines), and E10 \times 3 τ 240 (thin solid lines).

come reasonably good, and the low-level flow and reflectivity patterns as well as the shape of the cold pool also agree quite well with the true fields, again, especially in E10 \times 3 τ 168.

The accuracies of the ensemble mean analyses and forecasts are evaluated quantitatively by the rms errors defined in Eqs. (7) and (8). Since C30 has basically the same performance as the VrP experiment of TX05 (see their Fig. 4), it is not necessary to show the rms errors in C30. By using the results of C30 as benchmarks, the rms error for each variable in each of the remaining experiments relative to that in C30 can be conveniently examined by their ratio, that is, $\varepsilon_i(t_{j\pm})/\varepsilon_i(t_{j\pm})_{C30}$ [or $\varepsilon_i^f(t)/\varepsilon_i^f(t)_{C30}$] for the analysis (or forecast) mean for the i th variable. These ratios are similar to those defined in Eqs. (12) and (13) but not averaged (neither in time nor over the 11 variables). The rms error ratios of the ensemble means, $\varepsilon_i(t_{j\pm})/\varepsilon_i(t_{j\pm})_{C30}$, are plotted in Fig. 3 as functions of time over the entire assimilation period (16 cycles) for 10 variables (u , v , w , θ' , p' , q_w , q_c , q_r , q_s , and q_h) produced in C10, E10 \times 3 τ 90, E10 \times 3 τ 168, and E10 \times 3 τ 240. As shown, C10 has the largest rms errors over 16 analysis cycles for all variables. The experiments E10 \times 3 τ 90 and E10 \times 3 τ 240 perform better than C30 during the first 9 cycles, and E10 \times 3 τ 168 performs better than C30 in the first 10 and last 2 cycles.

The forecast ensemble mean rms errors, $\varepsilon_i^f(t)$, are

plotted in Fig. 4 for w , θ' , and q_r produced in C10, C30, and E10 \times 3 τ 168 as functions of time over the first hour of the forecast launched immediately after the last (16th) analysis (at $t = 100$ min). As shown, E10 \times 3 τ 168 continues performing better than C30 and much better than C10 during this forecast period. The results in Figs. 3 and 4 show again that the time-expanded sampling cannot only save CPU cost but also improve the analysis and forecast, especially when the sampling time interval is properly adjusted to optimal or nearly optimal.

c. Ensemble spread and consistency ratio

With a limited ensemble size, the conventional EnKF tends to underestimate the analysis uncertainty. Because of this, some empirical treatments, such as the covariance inflation (Anderson 2001) and “double” ensemble (Houtekamer and Mitchell 1998), have been introduced to improve the analysis. In this aspect, the time-expanded sampling approach appears to have an advantage because the sampling time interval can be adjusted to improve the ensemble spread not only initially but also in subsequent cycles. This may partially explain why the time-expanded sampling improves the filter performance, especially when the sampling time interval is properly selected.

Ideally, the ensemble spread should satisfy a consistency relationship with the ensemble mean forecast

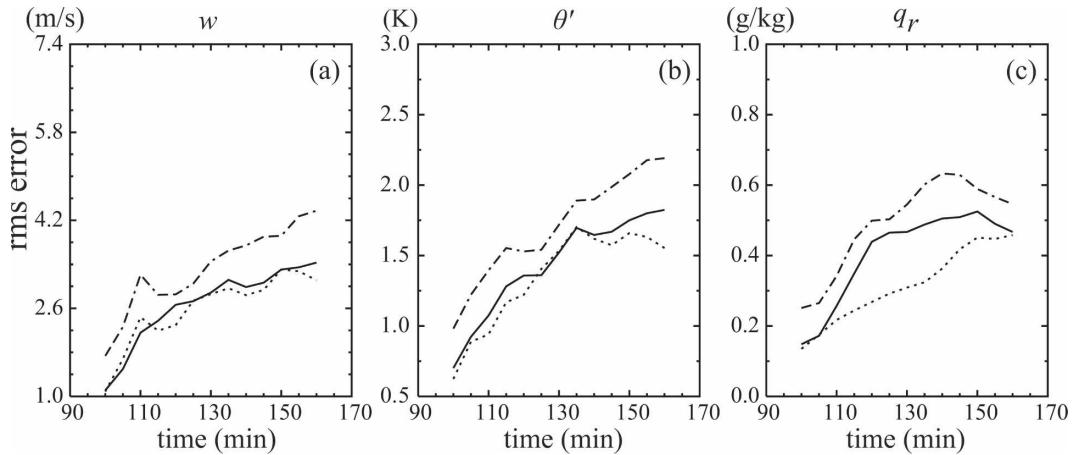


FIG. 4. Forecast ensemble mean rms errors, $\varepsilon_i^f(t)$, plotted as functions of forecast time ($160 \text{ min} \geq t > t_{16} = 100 \text{ min}$, beyond assimilation window) for (a) w , (b) θ' , and (c) q_r produced in C10 (dashed–dotted lines), C30 (thick solid lines), and E10 \times 3 τ 168 (dotted lines).

and analysis rms errors. This consistency relation requires the ratio between the ensemble spread and the ensemble mean rms error to be statistically equal to $\sqrt{N/(N + 1)}$ (Murphy 1988). We may call this ratio the consistency ratio for short and define it formally by

$$\text{consistency ratio} = s_i(t_{j\pm})/\varepsilon_i(t_{j\pm}), \quad (18)$$

where $s_i(t_{j\pm})$ is the ensemble spread and $\varepsilon_i(t_{j\pm})$ is the rms error of the ensemble mean averaged over the grid

points where the reflectivity $>10 \text{ dBZ}$, for the i th variable at the time immediately before (or after) each analysis. Here, as in Eq. (7), t_{j-} (or t_{j+}) denotes the time immediately before (or after) the j th analysis.

Figure 5 shows the consistency ratios for 10 variables ($u, v, w, \theta', p', q_v, q_c, q_r, q_s,$ and q_h) produced in C10, C30, and E10 \times 3 τ 168. As shown, the consistency ratios in C10 and C30 undergo similar variations and they are always below and often significantly below their respec-

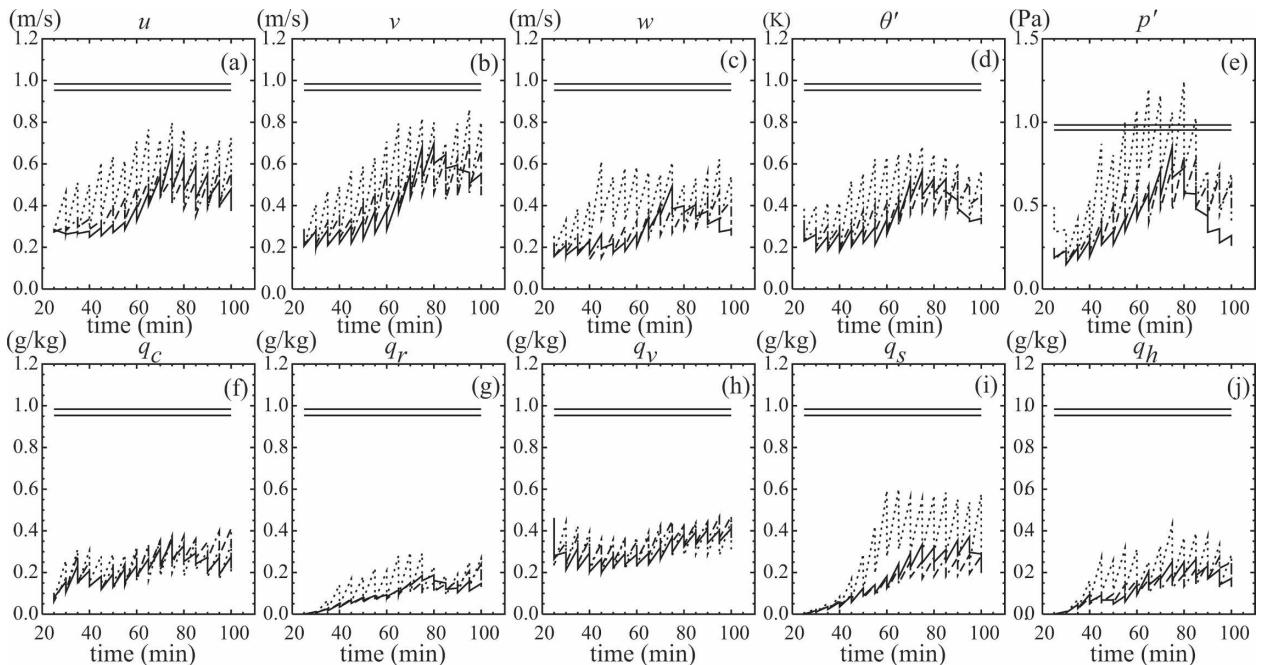


FIG. 5. Consistency ratios [see Eq. (18)] for (a) u , (b) v , (c) w , (d) θ' , (e) p' , (f) q_v , (g) q_c , (h) q_r , (i) q_s , and (j) q_h produced in C10 (dashed–dotted lines), C30 (thick solid lines), and E10 \times 3 τ 168 (dotted lines). The two straight solid lines are the values of $\sqrt{N/(N + 1)}$ for $N = 10$ and 30.

tive ideal values, that is, $\sqrt{N/(N+1)} = 0.984$ for $N = 30$ and $\sqrt{N/(N+1)} = 0.953$ for $N = 10$ as marked by the two straight solid lines in Fig. 5. The consistency ratios in E10×3τ168 are larger than those in C30 and C10, and they go up and down through each assimilation cycle more dramatically than those in C30 and C10. The sharp increases of the ratios in the forecast step are clearly due to the time-expanded sampling used in E10×3τ168. Because of this, the consistency ratios in E10×3τ168 become closer or much closer to the ideal value of $\sqrt{N/(N+1)}$ ($=0.984$ for $N = 30$) than in C30.

d. Experiments with $M = 2$

So far all the experiments are performed with $M \leq 1$. As explained in section 3a, the sampling time window is constrained by $2M\tau < 2T$. Because of this, the sampling time interval is constrained by $\tau < T/M$. Under this constraint, five test experiments are performed with M increased to 2. The results are listed in the last five rows of Table 1. Clearly, these five test experiments perform significantly better than C10, with two of them (E10×5τ84 and E10×5τ90) performing better than C30, and one of them (E10×5τ90) producing the best analysis, which is slightly better than the best analysis by E10×3τ168 using $M = 1$. The forecast RRE, however, becomes slightly larger in E10×5τ90 than in E10×3τ168. When $M = 2$, the optimal sampling time interval among the five test experiments is 90 s as shown in the last five rows of Table 1.

In E10×5τ90, the initial ensemble spread is enhanced 1.49 times relative to that in C30 (as indicated by the initial RS), and the initial SR is 0.93 and thus is still very close to 1. Note that these parameter values are very close to those listed for E10×3τ168 in Table 1, so the optimal value of τ ($=90$ s) obtained for $M = 2$ in E10×5τ90 can be explained in the same way as for $M = 1$ in E10×3τ168 in section 4a. These parameter values are about the same (i.e., initial RS ≈ 1.5 and initial SR ≈ 1) for $M = 1$ and 2.

When M is increased from 1 to 2, the time-expanded members are sampled more frequently (with reduced sampling time intervals) from the same run and thus become more similar to each other. Thus, there is a limit in improving the pdf represented by the time-expanded members by increasing M . This explains why E10×5τ90 (with $M = 2$) does not perform obviously better than E10×3τ168 (with $M = 1$). Note that N_s is set to 10 in all the above test experiments. As M is increased from 1 to 2 with this setting, the CPU cost is nearly doubled in the analysis step. Thus, increasing M to 2 and beyond provides no advantage. Decreasing N_s , say, from 10 to 5 can reduce the CPU cost, but the

TABLE 2. As in Table 1, but for imperfect model assimilation experiments.

Expt	Analysis RRE	Forecast RRE	Initial RS	Initial SR
C30	1.00	1.00	1.00	
C10	1.15	1.13	0.80	
E10×3τ90	1.06	1.07	0.96	0.56
E10×3τ120	1.01	0.96	1.08	0.70
E10×3τ168	1.02	0.90	1.34	0.92
E10×3τ186	1.01	0.92	1.47	1.02
E10×3τ210	0.99	0.90	1.66	1.16
E10×3τ240	0.98	0.86	2.02	1.44
E10×3τ270	1.12	0.88	2.60	2.00

performance (not shown) becomes worse than that of C10 regardless of the chosen values of M and τ , even though N can be larger or much larger than 30 by selecting $M > 2$. Thus, when the time-expanded sampling is used, there is a limit in reducing N_s from a value accepted in the conventional approach (such as $N_s = 30$ in C30).

5. Imperfect model experiments

In this section, nine imperfect model experiments are performed to further test and demonstrate the capabilities of the time-expanded sampling. As explained in sections 2a and 2b, the true state is simulated by the ARPS with an enhanced resolution ($\Delta x = \Delta y = 0.5$ km). The forecast model used by the assimilation, however, is the ARPS with the coarser 2-km horizontal resolution, as originally used in TX05 and in our perfect model experiments. Thus, the model is imperfect with respect to its high-resolution counterpart, and the model error is caused by the resolution deficiency with respect to the true state and process. For the imperfect model experiments, simulated observations are generated from the high-resolution true state in the polar coordinate according to the WSR-88D radar storm-mode scan configuration, after random noise is added [see Eq.(1)].

The results of the nine experiments are summarized in Table 2. As shown by the listed analysis and forecast RREs, C10 still performs worst among the nine experiments. Even though the model is imperfect now and there is no explicit treatment or estimation of the model error, the time-expanded sampling can improve the performance as shown by the test experiments in comparison with C10. Measured by the analysis and forecast RREs, two test experiments, that is, E10×3τ210 and E10×3τ240 perform even better than C30. Measured by the forecast RREs alone, six test experiments perform better than C30, with the best being

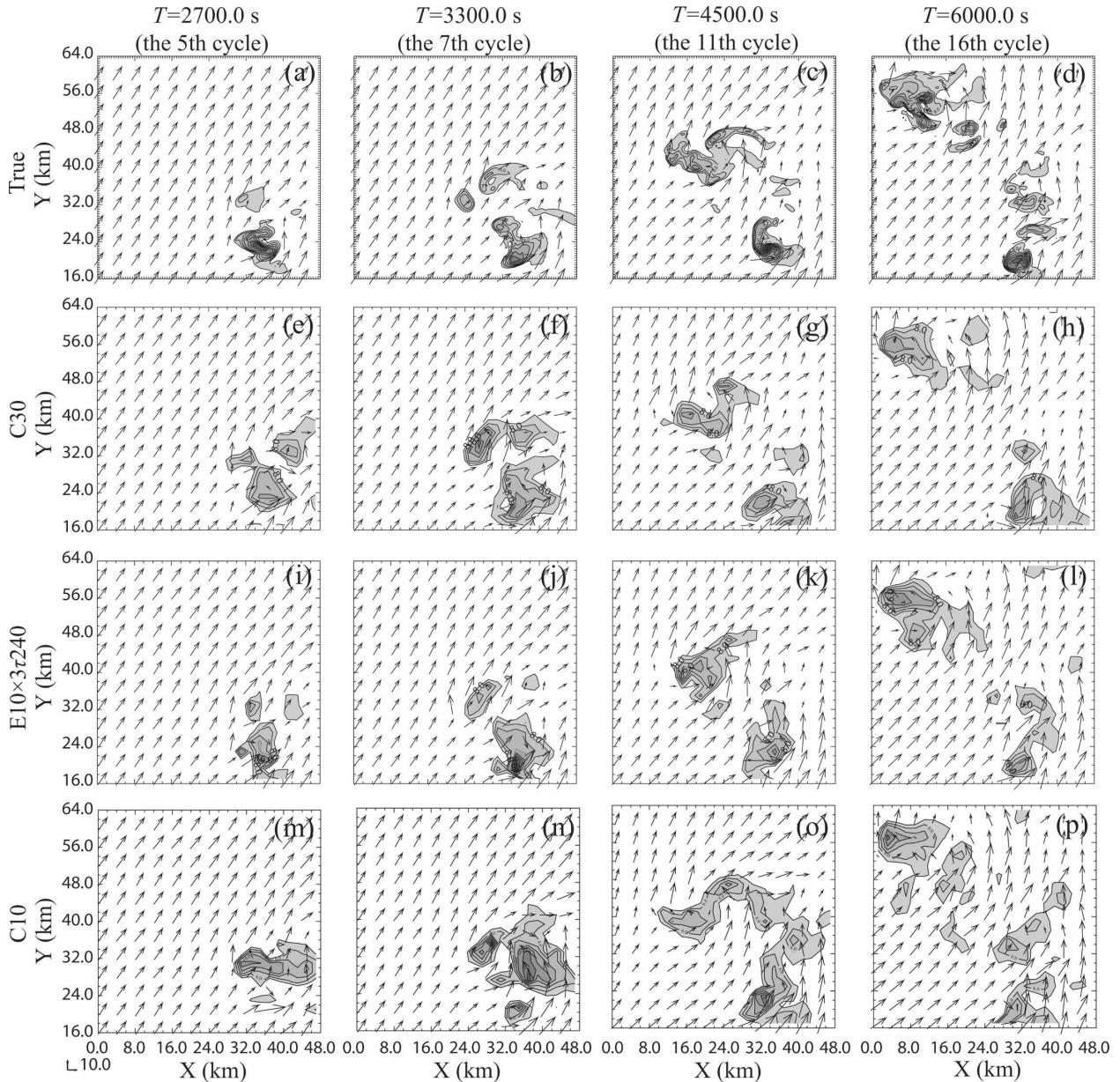


FIG. 6. As in Fig. 1 but for (a)–(d) high-resolution true simulation (vectors plotted every eight grid points) and analyses produced in the imperfect model experiments (e)–(h) C30, (i)–(l) E10×3τ240, and (m)–(p) C10 (vectors plotted every other grid point).

E10×3τ240. Therefore, the optimal sampling time interval is $\tau = 240$ s among the seven test experiments. This optimal value ($\tau = 240$ s) is larger than that ($\tau = 168$ s) in the perfect model case. This increase is expected, because it can increase the ensemble spread to account for the additional background forecast error due to model error, as discussed in the introduction.

Figure 6 shows the true fields of vertical velocity and horizontal wind at $z = 6$ km in a time series (first row) against the analyzed ensemble mean fields in C30 (second row), E10×3τ240 (third row), and C10 (last row).

As shown, the middle-level horizontal wind field is reasonably well produced by the analysis at the end of the fifth assimilation cycle in E10×3τ240. The major updraft pattern and intensity are also produced in E10×3τ240 better (closer to the true pattern) than in C30 and much better than in C10, although the updraft is slightly weaker than the true one. In the subsequent cycles, E10×3τ240 continuously performs well, and some of the detailed improvements in E10×3τ240 in comparison with C30 and C10 can be seen from Fig. 6. The low-level cold pool and the associated horizontal

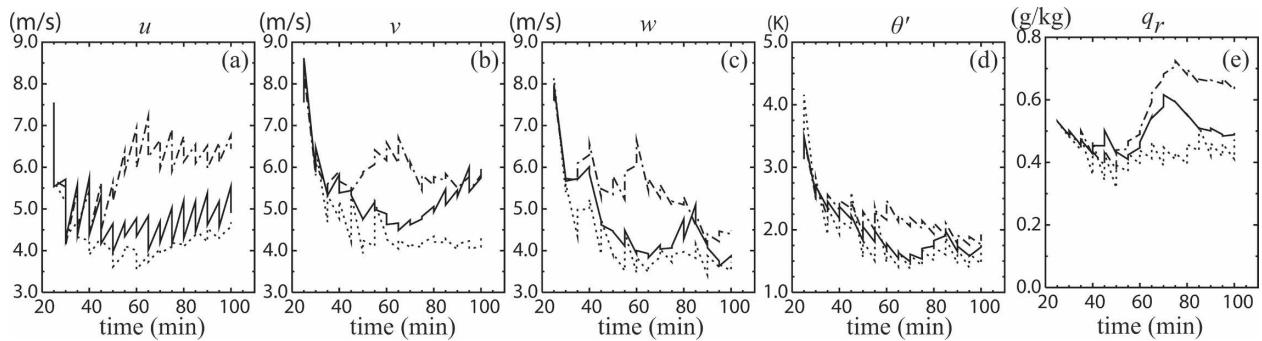


FIG. 7. As in Fig. 3 but for the ensemble mean rms errors, $\varepsilon_r(t_{\pm})$, plotted for (a) u , (b) v , (c) w , (d) θ' , and (e) q_r produced in the imperfect model experiments C10 (dashed-dotted lines), C30 (thick solid lines), and E10 \times 3 τ 240 (dotted lines).

wind divergence produced in E10 \times 3 τ 240 are also better than in C30 and much better than in C10 (not shown).

The rms errors [defined in Eq. (7)] of the analyzed mean fields of (u , v , w), perturbed potential temperature θ' , and rainwater q_r produced in C10, C30, and E10 \times 3 τ 240 are plotted as functions of time in Fig. 7. As shown, since the model is imperfect, the rms errors are not always reduced effectively and the analysis error reduction is not always able to overcome the forecast error growths. The rms errors undergo only limited reductions and remain significant after all the 16 cycles, especially in C10. Nevertheless, E10 \times 3 τ 240 has the smallest rms errors in all the variable fields during the entire assimilation period.

Figure 8 shows the rms errors [defined in Eq. (8)] of the ensemble mean forecasts of w , θ' , and q_r from C10, C30, and E10 \times 3 τ 240 for the first hour after the assimilation (from $t = 100$ to 160 min). As shown, the forecast rms errors in E10 \times 3 τ 240 are significantly smaller

than those in C30 and much smaller than those in C10. In comparison with the results presented in Fig. 4, the forecast error reductions made by the time-expanded sampling in E10 \times 3 τ 240 are clearly more significant in this imperfect model case than in the perfect model case.

The consistency ratios [defined in Eq. (18)] are shown in Fig. 9 for the same 10 model variables as in Fig. 5 but produced by the imperfect model experiments C10, C30, and E10 \times 3 τ 240. Note that the consistency ratios produced by each imperfect model experiment in Fig. 9 are slightly smaller than those produced by the perfect model counterpart in Fig. 5. As shown in Fig. 9, the consistency ratios are enhanced very significantly in E10 \times 3 τ 240 relative to those in C30 and C10. Because the optimal sampling time interval in this case is larger than that in the perfect model case, the overall ensemble spread is enhanced by the time-expanded sampling in E10 \times 3 τ 240 more significantly (toward the ideal value) than those in Fig. 5 for the perfect model

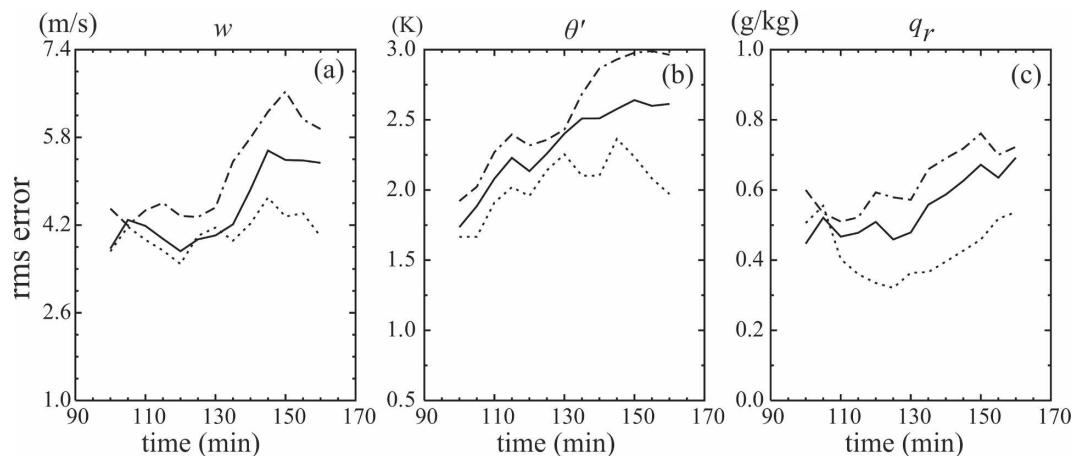


FIG. 8. As in Fig. 4 but for the imperfect model experiments C10 (dashed-dotted lines), C30 (thick solid lines), and E10 \times 3 τ 240 (dotted lines).

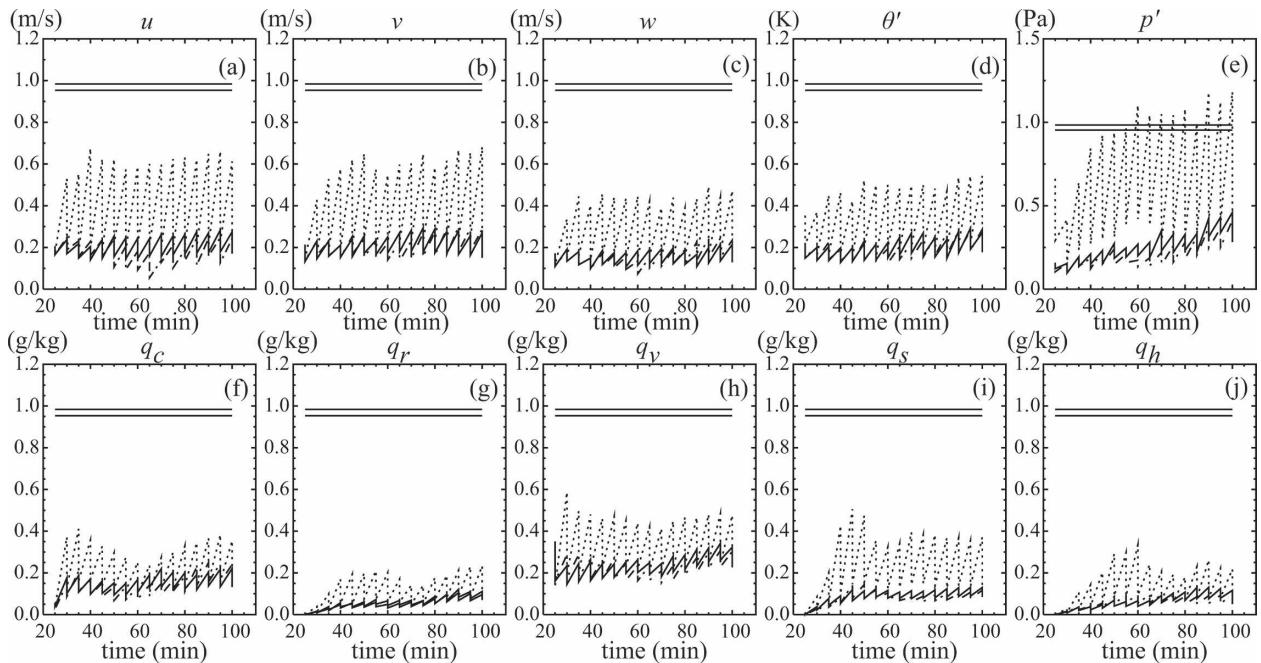


FIG. 9. As in Fig. 5 but for the imperfect model experiments C10 (dashed–dotted lines), C30 (thick solid lines), and E10×3τ240 (dotted lines).

case. This explains why the consistency ratios are enhanced more dramatically in Fig. 9 for E10×3τ240 than those in Fig. 5 for E10×3τ168 and why the ensemble mean rms errors are reduced in E10×3τ240 for the imperfect model case more significantly than those in E10×3τ168 for the perfect model case.

The results presented in this and previous sections suggest that the initial SR defined in Eq. (15) may be used to facilitate the initial selection of the sampling time interval τ (as it is calculated from the forecast ensemble before the assimilation), but the subsequent tuning of τ will require additional work especially if the model is significantly imperfect. For example, as shown by our experiments, τ may be preselected to make the initial SR close to 1 if the model is perfect or moderately larger than 1 if the model is moderately imperfect. With such an initially selected τ , assimilation experiments can be performed with time-expanded sampling to improve the filter performance. Then, innovation (observation minus background) data can be collected and used to estimate the spatially averaged background error variance (Xu et al. 2007). The estimated error variance can be used together with the ensemble spread to compute the consistency ratio [see Eq. (18)] and thus check the optimality of the preselected τ . Based on the computed consistency ratio, τ can be adjusted to improve the optimality (according to the discussions in section 4c). The above procedure may be used to facilitate applications of the proposed time-expanded

sampling to real radar data assimilation with an ensemble-based filter.

6. Conclusions

When the pdf of the model state is represented by an ensemble of state vectors in an ensemble-based filter, such as the EnKF or other hybrid filter (Lorenz 2003; Zupanski 2005), the covariance matrix is directly computed from the ensemble and this makes the filter much easier to implement than the four-dimensional variational data assimilation (4DVAR) technique (Lewis and Derber 1985; Le Dimet and Talagrand 1986; Bennett 1992; Talagrand 1997) and much cheaper to run than the classic extended Kalman filter (Jazwinski 1970; Cohn 1997). Ideally, the ensemble size should be sufficiently large to adequately represent the pdf (e.g., W. J. Martin and M. Xue 2007, unpublished manuscript). Practically, however, an ensemble-based filter has to use a limited or even small ensemble to reduce the computational cost. How to treat problems caused by limited or small ensembles is thus a primary issue for ensemble-based filters. Another primary issue for ensemble-based filters (also for 4DVAR) concerns how to treat or alleviate problems caused by unknown model errors. The time-expanded sampling proposed in this study is intended to deal with these issues. The results presented in this paper indicate that the proposed method can be a simple and effective way to improve

the filter performance and reduce the computational cost even in the presence of model error, although the method does not directly deal with model errors.

More specifically, by performing assimilation experiments with the EnSRF (Whitaker and Hamill 2002) and simulated radar observations by the ARPS (Xue et al. 2001), the proposed time-expanded sampling is shown to have the following merits: (i) It can reduce the number of prediction runs hence computational cost without reducing the effective ensemble size. (ii) When the sampling time interval is properly selected, it can improve the filter performance and alleviate spread-related problems caused by small ensemble sizes and unknown (but small or moderate) model errors more effectively than the conventional covariance inflation. It is also shown (see sections 4c and 5) that the improved filter performance can be explained by the enhanced ensemble spread and consistency ratio [see Eq. (18)] due to the time-expanded sampling. Furthermore, as the ensemble subspace spanned by the standard ensemble members is small and the ensemble space can be enlarged by the inclusion of time-expanded members, the ensemble structures can therefore be enriched by the time-expanded sampling. This may further explain why the time-expanded sampling can improve the filter performance more effectively than the optimized conventional inflation (see section 3a) when the ensemble size is small. The above summarized merits and interpretations are also supported by the results of assimilation experiments performed with the time-expanded sampling applied to a shallow-water equation model in Xu et al. (2008).

As explained in the introduction, the proposed time-expanded sampling was motivated by the consideration of timing error, but its applicability is not necessarily limited by this consideration. In fact, all ensemble predictions in this paper (as well as in Xu et al. 2008) were initialized with no systematic timing error and there was no obvious systematic timing error in the subsequent ensemble predictions. The above summarized merits and related interpretations also involve no systematic timing error. However, since the proposed approach relies on the anticipated development and/or propagation of the predicted weather system, it will become ineffective if the predicted system neither develops nor propagates. This is an obvious limitation, but the concerned scenario (no development and no propagation in the predicted system) is extremely rare and should never happen in the presence of active storms. On the other hand, as shown in Xu et al. (2008), the time-expanded sampling will become relatively ineffective when the ensemble size becomes large and the observations become complete. This appears to be a ma-

major limitation for the proposed time-expanded sampling. This limitation, however, does not reduce the potential usefulness of time-expanded sampling for real radar data assimilation, because radar observations are highly incomplete (limited to radial velocity and reflectivity only) and the practical ensemble size often has to be much (10^4 or 10^5 times) smaller than the state vector dimension in a full model (or the computations will be too expensive). As a matter of fact, the potential usefulness of time-expanded sampling for real radar data assimilation has been demonstrated by the recent study of Lu (2007).

In addition to the above summarized merits and interpretations, there could be fundamental reasons for the potential usefulness of time-expanded sampling in certain situations (including those mentioned above under the limitations discussed). In this regard, one may speculate two possible reasons (as suggested by one of the anonymous reviewers of this paper): (i) As the ensemble size is enlarged by the inclusion of time-expanded members, the time-expanded sampling could reduce the required localization and thus reduce imbalance problems that can be caused by extreme localization. (ii) Using time-expanded sampling could increase the number of possible locations where convection can be generated and thus improve the localized ensemble representation of the forecast pdf if some aspects of forecasts can be regarded as purely stochastic on the convective scale. These speculations are not yet verified in the study but deserve further investigation.

Acknowledgments. The authors are thankful to David Stensrud, Steven Weiss, and the two anonymous reviewers for their comments and suggestions that improved the presentation of the results. The authors also acknowledge the computational resources provided by the OU Supercomputing Center for Education and Research at the University of Oklahoma. The research was primarily supported by the ONR Grant N000140410312 to the University of Oklahoma and by FAA Contract IA DTFA03-01-X-9007 to NSSL. M. Xue and M. Tong were supported by NSF Grants ATM-0530814 and EEC-0313747 and M. Xue further by NSF ATM-0331756 and ATM-0331594. H. Lu, M. Xue, and S. Gao were also supported by Grant 40620120437 of the Chinese Natural Science Foundation. 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.

REFERENCES

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

- Bennett, A. F., 1992: *Inverse Method in Physical Oceanography*. Cambridge University Press, 346 pp.
- Bishop, C. H., B. J. Etherton, and S. J. Majumdar, 2001: Adaptive sampling with the ensemble transform Kalman filter. Part I: Theoretical aspects. *Mon. Wea. Rev.*, **129**, 420–436.
- Burgers, G., P. J. van Leeuwen, and G. Evensen, 1998: Analysis scheme in the ensemble Kalman filter. *Mon. Wea. Rev.*, **126**, 1719–1724.
- Cohn, S., 1997: An introduction to estimation theory. *J. Meteor. Soc. Japan*, **75**, 257–288.
- Colle, B. A., and M. Charles, 2007: Verification of extratropical cyclones within NCEP forecast models using an automated tracking algorithm. Preprints, *18th Conf. on Numerical Weather Prediction*, Park City, UT, Amer. Meteor. Soc., 6B.3. [Available online at <http://ams.confex.com/ams/pdfpapers/124612.pdf>.]
- , C. F. Mass, and D. Ovens, 2001: Evaluation of the timing and strength of MM5 and Eta surface trough passages over the eastern Pacific. *Wea. Forecasting*, **16**, 553–572.
- Dowell, D., F. Zhang, L. J. Wicker, C. Snyder, and N. A. Crook, 2004: Wind and temperature retrievals in the 17 May 1981 Arcadia, Oklahoma, supercell: Ensemble Kalman filter experiments. *Mon. Wea. Rev.*, **132**, 1982–2005.
- Evensen, G., 1994: Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *J. Geophys. Res.*, **99**, 10 143–10 162.
- , 2003: The ensemble Kalman filter: Theoretical formulation and practical implementation. *Ocean Dyn.*, **53**, 343–367.
- Gaspari, G., and S. E. Cohn, 1999: Construction of correlation functions in two and three dimensions. *Quart. J. Roy. Meteor. Soc.*, **125**, 723–757.
- Gregory, A. D., and R. H. Grumm, 2007: Forecast issues related to the unprecedented severe and high wind event of December 1, 2006. Preprints, *18th Conf. on Numerical Weather Prediction*, Park City, UT, Amer. Meteor. Soc., P1.14. [Available online at <http://ams.confex.com/ams/pdfpapers/123385.pdf>.]
- Houtekamer, P. L., and H. L. Mitchell, 1998: Data assimilation using an ensemble Kalman filter technique. *Mon. Wea. Rev.*, **126**, 796–811.
- , and —, 2001: A sequential ensemble Kalman filter for atmospheric data assimilation. *Mon. Wea. Rev.*, **129**, 123–137.
- Jazwinski, A. H., 1970: *Stochastic Processes and Filtering Theory*. Academic Press, 376 pp.
- LeDimet, F. X., and O. Talagrand, 1986: Variational algorithms for analysis and assimilation of meteorological observations: Theoretical aspects. *Tellus*, **38A**, 97–111.
- Lewis, J., and J. Derber, 1985: The use of adjoint equations to solve a variational adjustment problem with advective constraint. *Tellus*, **37A**, 309–322.
- Lorenc, A. C., 2003: The potential of the ensemble Kalman filter for NWP—A comparison with 4D-Var. *Quart. J. Roy. Meteor. Soc.*, **129**, 3183–3203.
- Lu, H., 2007: Efficient approaches in applying ensemble Kalman filter to storm-scale radar data assimilation. Ph.D. dissertation, Chinese Academy of Sciences, 115 pp.
- Manobianco, J., and P. A. Nutter, 1999: Evaluation of the 29-km Eta Model. Part II: Subjective verification over Florida. *Wea. Forecasting*, **14**, 18–37.
- Mass, C. F., D. Ovens, K. Westrick, and B. A. Colle, 2002: Does increasing horizontal resolution produce more skillful forecasts? *Bull. Amer. Meteor. Soc.*, **83**, 407–430.
- Mitchell, H. L., and P. L. Houtekamer, 2000: An adaptive ensemble Kalman filter. *Mon. Wea. Rev.*, **128**, 416–433.
- , —, and G. Pellerin, 2002: Ensemble size, balance and model-error representation in an ensemble Kalman filter. *Mon. Wea. Rev.*, **130**, 2791–2808.
- Murphy, J. M., 1988: The impact of ensemble forecasts on predictability. *Quart. J. Roy. Meteor. Soc.*, **114**, 463–493.
- Ray, P. S., B. Johnson, K. W. Johnson, J. S. Bradberry, J. J. Stephens, K. K. Wagner, R. B. Wilhelmson, and J. B. Klemp, 1981: The morphology of severe tornadic storms on 20 May 1977. *J. Atmos. Sci.*, **38**, 1643–1663.
- Snyder, C., and F. Zhang, 2003: Assimilation of simulated Doppler radar observations with an ensemble Kalman filter. *Mon. Wea. Rev.*, **131**, 1663–1677.
- Talagrand, O., 1997: Assimilation of observations, an introduction. *J. Meteor. Soc. Japan*, **75**, 81–99.
- Tippett, M. K., J. L. Anderson, C. H. Bishop, T. M. Hamill, and J. S. Whitaker, 2003: Ensemble square root filters. *Mon. Wea. Rev.*, **131**, 1485–1490.
- Tong, M., and M. Xue, 2005: Ensemble Kalman filter assimilation of Doppler radar data with a compressible nonhydrostatic model: OSS experiments. *Mon. Wea. Rev.*, **133**, 1789–1807.
- Weiss, S., and Coauthors, 2007: The NOAA Hazardous Weather Testbed: Collaborative testing of ensemble and convection-allowing WRF models and subsequent transfer to operations at the Storm Prediction Center. Preprints, *18th Conf. on Numerical Weather Prediction*, Park City, UT, Amer. Meteor. Soc., 6B.4. [Available online at <http://ams.confex.com/ams/pdfpapers/124772.pdf>.]
- Whitaker, J. S., and T. M. Hamill, 2002: Ensemble data assimilation without perturbed observations. *Mon. Wea. Rev.*, **130**, 1913–1924.
- Xu, Q., 1996: Generalized adjoint for physical processes with parameterized discontinuities. Part I: Basic issues and heuristic examples. *J. Atmos. Sci.*, **53**, 1123–1142.
- , 1997: Generalized adjoint for physical processes with parameterized discontinuities. Part IV: Problems in time discretization. *J. Atmos. Sci.*, **54**, 2722–2728.
- , K. Nai, and L. Wei, 2007: An innovation method for estimating radar radial-velocity observation error and background wind error covariances. *Quart. J. Roy. Meteor. Soc.*, **133**, 407–415.
- , L. Wei, H. Lu, C. Qiu, and Q. Zhao, 2008: Time-expanded sampling for ensemble-based filters: Assimilation experiments with a shallow-water equation model. *J. Geophys. Res.*, **113**, D02114, doi:10.1029/2007JD008624.
- Xue, M., and Coauthors, 2001: The Advanced Regional Prediction System (ARPS)—A multiscale nonhydrostatic atmospheric simulation and prediction tool. Part II: Model physics and applications. *Meteor. Atmos. Phys.*, **76**, 143–165.
- , M. Tong, and K. K. Droegemeier, 2006: An OSSE framework based on the ensemble square-root Kalman filter for evaluating impact of data from radar networks on thunderstorm analysis and forecast. *J. Atmos. Oceanic Technol.*, **23**, 46–66.
- , and Coauthors, 2007: CAPS realtime storm-scale ensemble and high-resolution forecasts as part of the NOAA Hazardous Weather Testbed 2007 Spring Experiment. Preprints, *18th Conf. on Numerical Weather Prediction*, Park City, UT, Amer. Meteor. Soc., 3B.1. [Available online at <http://ams.confex.com/ams/pdfpapers/124587.pdf>.]
- Zhang, F., C. Snyder, and J. Sun, 2004: Impacts of initial estimate and observations on the convective-scale data assimilation with an ensemble Kalman filter. *Mon. Wea. Rev.*, **132**, 1238–1253.
- Zupanski, M., 2005: Maximum likelihood Ensemble filter: Theoretical aspects. *Mon. Wea. Rev.*, **133**, 1710–1726.