Preprint, 9th Symposium on Integrated Observing and Assimilation Systems
for the Atmosphere, Oceans, and Land Surfaces
American Meteorological Society, San Diego, CA 2005

9.3. Impact of Radar Configuration and Scan Strategy on Assimilation of Radar Data using Ensemble Kalman Filter

Ming Xue, Mingjing Tong and Kelvin K. Droegemeier

School of Meteorology and Center for Analysis and Prediction of Storms
University of Oklahoma, Norman, OK 73019

1. Introduction

A new National Science Foundation Engineering Research Center, the Center for Collaborative Adaptive Sensing of the Atmosphere (CASA), was established in 2003 to develop innovative observing systems for high-resolution sensing of the lower atmosphere. The development of low-cost, high-spatial density (also short range) and dynamically adaptive networks of Doppler radars with polarimetric capabilities is one of the first goals. Such networks are to probe the lower atmosphere that is often missed by the existing operational WSR-88D Doppler radar network, so as to significantly improve the detection of hazardous weather events such as tornadoes, and to provide more complete data for the initialization of numerical weather prediction models. Future upgrade to using electronically steered phased array antennas in the vertical direction will permit even more dynamic scans and collaborations among the network radars.

To help with the design and operation of the first CASA radar test-bed to be deployed in southwestern Oklahoma, and to examine the potential impact of the data from the test-bed radars on storm-scale weather prediction through the assimilation of these data into the model initial conditions, a set of observing simulation system experiments (OSSEs) are conducted. A recently developed ensemble Kalman filter system is used for assimilating the data into a nonhydrostatic weather prediction model.

Since its first introduction by Evensen (1994), the ensemble Kalman filter (EnKF) technique for data assimilation has received much attention. In general, EnKF and related methods are designed to simplify and make possible the computation of flow-dependent error statistics. Rather than solving the equation for the time evolution of the probability density function of model state as in the traditional Kalman filter, EnKF methods apply the Monte Carlo method to estimate the forecast error statistics. A large ensemble of model states are integrated forward in time using the dynamic equations, the moments of the probability density function are then calculated from this ensemble for different times (Evensen 2003).

Very recently, EnKF has been applied to the assimilation of simulated Doppler radar data for modeled convective storms (Snyder and Zhang 2003; Zhang et al. 2004; Tong and Xue 2005, hereafter referred to as SZ03, ZSS04 and TX05, respectively) and of real radar data by Dowell et al. (2004). Very encouraging results are obtained in these studies in analyzing the state variables for convective storms, even though none of these state variables are directly observed by the radar. The first two studies assimilated only radial velocity data, while in Dowell et al. (2004), the use of reflectivity data is limited to the update of rainwater mixing ratio only. These studies all use the same anelastic cloud model with warm rain microphysics only.

In TX05, a general purpose compressible model is used that includes a multi-class ice microphysics parameterization. Different from SZ03 and ZSS04, both radial velocity and reflectivity data are assimilated. The study demonstrates the ability of EnKF in retrieving multiple microphysical species associated with a multi-class ice microphysics scheme, and in accurately retrieving the wind and thermodynamic variables as well. The relative impact of assimilating radial velocity and reflectivity data are also examined. In general, the assimilation system is able to establish the model storm not present in the initial guess extremely well after a number of assimilation cycles, and best results are obtained when both radial velocity and reflectivity data, including the reflectivity information outside precipitation regions, are used. This is so even though the observation operator of reflectivity is highly nonlinear. It is also shown in TX05 that dynamically consistent background error covariances develop in the system, especially in the later cycles, even in the unfavorable case in which only reflectivity information in the precipitation regions is assimilated. It is suggested that such flow-dependent background error covariances play a critical role in successful assimilation and retrieval.

The EnKF system of TX05 is used in this study with a number of differences and enhancements, for a set of OSSEs that assimilate simulated data from one WSR-88D radar located at Oklahoma City and/or the

* *Corresponding author address:*
Dr. Ming Xue, School of Meteorology, University of Oklahoma, SEC 1310, 100 E. Boyd, Norman OK 73019, mxue@ou.edu

planned Oklahoma test-bed CASA radars located near Chickasha and Lawton, about 80 to 100 km southwest of Oklahoma City. Details on the EnKF, the assumed configurations of CASA radars, the simulation of observations, and the design of OSSEs will be given in section 2. In section 3, the role of cross-covariances of the background errors between the observations and model state variables in the EnKF is examined by comparing the results of two experiments. In section 4, the results are reported on the OSSEs that examine the value added by one well-positioned (related to the storm location) CASA radar to the existing WSR-88D radar network, and the effectiveness in assimilating CASA radar data alone. The evaluation is performed for both quasi-stationary and fast-moving storm systems, and at different volume scan frequencies. Section 5 presents results of forecasts starting from selected EnKF analyses and section 6 gives a summary.

2. The OSSE System and EnKF Analysis Procedure

a. The prediction model and truth simulation

As in TX05, the May 20, 1977 Del City, Oklahoma supercell storm case (Ray et al. 1981) simulated by the ARPS model (Xue et al. 2000; 2001; 2003) is used in this study. The ARPS model is fully compressible and nonhydrostatic, and predicts 12 state variables, including three wind components u, v, w, potential temperature θ, pressure p, 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, plus the turbulence kinetic energy used by the 1.5-order subgrid-scale turbulence closure scheme. The microphysical processes are parameterized using the three-category ice scheme of Lin *et al.* (1983). More details on the model can be found in (Xue et al. 2000; 2001).

Different from TX05, a smaller horizontal grid spacing of 1.5 km is used in this study. Further, to better resolve the lower atmosphere, a vertically stretched grid with a minimum vertical resolution of 100 m is used, instead of a uniform 500m vertical resolution as used in TX05, SZ03 and ZSS04. The model domain is 16 km deep with 40 physical layers. Two domain sizes are used (see Fig. 1), the smaller one has 47×47 horizontal points, which, excluding two boundary points, is 66 km on each side. The larger domain, used by experiments with fast-moving storms and the corresponding cases of slow-moving ones, has 55×103 horizontal points and is 78 km by 150 km in physical size. The small domain is centered at 34.8º N and 98.1 º W and the large domain is centered at 34.75º N and 98.11º W and both use Lambert projection. The true latitudes of projection are 30 º and 60º N and the true longitude circle goes through the center of each grid.

The truth simulations or nature runs are initialized in the same way as in TX05. A sounding of 3300 J kg^{-1} CAPE (see Xue et al. 2001 for a skew-T plot of the sounding) is used to define the environmental condition and a 4 K ellipsoidal thermal bubble is used to initiate the storm. Three truth simulations were created; one for a slow-moving storm system with a quasi-stationary right-moving cell in the small domain (referred to as SMS), one for a fast-moving system in the large domain (FML), and one for a slow-moving system in the large domain (SML) for comparison with the latter.

The center location of the bubble is at x = 16 km, 0 km and 9 km, and y = -12 km, -62 km and -12 km, respectively for SMS, FML and SML. In the vertical, the bubble is centered at z = 1.5 km. The origin of horizontal coordinates is located at the domain center for both grids. The initial location of bubble for FML is chosen so that over most of the assimilation period, much of the storm system remains within the coverage of the four-radar CASA network (see Fig. 1). The choice for SML is such that at 60 min, the storm system is located at roughly the same location as that in FML so as to facilitate more direct comparisons.

Radiation condition is applied at the lateral boundaries and at the model top; the lower boundary is free slip. For the experiments with a slow-moving supercell system, a constant wind of u = 3 m s^{-1} and v = 14 m s^{-1} is subtracted from the original sounding to keep the right-moving cell near the center of model grid, as is done in TX05. For the experiments with a fast-moving storm system, the original wind profile is used.

During the truth simulation, the strength of the initial cell that develops out of the initial bubble increases quickly over the first 20 min then decreases over the next 30 min due to the splitting of the cell into two at around 55 min (see Fig. 5). The right moving (relative to the storm motion vector which is towards north-northeast) cell tends to dominate the system; the updraft reaches a peak value of 56 m s^{-1} at 101 min. The left moving cell starts to split again at 90 min. The initial cloud started to form at about 10 min, and rainwater formed shortly after. Significant ice phase fields appeared at about 15 min. The general evolution of the storm is similar to that documented in Xue et al (2001).

b. Simulation of radar observations

One WSR-88D and up to four CASA test-bed radars are involved in this OSSE study. For the WSR-88D radar, standard precipitation mode parameters are assumed, including 10 cm wavelength, 1º beam width, a total of 14 elevations with the lowest elevation at 0.5º and highest at 19.5º. The radial resolution is 250 m for radial velocity and 1 km for reflectivity (the radial resolution does not matter in this study because data are collected in Cartesian coordinates in the horizontal,

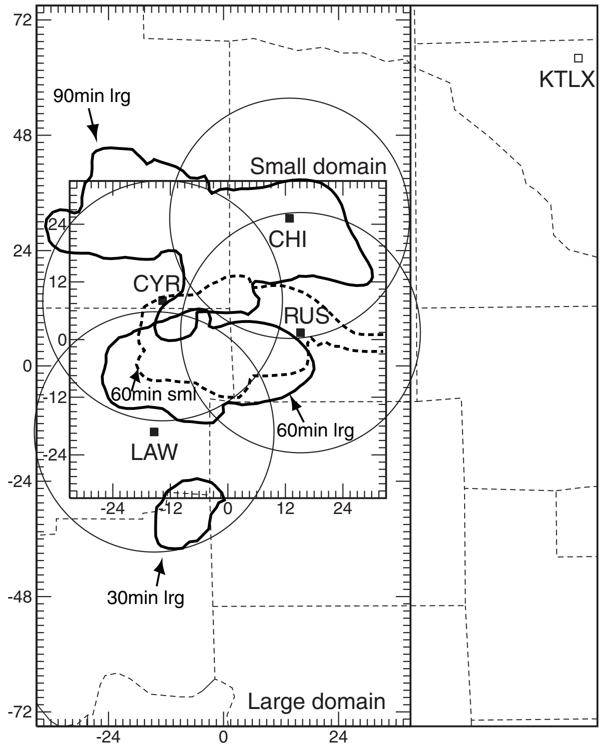

Fig. 1. A map with county borders showing the locations of first four planned CASA Oklahoma test bed radars, near Chickasha (CHI), Rush Springs (RUS), Lawton (LAW) and Cyris (CRY), together with the Twin Lake (KTLX) WSR-88D radar near Oklahoma City. The assumed 25 km maximum range of CASA radars are shown by the low-level 25 km range circles, and the two (large and small) analysis and forecast domains are shown as square and rectangular boxes with axis tick marks and labels. The origin of both domains is set at the center of the domain. Also plotted are the 10 dBZ low-level (50 m AGL) reflectivity contours of truth simulations FML (thick solid) and SMS (thick dashed), at the times labeled in the figure. The 10 dBZ contours for SML are similarly located as those of FML at 60 minutes, and are therefore not shown.

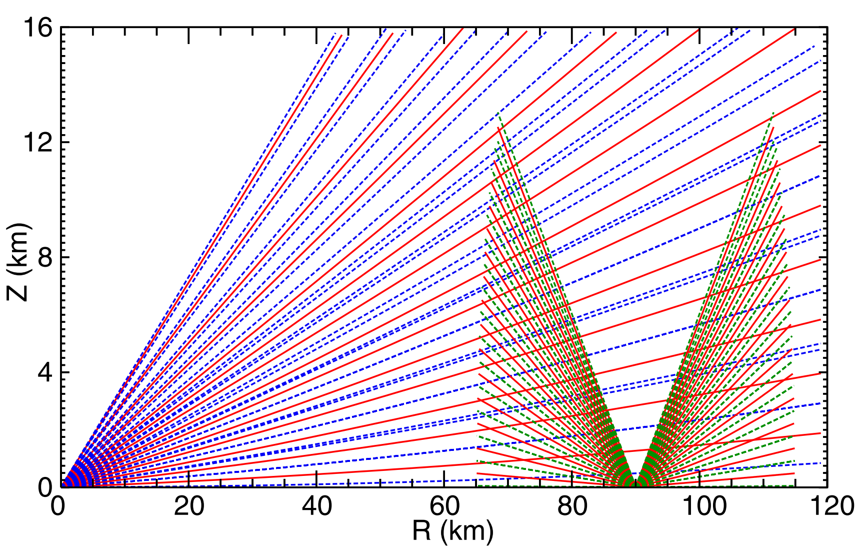

Fig. 2. The radial beams of a WSR-88D radar located at x=0 and of an assumed CASA radar located 90 km away. The red lines indicate the center of radar beams, the blue (green) dashed lines indicate the edge of these beams for the WSR-88D radar (CASA radar). The maximum range of the CASA radar is assumed to be 25 km.

see later); the maximum range is assumed to be 230 km (which is sufficient to cover the entire computational domain). For the first four CASA radars to be installed in the Oklahoma test-bed, whose primary goal is for the detection of low-level hazardous weather including tornadoes, and for improving weather prediction, the wavelength will be 3 cm (X-band) and beam width about 2º. The maximum range will be 25-30 km (we assume 25 km in this paper). Since the CASA radars will be dynamically configurable in real time in response to weather situation and user needs, their scanning strategies will remain flexible. For the purpose of this study, we assume a total of 15 elevations at 2º increment, with the elevation of the center of the lowest beam being at 1º. The impact of a variety of scanning strategies, including the vertical data coverage, on the quality of storm analysis will be the subject of a future paper.

The locations of the initial four Oklahoma test-bed radars to be installed as early as fall 2005 have been decided, and they are located near Chickasha, Rush Springs, Lawton and Cyril in Okalahoma, to the southwest of Twin Lake (KTLX) WSR-88D radar near Oklahoma City. The locations of the four CASA radars are plotted in Fig. 1 together with their assumed 25 km range circles at the low level.

Different from earlier OSSE studies of SZ03, ZSS04 and TX05, we assume that the simulated observations are available on the original radar elevation levels (or the data are on the radar plan position indicator (PPI)) rather than at the model grid points. We do assume that on each elevation level, radar observations are already interpolated from the radar polar coordinate to the Cartesian coordinate; in another word, the observations are found in the vertical columns co-located with the model scalar points. This assumption is reasonable since a horizontal interpolation to bring real radar data to the vertical columns is usually done before assimilating the data (e.g., in the 4DVAR work of Sun and Crook 2001). The main purpose of interpolation is to make the data distribution more uniform in the horizontal. Still, we plan to examine the effect of such horizontal interpolation by comparing with the analyses that use data in native radar coordinates.

The effects of the curvature of the earth and the beam bending due to vertical change of refractivity are taken into account by using the simply effective earth radius model discussed in Doviak and Zrnic (1993); it is mainly the earth curvature effect that prevents the radars from seeing the lowest atmosphere far away. The radar beams of one WSR-88D radar and one CASA radar located 90 km apart are illustrated in Fig. 2, together with the coverage by these beams up to their half-power width.

Since the observations are not at the model grid point, a forward observation operator is needed to bring the data from the model vertical levels to the radar elevation levels. This is achieved by using a simplified radar emulator that does power gain-based sampling only in the vertical direction:

$$\varphi_e = \frac{\sum G\varphi_g \Delta d}{\sum G\Delta d}\,, \tag{1}$$

where φ_e and φ_g are respectively the elevation level and grid point values of either radial velocity (V_r) or reflectivity (Z). Δd is the grid spacing in the direction perpendicular to the radar beam and d is the distance of grid point value from the center of radar beam. The power gain function G is assumed to be Gaussian and has the form of

$$G = \exp[-d^2/(2b^2)]\,, \tag{2}$$

where b is the beam half-width in terms of meters. For WSR-88D (CASA) radars, b corresponds to a 0.5 (1) degree elevation angle. This formulation is also used in Sun and Crook (2001) except for the approximation they make with the distance from the beam center.

For radial velocity, the grid point values involved in the numerator of Eq.(1) are first calculated from

$$V_{rg} = u\cos\alpha\sin\beta + v\cos\alpha\cos\beta + w\sin\alpha\,, \tag{3}$$

where α is the local elevation angle and β the azimuth angle of the radar beam that goes through the given grid point, and u, v and w are the model-simulated velocities interpolated to the scalar point of a staggered model grid. Subscript g of V_r denotes the grid point value. After V_r is sampled from the grid point values, random error drawn from a normal distribution with zero mean and standard deviation of 1 m s^{-1} are added as simulated observation errors. Since V_r is sampled directly from velocity fields, the effect of hydrometeor sedimentation is not involved in the data or assimilation.

The simulated reflectivity, Z, in dBZ, is calculated from the mixing ratios of rainwater, snow and hail hydrometeors, using the same formulations as in TX05. The formulations mostly follow those of Smith *et al* (1975) and are consistent with the ARPS ice microphysics. As with V_r, Z is first calculated at the grid points within the beam width then transferred to the beam elevations using Eq.(1). Random errors of zero mean and standard deviation of 5 dBZ are then added to the simulated reflectivity data.

When observation operator Eq.(1) is used to create simulated observations, no data is collected when no grid level is found within the beam width. This is equivalent to saying that data are discarded by the assimilation when the forward observation operator involves no grid levels within the beam width. For our vertically stretched analysis grids with high resolutions at the low levels, this does not happen often, when it does happen, the radar beam is usually very narrow therefore the atmosphere is already well sampled. Our procedure is therefore a natural way of thinning the data. For data sampling and for assimilation, we assume that the observation operators as well as the prediction model are perfect. This assumption is used in all earlier OSSE studies with radar data, and also in most atmospheric data assimilation systems. Model error will be an issue for future study.

c. The EnKF data assimilation procedure

Our EnKF system used in this study is similar to that described in Tong and Xue (2005) but with several differences. The first is that we now use the ensemble square root Kalman filter (EnSRF, Whitaker and Hamill 2002; Tippett et al. 2003) instead of the perturbed observation method (Evensen 1994; Burgers et al. 1998; Houtekamer and Mitchell 1998; Evensen 2003). Relatively small differences found between the two methods are reported in TX05 when 100 ensemble members are used. In this study, 40 members are used in all experiments; the EnSRF method is chosen because it generally performs better for limited ensemble sizes (Whitaker and Hamill 2002). With EnSRF, no additional perturbation is added to the observations except for the assumed observational errors added in the data sampling process.

As in TX05, we start the initial ensemble forecast at 20 min of model time when the first storm cell developing out of an initial bubble reaches peak intensity. The ensemble is initialized by adding random perturbations to a horizontally homogeneous ensemble mean defined by the environmental sounding and the random perturbations are sampled from Gaussian distributions with zero mean and standard deviation of 3 m s^{-1} for u, v, and w, 3 K for potential temperature θ and 0.5 g kg^{-1} for water vapor mixing ratio q_v. The pressure and microphysical variables are not perturbed. Similar to TX05, we apply the initial perturbations to the entire domain except for the outermost five rings of grid points near the lateral boundaries. We do not perturb u, v, θ, and q_v at the first grid level about ground either. SZ03 and Dowell et al (2004) found that limiting the initial perturbations to regions around the observed storms improves the analysis by avoiding the triggering by the perturbations spurious cells outside the regions of V_r data coverage (only V_r is assimilated in their studies). In our case, both V_r and Z data are assimilated and the assimilation of reflectivity information in non-precipitation regions helps suppress spurious cells. Our procedure allows for a more general application of the EnKF system, for, e.g., cases where data from platforms rather than radar are involved.

The observations are assimilated every 5 min and the first analysis is performed at 25 min, unless it is otherwise stated. Here, the CASA radars are assumed to be operating in the traditionally full volume scan mode and at a relatively low scan frequency of 5 min, but the impact of collecting and assimilating data at higher volume scan frequencies will also be examined.

As mentioned earlier, 40 ensemble members are used. The observation errors are assumed to be uncorrelated and the observations are analyzed sequentially one at a time (Houtekamer and Mitchell 2001). As in TX05, a covariance localization procedure following Houtekamer and Mitchell (2001) is employed that applies Schur product of the background error covariance calculated from the ensemble and a correlation function with local support. The correlation function follows Eq. (4.10) of Gaspari and Cohn (1999). An effective cutoff radius of 6 km is used in this study rather than 8 km as in TX05, which used 100 ensemble members and a 2 km horizontal resolution. The best choice of cutoff radius is found through numerical experimentation.

The covariance inflation procedure is necessary due to the typical underestimation of background error covariances from the limited-size ensemble. The procedure that we use is based on that of Anderson (2001) with an important modification; instead of applying everywhere, covariance inflation is limited to the grid points that are directly influenced during the analysis update by the observations found within the precipitation (where observed Z > 10 dBZ) regions. This modification is necessary to avoid amplifying spurious cells in precipitation-free regions. The inflation factor we use is 1.07 or the amount of inflation is 7 percent.

3. The Role of cross-covariances

In TX05, it is shown for essentially the worst case in which only reflectivity data in the precipitation regions are assimilated, that dynamically consistent background error covariances between the reflectivity in the updraft region and the model state variables in a significant portion of the domain can be obtained from the ensemble forecast during the later assimilation cycles. It is suggested that such flow-dependent cross (across different variables) covariance information is very important for a successful analysis of those state variables that are not directly observed and often not even related to the observed quantities (i.e., V_r and Z) through the observation operators. For these variables, analysis update based on cross covariances of background error and the adjustment through model dynamics during the prediction are the only ways by which they can be 'analyzed'. However, it is also shown that during the early cycles that follow the initialization of the ensemble from relatively poor initial guess of ensemble mean, the analysis update of variables not directly related to reflectivity (via the observation operator) actually hurts the analysis (i.e., the analysis increases rather than decreases the background error). It was found that best results are obtained when such updating is not done to indirectly related variables when assimilating Z during the first four cycles.

Before discussing the experiments on the impact of CASA radar data, we further examine in this section the role of cross covariance information and the related analysis update in assimilating radar data. Two experiments are performed, using only one WSR-88D radar. The first experiment, which we refer to as TLX (see Table 1), follows the standard procedure that assimilates V_r data in the precipitation region and Z data in all regions within the radar coverage. The procedure updates only directly related state variables (i.e., q_r, q_s and q_h) when assimilating Z data during the first 20 min (before the 4th analysis). All state variables are updated afterwards. In the second experiment (TLXNCov), throughout the assimilation, V_r is used to update only the wind components and Z is used to update q_r, q_s and q_h only. The indirectly related variables are not updated and have to rely on the model dynamics to adjust to the directly analyzed variables.

Fig. 3 shows the root-mean-square (*rms*) error curves of ensemble mean forecasts and analyses from these two experiments (together with those from TLXCYR to be discussed later) plotted against time or analysis cycle. The errors are obtained by averaging over those grid points where observed Z exceeds 10 dBZ. Despite a number of configuration differences, the behaviors of error reduction in TLX are very similar to those of experiment VrPZF reported in TX05, which used V_r and Z data in the same way. Errors in all fields are reduced rapidly up to about 55 min when they become stabilized for the next 30 min. Between 60 and 80 min, the *rms* errors in the wind components are about 1 ms^{-1} and the temperature error is less than 1 K. Such velocity errors are similar to the observational errors added to the radial velocity. This means that after eight to ten assimilation cycles, the EnKF system is producing a very good estimate of the state of simulated storm. The errors in some of the fields increase slightly after 80 min and such increase is apparently due to faster forecast error growth in the later stage of the storm system development when the left mover splits again into several smaller less organized updraft cores (see Fig. 5). The spreading of low-level cold pool not well captured by KTLX radar also contributes to the error increase. A similar behavior is also observed in TX05.

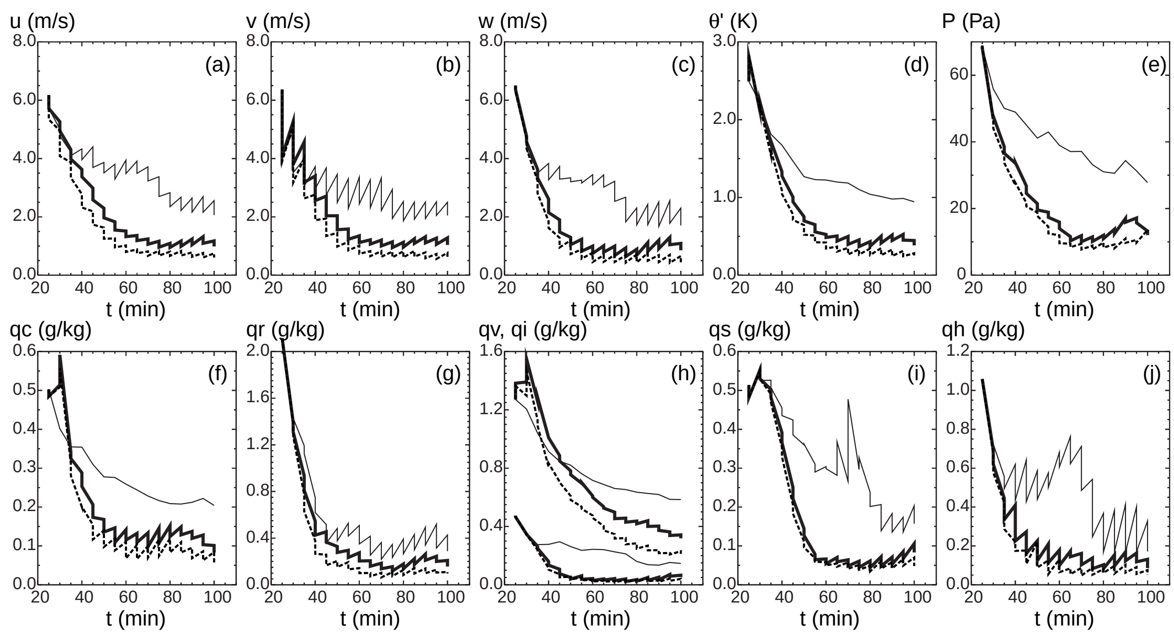

Fig. 3. The *rms* errors of ensemble-mean forecast and analysis, averaged over points at which the reflectivity is greater than 10dBZ for: (a) u, (b) v, (c) w, (d) θ', (e) p', (f) q_c, (g) q_r, (h) q_v (the curves with larger values) and q_i (the curves with lower values), (i) q_s and (j) q_h, for experiments TLX (thick black), TLXCYR (dashed) and TLXNCov (thin black). Units are shown in the plots. The drop of the error curves at specific times corresponds to the reduction of error by analysis.

Table 1 List of Assimilation Experiments

Experiment	Radars assimilated and configuration
TLX	TLX data only, standard assimilation procedure
TLXNCov	As TLX, but cross covariance information is not used for updating indirectly related state variables
TLXCYR	KTLX and Cyril radars
CASA1S	One CASA radar at Rush Springs in small domain
CASA3S	Three CASA radars at Cyril, Rush Spring and Lawton in small domain
CASA1L	Single Lawton radar in large domain
CASA4L	All 4 CASA radars in large domain
CASA1LM	As CASA1L, but with fast moving storm system
CASA4LM	As CASA4L, but with fast moving storm system
CASA1LF1/2	As CASA1L but with 1 and 2.5 min VSF and assimilation cycles, respectively
CASA1LMF1/2	As CASA1LM but with 1 or 2.5 min VSF and assimilation cycles
CASA4LF1/2	As CASA4L but with 1 or 2.5 min VSF and assimilation cycles
CASA4LMF1/2	As CASA4LM but with 1 or 2.5 min VSF and assimilation cycles

Table 2. Number of radar observations for radial velocity and reflectivity for each experiment at 30, 60 and 90 minutes

Experiment	30 min		60 min		90 min	
	Obs number		Obs number		Obs number	
	V_r	Z	V_r	Z	V_r	Z
TLX, TLXNCov	1458	17624	4199	17624	6616	17624
TLXCYR	2071	28777	7173	28777	12314	28777
CASA1S	1675	10834	6685	10834	8029	10834
CASA3S	2691	30960	10794	30960	14767	30960
CASA1L, CASA1LF1/2	1222	11730	2758	11730	2651	11730
CASA4L, CASA4LF1/2	4181	46855	14763	46855	18728	46855
CASA1LM, CASA1LMF1/2	1208	11730	2836	11730	348	11730
CASA4LM, CASA4LMF1/2	1963	46855	13674	46855	17420	46855

The error reduction in u, v, w, q_r and q_h in the first three assimilation cycles (up to 35 min) is very similar between TLX and TLXNCov (Fig. 3). The errors in θ', q_c and q_v are actually larger in TLX than in TLXNCov; the model state errors in these three variables are actually increased by the analysis during the first two cycles in TLX. Since during the first four cycles, only V_r data are used to update these variables, the cross-covariances between V_r and these indirectly related variables must have been unreliable, causing degradation of their state by the analysis. The same behavior was observed in TX05 when updating indirectly related variables using Z data in the first few cycles.

Despite the larger errors in the initial cycles, the errors in θ', q_c and q_v are quickly reduced in TLX to below those of TLXNCov starting around 35 min (Fig. 3d, Fig. 3f, Fig. 3h) and become significantly smaller in the later cycles, in a way similar to all other variables. For the velocity components, the *rms* error difference between the two experiments is as large as 2.5 ms^{-1} (3.5 v.s. 1 ms^{-1}, Fig. 3a and Fig. 3c), and at the end of the assimilation (100 min), the analysis error in the wind components is about twice as large in TLXNCov (2 v.s. 1 ms^{-1}, Fig. 3a-c). More disturbingly, the analyses of q_s and q_h that are directly related to the reflectivity observations are significantly degraded during the intermediate cycles. In fact, the *rms* error in q_s is increased by as much as 0.2 g kg^{-1} by the analysis at 70 min (Fig. 3i). This suggests that when the overall analysis is poor, even the covariance between the observation (Z is this case) and the directly related (through the observation operator) variable (q_s) can be bad, causing increase of error in the field by analysis

The plots of horizontal winds and w at the 6 km level from the truth, TLX and TLXNCov can be found in Fig. 5 for various analysis times; the corresponding surface fields are found in Fig. 6. We can see that at the 6 km level, the flow and storm updrafts are very poorly analyzed before 100 min (Fig. 5) in TLXNCov; in fact, the updrafts are almost completely missing at 60 min (Fig. 5j), resulting an *rms* error of over 3 ms^{-1} in w at this time (Fig. 3c). Fig. 6 shows that the analysis of low-level cold pool and precipitation patterns are very poor at 40 min in TLXNCov (Fig. 6i) while those of TLX (Fig. 6e) are significantly better. The agreement with the truth in these patterns, though better than at the higher levels, remains poorer than TLX case throughout the analysis period with TLXNCov, consistent with that we saw in the *rms* errors earlier.

Overall, it is clear that the analysis obtained in TLXNCov, in which cross-covariance information between the observations and indirectly related variables are not used for analysis update, is much poorer. We can therefore convincingly say that the cross-covariances among different variables are very important in producing accurate analysis of a convective storm using radar observations. We rely on such information in retrieving the variables that not directly observed.

Meanwhile, we also want to point out that the model dynamics also play an important role in the 'retrieval' process. This is indicated by the fact that in TLX, for most variables (all except for v and q_c and q_v in the first two cycles), the model forecast between analyses actually reduces the state error, sometimes significantly, during the earlier cycles when overall error is rapidly reduced. The y component of wind, v, and the hail mixing ratio q_h are the two variables whose errors tend to increase quickly during the forecast period while their analyses reduce the forecast errors most.

Apparently, what is happening is that because a significant component of wind in the y direction is directly observed by the radar (located to the north northwest of the storm at the early stage) and there is a large sensitivity of reflectivity to the hail concentration, v and q_h get corrected directly by the observations even before reliable background covariance structure is developed and when the overall state of the atmosphere is poorly estimated. In fact, the corrections to them tend to be 'overdone' relative to the state of the other variables, so that once set free (during the forecast), their errors grow back quickly. This happens when they try to restore back towards the state before the analysis so as to be more consistent with the other variables, which have been adjusted much less by the analysis. The other variables, on the other hand, are also adjusted during the forecast so as to be more consistent with v and q_h; in the process, their errors are reduced. For these reasons, we say that the model dynamics is also very important in producing a dynamically consistent estimate of the state of the atmosphere, in particular, of convective storms.

4. Assimilation of CASA Radar Data and Its Impact

a. Impact of single CASA radar in addition to one WSR-88D radar

As stated in the introduction, one of the key problems with the existing national network of WSR-88D radars is that the typical radar spacings of a few hundred kilometers preclude the observations of the lowest kilometers of the atmosphere at a distance from the radar due to the earth curvature effect and non-zero elevation of the lowest tilt. Important storm-scale features such as tornado, cold pool, gust front and downbursts are missed in such cases. It is well established that low-level cold pool is very important for the support and maintenance of convective systems. In this section, we compare the analysis and forecast TLXCYR, in which the data from a radar at Cyril site are add to the data from KTLX radar. The experiment is otherwise the same as TLX.

As can be seen from Fig. 2, at a 90 km range, the center of the lowest radar beam is over 1 km above the ground, and the lower edge of the 1 degree wide beam is about 500 m above ground, implying that the atmosphere blow 500 m is not illuminated by the radar beam hence not observed. The addition of CASA radars fills such gaps while at the same time increases the resolution of observations in the covered region. In certain regions, the CASA radars may even provide dual or multiple Doppler wind coverage; in our case, such coverage is very limited partly because V_r observations are only available in the precipitation regions (see Fig. 1). The total numbers of V_r and Z observations available and assimilated in each experiment are given in Table 2.

The *rms* error curves for the state variables of experiment TLXCYR are plotted as dashed lines in Fig. 3. It is immediately clear by comparing the curves with those of TLX that the additional CASA radar provides a consistently positive impact, on essentially all variables. For one thing, the slight error increase in the later cycles in TLX is essentially gone for most variables and this is believed to be due to a good low-level coverage of the cold pool associated with the left moving cells as well as a generally higher-resolution coverage provided by the Cyril radar (see Fig. 6). The decrease in the *rms* errors in the wind components is almost 0.5 ms^{-1}, making the analysis errors in the late cycles significantly lower than the observational error in V_r. Improvements in the analyses of other variables are also clearly evident from Fig. 3.

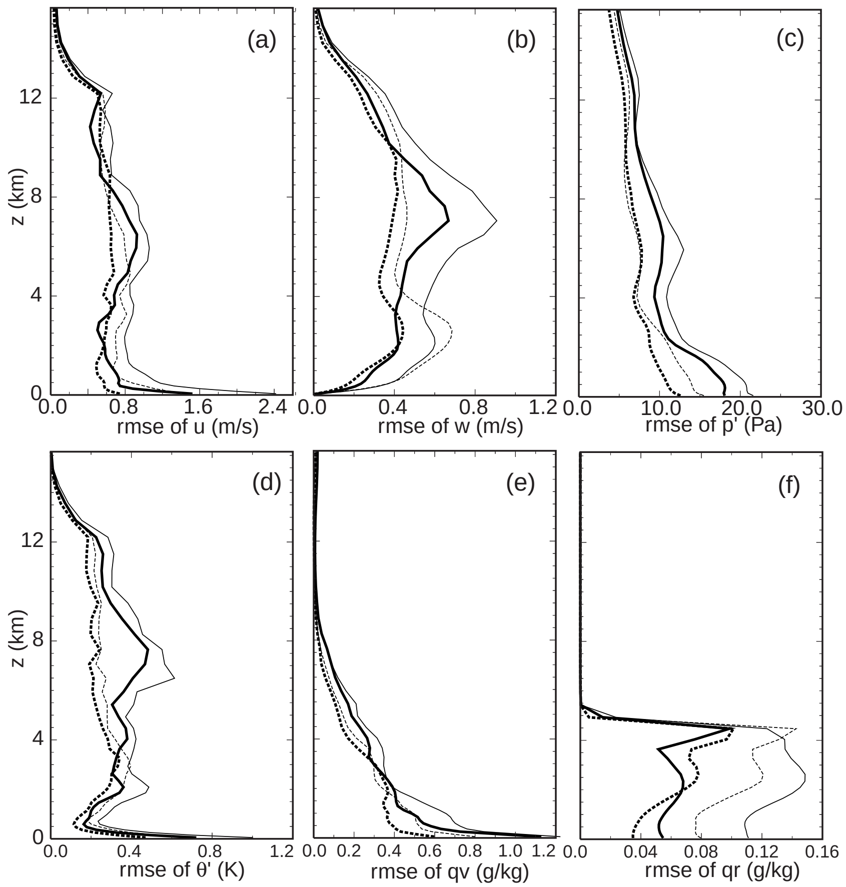

Fig. 4. Vertical profiles of *rms* errors of EnKF analyses (averaged over the entire horizontal domain) from experiments TLX (thin lines) and TLXCYR (thick lines) at 60 min (solid) and 80 min (dashed) for variables (a) u, (b) (w), (c) p', (d) θ', (e) q_v and (f) q_r.

The improvement as a function of height is revealed by the vertical profiles of *rms* errors. Fig. 4 shows that the largest differences in the *rms* errors between TLX and TLXCYR are found at the low levels. For u, the difference at the surface is about 0.9 ms^{-1} at 60 min and about 0.75 ms^{-1} at 80 min. For θ', the surface *rms* error differences are about 0.3 K and 0.23 K at 60 and 80 min, respectively. The *rms* error in q_r is about twice as large in TLX at the surface. The larger errors in u, θ' and q_r in TLX is a reflection of the poorer analysis of the low-level cold pool which is further driven by the poorer precipitation analysis (see Fig. 6 later). The error difference is generally larger at the earlier time (e.g., at 60 min); the difference decreases with the cycles as the analysis of TLX is also improved with the assimilation of more data and the buildup of the storm.

The 6-km level wind and w analyses from TLX (2nd row) and TLXCYR (4th row) during the assimilation period are plotted in Fig. 5, together the corresponding truth (first row). At this level, where the KTLX radar provides a rather good data coverage (see Table 2), the differences between the two experiments are relatively small though still identifiable, with the w fields from TLXCYR generally agreeing better with the truth. The differences at the low-levels are much larger. Fig. 6 shows the low-level flow and the simulated reflectivity, together with the cold pool as revealed by the negative θ perturbations. It is clear that the analyzed cold pool and the precipitation region are too small in area coverage in both experiments and are even more so in TLX at the earlier times, at, e.g., 40 min (Fig. 6e and Fig. 6m). The cold pool expands with time in both cases, but that in TLX never reaches the south boundary of the plotted domain by 100 min as the real one and that one in TLXCYR do. In fact, by 100 min, the structure and location of the cold pool boundary or gust front in TLXCYR agrees very well with the truth (Fig. 6p and Fig. 6d). The agreement in the reflectivity with the truth is very good in both cases.

To further examine the impact of Cyril radar data, we plot in Fig. 7 low-level vertical cross sections at 40 and 60 min through the cold pool along the lines indicated in Fig. 6. The cold pool is indicated by the contours of negative θ' which are plotted together with the winds projected to the plane of cross section. It can be seen by 40 min or after 4 analysis cycles, a updraft above 1 km level is reasonably well captured in both cases, but the low-level winds differ significantly from the truth due to an almost complete lack of cold pool in TLX and a much weaker but clearly identifiable one in TLXCYR. Another 4 analysis cycles later, at 60 min, the cold pool structure becomes much closer to the truth, but that in TLX remains noticeably narrower than the truth (25 km v.s. the true 30 km when the gust fronts are defined by -1 K θ' contours) while the gust front locations and cold pool width in TLXCYR are almost identical to the true ones (Fig. 7b and Fig. 7d). The gust front strength measured in terms of the horizontal θ' gradient is slightly weaker than the truth near the B' end of the cross section, which is part of rear flank gust front of the main cell. The depth of the cold pool is similar to that of the truth in both TLX and TLXCYR. The above results clearly indicate that the additional data from the Cyril radar are very helpful, especially during the earlier cycles, in establishing accurate low-level precipitation and cold pool structures as well as that of the associated winds in the model.

b. Assimilation of single or multiple CASA radars alone

In this section, we examine the ability of CASA radar(s) alone in producing a good analysis of the supercell storm system. Results from experiments CASA1S and CASA3S will be shown (Table 1). The analysis grid and procedure are the same as in TLX and TLXCYR except for the radar(s) used. In CASA1S, data from single radar at Rush Springs are used (see Fig. 1) while in CASA3S, data from three radars, at Cyril, Rush Springs and Lawton are used. The trailing 'S' in the names denotes the small domain used.

Fig. 8 shows that, when three CASA radars are used that provide a good spatial coverage of the storm system (see Fig. 1, Fig. 6 and Table 2), the quality of analysis is close to that of TLXCYR; in fact, for most variables, the error curves are between those of TLX and TLXCYR (c.f., Fig. 3), with the wind analysis errors at around or being lower than 1 ms^{-1} after 60 min. Similar conclusion can be drawn from the 6 km and surface plots shown in Fig. 5 and Fig. 6. The analysis of CASA3S can therefore be considered very good.

The errors of CASA1S are, however, consistently larger at all times (Fig. 8), and the errors start to increase significantly after 60 min, reaching over 2 ms^{-1} in the winds. This increase is clearly due to the lack of spatial coverage of the left-moving cells starting from 60 min (Fig. 6r - Fig. 6t), by the single Rush Springs radar. The lack of coverage in the western portion of the analysis domain is also responsible for the inability of the analysis to suppress spurious precipitation persistent in this part of domain (Fig. 6r - Fig. 6t).

The updraft core and the main precipitation regions of the right-moving cell did remain within the range of Rush Springs radar. By 100 min, the low-level flow within the radar range (Fig. 6t) and the mid-level updraft core and horizontal flow of the right mover (Fig. 5t) are well captured, but the updrafts of the left-moving cells are poorly analyzed (Fig. 5t); the low-level cold pool extends too far north, partly due to the merger with earlier spurious precipitation in the region (Fig. 7t).

The results show that when three CASA radars work together to provide a complete coverage of the storm system during the assimilation period, the EnKF analysis is almost as good as that from one well-positioned CASA radar plus one WSR-88D radar. When only one CASA radar is available and when a portion of the storm system is not covered by the radar, the quality of analysis deteriorates significantly. Spurious precipitation developed in part of the domain that could not be corrected by the analysis due to the lack of observations there.

c. Effect of storm motion

In all of our OSSEs reported so far, and in those of TX05 and ZSS04, a mean storm motion speed is first subtracted from the environmental sounding to keep the main storm cell quasi-stationary relative to the forecast and analysis grid. Doing so effectively reduces the local time tendency of model state and may have helped improving the quality of analysis. The use of a moving reference frame that follows the storm system is known to improve single Doppler wind analysis (Gal-Chen 1982; Zhang and Gal-Chen 1996; Liou and Luo 2001); traditional techniques that retrieve thermodynamic fields from the Doppler wind analyses (Gal-Chen 1978; Gal-Chen and Kropfli 1984) are also sensitive to the accuracy of time tendency estimate (Sun and Crook 1996). For general NWP applications, a moving reference frame is not easy, if at all possible, to implement.

In this section, we examine the effect of storm motion on the quality of EnKF analysis, by comparing experiments with and without subtracting the storm motion from the sounding. A larger grid, as shown in Fig. 1, is used to contain within the domain the fast moving storms for the entire period of analysis. The truth simulations used for the slow and fast-moving experiments are SML and FML, respectively.

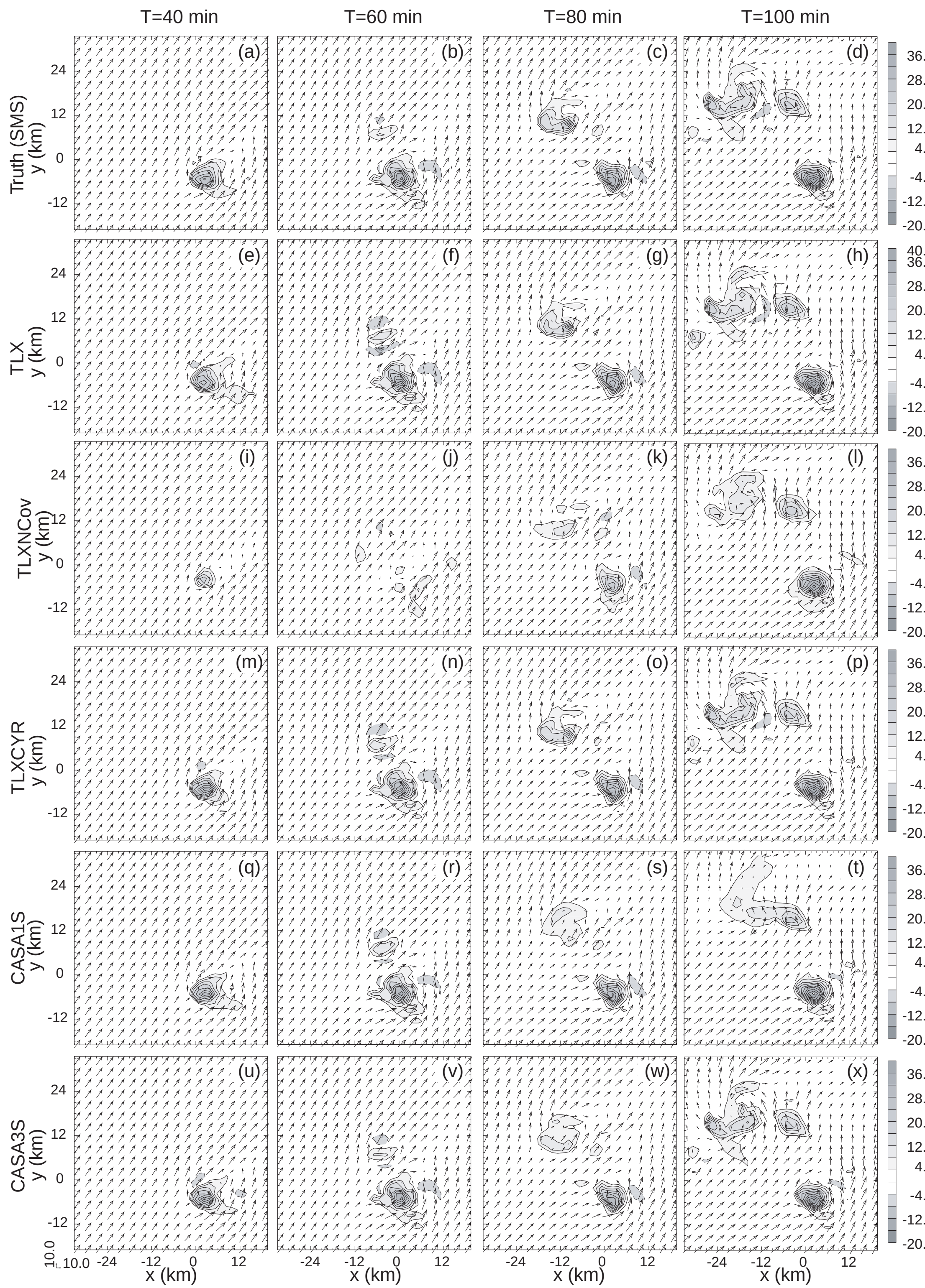

Fig. 5. Vertical velocity w (contours with shading at intervals of 4 m s^{-1}, negative contours are dashed) and horizontal wind (vectors, plotted every other grid point; ms^{-1}) in a subdomain, at z = 6 km: for truth simulation in the small domain (SMS) (a)-(d); analyses from experiments TLX (e)-(h), TLXNCov (i)-(l), TLXCYR (m)-(p), CASA1S (q)-(t), and CASA3S (u)-(x) at t = 40, 60, 80 and 100 min during the assimilation period.

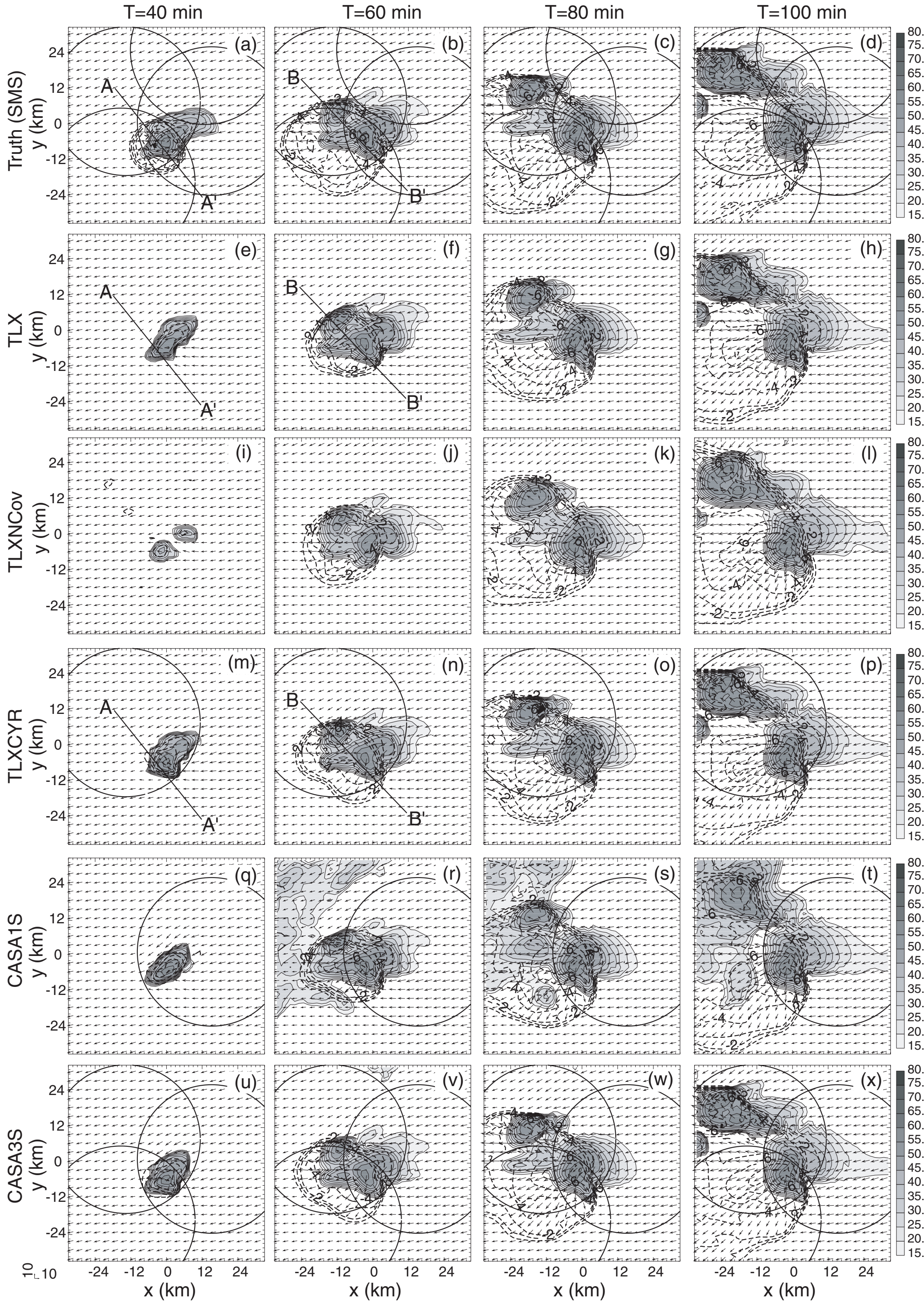

Fig. 6. As Fig. 5 but for horizontal winds (vectors; m s^{-1}), perturbation potential temperature (thick dashed lines at 1 K intervals) and simulated reflectivity (thin solid contours with shading at intervals of 5 dBZ, starting from 15 dBZ) at z = 50 m AGL. Lines A-A' and B-B' in the plots indicate the locations of vertical cross sections to be shown in Fig. 7.

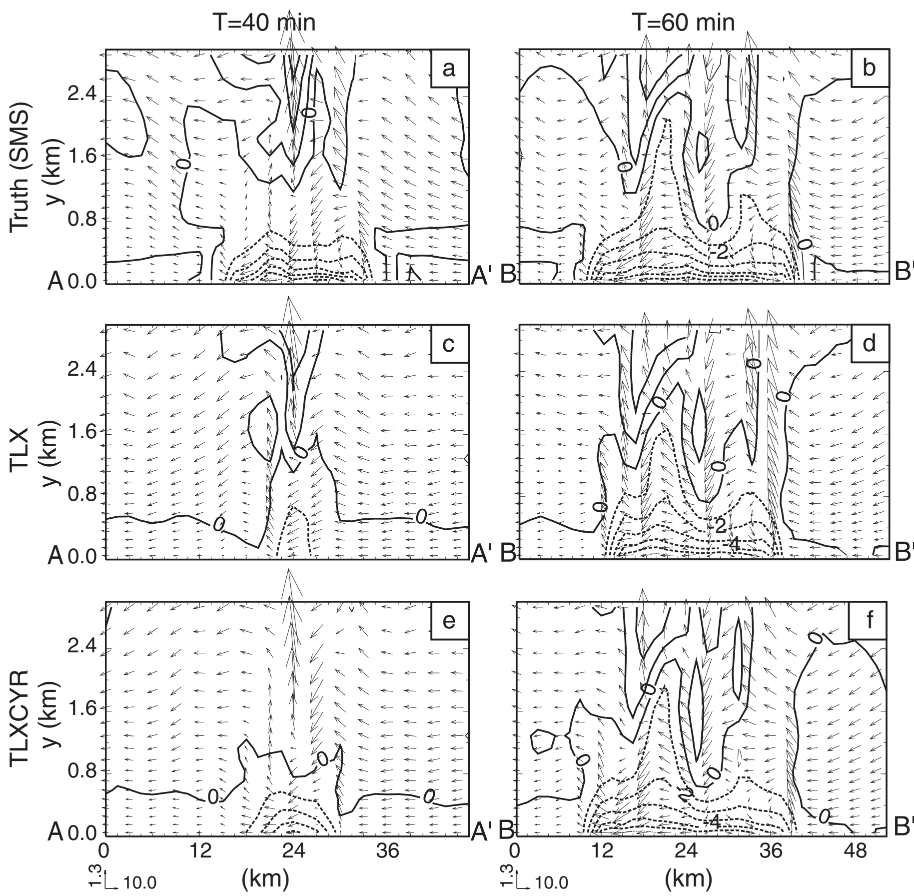

Fig. 7. Vertical cross-sections along lines A-A' (left panel) and B-B' (right panel) in Fig. 6 showing the analysis perturbation potential temperature (θ') contours at 1 K intervals, and the wind vectors projected to the cross section, for truth simulation SMS (a)-(b), experiments TLX (c)-(d), and TLCYR (e)-(f), at 40 (left panel) and 60 min (right panel) during the assimilation period.

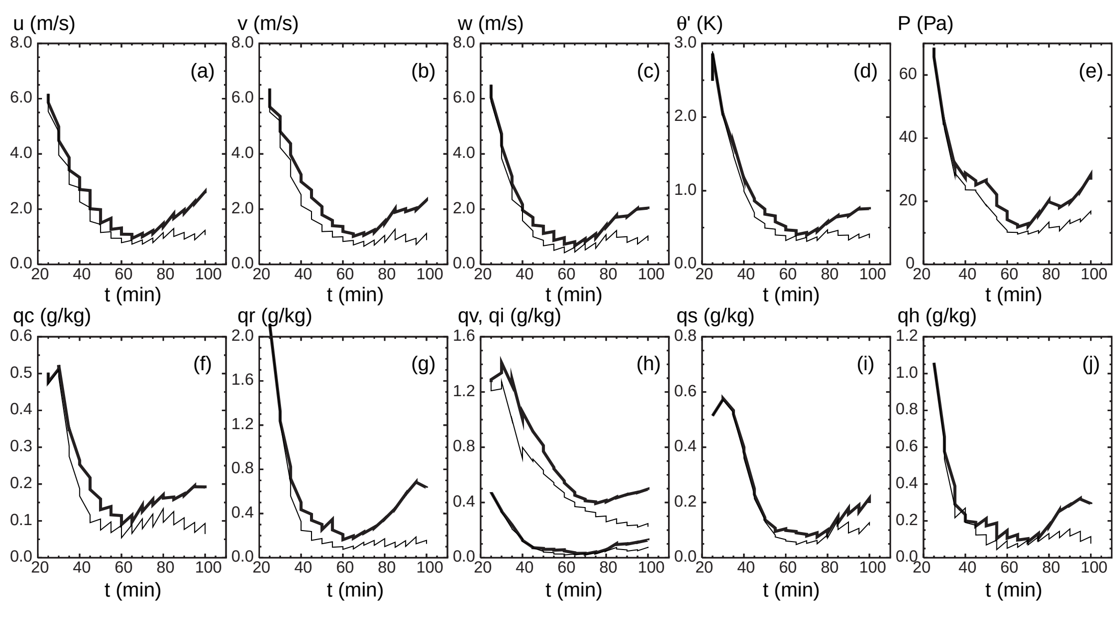

Fig. 8. As Fig. 3 but for experiments CASA1S (thick lines) and CASA3S (thin lines).

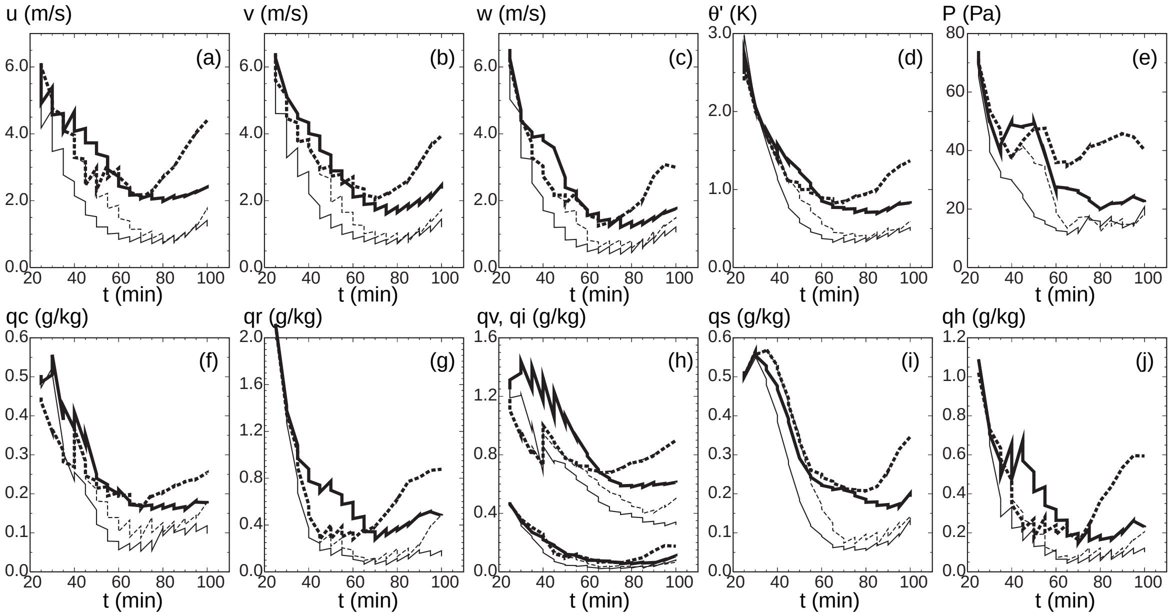

Fig. 9. As Fig. 3 but for experiments CASA1L (thick solid), CASA4L (thin solid), CASA1LM (thick dashed) and CASA4LM (thin dashed).

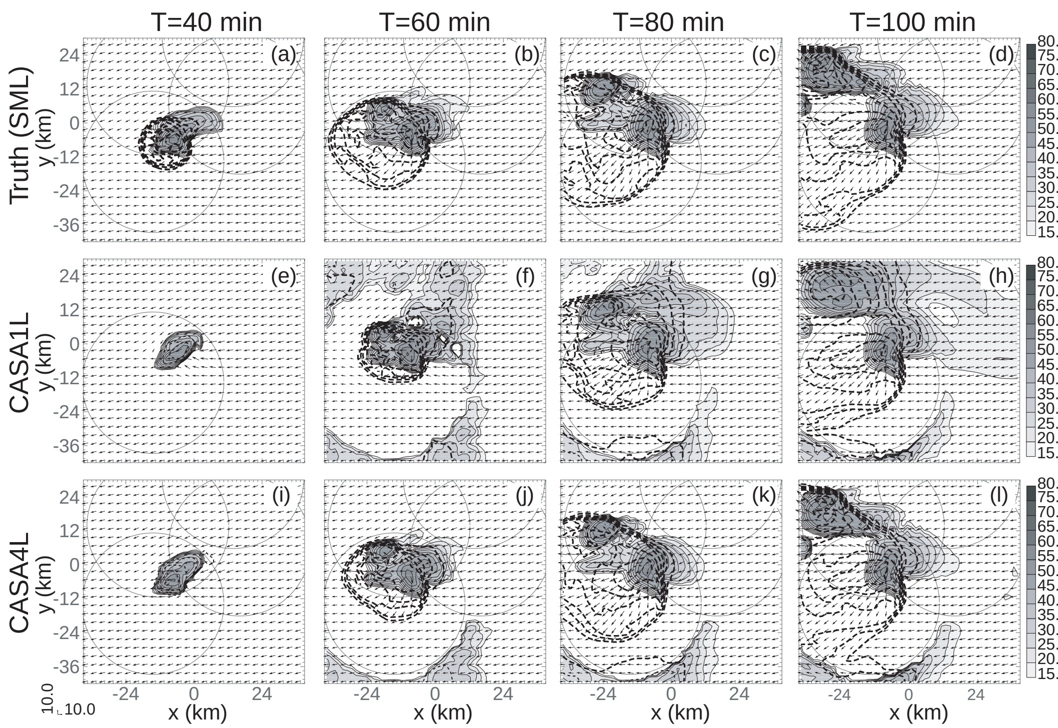

Fig. 10. Horizontal wind (vectors; m s^{-1}), θ' (thick dashed contours at 1 K intervals) and simulated reflectivity (thin solid contours with shading at intervals of 5 dBZ, starting from 15 dBZ) at z = 50 m AGL: for truth simulation SML (a)-(d); analyses from experiments CASA1L (e)-(h), and CASA4L (i)-(l), at t = 40, 60, 80 and 100 min during the assimilation period.

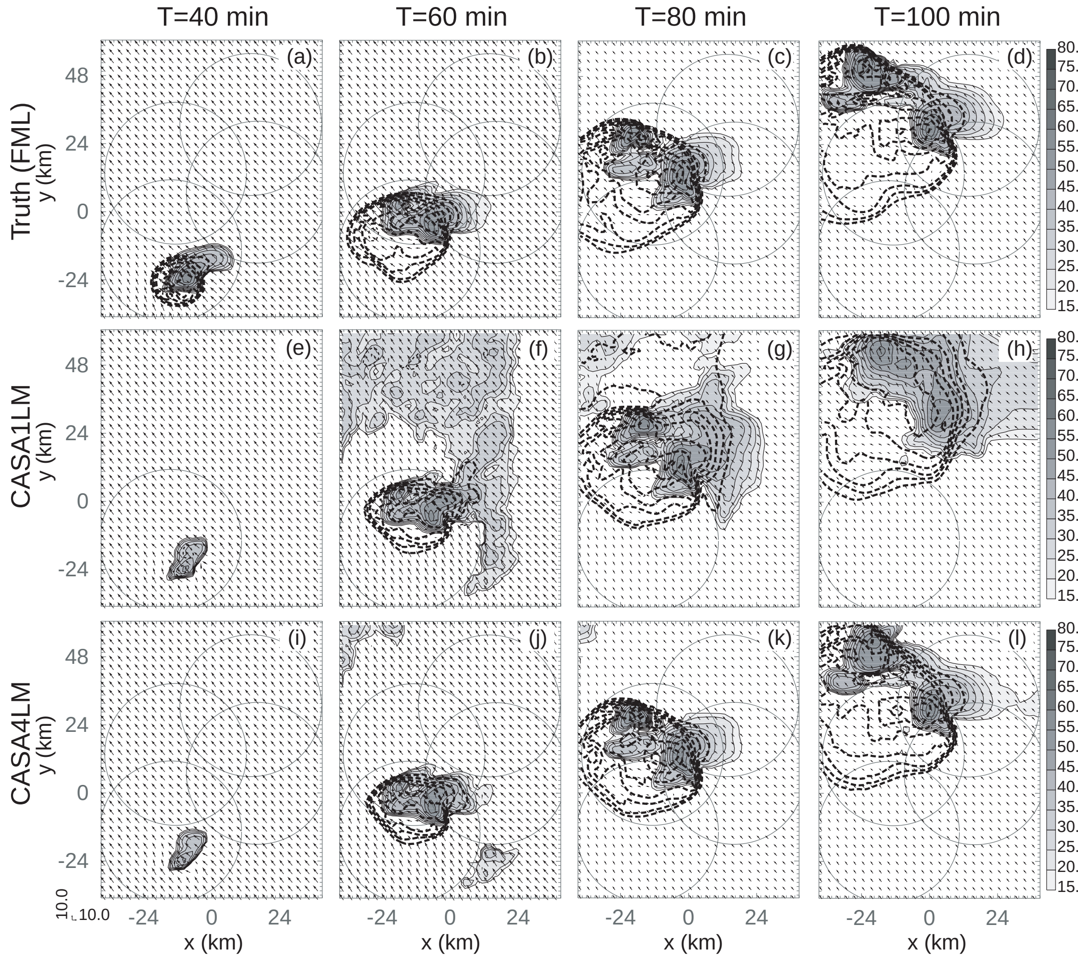

Fig. 11. As Fig. 10 but for truth simulation FML (a)-(d), and analyses from experiments CASA1LM (e)-(h), and CASA4LM (i)-(l).

Fig. 9 shows the forecast and analysis errors in the model state variables during the assimilation period for experiments CASA1L, CASA4L, CASA1LM and CASA4LM. The former two contain a slow moving storm system while the latter two contain a fast moving one. In CASA1L and CASA1LM, single CASA radar located at Lawton is assimilated while in CASA4L and CASA4LM, all four CASA radars are assimilated. Lawton radar provides the best coverage of the fast moving storm system during the early cycles. All four experiments use the same large domain.

It can be seen from the *rms* error plots (Fig. 9) that the analyses using four radars are consistently better than those of one-radar cases. Between 60 and 70 min where the errors are generally the smallest, the difference in the wind analysis errors is 1 ms^{-1} or more, and the difference in θ errors is larger than 0.5 K. The case with the largest errors during the later cycles is CASA1LM. In this case, the errors in all variables increase rapidly after 70 min from a level that is close to that of CASA1L at 70 min, when the precipitation regions of both left and right movers propagate out of the range of single Lawton radar (see Fig. 11). Before this time, spurious precipitation also exists in a significant portion of model domain (Fig. 11f) that is not corrected. As in some previous cases, such spurious precipitation is mainly caused by the initial random perturbations used to start the initial ensemble. Since the *rms* errors shown in Fig. 9 are calculated over the regions where observed *Z* exceeds 10 dBZ, most of the errors due to the spurious precipitation are not even reflected in the error plots.

The case with a slow moving storm system and 4 radars (CASA4L) produces the best analysis (Fig. 9). The errors of CASA4LM are close to though slightly larger in general than those of CASA4L, until 90 min. After this time, the errors also start to increase rapidly as the left mover propagates out of the range of all four radars (Fig. 11l). The analysis of the right mover remains rather good (Fig. 11l).

The 6-km level and the surface analysis fields from the four experiments are shown in Fig. 10 and Fig.

11, and some of them have already been referenced earlier. These fields help us understand the error evolutions shown in Fig. 9. In general, the data coverage appears to be the most significant factor that affects the quality of storm analysis. Once a storm or a portion of it moves out of the range of radar network, the model state error growth can no longer be controlled and the analysis deteriorates. The lack of data coverage in the entire analysis domain also negatively impacts the overall analysis as some spurious cells can be not suppressed. When the data coverage is similar, the analysis of a slow-moving storm system is slightly better than that of a fast moving one.

d. Impact of volume scan frequency

Another factor that can affect the accuracy of local time tendency estimate is the radar volume scan frequency (VSF). Faster scan tends to give a better estimate. Furthermore, freezing turbulence assumption made in certain single-Doppler velocity retrieval (SDVR) techniques becomes more valid between two scans of short time interval. The mean winds determined from successive volume scans based on the principle of tracking quasi-conserved quantities (e.g, Qiu and Xu 1992; Shapiro et al. 1995) are definitely more accurate using high-frequency data. Shapiro et al (2003) find that with a SDVR scheme based on a Lagrangian form of the radial component of the equation of motion, the wind retrieval error statistics are substantially improved as the volume scan intervals decreases from 5 min down to 1 min, using real Doppler-on-Wheels mobile radar data. ZSS04 find, however, with their EnKF system, that the analysis is only marginally better during the first few assimilation cycles when 2-min instead of 5-min volume scan data are assimilated, and the difference becomes minimal during the later cycles. The main storm in their case was quasi-stationary.

Since the CASA radars will be designed to operate with a variety of scan strategies that would respond in real time to user needs. When necessary, the radar can perform sector or even spatially or temporally interleaved scans at short time intervals. It is important to better understand the impact of scan frequency on the quality of thunderstorm analysis, so as to help design the control system of the network and to optimize the system operations. In this section, we attempt to answer some of the questions by comparing analyses from the WSR-88D-standard 5 min-interval data and data collected at 1 and 2.5 min intervals. To be fair, the analysis starts at 25 min in all experiments. Further, we will examine the impact for both slow-moving and fast moving storms. As in the previous subsection, truth simulations SML and FML are used.

The four large-domain experiments in the previous subsection are repeated assimilating 1 min and 2.5 min volume scan data instead. The general conclusion is that the faster volume scan does improve the quality of analysis, especially during the earlier cycles, but the sensitivity of the analysis to VSF is much smaller than the afore-mentioned retrieval techniques are, and the sensitivity decreases as the length of assimilation period increases but increases as the storms move faster. When the spatial coverage of the radar network is small, high VSF helps because the storm structure can be established quickly before the storm moves out of the range of the network. Actual results of these experiments will be presented at the conference.

5. Forecasts from ensemble-mean analyses

Since the goal of data assimilation is to provide a good initial condition for numerical weather prediction, in this section, we look at the quality of forecasts produced from the analyses. Fig. 12 shows the *rms* errors (averaged over the entire domain) of forecasts beginning from the ensemble-mean analyses of different times from experiments TLX, TLXCRY and CASA3S.

In general, more accurate estimate of the initial condition yields better forecast, but, because the differences in the analysis errors among the above three cases are relatively small, the forecast errors approach very similar values after 40 to 80 min, depending on the variable and the start time. The difference in errors is maintained for the longest for forecasts starting from 60 min. For example, the forecast error in u remains lower in TLXCRY than in TLX up to 165 min, or for over 100 min from the analysis time. The corresponding errors for w, θ' and q_r become indistinguishable by 140 min. The initial analysis errors of CASA3S are similar to those of TLXCYR, and its forecast errors generally oscillate between those of TLX and TLXCYR.

The error growth in w and q_r is fastest in the first 40 min or so; after that, the errors in w and q_r appear to reach saturation. The fast initial error growth in w and q_r is believed to be due to the fact that w and q_r represent smaller thunderstorm-scale disturbances that grow the fastest in the system while the *rms* errors of u and θ' measure the accuracy of the fields in both storm region and in the environment. For the latter, the error saturation is harder to reach. The overall error evaluations are similar for all three cases.

The forecast fields at the 6 km level and at the surface for TLX, TLXCYR and CASA3S as compared to the truth (not shown due to space limitation). It is found that the main updraft of the right mover is well forecasted up to 130 min, although a slight westward position error is seen in all three experiments. The features associated with the further split left movers are not predicted as well. By 190 min, the main updraft appears better positioned in TLXCYR although all three show northwestward position errors. The left moving cells are mostly out of the plotting domain at this time.

At the surface, the general patterns of cold pool and gust front, and the precipitation regions are reasonably well predicted even at 190 min. Errors in the details do exist. At 190 min, the reflectivity pattern of TLXCYR appears worst among the three cases, with spurious precipitation appearing within the cold pool west of the main precipitating downdraft. Such discrepancy shows up in the *rms* plotted in Fig. 12. At 190 min, the *rms* error in u is almost 1 ms^{-1} higher and in w 0.5 ms^{-1} higher in TLXCYR than in TLX or CASA3S. The error in θ' is about 0.3 K higher in TLXCYR. Because the error growth is nonlinear in such a convective system, the exact cause of such behaviors at this late stage is not clear. Still, all three analyses, as produced by TLX, TLXCYR and CASA3S, can be considered very good, because the ensuing forecasts remain good in terms of the storm morphology and *rms* errors for up to two hours. A significant portion of the *rms* errors at the later times is due to position errors. The forecasts starting from the analysis of TLXNCov or CASA1S are much poorer (not shown).

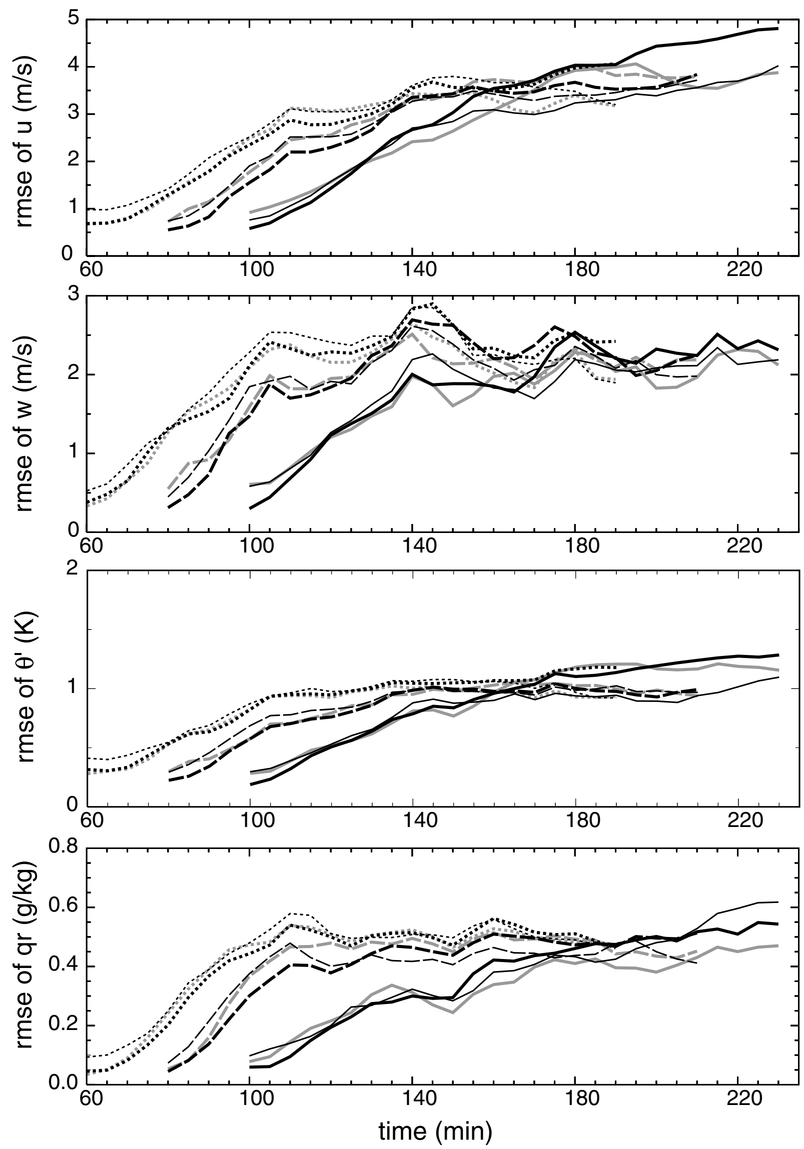

Fig. 12. The *rms* errors of forecasts averaged over the entire domain for: (a) u (m s^{-1}), (b) w (ms^{-1}), (c) θ' (K) and (d) q_r (g kg^{-1}). The forecasts begin from ensemble-mean analysis at t = 60 min (dotted), t = 80 min (dashed) and t = 100 min (solid) of experiments TLX (thin black curves), TLXCYR (thick black curves) and CASA3S (thick gray curves).

6. Summary

In this study we used the ensemble Kalman filter technique to assimilate simulated radial velocity and reflectivity data from an assumed WSR-88D radar and a network of four low-cost radars planned for the first Oklahoma test bed by the CASA project. Forecasts starting from selected analyses were also performed to further examine the quality of analysis and its impact on forecast.

It is shown that the assimilation of data from a CASA radar, in addition to data from one WSR-88D radar located about 90 km away, improves the analysis. Such improvement is most significant at the low levels where the WSR-88D radar does not observe. The subsequent forecast is also better for at least 40 min with the addition of CASA radar data. The experiments also show that when a single CASA radar is assimilated and when the radar does not provide full coverage of the storm system, significant errors can develop in the analysis that cannot be effectively corrected. The combination of several CASA radars effectively eliminates the problem when a complete coverage is available.

The impact of storm motion speed on the quality of EnKF analysis is also examined. In general, the analysis is better for a slower moving system although the quality of analysis of a fast moving storm is reasonable too when good data coverage is available. The quality of analysis can be improved by employing faster volume scans, especially for fast moving systems, but the sensitivity of the EnKF analysis of convective storms to the volume scan frequency is much less than that of more traditional single-Doppler velocity and thermodynamic retrieval schemes. In fact, very good analyses can be obtained even with the WSR-88D-standard 5-min volume scan frequency. For this reason, more versatile scan strategies may be developed for and employed by the CASA radars in response to user needs. For example, complete volume scans can be made by the radar network every 5 minutes in perhaps 1-2 minute periods while in between, the radars can be doing sector scans that focus on active local features such as tornado and microburst.

In general, the data coverage appears to be the most significant factor that affects the quality of storm analysis. Once a storm or a portion of it moves out of the range of radar network, the model error growth can no longer be controlled and the analysis deteriorates. The lack of data coverage in the entire analysis domain also negatively impacts the overall analysis as some spurious cells can be not suppressed. When the data coverage is similar, the analysis of a slow-moving storm system is slightly better than that of a fast moving one.

An additional experiment was also performed in which the V_r data are used to update = directly related wind components only and the Z data are used to update only the precipitating hydrometeor species. The updating of other variables based on the cross covariances of the background error is not performed. The analysis from this experiment is significantly poorer than the corresponding one using the cross covariance information, and the results demonstrate clearly that cross covariances play a key role in 'retrieving' unobserved fields in a storm system when assimilating radar data. At the same time, the perhaps equally important role of the model dynamics in the retrieval process is also discussed.

In the end, we point out that the CASA radar-related issues examined in this study are only a few of many. The dynamic adaptive systems in the CASA radar networks promise to establish a completely new paradigm for the sensing of the atmosphere, and the impact of data collected using a variety of possible scanning modes remain to be studied in a more systematic way. This study represents only the first step towards this direction.

Acknowledgement

This work was supported by NSF grants ATM0129892, EEC-0313747, and ATM-0331594. M. Xue was also supported by a DOT-FAA grant via DOC-NOAA NA17RJ1227 and a grant from Chinese Natural Science Foundation No. 40028504. The authors also benefited from discussions with Drs. Juanzhen Sun, Keith Brewster, and other members of the CASA project. The computations were performed at the Pittsburgh Supercomputing Center supported by NSF.

References

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

Burgers, G., P. J. v. Leeuwen, and G. Evensen, 1998: Analysis scheme in the ensemble Kalman filter. *Mon. Wea. Rev.*, **126,** 1719-1724.

Doviak, R. and D. Zrnic, 1993: *Doppler Radar and Weather Observations*. 2nd ed. Academic Press, 562 pp.

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.***,** 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(C5),** 10 143-10 162.

——, 2003: The ensemble Kalman filter: Theoretical

formulation and practical implementation. *Ocean Dynamics*, **53,** 343-367.
Gal-Chen, T., 1978: A method for the initialization of the anelastic equations: Implications for matching models with observations. *Mon. Wea. Rev.*, **106,** 587-606.
——, 1982: Errors in fixed and moving frame of references: Applications for convectional and Doppler radar analysis. *J. Atmos. Sci.*, **39,** 2279-2300.
Gal-Chen, T. and R. A. Kropfli, 1984: Buoyancy and pressure perturbation derived from dual-Doppler radar observations of the planetary boundary layer: Applications for matching models with observations. *J. Atmos. Sci.*, **41,** 3007-3020.
Gaspari, G. and S. E. Cohn, 1999: Construction of correlation functions in two and three dimensions. *Quart. J. Roy. Meteor. Soc.*, **125,** 723-757.
Houtekamer, P. L. and H. L. Mitchell, 1998: Data assimilation using an ensemble Kalman filter technique. *Mon. Wea. Rev.*, **126,** 796-811.
——, 2001: A sequential ensemble Kalman filter for atmospheric data assimilation. *Mon. Wea. Rev.*, **129,** 123-137.
Lin, Y.-L., R. D. Farley, and H. D. Orville, 1983: Bulk parameterization of the snow field in a cloud model. *J. Climate Appl. Meteor.*, **22,** 1065-1092.
Liou, Y. C. and I. S. Luo, 2001: An investigation of the moving-frame single-Doppler wind retrieval technique using Taiwan area mesoscale experiment low-level data. *J. Appl. Meteor.*, **40,** 1900-1917.
Qiu, C.-J. and Q. Xu, 1992: A simple adjoint method of wind analysis for single-Doppler data. *J. Atmos. Oceanic Tech.*, **9,** 588-598.
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.
Shapiro, A., S. Ellis, and J. Shaw, 1995: Single-Doppler radar retrievals with Phoenix II data: Clear air and microburst wind retrievals in the planetary boundary layer. *J. Atmos. Sci.*, **52,** 1265-1287.
Shapiro, A., P. Robinson, J. Wurman, and J. Gao, 2003: Single-Doppler velocity retrieval with rapid-scan radar data. *J. Atmos. Ocean. Tech.*, **20,** 1758-1775.
Smith, P. L., Jr., C. G. Myers, and H. D. Orville, 1975: Radar reflectivity factor calculations in numerical cloud models using bulk parameterization of precipitation processes. *J. Appl. Meteor.*, **14,** 1156-1165.
Snyder, C. and F. Zhang, 2003: Assimilation of simulated Doppler radar observations with an ensemble Kalman filter. *Mon. Wea. Rev.*, **131,** 1663-1677.
Sun, J. and N. A. Crook, 1996: Comparison of thermodynamic retrieval by the adjoint method with the traditional retrieval method. *Mon. Wea. Rev.*, **124,** 308–324.
——, 2001: Real-time low-level wind and temperature analysis using single WSR-88D data. *Wea. Forecasting*, **16,** 117-132.
Tippett, M. K., J. L. Anderson, C. H. Bishop, T. M. Hamill, and J. S. Whitaker, 2003: Ensemble Square Root Filters{ast}. *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.***,** Accepted.
Whitaker, J. S. and T. M. Hamill, 2002: Ensemble data assimilation without perturbed observations. *Mon. Wea. Rev.*, **130,** 1913-1924.
Xue, M., K. K. Droegemeier, and V. Wong, 2000: The Advanced Regional Prediction System (ARPS) - A multiscale nonhydrostatic atmospheric simulation and prediction tool. Part I: Model dynamics and verification. *Meteor. Atmos. Physics*, **75,** 161-193.
Xue, M., D.-H. Wang, J.-D. Gao, K. Brewster, and K. K. Droegemeier, 2003: The Advanced Regional Prediction System (ARPS), storm-scale numerical weather prediction and data assimilation. *Meteor. Atmos. Physics*, **82,** 139-170.
Xue, M., K. K. Droegemeier, V. Wong, A. Shapiro, K. Brewster, F. Carr, D. Weber, Y. Liu, and D.-H. Wang, 2001: The Advanced Regional Prediction System (ARPS) - A multiscale nonhydrostatic atmospheric simulation and prediction tool. Part II: Model physics and applications. *Meteor. Atmos. Phy.*, **76,** 143-165.
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.
Zhang, J. and T. Gal-Chen, 1996: Single Doppler wind retrieval in the moving frame of reference. *J. Atmos. Sci.*, **53,** 2609-2623.