Intercomparison of an Ensemble Kalman Filter with Three- and Four-Dimensional Variational Data Assimilation Methods in a Limited-Area Model over the Month of June 2003

MENG ZHANG AND FUQING ZHANG

Department of Meteorology, The Pennsylvania State University, University Park, Pennsylvania

XIANG-YU HUANG AND XIN ZHANG

National Center for Atmospheric Research, Boulder, Colorado

(Manuscript received 27 August 2010, in final form 15 October 2010)

ABSTRACT

This study compares the performance of an ensemble Kalman filter (EnKF) with both the three-dimensional and four-dimensional variational data assimilation (3DVar and 4DVar) methods of the Weather Research and Forecasting (WRF) model over the contiguous United States in a warm-season month (June) of 2003. The data assimilated every 6 h include conventional sounding and surface observations as well as data from wind profilers, ships and aircraft, and the cloud-tracked winds from satellites. The performances of these methods are evaluated through verifying the 12- to 72-h forecasts initialized twice daily from the analysis of each method against the standard sounding observations. It is found that 4DVar has consistently smaller error than that of 3DVar for winds and temperature at all forecast lead times except at 60 and 72 h when their forecast errors become comparable in amplitude, while the two schemes have similar performance in moisture at all lead times. The forecast error of the EnKF is comparable to that of the 4DVar at 12–36-h lead times, both of which are substantially smaller than that of the 3DVar, despite the fact that 3DVar fits the sounding observations much more closely at the analysis time. The advantage of the EnKF becomes even more evident at 48–72-h lead times; the 72-h forecast error of the EnKF is comparable in magnitude to the 48-h error of 3DVar/4DVar.

1. Introduction

Data assimilation (DA) is known as the process of creating the best estimate of the initial state for numerical weather prediction (NWP) models through combining all sources of information, including the first guess from previous short-term model forecasts and observations, along with the associated uncertainties in each source of information. Operational NWP models have predominantly used three- or four-dimensional variational (3DVar or 4DVar) data assimilation methods over the past couple of decades. Ensemble-based data assimilation methods, in various forms of ensemble Kalman filters (EnKFs) first proposed by Evensen (1994), which use a short-term ensemble to estimate the flow-dependent background error covariance, are becoming a popular alternative to the variational approaches.

Recently, systematic intercomparisons between the EnKF and 3DVar/4DVar have been conducted in various NWP models. For example, Houtekamer et al. (2005) compared the EnKF with 3DVar for the Canadian global model, and Whitaker et al. (2008) did so for the National Centers for Environmental Prediction's (NCEP's) Global Forecast System (GFS). Both studies showed convincingly the advantages of the EnKF over 3DVar in global models. For limited-area models, Meng and Zhang (2008a,b) also showed that the EnKF consistently outperformed the 3DVar within the Weather Research and Forecasting (WRF) model framework at the regional scales. More recently, Yang et al. (2009) compared the EnKF with 3D/4DVar using a quasigeostrophic model. Buehner et al. (2010a,b) compared the EnKF with 4DVar in an operational scenario of Canadian global model while Miyoshi et al. (2010) did so for the Japan Meteorological Agency's (JMA's) operational global model; both studies showed that the EnKF has comparable performance to

Corresponding author address: Dr. Fuqing Zhang, Department of Meteorology, The Pennsylvania State University, University Park, PA 16802.
E-mail: fzhang@psu.edu

DOI: 10.1175/2010MWR3610.1

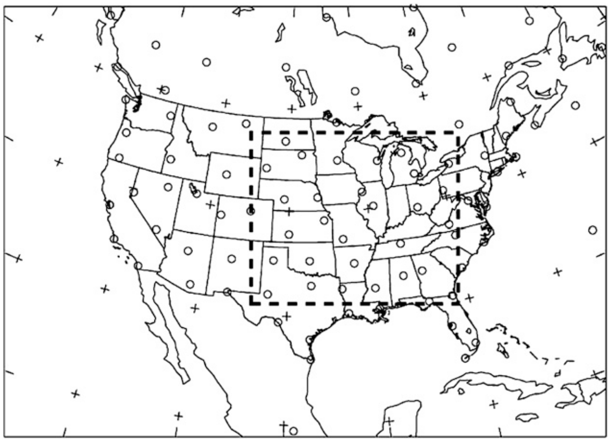

FIG. 1. Model domain configuration. The circles denote the locations of the radiosonde observations used for assimilation and verification. The dashed box shows the subset of the domain selected for verification statistics.

4DVar in global models. For limited-area models, Caya et al. (2005) compared the performance of the EnKF with 4DVar using synthetic observations for a supercell event in an anelastic cloud model and found that the EnKF and 4DVar performed comparably. Moreover, both Buehner et al. (2010a,b) and Caya et al. (2005) found that the EnKF-initialized models will have better forecast performance at a later stage of their forecasts after a few assimilation cycles even though 4DVar may fit the observations better at the analysis time and requires less than spinup time from a cold start. To the best of our knowledge, no study has directly compared the EnKF with 4DVar in a limited-area primitive model, let alone with real-data applications, likely because of the scarcity of limited-area 4DVar systems.

The current study compares the EnKF with the newly developed and released community 4DVar data assimilation system in the WRF framework (Huang et al. 2009) with the EnKF of Meng and Zhang (2008a,b). Also compared is the 3DVar system (Barker et al. 2004), which is also available in WRF. Experimental setup of the three DA systems is shown in section 2. Results of the intercomparison are discussed in section 3 and conclusions are presented in section 4.

2. Experimental design

In this study, the regional Advanced Research Weather Research and Forecasting (ARW-WRF) model (Skamarock et al. 2005) version 3.1.1 is employed as the platform to investigate all DA approaches in a realistic scenario. All experiments are conducted over a single domain covering

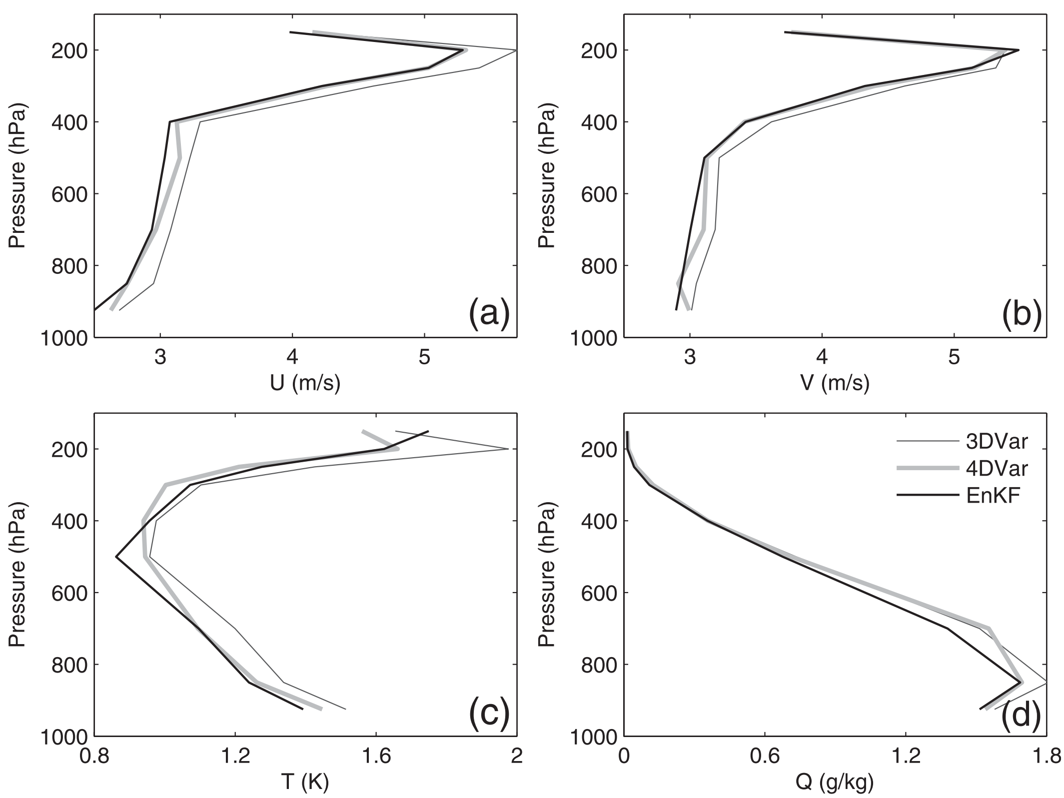

FIG. 2. Vertical profiles of the month-averaged 12-h forecast RMSEs of (a) U (m s^{-1}), (b) V (m s^{-1}), (c) T (K), and (d) Q (g kg^{-1}) for various DA methods.

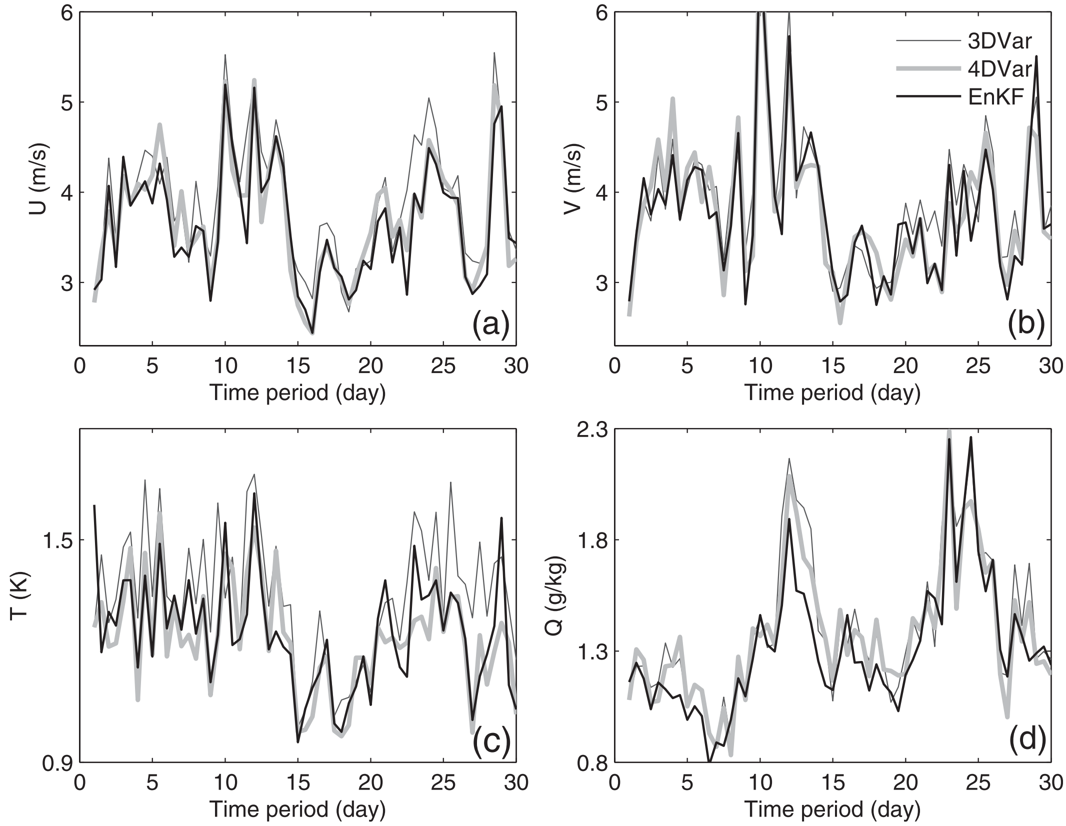

FIG. 3. Time evolution of the vertically averaged 12-h forecast RMSE of (a) U (m s^{-1}), (b) V (m s^{-1}), (c) T (K), and (d) Q (g kg^{-1}).

the contiguous United States and surrounding areas (Fig. 1), which has 71 × 51 horizontal mesh grids with 90-km spacing and 27 vertical levels up to 50 hPa. The Grell–Devenyi cumulus scheme, WRF single-moment six-class graupel microphysics scheme, and Yonsei State University (YSU) boundary layer scheme are used for all deterministic forecasts, whereas multiphysics parameterizations are considered for ensemble forecasts in the EnKF [as in Meng and Zhang (2008a,b)]. The first forecast cycle of this month-long experiment is initialized at 0000 UTC 1 June 2003, where the NCEP global final analysis (FNL) data are used to create the initial and lateral boundary conditions (IC and LBC). In subsequent cycles, the lateral boundary conditions are interpolated from the FNL analyses, whereas the initial conditions are updated and propagated from previous DA analyses.

The WRF variational data assimilation system (WRF-Var) version 3 (Huang et al. 2009) and the WRF EnKF system of Meng and Zhang (2008a,b) are used for all DA experiments. The WRF-Var has both 3DVar and 4DVar capabilities; the newly released 4DVar component is developed as an extension from the previous WRF 3DVar system (Barker et al. 2004). The background error covariance in 3DVar/4DVar experiments is static and prescribed by the National Meteorological Center (NMC) method (Parrish and Derber 1992), which assumes homogeneous and isotropic correlations for a set of independent control variables derived from the forecast difference between the 24- and 12-h lead time forecasts of the preceding month for the given domain (i.e., May 2003 in this case). The variance scale parameter is carefully tuned to 3.0 rather than the default value of 1.0 to give the best performance of the 3DVar among various experiments tested (not shown; this also improves the performance of 4DVar as compared to the default). The EnKF adopts settings similar to those in Meng and Zhang (2008a,b), that is, an ensemble of 40 members with diverse physics parameterization schemes, perturbed lateral boundary conditions from FNL analyses, a relaxation coefficient of 0.8 (Zhang et al. 2004), and the covariance localizations of Gaspari and Cohn (1999) using a radius of influence of 1800 km for radiosondes and profilers and 600 km for other observations. The initial ensemble perturbations are randomly generated at 0000 UTC 1 June 2003 through the CV5 option (Barker et al. 2004) of the WRF-Var system, as in Meng and Zhang (2008a,b).

Various types of meteorological observations from the archived Meteorological Assimilation Data Ingest System

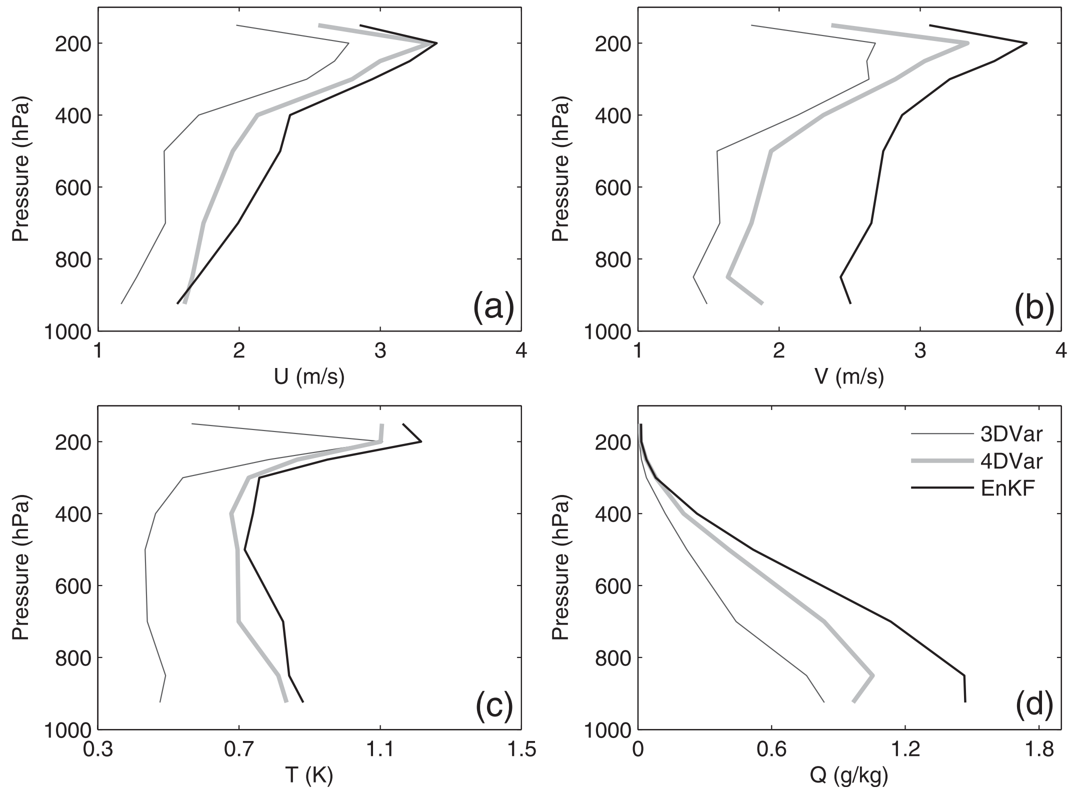

FIG. 4. As in Fig. 2, but for the analysis RMS fits to observations.

(MADIS) are assimilated during the month of June 2003, namely winds, temperature, moisture, and surface pressure from radiosondes, ships, and surface stations; winds from profilers; winds and temperature from aircraft; and cloud-tracked winds from satellites. Satellite radiances as well as other indirect observations are not used in this study because of the complexity of the observation operators. The observation preprocessing module (OBSPROC) of WRF-Var is implemented for data sorting, quality control, and observational error assignment (Barker et al. 2004). The first analysis time for all experiments is at 1200 UTC 1 June 2003; afterward, continuous DA cycles (analysis and forecast) with 6-h intervals are performed until the end of the month, valid at every 0000, 0600, 1200, and 1800 UTC. The assimilation window of 4DVar covers the period from -3 h to $+3$ h of each analysis time; therefore, all available observations distributed over such a 6-h window are assimilated at their exact time rather than at an approximate analysis time as in 3DVar and EnKF.

3. Intercomparison of 3DVar, 4DVar, and EnKF

The performance of all three DA systems over the one-month period is examined in this section. The root-mean-square error (RMSE) of horizontal winds (U, V), temperature T, and the mixing ratio of water vapor Q are calculated between model forecasts and radiosonde observations over a subset of the model domain (dashed box in Fig. 1). The same area is used for verification of forecast performance as in Meng and Zhang (2008b). There are a total of 59 individual 72-h WRF deterministic forecasts, initialized respectively from the variational and EnKF mean analyses at 0000 and 1200 UTC each day, used in all the verification statistics shown below.

We first examine the vertical distribution of the mean RMSE of the12-h WRF forecasts of U, V, T, and Q averaged over the entire month (Fig. 2). At a first glance, the largest RMSE of U, V, and T is around 200 hPa near the tropopause while the largest error of Q is in the mid to lower troposphere. Comparison of vertical profile of the mean 12-h forecast error from each DA experiment shows that the EnKF performs comparably to 4DVar for the horizontal winds and temperature fields while the EnKF has noticeably smaller RMSE than 4DVar for moisture throughout the troposphere, especially at mid levels. The comparable performance of the EnKF versus 4DVar for the wind and temperature fields and the advantage of the EnKF with regard to moisture are fairly persistent throughout the month, as can be seen from the time series of vertically averaged RMSE shown in

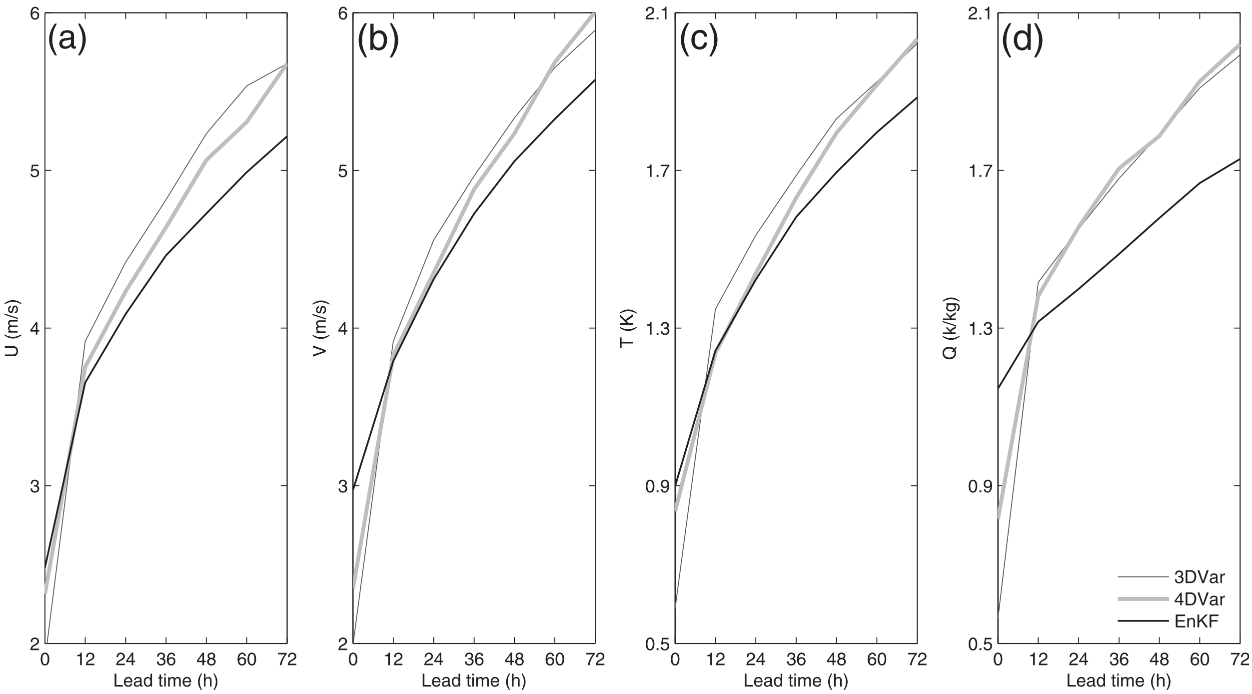

FIG. 5. Domain-averaged RMSEs further averaged over all 59 WRF forecasts of the month for each DA experiment at forecast lead times from 0 to 72 h evaluated every 12 h for (a) U (m s^{-1}), (b) V (m s^{-1}), (c) T (K), and (d) Q (g kg^{-1}).

Fig. 3. Both the EnKF and 4DVar have substantially smaller RMSEs than does 3DVar for all forecast variables at nearly all times (Figs. 2 and 3). The better performance of the EnKF over 3DVar is consistent with the findings of Meng and Zhang (2008b) despite the coarser resolution in this study. It is worth noting that despite using a coarse horizontal resolution, the RMSEs for both the EnKF and 3DVar in the current study are similar in magnitude to those in Meng and Zhang (2008b; see their Figs. 4 and 6), suggesting that the difference between variational and ensemble methods discussed herein is not sensitive to model resolution.

Figure 4 displays the vertical profiles of the RMS differences between the analysis of each DA scheme and the sounding observations averaged over all 59 analyses. Note that these differences are no longer called "errors" but are considered as a measure of the fit of the analysis to observations, since all verifying observations are assimilated by each scheme and thus no longer are considered to be independent verifications at the analysis time. In sharp contrast to Fig. 2 of the 12-h forecast error, Fig. 4 shows that the 3DVar scheme best fits the observations, followed by 4DVar, while the EnKF analysis has the largest difference between the radiosonde observations throughout the vertical domain. This result confirms that an analysis with a closer fit to observations does not necessarily lead to a better forecast.

Figure 5 summarizes the domain-averaged (in both vertical and horizontal directions) RMSEs further averaged over all 59 WRF forecasts of the month for each DA experiment at forecast lead times from 0 to 72 h evaluated every 12 h. Beyond what have been shown in Figs. 2 and 3, it is found that 4DVar has consistently smaller error than 3DVar from 12- to 48-h forecast lead times for horizontal winds and temperature fields, but their forecast error amplitude becomes comparable afterward (at 60 and 72 h). The differences in the RMSE of the moisture field between 4DVar and 3DVar are relatively small, since few moist processes are considered in the linear models of 4DVar (Huang et al. 2009). This implies that the 4DVar may not resolve more flow-dependent moist information through adjoint minimization than that in the static background error covariance as in 3DVar. On the other hand, the advantage of the EnKF over both 3DVar and 4DVar becomes more evident after the 36-h forecast time for all forecast fields, while the EnKF moisture forecast field is superior to those of both 3DVar and 4DVar at all lead times despite fitting less closely to the observations at the analysis time. This result is consistent with the time dependence of the relative performance of 4DVar and EnKF in previous works (Caya et al. 2005; Buehner et al. 2010a,b). It is rather remarkable that the 72-h forecast error of the EnKF is comparable in magnitude to the 48-h error of 3DVar/4DVar, a gain of nearly 1-day lead time in forecast accuracy.

Figure 6 shows the time evolution of the vertically averaged 72-h RMSEs from all DA experiments, which

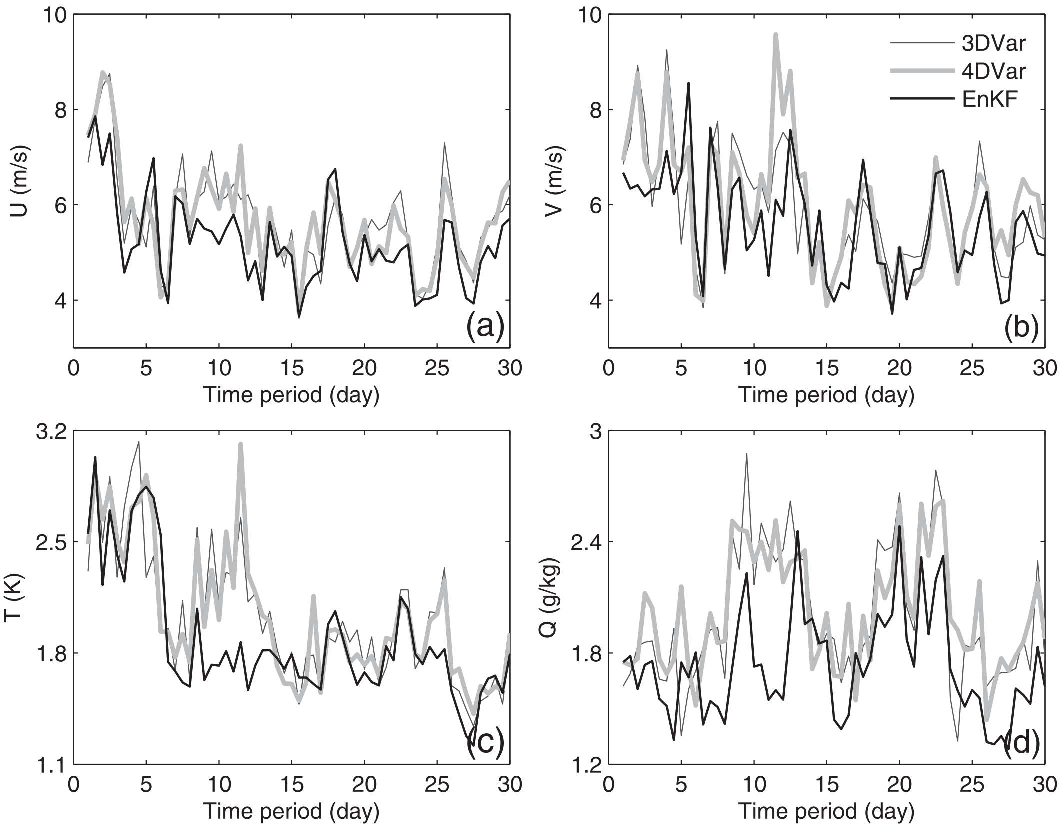

FIG. 6. As in Fig. 3, but for the 72-h forecast RMSEs.

displays an apparent fluctuation between different initialization times. The 72-h RMSE of the 4DVar-initialized forecasts is comparable to that of the 3DVar over all months while the advantage of the EnKF over the two variational schemes is most evident for a few episodes (e.g., 10–13 June) that feature the passage of strong mesoscale convective systems (Davis et al. 2004; Hawblitzel et al. 2007).

The clear advantage of the EnKF over 3DVar and 4DVar in the moisture field in all verifications signifies the benefit of using a flow-dependent background error covariance, since the balance used to estimate the background error covariance in WRF-Var (the same for both 3DVar and 4DVar) does not include the multivariate correlations for moisture field. The lack of multivariate balance may be one of the key factors impacting the DA performance. The EnKF also benefits from using a multiphysics ensemble mean for an improved forecast prior before the analysis (Meng and Zhang 2008a). However, it is unclear whether the faster growth of forecast error in winds and temperature of variational methods is a result of the larger error in moisture field initially or is due to the use of a static background error covariance that is not in balance with the flow-dependent moisture field, or a combination of both, which will be further examined in a future study.

4. Concluding remarks

The intercomparison of three state-of-the-art data assimilation approaches is presented on the platform of the regional-scale WRF model configurations, covering both variational (3DVar/4DVar) and ensemble-based (EnKF) methods over the month of June 2003. The accuracy of the 12-to-72 h forecasts initiated from various DA methods by assimilating surface and upper-air observations is examined. The EnKF and 4DVar generally have comparable performance, both of which are better than 3DVar, for the 12-to-36-h forecast skills of horizontal winds and temperature, while the EnKF has a substantially lower RMSE in moisture than 4DVar (and 3DVar) at all forecast lead times. The advantage of the EnKF for the moisture field is likely due to its use of multivariate flow-dependent background error covariance in contrast to the static covariance used in 3DVar/4DVar, as found in previous limited-area studies of Meng and Zhang (2008b) and Torn and Hakim (2008). The advantages of the EnKF over 4DVar (and 3DVar) become more pronounced for all forecast quantities examined at long forecast lead times (48–72 h). This advantage is particularly remarkable during active convective episodes of the month over the verifying domain.

Between the two variational schemes, the 4DVar consistently outperforms 3DVar for both horizontal winds and temperature up to 60-h forecast lead times but becomes comparable afterward. The 4DVar has similar amplitude of moisture RMSE to that of 3DVar at all lead times, suggesting that the simple moist physics used in the adjoint model of 4DVar is insufficient to adjust the moisture field to other fields.

The current study is based on a relatively coarse resolution because of the high computational cost required by the current 4DVar implementation. More systematic comparisons are planned in the future on the DA performances over a finer-resolution domain. On the other hand, since the advantage of 4DVar over 3DVar comes from the adjoint integration and minimization over the entire 6-h assimilation window while the advantage of the EnKF over 3DVar comes from the flow-dependent background error covariance, it is foreseeable to combine both of these advantages in a hybrid data assimilation system that couples the EnKF with 4DVar. A proof-of-concept study of such a hybrid system is shown to be promising in Zhang et al. (2009); there the short-term ensemble forecast from the EnKF provides the flow-dependent error covariance for 4DVar while the 4DVar analysis replaces the mean of the EnKF analysis. Our ongoing work investigates the coupling of an EnKF with 4DVar in the WRF framework that builds on the work of Zhang et al. (2009) as well as on the WRF EnKF–3DVar hybrid study of Wang et al. (2008); the preliminary results are very encouraging and will be reported in a separate paper.

Acknowledgments. We thank Zhiyong Meng at Peking University for her constructive comments. This work was supported by the Office of Naval Research Grants N000140410471 and N000140910526 and the National Science Foundation Grant ATM-0840651.

REFERENCES

Barker, D. M., W. Huang, Y.-R. Guo, A. Bourgeois, and Q. N. Xiao, 2004: A three-dimensional variational data assimilation system for MM5: Implementation and initial results. *Mon. Wea. Rev.,* **132,** 897–914.

Buehner, M., P. L. Houtekamer, C. Charette, H. L. Mitchell, and B. He, 2010a: Intercomparison of variational data assimilation and the ensemble Kalman filter for global deterministic NWP. Part I: Description and single-observation experiments. *Mon. Wea. Rev.,* **138,** 1550–1566.

——, ——, ——, ——, and ——, 2010b: Intercomparison of variational data assimilation and the ensemble Kalman filter for global deterministic NWP. Part II: One-month experiments with real observations. *Mon. Wea. Rev.,* **138,** 1567–1586.

Caya, A., J. Sun, and C. Snyder, 2005: A comparison between the 4DVAR and the ensemble Kalman filter techniques for radar data assimilation. *Mon. Wea. Rev.,* **133,** 3081–3094.

Davis, C., and Coauthors, 2004: The Bow-Echo and MCV Experiment: Observations and opportunities. *Bull. Amer. Meteor. Soc.,* **85,** 1075–1093.

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.

Gaspari, G., and S. E. Cohn, 1999: Construction of correlation functions in two and three dimensions. *Quart. J. Roy. Meteor. Soc.,* **125,** 723–757.

Hawblitzel, D. P., F. Zhang, Z. Meng, and C. A. Davis, 2007: Probabilistic evaluation of the dynamics and predictability of a mesoscale convective vortex of 10–13 June 2003. *Mon. Wea. Rev.,* **135,** 1544–1563.

Houtekamer, P. L., H. L. Mitchell, G. Pellerin, M. Buehner, M. Charron, L. Spacek, and B. Hansen, 2005: Atmospheric data assimilation with an ensemble Kalman filter: Results with real observations. *Mon. Wea. Rev.,* **133,** 604–620.

Huang, X.-Y., and Coauthors, 2009: Four-dimensional variational data assimilation for WRF: Formulation and preliminary results. *Mon. Wea. Rev.,* **137,** 299–314.

Meng, Z., and F. Zhang, 2008a: Tests of an ensemble Kalman filter for mesoscale and regional-scale data assimilation. Part III: Comparison with 3DVAR in a real-data case study. *Mon. Wea. Rev.,* **136,** 522–540.

——, and ——, 2008b: Tests of an ensemble Kalman filter for mesoscale and regional-scale data assimilation. Part IV: Comparison with 3DVAR in a month-long experiment. *Mon. Wea. Rev.,* **136,** 3671–3682.

Miyoshi, T., Y. Sato, and T. Kadowaki, 2010: Ensemble Kalman filter and 4D-Var intercomparison with the Japanese operational global analysis and prediction system. *Mon. Wea. Rev.,* **138,** 2846–2866.

Parrish, D. F., and J. C. Derber, 1992: The National Meteorological Center's spectral statistical-interpolation analysis system. *Mon. Wea. Rev.,* **120,** 1747–1763.

Skamarock, W. C., J. B. Klemp, J. Dudhia, D. O. Gill, D. M. Barker, W. Wang, and J. G. Powers, 2005: A description of the Advanced Research WRF Version 2. NCAR Tech. Note 468+STR, 88 pp.

Torn, R. D., and G. J. Hakim, 2008: Performance characteristics of a pseudo-operational ensemble Kalman filter. *Mon. Wea. Rev.,* **136,** 3947–3963.

Wang, X., D. M. Barker, C. Snyder, and T. M. Hamill, 2008: A hybrid ETKF–3DVAR data assimilation scheme for the WRF model. Part II: Real observation experiments. *Mon. Wea. Rev.,* **136,** 5132–5147.

Whitaker, J. S., T. M. Hamill, X. Wei, Y. Song, and Z. Toth, 2008: Ensemble data assimilation with the NCEP global forecast system. *Mon. Wea. Rev.,* **136,** 463–482.

Yang, S.-C., M. Corazza, A. Carrassi, E. Kalnay, and T. Miyoshi, 2009: Comparison of local ensemble transform Kalman filter, 3DVAR, and 4DVAR in a quasigeostrophic model. *Mon. Wea. Rev.,* **137,** 693–709.

Zhang, F., C. Snyder, and J. Sun, 2004: Impacts of initial estimate and observation availability on convective-scale data assimilation with an ensemble Kalman filter. *Mon. Wea. Rev.,* **132,** 1238–1253.

——, M. Zhang, and J. A. Hansen, 2009: Coupling ensemble Kalman filter with four-dimensional variational data assimilation. *Adv. Atmos. Sci.,* **26,** 1–8.