

An ensemble Kalman filter for WRF and a comparison with the WRF three-dimensional variational assimilation scheme

Alain Caya, Dale M. Barker, and Chris Snyder

National Center for Atmospheric Research, Boulder, Colorado, USA

E-mail: caya@ucar.edu

The ensemble Kalman filter (EnKF) and its variants are promising candidates for a variety of assimilation problems in atmospheric science. The EnKF employs flow-dependent forecast error covariances (estimated from an ensemble of forecasts), and produces an estimate of analysis uncertainty in the form of an ensemble of analyses. At the same time, the EnKF is relatively simple to implement for large atmospheric models, as it does not require minimization techniques or the model's adjoint. We have implemented an EnKF for the Weather Research and Forecasting (WRF) model. The feasibility of the scheme has been explored in experiments with simulated observations for both synoptic-scale domains covering North America and much smaller domains in which moist convection is explicitly resolved. The EnKF provides excellent estimates of the atmospheric state given, in the former case, observations broadly representative of the North-American radiosonde network or, in the later case, a single Doppler radar. The EnKF compares favorably to the three-dimensional variational assimilation scheme (3D-Var) available for WRF using simulated wind and temperature observations from the radiosonde network. However, the EnKF introduces somewhat more spurious gravity-wave noise than the 3D-Var for this assimilation problem.

1. Introduction

The Data Assimilation Initiative at NCAR is developing ensemble-based data assimilation schemes applicable to a variety of models. The list of supported models ranges from a 3-variable model, to full General Circulation Models. Other models can be added easily with minimal work. More details can be found at <http://www.image.ucar.edu/DAI/>. Results using the WRF model are presented here and evaluated by comparing with the WRF 3D-Var in the context of perfect model. The EnKF is also evaluated in terms of imbalance introduced by the assimilation procedure.

2. An EnKF

The general approach to ensemble-based data assimilation employed here is described in Anderson (2003). The specific algorithm, referred in the present study as EnKF, is a deterministic ensemble (square-root) filter (Anderson, 2001) and does not require perturbed observations. A 40-member ensemble is used in all EnKF experiments. Techniques to compensate for the sampling errors like covariance inflation are not used and the covariance localization is applied only in the horizontal directions.

3. The WRF 3D-Var

The 3D-Var assimilation system we are using here with the WRF model is described by Barker et al. (2004). The background-error covariance model is given by recursive filters. The variances and correlation lengths are obtained from the "NMC method" using global forecasts and are not specifically tuned for the present application. Specifically, the climatological background error covariances applied in this WRF-based study are derived from forecast differences of NCEP's (GFS) global model (Wu et al, 2002). Control variables are the balanced streamfunction ψ , unbalanced velocity potential χ , unbalanced temperature T , unbalanced surface pressure p_s , and a pseudo relative humidity variable.

4. Experimental setup

Synthetic observations are generated using the WRF model, the same model used for the assimilation with the 3D-Var and the EnKF assimilation schemes. This procedure eliminates the model error from the problem. The model has a 45x45 horizontal grid with 200 km grid spacing, covering essentially North America and parts of the surrounding oceans. The observations are generated at the locations indicated in Fig. 1 every 12 hours for a total assimilation period of 10 days. Observations consist of temperature and the horizontal wind components on the 10 pressure levels indicated in Fig. 2. The observation-error standard deviations for the wind

components are indicated in Fig. 2. The observation-error standard deviation for the temperature observations is independent of pressure and is set to 1 K.

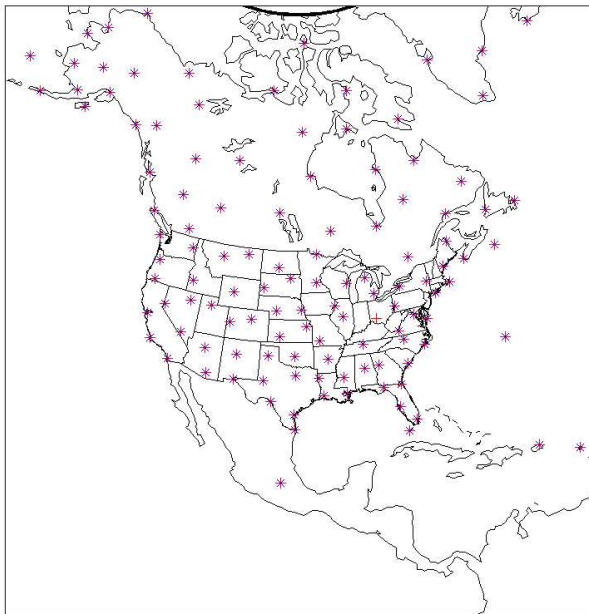


Figure 1. Model domain with the asterisks indicating the locations of the simulated temperature and wind soundings.

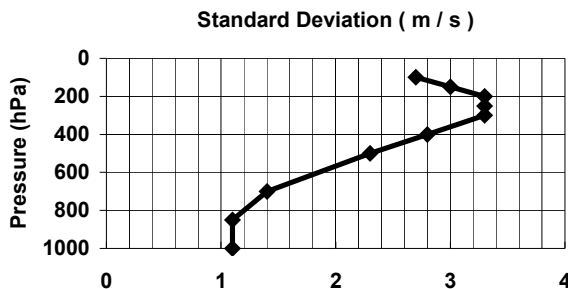


Figure 2. Observational-error standard deviations for the wind components as a function of pressure. The symbols indicate the 10 pressure levels at which the observations are simulated.

The first guess for 3D-Var and the initial ensemble mean for the EnKF is the 00 UTC January 1 2003 NCEP AVN analysis. The corresponding boundary conditions are specified every 6 hours for the 10-day assimilation period.

The initial and boundary conditions for the true state is obtained by superposing perturbations to the first guess. The perturbations are deviations of a randomly chosen 00 UTC January NCEP AVN analysis from the climatology, rescaled by 0.2. The ensemble boundary conditions for the EnKF are

generated in the same way, excluding the sample used for the true state. Thus, the boundary condition uncertainty is well represented by the ensemble of boundary conditions because the true boundary conditions are, by construction, drawn from the same distribution as the ensemble boundary conditions.

The initial ensemble members of the EnKF are generated by superposing perturbations to the ensemble mean first guess. These perturbations are drawn from a normal distribution having the 3D-Var background-error covariances. (Drawing from this distribution is straightforward as the WRF 3D-Var provides subroutines that calculate the product of the square root of the background error covariance matrix with a vector.) This procedure ensures that both assimilation schemes start with the same information about the background-error covariances.

5. Results

Fig. 3 shows the RMS error of the EnKF temperature analysis as a function of the localization half-width used in the assimilation. For localization half-widths between 1280 km and 2560 km, the EnKF temperature results are quite stable. With 640 km for the localization half-width, the EnKF results degrade but still improve over 3D-Var results after 3 assimilation cycles (1 day). Note that the 3D-Var errors are smaller during the first 12 hours (2 assimilation cycles); this is because the initial ensemble perturbations are drawn from the same normal distribution assumed in 3D-Var, and thus the first EnKF analysis simply approximates the 3D-Var analysis. In what follows, all EnKF experiments are performed with a localization half-width of 1280 km.

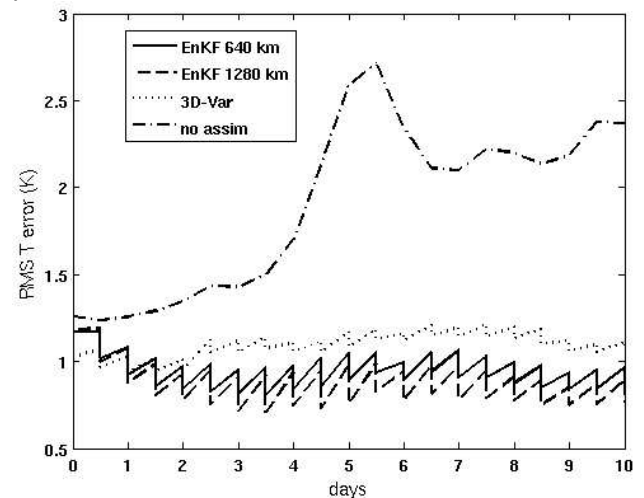


Figure 3. RMS EnKF temperature error as a function of time for 640 and 1280 km localization half-widths. As references, results from 3D-Var and a forecast without assimilation are also showed.

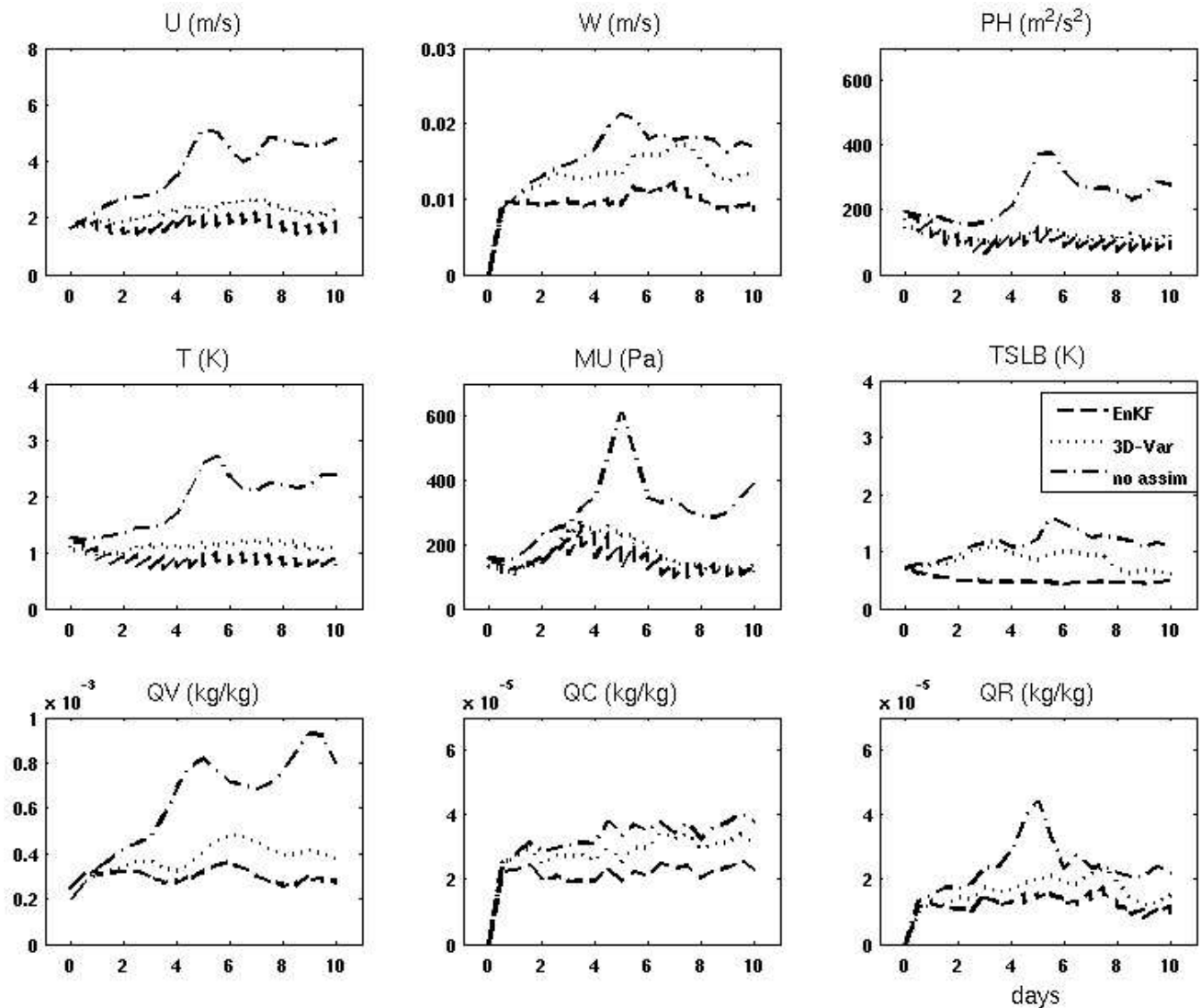


Figure 4. RMS error in the EnKF, 3D-Var, and a forecast without assimilation for: U, the zonal wind component; W, the vertical velocity; PH, the geopotential height; T, the temperature; MU, the total integrated mass; TSLB, the soil temperature; QV, QC, and QR are the specific humidity, cloud, and rain water, respectively.

In Fig. 4, results from the assimilation experiments are shown for 9 variables in the model, and compared to an experiment without assimilation of observations. The errors are generally much smaller with the assimilation of observations and are somewhat smaller for the EnKF than for 3D-Var for all variables. Notice that the errors are growing faster over 12 hours in the EnKF than in the 3D-Var analyses. However, an experiment where the assimilation is stopped after 5 days of assimilation reveals that errors stay smaller in the forecast from the EnKF than the one from 3D-Var analysis over the last five days (not shown).

We now turn to the aspect of imbalance introduced by the two assimilation techniques. Fig. 5 shows the RMS surface pressure tendency at every model time step, which is 10 minutes. Both assimilation schemes introduce imbalances, which are larger in the EnKF. The peak values are especially pronounced for the EnKF analyses, but these peaks appear to be associated with the external mode, which is explicitly damped in the simulations, and decay over two model time steps. The remaining undesirable gravity wave activity propagates and damps away, reaching levels comparable to that of a free forecast during the 12-hour integration before the next assimilation.

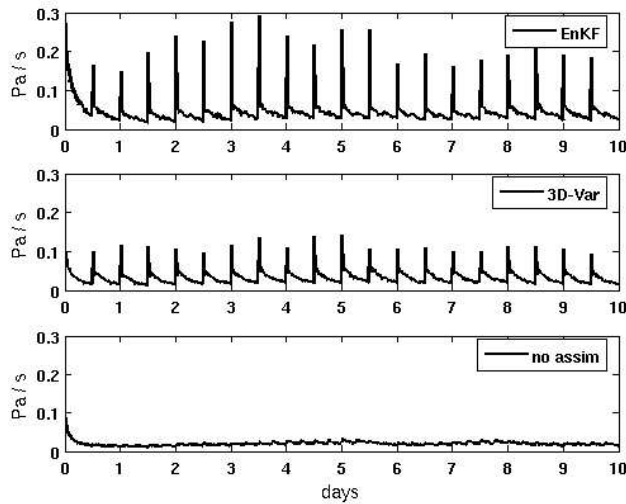


Figure 5. RMS of the surface pressure tendency calculated at each model time step (10 min.) in the EnKF and the 3D-Var analyses, and in the forecast without assimilation.

Fig. 6 shows the sensitivity of the result of Fig. 5 to the localization half-width in the EnKF. The mean of the peaks in Fig. 5, and the overall time average are shown in Fig. 6. For all localization half-widths, the peaks, associated with the external mode, are always higher in the EnKF. However, the overall time average of the surface pressure tendency decreases with increasing localization half-width in the EnKF, approaching the values obtained in the 3D-Var.

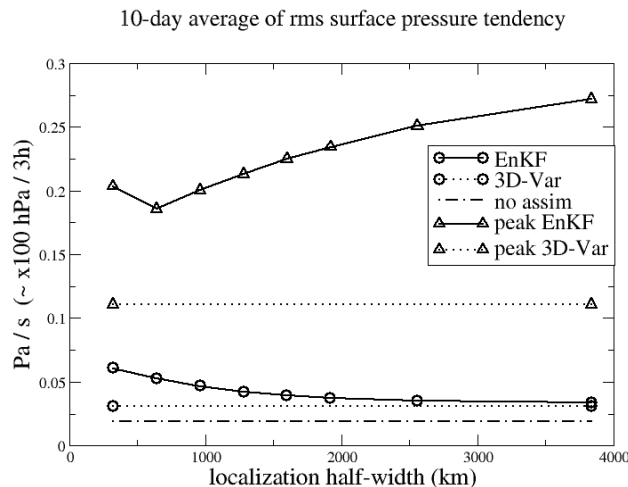


Figure 6. 10-day average of the RMS surface pressure tendency as a function of localization half-width in the EnKF (full line with circles). The 10-day average at the EnKF analysis times is shown by the full line with triangles. Corresponding results for the 3D-Var are indicated by the dotted lines, and the forecast without assimilation by a dash-dotted line.

6. Summary

By evolving the initial climatological forecast-error statistics, the EnKF is able to improve forecast accuracy with time. Thus, the non-stationary forecast-error statistics in the EnKF are beneficial. The RMS errors are smaller in the EnKF than in the 3D-Var analyses after a few assimilation cycles. However, the 3D-Var may improve with the use of more appropriate (i.e. WRF-based, flow-dependent) forecast-error statistics. Also, the difference between the two schemes may be smaller with a larger variety and number of observations. In the case where the model error is large and unknown, the benefit of non-stationary forecast-error statistics in the EnKF would be less apparent.

Examination of the surface pressure tendency reveals that the EnKF introduces somewhat more imbalance than 3D-Var. One obvious solution to this problem is to increase the ensemble size along with less severe covariance localization. We are also exploring more sophisticated techniques for covariance localization. In the case where these modifications to the EnKF are insufficient to reduce the spurious gravity-wave noise introduced during the EnKF assimilation, one may have to consider applying an initialization procedure to the EnKF analysis before the forecast. This initialization procedure could be a digital filter or balance constraints directly built into the EnKF scheme.

References:

- Anderson, J. L.: A local least squares framework for ensemble filtering. *Mon. Wea. Rev.*, **131**, 634-642, 2003.
- Anderson, J. L.: An ensemble adjustment Kalman filter for data assimilation. *Mon. Wea. Rev.*, **129**, 2884-2903, 2001.
- Barker, D. M., W. Huang, Y.-R. Guo, and Q. N. Xiao: A three-dimensional variational (3DVAR) data assimilation system for use with MM5: Implementation and initial results. *Mon. Wea. Rev.*, **132**, 897-914, 2004.
- Wu, W. -S., R. J. Purser, and D. F. Parrish: Three-dimensional variational analysis with spatially inhomogeneous covariances. *Mon. Wea. Rev.*, **130**, 2905-2916, 2002.