P 1.86 A COMPARISON OF THE HYBRID ENSEMBLE TRANSFORM KALMAN FILTER (ETKF)-3DVAR AND THE PURE ENSEMBLE SQUARE ROOT FILTER (EnSRF) ANALYSIS SCHEMES

Xuguang Wang*, Thomas M. Hamill, Jeffrey S. Whitaker
NOAA/CIRES Climate Diagnostics Center, Boulder, CO
Craig H. Bishop
Naval Research Laboratory, Monterey, CA

1. INTRODUCTION

Variational data analysis schemes, 3DVAR (e.g., Parrish and Derber 1992) and 4DVAR (e.g., Courtier et al. 1994). have dominated the operational atmospheric data assimilation for numerical weather prediction. Both 3DVAR and 4DVAR assimilations generally at each update cycle start with a crude background error covariance model which is *homogeneous, isotropic* and *stationary*.

Recently many studies have experimented with ensemble-based data assimilation schemes (e.g., Evensen 1994; Houtekamer and Mitchell 1998, 2001; Anderson 2001; Bishop et al. 2001; Whitaker and Hamill 2002; Ott et al. 2003; Snyder and Zhang 2003). Most of such experiments are based on a method known as the ensemble Kalman filter (EnKF). One potential advantage of the EnKF-based technique is that it provides a feasible way to produce *flow-dependent* estimate of the forecast error covariance. Another potential advantage is that the processes of ensemble forecast and data assimilation are unified. Tests of EnKF in real operational environment are still in progress. Recent experiments in an environment close to operations show that the ensemble square root filter (EnSRF) (Whitaker and Hamill 2005) and the EnKF with perturbed observations (EnKFPO) (Houtekamer et al. 2005), have comparable results with the operational 3DVAR.

With further comparisons of fully developed and tuned EnKF with 3DVAR still underway and given that the computational cost for implementing the full ensemble filter is still a concern for operational centers, it would be appealing to have an intermediate algorithm that can benefit from the flow-dependent information that ensembles provide and also can easily fit in the current operational 3DVAR framework without incurring much extra cost.

*Corresponding author address: Xuguang Wang, NOAA/CIRES Climate Diagnostics Center, R/CDC1, Boulder, CO, 80305. Email: xuguang.wang@noaa.gov

As discussed in this study, we envisage a hybrid data assimilation technique (Hamill and Snyder 2000; Etherton and Bishop 2004; Lorenc 2003; Buehner 2004) potentially has such attractive properties. In particular we study the hybrid ensemble transform Kalman filter (ETKF)-3DVAR scheme and explore its skills by comparing with the EnSRF, one of the intensively tested full ensemble filters.

2. THE HYBRID ETKF-3DVAR AND THE EnSRF ANALYSIS SCHEMES

2.1 The hybrid ETKF-3DVAR scheme

Figure 1 illustrates how the hybrid ETKF-3DVAR data assimilation cycle works. We start with an ensemble of K background forecasts at time t_0. We then repeat the following four steps for each data assimilation cycle. 1) Update the ensemble mean by the hybrid ETKF-3DVAR background error covariance, where the background error covariance is approximated by a linear combination of the sample covariance matrix of the ETKF forecast ensemble and the 3DVAR covariance matrix. 2) Update the forecast perturbations by the ETKF transformation matrix (Bishop et al. 2001; Wang and Bishop 2003; Wang et al. 2004) with no 3DVAR covariance involved. 3) Add the updated ensemble perturbations to the updated ensemble mean to generate K initial ensemble members. 4) Make K forecasts starting from the K initial ensemble members to the next analysis time.

2.2 The EnSRF analysis scheme

The EnSRF was first explored in Whitaker and Hamill (2002). Here we focus on the operational difference between the EnSRF and the hybrid ETKF-3DVAR schemes. In the hybrid ETKF-3DVAR, observations are assimilated simultaneously for the ensemble mean, and the perturbation is updated by the ETKF

transformation matrix at once. The EnSRF assimilates observations serially for each member. For the hybrid ETKF-3DVAR, since the ensemble mean update can be solved variationally (Lorenc 2003; Buhner 2004) and the ensemble perturbations are adjusted inexpensively with the ETKF transformation matrix in the ensemble subspace, the computational cost of the hybrid is expected not as expensive as in the EnSRF especially with the a huge number of observations. Unlike the hybrid ETKF-3DVAR where the background error covariance is estimated partly from the ensemble and partly from the 3DVAR covariance, the background-error covariance of the EnSRF is estimated fully from the ensemble and to avoid the problem of filter divergence due to sampling error, the background error covariance model is modified by a process called "covariance localization".

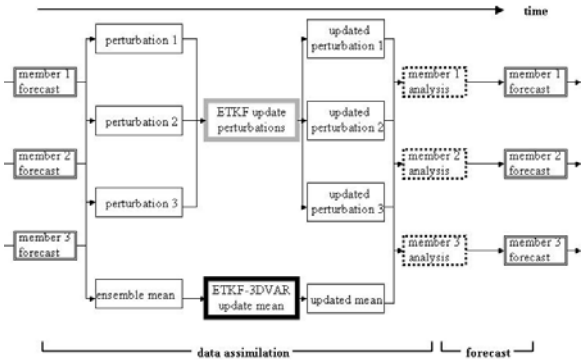

Figure 1. Illustration of the hybrid ETKF-3DVAR analysis and ensemble generation cycle for a hypothetical three-member ensemble.

3. EXPERIMENT DESIGN

3.1. Model, observation and ensemble configuration

In this study, we run a dry, global, two-layer primitive equation model (Zou et al. 1993). The model is spectral and the model state vector includes spectral coefficients of vorticity, divergence at two levels, and the layer thickness, $\Delta\pi$, where π is the Exner function. There is a simple, zonal wavenumber 2 terrain. The parameters chosen are the same as in Hamill and Whitaker (2004). The model is run at T31 resolution with the error doubling time about 4 days. . Perfect model assumption is made in the following experiment.

The observations of interface π and surface π, 362 of each, are taken at a set of nearly equally spaced locations on a spherical geodesic grid (Fig. 2). The observations consisted of the T31 true state plus errors drawn from a distribution with zero mean and standard deviation of 8.75 $Jkg^{-1}K^{-1}$ and 0.875 $Jkg^{-1}K^{-1}$ for interface π and surface π respectively. The number chosen are about a quarter of the globally averaged climatological spread of the nature runs. Observation errors are assumed to be independent spatially and temporally and observations are assimilated every 24h.

In all the experiments, the ensemble size is 50 members. The ensemble was initialized with random draws from the model climatology. The data assimilation is conducted for 150-day period and the statistics is accumulated over the last 100 days. The online estimate of the inflation factor is based on running previous two weeks' innovation statistics (Wang and Bishop 2003; Wang et al. 2004; Etherton and Bishop 2004).

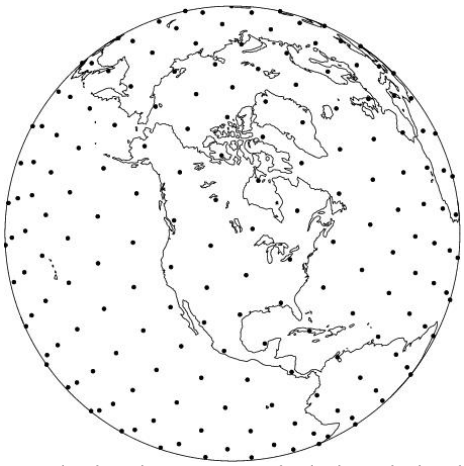

Fig. 2 Observation locations on 362 spherical geodesic grids.

3.2. 3DVAR error covariance

To simulate the static 3DVAR background error covariance. We first collect actual background error samples over time and then calculate the sample covariances. To avoid error divergence due to the limited sample size of the background error collected, covariance localization is applied. Also following Etherton and Bishop (2004) the maximal likelihood theory (Dee 1995) is applied to make online estimate of the magnitude for the 3DVAR error variance.

4. RESULTS

4.1. Analysis error

Fig. 3 shows the rms analysis error in the norm of Kinetic energy, second layer thickness $\Delta\pi_2$, and the surface pressure π. The black bars correspond to the results of the hybrid ETKF-3DVAR as a function of the weighting coefficient, α. The grey bars correspond to the rms analysis errors of the EnSRF with respect to different covariance localization length scale. The white bar is the result for the 3DVAR. The optimal linear combination coefficient for the hybrid ETKF-3DVAR is 0.4 for all three norms. The localization scale that produces the smallest rms for the EnSRF in this experiment is about 15000km ~ 25000km. In all three measures, the best performance of the hybrid ETKF-3DVAR is only slightly worse than that of the EnSRF. Quantitatively, 90% of the improvement of the EnSRF over the 3DVAR is achieved by the hybrid ETKF-3DVAR.

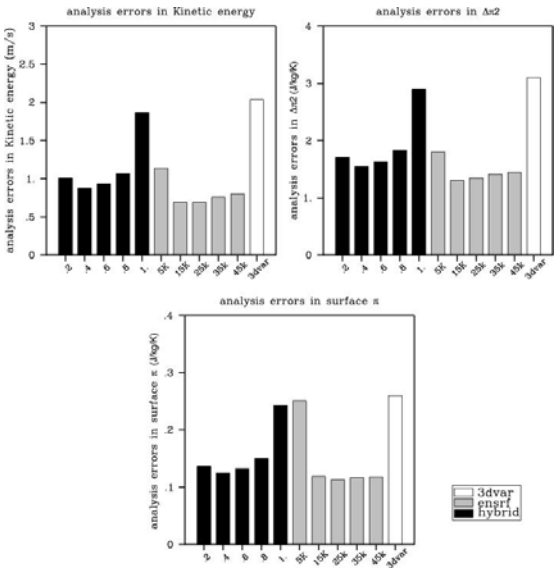

Fig. 3 Root mean square analysis error for Kinetic energy norm, second-layer thickness $\Delta\pi_2$ norm and surface π norm. The black bars are results for the hybrid ETKF-3DVAR scheme with different linear combination coefficient α = 0.2, 0.4, 0.6, 0.8, 1.0. The grey bars are results for the EnSRF with different covariance localization scales, 5000km, 15000km, 25000km, 35000km, 45000km. The white bar is for the 3DVAR.

4.2. Flow dependent covariance model

To demonstrate that with the ensemble covariance incorporated, the background error covariance is flow-dependent, we plot the analysis increment associated with assimilating a single observation. For illustration, we conduct the experiment based on the background ensemble at the 100th data assimilation cycle from each experiment above. Fig. 4 and Fig. 5 show the result for the hybrid ETKF-3DVAR with the optimal coefficient $\alpha = 0.4$ and EnSRF with the optimal covariance localization scale 25000km. The contours are the background ensemble mean for the 100th cycle. The shades are analysis increment by assimilating one interface π observation with the value of $3J/kg/K$ at $(47^\circ N, 108^\circ W)$. Both increments from the hybrid and the EnSRF follow the flow pattern at the 100th cycle and the wind increment is dynamically consistent with the interface height increment.

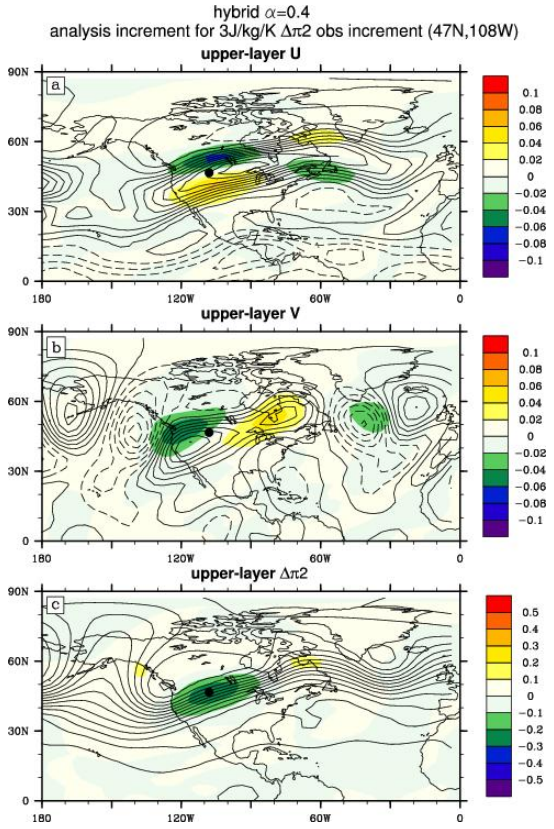

Fig. 4 A snap shot (the 100th analysis cycle) of the ensemble mean upper layer wind (U, V) and thickness ($\Delta\pi_2$) increments for single 3J/kg/K interface π observation increment at (47N, 108W) for the ETKF-3DVAR scheme with $\alpha = 0.4$. The black dot is the observation location. The contours are the background ensemble mean at the 100th cycle and the shades are the single observation increment.

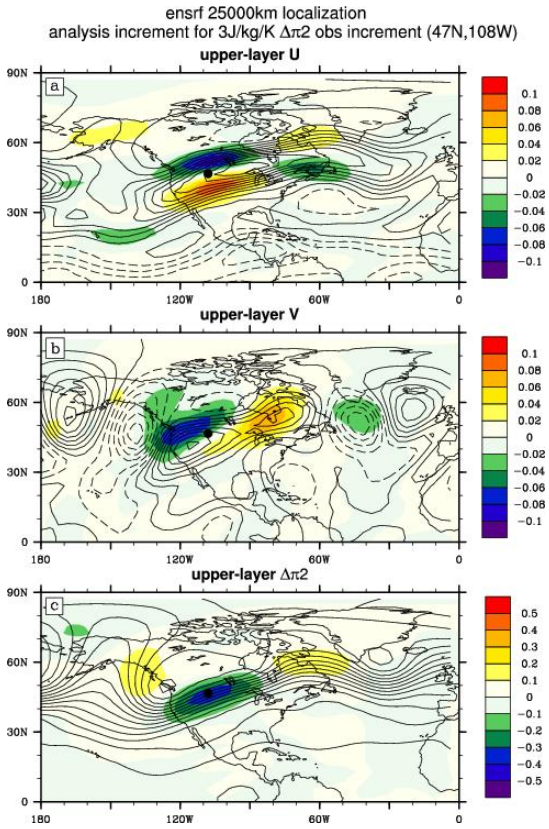

Fig. 5 Same as Fig. 4 except for the EnSRF with localization scale of 25000km.

4.3. Maximal perturbation growth in Kinetic energy norm

For using an ensemble to represent the background error covariance, one would prefer an ensemble that is able to reliably identify forecasts which is very likely to have large forecast errors. Because rapid amplification of analysis error can lead to large forecast error, it is desirable for the ensemble subspace to represent possible fast-growing analysis errors. We measure such an ability of an ensemble by calculating the fastest growth during the first 24h within the ensemble subspace under a certain norm.

Fig. 6 shows the 24h maximal growth in Kinetic energy norm within the ETKF and EnSRF ensemble subspaces. For the ETKF, the maximal growth corresponding to the optimal performance in rms analysis error measure (Fig. 3 $\alpha = 0.4$) is larger than that of the EnSRF (Fig. 3 localization 15000km~25000km). While the maximal growth for the ETKF with different linear combination coefficients is similar, the maximal growth of the EnSRF varies with the localization scale applied. In general the maximal growth decreases with more severe localization. It is speculated that the slower growth for the more severe localization may be due to the imbalanced perturbations induced by severe localization (e.g., Lorenz 2003).

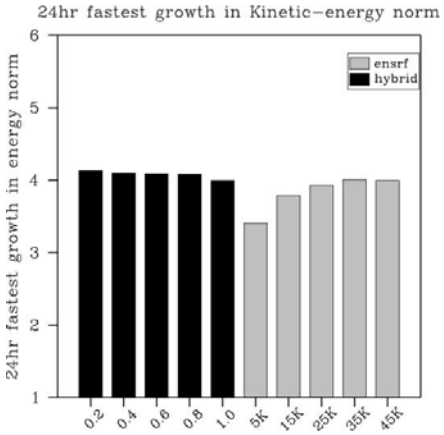

Fig. 6 Maximal 24hr perturbation growth in Kinetic energy norm within the ensemble perturbation subspace for the hybrid ETKF-3DVAR (black bars) with different linear combination coefficients α = 0.2, 0.4, 0.6, 0.8, 1.0 and the EnSRF (grey bars) with different covariance localization scales (5000km, 15000km, 25000km, 35000km, 45000km).

4.4. Balance

In operational data assimilation system, initialization is necessary because otherwise imbalance between mass, momentum and diabatic heating in the analysis can produce large-amplitude gravity waves. As discussed by Mitchell et al. (2002) and Lorenc (2003), the approximation adopted in EnKF, such as the covariance localization by Schur product can introduce imbalance, which compromises the benefits of avoiding filter divergence and bringing more accurate analysis. To evaluate the degree of balance/imbalance for an analysis generated by a data assimilation scheme, a useful global measure is the mean absolute tendency of surface pressure (Lynch and Huang 1992). A quantity equivalent to the surface pressure in the two-layer primitive equation model is the surface Exner function. Fig. 7 shows the mean absolute tendency of the hourly surface π averaged over global grids, all ensemble members, all time, for the first 24hr period. As expected, for the EnSRF, compared to the true tendency, imbalance is more substantial when the localization is more severe. For the hybrid, blending the ETKF ensemble covariance with no localization, the initial analyses are much more balanced than when the background covariance is purely from the 3DVAR. The

analysis ensemble for the hybrid with the optimal coefficient in rms analysis error measure ($\alpha = 0.4$) is more balanced than the EnSRF with optimal localization (15000km~25000km). This result suggests that it is more necessary for the EnSRF to have an explicit initialization step or to use a forecast with extra damping of gravity waves than the hybrid ETKF-3DVAR.

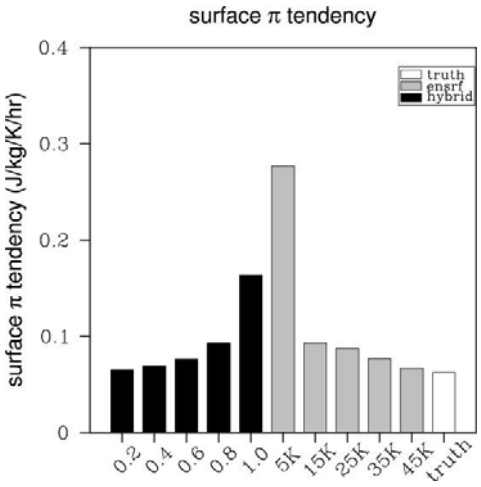

Fig. 7 Mean absolute surface π tendency (J/kg/K/hr) averaged globally, over the first 24hr period, and all ensemble members, for the first 24h. The black bars are for the hybrid ETKF-3DVAR with different linear combination coefficients α = 0.2, 0.4, 0.6, 0.8, 1.0 and the grey bars are for the EnSRF with different covariance localization scales (5000km, 15000km, 25000km, 35000km, 45000km). The white is for the truth.

4.5. Skill of ensemble spread

To measure the skill of the ensemble spread we use a method similar to Wang and Bishop (2003). We first consider a scatterplot of points whose ordinate and abscissa is respectively given by the squared analysis error and the analysis ensemble variance for a particular variable of interest. We collect such points over global grids and for all time samples. We then divide these points into four equally populated bins, arranged in order of increasing ensemble variance. Then we average the squared analysis error and analysis ensemble variance for each bin and take the square root. Connecting the points then yields a curve describing the relationship between the analysis ensemble spread and the rms analysis error. Fig. 8 shows an example of such curves of surface π for the hybrid with $\alpha = 0.4$ and the EnSRF with localization length scale equal to 15000km. There are two aspects of the curve that we are interested to exam. First, further averaging the four points we find that both schemes' ensemble spread is approximately equal to the rms analysis error, which means the ensemble spread is reliable. Second, how well can the ensemble spread distinguish an analysis with large error from that with small error, i.e., how precise is the ensemble spread? For a perfect ensemble, the curve constructed in the way above should follow a reference line with 45 degrees. Fig. 8 shows that the ensemble spread of both schemes are sub-optimal. It is negatively biased for small analysis error and positively biased for large analysis error. The extent to which the ensemble spread is biased can be seen by how much rotation it needs to fall parallel with the reference line. From Fig. 8, the hybrid and the EnSRF have similar skill in ensemble spread precision.

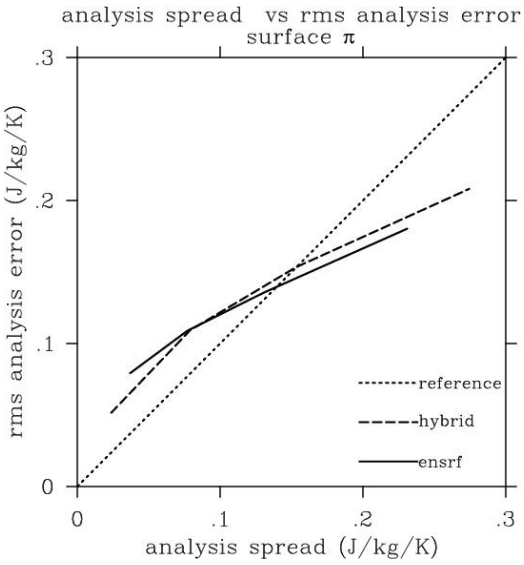

Fig. 8 The relationship between the surface π analysis spread and rms analysis error for the hybrid ETKF-3DVAR (dashed) and the EnSRF (solid). The dotted line is a reference line for a perfect ensemble.

5. CONCLUSIONS

In this paper, we compare the skill of the hybrid ETKF-3DVAR and the full EnSRF analysis schemes with an observation system simulation experiment run with the two-layer primitive equation model with a perfect model assumption. The hybrid ETKF-3DVAR updates the mean with a linear combination of the ensemble and 3DVAR covariance, which can fit in the current variational framework easily (Lorenc 2003; Buehner 2004). The ensemble perturbation is updated and generated by the ETKF, which is computationally

efficient (e.g., Bishop et al. 2001). The EnSRF (Whitaker and Hamill 2002) on the other hand runs relatively expensive parallel data assimilation for each ensemble member and costs scale with the number of observations as serial assimilation of observation is used. Experiments in this paper show promising results for the hybrid ETKF-3DVAR as the hybrid can achieve a large portion (90%) of the EnSRF's analysis improvement over the 3DVAR. The ensemble spread skill for the hybrid is similar to the EnSRF. To avoid filter divergence, the covariance localization is necessary for the EnSRF. For the hybrid ETKF-3DVAR experiment, no covariance localization is required as blending in the full rank 3DVAR background error covariance stabilizes the analysis cycles. Experiments show that applying inappropriate localization can induce imbalanced states and thus degrades the maximal growth in the ensemble perturbation subspace. The hybrid ETKF-3DVAR can be taken as an encouraging alternative method to improve the current variational data assimilation by incorporating ensemble information.

References

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

Bishop, C. H., and Z. Toth, 1999: Ensemble transformation and adaptive observations. *J. Atmos. Sci.*, **56**, 1748-1765.

—, B. J. Etherton, and S. J. Majumdar, 2001: Adaptive sampling with the ensemble transform Kalman filter. Part I: Theoretical Aspects. *Mon. Wea. Rev.*, **129**, 420-436.

Buehner, M., 2004: Ensemble-derived stationary and flow-dependent background error covariances: evaluation in a quasi-operational NWP setting. *Quart. J. Roy. Meteor. Soc.*, **128**, 1-999.

Courtier, P., J. N. Thepaut, and A. Hollingsworth, 1994: A strategy for operational implementation of 4D-VAR, using an incremental approach. *Quart. J. Roy. Meteor. Soc.*, **120**, 1367-1387.

Dee, D. P., 1995: On-line estimation of error covariance parameters for atmospheric data assimilation. *Mon. Wea. Rev*., **123**, 1128-1196.

Etherton, B. J. and C. H. Bishop, 2004: Resilience of Hybrid Ensemble/3DVAR Analysis Schemes to Model Error and Ensemble Covariance Error. *Mon. Wea. Rev*., **132**, 1065-1080.

Evensen, G., 1994: Sequential data assimilation with a nonlinear quasigeostrophic model using Monte Carlo methods to forecast error statistics. *J. Geophys. Res*., **99(C5)**, 10143-10162.

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

Hamill, T. M. and C. Snyder, 2000: A hybrid ensemble Kalman filter-3D variational analysis scheme. *Mon. Wea. Rev.,* **128**, 2905-2919.

—, J. S. Whitaker, and C. Snyder, 2001: Distance-dependent filtering of background error covariance estimates in an ensemble Kalman filter. *Mon. Wea. Rev*., **129**, 2776-2790.

—, and —, 2004: Accounting for the error due to unresolved scales in ensemble data assimilation: a comparison of different approaches *Mon. Wea. Rev*., accepted.

Houtekamer, P. L., and H. L. Mitchell, 1998: Data assimilation using an ensemble Kalman filter technique. *Mon. Wea. Rev.*, **126**, 796-811.

—, and —, 2001: A sequential ensemble Kalman filter for atmospheric data assimilation. *Mon. Wea. Rev.*, **129**, 123-137.

—, H. L. Mitchell, G. Pellerin, M. Buehner and M. Charron, 2005: Atmospheric data assimilation with an ensemble Kalman filter: results with real observations. *Mon. Wea. Rev*., **133**, 604-620.

Lorenc, A. C., 2003: The potential of the ensemble Kalman filter for NWP – a comparison with 4D-VAR. *Quart. J. Roy. Meteor. Soc.*, **129**, 3183-3203.

Lynch, P. and X.-Y. Huang, 1992: Initialization of the HIRLAM model using a digitial filter. *Mon. Wea. Rev*., **120**, 1019-1034.

Mitchell, H. L. and P. L. Houtekamer, 2002: Ensemble size, balance, and modle-error representation in an ensemble Kalman filter. *Mon. Wea. Rev*., **130**, 2791-2808.

Ott, E., B. R. Hunt, I. Szunyogh, A. V. Zimin, E. J. Kostelich, J. Corazza, E. Kalnay, D. J. Patil and J. A. Yorke, 2004: A local ensemble Kalman filter for atmospheric data assimilation. *Tellus*, **56 A**, 415-428.

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

Snyder, C., and F. Zhang, 2003: Assimilation of simulated Doppler radar observations with an ensemble Kalman filter. *Mon. Wea. Rev*., **131**, 1663-1677.

Whitaker, J. S. and T. M. Hamill, 2002: Ensemble data assimilation without perturbed observations. *Mon. Wea. Rev.*, **130**, 1913-1924.

— and —, 2005: Ensemble data assimilation with the NCEP GFS. Thorpex meeting proceeding.

Wang, X. C. H. Bishop, 2003: A comparison of breeding and ensemble transform Kalman filter ensemble forecast schemes. *J. Atmos. Sci*., **60**, 1140-1158.

—, —, and S.J. Julier, 2004: Which is better, an ensemble of positive-negative pairs or a centered spherical simplex ensemble? *Mon. Wea. Rev*., **132**, 1590-1605.

Zou, X., A. Barcilon, I. M. Navon, J. Whitaker, and D. G. Cacuci, 1993: An adjoing sensitivity study of blocking in a two-layer isentropic model. *Mon. Wea. Rev*., **121**, 2833-2857.