Q. J. R. Meteorol. Soc. (2004), **128**, pp. 1–999 doi: 10.1256/qj.04.n

Ensemble-derived stationary and flow-dependent background error covariances: Evaluation in a quasi-operational NWP setting

By Mark Buehner*

Data Assimilation and Satellite Meteorology Division, Meteorological Service of Canada, Canada

(Received February 2004)

SUMMARY

In this study several approaches for obtaining more accurate background error covariances for atmospheric data assimilation are evaluated. Experiments are conducted by replacing the covariances in the operational three-dimensional variational analysis system at the Canadian Meteorological Centre. In the current system, these covariances are computed using the so-called NMC method that is known to suffer from several deficiencies. The approaches evaluated in this study attempt to more realistically sample the probability distribution of background error by simulating (using a Monte Carlo approach) the error generated at each stage of the forecast-analysis process. The ensemble Kalman filter and a simpler approach applied to an existing forecast-analysis system are both used to generate these error samples. In addition, error samples are generated directly from the covariances of the operational system to allow the effects of sampling error to be quantified. Several strategies for estimating the full covariance matrix from a relatively small number of error samples are then employed. Approaches include the use of a spatially localized ensemble representation of the correlations that allows the usual assumptions of homogeneity and isotropy to be relaxed. In addition, the use of a weighted average between such a covariance matrix and a covariance matrix with homogeneous and isotropic correlations is evaluated. Several diagnostic results from the estimated background error covariances are presented in addition to verification statistics computed from two week forecast-analysis experiments. Modest forecast improvements are obtained by using the new background error covariance estimates, mostly in the southern hemisphere. However, additional results suggest that further improvements may be gained by increasing the number of error samples and a preliminary quantitative estimate of the expected gain is computed.

KEYWORDS: 3D-Var Correlation localization Ensemble Kalman filter

1. INTRODUCTION

Most data assimilation schemes used for numerical weather prediction (NWP) rely on a short-term forecast as a so-called background state. The purpose of the data assimilation procedure is to compute an "optimal" correction to this background state (referred to as the analysis increment) using the observations and estimates of the uncertainty associated with the background state and the observations. The uncertainty is typically characterized by covariance matrices for the error in the background state and the observations. These covariance matrices determine the level of influence each observation has on the analysis and how this influence is distributed both spatially and among the different types of analysis variables. Most schemes are derived from statistical estimation theory and produce an analysis increment that is optimal in the sense that it gives the analysis with the smallest error variance when a set of assumptions is satisfied. Critical among these assumptions is that the errors in the background state and observations are Gaussian with zero bias and precisely known covariances. The accurate specification of both covariance matrices represents a significant challenge in the field of NWP.

Currently, the background error covariances used for NWP are known to suffer from several deficiencies. The estimation of these covariances is a difficult problem due to both a lack of knowledge of the statistical properties of background error and the computational challenge of estimating the full covariance matrix of a random vector containing at least $O(10^6)$ elements. Partly to overcome these difficulties, it is often assumed that the correlations of background error are homogeneous and isotropic. While

* Corresponding author: Meteorological Service of Canada, 2121 Trans-Canada Hwy, Dorval, Quebec, H9P 1J3, Canada. e-mail: mark.buehner@ec.gc.ca

this assumption can be justified to some extent on a theoretical basis (Daley 1991), there is evidence that such correlations are not always appropriate, especially in baroclinically unstable conditions. Several approaches for efficiently representing the background error correlations without these assumptions have been proposed in the literature (e.g. Derber and Rosati 1989; Desroziers 1997; Weaver and Courtier 2001; Wu *et al.* 2002). To be practical, any such representation must involve a significant decrease in the number of parameters required to define the covariances. These parameters are then typically estimated from either a statistical study of the differences between forecasts and observations (Hollingsworth and Lönnberg 1986) or an ad hoc method such as the so-called NMC method described later (Parrish and Derber 1992). Another approach is the ensemble Kalman filter (EnKF), originally proposed by Evensen (1994). This approach is based on conducting Monte Carlo simulations to generate a set of random samples for the analysis and background state probability distributions. The covariances are then estimated from these random samples. However, additional assumptions must be introduced to reduce the sampling error that results from the necessity of using a relatively small sample size (Houtekamer and Mitchell 2001; Hamill and Whitaker 2001).

The goal of this study is to examine and evaluate several approaches for estimating the covariances of background error for use in variational data assimilation systems. Experiments were conducted to evaluate covariance matrices that differ according to two aspects:

1. the approach used to generate the ensemble of random samples of background error; and
2. the set of supplementary assumptions employed to obtain a useful estimate of the full covariance matrix from only $O(10^2)$ error samples.

While several approaches have been proposed and evaluated in various contexts in the literature, this study focuses on two approaches for generating samples and three strategies for estimating the full covariance matrix from the samples. The resulting six new types of covariance matrices are used to assimilate a complete set of meteorological observations and the results are compared with those obtained with the current operational system. Both approaches for generating the error samples rely on Monte Carlo simulations of the forecast-analysis process. The first approach is to produce samples of the stationary component of the background error distribution by perturbing the existing forecast-analysis system, similar to Fisher (1999). For the second approach, ensembles of background states are extracted from the EnKF of Houtekamer *et al.* (2003) to provide samples from the flow-dependent background error distribution. The first strategy for estimating the full covariance matrix from error samples employs the same set of assumptions used for the operational system, including the constraint that the correlations be homogeneous and isotropic in the horizontal (Gauthier *et al.* 1998). To avoid these constraints, which are unrealistic for many situations, a more general ensemble representation of the correlations is used as part of the second strategy. The third strategy is to form hybrid covariances by using a weighted average of the covariances estimated using the first two strategies, similar to Hamill and Snyder (2000). Examination of these six types of covariance matrices allows the importance of using flow-dependent covariances to be evaluated when using either homogeneous and isotropic correlations or a more general ensemble representation of the correlations. Similarly, the impact of using different strategies for estimating the full covariance matrix can be independently evaluated for stationary and flow-dependent error samples. Additional experiments are conducted to quantify the impact of the sampling error that results from estimating the background

error covariances from a small set of error samples. For these experiments the error samples are generated by randomly sampling the multi-variate Gaussian distribution defined by the full-rank covariances of the operational system.

The following section provides details on the role of background error covariances in data assimilation and outlines the approaches considered for both sampling background error and estimating the full covariance matrix using a small set of error samples. In Section 3, experimental details are given. Several diagnostics of the estimated background error covariances and results from analysis experiments are given in Section 4. Section 5 provides results from a series of realistic forecast-analysis experiments that span a two week period. Finally, conclusions are given in Section 6.

2. ESTIMATION OF BACKGROUND ERROR COVARIANCES

(*a*) *Role of background error covariances in data assimilation*

The importance of the background error covariances can be seen by examining the linear analysis equation (see e.g. Gelb 1974)

$$\begin{aligned} \Delta\mathbf{x} &= \mathbf{B}\mathbf{H}^T \left\{\mathbf{H}\mathbf{B}\mathbf{H}^T + \mathbf{R}\right\}^{-1} \left\{\mathbf{y} - H\left(\mathbf{x}_b\right)\right\} \\ &\equiv \mathbf{K}\left\{\mathbf{y} - H\left(\mathbf{x}_b\right)\right\}, \end{aligned} \tag{1}$$

where $\Delta\mathbf{x}$ is the analysis increment, $\mathbf{B}$ is the background error covariance matrix, $\mathbf{H}$ is the linearized version of the observation operator H that maps the model state vector into the space of the observations, $\mathbf{R}$ is the observation error covariance matrix, $\mathbf{y}$ is a vector containing the observations, $\mathbf{x}_b$ is the background state and the superscript T represents matrix transposition. The product of matrices on the right hand side of Eq. (1) is represented by the Kalman gain matrix, $\mathbf{K}$. If we take the case where only a single observation is assimilated, then the two bracketed quantities in Eq. (1) are scalars and the resulting analysis increment is proportional to $\mathbf{B}\mathbf{H}^T$, where $\mathbf{H}^T$ is a column vector. Consequently, for a given type of observation the spatial and multi-variate structure of the analysis increment depends strongly on the background error covariances.

In the variational approach, the analysis equation (1) is not explicitly solved, but instead the same analysis increment is obtained by minimizing the cost function

$$\begin{aligned} J &= J_b + J_o \\ &= \frac{1}{2}\Delta\mathbf{x}^T\mathbf{B}^{-1}\Delta\mathbf{x} + \frac{1}{2}\left(\mathbf{y}' - \mathbf{H}\Delta\mathbf{x}\right)^T \mathbf{R}^{-1}\left(\mathbf{y}' - \mathbf{H}\Delta\mathbf{x}\right), \end{aligned} \tag{2}$$

where J_b and J_o are the background and observation components of the cost function, respectively, and $\mathbf{y}' = \mathbf{y} - H\left(\mathbf{x}_b\right)$ is the innovation vector. The minimum of J is obtained by providing the value of J and its gradient for a given $\Delta\mathbf{x}$ to an iterative optimization algorithm. The preconditioning strategy used in the operational three-dimensional variational assimilation system (3D-Var) at the Canadian Meteorological Centre (CMC) is only performed with respect to the background component of the cost function. As a result, the square-root of $\mathbf{B}$ and its transpose are required instead of the inverse. The optimization problem is recast in terms of a control vector, $\boldsymbol{\gamma}$, related to the analysis increment according to

$$\Delta\mathbf{x} = \mathbf{B}^{1/2}\boldsymbol{\gamma}. \tag{3}$$

The adjoint of this operation is also required for computing the gradient of the cost function with respect to the control vector. Written in terms of the control vector, the background term of the cost function becomes $J_b = (1/2)\boldsymbol{\gamma}^T\boldsymbol{\gamma}$. Consequently, the

minimization is perfectly preconditioned with respect to the background cost function. If the background error covariance matrix has full rank, the dimension of the control vector can be as large as that of the analysis increment, denoted by N_x.

(*b*) *Approaches for sampling background error*

The exact statistics of the error in the short-term forecast used as the background state are not known since to obtain them would require precise knowledge of the real atmospheric state. However, several approximate approaches have been developed for sampling the background error. The extended Kalman filter provides a means of continually evolving the error covariances of the estimated state in a sequential forecast-analysis system. According to the extended Kalman filter, the background error covariances are given by (see e.g. Gelb 1974)

$$\mathbf{B} = \mathbf{M}\mathbf{P}^a\mathbf{M}^T + \mathbf{Q}, \tag{4}$$

where $\mathbf{M}$ is the linearized forecast model, $\mathbf{P}^a$ is the covariance matrix for the error in the previous analysis and $\mathbf{Q}$ is the model error covariance matrix that accounts for the additional error induced in the forecast by errors in the forecast model. Like the background error itself, an accurate statistical description of the model error is not available and remains a major challenge for all data assimilation approaches (Dee 1995). Assuming that the background and observation error covariances are correctly specified, the analysis error covariances are given by

$$\mathbf{P}^a = \left(\mathbf{I} - \mathbf{K}\mathbf{H}^T\right)\mathbf{B}. \tag{5}$$

For realistic problems, the solution of Eq. (4) and (5) is computationally infeasible due to the high dimensionality of the covariance matrices. Instead of manipulating the full covariance matrices, one common approach is to approximate the probability distributions by an ensemble of random samples drawn from the distribution. The following two approaches for generating such ensembles are examined in this study: (i) a simple Monte Carlo simulation approach applied to an existing data assimilation system (hereafter referred to as the perturbed 3D-Var) and (ii) the ensemble Kalman filter. These two approaches are compared with the NMC method as currently used in the operational system. A brief description of these three approaches follows.

Several NWP centres, including CMC, currently employ variational assimilation systems with stationary background error covariances estimated using the NMC method (Parrish and Derber 1992; Gauthier *et al.* 1998; Rabier *et al.* 1998; Derber and Bouttier 1999). Following this method, the differences between pairs of forecasts valid at the same time, but having different lead times, are taken to be representative of background error. Such forecast differences can easily be computed for a past period using the archived output of an operational forecasting system. At CMC, the differences between 48 and 24 hour forecasts taken over a period of 2-3 months are used. However, a lack of correspondence between these lagged forecast differences and 6 hour forecast error necessitates a tuning of the computed covariances.

The EnKF uses a Monte Carlo simulation strategy to render the standard Kalman filter algorithm feasible for even very high-dimensional systems (Evensen 1994). The analysis, background, model and observation error distributions in Eq. (4) and (5) are represented by an ensemble of random samples. To obtain these samples, a separate forecast-analysis experiment is run for each ensemble member in which the observations and background states are randomly perturbed in a way that is consistent with the specified uncertainties in the observations ($\mathbf{R}$) and the forecast model ($\mathbf{Q}$), respectively

(Burgers *et al.* 1998). The analysis for each ensemble member is performed using background error covariances estimated from the ensemble spread of forecasts valid for that specific analysis time. The approach can be made computationally feasible if the ensemble size is limited to $O(10^2)$. However, for sophisticated NWP models it remains unclear whether this is a sufficient number of error samples to adequately represent the probability distributions. It is also unclear, as already stated, how to specify the model error covariances. An implementation of the EnKF has been refined and applied to increasingly realistic models by a research group at the Canadian centre (Houtekamer and Mitchell 1998, 2001; Mitchell and Houtekamer 2000). The current version of this EnKF uses a complete set of observations and the operational forecast model with a full set of physical parametrisations (Houtekamer *et al.* 2003). The model error covariances are prescribed in the form of an idealized covariance matrix similar to the background error covariances used in the operational 3D-Var with the variances scaled by 0.25. In the present study, forecast ensembles from this system are used to estimate background error covariances for use in the variational assimilation system.

A simpler approach based on Monte Carlo simulation and similar to that described by Houtekamer *et al.* (1996) was recently used to recompute the stationary background error covariances in the variational analysis system at the European Centre for Medium-Range Weather Forecasts (Fisher 1999). Like the EnKF, an ensemble of forecast-analysis experiments are conducted with perturbed observations and background states, but with the analyses performed using prescribed stationary background error covariances. Also, instead of attempting to compute flow-dependent error statistics, the approach is used to estimate the stationary component of the error statistics over a period of several weeks. Due to the pooling of samples over time, only a small number of perturbed forecast-analysis experiments are required in addition to an unperturbed experiment. Differences between the 6 hour forecasts from the perturbed and unperturbed experiments are then computed and used to represent samples of background error. Again, the specification of the model error covariances used to compute perturbations to the background states remains the biggest challenge. For the present study, this approach was implemented using the 3D-Var with the operational background error covariances to perform the analyses. To partially overcome the difficulty of specifying the model error covariances, an adaptive tuning procedure was used. The tuning approach is based on a comparison between the temporally averaged innovation statistics from the unperturbed experiment and the simulated innovations from the perturbed experiment. The approach is simpler than that proposed by Dee (1995) and examined by Mitchell and Houtekamer (2000) in the EnKF context. The simplification is attained by assuming the model error covariances are proportional to the current operational background error covariances and therefore only the scaling factors applied to these covariances must be determined. The scaling factors for wind components and temperature are computed independently for each vertical level and for each of three latitude bands (total of 168 tuned parameters). The tuning procedure guarantees that the horizontally averaged innovation variances simulated by the perturbed experiment equal the true innovation variances. The computation of these scaling factors is outlined in Appendix A.

It should be noted that another class of ensemble-based approaches exists where the appropriate error covariances are obtained without the need to perturb the observations. While the use of perturbed observations leads to an ensemble of analyses with the correct error covariances in a mean sense (Burgers *et al.* 1998; Houtekamer and Mitchell 1998), for small ensembles the observation error distribution will be poorly represented by the perturbations. Approaches that instead use the observation error covariance matrix directly, are referred to generally as ensemble square root filters (Tippett *et al.* 2003).

In these approaches the ensemble of background states are linearly transformed to produce an ensemble of analyses with the correct ensemble spread covariances. This transformation is not unique and several different solutions have been proposed (Bishop *et al.* 2001; Anderson 2001; Whitaker and Hamill 2002).

(*c*) *Strategies for reducing sampling error*

For realistic NWP applications, a series of simplifying assumptions must be employed to accurately estimate the background error covariances from a small number of error samples. In many systems the error correlations are assumed to be stationary and globally homogeneous and isotropic. By making these assumptions, the number of independent parameters in the covariance matrix is vastly reduced and a full-rank matrix can be estimated from relatively few error samples. However, alternative assumptions may be used instead to reduce sampling error while retaining a more realistic representation of the covariances.

In the operational 3D-Var, the background error covariances are constrained to have homogeneous and isotropic correlations for the variables: streamfunction, unbalanced velocity potential, unbalanced temperature, natural logarithm of specific humidity, unbalanced surface pressure and surface skin temperature. In addition, the background error for each of these variables is assumed to be uncorrelated with the others and the correlations between the full analysis variables are modeled with linear balance operators. For example, the temperature increment is constructed by adding the increments of unbalanced and balanced temperature, where the latter is computed from the streamfunction increment via a balance operator for geostrophy. This implies correlations between the streamfunction and temperature increments that are consistent with geostrophy and the hydrostatic relationship. Additional balance operators are employed to create correlations between streamfunction and surface pressure (geostrophy) and between streamfunction and velocity potential near the surface (Ekman balance). Several parameters in the balance operators themselves are statistically estimated from the ensemble of error samples. The NMC method tends to underestimate variances in regions with fewer observations, such as over oceans, leading to unrealistic zonal variations. To address this problem, the background error variances in the operational system are constrained to be zonally invariant and are subjected to a tuning procedure. The complete covariance matrix can be expressed as

$$\mathbf{B} = \mathbf{G}\mathbf{V}^{1/2}\mathbf{C}_u\mathbf{V}^{T/2}\mathbf{G}^T, \tag{6}$$

where $\mathbf{C}_u$ is the correlation matrix for the set of independent variables listed above, $\mathbf{V}$ is a diagonal matrix containing the error variances, and $\mathbf{G}$ transforms the unbalanced variables into the full quantities for temperature, surface pressure, and velocity potential using the balance operators listed above. Due to the preconditioning strategy described in Section 2a, only the square-root of this covariance matrix is required. More details concerning the background error covariances in the operational system are given by Gauthier *et al.* (1998).

Alternatively, the error covariances can be estimated without constraining the correlations to be homogeneous and isotropic. However, correlations estimated directly from a small number of error samples often do not approach zero at long separation distances, but can be unrealistically large even on the opposite side of the globe. To overcome this problem, a procedure for spatially localizing the correlations was proposed by Gaspari and Cohn (1999) and examined in the context of an EnKF by Houtekamer and Mitchell (2001) and Hamill and Whitaker (2001). The technique for

efficiently employing a spatially localized ensemble representation of the background error correlations in a variational assimilation framework is described in Appendix B. Two approaches for utilizing the localized ensemble representation of the correlations were used. The first is to simply replace the correlation matrix in (6) with the localized correlation matrix computed for the independent analysis variables listed above. The second approach is very similar to how the covariances are estimated in the EnKF of Houtekamer *et al.* (2003). In that case no balance operators are used and the within- and between-variable correlations are estimated for: wind components, temperature, natural logarithm of specific humidity, surface pressure and skin temperature. There are two important differences between these approaches: (i) the between-variable correlations are estimated using simplified balance operators in the first and estimated directly from the error samples in the second; and (ii) localization is performed on streamfunction and velocity potential in the first and wind components in the second.

The difficulties with estimating the full correlation matrix from a small ensemble can also be addressed through the addition of a second correlation matrix for which the assumptions of homogeneity and isotropy are imposed. This hybrid approach was used in the context of an EnKF by Hamill and Snyder (2000). In the variational context where preconditioning is performed with respect to the background term, a convenient approach is to combine two such covariance matrices through augmentation of the state vector. The analysis increment is then computed using the relationship

$$\Delta \mathbf{x} = \beta_1 \mathbf{B}_1^{1/2} \boldsymbol{\gamma}_1 + \beta_2 \mathbf{B}_2^{1/2} \boldsymbol{\gamma}_2, \tag{7}$$

where $\mathbf{B}_1$ is the full-rank covariance matrix with homogeneous and isotropic correlations and $\mathbf{B}_2$ is the covariance matrix using a spatially localized ensemble representation of the correlations. The parameters β_i control the relative weight given to each covariance matrix. This approach for combining two matrices to form hybrid covariances provides the greatest flexibility since there is no requirement that one of the covariance matrices has low-rank. Also, the same level of preconditioning is retained as when employing either of the covariance components individually, though the dimension of the control vector is increased. Consequently, the augmented control vector used by the minimization algorithm is

$$\boldsymbol{\gamma} = \begin{bmatrix} \boldsymbol{\gamma}_1 \\ \boldsymbol{\gamma}_2 \end{bmatrix}$$

and the form of the background term of the cost function remains unchanged. It can be shown that this is equivalent to using the non-preconditioned cost function from Eq. (2) where $\mathbf{B}$ is replaced by the hybrid covariance matrix $\left(\beta_1^2 \mathbf{B}_1 + \beta_2^2 \mathbf{B}_2\right)$.

3. DESCRIPTION OF THE EXPERIMENTS

Several types of numerical experiments were performed using 3D-Var to address three primary goals. The first goal is to evaluate the impact of using the perturbed 3D-Var (PERT) or EnKF (ENKF) to generate background error samples versus the NMC method (CNTL). The second goal is to compare background error covariances estimated using either homogeneous and isotropic correlations (-HI), a spatially localized ensemble representation of the correlations (-ENS) or a weighted average of these two representations (-HYB). These two goals are closely related since the two new approaches for generating error samples are each combined with the three strategies for representing the correlations resulting in six types of covariances. The final goal is to estimate the

importance of the sampling error that results from using a finite number of samples to estimate the covariances.

To obtain error samples using the perturbed 3D-Var approach, several forecast-analysis experiments were conducted with the observations and background fields randomly perturbed at each analysis time. In addition, a similar forecast-analysis experiment was performed but without perturbations (CNTL experiment). In the perturbed experiments, the perturbations for the observations and background states were drawn from multi-variate Gaussian distributions with specified observation and model error covariances, respectively. The variances of the model error covariances were adaptively tuned, as described in Section 2b and Appendix A, to ensure that the the differences between the background states from the perturbed and unperturbed experiments are consistent with the forecast error variances estimated from the true innovations. These differences were then extracted and used as the error samples. A total of 153 error samples from three perturbed 3D-Var experiments were used for estimating stationary background error covariances (PERT experiments).

The time period for all data assimilation experiments was 19 May 2002 to 2 June 2002, chosen to coincide with the period used for the most recent EnKF study by Houtekamer *et al.* (2003). This allowed the use of the 128 member forecast ensembles produced for that study for estimating flow-dependent 3D-Var background error covariances (ENKF experiments).

To quantify the impact of the sampling error that results from using a finite size ensemble representation of the correlations, a set of random samples were directly generated from the full-rank background error covariances of the operational 3D-Var. Data assimilation experiments were then performed using background error covariances estimated from these random samples and employing a localized ensemble representation of the correlations (BSAMP experiments).

The combinations of approaches for generating error samples and representing the correlations results in nine different types of background error covariances that are evaluated in this study. A summary of these types of covariances is given in Table 1. The spatial localization procedure described in Section 2c and Appendix B was used for all covariance matrices for which the ensemble representation of the correlations was employed. The horizontal and vertical length scales for the localization vary between the experiments and are therefore given with the results of each experiment.

The background error variances from each covariance matrix are first examined to evaluate how the approach for generating error samples effects both the horizontal structure and the extent to which the wind and mass errors are in geostrophic balance. Next, to illustrate the combined effect of the correlations, the localization procedure and the balance operators in determining the covariances, a series of analysis experiments are performed in which only one observation is assimilated. These experiments illustrate how information from the single observation is spread both spatially and to the other variable types by the background error covariances. This type of experiment is performed to show the effect of applying the different correlation representations to the perturbed 3D-Var error samples. Single observation analysis experiments are also performed to illustrate the ability of the error samples from the EnKF to resolve a highly flow-dependent feature as compared with a typical 4D-Var analysis system. A series of 3D-Var analyses with a complete set of observations is then performed to examine the relationship between sampling error and both the number of error samples and the spatial localization used to estimate the correlations. To provide a more realistic evaluation of the impact of using flow-dependent error covariances with the ensemble representation of the correlations, a single analysis is performed with all available observations for a

TABLE 1. SUMMARY OF BACKGROUND ERROR COVARIANCE MATRICES EVALUATED.

Name	Description
CNTL	Control: covariance matrix from the operational 3D-Var
PERT-HI	Covariances estimated from the perturbed 3D-Var with homogeneous and isotropic correlations and balance operators
PERT-ENS	Same as PERT-HI, but with ensemble representation of the correlations, balance operators used for between-variable correlations
PERT-HYB	Hybrid covariances formed by computing the mean of the correlation matrices used in PERT-HI and PERT-ENS $\left(\beta_1^2 = \beta_2^2 = 0.5\right)$
ENKF-HI	Flow-dependent covariances from EnKF with homogeneous and isotropic correlations and using the balance operators as in PERT-HI
ENKF-ENS	Flow-dependent covariances from EnKF with correlations (including between-variable correlations) estimated directly from the forecast ensembles using an ensemble representation
ENKF-HYB	Hybrid covariances formed by computing the mean of the covariance matrices used in CNTL and ENKF-ENS $\left(\beta_1^2 \beta_2^2 = 0.5\right)$
BSAMP-32	Covariances estimated from 32 random samples generated using the operational 3D-Var covariance matrix, using an ensemble representation of the correlations and between-variables correlations estimated directly from the error samples (no balance operators)
BSAMP-128	Same as BSAMP-32, but using 128 random samples
BSAMP-512	Same as BSAMP-32, but using 512 random samples

Spatial localization is applied in all cases where the ensemble representation is used for the correlations.

particularly active case over the North Pacific. The analysis and subsequent forecast are compared with those from the operational system. Finally, forecast-analysis experiments are performed spanning the period from 19 May 2002 to 2 June 2002. These provide the most realistic evaluation of the overall impact that the new covariances would have in an operational setting. The experiments are used to evaluate and compare the six types of background error covariances obtained using the perturbed 3D-Var or EnKF to generate the error samples and the three strategies for representing the correlations. Forecast-analysis experiments are also conducted using three different ensemble sizes with samples randomly drawn from the multi-variate Gaussian distribution defined by the full-rank covariances of the operational system. These experiments complement the previously mentioned single analysis experiments by providing an evaluation of the effect of sampling error on both the analyses and background states in an operational setting.

4. ESTIMATED BACKGROUND ERROR COVARIANCES: DIAGNOSTICS

(*a*) *Model error tuning in perturbed 3D-Var*

Figure 1 shows the tuning coefficients adaptively computed for the perturbed 3D-Var experiments as described in Section 2b and Appendix A. The coefficients for the wind components (UV) and temperature (T) as a function of both vertical level and

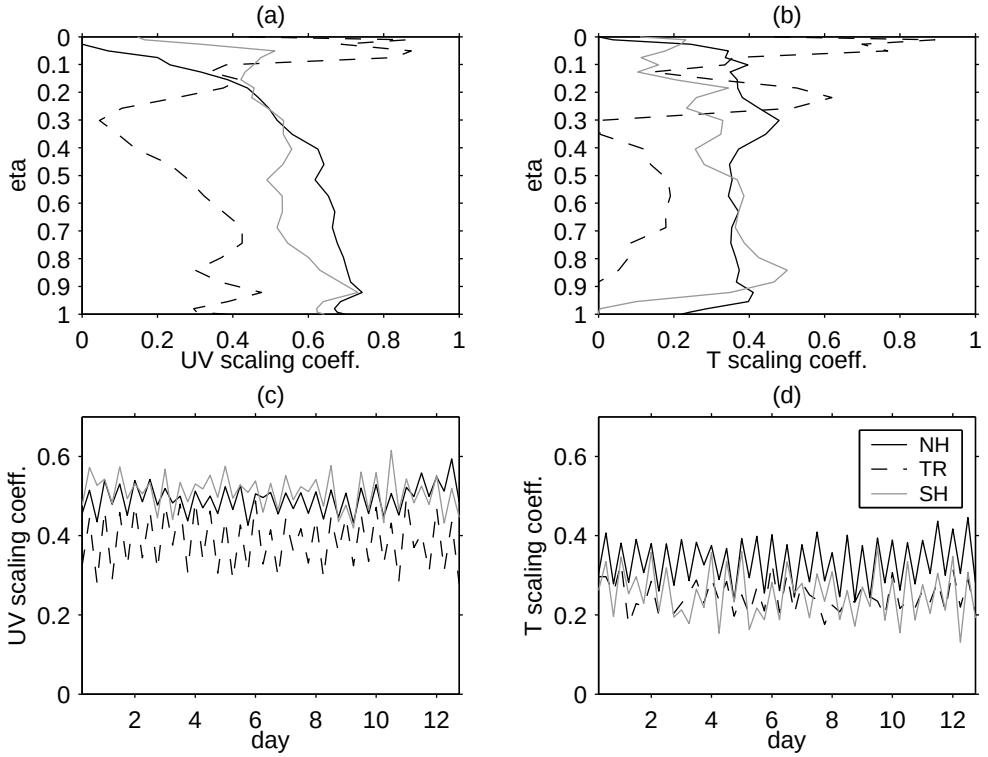

Figure 1. Adaptively tuned scaling coefficients applied to the operational background error covariances to obtain model error covariances for the perturbed 3D-Var experiment. The upper panels show the temporal mean for (a) wind components and (b) temperature as a function of vertical level. Lower panels show the vertical mean for (c) wind components and (d) temperature as a function of time. All panels show the coefficients for the northern extra-tropics, tropics and southern extra-tropics as denoted in the legend. Note that the zero values appearing in the vertical profiles correspond to a situation where a sufficient forecast error variance is simulated by the perturbed 3D-Var experiment without the need for model error perturbations.

time are shown. These coefficients were used to scale the standard deviations (std dev) of the operational background error covariances to obtain the appropriate model error covariances. The vertical structure of the coefficients differs substantially between the extra-tropical and tropical regions. In the extra-tropics the coefficients are largest in the troposphere and decrease above the tropopause. The coefficients for temperature are also reduced near the surface. In the tropics, the coefficients are generally larger above about 250 hPa. Temporally, the coefficients oscillate with values being relatively larger at 06 UTC and 18 UTC than at 00 UTC and 12 UTC. This is likely due to the higher number of radiosonde observations that are assimilated at 00 UTC and 12 UTC. The added observations, even though they are perturbed, act to reduce the simulated analysis error and therefore also decrease the simulated prediction error in the subsequent forecasts at 06 UTC and 18 UTC (before the addition of model error perturbations). Since the true innovation variances used to compute the tuning coefficients do not vary in time, the model error variances are increased to compensate for the decrease in the simulated prediction error.

In Fig. 2 the fraction of the estimated background error variance due to the model error perturbations for each region is shown for wind components and temperature. This shows that the background error for winds is most dominated by model error near the

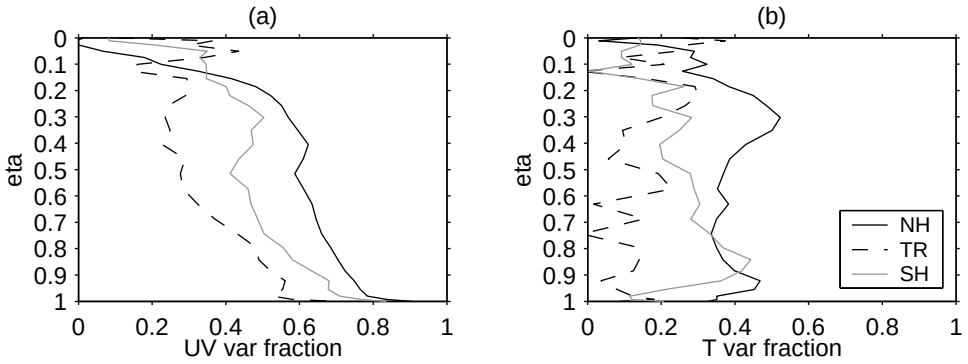

Figure 2. The fraction of the estimated background error variance from the perturbed 3D-Var experiments that is directly due to the model error perturbations for both (a) wind components and (b) temperature.

surface where it accounts for about 80% of the background error variance in the extra-tropics. This large contribution of model error to the surface wind background error may be due to the effects of unresolved orography and an overly diffusive boundary layer parametrisation on the surface winds. In contrast, the minimum contribution of model error is seen for temperature in the tropics where only about 15% of the background error variance is directly related to the addition of model error perturbations. The lower values for the southern extra-tropics relative to the northern extra-tropics may simply be related to larger values of background error std dev in the south (as discussed in the following section) whereas the absolute contribution from model error is more similar between the two regions.

(*b*) *Estimated background error variances*

The linear analysis equation (1) shows that, in a general sense, the ratio between the background error variances and the sum of the background and observation error variances gives an indication of the relative importance of a particular observation in the analysis. The observation error variances in the operational 3D-Var do not depend on horizontal location and therefore the variations in background error variance strongly determine relative differences in how observations will affect the analysis increment. The upper panels of Fig. 3 show the std dev for streamfunction at a level near the extra-tropical jet (left) and temperature in the middle troposphere (right) from the operational background error covariances. The lack of zonal variation is imposed when the covariances are computed from lagged forecast differences as described earlier. Similar std dev fields are shown in the middle panels of Fig. 3 as computed from the perturbed 3D-Var. In this case the zonal variations have been retained. In the bottom two panels of Fig. 3 the std dev is shown as computed by temporally averaging the EnKF background ensemble spread variances between 21 May 2002 and 2 June 2002. When using the error samples from the perturbed 3D-Var or the EnKF the horizontal variations in std dev appear to be realistic with larger values of std dev appearing in areas that are less well observed and more dynamically active such as in southern (winter) hemisphere for both variables and over the oceans for temperature. The higher level of small-scale variability in the perturbed 3D-Var versus the EnKF variances is due to the use of approximately 40 times less error samples from the perturbed 3D-Var. Differences from the two approaches are partly due to differences in the model error variances, where

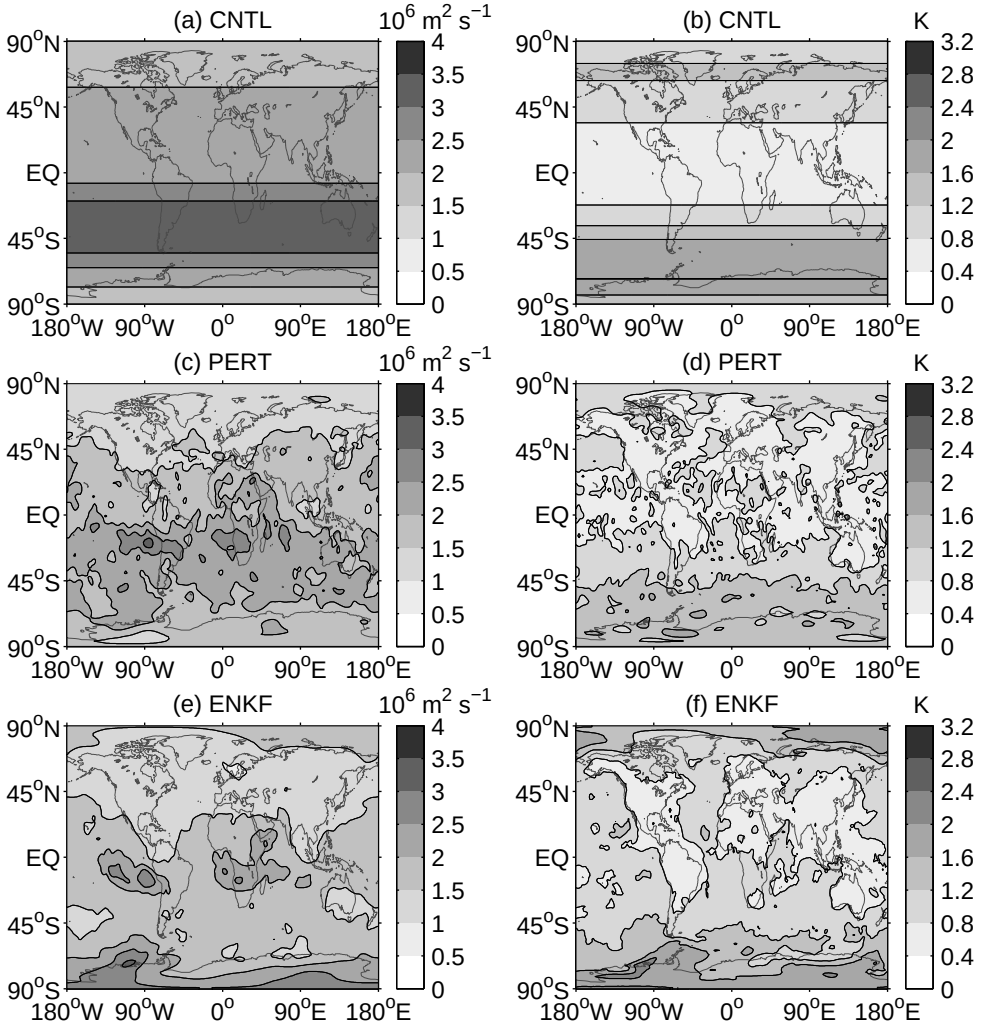

Figure 3. Estimated background error std dev of (a) streamfunction near 250 hPa ($\eta = 0.258$) and (b) temperature near 500 hPa ($\eta = 0.516$) computed using the "NMC method". Similarly, panels (c) and (d) show the same results estimated from the perturbed 3D-Var experiments and panels (e) and (f) show the results estimated by temporally averaging the background ensemble spread variances from the EnKF.

the EnKF uses globally constant model error variances and the perturbed 3D-Var uses an adaptively scaled version of the background error variances from the operational system.

The background error correlations between the mass and wind fields ensure that the analysis increment will be geostrophically balanced to some extent even when only wind or temperature observations are assimilated. Likely as a result of the relatively long 24 and 48 hour integrations used to compute the lagged forecast differences in the NMC method, the operational background error covariances have higher correlations between the mass and wind fields in the extra-tropical troposphere than the covariances computed from either the perturbed 3D-Var or the EnKF. Figure 4 shows the ratio of unbalanced temperature to the full temperature variance for the operational background error covariances (Fig. 4a) and for those computed from the perturbed 3D-Var (Fig. 4b).

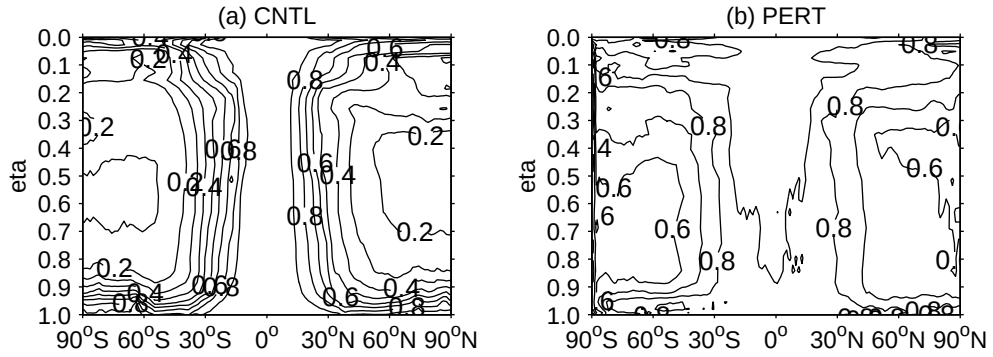

Figure 4. The zonally averaged unbalanced temperature variance normalized by the full temperature variance from (a) the operational background error covariances (CNTL) and (b) the covariances estimated from the perturbed 3D-Var (PERT).

Low values of this ratio correspond with high correlations between the mass and wind fields. The ratio reaches a minimum of 20% for the operational covariances, whereas for the perturbed 3D-Var it only reaches about 60%. A similar reduction in balance was also seen in the covariances computed from the forecast ensembles of the EnKF (not shown) and in the results of Fisher (1999). At this point, it is unclear if the reduced balance is more consistent with the true background error.

(*c*) *Single observation experiments*

The analysis increment resulting from the assimilation of a single observation provides a partial view of the background error covariances by showing how information from the observation is distributed both spatially and among the different analysis variables. From Eq. (1), the analysis increment is proportional to $\mathbf{BH}^T$, where $\mathbf{H}$ is reduced to a row vector representing the observation operator for the single assimilated observation. For observation types closely related to one of the variables represented in the background error covariances, the analysis increment is simply proportional to a column of $\mathbf{B}$.

First, the covariances from the operational 3D-Var are compared with the covariances estimated from the perturbed 3D-Var using the three representations for the background error correlations. In addition, the result of using the ensemble representation of the correlations without spatial localization is shown. The increments computed from assimilating a single zonal wind observation 1 m s^{-1} greater than the background wind over the Pacific Ocean near 250 hPa are shown when using each background error covariance matrix.

With the correlations constrained to be homogeneous and isotropic, the analysis increments are quite similar when the background error samples are generated using either the NMC method (Fig. 5, upper panels) or the perturbed 3D-Var (Fig. 5, lower panels). The horizontal and vertical correlations for wind at 250 hPa are slightly sharper and the geopotential height and wind increments decrease more rapidly north of 45°N when using error samples generated with the perturbed 3D-Var. The geopotential height increment is reduced when using the perturbed 3D-Var error samples, consistent with the reduction in the geostrophic balance discussed earlier. The analysis increment has a more complex structure when the ensemble representation of the correlations without spatial localization is used (Fig. 6, upper panels). Comparing the upper and middle panels of Fig. 6 shows that spatial localization has the expected effect of reducing

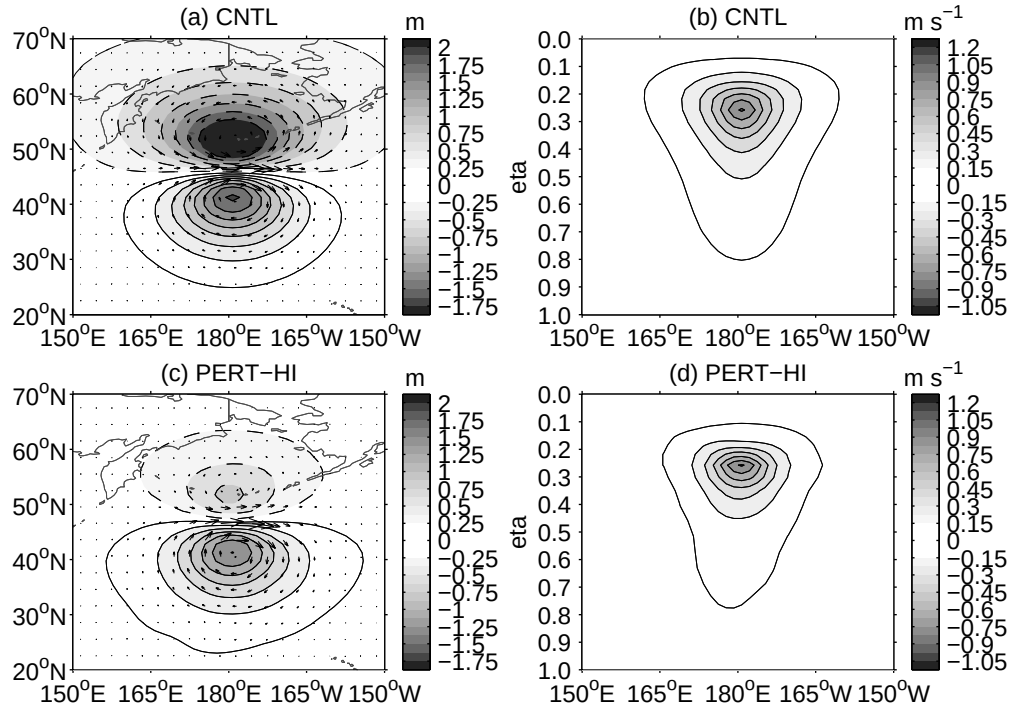

Figure 5. Analysis increment of wind (vectors in left panels and contours in right panels) and geopotential height (contours in left panels) from a single zonal wind observation located at 45° N, 180° E and near 250 hPa when using the CNTL background error covariances (upper panels) and the PERT-HI covariances (lower panels). Panels (a) and (c) show the analysis increment for wind and geopotential height near 250 hPa and panels (b) and (d) show the vertical-zonal cross-section of the wind increment along 45° N. Note that solid contours are used for positive values and dashed for negative.

the analysis increment for both wind and geopotential height the most for grid points furthest from the location of the observation. The analysis increment produced using the localized correlations resembles the increment produced using the homogeneous and isotropic correlations much more than when no spatial localization is applied. Using the mean of the localized ensemble representation and the homogeneous and isotropic representation of the correlations (Fig. 6, bottom panels) further increases the similarity of the analysis increments with those produced with only homogeneous and isotropic correlations while still retaining a degree of anisotropy in the general structure.

In addition, single observation experiments were conducted to evaluate the impact of estimating the between-variable correlations directly from the error samples without the use of balance operators (not shown). With the balance operators, the resulting analysis increments for geopotential height are constrained to be geostrophically balanced with the winds (as seen in the previous figure). Conversely, when the between-variable correlations are estimated directly from the error samples, the resulting geopotential height increment contains more small-scale structure and its contours are often far from being aligned with the wind vectors.

Obviously, the influence of spatially heterogeneous forcing on the correlation structure of background errors can not be resolved when the constraints of homogeneity and isotropy are imposed. However, the effects of such stationary influences as orography can be captured when using the ensemble representation of the correlations even when

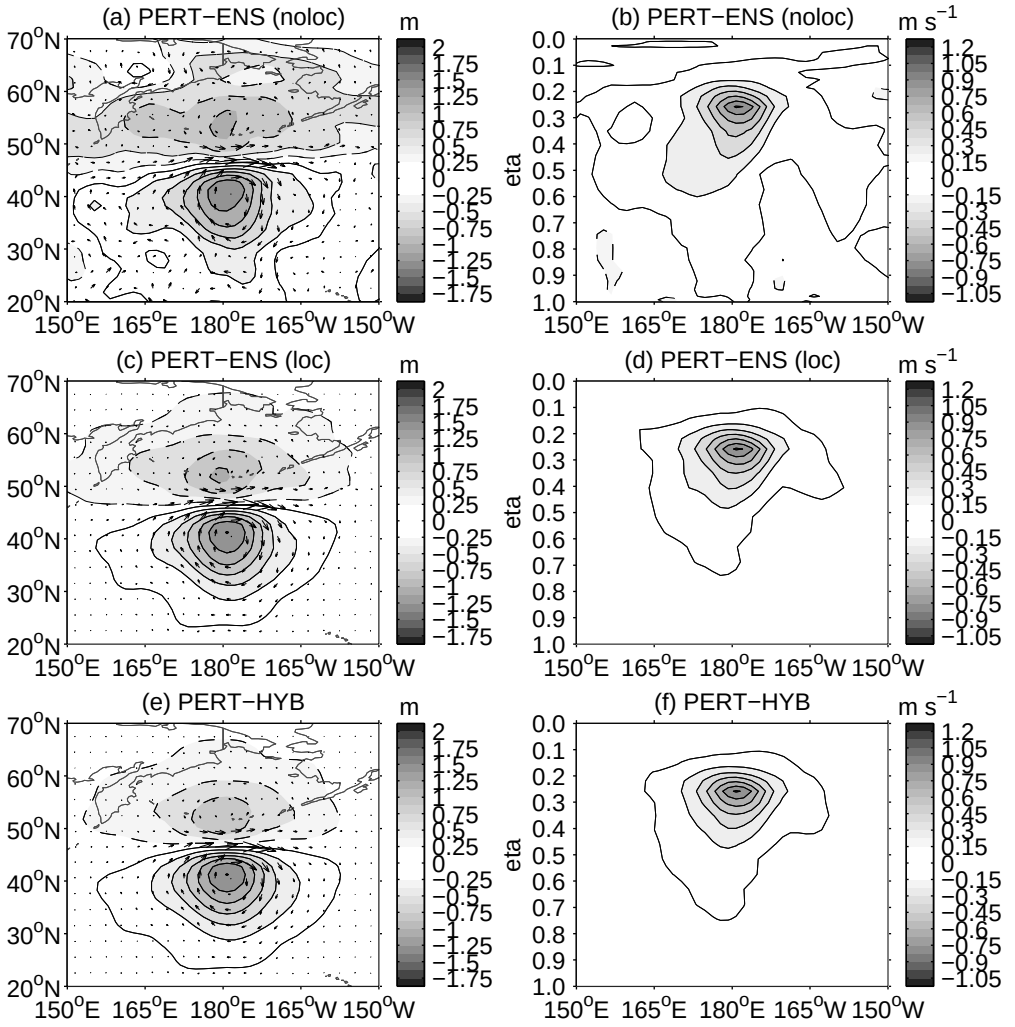

Figure 6. Same as the previous figure, but using the PERT-ENS background error covariances either without spatial localization (upper panels) or with localization (middle panels) and the PERT-HYB covariances (lower panels). Balance operators are used to represent the between-variable correlations.

the error covariances are assumed stationary and therefore estimated from error samples distributed over time. Figure 7 shows the analysis increment from assimilating a single zonal wind observation near 500 hPa over a mountainous region of North America with contours of the land elevation superimposed. When using correlations that are homogeneous and isotropic (Fig. 7a), the analysis increment is clearly not influenced by the presence of the mountains. Conversely, this influence can be seen when the ensemble representation of the correlations is used (Fig. 7b): the geostrophically balanced geopotential height increment exhibits a local increase where the wind increment is directed towards the west over the mountain peaks as can be seen around 50° N and 30° N. This results in a general equatorward deflection of the winds over the ascending slopes and a poleward deflection over the descending slopes.

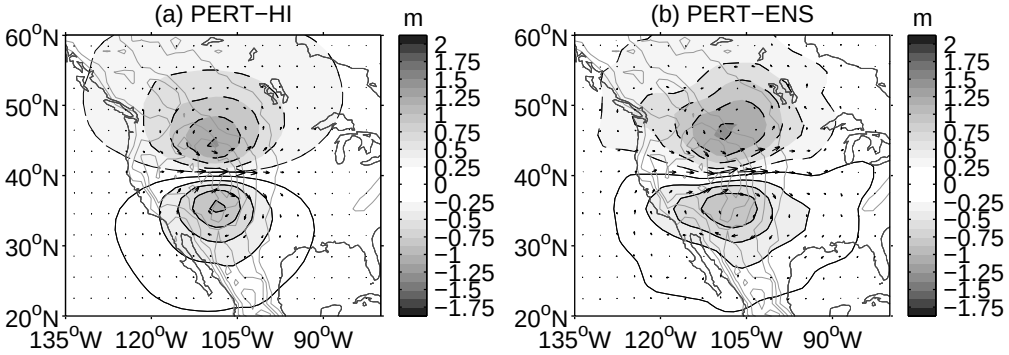

Figure 7. Analysis increment of wind (vectors) and geopotential height (shaded contours) from a single zonal wind observation located at 40° N, 110° W and near 500 hPa when using either (a) the PERT-HI covariances or (b) the PERT-ENS covariances with spatial localization. The orography is also shown in gray contours.

Nonstationary features such as strong horizontal gradients and regions of instability also influence the background error statistics. To capture these influences, however, requires a method for adequately sampling the instantaneous probability distribution of background error at each analysis time. The EnKF is designed to capture such flow-dependent error statistics. To demonstrate this ability of the EnKF, a single temperature observation 1 K greater than the background temperature near 900 hPa was assimilated within a strong near-surface temperature front that appeared over the North Pacific on 27 May 2002 at 12 UTC. The background error std dev for temperature and zonal wind estimated from the EnKF for this case (not shown) is elevated both in the vicinity of the front and close to the associated low pressure centre that is located to the north-west of the front at 45° N, 170° W. During the preceding two days this low pressure centre is seen to rapidly deepen as it progresses eastward across the Pacific Ocean. Eventually it encounters the west coast of North America where it stalls and begins to diminish in intensity.

The analysis increment produced using the background error covariances from the operational 3D-Var (Fig. 8) is clearly unaffected by the local meteorological conditions (the background temperature is shown in dark contours). The temperature increment decays in a nearly isotropic fashion away from the observation location and the wind increment is nearly zero at the location of the temperature observation. In contrast, when using the covariances estimated from the EnKF (Fig. 9) the temperature increment is slightly elongated along the front and the wind increment is larger with vectors oriented parallel with the background temperature gradient at the observation location. For comparison, the same experiment was performed using the EnKF covariances from the previous day (not shown). For that case the gradient in the background temperature field is oriented almost perpendicular to the original case and again the analysis increment for wind is parallel with the background temperature gradient. Finally, the background error covariances from the operational 3D-Var were propagated 6 hours using the tangent-linear and adjoint versions of the forecast model to simulate the implicit covariance propagation of a four-dimensional variational assimilation (4D-Var) over a 6 hour window. Using these propagated error covariances to assimilate the single temperature observation produces an analysis increment (Fig. 10) that is slightly modified relative to the control experiment (Fig. 8). The change in the wind increment demonstrates that the covariance propagation has introduced qualitatively similar correlations between the temperature and wind fields as in the EnKF covariances such that the winds are again

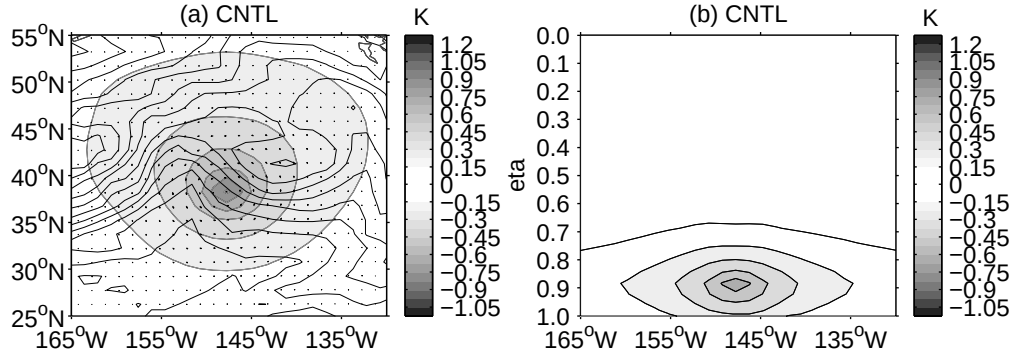

Figure 8. Analysis increment of temperature (shaded contours) and wind (vectors) from a single temperature observation near 900 hPa located in a strong near-surface temperature front at 12 UTC, 27 May 2002. Panel (a) shows temperature and wind increments near 900 hPa and (b) shows the vertical-zonal cross-section of the temperature increment along 40° N. The ensemble mean background temperature field is shown in the left panel as black unshaded contours with a contour interval 10 times larger than for the temperature increment. The CNTL background error covariances are used for generating this analysis increment.

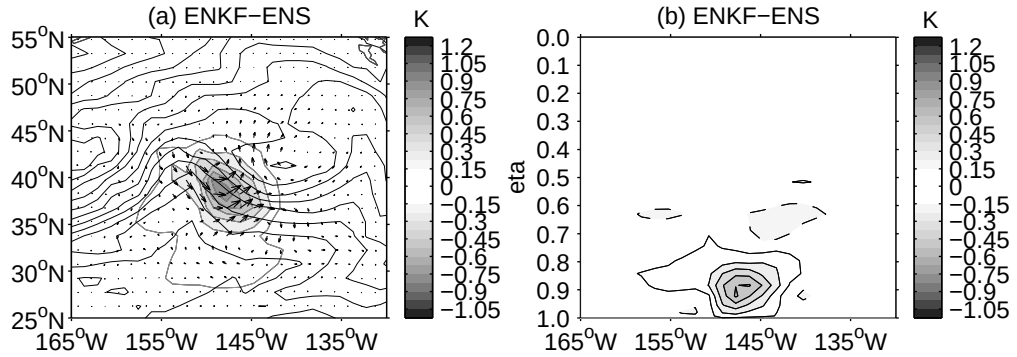

Figure 9. Same as Fig. 8, but using ENKF-ENS background error covariances estimated from the ensemble of EnKF background states valid at 12 UTC, 27 May 2002 with spatially localization.

parallel to the background temperature gradient. These background error correlations between temperature and wind appear to be related to temperature advection across the background temperature front. Further experiments to compare covariances from the EnKF and those implicitly propagated by 4D-Var are left for a future study.

(*d*) *3D-Var analysis sensitivity to ensemble size and spatial localization*

The covariances used in the operational 3D-Var form a full-rank matrix due to the imposition of homogeneity and isotropy on the correlations. A set of random samples were directly generated from these background error covariances to evaluate the sensitivity of the 3D-Var analysis to the number of error samples used to estimate the background error covariances and the spatial localization applied to the correlations. Without the application of spatial localization or even the assumption that the correlations are homogeneous and isotropic, the covariances estimated from these samples should converge to the full covariance matrix of the operational 3D-Var as the number of samples is increased. Therefore, since for this case the result of using the exact covariances is known, the use of these samples provides a practical means to quantify the effect of

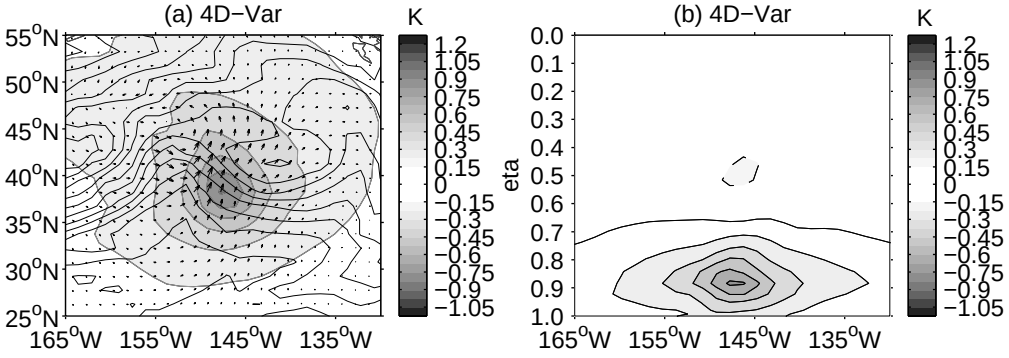

Figure 10. Same as Fig. 8, but with the CNTL background error covariances propagated 6 hours by tangent linear and adjoint versions of the forecast model to simulate the effective background error covariances used in a 4D-Var analysis.

sampling error on the analysis as a function of the ensemble size and the amount of spatial localization. If we assume the resulting covariances are affected by sampling error in a similar way as when the error samples are generated with the perturbed 3D-Var or EnKF, then the conclusions from these sensitivity tests should be applicable to those other cases for which the exact covariances are not known.

A series of 3D-Var analyses with all available data on 26 May 2002 at 12 UTC was performed using various ensemble sizes and length scales for the horizontal and vertical localizations to define an ensemble representation of the background error correlations. (The localization length scale defines the distance at which the function reaches zero.) The final value of the observation component of the cost function was used as a measure of the ability of the analysis to fit the data. For small ensemble sizes with no localization, it should be expected that the analysis can not fit the data well due to the limited number of degrees of freedom which can not exceed the ensemble size and therefore is much less than the number of independent observations. On the other extreme, localization with a very short length scale can eliminate background error correlations between closely spaced observations that are present in the full covariance matrix and consequently allow the analysis to fit these data too closely. The quality of the resulting analyses is evaluated by comparing the final value of the observation cost function with the result from using the exact covariance matrix (see Table 2).

Very little sensitivity with respect to ensemble size is seen when no spatial localization is applied to the correlations (denoted in Table 2 by length scales of ∞). In this case using 32, 128 or 512 error samples all give approximately the same fit to the data (maximum difference of 5%). Conversely, when applying horizontal localization with a length scale of 10000 km or smaller, the observation cost function is reduced by 34% or 46% (depending on the vertical localization used) when the ensemble size is increased from 32 to 512. The localization also affects the fit to the observations when applied to the true covariances, which is shown in the last column of Table 2. This demonstrates, for example, that a similar fit to the observations can be achieved by using 512 error samples with sufficient localization (length scales of 3500 km in the horizontal and 2 units of the natural logarithm of pressure in the vertical) as when using the full covariances, but that this amount of localization applied to the full covariances also decreases the final value of the observation cost function to 85% of the value when no localization is applied. This makes explicit the compromise involved in using spatial localization: the sampling error can not be removed without also suppressing the true correlations to some extent.

TABLE 2. NORMALIZED FINAL VALUE OF THE OBSERVATION COST FUNCTION AS A FUNCTION OF ENSEMBLE SIZE AND THE HORIZONTAL AND VERTICAL LOCALIZATION LENGTH SCALES (DISTANCE WHERE CORRELATION REACHES ZERO).

Localization length scales		Ensemble size			
Horizontal	Vertical	32	128	512	∞
∞	∞	3.15	3.10	2.98	1.00
10000 km	∞	2.72	2.30	1.77	0.96
5000 km	∞	2.42	1.83	1.35	0.91
3500 km	∞	2.22	1.60	1.19	0.87
2800 km	∞	2.09	1.46	1.12	0.84
2000 km	∞	1.89	1.28	-	0.79
1000 km	∞	1.53	1.04	-	0.65
10000 km	2	2.23	1.73	1.31	0.94
5000 km	2	1.82	1.35	1.08	0.89
3500 km	2	1.59	1.20	1.01	0.85
2800 km	2	1.47	1.11	0.97	0.82
2000 km	2	1.30	1.02	-	0.76
1000 km	2	1.04	0.88	-	0.63

The vertical length scale is specified in units of the natural logarithm of pressure. The column for ∞ ensemble size corresponds to using the full covariance matrix. The missing values could not be computed due to computer memory limitations.

5. RESULTS FROM REALISTIC FORECAST-ANALYSIS EXPERIMENTS

(*a*) *Verification of analysis increments from an individual case*

The 3D-Var was used to produce global analyses for 27 May 2002 at 12 UTC using the full set of operational observations and the background state from the operational system. This is the same case with a developing low pressure system over the North Pacific used to produce the analysis increments from a single temperature observation as shown in Fig. 8 to 10. Analyses were obtained using the CNTL and ENKF-ENS background error covariances and 5 day forecasts produced from each. Figure 11 shows the statistics for the difference between the analyses and also the differences between the two resulting forecasts for the North Pacific region from 20° N to 65° N and 140° E to 120° W. The differences in the analyses due to the change in background error covariances have a std dev near 1 m s^{-1} for the wind components, 0.5 K for the temperature field and less than 3 K for the dew-point depression (ES). After 2 days the maximum difference in the forecasts is seen for the winds near the level of the extra-tropical jet where the std dev has grown by about a factor of two. After 5 days the differences in all the variables have grown by about a factor of four, except for the winds at the jet level for which they have grown by about a factor of seven.

Figure 12 shows the statistics of the differences between forecast and analysis where the analysis is taken from a forecast-analysis experiment using the operational background error covariances. The statistics are again calculated for a region over the North Pacific as defined earlier. These differences can be considered to be reasonable estimates of the true error in the forecasts. Even though the difference between the two background error covariance matrices produces a large difference in the resulting forecasts (Fig. 11), the agreement of the forecasts with the verifying analyses is much more similar. As expected the largest difference for the winds occurs at the jet level where the std dev of the 5 day forecast from the ENKF-ENS analysis is about 1 m s^{-1} smaller than for the control. In general, the analysis using the ENKF-ENS background error

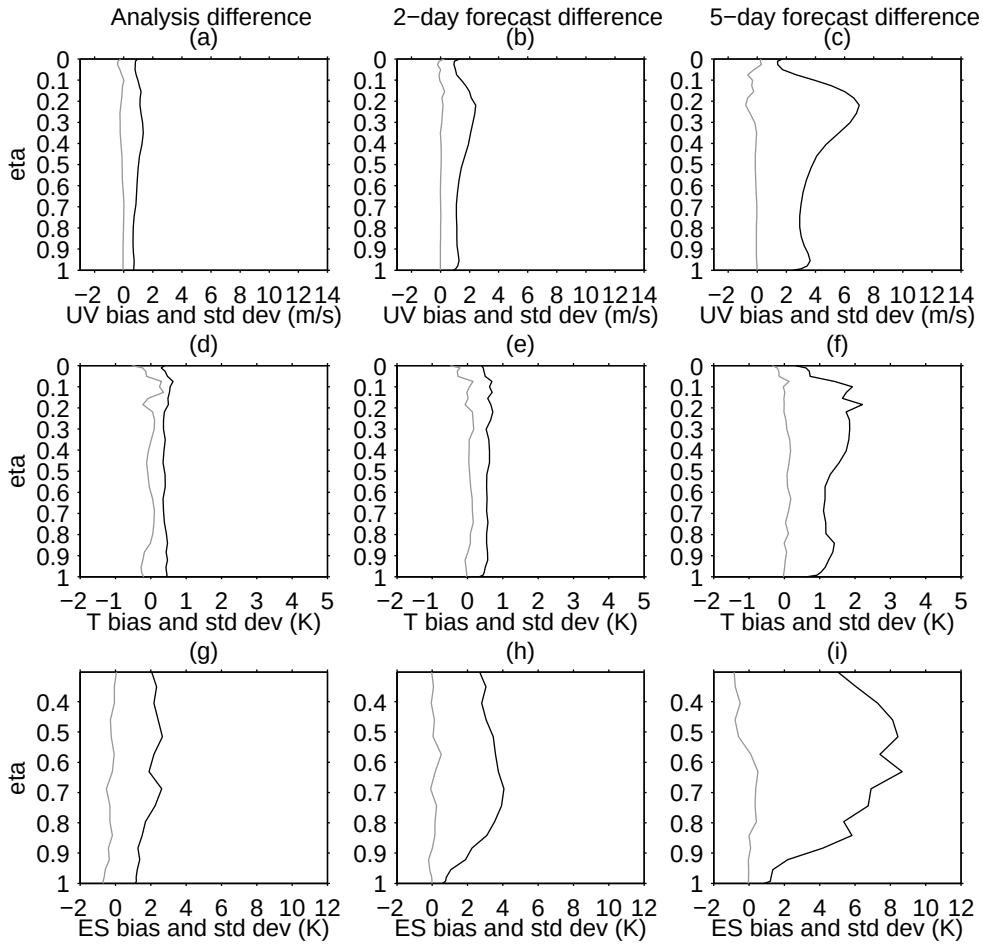

Figure 11. Bias (gray curves) and std dev (black curves) for the North Pacific region (20° N to 65° N and 140° E to 120° W) of the difference between the analyses (first column) valid at 12 UTC, 27 May 2002 and the resulting 2 and 5 day forecasts (second and third columns, respectively) computed using the CNTL and ENKF-ENS background error covariances.

covariances produces an improved forecast for this case and this region, except above about 200 hPa where a slight degradation is seen in the 5 day forecast.

(*b*) *Verification of two week forecast-analysis experiments*

A set of forecast-analysis experiments was performed to evaluate in a quasi-operational setting the impact of using each of the background error covariance matrices listed in Table 1. All experiments used the identical configuration of the Global Environmental Multiscale (GEM) model (Côté *et al.* 1998) to obtain the 6 hour forecasts used as the background states. The model horizontal resolution for the experiments was 1.5° with 28 eta levels in the vertical and the top set to 10 hPa. The same spatial resolution was also used for both the 3D-Var and EnKF analyses. All experiments also used a complete set of meteorological observations except that, to be consistent with the EnKF experiments of Houtekamer *et al.* (2003), surface wind and surface humidity observations were omitted. Instead of applying the normal quality control procedure to each

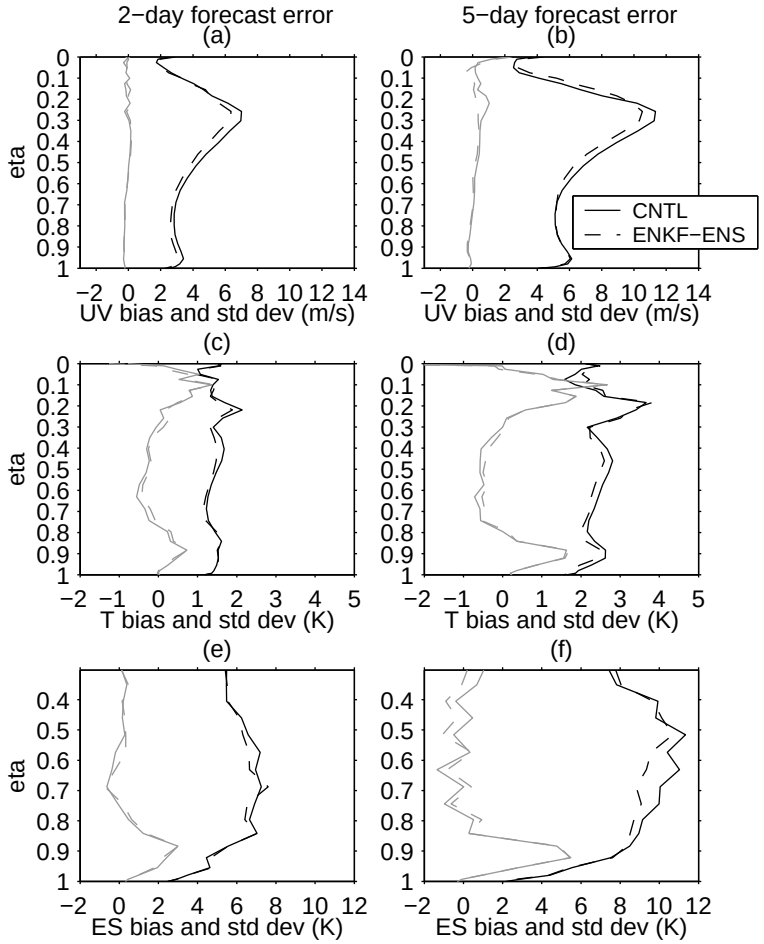

Figure 12. Bias (gray curves) and std dev (black curves) for the North Pacific region (20° N to 65° N and 140° E to 120° W) of the difference between the analyses from the CNTL forecast-analysis experiment and the 2 and 5 day forecasts from the analysis produced using either the operational background error covariances (CNTL) or the flow-dependent covariances estimated from the EnKF (ENKF-ENS) at 12 UTC, 27 May 2002.

experiment, only the observations accepted for assimilation in the operational analyses were used. This ensures that the identical set of observations was used for each experiment. A forecast-analysis experiment using the background error covariances from the operational 3D-Var served as the control (CNTL experiment). All forecast-analysis experiments spanned the period of 19 May 2002 to 2 June 2002.

The goal of the first set of experiments is to quantify the sensitivity of the analyses and background states from a complete forecast-analysis experiment to sampling error in the background error covariances. These covariances were estimated from either 32, 128 or 512 random samples generated directly from the multi-variate Gaussian distribution defined by the operational background error covariances. Spatial localization with length scales of 5000 km in the horizontal and 2 units of the natural logarithm of pressure in the vertical was applied to the ensemble representation of the correlations. The std dev of the differences between radiosonde observations and the 6 hour forecasts are shown in the upper panels of Fig. 13 from the forecast-analysis experiment using

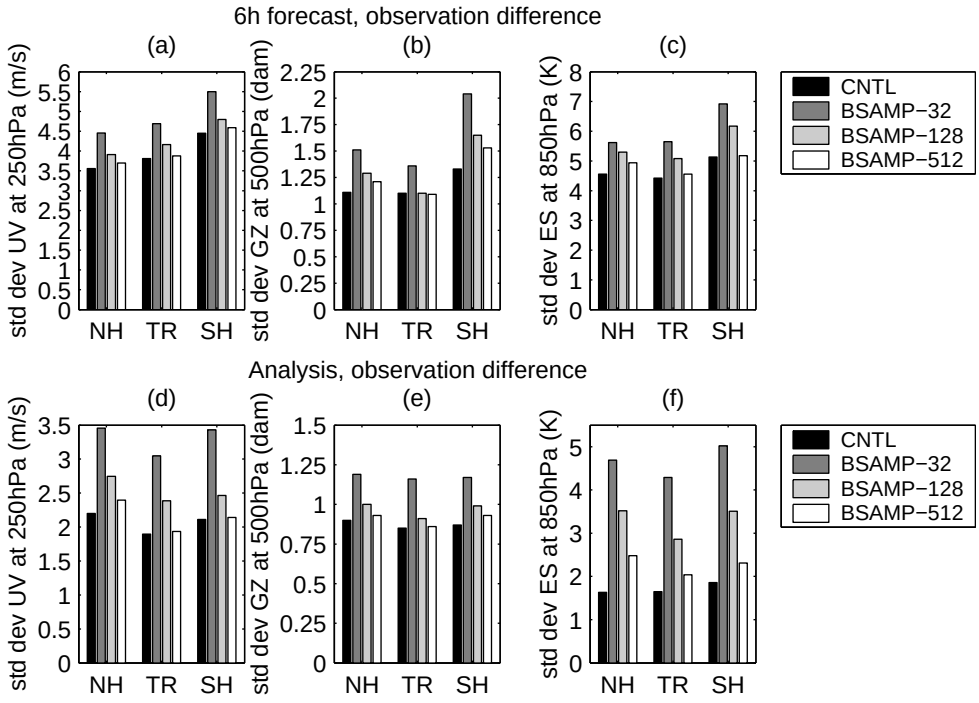

Figure 13. Std dev of the difference between the radiosonde observations and the 6 hour forecast (panels (a)-(c)) and the analyses (panels (d)-(f)) from the CNTL, BSAMP-32, BSAMP-128 and BSAMP-512 3D-Var experiments.

the CNTL, BSAMP-32, BSAMP-128 and BSAMP-512 background error covariances. Similarly, the std dev of the differences between radiosonde observations and the analyses are shown in the lower panels of Fig. 13 for the same experiments. Only the std dev for wind components at 250 hPa, geopotential height at 500 hPa and dew-point depression at 850 hPa are shown for the northern extra-tropical, tropical and southern extra-tropical regions. Consistent with Table 2, these results show a strong sensitivity in both the 6 hour forecasts and the analyses to the number of error samples used to estimate the background error covariances. For the specific correlation localization used in these experiments, even 512 error samples do not appear to be sufficient to estimate the full background error covariance matrix. The effects of sampling error on the 6 hour forecasts are largest for geopotential height in the extra-tropical regions and for the analyses the effects are largest for dew-point depression. Only in the tropics, the accuracy of the forecasts and analyses when using 512 error samples approaches that of the experiment with the full covariances for both wind components and geopotential height. Also, it is interesting to note that while the effect of sampling error on geopotential height remains approximately the same between the analyses and 6 hour forecasts, for winds and dew-point depression this effect is reduced substantially on the forecasts relative to the analyses.

In the next set of experiments the six types of covariance matrices estimated using the perturbed 3D-Var and EnKF error samples were evaluated. Again, spatial localization was applied when using the ensemble representation of the correlations (that is PERT-ENS, PERT-HYB, ENKF-ENS, and ENKF-HYB). For the covariances estimated from the EnKF forecast ensembles, the length scale for the localizing correlation

function was set to 2800 km in the horizontal and 2 units for the natural logarithm of pressure in the vertical. These parameters were chosen to match those used for the EnKF experiment of Houtekamer *et al.* (2003). For the covariances computed from the perturbed forecast-analysis experiments, the horizontal length scales were set for each vertical level to be proportional to the horizontal length scale of the homogeneous and isotropic correlations. The proportionality constant was set so that the length scale was close to 2800 km in the middle troposphere, but became substantially larger at upper levels.

The same verification statistics as shown in the previous figure were computed for the forecast-analysis experiments using the PERT-HI, PERT-ENS and PERT-HYB background error covariances. In the upper panels of Fig. 14 these statistics are shown for the 6 hour forecasts. When compared with the CNTL experiment, the use of the covariances estimated from the perturbed 3D-Var error samples results in 6 hour forecasts and analyses (not shown) of relatively similar quality. Employing the different approaches to estimate the correlations also has little impact on the results. Similar results were also obtained for the experiments using the error samples obtained from the EnKF as shown in the lower panels of Fig. 14. While the agreement of the 6 hour forecasts with the observations was similar to the PERT experiments, the difference between the analyses and the observations (not shown) was slightly larger for the wind observations and almost twice as large (in terms of std dev) for the dew-point depression. The decreased fit to the humidity observations is likely due to the lack of model error perturbations for humidity in the EnKF that results in the background error variances for humidity being underestimated. This lack of fit, however, is completely eliminated in the 6 hour forecast that has nearly identical std dev with respect to the radiosonde observations as the CNTL experiment.

The next figure was produced to provide a direct comparison of the quality of the error samples obtained from the EnKF, independent of sampling error. Results from the control experiment are shown together with those from an experiment using samples generated directly from the operational 3D-Var background error covariances (BSAMP-128) and an experiment using error samples from the EnKF (ENKF-ENS). The BSAMP-128 and ENKF-ENS experiments both use 128 error samples and the identical ensemble representation of the correlations, including identical spatial localization. A summary of the verification statistics with respect to radiosonde observations for the 6 hour forecasts is shown in Fig. 15. These results show that estimating the background error covariances from the EnKF-derived error samples generally leads to decreased 6 hour forecast error versus using the samples computed from the operational covariances. Only in the tropics, where the error from the three experiments are very similar, is this improvement not seen. If we assume the extent and impact of the sampling error is similar for both the BSAMP-128 and ENKF-ENS experiments, then this result also provides an estimate of the improvements that can be expected if the sampling error could be removed when estimating covariances from the EnKF by using a larger ensemble.

(*c*) *Verification of 5 day forecasts*

The impact of the new background error covariances on 5 day forecasts was also evaluated. The forecasts were produced from the analyses of the control experiment and of the six experiments with background error covariances estimated from the perturbed 3D-Var and EnKF. Analyses from 00 UTC and 12 UTC, 24 May 2002 to 2 June 2002 were used to initialize integrations of the operational configuration of the GEM model that uses a higher horizontal resolution (0.9°) than was used in the forecast-analysis experiments.

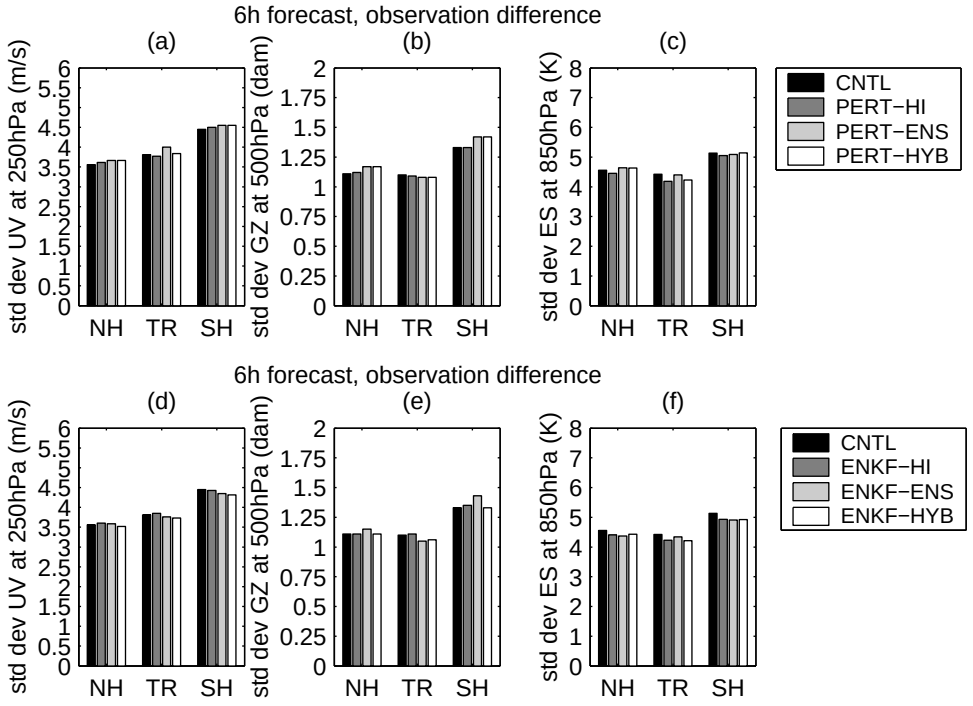

Figure 14. Std dev of the difference between the radiosonde observations and the 6 hour forecast from the CNTL, PERT-HI, PERT-ENS and PERT-HYB 3D-Var assimilation experiments are shown in panels (a)-(c). The same statistics are shown in panels (d)-(f) for the CNTL, ENKF-HI, ENKF-ENS and ENKF-HYB 3D-Var experiments.

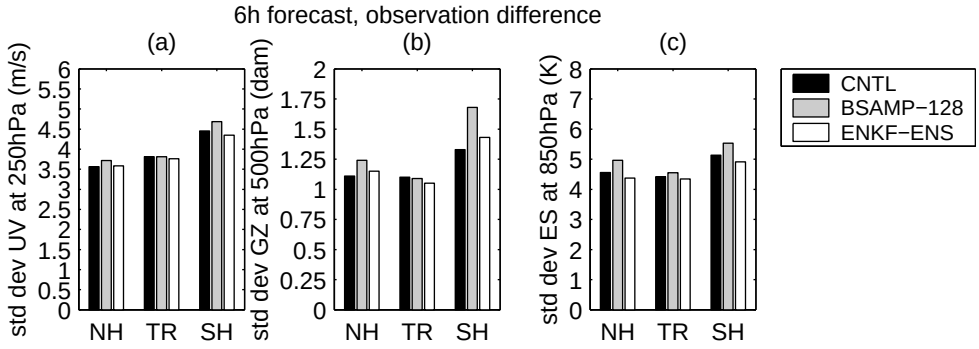

Figure 15. Std dev of the difference between the radiosonde observations and the 6 hour forecast from the CNTL, BSAMP and ENKF-ENS 3D-Var experiments.

The statistics of the differences between radiosonde observations and the 5 day forecasts from the CNTL, PERT-HI, PERT-ENS and PERT-HYB experiments are shown in the upper panels of Fig. 16. The impact of using background error covariances derived from the perturbed 3D-Var is positive on the std dev for winds at 250 hPa by about 1 m s^{-1} and for geopotential height at 500 hPa by about 0.8 dam in the southern extra-tropics. For these variables the hybrid covariances (PERT-HYB) lead to the most improvement.

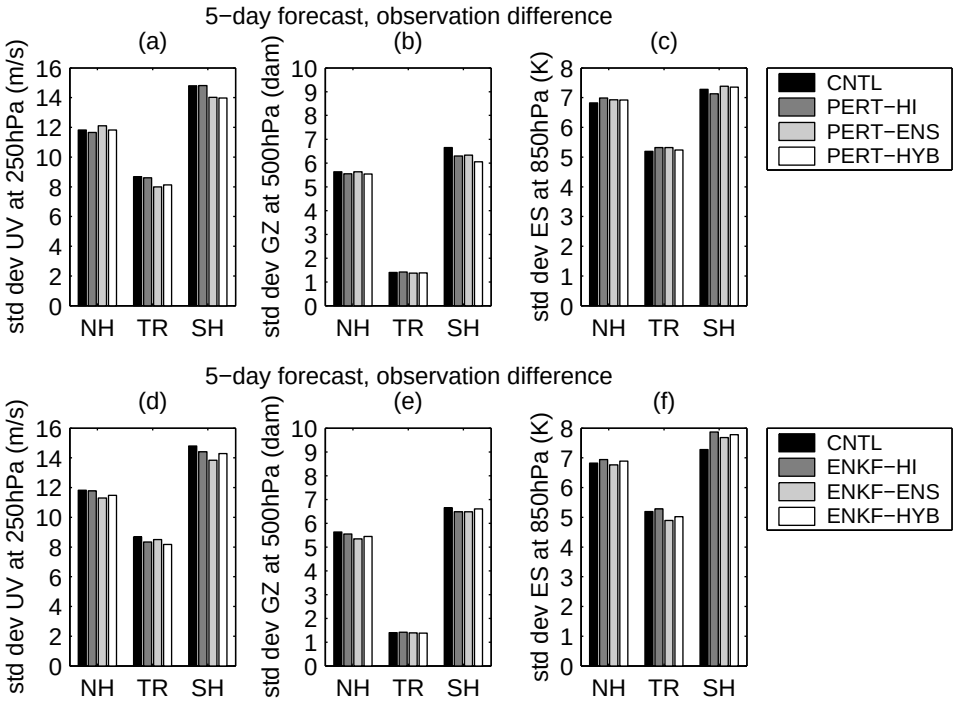

Figure 16. Std dev of the difference between the radiosonde observations and the 5 day forecast from the CNTL, PERT-HI, PERT-ENS and PERT-HYB 3D-Var assimilation experiments are shown in panels (a)-(c). The same results from the CNTL, ENKF-HI, ENKF-ENS and ENKF-HYB 3D-Var experiments are shown in panels (d)-(f).

Similar results are shown in the lower panels of Fig. 16 for the forecasts produced from the CNTL, ENKF-HI, ENKF-ENS and ENKF-HYB experiments. For these experiments the impact on the 5 day forecasts is generally a small improvement to the winds and geopotential heights. The maximum improvement to the winds is again about 1 m s^{-1} in the southern extra-tropics when the ensemble representation of the correlations is used (ENKF-ENS). For the geopotential heights, the maximum improvement is only about 0.2 dam in the northern extra-tropics, again when using the ENKF-ENS covariances.

6. CONCLUSIONS

By examining the estimated background error covariances, we find that the perturbed 3D-Var and EnKF approaches produce covariances substantially different from those used operationally. Some of these differences include: the horizontal structure of the background error variance is more realistic; the mass-wind correlations due to geostrophy are weaker; the effects of orography are resolved by using an ensemble representation of the correlations; and flow-dependent structures in a dynamically active meteorological situation are captured using EnKF-derived covariances.

Consistent with previous studies (e.g. Hamill and Whitaker 2001), we conclude that spatial localization of the background error correlations is essential to reduce the sampling error when the correlations are estimated directly from a small ensemble of error samples. From the experiments using random samples generated from the operational background error covariances (Section 4d) it is clear that without sufficient

localization the analysis is unable to fit the observations. As the localization length scale is decreased, however, the analysis draws towards the observations even more than when the true covariances are used. At this point, we can conclude that the localization length scale is too short. The length scale necessary to provide a similar fit to the observations as when using the true covariances depends on the number of error samples from which the covariances are estimated. However, applying this same localization to the true covariances causes them to be substantially altered as evidenced by the increased fit to the observations.

One could expect that the results from using the perturbed 3D-Var or EnKF-derived background error covariances to be different than when using the operational 3D-Var covariances only for dynamically unstable regions or periods. In these cases the dynamics amplify the errors in the initial conditions and as a consequence the relative importance of the simple estimate of the model error covariances is reduced. We saw that, on average, the model error perturbations in the perturbed 3D-Var experiments are very important, especially in the extra-tropical troposphere. Since the model error perturbations for both the perturbed 3D-Var and EnKF are generated using error covariances that resemble the background error covariances of the operational system, the model error component of the estimated background error covariances is very similar to those used in the BSAMP experiments. But, the results of Section 5b show that the use of the EnKF covariances produces better 6 hour forecasts than the BSAMP-128 experiment. Use of the covariances estimated from the perturbed 3D-Var error samples also improved the 6 hour forecasts for geopotential height and humidity in the extra-tropics relative to the BSAMP-128 experiment. Therefore we can conclude that the component of the error samples from the perturbed 3D-Var or EnKF that originates from the propagation of initial condition error improves the quality of the estimated error covariances beyond the quality of the covariances from the operational 3D-Var. Consequently, one can expect that a reduction in the sampling error (by increasing the ensemble size) will enable the perturbed 3D-Var or EnKF approaches to produce error covariances leading to improved forecasts relative to the operational system. To the extent that the differences between the BSAMP-128 and CNTL experiments are representative of the effects of sampling error in the experiments using error samples from the perturbed 3D-Var or EnKF, this provides an indication of the magnitude of the expected improvements. This represents a potential additional improvement to the 6 hour forecast error std dev (from using a much larger ensemble size) of up to 0.3 m s^{-1} for wind components, 4 m for geopotential height and 0.9 K for dew-point depression (from Fig. 15). However, additional research is required to confirm this preliminary result and to identify any effect that the simpler structure of the operational background error covariances may have on the approach.

This study focused on the use of background error covariances estimated from methods that attempt to more realistically sample and represent the covariances in a variational assimilation system. Comparisons with the results of Houtekamer *et al.* (2003) demonstrate that when using errors samples obtained from the EnKF, the 3D-Var could be made to closely reproduce the mean analysis produced by the EnKF system (not shown). However, one benefit of using a variational system is the flexibility to use various ways of representing the correlations including the use of hybrid covariances. Additionally, if the tangent-linear and adjoint versions of the forecast model are available, the extension to 4D-Var is relatively straightforward. This approach is thought to be an improvement over 3D-Var mostly due to the ability to use data closer to the measurement time. Even though 4D-Var implicitly propagates the covariances over the assimilation window, it is not practical to explicitly compute these flow-dependent background error covariances for use in the subsequent analysis. Therefore, it is envisaged that a system

using flow-dependent error covariances produced at regular intervals by the EnKF and assimilating all the data within the interval with 4D-Var could possibly produce better analyses than either the EnKF or 4D-Var (with stationary background error covariances) alone. The merits of such a system are discussed by Lorenc (2003). A similar type of system was evaluated at the European Centre for Medium-Range Weather Forecasts by Fisher and Andersson (2001), however with the flow-dependent background error covariances computed using singular vectors. They found no significant improvement over the use of 4D-Var with stationary background error covariances and therefore abandoned the approach due to the high computational cost of obtaining the singular vectors. At CMC, 4D-Var will soon replace 3D-Var as the operational analysis system and the EnKF is already being developed to initialize the operational probabilistic predictions. Consequently, the ensembles from the EnKF could be used to provide flow-dependent background error covariances to 4D-Var at no extra computational cost. This will be the subject of future research.

Finally, the two approaches examined in this study for generating samples of background error require both the observation and model error covariances as input. The tuning procedure implemented in the perturbed 3D-Var experiments allows the model error variances to be estimated from the innovations, however, this procedure itself relies on the accuracy of the observation error variances. This represents a fundamental limitation for all approaches: the innovations are the only source of information on the true error, but their is no way to separate the contributions of observation error and background error (where the background error is composed of model error and error originating from the previous analysis) without employing additional assumptions. For example, if the observation errors are assumed to be spatially uncorrelated and the background errors correlated, then given a sufficiently dense observing network the innovation variance can be partitioned between the observation and background components and the background error spatial correlations estimated (Hollingsworth and Lönnberg 1986). Though such assumptions are necessary to separate the two error components, they can not be independently verified from the innovations. More generally, if it is assumed that the covariances only depend on a sufficiently reduced set of parameters, then the approaches of Dee (1995), Talagrand (1999) and Desroziers and Ivanov (2001) may be useful for improving both the observation and model error covariances that will be necessary to obtain the maximum benefit from the perturbed 3D-Var and EnKF approaches.

ACKNOWLEDGEMENTS

The author thanks P. L. Houtekamer and H. L. Mitchell for providing the forecast ensembles from the EnKF for this study and P. Gauthier for many useful comments during the course of this work.

APPENDIX A

Adaptive tuning of model error variances

Here, the technique is described for computing the adaptively tuned scaling factors applied to the model error variances. The scaling factors for wind components and temperature are allowed to depend on vertical level and are independently computed for three latitude bands: northern extra-tropics (north of 20° N), tropics (between 20° N and 20° S) and southern extra-tropics (south of 20° S). Consequently a total of 168 parameters are estimated at each analysis time. First, the difference between the

forecasts from the perturbed and unperturbed experiments is computed for a particular analysis time and projected into observation space (with respect to only radiosonde wind and temperature observations). For each latitude band and vertical level the variances of these differences are added to those of the observation error covariance matrix. By adding the model error variances (also projected into observation space), the result is the simulated innovation variances. These simulated variances should equal the variances computed from real innovations, that is the difference between the observations and the background state from the unperturbed analysis-forecast experiment. All quantities in the resulting equation

$$\text{var}\left\{H\left(\mathbf{x}_{cntl}\right)-\mathbf{y}\right\}=\text{var}\left\{H\left(\mathbf{x}_{pert}\right)-H\left(\mathbf{x}_{cntl}\right)\right\}+\overline{\text{diag}\left(\mathbf{R}\right)}+\text{var}\left\{H\left(\boldsymbol{\epsilon}_{mod}\right)\right\}, \tag{A.1}$$

are known except the model error, $\boldsymbol{\epsilon}_{mod}$. The operator $\text{diag}(\mathbf{X})$ represents a vector containing the diagonal elements of the matrix $\mathbf{X}$. The variances (var) and average in Eq. (A.1) are computed by incorporating all observation locations for a given vertical level and latitude band. It should be noted that the innovation variances (the right side of Eq. (A.1)) are averaged over the 2 week period of the experiment since it is the stationary component of background error that we wish to sample. By making the substitution

$$\text{var}\left\{H(\boldsymbol{\epsilon}_{mod})\right\}=\alpha^2\left(j,k\right)\overline{\text{diag}\left(\mathbf{H}\mathbf{B}_{old}\mathbf{H}^T\right)}, \tag{A.2}$$

where $\mathbf{B}_{old}$ is the currently available background error covariance matrix, an equation for the scaling factor α as a function of the indices for latitude band (j) and vertical level (k) is obtained:

$$\alpha^2\left(j,k\right)=\frac{\text{var}\left\{H\left(\mathbf{x}_{cntl}\right)-\mathbf{y}\right\}-\text{var}\left\{H\left(\mathbf{x}_{pert}\right)-H\left(\mathbf{x}_{cntl}\right)\right\}-\overline{\text{diag}\left(\mathbf{R}\right)}}{\overline{\text{diag}\left(\mathbf{H}\mathbf{B}_{old}\mathbf{H}^T\right)}}, \tag{A.3}$$

under the constraint that $\alpha^2\left(j,k\right)\geqslant 0$. This calculation is performed at each analysis time for wind components and temperature. The model error perturbations for humidity are then computed using the temperature scaling coefficients. Similarly, the surface pressure perturbations are scaled using the coefficients computed for wind at the lowest level. This was done because the innovation statistics for surface pressure and the humidity variable used in the 3D-Var were not readily available with existing software.

APPENDIX B

A spatially localized ensemble representation of the background error correlations

In this appendix, we consider how an ensemble representation of the correlations with spatial localization can be efficiently implemented in a variational assimilation framework. The correlation matrix estimated directly from an ensemble of background error samples is

$$\mathbf{C}_{ens}=\frac{1}{N_e-1}\mathbf{V}^{-1/2}\left(\sum_{i=1}^{N_e}\mathbf{e}_i\mathbf{e}_i^T\right)\mathbf{V}^{-1/2}, \tag{B.1}$$

where $\mathbf{e}_i$ is the deviation of the ith ensemble member from the ensemble mean, $\mathbf{V}$ is a diagonal matrix containing the ensemble spread variances and N_e is the size of the ensemble. To precondition the minimization, as in the operational 3D-Var, only the square-root of the background error covariance matrix is required and is given by

$$\mathbf{C}_{ens}^{1/2}=\left(\frac{1}{N_e-1}\right)^{1/2}\mathbf{V}^{-1/2}\left[\mathbf{e}_1,\mathbf{e}_2,\ldots,\mathbf{e}_{N_e}\right], \tag{B.2}$$

where the object in brackets is the N_x by N_e matrix containing the deviations of all ensemble members from the ensemble mean. Therefore it is clear that, without imposing additional constraints, the rank of such an ensemble representation of the correlation matrix and the dimension of the control vector can not be greater than the size of the ensemble.

The localization procedure consists of simply computing the Schur product (element-wise matrix multiplication) between the estimated correlation matrix and a prescribed correlation matrix with correlations close to one for separation distances where the estimated correlations are deemed reliable and approaching zero outside this range (Gaspari and Cohn 1999). Houtekamer and Mitchell (2001) describe the implementation of localization in the context of an EnKF that explicitly solves the linear analysis equation for each of a series of observation batches. To implement localization in a variational analysis system with preconditioning, the square-root of the localized ensemble representation of the correlations can be expressed as

$$\mathbf{C}_{ens}^{1/2} = (N_e - 1)^{-1/2}\, \mathbf{V}^{-1/2} \left[\text{diag}\,(\mathbf{e}_1)\, \mathbf{L}^{1/2}, \text{diag}\,(\mathbf{e}_2)\, \mathbf{L}^{1/2}, \ldots, \text{diag}\,(\mathbf{e}_{N_e})\, \mathbf{L}^{1/2} \right], \tag{B.3}$$

where $\mathbf{L}$ is the prescribed correlation matrix used for localizing the ensemble correlations and the operator $\text{diag}(\mathbf{e})$ represents a matrix with the vector $\mathbf{e}$ along its diagonal, but otherwise containing zeros. For both horizontal and vertical localization the 5th-order piecewise rational function given as equation (4.10) in Gaspari and Cohn (1999) was used. Through the localization procedure, the rank of the correlation matrix is increased to at most $N_e \times \text{rank}(\mathbf{L})$. However, since the correlations in $\mathbf{L}$ are smooth and relatively large-scale, the dimension of the problem can be reduced by using a truncated spectral expansion to represent these correlations.

REFERENCES

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

Bishop, C. H., B. Etherton, and S. H. Majumdar 2001 Adaptive sampling with the ensemble transform Kalman filter. Part I: Theoretical aspects. *Mon. Weather Rev.,* **129,** 420–436

Burgers, G., P. J. Van Leeuwen, and G. Evensen 1998 Analysis scheme in the ensemble Kalman filter. *Mon. Weather Rev.,* **126,** 1719–1724

Côté, J., S. Gravel, A. Méthot, A. Patoine, M. Roch, and A. Staniforth 1998 The operational CMC-MRB global environmental multiscale (GEM) model. part I: Design considerations and formulation. *Mon. Weather Rev.,* **126,** 1373–1395

Daley, R. 1991 *Atmospheric Data Analysis,* Cambridge University Press, Cambridge, UK

Dee, D. P. 1995 On-line estimation of error covariance parameters for atmospheric data assimilation. *Mon. Weather Rev.,* **123,** 1128–1145

Derber, J., and F. Bouttier 1999 A reformulation of the background error covariance in the ECMWF global data assimilation system. *Tellus,* **51A,** 195–221

Derber, J., and A. Rosati 1989 A global oceanic data assimilation system. *J. Phys. Oceanogr.,* **19,** 1333–1347

Desroziers, G. 1997 A coordinate change for data assimilation in spherical geometry of frontal structures. *Mon. Weather Rev.,* **125,** 3030–3038

Desroziers, G., and S. Ivanov 2001 Diagnosis and adaptive tuning of observation-error parameters in a variational assimilation. *Q. J. R. Meteorol. Soc.,* **127,** 1433–1452

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

Fisher, M. 1999 Background error statistics derived from an ensemble of analyses, *ECMWF research department memorandum,* **79,** [available from ECMWF, Shinfield Park, Reading, RG2 9AX, UK]

Fisher, M. and E. Andersson	2001	Developments in 4D-Var and Kalman filtering, *ECMWF research department technical memorandum*, **347,** [available from ECMWF, Shinfield Park, Reading, RG2 9AX, UK]
Gaspari, G., and S. Cohn	1999	Construction of correlation functions in two and three dimensions. *Q. J. R. Meteorol. Soc.,* **125,** 723–757
Gauthier, P., M. Buehner, and L. Fillion	1998	'Background-error statistics modelling in a 3D variational data assimilation scheme.' Pp. 131–145 in Proceedings of the ECMWF workshop on diagnosis of data assimilation systems, 2–4 November, 1998, Reading, UK
Gelb, A., ed.	1974	*Applied Optimal Estimation,* M.I.T. Press, Cambridge, USA
Hamill, T. M., and C. Snyder	2000	A hybrid ensemble Kalman filter - 3D variational analysis scheme. *Mon. Weather Rev.,* **128,** 2905–2919
Hamill, T. M., and J. S. Whitaker	2001	Distance-dependent filtering of background error covariance estimates in an ensemble Kalman filter. *Mon. Weather Rev.,* **129,** 2776–2790
Hollingsworth, A., and P. Lönnberg	1986	The statistical structure of short-range forecast errors as determined from radiosonde data. Part i: The wind field. *Tellus,* **38A,** 111–136
Houtekamer, P. L., and H. L. Mitchell	1998	Data assimilation using an ensemble Kalman filter technique. *Mon. Weather Rev.,* **126,** 796–811
Houtekamer, P. L., and H. L. Mitchell	2001	A sequential ensemble Kalman filter for atmospheric data assimilation. *Mon. Weather Rev.,* **129,** 123–137
Houtekamer, P. L., L. Lefaivre, J. Derome, H. Ritchie, and H. L. Mitchell	1996	A system simulation approach to ensemble prediction. *Mon. Weather Rev.,* **124,** 1225–1242
Houtekamer, P. L., H. L. Mitchell, G. Pellerin, M. Buehner, M. Charron, L. Spacek, and B. Hansen	2003	Atmospheric data assimilation with the ensemble Kalman filter: results with real observations. *Mon. Weather Rev.,* (submitted)
Lorenc, A. C.	2003	The potential of the ensemble Kalman filter for NWP - a comparison with 4D-Var. *Q. J. R. Meteorol. Soc.,* **129,** 3183–3203
Mitchell, H. L., and P. L. Houtekamer	2000	An adaptive ensemble Kalman filter, *Mon. Weather Rev.,* **128,** 416–433
Parrish, D. F., and J. C. Derber	1992	The national meteorological center's spectral statistical interpolation analysis system. *Mon. Weather Rev.,* **120,** 1747–1763
Rabier, F., A. McNally, E. Andersson, P. Courtier, P. Undén, J. Eyre, A. Hollingsworth, and F. Bouttier	1998	The ECMWF implementation of three dimensional variational assimilation (3D-Var). Part ii: Structure functions. *Q. J. R. Meteorol. Soc.,* **124,** 1809–1830
Talagrand, O.	1999	'A posteriori evaluation and verification of analysis and assimilation algorithms'. Pp. 17–28 in Proceedings of the Workshop on Diagnosis of Data Assimilation Systems, November 2–4, 1999, Reading, UK
Tippett, M. K., J. L. Anderson, C. H. Bishop, T. M. Hamill, and J. S. Whitaker	2003	Ensemble square root filters. *Mon. Weather Rev.,* **131,** 1485–1490
Weaver, A. T., and P. Courtier	2001	Correlation modelling on the sphere using a generalized diffusion equation. *Q. J. R. Meteorol. Soc.,* **127,** 1815–1846
Whitaker, J., and T. M. Hamill	2002	Ensemble data assimilation without perturbed observations. *Mon. Weather Rev.,* **130,** 1913–1924
Wu, W.-S., R. J. Purser, and D. F. Parrish	2002	Three-dimensional variational analysis with spatially inhomogeneous covariances. *Mon. Weather Rev.,* **130,** 2905–2916