6.1 APPLICATION OF AN ENSEMBLE KALMAN FILTER WITH A DRY MULTI-LEVEL PRIMITIVE-EQUATION MODEL

Herschel L. Mitchell,* P. L. Houtekamer, and Gérard Pellerin

Meteorological Service of Canada, Dorval, Québec, Canada

1. INTRODUCTION

Evensen (1994) proposed a Kalman-filter based Monte Carlo approach that is potentially feasible for large atmospheric and oceanic applications. This approach, termed an ensemble Kalman filter (EnKF), uses a forecast model to integrate an ensemble of model states from one analysis time to the next and employs ensemble-based covariances in the analysis step. It is well suited for parallel computation: each ensemble member can simply be integrated on a different processor of a parallel computer (Evensen 1994, Keppenne 2000).

An efficient algorithm for the analysis step of the EnKF was proposed recently by Houtekamer and Mitchell (2001, hereafter HM01). That algorithm solves the Kalman filter equations by organizing the observations into batches which are assimilated sequentially. The small (and noisy) background-error covariances associated with remote observations are filtered using a Schur (elementwise) product of the covariances calculated from the ensemble and a correlation function having compact support (Gaspari and Cohn 1999). As in an earlier paper (Houtekamer and Mitchell 1998, hereafter HM98), the algorithm utilizes a pair of ensembles to deal with a problem of inbreeding. Having two ensembles allows the Kalman gain used for the assimilation of data into one ensemble to be computed from the other ensemble. In HM01, the sequential algorithm was shown to be computationally feasible for synoptic-scale analysis in an operational context, *if* the required number of ensemble members was O(100).

The purpose of this study is to test the sequential EnKF algorithm of HM01 in a primitive-equation context and to examine some of the issues that remain regarding its suitability for operational atmospheric data assimilation. In particular, the Kalman-filter framework includes a model-error term and its proper specification is crucial (Dee 1995). One focus of this study is how to account for model error in such a way that the balance in a primitive-equation context be maintained. A second concern relates to the use of the Schur product for localization. Will this produce imbalance in a primitive-equation context? Finally, how will the required number of ensemble members change as model-state vectors and numbers of observations approach values that occur in an operational context?

* *Corresponding author address*: Herschel L. Mitchell, Meteorological Research Branch, Meteorological Service of Canada, 2121 Trans-Canada Highway, Dorval, Québec, H9P 1J3, Canada, e-mail Herschel.Mitchell@ec.gc.ca.

2. EXPERIMENTAL ENVIRONMENT AND REPRESENTATION OF MODEL ERROR

To investigate the three issues enumerated above, the sequential EnKF of HM01 has been used to assimilate simulated radiosonde, satellite thickness, and aircraft reports into a dry, global, primitive-equation model. The model is a simplified version of the forecast model used operationally at the Canadian Meteorological Centre (Côté et al. 1998). The version of the model used here has 21 levels in the vertical, includes topography, and uses a 144×72 horizontal grid. (This implies a 2.5° grid spacing.) Our version of the model employs the simple forcing and dissipation proposed by Held and Suarez (1994). In total, about 80,000 observations are assimilated per day in the data assimilation experiments.

A method of accounting for model error in an EnKF context was proposed by Mitchell and Houtekamer (2000, hereafter MH) and utilized in the context of a 3-level quasigeostrophic model. Following Dee (1995), the method involved parameterizing the model error and using innovations to estimate the model-error parameters. An ensemble of streamfunction realizations with the specified statistical structure was then generated and added to the ensemble of model predictions, increasing the ensemble spread so as to represent the effect of model error.

It is known (Cohn and Parrish 1991) that if the model error is balanced then the Kalman-filter state estimate will also be balanced. This result motivates the approach taken in this study to extend the method of MH to a primitive-equation context. The method consists of parameterizing the model error, in terms of (i) a horizontal correlation function having a characteristic length scale and (ii) a cosine expansion in the vertical. Such an expansion is motivated by the cosine-like vertical structure of solutions to the linearized primitive equations under certain simplifying conditions (Simmons 1982). An ensemble of streamfunction perturbations, with this statistical structure, is now generated, as before. Each of these streamfunction perturbations is then used to generate an (approximately) balanced model perturbation, consisting of horizontal wind-component (u and v), temperature (T), and surface-pressure (p_s) perturbations. These perturbations are derived from the streamfunction so that the relationship between them is similar to that assumed for the balanced component of the background errors in 3D variational schemes (e.g., Parrish and Derber 1992). This same procedure and statistical description are also used for the generation of an ensemble of initial guess fields. In this study, the model-error statistics, like the observation-error statistics, are assumed to be known. This eliminates the need to (adaptively) estimate these statistics.

3. RESULTS AND CONCLUSIONS

A series of experiments is performed to examine to what extent the localization used in the EnKF produces imbalance. It is found that using a severe localization in the EnKF would cause substantial imbalance in the analyses, as expected. However, as the distance of imposed zero correlation increases to about 3000 km, the amount of imbalance becomes acceptably small.

A series of 14-day data assimilation cycles is performed with different configurations of the EnKF. The effect of varying the ensemble size and the distance of imposed zero correlation are of particular interest. The results are consistent with those obtained in HM98 and HM01. In particular, with respect to ensemble size they clearly indicate the benefits of increasing the ensemble size. With respect to localization, the results indicate that for a given ensemble size there is an optimal value of the localization parameter. The results indicate that the EnKF, with 2×32 ensemble members, performs well in the present context.

To investigate to what extent these encouraging results apply when real observations are assimilated, we intend to combine the sequential EnKF with a more complete version of the operational model. Model-error simulation is expected to be important in such an application.

Further details about the present study can be found in Mitchell, Houtekamer, and Pellerin (2001).

REFERENCES

Cohn, S.E., and D.F. Parrish, 1991: The behavior of forecast error covariances for a Kalman filter in two dimensions. *Mon. Wea. Rev.*, **119**, 1757–1785.

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. Wea. Rev.*, **126**, 1373–1395.

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

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

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

Held, I.M., and M.J. Suarez, 1994: A proposal for the intercomparison of the dynamical cores of atmospheric general circulation models. *Bull. Amer. Meteor. Soc.*, **75**, 1825–1830.

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

Houtekamer, P.L., and H.L. Mitchell, 2001: A sequential ensemble Kalman filter for atmospheric data assimilation. *Mon. Wea. Rev.*, **129**, 123–137.

Keppenne, C.L., 2000: Data assimilation into a primitive-equation model with a parallel ensemble Kalman filter. *Mon. Wea. Rev.*, **128**, 1971–1981.

Mitchell, H.L., and P.L. Houtekamer, 2000: An adaptive ensemble Kalman filter. *Mon. Wea. Rev.*, **128**, 416–433.

Mitchell, H.L., P.L. Houtekamer, and G. Pellerin, 2001: Ensemble size, balance, and model-error representation in an ensemble Kalman filter. *Mon. Wea. Rev.*, submitted.

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

Simmons, A.J., 1982: The forcing of stationary wave motion by tropical diabatic heating. *Quart. J. Roy. Meteor. Soc.*, **108**, 503–534.