4.8 ENSEMBLE KALMAN FILTER IN THE PRESENCE OF MODEL ERRORS

Hong Li[1,*], Eugenia Kalnay[1], Takemasa Miyoshi[2] and Christopher M. Danforth[1]

[1]University of Maryland, College Park, Maryland
[2]Numerical Prediction Division, Japan Meteorological Agency

1. INTRODUCTION

The accuracy of a data assimilation algorithm relies on error statistics of both the model forecast and the observations. The forecast error covariance has a crucial impact on the analysis but it is difficult to estimate. In 3DVAR, a data assimilation scheme used in many operational centers, the forecast error covariance is assumed to be isotropic and stationary. In contrast, ensemble Kalman filters (Anderson 2001, Houtekamer and Mitchell 2001, Whitaker and Hamill 2002, Ott et al. 2004, Hunt et al. 2006) include information on the flow-dependent error of day (both temporally and spatially variant) by estimating the forecast error covariance for each analysis cycle from the difference between the forecast ensemble members and the ensemble mean. The ensemble Kalman filters have been shown to be more accurate than 3D-Var under the assumption of a perfect model (Whitaker et al 2004, Szunyogh et al 2005). However, in reality, forecast errors derive not only from errors in the initial conditions but also from errors due to the model deficiencies. The latter type of error is usually called *model error.* The sources of model error can be due to lack of resolution, approximate parameterizations of physical processes, numerical dispersion, etc. For assimilation of real observations, the assumption of a perfect model should be relaxed. Therefore there is no guarantee that the EnKF will be still better than 3D-Var data assimilation systems when assimilating the real data. In fact, Miyoshi (2005) has shown that model error has a stronger negative influence on the performance of the EnKF than on the 3D-Var. Accounting for model errors associated with model deficiencies has become an important issue for all data assimilation systems, and especially for EnKF.

The Local Ensemble Transform Kalman Filter (LETKF) (Hunt et al 2006) is a relatively new ensemble based data assimilation scheme. It has been implemented to assimilate simulated observations in the NCEP GFS model (Szunyogh et al. 2005), and recently in the NASA fvGCM model (Liu et al. 2006). The results are excellent for a perfect model scenario. With real data, LETKF has been shown to be superior to the operational NCEP 3DVAR in southern Hemisphere and comparable to 3DVAR in northern Hemisphere by using a simple multiplicative inflation method to account for model error (Szunyogh et al. 2006). In order to develop the LETKF into a competitive, operationally applicable data assimilation system, it is necessary to investigate more advanced techniques for treating model errors to further improve the LETKF performance in the realistic weather forecast systems.

The main schemes for treating model errors in EnKF include the Dee and da Silva method (1998, DdS thereafter), 'covariance inflation' (Anderson and Anderson 1999), 'additive inflation' (Whitaker et al. 2006), the Baek et al (2006) high-order method, and the Danforth et al (2006) low-order method. In this study we will investigate DdS, 'covariance inflation', and Danforth et al (2006) low-order method, compare their performances on the LETKF. Though we focus on the LETKF, the results may also be applicable to other Ensemble Kalman Filters.

2. TECHNIQUES FOR TREATING MODEL ERRORS

2.1 Full dimensional augmented state method

Friedland (1969) proposed the "augmented state" method by augmenting the state vector by a model bias vector, to estimate both the state and bias variables. Building on this idea, Dee and da Silva (1998) developed a two-stage estimation algorithm, in which the estimation procedures for the bias and the state are carried out successively. At the first stage of the analysis process, the bias is estimated on every

* *Corresponding author address:* Hong Li 4339 CSS, College Park, MD 20742; e-mail: Lhhelen@atmos.umd.edu

model grid point by assimilating the observed-minus-forecast residuals $y^o - h(x^f)$ as the observed bias:

$$b^a = b^f - K_b[y^o - h(x^f) + h(b^f)] \quad (1)$$

$$K_b = P_{bb}^f H^T (HP_{bb}^f H^T + HP_{xx}^f H^T + R)^{-1} \quad (2)$$

where the matrix P_{bb}^f and P_{xx}^f are the forecast error covariance for the bias and the state variables, respectively. $h(\cdot)$ is the operator mapping the model state variables into observation space. At the second step, the analysis for the state variables is obtained using the standard analysis procedure with the unbiased forecast state $\tilde{x}^f$

$$\tilde{x}^f = x^f - b^a \quad (3)$$

$$x^a = \tilde{x}^f - K_x[y^o - h(\tilde{x}^f)] \quad (4)$$

$$K_x = P_{xx}^f H^T (HP_{xx}^f H^T + R)^{-1} \quad (5)$$

During the forecasting process, the forecast model for bias is assumed to be 'persistent', that is to say there are no dynamics for the bias itself, leading to

$$b_k^f = b_{k-1}^a \quad (6)$$

Estimating bias forecast error covariance P_{bb}^f is very difficult. It has been assumed to have the same spatial structure as the state forecast error covariance but with different covariance localization scales (Keppenne et al 2005). The cost of solving the equations above is roughly double that of without estimating bias since the updated equations are solved twice, first for the bias estimation and then for the state variables. A simplified version of the algorithm was described in Radakovich et al (2001) and Todling 2004, where the order of the bias estimation step is reversed with respect to that of the state analysis, and the bias gain is approximated by a multiple α (which is a tunable coefficient) of the analysis matrix. The simplified algorithm is as follows:

$$\tilde{x}^f = x^f - b^f \quad (7)$$

$$\delta x^a = K_x[y^o - h(\tilde{x}^f)] \quad (8)$$

$$b^a = b^f - \alpha \delta x^a \quad (9)$$

$$b_t^f = \mu * b_{t-1}^a \quad (10)$$

Equation (10) is the bias forecast model where μ is a tunable coefficient. μ =1.0 represents the 'persistent' model. Since there is no extra computation for the bias gain, the simplified DdS scheme is very efficient.

2.2 Low order bias estimation scheme

Danforth et al (2006) assume that model error is composed by the state-independent component and the state-dependent component.

$\varepsilon^f =$ state-independent error + state-dependent error.

The state-independent model error component can be represented by the mean bias $\mathbf{b}$ and the leading EOFs $\mathbf{e}_l$ from the anomalous model error field which is not represented in the mean bias. The state-dependent component is given by the leading SVDs $\mathbf{f}_m$ of the covariance of the coupled model state anomalies and corresponding measured errors.

$$\varepsilon^f = \mathbf{b} + \sum_{l=1}^{L} \beta_l \mathbf{e}_l + \sum_{m=1}^{M} \gamma_m \mathbf{f}_m \quad (11)$$

where L and M are the number of leading modes of EOFs and SVDs , respectively. The base fields $\mathbf{b}$, $\mathbf{e}_l$ and $\mathbf{f}_m$ are pre-computed using the samples in the training period. The variables need to be estimated on-line are the amplitudes β_l and γ_m which have a much lower dimension than the full model dimension. Danforth et al (2006) have performed this procedure in the 5-day forecast using 5 years as training period. In their case, the initial condition is from the NCEP reanalysis taken as "truth", therefore no data assimilation procedure is required.

In this study we extend their work to a more realistic case. Both the random errors in the initial condition and the model errors are considered and treated in each data assimilation cycle.

2.3 Covariance inflation

Covariance inflation has been widely used to prevent filter divergence due to an underestimation of the true forecast error covariance even with a perfect model. In the presence of model error, 'covariance inflation' is also a simple and straightforward method to account for model errors. In theory, the forecast error covariance should be given by:

$$\mathbf{P}_t^f = \mathbf{M}\mathbf{P}_{t-1}^a \mathbf{M}^T + \mathbf{Q} \quad (12)$$

where **M** is tangent linear model and **Q** is the model error covariance. In ensemble Kalman filters, the first term in the forecast error covariance above is estimated from the ensemble perturbation and the second term is ignored:

$$\mathbf{P}_t^f = \frac{1}{K-1}\sum_{k=1}^{K}(x^f - \bar{x}^f)*(x^f - \bar{x}^f)^T \quad (13)$$

When the model error is significant, ignoring the **Q** term will cause the forecast error covariance to underestimate the true forecast error and divert the analysis away from the truth. To account for the **Q** term in a simple way, we inflate the covariance matrix P_t^f, which is estimated from (13), by using $P_t^f = (1+\Delta)*P_t^f$, where Δ is a number called the inflation coefficient.

Comparing with the first two model error correction methods where the model error is directly removed from the forecast, the covariance inflation can be regarded as second order model error estimation scheme. In principle, if the model error can be exactly estimated, directly removing the model error should result in better analysis than parameterization of the model error covariance **Q** term in covariance inflation

3. IMPLEMENTATION ON THE LETKF IN THE PRESENCE OF THE MODEL ERRORS

3.1 The SPEEDY model

The SPEEDY model (Molteni 2003) is a recently developed atmospheric general circulation model (AGCM) with simplified physical parameterization schemes that are computationally efficient, but that maintain the basic characteristics of a state-of-the-art AGCM with complex physics. It has a spectral primitive-equation dynamics and triangular truncation T30 at 7 sigma levels.

3.2 Evidence of model error

First we investigate the SPEEDY model bias against the NCEP/NCAR reanalysis (NNR) fields (Kalnay et al 1996) as the “truth”. The SPEEY model is integrated from NNR initial conditions every 6 hours. Figure 1 shows the differences between the SPEEDY 6hr forecasts and the NNR verified at the same time, averaged over two months in the period from January 1, 1987 to February 28, 1987, for the zonal wind and height at 500 hPa. The largest model bias of the u-wind can be seen in the polar regions. Orographic effects are a major originator for the systematic errors in the height field.

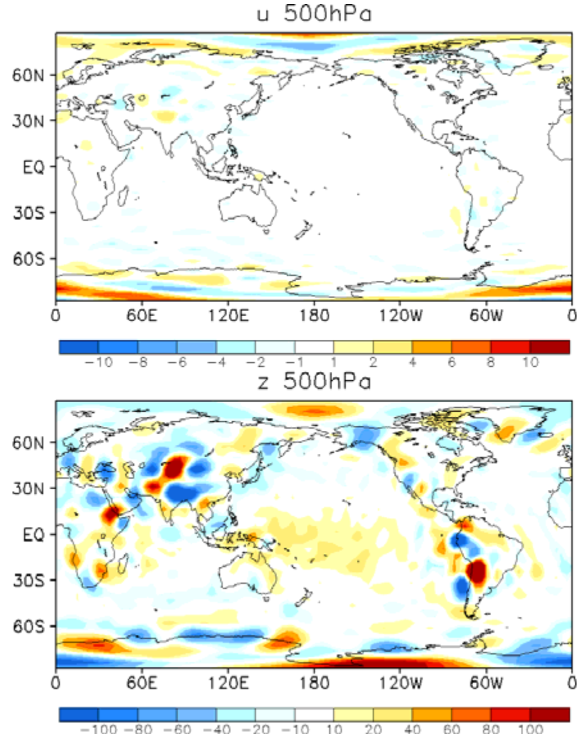

Fig. 1 Time-mean SPEEDY 6-hour model error of u-wind (top, unit: m/s) and height (bottom, unit: m) at 500hPa against NNR reanalysis fields for the period of January and February in 1987.

3.3 LETKF data assimilation scheme

LETKF is an ensemble square-root filter in which the observations are assimilated to update only the ensemble mean (14), while the ensemble perturbations are updated by transforming the forecast perturbations through a transform matrix (15) introduced by Bishop et al (2001). The basic formulas used in the LETKF (Hunt et al. 2006) are

$$\bar{x}^a = \bar{x}^b + X^b\tilde{P}^a(HX^b)^T R^{-1}[y^o - h(\bar{x}^b)] \quad (14)$$

$$X^a = X^b[(k-1)\tilde{P}^a]^{1/2} \quad (15)$$

Here X^a, X^b are the analysis and forecast ensemble perturbations, respectively (matrices whose columns are the difference between the ensemble members and the ensemble mean). The transform matrix $\tilde{P}^{a1/2}$ is the square-root of matrix $(k-1)\tilde{P}^a$ where

$\tilde{P}^a$, the analysis error covariance in ensemble space, is given by

$$\tilde{P}^a = \left[(k-1)I + (HX^b)^T R^{-1}(HX^b)\right]^{-1} \qquad (16)$$

It has dimension k by k, where k is the ensemble size, which is generally much smaller than both the dimension of the model and the number of observations. Thus, the LETKF performs the analysis in the space spanned by the forecast ensemble members, which greatly reduces the computational cost. Furthermore, since the analysis is computed independently at each grid point, the LETKF computation can be performed in parallel.

3.4 Experimental setup

The observations are obtained by adding zero mean normally distributed noise to the NNR fields. With respect to these observations, the SPEEDY model has significant model errors (shown in section 3.2). The observations are available on the model grid at every 2 grid points.

First the *control run* is performed, in which the LETKF is used to correct the random errors in the initial condition and no additional method is used to correct the model error. Then the LETKF is combined with each of several bias estimation methods to correct the forecast errors, which are due to both inaccurate initial conditions and model errors. To compare the performances of model error estimation methods, the initial one month is considered as the spin-up period, and the second month is used for verification.

For the low-order model error estimation scheme, only the first term in equation (11) has been considered. The time mean bias is estimated by the difference between 6-hour SPPEEDY forecast initiated by NNR and NNR at the verification time, averaged over the training period. Two different training periods have been tested. One is January and February, 1982-1986, the same period used by Danforth et al (2006), the other is the one month prior to the experimental month, for example December 1986 is used as the training period for the January 1987 experiment, and January 1987 as the training period for the experiment in February 1987.

4. PRELIMINARY RESULTS

Table 1 summarizes the ensemble-mean analysis error spatially averaged over the globe and temporally averaged over February 1987 for 500 hPa height using the dense observational network. All experiments use LETKF with 10 ensemble members. Unlike the other experiments, the "perfect model" experiment assimilates the observations generated from a nature run created with the same model, rather than the NNR. The RMSE in the control run, where no model error treatment is applied, is an order of magnitude larger than that in the perfect model experiment, indicating that LETKF is very vulnerable to model errors. In real data assimilation the situation may not be as bad because the model deficiencies in a sophisticated operational model should be much less dominant than that in SPEEDY.

Table 1 Global ensemble-mean analysis error of 500 hPa height fields from the control run (no model error estimator) and from applying different model error correction methods.

EXP	RMSE (m)
Control run	78.3
Approximate DdS (u=1.0,a=0.5)	Div
Approximate DdS (u=0.95, a=0.5)	67.1
20% covariance inflation	68.6
Low order (training Jan & Feb ,1982-86)	66.4
Low order (training 1 month prior)	50.1
Perfect model	5.5

If we correct the model error in each data assimilation cycle using three bias correction methods (covariance inflation, simplified DdS and low-order method), except for the simplified DdS algorithm with u=1.0, all the bias estimation methods succeed in reducing the analysis error at 500 hPa height. For the low-order method, the results show that removing the time-mean bias estimated from the recent training period is better than that from the long climatic period. The best result in table 1 shows a reduction of the RMSE error of 28 meters of 500 hPa geopotential height but it is still much worse than the perfect model case.

Fig. 2 shows the RMSE with time for other variables at 500 hPa. In general, the low-order method works better than the simplified DdS algorithm except for humidity. Covariance inflation works well also in temperature and humidity but not in geopotential height. The results at 500 hPa are true for other pressure levels below 200 hPa. Above 200 hPa, the analysis from the low-order method is bigger than that from the control run (Fig. 3).

5. DISCUSSON AND FUTURE WORK

We have tested three methods to treat model errors. In general, all of them can estimate and correct model error but just partially. Low-order method works better than simplified DdS and covariance inflation, but at high level it is worse.

For the DdS scheme, we have only tested its simplified version. We will try the original DdS scheme which is less efficient but should be more accurate than its simplified version.

For the low-order method, our results have shown that simply subtracting the constant mean bias from the background fields at every analysis cycle has a significant positive impact on the LETKF. In our future work, we will correct the diurnal bias and then the state-dependent bias. Retaining each component of the bias estimate, we should be able to get a good estimate of the true model errors. If successful, this approach could not only improve the performance of the LETKF, but also of the forecast, as well as providing information on model errors useful for diagnostic purposes.

REFERENCES

Anderson, J. L. and Anderson, S. L. 1999: A Monte Carlo implementation of the nonlinear filtering problem to produce ensemble assimilations and forecasts. *Mon. Wea. Rev.* **127,** 2741–2758.

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

Baek, S.J., B. R., Hunt, E. Kalnay, E. Ott, and I. Szunyogh 2006: Local ensemble Kalman filtering in the presence of model bias, *Tellus,* **58A,** 293-306.

Bishop, C. H., Etherton, B. J., and Majumdar, S. J., 2001: Adaptive Sampling with the Ensemble Transform Kalman Filter. Part I: Theoretical Aspects. *Monthly Weather Review*, **129**, 420-436.

Danforth, C., E. Kalnay, and T. Miyoshi, 2005: Estimating and Correcting Global Weather Model Error. In preparation.

Dee, D. P. and A. M. Da Silva, 1998: Data assimilation in the presence of forecast bias. *Quart. J. Roy. Meteor. Soc.*, **124**, 269-295.

Friedland, B. 1969: Treatment of Bias in Recursive Filtering. IEEE *Transaction on Automatic Control*, **AC-14**, 359-367.

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

Houtekamer, P.L, Mitchell, H.L., Pellerin, G., Buehner, M., Charron, M., Spacek, L., and Hansen, B., 2005: Atmospheric data assimilation with the ensemble Kalman filter: Results with real observations. *Monthly Weather Review*, **133**, 604-620.

Hunt, B. R., E. Kostelich, and I. Szunyogh, 2006: Efficient Data Assimilation for Spatiotemporal Chaos: a Local Ensemble Transform Kalman Filter, In preparation.

Kalnay, E., M. Kanamitsu, R. Kistler, W. Collins, D. Deaven, L. Gandin, M. Iredell, S. Saha, G. White, J. Woollen, Y. Zhu, M. Chelliah, W. Ebisuzaki, W. Higgins, J. Janowiak, K. C. Mo, C. Ropelewski, J. Wang, A. Leetmaa, R. Reynolds, Roy Jenne, and Dennis Joseph, 1996: The NMC/NCAR 40-Year Reanalysis Project. *Bull. Amer. Meteorol. Soc,* **77**, 437–471.

Liu, J., E. Fertig, H. Li, I. Szunyogh, B. R. Hunt, E. Kalnay, E. Kostelich, and R. Todling, 2006: Comparison between Local Ensemble Transform Kalman Filter and PSAS in NASA finite volume GCM. In preparation.

Miyoshi, T. 2005: Ensemble Kalman filter experiments with a primitive-equation global model, PhD Thesis at the University of Maryland.

Molteni, F., 2003: Atmospheric simulations using a GCM with simplified physical parametrizations. I: Model climatology and variability in multi-decadal experiments. *Climate Dyn.*, **20**, 175-191.

Ott, E., Hunt, B. R., Szunyogh, I., Zimin, A. V., Kostelich, E.J., Kostelich, M., Corazza, M., Sauer, T., Kalnay, E., Patil, D. J., and Yorke, J. A., 2004: A Local Ensemble Kalman Filter for Atmospheric Data Assimilation. *Tellus*, **56A**, 415-428.

Radakovich, J. D., P.R. Houser, A. M. da Silva, and M. G. Bosilovich, 2001: Results from global land-surface data assimilation methods. Pp. 132-134 in proceedings of the fifth symposium on integrated observing systems, 14-19 January 2001, Albuquerque, NM.

Szunyogh, I., E. J. Kostelich, G. Gyarmati, D. J. Patil, B. R. Hunt, E. Kalnay, E. Ott, and J. A. Yorke, 2005: Assessing a Local Ensemble Kalman Filter: Perfect Model

Experiments with the NCEP Global Model. *Tellus*, **57A**, 528-545.

Szunyogh, I., E. J. Kostelich, G. Gyarmati, E. Kalnay, B. R. Hunt, E. Ott, and J. A. Yorke, 2006: Assessing a Local Ensemble Kalman filter: Assimilating real observations with the NCEP global model. In preparation.

Todling R., 2004: Producing Unbiased Analyses in the presence of forecast bias. First LNCC meeting on computational modeling.

Whitaker, J. S., G. P. Compo, X. Wei, and T.H. Hammil, 2004: Reanalysis without Radiosondes using Ensemble Data Assimilation. *Monthly Weather Review*, **132**, 1190-1200.

Whitaker, J. S., Hamil, T. H., Wei, X., Song, Y. C., and Toth, Z., 2006: Ensemble data assimilation with NCEP global forecast system. Submitted to *Mon. Wea. Rew*

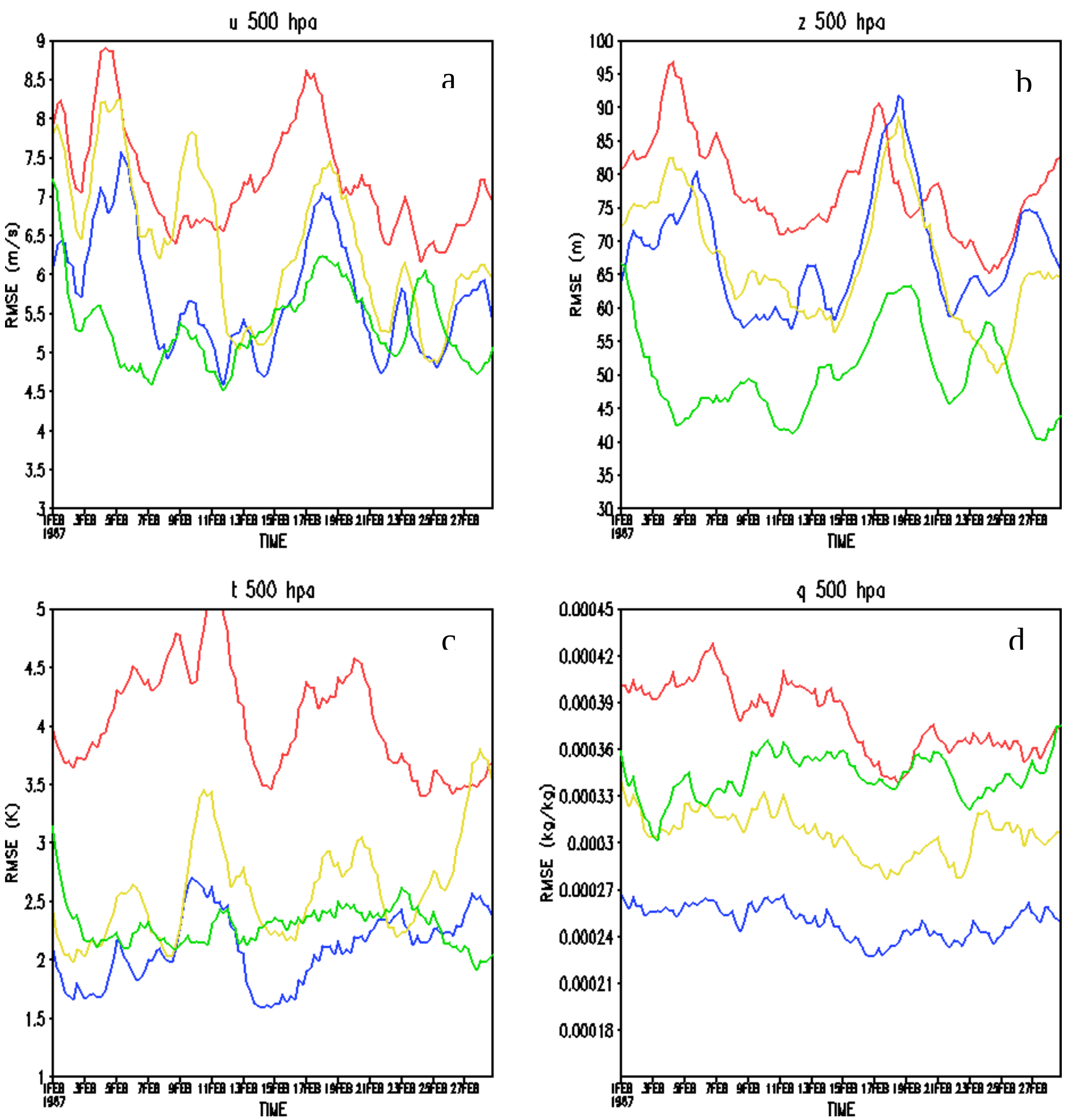

Fig. 2 Time series of global ensemble-mean analysis RMSE at 500 hPa using LETKF with 10 members to assimilate observations derived from the NNR. RMSE for different model error correction methods are shown in red (control run), blue (covariance inflation), yellow (simplified DdS) and green (low-order), for zonal wind (a), height (b), temperature (c) and specific humidity (d), respectively.

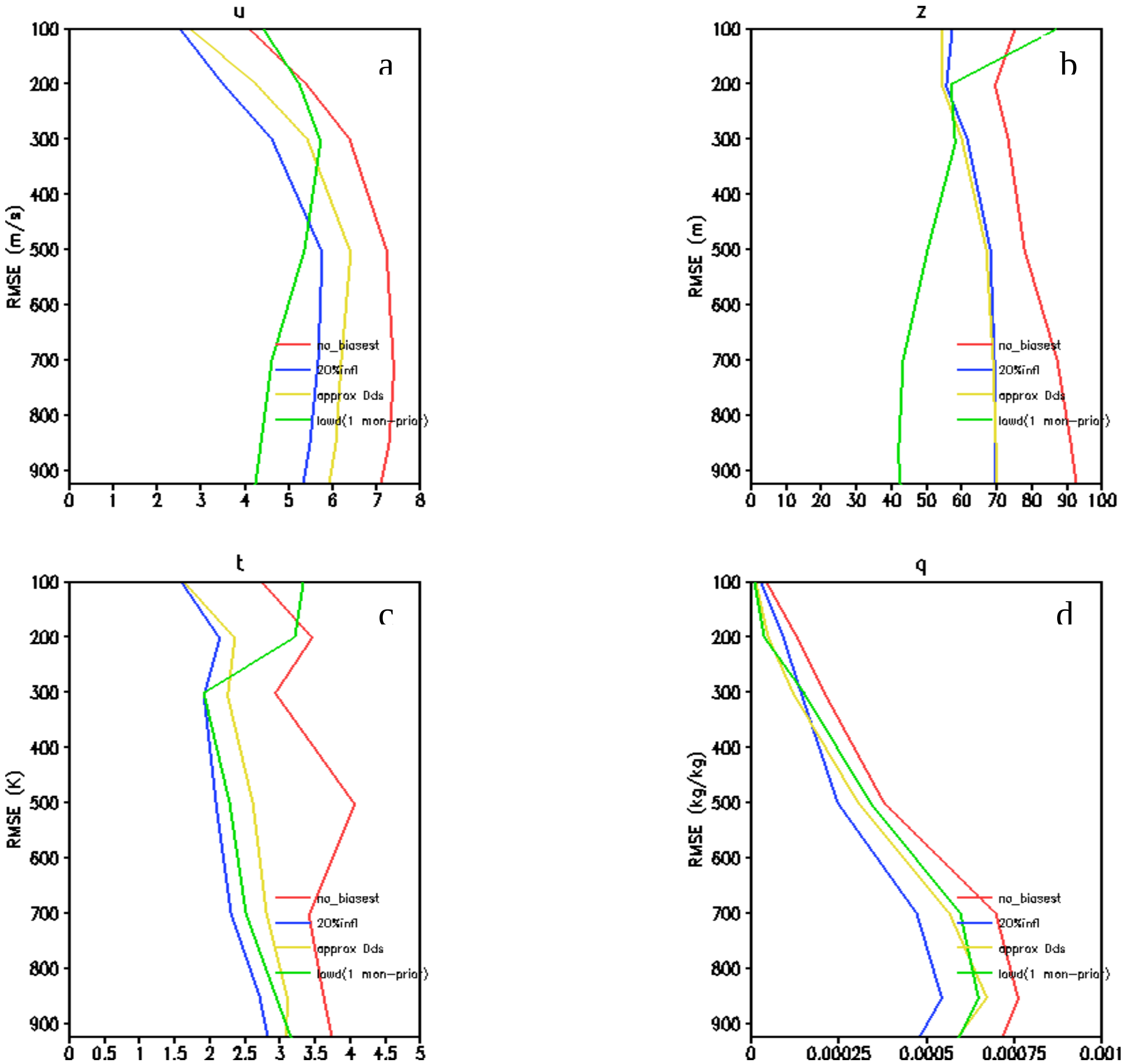

Fig.3 Global ensemble-mean analysis RMSE at all SPEEDY model levels using LETKF with 10 members to assimilate observations derived from the NNR. RMSE for different model error correction methods are shown in red (control run), blue (covariance inflation), yellow (simplified DdS) and green (low-order), for zonal wind (a), height (b), temperature (c) and specific humidity (d), respectively.