Optimization of Error Covariance Matrices and

Estimation of Observation Data Impact in the JMA Global 4D-Var System

Toshiyuki Ishibashi
Numerical Prediction Division, Japan Meteorological Agency
1-3-4 Otemachi, Chiyoda-ku, Tokyo 100-8122, Japan
E-mail: ishibasi@met.kishou.go.jp

1. Introduction

Since data assimilation theories based on maximum likelihood estimation such as 4D-Var have error covariance matrices (ECMs) as external parameters, real data assimilation systems (DASs) must have ECMs with a sufficient level of accuracy. The validity of ECMs is one of the main factors in determining the accuracy of analysis fields because they strongly influence the effects of observational data and background fields on analysis fields.

Two methods can be used for ECM optimization. The first is based on Desroziers and Ivanov (2001, QJRMS) and Chapnik (2006, QJRMS), and is referred to here as the DIC method. In DIC, ECMs are tuned to satisfy the theoretical relationships between ECMs and the cost function of 4D-Var, and the method includes evaluation of observation data impacts (ODIs) on analysis fields because it includes calculation of degrees of freedom for signals (DFS) (Rabier, 2002, QJRMS) used in ODI estimation. The second is a method based on Daescu (2008, MWR), and is referred to here as D08. Although this approach is not an ECM optimization method in itself, it indicates the direction in which ECMs should be tuned, since it calculates forecast error dependencies on ECMs. D08 also includes ODI estimation because it is an extension of Langland and Baker (2004) (LB04), which is an ODI estimation method involving the calculation of forecast error dependencies on observational data. Both of these ECM optimization methods therefore include ODI estimation.

This paper describes the development of ECM optimization and ODI estimation in the JMA global 4D-Var system.

2. ECM optimization and ODI estimation with DIC

DIC is an ECM optimization scheme that uses the theoretical relationships between ECMs and the cost function of 4D-Var as follows:

$$2J = N - TR(\mathbf{HK}) = N - DFS \qquad (1)$$

Here, $\mathbf{K}$ is the Kalman gain, $\mathbf{H}$ is a partial derivative of the observation operators, TR is a trace operator, N is the number of observation data, and J is the cost function of 4D-Var. This equation is valid for each block diagonal part of the observation ECM, defined as $\mathbf{R}$, and can be used to optimize $\mathbf{R}$ for each observation dataset. We implemented the DIC method on the JMA global 4D-Var system.

The results of DIC optimization show that diagonal components of the optimized $\mathbf{R}$ are about 30% of current settings for most of satellite radiance data, while are comparatively resemble the current settings for other conventional data. These results are consistent with departure value (observation minus guess) statistics, and the optimization recovers the theoretical relationships between cost functions and observation data numbers (Ishibashi, 2006). DIC includes ODI estimation as described in the introduction, since Equation (1) includes DFS. However, since this equation is correct only if DAS is optimal, DFS is shown after DIC optimization in Figure 1 (reproduced from Ishibashi, 2006). It can be seen that the contributions of radiance data and conventional data to analysis accuracy are about the same.

DIC enables determination of ECMs objectively rather than by trial and error. However, data assimilation cycle experiments with tuned ECMs using DIC (figures not shown) show that there are still several factors for consideration to improve analysis and forecast accuracy, including observation error correlation and the biases of the NWP model.

3. ECM optimization and ODI estimation with D08

D08 is a method used to calculate the dependencies of forecast errors on ECMs. It consists of three parts: the first is calculation of the forecast error sensitivity field (SF); the second is calculation of the dependencies of SF on observational data; the third is calculation of the dependencies of SF on ECMs. The first two parts are the same as the calculation of LB04, and the last one is a new addition in D08. This means that LB04 is included in the D08 approach.

LB04 is an ODI estimation method that calculates ODI as a forecast error reduction by assimilating observational data using adjoint operators of the forecast model and DAS. As the construction of an adjoint operator for DAS requires large changes to DAS, the implementation cost of LB04 is not small. We divide the adjoint operators of DAS into two-step linear problems and solve them using fixed original 4D-Var code (Tre'molet, 2008, TELLUS).

We evaluate the errors of a 15-hour forecast in terms of dry total energy (TE), and then calculate its sensitivity field (SF). To evaluate the validity of the calculated SF, we construct an approximate analysis error field from the SF through multiplication by a scalar coefficient, as a normalized SF is a good approximation of the analysis error field (Rabier, 1996, QJRMS). We then make an optimal initial field by extracting the approximate analysis error field from the original initial field. Figure 2 shows the error of the forecast from the optimal initial field and from the original initial field. As an explicit forecast error reduction can be seen from the optimal initial field, our SF calculation can be considered valid.

Figure 3 shows the results of LB04 using the SF. It indicates that the largest contributions to forecast error reduction are brought by AMSU-A sensors, with the next being radiosonde data in the JMA 4D-Var system. The contributions from the sum of all satellite radiance data and the sum of the remaining data (e.g., radiosondes and satellite winds) are compatible. However, comparing the results with those of ECMWF suggests that satellite contribution is rather small as hyper-spectral sounders (AIRS and IASI) and GPS data are not used (or are not enough used) in the JMA 4D-Var system, and also because the **R** settings for radiance data are too large. We also check the effects of the norm form used in the forecast error specification. If we use wet TE, the water vapor channel contributions increase slightly (figures not shown).

To check the ability of LB04 to detect erroneous observation data, we implemented a very small observation error setting for Channel 8 of AMSU-A/METOP in 4D-Var. LB04 detects forecast error increases from this data (Figure 4).

Finally, we construct a D08 system by extending the LB04 system, and evaluate the dependencies of forecast error on the observation ECM, **R**. These calculations are given by the following equation:

$$\partial J/\partial R_{p,q} = \partial J/\partial y_p \sum_s R^{-1}_{q,s}\left(Hdx-d\right)_s \quad (2)$$

Here, J is the forecast error in TE, p, q and s are indices for observational data, y is the observation value and *dx* is the analysis increment. Eventually, the calculation only multiplies the coefficients by the LB04 results. The results of D08 show that reduction of the observation error setting for most radiance data will reduce the level of forecast error (Figure 5). These results are consistent with those of the DIC approach described in Section 2. In fact, a 20-percent reduction of these observational data errors leads to the forecast error reduction shown in Figure 6.

4. Conclusions

Here we have reported on the development of ECM optimization and ODI estimation in the JMA global 4D-Var system. Further development and improvement of these methods, including measures such as the introduction of observation error correlations to DIC, are planned to reduce analysis and forecast errors in the future.

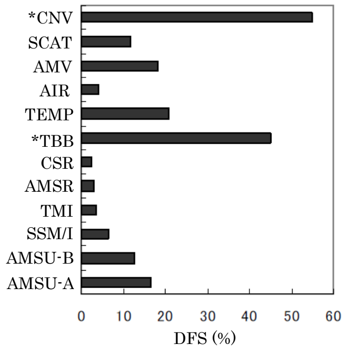

Figure 1. DFS (ratio to total DFS in %) for the JMA global 4D-Var system after DIC optimization. *TBB denotes all radiance data, *CNV denotes all data except *TBB, SCAT denotes QUIKSCAT, AIR denotes airplane, TEMP denotes radiosonde, CSR denotes geostationary satellite radiance, and the others denote individual satellite sensor names. This figure is adapted from Ishibashi (2006).

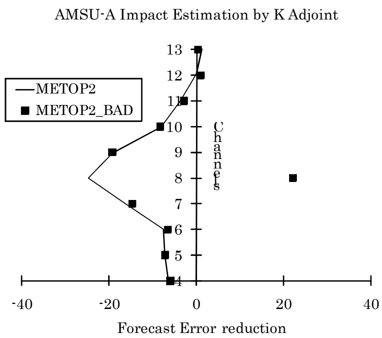

Figure 4. Example of erroneous channel detection by LB04. The solid black squares show the ODI estimation of LB04 for a case in which Channel 8 of AMSU-A/METOP2 is given an incorrectly small observation error setting. The solid line shows the case of routine setting for comparison.

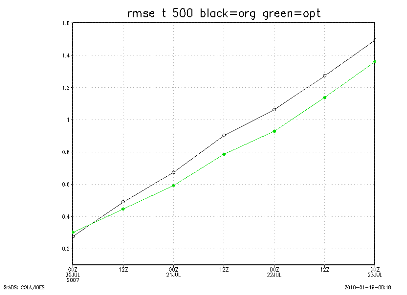

Figure 2. Root mean square error (RMSE) for 500-hPa temperature of forecasts from the original initial field (black line) and from the optimized initial field (green line)

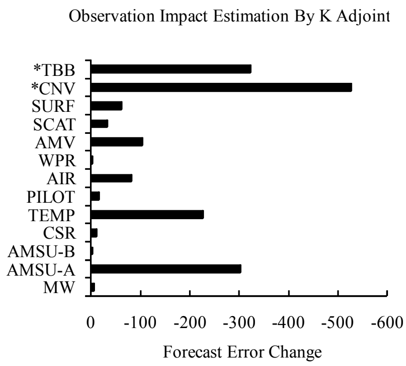

Figure 3. Results of LB04 ODI estimation in JMA's global 4D-Var DAS calculated for the initial, 00 UTC on July 20, 2007

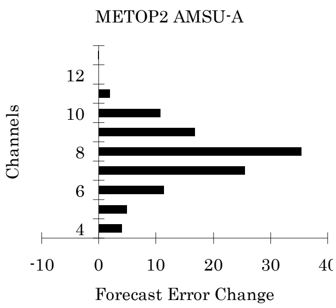

Figure 5. Results of D08 – forecast error change (measured in total energy) caused by perturbing 1 (K) of observation error standard deviation in the JMA global 4D-Var system for AMSU-A/METOP2

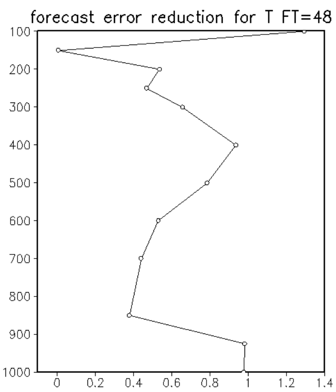

Figure 6. Forecast error reduction by **R** tuning according to the results of D08 estimation. The figure shows the global average of the normalized RMSE difference (routine setting minus tuned setting for temperature).