Tuning of observation and background error variances using innovation-based diagnostics and estimation

Environment Canada Environnement Canada

Y. Yang R. Ménard and Y. Rochon Air Quality Research Division, Environment Canada

Introduction

Accurate estimation of the observation and background error statistics plays an important role in data assimilation as they determine, at analysis time, the weight and spatial influence function of observations and possibly the impact on other variables. The observation error which is useful for data assimilation is best estimated within an assimilation cycle. On the other hand, the background error, or short term forecast error, is not independent of the observations and can also only be estimated within an assimilation cycle. In today's 3D-var or 4D-var assimilation systems, however, the observation error and background error are prescribed. An improper characterization of the observation and background error statistics will lead to a suboptimal assimilation scheme. Desrosiers and Ivanov(2001) and Desrosies et al(2005) developed a method to tune observation and background error using the diagnosis computed from analysis residuals. In this work we apply the method to the 3D-var assimilation system of Canadian Meteorological Center (CMC) and use it to tune the observation error iteratively for dynamic variables and then tune the background error for chemistry variables.

Innovation-based diagnostics and estimation

The χ^2 diagnostic is a measure of consistency between the variances of random variables. This diagnostic has been used in many applications such as geophysics (Tarantola, 1987), atmospheric retrievals (Rodgers 2000), and data assimilation (Bennett and Thornburn 1992, Talagrand 1999, Ménard and Chang 2000) where the random variable is a residual or innovation, i.e. the difference between observations and the model equivalent (at the same time and location). For data assimilation χ^2 is defined as

$$\chi^2 = \mathbf{d}^T \mathbf{\Gamma}^{-1}\mathbf{d} \quad , \quad \mathbf{d} = \mathbf{y} - \mathbf{H}\mathbf{x}^f \qquad (1)$$

where **d** is the innovation, and

$$\mathbf{\Gamma} = \mathbf{HBH}^T + \mathbf{R} \qquad (2)$$

is the a priori innovation covariance, **B** is the prescribed background error covariance and **R** is the prescribed observation error covariance and **H** is the observation operator. The expected value of χ^2 is given as

$$\langle \chi^2 \rangle = \langle \mathbf{d}^T \Gamma^{-1} \mathbf{d} \rangle = \mathrm{trace}\left(\mathbf{\Gamma}^{-1}\overline{\mathbf{\Gamma}}\right) \qquad (3)$$

Where $\overline{\mathbf{\Gamma}} = \langle \mathbf{d}\mathbf{d}^T \rangle$ is the sample covariance of the innovations. If the sample covariance of the innovation matches the given or prescribed innovation covariance, i.e. $\mathbf{\Gamma} = \overline{\mathbf{\Gamma}}$, then

$$\langle \chi^2 \rangle = m \qquad (4)$$

where m is the dimension of the observation space or the number of observations.

In 3D and 4D-Var, the value of χ^2 can be obtained directly from the value of the cost function at the minimum as follows

$$J_{\min} = J^b_{\min} + J^o_{\min} = \tfrac{1}{2}\mathbf{d}^T\left(\mathbf{HBH}^T + \mathbf{R}\right)\mathbf{d} = \tfrac{1}{2}\chi^2 \qquad (5)$$

Note that,

Condition 1: The sample covariance of the innovation matches the given or prescribed innovation covariance, i.e. $\mathbf{\Gamma} = \overline{\mathbf{\Gamma}}$

is a necessary condition to meet the $\chi 2$ diagnostic (4), but it is not a sufficient condition.

As an extension of the χ^2 diagnostics, Desrosiers and Ivanov (2001) developed a method to tune observation error and background error parameters by comparing the J^b and J^o at the minimum against what is estimated using a randomized trace method. Recently, Desrosiers et al. (2005) proposed a simpler more direct approach to estimate observation and background error parameters. It is noted that

$$\begin{aligned}\langle OmA\,(OmF)^T \rangle &= \langle (\mathbf{y} - \mathbf{Hx}^a)(\mathbf{y} - \mathbf{Hx}^f)^T \rangle \\ &= \langle (\mathbf{d} - \mathbf{HKd})\mathbf{d}^T \rangle = (\mathbf{I} - \mathbf{HK})\langle \mathbf{dd}^T \rangle \\ &= \mathbf{R}\mathbf{\Gamma}^{-1}\overline{\mathbf{\Gamma}} \end{aligned} \qquad (6)$$

$$\begin{aligned}\langle AmF\,(OmF)^T \rangle &= \langle (\mathbf{Hx}^a - \mathbf{Hx}^f)(\mathbf{y} - \mathbf{Hx}^f)^T \rangle \\ &= \langle \mathbf{HKd}\,\mathbf{d}^T \rangle = \mathbf{HK}\langle \mathbf{dd}^T \rangle \\ &= \mathbf{HBH}^T\mathbf{\Gamma}^{-1}\overline{\mathbf{\Gamma}} \end{aligned} \qquad (7)$$

If *Condition 1* is fulfilled, then

$$\langle OmA\,(OmF)^T \rangle = \mathbf{R} \qquad (8)$$

$$\langle AmF\,(OmF)^T \rangle = \mathbf{HBH}^T \qquad (9)$$

Also from (8) and (9) we have

$$\langle OmF\,(OmF)^T \rangle = \langle OmA\,(OmF)^T \rangle + \langle AmF\,(OmF)^T \rangle = \mathbf{R} + \mathbf{HBH}^T \qquad (10)$$

These diagnostics, (6-7), defined in observation space can be directly computed from the analysis residuals, and do not require extra computations.

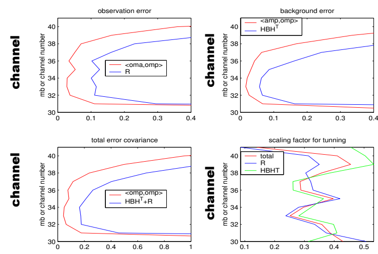
This algorithm can be applied iteratively. A scalar case study proved that the iteration scheme will always converge. However, if one only tunes the observation (background) error iteratively without tuning the other, the accuracy of the converged value depends on the accuracy of the prescribed background (observation) error. In a realistic system similar to an operational assimilation system, there is no known case of non-convergence except where **R** and **B** have the same correlation length. Here we will show the result of iterative tuning.

Model and experiments

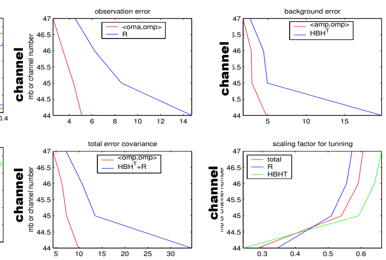
For dynamic assimilation, an earlier version of CMC's GEM-Strato model and 3D-Var system was used to assimilate dynamic measurements, including TOVS (AMSUA and AMSUB), RAOBS, and others. Background error statistics were obtained using the NMC method. AMSU observations were bias-corrected before assimilation. This run is for winter 2003.

For chemistry assimilation, we used an updated model of GEM-Strato, which incorporats the BIRA (Belgian Institute for Space Aeronomy) chemistry module. This experiment assimilates MIPAS measurements of CH4, with dynamic fields refreshed from another run that assimilated dynamics data mentioned above. This run was for summer 2003.

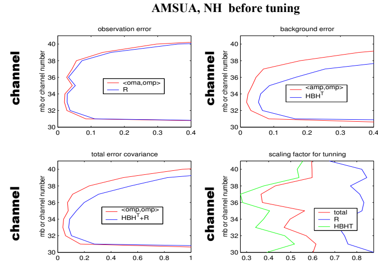
Diagnostics of dynamics assimilation runs before and after tuning observation error variances



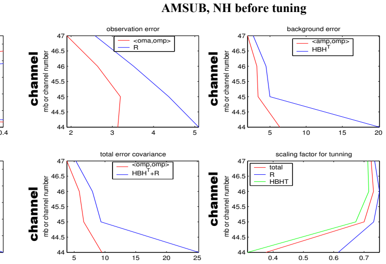
AMSUA, NH before tuning



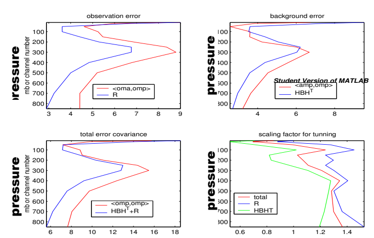
AMSUB, NH before tuning



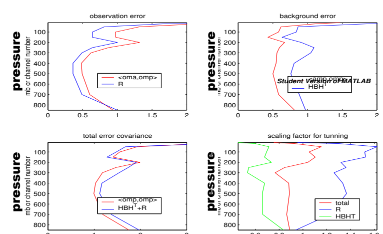
AMSUA, NH after tuning obs. error variances



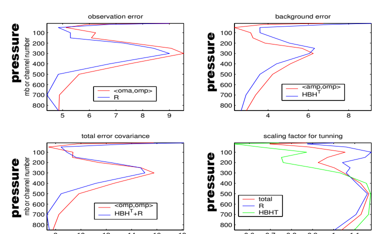
AMSUB, NH after tuning obs. error variances



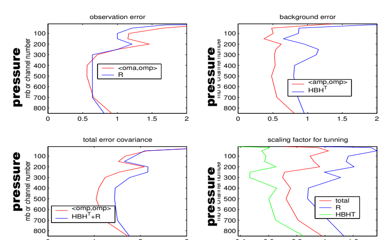
RAOBS UU, NH before tuning



RAOBS TT, NH before tuning

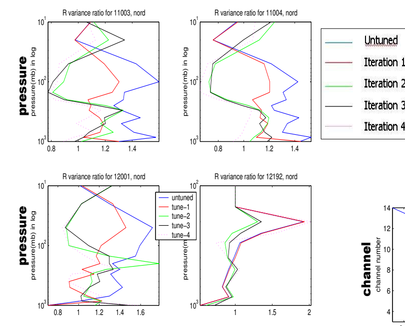


RAOBS UU, NH after tuning obs. error variances

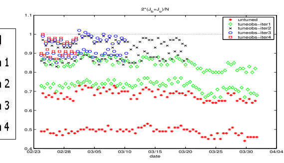


RAOBS TT, NH after tuning obs. error variances

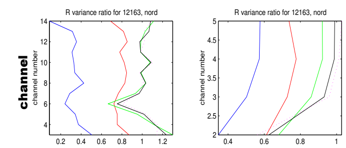
Results of iteratively tuning the observation error variances



Ratio of <OmA, OmP>/R for RAOBS

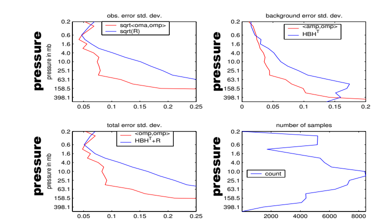


χ^2 diagnostic

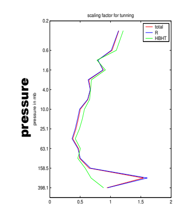


Ratio of <OmA, OmP>/R for AMSUA &B

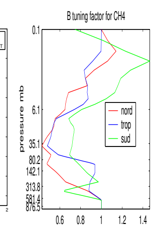
Tuning background error for MIPAS CH4



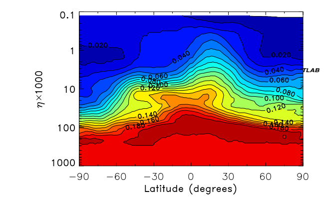
MIPAS CH4 before tuning



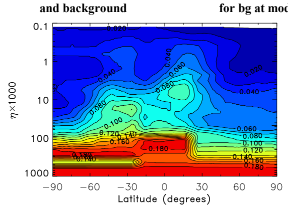
Tuning coefficient for obs. and background



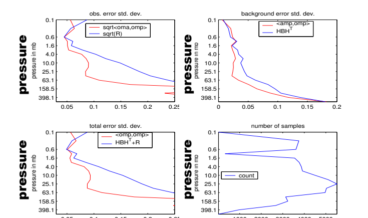
Tuning coefficient for bg at model level



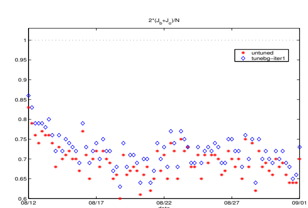
Zonal mean std. dev. of CH4 before tuning



Zonal mean std. dev. of CH4 after tuning



MIPAS CH4 after tuning



χ^2 diagnostic

Summary

Error statistics derived from innovations provide an estimate of the errors perceived by the assimilation system, with which error statistics prescribed in a 3Dvar system has to be consistent. Two diagnostic methods, namely the χ^2 test and Desrosiers et al's innovation-based consistency diagnostics are implemented and the diagnostic results are being used to tune the observation and background error variances for use with CMC's 3Dvar assimilation system.

Through iterative tuning of AMSU and RAOBS observations, the ratio of <OmA, OmP>/R are generally converging toward 1.0, and each iteration yields better consistency than the former one, especially for AMSU observations. For certain RAOBS variables at certain levels the iteration scheme does not seem to converge.

For dynamic observations, the iterative tuning consistently improves the χ^2 test result. The improvement is very consistent in time. The χ^2 values after the third and forth tuning are very close to 1.0. Whereas for CH4, by tuning the background error variances the improvement on χ^2 is consistent in time but the impact is vary limited.