# Convergence and stability of estimated error variances derived from assimilation residuals in observation space

**Richard Ménard, Yan Yang and Yves Rochon**

*Air Quality Research Division, Environment Canada*
*Richard.Menard@ec.gc.ca*

## 1. Introduction

The estimation of error statistics using assimilation residuals *A-F* and *O-A* in observation space along with *O-F* have received considerable attention in recent years. The method originally developed by Desrosiers and Ivanov (2001) and further developed by Chapnik et al. (2004, 2006) and Desrosiers et al. (2005) used randomization techniques to estimate the $\mathrm{Tr}(\mathbf{HK})$ and $\mathrm{Tr}(\mathbf{HK})$, and had the advantage of being capable of deriving estimates on subset of observations. A simpler estimation method was later found by Desrosier et al. (2005) which did not required the use of randomization techniques, deriving the error statistics directly from the covariances between *O-F* and *A-F* and *O-A*. Both variant of the method uses an iterative approach to estimate the observation and background error covariances (or variances), where the estimate of the $n^{\text{th}}$ iterate is used to update the error covariances prescribed in the analysis solver which generates a new analysis from which new statistics can be derived. The convergence of this iterated scheme is the object of this study.

The scheme involved a recalculation of the analysis for each iterate, and as such do not change the value of the forecast which is given as a prior. Using this iterative scheme online with an assimilation method provides then a mean to also change the forecast, and the issue of stability shall also be investigated.

In section 2, we derive the convergence properties for the scalar case based on the algorithm of Desrosiers et al. (2005), individually for the observation error variance, background error variances, and for both simultaneously. We then discuss in section 3 the convergence issue in a homogeneous 1D periodic domain with observation at each grid point, and where spectral analysis can be carried out. Finally in section 4 we present some results using a meteorology-chemistry model with a 3D Var-FGAT assimilation system.

## 2. The scalar case

### 2.1. Iteration on observation error variance

Let $\langle OmP, OmP\rangle = \mathbf{HBH}^T + \mathbf{R}$ and the Kalman gain be given by $\mathbf{K} = \overline{\mathbf{B}}\mathbf{H}^T\left(\mathbf{H}\overline{\mathbf{B}}\mathbf{H}^T + \overline{\mathbf{R}}\right)^{-1}$, where the overbar denotes the *prescribed* error covariances, and no obverbar denotes the true error covariances.

According to Desrosiers,

$$\langle OmA, OmP\rangle = \overline{\mathbf{R}}(\mathbf{H}\overline{\mathbf{B}}\mathbf{H}^T + \overline{\mathbf{R}})^{-1}(\mathbf{HBH} + \mathbf{R}) \tag{1}$$

We will conduct an analysis of convergence based on a scalar case, first with a correctly prescribed forecast error variance, and then with an incorrectly prescribed error variance.

### *2.1.1. Correctly prescribed forecast error variance*

Let

$$\overline{\mathbf{B}} = \mathbf{B} = \sigma_f^2 \tag{2}$$

$$\overline{\mathbf{R}} = \alpha\mathbf{R} = \alpha\sigma_o^2 \tag{3}$$

where $\alpha$ is a coefficient that represent the relative error of observation error variance. This is a tunning coefficient, and an optimal estimation is reached when $\alpha = 1$; note that $\sigma_f^2$, $\sigma_o^2$ are respectively the *true* forecast and observation error variance.

In this scalar case we have,

$$\langle OmA, OmP \rangle = \frac{\alpha\,\sigma_o^2}{\alpha\,\sigma_o^2 + \sigma_f^2}(\sigma_o^2 + \sigma_f^2) = \alpha\,\sigma_o^2\left(\frac{\gamma+1}{\alpha\,\gamma+1}\right) \tag{5}$$

where

$$\gamma = \frac{\sigma_o^2}{\sigma_f^2} \tag{6}$$

is the ratio of observation to forecast error variance. The purpose of this statistics is to estimate a new observation error variance. Suppose that in the r.h.s. of (5) the tunning coefficient is evaluated at $n^{\text{th}}$ iterate, i.e. $\alpha = \alpha_n$. In an iterative scheme, the l.h.s. of (5) represent the $n+1^{\text{th}}$ iterate, that is

$$\langle OmA, OmP \rangle = \alpha_{n+1}\,\sigma_o^2 \tag{7}$$

then we get an iterative scheme of the form,

$$\alpha_{n+1} = \alpha_n\left(\frac{\gamma+1}{\alpha_n\gamma+1}\right) = G(\alpha_n) \tag{8}$$

where the *mapping G* is defined as

$$G(\alpha) = \alpha\left(\frac{\gamma+1}{\alpha\,\gamma+1}\right) \tag{9}$$

The convergent point (or fixed point) $\alpha = \alpha^*$ is defined as

$$\alpha^* = G(\alpha^*) \tag{10}$$

substituting in (9) we get,

$$\alpha^* = 1 \tag{11}$$

which is the optimal value. The iterative scheme converges to the fixed point if

$$\left|G'(\alpha^*)\right| < 1 \tag{12}$$

After some math we find that,

$$\boxed{G'(\alpha^*) = \frac{1}{\gamma+1} = \frac{\sigma_f^2}{\sigma_o^2+\sigma_f^2} = K \le 1} \tag{13}$$

that is the scheme is *convergent* . In (13) $K$ is the Kalman gain.

### *2.1.2. Incorrectly prescribed forecast error variance*

In this case we consider that

$$\overline{\mathbf{B}} = \beta\mathbf{B} = \beta\sigma_f^2 \tag{14}$$

$$\overline{\mathbf{R}} = \alpha\mathbf{R} = \alpha\sigma_o^2 \tag{15}$$

where $\beta$ represents a suboptimal parameter of the forecast error variance. Carrying the same analysis as before, we get the following iterative map

$$\alpha_{n+1} = \alpha_n\left(\frac{\gamma+1}{\alpha_n\gamma+\beta}\right) = G(\alpha_n) \tag{16}$$

The fixed point is this case is

$$\alpha^* = 1 + \frac{1-\beta}{\gamma} = 1 + \frac{\left(\sigma_f^2 - \overline{\sigma}_f^2\right)}{\sigma_o^2} \tag{17}$$

and *does not correspond to the optimal value of the observation error variance*. For an underestimated forecast error variance, i.e. $0 \le \beta < 1$, the fixed point exceeds the true observation error. For an overestimated forecast error, i.e. $\beta > 1$, the fixed point is smaller than the true observation error.

The convergence at the fixed point is

$$\boxed{G'(\alpha^*) = \frac{\beta}{\gamma+1} = \frac{\beta\sigma_f^2}{\sigma_o^2+\sigma_f^2}} \tag{18}$$

When the forecast error variance is underestimated, i.e. $0 \le \beta < 1$, then the iterative scheme always converges, i.e. $\left|G'(\alpha^*)\right| < 1$. When the forecast error variance is overestimated, then the scheme will not converge if $\beta\sigma_f^2 = \overline{\sigma}_f^2 > \sigma_o^2 + \sigma_f^2$ , that is if the prescribed forecast error variance exceeds the innovation variance. Unless a mistake is made however, in practice this condition never occurs – whatever estimate of the forecast or observation error variance is obtained in an estimation procedure it is always checked to be smaller than the innovation variance.

Thus with the linear analysis made here, we conclude that in practice the observation error iterative scheme will always converge, but will lead to an overestimation of the observation error variance if the forecast error variance is underestimated, and will lead to an underestimation of the observation error variance if the forecast error variance is overestimated. There is thus a necessity to concurrently iterate on the forecast error variance.

## 2.2. Iteration on the forecast error variance

According to Desrosiers,

$$\langle AmP, OmP \rangle = \mathbf{H}\overline{\mathbf{B}}\mathbf{H}^T(\mathbf{H}\overline{\mathbf{B}}\mathbf{H}^T + \overline{\mathbf{R}})^{-1}(\mathbf{HBH} + \mathbf{R}) \tag{19}$$

### *2.2.1. Correctly prescribed observation error variance*

Suppose we have,

$$\overline{\mathbf{B}} = \beta\mathbf{B} = \beta\,\sigma_f^2 \tag{20}$$

$$\overline{\mathbf{R}} = \mathbf{R} = \sigma_o^2 \tag{21}$$

where $\beta$ is a coefficient that represent the relative error forecast error variance. In this scalar case, (19) becomes

$$\langle AmP, OmP \rangle = \frac{\beta\sigma_f^2}{\sigma_o^2 + \beta\sigma_f^2}(\sigma_o^2 + \sigma_f^2) = \beta\sigma_f^2\left(\frac{\gamma+1}{\gamma+\beta}\right) \tag{22}$$

As in the previous section we define an iterative scheme on the forecast error variance, which then takes the following form,

$$\beta_{n+1} = \beta_n\left(\frac{\gamma+1}{\gamma+\beta_n}\right) = F(\beta_n) \tag{23}$$

The fixed point of this iterative map is

$$\beta^* = 1 \tag{24}$$

which correspond to the optimal value of the forecast error variance. The scheme is also convergent, since

$$\boxed{F'(\beta^*) = \frac{\gamma}{\gamma+1} = \frac{\sigma_o^2}{\sigma_o^2 + \sigma_f^2} = I - K \le 1} \tag{25}$$

### *2.2.2. Incorrectly prescribed observation error variance*

Suppose we have,

$$\overline{\mathbf{B}} = \beta\mathbf{B} = \beta\,\sigma_f^2 \tag{26}$$

$$\overline{\mathbf{R}} = \alpha\mathbf{R} = \alpha\,\sigma_o^2 \tag{27}$$

Carrying the same analysis as before, we get the following iterative map

$$\beta_{n+1} = \beta_n\left(\frac{\gamma+1}{\alpha\gamma+\beta_n}\right) = F(\beta_n) \tag{28}$$

The fixed point is this case is

$$\beta^* = 1 + \gamma(1-\alpha) = 1 + \frac{\left(\sigma_0^2 - \overline{\sigma}_o^2\right)}{\sigma_f^2} \tag{29}$$

and *does not correspond to the optimal value of the forecast error variance*. For an underestimated observation error variance, i.e. $0 \le \alpha < 1$, the fixed point exceeds the true forecast error. For an overestimated forecast error, i.e. $\alpha > 1$, the fixed point is smaller than the true observation error.

The convergence at the fixed point is determined by

$$\boxed{F'(\beta^*) = \frac{\alpha\gamma}{\gamma+1} = \frac{\overline{\sigma}_o^2}{\sigma_o^2 + \sigma_f^2}} \tag{30}$$

When the observation error variance is underestimated, i.e. $0 \le \alpha < 1$, then the iterative scheme always converges, i.e. $\left|F'(\beta^*)\right| < 1$. When the observation error variance is overestimated, then the scheme will not converge if $\alpha\sigma_o^2 = \overline{\sigma}_o^2 > \sigma_o^2 + \sigma_f^2$ , that is if the prescribed observation error variance exceeds the innovation variance. Unless a mistake is made however, in practice this condition never occurs – whatever estimate of the forecast or observation error variance is obtained in an estimation procedure it is always checked to be smaller than the innovation variance.

Thus with the linear analysis made here, we conclude that in practice the forecast error iterative scheme will always converge, but will lead to an overestimation of the forecast error variance if the observation error variance is underestimated, and will lead to an underestimation of the forecast error variance if the observation error variance is overestimated. There is thus a necessity to concurrently iterate with the observation error variance.

## 2.3. Iteration on both observation and background error variances

When an update on both observation and background error variances are performed, the mapping takes the form

$$\begin{aligned} \alpha_{n+1} &= G(\alpha_n, \beta_n) = \frac{\alpha_n(\sigma_o^2 + \sigma_f^2)}{\alpha_n\sigma_o^2 + \beta_n\sigma_f^2} = \frac{\alpha_n(\gamma+1)}{\alpha_n\gamma + \beta_n} \\ \beta_{n+1} &= F(\alpha_n, \beta_n) = \frac{\beta_n(\sigma_o^2 + \sigma_f^2)}{\alpha_n\sigma_o^2 + \beta_n\sigma_f^2} = \frac{\beta_n(\gamma+1)}{\alpha_n\gamma + \beta_n} \end{aligned}. \tag{31}$$

We note immediately that the solution should obey

$$\frac{\alpha_n}{\beta_n} = \frac{\alpha_{n+1}}{\beta_{n+1}}. \tag{32}$$

The fixed point solution of the map (31) is given by

$$\begin{aligned} \alpha^* &= G(\alpha^*, \beta^*) \\ \beta^* &= F(\alpha^*, \beta^*) \end{aligned},$$

which reduces to a single equation

$$(\alpha^* - 1)\sigma_o^2 + (\beta^* - 1)\sigma_f^2 = 0, \tag{33}$$

which is a non-unique solution. The solution (33) actually says that the estimated innovation variance is equal to the true innovation variance. Also as it will be further demonstrated in the following section, once the solution is on the fixed-point set, there is no update on either observation or background estimated error variances – in other words these is no flow possible along the fixed-point attractor set.

The fixed-point set (33) is also an attractor, i.e. the scheme is convergent. Indeed, the Jacobian *J* of (31) at any value of $\alpha, \beta$ is

$$J = \frac{\sigma_o^2 + \sigma_f^2}{\alpha\sigma_o^2 + \beta\sigma_f^2} \begin{vmatrix} \beta\sigma_f^2 & -\alpha\sigma_f^2 \\ -\beta\sigma_o^2 & \alpha\sigma_o^2 \end{vmatrix} = 0 \tag{34}$$

equal to zero. So the scheme is convergent but converging to a non unique solution which will depend on the initial value of $\alpha$ and $\beta$.

The phase diagram $(\alpha, \beta)$ depicted in Figure 1 shows the fixed-point solution with the thick line, and iterations from $(\alpha_n, \beta_n)$ to $(\alpha_{n+1}, \beta_{n+1})$ as arrows.

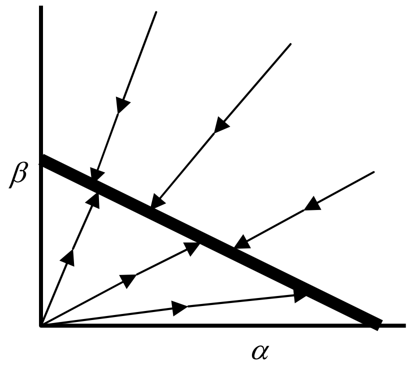


*Figure 1: Phase diagram of equ. (31)*

The scheme also converges to the fixed-point in a single iteration. Indeed consider iteration equation (31) for *n* = 0, multiply the alpha equation by $\sigma_o^2$ and multiply the beta equation by $\sigma_f^2$ and add the results, we get

$$\alpha_1 \sigma_o^2 + \beta_1 \sigma_f^2 = \sigma_o^2 + \sigma_f^2$$

that is the solution has reached the fixed-point set, and no further update on *α* and *β* can be achieved by iterating beyond *n* = 1.

Here are now a few important remarks

1 -

- If initially $\beta_0 \ll \alpha_0$, then the solution will converge to $\beta_1 = \beta^* \ll \alpha_1 = \alpha^*$ and from (33) we conclude that $\alpha^* \approx 1 + \sigma_f^2 / \sigma_o^2$.
- If initially $\alpha_0 \ll \beta_0$, then the solution will converge to $\alpha_1 = \alpha^* \ll \beta_1 = \beta^*$ and from (33) we conclude that $\beta^* \approx 1 + \sigma_o^2 / \sigma_f^2$.

2 – The following property is useful for the interpretation of the spatial correlation case discussed in the next section.

- If $\sigma_f^2 \ll \sigma_o^2$, then according to (33) the solution will converge to $\alpha^* \approx 1$, that is *it converges to the true observation error variance*.
- If initially $\sigma_o^2 \ll \sigma_f^2$, then according to (33) the solution converges to $\beta^* \approx 1$, that is *it converges to the true background error variance*.

3 - Also we note that an ensemble of solutions in which the estimated innovation variance is constant must lie parallel to the thick line in Fig 1, i.e. to the fixed-point solution. The iteration (32) is clearly not innovation variance preserving.

4 -It is also very important to note that along the fixed-point solution (33), there is no flow. If we substitute $\alpha(\beta) = 1 - (\beta - 1)\sigma_f^2 / \sigma_o^2$ into (31) we get a stationary point for *β*. Starting with an initial guess that has the total variance equal to the innovation variance will not evolve.

# 3. Spatial correlations

It has been argued that the failure of the scheme when both observation and background error parameters are estimated occurs because the correlation length scale of the observation error and background error are identical, which is in essence what is happening with the scalar case. In this section we will examine the scheme in a more general context allowing spatial correlations to be different. First, we will examine the problem from a completely general point of view to showing in the end that there is no new information in the *A-F* or *O-F* that there is already in the *O-F*. Second, we will consider a 1D periodic domain with observations at each grid points to obtain a spectral analysis of the problem and examine the uniqueness of the solution in that context.

## 3.1. General properties

As noted by Talagrand (2009), there is a one-to-one correspondence between the innovation vector and the data vector - a two-component vector composed of the *AmF* and the *OmA*. Since the observation error estimation equation (1) carries statistics of both *OmF* and *OmF*, it follows that equation (1) carries the same information content than the background error estimation equation (19). That is the reason why the simultaneous iteration on background and observation error fail to yield independent estimates of error variances, as was show in the scalar case, and more will be said in the multivariate case.

To derive the equivalence of the observation error estimation and background error estimation schemes, let's denote the statistics,

$$\begin{aligned} \left\langle (O-F)(O-F)^T \right\rangle &= \mathbf{O} \\ \left\langle (O-A)(O-F)^T \right\rangle &= \mathbf{A} \\ \left\langle (A-F)(O-F)^T \right\rangle &= \mathbf{P} \end{aligned} \qquad (35)$$

and re-write the observation and background error estimation scheme as,

$$\begin{aligned} \mathbf{A}\mathbf{O}^{-1}(\mathbf{H}\overline{\mathbf{B}}\mathbf{H}^T + \overline{\mathbf{R}}) &= \overline{\mathbf{R}} \\ \mathbf{P}\mathbf{O}^{-1}(\mathbf{H}\overline{\mathbf{B}}\mathbf{H}^T + \overline{\mathbf{R}}) &= \mathbf{H}\overline{\mathbf{B}}\mathbf{H}^T \end{aligned} \qquad (36)$$

It is easy to show that,

$$\mathbf{A} + \mathbf{P} = \mathbf{O} \tag{37}$$

which follows form the fact that *AmF* plus *OmA* has the same information content than *OmF*, and consequently,

$$\mathbf{AO}^{-1} = (\mathbf{O} - \mathbf{P})\mathbf{O}^{-1} = \mathbf{I} - \mathbf{PO}^{-1} \tag{38}$$

From the observation error estimation scheme in Eq. (36), we then get

$$\begin{aligned} \mathbf{AO}^{-1}(\mathbf{H}\overline{\mathbf{B}}\mathbf{H}^T + \overline{\mathbf{R}}) &= \overline{\mathbf{R}} \\ (\mathbf{I} - \mathbf{PO}^{-1})(\mathbf{H}\overline{\mathbf{B}}\mathbf{H}^T + \overline{\mathbf{R}}) &= \overline{\mathbf{R}} \\ \mathbf{H}\overline{\mathbf{B}}\mathbf{H}^T + \overline{\mathbf{R}} - \mathbf{PO}^{-1}\ (\mathbf{H}\overline{\mathbf{B}}\mathbf{H}^T + \overline{\mathbf{R}}) &= \overline{\mathbf{R}} \\ \mathbf{H}\overline{\mathbf{B}}\mathbf{H}^T &= \mathbf{PO}^{-1}\ (\mathbf{H}\overline{\mathbf{B}}\mathbf{H}^T + \overline{\mathbf{R}}) \end{aligned} \tag{39}$$

is the same as the background error estimation equation in Eq. (37).

The fixed point solution of the Desrosiers' scheme has also two important properties:

1 – At the fixed point solution for observation error $\overline{\mathbf{R}}^*$ and background error $\overline{\mathbf{B}}^*$,

$$\mathbf{H}\overline{\mathbf{B}}^*\mathbf{H}^T + \overline{\mathbf{R}}^* = \mathbf{HBH}^T + \mathbf{R} \tag{40}$$

that is sum of the observation and background error is equal to the innovation covariance.

2 – When at an iterate *k* the sum of observation and background error is equal to the innovation covariance,

$$\mathbf{H}\overline{\mathbf{B}}_k\mathbf{H}^T + \overline{\mathbf{R}}_k = \mathbf{HBH}^T + \mathbf{R} \tag{41}$$

no more updates on  can occur, that is the estimates has reached their fixed-points..

The property (2) also says that if the sum of the observation and background error covariances has reached the attractor set (41), and although the partitioning between observation and backbround error may be incorrect, there is no update possible within the attractor set (i.e. along the thick line of Fig 1).

### 3.2. 1D periodic domain

It has been argued in Chapnik et al (2004) that the failure of the scheme when both observation and background error parameters are estimated is expected when the correlation length scale of the observation error and background error are identical, which is in essence what is happening in the scalar case. To examine this issue we now consider a 1D periodic domain and allow the observation error to be spatially uncorrelated and the background error be spatially correlated so that there is a clear evidence of length scale separation. Furthermore we assume that the background error covariance is homogeneous and that we observe at each grid point of the model, i.e. $\mathbf{H} = \mathbf{I}$. In this case it is possible to diagonalize simultaneously the observation and background error covariance iteration equations,

$$\begin{aligned} \overline{\mathbf{R}}_{n+1} &= \overline{\mathbf{R}}_n(\overline{\mathbf{B}}_n + \overline{\mathbf{R}}_n)^{-1}(\mathbf{B} + \mathbf{R}) \\ \overline{\mathbf{B}}_{n+1} &= \overline{\mathbf{B}}_n(\overline{\mathbf{B}}_n + \overline{\mathbf{R}}_n)^{-1}(\mathbf{B} + \mathbf{R}) \end{aligned} \tag{42}$$

to become a set a scalar equations of the form (31), one set for each wavenumber *k*,

$$\begin{aligned}\bar{r}_{n+1}(k) &= \frac{\bar{r}_n(k)}{\bar{b}_n(k)+\bar{r}_n(k)}(b(k)+r(k)) \\ \bar{b}_{n+1}(k) &= \frac{\bar{b}_n(k)}{\bar{b}_n(k)+\bar{r}_n(k)}(b(k)+r(k))\end{aligned} \quad . \tag{43}$$

***a) Experiment with no constraint on error correlations***

In the experiment below we have considered a Gaussian background error correlation with L= 300 km on a periodic domain of length 40,000 km, and a spatially uncorrelated observation error, as in Desrosiers et al. (2009). Figure 2 display the observations error (left panel) and background error (middle panel) spectral variances. Both initial and true error covariances (in physical space) are based on the same correlation model but differ only in their (physical space) variances. The right panel represents the sum of observation and background error for each wavenumber *k*.

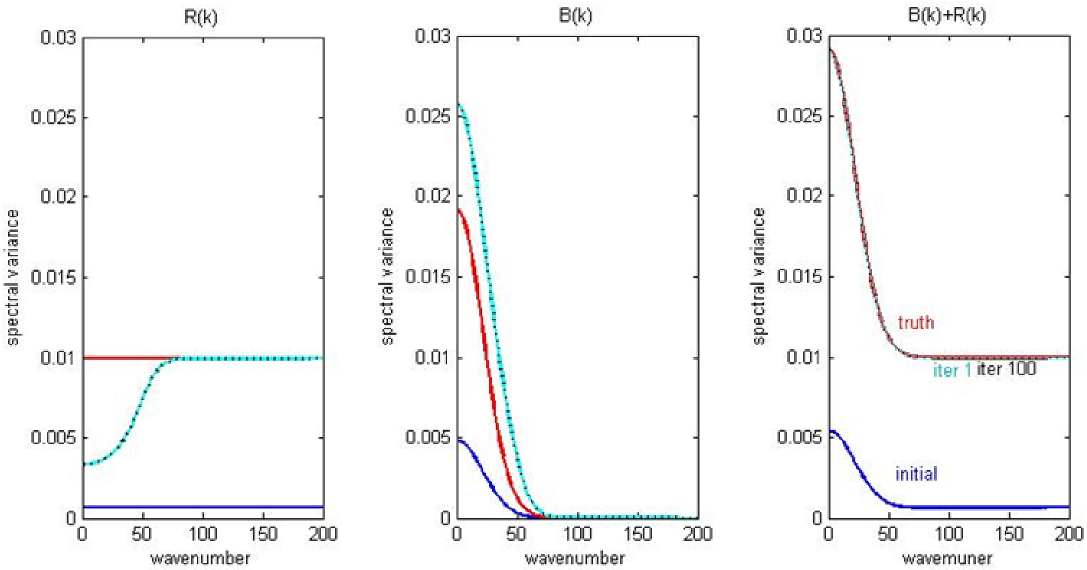


*Figure 2: Spectral convergence of observation error variance (left panel), background error variance (middle panel) and their sum (right panel). The red curve is the true spectrum, and the blue curve is the initial spectrum. The cyan curve is after a single iteration, and the dotted line after 100 iterations.*

Note from the red curves on the left and middle panels that for $k \gg 35$ the background error spectral variance is much smaller than the observation error spectral variance and thus the estimated observation error spectral variance converges to the truth. For wavenumber smaller than 35, the observation error clearly do not converges to the truth and thus the resulting observation error covariance in physical space is thus spatially correlated. We also note that the background error spectral variance is overestimated and in such a way that the sum of observation error and background error spectral variances (right panel) is equal to the truth. The convergence is also obtained in a single iteration.

***b) Experiment using the true error correlation as a constraint***

In the absence of additional information and/or constraints the Desrosier's scheme in a 1D-domain is not fundamentally different from the scalar case. Let us then examine a case where we constraint the observation error correlation to be spatially uncorrelated as in the truth and constraint the background error correlation to be the same as in the truth. The only differences are with respect to the initial error variances that are different from the truth error variances. Specifically, let's assume that;

$$r(k) = \frac{\sigma_o^2}{N} \quad ; \quad \bar{r}_n(k) = \frac{\alpha_n \sigma_o^2}{N} \quad ,$$
$$b(k) = \sigma_f^2 b^*(k) \; ; \; \bar{b}_n(k) = \beta_n \sigma_f^2 b^*(k) \tag{44}$$

where $N$ is the number of spectral components, and $b^*(k)$ is the spectrum of the background error correlation. The variance in physical space is computed from the sum of the spectral components, so that

$$\sum_{k=1}^{N} r(k) = \sigma_o^2 \quad ; \quad \sum_{k=1}^{N} b(k) = \sigma_f^2 \quad ; \quad \sum_{k=1}^{N} b^*(k) = 1$$
$$\sum_{k=1}^{N} \bar{r}_n(k) = \alpha_n \sigma_o^2 \; ; \quad \sum_{k=1}^{N} \bar{b}_n(k) = \beta_n \sigma_f^2 \tag{45}$$

Let's consider the first iteration starting with $\alpha_0 \neq 1$ and $\beta_0 \neq 1$. The first iteration on observation error can be re-written as

$$\alpha_1 = \alpha_0 \; \frac{\sigma_f^2 b^*(k) + \sigma_o^2 / N}{\beta_0 \sigma_f^2 b^*(k) + \alpha_0 \sigma_o^2 / N} \text{ or}$$

$$\alpha_1 \left( \beta_0 \sigma_f^2 b^*(k) + \frac{\alpha_0 \sigma_o^2}{N} \right) = \alpha_0 \left( \sigma_f^2 b^*(k) + \frac{\sigma_o^2}{N} \right) . \tag{46}$$

This equation can be summed up from $k$ = 1 to $N$,

$$\sum_{k=1}^{N} \alpha_1 \left( \beta_0 \sigma_f^2 b^*(k) + \frac{\alpha_0 \sigma_o^2}{N} \right) = \sum_{k=1}^{N} \alpha_0 \left( \sigma_f^2 b^*(k) + \frac{\sigma_o^2}{N} \right) ,$$

giving

$$\alpha_1 = \frac{\alpha_0 (\sigma_f^2 + \sigma_o^2)}{\alpha_0 \sigma_o^2 + \beta_0 \sigma_f^2} , \tag{47}$$

an identical equation to the scalar case. The first iteration on background error is slightly more complicated form

$$\beta_1 b^*(k) \left( \beta_0 \sigma_f^2 b^*(k) + \frac{\alpha_0 \sigma_0^2}{N} \right) = \beta_0 b^*(k) \left( \sigma_f^2 b^*(k) + \frac{\sigma_o^2}{N} \right). \tag{48}$$

Again summing up from $k$ = 1 to $N$, and letting

$$\sum_{k=1}^{N} [b^*(k)]^2 = \frac{1}{\delta} \tag{49}$$

yields the update equation for beta

$$\beta_1 = \frac{\beta_0 \left( \frac{\sigma_o^2}{N} + \frac{\sigma_f^2}{\delta} \right)}{\frac{\alpha_0 \sigma_o^2}{N} + \frac{\beta_0 \sigma_f^2}{\delta}} \tag{50}$$

The iteration is the same for any iterate, so the mapping takes the form

$$\alpha_{n+1} = \frac{\alpha_n(\sigma_o^2 + \sigma_f^2)}{\alpha_n \sigma_o^2 + \beta_n \sigma_f^2}$$

$$\beta_{n+1} = \frac{\beta_n\left(\dfrac{\sigma_o^2}{N} + \dfrac{\sigma_f^2}{\delta}\right)}{\dfrac{\alpha_n \sigma_o^2}{N} + \dfrac{\beta_n \sigma_f^2}{\delta}} \tag{51}$$

The fixed-point solution is this time *a unique point* equal to

$$\boxed{\alpha^* = \beta^* = 1} \tag{52}$$

that *is the true error statistics*. This results shows that if the error correlations are constrained with the true error correlations the Desrosiers' scheme to update the observation and background error variances converges to the true values.

It is also easy to show that *if the prescribed error correlation is not the true error correlation* then although the convergent solution is unique it *does not converge to the true solution*.

**Acknowledgements**: I would like to thanks G. Desrosiers for the numerous discussions we had about the convergence issue which has stimulated this analysis.

## References

Chapnik, B., G. Desrosiers, F. Rabier, and O. Talagrand, 2004: Properties and first application of an error-statistics tuning method in variational assimilation. *Q. J. R. Meteorol. Soc*., **130**, 2253-2275.

Chapnick, B., G. Desrosiers, F. Rabier, and O. Talagrand, 2006: Diagnosis and tuning of observational error statistics in a quasi operational data assimilation setting. *Q. J. R. Meteor. Soc*., *132*, 5430565.

Desrosiers, 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.

Desrosiers, G., L. Berre, B. Chapnik, and P. Poli, 2005: Diagnosis of observation, background, and analysis error statistics in observation space. Submitted to *Q. J. R. Meteorol. Soc*..

Desrosiers, G., L. Berre, and B. Chapnick, 2009: Objective validation of data assimilation systems: diagnosing sub-optimality. Proceedings of the *ECMWF Workshop on diagnostics of data assimilation system performance*. Reading June 15-17 2009.

Talagrand, O. 2009: A posteriori verification of analysis and assimilation algorithms. In *Data Assimilation: Making Sense of Observations* Eds: Lahoz W., B. Kahattatov and R. Ménard. Springer (to appear in 2009)