Environment Canada
Air Quality Research Division

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

Richard Ménard and Yan Yang
Air Quality Research Division – Environment Canada
Richard.Menard@ec.gc.ca

Abstract

The convergence of the Derosier's scheme to estimate observation and background error variances based on *OmF*, *OmA* and *AmF* is studied from a theoretical point of view. The general properties of the fixed point of the scheme are discussed and illustrated with a scalar, 1D domain, and in an operational assimilation system. Several iterated schemes are considered: the estimation of either observation or background error variances and the estimation of both variances either simultaneously or in sequence.

It is shown that for the simultaneous estimation of observation and background error variance the theoretical convergence is obtained in a single iteration, but the convergent value are incorrect although the sum of variances matches the innovation variance. Additional information (e.g. correlation model, lagged-innovation) is needed to resolve the estimation problem.

Iterative scheme

Starting with the innovation covariance

$$\left\langle (O-F)(O-F)^T \right\rangle = \mathbf{HBH}^T + \mathbf{R}$$

and a first estimate on observation and background error covariances

$$\overline{\mathbf{R}}_{k+1} = \overline{\mathbf{R}}_k (\mathbf{H}\overline{\mathbf{B}}_k\mathbf{H}^T + \overline{\mathbf{R}}_k)^{-1}(\mathbf{HBH}+\mathbf{R}) = \left\langle (O-A)(O-F)^T \right\rangle_{k+1}$$

$$\mathbf{H}\overline{\mathbf{B}}_{k+1}\mathbf{H}^T = \mathbf{H}\overline{\mathbf{B}}_k\mathbf{H}^T (\mathbf{H}\overline{\mathbf{B}}_k\mathbf{H}^T + \overline{\mathbf{R}}_k)^{-1}(\mathbf{HBH}+\mathbf{R}) = \left\langle (A-F)(O-F)^T \right\rangle_{k+1}$$

The overbar denotes the estimates, and k the iteration index

An iterated map $\mathbf{G}$

$$\mathbf{x}_{n+1} = \mathbf{G}(\mathbf{x}_n)$$

has a fixed point $\mathbf{x}^*$ if

$$\mathbf{x}^* = \mathbf{G}(\mathbf{x}^*).$$

The scheme is convergent if

$$\left| \frac{\partial \mathbf{G}(\mathbf{x}^*)}{\partial \mathbf{x}} \right| < 1$$

General properties of the Desrosier's scheme

(1) If $\overline{\mathbf{R}}^*$ and $\overline{\mathbf{B}}^*$ are fixed points, then

$$\mathbf{H}\overline{\mathbf{B}}^*\mathbf{H}^T + \overline{\mathbf{R}}^* = (\mathbf{HBH}+\mathbf{R})$$

(2) When the iterate k is such that

$$\mathbf{H}\overline{\mathbf{B}}_k\mathbf{H}^T + \overline{\mathbf{R}}_k = (\mathbf{HBH}+\mathbf{R})$$

no more updates on the individual components $\overline{\mathbf{R}}_k$ and $\overline{\mathbf{B}}_k$ can occur.

Convergence – scalar case

A - Iteration on observation error

$$\left\langle (O-A)(O-F)^T \right\rangle = \overline{\mathbf{R}}(\mathbf{H}\overline{\mathbf{B}}\mathbf{H}^T + \overline{\mathbf{R}})^{-1}(\mathbf{HBH}+\mathbf{R})$$

where $\left\langle (O-F)(O-F)^T \right\rangle = \mathbf{HBH}^T + \mathbf{R}$ is obtained from assimilation residuals and overbar denotes *prescribed* error covariances

i)- Correctly prescribed forecast error variance

$\overline{\mathbf{B}} = \mathbf{B} = \sigma_f^2$ $\quad \overline{\mathbf{R}} = \alpha\mathbf{R} = \alpha\sigma_o^2$ $\quad$ optimal value $\alpha = 1$

$$\left\langle (O-A)(O-F) \right\rangle = \frac{\alpha\sigma_o^2}{\alpha\sigma_o^2+\sigma_f^2}(\sigma_o^2+\sigma_f^2) = \alpha\sigma_o^2\left(\frac{\gamma+1}{\alpha\gamma+1}\right)$$

where $\gamma = \dfrac{\sigma_o^2}{\sigma_f^2}$

let $\left\langle (O-A)(O-F) \right\rangle = \alpha_{n+1}\sigma_o^2$ be the next iterate

so the iteration on α_n takes the form

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

Define a *mapping G*

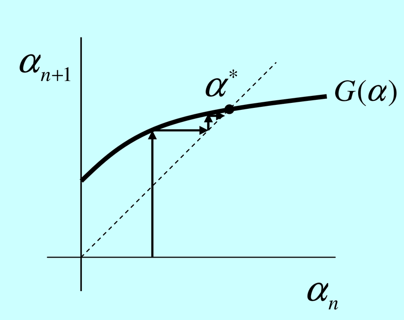
$$G(\alpha) = \alpha\left(\frac{\gamma+1}{\alpha\gamma+1}\right)$$

The fixed-point is

$$\alpha^* = G(\alpha^*)$$

condition for convergence

$$\left| G'(\alpha^*) \right| < 1$$



and so for this case we get $\alpha^* = 1$

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

the scheme is always convergent and converges to the true value, $\alpha = 1$

ii)- Incorrectly prescribed forecast error variance

$\overline{\mathbf{B}} = \beta\mathbf{B} = \beta\sigma_f^2$ $\quad \overline{\mathbf{R}} = \alpha\mathbf{R} = \alpha\sigma_o^2$

the mapping is now different

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

The fixed-point is

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

that is *not the true observation error* value.

- If forecast error variance is underestimated, obs error is overestimated
- If forecast error variance is overestimated, obs error is underestimated

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

Will not converge if: $\beta\sigma_f^2 = \overline{\sigma}_f^2 > \sigma_o^2 + \sigma_f^2$
In practice the estimated forecast error variance will never be larger than the innovation error variance, so for all practical cases the scheme converges.

B - Iteration on both observation and background error

Consider the case of tuning together α and β in each iteration

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

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

then the ratio

$$\mu_{n+1} = \frac{\alpha_{n+1}}{\beta_{n+1}} = \frac{\alpha_n}{\beta_n} = \mu_n = \cdots = \mu_0$$

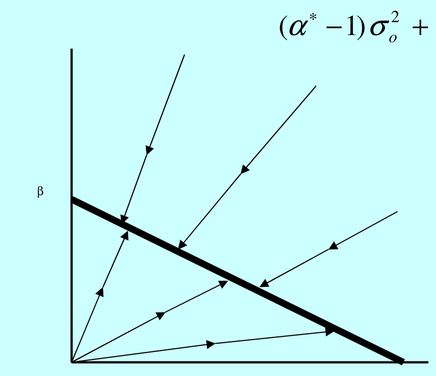
is constant.
The mapping $(\alpha_n, \beta_n) \leftrightarrow (\alpha_{n+1}, \beta_{n+1})$ is in fact ill-defined, since the Jacobian

$$\frac{\partial(G,F)}{\partial(\alpha_n,\beta_n)} = \frac{\gamma+1}{(\alpha_n\gamma+\beta_n)}\begin{pmatrix} \beta_n & -\alpha_n \\ -\beta_n\gamma & \alpha_n\gamma \end{pmatrix}$$

is rank deficient ! and its determinant is 0
So the scheme is strongly convergent

The fixed-point solution (thick black line)

$$(\alpha^*-1)\sigma_o^2 + (\beta^*-1)\sigma_f^2 = 0$$



- If $\sigma_f^2 \ll \sigma_o^2$ then the adjustment is mainly on the observation error variance
- If $\sigma_o^2 \ll \sigma_f^2$ then the adjustment is mainly on the background error variance
- The convergence occur in a single iteration, and the fixed-point depending on the departure point.

$$\overline{\mathbf{R}}_1 = \overline{\mathbf{R}}_0(\mathbf{H}\overline{\mathbf{B}}_0\mathbf{H}^T + \overline{\mathbf{R}}_0)^{-1}(\mathbf{HBH}+\mathbf{R})$$

$$\mathbf{H}\overline{\mathbf{B}}_1\mathbf{H}^T = \mathbf{H}\overline{\mathbf{B}}_0\mathbf{H}^T(\mathbf{H}\overline{\mathbf{B}}_0\mathbf{H}^T + \overline{\mathbf{R}}_0)^{-1}(\mathbf{HBH}+\mathbf{R})$$

$$\mathbf{H}\overline{\mathbf{B}}_1\mathbf{H}^T + \overline{\mathbf{R}}_1 = (\mathbf{HBH}+\mathbf{R})$$

Convergence – 1D domain - Simultaneous

Case where the background error covariance is *spatially correlated* and the observation error covariance is *spatially uncorrelated*

Assume an homogeneous **B** in a 1D periodic domain with observations at each grid points, **H = I** .
We can write the Fourier transform as a matrix **F**, and its inverse as $\mathbf{F}^T$

Then in the system

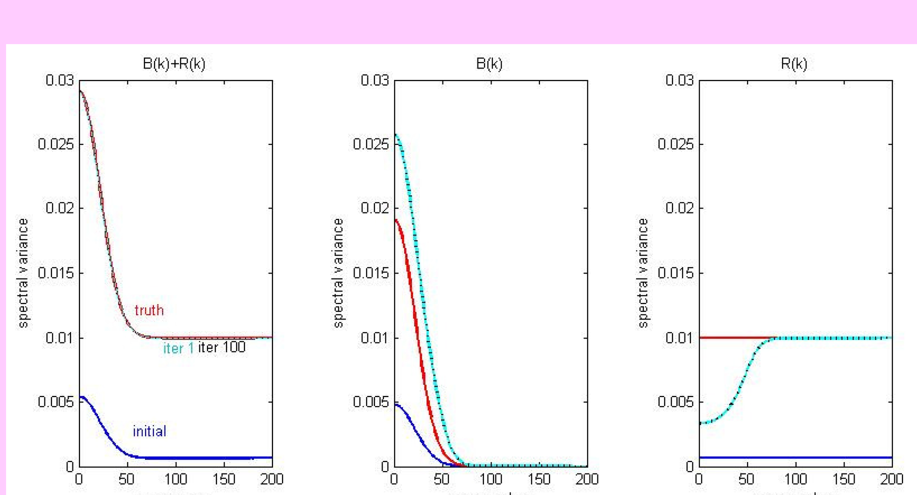
$$\mathbf{R}_{n+1} = \mathbf{R}_n(\mathbf{B}_n+\mathbf{R}_n)^{-1}\,\mathbf{O}$$

$$\mathbf{B}_{n+1} = \mathbf{B}_n(\mathbf{B}_n+\mathbf{R}_n)^{-1}\,\mathbf{O}$$

All matrices can be simultaneously diagonalized giving a N systems of scalar (variance) equations (one for each wavenumber *k*)

$$\hat{\mathbf{R}}_{n+1} = \hat{\mathbf{R}}_n(\hat{\mathbf{B}}_n+\hat{\mathbf{R}}_n)^{-1}\,\hat{\mathbf{O}} \qquad \hat{\mathbf{B}}_{n+1} = \hat{\mathbf{B}}_n(\hat{\mathbf{B}}_n+\hat{\mathbf{R}}_n)^{-1}\,\hat{\mathbf{O}}$$

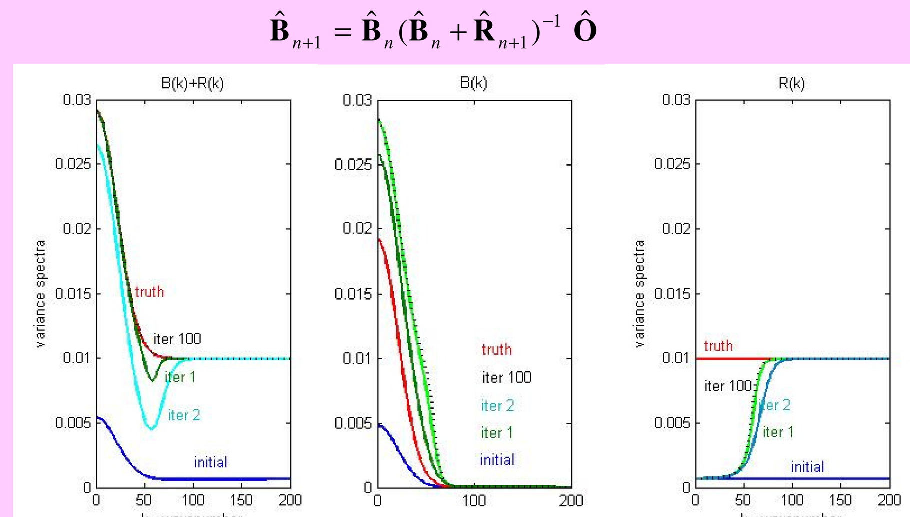
$$\Rightarrow \quad \hat{\alpha}_{n+1} = \hat{\alpha}_n\left(\frac{\gamma+1}{\hat{\alpha}_n\gamma+\hat{\beta}_n}\right) = G(\hat{\alpha}_n, \hat{\beta}_n) \qquad \hat{\beta}_{n+1} = \hat{\beta}_n\left(\frac{\gamma+1}{\hat{\alpha}_n\gamma+\hat{\beta}_n}\right) = F(\hat{\alpha}_n, \hat{\beta}_n)$$



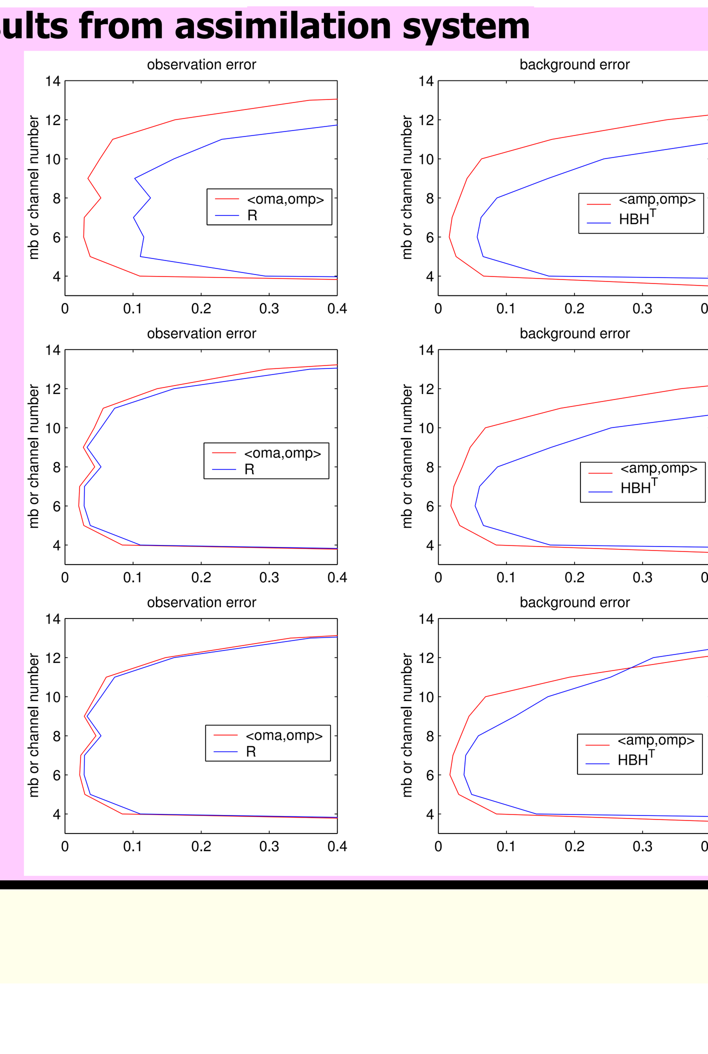
Sequential

$$\hat{\mathbf{R}}_{n+1} = \hat{\mathbf{R}}_n(\hat{\mathbf{B}}_n+\hat{\mathbf{R}}_n)^{-1}\,\hat{\mathbf{O}}$$

$$\hat{\mathbf{B}}_{n+1} = \hat{\mathbf{B}}_n(\hat{\mathbf{B}}_n+\hat{\mathbf{R}}_{n+1})^{-1}\,\hat{\mathbf{O}}$$



Results from assimilation system



Summary and Conclusions

- The convergence of the Desrosiers' et al (2005) scheme has been investigated from a theoretical context and from an assimilation cycle
- Iteration on either observation error variance or background error variance generally converges, but will converge to an overestimate if the counterpart in underestimated, and vice versa
- Iteration on both observation and background error variance converges in a single step, to a non-unique fixed-point solution. On the fixed point solution the innovation variance of the solution is the same as the innovation variance
- An analysis in a 1D-domain reveals the same behavior. While the spectral variance of the estimated innovation matches that of the spectral innovation variance, the individual component, i.e. the observation error and background error do not converge to the truth. In particular, the observation error becomes spatially correlated and the background error variance spectrum becomes more red.