Olivier Goux[1]
Anthony Weaver[1]
Oliver Guillet[2]
Selime Gürol[1]
[1]: CERFACS/CECI, Toulouse, France
[2]: Météo-France, Toulouse, France
goux@cerfacs.fr

METEO FRANCE

CERFACS
EUROPEAN CENTRE FOR RESEARCH AND ADVANCED TRAINING IN SCIENTIFIC COMPUTING

Accounting for correlated observation error in variational ocean data assimilation

In variational ocean data assimilation, diffusion operators are commonly used to model background error correlations. They are also suitable candidates to model observation error correlations as:

- their inverse can be accessed easily to apply $\boldsymbol{R}^{-1}$;
- they can be implemented with a finite-elements method to be applied to unstructured data [1];
- they allow spatially-varying and anisotropic correlations;
- their parameters provide some control on the convergence rate of the **B**-preconditioned conjugate gradient [2].

With spatially constant error covariance, grid and observation density, an error covariance matrix is circulant and diagonalized by a Fourier transform **F**. Its eigenvalue spectrum can be interpreted as the distribution of the error variance between all spatial scales.

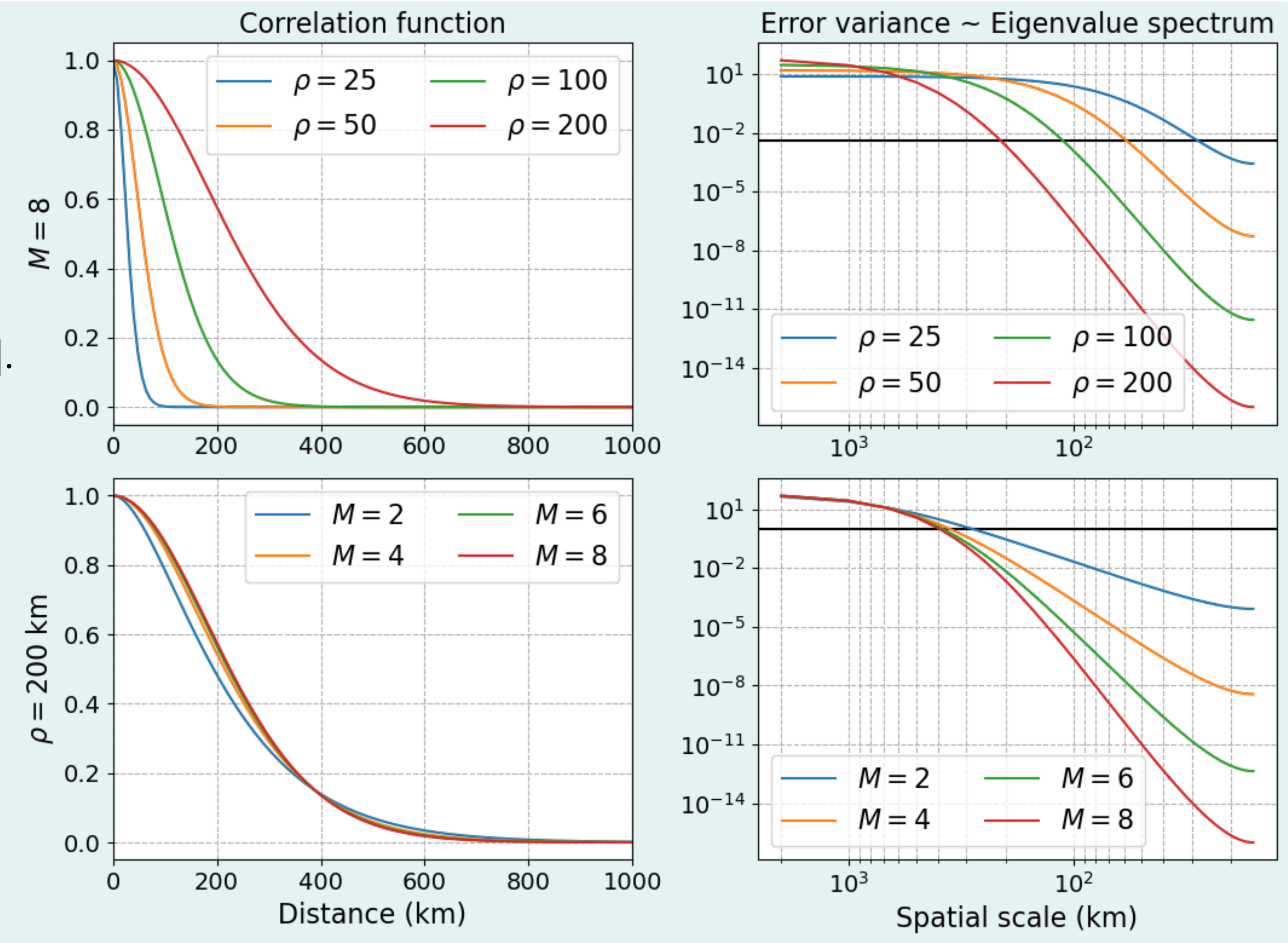
Observation error

$$\boldsymbol{R} = \mathbb{E}[\boldsymbol{\varepsilon_o}\boldsymbol{\varepsilon_o}^T]$$

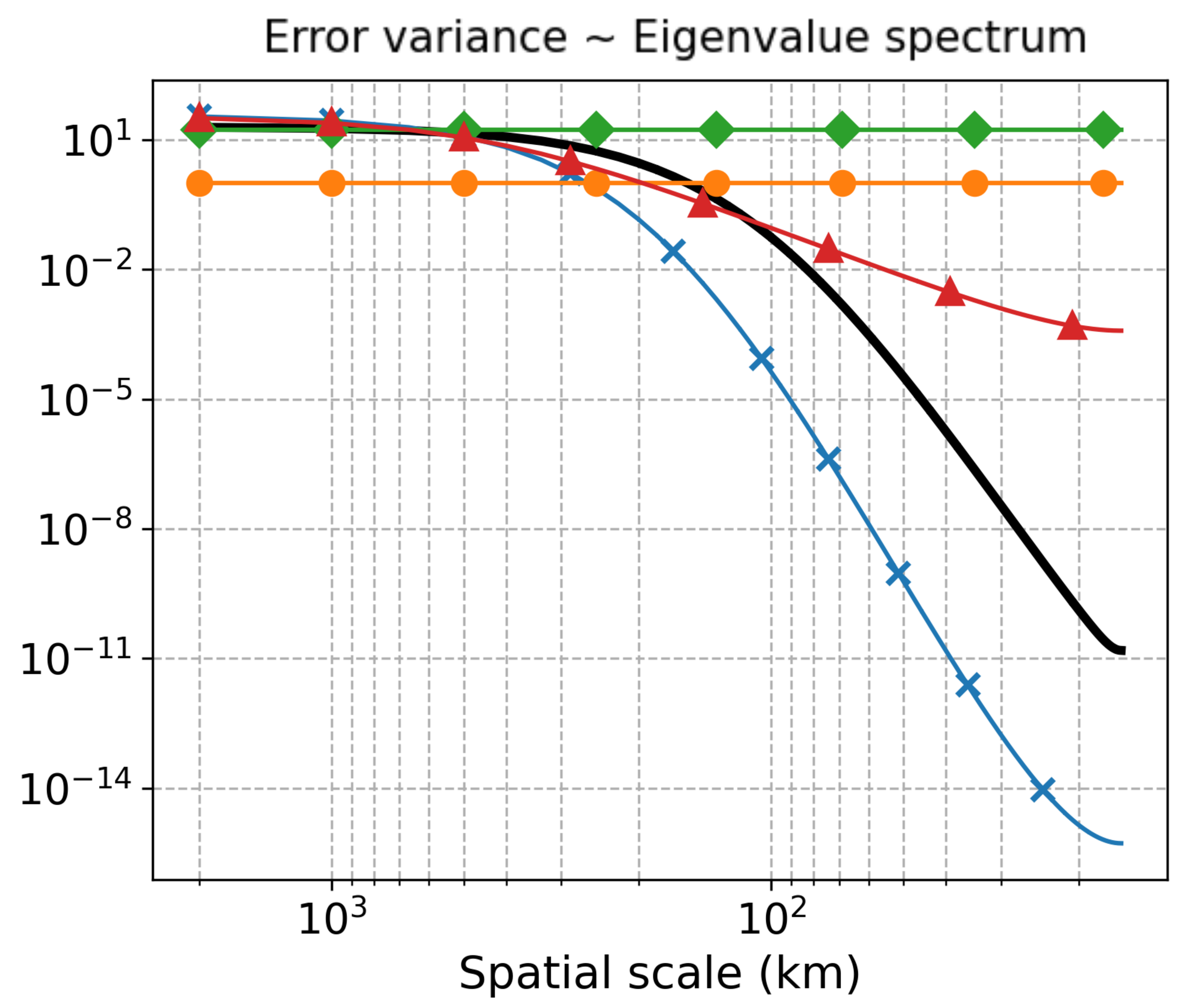
$$\boldsymbol{FRF}^T = \mathbb{E}[(\boldsymbol{F\varepsilon_o})(\boldsymbol{F\varepsilon_o})^T]$$

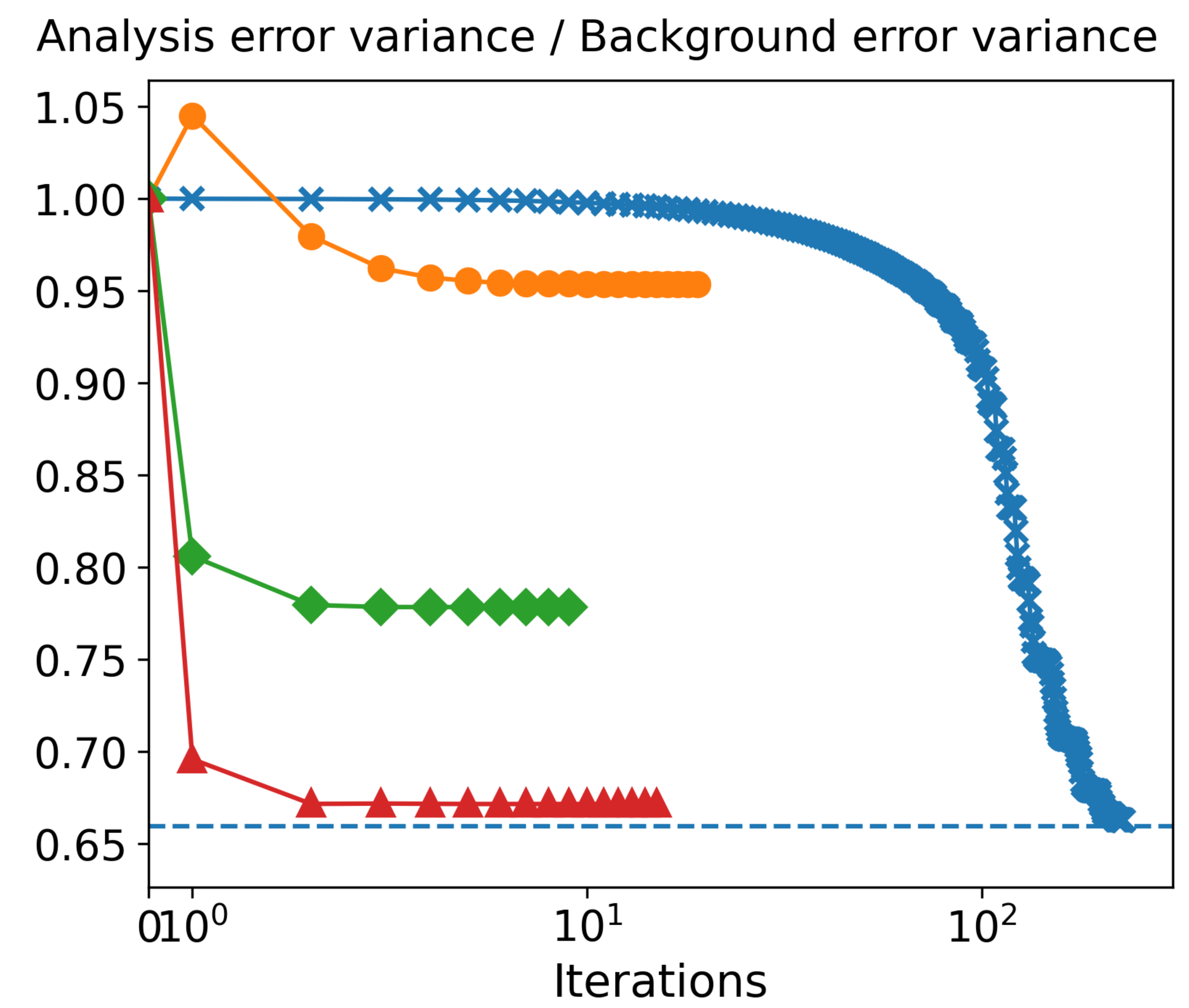
Diagonal — Fourier transform of the observation error

The main parameters of diffusion operators are a length scale ρ and a smoothness parameter M.
The cut-off of the error variance spectrum depends on ρ while the slope after the cut-off depends on M.



By comparing the background and observation error variance spectrums associated with $\boldsymbol{B}$ and $\boldsymbol{R}$, we can predict both the convergence rate of the minimization and the accuracy at full convergence, allowing us to find parameters for $\boldsymbol{B}$ and $\boldsymbol{R}$ leading both to a fast convergence and towards an accurate analysis. This is illustrated in the experiment below, where background states and observations are simulated with known error statistics defined by $\boldsymbol{B}$ and $\boldsymbol{R}$ and with respect to a known true state. The assimilation is then realized with different specifications of $\boldsymbol{R}$, and the total analysis error variance is computed at each iteration of the **B**-preconditioned conjugate gradient.





$\lambda_i(\boldsymbol{HBH^T})$: Background error variance spectrum projected onto the observation space — $\rho_b = 60\ km;\ M_b = 8;\ \sigma_b = 1$

To evaluate the impact of a specification of $\boldsymbol{R}$, we need to compare the associated observation error variance spectrum to the background error variance spectrum. The background error is correlated, and thus has a larger variance at large spatial scales than at small spatial scales.

$\lambda_i(\boldsymbol{R})$: True observation error variance spectrum — $\rho_o = 120\ km\ ;\ M_o = 10;\ \sigma_o = 1$

If a fully accurate $\boldsymbol{R}$ is specified, the theoretical minimum error (- - - -) is reached, but only after a prohibitive number of iterations. This poor convergence is associated with a large condition number. The condition number of the system (κ) increases if at any spatial scale, the observation error variance is smaller than the background error variance (projected onto the observation space) [2]:

$$\kappa = 1 + \max_i \frac{\lambda_i(\boldsymbol{HBH^T})}{\lambda_i(\boldsymbol{R})}$$

This occurs in this case at the smallest spatial scales here since $M_o > M_b$.

$\lambda_i(R_{diag})$: Observation error variance spectrum with neglected correlations — $\rho_o = 0\ km\ ;\ M_o = 0;\ \sigma_o = 1$

If observation error correlations are neglected, the convergence is accelerated as the problematic behaviour at small spatial scales is avoided. However, the observation error variance is underestimated at large spatial scales and overestimated at small spatial scales, leading to a sub-optimal solution at full convergence

$\lambda_i(R_{inflated})$: Observation error variance spectrum with variance inflation — $\rho_o = 0\ km\ ;\ M_o = 0;\ \sigma_o = 17$

Variance inflation reduces the overfit of the observations at large spatial scales at the expense of small spatial scales where the underfit is exacerbated. Prioritizing large spatial scales leads to an improvement of the analysis since the background and observation error variances are larger there than at small spatial scales.

$\lambda_i(R_{recond})$: Reconditioned observation error variance spectrum — $\rho_o = 120\ km\ ;\ M_o = 2;\ \sigma_o = 1$

Even if the 'true' observation error parameters lead to a slow convergence, a non-diagonal $\boldsymbol{R}$ can still be used with 'reconditioned' parameters. In particular, enforcing $M_o < M_b$ is advantageous even when it does not reflect the actual observation error statistics. It increases the specified observation error variance at the smallest spatial scales only, which accelerates the convergence without degrading significantly the analysis at full convergence, as large spatial scales are not affected.

Diffusion operators used for $\boldsymbol{B}$ typically use a large value of M_b, while observation error diagnostics for certain observation types suggest small values of M_o. In these cases where $M_o < M_b$, accounting for observation error correlations is likely to improve the convergence rate of the minimization. However, if future error diagnostics were to suggest larger values of M_o, the relation $M_o < M_b$ should still be enforced to preserve the minimization rate without inducing a significant degradation of the analysis at full convergence.

[1]: **Guillet, O, Weaver, AT, Vasseur, X, Michel, Y, Gratton, S, Gürol, S.** Modelling spatially correlated observation errors in variational data assimilation using a diffusion operator on an unstructured mesh. *Q J R Meteorol Soc* 2019
[2]: **Goux, O, Gürol, S, Weaver, AT, Diouane, Y, Guillet, O.** Impact of correlated observation errors on the conditioning of variational data assimilation problems. *Numer Linear Algebra Appl.* 2023

Work supported by the Copernicus Climate Change Service
ECMWF
Climate Change Service
climate.copernicus.eu
www.cerfacs.fr