Representing Ensemble Covariances in Variational Assimilation with a Diffusion Operator

Anthony T. Weaver and Andrea Piacentini
CERFACS/SUC URA 1875, Toulouse, France

Context

- We consider the problem of estimating the $N \times N$ background-error covariance matrix ($\mathbf{B}$) from an ensemble of $N_e \ll N$ error realizations.
- We consider localized, ensemble-based $\mathbf{B}$ models suitable for application with conjugate gradient (CG) algorithms in variational data assimilation.

Definitions

- Let $\{\boldsymbol{\epsilon}_p = \mathbf{x}_p - \mathbf{x}_0\}$ denote an ensemble of $p = 1, \ldots, N_e$ state error realizations with respect to an unperturbed control member $\mathbf{x}_0$.
- Let $\boldsymbol{\epsilon}'_p = \boldsymbol{\epsilon}_p - \overline{\boldsymbol{\epsilon}}$ where $\overline{\boldsymbol{\epsilon}} = \sum_{p=1}^{N_e} \boldsymbol{\epsilon}_p / N_e$, and $\mathbf{B}_{\mathrm{sam}} = \mathbf{X}' \mathbf{X}'^{\mathrm{T}}$ where $\mathbf{X}' = \left(\boldsymbol{\epsilon}'_1, \ldots, \boldsymbol{\epsilon}'_{N_e}\right) / \sqrt{N_e - 1}$.
- We define the *sample correlation matrix*

$$\mathbf{C}_{\mathrm{sam}} = \mathbf{D}^{-1/2}\, \mathbf{X}' \, \mathbf{X}'^{\mathrm{T}} \, \mathbf{D}^{-1/2} = \widehat{\mathbf{X}}' \, (\widehat{\mathbf{X}}')^{\mathrm{T}} \qquad (1)$$

where $\mathbf{D} = \mathbf{D}^{1/2}\,\mathbf{D}^{1/2} = \mathrm{diag}(\mathbf{B}_{\mathrm{sam}})$ and $\widehat{\mathbf{X}}' = \mathbf{D}^{-1/2}\,\mathbf{X}' = \left(\widehat{\boldsymbol{\epsilon}}'_1, \ldots, \widehat{\boldsymbol{\epsilon}}'_{N_e}\right)$.

Correlation Operators based on Implicit Diffusion

- Consider the operator $\widehat{\psi} \xmapsto{C_{\mathrm{dif}}} \psi$ on $\mathbb{R}^d$ defined by the solution of an M-step implicitly-formulated diffusion equation:

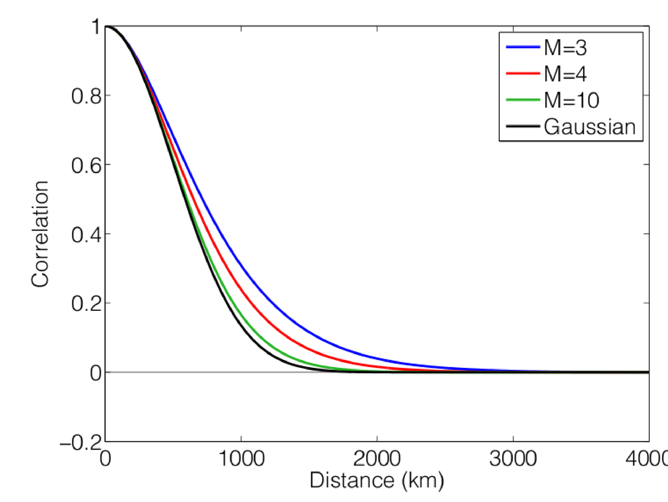
$$\left(1 - \nabla \cdot \boldsymbol{\kappa} \nabla\right)^M \psi(\boldsymbol{x}) = \widehat{\psi}(\boldsymbol{x}) \qquad (2)$$

- For constant $\boldsymbol{\kappa}$, the solution is a convolution of $\widehat{\psi}(\boldsymbol{x})$ with a covariance function from the Whittle-Matérn family ([3], [7]):

$$c_d(r) \sim r^{M-d/2} K_{M-d/2}(r)$$

where $K_{M-d/2}(r)$ is the modified Bessel function of the second kind, and

$$r = \sqrt{(\boldsymbol{x} - \boldsymbol{x}')^{\mathrm{T}} \boldsymbol{\kappa}^{-1} (\boldsymbol{x} - \boldsymbol{x}')}.$$



Method 1: Covariance modelling using Diffusion

- Represent the covariance matrix of the analysis variables as

$$\mathbf{B}_1 = \mathbf{D}^{1/2}\, \mathbf{C}_{\mathrm{dif}}\, \mathbf{D}^{1/2}$$

where $\mathbf{C}_{\mathrm{dif}}$ is a *full-rank* correlation matrix.

- Represent the correlation matrix as ([6])

$$\mathbf{C}_{\mathrm{dif}} = \mathbf{N}^{1/2}\, \mathbf{L}\, \mathbf{W}^{-1}\, \mathbf{N}^{1/2}$$

where $\mathbf{L}$ is the solution operator of Eq. (2), $\mathbf{W} = \mathrm{diag}(\text{metric coefficients})$ and $\mathbf{N} = \mathbf{N}^{1/2}\,\mathbf{N}^{1/2} = \mathrm{diag}(\text{normalization factors})$.

- Compute the *local* diffusion tensor from ([7])

$$\boldsymbol{\kappa}(\boldsymbol{x}) = \frac{1}{2M - d - 2}\, \boldsymbol{H}^{-1}(\boldsymbol{x}) \qquad (3)$$

where $\boldsymbol{H}(\boldsymbol{x})$, the correlation Hessian tensor, can be estimated from ([4])

$$\boldsymbol{H}_{\mathrm{sam}}(\boldsymbol{x}) = \sum_{p=1}^{N_e} \nabla \widehat{\boldsymbol{\epsilon}}'_p(\boldsymbol{x}) \left(\nabla \widehat{\boldsymbol{\epsilon}}'_p(\boldsymbol{x})\right)^{\mathrm{T}}. \qquad (4)$$

Method 2: Localization using Diffusion

- A common way to localize covariances in the EnKF is to compute the Schur (element-by-element) product $\circ$ of the sample correlation matrix (1) with a prescribed localized correlation matrix $\mathbf{C}_{\mathrm{loc}}$:

$$\mathbf{B}_2 = \mathbf{D}^{1/2} \left[\mathbf{C}_{\mathrm{sam}} \circ \mathbf{C}_{\mathrm{loc}}\right] \mathbf{D}^{1/2}. \qquad (5)$$

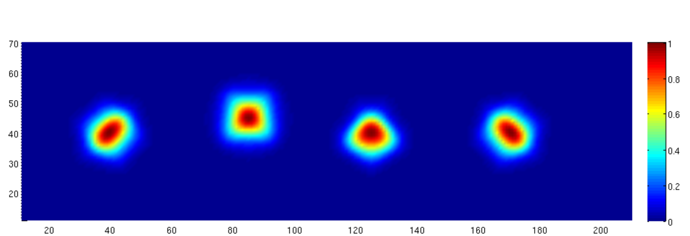
- A more convenient form of Eq. (5) for CG is ([2])

$$\mathbf{B}_2 = \mathbf{D}^{1/2} \left[\sum_{p=1}^{N_e} \mathbf{D}_{\widehat{\boldsymbol{\epsilon}}'_p}\, \mathbf{C}_{\mathrm{loc}}\, \mathbf{D}_{\widehat{\boldsymbol{\epsilon}}'_p}\right] \mathbf{D}^{1/2} \qquad (6)$$

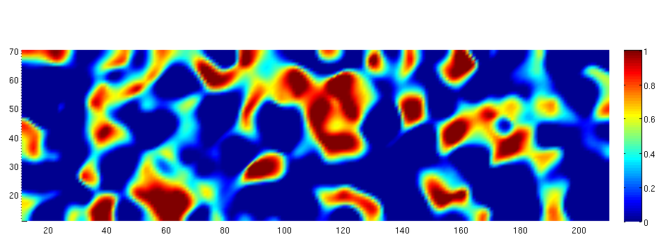
where $\mathbf{D}_{\widehat{\boldsymbol{\epsilon}}'_p} = \mathrm{diag}(\widehat{\boldsymbol{\epsilon}}'_p)$.

- The diffusion operator $\mathbf{C}_{\mathrm{dif}}$ can be used to define $\mathbf{C}_{\mathrm{loc}}$ in Eq. (6).
- The localization function can be made *adaptive* by setting the diffusion tensor in $\mathbf{C}_{\mathrm{loc}} = \mathbf{C}_{\mathrm{dif}}$ to $\alpha \boldsymbol{\kappa}_{\mathrm{sam}}$ where $\boldsymbol{\kappa}_{\mathrm{sam}}$ is estimated from Eqs (3)-(4) and $\alpha > 1$.

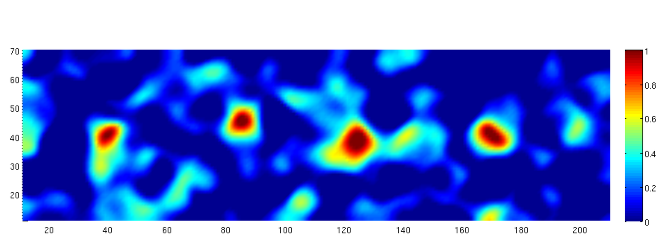
Idealized Example involving Anisotropic and Inhomogeneous Correlations



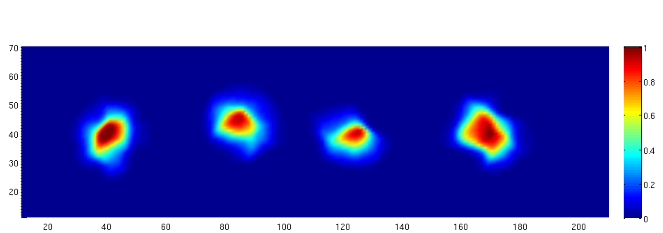
True correlations at selected points



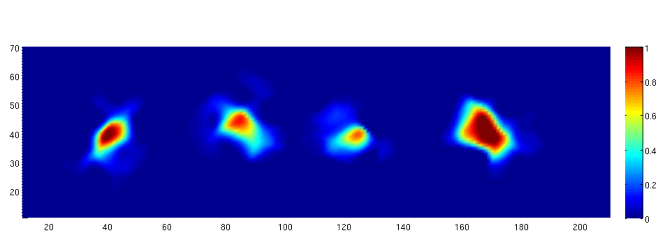
Raw correlations with $N_e = 10$



Raw correlations with $N_e = 100$



Estimated correlations using Method 1 with $N_e = 10$



Estimated correlations using Method 2 with $N_e = 10$ and $\alpha = 2$.

Remarks

- Method 1 requires the estimation of $O(N)$ diffusion tensor elements for each of the analysis variables and one application of $\mathbf{C}_{\mathrm{dif}}$ per CG iteration.
- Method 2 can be implemented with simpler (e.g., non-adaptive) diffusion operators, has greater flexibility for specifying multivariate covariances, but requires N_e applications of $\mathbf{C}_{\mathrm{dif}}$ per CG iteration.
- Spatial filtering of the variances and tensor elements can help reduce the effects of sampling error ([1]).
- Hybrid formulations that combine $\mathbf{B}_1$ and $\mathbf{B}_2$ are being considered for the NEMOVAR ocean data assimilation system ([5]).

References

[1] Berre, L. and G. Desroziers, 2010: *Mon. Wea. Rev.*, **138**, 3693–3720.
[2] Buehner, M., 2012: *Mon. Wea. Rev.*, **140**, 617–636.
[3] Guttorp, P. and T. Gneiting, 2006: *Biometrika*, **93**, 989–995.
[4] Michel, Y., 2012: *Q. J. R. Meteorol. Soc.*, DOI: 10.1002/qj.2007.
[5] Mogensen, K., M. A. Balmaseda and A. T. Weaver, 2012: ECMWF Tech. Memo., No. 668.
[6] Weaver, A. T. and P. Courtier, 2001: *Q. J. R. Meteorol. Soc.*, **127**, 1815–1846.
[7] Weaver, A. T. and I. Mirouze, 2012: *Q. J. R. Meteorol. Soc.*, DOI:10.1002/qj.1955.