Efficient parameterization of the observation error covariance matrix for square root or ensemble Kalman filters: application to ocean altimetry

Jean-Michel Brankart, Clément Ubelmann, Charles-Emmanuel Testut, Emmanuel Cosme, Pierre Brasseur, Jacques Verron

Laboratoire des Ecoulements Géophysiques et Industriels, Grenoble, France

Jean-Michel.Brankart@hmg.inpg.fr

Introduction

In the Kalman filter standard algorithm, the computational complexity of the observational update is proportional to the cube of the number y of observations (leading behaviour for large y). In realistic atmospheric or oceanic applications, this often leads to a prohibitive cost and to the necessity of simplifying the problem by aggregating or dropping observations. If the filter error covariance matrices are in square root form (as in square root or ensemble Kalman filters), the standard algorithm can be transformed to be linear in y, providing that the observation error covariance matrix is diagonal. It is an important drawback of this transformed algorithm often leading to assume uncorrelated observation errors for the sake of numerical efficiency. We show here that the linearity of the transformed algorithm in y can be preserved for other forms of the observation error covariance matrix. In particular, quite general correlation structures (with analytic asymptotic expression) can be simulated by adding gradient observations to the original observation vector.

Analysis update in square root or ensemble Kalman filters

Analysis schemes

If the Kalman filter forecast error covariance matrix is available in square root form, $\mathbf{P}^f = \mathbf{S}^f\mathbf{S}^{fT}$, the use of the Sherman-Morrison-Woodbury (SMW) formula leads (*Pham et al.*, 1998) to analysis equations in the forms ($\delta\mathbf{x}$ is the observational update increment):

$$\delta\mathbf{x} = \mathbf{S}^f[\mathbf{I} + (\mathbf{HS}^f)^T\mathbf{R}^{-1}(\mathbf{HS}^f)]^{-1}(\mathbf{HS}^f)^T\mathbf{R}^{-1}\delta\mathbf{y} \qquad (1)$$

$$\mathbf{S}^a = \mathbf{S}^f[\mathbf{I} + (\mathbf{HS}^f)^T\mathbf{R}^{-1}(\mathbf{HS}^f)]^{-1/2}. \qquad (2)$$

In the EnKF, an ensemble of corrections $\delta\mathbf{x}$ is computed from an ensemble of innovations $\delta\mathbf{y}$. The SMW formula is usually not used (*Evensen*, 2003):

$$\delta\mathbf{x} = \mathbf{S}^f(\mathbf{HS}^f)^T\left[(\mathbf{HS}^f)(\mathbf{HS}^f)^T + \mathbf{R}\right]^{-1}\delta\mathbf{y}. \qquad (3)$$

Computational complexities

We note x the size of the state vector, y the size of the observation vector, and r the number of error modes in the square root filter (ensemble size with the EnKF).

Formulas 1 and 2 are advantageous only if $\mathbf{R}$ can be inverted at low cost. For instance, if $\mathbf{R}$ is diagonal, the asymptotic computational complexity of formulas 1 and 2 are:

$$C_1 \sim yr^2 + \frac{r^3}{6} + xr \qquad \text{and} \qquad C_1^P \sim yr^2 + \frac{r^3}{2} + xr^2 \qquad (4)$$

These expressions are linear in y and x. Large observation vectors can be handled. The computational complexity of formula 3 (applied r times) is

$$C_0^E \sim \frac{y^3}{3} + 2ry^2 + 2rxy \qquad (5)$$

This expression is not linear in y. Formula 3 can then be applied only if y is quite small. A usual strategy consists in localizing the observational updates by subdomains (*Houtekamer and Mitchell*, 1998) and splitting the analysis updates into several steps, each step assimilating a subset of observations. SMW formula can be nonetheless applied in the localization strategy. If $\mathbf{R}$ is diagonal, the computational complexity of the EnKF analysis also becomes linear in y and x.

Linear transformation of the observation vector to simulate correlations

Rationale

The observational update given by formula 1 also minimizes

$$J = \delta\mathbf{x}^T\mathbf{P}^{f^{-1}}\delta\mathbf{x} + (\delta\mathbf{y} - \mathbf{H}\delta\mathbf{x})^T\mathbf{R}^{-1}(\delta\mathbf{y} - \mathbf{H}\delta\mathbf{x}) \qquad (6)$$

If we transform the observation vector by a regular (rank equal to y) linear transformation operator $\mathbf{T}$: $\delta\mathbf{y}^+ = \mathbf{T}\delta\mathbf{y}$, $\mathbf{H}^+ = \mathbf{TH}$, J remains unchanged if

$$\mathbf{R}^{-1} = \mathbf{T}^T\mathbf{R}^{+-1}\mathbf{T} \qquad (7)$$

Simple application: gradient operator in one dimension

Let us introduce the transformation $\mathbf{y}^+ = \mathbf{Ty} = \begin{bmatrix}\mathbf{T}_1\\ \mathbf{T}_2\end{bmatrix}\mathbf{y}$ where $\mathbf{T}_1$ is the identity matrix, $\mathbf{T}_2$ the gradient operator, $\mathbf{T}_{2,ij} = \frac{\delta_{ij}-\delta_{i-1,j}}{\Delta\xi}$. If $\mathbf{R}^+$ is homogeneous, i.e. $\mathbf{R}^+ = \begin{bmatrix}\sigma_0^2\mathbf{I} & 0\\ 0 & \sigma_1^2\mathbf{I}\end{bmatrix}$, then $\mathbf{R}$ verifies

$$\mathbf{R}^{-1} = \frac{1}{\sigma_0^2}\mathbf{I} + \frac{1}{\sigma_1^2\Delta\xi^2}\begin{bmatrix} 1 & -1 & 0 & \cdots & 0 & 0 & 0\\ -1 & 2 & -1 & \ddots & & 0 & 0\\ 0 & -1 & 2 & \ddots & \ddots & & 0\\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots\\ 0 & & \ddots & \ddots & 2 & -1 & 0\\ 0 & 0 & & \ddots & -1 & 2 & -1\\ 0 & 0 & 0 & \cdots & 0 & -1 & 1 \end{bmatrix} \qquad (8)$$

and it can be proven that this is a consistent discretization of the inverse of the covariance function

$$\mathcal{R}(\rho) = \frac{\sigma_0^2}{2}\exp\left(-\frac{|\rho|}{\ell}\right) \qquad \text{with} \qquad \ell = \frac{\sigma_0}{\sigma_1} \qquad (9)$$

Application to altimetry in the North Brazil current

Experiment: A 5-year simulation of the circulation in the North Brazil current region is performed with a regional configuration of the NEMO model. The 300 output snapshots (one every 6 days) determine the *true* states. The mean of this ensemble is taken as the *background* state (see Figure 1, upper panel). To parameterize the background error covariance matrix, we use the covariance of 59 snapshots (one per month over 5 years, except those that are less than 1 month away from the true state). Figure 1, bottom panel, illustrates what the square root of the matrix diagonal looks like.

As *observation*, Sea Surface Height (SSH) is observed over the full domain, with a 4 cm error standard deviation. **Two observation vectors** are generated from the true state: a first one, by adding uncorrelated observation noise, and a second one, by adding a correlated observation noise, with a covariance matrix given by equation 7, with a 2D gradient transformation. The noise is scaled to have a uniform standard deviation $\sigma = 0.04$ m. The observation error covariance is parameterized either with a diagonal matrix, or with a non-diagonal matrix, following the method described previously.

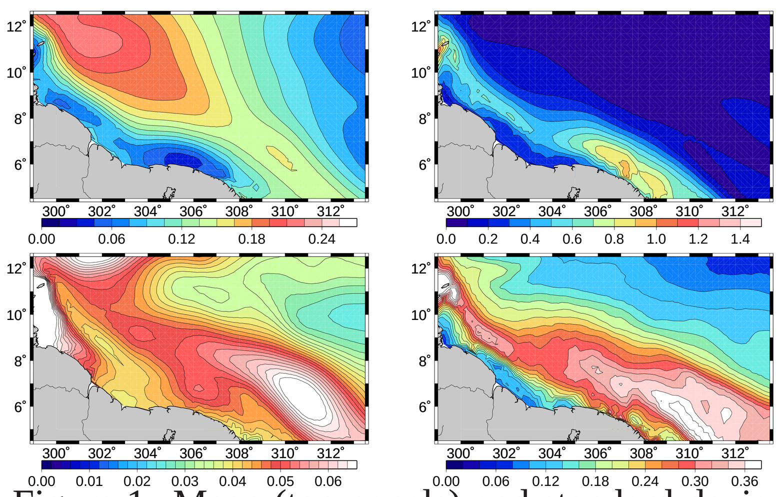
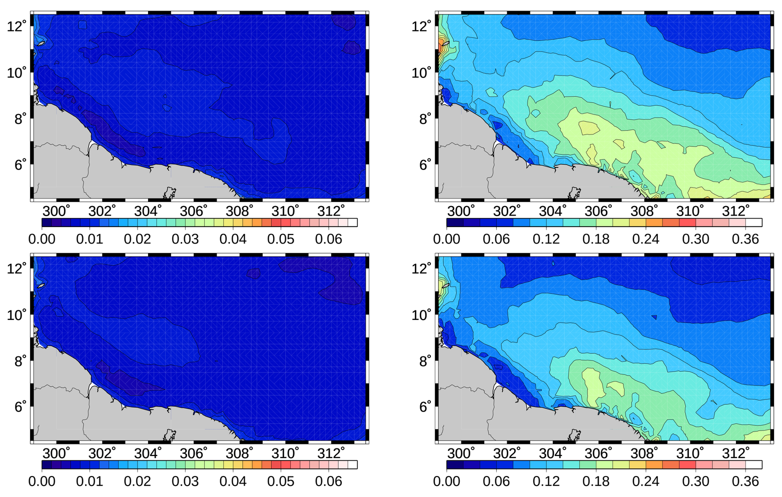


Figure 1: Mean (top panels) and standard deviation (bottom panels) of the 5 years simulation, for the sea surface height (in m, left panels), and sea surface velocity (in m/s, right panels).

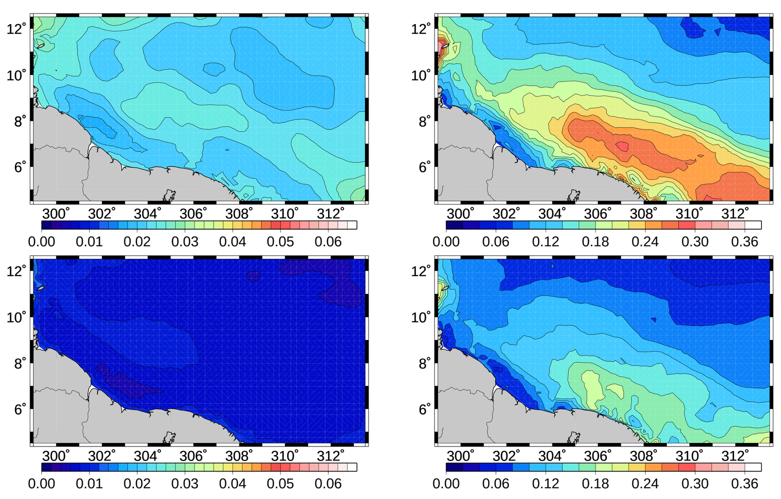
Uncorrelated errors



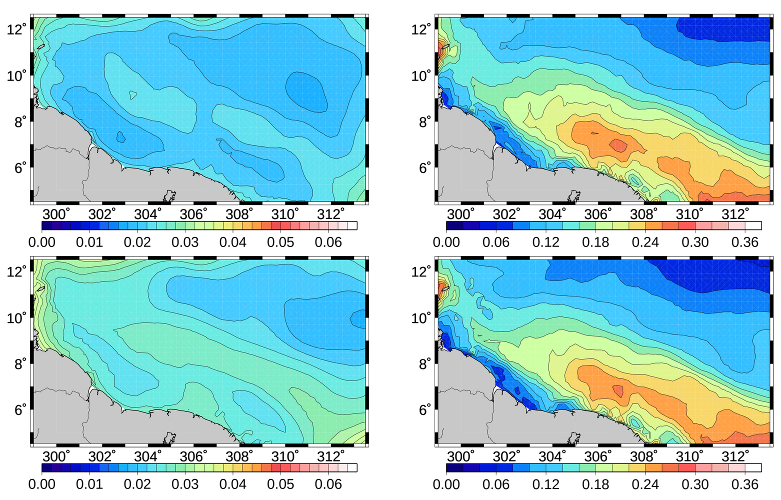
Observation errors are spatially uncorrelated, and $\mathbf{R}$ is diagonal. Figure 2 on the left shows the error standard deviation, as measured using the ensemble (top panels), and as estimated by the scheme (the square root of the diagonal of $\mathbf{P}^a$, bottom panels). It is shown for altimetry (in m, left panels) and for velocity (in m/s, right panels).

Correlated errors, with diagonal $\mathbf{R}$ parameterization

Observation errors are spatially correlated, but $\mathbf{R}$ is taken diagonal. Figure 3 on the right displays the same fields as Figure 2, for this experiment. The inappropriate parameterization of $\mathbf{R}$ leads to a significant discrepancy between the errors estimated by the ensemble and by the error modes.



Correlated errors, with consistent $\mathbf{R}$ parameterization



Observation errors are spatially correlated, and $\mathbf{R}$ is parameterized using the observation gradient method. The coherence between errors estimated with the ensemble and with the error modes is restored.

Conclusions

Classical algorithms to compute the observational update in Kalman filters are penalized by a computational complexity proportional to the cube of the number of observations. In square root or ensemble Kalman filters, this algorithm can be modified to become linear in the number of observations if the observation error covariance matrix is diagonal. Here it has been demonstrated that these benefits can be preserved with a non-diagonal parameterization of the observation error covariance matrix $\mathbf{R}$. The method simulates correlations by application of a linear transformation of the observation vector (with diagonal $\mathbf{R}$ in the transformed space). It is shown especially efficient to describe simple correlation structures if gradient observations can be added to the observation vector. This is possible for instance if the observations are distributed along lines or at the nodes of two-dimensional grids so that discrete gradients can be computed by substracting successive observations.

The method has been tested with the aim of reconstructing the circulation of the North Brazil current, as simulated by a $1/4°$ model of the Tropical Atlantic Ocean, using synthetic altimetric observations. Assuming a diagonal observation error covariance matrix in presence of a correlated noise leads to a non-optimal solution that underestimates the error variance. Optimal parameterizations of the observation error covariance matrix usually produce solutions that are close to minimizing the resulting error.

This work was conducted as part of the MERSEA project funded by the E.U. (Contract No. AIP3-CT-2003-502885), with additional support from CNES. The calculations were performed with the support of IDRIS/CNRS.

References

Evensen, G., The ensemble Kalman filter: Theoretical formulation and practical implementation, *Ocean Dynamics*, *53*, 343–367, 2003.

Houtekamer, P. L., and H. L. Mitchell, Data assimilation using an Ensemble Kalman Filter technique, *Monthly Weather Review*, *126*, 796–811, 1998.

Pham, D. T., J. Verron, and M. C. Roubaud, A singular evolutive extended Kalman filter for data assimilation in oceanography, *J. Marine. Sys.*, *16*, 323–340, 1998.