

1. Context

Kalman filter equation

The aim of data assimilation is to provide the estimation of an unknown state \mathbf{x}_q^t of a system (atmosphere, ocean,...) knowing observations \mathcal{Y}_q^o and the dynamics of the system. In the case where the estimation error is described by Gaussian distributions and for linear dynamics, the Kalman filter equations [Kalman, 1960] provide the time evolution of the mean and of the error covariance matrix, following

$$\begin{cases} \mathbf{A}_q = (\mathbf{I} - \mathbf{K}_q \mathbf{H}_q) \mathbf{B}_q, \\ \mathbf{B}_{q+1} = \mathbf{M}_{q+1 \leftarrow q} \mathbf{A}_q (\mathbf{M}_{q+1 \leftarrow q})^T, \end{cases} \quad (1)$$

Direct computation of equations Eq.(1) is impossible for large dimension geophysical application, and alternative strategies have been introduced to implement this algorithm, e.g. the Ensemble Kalman Filter (EnKF) [Evensen, 2009]. Present implementations of the EnKF relies on dozen of members, and in order to limit the sampling noise, filtering strategies are often introduced.

The parametric alternative

In this contribution we propose an alternative procedure to solve the Kalman filter equations. This relies on a parametric approximation of the covariance matrix \mathbf{B} and \mathbf{A} based on the description of their variance and local metric tensor fields.

2. Parametric formulation of the Kalman filter equations

Parametric approximation of covariance matrix

In this contribution we propose a new implementation that relies on a parametric description of covariance matrices in terms of variance field (diagonal of covariance matrix) and of local metric fields. The variance is defined as the diagonal of the covariance matrix, $V^b(\mathbf{x}) = \mathbf{B}(\mathbf{x}, \mathbf{x})$. The local metric \mathbf{g}_x of an error correlation function $\rho(\mathbf{x}, \mathbf{y}) = \mathbf{B}(\mathbf{x}, \mathbf{y}) / \sqrt{V_x V_y}$ is defined from the second order Taylor expansion

$$\rho(\mathbf{x}, \mathbf{x} + \delta \mathbf{x}) \underset{\delta \mathbf{x} \rightarrow 0}{\approx} 1 - \frac{1}{2} \|\delta \mathbf{x}\|_{\mathbf{g}_x}^2 + o(\|\delta \mathbf{x}\|^2), \quad (2)$$

where $\|\mathbf{x}\|_{\mathbf{g}}^2 = \mathbf{x}^T \mathbf{g} \mathbf{x}$. The construction of covariance matrix from the variance field and from the local metric field can be achieved considering the covariance model based on the diffusion equation [Weaver and Courtier, 2001]. In this case, the local diffusion tensor ν_x is related to the local metric tensor [Pannekoucke and Massart, 2008] following $\nu_x = \frac{1}{2} \mathbf{g}_x^{-1}$. Now in order to solve the Kalman filter equations Eq.(1) we describe how the variance and the diffusion tensor are evolving in time.

Parametric formulation of the Kalman filter analysis equation

From the assimilation of a single observation, we are able to feature the update formulation of variance and local metric tensor [Pannekoucke et al., 2016], and in particular

$$\nu_x^a = \frac{V_x^a}{V_x^b} \nu_x^b. \quad (3)$$

Then, by iteration of single observations, we deduce the algorithm 1

Require: Fields of ν^b and V^b , V^o and location \mathbf{x}_j of the p observations to assimilate

- 1: **for** $j = 1 : p$ **do**
- 2: 0- Initialization of intermediate quantities
- 3: $V_j^b = V_{\mathbf{x}_j}^b$, $V_j^o = V_{\mathbf{x}_j}^o$, $\nu_j = \nu_{\mathbf{x}_j}^b$
- 4: $\rho_j(\mathbf{x}) = \exp\left(-\frac{1}{4} \|\mathbf{x} - \mathbf{x}_j\|_{\nu_j^{-1}}^2\right)$
- 5:
- 6: 1- Computation of analysis statistics
- 7: $V_x^a = V_x^b \left(1 - \rho_j^2(\mathbf{x}) \frac{V_j^b}{V_j^b + V_j^o}\right)$
- 8: $\nu_x^a = \nu_x^b \left(1 - \rho_j^2(\mathbf{x}) \frac{V_j^b}{V_j^b + V_j^o}\right)$
- 9:
- 10: 2- Update of the background statistics
- 11: $V_x^b \leftarrow V_x^a$
- 12: $\nu_x^b \leftarrow \nu_x^a$
- 13: **end for**
- 14: Return fields ν^a and V^a

Algorithm 1: Iterated process building analysis covariance matrix at the leading order, under Gaussian shape assumption.

Parametric formulation of the Kalman filter forecast equation

For the particular linear advection-diffusion dynamics

$$\partial_t \alpha + c \partial_x \alpha = \kappa \partial_x^2 \alpha, \quad (4)$$

where c denotes the velocity and κ the diffusion rate, the time evolution of the variance and the diffusion tensor is given at the lead order by

$$\begin{cases} \partial_t \nu^b + \mathbf{u} \nabla \nu^b = \nu^b (\nabla \mathbf{u})^T + (\nabla \mathbf{u}) \nu^b + 2\kappa, \\ \partial_t V^b + \mathbf{u} \nabla V^b = -V^b \text{Tr} \left[(\nu^b)^{-1} \kappa \right]. \end{cases} \quad (5)$$

We reform here the conclusion of [Cohn, 1993], and extend Cohn's result by including the effect of the diffusion. The parametric covariance dynamics is then computed following the algorithm 2.

Require: Fields of ν^a and V^a . $\delta t = \tau/N$, $t = 0$

- 1: **for** $k = 1 : N$ **do**
- 2: 1- Pure advection
- 3: $D(\mathbf{x}) = \mathbf{x} + \mathbf{u}(\mathbf{x}, t) \delta t$
- 4: $\hat{\nu}_x^b(t + \delta t) = \mathbf{D}_x \nu_{D^{-1}(\mathbf{x})}^b(t) \mathbf{D}_x^T$
- 5: $\hat{V}_x^b(t + \delta t) = V^b[D^{-1}(\mathbf{x}), t]$
- 6:
- 7: 2- Pure diffusion
- 8: $\nu_x^b(t + \delta t) = \hat{\nu}_x^b(t + \delta t) + 2\kappa(\mathbf{x}) \delta t$
- 9: $V_x^b(t + \delta t) = \hat{V}_x^b(t + \delta t) \frac{V_x^b(t + \delta t)}{V_x^b(t + \delta t)^{1/2}}$
- 10:
- 11: 3- Update of the background statistics
- 12: $V_x^b(t) \leftarrow V_x^b(t + \delta t)$
- 13: $\nu_x^b(t) \leftarrow \nu_x^b(t + \delta t)$
- 14: $t \leftarrow t + \delta t$
- 15: **end for**
- 16: Return fields $\nu_x^b(\tau)$ and $V_x^b(\tau)$

Algorithm 2: Lagrangian algorithm that details the iteration process to forecast the background covariance matrix at time $t = \tau$ from the analysis covariance matrix given at time $t = 0$, under local homogeneity assumption.

3. Numerical experience

In order to mimic situations encountered in chemical transport model, the simple linear advection-diffusion transport Eq.(4) of a passive species α is considered.

For the simulation, the velocity is set to $c = 1$ and the time step δt is fixed to the advection time step $\delta t_{adv} = \delta x/c$. The diffusion rate is set so that the diffusion time step $\delta t_{diff} = \delta x^2/\kappa$ is equal to six times the advection time scale δt_{adv} .

The **observational network** considered here is set to measure half the domain from 180° to 360° , with one observation per grid-point.

In this particular one dimensional setting, the local metric tensor verifies $\mathbf{g}_x = \frac{1}{L_x^2}$ where L_x is called the **error correlation length-scale**, $L_x = \sqrt{2\nu_x}$. The initial background statistics are set following Fig.1, where the correlation is constructed from the model based on the diffusion equation, with the diffusion coefficient field $\nu_x = \frac{1}{2} L_x$.

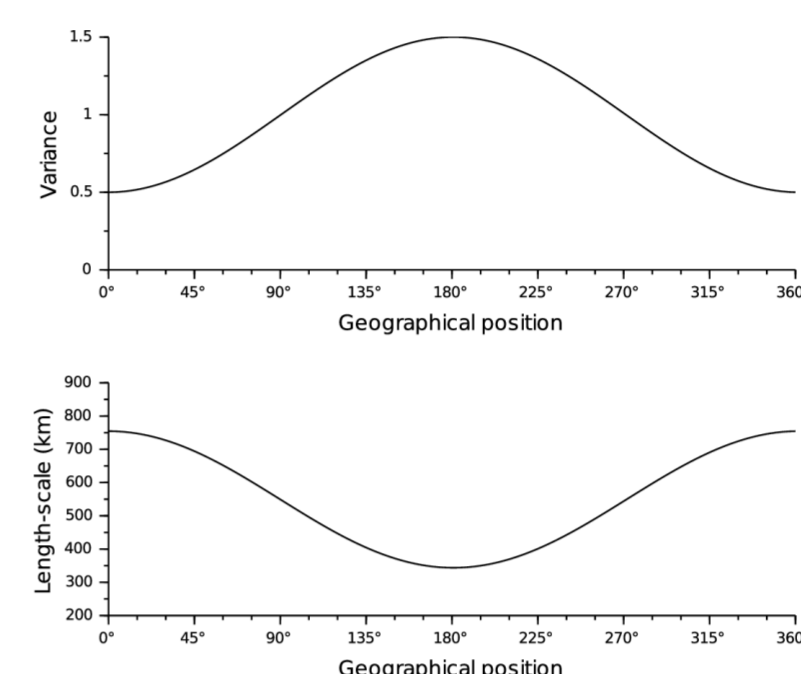


Figure -1- Initial condition for the background error variance field (top) and length-scale field (bottom).

Fig.2 shows the comparison between the variance and the length-scale fields computed from the Kalman filter versus the ones computed using the **Parametric Kalman Filter (PKF) described by Algorithms 1 and 2**. The results of the covariance dynamics as described from the Kalman filter equations are in continuous lines. The results obtained from PKF analysis and forecast steps is reproduced in dashed line. The KF and the PKF can be compared with the update process proposed by [Dee, 2003] (see [Menard et al., 2000]), thereafter called Parametric homogeneous Kalman filter (PhKF), where only the variance is updated while the correlation function is isotropic and remains of constant length-scale $L_h = 500\text{km}$ (dashed-dotted line).

This diffusive advection (Fig. 2) shows that the PKF is similar to the Kalman filter. The PKF provide a better approximation of the KF equations than the [Dee, 2003] scheme. For this case the PKF is able to reproduce the variance attenuation due to the diffusion process, while the [Dee, 2003] scheme fails to reproduce the KF reference. Comparing with the KF, the PKF is able to reproduce the time increase of the length-scale due to the physical diffusion term $\kappa > 0$.

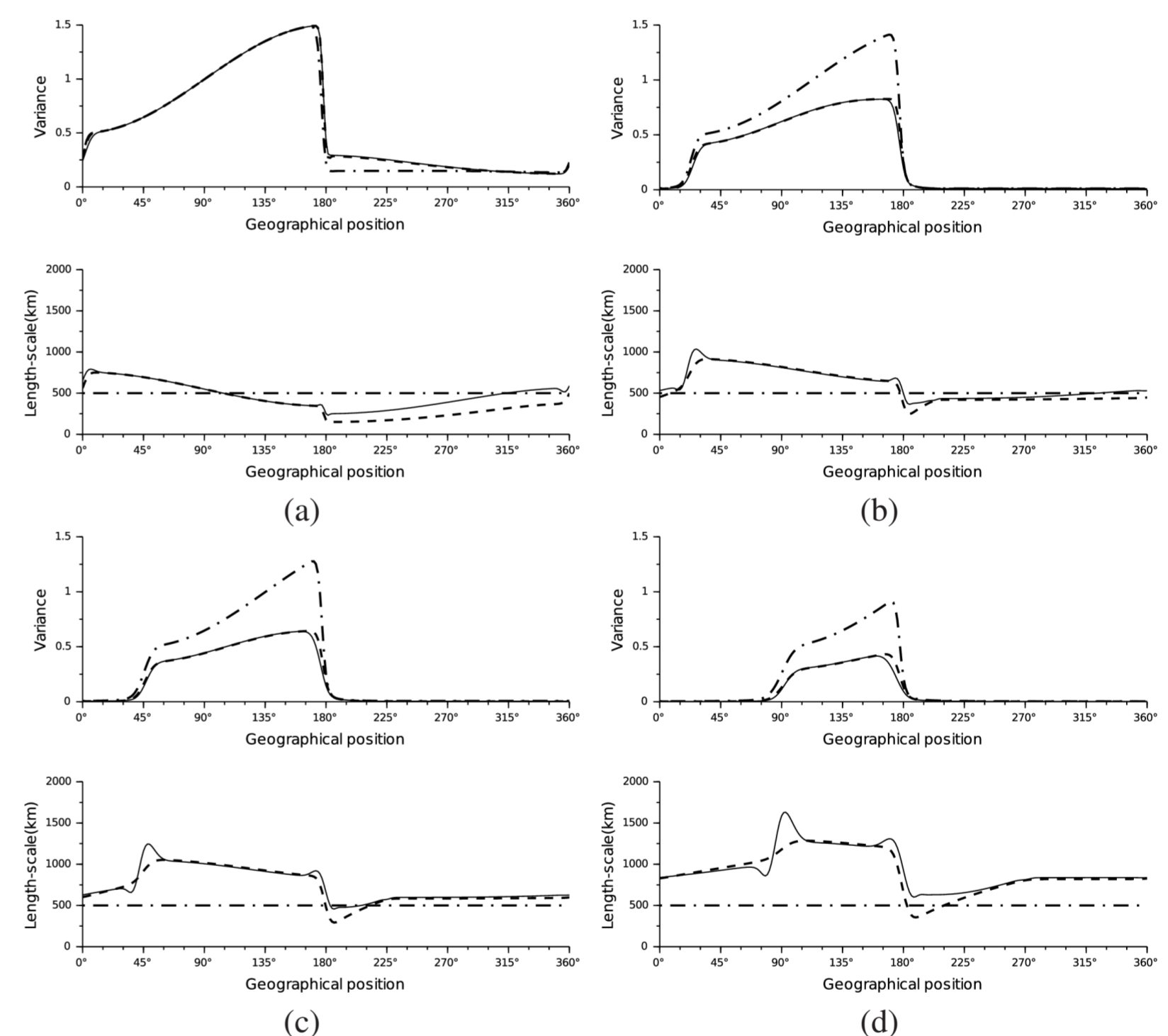


Figure -2- Diagnosis of the analysis covariance matrix at iterations 1 (a), 15 (b), 30 (c) and 60 (d) in case of an advection-diffusion dynamics. KF time evolution (continuous line), the PKF (dashed line) and the PhKF (dash dotted line).

4. Conclusions and Perspectives

A computational simplification of the Kalman filter is introduced - the Parametric Kalman filter. **The full covariance matrix dynamics of the Kalman filter**, that describes the evolution along the analysis and forecast cycle, **is replaced by the dynamics of the error variance and the diffusion tensor**, which is related to the correlation length-scales. The Parametric Kalman filter developed here has been applied to the simplified framework of advection-diffusion of a passive tracer, for its use in Chemical Transport Model assimilation. The Parametric Kalman filter is easy to compute and, in this setting, its computational cost is smaller than those of the EnKF. The validation of the method is presented for a simplified one-dimensional advection-diffusion dynamics.

Extension to the nonlinear framework should be investigated e.g. by considering the Burger equation, with **application to uncertainty quantification**.

5. References

- [1] Cohn, S. (1993). Dynamics of short-term univariate forecast error covariances. *Monthly Weather Review*, 121(11):3123-3149.
- [2] Dee, D. (2003). An adaptive scheme for cycling background error variances during data assimilation. In *ECMWF/GEWEX workshop on humidity analysis pp.15 (July 2002)*.
- [3] Evensen, G. (2009). *Data Assimilation: The Ensemble Kalman Filter*. Springer.
- [4] Kalman, R. E. (1960). A new approach to linear filtering and prediction problems. *Journal Basic Engineering*, 82:35-45.
- [5] Menard, R., Cohn, S., Chang, L.-P., and Lyster, P. M. (2000). Assimilation of stratospheric chemical tracer observations using a kalman filter. part I: Formulation. *Monthly Weather Review*, 128:2554-2571.
- [6] Pannekoucke, O. and Massart, S. (2008). Estimation of the local diffusion tensor and normalization for heterogeneous correlation modelling using a diffusion equation. *Q. J. R. Meteorol. Soc.*, 134:1425-1438.
- [7] Pannekoucke, O., Ricci, S., Barthelemy, S., Menard, R., and Thual, O. (2016). Parametric kalman filter for chemical transport model. *Tellus*, pages XX-XX.
- [8] Weaver, A. and Courtier, P. (2001). Correlation modelling on the sphere using a generalized diffusion equation (tech. memo. ecmwf, num. 306). *Quarterly Journal Royal Meteorological Society*, 127:1815-1846.