

Taking into account correlated observation errors by progressive assimilation of multiscale information

Vincent Chabot, Maëlle Nodet, Arthur Vidard

▶ To cite this version:

Vincent Chabot, Maëlle Nodet, Arthur Vidard. Taking into account correlated observation errors by progressive assimilation of multiscale information. American Geophysical Union Fall Meeting, Dec 2016, San Francisco, United States. 2016. hal-01402906

HAL Id: hal-01402906 https://inria.hal.science/hal-01402906

Submitted on 25 Nov 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers. L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0 International License

Taking into account correlated observation errors by progressive assimilation of multiscale information Univ. Grenoble Alpes, INRIA, France Vincent Chabot, Maëlle Nodet, Arthur Vidard Abstract Results 2: Inhomogeneous observation error Error covariance matrix representation The observation error is still Gaussian and spatially correlated, but ▶ The description of **correlated observation error statistics** is a Crucial choice for B and R: drive the way information is spread and these correlations are now inhomogeneous in space: how redundancy of error is delt with. challenge in data assimilation. $Y_{t_i} = Y_{t_i}^t + \epsilon$ with $\epsilon = A^T D_A^{1/2} \beta \quad \beta \sim \mathcal{N}(0, I)$ Currently, the observation errors are assumed uncorrelated (the ► So far: strong research effort for **B** modelling. covariance matrix is diagonal) which is a **severe approximation R** matrix: mostly assumed diagonal. so that D_A is the exact representation of **R** in the wavelet space. that leads to suboptimal results. ▷ Main reason for this: management greatly simplified. ▶ It is possible to use **multi-scale transformations** to retain the \triangleright Number of observations in general non constant over time \rightarrow Pixels ---diagonal matrix approximation while accounting for some correlation. prevents from using the same **R** matrix. Wavelet ----However this approach can lead to some convergence problems due to Consequelntly: at each assimilation cycle, a new **R** matrix should scale interactions. be formed and inverted. ► We propose an **online scale selection algorithm** that improves the convergence properties in such case. **Our work:** change of variable, which allows diagonal \mathbf{R} + non trivial correlations. Linear changes of variable A: assuming observation errors to be 10 Motivation: account for correlated satellite observation errors additive, unbiased and gaussian, i.e. $Y = Y^t + \epsilon$ with $\epsilon \sim \mathcal{N}(\mathbf{0}, \mathbf{R}), Y^t$ being the true signal, then $\mathbf{A}Y = \mathbf{A}Y^t + \beta$ with $\beta \sim \mathcal{N}(0, \text{ARA}')$. Numerical weather prediction requires the determination of the initial 10 100 20 ► Choice of **A** such that $D_A = \text{diag}(ARA^T) \simeq ARA^T \rightarrow \text{we retain}$ state of the system. Indeed, the true state, at a given moment and in Iterations all points of space, is not accessible. In order to retrieve an optimal a diagonal approximation: Figure caption: Ratio of residual errors initial condition one uses the so called data assimilation methods that $(Y - \mathcal{H}(X))^T R^{-1}(Y - \mathcal{H}(X))$ $r = (X^{t} - X)/(X^{t} - X^{b})$ along minimisation iterations for both combine information from observations, model equations and their $\simeq (Y - \mathcal{H}(X))^T A^T D_{\Delta}^{-1} A(Y - \mathcal{H}(X))$ pixel and wavelet based distances in presence of an inhomogeneous respective error statistics.

- Since the late 70s, satellites are a dominant source of information. Errors associated to such data are highly correlated in space, which can be detrimental if this is not properly accounted for. However their density in space allows for the efficient use of multi-scale transformation, which in turn permit a cheap but good approximation of said error statistics representation.
- For homogeneous spatially correlated Gaussian observation errors this approach is very efficient.
- For more complex errors, however, it can severely damage the convergence properties of the assimilation methods.
- Here we present, through a simple case mimicking a laboratory experiment, an illustration of the above-mentioned problem and a possible solution using scale selection during the assimilation process.

Data assimilation: Principle

What is data assimilation?

Combine at best different sources of information to estimate the state of a system:

- model equations
- observations, data
- background, a priori information
- errors statistics

What is data assimilation for?

- historically: initial state estimation, for weather forecasting.
- today, many other applications...
- ▶ ... and many other application domains.

Here: orthonormal wavelet transform for the operator A (see more options in 2015 reference).

Experimental settings

Experimental framework: mimics the drift of a vortex on a turntable. CORIOLIS experimental turntable (Grenoble, France): simulates the evolution of a vortex in the atmosphere. 1 rotation of the tank = 1 earth rotation.

Vortex creation: by stirring the water and made visible thanks to the addition of a passive tracer (fluorescein). Photographs of the vortex constitute the observed image sequence.

Numerical configuration

Shallow-water equations for the horizontal velocity $w(\mathbf{x}, t) = (u(\mathbf{x}, t), v(\mathbf{x}, t))$, and the water elevation $h(\mathbf{x}, t)$:

 $\begin{cases} \partial_t u - (f + \zeta)v + \partial_x B = -ru + \kappa \Delta u \\ \partial_t v + (f + \zeta)u + \partial_y B = -rv + \kappa \Delta v \\ \partial_t h + \partial_x (hu) + \partial_y (hv) = \mathbf{0}. \end{cases}$

with ζ the relative vorticity and B the Bernouilli potential:

$$\zeta = \partial_x v - \partial_y u, \quad B = g^* h + \frac{u^2 + v^2}{2}$$

where g^* is the reduced gravity. Coriolis parameter: β -plane, $f = f_0 + \beta y$, κ diffusion coeff., r bottom friction coeff.

Observation operator

Vortex temporal evolution: fluorescein concentration evolution, observed by an image sequence of the concentration of a passive tracer

observation error.

Comment: Even though the observation error is exactly represented in the Wavelet case, its minimisation is struggling to converge toward a minimum similar or better than for the pixel case. Main reason for this: some sort of aliasing in the small scales, as illustrated below:



Figure caption: Discrepancy between the background initial concentration and the successive observations $||Y_{t_i} - \mathcal{H}(X_0^b)||_X^2$ along time, as would be measured by the observation term of the cost function, with X = diag(R) for Pixel and $X = ARA^T$ for Wavelet.

Comments:

- Blue / pixel: It starts with a small value (the only difference comes from the noise) and, as time goes by, the vortex drifts and the difference with the initial concentration steadily increases.
- ► Green / wavelet: It shows a steep increase at the beginning, but then

Best Linear Unbiased Estimation (BLUE)

Least squares analysis: solve the inverse problem $\mathbf{y}^o = \mathcal{H}(\mathbf{x}^t) + \epsilon^o$, given a background estimate \mathbf{x}^f of the true input parameters \mathbf{x}^t , where:

- y^o incomplete observations, errors ε^o unbiased, covariance matrix R
 x^f = x^t + ε^f, ε^f background errors unbiased, covariance P^f
- \blacktriangleright observation operator $\mathcal{H}=\mathbf{H}$ maps linearly the input parameters to the observations

Solution: Best linear unbiased estimator (BLUE):

 $\begin{cases} x^{a} = (I - KH) x^{f} + K y^{o} = x^{f} + K(y^{o} - H(x^{f})) \\ K = P^{f} H^{T} (HP^{f} H^{T} + R)^{-1} \end{cases}$

Variational equivalence: x^a is the minimizer of $J(x) = (x - x^f)^T P^f (x - x^f) + (Hx - y^o)^T R(Hx - y^o)$

Variational Data Assimilation

Let \mathcal{M} be a dynamical model describing the evolution of the state variable X in space and time:

 $\| \partial_t X(X_0, \mathbf{x}, t) + \mathcal{M}(X(X_0, \mathbf{x}, t)) = 0 \\ \| X(X_0, \mathbf{x}, t_0) = X_0$

Let Y(t) be (partial) observations of this state variable. The aim of data assimilation is to estimate an optimal initial condition X_0^a (often called analysed state) so that it is not far from the first \boldsymbol{q} transported by the velocity field \boldsymbol{w} :

$$\begin{cases} \partial_t q + \nabla q \cdot w - \nu_T \Delta q = 0 \\ q(t_0) = q_0. \end{cases}$$

so that:

 $\mathcal{H}(X_{t_i}) = q(t_i).$

Twin Experiments Configuration

Synthetic observations: created thanks to a model simulation from a known "true state"; then an assimilation experiment is performed starting from another "background" state using the synthetic observations. Result can be compared with the synthetic truth.



Figure caption: "True" initial concentration of the passive tracer (first left) and noisy observations at initial time, after 90 minutes, 150 min and 270 min (right).

Results 1: Homogeneous observation error

The observation are obtained by adding a spatially correlated Gaussian noise to selected snapshot of the true trajectory

 $Y_{t_i} = Y_{t_i}^t + \epsilon$ with $\epsilon = \sim \mathcal{N}(\mathsf{R},\mathsf{I})$

In order to mimic the usual approach, only diagonal approximations of ${\bf R}$ are used:

$$R_{pix} = diag(R)$$
 $R_{wav} = D_A = diag(ARA^T)$

oscillate around a 'plateau'. This happens because, at this point, the norm is really dominated by the small scales. This is expected, since they are the least affected by the correlated noise, so their associated error variances are the smallest (i.e. one trusts more the small scales).

Red, black and purple / wavelets + removing the 1, 2 and 3 finest scales in the multi scale decomposition respectively. The problem appears later (i.e. for larger discrepancies) when removing the finest scales and even disappear for the purple one.

Results 3: Improvement by online scale selection

Idea: First assimilate the large scales, in order to reduce the difference between $\mathcal{H}(X)$ and Y, and to be in the monotonic region of the green curve, and then progressively includes the smaller scales.



Figure caption: Ratio of residual errors along minimisation iterations for pixel and wavelet-based and progressive wavelet-based distances in presence of an inhomogeneous observation error.

guess X_0^b (in general coming from a previous forecast), and that the model trajectory $X(X_0^a, x, t)$ is close to the observations Y(t).

This is done by defining X_0^a as the minimum of the cost function: $J(X_0) = J_b(X_0) + J_o(X_0)$ $= \frac{1}{2} ||X_0 - X_0^b||_V^2 +$ $\frac{1}{2} \sum_{t_i=t_0}^{t_f} ||Y(t_i) - \mathcal{H}(X(X_0, \mathbf{x}, t_i))||_{\mathcal{O}}^2$

where V is the model state space, \mathcal{O} the observation space and $\mathcal{H}: V \mapsto \mathcal{O}$ the observation operator. Usually, in variational data assimilation, the minimisation is done using a gradient descent type algorithm and the gradient is computed using adjoint methods. Typically in data assimilation one uses the Mahalanobis distance:

 $\|\cdot\|_{V}^{2} = \|\cdot\|_{B}^{2}, \quad \|\|_{\mathcal{O}}^{2} = \|\cdot\|_{R}^{2}, \quad \text{with } \|X\|_{K}^{2} = X^{T}K^{-1}X$ where **R** and **B** are the observation and background error covariance matrices respectively.



Figure caption: Ratio of residual errors

 $r = (X^t - X)/(X^t - X^b)$ along minimisation iterations for both pixel and wavelet based distances in presence of an homogeneous observation error.

Comment: Even though D_A is only an approximation of R, accounting for some part of the spatial correlation is clearly beneficial.

Conclusion: Using multi-scale transforms as changes of variables to represent observation error correlation in data assimilation is a cost effective and promising approach. However this can lead to convergence problems in some cases. A progressive assimilation of the finest scale can significantly improve the convergence, making it a more robust approach.

References

V. Chabot, M. Nodet, N. Papadakis, and A. Vidard, Accounting for observation errors in image data assimilation, Tellus A, vol. 67, no. 0, pp. 4117?19, Feb. 2015.

V. Chabot, M. Nodet, and A. Vidard, Progressive assimilation of multiscale observations, IEEE Transactions, 2016.

Papers online: https://team.inria.fr/airsea/maelle/#publi

maelle.nodet@univ-grenoble-alpes.fr, https://team.inria.fr/airsea/maelle