Modelling spatial observation error correlations using a diffusion operator on unstructured grids

Oliver Guillet

CERFACS / Météo France - CNRM

July 2018

Co-authors : S. Gratton[a], S. Gurol[b], Y. Michel[c] X. Vasseur[d], A. Weaver[b]

[a] INPT-IRIT, Toulouse, France

[b] CERFACS, Toulouse, France

[c] METEO-FRANCE, Toulouse, France

[d] ISAE-SUPAERO, Toulouse, France

Accounting for correlated observations in variational DA

- In variational DA, we aim to minimize the cost-function

$$\mathcal{J}(\mathbf{x}) = \frac{1}{2}\|\mathbf{x} - \mathbf{x}^b\|^2_{\mathbf{B}^{-1}} + \frac{1}{2}\|\mathcal{H}(\mathbf{x}) - \mathbf{y}^o\|^2_{\mathbf{R}^{-1}}$$

- Using a diagonal approximation for **R** is often done for algorithmic convenience but leads to a **sub-optimal solution**. Assumption justified through thinning and variance inflation.
- Work has been done to account for **temporal correlations** in surface pressure data (Jarvinen et al. (1999)) and **interchannel correlations** in satellite data (Bormann et al. (2010), Weston et al. (2014), Stewart et al. (2013)).
- But little work has been done to account for **spatial correlations** in observation error.

Accounting for correlated observation errors in VDA

- Brankart et al. (2008) proposed a method to model the inverse of the observation error correlation matrix based on the assimilation of the **successive derivatives** of the observed field.
- However, this method seems hard to generalize to **heterogeneously distributed data**.
- Fisher (2010) proposed interpolating the observations on a Cartesian grid. Approach revisited by Michel (2017).
- But getting access to R^{-1} is **numerically costly**.

Spatial correlation operators

We introduce a new method to model spatial correlations, designed for **large data sets** with a priori **unknown spatial distribution**.

Goal : Compute the product $\mathbf{Rz}$ for any $\mathbf{z}$, without storing explicitly $\mathbf{R}$ in memory.

Solution : Replace the computation of $\mathbf{Rz}$ with the solution of a **differential equation** with initial condition $\mathbf{z}$.

Continuous model (implicit diffusion) :

$$\frac{1}{\gamma_c}(1 - L^2\Delta)^m f_m(\mathbf{z}) = f_0(\mathbf{z})$$

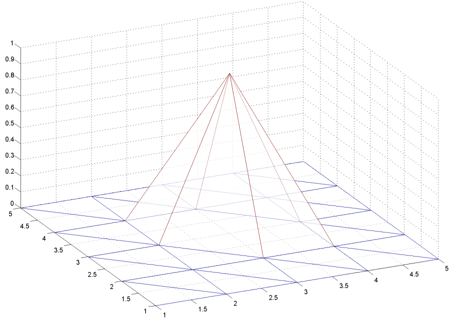
where γ_c is a normalization constant and L is a scale parameter.

Discretization with FEM

Discretization using standard P_1 finite elements :

$$(\mathbf{M} + \mathbf{K})\alpha_{n+1} = \mathbf{M}\alpha_n$$

where $0 \leq n < m$. **M** and **K** are called the mass matrix and the stiffness matrix, respectively.



Formulation of the correlation operators

The correlation operator is formulated as the product of the diffusion operator with the metric $\mathbf{M}^{-1}$:

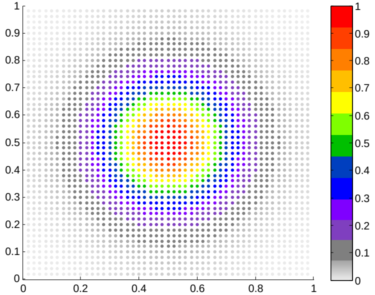
$$\mathbf{C} = \mathbf{\Gamma}[(\mathbf{M}+\mathbf{K})^{-1}\mathbf{M}]^m\mathbf{M}^{-1}\mathbf{\Gamma}$$

where $\mathbf{\Gamma}$ contains normalization factors. Inverting this equation leads to :

$$\mathbf{C}^{-1} = \mathbf{\Gamma}^{-1}\mathbf{M}[\mathbf{M}^{-1}(\mathbf{M}+\mathbf{K})]^m\mathbf{\Gamma}^{-1}$$

By including the standard deviation matrices $\mathbf{\Sigma}$, we get :

$$\mathbf{R} = \mathbf{\Sigma}\mathbf{C}\mathbf{\Sigma} \text{ and } \mathbf{R}^{-1} = \mathbf{\Sigma}^{-1}\mathbf{C}^{-1}\mathbf{\Sigma}^{-1}$$



Expected response to a Dirac impulse on a Cartesian mesh. Amplitude should be 1 at the origin.

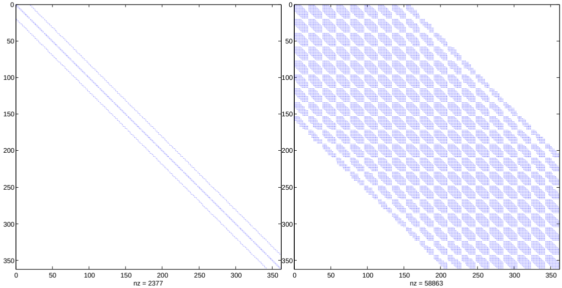
Advantage of sparsity

Advantages of the finite element method :

- The method is responsible for **the sparsity of matrices M and K** in the linear system :

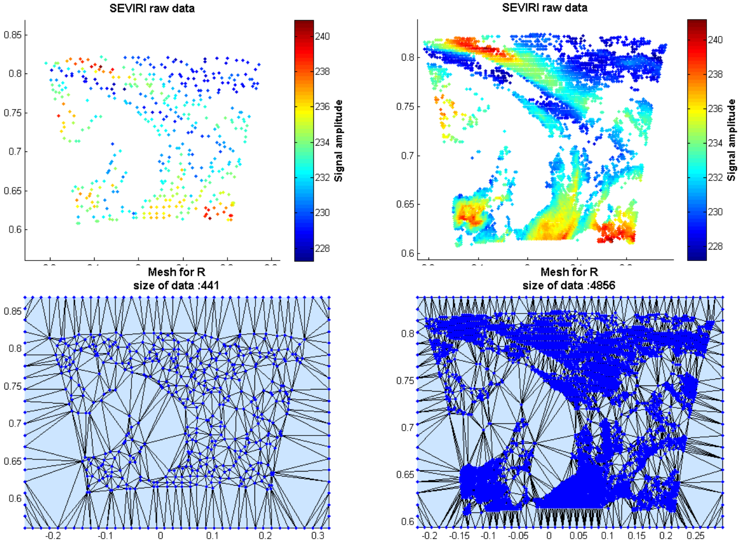
$$(\mathbf{M} + \mathbf{K})\alpha_{n+1} = \mathbf{M}\alpha_n$$

- Boundary conditions are easy to implement.



Left : profile of **M**. Right : profile of $\mathbf{R}^{-1}$

Varying resolutions

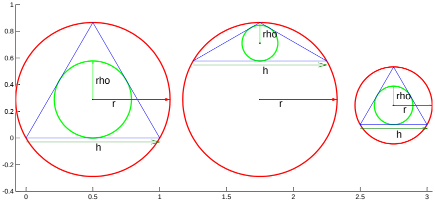


What is a good mesh ? (Shewchuk, 2002)

In the finite element method, the error is bounded by the mesh parameters :

$$\|u-u_h\|_{H^1(\Omega)} \leqslant C_\alpha h^k \|u\|_{H^{k+1}(\Omega)}$$

$$\forall \tau \quad \frac{h(\tau)}{\rho(\tau)} \leqslant \alpha$$

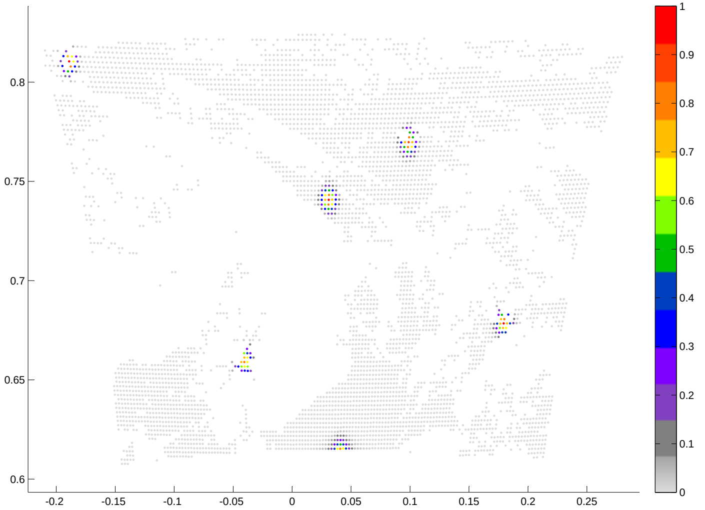


Impact of the quality of the mesh :

Large angles deteriorate the conditioning of the matrices due to **inaccurate representations of the gradients**.

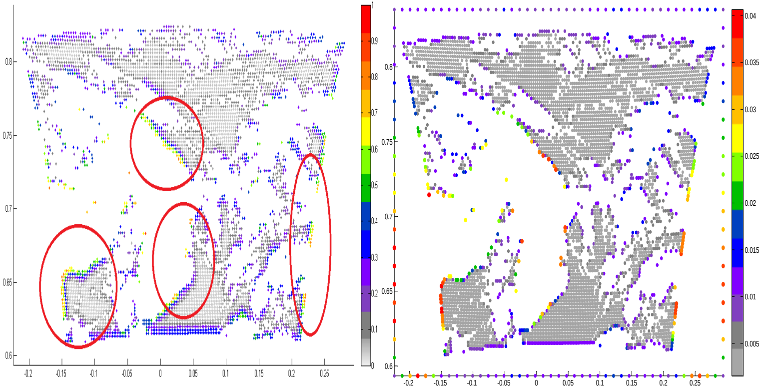
The mesh is also globally responsible for the stability of the computations.

Modelled correlation functions



Unitary response to several Dirac impulses.

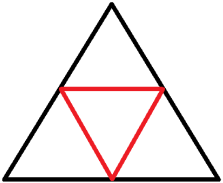
Limits of the approach



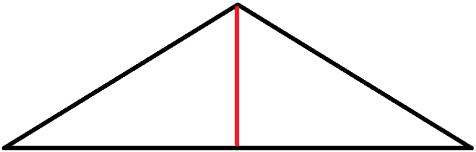
Left : Amplitude error at each location. Right : circumradius index.

Refinement strategies

- Recall that the FEM quality depends on the ratio $\frac{h(\tau)}{\rho(\tau)}$.
- This ratio can be improved through **local mesh refinement**.
- Several strategies are possible (reduction, bisection).



reduction



bisection

Summary

- $\mathbf{R}$ is not stored explicitly. Instead, it is represented as an operator.
- It is built from the implicit diffusion equation and a finite element discretization to allow treatments on unstructured grids.
- $\mathbf{R}^{-1}$ is easy to access.
- The amplitude of the correlation functions is equal to 1 almost everywhere.
- When the elements are ill-shaped, refining the mesh improves the accuracy of $\mathbf{R}$, at the expense of complicating $\mathbf{R}^{-1}$.
- Guillet et at. (2018) to be submitted to QJRMS.

Link with Brankart's method

- Consider the generalized diffusion equation :

$$(\alpha_0 I - \alpha_1 \Delta^1 + \ldots + (-1)^m \alpha_m \Delta^m) u^m = u^0$$

- Its finite element discretization leads to (normalization factors are omitted) :

$$\mathbf{C}^{-1} = \mathbf{M}[\alpha_0 \mathbf{I} + \alpha_1 \mathbf{M}^{-1}\mathbf{K} + \ldots + \alpha_m (\mathbf{M}^{-1}\mathbf{K})^m]$$

- For instance, take $m = 2$:

$$\mathbf{C}^{-1} = \underbrace{\begin{bmatrix} \mathbf{M}^{\frac{1}{2}} & \mathbf{K}^{\frac{1}{2}} & \mathbf{K}\mathbf{M}^{-\frac{1}{2}} \end{bmatrix}}_{\mathbf{T}^*} \underbrace{\begin{bmatrix} \alpha_0 \mathbf{I} & & \\ & \alpha_1 \mathbf{I} & \\ & & \alpha_2 \mathbf{I} \end{bmatrix}}_{(\mathbf{R}^+)^{-1}} \underbrace{\begin{bmatrix} \mathbf{M}^{\frac{1}{2}} \\ \mathbf{K}^{\frac{1}{2}} \\ \mathbf{M}^{-\frac{1}{2}}\mathbf{K} \end{bmatrix}}_{\mathbf{T}} \begin{matrix} 0^{th} \text{ order derivative} \\ 1^{st} \text{ order derivative} \\ 2^{th} \text{ order derivative} \end{matrix}$$

- Here, it can now adapt to any kind of spatial distribution of the data.