Issues specific to data sparse systems

Marc Bocquet

CEREA, joint lab École des Ponts ParisTech and EdF R&D, Université Paris-Est, France
Institut Pierre-Simon Laplace

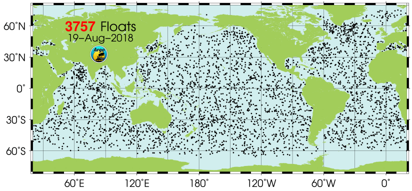
(marc.bocquet@enpc.fr)

With contributions from: L. Wu, M. R. Koohkan, T. Lauvaux, F. Chevallier, M. Krysta
and regular discussions over these topics with N. Bousserez

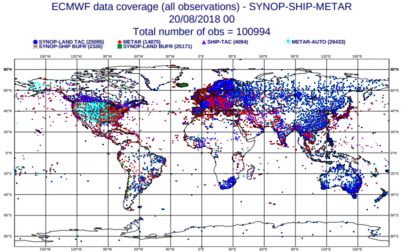
Outline

In situ, surface-based observations

- In situ, surface-based observations usually come from sparse monitoring networks.
- They are of high value because of their accuracy, their frequency, and the direct access to the instruments.



Ocean: Argo floats



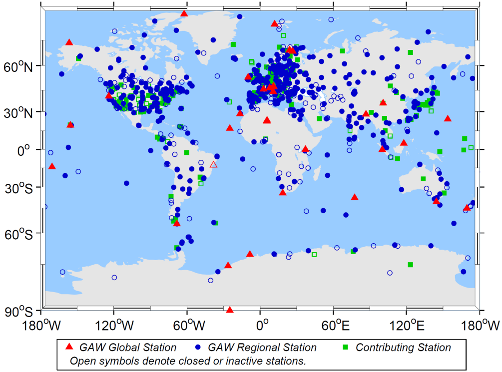
Atmosphere: conventional observations

- But afflicted by representation errors. Their measurements are not faulty, only our inability to simulate them. All observations are impacted by representation errors to some extend.

Janjić et al., 2018

In situ, surface-based observations

▶ The in-situ observations are still critical in atmospheric chemistry, especially air quality (boundary layer chemistry), in oceanography, etc, in meteorological reanalysis, boundary layer meteorology and micro-meteorology.



Gas: Global Atmosphere Watch network

Lidar: Earlinet network

▶ Intermediate instruments such as radar/lidar can form sparse networks too.

In situ, surface-based observations

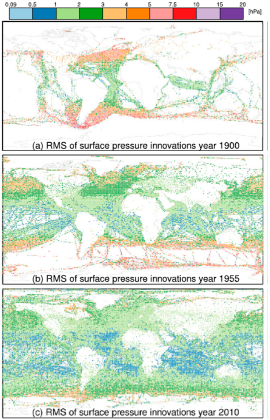
▶ Because of the contrasted scales of in-situ observations and models, sparsity calls for multiscale modelling and data assimilation.

▶ Sparsity of observations affects the balance of background statistics in a cycled data assimilation $\longrightarrow$ strong impact on most reanalysis endeavours, where the observation network can considerably evolve.

▶ Calls for:

- Fill-in data using geostatistics?
- Ensemble-based flow-dependent background error covariances (EnKF),
- Adjustable (ideally adaptive) background error covariances (EDA-based; diagnostics),
- Spatially adaptive inflation schemes: e.g. avoid inflating in data sparse regions,
- The problem gets tougher with coupled models with heterogeneous observation networks.

ERA-20C innovations RMS $\longrightarrow$



Karspeck et al., 2012; Whitaker et al., 2004; Poli et al., 2016; Laloyaux et al., 2018

Outline

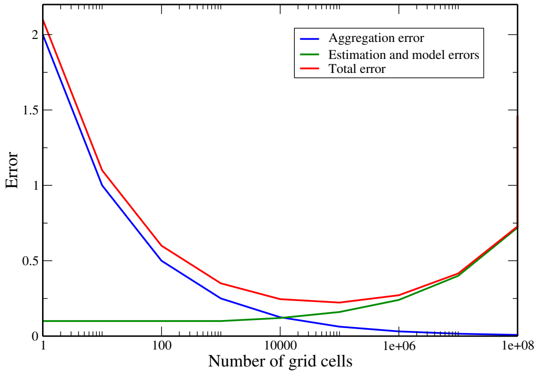
Dimensional reduction

- For most geophysical data assimilation / inverse problems, only a small fraction of the control space is actually informed by the observations.
- This could be for instance due to
  - the chaotic dynamics: it could be sufficient to control the unstable subspace and a few extra modes,
  - or to sparse observations (observation driven reduction?)
- Dimension reduction allows
  - faster computation of the solution and its uncertainty,
  - the use of sophisticated inference methods (non-linear sampling such as MCMC),
  - identification of surrogate models (Polynomial Chaos, machine learning techniques),

  and naturally applies to multiscale DA systems.
- How does the dimensional reduction impact the accuracy of the solution? Is there an optimal resolution?

Dimensional reduction

- There may be a competition between:
  - The aggregation errors
  - Errors that are due to scale-dependent modelling errors.



- Paradigm discussed by the greenhouse house gases inverse modelling community!

Peylin et al., 2001

Inverse modelling context

- Context: Inverse modelling of sources $\boldsymbol{\sigma}$ in atmospheric chemistry
  - Background $\mathbf{B}$ in control space. First guess $\boldsymbol{\sigma}_{\mathrm{b}}$.
  - $\mathbf{R}$ a priori on the observation/model errors
  - $\mathbf{H}$ Jacobian matrix of the problem (observation + model):

$$\boldsymbol{\mu} = \mathbf{H}\boldsymbol{\sigma} + \epsilon\,. \tag{1}$$

- BLUE analysis:

$$\begin{aligned} \boldsymbol{\sigma}_{\mathrm{a}} &= \boldsymbol{\sigma}_{\mathrm{b}} + \mathbf{B}\mathbf{H}^{\mathrm{T}}\left(\mathbf{R} + \mathbf{H}\mathbf{B}\mathbf{H}^{\mathrm{T}}\right)^{-1}\left(\boldsymbol{\mu} - \mathbf{H}\boldsymbol{\sigma}_{\mathrm{b}}\right)\,, \\ \mathbf{P}^{\mathrm{a}} &= \mathbf{B} - \mathbf{B}\mathbf{H}^{\mathrm{T}}\left(\mathbf{R} + \mathbf{H}\mathbf{B}\mathbf{H}^{\mathrm{T}}\right)^{-1}\mathbf{H}\mathbf{B}\,. \end{aligned} \tag{2}$$

- In the following a representation $\omega \in \mathcal{R}(\Omega)$ is a discretisation of the space-time domain of control (parameter) space Ω.

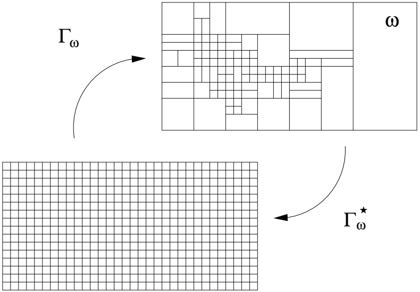
Up and down the scale ladder (1/4)

Restriction and prolongation

- Restriction operator : $\boldsymbol{\sigma} \underset{\text{coarse graining}}{\longrightarrow} \boldsymbol{\sigma}_\omega = \boldsymbol{\Gamma}_\omega \boldsymbol{\sigma}$, where $\boldsymbol{\Gamma}_\omega : \mathbb{R}^{N_{\rm fg}} \to \mathbb{R}^N$ defines the coarse graining operator (non-ambiguous).
- Prolongation operator : $\boldsymbol{\Gamma}^\star_\omega : \mathbb{R}^N \to \mathbb{R}^{N_{\rm fg}}$ refines $\boldsymbol{\sigma}_\omega$ into $\boldsymbol{\sigma}$ (ambiguous).

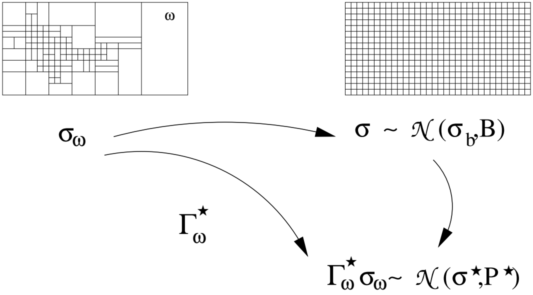
Scaling of errors

- Background error covariance matrix: $\mathbf{B}_\omega = \boldsymbol{\Gamma}_\omega \mathbf{B} \boldsymbol{\Gamma}^{\rm T}_\omega$,
- Observations/representativeness/model errors: $\mathbf{R}_\omega$, to be discussed later.



Bocquet et al., 2011

Up and down the scale ladder (2/4)



Bayesian choice of a prolongation operator

- Idea: Use prior $\boldsymbol{\sigma} \sim \mathcal{N}(\boldsymbol{\sigma}_\mathrm{b}, \mathbf{B})$ to refine the source. Knowing $\boldsymbol{\sigma}_\omega$ in representation ω, then from Bayes' rule, the most likely refined source is given by the mode of

$$q(\boldsymbol{\sigma}|\boldsymbol{\sigma}_\omega) = \frac{q(\boldsymbol{\sigma})}{q_\omega(\boldsymbol{\sigma}_\omega)} \delta\left(\boldsymbol{\sigma}_\omega - \boldsymbol{\Gamma}_\omega \boldsymbol{\sigma}\right) , \tag{3}$$

Bocquet et al., 2011

Up and down the scale ladder (3/4)

Bayesian choice of a prolongation operator

- Refinement is now a statistical process! But the prolongation operator will be defined as the most likely refinement operation.
- Thus the (estimate of the) refined source is

$$\boldsymbol{\sigma}^\star = \boldsymbol{\sigma}_\mathrm{b} + \mathbf{B}\boldsymbol{\Gamma}_\omega^\mathrm{T}\left(\boldsymbol{\Gamma}_\omega\mathbf{B}\boldsymbol{\Gamma}_\omega^\mathrm{T}\right)^{-1}\left(\boldsymbol{\sigma}_\omega - \boldsymbol{\Gamma}_\omega\boldsymbol{\sigma}_\mathrm{b}\right) , \tag{4}$$

which suggests the (affine) prolongation operator

$$\boldsymbol{\Gamma}_\omega^\star \equiv (\mathbf{I}_{N_\mathrm{fg}} - \boldsymbol{\Pi}_\omega)\boldsymbol{\sigma}_\mathrm{b} + \boldsymbol{\Lambda}_\omega^\star , \tag{5}$$

where the linear part of $\boldsymbol{\Gamma}_\omega^\star$ is

$$\boldsymbol{\Lambda}_\omega^\star \equiv \mathbf{B}\boldsymbol{\Gamma}_\omega^\mathrm{T}\left(\boldsymbol{\Gamma}_\omega\mathbf{B}\boldsymbol{\Gamma}_\omega^\mathrm{T}\right)^{-1} , \quad \text{and} \quad \boldsymbol{\Pi}_\omega \equiv \boldsymbol{\Lambda}_\omega^\star\boldsymbol{\Gamma}_\omega . \tag{6}$$

Bocquet et al., 2011

Up and down the scale ladder (4/4)

Up and down

- Must consistently satisfy $\mathbf{\Gamma}_\omega \mathbf{\Gamma}^\star_\omega = \mathbf{I}_N$.
- Down and up: $\mathbf{\Gamma}^\star_\omega \mathbf{\Gamma}_\omega = (\mathbf{I}_{N_{\mathrm{fg}}} - \mathbf{\Pi}_\omega)\boldsymbol{\sigma}_\mathrm{b} + \mathbf{\Pi}_\omega$

Properties of $\mathbf{\Pi}_\omega$

- $\mathbf{\Pi}_\omega$ is a projector since $\mathbf{\Pi}^2_\omega = \mathbf{\Pi}_\omega$.
- It is also $\mathbf{B}^{-1}$-symmetric: $\mathbf{\Pi}_\omega \mathbf{B} = \mathbf{B}\mathbf{\Pi}^\mathrm{T}_\omega$.

Observation equation in representation ω

- Then $\mathbf{H}$ becomes $\mathbf{H}_\omega = \mathbf{H}\mathbf{\Gamma}^\star_\omega$, and

$$\boldsymbol{\mu} = \mathbf{H}_\omega \boldsymbol{\sigma}_\omega + \epsilon_\omega = \mathbf{H}\mathbf{\Gamma}^\star_\omega \mathbf{\Gamma}_\omega \boldsymbol{\sigma} + \epsilon_\omega\ , \tag{7}$$

so that

$$\boldsymbol{\mu} = \mathbf{H}\boldsymbol{\sigma}_\mathrm{b} + \mathbf{H}\mathbf{\Pi}_\omega(\boldsymbol{\sigma} - \boldsymbol{\sigma}_\mathrm{b}) + \epsilon_\omega\ . \tag{8}$$

Bocquet et al., 2011

Accounting for aggregation/representation errors

Consistent observation equations:

$$\boldsymbol{\mu} = \mathbf{H}\boldsymbol{\sigma} + \epsilon = \mathbf{H}_\omega \boldsymbol{\sigma}_\omega + \epsilon_\omega \, . \tag{9}$$

Assuming aggregation is the only source of scale-dependent errors, one has $\mathbf{H}\boldsymbol{\sigma} + \epsilon = \mathbf{H}\boldsymbol{\sigma}_\mathrm{b} + \mathbf{H}\boldsymbol{\Pi}_\omega(\boldsymbol{\sigma} - \boldsymbol{\sigma}_\mathrm{b}) + \epsilon_\omega$, leading to the identification

$$\epsilon_\omega = \epsilon + \mathbf{H}\left(\mathbf{I}_{N_\mathrm{fg}} - \boldsymbol{\Pi}_\omega\right)(\boldsymbol{\sigma} - \boldsymbol{\sigma}_\mathrm{b}) \, . \tag{10}$$

Assuming independence of the error and source priors, the computation of the covariance matrix of these errors leads to

$$\begin{aligned} \mathbf{R}_\omega &= \mathbf{R} + \mathbf{H}\left(\mathbf{I}_{N_\mathrm{fg}} - \boldsymbol{\Pi}_\omega\right)\mathbf{B}\left(\mathbf{I}_{N_\mathrm{fg}} - \boldsymbol{\Pi}_\omega\right)\mathbf{H}^\mathrm{T} & (11) \\ &= \mathbf{R} + \mathbf{H}\left(\mathbf{I}_{N_\mathrm{fg}} - \boldsymbol{\Pi}_\omega\right)\mathbf{B}\mathbf{H}^\mathrm{T} \, . & (12) \end{aligned}$$

In that case, one checks that the innovation statistics are scale-independent.

Rodgers, 2000; Bocquet et al., 2011

The DFS criterion for the optimality of representations

- Idea: maximise the number of degrees of freedom in the signal (DFS), that come from the observations and is transferred to control space:

$$\begin{aligned}\mathcal{J} &= \mathrm{Tr}\left(\mathbf{I}_{\mathrm{N}} - \mathbf{P}^{\mathrm{a}}\mathbf{B}^{-1}\right) = \mathrm{Tr}\left(\mathbf{HK}\right) \\ &= \mathrm{Tr}\left(\mathbf{HBH}^{\mathrm{T}}\left(\mathbf{R} + \mathbf{HBH}^{\mathrm{T}}\right)^{-1}\right). \end{aligned} \tag{13}$$

  It also maximises the potential of BLUE.

- Bounded by the number of available observations: $0 \leq \mathcal{J} \leq d$.
  For a perfect observation/perfect model experiment: $\max_\omega \mathcal{J} = d$.
  Limited by the errors diagnosed in the observations $\max_\omega \mathcal{J} \leq d$.

- It also reads, for any representation ω:

$$\mathcal{J}_\omega = \mathrm{Tr}\left(\mathbf{H}_\omega\mathbf{B}_\omega\mathbf{H}_\omega^{\mathrm{T}}\left(\mathbf{R}_\omega + \mathbf{H}_\omega\mathbf{B}_\omega\mathbf{H}_\omega^{\mathrm{T}}\right)^{-1}\right). \tag{14}$$

Bocquet, 2009; Bocquet et al., 2011

Accounting for full scale-dependent errors

▶ Decomposition of errors from the scale analysis point of view:

- the scale-independent observation error ϵ^{o},
- an aggregation error: $\epsilon_\omega \equiv \epsilon + \epsilon^{\mathrm{c}}_\omega$, where $\epsilon^{c}_\omega = \mathbf{H}\left(\mathbf{I}_{N_{\mathrm{fg}}} - \mathbf{\Pi}_\omega\right)(\boldsymbol{\sigma} - \boldsymbol{\sigma}_{\mathrm{b}})$,
- the model error that would be scale-dependent ϵ^{m}_ω.

As a result:

$$\epsilon_\omega = \epsilon^{\mathrm{o}} + \epsilon^{\mathrm{c}}_\omega + \epsilon^{\mathrm{m}}_\omega \,. \tag{15}$$

The criterion for the design of representations may then be non-monotonic.

▶ Criteria under scale-covariant errors (i.e. without model error except representation errors)

$$\mathcal{J}_\omega = \mathrm{Tr}\left(\mathbf{\Pi}_\omega \mathbf{HBH}^{\mathrm{T}}\left(\mathbf{R} + \mathbf{HBH}^{\mathrm{T}}\right)^{-1}\right)\,. \tag{16}$$

Bocquet et al., 2011

Criteria under scale-covariant errors

▶ When considering scale-covariant errors, the dependence of the criteria in the representation ω can be simplified

- Fisher:

$$\mathcal{J}_\omega = \mathrm{Tr}\left(\mathbf{\Pi}_\omega \mathbf{B}\mathbf{H}^\mathrm{T}\mathbf{R}^{-1}\mathbf{H}\right) . \tag{17}$$

- DFS:

$$\mathcal{J}_\omega = \mathrm{Tr}\left(\mathbf{\Pi}_\omega \mathbf{H}\mathbf{B}\mathbf{H}^\mathrm{T}\left(\mathbf{R}+\mathbf{H}\mathbf{B}\mathbf{H}^\mathrm{T}\right)^{-1}\right) . \tag{18}$$

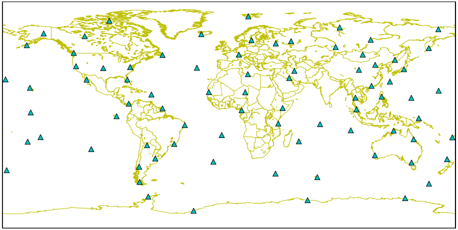
- Data-dependent:

$$\begin{aligned}\mathcal{J}_\omega =& \mathrm{Tr}\left(\mathbf{\Pi}_\omega \mathbf{B}\mathbf{H}^\mathrm{T}(\mathbf{R}+\mathbf{H}\mathbf{B}\mathbf{H}^\mathrm{T})^{-1}(\boldsymbol{\mu}-\boldsymbol{\mu}_\mathrm{b})\right.\\ &\left.\times(\boldsymbol{\mu}-\boldsymbol{\mu}_\mathrm{b})^\mathrm{T}(\mathbf{R}+\mathbf{H}\mathbf{B}\mathbf{H}^\mathrm{T})^{-1}\mathbf{H}\right) .\end{aligned} \tag{19}$$

▶ These objective functions can be proven to be increasing functions of the number of grid cells!

Bocquet et al., 2011

CTBTO IMS radionuclide network



▶ Objective: 80 radionuclide particle filters worldwide. 79 stations with designated location (treaty).

▶ Design network study Performance of the network assessed with detectability criteria.

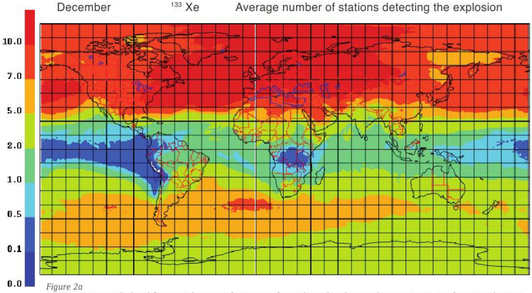
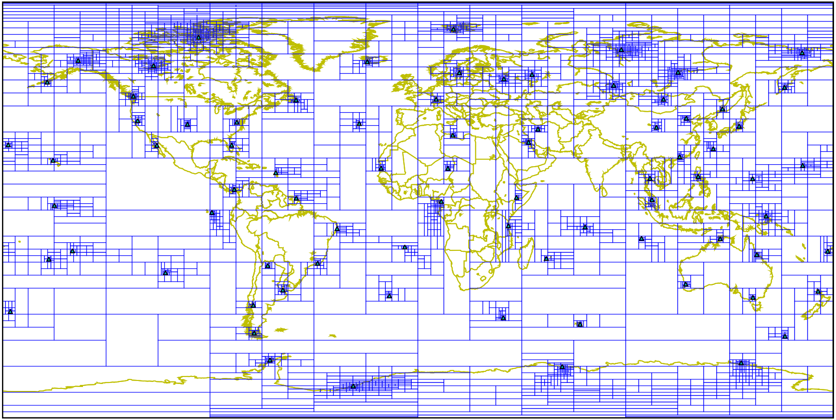


Figure 2a
Xenon coverage calculated for December using four years of actual weather data, real station sensitivity figures and xenon background information. In this graph, red represents the detection by many stations and blue depicts detection by a few stations of a one kt underground test with 10 percent leakage. The colours represent the average number of stations detecting the explosion.

Koohkan et al., 2012; Ringbom & Milley, 2009

Optimal adaptive grid for the IMS radionuclide network (1/3)

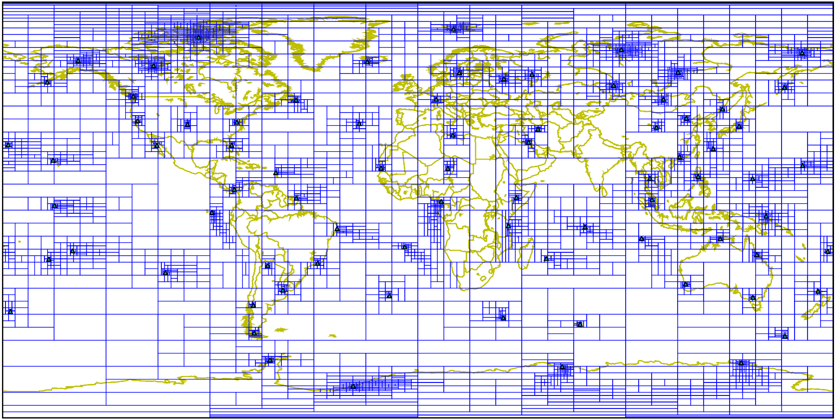
- Large error case. Equivalent to dfs/number of observations $\longrightarrow 0$. $N = 4096$.



Koohkan et al., 2012

Optimal adaptive grid for the IMS radionuclide network (2/3)

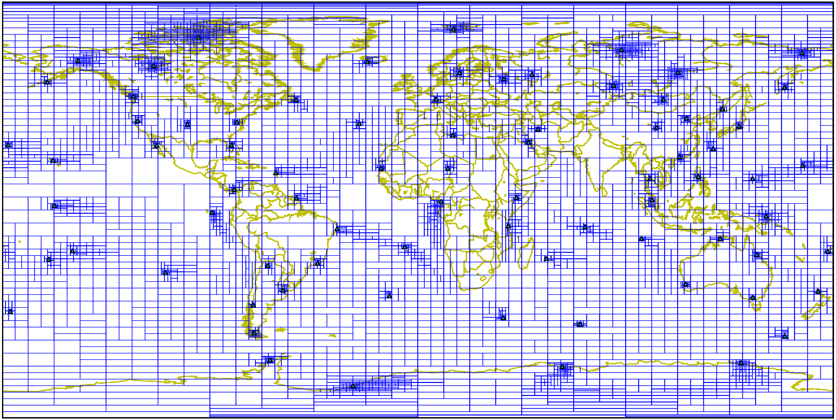
- Realistic case (optimistic): dfs/number of observations $\simeq 10\%$. $N = 4096$.



Koohkan et al., 2012

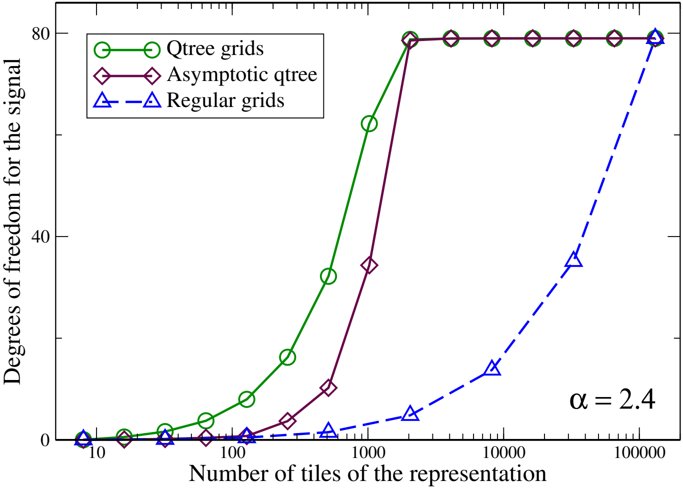
Optimal adaptive grid for the IMS radionuclide network (3/3)

▶ Small error case: dfs/number of observations $\simeq 90\%$. $N = 4096$.
Performance of distant future modelling and data assimilation systems.



Koohkan et al., 2012

Comparison of the asymptotic and exact designs performances



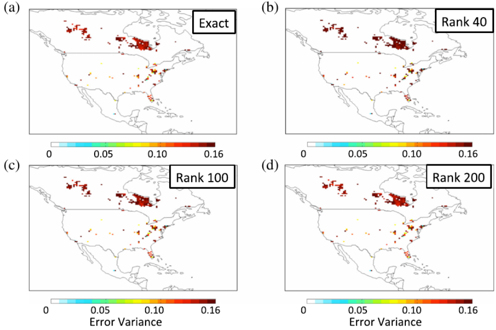
Optimised on the qtrees set, with the DFS criterion, on the CTBTO test case.

Bocquet & Wu, 2011

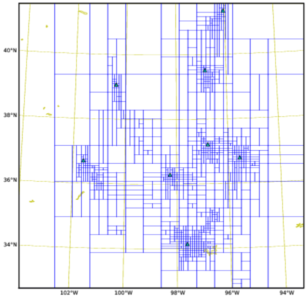
More general reduction techniques

▶ Instead of defining optimal adaptive control space grids, one can look for general low-rank representation and analyses within a truncated basis of the dominant modes.

▶ I invite you to read the recent and complete review on the topic by N. Bousserez and D. Henze, with further applications to greenhouse gas inverse problems.



Methane fluxes posterior variance



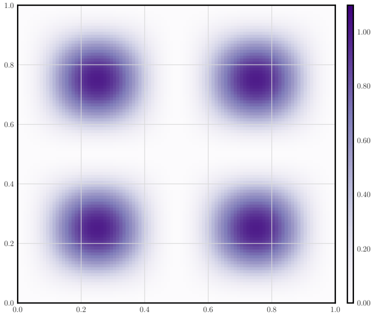
Reduced opt. grid for regional CO_2 inversion.

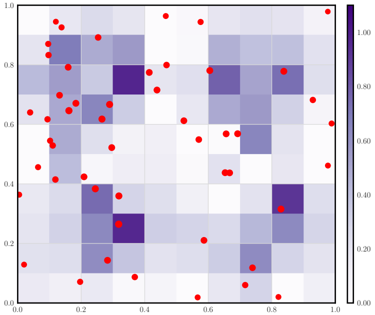
Bocquet & Wu, 2011; Spantini et al., 2015; Bousserez & Henze, 2018

Outline

Inverse problem in atmospheric chemistry with sparse networks

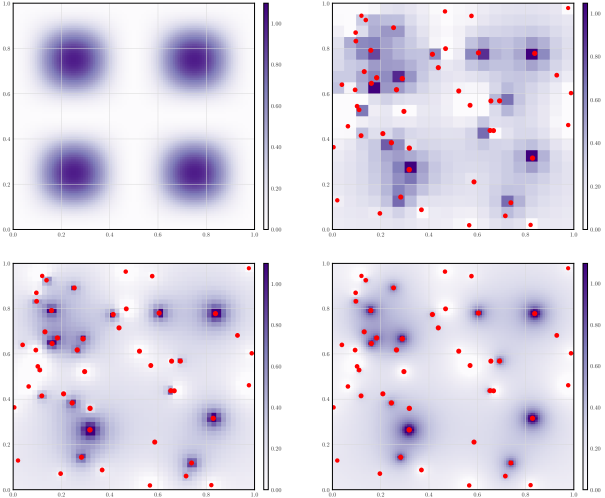
- Inverse modelling of emissions of trace gas (greenhouse gases, VOCs, gaseous reactive species, particulate matter, radionuclides, etc.)
- Sparse surface-based network remains key for the inversion of emissions/fluxes/source.
- Can be solved with traditional techniques of data assimilation/inverse problems assuming a fixed discretisation resolution of the model and of the emission prior and statistics.





The colocation issue

- Odd dependence of the solution on the resolution (colocation).



Issartel, 2003; Bocquet, 2005; Saide et al., 2011

Singular continuous limit of a DA system

- Assimilation of one concentration observation μ to retrieve an emission field σ

$$\mathcal{J}(\sigma) = \frac{1}{r^2}\left(\mu - \int \mathrm{d}x\, h(x)\sigma(x)\right)^2 + \frac{1}{m^2}\int \mathrm{d}x\, \sigma^2(x)$$

h is the forward model adjoint solution attached to μ. Thikonov regularisation.

- The solution is:

$$\sigma(x) = \frac{m^2 h(x)}{r^2 + m^2 \int \mathrm{d}x\, h^2(x)}\mu$$

- The physical model must have a proper continuum limit.

In particular $\int \mathrm{d}x\, h(x)\sigma(x)$ and $\int \mathrm{d}x\, h(x)$ are proper.

- However, the data assimilation system does not necessarily have one!

In particular $\int \mathrm{d}x\, h^2(x)$ is not necessarily proper. Its Riemann discretisation might diverge as $\Delta_x \longrightarrow 0$. In that case, when $\Delta_x \longrightarrow 0$, one has:

- $\sigma(x) \longrightarrow 0$, except maybe at the stations,
- $\mathrm{dfs} \propto \left[1 + \frac{r^2}{m^2}\left(\int \mathrm{d}x\, h^2(x)\right)^{-1}\right]^{-1} \longrightarrow 1.$

Bocquet, 2005

Singular continuous limit of a DA system

- The divergence depends on many critical factors:
  - the geometry of control space (where the background statistics are defined),
  - the geometry of the observation space,
  - the nature of the physics (advection, diffusion, convection, etc.) and the geometry of the physical space.

- Example of a diffusion problem $\boldsymbol{\mu} = \mathbf{H}\boldsymbol{\sigma}$:

$$\boldsymbol{\sigma}^\star = \mathbf{B}\mathbf{H}^{\mathrm{T}}\mathbf{G}^{-1}\boldsymbol{\mu} \quad \text{with} \quad \mathbf{G} = \mathbf{H}\mathbf{B}\mathbf{H}^{\mathrm{T}} .$$

The information matrix $\mathbf{G}$ is the Grammian of the adjoint solutions. A diagonal entry of $\mathbf{G}$ has the form

$$g \sim \sum_k v_k [\mathbf{c}_i^*]_k [\mathbf{c}_i^*]_k \tag{20}$$

We assume $\mathbf{c}_i^*$ has a diffusive behaviour close to the observation network:

$$c^*(\mathbf{r}, \mathbf{z}, t) = \frac{\exp\left\{-\frac{1}{t}\left(\frac{|\mathbf{r}|^2}{4K_h} + \frac{|\mathbf{z}|^2}{4K_z}\right)\right\}}{\sqrt{(4\pi t)^D K_h^d K_z^{D-d}}} . \tag{21}$$

Bocquet, 2005

Singular continuous limit of a DA system

- Geometry:
  - D: the dimension of the physical space in which diffusion takes place,
  - d: the dimension of control (source) space,
  - δ: the dimension of the observation embedding space.
- Asymptotically (taking first the limits Δ_t and Δ_z go to 0)

$$g \sim (D-2)!\frac{S_d}{2\pi^D}\frac{1}{K_h}\left(\frac{K_h}{K_z}\right)^{D-d}\frac{\Delta_x^{d-2D+2}}{2D-d-2}, \tag{22}$$

where $S_d = 2\pi^{d/2}/\Gamma(d/2)$ is the area of the unit sphere in dimension d.

- Two regimes:
  - If $d - 2D + 2 > 0$, g diverges when Δ_x goes to 0. Then there is degeneracy.
  - If $d - 2D + 2 \leq 0$, g converges. No degeneracy. Higher resolution is not detrimental to the inverse problem.

Bocquet, 2005

Singular continuous limit of a DA system

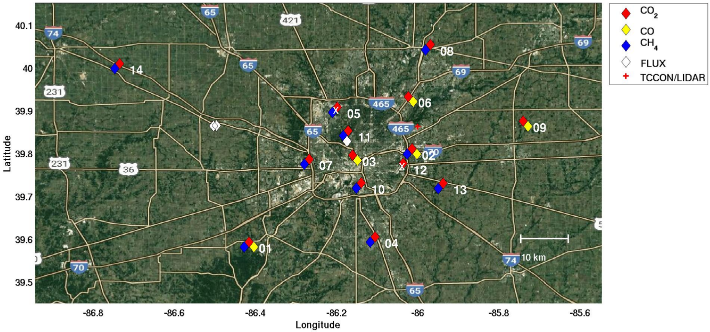
D	d	δ	divergence	$\\|\boldsymbol{\sigma}^\star\\|^2_{\mathbf{B}^{-1}}/\\|\boldsymbol{\sigma}\\|^2_{\mathbf{B}^{-1}}$	context
3	2	2	$g \sim \frac{1}{2\pi^2 K_z}\Delta_x^{-2}$	$\\|\boldsymbol{\mu}\\|^2\Delta_x^2$	surface obs. and source
2	2	2	$g \sim -\frac{1}{\pi K_h}\ln\Delta_x$	$-\frac{\\|\boldsymbol{\mu}\\|^2}{\ln\Delta_x}$	surface obs., source and transport
3	3	3	$g \sim \frac{2}{\pi^2 K_h}\Delta_x^{-1}$	$\\|\boldsymbol{\mu}\\|^2\Delta_x$	air obs., vol. source
3	2	3	no singularity	constant	air obs., surface source
3	0	2	no singularity	constant	surface obs., pointwise source

- Key findings :
  - For a 3D dispersion (diffusive-like close to the stations), $\|\boldsymbol{\sigma}^\star\|^2_{\mathbf{B}^{-1}}/\|\boldsymbol{\sigma}\|^2_{\mathbf{B}^{-1}}$ behaves like Δ_x^{4-d}. Always singular DA system.
  - Radiance height-resolved products and lidar observations used for inversion of emission on the ground does lead to the proper data assimilation system. No fundamental constraint on the spatial resolution.
- What is wrong with our setting of the inverse problem?
  - White noise has improper power spectrum; Tikhonov regularisation is unphysical!
  - But is coloured noise appropriate? Is there an intrinsic correlation length?

Bocquet, 2005

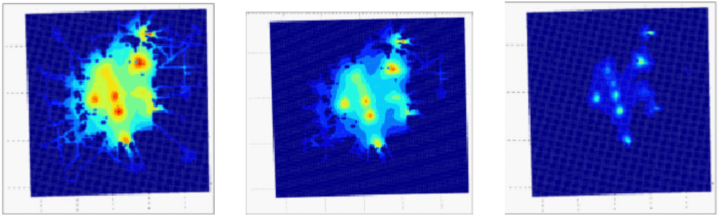
The INFLUX inverse problem

▶ INFLUX experiments in Indianapolis: inverse estimation of CO_2, CH_4 and CO urban fluxes using 12 towers, 5 NOA flasks samplers, 3 eddy flux towers, 1 Doppler lidar.

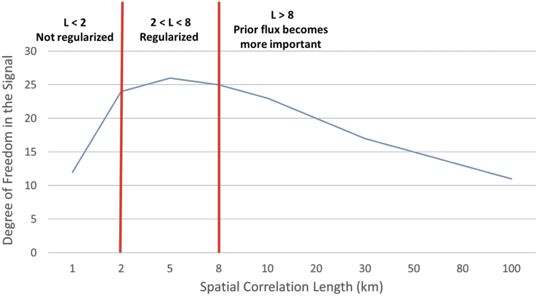


Lauvaux et al., 2016

The INFLUX inverse problem



- Error reduction in CO_2 emissions for a correlation length in **B** of 8, 5 and 2 kms.

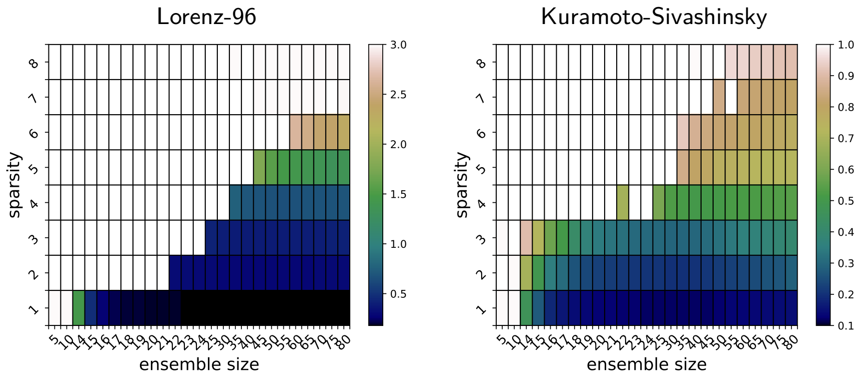


Wu et al., 2018

Outline

Required EnKF ensemble size with sparse observations

▶ RMSE of EnKF runs with sparser and sparser observations (without localisation, inflation optimally tuned)

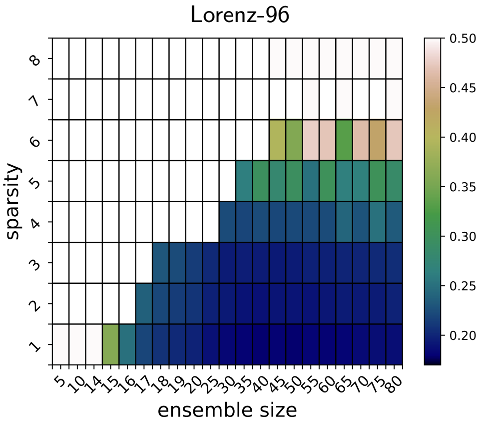


▶ *Catastrophic divergence* of the EnKF with a sparse network.

Bocquet et al., 2018

Required EnKF ensemble size with sparse observations

▶ RMSE of EnKF runs with sparser and sparser observations (without localisation, inflation optimally tuned) but with adjusted variance so as provide the same amount of information.



▶ *Catastrophic divergence* still present.

Catastrophic divergence

▶ Mechanism proposed by Gottwald and Majda (probably known by e.g., Anderson): Analysis at n for assimilating one observation y_m at site m:

$$x_n^{\mathrm{a}} = x_n^{\mathrm{b}} + P_{nm} \left(r + P_{mm} \right)^{-1} \left(y_m - x_n^{\mathrm{b}} \right) . \tag{23}$$

- P_{mm} fluctuates around the local variance as $\sim 1/N$.
- P_{nm} is of the order of $\sim 1/N$ whereas it should exponentially vanish with $d(n, m)$.
- As a consequence: spurious correlations yielding spurious update.
- Phenomenon amplified by a smaller r.

▶ Obviously localisation should be a remedy to catastrophic divergence

$$x_n^{\mathrm{a}} = x_n^{\mathrm{b}} + \rho_{nm} P_{nm} \left(r + P_{mm} \right)^{-1} \left(y_m - x_n^{\mathrm{b}} \right) . \tag{24}$$

But imbalance may be exacerbated by the lack of observations.

▶ Hybridisation of covariances is another option:

$$x_n^{\mathrm{a}} = x_n^{\mathrm{b}} + \left(\alpha P_{nm} + (1-\alpha) B_{nm} \right) \left(r + \alpha P_{mm} + (1-\alpha) B_{mm} \right)^{-1} \left(y_m - x_n^{\mathrm{b}} \right) . \tag{25}$$

Gottwald & Majda, 2013; Penny, 2014

Outline

Conclusions

▶ Optimal reduction techniques: Adapt control space in order to capture most of the system DFS and minimise the representation error (multiscale data assimilation).

▶ Sparse data assimilation systems pose several mathematical challenges:

- Sparse observations in a density-changing cycled DA system requires reliable, adaptive, possibly flow-dependent, background error statics; possibly spatially adaptive localisation, inflation and hybridisation for EnKF/hybrid/EDA.
- Depending on the geometry, statistics and physics of the data assimilation system, the continuous limit of the system might not exist!
  Traditional regularisation is inefficient or questionable. Alternatives?
- The EnKF used with sparse observations may lead to catastrophic divergence (another resurgence of sampling errors); that could be cured by (adaptive) localisation and/or hybridisation.

References I

[1] M. Bocquet, *Towards optimal choices of control space representation for geophysical data assimilation*, Mon. Wea. Rev., 137 (2009), pp. 2331–2348.

[2] M. Bocquet and A. Carrassi, *Four-dimensional ensemble variational data assimilation and the unstable subspace*, Tellus A, 69 (2017), p. 1304504.

[3] M. Bocquet, K. S. Gurumoorthy, A. Apte, A. Carrassi, C. Grudzien, and C. K. R. T. Jones, *Degenerate Kalman filter error covariances and their convergence onto the unstable subspace*, SIAM/ASA J. Uncertainty Quantification, 5 (2017), pp. 304–333.

[4] M. Bocquet and L. Wu, *Bayesian design of control space for optimal assimilation of observations. II: Asymptotics solution*, Q. J. R. Meteorol. Soc., 137 (2011), pp. 1357–1368.

[5] M. Bocquet, L. Wu, and F. Chevallier, *Bayesian design of control space for optimal assimilation of observations. I: Consistent multiscale formalism*, Q. J. R. Meteorol. Soc., 137 (2011), pp. 1340–1356.

[6] N. Bousserez and D. K. Henze, *Optimal and scalable methods to approximate the solutions of large-scale Bayesian problems: Theory and application to atmospheric inversion and data assimilation*, Q. J. R. Meteorol. Soc., 144 (2018), pp. 365–390.

[7] G. A. Gottwald and A. J. Majda, *A mechanism for catastrophic filter divergence in data assimilation for sparse observation networks*, Nonlin. Processes Geophys., 20 (2013), pp. 705–712.

[8] J.-P. Issartel, *Rebuilding sources of linear tracers after atmospheric concentration measurements*, Atmos. Chem. Phys., 3 (2003), pp. 2111–2125.

[9] T. Janjić, N. Bormann, M. Bocquet, J. A. Carton, S. E. Cohn, S. L. Dance, S. N. Losa, N. K. Nichols, R. Potthast, J. A. Waller, and P. Weston, *On the representation error in data assimilation*, Q. J. R. Meteorol. Soc., 0 (2018), pp. 0–0.

[10] M. R. Koohkan, M. Bocquet, L. Wu, and M. Krysta, *Potential of the international monitoring system radionuclide network for inverse modelling*, Atmos. Env., 54 (2012), pp. 557–567.

[11] T. Lauvaux, N. L. Miles, A. Deng, S. J. Richardson, M. O. Cambaliza, K. J. Davis, B. Gaudet, K. R. Gurney, J. Huang, D. O'Keefe, Y. Song, A. Karion, T. Oda, R. Patarasuk, I. Razlivanov, D. Sarmiento, P. Shepson, C. Sweeney, J. Turnbull, and K. Wu, *High-resolution atmospheric inversion of urban CO2 emissions during the dormant season of the Indianapolis Flux Experiment (INFLUX)*, J. Geophys. Res., 121 (2016), pp. 5213–5236.

[12] S. G. Penny, *The hybrid local ensemble transform Kalman filter*, Mon. Wea. Rev., 142 (2014), pp. 2139–2149.

[13] P. Poli, H. Hersbach, D. P. Dee, P. Berrisford, A. J. Simmons, F. Vitart, P. Laloyaux, D. G. H. Tan, C. Peubey, J.-N. Thépaut, et al., *ERA-20C: An atmospheric reanalysis of the twentieth century*, J. Clim., 29 (2016), pp. 4083–4097.

References II

[14] C. D. Rodgers, *Inverse methods for atmospheric sounding*, vol. 2, World Scientific, Series on Atmospheric, Oceanic and Planetary Physics, 2000.

[15] P. Saide, M. Bocquet, A. Osses, and L. Gallardo, *Constraining surface emissions of air pollutants using inverse modeling: method intercomparison and a new two-step multiscale approach*, Tellus B, 63 (2011), pp. 360–370.

[16] A. Spantini, A. Solonen, T. Cui, J. Martin, L. Tenorio, and Y. Marzouk, *Optimal low-rank approximations of Bayesian linear inverse problems*, SIAM Journal on Scientific Computing, 37 (2015), pp. A2451–A2487.

[17] A. J. Turner and D. J. Jacob, *Balancing aggregation and smoothing errors in inverse models*, Atmos. Chem. Phys., 15 (2015), pp. 7039–7048.

[18] J. S. Whitaker, G. P. Compo, X. Wei, and T. M. Hamill, *Reanalysis without radiosondes using ensemble data assimilation*, Mon. Wea. Rev., 132 (2004), pp. 1190–1200.

[19] K. Wu, T. Lauvaux, K. J. Davis, A. Deng, I. Lopez Coto, K. R. Gurney, and R. Patarasuk, *Joint inverse estimation of fossil fuel and biogenic CO2 fluxes in an urban environment: An observing system simulation experiment to assess the impact of multiple uncertainties*, Elem Sci Anth., 6 (2018), p. 17.

[20] L. Wu, M. Bocquet, T. Lauvaux, F. Chevallier, P. Rayner, and K. Davis, *Optimal representation of source-sink fluxes for mesoscale carbon dioxide inversion with synthetic data*, J. Geophys. Res., 116 (2011), p. D21304.