Grid resolution dependence in the reconstruction of an atmospheric tracer source

Marc Bocquet

Nonlinear Processes in Geophysics (2005) 12: 219–234
SRef-ID: 1607-7946/npg/2005-12-219
European Geosciences Union

Nonlinear Processes in Geophysics

Grid resolution dependence in the reconstruction of an atmospheric tracer source

M. Bocquet

CEREA, Research and Teaching Center in Atmospheric Environment,
Joint laboratory, École Nationale des Ponts et Chaussées/EDF R&D, avenue Blaise Pascal, 77455 Champs sur Marne, France

Received: 6 September 2004 – Revised: 14 December 2004 – Accepted: 17 January 2005 – Published: 8 February 2005

Part of Special Issue "Quantifying predictability"

Abstract. Methods for reconstructing sources of inert atmospheric tracers from ground measurements are currently studied, tested, and even implemented (in accident-type radionuclide release backtracking, for retrieval of carbon fluxes). Often the retrieved source exhibits a very strong and unrealistic (therefore unwanted) influence by the observation sites. This problem is shown not to be an intrinsic flaw of the reconstruction methods but rather due to the specifics of the atmospheric dispersion of a tracer, to the location of the receptors and to the expected source location. It is increasingly pronounced as the grid resolution for the source is improved, and we show how this translates mathematically. We rely on the general framework of inversion methods based on the maximum entropy on the mean principle. Those methods are well suited for accident-type tracer release problems. The dependence of the reconstruction on grid resolution is investigated both analytically and numerically, in conjunction with the issue of receptor influence. Two examples of synthetic experiments are given. The first one is a one-dimensional toy model which quantitatively validates the approach. The second one is based on the European Tracer Experiment and agrees well with the results obtained here. Finally, a generalization of the formalism is proposed so as to study the performance of reconstructions when observation and possibly model errors are present.

1 Introduction

Reconstructing pollutants sources is of increasing importance. An apparently simple form of this problem is the retrieval of sources of a passive (as opposed to reactive) tracer dispersed throughout the atmosphere. Many efforts have been devoted to the retrieval of fluxes of diffuse trace gas, such as methane, CO, and CO_2, and other types of semi-persistent pollutants like mercury. A few methodological works have emphasized the ill-conditioned nature of the inversion and identified how the physics of atmospheric dispersion bears on the technical difficulties of that type of reconstruction. Many contributions on the topics can be found in "Inverse Methods in Global Biogeochemical Cycles", Geophysical Monograph 114, American Geophysical Union, and more recently (Enting, 2002). Another class of problems concerns the retrieval of the source of pollutants following an accidental release. Typical examples are the monitoring of nuclear tests fostered by the test ban treaty (Pudykiewicz, 1998; Hourdin and Issartel, 2000) and the re-analysis of civil nuclear plant accidents such as Chernobyl. Although this paper mainly focuses on an accident type release, methods, analysis, and conclusions exposed here are still valid for extended sources and possibly diffusive sources. For an accidental source, one expects from inverse techniques to yield one or several locations of the release, the released mass of pollutant, and hopefully a temporal profile for the release.

Inversion techniques aim at extracting information on the source from a set of concentration measurements. To achieve this some knowledge from an adequate dispersion model is required. We rely on an Eulerian dispersion model, although some of the conclusions drawn here can possibly be extended to Lagrangian backtracking (see for example Stohl, 1998). Assimilating the concentration measurements is usually done through Kalman filtering or variational techniques (four-dimensional variational assimilation, see for example Elbern et al., 2000, or some under-constrained variants of three-dimensional variational assimilation/optimal interpolation). Under some circumstances, because of the linearity of the system, there are simpler ways to perform the inversion. These rely on adjoint techniques. They have been adapted to atmospheric problems in Marchuk (1995); Pudykiewicz (1998); Issartel and Baverel (2003). A variant known as the "representer method" (Bennett, 2002, and references therein) is used in oceanography and not only focuses on the source retrieval but also on the reconstruction of the complete dispersion event (full data assimilation system).

Correspondence to: M. Bocquet
(bocquet@cerea.enpc.fr)

The adjoint techniques are merely tools for inversion methods. For a high resolution reconstruction, the inversion is a severely ill-posed problem and a regularization is needed. The choice of the regularization and its motivation specify the inversion method. One type of regularization is by orthogonal projection onto some proper vector space of source fields. The most natural space is the space generated by the retroplumes, a set of adjoint solutions, each one of them attached to an observation.

In Bocquet (2005a) it was shown that this regularization can be viewed as an element from a larger class of regularizations, all relying on the principle of "maximum entropy on the mean". Those are the regularizations that are guaranteed to introduce less spurious information into the inversion. Each of them makes use of a probability density function ν for the prior distribution describing the source. It contains the statistical information garnered on the source before any measurement is made. The orthogonal projection technique is then seen as a regularization by entropy provided the prior distribution one assumes is Gaussian. These methods are expected to be efficient when the first moment of ν is little known or close to a black object (meaning the source is sparse), i.e. when the first guess for the source is not to be trusted or irrelevant (Bocquet, 2005b). However higher moments are also used. It is very natural in this framework to introduce priors such as positiveness, boundedness, extensity, etc. This prior information may constrain the source significantly. When there is a reliable first guess for the source, the method works as well but is not believed to perform significantly better than a least square variational inversion, where first and second moments are already used.

The success of a reconstruction using real measurement data depends on both the inversion technique used as well as the numerical transport model used. The two problems are decorrelated. For a perfect model (matching reality), the efficiency is only subject to the inversion technique. Conversely a poor model makes any sophisticated inversion technique pointless on real data. In this paper, the inversion technique aspect of the reconstruction problem is the main concern. In other words, it is assumed a good numerical dispersion model is used.

In recent works, it has been observed that the reconstruction is often strongly influenced by the detectors. The retrieved source may peak at the observation sites. Moreover the total mass retrieved for the source tends to be concentrated near a few receptors. This has been reported in the context of Lagrangian backtracking (Ashbaugh et al., 1985; Stohl, 1998) but also within an Eulerian approach (Issartel, 2003; Bocquet, 2005b). Recently it has been claimed that this is an artefact of the inversion which a proper inversion technique should be able to smooth out or eradicate (Issartel, 2003, 2005). In this paper, we argue that this feature of the reconstruction is not an artefact. It is due to the nature of the dynamics of the atmospheric transport of tracers, to the prior expectation of the source location and to the relative positions of the observation sites. We show when this failure in the reconstruction is prone to happen and the optimal conditions to avoid it, rather than ignoring it.

In Sect. 2, a detailed introduction and a reminder on recent developments are provided and the problem is exposed. In Sect. 3, the singular nature of the inversion method is examined. In Sect. 4, a numerical example based on the European Tracer Experiment (ETEX) is detailed. In Sect. 5, the error committed in the inversion is estimated. Its dependence on the mesh step and the observations is studied. In Sect. 6, we show how previous results should be generalized when the measurements are noisy and the model is imperfect, as it would be the case in an operational context. Conclusions are then given.

2 Principles of inversion

2.1 The forward transport

We are interested in the dispersion of a pollutant of concentration c, at regional/continental scale, over a domain $\Omega = \mathcal{D} \times [0, \tau]$, where $\mathcal{D}$ is the spatial domain and $[0, \tau]$ the time interval. The atmospheric dispersion equation is

$$\frac{\partial c}{\partial t} + \mathrm{div}\,(\boldsymbol{u}c) - \mathrm{div}\,(\boldsymbol{K}\nabla c) = \sigma\,, \tag{1}$$

where $\boldsymbol{K}$ parameterizes eddy diffusion, and $\boldsymbol{u}(\boldsymbol{x}, t)$ is the wind field. At continental scale, $\boldsymbol{K}$ is a diagonal tensor with components K_x, K_y, which are of limited influence, and vertical diffusion K_z, which needs a proper parameterization (Louis, 1979). σ is the forcing field, the source we are seeking to estimate. In the rest of the paper, it is assumed $\sigma(\boldsymbol{x}, t) \geq 0$ (positiveness of the source). This is not however a sine qua non condition for most of the conclusions of this work to stand. The air density is taken as approximately constant and homogeneous. As a consequence the air conservation equation simplifies to

$$\mathrm{div}(\boldsymbol{u}) = 0\,. \tag{2}$$

A complete generalization to a nonuniform air density is nevertheless possible (Issartel and Baverel, 2003; Hourdin et al., 2005). For an accidental release, it is appropriate to choose the following boundary conditions

$$\forall\,(\boldsymbol{x}, t) \in \partial\Omega_+, \quad c(\boldsymbol{x}, t) = 0\,, \tag{3}$$

where $\partial\Omega_+$ is the part of the domain boundary corresponding to the incoming wind field. In this context the initial conditions are

$$\forall\,\boldsymbol{x}, \quad c(\boldsymbol{x}, 0) = 0\,. \tag{4}$$

Moreover we assume the diffusive flux $\boldsymbol{K}\nabla c$ to be null (or negligible) at the boundary. Though it prevents pollutant to diffuse out of the domain, it does not impede (dominant) advection processes.

The measurement of concentrations is formalized through the equation:

$$\mu_i = \int_\Omega \mathrm{d}t\mathrm{d}\boldsymbol{x}\,\pi_i(\boldsymbol{x}, t)c(\boldsymbol{x}, t) \tag{5}$$

where μ_i ($i = 1, \cdots, p$) is a concentration value, indexed by i. π_i is the sampling function. It is measured per unit of volume and per unit of time. It describes how the averaging process of the concentration measurement is done and encodes the spatial and temporal extent of the observation. In order for μ_i to stand for a representative of the local concentration,

$$\int_\Omega \mathrm{d}t\mathrm{d}\boldsymbol{x}\, \pi_i(\boldsymbol{x}, t) = 1 \tag{6}$$

should be enforced. The observation equation (Eq. 5) introduces the L^2-scalar product over the functional space $L^2(\Omega)$:

$$(\Psi, \Phi) = \int_\Omega \mathrm{d}t\mathrm{d}\boldsymbol{x}\, \Psi(\boldsymbol{x}, t)\Phi(\boldsymbol{x}, t)\,, \tag{7}$$

which will be useful later.

2.2 The adjoint transport

Let us introduce a test field ϕ, to be multiplied against the transport equation (Eq. 1) over the whole domain Ω:

$$\int_\Omega \mathrm{d}t\mathrm{d}\boldsymbol{x}\, \phi \left(\frac{\partial c}{\partial t} + \mathrm{div}\,(\boldsymbol{u}c) - \mathrm{div}\,(\boldsymbol{K}\nabla c) - \sigma \right) = 0\,. \tag{8}$$

Integrations by part as well as boundary and initial conditions on c lead to the budget equation:

$$\begin{aligned} 0 = & \int_\Omega \mathrm{d}t\mathrm{d}\boldsymbol{x}\, c \left(-\frac{\partial \phi}{\partial t} - \mathrm{div}\,(\boldsymbol{u}\phi) - \mathrm{div}\,(\boldsymbol{K}\nabla\phi) \right) \\ & - \int_\Omega \mathrm{d}t\mathrm{d}\boldsymbol{x}\, \phi\,\sigma + \int_{\mathcal{D}} \mathrm{d}\boldsymbol{x}\, c(\tau)\phi(\tau) \\ & - \int_{\partial\mathcal{D}\times[0,\tau]} \mathrm{d}t\mathrm{d}\boldsymbol{S}\cdot(c\boldsymbol{K}\nabla\phi) + \int_{\partial\Omega_-} \mathrm{d}t\mathrm{d}\boldsymbol{S}\cdot(\phi c\boldsymbol{u})\,. \end{aligned} \tag{9}$$

To benefit from the dual formalism, it is then possible to define the following adjoint solution. Consider a measurement μ_i where $i \in [1, p]$. Its sampling function is π_i. Inspired by Eq. (9), one can introduce the adjoint equation:

$$-\frac{\partial c_i^*}{\partial t} - \mathrm{div}\left(\boldsymbol{u}\, c_i^*\right) - \mathrm{div}\left(\boldsymbol{K}\nabla\left(c_i^*\right)\right) = \pi_i\,, \tag{10}$$

where c_i^* therefore stands as the adjoint solution. In order to remove most of the terms of Eq. (9), we choose

$$\forall\,\boldsymbol{x}, \quad c_i^*(\boldsymbol{x}, \tau) = 0\,, \quad \forall\,(\boldsymbol{x}, t) \in \partial\Omega_-, \quad c_i^*(\boldsymbol{x}, t) = 0\,. \tag{11}$$

$\partial\Omega_-$ denotes the part of the domain boundary corresponding to the outgoing wind field. In addition, the diffusive flux $\boldsymbol{K}\nabla c_i^* = 0$ is imposed null at the boundaries. This completely defines the adjoint solution, c_i^*, called a "retroplume" in this context. Eq. (9) now becomes

$$\mu_i = \int_\Omega \mathrm{d}t\mathrm{d}\boldsymbol{x}\, \sigma(\boldsymbol{x}, t)c_i^*(\boldsymbol{x}, t)\,. \tag{12}$$

Previous results can be transposed to a discrete numerical model, where the domain Ω is composed of cells Ω_k, with $k \in [1, N]$. The concentration field c is discretized into c_k. However the discrete adjoint solution $c_{i,k}^*$ is not necessarily the discretization of c_i^* but the adjoint solution of the discrete numerical model. Generally, they do not coincide. $c_{i,k}^*$ remains in units of inverse spatial volume.

2.3 Principle of the inversion

Adopting a continuous form for the subsequent derivations would not always be mathematically safe. The discrete numerical point of view is therefore preferred (integrals on the domain Ω are replaced with sums on the grid).

In a source reconstruction, the main goal is to find a proper $\boldsymbol{\sigma}$, a vector of cell values σ_k, which satisfies the measurement equations

$$\forall\, i \in [1, p]\,, \quad \mu_i = \sum_{k=1}^{N} c_{i,k}^*\sigma_k\,, \tag{13}$$

which are a discrete version of Eq. (5). For convenience the observation matrix $[\boldsymbol{H}]_{i,k} = c_{i,k}^*$ is introduced. It encapsulates the model. If $\boldsymbol{\mu}$ is the vector of measurements, Eq. (13) can be recast as $\boldsymbol{\mu} = \boldsymbol{H}\boldsymbol{\sigma}$. To simplify formulas and interpretations, the cell volume elements are incorporated into the definition of σ_k. Therefore σ_k is in units of tracer mass and represents the total mass emitted in cell k.

For N significantly larger than p and in the absence of strongly restraining hypotheses on $\boldsymbol{\sigma}$, this problem is ill-posed. The underdetermination could only be lifted with a proper regularization. From a Bayesian point of view, this regularization must be equivalent to taking into account prior information on the source.

In addition to the experimental source of information, the prior information may often be significant. In this context, a reconstruction with finer resolution than what the mere data would allow is achievable. This is what is used in Issartel (2003), although the importance of the prior is not recognized. A construct introduced in Bocquet (2005a,b) devoted to that type of inversion is summarized in the following section.

Lastly, because $\boldsymbol{H}$ describes dispersion, it is known to be ill-conditioned as a matrix, adding to the difficulty. Short distance fluctuations of the source are cut-off by $\boldsymbol{H}$ with an efficiency that depends on the dispersion conditions (Enting, 2002).

2.4 Maximum entropy on the mean

Just before the inversion, one knows the set of concentration measurements. One possibly retains prior information on the source, such as very general features like positiveness, extensity, and boundedness but also specific features such as location and time-dependent information. All these pieces of information can be incorporated into a prior probability density function (pdf) $\nu : \boldsymbol{\sigma} \longrightarrow \nu(\boldsymbol{\sigma})$. $\boldsymbol{\sigma}$ belongs to the set of all potential source configurations. The entropy on the mean regularization is the method that maximizes ignorance (quantified as the entropy functional) on the source except for the set of μ_i and the prior ν. Therefore it guarantees that no spurious or unwanted information comes into the inversion. Details of the construction and its applications to atmospheric dispersion can be found in Bocquet (2005a). Let us summarize the main results needed for entropy techniques.

One expects the inversion to yield a posterior pdf $p(\boldsymbol{\sigma})$. The entropy to be maximized is

$$S = -\sum_{\boldsymbol{\sigma}} p(\boldsymbol{\sigma}) \ln\left(\frac{p(\boldsymbol{\sigma})}{\nu(\boldsymbol{\sigma})}\right) \tag{14}$$

provided the measurement constraints are satisfied on the mean, i.e.

$$\boldsymbol{\mu} = \langle \boldsymbol{H}\boldsymbol{\sigma} \rangle \equiv \sum_{\boldsymbol{\sigma}} p(\boldsymbol{\sigma}) \boldsymbol{H}\boldsymbol{\sigma} \,. \tag{15}$$

The symbolic sum on $\boldsymbol{\sigma}$ stands for an integral with respect to the measure $\prod_{k=1}^{N} \mathrm{d}\sigma_k$. It ranges over a domain included in $\mathbb{R}^N$. $-S$ is usually called the "Kullback-Leibler divergence or information": $\mathcal{K}(p,\nu) = -S$, which is a measure of the discrepancy between pdfs. Hence, introducing p Lagrange multipliers β_i, the following functional must be optimized

$$\mathcal{L} = \mathcal{K}(p,\nu) + \boldsymbol{\beta}^T \left(\boldsymbol{\mu} - \sum_{\boldsymbol{\sigma}} p(\boldsymbol{\sigma}) \boldsymbol{H}\boldsymbol{\sigma}\right) . \tag{16}$$

It is called the "level-2 primal problem". The problem can now be reduced by duality: it can be shown that it is equivalent to "minimize" the "secondary entropy"

$$\Psi = \ln Z(\boldsymbol{\beta}) - \boldsymbol{\beta}^T \boldsymbol{\mu} \;\text{ with }\; Z(\boldsymbol{\beta}) = \sum_{\boldsymbol{\sigma}} \nu(\boldsymbol{\sigma}) \exp\left(\boldsymbol{\beta}^T \boldsymbol{H}\boldsymbol{\sigma}\right), \tag{17}$$

with $Z(\boldsymbol{\beta})$, the "partition function". An average source can then be deduced through the estimator

$$\overline{\boldsymbol{\sigma}} = \langle \boldsymbol{\sigma} \rangle \equiv \sum_{\boldsymbol{\sigma}} p(\boldsymbol{\sigma}) \boldsymbol{\sigma} \,. \tag{18}$$

This estimator satisfies the measurement constraints. Minimizing Ψ over the β_i is called the "dual problem" of the level-2 primal problem.

Furthermore, it can be shown that solving this problem is equivalent to minimizing a cost function (the so-called "level-1 primal problem") $\mathcal{J}(\boldsymbol{\sigma}, \nu)$ [1], that depends on the prior ν and which is a functional of $\boldsymbol{\sigma}$. Its maximum is reached at $\overline{\boldsymbol{\sigma}}$. In the following sections the explicit form of $\mathcal{J}$, which depends on ν, will be given each time it is needed.

2.4.1 Factorization

When the prior pdf ν can be factorized according to the partition of the domain Ω into the cells Ω_k, then the partition function can also be factorized. The secondary entropy can be simplified accordingly. In this case

$$Z(\boldsymbol{\beta}) = \prod_{k=1}^{N} Z_k(\boldsymbol{\beta}) \,, \tag{19}$$

where

$$Z_k(\boldsymbol{\beta}) = \sum_{\sigma_k} \nu_k(\sigma_k) \exp\left(\left[\boldsymbol{\beta}^T \boldsymbol{H}\right]_k \sigma_k\right) . \tag{20}$$

As a result

$$\Psi = \sum_{k=1}^{N} \ln Z_k(\boldsymbol{\beta}) - \boldsymbol{\beta}^T \boldsymbol{\mu} \,. \tag{21}$$

This property will be used in the following sections.

[1] see (Gzyl, 1995), and Bocquet, M., unpublished work.

2.4.2 Poisson prior

The Bernoulli prior was shown to perform well on ETEX-like sources, and better than the Poisson law (Bocquet, 2005b). However, here, we are looking for cases when the retrieval performance is between success and failure. A reasonably efficient and simple prior like Poisson is therefore expected to suit our purpose.

When one only knows that the forcing $\boldsymbol{\sigma}$ is positive, a simple though non-trivial prior law like Poisson is appropriate. It is parameterized by a local average value θ_k. The prior pdf in cell k is given by

$$\nu_k(x_k) = e^{-\theta_k} \frac{\theta_k^{x_k}}{x_k!} \,. \tag{22}$$

x_k is an integer and $\nu_k(x_k)$ is interpreted as the probability density that a mass $m\,x_k$ is emitted at grid cell k, where m denotes a reference mass or level-spacing. Hence $m\theta_k$ appears as a first guess for the released mass in cell k. Since those laws are independent from one cell to another, factorization of the partition function is in order. The local partition function reads for any cell k

$$Z_k(\boldsymbol{\beta}) = \exp\left\{-\theta_k + \theta_k \exp\left(m\left[\boldsymbol{\beta}^T \boldsymbol{H}\right]_k\right)\right\} , \tag{23}$$

so that the secondary entropy is

$$\Psi = \sum_{k=1}^{N} \theta_k \left\{\exp\left(m\left[\boldsymbol{\beta}^T \boldsymbol{H}\right]_k\right) - 1\right\} - \boldsymbol{\beta}^T \boldsymbol{\mu} \,. \tag{24}$$

The reconstructed source is then

$$\overline{\sigma}_k = m\,\theta_k \exp\left(m\left[\boldsymbol{\beta}^T \boldsymbol{H}\right]_k\right) . \tag{25}$$

The level-1 primal cost function, which is useful when numerically computing the entropy of a source field, is easily derived:

$$\mathcal{J} = \sum_{k=1}^{N} \frac{\sigma_k}{m} \ln \frac{\sigma_k}{m\theta_k} + \theta_k - \frac{\sigma_k}{m} \,, \tag{26}$$

provided $\boldsymbol{\mu} = \boldsymbol{H}\boldsymbol{\sigma}$. The Poisson law will serve as a basis for the numerical examples of Sect. 4.

2.4.3 Gaussian prior

Gaussian laws are analytically tractable and that is why they will be used here, in addition to Poisson laws. They also allow for a clear connection to traditional variational data assimilation. Nevertheless it may not be a good choice for a positive source without a trusted first guess.

The source elements σ_k are supposed to be correlated according to $\mathrm{E}_\sigma\left[\sigma_k \sigma_l\right] = [\boldsymbol{B}]_{kl}$ and they follow a multivariate normal law. It is assumed there is no obvious first guess for the accidental release. The prior pdf is then

$$\nu(\boldsymbol{\sigma}) = \frac{1}{\sqrt{(2\pi)^N \det \boldsymbol{B}}} \exp\left(-\frac{1}{2}\,\boldsymbol{\sigma}^T \boldsymbol{B}^{-1} \boldsymbol{\sigma}\right) . \tag{27}$$

From it, a secondary entropy (level-2 dual problem) for the Gaussian prior is derived:

$$\Psi = \frac{1}{2}\boldsymbol{\beta}^T \boldsymbol{H}\boldsymbol{B}\boldsymbol{H}^T\boldsymbol{\beta} - (\boldsymbol{\beta}, \boldsymbol{\mu})\,. \tag{28}$$

Note that it is also possible to include observation errors in the entropy framework. If those errors are assumed Gaussian a priori, then Ψ corresponds to the 4D-Var PSAS cost function (Courtier, 1997), which will be mentioned in Sect. 6.

The solution to the minimization of the level-2 dual functional can be derived analytically:

$$\overline{\boldsymbol{\beta}} = \left(\boldsymbol{H}\boldsymbol{B}\boldsymbol{H}^T\right)^{-1}\boldsymbol{\mu}\,. \tag{29}$$

It gives the average source through $\overline{\boldsymbol{\sigma}} = \boldsymbol{B}\boldsymbol{H}^T\overline{\boldsymbol{\beta}}$, the problem solution.

This reconstruction is equivalent to the following projection method: a natural idea would be to suppose a tentative source of the form

$$\boldsymbol{\sigma}_\odot = \sum_{i=1}^{p} \lambda_i \boldsymbol{c}_i^*\,. \tag{30}$$

However this assumption hides the correlated nature of the prior description of the source encapsulated in $\boldsymbol{B}$. To overcome this, a transformation is applied on the vector space of the sources. By Choleskii decomposition there is a matrix $\boldsymbol{P}$ such that $\boldsymbol{B} = \boldsymbol{P}\boldsymbol{P}^T$. All potential sources $\boldsymbol{\sigma}$ are then transformed into $\boldsymbol{P}^{-1}\boldsymbol{\sigma}$. From the forward and backward transport equations, it can be verified that the adjoint discrete solutions should transform according to $\boldsymbol{c}_i^* \longrightarrow \boldsymbol{P}^T\boldsymbol{c}_i^*$. Once this transformation is performed, the components of a source are decorrelated. The solution $\boldsymbol{\sigma}_\odot$ is therefore a combination of the form

$$\boldsymbol{P}^{-1}\boldsymbol{\sigma}_\odot = \sum_{i=1}^{p} \lambda_i \boldsymbol{P}^T\boldsymbol{c}_i^*\,. \tag{31}$$

The scalar product of the members of this equation with $\boldsymbol{P}^T\boldsymbol{c}_j^*$ gives

$$\mu_j = \sum_{i=1}^{p} \lambda_i (\boldsymbol{c}_j^*, \boldsymbol{B}\boldsymbol{c}_i^*) \quad \text{or} \quad \boldsymbol{\mu} = (\boldsymbol{H}\boldsymbol{B}\boldsymbol{H}^T)\boldsymbol{\lambda}\,. \tag{32}$$

Therefore $\boldsymbol{\lambda}_\odot = \left(\boldsymbol{H}\boldsymbol{B}\boldsymbol{H}^T\right)^{-1}\boldsymbol{\mu}$ and the previous solution $\boldsymbol{\sigma}_\odot = \overline{\boldsymbol{\sigma}}$ is recovered.

Another way to characterize the projection state is to describe it as the minimal norm $||\boldsymbol{P}^{-1}\boldsymbol{\sigma}_\odot||$ state. That is why we define the canonical scalar product $(\boldsymbol{x}, \boldsymbol{y}) = \sum_{k=1}^{N} x_k y_k$ using the same notation as for the L^2 structure. Now the additional scalar product is defined by $\boldsymbol{x}, \boldsymbol{y} \longrightarrow (\boldsymbol{x}, \boldsymbol{y})_{\boldsymbol{B}^{-1}} = (\boldsymbol{x}, \boldsymbol{B}^{-1}\boldsymbol{y})$ and its related norm $x \longrightarrow ||\boldsymbol{x}||_{\boldsymbol{B}^{-1}} = (\boldsymbol{x}, \boldsymbol{x})_{\boldsymbol{B}^{-1}}$. So that $\boldsymbol{\sigma}_\odot$ is the minimal norm $||\boldsymbol{\sigma}_\odot||_{\boldsymbol{B}^{-1}}$ state.

As a consequence the projection method onto the space of retroplumes with cell to cell correlations properly assessed can be viewed as a particular case of the entropy method when all priors are Gaussian.

To go full circle, let us mention that the level-1 primal cost function for Gaussian prior is easily obtained:

$$\mathcal{J} = -\frac{1}{2}||\boldsymbol{\sigma}||^2_{\boldsymbol{B}^{-1}} \tag{33}$$

provided $\boldsymbol{\mu} = \boldsymbol{H}\boldsymbol{\sigma}$. Incidentally, this is a proof of the characterization of the projection by minimal norm.

A generalization to a non-zero first moment prior would be straightforward.

3 Singularity in the continuum limit

In the Gaussian case (equivalent to the projection method) the reconstruction is obtained though the inversion of a Gram matrix $\boldsymbol{G} = \boldsymbol{H}\boldsymbol{B}\boldsymbol{H}^T$ (the Hessian of the dual problem). For simplicity $\boldsymbol{B} = \boldsymbol{I}$ is assumed, so that $\boldsymbol{G} = \boldsymbol{H}\boldsymbol{H}^T$. $\boldsymbol{G}$ is well-defined on a grid, as its entries are discrete sums, although it may not be well-conditioned. However, as it was mentionned in Issartel (2003), this is not the case in the continuum limit. In the continuum limit $[\boldsymbol{G}]_{ij} = \int_\omega \mathrm{d}t\mathrm{d}\boldsymbol{x}\, c_i^*(\boldsymbol{x}, t) c_j^*(\boldsymbol{x}, t)$, where ω is the ground trace of the total domain Ω, outside of which the source is known to be null a priori. When $i = j$ this integral is singular in several cases of interest. Suppose $\pi_i(\boldsymbol{x}, t) = \delta(\boldsymbol{x} - \boldsymbol{x}_i)\delta(t - t_i)$ is a Dirac sampling function corresponding to a short period measurement at a site included in the domain ω (typically a surface observation). Then $\int_\omega \mathrm{d}t\mathrm{d}\boldsymbol{x}\, \left(c_i^*(\boldsymbol{x}, t)\right)^2$ diverges because the integrand is singular close to the observation site and non-integrable. As a consequence the Gram matrix is not properly defined and needs regularization, as was argued in Issartel (2003).

Operational problems are solved through numerics however, so that the Gram matrix is well-defined in practice (whether it is ill-conditioned is another issue). This objection does not discard the physical issue. Indeed as one makes the grid-cell smaller and smaller, the norm of $\boldsymbol{G}$ continues to diverge. An explanation for this phenomenon is exposed in the following sections.

To demonstrate the existence of the singularity as well as exhibit its dimensional dependence, a typical diagonal element of the Gram matrix must be estimated as the mesh step goes to zero.

3.1 Meteorological conditions influence

When it is present, the divergence is caused by the turbulent diffusion operator. This diffusion represents a sub-grid parameterization of advective stirring and mixing. Therefore diffusion, as an effective representation of dispersion at smaller scale, should always be present. A strong wind would lessen the importance of diffusion, and the Gram matrix divergence would be less dramatic.

A quantitative evaluation of the Gram matrix divergence is therefore bound to the meteorological conditions (see Enting, 2002, and references therein).

However, in the asymptotic limit when g diverges, the Gram matrix depends only on diffusion in a simpler way,

because advection is smooth in the vicinity of an observation site and does not contribute to the singular behaviour. This ultimately explains the behaviour of the Gram matrix (and ultimately the inversion) when the resolution increases.

In the reconstruction ahead, three typical scales characterizing the grid resolution are considered. They are the time-step Δ_t, the time scale related to horizontal diffusion Δ_x^2/K_h, and the time scale related to vertical diffusion Δ_z^2/K_z, where Δ_x is the horizontal resolution (mesh step) and Δ_z is the vertical resolution. Very rough estimations for continental transport studied in Sect. 4 indicate values of 1 h, 10 h (for the shortest) and 1/2 h, respectively. As a consequence, time and the vertical diffusion have the best resolution. Hence taking Δ_t and Δ_z to the continuum limit before Δ_x is sound, in this context.

3.2 Asymptotic divergence

Several types of inversion will be considered here. The inversion depends on three important factors: the spatial dimensionality of diffusion (usually three-dimensional and taking place in the domain Ω), the spatial dimensionality of the expected support of the source (denoted ω above), and the coverage of the observation network with respect to the source support. An additional factor is the way time is taken into account. For an accidental release, time is an extra dimension to cope with. For steady or time-averaged emissions, time is irrelevant, and diffusion has a different behaviour to take into account.

3.2.1 Accident-type reconstruction – first case

Consider a network of receptors dedicated to the detection and measurement of an accidental release of pollutant. The diffusion of this pollutant is a three-dimensional process $D = 3$. The support of the source (the space where the source is believed to be located a priori) is expected to lie within the space where the diffusion takes place. The source support is assumed to be a space of dimension d. In this space, the diffusion parameter is K_h. The diffusion parameter in the $D - d$ extra dimensions is denoted K_z. Typically, a ground network of receptors is used to reconstruct a $d = 2$ ground source from the $D = 3$-dimensional atmospheric dispersion.

Here it is assumed c^* is one of the adjoint solutions of Eq. (10), with $\pi(\boldsymbol{x}, t) = \delta(\boldsymbol{x} - \boldsymbol{x}_0)\delta(t - t_0)$. The reconstruction only requires the trace of c^* on the source support to be known. A diagonal entry of the Gram matrix is of the form

$$g = \int_{\omega_\circ} \mathrm{d}t\mathrm{d}\boldsymbol{x} \ \left(c^*(\boldsymbol{x}, t)\right)^2 . \tag{34}$$

$\omega_\circ$ is the expected support of the source (a surface representing the ground multiplied by the time interval) punctured by a cylinder around $(\boldsymbol{x}_0, t_0)$ of radius Δ_x and height $2\,\Delta_t$. g needs to be studied as Δ_t and Δ_x vanish.

A position vector $\boldsymbol{x}$ in the $D = 3$-dimensional space of diffusion is decomposed as $\boldsymbol{x} = \boldsymbol{z} \oplus \boldsymbol{r}$, where $\boldsymbol{r}$ belongs to the $d = 2$-dimensional space where the source is suspected to be. $\boldsymbol{z}$ is in the complementary $(D - d)$-dimensional vector. Assuming a close vicinity to the receptor site, advection is neglected and the diffusion parameters are supposed uniform and constant. Therefore the adjoint solution in the neighbourhood of the receptor is:

$$c^*(\boldsymbol{r}, \boldsymbol{z}, t) = \frac{\exp\left\{-\frac{1}{t}\left(\frac{|\boldsymbol{r}|^2}{4K_h} + \frac{|\boldsymbol{z}|^2}{4K_z}\right)\right\}}{\sqrt{(4\pi t)^D\, K_h^d\, K_z^{D-d}}} . \tag{35}$$

As justified before, the limits Δ_t and Δ_z going zero are taken first. The asymptotic behaviour of the matrix entry g can then be obtained as Δ_x goes to zero. After some algebra, one obtains that, if $d - 2D + 2 < 0$, g is divergent as

$$g \sim (D-2)!\frac{S_d}{2\pi^D}\frac{1}{K_h}\left(\frac{K_h}{K_z}\right)^{D-d} \frac{\Delta_x^{d-2D+2}}{2D-d-2} , \tag{36}$$

where $S_d = 2\pi^{d/2}/\,\Gamma(d/2)$ is the area of the unit sphere in dimension d. When $d - 2D + 2 > 0$, there is no divergence. The critical case $d - 2D + 2 = 0$ implies

$$g \sim (D-2)!\frac{S_d}{2\pi^D}\frac{1}{K_h}\left(\frac{K_h}{K_z}\right)^{D-d} \ln\frac{1}{\Delta_x} . \tag{37}$$

Consider the most obvious case: $D = 3$. Then the critical dimension for the source is $d = 4$, below which g is divergent. Therefore, in the case of interest $d = 2$, there is a divergence characterized by

$$g \sim \frac{1}{2\pi^2 K_z}\Delta_x^{-2} . \tag{38}$$

This behaviour is controlled by a fully developed vertical diffusion. In the absence of vertical diffusion $D = 2$, the actual behaviour would be critical

$$g \sim \frac{1}{\pi K_h}\ln\frac{1}{\Delta_x} . \tag{39}$$

It is reasonable to assert that for realistic meteorological conditions, the behaviour lies in between those two limiting behaviours. This issue is indirectly related to the question of how the source fluctuations are propagated to the receptor: it behaves like $1/|\boldsymbol{k}|$ (Newsam and Enting, 1988) whereas a purely two-dimensional dispersion would yield a $1/|\boldsymbol{k}|^2$ behaviour.

Equation (36) can be applied to other cases of interest. For example in situation where fields are averaged over one or more dimensions (zonal or vertical average).

3.2.2 Accident-type reconstruction – second case

Another interesting case is when the receptors are located outside the source support. A reconstruction of a temporal release profile while knowing the location of the release is a typical example. It can be shown that there is no divergence of g, when the resolution is increased.

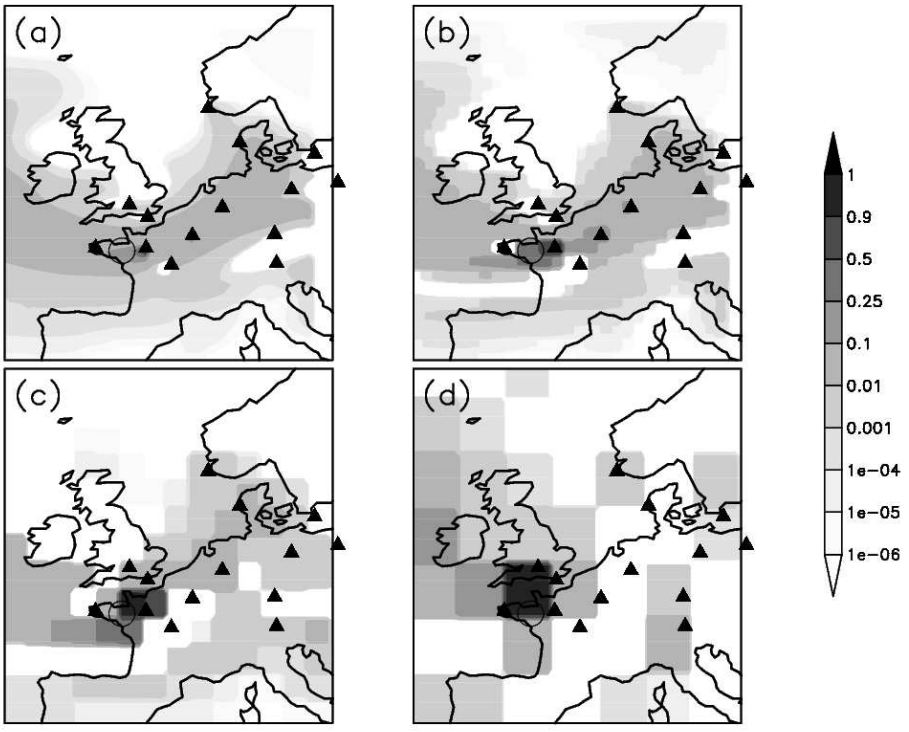

Fig. 1. Reconstructions with a Gaussian prior for four different spatial angular resolutions. Resolutions for panels **(a)**, **(b)**, **(c)**, and **(d)** are $0.5° \times 0.5°$, $1° \times 1°$, $2° \times 2°$, and $4° \times 4°$, respectively. All retrieved sources are normalized so that the maximum of the integrated field be 1. The ETEX release site is indicated by the center of the drawn circle.

3.2.3 Steady state release

The reconstruction method can also be applied to steady-state emission, with potential applications to CO_2 flux inversion, atmospheric mercury inversion, etc. In that case, an approximation is to consider the retroplumes as steady-state solutions with a typical sampling function of $\pi = \delta(\boldsymbol{x} - \boldsymbol{x}_0)$. The typical diagonal Gram entry is

$$g = \int_{\omega_\circ} \mathrm{d}\boldsymbol{x} \ \left(c^*(\boldsymbol{x})\right)^2 \ . \tag{40}$$

The steady state three-dimensional purely diffusive retroplume is characterized by a decrease from the ground receptor site of $1/|\boldsymbol{x}|$ (Seinfeld and Pandis, 1998). Then for $d < 2$ there is a divergence. g behaves like Δ_x^{d-2}. The critical dimension is therefore $d = 2$. Above $d = 2$, no divergence is expected.

Those results have a significant impact on the reconstruction. This will be detailed in the next sections, with an emphasis on the grid resolution used for the retrievals.

4 Numerical reconstructions for an accident-type release

Before any analytical account is given, examples of numerical reconstructions will be presented. They are based on the European Tracer EXperiment (ETEX). This experiment (here the first campaign will be emphasized) was conducted in 1994 by the European Joint Research Centre (Joint Research Centre, 1998). 340 kilograms (M) of Perfluoromethylcyclohexane (PMCH) were released uniformly from 23 October 1994 16:00 UTC to 24 October 1994 03:50 UTC, at Monterfil (Britanny, France, located 48°03′ N, 2°00′ W). Many research teams were involved in taking measurements to characterize the dispersion of the PMCH cloud over Europe. The results of the experiment were used to calibrate various atmospheric dispersion models (see the special issue of "Atmospheric Environment", 32, 24, 4089–4375 or the proceedings of the "ETEX symposium on long-range atmospheric transport, model verification and emergency response", 13–16 May 1997, Vienna (Austria), K. Nodop Ed.). Alternatively these measurements can be used to test source inversion methods.

In the following sections, twin experiments are performed for three sets of observations denoted A, B, and C of increasing size (48, 137 and 237 measurements, respectively) (see Bocquet, 2005b, for a complete description of those sets).

The resolution will be varied. The time-step is set to $\Delta_t = 1\,\mathrm{h}$, whereas the zonal and meridional angular steps take four different values ($0.5° \times 0.5°$, $1° \times 1°$, $2° \times 2°$, and $4° \times 4°$). The source which is numerically implemented mimics the ETEX-I source. It is contained in the cell to which the ETEX-I release site is attached, and uniformly distributed over 12 h.

The transport model used is POLAIR3D (Boutahar et al., 2004; Sartelet et al., 2002; Sportisse et al., 2002). ECMWF wind fields serve as inputs to the model.

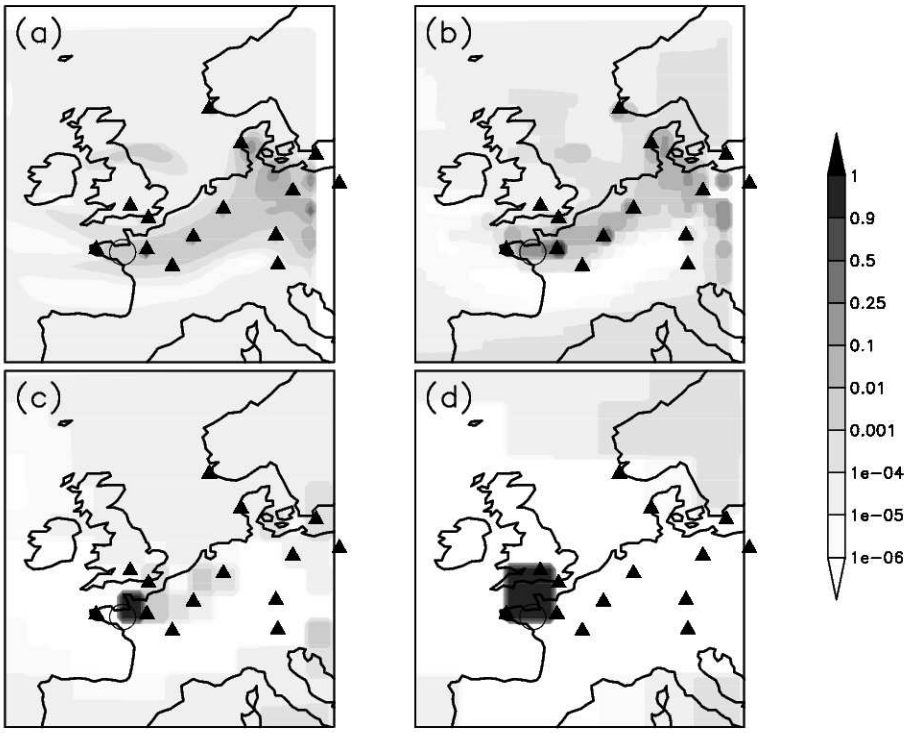

Fig. 2. Reconstructions with a Poisson prior. Features are the same as for the graphs related to the Gaussian case Fig. 1.

As suggested before, two types of prior will be used: Gaussian prior and Poisson prior. They are expected to perform poorly and fairly, respectively, with the observation sets and resolution at hand. They are interesting choices since they correspond to quite different outcomes of reconstructions and performances in the inversion (for optimal results of the inversion methods, see Bocquet, 2005b).

4.1 Gaussian prior

More precisely, the Gaussian prior is specified by the prior covariance matrix

$$\boldsymbol{B} = m^2 \boldsymbol{I} \,. \tag{41}$$

m^2 sets the scale of the local variance. In absence of noise, its actual value is of no consequence because it is easy to see in Eq. (29) that the retrieved source is independent of a rescaling of $\boldsymbol{B}$. In Fig. 1 four reconstructions are considered. In order for the various reconstructions to be comparable, the simplifications used in Bocquet (2005b) were dropped (rejection of water bodies cells, specific treatment of quasi-null measurements), so that the number of variables to invert is the same at constant resolution.

The reconstruction is performed within the following volume: from 12° W to 16° E, and 40° N to 66° N over a time-period of 75 h centered around the real release period.

The assimilated observations belong to set B. The angular resolution is set for both Δ_x and Δ_y to 0.5°, 1°, 2°, and 4° on graphs (a), (b), (c) and (d), respectively. Because it is not possible to give on paper a temporal account of the sources, they have been integrated in time. Since there is no prior positiveness with a Gaussian prior, negative values may be obtained. Cells with negative values appear white on those graphs. The reconstructed mass is 309 kg, 314 kg, 333 kg and 412 kg, respectively (to compare with 340 kg). Incidentally, this shows why the reconstructed mass cannot be retained as a score for the reconstruction (as opposed to ρ defined later on): those masses are actually differences between sources and sinks. To a limited extent, these graphs allow to check visually how the reconstruction degrades as the resolution increases. It is observed that the receptor prints on the retrieved sources get stronger as the resolution increases. For the higher resolution (graph a), higher values of the field are hidden by the triangles marking the receptors.

4.2 Poisson prior

The Poisson prior is defined by a reference mass m which will be taken to be $m_0 \equiv M/12$. Since M is not supposed to be known prior to the inversion, this choice should be considered optimal. It is unlikely but just as good as any other choice. Parameter θ_k is chosen to be uniform as θ. It depends on the resolution since $m\theta$ is the prior average released mass in a cell. The choice of $\theta = \theta_0$ for 0.5°, where $\theta_0 = 10^{-6}$, leads consistently to $\theta = 4\,\theta_0$ for 1°, $\theta = 16\,\theta_0$ for 2°, and $\theta = 64\,\theta_0$ for 4°. The explicit form of the Poisson prior should depend on the scale so that the physical prior assumptions remain scale invariant. However because we are looking for a near black source, the parameter θ remains small whatever the resolution tested here, and the effect of its variation on the reconstruction is expected to be weak.

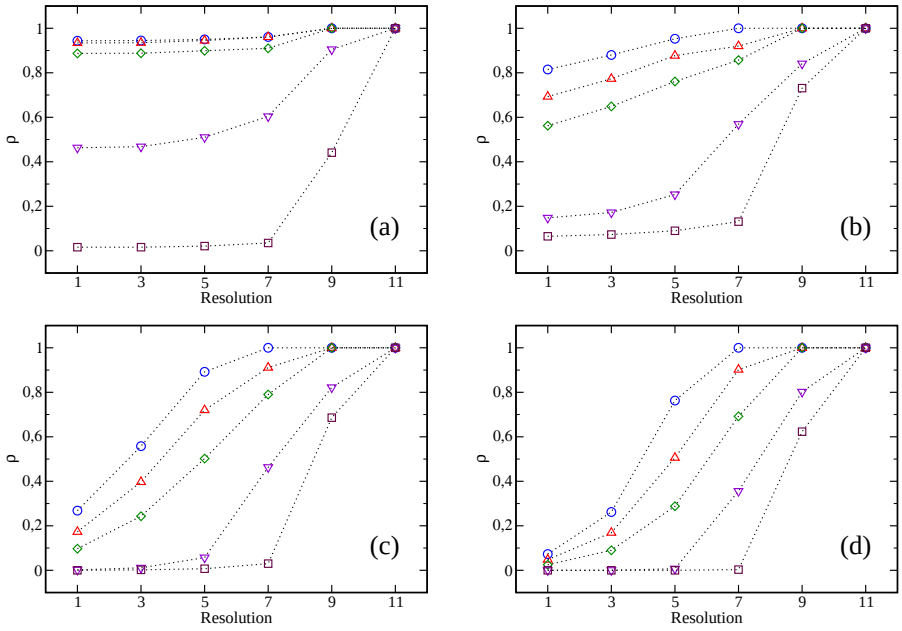

Fig. 3. These curves represent the score ρ for various reconstructions of the toy-model described in the text. Graph **(a)** corresponds to basis functions of Gaussian type. Graphs **(b)**, **(c)** and **(d)** correspond to power law functions of exponent $-1/2$, -1 and $-3/2$, respectively. The curves corresponding to $m = 1, 2, 3, 4$ and 5 (symbols $\bigcirc$, $\triangle$, $\Diamond$, $\triangledown$, and $\square$, respectively) are plotted with respect to the resolution index $n = 1, 3, 5, 7, 9$ and 11.

Four reconstructions with a spatial resolution of $0.5° \times 0.5°$, $1° \times 1°$, $2° \times 2°$, and $4° \times 4°$ are considered, with the time-integrated retrieved source shown on Figs. 2a, b, c and d, respectively. The conditions for the retrieval are the same as for the Gaussian case and therefore comparable. The reconstructed mass is 166 kg, 223 kg, 338 kg and 340 kg, respectively. The receptors imprints on the retrieved sources are visible on the upper graphs corresponding to higher resolution and degraded reconstruction.

The reconstructions are better in the Poisson case than in the Gaussian case. However it is difficult to judge by the mass or by the graphs for the finest resolutions (a and b). A quantitative indicator of the performance is therefore needed.

5 Estimating the analysis error and its dependence on the mesh step

In order to study the impact of resolution on the retrieval, the error made in the reconstruction of a discrete source $\boldsymbol{\sigma}$ must be estimated. To quantify this error, a visual account is often insufficient, all the more since $\overline{\boldsymbol{\sigma}}$ may represent a complex multidimensional field.

5.1 Error estimation for the Gaussian case

At first the prior is assumed Gaussian, although the first guess is not known and hence taken to be null. The reconstruction is equivalent to a projection. The first guess (actually the null source $\mathbf{0}$) is projected orthogonally with respect to the scalar product $(,)_{\boldsymbol{B}^{-1}}$ onto the vector space generated by $\boldsymbol{P}^T\boldsymbol{c}_i^*$. The projected state is $\overline{\boldsymbol{\sigma}}$. The true solution $\boldsymbol{\sigma}$ also belongs to this vector space (see Fig. 4). We need to estimate the distance between the two states using the metric defined by the scalar product. Building on an r.m.s. indicator suggests taking $||\boldsymbol{\sigma} - \overline{\boldsymbol{\sigma}}||_{\boldsymbol{B}^{-1}}$. Unfortunately its computation requires a known $\boldsymbol{\sigma}$. Additionally, for a complex correlation matrix, this computation is too numerically demanding for a fine grid. Those two obstacles can be overcome in the following way. Because of the projection, the Pythagoras theorem can be invoked:

$$||\boldsymbol{\sigma} - \overline{\boldsymbol{\sigma}}||^2_{\boldsymbol{B}^{-1}} = ||\boldsymbol{\sigma}||^2_{\boldsymbol{B}^{-1}} - ||\overline{\boldsymbol{\sigma}}||^2_{\boldsymbol{B}^{-1}} \,. \tag{42}$$

Obviously the norm $||\boldsymbol{\sigma}||_{\boldsymbol{B}^{-1}}$ is not known. But it is clear that $||\overline{\boldsymbol{\sigma}}||_{\boldsymbol{B}^{-1}}$ is enough to indicate if a reconstruction gets better with changing conditions. Note that as $||\overline{\boldsymbol{\sigma}}||_{\boldsymbol{B}^{-1}}$ grows, the reconstruction improves. This could be confusing because $\overline{\boldsymbol{\sigma}}$ can be seen alternatively as the state of the vector space generated by the $\boldsymbol{P}^T\boldsymbol{c}_i^*$ which has the minimal norm. The two statements are independent however and there is no contradiction.

Therefore $||\overline{\boldsymbol{\sigma}}||_{\boldsymbol{B}^{-1}}$ needs to be evaluated. Its numerical computation makes use of the dual approach. It was stated that the level-1 primal cost function is $\mathcal{J} = -\frac{1}{2}||\boldsymbol{\sigma}||_{\boldsymbol{B}^{-1}}$. In addition, at the optimum, the level-1 cost function equals the level-2 cost function: $\mathcal{J}(\overline{\boldsymbol{\sigma}}, \nu) = \mathcal{K}(p_{\overline{\sigma}}, \nu)$. $\mathcal{K}(p_{\overline{\sigma}}, \nu)$ is a natural outcome of the minimization on $\boldsymbol{\beta}$ however, and $||\overline{\boldsymbol{\sigma}}||_{\boldsymbol{B}^{-1}}$ can be computed efficiently through this procedure.

From the analytical perspective calculating $||\overline{\boldsymbol{\sigma}}||_{\boldsymbol{B}^{-1}}$ is a difficult problem-dependent task. However there are a few

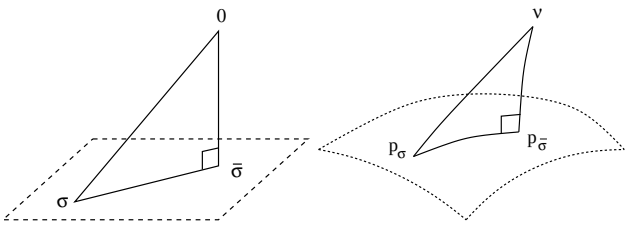

Fig. 4. The left figure depicts the optimization procedure in the Gaussian framework. The first guess (actually the null vector **0**) is projected orthogonally onto the space of sources which satisfy the measurement constraints. The true source $\boldsymbol{\sigma}$ belongs to this space. The right figure depicts the optimization for a general prior within the maximum entropy on the mean approach. The prior ν is "projected" onto the manifold of the exponential laws, the average of which satisfies the measurement constraints. The exponential law whose first moment is the true source belongs to this manifold.

simplifications in the asymptotic limit where the divergence takes place, i.e. when the spatial mesh step becomes smaller.

5.2 Score of the reconstruction (Gaussian case)

To give an account of the performance of a reconstruction in this context, and building on the previous remarks, we may introduce the ratio

$$\rho = ||\overline{\boldsymbol{\sigma}}||^2_{\boldsymbol{B}^{-1}} / ||\boldsymbol{\sigma}||^2_{\boldsymbol{B}^{-1}} \,. \tag{43}$$

Let us write a form of $||\overline{\boldsymbol{\sigma}}||^2_{\boldsymbol{B}^{-1}}$ that depends on the source $\boldsymbol{\sigma}$ only. Using $\overline{\boldsymbol{\sigma}} = \boldsymbol{B}\boldsymbol{H}^T\boldsymbol{G}^{-1}\boldsymbol{\mu}$ with $\boldsymbol{G} = \boldsymbol{H}\boldsymbol{B}\boldsymbol{H}^T$, we have

$$\begin{aligned} ||\overline{\boldsymbol{\sigma}}||^2_{\boldsymbol{B}^{-1}} &= \boldsymbol{\mu}^T\boldsymbol{G}^{-1}\boldsymbol{H}\boldsymbol{B}\boldsymbol{B}^{-1}\boldsymbol{B}\boldsymbol{H}^T\boldsymbol{G}^{-1}\boldsymbol{\mu} \\ &= \boldsymbol{\mu}^T\boldsymbol{G}^{-1}\boldsymbol{\mu} \\ &= \boldsymbol{\sigma}^T\boldsymbol{H}^T\boldsymbol{G}^{-1}\boldsymbol{H}\boldsymbol{\sigma}, \end{aligned} \tag{44}$$

because $\boldsymbol{\mu} = \boldsymbol{H}\boldsymbol{\sigma}$. Consider the matrix $\boldsymbol{H}$ as an operator between Hilbert spaces $(\mathbb{R}^N, \boldsymbol{B}^{-1})$ and $(\mathbb{R}^p, \boldsymbol{I})$ ($\boldsymbol{I}$, the identity matrix, stands for the canonical scalar product). Then there exists a singular value decomposition of $\boldsymbol{H}$

$$\boldsymbol{H} = \boldsymbol{U}\boldsymbol{D}\boldsymbol{V}^T \,. \tag{45}$$

$\boldsymbol{U}$ is a $p \times p$ orthogonal matrix, $\boldsymbol{D}$ is a $p \times p$ diagonal matrix and $\boldsymbol{V}$ is a $n \times p$ matrix satisfying $\boldsymbol{V}^T\boldsymbol{B}\boldsymbol{V} = \boldsymbol{I}$ (its column vectors are orthogonal with respect to the scalar product defined by $\boldsymbol{B}$). It is then easy to show that

$$\rho = \frac{(\boldsymbol{\sigma}, \boldsymbol{V}\boldsymbol{V}^T\boldsymbol{\sigma})}{(\boldsymbol{\sigma}, \boldsymbol{B}^{-1}\boldsymbol{\sigma})} = \frac{(\widetilde{\boldsymbol{\sigma}}, \widetilde{\boldsymbol{V}}\widetilde{\boldsymbol{V}}^T\widetilde{\boldsymbol{\sigma}})}{(\widetilde{\boldsymbol{\sigma}}, \widetilde{\boldsymbol{\sigma}})} \tag{46}$$

where $\widetilde{\boldsymbol{\sigma}} = \boldsymbol{P}^{-1}\boldsymbol{\sigma}$ and $\widetilde{\boldsymbol{V}} = \boldsymbol{P}^T\boldsymbol{V}$. $\widetilde{\boldsymbol{V}}\widetilde{\boldsymbol{V}}^T$ is a semi-definite operator. Moreover $\widetilde{\boldsymbol{V}}^T\widetilde{\boldsymbol{V}} = \boldsymbol{V}^T\boldsymbol{B}\boldsymbol{V} = \boldsymbol{I}$, so that the spectrum of $\widetilde{\boldsymbol{V}}\widetilde{\boldsymbol{V}}^T$ is a set of eigenvalues, 0 or 1. We conclude that $0 \leq \rho \leq 1$ (which is obvious from Eq. 42), and that ρ is a good indicator to evaluate the efficiency of the reconstruction (the closer to 1, the better).

5.3 Behaviours of the score

Although the form $\rho = (\widetilde{\boldsymbol{\sigma}}, \widetilde{\boldsymbol{V}}\widetilde{\boldsymbol{V}}^T\widetilde{\boldsymbol{\sigma}}) / (\widetilde{\boldsymbol{\sigma}}, \widetilde{\boldsymbol{\sigma}})$ is elegant, it is easier to use the form $\rho = \left(\boldsymbol{\mu}, \boldsymbol{G}^{-1}\boldsymbol{\mu}\right) / \left(\boldsymbol{\sigma}, \boldsymbol{B}^{-1}\boldsymbol{\sigma}\right)$ in the limit where the reconstruction score ρ decreases. Indeed as the resolution increases, the norm $||\boldsymbol{\sigma}||_{\boldsymbol{B}^{-1}}$ remains roughly constant. As do the measurements in the context of a synthetic experiment (for a real experiment, measurements are obviously model resolution independent). Therefore the behaviour of ρ is given by some properties of $\boldsymbol{G}^{-1}$ only. Depending on the choice of the projection space, essentially two kinds of behaviour are expected. Assume, as this is the case for retroplumes, that the projection basis is made of positive vectors. In the continuum limit, those vectors should become positive integrable functions. However they may be non square-integrable.

First suppose they are square-integrable, as the resolution increases the Gram matrix converges to a well-defined operator limit. It is therefore expected that ρ tends to a finite limit between 0 and 1. If they are not integrable, then all diagonal elements of the Gram matrix diverge, while all off-diagonal converge (we assume that the singularities responsible for the non square-integrability of two disctinct functions do not coincide, which is the case for retroplumes). Therefore the Gram matrix tends towards a diagonal matrix of increasing norm. This means that the functions forming the basis are less and less correlated. The norm of its inverse decreases, and the score ρ tends to 0. The reconstruction will eventually fails. The absence of correlation between basis functions implies that the projected source cannot span onto many functions but picks up a few of them localized on the true source support. This provides a mathematical explanation of why the solution is strongly influenced by the receptors in this limit for an atmospheric source reconstruction.

This was valid for a single set of measurements. Of course, insufficient observations will always make the reconstruction fail.

We have tested these ideas on a one-dimensional toy model, exhibiting several kinds of behaviour. The behaviour is determined by the nature of the basis functions which are not necessarily square integrable. The experiment takes place between date $t = 0$ and date $t = 1$. The continuous source is constant and equal to 1 between $t = 0.45$ and $t = 0.55$, and null elsewhere.The grid resolution is $\Delta t = 2^{n-13}$, with $n = 1, 3, 5, 7, 9, 11$. The projection space is spanned by a set of $p = 2^{7-m}$ functions, with $m = 1, 2, 3, 4, 5$. Four cases are considered. First Gaussians with equal root mean square but various centers are used. Next, singular functions of the form $t \to |t - t_i|^{-1/2}$, with t_i being their center ($i \in [1, p]$). Centers are chosen so that the corresponding vectors at a finite resolution be well defined. The square of these functions are not integrable. Eventually, we consider $t \to |t - t_i|^{-1}$ and $t \to |t - t_i|^{-3/2}$ functions, with similar properties though being steeper near t_i. The corresponding divergence of ρ in the time-step is expected to to be $-\ln \Delta_t$, Δ_t^{-1} and Δ_t^{-2} for the last three cases, respectively. Measurements performed are perfect.

The numerical implementation of this toy model confirms the results obtained so far. The score curves are plotted in Fig. 3. For the first case (graph a), the reconstruction stabilizes as the resolution increases, given a single set of measurements. In the second case (graph b), the reconstruction becomes poorer as the resolution increases. The reconstructed mass becomes smaller, and the shape of the retrieved object is increasingly sharper, strongly influenced by a restricted number of basis functions. A similar behaviour is observed in the last two cases (graphs c and d). However the quality of the reconstruction decreases much more rapidly as the resolution increases.

In the following sections, first applications will be considered. In particular the rate of decrease of ρ will be evaluated in the asymptotic limit.

5.4 Application to atmospheric source reconstruction

5.4.1 Ground source and ground observations

For the sake of simplicity, $\boldsymbol{B}$ is set to $\boldsymbol{I}$ from now on. A source of an atmospheric inert tracer, known to be surfacic, is to be retrieved. All p ground concentration measurements are performed over a time and space much smaller than Δ_t and Δ_x, so that they can be considered Dirac-like. Results of Sect. 3 can be used ($D = 3$, $d = 2$ case), Δ_x serving as cut-off. They correspond to the ETEX-type Gaussian reconstruction of Sect. 4. One then has $\boldsymbol{G} \propto \Delta_x^{-2}\boldsymbol{I}$ and

$$||\overline{\boldsymbol{\sigma}}||^2 \propto \left(\boldsymbol{\mu}^T\boldsymbol{\mu}\right)\Delta_x^2 \,. \tag{47}$$

As the mesh-step goes to zero, $||\overline{\boldsymbol{\sigma}}||$ will ultimately vanish. In particular in an ideal situation, the score would asymptotically behave like $\rho \propto \Delta_x^2$. This is alike case (d) of the toy-model. Therefore making the mesh-step too small degrades the reconstruction. This hints at the fact that the amount of tracer retrieved is smaller and smaller. However since L^1 and L^2 norms are distinct, there is no guarantee this statement is always true (especially for a Gaussian prior where positive and negative emission in a cell is allowed).

A physical argument can be proposed to explain the increasing unphysical importance of the receptor sites in the reconstruction. In the limit where the grid cells are smaller and smaller, it is not difficult to imagine a solution easily satisfying the measurement constraints. A small source is placed upwind of any receptor with the right amount of tracer to explain the measurement. The other receptors will be little affected. Such a solution is difficult to settle with a large Δ_x. In addition such a solution is preferred by the regularization (orthogonal projection here) to any ETEX-like one because it is likely to have a small norm. In this limit the reconstruction is bound to fail.

In Table 1 the scores for the reconstructions detailed in Sect. 4 for a Gaussian-type reconstruction are given. We observe that

- Although the resolution does not change over a large range, the score gets worse as the resolution improves, as expected. In any case the score is a decreasing function of the resolution.
- Four magnifications are insufficient to make serious comparisons with the analytical prediction $\rho \propto \Delta_x^2 \propto 4^{-n}$, but the trend is convincingly similar.
- As for the dependence of the quality of the reconstruction on the observation set, the monotony is clearly respected. Indeed, it is easy to prove that if measurement set B encompasses measurement set A, then $\rho_B > \rho_A$. Nevertheless, in this particular example, the dependence on the measurement set is not very relevant because sets A, B, and C are very peculiar.

Table 1. Values of the score ρ for ETEX-like reconstructions, when the prior ν is Gaussian.

set/resolution	$0.5^\circ \times 0.5^\circ$	$1^\circ \times 1^\circ$	$2^\circ \times 2^\circ$	$4^\circ \times 4^\circ$
A	0.007	0.033	0.295	0.661
B	0.007	0.034	0.303	0.862
C	0.008	0.035	0.330	0.913

On one hand it has been shown that for that kind of atmospheric inverse problem increasing the resolution will ultimately degrade the reconstruction. On the other hand the source description should be as precise as possible. Therefore, there exists an optimal resolution for that type of inverse problem.

So far we have assumed that the measurements were Dirac-like and as such responsible for the singularities. Yet, operational measurements are different. For example, ETEX measurements of PMCH concentration are integrated over 3 h. This is a way to regularize the singularities in the Gram matrix. However, it does not change basically the issue for a retrieval at continental scale. Indeed, even regularized retroplumes, should still look very steep close to the receptor area. Only truly non-local observation would change the analysis (for example a column measurement). However, strong correlations with a non-diagonal $\boldsymbol{B}$ may change conclusions, as it may be a way to regularize the Gram matrix. However we do not explore this topic here.

5.4.2 Other source types

In the case where the source is volumic (three spatial dimensions plus time) instead of surfacic, and the observations are surfacic and airborne, results of Sect. 3 yield $\rho \propto \Delta_x$. This is reminiscent of case (c) of the toy-model. The degradation with increasing resolution is less pronounced. Note that it does not imply the reconstruction is easier as the score depends on many other factors.

For an accidental release reconstruction with a single site release known a priori, i.e. when one is interested in retrieving the temporal profile of the accidental release, there is no divergence in the Gram matrix. ρ goes to a finite limit which

Table 2. Values of the score ρ for ETEX-like reconstructions, when the prior ν is Poisson.

set/resolution	$0.5^\circ \times 0.5^\circ$	$1^\circ \times 1^\circ$	$2^\circ \times 2^\circ$	$4^\circ \times 4^\circ$
A	0.082	0.223	0.700	0.996
B	0.325	0.515	0.994	0.998
C	0.475	0.917	0.998	0.998

is not necessarily 0, as the resolution increases. This was observed in Bocquet (2005b). The reconstruction is then easier since there is a resolution above which it cannot get worse. This is not an optimal resolution anymore (except from the point of view of sparing CPU time).

For surfacic sources and airborne observations, the conclusion should be similar. It suggests, at least theoretically, that using aircraft or satellite data prevents there being a resolution limit.

For time-averaging studies of steady ground emissions and sinks, observed by a ground network, one has

$$||\overline{\boldsymbol{\sigma}}||^2 \propto -\frac{\boldsymbol{\mu}^T\boldsymbol{\mu}}{\ln \Delta_x}\,, \tag{48}$$

provided the dispersion can be considered purely diffusive. This is comparable to case (b) of the toy-model. This might be of relevance for trace gas flux inversions. The degradation of the reconstruction as the resolution increases is expected to be very moderate here. Finding the optimal resolution is more difficult in this case, because the degradation is slow as the resolution increases (no sharp transition to help decide).

5.5 Score of the reconstruction (general case)

Let us consider the case of a general prior ν. Even if an explicit formula for the secondary entropy exists, analytical investigations are much more difficult since there is usually no explicit formula for $\overline{\boldsymbol{\beta}}$. That would require an expansion analysis in the moments of pdfs, or alternatively an expansion of order greater than two in the β_i in the secondary entropy. Yet, the dependence from the dimensionality put forward so far is not to change qualitatively. A similar behaviour of the reconstruction with respect to the resolution is still expected.

For numerical twin experiments, it is crucial to define a score that generalizes ρ. An r.m.s., similar to the score ρ in the Gaussian case, may look like a good indicator. Unfortunately, the usual Pythagoras equality does not hold here. Besides there is not a clear correspondence to the dual cost function anymore, which was very useful for numerical applications in the Gaussian case. Because the dual functional cannot be used to get ρ, the computing cost of this indicator could be prohibitive.

The solution stems from a similar identity to the Pythagorean equality for the general case. Instead of the scalar product metric, one should use the Kullback-Leibler divergence (see Gzyl, 1995, and references therein: the analog of the orthogonal projection is known as the Czizar theorem). Three probability density functions are to be considered. First is the prior ν. To define the last two pdfs, the following family of pdfs must be introduced

$$p_{\boldsymbol{v}}(\boldsymbol{s}) = \frac{\nu(\boldsymbol{s})\exp(\boldsymbol{\alpha}^T\boldsymbol{s})}{\sum_{\boldsymbol{s}'}\nu(\boldsymbol{s}')\exp(\boldsymbol{\alpha}^T\boldsymbol{s}')}\,, \tag{49}$$

where the N-vector $\boldsymbol{\alpha}$ parameterizes the family. $\boldsymbol{v}$ is related to $\boldsymbol{\alpha}$ through $\boldsymbol{v} = \nabla_{\boldsymbol{\alpha}}\left\{\ln\sum_{\boldsymbol{s}}\nu(\boldsymbol{s})e^{\alpha^T s}\right\}$. Equivalently, $\boldsymbol{v}$ is the mean of the pdf $p_{\boldsymbol{v}}$. By reciprocity, $\boldsymbol{\alpha}$ can be obtained from $\boldsymbol{v}$ by $\boldsymbol{\alpha} = \nabla_{\boldsymbol{s}}\mathcal{J}\,(\boldsymbol{s},\nu)_{|\boldsymbol{v}}$, where $\mathcal{J}$ is the level-1 primal cost function. The second pdf, $p_{\overline{\boldsymbol{\sigma}}}$, is the solution of the inversion (posterior pdf). It is a member of the exponential family with $\boldsymbol{\alpha} = \boldsymbol{H}^T\overline{\boldsymbol{\beta}}$. The third pdf, $p_{\boldsymbol{\sigma}}$, is characterizing the true source $\boldsymbol{\sigma}$. It is a member of the exponential family with $\boldsymbol{\alpha} = \nabla_{\boldsymbol{s}}\mathcal{J}\,(\boldsymbol{s},\nu)_{|\boldsymbol{\sigma}}$. $\boldsymbol{\sigma}$ designates this pdf because $p_{\boldsymbol{\sigma}}$ is the only member of the exponential family of densities whose mean is the true source $\boldsymbol{\sigma}$.

The Pythagoras identity can be proved along the lines:

$$\begin{aligned}\mathcal{K}(p_{\boldsymbol{\sigma}},\nu) &= \sum_{\boldsymbol{s}} p_{\boldsymbol{\sigma}}\ln\left(\frac{p_{\boldsymbol{\sigma}}}{\nu}\right)\\ &= \sum_{\boldsymbol{s}} p_{\boldsymbol{\sigma}}\ln\left(\frac{p_{\boldsymbol{\sigma}}}{p_{\overline{\boldsymbol{\sigma}}}}\right) + \sum_{\boldsymbol{s}} p_{\overline{\boldsymbol{\sigma}}}\ln\left(\frac{p_{\overline{\boldsymbol{\sigma}}}}{\nu}\right)\\ &\quad + \sum_{\boldsymbol{s}}(p_{\boldsymbol{\sigma}} - p_{\overline{\boldsymbol{\sigma}}})\ln\left(\frac{p_{\overline{\boldsymbol{\sigma}}}}{\nu}\right).\end{aligned} \tag{50}$$

Since

$$\begin{aligned}\sum_{\boldsymbol{s}}(p_{\boldsymbol{\sigma}} - p_{\overline{\boldsymbol{\sigma}}})\ln\left(\frac{p_{\overline{\boldsymbol{\sigma}}}}{\nu}\right) &= \sum_{\boldsymbol{s}}(p_{\boldsymbol{\sigma}} - p_{\overline{\boldsymbol{\sigma}}})\,\overline{\boldsymbol{\beta}}^T\boldsymbol{H}\boldsymbol{s}\\ &= \overline{\boldsymbol{\beta}}^T\boldsymbol{H}(\boldsymbol{\sigma} - \overline{\boldsymbol{\sigma}}) = 0,\end{aligned} \tag{51}$$

one eventually obtains

$$\mathcal{K}(p_{\boldsymbol{\sigma}}, p_{\overline{\boldsymbol{\sigma}}}) = \mathcal{K}(p_{\boldsymbol{\sigma}},\nu) - \mathcal{K}(p_{\overline{\boldsymbol{\sigma}}},\nu)\,, \tag{52}$$

which is adopted as a measure of the distance (though it is not in the mathematical sense) between the estimate and the true source. As for the Gaussian case, $\mathcal{K}(p_{\boldsymbol{\sigma}},\nu)$ cannot be known. Nevertheless the behaviour of $\mathcal{K}(p_{\overline{\boldsymbol{\sigma}}},\nu)$ can be studied. It is not difficult to check that when ν is normal, we retrieve Eq. (42). The geometrical analogy between the Gaussian prior and the general case is represented on Fig. 4.

It is therefore natural to define the following reconstruction score

$$\rho = \frac{\mathcal{K}(p_{\overline{\boldsymbol{\sigma}}},\nu)}{\mathcal{K}(p_{\boldsymbol{\sigma}},\nu)} \tag{53}$$

which is an analog of Eq. (43). They correspond when ν is Gaussian. Although there is no explicit analytical solution in the general case, formula Eq. (53) is of great help in numerics. Since $\mathcal{K} > 0$ and because of Eq. (52), one has $0 \leq \rho \leq 1$.

In Table 2 the scores for the reconstructions detailed in Sect. 4 are given for a Poisson-type reconstruction. General remarks made on the Gaussian prior case also apply to this

case. With the Poisson prior being better, it is not surprising that the scores are bigger. The scores range from excellent to poor and they ranged from good to bad in the normal case. However the trends are similar to those of the Gaussian case.

6 Taking model and observation errors into account

As was shown in Bocquet (2005a), errors (partly stemming from observation but mostly from the model) can be incorporated into the framework of inversion methods regularized by entropy, applied to the reconstruction of tracer sources. The PSAS 4D-Var method used in atmospheric data assimilation is recovered as one of those methods. When taking errors into account, the observation equation is

$$\boldsymbol{\mu} = \boldsymbol{H}\boldsymbol{\sigma} + \boldsymbol{\varepsilon}\,, \tag{54}$$

with $\boldsymbol{\varepsilon}$ a vector whose p components ε_i are the errors related to measurements μ_i (though it can also be related to model error). The level-2 primal problem is extended to

$$\mathcal{L} = \mathcal{K}(p,\nu) + \mathcal{K}(q,\zeta) + \boldsymbol{\beta}^T\,(\boldsymbol{\mu} - \boldsymbol{H}\boldsymbol{\sigma} - \boldsymbol{\varepsilon}) \tag{55}$$

where $\mathcal{K}(q,\zeta)$ is the Kullback-Leibler divergence of the error vector:

$$\mathcal{K}(q,\zeta) = \sum_{\boldsymbol{\varepsilon}} q(\boldsymbol{\varepsilon}) \ln\left(\frac{q(\boldsymbol{\varepsilon})}{\zeta(\boldsymbol{\varepsilon})}\right). \tag{56}$$

ζ is the prior pdf for the errors, whereas q denotes the posterior error pdf to be determined in the optimization of $\mathcal{L}$, along with the source pdf p. All derivations presented so far extend to this enlarged framework using this functional. In particular when the source prior and the errors prior are both Gaussian, the level-1 primal functional matches the cost function of the PSAS method. It reads

$$\mathcal{J} = -\frac{1}{2}||\boldsymbol{\sigma}||^2_{\boldsymbol{B}^{-1}} - \frac{1}{2}||\boldsymbol{\varepsilon}||^2_{\boldsymbol{R}^{-1}}\,. \tag{57}$$

Details can be found in Bocquet (2005a).

The score for the reconstruction can be generalized to the inversion problem with errors. A Pythagoras equality can be obtained, following the lines of derivation Eq. (50), and using the fact that $\boldsymbol{\mu} = \boldsymbol{H}\boldsymbol{\sigma} + \boldsymbol{\varepsilon} = \boldsymbol{H}\overline{\boldsymbol{\sigma}} + \overline{\boldsymbol{\varepsilon}}$ to obtain the analog of Eq. (51). One obtains

$$\begin{aligned}\mathcal{K}(p_{\boldsymbol{\sigma}}, p_{\overline{\boldsymbol{\sigma}}}) + \mathcal{K}(q_{\boldsymbol{\varepsilon}}, q_{\overline{\boldsymbol{\varepsilon}}}) = \mathcal{K}(p_{\boldsymbol{\sigma}},\nu) + \mathcal{K}(q_{\boldsymbol{\varepsilon}},\zeta)\\ -\mathcal{K}(p_{\overline{\boldsymbol{\sigma}}},\nu) - \mathcal{K}(q_{\overline{\boldsymbol{\varepsilon}}},\zeta).\end{aligned} \tag{58}$$

Therefore it is quite natural to define the ratio

$$\rho = \frac{\mathcal{K}(p_{\overline{\boldsymbol{\sigma}}},\nu) + \mathcal{K}(q_{\overline{\boldsymbol{\varepsilon}}},\zeta)}{\mathcal{K}(p_{\boldsymbol{\sigma}},\nu) + \mathcal{K}(q_{\boldsymbol{\varepsilon}},\zeta)}\,, \tag{59}$$

which satisfies $0 \leq \rho \leq 1$ because of Eq. (58). ρ is now an indicator for the reconstruction of both the source and the errors.

In the Gaussian case and when the error prior ζ is a normal pdf specified by the error covariance matrix $\boldsymbol{R}$, the ratio reads

$$\rho = \frac{||\overline{\boldsymbol{\sigma}}||^2_{\boldsymbol{B}^{-1}} + ||\overline{\boldsymbol{\varepsilon}}||^2_{\boldsymbol{R}^{-1}}}{||\boldsymbol{\sigma}||^2_{\boldsymbol{B}^{-1}} + ||\boldsymbol{\varepsilon}||^2_{\boldsymbol{R}^{-1}}} \tag{60}$$

with $\boldsymbol{\sigma}$ and $\boldsymbol{\varepsilon}$ as the true source and error vector, and $\overline{\boldsymbol{\sigma}}$ and $\overline{\boldsymbol{\varepsilon}}$ the retrieved source and error vector.

A rather complex study of the efficiency of the inversion in this more realistic framework should be undertaken. However this is far beyond the scope of this paper. An example is nevertheless given on ETEX-I. The set up is the same as in Sect. 4.

To ease the reconstruction, a smaller domain was considered: from 12° W to 12° E, and 40° N to 60° N over a time-period of 59 h centered around the real event. Simplifications which were not retained in Sect. 4 are now used (for instance very low measurements are discarded, but water bodies are not rejected). Set C was chosen, instead of B (but B is contained in C). The model is still assumed perfect. Measurements are generated with the model then perturbed normally with a standard deviation of 10% of the measurements.

The prior source is assumed to follow a Poisson law. We also assume a normal law for the errors. We choose for the error prior variances the square of 10% of the noisy measurements. Figure 5 gives a visual account of the reconstructions (integrated in time). The spatial resolution is $0.5° \times 0.5°$, $1° \times 1°$, $2° \times 2°$, and $4° \times 4°$ for graphs (a), (b), (c), and (d), respectively. The reconstructed mass is 248 kg, 265 kg, 303 kg, and 321 kg, respectively. Table 3 gives the related scores, in addition to scores for sets A and B. There is a clear degradation compared to the noiseless Poisson case. However not all scores are lower to their counterpart (for example set B, $0.5° \times 0.5°$). This is explained by the fact that the score not only marks the reconstruction of the source but also the reconstruction of the errors. Therefore a score related to a noisy reconstruction cannot be directly compared to a score related to a noiseless reconstruction.

In the case where the resolution is set to $4° \times 4°$, the score decreases with the data size. This anomaly can be explained in the following way. In this coarse resolution case, a few observation sites are located in the same cell. If no noise is generated (perfect twin experiment), the anomaly does not appear. When noise is taken into account, randomly generated errors could be different for two different observations at the same time, and different sites though in the same cell. This yields distinct measurements. The reconstruction procedure interprets this as a representativeness error, since it was expecting the same measurements from the model. The reconstruction turns more difficult as more observations share the same cell and same date. In that respect, this apparently paradoxical behaviour of the score is consistent. When all observations are separated, the effect disappears.

7 Conclusions

In this paper, we have investigated a few properties of the reconstruction of an atmospheric tracer source. Taking into account more observation data and significant prior information in the reconstruction has recently enabled high resolution reconstructions. In this paper, it was shown that the limiting behaviour of the reconstructed source with increasing

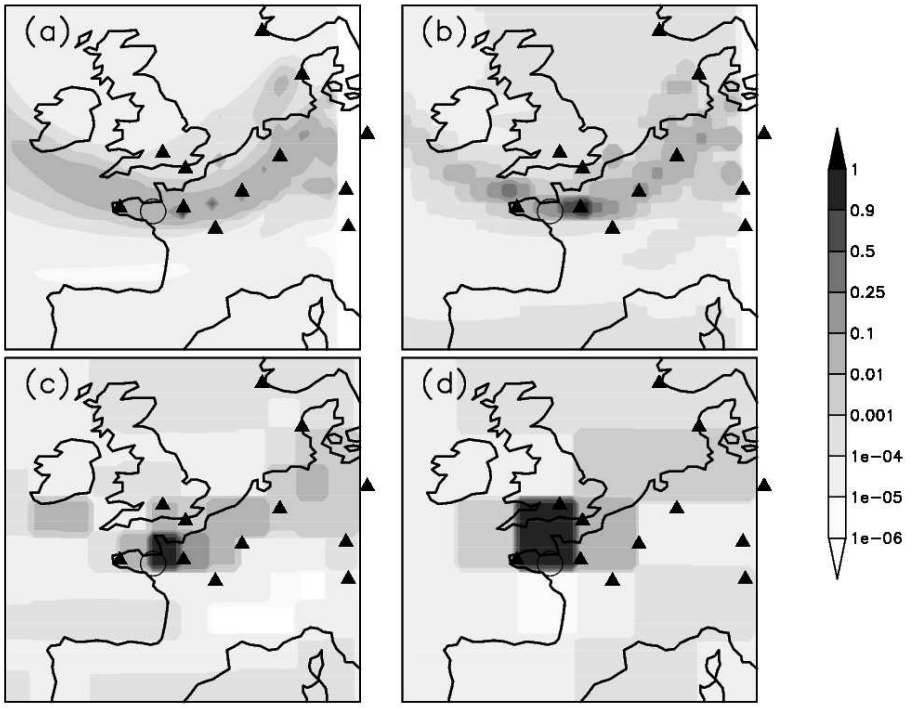

Fig. 5. These graphs show the reconstructed sources with a Poisson prior and a Gaussian prior for the observation errors for four different resolutions. Resolutions for panels (a), (b), (c), and (d) are $0.5° \times 0.5°$, $1° \times 1°$, $2° \times 2°$, and $4° \times 4°$, respectively. Features are the same as for the panels related to the Poisson case Fig. 2 and Gaussian case Fig. 1.

Table 3. Values of the score ρ for ETEX-like reconstructions, when the prior ν is Poisson and when observation error is taken into account as a Gaussian prior.

set/resolution	$0.5° \times 0.5°$	$1° \times 1°$	$2° \times 2°$	$4° \times 4°$
A	0.086	0.222	0.569	0.927
B	0.370	0.468	0.689	0.869
C	0.424	0.554	0.753	0.830

resolution is not simple. The origin of the singular behaviour in the reconstruction of a multidimensional source of an atmospheric inert tracer has been explained. This singularity translates visually into strong prints of the receptors sites onto the sources resulting from reconstruction. For atmospheric dispersion, the emergence of the problem depends on the dimensions of the prior source support, the effective dimension of the atmospheric turbulent diffusion, and how the observations network compares to the source support. The time dimension is also crucial to the analysis.

The problem is not an intrinsic flaw of the inversion methods. It is due to the physics of the dispersion which allows for this fragmented and peaked solutions when the information available (prior and observation) is insufficient and the when the grid resolution is too demanding. This effect disappears when the resolution is coarse enough, or when the background information is rich enough.

Generalized inversion methods were used. They are based on the maximum entropy principle and they include the classical least-square inversion. We have introduced a score to quantify the retrieval performance in the analytically tractable case where the source prior is Gaussian. It is directly related to a classical r.m.s. It has been generalized to more general priors. Strong arguments were given in favour of this score rather than more traditional indicators.

For a given set of observations on tracer concentration performed in the space where the source is expected to be, increasing the resolution of the reconstruction decreases the score in most cases of interest. The score was proved to vanish ultimately. For three-dimensional atmospheric dispersion, the score of an accidental release reconstruction behaves like Δ_x^{4-d} where d is the space dimension of the source support. When the resolution is too high (and therefore the score too low), the receptors influence on the reconstructed object is prominent. Low-dimensional reconstructions (for example a temporal profile retrieval of an emitting plant) do not exhibit this behaviour, since the score goes to a limiting finite value. This is generally true when the observation network is away from the place where the source is expected to be a priori. Airborne measurements for surfacic emissions may help in this respect.

These findings were supported by analytical derivations when the prior on the source is Gaussian. A toy model was used to corroborate these results. Finally these ideas were tested on the case of the ETEX-I experiment for both Gaussian and Poisson priors. The results agree qualitatively with the theory.

Eventually, an extended definition of the score was proposed to incorporate reconstruction based on noisy observations and an imperfect transport model. An example based on ETEX-I has been given.

To conclude, we have identified three interesting questions for future investigations.

This work shows that when inverting atmospheric sources in many problems, there is a resolution limit below which the reconstruction will be poor, and the retrieved source unphysically influenced by the receptor sites. For a specific problem (temporal profile) this issue was proven not to arise. The existence of the limit was proven to exist on semi-quantitative grounds. However calculating a priori the actual value of this threshold is a difficult quantitative issue worth studying. It depends on the meteorological conditions, in particular the relative influence of advection and diffusion.

The variation of the score was studied when considering the grid resolution and the size of the data set. However we have not considered changing the sites of observation and quantified its influence on the score. A valuable study, related to optimal design, would be to determine how to choose an optimal subset of observation sites among the stations which have participated in ETEX.

Finally, it was shown that the score depends on the eddy diffusion tensor. As a sub-grid parameterization, it actually depends on the resolution. This dependence was not taken into account in this work. Further investigations should be performed.

Acknowledgements. The author wishes to thank J.-P. Issartel and B. Sportisse for discussions and encouragements, and O. Talagrand for suggestions on the manuscript. Help from K. Fahey, V. Mallet and Y. Roustan was greatly appreciated.

Edited by: O. Talagrand
Reviewed by: two referees

References

Ashbaugh, L. L, Malm, W. C., and Sadeh, W. Z.: A residence time probability analysis of sulfur concentrations at Grand Canyon National park, Atmos. Environ., 19, 1263–1270, 1985.

Bennett, A. F.: Inverse Modeling of the Ocean and Atmosphere, Cambridge University Press, 2002.

Bocquet, M.: Reconstruction of an atmospheric tracer source using the principle of maximum entropy, I: Theory, Q. J. R. Meteorol. Soc., in press, 2005a.

Bocquet, M.: Reconstruction of an atmospheric tracer source using the principle of maximum entropy, II: Applications, Q. J. R. Meteorol. Soc., in press, 2005b.

Boutahar, J., Lacour, S., Mallet, V., Musson-Genon, L., Quélo, D., Roustan, Y., and Sportisse, B.: Development and validation of a fully modular platform for the numerical modeling of Air Pollution: POLAIR, Int. J. Environ. Poll., 22, 17–28, 2004.

Courtier, P.: Dual formulation of four-dimensional variational assimilation, Q. J. R. Meteorol. Soc., 123, 2449–2461, 1997.

Elbern, H., Schmidt, H., Talagrand, O., and Ebel, E.: 4D variational data assimilation with an adjoint air quality model for emission analysis, Environ. Mod. Softw., 15, 539–548, 2000.

Enting, I. G.: Inverse Problems in Atmospheric Constituent Transport, Cambridge Atmospheric and Space Science Series, edited by: Dessler, A. J., Houghton, J. H., and Rycroft, M. J., 2002.

Gzyl, H.: The method of maximum entropy, Series on Advances in Mathematics for Applied Sciences, Vol. 29, World Scientific Pub Co, 1995.

Hourdin, F. and Issartel, J.-P.: Sub-surface nuclear tests monitoring through the CTBT Xenon network, Geophys. Res. Lett., 27, 2245–2248, 2000.

Hourdin, F., Talagrand, O., and Idelkadi, A.: Eulerian backtracking of Atmospheric tracers, I: Adjoint derivation, parametrization of subgrid-scale transport, Q. J. R. Meteorol. Soc., in press, 2005.

Issartel, J.-P. and Baverel, J.: Inverse transport for the verification of the Comprehensive Nuclear Test Ban Treaty, Atmos. Chem. Phys., 3, 475–486, 2003,
SRef-ID: 1680-7324/acp/2003-3-475.

Issartel, J.-P.: Rebuilding sources of linear tracers after atmospheric concentration measurements, Atmos. Chem. Phys., 3, 2111–2125, 2003,
SRef-ID: 1680-7324/acp/2003-3-2111.

Issartel, J.-P.: Emergence of a linear tracer source from air concentration measurements, Atmos. Chem. Phys., 5, 249–273, 2005,
SRef-ID: 1680-7324/acp/2005-5-249.

Joint Research Centre: ETEX, The European tracer experiment, European Community, 1998.

Louis, J. F.: A parametric model of vertical eddy fluxes in the atmosphere, Boundary Layer Met., 17, 197–202, 1979.

Marchuk, G. I.: Adjoint equations and Analysis of Complex Systems. Mathematics and its Applications, edited by: Hazewinkel, Kluwer Academic Publisher, 295, 1995.

Newsam, G. N. and Enting, I. G.: Inverse problems in atmospheric constituent studies: I. Determination of surface sources under a diffusive transport approximation, Inverse problems, 4, 1037–1054, 1988.

Seinfeld, J. H. and Pandis, S. N.: Atmospheric Chemistry and Physics, From Air Pollution to Climate Change, Wiley-Interscience, 1998.

Pudykiewicz, J. A.: Application of adjoint transport tracer equations for evaluating source parameters, Atmos. Environ., 32, 3039–3050, 1998.

Sartelet, K., Boutahar, J., Quélo, D., Coll, I., Plion, P., and Sportisse, B.: Development and Validation of a 3D Chemistry-Transport Model, POLAIR3D, by Comparison with Data from ESQUIF Campaign, Proceedings of the 6th Gloream workshop: Global and regional atmospheric modeling, 140–146, 2002.

Sportisse, B., Boutahar, J., Debry, E., Quélo, D., and Sartelet, K.: Some tracks in Air Pollution Modeling, POLAIR: a numerical platform for air pollution modeling. RACSAM Journal of the Spanish Science Academy, Real Academia de Ciencias de Espana, 96, 507–528, 2002.

Stohl, A.: Computation, accuracy and applications of trajectories – a review and bibliography, Atmos. Environ., 32, 947–966, 1998.