

Constraining surface emissions of air pollutants using inverse modelling: method intercomparison and a new two-step two-scale regularization approach

Pablo Saide, Marc Bocquet, Axel Osses, Laura Gallardo

► To cite this version:

Pablo Saide, Marc Bocquet, Axel Osses, Laura Gallardo. Constraining surface emissions of air pollutants using inverse modelling: method intercomparison and a new two-step two-scale regularization approach. *Tellus B*, Co-Action Publishing/Blackwell, 2011, 63 (3), pp.360-370. <10.1111/j.1600-0889.2011.00529.x>. <inria-00632515>

HAL Id: inria-00632515

<https://hal.inria.fr/inria-00632515>

Submitted on 13 Nov 2013

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Constraining surface emissions of air pollutants using inverse modelling: method intercomparison and a new two-step two-scale regularization approach

By PABLO SAIDE^{1*}, MARC BOCQUET^{2,3}, AXEL OSSES^{4,5} and LAURA GALLARDO^{5,6},
¹CGRER, Center for Global and Regional Environmental Research, University of Iowa, Iowa City, IA 52242, USA;
²Université Paris-Est, CEREJA Joint Laboratory École des Ponts ParisTech and EDF R&D, Champs-sur-Marne, France; ³INRIA, Paris Rocquencourt Research Center, France; ⁴Departamento de Ingeniería Matemática, Universidad de Chile, Blanco Encalada 2120, piso 5, Santiago, Chile; ⁵Centro de Modelamiento Matemático, UMI 2807/Universidad de Chile-CNRS, Blanco Encalada 2120, piso 7, Santiago, Chile; ⁶Departamento de Geofísica, Universidad de Chile, Blanco Encalada 2002, Santiago, Chile

(Manuscript received 12 August 2010; in final form 8 February 2011)

ABSTRACT

When constraining surface emissions of air pollutants using inverse modelling one often encounters spurious corrections to the inventory at places where emissions and observations are colocated, referred to here as the colocalization problem. Several approaches have been used to deal with this problem: coarsening the spatial resolution of emissions; adding spatial correlations to the covariance matrices; adding constraints on the spatial derivatives into the functional being minimized; and multiplying the emission error covariance matrix by weighting factors. Intercomparison of methods for a carbon monoxide inversion over a city shows that even though all methods diminish the colocalization problem and produce similar general patterns, detailed information can greatly change according to the method used ranging from smooth, isotropic and short range modifications to not so smooth, non-isotropic and long range modifications. Poisson (non-Gaussian) and Gaussian assumptions both show these patterns, but for the Poisson case the emissions are naturally restricted to be positive and changes are given by means of multiplicative correction factors, producing results closer to the true nature of emission errors. Finally, we propose and test a new two-step, two-scale, fully Bayesian approach that deals with the colocalization problem and can be implemented for any prior density distribution.

1. Introduction

Observations of the atmosphere (and in general of any geophysical system) are typically unevenly distributed in space and time and have different precision and accuracy. Models, on the other hand, provide a self consistent framework but are subject to multiple errors due to uncertain estimates of parameters, errors in initial and/or boundary conditions, inadequate representation of processes, and lack of process understanding. The combination of both sources of incomplete and erroneous but complementary information provides a better description of the system state, and its evolution. To achieve this improvement, the combination must be made in an optimal sense, which either minimizes the weighted distance between model and observations (variational

approach) or minimizes the error variance of the system's predictor (Kalman Filter techniques). Both variational and Kalman filter methods have been widely used by the meteorological community over the last decades for weather forecasting. They were introduced as a way to avoid the uncontrolled propagation of errors due to the uncertain and incomplete description of initial conditions (e.g. Kalnay, 2003). In atmospheric chemistry, these methods are nowadays being increasingly used as observations are becoming more readily available. One such application of inverse modelling techniques aims at reducing high uncertainties diagnosed in spatially resolved emissions inventories (e.g. Lindley et al., 2000). Again, both variational (e.g. Elbern et al., 2007) and Kalman filter methods (e.g. Mulholland and Seinfeld, 1995; Peters et al., 2005) have been used, both methods assuming underlying Gaussian distributions over the parameters to be estimated. However, there are other methods that use non-Gaussian assumptions (Bocquet, 2005a) for the priors but have not been tested for distributed sources until now. These

*Corresponding author.

e-mail: pablo-saide@uiowa.edu

DOI: 10.1111/j.1600-0889.2011.00529.x

methods are attractive since they deal naturally with positive emissions and do not require the use of further constraints in the minimization (e.g. Davoine and Bocquet, 2007).

When doing inverse modelling of emissions, one must be aware that misfits between model and observations are not due only to emission inaccuracies, but also can be due to errors in meteorological fields and other model parameters, as well as errors in the representation of physical and chemical processes. For this reason, a strict model validation is often required previous to the inverse modelling stage (e.g. Saide et al., 2009) or one must account for these errors incorporating them into the methodology as an additional term into the minimized functional (e.g. Elbern et al., 2007).

Independently of the inversion method used, when dealing with a spatially and temporally distributed emission inventory one often encounters unrealistic corrections to the a priori inventory at places where ground observations are colocated with distributed sources. This is referred to here as the colocalization problem. These unrealistic corrections arise as the model sensitivity to the emissions is locally maximum at the grid-box where the observation is sampled and due to the scarcity of observations, leading to a correction that is overwhelmed by the observation signal. Moreover, it is possible to prove that the higher the resolution the worse this problem becomes (Bocquet, 2005b). This problem has also been identified by other authors as unrealistic small scale variations (Seibert, 2000) or the effect of not using correlations in the prior uncertainties (Houweling et al., 2004), among others. This problem can arise independently of the scale, from global (e.g. Rödenbeck et al., 2003) to local (e.g. Chang et al., 1997) scales and independently of the numerical dispersion model used.

One objective of this work is to describe and intercompare in a systematic way several approaches that have been used to solve the colocalization problem. Those approaches include: coarsening the spatial resolution of the emissions (e.g. Bocquet, 2005b); adding spatial correlations by means of introducing influence radii (e.g. Houweling et al., 2004; Peters et al., 2005); adding constraints on the spatial derivatives (e.g. Seibert, 2000; Carmichael et al., 2008; Dubovik et al., 2008); and adding weighting factors to the emission error covariance matrix (Saide et al., 2009).

All the approaches mentioned earlier modify the base methodology by making a posteriori corrections that might be more or less arbitrary, and most of them use Gaussian assumptions for emission errors, since they modify the non-diagonal structure of the error covariance matrices. Consequently, a second objective of our work is to propose and test a new fully Bayesian method that solves the problem by using a two-step two-scale approach. The functional to be minimized is theoretically developed taking the colocalization problem into account without making a posteriori modifications. It is based on previous studies of the phenomenon (Bocquet, 2005b) and it is derived from the Bayes formula (Bocquet, 2005a) by making a rigorous infer-

ence using observation and prior information. Also, the method can be implemented both for Gaussian and non-Gaussian prior distributions, without any approximation.

In the next section, some of the previous inverse modelling methodologies are presented. Then, in section 3, we present the colocalization problem and briefly summarize the methods used here to address it. Section 4 presents the fully Bayesian method we propose. Thereafter (Sections 5 and 6), we test these methods by applying them to the test problem of the improvement of a carbon monoxide inventory over the city of Santiago de Chile using real observational data. This makes it possible to present in a common framework to evaluate the advantages, disadvantages and similarities of the various methods. Finally, in Section 7, we summarize the main results and conclusions of this intercomparison study.

2. Inverse modelling methodology

The usual approach used for improving emission inventories or estimating a source of a tracer is a variational method with Gaussian assumptions for the mismatch between model and observations and mismatch between emissions and their priors, presented in the following functional:

$$\mathcal{L}_{\text{Gaussian}}(\boldsymbol{\sigma}) = \frac{1}{2}[\mathcal{H}(\boldsymbol{\sigma}) - \boldsymbol{\mu}]^T \mathbf{R}^{-1}[\mathcal{H}(\boldsymbol{\sigma}) - \boldsymbol{\mu}] + \frac{\alpha}{2}(\boldsymbol{\sigma} - \boldsymbol{\sigma}_b)^T \mathbf{B}^{-1}(\boldsymbol{\sigma} - \boldsymbol{\sigma}_b), \quad (1)$$

where $\boldsymbol{\sigma}$ is the vector of emissions, $\boldsymbol{\sigma}_b$ is the prior or first guess over the emissions, \mathcal{H} is the dispersion model composed with the sampling function over the measurement locations, $\boldsymbol{\mu}$ are the measurements, \mathbf{R} and \mathbf{B} are the covariance matrices of the errors in the measurements and in the emissions respectively and α is a regularization parameter (see Saide et al., 2009, for details). From now on we will consider that \mathcal{H} can be approximated by $\mathcal{H}(\boldsymbol{\sigma}) = \mathcal{H}(\boldsymbol{\sigma}_b) + \mathbf{H}(\boldsymbol{\sigma} - \boldsymbol{\sigma}_b)$, where \mathbf{H} is the Jacobian matrix that maps emissions into measurements, because we intend to apply the inverse modelling methods on carbon monoxide, whose physics is almost linear at the scales that will be considered.

Then the optimal emissions can be written as

$$\bar{\boldsymbol{\sigma}} = \boldsymbol{\sigma}_b + \alpha^{-1} \mathbf{B} \mathbf{H}^T (\mathbf{R} + \alpha^{-1} \mathbf{H} \mathbf{B} \mathbf{H}^T)^{-1} [\boldsymbol{\mu} - \mathcal{H}(\boldsymbol{\sigma}_b)]. \quad (2)$$

In some applications the emissions are restricted to be positive. This is not guaranteed by this method due to the Gaussian assumption for the errors. One way to solve this problem is to find the optimal emissions using eq. (1) with constraints (e.g. Sandu et al., 2005; Chai et al., 2009; Saide et al., 2009). Another path to solve this problem is to change the assumption for the emissions using strictly positive distributions. Bocquet (2005a) showed that functionals similar to eq. (1) can be derived using the principle of maximum entropy on the mean (MEM). For instance a Poisson distribution over the emissions can be assumed. It means that in cell k , the probability that the mass of pollutant

$m_k x_k$, where m_k is a local mass scale, and x_k an integer, is emitted is given by

$$p(x_k) = e^{-\theta_k} \frac{\theta_k^{x_k}}{x_k!}, \quad (3)$$

with θ_k a free parameter of the prior distribution. Considering the mean as $\sigma_b = \mathbf{m}\boldsymbol{\theta}$ and centred measurements $\bar{\boldsymbol{\mu}} = \boldsymbol{\mu} - \mathcal{H}(\sigma_b) + \mathbf{H}\sigma_b$ (observations subtracting the difference between the non-linear and linear models), the corresponding primal cost function (parameters to be minimized are in the emission space) is (Bocquet, 2005a)

$$\mathcal{L}_{Poisson}(\boldsymbol{\sigma}) = \frac{1}{2\alpha} (\mathbf{H}\boldsymbol{\sigma} - \bar{\boldsymbol{\mu}})^T \mathbf{R}^{-1} (\mathbf{H}\boldsymbol{\sigma} - \bar{\boldsymbol{\mu}}) + \sum_{k=1}^N \frac{1}{m_k} \left(\sigma_k \ln \frac{\sigma_k}{\sigma_{b,k}} + \sigma_{b,k} - \sigma_k \right). \quad (4)$$

Contrary to the Gaussian case, this functional does not have an analytic solution for $\bar{\boldsymbol{\sigma}}$ as in eq. (2). Thus it has to be solved numerically. The computational time required to solve these functionals is proportionally related to the number of parameters to optimize [in eq. (4) this is the number of emissions in the guess]. Usually, the number of surface emission parameters is higher than the number of observations and in this case, in order to save computing time, it is preferable to maximize the dual version of this functional (parameters to be optimized are in the observation space)

$$\widehat{\mathcal{L}}_{Poisson}(\boldsymbol{\beta}) = \boldsymbol{\beta}^T \bar{\boldsymbol{\mu}} - \frac{1}{2\alpha} \boldsymbol{\beta}^T \mathbf{R} \boldsymbol{\beta} - \sum_{k=1}^N \frac{\sigma_{b,k}}{m_k} \left\{ \exp(m_k [\boldsymbol{\beta}^T \mathbf{H}]_k) - 1 \right\}, \quad (5)$$

$$\bar{\sigma}_k = \sigma_{b,k} \exp(m_k [\boldsymbol{\beta}^T \mathbf{H}]_k) \quad (6)$$

where $\boldsymbol{\beta}$ is the corresponding dual vector with a dimension corresponding to the number of observations and $[\cdot]_k$ represents the k th component of the vector between brackets. The equivalence to this problem is ensured by the convexity of the MEM inference (Bocquet et al., 2010, and references within). It is interesting to note that in the case of Gaussian priors the optimal inventory is the guess plus a correction term (eq. 2), while when considering a Poisson prior the optimal emissions are the guess multiplied by a correction term (eq. 6). Consequently, the right estimator (Gaussian or Poisson) should be chosen accordingly with the a priori knowledge on the error distribution of the emissions.

3. The colocalization problem and methods for addressing it

Bocquet (2005b, 2009) showed that when recovering sources using ground measurements, the retrieved source often exhibits an unrealistic influence by the observation sites, which is increasingly pronounced as the grid resolution of the source is increased. It was shown to be due to the dispersive nature of

atmospheric transport and to the location of the observations in the manifold of the emission (in control space). Sensitivities stored in the \mathbf{H} matrix have local maxima where sources are colocalized with the observations. When applying the methods described in Section 2 without any modifications, the system has the tendency to overestimate changes in colocalized emissions creating this ‘colocalization problem’ (see Fig. 1a). Note that this phenomenon of colocalization also affects cells around the observation stations and not only the cells encompassing the observations.

This behaviour is independent from the type of distribution chosen to describe the emission inventory. However, when comparing Gaussian with non-Gaussian priors at the same scale, this problem is more obvious on the Gaussian results. Indeed, non-Gaussian priors are usually more informative than Gaussian priors because they usually imply additional assumptions for the emission fluxes (see Bocquet et al., 2010, for a quantitative illustration).

Several methods have been used to solve this problem, so before introducing the new two-scale method in the next section, we review four of the most commonly used methods in the following subsections.

3.1. Coarsening emission spatial resolution

The simplest way to avoid the colocalization problem is by coarsening the emission inventory resolution to a point where the problem is not apparent (Bocquet, 2005b). By choosing the right emission resolution from the beginning the methods presented in Section 2 can be applied directly without modifications. This coarsening can be made in many ways, for instance by grouping emissions by sectors (e.g. Bousquet et al., 1999; Pétron et al., 2002; Jorquera and Castro, 2010) or by just using a coarse model grid (e.g. Kaminski et al., 1999; Henze et al., 2009; Kopacz et al., 2010). It is important to note that many inverse modelling studies do not encounter the colocalization problem and do not have the need to modify their methodology, probably because their starting emission resolution is coarse enough.

3.2. Adding spatial correlations

Probably the most commonly used or standard methodology when the problem is detected and emission resolution is not modified consists in correlating emission cells with their neighbouring grids by specifying the non-diagonal elements of the \mathbf{B} matrix (e.g. Houweling et al., 2004; Peters et al., 2005). By doing this, high differences between neighbour cells are avoided preventing the appearance of colocalization. This method is often presented as enforcing correlations coming from errors in the calculation of the emission inventory. However, one must realize that it adds additional information on the regularity of the emission field which may be incorrect. Some examples on how to construct correlations can be found in Gaspari and Cohn

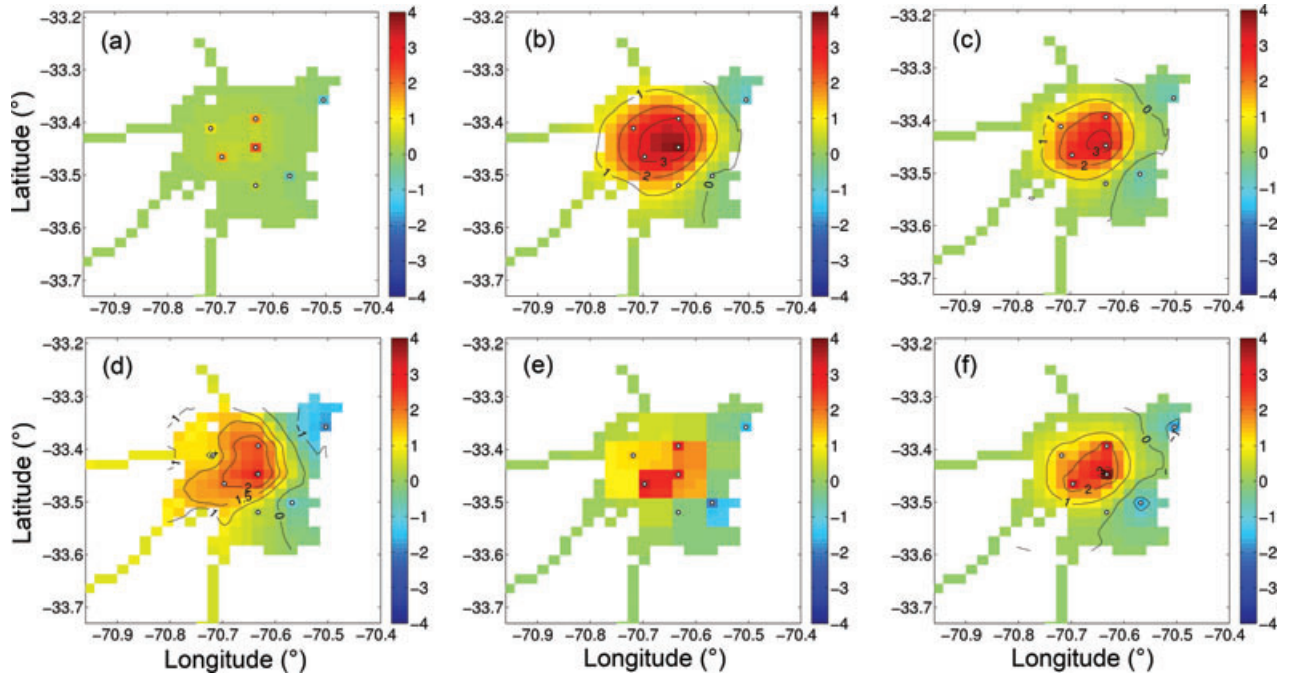


Fig. 1. Time average difference between the a priori (official) inventory and the optimized inventories after assimilating real observations using Gaussian estimates: grid-cell values and smoothed contours. (a) Using no modifications. (b) Using correlations with a radius of influence $r = 2.5$ cells. (c) Adding a constraint on the derivative in the functional. (d) Using weighting factors over \mathbf{B} . (e) Performing a two step multiscale inversion on a coarse grid of 3×3 for one of the possible grid positions and (f) the same as (e) after averaging the nine possible placements of the 3×3 grid. Note that positive values represent a decrease over the background emissions and negative values represent an increase over the background emissions. See the text for details. Units are in $\mu\text{g m}^{-2} \text{s}^{-1}$.

(1999). For instance, a commonly used method is to suppose a radius of influence with an exponential decrease with the distance between cells:

$$\mathbf{B}_{\text{new}} = \mathbf{B} : \mathbf{L}$$

$$L_{ij} = \exp\left(-\frac{d_{ij}}{l}\right) \quad (7)$$

with $(:)$ the Schur product (i.e. piecewise multiplication), d_{ij} some distance between emissions i and j and l a specified length scale. Figure 1b illustrates this method in the Gaussian case.

This method cannot be applied directly to non-Gaussian approaches since a multivariate law (like Gaussian) is needed to establish tractable spatial correlations.

To the knowledge of the authors, the use of anisotropic radii of influence have not been used for estimation of surface sources. However, some work has been done in this direction in data assimilation applications using anisotropic diffusion operators (e.g. Hoelzemann et al., 2001, 2009; Elbern et al., 2007) with promising results.

3.3. Constraining the spatial derivatives

This method has been used in studies such as Seibert (2000), Carmichael et al. (2008) and Dubovik et al. (2008) and consists in adding a supplementary term to the functional from eq. (1) to

constraint the n th order derivative of the difference between the guess emissions and the solution in order to obtain smoothness. The additional term in eq. (1) has the following structure:

$$\frac{\gamma}{2} [\mathbf{D}(\boldsymbol{\sigma} - \boldsymbol{\sigma}_b)]^T [\mathbf{D}(\boldsymbol{\sigma} - \boldsymbol{\sigma}_b)] \quad (8)$$

where γ is a regularization parameter similar to α in eq. (1) and where for instance $\mathbf{D} = \Delta$ (Laplacian $n = 2$, e.g. Seibert, 2000; Carmichael et al., 2008) or \mathbf{D} equals an operator that makes an implicit representation of the n th order derivative (Twomey, 1977). One can show that adding the additional term shown in eq. (8) to (1) is equivalent to replace $\alpha^{-1}\mathbf{B}$ in eq. (1) by

$$\alpha^{-1}\mathbf{B}_{\text{new}} = (\alpha\mathbf{B}^{-1} + \gamma\mathbf{D}^T\mathbf{D})^{-1}$$

$$= \alpha^{-1} \{ \mathbf{B} - \mathbf{B}[\mathbf{B} + \alpha\gamma^{-1}(\mathbf{D}^T\mathbf{D})^{-1}]^{-1}\mathbf{B} \}, \quad (9)$$

where the second equality is obtained using the Sherman–Morrison–Woodbury formula (Golub and Van Loan, 1996). If the original \mathbf{B} matrix is diagonal, eq. (9) can be seen as adding correlation to \mathbf{B} matrix, since the numerical computation of the derivative stored in \mathbf{D} uses spatial neighbours resulting in a non-diagonal square $\mathbf{D}^T\mathbf{D}$ matrix. In summary, this kind of methods results in a specific case of the methods presented in Section 3.2. Figure 1c illustrates this spatial derivative regularizing method in the Gaussian case. Note the similarity with Figure 1b which is the case of introducing spatial correlations.

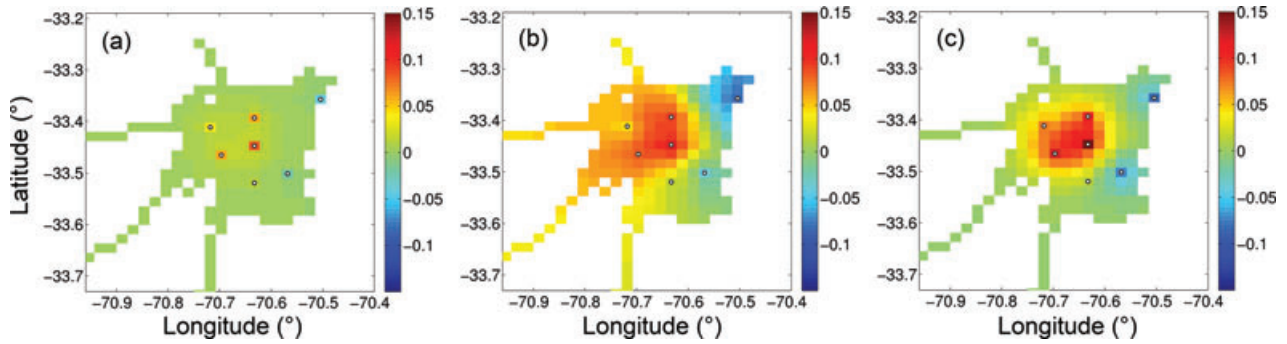


Fig. 2. Time average relative difference between the a priori (official) inventory and the optimized inventories after assimilating real observations using Poisson estimates. (a) Using no modifications. (b) Using factors over \mathbf{B} . (c) Performing a two step multiscale inversion on a coarse grid of 3×3 and averaging the nine possible placements of the grid. Units are in $\mu\text{g m}^{-2} \text{s}^{-1}$.

3.4. Adding a weighting factor in \mathbf{B} matrix

The method of the weighting factor consists in applying a factor \mathbf{F} that smooths singularities or peaks on the collocated spots in the *retroplumes* (Issartel, 2003), which in this study corresponds to the rows of the sensitivity matrix (\mathbf{H}). This can be equivalently achieved by weighting the \mathbf{B} matrix by a weighting factor \mathbf{F}^{-1} . The factor can be chosen as the relative sensitivity computed using some norm of the columns of the \mathbf{H} matrix, further preventing colocalization (Saide et al., 2009).

More precisely, to apply this method to a Gaussian case, \mathbf{B} is replaced by

$$\mathbf{B}_{\text{new}} = \mathbf{F}^{-\rho/2} \mathbf{B} \mathbf{F}^{-\rho/2} \quad (10)$$

for some $\rho > 0$ and with the diagonal of \mathbf{F} computed as (Saide et al., 2009)

$$[\mathbf{F}]_k = \frac{\|h_k\|_q}{\max_k \|h_k\|_q} \quad (11)$$

with $\|h_k\|_q$ the q -norm of the k th column of the \mathbf{H} matrix. Figure 1d illustrates the method for the Gaussian case using $q = 1$ and $\rho = 1$.

This method can be generalized to non-Gaussian non-multivariate distributions. Since the weighting factor is originally applied to the \mathbf{B} matrix, one assumes that the factor affects the variance, while keeping the mean constant. For the Poisson case, the vector of means is $\mathbf{m}\boldsymbol{\theta} = \boldsymbol{\sigma}_b$ and the vector of variances is $\mathbf{m}^2\boldsymbol{\theta} = \mathbf{m}\boldsymbol{\sigma}_b$. Then to fulfil the generalization \mathbf{m} and $\boldsymbol{\theta}$ must be divided and multiplied by the weighting factor, respectively. Figure 2b illustrates the method for the non-Gaussian case using $q = 1$ and $\rho = 1$. Although it has been shown to be efficient, this methodology is empirical and it has not been justified theoretically so far, but it is interpreted as a priori information as is also the case for (7) and (8).

4. Two-step two-scale inversion

As stated in Section 3.1, the colocalization problem can be solved by decreasing the resolution (coarsening) of the emissions in

order to exploit all the information from the observations. However, there should be finer residual information at the smaller scale in the first guess emissions ($\boldsymbol{\sigma}_b$) that can be utilized. Following this idea, once the inversion is performed at a coarser scale (first step), one still wants to infer how this result maps to a finer scale taking into account the finest component of the prior (second step). To do so, a projection operator Π is defined so as to map an emission field $\boldsymbol{\sigma} \in \mathbb{R}^N$ at fine scale (original emission inventory scale) to an emission field $\boldsymbol{\sigma}' \in \mathbb{R}^{N'}$ at a coarser scale. Π could be considered for instance, as the operator that averages fine cells into a coarser cell maintaining emission mass. Assume $\bar{\boldsymbol{\sigma}'}$ is the coarse scale estimate that results from the first step inversion.

We use maximum entropy on the mean because it will allow the use of primal and dual versions of a non-Gaussian formalism (Bocquet, 2005a). Define $\nu(\boldsymbol{\sigma})$ as the probability density prior on the emission field at the finer scale. Assume it can be split cell-wise ($\nu = \otimes_{k=1}^N \nu_k$), which means that the background errors are independent from one grid-cell to another (no covariances prescribed). Even if it is not a realistic assumption, we will assume it in order to obtain a first set of simplified equations. Then, the primal cost function that can help to define the emission field at fine scale knowing an estimation at a coarser scale reads

$$\mathcal{L}(\boldsymbol{\sigma}, \boldsymbol{\beta}, \boldsymbol{\lambda}) = \sum_{k=1}^N \hat{\nu}_k^*(\sigma_k) + \boldsymbol{\beta}^T (\bar{\boldsymbol{\mu}} - \mathbf{H}\boldsymbol{\sigma}) + \boldsymbol{\lambda}^T (\bar{\boldsymbol{\sigma}'} - \Pi\boldsymbol{\sigma}). \quad (12)$$

In this equation, the first constraint (second term on the right hand side) implemented by Lagrange multipliers $\boldsymbol{\beta} \in \mathbb{R}^d$ enforces that the finer emission field still satisfies the measurement equation $\bar{\boldsymbol{\mu}} = \mathbf{H}\boldsymbol{\sigma}$. For the sake of simplicity, we have not added observation errors but it can easily be done and will be included in the two following application cases. The second constraint (third term on the right hand side), implemented by Lagrange multipliers $\boldsymbol{\lambda} \in \mathbb{R}^{N'}$ enforces that the same mass obtained in the coarse inversion ($\bar{\boldsymbol{\sigma}'}$) is distributed into the fine emission field ($\boldsymbol{\sigma}$) block cell by block cell. This last constraint prevents the colocalization problem to happen.

The log-Laplace transform $\widehat{\nu}$ of ν is defined by

$$\widehat{\nu}(\boldsymbol{\omega}) = \sum_{\boldsymbol{\sigma}} \nu(\boldsymbol{\sigma}) \exp(\boldsymbol{\omega}^T \boldsymbol{\sigma}), \quad (13)$$

where the sum over $\boldsymbol{\sigma}$ is symbolical here and may be a real sum or an integral depending on the actual distributions. The Legendre–Fenchel transform $\widehat{\nu}^*$ of $\widehat{\nu}$ is defined by

$$\widehat{\nu}^*(\boldsymbol{\sigma}) = \sup_{\boldsymbol{\omega}} [\boldsymbol{\omega}^T \boldsymbol{\sigma} - \widehat{\nu}(\boldsymbol{\omega})]. \quad (14)$$

Since the constraints have been taken into account by the Lagrange multipliers, this cost function can be freely optimized on the $\boldsymbol{\sigma}$, $\boldsymbol{\beta}$ and $\boldsymbol{\lambda}$ vectors. The first step is to get rid of the $\boldsymbol{\sigma}$ set of variables. The optimal $\boldsymbol{\sigma}$, called $\boldsymbol{\sigma}^*$ from now on, should satisfy for all $k = 1, \dots, N$

$$\nabla_{\sigma_k} \widehat{\nu}_k^*(\sigma_k^*) - [\boldsymbol{\beta}^T \mathbf{H}]_k - [\boldsymbol{\lambda}^T \Pi]_k = 0. \quad (15)$$

Then the cost function becomes

$$\begin{aligned} \widehat{\mathcal{L}}(\boldsymbol{\beta}, \boldsymbol{\lambda}) &= \sum_{k=1}^{N'} \widehat{\nu}_k^*[\sigma_k^*(\boldsymbol{\beta}, \boldsymbol{\lambda})] \\ &\quad + \boldsymbol{\beta}^T [\bar{\boldsymbol{\mu}} - \mathbf{H} \boldsymbol{\sigma}^*(\boldsymbol{\beta}, \boldsymbol{\lambda})] + \boldsymbol{\lambda}^T [\bar{\boldsymbol{\sigma}}' - \Pi \boldsymbol{\sigma}^*(\boldsymbol{\beta}, \boldsymbol{\lambda})], \end{aligned} \quad (16)$$

with $\sigma_k^*(\boldsymbol{\beta}, \boldsymbol{\lambda}) = [\nabla_{\sigma_k} \widehat{\nu}_k^*]^{-1}([\boldsymbol{\beta}^T \mathbf{H}]_k + [\boldsymbol{\lambda}^T \Pi]_k)$.

By reciprocity of the Legendre–Fenchel transforms, $\widehat{\nu}(\boldsymbol{\omega}) = \sup_{\boldsymbol{\sigma}} (\boldsymbol{\sigma}^T \boldsymbol{\omega} - \widehat{\nu}_k^*(\boldsymbol{\sigma}))$, so that, if $\omega_k = \nabla_{\sigma_k} \widehat{\nu}_k^*(\sigma_k)$,

$$\widehat{\nu}_k(\omega_k) = \omega_k [\nabla_{\sigma_k} \widehat{\nu}_k^*]^{-1}(\omega_k) - \widehat{\nu}_k^*([\nabla_{\sigma_k} \widehat{\nu}_k^*]^{-1}(\omega_k)). \quad (17)$$

We set: $\omega_k = [\boldsymbol{\beta}^T \mathbf{H}]_k + [\boldsymbol{\lambda}^T \Pi]_k$. Therefore the dual cost function reads

$$\widehat{\mathcal{L}}(\boldsymbol{\beta}, \boldsymbol{\lambda}) = \boldsymbol{\beta}^T \bar{\boldsymbol{\mu}} + \boldsymbol{\lambda}^T \bar{\boldsymbol{\sigma}}' - \sum_{k=1}^N \widehat{\nu}_k([\boldsymbol{\beta}^T \mathbf{H}]_k + [\boldsymbol{\lambda}^T \Pi]_k). \quad (18)$$

Denote $\bar{\boldsymbol{\beta}}, \bar{\boldsymbol{\lambda}}$, the unique argument of the maximum of $\widehat{\mathcal{L}}(\boldsymbol{\beta}, \boldsymbol{\lambda})$. Then an estimator of the fine emission field is

$$\bar{\sigma}_k = (\nabla_{\omega_k} \widehat{\nu}_k) \left([\bar{\boldsymbol{\beta}}^T \mathbf{H}]_k + [\bar{\boldsymbol{\lambda}}^T \Pi]_k \right). \quad (19)$$

For ν we can typically choose a Gaussian law or a Poisson law.

4.1. Gaussian case

Here, it is assumed that the prior on the emission ν_{σ} is Gaussian: $\boldsymbol{\sigma} \sim \mathcal{N}(\boldsymbol{\sigma}_b, \mathbf{B})$ and that the fluxes can be correlated a priori and thus we do not make any splitting assumption. In addition, we will consider Gaussian observation errors distributed according to $\boldsymbol{\epsilon} \sim \mathcal{N}(\mathbf{0}, \mathbf{R})$, leading to the Gaussian prior ν_{ϵ} . The corresponding log-Laplace transforms are

$$\widehat{\nu}_{\sigma}(\boldsymbol{\chi}) = \frac{1}{2} \boldsymbol{\chi}^T \mathbf{B} \boldsymbol{\chi} \quad \widehat{\nu}_{\epsilon}(\boldsymbol{\delta}) = \frac{1}{2} \boldsymbol{\delta}^T \mathbf{R} \boldsymbol{\delta}, \quad (20)$$

and the related Legendre transforms of the latter are

$$\widehat{\nu}_{\sigma}^*(\boldsymbol{\sigma}) = \frac{1}{2} (\boldsymbol{\sigma} - \boldsymbol{\sigma}_b)^T \mathbf{B}^{-1} (\boldsymbol{\sigma} - \boldsymbol{\sigma}_b) \quad \widehat{\nu}_{\epsilon}^*(\boldsymbol{\epsilon}) = \frac{1}{2} \boldsymbol{\epsilon}^T \mathbf{R}^{-1} \boldsymbol{\epsilon}, \quad (21)$$

This leads to the dual cost function

$$\begin{aligned} \widehat{\mathcal{L}}(\boldsymbol{\beta}, \boldsymbol{\lambda}) &= \boldsymbol{\beta}^T (\bar{\boldsymbol{\mu}} - \mathbf{H} \boldsymbol{\sigma}_b) + \boldsymbol{\lambda}^T (\bar{\boldsymbol{\sigma}}' - \Pi \boldsymbol{\sigma}_b) - \frac{1}{2} \boldsymbol{\beta}^T \mathbf{R} \boldsymbol{\beta} \\ &\quad - \frac{1}{2} (\boldsymbol{\beta}^T \mathbf{H} + \boldsymbol{\lambda}^T \Pi) \mathbf{B} (\mathbf{H}^T \boldsymbol{\beta} + \Pi^T \boldsymbol{\lambda}), \end{aligned} \quad (22)$$

and to the estimators

$$\bar{\boldsymbol{\sigma}} = \boldsymbol{\sigma}_b + \mathbf{B} \mathbf{H}^T \bar{\boldsymbol{\beta}} + \mathbf{B} \Pi^T \bar{\boldsymbol{\lambda}} \quad \bar{\boldsymbol{\epsilon}} = \mathbf{R} \bar{\boldsymbol{\beta}}. \quad (23)$$

Analytically the $\bar{\boldsymbol{\beta}}$ and $\bar{\boldsymbol{\lambda}}$ are solutions of

$$\begin{pmatrix} \mathbf{R} + \mathbf{H} \mathbf{B} \mathbf{H}^T & \mathbf{H} \mathbf{B} \Pi^T \\ \Pi \mathbf{B} \mathbf{H}^T & \Pi \mathbf{B} \Pi^T \end{pmatrix} \begin{pmatrix} \bar{\boldsymbol{\beta}} \\ \bar{\boldsymbol{\lambda}} \end{pmatrix} = \begin{pmatrix} \bar{\boldsymbol{\mu}} - \mathbf{H} \boldsymbol{\sigma}_b \\ \bar{\boldsymbol{\sigma}}' - \Pi \boldsymbol{\sigma}_b \end{pmatrix}, \quad (24)$$

which leads to the explicit solutions for the Lagrange parameters

$$\begin{aligned} \bar{\boldsymbol{\beta}} &= [\mathbf{R} + \mathbf{H} (\mathbf{B} - \mathbf{B} \Pi^T (\Pi \mathbf{B} \Pi^T)^{-1} \Pi \mathbf{B}) \mathbf{H}^T]^{-1} \\ &\quad \times [\bar{\boldsymbol{\mu}} - \mathbf{H} \boldsymbol{\sigma}_b - \mathbf{H} \mathbf{B} \mathbf{H}^T (\Pi \mathbf{B} \Pi^T)^{-1} (\bar{\boldsymbol{\sigma}}' - \Pi \boldsymbol{\sigma}_b)], \\ \bar{\boldsymbol{\lambda}} &= [\Pi (\mathbf{B} - \mathbf{B} \mathbf{H}^T (\mathbf{R} + \mathbf{H} \mathbf{B} \mathbf{H}^T)^{-1} \mathbf{H} \mathbf{B}) \Pi^T]^{-1} \\ &\quad \times [\bar{\boldsymbol{\sigma}}' - \Pi \boldsymbol{\sigma}_b - \Pi \mathbf{B} \mathbf{H}^T (\mathbf{R} + \mathbf{H} \mathbf{B} \mathbf{H}^T)^{-1} (\bar{\boldsymbol{\mu}} - \mathbf{H} \boldsymbol{\sigma}_b)]. \end{aligned} \quad (25)$$

Then, by replacing $\bar{\boldsymbol{\beta}}$ and $\bar{\boldsymbol{\lambda}}$ into eq. (23) one obtains the explicit solution for the estimator of the source

$$\begin{aligned} \bar{\boldsymbol{\sigma}} &= \boldsymbol{\sigma}_b + \mathbf{B} \mathbf{H}^T (\mathbf{R} + \mathbf{H} \mathbf{B} \mathbf{H}^T)^{-1} (\bar{\boldsymbol{\mu}} - \mathbf{H} \boldsymbol{\sigma}_b) \\ &\quad + [\mathbf{B} - \mathbf{B} \mathbf{H}^T (\mathbf{R} + \mathbf{H} \mathbf{B} \mathbf{H}^T)^{-1} \mathbf{H} \mathbf{B}] \Pi^T \\ &\quad \times [\Pi (\mathbf{B} - \mathbf{B} \mathbf{H}^T (\mathbf{R} + \mathbf{H} \mathbf{B} \mathbf{H}^T)^{-1} \mathbf{H} \mathbf{B}) \Pi^T]^{-1} \\ &\quad \times [\bar{\boldsymbol{\sigma}}' - \Pi (\boldsymbol{\sigma}_b + \mathbf{B} \mathbf{H}^T (\mathbf{R} + \mathbf{H} \mathbf{B} \mathbf{H}^T)^{-1} (\bar{\boldsymbol{\mu}} - \mathbf{H} \boldsymbol{\sigma}_b))]. \end{aligned} \quad (26)$$

It can be noted that the first two terms in the equation corresponds exactly to the terms in eq. (2) (without the introduction of the α factor), then eq. (26) can be interpreted as the same solution as the base methodology plus a correction term to address the constraint coming from the coarse a posteriori estimate. It can also be shown that eq. (2) is a specific case of eq. (26) when the projection operator Π is the identity (no change in spatial resolution, $\bar{\boldsymbol{\sigma}}' = \bar{\boldsymbol{\sigma}}$). Figures 1e and f illustrate the two-scale method for the Gaussian case.

4.2. Poisson case

Assuming a Poisson distribution for the emission (see Section 2), the corresponding MEM primal cost function is (Bocquet, 2005a)

$$\begin{aligned} \mathcal{L}(\boldsymbol{\sigma}, \boldsymbol{\epsilon}, \boldsymbol{\beta}, \boldsymbol{\lambda}) &= \sum_{k=1}^N \frac{1}{m_k} \left(\sigma_k \ln \frac{\sigma_k}{\sigma_{b,k}} + \sigma_{b,k} - \sigma_k \right) + \frac{1}{2} \boldsymbol{\epsilon}^T \mathbf{R}^{-1} \boldsymbol{\epsilon} \\ &\quad + \boldsymbol{\beta}^T (\bar{\boldsymbol{\mu}} - \mathbf{H} \boldsymbol{\sigma} - \boldsymbol{\epsilon}) + \boldsymbol{\lambda}^T (\bar{\boldsymbol{\sigma}}' - \Pi \boldsymbol{\sigma}). \end{aligned} \quad (27)$$

Then the dual cost function to be maximized is:

$$\widehat{\mathcal{L}}(\boldsymbol{\beta}, \boldsymbol{\lambda}) = \boldsymbol{\beta}^T \bar{\boldsymbol{\mu}} + \boldsymbol{\lambda}^T \bar{\boldsymbol{\sigma}}' - \frac{1}{2} \boldsymbol{\beta}^T \mathbf{R} \boldsymbol{\beta} - \sum_{k=1}^N \frac{\sigma_{b,k}}{m_k} \{ \exp(m_k [\boldsymbol{\beta}^T \mathbf{H}]_k + m_k [\boldsymbol{\lambda}^T \boldsymbol{\Pi}]_k) - 1 \}, \quad (28)$$

and the estimators are:

$$\bar{\sigma}_k = \sigma_{b,k} \exp(m_k [\bar{\boldsymbol{\beta}}^T \mathbf{H}]_k + m_k [\bar{\boldsymbol{\lambda}}^T \boldsymbol{\Pi}]_k). \quad (29)$$

and

$$\bar{\boldsymbol{\epsilon}} = \mathbf{R} \bar{\boldsymbol{\beta}}. \quad (30)$$

Figure 2c illustrates the performance of the two-scale method for the non-Gaussian (Poisson) case.

5. Intercomparison test bed

The inversion is performed to improve a carbon monoxide (CO) emission inventory over the city of Santiago de Chile. Emissions correspond to the official 2002 emission inventory provided by the Chilean Environmental Agency (<http://www.conama.cl>) and are produced by a bottom-up approach (Corvalán and Osses, 2002). It has a spatial resolution of 2 km and a temporal resolution of one hour for a representative day of the week. Carbon monoxide observations correspond to Santiago air quality monitoring network (MACAM2 network, <http://www.asrm.cl/>). MM5 meteorological fields are used to feed the dispersion model (Grell et al., 1995), using 2 km horizontal resolution and 31 vertical levels as used for the same region in previous studies (Schmitz, 2005; Jorquera and Castro, 2010). Air quality simulations are performed using the Polyphemus platform (Mallet et al., 2007) for 2 km horizontal resolution and twelve vertical levels up to 6 km. The model horizontal domain is centred around Santiago (33.5S, 70.5W) and encompasses an area of 140 km × 126 km. Polyphemus is run in tracer mode, since CO at the city scale can be assumed so (residence time in the basin is not more than a couple of days and CO life time of order of months). CO boundary conditions are set to zero since emission rates in Santiago are very high and no relevant sources are found up-wind from the city. The sensitivity matrix (\mathbf{H}) is computed by running Polyphemus in adjoint mode. For a tracer, this is done by shifting winds in the opposite direction and running backwards in time (e.g. Davoine and Bocquet, 2007). Further details on model settings, validation of the meteorological and air quality models and assumptions made can be found in Saide et al. (2009).

The inversion experiment consists in improving the CO emission inventory using seven ground monitoring stations for a summer period going from January 15 to January 25, that was shown to be representative of the whole summer period for this region (Saide et al., 2009). Methods described in Sections 2–4 are used

to obtain improved emissions which are compared in order to obtain differences, similarities and advantages or disadvantages in between them. The method of coarsening the emission inventory (Section 3.1) is not applied since the objective is to compare a posteriori emissions under the same resolution.

The covariance matrices (when existent) are considered diagonal with a value of one and the weight of each term in the functional is controlled by the α parameter. Parameter α is considered the same for all methodologies (Gaussian and non-Gaussian) and is the same as in the previous work (1.4×10^7 , Saide et al., 2009), where it was estimated with the L-curve approach (Hansen and O’Leary, 1993; Davoine and Bocquet, 2007) given the lack of information on the uncertainty of the model and parameter errors. For the Poisson case, where there is no matrix \mathbf{B} , parameters \mathbf{m} and $\boldsymbol{\theta}$ need to be adjusted. For this case study it was found that better results are obtained when setting \mathbf{m} constant and computing $\boldsymbol{\theta}$ as $\theta_k = \sigma_{b,k}/m_k$ (distribution mean equation, see Section 3.4). Since \mathbf{m} is multiplying the emission term in eq. (5), it is assumed to be equal to 1 and the weighting of the terms is handled only by α .

For the case of adding spatial correlations (Section 3.2) eq. (7) is used for computing the \mathbf{B} matrix with $l = 2.5$ cells, which is the optimum found when doing synthetic inversion for this specific case (Saide et al., 2009). The Laplacian is chosen for the method that constrains the spatial derivatives (Section 3.3). Operator \mathbf{D} is computed using the centred differences. The γ parameter for this approach is chosen testing several values until no colocalization is found ($\gamma = 0.1\alpha$).

For the two-step two-scale method (Section 4) the emission resolution has to be decreased, and this can be done by several ways (e.g. Bocquet, 2009). In this study, the two scales are characterized by two regular grids with different resolutions. If a coarse grid-cell contains n fine grid-cells, then there exists n different ways to place the coarse grid with respect to the finest grid. Since there is no a priori preference for any of these choices, the inversion is performed for all the cases and the results are averaged to obtain a representative inventory for the chosen resolution. The same parameters are used for inversions at different scales (\mathbf{R} , \mathbf{B} , \mathbf{m} , α). For the case study, the resolution is coarsened in cells that group 3×3 grid-cells obtaining nine possible shifts of the grid.

As mentioned in Section 2, numerical minimization is required to obtain the optimized emissions for the non-Gaussian scheme. Also, to obtain positive emissions in the Gaussian case, eq. (1) is minimized using additional positivity constraints. The minimization routine used in both cases is the L-BFGS-B code (Zhu et al., 1997).

5.1. Colocalization index

To provide a quantitative measure of the colocalization problem to be able to compare the behaviour of the different methods,

the following colocalization index (CLI) index was constructed

$$CLI = \sum_{i=1}^{N_{est}} \left| \frac{\left(\sum_{j=1}^4 \delta\sigma_{i,j} \right) - 4\delta\sigma_i}{\frac{1}{5} \left[\left(\sum_{j=1}^4 \delta\sigma_{i,j} \right) + \delta\sigma_i \right]} \right|, \quad (30)$$

where $\delta\sigma = \sigma_b - \bar{\sigma}$ (difference between guess and analysis, plotted on Fig. 1), i loops on N_{est} total number of stations and j loops on the four spatial neighbours of each i station spot. The numerator in the right hand side of the equation is the numerical Laplacian estimator on the observation spots, which gives an estimation of the smoothness of the $\delta\sigma$ field on the observation places. The higher the Laplacian the less smooth is the solution, meaning that colocalization is present. The denominator on eq.(31) is the average of the $\delta\sigma_i$ and its neighbours. It is added in order to locally normalize the index and make it comparable in between methods (different methods generate different magnitude on $\delta\sigma$).

6. Results and discussion

The following results correspond to inversions that use only real observations. Results are classified by the use of Gaussian or non-Gaussian estimates.

6.1. Gaussian estimates

The inversions performed are: one without accounting for the colocalization problem (base run); one for each method presented in Section 3 (but coarsening spatial resolution); and one with the two-step inversion. Table 1 presents a summary of the following discussion and the CLI for each method.

Figure 1 shows the spatial distribution of the difference between the background and the optimized emissions. All methods agree in a decrease of emissions in the centre-western area and in an increase of emissions in the eastern area. However, several differences can be found depending on the method used. The base methodology (Fig. 1a) shows the colocalization problem with the highest CLI values (Table 1), where modifications are done preferably over the colocalized emissions and the rest remains almost the same.

Figures 1b and c present results from the methods that add correlation and the one that constrains the derivatives, respectively, and show really similar behaviour between them. Both solve the colocalization problem with smooth modifications on the emissions, showing low values on their CLI (Table 1). The derivative method shows the minimum CLI of all methods, which is reasonable since this method minimizes the Laplacian, which is how the CLI is constructed. It is interesting to note that changes in the emissions for both methods seem to be spatially isotropic (similar in all directions, see contours in Figs 1b and c). Another aspect that is notable is that there are regions far from the observations that are not changed at all (short influence range).

The results from the method that uses a weighting function over the B matrix is presented in Fig. 1d. The method is able to solve colocalization (CLI similar to the spatial correlation method, Table 1), but the pattern of modification is more spatially anisotropic, contrary to the previous methods (see contours in this figure). No covariance is built into the method, which permits changes that may not be as smooth as the previous results. Also, this method changes emissions far from the observation (long influence range). We cannot argue if this is favourable or unfavourable, since it depends mainly in where we think that emissions should be changed.

Finally, Figs 1e and f present the results from the two-step two-scale inversion, where the first one is the result of a single run of the method and the second the average of all possible grid positions for a determined coarse grid size, as mentioned in Section 5. The single run optimized emissions present marked influence from the coarse grid, but when averaging smoothness is achieved, the artefact disappears. Comparing to previous results, Fig. 1f presents a mixed behaviour. The modifications are not as isotropic as Figs 1b and c, but they do have a radius of influence where emissions at a certain distance from observations are not modified. The colocalization problem is solved, but not completely (CLI index is slightly higher than the spatial correlation or the weighting factor method, Table 1), as the maximum changes are at the colocalized spots. This method has an additional limitation of the necessity to chose an arbitrary total amount of cells to group, which is similar to choosing a radius of influence for the correlation methods. However, this could be solved by

Table 1. Summary of features of different methods for solving colocalization. CLI stands for the colocalization index (see Section 5.1). See text in section 6 for further explanations

Method	Bayesian	Non-Gaussian applicability	Isotropy/anisotropy	Influence range	CLI
No method (base)	Yes	Yes	–	Very short	56.5
Spatial correlations	No	No	i	Short	10.0
Constraint derivatives	No	No	i	Short	5.2
Weight factor to B	No	Yes	a	Long	10.6
Two-scale	Yes	Yes	i/a	Short	11.2

using optimal choices criteria (Bocquet, 2009). Another limitation is the necessity to do the two steps of the inversion and then repeat it several times to get a smooth solution which would cost more computational time than other methods.

Figure 3 shows the emissions temporal profiles in a selection of grid-cells for some of the methods. When comparing the results between Gaussian methods (Fig. 3, left column) many conclusions can be drawn. Since the Gaussian functional assumes an additive error in the mismatch of emissions, the difference between the optimized emissions and the guess tends to be similar for each parameter (Fig. 1). However, when looking at the plot with higher emissions (Fig. 3a3) it seems that the modifications are low compared with the plots with lower emissions (Fig. 3a1 and a4), but this is only due to use of different scales in the plots. As mentioned in Saide et al. (2009), during the morning, when the boundary layer is shallow and the winds are weak, emissions have higher sensitivity to the observations than the ones in the afternoon when the boundary layer deepens and winds are strong. This results in higher modifications of morning emissions and lower modifications in the afternoon ones when not applying a weighting factor. This behaviour can be found in Fig. 3a1 (colocated spot) and Fig. 3a2 (non-colocated spot).

Methods that add covariance to the background errors are prone to aggregation errors, although in a less obvious way than the methods based on straight coarsening of grid-cells. For instance, emissions in between the centre of the city and the north-east edge mainly affect the north-eastern station. When adding statistical correlations between emission flux errors (Figs 1b and c) these emission go down as they are correlated with emissions in the centre of the city. However, when looking at results where no additional correlations are enforced (Fig. 1d) they have the opposite behaviour, which helps to improve the fit in the north-east station. For details, Fig. 3a4 shows this behaviour for the afternoon hours. Other example of this behaviour is the pattern of modification of emissions colocated with the northwestern station. The misfits between observations and model in this station occur only during morning hours. However, correlation with neighbouring cells allows emission changes during the afternoon hours (line A in Fig. 3a1).

6.2. Poisson estimates

As mentioned in Section 3.2, methods that result in using a non-diagonal \mathbf{B} matrix cannot be used with the Poisson distribution. Hence results are presented for the base case, with weighting factor and for the two-step inversion. Fig. 2 shows the relative difference (difference divided by the first guess σ_b) between the guess and the optimized emissions. The behaviour is similar to the Gaussian estimates, but with relative differences, since the error is assumed multiplicative for this distribution (Section 2). This is desirable if the error in emissions are thought to be multiplicative. The two-step two-scale approach and the weighting factor method results show that the colocalization problem can

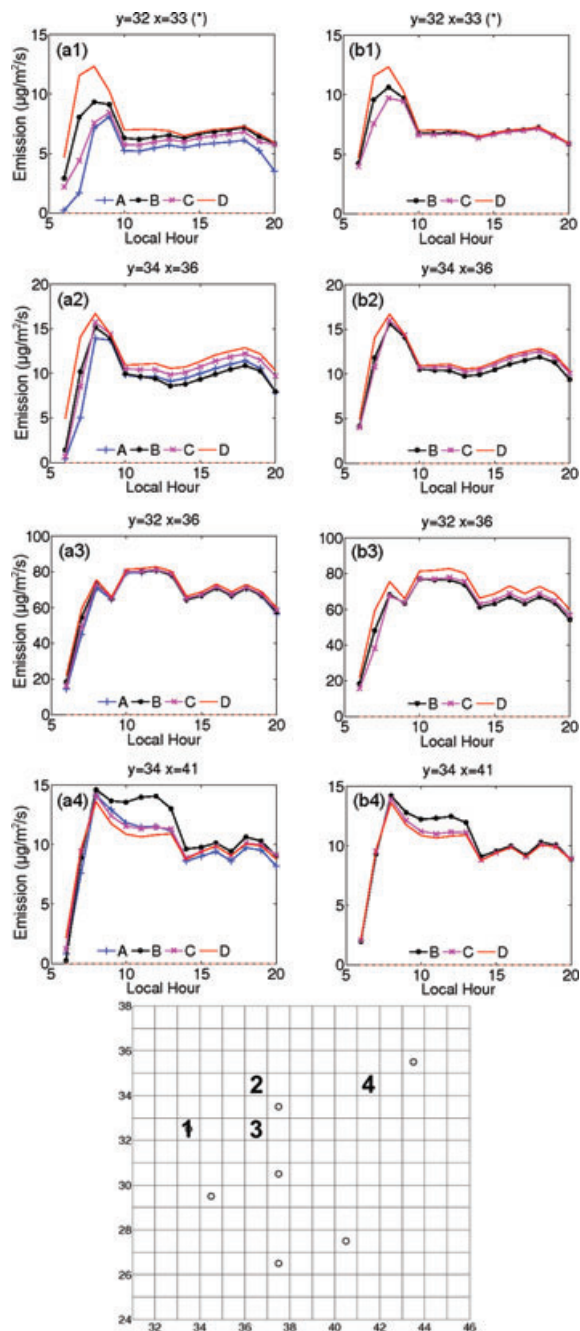


Fig. 3. Emission time series for a selection of grid-cells (1–4). Left (a) and right (b) columns present Gaussian and Poisson results respectively. Line A are the result using a non-diagonal \mathbf{B} matrix using a radius of influence $r = 2.5$ cells, line B using a weighting factor over \mathbf{B} matrix, line C applying the two-step method for 3×3 cells and averaging all possible results, and line D the initial emissions (guess). Bottom figure shows the grid-cell selection. Cell 1 is an emission colocated with an observation, cell 3 is the place where maximum emissions occur and cells 2 and 4 are cells selected to explain different behaviour of the inverse modelling methods. Numbers in the axis represent the cell number in the model grid. See the text for details. Units are in $\mu\text{g m}^{-2} \text{s}^{-1}$.

be solved even though no covariance matrices are specified in the methodology.

Figure 3 right column (b) shows the emission time series for Poisson estimates. Now, since changes are multiplicative, the change for high values first guess emissions is higher than in the Gaussian case (compare Fig. 3b3, where changes are close to $10 \mu\text{g m}^{-2} \text{s}^{-1}$, to Fig. 3a3 where changes are less notable). Also, due to the multiplicative changes, the low value emissions remain close to the background, which is not true for most of the Gaussian estimates that tend to approach 0 (compare first hours of the time series on the left and right columns of Fig. 3).

7. Summary and conclusions

Inverse modelling for improving spatially and temporary resolved emission inventories is being more extensively used by the atmospheric chemistry community. When performing the inversion one of the problems found is due to the fact that the observations lie in the space and time manifold of the emissions, resulting in spurious corrections to the emissions (colocalization problem).

Several methods have been used to solve this problem that consist in making modifications to a base inverse modelling technique (essentially 4D-Var). Several of these methods were reviewed and a new strictly Bayesian method was presented. They were all tested and qualitatively and quantitatively compared in an experiment that consists in improving a CO emission inventory at the scale of Santiago de Chile. Moreover, non-Gaussian (e.g. Poisson) probability density distributions of the background emission fluxes were introduced for this application and the different methods to solve colocalization were applied when possible.

Methods for solving colocalization that add correlations or constraint derivatives are similar in construction and showed similar results like an isotropic pattern of modification, smoothness and a short range radius of modification. But these methods are likely to be impacted by smooth aggregation errors: emission fluxes may be changed because of correlations with other emission flux errors and do not improve any misfit between model and observations. The method that considers the use of a weighting factor over the \mathbf{B} matrix showed a more anisotropic spatial pattern of modification and less smooth results. Also, it has the tendency to modify emissions that can be far from the observation locations, which can be preferable or not depending on the application.

A two-step two-scale fully Bayesian method was developed and applied with satisfactory results, solved the colocalization almost completely and showed a mixed behaviour compared to the previous methods used. It was shown to be a method independent of the emissions prior density distribution chosen by testing it successfully with Gaussian and Poisson priors. However, the method has some limitations like arbitrary estimation

of parameters and additional computational time required to execute it.

Results using Poisson estimates showed similar modification patterns as the Gaussian approach for all methods tested but consider multiplicative corrections for emissions, which is the most commonly used assumption for emissions inventories. This feature is reflected by avoiding the decrease in small emissions to zero values as in the Gaussian tests, and allowing high values to vary far from their original estimate.

Future research should be focused on finding the most suitable method depending on the application, based on the a priori knowledge on the emissions being optimized.

8. Acknowledgments

This work is being carried out with the aid of a grant from the Inter-American Institute for Global Change Research (IAI) CRN II 2017 which is supported by the US National Science Foundation (Grant GEO-0452325), the STIC-AMSUD project 'Air-quality prediction with data assimilation in Argentina and Chile' and Fulbright-CONICYT scholarship number 15093810. This paper is also a contribution to the MSDAG project supported by the *Agence Nationale de la Recherche*, grant ANR-08-SYSC-014. We appreciate the valuable comments from G.R. Carmichael and from two anonymous reviewers.

References

- Bocquet, M. 2009. Towards optimal choices of control space representation for geophysical data assimilation. *Mon. Wea. Rev.* **137**, 2331–2348.
- Bocquet, M. 2005a. Reconstruction of an atmospheric tracer source using the principle of maximum entropy. I: theory. *Q. J. R. Meteorol. Soc.* **131**, 2191–2208.
- Bocquet, M. 2005b. Grid resolution dependence in the reconstruction of an atmospheric tracer source. *Nonlin. Process. Geophys.* **12**, 219–233.
- Bocquet, M., Pires, C.A. and Wu, L. 2010. Beyond Gaussian statistical modelling in geophysical data assimilation. *Mon. Wea. Rev.* **138**, 2997–3023.
- Bousquet, P., Ciais, P., Peylin, P., Ramonet, M. and Monfray, P. 1999. Inverse modelling of annual atmospheric CO₂ sources and sink: 1. Method and control inversion. *J. Geophys. Res.* **104**(D21), 26161–26178.
- Carmichael, G.R., Sandu, A., Chai, T., Daescu, D., Constantinescu, E. and Tang, Y. 2008. Predicting air quality: improvements through advanced methods to integrate models and measurements. *J. Comput. Phys.* **227**, 3540–3571.
- Chai, T., Carmichael, G.R., Tang, Y., Sandu, A., Heckel, A., Richter, A. and co-authors. 2009. Regional NO_x emission inversion through a four-dimensional variational approach using SCIAMACHY tropospheric NO₂ column observations. *Atmos. Environ.* **43**, 5046–5055.
- Chang, M.E., Hartley, D.E., Cardelino, C., Haas-Laursen, D. and Chang, W.-L. 1997. On using inverse methods for resolving

- emissions with large spatial inhomogeneities. *J. Geophys. Res.* **102**(D13), 16023–16036.
- Corvalán, R. and Osses, M. 2002. Hot emission model for mobile sources: application to the metropolitan region of the city of Santiago-Chile. *J. Air Waste Manage. Assoc.* **52**, 167–174.
- Davoine, X. and Bocquet, M. 2007. Inverse modelling-based reconstruction of the Chernobyl source term available for long-range transport. *Atmos. Chem. Phys.* **7**, 1549–1564.
- Dubovik, O., Lapyonok, T., Kaufman, Y.J., Chin, M., Ginoux, P. and co-authors. 2008. Retrieving global aerosol sources from satellites using inverse modelling. *Atmos. Chem. Phys.* **8**, 209–250.
- Elbern, H., Strunk, A., Schmidt, H. and Talagrand, O. 2007. Emission rate and chemical state estimation by 4-dimensional variational inversion. *Atmos. Chem. Phys.* **7**, 3749–3769.
- Gaspari, G. and Cohn, S.E. 1999. Construction of correlation functions in two and three dimensions. *Q. J. R. Meteorol. Soc.* **125**, 723–757.
- Golub, G. and Van Loan, C.F. 1996. *Matrix Computation*, 3rd Edition John Hopkins University Press, Baltimore and London.
- Grell, G.A., Dudhia, J. and Stauffer, D.R. 1995. A description of the fifth-generation Penn State/NCAR mesoscale model (MM5). NCAR Technical Note, NCAR/TN–398+STR.
- Hansen, P.C. and O’Leary, D.P. 1993. The use of the L-curve in the regularization of discrete illposed problems. *SIAM J. Sci. Comput.* **14**(6), 1487–1503.
- Henze, D.K., Seinfeld, J.H., Shindell and D.T. 2009. Inverse modelling and mapping US air quality influences of inorganic PM_{2.5} precursor emissions using the adjoint of GEOS-Chem. *Atmos. Chem. Phys.* **9**, 5877–5903.
- Hoelzemann, J., Elbern, H. and Ebel, A. 2001. PSAS and 4D-var data assimilation for chemical state analysis by urban and rural observation sites. *Phys. Chem. Earth* **26**, 807–812.
- Hoelzemann, J.J., Longo, K.M., Fonseca, R.M., do Rosário, N.M.E., Elbern, H. and co-authors. 2009. Regional representativity of AERONET observation sites during the biomass burning season in South America determined by correlation studies with MODIS Aerosol Optical Depth. *J. Geophys. Res.* **114**, D13301.
- Houweling, S., Bréon, F.-M., Aben, I., Rödenbeck, C., Gloor, M. and co-authors. 2004. Inverse modelling of CO₂ sources and sinks using satellite data: a synthetic inter-comparison of measurement techniques and their performance as a function of space and time. *Atmos. Chem. Phys.* **4**, 523–538.
- Issartel, J.P. 2003. Rebuilding sources of linear tracers after atmospheric concentration measurements. *Atmos. Chem. Phys.* **3**, 2111–2125.
- Jorquera, H. and Castro, J. 2010. Analysis of urban pollution episodes by inverse modelling. *Atmos. Environ.* **44**, 42–54.
- Kalnay, E. 2003. *Atmospheric Modeling, Data Assimilation and Predictability*. Cambridge Univ. Press, Cambridge.
- Kaminski, T., Heimann, M. and Giering, R. 1999. A coarse grid three-dimensional global inverse model of the atmospheric transport: 2. Inversion of the transport of CO₂ in the 1980s. *J. Geophys. Res.* **104**(D15), 18555–18581.
- Kopacz, M., Jacob, D.J., Fisher, J.A., Logan, J.A., Zhang, L. and co-authors. 2010. Global estimates of CO sources with high resolution by adjoint inversion of multiple satellite datasets (MOPITT, AIRS, SCIAMACHY, TES). *Atmos. Chem. Phys.* **10**, 855–876.
- Lindley, S.J., Conlan, D.E., Raper, D.W. and Watson, A.F.R. 2000. Uncertainties in the compilation of spatially resolved emission inventories—evidence from a comparative study. *Atmos. Environ.* **34**, 375–388.
- Mallet, V., Quélo, D., Sportisse, B., Ahmed de Biasi, M., Debry, É. and co-authors. 2007. Technical note: the air quality modeling system Polyphemus. *Atmos. Chem. Phys.* **7**, 5479–5487.
- Mulholland, M. and Seinfeld, J.H. 1995. Inverse air pollution modeling of urban-scale carbon monoxide emissions. *Atmos. Environ.* **29**, 497–516.
- Peters, W., Miller, J.B., Whitaker, J., Denning, A.S., Hirsch, A. and co-authors. 2005. An ensemble data assimilation system to estimate CO₂ surface fluxes from atmospheric trace gas observations. *J. Geophys. Res.* **110**, D24304.
- Pétron, G., Granier, C., Khattatov, B., Lamarque, J.F., Yudin, Muller and co-authors. 2002. Inverse modeling of carbon monoxide surface emissions using CMDL network observations. *J. Geophys. Res.* **107**(D24), 4761.
- Ródenbeck, C., Houweling, S., Gloor, M. and Heimann, M. 2003. CO₂ flux history 1982–2001 inferred from atmospheric data using a global inversion of atmospheric transport. *Atmos. Chem. Phys.* **3**, 1919–1964.
- Saide, P., Osses, A., Gallardo, L. and Osses, M. 2009. Adjoint inverse modeling of a CO emission inventory at the city scale: Santiago de Chile’s case. *Atmos. Chem. Phys. Discuss.* **9**, 6325–6361.
- Sandu, A., Liao, W., Carmichael, G.R., Henze, D.K. and Seinfeld, J.H. 2005. Inverse modeling of aerosol dynamics using adjoints: theoretical and numerical considerations. *Aerosol Sci. Tech.* **39**, 677–694.
- Schmitz, R. 2005. Modelling of air pollution dispersion in Santiago de Chile. *Atmos. Environ.* **39**, 2035–2047.
- Seibert, P. 2000. Inverse modelling of sulfur emissions in Europe based on trajectories. In: *Inverse Methods in Global Biogeochemical Cycles* (eds P. Kasibhatla, M. Heimann, P. Rayner, N. Mahowald, R.G. Prinn and D.E. Hartley), AGU Geophysical Monograph Series, 147–154.
- Twomey, S. 1977. *Introduction to the Mathematics of Inversion in Remote Sensing and Indirect Measurements*, Elsevier Scientific Publishing Company, New York.
- Zhu, C., Byrd, R.H., Lu, P. and Nocedal, J. 1997. Algorithm 778: L-BFGS-B: Fortran subroutines for large-scale bound constrained-optimization. *ACM T. Math. Software* **23**, 550–560.