https://doi.org/10.5194/gmd-2020-407
Preprint. Discussion started: 18 December 2020

Geoscientific Model Development Discussions

The Community Inversion Framework v1.0: a unified system for atmospheric inversion studies

Antoine Berchet[1,*], Espen Sollum[2], Rona L. Thompson[2], Isabelle Pison[1], Joël Thanwerdas[1], Grégoire Broquet[1], Frédéric Chevallier[1], Tuula Aalto[3], Peter Bergamaschi[4], Dominik Brunner[5], Richard Engelen[6], Audrey Fortems-Cheiney[1], Christoph Gerbig[7], Christine Groot Zwaaftink[2], Jean-Matthieu Haussaire[5], Stephan Henne[5], Sander Houweling[8], Ute Kartens[9], Werner L. Kutsch[10], Ingrid T. Luijkx[11], Guillaume Monteil[9], Paul I. Palmer[12], Jacob C. A. van Peet[8], Wouter Peters[11,13], Philippe Peylin[1], Elise Potier[1], Christian Rödenbeck[7], Marielle Saunois[1], Marko Scholze[9], Aki Tsuruta[3] and Yuanhong Zhao[1]

[1]Laboratoire des Sciences du Climat et de l'Environnement, CEA-CNRS-UVSQ, Gif-sur-Yvette, France
[2]Norwegian Institute for Air Research (NILU), Kjeller, Norway
[3]Finnish Meteorological Institute (FMI), Helsinki, Finland
[4]European Commission Joint Research Centre, Ispra (Va), Italy
[5]Swiss Federal Laboratories for Materials Science and Technology (Empa), Dübendorf, Switzerland
[6]European Centre for Medium-Range Weather Forecasts, Reading, RG2 9AX, UK
[7]Max Planck Institute for Biogeochemistry, Jena, Germany
[8]Vrije Universiteit Amsterdam, Department of Earth Sciences, Earth and Climate Cluster, Amsterdam, the Netherlands
[9]Dep. of Physical Geography and Ecosystem Science, Lund University, Sweden
[10]Integrated Carbon Observation System (ICOS-ERIC), Helsinki, Finland
[11]Meteorology and Air Quality Group, Wageningen University and Research, Wageningen, the Netherlands
[12]School of GeoSciences, University of Edinburgh, Edinburgh, EH9 3FF, UK
[13]Centre for Isotope Research, University of Groningen, Groningen, the Netherlands

Correspondence: antoine.berchet@lsce.ipsl.fr

Abstract. Atmospheric inversion approaches are expected to play a critical role in future observation-based monitoring systems for surface greenhouse gas (GHG) fluxes. In the past decade, the research community has developed various inversion softwares, mainly using variational or ensemble Bayesian optimization methods, with various assumptions on uncertainty structures and prior information and with various atmospheric chemistry-transport models. Each of them can assimilate some or all of the available observation streams for its domain area of interest: flask samples, in-situ measurements or satellite observations. Although referenced in peer-reviewed publications and usually accessible across the research community, most systems are not at the level of transparency, flexibility and accessibility needed to provide the scientific community and policy makers with a comprehensive and robust view of the uncertainties associated with the inverse estimation of GHG fluxes. Furthermore, their development, usually carried out by individual research institutes, may in the future not keep pace with the increasing scientific needs and technical possibilities. We present here a Community Inversion Framework (CIF) to help rationalize development efforts and leverage

https://doi.org/10.5194/gmd-2020-407
Preprint. Discussion started: 18 December 2020


Geoscientific Model Development Discussions

the strengths of individual inversion systems into a comprehensive framework. The CIF is primarily a programming protocol to allow various inversion bricks to be exchanged among researchers. In practice, the ensemble of bricks makes a flexible, transparent and open-source python-based tool to estimate the fluxes of various GHGs both at global and regional scales. It will allow running different atmospheric transport models, different observation streams and different data assimilation approaches. This adaptability will allow a comprehensively assessment of uncertainty in a fully consistent framework. We present here the main structure and functionalities of the system, and demonstrate how it operates in a simple academic case.

1 Introduction

The role of greenhouse gases (GHGs) in climate change has motivated an exceptional effort over the last couple of decades to densify the observations of GHGs around the world (Ciais et al., 2014): from the ground, e.g., with the European Integrated Carbon Observation System (ICOS, https://www.icos-cp.eu/), from mobile platforms (e.g., from aircrafts, or balloons equipped with Aircore sampling; Filges et al., 2016; Karion et al., 2010), and from space (e.g., Crisp et al., 2018; Janssens-Maenhout et al., 2020), despite occasional budgetary difficulties (Houweling et al., 2012). These observations quantify the effect of exchange between the surface and the atmosphere on GHG concentrations (e.g., Ramonet et al., 2020) and can thus be used to determine the surface fluxes of GHGs through the inversion of atmospheric chemistry and transport (e.g., Peylin et al., 2013; Houweling et al., 2017). Alongside improved observation capabilities, national and international initiatives pave the way towards an operational use of atmospheric inversions to support emissions reporting to the United Nations Framework Convention on Climate Change (UNFCCC; e.g., Say et al., 2016; Henne et al., 2016; Bergamaschi et al., 2018a; Janssens-Maenhout et al., 2020, or the EU projects CHE – CO_2 Human emissions; che-project.eu – or VERIFY – verify.lsce.ipsl.fr).

In the past, research groups have developed various atmospheric inversion systems based on different techniques and atmospheric transport models, targeting specific trace gases or types of observations, and at various spatial and temporal scales, according to the particular scientific objectives of the study. All these systems have their own strengths and weaknesses and help explore the range of systematic uncertainty in the surface to atmosphere fluxes. Inter-comparison exercises are regularly conducted to assess the strengths and weaknesses of various inversion systems (e.g., Locatelli et al., 2013; Monteil et al., 2019; Bergamaschi et al., 2018b; Gurney et al., 2003; Saunois et al., 2020; Babenhauserheide et al., 2015; Peylin et al., 2013; Crowell et al., 2019; Schuh et al., 2019; Brunner et al., 2017; Chevallier et al., 2019). Inter-comparisons also provide an assessment of the systematic uncertainty on final flux estimates induced by the variety of options and choices in different inversion systems. However, although the inversion systems are referenced in peer-reviewed literature, and are usually accessible to the research community, they are typically not at the level of transparency, documentation, flexibility and accessibility required to provide both the scientific

https://doi.org/10.5194/gmd-2020-407
Preprint. Discussion started: 18 December 2020

community and policy makers with a comprehensive and robust view of the uncertainties associated with the inverse estimation of GHG fluxes. In particular, the differences between inversion systems (such as the atmospheric transport model, prior and observation space uncertainties, and inversion algorithm) make comparing their results particularly challenging, even when they are applied to the same problem. Moreover, research inversion systems are so far not ready for operational use, and their development, usually carried out by individual research institutes or limited consortia, may not keep pace with the scientific and technical needs to come, such as those linked to the increasing availability of high resolution satellite GHG observations (Janssens-Maenhout et al., 2020). A unified system would provide new possibilities to effectively and comprehensively assess GHG budgets, trends, and their uncertainties and quantify limitations and development needs related to different approaches, all which is needed in order to properly support emission reporting. Collaborative efforts towards unified systems are already happening in other data assimilation communities, with, e.g., the Object-Oriented Prediction System (OOPS; coordinated by the European Centre for Medium-range Weather Forecast, UK), or the Data Assimilation Research Testbed (DART; Anderson et al., 2009), dedicated to weather and marine forecasts. The Community Inversion Framework (CIF) is an initiative by members of the GHG atmospheric inversion community to bring together the different inversion systems used in the community, and is supported by the European Commission H2020 project VERIFY.

Despite their differences in methodology, application and implementation, almost all inversion systems rely on the same conceptual and practical bases: in particular, they use model-observation mismatches in a statistical optimization framework (based on Bayes' Theorem), and numerical atmospheric tracer transport and chemistry models to simulate mixing ratios of GHGs based on surface fluxes. The objectives of CIF are to develop a consistent input/model interface, to pool development efforts, and to have an inversion tool that is well-documented, open-source, and ready for implementation in an operational framework. The CIF is designed to be a flexible and transparent tool to estimate the fluxes of different GHGs (e.g. carbon dioxide CO_2, methane CH_4, nitrous oxide N_2O, or halocarbons) and other species based on atmospheric measurements. It is also designed to run at any spatial and temporal scales and with different atmospheric (chemistry-)transport models (global and regional, Eulerian and Lagrangian), with various observation data streams (ground-based, remote sensing, etc.), and a variety of data assimilation techniques (variational, analytical, ensemble methods, etc.). It will be possible to run it on multiple computing environments and corresponding set-ups and tutorials will be well documented. Community development will help in tackling the challenges in set-up and running, and accelerate adoption of the tool into wider use. One of the main foreseen advantages of the CIF is the capability to quantify and compare the errors due to the modeling of atmospheric transport and the errors due to the choice of a given inversion approach to solve a specific problem, in a fully consistent framework. The CIF will provide a common platform for quickly developing and testing new inversion techniques with several transport models, and it is hoped that with the combined community effort, it will be continuously improved and revised, keeping it state-of-the-art.

https://doi.org/10.5194/gmd-2020-407
Preprint. Discussion started: 18 December 2020


Geoscientific Model Development Discussions
Open Access
EGU

In the present paper, we lay out the basis of the CIF, giving details on its underlying principles and overall implementation. The proof-of-concept focuses on the implementation of several inversion methods, illustrated with a test case. We will dedicate a future paper to the evaluation of the system on a real-life problem with a number of interfaced atmospheric (chemistry-)transport models. At the time of writing the present article, the following models are interfaced with the CIF: the Global Circulation Models LMDZ (Chevallier et al., 2005) and TM5 (Krol et al., 2005; van der Laan-Luijkx et al., 2017), the regional chemistry-transport Eulerian model CHIMERE (Fortems-Cheiney et al., 2019) and the Lagrangian Particle Dispersion models FLEXPART (Pisso et al., 2019) and STILT (Trusilova et al., 2010). For the sake of simplicity, we refer to all types of (chemistry-)transport models generically as CTMs in the following. In Section 2, we describe the general theoretical framework for atmospheric inversions and how the CIF will include the theory in a flexible and general way. In Section 3, the practical implementation of the general design rules is explained, with details on the python implementation of the CIF. In Section 4, we demonstrate the capabilities of the CIF in a simple test case, applying various inversion techniques in parallel.

2 General principle

The version of the CIF presented here is implemented around Bayesian data assimilation methods with Gaussian assumptions, which constitute the main framework used in the atmospheric inversion systems for GHG fluxes (e.g., Enting, 2002; Bocquet et al., 2015). However, some studies have proposed possible extensions to more general probability density functions beyond the classical Gaussian case (e.g., truncated Gaussian densities, log-normal distributions, etc.; Michalak and Kitanidis, 2005; Bergamaschi et al., 2010; Miller et al., 2014; Zammit-Mangion et al., 2015; Lunt et al., 2016; Miller et al., 2019). Therefore, we propose here a general and flexible structure for our system that will be independent of limiting assumptions, as described in Sect 2.3, to allow future extensions to more general theoretical frameworks. In the following, mathematical formulas are written following notations based on Ide et al. (1997) and Rayner et al. (2019). We present the theoretical basis and several inversion methods that are implemented in the CIF as demonstrators.

2.1 General Bayesian data assimilation framework

The Bayesian approach consists in estimating the following conditional probability density function (pdf):

$$p^{\mathrm{a}}(\mathbf{x}) = p(\mathbf{x} \mid \mathbf{y}^{\mathrm{o}} - \mathcal{H}(\mathbf{x}^{\mathrm{b}}), \mathbf{x}^{\mathrm{b}}) \propto p(\mathbf{y}^{\mathrm{o}} - \mathcal{H}(\mathbf{x}) \mid \mathbf{x})\, p^{\mathrm{b}}(\mathbf{x}) \tag{1}$$

with $\mathbf{x}$ the target vector, $p^{\mathrm{a}}(\mathbf{x})$ the posterior distribution of the target vector, $p^{\mathrm{b}}(\mathbf{x})$ the prior knowledge of the target vector, $\mathbf{y}^{\mathrm{o}}$ the observation vector gathering all observations implemented in the inversion and $\mathcal{H}$ the observation operator linking the target vector to the observation vector. In the following, we also refer to $\mathcal{X}$ and $\mathcal{Y}$ as the target and observation spaces, respectively, from where the target and observation

https://doi.org/10.5194/gmd-2020-407
Preprint. Discussion started: 18 December 2020


Geoscientific Model Development Discussions

vectors are sampled. Classically, for atmospheric inversions, the observation vector $\mathbf{y}^{\mathrm{o}}$ includes ground-based measurements of GHG mixing ratios on fixed or mobile platforms, and remote sensing observations such as satellite observations. The target vector $\mathbf{x}$ includes the variables to be optimized by the inversion; it includes the main variables of interest, such as the surface fluxes, but also variables relating to atmospheric chemical sources and sinks, background concentrations in the case of limited-area transport models, model parameters, etc., which are required to make the inversion physically consistent. The observation operator $\mathcal{H}$ mainly includes the computation of atmospheric transport and chemistry (if relevant) by numerical Eulerian global circulation models (e.g., LMDZ, Chevallier et al. 2010; TM5, Houweling et al. 2014; GEOS-Chem, van der Laan-Luijkx et al. 2017; Liu et al. 2015; Palmer et al. 2019; Feng et al. 2017; NICAM, Niwa et al. 2017), regional Eulerian chemistry-transport models (e.g., CHIMERE, Broquet et al. 2011; Fortems-Cheiney et al. 2019; WRF-CHEM, Zheng et al. 2018; COSMO-GHG, Mizzi et al. 2016; LOTOS-EUROS, Curier et al. 2012) or Lagrangian Particle Dispersion models (e.g., FLEXPART, Thompson and Stohl 2014; STILT, Bagley et al. 2017; Brioude et al. 2013; Trusilova et al. 2010). It also includes pre- and post-processing operations required to project the target vector to a format compatible with the model input and the model outputs to the observation vector; these operations can be the applications of e.g., averaging kernels in the case of satellite operations, or interpolation of the target vector to higher resolution model inputs.

As errors in inversion systems come from a large variety of independent causes superimposing on each other, it is often assumed that the most relevant way of representing the distributions in Eq. 1 is to assume prior and observation spaces to be normal distributions. In addition, when assuming that the distributions in the state vector space and the observation space are independent from each other, it is possible to represent the distributions of Eq. 1 as follows:

$$\begin{cases} p^{\mathrm{b}}(\mathbf{x}) & \sim & \mathcal{N}(\mathbf{x}^{\mathrm{b}},\, \mathbf{B}) \\ p\left(\mathbf{y}^{\mathrm{o}} - \mathcal{H}(\mathbf{x}^{\mathrm{b}})\right) & \sim & \mathcal{N}(\mathbf{0},\, \mathbf{R}) \\ p^{\mathrm{a}}(\mathbf{x}) & \sim & \mathcal{N}(\mathbf{x}^{\mathrm{a}},\, \mathbf{A}) \end{cases} \tag{2}$$

with $\mathbf{B}$ and $\mathbf{A}$ the prior and posterior covariance matrix of uncertainties in the target vector, $\mathbf{x}^{\mathrm{b}}$ and $\mathbf{x}^{\mathrm{a}}$ the prior and posterior target vectors and $\mathbf{R}$ the covariance matrix of uncertainties in the observation vector and the observation operator.

2.2 Computation modes in the CIF

The present version of the CIF includes three main categories of inversion methods: 1) analytical, i.e. algebraic solution, 2) ensemble methods with the Ensemble Kalman Filter, and 3) variational with two examples of minimizing algorithms. Other types of data assimilation methods (e.g. direct sampling of probability density functions through Monte Carlo approaches) are also used by the community. The choice of implementing the three aforementioned methods first is motivated by their dominant use, and because these three

https://doi.org/10.5194/gmd-2020-407
Preprint. Discussion started: 18 December 2020


Geoscientific Model Development Discussions

use very different approaches for solving the Bayesian inversion problem: that is, with/without random sampling of probability distributions, and with/without the use of the adjoint of the observation operator. The adjoint of the observation operator, noted $\mathcal{H}^*$, is built following the mathematical definition of the adjoint; heuristically, it operates backwards compared to the observation operator in the sense that it determines the sensitivity to inputs (e.g. fluxes) given an incremental perturbation to outputs (e.g. concentrations) (e.g., Errico, 1997). In addition to the mentioned data assimilation methods, the CIF also includes the possibility to run forward simulations and to test the adjoint and the tangent linear of the observation operator for given inversion configurations. In the following we call all inversion methods and other types of computation in the CIF "computation modes". With these computation modes implemented in a flexible and general manner, it is anticipated that other inversion methods could be easily added to the CIF in the future (see Sect. 2.3).

2.2.1 Data assimilation methods

Analytical inversions

Analytical inversions compute the algebraic solution of the Gaussian Bayesian problem when it is linear and are used extensively at all scales (e.g., Stohl et al., 2009; Turner and Jacob, 2015; Kopacz et al., 2009; Bousquet et al., 2011; Wang et al., 2018; Palmer et al., 2006). When the observation operator is linear, $\mathcal{H}$ equals its Jacobian matrix $\mathbf{H}$, and conversely its adjoint $\mathcal{H}^*$ is the transpose of the Jacobian $\mathbf{H}^{\mathrm{T}}$. In that case, $\mathbf{x}^{\mathrm{a}}$ and $\mathbf{A}$ can be explicitly written as matrix products. There are two equivalent formulations of the matrix products for the solution of the problem (e.g., Tarantola and Valette, 1982):

$$\left\{\begin{array}{rcl}\mathbf{x}^{\mathrm{a}} & = & \mathbf{x}^{\mathrm{b}}+\mathbf{K}(\mathbf{y}^{\mathrm{o}}-\mathbf{H}\mathbf{x}) \\ \mathbf{A} & = & \mathbf{B}-\mathbf{K}\mathbf{H}\mathbf{B}\end{array}\right. \quad \text{or} \quad \left\{\begin{array}{rcl}\mathbf{x}^{\mathrm{a}} & = & \mathbf{x}^{\mathrm{b}}+(\mathbf{H}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{H}+\mathbf{B}^{-1})^{-1}\mathbf{H}^{\mathrm{T}}(\mathbf{y}^{\mathrm{o}}-\mathbf{H}\mathbf{x}) \\ \mathbf{A} & = & (\mathbf{H}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{H}+\mathbf{B}^{-1})^{-1}\end{array}\right. \tag{3}$$

with $\mathbf{K}$ the Kalman gain matrix: $\mathbf{K}=\mathbf{B}\mathbf{H}^{\mathrm{T}}(\mathbf{R}+\mathbf{H}\mathbf{B}\mathbf{H}^{\mathrm{T}})^{-1}$

The computation of an analytical inversion faces two main computational limitations. First, the matrix $\mathbf{H}$ representing the observation operator $\mathcal{H}$ must be built explicitly. This can be done either column by column, in the so-called response function method, by computing $\{\mathcal{H}(\mathbf{x}_i) \setminus \forall \mathbf{x}_i \in \mathcal{B}_\chi\}$ with $\mathcal{B}_\chi$ the canonical basis of the target space, or row by row, in the so-called footprint method, by computing $\{\mathcal{H}^*(\delta\mathbf{y}_i) \setminus \forall \delta\mathbf{y}_i \in \mathcal{B}_\mathcal{Y}\}$ with $\mathcal{B}_\mathcal{Y}$ the canonical basis of the observation space. Depending on the number of available observations or the size of the target vector, one of the two is preferred to limit the number of observation operator computations to be carried out explicitly. When the dimension of the target vector is relatively small, the response function is generally preferred, and conversely, when the observation vector is small, the footprint approach is preferred. The type of transport model used to compute the matrix $\mathbf{H}$ also plays a role in the choice of the approach: for Eulerian models, the response function approach is preferred for multiple reasons: (i) their adjoint is often much more costly than their forward, (ii) the adjoint may not be available for some models or

https://doi.org/10.5194/gmd-2020-407
Preprint. Discussion started: 18 December 2020

is difficult to generate, and (iii) the computation time of the forward is constant no matter how numerous the observations; for Lagrangian models, the footprint approach is preferred as they often compute backward transport simulations for each observation, allowing a straightforward computation of the adjoint (Seibert and Frank, 2004). In both cases, the explicit construction of the matrix $\mathbf{H}$ requires numerous independent simulations, which can be an insurmountable computational challenge.

The second obstacle consists in that the computation of the Kalman gain matrix in Eq. 3 (left) requires inverting of a matrix of the dimension of the observation vector, $\dim(\mathcal{Y})$, while for the other formulation (Eq. 3 right) a matrix of dimension $\dim(\chi)$, the dimension of the control space. If the dimensions of both the observation and the control spaces are very high, as in many inversion applications, the explicit computation of Eq. 3 with matrix products and inverses is not computationally feasible. For this reason, smart adaptations of the inversion framework (including approximations and numerical solvers) are often necessary to tackle even linear problems: the variational approach and the Ensemble Kalman filter are described below. Less frequently, intermediate adaptations of the analytical inversion also include sequential applications (e.g., Michalak, 2008; Bruhwiler et al., 2005; Brunner et al., 2012), that are a compromise between tackling the above-mentioned computational obstacles while maintaining the simplicity of the analytical inversion; however, such sequential analytical inversions are limited to specific linear, simple cases.

Ensemble methods

Ensemble methods are commonly used to tackle high-dimensional problems with limited non-linearity and to approximately characterized the optimal solution. In ensemble methods, such as Ensemble Kalman filters (EnKFs) or smoothers (e.g., Peters et al., 2005; Zupanski et al., 2007; Zupanski, 2005; Feng et al., 2009; Chatterjee et al., 2012), the issue of high dimensions in the system of Equations 3 is avoided using the two following main procedures:

- observations are assimilated sequentially in the system to reduce the dimension of the observation space, making it possible to explicitly compute matrix products and inverses; the overall inversion time window is processed incrementally using a smaller running computation window including a manageable number of observations; intermediate inversions are solved on the smaller running window that is gradually moved from the beginning to the end of the overall inversion window; the running assimilation window with so-called analysis and forecast steps introduces complex technical challenges to rigorously propagate errors from one iteration of the running window to the next one; for very dense observations, such as datasets from new-generation high-resolution satellites, the sequential assimilation of observations may not be sufficient; moreover, the sequential assimilation of observations is valid only under the assumption that observations for each assimilation window are not correlated with each others, which may prove incorrect for high-density data sets;

https://doi.org/10.5194/gmd-2020-407
Preprint. Discussion started: 18 December 2020

– matrix products in Eq. 1 involving the target vector covariance matrix $\mathbf{B}$ ($\mathbf{HBH}^{\mathrm{T}}$ and $\mathbf{BH}^{\mathrm{T}}$) are approximated by reducing the space of uncertainties to a low-rank representation; this is done in practice by using a Monte Carlo ensemble of possible control vectors sampling the distribution $\mathcal{N}(\mathbf{x}^{\mathrm{b}}, \mathbf{B})$; with such an approximation, matrix products can be written as follows:

$$\begin{cases} \mathbf{HBH}^{\mathrm{T}} & \simeq \dfrac{1}{N-1}\left(\mathcal{H}(\mathbf{x}_1),\, \mathcal{H}(\mathbf{x}_2),\, \ldots,\, \mathcal{H}(\mathbf{x}_N)\right) \,.\, \left(\mathcal{H}(\mathbf{x}_1),\, \mathcal{H}(\mathbf{x}_2),\, \ldots,\, \mathcal{H}(\mathbf{x}_N)\right)^{\mathrm{T}} \\ \mathbf{BH}^{\mathrm{T}} & \simeq \dfrac{1}{N-1}(\mathbf{x}_1,\, \mathbf{x}_2,\, \ldots,\, \mathbf{x}_N) \,.\, \left(\mathcal{H}(\mathbf{x}_1),\, \mathcal{H}(\mathbf{x}_2),\, \ldots,\, \mathcal{H}(\mathbf{x}_N)\right)^{\mathrm{T}} \end{cases} \tag{4}$$

where N is the size of the ensemble.

From there, Eq. 1 is solved analytically by replacing $\mathbf{HBH}^{\mathrm{T}}$ and $\mathbf{BH}^{\mathrm{T}}$ by their respective approximations. Using random sampling, ensemble methods are able to approximate large dimensional matrices at a reduced cost without using the adjoint of the observation operator (see variational inversion below) that can be challenging to implement. However, using too small ensembles causes degenerate approximations (under-estimating the uncertainty magnitude, or misrepresenting uncertainty structures), which limits the accuracy of the computed solution, and may require fixes as described in e.g., Bocquet (2011). In any case, the level of approximation necessary for this approach to work is strongly different for each problem, which requires preliminary studies before consistent application.

Variational inversions

Variational inversions are a numerical approximation to the solution of the inversion problem: they involve the gradient of the cost function in Eq. 5 and require to run forward and adjoint simulations iteratively (e.g., Meirink et al., 2008; Bergamaschi et al., 2010; Houweling et al., 2016, 2014; Fortems-Cheiney et al., 2019; Chevallier et al., 2010, 2005; Thompson and Stohl, 2014; Monteil and Scholze, 2019; Wang et al., 2019). In variational inversions, the solution $\mathbf{x}$ is defined as being that with maximum posterior probability. In the case of Gaussian assumptions, it is equivalent to computing the mode $\mathbf{x}^{\mathrm{a}}$ of the normal distribution. Computing $\mathbf{x}^{\mathrm{a}}$ in Eq. 1 is equivalent to finding the minimum of the cost function:

$$J(\mathbf{x}) = \frac{1}{2}(\mathbf{x}-\mathbf{x}^{\mathrm{b}})^{\mathrm{T}}\mathbf{B}^{-1}(\mathbf{x}-\mathbf{x}^{\mathrm{b}}) + \frac{1}{2}(\mathbf{y}^{\mathrm{o}}-\mathcal{H}(\mathbf{x}))^{\mathrm{T}}\mathbf{R}^{-1}(\mathbf{y}^{\mathrm{o}}-\mathcal{H}(\mathbf{x})) \tag{5}$$

The variational formulation does not require calculation of complex matrix products and inversions, contrary to the analytical inversion, and is thus not limited by vector dimensions. Still, the inverses of the uncertainty matrices $\mathbf{B}$ and $\mathbf{R}$ need to be computed, potentially prohibiting the use of very large and/or complex general matrices; this challenge is often overcome by reducing $\mathbf{B}$ and $\mathbf{R}$ to manageable combinations of simple matrices (e.g., Kronecker products of simple shape covariance matrices; see Sect. 2.3.1).

https://doi.org/10.5194/gmd-2020-407
Preprint. Discussion started: 18 December 2020


Geoscientific Model Development Discussions

In variational inversions, the minimum of the cost function in Eq. 5 is numerically estimated using quasi-Newtonian algorithms based on the gradient of the cost function:

$$\nabla J_{\mathbf{x}} = \mathbf{B}^{-1}.(\mathbf{x}-\mathbf{x}^{\mathrm{b}}) + \mathcal{H}^{*}\left(\mathbf{R}^{-1}.(\mathbf{y}^{\mathrm{o}}-\mathcal{H}(\mathbf{x}))\right) \tag{6}$$

Quasi-Newtonian methods are a group of algorithms designed to compute the minimum of a function iteratively. It should be noted that in high-dimension problems, it can take a very large number of iterations to reach the minimum of the cost function J, forcing the user to stop the algorithm before convergence, thus reaching only an approximation of $\mathbf{x}^{\mathrm{a}}$; in addition, iterative algorithms can reach local minima without ever reaching the global minimum, making it essential to thoroughly verify variational inversion results; this can happen in non linear cases, but also, due to numerical artefacts in linear cases (some points in the cost function have gradients so close to zero that the algorithm sees them as convergence points, whereas the only global minimum is somewhere else). In the community, examples of quasi-Newtonian algorithms commonly used are M1QN3 (Gilbert and Lemaréchal, 1989), the conjugate gradient CONGRAD (applicable only to linear or linearized problems; Fisher, 1998) algorithm (e.g., Chevallier et al., 2005) or the Broyden–Fletcher–Goldfarb–Shanno (BFGS) algorithm (Zheng et al., 2018; Bousserez et al., 2015). In general, quasi-Newtonian methods require an initial regularization of $\mathbf{x}$, the vector to be optimized, for better efficiency. In atmospheric inversions, such a regularization is generally made by optimizing $\chi = \mathbf{B}^{-1/2}.(\mathbf{x}-\mathbf{x}^{\mathrm{b}})$ instead of $\mathbf{x}$; we note $\mathfrak{A}$ the regularization space: $\chi \in \mathfrak{A}$. This transformation translates in Eq. 6 as follows:

$$\nabla J_{\chi} = \chi + \mathbf{B}^{1/2}.\mathcal{H}^{*}\left(\mathbf{R}^{-1}.(\mathbf{y}^{\mathrm{o}}-\mathcal{H}(\mathbf{x}))\right) \tag{7}$$

Solving Eq. 5 in the target vector space or Eq. 7 in the regularization space is mathematically fully equivalent, but the solution in the regularization space is often reached in fewer iterations. Moreover, in the regularization space, one can force the algorithm to focus on the main modes of the target vector space by filtering the smallest eigenvalues of the matrix $\mathbf{B}$. This reduces the dimension of χ and accelerates further the rate of convergence, although the solution of the reduced problem is only an approximation of the solution of the full problem. In the following we thus prefer calling the "regularization space" the "reduction space". The link between the two can be written as follows:

$$\begin{aligned} \chi_{\mathrm{full}} &= \mathbf{Q}\mathbf{\Lambda}^{1/2} \quad (\mathbf{x}-\mathbf{x}^{\mathrm{b}}) \\ \chi_{\mathrm{reduced}} &= \mathbf{Q}'\mathbf{\Lambda}'^{1/2} \quad (\mathbf{x}-\mathbf{x}^{\mathrm{b}}) \end{aligned} \tag{8}$$

with $\mathbf{B}^{1/2} = \mathbf{Q}\mathbf{\Lambda}^{1/2}\mathbf{Q}^{-1}$, $\mathbf{Q}$ and $\mathbf{\Lambda}$ being the matrices of the eigenvector and the matrix of the corresponding eigenvalues. $\mathbf{Q}'$ and $\mathbf{\Lambda}'$ are the reduced matrices of eigenvalues and eigenvectors with a given number of dominant eigenvalues.

When the observation operator is linear, the posterior uncertainty matrix $\mathbf{A}$ is equal to the inverse of the Hessian matrix at the minimum of the cost function. In most cases the Hessian cannot be computed

https://doi.org/10.5194/gmd-2020-407
Preprint. Discussion started: 18 December 2020


Geoscientific Model Development Discussions

explicitly, because of memory limitations, which is a major drawback of variational inversions. But some variational algorithms such as CONGRAD provide a coarse approximation of the Hessian: in the case of CONGRAD, leading eigenvectors of the Hessian can be computed, together with their eigenvalues (Fisher, 1998). Another approach to quantify the posterior uncertainty matrix $\mathbf{A}$, valid for both linear and non-linear cases, is to carry out a Monte Carlo ensemble of independent inversions with sampled prior vectors:

2.2.2 Auxiliary computation modes

Forward simulations

Forward simulations simply use the observation operator to compute simulated observation equivalents. It reads as:

$$(\mathbf{x}^{\mathrm{b}}, \mathbf{y}^{\mathrm{o}}) \rightarrow \mathcal{H}(\mathbf{x}^{\mathrm{b}}) \tag{9}$$

This mode is used to make quick comparisons between observations and simulations to check for inconsistencies before running a full inversion. It is also used by the analytical inversion mode to build response functions.

Test of the adjoint

The test of the adjoint is a crucial diagnostic for any inversion system making use of the adjoint of the observation operator. Such a test is typically required after making any edits to the code (to the forward observation operator or its adjoint) before running an inversion. Coding an adjoint is prone to errors and even small errors can have significant impacts on the computation of the gradient of the cost function in Eq. 6. Thus, one needs to make sure that the adjoint rigorously corresponds to the forward. This test consists in checking the definition of the mathematical adjoint of the observation operator. It writes as follows for a given target vector $\mathbf{x}$ and incremental target perturbation $\delta\mathbf{x}$:

$$< D\mathcal{H}(\mathbf{x}, \delta\mathbf{x}) \,|\, D\mathcal{H}(\mathbf{x}, \delta\mathbf{x}) > = < \delta\mathbf{x} \,|\, (\mathcal{H}^* \circ D\mathcal{H})(\mathbf{x}, \delta\mathbf{x}) > \tag{10}$$

$D\mathcal{H}(\mathbf{x}, \delta\mathbf{x})$ is the linearization of the observation operator $\mathcal{H}$ at the point $\mathbf{x}$ for a given increment $\delta\mathbf{x}$; it is computed with the tangent linear model, which is the numerical adaptation of $D\mathcal{H}(\mathbf{x}, \delta\mathbf{x})$. Then, $(\mathcal{H}^* \circ D\mathcal{H})(\mathbf{x}, \delta\mathbf{x})$ is calculated with the adjoint of the tangent-linear of $\mathcal{H}$ at the point $\mathbf{x}$.

In practice, the two terms of the equation are rarely exactly equal. Nevertheless, the difference should never exceed a few times the machine epsilon. Besides, Eq. 10 should be verified for any given target vector and increment. In practice, it is not possible to explicitly verify all possible combinations; but as the result

https://doi.org/10.5194/gmd-2020-407
Preprint. Discussion started: 18 December 2020


Geoscientific Model Development Discussions

of the test is highly sensitive to any error in the code, it is assumed that a few typical couples $(\mathbf{x}, \delta\mathbf{x})$ are sufficient to certify the validity of the adjoint.

2.3 Identification of common elementary transformations

2.3.1 General purpose operations

Every inversion algorithm and computation mode mentioned above can be decomposed into a pipeline of elementary transformations. These transformations are listed in Tab. 1 and include: the observation operator and its adjoint (their matrix representations in linear cases), matrix products with control and observation error covariance matrices and corresponding adjoints, and random sampling of Normal distributions. To limit redundancy in the CIF as much as possible, these elementary transformations are included in the CIF as generic transformation blocks on the same conceptual level. Overall, the decomposition of computation modes presently implemented in the CIF into elementary transformations leads to the structure in Fig. 1.

Avoiding redundancy makes the maintenance of the code much easier, and provides a clear framework for extensions to other inversion methods or features. For instance, inverse methods based on probability density functions other than Normal distributions could be easily implemented by updating the random ensemble generator, or by implementing new cost functions representing non Gaussian distributions, while keeping the remaining code unmodified. In particular, non-Gaussian cost functions still rely on the computation of the observation operator. New combinations of elementary transformations can also directly lead to new methods. For instance, ensemble variational inversion (e.g., Bousserez and Henze, 2018) is a direct combination of the available variational pipeline and the random sampling pipeline. Inversions estimating hyper-parameters through maximum-likelihood or hierarchical Bayesian techniques (e.g., Michalak et al., 2005; Berchet et al., 2014; Ganesan et al., 2014) could be integrated in the CIF by adapting the Gaussian cost function and by implementing a corresponding computation pipeline.

The complexity of the selected elementary transformations spans a wide range, from one-line straightforward codes to computationally expensive and complex code implementation. In small dimensional and/or linear problems, the computation of the observation operator using its Jacobian and matrix products may be computationally expensive, but is in principle rather straightforward to implement. For non-linear and/or high-dimensional problems, these transformations require simplifications and numerous intermediate steps. For instance, applying matrix products to the error covariance matrix $\mathbf{R}$ and $\mathbf{B}$ and computing their inverse is easy in small dimensions, but can be limiting in high dimensional problems; for that reason, the error covariance matrices are often reduced to particular decompositions; for instance, the error covariance matrix on the target vector $\mathbf{B}$ is often written as a Kronecker product of multiple spatial and/or temporal covariance matrices of lower dimensions, making matrix products manageable (e.g., Chevallier et al., 2005; Meirink et al., 2008; Yadav and Michalak, 2013).

https://doi.org/10.5194/gmd-2020-407
Preprint. Discussion started: 18 December 2020


Geoscientific Model Development Discussions
Open Access
EGU

In any case, the observation operator (see details in Sect. 2.3.2) appears as the center piece of any inversion method.

2.3.2 Observation operator

The observation operator is a key component of all inversion methods. It links the target space to the observation space, and conversely, its adjoint links the observation space to the target space. To do so, the observation operator projects its inputs through various intermediate spaces to the outputs. As atmospheric inversions need a representation of the atmospheric transport (and chemistry if relevant) to link the target vector (including surface fluxes, atmospheric sources and sinks, initial and boundary conditions for limited domains and time-windows, etc.) to the observation vector (including some form of atmospheric concentration measurements), the observation operator is built around a given (chemistry-)transport model in most cases: Eq. 11 illustrates the various projections in this case.

$$\mathbf{x} \xrightarrow{\mathbf{\Pi}^{\mathfrak{F}}_{\mathcal{X}}} \mathbf{f} \xrightarrow{\mathbf{\Pi}^{\mathcal{F}}_{\mathfrak{F}}} \text{inputs} \xrightarrow{\text{model}} \text{outputs} \xrightarrow{\mathbf{\Pi}^{\mathfrak{M}}_{\mathcal{C}}} \mathbf{c} \xrightarrow{\mathbf{\Pi}^{\mathcal{Y}}_{\mathfrak{M}}} \mathcal{H}(\mathbf{x}) \tag{11}$$

with $\mathbf{f}$ the target vector projected at the CTM's resolution (includes fluxes, but also other types of inputs required by the CTM), $\mathbf{c}$ the raw outputs extracted from the run of the CTM's executable (in general 4-dimensional concentration fields). $\mathbf{\Pi}$ operators are intermediate projectors: $\mathbf{\Pi}^{\mathfrak{F}}_{\mathcal{X}}$ projects the target vector at the spatial and temporal resolutions of the CTM's inputs, $\mathbf{\Pi}^{\mathcal{F}}_{\mathfrak{F}}$ dumps the input vector in files usable by the CTM's executable, $\mathbf{\Pi}^{\mathfrak{M}}_{\mathcal{C}}$ reads the CTM's outputs, $\mathbf{\Pi}^{\mathcal{Y}}_{\mathfrak{M}}$ reprojects the raw outputs at the observation vector resolution.

The targeted structure of the CIF should allow a full flexibility of observation operators, from the straightforward widely-used decomposition detailed in Eq.11 to more elaborated approaches including multiple transport models and complex super-observations (e.g., concentration gradients or aggregates; Bréon et al., 2015) and hyper-parameters (e.g., emission factors and model parameters used to produce emission maps; Rayner et al., 2010; Asefi-Najafabady et al., 2014). Therefore, the observation operator is designed as a pipeline of elementary interchangeable transformations with standardized input and output formats such that:

$$\mathcal{H} = \mathcal{H}_1 \circ \mathcal{H}_2 \circ \cdots \circ \mathcal{H}_N \tag{12}$$

https://doi.org/10.5194/gmd-2020-407
Preprint. Discussion started: 18 December 2020


Geoscientific Model Development Discussions

Such a design also allows us to rigorously separate transformations and thus to implement their respective adjoints more easily. Once adjoints for each individual operations are implemented, the construction of the general adjoint is straightforward by reversing the order of forward operations:

$$\mathcal{H}^* = \mathcal{H}_N^* \circ \mathcal{H}_{N-1}^* \circ \cdots \circ \mathcal{H}_1^* \tag{13}$$

Fig. 2 shows an example of a typically targeted observation operator. Operators from Eq. 11 are reported for the illustration. It includes two numerical models chained with each other; they can be for instance a coarse global CTM and a finer resolution regional CTM, such as in Rödenbeck et al. (2009) or Belikov et al. (2016). The system applies a series of transformations to the target vector, including spatial deaggregation for the optimization of emissions by regions, sector deaggregation to separate different activity sectors, reprojection to the CTM's resolution, application of temporal profiles (which is critical in air quality and anthropogenic CO_2 applications), unit conversions to the required inputs for the CTMs. On the observation vector side, observations can span multiple model time-steps, requiring posterior temporal averages, etc. In the case of super-observations (satellites retrievals, images, spatial gradients, etc.) in the observation vector, it is often necessary to combine multiple simulated point observations in given grid cells and time stamps into a single super-observation. This is the case for satellite observations being compared to all the model levels above a given location, or for concentration gradients comparing different time and locations, or also isotopic ratios that require to simulate separate isotopologues and recombine them after the simulation (as done in e.g., van der Velde et al., 2018; Peters et al., 2018).

3 Practical implementation

3.1 General rules

The Community Inversion Framework project follows the organisation scheme of Fig. 3. A centralized website is available at community-inversion.eu. The website includes all information given in the present paper, as well as further documentation details, practical installation instructions, tutorials and examples of possible set-ups. To foster the collaborative dynamics of our project, all scripts and codes are available in open-access on a GitLab server at git.nilu.no/VERIFY/CIF, where updates are published regularly. The frozen version of the code, documentation and data used for the present publication is available in Berchet et al. (2020). The repository includes the documentation, sources for the CTMs implemented in the CIF, as well as the Python library pyCIF. Our project is distributed as an open-source project under the CECILL-C licence of the French law (cecill.info). The license grants full rights for the users to use, modify and redistribute the original version of the CIF, conditional to the obligation to make their modifications available to the community and to properly acknowledge the original authors of the code. The authors of modifications own

https://doi.org/10.5194/gmd-2020-407
Preprint. Discussion started: 18 December 2020

intellectual property of their modifications, but under the same governing open license. Software that may be built around the CIF in the future can have different licensing, but all parts of the code originating from the CIF will be governed by the original CeCILL-C license. Similarly, some constituting pieces of the CIF can be adapted from other softwares governed by other licenses and simply interfaced to the CIF (e.g., transport models, minimizing algorithms, etc.); in that cases, the corresponding softwares keep their original license and their use and distribution in the CIF is subject to authorization by their owners (although open distribution and integration in the standard version of the CIF is encouraged). This is the case of the CONGRAD and M1QN3 algorithms which are used as minimizing algorithms in the variational inversions of the demonstration case in Sect. 4. The M1QN3 algorithm is distributed under the GNU General Public License, whereas CONGRAD is owned by ECMWF and is not open source; the later was interfaced with the CIF but is not openly distributed.

The pyCIF library is the practical embodiment of the CIF project. All theoretical operations described in Sect. 2 are computed by this module. It includes inversion computations, pre- and post-processing of CTM inputs and outputs, as well as target and observation vector reprojections, aggregation, etc., as written in Eq. 13. Python coding standards follow the community standards PEP-8 (python.org/dev/peps/pep-0008/).

Test cases (including the ones presented in Sect 4) are distributed alongside the CIF codes and scripts.

3.2 Plugin-based implementation

To reflect the theoretical flexibility required in the computation of various inversion methods and observation operators, we made the choice of implementing pyCIF following an abstract structure with a variety of so-called Python plugins, which are dynamically constructed and inter-connected depending on the set-up.

3.2.1 Objects and classes in pyCIF

General classes of objects emerge from the definition of the abstract structure of the inversion framework. These classes are defined by the data and metadata they carry, as well as by the methods they include and their interaction with other classes. The main classes are the following:

- computation modes: forward computations, the test of the adjoint, variational inversions, EnKF and analytical inversions are available (see details in Sect. 2.2);
- models: interfaces to CTMs; includes generation of input files, executing the code and post-processing outputs; included are a Gaussian model described in Sect. 4 for the demonstration of the system, as well as CHIMERE, LMDZ, FLEXPART, TM5, and STILT, all of which will be described in a dedicated future publication;
- platforms: deal with specific configurations on different clusters; it includes a standard platform as well as two supercomputers where the CIF was tested

https://doi.org/10.5194/gmd-2020-407
Preprint. Discussion started: 18 December 2020

- control vectors: store and apply operations related to the control vector, including spatial and temporal aggregation, deaggregation, regularization of the target vector;
- observation vectors: store and apply operations related to the observation vector, including application of observation errors
- observation operators: drive CTMs and apply elementary operations between the control and observation vectors
- transformations: elementary operations used to build the observation operator; includes temporal averaging or deaggregating of the target and observation vectors, projection of the target vector at the model input resolution, etc.
- data vectors: store all information on inputs for pyCIF; this vector is used by the observation and control vector classes to build themselves
- minimizers: algorithms used to minimize cost functions, including M1QN3 and CONGRAD algorithms so far
- simulators: cost functions to minimize in variational inversions; only includes the standard Gaussian cost function so far
- domains: store information about the CTM's grid, including coordinates of grid cell centers and corners, vertical levels, etc.
- fluxes, fields, meteos: fetch, read and write different formats of inputs for CTMs; so far includes only inputs specific to included CTMs, but will ultimately include standard data streams, such as widely used emission inventories or meteorological fields such as those from ECMWF
- measurements: fetch, read and write different types of observation data streams; only include the World Data Center for Greenhouse Gases so far (https://gaw.kishou.go.jp/), but classical data providers such as ICOS (icos-cp.eu) or ObsPack Masarie et al. (2014) will also be implemented in the CIF

Details on metadata and operations for each class are given in Supplements, Tab. **??**. Our objective was to design a code that is fully recursive in the sense that modifying some instance of a class does not require to update other classes calling or being called by the modified class. Thus, each class is built so that it only needs internal data, as well as data from the execution level just before and after it, in order to avoid complex dependencies while allowing proper recursive behaviour in building the transformation pipe. For instance, the observation operator applies a pipe of transformations from the target vector to the observation vector. Some transformations will use the model class to run the model, or the domain class to carry out reprojections, or the target vector to aggregate/de-aggregate target dimensions, etc. Despite the many complex

https://doi.org/10.5194/gmd-2020-407
Preprint. Discussion started: 18 December 2020


Geoscientific Model Development Discussions

transformations carried out under the umbrella of the observation operator, only the sub-transformations of the pipe are accessible at the observation operator level, which do not have to carry directly information about e.g., the model or other classes required at sub-levels. This makes the practical code of the observation operator much simpler and as easy to read as possible.

3.2.2 Automatic construction of the execution pipe

To translate the principle scheme of Fig. 1, pyCIF is not built in a sequential rigid manner. Plugins are interconnected dynamically at the initializing step of pyCIF depending on the chosen set-up (see Sect. 3.3 for details on the way to configure the CIF). The main strength of such a programming structure is the independence of all objects in pyCIF. They can be implemented separately in a clean manner. The developer only needs to specify what other objects are required to run the one being developed and pyCIF makes the links to the rest. It avoids unexpected impacts elsewhere in the code when modifying or implementing a feature in the system. In the following, we call this top-down relationship in the code a dependency.

For each plugin required in the configuration (primarily the computation mode), pyCIF initializes corresponding objects following simple rules. Following dependencies detailed in Tab. **??**, for every object to initialize, pyCIF will fetch and initialize required plugins and attach them to the original plugin. If the required plugin is explicitly defined in the configuration, pyCIF will fetch this one. In some cases, some plugins can be built on default dependencies that do not need to be defined. In that case, the required plugin can be retrieved using default plugin dependencies specified in the code itself and not needed in the configuration.

For instance, in the call graph in Fig. 1, "variational inversion" is a "computation mode" object in pyCIF. To execute, it requires a "minimizer" object (CONGRAD, M1QN3, etc.) that is initialized and attached to it. The minimizer requires a "simulator" object (the cost function) that itself will call functions in the "control vector" object and the "observation operator" object. Then the "observation operator" will initialize a pipeline of transformations including running the "model", and so on and so forth.

3.3 Definition of configurations in the CIF

In practice, pyCIF is configured using a YAML configuration file (yaml.org). This file format was primarily chosen for its flexibility and intuitive implementation of hierarchical parameters. In the YAML language, key words are specified with associated values by the user. Indentations indicate sub-levels of parameters, which makes it a consistent tool with the coding language python.

To set-up a pyCIF computation, the user needs to define the computation mode and all related requirements in the YAML configuration file. Every plugin has mandatory and optional arguments. The absence of one mandatory argument rises an error at initialization. Optional arguments are replaced by corresponding default values if not specified. Examples of YAML configuration files used to carry out the demonstration cases are given in Supplement.

https://doi.org/10.5194/gmd-2020-407
Preprint. Discussion started: 18 December 2020

Geoscientific Model Development Discussions

4 Demonstration case

In the following we describe a demonstration case based on a simple implementation of a Gaussian plume dispersion model and simple inversion set-ups. The purpose of this demonstration case is a proof-of-concept of the CIF, with various inversion methods. We comment and compare inversion set-ups and methods for the purpose of the exercise, but conclusions are not made to be generalized to any inversion case study due to the simplicity of our example. The test application with a simple Gaussian plume model allows users to quickly carry out the test cases themselves, even on desktop computers, to familiarize themselves with the system. Nevertheless, the Gaussian plume model is not only relevant for teaching purposes, but also for real applications, as it is used in many inversion studies from the scale of industrial sites with in-situ fixed or mobile measurements (e.g., Kumar et al., 2020; Foster-Wittig et al., 2015; Ars et al., 2017) to the global scale with satellite measurements (e.g., Nassar et al., 2017; Wang et al., 2020). Other models implemented in the CIF will be presented in a future paper evaluating the differences when using different transport models with all other elements of the configuration identical. The purpose of such an evaluation is to produce a rigorous inter-comparison exercise identifying the effect of transport errors in inversion systems.

4.1 Gaussian plume model

Gaussian plume models approximate real turbulent transport by a stable average Gaussian state (Hanna et al., 1982). Such models are not always suitable to compare with continuous measurements but can be adapted when using observations averaged over time. In the following, we consider the Gaussian plume assumption to be valid for comparing to hourly averaged observations. A simple application of the Gaussian plume model was implemented in the CIF as a testing and training utility. It is computationally easy to run, even on desktop computers. It includes the most basic Gaussian plume equations. In that application, concentrations $\mathcal{C}$ at location $(x_0,\, y_0,\, z_0)$ downwind from a source of intensity f at $(x_1,\, y_1,\, z_1)$ are given by:

$$\mathcal{C}(x_0,\, y_0,\, z_0) \;=\; \frac{f}{2\,\pi\,.\,\sigma_y\,.\,\sigma_z\,.\,\bar{u}} \exp\left(-\frac{y^2}{\sigma_y^2}\right)\,.\,\exp\left(-\frac{z^2}{\sigma_z^2}\right) \tag{14}$$

with

$$\begin{cases} \sigma_z &=& a\,.\,x^b \\ \sigma_y &=& |465.11628\times\; x\,.\,\tan(0.017653293\,(c - d\,.\,\ln x))| \\ x &=& < \dfrac{\mathbf{u}}{\bar{u}} \mid v_{\text{(source, receptor)}} > \\ y &=& \sqrt{||\,v_{\text{(source, receptor)}}\,||^2 - x^2} \end{cases} \tag{15}$$

x is the downwind distance between the source and receptor points along the wind axis, y is the distance between the wind axis and the receptor point, z is the difference between the source and the receptor alti-

https://doi.org/10.5194/gmd-2020-407
Preprint. Discussion started: 18 December 2020


Geoscientific Model Development Discussions

tudes, $\bar{u}$ is the average wind speed in the domain of simulation. (a, b, c, d) are parameters depending on the Pasquill-Gifford atmospheric vertical stability classes. There are 7 Pasquill-Gifford stability classes, from A extremely unstable (mostly in summer during the afternoon) to G very stable (occurring mostly during nighttime in winter). As the purpose of the demonstration case is primarily to work on coarsely realistic concentration fields, with a computational cost as low as possible, our implementation of the Gaussian plume model does not include any representation of particle reflection on the ground or on the top of the planetary boundary layer.

To illustrate atmospheric inversions, we use gridded surface fluxes to simulate concentrations using the Gaussian plume equation. Thus, the concentration at a given point and time t is the sum of Gaussian plume contributions from all individual grid points:

$$\begin{aligned} \mathcal{C}(x_0,\, y_0,\, z_0, t) &= \sum_{(x_1,\, y_1,\, z_1) \in \text{ grid}} \frac{f(x_1,\, y_1,\, z_1,\, t)}{2\,\pi\,.\,\sigma_y(t)\,.\,\sigma_z(t)\,.\,\bar{u}(t)} \exp\left(-\frac{y^2}{\sigma_y(t)^2}\right) \exp\left(-\frac{z^2}{\sigma_z(t)^2}\right) \\ &= \sum_{(x_1,\, y_1,\, z_1) \in \text{ grid}} \mathbf{H}_{(x_1,\, y_1,\, z_1,\, t)} \times f(x_1,\, y_1,\, z_1,\, t) \\ &= \mathbf{H}(t)\,.\,\mathbf{f}(t) \end{aligned} \tag{16}$$

This formulation highlights the linear relationship between concentrations and fluxes. As the concentrations can be expressed as a matrix product, the computation of the adjoint of the Gaussian model is straightforward and does not require extra developments:

$$\begin{aligned} \delta f(x_1,\, y_1,\, z_1,\, t) &= \sum_{(x_0,\, y_0,\, z_0) \in \text{ obs}} \frac{\delta \mathcal{C}(x_0,\, y_0,\, z_0,\, t)}{2\,\pi\,\sigma_y(t)\,\sigma_z(t)\,\bar{u}(t)} \exp\left(-\frac{y^2}{\sigma_y(t)^2}\right) \exp\left(-\frac{z^2}{\sigma_z(t)^2}\right) \\ &= \mathbf{H}(t)^T\,.\,\mathbf{C}(t) \end{aligned} \tag{17}$$

For the purpose of our demonstration cases, meteorological conditions (wind speed, wind direction, and stability class) are randomly generated for the simulation time-window. Fixed seeds are selected for the generation of random conditions in order to make them reproducible.

4.2 Configuration

4.2.1 Modelling set-up

Cases discussed in Sect. 4.3 are based on the Gaussian model computed on a domain of $2.5 \times 2\,\text{km}^2$ with a 100 m horizontal resolution. Surface point sources are located on a 100 m regular grid, with flux intensities as represented in Fig. 4. Fifty virtual measurement sites are randomly located in the domain with randomly selected altitudes above ground level as shown in Fig. 4. The inversion time-window spans a period of five days with hourly observations and meteorological forcing conditions. Meteorological conditions are a

https://doi.org/10.5194/gmd-2020-407
Preprint. Discussion started: 18 December 2020


Geoscientific Model Development Discussions

combination of a wind speed, a wind direction and a stability class applicable to the whole simulation domain for each hour. The three parameters are generated randomly for the period.

Truth observations are generated by running the Gaussian model in forward mode with the known true fluxes defined in Eq. 18, then prior fluxes are generated by perturbing the true fluxes as shown in Fig. 4 and following Eq. 19. A random noise of 1% of the standard deviation of the forward simulations was added to the truth observations to generate measurements. Prior fluxes and perturbation are generated following the equations below. Please note that the perturbation of the fluxes is generated using an explicit formula and not a random perturbation with a given covariance matrix. For that reason, we discuss results with different possible target vector and covariance matrices. This allows us to assess the sensitivity of an inversion method to the resolution and definition of the target vector and corresponding covariance matrix.

$$\mathbf{f} = f_0 \,.\, \left\{ \cos\left(2\pi \frac{x}{\sigma_x^1}\right) + \sin\left(2\pi \frac{y}{\sigma_y^1}\right) + \left(\frac{x}{\sigma_x^2}\right)^2 + \left(\frac{y}{\sigma_y^2}\right)^2 \right\} \tag{18}$$

$$\delta\mathbf{f} = 0.2 \times f_0 \,.\, \left\{ \cos\left(2\pi \frac{x}{\sigma_x^3}\right) + \sin\left(2\pi \frac{y}{\sigma_y^3}\right) \right\} \tag{19}$$

with f_0 an arbitrary reference flux, and scaling lengths σ_x^1, σ_x^2, σ_x^3, σ_y^1, σ_y^2, σ_y^3 equals 500, 1000, 200, 1000, 1000 and 300 m respectively.

4.2.2 Inversion set-ups

The objective of our test case is to prove that our system enables users to easily compare the behaviour of different inversion methods in various configurations. To do so, we conduct three sets of four inversions for the demonstration of our system. Each set includes one analytical inversion, one EnKF-based inversion and two variational inversions based on M1QN3 and CONGRAD minimization algorithms respectively. The sequential aspect of the EnKF is not implemented in the CIF, hence the comparison with the other inversion methods only includes the random sampling of the target vector distribution to solve Eq. 4. As the computation of posterior uncertainty matrices is not included for variational inversions at this point neither, posterior uncertainties are not discussed in the following.

The three sets of inversions differ by the dimension of the target vector and the spatial correlations of errors. The first set uses a target vector based on a grid of 3×3 pixels-aggregated regions independent from each other i.e. with no spatial error correlations. The target vectors of the second and third sets are defined at the grid cell's resolution with horizontal isotropic error correlations, following an exponential decay with a horizontal scale of 500 m and 10000 m respectively. For all inversion set-ups, the target vectors are defined as constant over time, i.e., only one coefficient per spatial dimension is optimized for the 5 days $\times$ 24 hours, computed by the model. In all set-ups, the magnitude of the observation noise used to generate "true"

https://doi.org/10.5194/gmd-2020-407
Preprint. Discussion started: 18 December 2020


Geoscientific Model Development Discussions

observations is chosen as observation errors in the matrix $\mathbf{R}$ for consistency. In all cases, the diagonal terms of the $\mathbf{B}$ matrix are set to 100%.

To assess the sensitivity of each set-up to the allocated computational resources, we computed the EnKF and the two variational inversions with varying numbers of simulations N. In the case of the EnKF, N simply depicts the size of the Monte Carlo ensemble. For variational inversions, each step i.e each computation of the cost function and its gradient requires one forward simulation and one adjoint simulation. The Gaussian model is a simple auto-adjoint model, which makes the adjoint simulations as long as the forward one. Therefore, N is equal to twice the number of computations of the cost function (one for the forward and one for the adjoint) in the minimization algorithm. Note that in many real application cases, the adjoint of a CTM is more costly than the forward, reducing the number of iterations possible in N times the cost of a forward. Indeed, despite the adjoint being mathematically as expensive as the forward, in practice, the computation of adjoint operations often requires the re-computation of intermediate forward computations, therefore increasing the computational burden of the adjoint model.

4.3 Results and discussion

Posterior increments for the four inversion methods in the three considered demonstration cases are presented in Fig. 5 (horizontally aggregated target vector), Fig. 6 (target vector at the pixel resolution with horizontal error correlations of 500 m) and Fig. 7 (same with correlations of 10000 m). Observation locations and heights are reported for information. The color scale of flux increments is the same as in Fig. 4 which represent the true "increments" to be retrieved. In Fig. 8, we present the evolution of the cost function of Eq. 5 depending on the number of simulations used for each inversion method for the three demonstration cases. The x-axis has been cropped at the origin as the EnKF value for small sizes of random ensembles diverges to infinity.

In the case with the target vector aggregated on groups of pixels, all inversion methods converge towards a very similar solution. In this case, the posterior increments reproduce the overall structure of the truth-prior difference, with four local maxima surrounding one local minimum in the center of the domain. However, the aggregated control vector results in too coarse patterns which are not representative of the actual true-prior difference. In the case with the target vector at the grid's resolution with spatial correlations of 500 m, the fully resolved analytical solution captures well the true-prior difference structure with four maxima surrounding a minimum. However, posterior increments are rather noisy compared to the truth. This is due to the spatial correlations being inconsistent with the smooth perturbation with fixed lengths scales in Eq. 19. Correlations help smoothing the posterior fluxes but not perfectly consistently with the truth. For the case with the target vector at the grid's resolution with spatial correlations of 10000 m, the analytical and variational inversions converge to very similar solutions. The EnKF inversion converges towards a relatively noisy

https://doi.org/10.5194/gmd-2020-407
Preprint. Discussion started: 18 December 2020


Geoscientific Model Development Discussions

solution, suggesting a possible degeneration of the algorithm; this is due to our very simple implementation of the EnKF, with no localization and a rather small number of samples in the random ensemble.

For the three cases, CONGRAD converges towards a cost function value similar to the analytical solution. It also converges at a faster pace than the two other methods and, after 20 and 50 simulations in the aggregated and full-resolution cases respectively, the convergence rates is close to zero, suggesting additional simulations do not provide significant additional information to CONGRAD. The variational inversion with M1QN3 converges towards the analytical solution only in the case of aggregated target vector. The EnKF inversion converges to the analytical solution in two of the three cases (aggregated and full-resolution with small correlation lengths). For the two full-resolution cases, M1QN3 converges to a local minimum instead of a global minimum, probably due to numerical artifacts in the algorithms which makes it mistake the global minimum with a local critical point with a gradient close to zero. This issue may be overcome by allowing the algorithm to carry on for more iterations. The convergence rate with M1QN3 is similar to the one for the EnKF inversion for full-resolution target vectors, but faster for the aggregated case.

Overall, CONGRAD appears to be the most cost-efficient algorithm to estimate the analytical solution in the case of a linear inversion in our very simple demonstration case. Though not as efficient, the EnKF method can approximate the analytical solution at reduced cost, but by design, its computation can be parallelized, which can allow a faster computation than CONGRAD when computational resources are available in parallel. M1QN3 proves not as efficient as its CONGRAD counterpart, but contrary to CONGRAD, it can accommodate non-linear cases.

5 Conclusions

We have introduced here a new generic inversion framework that aims at merging existing inversion systems together, in order to share development and maintenance efforts and to foster collaboration on inversion studies. It has been implemented in a way that is fully independent from the inversion configuration: from the application scales, from the species of interest, from the CTM used, from the assumptions for data assimilation, as well as from the practical operations and transformations applied to the data in pre- and post-processing stages. This framework will prevent redundant developments from participating research groups and will allow for a very diverse range of applications within the same system. New developments will be made in an efficient manner with limited risks of unexpected side effects, and thanks to the generic structure of the code, specific developments for a given application can be directly applied to other applications. For instance, new inversion methods implemented in the CIF can be directly tested with various transport models.

We have presented the first version of this *Community Inversion Framework* (CIF) alongside with its python-dedicated library *pyCIF*. As a first step, analytical inversions, variational inversions with M1QN3

https://doi.org/10.5194/gmd-2020-407
Preprint. Discussion started: 18 December 2020

and CONGRAD, and EnKF have been implemented to demonstrate the general applicability of the CIF. The four inversion techniques were tested here on a test case with a Gaussian Plume Model and with observations generated from known "true" fluxes. The impact of spatial correlations and of spatial aggregation, that drive the shape of the control vectors used in this scientific community, has been illustrated. The analytical inversion is the most accurate approach to retrieve the true fluxes, as expected, followed by variational inversions with the CONGRAD algorithm in our simple linear case. EnKF and M1QN3 are generally less accurate in capturing the true pattern of the fluxes in our examples, even though EnKF inversions have the advantage to be fully parallelizable, in contrast to variational inversions, that are sequential by design and therefore harder to parallelize (e.g., Chevallier, 2013).

The next step of the CIF is the implementation of a large variety of CTMs. CHIMERE, LMDz, TM5, FLEXPART, and STILT have already been implemented and a sequel paper will evaluate and compare their behaviour in similar inversion set-ups. COSMO-GHG and WRF-CHEM are also planned to be implemented in the near future to widen the developer/user community of the system. The use of various CTMs in identical inversion configurations (inversion method, observation and target vector, consistent interface, etc.) will allow extensive assessments of transport errors in inversions. Despite many past efforts put in inter-comparison exercises, such a level of inter-comparability has never been reached and will be a natural by-product of the CIF in the future. In addition, comparing posterior uncertainties from different inversion methods and set-ups will make it possible to fully assess the consistency of different inversion results.

The flexibility of the CIF allows very complex operations to be included easily. They include the use of satellite observations, that will be evaluated in a future paper, inversions using observations of isotopic ratios and optimizing both surface fluxes and source signatures (Thanwerdas et al., 2020). In addition, even though GHG studies have been the main motivation behind the development of the CIF, the system will also be tested for multi-species inversions including air pollutants.

Code and data availability. The codes, documentation pages (including installation instructions and tutorials) and demonstration data used in the present paper are registered under the following DOI: 10.5281/zenodo.4322372 (Berchet et al., 2020).

Author contributions. All authors contributed to the elaboration of the concept of Community Inversion Framework. Antoine Berchet designed the structure of the CIF. Antoine Berchet, Espen Sollum, Isabelle Pison and Joël Thanwerdas are the main developers of the python library pyCIF. Antoine Berchet, Espen Sollum and Isabelle Pison maintain the documentation and websites associated to the Community Inversion Framework.

https://doi.org/10.5194/gmd-2020-407
Preprint. Discussion started: 18 December 2020

Geoscientific Model Development Discussions
Open Access
EGU

Acknowledgements. The Community Inversion Framework is currently funded by the project VERIFY (verify.lsce.ipsl.fr), which received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no. 776810.

https://doi.org/10.5194/gmd-2020-407
Preprint. Discussion started: 18 December 2020

References

Anderson, J., Hoar, T., Raeder, K., Liu, H., Collins, N., Torn, R., and Avellano, A.: The Data Assimilation Research Testbed: A Community Facility, Bull. Amer. Meteor. Soc., 90, 1283–1296, https://doi.org/10.1175/2009BAMS2618.1, https://journals.ametsoc.org/bams/article/90/9/1283/59500/The-Data-Assimilation-Research-Testbed-A-Community, publisher: American Meteorological Society, 2009.

Ars, S., Broquet, G., Yver Kwok, C., Roustan, Y., Wu, L., Arzoumanian, E., and Bousquet, P.: Statistical atmospheric inversion of local gas emissions by coupling the tracer release technique and local-scale transport modelling: a test case with controlled methane emissions, Atmospheric Measurement Techniques, 10, 5017–5037, https://doi.org/https://doi.org/10.5194/amt-10-5017-2017, https://amt.copernicus.org/articles/10/5017/2017/, publisher: Copernicus GmbH, 2017.

Asefi-Najafabady, S., Rayner, P. J., Gurney, K. R., McRobert, A., Song, Y., Coltin, K., Huang, J., Elvidge, C., and Baugh, K.: A multiyear, global gridded fossil fuel CO2 emission data product: Evaluation and analysis of results, Journal of Geophysical Research: Atmospheres, 119, 10,213–10,231, https://doi.org/10.1002/2013JD021296, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2013JD021296, _eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1002/2013JD021296, 2014.

Babenhauserheide, A., Basu, S., Houweling, S., Peters, W., and Butz, A.: Comparing the CarbonTracker and TM5-4DVar data assimilation systems for CO_2 surface flux inversions, Atmospheric Chemistry and Physics, 15, 9747–9763, https://doi.org/https://doi.org/10.5194/acp-15-9747-2015, https://acp.copernicus.org/articles/15/9747/2015/, publisher: Copernicus GmbH, 2015.

Bagley, J. E., Jeong, S., Cui, X., Newman, S., Zhang, J., Priest, C., Campos-Pineda, M., Andrews, A. E., Bianco, L., Lloyd, M., Lareau, N., Clements, C., and Fischer, M. L.: Assessment of an atmospheric transport model for annual inverse estimates of California greenhouse gas emissions, J. Geophys. Res. Atmos., p. 2016JD025361, https://doi.org/10.1002/2016JD025361, http://onlinelibrary.wiley.com/doi/10.1002/2016JD025361/abstract, 2017.

Belikov, D. A., Maksyutov, S., Yaremchuk, A., Ganshin, A., Kaminski, T., Blessing, S., Sasakawa, M., Gomez-Pelaez, A. J., and Starchenko, A.: Adjoint of the global Eulerian–Lagrangian coupled atmospheric transport model (A-GELCA v1.0): development and validation, Geoscientific Model Development, 9, 749–764, https://doi.org/https://doi.org/10.5194/gmd-9-749-2016, https://gmd.copernicus.org/articles/9/749/2016/, publisher: Copernicus GmbH, 2016.

Berchet, A., Pison, I., Chevallier, F., Bousquet, P., Bonne, J.-L., and Paris, J.-D.: Objectified quantification of uncertainties in Bayesian atmospheric inversions, Geosci. Model Dev. Discuss., 7, 4777–4827, https://doi.org/10.5194/gmdd-7-4777-2014, http://www.geosci-model-dev-discuss.net/7/4777/2014/, 2014.

Berchet, A., Sollum, E., Pison, I., Thompson, R. L., Thanwerdas, J., Fortems-Cheiney, A., Peet, J. C. A. v., Potier, E., Chevallier, F., and Broquet, G.: The Community Inversion Framework: codes and documentation, https://doi.org/10.5281/zenodo.4322372, https://doi.org/10.5281/zenodo.4322372, 2020.

Bergamaschi, P., Krol, M., Meirink, J. F., Dentener, F., Segers, A., van Aardenne, J., Monni, S., Vermeulen, A. T., Schmidt, M., Ramonet, M., Yver, C., Meinhardt, F., Nisbet, E. G., Fisher, R. E., O'Doherty, S., and Dlugokencky, E. J.: Inverse modeling of European CH$_4$ emissions 2001-2006, J. Geophys.

https://doi.org/10.5194/gmd-2020-407
Preprint. Discussion started: 18 December 2020

Res., 115, https://doi.org/10.1029/2010JD014180, http://onlinelibrary.wiley.com.biblioplanets.gate.inist.fr/doi/10.1029/2010JD014180/abstract, 2010.

Bergamaschi, P., Danila, A., Weiss, R. F., Ciais, P., Thompson, R. L., Brunner, D., Levin, I., Meijer, Y., Chevallier, F., Janssens-Maenhout, G., Bovensmann, H., Crisp, D., Basu, S., Dlugokencky, E. J., Engelen, R., Gerbig, C., G\"unther, D., Hammer, S., Henne, S., Houweling, S., Karstens, U., Kort, E. A., Maione, M., Manning, A. J., Miller, J. B., Montzka, S. A., Pandey, S., Peters, W., Peylin, P., Pinty, B., Ramonet, M., Reimann, S., R\"ockmann, T., Schmidt, M., Strogies, M., Sussams, J., Tarasova, O. A., van Aardenne, J., Vermeulen, A. T., and Vogel, F.: Atmospheric monitoring and inverse modelling for verification of greenhouse gas inventories, EUR 29276 EN JRC111789, Publications Office of the European Union, Luxembourg, https://ec.europa.eu/jrc/en/publication/eur-scientific-and-technical-research-reports/atmospheric-monitoring-and-inverse-modelling-verification-greenhouse-gas-inventories, 2018a.

Bergamaschi, P., Karstens, U., Manning, A. J., Saunois, M., Tsuruta, A., Berchet, A., Vermeulen, A. T., Arnold, T., Janssens-Maenhout, G., Hammer, S., Levin, I., Schmidt, M., Ramonet, M., Lopez, M., Lavric, J., Aalto, T., Chen, H., Feist, D. G., Gerbig, C., Haszpra, L., Hermansen, O., Manca, G., Moncrieff, J., Meinhardt, F., Necki, J., Galkowski, M., O'Doherty, S., Paramonova, N., Scheeren, H. A., Steinbacher, M., and Dlugokencky, E.: Inverse modelling of European CH_4 emissions during 2006–2012 using different inverse models and reassessed atmospheric observations, Atmospheric Chemistry and Physics, 18, 901–920, https://doi.org/https://doi.org/10.5194/acp-18-901-2018, https://acp.copernicus.org/articles/18/901/2018/, publisher: Copernicus GmbH, 2018b.

Bocquet, M.: Ensemble Kalman filtering without the intrinsic need for inflation, Nonlin. Processes Geophys., 18, 735–750, https://doi.org/10.5194/npg-18-735-2011, http://www.nonlin-processes-geophys.net/18/735/2011/, 2011.

Bocquet, M., Elbern, H., Eskes, H., Hirtl, M., Žabkar, R., Carmichael, G. R., Flemming, J., Inness, A., Pagowski, M., Pérez Camaño, J. L., Saide, P. E., San Jose, R., Sofiev, M., Vira, J., Baklanov, A., Carnevale, C., Grell, G., and Seigneur, C.: Data assimilation in atmospheric chemistry models: current status and future prospects for coupled chemistry meteorology models, Atmospheric Chemistry and Physics, 15, 5325–5358, https://doi.org/https://doi.org/10.5194/acp-15-5325-2015, https://acp.copernicus.org/articles/15/5325/2015/, publisher: Copernicus GmbH, 2015.

Bousquet, P., Ringeval, B., Pison, I., Dlugokencky, E. J., Brunke, E. G., Carouge, C., Chevallier, F., Fortems-Cheiney, A., Frankenberg, C., Hauglustaine, D. A., et al.: Source attribution of the changes in atmospheric methane for 2006-2008, Atmos. Chem. Phys., 11, 3689–3700, http://www.atmos-chem-phys.org/11/3689/2011/acp-11-3689-2011.pdf, 2011.

Bousserez, N. and Henze, D. K.: Optimal and scalable methods to approximate the solutions of large-scale Bayesian problems: theory and application to atmospheric inversion and data assimilation, Quarterly Journal of the Royal Meteorological Society, 144, 365–390, https://doi.org/10.1002/qj.3209, https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/qj.3209, 2018.

Bousserez, N., Henze, D. K., Perkins, A., Bowman, K. W., Lee, M., Liu, J., Deng, F., and Jones, D. B. A.: Improved analysis-error covariance matrix for high-dimensional variational inversions: application to source estimation using a 3D atmospheric transport model, Quarterly Journal of the Royal Meteorological Society, 141, 1906–1921, https://doi.org/10.1002/qj.2495, https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/qj.2495, _eprint: https://rmets.onlinelibrary.wiley.com/doi/pdf/10.1002/qj.2495, 2015.

https://doi.org/10.5194/gmd-2020-407
Preprint. Discussion started: 18 December 2020

Brioude, J., Angevine, W. M., Ahmadov, R., Kim, S.-W., Evan, S., McKeen, S. A., Hsie, E.-Y., Frost, G. J., Neuman, J. A., Pollack, I. B., Peischl, J., Ryerson, T. B., Holloway, J., Brown, S. S., Nowak, J. B., Roberts, J. M., Wofsy, S. C., Santoni, G. W., Oda, T., and Trainer, M.: Top-down estimate of surface flux in the Los Angeles Basin using a mesoscale inverse modeling technique: assessing anthropogenic emissions of CO, NO_x and CO_2 and their impacts, Atmospheric Chemistry and Physics, 13, 3661–3677, https://doi.org/https://doi.org/10.5194/acp-13-3661-2013, https://acp.copernicus.org/articles/13/3661/2013/, publisher: Copernicus GmbH, 2013.

Broquet, G., Chevallier, F., Rayner, P., Aulagnier, C., Pison, I., Ramonet, M., Schmidt, M., Vermeulen, A. T., and Ciais, P.: A European summertime CO$_2$ biogenic flux inversion at mesoscale from continuous in situ mixing ratio measurements, Journal of Geophysical Research: Atmospheres, 116, n/a–n/a, https://doi.org/10.1029/2011JD016202, http://onlinelibrary.wiley.com/doi/10.1029/2011JD016202/abstract, 2011.

Bruhwiler, L. M. P., Michalak, A. M., Peters, W., Baker, D. F., and Tans, P.: An improved Kalman Smoother for atmospheric inversions, Atmospheric Chemistry and Physics, 5, 2691–2702, https://doi.org/10.5194/acp-5-2691-2005, http://atmos-chem-phys.net/5/2691/2005/acp-5-2691-2005.html, 2005.

Brunner, D., Henne, S., Keller, C. A., Reimann, S., Vollmer, M. K., O'Doherty, S., and Maione, M.: An extended Kalman-filter for regional scale inverse emission estimation, Atmospheric Chemistry and Physics, 12, 3455–3478, https://doi.org/https://doi.org/10.5194/acp-12-3455-2012, https://acp.copernicus.org/articles/12/3455/2012/, publisher: Copernicus GmbH, 2012.

Brunner, D., Arnold, T., Henne, S., Manning, A., Thompson, R. L., Maione, M., O'Doherty, S., and Reimann, S.: Comparison of four inverse modelling systems applied to the estimation of HFC-125, HFC-134a, and SF_6 emissions over Europe, Atmospheric Chemistry and Physics, 17, 10 651–10 674, https://doi.org/https://doi.org/10.5194/acp-17-10651-2017, https://acp.copernicus.org/articles/17/10651/2017/, publisher: Copernicus GmbH, 2017.

Bréon, F. M., Broquet, G., Puygrenier, V., Chevallier, F., Xueref-Remy, I., Ramonet, M., Dieudonné, E., Lopez, M., Schmidt, M., Perrussel, O., and Ciais, P.: An attempt at estimating Paris area CO2 emissions from atmospheric concentration measurements, Atmos. Chem. Phys., 15, 1707–1724, https://doi.org/10.5194/acp-15-1707-2015, http://www.atmos-chem-phys.net/15/1707/2015/, 2015.

Chatterjee, A., Michalak, A. M., Anderson, J. L., Mueller, K. L., and Yadav, V.: Toward reliable ensemble Kalman filter estimates of CO2 fluxes, Journal of Geophysical Research: Atmospheres, 117, https://doi.org/https://doi.org/10.1029/2012JD018176, http://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2012JD018176, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1029/2012JD018176, 2012.

Chevallier, F.: On the parallelization of atmospheric inversions of CO2 surface fluxes within a variational framework, Geosci. Model Dev., 6, 783–790, https://doi.org/10.5194/gmd-6-783-2013, https://www.geosci-model-dev.net/6/783/2013/, 2013.

Chevallier, F., Fisher, M., Peylin, P., Serrar, S., Bousquet, P., Bréon, F. M., Chédin, A., and Ciais, P.: Inferring CO$_2$ sources and sinks from satellite observations: Method and application to TOVS data, J. Geophys. Res., 110, D24 309, https://doi.org/10.1029/2005JD006390, http://www.agu.org/pubs/crossref/2005/2005JD006390.shtml, 2005.

Chevallier, F., Ciais, P., Conway, T. J., Aalto, T., Anderson, B. E., Bousquet, P., Brunke, E. G., Ciattaglia, L., Esaki, Y., Fröhlich, M., Gomez, A., Gomez-Pelaez, A. J., Haszpra, L., Krummel, P. B., Langenfelds, R. L., Leuenberger, M., Machida, T., Maignan, F., Matsueda, H., Morguí, J. A., Mukai, H., Nakazawa, T., Peylin, P., Ramonet, M., Rivier, L.,

https://doi.org/10.5194/gmd-2020-407
Preprint. Discussion started: 18 December 2020

Sawa, Y., Schmidt, M., Steele, L. P., Vay, S. A., Vermeulen, A. T., Wofsy, S., and Worthy, D.: CO2 surface fluxes at grid point scale estimated from a global 21 year reanalysis of atmospheric measurements, J. Geophys. Res., 115, D21 307, https://doi.org/10.1029/2010JD013887, http://onlinelibrary.wiley.com/doi/10.1029/2010JD013887/abstract, 2010.

Chevallier, F., Remaud, M., O'Dell, C. W., Baker, D., Peylin, P., and Cozic, A.: Objective evaluation of surface- and satellite-driven carbon dioxide atmospheric inversions, Atmospheric Chemistry and Physics, 19, 14 233–14 251, https://doi.org/https://doi.org/10.5194/acp-19-14233-2019, https://acp.copernicus.org/articles/19/14233/2019/, publisher: Copernicus GmbH, 2019.

Ciais, P., Dolman, A. J., Bombelli, A., Duren, R., Peregon, A., Rayner, P. J., Miller, C., Gobron, N., Kinderman, G., Marland, G., Gruber, N., Chevallier, F., Andres, R. J., Balsamo, G., Bopp, L., Bréon, F.-M., Broquet, G., Dargaville, R., Battin, T. J., Borges, A., Bovensmann, H., Buchwitz, M., Butler, J., Canadell, J. G., Cook, R. B., DeFries, R., Engelen, R., Gurney, K. R., Heinze, C., Heimann, M., Held, A., Henry, M., Law, B., Luyssaert, S., Miller, J., Moriyama, T., Moulin, C., Myneni, R. B., Nussli, C., Obersteiner, M., Ojima, D., Pan, Y., Paris, J.-D., Piao, S. L., Poulter, B., Plummer, S., Quegan, S., Raymond, P., Reichstein, M., Rivier, L., Sabine, C., Schimel, D., Tarasova, O., Valentini, R., Wang, R., van der Werf, G., Wickland, D., Williams, M., and Zehner, C.: Current systematic carbon-cycle observations and the need for implementing a policy-relevant carbon observing system, Biogeosciences, 11, 3547–3602, https://doi.org/https://doi.org/10.5194/bg-11-3547-2014, https://bg.copernicus.org/articles/11/3547/2014/, publisher: Copernicus GmbH, 2014.

Crisp, D., Meijer, Y., Munro, R., Bowman, K., Chatterjee, A., Baker, D., Chevallier, F., Nassar, R., Palmer, P. I., Agusti-Panareda, A., and others: A constellation architecture for monitoring carbon dioxide and methane from space, Prepared by the CEOS Atmospheric Constellation Greenhouse Gas Team, Version, 1, 2018.

Crowell, S., Baker, D., Schuh, A., Basu, S., Jacobson, A. R., Chevallier, F., Liu, J., Deng, F., Feng, L., McKain, K., Chatterjee, A., Miller, J. B., Stephens, B. B., Eldering, A., Crisp, D., Schimel, D., Nassar, R., O'Dell, C. W., Oda, T., Sweeney, C., Palmer, P. I., and Jones, D. B. A.: The 2015–2016 carbon cycle as seen from OCO-2 and the global in situ network, Atmospheric Chemistry and Physics, 19, 9797–9831, https://doi.org/https://doi.org/10.5194/acp-19-9797-2019, https://acp.copernicus.org/articles/19/9797/2019/, publisher: Copernicus GmbH, 2019.

Curier, R. L., Timmermans, R., Calabretta-Jongen, S., Eskes, H., Segers, A., Swart, D., and Schaap, M.: Improving ozone forecasts over Europe by synergistic use of the LOTOS-EUROS chemical transport model and in-situ measurements, Atmospheric Environment, 60, 217–226, https://doi.org/10.1016/j.atmosenv.2012.06.017, http://www.sciencedirect.com/science/article/pii/S1352231012005687, 2012.

Enting, I. G.: Inverse problems in atmospheric constituent transport, Cambridge University Press, http://books.google.fr/books?hl=fr&lr=&id=b61nA7D1XXUC&oi=fnd&pg=PP1&dq=enting+inverse+problem+atmospheric+constituent+transport+&ots=vc8hwiQ_YK&sig=u5npCEtl4WdBNLyRDn9V2EfD1SQ, 2002.

Errico, R. M.: What Is an Adjoint Model?, Bull. Amer. Meteor. Soc., 78, 2577–2592, https://doi.org/10.1175/1520-0477(1997)078<2577:WIAAM>2.0.CO;2, https://journals.ametsoc.org/bams/article/78/11/2577/55799/What-Is-an-Adjoint-Model, publisher: American Meteorological Society, 1997.

Feng, L., Palmer, P. I., Bösch, H., and Dance, S.: Estimating surface CO_2 fluxes from space-borne CO_2 dry air mole fraction observations using an ensemble Kalman Filter, Atmospheric Chemistry and Physics, 9, 2619–2633,

https://doi.org/10.5194/gmd-2020-407
Preprint. Discussion started: 18 December 2020

https://doi.org/https://doi.org/10.5194/acp-9-2619-2009, https://acp.copernicus.org/articles/9/2619/2009/, publisher: Copernicus GmbH, 2009.

Feng, L., Palmer, P. I., Bösch, H., Parker, R. J., Webb, A. J., Correia, C. S. C., Deutscher, N. M., Domingues, L. G., Feist, D. G., Gatti, L. V., Gloor, E., Hase, F., Kivi, R., Liu, Y., Miller, J. B., Morino, I., Sussmann, R., Strong, K., Uchino, O., Wang, J., and Zahn, A.: Consistent regional fluxes of CH_4 and CO_2 inferred from GOSAT proxy $XCH_4 : XCO_2$ retrievals, 2010–2014, Atmospheric Chemistry and Physics, 17, 4781–4797, https://doi.org/https://doi.org/10.5194/acp-17-4781-2017, https://acp.copernicus.org/articles/17/4781/2017/, publisher: Copernicus GmbH, 2017.

Filges, A., Gerbig, C., Chen, H., Franke, H., Klaus, C., and Jordan, A.: The IAGOS-core greenhouse gas package: a measurement system for continuous airborne observations of CO2, CH4, H2O and CO, Tellus B: Chemical and Physical Meteorology, 68, 27 989, https://doi.org/10.3402/tellusb.v67.27989, https://www.tandfonline.com/doi/abs/10.3402/tellusb.v67.27989%40zelb20.2016.68.issue-s1, publisher: Taylor & Francis _eprint: https://www.tandfonline.com/doi/pdf/10.3402/tellusb.v67.27989%40zelb20.2016.68.issue-s1, 2016.

Fisher, M.: Minimization algorithms for variational data assimilation, in: Seminar on Recent Developments in Numerical Methods for Atmospheric Modelling, 7-11 September 1998, pp. 364–385, ECMWF, Shinfield Park, Reading, https://www.ecmwf.int/node/9400, backup Publisher: ECMWF, 1998.

Fortems-Cheiney, A., Pison, I., Dufour, G., Broquet, G., Berchet, A., Potier, E., Coman, A., Siour, G., and Costantino, L.: Variational regional inverse modeling of reactive species emissions with PYVAR-CHIMERE, Geoscientific Model Development Discussions, pp. 1–22, https://doi.org/https://doi.org/10.5194/gmd-2019-186, https://www.geosci-model-dev-discuss.net/gmd-2019-186/, 2019.

Foster-Wittig, T. A., Thoma, E. D., and Albertson, J. D.: Estimation of point source fugitive emission rates from a single sensor time series: A conditionally-sampled Gaussian plume reconstruction, Atmospheric Environment, 115, 101–109, https://doi.org/10.1016/j.atmosenv.2015.05.042, http://www.sciencedirect.com/science/article/pii/S135223101530114X, 2015.

Ganesan, A. L., Rigby, M., Zammit-Mangion, A., Manning, A. J., Prinn, R. G., Fraser, P. J., Harth, C. M., Kim, K.-R., Krummel, P. B., Li, S., Mühle, J., O'Doherty, S. J., Park, S., Salameh, P. K., Steele, L. P., and Weiss, R. F.: Characterization of uncertainties in atmospheric trace gas inversions using hierarchical Bayesian methods, Atmos. Chem. Phys., 14, 3855–3864, https://doi.org/10.5194/acp-14-3855-2014, http://www.atmos-chem-phys.net/14/3855/2014/, 2014.

Gilbert, J. C. and Lemaréchal, C.: Some numerical experiments with variable-storage quasi-Newton algorithms, Mathematical Programming, 45, 407–435, https://doi.org/10.1007/BF01589113, https://link.springer.com/article/10.1007/BF01589113, 1989.

Gurney, K. R., Law, R. M., Denning, A. S., Rayner, P. J., Baker, D., Bousquet, P., Bruhwiler, L., Chen, Y.-H., Ciais, P., Fan, S., Fung, I. Y., Gloor, M., Heimann, M., Higuchi, K., John, J., Kowalczyk, E., Maki, T., Maksyutov, S., Peylin, P., Prather, M., Pak, B. C., Sarmiento, J., Taguchi, S., Takahashi, T., and Yuen, C.-W.: TransCom 3 CO2 inversion intercomparison: 1. Annual mean control results and sensitivity to transport and prior flux information, Tellus B, 55, 555–579, https://doi.org/10.1034/j.1600-0889.2003.00049.x, http://www.tellusb.net/index.php/tellusb/article/view/16728, 2003.

https://doi.org/10.5194/gmd-2020-407
Preprint. Discussion started: 18 December 2020

Hanna, S. R., Briggs, G. A., and Hosker, J.: Handbook on atmospheric diffusion, Tech. Rep. DOE/TIC-11223, National Oceanic and Atmospheric Administration, Oak Ridge, TN (USA). Atmospheric Turbulence and Diffusion Lab., https://doi.org/10.2172/5591108, https://www.osti.gov/biblio/5591108, 1982.

Henne, S., Brunner, D., Oney, B., Leuenberger, M., Eugster, W., Bamberger, I., Meinhardt, F., Steinbacher, M., and Emmenegger, L.: Validation of the Swiss methane emission inventory by atmospheric observations and inverse modelling, Atmos. Chem. Phys., 16, 3683–3710, https://doi.org/10.5194/acp-16-3683-2016, http://www.atmos-chem-phys.net/16/3683/2016/, 2016.

Houweling, S., Badawy, B., Baker, D. F., Basu, S., Belikov, D., Bergamaschi, P., Bousquet, P., Broquet, G., Butler, T., Canadell, J. G., Chen, J., Chevallier, F., Ciais, P., Collatz, G. J., Denning, S., Engelen, R., Enting, I. G., Fischer, M. L., Fraser, A., Gerbig, C., Gloor, M., Jacobson, A. R., Jones, D. B. A., Heimann, M., Khalil, A., Kaminski, T., Kasibhatla, P. S., Krakauer, N. Y., Krol, M., Maki, T., Maksyutov, S., Manning, A., Meesters, A., Miller, J. B., Palmer, P. I., Patra, P., Peters, W., Peylin, P., Poussi, Z., Prather, M. J., Randerson, J. T., Röckmann, T., Rödenbeck, C., Sarmiento, J. L., Schimel, D. S., Scholze, M., Schuh, A., Suntharalingam, P., Takahashi, T., Turnbull, J., Yurganov, L., and Vermeulen, A.: Iconic CO2 Time Series at Risk, Science, 337, 1038–1040, https://doi.org/10.1126/science.337.6098.1038-b, https://science.sciencemag.org/content/337/6098/1038.2, publisher: American Association for the Advancement of Science Section: Letters, 2012.

Houweling, S., Krol, M., Bergamaschi, P., Frankenberg, C., Dlugokencky, E. J., Morino, I., Notholt, J., Sherlock, V., Wunch, D., Beck, V., Gerbig, C., Chen, H., Kort, E. A., Röckmann, T., and Aben, I.: A multi-year methane inversion using SCIAMACHY, accounting for systematic errors using TCCON measurements, Atmos. Chem. Phys., 14, 3991–4012, https://doi.org/10.5194/acp-14-3991-2014, http://www.atmos-chem-phys.net/14/3991/2014/, 2014.

Houweling, S., Bergamaschi, P., Chevallier, F., Heimann, M., Kaminski, T., Krol, M., Michalak, A. M., and Patra, P.: Global inverse modeling of CH₄ sources and sinks: An overview of methods, Atmospheric Chemistry and Physics Discussions, pp. 1–30, https://doi.org/10.5194/acp-2016-572, http://www.atmos-chem-phys-discuss.net/acp-2016-572/, 2016.

Houweling, S., Bergamaschi, P., Chevallier, F., Heimann, M., Kaminski, T., Krol, M., Michalak, A. M., and Patra, P.: Global inverse modeling of CH_4 sources and sinks: an overview of methods, Atmospheric Chemistry and Physics, 17, 235–256, https://doi.org/https://doi.org/10.5194/acp-17-235-2017, https://acp.copernicus.org/articles/17/235/2017/, publisher: Copernicus GmbH, 2017.

Ide, K., Courtier, P., Ghil, M., and Lorenc, A. C.: Unified notation for data assimilation: operational, sequential and variational, J. Meteorol. Soc. Japan, pp. 181–189, http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.47.2607&rep=rep1&type=pdf, 1997.

Janssens-Maenhout, G., Pinty, B., Dowell, M., Zunker, H., Andersson, E., Balsamo, G., Bézy, J.-L., Brunhes, T., Bösch, H., Bojkov, B., Brunner, D., Buchwitz, M., Crisp, D., Ciais, P., Counet, P., Dee, D., Denier van der Gon, H., Dolman, H., Drinkwater, M. R., Dubovik, O., Engelen, R., Fehr, T., Fernandez, V., Heimann, M., Holmlund, K., Houweling, S., Husband, R., Juvyns, O., Kentarchos, A., Landgraf, J., Lang, R., Löscher, A., Marshall, J., Meijer, Y., Nakajima, M., Palmer, P. I., Peylin, P., Rayner, P., Scholze, M., Sierk, B., Tamminen, J., and Veefkind, P.: Toward an Operational Anthropogenic CO2 Emissions Monitoring and Verification Support Capacity, Bull. Amer. Meteor. Soc., 101, E1439–

https://doi.org/10.5194/gmd-2020-407
Preprint. Discussion started: 18 December 2020

E1451, https://doi.org/10.1175/BAMS-D-19-0017.1, https://journals.ametsoc.org/bams/article/101/8/E1439/345576/Toward-an-Operational-Anthropogenic-CO2-Emissions, publisher: American Meteorological Society, 2020.

Karion, A., Sweeney, C., Tans, P., and Newberger, T.: AirCore: An Innovative Atmospheric Sampling System, J. Atmos. Oceanic Technol., 27, 1839–1853, https://doi.org/10.1175/2010JTECHA1448.1, https://journals.ametsoc.org/jtech/article/27/11/1839/3375/AirCore-An-Innovative-Atmospheric-Sampling-System, publisher: American Meteorological Society, 2010.

Kopacz, M., Jacob, D. J., Henze, D. K., Heald, C. L., Streets, D. G., and Zhang, Q.: Comparison of adjoint and analytical Bayesian inversion methods for constraining Asian sources of carbon monoxide using satellite (MOPITT) measurements of CO columns, J. Geophys. Res., 114, D04 305, https://doi.org/10.1029/2007JD009264, http://onlinelibrary.wiley.com/doi/10.1029/2007JD009264/abstract, 2009.

Krol, M., Houweling, S., Bregman, B., van den Broek, M., Segers, A., van Velthoven, P., Peters, W., Dentener, F., and Bergamaschi, P.: The two-way nested global chemistry-transport zoom model TM5: algorithm and applications, Atmospheric Chemistry and Physics, 5, 417–432, https://doi.org/https://doi.org/10.5194/acp-5-417-2005, https://acp.copernicus.org/articles/5/417/2005/acp-5-417-2005.html, publisher: Copernicus GmbH, 2005.

Kumar, P., Broquet, G., Yver-Kwok, C., Laurent, O., Gichuki, S., Caldow, C., Cropley, F., Lauvaux, T., Ramonet, M., Berthe, G., Martin, F., Duclaux, O., Juery, C., Bouchet, C., and Ciais, P.: Mobile atmospheric measurements and local-scale inverse estimation of the location and rates of brief CH_4 and CO_2 releases from point sources, Atmospheric Measurement Techniques Discussions, pp. 1–27, https://doi.org/https://doi.org/10.5194/amt-2020-226, https://amt.copernicus.org/preprints/amt-2020-226/, publisher: Copernicus GmbH, 2020.

Liu, J., Bowman, K. W., and Henze, D. K.: Source-receptor relationships of column-average CO2 and implications for the impact of observations on flux inversions, Journal of Geophysical Research: Atmospheres, 120, 5214–5236, https://doi.org/10.1002/2014JD022914, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2014JD022914, _eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1002/2014JD022914, 2015.

Locatelli, R., Bousquet, P., Chevallier, F., Fortems-Cheney, A., Szopa, S., Saunois, M., Agusti-Panareda, A., Bergmann, D., Bian, H., Cameron-Smith, P., Chipperfield, M. P., Gloor, E., Houweling, S., Kawa, S. R., Krol, M., Patra, P. K., Prinn, R. G., Rigby, M., Saito, R., and Wilson, C.: Impact of transport model errors on the global and regional methane emissions estimated by inverse modelling, Atmospheric Chemistry and Physics Discussions, 13, 10 961–11 021, https://doi.org/10.5194/acpd-13-10961-2013, http://www.atmos-chem-phys-discuss.net/13/10961/2013/acpd-13-10961-2013.html, 2013.

Lunt, M. F., Rigby, M., Ganesan, A. L., and Manning, A. J.: Estimation of trace gas fluxes with objectively determined basis functions using reversible-jump Markov chain Monte Carlo, Geosci. Model Dev., 9, 3213–3229, https://doi.org/10.5194/gmd-9-3213-2016, http://www.geosci-model-dev.net/9/3213/2016/, 2016.

Masarie, K. A., Peters, W., Jacobson, A. R., and Tans, P. P.: ObsPack: a framework for the preparation, delivery, and attribution of atmospheric greenhouse gas measurements, Earth System Science Data, 6, 375–384, https://doi.org/https://doi.org/10.5194/essd-6-375-2014, https://www.earth-syst-sci-data.net/6/375/2014/essd-6-375-2014.html, 2014.

Meirink, J. F., Bergamaschi, P., and Krol, M. C.: Four-dimensional variational data assimilation for inverse modelling of atmospheric methane emissions: method and comparison with synthesis inversion, Atmospheric Chemistry and

Physics, 8, 6341–6353, https://doi.org/https://doi.org/10.5194/acp-8-6341-2008, https://acp.copernicus.org/articles/8/6341/2008/acp-8-6341-2008.html, publisher: Copernicus GmbH, 2008.

Michalak, A. M.: Technical Note: Adapting a fixed-lag Kalman smoother to a geostatistical atmospheric inversion framework, Atmospheric Chemistry and Physics, 8, 6789–6799, https://doi.org/https://doi.org/10.5194/acp-8-6789-2008, https://acp.copernicus.org/articles/8/6789/2008/acp-8-6789-2008.html, publisher: Copernicus GmbH, 2008.

Michalak, A. M. and Kitanidis, P. K.: A method for the interpolation of nonnegative functions with an application to contaminant load estimation, Stoch Environ Res Ris Assess, 19, 8–23, https://doi.org/10.1007/s00477-004-0189-1, http://link.springer.com/article/10.1007/s00477-004-0189-1, 2005.

Michalak, A. M., Hirsch, A., Bruhwiler, L., Gurney, K. R., Peters, W., and Tans, P. P.: Maximum likelihood estimation of covariance parameters for Bayesian atmospheric trace gas surface flux inversions, Journal of Geophysical Research: Atmospheres, 110, n/a–n/a, https://doi.org/10.1029/2005JD005970, http://onlinelibrary.wiley.com/doi/10.1029/2005JD005970/abstract, 2005.

Miller, S. M., Michalak, A. M., and Levi, P. J.: Atmospheric inverse modeling with known physical bounds: an example from trace gas emissions, Geosci. Model Dev., 7, 303–315, https://doi.org/10.5194/gmd-7-303-2014, http://www.geosci-model-dev.net/7/303/2014/, 2014.

Miller, S. M., Saibaba, A. K., Trudeau, M. E., Mountain, M. E., and Andrews, A. E.: Geostatistical inverse modeling with very large datasets: an example from the OCO-2 satellite, Geoscientific Model Development Discussions, pp. 1–25, https://doi.org/https://doi.org/10.5194/gmd-2019-185, https://www.geosci-model-dev-discuss.net/gmd-2019-185/, 2019.

Mizzi, A. P., Arellano Jr., A. F., Edwards, D. P., Anderson, J. L., and Pfister, G. G.: Assimilating compact phase space retrievals of atmospheric composition with WRF-Chem/DART: a regional chemical transport/ensemble Kalman filter data assimilation system, Geoscientific Model Development, 9, 965–978, https://doi.org/https://doi.org/10.5194/gmd-9-965-2016, https://gmd.copernicus.org/articles/9/965/2016/, publisher: Copernicus GmbH, 2016.

Monteil, G. and Scholze, M.: Regional CO_2 inversions with LUMIA, the Lund University Modular Inversion Algorithm, v1.0, Geoscientific Model Development Discussions, pp. 1–35, https://doi.org/https://doi.org/10.5194/gmd-2019-227, https://www.geosci-model-dev-discuss.net/gmd-2019-227/, 2019.

Monteil, G., Broquet, G., Scholze, M., Lang, M., Karstens, U., Gerbig, C., Koch, F.-T., Smith, N. E., Thompson, R. L., van der Laan-Luijkx, I. T., White, E., Meesters, A., Ciais, P., Ganesan, A. L., Manning, A., Mischurow, M., Peters, W., Peylin, P., Tarniewicz, J., Rigby, M., Rödenbeck, C., Vermeulen, A., and Walton, E. M.: The regional EUROpean atmospheric transport inversion COMparison, EUROCOM: first results on European wide terrestrial carbon fluxes for the period 2006–2015, Atmospheric Chemistry and Physics Discussions, pp. 1–40, https://doi.org/https://doi.org/10.5194/acp-2019-1008, https://acp.copernicus.org/preprints/acp-2019-1008/, publisher: Copernicus GmbH, 2019.

Nassar, R., Hill, T. G., McLinden, C. A., Wunch, D., Jones, D. B. A., and Crisp, D.: Quantifying CO2 Emissions From Individual Power Plants From Space, Geophysical Research Letters, 44, 10,045–10,053, https://doi.org/10.1002/2017GL074702, http://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2017GL074702, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/2017GL074702, 2017.

https://doi.org/10.5194/gmd-2020-407
Preprint. Discussion started: 18 December 2020

Niwa, Y., Tomita, H., Satoh, M., Imasu, R., Sawa, Y., Tsuboi, K., Matsueda, H., Machida, T., Sasakawa, M., Belan, B., and Saigusa, N.: A 4D-Var inversion system based on the icosahedral grid model (NICAM-TM 4D-Var v1.0) – Part 1: Offline forward and adjoint transport models, Geoscientific Model Development, 10, 1157–1174, https://doi.org/https://doi.org/10.5194/gmd-10-1157-2017, https://gmd.copernicus.org/articles/10/1157/2017/gmd-10-1157-2017.html, publisher: Copernicus GmbH, 2017.

Palmer, P. I., Suntharalingam, P., Jones, D. B. A., Jacob, D. J., Streets, D. G., Fu, Q., Vay, S. A., and Sachse, G. W.: Using CO2:CO correlations to improve inverse analyses of carbon fluxes, Journal of Geophysical Research: Atmospheres, 111, https://doi.org/https://doi.org/10.1029/2005JD006697, http://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2005JD006697, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1029/2005JD006697, 2006.

Palmer, P. I., Feng, L., Baker, D., Chevallier, F., Bösch, H., and Somkuti, P.: Net carbon emissions from African biosphere dominate pan-tropical atmospheric CO 2 signal, Nature Communications, 10, 3344, https://doi.org/10.1038/s41467-019-11097-w, https://www.nature.com/articles/s41467-019-11097-w, number: 1 Publisher: Nature Publishing Group, 2019.

Peters, W., Miller, J. B., Whitaker, J., Denning, A. S., Hirsch, A., Krol, M. C., Zupanski, D., Bruhwiler, L., and Tans, P. P.: An ensemble data assimilation system to estimate CO2 surface fluxes from atmospheric trace gas observations, Journal of Geophysical Research: Atmospheres, 110, https://doi.org/10.1029/2005JD006157, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2005JD006157, 2005.

Peters, W., van der Velde, I. R., van Schaik, E., Miller, J. B., Ciais, P., Duarte, H. F., van der Laan-Luijkx, I. T., van der Molen, M. K., Scholze, M., Schaefer, K., Vidale, P. L., Verhoef, A., Wårlind, D., Zhu, D., Tans, P. P., Vaughn, B., and White, J. W. C.: Increased water-use efficiency and reduced CO 2 uptake by plants during droughts at a continental scale, Nature Geoscience, 11, 744–748, https://doi.org/10.1038/s41561-018-0212-7, https://www.nature.com/articles/s41561-018-0212-7, number: 10 Publisher: Nature Publishing Group, 2018.

Peylin, P., Law, R. M., Gurney, K. R., Chevallier, F., Jacobson, A. R., Maki, T., Niwa, Y., Patra, P. K., Peters, W., Rayner, P. J., Rödenbeck, C., van der Laan-Luijkx, I. T., and Zhang, X.: Global atmospheric carbon budget: results from an ensemble of atmospheric CO$_2$ inversions, Biogeosciences, 10, 6699–6720, https://doi.org/10.5194/bg-10-6699-2013, http://www.biogeosciences.net/10/6699/2013/, 2013.

Pisso, I., Sollum, E., Grythe, H., Kristiansen, N. I., Cassiani, M., Eckhardt, S., Arnold, D., Morton, D., Thompson, R. L., Groot Zwaaftink, C. D., Evangeliou, N., Sodemann, H., Haimberger, L., Henne, S., Brunner, D., Burkhart, J. F., Fouilloux, A., Brioude, J., Philipp, A., Seibert, P., and Stohl, A.: The Lagrangian particle dispersion model FLEXPART version 10.4, Geoscientific Model Development, 12, 4955–4997, https://doi.org/https://doi.org/10.5194/gmd-12-4955-2019, https://gmd.copernicus.org/articles/12/4955/2019/, publisher: Copernicus GmbH, 2019.

Ramonet, M., Ciais, P., Apadula, F., Bartyzel, J., Bastos, A., Bergamaschi, P., Blanc, P. E., Brunner, D., Caracciolo di Torchiarolo, L., Calzolari, F., Chen, H., Chmura, L., Colomb, A., Conil, S., Cristofanelli, P., Cuevas, E., Curcoll, R., Delmotte, M., di Sarra, A., Emmenegger, L., Forster, G., Frumau, A., Gerbig, C., Gheusi, F., Hammer, S., Haszpra, L., Hatakka, J., Hazan, L., Heliasz, M., Henne, S., Hensen, A., Hermansen, O., Keronen, P., Kivi, R., Komínková, K., Kubistin, D., Laurent, O., Laurila, T., Lavric, J. V., Lehner, I., Lehtinen, K. E. J., Leskinen, A., Leuenberger, M., Levin, I., Lindauer, M., Lopez, M., Myhre, C. L., Mammarella, I., Manca, G., Manning, A., Marek, M. V., Marklund, P., Martin, D., Meinhardt, F., Mihalopoulos, N., Mölder, M., Morgui, J. A., Necki, J., O'Doherty, S., O'Dowd, C.,

https://doi.org/10.5194/gmd-2020-407
Preprint. Discussion started: 18 December 2020

Ottosson, M., Philippon, C., Piacentino, S., Pichon, J. M., Plass-Duelmer, C., Resovsky, A., Rivier, L., Rodó, X., Sha, M. K., Scheeren, H. A., Sferlazzo, D., Spain, T. G., Stanley, K. M., Steinbacher, M., Trisolino, P., Vermeulen, A., Vítková, G., Weyrauch, D., Xueref-Remy, I., Yala, K., and Yver Kwok, C.: The fingerprint of the summer 2018 drought in Europe on ground-based atmospheric CO2 measurements, Philosophical Transactions of the Royal Society B: Biological Sciences, 375, 20190 513, https://doi.org/10.1098/rstb.2019.0513, https://royalsocietypublishing.org/doi/abs/10.1098/rstb.2019.0513, publisher: Royal Society, 2020.

Rayner, P. J., Raupach, M. R., Paget, M., Peylin, P., and Koffi, E.: A new global gridded data set of CO2 emissions from fossil fuel combustion: Methodology and evaluation, Journal of Geophysical Research: Atmospheres, 115, https://doi.org/10.1029/2009JD013439, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2009JD013439, _eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2009JD013439, 2010.

Rayner, P. J., Michalak, A. M., and Chevallier, F.: Fundamentals of data assimilation applied to biogeochemistry, Atmospheric Chemistry and Physics, 19, 13 911–13 932, https://doi.org/https://doi.org/10.5194/acp-19-13911-2019, https://acp.copernicus.org/articles/19/13911/2019/, publisher: Copernicus GmbH, 2019.

Rödenbeck, C., Gerbig, C., Trusilova, K., and Heimann, M.: A two-step scheme for high-resolution regional atmospheric trace gas inversions based on independent models, Atmos. Chem. Phys, 9, 5331–5342, http://atmos-chem-phys.net/9/5331/2009/acp-9-5331-2009.pdf, 2009.

Saunois, M., Stavert, A. R., Poulter, B., Bousquet, P., Canadell, J. G., Jackson, R. B., Raymond, P. A., Dlugokencky, E. J., Houweling, S., Patra, P. K., Ciais, P., Arora, V. K., Bastviken, D., Bergamaschi, P., Blake, D. R., Brailsford, G., Bruhwiler, L., Carlson, K. M., Carrol, M., Castaldi, S., Chandra, N., Crevoisier, C., Crill, P. M., Covey, K., Curry, C. L., Etiope, G., Frankenberg, C., Gedney, N., Hegglin, M. I., Höglund-Isaksson, L., Hugelius, G., Ishizawa, M., Ito, A., Janssens-Maenhout, G., Jensen, K. M., Joos, F., Kleinen, T., Krummel, P. B., Langenfelds, R. L., Laruelle, G. G., Liu, L., Machida, T., Maksyutov, S., McDonald, K. C., McNorton, J., Miller, P. A., Melton, J. R., Morino, I., Müller, J., Murguia-Flores, F., Naik, V., Niwa, Y., Noce, S., O'Doherty, S., Parker, R. J., Peng, C., Peng, S., Peters, G. P., Prigent, C., Prinn, R., Ramonet, M., Regnier, P., Riley, W. J., Rosentreter, J. A., Segers, A., Simpson, I. J., Shi, H., Smith, S. J., Steele, L. P., Thornton, B. F., Tian, H., Tohjima, Y., Tubiello, F. N., Tsuruta, A., Viovy, N., Voulgarakis, A., Weber, T. S., van Weele, M., van der Werf, G. R., Weiss, R. F., Worthy, D., Wunch, D., Yin, Y., Yoshida, Y., Zhang, W., Zhang, Z., Zhao, Y., Zheng, B., Zhu, Q., Zhu, Q., and Zhuang, Q.: The Global Methane Budget 2000–2017, Earth System Science Data, 12, 1561–1623, https://doi.org/https://doi.org/10.5194/essd-12-1561-2020, https://essd.copernicus.org/articles/12/1561/2020/, publisher: Copernicus GmbH, 2020.

Say, D., Manning, A. J., O'Doherty, S., Rigby, M., Young, D., and Grant, A.: Re-Evaluation of the UK's HFC-134a Emissions Inventory Based on Atmospheric Observations, Environ. Sci. Technol., 50, 11 129–11 136, https://doi.org/10.1021/acs.est.6b03630, https://doi.org/10.1021/acs.est.6b03630, publisher: American Chemical Society, 2016.

Schuh, A. E., Jacobson, A. R., Basu, S., Weir, B., Baker, D., Bowman, K., Chevallier, F., Crowell, S., Davis, K. J., Deng, F., Denning, S., Feng, L., Jones, D., Liu, J., and Palmer, P. I.: Quantifying the Impact of Atmospheric Transport Uncertainty on CO2 Surface Flux Estimates, Global Biogeochemical Cycles, 33, 484–500, https://doi.org/https://doi.org/10.1029/2018GB006086, http://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2018GB006086, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1029/2018GB006086, 2019.

https://doi.org/10.5194/gmd-2020-407
Preprint. Discussion started: 18 December 2020

Seibert, P. and Frank, A.: Source-receptor matrix calculation with a Lagrangian particle dispersion model in backward mode, Atmospheric Chemistry and Physics, 4, 51–63, https://doi.org/10.5194/acp-4-51-2004, http://www.atmos-chem-phys.org/4/51/2004/acp-4-51-2004.html, 2004.

Stohl, A., Seibert, P., Arduini, J., Eckhardt, S., Fraser, P., Greally, B. R., Lunder, C., Maione, M., Mühle, J., O'Doherty, S., Prinn, R. G., Reimann, S., Saito, T., Schmidbauer, N., Simmonds, P. G., Vollmer, M. K., Weiss, R. F., and Yokouchi, Y.: An analytical inversion method for determining regional and global emissions of greenhouse gases: Sensitivity studies and application to halocarbons, Atmos. Chem. Phys., 9, 1597–1620, https://doi.org/10.5194/acp-9-1597-2009, http://www.atmos-chem-phys.net/9/1597/2009/, 2009.

Tarantola, A. and Valette, B.: Generalized nonlinear inverse problems solved using the least squares criterion, Reviews of Geophysics, 20, 219–232, https://doi.org/https://doi.org/10.1029/RG020i002p00219, http://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/RG020i002p00219, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1029/RG020i002p00219, 1982.

Thompson, R. L. and Stohl, A.: FLEXINVERT: an atmospheric Bayesian inversion framework for determining surface fluxes of trace species using an optimized grid, Geosci. Model Dev. Discuss., 7, 3751–3801, https://doi.org/10.5194/gmdd-7-3751-2014, http://www.geosci-model-dev-discuss.net/7/3751/2014/, 2014.

Trusilova, K., Rödenbeck, C., Gerbig, C., and Heimann, M.: Technical Note: A new coupled system for global-to-regional downscaling of CO2 concentration estimation, Atmos. Chem. Phys., 10, 3205–3213, https://doi.org/10.5194/acp-10-3205-2010, https://www.atmos-chem-phys.net/10/3205/2010/, 2010.

Turner, A. J. and Jacob, D. J.: Balancing aggregation and smoothing errors in inverse models, Atmos. Chem. Phys., 15, 7039–7048, https://doi.org/10.5194/acp-15-7039-2015, https://www.atmos-chem-phys.net/15/7039/2015/, 2015.

van der Laan-Luijkx, I. T., van der Velde, I. R., van der Veen, E., Tsuruta, A., Stanislawska, K., Babenhauserheide, A., Zhang, H. F., Liu, Y., He, W., Chen, H., Masarie, K. A., Krol, M. C., and Peters, W.: The CarbonTracker Data Assimilation Shell (CTDAS) v1.0: implementation and global carbon balance 2001–2015, Geosci. Model Dev., 10, 2785–2800, https://doi.org/10.5194/gmd-10-2785-2017, https://www.geosci-model-dev.net/10/2785/2017/, 2017.

van der Velde, I. R., Miller, J. B., van der Molen, M. K., Tans, P. P., Vaughn, B. H., White, J. W. C., Schaefer, K., and Peters, W.: The CarbonTracker Data Assimilation System for CO$_2$ and δ^{13}C (CTDAS-C13 v1.0): retrieving information on land–atmosphere exchange processes, Geoscientific Model Development, 11, 283–304, https://doi.org/10.5194/gmd-11-283-2018, https://www.geosci-model-dev.net/11/283/2018/, 2018.

Wang, F., Maksyutov, S., Tsuruta, A., Janardanan, R., Ito, A., Sasakawa, M., Machida, T., Morino, I., Yoshida, Y., Kaiser, J. W., Janssens-Maenhout, G., Dlugokencky, E. J., Mammarella, I., Lavric, J. V., and Matsunaga, T.: Methane Emission Estimates by the Global High-Resolution Inverse Model Using National Inventories, Remote Sensing, 11, 2489, https://doi.org/10.3390/rs11212489, https://www.mdpi.com/2072-4292/11/21/2489, 2019.

Wang, J. S., Kawa, S. R., Collatz, G. J., Sasakawa, M., Gatti, L. V., Machida, T., Liu, Y., and Manyin, M. E.: A global synthesis inversion analysis of recent variability in CO_2 fluxes using GOSAT and in situ observations, Atmospheric Chemistry and Physics, 18, 11 097–11 124, https://doi.org/https://doi.org/10.5194/acp-18-11097-2018, https://acp.copernicus.org/articles/18/11097/2018/, publisher: Copernicus GmbH, 2018.

Wang, Y., Broquet, G., Bréon, F.-M., Lespinas, F., Buchwitz, M., Reuter, M., Meijer, Y., Loescher, A., Janssens-Maenhout, G., Zheng, B., and Ciais, P.: PMIF v1.0: an inversion system to estimate the potential of

https://doi.org/10.5194/gmd-2020-407
Preprint. Discussion started: 18 December 2020

satellite observations to monitor fossil fuel CO_2 emissions over the globe, Geoscientific Model Development Discussions, pp. 1–27, https://doi.org/https://doi.org/10.5194/gmd-2019-326, https://gmd.copernicus.org/preprints/gmd-2019-326/, publisher: Copernicus GmbH, 2020.

Yadav, V. and Michalak, A. M.: Improving computational efficiency in large linear inverse problems: an example from carbon dioxide flux estimation, Geoscientific Model Development, 6, 583–590, https://doi.org/https://doi.org/10.5194/gmd-6-583-2013, https://gmd.copernicus.org/articles/6/583/2013/gmd-6-583-2013.html, publisher: Copernicus GmbH, 2013.

Zammit-Mangion, A., Cressie, N., Ganesan, A. L., O'Doherty, S., and Manning, A. J.: Spatio-temporal bivariate statistical models for atmospheric trace-gas inversion, Chemometrics and Intelligent Laboratory Systems, 149, Part B, 227–241, https://doi.org/10.1016/j.chemolab.2015.09.006, http://www.sciencedirect.com/science/article/pii/S0169743915002221, 2015.

Zheng, T., French, N. H. F., and Baxter, M.: Development of the WRF-CO2 4D-Var assimilation system v1.0, Geoscientific Model Development, 11, 1725–1752, https://doi.org/https://doi.org/10.5194/gmd-11-1725-2018, https://gmd.copernicus.org/articles/11/1725/2018/, publisher: Copernicus GmbH, 2018.

Zupanski, D., Denning, A. S., Uliasz, M., Zupanski, M., Schuh, A. E., Rayner, P. J., Peters, W., and Corbin, K. D.: Carbon flux bias estimation employing Maximum Likelihood Ensemble Filter (MLEF), Journal of Geophysical Research: Atmospheres, 112, https://doi.org/10.1029/2006JD008371, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2006JD008371, _eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2006JD008371, 2007.

Zupanski, M.: Maximum Likelihood Ensemble Filter: Theoretical Aspects, Mon. Wea. Rev., 133, 1710–1726, https://doi.org/10.1175/MWR2946.1, https://journals.ametsoc.org/mwr/article/133/6/1710/67374/Maximum-Likelihood-Ensemble-Filter-Theoretical, publisher: American Meteorological Society, 2005.

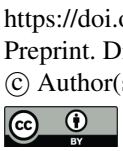
https://doi.org/10.5194/gmd-2020-407
Preprint. Discussion started: 18 December 2020

Geoscientific Model Development Discussions
Open Access
EGU

Table 1. Elementary operations required for each data assimilation method. An = Analytical inversion; EnKF = Ensemble Kalman filter; Var = Variational; Fwd = forward simulation; AdTest = Test of the adjoint. We note $\mathcal{X}$ and $\mathcal{Y}$ the target and observation spaces respectively, $\mathfrak{A}$ the regularization space in the minimization algorithm of variational inversions; the $(\cdot)^*$ symbol depicts the adjoint of corresponding spaces.

		Data assimilation method				
Elementary operation	Mathematical formulation	An	EnKF	Var	Fwd	AdTest
Forward observation operator	$\mathcal{X} \rightarrow \mathcal{Y}$ $\mathbf{x} \rightarrow \mathcal{H}(\mathbf{x})$ or $\mathbf{H}\mathbf{x}$	X	X	X	X	X
Adjoint observation operator	$\mathcal{Y}^* \rightarrow \mathcal{X}^*$ $\delta\mathbf{y} \rightarrow \mathcal{H}^*(\delta\mathbf{y})$ or $\mathbf{H}^T\delta y$	X		X		X
Normalisation of the observation increments	$\mathcal{Y}^* \rightarrow \mathcal{Y}^*$ $\delta\mathbf{y} \rightarrow \mathbf{R}^{-1}\delta\mathbf{y}$			X		
Regularization of the control space	$\mathfrak{A} \rightarrow \mathcal{X}$ $\chi \rightarrow \mathbf{x} = \mathbf{B}^{1/2}\chi + \mathbf{x}^{\mathrm{b}}$			X		
Adjoint of the control space regularization	$\mathcal{X}^* \rightarrow \mathfrak{A}^*$ $\delta\mathbf{x} \rightarrow \delta\chi \equiv \mathbf{B}^{1/2}\delta\mathbf{x}$			X		
Control space sampling	$\mathcal{X} \times \mathcal{X}^2 \rightarrow \mathcal{Y}$ $(\mathbf{x}, \mathbf{P}) \rightarrow (\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N)$		X			

https://doi.org/10.5194/gmd-2020-407
Preprint. Discussion started: 18 December 2020



Geoscientific Model Development
Discussions
Open Access
EGU

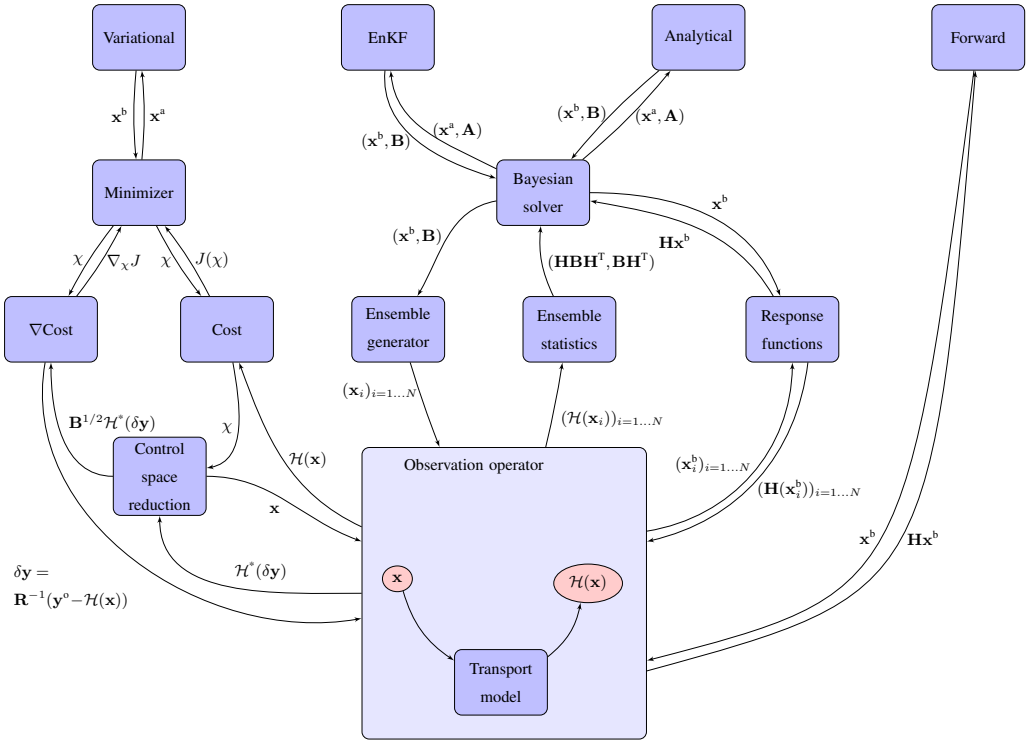


Figure 1. Call structure of the CIF.

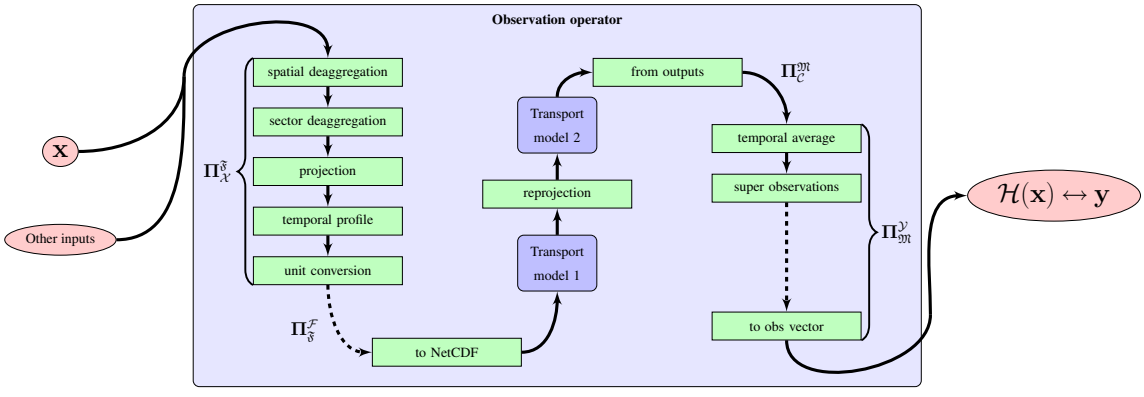


Figure 2. Observation operator structure.

https://doi.org/10.5194/gmd-2020-407
Preprint. Discussion started: 18 December 2020

Geoscientific Model Development Discussions
Open Access
EGU

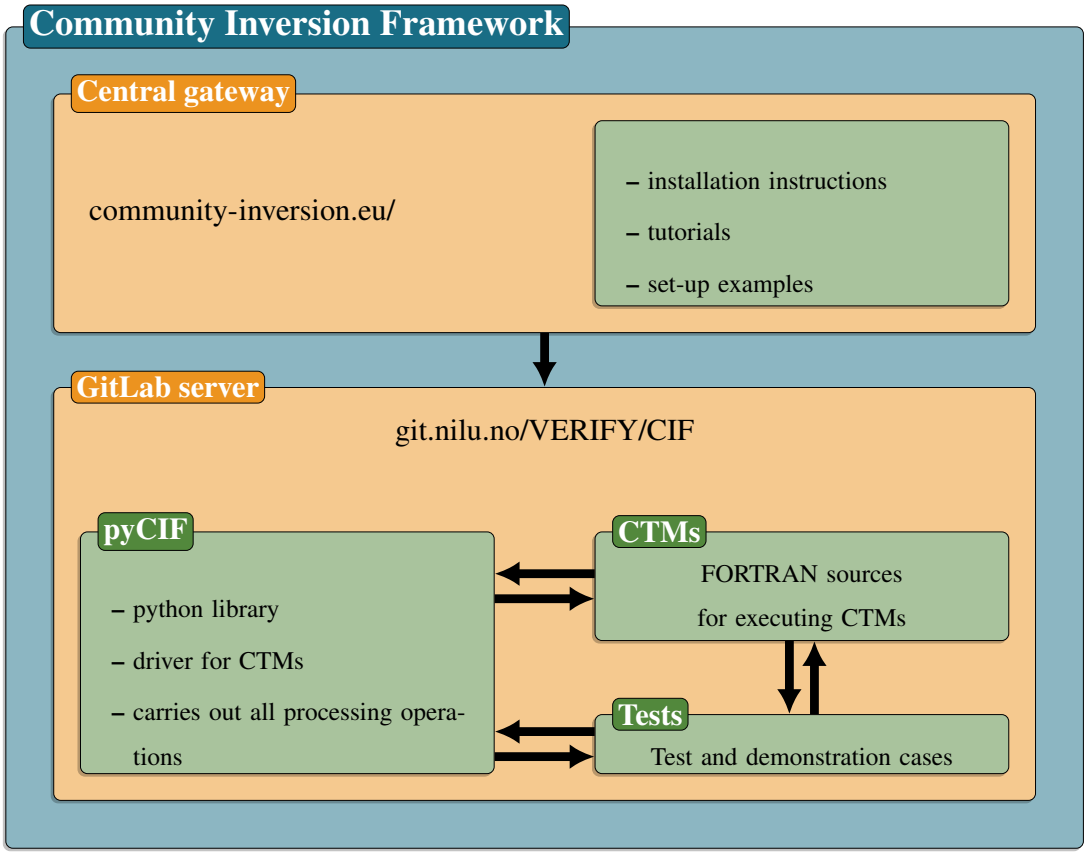


Figure 3. Organisation of the Community Inversion Framework

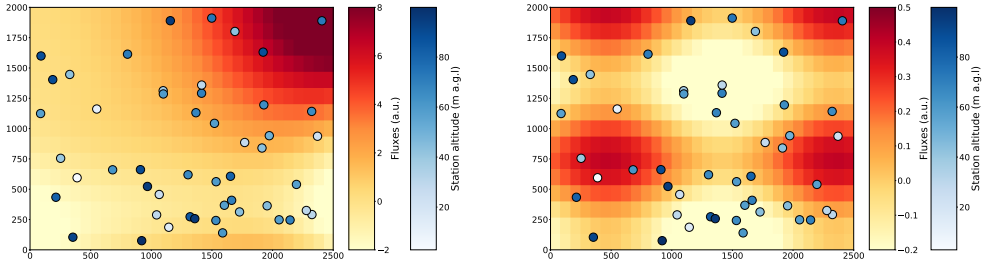


Figure 4. (left) True fluxes and observation sites. (right) Difference between prior fluxes and the truth.

https://doi.org/10.5194/gmd-2020-407
Preprint. Discussion started: 18 December 2020

Geoscientific Model Development Discussions

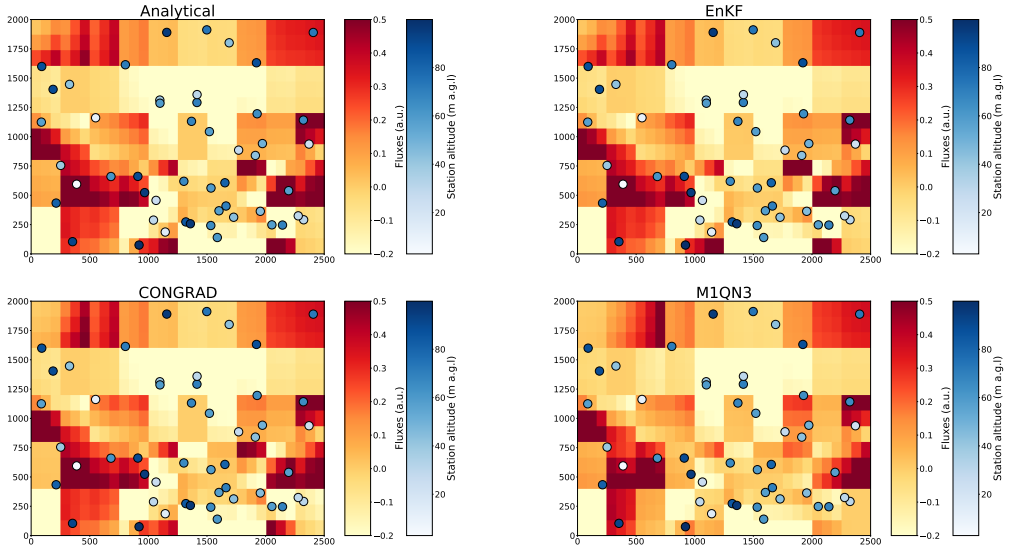


Figure 5. Posterior increments for analytical, EnKF, variational with CONGRAD and variational with M1QN3 (from top to bottom, left to right) for an inversion set-up with aggregated bands.

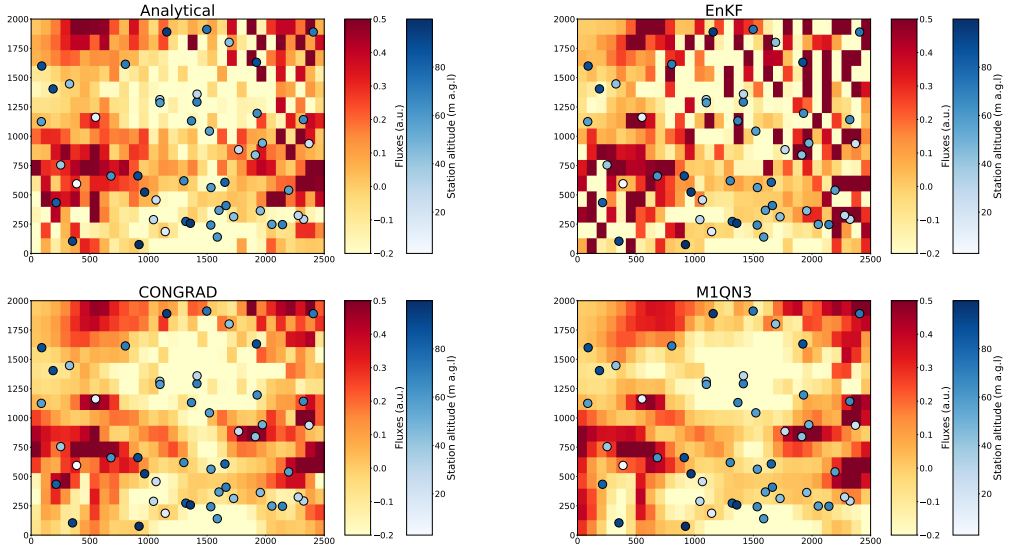


Figure 6. Same as Fig. 5 with an inversion set-up at the pixel resolution with horizontal correlation length of 500 m.

https://doi.org/10.5194/gmd-2020-407
Preprint. Discussion started: 18 December 2020

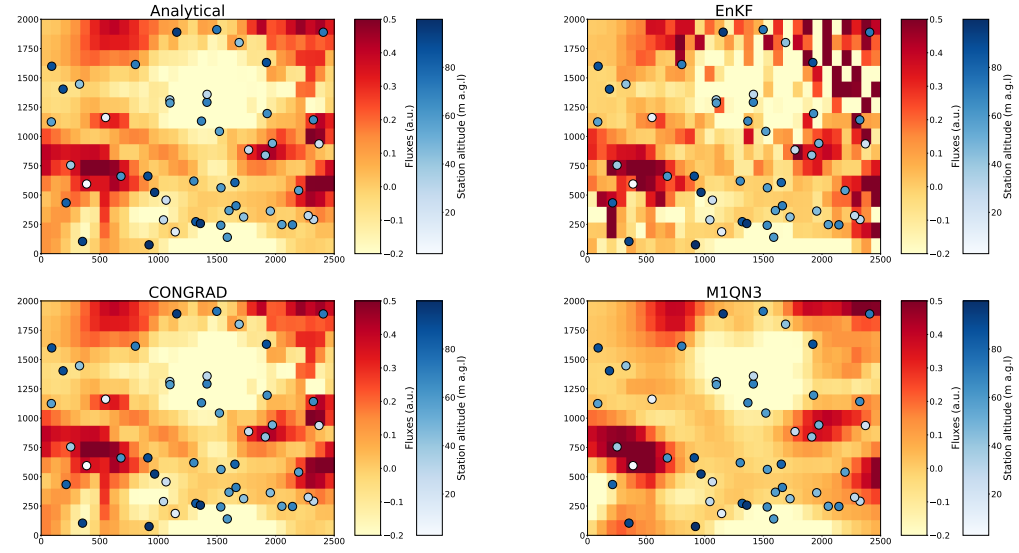


Figure 7. Same as Fig. 5 with an inversion set-up at the pixel resolution with horizontal correlation length of 10000 m.

https://doi.org/10.5194/gmd-2020-407
Preprint. Discussion started: 18 December 2020

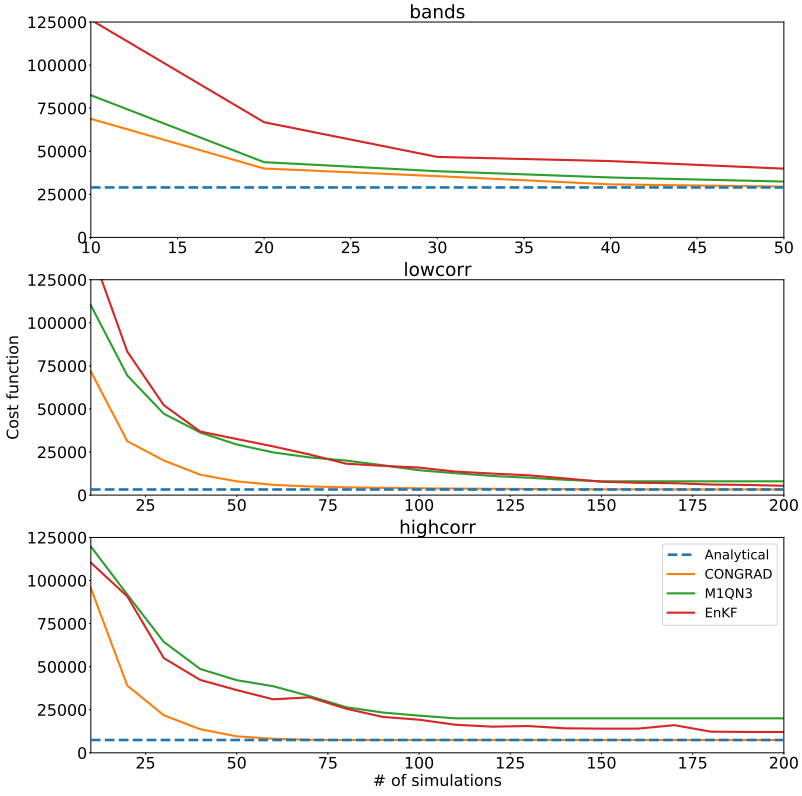


Figure 8. Sensitivity of the inversions to the number of computed simulations. (top) inversion set-up with aggregated regions of 3 pixels × 3 pixels ; (middle) inversion set-up at the pixel resolution with horizontal correlation length of 500 m; (bottom) inversion set-up at the pixel resolution with horizontal correlation length of 10000 m. For each sub-panel: (top) Distance between the true and posterior fluxes for a given number of simulations; (bottom) Evolution of the cost function depending on the number of simulations.