



**HAL**  
open science

## Variational inverse modeling within the Community Inversion Framework v1.1 to assimilate $\delta^{13}\text{C}(\text{CH}_4)$ and $\text{CH}_4$ : a case study with model LMDz-SACS

Joël Thanwerdas, Marielle Saunois, Antoine Berchet, Isabelle Pison, Bruce H. Vaughn, Sylvia Englund Michel, Philippe Bousquet

### ► To cite this version:

Joël Thanwerdas, Marielle Saunois, Antoine Berchet, Isabelle Pison, Bruce H. Vaughn, et al.. Variational inverse modeling within the Community Inversion Framework v1.1 to assimilate  $\delta^{13}\text{C}(\text{CH}_4)$  and  $\text{CH}_4$ : a case study with model LMDz-SACS. Geoscientific Model Development, 2022, 15, pp.4831-4851. 10.5194/gmd-15-4831-2022 . insu-03721924

**HAL Id: insu-03721924**

**<https://insu.hal.science/insu-03721924>**

Submitted on 13 Jul 2022

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

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



Distributed under a Creative Commons Attribution 4.0 International License



# Variational inverse modeling within the Community Inversion Framework v1.1 to assimilate $\delta^{13}\text{C}(\text{CH}_4)$ and $\text{CH}_4$ : a case study with model LMDz-SACS

Joël Thanwerdas<sup>1</sup>, Marielle Saunois<sup>1</sup>, Antoine Berchet<sup>1</sup>, Isabelle Pison<sup>1</sup>, Bruce H. Vaughn<sup>2</sup>, Sylvia Englund Michel<sup>2</sup>, and Philippe Bousquet<sup>1</sup>

<sup>1</sup>Laboratoire des Sciences du Climat et de l'Environnement, CEA-CNRS-UVSQ, IPSL, Gif-sur-Yvette, France

<sup>2</sup>INSTAAR, University of Colorado, Boulder, Boulder, CO, USA

**Correspondence:** Joël Thanwerdas (joel.thanwerdas@lsce.ipsl.fr)

Received: 1 April 2021 – Discussion started: 28 May 2021

Revised: 21 March 2022 – Accepted: 26 May 2022 – Published: 27 June 2022

**Abstract.** Atmospheric  $\text{CH}_4$  mole fractions resumed their increase in 2007 after a plateau during the 1999–2006 period, indicating relative changes in the sources and sinks. Estimating sources by exploiting observations within an inverse modeling framework (top-down approaches) is a powerful approach. It is, nevertheless, challenging to efficiently differentiate co-located emission categories and sinks by using  $\text{CH}_4$  observations alone. As a result, top-down approaches are limited when it comes to fully understanding  $\text{CH}_4$  burden changes and attributing these changes to specific source variations.  $\delta^{13}\text{C}(\text{CH}_4)_{\text{source}}$  isotopic signatures of  $\text{CH}_4$  sources differ between emission categories (biogenic, thermogenic, and pyrogenic) and can therefore be used to address this limitation. Here, a new 3-D variational inverse modeling framework designed to assimilate  $\delta^{13}\text{C}(\text{CH}_4)$  observations together with  $\text{CH}_4$  observations is presented. This system is capable of optimizing both the emissions and the associated source signatures of multiple emission categories at the pixel scale. To our knowledge, this represents the first attempt to carry out variational inversion assimilating  $\delta^{13}\text{C}(\text{CH}_4)$  with a 3-D chemistry transport model (CTM) and to independently optimize isotopic source signatures of multiple emission categories. We present the technical implementation of joint  $\text{CH}_4$  and  $\delta^{13}\text{C}(\text{CH}_4)$  constraints in a variational system and analyze how sensitive the system is to the setup controlling the optimization using the LMDz-SACS 3-D CTM. We find that assimilating  $\delta^{13}\text{C}(\text{CH}_4)$  observations and allowing the system to adjust isotopic source signatures provide relatively large differences in global flux estimates for wetlands ( $-5.7 \text{ Tg CH}_4 \text{ yr}^{-1}$ ), agriculture and

waste ( $-6.4 \text{ Tg CH}_4 \text{ yr}^{-1}$ ), fossil fuels ( $+8.6 \text{ Tg CH}_4 \text{ yr}^{-1}$ ) and biofuels–biomass burning ( $+3.2 \text{ Tg CH}_4 \text{ yr}^{-1}$ ) categories compared to the results inferred without assimilating  $\delta^{13}\text{C}(\text{CH}_4)$  observations. More importantly, when assimilating both  $\text{CH}_4$  and  $\delta^{13}\text{C}(\text{CH}_4)$  observations, but assuming that the source signatures are perfectly known, these differences increase by a factor of 3–4, strengthening the importance of having as accurate signature estimates as possible. Initial conditions, uncertainties in  $\delta^{13}\text{C}(\text{CH}_4)$  observations, or the number of optimized categories have a much smaller impact (less than  $2 \text{ Tg CH}_4 \text{ yr}^{-1}$ ).

## 1 Introduction

Methane ( $\text{CH}_4$ ) is a powerful greenhouse gas and is responsible for 23% (Etminan et al., 2016) of the radiative forcing induced by the well-mixed greenhouse gases ( $\text{CO}_2$ ,  $\text{CH}_4$ ,  $\text{N}_2\text{O}$ ). Atmospheric  $\text{CH}_4$  mole fractions have increased quasi-continuously since the preindustrial era and by about  $9 \text{ ppb yr}^{-1}$  (parts per billion per year) from 1984 to 1998 (Dlugokencky, 2021). After a plateau between 1999 and 2006 that still generates attention and controversy (e.g., Fujita et al., 2020; Thompson et al., 2018; McNorton et al., 2018; Turner et al., 2017; Schaefer et al., 2016; Schwietzke et al., 2016; Rice et al., 2016), the mole fractions resumed their increase at a large rate, exceeding  $10 \text{ ppb yr}^{-1}$  in 2014 and 2015. Trends in atmospheric  $\text{CH}_4$  are caused by a small imbalance between large sources and sinks. Assess-

ing their spatiotemporal characteristics is particularly challenging, considering the variety of  $\text{CH}_4$  emissions. Yet, identifying and quantifying the processes contributing to these changes is mandatory to formulate relevant  $\text{CH}_4$  mitigation policies that would contribute to meeting the target of the 2015 UN Paris Agreement on Climate Change and to limit climate warming to 2 °C.

Thanks to continuous efforts of surface monitoring networks, the spatial coverage and the accuracy of the atmospheric  $\text{CH}_4$  measurements provided to the scientific community have increased over the last decades. Consequently, top-down estimates using inversion methods emerged and became relevant, along with bottom-up estimates, for explaining and quantifying the recent sources and sinks variations. The first inverse modeling techniques were designed in the late 1980s and early 1990s for inferring greenhouse gas sources and sinks from atmospheric  $\text{CO}_2$  measurements (Enting and Newsam, 1990; Newsam and Enting, 1988). Without a regularization of the problem, e.g., providing prior information, the inverse problem is ill-conditioned (or ill-posed). It means that there is no unique solution to the problem but also that a small error in the assimilated data (here atmospheric observations) can result in large errors in the derived solution. Several inversion methods have been designed over the years, among which there are analytical (e.g., Bousquet et al., 2006; Gurney et al., 2002), ensemble (e.g., Zupanski et al., 2007; Peters et al., 2005), and variational methods (e.g., Chevallier et al., 2005). The variational formulation uses the adjoint equations of a specific model to compute the gradient of a cost function and then minimize it, for example, using a gradient descent method. Computational times and memory costs do not scale with the number of measurements and the number of variables to control, contrary to the analytical and ensemble methods, which can hardly accommodate very large observational datasets and control vectors at the same time. Thus, the variational formulation is preferred to the others when optimizing emissions and sinks at the pixel scale using large volumes of observational data, although its main limitation is the numerical cost for accessing posterior uncertainties when there is nonlinearity in the inversion problem (Berchet et al., 2021).

Inversion systems generally assimilate measurements from ground-based stations and/or satellites to constrain the global sources and sinks of  $\text{CH}_4$ , starting from a prior knowledge of these. These systems are very effective for providing total emission estimates (e.g., Saunio et al., 2020; Bergamaschi et al., 2018, 2013; Saunio et al., 2017; Houweling et al., 2017, and references therein). However, differentiating the contributions of multiple co-located  $\text{CH}_4$  source categories is challenging as it only relies on different seasonality cycles and on applied spatial distributions and error correlations (e.g., Bergamaschi et al., 2013, 2010). The atmospheric isotopic signal contains additional information on  $\text{CH}_4$  emissions that can help to separate emission categories based on their source origin. The atmospheric isotopic signal

$\delta^{13}\text{C}(\text{CH}_4)$  is defined as follows:

$$\delta^{13}\text{C}(\text{CH}_4) = \frac{R}{R_{\text{std}}} - 1, \quad (1)$$

where  $R$  and  $R_{\text{std}}$  denote the sample and standard  $^{13}\text{C}:\text{CH}_4$  ratios. We use the Vienna Pee Dee Belemnite (V-PDB) scale with  $R_{\text{std}} = 0.00112372$  (Craig, 1957) throughout this paper. The isotopic source signatures of  $\text{CH}_4$ , here denoted by  $\delta^{13}\text{C}(\text{CH}_4)_{\text{source}}$ , notably differ between emission categories ranging from  $^{13}\text{C}$ -depleted biogenic sources ( $-61.7 \pm 6.2\%$ ; 1 standard deviation) and thermogenic sources ( $-44.8 \pm 10.7\%$ ) to  $^{13}\text{C}$ -enriched thermogenic sources ( $-26.2 \pm 4.8\%$ ) (Sherwood et al., 2017; Schwietzke et al., 2016), although the distributions are very large and overlaps exist between the extreme values. Consequently,  $\delta^{13}\text{C}(\text{CH}_4)$  depends on both  $\text{CH}_4$  emissions and their isotopic signatures. Saunio et al. (2017) pointed out that many emission scenarios inferred from atmospheric inversions are not consistent with  $\delta^{13}\text{C}(\text{CH}_4)$  observations and that this constraint must be integrated into the inversion systems to avoid such inconsistencies. In addition, they highlighted the sensitivity of the atmospheric isotopic signal to the source partitioning and prescribed isotopic ratios. Since the 1990s,  $\delta^{13}\text{C}(\text{CH}_4)$  has been monitored at multiple sites, providing opportunities to use this constraint within an inversion framework. In addition, these values have been shifting towards more negative values since 2006 (Nisbet et al., 2019) when  $\text{CH}_4$  trends resumed their increase, suggesting that this isotopic data can help to understand the processes that contributed to the renewed growth. However, implementing the assimilation of such measurements into an inversion system is not straightforward and introduces additional complexity.

Hereinafter, the assimilation of  $\delta^{13}\text{C}(\text{CH}_4)$  observations to constrain the estimates of an inversion is referred to as the isotopic constraint. The implementation of such a constraint in an inversion system has already been attempted in previous studies focusing on  $\text{CH}_4$  (e.g., Thompson et al., 2018; McNorton et al., 2018; Rigby et al., 2017; Rice et al., 2016; Schaefer et al., 2016; Schwietzke et al., 2016; Rigby et al., 2012; Neef et al., 2010; Bousquet et al., 2006; Fletcher et al., 2004) but, to our knowledge, never in a variational system associated to a 3-D chemistry transport model (CTM). Adding this isotopic constraint to a variational inversion system is challenging as, in contrast to an analytic inversion in which the response functions of the model are precomputed, the isotopic constraints have to be considered both in the forward (simulated isotopic values) and the adjoint (sensitivity of isotopic observations to optimized variables) versions of the model.

This new system was implemented in the Community Inversion Framework (CIF), supported by the European Union H2020 project VERIFY (<http://community-inversion.eu>, last access: 3 June 2022) and required to implement new forward, tangent linear and adjoint operations. The forward operations were previously used to estimate the impact of the

Cl sink on the modeling of  $\text{CH}_4$  and  $\delta^{13}\text{C}(\text{CH}_4)$  in LMDz-SACS (Thanwerdas et al., 2019). The purpose of this study is to present the technical implementation of the isotopic constraint in a variational inversion system and to investigate the sensitivity of this new configuration to different parameters. Our aim is not to estimate trends in sectoral emissions over the last 2 decades: future studies will address the estimation of  $\text{CH}_4$  emissions over longer periods of time using this new system. The technical implementation and the various tested configurations are presented in Sect. 2. We analyze the results in Sect. 3. Section 4 presents our conclusions and recommendations on using such a multi-constraint variational system.

## 2 Methods

### 2.1 Theory of variational inversion

The notations introduced here follow the convention defined by Ide et al. (1997) and Rayner et al. (2019). The observation vector is called  $\mathbf{y}^o$ . It includes all available observations here, namely  $\text{CH}_4$  and  $\delta^{13}\text{C}(\text{CH}_4)$  measurements retrieved by surface stations, over the full simulation time window (see Sect. 2.4.2). The associated errors are assumed to be unbiased and Gaussian and are described within the error covariance matrix  $\mathbf{R}$ . This matrix accounts for all errors contributing to mismatches between simulated and observed values.  $\mathbf{x}$  is the control vector and includes all the variables (here  $\text{CH}_4$  surface fluxes, initial  $\text{CH}_4$  mole fractions, source signatures  $\delta^{13}\text{C}(\text{CH}_4)_{\text{source}}$ , and initial  $\delta^{13}\text{C}(\text{CH}_4)$  values) optimized by the inversion system. Hereinafter, these variables will be referred to as the control variables. Prior information about the control variables are provided by the vector  $\mathbf{x}^b$ . Its associated errors are also assumed to be unbiased and Gaussian and are described within the error covariance matrix  $\mathbf{B}$ .  $\mathcal{H}$  is the observation operator that projects the control vector  $\mathbf{x}$  into the observation space. This operator mainly consists of the 3-D CTM (here LMDz-SACS, as introduced in Sect. 2.2). Nevertheless, the CTM is followed by spatial and time operators, which interpolate the simulated fields to produce simulated equivalents of the assimilated observations at specific locations and times, making the simulations and observations comparable. An additional transformation operator, implemented in the new system, enables comparison between distinct simulated tracers, e.g.,  $^{12}\text{CH}_4$  and  $^{13}\text{CH}_4$ , and observations, e.g.,  $\delta^{13}\text{C}(\text{CH}_4)$  (see Sect. 2.3).

In a variational formulation of the inference problem that allows for  $\mathcal{H}$  nonlinearity, the cost function  $J$  is defined as follows:

$$\begin{aligned} J(\mathbf{x}) &= \frac{1}{2}(\mathbf{x} - \mathbf{x}^b)^T \mathbf{B}^{-1}(\mathbf{x} - \mathbf{x}^b) + \frac{1}{2}(\mathcal{H}(\mathbf{x}) - \mathbf{y}^o)^T \\ &\quad \times \mathbf{R}^{-1}(\mathcal{H}(\mathbf{x}) - \mathbf{y}^o) \quad (2) \\ &= J_b(\mathbf{x}) + J_o(\mathbf{x}). \quad (3) \end{aligned}$$

The cost function is therefore a sum of the following two parts:

1. The first part is induced by the differences between the posterior and prior variables ( $J_b$ ).
2. The second is induced by the differences between simulations and observations ( $J_o$ ).

The minimum of  $J$  can be reached iteratively with a descent algorithm that requires several computations of the gradient of  $J$  with respect to the control vector  $\mathbf{x}$ , as follows:

$$\nabla J_{\mathbf{x}} = \mathbf{B}^{-1}(\mathbf{x} - \mathbf{x}^b) + \mathcal{H}^*(\mathbf{R}^{-1}(\mathcal{H}(\mathbf{x}) - \mathbf{y}^o)). \quad (4)$$

$\mathcal{H}^*$  denotes the adjoint operator of  $\mathcal{H}$ . As in analytical and ensemble methods, the variational formulation necessitates the inversion of both error matrices  $\mathbf{R}$  and  $\mathbf{B}$ . In most applications,  $\mathbf{R}$  is considered diagonal as point observations are distant in time and space (i.e., uncorrelated observation errors), allowing for the inverse to be calculated easily, although that assumption should be revised with the increasing availability of satellite sources (Liu et al., 2020).  $\mathbf{B}$  is rarely diagonal due to spatial and temporal correlations of errors in the fluxes. However,  $\mathbf{B}$  is often decomposed as combinations of smaller matrices, e.g., using Kronecker products of sub-correlation matrices, which allows us to compute its inverse by blocks.

### 2.2 The chemistry transport model

The LMDz general circulation model (GCM) is the atmospheric component of the Institut Pierre-Simon Laplace Coupled Model (IPSL-CM) developed at the Laboratoire de Météorologie Dynamique (LMD) (Hourdin et al., 2006). The version of LMDz we use is an offline version dedicated to the inversion framework created by Chevallier et al. (2005). The precomputed air mass fluxes provided by the online version of LMDz are given as inputs to the transport model, significantly reducing the computational time. The model is set up at a horizontal resolution of  $3.8^\circ \times 1.9^\circ$  (96 grid cells in longitude and latitude) with 39 hybrid sigma pressure levels reaching an altitude up to about 75 km. About 20 levels are dedicated to the stratosphere and the mesosphere. The model time step is 30 min, and the output mole fractions are 3 h snapshots. The horizontal winds are nudged towards ECMWF meteorological analyses (ERA-Interim) in the online version of the model and then fed to the offline version. Vertical diffusion is parameterized by a local approach from Louis (1979), and deep convection processes are parameterized by the Tiedtke (1989) scheme.

The offline model LMDz is coupled with the Simplified Atmospheric Chemistry System (SACS; Pison et al., 2009). This chemistry system was previously used to simulate the oxidation chain of hydrocarbons, including  $\text{CH}_4$ , formaldehyde ( $\text{CH}_2\text{O}$ ), carbon monoxide (CO), and molecular hydrogen ( $\text{H}_2$ ). For the purpose of this study, this system was converted into a chemistry parsing system. It follows the same

principle as the one used by the regional model CHIMERE (Menut et al., 2013) and therefore allows for user-specific chemistry reactions. As a result, it generalizes the previous SACS module to any possible set of reactions. The adjoint code has also been implemented to allow variational inverse modeling. The different species are either prescribed (here OH, O(<sup>1</sup>D), and Cl) or simulated (here <sup>12</sup>CH<sub>4</sub> and <sup>13</sup>CH<sub>4</sub>). The prescribed species are neither transported in LMDz, nor are their mole fractions updated through chemical production or destruction. Such species are only used to calculate reaction rates to update the simulated species at each model time step. In this study, the isotopologues <sup>12</sup>CH<sub>4</sub> and <sup>13</sup>CH<sub>4</sub> are simulated as separate tracers, and CH<sub>4</sub> is defined as the sum of both isotopologues. Cl + CH<sub>4</sub> oxidation was implemented to complete the chemical removal of CH<sub>4</sub>, which previously only accounted for OH + CH<sub>4</sub> and O(<sup>1</sup>D) + CH<sub>4</sub> in the SACS scheme.

In the atmosphere, radicals (OH, O(<sup>1</sup>D), or Cl) react faster with <sup>12</sup>CH<sub>4</sub> than with <sup>13</sup>CH<sub>4</sub>. This effect is called the kinetic isotope effect (KIE) or the fractionation effect. Fractionation values are prescribed to the different sinks in SACS. Here, this value is defined by  $\text{KIE} = k_{12}/k_{13}$ , where  $k_{12}$  is the rate constant of a reaction between a radical and <sup>12</sup>CH<sub>4</sub>.  $k_{13}$  is the rate constant of the reaction between the same radical and <sup>13</sup>CH<sub>4</sub>. Additional information and prescribed KIE values are provided in Sect. S2 in the Supplement.

The chemistry transport LMDz-SACS is used to test the new variational inverse modeling system that is described in the next section.

### 2.3 Technical implementation of the isotopic constraint

The isotopic multi-constraint system was implemented in the CIF. The CIF has been designed to allow comparison of different approaches, models, and inversion systems used in the inversion community (Berchet et al., 2021). Different atmospheric transport models, regional and global and Eulerian and Lagrangian, are implemented within the CIF. The system presented in this paper was originally designed to run and be tested with LMDz-SACS but can theoretically be coupled with all models implemented in the CIF framework. The system is able to do the following:

- assimilate  $\delta^{13}\text{C}(\text{CH}_4)$  and CH<sub>4</sub> observations together,
- independently optimize fluxes and isotopic signatures for multiple emission categories, and
- optimize  $\delta^{13}\text{C}(\text{CH}_4)$  and CH<sub>4</sub> initial conditions.

Figure 1 shows the different steps of a minimization iteration of the cost function. Each iteration performed with the descent algorithm can be decomposed into the four main steps presented below. For clarity, here we only present the optimization of CH<sub>4</sub> fluxes and associated source signatures, but CH<sub>4</sub> and  $\delta^{13}\text{C}(\text{CH}_4)$  initial conditions can also be optimized by the system following the same process.

1. The process starts with a forward run. The different flux variables are extracted and converted into <sup>12</sup>CH<sub>4</sub> and <sup>13</sup>CH<sub>4</sub> mass fluxes for each category following Eqs. (5) and (6) below.

$$F_{12}^i = \frac{M_{12}}{M_T} \cdot \frac{1}{1 + A^i} \cdot F_T^i \quad (5)$$

$$F_{13}^i = \frac{M_{13}}{M_T} \cdot \frac{A^i}{1 + A^i} \cdot F_T^i, \quad (6)$$

with

$$A^i = (1 + \delta^{13}\text{C}(\text{CH}_4)_{\text{source}}^i) \cdot R_{\text{std}}. \quad (7)$$

$F_T^i$ ,  $F_{12}^i$  and  $F_{13}^i$  are the CH<sub>4</sub>, <sup>12</sup>CH<sub>4</sub>, and <sup>13</sup>CH<sub>4</sub> mass fluxes of a specific category  $i$ , respectively.  $M_T$ ,  $M_{12}$ , and  $M_{13}$  are the CH<sub>4</sub>, <sup>12</sup>CH<sub>4</sub>, and <sup>13</sup>CH<sub>4</sub> molar masses, respectively.  $\delta^{13}\text{C}(\text{CH}_4)_{\text{source}}^i$  is the isotopic signature of the category  $i$ .  $M_T$  should preferably depend on  $M_{12}$  and  $M_{13}$  when converting the mass fluxes, as follows:

$$M_T = \frac{M_{12} + A^i \cdot M_{13}}{1 + A^i}. \quad (8)$$

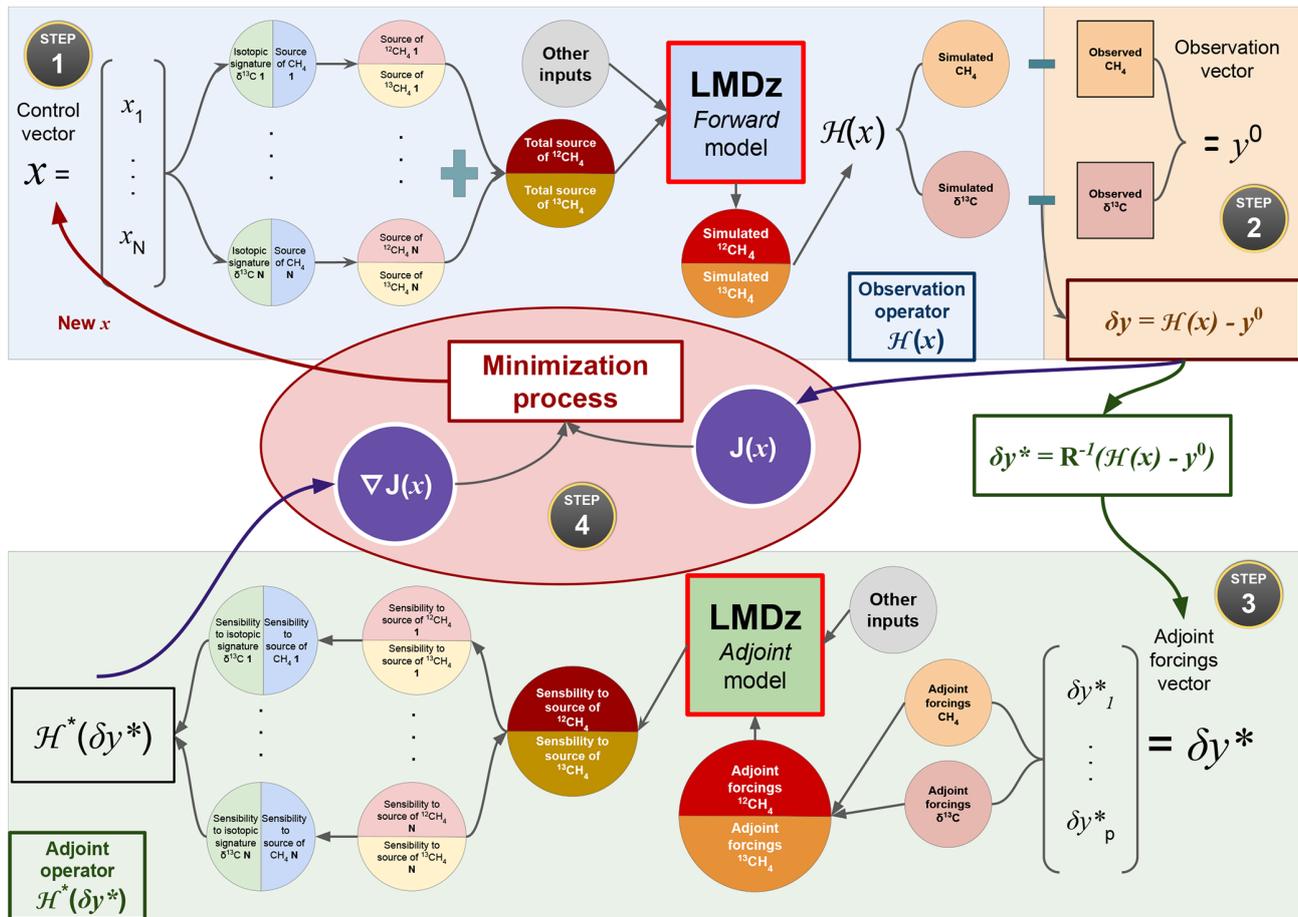
However, the complexity of the forward, tangent linear, and adjoint codes would be largely enhanced by such a relationship. The code structure would also be less generic, i.e., it could not be used for a joint assimilation of multiple isotopologues of CH<sub>4</sub>, such as both  $\delta^{13}\text{C}(\text{CH}_4)$  and  $\delta\text{D}(\text{CH}_4)$ . We choose to implement  $M_T$  as a constant that can be prescribed freely by the user, therefore without considering any influence of the  $M_{12}$  and  $M_{13}$  values, which are also prescribed by the user. As the observed isotopic source signatures roughly vary between  $-70\text{‰}$  and  $-10\text{‰}$ , a maximum variation of 0.004 % in  $M_T$  could be expected. It will very likely not affect the results of our study or that of any other inversion performed with our system.

The <sup>12</sup>CH<sub>4</sub> and <sup>13</sup>CH<sub>4</sub> total fluxes are then calculated by summing all categories and are used by the model LMDz-SACS to simulate the <sup>12</sup>CH<sub>4</sub> and <sup>13</sup>CH<sub>4</sub> atmospheric mole fractions over the time window considered. After the simulation, the simulated values are converted to the CH<sub>4</sub>- and  $\delta^{13}\text{C}(\text{CH}_4)$ -simulated equivalent of the assimilated observations using Eqs. (9) and (10) below, as follows:

$$[\text{CH}_4] = [^{12}\text{CH}_4] + [^{13}\text{CH}_4] \quad (9)$$

$$\delta^{13}\text{C}(\text{CH}_4) = \frac{[^{13}\text{CH}_4]}{[^{12}\text{CH}_4]} \cdot \frac{1}{R_{\text{std}}} - 1. \quad (10)$$

$[\text{CH}_4]$ ,  $[^{12}\text{CH}_4]$ , and  $[^{13}\text{CH}_4]$  are CH<sub>4</sub>, <sup>12</sup>CH<sub>4</sub>, and <sup>13</sup>CH<sub>4</sub> atmospheric mole fractions simulated by the model in moles per mole (hereafter mol mol<sup>-1</sup>), respectively.



**Figure 1.** The minimization iteration process in the newly designed system. The black circles with a gold border indicate the reading direction to follow. Step 1 (blue rectangle) refers to a forward run. Step 2 (orange rectangle) refers to the forward and adjoint operations required to compare observations and simulated values. Step 3 (green rectangle) refers to an adjoint run (this step must be read from the right to the left). Step 4 (red ellipse) refers to the minimization of the cost function operated by the dedicated minimization algorithm. Note that results of Step 2 are used in both the minimization process (red ellipse) and as inputs for Step 3. The minimization iteration process followed by the previous system is also illustrated in Fig. S1 in the Supplement.

2. These simulated values are then compared to the available observations in order to compute  $\mathcal{H}(x) - y^0$ , which is further used to infer the cost function and generate  $\text{CH}_4$  and  $\delta^{13}\text{C}(\text{CH}_4)$  adjoint forcings (indicated by the superscripted asterisk) that compose the vector  $\delta y^*$  as follows:

$$\delta y^* = \mathbf{R}^{-1}(\mathcal{H}(x) - y^0). \quad (11)$$

Although this vector is normally used directly as input to the adjoint model (see Eq. 4), the  $\text{CH}_4$  and  $\delta^{13}\text{C}(\text{CH}_4)$  adjoint forcings must first be converted into the  $^{12}\text{CH}_4$  and  $^{13}\text{CH}_4$  adjoint forcings in the new system.

3. The newly designed adjoint code that converts  $\text{CH}_4$  and  $\delta^{13}\text{C}(\text{CH}_4)$  adjoint forcings into  $^{12}\text{CH}_4$  and  $^{13}\text{CH}_4$  adjoint forcings is based on Eqs. (12)–(14), depending on

the type of the initial observation.

$$[^{12}\text{CH}_4]_{\text{CH}_4}^* = [^{13}\text{CH}_4]_{\text{CH}_4}^* = [\text{CH}_4]^* \quad (12)$$

$$[^{12}\text{CH}_4]_{\delta^{13}\text{C}}^* = -\frac{[^{13}\text{CH}_4]}{[^{12}\text{CH}_4]^2} \cdot \frac{1}{R_{\text{std}}} \cdot \delta^{13}\text{C}(\text{CH}_4)^* \quad (13)$$

$$[^{13}\text{CH}_4]_{\delta^{13}\text{C}}^* = \frac{1}{[^{12}\text{CH}_4]} \cdot \frac{1}{R_{\text{std}}} \cdot \delta^{13}\text{C}(\text{CH}_4)^*. \quad (14)$$

$[^{12}\text{CH}_4]_{\text{CH}_4}^*$  and  $[^{13}\text{CH}_4]_{\text{CH}_4}^*$  are adjoint forcings associated with  $\text{CH}_4$  observations.  $[^{12}\text{CH}_4]_{\delta^{13}\text{C}}^*$  and  $[^{13}\text{CH}_4]_{\delta^{13}\text{C}}^*$  are adjoint forcings associated with  $\delta^{13}\text{C}(\text{CH}_4)$  observations. The adjoint code of the CTM is then run with these adjoint forcings as inputs.

Outputs of the adjoint run provide the sensitivities of the adjoint forcings to the  $^{12}\text{CH}_4$  and  $^{13}\text{CH}_4$  mass fluxes of a specific category  $i$ , denoted by  $F_{12}^{*,i}$  and  $F_{13}^{*,i}$ . Equations (15) and (16) convert them back to sensitivities

to the initial control variables, denoted by  $F_{\text{T}}^{*,i}$  and  $\delta^{13}\text{C}(\text{CH}_4)_{\text{source}}^{*,i}$ .

$$F_{\text{T}}^{*,i} = \frac{1}{1+A} \cdot \left[ \frac{M_{12}}{M_{\text{T}}} \cdot F_{12}^{*,i} + \frac{M_{13}}{M_{\text{T}}} \cdot A \cdot F_{13}^{*,i} \right] \quad (15)$$

$$\delta^{13}\text{C}(\text{CH}_4)_{\text{source}}^{*,i} = R_{\text{std}} \cdot \frac{F_{\text{T}}}{(1+A)^2} \cdot \left[ \frac{M_{13}}{M_{\text{T}}} \cdot F_{13}^{*,i} - \frac{M_{12}}{M_{\text{T}}} \cdot F_{12}^{*,i} \right]. \quad (16)$$

4. The minimization algorithm utilizes these sensitivities to compute the gradient of the cost function. It then finds an optimized control vector which reduces the cost function and is used for the next iteration.

In order to confirm that the several adjoint operations have been correctly implemented, we also provide the results of multiple adjoint tests in Sect. S4 in the Supplement.

## 2.4 Setup of the reference simulation

The reference configuration (REF) is a variational inversion that optimizes the  $\text{CH}_4$  emission fluxes and  $\delta^{13}\text{C}(\text{CH}_4)$  isotopic source signatures of the following five different categories: biofuels–biomass burning (BB), agriculture and waste (AGW), fossil fuels (FFs), wetlands (WET), and other natural sources (NAT).  $\text{CH}_4$  and  $\delta^{13}\text{C}(\text{CH}_4)$  initial conditions are also optimized. The assimilation time-window is the 2012–2017 period. The five categories originate from an aggregation of 10 subcategories (see Table 1) and are chosen to be as isotopically consistent as possible. Sinks are not optimized here.

### 2.4.1 Control vector $\mathbf{x}$ and $\mathbf{B}$ matrix

We adopt the  $\text{CH}_4$  emissions compiled for inversions performed as part of the global methane budget (Saunio et al., 2020). Anthropogenic (including biofuels) and biomass burning emissions are based on the EDGARv4.3.2 database (Janssens-Maenhout et al., 2017) and the GFED4s databases (van der Werf et al., 2017), respectively. Statistics from British Petroleum (BP) and the Food and Agriculture Organization of the United Nations (FAO) have been used to extend the EDGARv4.3.2 database, which ended 2012, until 2017. The natural source emissions are based on averaged literature values; see Poulter et al. (2017) for wetlands, Kirschke et al. (2013) for termites, Lambert and Schmidt (1993) for oceanic sources, and Etiope (2015) for geological (onshore) sources. Emissions from geological sources have been scaled down to  $15 \text{ Tg CH}_4 \text{ yr}^{-1}$  in the prior emissions adopted in Saunio et al. (2020). All prior fluxes are prescribed at a monthly resolution and at the spatial resolution of LMDz. Globally averaged emissions over the 2012–2017 period are listed in Table 1.

Prior estimates of isotopic source signatures are provided either at the pixel scale (for wetlands), at the regional scale

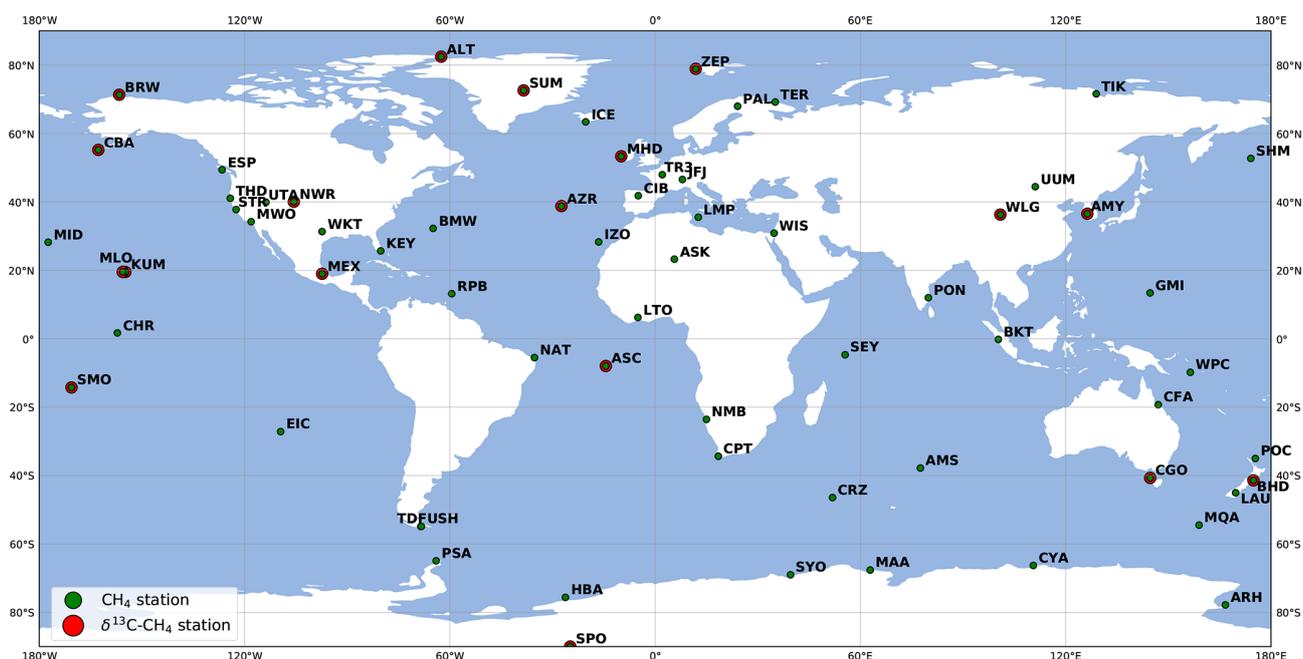
based on Atmospheric Tracer Transport Model Intercomparison Project (TransCom) regions (Patra et al., 2011) or at the global scale. The wetlands signature map is taken from Ganesan et al. (2018). Livestock isotopic source signatures are taken from Chang et al. (2019) and aggregated into the 11-region map by selecting region-specific values. Livestock source signatures have been likely decreasing over time since the 1990s due to changes in the  $\text{C}_3/\text{C}_4$  diet within the major livestock-producing countries, and therefore, annual values are prescribed. However, these estimates end in 2013, and we set the years 2014 to 2017 equal to the year 2013. Consequently, only the year 2012 has a different prescribed value from the other years. Coal and oil, gas, and industry (OGI) isotopic signature values are inferred from Sherwood et al. (2017) and Zazzeri et al. (2016) and aggregated into the same 11-region map. The EDGARv4.3.2 categories PRO\_OIL and PRO\_GAS (fugitive emissions during oil and gas exploitation) largely contribute ( $\sim 90\%$ ) to the total of the oil, gas, and industry subcategory. Therefore, we chose to neglect the influence of other sub-subcategories (such as industry) on the isotopic signature of the category. As for the biofuels–biomass burning category, we use region-specific signatures over the 11 regions. A global signature value is prescribed for each of the other categories. Except for the livestock category, all prior signatures are set constant over time. To infer the  $\delta^{13}\text{C}(\text{CH}_4)_{\text{source}}$  map of a category based on the subcategories, the  $^{12}\text{CH}_4$  and  $^{13}\text{CH}_4$  fluxes for each emission subcategory within a category are derived based on Eqs. (5) and (6) and added up. The resulting fluxes are then converted back to a  $\delta^{13}\text{C}(\text{CH}_4)_{\text{source}}$  map representing the aggregated isotopic signature of the category. Additional information regarding the chosen isotopic signatures and their references is provided in Sect. S1 in the Supplement.

In total, three values per month (10 d, 10 d, and the rest) for fluxes and their associated isotopic signatures are included in the control variables. Although the time variations in isotopic signatures are poorly constrained in the literature, we choose to include the same number of variables for fluxes and isotopic signatures in order to illustrate the full capabilities of the system and have it ready when more isotopic constraints will appear.

The portion of the diagonal of  $\mathbf{B}$  associated to prior  $\text{CH}_4$  emission fluxes is filled in with the variances set to 100 % of the square of the maximum of emissions over the cell and its eight neighbors during each month. Off-diagonal terms of  $\mathbf{B}$  (covariances) are based on correlation  $e$ -folding lengths (500 km over land and 1000 km over sea). The same method is applied for isotopic source signatures, although a specific percentage of uncertainties deduced from the global values of Sherwood et al. (2017) is used to infer each category's diagonal term (see Table 1). No temporal correlations are considered here. Finally, prior uncertainties in initial conditions are set to 10 % for  $\text{CH}_4$  ( $\sim 180$  ppb) and 3 % for  $\delta^{13}\text{C}(\text{CH}_4)$  ( $\sim 1.4\%$ ).

**Table 1.** Emissions and flux-weighted isotopic signatures of the  $\text{CH}_4$  sources averaged over 2012–2017 for different categories and their subcategories. Prior uncertainties in fluxes are set to 100 % for all categories and subcategories. Note: Unc. is the prior uncertainty in the isotopic signature prescribed to the category or the subcategory as a percentage of the signature. CH19 is Chang et al. (2019), GA18 is Ganesan et al. (2018), SH17 is Sherwood et al. (2017), WA16 is Warwick et al. (2016), ZA16 is Zazzeri et al. (2016), TO12 is Townsend-Small et al. (2012), KL10 is Klevenhusen et al. (2010), BO06 is Bousquet et al. (2006), BR01 is Bréas et al. (2001), SA01 is Sansone et al. (2001), CH00 is Chanton et al. (2000), HO00 is Holmes et al. (2000), CH99 is Chanton et al. (1999), BE98 is Bergamaschi et al. (1998), and LE93 is Levin et al. (1993).

Categories	Emissions ( $\text{Tgyr}^{-1}$ )	Signature (‰)	Unc. (%)	Subcategories	Emissions ( $\text{Tgyr}^{-1}$ )	Signature (‰)	Unc. (%)	Signature references
WET	180.3	−60.8	20	Wetlands	180.3	−60.8	20	GA18
AGW	226.4	−59.1	20	Rice cultivation	38.0	−63	20	SH17; BO06; BR01
				Livestock	117.8	−63.6	20	CH19
				Waste	70.6	−49.5	20	KL10; TO12; CH99; BE98; LE93
FFs	116.3	−43.4	25	Coal	38.4	−40.4	25	SH17; ZA16
				Oil, gas, industry	77.9	−44.9	25	SH17
BB	28.4	−22.5	40	Biofuels–biomass burning	28.4	−22.5	40	BO06; CH00
NAT	38.1	−49.9	15	Oceanic sources	14.4	−42.0	20	BR01; HO00; SA01
				Termites	8.7	−63.0	20	TH18; SH17; WA16
				Geological (onshore)	15.0	−50.0	20	BO06
Total	589.5	−54.1		Total	589.5	−54.1		



**Figure 2.** Locations of  $\text{CH}_4$  and  $\delta^{13}\text{C}(\text{CH}_4)$  surface stations. Affiliated networks are not displayed. More information can be found in Tables S3 and S4 in the Supplement.

## 2.4.2 Observation vector $\mathbf{y}$ and $\mathbf{R}$ matrix

$\text{CH}_4$  observations are taken from the data archived at the World Data Centre for Greenhouse Gases (WDCGG) of the World Meteorological Organization Global Atmospheric Watch (WMO-GAW) program. We selected 66 stations from 13 surface monitoring networks providing in situ measurements of  $\text{CH}_4$  mole fractions. The stations are displayed in Fig. 2.

$\delta^{13}\text{C}(\text{CH}_4)$  observations are taken from 18 surface stations from the Global Greenhouse Gas Reference Network (GGGRN), part of the NOAA Earth System Research Laboratory's (ESRL) Global Monitoring Laboratory (NOAA GML). Air samples were collected on an approximately weekly basis during the 2012–2017 period and analyzed by the Institute of Arctic and Alpine Research (INSTAAR) to provide  $\delta^{13}\text{C}(\text{CH}_4)$  isotope ratio measurements. The analytical uncertainty of the isotopic measurements, based on a surveillance cylinder, is 0.06‰. In this study, we focused on estimating monthly and annual flux variations rather than investigating daily or weekly variations. Prescribing error correlations in the  $\mathbf{R}$  matrix (introduced in Sect. 2.1) can be used to ensure that the inversion preferentially constrains the components we are interested in (i.e., long-term trend and seasonal cycle). In order to keep the  $\mathbf{R}$  matrix diagonal and to focus on monthly and annual variations of the signal, we chose to use  $\delta^{13}\text{C}(\text{CH}_4)$  observational data based on a curve fitting the original  $\delta^{13}\text{C}(\text{CH}_4)$  observations. The fitting curve is a function including three polynomial parameters (quadratic) and eight harmonic parameters, as in Masarie and Tans (1995). After the fitting, the pseudo-observations were sampled at the same time as the original observations. We also hypothesized that the convergence would be slightly faster if a smooth curve fitting the real observations was used instead of the real observations, which appeared to be false (see Sect. 3.1). One sensitivity inversion aims at estimating the error introduced by this simplification (simulation S2 in Table 2).

The  $\mathbf{R}$  matrix for both  $\text{CH}_4$  and  $\delta^{13}\text{C}(\text{CH}_4)$  is defined as diagonal, assuming that observation errors are not correlated, neither in space nor in time. This diagonal matrix can be decomposed into two parts, i.e., measurement and model error variances. Measurement errors account for instrumental errors, whereas model errors encompass transport and representativity errors induced by the model as follows:

$$\mathbf{R} = \mathbf{R}_{\text{measurement}} + \mathbf{R}_{\text{model}} \quad (17)$$

Here, we use the provided observation errors to fill the  $\mathbf{R}_{\text{measurement}}$  diagonal matrix. GlobalView Methane ( $\text{CH}_4$ ; GLOBALVIEW-CH4, 2009) values are used to represent model errors and prescribe variances at each station for  $\text{CH}_4$  mixing ratio measurements in order to fill the  $\mathbf{R}_{\text{model}}$  diagonal matrix. This simple approach has been used previously in atmospheric inversions (Locatelli et al., 2015, 2013; Yver et al., 2011; Bousquet et al., 2006; Rodenbeck et al., 2003).

Errors in GlobalView  $\text{CH}_4$  are computed at each site as the root mean square error (RMSE) of the measurements on a smooth curve fitting them. As GlobalView  $\text{CH}_4$  does not provide errors for GlobalView  $\text{CH}_4$  measurements, the same method has been applied here. The RMSE of the measurements on a smooth curve, fitting them over the 2012–2017 period, is prescribed as the standard deviation for each site providing  $\delta^{13}\text{C}(\text{CH}_4)$  measurements. These errors range between 3–19 ppb for  $\text{CH}_4$  observations and 0.11–0.20‰ for  $\delta^{13}\text{C}(\text{CH}_4)$  observations. Mean prescribed errors for each station are provided in Tables S3 and S4 in the Supplement.

## 2.4.3 Spin-up

Before starting the inversion, the model has been spun-up over 30 years, using constant emissions and recycling meteorology from the year 2012 in order to consider the long timescales for isotopic changes (Tans, 1997). At the end of the spin-up,  $\delta^{13}\text{C}(\text{CH}_4)$  values have been offset (+1.4‰) to fit the global mean  $\delta^{13}\text{C}(\text{CH}_4)$  in January 2012, and  $\text{CH}_4$  mole fractions have been scaled to fit the global mean  $\text{CH}_4$  mole fraction in January 2012. Due to the nonlinearity of transport and mixing, offsetting  $\delta^{13}\text{C}(\text{CH}_4)$  initial values in a forward run can generate errors. This impact is discussed later, using a configuration where  $\delta^{13}\text{C}(\text{CH}_4)$  initial conditions have not been offset (S1).

## 2.4.4 Sensitivity tests

A set of nine different configurations, including REF, has been designed to assess the impact of assimilating  $\delta^{13}\text{C}(\text{CH}_4)$  observations in addition to  $\text{CH}_4$  observations and also to evaluate the sensitivity of the inversion results to the system's setup.

Multiple parameters have been tested throughout the following configurations:

1. NOISO has no isotopic constraint. Therefore, this configuration only simulates  $\text{CH}_4$  and assimilates  $\text{CH}_4$  observations.
2. S1 uses  $\delta^{13}\text{C}(\text{CH}_4)$  initial conditions that are not offset and are therefore directly taken from the spin-up.
3. S2 assimilates the real  $\delta^{13}\text{C}(\text{CH}_4)$  observations instead of the fitting curve data.
4. In S3, the  $\delta^{13}\text{C}(\text{CH}_4)$  model uncertainties are divided by a factor 2.
5. T1 uses 10 subcategories instead of 5 aggregated categories, increasing the degrees of freedom.
6. In theory, the system is capable of optimally adjusting two source signatures if the assimilated information is sufficient. For instance, the system can choose to shift one signature downward and another upward in a given pixel in order to improve the fitting in this specific pixel.

The configuration T2 has been specifically designed to investigate whether the system would be able to retrieve a realistic distribution (similar to REF) starting from globally averaged signatures for each category.

7. In T3, the  $\delta^{13}\text{C}(\text{CH}_4)$  source signature uncertainties are set to a very low value (1 %) in order to prevent the system from optimizing them. In other words, all changes are put on  $\text{CH}_4$  emissions.
8. Finally, T4 applies both changes from T2 and T3.

Table 2 summarizes the different configurations and the associated changes. The configurations have been grouped into two sets to facilitate the analysis of results. On the one hand, S-group configurations (REF + S1–S4) have setup variations that are not expected to largely influence the results compared to REF. On the other hand, T-group configurations (T1–T4) alter parameters that are very likely to impact the results.

### 3 Results

#### 3.1 Minimization of the cost function

The minimization process is performed using the M1QN3 algorithm (Gilbert and Lemaréchal, 1989). One full simulation (forward + adjoint) with the isotopic constraint necessitates about 170 central processing unit (CPU) hours to run for 6 years, i.e., 2.4 CPU hours per month simulated. The computational burden is increased by a factor of 2 in comparison to an inversion without the isotopic constraint due to the doubling of simulated tracers ( $^{12}\text{CH}_4$  and  $^{13}\text{CH}_4$ ). One full simulation is generally enough to complete one iteration of the minimization process, but two or three simulations are sometimes required by M1QN3. Therefore, the number of simulations is slightly larger than the number of iterations. Figure 3 displays the minimization process of the cost function for all configurations.

Except for S1 and T1, the inversions were stopped when the gradient norm reduction exceeded 96 % for the third consecutive iteration. The number of iterations are compared to investigate the sensitivity of the computational cost to the setup. In total, 32 iterations (37 simulations) for NOISO, 43 iterations (47 simulations) for REF, and about 50 iterations for the others were necessary. Consequently, although assimilating  $\delta^{13}\text{C}(\text{CH}_4)$  observations requires at least 11 additional iterations, the setup has little influence on the number of iterations if the same convergence criteria is used. Also, using curve-fitted data instead of real observations do not reduce the computational burden as we first speculated.

S1 and T1 inversions were extended until their cost function reached the same reduction as REF in order to estimate the additional computational burden required to reach similar results when initial conditions are not offset (S1) and the number of categories is increased (T1). There were 10 and

21 additional iterations necessary for T1 and S1, respectively. For T1, it shows that increasing the degrees of freedom also increases the computational burden. For S1, it highlights the benefits of offsetting  $\delta^{13}\text{C}(\text{CH}_4)$  initial conditions.

As we assume no correlation of errors in  $R$ ,  $J_o$  (see Eq. 3) can be divided into  $\text{CH}_4$  and  $\delta^{13}\text{C}(\text{CH}_4)$  contributions. Figure 3 shows that all configurations lead to a fast reduction of the  $\delta^{13}\text{C}(\text{CH}_4)$  contribution. During the first 10 iterations, it decreased from 50 %–90 % (depending on the configuration) to 10 %–20 %. Conversely, the  $\text{CH}_4$  contribution increased from 10 %–50 % to 80 %–90 %. By adjusting the isotopic source signatures (all configurations besides T3–T4), the system was able to efficiently and rapidly reduce the discrepancies between simulated and observed  $\delta^{13}\text{C}(\text{CH}_4)$ . As a result, the  $\delta^{13}\text{C}(\text{CH}_4)$  RMSE decreased very rapidly during the first 10 iterations, while the  $\text{CH}_4$  RMSE decreased at a roughly constant rate. Consequently, the system is preferentially adjusting  $\delta^{13}\text{C}(\text{CH}_4)$  over  $\text{CH}_4$  values to reduce the cost function, presumably because the ratio of RMSE to the prescribed observational error for  $\delta^{13}\text{C}(\text{CH}_4)$  is, on average, about twice as large as for  $\text{CH}_4$ . In other terms, it is simpler for the system to adjust  $\delta^{13}\text{C}(\text{CH}_4)$  before attempting to modify  $\text{CH}_4$ . The ratio of the number of  $\delta^{13}\text{C}(\text{CH}_4)$  observations to the number of  $\text{CH}_4$  observations is not expected to play a significant role in the convergence process, although we did not study this influence rigorously. This ratio is only expected to affect the contribution of a component ( $\delta^{13}\text{C}(\text{CH}_4)$  or  $\text{CH}_4$ ) to the total cost function.

The decrease rate associated with  $\delta^{13}\text{C}(\text{CH}_4)$  RMSE can be increased by reducing the model uncertainties prescribed to the  $\delta^{13}\text{C}(\text{CH}_4)$  observations. S3 is an example of such an adjustment, as the model uncertainties have been divided by two. With this configuration, the system requires five fewer iterations than REF to reach a similar  $\delta^{13}\text{C}(\text{CH}_4)$  RMSE reduction but seven additional iterations to reach a similar  $\text{CH}_4$  RMSE reduction. T3 and T4 configurations constrain the isotopic signatures; thus, the reduction in the  $\delta^{13}\text{C}(\text{CH}_4)$  contribution necessitates 25 more iterations than REF to reach similar RMSE reduction. To summarize, the decrease rate associated with  $\delta^{13}\text{C}(\text{CH}_4)$  RMSE is highly dependent on the prescribed uncertainties in  $\delta^{13}\text{C}(\text{CH}_4)$  observations and the ability of the system to adjust source signatures.

#### 3.2 $\text{CH}_4$ and $\delta^{13}\text{C}(\text{CH}_4)$ fitting

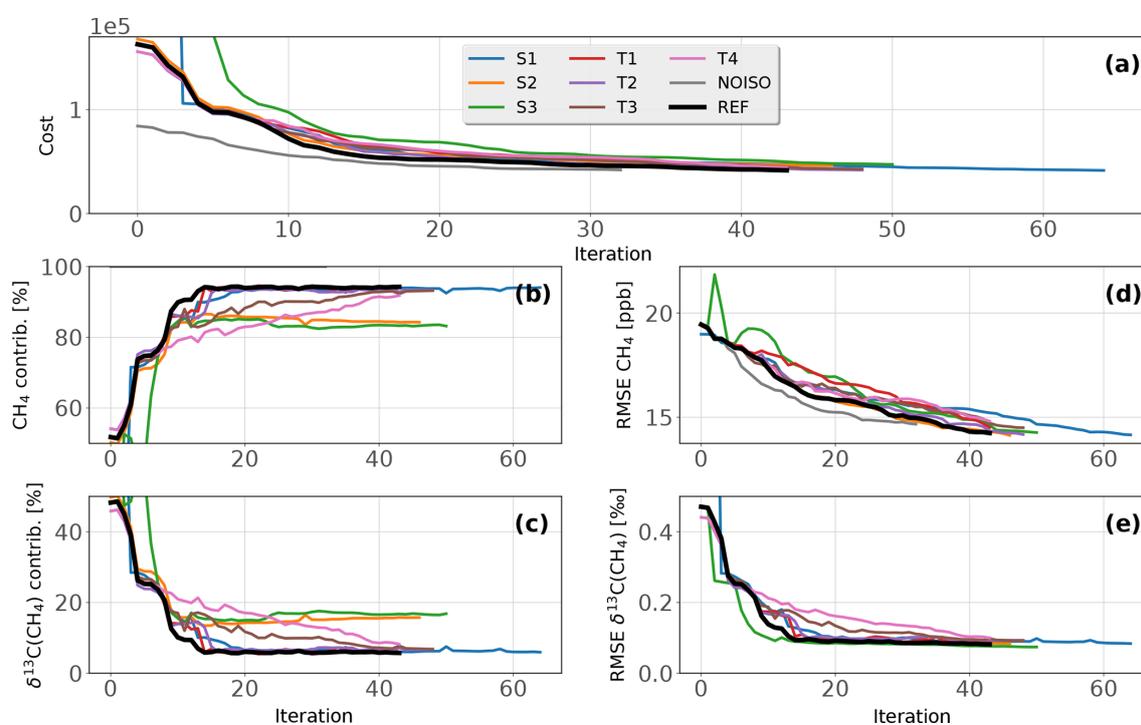
As expected, the assimilation process greatly improves the agreement between simulated and observed values for both  $\text{CH}_4$  and  $\delta^{13}\text{C}(\text{CH}_4)$ . Figure 4 shows the globally averaged time series of  $\text{CH}_4$  and  $\delta^{13}\text{C}(\text{CH}_4)$ .

$\text{CH}_4$  RMSE using prior estimates is 19.4 ppb and drops to  $14.3 \pm 0.2$  ppb ( $1\sigma$ ) on average over all the configurations using posterior estimates. Prior estimates capture the observed  $\text{CH}_4$  well, and the improvement is therefore relatively small. In addition, all configuration results regarding  $\text{CH}_4$  are very similar. In particular, NOISO is not performing much differ-

**Table 2.** Nomenclature and characteristics of the configurations. Details are provided in Sect. 2.4.4.

Name	$\delta^{13}\text{C}(\text{CH}_4)$ initial cond.	$\delta^{13}\text{C}(\text{CH}_4)$ observations	$\delta^{13}\text{C}(\text{CH}_4)$ model errors	$\delta^{13}\text{C}(\text{CH}_4)_{\text{source}}$ regional variability	$\delta^{13}\text{C}(\text{CH}_4)_{\text{source}}$ uncertainties	Number of categories
NOISO	Without isotopic constraint					5
REF	Offset	Curve fitting	RMSE obs fit	Regional variability	REF uncertainties	5
S1	No offset*	Curve fitting	RMSE obs fit	Regional variability	REF uncertainties	5
S2	Offset	Real obs	RMSE obs fit	Regional variability	REF uncertainties	5
S3	Offset	Curve fitting	RMSE obs fit/2	Regional variability	REF uncertainties	5
T1	Offset	Curve fitting	RMSE obs fit	Regional variability	REF uncertainties	10
T2	Offset	Curve fitting	RMSE obs fit	Global mean	REF uncertainties	5
T3	Offset	Curve fitting	RMSE obs fit	Regional variability	1 % for each cat.	5
T4	Offset	Curve fitting	RMSE obs fit	Global mean	1 % for each cat.	5

\* Prior uncertainties in initial  $\delta^{13}\text{C}(\text{CH}_4)$  conditions have been set to 10 %.



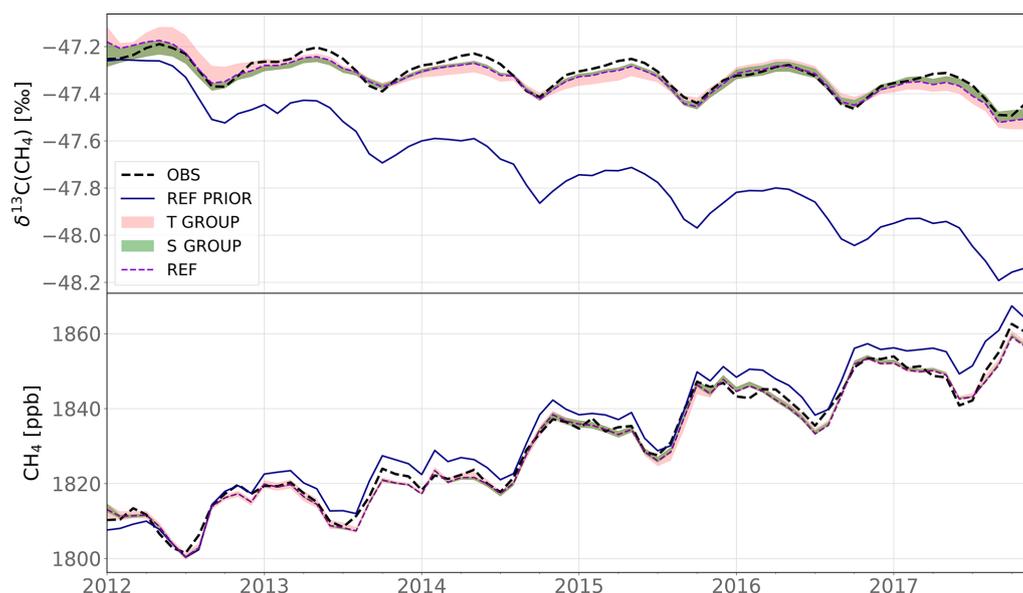
**Figure 3.** Minimization of the cost function for all configurations. (a) Value of the cost function with respect to the number of iterations. (b)  $\text{CH}_4$  contribution to  $J_0$ . (c)  $\delta^{13}\text{C}(\text{CH}_4)$  contribution to  $J_0$ . (d) RMSE associated to observed–simulated  $\text{CH}_4$ . (e) RMSE associated to observed–simulated  $\delta^{13}\text{C}(\text{CH}_4)$ . For clarity reasons, S1 and S3 initial values are not displayed because they are much larger than those of REF.

ently than the other configurations, indicating that the additional isotopic constraint does not affect the fitting to  $\text{CH}_4$  observations.

Prior  $\delta^{13}\text{C}(\text{CH}_4)$  prescribed in REF are continuously decreasing from  $-47.2\text{‰}$  to  $-48.2\text{‰}$  and thus agrees very poorly (RMSE is  $0.47\text{‰}$ ) with observed values. This can be due to an underestimation (values that are too negative) of some isotopic source signatures, an underestimation of the KIE values associated with the various sinks, an underestimation of the various sinks intensities (mostly CI and OH),

and/or a poor prior estimation of the source partitioning, i.e., an underestimation of  $^{13}\text{C}$ -enriched sources (FFs or BB) or an overestimation of  $^{13}\text{C}$ -depleted sources (WET or AGW). The data assimilation process reconciles simulated and observed  $\delta^{13}\text{C}(\text{CH}_4)$  (RMSE is  $0.086 \pm 0.008\text{‰}$ ) for all configurations, albeit that small differences, depending on the setup, emerge.

The S group provides a better match to  $\delta^{13}\text{C}(\text{CH}_4)$  observations than the T group ( $0.081 \pm 0.003\text{‰}$  versus  $0.091 \pm 0.007\text{‰}$ ). The fit is very similar within the S group.



**Figure 4.** Global monthly  $\delta^{13}\text{C}(\text{CH}_4)$  and  $\text{CH}_4$  means between 2012 and 2017. The dashed black and solid blue lines in each panel denote the observed and prior simulated values (REF), respectively. The red and green ranges show the maximum and minimum values of the T group and S group, respectively. The thick and dashed purple line denotes the posterior REF values. Globally averaged values are computed using a method similar to Masarie and Tans (1995), where a function including three polynomial parameters (quadratic) and eight harmonic parameters is fitted to each time series at available sites. The final value is obtained by performing a latitude-band weighted average over the marine boundary layer (MBL) sites. The latitude band width was set at  $30^\circ$ . The posterior NOISO lines were not included because (1) the posterior NOISO global source signature is  $-54.1\text{‰}$ , and the line would therefore reach lower values than the REF PRIOR, affecting the visual clarity of the upper plot. (2) The posterior NOISO  $\text{CH}_4$  values are extremely close to the REF values, and including it would also affect the clarity of the lower plot.

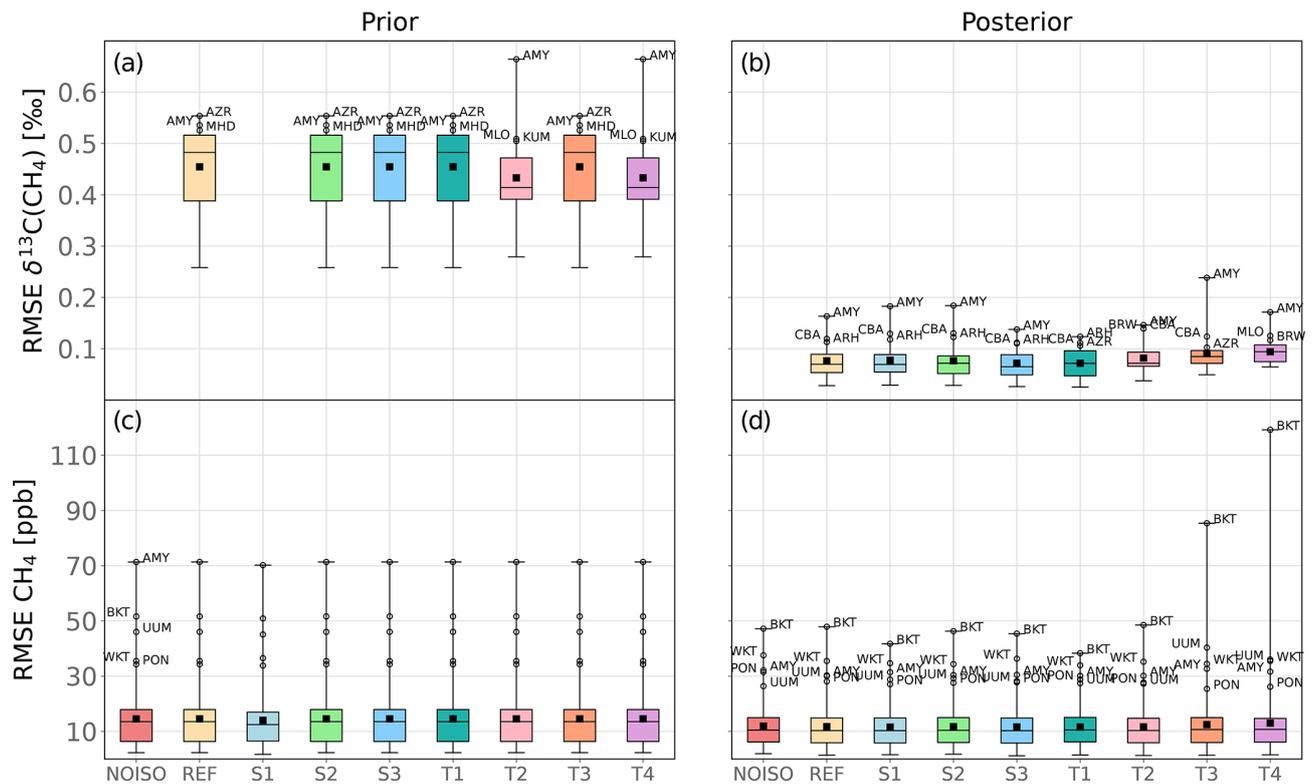
In contrast, the spread in the T group is larger, with  $\delta^{13}\text{C}(\text{CH}_4)$  RMSE being equal to  $0.093\text{‰}$ ,  $0.091\text{‰}$ , and  $0.099\text{‰}$ , respectively, for T2, T3, and T4. These results suggest that giving more freedom to the system to adjust the isotopic signatures and providing region-specific estimates of prior source signatures instead of global values may be key elements for reaching a better agreement. Best results (i.e., smallest RMSE) are obtained with T1 ( $0.079\text{‰}$ ). However, this configuration necessitates 10 additional iterations to reach better results than REF. Without these additional iterations, REF would have the best results ( $0.081\text{‰}$ ).

Figure 5 shows the RMSE distribution at all measurement sites for each configuration. All sites exhibit a RMSE reduction (from prior to posterior) for both  $\text{CH}_4$  and  $\delta^{13}\text{C}(\text{CH}_4)$ , except for Bukit Kototabang, Indonesia (BKT), with T3 and T4 configurations. Furthermore, BKT, Moody, USA (WKT), Ulaan Uul, Mongolia (UUM), Anmyeon-do, Republic of Korea (AMY), and Pondicherry, India (PON), exhibit a posterior  $\text{CH}_4$  RMSE above 25 ppb, showing that  $\text{CH}_4$  measurements retrieved at these stations are not properly reproduced by the model, despite the optimization. Prescribed observation errors are likely not the main cause because mean values for these stations are large (10–15 ppb) but not the largest among all the assimilated stations. It can also be due to transport error or misrepresentation of sources close

to the sites. Addressing this misfit is beyond the scope of this study, although the configuration influences the results. BKT and UUM fitting are notably deteriorated with T3 and T4 configurations. For example, BKT appears to be influenced by biomass burning sources in South East Asia, which are strongly dependent on the configuration (see Sect. 3.3). Moreover, T3 provides the poorest  $\delta^{13}\text{C}(\text{CH}_4)$  fitting at AMY ( $0.24\text{‰}$ ). Therefore, using global values for source signatures and preventing the system from optimizing them lead to poorer fitting. On the contrary, T1 improves the results, indicating that additional degrees of freedom can help to reconcile simulations with observations, especially in South East Asia and East Asia where these stations are located.

### 3.3 Global and regional emission increments

We are primarily interested in the additional information provided by the assimilation of  $\delta^{13}\text{C}(\text{CH}_4)$  data. Rather than discussing the regional and global  $\text{CH}_4$  emissions and comparing these results to previous estimates, we investigate the differences between emissions inferred from configurations with and without the additional isotopic constraint. Long-term inversions will be run in the future with this system to provide more robust estimates of  $\text{CH}_4$  emissions and compare them to the existing literature.



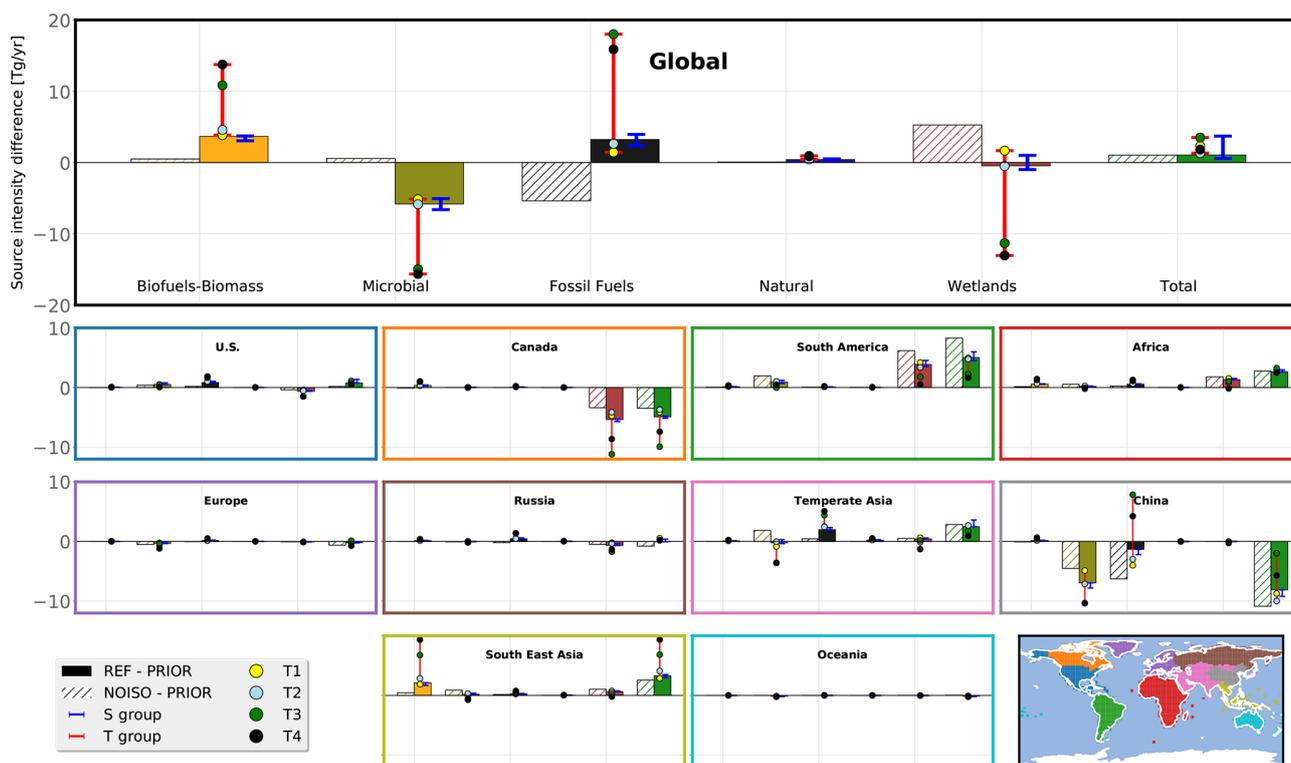
**Figure 5.** RMSE distribution over the surface stations. Panels (a, c) show the prior  $\text{CH}_4$  and  $\delta^{13}\text{C}(\text{CH}_4)$  RMSE, and panels (b, d) show the posterior RMSE. Panels (a, b) show  $\delta^{13}\text{C}(\text{CH}_4)$  RMSE, and panels (c, d) show  $\text{CH}_4$  RMSE. For clarity reasons, S1 prior is not shown for  $\delta^{13}\text{C}(\text{CH}_4)$  because the associated prior misfit is much larger than that of the other configurations. The box-and-whisker plots are covering the whole range of data. In panel (c), all station labels are identical; therefore, most of them are removed to improve the clarity.

The inversion time window is the 2012–2017 period. However, flux and source signature estimates in the 2012–2013 and 2016–2017 periods are not interpreted, as the system appears to require a 2-year spin-up (2012–2013) and a 2-year spin-down (2016–2017), over which the inversion problem is not sufficiently constrained, and isotopic signatures vary widely over time. Therefore, only the 2014–2015 estimates are analyzed in Sects. 3.3 and 3.4. Figure S2 in the Supplement shows the time series of isotopic signatures and illustrates this choice. These long effects are certainly caused by the relatively long relaxation timescales of isotopic ratios in the atmosphere (Tans, 1997) compared to that of total  $\text{CH}_4$ . Fully understanding this would require a lot of time and running multiple inversions (or possibly only tangent linear simulations), starting from different initial conditions spanning the prescribed uncertainty envelope, to infer until when the initial atmospheric isotopic ratios and/or isotopic source signatures can influence the time series of atmospheric isotopic ratios. This was too much work for this study but will certainly be addressed in future studies.

Figure 6 shows global and regional increments from the NOISO and REF inversions relative to prior estimates. Hereinafter, these differences will be referred to as REF increment (REF – PRIOR) and NOISO increment (NOISO –

PRIOR). The difference between both increments will be called an increment difference. Note that prior emissions are identical for all configurations. Posterior total emissions is  $594.6 \pm 1.2 \text{ Tg CH}_4 \text{ yr}^{-1}$  over all configurations, indicating that the isotopic constraint and setup configurations do not significantly affect the posterior global emissions. A higher discrepancy between the budgets would have indicated a malfunction in the system, as the prescribed sinks are identical. The small associated standard deviation is likely caused by a slight difference in the fitting to the observations and/or by the spatial variability in the prescribed sink coupled with a small relocation of emissions, depending on the configuration. For instance, OH concentrations are larger in the tropics, and a relocation of emissions from the tropics to higher latitudes would be compensated for by larger global emissions. Between REF and NOISO, there is only a difference of  $0.02 \text{ Tg CH}_4 \text{ yr}^{-1}$ . We can therefore conclude that the additional isotopic constraint either relocates the emissions or reallocates them between categories, as intended. All but one of the emission categories exhibit large changes between NOISO and REF, namely WET, FFs, AGW, and BB categories.

Overall, increments are large in regions with high emissions. Global increment differences (between REF and



**Figure 6.** REF and NOISO emission increments for the 2014–2015 period. Prior estimates (PRIOR) are identical for both configurations. The color-filled bars show the differences between REF posterior and prior estimates (REF increment). The hatched bars show the differences between NOISO posterior and prior estimates (NOISO increment). The upper panel refers to the global emissions. The lower panels refer to multiple regions of the globe. The regions are shown on the lower right panel. Red and blue error bars represent the minimum and maximum of the T group and S group, respectively. Circles on the red error bar show the results from the T group.

NOISO) in AGW ( $-6.4 \text{ Tg CH}_4 \text{ yr}^{-1}$ ) and FF emissions ( $+8.6 \text{ Tg CH}_4 \text{ yr}^{-1}$ ) are mainly due to regional increment differences in China and temperate Asia. AGW regional increment differences are equal to  $-2.1 \text{ Tg CH}_4 \text{ yr}^{-1}$  in temperate Asia and in China. Similarly, FF regional increment differences are equal to  $+1.5 \text{ Tg CH}_4 \text{ yr}^{-1}$  in temperate Asia and  $+5.0 \text{ Tg CH}_4 \text{ yr}^{-1}$  in China. The WET global increment difference ( $-5.7 \text{ Tg CH}_4 \text{ yr}^{-1}$ ) is mainly due to differences in Canada ( $-2.0 \text{ Tg CH}_4 \text{ yr}^{-1}$ ) and South America ( $-2.3 \text{ Tg CH}_4 \text{ yr}^{-1}$ ), but other regions such as Russia, temperate Asia, and South East Asia are involved. BB emissions are also modified when implementing the isotopic constraint. Their global increment difference is equal to  $+3.2 \text{ Tg CH}_4 \text{ yr}^{-1}$ , principally owing to regional increment differences in South East Asia ( $+1.7 \text{ Tg CH}_4 \text{ yr}^{-1}$ ), Canada ( $+0.4 \text{ Tg CH}_4 \text{ yr}^{-1}$ ), and Africa ( $+0.4 \text{ Tg CH}_4 \text{ yr}^{-1}$ ). The NAT category exhibits very few changes (less than  $1 \text{ Tg CH}_4 \text{ yr}^{-1}$ ), even in relative values.

S group configurations infer posterior results that are consistent with REF, with only small variations, depending on the category and the region (see Table S5 in the Supplement). In particular, S1 provides roughly the same results as REF, but with more iterations, highlighting again that offsetting the

initial conditions can help to reduce the computational burden without affecting the results. On the contrary, T-group configurations are affecting the increments, although T1 and T2 configurations are generally much closer to REF than T3 and T4. T1 (yellow dot; Fig. 6) and T2 (blue dot; Fig. 6) exhibit differences with the S group mainly in China, where WET and FF increments are modified ( $\sim -3 \text{ Tg CH}_4 \text{ yr}^{-1}$ ). More importantly, almost freezing the isotopic signatures to their prior values (T3 and T4) results in increment differences 3 to 4 times larger than with REF, i.e., more than  $10 \text{ Tg CH}_4 \text{ yr}^{-1}$  at the global scale. It highlights the dependence of the inferred  $\text{CH}_4$  emissions on the prior source signatures estimates. In other words, the quality of isotopic source signatures (values and uncertainties) appears to be critical for the robustness of emissions estimates.

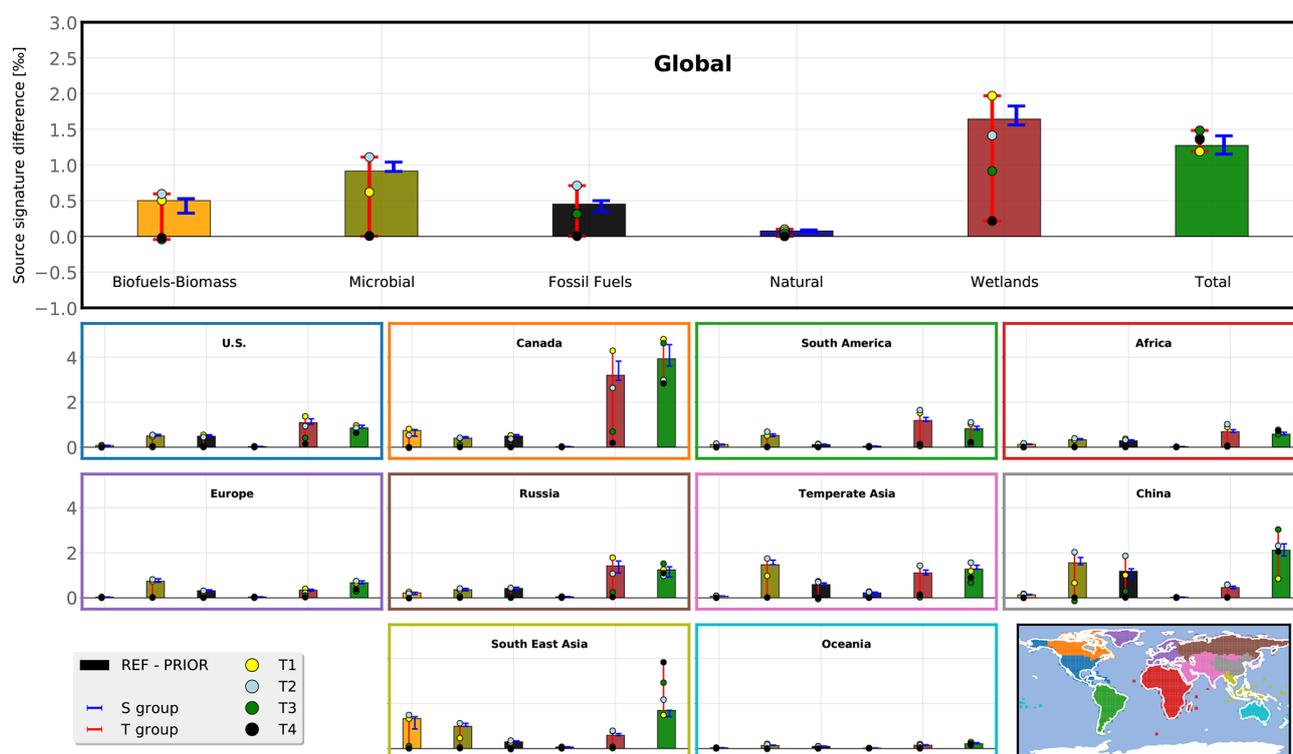
### 3.4 Global and regional source signature increments

Isotopic source signatures are also optimized by the system. Figure 7 provides the differences of flux-weighted source signatures between REF posterior and prior estimates for different regions and each emission category.

With configurations that allow the source signatures to be optimized, all source signatures are shifted upwards by the

**Table 3.** Global posterior  $\text{CH}_4$  emissions by source category and region ( $\text{Tg CH}_4 \text{ yr}^{-1}$ ) for the REF configuration. Uncertainties are reported as the min–max range of all configurations in square brackets.

	BB	AGW	FF	NAT	WET	Total
USA	1 [1–1]	22 [21–22]	14 [13–15]	2 [2–2]	17 [16–17]	56 [55–56]
Canada	2 [1–3]	2 [2–2]	2 [2–2]	1 [1–1]	21 [16–23]	29 [24–30]
South America	2 [2–3]	30 [29–31]	6 [6–6]	5 [5–5]	53 [50–55]	96 [93–99]
Africa	9 [8–10]	25 [25–26]	14 [13–15]	4 [4–4]	28 [26–28]	80 [80–80]
Europe	1 [1–1]	20 [19–20]	6 [6–7]	2 [2–2]	4 [4–4]	34 [33–34]
Russia	2 [2–2]	5 [5–5]	12 [12–13]	3 [3–3]	12 [11–13]	35 [34–36]
Temperate Asia	3 [3–3]	54 [51–56]	28 [27–31]	7 [7–7]	13 [11–13]	105 [104–106]
China	5 [5–5]	29 [26–32]	24 [19–33]	1 [1–1]	5 [5–5]	64 [61–70]
South East Asia	11 [9–18]	23 [22–23]	8 [7–8]	4 [3–4]	22 [21–23]	66 [66–72]
Oceania	1 [0–1]	4 [4–5]	2 [2–2]	1 [1–1]	3 [3–3]	11 [11–11]
Others	1 [1–1]	4 [4–4]	5 [5–5]	8 [8–8]	2 [2–2]	19 [19–19]
Global	37 [33–47]	220 [210–226]	119 [111–134]	38 [38–39]	180 [167–185]	594 [594–597]

**Figure 7.** REF flux-weighted source signature increments for the 2014–2015 period. The color-filled bars show the differences between REF posterior and prior estimates (REF increment). The upper panel refers to the global emissions. The lower panels refer to multiple regions of the globe. The regions are shown on the lower right panel. Red and blue error bars represent the minimum and maximum of the T group and S group, respectively. Circles on error bars show the results from the T group.

inversions in order to correct the excessively strong negative trend in  $\delta^{13}\text{C}(\text{CH}_4)$ . At the global scale, the flux-weighted source signatures of WET, FFs, AGW, BB, and NAT are increased by 1.7, 0.5, 0.9, 0.5, and 0.1‰, respectively. The global source signature is increased from  $-53.9\text{‰}$  (prior) to  $-52.6 \pm 0.2\text{‰}$  (posterior, with a standard deviation over the configurations). More information is provided in Table S6 in the Supplement. The posterior global signature is strongly

dependent on the KIE of atmospheric oxidation. This effect tends to deplete air in  $^{13}\text{CH}_4$ , shifting the  $\delta^{13}\text{C}(\text{CH}_4)$  to more positive values as the  $\text{CH}_4$  molecules emitted by the sources are removed from the atmosphere. The mean KIE in our simulations depends on (1) the prescribed OH,  $\text{O}(^1\text{D})$ , and Cl concentrations and (2) the prescribed KIE values associated to the individual sinks. As the mean KIE is the same for

all configurations, the posterior global source signatures are very close.

The WET global source signature, associated with REF posterior estimates, exhibits the larger upward shift compared to prior estimates, from a value of  $-60.8\%$  to  $-59.1\%$ . Large upward WET source signature shifts are located in boreal regions (North America and Russia) but also in South America and temperate Asia. The AGW source signature is increased by  $0.9\%$ , mainly due to changes in Asia. The FF source signature is increased by  $0.5\%$  due to a large increment in China ( $+1.2\%$ ). Finally, the BB source signature is modified in South East Asia ( $+1.4\%$ ) and Canada ( $+0.8\%$ ).

These changes are consistent within the S group (see blue error bars in Fig. 7), although small variations are visible (e.g.,  $\pm 0.3\%$  for WET in Canada). The source signature is therefore modified nearly to the same extent in all regions, no matter which configuration in the S group is analyzed. T1 (see the yellow dot in Fig. 7), with more optimized categories than the others, shows small differences at the global scale (less than  $0.3\%$  for all categories), although differences of more than  $1\%$  are found in China. Therefore, increasing the number of degrees of freedom leads to similar flux estimates but can affect the signatures at a regional scale. T2 estimates are shifted upward to reach a less negative global isotopic source signature without coming closer to the regional distribution of the S group. This is likely caused by the scarcity of  $\delta^{13}\text{C}(\text{CH}_4)$  stations, and correcting this behavior seems challenging without additional observations. The problem might be circumvented by using the regional scale rather than the pixel scale to optimize isotopic signature values. Future inversions will test this assumption.

These results must be interpreted with caution because the input data suffer from high uncertainties. The artificial increase in the source signatures by our system can hardly be related to the literature and former investigations. Consequently, it is challenging to conclude whether an increase in the source signatures would be more realistic (i.e., supported by observational data) than, for instance, only increasing the emissions of  $^{13}\text{C}$ -enriched sources such as BB. This system is only based on a mathematical and physical framework connecting the several groups of uncertainties (observational, prior fluxes, prior source signatures, and prior sinks) and finding the most likely solution. Better estimates of these uncertainties must be prescribed before obtaining robust results. In particular, the uncertainties on KIE values and sink intensities have not been tested here and could largely influence the results. Also, the uncertainties on source signatures are relatively smooth in REF compared to recent country-specific estimates (Sherwood et al., 2017). Assessing these uncertainties should be a key aspect for future studies using this new inversion system to quantify the global  $\text{CH}_4$  budget.

### 3.5 Posterior uncertainties

Formally, posterior uncertainties are given by the Hessian of the cost function. This matrix can hardly be computed at an achievable cost, considering the size of the inverse problem. Other means must be implemented to obtain the posterior uncertainty, such as estimating a lower-rank approximation of the Hessian, using Monte Carlo ensembles of the variational inversion to represent the prior uncertainties, or computing multiple configurations covering a given range of possibilities. Here, using multiple configurations provides insight into the posterior uncertainty associated with the posterior fluxes. We calculated the full uncertainty range using the minimum and maximum values among all the configurations, as in Saunio et al. (2020). WET, AGW, FF, and BB flux estimates (Table 3) exhibit an uncertainty of 10 %, 7 %, 19 %, and 38 %, respectively. BB is the most uncertain estimate relative to its intensity, although FF shows the largest absolute uncertainty ( $23 \text{ Tg CH}_4 \text{ yr}^{-1}$ ). These uncertainties are unlikely to be affected by the assimilation of additional  $\delta^{13}\text{C}(\text{CH}_4)$  data because we expect the uncertainties on the isotopic source signatures to have a much larger influence. However, this remains to be tested in future work if posterior uncertainties can be calculated.

At present, MIQN3 is not the only optimization algorithm that can be utilized to perform variational inversions in the CIF. The CONGRAD algorithm (Fisher, 1998), which follows a conjugate gradient method combined with a Lanczos algorithm, is also implemented. In particular, it facilitates the computation of posterior uncertainties considerably. Any change in algorithm is very easy and accessible to any CTM embedded in the CIF. However, CONGRAD has not yet been tested with  $\delta^{13}\text{C}(\text{CH}_4)$  data. As CONGRAD is only designed for linear problems, using this algorithm could radically change the results of the inversions performed with the isotopic constraints, and future work will focus on using CONGRAD to perform the inversions with isotopic constraints.

## 4 Conclusions and perspectives

We present here a new variational inversion system designed to assimilate the observations of both a specific trace gas and its isotopic data. This system allows us to optimize both tracer emissions and associated isotopic signatures for multiple source categories. To test this system, we have assimilated  $\text{CH}_4$  and  $\delta^{13}\text{C}(\text{CH}_4)$  data retrieved at different measurement sites over the globe.

Different configurations have been tested in order to assess the sensitivity of the system to the setup. We have shown that offsetting the  $\delta^{13}\text{C}(\text{CH}_4)$  initial conditions before the inversion (S1), using  $\delta^{13}\text{C}(\text{CH}_4)$  curve fitting data instead of the original observations (S2) and reducing the prescribed uncertainties in the  $\delta^{13}\text{C}(\text{CH}_4)$  observations (S3), has very

little effect on the inferred fluxes (less than  $2\text{ Tg CH}_4\text{ yr}^{-1}$  for each category at the global scale). However, offsetting  $\delta^{13}\text{C}(\text{CH}_4)$  initial conditions before the inversion results in a reduced computational time (21 less iterations).

Other setup choices have more influence on the results. Increasing the number of source categories (T1) requires more computational time (10 more iterations) to reach a cost function (and RMSE) reduction similar to REF. Moreover, although the global posterior emissions with an increased number of categories are very close to those inferred with REF (less than  $1\text{ Tg CH}_4\text{ yr}^{-1}$ ), the posterior isotopic signatures can be modified in some regions (more than 1‰ in China). Also, starting from globally averaged values for the source signatures (T2) makes the system unable to retrieve the region-specific isotopic signatures from REF. Increasing the number of  $\delta^{13}\text{C}(\text{CH}_4)$  observations could help us to cope with this issue. Finally, configurations constraining the source signatures (T3–T4) show differences in global flux estimates of more than  $10\text{ Tg CH}_4\text{ yr}^{-1}$ , compared to REF. This emphasizes the need for good prior source signature estimates.

The major drawback of this inversion system is undoubtedly the large computational burden of a full minimization process. At least 40 iterations appear to be necessary to reach a satisfying convergence state at the regional scale. For the LMDz-SACS model, a maximum of eight CPUs can be run in parallel, resulting in an elapsed time of 5–6 weeks to run one of the inversions of this study. A new generation of transport models, such as DYNAMICO (Dubos et al., 2015), could help to address this problem in the future by allowing more processors to run in parallel. Also, further developments will implement some parallelization methods to enable computational burden reduction (e.g., Chevallier, 2013). In addition, variational inversions, as implemented in the CIF, do not provide a quantification (even approximated) of the posterior uncertainties. Dedicated efforts need to be made to address this issue in the future and at an achievable numerical cost. In particular, using the CONGRAD algorithm instead of M1QN3 could be a solution, as both algorithms can be easily selected in the CIF. However, additional work is needed to ensure that switching the optimization algorithm does not affect the results inferred with our new system.

This system is implemented within the CIF framework and can therefore be used for inversions with the various CTMs embedded in the CIF, provided the adjoint codes of the models exist. As the operations developed for the purpose of this study are performed outside the model structure, forward, tangent linear, and adjoint codes from other CTMs do not require any modifications as long as the model is capable of simulating both  $^{12}\text{CH}_4$  and  $^{13}\text{CH}_4$  simultaneously. The prior input must be adapted to the new model (spatial and time resolution), but the format of the observational data and of the prescribed errors can be preserved. Also, due to the variational method benefits, the efforts dedicated to the preparation of inputs do not scale with either the size of the ob-

servational datasets or the length of the simulation time window. Therefore, this system is very powerful and is particularly relevant to study, in a consistent way, the influence of multiple physical parameters on atmospheric isotopic ratios such as the transport, the isotopic signatures, the emission scenarios, the KIE values, etc. We did not try to assess the sensitivity of the system to these parameters here as only the technical aspects of the system were tested. This will be part of future analyses.

As mentioned in the introduction, future work will address the estimation of  $\text{CH}_4$  emissions over longer periods of time using this new system. For instance, the 2000–2006  $\text{CH}_4$  stabilization period and the subsequent renewed growth are particularly interesting to study using the isotopic constraint as global  $\delta^{13}\text{C}(\text{CH}_4)$  started to decrease after 2006. These periods of time have already attracted considerable critical attention from many inversion studies (with or without the isotopic constraint) and comparing the results derived from such a complete 3-D variational inversion system with other recent estimates should be highly relevant. The most important limitation of assimilating  $\delta^{13}\text{C}(\text{CH}_4)$  lies in the fact that very limited  $\delta^{13}\text{C}(\text{CH}_4)$  data are available, and therefore, evaluating the posterior simulated  $\delta^{13}\text{C}(\text{CH}_4)$  is often challenging, if not impossible. However, satellite and balloon/AirCore data can easily be used to evaluate the posterior simulated  $\text{CH}_4$ .

$\delta^{13}\text{C}(\text{CH}_4)$  is not the only kind of isotopic data that can be assimilated in such a system. Many  $\delta\text{D}(\text{CH}_4)$  observations have also been retrieved during the 2004–2010 period at many different locations. These isotopic values can provide additional information that can further help to discriminate the co-emitted  $\text{CH}_4$  fluxes (Rigby et al., 2012). Moreover, ethane ( $\text{C}_2\text{H}_6$ ) is co-emitted with  $\text{CH}_4$  by fossil fuel extraction and distribution (Kort et al., 2016; Smith et al., 2015), and observations have been available at a multitude of sites since the early 1980s. Therefore, assimilating these data can provide additional constraints. The system will therefore be improved in the future in order to assimilate  $\delta^{13}\text{C}(\text{CH}_4)$ ,  $\delta\text{D}(\text{CH}_4)$  and  $\text{C}_2\text{H}_6$  observations together.

*Data availability.* The code files of the CIF version used in the present paper are registered under the following link: <https://doi.org/10.5281/zenodo.6304912> (Berchet et al., 2022). Prior anthropogenic fluxes (EDGARv4.3.2) can be downloaded from the EDGAR website ([https://edgar.jrc.ec.europa.eu/dataset\\_ghg432](https://edgar.jrc.ec.europa.eu/dataset_ghg432), last access: 23 February 2022). Biomass burning fluxes can be downloaded from the GFED website (<https://globalfiredata.org/pages/data/>, last access: 23 February 2022). Prior natural fluxes and other data are available upon request ([joel.thanwerdas@lsce.ipsl.fr](mailto:joel.thanwerdas@lsce.ipsl.fr)). The  $\text{CH}_4$  data used in the present paper were provided by many stations and measurement networks around the world (a comprehensive list can be found in the Supplement and in the acknowledgements). Their data are freely available upon request to the station maintainers or via dedicated websites. The  $\delta^{13}\text{C}(\text{CH}_4)$  observational data can be downloaded from the NOAA-GML website (<https://gml.noaa.gov/dv/data/>, last access: 23 February 2022).

*Supplement.* The supplement related to this article is available online at: <https://doi.org/10.5194/gmd-15-4831-2022-supplement>.

*Author contributions.* JT implemented the variational inversion system within the CIF, with the precious help of AB. JT designed, ran, and analyzed the tested configurations. AB, MS, IP, and PB provided scientific and technical expertise. They also contributed to the analysis of this work. BHV and SEM provided the  $\delta^{13}\text{C}(\text{CH}_4)$  data and scientific expertise regarding  $\delta^{13}\text{C}(\text{CH}_4)$  observations. JT prepared the paper, with contributions from all co-authors.

*Competing interests.* The contact author has declared that neither they nor their co-authors have any competing interests.

*Disclaimer.* Publisher's note: Copernicus Publications remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

*Acknowledgements.* This work has been supported by the CEA (Commissariat à l'Énergie Atomique et aux Énergies Alternatives). The study relies extensively on the meteorological data provided by the ECMWF. Calculations were performed using the computing resources of LSCE, maintained by François Marabelle and the LSCE IT team. The authors thank the reviewers (Peter Rayner and two anonymous referees) for their fruitful comments and suggestions on our work. We are also grateful to many station maintainers that provided the  $\text{CH}_4$  data. For the AGAGE network, we note that AGAGE is supported principally by NASA (USA) grants to MIT and SIO and also by BEIS (UK) and NOAA (USA) grants to Bristol University, CSIRO and BoM (Australia), FOEN grants to Empa (Switzerland), NILU (Norway), SNU (South Korea), CMA (China), NIES (Japan), and Urbino University (Italy). At the South African Weather Service Cape Point GAW station, we thank Casper Labuschagne, Thumeka Mkololo, and Warren Joubert. For the EC network, we thank Doug Worthy. For the MGO network, we thank Nina Paramonova and Victor Ivakhov. For the AMY network, we thank Haeyoung Lee and Sepyo Lee. The AMY data were funded by the Korea Meteorological Administration Research and Development Program (grant no. KMA2018-00522). For the ENEA network, we thank Alcide Disarra, Salvatore Piacentino, and Damiano Sferlazzo. For the EMPA network, we thank Martin Steinbacher and Brigitte Buchmann. The  $\text{CH}_4$  measurements at JFJ are run by EMPA in collaboration with FOEN and are also supported by ICOS-CH. For the BMKG-EMPA network, we thank Martin Steinbacher, Alberth Christian, and Sugeng Nugroho. The  $\text{CH}_4$  measurements at BKT are run by BMKG in collaboration with EMPA.

*Financial support.* This research has been supported by the Commissariat à l'Énergie Atomique et aux Énergies Alternatives (grant no. CFR 2018).

*Review statement.* This paper was edited by Richard Neale and reviewed by Peter Rayner and two anonymous referees.

## References

- Berchet, A., Sollum, E., Thompson, R. L., Pison, I., Thanwerdas, J., Broquet, G., Chevallier, F., Aalto, T., Berchet, A., Bergamaschi, P., Brunner, D., Engelen, R., Fortems-Cheiney, A., Gerbig, C., Groot Zwaaftink, C. D., Haussaire, J.-M., Henne, S., Houweling, S., Karstens, U., Kutsch, W. L., Lujckx, I. T., Monteil, G., Palmer, P. I., van Peet, J. C. A., Peters, W., Peylin, P., Potier, E., Rödenbeck, C., Saunio, M., Scholze, M., Tsuruta, A., and Zhao, Y.: The Community Inversion Framework v1.0: a unified system for atmospheric inversion studies, *Geosci. Model Dev.*, 14, 5331–5354, <https://doi.org/10.5194/gmd-14-5331-2021>, 2021.
- Berchet, A., Sollum, E., Pison, I., Thompson, R. L., Thanwerdas, J., Fortems-Cheiney, A., van Peet, J. C. A., Potier, E., Chevallier, F., Broquet, G., and Berchet, A.: The Community Inversion Framework: codes and documentation (v1.1), Zenodo [code], <https://doi.org/10.5281/zenodo.6304912>, 2022.
- Bergamaschi, P., Lubina, C., Königstedt, R., Fischer, H., Veltkamp, A. C., and Zwaagstra, O.: Stable isotopic signatures ( $\delta^{13}\text{C}$ ,  $\delta\text{D}$ ) of methane from European landfill sites, *J. Geophys. Res.-Atmos.*, 103, 8251–8265, <https://doi.org/10.1029/98JD00105>, 1998.
- 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  $\text{CH}_4$  emissions 2001–2006, *J. Geophys. Res.-Atmos.*, 115, D22309, <https://doi.org/10.1029/2010JD014180>, 2010.
- Bergamaschi, P., Houweling, S., Segers, A., Krol, M., Frankenberg, C., Scheepmaker, R. A., Dlugokencky, E., Wofsy, S. C., Kort, E. A., Sweeney, C., Schuck, T., Brenninkmeijer, C., Chen, H., Beck, V., and Gerbig, C.: Atmospheric  $\text{CH}_4$  in the first decade of the 21st century: Inverse modeling analysis using SCIAMACHY satellite retrievals and NOAA surface measurements, *J. Geophys. Res.-Atmos.*, 118, 7350–7369, <https://doi.org/10.1002/jgrd.50480>, 2013.
- Bergamaschi, P., Karstens, U., Manning, A. J., Saunio, 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  $\text{CH}_4$  emissions during 2006–2012 using different inverse models and reassessed atmospheric observations, *Atmos. Chem. Phys.*, 18, 901–920, <https://doi.org/10.5194/acp-18-901-2018>, 2018.
- Bousquet, P., Ciais, P., Miller, J. B., Dlugokencky, E. J., Hauglustaine, D. A., Prigent, C., Van der Werf, G. R., Peylin, P., Brunke, E.-G., Carouge, C., Langenfelds, R. L., Lathière, J., Papa, F., Ramonet, M., Schmidt, M., Steele, L. P., Tyler, S. C., and White, J.: Contribution of anthropogenic and natural sources to atmospheric methane variability, *Nature*, 443, 439–443, <https://doi.org/10.1038/nature05132>, 2006.

- Bréas, O., Guillou, C., Reniero, F., and Wada, E.: The Global Methane Cycle: Isotopes and Mixing Ratios, Sources and Sinks, *Isot. Environ. Health. S.*, 37, 257–379, <https://doi.org/10.1080/10256010108033302>, 2001.
- Chang, J., Peng, S., Ciais, P., Saunois, M., Dangal, S. R. S., Herrero, M., Havlík, P., Tian, H., and Bousquet, P.: Revisiting enteric methane emissions from domestic ruminants and their  $\delta^{13}\text{CCH}_4$  source signature, *Nat. Commun.*, 10, 3420, <https://doi.org/10.1038/s41467-019-11066-3>, 2019.
- Chanton, J. P., Rutkowski, C. M., and Mosher, B.: Quantifying Methane Oxidation from Landfills Using Stable Isotope Analysis of Downwind Plumes, *Environ. Sci. Technol.*, 33, 3755–3760, <https://doi.org/10.1021/es9904033>, 1999.
- Chanton, J. P., Rutkowski, C. M., Schwartz, C. C., Ward, D. E., and Boring, L.: Factors influencing the stable carbon isotopic signature of methane from combustion and biomass burning, *J. Geophys. Res.-Atmos.*, 105, 1867–1877, <https://doi.org/10.1029/1999JD900909>, 2000.
- Chevallier, F.: On the parallelization of atmospheric inversions of  $\text{CO}_2$  surface fluxes within a variational framework, *Geosci. Model Dev.*, 6, 783–790, <https://doi.org/10.5194/gmd-6-783-2013>, 2013.
- Chevallier, F., Fisher, M., Peylin, P., Serrar, S., Bousquet, P., Bréon, F.-M., Chédin, A., and Ciais, P.: Inferring  $\text{CO}_2$  sources and sinks from satellite observations: Method and application to TOVS data, *J. Geophys. Res.*, 110, D24309, <https://doi.org/10.1029/2005JD006390>, 2005.
- Craig, H.: Isotopic standards for carbon and oxygen and correction factors for mass-spectrometric analysis of carbon dioxide, *Geochim. Cosmochim. Ac.*, 12, 133–149, [https://doi.org/10.1016/0016-7037\(57\)90024-8](https://doi.org/10.1016/0016-7037(57)90024-8), 1957.
- Dlugokencky, E.: NOAA/GML, [https://www.esrl.noaa.gov/gmd/ccgg/trends\\_ch4/](https://www.esrl.noaa.gov/gmd/ccgg/trends_ch4/) (last access: 23 February 2022), 2021.
- Dubos, T., Dubey, S., Tort, M., Mittal, R., Meurdesoif, Y., and Hourdin, F.: DYNAMICO-1.0, an icosahedral hydrostatic dynamical core designed for consistency and versatility, *Geosci. Model Dev.*, 8, 3131–3150, <https://doi.org/10.5194/gmd-8-3131-2015>, 2015.
- Enting, I. G. and Newsam, G. N.: Atmospheric constituent inversion problems: Implications for baseline monitoring, *J. Atmos. Chem.*, 11, 69–87, <https://doi.org/10.1007/BF00053668>, 1990.
- Etiopie, G.: *Natural Gas Seepage: The Earth's Hydrocarbon Degassing*, Springer International Publishing, Switzerland, <https://www.springer.com/gp/book/9783319146003> (last access: 23 February 2022), 2015.
- Etminan, M., Myhre, G., Highwood, E. J., and Shine, K. P.: Radiative forcing of carbon dioxide, methane, and nitrous oxide: A significant revision of the methane radiative forcing, *Geophys. Res. Lett.*, 43, 12,614–12,623, <https://doi.org/10.1002/2016GL071930>, 2016.
- Fisher, M.: Minimization algorithms for variational data assimilation, <http://www.ecmwf.int/en/eLibrary/9400-minimization-algorithms-variational-data-assimilation> (last access: 23 August 2021), 1998.
- Fletcher, S. E. M., Tans, P. P., Bruhwiler, L. M., Miller, J. B., and Heimann, M.:  $\text{CH}_4$  sources estimated from atmospheric observations of  $\text{CH}_4$  and its  $^{13}\text{C}/^{12}\text{C}$  isotopic ratios: 2. Inverse modeling of  $\text{CH}_4$  fluxes from geographical regions, *Global Biogeochemical Cycles*, 18, GB4005, <https://doi.org/10.1029/2004GB002224>, 2004.
- Fujita, R., Morimoto, S., Maksyutov, S., Kim, H.-S., Arshinov, M., Brailsford, G., Aoki, S., and Nakazawa, T.: Global and Regional  $\text{CH}_4$  Emissions for 1995–2013 Derived From Atmospheric  $\text{CH}_4$ ,  $\delta^{13}\text{C-CH}_4$ , and  $\delta\text{D-CH}_4$  Observations and a Chemical Transport Model, *J. Geophys. Res.-Atmos.*, 125, e2020JD032903, <https://doi.org/10.1029/2020JD032903>, 2020.
- Ganesan, A. L., Stell, A. C., Gedney, N., Comyn-Platt, E., Hayman, G., Rigby, M., Poulter, B., and Hornibrook, E. R. C.: Spatially Resolved Isotopic Source Signatures of Wetland Methane Emissions, *Geophys. Res. Lett.*, 45, 3737–3745, <https://doi.org/10.1002/2018GL077536>, 2018.
- Gilbert, J. C. and Lemaréchal, C.: Some numerical experiments with variable-storage quasi-Newton algorithms, *Math. Program.*, 45, 407–435, <https://doi.org/10.1007/BF01589113>, 1989.
- GLOBALVIEW-CH4: Cooperative Atmospheric Data Integration Project – Methane, CD-ROM, also available on Internet via anonymous FTP to <ftp://ftp.cmdl.noaa.gov> (last access: 23 February 2022), Path: <cgg/ch4/GLOBALVIEW>, NOAA ESRL, Boulder, Colorado, 2009.
- 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., Maki, T., Maksyutov, S., Masarie, K., Peylin, P., Prather, M., Sarmiento, J., Taguchi, S., Takahashi, T., and Yuen, C.-W.: Towards robust regional estimates of  $\text{CO}_2$  sources and sinks using atmospheric transport models, *Nature*, 415, 626–630, <https://doi.org/10.1038/415626a>, 2002.
- Holmes, M. E., Sansone, F. J., Rust, T. M., and Popp, B. N.: Methane production, consumption, and air–sea exchange in the open ocean: An Evaluation based on carbon isotopic ratios, *Global Biogeochem. Cy.*, 14, 1–10, <https://doi.org/10.1029/1999GB001209>, 2000.
- Hourdin, F., Musat, I., Bony, S., Braconnot, P., Codron, F., Dufresne, J.-L., Fairhead, L., Filiberti, M.-A., Friedlingstein, P., Grandpeix, J.-Y., Krinner, G., LeVan, P., Li, Z.-X., and Lott, F.: The LMDZ4 general circulation model: climate performance and sensitivity to parametrized physics with emphasis on tropical convection, *Clim. Dynam.*, 27, 787–813, <https://doi.org/10.1007/s00382-006-0158-0>, 2006.
- Houweling, S., Bergamaschi, P., Chevallier, F., Heimann, M., Kaminski, T., Krol, M., Michalak, A. M., and Patra, P.: Global inverse modeling of  $\text{CH}_4$  sources and sinks: an overview of methods, *Atmos. Chem. Phys.*, 17, 235–256, <https://doi.org/10.5194/acp-17-235-2017>, 2017.
- Ide, K., Courtier, P., Ghil, M., and Lorenc, A. C.: Unified Notation for Data Assimilation: Operational, Sequential and Variational (gtSpecial Issue) Data Assimilation in Meteorology and Oceanography: Theory and Practice), *J. Meteorol. Soc. Jpn. Ser. II*, 75, 181–189, [https://doi.org/10.2151/jmsj1965.75.1B\\_181](https://doi.org/10.2151/jmsj1965.75.1B_181), 1997.
- Janssens-Maenhout, G., Crippa, M., Guizzardi, D., Muntean, M., Schaaf, E., Dentener, F., Bergamaschi, P., Pagliari, V., Olivier, J. G. J., Peters, J. A. H. W., van Aardenne, J. A., Monni, S., Doering, U., and Petrescu, A. M. R.: EDGAR v4.3.2 Global Atlas of the three major Greenhouse Gas Emissions for the period 1970–2012, *Earth Syst. Sci. Data Discuss.* [preprint], <https://doi.org/10.5194/essd-2017-79>, 2017.

- Kirschke, S., Bousquet, P., Ciais, P., Saunoy, M., Canadell, J. G., Dlugokencky, E. J., Bergamaschi, P., Bergmann, D., Blake, D. R., Bruhwiler, L., Cameron-Smith, P., Castaldi, S., Chevallier, F., Feng, L., Fraser, A., Heimann, M., Hodson, E. L., Houweling, S., Josse, B., Fraser, P. J., Krummel, P. B., Lamarque, J.-F., Langenfelds, R. L., Le Quééré, C., Naik, V., O'Doherty, S., Palmer, P. I., Pison, I., Plummer, D., Poulter, B., Prinn, R. G., Rigby, M., Ringeval, B., Santini, M., Schmidt, M., Shindell, D. T., Simpson, I. J., Spahni, R., Steele, L. P., Strode, S. A., Sudo, K., Szopa, S., van der Werf, G. R., Voulgarakis, A., van Weele, M., Weiss, R. F., Williams, J. E., and Zeng, G.: Three decades of global methane sources and sinks, *Nat. Geosci.*, 6, 813–823, <https://doi.org/10.1038/ngeo1955>, 2013.
- Klevenhusen, F., Bernasconi, S. M., Kreuzer, M., and Soliva, C. R.: Experimental validation of the Intergovernmental Panel on Climate Change default values for ruminant-derived methane and its carbon-isotope signature, *Anim. Prod. Sci.*, 50, 159, <https://doi.org/10.1071/AN09112>, 2010.
- Kort, E. A., Smith, M. L., Murray, L. T., Gvakharia, A., Brandt, A. R., Peischl, J., Ryerson, T. B., Sweeney, C., and Travis, K.: Fugitive emissions from the Bakken shale illustrate role of shale production in global ethane shift, *Geophys. Res. Lett.*, 43, 4617–4623, <https://doi.org/10.1002/2016GL068703>, 2016.
- Lambert, G. and Schmidt, S.: Reevaluation of the oceanic flux of methane: Uncertainties and long term variations, *Chemosphere*, 26, 579–589, [https://doi.org/10.1016/0045-6535\(93\)90443-9](https://doi.org/10.1016/0045-6535(93)90443-9), 1993.
- Levin, I., Bergamaschi, P., Dörr, H., and Trapp, D.: Stable isotopic signature of methane from major sources in Germany, *Chemosphere*, 26, 161–177, [https://doi.org/10.1016/0045-6535\(93\)90419-6](https://doi.org/10.1016/0045-6535(93)90419-6), 1993.
- Liu, X., Weinbren, A. L., Chang, H., Tadić, J. M., Moun-tain, M. E., Trudeau, M. E., Andrews, A. E., Chen, Z., and Miller, S. M.: Data reduction for inverse modeling: an adaptive approach v1.0, *Geosci. Model Dev.*, 14, 4683–4696, <https://doi.org/10.5194/gmd-14-4683-2021>, 2021.
- Locatelli, R., Bousquet, P., Chevallier, F., Fortems-Cheney, A., Szopa, S., Saunoy, 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, *Atmos. Chem. Phys.*, 13, 9917–9937, <https://doi.org/10.5194/acp-13-9917-2013>, 2013.
- Locatelli, R., Bousquet, P., Saunoy, M., Chevallier, F., and Cres-sot, C.: Sensitivity of the recent methane budget to LMDz sub-grid-scale physical parameterizations, *Atmos. Chem. Phys.*, 15, 9765–9780, <https://doi.org/10.5194/acp-15-9765-2015>, 2015.
- Louis, J.-F.: A parametric model of vertical eddy fluxes in the atmosphere, *Bound.-Lay. Meteorol.*, 17, 187–202, <https://doi.org/10.1007/BF00117978>, 1979.
- Masarie, K. A. and Tans, P. P.: Extension and integration of atmospheric carbon dioxide data into a globally consistent measurement record, *J. Geophys. Res.-Atmos.*, 100, 11593–11610, <https://doi.org/10.1029/95JD00859>, 1995.
- McNorton, J., Wilson, C., Gloor, M., Parker, R. J., Boesch, H., Feng, W., Hossaini, R., and Chipperfield, M. P.: Attribution of recent increases in atmospheric methane through 3-D inverse modelling, *Atmos. Chem. Phys.*, 18, 18149–18168, <https://doi.org/10.5194/acp-18-18149-2018>, 2018.
- Menut, L., Bessagnet, B., Khvorostyanov, D., Beekmann, M., Blond, N., Colette, A., Coll, I., Curci, G., Foret, G., Hodzic, A., Mailler, S., Meleux, F., Monge, J.-L., Pison, I., Siour, G., Turquety, S., Valari, M., Vautard, R., and Vivanco, M. G.: CHIMERE 2013: a model for regional atmospheric composition modelling, *Geosci. Model Dev.*, 6, 981–1028, <https://doi.org/10.5194/gmd-6-981-2013>, 2013.
- Neef, L., van Weele, M., and van Velthoven, P.: Optimal estimation of the present-day global methane budget, *Global Biogeochem. Cy.*, 24, GB4024, <https://doi.org/10.1029/2009GB003661>, 2010.
- Newsam, G. N. and Enting, I. G.: Inverse problems in atmospheric constituent studies. I. Determination of surface sources under a diffusive transport approximation, *Inverse Probl.*, 4, 1037–1054, <https://doi.org/10.1088/0266-5611/4/4/008>, 1988.
- Nisbet, E. G., Manning, M. R., Dlugokencky, E. J., Fisher, R. E., Lowry, D., Michel, S. E., Myhre, C. L., Platt, S. M., Allen, G., Bousquet, P., Brownlow, R., Cain, M., France, J. L., Hermansen, O., Hossaini, R., Jones, A. E., Levin, I., Manning, A. C., Myhre, G., Pyle, J. A., Vaughn, B. H., Warwick, N. J., and White, J. W. C.: Very Strong Atmospheric Methane Growth in the 4 Years 2014–2017: Implications for the Paris Agreement, *Global Biogeochem. Cy.*, 33, 318–342, <https://doi.org/10.1029/2018GB006009>, 2019.
- Patra, P. K., Houweling, S., Krol, M., Bousquet, P., Belikov, D., Bergmann, D., Bian, H., Cameron-Smith, P., Chipperfield, M. P., Corbin, K., Fortems-Cheiney, A., Fraser, A., Gloor, E., Hess, P., Ito, A., Kawa, S. R., Law, R. M., Loh, Z., Maksyutov, S., Meng, L., Palmer, P. I., Prinn, R. G., Rigby, M., Saito, R., and Wilson, C.: TransCom model simulations of  $\text{CH}_4$  and related species: linking transport, surface flux and chemical loss with  $\text{CH}_4$  variability in the troposphere and lower stratosphere, *Atmos. Chem. Phys.*, 11, 12813–12837, <https://doi.org/10.5194/acp-11-12813-2011>, 2011.
- 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  $\text{CO}_2$  surface fluxes from atmospheric trace gas observations, *J. Geophys. Res.-Atmos.*, 110, <https://doi.org/10.1029/2005JD006157>, 2005.
- Pison, I., Bousquet, P., Chevallier, F., Szopa, S., and Hauglustaine, D.: Multi-species inversion of  $\text{CH}_4$ ,  $\text{CO}$  and  $\text{H}_2$  emissions from surface measurements, *Atmos. Chem. Phys.*, 9, 5281–5297, <https://doi.org/10.5194/acp-9-5281-2009>, 2009.
- Poulter, B., Bousquet, P., Canadell, J. G., Ciais, P., Pregon, A., Saunoy, M., Arora, V. K., Beerling, D. J., Brovkin, V., Jones, C. D., Joos, F., Gedney, N., Ito, A., Kleinen, T., Koven, C. D., McDonald, K., Melton, J. R., Peng, C., Peng, S., Prigent, C., Schroeder, R., Riley, W. J., Saito, M., Spahni, R., Tian, H., Taylor, L., Viovy, N., Wilton, D., Wiltshire, A., Xu, X., Zhang, B., Zhang, Z., and Zhu, Q.: Global wetland contribution to 2000–2012 atmospheric methane growth rate dynamics, *Environ. Res. Lett.*, 12, 094013, <https://doi.org/10.1088/1748-9326/aa8391>, 2017.
- Rayner, P. J., Michalak, A. M., and Chevallier, F.: Fundamentals of data assimilation applied to biogeochemistry, *Atmos. Chem. Phys.*, 19, 13911–13932, <https://doi.org/10.5194/acp-19-13911-2019>, 2019.

- Rice, A. L., Butenhoff, C. L., Teama, D. G., Röger, F. H., Khalil, M. A. K., and Rasmussen, R. A.: Atmospheric methane isotopic record favors fossil sources flat in 1980s and 1990s with recent increase, *P. Natl. Acad. Sci. USA*, 113, 10791–10796, <https://doi.org/10.1073/pnas.1522923113>, 2016.
- Rigby, M., Manning, A. J., and Prinn, R. G.: The value of high-frequency, high-precision methane isotopologue measurements for source and sink estimation, *J. Geophys. Res.-Atmos.*, 117, D12312, <https://doi.org/10.1029/2011JD017384>, 2012.
- Rigby, M., Montzka, S. A., Prinn, R. G., White, J. W. C., Young, D., O'Doherty, S., Lunt, M. F., Ganesan, A. L., Manning, A. J., Simmonds, P. G., Salameh, P. K., Harth, C. M., Mühle, J., Weiss, R. F., Fraser, P. J., Steele, L. P., Krummel, P. B., McCulloch, A., and Park, S.: Role of atmospheric oxidation in recent methane growth, *P. Natl. Acad. Sci. USA*, 114, 5373–5377, <https://doi.org/10.1073/pnas.1616426114>, 2017.
- Rödenbeck, C., Houweling, S., Gloor, M., and Heimann, M.:  $\text{CO}_2$  flux history 1982–2001 inferred from atmospheric data using a global inversion of atmospheric transport, *Atmos. Chem. Phys.*, 3, 1919–1964, <https://doi.org/10.5194/acp-3-1919-2003>, 2003.
- Sansone, F. J., Popp, B. N., Gasc, A., Graham, A. W., and Rust, T. M.: Highly elevated methane in the eastern tropical North Pacific and associated isotopically enriched fluxes to the atmosphere, *Geophys. Res. Lett.*, 28, 4567–4570, <https://doi.org/10.1029/2001GL013460>, 2001.
- Saunois, M., Bousquet, P., Poulter, B., Peregón, A., Ciais, P., Canadell, J. G., Dlugokencky, E. J., Etiope, G., Bastviken, D., Houweling, S., Janssens-Maenhout, G., Tubiello, F. N., Castaldi, S., Jackson, R. B., Alexe, M., Arora, V. K., Beerling, D. J., Bergamaschi, P., Blake, D. R., Brailsford, G., Bruhwiler, L., Crevoisier, C., Crill, P., Covey, K., Frankenberg, C., Gedney, N., Höglund-Isaksson, L., Ishizawa, M., Ito, A., Joos, F., Kim, H.-S., Kleinen, T., Krummel, P., Lamarque, J.-F., Langenfelds, R., Locatelli, R., Machida, T., Maksyutov, S., Melton, J. R., Morino, I., Naik, V., O'Doherty, S., Parmentier, F.-J. W., Patra, P. K., Peng, C., Peng, S., Peters, G. P., Pison, I., Prinn, R., Ramonet, M., Riley, W. J., Saito, M., Santini, M., Schroeder, R., Simpson, I. J., Spahni, R., Takizawa, A., Thornton, B. F., Tian, H., Tohjima, Y., Viovy, N., Voulgarakis, A., Weiss, R., Wilton, D. J., Wiltshire, A., Worthy, D., Wunch, D., Xu, X., Yoshida, Y., Zhang, B., Zhang, Z., and Zhu, Q.: Variability and quasi-decadal changes in the methane budget over the period 2000–2012, *Atmos. Chem. Phys.*, 17, 11135–11161, <https://doi.org/10.5194/acp-17-11135-2017>, 2017.
- 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., Murguía-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., Rosentretter, 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., and Zhuang, Q.: The Global Methane Budget 2000–2017, *Earth Syst. Sci. Data*, 12, 1561–1623, <https://doi.org/10.5194/essd-12-1561-2020>, 2020.
- Schaefer, H., Fletcher, S. E. M., Veidt, C., Lassey, K. R., Brailsford, G. W., Bromley, T. M., Dlugokencky, E. J., Michel, S. E., Miller, J. B., Levin, I., Lowe, D. C., Martin, R. J., Vaughn, B. H., and White, J. W. C.: A 21st-century shift from fossil-fuel to biogenic methane emissions indicated by  $^{13}\text{CH}_4$ , *Science*, 352, 80–84, <https://doi.org/10.1126/science.aad2705>, 2016.
- Schwietzke, S., Sherwood, O. A., Bruhwiler, L. M. P., Miller, J. B., Etiope, G., Dlugokencky, E. J., Michel, S. E., Arling, V. A., Vaughn, B. H., White, J. W. C., and Tans, P. P.: Upward revision of global fossil fuel methane emissions based on isotope database, *Nature*, 538, 88–91, <https://doi.org/10.1038/nature19797>, 2016.
- Sherwood, O. A., Schwietzke, S., Arling, V. A., and Etiope, G.: Global Inventory of Gas Geochemistry Data from Fossil Fuel, Microbial and Burning Sources, version 2017, *Earth Syst. Sci. Data*, 9, 639–656, <https://doi.org/10.5194/essd-9-639-2017>, 2017.
- Smith, M. L., Kort, E. A., Karion, A., Sweeney, C., Herndon, S. C., and Yacovitch, T. I.: Airborne Ethane Observations in the Barnett Shale: Quantification of Ethane Flux and Attribution of Methane Emissions, *Environ. Sci. Technol.*, 49, 8158–8166, <https://doi.org/10.1021/acs.est.5b00219>, 2015.
- Tans, P. P.: A note on isotopic ratios and the global atmospheric methane budget, *Global Biogeochem. Cy.*, 11, 77–81, <https://doi.org/10.1029/96GB03940>, 1997.
- Thanwerdas, J., Saunois, M., Berchet, A., Pison, I., Hauglustaine, D., Ramonet, M., Crevoisier, C., Baier, B., Sweeney, C., and Bousquet, P.: Impact of atomic chlorine on the modelling of total methane and its  $^{13}\text{C} : ^{12}\text{C}$  isotopic ratio at global scale, *Atmos. Chem. Phys. Discuss.* [preprint], <https://doi.org/10.5194/acp-2019-925>, 2019.
- Thompson, R. L., Nisbet, E. G., Pisso, I., Stohl, A., Blake, D., Dlugokencky, E. J., Helmig, D., and White, J. W. C.: Variability in Atmospheric Methane From Fossil Fuel and Microbial Sources Over the Last Three Decades, *Geophys. Res. Lett.*, 45, 11,499–11,508, <https://doi.org/10.1029/2018GL078127>, 2018.
- Tiedtke, M.: A Comprehensive Mass Flux Scheme for Cumulus Parameterization in Large-Scale Models, *Mon. Weather Rev.*, 117, 1779–1800, [https://doi.org/10.1175/1520-0493\(1989\)117<1779:ACMFSF>2.0.CO;2](https://doi.org/10.1175/1520-0493(1989)117<1779:ACMFSF>2.0.CO;2), 1989.
- Townsend-Small, A., Tyler, S. C., Pataki, D. E., Xu, X., and Christensen, L. E.: Isotopic measurements of atmospheric methane in Los Angeles, California, USA: Influence of “fugitive” fossil fuel emissions, *J. Geophys. Res.-Atmos.*, 117, D07308, <https://doi.org/10.1029/2011JD016826>, 2012.
- Turner, A. J., Frankenberg, C., Wennberg, P. O., and Jacob, D. J.: Ambiguity in the causes for decadal trends in atmospheric methane and hydroxyl, *P. Natl. Acad. Sci. USA*, 114, 5367–5372, <https://doi.org/10.1073/pnas.1616020114>, 2017.
- van der Werf, G. R., Randerson, J. T., Giglio, L., van Leeuwen, T. T., Chen, Y., Rogers, B. M., Mu, M., van Marle, M. J. E., Morton, D. C., Collatz, G. J., Yokelson, R. J., and Kasibhatla, P. S.: Global

- fire emissions estimates during 1997–2016, *Earth Syst. Sci. Data*, 9, 697–720, <https://doi.org/10.5194/essd-9-697-2017>, 2017.
- Warwick, N. J., Cain, M. L., Fisher, R., France, J. L., Lowry, D., Michel, S. E., Nisbet, E. G., Vaughn, B. H., White, J. W. C., and Pyle, J. A.: Using  $\delta^{13}\text{C}-\text{CH}_4$  and  $\delta\text{D}-\text{CH}_4$  to constrain Arctic methane emissions, *Atmos. Chem. Phys.*, 16, 14891–14908, <https://doi.org/10.5194/acp-16-14891-2016>, 2016.
- Yver, C. E., Pison, I. C., Fortems-Cheiney, A., Schmidt, M., Chevalier, F., Ramonet, M., Jordan, A., Søvde, O. A., Engel, A., Fisher, R. E., Lowry, D., Nisbet, E. G., Levin, I., Hammer, S., Necki, J., Bartyzel, J., Reimann, S., Vollmer, M. K., Steinbacher, M., Aalto, T., Maione, M., Arduini, J., O’Doherty, S., Grant, A., Sturges, W. T., Forster, G. L., Lunder, C. R., Privalov, V., Paramonova, N., Werner, A., and Bousquet, P.: A new estimation of the recent tropospheric molecular hydrogen budget using atmospheric observations and variational inversion, *Atmos. Chem. Phys.*, 11, 3375–3392, <https://doi.org/10.5194/acp-11-3375-2011>, 2011.
- Zazzeri, G., Lowry, D., Fisher, R. E., France, J. L., Lanoisellé, M., Kelly, B. F. J., Necki, J. M., Iverach, C. P., Ginty, E., Zimnoch, M., Jasek, A., and Nisbet, E. G.: Carbon isotopic signature of coal-derived methane emissions to the atmosphere: from coalification to alteration, *Atmos. Chem. Phys.*, 16, 13669–13680, <https://doi.org/10.5194/acp-16-13669-2016>, 2016.
- Zupanski, D., Hou, A. Y., Zhang, S. Q., Zupanski, M., Kummerow, C. D., and Cheung, S. H.: Applications of information theory in ensemble data assimilation, *Q. J. Roy. Meteorol. Soc.*, 133, 1533–1545, <https://doi.org/10.1002/qj.123>, 2007.