Manuscript prepared for Biogeosciences
with version 2014/09/16 7.15 Copernicus papers of the LaTeX class copernicus.cls.
Date: 30 April 2015

Natural and anthropogenic methane fluxes in Eurasia: a meso-scale quantification by generalized atmospheric inversion

Berchet A.[1,*], Pison I.[1], Chevallier F.[1], Paris J.-D.[1], Bousquet P.[1], Bonne J.-L.[1], Arshinov M. Yu.[2], Belan B. D.[2], Cressot C.[1], Davydov D. K.[2], Dlugokencky E. J.[3], Fofonov A. V.[2], Galanin A.[4], Lavrič J.[5], Machida T.[6], Parker R.[7], Sasakawa M.[6], Spahni R.[8], Stocker B. D.[8] and Winderlich J.[9]

[1]Laboratoire des Sciences du Climat et de l'Environnement, CEA-CNRS-UVSQ, IPSL, Gif-sur-Yvette, France.
[2]V. E. Zuev Institute of Atmospheric Optics, SB-RAS, Tomsk, Russia.
[3]NOAA Earth System Research Laboratory, Boulder, Colorado, USA.
[4]P. E. Melnikov Permafrost Institute SB-RAS, Yakutsk, Russia.
[5]Max Planck Institute for Biogeochemistry, Jena, Germany.
[6]Center for Global Environmental Research, National Institute for Environmental Studies, Tsukuba, Japan.
[7]University of Leicester, Leicester, United Kingdom.
[8]Climate and Environmental Physics, Physics Institute, and Oeschger Centre for Climate Change Research, University of Bern, Bern, Switzerland.
[9]Max Planck Institute for Chemistry, Mainz, Germany.
[*]Now at Laboratory for Air Pollution/Environmental Technology, Swiss Federal Laboratories for Materials Science and Technology, Empa, Dübendorf, Switzerland.

Correspondence to: A. Berchet (antoine.berchet@empa.ch)

Abstract. Eight surface observation sites providing quasi-continuous measurements of atmospheric methane mixing ratios have been operated since the mid-2000's in Siberia. For the first time in a single work, we assimilate all of these in-situ data in an atmospheric inversion. Our objective is to quantify methane surface fluxes from anthropogenic and wetland sources at the meso-scale in the Siberian Lowlands for the year 2010. To do so, we first inquire into the way the inversion uses the observations and the way the fluxes are constrained by the observation sites. As atmospheric inversions at the meso-scale suffer from mis-quantified sources of uncertainties, we follow recent innovations in inversion techniques and use a new inversion approach which quantifies the uncertainties more objectively than the previous inversions. We find that, due to errors in the representation of the atmospheric transport and redundant pieces of information, only one observation every few days is found valuable by the inversion. The remaining high-resolution quasi-continuous signal is representative of very local emission patterns. An analysis of the use of information by the inversion also reveals that the observation sites constrain methane emissions within a radius of 500 km. More observation sites than the ones currently in operation are then necessary to constrain the whole Siberian Lowlands. Still, the fluxes within the constrained areas are quantified with objectified uncertainties. At the end, the tolerance intervals for posterior methane fluxes are of roughly 20% (resp. 50%) of the fluxes for anthropogenic (resp. wetland) sources. About 50–70% of emissions are constrained by the

inversion on average on an annual basis. Extrapolating the figures on the constrained areas to the whole Siberian Lowlands, we find a regional methane budget of 5–28 $TgCH_4$ for the year 2010, i.e. 1–5% of the global methane emissions. As very few in-situ observations are available in the region of interest, observations of methane total columns from the Greenhouse Gas Observing SATellite (GOSAT) are used for the evaluation of the inversion results, but they exhibit marginal signal from the fluxes within the region of interest.

1 Introduction

Methane (CH_4) in the atmosphere contributes to climate forcing as a greenhouse gas and is involved in the atmospheric oxidizing capacity (Forster et al., 2007). Characterizing the variability of the atmospheric CH_4 composition requires accurate understanding of the methane biogeochemical cycle, in particular of the surface-atmosphere fluxes, of their spatial distribution and of their temporal variability. The quantification of these contributions to the methane cycle still experiences high uncertainties (Kirschke et al., 2013). The global surface to atmosphere CH_4 fluxes range between 500 and 600 $TgCH_4.y^{-1}$ (1 Tg = 10^{12} g). Two of the main contributors to the global CH_4 budget are natural emissions from inundated areas and anthropogenic sources from coal, oil and gas extraction and distribution. Inundated areas are responsible for 145–260 $TgCH_4.y^{-1}$, i.e. 25–50% of total emissions with a very high heterogeneous spatial distribution and year-to-year variability (e.g., Bousquet et al., 2006; Dlugokencky et al., 2009; Bergamaschi et al., 2009). The anthropogenic sources from fossil fuel burning and leaks account for 100-150 $TgCH_4.y^{-1}$, i.e. 20–30% of total emissions, according to the EDGAR inventory (depending on the year and the version of the inventory; http://edgar.jrc.ec.europa.eu).

The West Siberian plain concentrates significant sources of CH_4 of both wetland and anthropogenic types (Lechtenböhmer et al., 2005; Spahni et al., 2011). On one side, with 50–70% of its area covered by peatlands (Peregon et al., 2009), about 13% of global wetlands are located in the West Siberian plain. On the other side, Russia produces 20% of the natural gas in the world, mostly extracted in Siberia; and 0.1 to 10% (i.e. 0.5–40 $TgCH_4.y^{-1}$) of this gas are estimated to leak into the atmosphere (e.g., Hayhoe et al., 2002). Large amounts of methane are also released during the oil welling, of which Russia is also a major producer ($\sim$10–15% of the global production), with 1-2% of the oil production leaked into the atmosphere as methane (e.g., Agency, 2011). Documenting the emissions of methane in the West Siberian plain is thus critical to reduce the uncertainties on the global methane budget.

However, accurately quantifying the wetland and anthropogenic emissions in the West Siberian plain is challenging. On the one hand, wetland emissions at high latitudes like in Siberia exhibit a clear year-to-year variability (Bergamaschi et al., 2013) and a distinct seasonality (e.g., Pickett-Heaps et al., 2011) due to high sensitivity to the soil temperature and humidity, to the water table

depth and to the total inundated surface which can vary up to 25% from year to year (Ringeval et al., 2010). This high sensitivity to the local meteorological parameters could cause still unobserved drastic increases of CH_4 emissions from boreal wetlands with climate change (Bohn et al., 2007). On the other hand, anthropogenic sources are mainly related to uncontrolled leakage which is difficult to estimate. Quantifying these leaks raises many issues: pipelines with tiny leaks from chemical permeability span over thousands of kilometres, single leaks range on a spectrum of several orders of magnitude (from the lower with pipeline permeability leaks, to the higher during the welling), and leaks can appear (and disappear when detected and repaired) very quickly.

Despite the importance of quantifying and understanding the contribution of the West Siberian plain to the global CH_4 budget, few studies have been dedicated to this region. Glagolev et al. (2011) carried out extensive field measurements of local wetland CH_4 emissions in order to characterize the emission patterns of each different environment. They upscaled their results to the whole region using wetland distribution maps. However, considering the discrepancies on the wetland distribution (Peregon et al., 2009; Frey and Smith, 2007) and their extension variability (Ringeval et al., 2010), numerous hard-to-quantify uncertainty sources are expected from this approach. For the quantification of the anthropogenic sources, for instance, Dedikov et al. (1999) measured mixing ratios of methane close to gas lines and gas facilities to deduce emission factors. They upscaled their figures to the Russian territory and got an emission factor for CH_4 lower than 1% of the total production of natural gas. Reshetnikov et al. (2000) reviewed the existing literature about Siberia and found emission factors from 0.4 to 14% of the total production. Another approach is the analysis at different places and dates of the local variations of the atmospheric composition (mixing ratios and isotopic fraction) in CH_4 (and related species). The variations of the atmospheric composition provides information on the relative contribution of the different local processes in the Siberian budget. Such analysis has been carried out using observations from mobile platforms, such as aircraft profiles (Yamada et al., 2005; Umezawa et al., 2012) or train and ship measurements (Tarasova et al., 2006), and fixed stations (Sasakawa et al., 2012). Nevertheless, these studies are not systematic and comprehensive: they give local instantaneous information and only knowledge about the relative contributions of particular processes to the total regional fluxes. High spatial and temporal resolutions, absolute and consistent quantification and separation of the different main sources is essential to better characterizing and estimating the contribution of Siberian emissions to the global methane budget.

A first step towards this goal was done by Winderlich (2012). In this work, a systematic analysis of the variability of the atmospheric composition at the ZOTTO tall tower (described in Winderlich et al., 2010) and at a small set of auxiliary sites was carried out. Their approach relied on atmospheric inversion techniques based on the data assimilation Bayesian theory (Enting et al., 2012; Tarantola, 1987) and in principle allows an objective use of the information (here the variability of the atmospheric composition). This provided insights on, e.g., the likely under-estimation of the local

anthropogenic sources from gas leak in some inventories (EDGAR v4.1 in Winderlich, 2012). However, regional atmospheric inversions with only a few sites give only local constraints on the surface fluxes. More critically, a regional inversion with a small number of observations experiences difficulties in identifying and separating the different contributions to CH_4 emissions. This mis-separation can be related to mis-quantified sources of errors in the atmospheric inversion in addition to the lack of information. This issue is dominant in the West-Siberian plain because of the co-located wetland and anthropogenic emissions. The objective of quantifying and separating the regional Siberian sources requires an inversion based on numerous observation sites, with a comprehensive approach of quantification of the uncertainties.

Recent literature highlights the need for precisely and objectively quantifying all errors in the inversion (transport, representation, flux distribution, etc.) in order to evaluate their impact on the inversion results (Lauvaux et al., 2009; Winiarek et al., 2012; Berchet et al., 2013b; Ganesan et al., 2014). In Berchet et al. (2014), we proposed a general method in order to objectively quantify most of the critical sources of errors in the inversion. This improved algorithm is based on a Monte Carlo approach superimposed to maximum likelihood estimators (Chapnik et al., 2004; Michalak and Kitanidis, 2005).

For the first time in Eurasia, we use this improved algorithm on a network of 8 surface sites (Sasakawa et al., 2010; Winderlich et al., 2010) covering a large part of the Siberian lowlands and of 5 remote sites that constrain the air masses coming into the domain and getting out of it. These sites, which have been operated since the mid-2000's, are implemented into the inversion system with objectified uncertainty quantification from Berchet et al. (2014). Here, our goal is to deduce an accurate quantification of the fluxes at the meso-scale with a temporal resolution of a few days from the variability of the atmospheric CH_4 composition at the 8+5 observation sites.

We explain the theoretical background used in our study in Sect. 2. In Sect. 3, the datasets and models used in the inversion are introduced. We then present the results on the fluxes and the limitations of the inversion in Sect. 4 and Sect. 5. Our inversion is then evaluated in Sect. 6 by using in-situ measurements and independent satellite observations, as the very few available surface observations do not allow to keep enough for evaluation.

2 Marginalized inversion framework

2.1 Motivations towards marginalizing

As the atmosphere mixes irreversibly air masses from different CH_4 sources, using the atmospheric signal as in an atmospheric inversion cannot lead to a deterministic characterization of the surface fluxes. In the classical Bayesian framework (Tarantola, 1987), the objective of the inversion is to inquire into the probability density function (pdf) of the surface fluxes, or more generally of the

state of the system, with some knowledge about the atmospheric composition and on the prior state distribution. The sought *pdf* can be written: $p(\mathbf{x}|\mathbf{y}^{o} - H(\mathbf{x}^{b}), \mathbf{x}^{b})$.

In this formula, the vector $\mathbf{y}^{o}$ gathers all the available observations; H is the observation operator converting the information in the state vector to the observation space; the vector $\mathbf{x}$ depicts the state of the system (mainly CH_4 surface fluxes in our case) and $\mathbf{x}^{b}$ is the background vector including the prior knowledge on the state $\mathbf{x}$ of the system. H typically represents the discretization of the problem to be computed numerically and the atmospheric transport from the emission areas to the observation sites. In the following, we consider H linear and associate it to its Jacobian matrix $\mathbf{H}$. In our case, we simulate the atmospheric transport only on a domain of limited area as detailed in Sect. 3.3 and 3.4. Additional information about the atmospheric composition at the boundaries of the domain of interest is then necessary to compare observed with simulated atmospheric composition. Therefore, the state vector $\mathbf{x}$ encompasses the surface fluxes, but also the lateral boundary conditions related to the observed baselines at each observation site.

Thus, the inversion computes the pieces of information contained within the observations, the prior state and the representation of the transport with their associated uncertainties (e.g., measurement errors, uncertainties in the flux inventories, etc.). With the usual Gaussian assumption, all the uncertainties are considered as normal *pdf*s and can be described with modes and uncertainty covariance matrices. The inversion then deduces an optimal posterior state vector $\widehat{\mathbf{x}^{a}}$ and posterior uncertainties $\mathbf{P}^{a}$. Within the Gaussian assumption, the posterior state vector and uncertainty matrix can thus be explicitly defined:

$$\begin{cases} \widehat{\mathbf{x}^{a}} &= \mathbf{x}^{b} + \mathbf{K}(\mathbf{y}^{o} - \mathbf{H}\mathbf{x}^{b}) \\ \mathbf{P}^{a} &= \mathbf{B} - \mathbf{K}\mathbf{H}\mathbf{B} \end{cases} \tag{1}$$

In Eq. 1, the matrix $\mathbf{K} = \mathbf{B}\mathbf{H}^{T}(\mathbf{R} + \mathbf{H}\mathbf{B}\mathbf{H}^{T})^{-1}$ is the Kalman gain matrix. The matrices $\mathbf{R}$ and $\mathbf{B}$ are the covariance matrices describing the observation and background uncertainties. Observation uncertainties encompass measurement, discretization and transport errors. Background uncertainties include the uncertainties in the spatial distribution of the fluxes, in their temporal variability and in their absolute value.

As long as these uncertainty matrices are known, the inversion only faces technical issues (e.g., matrix inverses and products in large dimension problems) for numerical implementation. However, only the uncertainties in the measurements are objectively quantified during the calibration process. The errors in the transport or in the prior fluxes are not perfectly known and, in most case, they are built relying on some expert knowledge about the system. But this subjective knowledge can lead to ill-specified matrices, which have a dramatic impact on the inversion results (e.g., Cressot et al., 2014). Recent studies inquired into objectified ways of specifying these matrices (e.g., Michalak and Kitanidis, 2005; Winiarek et al., 2012; Berchet et al., 2013b). The approach in these papers was to find optimal uncertainty matrices $\mathbf{R}$ and $\mathbf{B}$ along an objective statistical criterion: the maxi-

mum likelihood. The implementation of the method gave encouraging results, but the impact of the uncertainties within the maximum likelihood computation were complicated to evaluate.

Berchet et al. (2014) used a general marginalization approach in order to quantify objectively all the uncertainties impacting on the posterior fluxes. In the following, we summarize this approach.

2.2 Method outline

The marginalization consists in computing the complete *pdf* $p(\mathbf{x}|\mathbf{y}^{\mathrm{o}}-\mathbf{H}\mathbf{x}^{\mathrm{b}},\mathbf{x}^{\mathrm{b}})$ as a weighted sum of the Gaussian *pdf*s $p(\mathbf{x}|\mathbf{y}^{\mathrm{o}}-\mathbf{H}\mathbf{x}^{\mathrm{b}},\mathbf{x}^{\mathrm{b}},\mathbf{R},\mathbf{B})$ calculated for each possible uncertainty matrices $\mathbf{R}$ and $\mathbf{B}$. This can be written as follows:

$$\begin{aligned} & p(\mathbf{x}|\mathbf{y}^{\mathrm{o}}-\mathbf{H}\mathbf{x}^{\mathrm{b}},\mathbf{x}^{\mathrm{b}}) \\ & = \int_{(\mathbf{R},\mathbf{B})} p(\mathbf{x}|\mathbf{y}^{\mathrm{o}}-\mathbf{H}\mathbf{x}^{\mathrm{b}},\mathbf{x}^{\mathrm{b}},\mathbf{R},\mathbf{B}) \\ & \quad \times p(\mathbf{R},\mathbf{B}|\mathbf{y}^{\mathrm{o}}-\mathbf{H}\mathbf{x}^{\mathrm{b}},\mathbf{x}^{\mathrm{b}}) \quad \mathrm{d}(\mathbf{R},\mathbf{B}) \end{aligned} \tag{2}$$

To compute the marginalized integral in Eq. 2, a large number (60 000 in our case) of posterior vectors $\widehat{\mathbf{x}^{\mathrm{a}}}$ and posterior uncertainty matrices $\widehat{\mathbf{P}^{\mathrm{a}}}$ is computed through individual inversions as in Eq. 1 with different uncertainty matrices $\widehat{\mathbf{R}}$ and $\widehat{\mathbf{B}}$. This Monte Carlo sampling is carried out based on an estimate of the *pdf*s of the uncertainty matrices, $p(\mathbf{R})$ and $p(\mathbf{B})$. This estimate is deduced from the objectified maximum likelihood approach.

Posterior uncertainties and correlations are in the end computed from the ensemble of 60 000 computed individual inversion results as follows:

$$\left\{ \begin{array}{lcl} \mathbf{x}^{\mathrm{a}} & = & \frac{1}{N}\sum_{i=1..N}\widehat{\mathbf{x}^{\mathrm{a}}}^{i} \\ \mathbf{P}^{\mathrm{a}} & = & \frac{1}{N-1}\sum_{i=1..N}\left(\widehat{\mathbf{x}^{\mathrm{a}}}^{i}-\mathbf{x}^{\mathrm{a}}\right)\left(\widehat{\mathbf{x}^{\mathrm{a}}}^{i}-\mathbf{x}^{\mathrm{a}}\right)^{\mathrm{T}} \end{array} \right. \tag{3}$$

with N the number of Monte Carlo draws.

As detailed in Sect. 2.1, the main motivation for marginalizing the classical inversion framework is to use the available information (in-situ observations, flux inventory, transport model) in a way which is as objective as possible. The marginalized inversion gives an explicit and objectified access to pieces of information (described in details in Sect. 2.3) required:

1. to evaluate the efficiency of the observation network for constraining regional emissions and to give guidelines for improving monitoring deployment;

2. to inquire into atmospheric inversion skills in terms of resolved temporal and spatial resolution and of emission process separation, and to deduce observation and modeling requirements for future better inversions; and

3. to assess the robustness of emission inventories and process-based surface-atmosphere exchange models

These three key points are further discussed in Sect. 4.

One of the main drawbacks in our method we use is its numerical cost. Computing 60 000 individual inversions is cumbersome and requires extensive amounts of memory. Smartly chosen filtering criteria, observation sampling and flux aggregation patterns must be carried out before the marginalized inversion in order to reduce its complexity without degrading the inversion optimality. As recent efforts have been made to reduce the subjectivity in the inversion, we rely on objective criteria (though not computed in an exact manner due to computational limitations) to complete the observation sampling, flux aggregation and filtering. These criteria are shortly developed in Sect. 2.4 and detailed in Berchet et al. (2014).

2.3 Output analysis

As described above, the marginalized regional inversion as we compute it answers three main questions. Below are given details on how these questions are treated.

2.3.1 Network efficiency

The evaluation of the network coverage is carried out through the explicit computation at the maximum likelihood of the influence $\mathbf{K}_{\text{max}}\mathbf{H}$ and sensitivity matrices $\mathbf{H}\mathbf{K}_{\text{max}}$ following Cardinali et al. (2004). With these two matrices, we can compute the weight ω of each observation site in the inversion and the unconstrained emissions ϵ as follows:

$$\omega_i = \sum_{j \text{ at } i} \left(\mathbf{H}\mathbf{K}_{\text{max}}\right)_{j,j} \text{ for each site } i \tag{4a}$$

$$\epsilon_i = \left(1 - \mathbf{K}_{\text{max}}\mathbf{H}\right) \times \phi_i \text{ for each emission pixel } i \tag{4b}$$

with ϕ_i the emissions at the pixel i.

In Eq. 4a, the higher the score ω_i the more the inversion uses the site i. Observation sites downwind emissions will have a strong impact on the inversion, but observations constraining air masses coming into the domain of interest are also key stones for regional inversions. Eq. 4b depicts emitted CH_4 that is not seen or constrained by the inversion. Regions with the highest ϵ_i are unseen areas with strong emissions (as considered by inventories), where additional observation sites would be required.

2.3.2 Solved spatial and temporal scales

The spatial and temporal resolutions that the inversion can solve are described by posterior error covariances. Mis-separated regions are usually detected through so-called flux dipoles (e.g., Rödenbeck et al., 2003). In our case, as we explicitly and objectively compute the posterior matrix $\mathbf{P}^{\text{a}}$ from the Monte Carlo ensemble $\left(\widehat{\mathbf{x}^{\text{a}}}\right)$, posterior correlations are used for post-processing groups of

ill-separated regions (these regions are defined through a dynamically-chosen aggregation pattern prior to the inversion; see Sect. 2.4.2). Both strongly positive and negative correlations $r_{i,j}$ point to ill-separated regions i and j. Following Observing System Simulation Experiments (OSSEs) carried out in Berchet et al. (2014), we group two posterior flux regions i and j when:

$$|r_{i,j}| = |\mathbf{P}^{\mathrm{a}}_{i,j}| > 0.5 \tag{5}$$

For instance, some neighbouring emissions or successive ones could be mixed in the atmospheric signal and then be not separated from the inversion point of view. This post-processing makes it possible to group posterior fluxes in order to avoid dipoles and reduce uncertainties on bigger regions. Thus, it is also possible to filter out regional fluxes that are not separated from the boundary conditions. Every group of correlated regions including some contribution from the boundary conditions are excluded from the analysis of inversion results. Indeed, errors on the lateral boundary conditions can mis-lead the inversion about the regional fluxes.

Through this post-grouping, we can assess the typical size of aggregated regions that the inversion can constrain with our set-up (observations and transport model resolution). The typical time-scale that the inversion can detect for a change in emission can also be assessed with this post-processing. As we want to separate contributions from anthropogenic and wetland emissions, the post-grouping may group or not co-located emissions from different processes and then give insights into the separation ability of our regional inversion.

2.3.3 Posterior flux analysis

Ideally, an atmospheric inversion provides insights about emissions. From the correlation grouping applied to the Monte Carlo ensemble of posterior fluxes, we can compute tolerance intervals of posterior fluxes so that 68.27% of the ensemble is within the interval. The number 68.27% makes the tolerance intervals equivalent to the $\pm\sigma$ interval in the Gaussian framework, as the Monte Carlo posterior ensemble does not necessarily a Gaussian distribution. The inversion thus indicates that the fluxes we are inquiring into are very likely in the defined posterior tolerance interval. Deviating posterior tolerance intervals compared to prior fluxes point to required updates in the used prior database. Below we present our visualization approach to control posterior fluxes.

Figure 5 synthesizes the inverted methane fluxes for Siberian lowlands (hatched domain in Fig. 1). As detailed in Sect. 2.3.2, anthropogenic emissions (inverted at a monthly scale) can be grouped with wetland emissions (considered at the weekly scale). So, the lowest common multiple on which the fluxes can be analysed is the monthly scale. Then, for the Siberian Lowlands, for each month, we define the proportion of the fluxes that are within regions constrained through the inversion (written on the left of each pie chart in Fig. 5). The proportions of anthropogenic and wetland emissions that are constrained, mixed with another type of emission or unseen by the observation network are represented in the pie charts in Fig. 5. Finally, within the proportion of constrained regions, we

analyse the inversion correction on anthropogenic, wetland and mixed emissions (bar plots in Fig. 5). For each type of emission (anthropogenic, wetland and mixed), we present in the bar diagrams of Fig. 5 the tolerance intervals of posterior fluxes and the prior uncertainties as calculated by the maximum likelihood algorithm (that is to say $\mathbf{B}_{\text{max}}$; see Sect. 2.2). For each month, the tolerance intervals on the total prior and posterior methane budget in the Siberian Lowlands is also given in $TgCH_4$.

2.4 Size reduction and filters

As suggested in Sect. 2.2, the marginalized inversion requires some filtering, observation sampling and flux aggregation, so it can be numerically computed. Below, we explain how we carry out these pre-processing in a way that do not dampen the advantages of the marginalization.

2.4.1 Observation sampling

Here we try to reduce the dimension of the observation space. At the regional scale, considering the spatial resolution of our transport model, only the synoptic variability of the observed signal is relevant. We then decide to keep only one piece of information per site and per day as it is commonly done at the global and continental scales. In addition, simulated vertical mixing close to the surface where observations are carried out is known to be flawed when the planetary boundary layer (PBL) is shallow (typically at night and in winter in Siberia). We then sample the observations during the afternoon when the PBL is higher than 500 m as suggested by prior studies (e.g., Berchet et al., 2013b) and we pick the observed and simulated mixing ratios at the time when the observations are minimum.

As surface emissions dominate on surface sinks for CH_4, keeping the minimum observed mixing ratio by afternoon is equivalent to detecting the time when the PBL is at its maximum, hence when the atmospheric model is the more accurate. This criterion filters out outliers generated by local influences which cannot be reproduced by an atmospheric transport model with a resolution larger than 25 km, and which only add noise to the system.

For our case study, out of 127000 hourly measurements available in 2010, 30000 pass through the PBL height and night filters (see black dots in Fig. 2). Out of these 30000 data points, about 2000 daily aggregates are selected. Details by observation site are given in Tab. 2.

2.4.2 Flux aggregation and constraints

The following procedures are bound to decrease the state space dimension. To define aggregation pattern, we divide our domain into 35 physical regions for each emission process (according to vegetation types, demography, industrial activity, etc.) as a basic pattern. This basic pattern is chosen so that the mesh gets finer closer to the observation network. The resolution of the transport representation is also chosen finer close to the observation network. To further reduce the number of

aggregated regions, under-constrained regions are grouped together. This is carried out by analysing the observation network footprints estimated with a Lagrangian model (Sect. 3.2) which offers an efficient way to compute them.

In addition to the footprint aggregation, the influence matrix $\mathbf{K}_{\text{max}}\mathbf{H}$ makes it possible to quantify observational constraints on the fluxes. Below a given threshold of constraint for a flux (related to the flux contribution to the atmospheric signal), the inversion cannot deduce any valid information on the flux. For this reason, we also filter out a region i with very low constraints from the marginalization if:

$$(\mathbf{K}_{\text{max}}\mathbf{H})_i^{\text{a}} < 0.5 \tag{6}$$

2.4.3 Plume filtering

It is known that atmospheric transport models suffer from temporal and spatial mismatches when simulating air masses with strong mixing ratio gradients (e.g., for plumes well-delimited from the background air masses). When a plume is transported to the wrong place and time, as we sample air masses at a given point location, the differences between simulated and observed mixing ratios can reach unrealistic values. Such strong model-observation differences have a significant impact on inversion corrections on fluxes.

To dampen such undesirable effects, we introduce a procedure to filter out plume-like air masses from the inversion input. As explained in Sect. 2.2, a maximum likelihood estimation is computed prior to the marginalization. We take advantage of the maximum likelihood computation to detect air masses critically ill-reproduced by the transport model. As the maximum likelihood estimation computes optimal uncertainty matrices $\mathbf{R}_{\text{max}}$ and $\mathbf{B}_{\text{max}}$, we filter out observations with a too high computed uncertainty. That is to say, for each observation i, the data point is excluded if:

$$(\mathbf{R})_{\text{max}}^{i,i} > 20\,\text{ppb} \tag{7}$$

This criterion does not necessarily exclude only plumes generating a big signal, but also the ones that are very poorly reproduced by the transport model.

One should notice that this criterion is computed in association with the low constraint criterion of Eq. 5. That is to say, a region which always illuminates the observation network through plume-like air masses will have all its constraining observations filtered out. As a consequence, it will be considered poorly constrained and then be itself eliminated from the inversion.

Taken all together, the criteria defined above filter out a large part of the available observations and regional fluxes. In particular, hot spot emissions are largely removed from the inversion (except for some exception when the observations are far enough, so that the transport manages to well reproduce the plume generated by the emissions).

3 Set up for an Eurasian domain

We are interested in the Eurasian surface-atmosphere budget of methane. As developed in Sect. 2, the computation of the marginalized inversion needs measurements ($\mathbf{y}^o$; Sect. 3.1), prior fluxes ($\mathbf{x}^b$; Sect. 3.3) and an observation operator computed with an atmospheric transport model ($\mathbf{H}$; Sect. 3.4). The footprints of the observation network required for choosing the aggregation patterns (necessary to the computation of $\mathbf{H}$) are estimated with a Lagrangian model (Sect. 3.2). The observations are compared to the fluxes through the observation operator $\mathbf{H}$. The operator $\mathbf{H}$ is computed with a transport model, which simulates mixing ratios, whereas the observation sites provide dry air mole fractions. For trace gases such as CH_4, dry air mole fractions and mixing ratios are equivalent. In all of the following, we affiliate dry air mole fractions to mixing ratios.

3.1 The observation network: $\mathbf{y}^o$

The inversion is based on measurements of atmospheric CH_4 dry air mole fractions at 14 Eurasian surface sites. The Eurasian sites are situated all over the Eurasian continent, from South Korea to Scandinavia (see Fig. 1). They are maintained by the National Institute for Environmental Studies (NIES, Tsukuba, Japan), the Institute for Atmospheric Optics (IAO, Tomsk, Russian Federation), the Max Planck Institute (MPI, Iena, Germany), the P. E. Melnikov Permafrost Institute (Yakutsk, Russian Federation), the Finnish Meteorological Institute (FMI, Helsinki, Finland) and NOAA Earth System Research Laboratories (Boulder, United States of America). Most stations provide quasi-continuous measurements of methane atmospheric composition. A few stations only collect flasks at a weekly scale (see Tab. 1). The location, measurement type, maintaining institute and 3-letter symbol of each observation site are described in Tab. 1.

Flasks samples are collected at SDZ, TAP and UUM sites as partners of the NOAA global cooperative air sampling network (Dlugokencky et al., 1994). These flask sampling sites are designed to monitor large-scale and global variability. They are then placed at locations, which are most of the time not illuminated by local and regional sources. Thus, the observations from these sites generally pass the filtering criteria of Sect. 2.4. Flasks samples are sent to and analysed at NOAA-ESRL by gas chromatography with flame ionization detection against the NOAA 2004 CH_4 standard scale (Dlugokencky et al., 2005).

NIES sites (AZV, BRZ, DEM, IGR, KRS, NOY, VGN and YAK) are part of the Japan–Russia Siberian Tall Tower Inland Observation Network (JR-STATION; Sasakawa et al., 2010) and are equipped with CH_4 semiconductor sensors based on a tin dioxide natural gas leak detector developed by Suto and Inoue (2010). The instruments are calibrated on tanks traceable to NIES 94 CH_4 scale. The NIES 94 CH_4 scale is higher than the NOAA 04 scale by 3.5–4.6 ppb in the range between 1750 and 1840 ppb (Zhou et al., 2009). ZOT site is operated by MPI since April 2009 (Winderlich et al., 2010) and air is analysed by an EnviroSense 3000i analyser (Picarro Inc., USA, CFADS-17) based

on the cavity ring-down spectroscopy technique (CRDS Crosson, 2008). The calibration system uses tanks traceable to NOAA 04 scale. All the mixing ratios are reported to the NOAA 04 scale before being implemented into the inversion system.

JR-STATION and ZOT sites are located in the vicinity of anthropogenic and wetland sources. These local sources strongly influence the Siberian network, which was designed to monitor regional emissions. As a consequence, numerous observations from the Siberian network are ill-reproduced by our transport model and then are filtered out from the inversion as "plume" observations (following criteria in Sect. 2.4; see also Fig. 2 and Tab. 2 in Sect. 6.1). In particular, measured mixing ratios at BRZ site are not well simulated by our model, possibly due to missing local emissions in the prior. Most BRZ observations are thus filtered out from the inversion.

Due to logistical issues or instrument dysfunctions, observation sites do not provide measurements all year round. Fig. 2, described in Sect. 4.1, details the temporal availability of the observations. The sampling bias is known to impact the inversion results (Villani et al., 2010). The issue is discussed in Sect. 4.1, but the general method developed in Sect. 2 consistently takes into account such a bias into increased posterior tolerance intervals and constraints on the emissions.

The observation vector $\mathbf{y}^{o}$ is defined after Sect. 2.4.1 sampling method. The final size of $\mathbf{y}^{o}$ implemented in the inversion is 2000. On average, 0.4 observations per station per day are validated for the inversion.

3.2 Estimates of the network footprints

As the observations that will be implemented in the system are known, the observation network footprints, necessary to choose the aggregation patterns in order to define the prior state vector $\mathbf{x}^{b}$ and the observation operator $\mathbf{H}$ (as detailed in Sect. 2.4.2), can be computed. As we do not carry out a quantitative analysis of the footprints, we only need a rough estimation of the footprint patterns. Thus, we compute simulations with the Lagrangian dispersion model FLEXPART version 8.2.3 (Stohl et al., 2005) to get such an estimation. To build the footprints, we compute numerous back-trajectories of virtual particles from the observation sites at the times when measurements are available and valid for the inversion.

The model is forced by ECMWF ERA-Interim data at an horizontal resolution of $1^{\circ} \times 1^{\circ}$, with 60 vertical levels and 3 hours temporal resolution (Uppala et al., 2005). Virtual particles are released in a 3D box (10 km per side and 1000 m high) centered around each observation site with 10-day lifetime backwards in time. The footprints are computed on a $0.5^{\circ} \times 0.5^{\circ}$ horizontal grid, following the method of Lin et al. (2003), taking into account the boundary layer height at each particle location. This method considers that only the particles within the boundary layer are influenced by surface emissions and that the boundary layer is well-enough mixed to be considered as uniform.

Table 1: Sites characteristics. The altitudes of the sites are given as m above sea level (a.s.l.) and the inlet height is in m above ground level (a.g.l.). The frequency column depict the type of instrument in the site: C=quasi-continuous, F=flask sampling.

Station	ID	Location			Inlet	Frequency	Network
		Lon	Lat	Alt	height		/ Institute
		(° E)	(° N)	(m a.s.l)	(m a.g.l.)		
Azovo	AZV	73.03	54.71	100	50	C	JR-STATION
Berezorechka	BRZ	84.33	56.15	150	80	C	JR-STATION
Demyanskoe	DEM	70.87	59.79	75	63	C	JR-STATION
Igrim	IGR	64.42	63.19	25	47	C	JR-STATION
Karasevoe	KRS	82.42	58.25	50	67	C	JR-STATION
Noyabrsk	NOY	75.78	63.43	100	43	C	JR-STATION
Pallas	PAL	24.12	67.97	560	7	C	FMI
Shangdianzi	SDZ	117.12	40.65	287	0	F	NOAA/ESRL
Tae-ahn Peninsula	TAP	126.12	36.72	20	0	F	NOAA/ESRL
Ulaan Uul	UUM	11.08	44.45	914	0	F	NOAA/ESRL
Vaganovo	VGN	62.32	54.50	200	85	C	JR-STATION
Yakutsk	YAK	129.36	62.09	210	77	C	JR-STATION
Zotino	ZOT	89.35	60.80	104	301	C	MPI

3.3 Prior fluxes and state vector: $\mathbf{x}$

The inversion system optimizes prior fluxes grouped in the regions aggregated by the pre-processing procedure (see Sect. 2.2). The prior spatial distribution and temporal variability of the fluxes are deduced from: 1) EDGAR database v4.2 FT2010 (http://edgar.jrc.ec.europa.eu) for year 2010 for anthropogenic emissions, 2) LPX-Bern v1.2 process model (Stocker et al., 2014) at a monthly scale for wetland emissions 3) GFED v4 database at a daily scale for wildfires. In Fig. 1, the distributions of the anthropogenic hot spots of emissions and of the wetland regions are represented, superimposed over the regional topography. Anthropogenic emissions of methane in the region are mainly hot spots related to the intense oil and gas industry in the Siberian Lowlands and to the leaks in the distribution system in population centres in the vicinity of the Trans-Siberian Railway in the Southern part of the Siberian plain. Wetland emissions are mainly confined to the lower part of Siberia in the West Siberian plain, half of which is lower than 100 meters above sea level. Wildfires occur mainly in spring and summer in the Eurasian forest-covered areas; they emit CH_4 as intense hot spots.

The EDGAR inventory uses economic activity maps by sectors and convolved with emission factors calculated in laboratories or with statistical studies (Olivier et al., 2005). The Bern based

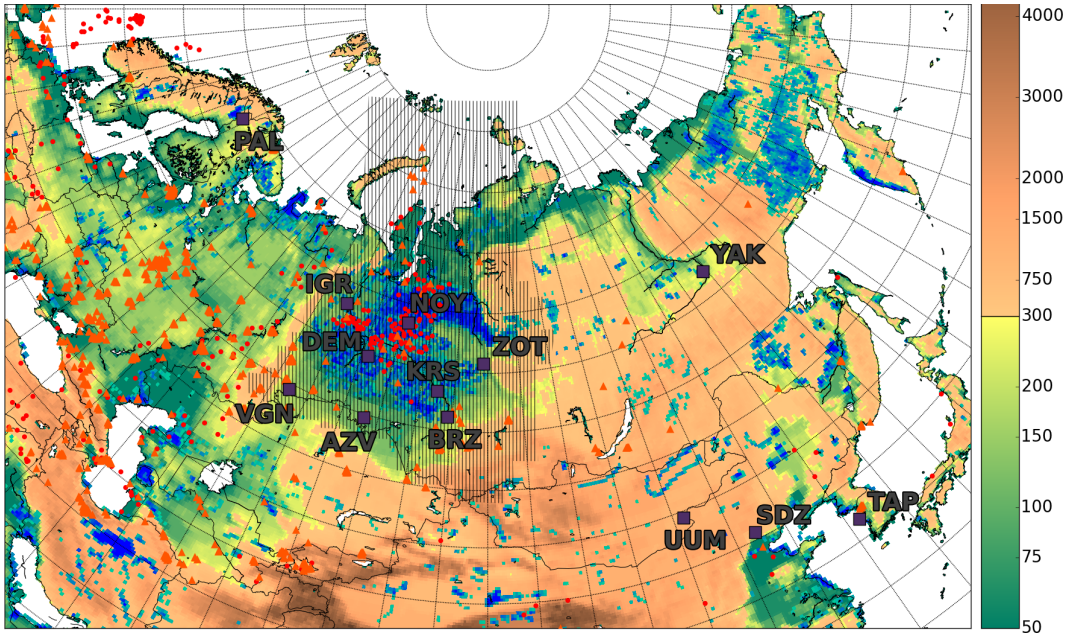

Figure 1: Topographic map of the domain of interest. The colour bar paints the altitude above sea level (from ETOPO1 database; Amante and Eakins, 2009). Red dots (resp. orange triangle) depict hot spots of CH_4 emissions (based on EDGAR v4.2 inventory; see Sect. 3.3) related to oil welling and refineries (resp. gas extraction and leaks during distribution in population centres). Purple squares highlight the observation site localizations. Blueish shaded areas represent inundated regions, wetlands and peatlands (from the Global Lakes and Wetlands Database; Lehner and Döll, 2004). The Siberian budget in Sect. 5.2 is calculated on the hatched area.

Land surface Processes and eXchanges (LPX-Bern v1.2) model is an update of the dynamic global vegetation model LPJ (Spahni et al., 2011). It includes a dynamical simulation of inundated wetland areas (Stocker et al., 2014), dynamic nitrogen cycle (Stocker et al., 2013), and dynamic evolution of peatlands (Spahni et al., 2013; Stocker et al., 2014). The model uses CRU TS 3.21 input data (temperature, precipitation rate, cloud cover, wet days), observed atmospheric CO_2 and prescribed nitrogen deposition (Lamarque et al., 2011) for each year for the simulation of dynamic forest and peatland vegetation growth. The GFED v4 database is built from the 500 m Collection 5.1 MODIS DB burned-area mapping algorithm (Giglio et al., 2009) . CH_4 emissions at monthly and daily scales are deduced from the burnt areas using Carnegie-Ames-Stanford-Approach (CASA model; Potter et al., 1993) and emission factors (van der Werf et al., 2010).

We are aiming at a separation of the types of emissions at the meso-scale. We therefore aggregate the emissions along the three different types of sources, with specific spatial patterns and temporal profiles for each type of emissions. Anthropogenic sources are hot spots emitting all year round. Wetlands are responsible for diffuse fluxes on large areas, with high temporal variability depending on

the local weather conditions (typically temperature or water table depth). The emissions of CH_4 from wildfires comes from point sources and occur on relatively short periods (Kasischke and Bruhwiler, 2002). Consequently, we do not aggregate the different types of emissions along the same spatial patterns and temporal intervals. Anthropogenic emissions are aggregated by month, while wetlands and wildfires, which have quicker time responses to meteorological changes, are grouped by periods of 10 days. In the following, we discuss the inversion results only in term of anthropogenic and wetland emissions. Indeed, as the wildfire emissions generate plumes relatively well-defined from the ambient air, the marginalized inversion exclude from the system all the emission contribution related to fires (according to the procedures described in Sect. 2.2).

For the computation of the observation operator $\mathbf{H}$ (see Sect. 3.4), we use a regional chemistry-transport model with a domain limited in space and time. Initial and lateral boundary conditions (hereafter IC and LBC) are then also to be optimized in the system. Prior lateral mixing ratios are deduced from simulations at the global scale by the general circulation model LMDz with the assimilation of surface observations outside the domain of interest (Bousquet et al., 2011). LBC are assimilated by periods of 10 days. We arbitrarily aggregate LBC along 4 horizontal components (by side of the domain) and 2 vertical ones (1013–600 hPa and 600–300 hPa). Though we are mainly focused on Siberian lowlands, the domain of model computation has been chosen spanning over all Eurasia. This is expected to attenuate the impact of the rough global resolution in LMDz boundary conditions on the simulated variations of mixing ratios at the observation sites. Indeed, the central region is thousands of kilometres away from the sides of the domain.

To summarize, all the pieces of information in the observations are assimilated to constrain 1700 aggregated regions of flux and boundary conditions: 10×12 month regions for anthropogenic emissions, 10×36 10-day period for wildfires, 25×36 10-day period for wetlands, 9 (4 sides $\times$ 2 horizontal levels + roof top) $\times36$ 10-day period for the lateral boundary conditions. After the filtering of Sect. 2.4, the dimension of the state space is reduced from 1700 to 275.

3.4 The observation operator: H

We explicitly define the observation operator $\mathbf{H}$ by computing the forward atmospheric transport from the regions of aggregated emissions (defined in Sect. 3.3) to the observation sites. As CH_4 is a reactive species, the observation operator should include the oxidation by OH radicals. However, as the residence time of the air masses in the domain of simulation is short (a few days to a few weeks) compared to CH_4 atmospheric life time (8–10 years; Dentener et al., 2003), ignoring OH sink only generates small differences in the simulated mixing ratios. Additionally, OH sink is mostly responsible for large scale gradients while the regional inversion focuses on the synoptic scale. Thus, the regional inversion system attributes OH sink to global boundary conditions, and not to regional fluxes.

Thus, for each aggregated region, we calculate the so-called response functions using the transport module of the Eulerian meso-scale non-hydrostatic chemistry transport model CHIMERE (Vautard et al., 2001; Menut et al., 2013). This model was developed in a framework of air quality simulations (Schmidt et al., 2001; Pison et al., 2007), but is also used for greenhouse gas studies (Broquet et al., 2011; Berchet et al., 2013b). We use a quasi-regular horizontal grid zoomed near the observation sites after Sect. 2 considerations. The domain of interest is of limited area and spans over the mainland of the Eurasian continent (see Fig. 3). As we are interested in meso-scale fluxes, we take a spatial resolution larger than 25 km. The average side length of the grid cells is 25 km close to the West-Siberian stations and 150 km away of the centre of the domain. The 3D-domain embraces roughly all the troposphere, from the surface to 300 hPa (~ 9000 m), with 29 vertical layers geometrically spaced. The model time step varies dynamically from 4 to 6 min depending on the maximum wind speed in the domain. The model is an off-line model which needs meteorological fields as forcing. The forcing fields are deduced from interpolated meteorological fields from the European Centre for Medium-range Weather Forecast (ECMWF) with a horizontal resolution of $0.5° \times 0.5°$ every 3 h (Uppala et al., 2005).

3.5 Independent observations for evaluation

Any inversion has to be confronted by independent data in order to evaluate its results. Few in-situ measurements of CH_4 mixing ratios are available in Siberia. We choose to assimilate all surface observations described in Sect. 3.1 for the optimization of CH_4 fluxes. NIES and LSCE carry out aircraft measurements in the region (Paris et al., 2010; Umezawa et al., 2012; Berchet et al., 2013a), but these measurements are still difficult to compare to meso-scale models. In addition, their spatial and temporal coverage is poor for the year 2010 and they are not numerous enough to get significant validation insights.

For year 2010, the only remaining observations with sufficient spatial coverage and temporal availability are the total columns retrieved by the Greenhouse Gas Observing SATellite (GOSAT). In Sect. 6.2, we evaluate the results of the inversion against GOSAT data. The Japanese satellite GOSAT was launched by the Japan Aerospace Exploration Agency (JAXA), NIES and the Japanese Ministry of the Environment (MOE) in January 2009. It has a polar sun-synchronous orbit at 667 km and provides a full coverage of the Earth every 3 days with a swath of 750 km and a ground pixel resolution of 10.5 km at nadir. The TANSO-FTS instrument observes the solar radiation reflected at the surface and the top of the atmosphere in the short wave infrared (SWIR) domain that allows deducing total columns of methane (XCH_4) in cloud-free and sunlight conditions. The typical size of GOSAT pixels is $\sim$10 km. We use version 3.2 of the TANSO-FTS bias-corrected XCH_4 proxy retrievals performed at the University of Leicester (Parker et al., 2011). The XCH_4 retrieval algorithm uses an iterative retrieval scheme based on Bayesian optimal estimation and associated to averaging kernels and a priori profiles. The retrieval accuracy is estimated to be about 0.6% (i.e. $\sim$10ppb).

The retrieval algorithm needs CO_2 mixing ratios as a proxy for the light path. We use the 4-D CO_2 analysis from the surface air-sample inversion by Chevallier et al. (2010) (MACC v10.2). We obtain ∼25000 GOSAT observations in 2010 over the domain of interest.

In order to compare the observations of the total columns to the model, we use the averaging kernels to compute prior and posterior model equivalents. The regional CTM CHIMERE has a top layer at 300 hPa in our set-up. The stratospheric contribution to the total columns is deduced from the global model LMDz used for the initial and lateral boundary conditions (described in Sect. 3.3). The average observed XCH_4 is $\sim$ 1775 ppb over the domain throughout the year. The prior average in LMDz XCH_4 is $\sim$ 1820 ppb. This bias is attributable to the excessive injection of tropospheric air into the stratosphere in our version of LMDz. In Sect. 6.2, the bias on XCH_4 of 45 ppb has been corrected to allow the observation-model comparison.

4 Diagnostics of the marginalized inversion

The marginalized inversion described in Sect. 2 provides us with tolerance intervals of posterior fluxes, posterior correlations of errors and influence indicators. As the marginalized inversion filters out some data and regions, we present and analyse here the overall performance of our inversion, the effects of the data selection in the inversion and the implied limitations.

4.1 Observation weight in the inversion

4.1.1 Temporal monitoring constraints

As explained in Sect. 2, the method developed by Berchet et al. (2014) filters out numerous observations and emission regions. Some observations available in the domain in 2010 are set aside before the inversion because of known flaws in CTMs. But the marginalized inversion also flags out additional observations when they are measured within plumes difficult to inverse. The remaining pieces of data do not have the same weight in the inversion. Contrary to most classical inversion methods which cannot afford the computation of the explicit sensitivity matrix (see Sect. 2.3.1) informing in the weight of individual observations, the marginalized inversion allows us to explicitly analyse the use of the observations in the system. In Fig. 2, we represent the observations filtered out along the PBL height criterion before the inversion (grey points), the ones flagged out during the inversion (black dots) and the relative weight of the remaining used observations (coloured circles).

Many observations cannot be assimilated (black dots), especially in winter, when the very cold conditions (temperatures lower than -20°C in average) related to the Siberian High generate very stable atmospheric conditions. In these conditions, the local emissions, which cannot be well assimilated in our inversion system, significantly influence the observations. In addition to the numerous not-assimilated observations, all daily observations that are not filtered do not necessarily convey the same amount of information: all the blue circles in Fig. 2 depict pieces of data with a negligible

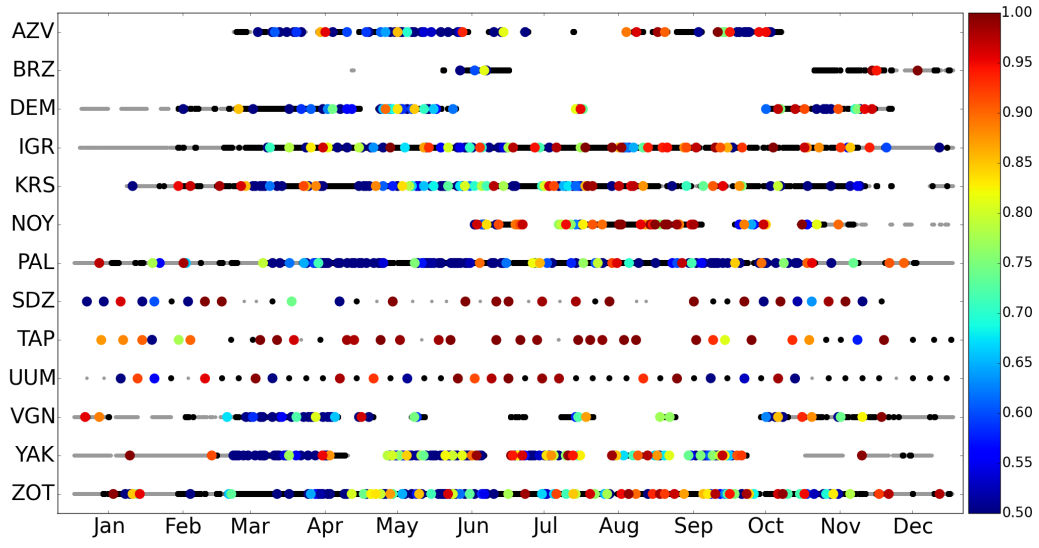

Figure 2: Observation availability and contribution to the inversion for each site. Grey dots stand for available data flagged out because of a planetary boundary layer height smaller than 500 m. Black dots denote the data points that are filtered during the computation of the algorithm described in Sect. 2 and 3.1. Colour circles represent the pieces of information carried out by each selected piece of data (see Sect. 2.3).

influence on the inversion. The two main explanations for this inability to assimilate all the available pieces of data is the chosen scale of interest and the integrating character of the atmosphere. First, as we are interested in meso-scale fluxes, the system has been chosen with a spatial resolution of 25 to 100 km. All the variability in CH_4 mixing ratios driven by single local plumes cannot be reproduced in the system. Second, the limitation of the atmospheric inversion comes from the fact that the atmosphere behaves as an integrator, hence attenuating some information in the atmospheric signal. Tracking back the atmospheric signal to the fluxes then has intrinsic limitations. This limited capability of the system drastically reduces the number of usable pieces of information. Out of the 127000 hourly measurements available in 2010, the pre-processing (as defined in Sect. 3.1) only retains 2000 daily aggregates into the inversion system. The system then excludes some observations and, at the end, only 800 data points remains, with 460 pieces of information (i.e. the trace of the sensitivity matrix; see details in Cardinali et al., 2004) carried by the atmospheric signal. Many observations give redundant information in our specific inversion framework at the meso-scale. The observations that cannot be processed by the meso-scale marginalized inversion carry information about local emissions.

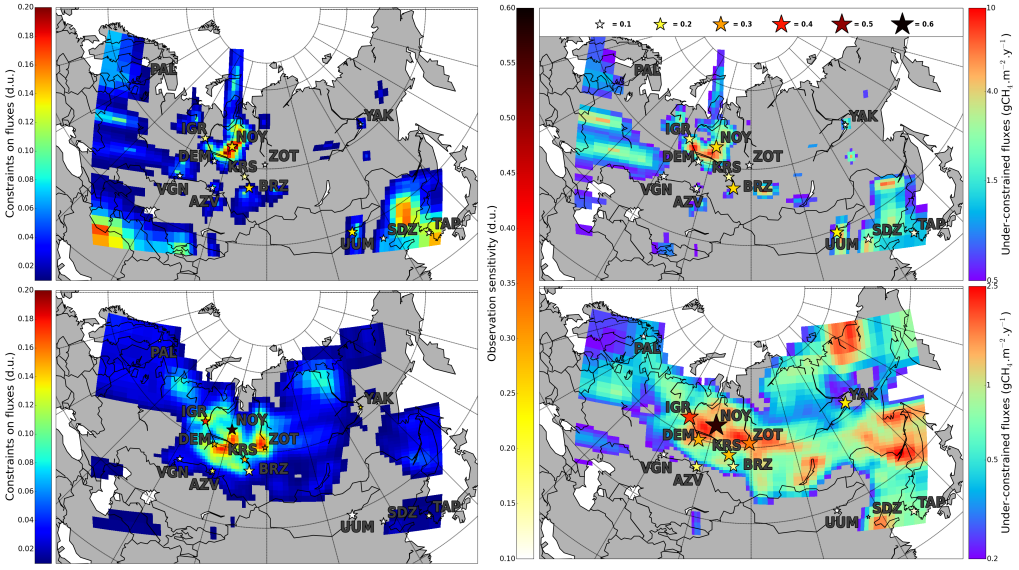

Figure 3: Maps of the constraints on the fluxes (top: anthropogenic fluxes; bottom: wetland emissions) from the observations in the inversion. (left column) Constraints on the fluxes as calculated from the influence matrix (see Sect. 2.3.1) and sensitivity of the inversion to observation sites as defined by their average weight ω in Eq. 4a of Sect. 2.3.1. (right column) Non-constrained fluxes as computed from the vector ϵ defined in Eq. 4b of Sect. 2.3.1 in $gCH_4.m^{-2}.y^{-1}$.

4.1.2 Network range of constraints

The temporal use of the available observations matters in the inversion, as much as the relative use of the different observation sites. Fig. 3 displays the location and the average weight of each observation site (the coloured background is described and discussed further in Sect. 4.2). It is divided into four panels in order to separate the use of information related to anthropogenic (top panels) and wetland (bottom panels) emissions. As the marginalized inversion raw results are also processed in order to detect the regions that are mis-separated from the boundary conditions (see Sect. 2.3.2), data from side observation sites is noticeably less used than for central sites.

Comparing anthropogenic (top panel) and wetland (bottom panel) maps, we notice that the weights of the observations are smaller for anthropogenic hot spot emissions. As expected, the inversion experiences difficulties in constraining emission hot spots, compared to diffuse fluxes. Concerning the spatial distribution of observation weights, wetland-related constraints follow the heuristics that the closer the observation site is to the fluxes, the higher the constraints to the inversion system is. Anthropogenic-related constraints do not exhibit such a pattern. For instance, NOY, close to the main oil extraction fields, has a lower observational influence than BRZ, remote from hot spots. Looking at wetland-related influences, BRZ has a bigger influence than NOY, while, as for anthropogenic emissions, wetlands emit more intensely in the vicinity of NOY. Then, a surface observation site

must be not too close, but not too far, from an emission hot spot to optimally constrain it. There is no generic criterion for this optimal distance to the observation sites as it depends on the atmospheric transport or the intensity of the hot spots.

Looking at the differences in the relative weights of the observation sites between the raw inversion results and the LBC-separated ones (not shown here), one can notice that the sites at the edge of the domain of interest are logically dedicated to constraining the LBC. Even the relative weight of the observation sites surrounding the Siberian Lowlands is significantly reduced. Additional observations away from the region of interest would be necessary to overcome the influence of the LBC.

4.2 Constrained regions

The spatial distribution of the observational constraints on the fluxes is calculated from the sensitivity matrix (see Sect. 2.3). The information in this matrix is convolved with the prior distribution of the fluxes to deduce the maps in the right column of Fig. 3. The spatial distribution of the constraints on the fluxes depends on the intensity of the emissions and the distance to the observation sites. In the left column of Fig. 3, the constraints are higher close to the Siberian network than away from it independent of the intensity of emissions. For example, the Western part of the Russian Federation contains most of the anthropogenic emissions of the country (roughly 20 $TgCH_4/y$ according to EDGAR FT2010). But the constraints are lower for this region than for the Siberian Lowlands with smaller emissions (8 $TgCH_4/y$ according to EDGAR FT2010). The post-processing excluding the regions that are mis-separated from the LBC by the inversion highlights this pattern. The observation sites within the denser part of the network seem to constrain emissions within a radius of roughly 500 km.

In the right column of Fig. 3, we display the average fluxes which are not considered as constrained by the inversion (as detailed in Sect. 2.3.1). Despite the limited range of the observation sites and the high number of filtered out data points, wetlands in the Siberian Lowlands are significantly constrained. The remaining unconstrained fluxes are of the same order of magnitude as minor wetland emitting region, such as in the far Eastern parts of Russian shores on the Arctic ocean and the Pacific ocean. As a consequence, with the existing network constraining major wetland areas, minor wetland regions now contribute equally to the uncertainties on Siberian CH_4 budget. This points to a needed extension of the monitoring network toward these minor emitting wetland areas. For anthropogenic emissions, the constraints on oil and gas related emissions are still too low. Anthropogenic emissions in Siberian Lowlands are still to be inquired into to reduced uncertainties on Russian CH_4 emissions.

In Fig. 5 described in Sect. 2.3.3, we also explicitly compute the portion of constrained emissions per month. On average, the major part of the emissions is not constrained by the inversion. The maximum proportion of constrained emissions is reached in summer with 50–70% of the emissions

constrained. In contrast, in winter, only a small part of the emissions are constrained. The proportion is critical in January and November when only 0–9% of the emissions are constrained. In Sect. 4.1.1, we noticed that most of the observations are flagged out in winter because of the very low boundary layer. Consequently, emissions in winter months are not well constrained.

Overall, the configuration of the network gives valuable insights on Siberian Lowland emissions, but it is not entirely adapted to our objective of constraining the Siberian Lowland CH_4 budget, even during the period when supply issues do not prevent the acquisition of observations and when the atmosphere is mixed enough for the CTM to accurately reproduce the transport patterns. Additional observation sites would be needed for a complete resolution of the regional fluxes.

4.3 Solved time and space resolution

In order to inquire into possible improvements in the regional inversion, we compute the typical temporal and spatial resolutions the inversion can solve. These scales are plotted in Fig. 4.

For each pixel of the domain of interest, we consider the groups the pixel is within at the different periods of the year (months for anthropogenic pixels, 10-day periods for wetland emissions). We then average the size of the selected groups along the year. We do the same for the duration of each period. The lowest common time step for anthropogenic and wetland emissions is month. Anthropogenic are thus consistently solved at a monthly scale. As wetland emissions can be grouped with anthropogenic ones, their solved time resolution can be increased from the 10-day basis. The temporal resolution is then computed only for wetland emissions in Fig. 4.

As expected, the resolved spatial resolution is better close to the network. Thus, most of Siberian Lowlands are constrained with a typical resolution below 2×10^6 km^2. The best resolution in our system configuration is roughly 700000 km^2. As we chose aggregation patterns with a mesh of about 300000 km^2 in Sect. 2.4.2, this confirms that our aggregation procedures are not too coarse.

However, the resolved spatial resolution suggests that numerous additional monitoring sites would be required to identify emission patterns in relation to hydrological and meteorological parameters as wetlands can react quickly and with high gradients to changes in the weather or in the water table depth. A high temporal resolution would be required in the inversion to link wetland emissions to these physical parameters. We see in Fig. 4 that most wetland fluxes are constrained with a temporal resolution of typically two weeks, which is too long to resolve quick changes in emissions. Wetlands along the Yenissei river, far from anthropogenic emissions are solved with a better temporal resolution than wetlands in central Lowlands. Unfortunately, this encouraging resolution is compensated by very high posterior uncertainties in the Yenissei flux. Overall, wetland emissions are resolved at

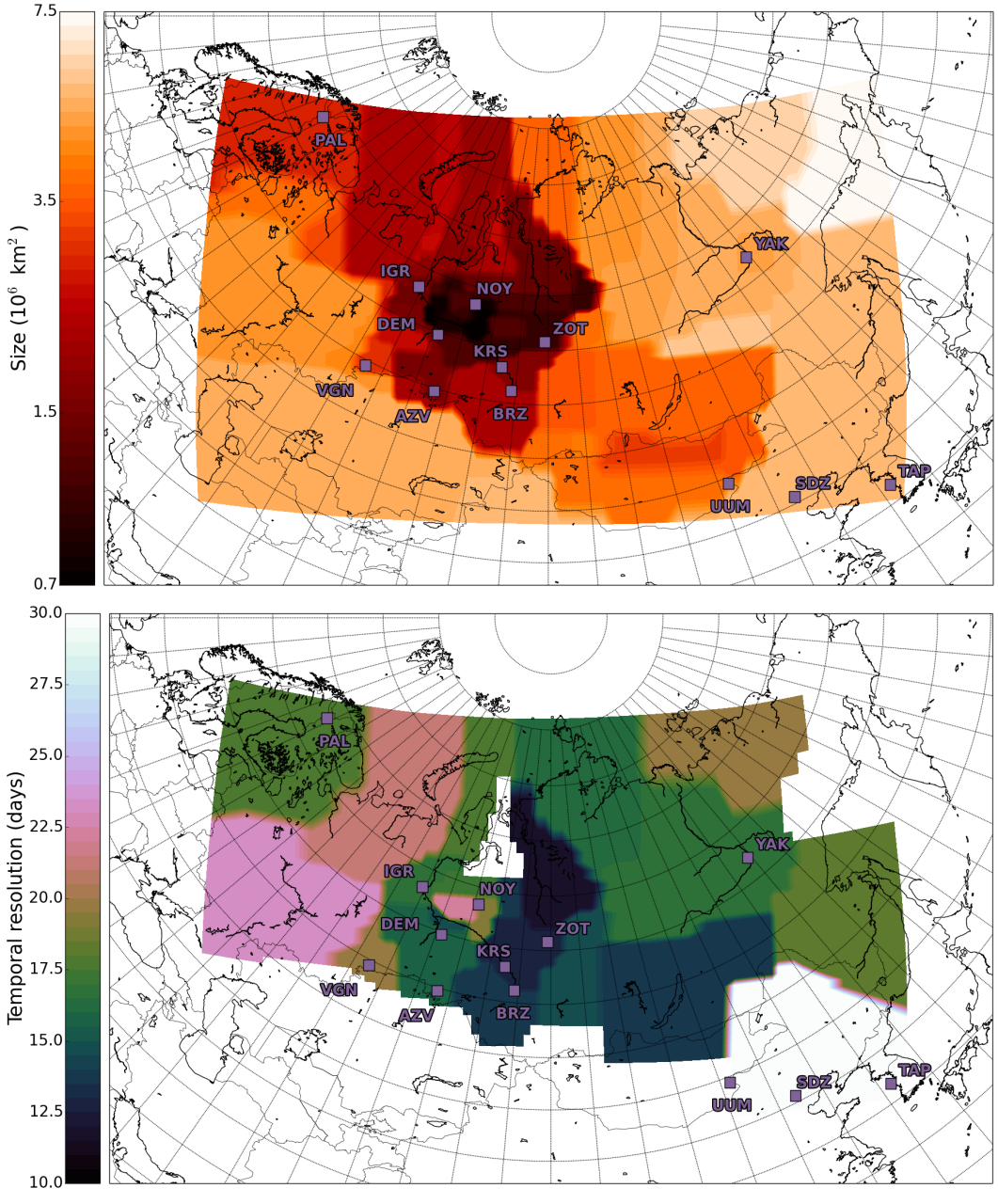

Figure 4: Temporal and spatial resolutions solved by the inversion. (top) Typical sizes of aggregation groups (as defined in Sect. 2.3.2); for each pixel, the size of the group the pixel belongs to is averaged along the whole period of study. (bottom) Time-scales resolved for wetland emissions; the period covered by pixel groups is averaged along the year.

a temporal resolution that allows the detection of the seasonal cycle, but not sufficient for linking wetland emissions to physical parameters varying at the synoptic scale.

To summarize, the inversion approach that we developed allows a precise quantification of the use of the observations. We also can deduce where the inversion results are the most reliable from the spatial influence of the network. At the meso-scale and in the Siberian framework, it appears that i) hourly and even daily measurements are difficult to assimilate, ii) anthropogenic emission hot spots require observation sites remote from them to be inverted, iii) diffuse wetland emissions can be constrained with sites located close to them, iv) the observation sites constrain fluxes within a radius of ~ 500 km in our meso-scale inversion. This knowledge could find applications in network design and in the choice of the type of measurements.

5 Results of the marginalized inversion

In Sect. 5.1, we describe the methane fluxes resulting from the inversion system that are explicitly constrained and not mis-separated from the lateral boundary conditions. We then extrapolate the fluxes on the constrained regions to the Siberian Lowlands and discuss the total budget of CH_4 of the region (Sect. 5.2).

5.1 Inverted fluxes

We defined successive filtering and aggregating procedures in Sect. 2. At the end of the inversion, out of the 275 constrained aggregated regions, only 166 groups of regions are kept: 35 anthropogenic groups, 101 wetland groups and 30 mixed groups. Constraints, ill-separated regions and posterior fluxes are summarized in Fig. 5 described in Sect. 2.3.3.

In summer, when wetlands are active, only August and September provide dominant wetland signals whereas winter months (December to April) show dominant anthropogenic sources. The rest of the year (May to July and October to November), the mixed source dominates, indicating that the inversion meets with difficulties in separating anthropogenic from wetland emissions. In the configuration we use, as expected with a sparse network, the usable pieces of information are not sufficient to fully reach the objective to separate emission types. In general, as we filtered out hot spots too close from the observation sites, only distant emissions and diffuse ones are constrained. Thus, the inversion system tries to separate similar mixed atmospheric signals from different sources. With no reliable information on the temporal profile of the different co-located emissions or any measurements of the isotopic atmospheric composition, separating co-located emissions with atmospheric meso-scale inversion is not possible.

In almost all cases, the intensity of the prior fluxes is close to the middle of the posterior tolerance interval. This means that the inventories and models used to get the prior fluxes in the region are reasonably close to the physical fluxes. The main output of the marginalized inversion compared

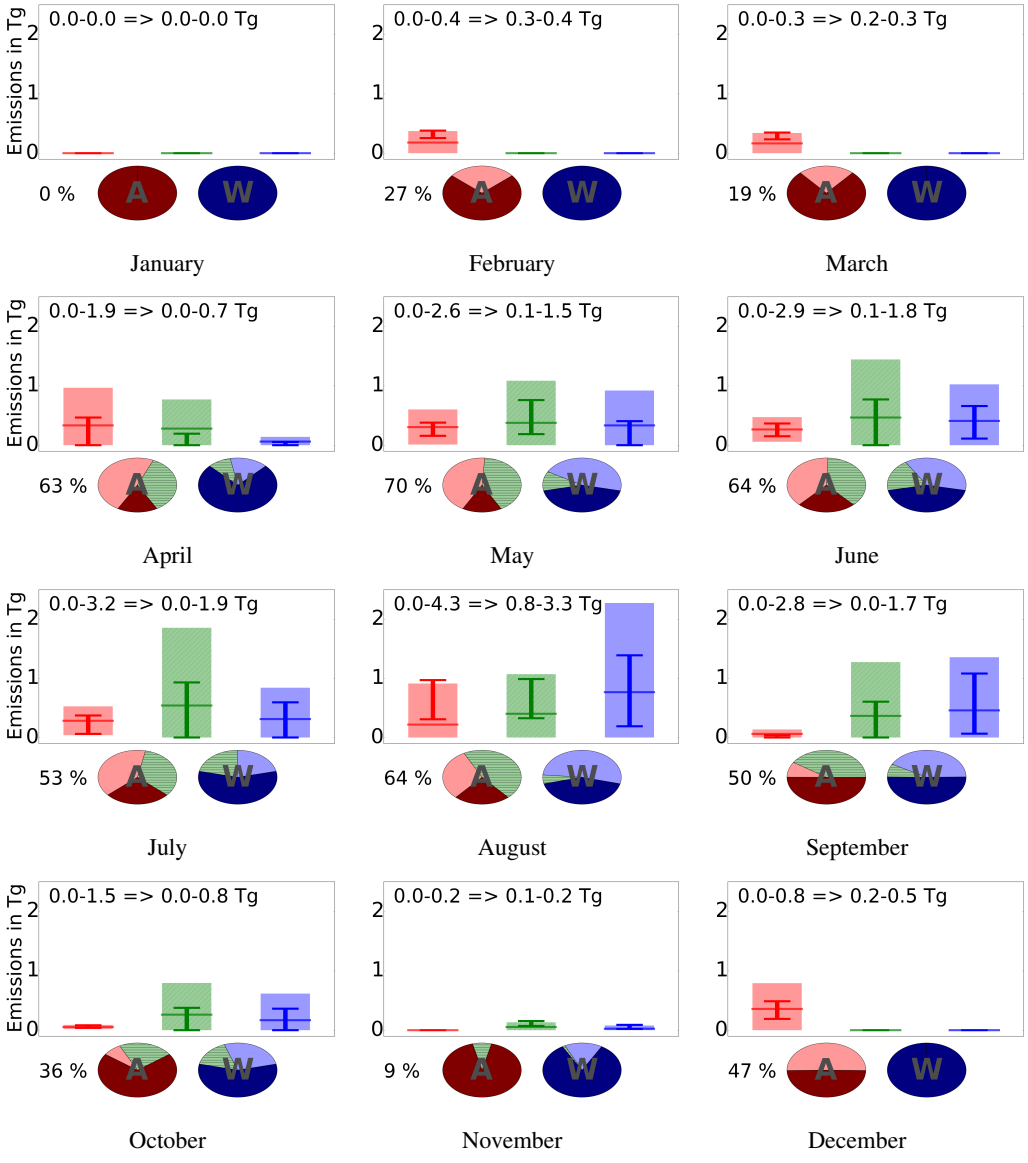

Figure 5: Inversion results for the twelve months of 2010 in the Siberian Lowlands (hatched area in Fig. 1). For each month, the bar plots show the prior (shade) and posterior (bar) tolerance intervals for anthropogenic (left), mixed (middle) and wetland (right) emissions. The pie charts depict the proportion of prior emissions within regions constrained by the inversion: (left, A) anthropogenic, (right, W) wetlands. Dark portions mean no constraint; lighter shade shows constrained regions; the hatched green portions are for the emissions the type of which the inversion cannot separate. The percentage indicates the fraction of the total emissions that are constrained each month. The given intervals in Tg are for the prior and posterior constrained Siberian budgets. See Sect. 2.3.3 for details.

to model or bottom-up approaches is the explicit computation of tolerance intervals of posterior fluxes on the regions that are explicitly constrained. The posterior tolerance intervals are significantly reduced compared to the prior uncertainties computed from the maximum likelihood. In most cases, the maximum likelihood algorithm suggests prior uncertainties more than 100% of the prior emissions. This is consistent with the uncertainties in the wetland models and in the anthropogenic emission inventories in the region (e.g., Hayhoe et al., 2002; Melton et al., 2013). After the inversion, on the constrained regions, the tolerance intervals are equivalent to uncertainties below 100%. For anthropogenic emissions, the posterior uncertainties are in the range of 20–50%, while for wetland emissions, the uncertainties span from 60 to 120%. The large ranges of uncertainties attributed to wetland emissions can be explained by the high temporal and spatial variability of the fluxes. The real temporal profiles and spatial patterns of wetland emissions differ from the ones used by the inversion system. The inversion thus experiences difficulties in precisely reproducing wetland fluxes, but consistently compute high posterior uncertainties. Despite the mis-separation of the co-located emissions, the marginalized inversion thus provides reliable estimates of fluxes uncertainties at the regional scale.

5.2 Siberian Lowland CH_4 budget

5.2.1 Seasonal cycle and yearly emissions

The marginalized inversion gives explicit tolerance intervals only on the constrained regions, that embody less than two thirds of the regional emissions. We then do not have more information than the prior emissions about the regions that are not upwind of the the observation sites or that are filtered out because of the plume criterion. Nevertheless, we are inquiring into the regional budget of methane. Extrapolations are necessary to infer regional emissions from the partial vision given by Fig. 5.

Here, for each month of the year 2010, we decide to extrapolate the corrections applied in Sect. 5.1 to all the fluxes only when a significant part of the prior emissions are constrained (hereafter 20%). As the prior emissions are close to the middle of the posterior tolerance intervals, this extrapolation does not radically change the inverted budget, but it allows the computation of regional uncertainties. The time series of the extrapolated anthropogenic, wetland and total emissions in the Siberian Lowlands (hatched area in Fig. 1) are displayed in Fig. 6. For each type of emissions, the posterior tolerance intervals are computed only for months with more than 20% of constrained emissions for the associated type. The total annual budgets in the legend of Fig. 6 are computed from the annual corrections on the constrained region for anthropogenic, wetland and total emissions.

Wetland emissions are negligible in winter and intense in summer. In the prior emission, the seasonal cycle of the regional emissions is smooth, the emitting season beginning in March. The marginalized inversion suggests a shift to April-May for the start of the CH_4 emissions from wet-

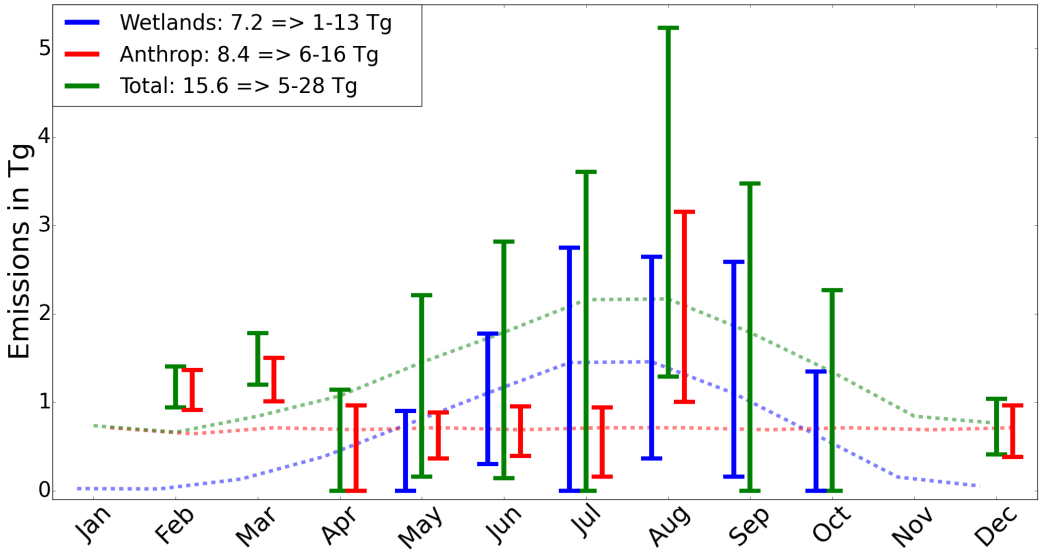

Figure 6: Time series of the posterior emissions on the Siberian Lowlands. The emissions are reported in $TgCH_4$ for each month. The correction factors computed on the constrained regions (in Sect. 5) are extrapolated to the Siberian Lowlands (see Fig. 1). The extrapolation is carried out only when more than 20% of the prior emissions are constrained. The error bars depict the tolerance interval of posterior emissions. The dotted lines denote the prior fluxes. The yearly prior emissions and posterior tolerance intervals in $TgCH_4$ are noted in the legend.

lands. Anthropogenic emissions are supposed to be constant in the prior. Posterior emissions in February and March are 20–30% higher than the average posterior emissions. In a region of dense oil and gas extraction, the emissions can be partly correlated to the production, but also to household heating consumption. As the demand of gas is higher in winter, the production is 15–20% bigger during the winter quarter (as suggested by GAZPROM annual reports, the main Russian gas company; www.gazprom.com). This could explain the higher posterior estimates in winter. However, local biogenic processes emitting CH_4 have also been observed in winter, which perturb the measured surface mixing ratios (Arshinov et al., 2009; Winderlich, 2012). Limited bacterial activity below the snow cover can produce CH_4, which perspires to the atmosphere. As the atmosphere is very stable in winter, this phenomenon can generate significant local increases of CH_4 mixing ratios close to the observation sites. As the model does not account for such emissions, the inversion would attribute them to the anthropogenic emissions. This should be furtherly inquired into with isotopic measurements. Nevertheless, most problematic observations with possible local biogenic influence within very stable planetary boundary layers have been flagged out from the inversion due to the mis-representation of the vertical mixing in the model. Therefore, the winter increase of CH_4 emissions in the inversion can be attributed to a real increase in anthropogenic emissions due to the cumulated increased production and consumption of gas for heating.

Overall, on a yearly basis, posterior anthropogenic and wetland emissions are roughly equal as in the prior. The tolerance interval is 1–13 $TgCH_4$/y for wetland emissions, and 6–16 $TgCH_4$/y for anthropogenic emissions. The ranges of uncertainties are reduced by 40% for wetland emissions and 57% for anthropogenic emissions from prior to posterior fluxes. Siberian oil and gas extracting activity and population centres are then responsible of 1.5–4.2% of the global anthropogenic emissions. As a large portion of the oil and gas extracted in Siberia is exported to Europe, the inclusion of the emissions from the Siberian extraction process would have a significant impact on the total European emissions of $\sim 27\,TgCH_4$ (as computed from EDGAR v4.2 for all Europe apart from Russia, Ukraine and Belarus). The uncertainties of the wetland emissions are still too high to provide valuable insights for the modelling of these emissions, such as the start of the emitting season, or correlation to the precipitation rates or temperatures.

5.2.2 Wildfire influence

A peak of anthropogenic emissions with large uncertainties occurs in August. The large tolerance interval is still compatible with the prior scenario. Though, analysing into details posterior fluxes in August aggregation groups, we can propose two plausible explanation to this increase un-captured by prior data sets.

First, an increase of $0.7\,TgCH_4$ in August anthropogenic emissions is suggested in oil and gas extraction regions. This could be explained by the activities of prospecting companies. They make use of the decrease in the household demand in summer to carry out maintenance and welling operations on the oil and gas welling sites. These operations are related to punctual leaks and purging releases of gas.

Second, August is also a month with numerous forest fires. In particular, very large wildfires occurred in Western Russia in August 2010, upwind our observation network. As said in Sect. 3.3, wildfires are systematically excluded from the marginalized inversion because they generate intense plumes difficult to take into account in our model. But, as West Russian wildfires took place far away of the observing sites, emitted plumes could have been well mixed before reaching these sites. Thus, the inversion system would have failed in eliminating observations very influenced by wildfires. The increase in the mixing ratios could then have been wrongly attributed by the inversion to an increase in anthropogenic emissions. As a confirmation, an aggregation group in August embraces anthropogenic emissions from Western Russia, grouped with anthropogenic and wetland emissions in the Siberian Lowlands. The inversion suggests that the emissions are increased by $+150\%$ for this group.

6 Evaluation of the inverion

6.1 Performance on filtered out data

For any inversion system, the inversion results are to be evaluated with independent datasets. A usual way to evaluate inversion results is to carry out a leave-one-out experiment. It usually consists in 1) flagging out all the observations from a site, 2) computing the inversion with the reduced data set, and 3) comparing the prior and posterior simulated mixing ratios to the observed ones at the left-out site. In principle, the leave-one-out inversion should improve the simulated mixing ratios at the left-out sites. That is to say, the posterior mixing ratios should be closer to and more consistent with the observations at the left-out site than the prior ones, even if not assimilated in the inversion. In our case, the observation sites are far from each other and not numerous in regard to the size of the domain. Moreover, as developed in Sect. 4.2, the marginalized inversion explicitly informs about the constrained regions. Flagging out one site modifies the inversion results mainly in the surrounding regions where the uncertainty reduction becomes negligible compared to the complete dataset. Therefore, the optimized fluxes in our leave-one-out experiments (not shown) remain within the range of uncertainty of another one for constrained regions. This confirms that the method we use consistently accounts for the uncertainties, but is not sufficient to quantitatively evaluate the optimized fluxes.

We can also use the data points which have been filtered out by our system in order to evaluate the inversion results. As the number of filtered out observations is high, sampling biases may be expected from the filtering procedures. In addition, as only a few data points are assimilated, unrealistic fluxes could have been inferred by the inversion to fit the assimilated data leading to a flawed reproduction of the remaining observations. As one can see in Tab. 2, the marginalized inversion significantly improves the simulated mixing ratios at the sites where data is used as expected. As a proof on realistic flux prescription despite the filter on the observations, for all the remaining data, the model results are also well improved for unused data.

This confirms that our method does not create sampling biases despite the high number of filtered data points. It also confirms that the increments on the fluxes are realistic from the point of view of our network.

6.2 GOSAT evaluation

Genuinely independent observations are required for a better evaluation of the inversion results. Long-term monitoring surface sites are scarce in Siberia, and airborne (e.g., Paris et al., 2010; Berchet et al., 2013a) or train (e.g., Tarasova et al., 2006) measurement campaigns only provide snapshots of the Siberian atmospheric composition. Therefore, we try to evaluate the marginalized inversion results with satellite data. We choose GOSAT total column biased-corrected retrievals (see Sect. 3.5) and compare them to their simulated equivalents. In Fig. 7, we compare the prior

Table 2: Correlations of observed and simulated mixing ratios, prior and posterior to the inversion. Correlation coefficient r is presented for only the filtered data points used in the inversion, but also for the data points filtered out from the inversion.The number of daily available observations is also reported alongside with the number of data used in the inversion.

Site ID	Correlations			Number of observations	
	Prior	Posterior		Daily raw	Used
		Used	Filtered		
AZV	0.70	0.89	0.89	131	54
BRZ	0.79	0.77	0.71	66	7
DEM	0.69	0.84	0.84	184	49
IGR	0.58	0.78	0.77	349	71
KRS	0.62	0.79	0.89	307	87
NOY	0.56	0.82	0.78	122	44
PAL	0.53	0.76	0.75	362	102
SDZ	0.58	0.99	0.99	32	26
TAP	0.70	0.99	0.99	41	31
UUM	0.09	0.65	0.69	51	21
VGN	0.64	0.82	0.81	204	44
YAK	0.17	0.91	0.90	272	115
ZOT	0.64	0.90	0.92	363	107

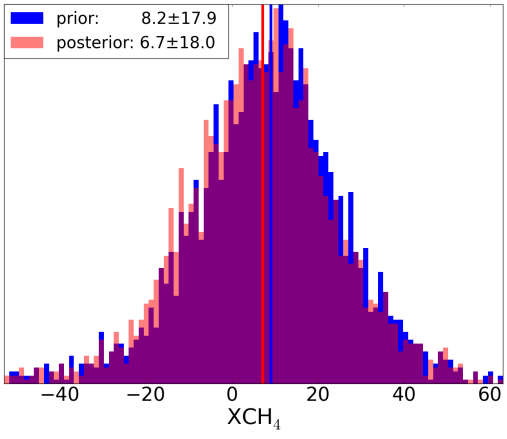

Figure 7: Distribution of the differences between the observed and simulated prior and posterior GOSAT total columns. The differences are computed only for the GOSAT observations downwind the emissions constrained by the surface sites, i.e. with a simulated contribution from these regions to the total column bigger than 2 ppb.

and posterior simulated total columns to the observed columns. Since the inversion only constrains the emissions that are upwind the assimilated surface observation sites, we focus on the GOSAT columns that are gathered downwind of the regions constrained by the surface sites. To do so, we select only the CH_4 columns, of which the model equivalent is computed with a contribution from the constrained regions larger than 2 ppb. This criterion keeps 3000 of the 25000 available GOSAT measurements in the domain of interest during the year 2010. More than 96% of the selected prior and posterior simulated columns are less than 10 ppb different from each other. This shift is not significant compared with the $\sim$10 ppb of uncertainties in the GOSAT retrievals.

GOSAT zonal coverage is rather uniform. GOSAT columns thus carry information on the spatial gradients of the atmospheric methane composition. As analysing the individual total columns does not provide sufficient information about the regional emissions, we therefore now focus on the spatial gradients of the total columns. In the observed CH_4 columns, we notice large-scale negative Eastward (from Europe to West Siberia) and Northward (from China to North-East Siberia) gradients within measured total columns. These gradients of about 5 ppb per 1000 km are related to the large-scale mixing of the polluted air masses from Europe and China into clean background air masses. The contribution of the local emissions to the observed total columns is thus computed here as the local increase from the regional background in order to bypass the very large scale gradients. In Fig. 8, we represent the simulated and observed contributions of the Siberian local emissions to the GOSAT total columns. As the wetlands emit small amounts of CH_4 ($< 10^{-3}$ kg.m^{-2}.month^{-1}) spread over very large areas ($\sim 2 \times 10^6$ km^2), their contribution to CH_4 columns is very smooth. In the model, the contribution of wetlands to the total columns does not exceed 3 ppb over the Siberian Lowlands. Therefore, the gradients in the CH_4 columns due to regional wetland emissions cannot be separated from the large scale gradients of 5 ppb per 1000 km related to the emissions outside the region of interest.

The anthropogenic hot spots are observable in both measured and simulated total columns. Above the hot spots in the Siberian Lowlands, the observed total columns shift from the observed background by up to 15 ppb. In the model, the local contribution of anthropogenic hot spots to the total columns is significant only close to the sources, with local increases in the total column up to 6 ppb. With the spatial resolution that we chose for atmospheric transport (with grid cells of more than 25 km to be compared with GOSAT pixel size of $\sim$ 10 km), the emissions from the local hot spots mix quickly into the background. This fast dilution can explain at least part of the factor of difference of 2–3 between observed and simulated CH_4 columns.

6.3 Toward using satellite measurements in regional frameworks

Satellite data, as a tool for filling the lack of observations in Siberia to evaluate the inversion results, do not seem to be suited for our regional configuration. As satellite observations offer a very good

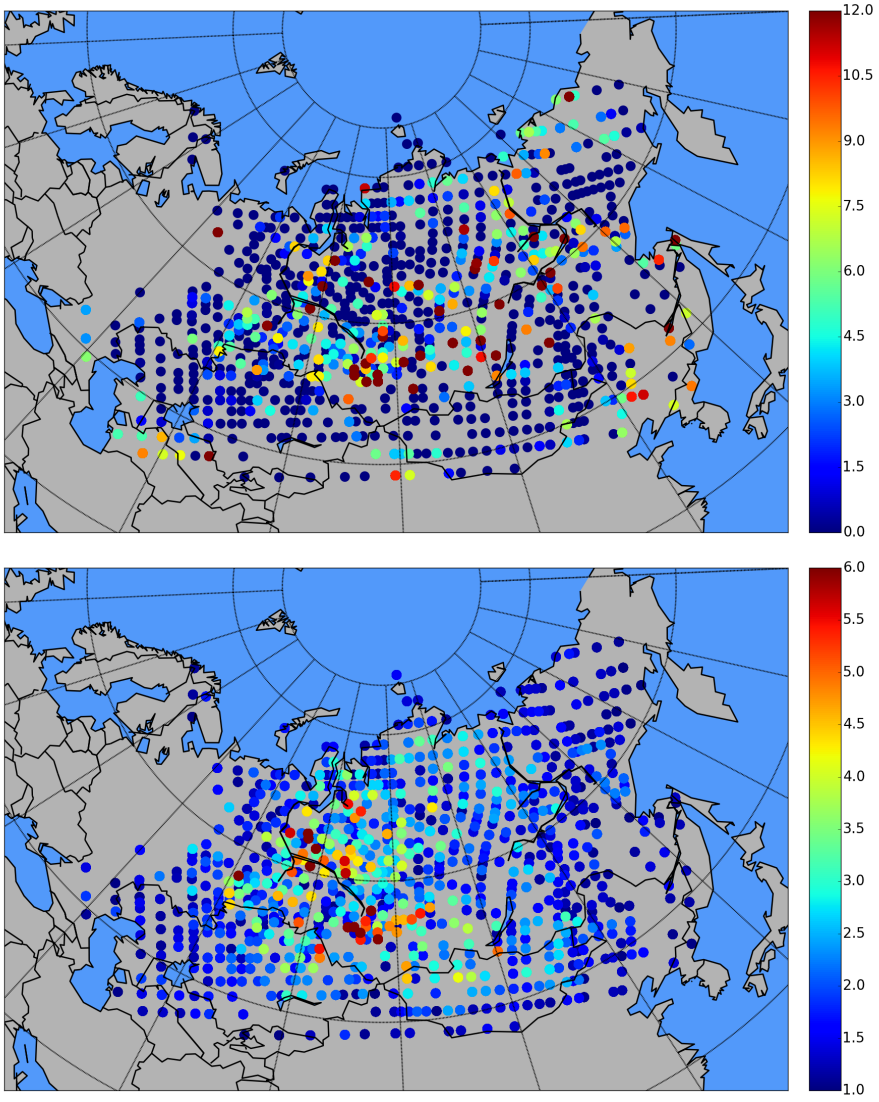

Figure 8: Observed (top) and simulated (bottom) contributions of the local Siberian sources to GOSAT total columns. The observed contribution of local sources are selected as excesses from the regional measured background (defined as the average observed total column in a zone of roughly 2000 km in the zonal direction and 1000 km in the meridional one). The simulated contributions are computed from CHIMERE forward simulations of the emissions of the Siberian Lowlands. The colour scale is not the same between the two panels in order to facilitate the readability of the figure.

potential in term of spatial coverage, in-situ observations difficult to maintain and extend in the Siberian framework should be complemented by new satellite data sets.

The inversion system could be developed in order to use GOSAT data as a proxy of the large scale gradients given here by a global model. With such additional observations, the LBC, representative of the emissions outside the domain of interest, would be better constrained in the inversion. As a consequence, fewer regions would be expected to be mis-separated from the LBC and the assimilated surface sites at the edge of the network would then provide more information about the emissions within the domain of interest. In addition, high resolution transport simulations close to the hot spots would better represent the dilution of the plumes. The model-observation comparison for quantifying anthropogenic sources would then be more suitable. However, the plumes generated by the hot spots are not necessarily simulated at the right location and date. Indeed, the location of emission hot spots is not always well known a priori and temporal or spatial transport mismatches can occur. Direct comparisons between the observed and simulated CH_4 columns would face the same issue as for surface sites when assimilating hot spots plumes. Recent developments (e.g., Krings et al., 2013) point to the possibility of using 1D or 2D high resolution snapshots of the hot spot plumes to infer information about very local emissions. Integrated comparison of the observed and simulated plumes in the CH_4 columns could then be implemented in a meso-scale inversion system.

With such techniques, future satellite missions with active remote sensing (e.g., the joint French-German cooperation Methane Remote Sensing LIDAR Mission, MERLIN) providing high resolution accurate 1D or 2D products could be used in regional inversions; the spatial resolution of the products to be used in such an inversion system should of the same magnitude as the meso-scale transport model, i.e. at least < 50 km.

7 Conclusions

We assimilated the data collected in 2010 at 8 surface observation sites measuring atmospheric CH_4 mixing ratios in the West-Siberian plain into a regional atmospheric inversion. It was the first time all these observations were used in a single study. As regional inversions suffer from mis-specified uncertainties, we implemented an enhanced Bayesian method developed by Berchet et al. (2014) in order to get reliable results at the regional scale with an objectified quantification of the uncertainties in the system. This new method allows us to consistently evaluate the local spatial distribution of the sensitivity of the emission areas to the inversion and the usefulness of each available observation. The inversion seems to be able to primarily constrain the emissions close to the observation sites (within a radius of roughly 500 km). The inversion system assimilate daily observation aggregates to constrain the emissions. Amongst all the observation aggregates, and despite the efforts to provide precise and quasi-continuous measurements, our meso-scale inversion system properly uses only one piece of information every few days. This is mainly caused by atmospheric limitations related

to transport and mixing; the fewer the assimilated data, the higher the uncertainties keep after the inversion. Even in the regions close to the stations, the posterior uncertainties (objectively quantified) thus remain larger than 20% of the prior fluxes for anthropogenic emissions and 50% for wetlands, although an important error reduction is achieved. Nevertheless, objectified uncertainties allow a robust evaluation of the wide range of proposed wetland and anthropogenic emissions in Siberia. On average, the posterior tolerance interval (defined so that 68.27% of the Monte Carlo marginalized ensemble is within the interval) on the West-Siberian plain methane budget is 5–28 $TgCH_4$ for the year 2010.

The year 2010 is the first year when most of the used observation sites are functional. Reproducing our set-up to subsequent years would provide a more robust estimation of the regional fluxes and possibly valuable information about the year-to-year variability of Siberian methane fluxes. Finally, satellite platforms provide an extensive spatial coverage of observational constraints. Implementing such rather uniform observation coverage in a regional framework with few surface sites is tempting. However, with the inversion framework used here, satellite data would be useful only for constraining large scale gradients, hence the lateral boundary conditions. Further work on inversion systems is required so that satellite observations can be used to quantify local emissions in a regional framework like this one.

Acknowledgements. We thank all the PIs and supporting staff from the sites we used for maintaining the instruments and providing quasi-continuous and precise measurements. We thank Anmyeon-do's (South Korea) PI for contribution to early developments of the inversion system. We thank the reviewers for their fruitful remarks and suggestions to significantly improve our first manuscript. This research was supported by the Commissariat à l'Énergie Atomique et aux Énergies Renouvelables, by the French ANR-CLIMSLIP and the YAK-AEROSIB project. The maintenance of the network is also supported by state contracts of the Ministry of Education and Science of Russia No. 14.613.21.0013 (RFMEFI61314X0013).

References

Agency, U. E. P.: Inventory of U.S. greenhouse gas emissions and sinks: 1990-2009, Tech. Rep. 430-R-11-005, U.S. EPA, Washington D.C., 2011.

Amante, C. and Eakins, B. W.: ETOPO1 1 arc-minute global relief model: procedures, data sources and analysis, US Department of Commerce, National Oceanic and Atmospheric Administration, National Environmental Satellite, Data, and Information Service, National Geophysical Data Center, Marine Geology and Geophysics Division, 2009.

Arshinov, M. Y., Belan, B. D., Davydov, D. K., Inouye, G., Krasnov, O. A., Maksyutov, S., Machida, T., Fofonov, A. V., and Shimoyama, K.: Spatial and temporal variability of CO2 and CH4 concentrations in the surface atmospheric layer over West Siberia, Atmos Ocean Opt, 22, 84–93, doi:10.1134/S1024856009010126, http://link.springer.com/article/10.1134/S1024856009010126, 2009.

Berchet, A., Paris, J.-D., Ancellet, G., Law, K. S., Stohl, A., Nédélec, P., Arshinov, M. Y., Belan, B. D., and Ciais, P.: Tropospheric ozone over Siberia in spring 2010: remote influences and stratospheric intrusion, Tellus B, 65, doi:10.3402/tellusb.v65i0.19688, http://www.tellusb.net/index.php/tellusb/article/view/19688/0, 2013a.

Berchet, A., Pison, I., Chevallier, F., Bousquet, P., Conil, S., Geever, M., Laurila, T., Lavrič, J., Lopez, M., Moncrieff, J., Necki, J., Ramonet, M., Schmidt, M., Steinbacher, M., and Tarniewicz, J.: Towards better error statistics for atmospheric inversions of methane surface fluxes, Atmos. Chem. Phys., 13, 7115–7132, doi:10.5194/acp-13-7115-2013, http://www.atmos-chem-phys.net/13/7115/2013/, 2013b.

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

Bergamaschi, P., Frankenberg, C., Meirink, J. F., Krol, M., Villani, M. G., Houweling, S., Dentener, F., Dlugokencky, E. J., Miller, J. B., Gatti, L. V., Engel, A., and Levin, I.: Inverse modeling of global and regional CH$_4$ emissions using SCIAMACHY satellite retrievals, Journal of Geophysical Research: Atmospheres, 114, doi:10.1029/2009JD012287, http://onlinelibrary.wiley.com.biblioplanets.gate.inist.fr/doi/10.1029/2009JD012287/abstract, 2009.

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 CH4 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, doi:10.1002/jgrd.50480, http://onlinelibrary.wiley.com/doi/10.1002/jgrd.50480/abstract, 2013.

Bohn, T. J., Lettenmaier, D. P., Sathulur, K., Bowling, L. C., Podest, E., McDonald, K. C., and Friborg, T.: Methane emissions from western Siberian wetlands: heterogeneity and sensitivity to climate change, Environ. Res. Lett., 2, 045 015, doi:10.1088/1748-9326/2/4/045015, http://iopscience.iop.org/1748-9326/2/4/045015, 2007.

Bousquet, P., Ciais, P., Miller, J. B., Dlugokencky, E. J., Hauglustaine, D. A., Prigent, C., Werf, G. R. V. d., 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, doi:10.1038/nature05132, http://www.nature.com/nature/journal/v443/n7110/full/nature05132.html, 2006.

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

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

Cardinali, C., Pezzulli, S., and Andersson, E.: Influence-matrix diagnostic of a data assimilation system, Q. J. R. Meteorol. Soc., 130, 2767–2786, doi:10.1256/qj.03.205, http://onlinelibrary.wiley.com/doi/10.1256/qj.03.205/abstract, 2004.

Chapnik, B., Desroziers, G., Rabier, F., and Talagrand, O.: Properties and first application of an error-statistics tuning method in variational assimilation, Q. J. R. Meteorol. Soc., 130, 2253–2275, doi:10.1256/qj.03.26, http://onlinelibrary.wiley.com/doi/10.1256/qj.03.26/abstract, 2004.

Chevallier, F., Feng, L., Bösch, H., Palmer, P. I., and Rayner, P. J.: On the impact of transport model errors for the estimation of CO$_2$ surface fluxes from GOSAT observations, Geophysical Research Letters, 37, doi:10.1029/2010GL044652, http://onlinelibrary.wiley.com/doi/10.1029/2010GL044652/abstract, 2010.

Cressot, C., Chevallier, F., Bousquet, P., Crevoisier, C., Dlugokencky, E. J., Fortems-Cheiney, A., Frankenberg, C., Parker, R., Pison, I., Scheepmaker, R. A., Montzka, S. A., Krummel, P. B., Steele, L. P., and Langenfelds, R. L.: On the consistency between global and regional methane emissions inferred from SCIAMACHY, TANSO-FTS, IASI and surface measurements, Atmos. Chem. Phys., 14, 577–592, doi:10.5194/acp-14-577-2014, http://www.atmos-chem-phys.net/14/577/2014/, 2014.

Crosson, E. R.: A cavity ring-down analyzer for measuring atmospheric levels of methane, carbon dioxide, and water vapor, Appl. Phys. B, 92, 403–408, doi:10.1007/s00340-008-3135-y, http://link.springer.com/article/10.1007/s00340-008-3135-y, 2008.

Dedikov, J., Akopova (VNIIGaz), G., Gladkaja (VNIIGaz), N., Piotrovskij (Tyumentransgaz), A., Markellov (Volgotransgaz), V., Salichov (Yamburggazdabuicha), S., Kaesler, H., Ramm, A., Müller von Blumencron, A., and Lelieveld, J.: Estimating methane releases from natural gas production and transmission in Russia, Atmospheric Environment, 33, 3291–3299, doi:10.1016/S1352-2310(98)00388-4, http://www.sciencedirect.com/science/article/pii/S1352231098003884, 1999.

Dentener, F., Peters, W., Krol, M., Van Weele, M., Bergamaschi, P., and Lelieveld, J.: Interannual variability and trend of CH$_4$ lifetime as a measure for OH changes in the 1979-1993 time period, J. Geophys. Res., 108, 4442, doi:10.1029/2002JD002916, http://www.agu.org/pubs/crossref/2003/2002JD002916.shtml, 2003.

Dlugokencky, E. J., Steele, L. P., Lang, P. M., and Masarie, K. A.: The growth rate and distribution of atmospheric methane, Journal of Geophysical Research: Atmospheres, 99, 17 021–17 043, doi:10.1029/94JD01245, http://onlinelibrary.wiley.com.biblioplanets.gate.inist.fr/doi/10.1029/94JD01245/abstract, 1994.

Dlugokencky, E. J., Myers, R. C., Lang, P. M., Masarie, K. A., Crotwell, A. M., Thoning, K. W., Hall, B. D., Elkins, J. W., and Steele, L. P.: Conversion of NOAA atmospheric dry air CH$_4$ mole fractions to a gravimetrically prepared standard scale, J. Geophys. Res., 110, D18 306, doi:10.1029/2005JD006035, http://europa.agu.org/?uri=/journals/jd/jd0518/2005JD006035/2005JD006035.xml&view=article, 2005.

Dlugokencky, E. J., Bruhwiler, L., White, J. W. C., Emmons, L. K., Novelli, P. C., Montzka, S. A., Masarie, K. A., Lang, P. M., Crotwell, A. M., Miller, J. B., and Gatti, L. V.: Observational constraints on recent increases in the atmospheric CH$_4$ burden, Geophys. Res. Lett., 36, L18 803, doi:10.1029/2009GL039780, http://onlinelibrary.wiley.com/doi/10.1029/2009GL039780/abstract, 2009.

Enting, I. G., Rayner, P. J., and Ciais, P.: Carbon Cycle Uncertainty in REgional Carbon Cycle Assessment and Processes (RECCAP), Biogeosciences, 9, 2889–2904, doi:10.5194/bg-9-2889-2012, http://biogeosciences.net/9/2889/2012/bg-9-2889-2012.html, 2012.

Forster, P., Ramaswamy, V., Artaxo, P., Berntsen, T., Betts, R., Fahey, D. W., Haywood, J., Lean, J., Lowe, D. C., and Myhre, G.: Changes in atmospheric constituents and in radiative forcing, Climate change, 20, http://cfpub.epa.gov/ncea/hero/index.cfm?action=search.view=92936, 2007.

Frey, K. E. and Smith, L. C.: How well do we know northern land cover? Comparison of four global vegetation and wetland products with a new ground-truth database for West Siberia, Global Biogeochemical Cycles, 21, doi:10.1029/2006GB002706, http://onlinelibrary.wiley.com/doi/10.1029/2006GB002706/abstract, 2007.

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

Giglio, L., Loboda, T., Roy, D. P., Quayle, B., and Justice, C. O.: An active-fire based burned area mapping algorithm for the MODIS sensor, Remote Sensing of Environment, 113, 408–420, doi:10.1016/j.rse.2008.10.006, http://www.sciencedirect.com/science/article/pii/S0034425708003180, 2009.

Glagolev, M., Kleptsova, I., Filippov, I., Maksyutov, S., and Machida, T.: Regional methane emission from West Siberia mire landscapes, Environ. Res. Lett., 6, 045 214, doi:10.1088/1748-9326/6/4/045214, http://iopscience.iop.org/1748-9326/6/4/045214, 2011.

Hayhoe, K., Kheshgi, H. S., Jain, A. K., and Wuebbles, D. J.: Substitution of Natural Gas for Coal: Climatic Effects of Utility Sector Emissions, Climatic Change, 54, 107–139, doi:10.1023/A:1015737505552, http://link.springer.com/article/10.1023/A%3A1015737505552, 2002.

Kasischke, E. S. and Bruhwiler, L. P.: Emissions of carbon dioxide, carbon monoxide, and methane from boreal forest fires in 1998, Journal of Geophysical Research: Atmospheres (1984–2012), 107, FFR–2, http://www.agu.org/journals/jd/jd0223/2001JD000461/, 2002.

Kirschke, S., Bousquet, P., Ciais, P., Saunois, 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 Quéré, 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, Nature Geosci, 6, 813–823, doi:10.1038/ngeo1955, http://www.nature.com/ngeo/journal/v6/n10/full/ngeo1955.html, 2013.

Krings, T., Gerilowski, K., Buchwitz, M., Hartmann, J., Sachs, T., Erzinger, J., Burrows, J. P., and Bovensmann, H.: Quantification of methane emission rates from coal mine ventilation shafts using airborne remote sensing data, Atmos. Meas. Tech., 6, 151–166, doi:10.5194/amt-6-151-2013, http://www.atmos-meas-tech.net/6/151/2013/, 2013.

Lamarque, J.-F., Kyle, G. P., Meinshausen, M., Riahi, K., Smith, S. J., Vuuren, D. P. v., Conley, A. J., and Vitt, F.: Global and regional evolution of short-lived radiatively-active gases and aerosols in the Representative Concentration Pathways, Climatic Change, 109, 191–212, doi:10.1007/s10584-011-0155-0, http://link.springer.com/article/10.1007/s10584-011-0155-0, 2011.

Lauvaux, T., Pannekoucke, O., Sarrat, C., Chevallier, F., Ciais, P., Noilhan, J., Rayner, P. J., et al.: Structure of the transport uncertainty in mesoscale inversions of CO$_2$ sources and sinks using ensemble model simulations, Biogeosciences, 6, 1089–1102, http://content.imamu.edu.sa/Scholars/it/net/bg-6-1089-2009.pdf, 2009.

Lechtenböhmer, S., Dienst, C., Fischedick, M., Hanke, T., Langrock, T., Assonov, S. S., and Brenninkmeijer, C.: Greenhouse gas emissions from the Russian natural gas export pipeline system : results and extrapolation of measurements and surveys in Russia, p. 40, http://epub.wupperinst.org/frontdoor/index/index/docId/2136, 2005.

Lehner, B. and Döll, P.: Development and validation of a global database of lakes, reservoirs and wetlands, Journal of Hydrology, 296, 1–22, doi:10.1016/j.jhydrol.2004.03.028, http://www.sciencedirect.com/science/article/pii/S0022169404001404, 2004.

Lin, J. C., Gerbig, C., Wofsy, S. C., Andrews, A. E., Daube, B. C., Davis, K. J., and Grainger, C. A.: A near-field tool for simulating the upstream influence of atmospheric observations: The Stochastic Time-Inverted Lagrangian Transport (STILT) model, Journal of Geophysical Research: Atmospheres, 108, doi:10.1029/2002JD003161, http://onlinelibrary.wiley.com/doi/10.1029/2002JD003161/abstract, 2003.

Melton, J. R., Wania, R., Hodson, E. L., Poulter, B., Ringeval, B., Spahni, R., Bohn, T., Avis, C. A., Beerling, D. J., Chen, G., Eliseev, A. V., Denisov, S. N., Hopcroft, P. O., Lettenmaier, D. P., Riley, W. J., Singarayer, J. S., Subin, Z. M., Tian, H., Zürcher, S., Brovkin, V., van Bodegom, P. M., Kleinen, T., Yu, Z. C., and Kaplan, J. O.: Present state of global wetland extent and wetland methane modelling: conclusions from a model inter-comparison project (WETCHIMP), Biogeosciences, 10, 753–788, doi:10.5194/bg-10-753-2013, http://www.biogeosciences.net/10/753/2013/, 2013.

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, doi:10.5194/gmd-6-981-2013, http://www.geosci-model-dev.net/6/981/2013/, 2013.

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

Olivier, J. G. J., Van Aardenne, J. A., Dentener, F. J., Pagliari, V., Ganzeveld, L. N., and Peters, J. A. H. W.: Recent trends in global greenhouse gas emissions:regional trends 1970–2000 and spatial distributionof key sources in 2000, Environmental Sciences, 2, 81–99, doi:10.1080/15693430500400345, http://dx.doi.org/10.1080/15693430500400345, 2005.

Paris, J. D., Ciais, P., Nédélec, P., Stohl, A., Belan, B. D., Arshinov, M. Y., Carouge, C., Golitsyn, G. S., and Granberg, I. G.: New insights on the chemical composition of the Siberian air shed from the YAK-AEROSIB aircraft campaigns, B. Am. Meteorol. Soc., 91, 625–641, http://zardoz.nilu.no/~andreas/publications/168.pdf, 2010.

Parker, R., Boesch, H., Cogan, A., Fraser, A., Feng, L., Palmer, P. I., Messerschmidt, J., Deutscher, N., Griffith, D. W. T., Notholt, J., Wennberg, P. O., and Wunch, D.: Methane observations from the Greenhouse Gases Observing SATellite: Comparison to ground-based TCCON data and model calculations, Geophysical Research Letters, 38, n/a–n/a, doi:10.1029/2011GL047871, http://onlinelibrary.wiley.com/doi/10.1029/2011GL047871/abstract, 2011.

Peregon, A., Maksyutov, S., and Yamagata, Y.: An image-based inventory of the spatial structure of West Siberian wetlands, Environ. Res. Lett., 4, 045 014, doi:10.1088/1748-9326/4/4/045014, http://iopscience.iop.org/1748-9326/4/4/045014, 2009.

Pickett-Heaps, C. A., Jacob, D. J., Wecht, K. J., Kort, E. A., Wofsy, S. C., Diskin, G. S., Worthy, D. E. J., Kaplan, J. O., Bey, I., and Drevet, J.: Magnitude and seasonality of wetland methane emissions from the Hudson Bay Lowlands (Canada), Atmos. Chem. Phys., 11, 3773–3779, doi:10.5194/acp-11-3773-2011, http://www.atmos-chem-phys.net/11/3773/2011/, 2011.

Pison, I., Menut, L., and Bergametti, G.: Inverse modeling of surface NOx anthropogenic emission fluxes in the Paris area during the Air Pollution Over Paris Region (ESQUIF) campaign, J. Geophys. Res., 112, D24 302, doi:10.1029/2007JD008871, http://www.agu.org/pubs/crossref/2007/2007JD008871.shtml, 2007.

Potter, C. S., Randerson, J. T., Field, C. B., Matson, P. A., Vitousek, P. M., Mooney, H. A., and Klooster, S. A.: Terrestrial ecosystem production: A process model based on global satellite and surface data, Global Biogeochem. Cycles, 7, 811–841, doi:10.1029/93GB02725, http://onlinelibrary.wiley.com.biblioplanets.gate.inist.fr/doi/10.1029/93GB02725/abstract, 1993.

Reshetnikov, A. I., Paramonova, N. N., and Shashkov, A. A.: An evaluation of historical methane emissions from the Soviet gas industry, Journal of Geophysical Research: Atmospheres, 105, 3517–3529, doi:10.1029/1999JD900761, http://onlinelibrary.wiley.com/doi/10.1029/1999JD900761/abstract, 2000.

Ringeval, B., de Noblet-Ducoudré, N., Ciais, P., Bousquet, P., Prigent, C., Papa, F., and Rossow, W. B.: An attempt to quantify the impact of changes in wetland extent on methane emissions on the seasonal and interannual time scales, Global Biogeochemical Cycles, 24, doi:10.1029/2008GB003354, http://onlinelibrary.wiley.com.biblioplanets.gate.inist.fr/doi/10.1029/2008GB003354/abstract, 2010.

Rödenbeck, C., Houweling, S., Gloor, M., and Heimann, M.: CO$_2$ flux history 1982-2001 inferred from atmospheric data using a global inversion of atmospheric transport, Atmos. Chem. Phys., 3, 1919–1964, doi:10.5194/acp-3-1919-2003, http://www.atmos-chem-phys.net/3/1919/2003/, 2003.

Sasakawa, M., Shimoyama, K., Machida, T., Tsuda, N., Suto, H., Arshinov, M., Davydov, D., Fofonov, A., Krasnov, O., Saeki, T., Koyama, Y., and Maksyutov, S.: Continuous measurements of methane from a tower

network over Siberia, Tellus B, 62, 403–416, doi:10.1111/j.1600-0889.2010.00494.x, http://onlinelibrary.wiley.com/doi/10.1111/j.1600-0889.2010.00494.x/abstract, 2010.

Sasakawa, M., Ito, A., Machida, T., Tsuda, N., Niwa, Y., Davydov, D., Fofonov, A., and Arshinov, M.: Annual variation of CH$_4$ emissions from the middle taiga in West Siberian Lowland (2005–2009): a case of high CH$_4$ flux and precipitation rate in the summer of 2007, Tellus B, 64, doi:10.3402/tellusb.v64i0.17514, http://www.tellusb.net/index.php/tellusb/article/view/17514, 2012.

Schmidt, H., Derognat, C., Vautard, R., and Beekmann, M.: A comparison of simulated and observed ozone mixing ratios for the summer of 1998 in Western Europe, Atmos. Environ., 35, 6277–6297, http://www.sciencedirect.com/science/article/pii/S1352231001004514, 2001.

Spahni, R., Wania, R., Neef, L., van Weele, M., Pison, I., Bousquet, P., Frankenberg, C., Foster, P. N., Joos, F., Prentice, I. C., and van Velthoven, P.: Constraining global methane emissions and uptake by ecosystems, Biogeosciences, 8, 1643–1665, doi:10.5194/bg-8-1643-2011, http://www.biogeosciences.net/8/1643/2011/, 2011.

Spahni, R., Joos, F., Stocker, B. D., Steinacher, M., and Yu, Z. C.: Transient simulations of the carbon and nitrogen dynamics in northern peatlands: from the Last Glacial Maximum to the 21st century, Clim. Past, 9, 1287–1308, doi:10.5194/cp-9-1287-2013, http://www.clim-past.net/9/1287/2013/, 2013.

Stocker, B. D., Roth, R., Joos, F., Spahni, R., Steinacher, M., Zaehle, S., Bouwman, L., Xu-Ri, and Prentice, I. C.: Multiple greenhouse-gas feedbacks from the land biosphere under future climate change scenarios, Nature Clim. Change, 3, 666–672, doi:10.1038/nclimate1864, http://www.nature.com/nclimate/journal/v3/n7/full/nclimate1864.html, 2013.

Stocker, B. D., Spahni, R., and Joos, F.: DYPTOP: a cost-efficient TOPMODEL implementation to simulate sub-grid spatio-temporal dynamics of global wetlands and peatlands, Geosci. Model Dev. Discuss., 7, 4875–4930, doi:10.5194/gmdd-7-4875-2014, http://www.geosci-model-dev-discuss.net/7/4875/2014/, 2014.

Stohl, A., Forster, C., Frank, A., Seibert, P., Wotawa, G., et al.: Technical note: The Lagrangian particle dispersion model FLEXPART version 6.2, Atmos. Chem. Phys., 5, 2461–2474, http://hal-insu.archives-ouvertes.fr/hal-00301615/, 2005.

Suto, H. and Inoue, G.: A New Portable Instrument for In Situ Measurement of Atmospheric Methane Mole Fraction by Applying an Improved Tin Dioxide–Based Gas Sensor, Journal of Atmospheric and Oceanic Technology, 27, 1175–1184, doi:10.1175/2010JTECHA1400.1, http://journals.ametsoc.org/doi/abs/10.1175/2010JTECHA1400.1, 2010.

Tarantola, A.: Inverse Problem Theory, Elsevier, New York, 1987.

Tarasova, O., Brenninkmeijer, C., Assonov, S., Elansky, N., Röckmann, T., and Brass, M.: Atmospheric CH$_4$ along the Trans-Siberian railroad (TROICA) and river Ob: Source identification using stable isotope analysis, Atmospheric Environment, 40, 5617–5628, doi:10.1016/j.atmosenv.2006.04.065, http://www.sciencedirect.com/science/article/pii/S1352231006004614, 2006.

Umezawa, T., Machida, T., Aoki, S., and Nakazawa, T.: Contributions of natural and anthropogenic sources to atmospheric methane variations over western Siberia estimated from its carbon and hydrogen isotopes, Global Biogeochemical Cycles, 26, doi:10.1029/2011GB004232, http://onlinelibrary.wiley.com.biblioplanets.gate.inist.fr/doi/10.1029/2011GB004232/abstract, 2012.

Uppala, S. M., Kållberg, P. W., Simmons, A. J., Andrae, U., Bechtold, V. D. C., Fiorino, M., Gibson, J. K., Haseler, J., Hernandez, A., Kelly, G. A., Li, X., Onogi, K., Saarinen, S., Sokka, N., Allan, R. P., Andersson, E., Arpe, K., Balmaseda, M. A., Beljaars, A. C. M., Berg, L. V. D., Bidlot, J., Bormann, N., Caires, S., Chevallier, F., Dethof, A., Dragosavac, M., Fisher, M., Fuentes, M., Hagemann, S., Hólm, E., Hoskins, B. J., Isaksen, L., Janssen, P. a. E. M., Jenne, R., Mcnally, A. P., Mahfouf, J.-F., Morcrette, J.-J., Rayner, N. A., Saunders, R. W., Simon, P., Sterl, A., Trenberth, K. E., Untch, A., Vasiljevic, D., Viterbo, P., and Woollen, J.: The ERA-40 re-analysis, Q.J.R. Meteorol. Soc., 131, 2961–3012, doi:10.1256/qj.04.176, http://onlinelibrary.wiley.com/doi/10.1256/qj.04.176/abstract, 2005.

van der Werf, G. R., Randerson, J. T., Giglio, L., Collatz, G. J., Mu, M., Kasibhatla, P. S., Morton, D. C., DeFries, R. S., Jin, Y., and van Leeuwen, T. T.: Global fire emissions and the contribution of deforestation, savanna, forest, agricultural, and peat fires (1997–2009), Atmos. Chem. Phys., 10, 11 707–11 735, doi:10.5194/acp-10-11707-2010, http://www.atmos-chem-phys.net/10/11707/2010/, 2010.

Vautard, R., Beekmann, M., Roux, J., and Gombert, D.: Validation of a hybrid forecasting system for the ozone concentrations over the Paris area, Atmospheric Environment, 35, 2449–2461, doi:10.1016/S1352-2310(00)00466-0, http://www.sciencedirect.com/science/article/pii/S1352231000004660, 2001.

Villani, M. G., Bergamaschi, P., Krol, M. C., Meirink, J. F., and Dentener, F.: Inverse modeling of European CH$_4$ emissions: sensitivity to the observational network, Atmos. Chem. Phys., 10, 1249–1267, http://igitur-archive.library.uu.nl/phys/2011-0323-200350/UUindex.html, 2010.

Winderlich, J.: Setup of a CO$_2$ and CH$_4$ measurement system in Central Siberia and modeling of its results, Ph.D. thesis, Univ. of Hamburg, Germany, http://ediss.sub.uni-hamburg.de/volltexte/2012/5533/, 2012.

Winderlich, J., Chen, H., Gerbig, C., Seifert, T., Kolle, O., Lavrič, J. V., Kaiser, C., Höfer, A., and Heimann, M.: Continuous low-maintenance CO$_2$/CH$_4$/H$_2$O measurements at the Zotino Tall Tower Observatory (ZOTTO) in Central Siberia, Atmospheric Measurement Techniques, 3, 1113–1128, doi:10.5194/amt-3-1113-2010, http://www.atmos-meas-tech.net/3/1113/2010/amt-3-1113-2010.html, 2010.

Winiarek, V., Bocquet, M., Saunier, O., and Mathieu, A.: Estimation of errors in the inverse modeling of accidental release of atmospheric pollutant: Application to the reconstruction of the cesium-137 and iodine-131 source terms from the Fukushima Daiichi power plant, J. Geophys. Res., 117, D05 122, doi:10.1029/2011JD016932, http://onlinelibrary.wiley.com/doi/10.1029/2011JD016932/abstract, 2012.

Yamada, K., Yoshida, N., Nakagawa, F., and Inoue, G.: Source evaluation of atmospheric methane over western Siberia using double stable isotopic signatures, Organic Geochemistry, 36, 717–726, doi:10.1016/j.orggeochem.2005.01.016, http://www.sciencedirect.com/science/article/pii/S0146638005000355, 2005.

Zhou, L. X., Kitzis, D., and Tans, P. P.: Report of the fourth WMO round-robin reference gas intercomparison, 2002-2007, in: Report of the 14th WMO Meeting of Experts on Carbon Dioxide Concentration and Related Tracer Measurement Techniques, WMO/GAW Rep, vol. 186, pp. 40–43, 2009.