Biogeosciences Discuss., 11, 14587–14637, 2014
www.biogeosciences-discuss.net/11/14587/2014/
doi:10.5194/bgd-11-14587-2014

Biogeosciences
Discussions
Open Access

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

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

[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

Received: 17 September 2014 – Accepted: 29 September 2014 – Published: 13 October 2014

Correspondence to: A. Berchet (aberchet@lsce.ipsl.fr)

Published by Copernicus Publications on behalf of the European Geosciences Union.

Discussion Paper | Discussion Paper | Discussion Paper | Discussion Paper |

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 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 signals are 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 are 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 Tg CH_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 forcings as a greenhouse gas and is involved in the atmospheric oxydizing capacity as an ozone precursor (Forster et al., 2007). Characterizing the variability of the atmospheric CH_4 composition requires an accurate understanding of the methane biogeochemical cycle, in particular of the surface-atmosphere fluxes, their spatial distribution and 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 Tg CH_4 yr^{-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 Tg CH_4 yr^{-1}, i.e. 25–50 % of total emissions with a very high heterogeneous spatial distribution and year-to-year variability (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 Tg CH_4 yr^{-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, 2005; Spahni et al., 2011). 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. Russia produces 20 % of the natural gas in the world, mostly extracted in Siberia. 0.1–10 % (i.e. 0.5–40 Tg CH_4 yr^{-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 (~ 10–15 % of the global production), with 1–2 % of the oil production leaked into the atmosphere as methane (e.g., US Environment Protection 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 kilometers, 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 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 gaslines and gas facilities to deduce emission factors. They upscaled their figure to the Russian territory and got an emission factor lower than 1 % of the total production. Reshetnikov et al. (2000) reviewed the existing litterature 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 absolute and isotopic composition in CH_4 (and related species).

The variations of the atmospheric composition provides information on the relative contribution of the different local process 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 to the 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) was carried out. Their approach relied on atmospheric inversion techniques based on the data assimiliation 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. However, atmospheric inversions with one observation site give only local constraints on the surface fluxes. More critically, a one-site atmospheric inversion 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.

The atmospheric “top–down” inversion approach relies on the simulation of the atmospheric transport of air-masses in order to track-back their origin. Chemistry-Transport Models used for simulating the atmospheric transport suffer from

uncertainties in reproducing the real atmospheric transport. The uncertainties come for instance from the transport errors (Prather et al., 2008; Locatelli et al., 2013). Recent literature highlights the need for precisely and objectively quantifying all of these errors 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. These improved algorithms are based on a Monte Carlo approach superimposed to Maximum of Likelihood estimators (Chapnik et al., 2004; Michalak and Kitanidis, 2005).

For the first time in Eurasia, we use these improved algorithms 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 airmasses in-/out-coming to/from the region. These sites, which have been operated since the mid-2000's, are implemented into the inversion system with objectified uncertainty quantification of Berchet et al. (2014). 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 improvements used in our study in Sect. 2.1. In Sect. 2.2, 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. 3. Our inversion is then evaluated in Sect. 4 by using independent satellite observations, as the very few available surface observations do not allow to keep some for evaluation.

2 Methods

In Sect. 2.1, we describe the atmospheric inversion approach we use for the quantification of Siberian CH_4 fluxes. In Sect. 2.2, we present the Siberian set-up for the inversion.

2.1 Marginalized Bayesian inversion

2.1.1 Method outline

As the atmosphere mixes irreversibly air masses from different CH_4 sources, using the atmospheric signal 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. To do so, we look for the pdf of the state of the system with some knowledge about the atmospheric composition and on the state distribution. More precisely, the pdf can be written: $p(\boldsymbol{x}|\boldsymbol{y}^{\mathrm{o}} - H(\boldsymbol{x}^{\mathrm{b}}), \boldsymbol{x}^{\mathrm{b}})$. The vector $\boldsymbol{y}^{\mathrm{o}}$ gathers all the available observations; H is the observation operator converting the information in the state vector to the observation space; the vector $\boldsymbol{x}$ depicts the state of the system and $\boldsymbol{x}^{\mathrm{b}}$ is the background vector including the prior knowledge on the state 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. 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 $\boldsymbol{x}$ encompasses the surface fluxes and the lateral boundary conditions.

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.). The inversion then deduces an optimal posterior state vector $\widehat{\boldsymbol{x}^{\mathrm{a}}}$ and posterior uncertainties. With the usual Gaussian assumption, all the uncertainties are considered as normal pdfs and can be described with uncertainty covariance matrices.

Then, the posterior state vector and uncertainty matrice can be explicitly defined:

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

In Eq. (1), the matrix $\mathbf{K} = \mathbf{BH}^{\mathrm{T}}(\mathbf{R} + \mathbf{HBH}^{\mathrm{T}})^{-1}$ is the Kalman gain matrix. The matrices **R** and **B** are the covariance matrices describing the observation and background uncertainties. Observation uncertainties encompass measurement, discretization and transport errors. Background uncertainties include the uncertainties about the spatial distribution of the fluxes, their temporal varibility and 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 implementation on computers. 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 speciying 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 **R** and **B** along an objective statistical criterion: the Maximum of Likelihood. The implementation of the method gave encouraging results, but the impact of the uncertainties within the Maximum of Likelihood algorithm 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.

The marginalization consists in computing the pdf $p(\boldsymbol{x}|\boldsymbol{y}^{\mathrm{o}} - \mathbf{H}\boldsymbol{x}^{\mathrm{b}}, \boldsymbol{x}^{\mathrm{b}})$ as a sum of the pdfs calculated for each possible uncertainty matrices **R** and **B**. As the sum of all the local pdfs has a complex non-Gaussian shape, the marginalization is computed through a Monte Carlo approach. To do so, a large number of posterior vectors $\widehat{\boldsymbol{x}^{\mathrm{a}}}$ is

computed through individual inversions as in Eq. (1) with different uncertainty matrices **R** and **B**.

Computing such numerous individual inversions is costly and requires extensive amounts of memory. Observation sampling and flux aggregation must be carried out before the marginalized inversion in order to reduce its complexity without degrading the inversion optimality. The observation sampling is chosen as the minimum measured mixing ratio per day per observation site. 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 only adding noise to the system. The patterns for the aggregation of emissions are chosen so that the mesh goes finer closer to the observation network. The resolution of the transport representation is also chosen finer close to the observation network. Additionaly, regions upwind of observation sites must be separated from the unseen areas. Separating seen from unseen areas and choosing the length scale for mesh resolution reduction requires a pre-processing on the fluxes before the inversion. The pre-processing of the prior fluxes is based on the footprints of the observation network. These footprints are estimated with a Lagrangian model (Sect. 2.2.2) which offers an efficient way to compute them. Deeper details on the aggregation choice are given in the method description paper (Berchet et al., 2014).

The marginalized inversion applies to the pre-processed aggregated fluxes. However, a last operation is necessary to perform the inversion. Indeed, in the ensemble of posterior vectors ($\widehat{\boldsymbol{x}^{\mathrm{a}}}$), only the regions that are reliably constrained and optimized during the marginalized inversion should remain. Some of the pre-processed fluxes are either not seen by the observations because they are negligible compared to other contributions or ill-inverted because plumes emitted by hot spots are erroneously reproduced in the observation operator due to temporal and spatial mismatches. These areas are difficult to optimize in the single atmospheric inversions. These problematic components of the state vector are dynamically detected during the marginalized inversion (see details in Berchet et al., 2014) and filtered out.

To summarize, the marginalized inversion used in this study relies on the same inputs as other atmospheric inversions: atmospheric measurements, prior fluxes and a representation of the atmospheric transport. A pre-processing procedure is applied on the observation and fluxes in order to reduce the complexity of the problem. Then, a Monte Carlo ensemble of inversions is carried out with first-guess uncertainty matrices **R** and **B**. The distribution of the matrices **R** and **B** is deduced from the Maximum of Likelihood computed with a pseudo-Newtonian algorithm. We choose the distribution of **R** and **B** as a χ^2-like distribution of diagonal matrices, the average of which is the Maximum of Likelihood ($\mathbf{R}_{max}, \mathbf{B}_{max}$).

2.1.2 Marginalized inversion outputs

The final outputs of the marginalized inversion are: (i) an ensemble of posterior vectors ($\widehat{\boldsymbol{x}^a}$), (ii) the influence matrix $\mathbf{K}_{max}\mathbf{H}$ and (iii) the sensitivity matrix $\mathbf{H}\mathbf{K}_{max}$. The influence and sensitivity matrices here depends on ($\mathbf{R}_{max}, \mathbf{B}_{max}$), the tuple of uncertainty matrices at the Maximum of Likelihood. Following Cardinali et al. (2004), the diagonal terms of the influence (resp. sensitivity) matrix give explicit diagnostics about the regions that are constrained by the observation network (resp. about the constraints to the inversion given by each individual observation).

The ensemble of posterior vectors does not follow a normal distribution as detailed in Berchet et al. (2014). We then 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. With the ensemble of posterior vectors, we can also compute posterior covariances. These covariances indicate when posterior state components are mis-separated by the inversion. Classical inversions require a cautious examination to detect such mis-separated regions. For instance, mis-separated regions can induce artificial flux dipoles (e.g., Rödenbeck et al., 2003). The explicit posterior covariances allow us to define groups of regions that are mis-separated. The physical relevance and confidence on the grouped results is then improved.

Moreover, as detailed in Sects. 2.2.3 and 2.2.4, we focus on a domain of limited area. We then need information on the global CH_4 mixing ratios entering the domain. The global mixing ratios come from a Global Circulation Model, which suffer uncertainties. Errors on the lateral boundary conditions can mis-lead the inversion on the regional fluxes. From the posterior covariances, we can inquire into the mis-separation of inside regional fluxes with global mixing ratios and exclude from the final results the regions which are impacted by the uncertainties on the lateral boundary conditions.

In Sect. 3, we comment and discuss the posterior tolerance intervals, correlations and the influence indicators.

2.2 Set up for an Eurasian domain

We are interested in the Eurasian surface–atmosphere budget of methane. As developed in Sect. 2.1, the computation of the marginalized inversion needs measurements ($\boldsymbol{y}^o$; Sect. 2.2.1), prior fluxes ($\boldsymbol{x}^b$; Sect. 2.2.3) and an observation operator computed with an atmospheric transport model (**H**; Sect. 2.2.4). The footprints of the observation network required for choosing the aggregation patterns (necessary to the computation of **H**) are estimated with a Lagrangian model (Sect. 2.2.2). The observations are compared to the fluxes through the observation operator **H**. The operator **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.

2.2.1 The observation network: $\boldsymbol{y}^o$

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) and the Finnish Meteorological Institute (FMI, Helsinki, Finland) and NOAA-ESRL (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 Table 1). The location, measurement type, maintaining institute and 3-letter symbol of each observation site are described in Table 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). They are sent to and analyzed 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-STATIONS; Sasakawa et al., 2010) and are equiped 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). The ZOT site is operated by MPI since April 2009 (Winderlich et al., 2010) and air is analyzed by an EnviroSense 3000i analyzer (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.

Due to possible logistical issues or instrument dysfunctions, observation sites do not provide measurements all year round. Figure 2, described in Sect. 3.1.1, details the temporal availability of the observations. The sampling bias is known to impact the inversion results (Villani et al., 2010). The issue is detailed in Sect. 3.1.1, but the general method developed in Sect. 2.1 consistently takes into account such a bias into the posterior tolerance intervals.

In an earlier paper (Berchet et al., 2013b), it was suggested that assimilating observations by night or with a planetary boundary layer (PBL) thiner than 500 m may deteriorate the inversion results. Indeed, under these conditions with very limited vertical mixing close to the surface, the transport model cannot reproduce the atmospheric mixing ratios at the surface observation sites properly. For that reason, we flag out all data according to these criteria. Out of 125 000 hourly measurements in 2010, 30 000 passes through the PBL and night filter (see black dots in Fig. 2).

Additionaly, as explained in Sect. 2.1, the marginalized inversion computation is limited by the size of the observation vector. To reduce the number of observation fed into the system, we sample the minimum measurement of mixing ratios by day and site. This sampling protocol flags out local plumes that cannot be properly represented by the transport model. Moreover, in most cases, the daily minimum occurs at the time when the PBL height is maximum, mostly in the afternoon, hence when local effects are vertically well mixed and then better represented in the transport model.

The final size of $\boldsymbol{y}^{o}$ implemented in the inversion is 2000. On average, 0.4 observations per station per day are validated for the inversion.

2.2.2 Estimate 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 $\boldsymbol{x}^{b}$ and the observation operator $\mathbf{H}$, can be computed. 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° × 1°, with 60 vertical levels and 3 h temporal resolution. Virtual particles are released

in a 3-D box 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.

2.2.3 Prior fluxes and state vector: *x*

The inversion system optimizes prior fluxes grouped in the regions aggregated by the pre-processing procedure (see Sect. 2.1.1). 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 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 Transiberian Railroad 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 m a.s.l. 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 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, wetdays), 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 spot 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 thin plumes, the marginalized inversion exclude from the system all the emission contribution related to fires.

For the computation of the observation operator **H** (see Sect. 2.2.4), we use a regional chemistry-transport model with a domain limited in space and time. Initial and lateral boundary conditions (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 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 kilometers 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 × 2 horizontal levels + roof top) ×36 10 day period for the lateral boundary conditions.

2.2.4 The observation operator: H

We explicitly define the observation operator **H** by computing the forward atmospheric transport from the regions of aggregated emissions (defined according to Sect. 2.2.3 criteria) to the observation sites. As CH_4 is a reactive species, the observation operator should include the oxidation by OH radicals. However, the typical residence time of the air masses in the domain of interest is of a few days to weeks. As the life time of CH_4 in the atmosphere is estimated to be 8–10 years (Dentener et al., 2003), CH_4 chemistry can be neglected in the computation of the observation operator at our temporal and spatial scales.

Thus, for each aggregated region, we calculate the so-called response functions using the transport module of the Eulerian mesoscale 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. 2.1 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 3-D-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° × 0.5° every 3 h.

2.2.5 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. 2.2.1 for the optimization of CH_4 fluxes. 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. 4, 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 ∼ 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. ∼ 10 ppb). 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 ∼ 25 000 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. 2.2.3). The average observed XCH_4 is ∼ 1775 ppb over the domain throughout the year. The prior average in LMDz XCH_4 is ∼ 1820 ppb. This bias is attributable to the excessive injection of tropospheric air into the stratosphere in our version of LMDz. In Sect. 4, the bias on XCH_4 of 45 ppb has been corrected to allow the observation-model comparison.

3 Results

The marginalized inversion described in Sect. 2.1 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 first present and analyze the effects of the data selection in the inversion and the implied limitations (Sect. 3.1). We then describe and discuss the inverted fluxes in the regions that are constrained by the observation network (Sect. 3.2). Siberian Lowlands are partially constrained by the observations. We extrapolate the fluxes on the constrained regions to the Siberian Lowlands and discuss the total budget of CH_4 of the region (Sect. 3.3).

3.1 Diagnostics of the marginalized inversion

3.1.1 Observation weight in the inversion

As explained in Sect. 2.1, 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 flags out additional observations when they are measured during thin plumes difficult to inverse. The remaining pieces of data do not have the same weight in the inversion. Contrary to classical inversion methods, the marginalized inversion allows us to explicitly analyze the use of the remaining observations in the system. This diagnostics on the use of the observation is deduced from the diagonal terms of the sensitivity matrix (see Sect. 2.1.2). 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 (colored 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 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 25–100 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 127 000 hourly measurements available in 2010, the pre-processing (as defined in Sect. 2.2.1) 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.

The temporal use of the available observations matters in the inversion, as much as the relative use of the different observation sites. Figure 3 displays the location and the average weight of each observation site. It is divided into four panels in order to separate the use of information related to anthropogenic (top panel) and wetland (bottom panel) emissions. The marginalized inversion raw results are also processed in order to detect the regions that are mis-separated from the boundary conditions. The left column presents the raw inversion diagnostics while the right column displays the LBC-separated ones.

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 to the fluxes, the higher the constraints it provides to the inversion system. 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 intensly 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, 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.

3.1.2 Constrained regions

The spatial distribution of the observational constraints on the fluxes is calculated from the sensitivity matrix (see Sect. 2.1.2). The information in this matrix is convolved with the prior distribution of the fluxes to deduce the maps in 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 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 $20\,\mathrm{Tg\,CH_4\,yr^{-1}}$ according to EDGAR FT2010). But the constraints are lower for this region than for the Siberian Lowlands with smaller emissions ($8\,\mathrm{Tg\,CH_4\,yr^{-1}}$ according to EDGAR FT2010). The post-processing excluding the regions that are mis-separated from the LBC by the inversion accentuates this pattern. In the left panel, the constraints on the regions close to the side of the domain are significantly reduced. The observation sites at the edge of the network also decrease their influence on the inversion during the post-processing. The observation sites within the denser part of the network seem to constrain emissions within a radius of roughly 500 km.

To summarize, the inversion approach that we use 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 measurements are difficult to assimilate, (ii) anthropogenic emission hot spots require observation sites remote from them to be inverted, (iii) diffuse wetland emissions can use close sites to be constrained, and (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.

In the following, we develop the methane fluxes resulting from this inversion system that are explicitly constrained and not mis-separated from the lateral boundary conditions.

3.2 Results of the marginalized inversion

Figure 4 synthetizes the inverted methane fluxes for Siberian Lowlands (hatched domain in Fig. 1) following different steps. First, analyzing the posterior matrix of covariance defined in Sect. 2.1, we flag out the regions that are not separated from the LBC. Then, amongst the remaining components of the state space, we group the regions, the signal of which cannot be separated from each other. This grouping tends to associate neighbouring regions and co-located sources, but also emissions from the same region but from successive periods. As anthropogenic emissions (inverted at a monthly scale) can also be grouped with wetland emissions (considered at the weekly scale), the lowest common multiple on which the fluxes can be analyzed is the monthly scale. At the end, out of the 275 constrained aggregated regions, only 166 groups of regions are kept: 35 anthropogenic groups, 101 wetland groups and 30 mixed groups. For the Siberian Lowlands, for each month, we define the proportion of the prior fluxes that are within regions constrained through the inversion (written on the left of each pie chart in Fig. 4). 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. 4. Finally, within the proportion of constrained regions, we analyze the inversion correction on anthropogenic, wetland

and mixed emissions (bar plots in Fig. 4). For each type of emission (anthropogenic, wetland and mixed), we present in the bar diagrams of Fig. 4 the tolerance intervals of posterior fluxes and the prior uncertainties as calculated by the Maximum of Likelihood algorithm (that is to say $\mathbf{B}_{max}$; see Sect. 2.1.1). For each month, the tolerance intervals on the total prior and posterior methane budget in the Siberian Lowlands is also given in $Tg\,CH_4$.

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, 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 seems complicated.

Concerning the value of the fluxes, in almost all cases, the prior fluxes are 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 real fluxes. The main output of the marginalized inversion compared 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 of Likelihood. In most cases, the Maximum of 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 provides accurate estimates of the fluxes at the regional scale.

In addition to separating the types of emissions, one of the objective was to consistently constrain the methane budget of the Siberian Lowlands. 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 figure is critical in January and November when only 0–9 % of the emissions are constrained. In Sect. 3.1.1, 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 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. Extrapolations are necessary to infer regional emissions from the partial vision given by Fig. 4. Additional observation sites would be needed for a complete resolution of the regional fluxes.

3.3 Siberian Lowland CH_4 budget

The marginalized inversions provide 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 issue. Nevertheless, we are inquiring into the regional budget of methane. Classical inversion systems keep the prior emissions in unseen regions. Here, for each month of the year 2010, we decide to extrapolate the corrections applied in Sect. 3.2 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. 5. 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. 5 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 wetlands. 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 the annual reports of GASPROM, 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. Nevertheless, observations 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.

A peak of anthropogenic emissions with large uncertainties occurs in August. The large tolerance interval is compatible with a smooth scenario. However, extracting companies make use of the decrease in the household demand in summer to carry out maintenance operations on the oil and gas welling sites. These operations are related to punctual leaks and purging releases of gas. This could explain an increase of emissions in August. August is also a month with numerous forest fires. As said in Sect. 2.2.3, wildfires are systematically excluded from the marginalized inversion because they generate steep plumes difficult to take into account. The inversion system could have failed in eliminating observations acquired within some wildfire plumes. The increase in the mixing ratios could then have been wrongly attributed to an increase in anthropogenic emissions.

Overall, on a yearly basis, posterior anthropogenic and wetland emissions are roughly equal, as in the prior state vector. The tolerance interval is 1–13 Tg CH_4 yr^{-1} for wetland emissions, and 6–16 Tg CH_4 yr^{-1} for anthropogenic emissions. The ranges of uncertainties are reduced by 40 % for wetland emissions and 57 % for anthropogenic emissions. Siberian oil and gas extracting activity and population centers 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 not negligible impact on the total European emissions of 27 Tg CH_4. 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.

4 Comparison with GOSAT total columns

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. 3.1.2, 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.

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. 2.2.5) and compare them to their simulated equivalents. In Fig. 6, we compare the prior 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 25 000 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. We observe a mean shift of 1.6 ± 3.9 ppb from the prior to posterior simulated CH_4 columns in the direction of an improvement. However, this shift is not significant compared with the ~ 10 ppb of uncertainties in the GOSAT retrievals.

The GOSAT zonal coverage is rather uniform. GOSAT columns thus carry information on the spatial gradients of the atmospheric methane composition. As analyzing 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. 7, 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 ~ 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.

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. 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 1-D or 2-D 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 teledetection (e.g., the joint French–German cooperation Methane Remote Sensing Lidar Mission, MERLIN) providing high resolution accurate 1-D or 2-D products could be used in regional inversions.

5 Conclusions

We assimilated the data collected in 2010 at 8 surface observation sites measuring atmospheric CH_4 mixing ratios in the West-Siberian plain. It was the first time all these observations were used in a single study. 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. A direct consequence of this inability to assimilate high-frequency data is the relatively high posterior uncertainties in the optimized fluxes. Even in the regions close to the stations, the posterior uncertainties (objectively quantified) 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 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 Tg CH_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 developements of the inversion system. This research was supported by the Commissariat à l'Énergie Atomique et aux Énergies Renouvelables, by the ANR-CLIMSLIP and the YAK-AEROSIB project.

References

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, Boulder, Colorado, 2009. 14631

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 CO_2 and CH_4 concentrations in the surface atmospheric layer over West Siberia, Atmospheric and Oceanic Optics, 22, 84–93, 2009. 14613

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, 19688, doi:10.3402/tellusb.v65i0.19688, 2013a. 14614

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, 2013b. 14593, 14595, 14600, 14603

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, 2014. 14593, 14595, 14596, 14597, 14606, 14617

Bergamaschi, P., Krol, M., Dentener, F., Vermeulen, A., Meinhardt, F., Graul, R., Ramonet, M., Peters, W., and Dlugokencky, E. J.: Inverse modelling of national and European CH_4 emissions using the atmospheric zoom model TM5, Atmos. Chem. Phys., 5, 2431–2460, doi:10.5194/acp-5-2431-2005, 2005.

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, J. Geophys. Res., 114, D22301, doi:10.1029/2009JD012287, 2009. 14590

Bergamaschi, P., Krol, M., Meirink, J. F., Dentener, F., Segers, A., van Aardenne, J., Monni, S., Vermeulen, A. T., Schmidt, M., Ramonet, M., Yver, C., Meinhardt, F., Nisbet, E. G., Fisher, R. E., O'Doherty, S., and Dlugokencky, E. J.: Inverse modeling of European CH_4 emissions 2001–2006, J. Geophys. Res., 115, D22309, doi:10.1029/2010JD014180, 2010.

Bergamaschi, P., Houweling, S., Segers, A., Krol, M., Frankenberg, C., Scheepmaker, R. A., Dlugokencky, E., Wofsy, S. C., Kort, E. A., Sweeney, C., Schuck, T., Brenninkmeijer, C., Chen, H., Beck, V., and Gerbig, C.: Atmospheric CH_4 in the first decade of the 21st century: inverse modeling analysis using SCIAMACHY satellite retrievals and NOAA surface measurements, J. Geophys. Res.-Atmos., 118, 7350–7369, doi:10.1002/jgrd.50480, 2013. 14591

Bocquet, M., Wu, L., and Chevallier, F.: Bayesian design of control space for optimal assimilation of observations. Part I: Consistent multiscale formalism, Q. J. Roy. Meteor. Soc., 137, 1340–1356, 2011.

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, 045015, doi:10.1088/1748-9326/2/4/045015, 2007. 14591

Bousquet, P., Peylin, P., Ciais, P., Ramonet, M., and Monfray, P.: Inverse modeling of annual atmospheric CO_2 sources and sinks: 2. Sensitivity study, J. Geophys. Res., 104, 26179–26193, 1999.

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, 2006. 14590

Bousquet, P., Ringeval, B., Pison, I., Dlugokencky, E. J., Brunke, E.-G., Carouge, C., Chevallier, F., Fortems-Cheiney, A., Frankenberg, C., Hauglustaine, D. A., Krummel, P. B., Langenfelds, R. L., Ramonet, M., Schmidt, M., Steele, L. P., Szopa, S., Yver, C., Viovy, N., and Ciais, P.: Source attribution of the changes in atmospheric methane for 2006–2008, Atmos. Chem. Phys., 11, 3689–3700, doi:10.5194/acp-11-3689-2011, 2011. 14602

Brodzik, M. J., Billingsley, B., Haran, T., Raup, B., and Savoie, M. H.: EASE-Grid 2.0: Incremental but significant improvements for earth-gridded data sets, ISPRS, International Journal of Geo-Information, 1, 32–45, 2012.

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, J. Geophys. Res., 116, D23303, doi:10.1029/2011JD016202, 2011. 14603

Cardinali, C., Pezzulli, S., and Andersson, E.: Influence-matrix diagnostic of a data assimilation system, Q. J. Roy. Meteor. Soc., 130, 2767–2786, 2004. 14597, 14607

Carouge, C., Rayner, P. J., Peylin, P., Bousquet, P., Chevallier, F., and Ciais, P.: What can we learn from European continuous atmospheric CO_2 measurements to quantify regional fluxes – Part 2: Sensitivity of flux accuracy to inverse setup, Atmos. Chem. Phys., 10, 3119–3129, doi:10.5194/acp-10-3119-2010, 2010.

Chapnik, B., Desroziers, G., Rabier, F., and Talagrand, O.: Properties and first application of an error-statistics tuning method in variational assimilation, Q. J. Roy. Meteor. Soc., 130, 2253–2275, 2004. 14593

Chatterjee, A. and Michalak, A. M.: Technical Note: Comparison of ensemble Kalman filter and variational approaches for CO_2 data assimilation, Atmos. Chem. Phys., 13, 11643–11660, doi:10.5194/acp-13-11643-2013, 2013.

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, Geophys. Res. Lett., 37, L21803, doi:10.1029/2010GL044652, 2010. 14605

Chorin, A. J. and Morzfeld, M.: Conditions for successful data assimilation, J. Geophys. Res.-Atmos., 118, 11522–11533, 2013.

Courtier, P., Thèpaut, J.-N., and Hollingsworth, A.: A strategy for operational implementation of 4D-Var, using an incremental approach, Q. J. Roy. Meteor. Soc., 120, 1367–1387, 1994.

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, 2014. 14595

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, 2008. 14599

Dedikov, J., Akopova, G., Gladkaja, N., Piotrovskij, A., Markellov, V., Salichov, S., Kaesler, H., Ramm, A., Müller von Blumencron, A., and Lelieveld, J.: Estimating methane releases from natural gas production and transmission in Russia, Atmos. Environ., 33, 3291–3299, 1999. 14591

Denman, K. L., Brasseur, G., Chidthaisong, A., Ciais, P., Cox, P. M., Dickinson, R. E., Hauglustaine, D., Heinze, C., Holland, E., and Jacob, D.: Couplings between changes in the climate system and biogeochemistry, in: Climate Change 2007, Cambridge University Press, Cambridge, 541–584, 2007.

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, 2003. 14603

Dlugokencky, E. J., Steele, L. P., Lang, P. M., and Masarie, K. A.: The growth rate and distribution of atmospheric methane, J. Geophys. Res.-Atmos., 99, 17021–17043, 1994. 14599

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, D18306, doi:10.1029/2005JD006035, 2005. 14599

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, L18803, doi:10.1029/2009GL039780, 2009. 14590

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, 2012. 14592

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, in: Climate Change 2007, Cambridge University Press, Cambridge, 129–234, 2007. 14590

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 Biogeochem. Cy., 21, GB1016, doi:10.1029/2006GB002706, 2007. 14591

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, 2014. 14593

Giglio, L., Descloitres, J., Justice, C. O., and Kaufman, Y. J.: An enhanced contextual fire detection algorithm for MODIS, Remote Sens. Environ., 87, 273–282, 2003.

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 Sens. Environ., 113, 408–420, 2009. 14602

Glagolev, M., Kleptsova, I., Filippov, I., Maksyutov, S., and Machida, T.: Regional methane emission from West Siberia mire landscapes, Environ. Res. Lett., 6, 045214, doi:10.1088/1748-9326/6/4/045214, 2011. 14591

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, 2002. 14590, 14611

Howarth, R. W., Shindell, D., Santoro, R., Phillips, N., and Townsend-Small, A.: Methane Emissions from Natural Gas Systems, Background paper prepared for the National Climate Assessment, Office of Science & Technology Policy Assessment, Washington, DC, 2012.

Kaminski, T., Rayner, P. J., Heimann, M., and Enting, I. G.: On aggregation errors in atmospheric transport inversions, J. Geophys. Res., 105, 4703–4715, 2001.

Kasischke, E. S. and Bruhwiler, L. P.: Emissions of carbon dioxide, carbon monoxide, and methane from boreal forest fires in 1998, J. Geophys. Res.-Atmos., 107, 8146, doi:10.1029/2001JD000461, 2002. 14602

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, Nat. Geosci., 6, 813–823, 2013. 14590

Konovalov, I. B., Beekmann, M., Kuznetsova, I. N., Yurova, A., and Zvyagintsev, A. M.: Atmospheric impacts of the 2010 Russian wildfires: integrating modelling and measurements of an extreme air pollution episode in the Moscow region, Atmos. Chem. Phys., 11, 10031–10056, doi:10.5194/acp-11-10031-2011, 2011.

Koohkan, M. R. and Bocquet, M.: Accounting for representativeness errors in the inversion of atmospheric constituent emissions: application to the retrieval of regional carbon monoxide fluxes, Tellus B, 64, 19047, doi:10.3402/tellusb.v64i0.19047, 2012.

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, 2013. 14616

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, 2011. 14602

Lauvaux, T., Pannekoucke, O., Sarrat, C., Chevallier, F., Ciais, P., Noilhan, J., and Rayner, P. J.: Structure of the transport uncertainty in mesoscale inversions of CO_2 sources and sinks using ensemble model simulations, Biogeosciences, 6, 1089–1102, doi:10.5194/bg-6-1089-2009, 2009. 14593

Lauvaux, T., Schuh, A. E., Uliasz, M., Richardson, S., Miles, N., Andrews, A. E., Sweeney, C., Diaz, L. I., Martins, D., Shepson, P. B., and Davis, K. J.: Constraining the CO_2 budget of the corn belt: exploring uncertainties from the assumptions in a mesoscale inverse system, Atmos. Chem. Phys., 12, 337–354, doi:10.5194/acp-12-337-2012, 2012.

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, A Project on behalf of EON Ruhrgas AG Wuppertal/Mainz, Wuppertal Institute for Climate, Environment, Energy/Max-Planck-Institute for Chemistry, Wuppertal, 2005. 14590

Lehner, B. and Döll, P.: Development and validation of a global database of lakes, reservoirs and wetlands, J. Hydrol., 296, 1–22, 2004. 14631

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, J. Geophys. Res., 108, 4493, doi:10.1029/2002JD003161, 2003. 14601

Locatelli, R., Bousquet, P., Chevallier, F., Fortems-Cheney, A., Szopa, S., Saunois, M., Agusti-Panareda, A., Bergmann, D., Bian, H., Cameron-Smith, P., Chipperfield, M. P., Gloor, E., Houweling, S., Kawa, S. R., Krol, M., Patra, P. K., Prinn, R. G., Rigby, M., Saito, R., and Wilson, C.: Impact of transport model errors on the global and regional methane emissions estimated by inverse modelling, Atmos. Chem. Phys., 13, 9917–9937, doi:10.5194/acp-13-9917-2013, 2013. 14593

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, 2013. 14611

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, 2013. 14603

Michalak, A. M. and Kitanidis, P. K.: A method for the interpolation of nonnegative functions with an application to contaminant load estimation, Stoch. Env. Res. Risk A., 19, 8–23, 2005. 14593, 14595

Michalak, A. M., Hirsch, A., Bruhwiler, L., Gurney, K. R., Peters, W., and Tans, P. P.: Maximum likelihood estimation of covariance parameters for Bayesian atmospheric trace gas surface flux inversions, J. Geophys. Res., 110, D24107, doi:10.1029/2005JD005970, 2005.

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

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, Environ. Sci., 2, 81–99, 2005. 14601

Paris, J.-D., Stohl, A., Nédélec, P., Arshinov, M. Yu., Panchenko, M. V., Shmargunov, V. P., Law, K. S., Belan, B. D., and Ciais, P.: Wildfire smoke in the Siberian Arctic in summer: source characterization and plume evolution from airborne measurements, Atmos. Chem. Phys., 9, 9315–9327, doi:10.5194/acp-9-9315-2009, 2009.

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, 2010. 14614

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, Geophys. Res. Lett., 38, L15807, doi:10.1029/2011GL047871, 2011. 14604

Peregon, A., Maksyutov, S., and Yamagata, Y.: An image-based inventory of the spatial structure of West Siberian wetlands, Environ. Res. Lett., 4, 045014, doi:10.1088/1748-9326/4/4/045014, 2009. 14590, 14591

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, 2011. 14591

Pison, I., Menut, L., and Bergametti, G.: Inverse modeling of surface NO_x anthropogenic emission fluxes in the Paris area during the Air Pollution Over Paris Region (ESQUIF) campaign, J. Geophys. Res., 112, D24302, doi:10.1029/2007JD008871, 2007. 14603

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. Cy., 7, 811–841, 1993. 14602

Prather, M. J., Zhu, X., Strahan, S. E., Steenrod, S. D., and Rodriguez, J. M.: Quantifying errors in trace species transport modeling, P. Natl. Acad. Sci USA, 105, 19617–19621, 2008. 14593

Reichle, R. H., Crow, W. T., and Keppenne, C. L.: An adaptive ensemble Kalman filter for soil moisture data assimilation, Water Resour. Res., 44, W03423, doi:10.1029/2007WR006357, 2008.

Reshetnikov, A. I., Paramonova, N. N., and Shashkov, A. A.: An evaluation of historical methane emissions from the Soviet gas industry, J. Geophys. Res.-Atmos., 105, 3517–3529, 2000. 14591

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 Biogeochem. Cy., 24, GB2003, doi:10.1029/2008GB003354, 2010. 14591

Ringeval, B., Houweling, S., van Bodegom, P. M., Spahni, R., van Beek, R., Joos, F., and Röckmann, T.: Methane emissions from floodplains in the Amazon Basin: challenges in developing a process-based model for global applications, Biogeosciences, 11, 1519–1558, doi:10.5194/bg-11-1519-2014, 2014.

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, 2003. 14597

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, 2010. 14593, 14599

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, 17514, doi:10.3402/tellusb.v64i0.17514, 2012. 14592

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, 2001. 14603

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, 2011. 14590, 14601

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, 2013. 14601

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 Climate Change, 3, 666–672, 2013. 14601

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, 2014. 14601

Stohl, A., Forster, C., Frank, A., Seibert, P., and Wotawa, G.: Technical note: The Lagrangian particle dispersion model FLEXPART version 6.2, Atmos. Chem. Phys., 5, 2461–2474, doi:10.5194/acp-5-2461-2005, 2005. 14600

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, J. Atmos. Ocean. Tech., 27, 1175–1184, 2010. 14599

Tarantola, A.: Inverse Problem Theory, Elsevier, New York, 1987. 14592, 14594

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, Atmos. Environ., 40, 5617–5628, 2006. 14592, 14614

Tolk, L. F., Dolman, A. J., Meesters, A. G. C. A., and Peters, W.: A comparison of different inverse carbon flux estimation approaches for application on a regional domain, Atmos. Chem. Phys., 11, 10349–10365, doi:10.5194/acp-11-10349-2011, 2011.

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 Biogeochem. Cy., 26, GB4009, doi:10.1029/2011GB004232, 2012. 14592

Urbanski, S. P., Hao, W. M., and Nordgren, B.: The wildland fire emission inventory: western United States emission estimates and an evaluation of uncertainty, Atmos. Chem. Phys., 11, 12973–13000, doi:10.5194/acp-11-12973-2011, 2011.

US Environment Protection Agency: Inventory of U.S. greenhouse gas emissions and sinks: 1990–2009, Tech. Rep. 430-R-11-005, US EPA, Washington DC, 2011. 14590

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, 11707–11735, doi:10.5194/acp-10-11707-2010, 2010. 14602

Vautard, R., Beekmann, M., Roux, J., and Gombert, D.: Validation of a hybrid forecasting system for the ozone concentrations over the Paris area, Atmos. Environ., 35, 2449–2461, 2001. 14603

Villani, M. G., Bergamaschi, P., Krol, M., Meirink, J. F., and Dentener, F.: Inverse modeling of European CH_4 emissions: sensitivity to the observational network, Atmos. Chem. Phys., 10, 1249–1267, doi:10.5194/acp-10-1249-2010, 2010. 14599

Winderlich, J.: Setup of a CO_2 and CH_4 measurement system in Central Siberia and modeling of its results, Ph.D. thesis, University of Hamburg, Hamburg, Germany, 2012. 14592, 14613

Winderlich, J., Chen, H., Gerbig, C., Seifert, T., Kolle, O., Lavrič, J. V., Kaiser, C., Höfer, A., and Heimann, M.: Continuous low-maintenance $CO_2/CH_4/H_2O$ measurements at the Zotino Tall Tower Observatory (ZOTTO) in Central Siberia, Atmos. Meas. Tech., 3, 1113–1128, doi:10.5194/amt-3-1113-2010, 2010. 14592, 14593, 14599

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, D05122, doi:10.1029/2011JD016932, 2012. 14593, 14595

Wu, L., Bocquet, M., Lauvaux, T., Chevallier, F., Rayner, P., and Davis, K.: Optimal representation of source-sink fluxes for mesoscale carbon dioxide inversion with synthetic data, J. Geophys. Res, 116, D21304, doi:10.1029/2011JD016198, 2011.

Yamada, K., Yoshida, N., Nakagawa, F., and Inoue, G.: Source evaluation of atmospheric methane over western Siberia using double stable isotopic signatures, Org. Geochem., 36, 717–726, 2005. 14592

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, 10–13 September 2007, Helsinki, Finland, WMO/GAW Report, 186, 40–43, 2009. 14599

Table 1. Sites characteristics. The altitudes of the sites are given as m a.s.l. and the inlet height is in m 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 (°E)	Lat (°N)	Alt (m a.s.l.)	height (m a.g.l.)		Institute
Azovo	AZV	73.03	54.71	100	50	C	JR-STATIONS
Berezorechka	BRZ	84.33	56.15	150	80	C	JR-STATIONS
Demyanskoe	DEM	70.87	59.79	75	63	C	JR-STATIONS
Igrim	IGR	64.42	63.19	25	47	C	JR-STATIONS
Karasevoe	KRS	82.42	58.25	50	67	C	JR-STATIONS
Noyabrsk	NOY	75.78	63.43	100	43	C	JR-STATIONS
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-STATIONS
Yakutsk	YAK	129.36	62.09	210	77	C	JR-STATIONS
Zotino	ZOT	89.35	60.80	104	301	C	MPI

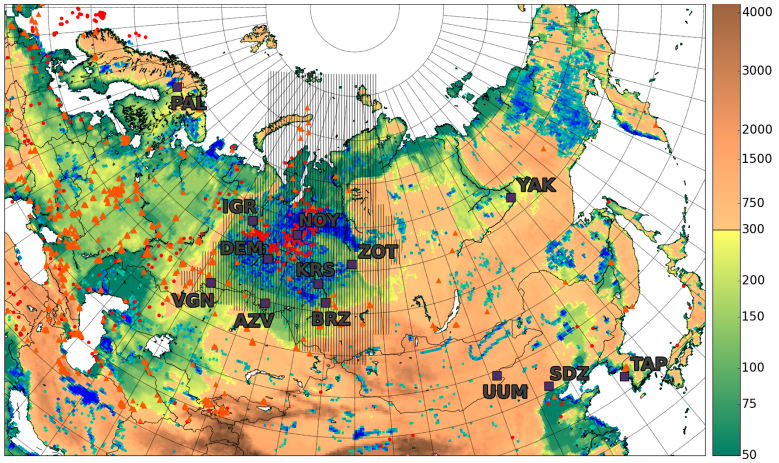

Figure 1. Topographic map of the domain of interest. The colorbar paints the altitude a.s.l. (from ETOPO1 database; Amante and Eakins, 2009). Red dots (resp. orange triangle) depicts hot spots of CH_4 emissions (based on EDGAR v4.2 inventory; see Sect. 2.2.3) related to oil welling and refineries (resp. gas extraction and leaks during distribution in population centers). Purple squares highlight the observation site localization. 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. 3.3 is calculated on the hatched area.

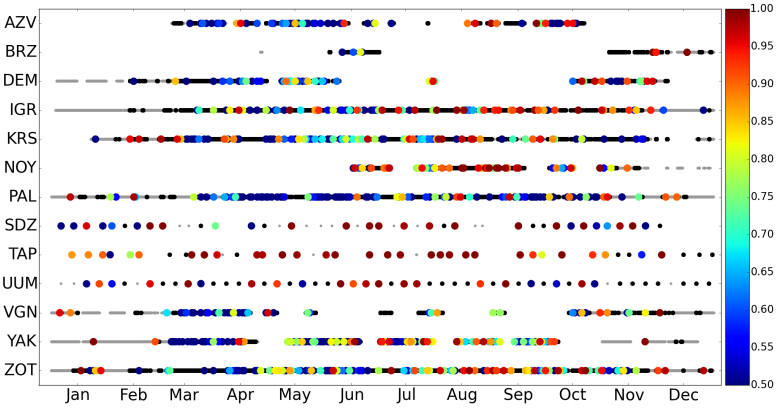

Figure 2. Observation availability and contribution to the inversion for each site. Grey dots stands for available data flagged out because of a planetary boundary layer height smaller than 500 m. Black dots denote the data points that are filtered during the computation of the algorithm described in Sects. 2.1 and 2.2.1. Color circles represent the pieces of information carried out by each selected piece of data (see Sect. 2.1.2).

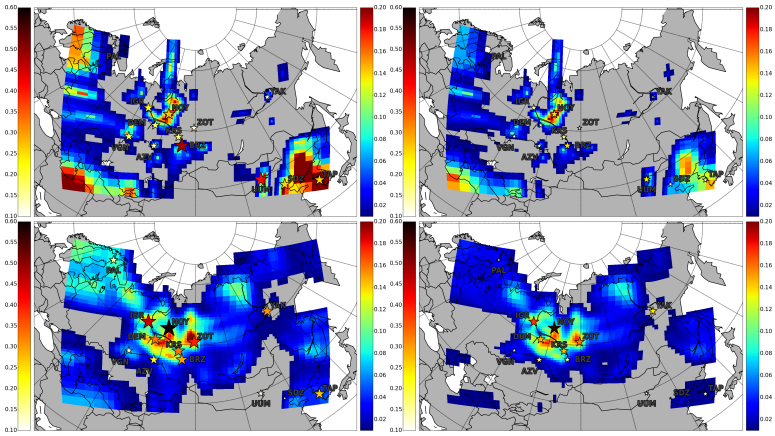

Figure 3. Maps of the constraints to the fluxes from the observations in the inversion. In the top (resp. bottom) panel, the constraints to the anthropogenic (resp. wetland) fluxes are displayed. On the left column, all the emissions are included, while on the right column, are taken into account only the regions that are not mis-separated from the boundary conditions. The average contribution (as defined in Sect. 2.1.2) of each station is represented by the size of the marker and its color (left color bar). The constraints on the fluxes (as defined in Sect. 2.1.2) are displayed according to the right color bar.

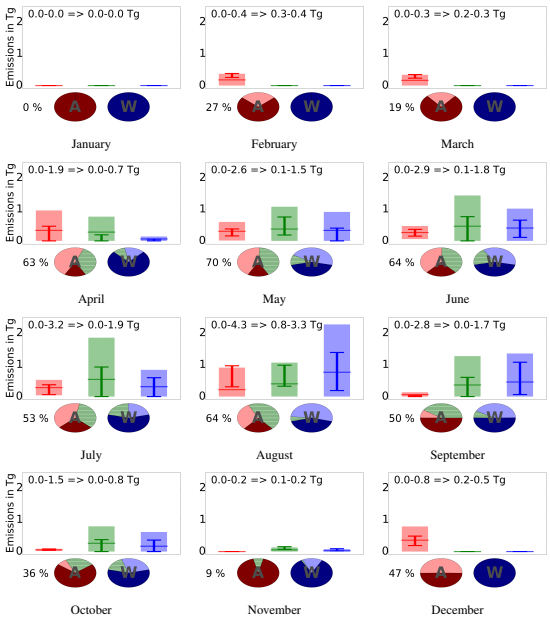

Figure 4. Inversion results for the twelve months of 2010 in the Siberian Lowlands (hatched area in Fig. 1). For each month, The bar plots show the prior (shade) and posterior (cap) tolerance intervals for anthropogenic (left), mixed (middle) and wetland (right) emissions. The pie charts depict the proportion of prior emissions within regions constrained by the inversion: (left, A) anthropogenic, (right, W) wetlands. Dark portions means 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 budget.

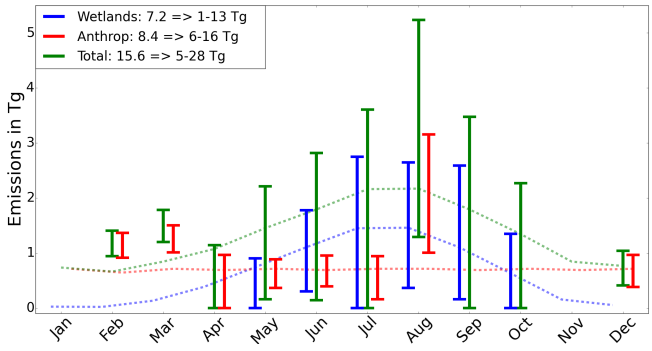

Figure 5. Time series of the posterior emissions on the Siberian Lowlands. The emissions are reported in Tg CH_4 for each month. The correction factors computed on the constrained regions (in Sect. 3.2) are extrapolated to the Siberian Lowlands (see Fig. 1). The extrapolation is carried out only when 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 Tg CH_4 are noted in the legend.

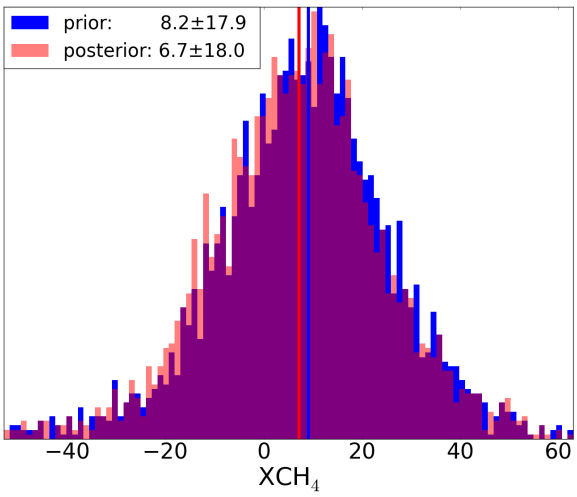

Figure 6. Distribution of the differences between the observed and simulated prior and posterior GOSAT total columns. The differences are computed only for the GOSAT observations downwind the emissions constrained by the surface sites, i.e. with a simulated contribution from these regions to the total column bigger than 2 ppb.

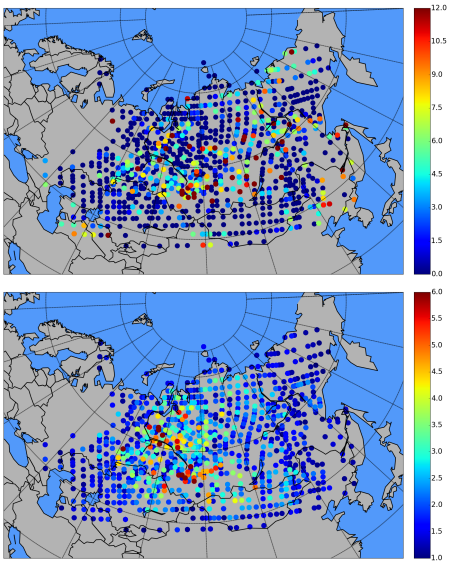

Figure 7. Observed (top) and simulated (bottom) contributions of the local Siberian sources to GOSAT total columns. The observed contribution of local sources are selected as excesses from the regional measured background (defined as the average observed total column in a zone of roughly 2000 km in the 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 color scale is not the same between the two panels in order to facilitate the readibility of the figure.