OPEN ACCESS

University of BRISTOL

Link to published version (if available):
10.5194/acp-11-9887-2011

Link to publication record in Explore Bristol Research
PDF-document

**Atmospheric
Chemistry
and Physics**

Inversion of long-lived trace gas emissions using combined Eulerian and Lagrangian chemical transport models

M. Rigby[1]**, A. J. Manning**[2]**, and R. G. Prinn**[1]

[1]Center for Global Change Science, Massachusetts Institute of Technology, 77 Massachusetts Ave., Cambridge, MA, 02139, UK
[2]Atmospheric Dispersion Group, UK Met. Office, Exeter, EX1 3PB, UK

Abstract. We present a method for estimating emissions of long-lived trace gases from a sparse global network of high-frequency observatories, using both a global Eulerian chemical transport model and Lagrangian particle dispersion model. Emissions are derived in a single step after determining sensitivities of the observations to initial conditions, the high-resolution emissions field close to observation points, and larger regions further from the measurements. This method has the several advantages over inversions using one type of model alone, in that: high-resolution simulations can be carried out in limited domains close to the measurement sites, with lower resolution being used further from them; the influence of errors due to aggregation of emissions close to the measurement sites can be minimized; assumptions about boundary conditions to the Lagrangian model do not need to be made, since the entire emissions field is estimated; any combination of appropriate models can be used, with no code modification. Because the sensitivity to the entire emissions field is derived, the estimation can be carried out using traditional statistical methods without the need for multiple steps in the inversion. We demonstrate the utility of this approach by determining global SF_6 emissions using measurements from the Advanced Global Atmospheric Gases Experiment (AGAGE) between 2007 and 2009. The global total and large-scale patterns of the derived emissions agree well with previous studies, whilst allowing emissions to be determined at higher resolution than has previously been possible, and improving the agreement between the modeled and observed mole fractions at some sites.

1 Introduction

Atmospheric mole fraction measurements provide valuable information on the sources and sinks of long-lived trace gases, provided that an appropriate relationship between the surface flux and the observations can be estimated using chemical transport models (CTMs). For the majority of long-lived greenhouse gases and ozone depleting species, these observations are primarily made at the surface, either at high-frequency in the case of in situ stations, or at weekly to monthly intervals in the case of flask sampling sites. When attempting to extract regional surface flux information, high-frequency observations are particularly powerful. However, such in situ networks currently have a low spatial density, and will likely continue to do so in the near future. Therefore, the challenge for any inverse method that uses this data is to maximize the information that can be extracted from such a network.

This paper will primarily focus on the Advanced Global Atmospheric Gases Experiment (AGAGE, Prinn et al., 2000) and affiliated networks, which measure over 40 trace gases at high-frequency and precision at eleven locations worldwide. However, the methods developed will be applicable to any global, high-frequency network measuring long-lived species (e.g. the National Oceanic and Atmospheric Administration, in situ program, Geller et al., 1997).

Sources of many gases have been derived using the AGAGE network with two types of CTM: Eulerian models in which atmospheric transport and chemistry is calculated relative to some coordinate system (e.g. a fixed 3-D grid), or Lagrangian models, where diffusion and chemistry are calculated from the perspective of air parcels that are carried by the wind.

Global Eulerian models have been used extensively for the inversion of global emissions and emissions from large regions using AGAGE and other data (e.g. Gurney et al., 2002; Chen and Prinn, 2006; Rigby et al., 2008, 2010). Source-receptor relationships in these inversions are derived using perturbed emissions from large regions of the globe and multiple forward model runs. This approach is simple to implement, but the computational cost prevents disaggregation of the global emissions field into a large number of emissions elements (e.g. grid-scale elements). A common problem is therefore "aggregation error", in which spurious optimized emissions may be produced by an incorrect spatial distribution within each bulk region. To reduce the influence of aggregation errors in an Eulerian modeling framework, two approaches have been used: coding of an adjoint of the CTM, or the approximation of the a posteriori solution and covariance with the use of ensembles of randomly perturbed emissions fields. The adjoint method allows a cost function to be minimized through the production of computer code that directly determines the sensitivity of observations to the emissions field (e.g. Kaminski et al., 1999; Meirink et al., 2008; Kopacz et al., 2010). However, for most CTMs an adjoint is not available and represents a significant technical challenge to develop. The ensemble method, in contrast, is relatively simple to implement as no code modification is required, instead source-receptor relationships are approximated using multiple random realizations of the emissions field (e.g. Peters et al., 2005). A limitation of the ensemble approach is that, in general, a large number of perturbed fields must be propagated through the CTM in order to prevent significant sampling uncertainties in the derived sensitivities. One problem common to most inversions using global Eulerian CTMs is that very high resolution simulations can become highly computationally expensive. It is particularly desirable to have increased resolution close to a measurement site, where small transport uncertainties can have a large impact on the derived emissions field. To address this problem, some Eulerian CTMs now have a "zoom" capability in which high-resolution regional grids are nested within a coarse global grid (e.g. Krol et al., 2005). However, such schemes have only been developed for a small subset of CTMs at present.

Lagrangian particle dispersion models (LPDMs) calculate the sensitivity of measurements to the surrounding emissions field by tracking a large number of "particles" backwards in time for a limited number of days and determining where they intercept the surface layer. The sensitivity of each observation to the emissions field is often referred to as the "footprint" of the measurement. Regional emissions have been determined using AGAGE measurements and several LPDMs (e.g. O'Doherty et al., 2004; Stohl et al., 2009; Manning et al., 2011). Whilst such regional approaches largely avoid aggregation errors (since sensitivities to each grid cell surrounding a measurement are directly calculated), a limitation is that boundary conditions must be estimated to account for emissions into the atmosphere farther back in time

than the temporal extent of the back-trajectories. This is usually achieved by identifying observations that represent background air and interpolating between them (e.g. Manning et al., 2011), or by statistically estimating some offset that must be applied to the observations over some time period (e.g. Stohl et al., 2009; Vollmer et al., 2009). A limitation of the first approach is that background mole fractions that are lower than the observed values cannot be identified, whilst one problem with the second is that constant background levels must be assumed over some time period. One of the major strengths of LPDMs is that the calculated footprints can be used for any gas, provided that its lifetime is much longer than the back-trajectory timescale.

In this paper, we present a method whereby outputs from Lagrangian and Eulerian CTMs can be coupled in order to make use of the strongest aspects of both for simultaneous global and regional emissions estimation. This will be achieved by combining sensitivities calculated using both models. Previous studies have coupled these two types of model in the time-domain to simulate mole fractions using high-resolution emissions fields (e.g. Koyama et al., 2011). The approach that we present here has been specifically developed to address the problem of extracting global and regional emissions information from a spatially-sparse, high-frequency monitoring network such as AGAGE. The advantages of our method are that:

– No code modification of either CTM is required, making it applicable to any combination of global Eulerian model and LPDM

– Emissions from individual grid cells can be derived close to the high-frequency monitoring sites (using the LPDM) and from large regions further from them (using the Eulerian model), reducing aggregation errors at minimal computation expense

– High-resolution simulations can be performed over limited regions close to the monitoring sites, and at coarse resolution for the rest of the globe, minimizing the computational load of such a high resolution inversion (as was identified by Roedenbeck et al., 2009)

– Once run, the LPDM output can be applied to any long-lived gas

– Time varying boundary conditions to the LPDM are implicitly estimated in the scheme

We outline a method to achieve a global inversion with regional high-resolution emissions estimates in a single step. This is presented as an alternative to the recent work by Roedenbeck et al. (2009), who outlined an inverse method in which a course global inversion was first solved, and then a second high-resolution inversion was performed in one region using the optimized emissions from the first. Our one-step approach has the advantage that covariances can be
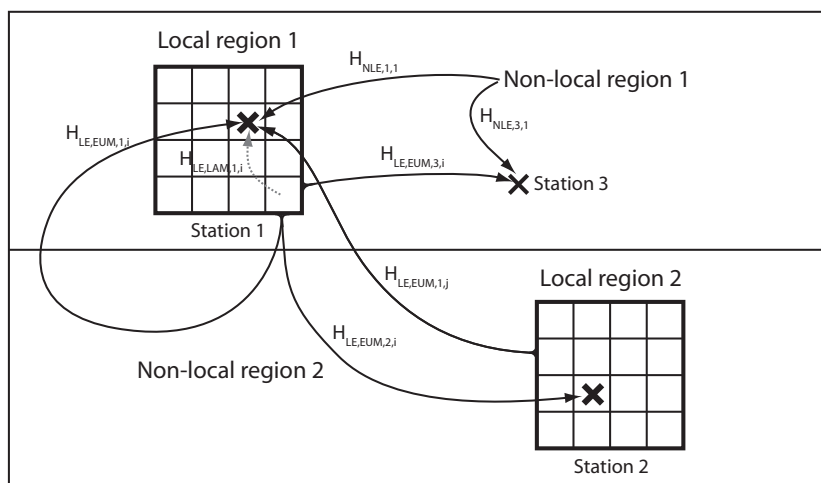
Fig. 1. A schematic of emissions regions and sensitivities. Two large "non-local" regions are shown from which bulk emissions are estimated, and two "local" regions close to monitoring sites, within which small-scale emissions can be estimated. The arrows show examples of the sensitivity components, with each arrow originating from an emissions element (cell i in local region 1 and j in local region 2) and ending at a measurement site. The grey dashed arrow indicates the sensitivity component calculated by the LPDM and the solid black arrows show the sensitivity components calculated by the Eulerian model.

easily retained between emissions within the high-resolution regions and those outside, and that independence between prior information and the observations can be ensured by only using the observations once. Further, we outline a method by which we can explicitly track sensitivities of *all* observations to emissions from multiple "high-resolution regions".

2 A one-step inverse method using a particle dispersion model and global Eulerian model

For most long-lived trace gases (with lifetimes of years or longer), the assumption that atmospheric mole fractions respond linearly to changes in emissions holds extremely well at current global emission rates. By using this linearity, we can relate a vector of observations (y) to a vector of parameters, which here consists of emissions and initial conditions (x), using the equation (e.g. Tarantola, 2005):

$$y = \mathbf{H}x + \text{residual} \tag{1}$$

Here $\mathbf{H}$ is a matrix of sensitivities of the observations to changes in emissions or initial conditions (a "Jacobian" matrix of partial derivatives). $\mathbf{H}$ is to be estimated using the CTMs. We further exploit the linearity of the system by decomposing the sensitivity matrix and parameters vector into three components:

$$\mathbf{H} = (\mathbf{H}_{\text{IC}}, \mathbf{H}_{\text{NLE}}, \mathbf{H}_{\text{LE}}) \tag{2}$$

$$x = (x_{\text{IC}}, x_{\text{NLE}}, x_{\text{LE}})^T \tag{3}$$

where $\mathbf{H}_{\text{IC}}$ is the sensitivity of all the observations to some initial conditions, $\mathbf{H}_{\text{NLE}}$ is the sensitivity of all the observations to bulk emissions regions further from the monitoring sites, and $\mathbf{H}_{\text{LE}}$ is the sensitivity of all the observations to each emissions element (or small aggregated regions) close to the monitoring sites. The distinction between the non-local and local emissions is shown schematically in Fig. 1. We will estimate $\mathbf{H}_{\text{IC}}$ and $\mathbf{H}_{\text{NLE}}$ using a global Eulerian CTM (see below). As noted by Roedenbeck et al. (2009), the sensitivity of the observations to the local emissions field must be further decomposed into two parts:

$$\mathbf{H}_{\text{LE}} = \mathbf{H}_{\text{LE,LAM}} + \mathbf{H}_{\text{LE,EUM}} \tag{4}$$

The term $\mathbf{H}_{\text{LE,LAM}}$ contains the sensitivity of the observations at a particular site to emissions immediately surrounding the monitoring site as calculated by the LPDM (or Lagrangian model, which we refer to as LAM in the equations). However, since the LPDM trajectories are finite in time, we must account for both the long-term fate of the emissions as they are mixed into the global background, and the impact of the emissions from one LPDM region on observations at another. These factors are contained in the matrix $\mathbf{H}_{\text{LE,EUM}}$ and will be estimated using the Eulerian CTM (EUM).

Methods for calculating each term in the sensitivity matrix are discussed in the subsections that follow. Once the combined sensitivity matrix ($\mathbf{H}$) has been derived, Eq. (1) can be solved in a variety of ways (for details see Khasibatla et al., 2000; Enting, 2002; Tarantola, 2005; Wunsch, 2006, and Sect. 3 of this paper).

2.1　Sensitivity to initial conditions (H_{IC})

In the above scheme the initial mole fraction in the atmosphere must be accounted for, since a reference run of one model cannot be performed (as with inversions involving only one model, e.g. Chen and Prinn, 2006; Rigby et al., 2008, 2010, and others). It is unlikely that the entire 3-D initial mole fraction field can be estimated for every model grid cell in the inversion (both because of the lack of an adjoint in general, and because of the highly under-determined nature of such a problem). Therefore some method must be used to account for the initial mole fractions at each measurement site. In the example below, in which trace gas emissions for the period 2007 to 2009 were determined, the initial conditions were estimated by including four terms in the parameters vector x_{IC} and four columns in the matrix H_{IC}: one describing the sensitivity of the observations to a uniform atmospheric mole fraction at the beginning of 2004 and three for global total emissions in 2004, 2005 and 2006. The sensitivity of all of the observations to each of these terms was estimated, along with their influence on observations between 2004 and 2006. A three-year period was chosen to allow realistic inter-hemispheric and stratospheric-tropospheric gradients to be set up before the start of 2007. Sensitivities were estimated by perturbing each quantity and tracking the resulting mole fractions with the Eulerian model.

2.2　Sensitivity to non-local emissions (H_{NLE})

The term "non-local emissions" used here refers to estimates of emissions from large areas outside of the regions for which the Lagrangian model is used to estimate sensitivities (see Fig. 1). Sensitivities of the observations to emissions from these regions can easily be estimated by perturbing emissions from each region, relative to some reference emissions field, and tracking the modeled mole fractions at the monitoring locations (c.f. Chen and Prinn, 2006; Rigby et al., 2010). An example of the sensitivity of measurements at Gosan, South Korea to several non-local emissions regions is shown in Fig. 2a. Sensitivities to emissions from each continent (North America, South America, Europe, Asia, Africa, Oceania) are shown in the figure, excluding emissions from four local regions used in Sect. 3.

2.3　Short-timescale sensitivity to local emissions ($H_{LE,LAM}$)

Short-timescale sensitivity refers to the impact on the measurements of emissions close to monitoring sites over timescales of the order of days (e.g. "pollution events"). This component is directly estimated by the LPDM (see Fig. 2b).
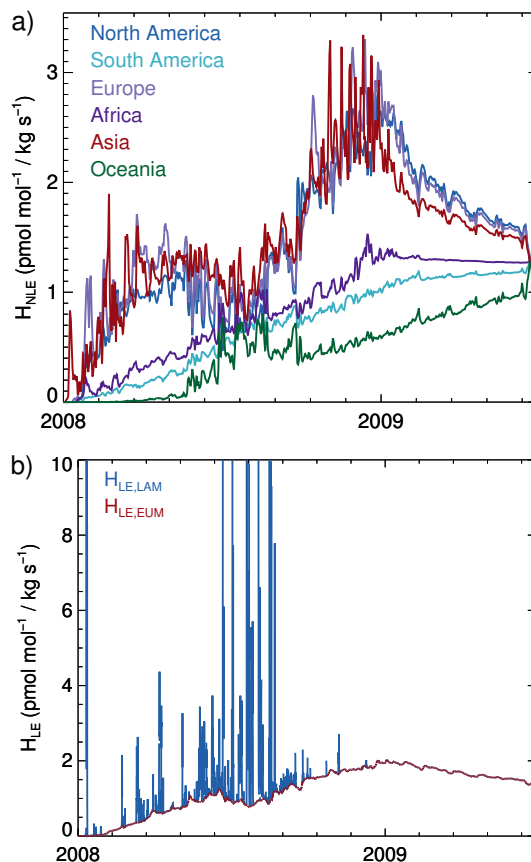

Fig. 2. The sensitivity of SF_6 observations at Gosan, South Korea to: **(a)** perturbations in 2008 emissions from the six non-local regions used in the SF_6 inversion in Sect. 3; **(a)** changes in 2008 emissions from one grid cell close to the monitoring site. In (b) the red line shows the contribution that this emissions element makes to the background mole fraction, as calculated by the Eulerian model ($H_{LE,EUM}$), whilst the blue line shows sensitivities calculated by the LPDM ($H_{LE,LAM}$). The blue line has been added to the red line to show the overall sensitivity due to this grid cell (Equation 4).

2.4　Long-timescale sensitivity to local emissions ($H_{LE,EUM}$)

Given the finite length of time that back-trajectories can be realistically calculated in the LPDM, the long-term fate of emissions originating from regions close to the monitoring sites must be estimated. This can be achieved using the Eulerian model. There are two types of long-timescale sensitivity that must be estimated for emissions close to a monitoring site: the impact of a change in emissions close to one site on the mole fractions measured at another, and the long-term impact of emissions close to a site on that site.

As shown in Fig. 1, imagine we have two measurement stations (Station 1 and 2), surrounded by two "local" regions (Regions 1 and 2). To estimate the sensitivity of mole fractions at Station 2 due to a change in emissions close to
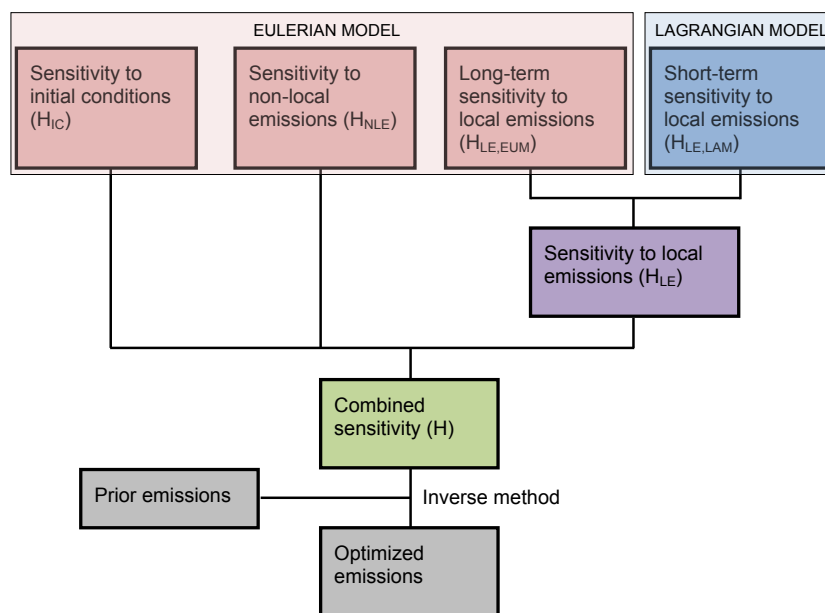
Fig. 3. Flow chart outlining the estimation procedure.

Station 1, the Eulerian model can be run with perturbed emissions for the whole of Region 1 (provided the sites are not too close together). The sensitivity can be estimated from the difference between the reference and perturbed runs. If the sites are close together, however, aggregation errors may result, in which case the perturbation to emissions Region 1 may need to be sub-divided, or if the sites are very close, the local region could be extended to incorporate both stations. If the region is very large, the perturbation may also need to be sub-divided, so that changes to the background mole fraction due to emissions from different parts of the local region are more accurately modeled.

The more difficult sensitivity to estimate is the long-term impact of near-field emissions on the local measurement site (e.g. the contribution of emissions from near Region 1 to the background mole fraction measured at Station 1). This is because an Eulerian model run with a perturbed local emissions field will contain both the change in the background mole fraction that is desired, along with short-term fluctuations that are already being accounted for by the LPDM. The challenge is therefore to retain the long-timescale part of the perturbed local emissions signal, and remove the short-term "pollution events". The way that this may be achieved in the most efficient way will be dependent on model architecture. The method we will use in Sect. 3 is to run two sets of perturbed emissions for each local region. One set of perturbed emissions propagates through the model as usual (the same runs as were used to estimate inter-site sensitivity), and a second exhibits the usual transport and chemistry until it leaves the local region at which point it encounters a very powerful reactant that destroys the gas. This second run will therefore

only contain the short-timescale fluctuations. In practice, the domain in which the pollutant is free from the influence of the "reactant" should be larger than the local region, to allow for instances where the LPDM has tracked air parcels that leave the local region and re-enter after some short time period (hours-days). Once obtained, the second run can be subtracted from the first, in order to retain the longer-timescale variations brought about by a change in emissions from the local regions. An example of this term is shown in Fig. 2b.

Figure 2 illustrates the value of using the Eulerian model to account for background fluctuations at a site like Gosan, South Korea, in which the background mole fraction can change rapidly, due to the occasional intrusion of southern hemispheric air (these changes are typically seen during the summer months). The assumption of constant background mole fractions during some period, or the identification of background measurements, may be difficult during such rapidly-fluctuating large-scale transport events (e.g. Stohl et al., 2009; Manning et al., 2011).

2.5 Summary of sensitivity estimation procedure

Here we will briefly summarize the sensitivity estimation procedure, explaining the required model runs. The procedure is also outlined in Fig. 3.

1. Obtain LPDM footprints of sufficient back-trajectory simulation time to more than cover the local emissions field close to each monitoring site

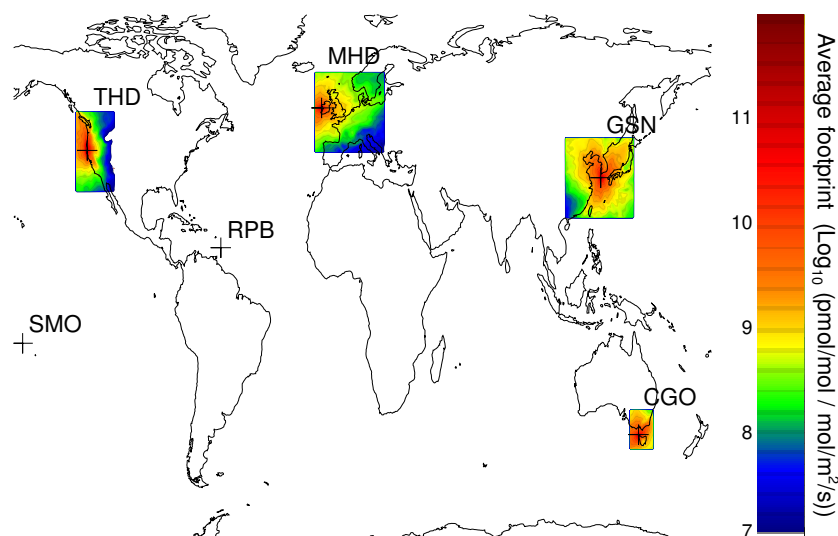2. Run the Eulerian model with perturbed initial conditions

Fig. 4. Average footprints at four AGAGE sites (THD=Trinidad Head, California, MHD = Mace Head, Ireland, GSN = Gosan, South Korea, CGO = Cape Grim, Tasmania) for the year 2008. Also shown are the locations of two sites (RPB = Ragged Point, Barbados, SMO = Cape Matatula, Samoa) used to constrain background emissions. The boxes show the extent of the "local" regions in which the LPDM was used to estimate sensitivities.

3. Run the Eulerian model using a reference emission field and perturbed emissions from:

   – Each non-local region

   – The whole (or some sub-divisions) of each local region with realistic chemistry

   – The whole (or some sub-divisions) or each local region with realistic chemistry within the local region, and a powerful reactant outside the region

3   Application: estimation of global SF$_6$ emissions

Using the scheme outlined above, we derived global emission rates of sulfur hexafluoride (SF$_6$) for each year between 2007 and 2009 using 6 sites from the AGAGE network. SF$_6$, which is a powerful and long-lived greenhouse gas, has been the focus of several recent papers, since its emission rate has dramatically increased since 2001 (Rigby et al., 2010; Levin et al., 2010). Rigby et al. (2010) derived regional emissions from 2004 to 2008 using AGAGE and NOAA measurements. However, they were only able to constrain emissions from continent-scale regions. They also noted that, because of this inversion setup and the relatively coarse-resolution global Eulerian model used, aggregation- and model resolution-errors were likely to be leading to poorly modeled mole fractions at some measurement sites such as Gosan, South Korea.

The sensitivities of SF$_6$ observations to initial conditions and emission rates were estimated using the Eulerian Model for Ozone and Related Tracers (MOZART v4.5, Emmons et al., 2010) and the Lagrangian Numerical Atmospheric

Modelling Environment (NAME v3, Ryall et al., 1998; Manning et al., 2011). Prior estimates of global emissions were taken from the Emissions Database for Global Atmospheric Research (EDGAR v4.0, JRC/PBL, 2009). MOZART was run at approximately $2.8° \times 2.8°$ resolution, and NAME at $0.38° \times 0.56°$ resolution. 36 000 particles were tracked in the NAME model for 13-days prior to 3-h-average measurement periods.

Six AGAGE stations were used in the inversion: Mace Head, Ireland (MHD), Trinidad Head, California (THD), Gosan, South Korea (GSN), Cape Grim, Tasmania (CGO), Ragged Point, Barbados (RPB) and Cape Matatula, American Samoa (SMO). The 2008-average footprints for the first four stations are shown in Fig. 4. Emissions were derived at high-resolution inside the boxes shown in the figure. The size of these regions was chosen based on the average footprint and the extent of significant emissions, as predicted by EDGAR. High-resolution emissions were not estimated around the RPB and SMO sites, since they do not regularly intercept polluted air (with respect to SF$_6$). These sites were included in the inversion to help constrain the global background.

Sensitivities to initial conditions ($\mathbf{H}_{IC}$) were estimated using MOZART. As described in Sect. 2.1, we estimated the sensitivity to a change in a uniform mole fraction field at the start of 2004, and to annual emissions from 2004 to 2006. Monthly-average mole fractions at each AGAGE site between 2004 and 2006 were included in the inversion to constrain emissions in these years (however, the derived emissions between 2004 and 2006 were assumed to have little physical meaning, and are therefore not presented below.

They serve only to set up a realistic mole fraction field at the start of 2007).

The sensitivity of the measurements to changes in emissions from the six non-local emission regions, excluding the areas within the boxes in Fig. 4, were estimated using the MOZART model ($\mathbf{H}_{NLE}$). This was achieved by increasing the EDGAR emissions by 100% in these regions and tracking the resultant change in mole fraction at the AGAGE stations, compared to a reference run (see Fig. 2a). Six continental regions were used: North America, South America, Europe, Africa, Asia and Oceania.

To estimate the sensitivity of the measurements to local emissions, both models were required. The Lagrangian model term $\mathbf{H}_{LE,LAM}$ was obtained directly from the NAME model footprints, which were calculated every 3 h. The term ($\mathbf{H}_{LE,EUM}$) was estimated using two perturbed runs of the MOZART model for each local region as described in Sect. 2.4. The first of these perturbed runs approximated the sensitivity of every measurement to a 1 kg change in emissions originating somewhere within each of the "local" regions. The second, in which the SF_6 was destroyed upon leaving a region 5-degrees larger on each side than the local regions, served only to remove the pollution events from the sensitivity to local emissions calculated using the Eulerian model (since we only wanted to model the pollution events using the LPDM). The total sensitivity of the observations to emissions at each grid cell within the local regions was calculated as the sum of these short- and long-timescale components (Eq. 4). Figure 2b shows an example of the derived sensitivities.

In order to reduce the computational load of the inversion, sensitivities to grid-scale emissions within the LPDM regions were aggregated into a smaller set of emissions regions. Grid cells were aggregated into these sub-regions by considering: (a) the product of the average footprint and the EDGAR emissions; (b) the country that they resided in. The spatial distribution of emissions within each sub-region was obtained also from EDGAR. Details of the grid cell aggregation scheme are outlined in the Supplement. Emissions from approximately 100 local sub-regions were estimated in the inversion.

Measurements at the six AGAGE sites were used to constrain emissions from the aggregated local sub-regions and background emissions regions. AGAGE measures SF_6 at approximately 2-hourly intervals. However, since we would expect that model uncertainties might be correlated with a timescale similar to synoptic variability, measurements were averaged into 5-day periods. These 5-day average observations were assumed to be independent in the inversion, and no uncertainty correlation between the sites was assumed. The influence of choosing an alternative averaging period was explored in the emissions uncertainty calculation, outlined below.

Following these steps, the vector $\boldsymbol{x}$ contained 4 initial condition elements, 18 background emissions elements
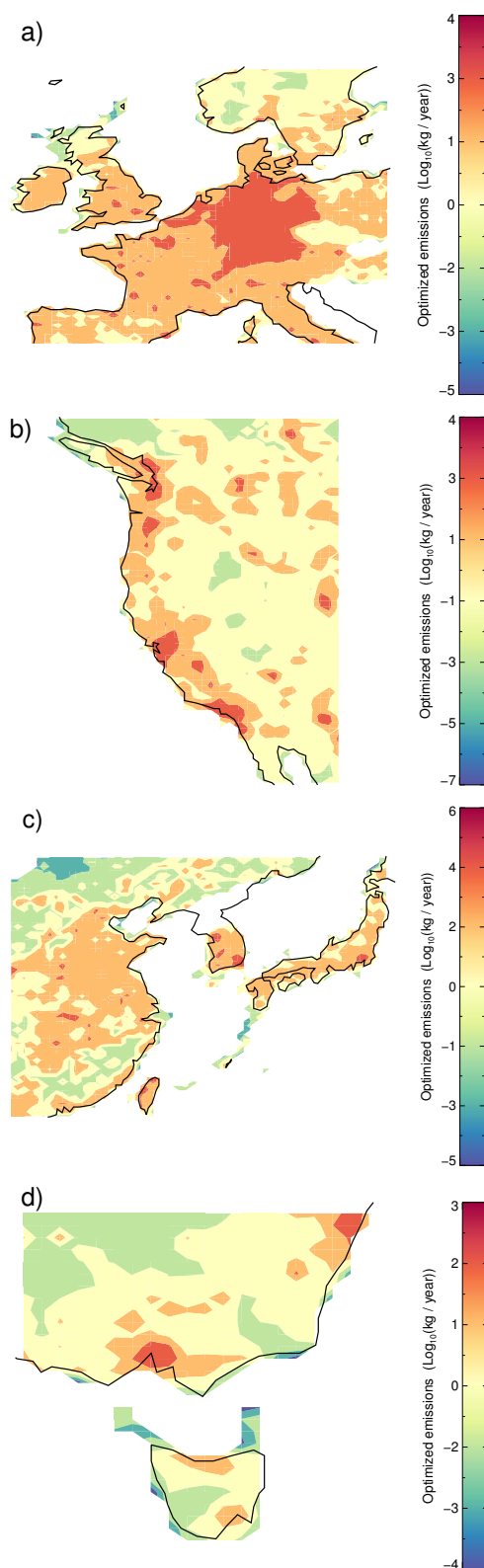

Fig. 5. Optimized SF_6 emissions surrounding **(a)** Mace Head, Ireland, **(b)** Trinidad Head, California, **(c)** Gosan, South Korea, **(d)** Cape Grim, Tasmania.
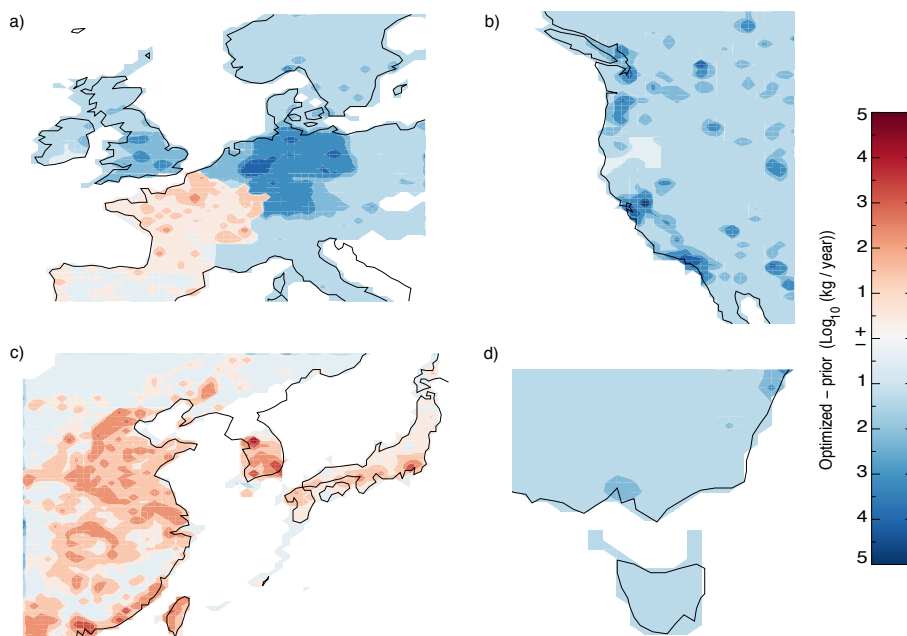
Fig. 6. Difference between optimized and prior SF_6 emissions surrounding **(a)** Mace Head, Ireland, **(b)** Trinidad Head, California, **(c)** Gosan, South Korea, **(d)** Cape Grim, Tasmania. Red and blue colors indicate that the optimized values are higher and lower than the prior respectively.

(6 regions for each year) and approximately 300 local emissions sub-regions (around 100 regions in each year). The measurement vector $\mathbf{y}$ contained approximately 500 5-day averaged measurements at the six sites for the 2007 to 2009 period, along with monthly averages at the stations between 2004 and 2006 used only to constrain the initial conditions.

Emissions were derived using EDGAR v4.0 a priori. Emissions from each sub-region within the local regions were found to follow an exponential probability density function (PDF) in the inventory (see Supplement). Therefore, an inversion that assumes that emissions follow a Gaussian PDF (as in Rigby et al. (2010); Stohl et al. (2009) and many others) may not be ideally suited to the problem, and leads to negative emission rates at some grid cells. To optimally combine the observations, which we assumed had Gaussian uncertainties, and the prior that followed an exponential PDF, we used a Markov Chain Monte-Carlo approach.

Following the discussion in Tarantola (2005, 41–55), the joint PDF of the prior and the observations is given by:

$$\rho_a(\mathbf{x}) = c\rho_f(\mathbf{x})\exp(-\frac{1}{2}(\mathbf{y}-\mathbf{H}\mathbf{x})^T\mathbf{R}^{-1}(\mathbf{y}-\mathbf{H}\mathbf{x})) \quad (5)$$

Where $\rho_f(\mathbf{x})$ is the prior probability density of $\mathbf{x}$ and the exponent term is the likelihood function of model–measurement mismatch, (assuming that the model-measurement uncertainties are normally distributed). The constant c normalizes the posterior distribution. The model-measurement error covariance is described by the $N \times N$ matrix $\mathbf{R}$. These uncertainties were assumed to be uncorrelated,

so that $\mathbf{R}$ was diagonal. The diagonal elements of $\mathbf{R}$ were calculated as the quadratic sum of the measurement repeatability and the scale propagation uncertainty (each assumed to be 0.05 pmolmol^{-1}, following Rigby et al. (2010)) and an estimate of the model representation error. The latter was estimated as the standard deviation of the variability in the local influence on the measurements within the measurement averaging period (as predicted by the LPDM and the prior emissions field). We assumed a priori that the initial condition and non-local emissions elements of $\mathbf{x}$ follow Gaussian distributions, and have a standard deviation equal to their mean (i.e. 100 % uncertainty). The emissions from the small regions within the local domains were assumed to follow an exponential distribution. For the ith element of $\mathbf{x}_{LE}$:

$$\rho_f(x_{LE,i}) = \begin{cases} \lambda_i e^{-\lambda_i x_{LE,i}}, & x_{LE,i} \geq 0 \\ 0, & x_{LE,i} < 0 \end{cases} \quad (6)$$

The parameter λ_i is equal to the inverse of the mean of the distribution (and also the inverse of the standard deviation), which was obtained from the EDGAR database.

The Metropolis-Hastings algorithm was used to explore the posterior probability density (Metropolis et al., 1953; Hastings, 1970; Tarantola, 2005). In this method, the posterior PDF is traversed by a "chain" that propagates according to some simple rules. Firstly, assume that we are at some point $\mathbf{x}_k$ in the state space. We select some new point, $\mathbf{x}_{k+1}$, to attempt to transition to, using a jumping distribution that must be symmetrical (a Gaussian distribution was used). The proposed transition takes place in only one dimension of the
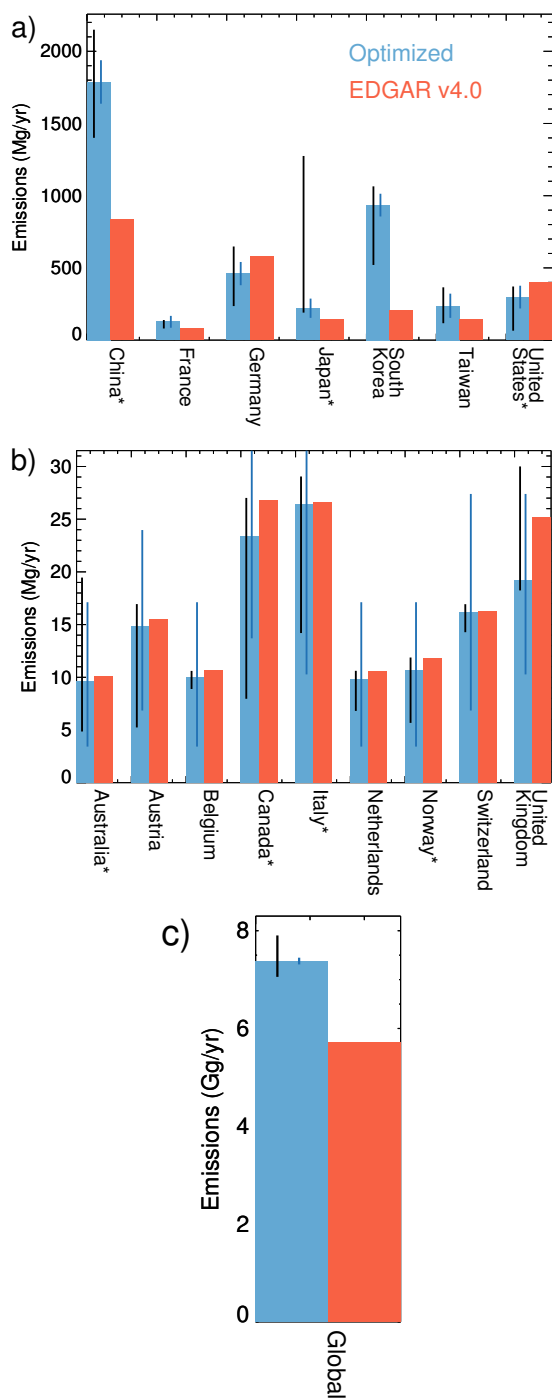
Fig. 7. Optimized 2007–2009 average emissions (blue bars) and EDGAR v4.0 2005 emissions (red bars) for: **(a)** countries with large emissions within "local" regions; **(b)** countries with smaller emissions within "local" regions; **(c)** global totals. The blue error bars show the 16th to 84th percentile ranges for the posterior distribution and black error bars show the range of average emission rates, obtained in the different inversions performed in Sect. 3. Asterisks indicate countries that do not entirely reside within the local regions. The emissions shown for these countries are a fraction of the national total emissions.

space at a time. We accept the move to the new state, provided that:

$$\ln(U) \le \ln\left(\frac{\rho_f(\boldsymbol{x}_{k+1})}{\rho_f(\boldsymbol{x}_k)}\right) \tag{7}$$
$$-\frac{1}{2}\left((\boldsymbol{n}_{k+1})^T \mathbf{R}^{-1}(\boldsymbol{n}_{k+1}) - (\boldsymbol{n}_k)^T \mathbf{R}^{-1}(\boldsymbol{n}_k)\right)$$

where U is a uniformly-distributed random number between 0 and 1 and $\boldsymbol{n}_k$ is a vector of model-measurement residuals ($\boldsymbol{y} - \mathbf{H}\boldsymbol{x}_k$). Equation (7) is written as the logarithm of the usual form, to prevent numerical truncation errors that would otherwise result from taking the exponent of the second term on the right-hand side, which can become very large.

It was found that a chain length of 10^5 was long enough to produce a solution with a sampling error that was much smaller than the derived posterior covariance. A "burn-in" chain of length 10^5 was used to initiate the solution chain. The variance of the jumping distribution was adjusted to obtain an acceptance ratio (the fraction of the total chain length in which a transition was successful) of between 0.25 and 0.5, to most efficiently sample the posterior distribution (Roberts et al., 1997).

In addition to the uncertainty estimate provided by the posterior PDF, we investigated the influence of some assumptions used in the inversion, on the derived emissions uncertainties. We performed multiple inversions aimed at testing the influence of: (a) the measurement averaging period (periods of 1 to 30 days were used); (b) using only afternoon observations, rather than observations throughout the day and night; (c) doubling the number of aggregated emission regions within the local domains; (d) SF_6 calibration scale uncertainties, by running the inversion with the observations increased, and then decreased by 2 % (the estimated scale uncertainty, Rigby et al., 2010). In the discussion that follows we quote both the uncertainty derived in the inversion (the optimal combination of prior and measurement-model uncertainty), and the range of mean emissions that was obtained in the many separate inversions performed.

4 Results and discussion

The derived emissions fields within the LPDM regions for the period 2007 - 2009 are shown in Fig. 5, and the deviations from the EDGAR v4.0 prior, which was compiled for 2005, are shown in Fig. 6. The optimized fields contain a very large amount of information. However, the most significant signals derived in the inversion show that: (a) emissions from most of the East Asian region appear to be significantly higher than the inventory, (b) German emissions are lower than in the inventory, (c) emissions from the West coast of the USA and Canada, and from Southern Australia are somewhat lower than in the inventory. A summary of the derived emissions from countries that emit significant quantities of SF_6 within the AGAGE station footprints are shown in Fig. 7, along with
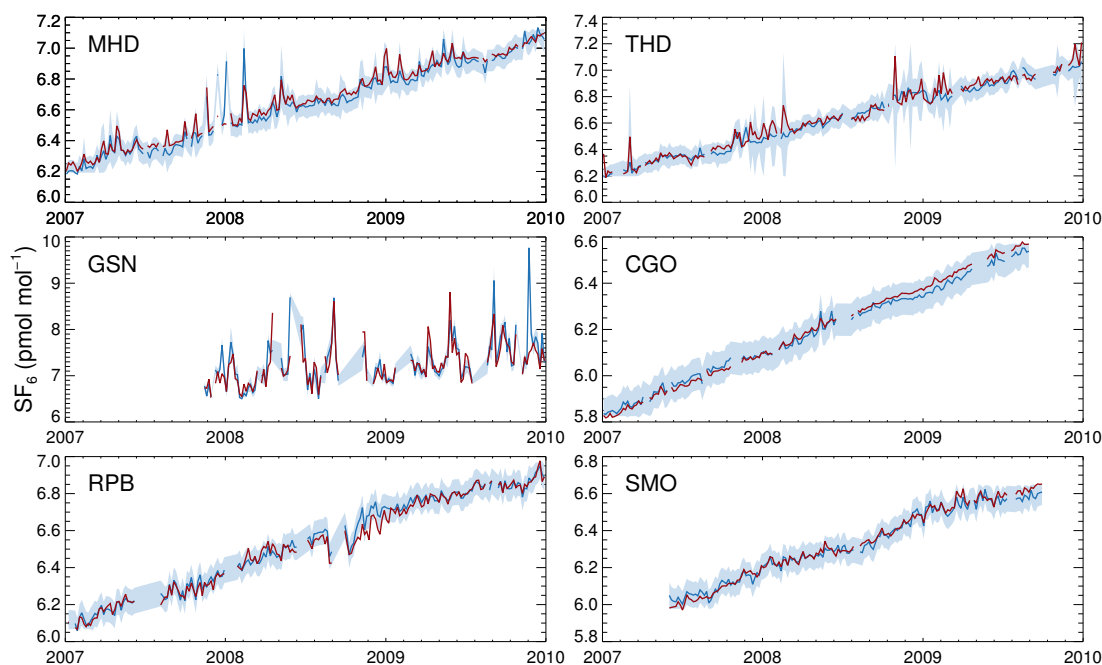
Fig. 8. Optimized 5-day average SF$_6$ mole fractions (red) and observations (blue) at six AGAGE sites used to derive global emissions. Shading indicates 1-sigma model–measurement uncertainty.

Table 1. Model performance metrics. The first two columns show the root-mean-square error (RMSE) for the prior and posterior models compared to the observations. The middle columns show the mean bias between the model and observations. The final two columns show the squared correlation coefficient for the de-trended models and observations. In all cases the a posteriori initial conditions have been added to the derived mole fractions.

Station	RMSE (pmol mol^{-1})		Bias (pmol mol^{-1})		R^2	
	Prior	Posterior	Prior	Posterior	Prior	Posterior
MHD	0.14	0.05	0.116	0.020	0.39	0.48
THD	0.16	0.06	0.117	0.016	0.04	0.11
GSN	0.69	0.33	0.547	0.050	0.37	0.52
CGO	0.06	0.02	0.056	0.007	0.53	0.53
RPB	0.16	0.03	0.140	0.011	0.44	0.56
SMO	0.10	0.02	0.089	0.005	0.62	0.64

their derived uncertainties (tabulated emissions, their uncertainties and fractions of countries falling within local regions are given in the Supplement).

Similar to previous studies, we find that emissions from all countries close to the Gosan monitoring site were higher during 2007–2009 than the EDGAR 2005 estimates (Vollmer et al., 2009; Kim et al., 2010; Li et al., 2011). Our 2007–2009 estimates for emissions from mainland China are somewhat higher than those estimated by Kim et al. (2010) for 2008, and are substantially higher than those found using the Shandianzi station to the North of Beijing by Vollmer et al. (2009) for 2007. Release rates from China derived by Li et al. (2011) are also somewhat lower than our estimates, whilst their Japanese emissions are higher, but within our un-

certainty range. Our South Korean emissions are substantially higher than the Li et al. (2011) estimates. The finding that East Asian emissions are substantially higher than the prior, and that European emissions may be lower, are qualitatively consistent with the Rigby et al. (2010) regional inversion. The differences in emissions at the national scale found in these studies is not entirely surprising given the sparse measurement network. To better constrain these emissions, expanded regional monitoring networks are clearly required.

The global total emission rate of $7.38^{7.45}_{7.31}$ $\binom{7.90}{7.06}$ Ggyr^{-1} for 2007–2009 (the first uncertainty is the 16th and 84th percentile obtained in the inversion, the figures in brackets are range in the mean obtained from the ensemble of inversions) agrees well with the findings of Rigby et al. (2010), who

derived a value of $7.27\pm0.6\,\mathrm{Gg\,yr^{-1}}$ (1-σ uncertainty) for 2008.

The measurements and optimized mole fractions at the six AGAGE sites used in this work are shown in Fig. 8. Root-mean-square differences between the prior and posterior models are shown in Table 1, along with mean bias and correlation coefficients for de-trended model and data. The variability is found to be very well reproduced at most sites. Further, the a posteriori mean bias and root-mean-square errors are generally smaller than the estimated measurement uncertainty. It is particularly striking that excellent agreement is obtained at Gosan, Korea (posterior de-trended squared correlation coefficient of 0.52). This site was found to be poorly modeled using a global Eulerian model by Rigby et al. (2010, supplementary material). We propose that the use of high-resolution meteorological fields in this region, along with reduced aggregation error close to the site is responsible for this improved agreement. Poorer agreement is obtained at Trinidad Head, California, which may indicate transport model inaccuracies in that region, or significant inaccuracies in the spatial distribution of the local prior emissions field. No attempt was made here to account for systematic transport model biases, which could be substantial. This important term can be estimated by substitution of different Eulerian or Lagrangian models in the framework, and will be the subject of future work.

Although no code modification is required for either of the CTMs, the implementation of this framework can be quite involved. In light of this, and to aid future researchers, the Interactive Data Language (IDL) code used to combine NAME and MOZART sensitivities will be made available upon request to the corresponding author.

5 Conclusions

We have shown that global and regional emissions of long-lived trace gases can be derived efficiently by combining Eulerian and Lagrangian chemical transport models. A method is outlined that allows these emissions to be determined in a single step. This method has significant advantages over previous approaches to deriving emissions from a sparse monitoring network in that: aggregation errors are minimized by the use of a Lagrangian particle dispersion model to simulate near-field emissions; high-resolution simulations can be efficiently performed over limited regions close to the monitoring sites; uncertainties can be estimated, along with covariances between the far-field and near-field emissions.

Supplement related to this article is available online at:
http://www.atmos-chem-phys.net/11/9887/2011/
acp-11-9887-2011-supplement.zip.

References

Chen, Y.-H. and Prinn, R. G.: Estimation of atmospheric methane emissions between 1996 and 2001 using a three-dimensional global chemical transport model, J. Geophys. Res., 111, D10307, doi:10.1029/2005JD006058, 2006.

Emmons, L. K., Walters, S., Hess, P. G., Lamarque, J.-F., Pfister, G. G., Fillmore, D., Granier, C., Guenther, A., Kinnison, D., Laepple, T., Orlando, J., Tie, X., Tyndall, G., Wiedinmyer, C., Baughcum, S. L., and Kloster, S.: Description and evaluation of the Model for Ozone and Related chemical Tracers, version 4 (MOZART-4), Geosci. Model Dev., 3, 43–67, doi:10.5194/gmd-3-43-2010, 2010.

Enting, I.: Inverse Problems in Atmospheric Constituent Transport, Cambridge University Press, Cambridge, UK, 2002.

Geller, L. S., Elkins, J. W., Lobert, J. M., Clarke, A. D., Hurst, D. F., Butler, J. H., and Myers, R. C.: Tropospheric SF6: Observed latitudinal distribution and trends, derived emissions and interhemispheric exchange time, Geophys. Res. Lett., 24, 675–678, 1997.

Gurney, K. R., Law, R. M., Denning, A. S., Rayner, P. J., Baker, D., Bousquet, P., Bruhwilerk, L., Chen, Y.-h., Ciais, P., Fan, S., Fung, I., Gloor, M., Heimann, M., Higuchi, K., John, J., Maki, T., Maksyutov, S., Masarie, K., Peylin, P., Prather, M., Pak, B., Randerson, J., Sarmiento, J., Taguchi, S., Takahashi, T., and Yuen, C.: Towards robust regional estimates of CO_2 sources and sinks using atmospheric transport models, Nature, 415, 626–630, 2002.

Hastings, W. K.: Monte Carlo sampling methods using Markov chains and their applications, Biometrika, 57, 97–109, 1970.

JRC/PBL: Emission Database for Global Atmospheric Research (EDGAR), release version 4.0, 2009.

Kaminski, T., Heimann, M., and Giering, R.: A coarse grid three-dimensional global inverse model of the atmospheric transport 1. Adjoint model and Jacobian matrix, J. Geophys. Res., 104, 18535–18553, 1999.

Khasibatla, P., Heimann, M., Rayner, P., Mahowald, N., Prinn, R., and Hartley, D., eds.: Inverse Methods in Global Biogeochemical Cycles, American Geophysical Union, Washington DC, USA, 2000.

Kim, J., Li, S., Kim, K. R., Stohl, A., Mühle, J., Kim, S. K., Park, M. K., Kang, D.-J., Lee, G., Harth, C. M., Salameh, P. K., and Weiss,R. F.: Regional atmospheric emissions determined from measurements at Jeju Island, Korea?: Halogenated compounds from China, J. Geophys. Res., 37, L12801, doi:10.1029/2010GL043263.

Kopacz, M., Jacob, D. J., Fisher, J. a., Logan, J. a., Zhang, L., Megretskaia, I. a., Yantosca, R. M., Singh, K., Henze, D. K., Burrows, J. P., Buchwitz, M., Khlystova, I., McMillan, W. W., Gille, J. C., Edwards, D. P., Eldering, a., Thouret, V., and Nedelec, P.: Global estimates of CO sources with high resolution by adjoint inversion of multiple satellite datasets (MOPITT, AIRS, SCIAMACHY, TES), Atmos. Chem. Phys., 10, 855–876, doi:10.5194/acp-10-855-2010, 2010.

Koyama, Y., Maksyutov, S., Mukai, H., Thoning, K., and Tans, P.: Simulation of variability in atmospheric carbon dioxide using a global coupled Eulerian – Lagrangian transport model, Geosci. Model Dev., 4, 317–324, doi:10.5194/gmd-4-317-2011, 2011.

Krol, M., Houweling, S., Bregman, B., van Den Broek, M., Segers, A., van Velthoven, P., Peters, W., Dentener, F., and Bergamaschi, P.: The two-way nested global chemistry-transport zoom model TM5: algorithm and applications, Atmos. Chem. Phys., 5, 417–432, doi:10.5194/acp-5-417-2005, 2005.

Levin, I., Naegler, T., Heinz, R., Osusko, D., Cuevas, E., Engel, A., Ilmberger, J., Langenfelds, R. L., Neininger, B., Rohden, C. V., Steele, L. P., Weller, R., Worthy, D. E., and Zimov, S. A.: The global SF6 source inferred from long-term high precision atmospheric measurements and its comparison with emission inventories, Atmos. Chem. Phys., 10, 2655–2662, doi:10.5194/acp-10-2655-2010,, 2010.

Li, S., Kim, J., Kim, K.-R., Mühle, J., Kim, S.-K., Park, M.-K., Stohl, A., Kang, D.-J., Arnold, T., Harth, C. M., Salameh, P. K., and Weiss, R. F.: Emissions of Halogenated Compounds in East Asia Determined from Measurements at Jeju Island, Korea, Environ. Sci. Technol., 45, 5668–5675, 2011.

Manning, A. J., O'Doherty, S., Jones, A. R., Simmonds, P. G., and Derwent, R. G.: Estimating UK methane and nitrous oxide emissions from 1990 to 2007 using an inversion modeling approach, J. Geophys. Res., 116, D02305, doi:10.1029/2010JD014763, 2011.

Meirink, J. F., Bergamaschi, P., and Krol, M. C.: Four-dimensional variational data assimilation for inverse modelling of atmospheric methane emissions: method and comparison with synthesis inversion, Atmos. Chem. Phys., 8, 6341–6353, doi:10.5194/acp-8-6341-2008, 2008.

Metropolis, N., Rosenbluth, A., Rosenbluth, M., Teller, A., and Teller, E.: Equation of State Calculations by Fast Computing Machines, J. Chem. Phys., 21, 1087–1092, 1953.

O'Doherty, S., Cunnold, D. M., Manning, A., Miller, B. R., Wang, R. H. J., Krummel, P. B., Fraser, P. J., Simmonds, P. G., McCulloch, A., Weiss, R. F., Salameh, P., Porter, L. W., Prinn, R. G., Huang, J., Sturrock, G., Ryall, D., Derwent, R. G., and Montzka, S. A.: Rapid growth of hydrofluorocarbon 134a and hydrochlorofluorocarbons 141b, 142b, and 22 from Advanced Global Atmospheric Gases Experiment (AGAGE) observations at Cape Grim, Tasmania, and Mace Head, Ireland, J. Geophys. Res., 109, 1–16, 2004.

Peters, W., Miller, J. B., Whitaker, J., Denning, A. S., Hirsch, A., Krol, M. C., Zupanski, D., Bruhwiler, L., and Tans, P. P.: An ensemble data assimilation system to estimate CO2 surface fluxes from atmospheric trace gas observations, J. Geophys. Res., 110, D24304, doi:10.1029/2005JD006157, 2005.

Prinn, R. G., Weiss, R. F., Fraser, P. J., Simmonds, P. G., Cunnold, D. M., Alyea, F. N., O'Doherty, S., Salameh, P., Miller, B. R., Huang, J., Wang, R. H. J., Hartley, D. E., Harth, C., Steele, L. P., Sturrock, G., Midgley, P. M., and McCulloch, a.: A history of chemically and radiatively important gases in air deduced from ALE/GAGE/AGAGE, J. Geophys. Res., 105, 17751–17792 doi:10.1029/2000JD900141, 2000.

Rigby, M., Prinn, R. G., Fraser, P. J., Simmonds, P. G., Langenfelds, R. L., Huang, J., Cunnold, D. M., Steele, L. P., Krummel, P. B., Weiss, R. F., O'Doherty, S., Salameh, P. K., Wang, H. J., Harth, C. M., Mühle, J., and Porter, L. W.: Renewed growth of atmospheric methane, Geophys. Res. Lett., 35, L22805, doi:10.1029/2008GL036037, 2008.

Rigby, M., Mühle, J., Miller, B. R., Prinn, R. G., Krummel, P. B., Steele, L. P., Fraser, P. J., Salameh, P. K., Harth, C. M., Weiss, R. F., Greally, B. R., O'Doherty, S., Simmonds, P. G., Vollmer, M. K., Reimann, S., Kim, J., Kim, K.-R., Wang, H. J., Olivier, J. G. J., Dlugokencky, E. J., Dutton, G. S., Hall, B. D., and Elkins, J. W.: History of atmospheric SF6 from 1973 to 2008, Atmos. Chem. Phys., 10, 10305–10320, doi:10.5194/acp-10-10305-2010, 2010.

Roberts, G. O., Gelman, A., and Gilks, W. R.: Weak Convergence And Optimal Scaling Of Random Walk Metropolis Algorithms, Ann. Appl. Probab., 7, 110–120, 1997.

Roedenbeck, C., Gerbig, C., Trusilova, K., and Heimann, M.: A two-step scheme for high-resolution regional atmospheric trace gas inversions based on independent models, Atmos. Chem. Phys., 9, 5331–5342, doi:10.5194/acp-9-5331-2009, 2009.

Ryall, D. B., Maryon, R. H., Derwent, R. G., and Simmonds, P. G.: Modelling long-range transport of CFCs to Mace Head, Ireland, Q. J. Roy. Meteorol. Soc., 124, 417–446, doi:10.1002/qj.49712454604, 1998.

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

Tarantola, A.: Inverse Problem Theory, Society for Industrial and Applied Mathematics, Philadelphia, USA, 2005.

Vollmer, M. K., Zhou, L. X., Greally, B. R., Henne, S., Yao, B., Reimann, S., Stordal, F., Cunnold, D. M., Zhang, X. C., Maione, M., Zhang, F., Huang, J., and Simmonds, P. G.: Emissions of ozone-depleting halocarbons from China, Geophys. Res. Lett., 36, L15823, doi:10.1029/2009GL038659, 2009.

Wunsch, C.: Discrete Inverse and State Estimation Problems: With Geophysical Fluid Applications, Cambridge University Press, Cambridge, UK, 2006.