A new inversion method to calculate emission inventories without a prior at mesoscale: Application to the anthropogenic CO₂ emission from Houston, Texas

J. Brioude,1,2 G. Petron,2,3 G. J. Frost,1,2 R. Ahmadov,1,2 W. M. Angevine,1,2 E.-Y. Hsie,1,2 S.-W. Kim,1,2 S.-H. Lee,4 S. A. McKeen,1,2 M. Trainer,1 F. C. Fehsenfeld,1,2 J. S. Holloway,1,2 J. Peischl,1,2 T. B. Ryerson,1 and K. R. Gurney5

Received 23 September 2011; revised 13 January 2012; accepted 17 January 2012; published 10 March 2012.

[1] We developed a new inversion method to calculate an emission inventory for an anthropogenic pollutant without a prior emission estimate at mesoscale. This method employs slopes between mixing ratio enhancements of a given pollutant (CO₂, for instance) with other co-emitted tracers in conjunction with the emission inventories of those tracers (CO, NOy, and SO₂ are used in this example). The current application of this method employed in situ measurements onboard the NOAA WP-3 research aircraft during the 2006 Texas Air Quality Study (TexAQS 2006). We used 3 different transport models to estimate the uncertainties introduced by the transport models in the inversion. We demonstrated the validity of the new inversion method by calculating a 4 × 4 km² emission inventory of anthropogenic CO₂ in the Houston area in Texas, and comparing it to the 10 × 10 km² Vulcan emission inventory for the same region. The calculated anthropogenic CO₂ inventory for the Houston Ship Channel, home to numerous major industrial and port emission sources, showed excellent agreement with Vulcan. The daytime CO₂ average flux from the Ship Channel is the largest urban CO₂ flux reported in the literature. Compared to Vulcan, the daytime urban area CO₂ emissions were higher by 37% ± 6%. Those differences can be explained by uncertainties in emission factors in Vulcan and by increased emissions from point sources and on-road emitters between 2002, the reference year in Vulcan, and 2006, the year that the TexAQS observations were made.


1. Introduction

[2] Emission inventories are used in models to predict or estimate the atmospheric chemical composition at mesoscale (air quality forecasts) and global scales (climate system). Emission inventories are difficult to estimate at mesoscale because observations with good spatial (~km) and temporal (~hourly) resolutions covering a large domain are required. Emission inventories are typically calculated using a “bottom-up” estimate by summing emissions from different sources using emission factors appropriate to each source. Those sources can be anthropogenic, oceanic or biogenic (including biomass burning). However, a bottom-up estimate is often associated with uncertainties that reduce the forecast skill of chemical-transport models.

[3] To evaluate and improve an existing emission inventory (the “prior”), it is common to utilize observations and an inversion method. When a prior is not available for a given species, the inversion is calculated assuming a homogenous surface flux of an arbitrary value. For instance, Bergamaschi et al. [2010] showed that 1 × 1₀⁶CH₄ emissions in Europe could be constrained using continuous observations from 17 surface stations during 6 years without a prior. The flux obtained at a country scale was close to the flux of a reference inversion that used a prior. Such a method is usually applied on a coarse resolution grid, because a large number of observations are needed to constrain the surface flux in the inversion process. One major problem is that the spatial
distribution of surface sources in the posterior can be biased by uncertainties from the transport model, requiring a careful analysis of the inversion results using additional data.

[4] In this paper, we present a new inversion method that allows the calculation of an emission inventory without a prior at mesoscale. This method is based on the linear relationship between a given species and co-emitted tracers of the same anthropogenic origin. Linear relationships between species have been used in different ways in the literature. Palmer et al. [2006] and Wang et al. [2009] used aircraft or satellite measurements of correlations between CO and CO2 in conjunction with prior estimates of CO and CO2 fluxes to optimize CO2 surface fluxes either at regional or global scale using a chemical transport model. They found that the additional information provided by CO reduced the uncertainties in the posterior estimate of CO2 flux compared with a classic inversion alone. Based on measurements only, Wunch et al. [2009] used CO2 as a tracer to infer surface fluxes of CH4 and CO at the scale of the city of Los Angeles. By assuming that the spatial distribution of emissions is the same for the tracer (CO2) and the other chemical species of interest, Wunch et al. [2009] were able to infer the average flux of CO and CH4 in Los Angeles based on the average flux of CO2 from the California Air Resources Board (CARB) and the Emission Database for Global Atmospheric Research (EDGAR) inventories. The calculated CO and CH4 fluxes were in agreement with the fluxes found in CARB and EDGAR. Likewise, van der Laan et al. [2009] estimated emissions of CH4 and N2O in Netherlands using their linear relationships with 222Radon, for which surface flux emissions are known. They found a good agreement with the existing national inventories of CH4 and N2O emissions.

[5] Those methods allow the calculation of surface fluxes integrated over a large spatial domain, like a megacity or a country, with a single dominant source of emissions. However, improvement of surface inventories for a region encompassing a mix of emissions sectors (e.g., urban and industrial) requires knowledge of emission ratios of the desired pollutants to CO2 for each source type, and the use of a transport model if a gridded inventory is to be calculated.

[6] In section 2 of this paper we demonstrate that, based on a new inverse modeling approach, measured slopes of linear relationships between tracer mixing ratios can be translated into surface flux ratios in order to infer the surface inventory of a given species without a prior estimate. We call this new inversion approach the “flux ratio inversion method.” Brioude et al. [2011] applied inverse modeling techniques using in situ measurements of the NOAA WP-3 aircraft over Houston during the Texas Air Quality Study (TexAQS) 2006 campaign [Parrish et al., 2009, and references therein] and 3 transport models to optimize CO, NOy and SO2 inventories. In section 3, we use those optimized inventories to calculate a 4 x 4 km2 CO2 emission inventory in Houston without a CO2 prior based on the slopes of CO, NOy and SO2 with CO2. Figure 1 shows examples of the linear relationships between mixing ratios of CO2, CO, and NOy seen downwind of the Houston area.
It is known that urban areas are net sources of CO₂ despite CO₂ uptake by vegetation [Helfter et al., 2011 and references therein]. To date, the largest CO₂ fluxes reported in urban areas are for Montreal (0.6 × 10⁻⁹ kg/m²/s [Bergeron and Strachan, 2011]), London (1.1 × 10⁻⁶ kg/m²/s [Helfter et al., 2011]) and Edinburgh (1.1 × 10⁻⁶ kg/m²/s [Nemitz et al., 2002]). Suburban sites have lower net CO₂ fluxes because of smaller anthropogenic emissions and larger vegetation cover, resulting in larger uptake of CO₂ by vegetation [e.g., Bergeron and Strachan, 2011]. An exponential decrease relationship exists between net emissions of CO₂ and vegetation cover in urban areas [e.g., Helfter et al., 2011; Ramamurthy and Pardyjak, 2011].

In order to estimate uncertainties in the flux ratio inversion method due to biogenic processes that cannot be addressed by correlations between anthropogenic tracers, our sensitivity tests showed that the constructed CO₂ emission inventory from the flux ratio inversion must be used as a prior in a classic inverse calculation to derive an optimized CO₂ posterior. In section 4, a CO₂ posterior is calculated and compared to the constructed CO₂ inventory by the flux ratio inversion and to the 10 × 10 km² Vulcan anthropogenic CO₂ inventory for the U.S. [Gurney et al., 2009]. Our conclusions are given in section 5.

Houston was selected as the site for the initial application of this approach for several reasons. Houston, the 4th most populous urban area in the U.S., has serious and frequent problems with non-attainment of air quality standards that lead to detrimental human health effects. Houston’s poor air quality results from an unusual variety of pollutant sources, including the petrochemical industry, power plants, high automobile usage, and a major shipping terminal, which makes development of bottom-up inventories particularly difficult. About 50% of the U.S. petrochemical refining capability is located in and around the Houston metropolitan area. This atypical combination of large sources in Houston also leads to higher per capita CO₂ emissions than other U.S. cities [Gurney et al., 2009]. The complexity of emissions sources and their impact on air quality in Houston and other cities of east Texas was a primary motivation for the Texas Air Quality Studies in 2000 and 2006; the latter campaign is the source of the observations used in our inverse modeling.

2. Method

2.1. Models and Observations

In this work we used the same 2006 observations and the same three transport models as in the work of Brioude et al. [2011]. Brioude et al. [2011] present a detailed description of the observations and models and an in-depth discussion of their uncertainties, which we summarize here. We used CO, NO₃, SO₂ and CO₂ mixing ratios measured in September 2006 on board the NOAA WP-3 research aircraft during the TexAQS 2006 campaign over southeastern Texas. Several WP-3 flight plans were designed to characterize the pollutant emission inventories upward and downward of the Houston area. The 8 flights used in this study (from September 15th to 29th) were selected to avoid cloudy and rainy days. Among them, 5 flights were associated with clear sky conditions, and 3 flights had a small amount of scattered clouds. Those flights were weekday flights, and therefore our estimate of CO₂ surface emissions might be biased toward weekdays. CO was measured using vacuum ultraviolet resonance fluorescence [Holloway et al., 2000] with an uncertainty of ± (1 ppbv +0.05% CO). NOy (total reactive nitrogen [Fahey et al., 1986]) was measured by ozone chemiluminescence with an uncertainty of ± (0.20 ppbv + 0.12 NOy) [Ryerson et al., 2000]. SO₂ was measured by UV fluorescence with an uncertainty of ± (0.3 ppbv + 0.15% SO₂) [Ryerson et al., 1998]. CO₂ was measured by a nondispersive IR absorption instrument with a precision of ± (0.08 ppmv + 0.3% CO₂) [Peischl et al., 2010]. All those species were measured once per second.

In the work of Brioude et al. [2011], we used the chemical measurements in the atmospheric boundary layer (ABL) to evaluate and improve the 4 × 4 km² U.S. Environmental Protection Agency’s (EPA’s) National Emission Inventory (NEI) 2005 of CO, NOy and SO₂. The optimized CO, NOy and SO₂ emission inventories for Houston calculated by Brioude et al. [2011] were used in this paper. For each flight, a chemical background level was subtracted from the measured mixing ratios. We defined the chemical background of CO, NOy and SO₂ as the lowest mixing ratio found in the ABL upwind of Houston. We treated CO, NOy and SO₂ as passive tracers. The measurements were not far from the emission sources (within 150 km, Figure 1), so that secondary production of CO can be neglected. We found that the uncertainty of assuming that SO₂ and NOy are passive tracers on the surface flux estimate in the posterior was probably lower than other sources of uncertainties, as described by Brioude et al. [2011]. The chemical background value of CO₂ was estimated from the vertical distribution of CO₂ mixing ratios obtained during each flight, as exemplified in Figure 1d. During TexAQS 2006, the average CO₂ uptake by the biosphere measured on the NOAA WP-3 was small in magnitude compared to anthropogenic CO₂ fluxes over the Houston urban area. We discuss sensitivities to these aspects of the methodology further in section 4.

We used meteorological data simulated by 3 independent meteorological models (operational ECMWF and 2 configurations of the WRF mesoscale model) to run the FLEXPART Lagrangian particle dispersion model [Stohl et al., 2005], which simulated atmospheric transport over 24 h in order to focus on local anthropogenic source contributions to ambient concentrations. Each combination of simulated meteorological fields with FLEXPART can be considered as an independent transport model. The ensemble of those 3 transport models was used to estimate model uncertainties. 20,000 particles were released every 20 s (every 2km based on a typical aircraft ground speed of 100 m/s) from FLEXPART boxes located along the aircraft flight tracks. The chemical measurements were averaged in each FLEXPART box time window. The FLEXPART output had a resolution of 0.045° × 0.05°. When this output is combined with a surface flux emission inventory, one can calculate chemical mixing ratios along the aircraft flight tracks. In this way, FLEXPART calculates source-receptor relationships (matrix H in section 2.2) between surface emission inventories and aircraft measurements, which is used in our inverse modeling technique. Having 3 transport models allowed us to estimate model uncertainties in the simulation of chemical mixing ratios using the inversion method. We used additional large-scale FLEXPART
simulations to estimate regional and continental transport of pollution into the Houston area, in order to constrain the observational data to those with local emission influences only (for further details about the large scale FLEXPART runs, see Brioude et al. [2007]). Observations measured within the ABL were used in this study. No clear bias between transport models was found in the simulations by Brioude et al. [2011]. However, nocturnal atmospheric transport was poorly represented by the meteorological models. Therefore, the CO2 surface fluxes calculated in this paper were compared to the Vulcan anthropogenic CO2 fluxes for daytime only (from 09:00 to 18:00 local time (LT)). To reduce the size of the matrices involved in the inversion method, we used surface fluxes for grid cells that have significant anthropogenic emissions.

We compare our inventory to the 10 \times 10 \text{km}^2 bottom-up Vulcan anthropogenic CO2 emission inventory [Gurney et al., 2009]. As a starting point for developing our new approach, Figure 2b presents the posterior CO2 emission inventory calculated using a standard 3D inverse modeling technique and the aircraft measurements of CO2 mixing ratios from TexAQS 2006. No prior was used in this inversion method. The matrix H involved in the inversion had a dimension of 2817 lines (number of observations) by 791 columns (number of surface grid cells in 3D). To obtain the most realistic posterior, we used an iterative gradient method. Using an analytical solution gives unrealistically large variation between contiguous grid cells. The fluxes shown in Figure 2b represent the average fluxes using the ensemble of 3 transport models. Compared to the 10 \times 10 \text{km}^2 Vulcan anthropogenic CO2 emission inventory (Figure 2c), the posterior fluxes have a similar range of values, from \( -0.5 \) to \( +4 \times 10^{-6} \text{kg/m}^2\text{s} \). However, the spatial distribution is somewhat random, with large emissions on the western border of Houston and near Freeport to the south.

[14] We compare surface fluxes between the CO2 posterior and Vulcan in 3 different regions in Houston (Figure 2a): a large area called Greater Houston, the Houston urban core area that is dominated by urban emissions, and the Houston Ship Channel area with strong industrial, port, and motor vehicle emissions. The spatial distribution of the posterior’s largest emission values is roughly in agreement with that in Vulcan, with the maximum fluxes found in the Ship Channel (Figure 2). The average CO2 flux in the Ship Channel (1.8 \times 10^{-6} \text{kg/m}^2\text{s}) and the urban area (0.87 \times 10^{-6} \text{kg/m}^2\text{s}) are close to the anthropogenic emissions in Vulcan (2.17 and 0.7 \times 10^{-6} \text{kg/m}^2\text{s}, respectively). But the CO2 spatial distribution is quite noisy and does not seem to be consistently related to existing anthropogenic emissions of CO, NOy or SO2 (Figure 3). Therefore, this emission inventory has a highly uncertain spatial distribution and should not be used in a chemical-transport model. To improve this posterior inventory, one would need many more observations (i.e., more observations than surface grid cells) to constrain the surface fluxes of the posterior in a 4D inversion. However, due to the inherent uncertainties in the transport models, the spatial distribution of fluxes in the posterior would always be uncertain, regardless of the number of observations used in the inversion. Therefore, a different method that uses additional constraints besides the absolute mixing ratio observations is necessary to calculate a realistic posterior inventory.

2.2. Theoretical Basis of the Flux Ratio Inversion Method

[15] This section presents the mathematical basis of the method to calculate a realistic emission inventory without a prior. A validation of the method is given in section 3.1. The method is based on the relationship between tracers of the same origin (anthropogenic, biogenic, etc). For example, Figure 1 shows that linear relationships exist between the mixing ratios of CO2 with CO and NOy, with a different slope depending on whether the source is solely urban emissions or a combination of urban and industrial emissions. The method allows the calculation of the emissions of a given species \( S \) using the optimized gridded emission inventories of species that are known to be co-emitted; for
example, CO, NOy and SO2 are co-emitted with anthropogenic CO2. The inventories of CO ($x_{CO}$), NOy ($x_{NOy}$) and SO2 ($x_{SO2}$) are optimized using a Bayesian 4D inverse modeling method [Brioude et al., 2011]. The matrix $H$ involved in the 4D inversion has a dimension of 2817 lines (number of observations) by 5537 columns (number of surface grid cells in 4D, with a first time step between 19:00 and 04:00 LT, a second time step between 05:00 and 08:00 LT, and 5 time steps every 2 h between 09:00 and 18:00 LT, resulting in 7 different time steps).

[16] An inverse modeling method is designed to minimize the error between measured and simulated observations ($\varepsilon_S$) which is equivalent to minimizing a cost function ($J$) written as

$$J_S = \frac{1}{2} \varepsilon_S^T R_S^{-1} \varepsilon_S$$

(1)

with $R_S$ being the covariance matrix of the error in the observations. In the case of mixing ratios of chemical species, $\varepsilon_S$ is written:

$$\varepsilon_S = y_S^o - y_S^m$$

(2)

with $y_S^o$ being the measured mixing ratio enhancement above background level, and $y_S^m$ the simulated mixing ratio enhancement.

[17] The simulated observation of $S$, $y_S^m$, by a linear model $H$ using an inventory $x_S$ can be written:

$$y_S^m = H x_S$$

(3)

Hence, the cost function $J_S$ can be written:

$$J_S(x_S) = \frac{1}{2} \varepsilon_S^T R_S^{-1} \varepsilon_S$$

(4)

In our inversion problem, $x_S$ is the variable input of the cost function $J_S$.

[18] We can estimate $y_S^o$ and $x_S$ using information from another tracer $t$. A linear relationship between the $i$th observation of a species $S$ with the $i$th observation of a tracer $t$ can be written

$$y_S^o(i) = a_t^S(i)y_t^o(i) + c(i)$$

(5)

with $a_t^S$ being the time series of slope between the mixing ratio of $S$ and that of the tracer $t$ for each observation, and $c$
being a term that varies with the chemical background of $S$ and tracer $t$. In our case, $y^o_S$ and $y^o_t$ are background enhancements above background, and we assume that when $y^o_S$ equals 0, $y^o_t$ equals 0. Therefore the term $c$ is equal to zero, and $a^o_t$ can be calculated from the ratios between the mean mixing ratio (over 20 s) above background of $S$ and that of the tracer $t$ for each observation. However, equation (22), the main result of the mathematical development of the flux ratio inversion method, would be the same if absolute mixing ratios were used instead of mixing ratio enhancements above background and if both the terms $a$ and $c$ had to be estimated.

[19] The ratio $a^o_t$ varies with time (because the measurements are downwind of a mixture of sources for which the origin varies with time) and has the same number of elements as $y^o_S$ and $y^o_t$. These attributes mean that 2 sources of information ($a^o_t$ and $y^o_t$) can be used instead of only one ($y^o_S$)

In each 20-s time period, we assume that $y^o_S$ and $y^o_t$ are correlated with a correlation $\rho$ equal to 1. Based on classic error propagation estimates of correlated variables, one can write the $i$th element in $\sigma_a$ as

$$a^o_t(i) = y^o_S(i) \times y^o_t(i)$$

(6)

In each 20-s time period, we assume that $y^o_S$ and $y^o_t$ are correlated with a correlation $\rho$ equal to 1. Based on classic error propagation estimates of correlated variables, one can write the $i$th element in $\sigma_a$ as

$$\left(\frac{\sigma_a(i)}{a^o_t(i)}\right)^2 = \left(\frac{\sigma_S(i)}{y^o_S(i)}\right)^2 + \left(\frac{\sigma_t(i)}{y^o_t(i)}\right)^2 - 2 \frac{\sigma_S(i) \sigma_t(i)}{y^o_S(i) y^o_t(i)} \rho$$

(7)

$$\left(\frac{\sigma_a(i)}{a^o_t(i)}\right)^2 = \left(\frac{\sigma_S(i)}{y^o_S(i)} - \frac{\sigma_t(i)}{y^o_t(i)}\right)^2$$

(8)

$$\sigma_a(i) = a^o_t(i) \left(\frac{\sigma_S(i)}{y^o_S(i)} - \frac{\sigma_t(i)}{y^o_t(i)}\right)$$

(9)

$$\sigma_a(i) = \left| \frac{\sigma_S(i)}{y^o_S(i)} - \frac{\sigma_t(i)}{y^o_t(i)} \right|$$

(10)

This definition of $\sigma_a$ is used below to calculate covariance matrices.

[20] If one has $n$ tracers, $y^o_S$ can be written $n$ different ways based on the measured $a^o_t$ and measured $y^o_t$.

[21] We also assume that a relationship can be found between the surface inventory of $S$ ($x_S$) and the surface inventories of each tracer $t$ ($x_t$). For $n$ tracers, we can then write

$$x_s = \sum_{k=1}^{n} B_k x_k$$

(11)

where $B_k$ represents the linear relationship between surface fluxes of $S$ and the $n$ different tracers. $B_k$ is a diagonal matrix and diag($B_k$) has the same dimension as $x_s$ and the $x_t$ (in space and time). Here the unknown $x_s$ is based on $n$ unknown diag ($B_k$) and $n$ known $x_t$. $B_k$ is equivalent to a flux ratio between the surface fluxes $x_s$ and $x_t$. $B_k$ depends on the nature of the source and varies if the emission ratio between tracers varies. For instance, if the $i$th element $x_t(i)$ varies with time due to variation in emission intensity, $x_t(i)$ will also vary, but $B_k(i,i)$ will remain constant assuming that the emission ratios don’t vary over the time period of the observations (in our case, about 15 days).

[22] Based on equations (4) and (5), one can define $J_S$ using $n$ different variables $x_t$ instead of one $x_s$

$$J_S(x_t) = J_S(x_t, t \in [1, n])$$

(12)

For a given tracer $t$,

$$J_S = \frac{1}{2} \left( A^o_t y^o_t - H \sum_{k=1}^{n} B_k x_k \right)^T R_S^{-1} \left( A^o_t y^o_t - H \sum_{k=1}^{n} B_k x_k \right)$$

(13)

$A^o_t$ is a diagonal matrix with diag($A^o_t$) being the elements of the measured slopes $a^o_t$.

[23] $J_S$ can be written $n$ different ways using the $n$ different observed $a^o_t$ and $y^o_t$.

$$J_S = \frac{1}{2} \left( A^o_t y^o_t - H \sum_{k=1}^{n} B_k x_k \right)^T R_S^{-1} \left( A^o_t y^o_t - H \sum_{k=1}^{n} B_k x_k \right)$$

(14)

$J_S$ is written with $2n$ independent observations (diag($A^o_t$) and $y^o_t$) and $n$ unknowns diag($B_k$).

[24] The simulated observation of a given tracer $t$, $y^m_t$, by a linear model $H$ using an inventory $x_t$ can be written

$$y^m_t = H x_t$$

(15)

with $\epsilon_t$ being the unbiased error between $y^m_t$ and $y^o_t$

$$\epsilon_t = y^m_t - y^o_t$$

$$y^o_t = H x_t + \epsilon_t$$

Therefore,

$$y^o_S = A^o_t y^o_t$$

$$y^o_S = A^o_t (H x_t + \epsilon_t)$$

(16)

Hence, one can write equation (14) as

$$\forall t \in [1, n], J_S = \frac{1}{2} \left( A^o_t (H x_t + \epsilon_t) - H \sum_{k=1}^{n} B_k x_k \right)^T \cdot R_S^{-1} \left( A^o_t (H x_t + \epsilon_t) - H \sum_{k=1}^{n} B_k x_k \right)$$

(17)
Now, $J_S$ is written with the observed slopes $\text{diag}(A_t')$ between mixing ratios and the unknown $\text{diag}(B_t)$ ratios between inventories. 

To estimate the diagonal elements in $R_S$, we estimate the $i$th element of standard deviation in $\varepsilon_t$, $\delta_S$, based on the standard deviation in the tracer mixing ratio $\varepsilon_t$, $\delta_t$, and the uncertainty $\sigma_t$ associated with the slope $a_t'$:

\[
y_t'(i) = a_t'(i)y_t(i)
\]

Let’s assume that $y_t'(i) \approx a_t'(i)y_t'(i)$

Then

\[
y_t'(i) - y_t'(i) \approx a_t'(i)(y_t'(i) - y_t'(i))
\]

\[
\varepsilon_S(i) = a_t'(i)\varepsilon_t(i) \quad (18)
\]

$\varepsilon_S$, $\sigma_a$, and $\varepsilon_t$ are uncorrelated variables. Hence, based on error propagation of the product of two uncorrelated variables, the $i$th element of $\delta_S$ can be written:

\[
\left( \frac{\delta_S(i)}{\varepsilon_S(i)} \right)^2 = \left( \frac{\sigma_S(i)}{a_t'(i)} \right)^2 + \left( \frac{\delta_t(i)}{\varepsilon_S(i)} \right)^2
\]

\[
\delta_S(i) = \varepsilon_S(i)^2 \left( \frac{\sigma_S(i)}{a_t'(i)} \right)^2 + \varepsilon_S(i)^2 \left( \frac{\delta_t(i)}{\varepsilon_S(i)} \right)^2
\]

\[
\delta_S(i)^2 = \varepsilon_S(i)^2 \sigma_S(i)^2 \quad (19)
\]

$\varepsilon_t$, $\sigma_a$, and $\varepsilon_t$ can be estimated by taking the difference between the observed and simulated mixing ratios of tracer $t$ using an optimized inventory $x_t$. $\sigma_S(i)$ is calculated using equation (10). $a_t'(i)$ is measured. $\delta_t(i)^2$ is equal to the sum of the variance in the measurements and the variance of the simulated mixing ratio of tracer $t$ by the model. Brioude et al. [2011] used 3 transport models to partially assess the variance from models. We used these results to calculate $\delta_t(i)^2$.

If such an estimate is unavailable, one can assume a constant uncertainty from the model. We assume that the covariance matrix of the uncertainty in the observations $R_S$ is diagonal and each element is equal to $\delta_S(i)^2$. We use the notation $R_t$ when $\delta_S(i)^2$ is calculated using information from tracer $t$.

$J_S$ is minimum when the Jacobian of $J_S$ is a null vector. As explained before (see discussion of equation (12)), the $x_t$ are the input variables of $J_S$ while $B_t$ are parameters. Therefore,

\[
\nabla_j J = 0 \quad \forall t \in [1,n], \left( \frac{\partial J}{\partial x_t} \right) = 0 \quad (20)
\]

For each partial derivative against $x_t$, we write $J_S$ using equation (17)

\[
\forall t \in [1,n], \left( \frac{\partial J}{\partial x_t} \right) = \frac{\partial}{\partial x_t} \left( \frac{1}{2} \left( a_t'(Hx_t + \varepsilon_t) - H \sum_{k=1}^{n} B_k x_k \right)^T \right)
\]

\[
\cdot R^{-1}_t \left( a_t'(Hx_t + \varepsilon_t) - H \sum_{k=1}^{n} B_k x_k \right) = 0 \quad (21)
\]

Based on equation (16), we rewrite for simplicity:

\[
a_t'(Hx_t + \varepsilon_t) = y_t'
\]

Hence, equation (21) becomes

\[
\forall t \in [1,n], \left( (A_t'H)^T - (HB_t)^T \right) R^{-1}_t \left( y_t' - H \sum_{k=1}^{n} B_k x_k \right) = 0
\]

\[
\forall t \in [1,n], \left( (A_t'H)^T - (HB_t)^T \right) R^{-1}_t y_t' - H \sum_{k=1}^{n} B_k x_k = 0
\]

\[
\forall t \in [1,n], \left( (A_t'H)^T R^{-1}_t y_t' - (A_t'H)^T R^{-1}_t H \sum_{k=1}^{n} B_k x_k \right) = 0
\]

\[
\forall t \in [1,n], \left( (A_t'H)^T R^{-1}_t y_t' - (A_t'H)^T R^{-1}_t H \sum_{k=1}^{n} B_k x_k \right) = 0
\]

In our study, $x_t$ is calculated using one tracer (either CO, NOy or SO2), two tracers (CO and NOy), or 3 tracers (CO, NOy and SO2). The chemical background for each chemical species is subtracted for each flight, and the slopes $a_{CO}', a_{NOy}'$, and $a_{SO2}'$ are calculated for each FLEXPART box along the flight path by simply dividing the 20-s average CO2 mixing ratio with the 20-s average mixing ratios of CO, NOy and SO2.

Each element in the $n$ different $B_k$ can be solved independently. For instance, using 2 tracers (CO and NOy), equation (22) can be solved for each jth element $b_{CO}$ and $b_{NOy}$ in $\text{diag}(B_{CO})$ and $\text{diag}(B_{NOy})$ as:

\[
\frac{\partial J}{\partial B_{CO}^j} = \left( (A_{CO,H})^T R_{CO}^{-1} y_{CO} \right) - \left( (A_{CO,H})^T R_{CO}^{-1} H_{CO} x_t \right) b_{CO} = \left( (A_{CO,H})^T R_{CO}^{-1} H_{CO} x_t \right) b_{CO}
\]

\[
\frac{\partial J}{\partial B_{NOy}^j} = \left( (A_{NOy,H})^T R_{NOy}^{-1} y_{NOy} \right) - \left( (A_{NOy,H})^T R_{NOy}^{-1} H_{NOy} x_t \right) b_{NOy} = \left( (A_{NOy,H})^T R_{NOy}^{-1} H_{NOy} x_t \right) b_{NOy}
\]

\[
\frac{\partial J}{\partial B_{SO2}^j} = \left( (A_{SO2,H})^T R_{SO2}^{-1} y_{SO2} \right) - \left( (A_{SO2,H})^T R_{SO2}^{-1} H_{SO2} x_t \right) b_{SO2} = \left( (A_{SO2,H})^T R_{SO2}^{-1} H_{SO2} x_t \right) b_{SO2}
\]

\[
\frac{\partial J}{\partial B_{CO}^j} = \left( (A_{CO,H})^T R_{CO}^{-1} y_{CO} \right) - \left( (A_{CO,H})^T R_{CO}^{-1} H_{CO} x_t \right) b_{CO} = \left( (A_{CO,H})^T R_{CO}^{-1} H_{CO} x_t \right) b_{CO}
\]

\[
\frac{\partial J}{\partial B_{NOy}^j} = \left( (A_{NOy,H})^T R_{NOy}^{-1} y_{NOy} \right) - \left( (A_{NOy,H})^T R_{NOy}^{-1} H_{NOy} x_t \right) b_{NOy} = \left( (A_{NOy,H})^T R_{NOy}^{-1} H_{NOy} x_t \right) b_{NOy}
\]

\[
\frac{\partial J}{\partial B_{SO2}^j} = \left( (A_{SO2,H})^T R_{SO2}^{-1} y_{SO2} \right) - \left( (A_{SO2,H})^T R_{SO2}^{-1} H_{SO2} x_t \right) b_{SO2} = \left( (A_{SO2,H})^T R_{SO2}^{-1} H_{SO2} x_t \right) b_{SO2}
\]
Rearranging the terms, the system of equations can be written:

\[
\begin{align*}
\frac{\partial J}{\partial x_{CO,j}} &= \left( (ACO_H)R_{COj}^{1+} \right)_{CO} - \left( (ACO_H)R_{COj}^{1+} \right)_{HxCO} + \left( (ACO_H)R_{COj}^{1+} \right)_{HxNO} - \left( (ACO_H)R_{COj}^{1+} \right)_{HxSO} - \left( (ACO_H)R_{COj}^{1+} \right)_{HxCO} + \left( (ACO_H)R_{COj}^{1+} \right)_{HxNO} + \left( (ACO_H)R_{COj}^{1+} \right)_{HxSO} + \left( (ACO_H)R_{COj}^{1+} \right)_{HxCO} + \left( (ACO_H)R_{COj}^{1+} \right)_{HxNO} + \left( (ACO_H)R_{COj}^{1+} \right)_{HxSO} + \left( (ACO_H)R_{COj}^{1+} \right)_{HxCO} + \left( (ACO_H)R_{COj}^{1+} \right)_{HxNO} + \left( (ACO_H)R_{COj}^{1+} \right)_{HxSO} = 0 \\
\frac{\partial J}{\partial x_{NOj}} &= \left( (ACO_H)R_{NOj}^{1+} \right)_{NOy} - \left( (ACO_H)R_{NOj}^{1+} \right)_{HxCO} + \left( (ACO_H)R_{NOj}^{1+} \right)_{HxNO} + \left( (ACO_H)R_{NOj}^{1+} \right)_{HxSO} + \left( (ACO_H)R_{NOj}^{1+} \right)_{HxCO} + \left( (ACO_H)R_{NOj}^{1+} \right)_{HxNO} + \left( (ACO_H)R_{NOj}^{1+} \right)_{HxSO} + \left( (ACO_H)R_{NOj}^{1+} \right)_{HxCO} + \left( (ACO_H)R_{NOj}^{1+} \right)_{HxNO} + \left( (ACO_H)R_{NOj}^{1+} \right)_{HxSO} + \left( (ACO_H)R_{NOj}^{1+} \right)_{HxCO} + \left( (ACO_H)R_{NOj}^{1+} \right)_{HxNO} + \left( (ACO_H)R_{NOj}^{1+} \right)_{HxSO} + \left( (ACO_H)R_{NOj}^{1+} \right)_{HxCO} + \left( (ACO_H)R_{NOj}^{1+} \right)_{HxNO} + \left( (ACO_H)R_{NOj}^{1+} \right)_{HxSO} = 0 \\
\frac{\partial J}{\partial x_{SOj}} &= \left( (ACO_H)R_{SOj}^{1+} \right)_{SO2} - \left( (ACO_H)R_{SOj}^{1+} \right)_{HxCO} + \left( (ACO_H)R_{SOj}^{1+} \right)_{HxNO} + \left( (ACO_H)R_{SOj}^{1+} \right)_{HxSO} + \left( (ACO_H)R_{SOj}^{1+} \right)_{HxCO} + \left( (ACO_H)R_{SOj}^{1+} \right)_{HxNO} + \left( (ACO_H)R_{SOj}^{1+} \right)_{HxSO} + \left( (ACO_H)R_{SOj}^{1+} \right)_{HxCO} + \left( (ACO_H)R_{SOj}^{1+} \right)_{HxNO} + \left( (ACO_H)R_{SOj}^{1+} \right)_{HxSO} + \left( (ACO_H)R_{SOj}^{1+} \right)_{HxCO} + \left( (ACO_H)R_{SOj}^{1+} \right)_{HxNO} + \left( (ACO_H)R_{SOj}^{1+} \right)_{HxSO} + \left( (ACO_H)R_{SOj}^{1+} \right)_{HxCO} + \left( (ACO_H)R_{SOj}^{1+} \right)_{HxNO} + \left( (ACO_H)R_{SOj}^{1+} \right)_{HxSO} = 0
\end{align*}
\]

Each element at position \(j\) in \(b_{CO}\) and \(b_{NOy}\) can be calculated by solving a system of nonlinear equations of the form

\[
F(x, y) = 0 \Rightarrow \begin{cases} 
\alpha_1 + \alpha_2x + \alpha_3y + \alpha_4xy + \alpha_5xy = 0 \\
\beta_1 + \beta_2x + \beta_3y + \beta_4xy + \beta_5xy = 0 \\
\gamma_1 + \gamma_2x + \gamma_3y + \gamma_4xy + \gamma_5xy + \gamma_6xy + \gamma_7xy = 0
\end{cases}
\]

where

\[
x = b_{CO} \\
y = b_{NOy}
\]

Since an exact solution is not necessarily possible, we instead find the minimum of the norm \(F(x, y)\) using a built-in function that solves nonlinear least squares problems (the function \(lsqnonlin\) in the software Matlab was used in this study).

[28] Using 3 tracers (CO, NOy and SO2), equation (22) becomes

\[
B_R^i H^j R_{ij}^{-1} \left( y_{ij} - H \sum_{k=1}^{n} B_k x_k \right) = 0
\]

Because \(H\) has associated uncertainties, the vector \(y_{ij}\) is never equal to the vector \(H \sum_{k=1}^{n} B_k x_k\). Thus \(B_k\) must be equal to 0. Hence, if a tracer \(t\) has no correlation with \(S\), no information will be added to the system of equations. Similarly, adding a tracer that is a linear combination of other tracers in the system of equations will not add any information or bias the solution.

Equation (26) is a nonlinear system of equations of the form

\[
F(x, y, z) = 0 \Rightarrow \begin{cases} 
\alpha_1 + \alpha_2x + \alpha_3y + \alpha_4zx + \alpha_5zy + \alpha_6zx + \alpha_7zy = 0 \\
\beta_1 + \beta_2x + \beta_3y + \beta_4zx + \beta_5zy + \beta_6zx + \beta_7zy = 0 \\
\gamma_1 + \gamma_2x + \gamma_3y + \gamma_4zx + \gamma_5zy + \gamma_6zx + \gamma_7zy + \gamma_8zx = 0
\end{cases}
\]

The elements in \(b_{CO}\), \(b_{NOy}\), and \(b_{SO2}\) are calculated by finding the minimum of the norm \(F(x, y, z)\). The success of this method will depend on how well the model \(H\) and how variable are the time series of the slopes \(a_{CO}, a_{NOy}\) and \(a_{SO2}\).

[29] This method can be used with any number of tracers. If a tracer \(t\) has no correlation with the species \(S\), then \(a_t\) equals zero or is negligible. Therefore equation (22) becomes:

\[
B_R^i H^j R_{ij}^{-1} \left( y_{ij} - H \sum_{k=1}^{n} B_k x_k \right) = 0
\]

3. Validation of the Flux Ratio Inversion Method

3.1. Construction of a CO2 Prior

[30] The flux ratio inversion method was first applied to estimating CO2 emissions in Houston using a single tracer. Figure 3 shows the optimized priors (left panels) of CO, NOy and SO2 from Brioude et al. [2011] that are used in the method. CO is mainly emitted by urban sources, particularly on-road vehicles, with large emissions in the Houston urban area and Ship Channel. NOy is mainly emitted by on-road sources throughout the city and industry facilities in the Ship Channel. SO2 is emitted mainly by industry facilities, with
most of the emissions from a number of large sources in the Ship Channel. The constructed CO\textsubscript{2} flux using each single tracer (Figure 3, middle) has similar spatial distribution as the tracer used in the method. Each CO\textsubscript{2} inventory is an average based on 100 realizations per transport model applying a random term on the slopes that are used in the flux ratio inversion method. The CO\textsubscript{2} flux constructed using CO has large emissions in the urban area, that constructed using NO\textsubscript{y} has large emissions in the Ship Channel, and that constructed from SO\textsubscript{2} has large emissions from point sources in the Ship Channel. Even though each constructed CO\textsubscript{2} inventory has a different surface distribution, they do have similar absolute magnitudes, with a standard deviation varying from 0.11 to 0.43 × 10\textsuperscript{-6} kg/m\textsuperscript{2}/s. Based on the ensemble of the results with each of the tracers, the uncertainty in CO\textsubscript{2} flux is about 20\% in the Houston urban area.

It is interesting to assess how the measured slopes between tracer mixing ratios are converted into surface flux ratios by the flux ratio inversion method. Figure 4 presents the flux ratio between each CO\textsubscript{2} inventory constructed using a single tracer and the tracer itself (CO, NO\textsubscript{y} or SO\textsubscript{2}). The geographical distribution of flux ratio (CO/CO\textsubscript{2}, NO\textsubscript{y}/CO\textsubscript{2}, SO\textsubscript{2}/CO\textsubscript{2}) is non-uniform. CO/CO\textsubscript{2} and NO\textsubscript{y}/CO\textsubscript{2} have larger ratios in the urban area than in the Ship Channel, implying that NO\textsubscript{y} or CO emissions per unit of CO\textsubscript{2} emitted (or per mass of fuel burned) in the urban area are higher than in the Ship Channel. The SO\textsubscript{2}/CO\textsubscript{2} flux ratio has a maximum near the Parish Power Plant southwest of the urban center. Figures 3 and 4 confirm that the flux ratio inversion method is able to use the information available in the measured slopes and existing priors of CO, NO\textsubscript{y} and SO\textsubscript{2} to constrain surface flux ratios with CO\textsubscript{2} in Houston.

Figure 5 shows the average constructed CO\textsubscript{2} emission inventory using 2 (Figure 5a) and 3 tracers (Figure 5c) in the flux ratio inversion method. Using 2 and 3 tracers, the CO\textsubscript{2} flux converged to a solution with larger emissions in the Ship Channel, where most of the large industrial sources are located, than in the urban area. The CO\textsubscript{2} fluxes are negligible at the border of the domain. The uncertainties are smaller in regions with large anthropogenic emissions. The average uncertainty in the constructed CO\textsubscript{2} emissions in the Houston urban area and Ship Channel is about 30 to 40\% (Figures 5b and 5d). The CO\textsubscript{2} emission inventory constructed with 3 tracers is theoretically the most precise among the CO\textsubscript{2} emission inventories in Figures 3 and 5. The spatial distribution and absolute magnitudes of the constructed CO\textsubscript{2} emissions are now in better agreement with Vulcan than the CO\textsubscript{2} inventory calculated without a prior (compare Figures 2 and 5).

3.2. Sensitivity to Biospheric CO\textsubscript{2} Flux

In section 3.1, we constructed an anthropogenic CO\textsubscript{2} emission inventory based on anthropogenic tracers. However, the CO\textsubscript{2} fluxes at the surface should also include impacts from biogenic sources in the form of uptake and ecosystem respiration. In each 4 × 4 km\textsuperscript{2} grid cell, vegetation could potentially be collocated with anthropogenic sources. In this section, we present sensitivity tests using the CO\textsubscript{2} emission inventory calculated with 3 tracers in section 3.1 (Figure 5c).

In the first test, we prescribed an uptake CO\textsubscript{2} flux with a magnitude of −2 × 10\textsuperscript{-6} kg/m\textsuperscript{2}/s, which is roughly the magnitude of anthropogenic emissions, in a portion of the urban area region (Figure 6a) of the constructed CO\textsubscript{2} inventory using 3 tracers (Figure 5c). After adding this uptake flux, the average constructed CO\textsubscript{2} emissions in the urban area were reduced by 50\%, and by 16\% in the Greater Houston region. We call this inventory the “true CO\textsubscript{2} inventory with uptake” in subsequent tests. The optimized CO, NO\textsubscript{y} and SO\textsubscript{2} inventories were unmodified. This sensitivity test used an unrealistically large uptake flux, and the error on the constructed CO\textsubscript{2} flux should not be used as an uncertainty estimate of the flux ratio inversion method. We calculated hypothetical observations of CO\textsubscript{2}, CO, NO\textsubscript{y} and SO\textsubscript{2} using the transport model output (H), the CO\textsubscript{2} emission inventory calculated in section 3.1 and including the area where uptake dominates, and the optimized tracer emission inventories for CO, NO\textsubscript{y} and SO\textsubscript{2}. We then applied the flux ratio inversion method using the model and the hypothetical
observations to construct a CO2 surface flux. Therefore, our sensitivity test was valid assuming perfect transport models.

After applying the flux ratio inversion method, the new constructed CO2 flux doesn’t show any area where uptake dominates (Figure 6b). In this case, the CO2 flux in the urban region is overestimated by 90% (Table 1), while the CO2 flux in the Ship Channel is underestimated by 30%, due to the method’s uncertainty. The average CO2 flux in the Greater Houston region is underestimated by 10%. The grid cells in the urban region that are not associated with the uptake are slightly modified. Some grid cells have larger CO2 emissions than the true CO2 inventory with uptake. Uptake areas, being a biogenic source, cannot be resolved using anthropogenic tracers in the flux ratio inversion method. Because there is an overestimation in the urban area region and an underestimation in the Ship Channel, the average difference in the Greater Houston area is fairly small (±10%). It seems that the uncertainty in the flux ratio inversion method due to uptake is redistributed between the urban area region and the Ship Channel.

A second test used the newly constructed CO2 flux (with larger uncertainties in the urban area and Ship Channel due to uptake) as a prior to run a classic inversion using the hypothetically measured CO2 mixing ratio. We used a set of measurements independent of the one used to calculate the CO2 prior. We assembled an ensemble of 100 solutions by applying a random term on the prior and using a random subset of observations to reduce the error correlation in the posterior. The mean solution (Figure 6c and Table 1) shows large differences with the prior in the urban area. The area where uptake dominates is clear in the CO2 posterior, and the magnitude of the uptake is in good agreement with the true CO2 emission inventory with uptake. Compared to the truth, the posterior CO2 flux agrees extremely well in the Greater Houston region (essentially no difference), urban area (within 1%) and Ship Channel (within 3%). This test shows that even though an area where uptake dominates can increase the uncertainty in a constructed anthropogenic CO2 inventory obtained by the flux ratio inversion method, the constructed CO2 flux can be adequately used as a prior in a classic inversion to obtain a realistic optimized CO2 posterior. Furthermore, the sensitivity of the constructed anthropogenic CO2 inventory to biogenic fluxes can be estimated by comparing its anthropogenic mean flux in the Greater Houston, Urban area and Ship Channel to those found in the CO2 posterior inventory.

4. Results: Anthropogenic CO2 Estimate and Comparison With Vulcan Emission Inventory

To estimate the sensitivity of the flux ratio inversion method to biogenic fluxes, we applied a classic inversion method.
method to the constructed anthropogenic CO₂ emission inventory calculated in section 3.1 with 3 tracers (Figure 5c).

We used an independent set of observations from those used to calculate the anthropogenic CO₂ inventory with the flux ratio inversion, by using half of the measurement data set to calculate the posterior CO₂ emission inventory, the second half of the measurement data set being used in the flux ratio inversion method. Because the covariance matrix of the observations is unknown, we applied the L-curve criterion [Henze et al., 2009; Brioude et al., 2011] to estimate the best balance between the covariance matrices of the observations and the prior, rather than making an arbitrary choice. Using the L-curve criterion, we found that the best choice was to use the same relative uncertainty in both covariance matrices (not shown). We calculated an ensemble of 100 solutions for each of the 3 transport models by using a random term on the prior and using a random subset of observations to reduce the cross correlation of the error in the posterior. The uncertainty estimate of the results in Table 2 was based on the variability of the 300 solutions.

[38] The posterior total CO₂ flux inventory (Figure 7d) changes significantly compared to the anthropogenic CO₂ prior calculated by the flux ratio inversion method (Figure 7c). The average total CO₂ flux in the posterior is lower in the urban area and the Ship Channel by 19% (−0.15 × 10⁻⁶ kg/m²/s) and 6% (−0.12 × 10⁻⁶ kg/m²/s), respectively, than the constructed anthropogenic CO₂ prior. Grid cells with negligible anthropogenic emission are not used in the inversion, and are likely to have significant biogenic fluxes, with dominant uptake during daytime. Removing those grid cells from the classic inversion may increase uncertainties in the CO₂ emissions estimate in the posterior inventory. The spatial distribution of CO₂ fluxes in the Ship Channel has also changed, with the largest fluxes being shifted to the middle of the Ship Channel instead of the western part. The spatial distribution around the Ship Channel is now closer to the 10 × 10 km² Vulcan anthropogenic CO₂ inventory. Areas where uptake dominates are found south of the urban area and in a few patches in the northern part of the city. The Normalized Difference Vegetation Index (NDVI) from the Moderate resolution Imaging Spectroradiometer (MODIS) is used as a crude estimate of vegetation uptake. A map of NDVI (Figure 7b) shows that those areas where uptake dominates are associated with higher NDVI than the main part of the urban area. The CO, NOy and SO₂ fluxes from those grid cells are negligible, confirming that those grid cells contain few anthropogenic sources, which is to be expected from areas dominated by biogenic uptake.

[39] To validate the CO₂ fluxes in the constructed inventory from the flux ratio inversion and in the posterior, we compared the average CO₂ fluxes to the values from the 10 × 10 km² Vulcan anthropogenic CO₂ emission inventory for the 3 different regions defined on the map (Table 2 and Figure 7). The 10 × 10 km² grid cells were linearly interpolated on to the regular latitude-longitude grid used in the inversion to estimate the average Vulcan anthropogenic CO₂ fluxes in the Houston urban area and Ship Channel regions. The constructed, posterior and Vulcan inventories agree on the spatial distribution of CO₂ fluxes, with large emissions in the Ship Channel, significant emissions in the urban area, and fluxes close to zero at the border of the urban area of Houston. Compared to Vulcan, the average total CO₂ flux in the posterior is lower by 2% ± 4% in the Ship Channel, while the average anthropogenic CO₂ flux in the constructed inventory is higher by 4% ± 7%. The uncertainty estimates represent the 95% confidence interval on the difference between the mean flux in Vulcan and the mean fluxes in the posterior or constructed CO₂ inventory based on a statistical significance in a t-test. Those differences are statistically insignificant. The differences for posterior – Vulcan and
Table 1. Average Surface CO2 Flux From the Greater Houston, the Houston Urban Area and the Houston Ship Channel Regions From a Hypothetical True CO2 Emission Inventory With Biogenic Uptake (Figure 6)*

<table>
<thead>
<tr>
<th></th>
<th>Hypothetical True CO2 Emission Inventory With Uptake (10^-6 kg/s/m²)</th>
<th>Constructed CO2 With Uptake</th>
<th>Posterior Using Constructed CO2 With Uptake as a Prior</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Emission</td>
<td>Difference</td>
<td>Emission</td>
</tr>
<tr>
<td></td>
<td>(10^-6 kg/s/m²)</td>
<td>(10^-6 kg/s/m²)</td>
<td>(10^-6 kg/s/m²)</td>
</tr>
<tr>
<td>Greater Houston</td>
<td>0.364</td>
<td>0.326</td>
<td>−10%</td>
</tr>
<tr>
<td>Urban area</td>
<td>0.447</td>
<td>0.847</td>
<td>−90%</td>
</tr>
<tr>
<td>Ship channel</td>
<td>2.25</td>
<td>1.58</td>
<td>−30%</td>
</tr>
</tbody>
</table>

*See Figure 2 (bottom) for definition of these regions. Also shown are the emissions for these regions from the constructed CO2 inventory from the flux ratio inversion method and from a CO2 posterior inventory using the constructed CO2 inventory as a prior, as well as the differences in these two inventories from the hypothetical true inventory in the left-hand column. See section 3.2 for details.

constructed – Vulcan are +16% ± 4% and +37% ± 6%, respectively, in the urban area, and +19% ± 6% and +23% ± 10%, respectively, in the Greater Houston region.

[40] The total CO2 fluxes in the posterior are not exclusively coming from anthropogenic sources. They include not only anthropogenic sources but also uptake by vegetation and respiration. It is necessary to estimate the biogenic fluxes and remove them from the posterior total CO2 fluxes to compare anthropogenic emission estimates from the posterior, the constructed inventory and Vulcan. The average NDVI index measured by MODIS (Figure 7b) is equal to 0.33 in the Ship Channel, 0.38 in the urban area region and 0.41 in the Houston region. The average uptake calculated in the CO2 posterior is associated with negligible anthropogenic sources and is equal to −0.06 × 10^-6 kg/m²/s with an average NDVI of 0.47. According to the 1 × 1° NOAA assimilation system CarbonTracker [Peters et al., 2007], the average daytime biogenic uptake of CO2 in the Greater Houston region is −0.08 × 10^-6 kg/m²/s. However, the 1 × 1° grid cell in CarbonTracker is roughly centered over Houston and includes a significant fraction of land surface outside urban areas and is probably not adapted for biogenic fluxes in urban areas. Based on the analysis of CO2-CO scatterplots of the flight on September 26th (not shown), we estimated that the strongest uptake reduces the CO2 concentration by 1.2 ppmv, without coincident CO variation, while the urban emissions increase the CO2 concentration by 7 ppmv, so uptake is equal to ~15% of the anthropogenic urban emissions. Based on our flux estimate in the constructed anthropogenic CO2 inventory, this percentage translates into a surface uptake flux of −0.15 × 10^-6 kg/s/m². Interestingly, this estimate is also equal to the average difference between the average anthropogenic CO2 flux in the anthropogenic CO2 inventory calculated by the flux ratio inversion and the average total CO2 flux from the posterior CO2 inventory in the Houston urban area region. Each of those uptake flux estimates is associated with uncertainties. Because uptake fluxes are expected to be heterogeneous in an urban area, we estimate that the uncertainty due to surface uptake flux in our estimates of anthropogenic CO2 flux in the posterior ranges from −0.06 to −0.15 × 10^-6 kg/m²/s. This translates into a difference between Vulcan and the anthropogenic CO2 posterior of +24% to +37% in the urban area. In the Ship Channel, the difference varies from +1% to +5% and is still statistically insignificant. Compared to the constructed anthropogenic CO2 inventory, the anthropogenic CO2 posterior inventory is different by 0% to −10% in the urban area, and between −3% to +1% in the Ship Channel. The constructed anthropogenic CO2 inventory is in agreement with the anthropogenic CO2 posterior inventory. These results show that the emission estimates from the flux ratio inversion method are robust.

[41] The average flux values from the Greater Houston and urban area regions are typical of emissions from other cities [Helfter et al., 2011; Ramamurthy and Pardyjak, 2011, and references therein]. But to date, the average CO2 flux found in the Ship Channel (a surface area of 288 km² in our analysis) with our top-down analysis is the largest reported in the literature.

[42] The discrepancies between Vulcan and the constructed or the posterior anthropogenic CO2 inventory (+37% and +30%, respectively, on average in the urban area, and

Table 2. Average Surface CO2 Flux From the Greater Houston, the Houston Urban Area and the Houston Ship Channel Regions From the Vulcan Anthropogenic CO2 Inventory, From the Constructed CO2 Inventory From the Flux Ratio Inversion Method With 3 Tracers, From the Total CO2 Posterior Flux Calculated Using the Constructed CO2 Inventory From the Flux Ratio Inversion Method With 3 Tracers as a Prior, and a Range of Anthropogenic CO2 Flux Values Depending on Estimates of Uptake, Along With The Differences in the Posterior Inventories With Vulcan

<table>
<thead>
<tr>
<th>Daytime CO2 Surface Fluxes</th>
<th>Vulcan Anthropogenic CO2 Emission Calculated by the Flux Ratio Inversion Method</th>
<th>Anthropogenic CO2 Posterior Emission with Estimate of Values of Uptake</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Emission (10^-6 kg/s/m²)</td>
<td>Relative Difference</td>
</tr>
<tr>
<td>Greater Houston</td>
<td>0.31</td>
<td>0.38 ± 0.03</td>
</tr>
<tr>
<td>Urban area</td>
<td>~0.7</td>
<td>0.96 ± 0.04</td>
</tr>
<tr>
<td>Ship channel</td>
<td>~2.17</td>
<td>2.25 ± 0.15</td>
</tr>
</tbody>
</table>

...
+23% and +54% in the Greater Houston) are probably due to emission factor uncertainties and uncertainties in the inversion method. For instance, previous studies have found that natural gas combustion in urban areas was larger than reported in surface inventories [Pataki et al., 2009]. The discrepancies could also come from differences in emissions between 2002, the reference year in Vulcan, and 2006, the year of the observations used in our calculations. For example,

Table 3. Differences Between the Measured CO₂ Mixing Ratio and That Simulated by FLEXPART Using the Vulcan Anthropogenic CO₂ Inventory, Using the Constructed CO₂ Inventory Derived From the Flux Ratio Inversion Method With 3 Tracers, and Using the CO₂ Posterior Inventory

<table>
<thead>
<tr>
<th></th>
<th>Vulcan Emission Inventory</th>
<th>Constructed CO₂ Inventory by the Flux Ratio Inversion Method</th>
<th>CO₂ Posterior Inventory</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average error with measured CO₂ mixing ratio</td>
<td>-2.21 ± 0.3 ppm</td>
<td>-0.9 ± 0.3 ppm</td>
<td>-0.3 ± 0.1 ppm</td>
</tr>
<tr>
<td>Linear correlation</td>
<td>0.38</td>
<td>0.50</td>
<td>0.57</td>
</tr>
</tbody>
</table>
according to Federal Highway Administration on-road fuel use statistics, the consumption of gasoline in Texas increased by 8% from 2002 to 2006. Assuming fuel consumption also increased in other sectors of the economy between 2002 and 2006, it is reasonable that the constructed or posterior CO₂ emissions in Houston are somewhat larger than those of Vulcan.

[45] To confirm that the flux ratio inversion method can be used to estimate CO₂ anthropogenic emissions in Houston, we present the errors (based on the ensemble of 3 transport models) between the measured CO₂ mixing ratio (above background) and that simulated by FLEXPART (Table 3). Using the Vulcan anthropogenic CO₂ emission inventory, the average error between simulated and measured CO₂ mixing ratios is $-2.21 \pm 0.3$ ppmv (the linear correlation coefficient, r, is 0.38). Using the constructed anthropogenic CO₂ inventory obtained from the flux ratio inversion method, the average simulated-measured error is reduced to $-0.9 \pm 0.3$ ppmv (r = 0.50). After applying the classic inversion with the constructed CO₂ as a prior, the average error is reduced further, to $-0.3 \pm 0.1$ ppmv (r = 0.57). The average error using Vulcan is larger than those using the constructed anthropogenic CO₂ inventory and the total flux CO₂ posterior mostly because urban emissions are underestimated and the spatial distribution of surface fluxes are uncertain. The average error is significantly lower using the posterior than using the constructed anthropogenic CO₂ inventory by the flux ratio inversion because uncertainties inherent to the flux ratio inversion (inversion of slopes, assumption that CO, NOy and SO₂ inventories are optimum) are significantly reduced.

[46] These results show that the anthropogenic CO₂ emission inventory calculated by the flux ratio inversion method has a similar spatial distribution and magnitude of emission as a bottom-up inventory like Vulcan, and that one can calculate an inventory at mesoscale without a prior by a top-down method if existing information from other tracers is available.

5. Conclusions

[45] A new inversion method that allows the calculation at mesoscale of a surface emission inventory without a prior has been successfully applied to calculate an anthropogenic CO₂ emission inventory over Houston. The new inversion method, which we call “flux ratio inversion,” constrains the ratios of surface emissions between a given chemical species and various chemical tracers by using the measured ratios of those chemical species’ atmospheric concentrations. The larger the number of tracers used, the more precise the constructed emission inventory derived. We used in situ measurements of CO₂, CO, NOy and SO₂ from the NOAA WP-3 aircraft during the TexAQS 2006 campaign to calculate an anthropogenic CO₂ emission inventory, based on optimized emission inventories of CO, NOy, and SO₂.

[46] Sensitivity tests showed that a classic inversion, using CO₂ in situ measurements and the CO₂ emission inventory calculated by the flux ratio inversion as a prior, is necessary to 1) estimate the numerical uncertainty of the flux ratio inversion method, and 2) estimate uptake fluxes where biospheric sources are predominant. In our example, only anthropogenic tracers were used to calculate CO₂ fluxes via the flux ratio inversion method, but CO₂ biogenic uptake fluxes are not represented by anthropogenic tracers. After applying a classic inversion using the anthropogenic CO₂ inventory calculated by the flux ratio inversion as a prior, the total CO₂ posterior fluxes were 19% lower in the Houston urban area and 6% lower in the Ship Channel than in the constructed anthropogenic CO₂ prior inventory. The CO₂ posterior showed net uptake in areas where anthropogenic sources are negligible. Ideally, $^{14}$CO₂ measurements could more precisely constrain fossil fuel combustion sources, and the combination of $^{14}$CO₂ with other anthropogenic tracers would allow a direct inversion calculation of the anthropogenic CO₂ inventory. The spatial distribution of surface CO₂ flux also was significantly improved in the Ship Channel by the inversion.

[47] We estimated a range of uptake fluxes for Houston and removed them from the total CO₂ posterior fluxes. We found that the CO₂ emissions in the constructed anthropogenic CO₂ flux from the flux ratio inversion were in good agreement with the anthropogenic posterior CO₂ flux in the Houston urban area and the Ship Channel. This result confirms that the emission estimates of anthropogenic CO₂ emissions from the flux ratio inversion method are robust.

[48] The anthropogenic CO₂ emission estimates in the Ship Channel from the constructed anthropogenic CO₂ inventory were also in agreement with the anthropogenic Vulcan CO₂ emission inventory. The daytime anthropogenic CO₂ flux in the constructed inventory from the flux ratio inversion was higher than Vulcan by $+37\% \pm 6\%$ in the urban area and by $23\% \pm 10\%$ in the Greater Houston area. The differences in the urban area between the constructed inventory and Vulcan are attributed to emission factor uncertainties in Vulcan and to increases in anthropogenic surface emissions between 2002, the reference year in Vulcan, and 2006, the year of the TexAQS observations driving the inversion.

[49] We have proven that our new inversion method can be applied to the estimation of emission Inventories at mesoscale. We assert that this methodology can generally be used for any inversion problem that requires prior knowledge to constrain the spatial distribution of surface fluxes, as long as co-emitted tracers measurements with existing (and reliable) inventories are available.

[50] Acknowledgments. We would like to thank Gunnar Schade and two anonymous referees for their helpful comments. CarbonTracker results are provided by NOAA/ESRL Boulder, Colorado, USA, http://carbontracker.noaa.gov.

References


Fahey, D., G. Hübner, D. Parrish, E. Williams, R. Norton, B. Ridley, H. Singh, S. Liu, and F. Fehsenfeld (1986), Reactive nitrogen species in the troposphere: Measurements of NO, NO₂, HNO₃, particulate nitrate, peroxyacetyl nitrate (PAN), O₃, and total reactive odd nitrogen (NOₓ).


Parrish, D. D., et al. (2009), Overview of the second Texas Air Quality Study (TexAQS II) and the Gulf of Mexico Atmospheric Composition and Climate Study (GoMACCS), J. Geophys. Res., 114, D00F13, doi:10.1029/2008JD011842.


