CoCO2

Prototype system for a Copernicus CO_2 service

Benchmarking of plume detection and quantification methods

coco2-project.eu

CoCO2

Prototype system for a Copernicus CO_2 service

D4.4 Benchmarking of plume detection and quantification methods

Dissemination Level:	Public
Authors:	Janne Hakkarainen, Johanna Tamminen, Janne Nurmela, Hannakaisa Lindqvist (FMI); Diego Santaren, Grégoire Broquet, Frédéric Chevallier (LSCE); Erik Koene, Gerrit Kuhlmann, Dominik Brunner (Empa)
Date:	17/02/2023
Version:	1.0
Contractual Delivery Date:	31/12/2022
Work Package/ Task:	WP4/ T4.2
Document Owner:	FMI
Contributors:	FMI/LSCE/Empa
Status:	Final


ECMWF

Empa
Materials Science and Technology

ICOS


cea

Max Planck Institute
for Biogeochemistry

TNO innovation
for life

LUND UNIVERSITY

VU
VRIJE
UNIVERSITEIT
AMSTERDAM

WAGENINGEN
UNIVERSITY & RESEARCH

Barcelona
Supercomputing
Center
Centro Nacional de Supercomputación

CICERO

cmcc

cnrs

Deutscher Wetterdienst
Wetter und Klima aus einer Hand
DWD

École des Ponts
ParisTech

FMI

FORTH

iLab

MERCATOR
OCEAN
INTERNATIONAL

METEO
FRANCE


FCiências

THE CYPRUS
INSTITUTE

DLR
Deutsches Zentrum
für Luft- und Raumfahrt

CoCO2: Prototype system for a Copernicus CO_2 service

Coordination and Support Action (CSA)
H2020-IBA-SPACE-CHE2-2019 Copernicus evolution – Research activities in support of a European operational monitoring support capacity for fossil CO2 emissions

Project Coordinator: Dr Richard Engelen (ECMWF)
Project Start Date: 01/01/2021
Project Duration: 36 months

Published by the CoCO2 Consortium

Contact:
ECMWF, Shinfield Park, Reading, RG2 9AX,
richard.engelen@ecmwf.int

The CoCO2 project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 958927.

Figures

Tables

1 Executive Summary

This report synthetizes the benchmarking and the evaluation of the main current plume inversion methods that has been conducted in Task 4.2 of the CoCO2 project. The methods have been identified based on their ability to quantify the carbon dioxide (CO_2) emissions from satellite images of plumes downwind of sources with relatively low computation cost. The methods are either based on the analysis of individual plume images or on averaging the images over long periods of time.

Five computationally light methods for emission detection and quantification have been tested: the Cross-sectional flux method (CS), Gaussian Plume model fitting (GP), the Integrated Mass Enhancement method (IME), Light Cross Sectional flux methods (LCS) and the Divergence method (Div). The methods have been described in detail in Deliverable 4.3, and here we focus on their implementation and performance. The computation time for estimating emissions of 16 sources using one month of (cloud-free) data from three satellites varied between 30 – 120 min depending on the method, noting also that further optimization could most probably be made for operational processing. All these methods are already semi-automatic in the sense that they do not require specific parameterization to tackle new images, estimate uncertainties and implement automatic checks to avoid erroneous estimates by flagging bad data. They have been implemented in Python library for data-driven emission quantification (ddeq) and were applied to both simulated data sets and real observations covering temporally one full year.

Three datasets were used for analyzing the performance of the methods. The first and largest benchmarking setup consisted of the SMARTCARB dataset of synthetic satellite observations, which was produced to closely mimic the CO_2 and nitrogen dioxide (NO_2) observations of the upcoming Copernicus CO_2 Monitoring Mission (CO2M). A quantitative assessment of the methods has been based on this first application to synthetic images with the knowledge of the corresponding synthetic “true” emissions. Second, we also apply the methods to the library plumes generated in Task 4.1 of the CoCO2 project, and specifically focus on the high-resolution MicroHH model simulation of the Matimba power plant, where synthetic data including nitrogen oxides (NO_X) chemistry is available. Third, we have tested the methods on real NO_2 observations from the Copernicus Sentinel 5 Precursor TROPOMI instrument over the Matimba/Medupi power station in South Africa to estimate NO_X emissions. Real NO_2 observation were chosen to better mimic the future CO2M image data than current CO_2 observation systems. This third set of experiments support the evaluation of the robustness of the methods and of the statistics of uncertainties derived from the pseudo data experiments by tackling real data with additional sources of uncertainties and complexity. The report also reminds the extensive application of one of the methods to OCO-2 and OCO-3 XCO_2 observations, which is documented in other deliverables (D6.4, D6.5 and D6.6) but which also feeds the evaluation with real data.

According to the analysis presented here, each of the methods seemed to have their own strengths and weaknesses and none of the methods clearly outperformed (see Appendix for summary tables). In most realistic cloudy case, the median relative agreement of annual emissions of the 16 sources studied was -40% (GP), -23% (CS), -36% (IME) and 21% (LCS). LCS and GP methods were found overall most robust in SMARTCARB simulation cases and IME seemed to have the largest dispersion in emission estimates. Underestimation of the strongest emissions was typically observed for methods based on processing individual images (i.e., single image methods) especially for LCS (median relative bias around -40%). Such systematic underestimation of strong emission sources was not observed in the annual emission estimates of Div method. On the other hand, Div method succeeded to estimate

emissions of only 10 out of 16 sources, annual median relative agreement of 10 sources being 20% (Div).

MicroHH simulations and TROPOMI NO_2 real data case demonstrated the advantage of having a model-based conversion factor for NO_X-to-NO_2 instead of commonly used fixed value 1.32 when NO_X emissions are estimated. The simulated MicroHH tests revealed that the concept of effective wind probably requires re-visiting since systematic overestimation of NO_X and CO_2 emissions were observed.

2 Introduction

2.1 Background

The Paris Agreement, adopted in 2015, calls a need for monitoring anthropogenic greenhouse gas (GHG) emissions and space-based observations offer new opportunities for improving such assessments. Several space agencies have responded to this call by designing new carbon dioxide (CO_2) monitoring missions.

In Europe, ESA and EU have planned and developed the Copernicus CO_2 Monitoring (CO2M) mission to monitor anthropogenic CO_2 emissions (Janssens-Maenhout et al., 2020; Meijer et al., 2020). Currently, a constellation of up to three satellites is planned. The first two satellites are expected to be launched by 2026. In addition to CO_2, the CO2M instrument will measure nitrogen dioxide (NO_2) and methane (CH_4). The instrument will have a spatial resolution of 4 km^2 and an imaging swath of 250 km. To support achieving the strict accuracy requirements of the GHG measurements, dedicated aerosol and cloud instruments are added to the payload (Durand et al. 2022).

Several “plume detection/quantification” techniques for inferring emissions from satellite images or transects of the plumes downwind to the targeted sources have been proposed (e.g., Varon et al, 2018) and many of the methods (e.g., Gaussian plume method, cross-section flux method) have already been tested with real CO_2 space-based data mainly from NASA’s OCO-2 and OCO-3 satellites (e.g., Nassar et al., 2017, 2022; Reuter et al., 2019; Hakkarainen et al., 2021, 2023; Chevallier et al., 2022). Several simulation-based studies with synthetic CO_2 images exist as well (e.g., Kuhlmann et al., 2021).

Nitrogen dioxide (NO_2) is often co-emitted with carbon dioxide. In general, the NO_2 plumes are easier to detect than CO_2 plumes as the signal-to-noise ratio (SNR) is higher for NO_2 and NO_2 is less affected by clouds than CO_2. Moreover, the lifetime of NO_2 is short (only a few hours), which means NO_2 background concentrations are typically low compared to the plume signal. Thus, the observations of NO_2 can be used to guide the detection of CO_2 emission plumes and their shape (e.g., Kuhlmann et al., 2019; Reuter et al., 2019; Hakkarainen et al., 2021, 2023).

To estimate nitrogen oxides (NO_X) emissions from space-based NO_2 observations, methods based on the temporal averages of spatially co-located observations have often been applied (Fioletov et al., 2015; Beirle et al., 2011, 2019; de Foy et al., 2014). Historically, emission estimation methods based on temporal averaging have not yet been successfully applied to satellite-based CO_2 observations, although this option has been discussed by Hakkarainen et al. (2016) and Hill and Nassar (2019). During the CoCO2 project, we have adapted the divergence method, developed originally for NO_2 (Beirle et al., 2019, 2021) to estimate CO_2 emissions (Hakkarainen et al., 2022). This method will be also benchmarked here.

2.2 Scope of this deliverable

2.2.1 Objectives of this deliverable

The objective of this report is to benchmark the plume detection and emission quantification methods identified in Deliverable 4.3 (Koene et al., 2021).

2.2.2 Work performed in this deliverable

In order to support an efficient application and easy comparison, all methods were integrated in a common Python package **ddeq**. We have created dedicated benchmarking protocols for both the experiments with the synthetic and the real satellite images, run the benchmarking cases for all the methods and compiled the results. The benchmarking relies on the evaluation of the chosen methods with synthetic satellite observations for which the underlying emissions are known. For this task, we use the synthetic SMARTCARB dataset developed by Kuhlman et al. (2019) in the homonymous ESA-funded project. The benchmarking also relies on the application of the methods to real satellite data. Since the performance of the methods often rely on the process of wide (one to few hundred km swath-) images, and since there are no CO_2 observing satellites providing such wide images currently, we use the tropospheric NO_2 retrievals from the Copernicus Sentinel 5P/TROPOMI instrument. The intensive application of the LCS method to the OCO-2 and OCO-3 data (see D6.4, D6.5 and D6.6), does not directly participate to the benchmarking of the methods but it confirms the robustness of the application of the methods to real data.

2.2.3 Deviations and counter measures

Not applicable.

3 Methods for plume detection and emission quantification

In this section we describe the methods selected for the benchmarking. A more comprehensive overview is given in Deliverable 4.3 “Documentation of plume detection and quantification methods” by Koene et al. 2021. A short summary is given in **Error! Reference source not found.**.

Table 1 Summary of the methods for plume detection and emission quantification.

Method	Temporal treatment	Availability of joint CO_2 & NO_2 Imple-mentation (NO_2 supporting plume detection)	Background treatment and noise reduction	Method for quality indication	Computational cost (1)	Level of automatization (automatic up to known source locations/ fully automatic)
Cross-sectional flux method (CS)	Individual over-passes	Yes	Background estimation using sliding median filter and normalized convolution. No noise reduction.	Uncer-tainty of the emission estimation	Low 1 and 5 min plus 20 min for pre-processing incl. plume detection (2)	Automatic up to knowing source locations
Gaussian plume approach (GP)	Individual over-passes	Yes	Background estimation using sliding median filter and normalized convolution. No noise reduction.	Uncer-tainty of the emission estimation	Medium 42 and 90 min plus 20 min for pre-processing incl. plume detection (2)	Automatic up to knowing source locations & a priori estimate of emission rates

Integrated mass enhancement (IME)	Individual over-passes	Yes	Background estimation using sliding median filter and normalized convolution. No noise reduction.	Uncer-tainty of the emission estimation	Low 30 s plus 20 min for pre-processing incl. plume detection	Automatic up to knowing source locations
Light cross-sectional flux method (LCS)	Individual over-passes	Yes	Background estimation using a linear fit over 100-km wide windows. No noise reduction.	R^2 score of line densities	Low 3 and 9 min plus 3 min pre-processing	Automatic up to knowing source locations
Divergence method (DIV)	Applied to averages of overpasses over moths to years	No	For CO_2 local/regional background removal and denoising	Uncer-tainty of the emission estimation. MCMC for peak fitting. Additional visual inspection.	Low 3 min for divergence map, 1 s for optimization plus 20 min with MCMC sampling (chain length dependent)	Automatic up to knowing source locations

(1) Computation time was estimated by processing one month of cloud-free SMARTCARB orbits on the ICOS-CP Jupyter server (see Section 3.8 for details); (2) First number when using only CO_2 and second number when using both CO_2 and NO_2.

3.1 Cross-sectional (CS) flux method

3.1.1 General method

The mass continuity equation states that the net addition of mass through emission sources (E) and sinks (S), minus their temporal accumulation in a volume ($\hat{\rho}$) is balanced by the (net) mass flux over the boundaries of a domain. Considering the case of a vertically integrated column, the mass continuity equation can be shown to be exactly[1] equal to

$$E - S - \hat{\rho} = \oint_A \vec{F} \cdot \vec{n}\, dA,$$

with $\vec{F} = (\mathrm{TVCD} - \mathrm{BG})[U_{\mathrm{eff}} \;\; V_{\mathrm{eff}}]^T$, corresponding to the (enhanced mass) flux through any column with basal area differential dA, using the total vertical column density (TVCD) minus its background (BG), multiplied with the effective horizontal wind speeds U and V. For example, the east-west directed effective wind is defined as

$$U_{\mathrm{eff}} = \int_0^T C_0\, U\, dz,$$

where C_0 describes the vertical distribution of the gas concentrations in the column enhancement ($\mathrm{TVCD} - \mathrm{BG}$), such as the emission height, and range $z \in [0, T]$ describes the integration interval from the ground, up to any point above the enhanced mass.

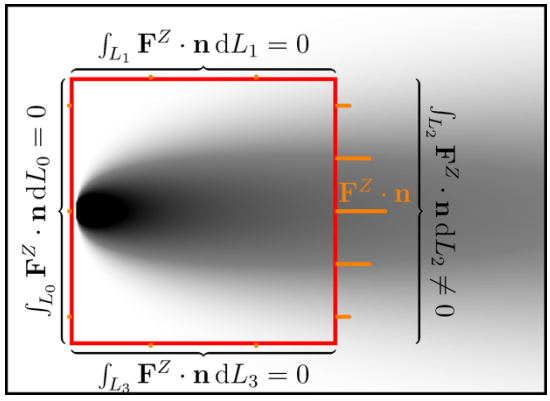
3.1.2 Application

To estimate the emission sources (E) using the CS method, we make further assumptions:

[1] Under the assumptions that transport takes place purely through linear advection, and that there is no mass flux into the ground or out into space.

1. We ignore the sink component for CO_2 (S=0), while we assume that for NO_2 the emissions decay exponentially away from the source, which we can estimate by fitting an exponentially decaying curve to the observed estimates.
2. We assume steady-state conditions, so assume $\hat{\rho} = 0$ in a given volume. This assumption is not strictly correct (e.g., non-uniform winds can yield puff-like structures of accumulating mass in specific areas). However, the net accumulation should have a zero mean over time and space, so the assumption is valid as long as we average over many estimates in space and time.
3. Assuming plume emissions come from (isolated) point sources, we can simplify the closed line integral of the mass continuity equation to a single line integral in the downwind direction, i.e., we assume $\oint_A \vec{F} \cdot \vec{n}\, dA \approx \int_l F_n\ dl$. The assumption is valid as long as we satisfy: (*a*) that our line segment l transects the entire plume in its cross-sectional direction, and (*b*) that upwind of the plume the enhanced concentrations are negligible $(\mathrm{TVCD} - \mathrm{BG} \approx 0)$. (See the figure below, where the integral cross-sectional to the flux is non-zero, while all other line integrals making up the closed contour vanish).
4. The closed line integral $\int_l F_n\ dl$ is replaced by a Gaussian distribution fitting procedure, as this will be the typical distribution of a plume when sampled in its cross-sectional direction (see again the figure below). This furthermore helps filling data gaps, for example due to clouds. To make the fit, both the CO_2 and NO_2 measurements are used simultaneously to obtain the mean position and standard deviation of the Gaussian cross-section. The uncertainty of the CO_2 measurements will typically be higher than the NO_2 measurements, so there is a transfer of information from the NO_2 to the CO_2 line integral.
5. To improve the Gaussian fits, we split the plume into “sub-polygons” of about 5 km intervals in the downstream direction (about 2 to 3 pixels deep, for the upcoming CO2M satellites) and fit a Gaussian to each of the sub-polygons.

For more details, see Kuhlmann et al. (2019, 2021).



3.1.3 Uncertainty and quality flags

The uncertainty of the cross-sectional flux method is computed by propagation of uncertainty from the single sounding precision. The uncertainty of the estimated emissions includes the uncertainty of the estimate and the wind speed,

$$\sigma_E = \sqrt{\sigma^2_{q|x=0|} \cdot u^2 + \sigma^2_u \left(\frac{E_{est}}{u}\right)^2}$$

where $\sigma^2_{q|x=0|} \cdot u^2$ corresponds to the estimated uncertainty of the line densities projected towards the source ($x = 0$), combined with the downwind wind speed u, plus the estimated error in the wind speed (we assume $\sigma_u = 0.5$ m/s).

To remove bad estimates, emissions are not computed when they are negative or very large. Estimates are further excluded if the angle between wind speed and centre curve is larger than 45 degrees, which often indicates false plume detections. Estimates are also rejected, if more than 5 pixels were detected upwind of the plume.

3.2 Gaussian plume (GP) method

3.2.1 General method

The analytical solution for a model of constant emissions (E) being released under a uniform wind field with steady-state conditions and isotropic diffusion is the Gaussian plume model. By integrating this model vertically, we may compute what a total column could look like for a given set of parameters; and by fitting this model to the actually observed (enhanced) total column, we can solve an inverse problem for the parameters that best describe the total column. The equation fitted to the observations is

$$I = \frac{E}{\sqrt{2\pi}\sigma(x)} e^{-\frac{y^2}{2\sigma^2(x)}},\ \sigma(x) = \sqrt{\frac{2Kx^b}{u}},$$

where I is the image (e.g., $\mathrm{TVCD} - \mathrm{BG}$), E is still the emission flux, x and y are the downwind and cross-wind plume directions respectively, $\sigma(x)$ describes the spread of the plume in the downwind direction, K describes the eddy diffusivity coefficients, and u describes the effective downwind wind speed, and the exponent b on x in the equation for $\sigma(x)$ allows for further non-linear plume dispersion along the plume, depending on meteorological conditions.

We note here that the Gaussian plume model used here is more flexible than a typical use of the Gaussian plume fit: we allow further non-linearity through the exponential factor 'b' in the model, and the x and y directions are plume-following coordinates. The latter feature allows the Gaussian plume method to contain curvature. This is in in contrast to a more typical application of a Gaussian plume fit, where the plume is typically just straight.

3.2.2 Application

The plume centre curve is described by a 2nd order Bezier curve which has three control points (one centred at the known plume source location, the other two are fitted parameters for the Gaussian curve, initialized along the curve as already obtained using the same method as, e.g., the above cross-sectional flux approach). To model the NO_2 image with the Gaussian plume model, we multiply the Gaussian plume model with a factor $e^{-\frac{x}{u\tau}}$ for half-life time τ, which describes the exponential decay of the NO_2 field downwind of the plume. All parameters are fitted in the inversion (i.e., $E,\ K,\ b,\ u,\ \tau$, and two points to describe the plume center line, apart from the source point which is kept fixed). The inversion is a simple Levenberg-Marquardt least-squares optimization. The initial conditions are the same in all cases:

- E is initialized to the mean summer emission rate for a given source,
- K is initialized to 400 m^2/s, allowed to range from 200 to 1600 m^2/s,
- b is initialized to 1, allowed to range from 1 to 2,
- u is initialized as $\sqrt{U_{\mathrm{eff}}^2 + V_{\mathrm{eff}}^2}$ following the definitions in the previous section, allowed to range an extra ± 2 m/s,
- τ is initialized to 4 hours but allowed to range from 2 to 16 hours for NO_2, while $\tau = \infty$ for CO_2,
- The two points describing the Bezier curve of the plume are initialized along the plume centreline as in the previous section.

If NO_2 data and CO_2 data are present, we first fit the Gaussian plume to the NO_2 observations; then, we set $\tau = \infty$ and use the other obtained variables as initial conditions (except for E,

which is reinitialized to the mean CO_2 summer emission rate for that source). We only allow a small amount of deviation around the previously obtained Gaussian plume parameters.

We note that our use of a prior emission expectation (E) corresponding to the mean summer emission values, is unlike any of the other methods presented in this deliverable. A good prior emission estimate was required to get the least-squares fitting procedure to converge onto a global (rather than local) minimum – as the Gaussian plume model has many factors subtly trading off against each other. We note that, alternatively, one could for example obtain a prior emission expectation using another lightweight method such as discussed in this deliverable.

3.2.3 Uncertainty and quality flags

We compute the uncertainty of the Gaussian plume estimate as

$$\sigma_Q = \sqrt{\sigma_{E,fit}^2 + \sigma_u^2 \left(\frac{E_{fit}}{u}\right)^2},$$

where $\sigma_{E,fit}$ is the estimated standard deviation of the fitted emission data, while the rest follows the uncertainty estimate of the Cross-sectional flux method described above.

Data was rejected when no fit was found, no standard deviation was found (i.e., if no good fit was possible), or when the emission rate was below 1/4 or beyond 4 times the prior expected emission rate. Again, we note that our use of a prior emission estimate is unlike any of the other methods presented in this deliverable.

3.3 Integrated mass enhancement (IME)

3.3.1 General method

The integrated mass enhancement (IME) method computes the emission flux from the integrated mass above the background and a residence time (τ) (Frankenberg et al. 2016):

$$Q = \frac{1}{\tau} IME.$$

Varon et al. (2018) used the method to estimate methane emissions. They trained a model for estimating the residence time from 10-m winds using LES simulations.

Here, we derive the IME method from a Gaussian plume model following Frankenberg et al. (2016). For Gaussian plume model, the integrated mass enhancement is given as

$$IME = \int_{-\infty}^{+\infty} \int_{0}^{L} \frac{Q}{\sqrt{2\pi}\sigma(x)U} \exp\left(-\frac{y^2}{2\sigma(x)^2}\right) dx\ dy.$$

Solving the integral results in

$$IME = \frac{Q \cdot L}{U} \Leftrightarrow Q = \frac{U}{L} IME.$$

This version is valid for inert gases like CO_2 or methane. For NO_2, the Gaussian plume model needs to be multiplied with a decay term:

$$D = e^{-\frac{x}{L_0}}$$

with decay length $L_0 = U\,\tau$.

For the integral, it is then possible to compute the emissions as follows

$$Q = \frac{f_{1.32}}{c} \frac{U}{L} IME \quad \text{with} \quad c = \frac{L_0}{L} \cdot \left(\exp\left(-\frac{L_{min}}{L_0}\right) - \exp\left(-\frac{L_{max}}{L_0}\right)\right)$$

with NO_2-to-NO_x conversion factor f. L_{min} and L_{max} is the along-plume interval over which the IME is computed.

3.3.2 Application

To apply the IME method, we use the plume detection algorithm to identify the location of the plume from either CO_2 or NO_2 observations using the same settings as for the cross-sectional flux and Gaussian plume approach. The algorithm also computes the CO_2 background field, the centre line and the along- and across-plume distance for each pixel.

The integration area needs to include pixels with CO_2 enhancements below the detection limit of the plume detection algorithm. The area is obtained by applying a binary dilation to the binary mask of detected pixels using a circular kernel with radius of 25 km. The dilation increases the area of the detected plume to include pixels below the detection limit. In along plume direction, the integration is done from the source ($L_{min} = 0$) to a distance L_{max} . L_{max} is the distance of the most distant pixel in the integration area minus 10 km.

The wind speed is taken either from SMARTCARB or from ERA5. For SMARTCARB, wind speed is computed using the mean of 5×5 grid cells in the COSMO grid at the source location.

3.3.3 Uncertainty and quality flags

The uncertainty of the IME is computed by propagation of uncertainty from the single sounding precision. The uncertainty of the estimated emissions includes the uncertainty of IME and the wind speed. The wind speed uncertainty is taken from the standard deviation of the 5×5 grid cells used for computing the mean, but at least 1.0 m/s. Methodological uncertainties, for example, from the assumption of steady-state conditions are not considered in the estimated uncertainty.

To remove bad estimates, emissions are only computed when at least 10 pixels are detected, and the plume does not overlap with other plumes. For power plants, we remove estimates when more than five pixels were detected 2 km upstream of the source. We exclude estimates if the angle between wind speed and centre curve is larger than 45 degrees, which often indicates false detections.

3.4 Light cross-sectional flux method

3.4.1 Initial algorithm

The light cross-sectional flux method used in this study is derived from the method originally developed by Zheng et al. (2020) to estimate the CO_2 emissions of Chinese cities and industrial areas that produce atmospheric plumes clearly detectable in transects by OCO-2. This method has then been adapted to the routine and automatic estimation of isolated clusters of CO_2 emissions worldwide (Chevallier et al., 2020) and used to study the temporal variability of the emissions by using several years of OCO-2 and OCO-3 data (Chevallier et al., 2022).

The initial inversion approach was thus designed for the use of OCO-2 data which are characterized by a resolution of few km^2 and by a narrow swath about 10 km wide. These data show mostly variations along the direction of the satellite track and the inversion approach considers the data as one-dimensional. Like the other cross-sectional flux method, this approach computes line densities associated to the increase of XCO_2 above the background in sections of the data corresponding to plume transects. Here, the local XCO_2 enhancements above the background and thus the line densities are extracted by fitting locally OCO-2 data segments with a one-dimensional function that includes a Gaussian term for the plume transect and a linear term for the background depending on the distance along the satellite track (Zheng et al., 2020). The observations associated to the plume from a given source often describe a Gaussian bell function along the satellite track (Figure 1). However, the choice of the Gaussian form is primarily motivated by its neutrality in the sense of the principle of

maximum entropy. An important aspect of the method compared to other cross-sectional flux methods is its relative simplicity and low computational cost, which allowed applying it to the full OCO-2 and OCO-3 multi-annual and global datasets.

The method was easily applied to OCO-3 XCO$_2$ observations in nominal mode since this mode is very close to the nominal XCO$_2$ observation mode of OCO-2 (Chevallier et al. 2022). When applying the method to OCO-3 Snapshot Area Maps (SAMs), Chevallier et al. (2022) simply processed independently each band of narrow observations composing, in practice, these small images.

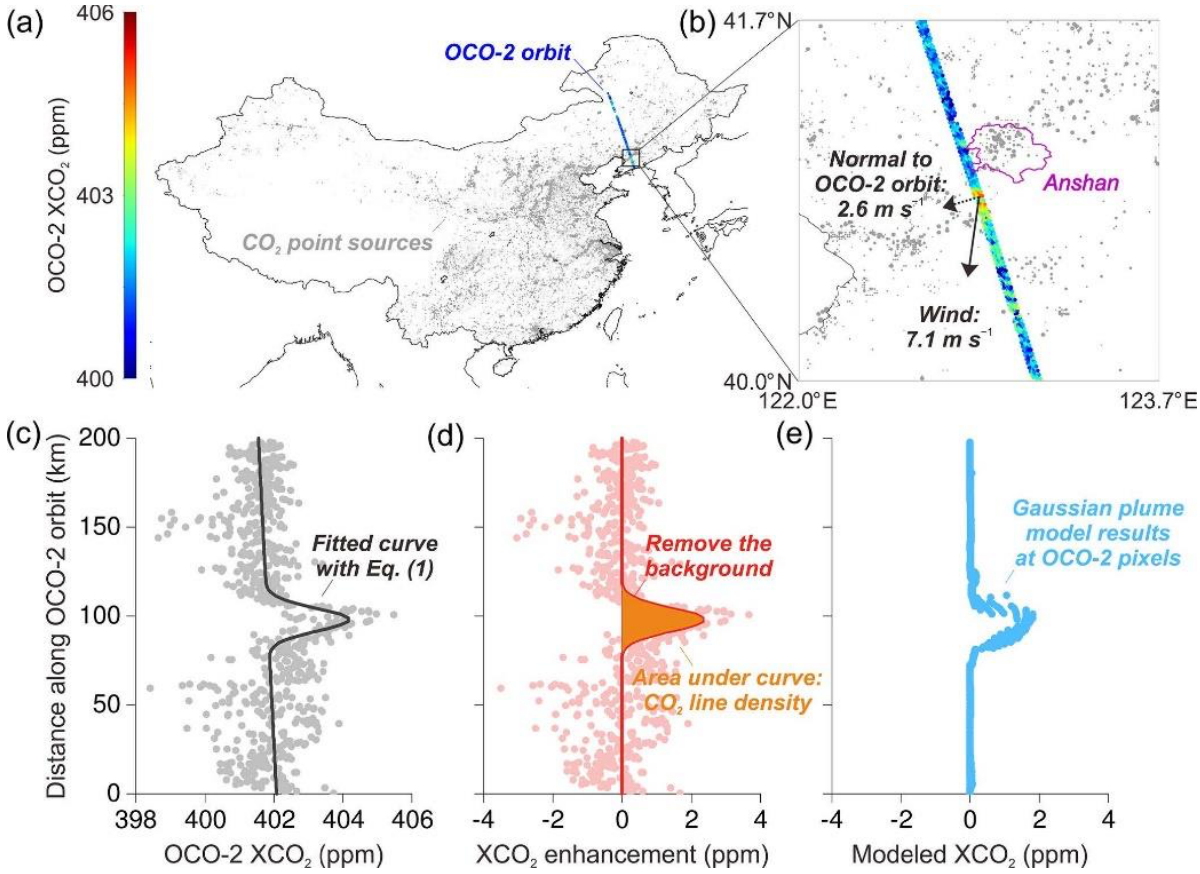


Figure 1 Description of the algorithm used by Zheng et al. (2020) to derive CO$_2$ emissions with OCO-2 data (a). Line densities (c) are extracted from the fit of a function (b) around XCO$_2$ enhancements. The retrieved emissions are computed at the selected enhancements as the line densities multiplied by the orthogonal components to the satellite track of the effective winds (a). Figure from Zheng et al. (2020).

3.4.2 Adaptation and implementation of the method to tackle wide synthetic XCO$_2$ images

The here described cross-sectional method was designed to process quasi-one-dimensional data such as provided by the OCO-2 mission to derive emissions approximately corresponding the time of the plume transect. In this comparative study, this initial cross-sectional method has been adapted to the process of CO2M-like images to retrieve average emissions over the hour preceding CO2M-like data: i.e., between 11:00 and 12:00 in the experiments with SMARTCARB data. Moreover, it has also been enhanced by the ability to process column-averaged NO$_2$ images in conjunction with XCO$_2$ images to estimate CO$_2$ emissions. Analyses have been conducted to ensure that all the corresponding modifications preserved the low cost of the computations for the method. As described hereinafter, the

modified cross-sectional method follows three main steps to estimate the emissions of a given source from a given image.

First, the inversion task of XCO_2 images is converted into a one-dimensional problem: a data slice is extracted from the image: *dwind* traveled by the wind from the source in one hour (see Figure 2). The slice extends then 50 km on either side of the source. If there are more than 100 available data pixels, the subsequent steps of the algorithm will process the data contained within this slice. Within the distance *dwind* from the source, the method then selects the pixels of the data slice where local XCO_2 enhancements, defined as the difference between the local XCO_2 value and the average of the data surrounding the pixel within a 100 km wide window centered on it, are greater than the spatial variability of the data contained within this window. As selected local enhancements are within the distance *dwind* of the source, they could be generally associated to the targeted source except when the plume generated by a nearby source comes on top of the plume of the targeted source.

Then, for each selected local enhancement, and if there are more than 100 retrievals contained within the 100 km wide window centered on it, a function containing a linear term and a Gaussian function is fitted to the data contained within the centered window:

$$y = m * x + b + \frac{A}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad (1)$$

where y is XCO_2 (ppm), x the distance (km) along the data slice contained within the centered window. The linear term accounts for the fit of the XCO_2 background and the Gaussian function for the fit of the potential plume associated to the selected enhancement. The fit parameters m, b, A, μ and σ are estimated through a non-linear least-squares procedure of the Levenberg-Marquardt type. In order to prevent the algorithm from diverging or converging towards local minima, it is initialized with an ensemble of first guesses for its parameters and data undergo a Z-norm normalization. The quality of the fits is assessed with the R^2 regression score and only the parameters associated with the fit of largest R^2 are kept for the subsequent steps. If the least-squares algorithm manages to successfully fit the data and to produce a set of optimal parameters, additional criteria are applied in order to discriminate the local enhancements associated to a robust curve fit: 1) as the resolution of the data is of 2 km, the standard deviation (σ) of the fitted Gaussian function should be greater than 1 km. 2) For point sources, σ should be lower than 5 km in order to discard Gaussian fitted curves associated to unrealistic wide plumes. For each local XCO_2 enhancement satisfying all these criteria, the value of the CO_2 line density, approximated here by the area between the Gaussian peak curve and the background line, is equal to the parameter A.

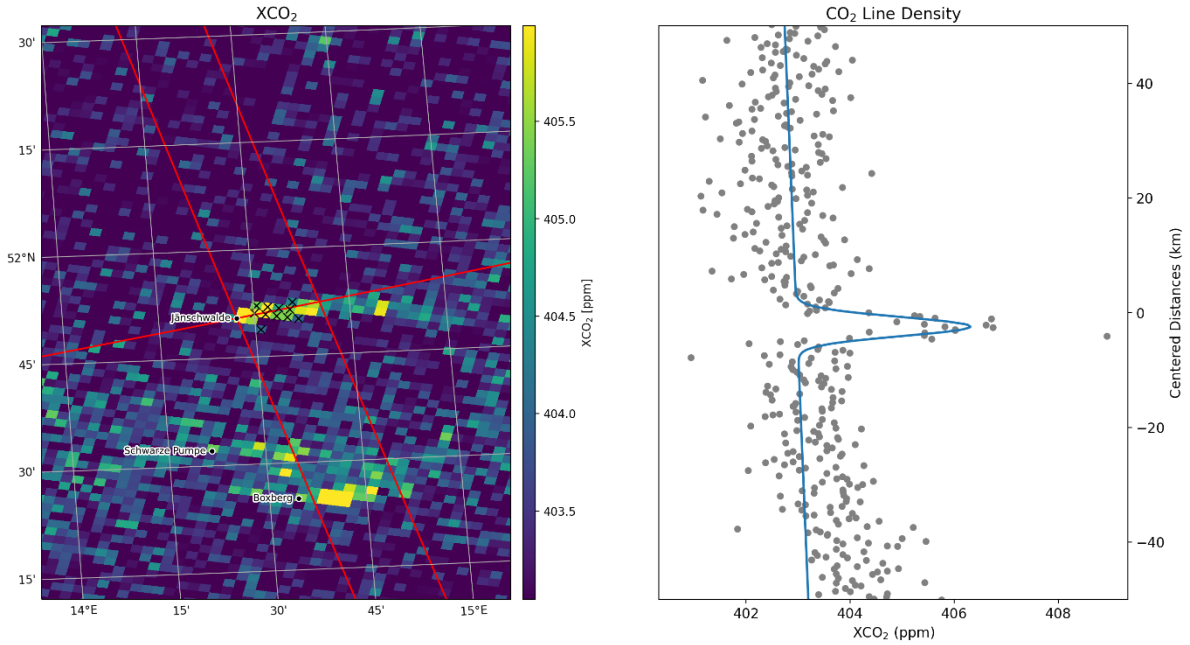


Figure 2 Illustration of the light cross-sectional method on the example of the Janschwalde power plant with a CO2M-like image of XCO_2 extracted from a COSMO-CHG simulation (June 22nd, 2015). A data slice downwind of the source is extracted from the XCO_2 image in order to capture most of the plume emitted by the source for one hour (left panel, parallel red lines orthogonal to the along-wind direction at the source also in red). Within this data slice, a curve fitting is performed with the XCO_2 data points contained within 200–km wide windows centered on the enhancements above the background that are close to the source (right panel). Under certain conditions (see text), some enhancements are selected (crosses in the left panel) and their associated line densities used for the estimation of the emission amplitude of the source.

Finally, for each local XCO_2 enhancement from which a CO_2 line density could be extracted at the former step, the estimated cross-sectional flux, *i.e.,* the amount of mass of CO_2 passing per unit time through the section of the plume intersecting at the local enhancement with the data slice, is obtained by multiplying the line density by the component of the effective wind orthogonal to the data slice. This effective wind should represent the average wind according to the vertical density of CO_2 that crosses the section defined by the density line but unfortunately the vertical distribution of CO_2 in the plume is usually not known. Some studies have then chosen as proxies the winds vertically averaged according to emission height profiles (Kuhlmann et al., 2021), or the winds at, or averaged below, a given altitude (Zheng et al., 2020; Chevallier et al., 2020). Here, we have chosen to compute the emissions with the winds taken at the altitudes of 100, 200 and 300 m. The fact of considering the effective wind as a set of winds taken at three different altitudes combines the simplicity (low cost) of the calculation related to a wind taken at a constant altitude with the consideration of different atmospheric conditions and different heights of emissions characterizing the transport signatures produced by sources. Moreover, in order to prevent the algorithm from producing unrealistic values, the estimated cross-sectional fluxes cannot take values outside the range [0 – 90 $MtCO_2/yr$].

Finally, under steady-state atmospheric conditions, each of the cross-sectional CO_2 fluxes derived at the selected XCO_2 enhancements is almost equivalent to the upwind source emissions. Therefore, as several enhancements belonging to the atmospheric signature of the source are generally processed for each of the three effective wind heights used for the computation, the algorithm produces generally several individual estimates of the source emission for a given image. We will consider then that the estimate computed by the method for a given source and for a given image is the median value of these individual estimates. And, as each individual estimate is characterized by the R^2 value associated to the fit of the

corresponding line density, we will consider then that the estimate computed by the method has an associated R^2 equal to the median value of the individual R^2s.

Compared to the cross-sectional method described previously (Section 3.1), the cross-sectional method described here is characterized by less complexity, and it does not need the complex geometrical characterization of the plume required by the previously described method which requires the centrelines and the sub-polygons where the flux estimations are performed. The relatively low computing cost of this method (see section of the computing cost) justifies its qualification as a *light* cross-sectional method. Also, the relative simplicity and inexpensiveness of this method allows it to be easily adapted to an automatic monitoring of CO_2 emissions worldwide (Chevallier et al., 2020).

3.4.3 Adaptation and implementation of the method in order to use NO_2 data from satellite images

If column-averaged NO_2 observations are used instead of observations of XCO_2 to estimate NO_X emissions, the same process of determination and selection than that of CO_2 is applied in order to determine the line densities associated to NO_2 observations and the corresponding cross-sectional NO_2 fluxes at the selected enhancements. But, the determination of the NO_X emissions at the corresponding source has to take into account the fact that sources emit NO_2 with a NO_2:NO_X ratio and that this chemical specie undergoes chemical reactions and thus a depletion with time downwind of sources. Following the paper of Kuhlmann et al. (2021), these two effects are modeled by an exponential function which decays with time, and which scales the cross-sectional NO_2 flux at the selected enhancements:

$$\mathrm{NO_x}(0) = f\mathrm{NO_2}(x)e^{\frac{x}{u\times\tau}}, \quad (2)$$

Where $NO_X(0)$ is the NO_X emission amplitude at the source, $NO_2(x)$ the cross-sectional flux of NO_2 assessed at the distance x from the source, f is the NO_2:NO_X ratio set to a constant value of 1.32 when inverting SMARTCARB data, u the effective wind speed and τ the decay time set to a constant value of 4 h. The constant values of f and τ are the same that were assumed in the simulations of XCO_2 extracted from the COSMO-GHG model used in this comparison study (Kuhlmann et al., 2019).

If column-averaged NO_2 observations are used in conjunction with observations of XCO_2 to estimate CO_2 emissions, the line densities associated to NO_2 observations are first determined. Then, at the locations of the selected NO_2 enhancements, the curve fitting producing the CO_2 line densities from XCO_2 data is performed by keeping the standard deviation (σ) and the mean position (μ) of the Gaussian term fixed to their value deduced from the computation of the NO_2 line densities (Eq. 1). For some selected NO_2 enhancements, the curve fitting of the XCO_2 data at their location may not however produce relevant CO_2 line densities and, when using XCO_2 data only, additional relevant CO_2 enhancements could be determined at different locations of the selected NO_2 enhancements. Therefore, we consider that the estimated emissions obtained by the joint use of NO_2 and XCO_2 observations consist of the emissions determined from both observations but also of the emissions derived from the use of XCO_2 data only, at the enhancements which have not been selected when estimating NO_2 line densities.

3.4.4 Uncertainty and quality flags

The uncertainty in the estimates provided by the LCS method can be computed by propagation of the error assessed by the fitting algorithm on the amplitude of the Gaussian function (Eq. 2). However, this uncertainty is more prone to computing errors than the R^2 score associated to the curve fitting which justifies the use of this indicator to assess the quality of the estimates.

As described above, several quality checks remove potential unrealistic estimates of a given source: the fitting window should contain enough data pixels, the selected enhancements should have sufficient amplitude and be close to the source, the standard deviations associated to the fit of the Gaussian function should be larger than 1 km and smaller than 5 km, and the estimated emissions should be positive and lower than 90 $MtCO_2$/yr.

3.5 Divergence method

Beirle et al. (2019, 2021) introduced the divergence method to estimate the NO_X emissions from TROPOMI NO_2 observations. During the CoCO2 project, we have adapted the divergence method to estimate CO_2 emissions (Hakkarainen et al., 2022). A comprehensive theoretical discussion about this approach is given in the supplementary material of Hakkarainen et al. (2022).

The divergence method is based on the continuity equation (Jacob, 1999) at the steady state, where the divergence of vector field F (flux) is defined as the difference between emissions E and sinks S:

$$\nabla \cdot F = \mathrm{E} - \mathrm{S}$$

The flux F is defined as $F = (F_x, F_y) = (I \cdot U_{\mathrm{eff}}, I \cdot V_{\mathrm{eff}})$, where $I = \mathrm{TVCD}$ is the vertical column density observed by satellite, and U_{eff} and V_{eff} are the eastward and northward winds, respectively, at the level of the enhanced concentrations (as described in Section 3.1). The NO_X sink can be calculated from the NO_2 columns as $\mathrm{S} = \mathrm{LI}/\tau$, where τ is the NO_X lifetime generally assumed as four hours (as used also in the SMARTCARB simulations) and L is the constant NO_X-to-NO_2 ratio (typically assumed as 1.32). The divergence method can also be applied to CO_2 but since its lifetime is extremely long (in the order of centuries) as compared to NO_X, the sink term can be neglected. To adapt the original divergence approach to long-lived gases, such as CO_2, we remove the atmospheric background (e.g., Hakkarainen et al., 2016) before calculating the divergence as the flux is not linear with the column I due to the changing wind speed.

3.5.1 Implementation

Here we define the XCO_2 background for each orbit as the median over the area of interest. We implanted two approaches: A regional one, where the background is defined over the SMARTCARB area and a local one where the only the area near the emission source is used. Here we also use mean filter to denoise the CO_2 observations.

In order to calculate source-specific emissions from the enhancements in the averaged divergence/emission fields, the peak fitting approach (as in Beirle et al, 2021) is applied by fitting a function including a Gaussian and a linear term.

3.5.2 Uncertainty and quality flags

The divergence method is based on temporal averaging of the satellite observations. Typically, one year (or at least several months) of data are needed to clearly detect point sources. We average over all the available observations to create the seasonal/annual divergence map. We consider the emissions estimated from the seasonal/annual emissions maps as final estimates, although the sampling of the satellite observations might be uneven. The fit can be analyzed via typical error statistics, but also a visual examination of the fit is good practice. We note that not necessarily all sources produce strong enough divergence peaks that can be fitted.

In the divergence method, the emission estimation is performed in two parts. In the first part the divergence/emission map is created. This process is fully automated and can be done without any supervision. In the second part, peak fitting is applied to each point source identified in the maps. We apply the adaptive Markov chain Monte Carlo (MCMC) method to

sample the parameter distribution. We also estimate the statistical noise of the averaged divergence field using MCMC. The fitted parameters are estimated as the mean values of the posterior distribution and the fitting uncertainties as the standard deviation. Finally, for all the sources estimated we also inspect visually the quality of the fit.

3.6 Quality Index (QI) for the instant estimates: accuracy vs number of estimates

Different performance indicators or error estimates can be derived from the application of the inversion methods. Such indicators can be used to identify and select the most reliable estimates. However, there is no objective criteria to impose a threshold on the quality of the estimates. Higher quality comes with smaller sets of estimates and optimal values may also depend on the application. Therefore, a specific aim of our analysis of the experiments with synthetic XCO_2 images is to (i) evaluate quality indices generated by the different methods through comparisons with the corresponding actual errors in the emission estimates (ii) vary the threshold on these quality indices to analyze the variations of the statistics of errors in the emission estimates as a function of the number of cases for which estimates are provided. This analysis is conducted only for methods providing instant estimates from individual images (IME, CS, GP and LCS) since all methods provide annual emission estimates for the full set of sources whatever the level of selection applied to individual images.

The quality indices selected for the different methods are associated to:

- the R^2 score associated to the computation of the line-densities for the LCS method (Section 3.4);
- the derived uncertainties in the emission estimates for the IME, GP and CS methods (Sections 3.3, 3.1 and 3.2 , respectively).

The analysis is expected to demonstrate the complementarities between the different methods since these methods behave differently depending on whether they tackle complex of simple cases, cases with high or low emission rates etc. For example, if the minimal R^2 score, Figure 3). The same behavior is observed for the IME, CS and GP methods when decreasing the computed uncertainty in the estimates below which they are selected. Therefore, these thresholds not only define the accuracy but also the associated number of estimates that can be retrieved. And, by varying the value of one of these thresholds, each method can be thus associated to a range of accuracies with their respective number of estimates for a given inversion configuration (e.g., cloudy or cloud-free). This allows to assess the quantity and quality of the estimates that are associated to a given accuracy of the method.

In order to simplify the presentation of the results and preparations of the images, a common normalized indicator of the quality of the emission estimates is used for all the methods. The Q_{method} value of 1 corresponds to the maximal accuracy that the method can provide, with a minimum number of estimates. Conversely, a value of 0 corresponds to the minimal accuracy, providing the maximum number of estimates. The Q_{method} value of 0.5 represents a compromise between accuracy and number of estimates.

The quality index Q_{LCS} associated with the light cross-sectional method is defined as the linear projection of a range of R^2 values to the interval [0, 1]. The maximal number of estimates (Q_{LCS} = 0) is obtained when no selection of the estimated emissions against their respective R^2 is made. The value of R^2 above which estimates are selected corresponding to Q_{LCS} = 0 is thus 0. In order to get a representative number of estimates, the value of R^2 corresponding to Q_{LCS} = 1 is set to the 95th percentile of the ensemble of the R^2 values. For the IME, CS and GP methods, the quality index is defined as the linear projection of a range of uncertainty estimates to the interval [0, 1]. The value of uncertainty below which estimates are selected corresponding to Q_{method} = 0 is the maximum value of the ensemble of uncertainty estimates. The value of uncertainty corresponding to Q_{method} = 1 is set to the 5th percentile of this ensemble in order to retrieve a significant number of high-quality estimates.

3.7 Estimates of annual and monthly averages

Estimates for long time periods (months or a full year), from the instantaneous emission estimates corresponding to individual images are computed in three different ways. Momentarily defining notation for the emission rate of a given source estimated at time t_i as $E_i = E(t = t_i)$ and the associated uncertainty $\sigma_{Q,i} = \sigma_Q(t = t_i)$, and assuming that there are N estimates available, then the three estimates are defined as:

- Arithmetic mean: $\frac{1}{N}\sum_n E_i = \frac{\sum_n E_i}{\sum_n 1}$,
- Inverse variance weighted mean: $\frac{\sum_n E_i/\sigma_{Q,i}^2}{\sum_n 1/\sigma_{Q,i}^2}$,
- The median emission estimate over all considered time.

We note that these estimates are subject to temporal sampling biases when data with clouds are used by the inversion methods. As the cloud cover is denser during winter, emissions estimated during summer are indeed over-represented and more advanced methods (e.g., fitting periodic curves that can represent seasonal cycles as by Kuhlmann et al. (2021)) could be used instead to improve the estimates, but they are not considered here.

As the divergence method is based on temporal averaging of the satellite observations (Section 3.5), we consider the emissions estimated from the seasonal/annual emissions maps as final estimates. Also with the divergence method, the final estimates are subject for temporal sampling biases.

3.8 Computational cost

Here we provide details on the computational cost of the different methods. We note that algorithms are here mainly optimized for conveniently developing and testing the methods in a research setting (i.e., being able to process one orbit within <1 minute) and that all algorithms can be optimized for reduced processing times in an operational system.

The CS, GP, and IME methods use two steps. In the first steps, the plume detection algorithm detects all plumes in the swath. For each detected plume, a local coordinate system (along- and across-plume direction) is computed using a centre line, the background is estimated, and columns are converted to kg/m². Furthermore, an image showing the plumes and centre lines is generated for each swath. These first steps take about 20 minutes for processing one month of cloud-free SMARTCARB data (66 orbits) on the ICOS-CP. In the second step, the three different methods are applied for estimating the emissions. The cross-sectional flux method requires 1.5 minutes (CO_2 only) or 5.5 minutes (CO_2 and NO_2) for processing one month. The GP method needs 42 minutes (CO_2 only), or 90 minutes (CO_2 and NO_2) and the IME method needs about 30 seconds. The LCS method requires 3 minutes (CO_2 only) and 9 minutes (CO_2 and NO_2) plus 3 minutes for pre-processing for one month of SMARTCARB data.

The computational cost of the divergence method is quite low. The first part of the method is essentially computational cost of the averaging that is done in many typical satellite data applications. The divergence calculation per se, can be done in a fraction of the second. The final part, peak fitting is the most time-consuming part. The fitting can be done via typical methods where the cost function is minimized. We perform here adaptive MCMC calculations, which in principle is a quite computationally costly method. However, the peak fitting function is just a simple algebraic equation which is fast to evaluate. This means that even with hundreds of thousands of samples, an individual source can be estimated within minutes. Some human inspection of the results might be needed at the end.

4 Results of SMARTCARB benchmarking

4.1 SMARTCARB dataset

The synthetic observations used in this section were created within the ESA-funded SMARTCARB project to prepare for the upcoming CO2M mission. The dataset has been extensively described and used in previous works (Brunner et al., 2019; Kuhlmann et al., 2019; 2020; 2021), and is openly available from https://doi.org/10.5281/zenodo.4048227 (Kuhlmann et al., 2020b).

The synthetic NO_2 and CO_2 vertical columns are based on atmospheric transport model simulations obtained with the COSMO-GHG model at 1 km by 1 km resolution. The model domain covers parts of Germany, Poland, and Czechia for the year 2015. The synthetic data were further averaged to 2 km by 2 km satellite pixels along the 250 km wide swath for a constellation of up to six CO2M satellites. The simulations used a simplified NO_x chemistry with a fixed NO_x decay time of 4 hours. NO_x concentrations were converted to NO_2 concentrations using an empirical equation.

4.2 Benchmarking setup

To perform the benchmarking, we use the SMARTCARB dataset with a synthetic CO2M constellation consisting of three satellites. We evaluate the methods using one year (2015) of simulations. We test the emission estimation methods both with realistic cloud scenarios (1% threshold for CO_2 and 30% threshold for NO_2 retrievals) and assuming persistent clear sky conditions. All the methods are tested using a three-satellite constellation (0000, 0805, 1610).

We apply random noise to the synthetic XCO_2 and NO_2 observations. For NO_2 retrievals we selected noise scenario 'high', which corresponds to white additive noise with a standard deviation of 2×10^{15} molec./cm^2. For the scenario with clouds, NO_2 random noise increases with cloud fraction and roughly doubles at 30% cloud fraction. For the CO_2 we use the medium noise scenario which corresponds to white additive noise with a standard deviation of 0.7 ppm for the VEG50 scenario (vegetation albedos and 50° solar zenith angle). The CO2M mission requirements indicate that the CO_2 precision shall be better than 0.7 ppm for vegetation scenario at solar zenith angle of 50 degrees and the NO_2 precision better than 1.5×10^{15} molec./cm^2 (Meijer et al., 2020).

Table 2 summarizes the SMARTCARB benchmarking cases carried out here. We ran all the setups using two sets of winds: COSMO and ERA5. The LCS method uses the winds at the altitudes of 100, 200 and 300 m (Section 3.4.2) while for the other methods, the winds are vertically averaged using the Gridded Nomenclature for Reporting - sector A Public Power (GNFR-A/SNAP-1) emission profile as used by Brunner et al. (2019). The different benchmarking cases that have been run are summarized in Table 2.

Table 2: Benchmark cases.

Benchmark case	Wind dataset	Cloud fraction thresholds	Joint NO_2 and CO_2
Case 1	COSMO	100% (no clouds)	Yes
Case 2	COSMO	1% for CO_2, 30% for NO_2	No
Case 3	COSMO	100%	No
Case 4	COSMO	1% for CO_2, 30% for NO_2	Yes

Case 5	ERA5	100%	Yes
Case 6	ERA5	1% for CO_2, 30% for NO_2	No
Case 7	ERA5	100% (no clouds)	No
Case 8	ERA5	1% for CO_2, 30% for NO_2	Yes

As benchmarking criteria, we used following metrics:

- Number of individual emissions estimated (i.e., successful processing of overpass image).
- Distribution of the differences between the true emission and the estimated emission.
- Distribution of the relative difference between the true emission and the estimated emission.
- Distribution of the absolute values of the difference between the true emission and the estimated emission.
- Annual emission estimate and their relative deviation from the true emission.
- Monthly emission estimates and their relative deviation from the true emissions.
- RMSE (root mean square error) of annual emission estimates.

4.3 Results on instant emission estimates based on individual images

In the following subsections we discuss the benchmarking results and sensitivity studies of CS, GP, IME and LCS methods for estimating emissions from single images. For single image estimates, inversion results can be represented by curves of accuracy *vs* number of estimates which are derived by varying the QIs across their range from highest to lowest accuracy. This type of figure gives a complete overview for each inversion method of the performance in terms of accuracy and number of estimates but for the sake of clarity, inversion results for the image estimates are also represented by box plots of the deviation distributions for all the estimates without any selection.

4.3.1 Performance vs the number of estimates

To assess the inherent performance of the methods without considering the impact of the cloud cover or the uncertainty in the winds, inversion results are analyzed for Case 1 configuration using XCO_2 and NO_2 cloud-free data and the COSMO winds, ie., the winds which were used to generate the synthetic XCO_2 and NO_2 observations.

For most QI values, the inversion methods are ranked according to their median accuracy in the following descending order: GP, LCS, CS and IME (Figure 3). More precisely, the GP and LCS methods are characterized by similar accuracies which are significantly higher than that of the CS and IME methods. For example, for 1,000 estimates, the median deviations are ~ 31% and ~35% for the former methods whereas they are ~48% and ~73% for the latter. The relative differences between the methods increase when the QI tends towards 0: the median deviations are ~37%, ~47%, ~64% and ~102% for the maximum number of estimations reached by the GP, LCS, CS and IME methods respectively.

The maximum number of estimates varies significantly depending on the type of inversion method: from 1,661 estimates for the IME method to 2,295 for the LCS method. The GP and CS methods, based on the same algorithm of plume detection as the IME method, produce up to 1,776 and 2,012 estimates respectively. These differences come from the fact that the number of detected pixels below which the algorithm rejects plumes is different depending on the method. Moreover, the rejection criteria are also different depending on the emission

quantification algorithms used by the different methods. Finally, the overall complexity of the IME, CS and GP methods which are characterized by several rejection criteria for estimates may explain why these three methods deliver a maximum number of estimates lower than the LCS method.

Most of the accuracy indicators associated to the GP and LCS methods are only slightly correlated to the number of estimates (Figure 3). For example, over the entire range of the number of estimates, the 3rd quartiles of the distributions of the relative absolute deviations only increase from 53% to 61% and from 51% to 71% for the GP and LCS methods respectively. On the opposite, the IME and CS methods show increases for the 3rd quartiles that are clearly larger: from 81% to 231% and from 65% to 154% respectively. For these latter methods the selection of estimates depending on their quality index is more efficient than for the GP and LCS methods in the sense that the 3rd quartiles and the 95th percentiles, i.e., the proportion of "poor" estimates, significantly decreases when the quality index increases, i.e., when the number of estimates decreases. This suggests that the IME and CS methods provide more reliable uncertainty estimates in the individual emission estimates than the GP and LCS methods. Their definition and derivation of the quality index better reflects the level of error than for the GP and LCS methods. Nevertheless, the inter-quartile ranges of errors for the GP and LCS methods are relatively narrow, so that for these methods, the tradeoff between performance and number of estimates is thus not a crucial question and retrieving an important number of estimates does not imply a significant deterioration in performance for these methods. On the other hand, this also indicates that the present quality indicators of GP and LCS could probably be improved.

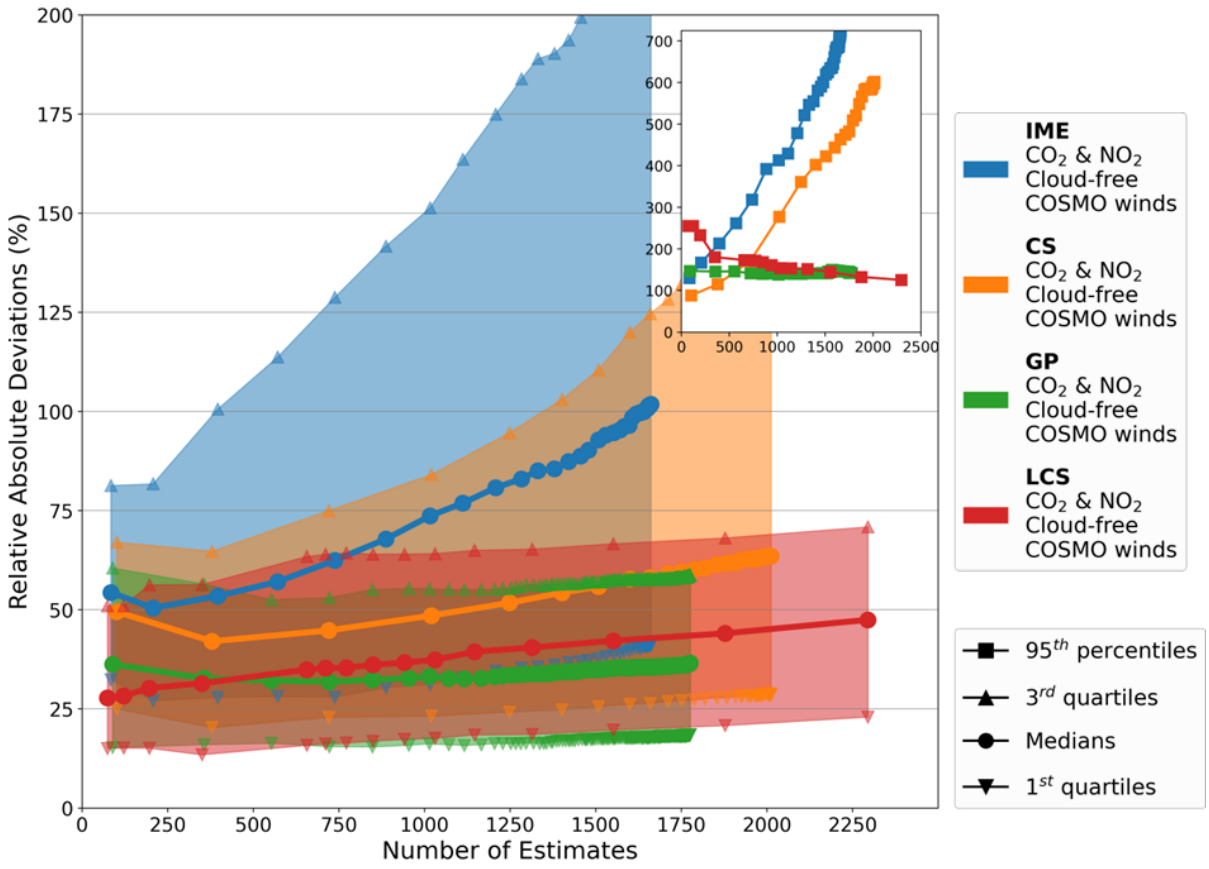


Figure 3 Performance of inversions vs number of estimates. The inversion methods showed here use CO_2 and NO_2 cloud-free data and COSMO winds. The filled areas represent the inter-quartiles of the distributions of the relative absolute deviations depending on the number of estimates. Points belonging to a same curve are associated to a different QI and from left to right along curves, points are associated to a decreasing QI; the points at the left and right ends of the curves are associated to QIs of 1 and 0 respectively.

4.3.2 Sensitivity to the emission strengths of the sources

IME, CS and GP methods globally produce better estimates for strong rather than for weak sources and this fact is particularly pronounced for the IME and CS methods (Figure 4). For these methods, the median values of the absolute deviations for the first quartile of true emissions (0–6.9 MtCO$_2$/yr, 1st row) are ~206% and ~121% respectively whereas they are ~48% and ~37% respectively for the fourth quartile of true emissions (15.6–53.2 MtCO$_2$/yr, 4th row). For low-emitting sources, the performance of these methods is hampered by the small number of enhanced pixels than can be detected within an image with noise. This can strongly hamper, for instance, the identification of the center line of the plume. The present implementation of GP method uses prior knowledge of the source strength and rejects estimates that are outside 25% - 400% rage from the prior. This screening probably improves the agreement of GP, especially for weak sources.

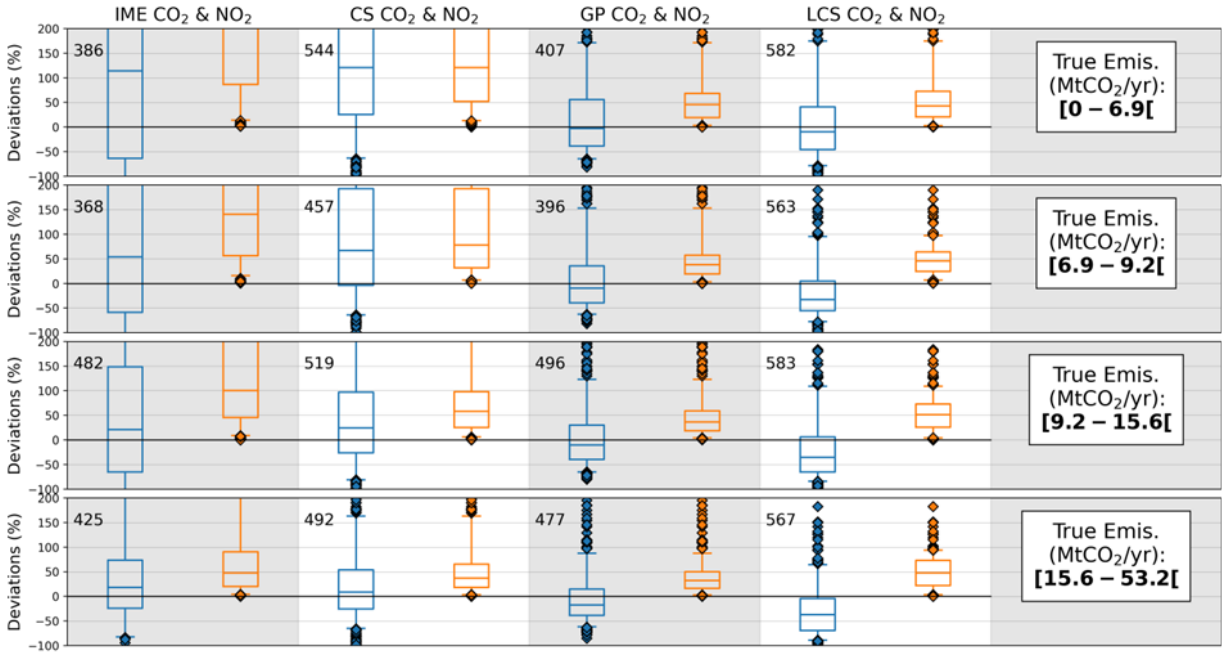


Figure 4 Performance of inversion methods for image estimates and for different ranges of true emissions, i.e., the ones used to generate the CO$_2$ and NO$_2$ data. Inversions are performed with CO$_2$ and NO$_2$ data in the cloud-free case. Results are shown for a QI of 0 for all methods (no selection of estimates). The boxes represent the inter-quartiles of the distributions for the relative deviations (blue boxes) and for the relative absolute deviations (orange boxes), the whiskers the 5th and 95th percentiles, and the lines within boxes the medians. The number next to the boxes is the number of estimates corresponding to the true emissions range and the inversion method of the box.

The sensitivity to the true emission strength is weak and uncertain for the LCS method when considering the 1st quartiles, the medians and the 3rd quartiles of the distributions of the absolute deviations. However, for the 95th percentiles, there is a sensitivity in all methods: for strong sources, the 95th percentile accuracy indicator is 225%, 166%, 90% and 94% whereas for weak sources, it is 1129%, 946%, 173% and 175% for the IME, CS, GP and LCS inversion models respectively. Likely, atmospheric signals produced by strong sources stand out more against the background than for weak sources leading to better signal-to-noise ratio in the XCO$_2$ and NO$_2$ images, and thus also reducing the uncertainty.

Biases in the estimates can also depend on the strength of the source, as in the case of the IME and CS methods, which significantly overestimate the emissions of weak sources compared to strong sources: for the former sources, the medians of the deviation distributions are 114% and 121% for the IME and CS models (blue bars, 1st row) respectively whereas they are 18% and 9% for the strong sources (blue bars, 4th row). This is likely caused by the plume detection algorithm which, for weak sources, wrongly assigns enhancements from other sources in the vicinity to the source of interest. Conversely, the LCS model significantly underestimates the emissions of strong sources compared to weak sources: the median of the deviation distribution is -37% for the former sources whereas it is -9% for the latter.

4.3.3 Impact of the use of NO_2 images for the detection of the plume

For all the inversion methods, the use of NO_2 data to detect and characterize the plumes increases the number of estimates with respect to CO_2-only inversions (Figure 5). The increase is dramatic for the IME and GP methods (~93% and ~70%), important for the CS method (~35%) and weak for the LCS method (~3%). NO_2 data considerably increase the number of estimates for the IME, GP and CS approaches because they use a detection algorithm to identify the location of the plume which is less reliable when using CO_2 observations only (Kuhlmann et al. 2019). As the CS method needs fewer pixels to process plumes, the number of estimates provided by this method is less impacted by using CO_2 data only compared to the IME and GP methods which reject more cases. Finally, for the LCS method, the detection of the plume is made on a slice across wind close to the city, *i.e.,* where the signal of XCO_2 enhancements due to the emissions are relatively large. This focus probably explains the smaller impact of using NO_2 images for this method.

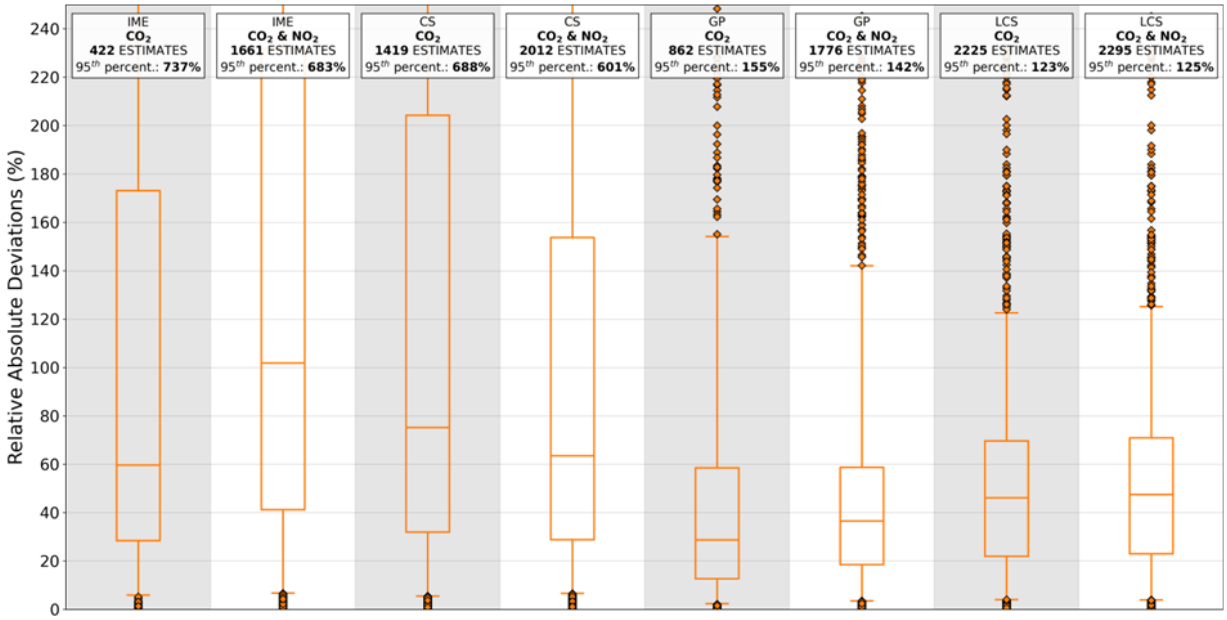


Figure 5 Performance of the inversion methods for image estimates when using CO_2 data alone or with NO_2. The inversion methods use COSMO winds and cloud-free CO_2/NO_2 data. Results are shown for a QI of 0 for all methods (no selection of estimates). The boxes are the inter-quartiles of the distributions of the relative absolute deviations, the whiskers are the 5th and 95th percentiles, and the lines within boxes are the medians.

For the GP and LCS methods, the errors in the emission estimates are similar between the CO_2 & NO_2 and the CO_2-only inversions while it increases when using NO_2 data for the CS method; In particular, the 95th percentile of the absolute residuals associated to this method is reduced from ~205% to ~155% which would suggest that the use of NO_2 data decreases the number of outliers for the CS method. Contrarily, the performance of the IME method decreases when adding NO_2 data, but this fact could be related to a numerical artifact: the IME method performs much better for high-emitting sources than for low-emitting sources (see

Section 4.3.2) and the use of NO_2 data likely allows constraining these sources more efficiently than with CO_2 data only. Therefore, when adding NO_2 data, the ratio low-emitting vs high-emitting sources increases and then the performance degrades. This bias associated to the relative bad estimation of low-emitting sources is confirmed when performance is assessed in terms of deviations instead of relative deviations (Figure 6).

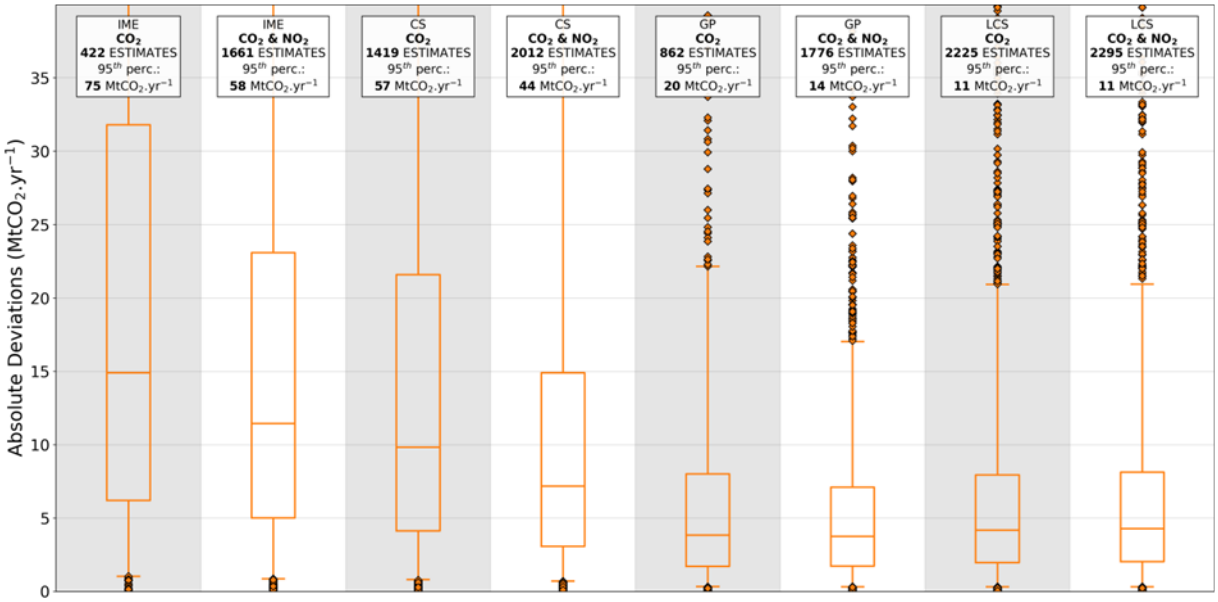


Figure 6 Performance of inversion methods for image estimates when using CO_2 data alone or with NO_2. The inversion methods use COSMO winds and cloud-free CO_2/NO_2 data. Results are shown for a QI of 0 for all methods (no selection of estimates). The boxes represent the inter-quartiles of the distributions of the absolute deviations, the whiskers are the 5th and 95th percentiles, and the lines within boxes are the medians.

4.3.4 Impact of uncertainties in the wind

To assess the impact of potential uncertainties in the wind, inversions are carried out with a different wind product than the one used by the COSMO model to generate the synthetic XCO_2 and NO_2 data which is in total agreement with the images used in the inversions. To this purpose, COSMO winds are replaced by ERA-5 winds, which in particular have a much coarser spatial resolution: 0.25° compared to 0.01°.

For all methods, inversion results are weakly impacted when using the ERA-5 winds instead of the COSMO winds, and they can even improve results such as in the CS method. This fact suggests that the main uncertainty related to the wind lies in the computation by the inversion models of the effective wind used to estimate the emissions (Figure 7) rather in the spatial uncertainty of the wind components. Therefore, by compensation of errors, inversion results can be better when using a different wind product than the one in agreement with the XCO_2 and column-averaged NO_2 data.

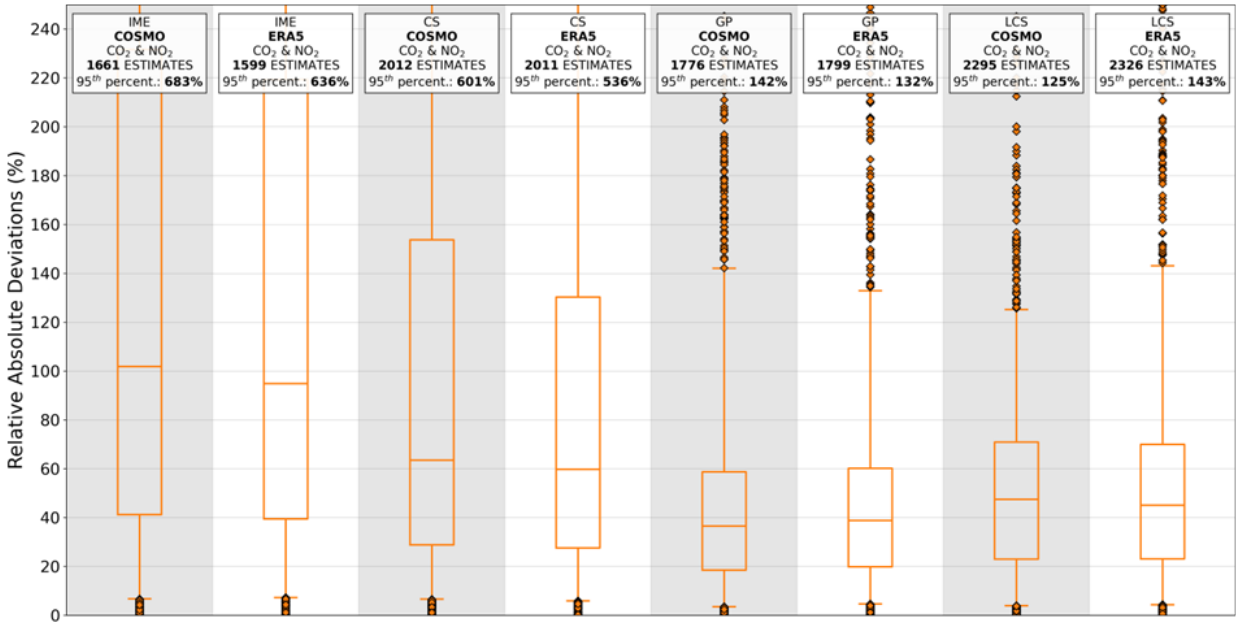


Figure 7 Performance of inversion methods for image estimates when using COSMO winds or ERA-5 winds. The inversion methods use CO_2 and NO_2 cloud-free data. Results are shown for a QI of 0 for all methods (no selection of estimates). The boxes represent the inter-quartiles of the distributions for the absolute relative deviations, the whiskers the 5th and 95th percentiles, and the lines within boxes the medians.

4.3.5 Impact of the cloud cover

The impact of clouds is studied by comparing cloud free and cloudy images as described in Table 2. When integrating the cloud cover in the XCO_2 and column-averaged NO_2 data, the number of estimates is dramatically reduced for all the inversions with a decrease of 94%, 86%, 85% and 87% for the IME, CS, GP and LCS methods respectively (Table 3). The number of estimates that can be provided for the cloudy configuration is at the maximum close to 300 (LCS) and can decrease to 96 for the IME method which cannot provide robust estimates for images with broken clouds. This overall important decrease in the number of estimates is likely consistent with an important reduction of the number of images that can be processed due to the overlay of cloud cover on the images.

When using cloudy data, not only the number of estimates is impacted but also the performance of the methods, albeit to a much a lesser extent. When comparing the results for estimates retrieved for the same images, the “cloud-free” inversions perform better than the “cloudy” ones (Figure 8). For a given image, the partial masking by the cloud cover probably removes pixels whose information helps to better determine the emissions as it is the case with cloud-free data

Table 3 Number of estimates for each inversion method when data with and without clouds are used.

Inversion method	Cloud-free data	Cloudy data
IME	1661	96
CS	2012	287
GP	1776	266
LCS	2295	298
Div (local)	~260 (images per source, 16 sources)	~100 (images per source, 16 sources)

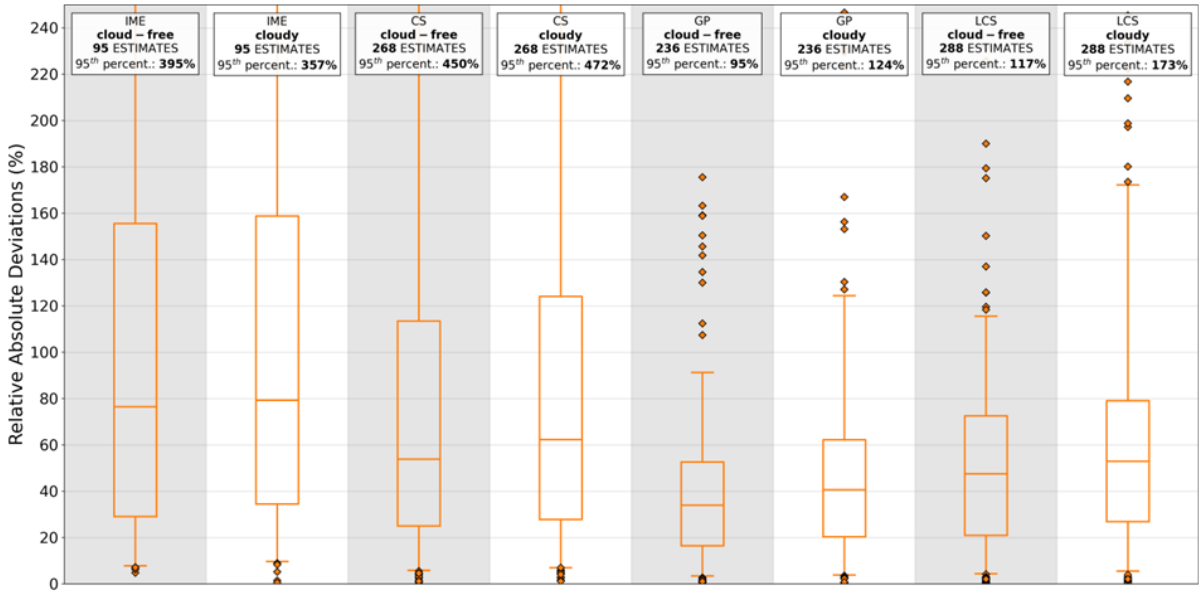


Figure 8 Performance of inversion methods when using data with or without clouds for the emissions estimated from the same images. The inversion methods use CO_2 and NO_2 data and COSMO winds. Results are shown for a QI of 0 for all methods (no selection of estimates). The boxes represent the inter-quartiles of the distributions for the absolute relative deviations, the whiskers the 5th and 95th percentiles, and the lines within boxes the medians.

4.3.6 Synthesis

The sensitivities studies that have been analyzed above can be also illustrated by the variations of accuracy with number of estimates (Figure 9) when the true emissions of the sources are below (1st row) and above (2nd column) 10 $MtCO_2$/yr and, for the inversion configurations for which cloud-free (1st column) and cloudy (2nd column) data of CO_2 and NO_2 are used. Moreover, in order to get closer to a realistic scenario, Figure 9 displays results from the inversions performed with the ERA-5 winds. Summary tables are presented in Appendix A.

The main conclusions that have been drawn in Sections 4.3.2 and 4.3.5 are still valid across the range of QIs: 1) The IME and CS methods perform much better for strong than for weak sources, 2) Taking into account the cloud cover dramatically decreases the number of estimates. Finally, for the most realistic scenario, i.e., cloudy data with ERA-5 winds, one can expect when using the GP and LCS methods a median accuracy of around 45% with an inter-quartile range of around 25–65% for a total number of estimates of around 250.

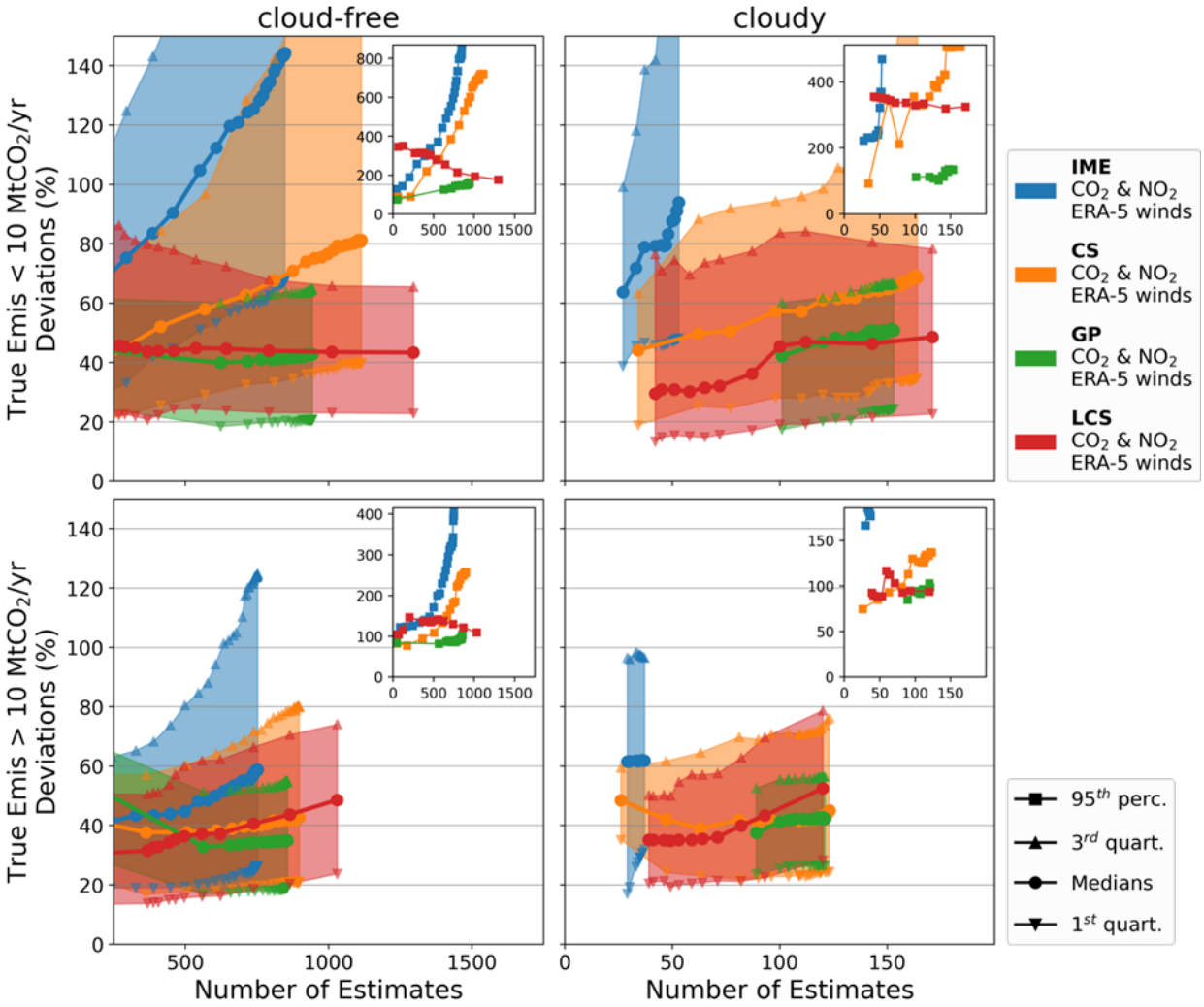


Figure 9 Performance of inversions vs number of estimates. The inversion methods showed here use CO_2 and NO_2 data and ERA5 winds for the cloud-free (1st column) and cloudy (2nd column) cases. Results are shown for the cases where the true emissions of the sources are below (1st row) and above (2nd row) 10 $MtCO_2$/yr. The filled areas represent the inter-quartiles of the distributions of the relative absolute deviations depending on the number of estimates. Each point belonging to the same curve is associated to a different quality index and from left to right along the same curve, points are associated to a decreasing quality.

4.4 Results of annual and monthly averages of the emissions

To assess the performance of the inversion methods at the annual scale, we consider all the image estimates that are generated by the inversion methods regardless of their quality. Annual estimates for a given source are then computed in three ways as described in Section 3.7: 1) the arithmetic means of the image estimates that are available for the source across the entire year, 2) the means of the image estimates weighted by their uncertainty for the IME, CS and GP models and by the inverse of their R^2 value for the LCS model, and 3) the median values of the ensembles of the image estimates that are available for the source across the entire year. The annual means are common statistical indicators of the distributions of image estimates across the year for a given source, the annual weighted means are statistical indicators which integrate the information provided by the methods on the quality of the estimates when averaging, the annual medians are statistical indicators more robust to outliers than the means. Since the divergence method is applied by temporally averaging satellite observations over the year, it produces only a single estimate per source; the three types of annual estimates are thus all three equal to this single estimate. Finally, to compute the relative deviations between estimated and true annual values, the latter are defined as the arithmetic

means of the true values of the emissions of the source across the 365 days of the year. Monthly estimates are computed in the same three ways as the annual estimates but consider the distributions of the image estimates month by month.

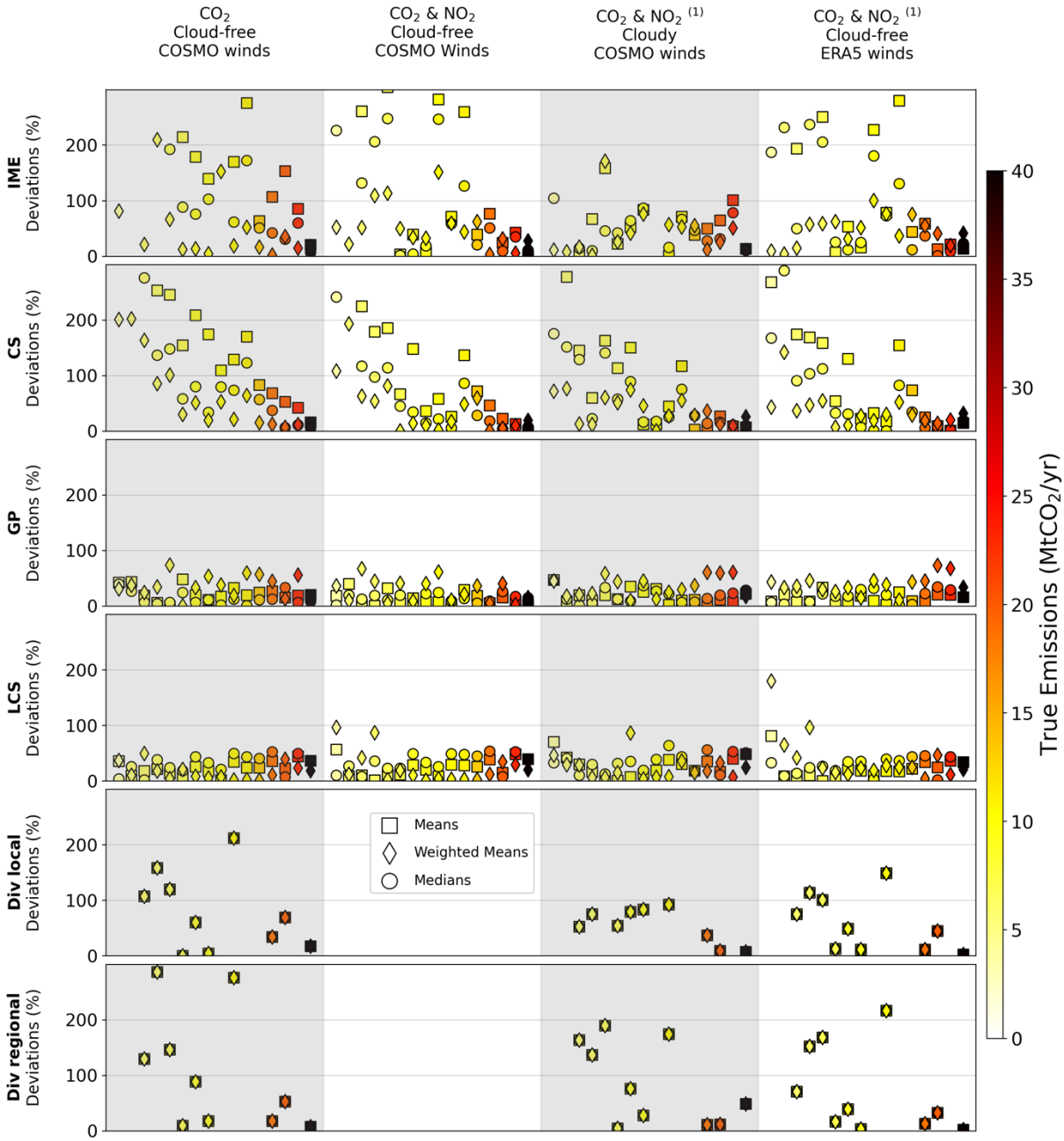


Figure 10 Performance of the inversion methods for annual estimates. The markers represent the relative absolute residual for a given source between the arithmetic means (squares), the weighted means (diamonds) and the medians (circles) of the estimates and of the true emissions over a year. The Inversions are performed with CO_2 cloud-free data and COSMO winds (1st column), with CO_2 and NO_2 cloud-free data and with COSMO winds (2nd column), with CO_2 and NO_2 cloudy data and with COSMO winds (3rd column), and with CO_2 and NO_2 cloud-free data and with ERA-5 winds (4th column). (1) For the Divergence methods, the inversions of the 3rd and 4th columns are performed with CO_2 data only.

Some results at the annual scale are consistent with the findings from the analysis of instant estimates. For example, annual estimates are on average weakly impacted by the use of ERA-5 winds instead of COSMO winds (comparison between the 2nd and 4th columns of Figure 10).

Moreover, the IME and CS methods have poorer performance for weak sources than the other methods also at the annual scale (1st and 2nd rows of the Figure 10). However, deviations from true values are mitigated when we consider the annual medians and especially the annual weighted means which supports the reliability of the uncertainties in the estimates derived by these methods (see Section 3.7). Results for the LCS and GP methods hardly depend on the strength of the sources and are hardly improved by averaging with the weights and can be even degraded for some low-emitting sources in the cloud-free ERA-5 case (4th column, Figure 10).

An interesting point is that annual estimates of most of inversion methods are similar between inversions using data with or without clouds and even better for the IME (comparison between the 2nd and 3rd columns, Figure 10) and divergence approaches (comparison between the 1st and 3rd columns, Figure 10). The important differences in the amount of image estimates between these two inversion configurations do not bias the annual estimates, at least for the year and the sources considered in this study. Furthermore, in the configuration with clouds, the relatively small amount of image estimates does not prevent most of the methods from determining annual emissions for all the sources. The only exceptions are the divergence approaches which estimate the annual emissions of 10 out of 16 sources. This limitation, also present for cloud-free data configurations, is related to the fact that some sources don't produce strong enough divergence peaks from which annual estimates can be assessed. Summary tables are presented in Appendix A.

For the most realistic inversion configuration, i.e., considering data with clouds and ERA-5 winds used, the inversion methods globally deliver annual estimates close to the truth for sources whose true annual emissions are lower than 13 MtCO$_2$/yr: the estimates averaged across the methods are close to the 1:1 line for these sources in Figure 11 ("Ensemble" method). For several of the sources whose annual emissions are above 13 MtCO$_2$/yr (Janschwalde, Boxberg, Lippendorf and Prunerov), most methods systematically underestimate the annual emissions leading to a relatively important underestimation of the annual emissions also by the ensemble model. These biases could be partly related to the fact that, due to a denser cloud cover during winter, most of methods do not manage to estimate emissions for winter months when emissions are the highest (see Section 3.7). The city of Berlin which emits ~20 MtCO$_2$/yr is the only example of a strong source where the error compensation between the different methods leads to a correct agreement between the ensemble method and the truth.

The ensemble method performs better than all the methods with a RMSE that is generally less than half as high except for the GP method whose RMSE is close by (legend of Figure 11). Nevertheless, the underestimation of the annual emissions for strong sources leads to an important RMSE of the ensemble method: its value is equal to 4.16 MtCO$_2$/yr which, relatively to the median of the true annual emissions, leads to a relative RMSE of ~43%.

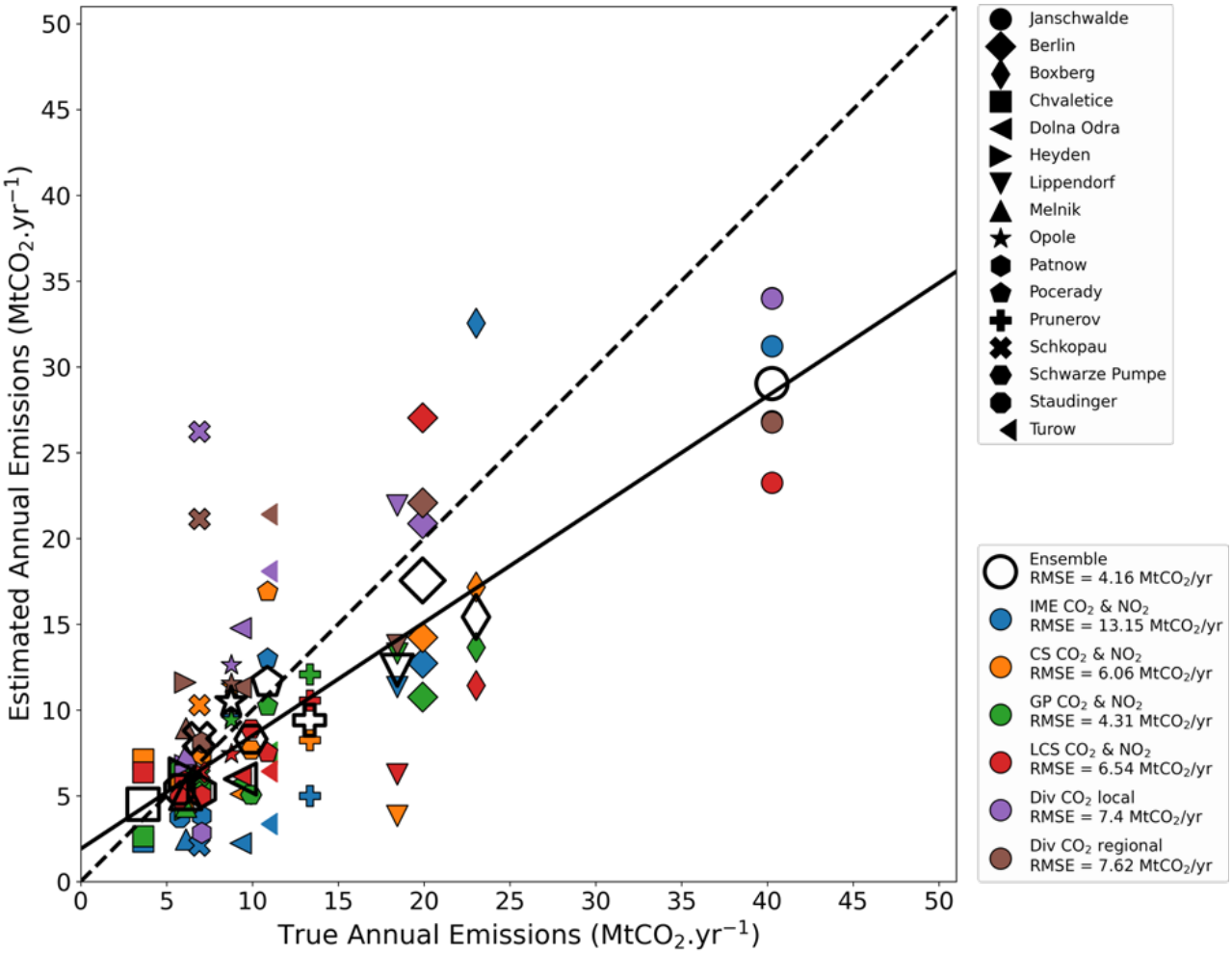


Figure 11 Annual weighted means (IME and CS methods) and medians (GP, LCS and Divs methods) of the estimated emissions vs annual means of the true emissions. Each marker represents a given emission source and each color a given inversion method. The Ensemble method aggregates the values provided by all inversion methods. The divergence inversion methods (Div CO_2 local and regional) use CO_2 cloudy data with ERA-5 winds; the other methods use CO_2 and NO_2 cloudy data with ERA-5 winds. The plain line represents the linear fit of the Ensemble data and the dashed line the 1:1 line. The bottom legend displays the RMSE between estimated and true emissions for each inversion method.

Next, we analyse monthly emission estimates to study how well seasonal cycle can be captured by the light emission quantification methods. All methods of inversion fail to reproduce the seasonal cycle of the true emissions as illustrated in Figure 12. Even though a method correctly estimates annual emissions, its monthly estimates can be in important disagreement with the truth as it is the case for the CS method on the Heyden source or for the LCS method on the Melnik source. Not only are the amplitudes of the monthly true emissions not reproduced but the monthly estimates also do not follow any cycle and show an erratic monthly evolution. Furthermore, the methods generally fail to produce estimates for all the months of the year due to the temporal sparsity of data when the impact of the cloud cover is taken into account.

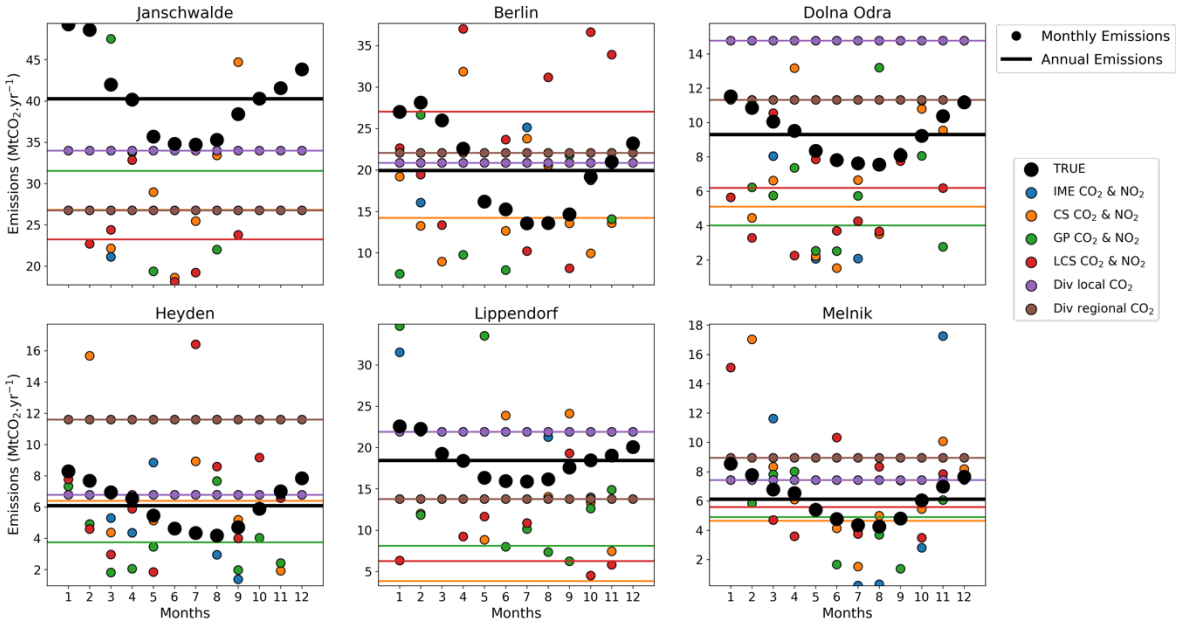


Figure 12 Annual and monthly estimates of the true and estimated emissions for different sources. Each sub-plot is associated to a given source. Lines represent annual averages and points represent monthly averages. Colours are associated to the different inversion methods (The true emissions are in black). Annual and monthly estimates for the IME and CS methods are weighted means of the image estimates. Annual and monthly estimates for the GP and LCS are medians of distributions of image estimates while for divergence method we use annual estimate also for monthly estimates. All inversion methods use CO$_2$ and NO$_2$ cloudy data (CO$_2$ data only for the Divs methods) with ERA-5 winds.

5 Results of benchmarking, library of plumes

5.1 Introduction and method

In this section, we apply the light-weight methods to MicroHH simulations conducted for 48 hours with the Matimba power plant. Details of this LES model and the Matimba power plant simulation are described in CoCO2 deliverable D4.2 (Koene & Brunner, 2023). The motivation here is to analyze, in a controlled environment, how CO_2 inversions can be supported by column observations of NO_2 and specifically how the NO_x chemistry affects the interpretation. The simulations contain full-chemistry and have fields for CO_2, NO and NO_2 that were converted to synthetic CO2M observations. We added random noise of 0.7 ppm for XCO_2 and 2×10^{15} molecules/cm^2 for NO_2 and NO_X column densities. Figure 13 shows an example of the synthetic CO2M observations of CO_2 and NO_2.

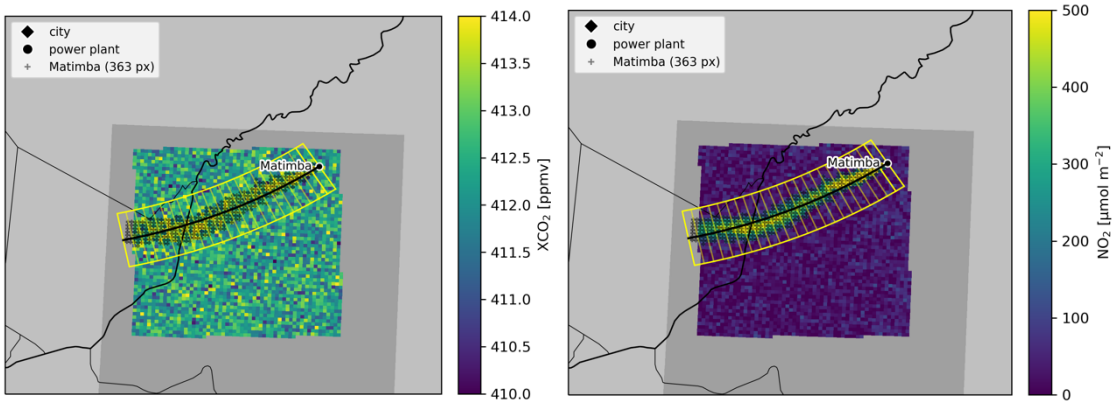


Figure 13 Example of Matimba plume identified by plume detection algorithm.

The synthetic dataset is used for testing how well the different methods can quantify CO_2 and NO_X emissions. For NO_X, we specifically analyse how well NO_X emissions can be estimated using either the NO_X fields or the NO_2 fields from the simulations. This makes it possible to study if the typically used values of the NO_2-to-NO_X conversion factor (f = 1.32) and the NO_X decay time (τ = 4 h) are suitable or need to be modified.

5.2 Results and discussions

Figure 14 shows the time series of estimated CO_2 and NO_X emissions as well as NO_X decay times by the cross-sectional flux method (CS) and Gaussian plume inversion (GP). The time series for the IME method is shown in Figure 15. All methods overestimate the CO_2 and NO_X emissions by 33% to 75% (Table 4), which is likely caused by a too high wind speed that is computed from the wind speed in the MicroHH simulation weighted by the vertical emission profile. Both CO_2 and NO_X emission estimates show a clear diurnal cycle with higher estimates at night and lower estimates during daytime – while the simulation used a constant emission rate. The strong diurnal variability results in a large scatter of about 40%. Note that the biases are close to zero during daytime (9-15 UTC) when the planet boundary layer is well mixed, showing that the used vertically weighted wind speeds should provide an unbiased estimate during satellite overpass times (CO2M: 11:30 local time, TROPOMI: 13:30 local time).

The decay time was estimated for the NO_X columns by the cross-sectional flux and the Gaussian plume inversion with a median of 5.1 and 3.2 hours, which is consistent with the typical value of 4 hours used in previous studies. For the NO_2 columns, estimated decay times are higher (13.5 and 41.1 h for CS and GP method). This is likely caused by the increase of NO_2 columns near the source due to the conversion of NO to NO_2, which increases the lifetime of NO_2 near the source.

The NO_2 to NO_X conversion factor f is estimated for each method using the median ratio of estimated NO_X to NO_2 emissions. The estimated conversion factors are 2.41±1.09, 1.96±0.81 and 2.03±0.50, 3.1 ±0.9 for CS, GP, IME and LCS approach, which is significantly higher than the literature values of 1.32 used in many studies. The conversion factor at overpass time does not vary strongly from the daily mean. The conversion factor depends on plume length as shown for the IME approach in Figure 15. The factor is about 6.0 for a 15-km long plume and drops rapidly reaching <2.1 for plumes 60-km long plumes.

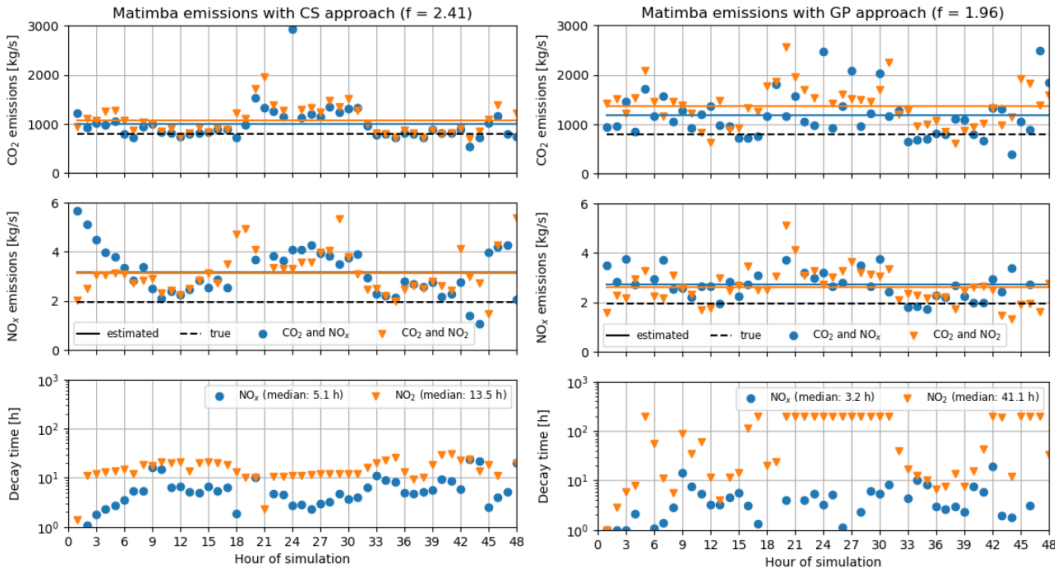


Figure 14 Time series of estimated CO_2 and NO_x emissions and estimated decay times of Matimba using the MicroHH simulations for CS and GP method. The optimal NO_2 to NO_x

conversion factor (f) was calculated from the median ratio of estimated NO_2 and NO_x estimates.

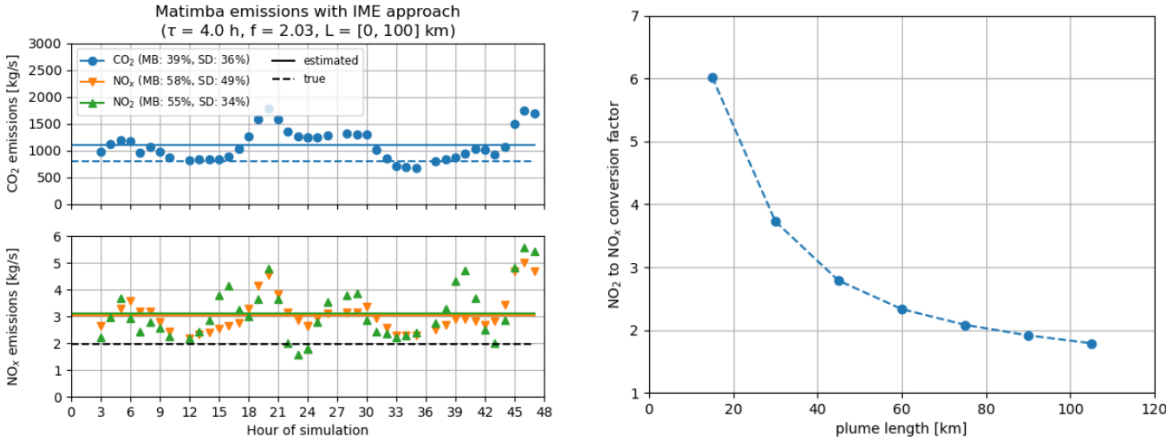


Figure 15 (left) Time series of estimated CO_2 and NO_x emissions of Matimba by applying the IME method to the MicroHH simulations. (right) Dependency of the conversion factor on plume length estimated from the IME method.

Table 4: Computed relative mean bias (MB) and standard deviation (SD) of the different method for estimating CO_2 and NO_x emissions from MicroHH simulations. MB and SD are computed as mean and standard deviation of the difference between estimated and true emissions divided by the true emissions.

	CS		GP		IME	
	CO_2	NO_X	CO_2	NO_X	CO_2	NO_X
MB [%]	34%	60%	74%	33%	39%	55%
SD [%]	33%	44%	51%	36%	36%	34%

Figure 16 shows the NO_X and NO_2 emission estimation fit using exponentially modified Gaussian (EMG) method. The hourly data is taken between 9-15 UTC. Using this approach, we obtain f = 2.18.

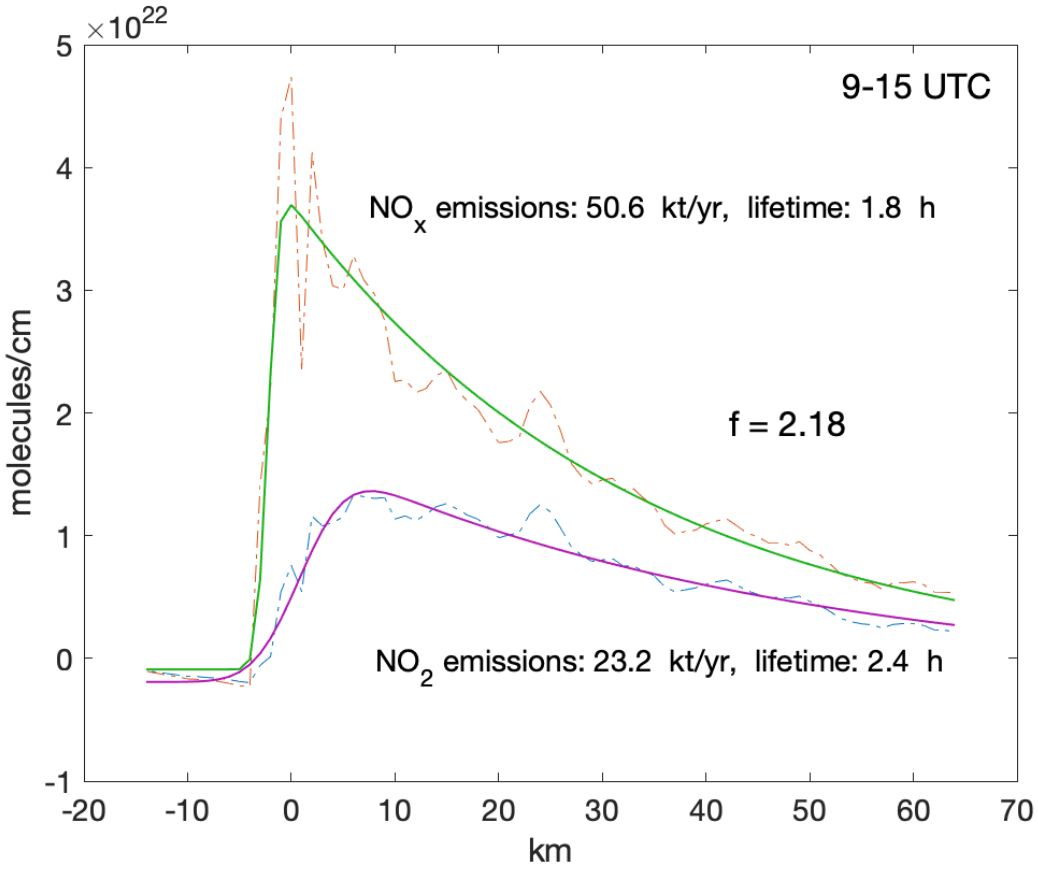


Figure 16 MicroHH NO_X and NO_2 data fitted with the EMG model.

6 Results of benchmarking, satellite data

6.1 Tests with OCO-2 and OCO-3

As described in Section 3.4, the LCS method presented here was initially developed based on real observations from the OCO-2 satellite (Zheng et al. 2020). The application to synthetic data was only introduced at a later stage to participate in benchmarking tests and to prepare this report. In parallel, Deliverables D6.4, D6.5 and D6.6 of CoCO2 have continued the application and the extensive testing with real observations from OCO-2 and OCO-3. The reader is referred to them for further details.

6.2 Benchmarking setup with TROPOMI

For the real satellite benchmarking case, we selected the Matimba and Medupi power stations (Figure 17). These stations have also been analyzed in a previous study (Hakkarainen et al., 2021). The two coal-fired power stations are in South Africa (23.67°S, 27.61°E) and an optimal case study as they are a large and isolated emission source with several TROPOMI/OCO-2 collocations. Figure 18 shows examples of plumes originating from the stations. Here we analyse the TROPOMI NO_2 observations from the reference year 2021. We use the vertically averaged ERA5 winds using the GNFR-A emission profile. The data is taken between 11–13 UTC near the power stations. Here our goal is to benchmark the methods described in Section 3 for NO_X emission estimation.

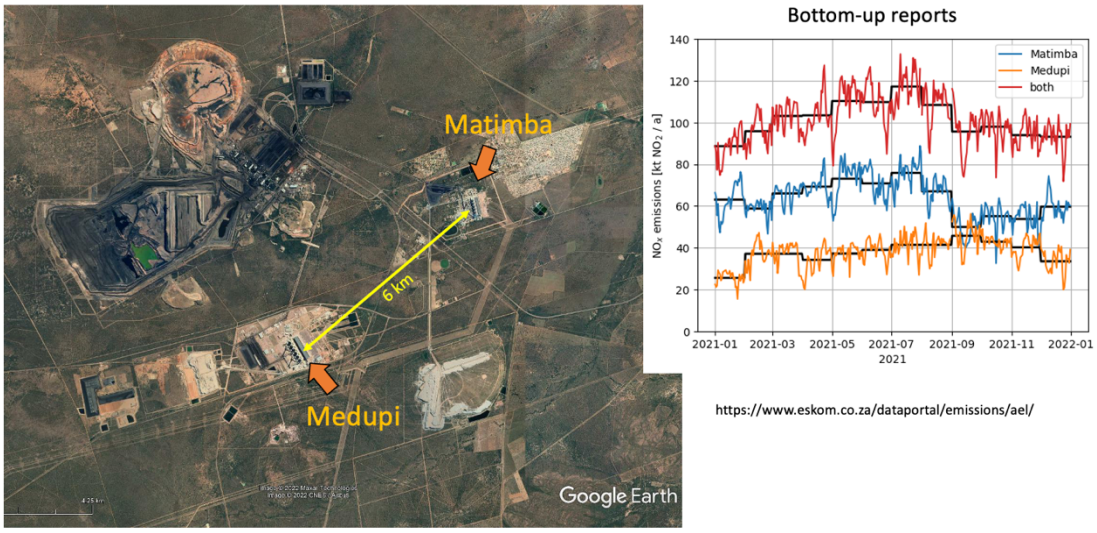


Figure 17 Matimba and Medupi power stations in South Africa (© Google Earth 2023). The Eskom-reported NO_X emissions are illustrated on the right (https://www.eskom.co.za/dataportal/emissions/ael/).

Figure 19 shows the previous OCO-2 and TROPOMI emission estimation results for the years 2018–2020. Hakkarainen et al. (2021) found NO_X-to-CO_2 emission ratio of (2.6 ± 0.6) × 10^{-3} for Matimba/Medupi power station. This emission ratio can be used for converting the estimated NO_X emissions to CO_2 emissions.

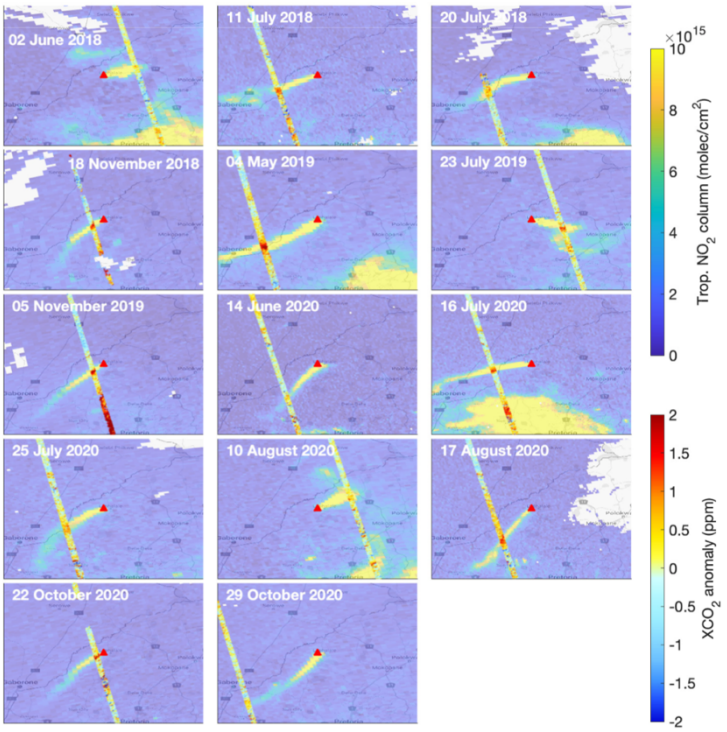


Figure 18 OCO-2 and TROPOMI observations near Matimba/Medupi power station (red triangle) in South Africa between May 2018 and November 2020. Figure taken from (Hakkarainen et al., 2021).

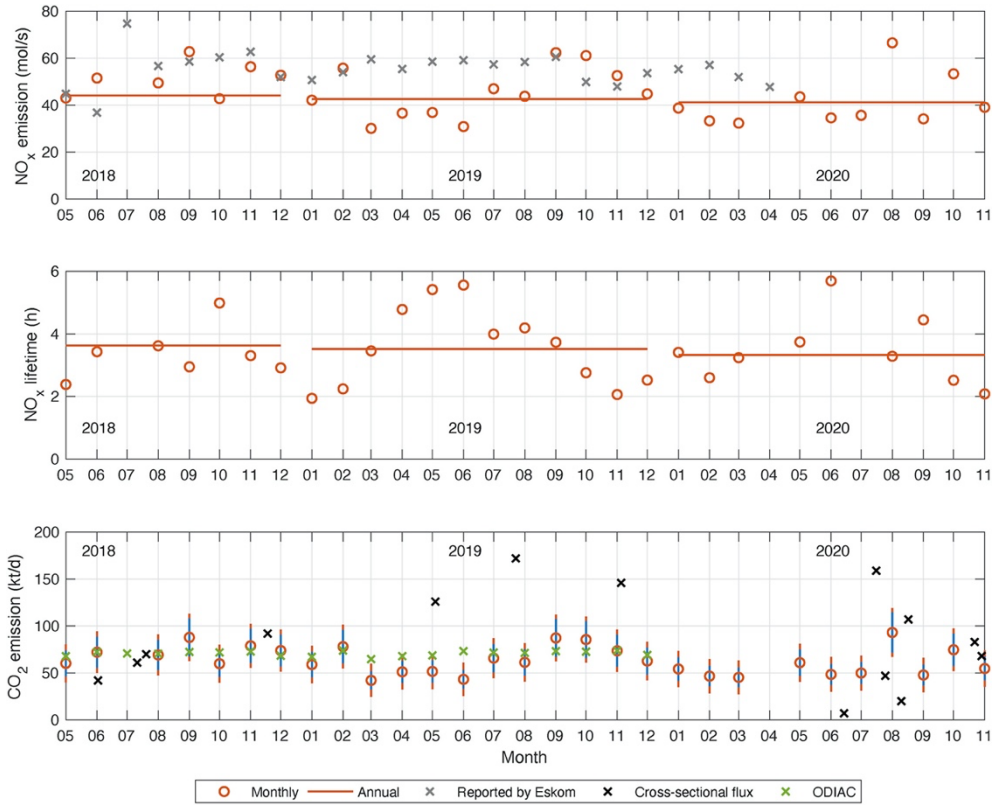


Figure 19 OCO-2 and TROPOMI results for Matimba/Medupi power station. Figure taken from (Hakkarainen et al., 2021).

6.3 Sentinel 5P/TROPOMI dataset

The TROPOspheric Monitoring Instrument (Veefkind et al., 2012) was launched onboard the Copernicus Sentinel-5P satellite on 13 October 2017. The satellite has an equatorial crossing time of 13:30 LT. The current spatial resolution is 5.5 km by 3.5 km at nadir and it covers about 2600 km wide swath. Here we use the tropospheric NO_2 vertical columns with quality flag screening. Due to changes in the operational algorithm, we use the version v02.03.01 intermediate reprocessing on the S5P-PAL system https://data-portal.s5p-pal.com/, which provides a seamless connection with the operational version 2.3.1 data product. Technical details can be found from the readme file: https://data-portal.s5p-pal.com/product-docs/no2/PAL_reprocessing_NO2_v02.03.01_20211215.pdf.

6.4 Results

We first analyze the emissions using the methods based on temporal averaging. Figure 20 shows the results using the divergence method. The first panel shows the sinks calculated using the assumed lifetime of four hours, the second panel is the divergence, and the third panel is the emissions, i.e., $E = S + D$. The emissions from Matimba/Medupi power stations are clearly visible and are easy to fit. When a peak function is fitted to the emission map, we obtain the NO_X emissions of 41 kton/year.

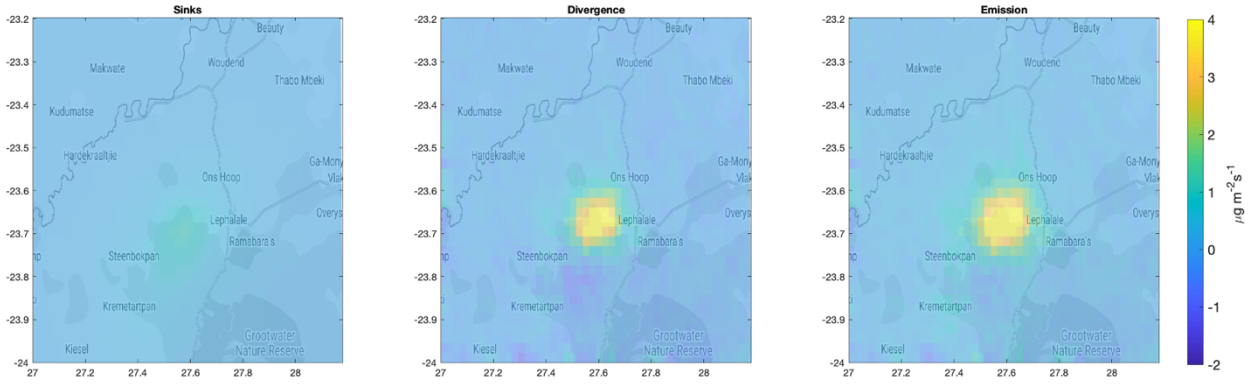


Figure 20 Illustration of the divergence method for Matimba/Medupi power stations in 2021 based on Sentinel-5P/TROPOMI NO_2 observations.

Another typical method for estimating the NO_X emissions from satellite observations is the exponentially modified Gaussian (EMG) approach (Beirle et al, 2011; De Foy et al., 2014). This method has been applied to the Matimba power station in recent studies (Hakkarainen et al., 2021; Potts et al., 2022). Figure 21 shows the NO_2 data and the fit using the EMG approach. We obtain the emissions of 63 kton/year and the lifetime of 3.5 hours.

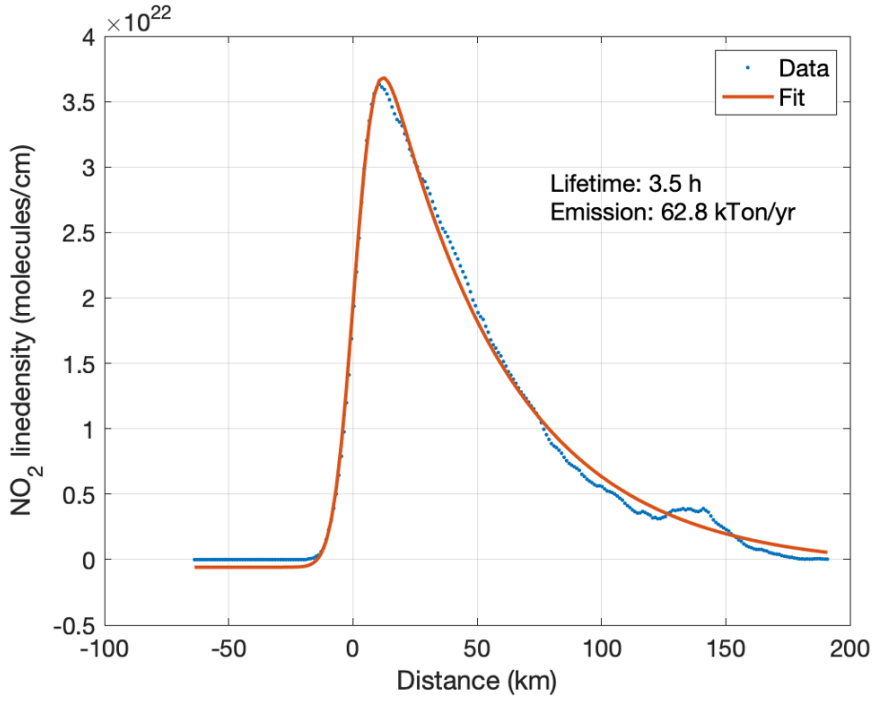


Figure 21 Exponentially modified Gaussian (EMG) method for Matimba/Medupi power stations in 2021.

These values obtained with the EMG method are in excellent agreement with the those reported in recent studies (Hakkarainen et al., 2021; Potts et al., 2022). The emission estimates derived using the divergence method show some larger differences as compared to previous studies. Beirle et al. (2019) obtained the emissions of 54 kton/year, whereas Beirle et al. (2021) obtained the emission of 21 kton/year (after applying correction for different factors such as AMF, lifetime, and winds). As discussed above, our estimate is 41 kton/year.

The EMG NO_X emissions are about 50% higher than the values obtained with the divergence method. Figure 22 shows the EMG and divergence emission estimates compared to the daily emissions as reported by Eskom (https://www.eskom.co.za/dataportal/emissions/ael/). We note that EMG emissions are in good agreement with reported Matimba emissions as noted also by previous studies. However, both EMG and divergence method underestimate the combined Matimba/Medupi emissions of about 100 kton/year. Part of the reason is that we have used NO_X-to-CO_2 emission ratio of 1.32 that is likely too low as discussed in Section 5.

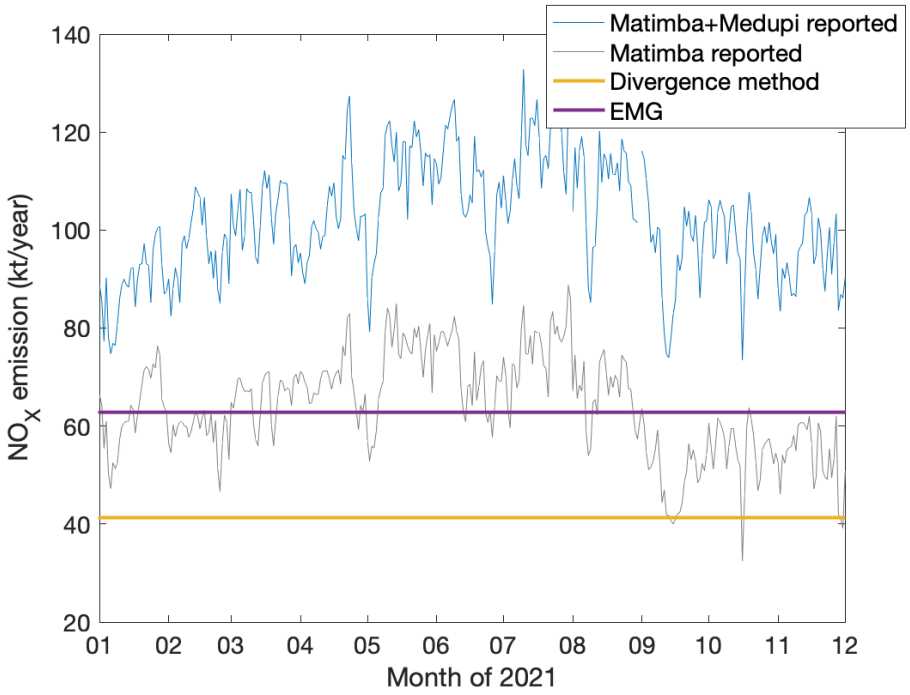


Figure 22 NO_X emission estimates obtained with divergence and EMG method compared against reported emissions.

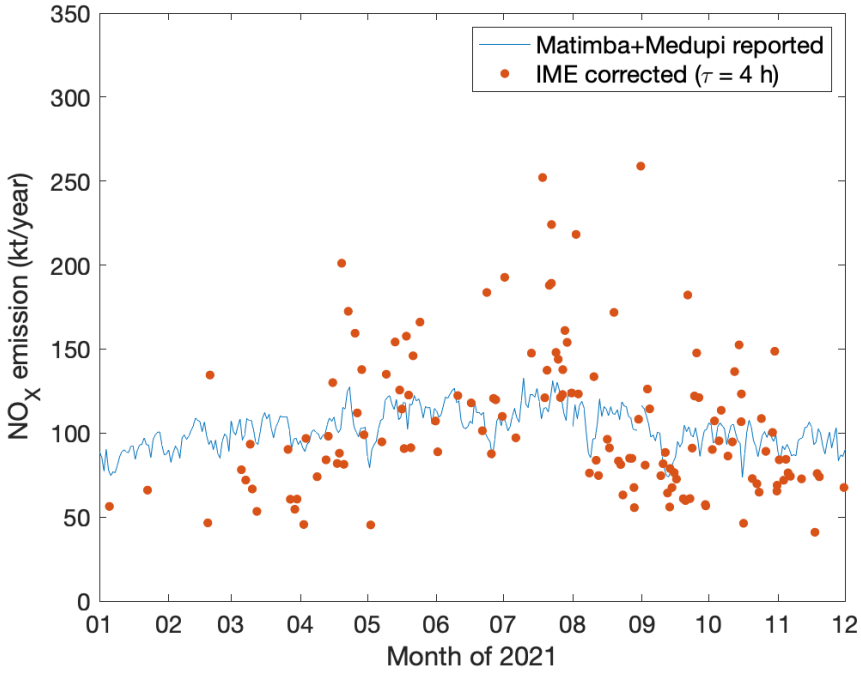


Figure 23 NO_X emission estimates obtained with the IME method (assuming lifetimes of 4 h) compared against reported emissions.

Figure 23 shows the NO_X emissions derived using the integrated mass enhancement (IME) method assuming the NO_X lifetime of 4 hours. We use the conversion factor f = 2.03 as found in Section 5 instead of the conversion factor 1.32 (as used with EMG and divergence methods). We note that during austral summer we obtain lower NO_X estimates than during austral winter. Using the conversion factor 1.32, the annual arithmetic mean is 68 kton/year and the median 59 kton/year, which are in good agreement with the EMG method. Using the conversion factor 2.03, we obtain mean and median of 104 kton/year and 91 kton/year, respectively. These values are in good agreement with the reported emissions.

Figure 24 shows the NO_X emissions estimates using the Gaussian plume (GP) and cross-sectional flux method (CS) using conversion factors f = 1.96 and f= 2.41, respectively. We obtain similar patterns as with the IME method, i.e., higher estimates during austral winter. The annual arithmetic means are 102 and 112 kton/year, respectively. For the GP method, the weighted annual mean and median are 55 and 73 kton/year and for the CS method 113 and 81 kton/year, respectively. For both the GP and CS methods, we do not assume fixed lifetime, but the decay times are fitted (Figure 25 and Figure 26). For both methods, estimated decay times are larger than the 3.5 hours estimated using the EMG method, which is expected since the individual NO_2 plumes do not necessarily follow an exponential decay. The arithmetic and inverse variance weighted mean lifetimes based on the CS and GP methods are reported in the legend of Figure 25 and Figure 26, respectively. The median values are about 20 and 7 hours. We note that the maximum allowed decay time for the Gaussian plume method was set to 200 hours.

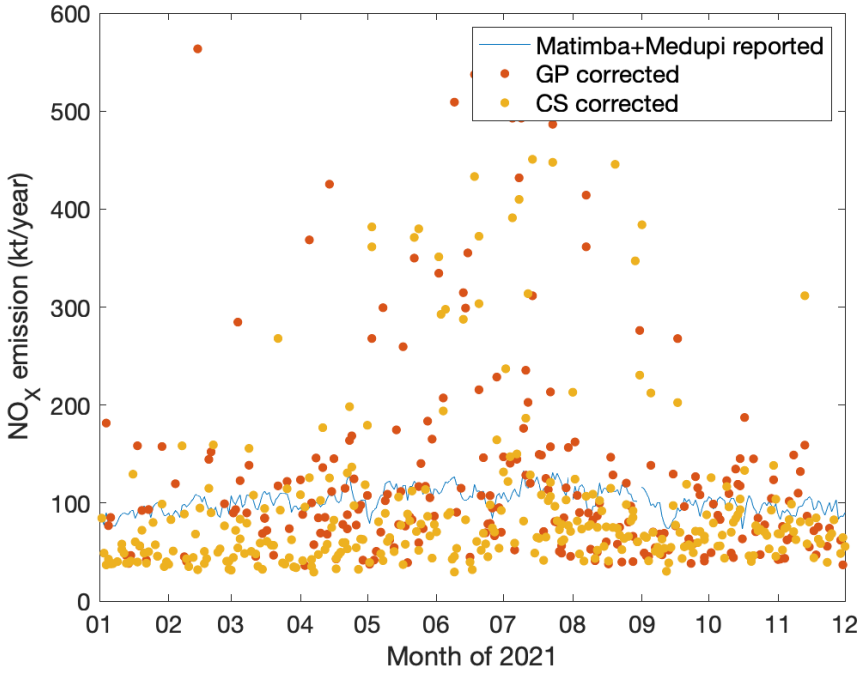


Figure 24 NO_X emission estimates obtained with Gaussian plume (GP) and cross-section flux (CS) method compared against reported emissions.

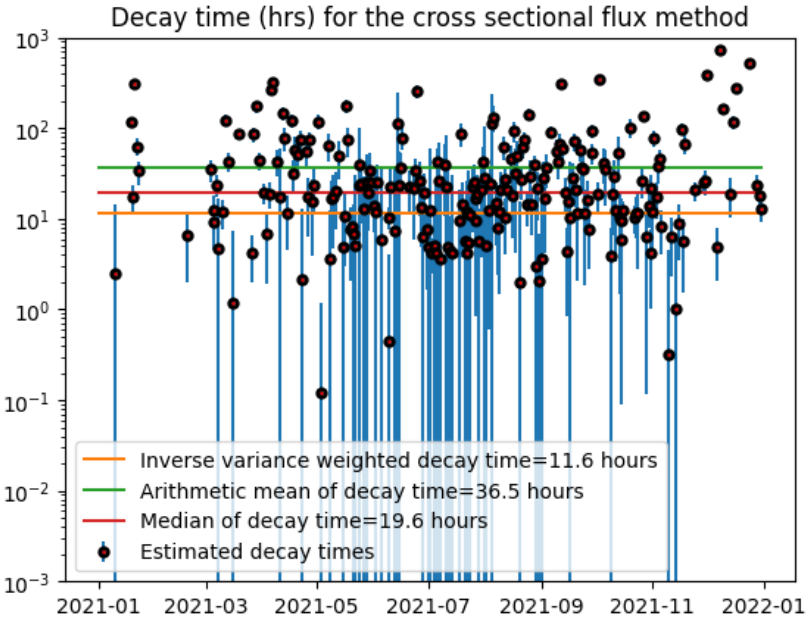


Figure 25 Estimated decay time (in units of hours) for the Matimba/Medupi emissions as estimated with the cross-sectional flux method.

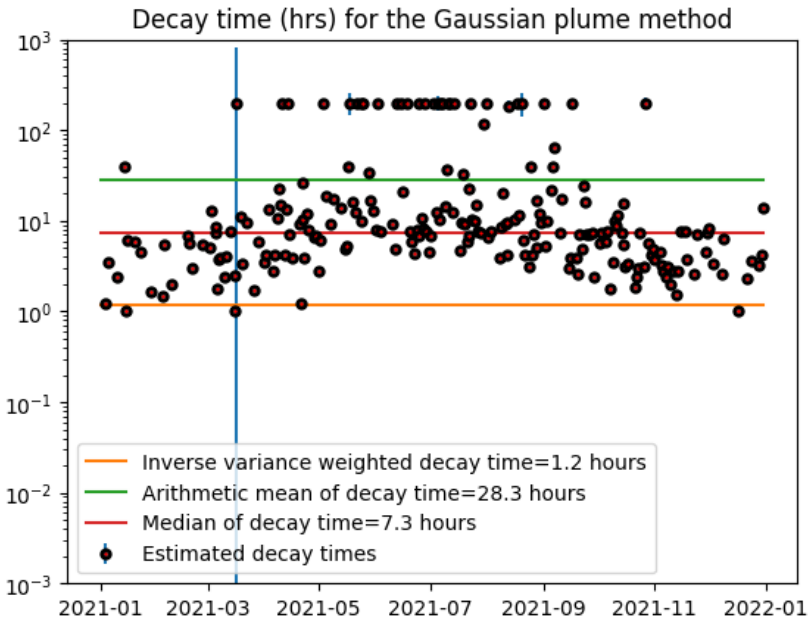


Figure 26 Estimated decay time (in units of hours) for the Matimba/Medupi emissions as estimated with the Gaussian plume method (note the different vertical axis limits compared to the previous image). The maximum allowed decay time for the Gaussian plume method was set to 200 hours.

Figure 27 shows the NO_X emission estimates using the light cross-sectional flux method (LCS) with the conversion factor f = 6. We note that the overall seasonality of the NO_X emissions is similar to the ones observed with the CS, GP and IME methods, i.e., the emissions are higher during austral winter. The median value is 64 kton/year. Using the standard conversion factor f = 1.32, the LCS estimations would be substantially lower. The reason for this high conversion factor value is that with the LCS method most of the information comes from the vicinity of the emission source. Using the LCS method, we were only able to estimate the emissions for 36 days which is relatively low compared to the IME method (147 days), GP method (245 days) and CS method (335 days).

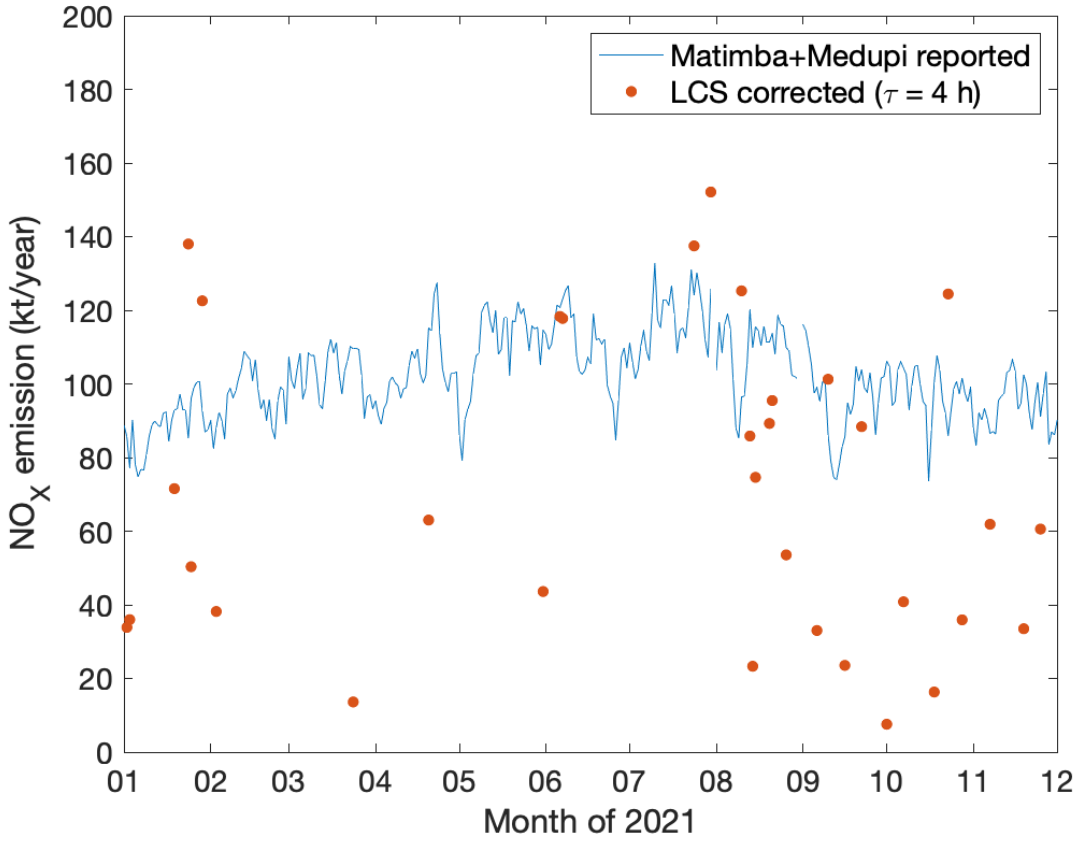


Figure 27 NO_X emission estimates obtained with light cross-sectional flux method (LCS) using the lifetime of four hours and compared against reported emissions.

Figure 28 summarizes the annual emission estimates of different methods. As we don't have specific conversion factors for all the methods, we have used the standard conversion factor 1.32 for all the methods. With this conversion factor, all the methods systematically underestimate the combined reported Matimba/Medupi annual NO_X emissions of about 100 kton/year. The estimated annual emissions vary between 39 and 71 kton/year. The reported values are roughly constant throughout the year. The daily NO_X emission estimates typically underestimate the reported values, although values higher than those reported can be found during the austral winter. This is in-line with previous EMG monthly estimates shown in Figure 19. The arithmetic mean is typically between 60 and 70 kton/year. Medians and weighted means are typically lower. The annual emission estimate obtained with the divergence method (41 kton/year) is comparable with the weighted mean estimate of the GP method (39 kton/year), and also quite close to the typical daily values obtained with the CS method.

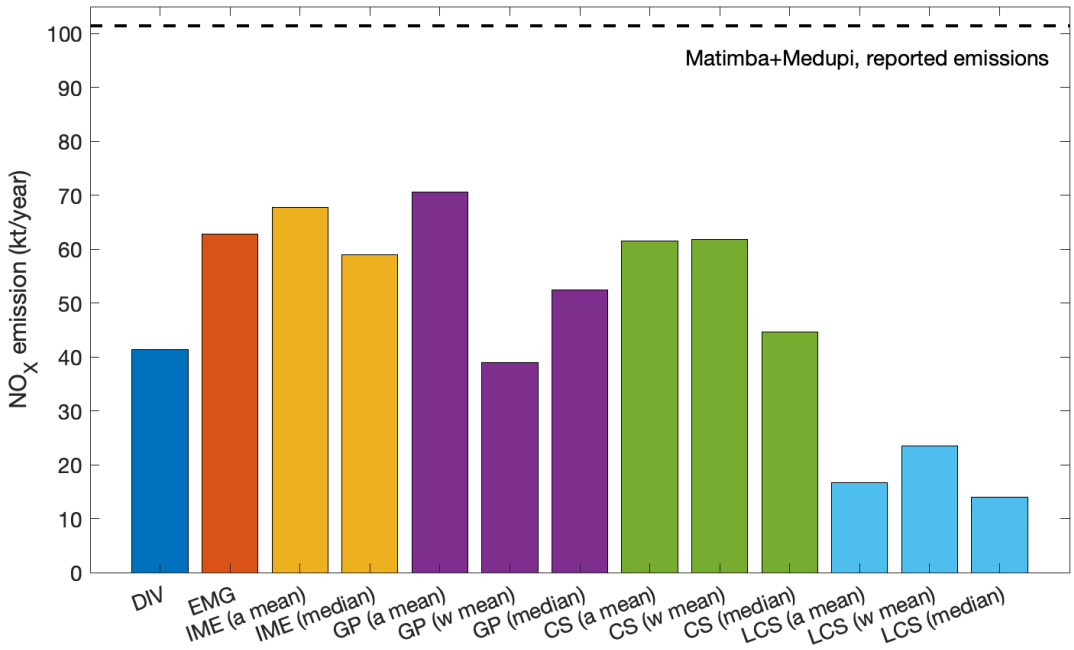


Figure 28 Summary of the annual NO_X emission estimates for Matimba and Medupi power stations using a NO_2-to-NO_X conversion factor f of 1.32.

As noted in Section 5, large part of the reason for the underestimation when using the conversion factor f = 1.32 is that NO is not yet fully oxidized and, as most of the emissions are emitted in the form of NO, this can yield the underestimation of the true emissions. Figure 29 summarizes the annual median emissions for EMG, IME, GP, CS and LCS methods using case specific conversion factors. With these conversion factors, the annual median values are substantially closer to the reported emissions (and the arithmetic means even closer). Assuming that the reported values are correct, these results indicate that the model-based conversions factors are needed to better calculate the true NO_X emissions for the strong point sources like Matimba/Medupi power station. As the MicroHH simulation (analysed in Section 5 was only 48 hours rather than a full year long, we could not calculate specific conversion factors for the divergence method. However, if similar conversion factor (e.g., f = 2.5) would be used, we would obtain results similar to the other methods as also indicated in Figure 28.

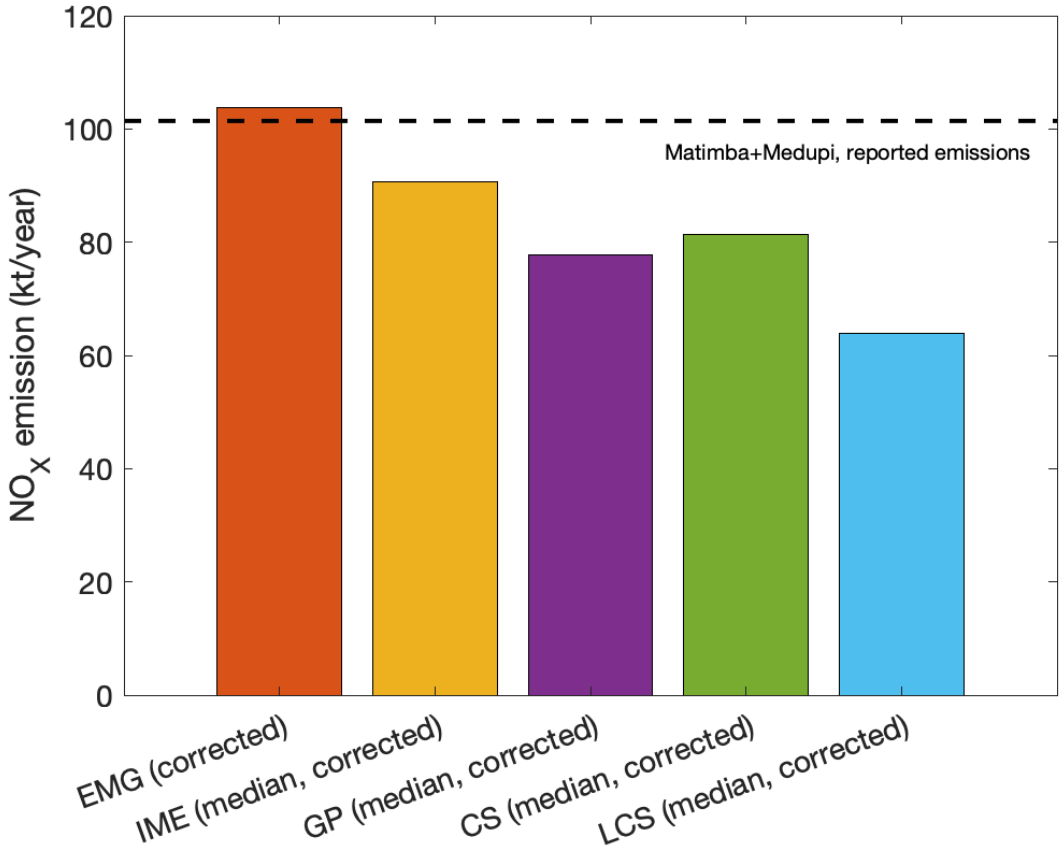


Figure 29 Summary of the annual NO_X emission estimates with specific conversion factors for Matimba and Medupi power stations. See Section 5 for details.

7 Conclusion

In CoCO2 Task 4.2 we have tested and benchmarked computationally light emission estimation methods to quantify local and city scale plumes based on satellite overpass data. The aim is to support the prototype development of the CO_2 Monitoring and Verification Support system which utilizes upcoming CO2M satellite data with 2×2 km pixel resolution. Five computationally light methods for emission detection and quantification have been tested: the Cross-sectional flux method (CS), Gaussian Plume model fitting (GP), the Integrated Mass Enhancement method (IME), Light Cross Sectional flux methods (LCS) and the Divergence method (Div). The main outcomes of the tests and recommendations for future work are discussed below.

7.1 Conclusions based on SMARTCARB benchmarking

One year of synthetic SMARTCARB simulations and synthetic XCO_2 and tropospheric column NO_2 data as observed by CO2M were used to benchmark inversion methods for 16 sources in central Europe in various conditions. The overall performance of each method is linked how 'successful' estimates are selected. More strict screening of acceptable cases improves considerably the performance for individual images especially for IME and CS methods, but it can strongly decrease the number of estimates, e.g., for IME the relative median absolute difference decreases from over 100% to 50% if only 15% of the best estimates are used and CS median decreases from 60% to below 40% when 40% of the best estimates are included.

This also confirms that the error estimates of CS and IME are useful for quality screening. On the contrary, GP and LCS seem to both perform robustly for wide range of estimates but the maximum number of successful estimates for GP remains 20% lower than for LCS. According to the test performed, the accuracy of CS and IME methods seems to depend on the emission source strength – when the source is strong the methods provide better (relative) agreements. This is not the case for LCS and GP methods which seem to work robustly for all emission sources. The advantage of using co-emitted NO_2 signal for plume identification was clear for CS, IME and GP methods for which the number of estimated emissions more than doubled. Adding NO_2 data increased only slightly the already high number of estimates of LCS method.

Clouds strongly hamper the number of successful emission estimates of all methods (based on single images). The number of successful estimates reduces by over 85% for all methods when clouds are considered. The deviation of the estimates from the truth also increases when clouds are considered, however not dramatically; the median of relative absolute deviation increases about 10% for single image methods. On annual level the quality of the estimates decreases also slightly when clouds are considered.

Accurate wind information is known to be crucial for plume estimation techniques. However, sensitivity of the methods on the spatial wind fields was found here insignificant for all methods. This was studied by replacing the SMARTCARB/COSMO winds, used in the simulation, with ERA5 winds.

GP and LCS methods seem to be most robust single image techniques in terms of number of successful emission estimates, low sensitivity of the emission estimate quality (relative absolute difference) on source strength and the impact of clouds. On the other hand, the bias (relative difference) of the LCS method indicates more systematic low bias in the case of strong emission sources (median relative difference of LCS is around -40% while for other methods it is roughly between -25% and 20%). Furthermore, it is worth noting that the current implementation of GP uses as first guess emissions relatively close to the actual emissions (here, the average of the actual emissions over summer is used as a first guess).

In the most realistic benchmarking case (ERA5 winds, cloudy case, using both XCO_2 and NO_2 data) the median relative agreement of annual emissions of the 16 sources was -40% (GP), -23% (CS), -36% (IME) and 21% (LCS). As the relative uncertainty is larger in case of smaller sources these summary values tend to weight more low-emitting sources. Analysis of the annual averages allows also benchmarking Div methods. Out of the 16 sources Div method succeeds to provide estimates for 10 sources with median relative agreement 20% (Div local) and 26% (Div regional).

RMSE analysis of annual weighted mean estimates of the 16 sources indicates best agreement for GP and somewhat lower for CS and LCS methods while IME suffers from lowest overall performance. Interestingly, ensemble estimate (mean of the studied methods) provides the best estimate with lowest RMSE (4.16 $MtCO_2$/yr), which corresponds roughly relative RMSE of 43%.

To our disappointment, none of the methods were able to detect even qualitatively the monthly seasonal cycle of the emissions which suggests that further methodological development (applying, e.g., time series modeling methods) is needed to capture such temporal variability in the emissions.

7.2 Conclusions based on the MicroHH model tests

In CoCO2 Task 4.1 library of plumes was developed. We used data from 48-h simulation with MicroHH model for Matimba power plant in South Africa. This experiment allowed studying how nitrous oxides chemistry affect the emission estimates in a controlled experiment. Emission estimates of CO_2 and NO_x of all three methods studied (CS, GP, IME) overestimate the emissions compared to the truth. This 34–74% overestimation (mean bias), depending on

the method, is probably caused by too high wind speed used. However, during daytime (9–15 UTC) biases vanish suggesting that using an effective wind speed computed from the vertical emission profile works well during daytime when the planet boundary layer is well mixed. The estimated decay times of NO_x estimated with the GP and CS method were consistent with commonly used 4h. However, the decay time for NO_2 columns were significantly higher (from 13 to 41h) which is probably caused by the conversion of NO to NO_2 near the source, which is not accounted for in methods, which increases the decay time close to the source.

The estimated conversion factors from NO_2-to-NO_x varied between 2 and 6 for plume length ranging from 100 km to 15 km. This is considerably higher than typically used 1.32. This is specifically relevant if column observations of NO_2 are used to estimate NO_x fluxes or further used to approximate CO_2 fluxes. Further work on understanding the variability of NO_2-to-NO_x conversion factors could be beneficial.

7.3 Conclusions based on Sentinel 5 Precursor / TROPOMI NO_2 tests

All the five light inversion methods were tested with one year of Sentinel 5P / TROPOMI tropospheric column NO_2 data to derive NO_X emission estimates of Matimba-Medupi power plant area in South Africa. Compared to Eskom-reported emissions, each of the methods strongly underestimated the magnitude of the source if commonly used NO_2-to-NO_X conversion factor 1.32 was used. The annual emissions varied between 39 and 71 kton/yr while the reported value was closer to 100 kton/yr. By applying the conversion factors obtained from the simulated MicroHH tests in Matimba the agreement was clearly improved, resulting however still in underestimation of about 10–35% (when annual median values are considered), IME method being closest to the reported values. This shows that model-based conversion factors are beneficial for estimating NO_X emissions. We also note that from the short 48 h MicroHH simulation the conversion factors for Div method were not possible to compute. It should also be noted that the conversion factor and the NOx lifetime will have a seasonal cycle that could not be analyzed with the short period available from MicroHH simulations. Somewhat surprisingly, we also found out that the number of estimates for LCS method was considerably lower than for CS, IME and GP methods and the emissions were also about half compared to the other methods. This is most probably linked to the fact that the method relies on the emission signal close to the source where NO has not yet converted to NO_2.

7.4 Recommendations for future work

The benchmarking tests conducted here turned out to be very beneficial for the development of all the methods. Several rounds of iterations improved the results considerably. We expect also that further improvement of the methods can be obtained, specifically by optimizing the identification of successful inversions (flagging, improved quality indices and error estimation).

Here we studied how NO_2 benefits the CO_2 plume identification, but information of the co-emitted gases could also be utilized more comprehensively for CO_2 emission quantification.

Co-emitted aerosols, e.g., in cities, can also complicate the CO_2 emission quantification. In this benchmarking study aerosols were not considered but it would be important to analyse their effects as well.

As none of the methods was clearly superior compared to the other methods, it may be wise to consider ensemble methods for achieving robustness. The analysis of all the methods naturally needs more resources, but as computational cost of all the methods is quite light not much computational burden is added. Further work is also needed to improve detecting seasonal cycle in the emissions which remained here a challenge. If known, *a priori*, plant-based seasonal and diurnal cycles can be used in the analysis. However, this option is not available for all the emission sources.

8 References

Beirle, S., Boersma, K. F., Platt, U., Lawrence, M. G., and Wagner, T. (2011). Megacity emissions and lifetimes of nitrogen oxides probed from space. Science 333, 1737–1739. doi:10.1126/science.1207824

Beirle, S., Borger, C., Dörner, S., Eskes, H., Kumar, V., de Laat, A., et al. (2021). Catalog of NOx emissions from point sources as derived from the divergence of the NO_2 flux for TROPOMI. Earth System Science Data 13, 2995–3012. doi:10.5194/essd-13-2995-2021

Beirle, S., Borger, C., Dörner, S., Li, A., Hu, Z., Liu, F., et al. (2019). Pinpointing nitrogen oxide emissions from space. Science Advances 5. doi:10.1126/sciadv.aax9800

Bovensmann, H., Buchwitz, M., Burrows, J. P., Reuter, M., Krings, T., Gerilowski, K., et al. (2010). A remote sensing technique for global monitoring of power plant CO_2 emissions from space and related applications. Atmospheric Measurement Techniques 3, 781–811. doi:10.5194/amt-3-781-2010

Brunner, D., Kuhlmann, G., Marshall, J., Clément, V., Fuhrer, O., Broquet, G., et al. (2019). Accounting for the vertical distribution of emissions in atmospheric CO_2 simulations. Atmospheric Chemistry and Physics 19, 4541–4559. doi:10.5194/acp-19-4541-2019

Chevallier, F., Zheng, B., Broquet, G., Ciais, P., Liu, Z., Davis, S. J., et al. (2020). Local anomalies in the column-averaged dry air mole fractions of carbon dioxide across the globe during the first months of the coronavirus recession. Geophysical Research Letters, 47, 2020GL090244. doi:10.1029/2020GL090244.

Chevallier, F., Broquet, G., Zheng, B., Ciais, P., Eldering, A., 2022. Large CO_2 emitters as seen from satellite: Comparison to a gridded global emission inventory. Geophysical Research Letters 49, doi:10.1029/2021GL097540.

Durand, Y., G. Bazalgette Courrèges-Lacoste, C. Pachot, A. Pasquet, A. Chanumolu, Y. Meijer, V. Fernandez, S. Lesschaeve, D. Spilling, A. Dussaux, D. Serre, and F. te Hennepe (2022). Copernicus CO2M mission for monitoring anthropogenic carbon dioxide emissions from space: payload status. Proc. SPIE 12264, Sensors, Systems, and Next-Generation Satellites XXVI, 1226405. doi:10.1117/12.2636158

de Foy, B., Wilkins, J. L., Lu, Z., Streets, D. G., and Duncan, B. N. (2014). Model Evaluation of Methods for Estimating Surface Emissions and Chemical Lifetimes from Satellite Data. Atmos. Environ. 98, 66–77. doi:10.1016/j.atmosenv.2014.08.051

Frankenberg, C., A. K. Thorpe, D. R. Thompson, G. Hulley, E. A. Kort, N. Vance, J. Borchardt, T. Krings, K. Gerilowski, C. Sweeney, S. Conley, B. D. Bue, A. D. Aubrey, S. Hook, R. O. Green (2016). Airborne methane remote measurements reveal heavy-tail flux distribution in Four Corners region. Proc. Natl. Acad. Sci. U.S.A. 113, 9734–9739. doi: 10.1073/pnas.1605617113

Hakkarainen, J., Ialongo, I., Maksyutov, S., and Crisp, D. (2019). Analysis of Four Years of Global XCO_2 Anomalies as Seen by Orbiting Carbon Observatory-2. Remote Sensing 11, 850. doi:10.3390/rs11070850

Hakkarainen, J., Ialongo, I., and Tamminen, J. (2016). Direct space-based observations of anthropogenic CO_2 emission areas from OCO-2. Geophysical Research Letters 43, 11,400–11,406. doi:10.1002/2016GL070885

Hakkarainen, J., Ialongo, I., Koene, E., Szeląg, M., Tamminen, J., Kuhlmann, G., and Brunner, D. (2022). Analyzing local carbon dioxide and nitrogen oxide emissions from space using the divergence method: An application to the synthetic SMARTCARB dataset. Frontiers in Remote Sensing 3. doi:10.3389/frsen.2022.878731.

Hakkarainen, J., Ialongo, I., Oda, T., Szeląg, M. E., O'Dell, C. W., Eldering, A., and Crisp, D. (2023). Building a bridge: Characterizing major anthropogenic point sources in the South African Highveld region using OCO-3 carbon dioxide Snapshot Area Maps and Sentinel-5P/TROPOMI nitrogen dioxide columns. Environmental Research Letters. doi:10.1088/1748-9326/acb837

Hakkarainen, J., Szeląg, M. E., Ialongo, I., Retscher, C., Oda, T., and Crisp, D. (2021). Analyzing nitrogen oxides to carbon dioxide emission ratios from space: A case study of Matimba Power Station in South Africa. Atmospheric Environment: X 10, 100110. doi:10.1016/j.aeaoa.2021.100110

Hill, T. and Nassar, R. (2019). Pixel size and revisit rate requirements for monitoring power plant CO_2 emissions from space. Remote Sensing 11. doi:10.3390/rs11131608

Jacob, D. J. (1999). Introduction to Atmospheric Chemistry (Princeton University Press)

Janssens-Maenhout, G., Pinty, B., Dowell, M., Zunker, H., Andersson, E., Balsamo, G., et al. (2020). Toward an operational anthropogenic CO_2 emissions monitoring and verification support capacity. Bulletin of the American Meteorological Society 101, E1439–E1451. doi:10.1175/BAMS-D-19-0017.1

Koene, E. et al. (2021). Documentation of plume detection and quantification methods. Tech. rep., Empa. CoCO2: Prototype system for a Copernicus CO_2 service. https://coco2-project.eu/node/329

Koene, E. and Brunner, D. (2023). Assessment of plume model performance. Tech. Rep., Empa. CoCO2: Prototype system for a Copernicus CO_2 service. https://coco2-project.eu/node/357

Kuhlmann, G., Broquet, G., Marshall, J., Clément, V., Löscher, A., Meijer, Y., et al. (2019). Detectability of CO2 emission plumes of cities and power plants with the Copernicus Anthropogenic CO2 Monitoring (CO2M) mission. Atmospheric Measurement Techniques 12, 6695–6719. doi:10.5194/amt-12-6695-2019

Kuhlmann, G., Brunner, D., Broquet, G., and Meijer, Y. (2020a). Quantifying CO_2 emissions of a city with the Copernicus Anthropogenic CO2 Monitoring satellite mission. Atmospheric Measurement Techniques 13, 6733–6754. doi:10.5194/amt-13-6733-2020

[Dataset] Kuhlmann, G., Clément, V., Marshall, J., Fuhrer, O., Broquet, G., Schnadt-Poberaj, C., et al. (2020b). Synthetic XCO_2, CO and NO_2 Observations for the CO2M and Sentinel-5 Satellites. doi:10.5281/zenodo.4048228

Kuhlmann, G., Henne, S., Meijer, Y., and Brunner, D. (2021). Quantifying CO_2 Emissions of Power Plants With CO_2 and NO_2 Imaging Satellites. Frontiers in Remote Sensing 2, 14. doi:10.3389/frsen.2021.

Meijer, Y. et al. (2020). Copernicus CO_2 Monitoring Mission Requirements Document. Tech. rep., European Space Agency. Issue 3.0, EOP-SM/3088/YM-ym

Nassar, R., Hill, T. G., McLinden, C. A., Wunch, D., Jones, D. B. A., and Crisp, D. (2017). Quantifying CO_2 Emissions From Individual Power Plants From Space. Geophysical Research Letters 44, 10,045–10,053. doi:10.1002/2017GL074702

Nassar R., Moeini O., Mastrogiacomo J.-P., O'Dell C.W., Nelson R.R., Kiel M., Chatterjee A., Eldering A. and Crisp D. (2022). Tracking CO_2 emission reductions from space: A case study at Europe's largest fossil fuel power plant. Frontiers in Remote Sensing 3:1028240. doi: 10.3389/frsen.2022.1028240

Potts, D. A., Timmis, R., Ferranti, E. J. S., and Vande Hey, J. D. (2022). Identifying and accounting for the Coriolis Effect in satellite NO_2 observations and emission estimates. Atmos. Chem. Phys. Discuss. [preprint]. doi:10.5194/acp-2022-599

Reuter, M., Buchwitz, M., Schneising, O., Krautwurst, S., O'Dell, C. W., Richter, A., et al. (2019). Towards monitoring localized CO_2 emissions from space: co-located regional CO_2 and NO_2 enhancements observed by the OCO-2 and S5P satellites. Atmospheric Chemistry and Physics 19, 9371–9383. doi:10.5194/acp-19-9371-2019

Varon, D. J., Jacob, D. J., McKeever, J., Jervis, D., Durak, B. O. A., Xia, Y., et al. (2018). Quantifying methane point sources from fine-scale satellite observations of atmospheric methane plumes. Atmospheric Measurement Techniques 11, 5673–5686. doi:10.5194/amt-11-5673-2018

Zheng, B., Chevallier, F., Ciais, P., Broquet, G., Wang, Y., Lian, J., & Zhao, Y. (2020). Observing carbon dioxide emissions over China's cities and industrial areas with the Orbiting Carbon Observatory-2. *Atmospheric Chemistry and Physics*, *20*, 8501–8510. doi:10.5194/acp-20-8501-2020

9 Appendix A

Tables below complement and summarize the analysis of Section 4: Results of SMARTCARB benchmarking. For each cell, the first number is the median then the IQR and finally the number of estimates. All the inversions are with ERA-5 winds and the results are for all the estimates without any selection (QI = 0).

Table 5 Image estimates: absolute differences.

Method	Cloud-free CO_2	Cloud-free CO_2 & NO_2	Cloudy CO_2 & NO_2
IME	59%, [33%, 157%], 382	94%, [39%, 219%], 1599	73%, [40%, 142%], 90
CS	71%, [29%, 192%], 1419	59%, [27%, 130%], 2011	57%, [28%, 98%], 287
GP	30%, [13%, 61%], 892	38%, [19%, 60%], 1799	46%, [25%, 62%], 274
LCS	44%, [22%, 67%], 2246	45%, [23%,69%], 2326	50%, [26%, 78%], 291

Table 6 Image estimates: differences.

Method	Cloud-free CO_2	Cloud-free CO_2 & NO_2	Cloudy CO_2 & NO_2
IME	45%, [-22%, 155%], 382	21%, [-56%, 154%], 1599	-6%, [-71%, 73%], 90
CS	55%, [-9%, 192%1, 1419	31%, [-24%, 130%], 2011	15%, [-37%, 98%], 287
GP	5%, [-23%, 39%], 892	-19%, [-44%, 21%], 1799	-26%, [-52%, 23%], 274
LCS	-21%, [-52%, 25%], 2246	-23%, [-54%, 22%], 2326	-21%, [-55%, 34%], 291

Table 7 Annual estimates: absolute differences.

Method	Cloud-free CO_2	Cloud-free CO_2 & NO_2	Cloudy CO_2 & NO_2
IME	50%, [28%, 100%], 16	46%. [28%, 59%], 16	39%, [34%, 63%], 16
CS	24%, [10%, 83%], 16	29%, [19%, 44%], 16	29%, [23%, 46%], 16
GP	36%, [7%, 53%], 16	37%, [32%, 45%], 16	40%, [21%, 58%], 16
LCS	17%, [15%, 42%], 16	23%, [18%, 43%], 16	35%, [18%, 41%], 16
DIV local	46%, [11%, 94%], 10	46%, [11%, 94%], 10	32%, [16%, 59%], 10
DIV regional	35%, [13%, 131%], 10	35%, [13%, 131%], 10	32%, [22%, 79%1, 10

Table 8 Annual estimates: differences.

Method	Cloud-free CO_2	Cloud-free CO_2 & NO_2	Cloudy CO_2 & NO_2
IME	6%, l-30%, 100%], 16	-25%, [-52%, 12%], 16	-36%, [-60%, -14%], 16
CS	24%, [-4%, 83%], 16	-10%, [-22%, 44%], 16	-23%, [-30%, 11%], 16
GP	-22%, [-53%, 0%], 16	-37%, [-45%, -32%], 16	-40%, [-58%, -21%], 16
LCS	17%, [11%, 42%], 16	23%, [14%, 43%], 16	21%, [0%, 41%], 16
DIV local	46%, [4%, 94%] , 10	46%, [4%, 94%], 10	20%, [6%, 55%], 10
DIV regional	35%, [5%, 131%], 10	35%. [5%. 131%1. 10	26%, [11%, 79%], 10

Document History

Version	Author(s)	Date	Changes
0.1	FMI/LSCE/Empa	02/02/2023	Initial version
1.0	FMI/LSCE/Empa	20/02/2023	Final version based on reviews

Internal Review History

Internal Reviewers	Date	Comments
Marc Guevara Vilardell (BSC)	10/02/2023	Please find my remarks attached. There are not many, mainly in the summary and conclusions.

Estimated Effort Contribution per Partner

Partner	Effort
Organisation	effort in person month
FMI	4
Empa	5
LSCE	5
Total	**14**