CoCO2

Prototype system for a Copernicus CO_2 service

Documentation of plume detection and quantification methods

Erik Koene, Dominik Brunner

& Gerrit Kuhlmann

coco2-project.eu

Co-ordinated by ECMWF

CoCO2

Prototype system for a Copernicus CO_2 service

D4.3 Documentation of plume detection and quantification methods

Dissemination Level:	Public
Author(s):	Erik Koene, Dominik Brunner & Gerrit Kuhlmann (Empa), Janne Hakkkarainen (FMI), Joffrey Dumont Le Brazidec (ENPC), Grégoire Broquet (CEA)
Date:	20/12/2021
Version:	1.0
Contractual Delivery Date:	31/12/2021
Work Package/ Task:	WP4/ T4.3
Document Owner:	Empa
Contributors:	Empa/CEA/FMI/ENPC
Status:	Final

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.

Table of Contents

Figures

1 Executive Summary

This document illustrates a family of approaches to quantify CO_2 and NO_2 emissions from point sources using CO_2 and NO_2 images retrieved from the upcoming Copernicus Carbon Dioxide Monitoring (CO2M) satellite constellation, in an automatic and lightweight fashion (i.e., without requiring atmospheric transport simulations, and being computationally cheap to apply globally). The *Background* section covers the scope and materials used for this report in more detail. The *Plume detection and quantification methods* section describes promising methods from the literature in more detail, which either work on images from a single satellite overpass, or on the temporal average of multiple overpasses of a single source. The section *Potential issues and suggested solutions* discusses a number of identified issues that can hinder the detection or quantification steps of these algorithms, as well as steps to mitigate these problems (if possible). Significant issues that are addressed at this stage are overcoming low signal-to-noise ratios and obtaining more accurate estimates of the wind speed and direction. Issues that are harder to overcome are the low annual number of expected plume detections for a given source, and providing good estimates of the background fields necessary to determine the plume enhancements. The approaches described here have been tested on a limited set of synthetic data but will be applied to a larger library of plumes developed within task 4.1, which in turn will be described in Deliverable 4.2. We would like to point out that this document only describes the various methods, while an evaluation of the performance of the different methods will be provided in Deliverable 4.4.

2 Introduction

2.1 Background

2.1.1 Context of this report

To support the ambition of national and EU legislators to substantially lower greenhouse gas (GHG) emissions as ratified in the Paris Agreement on Climate Change, an observation-based "top-down" GHG monitoring system is needed to complement and support the legally binding "bottom-up" reporting in national inventories. For this purpose, the European Commission is establishing an operational anthropogenic GHG emissions Monitoring and Verification Support (MVS) capacity as part of its Copernicus Earth observation programme. For the MVS system to relevantly support policy implementations, it must have the following four capabilities (cf. the EU report by Pinty et al., 2017): (1) detection of emitting hot spots such as megacities or power plants, (2) monitoring the hot spot emissions to assess emission reductions of the activities, (3) assessing emission changes against local reduction targets to monitor impacts of the nationally determined contributions, (4) assessing the national emissions and changes in 5-year time steps to estimate the global stock take. A constellation of two to three CO_2, NO_2, and CH_4 monitoring satellites (CO2M) will be at the core of this MVS system. The satellites, to be launched in 2025, will provide imaging of CO_2, NO_2, and CH_4 at a resolution of about 2 km × 2 km along a 250-km wide swath. This will not only allow the large-scale distribution of the two most important GHGs (CO_2 and CH_4) to be observed, but will also capture the plumes of individual large point sources and cities. The CoCO2 project was established to develop prototype methods that can realize the four envisioned capabilities for the MVS system based on the satellite images.

In Work Package 4.2 of the CoCO2 project, methods are investigated to quantify CO_2 emissions from large point sources and megacities using images of their plumes collected by the future CO2M satellite constellation, in line with the MVS capability steps 1, 2, and 3 as outlined above. The methods are supposed to be lightweight, i.e., they should not require too large computing resources. It is expected that there are around 900 strong point sources with emissions higher than 3.5 Mt CO_2/year, and 150 to 300 cities worldwide that produce emission

plumes large enough to be observed by a CO2M constellation of three satellites about 50 times per year (Wang et al, 2019; Kuhlmann et al., 2021). If atmospheric transport models were used to provide constraints on the estimated emissions for all the images of all individual point sources and cities (e.g., Broquet et al., 2018; Wu et al., 2018; Ye et al., 2020; Lei et al., 2021), up to 60 000 plume-resolving simulations would be required every year (an average of about 7 per hour), which is a challenging computational requirement. Instead, the lightweight methods considered in this document are *not* based on atmospheric transport models. Rather, they take a data-driven approach based on direct observation of signals in the satellite images themselves, and at best use simplified assumptions of plume physics and dynamics along with an estimate of the wind field. As a result, the methods presented here can be applied to large power plants and megacities only. We remark that efforts using atmospheric transport models for localized inversions are undertaken in CoCO2 Work Packages 4.3, 5.3, and 6.

2.1.2 Satellite measurements & SMARTCARB dataset

The CO2M mission is an upcoming constellation of at least two satellites flying in a sun-synchronous low-Earth orbit with equator crossing times around 11:30 local time. CO_2 and NO_2 images will be recorded at a resolution of about 2×2 km^2, with a 250 km swath. The CO_2 observations will have a precision (1-sigma) of 0.7 ppm or better for a vegetation scenario at a solar zenith angle (SZA) of 50 degrees (VEG50 scenario) and a systematic error smaller than 0.5 ppm. The NO_2 observations will have a precision (1-sigma) of 1.5×10^{15} molecules cm^{-2} or better (CO2M MRD v3: Earth and Mission Science Division, 2021).

Synthetic CO2M observations to illustrate the potential and challenges of the type of approaches described in this report are available from an Observing System Simulation Experiment (OSSE) from high-resolution atmospheric transport simulations with the COSMO-GHG model for a domain covering eastern Germany, Poland, and parts of the Czech Republic. The dataset was created for the year 2015 in the SMARTCARB project (Brunner et al., 2019; Kuhlmann et al., 2019), and is publicly available from https://doi.org/10.5281/zenodo.4048227.

2.2 Scope of this deliverable

2.2.1 Objectives of this deliverable

The objective of this report is to document promising plume detection and quantification methods, based on existing methods found in the literature, demonstrating their potential, and identifying possible improvements. Promising "candidate methods" will be further developed and evaluated in Task 4.2 and documented in the upcoming Deliverable 4.4, both with respect to their ability to quantify emissions accurately and their potential to be applied operationally in a future MVS system.

2.2.2 Work performed in this deliverable

This report is based on plume detection and quantification methods identified in the literature. Additional work performed in this deliverable involved testing two promising approaches for plume detection and quantification, highlighting possible strengths and deficiencies, and identifying areas that may require improvement.

2.2.3 Deviations and counter measures

None.

3 Review of plume detection and quantification methods

Two families of approaches for plume detection and quantification were identified in the literature and reviewed. The first type of approach applies to "instantaneous plumes", obtained from single satellite overpasses. We describe three possible plume detection methods for such cases, and four different applicable emission quantification methods. The second family of approaches considers the temporal average of plume fluxes as estimated from multiple images taken at different times. We describe a single plume detection and quantification method for this type of approach.

The family of approaches described below can be used to detect plumes from yet unknown point sources (in an exploratory manner). Detecting unknown point sources is more relevant for methane than CO2, so is not the focus of this report. Instead, the primary envisioned use case for these methods is to systematically apply them to the CO2M images to detect and process all plumes related to known large point sources and cities.

3.1 Plume detection & quantification for instantaneous plume interpretation

3.1.1 Plume detection

3.1.1.1 Using a convolutional neural network

Emission plumes in satellite images are CO_2 or NO_2 enhancements over a background field due to CO_2 or NO_x emissions from cities, power plants, or other facilities. One way to detect (learn the presence, position, and boundaries of) CO_2 plumes is based on supervised machine learning methods that fall within the field of remote sensing image detection (Lary et al. 2016; Maxwell et al. 2018), which solve this task with convolutional neural networks.

A natural first step in plume identification is the construction of a plume presence detection algorithm. A dataset comprising a representative set of so-called "positive" images (with the presence of a plume) and so-called "negative" images (without the presence of a plume) is used as input to a binary convolutional neural network classifier. Such a neural network then learns to recognise the characteristics of an image with a plume, and its presence. This work was done for example with TROPOMI satellite images for NO_2 plumes by Finch et al. (2021), or for power plant smoke plumes by Mommert et al. (2020). To enhance and help the neural network learn the characteristics of a field, additional input data can be used to feed the algorithm, such as meteorological or dynamical information.

A second step in plume identification is the construction of a plume contour, which can be done in several ways. A first method is to build an algorithm based on the plume presence algorithm (Finch et al. 2021). This method has the advantage of relying solely on the use of a cheap labelling that indicates whether an image is positive or not. An alternative method is to directly use a segmentation algorithm such as a U-Net style neural network (Ronneberger et al. 2015) which has already proven effective for other types of data such as detecting power plant smoke plumes (Mommert et al., 2020). However, using a convolutional neural network relies on the use of labelled segmented images (as the neural network must be trained on images which contain the segmentation mask of the plume) which can be very difficult to obtain, and/or which requires a lot of manual work to generate.

3.1.1.2 Using image thresholding and segmentation

An alternative to the plume detection with a neural network described in the previous section is to identify the plumes using image thresholding and segmentation, as described below.

3.1.1.2.1 Forming a plume pixel mask

The first step in the detection of plumes in this approach is the generation of a 'mask' that indicates whether a pixel value is likely enhanced over the background or not. We denote a

satellite image as $I(i,j)$, considering it a 2-D array of pixel values indexed at positions (i,j). The mask $M(i,j)$ indicates whether a value is likely enhanced or not. It can be constructed as

$$M(i,j) = \begin{cases} \text{True}, & \dfrac{F_s[I](i,j) - F_b[I](i,j)}{N(i,j)} > z_0, \\ \text{False}, & \text{otherwise}, \end{cases}$$

where F_s is an operator enhancing the plume signal (e.g., a Gaussian filter which denoises the signal), F_b an operator that suppresses the signal but extracts the background field (e.g., a median filter), N is some scaling factor to suppress noisy signals (e.g., related to the quality of the data), and z_o represents a cut-off value beyond which we consider a pixel to be more likely to be in the plume than not. The masking test can be carried out on small 5×5 patches of pixels compared to a background, such that z_0 can be a test-statistic taken from Student's t-test (Varon et al., 2018), or by comparing individual samples to the background mean, such that z_0 may be taken from a one-sample z-test instead (Kuhlmann et al., 2019). The value for z_0 corresponds to a one-sided confidence level of, e.g., 0.99, that is, a likelihood of 99% or more that a pixel with that value does *not* come from the same distribution as the background.

Variation between these thresholding-based plume detection methods generally comes from the exact definition of the plume-enhancing and background-extracting operators F_s and F_b. These operators are not necessarily linear operators, and may include chaining together separate operations (e.g., a Gaussian smoother followed by a median filter).

In Varon et al. (2018), the signal-coherency-enhancing operator F_s is applied not on the input fields but on the output mask, by smoothing and median-filtering the formed mask (see Figure 1). However, it is preferable to apply the smoothing first for a simple reason: consider a recorded pixel $I(i,j)$ that is lower than its true value due to noise. It may then, erroneously, "fail" the statistical test, and this not be taken into account when applying smoothing on the output mask. Conversely, by first applying a sufficient denoising step, the pixel may correctly "pass" the statistical test, and no smoothing of the output would be required anymore. In this way, the spatial extent of the plume is primarily constrained by the available data, and not by a post-processing filter.

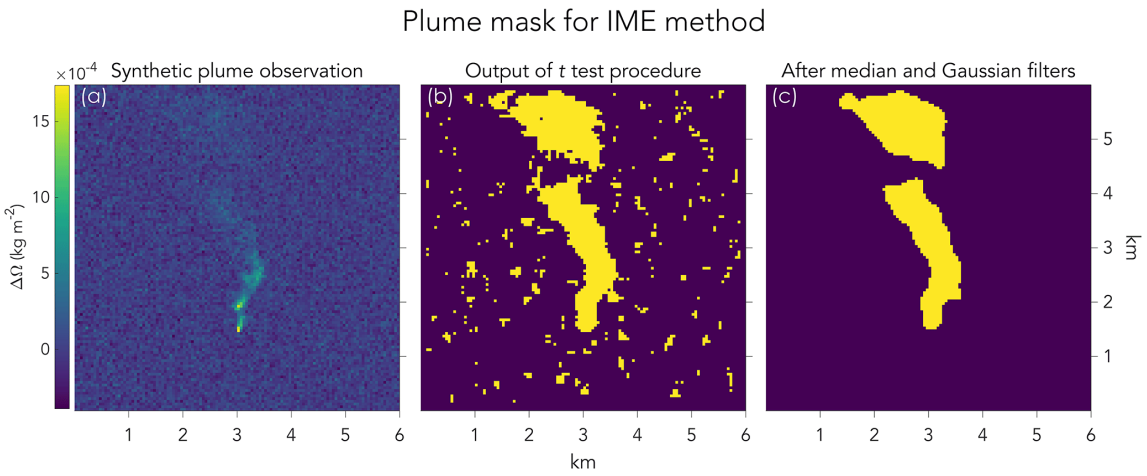

Figure 1: Example of a plume segmentation procedure from Varon et al. (2018). A test of the signal enhancement over the background is performed, and only pixels signficantly enhanced over the background mean are kept.

An advantage of the CO2M mission is that it also provides NO_2 observations. Since NO_2 is co-emitted with CO_2 in the case of high-temperature combustion of fossil and other fuels, these observations are already in widespread use for detecting co-located CO_2 and NO_2 emission plumes and their shape (e.g., Reuter et al, 2019; Hakkarainen et al., 2021). The NO_2 signal is generally easier to detect with current systems due to the better signal-to-noise ratio and lower background values as compared with CO_2. Since the launch of TROPOMI/Sentinel 5P (S5P)

in 2017, it has been possible to observe individual NO_2 emission plumes from single satellite overpasses. For the future CO2M mission, it is expected that a plume will be at least twice as likely to be detected in an NO_2 image compared to its counterpart CO_2 image due to better signal-to-noise properties and less adverse effects due to cloud cover (Kuhlmann et al., 2019). The NO_2 images form a good complement because they are an excellent tracer for CO_2 from anthropogenic high-temperature combustion emissions and, in contrast to CO_2, are hardly affected by biospheric fluxes. An important consequence is that the plume mask can thus be formed based on the NO_2 image alone, which can be used for quantification in both the NO_2 *and* CO_2 plumes at a later stage.

3.1.1.2.2 Segmenting and assigning sources to plumes

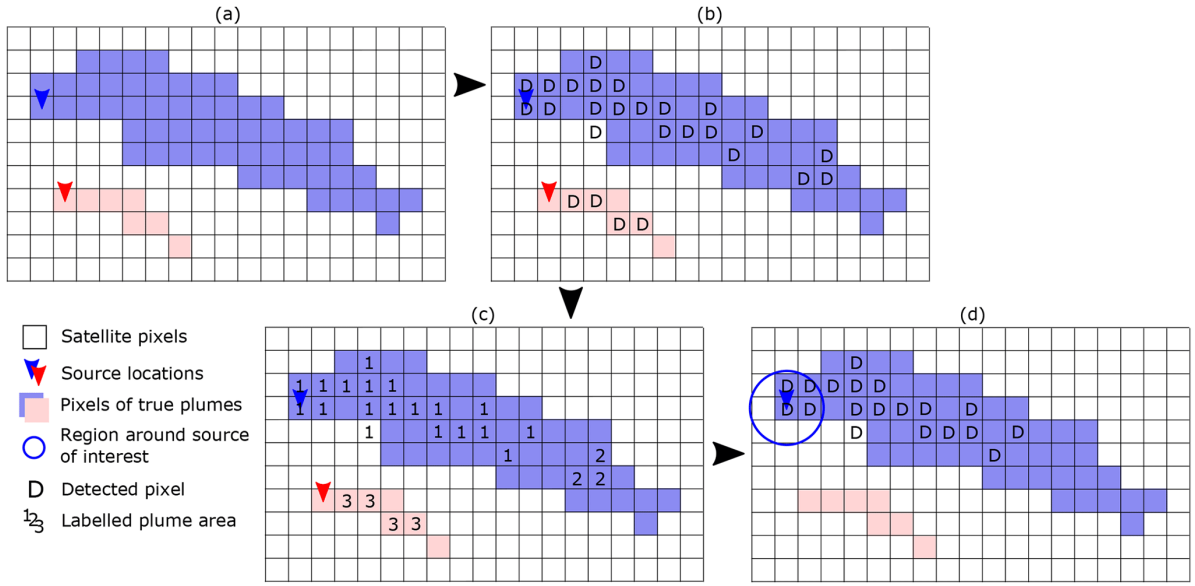

Figure 2: Example of plume segmentation and assignment from Kuhlmann et al. (2019). Panel (a) shows an idealized scenario of two plumes by coloring corresponding pixels blue or red. Panel (b) shows a hypothetical plume mask $M(i,j)$ formed based on a possible noisy realization of this plume scenario. Panel (c) shows how laterally and diagonally connected pixels are assigned a label, restarting each time for a new pixel body. Panel (d) shows how a know source location is used (indicated with the blue arrow and circle around it) to assign plume pixels to a source, only if the source is connected to the plume pixels.

Once the plume mask $M(i,j)$ is in place, one needs to assign pixels to specific plumes. For this purpose, the plumes are segmented (i.e., the image is partitioned into sets of pixels that are connected laterally or diagonally, to form connected plume bodies). Once this process is complete, a list of known locations for point or area sources can be used to determine which plume can be assigned to which source. A schematic example of this procedure is shown in Figure 2. Note that detected plumes, which are *not* connected to a known source location, are discarded at this stage. Furthermore, plume bodies that overlap with multiple possible sources are currently dismissed; see Section 4.1.3 for more details.

3.1.1.3 Using Gaussian model fitting

Next to the neural network approach or the image thresholding and segmentation, a final method identified in the literature of lightweight plume detection methods is fitting a model of an expected plume to the instantaneous images. Figure 3 shows an example of such a method with an application to narrow swath observations from the OCO-2 mission. A model of the form

$$\left\| \underbrace{mx+b}_{linear\ term} + \underbrace{\frac{A}{\sigma\sqrt{2\pi}} e^{-\frac{(x-x_0)^2}{2\sigma^2}}}_{Gaussian\ term} - \underbrace{I(x,y)}_{observed\ image} \right\|_2,$$

is fitted to the observations, where x here corresponds to the along track distance. By obtaining best parameters for m, b, A, σ and x_0 (e.g., in a least-squares error sense), one obtains an idealized description of the plume enhancement (the Gaussian term) over the background (the linear term). Gaussian plume models (Bovensmann et al, 2010) can be used to check whether some of the known emission sources in the vicinity of the measured plume transect are likely to have produced the observed signal. A method such as the cross-sectional flux method (see next section) can then be used to turn the Gaussian model fit into an emission estimate.

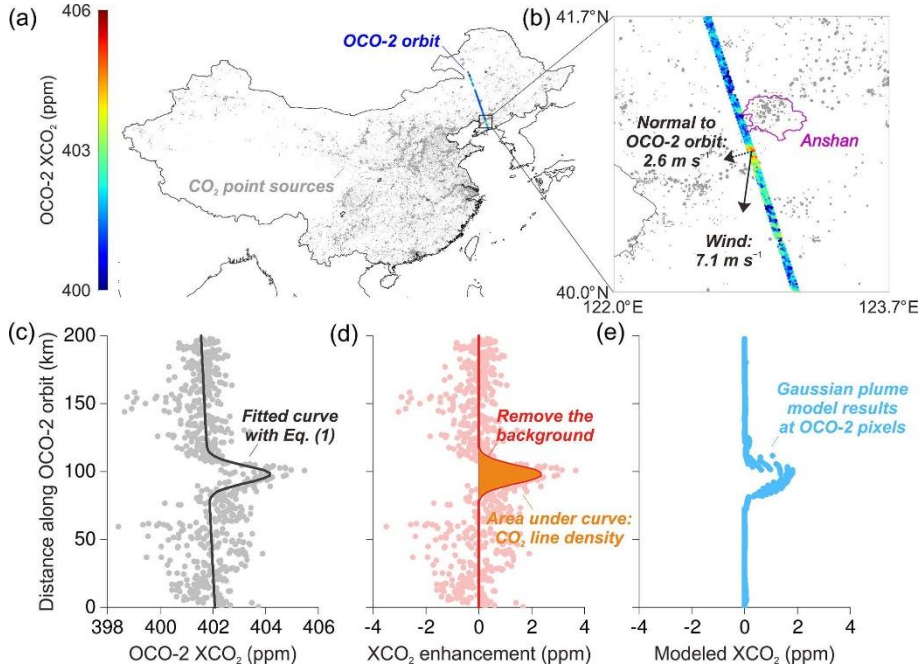

Figure 3: Fitting a linear function plus a Gaussian to detect the cross-section of a plume (not necessarily perpendicular to the along-plume or wind direction) which can be used together with a cross-sectional flux method to estimate emissions at a later point. Figure from Zheng et al. (2020).

3.1.2 Plume quantification

Once a plume is detected in a satellite image, and the plumes have been assigned to emission sources, the next step is to quantify the emission fluxes of the identified plumes.

3.1.2.1 Background estimation

In order to estimate the plume emissions, we require the 'column mass enhancement' of the plume over the present background field. The enhanced quantity $\Delta I(i,j)$ is extracted from the satellite image $I(i,j)$ in the generalized form

$$\Delta I(i,j) = G_s[I](i,j) - G_b[I](i,j),$$

where G_s corresponds to an operator that extracts a denoised estimate of the signal (e.g., a Gaussian filter), and G_b corresponds to an operator that extracts purely the background of the signal (e.g., normalized convolution after removing all plume pixels from the images). These operators are not necessarily the same as those used in the plume detection steps (referred to as F_s and F_b), where it was important to avoid false negatives (i.e., to avoid *not* finding a

plume where one was present). Here, a higher accuracy is required to accurately estimate the background, as an underestimation of the background leads to an overestimation of the column mass enhancement, and vice versa (see Section 4.2.2 for more details). Normalized convolution and/or median filtering applied to all pixels *not* masked as plume pixels are two viable options. The formulation of more advanced operators that estimate the background fields with higher accuracy is still an open research problem.

3.1.2.2 Methods for estimating source rates from plume images

The following four methods for plume emission estimation were outlined in a review by Varon et al. (2018):

1. **Gaussian plume inversion method**. This method corresponds to finding a *Gaussian plume model* that best fits the measured plume. Defining the local column enhancement over the background as ΔI, we attempt to fit a model of the form
$$\Delta I(x,y) = \frac{Q}{U}\frac{e^{-\frac{y^2}{2\sigma_y(x)^2}}}{\sqrt{2\pi}\sigma_y(x)}$$
to the data, where Q is the source flux (thus the quantity of interest), U the wind speed in the downstream direction (denoted with the coordinate x), and $\sigma_y(x)$ describes the horizontal spread of the plume perpendicular to the downstream direction, denoted with coordinate y. The method naturally does not work well for plumes that depart from a Gaussian behaviour, e.g., due to turbulence or other departures from steady-state conditions.
2. **Source pixel method**. This method estimates the source rate by looking merely at the pixel enhancement of the source location, not at the downwind part of the plume. Then we can compute the source rate Q in units of kg/s due to a measurement at the plume pixel,
$$Q = \frac{UWp}{gI_a}\Delta I,$$
where U is the instantaneous wind speed at the source pixel, W represents the pixel dimensions, p the surface pressure, g the gravitational acceleration, I_a represents the column of dry air (kg m^{-2}), and ΔI represents the local column enhancement over the background. The downside of this method is that it ignores a lot of additional information from the plume downwind, and it is more vulnerable to noise and systematic errors in the estimation of the column mass enhancement than the other methods.
3. **Cross-sectional flux method**. This method estimates the flux through one (or more) plume cross-section(s). The cross-sectional flux is computed as a line integral of the column mass enhancement in the direction perpendicular to the plume's long axis (line density) multiplied with the effective wind speed. The integration is, for example, carried out with a discrete summation over the detectable width of the plume.
4. **Integrated mass enhancement (IME) method**. This method estimates the source rate based on integrating the total mass of the plume and dividing by the residence time of the species within the detectable plume. The latter variable is a function of dissipation and turbulent diffusion, for example expressed in terms of an effective wind speed and plume size. This method is non-linear and dependent on relationships drawn between meteorological circumstances and plume properties, perhaps also on the chemical species.

A major challenge common to all methods is finding an appropriate effective wind-speed for the plume, as wind speed may vary in the vertical direction, but the satellite images contain integrated column measurements with no meaningful vertical resolution. This challenge is discussed in more detail in Section 4.2.3.

In the review of Varon et al. (2018), only the cross-sectional flux and integrated mass enhancement methods were considered to be viable methods, free of large errors and robust against plumes in varying meteorological and noisy conditions (e.g., in presence of instrument noise). For the CO2M mission, given the fact that we need global and unbiased emission estimates for a large variety of plumes, the cross-sectional flux method appears to be a well-suited method, except for situations where no clearly identifiable elongated plume evolves due to low winds, a complex clustering of sources, or due to specific flow conditions such as in basin cities like Los Angeles (e.g., Schwandner et al. 2017).

A Gaussian plume inversion method remains a possible alternative to the cross-sectional flux method, given this Work Package focusses primarily on large plumes which should fit a Gaussian shape relatively well. An advantage of the Gaussian plume inversion is that it may be better able to deal with overlapping plumes, by simply fitting multiple Gaussian plumes based on the known point sources. Examples may be found in Nassar et al. (2017) for small satellite swath observations. A general framework for plume inversion using Gaussian plumes may be found in Wang et al. (2020). The performance of this method requires more attention.

Two specific advances in the cross-sectional flux method that can be found in the literature are worth highlighting here. The first is that the line density does not have to be computed for a (set of) single line(s), but can instead be computed for a set of boxes (Kuhlmann et al., 2019). By computing the average line density within the boxes, more pixels become available in each plume to estimate the line densities. See Figure 4 for an example. The second advance is the assumption that the line density in the cross-sectional direction is approximately Gaussian (Reuter et al., 2019). This assumption (as in plume quantification method 1 above) is not valid for complex sources like cities and small turbulent plumes. On the other hand, it allows one to find a best-fitting Gaussian that fits both the CO_2 *and* NO_2 data simultaneously in terms of the location of the centre and width of the Gaussian as expressed through a standard deviation, thus again allowing us to leverage the higher signal-to-noise ratio of plumes in the the NO_2 data to better interpret the CO_2 data. An example of this is also shown in Figure 4.

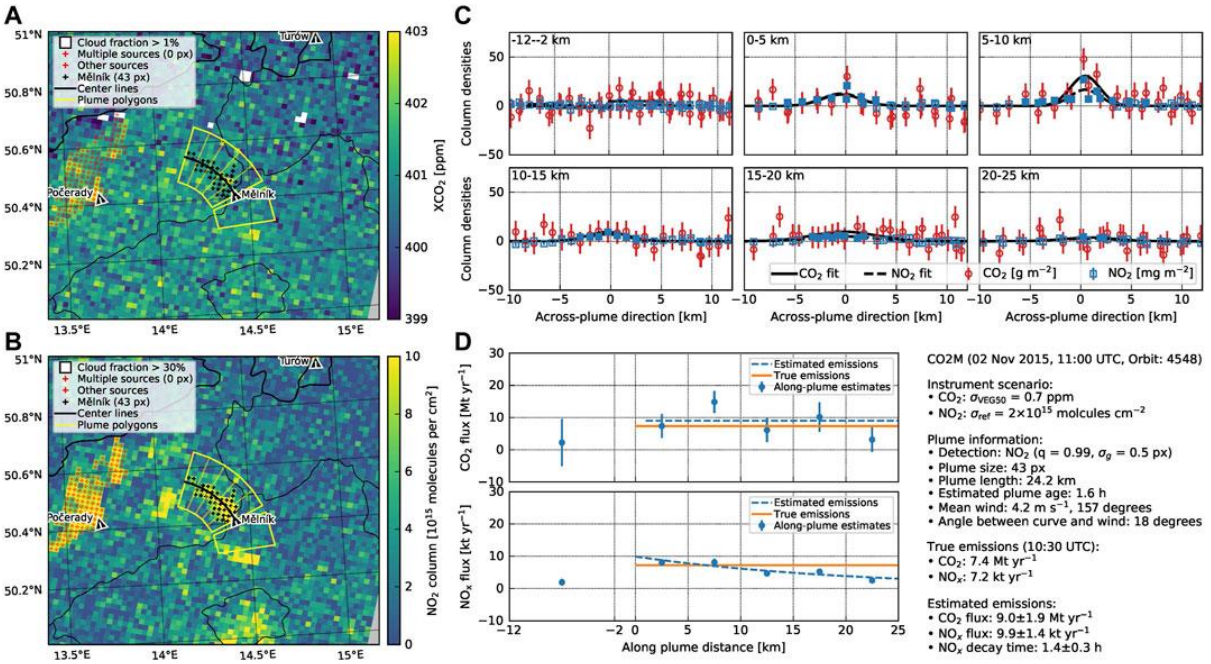

Figure 4: Panel (a) and (b) show examples of detected plumes (black dotted pixels) in the CO_2 and NO_2 images, along with cross-sectional boxes (yellow lines) which are used to estimate the line densities along the plume. Panel (c) shows the cross-sectional line densities along the length of the plume, as estimated from the (noisy) input data. Note how a plume starts out more concentrated close to the source, but due to diffusion becomes wider and flatter along the length of the plume. Panel (d) shows the estimated emissions along the length of the plume for both

CO_2 and NO_2 data. Note how the NO_2 concentration decreases along the plume, as it chemically reacts and thus decays, while the CO_2 concentration remains constant. Figure from Kuhlmann et al. (2021).

By averaging over the estimated fluxes along the plume length, the average flux due to the source may be obtained. Note that for NO_2 data, the flux will decrease along the plume axis as it is chemically depleted, which must be corrected for. It is also necessary to find a suitable NO_2 to NO_x conversion factor. These steps are not required for CO_2 emission estimates. Figure 4 shows an example of the described plume quantification applied to the SMARTCARB dataset.

3.1.2.3 Temporally averaging the data

The methods described thus far allow us to quantify emissions (in units of kg/s) for a single satellite image. To make statements about emissions over longer periods of time, and to take advantage of the detection of a single source in multiple satellite images, one can compute a temporal average of the various computed fluxes, taking into account their expected temporal variations. An example of this principle is shown in Figure 5 with the SMARTCARB dataset, in which plume detection and quantification was carried out for synthetic satellite images for two power plants over a period of one year. A periodic low-order C-spline was used to approximate the time-varying emissions, shown in red, and to compute the annual mean emissions (Kuhlmann et al., 2021). The figure also displays the true emission profile in black, which contains a notable weekly and monthly signal that is not well estimated by the individual estimates. However, the integral of estimated emissions over a full year comes remarkably close to the true emissions.

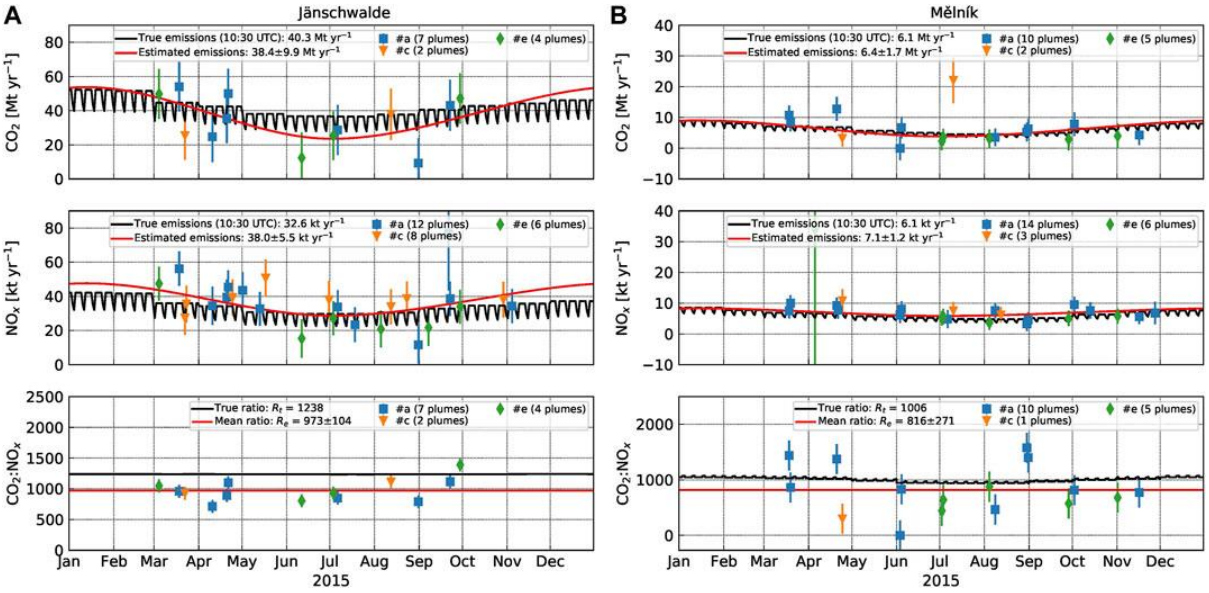

Figure 5: Time series of estimated CO_2 and NO_x emissions and CO_2:NO_x emission ratios at (A) Jänschwalde and (B) Mělník for a constellation of three satellites for high-noise CO_2 and NO_2 observations. The different coloured markers denote the satellite in the constellation. From Kuhlmann et al. (2021).

To better quantify the uncertainties of the final annual estimate of fluxes, one can approximate the standard deviation for the annual average, σ_Q, as

$$\sigma_Q = \sqrt{\frac{1}{n}\left(\sigma_{q0}^2 u^2 + \sigma_u^2 \left(\frac{Q_e}{u}\right)^2\right) + (Q_a m + b)^2 + \frac{s_d^2}{n} + \frac{s_h^2}{n}},$$

where n represents the number of successful estimates, σ_{q0} corresponds to the uncertainty of the line density estimates, u to the wind speed, σ_u the uncertainty of the wind speed, Q_e the estimated flux as computed previously, Q_a is the mean annual flux from the bottom-up inventory, m and b are factors to tune the standard error to the emission strength (Kuhlmann et al., 2021), s_d and s_h correspond to the day-to-day and hour-to-hour uncertainties, respectively, and n corresponds to the number of successful estimates.

An alternative to the method described above is to fit a prior time profile with prior uncertainties and temporal error correlations to the observation-based emission estimates (with their own uncertainties). Prior profiles and their uncertainties can, for example, be taken from Crippa et al. (2020).

3.1.3 ddeq Python package

To illustrate and test the image thresholding and plume quantification methods described thus far, as well as later in this document, we used an implementation of these methods that was prepared by Gerrit Kuhlmann in the programming language Python. The implementation is accessible under https://gitlab.com/empa503/remote-sensing/ddeq. Project partners are invited to contribute to this open source package, e.g. by implementing their plume detection or quantification method as an add-on. Practical examples of the application of the package are illustrated in a Jupyter notebook accessible in the subfolder "notebooks".

3.2 Plume detection & quantification for time-averaged plume interpretation

3.2.1 Theory of the divergence method

In contrast to CO_2, which has a long atmospheric lifetime and large background compared to plume signal enhancements, the lifetime of NO_2 is very short (only a few hours) and the atmospheric background is close to zero. For these reasons, averaging satellite-based NO_2 concentrations over a period of time gives information on the nitrogen oxide (NO_x=NO+NO_2) emission sources. To estimate NO_x emissions from averaged NO_2 columns, statistical methods (i.e., based on multiple spatially co-located observations) are often applied (e.g., Fioletov et al., 2015; Beirle et al., 2011, 2019; de Foy et al., 2014). The advantage of these statistical methods is that they do not require complex atmospheric model runs, and they generally provide more robust emission estimates compared to individual satellite overpasses. In addition, these methods have been successfully applied to instruments and locations where the individual plumes are not detectable on single images, but the emission signal becomes visible when multiple scenes are averaged (e.g., Ialongo et al., 2021). As NO_x and CO_2 are often co-emitted, in some cases it might be advantageous to convert NO_x to CO_2 emissions by using NO_x-to-CO_2 emission ratios. For example, Hakkarainen et al. (2021) proposed a method to obtain source-specific NO_x-to-CO_2 emission ratios from TROPOMI and OCO-2 satellite observations. The method was applied to convert satellite-based monthly NO_x emission estimates to CO_2 emissions for the Matimba Power Station in South Africa. However, to instead apply direct statistical methods to the satellite CO_2 observations to detect anthropogenic sources, we define CO_2 anomalies as the difference to a certain background that accounts for the increasing CO_2 levels in the atmosphere and its spatio-temporal variability (Hakkarainen et al., 2016, 2019). Due to the scarce coverage of current measurements systems, a CO_2 emission estimation with statistical methods has not been successful at present, although options have been discussed by Hakkarainen et al. (2016) and Hill and Nassar (2019).

Recently, the divergence method (Beirle et al., 2019, 2021) has shown great potential in providing robust NO_x emission estimates based on satellite observations. The divergence method is based on the continuity equation (Jacob, 1999) at the steady state, which says that a divergence operator D applied to the vector flux field F corresponds to the difference between emissions E and sinks S:

$$D[F] = \nabla \cdot F = E - S.$$

As discussed by Beirle et al. (2019), the NO_2 sink can be calculated from the NO_2 image I as $S = LI/\tau$, 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 as in Beirle et al., 2011, 2019). In a follow-up paper (Beirle et al., 2021), the sink term was neglected due to the uncertainties in the assumed NO_x lifetime, and only the divergence D was related to the emissions. Since the lifetime of CO_2 is extremely long as compared to NO_x, the sink term can be neglected also for CO_2.

If we apply a temporal averaging operator A to the above relation, we can compute the temporal average of the emissions, while thus neglecting the sink term,

$$A\big[D[F(t)]\big] \approx A[E],$$

where $F(t)$ corresponds to a flux field for specific times, over which the operator A sums.

In practice, for calculating the fluxes, wind information from a meteorological data assimilation product should be used. For the analysis presented in the following, we used the wind information from the European Centre for Medium-Range Weather Forecasts (ECMWF) next-generation reanalysis ERA5 dataset (Hershbach and Dee, 2017; Hoffmann et al., 2019) given at $0.1° \times 0.1°$ grid size resolution and hourly temporal resolution. Following the approach by Fioletov et al. (2015), we used the mean value from 900, 950 and 1000 hPa layers. An alternative could be to use the ERA5 output at the model layers. For each data point, we take the closest point from the wind grid and then temporally interpolate the wind values to the measurement time. The flux F is defined as

$$F(t) = \begin{pmatrix} F_x(t) \\ F_y(t) \end{pmatrix} = \begin{pmatrix} \Delta I(t) u(t) \\ \Delta I(t) v(t) \end{pmatrix},$$

where ΔI is the vertical column enhancement over the background at a specific time t, and where u and v are the eastward and northward winds, respectively. We then regrid F_x and F_y to a $0.05° \times 0.05°$ resolution, and compute the divergence field $D[F(t)] = \frac{\partial F_x(t)}{\partial x} + \frac{\partial F_y(t)}{\partial y} \approx E(t)$. The average emissions are finally computed from averaging all available divergence computations, $A[D[F(t)]] = A[E(t)]$. The partial derivatives are calculated using a second-order accurate central finite difference approximation, e.g.,

$$\left.\frac{\partial F_x}{\partial x}\right|_{i,j} \approx \frac{F(i+1,j) - F(i-1,j)}{x(i+1,j) - x(i-1,j)},$$

where $x(i+1,j)$ and $x(i-1,j)$ correspond to the locations of computed fluxes one longitudinal gridbox east or west, respectively, than the point for which the divergence is computed. A one-sided difference can be used at the edges of the domain.

We note that the averaging operator and divergence operator are linear operators, but their output is not linearly related to the vertical column enhancement over the background. That is, $A[F] \neq A[\Delta I] A[(u,v)^T]$ and $D[F] \neq D[\Delta I] D[(u,v)^T]$. For the divergence operator, the following relation holds due to the product rule,

$$D[F] = \nabla \cdot F = \frac{\partial(\Delta I u)}{\partial x} + \frac{\partial(\Delta I v)}{\partial y} = (\nabla(\Delta I) + (\Delta I)\nabla) \cdot \begin{pmatrix} u(t) \\ v(t) \end{pmatrix},$$

which means that a remaining background field in vertical column enhancement over the background (e.g., $\Delta I + 400$) would influence the result of the computed fluxes. This is why we must not use the total column image I for the divergence method, but the total column image with the background subtracted, $\Delta I = G_s[I] - G_b[I]$, as in Section 3.1.2.1. For NO_2 images, one could assume that the background field is zero, such that the signal enhanced over the background is equal to the measured signal itself. For CO_2 images, we must apply this correction explicitly.

3.2.2 Example using NO_2 column data

Figure 6 illustrates an example of the divergence calculation in the Highveld area (South Africa) based on TROPOMI NO_2 data. The emissions caused by several power stations in the area are visible as enhancements in the divergence map (Figure 6, middle panel). The divergence approach enables the separation between several strong neighboring sources that might appear as mixed in the NO_2 mean fields.

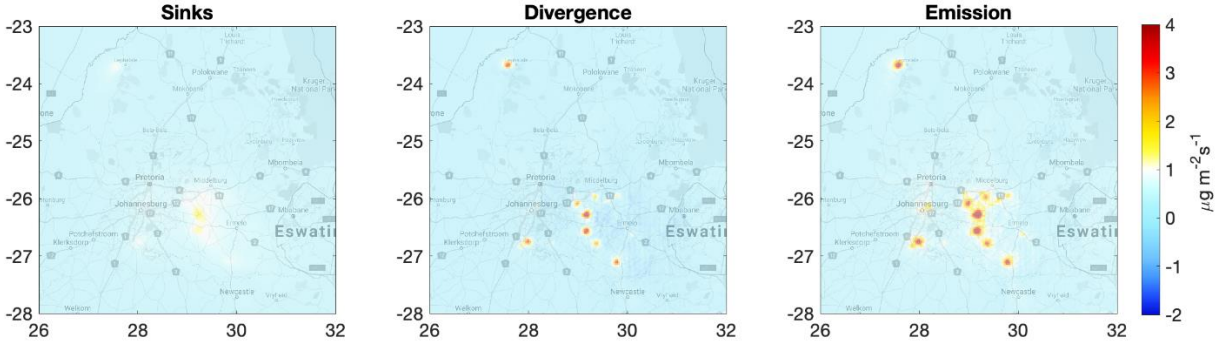

Figure 6: TROPOMI NO_x divergence calculated in the Highveld area (South Africa) for the year 2019 with $0.03° \times 0.03°$ grid size.

To calculate source-specific emissions from the divergence fields, a peak fitting approach (as in Beirle et al., 2021) can be applied by obtaining the best-fitting parameters between sources in the computed field $A\big[D[F(t)]\big]$ and the following Gaussian function:

$$G(x,y) = \frac{Q}{2\pi\sigma_x\sigma_y}\exp\left(-\frac{(x-x_0)^2}{2\sigma_x^2}\right)\exp\left(-\frac{(y-y_0)^2}{2\sigma_y^2}\right) + m_x(x-x_0) + m_y(y-y_0) + b,$$

where Q is the emission flux in units of kg/s, σ_x and σ_y describe the width of the 2-dimentional Gaussian function in the respective directions, and x_0 and y_0 indicate the location of the source. The terms m_x, m_y and b are used to define a linear background field.

3.2.3 Example using CO_2 column data

Similar to NO_2, the divergence map can be calculated also for CO_2. We make a small modification to the algorithm from Beirle et al. (2019, 2021) for the CO_2 case, which has to do with the order in which the divergence operator D and averaging operator A are applied to the data. Both D and A are linear operators and thus commute, i.e., $A\big[D[F(t)]\big] = D[A[F(t)]]$. This property is used by Beirle et al. (2019, 2021) to first average all flux fields before computing the final divergence image. However, we found that the discrete operators are not commutative in the presence of missing pixels (due to, e.g., cloud cover). The issue is that the operator A is defined to ignore missing data as denoted with NaN, e.g., $A[\{1,2,3,\mathrm{NaN}\}] = 2$. This is a sensible choice for A, as it means that we only average those data where and when they are available. However, A and D cease to commute with this definition of the average, as they can sample different data distributions. We illustrate this property in two examples below.

First, we consider a case where all data are present. Consider a flux F at three consecutive locations x_1, x_2, and x_3 at three separate instances in time t_1, t_2, and t_3. We compute the divergence of the fluxes in the last column, and the average of the measurements in the last row. The bottom-right cell then computes the divergence based on the values in the last row, $D[A[F(t)]]$, as well as the average of the computed divergence values as present in the last column, $A[D[F(t)]]$. Both computed values are equal, as expected.

	$F(x_1)$	$F(x_2)$	$F(x_3)$	$D[F] = \frac{F(x_3) - F(x_1)}{2}$
t_1	0	15	14	7
t_2	0	13	11	5.5
t_3	3	17	16	6.5
$A[F(t)]$	1	15	13.66..	$D[A[F(t)]] = 6.33..$ $A[D[F(t)]] = 6.33..$

Second, we consider a case where $F(x_3, t_3)$ is missing. Again, we compute the divergence of the average versus the average of the divergence. Now, the values are not identical.

	$F(x_1)$	$F(x_2)$	$F(x_3)$	$D[F] = \frac{F(x_3) - F(x_1)}{2}$
t_1	0	15	14	7
t_2	0	13	11	5.5
t_3	3	17	NaN	NaN
$A[F(t)]$	1	15	12.5	$D[A[F(t)]] = 5.75$ $A[D[F(t)]] = 6.25$

The reason for the difference in the computed values is simply that $A[F(x_1, t)]$ is computed using three values, whereas $A[F(x_3, t)]$ is computed using only two values. In other words, the average is computed over two slightly different distributions before computing $D[A[F(t)]]$. By computing $A[D[F(t)]]$ instead, the average is always computed over the same distribution of points, not affected by the missing data. Hence, unlike Beirle et al. (2019, 2021), we will compute results for the CO_2 emissions by computing the divergence over the flux fields first, which is then averaged over time. Tests have shown that this improves the divergence images.

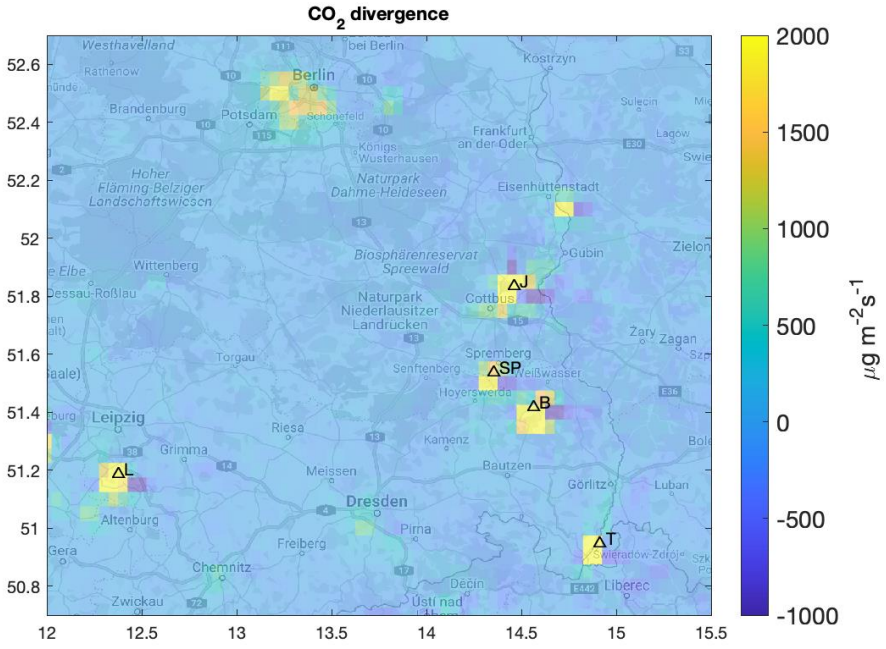

Figure 7: CO_2 divergence calculated from the COSMO-GHG model simulations with $0.05° \times 0.05°$ grid size. Only anthropogenic enhancements are considered. Positive values correspond with strong emissions sources like power stations (Boxberg, Jänschwalde, Lippendorf, Schwarze Pumpe, and Turów marked with B, J, L, SP, and T, respectively) and the city of Berlin.

Figure 7 illustrates an ideal case where the CO_2 divergence is calculated from the CO_2 fields for the year 2015 simulated by the COSMO-GHG model without added noise or a background field. Model simulations at 11:00 UTC are considered, assuming clear-sky conditions. The largest point sources, such as individual power stations (Boxberg, Jänschwalde, Lippendorf, Schwarze Pumpe, and Turów) and the city of Berlin are visible as enhancements in the divergence map.

In general, the situation is more complicated than illustrated in Figure 7, as several aspects affect the divergence calculation. For example, the amount of available data is reduced due to the limited coverage of satellite observations and the persistence of cloudy conditions. The effect of clouds is more restricting for CO_2 compared to NO_2, since completely clear sky conditions are needed for a successful CO_2 retrieval, while partially cloudy conditions (cloud fraction smaller than 30%) are considered suitable for reliable NO_2 retrievals. In addition, the calculation of the CO_2 divergence requires the removal of the background (about 400 ppm, but varying over time) to retain only the signal enhancement over the background. The largest challenge is posed by the instrument noise, ranging between 0.5 and 1.0 ppm (Kuhlmann et al., 2020).

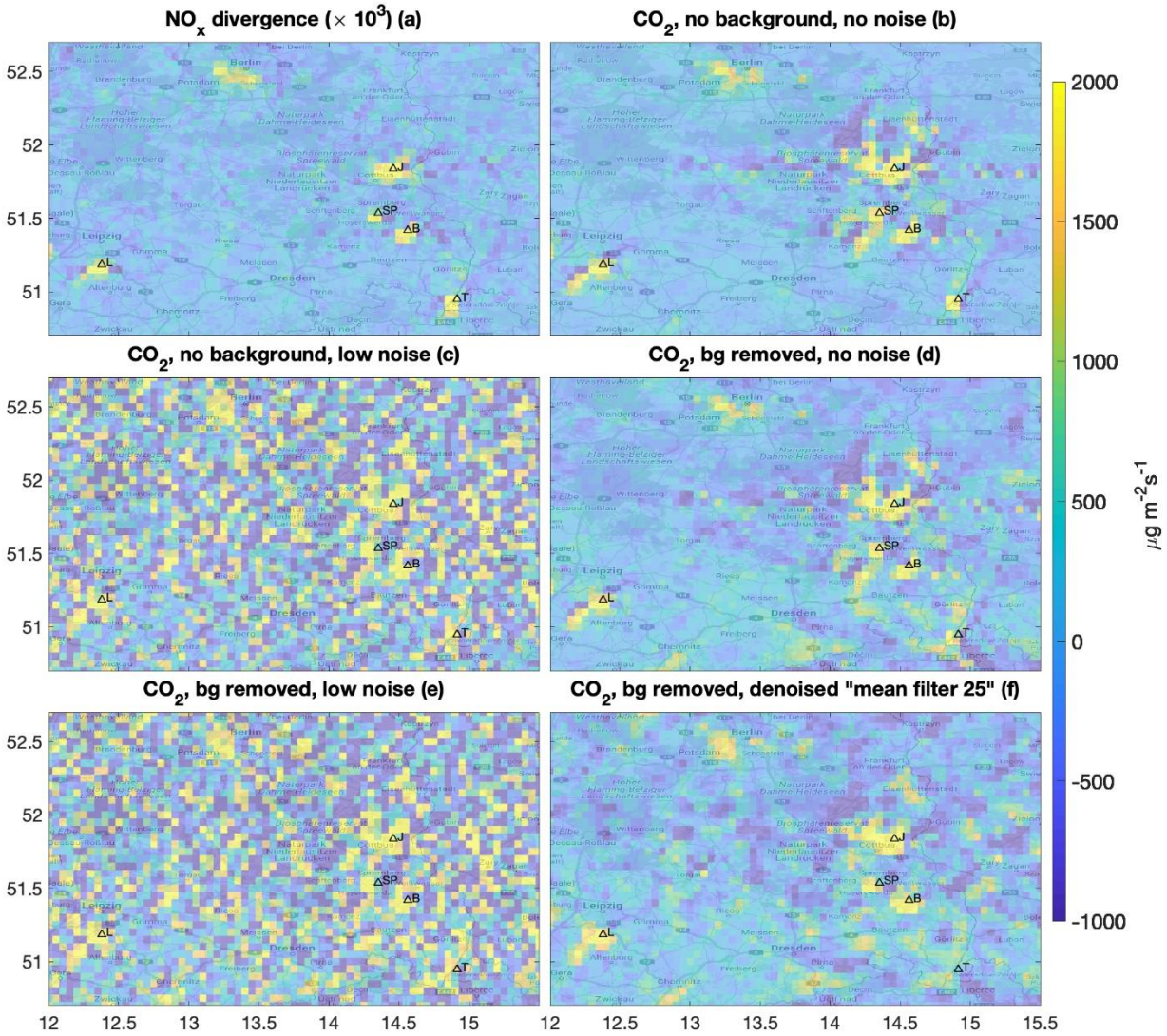

Figure 8: Divergence calculation with different setups using SMARTCARB dataset. See text for an explanation.

Figure 8 illustrates the divergence calculated using the SMARTCARB dataset with different setups. The divergence is calculated using a constellation of two satellites (Kuhlmann et al.,

2019) for the full year 2015, with data filtering for cloud free conditions. For reference, Figure 8a shows the NO_x divergence calculated from simulations with noise (standard deviation 1.5×10^{15} molecules/cm^2), without any background removal and with cloud fraction smaller than 0.3. Figure 8b shows the CO_2 divergence based on simulations without background or noise, and with a cloud fraction limit of 0.01. Both NO_x and CO_2 divergence maps have similar spatial features, with enhancements close to the main emission sources, but the CO_2 fields appear noisier and less sharp, as expected due to the longer lifetime and the more restrictive cloud fraction limit, which reduces the number of available observations. With these setups, there are on average about 32 NO_2 and 16 CO_2 observations per $0.05° \times 0.05°$ grid cell for the whole year. Figure 8c shows the CO_2 divergence map after adding artificial noise (standard deviation 0.5 ppm) to the simulations. Figure 8d shows the CO_2 divergence after adding a simulated background and subsequently removing an approximation of it by calculating the anomaly as

$$\mathrm{XCO_2(anomaly)} = \mathrm{XCO_2(observed)} - \mathrm{XCO_2(background)},$$

where the background is calculated as the daily median over the area covered in Figure 8. The background definition can be tuned case-by-case. Visually, the noise addition has a much larger effect on the divergence than the background addition (and removal). Figure 8e combines Figure 8c and d. Finally, in Figure 8f, the XCO_2 data are denoised by using a 2D convolution with constant 5-by-5 kernel before removing the background and calculating the CO2 divergence. After denoising the data, the CO_2 divergence in Figure 8f is similar to Figure 8b and Figure 8d. Denoising thus appears to be an essential component in the CO_2 divergence calculations. Different denoising strategies are further discussed in Section 4.1.1.

4 Potential issues and suggested solutions

4.1 Difficulties in detection

4.1.1 Improving the signal-to-noise ratio / Denoising

The quality of the input satellite images is of paramount importance for the plume detection and quantification algorithms. Obviously, lower noise levels require less filtering to enhance the signal, and allow more plumes to be detected. Regardless, in Section 3, it was mentioned that typically methods are used to suppress the noise. Below, some of these options are given in more detail.

4.1.1.1 Mean filter

If pixels are *on average* noise-free, but simply contain some amount of random noise added to each pixel, then taking the *local mean* may partially counteract this noise on pixels too. Some examples of local surrounding pixel shapes are given in Figure 9. A cheap way to compute the (local) mean filtered images is to define a small array which sums to one, and convolve this 'kernel' with the image. For example, convolution of an image with the kernel

$$C = \begin{bmatrix} 0 & 0 & 1/13 & 0 & 0 \\ 0 & 1/13 & 1/13 & 1/13 & 0 \\ 1/13 & 1/13 & 1/13 & 1/13 & 1/13 \\ 0 & 1/13 & 1/13 & 1/13 & 0 \\ 0 & 0 & 1/13 & 0 & 0 \end{bmatrix},$$

corresponds to the average of the centre pixel and 12 pixels surrounding it (corresponding to the case '13' in Figure 9).

Unfortunately, mean filters have a drawback, which is that they may average out the "noise", but also the "signal". For example, if the local mean is computed over a patch that contains

both the plume and the background, the computed value corresponds neither accurately to the plume, nor accurately to the background.

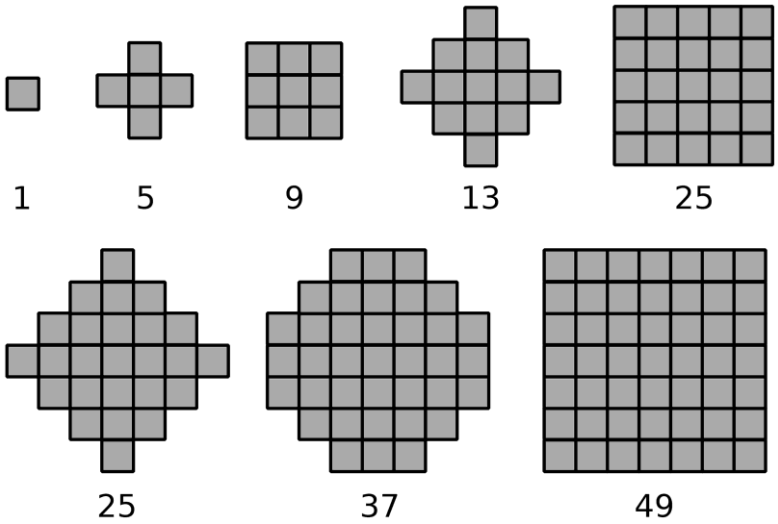

Figure 9: A selection of example shapes that can be used to estimate the local mean or local median based on surrounding pixels. Figure from Kuhlmann et al. (2019).

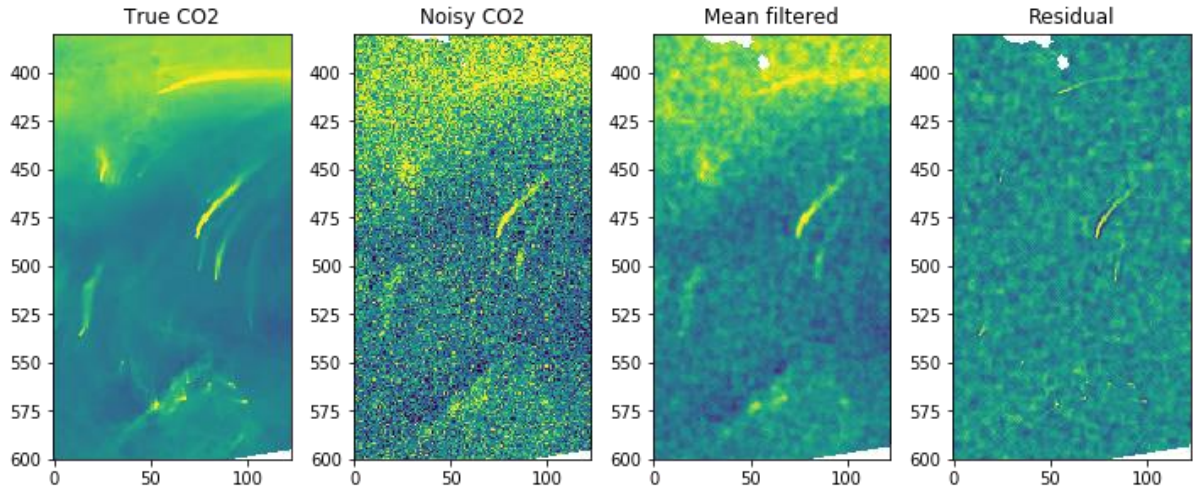

Figure 10: Effect of mean filtering, with the left panel showing the true CO_2 field, the middle-left panel showing a noisy realization of the true field (with discarded data in white), and the middle-right panel shows the mean filtered solution using shape '13' from Figure 9. The rightmost panel shows the residual between the mean filtered and true CO_2 fields.

4.1.1.2 Gaussian smoothing

As an alternative to a mean filter, one can convolve the image with a kernel that weighs locations closer to the centre higher than pixels further away from the centre. For example, this can be done using weights drawn from a 2D Gaussian distribution and normalized again such that the total sum of the kernel equals one.

Convolution with a Gaussian kernel is identical to application of a digital low-pass filter. This clarifies its effect immediately: high-frequency noise is removed, but any high-frequency signal is also affected. For example, if a plume is only a few pixels wide, it may be smoothed out considerably by such a Gaussian based filter. Again, this technique can destroy significant "signal" in the process of removing the "noise". This can be observed in Figure 11, which shows a better denoised field compared to Figure 10, as can for example be seen in the lower

fifth of the image, but the resulting field still appears considerably more smoothed compared to the 'true' CO_2 field.

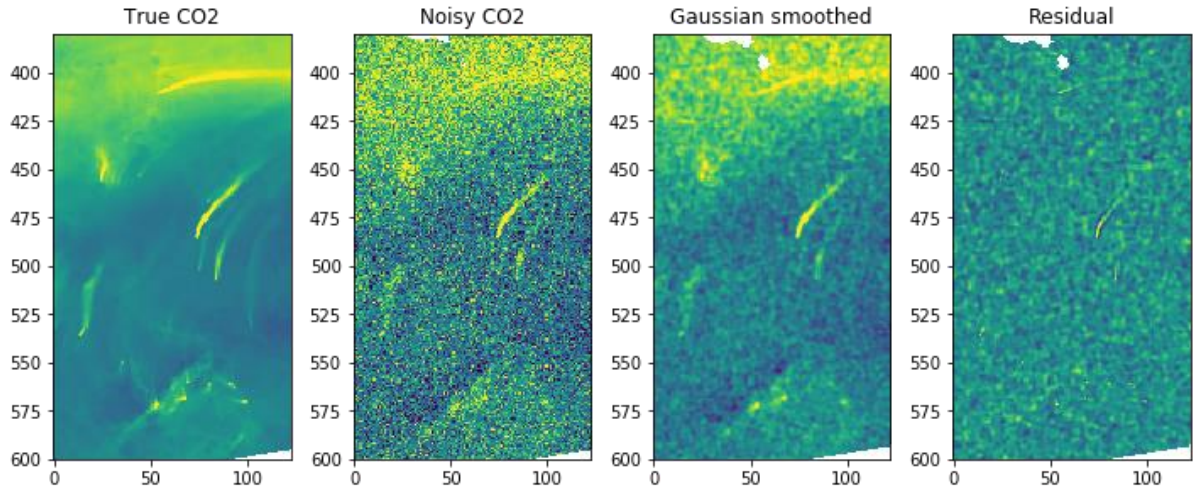

Figure 11: Effect of Gaussian smoothing, with the left panel showing the true CO_2 field, the middle-left panel showing a noisy realization of the true field (with discarded data in white), and the middle-right panel shows the Gaussian smoothed solution for a standard deviation in both horizontal and vertical directions of $\sigma = 1$ pixel; the convolution array has values following shape '13' from Figure 9. The rightmost panel shows the residual between the Gaussian smoothed and true CO_2 fields.

4.1.1.3 Median filtering

Mean filtering and Gaussian smoothing have an additional drawback, which is that they are not robust statistics. The presence of a single outlier (a very high or very low value) may significantly skew the result of the mean or Gaussian filtering operation. A more robust operation is the median, which will not change much if one of the values is replaced with a very high or low value. This property also has its drawback, as it means that application of a median filter will filter out small features, such as small plumes. An example is shown in Figure 12, where the features on the bottom fifth part of the image are less clearly visible than in Figure 11. Median filtering, on the other hand, could be a viable option to estimate a background image, if a filter width is chosen that is several times larger than the size of individual plumes.

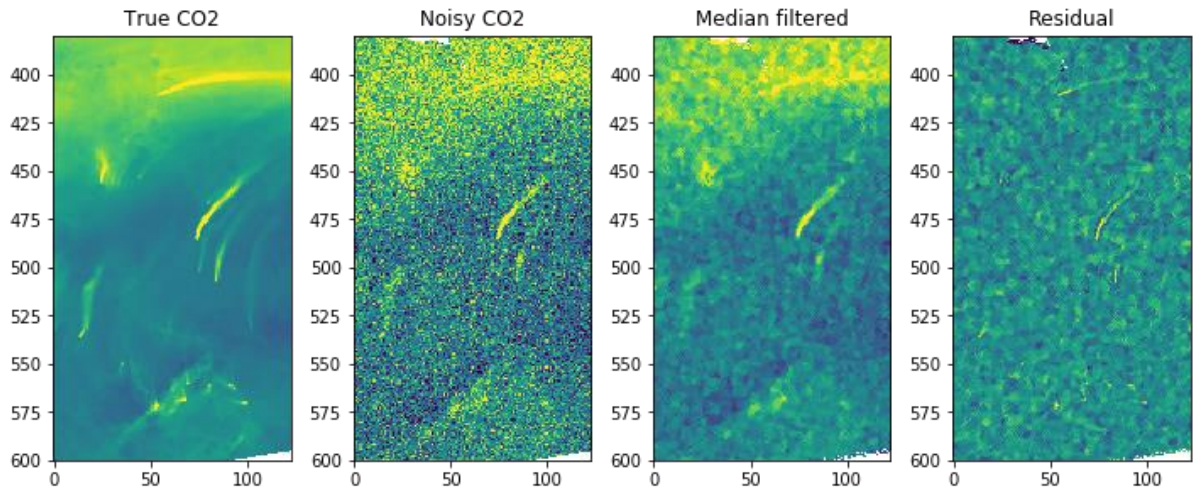

Figure 12: Effect of median filtering, with the left panel showing the true CO_2 field, the middle-left panel showing a noisy realization of the true field (with discarded data in white), and the

middle-right panel shows the median filtered output; the local array used for the median filter has values following shape '13' from Figure 9. The rightmost panel shows the residual between the median filtered and true CO_2 fields.

4.1.1.4 K-SVD

As an alternative to these classical methods of suppressing noise, two more modern denoising methods from the field of Computer Vision were explored. These methods require that the satellite image can be represented using combinations of a small set of base patches (e.g., 8×8 pixels). The assumption is that the patches in the figure repeat multiple times, whereas the noise is not repeating. This assumption is well founded for many natural pictures, and holds for plumes as well: a plume may contain multiple patches where the plume is oriented in the same direction; or multiple plumes are oriented in the same direction.

For K-SVD (which stands for K-means clustering followed by singular value decomposition [SVD]), the steps come down to trying to establish a small "dictionary" of image patches which closely represent the original input signal (Aharon, Elad, & Bruckstein, 2006). The K-means step is used to find clusters of similar-looking patches, to select possible entries for the "dictionary" of patches, while the SVD step is used to reconstruct the image with a small set of dictionary patches.

We use the algorithm in multichannel form, to take the CO_2 and NO_2 images into account jointly. What that means is that it learns to find a sparse set of $8 \times 8 \times 2$ patches in both the CO_2 and NO_2 images. This means that the more information-rich selection of NO_2 patches (due to the higher signal-to-noise ratio of these measurements) is carried over onto the corresponding CO_2 patches.

An example of the procedure is shown in Figure 13. Note how plumes appear sharper in this image compared to the other techniques. For example, the plume with its source around coordinate 75 in the horizontal direction and 500 in the vertical direction can now be distinguished, while this plume was not visible using the previous techniques for improving the signal-to-noise ratio. A test carried out on the entire SMARTCARB dataset shows that the peak signal-to-noise ratio (PSNR) compared to the noise-free inputs increases by +7 dB when moving from the noisy CO_2 image to the K-SVD reconstructed image using both the CO_2 and NO_2 images as input.

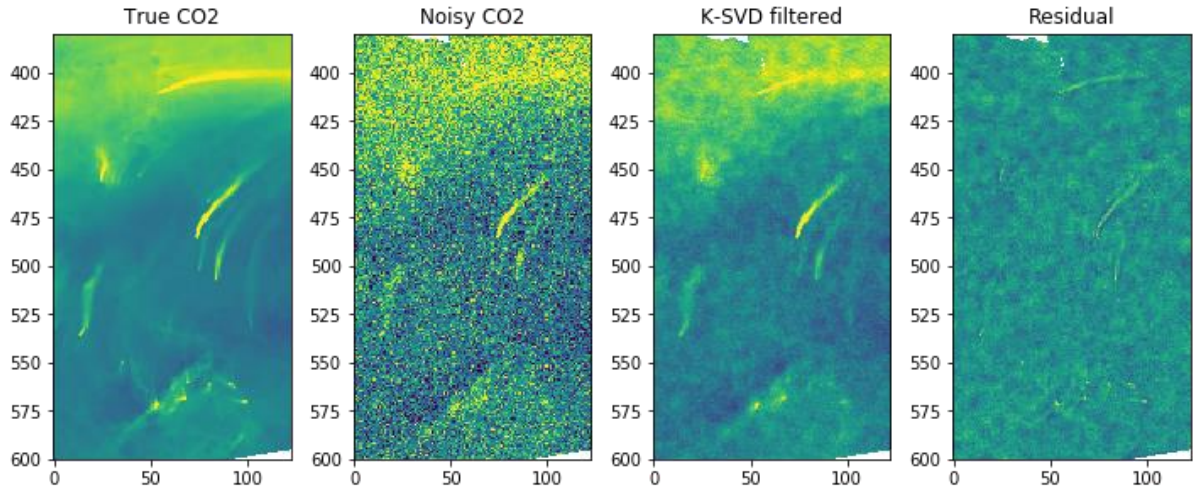

Figure 13: Effect of K-SVD, with the left panel showing the true CO_2 field, the middle-left panel showing a noisy realization of the true field (with discarded data in white), and the middle-right panel shows the K-SVD output (with setting $\sigma = 20$). The noisy NO_2 image used along with the noisy CO_2 image is not shown. The rightmost panel shows the residual between the K-SVD output and true CO_2 fields.

4.1.1.5 BM3D

The second of two Computer Vision techniques that was tested is BM3D (which stands for Block Matching and 3D filtering; Dabov et al., 2007). In this method, N similar image patches (e.g., of size 8×8 pixels) are extracted from the input image to form a block (e.g., of size $8 \times 8 \times N$). These blocks are jointly denoised following three steps:

1. The blocks of patches are domain transformed (e.g., a 3D wavelet transform, a 3D Fourier transform, …);
2. A thresholding step is applied to the transformed blocks, i.e., coefficients falling below a threshold are removed, to retain only the strong components;
3. The resulting block is transformed back to the image domain.

This step is carried out for each patch around each input pixel. The jointly denoised image patches (from steps 1-3 above) are then aggregated into a first estimate of the denoised image. Then, the steps are carried out again, selecting similar image patches now based on the first estimate of the denoised image, and applying a Wiener filter rather than a hard thresholding step. The full algorithm is diagrammatically shown in Figure 14.

We again used a multichannel implementation of this algorithm to exploit the joint information in the CO_2 and NO_2 images. In this instance, that corresponds to selecting the *locations* of similar looking patches from a weighed image of NO_2+CO_2, where both images have first been normalized to cover the same dynamic range. The denoising steps are then performed on the combined image as well as the individual CO_2 or NO_2 images, which are combined at the end to estimate the denoised CO_2 and NO_2 images. As with K-SVD before, the selection of information-rich patches is helped by the higher signal-to-noise ratio of the NO_2 images.

An example of the procedure is shown in Figure 15. We can see that the BM3D algorithm obtains a greatly denoised image, considering the high level of noise when compared to the input image. A test carried out on the entire SMARTCARB dataset shows the PSNR improvement is +13 dB when compared the noisy input images.

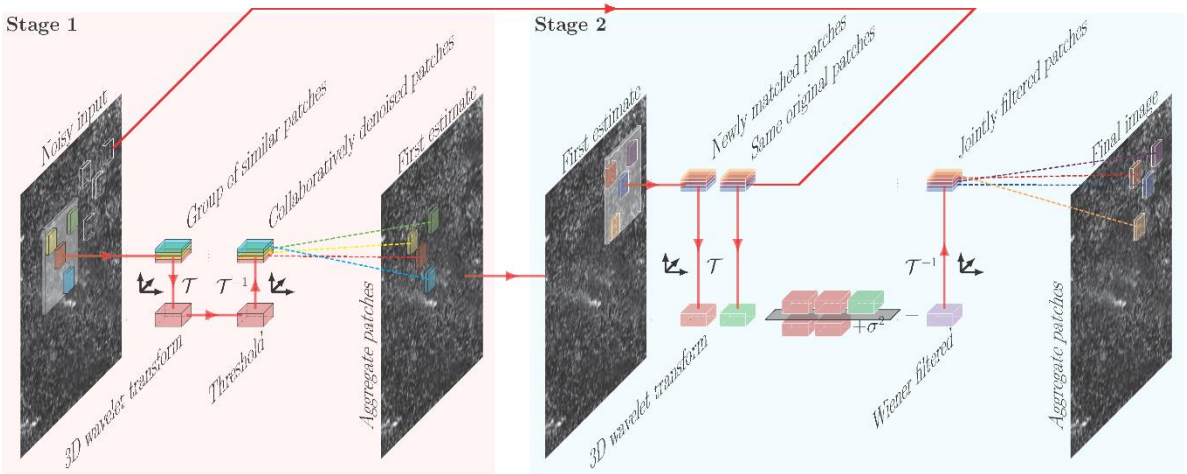

Figure 14: A schematic explanation of BM3D. In stage 1, similar looking patches are collaboratively denoised to produce a first denoised estimate image. In stage 2, similar looking patches are selected from the first estimate, and together with the corresponding patches from the original input form two blocks. Using a Wiener filter, the original image patches are denoised, leading to the final denoised image. The steps are carried out for all patches in the image.

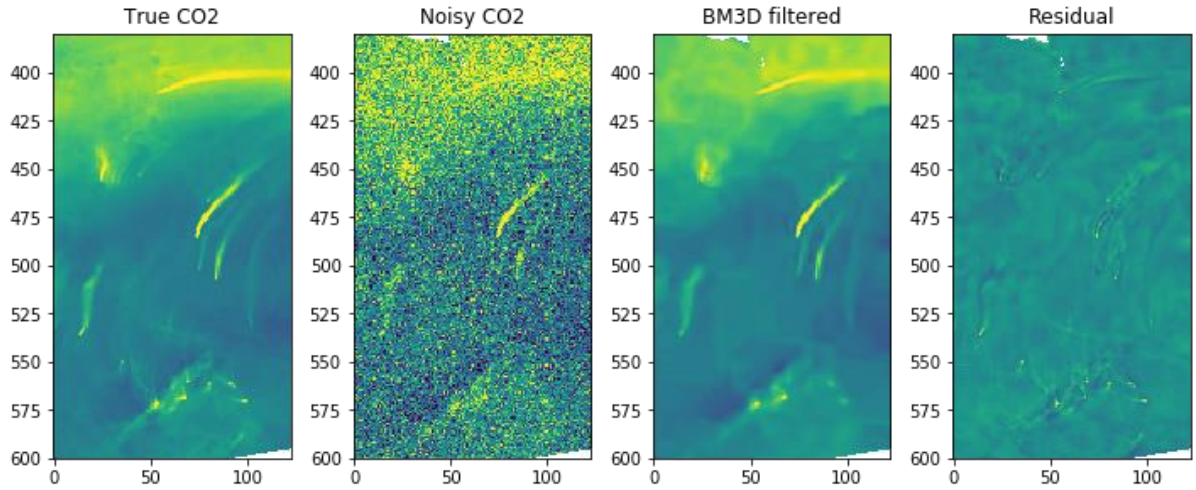

Figure 15: Effect of BM3D, with the left panel showing the true CO_2 field, the middle-left panel showing a noisy realization of the true field (with discarded data in white), and the middle-right panel shows the BM3D output (with $\sigma = 20$). The noisy NO_2 image used along with the noisy CO_2 image is not shown. The rightmost panel shows the residual between the BM3D output and true CO_2 fields.

As a final remark, we will of course not be able to compute the difference between the 'true' and denoised CO_2 images in practice. Instead, we can subtract the noisy CO_2 and filtered images, which should contain instrument noise primarily, not plume-shaped residuals.

4.1.2 Cloud cover (missing data interpolation)

When a satellite image pixel is covered by a cloud cover fraction higher than 1% (for CO_2 satellite images) or 30% (for NO_2 satellite images), no reliable measurement can be made of the corresponding chemical species at that pixel. In practice, this has severe consequences, as was also shown in the SMARTCARB project (Kuhlmann et al., 2019, 2020, 2021): for a constellation of two satellites that operates for a full year, on average only 32 NO_2 and 16 CO_2 observations per $0.05^\circ \times 0.05^\circ$ grid cell are available. This means that plumes will not always be measured fully. Linking back to the plume detection algorithms described in Section 3.1.1, we then also remark that localized cloud cover can split one plume body into two plume bodies; and the plume detection algorithm will discard the plume body not connected to the source point. For this purpose, we have studied modern Computer Vision techniques to connect the plumes where data is otherwise missing, also known as 'inpainting'. Promising results were obtained with a neural network approach following that described in Zeng et al. (2021). The network is trained on SMARTCARB images of NO_2 (as this species contains the most usable pixels available, thus contains the most data in each image to base the inpainting on). Additional input is the cloud cover modelled by the SMARTCARB dataset, and an average of the 10 and 100 m altitude wind speeds from ERA-5, which is turned into streamlines (as the neural network appeared to respond better to image inputs rather than the raw wind data fields). For other factors, we refer to Zeng et al. (2021) to explain the general setup of the neural network and training procedure.

To assess the inpainting performance, we compute the error by comparing the true image I_{true} to the inpainted image $I_{inpainted}$, weighed by the number of cloudy pixel points,

$$E = \frac{\sum_{i,j} \left| I_{true}(i,j) - I_{inpainted}(i,j) \right|}{10^{-10} + \sum_{i,j} cloudcover(i,j)},$$

where $cloudcover$ is an image of the same size as $I_{inpainted}$, which equals 1 for cloudy (inpainted) pixels, and 0 otherwise. Note that I_{true} and $I_{inpainted}$ differ only at the cloudy pixels,

and are identical otherwise. By summing the above error functional for 1528 extracts from the SMARTCARB dataset, and dividing through 1528, we obtain a mean error of 5.9176×10^{15} molecules/cm^2. When not using the wind dataset in the same neural network, the mean absolute error doubles. Hence, the wind field carries relevant information for the neural network to fill the missing data sections. An example is shown in Figure 16.

We note that, at this stage, we do not intend to use the inpainted images *quantitatively*, i.e., to use the inpainted data for estimation of emissions. We merely propose this as a technique to connect otherwise disjointed plume bodies in the plume detection algorithm described thus far. The effect of this technique on the plume detection and quantification performance has not been tested yet, but this will be done in the future.

We note that training the neural network requires up to a day; but using the neural network for inpainting of, e.g., a 123×123 pixel image as shown in Figure 16 below takes much less than a second. Hence, it still is still relatively lightweight computationally.

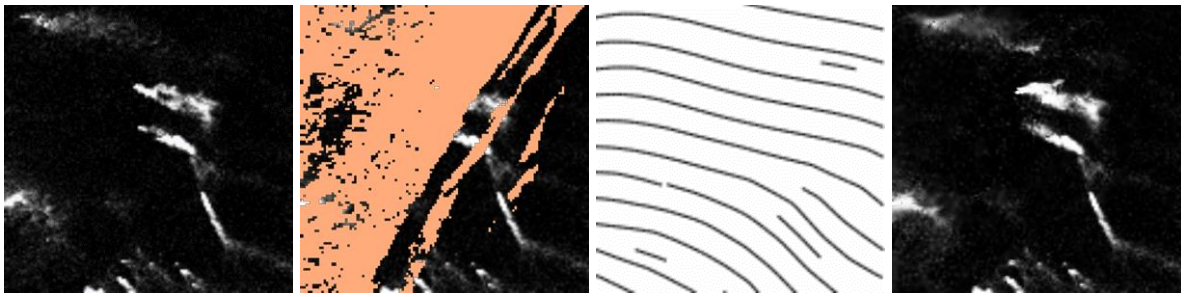

Figure 16: An example of inpainting. The left-most image shows a ground-truth NO_2 image without cloud cover. The second-left image shows this ground truth image covered with clouds (shown in beige); hence, all the beige parts of the image must be filled in by the inpainting algorithm. The second-right image shows the wind field as taken from ERA-5. The right-most image corresponds to the inpainted image. We note that the algorithm has correctly connected previously disjoint plume bodies, despite the fact that a large portion of the image was covered by clouds.

4.1.3 Designation of connected plumes

The plume detection & quantification algorithm described in Section 3.1.1.2 will dismiss plume bodies that are connected to two or more sources rather than to a single source. These plumes may be overlapping while located next to each other (i.e., oriented in parallel, mixing into each other), or a plume may be detected for such a long distance that it overlaps with a plume located (much) further downstream. The reason for the rejection of these plume bodies is clear: the fluxes for such a connected plume body cannot be ascribed to a single source and, furthermore, assumptions of constant fluxes over the entire plume body or approximately Gaussian-shaped cross-sections are not met in these instances.

We plan to address this issue in the coming year, to maximize the number of plume measurements that can be realized by the algorithm. For example, if one plume body overlaps with two or more known source locations, we can use an estimate of the wind field to, at the very least, quantify the upstream part of one or more plumes, before they mix (see Figure 17).

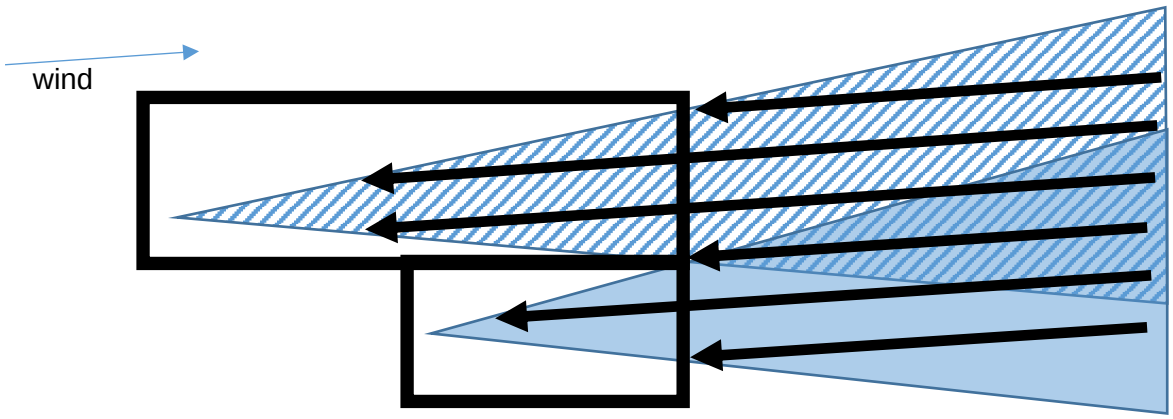

Figure 17: A schematic diagram showing two plumes (in blue hashed form and blue filled form). The two plume bodies are 'connected', in the sense that a plume detection algorithm that would detect both the hashed and filled blue bodies entirely would not detect two separate plume bodies, but the union of them both. However, when moving upstream from points in the plume body (shown with the black arrows), we end up at two separate plume sources. One can at least quantify the unmixed parts of the plumes in these upstream regions, indicated with the two thick black boxes.

When it comes to detecting and quantifying the individual contributions of the plumes once they have mixed, some assumptions begin to break down (such as an approximately Gaussian-shaped cross-sectional profile). However, we will explore whether or not the mixed portion of the plumes can also be used for plume detection and quantification.

4.2 Potential sources of error

4.2.1 Bias through use of synthetic data

The methods described in Section 3.1.1.2 and Section 3.2, as well as the denoising methods of Section 4.1.1 and the inpainting method of Section 4.1.2 have been tested extensively on the SMARTCARB dataset. Parameters have been specifically tuned to perform well on this dataset. There is thus a chance that the methods will have a different performance when used on other data. This assertion can be partially tested in the coming year when the performance of the described plume detection and quantification methods will be tested on a library of plumes under development in Work Package 4.1. A larger OSSE dataset with an accurate noise model, including systematic errors not included in the SMARTCARB dataset, would be useful to test how well the method will work on the future CO2M data, but currently a full error parametrization considering the aerosol information from the multi-angle polarimeter (MAP) is available for only a few days for a domain around the city of Berlin.

4.2.2 Background estimation

As described in Section 3.1.1.2 and Section 3.2, the quantification of plumes is heavily dependent on an accurate estimate of the background field, such that only the *signal enhancement over the background* is used to estimate the emission fluxes. The complexity of this problem is shown in Figure 18: the background field of CO_2 or NO_2 'underneath' the plume enhancement can be strongly varying and strongly heterogeneous. As the satellite image consists only of the combined data (enhancement + background), it essentially forms an underdetermined system: there are more unknowns than there are equations to link them.

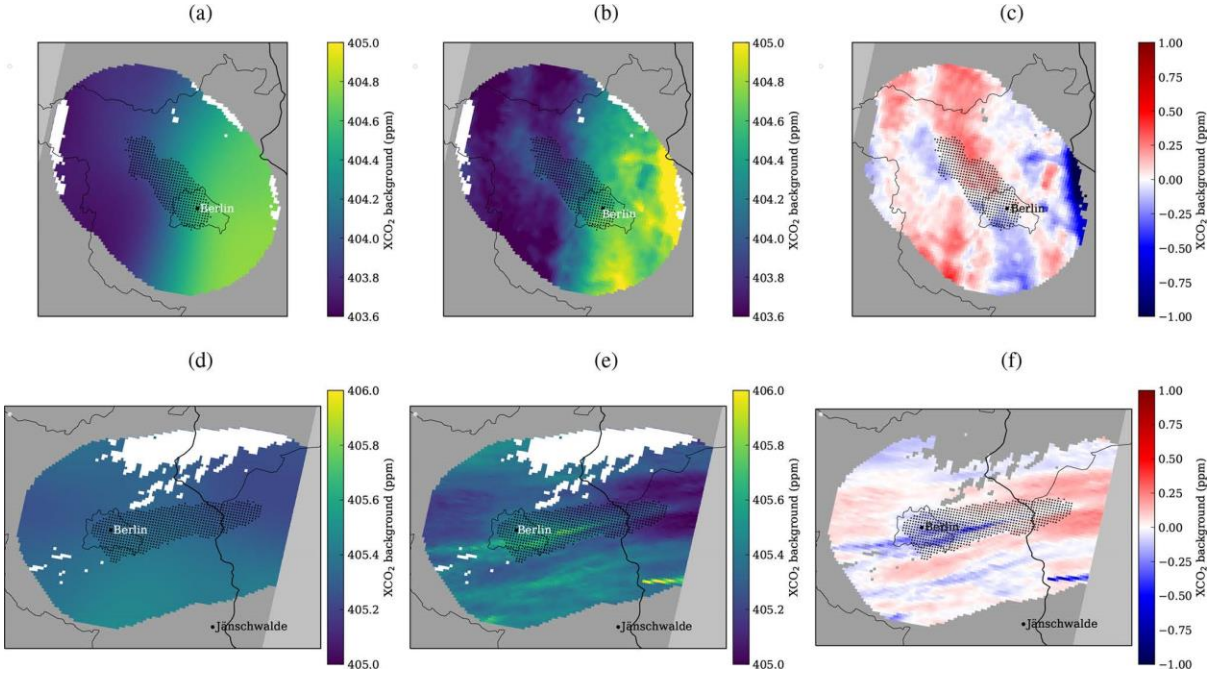

Figure 18: This figure shows two examples of estimated backgrounds using a median filter (a, d) and true modelled backgrounds (b, e), and the difference between the two background fields (c-f). Note how the true background fields can contain strong CO_2 gradients, or plume-like features that can be smaller or thinner than plumes under consideration. Conversely, the estimated background is a smooth field. This, effectively, means that the signal enhancement over the background can be over- or underestimated.

A large median filter is one option to estimate the background, which as outlined above, is a robust statistical technique to, at the very least, get a coarse idea of the background. After all, a typical image will contain more background pixels than plume pixels. Other techniques from computer vision were tested, such as a Rolling Ball algorithm, but this method tends to overestimate the background as it is still influenced by the presence of plumes. Smoothing methods such as a low-pass filter suffer from the same problem. Hence, this is not a problem that can be fully solved in an unbiased way, and requires further research. Ideally, deviations from a smooth background are not systematic, so that errors in estimated emissions average out over many quantifications. However, there could also be systematic influences, for example emissions from a nearby source that often interfere with the plume of the targeted source. In case of a city, biospheric fluxes could also introduce systematic errors in the estimated background since vegetation density is typically different inside a city from its surroundings. This may result in biospheric CO_2 signals that are spatially correlated with the anthropogenic CO_2 plume.

4.2.3 Uncertainty in wind properties

All data-driven emission quantification methods depend heavily on having accurate data regarding the wind speed and, to a lesser degree, wind direction. For methods applied to images from single overpasses, wind direction can be estimated from the direction of the column plume, but wind speed needs to be obtained from other sources of information. Unless there are observations of winds available directly at the source (e.g. from a wind-LiDAR), wind speeds will have to be obtained from a weather analysis product such as ECMWF's ERA5 reanalysis.

In order to obtain an unbiased emission estimate, an effective wind speed needs to be determined for each plume, which corresponds to the mean transport speed of the plume. If

the satellite total column measurement is equally sensitive to CO_2 molecules at all altitudes in the atmosphere, the effective wind speed can be calculated as

$$\int_0^{z_T} UCdz \Big/ \int_0^{z_T} Cdz'$$

where *U* is the along-plume wind speed, *C* is the cross-wind integrated plume concentration (i.e. concentration [µg/m^3] of CO_2 above background), and z_T is a freely selected altitude above the top of the plume. The effective wind speed is thus the vertically averaged wind speed weighted by the vertical profile of the CO_2 concentration in the plume above background. If the vertical sensitivity of the measurements is not uniform, the product of *C* times the averaging kernel has to be used instead of *C*. The effective wind speed can change in the across-plume direction also, in which case also horizontal weighing should be used.

Unfortunately, the vertical distribution of CO_2 in the plume is usually not known. Some information on plume height may be obtained by comparing the direction of the plume with the direction of the analysed winds at different altitudes, but this is only possible if there is a strong directional shear in the winds and if the analysed winds are accurate. For point sources, a better approach is to estimate the plume altitude based on plume rise calculations. The effective emission height can be much higher than the geometric height of a stack because of the momentum and buoyancy of the flue gas. In general, plume rise depends on stack geometry (height and diameter), flue gas properties (temperature, humidity, exit velocity) and on meteorological conditions (wind speed, atmospheric stability) (Bieser et al., 2011; Brunner et al., 2019). When this information is not available, typical stack parameters according to Pregger and Friedrich (2009) may be used. In case of a well-mixed atmospheric boundary layer, the exhaust plume will become mixed rather uniformly throughout the depth of the boundary layer with increasing distance from the source. In this case, a pressure-weighted mean wind speed in the boundary layer might be adequate. For emissions from a city, this could also be a viable approach.

The uncertainty in the effective wind speed will be determined by the uncertainty in the analysed winds and by the uncertainty in the vertical distribution of CO_2 in the plume. Clearly more research is needed to quantify these uncertainties and to make optimal choices for different types of emission sources and for different weather conditions.

4.2.4 Inability to deal with the temporal component of the plume

The downstream part of the plume was not emitted at the same time as it was captured on a satellite image, but at an earlier time. In this sense, time and space are interlinked dimensions, and the plume thus contains both a spatial and temporal component. If emissions vary over time, we should thus not necessarily expect constant fluxes over the length of the plume, which is however exactly what is done in instantaneous plume interpretation methods (Section 3.1). To fix this problem, atmospheric transport models are required to link the spatial and temporal dimensions to each other. Such atmospheric transport models are however explicitly dismissed in the methods described in this Deliverable, as we restrict ourselves to lightweight methods only. Thus, this problem will not be dealt with. We note that the time-averaged methods (Section 3.2) do not suffer from this problem, as in the end only the fluxes at the location of the source will in theory average out in such a way that they produce a non-zero signal at the source, and not elsewhere. An approach based on neural networks will not necessarily suffer from this problem either, if it can learn to infer the atmospheric transport dynamics from the input data.

4.2.5 A limited ability to detect spatially structured errors in the images that correlate with surface and atmospheric conditions

Satellite XCO_2 or NO_2 images can contain errors with spatial patterns that are correlated with uncertainties in atmospheric conditions (molecular species, other scatterers, aerosols, water, pressure, temperature, cloud cover, etc.) or surface properties (albedo, most importantly). These are due to uncertainties in the assumed atmospheric and surface conditions used in the retrievals, and cause so-called systematic errors in the retrieval process (in the inversion of the radiative transfer). At present, these spatially correlated errors are neither estimated nor provided in the satellite data products. The lightweight methods presented here rely on the data present in the satellite image only and are therefore unable to detect or correct for this type of error in the data.

The meteorological and/or chemistry-transport models (meteo-CTM) used in a model-based approach can estimate some information about these surface and atmospheric conditions, and can thus partly support the identification of spatial patterns in the XCO_2 and NO_2 images which could be erroneous and which should thus probably be discarded. However, the patterns depend also on the retrieval assumptions and the radiative transfer computations and therefore the identification of the systematic errors using only meteo-CTMs remains limited. Therefore, at present this does not appear to be a critical advantage of the model-based approach compared to the lightweight methods.

5 Conclusion

This document outlined a series of promising and very different "lightweight" (without requiring atmospheric transport model runs) methods to detect and quantify plumes in satellite images that are expected to be produced by the future CO2M satellite constellation, which either work on single overpass satellite images, or on the temporal average of multiple overpasses over a single source. After a detailed description of the methods, a set of issues was discussed that are being actively addressed within the CoCO2 project to improve the performance of the methods. Notable problems that are actively worked on are improving the signal-to-noise ratio and improving estimates of the wind speed and direction, which are required for accurately computing the emission fluxes. Issues that are harder to overcome are the low number of expected plume images due to full cloud coverage, and a difficulty in estimating the background concentration fields with high spatial resolution. The limitation that these examples have mostly been based on one synthetic dataset may be overcome with the use of a library of plumes developed in Task 4.1, which will be detailed in Deliverable 4.2 due at the end of 2022. Furthermore, we remark that this document only describes the various methods, while an evaluation of the performance of the different methods will be provided in Deliverable 4.4, also due at the end of 2022.

The methods developed here are intended to be used eventually in the operational monitoring and verification system (MVS) of the Copernicus programme, to quickly provide estimates of plume and city emissions when they can be observed.

6 References

Aharon, M., Elad, M. & Bruckstein, A. (2006). K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation. *IEEE Transactions on signal processing*, 4311-4322.

Beirle, S., Boersma, K.F., Platt, U., Lawrence, M.G., & Wagner, T. (2011). Megacity emissions and lifetimes of nitrogen oxides probed from space. *Science*,1737-1739.

Beirle, S., Borger, C., Dörner, S., Eskes, H., Kumar, V., de Laat, A., & Wagner, T. (2021). Catalog of NOx emissions from point sources as derived from the divergence of the NO2 flux for TROPOMI. *Earth System Science Data*, 2995-3012.

Beirle, S., Borger, C., Dörner, S., Li, A., Hu, Z., Liu, F., Wang, Y., & Wagner,T. (2019). Pinpointing nitrogen oxide emissions from space. *Science Advances*, eaax9800.

Bieser, J., Aulinger, A., Matthias, V., Quante, M., & Van Der Gon, H. D. (2011). Vertical emission profiles for Europe based on plume rise calculations. *Environmental Pollution*, 2935-2946.

Bovensmann, H., Buchwitz, M., Burrows, J. P., Reuter, M., Krings, T., Gerilowski, K., Schneising, O., Heymann, J., Tretner, A., & Erzinger, J. (2010). A remote sensing technique for global monitoring of power plant CO 2 emissions from space and related applications. *Atmospheric Measurement Techniques*, 781-811.

Broquet, G., Bréon, F.M., Renault, E., Buchwitz, M., Reuter, M., Bovensmann, H., Chevallier, F., Wu, L., & Ciais, P. (2018). The potential of satellite spectro-imagery for monitoring CO 2 emissions from large cities. *Atmospheric Measurement Techniques*, 681-708.

Brunner, D., Kuhlmann, G., Marshall, J., Clément, V., Fuhrer, O., Broquet, G., Löscher, A., & Meijer, Y. (2019). Accounting for the vertical distribution of emissions in atmospheric CO 2 simulations. *Atmospheric Chemistry and Physics*, 4541-4559.

Crippa, M., Solazzo, E., Huang, G., Guizzardi, D., Koffi, E., Muntean, M., Schieberle, C., Friedrich, R., & Janssens-Maenhout, G. (2020). High resolution temporal profiles in the Emissions Database for Global Atmospheric Research. *Scientific data*, 1-17.

Dabov, K., Foi, A., Katkovnik, V., & Egiazarian, K. (2007). Image denoising by sparse 3-D transform-domain collaborative filtering. *IEEE Transactions on image processing*, 2080-2095.

Earth and Mission Science Division (2020). Copernicus CO_2 Monitoring Mission Requirements Document, v3. Accessed from https://esamultimedia.esa.int/docs/EarthObservation/CO2M_MRD_v3.0_20201001_Issued.pdf at November 23, 2021.

Finch, D., Palmer, P., & Zhang, T. (2021). Automated detection of atmospheric NO2 plumes from satellite data: a tool to help infer anthropogenic combustion emissions. *Atmospheric Measurement Techniques Discussions*, 1-21.

Fioletov, V.E., McLinden, C.A., Krotkov, N., & Li, C. (2015). Lifetimes and emissions of SO2 from point sources estimated from OMI. *Geophysical Research Letters*, 1969–1976.

de Foy, B., Wilkins, J.L., Lu, Z., Streets, D.G., & Duncan, B.N. (2014). Model evaluation of methods for estimating surface emissions and chemical lifetimes from satellite data. *Atmospheric Environment*, 66-77.

Hakkarainen, J., Ialongo, I., Maksyutov, S., & Crisp, D. (2019). Analysis of FourYears of Global XCO2 Anomalies as Seen by Orbiting Carbon Observatory-2. *Remote Sensing*, 850.

Hakkarainen, J., Ialongo, I., & Tamminen, J. (2016). Direct space-based observations of anthropogenic CO2 emission areas from OCO-2. *Geophysical Research Letters*, 11400-11406.

Hakkarainen, J., Szeląg, M.E., Ialongo, I., Retscher, C., Oda, T., & 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.

Hersbach, H., & Dee, D. (2017). ERA5 reanalysis is in production. *ECMWF Newsletter*, 147, 7.

Hoffmann, L., Günther, G., Li, D., Stein, O., Wu, X., Griessbach, S., Heng, Y., Konopka, P., Müller, R., Vogel, B., & Wright, J. S. (2019). From ERA-Interim to ERA5: the considerable impact of ECMWF's next-generation reanalysis on Lagrangian transport simulations. *Atmospheric Chemistry and Physics*, 19, 3097-3124.

Hill, T., & Nassar, R. (2019). Pixel size and revisit rate requirements for monitoring power plant CO2 emissions from space. *Remote Sensing*, 1608.

Ialongo, I., Stepanova, N., Hakkarainen, J., Virta, H., & Gritsenko, D. (2021). Satellite-based estimates of nitrogen oxide and methane emissions from gas flaring and oil production activities in Sakha Republic, Russia. *Atmospheric Environment: X 11*, 100114.

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

Kuhlmann, G., Clément, V., Marshall, J., Fuhrer, O., Broquet, G., Schnadt-Poberaj, C., Löscher, A. Meijer, Y., & Brunner, D. (2019). SMARTCARB – Use of Satellite Measurements of Auxiliary Reactive Trace Gases for Fossil Fuel Carbon Dioxide Emission Estimation. Final report of ESA study contract n°4000119599/16/NL/FF/mg.

Kuhlmann, G., Broquet, G., Marshall, J., Clément, V., Löscher, A., Meijer, Y., & Brunner, D. (2019). Detectability of CO2 emission plumes of cities and power plants with the Copernicus Anthropogenic CO2 Monitoring (CO2M) missions. *Atmospheric Measurement Techniques*, 6695-6719.

Kuhlmann, G., Brunner, D., Broquet, G., & Meijer, Y. (2020). Quantifying CO2 emissions of a city with the Copernicus Anthropogenic CO2 Monitoring satellite mission. *Atmospheric Measurement Techniques*, 6733–6754.

Kuhlmann, G., Henne, S., Meijer, Y., & Brunner, D. (2021). Quantifying CO2 Emissions of Power Plants With CO2 and NO2 Imaging Satellites. *Frontiers in Remote Sensing*, 1-18.

Lary, D. J., Alavi, A. H., Gandomi, A. H., & Walker. A. L. (2016). Machine learning in geosciences and remote sensing. *Geoscience Frontiers*, 3-10.

Lei, R., Feng, S., Danjou, A., Broquet, G., Wu, D., Lin, J., O'Dell, C., & Lauvaux, T. (2021). Fossil fuel CO2 emissions over metropolitan areas from space: A multi-model analysis of OCO-2 data over Lahore, Pakistan. *Remote Sensing of Environment*, 112625.

Maxwell, A. E., Warner, T. A., & Fang, F. (2018). Implementation of machine-learning classification in remote sensing: An applied review. *International Journal of Remote Sensing*, 2784-2817.

Mommert, M., Sigel, M., Neuhausler, M., Scheibenreif, L., & Borth, D. (2020). Characterization of Industrial Smoke Plumes from Remote Sensing Data. *arXiv preprint* arXiv:2011.11344.

Nassar, R., Hill, T. G., McLinden, C. A., Wunch, D., Jones, D. B., & Crisp, D. (2017). Quantifying CO2 emissions from individual power plants from space. *Geophysical Research Letters*, 10045-10053.

Pinty B., G. Janssens-Maenhout, M. Dowell, H. Zunker, T. Brunhes, P. Ciais, D. Dee, H. Denier van der Gon, H. Dolman, M. Drinkwater, R. Engelen, M. Heimann, K. Holmlund, R. Husband, A. Kentarchos, Y. Meijer, P. Palmer and M. Scholze (2017). An Operational Anthropogenic CO_2 Emissions Monitoring & Verification Support capacity - Baseline Requirements, Model Components and Functional Architecture. European Commission Joint Research Centre, EUR 28736 EN.

Pregger, T., & Friedrich, R. (2009). Effective pollutant emission heights for atmospheric transport modelling based on real-world information. *Environmental Pollution*, 552-560.

Reuter, M., Buchwitz, M., Schneising, O., Krautwurst, S., O'Dell, C.W., Richter, A., Bovensmann, H., & Burrows, J.P. (2019). Towards monitoring localized CO2 emissions from space: co-located regional CO2 and NO2 enhancements observed by the OCO-2 and S5P satellites. *Atmospheric Chemistry and Physics*, 9371–9383.

Ronneberger, O., Fischer, P., & Brox, T. (2015). U-Net: Convolutional networks for biomedical image segmentation. *International Conference on Medical image computing and computer-assisted intervention*, 234-241.

Schwandner, F.M., Gunson, M.R., Miller, C.E., Carn, S.A., Eldering, A., Krings, T., Verhulst, K.R., Schimel, D.S., Nguyen, H.M., Crisp, D., & O'Dell, C.W. (2017). Spaceborne detection of localized carbon dioxide sources. *Science*, 358(6360).

Varon, D.J., Jacob, D.J., McKeever, J., Jervis, D., Durak, B.O.A., Xia, Y., & Huang, Y. (2018). Quantifying methane point sources from fine-scale satellite observations of atmospheric methane plumes. *Atmospheric Measurement Techniques*, 5673-5686.

Wang, Y., Broquet, G., Bréon, F.M., Lespinas, F., Buchwitz, M., Reuter, M., Meijer, Y., Loescher, A., Janssens-Maenhout, G., Zheng, B., & Ciais, P. (2020). PMIF v1. 0: assessing the potential of satellite observations to constrain CO 2 emissions from large cities and point sources over the globe using synthetic data. *Geoscientific Model Development*, 5813-5831.

Wu, D., Lin, J., Fasoli, B., Oda, T., Ye, X., Lauvaux, T., Yang, E., & Kort, E. (2018). A Lagrangian approach towards extracting signals of urban CO2 emissions from satellite observations of atmospheric column CO2 (XCO2): X-Stochastic Time-Inverted Lagrangian Transport model ("X-STILT v1"). *Geoscientific Model Development*, 4843-4871.

Ye, X., Lauvaux, T., Kort, E., Oda, T., Feng, S., Lin, J., Yang, E., & Wu, D. (2020). Constraining Fossil Fuel CO2 Emissions From Urban Area Using OCO-2 Observations of Total Column CO2. *Journal of Geophysical Research: Atmospheres*, 1-29.

Zeng, Y., Lin, Z., Lu, H., & Patel, V. M. (2021). CR-Fill: Generative Image Inpainting With Auxiliary Contextual Reconstruction. *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 14164-14173.

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(14), 8501-8510.

Document History

Version	Author(s)	Date	Changes
0.1	Erik Koene, Dominik Brunner & Gerrit Kuhlmann (Empa), Janne Hakkkarainen (FMI), Joffrey Dumont Le Brazidec (ENPC), Grégoire Broquet (CEA)	30/11/2021	Initial version
1.0	""	15/12/2021	Final version based on reviews.

Internal Review History

Internal Reviewers	Date	Comments
Julia Marshall (DLR)	14/12/2021	Overall quite good, see comments and requests for clarification/context in the comments throughout.
Paul Palmer (UEDIN)	01/12/2021	-

Estimated Effort Contribution per Partner

Partner	Effort
Empa	3
FMI	2
ENPC	1
CEA	1
Total	**7**