CO2MVS RESEARCH ON SUPPLEMENTARY OBSERVATIONS

D2.2 Time series of NO_x and CO emissions of hot spots in Africa, Europe and SE Asia in reference year

Due date of deliverable	31 July 2025
Submission date	13 August 2025
File Name	CORSO-D2-2-V1.1
Work Package / Task	WP-2 / Task 2.1
Organisation Responsible of Deliverable	EMPA
Author name(s)	Gerrit Kuhlmann, Erik Koene, Sandro Meier, Gijs Leguijt, Arjo Segers, Hugo Denier van der Gon, Chloe Schooling, Douglas Finch, Paul Palmer
Revision number	1.1
Status	final
Dissemination Level / location	Public

Funded by the European Union

The CORSO project (grant agreement No 101082194) is funded by the European Union.
Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the Commission. Neither the European Union nor the granting authority can be held responsible for them.

1 Executive Summary

This work was carried out within Work Package 2 (WP2), which aims to improve the estimation of anthropogenic emissions by leveraging co-emitted species in support of the future CO_2 Monitoring and Verification Support (CO2MVS) system. Nitrogen dioxide (NO2) and carbon monoxide (CO) observations are available with higher spatial and temporal coverage than CO_2 observations from instruments like TROPOMI on Sentinel-5P. NO_2 is also observed by satellites with higher precision than carbon dioxide (CO_2), making it a valuable tracer for anthropogenic emissions.

This report supports WP2 objectives by detailing the methodology and results of using TROPOMI satellite observations of NO_2 and CO columns to quantify emissions from urban areas, power generation facilities, and iron and steel plants. Emission plumes of NO_2 and CO are identified in individual satellite overpasses using plume detection algorithms. Subsequently, NO_x and CO emissions are estimated through data-driven mass balance approaches and analytical inversion methods.

The report is accompanied by several NetCDF files containing 2021 time series of estimated NO_x and CO emissions for selected megacities, power plants, and iron and steel production sites across Africa and Europe.

Table of Contents

2 Introduction

2.1 Background

To enable the European Union (EU) to move towards a low-carbon economy and implement its commitments under the Paris Agreement, a binding target was set to cut emissions in the EU by at least 40% below 1990 levels by 2030. European Commission (EC) President von der Leyen committed to deepen this target to at least 55% reduction by 2030. This was further consolidated with the release of the Commission's European Green Deal on the 11th of December 2019, setting the targets for the European environment, economy, and society to reach zero net emissions of greenhouse gases in 2050, outlining all needed technological and societal transformations that are aiming at combining prosperity and sustainability. To support EU countries in achieving the targets, the EU and European Commission (EC) recognised the need for an objective way to monitor anthropogenic CO_2 emissions and their evolution over time.

Such a monitoring capacity will deliver consistent and reliable information to support informed policy- and decision-making processes, both at national and European level. To maintain independence in this domain, it is seen as critical that the EU establishes an observation-based operational anthropogenic CO_2 emissions Monitoring and Verification Support (MVS) (CO2MVS) capacity as part of its Copernicus Earth Observation programme.

The CORSO research and innovation project will build on and complement the work of previous projects such as CHE (the CO2 Human Emissions), and CoCO2 (Copernicus CO2 service) projects, both led by ECMWF. These projects have already started the ramping-up of the CO2MVS prototype systems, so it can be implemented within the Copernicus Atmosphere Monitoring Service (CAMS) with the aim to be operational by 2026. The CORSO project will further support establishing the new CO2MVS addressing specific research & development questions.

The main objectives of CORSO are to deliver further research activities and outcomes with a focus on the use of supplementary observations, i.e., of co-emitted species as well as the use of auxiliary observations to better separate fossil fuel emissions from the other sources of atmospheric CO_2. CORSO will deliver improved estimates of emission factors/ratios and their uncertainties as well as the capabilities at global and local scale to optimally use observations of co-emitted species to better estimate anthropogenic CO_2 emissions. CORSO will also provide clear recommendations to CAMS, ICOS, and WMO about the potential added-value of high-temporal resolution $^{14}CO_2$ and APO observations as tracers for anthropogenic emissions in both global and regional scale inversions and develop coupled land-atmosphere data assimilation in the global CO2MVS system constraining carbon cycle variables with satellite observations of soil moisture, LAI, SIF, and Biomass. Finally, CORSO will provide specific recommendations for the topics above for the operational implementation of the CO2MVS within the Copernicus programme.

2.2 Scope of WP-2

The work presented in this report is part of WP2 of CORSO, which deals with “Use of co-emitted species (correlations, improved emission ratios, uncertainties) in data assimilation systems". The aim of WP2 is to enhance the use of observations of co-emitted species (NO_2, CO) to better estimate anthropogenic CO_2 emissions in the future CO2MVS capacity. This is based on the recognition that anthropogenic CO_2 emissions cannot completely be constrained with CO_2 concentration observations alone, and the detectability of the anthropogenic signal of co-emitted species is often much better than that of CO_2. For the emission estimation development at local scale, this WP focuses on the development of methods to increase the accuracy of annual CO_2 emission estimates of hot spots, industrial and urban areas by integrating satellite observations of co-emitted species (NO_2 and CO) in data assimilation

systems. Since CO_2 satellite observations are temporally sparse (even with the future CO2M constellation), temporal sampling biases are a significant source of uncertainty in annual CO_2 emission estimates of hot spots. Co-emitted species such as CO and NO_2 are and will be available at sub-diurnal temporal coverage from current and future LEO and GEO satellites. They can therefore be used to improve the constraint on the temporal variability of CO_2 emissions and hence for reducing the uncertainty in annual estimates. The local and regional studies will focus on three regions: Europe, Africa, and Southeast Asia.

2.2.1 Objectives of this deliverable

The objectives of this deliverable are to provide estimates of CO and NO_x emissions of hot spots for Africa, Europe and Southeast Asia using Sentinel-5P/TROPOMI NO_2 and CO observations and Geostationary Environment Monitoring Spectrometer (GEMS) NO_2 observations. Since GEMS currently does not provide NO_2 observations for 2021, the current version of the deliverable only uses TROPOMI observations. The dataset will be used in WP1 (Task 1.3) for validating the bottom-up inventories.

The title of this deliverable is “Time series of NO_x and CO emissions of hot spots in Africa, Europe and SE Asia in reference year”. The deliverable contains this report, which describes the methodology and several dataset files, which is given separately as a zip file, containing the time series of NO_x and CO emissions for different hot spots.

2.2.2 Work performed in this deliverable

This deliverable was accomplished through a series of activities detailed in Section 3:

- Gathering of NO_2 and CO observations from the TROPOMI instrument.
- Gathering of wind fields from the ERA-5.
- Development and improvement of plume detection and quantification methods for determining CO and NO_x emissions of hot spots.
- Application of the methods to selected megacites, power plants and iron/steel plants to generate a time series of CO and NO_x emissions for 2021.
- Compiling the files containing time series of emissions.

2.2.3 Deviations and counter measures

Access to GEMS NO_2 data has been delayed and thus only TROPOMI NO_2 data were used in this version.

2.3 Project partners

Partners	
EUROPEAN CENTRE FOR MEDIUM-RANGE WEATHER FORECASTS	ECMWF
AKADEMIA GORNICZO-HUTNICZA IM. STANISLAWA STASZICA W KRAKOWIE	AGH
BARCELONA SUPERCOMPUTING CENTER - CENTRO NACIONAL DE SUPERCOMPUTACION	BSC
COMMISSARIAT A L ENERGIE ATOMIQUE ET AUX ENERGIES ALTERNATIVES	CEA
KAMINSKI THOMAS HERBERT	iLab
METEO-FRANCE	MF
NEDERLANDSE ORGANISATIE VOOR TOEGEPAST NATUURWETENSCHAPPELIJK ONDERZOEK TNO	TNO
RIJKSUNIVERSITEIT GRONINGEN	RUG
RUPRECHT-KARLS-UNIVERSITAET HEIDELBERG	UHEI

LUNDS UNIVERSITET	ULUND
UNIVERSITE PAUL SABATIER TOULOUSE III	UT3-CNRS
WAGENINGEN UNIVERSITY	WU
EIDGENOSSISCHE MATERIALPRUFUNGS- UND FORSCHUNGSANSTALT	EMPA
EIDGENOESSISCHE TECHNISCHE HOCHSCHULE ZUERICH	ETHZ
UNIVERSITY OF BRISTOL	UNIVBRIS
THE UNIVERSITY OF EDINBURGH	UEDIN

3 Data and methods

This deliverable describes the input data and methods used for estimating time series of CO and NO_x emissions from hot spots using Sentinel-5P/TROPOMI CO and NO_2 observations for year 2021. The dataset includes hot spots in Africa, Europe and Southeast Asia. The data product will be used in Task 1.3 for evaluating the bottom-up emission estimates prepared for the same year.

3.1 Plume detection using machine-learning algorithm

3.1.1 Image segmentation

This work builds on a previous NO_2 plume detection model (Finch et al., 2022), which used a convolutional neural network (CNN) to classify images of TROPOMI tropospheric NO_2 column data as to whether or not they contained an emission plume. This model had limitations as there was no information on the location or size of the plume, or whether the image contained multiple plumes. The model was only successful if a plume was contained within a pre-defined image size; therefore, any plumes larger than the image size, or crossing between two or more images introduced inaccuracies in the detection. To address these issues, a segmentation model was developed. This generates a mask indicating the probable location of the plume and includes a method for merging plumes that span multiple images

A U-Net style model is used, consisting of a series of down-sampling blocks, each containing a double convolutional layer, a max pooling layer and a dropout layer set at 20%. This is followed by a double convolutional layer, which determines the important features of the input image. Subsequently, up-sampling blocks reconstruct the image to its original dimensions, producing a mask of the predicted plume. A schematic of the model is shown in Figure 1.

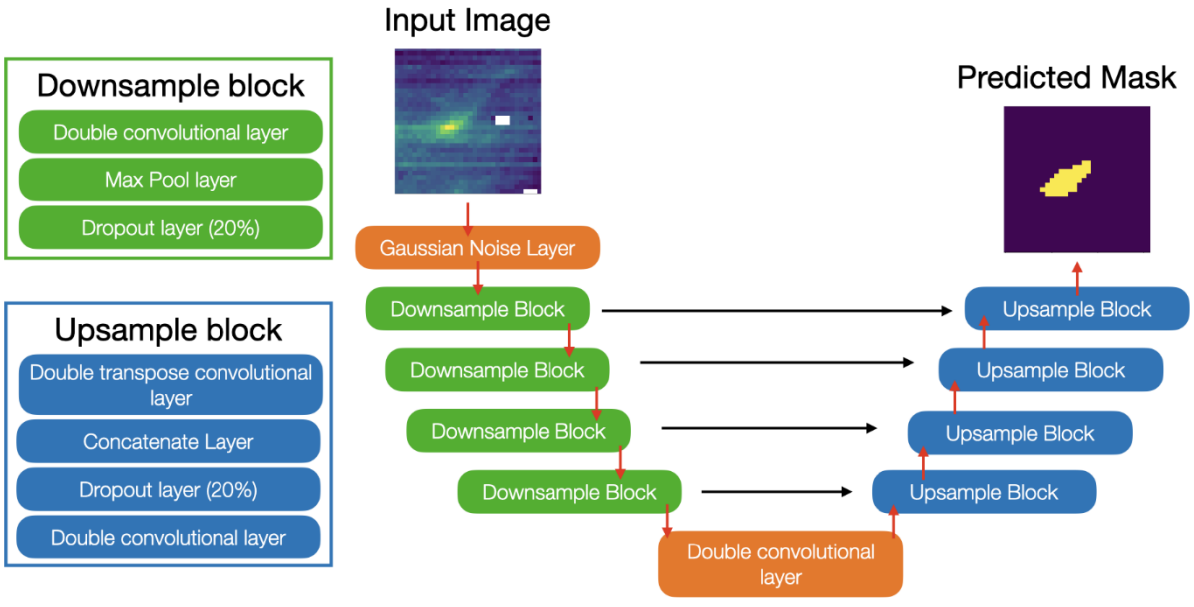

Figure 1: Schematic of the U-Net architecture used to develop the plume detection models. An input image (top left) goes through four downsampling blocks (green) made from convolutional, max pooling and dropout layers. The image is then put through a double convolutional layer (orange) to detect patterns, and a mask (top right) is built by passing the image through four upsample blocks (blue) consisting of a transpose convolutional, concatenation, dropout and a convolutional layer.

The model input is a normalised 32 x 32-pixel section of the TROPOMI swath with an output of the same size. To manage plumes that are larger than the image or straddling multiple images, each TROPOMI swath is split into overlapping images created from a rolling window

of four pixels in both the along-swath and across-swath directions. This process generates approximately 100,000 images per swath, which are fed into the model. Each image is then passed through the plume detection model and the swath is rebuilt from the predicted output based on the median of the mask prediction. The final product is an array the same size as the original swath containing the predicted masks. Finally, OpenCV is used to detect each individual plume mask in the swath regardless of the size or shape.

To train the plume detection model, plume masks were drawn for 702 images, where approximately 30% of these had multiple plumes within them. The dataset was augmented by flipping and rotating the images creating a final training dataset of 4,914 images. This was then randomly split into 80% training and 20% testing data. The plumes chosen to train the model were taken from random swaths from anytime during 2019 from across the globe. This reduces the chance of selecting neighbouring images and therefore reduces the chances of auto-correlation between images in space or time.

To get an indication of whether the detected plume is likely to be from a biomass burning source, the plumes are compared to the locations of fire detections in VIIRS and MODIS data. A plume is marked with a biomass burning flag if there are fire detections from either VIIRS or MODIS within the plume boundary on the same date as the plume.

3.1.2 Emission quantification

To estimate an emission Q rate associated with a plume we use the integrated mass enhancement (IME) method following equation:

$$Q = \frac{\Delta M \times U}{L}$$

where ΔM is the integrated mass enhancement of the plume compared to the background (in kg), U is the wind speed (in m s^{-1}) and L is the length of the plume (in metres). To calculate the mass enhancement of the plume in relation to the background, first the boundary of the plume is determined by finding the outline of the predicted plume mask. Figure 2 shows a typical plume found in TROPOMI data with associated plume boundary.

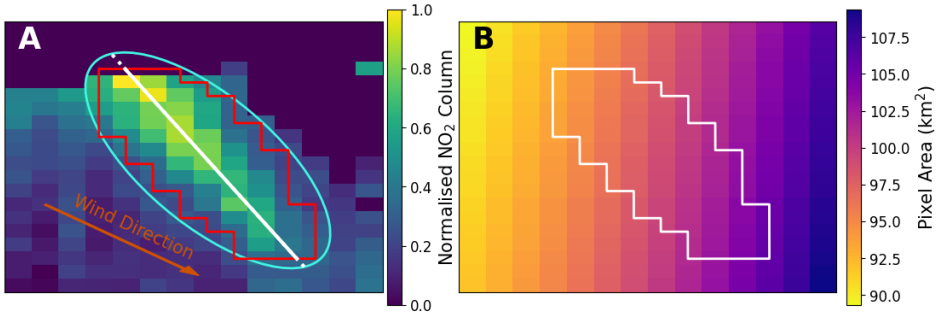

Figure 2: Normalized plume from the TROPOMI data, with predicted mask boundary (red), fitted ellipse (light blue), and plume primary axis (white) on the left, and pixel area for the same image with the plume boundary (white) shown for context.

Using the pixel area, the mass of gas within each pixel is calculated from the TROPOMI column observations (molecules per m^2). The mass of the plume within the boundary can be calculated, and the background mass is taken as the median mass of all pixels within the image but outside the plume boundary. The mass enhancement is the difference between these two values. The length of the plume is determined by fitting an ellipse to the plume using the OpenCV package in python.

The wind speed is taken as the median wind speed within the plume boundary. The wind speed used is the ECMWF 10 m U and V wind vectors. The angle between wind speed and the direction of the plume is calculated and can be used for filtering of plumes, since, if the direction of a plume does not align with the wind direction, then it may not be a genuine plume.

This filtering was not applied for the results shown here as more refining and quality checking is needed to be confident in the results.

3.2 CO emissions of hot spots from TROPOMI CO observations

Two types of CO hotspots have been investigated. Emissions of urban areas are estimated through the cross-section flux (CSF) method. Iron & steel plants, which tend to have lower emission rates than the urban areas, have their emissions quantified using an analytical inversion.

3.2.1 Cross-sectional flux analysis for African mega-cities

Here we give a short overview of the CSF method as applied to mega-cities in Africa, a detailed description of the algorithm is given in Leguijt et al. (2023). Using three years of quality-filtered operational TROPOMI CO data (Landgraf et al., 2018), the emission rates of 29 mega-cities in Africa are estimated. Each overpass, pixels with enhanced CO concentrations downwind of the city are selected as part of the plume. Despite using multiple wind products (10 meter altitude wind fields from NASA/GMAO GEOS-FP reanalysis data, planetary boundary layer averaged GEOS-FP reanalysis data and the 10 meter altitude wind fields from the ECMWF ERA5 product – (Molod et al., 2012; Hersbach et al., 2020) wind field information is a major contributor of uncertainty to the final emission estimate. Therefore, the wind direction is redetermined based on the plume direction, under the assumption that the enhanced concentrations move with the wind.

Following the determination of plume pixels, a 2D spline is fit through the plume, perpendicular to which cross-sections are drawn. In the absence of a clear plume, a rectangular box is used to draw down-wind cross-sections. Integration of enhancements along the cross-sections and multiplication with the local wind-speed yields daily emission estimates. Enhancements are calculated by subtracting the mean concentration in an up-wind box. The choice for an upwind background estimation allows application to a large number of sources, despite large variation in observed plume shapes, plume widths and plume sizes.

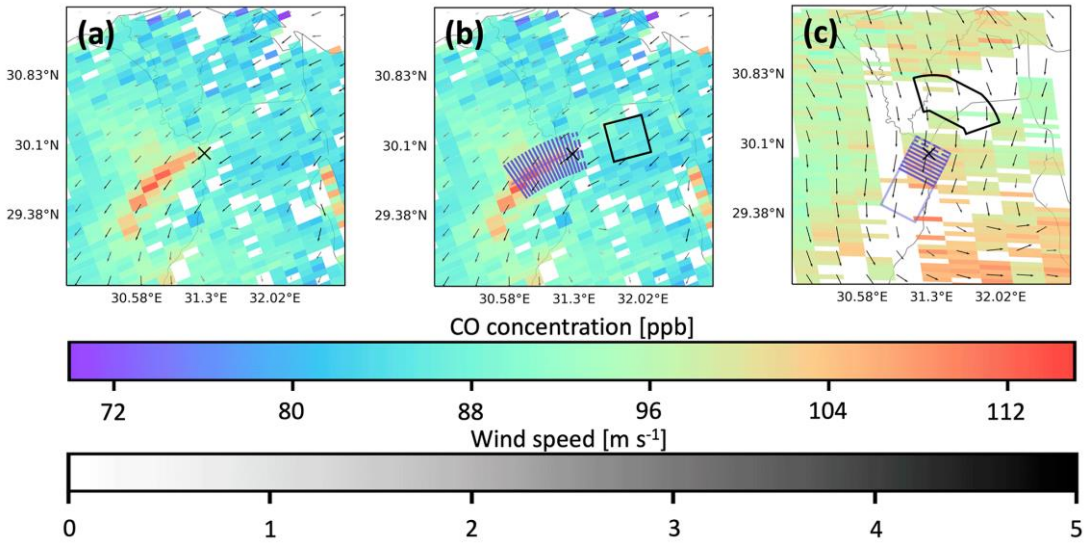

Figure 3: Example of the CSF method applied to estimate CO emissions from Cairo on two individual days from Leguijt et al. (2023): (a) TROPOMI CO image on 7 April 2019 The arrows show 10-meter altitude winds from the GEOS-FP reanalysis data, with wind speed defined by length as well as gray shading . (b) Fitting a 2D spline through enhanced pixels, the shape of the plume is determined. The spline, and perpendicular cross-sections are shown in blue. The background is determined using the region indicated with the black box. (c) CO concentrations over Cairo measured by TROPOMI on the 27 March 2020. Due to the absence of a clearly observable plume, a spline fit is not possible, and a rectangular box is used instead.

The final CSF algorithm was tested on plumes simulated with the Weather Research and Forecasting (WRF) model as explained in Leguijt et al. (2023).

3.2.2 Analytical inversion over Europe's iron & steel industry

Integrated iron & steel plants, which are plants covering the entire conversion from iron ore to steel, are the largest CO sources in Europe. However, although few plants report emissions above 100 Gg per year, the majority of the plants has emission rates below what can be quantified using the CSF as implemented in Section 3.2.1. To accurately quantify the emission rates of the individual plants, an analytical inversion was used, as described in (Leguijt et al., 2025). Here we will give a short summary of the implemented method.

Members of the European Union are required to report facility-level emission rates to the European Pollutant Release and Transfer Register (E-PRTR, 2023), this framework was used as prior emission estimates. Using WRF as forward model, 3D concentration fields were simulated for the 21 largest iron & steel plants of Europe. The prior emissions were supplemented with the TNO GHGco inventory version 4 (Kuenen et al., 2022) for the anthropogenic emissions and boundary conditions from the air pollutant forecast product of the Copernicus Atmosphere Monitoring Service (CAMS). The resulting 3D fields are converted into total columns using the TROPOMI averaging kernel to allow proper comparison to the columns measured by TROPOMI.

Optimizing the cost function $J(x)$, given by

$$J(x) = (x - x_A)^T S_A^{-1}(x - x_A) + \gamma(y - Kx)^T S_O^{-1}(y - Kx).$$

Here, x_A represents the prior emissions, S_A the prior error covariance matrix, γ the regularization parameter, y the observed concentrations, K the Jacobian and S_O the observational error covariance matrix. The two terms comprising $J(x)$ represent departure from the prior and difference between simulation and observation.

As $y - Kx$, is evaluated per pixel, spatial mismatches between simulation and observation will result in underestimation of the emission rate. This effect was remedied by aggregating the observations. In addition, an ensemble of plumes was simulated by running the WRF model with different boundary layer schemes. The optimal simulation was determined on a daily basis.

To verify the inversion emission estimates with prior-independent methods, a CSF, as discussed in Section 3.2.1, was applied to the 7 largest plants, which all report emission rates above the 100 Gg threshold. In addition, the simulations were extended to the year 2020 for 4 plants. Differences in emissions between 2019 and 2020 for these plants were compared to trends in oversampled wind-rotated data as in Clarisse et al. (2019) .

3.3 NO_x emissions of hot spots from TROPOMI NO_2 observations

NO_x emissions of hot spots are estimated using the divergence (DIV) method and the cross-sectional flux (CSF) method implemented in the Python library for data-driven emission quantification (ddeq; Kuhlmann et al., 2024). ddeq is an open-source library that was originally developed for the ESA SMARTCARB project for estimating CO_2 emissions from synthetic CO2M data (Kuhlmann et al., 2019; Kuhlmann et al., 2021). In the CoCO2 project, the library was extended with additional methods and used for benchmarking different methods for emission quantification of hot spots (Hakkarainen et al., 2024; Santaren et al., 2025).

3.3.1 Input data

The Sentinel-5P/TROPOMI NO_2 product (Version 3.6.2) for 2021 was downloaded from the Copernicus Dataspace (https://dataspace.copernicus.eu/). For wind data, the global ERA-5 reanalysis was obtained on pressure levels (Hersbach et al., 2020). The effective wind speed

for each source was computed using the standard emission profile for power plants (GNFR-A profile, Brunner et al. (2019)).

3.3.2 Divergence method

Emission estimates using the divergence method (Beirle et al., 2019; Koene et al., 2024) were obtained by first computing a global flux map $\overline{F}$ at a common resolution of 0.03 degrees ($F = \nabla \cdot \left(\vec{U}\,\mathrm{VCD}\right) + \frac{\mathrm{VCD}}{\tau}$, where $\vec{U}$ is the GNFR-A weighed wind, VCD is the tropospheric NO_2 column, and τ was set to 4 hours representing the expected lifetime; $\overline{F}$ is the yearly-averaged version of this dataset). The divergence operation was computed on the native TROPOMI grid before conservatively remapping the result onto a common grid.

Air mass factors (AMF) are used to convert slant column densities (SCD) obtained from the spectral fitting to vertical column densities (VCD):

$$VCD = \frac{SCD}{AMF} \quad \text{with} \quad AMF = \frac{\sum_l VCD_l \cdot AMF_l}{\sum_l VCD_l}$$

where AMF_l are the 1D-layer AMFs and VCD_l are the layer NO_2 VCDs taken from low-resolution global TM5 simulations. Following the TROPOMI user manual, we recalculated the VCD using an updated expected profile matched with the GNFR-A profile, allowing us to compute

$$\mathrm{VCD}_{new} = \frac{\sum_i (x_i)}{\frac{M}{M_T}\sum_i (A_T x_i)}\mathrm{VCD}_{TROPOMI},$$

where x_i is the new profile defined on the same pressure grid as the averaging kernel (AK) A_T, M is the *total* air mass factor and M_T is the *tropospheric* air mass factor, and $\mathrm{VCD}_{TROPOMI}$ is the TROPOMI tropospheric NO_2 VCD. The ratio $\sum_i (x_i) / \frac{M}{M_T}\sum_i (A_T x_i)$ is referred to as the AMF correction factor, which is used to correct the estimated emissions at a source (Beirle et al., 2023)

After computing the flux map $\overline{F}$ we estimated emissions for individual sources by integrating the mass within geodesic circles centred on each source. The circle radii ranged from 1 to 60 km in 1 km increments. For each radius, the integrated NO_2 mass is multiplied with an estimate of the NO_2-to-NO_x ratio of 1.65 + $0.5 \times VCD/VCD_{max}$ (i.e., the conversion factor increases when there is a lot of NO_2), which was chosen to provide reasonable factors for power plants where conversion factors are available from MicroHH simulations (Belchatow, Jänschwalde and Matimba/Medupi) (Meier et al., 2024; Krol et al., 2024). Next, the local 5x5 mean of the AMF correction factor is applied to scale the emissions. An example of this procedure is seen in Figure 4. To determine the optimal integration radius (i.e., the point where additional area no longer meaningfully corresponds to the selected source), we analysed the shape of the cumulative curve. The selected radius was the smallest radius at which (A) the first derivative of the cumulative curve became negative (i.e., the curve peaked), or (B) the second derivative crossed from negative to positive (i.e., the curve began to bend upward again, indicating another source was integrated too). This curvature-based method proved effective in isolating the emissions from a single source while avoiding contributions from nearby sources.

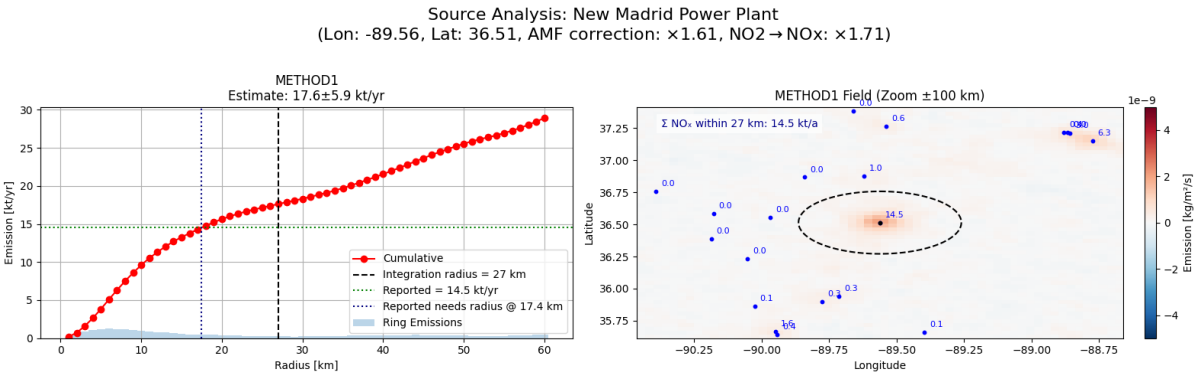

Figure 4: Application of the divergence method to estimate NO_x emissions from the New Madrid Power Plant, USA. *Left*: Cumulative integration of NO_2 mass as a function of increasing radius from the source. *Right*: Computed flux map with selected integration radius shown by the circle. The total estimated emissions from all bottom-up sources (shown as a blue scatter plot) within the selected radius are also indicated. See main text for details.

Finally, after estimating annual emissions, the same approach was applied on a monthly scale – by recomputing $\overline{F}$ on a monthly scale, but while re-using the AMF correction factor and selected integration radius from the annual data. Only the NO_2-to-NO_x conversion factor varies for one month of data at a time.

3.3.2.1 Uncertainty

The uncertainty of the estimates is computed using the principle of quadratic error propagation (also known as root-sum-squares error propagation). It is calculated using the heuristic formula for $\Delta Q_{1\sigma}$ which we define as an estimate of the '1σ' uncertainty

$$\Delta Q_{1\sigma} = Q \times \sqrt{[0.1(1-A)]^2 + \left(0.1\frac{365}{\mathrm{d}}\right)^2 + \left(0.1\frac{\sqrt{\overline{U}^2+\overline{V}^2}}{d}\right)^2 + 0.2^2}$$

where Q is the emission estimate (i.e., the uncertainty is proportional to the size of the estimate), A is the air mass factor correction (i.e., this is a small error term scoring the fact that we modified our data with a number that has its own uncertainty), d is the number of valid overpasses for which the quality factor was 0.75 or larger (i.e., if we have daily good quality overpasses, we would have $d = 365$ and the error term would correspondingly be small; if we have fewer overpasses the error term increases proportionally), $\sqrt{\overline{U}^2+\overline{V}^2}$ is the yearly mean effective wind speed in meters per second (i.e., if there is a *net* yearly wind in a particular direction, the divergence map exhibits certain artifacts as we have observed in the data). The final term adds a final fixed expected error. Thus, in the ideal case where the AMF correction is not large ($A \approx 1.7$), there is a lot of good data ($d \approx 365$), and we have no net yearly wind in a particular direction, we end up with

$$\Delta Q_{1\sigma,\text{best case}} \approx Q \times \sqrt{[0.1 \cdot -0.7]^2 + 0.1^2 + 0.2^2} \approx Q \times 0.23.$$

Hence, even in the ideal case expect an around 23% error on the estimate. As the data quality worsens (say, $A \approx 2$, $d = 50$, $\sqrt{\overline{U}^2+\overline{V}^2} \approx 10$), we end up with

$$\Delta Q_{1\sigma,\text{worse case}} \approx Q \times \sqrt{[0.1 \cdot -1]^2 + (0.1 \cdot 7.3)^2 + (0.1 \cdot 0.5)^2} \approx Q \times 0.77.$$

Hence, in this 'bad' case, the error goes up to 77%, but as the number of overpasses decreases this number can in principle grow to over 100%.

The uncertainty model is itself highly uncertain as typically we assume some value for an error (e.g., 10% for the AMF correction as also done by Beirle et al., 2019, 2023), and formulating a more representative or precise uncertainty model is a topic of active research. Many more terms could in principle be added, e.g., those related to the NO_2-to-NO_x ratio, the ERA-5 wind speed, and the lifetime. Some more complicated terms which should, in principle, also be added are: (1) if there is a sampling bias towards summer months where emissions may be lower our estimate may be too low, (2) the GNFR-A profile used to weight the winds and the AMF profile is only a rough guess for the plume height but in reality this varies along with the meteorological conditions, (3) the divergence method assumes steady-state conditions and horizontally homogeneous conditions, which is not a valid assumption when meteorological conditions are (often) changing around overpass time, (4) we ascribe all the emissions within the integration radius to the (known) source at its center, but if other sources are nearby (urban, industrial, biomass burning, …) the estimate naturally doesn't apply to merely the point source of interest and we likely overestimate the emissions.

3.3.3 Cross sectional flux method

3.3.3.1 Identification of the plume area

As a first step, the location of the emission plume is determined using the ERA-5 wind direction. Identifying the plume area using the wind vector is more robust than a thresholding approach, resulting in less misdetections especially for weak plumes or for emission clusters with nearby plume (Santaren et al., 2025). The disadvantage of the approach is that it cannot be used for estimating the NO_x lifetime directly from the measurements.

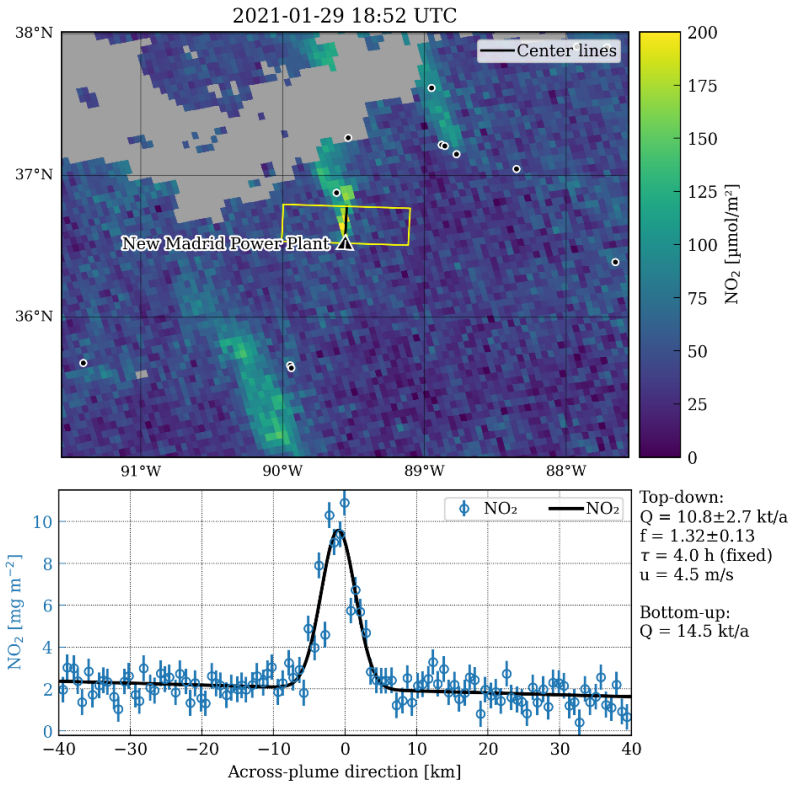

Figure 5: Example of the CSF method applied to estimate NO_x emissions from the New Madri Power Plant in the United States of America. (upper) TROPOMI NO_2 image on 29 January 2021 (18:52 UTC) with plume region marked by yellow polygon. (lower) Across-plume columns in across-plume direction with Gaussian curve fitted to estimate line densities. Emission rate (Q) is computed from line density and wind speed (u=4.5 m/s) using a NO_2-to-NO_x conversion factor (f=1.32) and a NO_x lifetime (τ) of 4 hours. In this example, the NO_x emission rate at overpass was estimated as 10.8 kt NO_2 / a, while the bottom-up inventory states an annual mean emission of 14.5 kt/a.

To define the plume area, a Bezier curve following the wind direction up to 30 km downwind is used for computing along- and across-plume coordinates. The plume area used for estimating the emissions is 60 km wide perpendicular to the wind speed.

3.3.3.2 Line densities

The NO_2 line densities are computed by fitting a Gaussian curve with a linear background to all pixels in the plume area:

$$g(y) = \frac{q}{\sqrt{2\pi}\sigma}\exp\left(-\frac{(y-\mu)^2}{2\sigma^2}\right) + my + b$$

where q is the line density (in kg/m), μ and σ are centre position and standard width of the curve (in m), and *m* and *b* are slope and intercept of a linear background.

3.3.3.3 Air mass factor

Air mass factors (AMF) are recomputed using the averaging kernels (AK) provided by the NO_2 product and the NO_2 profile from the TM5 simulations, where the NO_2 enhancement from an initial fit of the Gaussian curve was added to the profile assuming a GNFR-A emission distribution. NO_2 column densities are updated following the AMF correction and the line densities are fitted again with the updated values to obtain the AMF-corrected line density.

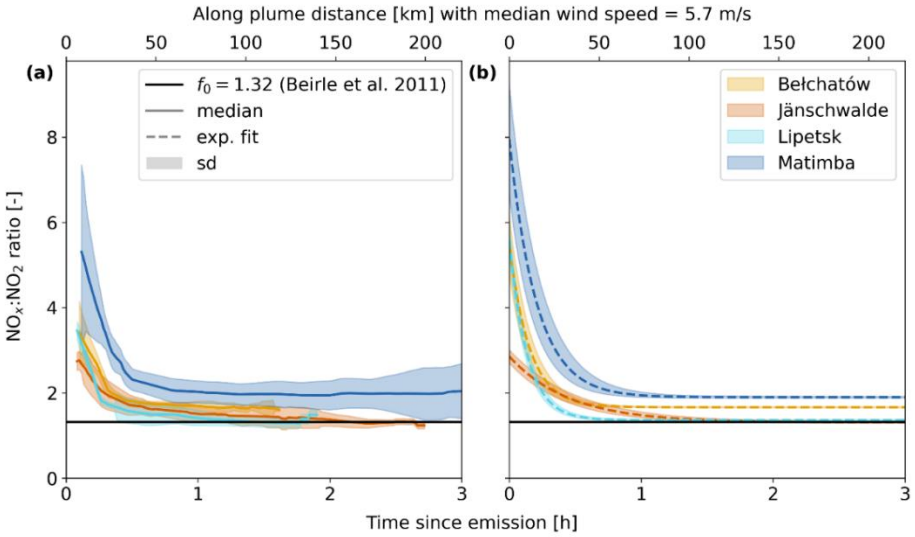

Figure 6: NO_x:NO_2 ratios as a function of time since emission: (a) Median and standard deviation in the MicroHH simulations and (b) fitted negative exponential function and corresponding standard deviation (from Meier et al. 2024).

3.3.3.4 NO_x chemistry

To estimate NO_x emissions from NO_2 observations, it is necessary to convert the NO_2 to NO_x columns using an NO_2-to-NO_x conversion model and to account for the NO_x decay since emission. We used three methods for estimating the NO_2 to NO_x conversion factor:

1. **Outside Europe**: Standard values using f = 1.32 and τ = 4 hours.
2. **Inside Europe**: Random forest model trained with GEOS-Chem simulations to predict NO_2:NO_x ratios and NO_x lifetimes from geographical and meteorological parameters. The random forest model was only available for Europe for this deliverable (as planned). It will be possible to reprocess the emission estimates as soon as a global model is available (CORSO D2.5 and Schooling et al. 2025),
3. **Selected plants**: MicroHH simulations with chemistry were used to determine NO_2 to NO_x conversion for selective power plants. The line densities are converted from NO_2

to NO_x by accounting for the NO_2 to NO_x conversion formula $f(t)$, which is computed as time since of emissions (using the effective wind speed):

$$f(t) = m \cdot \exp\left(-\frac{t}{r}\right) + f_0$$

where m, τ and f_0 are parameters (Meier et al 2024).

3.3.3.5 NO_x emissions

The emission rate Q is computed as

$$Q = \frac{f \cdot u \cdot q}{\exp\left(-\frac{x}{u\tau}\right)}$$

where *f* is the NO_2 to NO_x conversion factor, *u* is the effective wind speed, *q* is the line density, *x* is the distance from the source (i.e., 15 km), and τ is the NO_x lifetime.

3.3.3.6 Uncertainties

The uncertainties of emissions are computed by the *ddeq* library, accounting for the uncertainty of the NO_2 column densities, the wind speed (using 1 m/s) and the NO_2-to-NO_x conversion model. The uncertainty of background field and the cross-sectional flux method due to assumptions such as steady-state conditions are currently not included in the uncertainty budget.

4 Results

In this section, we show some results from the dataset described in this document. The data are attached and were used in Deliverable 1.3 for a detailed comparison between top-down and bottom-up estimates of CO and NO_x emissions.

4.1 Plume detection using machine-learning algorithm

Figure 7 shows location of anthropogenic and biomass burning plumes for 2021 over the Iberian Peninsula and Nigeria obtained from TROPOMI NO_2 observations using the plume detection algorithm. The figure indicates all locations where the outline of a detected plume contains a fire detected by either VIIRS or MODIS on the same day. To consider a detected plume as a biomass burning fire, we filter the VIIRS and MODIS datasets to only include presumed vegetation fire and to have a medium or high confidence (for VIIRS) or greater than 50% (for MODIS). This is a crude characterisation of whether the plume originates from anthropogenic combustion or biomass burning. As can be seen in the figure, this method shows that most plumes characterised as biomass burning are in locations where inventories suggest substantial biomass burning (tropics, scrubland etc). However, there are some plumes labelled as biomass burning which are located over regions which tend not to have biomass burning emissions (e.g. urban areas). Plume labelling in these regions may be inaccurate. This may reflect a false detection or labelling from the VIIRS or MODIS datasets, or where the active fire may be located anywhere in the plume boundary so that a detected plume may straddle the location of anthropogenic and biomass burning sources. Alternatively, the label for these detected plumes is accurate and there was a vegetation fire in these unlikely locations. These results have not been compared to GFAS or GFED and further analysis using these datasets could refine the plume characterisation for future iterations of this product.

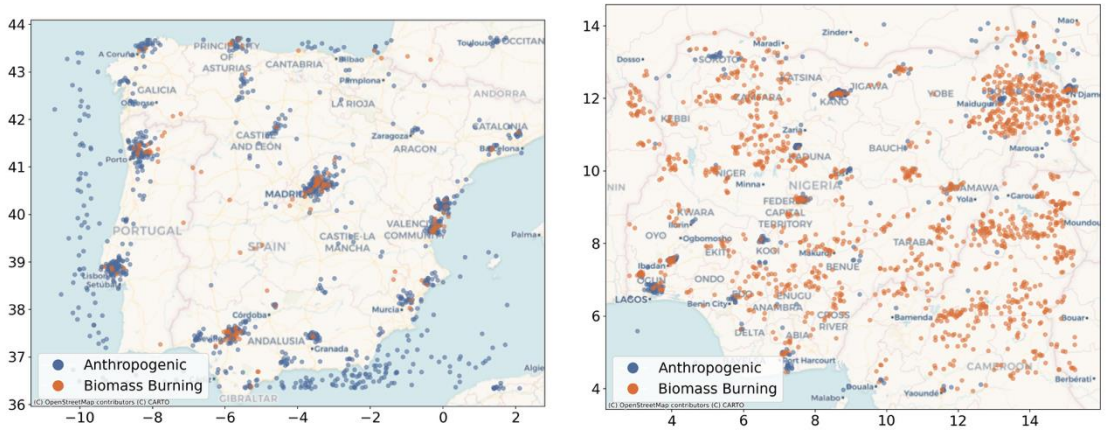

Figure 7: Map showing locations and outlines of plumes found in 2021 over the Iberian Peninsula (left) and Nigeria (right). Plumes flagged as anthropogenic are shown in the blue and biomass burning in orange (see text for details).

4.2 Examples for CO emissions

4.2.1 Cross-sectional flux analysis for African mega-cities

After verification of the validity and calibration of the method, we apply it to 29 of the largest cities in Africa. These cities are chosen based on their population or because they are emitting above the CSF's quantification threshold in the Dynamics–Aerosol–Chemistry–Cloud Interactions in West Africa (DACCIWA) inventory (Keita et al., 2021). Figure 8 shows the results of our TROPOMI quantification and a comparison with the DACCIWA inventory (Keita et al., 2021) and Emissions Database for Global Atmospheric Research (EDGAR) inventory

(Janssens-Maenhout et al., 2019). Note that we used different sizes for the city masks applied to the bottom-up inventories to ensure a fair comparison to the satellite-based emission estimates and found that the choice of city mask did not impact our conclusions.

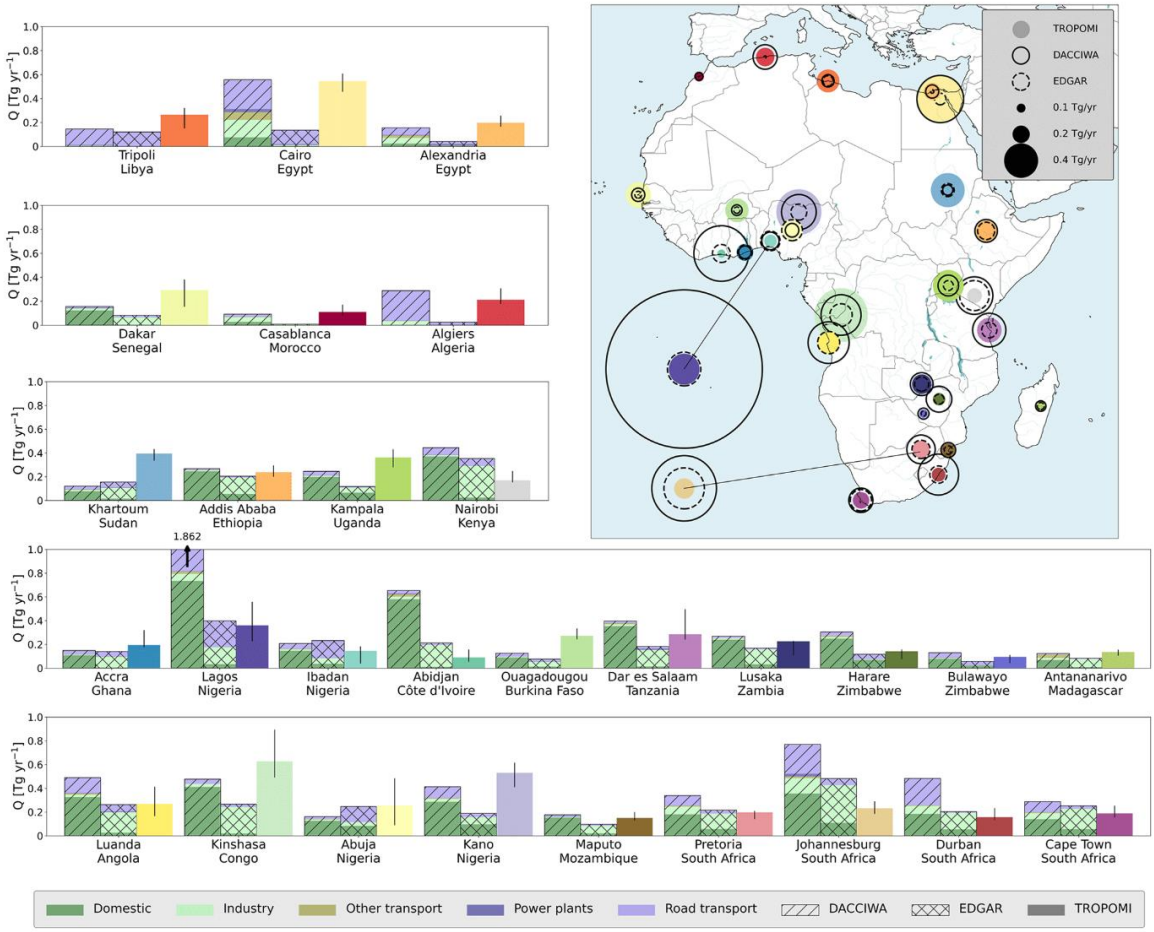

Figure 8: CSF emission quantifications for the largest African cities. Comparison between TROPOMI CO emission estimates averaged for 2019–2021 (shown as colored circles) and the DACCIWA and EDGAR v5 emission inventories for 2015 shown by the black (dashed) rings. The emission strength is indicated by the size of the circles or rings. The same comparison is made in bar plots, where the first two bars show the emission rates from DACCIWA and EDGAR respectively including the sectoral breakdown. The third bar gives the corresponding TROPOMI estimate, where the uncertainty is given by the range of the ensemble. The cities are ordered by geographical location. The emission estimate for Lagos in DACCIWA extends beyond the figure boundary.

Comparison of the TROPOMI-based emission estimates to DACCIWA and EDGAR bottom-up inventories shows that CO emission rates in northern Africa are underestimated in EDGAR, suggesting overestimated combustion efficiencies. We see the opposite when comparing TROPOMI to the DACCIWA inventory in South Africa and Côte d'Ivoire, where CO emission factors appear to be overestimated. Over Lagos and Kano (Nigeria), we find that potential errors in the spatial disaggregation of national emissions cause errors in DACCIWA and EDGAR, respectively. Finally, we show that our computationally efficient quantification method combined with the daily TROPOMI observations can identify a weekend effect in the road-transport-dominated CO emissions from Cairo and Algiers. For more details and further discussion, we refer to Leguijt et al. (2023).

4.2.2 Analytical inversion over Europe's iron & steel industry

We performed analytical inversions with 2019 TROPOMI CO satellite data to determine annual carbon monoxide emission rates for 21 European integrated iron & steel plants. These plants

are the largest carbon monoxide point sources in Europe. Their reported facility level emissions to the E-PRTR are used as prior estimates in our inversions. Per site, the inversion uses one of eight simulations with different meteorology for each day to reach optimal spatial agreement between observation and simulation. We allow for further freedom in the inversion by optimizing the CAMS-based background daily, rather than performing annual scaling. For a selected number of plants also a year 2020 estimate was derived (see Leguijt et al., 2025).

In addition to making use of inversions, TROPOMI CO data can be used to determine emission rates using simpler "mass balance" methods like the Cross-Sectional Flux (CSF) method. Leguijt et al. (2023) found a lower limit of 100 Gg/year to let the CSF method be able to estimate these emissions when applied to synthetic (model simulated) TROPOMI data. Seven of the studied plants; Arcelor Dunkerque, Gent, Dubrowie, Port Talbot S Works, Thyssen Schwelgern, Huttenwerke Duisburg, and U.S. Steel s.r.o., have reported annual emission rates above 100 Gg and therefore merit the application of the CSF. The results are shown in Figure 9. The comparison of the results with reported emissions data will be further analyzed as part of CORSO Task 1.3, where the intention is to also include the co-emitted NO_x emissions.

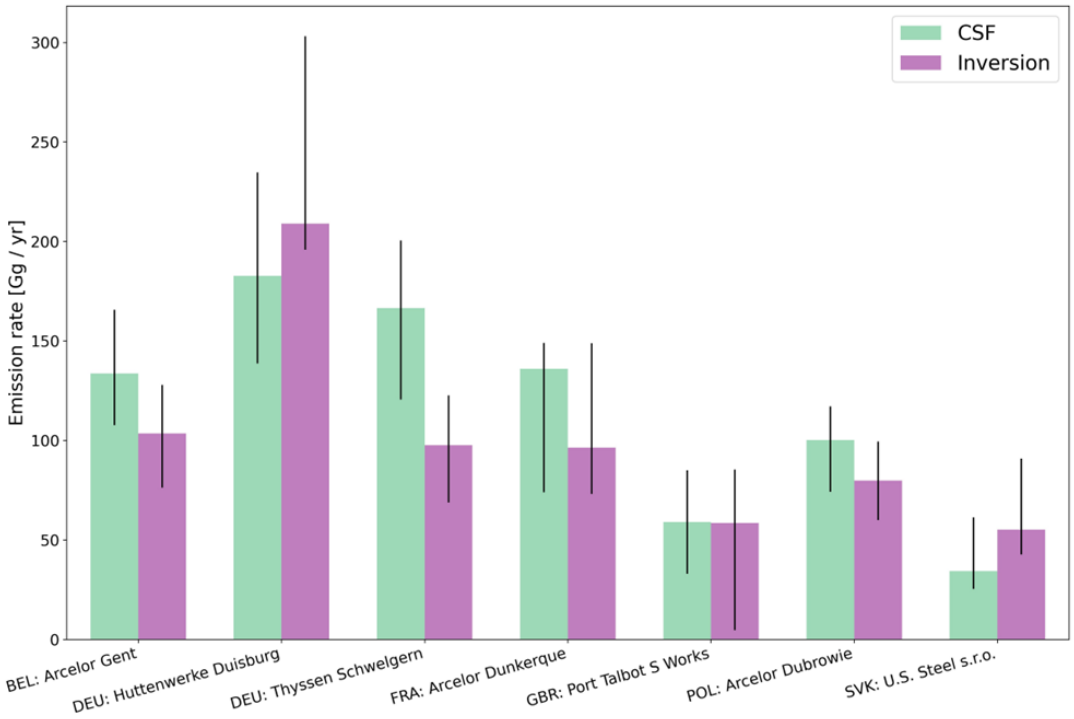

Figure 9: Comparison of year 2019 inversion-based TROPOMI CO emission estimates with estimates from the mass-balance cross sectional flux (CSF) method. As the lower limit for the TROPOMI-based CSF method was estimated at 100 Gg per year, only the plants with prior or posterior estimates above this value are compared. The error bars show the spread in estimated emissions over an ensemble of inversions with perturbed configuration parameters such a meteorological data or selection criteria.

4.3 Examples for NO_x emissions

The CSF and DIV method for quantifying NO_x emissions of point sources was applied to the strongest 100 point sources in the CORSO point source database, considering only sources that are not near other sources. The list includes Belchatow and Jänschwalde power plants for which MicroHH simulations are available. Furthermore, we made sure that at least ten sources were in Europe and the USA, where bottom-up estimates are expected to be of high quality. For Europe, we specially processed 20 iron/steel plants for which also CO emissions were estimated. For USA, we processed 12 power plants where daily NO_x emission reports are available.

Figure 10 compares annual emission of bottom-up and top-down methods for 20 point sources in Europe (EU27+UK) and USA. Overall, the agreement is very good for sources located in countries where bottom-up emissions are known well, providing confidence that satellite-based NO_x emission estimates provide sufficiently accurate estimates for regions where bottom-up estimates are less accurate.

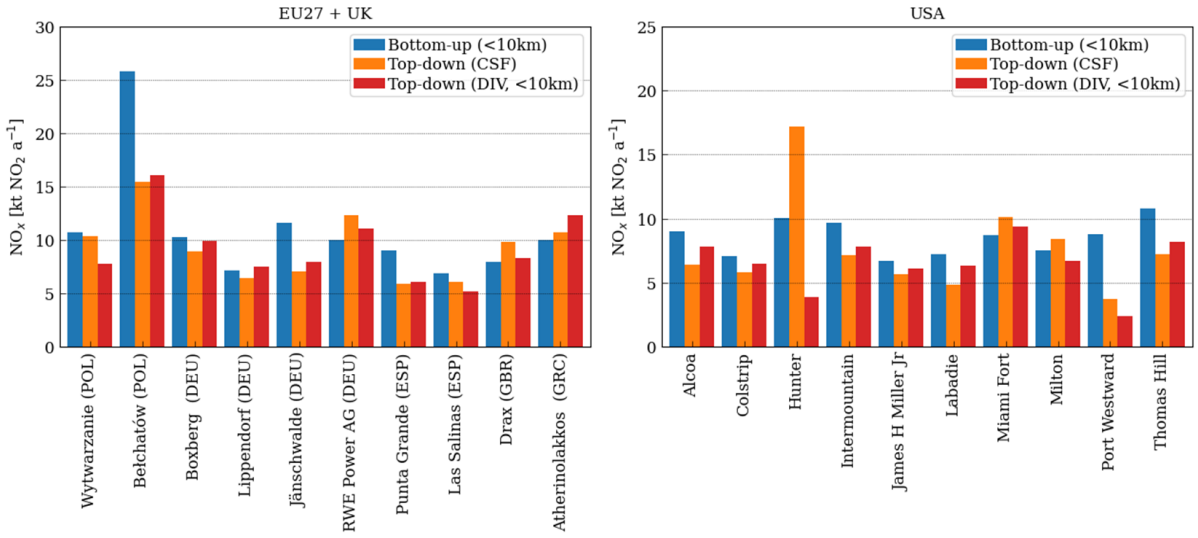

Figure 10: Example of annual estimated of NO_x emissions of point sources in Europe and USA, comparing bottom-up estimates from CORSO point source database (Version 1.0) with top-down estimates from the cross-sectional flux and divergence method. The deviations between bottom-up and top-down are analyzed and discussed in CORSO Deliverable 1.3. Countries are shown using ISO3 abbreviations.

Figure 11 compares time series for two European power plants, Bełchatów and Jänschwalde, and one U.S. power plant, Miami Fort, illustrating the impact of NO_2-to-NO_x conversion factors and temporal variability of emissions.

The Miami Fort time series includes daily bottom-up emission estimates, which reveal high emissions during winter and very low emissions from May to October. NO_x emissions derived from TROPOMI satellite observations show similar seasonal patterns, particularly when comparing monthly averages. This suggests that the satellite-based approach is well-suited for capturing the temporal variability of NO_x emissions.

A large uncertainty of the top-down NO_x emission estimates is related to the correction for NO_x chemistry, which includes the NO_2-to-NO_x conversion factor f and the NO_2 lifetime τ. The timeseries shows the estimates using the machine-learning (ML) model trained with GEOS-Chem simulation. The annual top-down estimates are shown using three different methods. First, default literature values for f (=1.32) and τ (=4 h) as used outside Europe. Second, the values from the ML model as used inside Europe (Schooling et al., 2025). Finally, values from plume-resolving chemistry simulations with the MicroHH model are used that are only available for selected power plants (Meier et al., 2024; Krol et al., 2024). The top-down estimates of annual NO_x emissions are similar using the default literature values and the ML model, because annual averages of NO_2-to-NO_x conversion factor and NO_x lifetime from the model are quite consistent with the literature values. The NO_2-to-NO_x conversion factors derived from MicroHH simulations are higher, slightly increasing the annual estimates. In addition, we used a lifetime of 2 h, which is the median found by Meier et al. (2024), when applying the CSF method to several cross sections downstream of the power plants, which allows for estimating the lifetime.

This analysis suggests that top-down estimates using the CSF method can underestimate NOx emissions by approximately 20-50%. We therefore provide conversion factors and

additional parameters with the dataset, enabling users to update NO_x chemistry corrections when more reliable data are available.

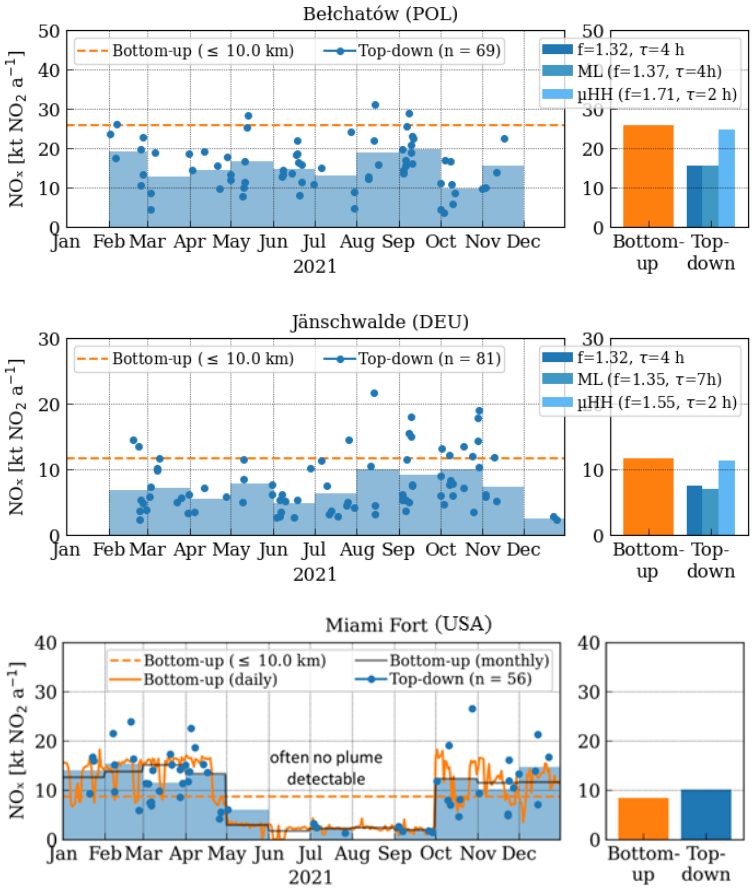

Figure 11: Example of time series of estimated NO_x emissions from cross-sectional flux method compared to bottom-up CORSO point source database (Version 1.0). For Belchatow and Jänschwalde power plant, top-down estimates are given for the three different methods used for accounting for NOx chemistry. For Miami Fort power plant, daily bottom-up estimates are available, showing good agreement with top-down estimates. The deviations between bottom-up and top-down are analyzed and discussed in CORSO Deliverable 1.3.

5 Future work

The dataset will be used in Task 1.3 for evaluating the time profiles.

The dataset has reference DOI 10.5281/zenodo.16838558

6 Acknowledgements

We acknowledge Ilse Aben (SRON) and Bram Maasakkers (SRON) for their support in the quantification of CO emissions from TROPOMI observations.

7 References

E-PRTR: Industrial Reporting under the Industrial Emissions Directive 2010/75/EU and European Pollutant Release and Transfer Register Regulation (EC) No 166/2006 - ver. 8.0 Mar 2023 (Tabular data), [dataset], 2023.

Beirle, S., Borger, C., Jost, A., and Wagner, T.: Improved catalog of NO2 point source emissions (version 2), Earth Syst Sci Data, 15, 3051-3073, 10.5194/essd-15-3051-2023, 2023.

Beirle, S., Borger, C., Dörner, S., Li, A., Hu, Z. K., Liu, F., Wang, Y., and Wagner, T.: Pinpointing nitrogen oxide emissions from space, Sci Adv, 5, ARTN eaax9800

10.1126/sciadv.aax9800, 2019.

Brunner, D., Kuhlmann, G., Marshal, J., Clement, V., Fuhrer, O., Broquet, G., Loscher, A., and Meijer, Y.: Accounting for the vertical distribution of emissions in atmospheric CO2 simulations, Atmospheric Chemistry and Physics, 19, 4541-4559, 10.5194/acp-19-4541-2019, 2019.

Clarisse, L., Van Damme, M., Clerbaux, C., and Coheur, P. F.: Tracking down global NH_3 point sources with wind-adjusted superresolution, Atmospheric Measurement Techniques, 12, 5457-5473, 10.5194/amt-12-5457-2019, 2019.

Finch, D. P., Palmer, P. I., and Zhang, T.: Automated detection of atmospheric NO2 plumes from satellite data: a tool to help infer anthropogenic combustion emissions, Atmos. Meas. Tech., 15, 721-733, 10.5194/amt-15-721-2022, 2022.

Hakkarainen, J., Kuhlmann, G., Koene, E., Santaren, D., Meier, S., Krol, M. C., van Stratum, B. J. H., Ialongo, I., Chevallier, F., Tamminen, J., Brunner, D., and Broquet, G.: Analyzing nitrogen dioxide to nitrogen oxide scaling factors for data-driven satellite-based emission estimation methods: A case study of Matimba/Medupi power stations in South Africa, Atmos Pollut Res, 15, 10.1016/j.apr.2024.102171, 2024.

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., De Chiara, G., Dahlgren, P., Dee, D., Diamantakis, M., Dragani, R., Flemming, J., Forbes, R., Fuentes, M., Geer, A., Haimberger, L., Healy, S., Hogan, R. J., Hólm, E., Janisková, M., Keeley, S., Laloyaux, P., Lopez, P., Lupu, C., Radnoti, G., de Rosnay, P., Rozum, I., Vamborg, F., Villaume, S., and Thépaut, J. N.: The ERA5 global reanalysis, Quarterly Journal of the Royal Meteorological Society, 146, 1999-2049, 10.1002/qj.3803, 2020.

Janssens-Maenhout, G., Crippa, M., Guizzardi, D., Muntean, M., Schaaf, E., Dentener, F., Bergamaschi, P., Pagliari, V., Olivier, J. G. J., Peters, J. A. H. W., van Aardenne, J. A., Monni, S., Doering, U., Petrescu, A. M. R., Solazzo, E., and Oreggioni, G. D.: EDGAR v4.3.2 Global Atlas of the three major greenhouse gas emissions for the period 1970-2012, Earth Syst Sci Data, 11, 959-1002, 10.5194/essd-11-959-2019, 2019.

Keita, S., Liousse, C., Assamoi, E. M., Doumbia, T., N'Datchoh, E. T., Gnamien, S., Elguindi, N., Granier, C., and Yoboué, V.: African anthropogenic emissions inventory for gases and particles from 1990 to 2015, Earth Syst Sci Data, 13, 3691-3705, 10.5194/essd-13-3691-2021, 2021.

Koene, E. F. M., Brunner, D., and Kuhlmann, G.: On the Theory of the Divergence Method for Quantifying Source Emissions From Satellite Observations, J Geophys Res-Atmos, 129, 10.1029/2023JD039904, 2024.

Krol, M., van Stratum, B., Anglou, I., and Boersma, K. F.: Evaluating NOx stack plume emissions using a high-resolution atmospheric chemistry model and satellite-derived NO2 columns, Atmos. Chem. Phys., 24, 8243-8262, 10.5194/acp-24-8243-2024, 2024.

Kuenen, J., Dellaert, S., Visschedijk, A., Jalkanen, J. P., Super, I., and van der Gon, H. D.: CAMS-REG-v4: a state-of-the-art high-resolution European emission inventory for air quality modelling, Earth Syst Sci Data, 14, 491-515, 10.5194/essd-14-491-2022, 2022.

Kuhlmann, G., Henne, S., Meijer, Y., and Brunner, D.: Quantifying CO2 Emissions of Power Plants With CO2 and NO2 Imaging Satellites, Frontiers in Remote Sensing, 2, 10.3389/frsen.2021.689838, 2021.

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

Kuhlmann, G., Koene, E. F. M., Meier, S., Santaren, D., Broquet, G., Chevallier, F., Hakkarainen, J., Nurmela, J., Amorós, L., Tamminen, J., and Brunner, D.: The ddeq Python library for point source quantification from remote sensing images (Version 1.0), EGUsphere, 2024, 1-27, 10.5194/egusphere-2023-2936, 2024.

Landgraf, J., de Brugh, J., Scheepmaker, R., Borsdorff, T., Houweling, S., and Hasekamp, O.: Algorithm theoretical baseline document for Sentinel-5 precursor: Carbon monoxide total column retrieval, Netherlands Institute for Space Research, the Netherlands, 2018.

Leguijt, G., Maasakkers, J. D., van der Gon, H. A. C. D., Segers, A. J., Borsdorff, T., and Aben, I.: Quantification of carbon monoxide emissions from African cities using TROPOMI, Atmospheric Chemistry and Physics, 23, 8899-8919, 10.5194/acp-23-8899-2023, 2023.

Leguijt, G., Maasakkers, J. D., van der Gon, H. A. C. D., Segers, A. J., Borsdorff, T., van der Velde, I. R., and Aben, I.: Comparing space-based to reported carbon monoxide emission estimates for Europe's iron and steel plants, Atmospheric Chemistry and Physics, 25, 555-574, 10.5194/acp-25-555-2025, 2025.

Meier, S., Koene, E. F. M., Krol, M., Brunner, D., Damm, A., and Kuhlmann, G.: A lightweight NO_2-to-NO_x conversion model for quantifying NO_x emissions of point sources from NO_2 satellite observations, Atmospheric Chemistry and Physics, 24, 7667-7686, 10.5194/acp-24-7667-2024, 2024.

Molod, A., Takacs, L., Suarez, M., Bacmeister, J., Song, I.-S., and Eichmann, A.: The GEOS-5 atmospheric general circulation model: Mean climate and development from MERRA to Fortuna, 11-37, 2012.

Santaren, D., Hakkarainen, J., Kuhlmann, G., Koene, E., Chevallier, F., Ialongo, I., Lindqvist, H., Nurmela, J., Tamminen, J., Amorós, L., Brunner, D., and Broquet, G.: Benchmarking data-driven inversion methods for the estimation of local CO2 emissions from synthetic satellite images of XCO2 and NO2, Atmos. Meas. Tech., 18, 211-239, 10.5194/amt-18-211-2025, 2025.

Schooling, C. N., Palmer, P. I., Visser, A., and Bousserez, N.: Development of a parametrised atmospheric NOx chemistry scheme to help quantify fossil fuel CO2 emission estimates, EGUsphere, 2025, 1-27, 10.5194/egusphere-2024-3949, 2025.

CORSO

Document History

Version	Author(s)	Date	Changes
1.0	Gerrit Kuhlmann, Erik Koene, Sandro Meier, Gijs Leguijt, Arjo Segers, Hugo Denier van der Gon, Chloe Schooling, Douglas Finch, Paul Palmer	18.7.2025	
1.1	Gerrit Kuhlmann, Erik Koene, Sandro Meier, Gijs Leguijt, Arjo Segers, Hugo Denier van der Gon, Chloe Schooling, Douglas Finch, Paul Palmer	13.8.2025	Addressing comments from internal review.

Internal Review History

Internal Reviewers	Date	Comments
Mattew Rigby (UNIBRIS) and Jaroslaw Necki (AGH)	July 2025	Uncertainty section is insufficiently described or lacking for some methods. Chapter 4 is definitely written with other language and some rephrasing would be helpful. Comments added aside of the text. Proposition for executive summary delivered.