Plume detection and estimate emissions for biomass burning plumes from TROPOMI Carbon monoxide observations using APE v1.1

Manu Goudar[1], Juliëtte Anema[2], Rajesh Kumar[3], Tobias Borsdorff[1], and Jochen Landgraf[1]

[1]SRON Netherlands Institute for Space Research, Leiden, The Netherlands
[2]Royal Netherlands Meteorological Institute (KNMI), The Netherlands
[3]National Center for Atmospheric Research (NCAR), USA

Correspondence: Manu Goudar (manugv@sron.nl)

Abstract.

This paper presents the Automated Plume Detection and Emission Estimation Algorithm (APE), developed to detect CO plumes from isolated biomass burning events and to quantify the corresponding CO emission rate. APE uses the CO product of the Tropospheric Monitoring Instrument (TROPOMI) aboard the Copernicus Sentinel-5 Precursor (S-5P) satellite, launched in 2017 and collocated active fire data of the Visible Infrared Imaging Radiometer Suite (VIIRS), the latter flying 3 min ahead of S-5P. After identifying appropriate fire events using VIIRS data, an automated plume detection algorithm based on traditional image processing algorithms selects plumes for further data interpretation. The approach is based on several thresholds which are tuned for data over the US on September 2020. Subsequently, the CO emission rate is estimated using cross-sectional flux method which requires the horizontal wind fields at the height of the plume. To infer proper wind fields from ECMWF reanalysis 5 data, we test three different plume heights. We consider a constant plume height at 100 m and at the injection height provided by Global Fire Assimilation System (GFAS) and we simulate the plume height in downwind direction with a 3D Lagrangian model. APE is configured to optimize performance of the US and its performance is verified for observations over Australia and Siberia. APE identified 5562 fire clusters and only 1327 cases was deemed to have good TROPOMI CO data. 882 plumes were detected in 1327 cases and only 378 plumes were considered as 309 plumes were short and about 195 had multiple sources of fire in them. Finally, the emissions were estimated for 226 cases in 378 cases out of which 5 were further visually filtered leading to 221 cases. The constant plume height at 100 m provides unacceptable emission estimates. The difference in emission estimates using a plume height at the GFAS injection height and using Lagrangian simulation, was found < 4% for Siberia but is significant for several cases for US and Australia. Therefore, APE will employ the simulated varying plume height in downwind direction for the automated algorithm. Finally, we provide a first estimate of the emission uncertainties. The assumption of a constant emission over the time of the plume formation and the spatial under-sampling of the CO column concentration by TROPOMI yields error < 20 %Mostly, an emission plume created by a burning and uncertainties in the GFAS injection height may cause errors < 100 %. Errors in the wind data are considered to be significant but could not the quantified because of the lack of uncertainty information in the used wind fields.

1 Introduction

25 Carbon monoxide (CO) is an air pollutant and in high concentrations, it causes harmful health effects. Moreover, it is an indirect greenhouse gas and a contributor to increase in several greenhouse gases in the atmosphere (Spivakovsky et al., 2000). CO is produced mainly due to incomplete combustion, for example, Andreae et al. (1988); Watson et al. (1990) have showed that the biomass burning increases CO in atmosphere. Shi et al. (2015) quantified the total CO emissions from vegetation burning, fuel-wood combustion, and human waste in different tropical regions. Also, Granier et al. (2011); Crippa et al. (2018);

30 Hoesly et al. (2018) show that the CO emission due fossil fuel burning has increased since 2000. Additionally, CO emitted by localized sources at the ground leads to a prominent footprint in the atmosphere, namely plumes, due to its lifetime from days to several months (Holloway et al., 2000). Thus, it becomes essential to understand the effect of CO on air-quality and climate by measuring and quantifying it accurately on a global and local scales.

The TROPOspheric Monitoring Instrument (TROPOMI) aboard the Sentinel-5 Precursor (S5P) satellite launched in 2017,

35 monitors CO daily on global scale (Borsdorff et al., 2018) and at a high spatial resolution of 7×7 km^2, improved to 5.5×7 km^2 in August 2019. Due to it's high spatial resolution and the daily coverage of the CO Level-2 dataset, the CO emissions by cities (Borsdorff et al., 2019a, 2020; Lama et al., 2020), wildfires (Schneising et al., 2020; Li et al., 2020; Magro et al., 2021; van der Velde et al., 2021) and industrial sources (Tian et al., 2021) have been quantified and studied. Rowe et al. (2022) compared the TROPOMI CO measurements to in-situ aircraft measurements for different biomass burning plumes in 2018 and

40 found that the difference between the two measurements is about 7.2%. This shows that the TROPOMI CO measurements can be use to quantify the biomass burning events.

Furthermore, most of the above studies estimate CO flux on large-scale regions (Schneising et al., 2020; Magro et al., 2021; van der Velde et al., 2021) and mega-city scales (Borsdorff et al., 2019a, 2020; Lama et al., 2020). However, not many single point emissions are quantified from the TROPOMI CO dataset. Tian et al. (2021) showed CO emissions based on TROPOMI

45 for single point industrial sources from India and China. They were able to perform a statistical study for three years as the geo-location is known for the industrial source. A similar study to quantify emissions from single point biomass burning (fires) using TROPOMI CO data has not been shown in the literature as geo-locations for fires are not known a priori. The geo-locations for fires can be detected using the Visible Infrared Imaging Radiometer Suite (VIIRS) 375m thermal anomalies/active fire product (Schroeder et al., 2014). The VIIRS instrument is aboard the joint NASA/NOAA Suomi National Polar-orbiting Partnership

50 (Suomi NPP) satellites and flies in the same orbit as S5P in loose formation with a temporal separation of 3.5 minutes between them. This short time difference helps in collocating observations of TROPOMI CO data and VIIRS active fire data product for this study.

The CO plumes in the TROPOMI data can be used to estimate CO emission by fires and different methods are discussed in the literature, namely, the inversion methods coupled with Gaussian dispersion models (Krings et al., 2011; Nassar et al., 2017;

55 Lee et al., 2019) or different Chemical Transport Models (CTM) (Brasseur and Jacob, 2017), Cross-sectional Flux Methods (CFM) (White et al., 1976; Beirle et al., 2011; Cambaliza et al., 2014, 2015; Kuhlmann et al., 2020) and integrated mass enhancement (IME) method (Frankenberg et al., 2016). The inversion coupled with a Gaussian plume model can be used for

flux inversion of isolated single plume event assuming steady and uniform wind conditions. The method fits an analytically computed Gaussian plume to TROPOMI CO column observations. The method can only be applied to observations under very specific wind conditions (Varon et al., 2018). The IEM uses an empirical linear relation between the emission and the integrated mass in the observed plume that was established based on aircraft data (Frankenberg et al., 2016). Current implementations are only applicable for single emission events for methane plumes and no such relationship has been established for CO measurements in fires. Thus, we do not consider IME for the present work. The inversion methods using CTMs, such as Weather Research and Forecasting model coupled to Chemistry (WRF-Chem) (Grell et al., 2005), GEOS-Chem (Bey et al., 2001), and others can reduce uncertainties thereby predicting emissions more accurately. They can be applied to complex emission events, but the corresponding simulations are complex, computationally expensive and difficult to automate, in particular for a large number of fires with different geolocations which is the objective of this study. The CFM is suited for the interpretation of isolated plumes but does not assume a particular plume shape. It needs comparatively less computational power and is easier to automate. It is based on the mass conservation of the pollutant transport in downwind direction of the plume. The CO emission is estimated from corresponding fluxes across different planes perpendicular to the direction of plume using the wind velocity at the plume height. The plume height depends upon different aspects, namely, meteorology, emission height, etc (Brunner et al., 2019) and may not be explicitly available. Moreover, the CFM breaks down when diffusion is dominant, i.e, when the wind velocity is $< 2\,\mathrm{ms}^{-1}$ (Varon et al., 2018).

The present work aims at developing an automated scheme to detect single and spatially isolated emission of biomass burning events on the spatial scale of TROPOMI observations and to estimate corresponding CO emissions. For this purpose, we employ and improve the CFM as it has the potential to be applied in an operational data processing. First, VIIRS fire data and satellite data are prepared for automated plume detection which is discussed in section 2.1. Plume detection algorithm from a single point source using VIIRS fire counts is the subject of section 2.2. Section 2.3 describes emission estimation using cross-sectional flux method where an appropriate choice of the plume height and the wind fields are discussed. The study results are discussed in Section 3 and finally Section 4 concludes our study and sets recommendations for the future work.

2 Methodology

Figure 1 illustrates a high-level flowchart of APE. Corresponding pseudo-code is given in Appendix B algorithm description 1. It is divided into three parts, namely data preparation, automatic plume detection, and emission estimation. During data preparation, the algorithm identifies single point fire sources from VIIRS 375 m active fire data product (Schroeder et al., 2014) and subsequently selects and extracts TROPOMI CO data around every located fire source. The plume detection algorithm searches for a plume in the extracted CO data which is required as an input for emission estimation. The emission estimation algorithm initially computes the background CO which is the usual observed CO concentration at that location without any CO emissions due to the fire. Subtracting the background allows us to obtain the enhanced CO which further is used to estimate the emissions by cross-sectional flux method. These three parts of the algorithm are discussed in the following sections.
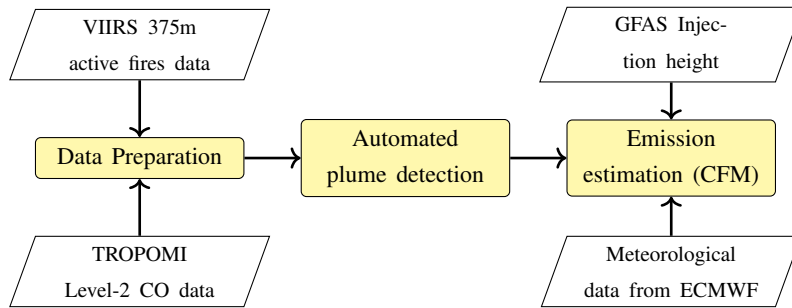
Figure 1. High-level flow chart of the APE algorithm.

 ## 2.1 Data preparation

2.1.1 Selection of fire events

Fire events are inferred from the VIIRS 375m active fire data product (Schroeder et al., 2014) provided by Fire Information for Resource Management System (FIRMS) which is operated by NASA's Earth Science Data and Information System, (https://earthdata.nasa.gov/active-fire-data). The data contains different parameters such as fire radiative power (FRP), temperature and the time of measurement defined at latitude-longitude coordinates which correspond to the center of a 375×375 m^2 ground pixel. From now on, each of these latitude-longitude coordinates will be referred to as a fire count in this paper. Mostly, an emission plume created by burning a single VIIRS pixel cannot be detected by TROPOMI with its pixel size of 7×7 or 5.5×7 km^2. Only clusters of VIIRS fire counts can lead to a detectable CO plume in the observations. To identify appropriate fire clusters, we employ the Density-Based Spatial Clustering of Applications with Noise (DBSCAN) (Ester et al., 1996; Schubert et al., 2017) algorithm from the scikit-learn library (Pedregosa et al., 2011). It separates the areas which are densely packed with fire counts from the low density areas and has an ability to detect arbitrarily-shaped clusters. DBSCAN requires two inputs, first the maximum search radius, r_{max}, around a fire count and second the minimum number of fire counts within the area, n_{min}. r_{max} is set to 4 km which is approximately half of the TROPOMI pixel size. The minimum number of fire counts has been empirically set to $n_{min} = 10$. For further analyses, we convert each cluster to a single point source using fire radiative power (FRP) as weights of the individual fire counts. This single point source will be referred to as fire source from now on and serves as an input to TROPOMI CO data preparation.

2.1.2 TROPOMI CO data preparation

For the identified fire sources from VIIRS data, the corresponding TROPOMI orbit (See Table B1 for L2 product version) is selected. Figure 2 illustrates the collocated information for part of TROPOMI orbit number 10254 over Australia. The orbit is corrected for stripes (see Fast Fourier Transformation algorithm of Borsdorff et al. (2019b)). For each fire source, we extract a data granule of 41×41 TROPOMI CO pixels centered around the source. The granule size is minimum 220 km is chosen

to capture the distance of the emission traveled from the source to the granule edges within 6h with an average velocity of 5 ms^{-1}. After extraction, two data quality filters are applied.

DP-1 The maximum pixel size due to distortion in swath direction is restricted to < 12 km to avoid large pixel size and its variation within the granule.

DP-2 A sufficient number of CO pixels with data quality $q_a > 0.5$ (Apituley et al., 2018). $q_a > 0.5$ corresponds to clear-sky, clear-sky like and mid-level clouds observations. Within the entire data granule 80% of all pixels must fulfil this quality criteria. Additionally we require 85 % of 'good' pixels in an area of 7×7 pixels centered around the fire source. The more usable pixels around plume the better the plume can be disentangled from the atmospheric background (see discussion in Sec 2.3.1).

The threshold values are empirically determined for a reference data set from September 2020 over the US and verified for two other data sets over Australia and Siberia (see Tab. B1 in Appendix B for a detailed specification of the data sets). Finally, the selected CO scene corresponding to a fire source is forwarded as an input to the plume detection algorithm.
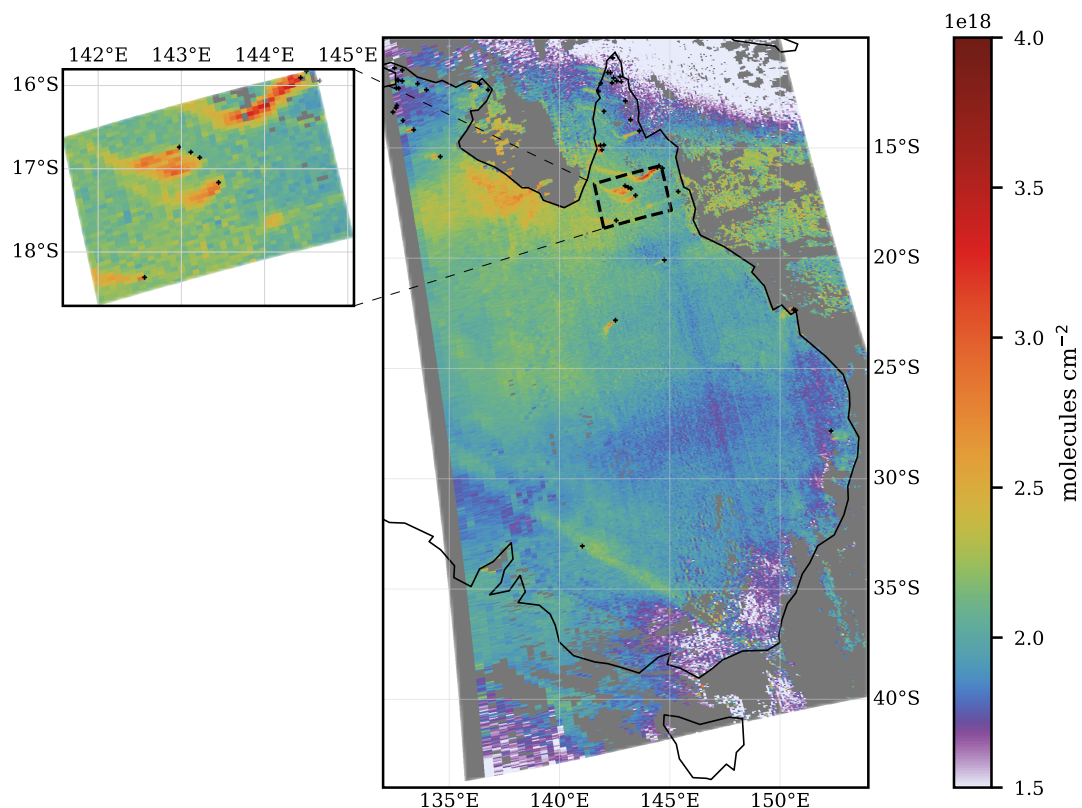
Figure 2. 49 detected fire sources represented by black '+' on 2019-10-06 overlapped with the TROPOMI level 2 CO data for orbit 10254. The dashed region represents a 41×41 pixel granule.

2.2 Plume detection algorithm

125 The next step of APE identifies plumes within each selected CO data granule. Kuhlmann et al. (2019) developed a plume detection algorithm based on statistical methods and Finch et al. (2021) used machine learning to detect plumes. A machine learning approach is not considered in the present study mainly due to non-availability of data containing detected plumes and their sources for training. Instead, our plume detection approach is based on traditional image processing algorithms (van der Walt et al., 2014).

130 Using the extracted CO TROPOMI data, a plume is detected by a region-based segmentation algorithm where pixels with similar properties are clustered together to form a homogeneous region. One of the most commonly used and classic region-based segmentation algorithm is the 'marker based watershed transform method' (Beare, 2006; Gao et al., 2004). The CO column concentration represents metaphorically the altitude of a topographic map. Thus, the watershed algorithm segments the regions into valleys and mountains (CO enhancements) based on a given marker and a gradient map. In the following 135 paragraphs, we describe the plume detection in more detail using an example.

The marker-based watershed algorithm in the scikit-image package (van der Walt et al., 2014) takes two inputs to segment an image, one is the 'gradient map' $\mathbf{I}_{grad}$ where the changes in altitude are emphasized and homogeneous regions are dampened. The second input is a marker image $\mathbf{I}_{mark}$ which provides the seed points referred by an integer label for the algorithm. The definition of both inputs is discussed in the following paragraphs.

140 We start with the extracted CO TROPOMI granule of a 41×41 pixel size $\mathbf{I}(i,j)$ with $i,j = 1, \cdots, 41$. An example is shown in Figure 3(a). First, high frequency components or noise of the CO-image are reduced by a 2D Gaussian filter with standard deviation $\sigma = 0.5$ (pixel), which was chosen empirically. The smoothened image is referred to as $\mathbf{I}_s$. From this image, the gradient map $\mathbf{I}_{grad}$ is computed using a Sobel operator (Sobel and Feldman, 1990; van der Walt et al., 2014) as shown below

$$\mathbf{I}_{grad} = \sqrt{\mathbf{G}_x + \mathbf{G}_y} \tag{1}$$

145 with

$$\mathbf{G}_x = \begin{bmatrix} 1 & 0 & -1 \\ 2 & 0 & -2 \\ 1 & 0 & -1 \end{bmatrix} * \mathbf{I}_s \qquad \mathbf{G}_y = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix} * \mathbf{I}_s, \tag{2}$$

where $*$ represents the convolution operator. Here, the gradient $\mathbf{I}_{grad}$ emphasizes the edges of a plume as shown in Figure 3b.

By default the marker image ($\mathbf{I}_{mark}$) is initialized to zero and subsequently two different seeds are defined. One seed indicates regions which do not have CO enhancements and another refers to regions of clear CO enhancements given by $\mathbf{I}_{mark}(i,j) = 1$ 150 and $\mathbf{I}_{mark}(i,j) = 2$, respectively. The seeds are defined as follows:

1. The regions of no CO enhancements: A pixel $\mathbf{I}_{mark}(i,j)$ has no CO enhancement if it is either below the median of $\mathbf{I}_S$ or below the mean of the 15x15 pixels centered at $\mathbf{I}_S(i,j)$. The size of 15x15 pixels is empirically chosen to account for the background variability. The pixels corresponding to no enhancement can be seen in Figure 3c represented by label '1' and is referred to as preliminary marker image.
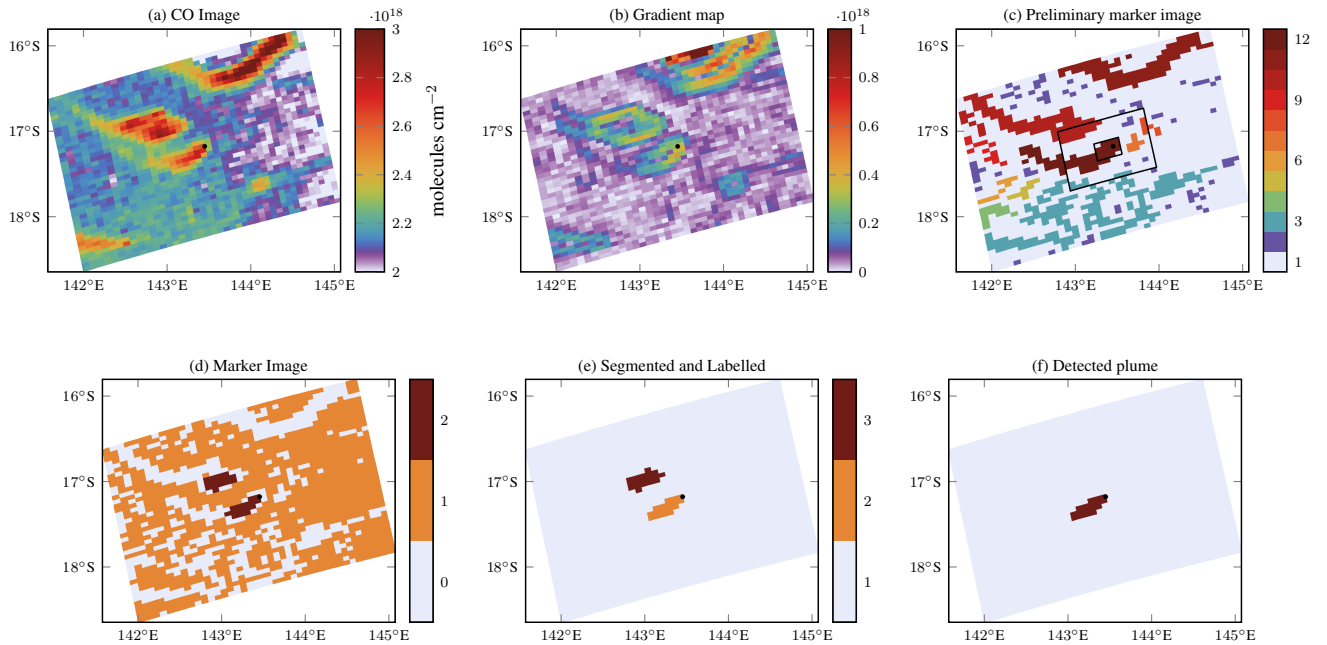
Figure 3. Plume detection algorithm. Figure shows different steps of the algorithm for an example plume for a fire source on 6th October 2019 in Australia.

2. The regions with CO enhancements: Using the preliminary marker image with labels '0' and '1', we identify all connected pixels with same marker value (referred to as connected regions hereafter) using the 'label' algorithm (Fiorio and Gustedt, 1996) in the scikit-image (van der Walt et al., 2014). Each connected region is identified by an unique integer value per pixel (not to be confused with the seed marker). Next, we zoom in an 5×5 pixel area around the fire source, and extract all connected regions as a potential plumes. We extend the potential plumes by going to a 15×15 pixels around the fire source using pixels of the same label. We extract all labeled CO data in this area and calculate a CO threshold as their mean CO value. We mark all pixels within the 15×15 pixel area with a seed '2' which have CO values above this threshold, which further yields the remaining seed points which are defined only in a 15×15 pixel area around the fire source.

The above selection process is illustrated in Fig. 3c with the different labels of connected regions. The final marker image is shown in Figure 3d. Finally, the watershed algorithm calculates a segmented image for the entire domain using the gradient map image $\mathbf{I}_{\mathrm{grad}}$ and the marker image $\mathbf{I}_{\mathrm{mark}}$. Figure 3e gives an example of a segmented image. Using the gradient map, the watershed algorithm decided that the two areas of enhanced CO values are not connected and so do not belong to the same plume. From the segmented image, we extract the correct plume which should originate from emissions at the source location. Therefore, we consider only those labeled areas that overlap with the center 7×7 pixels. Figure 3f shows the detected plume. The detected plume in this case seem to be shorter, however the tail end of the plume i.e, $< 143^0$E, will fail background

subtraction due to similar enhancements as the background. This can also be seen in gradient map where no gradient is detected on top side of the plume.

Finally, the extracted plume is evaluated regarding its suitability for further processing and the plume length is calculated for each extracted plume. Only the data are provided to the emission estimation module if:

175   PD-1 the plume length > 25km.

If the plume length is < 25 km then the detected plume is flagged as short plume and will be ignored for further processing. The short plumes are difficult to quantify in an automated way as they can have different shapes which makes it difficult to identify the plume direction. Before starting the emission estimate module, we remove all plumes with multiple fire sources, as the aim of this paper is to quantify fires with single sources. This is done by the following filter:

180   PD-2 Check if there are not more than 9 non-clustered fire counts or any other identified fire cluster within $0.05°$ distance from or in the identified plume.

2.3   Emission estimation

For the detected plumes, the emissions were estimated using the cross-sectional flux method (CFM) (White et al., 1976; Beirle et al., 2011; Cambaliza et al., 2014, 2015; Kuhlmann et al., 2020). The CO emission E is defined as the mean flux through n

185 cross sections perpendicular to the downwind direction of the plume, namely

$$E = \frac{1}{n} \sum_{i=1}^{n} Q_i$$
$$Q_i = \int \delta C_{co}^i(s, t_0) \cdot v^i(z_i, s, t_0) \cdot ds \tag{3}$$

where Q_i (in kgs^{-1}) is the CO flux through cross section i, δC_{co}^i (in kgm^{-2}) is the background subtracted CO values along a cross-section i and v^i (in ms^{-1}) is the velocity perpendicular to the cross-section i. The wind velocity $v(z, s, t_0)$ at the plume

190 height z, the cross-section position s and at the observation time t_0 is obtained from the European Center for Medium range Weather Forecasts (ECMWF) Reanalysis v5 (ERA5) data (Hersbach et al., 2018). For error characterization, we define the standard error (σ_{E}) as

$$\sigma_{\mathrm{E}} = \frac{1}{n} \sqrt{\sum_{i=1}^{n} (E - Q_i)^2} \tag{4}$$

$$\tag{5}$$

195 To determine the cross-sections, hereafter referred to as transects, we first calculate the direction of the plume in downwind direction. The plume line results from a second order curve fit through the pixel centers of the detected plume (see e.g. the black solid line in Fig. 4a). Next, we calculate the transects at every 2.5 km perpendicular to the plume line and are illustrated as dashed lines in Figure 4a. Transects are sampled at 2.5 km to reduce the errors due to interpolation discussed in next paragraph.
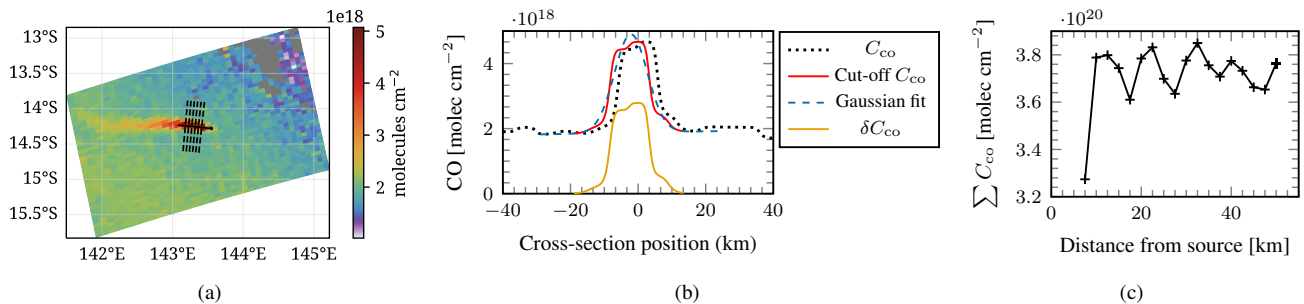
Figure 4. Plume on 2019-05-19 at 04:55 UTC. (a) Plume and every second transact lines drawn based on the detected plume separated by 5 km in downwind direction. (b) The black dotted line corresponds to CO column along a transect in (a) and the red line shows re-centered and cut-off CO used for Gaussian fitting. Blue-dash dotted line corresponds to the Gaussian fit and orange line represents the enhanced CO along the transect. (c) $\sum C_{co}$ along a transect against the distance from source.

To calculate Q_i in Eq. 3, each transect is sampled at distances of 500 m. The points over an transects are over-sampled to get a smoother CO distribution which further helps in background subtraction discussed in Sec. 2.3.1. Along each transect, the CO column (C_{co}) is extracted by linear interpolation of the original CO data and is illustrated by dotted black line in Figure 4b. This CO column is further used to compute δC_{co} in Eq. 3. During the diagnostics tests of our interpolation algorithm, an oscillation was observed in the CO columns integrated along the transects as a function of the downwind distance from the fire source (see Figure 4c). The oscillation is due to the under-sampling of the CO distribution by the TROPOMI instrument. The distance between two minima is approximately equal to the TROPOMI pixel size. This error is found to propagate further into the CO enhancement δC_{co} which is computed from background subtraction algorithm.

2.3.1 Background Subtraction

To determine the atmospheric background of CO per transect, first we re-center the C_{co} such that the maximum is at the origin to facilitate the Gaussian fit. The transect line is truncated at the first minima of CO on either side of the origin as illustrated by red line in the Fig. 4b. To determine the background for each transect (red line), we assume that the column CO along the transect can be expressed as

$$C_{co} = H_0 + H_1 \cdot s + A_0 G(s) \tag{6}$$

where H_0 and H_1 represents background and the slope of change in background over the transect, respectively. A_0 is the amplitude of the Gaussian distribution (G). We determine the background by fitting Eq. 6 through the CO data, which we subsequently subtract from the C_{co} data to calculate the CO enhancement

$$\delta C_{co} = \max\{0, \, C_{co} - H_0 + H_1 \cdot s\}. \tag{7}$$

Here, the negative CO column enhancements are ignored. The blue dashed line and the orange line in Fig. 4b represent Gaussian fit and δC_{co}, respectively.

2.3.2 Filtering during background subtraction

The background subtraction includes an important filtering mechanism to remove overlapping plumes. This is done during the background subtraction after the transect line is truncated. The filter criteria

EE-1 Difference between the two minima on either side of a truncated transect < 10% of the smallest value ensures a smooth background and so the absence of any interference with adjacent emission events. It should be noted that this filter also account for background situation which are too demanding for the CFM.

2.3.3 Plume height

The plume height z_i at a transect/cross-section i is used to extract the appropriate wind velocity $\upsilon(z_i, s, t_0)$. For wildfires, Rémy et al. (2017) showed that the an Integrated monitoring and modeling System for wild-land Fires (IS4FIRES) injection height, z_{inj}, from the Global Fire Assimilation System (GFAS) database is in a good agreement with the observations. Sofiev et al. (2012) showed the IS4FIRES injection height deviated by less than 500 m compared to MISR Plume Height Project (MPHP), thus we consider 500m as plume height uncertainty. First, we assume that the plume height is equal to z_{inj} and is constant throughout the plume, which may hold true for stable meteorological conditions. The constant plume height will be referred as z_c and uncertainty at this plume height is given as z_c^p and z_c^m which correspond of $z_{inj} + 500$ m and $z_{inj} - 500$ m, respectively.

It should be noted that the injection height computed in GFAS is for 24 h and may not be appropriate for a satellite plume which is a snapshot at time t_0. Additionally, the plume height might vary due to meteorology in downwind direction. Therefore, we simulated alternatively particle trajectories starting at the fire location around the injection height with a three dimensional Lagrangian tracer dispersion model. This allows us to estimate the local plume height z_i by a vertical averaging of tracers along the downwind direction. The estimated plume height is referred as z_{lag} in the following, which captures the change in height in the downwind direction.

The Lagrangian simulations are performed using tracer particles. The motion of tracers is simulated according to

$$\frac{d\mathbf{x}_p(t)}{dt} = \mathbf{v}(\mathbf{x}_p(t)) \tag{8}$$

where $\mathbf{v}(\mathbf{x}_p)$ represents the fluid velocity at the instantaneous particle position $\mathbf{x}_p$. The explicit forward Euler scheme (Butcher, 2003, p. 45) is employed to integrate the equation in time. The velocity on the right-hand side of the Equation. (8) is calculated by the tri-linear interpolation of the ERA5 velocity fields. The source locations for the Lagrangian simulations are based on the fire counts described in Section 2.1.1. At each source location, 3 tracer particles are released at z_{inj} and $z_{inj} \pm 500$ m and the particle trajectories are simulated. Particles are released at $z_{inj} \pm 500$ m for uncertainty analysis. The end time of the simulations is the TROPOMI measurement time t_0, which is about 13:30h local time (Veefkind et al., 2012), and the simulation starts at $t_0 - 6$ h. The particles are released from the source locations every 2 minutes. Figure 5(a) shows the tracer particle simulation for a plume and the grayish band indicates them at TROPOMI measurement time. The contribution to emissions from fires is low in early morning as shown in the ecosystem-specific diurnal cycles by Li et al. (2019) and so we ignore trajectory
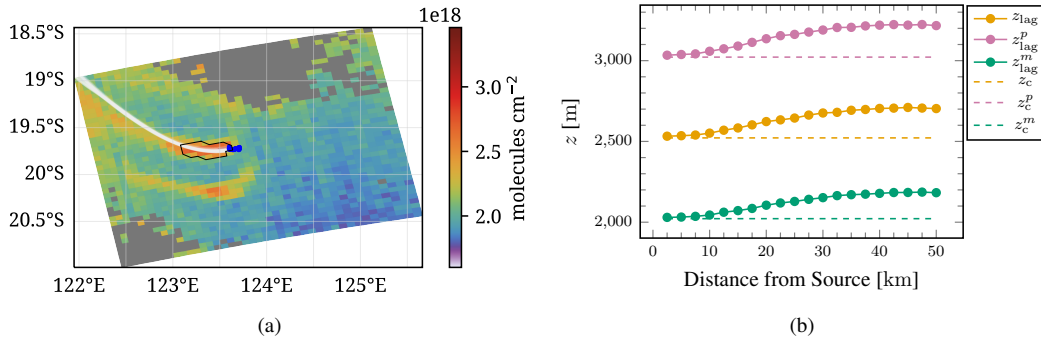
Figure 5. (a) The white band shows all tracer particles at the end of Lagrangian simulation and the blue dots show the fire counts on the detected plume. (b) Shows the plume height computed for different transects from Lagrangian simulations. The constant plume height (z_c) represented by orange dashed line is 2521.87 m.

simulations before $t_0 - 6$ h. Additionally the process of heating due to fires is not accounted in our Lagrangian simulation as we assume the ERA5 velocity fields contain some aspect of it as ERA5 assimilates skin surface temperatures.

At each intersect, the height of the tracer particles released at z_{inj} are extracted and the mean height, $z_{\mathrm{lag},i}$ is computed and assumed to be constant along the transect. Figure 5(b) shows the plume height for different transects from the fire source which is used to compute velocity, v, in Equation 3. The uncertainty in plume height is defined as z_{lag}^p and z_{lag}^m and are computed from tracer particles that were released at height $z_{\mathrm{inj}} + 500$ m and $z_{\mathrm{inj}} - 500$ m, respectively and can be observed in Figure 5(b). Finally, the velocity, v, is used to compute emissions.

2.3.4 Filtering during Lagrangian simulations

Related to the Lagrangian simulation, we apply three filters:

EE-2 The injection height from GFAS must be available.

EE-3 If the simulated trajectories are not aligned in the direction of the plume then the plume is rejected.

EE-4 If the wind velocity at the TROPOMI measurement time which is used to compute emissions is less than $2 \mathrm{~ms}^{-1}$ then the plume is rejected.

Filter EE-2 may become relevant due to false detection of plume or false fire in VIIRS active fire database or just the data is missing in GFAS database. Reason for filter EE-3 might be a rotation or the errors in the ERA5 velocities or due to the spatial and temporal resolution of velocity fields or inaccurate injection height. Finally, if the wind speed is below the specified value in EE-4, diffusion dominates the pollutant transport and CFM is not appropriate to estimate the CO emission.

Table 1. Results for automated plume detection and emission estimation algorithm (APE v1.1) for four months in US, Australia and Siberia

Regions	Fire Clusters	CO data	Plume detection	Emission estimation	Visual Inspection
US	1081	213	130	37	35
AU	2013	385	266	129	128
Sib Jun	416	130	83	35	34
Sib Jul	2052	599	94	25	24
All Regions	5562	1327	378	226	221

3 Algorithm application

Our APE algorithm targets at global performance and includes several threshold values, which needs to be determined carefully for optimal performance. For the current version APE V1.1. we decided to determine the thresholds using the region encapsu-
270 lating the United States of America (US) on Sept 2020. The algorithm is verified by applying it to other regions encapsulating Australia (AU) on Oct 2019 and Siberia (Sib) on Jun and Jul 2021 (see Table. B1 for more details on the selected regions). It is important to note that these regions are not used to configure APE and so can be used to test the overall performance of the algorithm. The different time periods were chosen to focus on the regional burning season and so to maximize number of observed fires. Table 1 shows the number of plume evaluated in the different modules of APE. *Fire cluster* and *CO data*
275 columns highlight the results of data preparation (see Section. 2.1) part. The columns *plume detection* and *emission estimation* show the results for plume detection (sec. 2.2) and emission estimation (sec. 2.3) parts. Furthermore, the details corresponding to filtering can be found in Tables B2, B3 and B4 in appendix.

A total of 5562 fire clusters (see Table. 1) were identified in the VIIRS active fire data product for all regions based on the clustering method discussed in Section 2.1.1. For each fire cluster, TROPOMI CO data was filtered for maximum pixel size
280 and quality (see Sec. 2.1.2). The filter on the TROPOMI pixel size (DP-1) filtered about 1533 cases out of 5562 cases which mostly belonged to Australia and US. The quality of TROPOMI CO data (DP-2) was found to be bad for about 2553 cases in 5562 cases. For the Siberian region on July 2021 more than 50% of all fire clusters are flagged as bad quality data because of more cloudy observations at these latitudes. Finally the data preparation part yielded a total of 1327 good CO data granules for all regions for further processing.

285 The plume detection algorithm described in Sec. 2.2 identified a total of 378 plumes for all regions from available 1327 good CO data cases. A total of 445 cases were found to have no enhancements (see Table B3) which means that the CO enhancement from these fires was not significant enough to be measured by TROPOMI. In addition to this, about 309 identified plumes were flagged as short as the length of these plumes was found to be shorter than 25 km (PD-1 filter). Further, the PD-2 filter identified a total of 195 cases where other fire sources and clusters were present in the detected plumes.

290 The emission estimation algorithm takes 378 plumes as an input and computes emissions using CFM for a total of 226 cases. So a total of 152 plumes were rejected by filter EE-1, EE-2, EE-3, and EE-4 during the emission estimation (See Table. B4 for details). The injection height from GFAS database was not available for 57 cases (EE-2) and 29 plumes failed because
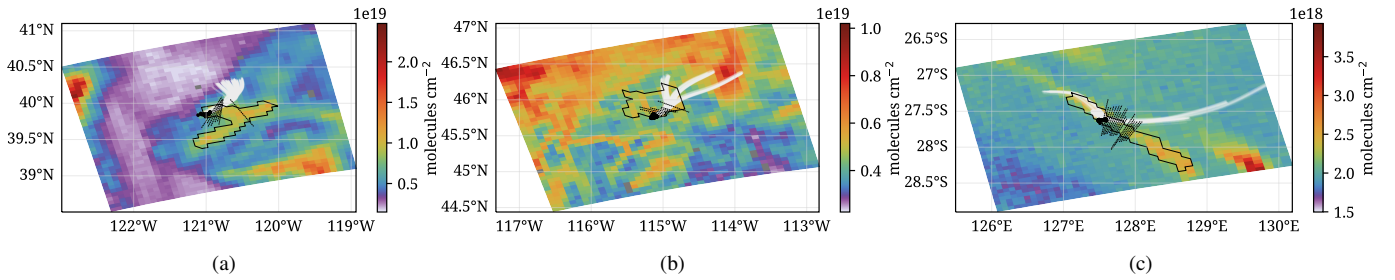
Figure 6. False positive detection of plumes. Black dots indicate the fire counts, white bands the tracer particles and the black polygon depicts the detected plume, respectively. The black dashed lines are transaction lines.

overlapping with other plumes (EE-1). Furthermore, particles-plume alignment EE-3 removed a total of 51 cases. This is mostly due to bad plume detection, inaccurate velocities or injection heights. Lastly, the velocity filter EE-4 rejects about 15 cases. Finally, the automated plume selection can be verified with a visual inspection. For the US region difference are small by construction as the threshold values were tuned for this data ensemble. For the verification areas Australia and Siberia we see similar good performance of the detection algorithm, and in total only 5 of the 226 plumes were flagged as false identifications. Fig. 6 gives three examples of the false positive detection. Thus, our analysis confirms the applicability of our algorithm to other areas with a confidence of 97.7% of the cases.

3.1 Cross-Sectional Massflux method (CFM)

The CFM computed emissions for 221 cases. To compare the effect of plume heights, two variables are defined, namely, mean plume height $\overline{z}_{\text{lag}}$, which is mean of z_{lag} of all transects along the downwind direction of the plume and the maximum rise in plume height (δz) with respect to z_c. They are given as

$$\overline{z}_{\text{lag}} = \frac{1}{m} \sum_{i=0}^{m} z_{\text{lag},i} \tag{9}$$

$$\delta z = \max\{z_{\text{lag}}\} - z_c \tag{10}$$

The mean plume height $\overline{z}_{\text{lag}}$ is plotted against the constant plume height (z_c) in Figure 7(a) and (c) for US and Australia, and Siberian region, respectively. The δz decreased and increased in downwind direction for about 43 and 178 fires, respectively. Additionally, the δz in the downwind direction is found to vary > 500 m for 30 fires in regions encapsulating Australia and US as represented by the orange color in Figure 7(a). However, no such cases were found in Siberia (see Fig 7(c)). Among these 30 fires, about 11 fires had $\delta z > 1000$ m. As the total fire radiative power (FRP) and fire counts represent the heat generated and the area burnt, a relation between plume height rise and these two variables can be expected as higher FRP means higher temperature which heats up the air, leading rise of the warm air. However, no such relation was observed and there were cases with low FRP or low fire counts where the $\delta z > 1000$ m. Additionally, it was difficult to find an appropriate reason for large rise
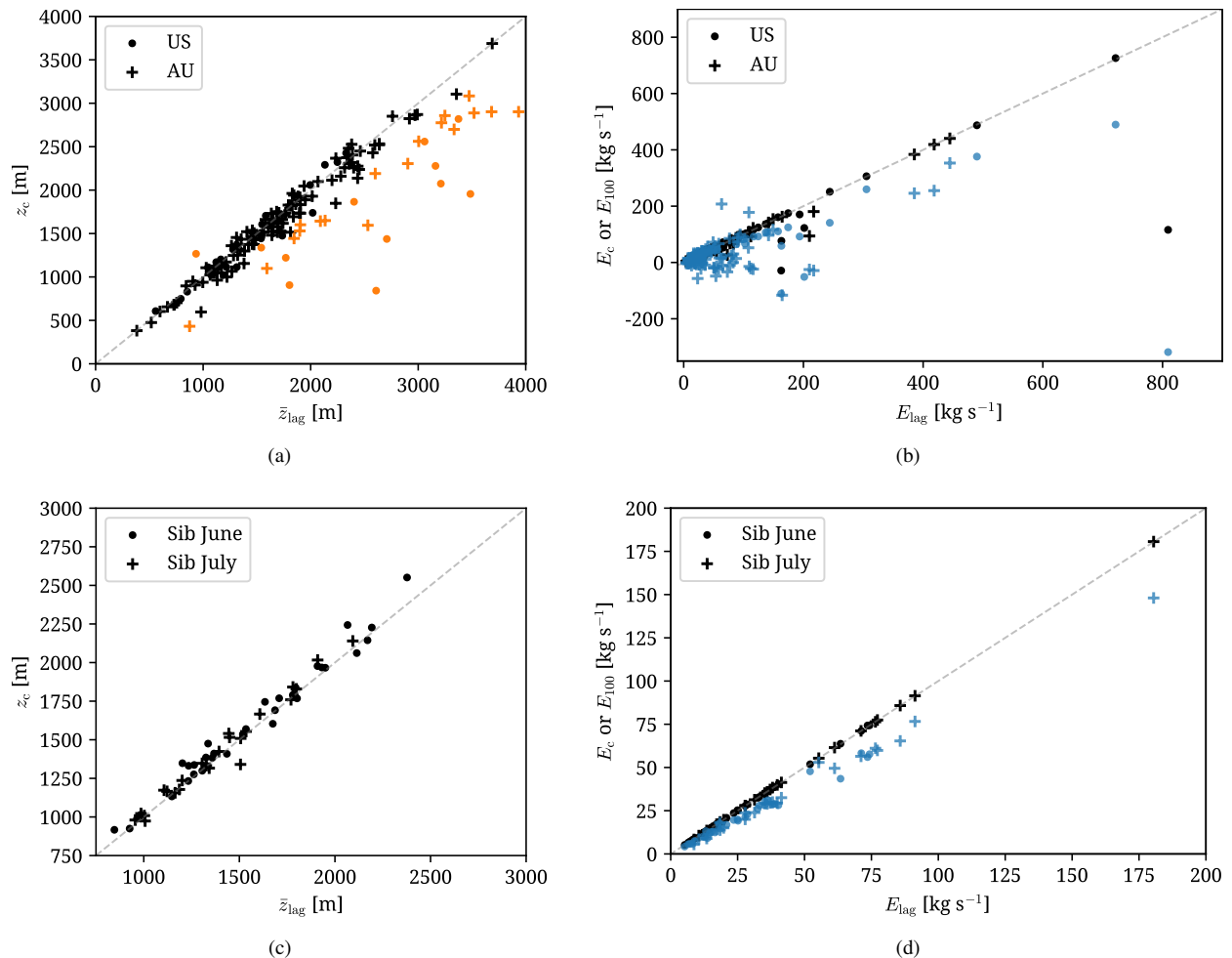
Figure 7. Plume height variation and emissions for regions encapsulating US and Australia (top figures) and Siberia (bottom figures). (Panel a and c) The mean plume height (see Eq. (9)) versus the constant plume height for each fire. The black color represents $\delta z \leq 500$m and orange color indicates $\delta z > 500$m (Panel b and d) Comparison between the emissions computed at plume height z_{lag} versus z_{c} represented by black color and z_{lag} versus a constant plume height of 100 m represented by blue color.

in plume height in downwind direction. Obviously, this plume height variation can influence the emissions due to the change in the velocity with height.

Figure 7(b) and (d) compare the emissions computed from Lagrangian plume height (E_{lag}) with the emissions computed from constant plume height (E_{c}) represented by black color and 100 m plume height (E_{100}) represented by blue color. Combination of all cases in figures 7(b) and (d) show that the E_{c} varied less than 10% from the E_{lag} for a total of 198 cases. And the fires in Siberia varied by less than 4%. However, 23 cases in US and AU had variation larger than 10%. Thus, the overall effect of the Lagrangian plume height to the constant plume height on the emission estimate is considered minor, however we

could identify several cases where the emissions estimate from Lagrangian plume height become more reliable. For example, a US fire (black color) on the bottom right of Figure 7(b) was found to have high $E_{\text{lag}} = 809 \text{ kgs}^{-1}$ and low $E_{\text{c}} = 115.9 \text{ kgs}^{-1}$. The total fire radiative power (FRP) for this case was found to be the highest among all the detected plumes and the burnt area (number of fire counts in VIIRS data) was third highest among all the detected cases. Additionally, the CO enhancement was large, thus, a high emission estimate is expected. Furthermore, the high FRP is correlated with higher temperatures, so an increase in the plume height in downwind direction is normal. It should be noted that the Lagrangian simulations do not consider the heating. However, we assume that the velocities in ERA5 cover this as it assimilates the surface temperature. The increase in plume height is observed in the Lagrangian simulations as the z_{lag} increases by 1350 m compared to constant plume height in the downwind direction at 32.5 km from the fire source. From this, one can conclude that the E_{lag} can be more appropriate than E_{c}. A similar reasoning can be used to explain why the E_{lag} was higher compared to E_{c} for the fires where the FRP on average was higher.

Figures 7(b) and (d) also compare emissions from Lagrangian plume height to a constant 100 m plume height. We considered 100 m plume height as three-dimensional velocity fields, which are required to compute the CO emissions based on the plume heights z_{c} and z_{lag}, amount to a large quantity of data. Furthermore, computing emissions by scaling 100 m winds would simplify the the approach to a large extent. However, we found no correlation between the difference in the emissions ($E_{100} - E_{\text{lag}}$ or E_{c}) and the variation in plume heights. Additionally a total of 37 fires were found to negative values for E_{100} due to a negative velocity at 100 m. This makes it difficult to find an appropriate scaling to obtain emissions at z_{lag} from the velocities at 100 m. Thus, highlighting the importance of using three-dimensional velocity fields rather than surface near wind fields at a fixed altitude. From all these observations, we conclude that the varying plume height is more reliable to compute emissions by an automated algorithm.

It should be noted that the emission estimations in this paper are not compared to any existing database as we are not aware of any in-situ studies except by Rowe et al. (2022) which compares the CO column from aircraft measurements to TROPOMI data.

3.2 Emission uncertainty

We estimate two different contributions to the uncertainty of the estimated emission, where we assumed constant emissions in time for each plume. First, we consider the relative variation of the CO fluxes through the different intersects Q_i. Different error sources may cause this variation and the corresponding error on the flux estimate can be characterized by the standard error σ_{E} in Eq. 4. Second, errors that affect the different fluxes Q_i equally cannot be address by this approach, and so require a dedicated discussion. These errors are TROPOMI bias, ERA5 velocity bias and injection height uncertainty. In addition to these two errors, we verify the emission uncertainty in APE, i.e., the standard error, using data from WRF simulations where the wind velocities, CO and injection height are known.

Table 2. Maximum values of standard error and emission uncertainties due to plume height for different regions among all fires.

Region	$\sigma_{\mathrm{E}}(z_{\mathrm{lag}})\%$	$\sigma_{\mathrm{E}}(z_{\mathrm{c}})\%$	$\Delta E_{\mathrm{lag}}^{p}\%$	$\Delta E_{\mathrm{lag}}^{m}\%$	$\Delta E_{c}^{p}\%$	$\Delta E_{c}^{m}\%$
US	15.11	35.51	71.28	82.27	246.93	163.54
AU	18.79	28.10	94.78	95.41	130.17	170.37
Sib Jun	18.21	18.13	17.34	14.95	17.37	8.84
Sib Jul	19.72	19.57	14.88	13.12	15.47	12.1

3.2.1 Standard errors

The standard error encompass various uncertainty sources, e.g., the interpolation error due to the under-sampling of the CO field by TROPOMI (shown in Fig. 4(c)), relative wind errors with respect to the ERA5 velocity fields, the precision of the TROPOMI CO data, the uncertainty variation in defining the atmospheric CO background per intersect, and the temporal variation of the emission around its mean. σ_{E} does not allow us to disentangle these error sources, expect for the TROPOMI CO precision. This is $< 10\%$ even for dark scenes over land (Landgraf et al., 2016). For the flux estimate, this yields an negligible error contribution. To compare the standard error for different fires, Table. 2 reports the maximum relative standard error for the four regions using the Lagrangian plume height and the constant plume height z_c. For individual fires errors can be accessed in the data base (Goudar et al., 2023). The data shows that the maximum standard error for the Lagrangian plume height is significantly smaller than for z_c for both the US and Australia data. This is another indication to use the Lagrangian plume height as baseline for APE. For the Siberian region there is no difference between the two methods because the plume height does not vary much as depicted in Figure. 7c. Overall, the standard error of the emission estimate is < 20 %.

3.2.2 Other errors

One potential error which cannot be addressed with the standard error is an overall bias in the TROPOMI CO product. Borsdorff et al. (2019b) reported a CO bias of 3.4ppb for the TROPOMI product compared to the Total Carbon Column Observing Network (TCCON). This corresponds to an typical relative error < 1.7 % for a plume concentration of about 200 ppb in a plume. Assuming the CO bias to be constant over the plume, it yields the same relative error on the emission estimate and so can be neglected.

Another error of this category is the emission uncertainty because of the uncertainty in IS4FIRES injection height of ± 500 m Sofiev et al. (2012). For each fire, we calculate the emission uncertainty

$$\Delta E_{\mathrm{lag}}^{p/m} = \left| \frac{E_{\mathrm{lag}}^{p/m} - E_{\mathrm{lag}}}{E_{\mathrm{lag}}} \right| \tag{11}$$

using plume heights z_{lag}^{p} and z_{lag}^{m}, respectively (see Fig 5b). Analogously, the uncertainties $\Delta E_{\mathrm{c}}^{p}$ and $\Delta E_{\mathrm{c}}^{m}$ for E_{c}, are computed and the maximum number per region is again given in Table 2. The uncertainties changes from fire to fire and can be found in the data (Goudar et al., 2023). For the Siberian region, the maximum uncertainties are small indicating little vertical variation in the velocity. For US and Australia region, the corresponding uncertainties are much larger where the uncertainties

Table 3. Comparison of actual emissions to the emissions computed at plume height z_{lag} for the three selected plumes shown in Fig. 8. The uncertainty in the table is computed as 100*(Actual - Computed)/Actual.

Time in UTC (H:M)	Actual (kgs^{-1})	1 km grid (kgs^{-1}); Uncertainty	TROPOMI grid (kgs^{-1}); Uncertainty
17:00	28.45	20.26; 28.8%	20.67; 27.3%
18:00	56.84	34.92; 38.5%	34.52; 37.5%
19:00	97.86	99.15; -1.5%	99.36; -1.53%

for the constant emission height exceeds those using the Lagrangian simulation by a factor of 2-3. This hints at a more variable wind field for these regions. Overall, we estimate this APE error term to be the largest error contribution with error < 100 % for each fire.

380   Finally, we consider errors in the wind velocity that are constant over the plume domain. The error propagate one-to-one into error of the flux estimate. Uncertainties of the ERA5 wind fields in the tropospheric boundary layer are not reported. Gualtieri (2022) derived surface near wind errors of 1.76 m/s (root-mean-square error) for ERA5 data. A typical wind speed at the plume height is 3-11 m/s and although at the plume height the wind speed error might be smaller, we consider this error as a significant error contribution. However, we refrain from quantifying this error because of lack of reliable the knowledge on
385   the wind.

3.2.3   Verification of emission uncertainty

We verify our uncertainty estimates evaluating WRF simulations of a CO plume using APE. The WRF simulation was performed using real atmospheric forcing at 1 km resolution for a fire with the highest FRP (USA, September 12, 2020, see Sec. 3.1). The details on the WRF simulation can be found in the Appendix A. Three plumes at three different UTC times
390   shown in Figure 8a-c were selected and emissions were estimated by our algorithm. It should be noted that the averaging kernels were not used to degrade to TROPOMI data and only the enhancements were simulated in the model, thus the background is set to zero by simulation. The plume height (z_{lag}) was computed as the maximum height where the concentration became zero and fire sources were same as the sources used in WRF simulation. The velocity used in both Lagrangian simulations and emission estimations was WRF velocity data. The emissions for these plumes were estimated by APE assuming a constant
395   emission in time and are presented in Table 3. Here, the actual emission is the mean of the known total CO emission from all fire sources with time. This considered time is based on the release time of the particles around the final transect used to compute emission. Additionally, we degraded the simulation grid to the TROPOMI grid shown in Figs. 8e-g.

The uncertainties of the APE emission estimate ranges between -1.5 % and +38.5% (Table 3). In all three plumes, the velocity and plume height used by APE are appropriate, however the emissions computed by our algorithm differed from the
400   actual emissions. This is attributed to the error in the cross-sectional flux method due to the assumption of a constant emissions which might not be the case for a fire. It should be noted that this uncertainty is for one particular case and it can vary depending
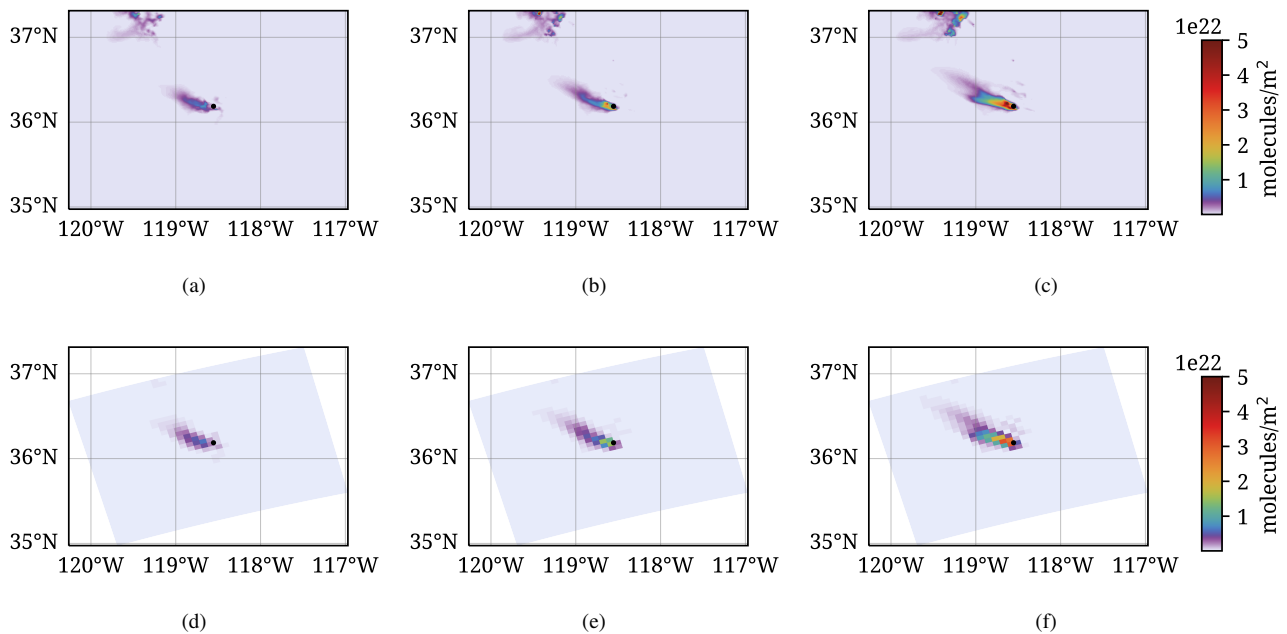
Figure 8. Three selected plumes at three different UTC times (a, d) at 17:00, (b, e) at 18:00 and (c, f) at 19:00. Top (a-c) represent plumes at 1 km grid resolution and bottom (d-f) represent TROPOMI grid resolution.

upon the case. For the three selected cases, the CFM method leads to an error of 28.8, 38.5 and 1.5% which is in the range of the derived standard error. The difference in emissions between high resolution (1 km grid) and low resolution (TROPOMI grid) was found to be less than 2%. If the velocity is accurate, then it can be concluded that having higher resolution data does not have much effect on the cross-sectional flux method. Overall this analysis suggests that the assumption of a constant emission is the major error source next to errors in the wind field and uncertainties in the injection height.

4 Conclusions and recommendations

An automated plume detection and emission estimation scheme for CO flux inversion for single point fires was developed by integrating four freely available data sources, such as, VIIRS active fire dataset, TROPOMI CO dataset, injection height from GFAS and ERA5 meteorological data. The automated plume detection and emission estimation algorithm (APE v1.1) was optimized for one region and its performance is verified for three months data for two other regions, Australia and Siberia. For all the regions, 1327 fires had clear sky TROPOMI CO data of sufficient quality among 5562 identified fire clusters. Even though 882 plumes were identified in 1327 cases, only 378 plumes were considered as 309 plumes were too short and about 195 had multiple sources of fire in them. Lastly, the emissions were estimated for only 226 cases among 378 plumes. We can conclude that APE can reliably detect and estimate emissions automatically for 97.7% of the cases.

The key to automatically detect fire plumes in TROPOMI CO dataset and estimate emissions was knowing the fire source a priori. In the present study, VIIRS active fire data product from Suomi NPP helped in identifying fire source which was used to detect plumes in TROPOMI CO data. This highlights the potential for flying the satellites Suomi NPP and SP5 in the same orbit and in formation with a temporal separation of 3.5 minutes to identify fires.

We considered three different assumption on plume heights, first a constant plume height at 100 m altitude, second a constant plume height at the GFAS injection height and third a varying plume height using a Lagrangian model. The varying plume height was chosen based on the characteristics of a fire. If a fire is at its peak then the air around is heated and rises in the atmosphere and at the same time, it will be transported away from the fire source. Thus, one can observe the CO rise with distance from the source. It should be noted that we assume the ERA5 velocity fields incorporate this heating effect to a certain extent as it assimilates the surface temperature from satellites. In our simulations, the plume height varied by more than 500 m in downwind direction for 30 cases out of 221 cases and all 30 cases were in US and Australia. The plume height variation was found to be minimal in Siberia.

The assumption of plume height at 100 m led to unreliable emission estimates and was discarded. The difference in estimation emission for the constant injection height from GFAS and varying plume height, was observed to be less than 4% for Siberian region. Larger differences we observed for the US and Australia, where the maximum uncertainty using the varying plume height is half of that using a constant plume height. However, for many plumes the difference in the uncertainty estimate was minor. Based on this findings, we decided to use the Lagrangian model for the plume height as processing baseline of APE.

Overall, we estimated the uncertainty of our product with the standard error < 20 %, which mainly accounts for errors in the spatial under-sampling of the CO field by TROPOMI and the assumption of a constant emission for the time frame relevant for the plume formation. The TROPOMI CO data is of high quality regarding precision and bias and does not provide any significant contribution to the CO emission estimate of APE. Additionally, we analyzed emission errors due to the GFAS uncertainty in the injection height. Depending on the meteorological situation at the different regions, errors are < 100 %. Errors due to wrong wind information is also considered to be important but could not be specified as the ERA5 data product does not provide an uncertainty estimate of the provided wind field at the plume height.

Finally, for the first time the presented algorithm is appropriate to estimate CO emissions from fires from TROPOMI/VIIRS data by a fully automated algorithm. It is considered as a baseline for future APE upgrades to optimize automated emission estimates of CO point sources. As a next step, we consider (1) the processing of entire CO TROPOMI data set, (2) expanding emission estimations for multiple fire sources and (3) develop an improved inversion scheme. This can done by developing algorithms that maps the simulated tracer particles from Lagrangian simulations to the TROPOMI CO concentrations to compute emissions, and (4) verifying emissions predicted by APE to available in-situ emissions databases.

Appendix A: The WRF Model description

The WRF model configured in a two-domain configuration is applied in the tracer mode to simulate the transport and dispersion of CO emitted by a wildfire in the US. The outer and inner domains are run at a horizontal grid spacing of 5×5 km^2 and 1×1 km^2, respectively. The model domains are centered at 36.16225^0N, 119.1528^0E and have 43 vertical levels stretching from the surface to a model top of 50 hPa. The outer domain has 200×200 grid points while the inner domain has 400×350 grid points in the west-east and north-south directions. The meteorological initial and boundary conditions for the outer domain are based on the Global Forecast System (GFS) forecasts available every 3 hours at a horizontal grid spacing of $0.25^0 \times 0.25^0$. The static geographical fields and the GFS output are mapped onto the WRF domains using the WRF pre-processing system (WPS). The physical parameterizations follow Kumar et al. (2021) except the cumulus parameterization that is turned off in the inner domain.

Biomass burning emissions are obtained from the Fire Inventory from NCAR (FINN; Wiedinmyer et al., 2011) version 2.5 and are distributed vertically online using a plume rise parameterization developed by Freitas et al. (2007). This parameterization selects fire properties appropriate for the land use in every grid box containing fire emissions and simulates the plume rise explicitly using the environmental conditions simulated by WRF. Since we are using the model in the tracer mode, the chemical evolution of the plume is not simulated. To describe the loss of CO in the model, we allow the CO fire emissions to decay with an e-folding lifetime of 30 days. No other source (anthropogenic emissions, biogenic emissions or photo-chemical production from hydrocarbons) is included in the simulation. The model run started on 12 Sep 2020 at 12 UTC and stopped at 13 Sep 00 UTC. We used a time step of 20 s for the outer domain and 4 s for the inner domain. The model output is saved every min and used for further analysis.

Appendix B: Algorithm and Simulation details

Table B1. Considered region and time. The region is rectangular and is constructed based on the origin and the width and height. The origin is always the south-western point of the region.

Label	Region Origin [(lon, lat)]	Region Size (Width, Height)	Time	L2 product version	VIIRS, ERA5 and GFAS Data access
US	140^0W, 20^0N	80^0, 45^0	Sept 2020	1.03.02	10 Oct 2020
AU	70^0E, 53^0S	55^0, 27^0	Oct 2019	1.03.02	10 Oct 2020
Sib	113^0E, 44^0N	41^0, 34^0	June 2021	1.04.00	5 Feb 2023
Sib	113^0E, 44^0N	41^0, 34^0	July 2021	2.02.00	5 Feb 2023

Code availability. APE v1.1 code is archived on Zenodo (https://doi.org/10.5281/zenodo.7361559).

Algorithm 1 APE algorithm: Pseudo-code

Require: region and time
  for region and time **do**
    Find Fire sources from VIIRS data (Sec. 2.1.1)
    for Each fire source **do**
      Extract TROPOMI CO data granule (Sec. 2.1.2)
      if Data is good **then**
        Detect plume by plume detection algorithm (Sec. 2.2)
        if Plume is detected **then**
          Estimate emission (Sec. 2.3)
        end if
      end if
    end for
  end for

Table B2. Filtering from fire clusters to good CO data. The total column is same as the fire clusters in Table 1.

Region	Grid size	Quality	Multiple Clusters	Good Data	Total
US	442	373	53	213	1081
AU	1020	512	87	385	2013
Sib Jun	37	249		130	416
Sib Jul	34	1419		599	2052
All Regions	1533	2553	140	1327	5562

Table B3. Filtering from good data to plume detection Table 1. The total should represent the good CO data available.

Region	No enhancements	Short plumes	Other clusters	Detected plumes	Total
US	42	41	51	79	213
AU	57	62	94	172	385
Sib Jun	22	25	12	71	130
Sib Jul	324	181	38	56	599
All Regions	445	309	195	378	1327

Data availability. The TROPOMI CO dataset of this study is available for download at ftp://ftp.sron.nl/open-access-data-2/TROPOMI/tropomi/co/ (last access: 5 Feb October 2023). The IS4FIRES injection height and the 3-d velocities at 127 model levels were obtained from the Global Fire Assimilation System (GFAS) database and the European Center for Medium range Weather Forecasts (ECMWF) Reanalysis v5 (ERA5), respectively on 10 October 2020. The Visible Infrared Imaging Radiometer Suite 375m thermal anomalies/active fire product was also accessed on October 2020 (https://firms.modaps.eosdis.nasa.gov/active_fire/). The processed data is available DOI:10.5281/zenodo.7728874.

470

Table B4. Filtering from plume detection to emission estimation Table 1. The total should represent the plume detection cases.

Region	No Injection height	Background Subtraction	Plume alignment	Velocity < 2 ms^{-1}	Emission estimation	Total
US	14	14	11	3	37	79
AU	20	4	13	6	129	172
Sib Jun	15	2	15	4	35	71
Sib Jul	8	9	12	2	25	56
All Regions	57	29	51	15	226	378

Author contributions. MG developed the code and performed analysis with the inputs from TB and JL. JA master thesis served as feasibility study for this work. RK performed WRF simulations and gave inputs on its data analysis. All co-authors commented and improved the paper with a special mention to JL and TB.

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. The National Center for Atmospheric Research is sponsored by the National Science Foundation.

475

References

Andreae, M. O., Browell, E. V., Garstang, M., Gregory, G. L., Harriss, R. C., Hill, G. F., Jacob, D. J., Pereira, M. C., Sachse, G. W., Setzer, A. W., Dias, P. L. S., Talbot, R. W., Torres, A. L., and Wofsy, S. C.: Biomass-burning emissions and associated haze layers over Amazonia, Journal of Geophysical Research, 93, 1509, https://doi.org/10.1029/jd093id02p01509, 1988.

Apituley, A., Pedergnana, M., Sneep, M., Veefkind, J. P., Loyola, D., Landgraf, J., and Borsdorff, T.: Sentinel-5 precursor/TROPOMI Level 2 Product User Manual Carbon Monoxide, User Manual SRON-S5P-LEV2-MA-002, 1.0.0, SRON Netherlands Institute for Space Research, Leiden, The Netherlands, 2018.

Beare, R.: A locally constrained watershed transform, IEEE Transactions on Pattern Analysis and Machine Intelligence, 28, 1063–1074, https://doi.org/10.1109/tpami.2006.132, 2006.

Beirle, S., Boersma, K. F., Platt, U., Lawrence, M. G., and Wagner, T.: Megacity Emissions and Lifetimes of Nitrogen Oxides Probed from Space, Science, 333, 1737–1739, https://doi.org/10.1126/science.1207824, 2011.

Bey, I., Jacob, D. J., Yantosca, R. M., Logan, J. A., Field, B. D., Fiore, A. M., Li, Q., Liu, H. Y., Mickley, L. J., and Schultz, M. G.: Global modeling of tropospheric chemistry with assimilated meteorology: Model description and evaluation, Journal of Geophysical Research: Atmospheres, 106, 23 073–23 095, https://doi.org/10.1029/2001jd000807, 2001.

Borsdorff, T., de Brugh, J. A., Hu, H., Aben, I., Hasekamp, O., and Landgraf, J.: Measuring Carbon Monoxide With TROPOMI: First Results and a Comparison With ECMWF-IFS Analysis Data, Geophysical Research Letters, 45, 2826–2832, https://doi.org/10.1002/2018gl077045, 2018.

Borsdorff, T., aan de Brugh, J., Pandey, S., Hasekamp, O., Aben, I., Houweling, S., and Landgraf, J.: Carbon monoxide air pollution on sub-city scales and along arterial roads detected by the Tropospheric Monitoring Instrument, Atmospheric Chemistry and Physics, 19, 3579–3588, https://doi.org/10.5194/acp-19-3579-2019, 2019a.

Borsdorff, T., aan de Brugh, J., Schneider, A., Lorente, A., Birk, M., Wagner, G., Kivi, R., Hase, F., Feist, D. G., Sussmann, R., Rettinger, M., Wunch, D., Warneke, T., and Landgraf, J.: Improving the TROPOMI CO data product: update of the spectroscopic database and destriping of single orbits, Atmospheric Measurement Techniques, 12, 5443–5455, https://doi.org/10.5194/amt-12-5443-2019, 2019b.

Borsdorff, T., García Reynoso, A., Maldonado, G., Mar-Morales, B., Stremme, W., Grutter, M., and Landgraf, J.: Monitoring CO emissions of the metropolis Mexico City using TROPOMI CO observations, Atmospheric Chemistry and Physics Discussions, pp. 1–21, https://doi.org/10.5194/acp-2020-238, 2020.

Brasseur, G. P. and Jacob, D. J.: Modeling of Atmospheric Chemistry, Cambridge University Press, https://doi.org/10.1017/9781316544754, 2017.

Brunner, D., Kuhlmann, G., Marshall, J., Clément, V., Fuhrer, O., Broquet, G., Löscher, A., and Meijer, Y.: Accounting for the vertical distribution of emissions in atmospheric CO_2 simulations, Atmospheric Chemistry and Physics, 19, 4541–4559, https://doi.org/10.5194/acp-19-4541-2019, 2019.

Butcher, J. C.: Numerical Methods for Ordinary Differential Equations, Wiley, https://doi.org/10.1002/0470868279, 2003.

Cambaliza, M. O. L., Shepson, P. B., Caulton, D. R., Stirm, B., Samarov, D., Gurney, K. R., Turnbull, J., Davis, K. J., Possolo, A., Karion, A., Sweeney, C., Moser, B., Hendricks, A., Lauvaux, T., Mays, K., Whetstone, J., Huang, J., Razlivanov, I., Miles, N. L., and Richardson, S. J.: Assessment of uncertainties of an aircraft-based mass balance approach for quantifying urban greenhouse gas emissions, Atmospheric Chemistry and Physics, 14, 9029–9050, https://doi.org/10.5194/acp-14-9029-2014, 2014.

Cambaliza, M. O. L., Shepson, P. B., Bogner, J., Caulton, D. R., Stirm, B., Sweeney, C., Montzka, S. A., Gurney, K. R., Spokas, K.,
Salmon, O. E., Lavoie, T. N., Hendricks, A., Mays, K., Turnbull, J., Miller, B. R., Lauvaux, T., Davis, K., Karion, A., Moser, B., Miller,
C., Obermeyer, C., Whetstone, J., Prasad, K., Miles, N., and Richardson, S.: Quantification and source apportionment of the methane
emission flux from the city of Indianapolis, Elementa: Science of the Anthropocene, 3, https://doi.org/10.12952/journal.elementa.000037,
2015.

Crippa, M., Guizzardi, D., Muntean, M., Schaaf, E., Dentener, F., van Aardenne, J. A., Monni, S., Doering, U., Olivier, J. G. J., Pagliari, V.,
and Janssens-Maenhout, G.: Gridded emissions of air pollutants for the period 1970–2012 within EDGAR v4.3.2, Earth System Science
Data, 10, 1987–2013, https://doi.org/10.5194/essd-10-1987-2018, 2018.

Ester, M., Kriegel, H.-P., Sander, J., and Xu, X.: A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise,
pp. 226–231, From: KDD-96 Proceedings. Copyright © 1996, AAAI (www.aaai.org). All rights reserved., 1996.

Finch, D., Palmer, P., and Zhang, T.: Automated detection of atmospheric NO_2 plumes from satellite data: a tool to help infer anthropogenic
combustion emissions, Atmospheric Measurement Techniques Discussions, 2021, 1–21, https://doi.org/10.5194/amt-2021-177, 2021.

Fiorio, C. and Gustedt, J.: Two linear time Union-Find strategies for image processing, Theoretical Computer Science, 154, 165–181,
https://doi.org/10.1016/0304-3975(94)00262-2, 1996.

Frankenberg, C., Thorpe, A. K., Thompson, D. R., Hulley, G., Kort, E. A., Vance, N., Borchardt, J., Krings, T., Gerilowski, K.,
Sweeney, C., Conley, S., Bue, B. D., Aubrey, A. D., Hook, S., and Green, R. O.: Airborne methane remote measurements re-
veal heavy-tail flux distribution in Four Corners region, Proceedings of the National Academy of Sciences, 113, 9734–9739,
https://doi.org/10.1073/pnas.1605617113, 2016.

Freitas, S. R., Longo, K. M., Chatfield, R., Latham, D., Dias, M. A. F. S., Andreae, M. O., Prins, E., Santos, J. C., Gielow, R., and Carvalho,
J. A.: Including the sub-grid scale plume rise of vegetation fires in low resolution atmospheric transport models, Atmospheric Chemistry
and Physics, 7, 3385–3398, https://doi.org/10.5194/acp-7-3385-2007, 2007.

Gao, H., Xue, P., and Lin, W.: A new marker-based watershed algorithm, in: 2004 IEEE International Symposium on Circuits and Systems
(IEEE Cat. No.04CH37512), vol. 2, IEEE, IEEE, https://doi.org/10.1109/iscas.2004.1329213, 2004.

Goudar, M., Anema, J., Kumar, R., Borsdorff, T., and Landgraf, J.: APE generated Data, https://doi.org/10.5281/zenodo.7728874, 2023.

Granier, C., Bessagnet, B., Bond, T., D'Angiola, A., van der Gon, H. D., Frost, G. J., Heil, A., Kaiser, J. W., Kinne, S., Klimont, Z., Kloster, S.,
Lamarque, J.-F., Liousse, C., Masui, T., Meleux, F., Mieville, A., Ohara, T., Raut, J.-C., Riahi, K., Schultz, M. G., Smith, S. J., Thompson,
A., van Aardenne, J., van der Werf, G. R., and van Vuuren, D. P.: Evolution of anthropogenic and biomass burning emissions of air
pollutants at global and regional scales during the 1980–2010 period, Climatic Change, 109, 163–190, https://doi.org/10.1007/s10584-
011-0154-1, 2011.

Grell, G. A., Peckham, S. E., Schmitz, R., McKeen, S. A., Frost, G., Skamarock, W. C., and Eder, B.: Fully coupled "online" chemistry
within the WRF model, Atmospheric Environment, 39, 6957–6975, https://doi.org/10.1016/j.atmosenv.2005.04.027, 2005.

Gualtieri, G.: Analysing the uncertainties of reanalysis data used for wind resource assessment: A critical review, Renewable and Sustainable
Energy Reviews, 167, 112 741, https://doi.org/10.1016/j.rser.2022.112741, 2022.

Hersbach, H., Bell, B., Berrisford, P., Biavati, G., Horányi, A., Muñoz Sabater, J., Nicolas, J., Peubey, C., Radu, R., Rozum, I., Schepers, D.,
Simmons, A., Soci, C., Dee, D., and Thépaut, J.-N.: ERA5 model level parameters: instantaneous., Copernicus Climate Change Service
(C3S) Climate Data Store (CDS), accessed on 10-10-2020, 2018.

Hoesly, R. M., Smith, S. J., Feng, L., Klimont, Z., Janssens-Maenhout, G., Pitkanen, T., Seibert, J. J., Vu, L., Andres, R. J., Bolt, R. M., Bond,
T. C., Dawidowski, L., Kholod, N., ichi Kurokawa, J., Li, M., Liu, L., Lu, Z., Moura, M. C. P., O'Rourke, P. R., and Zhang, Q.: Historical

(1750–2014) anthropogenic emissions of reactive gases and aerosols from the Community Emissions Data System (CEDS), Geoscientific Model Development, 11, 369–408, https://doi.org/https://doi.org/10.5194/gmd-11-369-2018, 2018.

Holloway, T., Levy, H., and Kasibhatla, P.: Global distribution of carbon monoxide, Journal of Geophysical Research: Atmospheres, 105, 12 123–12 147, https://doi.org/10.1029/1999jd901173, 2000.

Krings, T., Gerilowski, K., Buchwitz, M., Reuter, M., Tretner, A., Erzinger, J., Heinze, D., Pflüger, U., Burrows, J. P., and Bovensmann, H.: MAMAP – a new spectrometer system for column-averaged methane and carbon dioxide observations from aircraft: retrieval algorithm and first inversions for point source emission rates, Atmospheric Measurement Techniques, 4, 1735–1758, https://doi.org/https://doi.org/10.5194/amt-4-1735-2011, 2011.

Kuhlmann, G., Broquet, G., Marshall, J., Clément, V., Löscher, A., Meijer, Y., and Brunner, D.: Detectability of CO_2 emission plumes of cities and power plants with the Copernicus Anthropogenic CO_2 Monitoring (CO2M) mission, Atmospheric Measurement Techniques, 12, 6695–6719, https://doi.org/https://doi.org/10.5194/amt-12-6695-2019, 2019.

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

Kumar, R., Bhardwaj, P., Pfister, G., Drews, C., Honomichl, S., and D'Attilo, G.: Description and Evaluation of the Fine Particulate Matter Forecasts in the NCAR Regional Air Quality Forecasting System, Atmosphere, 12, 302, https://doi.org/10.3390/atmos12030302, 2021.

Lama, S., Houweling, S., Boersma, K. F., Eskes, H., Aben, I., Denier van der Gon, H. A. C., Krol, M. C., Dolman, H., Borsdorff, T., and Lorente, A.: Quantifying burning efficiency in megacities using the NO_2/CO ratio from the Tropospheric Monitoring Instrument (TROPOMI), Atmospheric Chemistry and Physics, 20, 10 295–10 310, https://doi.org/10.5194/acp-20-10295-2020, 2020.

Landgraf, J., aan de Brugh, J., Scheepmaker, R., Borsdorff, T., Hu, H., Houweling, S., Butz, A., Aben, I., and Hasekamp, O.: Carbon monoxide total column retrievals from TROPOMI shortwave infrared measurements, Atmospheric Measurement Techniques, 9, 4955–4975, https://doi.org/10.5194/amt-9-4955-2016, 2016.

Lee, B., Cho, S., Lee, S.-K., Woo, C., and Park, J.: Development of a Smoke Dispersion Forecast System for Korean Forest Fires, Forests, 10, 219, https://doi.org/10.3390/f10030219, 2019.

Li, F., Zhang, X., Roy, D. P., and Kondragunta, S.: Estimation of biomass-burning emissions by fusing the fire radiative power retrievals from polar-orbiting and geostationary satellites across the conterminous United States, Atmospheric Environment, 211, 274–287, https://doi.org/https://doi.org/10.1016/j.atmosenv.2019.05.017, 2019.

Li, F., Zhang, X., Kondragunta, S., and Lu, X.: An evaluation of advanced baseline imager fire radiative power based wildfire emissions using carbon monoxide observed by the Tropospheric Monitoring Instrument across the conterminous United States, Environmental Research Letters, 15, 094 049, https://doi.org/10.1088/1748-9326/ab9d3a, 2020.

Magro, C., Nunes, L., Gonçalves, O., Neng, N., Nogueira, J., Rego, F., and Vieira, P.: Atmospheric Trends of CO and CH4 from Extreme Wildfires in Portugal Using Sentinel-5P TROPOMI Level-2 Data, Fire, 4, 25, https://doi.org/10.3390/fire4020025, 2021.

Nassar, R., Hill, T. G., McLinden, C. A., Wunch, D., Jones, D. B. A., and Crisp, D.: Quantifying CO_2 Emissions From Individual Power Plants From Space, Geophysical Research Letters, 44, https://doi.org/10.1002/2017gl074702, 2017.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., and Duchesnay, E.: Scikit-learn: Machine Learning in Python, Journal of Machine Learning Research, 12, 2825–2830, 2011.

Rémy, S., Veira, A., Paugam, R., Sofiev, M., Kaiser, J. W., Marenco, F., Burton, S. P., Benedetti, A., Engelen, R. J., Ferrare, R., and Hair, J. W.: Two global data sets of daily fire emission injection heights since 2003, Atmospheric Chemistry and Physics, 17, 2921–2942, https://doi.org/10.5194/acp-17-2921-2017, 2017.

Rowe, J. P., Zarzana, K. J., Kille, N., Borsdorff, T., Goudar, M., Lee, C. F., Koenig, T. K., Romero-Alvarez, J., Campos, T., Knote, C., Theys, N., Landgraf, J., and Volkamer, R.: Carbon Monoxide in Optically Thick Wildfire Smoke: Evaluating TROPOMI Using CU Airborne SOF Column Observations, ACS Earth and Space Chemistry, 6, 1799–1812, https://doi.org/10.1021/acsearthspacechem.2c00048, 2022.

Schneising, O., Buchwitz, M., Reuter, M., Bovensmann, H., and Burrows, J. P.: Severe Californian wildfires in November 2018 observed from space: the carbon monoxide perspective, Atmospheric Chemistry and Physics, 20, 3317–3332, https://doi.org/10.5194/acp-20-3317-2020, 2020.

Schroeder, W., Oliva, P., Giglio, L., and Csiszar, I. A.: The New VIIRS 375 m active fire detection data product: Algorithm description and initial assessment, Remote Sensing of Environment, 143, 85–96, https://doi.org/10.1016/j.rse.2013.12.008, 2014.

Schubert, E., Sander, J., Ester, M., Kriegel, H. P., and Xu, X.: DBSCAN Revisited, Revisited, ACM Transactions on Database Systems, 42, 1–21, https://doi.org/10.1145/3068335, 2017.

Shi, Y., Matsunaga, T., and Yamaguchi, Y.: High-Resolution Mapping of Biomass Burning Emissions in Three Tropical Regions, Environmental Science & Technology, 49, 10 806–10 814, https://doi.org/10.1021/acs.est.5b01598, 2015.

Sobel, I. and Feldman, G. M.: An Isotropic 3×3 image gradient operator, 1990.

Sofiev, M., Ermakova, T., and Vankevich, R.: Evaluation of the smoke-injection height from wild-land fires using remote-sensing data, Atmospheric Chemistry and Physics, 12, 1995–2006, https://doi.org/10.5194/acp-12-1995-2012, 2012.

Spivakovsky, C. M., Logan, J. A., Montzka, S. A., Balkanski, Y. J., Foreman-Fowler, M., Jones, D. B. A., Horowitz, L. W., Fusco, A. C., Brenninkmeijer, C. A. M., Prather, M. J., Wofsy, S. C., and McElroy, M. B.: Three-dimensional climatological distribution of tropospheric OH: Update and evaluation, Journal of Geophysical Research: Atmospheres, 105, 8931–8980, https://doi.org/10.1029/1999jd901006, 2000.

Tian, Y., Sun, Y., Borsdorff, T., Liu, C., Liu, T., Zhu, Y., Yin, H., and Landgraf, J.: Quantifying CO emission rates of industrial point sources from TROPOMI observations, Environmental Research Letters, https://doi.org/10.1088/1748-9326/ac3b1a, 2021.

van der Velde, I. R., van der Werf, G. R., Houweling, S., Maasakkers, J. D., Borsdorff, T., Landgraf, J., Tol, P., van Kempen, T. A., van Hees, R., Hoogeveen, R., Veefkind, J. P., and Aben, I.: Vast CO2 release from Australian fires in 2019–2020 constrained by satellite, Nature, 597, 366–369, https://doi.org/10.1038/s41586-021-03712-y, 2021.

van der Walt, S., Schönberger, J. L., Nunez-Iglesias, J., Boulogne, F., Warner, J. D., Yager, N., Gouillart, E., and Yu, T.: scikit-image: Image processing in Python, PeerJ, 2, e453, https://doi.org/10.7717/peerj.453, 2014.

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

Veefkind, J., Aben, I., McMullan, K., Förster, H., de Vries, J., Otter, G., Claas, J., Eskes, H., de Haan, J., Kleipool, Q., van Weele, M., Hasekamp, O., Hoogeveen, R., Landgraf, J., Snel, R., Tol, P., Ingmann, P., Voors, R., Kruizinga, B., Vink, R., Visser, H., and Levelt, P.: TROPOMI on the ESA Sentinel-5 Precursor: A GMES mission for global observations of the atmospheric composition for climate, air quality and ozone layer applications, Remote Sensing of Environment, 120, 70–83, https://doi.org/https://doi.org/10.1016/j.rse.2011.09.027, 2012.

Watson, C. E., Fishman, J., and Reichle, H. G.: The significance of biomass burning as a source of carbon monoxide and ozone in the southern hemisphere tropics: A satellite analysis, Journal of Geophysical Research, 95, 16 443, https://doi.org/https://doi.org/10.1029/JD095iD10p16443, 1990.

White, W. H., Anderson, J. A., Blumenthal, D. L., Husar, R. B., Gillani, N. V., Husar, J. D., and Wilson, W. E.: Formation and Transport of Secondary Air Pollutants: Ozone and Aerosols in the St. Louis Urban Plume, Science, 194, 187–189, https://doi.org/10.1126/science.959846, 1976.

Wiedinmyer, C., Akagi, S. K., Yokelson, R. J., Emmons, L. K., Al-Saadi, J. A., Orlando, J. J., and Soja, A. J.: The Fire INventory from NCAR (FINN): a high resolution global model to estimate the emissions from open burning, Geoscientific Model Development, 4, 625–641, https://doi.org/10.5194/gmd-4-625-2011, 2011.

630