Geosci. Model Dev., 18, 3607–3622, 2025
https://doi.org/10.5194/gmd-18-3607-2025

Geoscientific Model Development
Open Access

Quantification of CO_2 hotspot emissions from OCO-3 SAM CO_2 satellite images using deep learning methods

Joffrey Dumont Le Brazidec[1,4], **Pierre Vanderbecken**[1], **Alban Farchi**[1], **Grégoire Broquet**[2], **Gerrit Kuhlmann**[3], and **Marc Bocquet**[1]

[1]CEREA, École des Ponts, EDF R&D, Île-de-France, France
[2]Laboratoire des Sciences du Climat et de l'Environnement, LSCE/IPSL, CEA-CNRS-UVSQ, Université Paris-Saclay, Gif-sur-Yvette, France
[3]Laboratory for Air Pollution/Environmental Technology, Empa, Dübendorf, Switzerland
[4]currently at: European Centre for Medium-Range Weather Forecasts (ECMWF), Bonn, Germany

Correspondence: Joffrey Dumont Le Brazidec (joffrey.dumont@ecmwf.int)

Received: 24 August 2024 – Discussion started: 8 October 2024
Revised: 16 February 2025 – Accepted: 13 March 2025 – Published: 18 June 2025

Abstract. This paper presents the development and application of a deep-learning-based method for inverting CO_2 atmospheric plumes from power plants using satellite imagery of the CO_2 total column mixing ratios (XCO_2). We present an end-to-end convolutional neural network (CNN) approach, processing the satellite XCO_2 images to derive estimates of the power plant emissions, that is resilient to missing data in the images due to clouds or to the partial view of the plume owing to the limited extent of the satellite swath.

The CNN is trained and validated exclusively on CO_2 simulations from eight power plants in Germany in 2015. The evaluation on this synthetic dataset shows an excellent CNN performance with relative errors close to 20 %, which is only significantly affected by substantial cloud cover. The method is then applied to 39 images of the XCO_2 plumes from nine power plants, acquired by the Orbiting Carbon Observatory-3 Snapshot Area Maps (OCO3 SAMs), and the predictions are compared to average annual reported emissions. The results are very promising, showing a relative difference in the predictions to reported emissions only slightly higher than the relative error diagnosed from the experiments with synthetic images. Furthermore, analysis of the area of the images in which the CNN-based inversion extracts the information for the quantification of the emissions, based on integrated-gradient techniques, demonstrates that the CNN effectively identifies the location of the plumes in the OCO-3 SAM images. This study demonstrates the feasibility of applying neural networks that have been trained on synthetic datasets for the inversion of atmospheric plumes in real satellite imagery from XCO_2 and provides the tools for future applications.

1 Introduction

Satellite imagery of the total column average dry-air mole fraction of carbon dioxide (XCO_2) from the Snapshot Area Map (SAM) mode of the Orbiting Carbon Observatory-3 (OCO-3) (Eldering et al., 2019) or the forthcoming CO2M mission (Janssens-Maenhout et al., 2020; Meijer et al., 2023) is pivotal for monitoring carbon dioxide (CO_2) emissions. In the vicinity of large CO_2 anthropogenic sources, such as power plants, satellite images may include CO_2 atmospheric plumes emanating from these sources. From these images, atmospheric inversion approaches can estimate the CO_2 emissions of the sources by analysing the signal intensity of the detected plumes (Nassar et al., 2017; Reuter et al., 2019; Chevallier et al., 2019; Wu et al., 2020; Zheng et al., 2020; Nassar et al., 2022; Chevallier et al., 2022; Cusworth et al., 2021).

Various approaches can be used to determine the emissions underlying the XCO_2 plumes in the satellite imagery. The first technique category relies on traditional atmospheric inversion methods that minimize the misfits between the satellite observations and simulations of the plumes with relatively expensive (Eulerian or Lagrangian) transport models to identify the optimal emission estimate (Pillai et al., 2016;

Broquet et al., 2018). The second technique category includes light-weight methods that apply the principle of mass conservation to compute the emissions from the CO_2 enhancement of the emission plume (such as integrated mass enhancements, divergence methods, and cross-sectional flux methods) or compare the observed plume with a Gaussian plume model (Gaussian plume inversions). Light-weight methods rely on wind fields taken, for example, from meteorological reanalysis products. These light-weight methods have been evaluated in several studies (e.g. Varon et al., 2018; Hakkarainen et al., 2024; Danjou et al., 2025; Santaren et al., 2025; Kuhlmann et al., 2024; Danjou et al., 2024). Despite the advancements in CO_2 plume inversion techniques, significant challenges remain, notably, (1) the extraction of plumes from XCO_2 backgrounds, which is hindered by low signal-to-noise ratios due to the large amplitude of background variations associated with the CO_2 natural fluxes and to relatively high noise in the image (due to instrumental errors and to uncertainties in the retrieval of mole fractions from the satellite measurements); (2) the complex process of deducing the source emissions from clearly delineated plumes, which is marred by uncertainties in the corresponding transport and dispersion (i.e. in either the transport modelling or in the wind field and assumptions regarding the vertical structure of the 3D CO_2 plume for the derivation of the effective wind driving the 2D XCO_2 plume in the light-weight analysis; Dumont Le Brazidec et al., 2023); and (3) the reconstruction of emissions from images with a partial view of the plumes due to missing data where there are clouds or gaps in satellite coverage.

Machine learning models have been suggested in response to these obstacles and have been primarily applied to CH_4 and NO_2 images (e.g. Lary et al., 2016; Finch et al., 2022; Jongaramrungruang et al., 2021; Joyce et al., 2023; Kumar et al., 2023). Our previous work (Dumont Le Brazidec et al., 2023, 2024a) pioneered the use of deep learning methodologies, specifically convolutional neural networks (CNNs), for the segmentation and inversion of CO_2 plumes for the estimate of point sources. This approach has demonstrated its efficacy in addressing these challenges when tackling synthetic satellite images with a full coverage of the plumes, i.e. without the loss of observations due to cloud cover or quality control in the limited satellite field of view. This paper is a direct continuation of Dumont Le Brazidec et al. (2024a). Specifically, our approach involves developing a supervised-learning CNN system designed to predict CO_2 emissions using XCO_2 images and ancillary data (such as wind fields, time, and NO_2 images which will be measured by CO2M). This CNN is trained on a synthetic dataset, constructed from model simulations, comprising synthetic XCO_2 fields and the corresponding true emissions. Through this training process, the CNN learns to correlate specific features within the input images covering the plume from a targeted point source with certain output values, namely, the emissions from the point source. The CNN's capability to generalize is subsequently assessed using a new, unseen dataset during the training phase. In particular, this assessment is based on tests targeting a source that was not covered by the synthetic images used for the training phase.

Our previous research (Dumont Le Brazidec et al., 2023) evaluated the models using only synthetic images without missing data, comparing them against light-weight alternative methods for which they demonstrated better performance, with an absolute error about half that of the cross-sectional flux method. In the current study, we extend our approach by analysing actual satellite data, specifically examining 39 OCO-3 SAM observations to quantify emissions. These images encompass 64 km^2 and cover nine power plants located in the USA (seven images), Europe (one image), and China (one image). To make this possible, this paper introduces a new upgrade of the CNN approach to address the third principal challenge in CO_2 plume inversion: handling images with a partial cover of the plumes due to the loss of observations associated with clouds or due to the limited extent of the satellite swath. Furthermore, the training of the CNN involves a novel data augmentation strategy, specifically the incorporation of beta or uniform distribution mappings for plumes and the corresponding emissions. This enhancement aims to improve the robustness and stability of the CNN with respect to predicting CO_2 emissions under various conditions.

The structure of this paper is as follows: Sect. 2 introduces the synthetic dataset, which bears a significant resemblance to that described by Dumont Le Brazidec et al. (2024a) and Santaren et al. (2025), the OCO-3 SAMs utilized exclusively for evaluation, and the dataset's training–validation–test split strategy. Section 3 details the model, the developed data augmentation approach aimed at stabilizing CNN training, the methodology for addressing the problem of clouds, and the training parameterization. In Sect. 4, we successively present the CNN's emission estimations for plumes across the synthetic and OCO-3 SAM datasets. Special attention is given to analysing the model's OCO-3 SAM predictions through the lens of integrated gradients, a method that elucidates the contribution of each input feature to the model's predictions, enhancing interpretability. Finally, Sect. 5 discusses cogent future directions, before we conclude.

2 Dataset

2.1 Synthetic dataset

The synthetic dataset employed in this study is very similar to that used by Dumont Le Brazidec et al. (2024a). The dataset consists of hourly XCO_2 and NO_2 fields from the SMARTCARB project (Brunner et al., 2019; Kuhlmann et al., 2019), which generated 1 year of synthetic CO2M observations from high-resolution CO_2 and NO_2 transport simulations covering power plants in Germany, Poland, and the Czech Republic. The SMARTCARB dataset has been used in various studies for assessing emission quantification methods (e.g. Kuhlmann et al., 2020a, 2021; Hakkarainen et al., 2021; Santaren et al., 2025). For this study, we extracted 32 pixel × 32 pixel (2 km resolution) fields centred on different power plants. For comparison, in Dumont Le Brazidec et al. (2024a), the image size was chosen as 64 pixel × 64 pixel. The transition to focusing the analysis on a more confined area surrounding the hotspots (power plants) is driven by several factors: (i) the critical portion of the plume influencing emission reconstruction typically lies within this central area, as noted in Dumont Le Brazidec et al. (2024a); (ii) satellite swath limitations – the limited spatial extent of a swath and temporal constraints between two swaths makes it unlikely that satellite imaging will consistently capture 128 km^2 areas centred over emission sources; and (iii) this more focused approach demonstrates a stabilizing effect on neural network training, likely due to the reduction in superfluous information.

To account for the inherent noise associated with satellite instruments, we introduce Gaussian random noise with a standard deviation of 0.7 ppm to the XCO_2 images, reflecting typical noise levels expected for OCO-3 and CO_2M snapshots as reported by Meijer (2020), Taylor et al. (2023), and Danjou et al. (2024). Given the observed strong correlation between NO_2 and CO_2 plumes and CO_2M's capability to measure NO_2, we incorporate noisy NO_2 fields in our analysis, characterized by Gaussian noise with a variance of 1×10^{15} molec. cm^{-2} (Kuhlmann et al., 2019).

Similarly to Dumont Le Brazidec et al. (2024a), we integrate ERA5 wind data as additional input to the CNN model, aligning their original resolution of 28 km with the 2 km resolution used for the CO_2 and NO_2 images. Specifically, we employ 2D u and v wind fields, representing the average zonal and meridional winds, respectively, across the five lowest model levels of ERA5. This averaging process approximates the atmospheric conditions below 100 m.

To include the impact of cloud cover in the inversion of CO_2 plumes, we use the simulated cloud cover fractions extracted from the SMARTCARB dataset to mask pixels where retrievals are not available due to the high cloud fraction. Following Kuhlmann et al. (2019), we use a cloud threshold of 1 % for CO_2 images and 30 % for NO_2 images.

Moreover, we study the impact of introducing temporal information to our CNN inputs, by incorporating the hour of the day, day of the week, and day of the year. To capture the cyclical nature of time, these features are transformed into cosine and sine representations, ensuring proximity between temporally adjacent data points (e.g. the last and first hours of the day). Consequently, each XCO_2 field is associated with a vector of six scalar values encoding the temporal context of the observation. In Fig. 1, we present typical input data used by the CNN to predict the emissions of the local hotspot.

2.2 Set of OCO-3 SAMs

The OCO-3 SAM mode is an observation strategy designed to monitor CO_2 emissions from specific emission hotspots (large urban areas and large industrial point sources). Unlike OCO-3's standard observation mode, which conducts continuous scans of the Earth's atmosphere in nadir or glint mode, the SAM mode is a targeting mode that provides high-resolution XCO_2 images around such emission hotspots. In this study, we selected 39 OCO-3 SAMs at nine power plants to evaluate the applicability and reliability of our CNN model trained on synthetic datasets. We selected OCO-3 SAM images corresponding to (1) power plants for which reports of the emissions are available and have been studied in the scientific literature and (2) a sufficient number of cloud-free XCO_2 retrievals of good quality are available. The list of power plants selected are described, including the average reported emissions and number of collected SAMs, in Table 1.

To adapt the raw OCO-3 SAM data for CNN analysis, we first construct a 32 pixel × 32 pixel grid with a resolution of 2 km (similar to the resolution of OCO3 SAMs or CO2M) centred on the power plant. Each grid cell is populated through a weighted interpolation of surrounding OCO-3 SAM data pixel centres, considering only those within a distance of less than 0.66 times the new grid resolution. This specific distance threshold was determined through experimentation to optimally preserve information from the original dataset. Although this mapping strategy provides a straightforward means of converting OCO-3 SAM data into a format compatible with our CNN, it is acknowledged that this approach has limitations and that the observation information might not be perfectly conserved. Additionally, as most OCO3 SAMs used in this study were taken in 2021 or 2022, the synthetic images are adjusted to account for the effect of climate change and the general increase in the CO_2 concentration of 2.3 ppm yr^{-1} since 2015 (SMARTCARB synthetic dataset year). Using eight examples, Fig. 2 illustrates the process of transforming original OCO-3 SAM data into an XCO_2 field suitable for CNN reconstruction.

2.3 Training, validation, and test split choices

For tests on synthetic and real data, to avoid data leakage, a rigorous geographical separation is maintained between the power plants used in the training and validation datasets and those used in the test dataset. For instance, when training a model to predict emissions from the Boxberg power plant, Boxberg plumes are excluded from the training set. The validation dataset comprises plumes from a different power plant, Dolna Odra, which is neither used to train nor test the CNNs. This splitting strategy is outlined in Fig. 3.

This approach mirrors the strategy adopted in Dumont Le Brazidec et al. (2024a) for the analysis of synthetic images. We focus on the same three power plants for the tests – Lippendorf, Boxberg, and Turów – and train distinct mod-

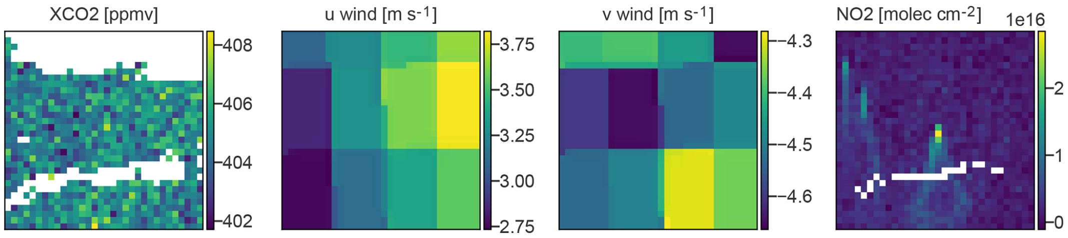

Figure 1. Examples of inputs used by the CNN model. The first, second, third, and fourth columns represent the XCO_2 images, vertically averaged u winds, vertically averaged v winds, and NO_2 images, respectively. The power plant of interest is always located in the middle of the image.

Table 1. List of power plants selected for this study, along with their annual reported emissions, coordinates, and number of observations used. The data span from 2020 to 2023. Emission statistics are sourced from Nassar et al. (2021), Grant et al. (2021), and Lin et al. (2023), which are based on the US Environmental Protection Agency (EPA, https://www.epa.gov/airmarkets/power-sector-emissions-data, last access: 15 June 2025) and the European Pollutant Release and Transfer Register (E-PRTR, https://www.eea.europa.eu/en, last access: 15 June 2025).

Power plant	Coordinates (lat, long)	Reported emissions ($Mt\,CO_2\,yr^{-1}$)	Number of OCO-3 images
Colstrip	45.88, −106.61	13.6	4
Bełchatów	51.27, 19.33	37.6	4
Tuoketuo	40.20, 111.36	29.5	8
Cumberland	36.39, −87.65	12.4	3
Labadie	38.56, −90.84	15.0	6
Intermountain	39.51, −112.58	5.0	2
Hunters	39.17, −111.03	7.2	7
Parish	29.48, −95.63	13.2	4
Conemaugh	40.38, −79.06	16.9	1

els for each to predict their emissions. These models share the same architectural framework, hyperparameters, CNN structure, and preprocessing layers, but they are trained on a dataset excluding plumes from the target power plant.

The rationale behind selecting Lippendorf, Boxberg, and Turów as test power plants is thoroughly discussed by Dumont Le Brazidec et al. (2024a). Briefly, these power plants were chosen for their distinct characteristics: Lippendorf's average emissions are equal to 15.2 $Mt\,CO_2\,yr^{-1}$; Boxberg's plume is often located close to other power plant plumes and its average emission amount to 19.0 $Mt\,CO_2\,yr^{-1}$; and Turów is characterized by low emissions of 8.7 $Mt\,CO_2\,yr^{-1}$. These selection criteria ensure an evaluation of the proposed CNN architecture across various emission scenarios.

It is critical to underline that while the test dataset for one experiment becomes part of the training dataset for another, each experiment was conducted independently, ensuring that model tuning was not optimized by outcomes derived from the test datasets. Finally, in our assessment of CNNs against OCO-3 SAM data, the training was based exclusively on synthetic data.

3 Deep learning methodology

The goal of this study is to determine the CO_2 emission rate (in $MtCO_2\,yr^{-1}$) of the hotspot in the centre of a XCO_2 image using a CNN model, which takes the XCO_2 image (alongside other data) as input. This section describes the CNN model and the data augmentation strategy, with a particular focus on the method to address cloud interference, and discusses training parameters.

3.1 CNN model and preprocessing layers

This subsection describes the CNN-based inversion system (the CNN model with its preprocessing layers that estimates emissions from images) and how it is trained. The CNN-based inversion system is a compound of preprocessing layers and a core CNN model. Preprocessing layers are operations successively applied to the XCO_2 fields and ancillary data before they are processed by the core CNN model. The core CNN model is a statistical model whose parameters (or neurons) are optimized during the training phase; its function is to identify and extract features from the input data, which it learns to associate with specific levels of emissions. The training phase of the CNN-based inversion system consists

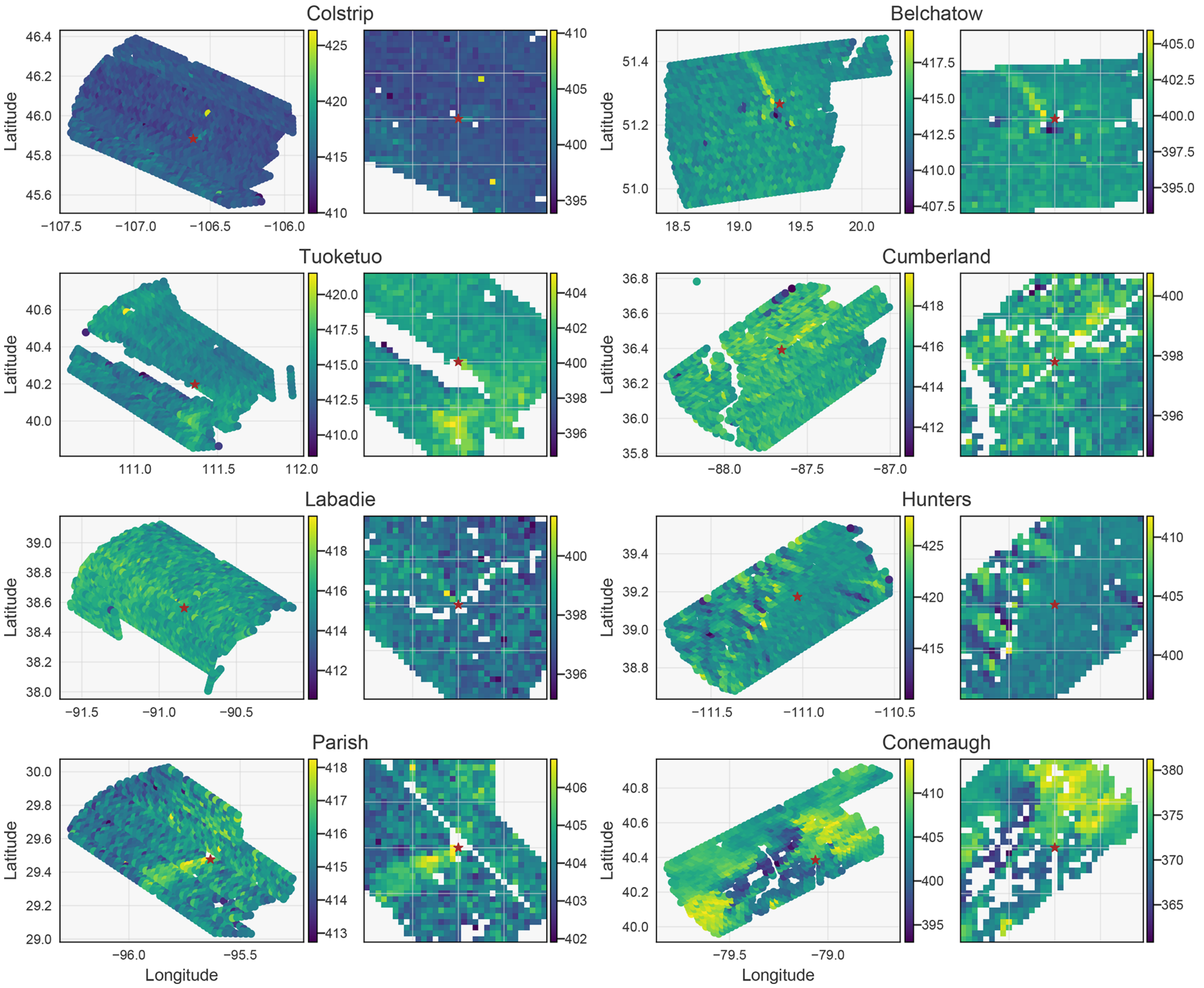

Figure 2. Examples of OCO-3 SAM observations for eight different power plants and the transformation of these observations into CNN-compatible images. For each case, the original OCO-3 SAM data are described on the left, whereas the corresponding CNN-compatible mapping (in which the power plant is always in the centre of the image) is shown on the right. All values are in parts per million by volume (ppmv).

of a series of five steps, depicted in Fig. 4 and outlined in the following paragraphs.

3.1.1 Data augmentation

The data augmentation process creates an artificially infinite dataset from the SMARTCARB dataset to prevent the model from overfitting due to the SMARTCARB dataset's limitations. Specifically, instead of using the XCO_2 field directly from the SMARTCARB dataset to train the core model, we use a composition of five different elements:

1. The principal component is a synthetic image centred on a major power plant of interest, exclusively containing the XCO_2 plume from that facility and the other major power plants. The SMARTCARB dataset composition facilitates isolating this field from all other anthropogenic and biogenic fluxes. This first component undergoes a distribution mapping, whereby an emission level is randomly drawn from a probability distribution, either a beta (to mitigate the training on extreme emissions) or a uniform distribution, as shown in Fig. 5, and the plume image is adjusted accordingly (we trained separate CNN models for each distribution choice). Simultaneously, the CNN output is also adjusted at the emission level.

2. A randomly drawn XCO_2 background, which is augmented by summing it with a random number $b \sim$

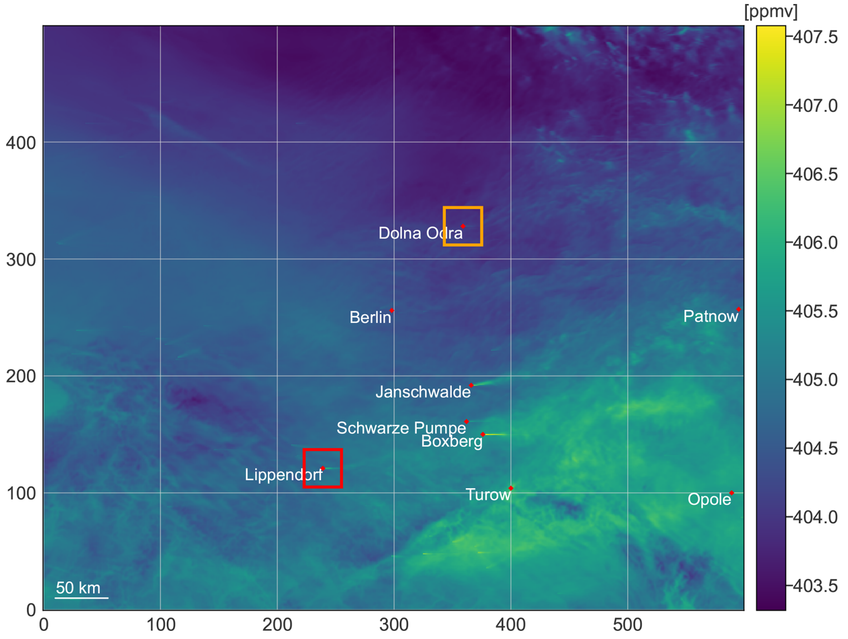

Figure 3. Map of XCO_2 concentrations within the complete SMARTCARB domain on 12 December 2015 at 03:00 UTC. The depicted XCO_2 fields are devoid of synthetic satellite noise for visibility. When constructing a model to predict emissions from Lippendorf, based on Lippendorf-centred fields (indicated by the red square), images from Dolna Odra (indicated by the orange square) are used for validation, whereas images from the remaining power plants serve as training data.

$U(-3.5, 3.5)$ (in ppmv), is added, uniformly across the field, to this first component in a manner analogous to Dumont Le Brazidec et al. (2024a). The selection of the background (and all subsequently described elements) through uniform random drawing is independent of the position of the main plume component.

3. Other anthropogenic XCO_2 plumes identified in the SMARTCARB area, each scaled by a random factor ranging from 0.33 to 3, are added to this.
4. The application of cloud cover constitutes the fourth component. A random selection of cloud cover from the SMARTCARB area is made, independent of the selections for other fields. The XCO_2 pixels are deemed unobserved when cloud cover exceeds 0.01, leading to replacement with NaN, and are subsequently replaced with the minimum value across all XCO_2 fields. This is done in order for the CNN model to learn to ignore this non-informative constant value.
5. The fifth component adds random Gaussian noise with a variance of 0.7 ppmv to the other fields.

3.1.2 Concatenation

The concatenation of the main XCO_2 field with the ancillary data represents the second step. The ancillary data may include wind conditions, the time and date of the observation, and the NO_2 field. In instances where the NO_2 field is incorporated, it also undergoes a data augmentation process (not depicted in Fig. 4). Initially, the NO_2 plume is scaled by a random factor drawn from a uniform distribution ranging from 0.75 to 2 to ensure that the NO_2 plume amplitude is decorrelated from that of the CO_2 plume, thereby preventing the core CNN model from relying on the tight correlation between the NO_x and CO_2 emissions for the inversion. In principle, due to the large variations and uncertainties in the CO_2-to-NO_x emission ratios and the lifetime of NO_x, NO_2 should primarily support the plume detection in the overall inversion process. Subsequently, the NO_2 field is partially masked due to cloud cover. For this, we adopt the criterion from Kuhlmann et al. (2019): an NO_2 pixel is marked as NaN if its cloud cover fraction exceeds 0.3. Furthermore, the NO_2 field is subject to Gaussian noise with a variance of 1×10^{15} molec. cm^{-2} (Kuhlmann et al., 2019).

3.1.3 Normalization

Z-score normalization of each physical field within the concatenated input data constitutes the third step and is performed independently for each channel.

3.1.4 Processing

The fourth step is the core CNN model mapping from XCO_2 and ancillary fields to a scalar emission value. This model, consistent with that described in Dumont Le Brazidec et al. (2024a), features a series of convolutional, max pooling,

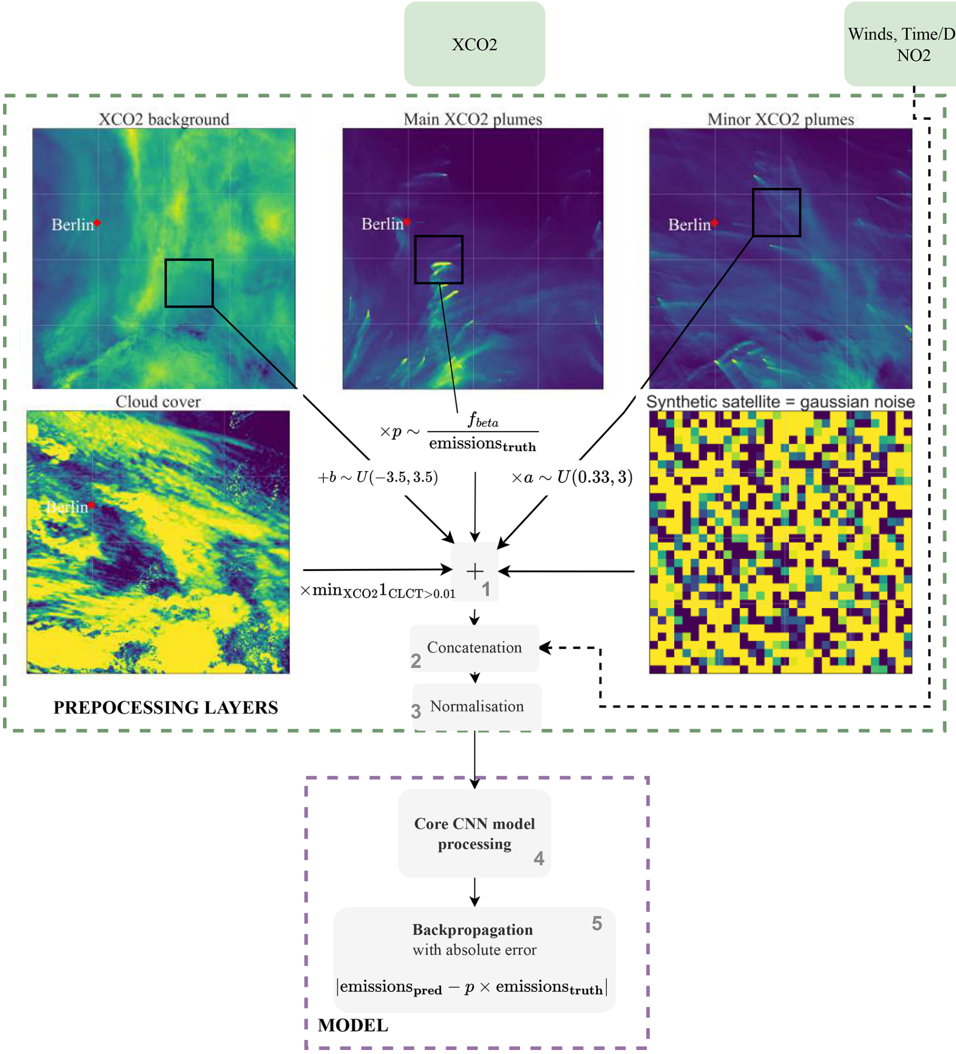

Figure 4. Description of the inversion system at the training time as a compound of preprocessing layers and the model. The CNN-based inversion system consists of five steps. First, in step (1), an XCO_2 field is constructed as the sum of the background XCO_2, major and minor CO_2 plumes, cloud cover, and synthetic satellite noise, with the latter represented by Gaussian noise with a standard deviation of 0.7 ppmv and a 0 mean. The major plumes are scaled using either a beta or uniform distribution (with only the beta distribution shown in the figure), a uniform random field of between −3.5 and 3.5 ppmv is added to the background, and the remaining plumes are scaled by random factors drawn from a uniform distribution of between 0.33 and 3. Pixels in the reconstructed XCO_2 field are marked as missing (NaN) when cloud cover exceeds a specified threshold. The plumes of interest are augmented by a beta or uniform distribution (only beta is represented in the figure), the background is added to uniformly drawn fields of between −3.5 and 3.5 ppmv, and the other plumes are multiplied by a uniform distribution of between 0.33 and 3. The reconstructed XCO_2 field pixels are considered to be NaN or not according to the clouds based on a given threshold. In step (2), concatenation with ancillary data (winds, time, and NO_2) is undertaken. Step (3) involves the standardization of the fields. Step (4) entails processing by the core CNN model. Finally, step (5) involves backpropagation.

batch normalization, and dropout layers, with a total of 186 000 trainable parameters. Specifically, if time and date features are used, they are integrated into the CNN following the feature extraction (following the last dense layer).

3.1.5 Backpropagation

The final step entails computing the loss gradient, enabling neuron adjustments within the core CNN model through backpropagation.

In contrast to the training phase, the inversion system in the evaluation phase consists only of the concatenation, normalization, and processing by the core CNN model. The synthetic test dataset consists of preconstructed, physically

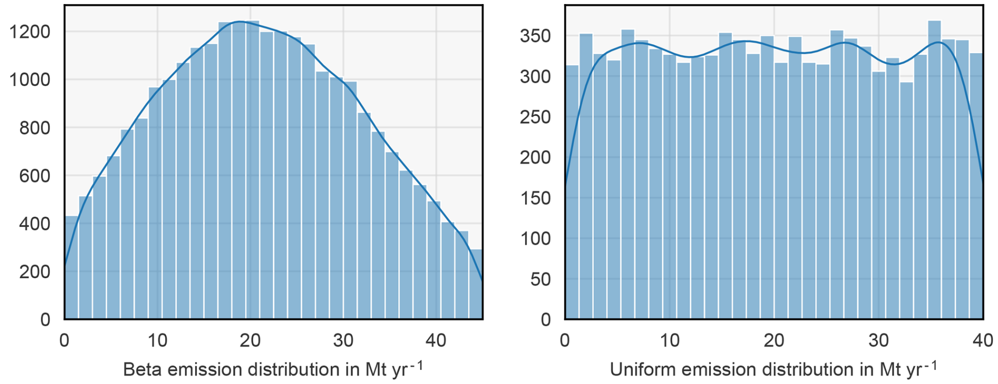

Figure 5. At training time, the hotspot emission and corresponding plume are adjusted based on a random draw from either a beta or uniform distribution.

consistent simulated data (except for clouds, as explained in Sect. 3.2), maintaining consistency with the methodology outlined in Dumont Le Brazidec et al. (2024a).

3.2 Clouds

To assess the impact of cloud cover on CNN performance, we consider models trained and tested on varied datasets distinguished by varying degrees of fraction of cloudy pixels in the XCO_2 images:

- A first series of models are trained and tested on XCO_2 images under clear-sky conditions.
- A second series of models are trained and tested on XCO_2 images with cloud cover ranging from 0 % to 25 %.
- A third series of models are trained on XCO_2 images with cloud cover from 0 % to 50 % but are tested on images with cloud cover from 25 % to 50 %.
- A final series of models are trained on XCO_2 images with cloud cover from 0 % to 75 % but are tested on images with cloud cover from 50 % to 75 %.

These varying degrees of cloud cover are constructed through random sampling of cloud cover over the SMARTCARB domain. This method of training and testing models under varying cloud conditions allows us to compare the degradation of model performance with increased cloud cover. Additionally, training the model tested on cloud cover between 50 % and 75 % on a range from 0 % to 75 % ensures the maintenance of a "universal" model capable of inverting plumes in scenarios with both low and high cloud cover.

3.3 Training parameterization

We configure the training hyperparameters as follows: the model uses the Adam optimizer, with an initial learning rate of 1×10^{-3}, which is adjusted according to a reduce-on-plateau strategy down to 1×10^{-5} with a patience parameter set to 20. The batch size is established at 128, and the training process spans 750 epochs. These parameters were selected based on a rigorous experimental process, combined with adherence to established practices in the field. For the loss function, the mean absolute error (MAE) was chosen.

4 Application to synthetic and OCO-3 SAM observations

4.1 Application to synthetic dataset

Similarly to Dumont Le Brazidec et al. (2024a), we investigate the performance of various CNN models with respect to predicting the emissions of the Lippendorf, Turów, or Boxberg power plants. A collection of CNNs undergo training on subsets of power plants, each excluding one for evaluation. For each power plant, the collection corresponds to models that are trained and tested on images affected by varying levels of cloud cover. In addition, the models are trained with two different input configurations: one that includes XCO_2, wind, time, and NO_2 data and another that includes all of these variables except for NO_2. As a result, a total of 3 (number of target power plants) × 4 (cloud cover scenarios) × 2 (input configurations) = 24 CNNs are trained and evaluated.

Figures 6 and 7 show kernel density estimation (KDE) plots for the absolute relative error and the algebraic difference between the model predictions and the true emissions for the configuration without and with NO_2 input. A comprehensive summary of the results is also provided in Table 2.

The novel data augmentation strategy presented in Sect. 3.1 improves the stability of the performance of the CNNs in comparison to Dumont Le Brazidec et al. (2024a), making it unnecessary to train a CNN ensemble to achieve satisfactory and consistent results. Significantly, in comparison to Dumont Le Brazidec et al. (2024a), the Boxberg

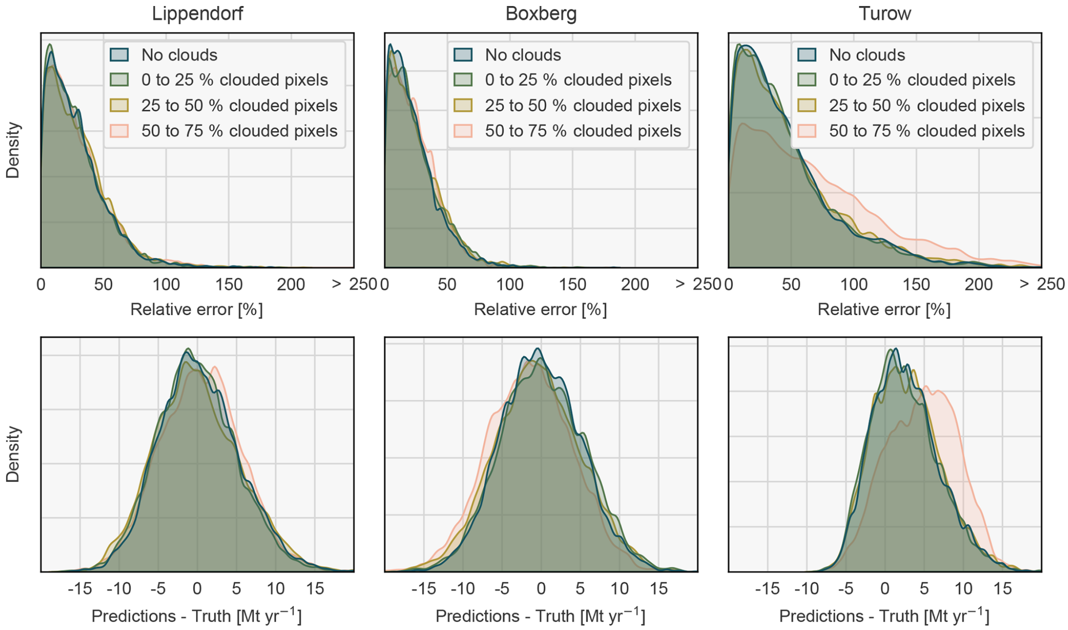

Figure 6. Density plots of the absolute relative error and of the algebraic difference between the CNN predictions and the reported emissions. The CNN models are trained and evaluated with XCO_2, wind, and time input data, affected by varying levels of cloud cover. Predictions with absolute relative errors greater than 250 % or absolute errors greater than 15 Mt yr^{-1} were set to 250 or 15, respectively, to increase visibility.

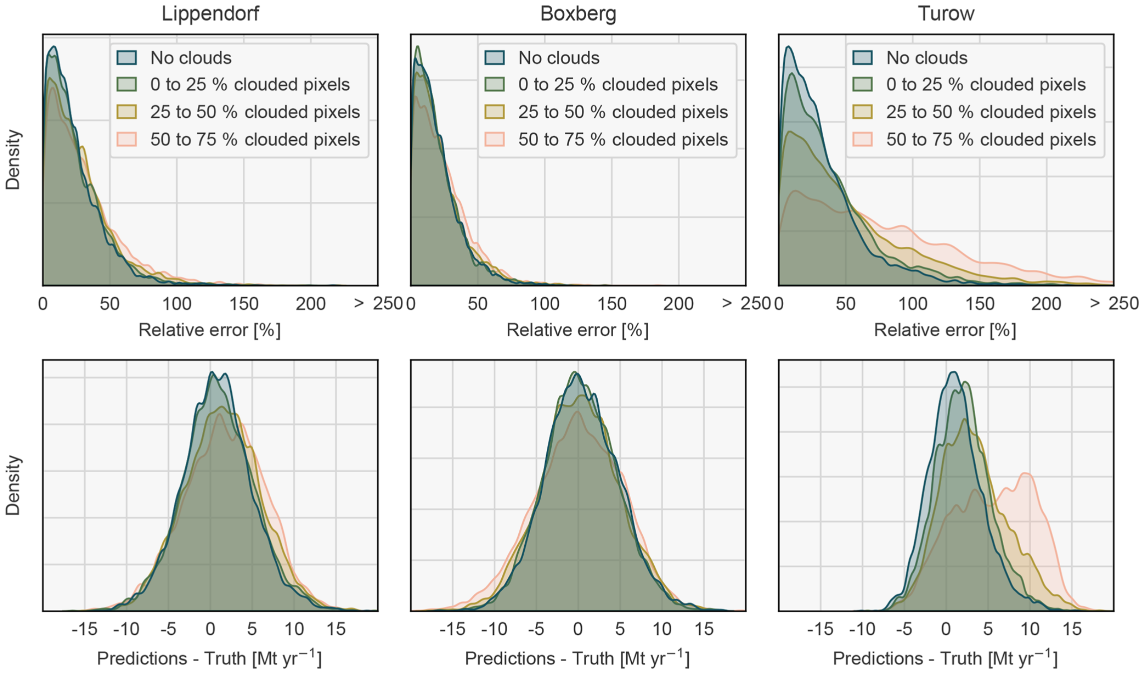

Figure 7. Density plots of the absolute relative error and of the algebraic difference between the CNN predictions and the reported emissions. The CNN models are trained and evaluated with XCO_2, wind, time, and NO_2 input data, affected by varying levels of cloud cover. Predictions with absolute relative errors greater than 250 % or absolute errors greater than 15 Mt yr^{-1} were set to 250 or 15, respectively, to increase visibility.

Table 2. Median of the relative error between the CNN predictions and the true emissions for Lippendorf, Boxberg, and Turów for varying levels of clouds and input configurations. Entries in the table represent relative errors expressed as percentages of the true emissions.

		Cloud fraction			
Power plant	With NO_2	0 %	0 % to 25 %	25 % to 50 %	50 % to 75 %
Lippendorf	x	23.1	22.8	25.1	23.7
Lippendorf	✓	17.9	18.7	21.6	23.5
Boxberg	x	18.5	19.5	20.2	21.1
Boxberg	✓	15.6	15.5	16.7	19.5
Turów	x	35.8	35.4	37.6	57.3
Turów	✓	24.8	28.9	37.9	65.2

median relative error with NO_2 decreased from 36.9 % to 15.6 %. Furthermore, an improvement in the results is observed when using the mean of the emissions predicted by applying the CNN to an ensemble of images with added Gaussian noise (not shown). Specifically, for added Gaussian noise with a standard deviation of 0.3, the Lippendorf median relative error (without additional inputs) decreases from 23.1 % to 18.5 %. Finally, incorporating time or wind as a feature yields no significant benefit with respect to the performance of the CNNs.

Concerning the influence of clouds, in the cases of Lippendorf and Boxberg, the accuracy of plume emission predictions is not significantly compromised by their introduction, even with a high cloud cover exceeding 50 %. This observation is valid whether or not NO_2 is factored into the analysis. However, for Turów, a power plant with lower emissions, the performance of CNN predictions degrades progressively with an increase in cloud cover, notably when cloud cover exceeds 50 %. The specific decline in prediction accuracy for Turów can likely be traced back to the fact that Turów's plume is mostly indistinguishable from the background. Consequently, the CNN's capacity to accurately estimate Turów's emissions is inherently based on limited information, even in the absence of clouds. The introduction of cloud cover exacerbates this issue by further diminishing the available information.

4.2 Application to OCO-3 SAM observations

In this section, we assess the ability of CNNs trained on power plant plumes from the SMARTCARB synthetic dataset encompassing the power plants of Jänschwalde, Schwarze Pumpe, Boxberg, Lippendorf, Turów, Pątnów, and Opole, to estimate emissions from real plumes observed at power plants by OCO-3 SAM, along with ERA5 wind fields and time information. A total of 39 observations of OCO-3 SAM data for nine power plants are examined.

To obtain meaningful statistics from the small number of images, we use two different methods to increase the number of predictions for each image:

1. an ensemble of 100 images $x_i^1, \ldots, x_i^{100}$ for each normalized OCO-3 SAM observation x_i, where $x_i^j \sim \mathcal{N}(x_i, 0.3)$;

2. an ensemble of 16 neural networks, all trained with slightly different hyperparameters considering various levels of cloud cover and with either uniform or beta distribution used for augmentation.

Each neural network generates 100 predictions from the 100 images. The ensemble mean should give a more accurate estimate than a single prediction for x_i, as seen with the synthetic data (see Sect. 4.1). Together with the 16 networks, we obtain 1600 predictions for each image, enhancing the robustness and reliability of our statistical analysis.

Figure 8 shows the ensemble of predictions for each power plant compared to the annual reported emissions. The median absolute and median absolute relative differences between the ensemble average predictions and the reported emissions are 7 $\mathrm{Mt\,yr^{-1}}$ and 29 %, respectively. This relative difference of 29 % is only slightly higher than what was observed on synthetic satellite imagery. Specifically, CNNs exhibit a good match with reported emissions for power plants with emissions ranging between 10 and 20 $\mathrm{Mt\,yr^{-1}}$ (e.g. Colstrip, Cumberland, Labadie, Parish, and Conemaugh). However, the discrepancy with reported emissions largely increases for power plants at the extremes of low or high emissions (e.g. Bełchatów, Tuoketuo, Hunters, and Intermountain). The emissions estimated by the CNNs range from 6.7 to 30.4 $\mathrm{Mt\,yr^{-1}}$ (considering the 5 % and 95 % quantile predictions), whereas Bełchatów's reported emissions stand at 37.6 $\mathrm{Mt\,yr^{-1}}$ and Intermountain's stand at 5 $\mathrm{Mt\,yr^{-1}}$. This indicates that the variance in CNN predictions is significantly lower than that of the reported emissions. Given that the CNNs were trained on plumes with emission levels spanning from 0 to 45 $\mathrm{Mt\,yr^{-1}}$, it was initially anticipated that they could accurately predict plumes akin to those from Bełchatów or Intermountain. Furthermore, in Sect. 4.1, we show that the CNNs reliably recover the low emissions (8.7 $\mathrm{Mt\,yr^{-1}}$ on average) of the Turów power plant. The subsequent analyses will explore the causes of the observed dis-

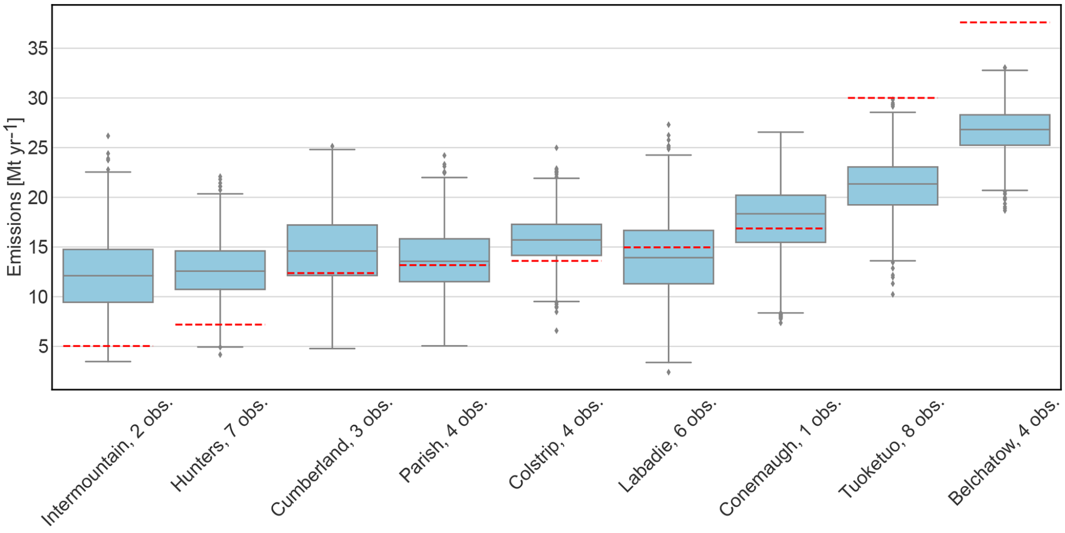

Figure 8. Box plots of the ensembles of predictions based on the OCO-3 SAM observations for various power plants. Comparison with the reported annual emissions of the corresponding power plants (dashed red lines). Boxes span the quartiles (25th to 75th percentiles), whiskers extend to the last points within 1.5 × IQR (interquartile range), and points beyond the whiskers represent outliers

crepancies between extreme reported emissions and CNN predictions.

Figure 9 shows the predictions of a randomly selected CNN model from the ensemble for eight specific OCO-3 SAM images. These images were chosen after a thorough inspection of all 39 snapshots in our dataset to illustrate the key patterns that we identified. For each OCO-3 SAM image, we show a sensitivity map obtained by the integrated-gradient method, which computes the gradient of the model's output (the emissions) relative to its input pixels, indicating how emissions are expected to increase or decrease with changes in pixel values (see Dumont Le Brazidec et al., 2024a, for details). Assuming that emission estimates are directly correlated with the detection of plume pixels, the integrated-gradient maps are anticipated to highlight a collection of positive pixels that effectively reconstruct the plume.

Figure 9a, b, c, and d are four instances of clear identification of the plume by the CNN. The integrated-gradient method in each of these cases reveals collections of positive pixels forming a discernible plume shape. These positive pixels are encircled by negatives, suggesting that if the surrounding pixels intensified to match the plume's pixel values, the CNN would be less likely to recognize these as plume pixels, interpreting the aggregate as elevated background values instead. Predictions closely match reported emissions in each scenario, barring the anomaly of Intermountain. This discrepancy is logical, given that the Intermountain plume is visually detectable in the image, whereas plumes corresponding to emissions of 5 Mt yr^{-1} in the SMARTCARB dataset are typically obscured by the background.

Figure 9e and f illustrate scenarios where clouds obscure the central portion of the image, thereby concealing a major part of the plume. These examples allow us to investigate how the CNN adapts to such conditions, making inferences based on the limited information available. In Fig. 9e, the CNN identifies a plume adjacent to the obscured area and bases its emission estimate on this collection of pixels. In Fig. 9g, the CNN interprets a significant cluster of high-value pixels as the tail end of the concealed plume and calculates emissions based on this inferred section of the plume.

Figure 9g and h shed light on a primary cause of the supposed overestimation of emissions from low-emission power plants. These images feature barely discernible plumes alongside significant patterns (potentially systematic satellite errors) appearing on the left side of the images in both cases. The CNN mistakenly identifies these patterns as part of a plume in each case. Consequently, the model infers disproportionately high emissions based on this noise, leading to a substantial overestimation of the emissions of these power plants.

In Fig. 10, we propose a first possible explanation for the underestimation of high-emission plumes by the CNN. Three observed high-emission plumes – one from Bełchatów and two from Tuoketuo – are compared against SMARTCARB plumes at the Bełchatów- or Tuoketuo-reported emission levels. Specifically, the SMARTCARB simulations are chosen to represent emissions of 37 and 29.5 Mt yr^{-1}, aligning with the reported emissions for Bełchatów and Tuoketuo, respectively, and have comparable ERA5 wind speeds to those of the Bełchatów or Tuoketuo snapshot OCO3 SAM images. The simulated plumes appear more pronounced against the background than their real counterparts, suggesting a higher emission magnitude. This observation might account for the model's tendency to estimate lower emissions than

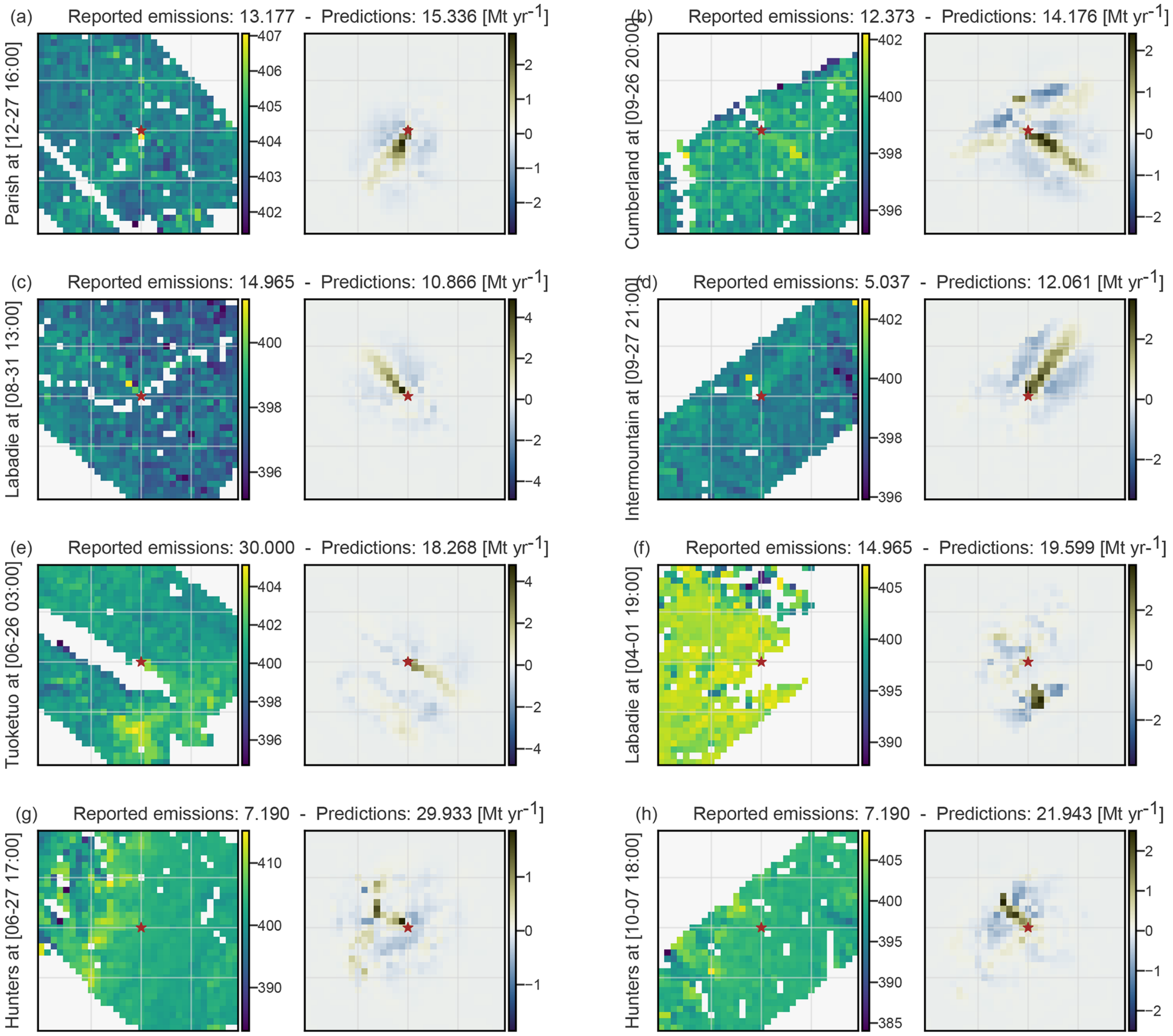

Figure 9. Analysis of the predictions of a CNN (chosen randomly from the ensemble) on eight specific OCO-3 SAM images. Each of the images is presented alongside the resulting map (on its right) from the application of the integrated-gradient method and the reported and predicted emissions in megatonnes per year (Mt yr^{-1}). The power plant is indicated by a brown star. The CO$_2$ concentration fields on the left are shown in parts per million by volume (ppmv). The integrated-gradient images on the right represent dimensionless scores that indicate the contribution of each pixel to the predicted concentration.

Bełchatów- and Tuoketuo-reported emissions. Further validation comes from the integrated-gradient analysis, indicating accurate plume contour predictions by the model and affirming that the relevant information was used for its estimations.

Another reason for the CNN underestimating high-emission plumes could be regression towards the mean in scenarios with high cloud levels. To show this, we train a CNN with a dataset of power plants with uniformly distributed emissions between 0 and 40 Mt yr^{-1} and a low cloud cover (< 25 % of the image covered with clouds). In Fig. 10, we plot the distribution of the predictions of the CNN for synthetic images of the Dolna Odra power plant with uniformly distributed emissions against the truth at different levels of cloud cover.

We observe a convergence to average values correlated with cloud cover intensity. When the CNN lacks sufficient information in the image to infer emissions, it tends to average its predictions to minimize loss. A second observation is that, even for low cloud cover, the CNN struggles with emission levels higher than 33 Mt yr^{-1}, while it is trained with emissions uniformly distributed between 0 and 40 Mt yr^{-1} (note that the CNNs trained in previous sections were trained for emission levels between 0 and 45 Mt yr^{-1}). Increasing the number of high-emission plumes in the training dataset

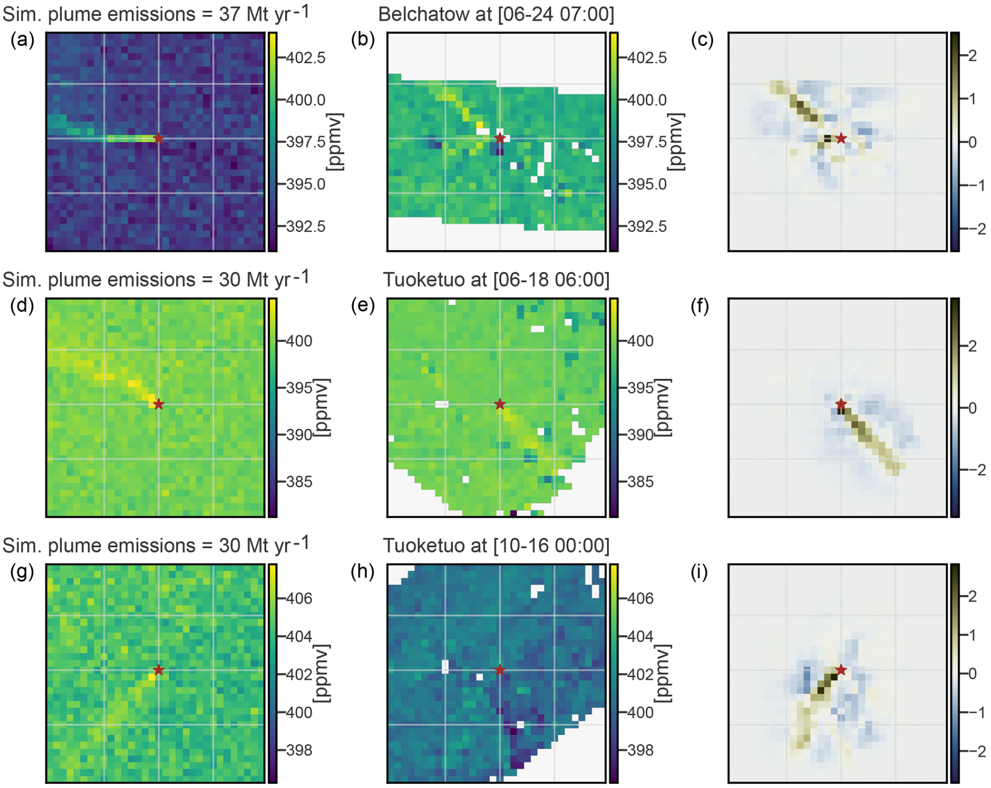

Figure 10. Analysis of the predictions of a CNN (chosen randomly from the ensemble) for three high-emission plumes. One plume from Bełchatów and two from Tuoketuo **(b, e, h)** are compared against equivalent high-emission plumes from the SMARTCARB dataset **(a, d, g)**. For each OCO-3-SAM-based plume, the integrated-gradient approach is applied and presented in panels **(c)**, **(f)**, and **(i)**. To ensure a fair comparison between columns 1 and 2, identical colour bars have been used. Power plants are indicated by a brown star.

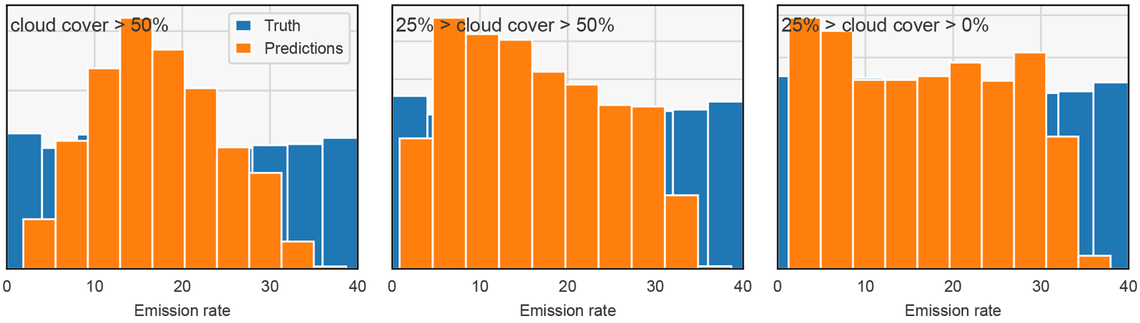

Figure 11. Distribution of the predictions of a CNN trained with uniformly distributed emissions for synthetic images of the Dolna Odra power plant with uniformly distributed emissions against the truth at different levels of cloud cover.

would likely reduce the CNN's bias towards emissions near the upper limit defined in the training data.

5 Discussions and limitations

The ability of CNNs to estimate CO_2 emissions from power plant plumes was validated on synthetic satellite images. The presence of cloud cover does not significantly affect the CNNs performance, except in instances of substantial cloud presence. CNNs demonstrate adaptability, leveraging residual information to accurately estimate emissions under heavily clouded conditions. The inclusion of NO_2 data proves slightly beneficial, enhancing the CNN efficacy under all-sky conditions.

Once trained on simulated XCO_2 images, the CNNs can be directly applied to real-world data with high accuracy, unlike traditional methods, which struggle to detect plumes and distinguish them from the background due to the low signal-to-noise ratio of CO_2 plumes. Nevertheless, it is observed that the spread of the CNN predictions is lower than the spread of the OCO3-SAM-reported emissions. Predictions are significantly higher for low-emission power plants, due to the presence of systematic errors in the image that are falsely identified as plumes, and significantly lower for high-emission power plants. Furthermore, predictions are lower than the reported annual emissions for high-emission power plants. This is likely due to regression towards the mean in weakly informative images and/or discrepancies between the training and evaluation datasets. Finally, it is acknowledged that comparing instantaneous emissions measured during satellite overpasses with reports of annual average emissions from the EPA and E-PRTR inventories presents challenges, owing to the variability and intermittent nature of power production and CO_2 emissions.

The divergence between the distributions of real XCO_2 observations and those of the simulations observed in Sect. 4.2, particularly in terms of systematic satellite errors, creates a domain shift between training and test conditions that likely leads to systematic errors in CNN predictions, necessitating CNN adaptation. To account for systematic satellite errors, a promising approach involves mingling real and simulated data during the training phase, such as overlaying a simulated plume of known emissions onto a real background. This method would introduce systematic errors typical of real satellite data while maintaining a controlled environment for supervised learning.

6 Conclusions and perspectives

In this paper, we improve the CNN model for the inversion of CO_2 plumes from Dumont Le Brazidec et al. (2024a) through the introduction of a novel data augmentation strategy and a dedicated approach to deal with clouds. This methodology was validated using the synthetic CO2M observations from the SMARTCARB dataset, demonstrating its efficacy in handling cloud-covered scenarios. Our findings indicate that, on average, clouds do not pose a significant challenge for CNNs, which maintain high performance levels under both sparse- and dense-cloud conditions. An exception is observed in the case of the Turów power plant, where performance significantly drops. This decline is likely attributable to Turów's relatively low emission levels, which result in its plumes being inherently less distinguishable from the background.

Following its validation, the methodology is applied to OCO-3 SAM observations. In total, 39 observations across nine power plants, adjusted for resolution and shape to match CNN input requirements, are analysed. For each observation, an ensemble of predictions is produced by CNNs trained on the SMARTCARB synthetic dataset. The results are promising, exhibiting a relative difference with the reported emissions only slightly superior to the relative error observed with the synthetic dataset. Specifically, predicted emissions for images from power plants with mid-level emissions, such as Colstrip and Parish, correspond very accurately to reported emissions. Moreover, through the application of integrated-gradient techniques, it is demonstrated that the CNNs effectively identify plumes in the OCO-3 SAM images and accurately estimate emissions from the plumes' physical locations.

However, we observed that images capturing low- and high-emission power plants' plumes are prone to overestimation and underestimation, respectively, in comparison to the reported emissions. Systematic satellite retrieval errors are identified as a frequent cause of overestimation in the images of low-emission power plants. These errors, often non-Gaussian and absent in the synthetic training dataset, lead to significant inaccuracies.

This study demonstrates the feasibility of applying neural networks to real satellite imagery of XCO_2 following training on simulated datasets. Although we advocate for the integration of a hybrid training approach that incorporates both real and simulated images in order to improve the robustness and accuracy of the model, we provide a ready-to-use CNN CO_2 plume inversion tool based on satellite imagery.

Code and data availability. The datasets used in this paper are available from a compliant repository (https://doi.org/10.5281/zenodo.12788520, Dumont Le Brazidec, 2024b) and originate from https://doi.org/10.5281/zenodo.4048228 (Kuhlmann et al., 2020b). The weights of the CNNs are available from https://doi.org/10.5281/zenodo.12788520 (Dumont Le Brazidec, 2024b) The codes for training the CNN are available from https://doi.org/10.5281/zenodo.14013176 (Dumont Le Brazidec, 2024c).

Author contributions. JDLB contributed to conceptualization, developed the methodology, implemented the software, conducted the investigation, performed formal analysis, created the visualizations, managed resources, and administered the project. PV contributed to the investigation and formal analysis. AF contributed to conceptualization, methodology, and project administration. MB contributed to the conceptualization, methodology, administered the project, and secured funding. GB contributed to conceptualization and methodology. GK provided resources. JDLB wrote the original draft of the manuscript; GB, GK, AF, MB, and PV contributed by reviewing the manuscript.

Competing interests. The contact author has declared that none of the authors has any competing interests.

Acknowledgements. This project has been funded by the European Union's Horizon 2020 Research and Innovation programme under grant agreement no. 958927 (Prototype system for a Copernicus CO_2 service). CEREA is a member of Institut Pierre-Simon Laplace (IPSL). The authors are grateful to Robert R. Nelson for his assistance with downloading the OCO-3 SAM images.

Financial support. This research has been supported by the European Union's Horizon 2020 Research and Innovation programme (grant no. 958927).

Review statement. This paper was edited by Le Yu and reviewed by two anonymous referees.

References

Broquet, G., Bréon, F.-M., Renault, E., Buchwitz, M., Reuter, M., Bovensmann, H., Chevallier, F., Wu, L., and Ciais, P.: The potential of satellite spectro-imagery for monitoring CO_2 emissions from large cities, Atmos. Meas. Tech., 11, 681–708, https://doi.org/10.5194/amt-11-681-2018, 2018.

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, Atmos. Chem. Phys., 19, 4541–4559, https://doi.org/10.5194/acp-19-4541-2019, 2019.

Chevallier, F., Remaud, M., O'Dell, C. W., Baker, D., Peylin, P., and Cozic, A.: Objective evaluation of surface- and satellite-driven carbon dioxide atmospheric inversions, Atmos. Chem. Phys., 19, 14233–14251, https://doi.org/10.5194/acp-19-14233-2019, 2019.

Chevallier, F., Broquet, G., Zheng, B., Ciais, P., and Eldering, A.: Large CO_2 Emitters as Seen From Satellite: Comparison to a Gridded Global Emission Inventory, Geophys. Res. Lett., 49, e2021GL097540, https://doi.org/10.1029/2021GL097540, 2022.

Cusworth, D. H., Duren, R. M., Thorpe, A. K., Eastwood, M. L., Green, R. O., Dennison, P. E., Frankenberg, C., Heckler, J. W., Asner, G. P., and Miller, C. E.: Quantifying Global Power Plant Carbon Dioxide Emissions With Imaging Spectroscopy, AGU Advances, 2, e2020AV000350, https://doi.org/10.1029/2020AV000350, 2021.

Danjou, A., Broquet, G., Lian, J., Bréon, F.-M., and Lauvaux, T.: Evaluation of light atmospheric plume inversion methods using synthetic XCO satellite images to compute Paris CO emissions, Remote Sens. Environ., 305, 113900, https://doi.org/10.1016/j.rse.2023.113900, 2024.

Danjou, A., Broquet, G., Schuh, A., Bréon, F.-M., and Lauvaux, T.: Optimal selection of satellite XCO_2 images for urban CO_2 emission monitoring, Atmos. Meas. Tech., 18, 533–554, https://doi.org/10.5194/amt-18-533-2025, 2025.

Dumont Le Brazidec, J., Vanderbecken, P., Farchi, A., Bocquet, M., Lian, J., Broquet, G., Kuhlmann, G., Danjou, A., and Lauvaux, T.: Segmentation of XCO_2 images with deep learning: application to synthetic plumes from cities and power plants, Geosci. Model Dev., 16, 3997–4016, https://doi.org/10.5194/gmd-16-3997-2023, 2023.

Dumont Le Brazidec, J., Vanderbecken, P., Farchi, A., Broquet, G., Kuhlmann, G., and Bocquet, M.: Deep learning applied to CO_2 power plant emissions quantification using simulated satellite images, Geosci. Model Dev., 17, 1995–2014, https://doi.org/10.5194/gmd-17-1995-2024, 2024a.

Dumont Le Brazidec, J.: Quantification of CO_2 hotspot emissions from OCO-3 SAM CO_2 satellite images using deep learning methods – data and weights, Zenodo [data set], https://doi.org/10.5281/zenodo.12788520, 2024b.

Dumont Le Brazidec, J.: JoffreyDumontLeBrazidec/co2-oco3-inv-dl: Quantification of CO_2 hotspot emissions from OCO-3 SAM CO2 satellite images using deep learning methods – code, Zenodo [code], https://doi.org/10.5281/zenodo.14013176, 2024c.

Eldering, A., Taylor, T. E., O'Dell, C. W., and Pavlick, R.: The OCO-3 mission: measurement objectives and expected performance based on 1 year of simulated data, Atmos. Meas. Tech., 12, 2341–2370, https://doi.org/10.5194/amt-12-2341-2019, 2019.

Finch, D. P., Palmer, P. I., and Zhang, T.: Automated detection of atmospheric NO_2 plumes from satellite data: a tool to help infer anthropogenic combustion emissions, Atmos. Meas. Tech., 15, 721–733, https://doi.org/10.5194/amt-15-721-2022, 2022.

Grant, D., Zelinka, D., and Mitova, S.: Reducing CO_2 emissions by targeting the world's hyper-polluting power plants, Environ. Res. Lett., 16, 094022, https://doi.org/10.1088/1748-9326/ac13f1, 2021.

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

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, 102171, https://doi.org/10.1016/j.apr.2024.102171, 2024.

Janssens-Maenhout, G., Pinty, B., Dowell, M., Zunker, H., Andersson, E., Balsamo, G., Bézy, J.-L., Brunhes, T., Bösch, H., Bojkov, B., Brunner, D., Buchwitz, M., Crisp, D., Ciais, P., Counet, P., Dee, D., Gon, H. D. v. d., Dolman, H., Drinkwater, M. R., Dubovik, O., Engelen, R., Fehr, T., Fernandez, V., Heimann, M., Holmlund, K., Houweling, S., Husband, R., Juvyns, O., Kentarchos, A., Landgraf, J., Lang, R., Löscher, A., Marshall, J., Meijer, Y., Nakajima, M., Palmer, P. I., Peylin, P., Rayner, P., Scholze, M., Sierk, B., Tamminen, J., and Veefkind, P.: Toward an Operational Anthropogenic CO_2 Emissions Monitoring and Verification Support Capacity, B. Am. Meteorol. Soc., 101, E1439–E1451, https://doi.org/10.1175/BAMS-D-19-0017.1, 2020.

Jongaramrungruang, S., Matheou, G., Thorpe, A. K., Zeng, Z.-C., and Frankenberg, C.: Remote sensing of methane plumes: instrument tradeoff analysis for detecting and quantifying local sources at global scale, Atmos. Meas. Tech., 14, 7999–8017, https://doi.org/10.5194/amt-14-7999-2021, 2021.

Joyce, P., Ruiz Villena, C., Huang, Y., Webb, A., Gloor, M., Wagner, F. H., Chipperfield, M. P., Barrio Guilló, R., Wilson, C., and Boesch, H.: Using a deep neural network to detect methane point sources and quantify emissions from PRISMA hyperspectral satellite images, Atmos. Meas. Tech., 16, 2627–2640, https://doi.org/10.5194/amt-16-2627-2023, 2023.

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, Atmos. Meas. Tech., 12, 6695–6719, https://doi.org/10.5194/amt-12-6695-2019, 2019.

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

Kuhlmann, G., Clément, V., Marshall, J., Fuhrer, O., Broquet, G., Schnadt-Poberaj, C., Löscher, A., Meijer, Y., and Brunner, D.: Synthetic XCO_2, CO and NO_2 observations for the CO2M and Sentinel-5 satellites, Zenodo [data set], https://doi.org/10.5281/zenodo.4048228, 2020b.

Kuhlmann, G., Henne, S., Meijer, Y., and Brunner, D.: Quantifying CO_2 Emissions of Power Plants With CO_2 and NO_2 Imaging Satellites, Front. Remote Sens., 2, 689838, https://www.frontiersin.org/article/10.3389/frsen.2021.689838 (last access: 15 June 2025), 2021.

Kuhlmann, G., Koene, E., 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), Geosci. Model Dev., 17, 4773–4789, https://doi.org/10.5194/gmd-17-4773-2024, 2024.

Kumar, S., Arevalo, I., Iftekhar, A. S. M., and Manjunath, B. S.: MethaneMapper: Spectral Absorption Aware Hyperspectral Transformer for Methane Detection, IEEE, 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Vancouver, Canada, 18–22 June 2023, 17609–17618, https://openaccess.thecvf.com/content/CVPR2023/html/Kumar_MethaneMapper_Spectral_Absorption_Aware_Hyperspectral_Transformer_for_Methane_Detection_CVPR_2023_paper.html (last access: 15 June 2025), 2023.

Lary, D. J., Alavi, A. H., Gandomi, A. H., and Walker, A. L.: Machine learning in geosciences and remote sensing, Geosci. Front., 7, 3–10, https://doi.org/10.1016/j.gsf.2015.07.003, 2016.

Lin, X., van der A, R., de Laat, J., Eskes, H., Chevallier, F., Ciais, P., Deng, Z., Geng, Y., Song, X., Ni, X., Huo, D., Dou, X., and Liu, Z.: Monitoring and quantifying CO_2 emissions of isolated power plants from space, Atmos. Chem. Phys., 23, 6599–6611, https://doi.org/10.5194/acp-23-6599-2023, 2023.

Meijer, Y.: Copernicus CO_2 Monitoring Mission Requirements Document, Earth and Mission Science Division, 84 pp., https://esamultimedia.esa.int/docs/EarthObservation/CO2M_MRD_v3.0_20201001_Issued.pdf (last access: 15 June 2025), 2020.

Meijer, Y., Andersson, E., Boesch, H., Dubovik, O., Houweling, S., Landgraf, J., Lang, R., and Lindqvist, H.: Editorial: Anthropogenic emission monitoring with the Copernicus CO_2 monitoring mission, Frontiers in Remote Sensing, 4, 1217568, https://doi.org/10.3389/frsen.2023.1217568, 2023.

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, Geophys. Res. Lett., 44, 10045–10053, https://doi.org/10.1002/2017GL074702, 2017.

Nassar, R., Mastrogiacomo, J.-P., Bateman-Hemphill, W., McCracken, C., MacDonald, C. G., Hill, T., O'Dell, C. W., Kiel, M., and Crisp, D.: Advances in quantifying power plant CO_2 emissions with OCO-2, Remote Sens. Environ., 264, 112579, https://doi.org/10.1016/j.rse.2021.112579, 2021.

Nassar, R., Moeini, O., Mastrogiacomo, J.-P., O'Dell, C. W., Nelson, R. R., Kiel, M., Chatterjee, A., Eldering, A., and Crisp, D.: Tracking CO_2 emission reductions from space: A case study at Europe's largest fossil fuel power plant, Frontiers in Remote Sensing, 3, https://www.frontiersin.org/articles/10.3389/frsen.2022.1028240 (last access: 15 June 2025), 2022.

Pillai, D., Buchwitz, M., Gerbig, C., Koch, T., Reuter, M., Bovensmann, H., Marshall, J., and Burrows, J. P.: Tracking city CO_2 emissions from space using a high-resolution inverse modelling approach: a case study for Berlin, Germany, Atmos. Chem. Phys., 16, 9591–9610, https://doi.org/10.5194/acp-16-9591-2016, 2016.

Reuter, M., Buchwitz, M., Schneising, O., Krautwurst, S., O'Dell, C. W., Richter, A., Bovensmann, H., and Burrows, J. P.: Towards monitoring localized CO_2 emissions from space: co-located regional CO_2 and NO_2 enhancements observed by the OCO-2 and S5P satellites, Atmos. Chem. Phys., 19, 9371–9383, https://doi.org/10.5194/acp-19-9371-2019, 2019.

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 CO_2 emissions from synthetic satellite images of XCO_2 and NO_2, Atmos. Meas. Tech., 18, 211–239, https://doi.org/10.5194/amt-18-211-2025, 2025.

Taylor, T. E., O'Dell, C. W., Baker, D., Bruegge, C., Chang, A., Chapsky, L., Chatterjee, A., Cheng, C., Chevallier, F., Crisp, D., Dang, L., Drouin, B., Eldering, A., Feng, L., Fisher, B., Fu, D., Gunson, M., Haemmerle, V., Keller, G. R., Kiel, M., Kuai, L., Kurosu, T., Lambert, A., Laughner, J., Lee, R., Liu, J., Mandrake, L., Marchetti, Y., McGarragh, G., Merrelli, A., Nelson, R. R., Osterman, G., Oyafuso, F., Palmer, P. I., Payne, V. H., Rosenberg, R., Somkuti, P., Spiers, G., To, C., Weir, B., Wennberg, P. O., Yu, S., and Zong, J.: Evaluating the consistency between OCO-2 and OCO-3 XCO_2 estimates derived from the NASA ACOS version 10 retrieval algorithm, Atmos. Meas. Tech., 16, 3173–3209, https://doi.org/10.5194/amt-16-3173-2023, 2023.

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, Atmos. Meas. Tech., 11, 5673–5686, https://doi.org/10.5194/amt-11-5673-2018, 2018.

Wu, D., Lin, J. C., Oda, T., and Kort, E. A.: Space-based quantification of per capita CO_2 emissions from cities, Environ. Res. Lett., 15, 035004, https://doi.org/10.1088/1748-9326/ab68eb, 2020.

Zheng, B., Chevallier, F., Ciais, P., Broquet, G., Wang, Y., Lian, J., and Zhao, Y.: Observing carbon dioxide emissions over China's cities and industrial areas with the Orbiting Carbon Observatory-2, Atmos. Chem. Phys., 20, 8501–8510, https://doi.org/10.5194/acp-20-8501-2020, 2020.