



## RESEARCH LETTER

10.1029/2021GL097540

Large CO<sub>2</sub> Emitters as Seen From Satellite: Comparison to a Gridded Global Emission InventoryFrédéric Chevallier<sup>1</sup> , Grégoire Broquet<sup>1</sup> , Bo Zheng<sup>2</sup>, Philippe Ciais<sup>1</sup> , and Annmarie Eldering<sup>3</sup>

<sup>1</sup>Laboratoire des Sciences du Climat et de l'Environnement, LSCE/IPSL, CEA-CNRS-UVSQ, Université Paris-Saclay, Gif-sur-Yvette, France, <sup>2</sup>Institute of Environment and Ecology, Tsinghua Shenzhen International Graduate School, Tsinghua University, Shenzhen, China, <sup>3</sup>Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA, USA

## Key Points:

- We have retrieved some instantaneous CO<sub>2</sub> emissions for one third of the large emission cells of a global high-resolution hourly inventory
- The emission retrievals explain more than one third of the inventory variance at the corresponding cells and hours
- Consistent temporal variations of median emissions suggest that trends can be robustly calculated when more data become available

## Correspondence to:

F. Chevallier,  
[frederic.chevallier@lscce.ipsl.fr](mailto:frederic.chevallier@lscce.ipsl.fr)

## Citation:

Chevallier, F., Broquet, G., Zheng, B., Ciais, P., & Eldering, A. (2022). Large CO<sub>2</sub> emitters as seen from satellite: Comparison to a gridded global emission inventory. *Geophysical Research Letters*, 49, e2021GL097540. <https://doi.org/10.1029/2021GL097540>

Received 17 DEC 2021  
Accepted 22 FEB 2022

## Author Contributions:

**Conceptualization:** Frédéric Chevallier, Grégoire Broquet  
**Data curation:** Frédéric Chevallier  
**Formal analysis:** Frédéric Chevallier, Grégoire Broquet  
**Funding acquisition:** Frédéric Chevallier  
**Investigation:** Frédéric Chevallier, Grégoire Broquet, Bo Zheng, Philippe Ciais  
**Methodology:** Frédéric Chevallier, Grégoire Broquet, Bo Zheng, Philippe Ciais  
**Project Administration:** Frédéric Chevallier  
**Resources:** Frédéric Chevallier  
**Software:** Frédéric Chevallier, Bo Zheng, Annmarie Eldering  
**Supervision:** Frédéric Chevallier

© 2022. The Authors.

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial-NoDerivs License](#), which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

**Abstract** Using the multiyear archive of the two Orbiting Carbon Observatories (OCO) of NASA, we have retrieved large fossil fuel CO<sub>2</sub> emissions (larger than 1.0 ktCO<sub>2</sub> h<sup>-1</sup> per 10<sup>-2</sup> square degree grid cell) over the globe with a simple plume cross-sectional inversion approach. We have compared our results with a global gridded and hourly inventory. The corresponding OCO emission retrievals explain more than one third of the inventory variance at the corresponding cells and hours. We have binned the data at diverse time scales from the year (with OCO-2) to the average morning and afternoon (with OCO-3). We see consistent variations of the median emissions, indicating that the retrieval-inventory differences (with standard deviations of a few tens of percent) are mostly random and that trends can be calculated robustly in areas of favorable observing conditions, when the future satellite CO<sub>2</sub> imagers provide an order of magnitude more data.

**Plain Language Summary** In the wake of the Paris Climate Agreement, there is an increasing need to monitor emissions from fossil fuel combustion around the world. For CO<sub>2</sub> in particular, satellite imagers are being designed to observe the emission plumes from large point sources and intense urban area sources. In order to assess their potential, we have tested a simple emission retrieval scheme on the multi-year archive of the two NASA Orbiting Carbon Observatories which provide dense observations along the orbit line, but with a narrow swath. We have compared our results with a global gridded and hourly inventory. The corresponding emission retrievals explain a large part of the inventory variability, despite uncertainty in both datasets. We also see consistent variations in the middle emission values at different time scales. These results suggest that the differences between retrievals and inventory are mostly random and that the trends can be calculated robustly in areas of favorable observation conditions, when future satellite CO<sub>2</sub> imagers provide an order of magnitude more data.

## 1. Introduction

With Article 4 of the 2015 Paris Climate Agreement, signatories (called Parties) aim to reach the peak of global greenhouse gas (GHG) emissions as quickly as possible, and then rapidly undertake their reduction. This ambition is not legally binding, but the Parties commit their public notoriety within the Enhanced Transparency Framework (ETF): from 2024 onwards, they will report every two years, under international peer review, on their progress in implementing their Nationally Determined Contributions (NDCs). Any evidence of a change in emission or removal of GHGs will be used either upstream to help Parties prepare accurate biennial transparency reports (BTRs) or downstream to facilitate peer review and public scrutiny.

By observing most of the Earth in a coherent manner, satellites offer solid advantages for this purpose, supported by new or improving spectral imaging capabilities for the surfaces or the atmosphere (e.g., Janssens-Maenhout et al., 2020). Carbon dioxide (CO<sub>2</sub>), being the main GHG emitted by human activities, is a major target for observation in the atmosphere, but it requires exceptionally fine detection capabilities and processing quality, which have only recently emerged (see, e.g., Chevallier et al., 2019, and references therein). Within the ETF, observations of CO<sub>2</sub> can inform about anthropogenic CO<sub>2</sub> emissions from the “slow carbon cycle” (emissions from the burning of fossil fuels or the production of cement, called FF hereafter) and from the “fast carbon cycle” (CO<sub>2</sub> emissions or removals from agriculture, forestry and other land use activities over managed lands). We refer to Chevallier (2021) and Deng et al. (2021) for a discussion about the latter. Here, we discuss the current capability to estimate the former with satellite retrievals of the column-average CO<sub>2</sub> dry air-mole fraction (XCO<sub>2</sub>).

**Validation:** Frédéric Chevallier,

Annamarie Eldering

**Visualization:** Frédéric Chevallier

**Writing – original draft:** Frédéric Chevallier

**Writing – review & editing:** Frédéric Chevallier, Grégoire Broquet, Bo Zheng, Philippe Ciais, Annmarie Eldering

We are studying the case of XCO<sub>2</sub> observations from NASA's second and third Orbiting Carbon Observatories (OCO-2 and -3). Our motivation here vis-à-vis these two spectrometers is twofold. First, their continuous sampling of the Earth along their orbit line with a resolution of 3 km<sup>2</sup>, makes them the closest existing instruments to a CO<sub>2</sub> imager like the future Copernicus CO<sub>2</sub> Monitoring Mission (CO2M; ESA, 2019): they would only need a wide swath to correspond to such an imager, because theirs is narrow (between 2 and 10 km). Second, the XCO<sub>2</sub> retrievals based on their measurements achieve exceptional accuracy by today's standards with root-mean-squared difference to the reference data from the Total Carbon Column Observing Network (TCCON, Wunch et al., 2011) close to 1 μmol mol<sup>-1</sup> (micromole per mole, abbreviated as ppm, for parts per million, in what follows). OCO-2 started its science measurements in September 2014 and is still operated at the time of writing, which extends its data record to more than seven years. OCO-3 science measurements started in August 2019, but in a different orbital configuration that allows sampling diverse local daytime hours, while OCO-2 science data are all close to local midday.

We retrieve FF CO<sub>2</sub> emissions from the transects of XCO<sub>2</sub> provided by OCO-2, closely following the approach of Zheng et al. (2020): after Nassar et al. (2017) or Reuter et al. (2019), these authors detected the isolated XCO<sub>2</sub> enhancements along the OCO-2 orbits that could correspond to transects of plumes from nearby upwind sources, estimated the corresponding XCO<sub>2</sub> line densities and multiplied them by the wind speed in the direction normal to the OCO-2 tracks to estimate the corresponding FF CO<sub>2</sub> emissions. Zheng et al. (2020) tested the approach for all orbits over or around China between September 2014 and August 2019. They identified 60 good cases of FF plume transects and found good consistency of their emission retrieval with a detailed gridded regional inventory (with a coefficient of determination  $r^2 = 0.5$  after wind direction adjustment). Chevallier et al. (2020) automated the processing to extend it to the whole globe, with a focus on the first months of the coronavirus recession in 2020, compared to the rest of the archive: they found that the main FF CO<sub>2</sub> emission anomalies between February and May 2020 were obscured by frequent or persistent cloud conditions in the regions where they occurred. Here, we further extend this work to seven years of OCO-2 data (September 2014–September 2021) and, for the first time, to OCO-3 data (August 2019–September 2021). We return to the question of temporal variability, which we did not address in Zheng et al. (2020) and that we left on a failure (concerning the detection of the anomaly corresponding to the coronavirus recession) in Chevallier et al. (2020). For that, we compare the corresponding emission retrievals to the Emissions Database for Global Atmospheric Research (EDGAR) v6.0, a global bottom-up emission inventory that is available at resolution 0.1° × 0.1° globally and down to the hourly scale for the whole globe until December 2018. We look for the interannual, seasonal, weekly and daily FF CO<sub>2</sub> emission signals from large CO<sub>2</sub> emitters, that are defined here as 0.1° × 0.1° inventory cells where an emission greater than or equal to 1.0 ktCO<sub>2</sub> h<sup>-1</sup> is reported.

The comparison is presented in Section 3, after the description of our data and of our retrieval technique in Section 2. General conclusions are drawn in Section 4.

## 2. Retrievals and Inventories

### 2.1. XCO<sub>2</sub> Retrievals

The name “OCO-2” designates both a specific spacecraft in sun-synchronous polar orbit and the single sounding instrument it carries. This instrument incorporates three boresighted spectrometers. “OCO-3” designates a similar instrument to that of OCO-2, but is installed as an external payload on the International Space Station (ISS). OCO-3 is therefore not autonomous, and follows viewing and operation constraints from the ISS. This disadvantage is offset by a new pointing mirror assembly which allows for much more agile vision than OCO-2. The ISS has an orbital inclination of 51.6° and does not cross the Equator at a fixed local solar time (LST), while the ascending node of OCO-2 crosses it systematically at around 1:30 p.m. LST. The LST of the measurement is important because the two OCO instruments measure reflected sunlight, but also because XCO<sub>2</sub> varies during the day mostly due to changes in natural emissions and absorptions, and in FF emissions. The ISS also has a 12% faster ground-track velocity than OCO-2 (7.667 km s<sup>-1</sup> vs. 6.75 km s<sup>-1</sup>; see Eldering et al., 2019, and Baker et al., 2021), resulting in slightly less spatially dense retrievals, except when a specific scanning mode is enabled to perform side-by-side scans and provide snapshot area maps (SAMs) of ~80 × 80 km<sup>2</sup>. The soundings of the two OCOs are therefore distributed differently in space and time, but their footprint and precision are similar (~3 km<sup>2</sup> and better than 1 ppm, respectively).

For the two OCO instruments, we use all operational quality-controlled bias-corrected XCO<sub>2</sub> retrievals with associated uncertainty over both land and ocean made by the Atmospheric CO<sub>2</sub> Observations from Space (ACOS) algorithm of NASA, version 10, described by Osterman et al. (2020). We neglect the impact of the deviation of the column averaging kernel from an ideal vector of ones. When computing the XCO<sub>2</sub> line densities (see Section 2.2), our emission retrieval also uses the a priori surface pressure and the a posteriori total column water vapor of the ACOS data.

## 2.2. FF CO<sub>2</sub> Emission Retrievals

Our approach to retrieve the CO<sub>2</sub> emissions from the anthropogenic plume transects seen by the OCO instruments closely follows the detailed description made in Zheng et al. (2020) and Chevallier et al. (2020). We therefore briefly summarize it and explain the few refinements that we have brought to it since then. We process data from both instruments in exactly the same way.

Each OCO orbit is analyzed with a 200 km moving window successively centered on each of the validated retrieval. The width of the moving window defines the maximum size of enhancements allowed. If a retrieval value stands out of the variability of the retrievals in the window, the procedure attempts to fit a function that represents a bell curve on top of a linear background, on the soundings located over the dominant surface type in the window. For the SAM observation mode of OCO-3, we exclude from the fit the retrievals that are not in the same scan line, in order to stay within a cross-sectional vision. The soundings are represented by their XCO<sub>2</sub> retrievals and by their position  $x$  along the satellite track: their across-track position is ignored, the satellite narrow swath being used only to damp retrieval noise. The restriction to the soundings of the dominant surface type (land or water) mainly skips the artificial discontinuities in the retrievals linked to the corresponding changes in the surface albedo. The adjustable parameters are the standard deviation  $\sigma$  and the height  $A$  of the Gaussian ( $A \cdot e^{\left[\frac{-x^2}{2\sigma^2}\right]}$ ), and the slope  $m$  and offset  $b$  of the line ( $m \cdot x + b$ ). The center of the peak is positioned in the middle of the 200 km moving window and is not adjusted. The choice of the Gaussian form is linked to its neutrality (in the sense of the principle of maximum entropy) rather than to a hypothesis on the exact shape of the plume. In particular, a visual inspection of the orbits reveals that many XCO<sub>2</sub> enhancements take the form of Gaussian mixtures that can still be well adjusted by a single Gaussian: these ones are retained by the automatic process.

The quality of the fit is evaluated by the following criteria (numerical values given only when they differ from Chevallier et al., 2020): (a) its coefficient of determination  $r^2$  (it has to be larger than 0.7<sup>2</sup>), (b) the density of the retrievals within  $1\sigma$  of the Gaussian center, and between  $2\sigma$  and  $3\sigma$  of it, (c) the value of  $\sigma$  (it must be between 2 and 200/6 km, except for the 80 × 80 km<sup>2</sup> SAM observation mode of OCO-3, for which the maximum acceptable  $\sigma$  is reduced to 20 km), (d) the value of the Gaussian height relative to the retrieval variability in the window. A validated fit is interpreted as reflecting the potential crossing of a CO<sub>2</sub> plume from an emitting area or point source. If several validated fitting functions, centered on different retrievals, overlap, only the one with the largest  $r^2$  is kept. Such an overlap can happen for instance between estimates made from different scan lines of a SAM.

By definition, the area under the Gaussian is the XCO<sub>2</sub> line density (in ppm  $m$ , then converted to kg m<sup>-1</sup>), which reflects the distribution of CO<sub>2</sub> abundance in the CO<sub>2</sub> plume. We multiply it with the wind speed in the direction normal to the OCO track to retrieve the corresponding emission under the assumption of a steady wind. The wind speed is taken from the fifth generation of ECMWF atmospheric reanalyses of the global climate (ERA5, Hersbach et al., 2020) in the model ninth level from the surface, that corresponds to a geometric altitude of about 250 m (after Brunner et al., 2019). The few cases with retrieved emissions less than 0.5 ktCO<sub>2</sub> h<sup>-1</sup> are judged as unreliable and are left out. This detection threshold is comparable to the one found by Lespinas et al. (2020).

This set of emission retrievals is then filtered to keep only the fresh plumes, defined here as plumes which are mostly less than 3 hr old, because older ones have a more complex transport history. In Chevallier et al. (2020), we used a loose criterion based on satellite retrievals of the nitrogen dioxide column. Here we use a stricter criterion, in which only OCO-observed plume transects that are either downwind of, or less than 30 km from, a 0.1° × 0.1° cell of EDGAR where an emission of at least 1.0 ktCO<sub>2</sub> h<sup>-1</sup> is reported (see Section 2.4 below for a definition of “downwind” and for an explanation of the 30 km criterion). Note that we do not attempt to optimize the wind direction given the uncertainty in the emitter location in EDGAR and given the relatively large area of the cells of this inventory (10<sup>-2</sup> square degree). Other sources of information would be needed to do it properly.

### 2.3. FF CO<sub>2</sub> Emission Inventory

EDGAR is a scientific emission inventory for air pollutants, aerosols and GHGs over the globe, that synthesizes international and national activity data and combines them with emission factors. Version 6.0 provides, among other data, monthly FF CO<sub>2</sub> emission values in 16 sectors (plus others for ship and aviation that are ignored here) and gridded globally at resolution  $0.1^\circ \times 0.1^\circ$  from 2000 until 2018 ([https://edgar.jrc.ec.europa.eu/index.php/dataset\\_ghg60](https://edgar.jrc.ec.europa.eu/index.php/dataset_ghg60), accessed 17 December 2021, Crippa et al., 2021). The seasonal variations in FF CO<sub>2</sub> emissions are mostly for residential heating, with smaller variations for the road-transport and power-industry sectors. The subnational spatial distribution of the emissions corresponds as much as possible to the places where the corresponding activity is likely located, as defined from various auxiliary databases (Janssens-Maenhout et al., 2019).

We further downscale the monthly emission maps at the hourly level using typical daily and weekly temporal profiles in each country and sector provided by Crippa et al. (2020). For the years 2019, 2020 and 2021, we simply take the EDGAR values valid for the year 2018. This neglects the large 2020 anomaly of FF emissions, but OCO-2 barely saw it in the total column XCO<sub>2</sub> (see Section 1).

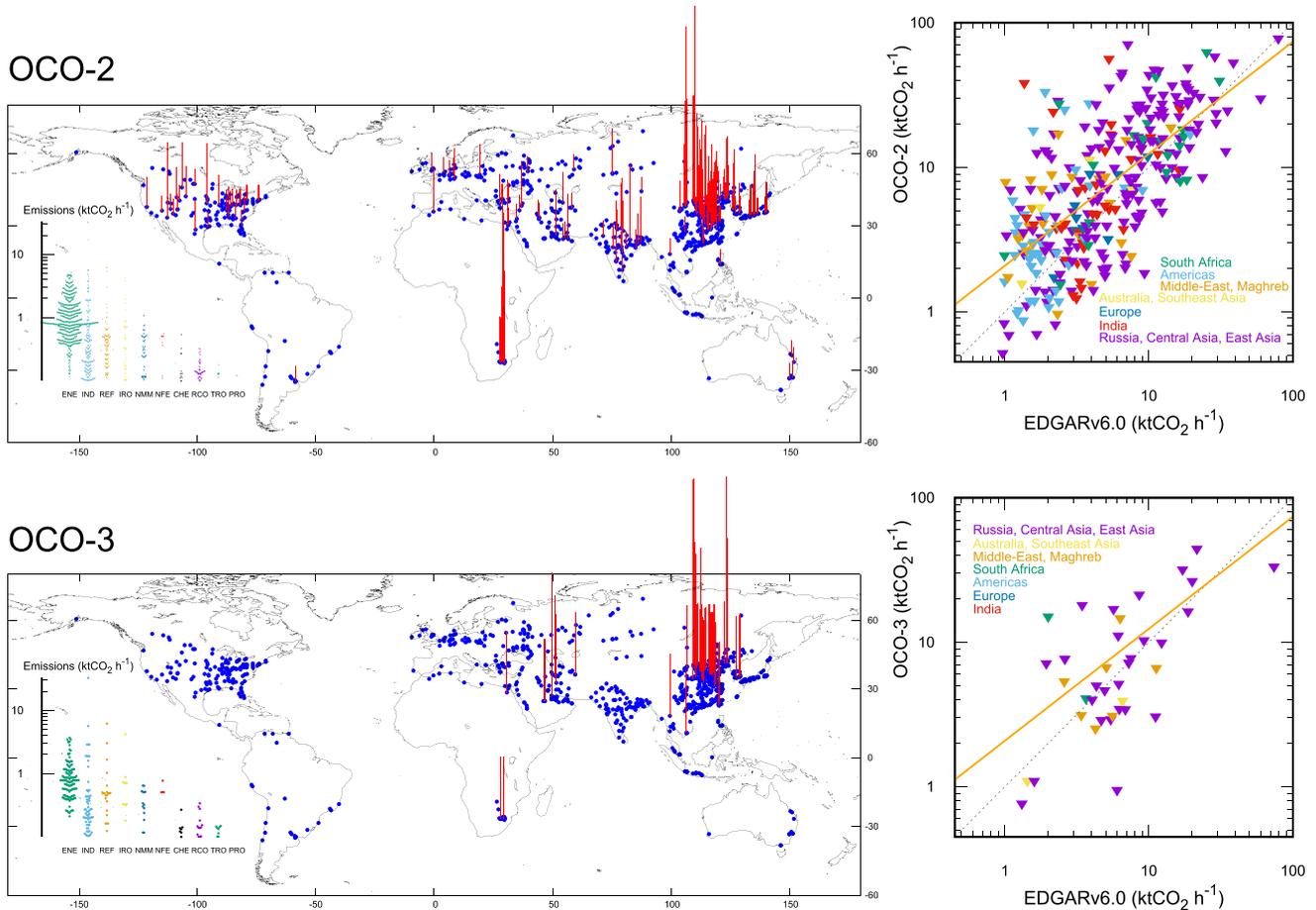
### 2.4. Definition of the Footprint of the FF CO<sub>2</sub> Emission Retrievals

The uncertainty in the wind field and the lack of information on the vertical distribution of the plume leaves a large margin for attribution of the plume to specific emitters, beyond the large nearby one that allows selecting the emission retrievals, as explained in Section 2b. We define the plume footprint ad hoc as follows. First, we only consider EDGAR cells whose emission plumes could reach the orbit within less than 7 hr in a straight line, given the wind at the peak (as defined in Section 2.2). Older plumes are assumed to be too diffused to contribute to the observed enhancement (Broquet et al., 2018). Second, we exclude the cells whose emission would pass through the orbit in a straight line (given the wind at the peak) more than  $5\sigma$  from the Gaussian center. The loose  $5\sigma$  criterion accounts for the uncertainty in the wind direction. For the same reason, we reintegrate EDGAR-estimated emissions from the cells that are “downwind” of the orbit but less than 30 km from it. In the following, we compare the emission retrievals for a given date and time with the sum of the emissions from the selected cells at the time in the past consistent with the wind and the distance to the orbit.

## 3. Results

The automatic algorithm of Section 2.2 selected 313 XCO<sub>2</sub> enhancements in the OCO-2 seven-year record and 38 in the two-year OCO-3 record. The resulting high computational cost (about 100,000 hr Central Processing Unit time on AMD Rome compute nodes at 2.6 GHz with 256 GB memory per node) limited us to a one-off analysis of the OCO archive. The quality of the Gaussian fits is overall similar in both cases with median  $r^2$  of 0.59 and 0.55, respectively for the OCO-2 collection and the OCO-3 collection. With 38 emission retrievals, the average OCO-3 yield of emission retrievals is 40% of its twin. It is related to the many interruptions of data collection for OCO-3 further to the ramp-up of the mission during its first months and to the operation constraints aboard the ISS Taylor et al. (2020). It is also linked to the acquisition of ~20 SAMs usually per day on land, that target the same areas as our method while reducing the data density on both sides along the orbit (in transition observation mode).

As seen in Figure 1, the XCO<sub>2</sub> enhancements selected for OCO-2 are mostly located in China (144 cases), the USA (40), India (35), South Africa (20), South Korea (10), Japan (10), Turkey (7), and Germany (5). The 313 OCO-2 enhancements are attributed to 329 large-emission cells of EDGAR, that is, one third of the 1050 cells the emission of which exceeds  $1.0 \text{ ktCO}_2 \text{ h}^{-1}$  at some stage during the period September 2014 – December 2018 (note that the total of the maximum values on these 1050 cells is  $2.5 \text{ MtCO}_2 \text{ h}^{-1}$ , but the sum of their annual-mean emission corresponds to  $1.6 \text{ MtCO}_2 \text{ h}^{-1}$  only). On average, most assigned cells are visited twice during the period, but the cell of the Dalate coal-fired power station in Inner Mongolia, China, is seen 13 times and the one of the Ningxia Lingwu coal-fired power station in Ningxia, China, is seen 11 times. Also noticeable are the 10 attributions to each of the three cells of the Tutuka, Kendal and Kriel coal-fired power plants, southeast of Johannesburg, South Africa. More generally, according to the inventory (see the inset of the map in Figure 1), emissions in the assigned cells are mostly from large emitters in the power-industry sector (ENE) and, to a smaller extent, in the combustion-for-manufacturing sector (IND); the other sectors, like road transport, have a smaller share. We note that few coastal cells are selected regardless of the geographical orientation of the coast, likely because of

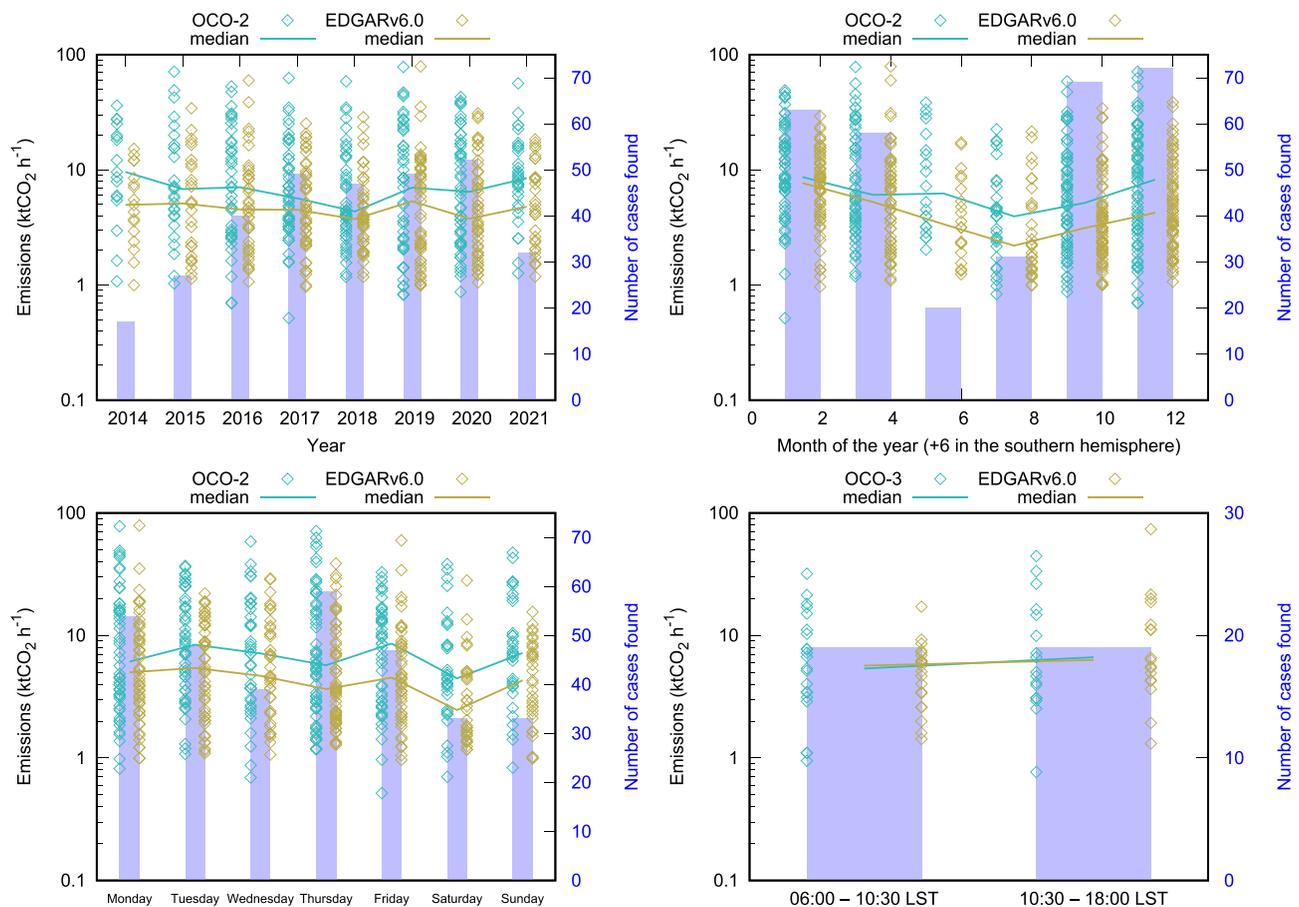


**Figure 1.** In the maps on the left, the blue dots represent the EDGARv6.0 cells the emission of which exceeds  $1.0 \text{ ktCO}_2 \text{ h}^{-1}$  at some stage during the period September 2014–December 2018. Red impulses on these maps illustrate the number of times when retrieved emissions are attributed to these cells for OCO-2 (top) and OCO-3 (bottom). The maximum height corresponds to 13 times and 3 times, respectively for OCO-2 and OCO-3. A jitter plot on a vertical logarithmic scale is superimposed in the eastern Pacific part of each map. It shows the distribution of the emissions per sector in the assigned cells: power industry (ENE), combustion for manufacturing (IND), oil refineries (REF), production of iron and steel (IRO), production of non-metallic minerals (NMM), production of non-ferrous metals (NFE), production of chemicals (CHE), residential (RCO), road transport (TRO) and fuel production/transmission (PRO). The scatter plots on the right are for the retrieved emission values versus the inventory values for OCO-2 (top) and OCO-3 (bottom), with colors reflecting the geographical location. The orange line is the regression line in base 10 logarithm. The black dots form the bisector.

the specific surface albedo, wind and cloud configurations in these areas in addition to taking into account only soundings over the dominant type of surface in the orbit segment.

A majority (22) of the OCO-3 enhancements correspond to the nominal observation modes (nadir or glint), for which the inversion method is the most suitable. The other 16 cases correspond to the SAM observation mode. The OCO-3 enhancements follow the same distribution in space and sectors as OCO-2 (Figure 1 bottom), but no enhancement is selected in Western countries.

The scatter plot on the top right of Figure 1 displays the collection of 313 OCO-2 emission retrievals as a function of the EDGAR values at emission time (rather than enhancement time), given the wind speed. The fact that the cases have been selected on the condition that there is an emission hot spot in EDGAR greater than or equal to  $1.0 \text{ ktCO}_2 \text{ h}^{-1}$  in the vicinity of the  $\text{XCO}_2$  enhancement (in order to select fresh plumes, see Section 2.2) makes the cloud of points asymmetric for small emission values. Still, the cloud of points overall stretches along the bisector, with an  $r^2$  of 0.37 ( $p = 3 \cdot 10^{-33}$  for a hypothesis test whose null hypothesis is that the slope is zero). The standard deviation of the relative differences correspondingly decreases with increasing retrieval values: 80% for the retrieval values less than  $2 \text{ ktCO}_2 \text{ h}^{-1}$  (including the artifact caused by the threshold on the inventory data), 68% for  $2\text{--}5 \text{ ktCO}_2 \text{ h}^{-1}$ , down to 38% for  $30\text{--}80 \text{ ktCO}_2 \text{ h}^{-1}$ . Figure 1 displays the cloud of points with both axes



**Figure 2.** Retrieved emission values and inventory values as a function of the year, the month of the year (+6 in the southern hemisphere) averaged 2 by 2, the day of the week and the time of the day (of the plume, not of the emissions). In each graph, the number of cases per bin is reported in violet (right axis). The median in each bin is also reported as a curve. Note that the graphs are for OCO-2, except the one in the bottom right quarter which is for OCO-3.

on logarithmic scales because the retrieval values are distributed heterogeneously over three orders of magnitude. If we take the logarithm of the emissions, as is done in practice in the figure,  $r^2$  increases up to 0.44. The slope is then  $0.78 \pm 0.10$  (95% confidence interval) and the power law deduced from the regression line is  $y = 2.08 x^{0.78}$ . There is usually more than one large emitter in the plume footprint and the agreement is for the total, not for one specifically. The point in the upper right corner is for an enhancement West of Anshan, China. In this case, the agreement between retrieval and inventory is better than 2%, but we note that the spatial and sectoral distribution of the corresponding emissions in EDGAR is very different from the recent Global Carbon Grid v1.0 from Tsinghua University which benefits from detailed information about Chinese emitters (<http://gidmodel.org.cn/>, accessed 17 December 2021).

The statistics for the 38 OCO-3 enhancements show similar linear correlation:  $r^2$  is 0.41, either without or with logarithm. Given their lesser representativity, we will focus our analysis of the OCO-3 emission collection on the issue of local time dependency, that OCO-2 cannot address.

Figure 2 shows the emission retrievals and the corresponding inventory values as a function of (for OCO-2 only) the year, the month of the year, the day of the week and (for OCO-3 only) the time of the day.

The number of enhancements per year for OCO-2 varies from single to double even when excluding incomplete annual time series (2014, 2021 or year 2017 when the instrument was down for 1.5 months). The interannual variability of EDGAR in the enhancement footprint is relatively small with median values oscillating between 3.7 and 5.3  $\text{ktCO}_2 \text{ h}^{-1}$  (remember that the EDGAR maps after 2018 are those of 2018). The retrieval variation has a slightly larger amplitude (4.3–8.4  $\text{ktCO}_2 \text{ h}^{-1}$  in 2015–2021) and a simpler V shape with 2018 at the minimum. As noted above, we do not expect the large emission reduction that occurred at the start of the Covid-19 pandemic to

be reflected in the emission retrievals because this anomaly mostly occurred in the presence of clouds (in China, the USA or India) or too high in the north during that time of the year for OCO-2 (Europe).

The median seasonal cycle shows rather consistent amplitude (3.9–8.6 ktCO<sub>2</sub> h<sup>-1</sup> for the retrieval median vs. 2.2–7.7 ktCO<sub>2</sub> h<sup>-1</sup> for the inventory) with a minimum around August and a maximum in February (after a correction of 6 months in the southern hemisphere). The summer minimum is also seen in the number of cases. The two cycles may be related indeed: less emissions in summer mechanically reduce the number of detected enhancements; a similar feature is seen for Saturday cases below. The two cycles could be accidentally correlated as well due to more cloudiness in the summer at the time of the day when OCO-2 passes.

During the week, the retrievals and the inventory describe relatively small variations from Monday to Friday (5.7–8.6 and 3.6–5.4 ktCO<sub>2</sub> h<sup>-1</sup> for the retrievals and the inventory medians, respectively). A reduction is seen on Saturday (4.5 and 2.5 ktCO<sub>2</sub> h<sup>-1</sup>, respectively). The number of enhancements is also small on that day. Sunday only stands out by the low number of cases: the two medians are similar to the other days of the week because the detected cases on Sunday are much less distributed over the globe, even compared to Saturday: nearly all emission retrievals on Sunday are in China and South Africa (not shown), which drives both averages upward.

To study the diurnal cycle of the inferred emissions, we try to take advantage of the small number of OCO-3 enhancements. By dividing the day around 10:30 a.m. LST, we get 19 cases on both sides. The time is of the enhancement, not of the emissions: emissions are spread within the few hours before data acquisition (consistent with our processing of the inventory for the comparison, see Section 2.4). The inventory and the retrievals agree on an increase of the median emission from the first time slot to the second one (+11% and +24%, respectively). There are too few points to draw a general conclusion about the emission diurnal cycle in general, and we only retain here the agreement on the hierarchy between the two groups.

#### 4. Summary and Discussion

We have systematically looked for the imprint of large neighboring emitters in OCO-2 and OCO-3 retrievals with a plume cross-sectional inversion method and we have compared the results with the corresponding values of the EDGAR 6.0 global inventory disaggregated at hourly scale. We find a large correlation between the two, with more than one third of the inventory variance explained by the OCO retrievals. We interpret the similar statistics for OCO-2 and OCO-3 differences with the inventory as a sign of robustness. The standard deviation of the differences decreases with increasing retrieval values from 80% to 40%, likely because the signal-to-noise ratio for the retrieval increases with the emission level, and because larger emitters are usually better represented in the inventories than smaller ones. Interannual variations are marginal in both the OCO-2 retrievals and the inventory (in terms of median values and in log scale), but the seasonal cycle is marked in both, with a summer minimum. A Saturday minimum is consistently inferred from both, while the rest of the week shows less variations, including Sunday when significantly fewer large emitters are detected outside of China and South Africa. Finally, the few OCO-3 cases display a consistent increase in emissions from morning plumes to afternoon plumes.

These results indicate that the retrievals and the inventory distinguish the emission levels in a consistent manner. The retrievals and the inventory agree on the order of magnitude of the emissions, a feature which is highlighted by the better correlation in logscale for the 313 OCO-2 plumes. Much of their difference is random, so averaging can reveal consistent variations over time. However, the systematic differences, whether from the retrievals or from the inventory, are different for different emission levels, the retrievals being usually larger (smaller) than the inventory for values less (more) than 40 ktCO<sub>2</sub> h<sup>-1</sup>.

The difference between retrievals and inventory (with a 40%–80% standard deviation depending on the order of magnitude of the retrieval) provides an upper bound for the statistical uncertainty of the emission retrievals.

We expect a large contribution from the emission retrievals in these differences. Indeed, in addition to the uncertainty in the input XCO<sub>2</sub> retrievals, our cross-sectional inversion approach smooths the XCO<sub>2</sub> transect with four parameters only, neglects any vertical heterogeneity in the plume or in the XCO<sub>2</sub> averaging kernel, and relies on uncertain wind information without any adjustment of the wind direction.

The uncertainty in the 3D wind field also affects the way we relate the emission retrievals to the inventory. Further, we base our definition of a large emitter on a threshold on the value of the EDGAR v6.0 inventory in the

neighboring cells and in the hours preceding the passage of the satellite. The selection therefore relies heavily on the quality of the spatial and temporal disaggregation of national totals used in EDGAR. To give a simple example of the challenge, the Matimba coal-fired power plant in South Africa does not correspond to an emission hotspot in EDGAR for the years of study, while its plume is clearly visible in some OCO-2 orbits (Nassar et al., 2021). In the iron-and-steel-production sector, the Anshan Baode factory in Anshan, China, appears to be out of place by a few tens of kilometers, and nearby large sources of industrial emissions appearing in the recent Global Carbon Grid v1.0 are missing. With more precise spatial proxies, we would most likely select a larger number of plumes in the OCO-2 and OCO-3 records. In addition, if we knew the location of all large point sources well, it would be possible to optimize the effective wind direction (Nassar et al., 2017), even though uncertainty in the wind speed would remain problematic for the emission estimation.

The uncertainty in the spatial and temporal proxies of the inventory affects our emission attribution but should also explain another important share of the differences with the retrievals. In addition, we had to rely on inventory values for the year 2018 when comparing with observations from later years.

Given these three uncertainty terms (the retrievals, the inventory and the way we relate them together), actual differences with standard deviations of a few tens of percent are encouraging.

1050 EDGAR cells pass the emission floor value of “large emitters” at some time during the study period and one third of them are identified as contributing to an XCO<sub>2</sub> enhancement in the OCO-2 orbits: on average, each of these is identified twice in the seven-year record. The yield is therefore relatively small (it could perhaps be increased by reducing the size of the analysis window), but considering the narrow swath of the instrument that only allows for a small portion of the globe to be sampled each day in the absence of cloud, it gives factual evidence based on several hundreds of real cases over the globe, of the capacity of future CO<sub>2</sub> imagers like CO2M that has been documented mostly based on simulations (e.g., Lespinas et al., 2020) or a few tens of real cases (e.g., Zheng et al., 2020; Nassar et al., 2021). An implicit assumption is that future CO<sub>2</sub> imagers will allow retrieving XCO<sub>2</sub> with comparable quality to OCO-2, which is not necessarily granted considering different mission tradeoffs (Sierk et al., 2019). Restrictions on insolation (that affect observation in the high latitudes of the northern hemisphere and exclude it long before dusk and well after dawn), on cloud and aerosol loading (that affected, e.g., the detection of Covid-19-related patterns) with passive measurements in the near infrared/short-wave infrared spectral regions will stay. Also we note that very few emitters along the coasts could be identified in our study. Similarly, related to our analysis of the hourly version of EDGAR v6.0, some large emitters are not “large” all the time and may go under the detection threshold for a few hours every day or a few months every year. Last, the origin of some of the CO<sub>2</sub> within observed plumes will remain ambiguous, with a possible contribution from biofuels, the managed urban biosphere and human respiration over cities (Ciais et al., 2020; Miller et al., 2020).

Our method targets isolated plumes within a 200 km-wide orbit segment, but visual examination of the selected enhancement transects suggests that many of them are made of several plumes. This is confirmed by the fact that half of the enhancements correspond to more than one large emitter in EDGAR within a few hours windward of the enhancements. Our emission retrievals therefore correspond to values aggregated in space (typically over a few thousand km<sup>2</sup>) and along sectors, with a composition that varies with the wind. There is temporal aggregation as well, but limited to a few hours (because of the rapid dispersion of anthropogenic emission plumes; see, e.g., Broquet et al., 2018). The product uncertainty is large, but appears to be mostly random, so that trends can be robustly calculated if enough data is available. At this stage, this emerging asset is guiding the evolution of such a computationally light XCO<sub>2</sub>-driven emission product toward a transparent indicator of trends in large FF emissions for well-observed administrative regions of the size of Kuwait or larger (>10<sup>4</sup> km<sup>2</sup>) and for an appropriate range of hours in the day. It will not fit the reporting to the United Nations Framework Convention on Climate Change (UNFCCC) directly, but can serve both the scientific debate and the general public.

### Data Availability Statement

OCO-2 and OCO-3 retrievals were obtained from the data archive maintained at the NASA Goddard Earth Science Data and Information Services Center ([https://oco2.gesdisc.eosdis.nasa.gov/data/OCO2\\_DATA/OCO2\\_L2\\_Lite\\_FP.10r/](https://oco2.gesdisc.eosdis.nasa.gov/data/OCO2_DATA/OCO2_L2_Lite_FP.10r/) and [https://oco2.gesdisc.eosdis.nasa.gov/data/OCO3\\_DATA/OCO3\\_L2\\_Lite\\_FP.10r/](https://oco2.gesdisc.eosdis.nasa.gov/data/OCO3_DATA/OCO3_L2_Lite_FP.10r/)). Wind data on ERA5 model levels is available from <https://confluence.ecmwf.int/display/CKB/How+to+download+ERA5>. The EDGAR inventory can be downloaded from <http://data.europa.eu/89h/97a67d67-c62e-4826-b873-9d972c4f670b>.

### Acknowledgments

This study 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<sub>2</sub> service). It was granted access to the HPC resources of TGCC under the allocation A0110102201 made by GENCI. OCO-2 and OCO-3 retrievals were produced by the OCO-2 and OCO-3 project at the Jet Propulsion Laboratory, California Institute of Technology. We thank Greet Janssens-Maenhout for her encouraging feedback throughout this study, François-Marie Bréon and two anonymous reviewers for their careful reading of an intermediate version of the text, and Alexandre Danjou for interesting discussions on this topic.

### References

- Baker, D. F., Bell, E., Davis, K. J., Campbell, J. F., Lin, B., & Dobler, J. (2021). A new exponentially-decaying error correlation model for assimilating OCO-2 column-average CO<sub>2</sub> data, using a length scale computed from airborne lidar measurements [preprint]. *Geoscientific Model Development Discussions*. <https://doi.org/10.5194/gmd-2020-444>
- Broquet, G., Bréon, F.-M., Renault, E., Buchwitz, M., Reuter, M., Bovensmann, H., et al. (2018). The potential of satellite spectro-imagery for monitoring CO<sub>2</sub> emissions from large cities. *Atmospheric Measurement Techniques*, *11*, 681–708. <https://doi.org/10.5194/amt-11-681-2018>
- Brunner, D., Kuhlmann, G., Marshall, J., Clément, V., Fuhrer, O., Broquet, G., et al. (2019). Accounting for the vertical distribution of emissions in atmospheric CO<sub>2</sub> simulations. *Atmospheric Chemistry and Physics*, *19*, 4541–4559. <https://doi.org/10.5194/acp-19-4541-2019>
- Chevallier, F. (2021). Fluxes of carbon dioxide from managed ecosystems estimated by national inventories compared to atmospheric inverse modeling. *Geophysical Research Letters*, *48*, e2021GL093565. <https://doi.org/10.1029/2021gl093565>
- Chevallier, F., Remaud, M., O'Dell, C. W., Baker, D., Peylin, P., & Cozic, A. (2019). Objective evaluation of surface- and satellite-driven carbon dioxide atmospheric inversions. *Atmospheric Chemistry and Physics*, *19*, 14233–14251. <https://doi.org/10.5194/acp-19-14233-2019>
- Chevallier, F., Zheng, B., Broquet, G., Ciais, P., Liu, Z., Davis, S. J., et al. (2020). Local anomalies in the column-averaged dry air mole fractions of carbon dioxide across the globe during the first months of the coronavirus recession. *Geophysical Research Letters*, *47*, e2020GL090244. <https://doi.org/10.1029/2020gl090244>
- Ciais, P., Wang, Y., Andrew, R., Bréon, F.-M., Chevallier, F., Broquet, G., et al. (2020). Biofuel burning and human respiration bias on satellite estimates of fossil fuel CO<sub>2</sub> emissions. *Environmental Research Letters*, *15*, 074036. <https://doi.org/10.1088/1748-9326/ab7835>
- Crippa, M., Guizzardi, D., Muntean, M., Schaaf, E., Lo Vullo, Eleonora, et al. (2021). EDGAR v6.0 greenhouse gas emissions [Dataset]. European Commission, Joint Research Centre (JRC) PID. Retrieved from <http://data.europa.eu/89h/97a67d67-c62e-4826-b873-9d972c4f670b>
- Crippa, M., Solazzo, E., Huang, G., Guizzardi, D., Koffi, E., Muntean, M., et al. (2020). High resolution temporal profiles in the emissions database for global atmospheric research. *Scientific Data*, *7*, 121. <https://doi.org/10.1038/s41597-020-0462-2>
- Deng, Z., Ciais, P., Tzompa-Sosa, Z. A., Saunio, M., Qiu, C., Tan, C., et al. (2021). Comparing national greenhouse gas budgets reported in UNFCCC inventories against atmospheric inversions [preprint]. *Earth System Science Data Discussions*. <https://doi.org/10.5194/essd-2021-235>
- Eldering, A., Taylor, T. E., O'Dell, C. W., & Pavlick, R. (2019). The OCO-3 mission: Measurement objectives and expected performance based on 1 year of simulated data. *Atmospheric Measurement Techniques*, *12*, 2341–2370. <https://doi.org/10.5194/amt-12-2341-2019>
- ESA. (2019). Copernicus CO<sub>2</sub> monitoring mission requirements document, version 2.0 of 27/09/19, ESA Earth and Mission Science Division document ref. EOP-SM/3088/YM-ym. Retrieved from [https://esamultimedia.esa.int/docs/EarthObservation/CO2M\\_MRD\\_v2.0\\_Issued20190927.pdf](https://esamultimedia.esa.int/docs/EarthObservation/CO2M_MRD_v2.0_Issued20190927.pdf)
- Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., et al. (2020). The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, *1*, 51. <https://doi.org/10.1002/qj.3803>
- Janssens-Maenhout, G., Crippa, M., Guizzardi, D., Muntean, M., Schaaf, E., Dentener, F., et al. (2019). EDGAR v4.3.2 Global Atlas of the three major greenhouse gas emissions for the period 1970–2012. *Earth System Science Data*, *11*, 959–1002. <https://doi.org/10.5194/essd-11-959-2019>
- Janssens-Maenhout, G., Pinty, B., Dowell, M., Zunker, H., Andersson, E., Balsamo, G., et al. (2020). Towards an operational anthropogenic CO<sub>2</sub> emissions monitoring and verification support capacity. *Bulletin of the American Meteorological Society*, *101*. <https://doi.org/10.1175/BAMS-D-19-0017.1>
- Lespinas, F., Wang, Y., Broquet, G., et al. (2020). The potential of a constellation of low earth orbit satellite imagers to monitor worldwide fossil fuel CO<sub>2</sub> emissions from large cities and point sources. *Carbon Balance and Management*, *15*, 18. <https://doi.org/10.1186/s13021-020-00153-4>
- Miller, J. B., Lehman, S. J., Verhulst, K. R., Miller, C. E., Duren, R. M., yadav, V., et al. (2020). Large and seasonally varying biospheric CO<sub>2</sub> fluxes in the Los Angeles megacity revealed by atmospheric radiocarbon. *Proceedings of the National Academy of Sciences of the United States of America*, *117*, 26681–26687. <https://doi.org/10.1073/pnas.2005253117>
- Nassar, R., Hill, T. G., McLinden, C. A., Wunch, D., Jones, D. B. A., & Crisp, D. (2017). Quantifying CO<sub>2</sub> emissions from individual power plants from space. *Geophysical Research Letters*, *44*, 10045–10053. <https://doi.org/10.1002/2017gl074702>
- Nassar, R., Mastrogiacomo, J.-P., Bateman-Hemphill, W., McCracken, C., MacDonald, C. G., Hill, T., et al. (2021). Advances in quantifying power plant CO<sub>2</sub> emissions with OCO-2. *Remote Sensing of Environment*, *264*. <https://doi.org/10.1016/j.rse.2021.112579>
- Osterman, G., O'Dell, C., Eldering, A., Fisher, B., Crisp, D., Cheng, C., et al. (2020). *Orbiting Carbon Observatory-2 & 3 (OCO-2 & OCO-3)—data product user's guide, operational level 2 data versions 10 and lite file version 10 and VEarly. Version 1.0, revision A, 8 June 2020*. Jet Propulsion Laboratory. Retrieved from [https://docserv01.gesdisc.eosdis.nasa.gov/public/project/OCO/OCO2\\_OCO3\\_B10\\_DUG.pdf](https://docserv01.gesdisc.eosdis.nasa.gov/public/project/OCO/OCO2_OCO3_B10_DUG.pdf)
- Reuter, M., Buchwitz, M., Schneising, O., Krautwurst, S., O'Dell, C. W., Richter, A., et al. (2019). Towards monitoring localized CO<sub>2</sub> emissions from space: Co-located regional CO<sub>2</sub> and NO<sub>2</sub> enhancements observed by the OCO-2 and S5P satellites. *Atmospheric Chemistry and Physics*, *19*, 9371–9383. <https://doi.org/10.5194/acp-19-9371-2019>
- Sierk, B., Bézy, J.-L., Löscher, A., & Meijer, Y. (2019). The European CO<sub>2</sub> monitoring mission: Observing anthropogenic greenhouse gas emissions from space. In *International Conference on Space Optics — ICSSO 2018*, Proc. SPIE 11180, 111800M. <https://doi.org/10.1117/12.2535941>
- Taylor, T. E., Eldering, A., Merrelli, A., Kiel, M., Somkuti, P., Cheng, C., et al. (2020). *OCO-3 early mission operations and initial (vEarly) XCO2 and SIF retrievals*, *251*, 112032. <https://doi.org/10.1016/j.rse.2020.112032>
- Wunch, D., Toon, G. C., Blavier, J. F. L., Washenfelder, R. A., Notholt, J., Connor, B. J., et al. (2011). The total carbon column observing Network. *Philosophical Transactions of the Royal Society A*, *369*, 2087–2112. <https://doi.org/10.1098/rsta.2010.0240>
- Zheng, B., Chevallier, F., Ciais, P., Broquet, G., Wang, Y., Lian, J., & Zhao, Y. (2020). Observing carbon dioxide emissions over China's cities and industrial areas with the Orbiting Carbon Observatory-2. *Atmospheric Chemistry and Physics*, *20*, 8501–8510. <https://doi.org/10.5194/acp-20-8501-2020>