EYE-CLIMA
Verifying emissions of climate forcers

Report on CO_2 and CH_4 Satellite Datasets: OCO-2 CO_2 Comparison

DELIVERABLE 1.3

Author(s):	M. Hilker, H. Bösch
Date of submission:	25-09-2025
Version:	1.0
Responsible partner:	University of Bremen
Deliverable due date:	30-06-2025
Dissemination level:	Public

Call:	HORIZON-CL5-2022-D1-02
Topic:	Climate Sciences and Responses
Project Type:	Research and Innovation Action
Lead Beneficiary:	NILU - Norsk Institutt for Luftforskning

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101081395

Document History

Version	Date	Comment	Modifications made by
0.1	05-09-2025	First Draft	M. Hilker, H. Bösch
0.2	11-09-2025	Internal review	A. Fortems-Cheiney
1.0	25-09-2025	Final revised version	M. Hilker, H. Bösch

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101081395

Summary

Global measurements of the atmospheric CO_2 column are available since 2002 when SCIAMACHY onboard ESA ENVISAT was launched, this was followed by the Japanese GOSAT satellite in 2009 (and its successor GOSAT-2 in 2018), and in 2014 by the NASA Orbiting Carbon Observatory (OCO-2). In comparison to previous missions, OCO-2 has the best performance in terms of coverage, spatial resolution and uncertainties. In this report, we have evaluated the available CO_2 data products of the OCO-2 mission: the official NASA product and the IUP FOCAL product. The evaluation includes an intercomparison of the data products and an assessment against ground-based reference measurements from the TCCON network for the global and the European domain.

Globally, both data products have roughly the same number of data points but there are clear differences in coverage with the NASA product having more datapoints over most land masses with the exception of North Africa and the Arabian Peninsula. Over Europe, both data products have good coverage in summer but, as expected, low/no coverage in high latitude winter. The NASA product has distinctively more data over European land masses. Over Europe, we obtain a clear land sea difference in XCO_2 between both datasets with FOCAL being lower than the NASA dataset over water. The existence of small but systematic differences between both datasets is confirmed with correlation plots. XCO_2 time series of both datasets for the time period from 2014 to 2025 agree well in general, with some differences in the seasonal amplitude, which mostly disappear when using only the same data and the same a priori, although a small tendency remains for the NASA product to show a slightly larger seasonal amplitude. Comparison to the TCCON network for sites in Europe show a good match of both datasets with TCCON data with correlation coefficients almost always above 0.9. Biases for the different stations are mostly below 0.5 ppm but can be larger for the Garmisch, Harwell and Nicosia sites. We find that the NASA dataset has lower biases compared to the FOCAL product for all sites except for Harwell. The NASA product has also a slightly lower scatter for all sites and a noticeable larger number of datapoints over each TCCON site.

Overall, we find that two high quality data products exist for OCO-2. When we focus on Europe, we find higher data density over European land masses by the NASA product and a better agreement with European TCCON reference sites. We also find a systematic difference in land-sea contrast in the XCO_2 datasets which is difficult to assess with TCCON data and we cannot tell which data product is a better choice in this aspect.

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101081395

TABLE OF CONTENTS

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101081395

1. Introduction

CO_2 is an important greenhouse gas and measurement of its atmospheric distribution provides information on surface-atmosphere fluxes. Global measurements of total atmospheric CO_2 columns are available since 2002 when SCIAMACHY onboard ESA ENVISAT was launched (Bovensmann et al., 1999). This has been followed by the Japanese GOSAT satellite in 2009 (and its successor GOSAT-2 in 2018) (Kuze et al., 2009) and in 2014 by the NASA Orbiting Carbon Observatory (OCO-2) (Crisp et al., 2008).

Very recently, the CNES/UKA MicroCarb (Cansot et al., 2023) mission and the Japanese GOSAT-GW (Tanimoto and Matsunaga, 2024) mission have been launched but no data products are yet available for either mission.

In comparison to previous missions, OCO-2 has the best performance in terms of coverage, spatial resolution and uncertainties and has been successfully used for a number of regional scale flux inversion studies (e.g., Byrne et al., 2023; Crowell et al., 2019; Wang et al., 2020) as well as in EYE-CLIMA (D3.1 and D3.2).

In this report, we focus exclusively on evaluating the available CO_2 data products from the OCO-2 mission. The evaluated CO_2 data products for OCO-2 are the official NASA product and the IUP FOCAL product. The evaluation includes an intercomparison of the data products and an assessment against ground-based reference measurements from the TCCON network for a global and the European domain.

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101081395

2. Datasets

2.1 The NASA OCO-2 dataset

2.1.1 Description

The NASA OCO-2 data product (OCO-2/OCO-3 Science Team et al., 2024) is inferred with an optimal estimation–based full-physics retrieval algorithm that simultaneously retrieves CO_2 profiles along with aerosol and surface parameters. The algorithm employs a detailed forward model to describe the transfer of light through the atmosphere and employs the Low-Stream-Interpolation method to accelerate the computationally intensive multiple scattering radiative transfer calculations (O'Dell, 2010).

For land measurements, the retrieval retrieves 64 parameters that form the so-called state vector, which includes a 20-layer CO_2 profile. Ocean measurements use a 60-element state vector with a different surface reflectance parameterization.

The aerosol a priori is set-up according to the data from GEOS-5 Forward Product for Instrument Teams (FP-IT) product. The CO_2 a priori profiles used are similar to those in the GGG2020 TCCON retrieval. These profiles are based on flask measurements from the Mauna Loa and American Samoa NOAA sites, representing tropical CO_2 concentrations adjusted for atmospheric transport time using an empirical age-of-air approximation.

In post-processing, quality filtering and bias correction is applied to the retrieved data (O'Dell et al., 2018). The quality filter is threshold-based, and the bias correction uses a multivariate linear regression. Parameters for both steps are derived by comparison against a truth proxy dataset consisting of TCCON measurements and a multi-model ensemble.

2.1.2 Used Data

In this report, we use the OCO2_L2_Lite_FP v11.2r data product, which was downloaded from https://oco2.gesdisc.eosdis.nasa.gov/data/OCO2_DATA/OCO2_L2_Lite_FP.11.2r/. The data is stored in netCDF4 files and includes bias-corrected XCO_2 values, uncertainties, quality flags, a priori values, and averaging kernels. In accordance with the user guide, only soundings with xco2_quality_flag = 0 are used. Only data acquired before 2024-03-01 are used.

2.2 The FOCAL IUP OCO-2 dataset

2.2.1 Description

The FOCAL OCO-2 data product (Reuter et al., 2025) also uses an optimal estimation–based full-physics retrieval algorithm to infer XCO_2. Atmospheric scattering is approximated by a single thin scattering layer, enabling an analytical and computationally efficient solution to the radiative transfer equation.

A 40-element state vector is used for all measurements, including a 5-layer CO_2 profile. A priori CO_2 profiles are provided by the $SLIMCO_2$ model (Noël et al., 2022).

In post-processing, both quality filtering and bias correction are applied. The quality filter is threshold-based, with thresholds derived from a two-month subset of retrievals. Bias correction is performed using a Random Forest Regressor, trained on a combination of TCCON and $SLIMCO_2$ data as reference truth.

2.2.2 Used Data

In this report, we use FOCAL OCO-2 version 11, which was downloaded from https://www.iup.uni-bremen.de/~mreuter/focal.php. The product is stored in netCDF4 files containing bias-corrected XCO_2 values, uncertainties, a priori profiles, and averaging kernels. Only soundings that pass the quality filters

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101081395

are included in the files, so no additional filtering is necessary. The dataset includes OCO-2 measurements up to 2024-03-01.

2.3 TCCON Data

2.3.1 Description

The Total Carbon Column Observing Network (TCCON) is a global network of ground-based Fourier Transform Spectrometers. TCCON measures direct solar spectra in the near- to shortwave-infrared (NIR to SWIR) range and provides column-averaged dry-air mole fractions of CO_2, CH_4, CO, N_2O, H_2O, HDO, and HF (Wunch et al., 2011). For CO_2, the GGG2020 release reports a total measurement uncertainty of approximately 0.16% (~0.6 ppm) (Laughner et al., 2024).

2.3.2 Used Data

We use data from version GGG2020 in this report. The TCCON data were obtained from the TCCON Data Archive hosted by CaltechDATA at https://tccondata.org. The data is provided in netCDF format and contains the retrieved values, and ancillary data like surface pressure, temperature, averaging kernels and a priori profiles. For this analysis, only European sites were selected, as summarised in Table 1.

Station	**Reference**
Bialystok, Poland	(Petri et al., 2024)
Bremen, Germany	(Notholt et al., 2022)
Garmisch, Germany	(Sussmann and Rettinger, 2025)
Harwell, UK	(Weidmann et al., 2023)
Karlsruhe, Germany	(Hase et al., 2024)
Nicosia, Cyprus	(Petri et al., 2024)
Ny-Ålesund, Norway	(Buschmann et al., 2022)
Orléans, France	(Warneke et al., 2024)
Paris, France	(Té et al., 2022)
Sodankylä, Finland	(Kivi et al., 2022)

Table 1: List of TCCON stations used in this report.

2.4 SLIMCO$_2$

2.4.1 Description

The SLIMCO$_2$ model (Noël et al., 2022) is designed to provide computationally efficient estimates of atmospheric CO_2 profiles, column averages and a corresponding covariance estimate for use as a priori inputs in remote sensing retrievals. SLIMCO$_2$ is based on a climatology database derived from long-term model data. The climatology covers one year with 36 timesteps and is spatially gridded on a 3° by 2° grid with 20 vertical layers. The database also includes a NOAA-derived annual growth rate table for CO_2 to extend applicability beyond the original data period.

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101081395

2.4.2 Used Data

The $SLIMCO_2$ version used for the analysis in this report is based on the CAMS CO2 v22r1 (Copernicus Atmosphere Monitoring Service, 2020). This $SLIMCO_2$ differs from the version used in the FOCAL retrieval, which is based on CAMS CO2 v20r2 (Copernicus Atmosphere Monitoring Service, 2020).

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101081395

3. Intercomparison of NASA and FOCAL OCO-2 data products

3.1 Methodology

The NASA and FOCAL OCO-2 data products, introduced in section 2, were compared against each other using three prepared datasets, referred to in this report as “all soundings”, “same sampling” and “same sampling & a priori.”

- “All soundings” includes all quality-filtered soundings from each product.
- “Same sampling” consists only of soundings that are present in both products.
- “Same sampling & a priori” uses the same soundings as "same sampling” but with a common a priori applied to both datasets.

The replacement of the original a priori profiles in both datasets was carried out using the method described by Wunch et al. (2011), which utilizes the averaging kernels and a priori information provided in the data files.

SLIMCO_2 was selected as the common a priori, as it allows for straightforward generation of an CO_2 profile for each individual sounding.

3.1.1 Maps

Several maps were produced to analyse the datasets. These include monthly and yearly average XCO_2 column maps for all three datasets, as well as maps showing the monthly and yearly number of measurements for the "all soundings" dataset. All values were averaged or counted on a 2° × 2° grid.

Maps are available for both global coverage and a European subset. The European region is defined as spanning 30°N to 72°N in latitude and 20°W to 50°E in longitude.

3.1.2 Scatter plots

The “same sampling” dataset was used to generate scatter plots for a direct comparison of the NASA and FOCAL XCO_2 data. A linear regression was performed on the individual soundings, but the results are visualized as 2D histograms to highlight data density. Each figure includes the regression equations and the correlation coefficients.

3.1.3 Time series

Monthly XCO_2 averages were calculated for both global and European domains across all three datasets to analyse the temporal evolution of the two products. For the “all soundings” dataset, additional time series were generated for the Northern Hemisphere, Southern Hemisphere, and the Tropics.

The regions are defined as follows:

- Southern Hemisphere: Latitude < 0°
- Northern Hemisphere: Latitude > 0°
- Tropics: 23.4°S < Latitude < 23.4°N
- Europe: 30°N < Latitude < 72°N and 20°W < Longitude < 50°E

3.2 Results

This report uses the year 2020 as a representative example. To illustrate seasonal variations, the months of February and August have been selected.

This project has received funding from the European Union’s Horizon Europe research and innovation programme under grant agreement No 101081395

While data has been generated for all months and years from September 2014 to February 2024, only a subset is presented here for brevity. The full dataset is available at: https://nc.uni-bremen.de/index.php/s/t4R9Wg9sLjM8eQS

3.2.1 Coverage

Temporal and spatial coverage is a key aspect of the XCO_2 data products and can vary significantly across both space and time. Although both products are derived from the same OCO-2 measurements, the retrieval algorithms they respond differently to the specific observation geometry and atmospheric conditions of each sounding. In some cases, the retrieval may fail to converge, and each product applies distinct quality filtering strategies, which remove different subsets of the soundings. As a result, the overall coverage differs between the two products.

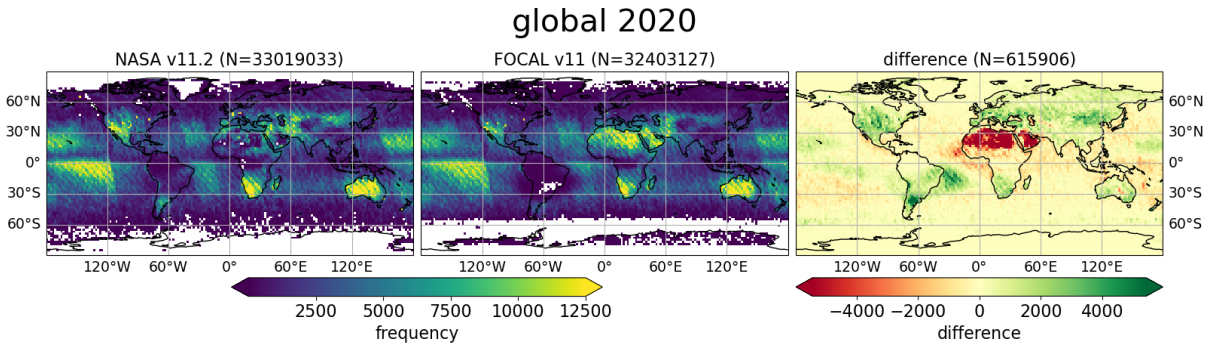

Figure 1: Comparison of global coverage in 2020 for both OCO-2 data products. Left: NASA v11.2, centre: FOCAL v11, right: Difference (NASA – FOCAL).

Figure 1 displays the total number of OCO-2 measurements available in 2020 after applying quality filters. The NASA product includes 33,019,033 valid measurements, 1.9% (or 615,906 soundings) more than FOCAL. Overall, the NASA product has more soundings over land masses, except in North Africa and the Arabian Peninsula, where FOCAL has noticeably more data.

Coverage is low for both products at latitudes above 60°N and below 60°S. In the Northern Hemisphere, the coverage is generally similar, although NASA has no data over Greenland. In the Southern Hemisphere, FOCAL's ocean coverage ends around 55°S, but it includes some data over Antarctica. NASA, on the other hand, provides sparse ocean coverage and almost no data over the Antarctic continent.

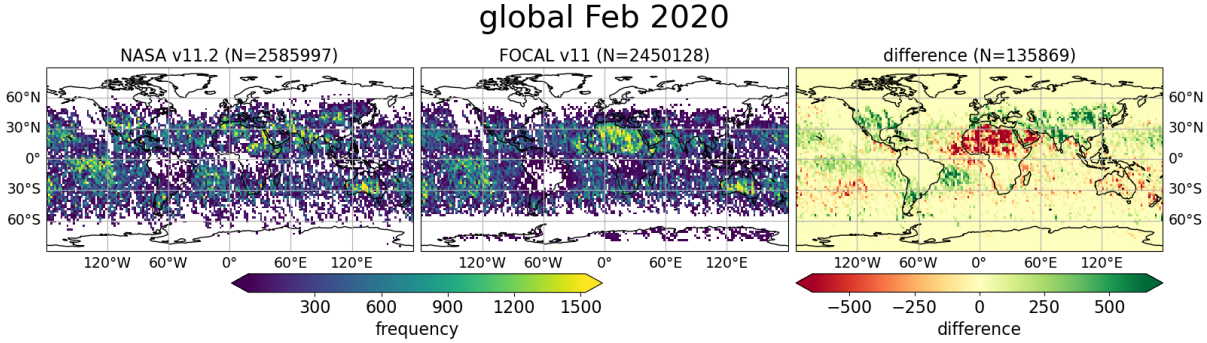

Figure 2: Same as Figure 1 but for measurements in February 2020.

Monthly variations follow similar spatial patterns but are influenced by seasonal effects that impact the availability of soundings. Figure 2 shows data for February 2020, while Figure 3 shows August 2020. August has a higher total number of measurements, especially at latitudes north of 60°N. However, fewer measurements are available near the equator in Africa during August. FOCAL and NASA have a

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101081395

similar number of total data points with FOCAL has a slightly lower number in February and a larger number in August compared to NASA.

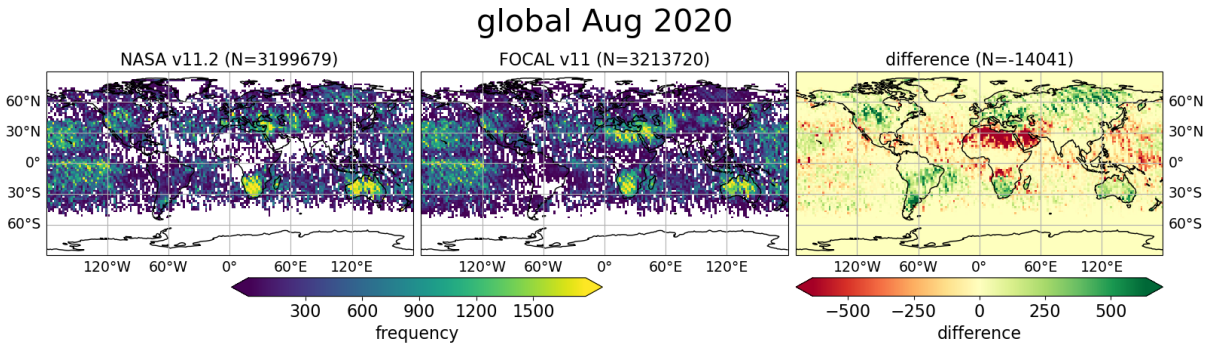

Figure 3: Same as Figure 1 but for measurements in August 2020.

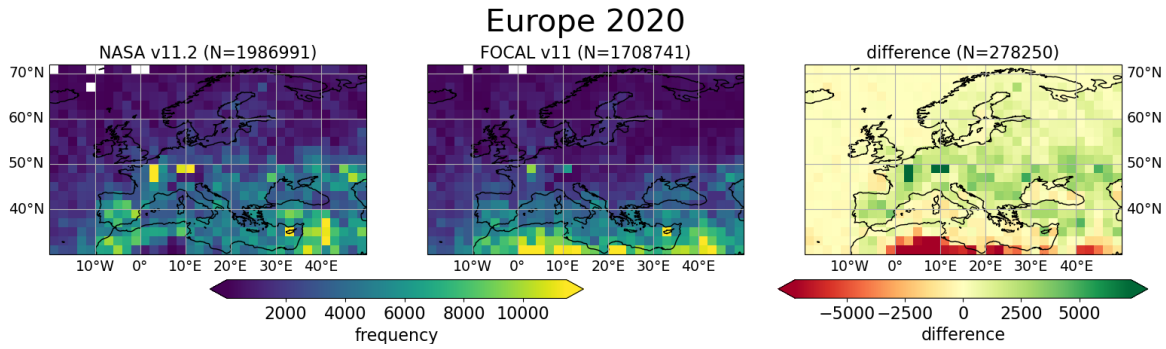

Figure 4: Same as Figure 1 but only for measurements in Europe.

Figure 4 focuses on the European region, which also includes the northern coast of Africa. The pattern mirrors the global data shown in Figure 1. NASA generally provides more measurements over land, except in northern Africa where FOCAL has substantially more data. In total, NASA reports 1,986,991 soundings in this region, 16.3% more than FOCAL, a significantly larger margin than observed globally. This may be attributed to the higher land fraction in this region and the inclusion of a small part of northern Africa.

In February 2020 (Figure 5), NASA covers more of the region between 50°N and 60°N, while FOCAL's coverage mostly ends at 50°N. FOCAL retains more data over Algeria, whereas NASA has more data elsewhere. In August (Figure 6), both products offer broad coverage across most of Europe.

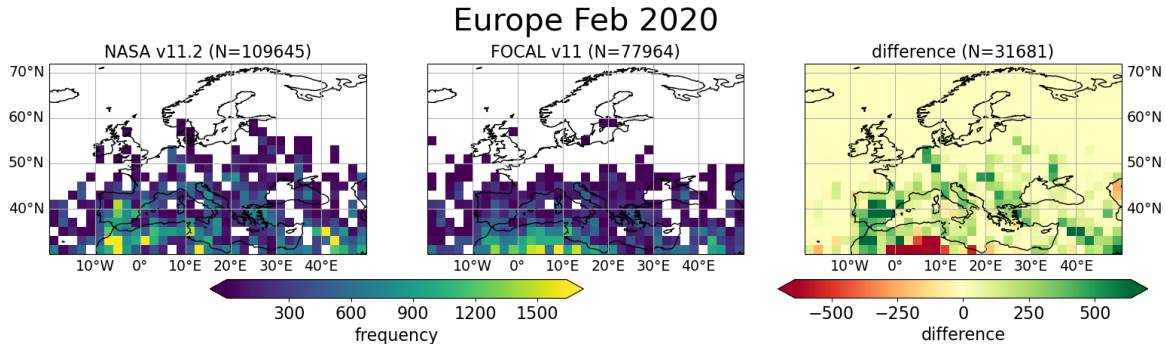

Figure 5: Same as Figure 1 but only for measurements in Europe in February 2020.

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101081395

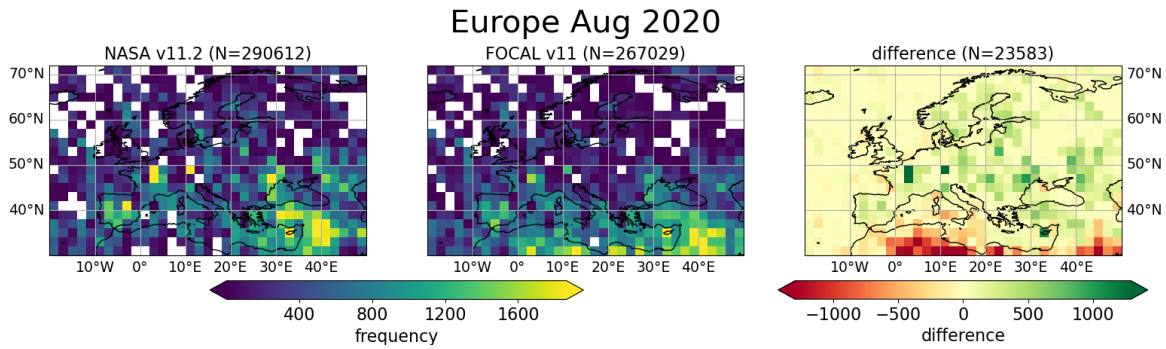

Figure 6: Same as Figure 1 but only for measurements in Europe in August 2020.

Figure 7 shows the monthly time series of the number of measurements from both the NASA and FOCAL data products, covering the period from September 2014 to February 2024. Over this time range, FOCAL includes 299,656,320 soundings globally, 2.8% more than NASA. Most of this extra data stems from measurements taken before 2020. In the European region, however, NASA has a total of 15,990,030 soundings, 11.8% more than FOCAL.

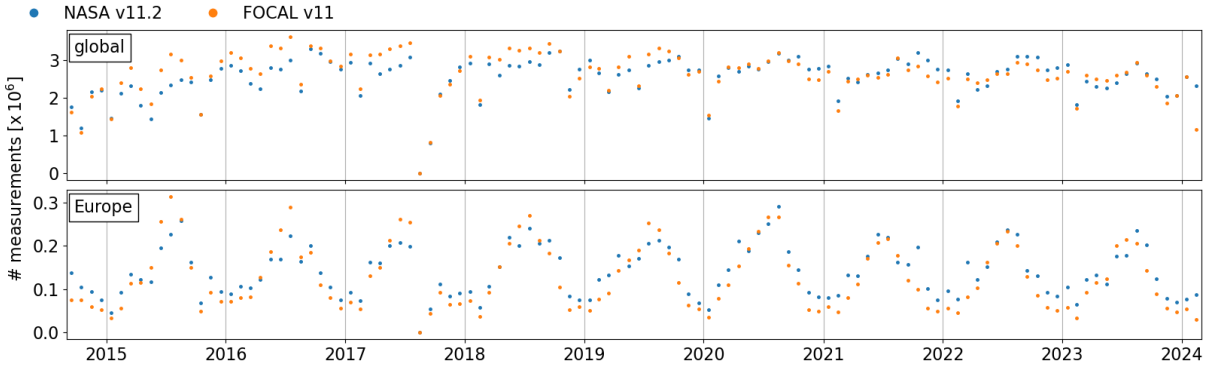

Figure 7: Time series of monthly number of measurements for the NASA and FOCAL data products. Top: global, bottom: Europe subset.

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101081395

3.2.2 XCO_2 Maps

Figure 8 shows global average XCO_2 maps for the two datasets introduced in section 3. The "all soundings" maps reveal substantial regional differences between the NASA and FOCAL products. FOCAL shows significantly higher XCO_2 values south of the Sahara in Africa and over India compared to NASA.

This pattern does not directly match the coverage shown in Figure 1 but can still be impacted by the coverage. Using the "same sampling" dataset eliminates sampling as cause for differences in the gridded maps by ensuring identical soundings are used. The NASA map shows only small changes, while FOCAL values over Africa and India are reduced. This indicates, that soundings which cause these differences are mostly missing in the NASA data. This may suggest that FOCAL does not flag all soundings of poor quality in this region but regional validation sites would be needed to assess this. The difference maps for the "same sampling" case exhibits a significantly reduced standard deviation of 0.53 ppm (1.08 ppm for "all soundings").

The last row in Figure 8 illustrates the impact of the different a priori profiles used by both products. Applying common a priori profiles in the "same sampling & a priori" dataset reduces the mean global difference from 0.24 ppm to 0.08 ppm and the variability slightly reduces to 0.49 ppm.

Maps for February 2020 (Figure 9) display a similar pattern to the full-year maps in Figure 8. The characteristic differences in the "all soundings" dataset is visible again, but with smaller reductions in variability for the "same sampling" and "same sampling & a priori" cases. However, Figure 10 (August 2020) shows a different picture with a different XCO_2 pattern over Africa for FOCAL. Application of "same sampling" and "same sampling & a priori" conditions lead to a smaller reduction in the standard deviation (from 1.01 ppm to 0.72 ppm).

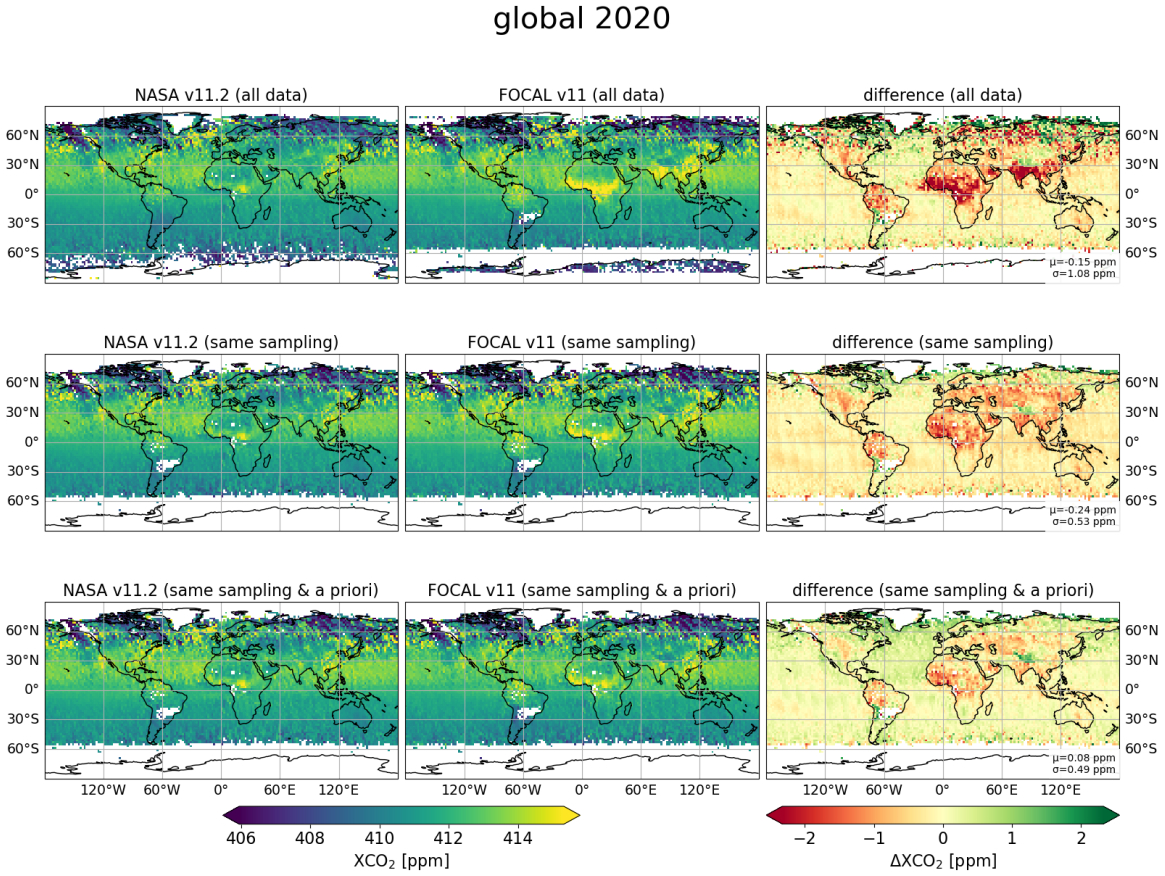

Figure 8: Maps of global average XCO2 concentrations for 2020. First row: all data for each product after quality filtering, second row: only soundings that appear in both products, third row: additionally corrected to use same a priori. First column:

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101081395

NASA, second column: FOCAL, third column: difference (NASA – FOCAL). Data are averaged on a 2° x 2° grid. μ and σ are mean and standard deviation of the gridded differences.

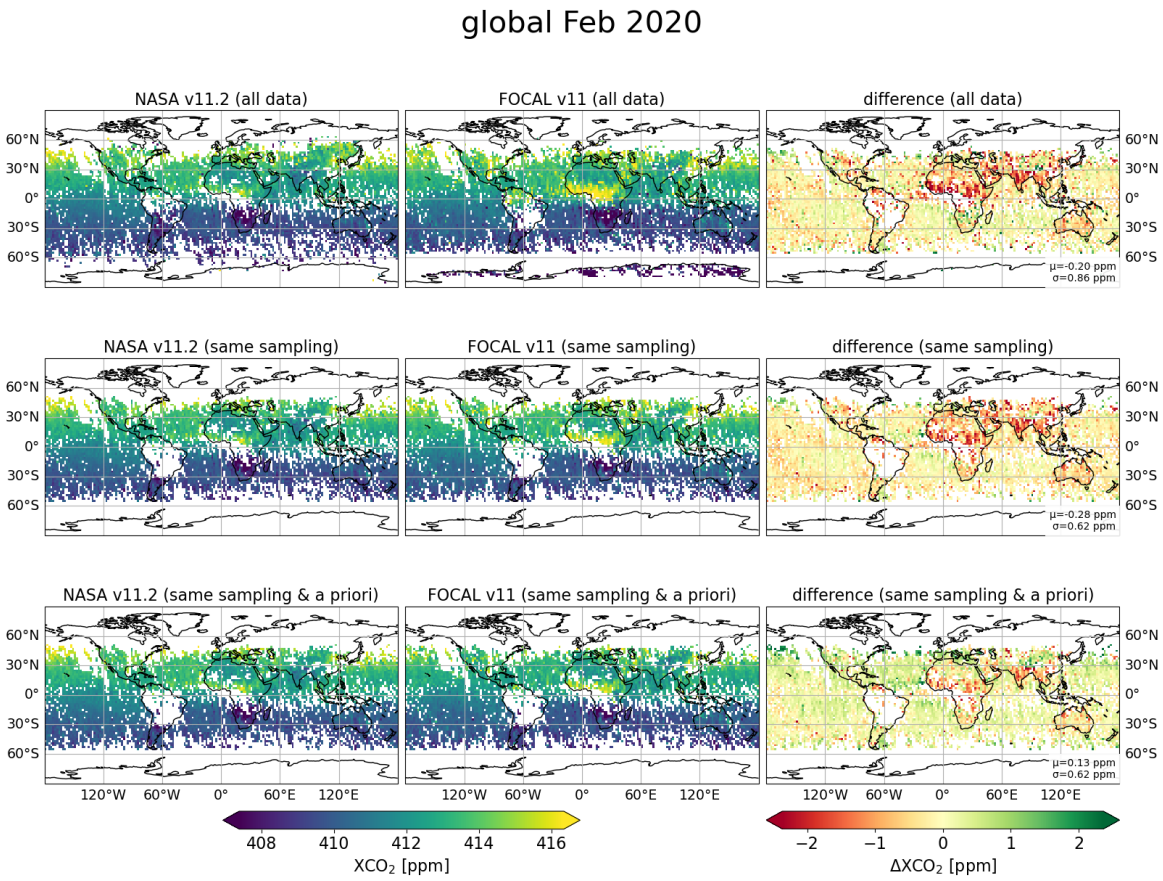

Figure 9: Same as Figure 8 but for soundings in February 2020.

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101081395

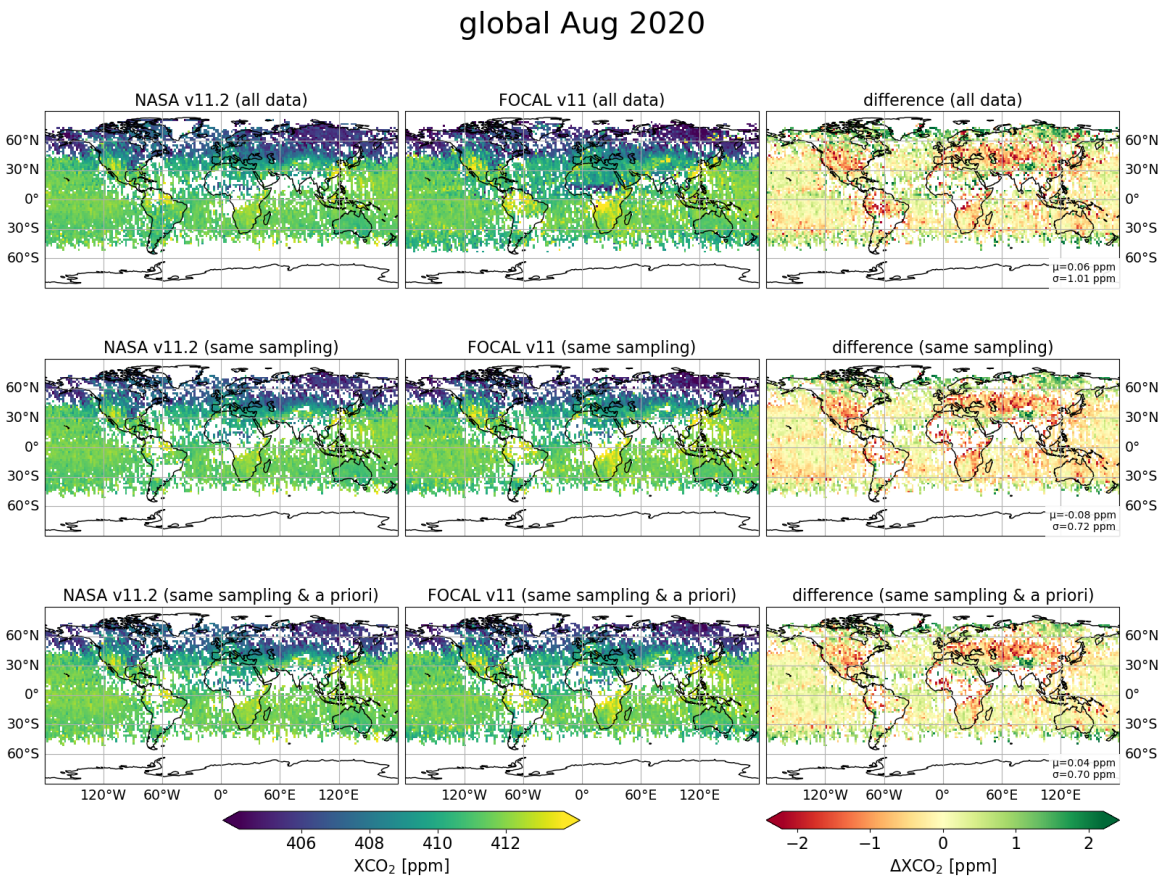

Figure 10: Same as Figure 9 but for soundings in August 2020.

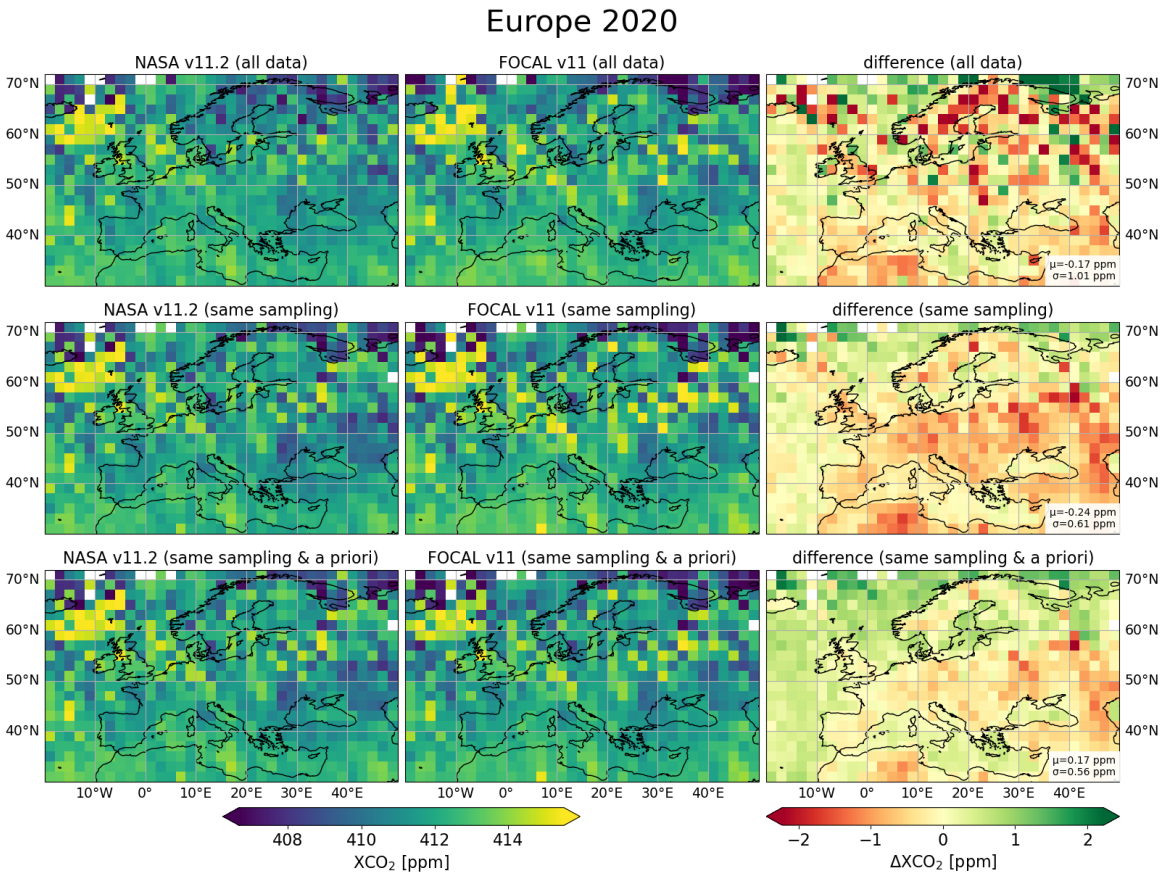

Figure 11: Same as Figure 8 for soundings in Europe.

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101081395

Figure 11 presents mean XCO_2 maps for the Europe region in 2020. The XCO_2 fields for "all soundings" show a characteristic region with high values between Iceland and Britain. The "all soundings" difference map also appears relatively noisy above 50°N with large fluctuations from pixel to pixel. The "same data" difference map shows smoother differences and a small land-sea bias becomes visible. At the same time, the region of high values between Iceland and Britain becomes less pronounced. FOCAL has on average 0.22 ppm higher values over land while NASA has 0.62 higher values over water. The "same sampling & a priori" maps reduce the differences and a land-sea contrast is visible. The standard deviation shows a similar reduction as the global data.

In February 2020 (Figure 12), both products provide coverage over only a small area of Europe, making it difficult to identify meaningful patterns. Similar to the August 2020 maps in Figure 10, the Europe region in August shows no significant improvement when using the "same sampling" or "same sampling & a priori" datasets. FOCAL still reports higher XCO_2 over land, and NASA retains higher values over water.

The maps in this section show that there are systematic regional XCO_2 differences between the NASA and FOCAL OCO-2 products. A portion of these differences is caused by different sampling. Soundings that cause the biggest differences are mostly missing in the NASA product. Using a common a priori also results in a reduced variability of the differences.

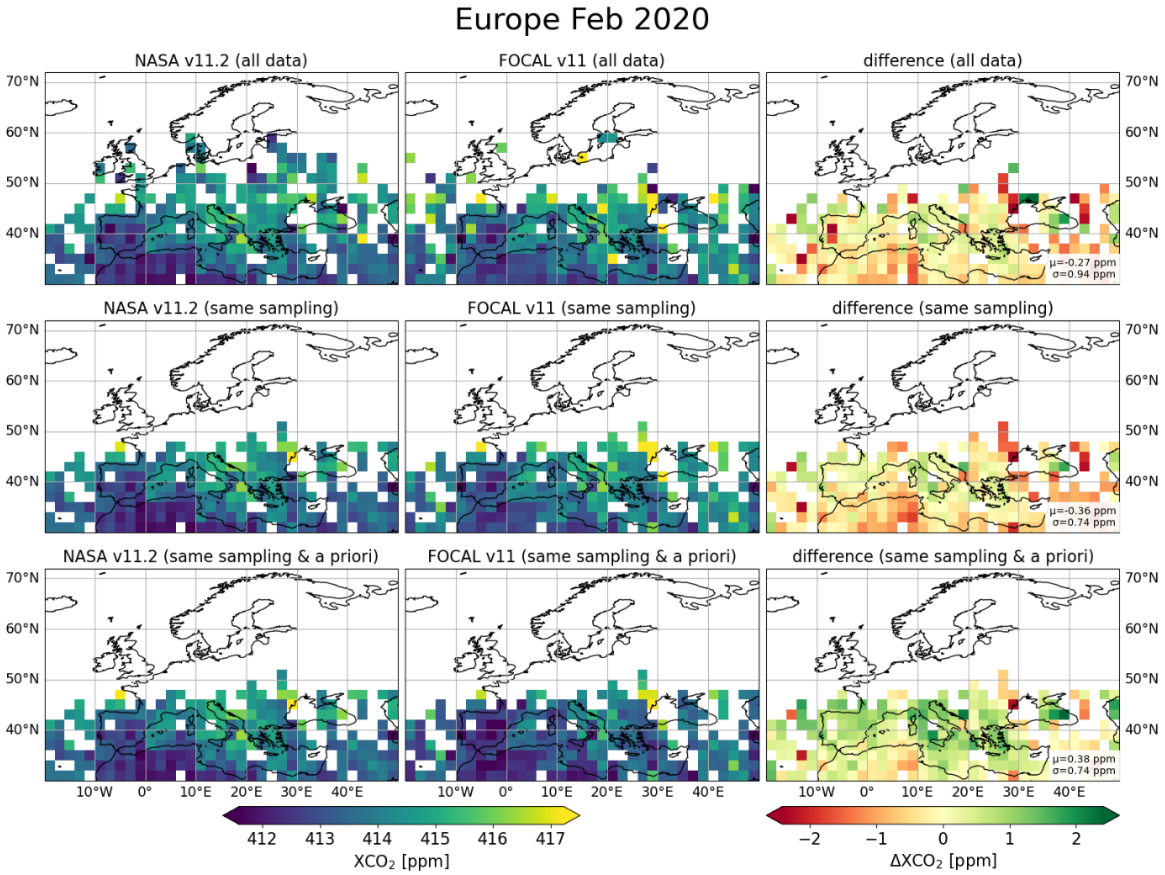

Figure 12: Same as Figure 9 but for soundings in Europe in February 2020.

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101081395

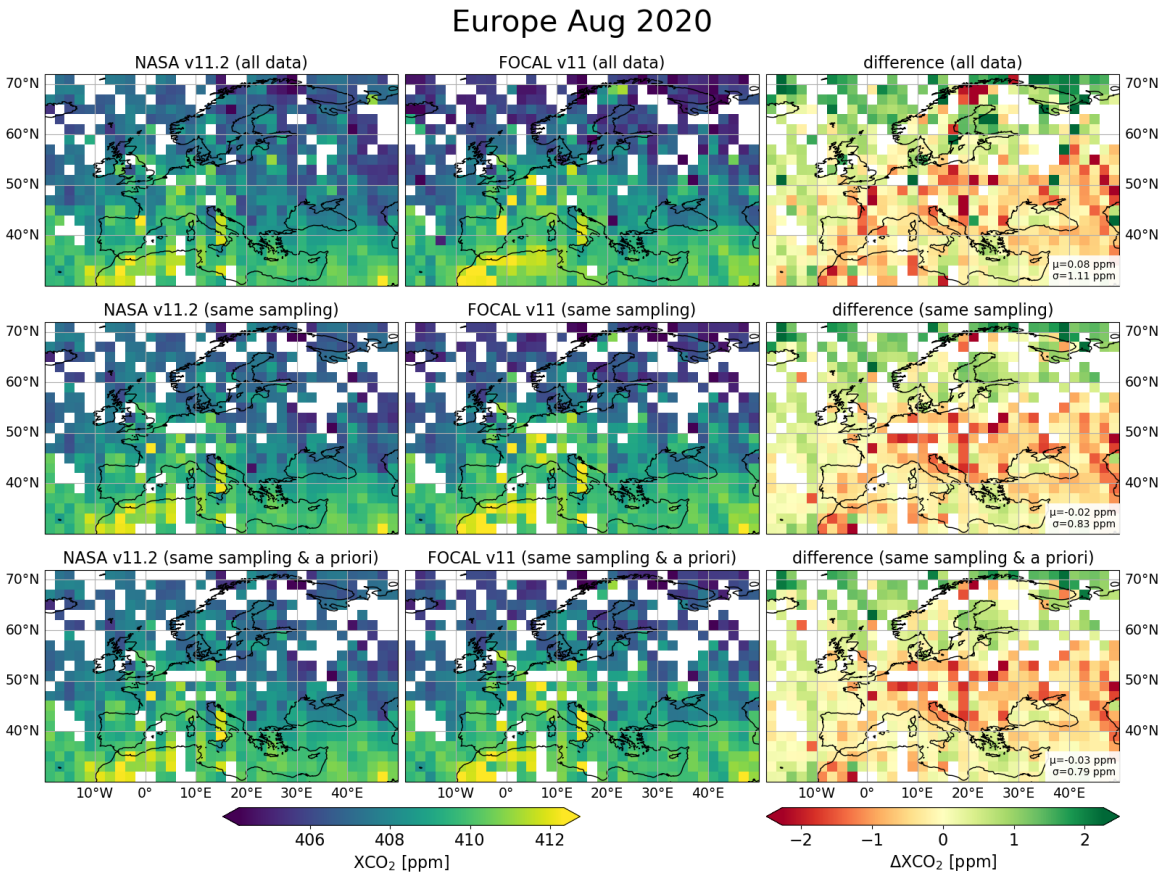

Figure 13: Same as Figure 9 but for soundings in Europe in August 2020.

3.2.3 Scatter plots

Figure 14 presents a scatter plot comparing both products using data from 2020. The correlation between the NASA and FOCAL "same sampling" datasets is strong (r = 0.93). A linear regression of individual soundings yields a slope of 0.85, indicating systematic differences between both products. Figures 15 (February) and 16 (August) show very similar results.

The Europe region also shows consistent results for the full year, with a slightly higher correlation (r = 0.94) and a regression slope of 0.86. However, for February 2020, the agreement deteriorates: the correlation drops to r = 0.71, and the slope decreases to 0.65. This reduction is primarily due to the limited range of data values covered in the regression for that month. The mean bias is still relatively small.

When using the "same sampling & a priori" dataset, NASA and FOCAL agree better. The regression slope increases (from 0.85 to 0.90), while the correlation slightly decreases (from 0.93 to 0.92). This trend is also observed in the regional and monthly datasets.

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101081395

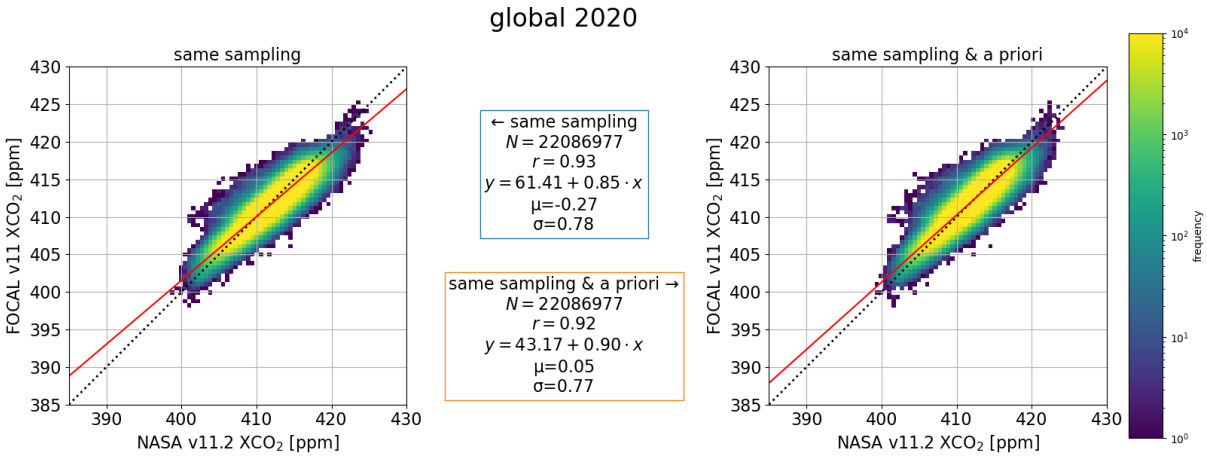

Figure 14: Comparison of global NASA and FOCAL data for 2020. Left: all soundings that appear in both products, right: additionally corrected for same a priori. The parameters of both linear regression (red line) are listed in the centre. μ and σ are the mean and standard deviation of the difference (NASA – FOCAL).

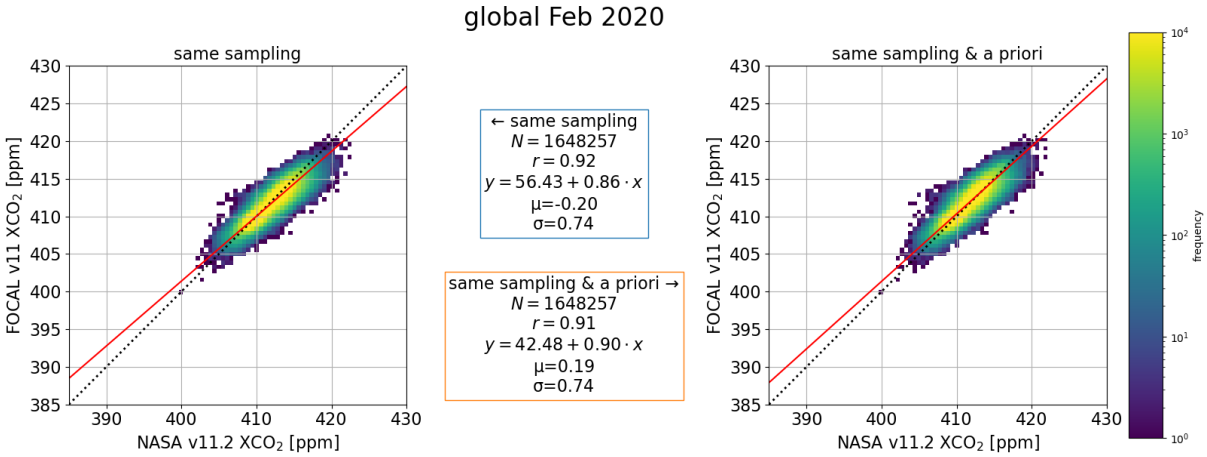

Figure 15: Same as Figure 15 but for soundings in February 2020

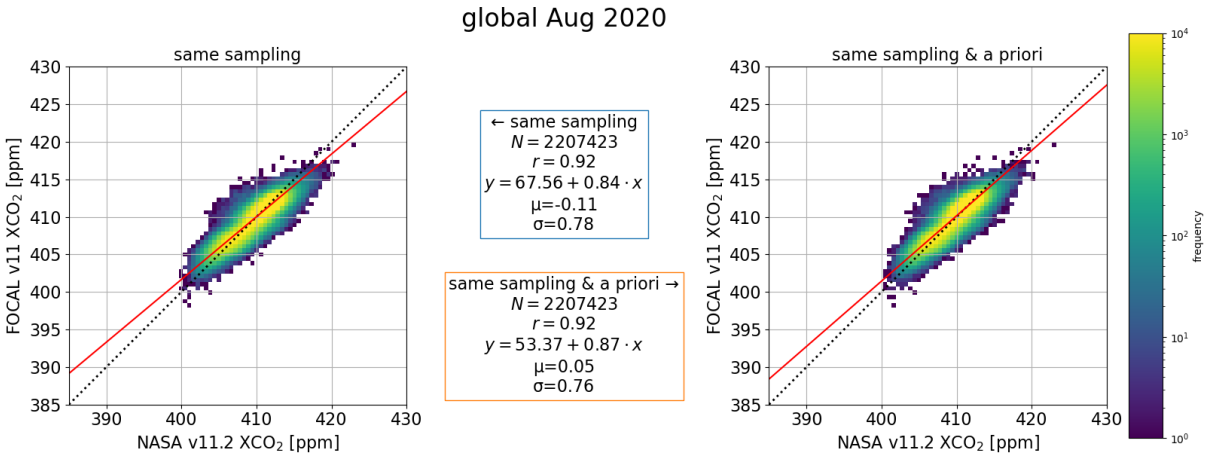

Figure 16: Same as Figure 15 but for soundings in August 2020.

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101081395

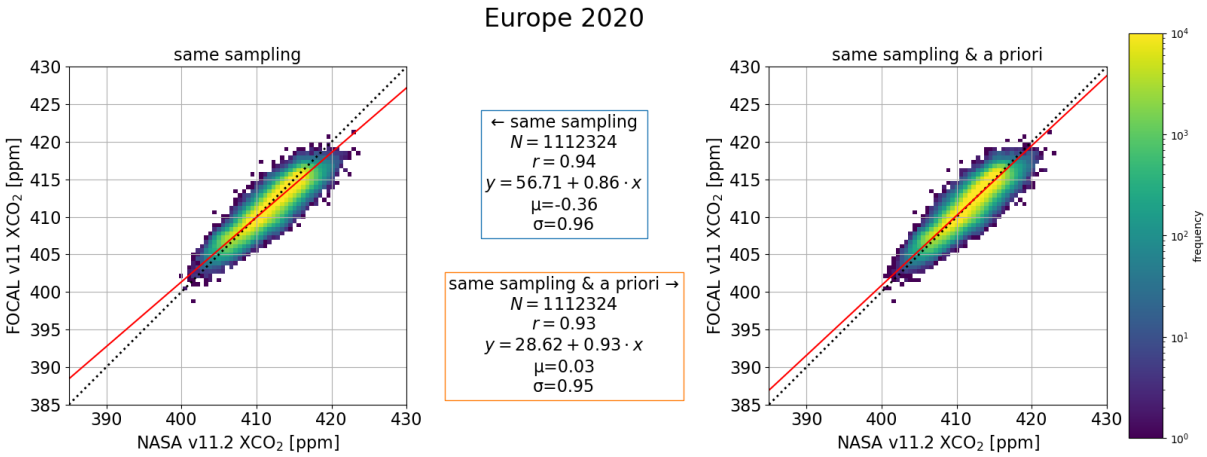

Figure 17: Same as Figure 15 but for soundings in Europe.

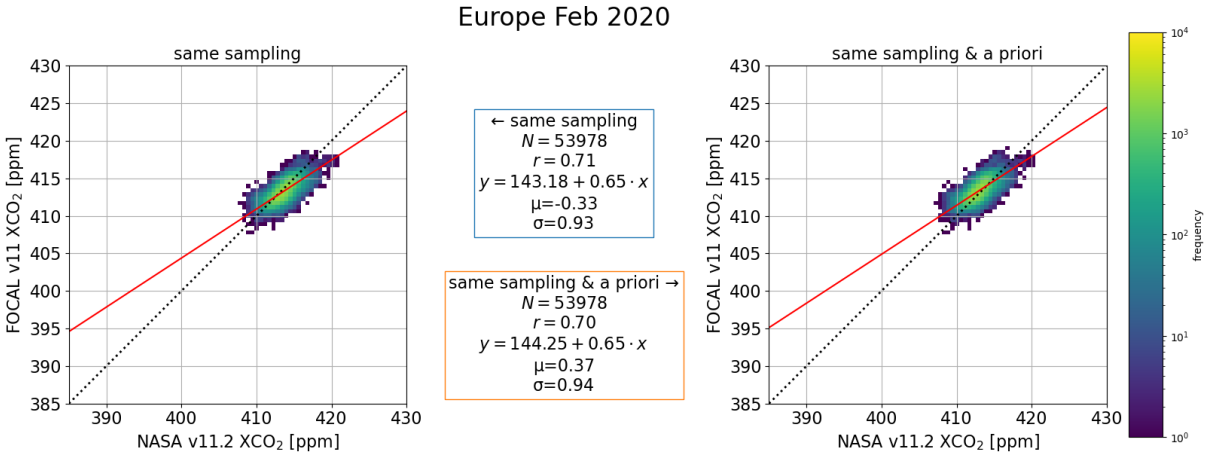

Figure 18: Same as Figure 15 but for soundings in Europe in February 2020.

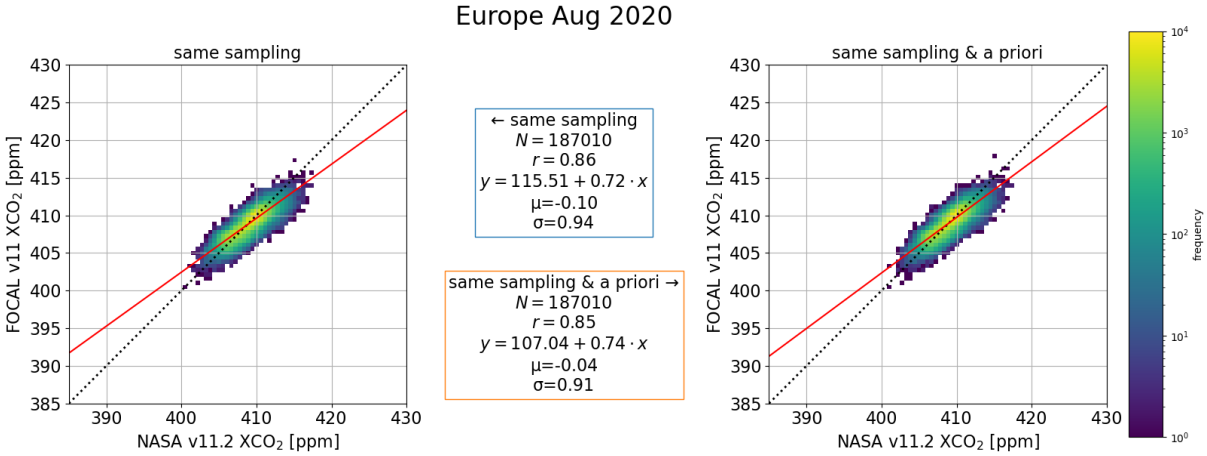

Figure 19: Same as Figure 15 but for soundings in Europe in August 2020.

3.2.4 Timeseries

The time series of monthly global XCO_2 averages shown in Figure 20 reveal only small differences between the NASA and FOCAL products. A consistent seasonal pattern is observed every year from approximately April to June, during which FOCAL reports slightly higher XCO_2 values (0.41 ppm). This difference is reduced to 0.34 ppm when using the “same sampling” dataset and is nearly eliminated (0.01 ppm) in the “same sampling & a priori” dataset.

This project has received funding from the European Union’s Horizon Europe research and innovation programme under grant agreement No 101081395

For the Europe region (Figure 21), the situation is similar, but the systematic differences are more pronounced and cover a larger portion of each year. While the "same sampling & a priori" dataset reduces the discrepancy on average (0.00 ppm), it does not completely remove it. Periodic differences stay visible.

Regional breakdowns in Figure 22 for the Northern Hemisphere, Tropics, and Southern Hemisphere show that the systematic differences are present across all regions, with Europe displaying the largest deviations.

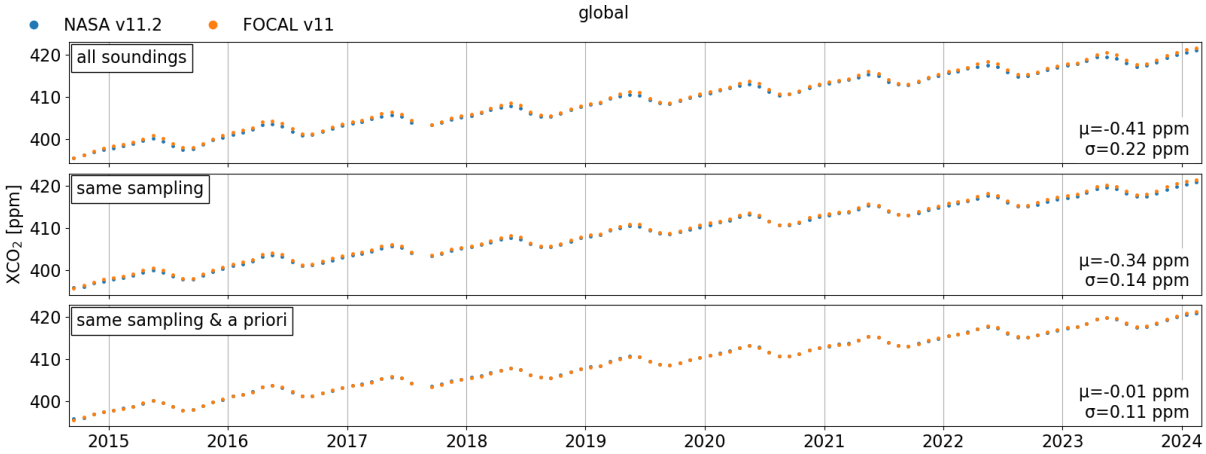

Figure 20: Timeseries of monthly global XCO2 averages. First row: all data/ different sampling, second row: only data in both products / same sampling, third row: additionally corrected for same a priori. μ and σ are the mean and standard deviation of the difference (NASA – FOCAL).

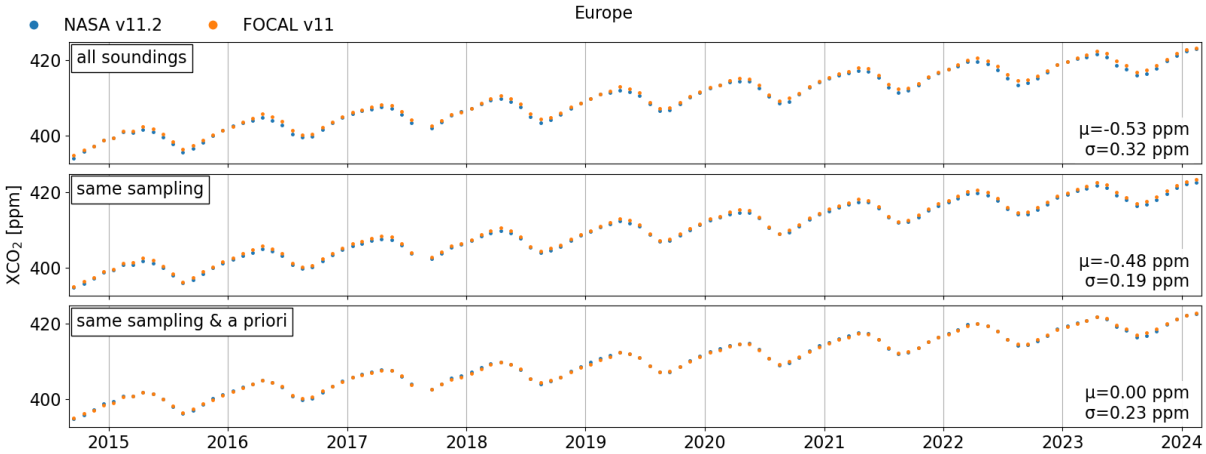

Figure 21: Same as Figure 20 but for average XCO2 in Europe.

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101081395

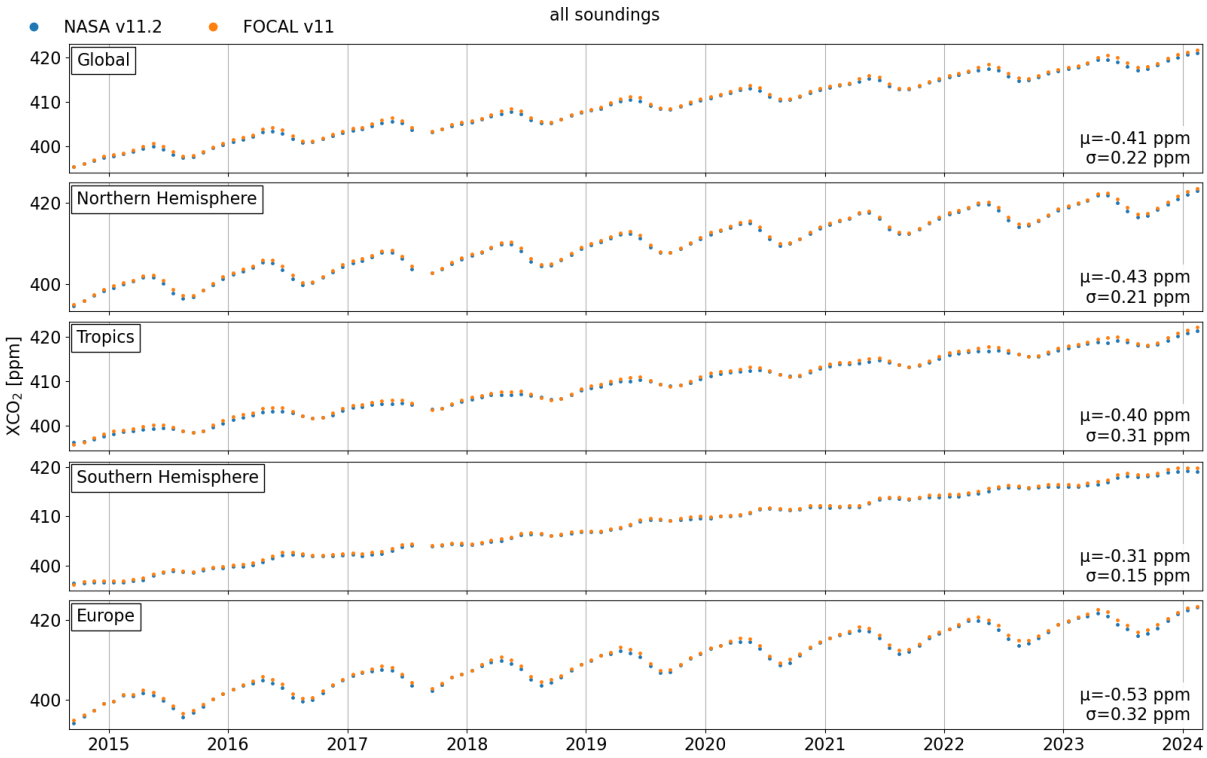

Figure 22: Timeseries of monthly XCO2 averages for different regions. Global and Europe data are identical to the first rows in Figure 20 and Figure 21 respectively.

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101081395

4. Comparison with TCCON

4.1 Methodology

The comparison with TCCON data focuses on Europe, and only European sites, as listed in Table 1, were considered.

Co-location criteria were defined as follows:

- a maximum time difference of two hours,
- a maximum spatial separation of 100 km,
- a maximum surface elevation difference of 250 m.

For each satellite sounding, all co-located TCCON measurements were averaged.

As for the intercomparison described in this section, all measurements were adjusted to a common a priori profile to correct for the a priori contribution in the smoothing equation. SLIMCO$_2$ was again used as the common a priori profile for all datasets.

Results are presented as time series of individual soundings from both OCO-2 products alongside their co-located TCCON observations. To investigate the relationship between satellite and TCCON data, scatter plots were generated. A linear regression was performed on the individual soundings, but results are shown as 2D histograms to reflect data density. Each figure includes the regression equation and the correlation coefficient.

4.2 Results

Figures 23 and 24 present time series of both data products alongside their TCCON colocations. For some stations, the colocations only partially cover the full temporal extent of the OCO-2 data products. For example, colocations at Białystok end in October 2018 when the station was decommissioned. At Bremen, although the station is still active, data are currently only available up to the end of 2021, and there are no colocations for 2017. At Harwell, colocations begin in 2022, and at Nicosia in 2019. In general, there are few colocations during the winter months, with the exception of Nicosia, which maintains more consistent winter coverage. Although both satellite products span a similar time range, the number of colocations with FOCAL is significantly lower than with NASA, less than is expected based on the data discussed in section 3.

Figures 25 and 26 show scatter plots comparing each satellite product directly with their respective TCCON colocations. The correlation is strong for most stations. The weakest agreement is observed at Harwell, with correlation coefficients of r = 0.90 for NASA and r = 0.87 for FOCAL, and linear regression slopes of 0.85 and 0.80, respectively. The short time period of available data at Harwell results in a limited range of XCO_2 values, making the regression fit less robust and more sensitive to outliers. The best agreement is at the Karlsruhe station, with r = 0.99 for NASA and r = 0.98 for FOCAL, and regression slopes of 1.00 and 1.01.

Figure 27 summarizes the statistics for all stations, excluding Ny-Ålesund, which has too few colocations with the FOCAL dataset. The bias at each station is calculated as the mean difference between the retrieved satellite XCO_2 and the corresponding TCCON measurements. Biases at Białystok, Karlsruhe, Orléans, and Paris are relatively small for both satellite products. The largest bias is observed at the Garmisch-Partenkirchen station, where both FOCAL and NASA agree closely. The most significant difference between the two satellite products at a single station is found at Nicosia, where both datasets also have the highest number of colocations with TCCON. The bias values for Nicosia are 0.01 ppm for NASA and 0.92 ppm for FOCAL.

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101081395

Overall, the NASA product shows slightly better agreement with TCCON compared to FOCAL. The station-to-station bias is 0.45 ppm for NASA and 0.56 ppm for FOCAL. The mean scatter is 0.88 ppm for NASA and 1.11 ppm for FOCAL.

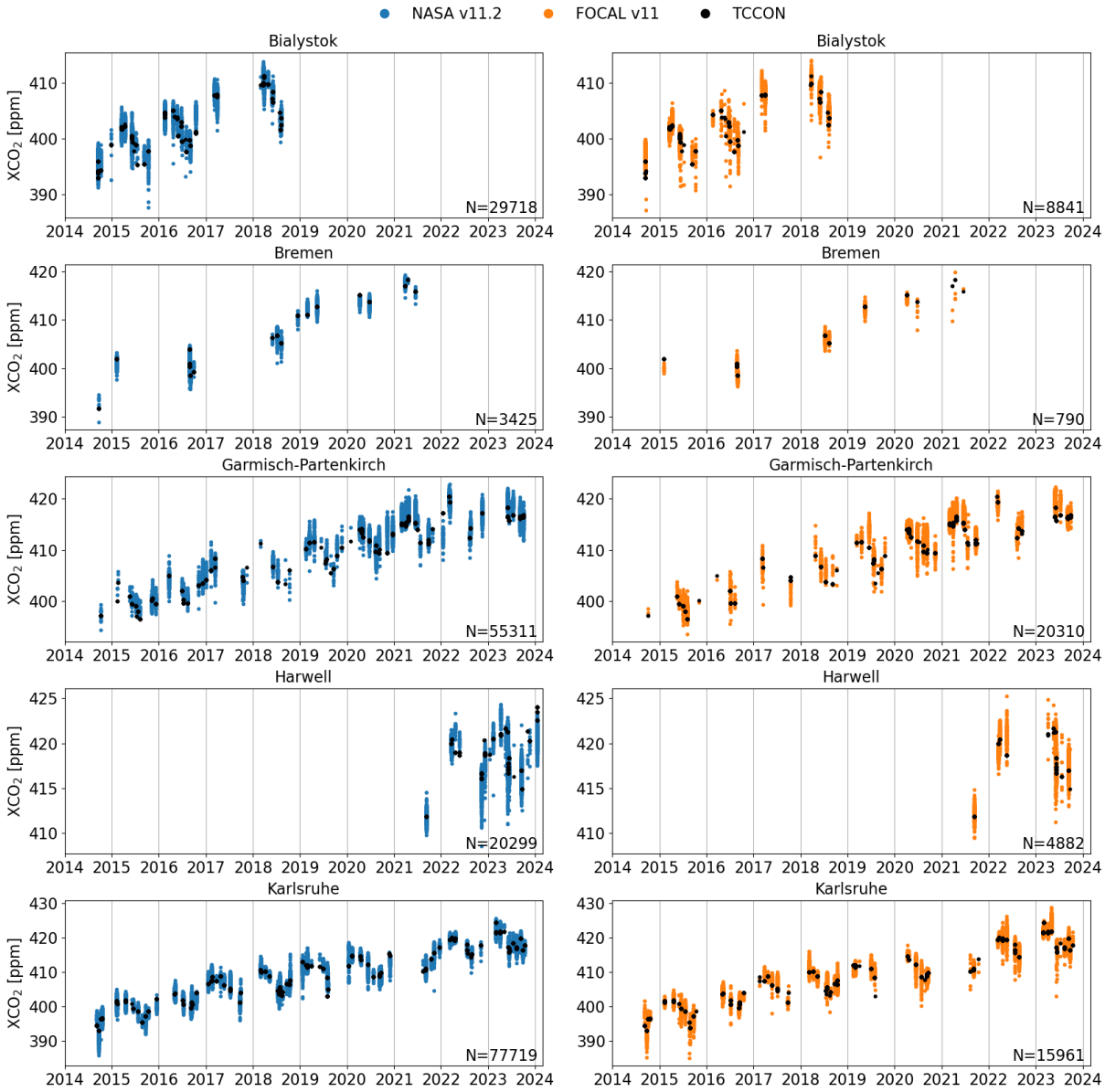

Figure 23: Timeseries of XCO2 for OCO-2 soundings and their TCCON colocations for different TCCON stations. Left: NASA, right FOCAL.

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101081395

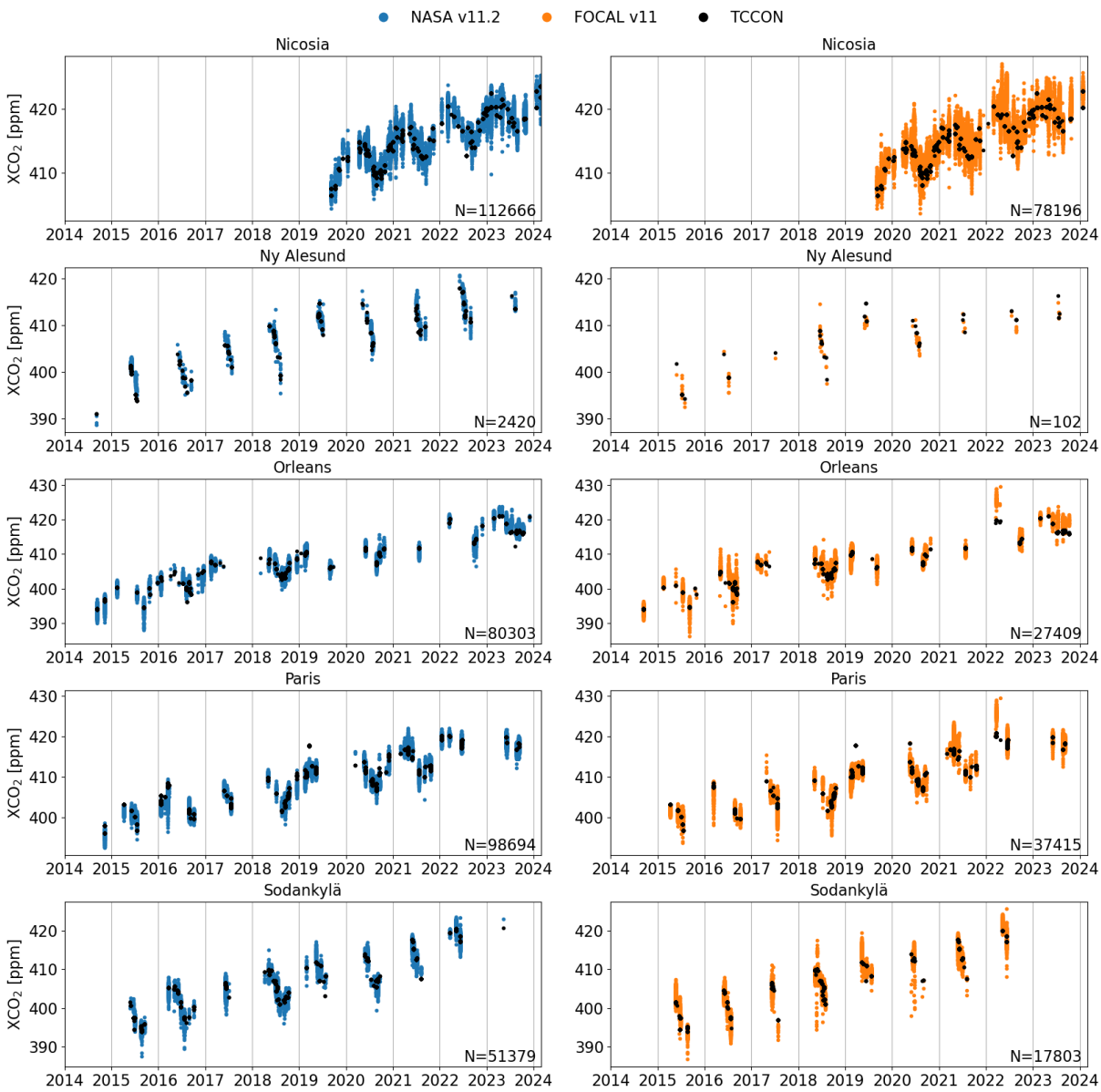

Figure 24: Same as Figure 23 but for more stations.

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101081395

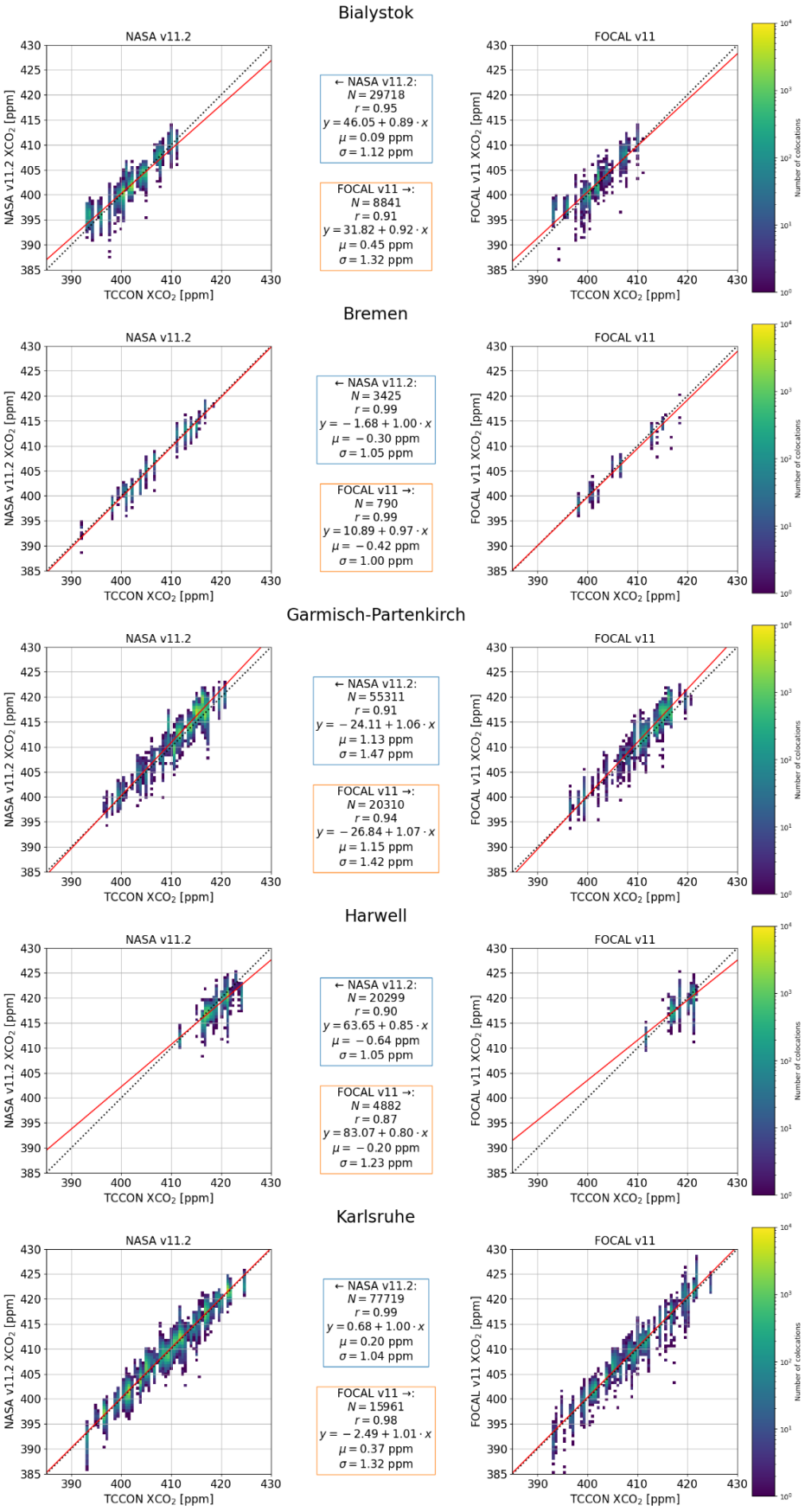

Figure 25: Comparison of OCO-2 XCO2 with TCCON colocations for different TCCON stations. Left: NASA, right: FOCAL. The parameters of a linear regression (red line) are listed in the centre.

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101081395

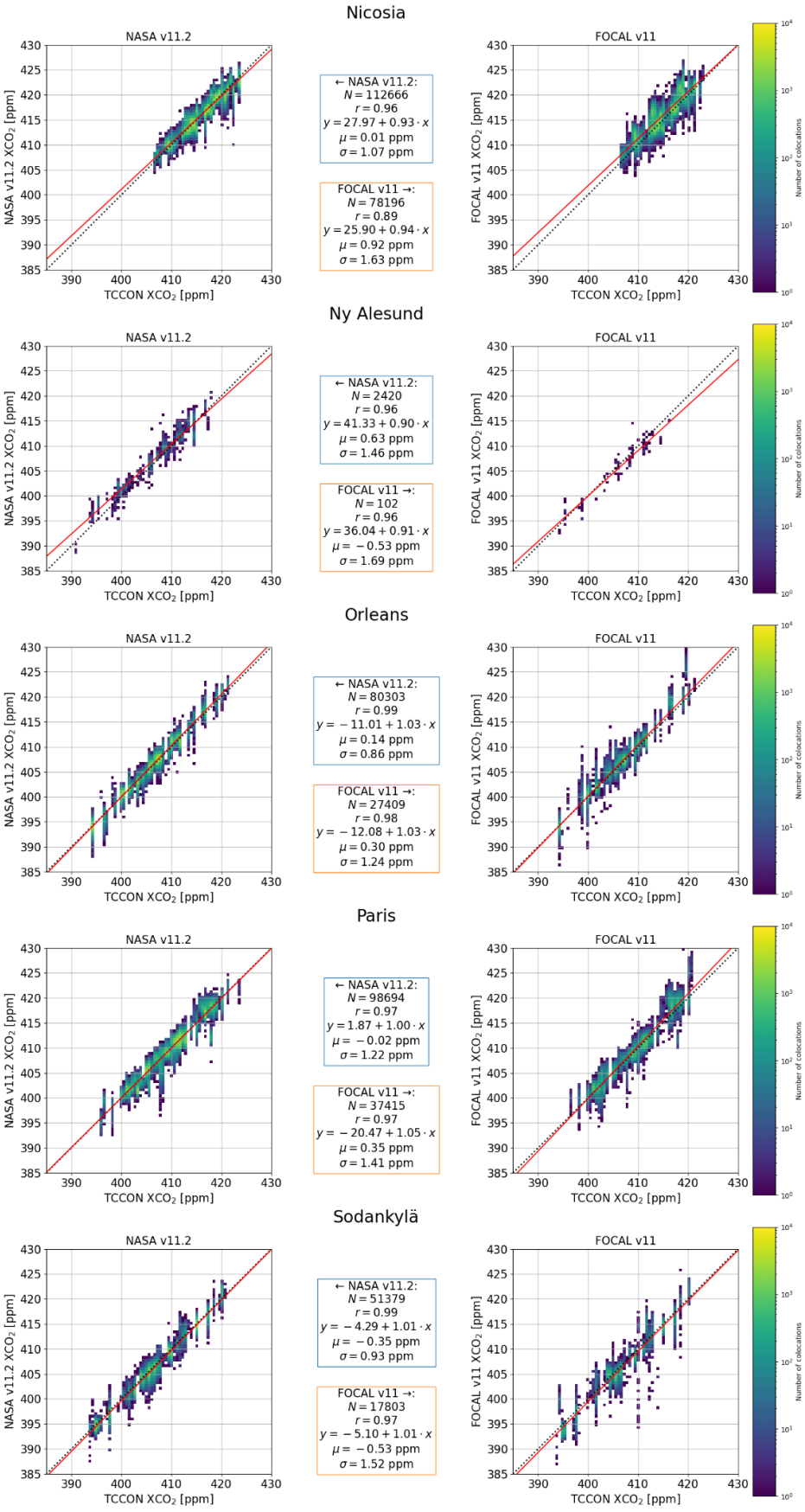

Figure 26: Same as Figure 25 but for more stations

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101081395

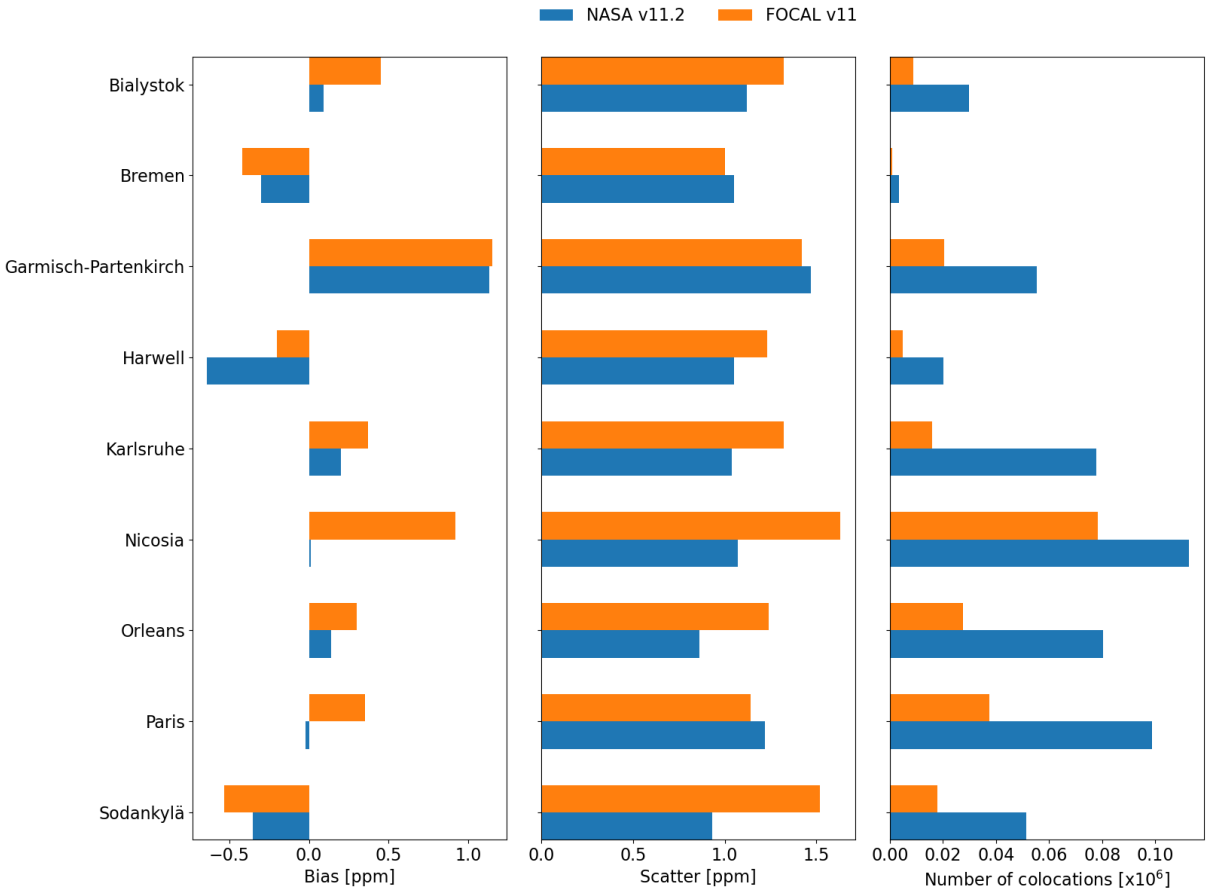

Figure 27: Comparison of OCO-2 XCO2 data products with TCCON. Left: Station bias, centre: scatter, right: number of colocations. Ny-Ålesund was excluded due to insufficient colocations with FOCAL.

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101081395

5. Conclusion

We have evaluated the two available XCO_2 data products for the OCO-2 mission: the official NASA product and the IUP FOCAL product.

One aspect that we evaluated is coverage. Globally, both data products have roughly the same number of data points but there are clear differences in coverage with the NASA product having more datapoints over most land masses with the exception of North Africa and the Arabian Peninsula. Over Europe, both data products have good coverage in summer but, as expected, low/no coverage in high latitude winter. The NASA product has distinctively more data over European land masses.

Furthermore, we have assessed difference in the spatial distribution of XCO_2 of both products. Here, we distinguish between comparisons where (1) all quality-filtered soundings from each product are used, (2) where only of soundings that are present in both products are used and (3) where the same soundings as for (2) are used but with a common a priori applied to both datasets.

Globally, we find substantial regional differences between the NASA and FOCAL products, particularly south of the Sahara in Africa and over India. These are also regions where the FOCAL product has more data points. When using only the same soundings, these XCO_2 differences are reduced, and they decrease further when a common a priori is applied although residual differences of up to 1 ppm remain. This reaffirms that sampling and the used a priori matters for such comparisons. The comparison over Europe shows large XCO_2 difference between both datasets with high variability. This is likely related to the variable number of datapoint of both datasets. Using the same data points in both datasets shows a clearer picture with FOCAL data being higher over European land while both datasets agree better over water. Using the same a priori shifts this picture somewhat, and we obtain better agreement over land with a clear land sea difference between both datasets with FOCAL being lower than the NASA dataset over water. The existence of small but systematic differences between both datasets is confirmed with correlation plots.

Beside the spatial comparison, we also consider a temporal comparison between both datasets by assessing time series. In general, we find for both dataset consistent XCO_2 time series for the time period from 2014 to 2025. Some differences in the seasonal amplitude are visible when using all data but this mostly vanishes when using only the same soundings with the same a priori, although a small tendency remains of the NASA product to show a slightly larger seasonal amplitude.

Finally, we compared both datasets to measurements from the TCCON network for sites in Europe. Overall, we find a good match of both datasets with TCCON data with correlation coefficients almost always above 0.9. Biases for the different stations are mostly below 0.5 ppm but can be larger for the Garmisch, Harwell and Nicosia sites. We find that the NASA dataset has lower biases compared to the FOCAL product for all sites except for Harwell. The NASA product has also a slightly lower scatter for all sites and a noticeable larger number of datapoints over each TCCON site.

In summary, we find that two high quality data products exist for OCO-2. Both datasets have noticeable differences in their data density over Europe and a small but systematic difference exist in land-sea contrast between both datasets which will have an impact when used in surface flux inversion studies. Since the NASA data product shows better agreement with TCCON reference sites in Europe, there is a recommendation towards this product. However, the differences in the land-sea contrast of the data products cannot be evaluated with the TCCON data and we cannot tell which data product is a better choice under this aspect.

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101081395

Acknowledgements

The FOCAL OCO-2 XCO_2 product has been generated with funding from:

ESA (GHG-CCI project)

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101081395

References

Bovensmann, H., Burrows, J.P., Buchwitz, M., Frerick, J., Noël, S., Rozanov, V.V., Chance, K.V., Goede, A.P.H., 1999. SCIAMACHY: Mission Objectives and Measurement Modes. Journal of the Atmospheric Sciences 56, 127–150. https://doi.org/10.1175/1520-0469(1999)056%253C0127:SMOAMM%253E2.0.CO;2

Buschmann, M., Petri, C., Palm, M., Warneke, T., Notholt, J., 2022. TCCON data from Ny-Ålesund, Svalbard (NO), Release GGG2020.R0.

Byrne, B., Baker, D.F., Basu, S., Bertolacci, M., Bowman, K.W., Carroll, D., Chatterjee, A., Chevallier, F., Ciais, P., Cressie, N., Crisp, D., Crowell, S., Deng, F., Deng, Z., Deutscher, N.M., Dubey, M.K., Feng, S., García, O.E., Griffith, D.W.T., Herkommer, B., Hu, L., Jacobson, A.R., Janardanan, R., Jeong, S., Johnson, M.S., Jones, D.B.A., Kivi, R., Liu, J., Liu, Z., Maksyutov, S., Miller, J.B., Miller, S.M., Morino, I., Notholt, J., Oda, T., O'Dell, C.W., Oh, Y.-S., Ohyama, H., Patra, P.K., Peiro, H., Petri, C., Philip, S., Pollard, D.F., Poulter, B., Remaud, M., Schuh, A., Sha, M.K., Shiomi, K., Strong, K., Sweeney, C., Té, Y., Tian, H., Velazco, V.A., Vrekoussis, M., Warneke, T., Worden, J.R., Wunch, D., Yao, Y., Yun, J., Zammit-Mangion, A., Zeng, N., 2023. National CO_2 budgets (2015–2020) inferred from atmospheric CO_2 observations in support of the global stocktake. Earth System Science Data 15, 963–1004. https://doi.org/10.5194/essd-15-963-2023

Cansot, E., Pistre, L., Castelnau, M., Landiech, P., Georges, L., Gaeremynck, Y., Bernard, P., 2023. MicroCarb instrument, overview and first results, in: International Conference on Space Optics — ICSO 2022. Presented at the International Conference on Space Optics — ICSO 2022, SPIE, pp. 1341–1353. https://doi.org/10.1117/12.2690330

Copernicus Atmosphere Monitoring Service, 2020. CAMS global inversion-optimised greenhouse gas fluxes and concentrations. https://doi.org/10.24381/ed2851d2

Crisp, D., Miller, C.E., DeCola, P.L., 2008. NASA Orbiting Carbon Observatory: measuring the column averaged carbon dioxide mole fraction from space. JARS 2, 023508. https://doi.org/10.1117/1.2898457

Crowell, S., Baker, D., Schuh, A., Basu, S., Jacobson, A.R., Chevallier, F., Liu, J., Deng, F., Feng, L., McKain, K., Chatterjee, A., Miller, J.B., Stephens, B.B., Eldering, A., Crisp, D., Schimel, D., Nassar, R., O'Dell, C.W., Oda, T., Sweeney, C., Palmer, P.I., Jones, D.B.A., 2019. The 2015–2016 carbon cycle as seen from OCO-2 and the global in situ network. Atmospheric Chemistry and Physics 19, 9797–9831. https://doi.org/10.5194/acp-19-9797-2019

Hase, F., Herkommer, B., Groß, J., Blumenstock, T., Kiel, M.ä., Dohe, S., 2024. TCCON data from Karlsruhe (DE), Release GGG2020.R2.

Kuze, A., Suto, H., Nakajima, M., Hamazaki, T., 2009. Thermal and near infrared sensor for carbon observation Fourier-transform spectrometer on the Greenhouse Gases Observing Satellite for greenhouse gases monitoring. Appl. Opt., AO 48, 6716–6733. https://doi.org/10.1364/AO.48.006716

Laughner, J.L., Toon, G.C., Mendonca, J., Petri, C., Roche, S., Wunch, D., Blavier, J.-F., Griffith, D.W.T., Heikkinen, P., Keeling, R.F., Kiel, M., Kivi, R., Roehl, C.M., Stephens, B.B., Baier, B.C., Chen, H., Choi, Y., Deutscher, N.M., DiGangi, J.P., Gross, J., Herkommer, B., Jeseck, P., Laemmel, T., Lan, X., McGee, E., McKain, K., Miller, J., Morino, I., Notholt, J., Ohyama, H., Pollard, D.F., Rettinger, M., Riris, H., Rousogenous, C., Sha, M.K., Shiomi, K., Strong, K., Sussmann, R., Té, Y., Velazco, V.A., Wofsy, S.C., Zhou, M., Wennberg, P.O., 2024. The Total Carbon Column Observing Network's GGG2020 data version. Earth System Science Data 16, 2197–2260. https://doi.org/10.5194/essd-16-2197-2024

Noël, S., Reuter, M., Buchwitz, M., Borchardt, J., Hilker, M., Schneising, O., Bovensmann, H., Burrows, J.P., Di Noia, A., Parker, R.J., Suto, H., Yoshida, Y., Buschmann, M., Deutscher, N.M., Feist, D.G., Griffith, D.W.T., Hase, F., Kivi, R., Liu, C., Morino, I., Notholt, J., Oh, Y.-S., Ohyama, H., Petri, C., Pollard, D.F., Rettinger, M., Roehl, C., Rousogenous, C., Sha, M.K., Shiomi, K., Strong, K., Sussmann, R., Té, Y., Velazco, V.A., Vrekoussis, M., Warneke, T., 2022. Retrieval

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101081395

of greenhouse gases from GOSAT and GOSAT-2 using the FOCAL algorithm. Atmospheric Measurement Techniques 15, 3401–3437. https://doi.org/10.5194/amt-15-3401-2022

Notholt, J., Petri, C., Warneke, T., Buschmann, M., 2022. TCCON data from Bremen (DE), Release GGG2020.R0.

OCO-2/OCO-3 Science Team, Payne, V., Chatterjee, A., 2024. OCO-2 Level 2 bias-corrected XCO2 and other select fields from the full-physics retrieval aggregated as daily files, Retrospective processing V11.2r (OCO2_L2_Lite_FP) at GES DISC. https://doi.org/10.5067/70K2B2W8MNGY

O'Dell, C.W., 2010. Acceleration of multiple-scattering, hyperspectral radiative transfer calculations via low-streams interpolation. Journal of Geophysical Research: Atmospheres 115. https://doi.org/10.1029/2009JD012803

Petri, C., Vrekoussis, M., Rousogenous, C., Warneke, T., Sciare, J., Notholt, J., 2024. TCCON data from Nicosia (CY), Release GGG2020.R1.

Reuter, M., Hilker, M., Noël, S., Buchwitz, M., Schneising, O., Bovensmann, H., Burrows, J.P., Bösch, H., 2025. ESA Greenhouse Gases Climate Change Initiative (GHG_cci): Column averaged carbon dioxide from OCO-2 generated with the FOCAL algorithm, version 11.0. https://doi.org/10.5285/198D56F5FBBF4144B8FD4932328BE462

Sussmann, R., Rettinger, M., 2025. TCCON data from Garmisch (DE), Release GGG2020.R1.

Tanimoto, H., Matsunaga, T., 2024. The GOSAT-GW greenhouse gas observing mission: Concept. Journal of The Remote Sensing Society of Japan 44, 96–101. https://doi.org/10.11440/rssj.2024.013

Té, Y., Jeseck, P., Janssen, C., 2022. TCCON data from Paris (FR), Release GGG2020.R0.

Wang, J., Feng, L., Palmer, P.I., Liu, Y., Fang, S., Bösch, H., O'Dell, C.W., Tang, X., Yang, D., Liu, L., Xia, C., 2020. Large Chinese land carbon sink estimated from atmospheric carbon dioxide data. Nature 586, 720–723. https://doi.org/10.1038/s41586-020-2849-9

Warneke, T., Petri, C., Notholt, J., Buschmann, M., 2024. TCCON data from Orléans (FR), Release GGG2020.R1.

Weidmann, D., Brownsword, R., Doniki, S., 2023. TCCON data from Harwell, Oxfordshire (UK), Release GGG2020.R0.

Wunch, D., Wennberg, P.O., Toon, G.C., Connor, B.J., Fisher, B., Osterman, G.B., Frankenberg, C., Mandrake, L., O'Dell, C., Ahonen, P., Biraud, S.C., Castano, R., Cressie, N., Crisp, D., Deutscher, N.M., Eldering, A., Fisher, M.L., Griffith, D.W.T., Gunson, M., Heikkinen, P., Keppel-Aleks, G., Kyrö, E., Lindenmaier, R., Macatangay, R., Mendonca, J., Messerschmidt, J., Miller, C.E., Morino, I., Notholt, J., Oyafuso, F.A., Rettinger, M., Robinson, J., Roehl, C.M., Salawitch, R.J., Sherlock, V., Strong, K., Sussmann, R., Tanaka, T., Thompson, D.R., Uchino, O., Warneke, T., Wofsy, S.C., 2011. A method for evaluating bias in global measurements of CO_2 total columns from space. Atmospheric Chemistry and Physics 11, 12317–12337. https://doi.org/10.5194/acp-11-12317-2011

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101081395

https://eyeclima.eu

BRUSSELS, 25 09 2025

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union. Neither the European Union nor the granting authority can be held responsible for them.

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101081395