ESA Climate Change Initiative “Plus” (CCI+)

End-to-End ECV Uncertainty Budget Version 3 (E3UBv3)

-

for the FOCAL XCO_2 OCO-2 Data Product

CO2_OC2_FOCA (v10)

for the Essential Climate Variable (ECV)

Greenhouse Gases (GHG)

Authors:
Maximilian Reuter (mreuter@iup.physik.uni-bremen.de),
Michael Hilker, Malin Ingendahl, and Michael Buchwitz
Institute of Environmental Physics (IUP) / Institute of Remote Sensing (IFE),
University of Bremen (UB), Bremen, Germany

The further development of the FOCAL retrieval algorithm and corresponding OCO-2 data processing and analysis is co-funded by:

ESA via CCI project **GHG-CCI+**

EUMETSAT via the **FOCAL-CO2M** study

The **European Commission** via the H2020 project **VERIFY** (Grant Agreement No. 776810)

Change log

Version Nr.	Date	Status	Reason for change
Version 1 draft	29. Nov. 2019	Submitted for review	New document
Version 1	06. Jan. 2020	Final / submitted	Improved Fig. 5
Version 2 draft	16. Nov. 2020	Submitted for review	New data set FOCAL v09 2015-2019
Version 2	25. Jan. 2021	Final	Modified discussion, new Fig. 4.1
Version 3	06. Feb. 2022	As submitted	New data set FOCAL v10 2014-2021, improved validation

Table of Contents

1 Purpose of document

The purpose of this document is to describe the End-to-End ECV Uncertainty Budget (E3UB) of the FOCAL OCO-2 (Orbiting Carbon Observatory-2) XCO_2 (column-averaged dry-air mole fraction of CO_2) product CO2_OC2_FOCA v10. The E3UB is intended for experienced users, as for instance in the field of inverse modeling of surface fluxes.

The CO2_OC2_FOCA retrieval algorithm is developed and operated by the University of Bremen and it analyzes radiance measurements (level 1 data) of OCO-2 in order to retrieve XCO_2. FOCAL's radiative transfer and retrieval technique has initially been described by **/Reuter et al., 2017a/** and its first application to OCO-2 data by **/Reuter et al., 2017b/**. Since then, many minor and major algorithm improvements have been implemented and documented in the most recent version of the CO2_OC2_FOCA algorithm theoretical basis document **/ATBDv3, 2021/**. Additional information on FOCAL can also be obtained from the FOCAL website (http://www.iup.uni-bremen.de/~mreuter/focal.php).

Reliable uncertainty estimates of the XCO_2 retrieval are, e.g., required to translate remotely sensed data into estimations of surface fluxes with a known degree of confidence. The GHG-CCI user requirements document **/URDv3.0, 2020/** defines strict measurement accuracy and precision requirements, allowing the identification of minute changes in magnitude and sign of the XCO_2 concentration.

This E3UB document provides important information on the data reliability of the FOCAL OCO-2 XCO_2 level 2 (L2, i.e., individual soundings) product. This includes an assessment of stochastic and potential systematic uncertainties based on results of a validation study, analyses of simulated observations, a model comparison, and a comparison with the operational NASA OCO-2 XCO_2 data product.

2 Uncertainty analysis based on simulations

In order to asses FOCAL's theoretical capabilities in retrieving XCO_2, we confronted it with radiance measurements simulated with the accurate RT (radiative transfer) code SCIATRAN **/Rozanov et al., 2014/**. The performed analyses can also be understood as test of the suitability of the approximations made in FOCAL's RT and of the retrieval setup. Hereby, we primarily concentrate on the influence of light scattering and analyze the systematic errors and stochastic a posteriori uncertainties of several different geophysical scenarios.

We use the FOCAL OCO-2 XCO_2 retrieval algorithm v10 as described by **/ATBDv3, 2021/** for these experiments. Its implemented post filtering and bias correction methods account not only for systematic errors, e.g., due to imperfections of FOCAL's RT or due to the underdetermined nature of the retrieval problem. These data driven methods also account for unknown unknowns such as potential instrumental effects. As it is not possible to realistically simulate the spectral error introduced by unknown effects, it cannot be expected that FOCAL's post filtering and bias correction methods produce meaningful results when applied to simulated spectra. Therefore, we omit FOCAL's entire post processing for these experiments.

2.1 Simulations

Our set of simulations is the same as that used by **/Reuter et al., 2017a/**. It is not designed to comprehensively cover the majority of potential geophysical scenarios, because the final quality depends on the full retrieval scheme including, e.g., potential instrument and forward model errors as well as post filtering and bias correction methods. The aim of these retrieval experiments is rather to get an impression which geophysical scenarios may require post filtering or bias correction and on the to be expected magnitude of the bias correction.

We defined a geophysical *baseline* scenario that has a spectrally flat albedo of 0.2, 0.2, 0.1, and 0.05 in the SIF, O2, wCO2, and sCO2 fit window (values which have also been used by, e.g., **/Bovensmann et al., 2010/**). It does not include chlorophyll fluorescence, scattering by aerosols, clouds, or Rayleigh. Its temperature, pressure, and water vapor (XH2O = 3031ppm) profiles are taken from an ECMWF analysis of

August 28, 2015, 12:00 UTC, 9°E, 53°N. Its CO_2 profile is calculated with the simple empirical carbon model SECM2016 and corresponds to an XCO_2 value of about 395ppm. All other scenarios are descendants of the *baseline* scenario.

Each scenario is analyzed for three solar zenith angles (20°, 40°, and 60°) and for two directions of polarization (parallel and perpendicular to the solar principle plane (SPP)). The satellite zenith angle is set to 0°(nadir).

The *SIF* scenario adds 1mW/m2/sr/nm chlorophyll fluorescence at 760nm to the simulated measurement of the *baseline* scenario. The *XCO2+6ppm* scenario has an increased CO_2 concentration of 15ppm, 10ppm, and 5ppm in the three lowermost layers, so that the column-average concentration is enhanced by 6ppm.

All scattering related scenarios are more complex for the retrieval because of FOCAL's scattering approximations. The *Rayleigh* scenario adds Rayleigh scattering to the *baseline* scenario; the Rayleigh optical thickness at 760nm for this scenario is about 0.026. *Rayleigh+Aerosol BG* additionally includes a (primarily) stratospheric background aerosol with an AOT (aerosol optical thickness at 760nm) of 0.019 (0.003 at 1600nm and 0.001 at 2050 nm). *Rayleigh+Aerosol cont* adds a continental aerosol to the boundary layer so that the total AOT becomes 0.158 (0.060 at 1600nm and 0.037 at 2050nm). *Rayleigh+Aerosol urban* adds a strong contamination with urban aerosol to the boundary layer and the total AOT becomes 0.702 (0.245 at 1600nm and 0.151 at 2050nm).

The scenarios *Rayleigh+Dark surface*, *Rayleigh+Bright surface*, and *Rayleigh+Ocean glint* distinguish from the *Rayleigh* scenario only by their surface reflection properties. *Rayleigh+Dark surface* and *Rayleigh+Bright surface* correspond to the *Rayleigh* scenario but with an albedo multiplied with 0.7 and 1.4, respectively. The *Rayleigh+Ocean glint* scenario deviates from the assumption of a Lambertian surface bidirectional reflectance distribution function (BRDF); it includes an ocean surface and a wind speed of 5m/s, 37° relative to the SPP. Additionally, the satellite zenith angle of this scenario is set to 0.75 times the solar zenith angle so that the satellite looks near the glint spot of specular reflectance.

Two cloud scenarios (*Rayleigh+AerosolBG+Water cloud* and *Rayleigh+Aerosol BG+Ice cloud*) add a sub-visible water or ice cloud to the *Rayleigh+Aerosol BG* scenario. The water cloud has a height of 3km, droplets with an effective radius of 12µm, and a COT (cloud optical thickness at 500nm) of 0.039. The ice cloud is made of fractal particles with an effective radius of 50µm, has a height of 8km, and a COT of 0.033.

2.2 Results

Primarily, we are interested in XCO_2 retrieval results of high quality; the correct retrieval of other state vector elements is less important as long as the XCO_2 quality is not affected. Figure 2.1 summarizes the systematic errors and stochastic uncertainties of the retrieved XCO_2 for all analyzed geophysical scenarios, solar zenith angles, and polarizations.

The *baseline* scenario is mainly to ensure consistency of the RT used to simulate the measurements (SCIATRAN) and the RT of the retrieval (FOCAL). Additionally, the *baseline* scenario allows estimates of the retrieval's noise error. With SCIATRAN, it is not simply possible to simulate FOCAL's scattering approximations, that is why this scenario excludes scattering. The systematic errors of the *baseline* scenario are always very small (0.0025ppm at maximum), which confirms the RT's consistency in the absorption-only case and ensures that, e.g., the number of particles is basically identical in the SCIATRAN and the FOCAL "world".

A more complex case for FOCAL is the Rayleigh scenario, because Rayleigh scattering takes place in the entire atmospheric column with a peanut-shaped scattering phase function (SPF). This means, due to the approximations of FOCAL's RT, it cannot be expected that FOCAL is able to perfectly fit the simulated spectra. Figure 2.2 (top) shows a spectral fit in all fit windows but with a state vector not including any scattering parameter. Not surprisingly, the residual in the O_2 fit window becomes large compared to the simulated measurement noise. The residuals in the CO_2 fit windows are already small compared to the instrumental noise even without fitting scattering parameters. This is only partly explained by Rayleigh scattering having an Ångström exponent of four and, therefore, a much smaller scattering optical thickness at longer wavelengths. It also indicates that disentangling scattering parameters and CO_2 concentration from measurements in the CO_2 fit windows may be difficult. In other words, most of the scattering information must be imprinted in the residual of the O_2 fit window.

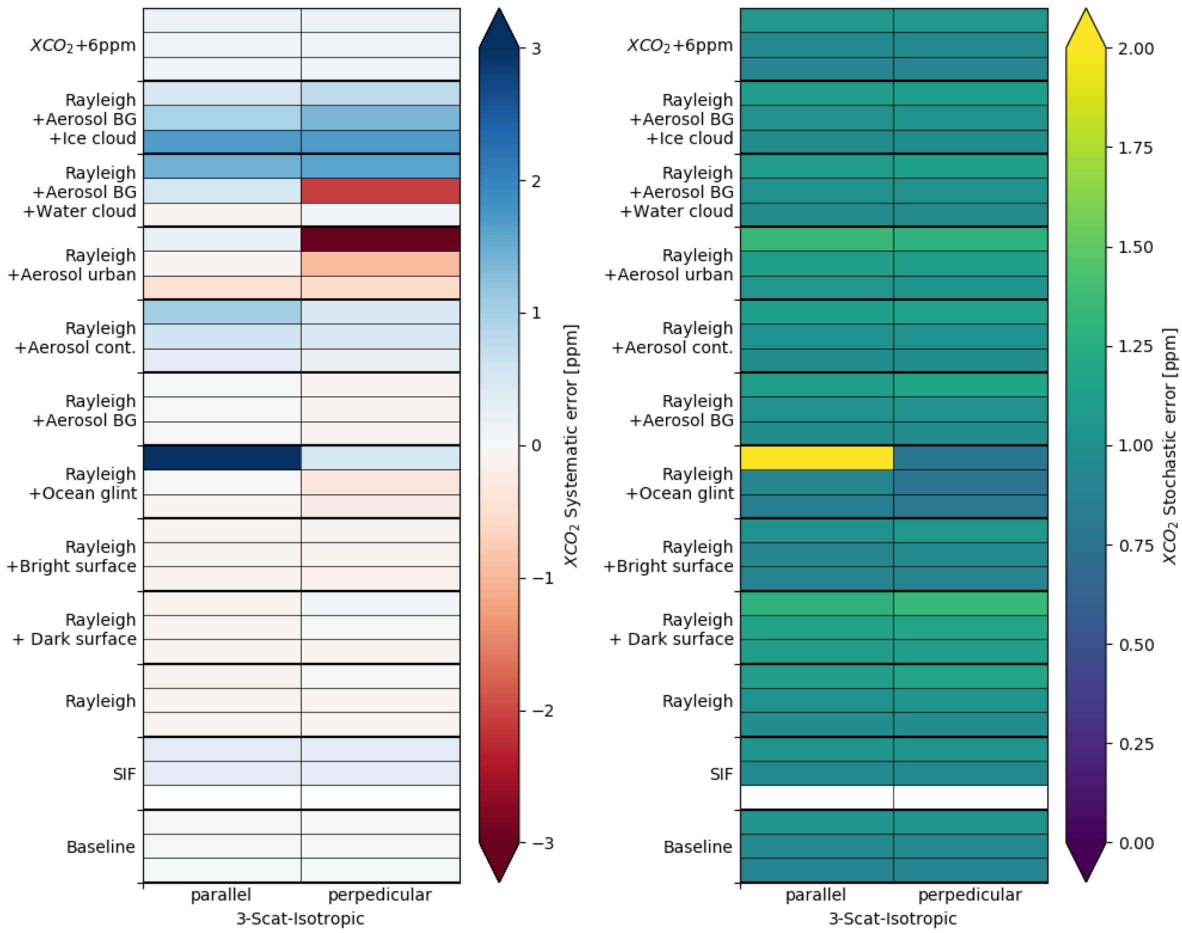

Figure 2.1: Error characteristics of twelve geophysical scenarios. Each scenario has been analyzed for polarization parallel (left) and perpendicular (right) to the SPP as well as for three solar zenith angles (20°, 40°, and 60°, from bottom to top). White boxes (right) represent not converging retrievals. Left: Systematic error (retrieved minus simulated XCO_2). Right: Stochastic uncertainty as reported by the optimal estimation retrieval. See also **/Reuter et al., 2017a/** for details on the geophysical scenarios and **/ATBDv3, 2021/** for details on the retrieval setup.

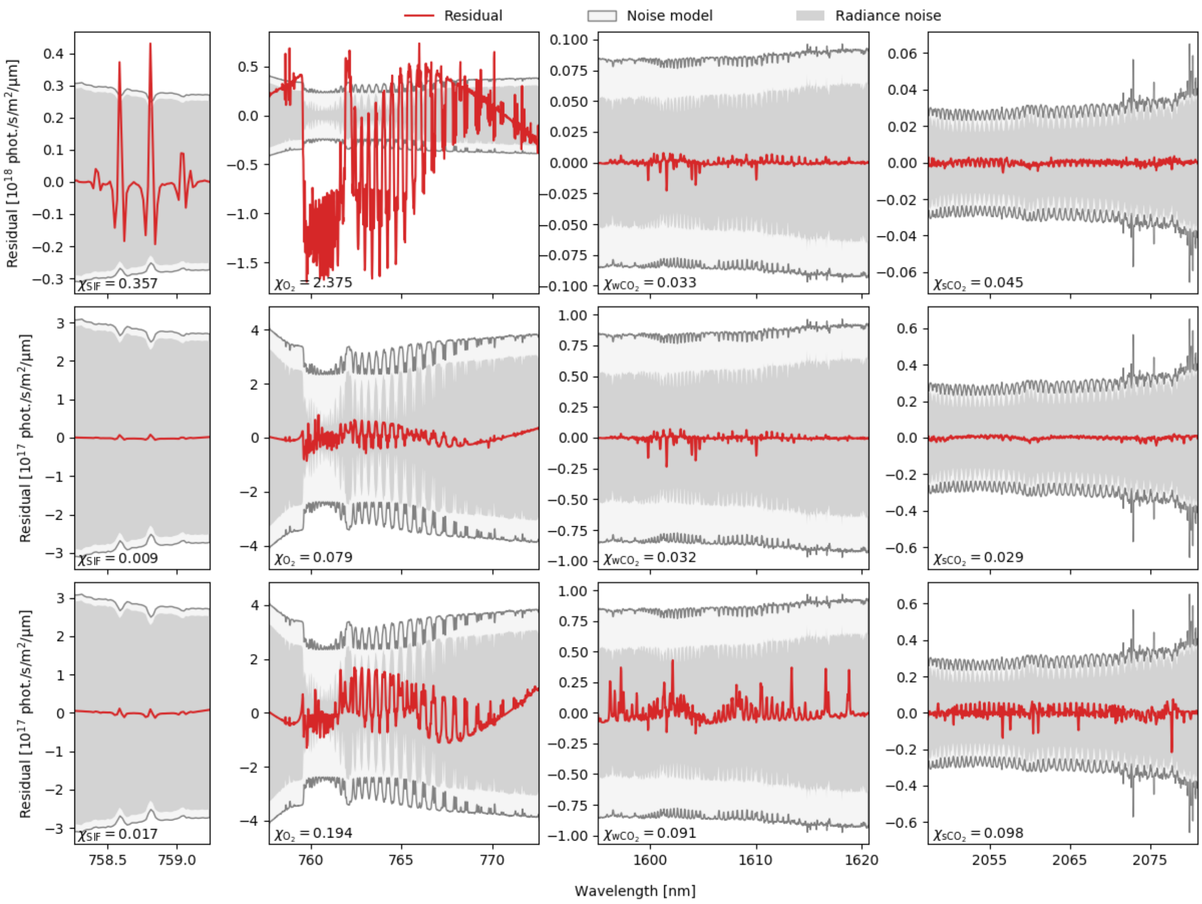

Figure 2.2: Residuals (fit minus measurement, red) and measurement noise (gray area = noise expected from the OCO-2 instrument, gray line = FOCAL's noise model used for the retrieval **/ATBDv3, 2021/**) of SCIATRAN simulated OCO-2 measurements fitted with FOCAL (solar zenith angle = 40°, parallel polarization). Top: *Rayleigh* scenario but with disabled fitting of scattering parameters. Middle: *Rayleigh* scenario. Bottom: *Rayleigh+Aerosol BG+Water cloud* scenario. See also **/Reuter et al., 2017a/** for details on the geophysical scenarios and **/ATBDv3, 2021/** for details on the retrieval setup.

Allowing the retrieval to fit FOCAL’s scattering parameters reduces the Chi square of the O_2 residual by a factor of 30, so that it becomes considerably smaller than expected from instrumental noise.

All other scattering related scenarios are even more “complicated” for FOCAL because different particles contribute to scattering. For example, cloud particles have different properties like height or Ångström exponent as aerosol particles, but FOCAL can only retrieve one effective height and one effective Ångström exponent. Additionally, the SPFs of aerosols and clouds are less isotropic. Therefore, the residuals (Figure 2.2, bottom) and more importantly, the systematic errors typically increase for these scenarios (Figure 2.1, left).

Applying FOCAL to the *Rayleigh+Ocean glint* scenario with a highly non-Lambertian surface BRDF results in systematic XCO_2 errors usually comparable to those of the *Rayleigh* scenario (Figure 2.1, left) except for a solar zenith angle of 60°and polarization parallel to the SPP. In near-glint geometry, specular reflectance dominates the radiation field but with increasing solar zenith angle, the reflected radiation becomes more and more polarized. As a result, the direct photon path often dominates (if not observing parallel polarization at large solar zenith angles) and an imperfect parameterization of scattering becomes less important. The domination of the direct photon path also results in a larger total radiance and, correspondingly, smaller stochastic errors in perpendicular polarization (Figure 2.1, right). The larger systematic XCO_2 error of about 3ppm at 60°and parallel polarization is a result of the poor surface reflectivity in this observation geometry and associated with a large stochastic uncertainty of about 6ppm and little error reduction. This means, applied to real measurements, such retrievals would most certainly be filtered out by the post-processing. Note that due to the non-Lambertian surface, the retrieved albedo may have values larger than one.

The systematic errors of the *SIF* scenario are approximately 0.3ppm but the retrieval did not converge for both polarization directions when the solar zenith angle was 20°.

FOCAL’s state vector composition does not allow to change the number of dry-air particles in the atmospheric column, e.g., by fitting the surface pressure, or a shift of the temperature profile. As a result, relative errors of the number of dry-air particles computed from the meteorological profiles directly translate into relative errors of the retrieved XCO_2. For example, a 1hPa error of the surface pressure will result in a XCO_2 error of about 0.4ppm.

2.3 Summary

As accurate XCO_2 retrievals will probably always require a rigorous cloud and aerosol screening, we concentrated on scenarios with scattering optical thicknesses in the range of about 0.03 and 0.70. The quality of the spectral fits is usually considerably better than expected from the instrumental noise which is particularly the case in the SIF and the CO_2 fit windows. Figure 2.2 shows some example fit residuals.

Systematic errors of XCO_2 ranged from -3.4ppm to 3.0ppm and were usually smaller than ±0.3ppm (for the tested scenarios). The stochastic uncertainty of XCO_2 was typically about 1.0ppm (Figure 2.1).

3 Validation with TCCON

The validation results shown in this section are valid for FOCAL v10. The applied methods are similar to those described in BESD's comprehensive error characterization Report **/CECRv3, 2017/** and the product validation and inter-comparison Report **/PVIRv5, 2017/** of ESA's GHG CCI project and partly also in the publication of **/Reuter et al., 2020/**. For all comparisons, averaging kernels have been applied and the influence of the smoothing error reduced as described in Section 5.2 of ESA's GHG CCI+ product user guide version 4 (PUGv4) for the FOCAL XCO_2 OCO-2 data product CO2_OC2_FOCA **/PUGv4, 2022/**.

3.1 Co-location

FOCAL's XCO_2 has been validated with TCCON GGG2014 measurements **/Wunch et al., 2011/**. The co-location criteria are defined by a maximum time difference of two hours, a maximum spatial distance of 500km, and a maximum surface elevation difference of 250m. Additionally, only TCCON sites with at least 1000 co-locations (4 in the case of daily, weekly, or monthly averages) covering a time period of at least two years are taken into account.

Figure 3.1 shows all 2331159 co-located FOCAL and TCCON XCO_2 retrieval results used for the validation study. One can see that the temporal sampling differs from site to site and that FOCAL captures the year-to-year increase and the seasonal features well.

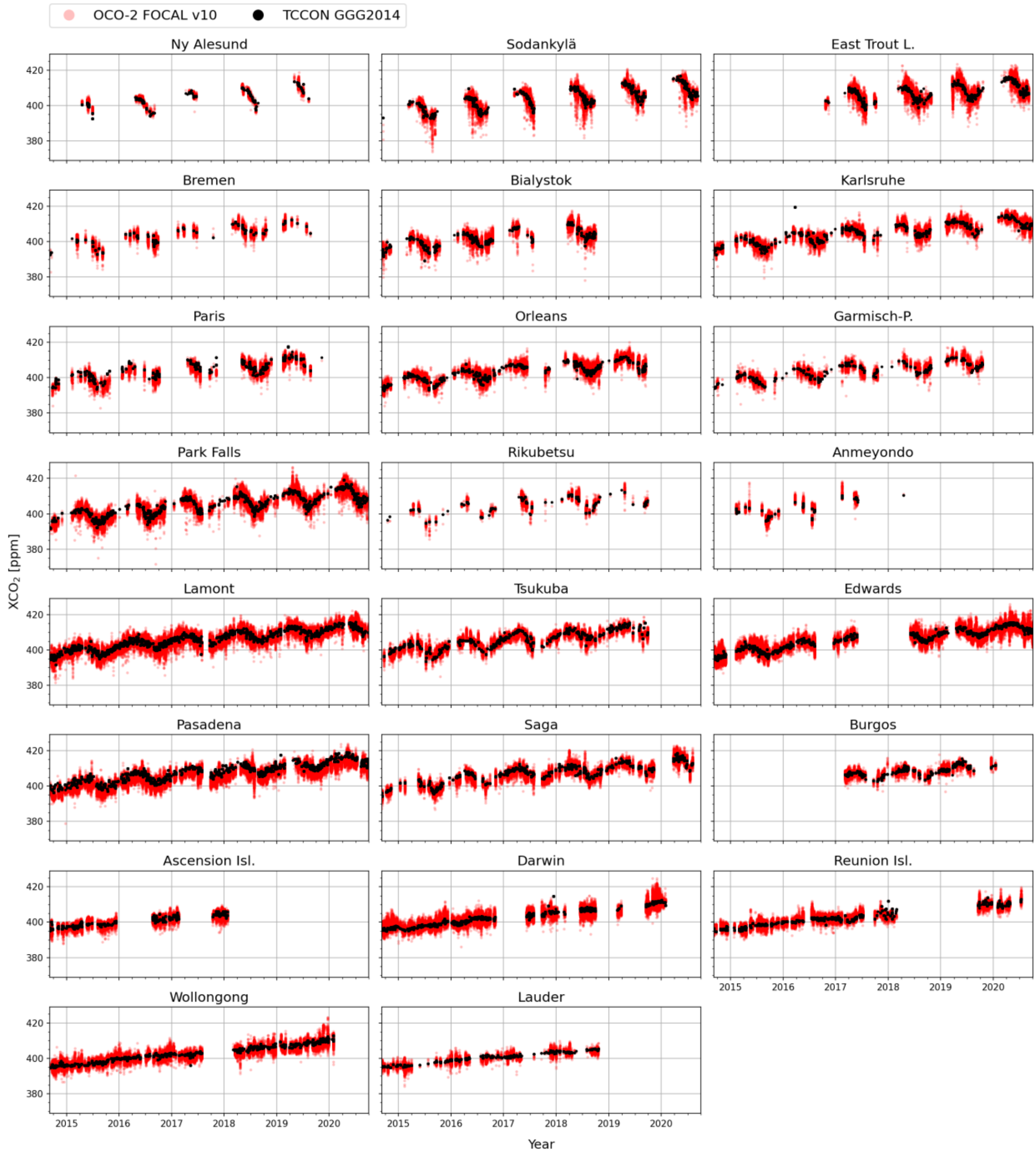

Figure 3.1: Co-located FOCAL and TCCON XCO_2 retrieval results used for the validation study. The TCCON sites are order from top/left to bottom/right by average latitude of the co-located satellite soundings.

3.2 Daily, weekly, and monthly averages

For some applications, it is expected that FOCAL XCO_2 data will be aggregated to "super soundings" averaging, e.g., all soundings of an orbit in a surrounding of a target. Also FOCAL XCO_2 data might be used to compute L3 (level 3) products, e.g., in the manner of gridded monthly averages. With such application in the mind, we computed daily, weekly, and monthly averages of the FOCAL and TCCON co-locations at each TCCON site. In order to improve the robustness, daily, weekly, and monthly averages are only calculated when averaging at least 10, 30, or 50 individual soundings, respectively. As an example, Figure 3.2 shows the daily, weekly, and monthly FOCAL XCO_2 averages for the Lamont and Darwin TCCON sites. Due to OCO-2's data density, it is often the case that one overpass generates many co-colocations. This considerably reduces the scatter of the daily averages compared to the individual soundings.

Note that FOCAL reports only on the stochastic uncertainty of the individual soundings. In the case of daily, weekly, and monthly averages we computed the corresponding uncertainties by applying the rules of error propagation under the assumption of uncorrelated errors.

3.3 General overview

The overall agreement of the FOCAL data (and its averages) with TCCON data at all sites is illustrated in Figure 3.3. The histograms of the difference (FOCAL – TCCON) show in all cases a near Gaussian distribution with a center between -0.43ppm and -0.29ppm. The standard deviation of the difference reduces from 1.88ppm for individual soundings to 1.14ppm for monthly averages. The FOCAL vs. TCCON heat maps show a pronounced clustering along the one-to-one line for all cases. This is supported by a good agreement of the orthogonal distance regression with the one-to-one line and high Pearson correlation coefficients between 0.94 for individual soundings and 0.97 for monthly averages.

These results provide a first rough overview of FOCAL's agreement with TCCON. However, except for an average bias, they do not allow to separate systematic and stochastic error components.

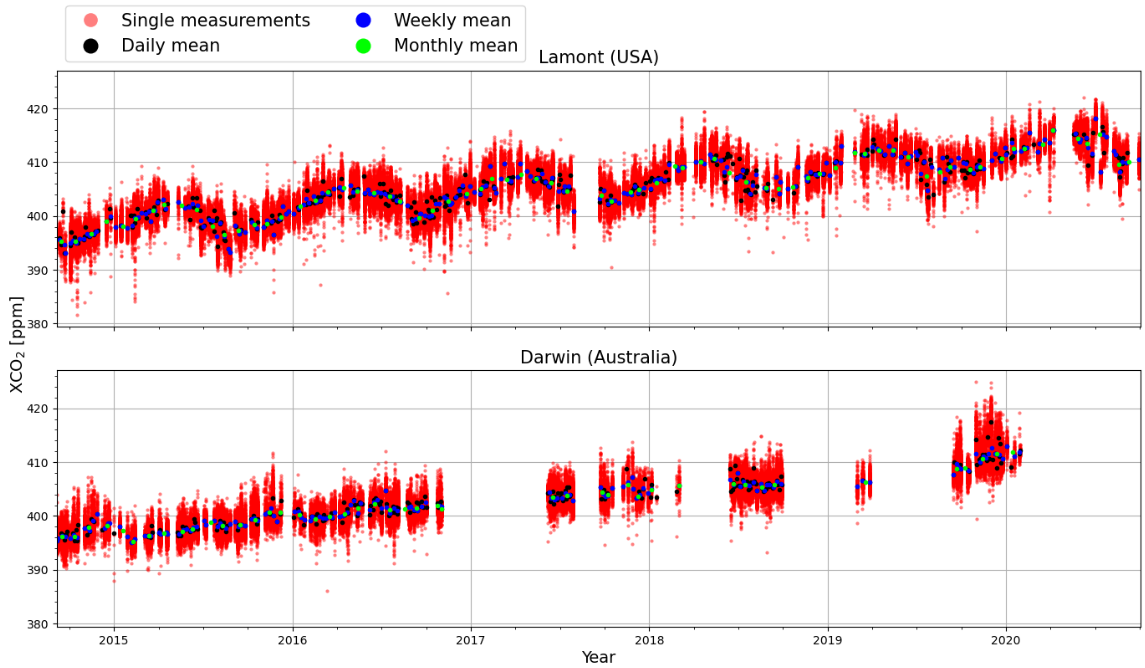

Figure 3.2: Co-located FOCAL XCO_2 retrieval results and their daily, weekly, and monthly averages at the TCCON sites Lamont (top) and Darwin (bottom) used for the validation study.

3.4 Stochastic and systematic error components

The method described in the following allows us to separate the stochastic errors from potential regional or seasonal biases as well as from a linear drift.

3.4.1 Per site performance statistics

For the co-locations of each site, we compute the FOCAL minus TCCON differences ΔX and fit the following bias model:

3-1
$$\Delta X = a_0 + a_1 t + a_2 \sin(2\pi t + a_3) + \varepsilon$$

Here, t is the time of the measurements in fractional years, a_{0-3} the free fit parameters from which we compute the systematic error components, and ε the fit residuum. Figure 3.4 shows at the example of the TCCON sites Lamont and Darwin the fitted bias functions for the individual soundings, daily, weekly, and monthly averages.

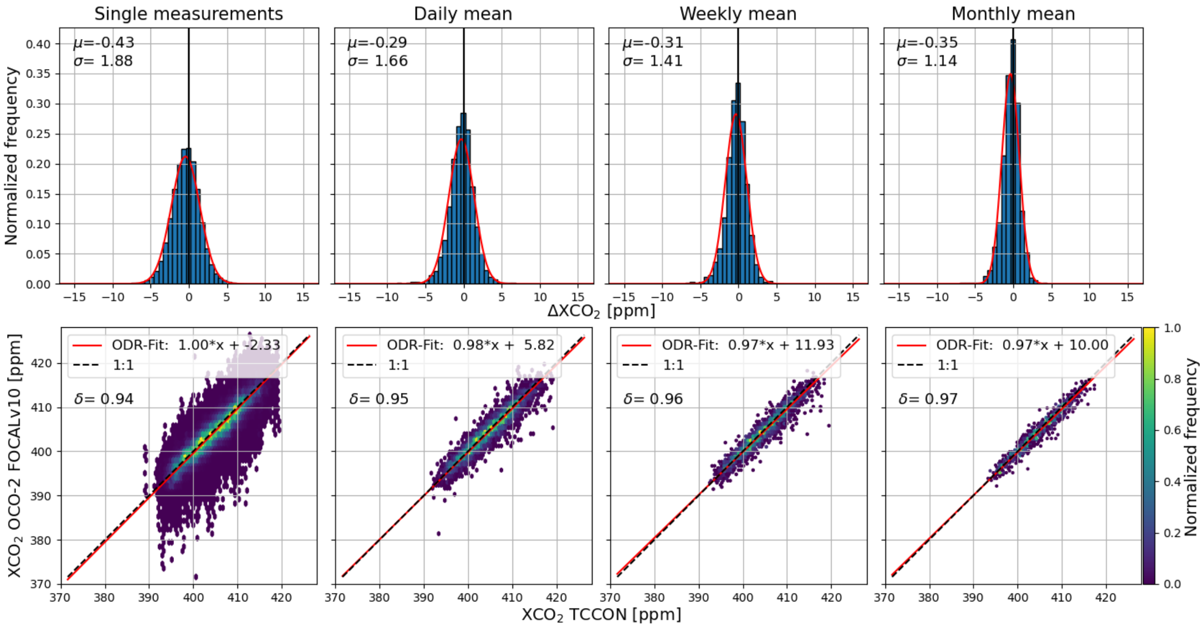

Figure 3.3: Overall overview on the agreement of the FOCAL data (and its averages) with TCCON data at all sites. Top: Normalized histograms of the difference FOCAL – TCCON. Bottom: Heat maps TCCON vs. FOCAL including one-to-one line, orthogonal distance regression (ODR), and Pearson correlation coefficient δ.

We compute the station or regional bias Δ_{reg} from the average (ave) of the fit values:

3-2 $$\Delta_{reg} = \text{ave}[a_0 + a_1 t + a_2 \sin(2\pi t + a_3)]$$

The seasonal bias Δ_{sea} is computed from the standard deviation (std) of the seasonal component of the fit:

3-3 $$\Delta_{sea} = \text{std}[a_2 \sin(2\pi t + a_3)]$$

It shall be noted that the vector t consists only of the time of the measurements. This means, Δ_{sea} is only computed from those parts of the seasonal cycle actually covered by observations.

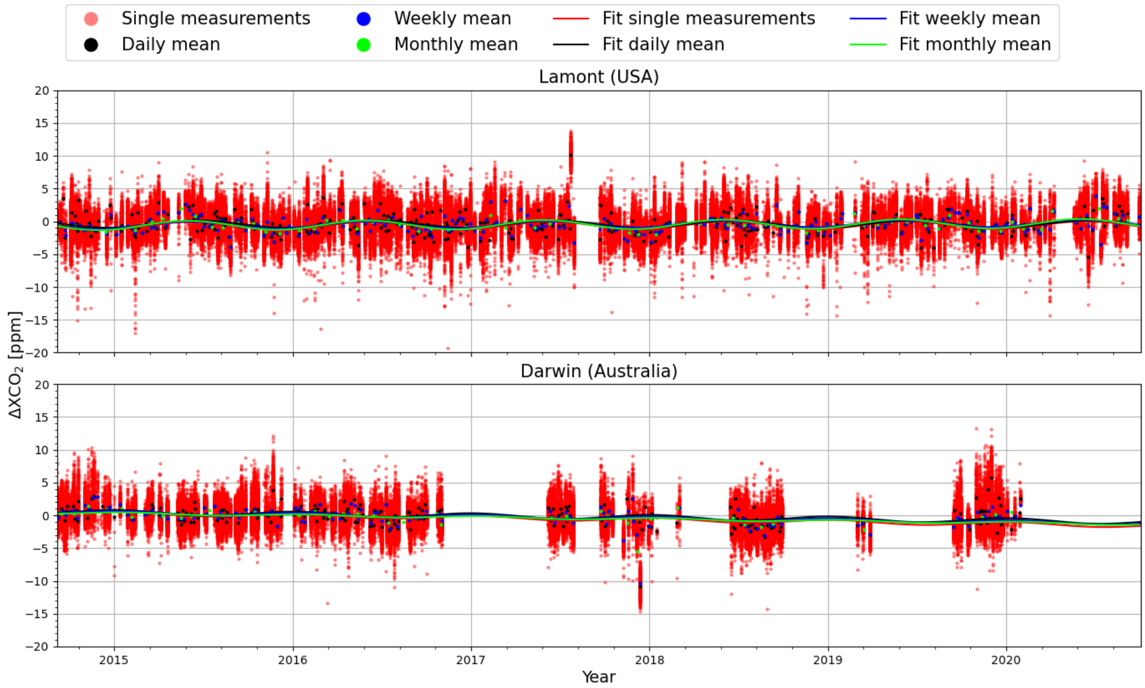

Figure 3.4: ΔXCO_2 (FOCAL – TCCON) for the co-locations of the single measurements, daily, weekly, and monthly averages at the TCCON sites Lamont (top) and Darwin (bottom). Additionally, the corresponding fits of the bias model (Eq. 3-1) are shown.

The linear drift corresponds to the fit parameter $\Delta_{dri} = a_1$, and the single sounding precision, i.e., the stochastic retrieval uncertainty σ, is computed from the standard deviation of the residuum.

3-4 $$\sigma = \mathrm{std}[\varepsilon]$$

We define the spatiotemporal bias Δ_{spt} as combination of regional and seasonal bias.

3-5 $$\Delta_{spt} = \sqrt{\Delta_{reg}{}^2 + \Delta_{sea}{}^2}$$

The FOCAL retrieval algorithm reports on the XCO_2 stochastic uncertainty σ'_{rep} for each sounding. From these values, we compute the average reported uncertainty σ_{rep} per station by:

3-6 $$\sigma_{rep} = \sqrt{\mathrm{ave}\left({\sigma'_{rep}}^2\right)}$$

3.4.2 Summarizing performance statistics

Based on the per site statistics, the following summarizing performance statistics are calculated.

The average site bias $\overline{\Delta_{reg}}$ and the site-to-site variability is computed from the mean and the standard deviation of the individual site biases:

3-7 $$\overline{\Delta_{reg}} = \mathrm{ave}(\Delta_{reg}) \pm \mathrm{std}(\Delta_{reg})$$

The average seasonal bias $\overline{\Delta_{sea}}$ is computed by:

3-8 $$\overline{\Delta_{sea}} = \mathrm{avg}(\Delta_{sea})$$

The overall spatiotemporal bias $\overline{\Delta_{spt}}$ is computed by:

3-9 $$\overline{\Delta_{spt}} = \sqrt{\overline{\Delta_{reg}}^2 + \overline{\Delta_{sea}}^2}$$

The average drift and the drift uncertainty is computed by:

3-10 $$\overline{\Delta_{dri}} = \mathrm{ave}(\Delta_{dri}) \pm \mathrm{std}(\Delta_{dri})$$

As the linear drift can be assumed to be globally constant, the station-to-station standard deviation of the linear drift can be considered a measure of its uncertainty. The overall single sounding precision and reported uncertainty are computed by:

3-11 $$\overline{\sigma} = \sqrt{\mathrm{ave}(\sigma^2)}$$

3-12 $$\overline{\sigma_{rep}} = \sqrt{\mathrm{ave}\left({\sigma_{rep}}^2\right)}$$

3.5 Results

The results of all site performance statistics as well as the summarizing performance statistics for individual soundings, daily, weekly, and monthly averages are illustrated in Figure 3.5. Based on this figure, it can first be noted that averaging does not have a substantial impact on the validation results for the systematic error components. This is especially the case for the summarizing performance statistics which are similar for individual soundings, daily, weekly, and monthly averages. Therefore, it is sufficient that we primarily concentrate on the results for individual soundings from now on and Table 3.1 lists only values of the statistics for individual soundings.

However, the results for the stochastic error component show some important differences. The overall result for the stochastic error of the individual soundings amounts to 1.69ppm which agrees well with the corresponding reported uncertainty of 1.69ppm. This is not the case for the results of the averages. The actual stochastic error reduces for daily (1.39ppm), weekly (1.11ppm), and monthly (0.80ppm) averages, but the reduction is far less pronounced as for the reported uncertainty which has been computed under the assumption of uncorrelated errors. Therefore, it has to be expected that the separation of systematic and stochastic errors by Eq. 3-1 is incomplete at least for the individual soundings. In other words, it can be expected that parts of the residuum ε of Eq. 3-1 for the individual soundings are actually of systematic origin.

For this reason, we grouped the residuum into bins consisting of $n = 1, 2, 3, \cdots$ elements and analyzed its standard deviation as function of the bin size. As the reported retrieval precision is usually relatively constant at one TCCON site, it should be expected that the standard deviation of the binned residuum scales approximately with $1/\sqrt{n}$. We performed this experiment for the TCCON site Lamont because of the large number of co-locations. As shown in Figure 3.6 (top/left), the actual precision (standard deviation of the binned residuum) of the individual soundings does not follow the curve expected for uncorrelated errors. In contrast, the actual precision of daily (Figure 3.6, top/right), weekly (Figure 3.6, bottom/left), and monthly averages (Figure 3.6, bottom/right) agrees well with the expectation for uncorrelated errors. These results may differ in detail from TCCON site to TCCON site, but indicates that the errors of the individual soundings may have additional systematic components not covered by the seasonal component of Eq. 3-1.

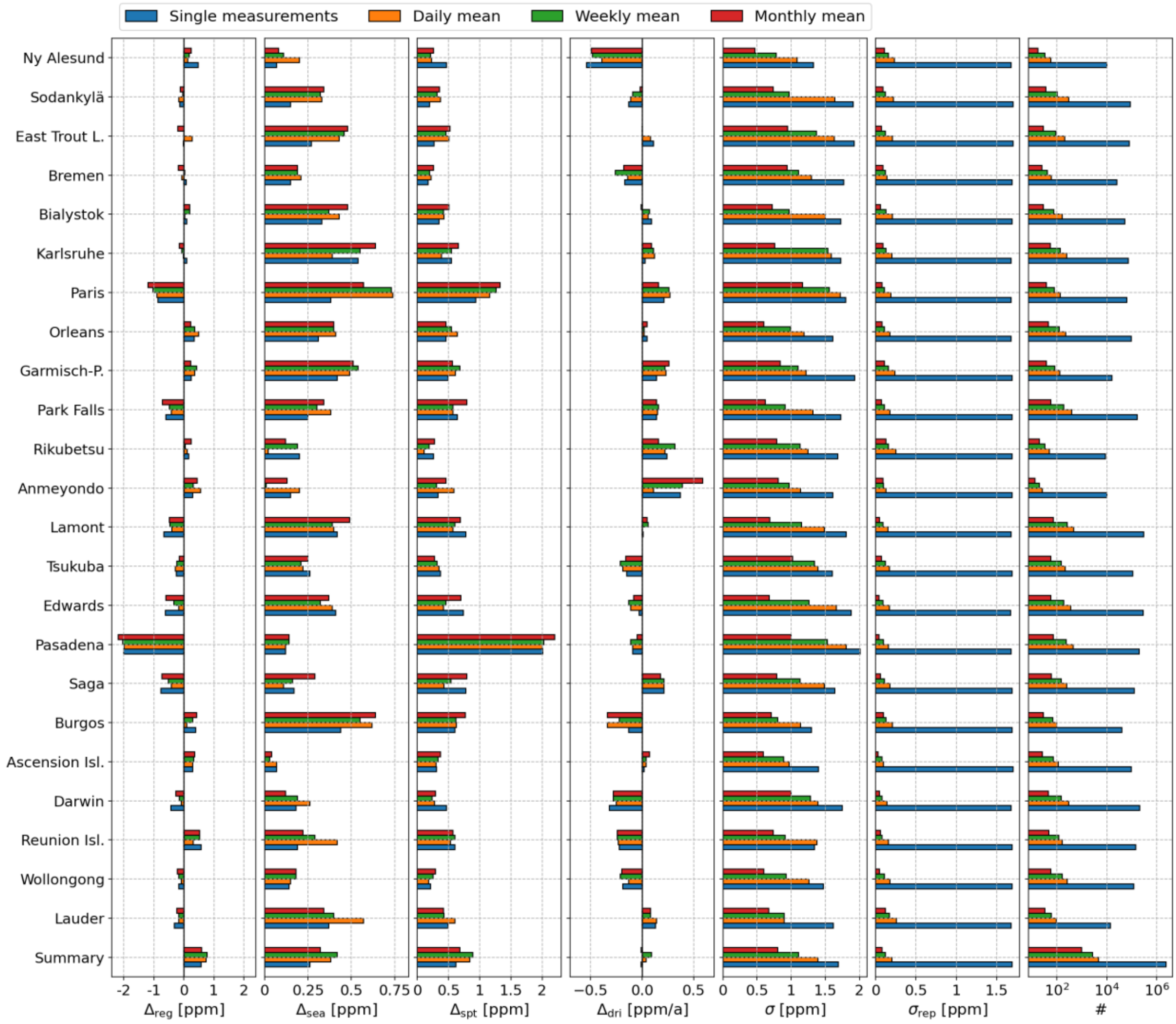

Figure 3.5: Validation results for FOCAL single measurements, daily, weekly, and monthly averages. From left to right, the figure shows the per site performance statistics (Section 3.4.1) regional (Δ_{reg}), seasonal (Δ_{sea}), and spatiotemporal bias (Δ_{spt}), the linear drift (Δ_{dri}), the actual (σ) and reported precision (σ_{rep}), and the number of soundings (#). TCCON sites are order from top to bottom by average latitude of the co-located satellite soundings. The last row includes the summarizing performance statistics as defined in Section 3.4.2.

Table 3.1: Validation results for FOCAL single measurements. From left to right, the table lists the per site performance statistics (Section 3.4.1) regional (Δ_{reg}), seasonal (Δ_{sea}), and spatiotemporal bias (Δ_{spt}), the linear drift (Δ_{dri}), the actual (σ) and reported precision (σ_{rep}), and the number of soundings (#). TCCON sites are order from top to bottom by average latitude of the co-located satellite soundings. The last row includes the summarizing performance statistics as defined in Section 3.4.2.

Station	Δ_{reg} [ppm]	Δ_{sea} [ppm]	Δ_{spt} [ppm]	Δ_{dri} [ppm/a]	σ [ppm]	σ_{rep} [ppm]	#
Ny Alesund	0.47	0.07	0.47	-0.54	1.33	1.68	10223
Sodankylä	-0.13	0.15	0.20	-0.13	1.91	1.70	86093
East Trout Lake	-0.02	0.27	0.27	0.11	1.92	1.70	80390
Bremen	0.08	0.15	0.17	-0.17	1.77	1.69	25681
Bialystok	0.09	0.33	0.35	0.09	1.73	1.69	51219
Karlsruhe	0.09	0.54	0.55	0.03	1.73	1.68	73437
Paris	-0.86	0.38	0.94	0.21	1.80	1.68	63569
Orleans	0.34	0.31	0.46	0.05	1.61	1.68	95043
Garmisch-P.	0.25	0.42	0.49	0.14	1.93	1.69	15863
Park Falls	-0.59	0.25	0.64	0.14	1.73	1.69	167324
Rikubetsu	0.16	0.20	0.26	0.24	1.68	1.69	8879
Anmeyondo	0.29	0.15	0.33	0.37	1.61	1.69	9904
Lamont	-0.66	0.42	0.78	0.01	1.81	1.68	301063
Tsukuba	-0.25	0.26	0.37	-0.15	1.60	1.69	110964
Edwards	-0.61	0.41	0.74	-0.03	1.88	1.67	287748
Pasadena	-2.00	0.12	2.01	-0.09	2.01	1.68	197002
Saga	-0.76	0.17	0.78	0.21	1.64	1.69	127152
Burgos	0.40	0.44	0.60	-0.13	1.30	1.69	39354
Ascension Isl.	0.30	0.07	0.31	0.02	1.40	1.70	95046
Reunion Isl.	-0.44	0.18	0.47	-0.32	1.75	1.68	210700
Darwin	0.57	0.19	0.60	-0.22	1.34	1.69	140479
Wollongong	-0.16	0.14	0.21	-0.19	1.47	1.69	120422
Lauder	-0.32	0.37	0.49	0.13	1.62	1.68	13604
Summary	**-0.16±0.57**	**0.26**	**0.62**	**-0.01±0.20**	**1.69**	**1.69**	**2331159**

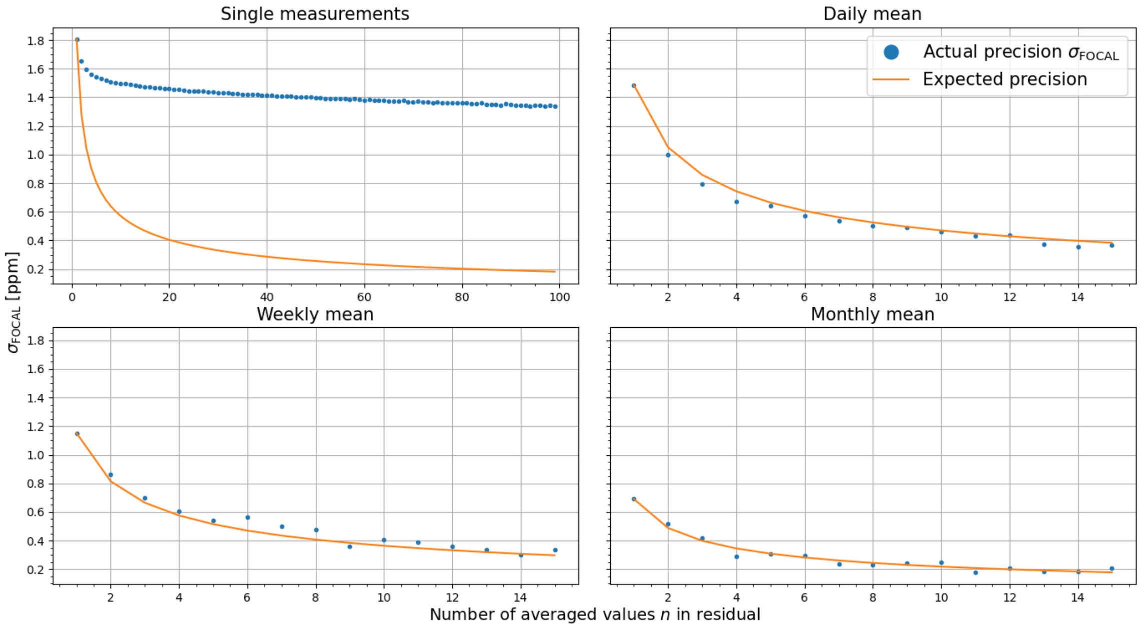

Figure 3.6: Actual and expected retrieval precision of FOCAL computed from residuals with increasing bin size for the TCCON site Lamont for single measurements (top/left), daily (top/right), weekly (bottom/left), and monthly averages (bottom/right).

The validation results for the individual soundings (Table 3.1, Figure 3.5) show that there is only a small overall average bias of -0.16ppm. Regional biases estimated from the site-to-site bias variability amount to 0.57ppm and are strongly influenced by the relatively large negative bias of -2ppm at the TCCON site Pasadena. The average seasonal and spatiotemporal bias amounts to 0.26ppm and 0.62ppm, respectively. The overall linear drift of 0.01ppm/a is much smaller than its site-to-site variability of 0.2ppm and, therefore, considered not significant.

Additionally, a measure for the year-to-year stability is computed as follows. For each TCCON site, the residual ε of the bias fit (Eq. 3-1) is smoothed by a running average of 365 days. Only days where more than 10 co-locations contribute to the running average of at least 5 TCCON sites are further considered. At these days, the station-to-station average is calculated (Figure 3.7, black line).

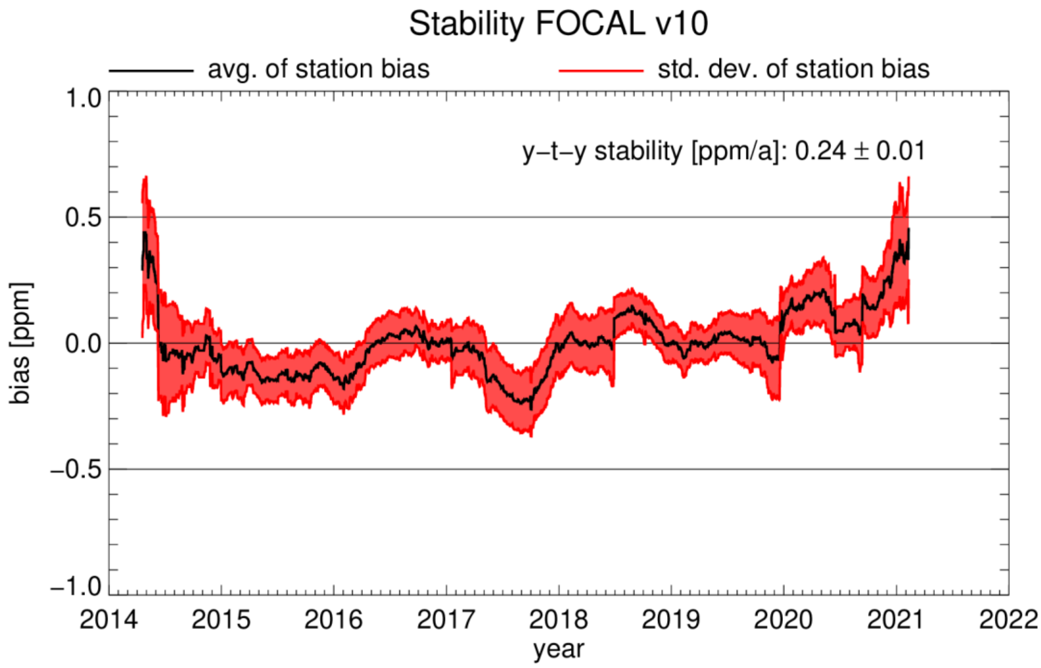

Figure 3.7: Stability analyses for FOCAL. The black curve shows the average station bias and the red curves its uncertainty represented by the station-to-station standard deviation

The corresponding expected uncertainty is computed from the standard error of the mean (derived from the station-to-station standard deviation and the number of stations) and by error propagation of the reported single sounding uncertainties (Figure 3.7, red line). For FOCAL, the average is always between about -0.25ppm and 0.45ppm with an uncertainty of typically about 0.15ppm. Most of the time, the average is not significantly different from zero, i.e., its two sigma uncertainty is larger than its absolute value. Due to the relatively large uncertainty, we decided to compute not the maximum minus minimum as a measure for the year-to-year stability because this quantity can be expected to increase with length of the time series simply due to statistics. Therefore, we estimate the year-to-year stability by randomly selecting pairs of dates with a time difference of at least 365 days. For each selection we computed the difference modified by a random component corresponding to the estimated uncertainty. From 1000 of such pairs we compute the standard deviation as estimate for the year-to-year stability. We repeat this experiment 1000 times and compute the average (0.24ppm) and standard deviation

(0.01ppm). From this, we conclude that the year-to-year stability is 0.24ppm/a (Figure 3.7).

3.6 Summary

We validated the FOCAL v10 XCO_2 data product with TCCON GGG2014 data of the years 2014 – 2021. The validation has been performed for daily, weekly, and monthly averages as well as for single soundings. Analyzing the single soundings without temporal averaging, we find that the overall bias of the FOCAL data amounts to -0.16ppm. Regional biases vary from site to site by 0.57ppm. Seasonal and spatiotemporal biases amount on average to 0.26ppm and 0.62ppm, respectively. We found no significant linear drift (-0.01±0.20ppm). In the context of the systematic error characteristics, it shall be noted that **/Wunch et al., 2010, 2011/** specifies the accuracy (1σ) of TCCON to be about 0.4ppm. This means, e.g., that it cannot be expected to find regional biases considerably less than 0.4ppm using TCCON as reference. We find that the inferred systematic errors, i.e., regional, seasonal, and spatiotemporal biases as well as linear drift, do not critically depend on averaging. The year-to-year stability has been estimated to be 0.24ppm/a. The overall precision of the individual soundings is 1.69ppm which agrees well with the corresponding reported uncertainty of 1.69ppm. The overall precision improves for daily (1.39ppm), weekly (1.11ppm), and monthly (0.80ppm) averages. We find indications that the estimated precision of the individual soundings does actually comprise not only purely stochastic but also residual unknown systematic components. No such indications were found for the daily, weekly, and monthly averages.

4 Comparison with CAMS model results

In this section we compare six months in 2015 of post-filtered and bias corrected FOCAL v10 XCO_2 results with corresponding values of the CAMS v15r4 model accounting for FOCAL's column averaging kernels as explained in Section 5 of ESA's GHG CCI+ product user guide version 4 (PUGv4) for the FOCAL XCO_2 OCO-2 data product CO2_OC2_FOCA **/PUGv4, 2022/**.

Figure 4.1 shows 5°×5°monthly gridded values for six months (Feb., Apr., Jun., Aug., Oct., and Dec. 2015) of FOCAL data and Figure 4.2 shows corresponding values of CAMS v15r4 data. The main large scale spatial and temporal patterns are similar for FOCAL and CAMS with largest and smallest values in the northern hemisphere in April and August, respectively. Differences become larger at smaller scales, e.g., FOCAL sees larger values in natural and anthropogenic source regions of Sub-Saharan Africa and East Asia, e.g., in April but also above the Sahara, e.g., in August. However, it shall be noted that sometimes only few data points are in the corresponding grid boxes.

In grid boxes with more than 100 soundings, the standard error of the mean becomes negligible (~0.1ppm). Therefore, the difference between FOCAL and CAMS in such grid boxes can be interpreted as systematic temporal and regional mismatch or bias. The heat map shown in Figure 4.3 bases on these grid boxes. The standard deviation of this systematic mismatch (including also representation errors) amounts to 0.97ppm and the correlation between FOCAL and CAMS is 0.90.

The standard deviation of the single sounding mismatch after subtracting the systematic mismatch amounts to 1.3ppm which is somewhat smaller than the average reported uncertainty of 1.7ppm. The overall average offset (FOCAL – CAMS) is 0.08ppm.

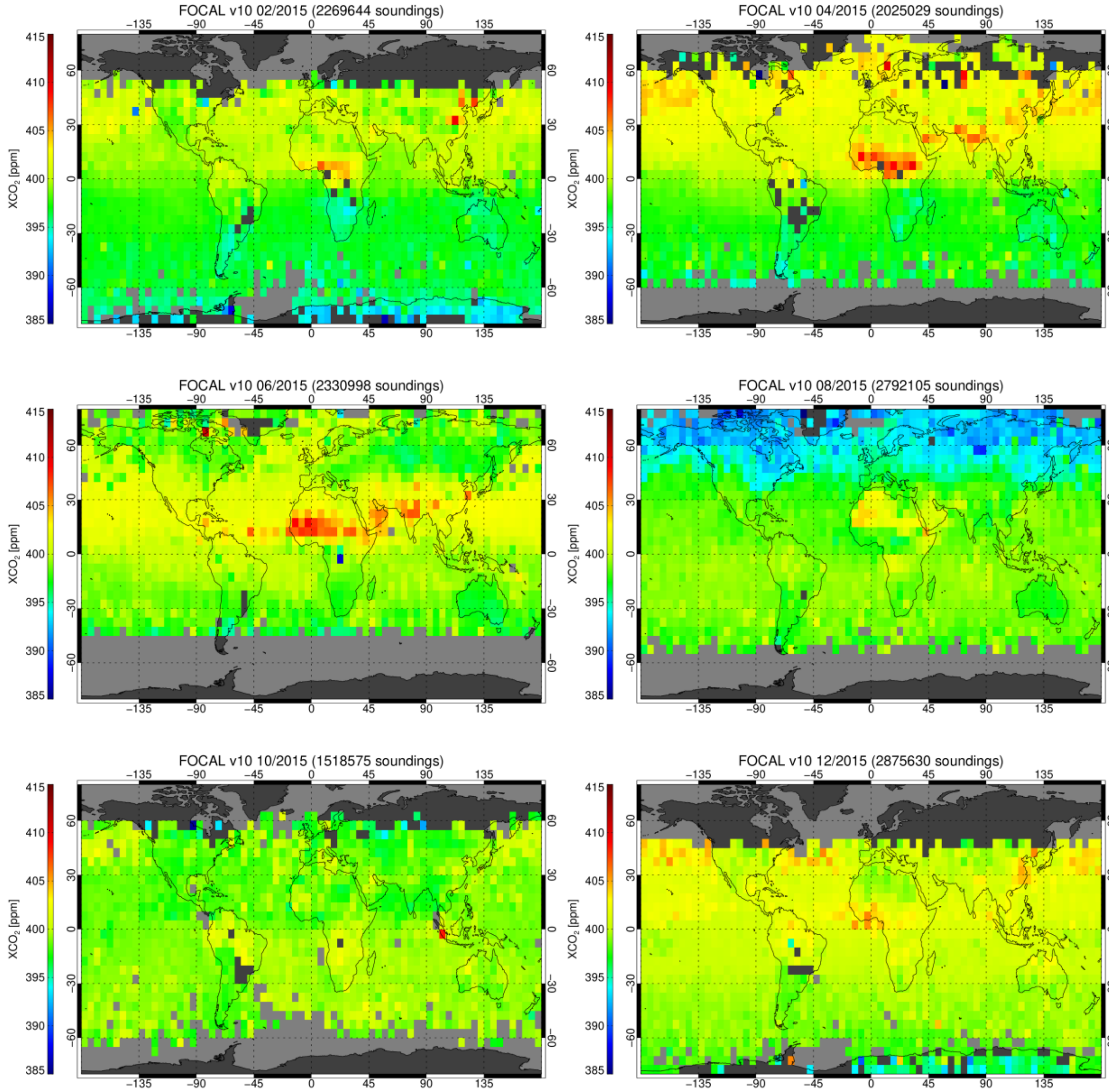

Figure 4.1: FOCAL v10 monthly mean XCO_2 gridded to 5°×5°. From top/left to bottom/right: Feb., Apr., Jun., Aug., Oct., and Dec. 2015.

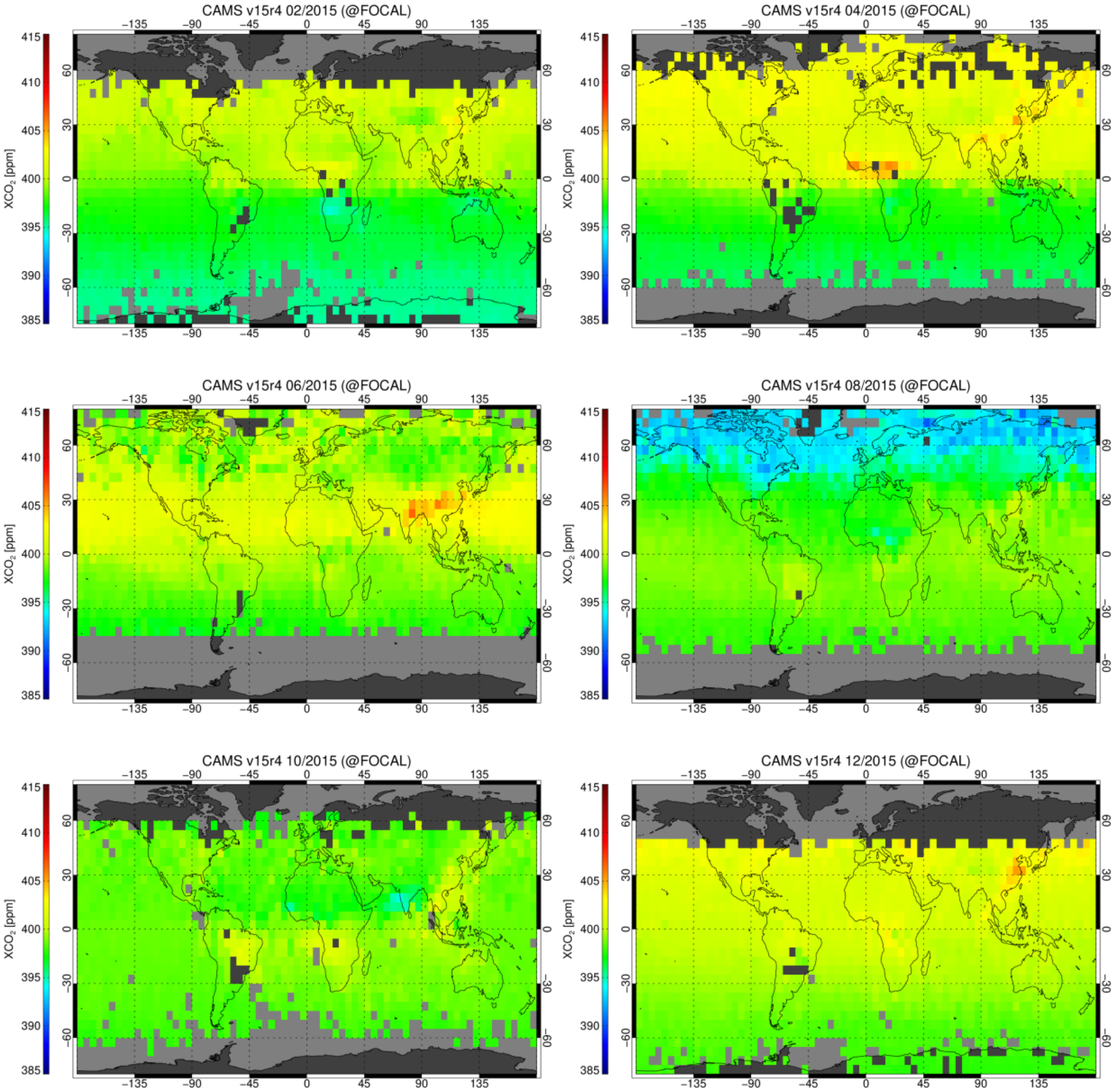

Figure 4.2: CAMS v15r4 monthly mean XCO_2 sampled as FOCAL and gridded to 5°×5°. From top/left to bottom/right: Feb., Apr., Jun., Aug., Oct., and Dec. 2015.

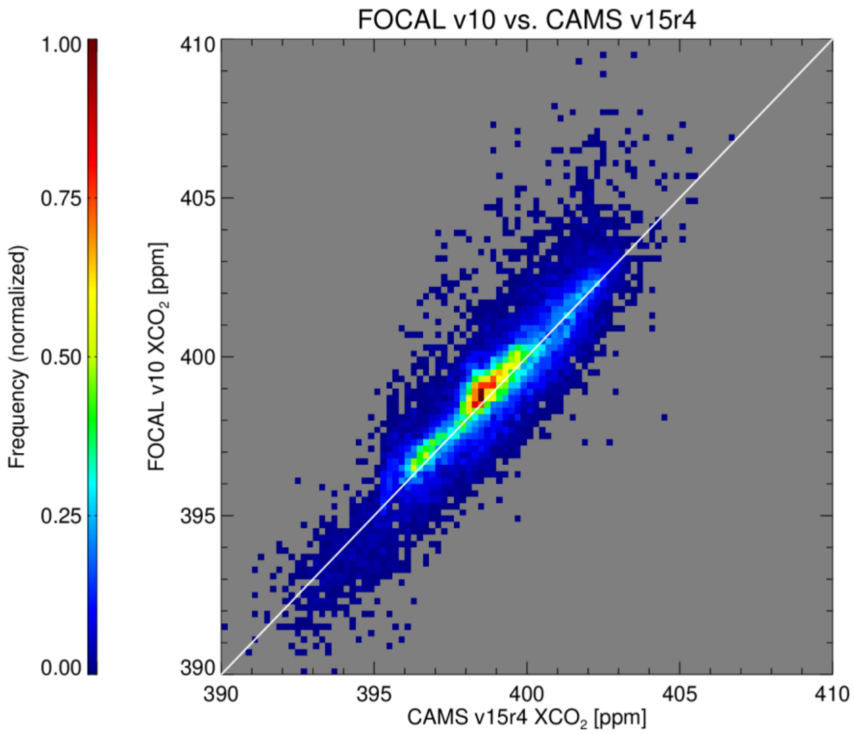

Figure 4.3: Heat map of FOCAL vs. CAMS XCO_2 data on the basis of monthly mean 5°×5° grid boxes including more than 100 data points.

5 Comparison with NASA's operational OCO-2 XCO_2 L2 product

In this section we compare the same months of post-filtered and bias corrected FOCAL v10 XCO_2 results with NASA's operational OCO-2 L2 product v10.2 (**/O'Dell et al., 2018/**, **/Kiel et al., 2019/**). Our comparison method is similar to what has been done in Section 4. However, as FOCAL and the NASA product feature different samplings, we first gridded the NASA product and compared FOCAL with corresponding grid box averages. In order to improve the comparability, both data products have been adjusted for a common a priori as explained in Section 5 of ESA's GHG CCI+ product user guide version 4 (PUGv4) for the FOCAL XCO_2 OCO-2 data product CO2_OC2_FOCA **/PUGv4, 2022/**.

Comparing Figure 4.1 with Figure 5.1 shows similar large scale temporal and spatial patterns and also the enhancements due to the anthropogenic source regions of East Asia in April are somewhat similar. Additionally, the total number of soundings used to compute the shown maps are very similar.

The most obvious differences between the gridded FOCAL and the gridded NASA product are the somewhat larger variability of FOCAL and the larger values in the Sub-Saharan biomass burning region seen by FOCAL. In this context it shall be noted that the FOCAL L2 and the NASA L2 product are sampled differently.

Similarly, as done for the model comparison, we concentrate only on grid boxes with more than 100 FOCAL and NASA soundings so that the standard error of the mean becomes negligible (~0.1ppm). Therefore, the difference between FOCAL and NASA in such grid boxes can be interpreted as systematic temporal and regional mismatch or bias. The heat map shown in Figure 5.2 bases on these grid boxes. The standard deviation of the systematic mismatch (including also representation errors) amounts to 0.83ppm and the correlation between FOCAL and NASA is 0.92.

FOCAL scatters within the grid boxes with a standard deviation of 1.4ppm which is somewhat smaller than the average reported uncertainty of 1.7ppm. The overall average offset (FOCAL – NASA) is -0.09ppm.

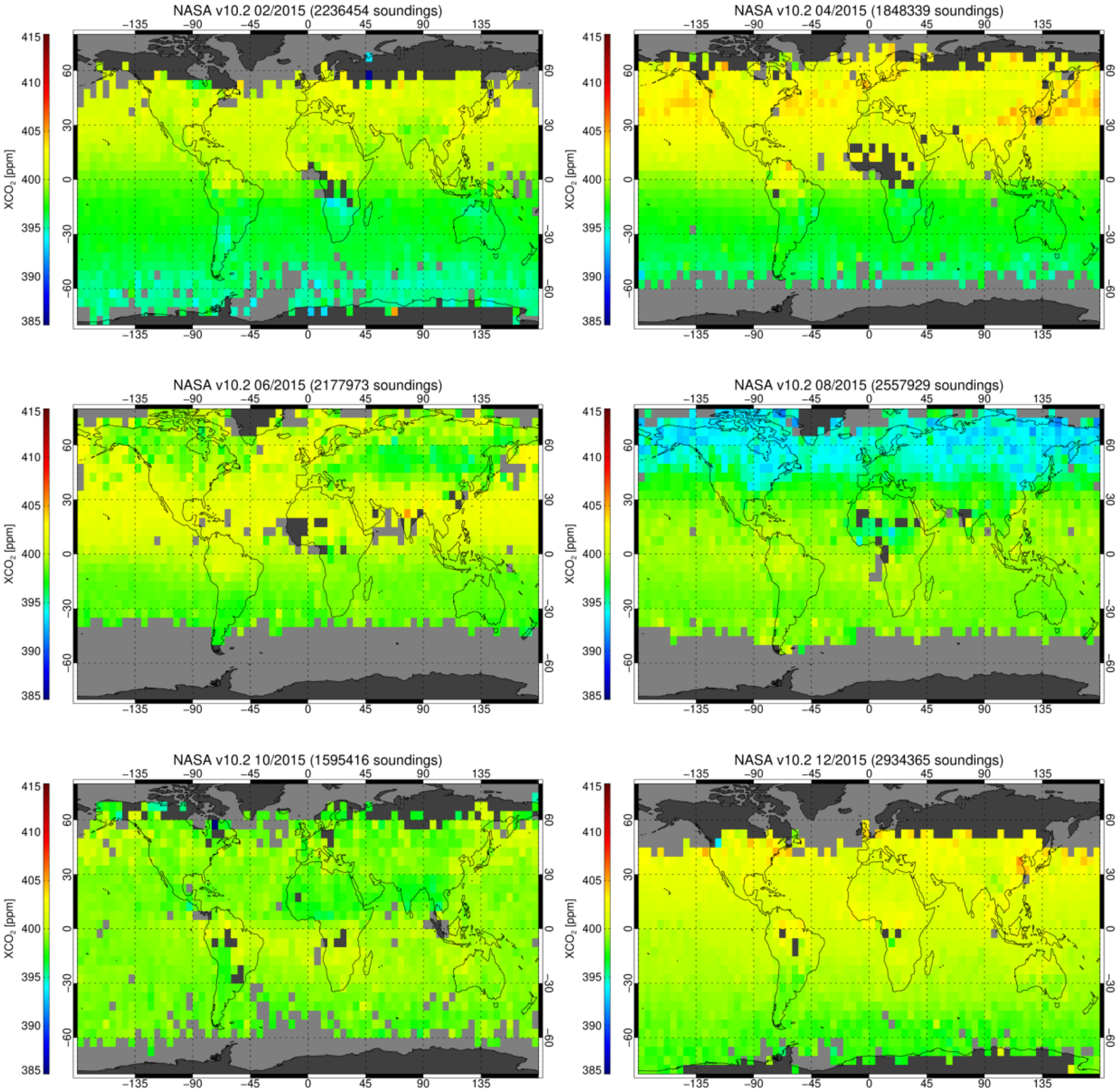

Figure 5.1: NASA's operational OCO-2 v10.2 XCO_2 L2 product gridded to 5°×5° monthly means. From top/left to bottom/right: Feb., Apr., Jun., Aug., Oct., and Dec. 2015.

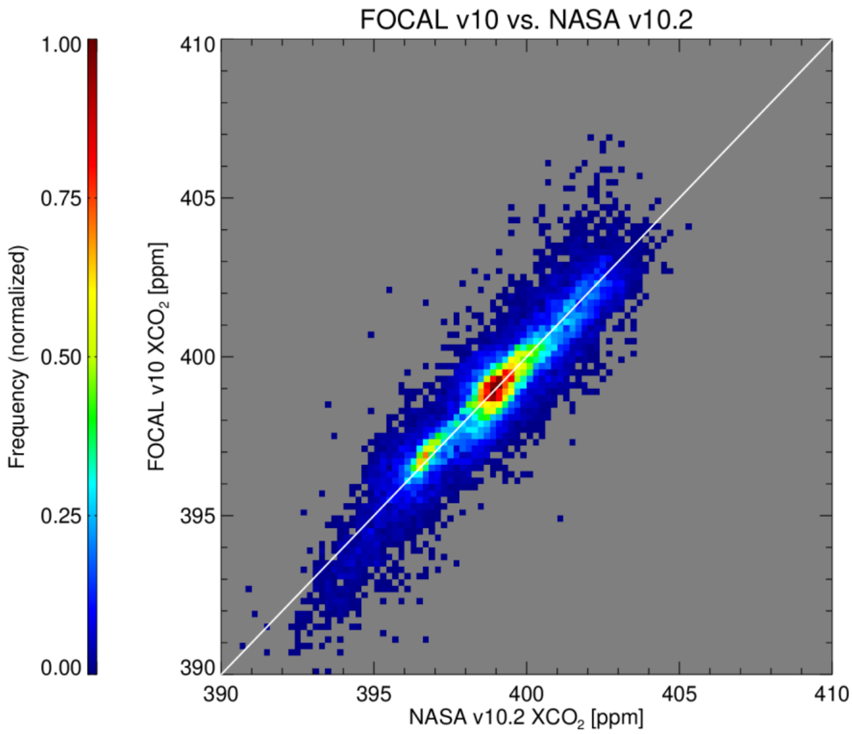

Figure 5.2: Heat map of FOCAL vs. NASA OCO-2 v10.2 XCO_2 data on the basis of monthly mean 5°×5° grid boxes including more than 100 data points.

6 Acknowledgements

The FOCAL OCO-2 XCO_2 retrieval development, data processing and analysis has received co-funding from ESA's Climate Change Initiative (CCI+) via project GHG-CCI+ (contract 4000126450/19/I-NB, https://climate.esa.int/en/projects/ghgs), EUMETSAT via the FOCAL-CO2M study (contract EUM/CO/19/4600002372/RL), the European Union via the Horizon 2020 (H2020) projects VERIFY (Grant Agreement No. 776810, http://verify.lsce.ipsl.fr) and CHE (Grant Agreement No. 776186, https://www.che-project.eu), and by the State and the University of Bremen.

The following data sets have been used for the FOCAL OCO-2 XCO_2 retrieval development, data processing and analysis. The OCO-2 data were produced by the OCO-2 project at the Jet Propulsion Laboratory, California Institute of Technology, and obtained from the OCO-2 data archive maintained at the NASA Goddard Earth Science Data and Information Services Center. NASA provided the MODIS L2 collection 6 MYD35 cloud mask data. ABSCO cross sections were provided by NASA and the ACOS/OCO-2 team. ECMWF provided the used meteorological profiles. The used solar spectra were made available by R. L.Kurucz and G.C. Toon. The used chlorophyll fluorescence spectrum has been published by U.Rascher. TCCON data were obtained from the TCCON Data Archive, hosted by the Carbon Dioxide Information Analysis Center (CDIAC)

7 References

/ATBDv3, 2021/ M. Reuter, M. Hilker, S. Noël, M. Buchwitz, O. Schneising, H. Bovensmann, and J. P. Burrows: Algorithm Theoretical Basis Document Version 3 (ATBDv3) - Retrieval of XCO2 from the OCO-2 satellite using the Fast Atmospheric Trace Gas Retrieval (FOCAL), ESA Climate Change Initiative "Plus" (CCI+) for the Essential Climate Variable (ECV) Greenhouse Gases (GHG), ESA GHG CCI+, 2021

/Bovensmann et al., 2010/ Bovensmann, H., Buchwitz, M., Burrows, J. P., Reuter, M., Krings, T., Gerilowski, K., Schneising, O., Heymann, J., Tretner, A., and Erzinger, J.: A remote sensing technique for global monitoring of power plant CO2 emissions from space and related applications, Atmospheric Measurement Techniques, 3, 781–811, doi:10.5194/amt-3-781-2010, URL http: //www.atmos-meas-tech.net/3/781/2010/, 2010.

/CECRv3, 2017/ M. Reuter, O. Schneising, M. Buchwitz, J. Heymann: Comprehensive Error Characterisation Report: BESD full-physics retrieval algorithm for XCO2 for the Essential Climate Variable (ECV) Greenhouse Gases (GHG), ESA GHG-CCI project, 2017

/Kiel et al., 2019/ Kiel, M., O'Dell, C. W., Fisher, B., Eldering, A., Nassar, R., MacDonald, C. G., and Wennberg, P. O.: How bias correction goes wrong: measurement of XCO2 affected by erroneous surface pressure estimates, Atmos. Meas. Tech., 12, 2241–2259, https://doi.org/10.5194/amt-12-2241-2019, 2019

/O'Dell et al., 2018/ O'Dell, C. W., Eldering, A., Wennberg, P. O., Crisp, D., Gunson, M. R., Fisher, B., Frankenberg, C., Kiel, M., Lindqvist, H., Mandrake, L., Merrelli, A., Natraj, V., Nelson, R. R., Osterman, G. B., Payne, V. H., Taylor, T. E., Wunch, D., Drouin, B. J., Oyafuso, F., Chang, A., McDuffie, J., Smyth, M., Baker, D. F., Basu, S., Chevallier, F., Crowell, S. M. R., Feng, L., Palmer, P. I., Dubey, M., García, O. E., Griffith, D. W. T., Hase, F., Iraci, L. T., Kivi, R., Morino, I., Notholt, J., Ohyama, H., Petri, C., Roehl, C. M., Sha, M. K., Strong, K., Sussmann, R., Te, Y., Uchino, O., and Velazco, V. A.: Improved retrievals of carbon dioxide from Orbiting Carbon Observatory-2 with the version 8 ACOS algorithm, Atmos. Meas. Tech., 11, 6539–6576, https://doi.org/10.5194/amt-11-6539-2018, 2018

/PUGv4, 2022/ M.Reuter, M.Hilker, and M.Buchwitz: Product User Guide Version 4 (PUGv4) for the FOCAL XCO2 OCO-2 Data Product CO2_OC2_FOCA (v10) for the Essential Climate Variable (ECV) Greenhouse Gases (GHG) ESA Climate Change Initiative “Plus” (CCI+), 2022

/PVIRv5, 2017/ Buchwitz et al., Product Validation and Intercomparison Report (PVIR), version 5, ESA GHG-CCI project, 2017

/Reuter et al., 2017a/ Reuter, M., M. Buchwitz, O. Schneising, S. Noël, V. Rozanov, H. Bovensmann, J. P. Burrows, A Fast Atmospheric Trace Gas Retrieval for Hyperspectral Instruments Approximating Multiple Scattering - Part 1: Radiative Transfer and a Potential OCO-2 XCO_2 Retrieval Setup, Remote Sens., 9, 1159, doi:10.3390/rs9111159, 2017.

/Reuter et al., 2017b/ Reuter, M., M. Buchwitz, O. Schneising, S. Noël, H. Bovensmann, J. P. Burrows, A Fast Atmospheric Trace Gas Retrieval for Hyperspectral Instruments Approximating Multiple Scattering - Part 2: Application to XCO_2 Retrievals from OCO-2, Remote Sens., 9, 1102, doi:10.3390/rs9111102, 2017.

/Reuter et al., 2020/ M. Reuter, M. Buchwitz, O. Schneising, S. Noël, H. Bovensmann, J.P. Burrows, H. Boesch, A. Di Noia, J. Anand, R.J. Parker, P. Somkuti, L. Wu, O.P. Hasekamp, I. Aben, A. Kuze, H. Suto, K. Shiomi, Y. Yoshida, I. Morino, D. Crisp, C.W. O'Dell, J. Notholt, C. Petri, T. Warneke, V.A. Velazco, N.M. Deutscher, D.W.T. Griffith, R. Kivi, D.F. Pollard, F. Hase, R. Sussmann, Y.V. Té, K. Strong, S. Roche, M.K. Sha, M. De Mazière, D.G. Feist, L.T. Iraci, C.M. Roehl, C. Retscher, and D. Schepers: Ensemble-based satellite-derived carbon dioxide and methane column-averaged dry-air mole fraction data sets (2003-2018) for carbon and climate applications, Atmos. Meas. Tech., https://www.atmos-meas-tech.net/13/789/2020, 2020

/Rozanov et al., 2014/ Rozanov, V., Rozanov, A., Kokhanovsky, A., and Burrows, J.: Radiative transfer through terrestrial atmosphere and ocean: Software package SCIATRAN, Journal of Quantitative Spectroscopy and Radiative Transfer, 133, 13 – 71, doi:https://doi.org/10.1016/j.jqsrt.2013.07.004,
URL http://www. sciencedirect.com/science/article/pii/S0022407313002872, 2014

/URDv3.0, 2020/ Chevallier et al., GHG-CCI: User Requirements Document for the GHG-CCI+ project of ESA’s Climate Change Initiative, pp. 42, version 3.0, 17

February, 2020. Link: https://www.iup.uni-bremen.de/carbon_ghg/docs/GHG-CCIplus/URD/URDv3.0_GHG-CCIp_Final.pdf

/Wunch et al., 2010/ Wunch, D., Toon, G. C., Wennberg, P. O., Wofsy, S. C., Stephens, B. B., Fischer, M. L., Uchino, O., Abshire, J. B., Bernath, P., Biraud, S. C., Blavier, J.-F. L., Boone, C., Bowman, K. P., Browell, E. V., Campos, T., Connor, B. J., Daube, B. C., Deutscher, N. M., Diao, M., Elkins, J. W., Gerbig, C., Gottlieb, E., Griffith, D. W. T., Hurst, D. F., Jiménez, R., Keppel-Aleks, G., Kort, E. A., Macatangay, R., Machida, T., Matsueda, H., Moore, F., Morino, I., Park, S., Robinson, J., Roehl, C. M., Sawa, Y., Sherlock, V., Sweeney, C., Tanaka, T., and Zondlo, M. A.: Calibration of the Total Carbon Column Observing Network using aircraft profile data, Atmospheric Measurement Techniques, 3, 1351–1362, doi:10.5194/amt-3-1351-2010, URL http://www.atmos-meas-tech.net/3/1351/2010/, 2010.

/Wunch et al., 2011/ Wunch, D., Toon, G. C., Blavier, J.-F. L., Washenfelder, R. A., Notholt, J., Connor, B. J., Griffith, D. W. T., Sherlock, V., and Wennberg, P. O.: The Total Carbon Column Observing Network (TCCON), Philosophical Transactions of the Royal Society of London, Series A: Mathematical, Physical and Engineering Sciences, 369, 2087–2112, doi:10.1098/rsta.2010.0240, 2011.

*** end of document