




Article

A Fast Atmospheric Trace Gas Retrieval for Hyperspectral Instruments Approximating Multiple Scattering—Part 2: Application to XCO₂ Retrievals from OCO-2

Maximilian Reuter ^{*} , Michael Buchwitz, Oliver Schneising, Stefan Noël, Heinrich Bovensmann  and John P. Burrows 

Institute of Environmental Physics, University of Bremen, P.O. Box 330440, 28334 Bremen, Germany; Michael.Buchwitz@iup.physik.uni-bremen.de (M.B.); schneising@iup.physik.uni-bremen.de (O.S.); Stefan.Noel@iup.physik.uni-bremen.de (S.N.); heinrich.bovensmann@iup.physik.uni-bremen.de (H.B.); burrows@iup.physik.uni-bremen.de (J.P.B.)

* Correspondence: mail@maxreuter.org

Received: 31 August 2017; Accepted: 24 October 2017; Published: 28 October 2017

Abstract: Satellite retrievals of the atmospheric dry-air column-average mole fraction of CO₂ (XCO₂) based on hyperspectral measurements in appropriate near (NIR) and short wave infrared (SWIR) O₂ and CO₂ absorption bands can help to answer important questions about the carbon cycle but the precision and accuracy requirements for XCO₂ data products are demanding. Multiple scattering of light at aerosols and clouds can be a significant error source for XCO₂ retrievals. Therefore, so called full physics retrieval algorithms were developed aiming to minimize scattering related errors by explicitly fitting scattering related properties such as cloud water/ice content, aerosol optical thickness, cloud height, etc. However, the computational costs for multiple scattering radiative transfer (RT) calculations can be immense. Processing all data of the Orbiting Carbon Observatory-2 (OCO-2) can require up to thousands of CPU cores and the next generation of CO₂ monitoring satellites will produce at least an order of magnitude more data. For this reason, the Fast atmospheric trace gas retrieval FOCAL has been developed reducing the computational costs by orders of magnitude by approximating multiple scattering effects with an analytic solution of the RT problem of an isotropic scattering layer. Here we confront FOCAL for the first time with measured OCO-2 data and protocol the steps undertaken to transform the input data (most importantly, the OCO-2 radiances) into a validated XCO₂ data product. This includes preprocessing, adaptation of the noise model, zero level offset correction, post-filtering, bias correction, comparison with the CAMS (Copernicus Atmosphere Monitoring Service) greenhouse gas flux inversion model, comparison with NASA's operational OCO-2 XCO₂ product, and validation with ground based Total Carbon Column Observing Network (TCCON) data. The systematic temporal and regional differences between FOCAL and the CAMS model have a standard deviation of 1.0 ppm. The standard deviation of the single sounding mismatches amounts to 1.1 ppm which agrees reasonably well with FOCAL's average reported uncertainty of 1.2 ppm. The large scale XCO₂ patterns of FOCAL and NASA's operational OCO-2 product are similar and the most prominent difference is that FOCAL has about three times less soundings due to the inherently poor throughput (11%) of the MODIS (moderate-resolution imaging spectroradiometer) based cloud screening used by FOCAL's preprocessor. The standard deviation of the difference between both products is 1.1 ppm. The validation of one year (2015) of FOCAL XCO₂ data with co-located ground based TCCON observations results in a standard deviations of the site biases of 0.67 ppm (0.78 ppm without bias correction) and an average scatter relative to TCCON of 1.34 ppm (1.60 ppm without bias correction).

Keywords: CO₂; Satellite; remote sensing; OCO-2; validation

1. Introduction

Satellite retrievals of the atmospheric dry-air column-average mole fraction of CO₂ (XCO₂) based on hyperspectral measurements in appropriate near (NIR) and short wave infrared (SWIR) O₂ and CO₂ absorption bands can help to answer pressing questions about the carbon cycle (e.g., [1]). However, the precision and even more the accuracy requirements for applications like surface flux inversion or emission monitoring are demanding (e.g., [2–4]). As an example, large scale biases of a few tenths of a ppm can already hamper an inversion with mass-conserving global inversion models [2,3].

The Scanning Imaging Absorption Spectrometer for Atmospheric Chartography (SCIAMACHY, [5,6]) became operational in 2002 and its radiance measurements allowed to start the time series of NIR/SWIR XCO₂ retrievals. With an overlap of about three years, the Greenhouse Gases Observing Satellite (GOSAT, [7]) allowed complementation and continuation of this time series in 2009.

The Orbiting Carbon Observatory-2 (OCO-2) was launched in 2014 also aiming at continuing and improving XCO₂ observations from space. As part of the A-train satellite constellation, OCO-2 flies in a sun-synchronous orbit crossing the equator at 13:36 local time. It measures one polarization direction of the solar backscattered radiance in three independent wavelength bands: the O₂-A band at around 760 nm (band1) with a spectral resolution of about 0.042 nm and a spectral sampling of about 0.015 nm, the weak CO₂ band at around 1610 nm (band2) with a spectral resolution of about 0.080 nm and a spectral sampling of about 0.031 nm, and the strong CO₂ band at around 2060 nm (band3) with a spectral resolution of about 0.103 nm and a spectral sampling of about 0.040 nm. OCO-2 is operated in a near-push-broom fashion and has eight footprints across track and an integration time of 0.333 s. The instrument's spatial resolution at ground is 1.29 km across track and 2.25 km along track. More information on the OCO-2 instrument can be obtained from the publications of Crisp et al. [8,9].

Several XCO₂ retrieval algorithms exist for the SCIAMACHY and GOSAT observations covering the whole mission periods (e.g., [10–18]). Because of the sparse ground based validation sites, the analyses of an ensemble of independently developed algorithms can give important insights in the quality of the retrievals, especially, remote from the validation sites [19] and strengthen the geophysical interpretation of the data [20]. Nevertheless, so far only the operational NASA algorithm exists for OCO-2 [21,22] covering the whole mission period.

One reason for this is that many of the existing algorithms require computationally expensive multiple scattering radiative transfer (RT) calculations and the data rate of OCO-2 is about two orders of magnitude larger than for GOSAT. In part 1 of this publication [23], the Fast atmospheric trace gas retrieval (FOCAL) has been introduced. It approximates multiple scattering effects with an analytic solution of the scalar RT problem of an isotropic scattering layer and a Lambertian surface. This enhances the computational efficiency by orders of magnitude making it possible to process the whole OCO-2 data stream on few CPU cores in real time and will allow also the analyses of future satellite missions which will provide at least an order of magnitude more data.

In part 1, Reuter et al. [23] tested various retrieval setups with simulated OCO-2 measurements and concluded that their 3-Scat setup is a promising candidate for further studies with measured OCO-2 data. The 3-Scat retrieval setup fits the OCO-2 measured radiance simultaneously in four fit windows: SIF (~758.26–759.24 nm), O₂ (~757.65–772.56 nm), wCO₂ (~1595.0–1620.6 nm), and sCO₂ (~2047.3–2080.9 nm). This is achieved by iteratively optimizing the state vector including the following geophysical parameters: five layered CO₂ and H₂O concentration profiles, the pressure (i.e., height), scattering optical thickness at 760 nm, and the Ångström exponent of a scattering layer, solar induced chlorophyll fluorescence (SIF), and polynomial coefficients describing the spectral albedo in each fit window.

In this publication, we confront the *3-Scat* retrieval setup with actually measured OCO-2 data and protocol the steps undertaken to transform the input data (most importantly, the OCO-2 radiances) into a validated XCO₂ data product. This includes preprocessing (Section 2), adaptation of the noise model (Section 3.1), zero level offset correction (Section 3.2), post-filtering (Section 4.1), bias correction (Section 4.2), comparison with CAMS (Copernicus Atmosphere Monitoring Service) greenhouse gas flux inversion model (Section 5.1), comparison with NASA's operational OCO-2 XCO₂ product (Section 5.2), and validation with ground based Total Carbon Column Observing Network (TCCON) data (Section 5.3). The complete retrieval chain including adaptations, pre-, and postprocessing will be referred to as FOCAL v06 in the following.

2. Preprocessing

During preprocessing, we collect all datasets that are needed to run the retrievals and pre-filter soundings with potentially degraded quality or potential cloud or aerosol contamination. Due to the demanding precision and accuracy requirements for XCO₂ retrievals (e.g., [2–4]) and the large amount of OCO-2 data, we prioritize quality over quantity in the course of pre-filtering.

The primary input data used for this publication are global OCO-2 L1b calibrated radiances (i.e., a single linear component of the polarization of the incoming light) version 7r [9,24] of the year 2015 in glint (GL), nadir (ND), target (TG), and transition (XS) mode which have been obtained from <https://daac.gsfc.nasa.gov>. Each of OCO-2's three bands consists of 1016 spectral pixels which we group into the four fit windows illustrated in Figure 1 showing a typical OCO-2 measurement fitted with FOCAL.

Each L1b orbit file includes information on spectral pixels with potentially reduced quality, e.g., due to radiometric problems. Based on the last nadir orbit in 2015 (oco2_L1bScND_07974a_151231_B7200r_160121043229.h5), we generated a dead or bad pixel mask which we use for the O₂ and both CO₂ fit windows. For the SIF fit window, we ignore the dead or bad pixel mask because it is located in a spectral region generally flagged as potentially bad.

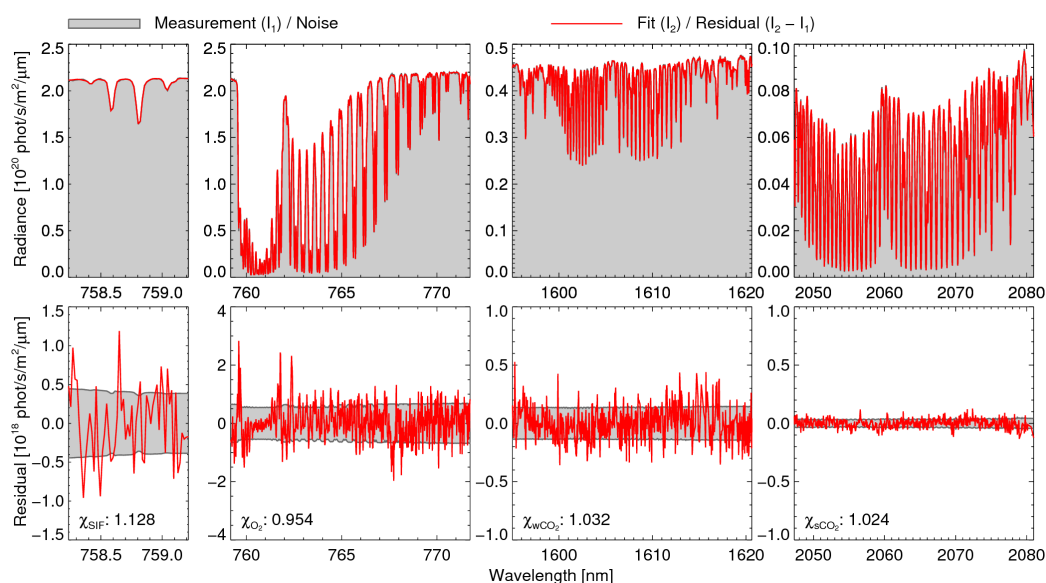


Figure 1. OCO-2 measurement of June 6, 2015, 12:01 UTC near Hamburg, Germany (sounding ID: 2015060512011938) fitted with FOCAL. (**Top**) Simulated and fitted radiance measurement in gray and red, respectively. (**Bottom**) Adapted measurement noise (see Section 3.1) and fit residual (fit minus measurement) in gray and red, respectively; χ_j is an estimate of the goodness of fit (relative to the noise) in fit window j and is computed as defined in part 1 [23].

We reject all soundings flagged to have potentially reduced quality (quality flag $\neq 0$) or failing a data integrity test (e.g., unreasonable sounding ID or time). We filter out potentially “tricky” scenes with solar or satellite zenith angles greater than 70° , latitudes beyond $\pm 70^\circ$, or extreme surface roughnesses (standard deviation of the surface elevation) greater than 1000 m. In Figure 2, this filter is referred to as LAT/SUZ/SAZ/ σ ALT.

We use the spike EOF analysis provided with the OCO-2 L1b data [24] and accept only soundings with less than 60 spectral pixels with potentially poor quality (referred to as bad colors) in the O_2 and no bad colors in the CO_2 bands. This primarily filters out soundings above South America and the South Atlantic because of contamination by cosmic rays within the South Atlantic Anomaly (SAA) caused by the shape of the inner Van Allen radiation belt (Figure 2).

Potentially aerosol contaminated scenes are filtered using OMI (Ozone Monitoring Instrument aboard Aura lagging OCO-2 by 30 min) L3 global daily gridded $1^\circ \times 1^\circ$ UV aerosol index data (OMAERUVd v003 obtained from <https://daac.gsfc.nasa.gov>) with a filtering threshold of one. As described by Stammes [25], the UV aerosol index is derived by comparing the measured reflectance ratio at two wavelengths (342.5 nm and 388.0 nm) to the calculated reflectance ratio using a Rayleigh atmosphere with an assumed surface albedo. The UV index is relatively insensitive to scattering aerosol layers or clouds, because it is mainly determined by the reduction of Rayleigh multiple scattering due to aerosol absorption. As illustrated in Figure 2, this filter most prominently impacts regions contaminated with desert dust aerosols.

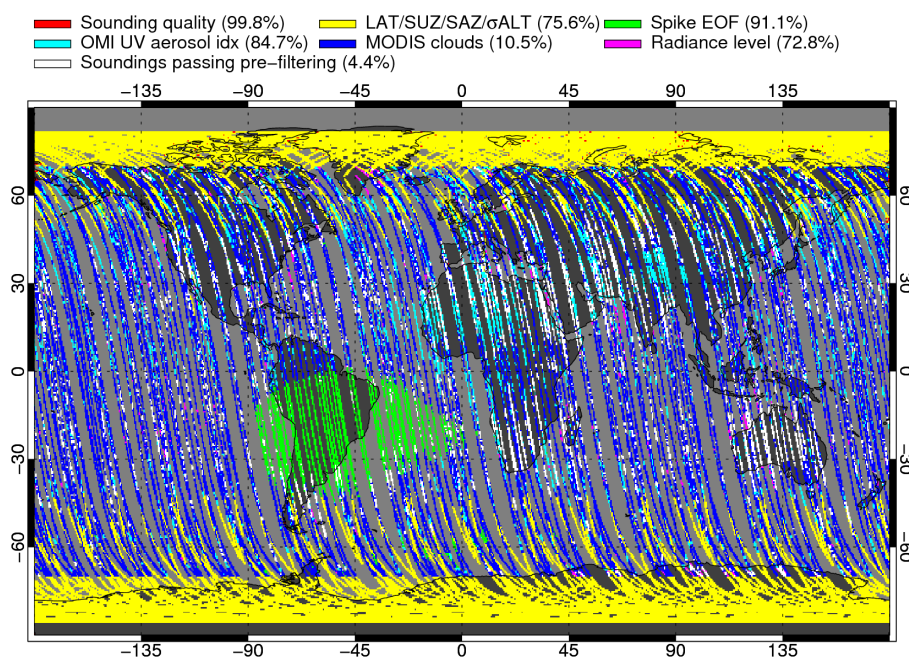


Figure 2. Pre-filtering statistics of the 24 days data subset used for the noise model analysis (Section 3.1). The filters are applied in the order: Sounding quality, LAT/SUZ/SAZ/ σ ALT, Spike EOF, OMI UV aerosol idx, MODIS clouds, and Radiance level (see main text for a description). The colors represent filter activity and soundings passing all filters are shown in white. Numbers in brackets represent filter throughputs.

Potentially cloud contaminated scenes are filtered using MODIS Aqua (moderate-resolution imaging spectroradiometer aboard Aqua) L2 cloud mask data [26] with about $1 \text{ km} \times 1 \text{ km}$ resolution (collection 6, MYD35, obtained from <https://ladsweb.modaps.eosdis.nasa.gov>). All MODIS ground pixels which are not flagged as clear or probably clear are considered as potentially cloudy. Aqua is lagging OCO-2 by 15 min and in order to account for the parallax effect and potential cloud movements, we use only OCO-2 data with at least 10 km distance to the nearest MODIS cloud. Even though 10 km

is not overly conservative, this filter has a throughput of only about 11% and dominates the total pre-filtering throughput of about 4% (Figure 2).

Additionally, we filter out very dark or bright scenes, i.e., extreme detector fillings. Specifically, we ensure that the continuum radiance in each band is between 5% and 95% of the maximum band radiance as specified in the OCO-2 L1b ATBD (Algorithm Theoretical Basis Document, [24]).

Meteorological profiles come from ECMWF operational analysis data (<http://apps.ecmwf.int>) and have a resolution of six hours, $0.75^\circ \times 0.75^\circ$, and 137 height layers. As part of the preprocessor, these profiles are corrected for the actual surface height of the OCO-2 soundings and split into 20 layers containing the same number of dry-air particles.

3. Retrieval Adaptations

The theoretical bases of FOCAL's retrieval method is described in detail in part 1 of this publication [23]. However, in order to analyze actually measured data instead of simulations, we made the following adaptations.

3.1. Noise Model

The measurement error covariance matrix [23] has to account not only for the measurement noise but for the total error including also the forward model error. The measurement noise of the instrument is well known from laboratory measurements and in-flight estimates. In theoretical studies (as in part 1, [23]), it is often assumed for convenience, that the measurement noise dominates and that other error components can be neglected, i.e., the noise model is approximated by the measurement noise.

Especially when analyzing measured data, unknown inaccuracies of the forward model can violate this assumption and lead to larger fit residuals and unrealistic results (and error estimates) because the optimal estimation retrieval puts too much trust in the measurement. This may happen, e.g., due to imperfect knowledge of the instrumental line shape function (ILS), unconsidered spectroscopic effects such as Raman scattering, inaccuracies of the spectroscopic data bases, approximations of the radiative transfer model, or imperfect meteorology.

Ideally, one would reduce the fit residuals to the instrument's noise level by improving the forward model, but this is often not possible. A potential solution is to fit parts of the residuum by empirical orthogonal functions (EOF) computed from a representative set of measurements [21]. Another approach is to adjust the noise model so that it accounts for measurement noise plus forward model error (e.g., [16,17,27]) and a variant of this approach is also used by us.

Most forward model errors can be interpreted to result from inaccuracies of the computed (effective) atmospheric transmittance. However, the largest scene-to-scene variability of the simulated radiance is due to changes of, e.g., albedo and solar zenith angle. Therefore, it is reasonable to assume forward model errors to be approximately proportional to the continuum signal I_{cont} which we obtain from up to nine spectral pixels at the fit windows' lower wavelength length ends.

We model the root mean square residual to continuum signal ratio RSR by

$$RSR = \sqrt{NSR^2 + \delta F^2}, \quad (1)$$

where NSR represents the root mean square of the spectral 1σ radiance noise (as reported in the OCO-2 L1b data) to continuum signal ratio and δF the relative forward model error.

In order to estimate the free parameter δF , we analyzed a representative set of pre-filtered soundings (Figure 3) with a modified FOCAL setup for which we (quadratically) added 2% of the continuum radiance to the measurement noise. This overestimation of the expected total error effects that the retrieval usually converges towards values being not very far away from the a priori, i.e., values being more or less realistic. Additionally, we switched off the SIF retrieval (which is basically identical to a zero level offset retrieval in the SIF fit window) and switched on the retrieval of zero

level offsets in all four fit windows.

If the instrument noise would dominate the total error, RSR and NSR would (statistically) lie on a 1:1 line. After the removal of outliers (Figure 4, gray dots), this is basically the case for the SIF fit window with forward model errors estimated to be about 0.5‰ of the (continuum) signal (Figure 4). The forward model error within the other fit windows is estimated to be between 2.5‰ and 3.2‰ (Figure 4). This means, the total error in dark scenes (large NSR) is still dominated by the instrumental noise but in bright scenes (small NSR), the forward model error dominates.

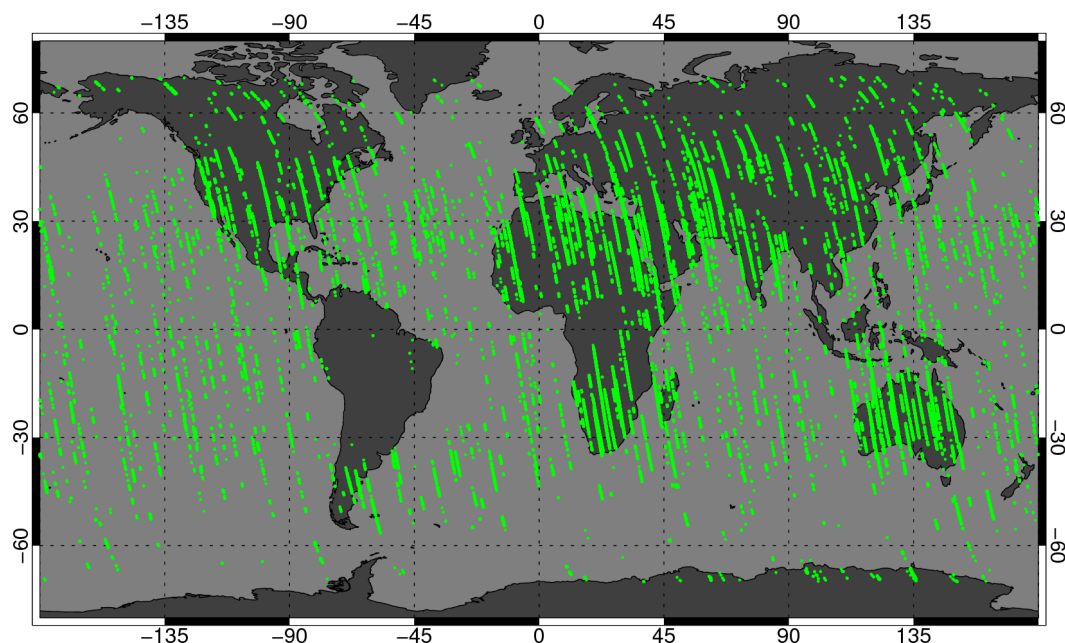


Figure 3. Sampling of all pre-filtered soundings analyzed in order to determine the noise model. The data set consists of 10% of all pre-filtered OCO-2 soundings (randomly selected) of 24 days in 2015 (13.01., 15.01., 14.02., 16.02., 10.03., 20.03., 03.04., 19.04., 08.05., 23.05., 08.06., 24.06., 15.07., 16.07., 15.08., 16.08., 15.09., 16.09., 15.10., 16.10., 15.11., 17.11., 12.12., 18.12.). This results in a manageable but still representative data set with respect to nadir/glint observation geometry, season, and spatial distribution.

Outliers are removed as follows: The data set is grouped in 35 NSR bins. Only bins with more than 500 samples are further considered. Within each bin, RSR should follow a χ^2 -distribution with as many degrees of freedom as spectral pixels of the fit window. The number of spectral pixels is always large enough to approximate the χ^2 -distribution with a Gaussian distribution. Outliers represent poor fits, e.g., due to complicated atmospheric conditions which cannot be well described by the forward model. As they usually enhance the RSR, we have to approach the expectation value of RSR from the lowermost values. The 2.28th and 15.9th percentile (Figure 4, red and orange points) of the Gaussian distribution are two and one standard deviations smaller than the expectation value. We used this to estimate the expectation value (Figure 4, green points) from which we determined the free fit parameter δF of Equation (1) (numerical values are shown in Figure 4). Note that adding 4% instead of 2% of the continuum radiance to the measurement noise gave similar results (not shown here).

Soundings with a RSR being more than two standard deviations larger than expected from Equation (1) are considered outliers. For this purpose, we fitted the second order polynomial

$$2\sigma = a_0 + a_1 NSR + a_2 NSR^2 \quad (2)$$

and use it as threshold for the maximal allowed deviation from the RSR model (Figure 4, gray lines).

We define the noise model which modifies the reported OCO-2 L1b radiance noise N analog to Equation (1):

$$N' = \sqrt{N^2 + I_{\text{cont}}^2 \delta F^2}. \quad (3)$$

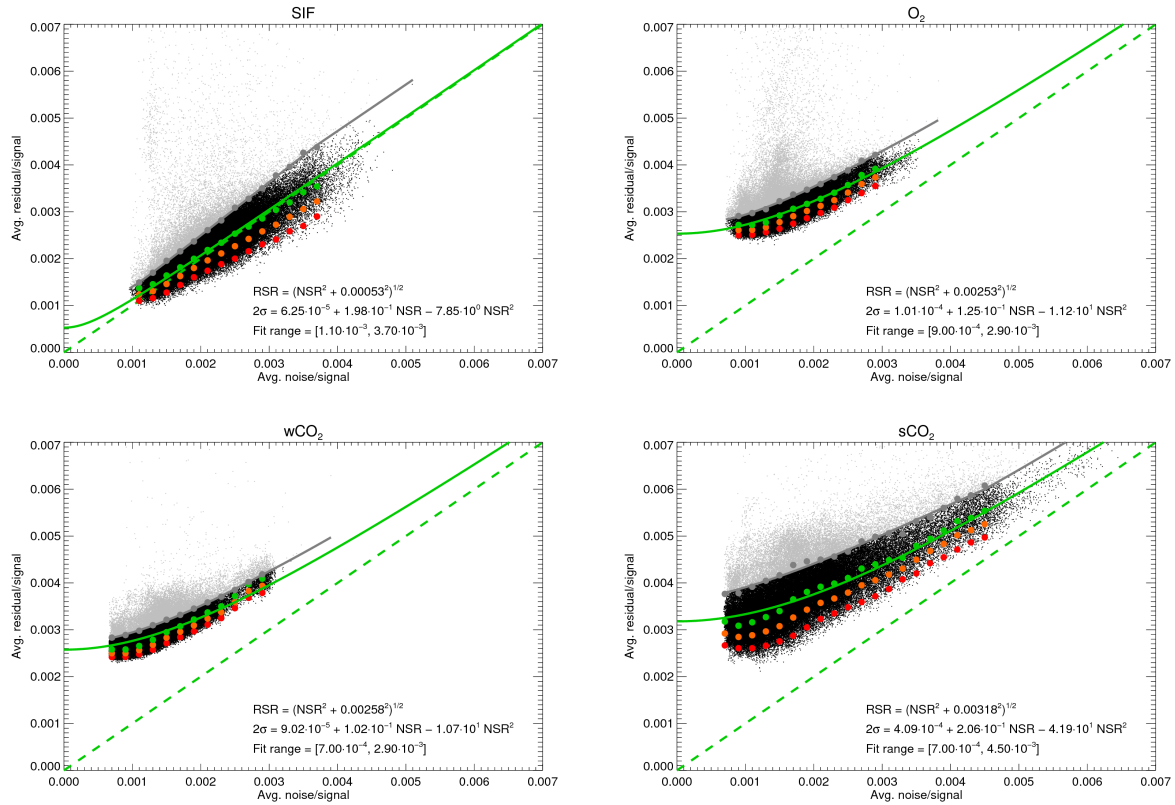


Figure 4. Root mean square noise to signal ratio NSR versus root mean square residual to signal ratio RSR for all four fit windows. **red points:** 2.28th percentile within bins with more than 500 samples (35 bins in total). **orange points:** 15.9th percentile. **green points:** expectation value estimated from the 2.28th and 15.9th percentile. **solid green line:** RSR as computed from the RSR model (Equation (1)). **gray points:** RSR model plus 2σ estimated from the 2.28th and 15.9th percentile. **gray line:** outlier threshold. **gray dots:** potential outliers. **dashed green line:** one-to-one line.

3.2. Zero Level Offset Correction

We define as zero level offset (ZLO) an additive fit window-wide radiance offset. An apparent or effective ZLO can have various reasons such as residual calibration errors or unconsidered spectroscopic effects. Many of these effects can be expected to result in ZLOs being approximately proportional to the fit window's continuum radiance. In order to study potential ZLOs, we used the same modified FOCAL setup as in the last section but with the just defined noise model. The simultaneous retrieval of ZLOs reduce the uncertainty reduction for XCO₂ and renders the SIF retrieval impossible. Therefore, we aimed at a ZLO correction rather than a ZLO retrieval per sounding. We analyzed the same 24 days of OCO-2 data as in the last section but filtered for potential contamination with chlorophyll fluorescence because in the SIF fit window it is not possible to disentangle ZLO and SIF (Figure 5). For this purpose, we used monthly L3 MODIS Aqua chlorophyll-a data (obtained from https://modis.gsfc.nasa.gov/data/dataproduct/chlor_a.php, [28]) over ocean and normalized difference vegetation index (NDVI) data over land (obtained from <https://modis.gsfc.nasa.gov/data/dataproduct/mod13.php>).

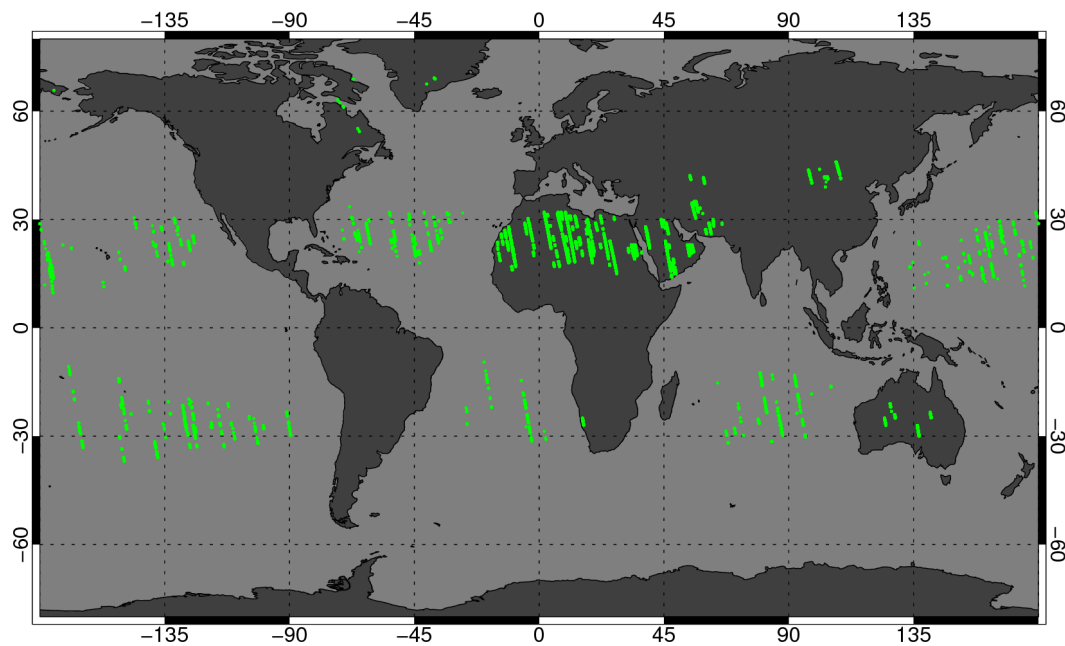


Figure 5. Sampling of all pre-filtered soundings analyzed in order to determine the ZLO correction. The data set consists of all pre-filtered OCO-2 soundings of 24 days in 2015 (13.01., 15.01., 14.02., 16.02., 10.03., 20.03., 03.04., 19.04., 08.05., 23.05., 08.06., 24.06., 15.07., 16.07., 15.08., 16.08., 15.09., 16.09., 15.10., 16.10., 15.11., 17.11., 12.12., 18.12.) additionally filtered for potential contamination with chlorophyll fluorescence (see main text).

Figure 6 shows that we find a reasonably linear relationship (with correlations around 0.9) between the retrieved ZLO and the continuum radiance within the SIF and both CO₂ fit windows hinting at ZLOs in the range of 0.8–1.8% of the continuum radiance. In the following, we use the fitted linear relationship as ZLO correction for these three fit windows. In the O₂ fit window, the correlation between ZLO and continuum radiance is poor and the linear fit suggests a small negative slope. Therefore, we decided to not apply a ZLO correction for this fit window.

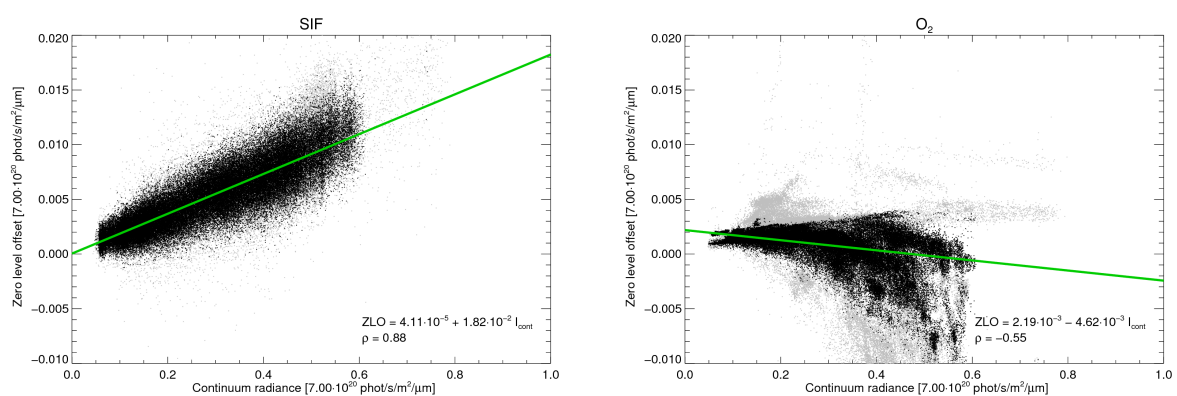


Figure 6. Cont.

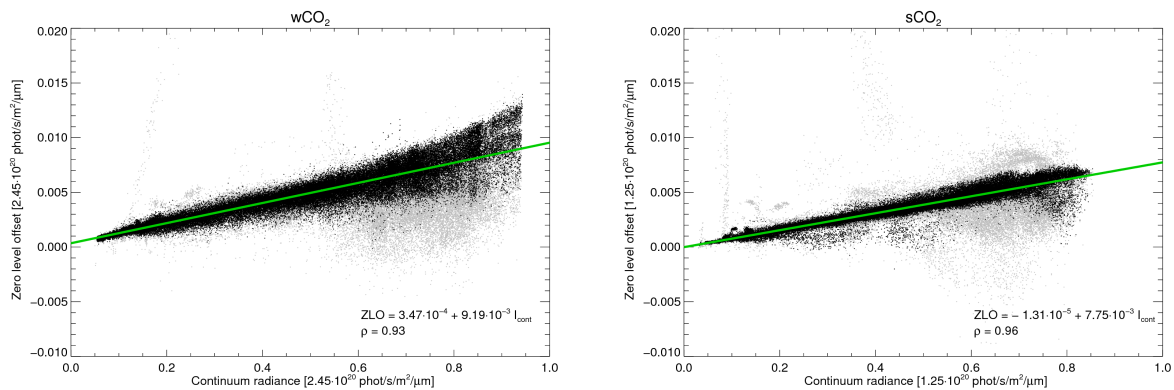


Figure 6. Retrieved zero level offset (ZLO) versus continuum radiance (I_{cont}) for all four fit windows. **gray dots:** potential outliers, (i.e., no convergence, $\chi^2 > 2$, or RSR exceeding threshold (see Figure 4)). **green line:** linear fit.

4. Postprocessing

4.1. Filtering

First of all, we check for convergence, i.e., the state vector increment has to be small compared to the a posteriori uncertainty, the maximum number of iterations must not exceed 15, and χ^2 must not exceed 2 (for more details, see part 1 [23]). Convergence is achieved in about 74% of all pre-filtered OCO-2 soundings. Many non-converging soundings can be found near the SAA, the Saharan desert, and the Arabian peninsula (Figure 7).

In the next step, we check for each fit window if the RSR is smaller than the threshold for potential outliers defined in Section 3.1. The throughput of this filter, which is most active above the tropical oceans (Figure 7), is about 68%.

Additionally, we filter for potential outliers by parameters that have a unexpectedly large influence on the retrieved local XCO₂ variability. For the example data shown in Figure 7, this filter is most active in high latitudes and has a throughput of about 84%.

This filter bases on the idea that XCO₂ outliers increase the local retrieved XCO₂ variability and are likely correlated with extreme values of some of the candidate parameters: XCO₂ uncertainty σ_{XCO_2} , lowermost layer of the CO₂ column averaging kernel, XH₂O, XH₂O uncertainty $\sigma_{\text{XH}_2\text{O}}$, XH₂O difference to the a priori, continuum radiance in the O₂ ($I_{\text{cont}}^{\text{O}_2}$), wCO₂ ($I_{\text{cont}}^{\text{wCO}_2}$), and sCO₂ ($I_{\text{cont}}^{\text{sCO}_2}$) fit window, gradient between first and second CO₂ layer ∇CO_2 , albedo difference between the O₂ and sCO₂ fit window, and all non CO₂ and H₂O state vector elements [23].

For a representative two months data set (April and August 2015), we estimated the local retrieved XCO₂ variability $\text{VAR}(\Delta \text{XCO}_2)$ as follows: For each sounding, we computed the difference ΔXCO_2 between XCO₂ and its $5^\circ \times 5^\circ$ daily median and subsequently, we computed the variance of all ΔXCO_2 values falling in grid boxes with more than 100 samples. Now we searched for an upper or lower threshold for that candidate parameter which reduces $\text{VAR}(\Delta \text{XCO}_2)$ most when removing 1‰ of all data points. We repeated this until 15% of all data points were removed. In order to reduce the complexity of the postprocessing filter procedure, we now identified the 10 most promising parameters separately for land and ocean and repeated the whole exercise to find filter thresholds for these 10 parameters.

Above land (Figure 8, left), the potential outliers filter reduces the variance of ΔXCO_2 from 2.05 ppm² to 1.68 ppm². The Ångström exponent \AA is the dominant parameter, contributing 38% to the variance reduction. All parameter thresholds found for the potential outliers filter above land are listed in Table 1 (left).

Above sea (Figure 8, right), this filter reduces the variance of ΔXCO_2 from 1.22 ppm² to 1.09 ppm². In glint geometry, scattering is less important and the dominant parameter is the wavelength squeeze in the wCO₂ fit window $\lambda_{\text{sq}}^{\text{wCO}_2}$, contributing 32% to the variance reduction. All parameter thresholds found for the potential outliers filter above land are listed in Table 1 (right).

The combined throughput of all three post-filters (convergence, residual, and potential outliers) is about 42%.

Table 1. Thresholds and prorated variance reduction of the 10 parameters of the potential outliers filter for soundings above land (**top**) and sea (**bottom**). In total, the variance of ΔXCO_2 is reduced from 2.05 ppm² to 1.68 ppm² above land and reduced from 1.22 ppm² to 1.09 ppm² above sea. See part 1 [23] and the main text for a description of the individual parameters.

	Parameter	Lower Threshold	Upper Threshold	Variance Reduction [%]
Land	\AA	1.6669	-	38
	σXCO_2 [ppm]	-	1.2963	17
	$\text{ILS}_{\text{sq}}^{\text{wCO}_2}$	-	1.0022	16
	p_s [p ₀]	-1.6435×10^{-1}	2.2603×10^{-1}	11
	∇CO_2 [ppm]	5.2509	5.9995	11
	$\lambda_{\text{sq}}^{\text{wCO}_2}$ [nm]	-5.2186×10^{-4}	3.9367×10^{-5}	2
	$\text{ILS}_{\text{sq}}^{\text{wCO}_2}$	-	1.0041	1
	$\lambda_{\text{sq}}^{\text{O}_2}$ [nm]	-	-2.5907×10^{-5}	2
	$\lambda_{\text{sh}}^{\text{O}_2}$ [nm]	-6.6146×10^{-4}	9.2043×10^{-4}	1
	$\sigma\text{XH}_2\text{O}$ [ppm]	-	15.705	1
Sea	$\lambda_{\text{sq}}^{\text{wCO}_2}$ [nm]	-3.1372×10^{-4}	-8.2869×10^{-5}	32
	$\alpha\text{P}_{\text{SIF}}$	-2.3184×10^{-3}	3.4846×10^{-3}	20
	$\alpha\text{P}_{\text{O}_2}$	-	1.7900×10^{-4}	19
	$\lambda_{\text{sh}}^{\text{wCO}_2}$ [nm]	-	2.1023×10^{-3}	9
	$\text{ILS}_{\text{sq}}^{\text{O}_2}$	-	1.0175	4
	$\alpha\text{P}_{\text{wCO}_2}^2$	-	-2.1247×10^{-4}	5
	$\alpha\text{P}_{\text{wCO}_2}^1$	3.2736×10^{-4}	-	4
	$\text{I}^{\text{wCO}_2}_{\text{cont}}$ [1.25×10^{20} Ph/s/m ² /μm]	5.6468×10^{-2}	-	4
	$\lambda_{\text{sh}}^{\text{O}_2}$ [nm]	-3.4860×10^{-4}	-	2
	\AA	1.9014	-	2

4.2. Bias Correction

The basic assumption of the bias correction scheme is that on average XCO₂ has little variations on small scales so that correlations to more variable parameters can be used to quantify biases. As a consequence, the bias correction does not require any ground truth data except for the quantification of a globally constant offset.

The final bias model consists of four components: a footprint bias B_f , a land/sea bias B_{ls} , a linear bias model B_{lin} , and a globally constant bias B_g :

$$B = B_f + B_{\text{ls}} + B_{\text{lin}} + B_g. \quad (4)$$

These four components are successively derived from analyses of the same two months data set (April and August 2015) used to determine the thresholds of the potential outliers filter (Section 4.1).

The swath of OCO-2 consists of eight neighboring footprints across track. In order to determine the mean footprint anomaly, we used only soundings belonging to complete sets of eight neighboring soundings which all passed the post-filtering and which were entirely over land or sea. Figure 9 (right) shows the sampling of the roughly 180,000 soundings where this is the case. For each of these sets of

eight soundings, we compute the footprint anomaly and subsequently the average footprint anomaly of all sets which we then use as bias function. Figure 9 (left) shows the footprint bias pattern at the example of all soundings passing the post-filtering in August 2015 and the corresponding bias function depending on the footprint f is:

$$B_f(f) = \begin{cases} -0.974 \text{ ppm} & \text{if } f = 1, \\ -0.336 \text{ ppm} & \text{if } f = 2, \\ -0.234 \text{ ppm} & \text{if } f = 3, \\ -0.315 \text{ ppm} & \text{if } f = 4, \\ -0.856 \text{ ppm} & \text{if } f = 5, \\ 1.013 \text{ ppm} & \text{if } f = 6, \\ 0.484 \text{ ppm} & \text{if } f = 7, \\ 1.219 \text{ ppm} & \text{if } f = 8. \end{cases} \quad (5)$$

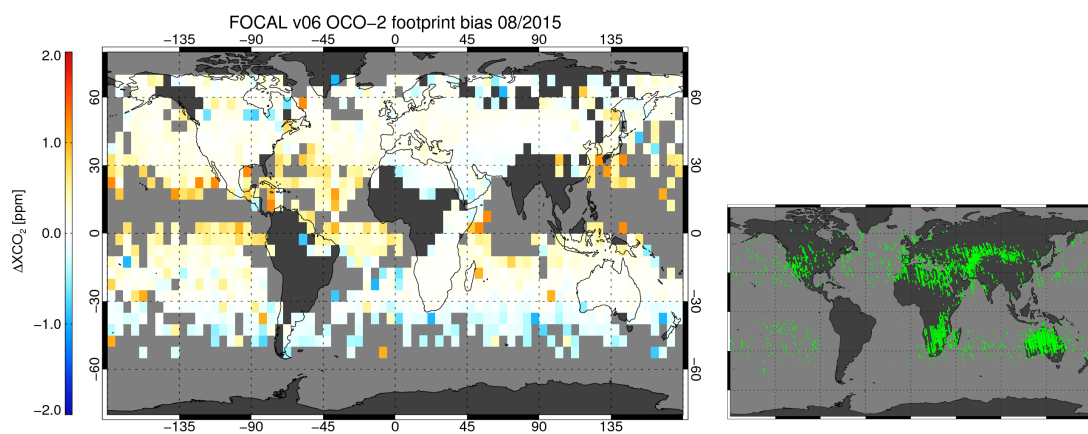


Figure 9. FOCAL v06 OCO-2 footprint bias pattern (Equation (5)) at the example of August 2015 (**left**) and sampling of soundings used to determine the footprint bias (**right**).

In order to determine the land/sea bias, we corrected all post-filtered results for the footprint bias and analyzed all coastline overpasses with a maximum duration of 120 s (≈ 800 km along track), at least 100 soundings, and a land fraction between 30% and 70%. For each of these coastline overpass (162 with about 102,000 soundings, Figure 10, right) we computed the land/sea anomaly and hence the average land/sea anomaly (± 0.8986 ppm). Figure 10 (left) shows the land/sea bias pattern corresponding to the bias function

$$B_{ls}(l) = 0.8986 \text{ ppm} (2l - 1) \quad (6)$$

with l being the land/sea fraction.

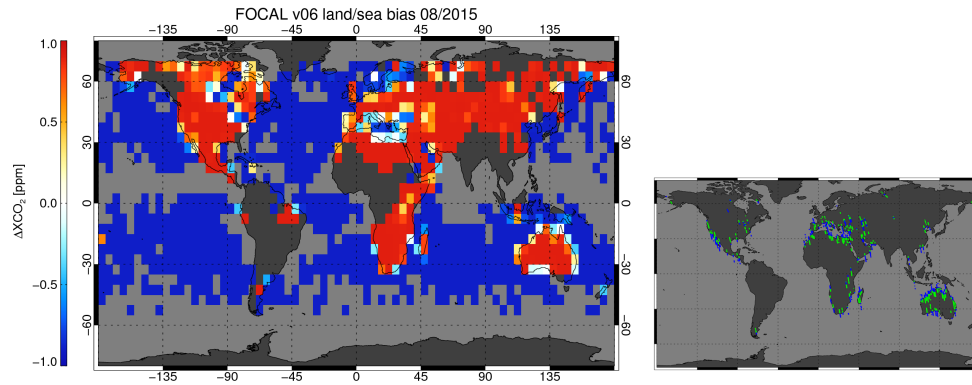


Figure 10. FOCAL v06 land/sea bias pattern (Equation (6)) at the example of August 2015 (left) and sampling of soundings used to determine the land/sea bias (right).

In addition to the footprint and land/sea biases, we found a small correlation ($\rho = 0.16$) between small scale anomalies of XCO_2 and the retrieved ILS squeeze in the wCO_2 fit window $ILS_{sq}^{wCO_2}$. The small scale anomalies have been computed from 60 s chunks (≈ 400 km) of post-filtered and footprint and land/sea bias corrected OCO-2 orbit data with more than 10 soundings. In total about 5000 chunks with one million soundings (Figure 11, right) have been analyzed to compute the linear bias model (Figure 11, left):

$$B_{lin}(ILS_{sq}^{wCO_2}) = 107.936 \text{ ppm} \cdot ILS_{sq}^{wCO_2} - 107.862 \text{ ppm} . \quad (7)$$

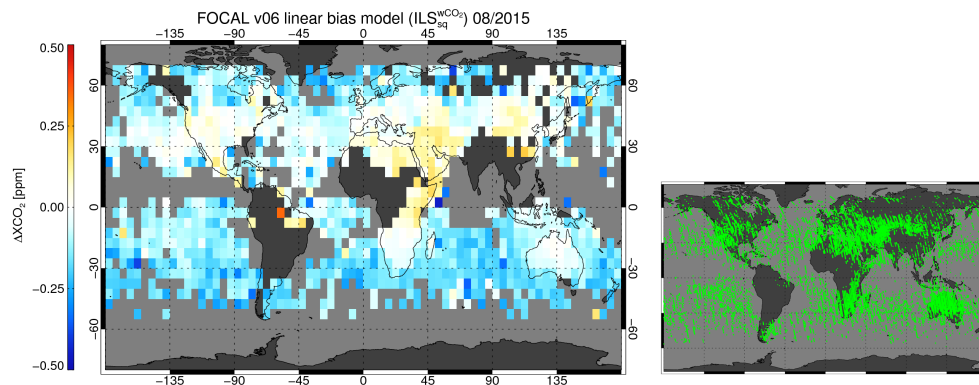


Figure 11. FOCAL v06 bias pattern of the linear bias model (Equation (7)) at the example of August 2015 (left) and sampling of soundings used to determine the linear bias model (right).

Finally, we correct for a global offset in respect to the optimized CO_2 concentration fields of the CAMS (Copernicus Atmosphere Monitoring Service) greenhouse gas flux inversion model [29] obtained from <http://apps.ecmwf.int>. CAMS is the CO_2 atmospheric inversion product of the European Union programme Copernicus that develops information services based on satellite Earth observation and other data (<http://www.copernicus.eu/>). The product is released twice per year and each time covers the full period from 1979 until the year before the release. It results from an analysis of CO_2 surface air sample measurements over the corresponding period and consists of optimized CO_2 surface fluxes over the globe and of associated 3D CO_2 concentrations. Version 15r4 used here analyzed 37 years of surface measurements (1979–2015). Its spatial resolution is of 3.75° in longitude and 1.875° in latitude, with 39 hybrid layers in the vertical. A full description of v15r4 is given by Chevallier [30], showing also, among other validation results, that its root-mean-square fit to TCCON

measurements is usually close to 1 ppm. The global offset of FOCAL v06 relative to the CAMS model amounts:

$$B_g = -1.673 \text{ ppm} . \quad (8)$$

Figure 12 (left) shows an example of the total bias pattern consisting of all four components (footprint bias, land/sea bias, linear bias model, global bias). As can be seen, the large scale pattern of the total bias is dominated by the land/sea bias followed by the footprint bias and the linear bias model plays only a minor role. For comparison, Figure 12 (right) shows the total bias pattern of NASA's operational OCO-2 L2 product v7.3.05b [21,22] obtained from <https://daac.gsfc.nasa.gov> and in the following referred to as NASA v7.3.05b. The overall variability is similar (0.82 ppm and 0.71 ppm for FOCAL v06 and NASA v7.3.05b, respectively) and the NASA product also has a distinct land/sea bias but with opposite sign, i.e., with largest values over sea (note the reversed color bar in Figure 12, right).

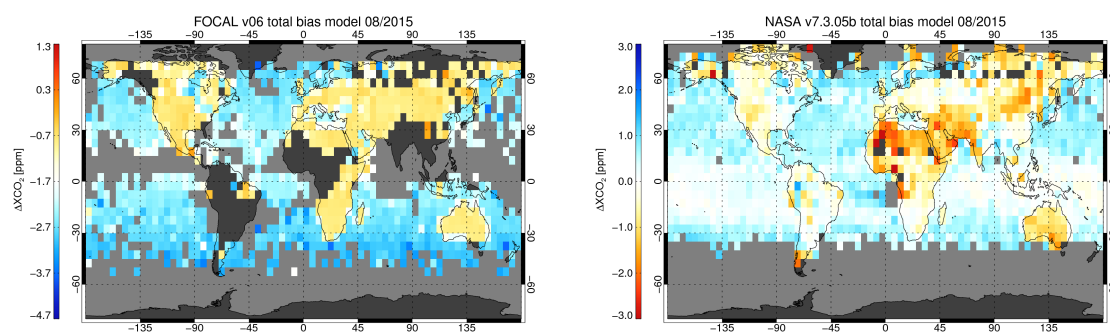


Figure 12. FOCAL v06 total bias pattern (Equation (4)) at the example of August 2015 (left) and NASA OCO-2 v7.3.05b total bias pattern (right).

5. Comparison and Validation

5.1. Model Comparison

In this section, we compare two months (April and August 2015) of post-filtered and bias corrected FOCAL v06 XCO₂ results with corresponding values of the CAMS v15r4 model accounting for FOCAL's column averaging kernels (e.g., [31]). Figure 13 shows 5° × 5° monthly gridded values for both months, FOCAL, and CAMS. The main 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 in April but also above the Sahara in August. However, it shall be noted that often only few data points are in the corresponding grid boxes.

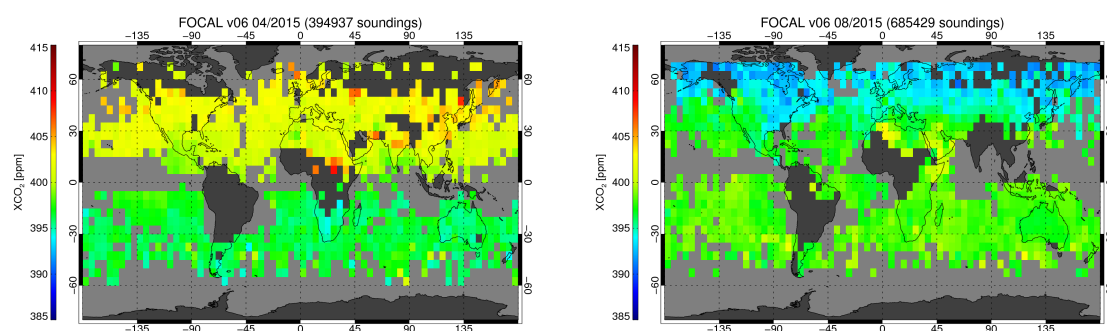


Figure 13. Cont.

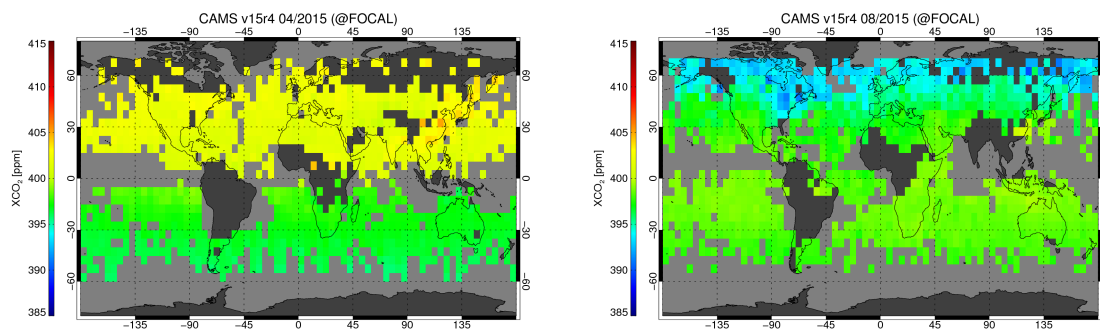


Figure 13. Monthly mean XCO_2 at $5^\circ \times 5^\circ$. (Top) FOCAL v06. (Bottom) CAMS v15r4 sampled as FOCAL. (Left) April 2015. (Right) August 2015.

In grid boxes with more than 100 soundings, the standard error of the mean becomes negligible (≈ 0.1 ppm). Therefore, the difference between FOCAL and CAMS in such grid boxes can be interpreted as systematic temporal and regional mismatch or bias. The standard deviation of this systematic mismatch (including also representation errors) amounts to 1.0 ppm. The standard deviation of the single sounding mismatch after subtracting the systematic mismatch amounts to 1.1 ppm which agrees reasonably well with the average reported uncertainty of 1.2 ppm.

5.2. Comparison with NASA's Operational OCO-2 L2 Product

In this section, we compare the same two months (April and August 2015) of post-filtered and bias corrected FOCAL v06 XCO_2 results with NASA's operational OCO-2 L2 product. Comparing Figure 14 with Figure 13 (top) shows similar large scale temporal and spatial patterns and also the relative enhancement in the anthropogenic source regions of East Asia in April are similar. The most obvious difference is that the NASA product has about three times more soundings. The primary reason for this is the inherently poor throughput (11%) of the MODIS based cloud screening of the preprocessor (see discussion in Section 2).

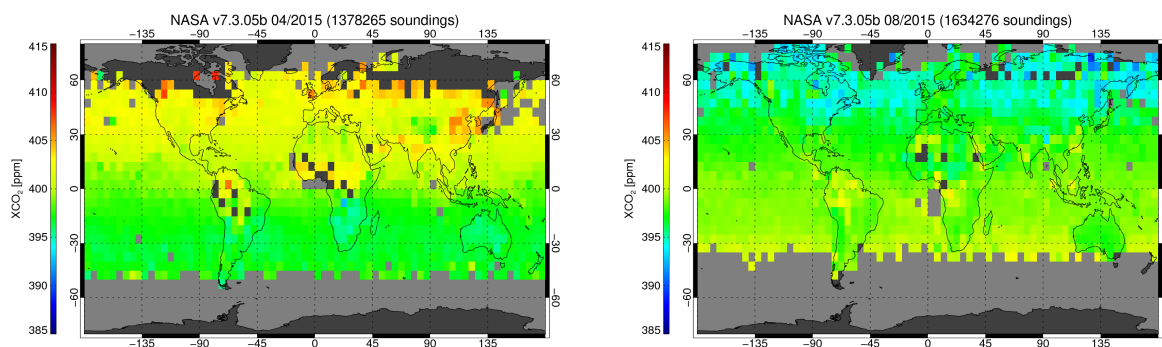


Figure 14. NASA v7.3.05b monthly mean XCO_2 at $5^\circ \times 5^\circ$. Left: April 2015. Right: August 2015.

Analyzing only the same soundings in both data sets and considering the column averaging kernels, the NASA product has on average 0.7 ppm larger values than FOCAL which is (due to the used color table) most noticeable in the northern hemisphere of Figure 14 (right). The standard deviation of the difference is 1.1 ppm. As done in the last section, we separate the systematic mismatch from the stochastic mismatch by analyzing grid boxes with more than 100 co-locations. The standard deviation of the stochastic and the systematic mismatch amounts 0.91 ppm and 0.83 ppm, respectively. It is no surprise, that the stochastic mismatch is smaller than expected from the combined

reported uncertainties because both data products base on the same L1b input data including the same noise spectra.

5.3. Validation with TCCON

In this section, we show validation results for one year (2015) of FOCAL v06 and NASA's operational OCO-2 L2 data and analyze the influence of the bias corrections. We used ground based TCCON [32] GGG2014 data obtained from <http://tcon.ornl.gov> as reference data set and a similar validation protocol as Reuter et al. [19,33]. We considered the column averaging kernels of all data products and co-located OCO-2 and TCCON measurements with a maximum time difference of 2 h, a maximum distance of 500 km, and a maximum surface elevation difference of 250 m. In cases with multiple TCCON measurements of the same site co-locating with an OCO-2 sounding, we averaged the TCCON measurements. In total we found about 179,000 and 378,000 co-locations for FOCAL and the NASA product, respectively.

Figure 15 shows the co-locations of all 19 sites with more than 250 co-locations. Per site statistics (bias and scatter, i.e., single sounding precision measured by the standard deviation of the difference to TCCON) are shown from north to south in Figure 16. Note that global offsets have been removed for both figures (-0.29 ppm, 0.35 ppm, -0.62 ppm, and -0.94 ppm for NASA, NASA not bias corrected, FOCAL, and FOCAL not bias corrected).

Both algorithms show a somewhat similar bias site-to-site pattern regardless whether the bias correction is applied or not. The largest differences of the bias corrected satellite products can be found in Sodankylä and Tsukuba with larger than 1 ppm biases of the NASA product and FOCAL. The standard deviations of the site biases are 0.82 ppm and 0.67 ppm for the NASA product and FOCAL (0.69 ppm and 0.78 ppm if no bias correction is applied). These algorithm-to-algorithm differences are barely significant because TCCON's per site accuracy is about 0.4 ppm (1σ) [32].

The analyzed algorithms also show a similar site-to-site pattern for the scatter with lowest values for the southern hemispheric sites probably due to smaller natural variability and, consequently, smaller representation errors. Both algorithms have a similar average scatter relative to TCCON before bias correction (1.62 ppm and 1.60 ppm for NASA and FOCAL) and after bias correction (1.31 ppm and 1.34 ppm for NASA and FOCAL). This means, both bias corrections primarily reduce the scatter rather than the site biases. However, according to Figure 12, the influence of the bias correction on the spatial bias pattern may be larger elsewhere.

FOCAL's retrieved XH_2O has also been initially compared with TCCON in the same manner. However, due to the much larger natural variability of water vapor (typically spanning a range from 500 ppm to 7000 ppm), we used stricter co-location criteria (1h maximum time difference and 150 km maximum distance) reducing the number of co-locations roughly by a factor of five. The global offset amounts to -150 ppm, the standard deviation of the site biases is 206 ppm, and the average single sounding precision is 293 ppm. It should be mentioned that, in contrast to XCO_2 , the agreement significantly reduces when relaxing the co-location criteria. Conversely, a significant part of the observed deviations could still be due to representation errors which are expected to reduce for even stricter co-location criteria. This, however, would also further reduce the number of co-locations.

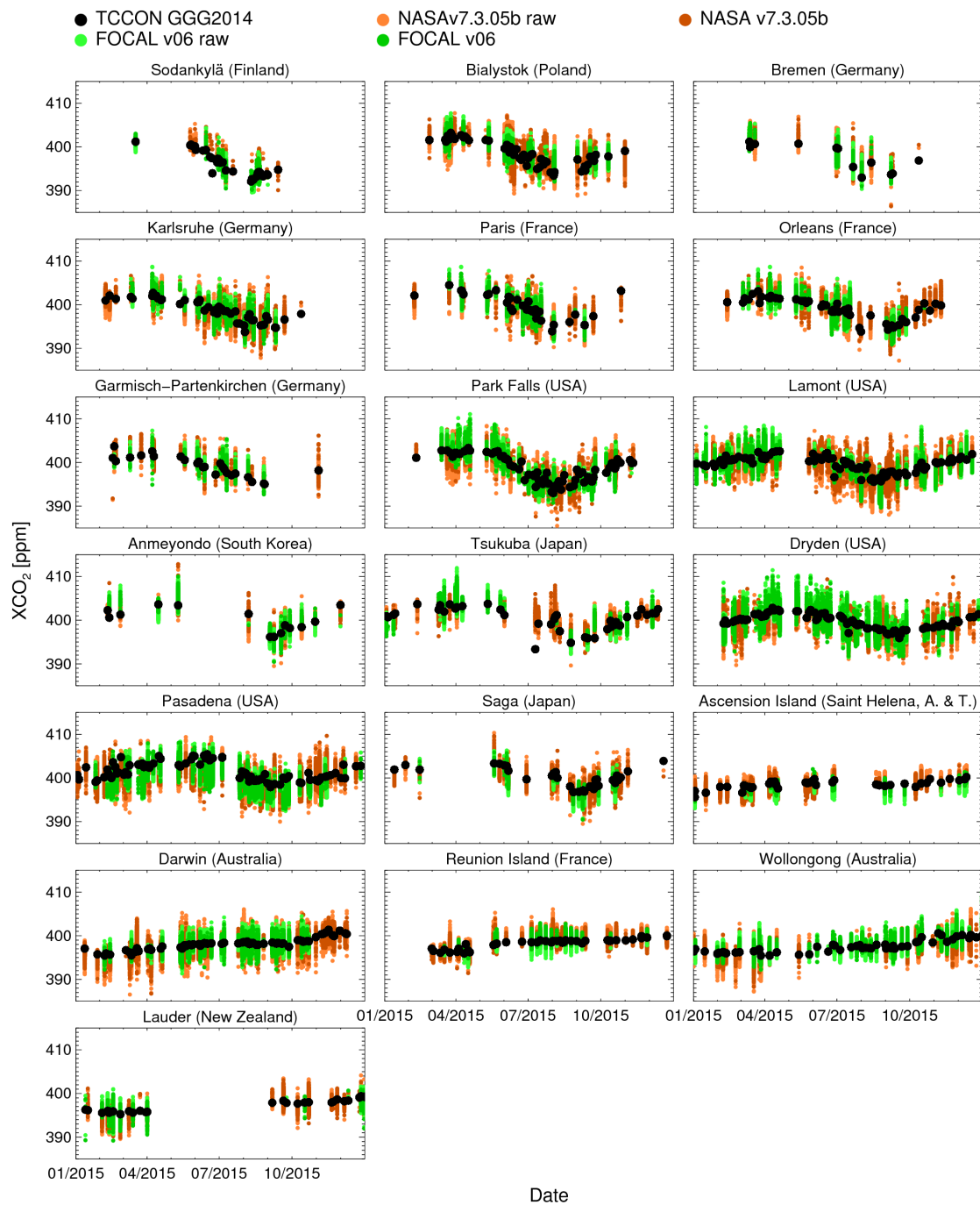


Figure 15. Validation of FOCAL v06 and NASA's operational OCO-2 L2 product (both with and without bias correction) with TCCON data from sites with more than 250 co-locations. The sites are ordered from north (top/left) to south (bottom/right): Sodankylä [34], Bialystok [35], Bremen [36], Karlsruhe [37], Paris [38], Orleans [39], Garmisch-Partenkirchen [40], Park Falls [41], Lamont [42], Anmeyondo [43], Tsukuba [44], Dryden [45], Pasadena [46], Saga [47], Ascension Island [48], Darwin [49], Reunion Island [50], Wollongong [51], and Lauder [52].

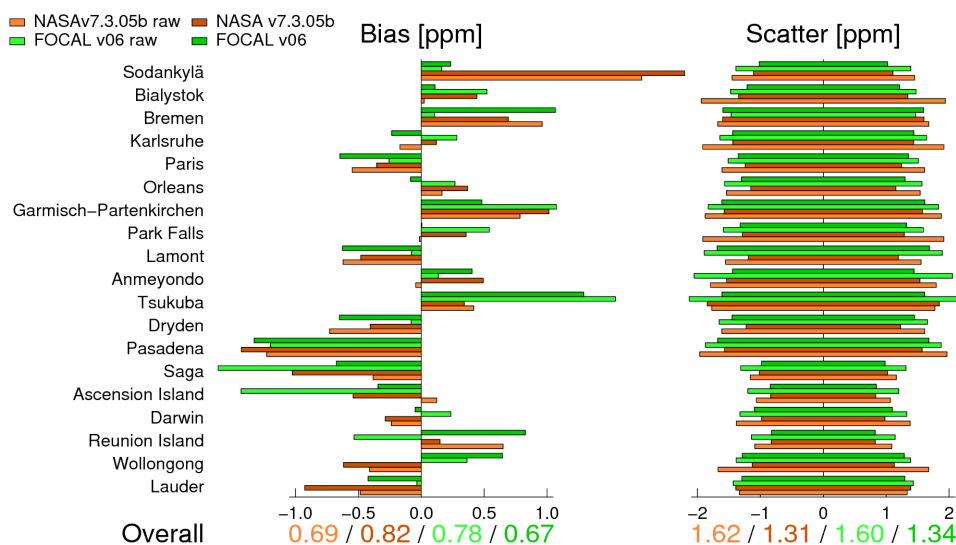


Figure 16. Validation statistics bias and scatter per TCCON site with more than 250 co-locations for FOCAL v06 and NASA’s operational OCO-2 L2 product (both with and without bias correction). The summarizing values (“overall”) represent the standard deviation of the site biases and the average scatter relative to TCCON, respectively. The sites are ordered from north (top) to south (bottom).

6. Conclusions

We used the fast atmospheric trace gas retrieval FOCAL v06 to retrieve XCO₂ from one year (2015) of OCO-2 measurements and presented the applied pre- and post-processing methods including pre-filtering, noise model, zero level offset correction, post-filtering, and bias correction.

The strict pre-filtering bases on sounding quality, NASA’s spike EOF analyses, OMI UV aerosol index, and MODIS Aqua cloud coverage. Due to the wider swath, MODIS cloud masking has the potential advantage to better account for 3D-effects caused by neighboring cloud contamination. However, as Aqua is lagging OCO-2 by 15 min, we chose a cloud filtering radius of 10 km, to prevent potential cloud movement from introducing cloud contamination. As a result this filter has a throughput of only about 11% and dominates the total pre-filtering throughput of about 4%. This makes this filter the main reason for FOCAL v06 having about three times less data points than NASA’s operational product and an OCO-2 based cloud filtering as also done by NASA [53] is a potential solution for future FOCAL versions.

In order to consider not only instrumental noise but also (pseudo) noise of the forward model, we set up a noise model that depends on the instrument noise and one free fit parameter which we determined from the residuals of a set of relatively unconstrained retrievals. The noise model suggests that forward model errors (plus potential pseudo noise of the instrument) have a magnitude of 0.5–3.2‰ of the continuum radiance. This means that in dark scenes the mismatch of simulated and measured radiance is still dominated by the noise of the instrument but in bright scenes (e.g., above deserts) the forward model error dominates.

Apparent or effective zero level offsets can have various reasons such as residual calibration errors or unconsidered spectroscopic effects. For the SIF, and both CO₂ fit windows, we found linear relationships between the retrieved zero level offsets and the continuum radiances with slopes between 0.8% and 1.8%. As FOCAL v06 usually does not retrieve the ZLO per sounding, we correct the measured radiance with the derived linear relationships before the retrieval.

Post-filtering checks for convergence, for fit window residuals being smaller than the thresholds derived from the noise model analyses, and for potential outliers. Non converging soundings are often found near the SAA, the Saharan desert, and the Arabian peninsula. Soundings with too large residuals are often found above the tropical oceans. The filter for potential outliers is most active in

high latitudes and its dominant input parameter (above land) is the retrieved Ångström exponent. The total post-filtering throughput is about 42%. The average RSR is 2.2%, 3.0%, 2.8%, and 3.4% for the SIF, the O₂, the wCO₂ and the sCO₂ fit window, respectively.

A bias correction has been applied to the post-filtered results which primarily bases on the assumption that XCO₂ has (on average) little variations on small scales so that correlations to more variable parameters can be used to quantify biases. As a consequence, the bias correction does not require any ground truth data except for a globally constant offset of −1.67 ppm. We found a distinct OCO-2 footprint dependent bias in the range between −0.97 ppm and 1.22 ppm but the most prominent global bias pattern results from the land/sea bias of 1.80 ppm. One could speculate that the land/sea bias has its origin in FOCAL's assumption of Lambertian surfaces. However, retrievals of simulated ocean observations done in part 1 [23] do not support this hypothesis.

We compared FOCAL v06 XCO₂ results with co-located values of the CAMS v15r4 model. Both data sets show similar large scale spatial patterns and the systematic temporal and regional differences have a standard deviation of 1.0 ppm. The standard deviation of the single sounding mismatches amounts to 1.1 ppm which agrees reasonably well with the average reported uncertainty of 1.2 ppm.

We also compared FOCAL's v06 XCO₂ with the operational NASA OCO-2 product. Large scale patterns of both data sets are similar and the most prominent difference is that the NASA product has about three times more soundings. The primary reason for this is the inherently poor throughput (11%) of the MODIS based cloud screening of FOCAL's preprocessor. The NASA product has on average 0.7 ppm larger values than FOCAL v06. The standard deviation of the difference between both products is 1.1 ppm.

Finally, we validated one year (2015) of FOCAL v06 XCO₂ data with and without bias correction as well as NASA's operational OCO-2 XCO₂ product with and without bias correction with co-located ground based TCCON observations. The algorithms show similarities in the site-to-site patterns of bias and scatter. The standard deviations of the site biases are 0.82 ppm and 0.67 ppm for the NASA product and FOCAL, respectively (0.69 ppm and 0.78 ppm without bias correction). These algorithm-to-algorithm differences are barely significant because TCCON's per site accuracy is about 0.4 ppm (1 σ) [32]. The average scatter relative to TCCON is 1.31 ppm and 1.34 ppm for NASA and FOCAL, respectively (1.62 ppm and 1.60 ppm without bias correction).

Additionally, we performed an initial validation of one year (2015) of FOCAL v06 XH₂O data with co-located ground based TCCON observations and found site-to-site biases with a standard deviation of 206 ppm and an average single sounding precision of 293 ppm. However, due to the much larger natural variability of XH₂O compared to XCO₂, future studies are needed to quantify or minimize the influence of representation errors.

Processing an entire year of OCO-2 data with FOCAL v06 took about two weeks on a small cluster with 8 Intel Xeon E5-2687W CPUs with eight cores running at 3.1 GHz (released in 2012). This means that FOCAL is fast enough to process data from current and future satellites similar to CarbonSat [4,54] providing at least an order of magnitude more data with a reasonable amount of CPU cores - especially, when taking into account the to be expected CPU developments until launch date. Additionally, FOCAL's computations are simple enough for an adaptation to GPU architecture with reasonable effort which has the potential for a further substantial acceleration.

Acknowledgments: This work was in part funded by ESA/ESRIN (GHG-CCI II project), the European Union (Copernicus Climate Change Service led by ECMWF), and the State and the University of Bremen. ECMWF provided the used meteorological profiles. 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. Additionally, NASA provided the MODIS L2 collection 6 MYD35 cloud mask data, the OMI L3 OMAERUVd v003 UV aerosol index data, and the absorption cross section database ABSCO. We used data of the CAMS greenhouse gas flux inversion model and thank F. Chevallier for his comments. The used solar spectra were made available by R.L. Kurucz and G.C. Toon, and we used a chlorophyll fluorescence spectrum published by U. Rascher. TCCON data were obtained from the TCCON Data Archive, hosted by the Carbon Dioxide Information Analysis Center (CDIAC).

Author Contributions: Maximilian Reuter: Design, implementation, and operation of the XCO₂ satellite retrieval; design, implementation, conduction, and interpretation of the comparison and validation experiments; writing the paper. Michael Buchwitz, Oliver Schneising, Stefan Noël, Heinrich Bovensmann, John P. Burrows: Support in design of the XCO₂ satellite retrieval; support in interpretation of the comparison and validation experiments; support in writing the paper.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Reuter, M.; Buchwitz, M.; Hilker, M.; Heymann, J.; Bovensmann, H.; Burrows, J.P.; Houweling, S.; Liu, Y.Y.; Nassar, R.; Chevallier, F.; et al. How much CO₂ is taken up by the European Terrestrial Biosphere? *Bull. Am. Meteorol. Soc.* **2017**, *98*, 665–671.
2. Miller, C.E.; Crisp, D.; DeCola, P.L.; Olsen, S.C.; Randerson, J.T.; Michalak, A.M.; Alkhaled, A.; Rayner, P.; Jacob, D.J.; Suntharalingam, P.; et al. Precision requirements for space-based XCO₂ data. *J. Geophys. Res.* **2007**, *112*, doi:10.1029/2006JD007659.
3. Chevallier, F.; Bréon, F.M.; Rayner, P.J. Contribution of the Orbiting Carbon Observatory to the estimation of CO₂ sources and sinks: Theoretical study in a variational data assimilation framework. *J. Geophys. Res.* **2007**, *112*, doi:10.1029/2006JD007375.
4. Bovensmann, H.; Buchwitz, M.; Burrows, J.P.; Reuter, M.; Krings, T.; Gerilowski, K.; Schneising, O.; Heymann, J.; Tretner, A.; Erzinger, J. A remote sensing technique for global monitoring of power plant CO₂ emissions from space and related applications. *Atmos. Meas. Tech.* **2010**, *3*, 781–811.
5. Burrows, J.P.; Hölzle, E.; Goede, A.P.H.; Visser, H.; Fricke, W. SCIAMACHY—Scanning imaging absorption spectrometer for atmospheric chartography. *Acta Astronaut.* **1995**, *35*, 445–451.
6. Bovensmann, H.; Burrows, J.P.; Buchwitz, M.; Frerick, J.; Noël, S.; Rozanov, V.V.; Chance, K.V.; Goede, A. SCIAMACHY—Mission objectives and measurement modes. *J. Atmos. Sci.* **1999**, *56*, 127–150.
7. Kuze, A.; Suto, H.; Nakajima, M.; Hamazaki, T. Thermal and near infrared sensor for carbon observation Fourier-transform spectrometer on the Greenhouse Gases Observing Satellite for greenhouse gases monitoring. *Appl. Opt.* **2009**, *48*, 6716.
8. Crisp, D.; Atlas, R.M.; Bréon, F.M.; Brown, L.R.; Burrows, J.P.; Ciais, P.; Connor, B.J.; Doney, S.C.; Fung, I.Y.; Jacob, D.J.; et al. The Orbiting Carbon Observatory (OCO) mission. *Adv. Space Res.* **2004**, *34*, 700–709.
9. Crisp, D.; Pollock, H.R.; Rosenberg, R.; Chapsky, L.; Lee, R.A.M.; Oyafuso, F.A.; Frankenberg, C.; O'Dell, C.W.; Bruegge, C.J.; Doran, G.B.; et al. The on-orbit performance of the Orbiting Carbon Observatory-2 (OCO-2) instrument and its radiometrically calibrated products. *Atmos. Meas. Tech.* **2017**, *10*, 59–81.
10. Buchwitz, M.; Rozanov, V.V.; Burrows, J.P. A near-infrared optimized DOAS method for the fast global retrieval of atmospheric CH₄, CO, CO₂, H₂O, and N₂O total column amounts from SCIAMACHY Envisat-1 nadir radiances. *J. Geophys. Res.* **2000**, *105*, 15231–15245.
11. Reuter, M.; Buchwitz, M.; Schneising, O.; Heymann, J.; Bovensmann, H.; Burrows, J.P. A method for improved SCIAMACHY CO₂ retrieval in the presence of optically thin clouds. *Atmos. Meas. Tech.* **2010**, *3*, 209–232.
12. Schneising, O.; Buchwitz, M.; Reuter, M.; Heymann, J.; Bovensmann, H.; Burrows, J.P. Long-term analysis of carbon dioxide and methane column-averaged mole fractions retrieved from SCIAMACHY. *Atmos. Chem. Phys.* **2011**, *11*, 2863–2880.
13. Heymann, J.; Bovensmann, H.; Buchwitz, M.; Burrows, J.P.; Deutscher, N.M.; Notholt, J.; Rettinger, M.; Reuter, M.; Schneising, O.; Sussmann, R.; et al. SCIAMACHY WFM-DOAS XCO₂: Reduction of scattering related errors. *Atmos. Meas. Tech.* **2012**, *5*, 2375–2390.
14. Butz, A.; Guerlet, S.; Hasekamp, O.; Schepers, D.; Galli, A.; Aben, I.; Frankenberg, C.; Hartmann, J.M.; Tran, H.; Kuze, A.; et al. Toward accurate CO₂ and CH₄ observations from GOSAT. *Geophys. Res. Lett.* **2011**, *38*, doi:10.1029/2011GL047888.
15. Cogan, A.J.; Boesch, H.; Parker, R.J.; Feng, L.; Palmer, P.I.; Blavier, J.F.L.; Deutscher, N.M.; Macatangay, R.; Notholt, J.; Roehl, C.; et al. Atmospheric carbon dioxide retrieved from the Greenhouse gases Observing SATellite (GOSAT): Comparison with ground-based TCCON observations and GEOS-Chem model calculations. *J. Geophys. Res. Atmos.* **2012**, *117*, doi:10.1029/2012JD018087.
16. O'Dell, C.W.; Connor, B.; Bösch, H.; O'Brien, D.; Frankenberg, C.; Castano, R.; Christi, M.; Eldering, D.; Fisher, B.; Gunson, M.; et al. The ACOS CO₂ retrieval algorithm—Part 1: Description and validation against synthetic observations. *Atmos. Meas. Tech.* **2012**, *5*, 99–121.

17. Yoshida, Y.; Kikuchi, N.; Morino, I.; Uchino, O.; Oshchepkov, S.; Bril, A.; Saeki, T.; Schutgens, N.; Toon, G.C.; Wunch, D.; et al. Improvement of the retrieval algorithm for GOSAT SWIR XCO₂ and XCH₄ and their validation using TCCON data. *Atmos. Meas. Tech.* **2013**, *6*, 1533–1547.
18. Buchwitz, M.; Reuter, M.; Schneising, O.; Hewson, W.; Detmers, R.G.; Boesch, H.; Hasekamp, O.P.; Aben, I.; Bovensmann, H.; Burrows, J.P.; et al. Global satellite observations of column-averaged carbon dioxide and methane: The GHG-CCI XCO₂ and XCH₄ CRDP3 data set. *Remote Sens. Environ.* **2017**, doi:10.1016/j.rse.2016.12.027.
19. Reuter, M.; Bösch, H.; Bovensmann, H.; Bril, A.; Buchwitz, M.; Butz, A.; Burrows, J.P.; O'Dell, C.W.; Guerlet, S.; Hasekamp, O.; et al. A joint effort to deliver satellite retrieved atmospheric CO₂ concentrations for surface flux inversions: The ensemble median algorithm EMMA. *Atmos. Chem. Phys.* **2013**, *13*, 1771–1780.
20. Reuter, M.; Buchwitz, M.; Hilker, M.; Heymann, J.; Schneising, O.; Pillai, D.; Bovensmann, H.; Burrows, J.P.; Bösch, H.; Parker, R.; et al. Satellite-inferred European carbon sink larger than expected. *Atmos. Chem. Phys.* **2014**, *14*, 13739–13753.
21. Boesch, H.; Brown, L.; Castano, R.; Christi, M.; Connor, B.; Crisp, D.; Eldering, A.; Fisher, B.; Frankenberg, C.; Gunson, M.; et al. *Orbiting Carbon Observatory-2 (OCO-2) Level 2 Full Physics Retrieval Algorithm Theoretical Basis; Version 2.0 Rev 2*; National Aeronautics and Space Administration, Jet Propulsion Laboratory, California Institute of Technology: Pasadena, CA, USA, 2015.
22. Eldering, A.; O'Dell, C.W.; Wennberg, P.O.; Crisp, D.; Gunson, M.R.; Viatte, C.; Avis, C.; Braverman, A.; Castano, R.; Chang, A.; et al. The Orbiting Carbon Observatory-2: First 18 months of science data products. *Atmos. Meas. Tech.* **2017**, *10*, 549.
23. Reuter, M.; Buchwitz, M.; Schneising, O.; Noël, S.; Rozanov, V.; Bovensmann, H.; Burrows, J.P. A fast atmospheric trace gas retrieval for hyperspectral instruments approximating multiple scattering—Part 1: Radiative transfer and a potential OCO-2 XCO₂ retrieval setup. *Remote Sens.* **2017**, submitted.
24. Eldering, A.; Pollock, R.; Lee, R.; Rosenberg, R.; Oyafuso, F.; Crisp, D.; Chapsky, L.; Granat, R. *Orbiting Carbon Observatory-2 (OCO-2) - LEVEL 1B - Algorithm Theoretical Basis; Version 1.2 Rev 1*; National Aeronautics and Space Administration, Jet Propulsion Laboratory, California Institute of Technology: Pasadena, CA, USA, 2015.
25. Stammes, P. *OMI Algorithm Theoretical Basis Document, Volume III, Clouds, Aerosols, and Surface UV Irradiance (ATBD-OMI-03)*; Royal Netherlands Meteorological Institute (KNMI): De Bilt, The Netherlands, 2002.
26. Ackerman, S.; Frey, R.; Strabala, K.; Liu, Y.; Gumley, L.; Baum, B.; Menzel, P. *Discriminating Clear-Sky from Cloud with MODIS—Algorithm Theoretical Basis Document (MOD35)*; Version 6.1; Cooperative Institute for Meteorological Satellite Studies, University of Wisconsin—Madison: Madison, WI, USA, 2010.
27. Heymann, J.; Reuter, M.; Hilker, M.; Buchwitz, M.; Schneising, O.; Bovensmann, H.; Burrows, J.P.; Kuze, A.; Suto, H.; Deutscher, N.M.; et al. Consistent satellite XCO₂ retrievals from SCIAMACHY and GOSAT using the BESD algorithm. *Atmos. Meas. Tech.* **2015**, *8*, 2961–2980.
28. Hu, C.; Lee, Z.; Franz, B. Chlorophyll-a algorithms for oligotrophic oceans: A novel approach based on three-band reflectance difference. *J. Geophys. Res. Oceans* **2012**, *117*, doi:10.1029/2011JC007395.
29. Chevallier, F.; Ciais, P.; Conway, T.J.; Aalto, T.; Anderson, B.E.; Bousquet, P.; Brunke, E.G.; Ciattaglia, L.; Esaki, Y.; Fröhlich, M.; et al. CO₂ surface fluxes at grid point scale estimated from a global 21 year reanalysis of atmospheric measurements. *J. Geophys. Res.* **2010**, *115*, doi:10.1029/2010JD013887.
30. Chevallier, F. *Validation report for the inverted CO2 fluxes, v16r1*; CAMS deliverable CAMS73_2015S2_D73.1.4.2-1979-2016-v1_201707; 2017. Available online: https://atmosphere.copernicus.eu/sites/default/files/FileRepository/Resources/Validation-reports/Fluxes/CAMS73_2015SC2_D73.1.4.2-1979-2016-v1_201707_final.pdf (accessed on 29 October 2017).
31. Rodgers, C.D. *Inverse Methods for Atmospheric Sounding: Theory and Practice*; World Scientific Publishing: Singapore, 2000.
32. Wunch, D.; Toon, G.C.; Blavier, J.F.L.; Washenfelder, R.A.; Notholt, J.; Connor, B.J.; Griffith, D.W.T.; Sherlock, V.; Wennberg, P.O. The Total Carbon Column Observing Network (TCCON). *Philos. Trans. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.* **2011**, *369*, 2087–2112.
33. Reuter, M.; Bovensmann, H.; Buchwitz, M.; Burrows, J.P.; Connor, B.J.; Deutscher, N.M.; Griffith, D.W.T.; Heymann, J.; Keppel-Aleks, G.; Messerschmidt, J.; et al. Retrieval of atmospheric CO₂ with enhanced accuracy and precision from SCIAMACHY: Validation with FTS measurements and comparison with model results. *J. Geophys. Res.* **2011**, *116*, doi:10.1029/2010JD015047.

34. Kivi, R.; Heikkinen, P.; Kyro, E. TCCON data from Sodankyla, Finland, Release GGG2014R0. In *TCCON Data Archive*; Carbon Dioxide Information Analysis Center (CDIAC): Sacramento, CA, USA, 2014.
35. Deutscher, N.; Notholt, J.; Messerschmidt, J.; Weinzierl, C.; Warneke, T.; Petri, C.; Grupe, P.; Katrynski, K. TCCON data from Bialystok, Poland, Release GGG2014R1. In *TCCON Data Archive*; Carbon Dioxide Information Analysis Center (CDIAC): Sacramento, CA, USA, 2014.
36. Notholt, J.; Petri, C.; Warneke, T.; Deutscher, N.; Buschmann, M.; Weinzierl, C.; Macatangay, R.; Grupe, P. TCCON data from Bremen, Germany, Release GGG2014R0. In *TCCON Data Archive*; Carbon Dioxide Information Analysis Center (CDIAC): Sacramento, CA, USA, 2014.
37. Hase, F.; Blumenstock, T.; Dohe, S.; Groß, J.; Kiel, M. TCCON data from Karlsruhe, Germany, Release GGG2014R1. In *TCCON Data Archive*; Carbon Dioxide Information Analysis Center (CDIAC): Sacramento, CA, USA, 2014.
38. Te, Y.; Jeseck, P.; Janssen, C. TCCON data from Paris, France, Release GGG2014R0. In *TCCON Data Archive*; Carbon Dioxide Information Analysis Center (CDIAC): Sacramento, CA, USA, 2014.
39. Warneke, T.; Messerschmidt, J.; Notholt, J.; Weinzierl, C.; Deutscher, N.; Petri, C.; Grupe, P.; Vuillemin, C.; Truong, F.; Schmidt, M.; et al. TCCON data from Orleans, France, Release GGG2014R0. In *TCCON Data Archive*; Carbon Dioxide Information Analysis Center (CDIAC): Sacramento, CA, USA, 2014.
40. Sussmann, R.; Rettinger, M. TCCON data from Garmisch, Germany, Release GGG2014R0. In *TCCON Data Archive*; Carbon Dioxide Information Analysis Center (CDIAC): Sacramento, CA, USA, 2014.
41. Wennberg, P.O.; Roehl, C.; Wunch, D.; Toon, G.C.; Blavier, J.F.; Washenfelder, R.; Keppel-Aleks, G.; Allen, N.; Ayers, J. TCCON data from Park Falls, Wisconsin, USA, Release GGG2014R0. In *TCCON Data Archive*; Carbon Dioxide Information Analysis Center (CDIAC): Sacramento, CA, USA, 2014.
42. Wennberg, P.O.; Wunch, D.; Roehl, C.; Blavier, J.F.; Toon, G.C.; Allen, N.; Dowell, P.; Teske, K.; Martin, C.; Martin, J. TCCON data from Lamont, Oklahoma, USA, Release GGG2014R1. In *TCCON Data Archive*; Carbon Dioxide Information Analysis Center (CDIAC): Sacramento, CA, USA, 2014.
43. Goo, T.Y.; Oh, Y.S.; Velazco, V.A. TCCON data from Anmeyondo, South Korea, Release GGG2014R0. In *TCCON Data Archive*; Carbon Dioxide Information Analysis Center (CDIAC): Sacramento, CA, USA, 2014.
44. Morino, I.; Matsuzaki, T.; Shishime, A. TCCON data from Tsukuba, Ibaraki, Japan, 125HR, Release GGG2014R1. In *TCCON Data Archive*; Carbon Dioxide Information Analysis Center (CDIAC): Sacramento, CA, USA, 2014.
45. Iraci, L.; Podolske, J.; Hillyard, P.; Roehl, C.; Wennberg, P.O.; Blavier, J.F.; Landeros, J.; Allen, N.; Wunch, D.; Zavaleta, J.; et al. TCCON data from Armstrong Flight Research Center, Edwards, CA, USA, Release GGG2014R1. In *TCCON Data Archive*; Carbon Dioxide Information Analysis Center (CDIAC): Sacramento, CA, USA, 2014.
46. Wennberg, P.O.; Wunch, D.; Yavin, Y.; Toon, G.C.; Blavier, J.F.; Allen, N.; Keppel-Aleks, G. TCCON data from Jet Propulsion Laboratory, Pasadena, California, USA, Release GGG2014R0. In *TCCON Data Archive*; Carbon Dioxide Information Analysis Center (CDIAC): Sacramento, CA, USA, 2014.
47. Shiomi, K.; Kawakami, S.; Ohyama, H.; Arai, K.; Okumura, H.; Taura, C.; Fukamachi, T.; Sakashita, M. TCCON data from Saga, Japan, Release GGG2014R0. In *TCCON Data Archive*; Carbon Dioxide Information Analysis Center (CDIAC): Sacramento, CA, USA, 2014.
48. Feist, D.G.; Arnold, S.G.; John, N.; Geibel, M.C. TCCON data from Ascension Island, Saint Helena, Ascension and Tristan da Cunha, Release GGG2014R0. In *TCCON Data Archive*; Carbon Dioxide Information Analysis Center (CDIAC): Sacramento, CA, USA, 2014.
49. Griffith, D.W.T.; Deutscher, N.; Velazco, V.A.; Wennberg, P.O.; Yavin, Y.; Aleks, G.K.; Washenfelder, R.; Toon, G.C.; Blavier, J.F.; Murphy, C.; et al. TCCON data from Darwin, Australia, Release GGG2014R0. In *TCCON Data Archive*; Carbon Dioxide Information Analysis Center (CDIAC): Sacramento, CA, USA, 2014.
50. De Maziere, M.; Sha, M.K.; Desmet, F.; Hermans, C.; Scolas, F.; Kumps, N.; Metzger, J.M.; Dufлот, V.; Cammas, J.P. TCCON data from Reunion Island (La Reunion), France, Release GGG2014R0. In *TCCON Data Archive*; Carbon Dioxide Information Analysis Center (CDIAC): Sacramento, CA, USA, 2014.
51. Griffith, D.W.T.; Velazco, V.A.; Deutscher, N.; Murphy, C.; Jones, N.; Wilson, S.; Macatangay, R.; Kettlewell, G.; Buchholz, R.R.; Riggensbach, M. TCCON data from Wollongong, Australia, Release GGG2014R0. In *TCCON Data Archive*; Carbon Dioxide Information Analysis Center (CDIAC): Sacramento, CA, USA, 2014.
52. Sherlock, V.; Connor, B.; Robinson, J.; Shiona, H.; Smale, D.; Pollard, D. TCCON data from Lauder, New Zealand, 125HR, Release GGG2014R0. In *TCCON Data Archive*; Carbon Dioxide Information Analysis Center (CDIAC): Sacramento, CA, USA, 2014.

53. Taylor, T.E.; O'Dell, C.W.; Partain, P.T.; Cronk, H.Q.; Nelson, R.R.; Rosenthal, E.J.; Chang, A.Y.; Osterman, G.B.; Pollock, R.H.; Gunson, M.R. Orbiting Carbon Observatory-2 (OCO-2) cloud screening algorithms: Validation against collocated MODIS and CALIOP data. *Atmos. Meas. Tech.* **2016**, *9*, 973.
54. Buchwitz, M.; Reuter, M.; Bovensmann, H.; Pillai, D.; Heymann, J.; Schneising, O.; Rozanov, V.; Krings, T.; Burrows, J.P.; Boesch, H.; et al. Carbon Monitoring Satellite (CarbonSat): Assessment of atmospheric CO₂ and CH₄ retrieval errors by error parameterization. *Atmos. Meas. Tech.* **2013**, *6*, 3477–3500.



© 2017 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).