ESA Climate Change Initiative “Plus” (CCI+)

End-to-End ECV Uncertainty Budget
Version 1 (E3UBv1)

for the Essential Climate Variable (ECV)
Greenhouse Gases (GHG)

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

ESA CCI via project GHG-CCI+
The European Commission via the H2020 projects
CHE (Grant Agreement No. 776186)
VERIFY (Grant Agreement No. 776810)
## Change log

<table>
<thead>
<tr>
<th>Version Nr.</th>
<th>Date</th>
<th>Status</th>
<th>Reason for change</th>
</tr>
</thead>
<tbody>
<tr>
<td>Version 1 draft</td>
<td>29. Nov. 2019</td>
<td>Submitted for review</td>
<td>New document</td>
</tr>
<tr>
<td>Version 1</td>
<td>6. Jan. 2020</td>
<td>Final / submitted</td>
<td>Improved Fig. 5</td>
</tr>
</tbody>
</table>
## Table of Contents

1 Introduction .................................................................................................................... 4  
1.1 Purpose of document .............................................................................................. 4  
1.2 Intended audience .................................................................................................. 4  
2 Uncertainty analysis based on simulations ............................................................... 5  
3 Comparison with CAMS model results ...................................................................... 9  
4 Comparison with NASA’s operational OCO-2 L2 product ......................................... 11  
5 Validation with TCCON ............................................................................................... 12  
5.1 XCO₂ .................................................................................................................... 12  
5.2 XCO₂ uncertainty ................................................................................................. 18  
Acknowledgements ...................................................................................................... 20  
References ...................................................................................................................... 21
1 Introduction

1.1 Purpose of document

This document describes the End-to-End ECV Uncertainty Budget (E3UB) of the FOCAL OCO-2 XCO₂ product. It provides an overview of random and systematic errors affecting the FOCAL OCO-2 XCO₂ L2 product. Reliable uncertainty estimates of the retrieval are required to translate remotely sensed data into modelled estimations with a known degree of confidence, allowing, e.g., the detection of climate change impacts additional to the natural variability of greenhouse gases. In particular, the GHG-CCI User Requirements have placed strict measurement accuracy and precision requirements on the participating GHG retrievals, allowing identification of minute changes in magnitude and sign of XCO₂ concentration change (/URDv2.1, 2017/).

1.2 Intended audience

This document is intended for users in the modelling community applying IUPI’s FOCAL OCO-2 XCO₂ L2 product for CO₂ inversions, as well as remote sensing experts interested in atmospheric soundings of XCO₂. In both cases, the work presented here will give the user a more thorough understanding of uncertainties implicit in this GHG-CCI+ product.
2 Uncertainty analysis based on simulations

This section summarizes the results shown in section 7 of FOCAL’s /ATBDv1/ which, in turn, bases on the publication of /Reuter et al., 2017a/, in parts.

In order to assess FOCAL’s theoretical capabilities (primarily in retrieving XCO₂, XH₂O, and SIF), /Reuter et al., 2017a/ confronted it with radiance measurements simulated with the accurate RT code SCIATRAN /Rozanov et al., 2014/. The performed analyses can be understood also as test of the suitability of the approximations made in FOCAL’s RT and of the retrieval setup. Hereby, they primarily concentrated on scattering related errors and analyzed the systematic and stochastic, i.e., the a posteriori errors of several different geophysical scenarios and retrieval setups. They were not aiming 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 and different post-filtering capabilities. The aim of their retrieval experiments, summarized in this section, was rather to identify a retrieval setup which is a promising candidate for a full retrieval scheme and its application to actually measured OCO-2 data. Their 3-Scat retrieval setup (see /Reuter et al., 2017a/ for details) showed often relatively low systematic XCO₂ and XH₂O errors with low polarization dependency, well controlled retrieved profiles, lowest CO₂ smoothing errors, a relatively realistic a priori error correlation matrix for CO₂, and advantageous AKs. Therefore, they identified this retrieval setup as promising candidate which, in turn, became the basis for FOCAL v8.

The 3-Scat setup fits the OCO-2 measured radiance in four fit windows by simultaneously retrieving 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, SIF, and polynomial coefficients describing the spectral albedo in each fit window.

As accurate XCO₂ retrievals will probably always require a rigorous cloud and aerosol screening, /Reuter et al., 2017a/ concentrated on scenarios with scattering optical thicknesses in the range of about 0.03 and 0.70. The quality of the spectral fits in the O₂ fit window was usually 2.5 to 4 times better than expected from instrumental noise. In the CO₂ fit windows, the quality of the spectral fits was usually at least 7 times better than expected from instrumental noise and even smaller fit residuals were obtained in the SIF fit window. Figure 1 shows some example fit residuals.
Figure 1: Residuals (fit minus measurement, red) and measurement noise (gray) of SCIATRAN simulated OCO-2 measurements fitted with FOCAL (solar zenith angle = 40°, parallel polarization). Top: Rayleigh scenario and the 0-Scat setup (allowing not the fitting of scattering parameters). Middle: Rayleigh scenario and the 3-Scat setup. Bottom: Rayleigh+AerosolBG+Water cloud scenario and the 3-Scat setup. See /Reuter et al., 2017a/ for details on the geophysical scenarios and retrieval setups.
Systematic errors of XCO₂ ranged from -2.5 ppm to 3.0 ppm and were usually smaller than ±0.3ppm (for the tested scenarios). The stochastic uncertainty of XCO₂ was typically about 1.0ppm (Figure 2). Systematic errors of XH₂O ranged from -243 ppm to 0 ppm and were usually smaller than ±6 ppm. The stochastic uncertainty of XH₂O was typically about 9ppm. Note, 1000ppmb correspond to 6.44 kg/m² for the analyzed H₂O profiles. The degree of freedom for the retrieved five-layered CO₂ and H₂O profiles was typically 2.2. As SIF is retrieved from Fraunhofer lines in a spectral region with negligible gaseous absorption features, it can be retrieved without significant interferences with the retrieved scattering properties. The systematic SIF errors were always below 0.02 mW/m²/sr/nm. Therefore, /Reuter et al., 2017a/ expected that instrumental or forward model effects causing an infilling (a reduction of the line depths) of the used Fraunhofer lines will dominate the systematic errors when analyzing actually measured data. The stochastic uncertainty of SIF was usually below 0.3 mW/m²/sr/nm.
Figure 2: Error characteristics of nine retrieval setups and twelve geophysical scenarios. Each box includes six sub-boxes representing polarization parallel (left) and perpendicular (right) to the SPP as well as three solar zenith angles (20°, 40°, and 60°, from bottom to top). Gray boxes represent not converging retrievals. **Left:** Systematic error (retrieved minus true XCO2). **Right:** Stochastic uncertainty as reported by the optimal estimation retrieval. See [Reuter et al., 2017a](#) for details on the geophysical scenarios and retrieval setups.
3 Comparison with CAMS model results

This section bases on section 8.1 of FOCAL’s /ATBDv1, 2019/ which, in turn, summarizes results of a comparison of FOCAL v06 with the CAMS model done by /Reuter et al., 2017b/.

Here 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., Rodgers, 2000). **Figure 3** 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.

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 standard deviation of this systematic mismatch (including also representation errors) amounts to 1.0ppm. 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.
Figure 3: Monthly mean XCO₂ gridded to 5°×5°. Top: FOCAL v06. Bottom: CAMSv15r4 sampled as FOCAL. Left: April 2015. Right: August 2015.
4 Comparison with NASA’s operational OCO-2 L2 product

This section bases on section 8.2 of FOCAL’s /ATBDv1, 2019/ which, in turn, summarizes results of a comparison of FOCAL v06 with the CAMS model done by /Reuter et al., 2017b/.

In this section we compare the same two months (April and August 2015) of post-filtered and bias corrected FOCAL v06 XCO₂ results with NASA’s operational OCO-2 L2 product. Comparing Figure 4 with Figure 3 (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 Sec. 5).

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 4 (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.

Figure 4: NASA v7.3.05b monthly mean XCO₂ at 5°×5°. Left: April 2015. Right: August 2015.
5 Validation with TCCON

The validation results shown in this section are valid for FOCAL v08. The applied methods are similar to those described in BESD's Comprehensive Error Characterisation Report /CECRv3, 2017/ and the Product Validation and Intercomparison Report /PVIRv5, 2017/ of ESA’s GHG CCI project and partly also in the publication of /Reuter et al., 2011/. For all comparisons, averaging kernels have been applied as described in the C3S GHG Product User Guide and Specification /PUGS, 2017/.

5.1 XCO2

FOCAL’s XCO2 has been validated with TCCON GGG2014 measurements. The co-location criteria are defined by a maximal time difference of two hours, a maximal spatial distance of 500 km, and a maximal surface elevation difference of 250 m. Figure 5 shows all co-located FOCAL and TCCON retrievals of the years 2015-2018 for TCCON sites with more than 250 co-locations and covering a time period of at least one year. One can see that FOCAL captures the year-to-year increase and the seasonal features. For each station, the performance statistics number of co-locations, station bias, seasonal bias, linear drift, and single measurement precision were calculated.

We define the station bias as average difference to TCCON. Seasonal bias, linear drift, and single sounding precision have been derived by fitting the following trend model:

\[ \Delta X = a_0 + a_1 t + a_2 \sin(2\pi t + a_3) + \varepsilon \]

Here, \( \Delta X \) represents the difference satellite minus TCCON, and \( a_{0-3} \) the free fit parameters. Specifically, \( a_1 \) represents the linear drift and \( a_2 \) the amplitude of the seasonal bias. The single sounding precision is computed by the standard deviation of the residua \( \varepsilon \).

Based on the per station statistics, the following summarizing statistics have been calculated: Total number of co-locations used for validation, averaged single measurement precision, station-to-station bias (standard deviation of the station biases), average seasonal bias (standard deviation of the seasonal bias term), and average linear drift. As the linear drift can be assumed to be globally constant, the station-to-station standard deviation of the linear drift is a measure for its uncertainty. Per station statistics and overall performance estimates are listed in Table 1.

In total, \(~600000\) co-located FOCAL measurements have been used for the validation exercise. The overall single measurement precision is 1.52 ppm and station-to-station biases amount to 0.64 ppm.

In the context of station-to-station biases, it shall be noted that /Wunch et al., 2010, 2011/ specifies the accuracy (1\(\sigma\)) of TCCON to be about 0.4 ppm. This means it cannot be expected to find regional biases considerably less than 0.4 ppm using TCCON as reference.
Seasonal cycle biases amount to 0.38 ppm on average and no significant (temporally linear) drift can be found (0.00±0.75 ppm/a).

Additionally, a measure for the year-to-year stability has been computed as follows. For each TCCON site, the residual difference (satellite - TCCON) which is not explained by station bias, seasonal bias, and/or linear drift has been derived by subtracting the fit of the bias model $\Delta X$ from the satellite minus TCCON difference. These time series were smoothed by a running average of 365 days. Only days where more than 10 co-locations contributed to the running average of at least 5 TCCON sites have been further considered. At these days, the station-to-station average has been calculated. The corresponding expected uncertainty has been 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. For FOCAL, the average is always between about -0.3ppm and 0.2ppm (Figure 6) with an uncertainty of typically about 0.15ppm. Most of the time, the average is not significantly different from zero, i.e., its one 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 (0.21ppm) and standard deviation (0.01). From this, we conclude that the year-to-year stability is 0.21 ppm/a (Figure 6).
Figure 5: Validation of single soundings of FOCAL (green) with co-located TCCON measurements (black) at all TCCON sites with more than 250 co-locations and covering a time period of at least one year. Numbers in the figures: Δ = station bias, i.e., average of the difference; σ = single measurement precision, i.e., standard deviation of the difference; N = number of co-locations.
Figure 6: 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 and error propagation from single sounding measurement noise.

Stability FOCAL v08

avg. of station bias  std. dev. of station bias

y–t–y stability [ppm/a]: 0.21 ± 0.01
### Table 1: Validation statistics for all TCCON sites with more than 250 co-locations and covering a time period of at least one year with number of co-locations (#), single measurement precision (σ), station bias (Δ), seasonal bias (s) and linear drift (d). The last row contains the overall statistics. In this row σ represents the (quadratic) average single measurement precision, Δ the station-to-station bias (i.e., the standard deviation of the station biases), s the average seasonal bias, and d the average drift plus minus its standard deviation.

<table>
<thead>
<tr>
<th>Station</th>
<th>#</th>
<th>σ [ppm]</th>
<th>Δ [ppm]</th>
<th>s [ppm]</th>
<th>d [ppm/a]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sodankylä</td>
<td>6270</td>
<td>1.16</td>
<td>0.23</td>
<td>0.24</td>
<td>-0.08</td>
</tr>
<tr>
<td>East Trout Lake</td>
<td>7094</td>
<td>1.44</td>
<td>0.46</td>
<td>0.64</td>
<td>0.46</td>
</tr>
<tr>
<td>Bialystok</td>
<td>16546</td>
<td>1.38</td>
<td>0.14</td>
<td>0.12</td>
<td>0.00</td>
</tr>
<tr>
<td>Bremen</td>
<td>9415</td>
<td>1.71</td>
<td>0.13</td>
<td>0.40</td>
<td>-0.34</td>
</tr>
<tr>
<td>Karlsruhe</td>
<td>27916</td>
<td>1.49</td>
<td>0.39</td>
<td>0.70</td>
<td>0.16</td>
</tr>
<tr>
<td>Paris</td>
<td>21642</td>
<td>1.36</td>
<td>-0.80</td>
<td>0.48</td>
<td>-0.08</td>
</tr>
<tr>
<td>Orleans</td>
<td>31232</td>
<td>1.31</td>
<td>0.50</td>
<td>0.20</td>
<td>0.24</td>
</tr>
<tr>
<td>Garmisch-P.</td>
<td>3350</td>
<td>1.53</td>
<td>0.52</td>
<td>0.52</td>
<td>0.23</td>
</tr>
<tr>
<td>Park Falls</td>
<td>33631</td>
<td>1.34</td>
<td>-0.04</td>
<td>0.56</td>
<td>0.15</td>
</tr>
<tr>
<td>Rikubetsu</td>
<td>894</td>
<td>1.78</td>
<td>0.45</td>
<td>0.58</td>
<td>0.76</td>
</tr>
<tr>
<td>Lamont</td>
<td>81746</td>
<td>1.62</td>
<td>-0.13</td>
<td>0.18</td>
<td>-0.16</td>
</tr>
<tr>
<td>Anmeyondo</td>
<td>3674</td>
<td>1.37</td>
<td>0.39</td>
<td>0.29</td>
<td>-0.25</td>
</tr>
<tr>
<td>Tsukuba</td>
<td>31374</td>
<td>1.60</td>
<td>0.20</td>
<td>0.19</td>
<td>0.25</td>
</tr>
<tr>
<td>Dryden</td>
<td>67046</td>
<td>1.59</td>
<td>0.14</td>
<td>0.39</td>
<td>-0.04</td>
</tr>
<tr>
<td>Pasadena</td>
<td>14361</td>
<td>2.01</td>
<td>-0.23</td>
<td>1.01</td>
<td>-3.19</td>
</tr>
<tr>
<td>Pasadena</td>
<td>71977</td>
<td>1.68</td>
<td>-1.46</td>
<td>0.22</td>
<td>-0.12</td>
</tr>
<tr>
<td>Saga</td>
<td>25614</td>
<td>1.66</td>
<td>-1.07</td>
<td>0.18</td>
<td>0.14</td>
</tr>
<tr>
<td>Heifei</td>
<td>4946</td>
<td>2.21</td>
<td>-1.51</td>
<td>0.25</td>
<td>1.01</td>
</tr>
<tr>
<td>Burgos</td>
<td>4639</td>
<td>1.10</td>
<td>0.31</td>
<td>0.15</td>
<td>0.60</td>
</tr>
<tr>
<td>Ascension Island</td>
<td>10977</td>
<td>1.11</td>
<td>0.40</td>
<td>0.23</td>
<td>0.17</td>
</tr>
<tr>
<td>Darwin</td>
<td>67529</td>
<td>1.41</td>
<td>-0.01</td>
<td>0.24</td>
<td>-0.06</td>
</tr>
<tr>
<td>Reunion Island</td>
<td>17988</td>
<td>0.94</td>
<td>1.04</td>
<td>0.30</td>
<td>0.00</td>
</tr>
<tr>
<td>Wollongong</td>
<td>30508</td>
<td>1.24</td>
<td>0.37</td>
<td>0.31</td>
<td>-0.16</td>
</tr>
<tr>
<td>Lauder</td>
<td>9805</td>
<td>1.84</td>
<td>-0.40</td>
<td>0.81</td>
<td>0.35</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>600174</td>
<td>1.52</td>
<td>0.64</td>
<td>0.38</td>
<td>0.00±0.75</td>
</tr>
</tbody>
</table>
5.2 XCO₂ uncertainty

Especially for the application of flux inversion, reliable information on the uncertainty of each individual sounding is necessary. For this purpose, we analyzed the same validation dataset of co-located FOCAL and TCCON measurements also used in Section 5.1.

For each co-location used for the validation shown in Section 5.1, we have a residual $\varepsilon$ of the bias model $\Delta X$. From this residual, we computed our best estimate for the stochastic uncertainty (precision) as it does not include the analyzed systematic biases (trend, seasonal cycle, station-to-station).

For each $\varepsilon$, we have a corresponding uncertainty reported by FOCAL’s optimal estimation retrieval. We pooled the entire data set of about 600000 co-locations into 20 bins with increasing reported uncertainty in a way that each bin included the same number of co-locations (about 30000). In each bin, we computed the (quadratic) average reported uncertainty and the standard deviation of the residual $\varepsilon$ (true precision).

Figure 7 shows that both quantities are connected by a more or less linear relationship, except for an outlier in the bin of the largest reported uncertainties. The reported uncertainty is mainly driven by the instrumental noise which is in turn driven by the radiance so that the darkest scenes usually have the largest reported uncertainties. This means, especially the bins including the largest (or smallest) reported uncertainties may be dominated by an individual validation site with especially dark (or bright) albedo, while the other bins usually consist of data from a larger mixture of TCCON sites.

The linear fit shown in Figure 7 shows that FOCAL’s reported uncertainties has a positive correlation with the true precision but it shows also that FOCAL’s reported uncertainty is somewhat to optimistic. However it shall be noted that the residual $\varepsilon$ does not only include instrumental noise but also pseudo noise from representation errors.

In summary, we suggest that users who are interested in more realistic uncertainty estimates, shall apply the following error parameterization derived from the linear fit shown in Figure 7:

$$
\sigma_{\text{corrected}}^{\text{XCO}_2} = \sigma_0^{\text{XCO}_2} \cdot 1.128 + 0.128 \text{ppm}
$$
Figure 7: Reported uncertainty of FOCAL’s optimal estimation retrieval vs. true precision computed from the residual $\varepsilon$ of the bias model of Section 5.1.
Acknowledgements

The FOCAL development and OCO-2 data processing and analysis is co-funded by ESA’s Climate Change Initiative (CCI+) via project GHG-CCI+ (http://www.esa-ghg-cci.org/), the European Union via the Horizon 2020 (H2020) projects VERIFY (Grant Agreement No. 776810, see also: http://verify.lsce.ipsl.fr) and CHE (Grant Agreement No. 776186, see also: https://www.che-project.eu), and by the State and the University of Bremen.

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.

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).
References

/[ATBDv1, 2019]/ Reuter et al.: Algorithm Theoretical Basis Document Version 1 (ATBDv1) - Retrieval of XCO2 from the OCO-2 satellite using the Fast Atmospheric Trace Gas Retrieval (FOCAL), ESA GHG-CCI project, 2019


