



# **Product Quality Assessment Report (PQAR) – ANNEX D** for products XCO2\_EMMA, XCH4\_EMMA, XCO2\_OBS4MIPS, XCH4\_OBS4MIPS (v4.3, 01/2003-06/2020)

## **C3S\_312b\_Lot2\_DLR – Atmosphere**

Issued by: Maximilian Reuter, University of Bremen,

Institute of Environmental Physics (IUP)

Date: 18/02/2021

Ref: C3S\_D312b\_Lot2.2.3.2-v3.0\_PQAR-GHG\_ANNEX-D\_v5.0

Official reference number service contract: 2018/C3S\_312b\_Lot2\_DLR/SC1



*This document has been produced in the context of the Copernicus Climate Change Service (C3S).*

*The activities leading to these results have been contracted by the European Centre for Medium-Range Weather Forecasts, operator of C3S on behalf of the European Union (Delegation Agreement signed on 11/11/2014). All information in this document is provided "as is" and no guarantee or warranty is given that the information is fit for any particular purpose.*

*The user thereof uses the information at its sole risk and liability. For the avoidance of all doubts, the European Commission and the European Centre for Medium-Range Weather Forecasts has no liability in respect of this document, which is merely representing the authors view.*



## Contributors

**INSTITUTE OF ENVIRONMENTAL PHYSICS (IUP),  
UNIVERSITY OF BREMEN, BREMEN, GERMANY  
(IUP)**

M. Reuter

M. Buchwitz

O. Schneising-Weigel



## 1 Table of Contents

<b>History of modifications</b>	<b>5</b>
<b>Related documents</b>	<b>6</b>
<b>Acronyms</b>	<b>7</b>
<b>General definitions</b>	<b>11</b>
<b>Scope of document</b>	<b>12</b>
<b>Executive summary</b>	<b>13</b>
<b>Product validation methodology</b>	<b>14</b>
<b>Validation Results</b>	<b>16</b>
<b>1.1 XCO<sub>2</sub>_EMMA</b>	<b>16</b>
1.1.1 Validation	16
1.1.2 Summary	25
<b>1.2 XCH<sub>4</sub>_EMMA</b>	<b>26</b>
1.2.1 Validation	26
1.2.2 Summary	33
<b>1.3 XCO<sub>2</sub>_OBS4MIPS</b>	<b>34</b>
<b>1.4 XCH<sub>4</sub>_OBS4MIPS</b>	<b>34</b>
<b>References</b>	<b>35</b>



## History of modifications

Version	Date	Description of modification	Chapters / Sections
1.1	20-October-2017	New document for data set CDR1 (2003-2016)	All
2.0	4-October-2018	Update for CDR2 (2003-2017)	All
3.0	12-August-2019	Update for CDR3 (2003-2018)	All
3.1	03-November-2019	Update after review by Assimila: Correction of typos and figure captions moved to above figures	All
4.0	18-August-2020	Update for CDR4 (2003-2019)	All
5.0	18-February-2021	Update for CDR5 (01/2003-06/2020)	All



## Related documents

Reference ID	Document
D1	<p>Main PQAR:</p> <p><b>Buchwitz, M., et al.:</b> Product Quality Assessment Report (PQAR) – Main document for Greenhouse Gas (GHG: CO<sub>2</sub> &amp; CH<sub>4</sub>) data set CDR 5 (0.2003-06.2020), project C3S_312b_Lot2_DLR – Atmosphere, v5.0, 2021.</p> <p><b>Important Note:</b></p> <p><i>This document is an ANNEX to the Main PQAR document and contains the quality assessment results of the data provider.</i></p> <p><b><i>For the final overall quality assessment results of the data products described in this document see the Main PQAR document.</i></b></p>
D2	<p><b>Reuter, M., et al.:</b> Algorithm Theoretical Basis Document (ATBD) – ANNEX D for products XCO2_EMMA, XCH4_EMMA, XCO2_OBS4MIPS, XCH4_OBS4MIPS (v4.3, 01/2003-06/2020) , project C3S_312b_Lot2_DLR – Atmosphere, v5.0, 2021.</p>



## Acronyms

Acronym	Definition
AIRS	Atmospheric Infrared Sounder
AMSU	Advanced Microwave Sounding Unit
ATBD	Algorithm Theoretical Basis Document
BESD	Bremen optimal ESTimation DOAS
CAR	Climate Assessment Report
C3S	Copernicus Climate Change Service
CCDAS	Carbon Cycle Data Assimilation System
CCI	Climate Change Initiative
CDR	Climate Data Record
CDS	(Copernicus) Climate Data Store
CMUG	Climate Modelling User Group (of ESA's CCI)
CRG	Climate Research Group
D/B	Data base
DOAS	Differential Optical Absorption Spectroscopy
EC	European Commission
ECMWF	European Centre for Medium Range Weather Forecasting
ECV	Essential Climate Variable
EMMA	Ensemble Median Algorithm
ENVISAT	Environmental Satellite (of ESA)
EO	Earth Observation
ESA	European Space Agency
EU	European Union
EUMETSAT	European Organisation for the Exploitation of Meteorological Satellites
FCDR	Fundamental Climate Data Record
FoM	Figure of Merit
FP	Full Physics retrieval method



FTIR	Fourier Transform InfraRed
FTS	Fourier Transform Spectrometer
GCOS	Global Climate Observing System
GEO	Group on Earth Observation
GEOSS	Global Earth Observation System of Systems
GHG	GreenHouse Gas
GOME	Global Ozone Monitoring Experiment
GMES	Global Monitoring for Environment and Security
GOSAT	Greenhouse Gases Observing Satellite
IASI	Infrared Atmospheric Sounding Interferometer
IMAP-DOAS (or IMAP)	Iterative Maximum A posteriori DOAS
IPCC	International Panel in Climate Change
IUP	Institute of Environmental Physics (IUP) of the University of Bremen, Germany
JAXA	Japan Aerospace Exploration Agency
JCGM	Joint Committee for Guides in Metrology
L1	Level 1
L2	Level 2
L3	Level 3
L4	Level 4
LMD	Laboratoire de Météorologie Dynamique
MACC	Monitoring Atmospheric Composition and Climate, EU GMES project
NA	Not applicable
NASA	National Aeronautics and Space Administration
NetCDF	Network Common Data Format
NDACC	Network for the Detection of Atmospheric Composition Change
NIES	National Institute for Environmental Studies
NIR	Near Infra Red
NLIS	LMD/CNRS <i>neural</i> network mid/upper tropospheric CO <sub>2</sub> and CH <sub>4</sub> retrieval algorithm
NOAA	National Oceanic and Atmospheric Administration





Obs4MIPs	Observations for Climate Model Intercomparisons
OCO	Orbiting Carbon Observatory
OE	Optimal Estimation
PBL	Planetary Boundary Layer
ppb	Parts per billion
ppm	Parts per million
PR	(light path) PROxy retrieval method
PVIR	Product Validation and Intercomparison Report
QA	Quality Assurance
QC	Quality Control
REQ	Requirement
RMS	Root-Mean-Square
RTM	Radiative transfer model
SCIAMACHY	SCanning Imaging Absorption spectroMeter for Atmospheric ChartographY
SCIATRAN	SCIAMACHY radiative transfer model
SRON	SRON Netherlands Institute for Space Research
SWIR	Short Wava Infra Red
TANSO	Thermal And Near infrared Sensor for carbon Observation
TANSO-FTS	Fourier Transform Spectrometer on GOSAT
TBC	To be confirmed
TBD	To be defined / to be determined
TCCON	Total Carbon Column Observing Network
TIR	Thermal Infra Red
TR	Target Requirements
TRD	Target Requirements Document
WFM-DOAS (or WFMD)	Weighting Function Modified DOAS
UoL	University of Leicester, United Kingdom
URD	User Requirements Document
WMO	World Meteorological Organization



---

Y2Y	Year-to-year (bias variability)
-----	---------------------------------



## General definitions

Table 1 lists some general definitions relevant for this document.

**Table 1:** General definitions.

Item	Definition
XCO <sub>2</sub>	Column-averaged dry-air mixing ratios (mole fractions) of CO <sub>2</sub>
XCH <sub>4</sub>	Column-averaged dry-air mixing ratios (mole fractions) of CH <sub>4</sub>
L1	Level 1 satellite data product: geolocated radiance (spectra)
L2	Level 2 satellite-derived data product: Here: XCO <sub>2</sub> and XCH <sub>4</sub> information for each ground-pixel
L3	Level 3 satellite-derived data product: Here: Gridded XCO <sub>2</sub> and XCH <sub>4</sub> information, e.g., 5°x5°, monthly
L4	Level 4 satellite-derived data product: Here: Surface fluxes (emission and/or uptake) of CO <sub>2</sub> and CH <sub>4</sub>



## Scope of document

This document is a Product Quality Assessment Report (PQAR) for the Copernicus Climate Change Service (C3S, <https://climate.copernicus.eu/>) greenhouse gas (GHG) component as covered by project C3S\_312b\_Lot2.

Within this project satellite-derived atmospheric carbon dioxide (CO<sub>2</sub>) and methane (CH<sub>4</sub>) Essential Climate Variable (ECV) data products will be generated and delivered to ECMWF for inclusion into the Copernicus Climate Data Store (CDS) from which users can access these data products and the corresponding documentation.

The GHG satellite-derived data products are:

- Column-averaged dry-air mixing ratios (mole fractions) of CO<sub>2</sub> and CH<sub>4</sub>, denoted XCO<sub>2</sub> (in parts per million, ppm) and XCH<sub>4</sub> (in parts per billion, ppb), respectively.
- Mid/upper tropospheric mixing ratios of CO<sub>2</sub> (in ppm) and CH<sub>4</sub> (in ppb).

This document describes the validation and quality assessment of the C3S products XCO<sub>2</sub>\_EMMA, XCH<sub>4</sub>\_EMMA, XCO<sub>2</sub>\_OBS4MIPS and XCH<sub>4</sub>\_OBS4MIPS.

These products are merged multi-sensor XCO<sub>2</sub> and XCH<sub>4</sub> Level 2 and Level 3 products generated using algorithms developed at University of Bremen, Germany.

For an overview of these merged Level 2 data products XCO<sub>2</sub>\_EMMA and XCH<sub>4</sub>\_EMMA and of these merged Level 3 data products XCO<sub>2</sub>\_OBS4MIPS and XCH<sub>4</sub>\_OBS4MIPS see also *Reuter et al., 2020*.



## Executive summary

This Product Quality Assessment Report (PQAR) describes the validation of the EMMA v4.3 CO<sub>2</sub> and EMMA v4.3 CH<sub>4</sub> products (in the following also referred to as XCO<sub>2</sub>\_EMMA and XCH<sub>4</sub>\_EMMA) with TCCON ground based measurements. Originally, the EMMA algorithm (v1.3) was described and validated in the publication of *Reuter et al. (2013)*. More recently, *Reuter et al. (2020)* described the latest EMMA CO<sub>2</sub> and CH<sub>4</sub> validation and developments. These publications are the blueprint for this PQAR and several of the shown figures are updated versions of figures shown in them. EMMA is composed of an ensemble of individual SCIAMACHY, GOSAT, and OCO-2 L2 algorithms and this document also contributes to the inter-comparison of the contributing algorithms.

For XCO<sub>2</sub> we find that the individual algorithms have a single measurement precision in the range of 1.27 ppm (OCO-2 NASA) to 2.09 ppm (GOSAT RemotTeC and NIES operational). EMMA has a single measurement precision of 1.59 ppm. EMMA's combined regional and seasonal biases (0.41 ppm) are on the lower end of the range of the individual algorithms. The found linear drifts are small and not significant, i.e., the trend is smaller than twice its uncertainty. The linear drift found for EMMA's XCO<sub>2</sub> is  $-0.03 \pm 0.03$  ppm/a. The year-to-year stability is in the range of 0.21 ppm/a (OCO-2 NASA) and 0.41 ppm/a (GOSAT UoL-FP). EMMA's year-to-year stability is 0.31 ppm/a.

For XCH<sub>4</sub> we find that the individual algorithms have a single measurement precision in the range of 12.90 ppb (UoL-FP) to 13.85 ppb (RemoTeC-PR) except for WFMD which has a single measurement precision of 99.30 ppb. EMMA has a single measurement precision of 13.42 ppb (excluding the period till 04/2010). EMMA's combined regional and seasonal biases (3.62 ppb) are at the lower end of the range of the individual algorithms (3.56 ppb for RemoTeC-PR to 14.08 ppb for WFMD). The found linear drifts are usually small and not significant, i.e., the trend is smaller than twice its uncertainty. The linear drift found for EMMA's XCH<sub>4</sub> is  $-0.26 \pm 0.40$  ppb/a. The year-to-year stability is in the range of 1.3 ppb/a (UoL-PR) and 7.9 ppb/a (WFMD). EMMA's year-to-year stability of 1.6 ppb/a is at the lower end of the range.

The TCCON-validation of the XCO<sub>2</sub>\_OBS4MIPS and XCH<sub>4</sub>\_OBS4MIPSLlevel 3 products is based on comparisons of monthly mean data and is described in the main PQAR document (D1).

The validation of the Level 3 product XCO<sub>2</sub>\_OBS4MIPS can be summarized as follows: The overall monthly mean uncertainty is 1.2 ppm and the mean bias is -0.11 ppm. Relative systematic errors, i.e., spatial and temporal biases amount to  $0.7 \pm 0.6$  ppm. The computed linear drift of  $-0.07 \pm 0.20$  ppm is small and not significant.

The validation of the Level 3 product XCH<sub>4</sub>\_OBS4MIPS can be summarized as follows: The overall monthly mean uncertainty is 8.8 ppb and the mean bias is -3.3 ppb. Relative systematic errors, i.e., spatial and temporal biases amount to  $5 \pm 6$  ppb. The computed linear drift of  $0.1 \pm 1.0$  ppb is small and not significant.



## Product validation methodology

As described in D2, EMMA v4.3 CO<sub>2</sub> and CH<sub>4</sub> make use of the following satellite data products: SCIAMACHY BESD v02.01.02 (*Reuter et al., 2016*), GOSAT ACOS v9r (*O'Dell et al., 2012*), GOSAT FOCAL v1.0 (*Noël et al., 2020*), GOSAT RemoTeC v2.3.8 (*Detmers et al., 2017a*), GOSAT UoL-FP v7.3 (*Boesch and Anand, 2017*), GOSAT NIES v02.9xbc (*Yoshida et al., 2013*), GOSAT NIES PPDF-S v02.xx (*Bril et al., 2012*), OCO-2 NASA v10.2 (*Kiel et al., 2019*), and OCO-2 FOCAL v09 (*Reuter et al., 2017a,b*) for XCO<sub>2</sub> and WFMD v4.0 (*Schneising et al., 2016*), RemoTeC-FP v2.3.8 (*Detmers et al., 2017a*), RemoTeC-PR v2.3.9 (*Detmers et al., 2017b*), UoL-FP v7.3 (*Boesch and Anand, 2017*), UoL-PR v9.0 (*Boesch and Anand, 2017*), NIES v02.9xbc (*Yoshida et al., 2011*), and PPDF-S v02.xx (*Bril et al., 2012*) for XCH<sub>4</sub>.

The EMMA CO<sub>2</sub> and CH<sub>4</sub> data products are validated with TCCON (using version GGG2014) measurements in a similar way as done by *Reuter et al. 2011*. 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.

For each TCCON site with more than 250 co-locations and covering a time period of at least one year, the performance statistics number of co-locations, station bias, seasonal bias, linear drift, and single measurement precision are calculated. The validation period ranges from 01/2009 to 06/2020 for CO<sub>2</sub> and 04/2010 to 06/2020 for CH<sub>4</sub>.

The main validation results are computed by fitting the following bias model to the difference between the satellite retrievals and the TCCON measurements at each TCCON site.

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

Here,  $\Delta X$  represents the difference satellite minus TCCON,  $\varepsilon$  the residual, and  $a_{0-3}$  the free fit parameters. Specifically,  $a_0$  represents the average site bias (i.e., the regional bias),  $a_1$  the linear drift and  $a_2$  the amplitude of the seasonal bias at a TCCON site. The seasonal bias is computed from the standard deviation of the seasonal term and the single measurement precision from the standard deviation of the residual.

Based on the per station statistics, the following summarizing statistics are calculated: Total number of co-locations used for validation, (quadratic) average single measurement precision, station-to-station bias (standard deviation of the station biases), average seasonal bias, 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.

Additionally, a measure for the year-to-year stability is computed: For each TCCON site, the residual difference (satellite - TCCON) which is not explained by station bias, seasonal bias, and/or linear drift is derived by subtracting the fit of the trend model  $\Delta X$  from the satellite minus TCCON difference. These time series are smoothed by a running average of 365 days. Only days with more than 10 co-locations contributing to the running average of at least 5 TCCON sites are further considered. At these days, the station-to-station average is calculated. 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.

Due to the relatively large uncertainty, we do not compute 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 compute 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 and standard deviation.

As EMMA is constructed from an ensemble of individual L2 algorithms, the purpose of this document is not only product validation but also inter-comparison. Therefore, all individual algorithms contributing to EMMA are validated in the exact same way as EMMA.

Additionally, this document shows assessments of temporal and spatial bias patterns based on  $10^{\circ} \times 10^{\circ}$  monthly gridded level 3 data sets (not to be confused with the XCO<sub>2</sub>\_OBS4MIPS and XCH<sub>4</sub>\_OBS4MIPS data sets).

We calculate the fraction of potential outliers according to unrealistically large spatial gradients ( $>3\text{ppm}/10^{\circ}$  for XCO<sub>2</sub> and  $>20\text{ppb}/10^{\circ}$  for XCH<sub>4</sub>), unrealistically large deviations from CarbonTracker (*Jacobson et al., 2020*) CT2019 ( $>3\text{ppm}$  for XCO<sub>2</sub> and  $>20\text{ppb}/10^{\circ}$  for XCH<sub>4</sub>), and larger deviations from EMMA ( $>2.5\text{ppm}$  for XCO<sub>2</sub> and  $>10\text{ppb}/10^{\circ}$  for XCH<sub>4</sub>).

We also compare the north/south (N/S) gradient of each month with CT2019 and TCCON by averaging all northern and southern hemispheric grid boxes (using the same sampling). However, it shall be noted that the statistics in comparing to TCCON are less robust because only few grid boxes include TCCON stations.

Additionally, we compare the seasonal (peak-to-peak) amplitude of each grid box with CT2019 and TCCON by calculating the difference between annual maximum and minimum. We consider only those grid boxes with at least six valid months and use the same sampling. Again, the TCCON statistics are probably not very robust because they rely on few grid boxes with seasonal cycles.

The TCCON-validation of the XCO<sub>2</sub>\_OBS4MIPS and XCH<sub>4</sub>\_OBS4MIPS level 3 products is based on comparisons of monthly mean data and is described in the main PQAR document (D1)



## Validation Results

### 1.1 XCO<sub>2</sub>\_EMMA

#### 1.1.1 Validation

**Figure 1** shows all co-located EMMA and TCCON retrievals used for validation. Additionally, it includes all co-locations of the individual algorithms contributing to EMMA. The overall statistics per contributing algorithm are summarized in **Table 2**. **Table 3** shows the validation summary specifically for EMMA v4.3 CO<sub>2</sub>, i.e., the XCO<sub>2</sub>\_EMMA product. The results are valid for the time periods covered by the individual algorithms but earliest starting in 2009, respectively.

The individual algorithms have a single measurement precision in the range of 1.27 ppm (OCO-2 NASA) – 2.09 ppm (GOSAT RemoTeC and NIES operational). EMMA has a single measurement precision of 1.59 ppm. EMMA's combined regional and seasonal biases (0.41 ppm) are on the lower end of the range of the individual algorithms (0.38 ppm for OCO-2 NASA to 0.88 ppm for GOSAT PPDF-S).

**Figure 2** (left) shows the anomaly of station biases of the used algorithms. One can see that most satellite retrievals have a high bias of about 0.3 ppm – 1.0 ppm at the Sodankylä, Bremen, Karlsruhe, Orleans, and Garmisch-Partenkirchen TCCON sites and low bias of similar magnitude at Park Falls, Lamont and the southern hemispheric sites Darwin and Wollongong. This feature considerably contributes to the algorithms station-to-station bias statistics. Currently, it is unclear whether this discrepancy comes from the satellite retrievals or the TCCON.

The drift analysis in **Figure 2** (right) shows more or less small negative trends (typically below -0.1 ppm/a) for most algorithms at the sites Darwin and Wollongong and positive trends of the same magnitude for most algorithms at Orleans and Karlsruhe. This is a bit surprising because some of the sites are located in similar latitude bands.

**Figure 3** shows the smoothed average residual difference (satellite - TCCON) which is not explained by station bias, seasonal bias, and/or linear drift. The year-to-year stability computed from the variability of the average is in the range of 0.17 ppm/a (OCO-2 NASA) and 0.41 ppm/a (UoL-FP). EMMA's year-to-year stability is 0.31 ppm/a.

Analyses of gridded L3 data show that all algorithms reproduce large scale features well, however, there are still differences of a few ppm when looking into the details (**Figure 4**).

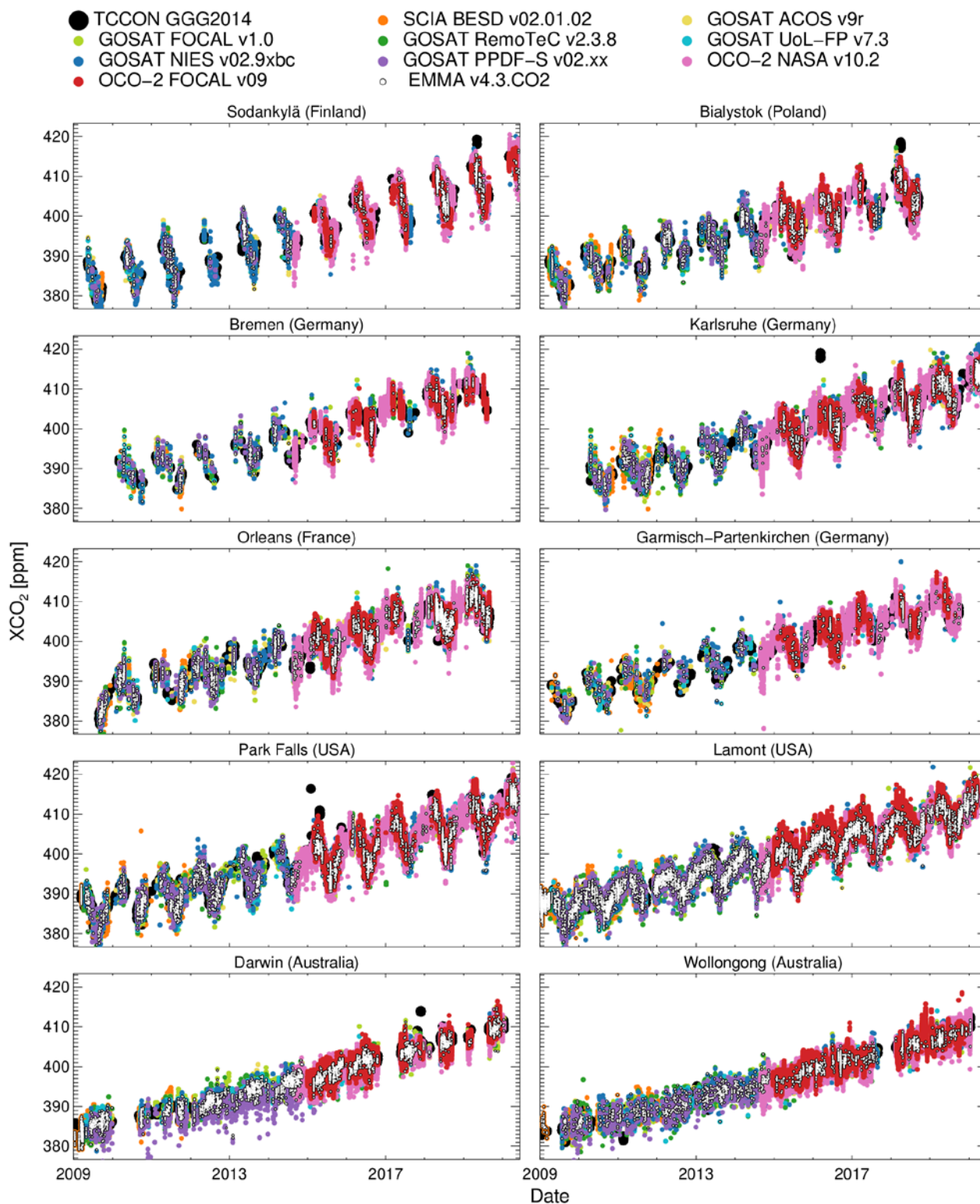
The satellite retrieved seasonal amplitudes are in similarly good agreement with TCCON and CarbonTracker (**Figure 5**, top left).

Comparison of the north/south gradients show also similar performances when comparing against CarbonTracker and TCCON. However, this should not be over interpreted because TCCON contributes only to few grid boxes especially on the southern hemisphere (**Figure 5**, top right).





**Figure 1:** Validation of individual XCO<sub>2</sub> algorithms and EMMA v4.3 CO<sub>2</sub> with TCCON GGG2014.

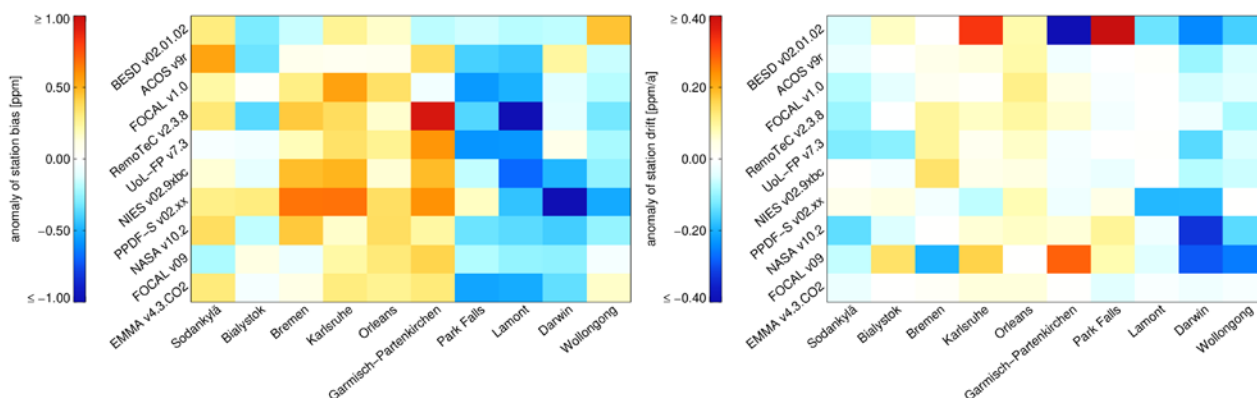




**Table 2:** Summarizing XCO<sub>2</sub> validation statistics for all TCCON sites that have been used for the validation. Listed are the number of co-locations (#), average single measurement precision, regional and seasonal accuracy, linear trend, year-to-year stability, and the probability that the accuracy and stability TR are met.

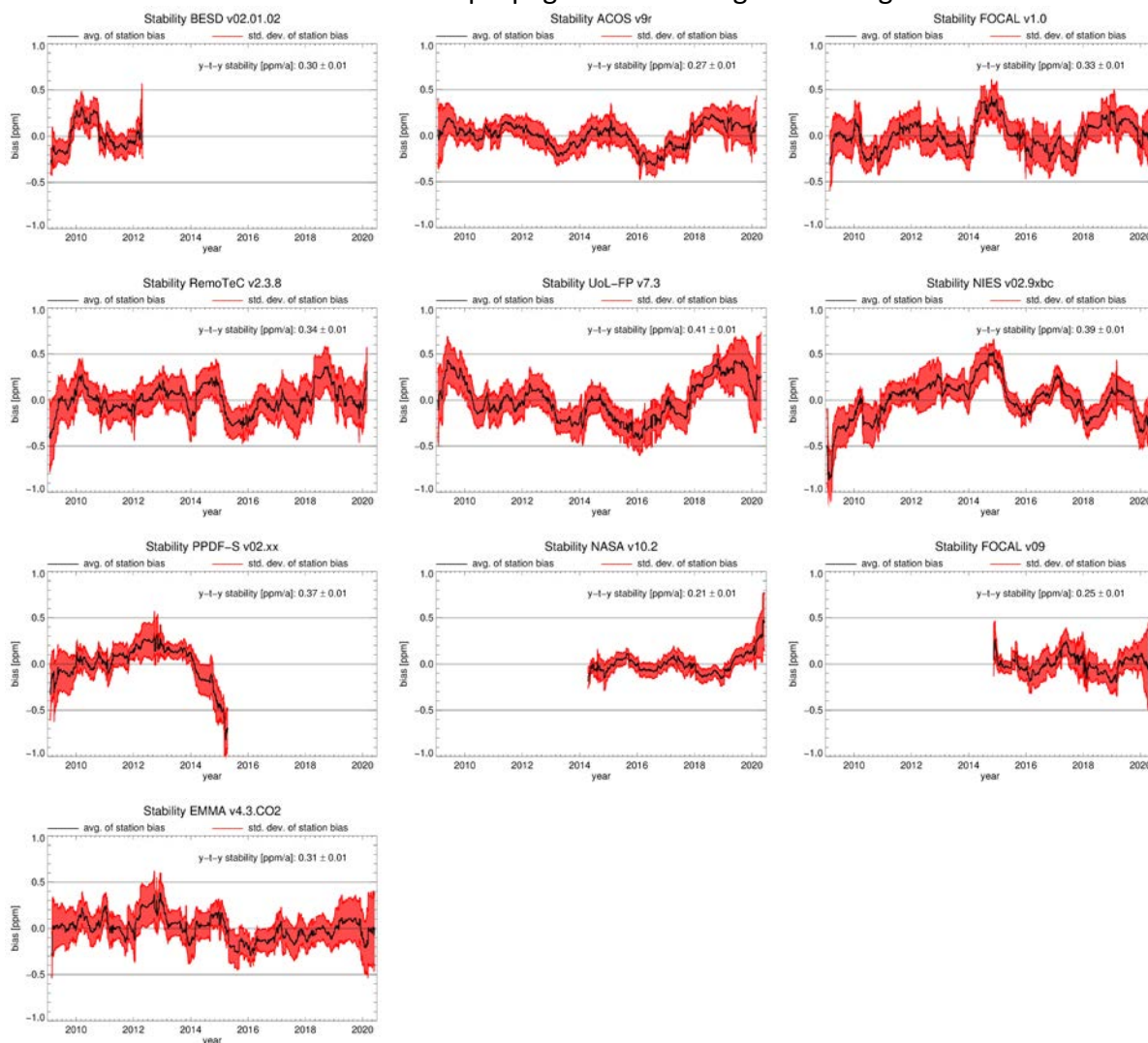
Algorithm	#	Precision [ppm]	Accuracy [ppm]		Stability [ppm/a]		Probability that TR is met [%]	
			Regional	Seasonal	Trend	Year2Year	Accuracy	Stability
SCIAMACHY BESD v02.01.02	18611	1.79	0.27	0.36	0.17±0.26	0.30	-	-
GOSAT ACOS v9r	19707	1.65	0.33	0.21	-0.00±0.05	0.27	-	-
GOSAT FOCAL v1.0	15412	1.90	0.36	0.19	-0.04±0.06	0.33	-	-
GOSAT RemoTeC v2.3.8	14707	2.09	0.56	0.31	-0.00±0.07	0.34	53	98
GOSAT UoL-FP v7.3	14680	1.80	0.38	0.39	-0.06±0.08	0.41	64	97
GOSAT NIES v02.9xbc	18359	2.09	0.41	0.29	-0.01±0.06	0.39	-	-
GOSAT PPDF-S v02.xx	8478	1.89	0.84	0.25	-0.17±0.09	0.37	-	-
OCO-2 NASA v10.2	1823829	1.27	0.34	0.18	-0.00±0.13	0.21	-	-
OCO-2 FOCAL v09	396111	1.45	0.25	0.36	-0.03±0.19	0.25	-	-
EMMA v4.3 CO2	34623	1.59	0.35	0.23	-0.03±0.03	0.31	76	98

**Figure 2:** Anomaly of station biases (left) and station drift (right).





**Figure 3:** Stability analyses for EMMA and the contributing individual algorithms. 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.



In terms of the frequency of potential outliers and standard deviation of the difference to CarbonTracker and TCCON, EMMA lies in the midfield of the algorithms (**Figure 5**, bottom).

ACOS usually provides the largest part of the relative data weight of the GOSAT algorithms in EMMA (**Figure 6**). The relative data rate drastically increases in 2015 when both OCO-2 algorithms provide data.

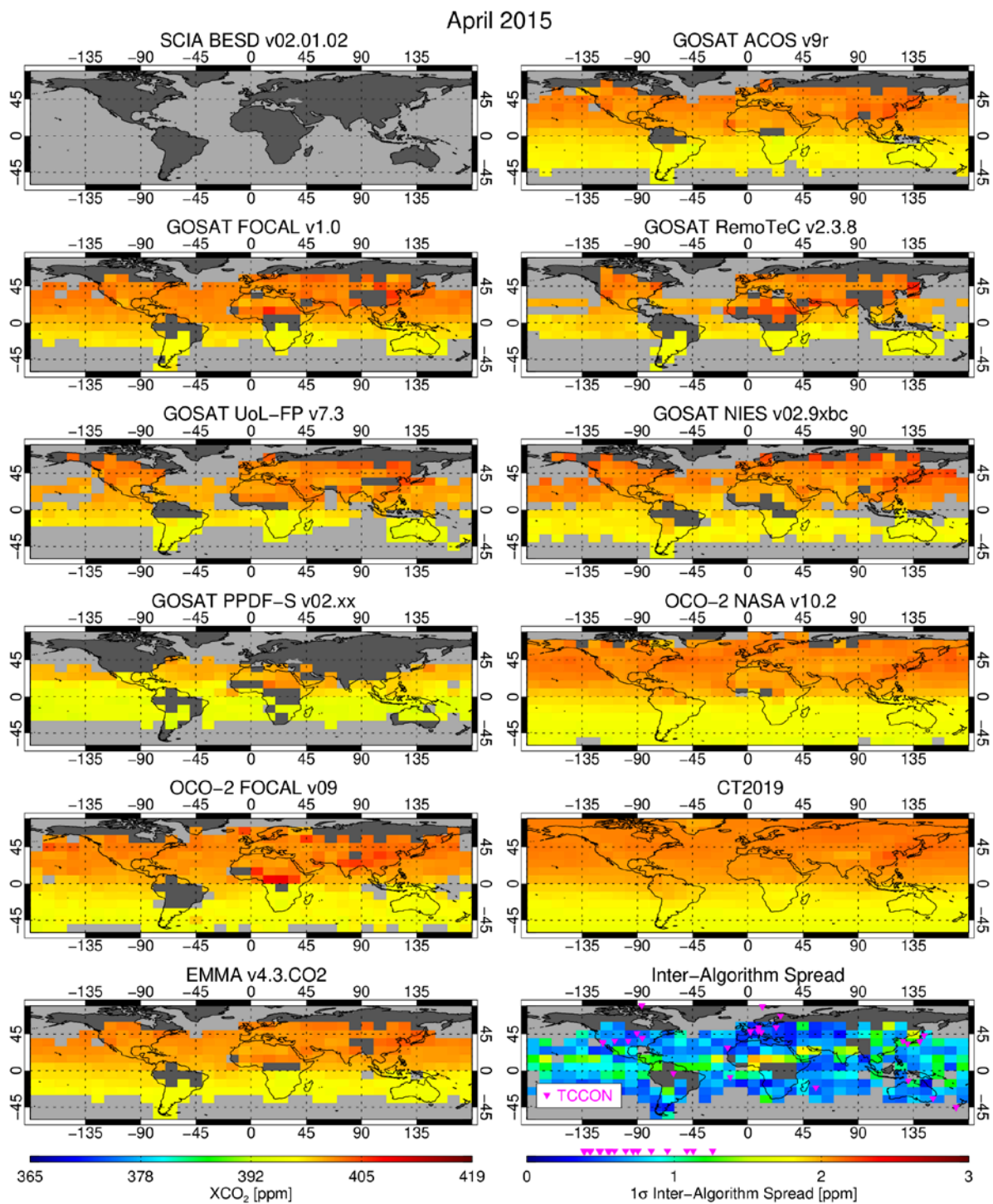


The average inter-algorithm spread has values most times between 0.4 ppm and 1.6 ppm and is typically below 1.0 ppm (**Figure 7**, left). The largest inter-algorithm spreads are observed in the tropics, Asia, and in high latitudes. Only a small fraction of the inter-algorithm spread can be explained with differences expected due to measurement noise so that most of the differences can be considered systematic. Only in high latitudes and at some coast-lines measurement noise is expected to explain a significant part of the inter-algorithm spread (**Figure 7**, right).

It is interesting to note that the average inter-algorithm spread usually reduced with every new EMMA version (always including the most recent algorithm versions, **Figure 8**). This means that EMMA observed kind of a convergence among the individual algorithms. It is not entirely clear where this convergence is coming from and many effects may contribute to the explanation: algorithms are improved and bugs are removed but algorithms may also become more similar by using the same input data (e.g., spectroscopy, elevation model). However, EMMA v3.1 did not follow this trend because the average inter-algorithm spread increased slightly. Most probably, this was caused by adding the not bias corrected operational NIES product and NIES' PPDF-S product. Additionally, the EMMA time period enhanced so that small drifts in the data sets, which are not corrected by EMMA's overall offset correction, can contribute to a larger extend.



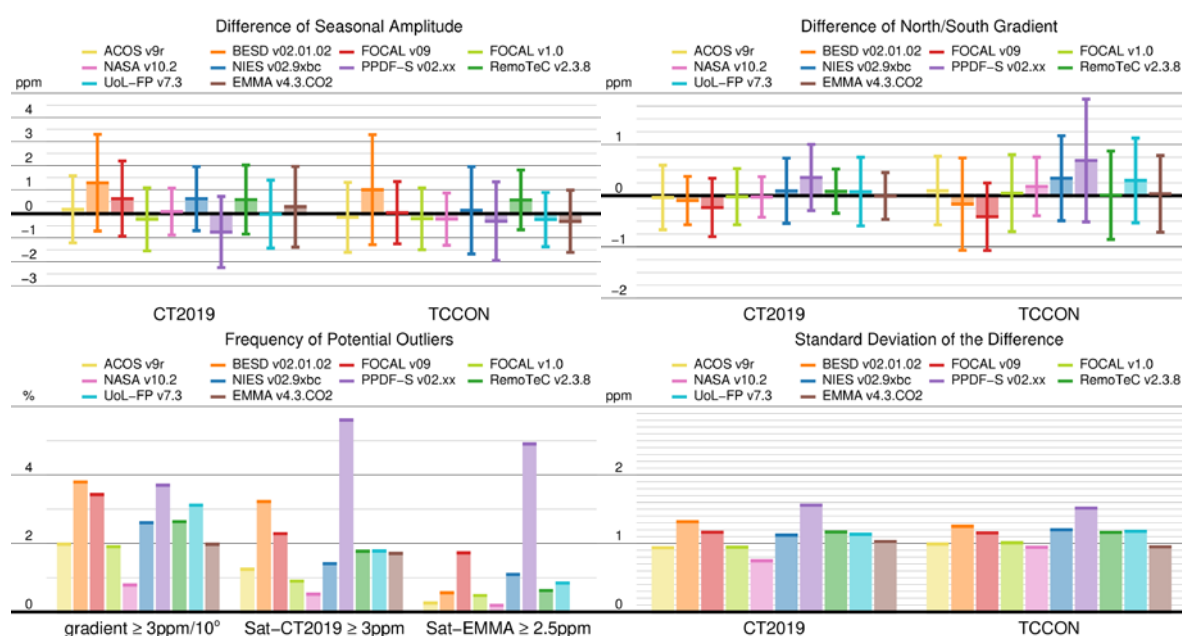
**Figure 4:** Monthly gridded XCO<sub>2</sub> averages and inter-algorithm spread at the example of April 2015 for EMMA v4.3 CO<sub>2</sub>.





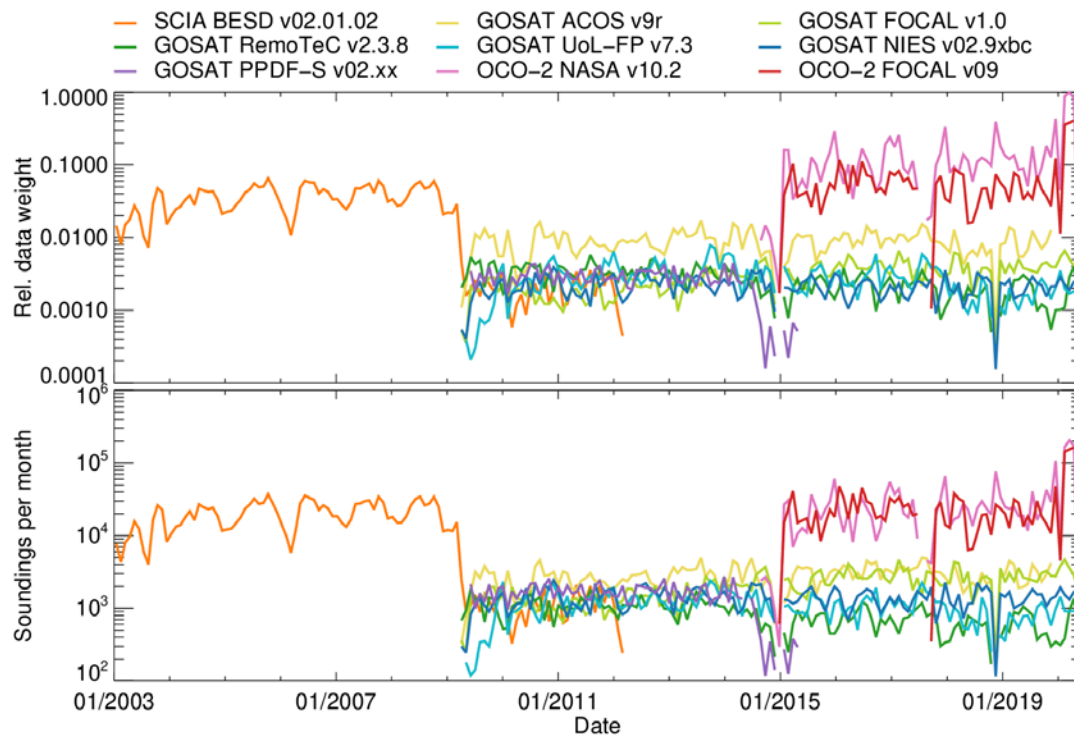


**Figure 5:** Top left: Difference of seasonal cycle amplitude of all individual algorithms as well as EMMA compared with CarbonTracker v2017 and TCCON GGG2014. Top right: Difference of north/south gradient of all individual algorithms as well as EMMA compared with CarbonTracker v2017 and TCCON. Bottom left: Frequency of potential outliers estimated by large gradients, large differences to CarbonTracker, and large differences to EMMA. Bottom right: Standard deviation of the difference of all algorithms and EMMA to CarbonTracker and TCCON.

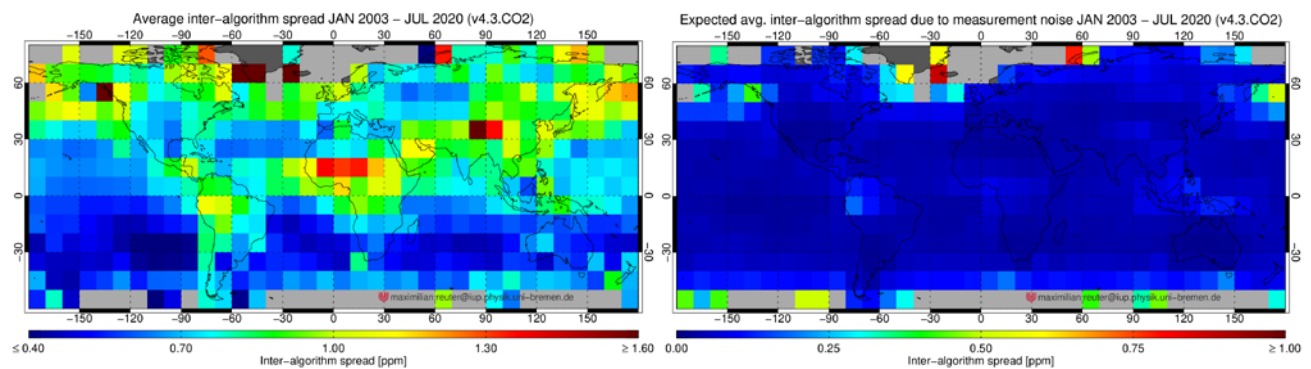




**Figure 6:** EMMA's normalized relative data weight proportional to  $\sum 1/\sigma_i^2$  (top) and number of soundings (bottom) per algorithm and month.

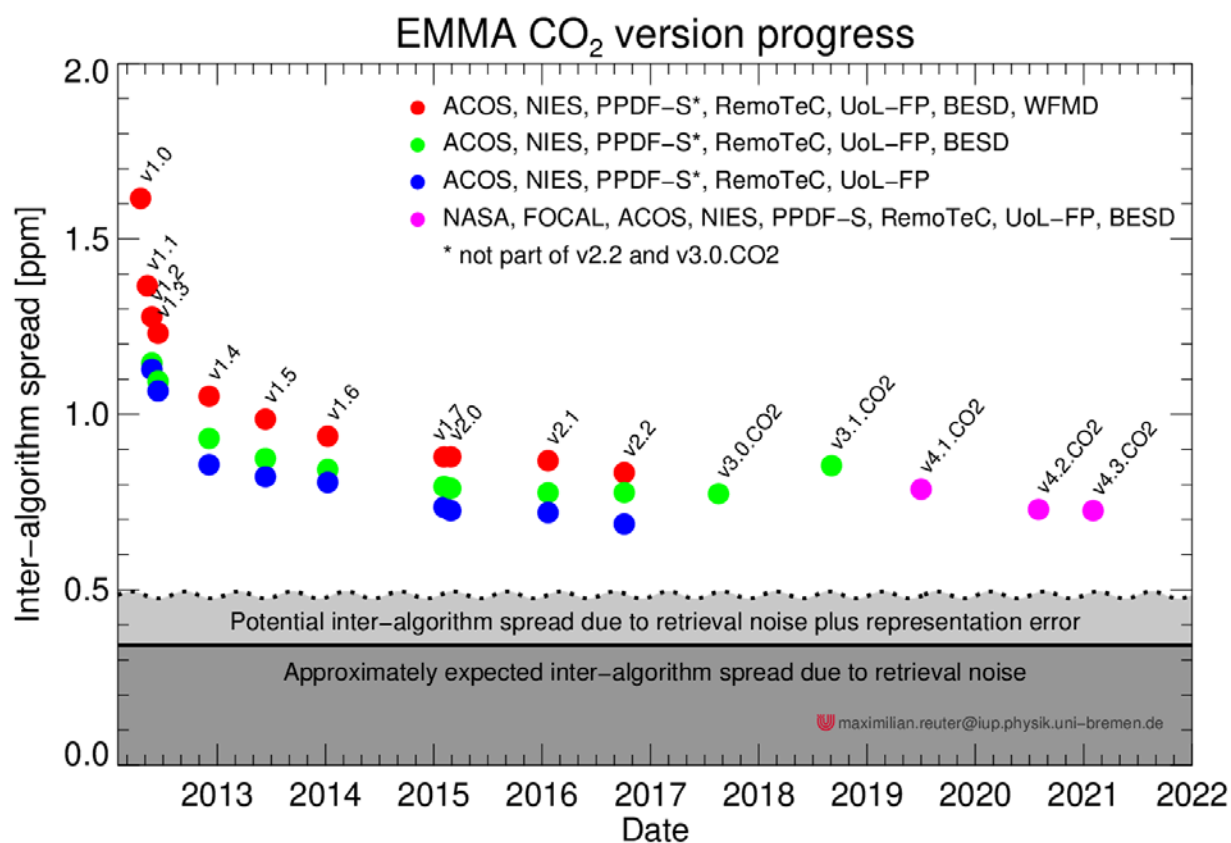


**Figure 7:** Average inter-algorithm scatter of monthly 10°x10° averages from January 2003 till July 2020 (left) and corresponding expected contribution of measurement noise (right).





**Figure 8:** Average inter-algorithm spread of all EMMA versions compared with the approximately expected contribution of retrieval noise and a rough estimate of the representation error.







### 1.1.2 Summary

The validation results are summarized in **Table 3**.

**Table 3:** Product Quality Summary Table for product XCO<sub>2</sub>\_EMMA.

Product Quality Summary Table for Product: XCO <sub>2</sub> _EMMA Level: 2, Version: 4.3, Time period covered: 01.2003 – 06.2020				
Parameter [unit]	Achieved performance	Requirement	TR	Comments
Single measurement precision (1-sigma) in [ppm]	1.59	< 8 (T) < 3 (B) < 1 (G)	-	-
Uncertainty ratio in [-]: Ratio reported uncertainty to standard deviation of satellite-TCCON difference	1.00	-	-	No requirement but value close to unity expected for a high quality data product.
Mean bias [ppm]	0.28	-	-	No requirement but value close to zero expected for a high quality data product.
Accuracy: Relative systematic error [ppm]	Spatial – spatiotemporal: 0.35 – 0.41	< 0.5	Probability that accuracy TR is met: 76%	-
Stability: Drift [ppm/year]	-0.03±0.03 (1-sigma)	< 0.5	Probability that stability TR is met: 98%	-
Stability: Year-to-year bias variability [ppm/year]	0.31 (1-sigma)	< 0.5	-	-



## 1.2 XCH<sub>4</sub>\_EMMA

### 1.2.1 Validation

**Figure 9** shows all co-located EMMA and TCCON retrievals used for validation. Additionally, it includes all co-locations of the individual algorithms contributing to EMMA. The overall statistics per contributing algorithm are summarized in **Table 4**. **Table 5** shows the validation summary specifically for EMMA v4.3 CH<sub>4</sub> i.e., the XCH<sub>4</sub>\_EMMA product. The results are valid for the time periods covered by the individual algorithms but starting in 04/2010 (i.e., after the WFMD contribution to EMMA ended), respectively. Note that this significantly limits the validation period for WFMD which ends at the beginning of 2012. Therefore, the validation results for WFMD are less robust and do not cover the years 2003-2005 when the SCIAMACHY instrument performed best in respect to XCH<sub>4</sub>.

The individual algorithms have a single measurement precision in the range of 12.90 ppb (UoL-FP) – 13.85 ppb (RemoTeC-PR) except for WFMD which has a single measurement precision of 99.30 ppb. EMMA has a single measurement precision of 13.42ppb which is similar as for the GOSAT algorithms. EMMA's combined regional and seasonal biases (3.62 ppb) are at the lower end of the range of the individual algorithms (3.56 ppb for RemoTeC-PR to 14.08 ppb for WFMD).

**Figure 10** (left) shows the anomaly of station biases of the used algorithms. One can see that most satellite retrievals have a high bias of typically 4 ppb at the Sodankylä and Park Falls TCCON sites and low bias of similar magnitude at the southern hemispheric sites Darwin and Wollongong. This feature is somewhat similar to the feature observed for XCO<sub>2</sub> and considerably contributes to the algorithms station-to-station bias statistics. Currently, it is unclear whether this discrepancy comes from the satellite retrievals or the TCCON.

The drift analysis in **Figure 10** (right) shows more or less small but consistent negative trends (typically below -0.5 ppb/a) at the sites Lamont and Park Falls, and Wollongong.

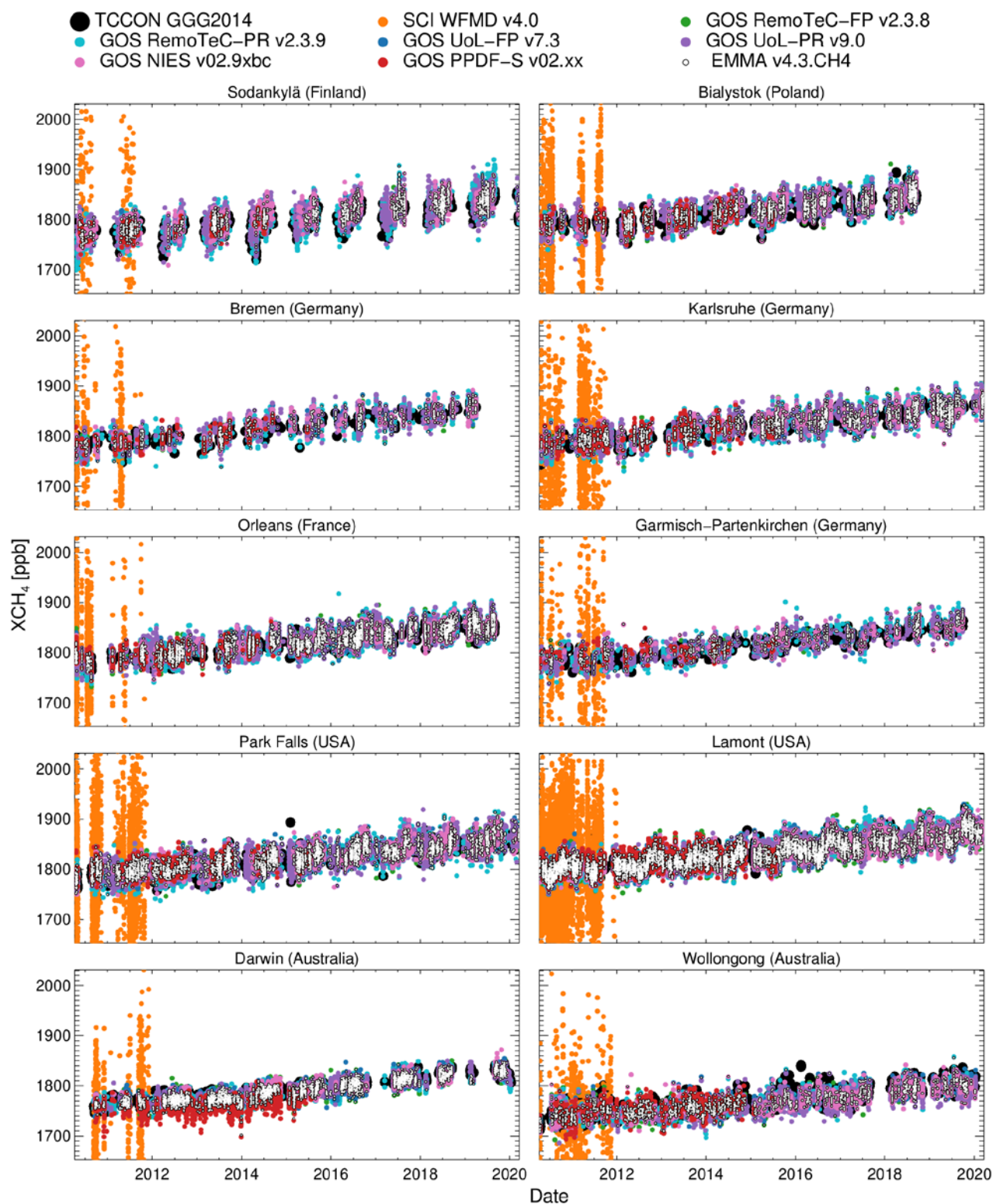
**Figure 11** shows the smoothed average residual difference (satellite - TCCON) which is not explained by station bias, seasonal bias, and/or linear drift. The year-to-year stability computed from the variability of the average is in the range of 1.3 ppb/a (UoL-PR) to 2.3 ppb/a (PPDF-S) for the GOSAT algorithms and 8.0 ppb/a for WFMD. EMMA's year-to-year stability of 1.9 ppb/a is at the lower end of the range of GOSAT algorithms.

Analyses of gridded L3 data show that all algorithms reproduce large scale features well, however, there are still differences of a some 10 ppb when looking into the details (**Figure 12**).

All algorithms see a slightly but consistently larger north/south gradient than TCCON (**Figure 13**, left). The same is true for the seasonal amplitude (**Figure 13**, left). In terms of the standard deviation of the difference to TCCON and the frequency of potential outliers, EMMA performs similarly as the best individual algorithms (**Figure 13**).



**Figure 9:** Validation of individual XCH<sub>4</sub> algorithms and EMMA v4.3 CH<sub>4</sub> with TCCON GGG2014.

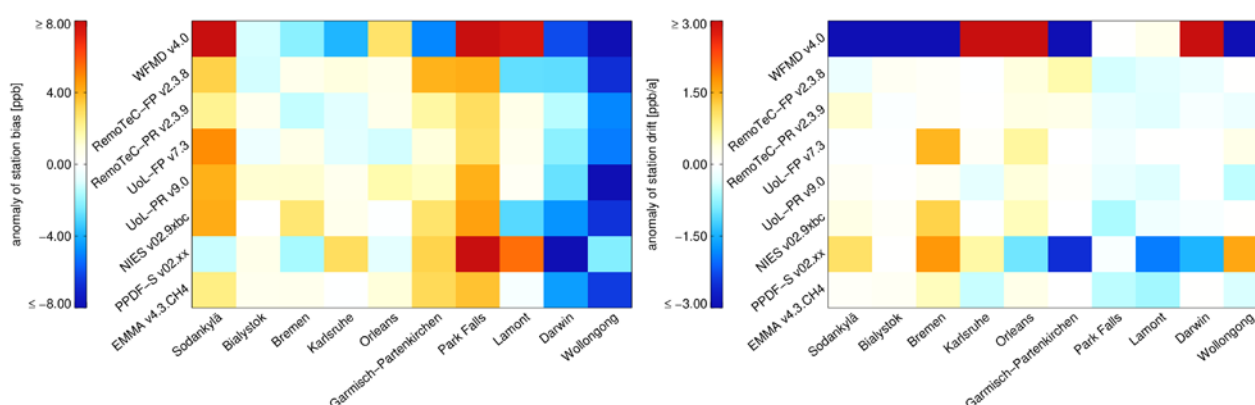




**Table 4:** Summarizing XCH<sub>4</sub> validation statistics for all TCCON sites that have been used for the validation. Listed are the number of co-locations (#), average single measurement precision, regional and seasonal accuracy, linear trend, year-to-year stability, and the probability that the accuracy and stability TR are met. Note, due to the short validation period of WFMD and the instrumental issues of SCIAMACHY during this period, the WFMD results are less robust and trend and probabilities for meeting the TRs have not been assessed.

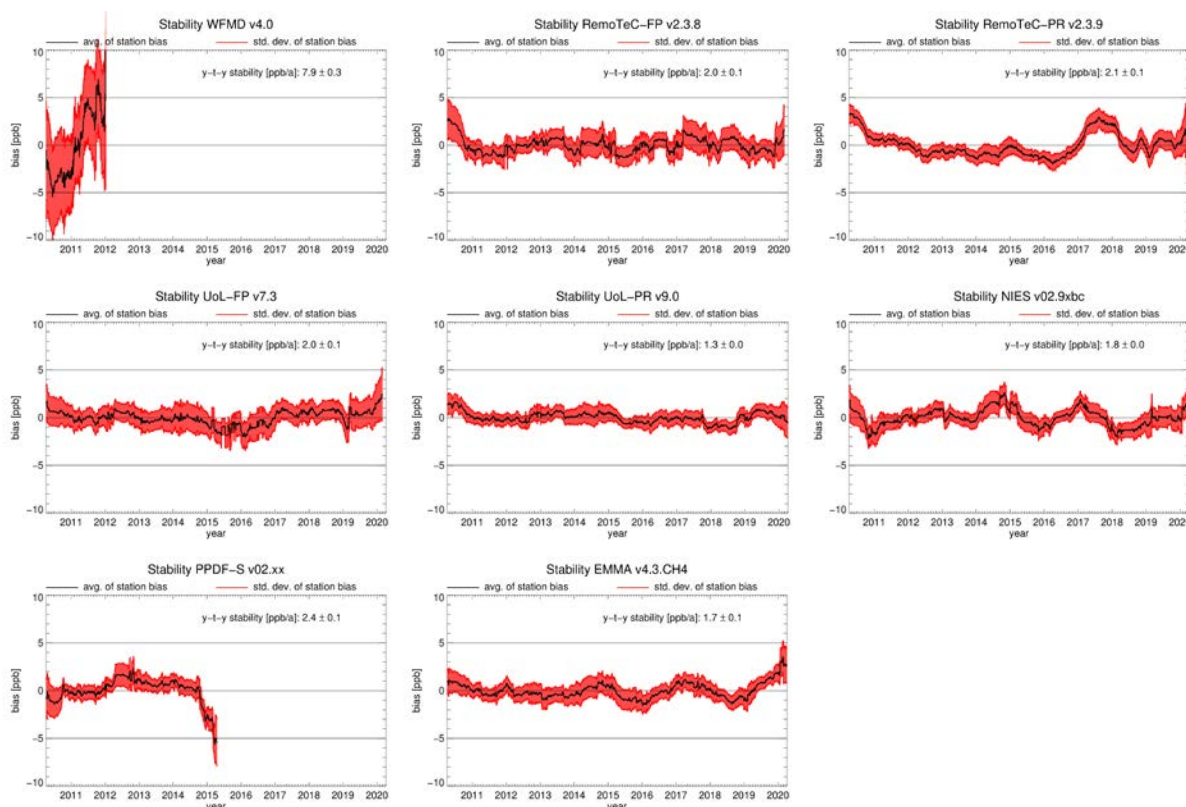
Algorithm	#	Precision [ppb]	Accuracy [ppb]		Stability [ppb/a]		Probability that TR is met [%]	
			Regional	Seasonal	Trend	Year2Year	Accuracy	Stability
WFMD v4.0	13258	99.30	8.25	11.41	-	7.9	-	-
RemoTeC-FP v2.3.8	13280	13.51	3.47	2.51	-0.18±0.33	2.0	90	99
RemoTeC-PR v2.3.9	42803	13.85	2.29	2.72	0.19±0.26	2.1	92	99
UoL-FP v7.3	13242	12.90	2.74	2.76	-0.65±0.50	2.0	92	98
UoL-PR v9.0	39602	13.35	4.48	2.20	-0.20±0.28	1.3	88	99
NIES v02.9xbc	17103	12.91	3.74	2.45	-0.45±0.50	1.8	-	-
PPDF-S v02.xx	7523	13.59	6.47	3.58	-1.53±1.49	2.4	-	-
EMMA v4.3 CH4	24794	13.42	3.17	1.75	-0.26±0.40	1.7	92	99

**Figure 10:** Anomaly of station biases (left) and station drift (right).





**Figure 11:** Stability analyses for EMMA and the contributing individual algorithms. 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.



Once the GOSAT algorithms kick in in 2009, the proxy algorithms RemoTeC-PR and UoL-PR clearly provide the largest part of the relative data weight in EMMA (**Figure 14**). Note also the drop in the relative data weight of WFMD end of 2005 due to instrument degradation.

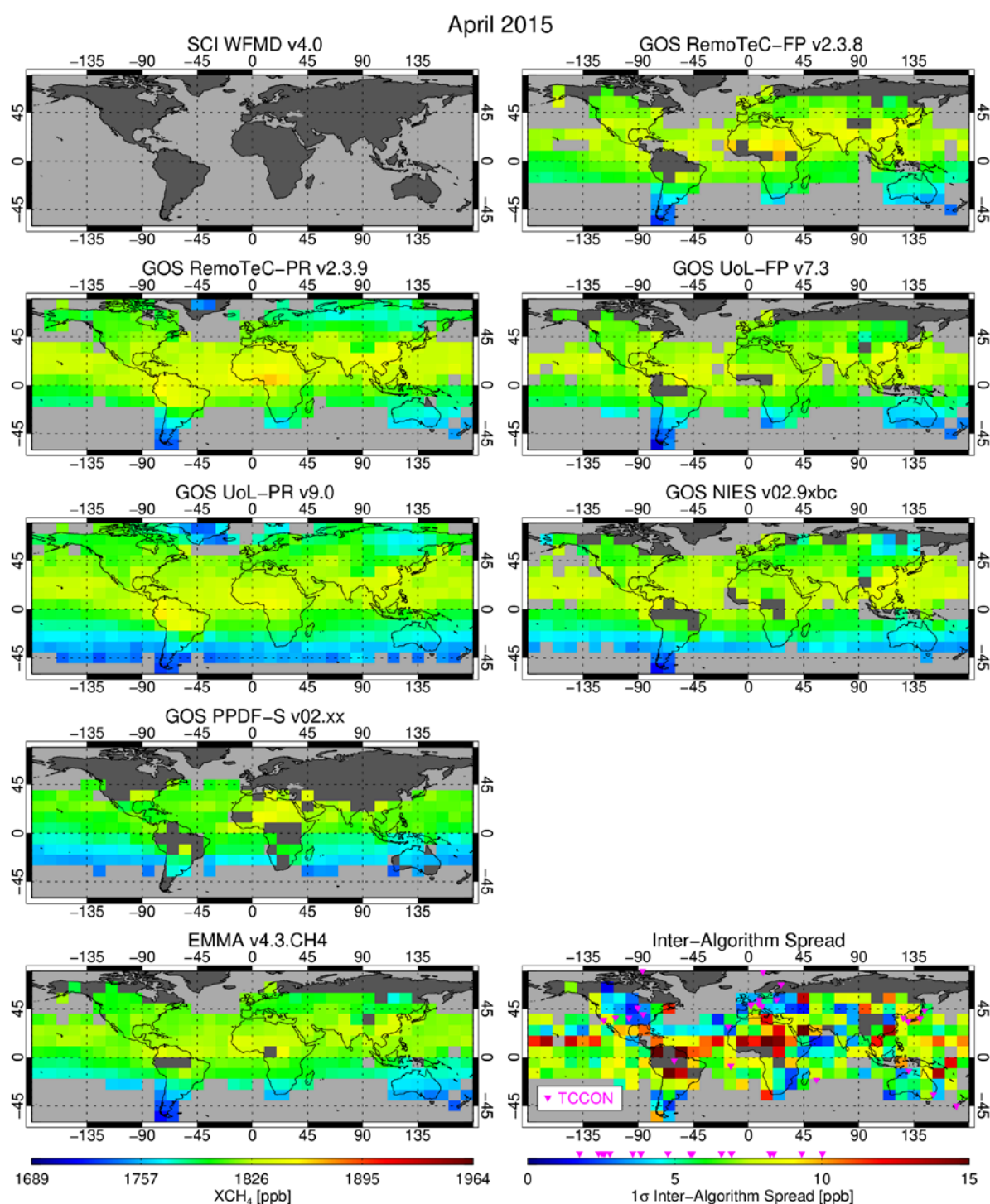
The average inter-algorithm spread has values mostly between 2 ppb and 12 ppb and is typically below 9 ppb (**Figure 15**, left). The largest inter-algorithm spreads are observed in the tropics, Asia, and in high latitudes. Only a small fraction of the inter-algorithm spread can be explained with differences expected due to measurement noise so that most of the differences can be considered systematic errors. Only in high latitudes and at some coast-lines measurement noise is expected to explain a significant part of the inter-algorithm spread (**Figure 15**, right).

The average inter-algorithm spread enhanced till EMMA v3.1 and slightly reduced since then (**Figure 16**). Part of the enhancement is most probably caused by adding the not bias corrected operational NIES product and NIES' PPDF-S product in v3.1. Additionally, the EMMA time period enhanced so that small drifts in the data sets, which are not corrected by EMMA's overall offset correction, can contribute to a larger extend.



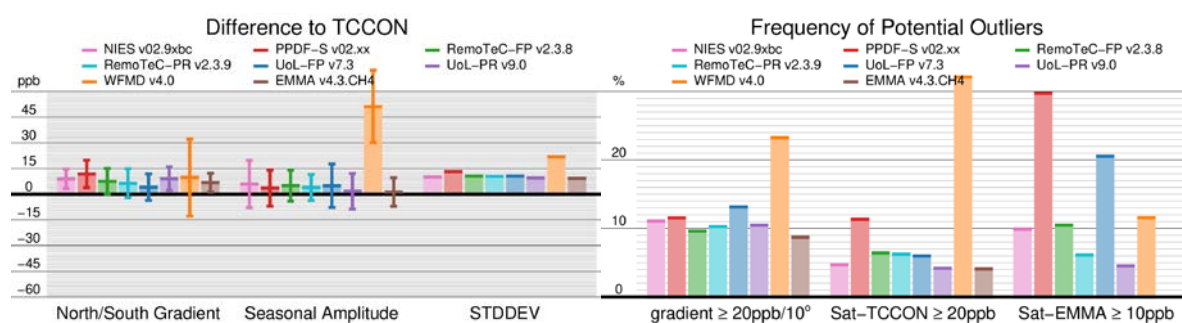


**Figure 12:** Monthly gridded XCH<sub>4</sub> averages and inter-algorithm spread at the example of April 2015 for EMMA v4.3 CH<sub>4</sub>.

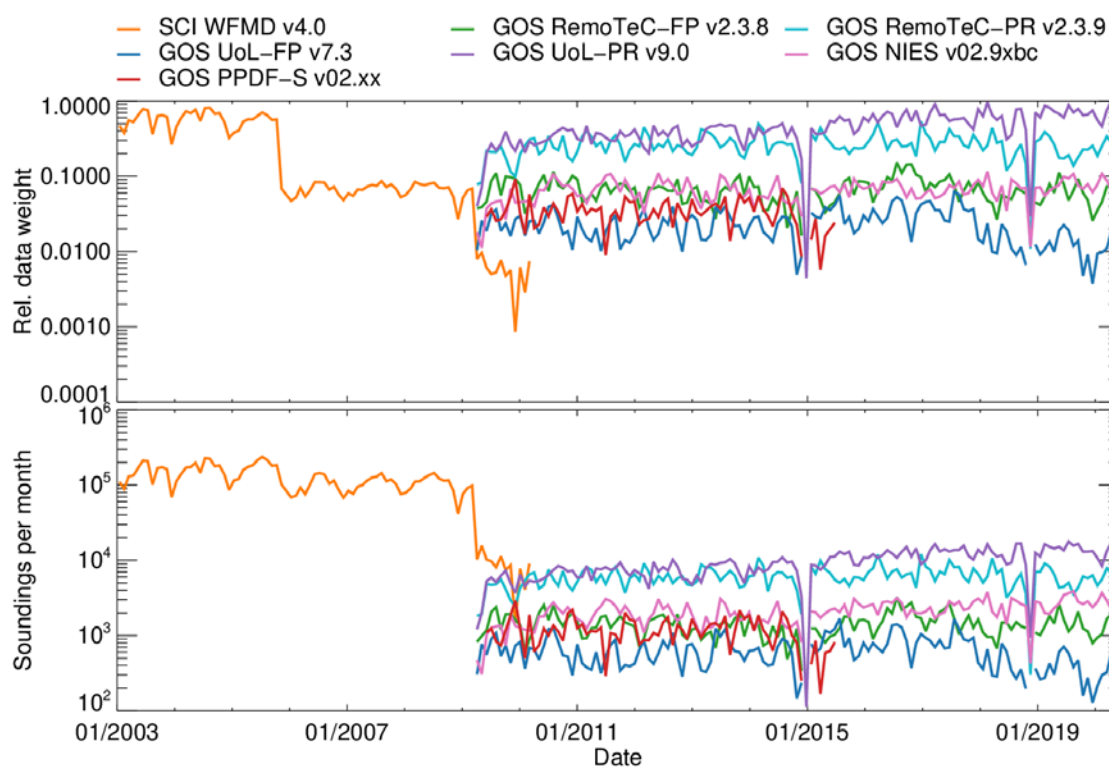




**Figure 13: Left:** Difference of the north/south gradient, difference of the seasonal cycle amplitude, and standard deviation of the difference in respect to TCCON for all individual algorithms as well as EMMA. **Right:** Frequency of potential outliers estimated by large gradients, large differences to TCCON, and large differences to EMMA.

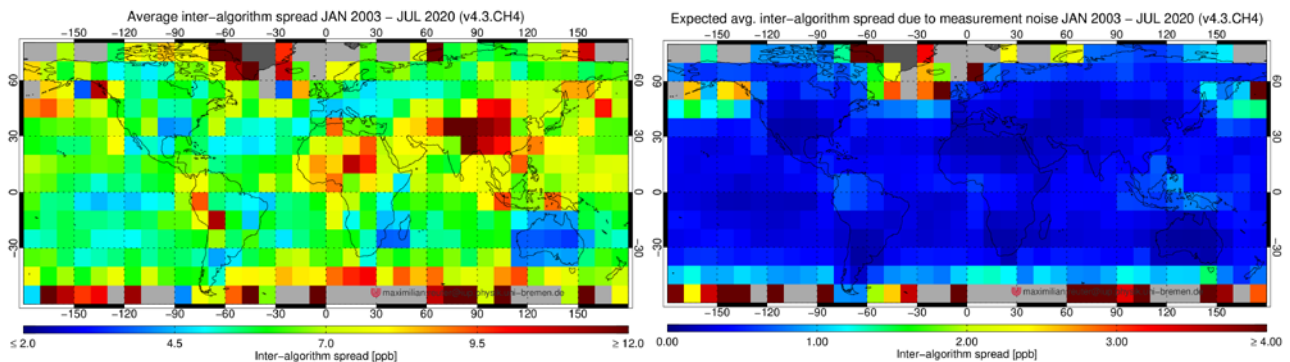


**Figure 14:** EMMA's normalized relative data weight proportional to  $\sum 1/\sigma_i^2$  (**top**) and number of soundings (**bottom**) per algorithm and month.

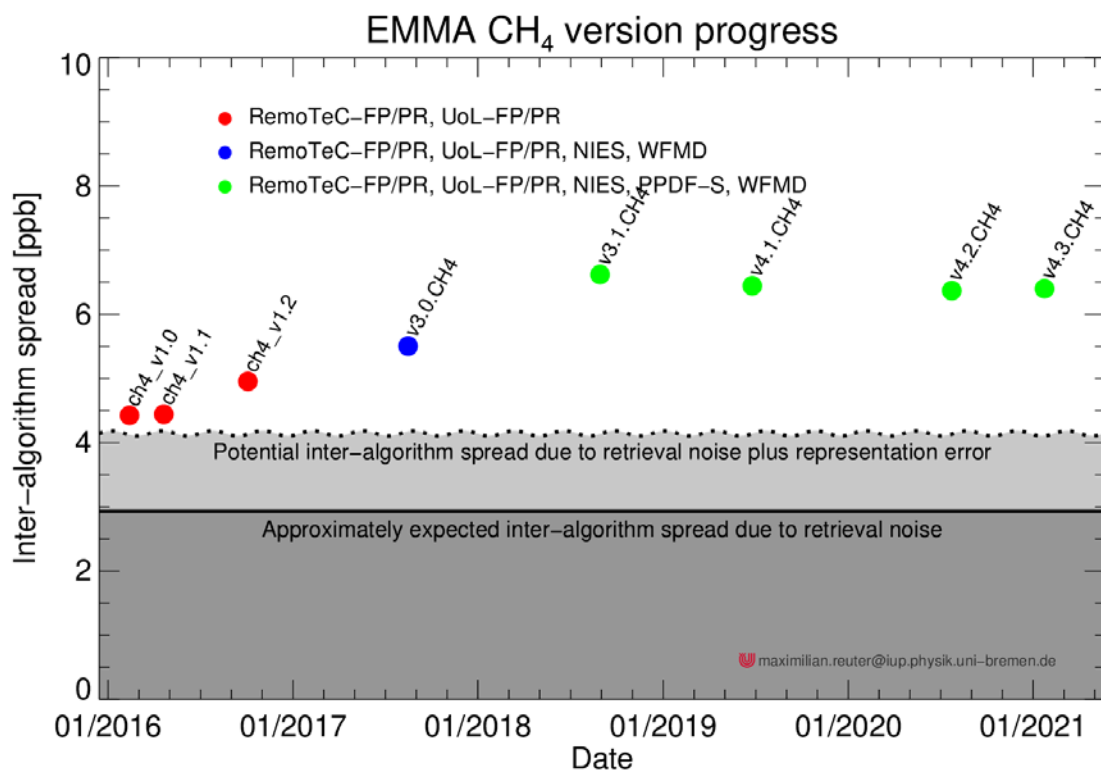




**Figure 15:** Average inter-algorithm scatter of monthly  $10^\circ \times 10^\circ$  averages from January 2003 till July 2020 (left) and corresponding expected contribution of measurement noise (right).



**Figure 16:** Average inter-algorithm spread of all EMMA versions compared with the approximately expected contribution of retrieval noise and a rough estimate of the representation error.







### 1.2.2 Summary

The validation results are summarized in **Table 5**.

**Table 5:** Product Quality Summary Table for product XCH4\_EMMA. Note that the achieved performance corresponds to the time period after 03/2010 when only GOSAT algorithms are part of EMMA.

Product Quality Summary Table for Product: XCH4_EMMA Level: 2, Version: 4.3, Time period covered: 01.2003 – 06.2020				
Parameter [unit]	Achieved performance	Requirement	TR	Comments
Single measurement precision (1-sigma) in [ppb]	13.42	< 34 (T) < 17 (B) < 9 (G)	-	
Uncertainty ratio) in [-]: Ratio reported uncertainty to standard deviation of satellite-TCCON difference	1.01	-	-	No requirement but value close to unity expected for a high quality data product.
Mean bias [ppb]	-7.10	-	-	No requirement but value close to zero expected for a high quality data product.
Accuracy: Relative systematic error [ppb]	Spatial – spatiotemporal: 3.17 – 3.62	< 10	Probability that accuracy TR is met: 92%	-
Stability: Linear bias trend [ppb/year]	-0.26±0.40 (1-sigma)	< 3	Probability that stability TR is met: 99%	-
Stability: Year-to-year bias variability [ppb/year]	1.7 (1-sigma)	< 3	-	



### 1.3 XCO<sub>2</sub>\_OBS4MIPS

The TCCON-validation of the XCO<sub>2</sub>\_OBS4MIPS Level 3 product is based on comparisons of monthly mean data and is described in the main PQAR document (D1).

The validation of Level 3 product XCO<sub>2</sub>\_OBS4MIPS can be summarized as follows: The overall monthly mean uncertainty is 1.2 ppm and the mean bias is -0.11 ppm. Relative systematic errors, i.e., spatial and temporal biases amount to  $0.7 \pm 0.6$  ppm. The computed linear drift of  $-0.07 \pm 0.20$  ppm is small and not significant.

### 1.4 XCH<sub>4</sub>\_OBS4MIPS

The TCCON-validation of the XCH<sub>4</sub>\_OBS4MIPS Level 3 product is based on comparisons of monthly mean data and is described in the main PQAR document (D1).

The validation of Level 3 product XCH<sub>4</sub>\_OBS4MIPS can be summarized as follows: The overall monthly mean uncertainty is 8.8 ppb and the mean bias is -3.3 ppb. Relative systematic errors, i.e., spatial and temporal biases amount to  $5 \pm 6$  ppb. The computed linear drift of  $0.1 \pm 1.0$  ppb is small and not significant.



## References

**Boesch and Anand, 2017:** H. Boesch and J. Anand, Algorithm Theoretical Basis Document (ATBD) – ANNEX A for products CO<sub>2</sub>\_GOS\_OCFP, CH<sub>4</sub>\_GOS\_OCFP & CH<sub>4</sub>\_GOS\_OCFR, Copernicus Climate Change Service (C3S) project on satellite-derived Essential Climate Variable (ECV) Greenhouse Gases (CO<sub>2</sub> and CH<sub>4</sub>) data products (project C3S\_312a\_Lot6), Version 1 (21/08/2017), 2017

**Bril et al., 2012:** Bril, A.; Oshchepkov, S.; Yokota, T. Application of a probability density function-based atmospheric light-scattering correction to carbon dioxide retrievals from GOSAT over-sea observations. *Remote Sensing of the Environment* 2012, 117, 301–306

**Buchwitz et al., 2021:** Buchwitz, M., et al., Product Quality Assessment Report (PQAR) – Main document for Greenhouse Gas (GHG: CO<sub>2</sub> & CH<sub>4</sub>) data set CDR 5 (01.2003-06.2020), project C3S\_312b\_Lot2\_DLR – Atmosphere, v5.0, 2021.

**Detmers, 2017a:** R. Detmers, Algorithm Theoretical Basis Document (ATBD) – ANNEX B for products CO<sub>2</sub>\_GOS\_SRFP & CH<sub>4</sub>\_GOS\_SRFP, Copernicus Climate Change Service (C3S) project on satellite-derived Essential Climate Variable (ECV) Greenhouse Gases (CO<sub>2</sub> and CH<sub>4</sub>) data products (project C3S\_312a\_Lot6), Version 1 (21/08/2017), 2017

**Detmers, 2017b:** R. Detmers, Algorithm Theoretical Basis Document (ATBD) – ANNEX C for product CH<sub>4</sub>\_GOS\_SRPR, Copernicus Climate Change Service (C3S) project on satellite-derived Essential Climate Variable (ECV) Greenhouse Gases (CO<sub>2</sub> and CH<sub>4</sub>) data products (project C3S\_312a\_Lot6), Version 1 (21/08/2017), 2017

**Jacobson et al., 2020:** Jacobson, A. R., Schuldt, K. N., Miller, J. B., Oda, T., Tans, P., Arlyn Andrews, Mund, J., Ott, L., Collatz, G. J., Aalto, T., Afshar, S., Aikin, K., Aoki, S., Apadula, F., Baier, B., Bergamaschi, P., Beyersdorf, A., Biraud, S. C., Bollenbacher, A., Bowling, D., Brailsford, G., Abshire, J. B., Chen, G., Huilin Chen, Lukasz Chmura, Sites Climadat, Colomb, A., Conil, S., Cox, A., Cristofanelli, P., Cuevas, E., Curcoll, R., Sloop, C. D., Davis, K., Wekker, S. D., Delmotte, M., DiGangi, J. P., Dlugokencky, E., Ehleringer, J., Elkins, J. W., Emmenegger, L., Fischer, M. L., Forster, G., Frumau, A., Galkowski, M., Gatti, L. V., Gloor, E., Griffis, T., Hammer, S., Haszpra, L., Hatakka, J., Heliasz, M., Hensen, A., Hermanssen, O., Hintsa, E., Holst, J., Jaffe, D., Karion, A., Kawa, S. R., Keeling, R., Keronen, P., Kolari, P., Kominkova, K., Kort, E., Krummel, P., Kubistin, D., Labuschagne, C., Langenfelds, R., Laurent, O., Laurila, T., Lauvaux, T., Law, B., Lee, J., Lehner, I., Leuenberger, M., Levin, I., Levula, J., Lin, J., Lindauer, M., Loh, Z., Lopez, M., Myhre, C. L., Machida, T., Mammarella, I., Manca, G., Manning, A., Manning, A., Marek, M. V., Marklund, P., Martin, M. Y., Matsueda, H., McKain, K., Meijer, H., Meinhardt, F., Miles, N., Miller, C. E., Mölder, M., Montzka, S., Moore, F., Josep-Anton Morgui, Morimoto, S., Munger, B., Jaroslaw Necki, Newman, S., Nichol, S., Niwa, Y., O'Doherty, S., Mikael Ottosson-Löfvenius, Paplawsky, B., Peischl, J., Peltola, O., Jean-Marc Pichon, Piper, S., Plass-Dölmer, C., Ramonet, M., Reyes-Sanchez, E., Richardson, S., Riris, H., Ryerson, T., Saito, K., Sargent, M., Sasakawa, M., Sawa, Y., Say, D., Scheeren, B., Schmidt, M., Schmidt, A., Schumacher, M., Shepson, P., Shook, M., Stanley, K., Steinbacher, M., Stephens, B., Sweeney, C., Thoning, K., Torn, M., Turnbull, J., Tørseth, K., Bulk, P. V. D., Laan-Luijkx, I. T. V. D., Dinter, D. V., Vermeulen, A., Viner, B., Vitkova, G., Walker, S., Weyrauch, D., Wofsy, S., Worthy, D., Dickon Young,



and Mirosław Zimnoch: CarbonTracker CT2019, <https://doi.org/10.25925/39M3-6069>, <https://www.esrl.noaa.gov/gmd/ccgg/carbontracker/CT2019/>, 2020

**Noël et al., 2020:** XCO<sub>2</sub> retrieval for GOSAT and GOSAT-2 based on the FOCAL algorithm, 2020 (in preparation)

**O'Dell et al., 2012:** O'Dell, C. W., Connor, B., Bösch, H., O'Brien, D., Frankenberg, C., Castano, R., Christi, M., Eldering, D., Fisher, B., Gunson, M., McDuffie, J., Miller, C. E., Natraj, V., Oyafo, F., Polonsky, I., Smyth, M., Taylor, T., Toon, G. C., Wennberg, P. O., and Wunch, D.: The ACOS CO<sub>2</sub> retrieval algorithm – Part 1: Description and validation against synthetic observations, *Atmos. Meas. Tech.*, 5, 99–121, doi:10.5194/amt-5-99-2012, 2012

**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 XCO<sub>2</sub> affected by erroneous surface pressure estimates, *Atmos. Meas. Tech.*, 12, 2241–2259, <https://doi.org/10.5194/amt-12-2241-2019>, 2019

**Reuter et al., 2011:** M. Reuter, H. Bovensmann, M. Buchwitz, J. P. Burrows, B. J. Connor, N. M. Deutscher, D. W. T. Griffith, J. Heymann, G. Keppel-Aleks, J. Messerschmidt, J. Notholt, C. Petri, J. Robinson, O. Schneising, V. Sherlock, V. Velasco, T. Warneke, P. O. Wennberg, D. Wunch: Retrieval of atmospheric CO<sub>2</sub> with enhanced accuracy and precision from SCIAMACHY: Validation with FTS measurements and comparison with model results. *Journal of Geophysical Research - Atmospheres*, 116, D04301, doi: 10.1029/2010JD015047, 2011

**Reuter et al., 2013:** M. Reuter, H. Bösch, H. Bovensmann, A. Bril, M. Buchwitz, A. Butz, J. P. Burrows, C. W. O'Dell, S. Guerlet, O. Hasekamp, J. Heymann, N. Kikuchi, S. Oshchepkov, R. Parker, S. Pfeifer, O. Schneising, T. Yokota, and Y. Yoshida: A joint effort to deliver satellite retrieved atmospheric CO<sub>2</sub> concentrations for surface flux inversions: the ensemble median algorithm EMMA. *Atmospheric Chemistry and Physics*, doi:10.5194/acp-13-1771-2013, 13, 1771-1780, 2013

**Reuter et al., 2017a:** M. Reuter, M. Buchwitz, O. Schneising, S. Noël, V. Rozanov, H. Bovensmann and 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<sub>2</sub> Retrieval Setup, *Remote Sensing*, 9(11), 1159; doi:10.3390/rs9111159, 2017

**Reuter et al., 2017b:** M. Reuter, M. Buchwitz, O. Schneising, S. Noël, H. Bovensmann and J. P. Burrows: A Fast Atmospheric Trace Gas Retrieval for Hyperspectral Instruments Approximating Multiple Scattering - Part 2: Application to XCO<sub>2</sub> Retrievals from OCO-2, *Remote Sensing*, 9(11), 1102; doi:10.3390/rs9111102, 2017

**Reuter et al., 2020:** M. Reuter, M. Buchwitz, O. Schneising, S. Noël, H. Bovensmann, J. P. Burrows, H. Bösch, 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. Velasco, 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.

**Yoshida et al., 2011:** Yoshida, Y., Ota, Y., Eguchi, N., Kikuchi, N., Nobuta, K., Tran, H., Morino, I., and Yokota, T.: Retrieval algorithm for CO<sub>2</sub> and CH<sub>4</sub> column abundances from short-wavelength infrared spectral observations by the Greenhouse gases observing satellite, Atmos. Meas. Tech., 4, 717–734, doi:10.5194/amt-4-717-2011, 2011

