ECMWF COPERNICUS REPORT

Copernicus Climate Change Service

Product Quality Assessment Report (PQAR) – ANNEX D for products XCO2_EMMA, XCH4_EMMA, XCO2_OBS4MIPS, XCH4_OBS4MIPS (v4.5, 01/2003-12/2022)

C3S2_312a_Lot2_DLR – Atmosphere

Issued by: Maximilian Reuter, University of Bremen,

Institute of Environmental Physics (IUP)

Date: 14/12/2023

Ref: C3S2_312a_Lot2_D-WP2_PQAR-2023-GHG_ANNEX-D_v7.2

Official reference number service contract: 2021/C3S2_312a_Lot2_DLR/SC1

PROGRAMME OF THE EUROPEAN UNION

IMPLEMENTED BY ECMWF

Contributors

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

M. Reuter

M. Buchwitz

O. Schneising-Weigel

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
6.0	4-August-2022	Update for CDR6 (01/2003-12/2021)	All
6.1	14-December-2022	Update after review (use of new template, several improvements at various places)	All
6.2	14-February-2023	Update after 2nd review. Several improvements at various places.	All
6.3	02-March-2023	Minor updates after 2nd review to generate clean version.	All
7.0	24-August-2023	Update for data set CDR7 (temporal coverage: 2003-2022)	All
7.1	18-October-2023	Update after review	All
7.2	17-November-2023	Minor update after review	All
7.2	14-December-2023	2 formatting issues corrected and some missing numbers added.	Pages 16, 35, 51 of version 17/11/2023.

List of datasets covered by this document

Deliverable ID	Product title	Product type (CDR, ICDR)	Version number	Delivery date
WP2-FDDP-GHG-v2	XCO2_EMMA, XCH4_EMMA, XCO2_OBS4MIPS, XCH4_OBS4MIPS	CDR 7	v4.5	31-Aug-2023

Related documents

Reference ID	Document
D1	Main PQAR: **Buchwitz, et al., 2023:** Product Quality Assessment Report (PQAR) – Main document for Greenhouse Gas (GHG: CO_2 & CH_4) data set CDR 7 (2003-2022), project C3S2_312a_Lot2_DLR – Atmosphere, 2023. ***Important Note:*** *This document is an ANNEX to the Main PQAR document and contains the quality assessment results of the data provider.* ***For the final overall quality assessment results of the data products described in this document see the Main PQAR document.***
D2	**Reuter, et al., 2023a:** Algorithm Theoretical Basis Document (ATBD) – ANNEX D for products XCO2_EMMA, XCH4_EMMA, XCO2_OBS4MIPS, XCH4_OBS4MIPS (v4.5, 01/2003-12/2022), project C3S2_312a_Lot2_DLR – Atmosphere, 2023.
D3	**TRD GAD GHG, 2021:** Buchwitz, M., Reuter, M., Schneising-Weigel, O., Aben, I., Wu, L., Hasekamp, O. P., Boesch, H., Di Noia, A., Crevoisier, C., Armante, R.: Target Requirement and Gap Analysis Document, Copernicus Climate Change Service (C3S) project on satellite-derived Essential Climate Variable (ECV) Greenhouse Gases (CO_2 and CH_4) data products, Version 3.1, 19-February-2021, pp. 81, 2021. Latest version: http://wdc.dlr.de/C3S_312b_Lot2/Documentation/GHG/C3S2_312a_Lot2_TRD-GAD_GHG_latest.pdf
D4	**Reuter, et al., 2023b:** Product User Guide and Specification (PUGS) – ANNEX D for products XCO2_EMMA, XCH4_EMMA, XCO2_OBS4MIPS, XCH4_OBS4MIPS (v4.5, 01/2003-12/2022) C3S_312a_Lot2_DLR – Atmosphere, 2023.

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 InfraRed
NLIS	LMD/CNRS *neuronal* network mid/upper tropospheric CO2 and CH4 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 InfraRed
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 InfraRed
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

Essential climate variable (ECV)

An ECV is a physical, chemical, or biological variable or a group of linked variables that critically contributes to the characterization of Earth’s climate.

Climate data record (CDR)

The US National Research Council (NRC) defines a CDR as a time series of measurements of sufficient length, consistency, and continuity to determine climate variability and change.

Fundamental climate data record (FCDR)

A fundamental climate data record (FCDR) is a CDR of calibrated and quality-controlled data designed to allow the generation of homogeneous products that are accurate and stable enough for climate monitoring.

Thematic climate data record (TCDR)

A thematic climate data record (TCDR) is a long time series of an essential climate variable (ECV).

Intermediate climate data record (ICDR)

An intermediate climate data record (ICDR) is a TCDR which undergoes regular and consistent updates, for example because it is being generated by a satellite sensor in operation.

Satellite data processing levels

The NASA Earth Observing System (EOS) distinguishes six processing levels of satellite data, ranging from Level 0 (L0) to Level 4 (L4) as follows.

L0	Unprocessed instrument data
L1A	Unprocessed instrument data alongside ancillary information
L1B	Data processed to sensor units (geo-located calibrated spectral radiance and solar irradiance)
L2	Derived geophysical variables (e.g., XCO_2) over one orbit
L3	Geophysical variables averaged in time and mapped on a global longitude/latitude horizontal grid
L4	Model output derived by assimilation of observations, or variables derived from multiple measurements (or both)

Additional definitions as relevant for this document:

Systematic error: component of measurement error that in replicate measurements remains constant or varies in a predictable manner

Note: “Systematic error” = “Absolute systematic error” (in contrast to “Relative systematic error” defined below).

For satellite GHG ECV products especially the “Relative systematic error” is important. The definition as used here is as follows:

Relative systematic error: Identical with “Systematic error” but after bias correction and without considering a possible “global offset” (overall mean bias). Reflects the importance of spatially and temporally correlated errors (“spatio-temporal biases”). Computed from standard deviations of spatial and temporal biases.

Bias: estimate of a systematic measurement error (JCGM, 2008).

Precision is the measure of reproducibility or repeatability of the measurement without reference to an international standard so that precision is a measure of the random and not the systematic error. Suitable averaging of the random error can improve the precision of the measurement but does not establish the systematic error of the observation (CMUG-RBD, 2010).

Note: Precision (as explained in TRD (D3)) is quantified with the standard deviation (1-sigma) of the error distribution.

Stability is a term often invoked with respect to long-term records when no absolute standard is available to quantitatively establish the systematic error - the bias defining the time-dependent (or instrument-dependent) difference between the observed quantity and the true value (CMUG-RBD, 2010).

Note: Stability requirements cover inter-annual error changes. If the change in the average bias from one year to another is larger than the defined values, the corresponding product does not meet the stability requirement.

Representativity is important when comparing with or assimilating in models. Measurements are typically averaged over different horizontal and vertical scales compared to model fields. If the measurements are smaller scale than the model it is important. The sampling strategy can also affect this term (CMUG-RBD, 2010).

Threshold requirement: The threshold is the limit at which the observation becomes ineffectual and is not of use for climate-related applications (CMUG-RBD, 2010).

Goal requirement: The goal is an ideal requirement above which further improvements are not necessary (CMUG-RBD, 2010).

Breakthrough requirement: The breakthrough is an intermediate level between the “threshold” and “goal“requirements, which - if achieved - would result in a significant improvement for the targeted application. The breakthrough level may be considered as an optimum, from a cost-benefit point of view when planning or designing observing systems (CMUG-RBD, 2010).

Horizontal resolution is the area over which one value of the variable is representative of (CMUG-RBD, 2010).

Vertical resolution is the height over which one value of the variable is representative of. Only used for profile data (CMUG-RBD, 2010).

Observing Cycle (or **Revisit Time)** is the temporal frequency at which the measurements are required (CMUG-RBD, 2010).

Table of Contents

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 the project C3S2_312a_Lot2.

Within this project satellite-derived atmospheric carbon dioxide (CO_2) and methane (CH_4) 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_2 and CH_4, denoted XCO_2 (in parts per million, ppm) and XCH_4 (in parts per billion, ppb), respectively.
- Mid/upper tropospheric mixing ratios of CO_2 (in ppm) and CH_4 (in ppb).

This document describes the validation and quality assessment of the C3S products XCO2_EMMA (v4.5), XCH4_EMMA (v4.5), XCO2_OBS4MIPS (v4.5) and XCH4_OBS4MIPS (v4.5).

These products are merged multi-sensor XCO_2 and XCH_4 level 2 and level 3 products generated using algorithms developed at the University of Bremen, Germany (see D2 and Reuter et al. (2013, 2020)).

Executive summary

This Product Quality Assessment Report (PQAR) describes the validation of the EMMA v4.5 CO_2 and EMMA v4.5 CH_4 products (in the following also referred to as XCO2_EMMA and XCH4_EMMA) with TCCON (Total Carbon Column Observing Network; Wunch et al., 2011) 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_2 and CH_4 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, GOSAT-2, and OCO-2 L2 algorithms and this document also contributes to the inter-comparison of the contributing algorithms.

For XCO_2 we find that the individual algorithms have a single measurement precision in the range of 1.29 ppm (OCO-2 NASA) to 2.91 ppm (GOSAT-2 RemotTeC). EMMA has a single measurement precision of 1.78 ppm. EMMA's combined regional and seasonal biases (0.47 ppm) lie at the lower end of the range of the individual algorithms (0.37 ppm for GOSAT ACOS and GOSAT FOCAL to 0.98 for GOSAT-2 NIES). The found linear drifts are small and almost always not significant, i.e., the trend is smaller than twice its uncertainty. The linear drift found for EMMA's XCO_2 is 0.02±0.09 ppm/a. The year-to-year stability is in the range of 0.16 ppm/a (OCO-2 NASA) and 0.71 ppm/a (GOSAT RemoTeC). EMMA's year-to-year stability is 0.36 ppm/a.

For XCH_4 we find that the individual algorithms have a single measurement precision in the range of 11.51 ppb (GOSAT-2 FOCAL-FP) to 17.80 ppb (GOSAT-2 RemoTeC-PR) except for WFMD which has a single measurement precision of 104.27 ppb. EMMA has a single measurement precision of 13.51 ppb (excluding the WFMD including period till 04/2010). EMMA's combined regional and seasonal biases (2.59 ppb) are at the lower end of the range of the individual algorithms (2.00 ppb for GOSAT FOCAL FP to 20.39 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_4 is 0.11±0.22 ppb/a. The year-to-year stability is in the range of 1.4 ppb/a (GOSAT-2 RemoTeC-PR) and 3.5 ppb/a (GOSAT RemoTeC-PR). EMMA's year-to-year stability of 1.8 ppb/a is at the lower end of that range. The linear drift and year-to-year stability of WFND could not be reliably determined due to the relatively short time span and few station data within the time period.

The TCCON-validation of the XCO2_OBS4MIPS and XCH4_OBS4MIPSLlevel 3 products is based on comparisons of monthly mean data and is described in the main PQAR document (D1).

The validation of Level 3 product XCO2_OBS4MIPS can be summarized as follows:

The overall monthly mean uncertainty is 1 ppm and the mean bias is 0.39 ppm. Relative systematic error, i.e., the spatio-temporal bias, is 0.5±0.6 ppm (1-sigma). The computed linear drift of 0.09±0.23 ppm (1-sigma) is small and not significant. The probability that the 0.5 ppm accuracy requirement is met is 66%. The probability that the 0.5 ppm/year stability requirement is met is 97%. Overall, this product has therefore reasonable accuracy and high stability.

The validation of Level 3 product XCH4_OBS4MIPS can be summarized as follows:

The overall monthly mean uncertainty is 8.1 ppb and the mean bias is -0.55 ppb. Relative systematic error, i.e., the spatio-temporal bias, is 4.7±6 ppb (1-sigma). The computed linear drift of 0.68±1.1 ppb (1-sigma) is small and not significant. The probability that the 10 ppb accuracy requirement is met is 89%. The probability that the 3 ppb/year stability requirement is met is 98%. Overall, this product has therefore very good accuracy and high stability.

1. Product validation methodology

As described in D2, EMMA v4.5 CO_2 and CH_4 make use of the satellite data products listed in Table 1 and Table 2, respectively.

The EMMA CO_2 and CH_4 data products are validated with TCCON (Wunch et al., 2011; D1) version GGG2020 measurements in a similar way as done by Reuter et al. 2011. For all comparisons, averaging kernels have been applied and the influence of the smoothing error reduced as described by Wunch 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 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 12/2022 for CO_2 and 04/2010 to 12/2020 for CH_4.

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.

$$\Delta X = a_0 + a_1 t + a_2 \sin(2\pi t + a_3) + \varepsilon \quad \text{Eq. 1}$$

Here, ΔX represents the difference satellite minus TCCON, ε the residual, and a_{0-3} the free fit parameters. Specifically, a_0 represents an offset, a_1 the linear drift and a_2 the amplitude of the seasonal bias at a TCCON site. The bias at a station is computed from the average of the fit model, 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 ΔX 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.

Table 1: L2 algorithms used in EMMA v4.5 CO_2.

Satellite/Instrument	Algorithm	Institution	ID	Reference
SCIAMACHY	BESD v02.01.02	IUP	2	Reuter et al. (2010, 2011, 2016)
GOSAT	NIES v02.9xbc (bias corrected)	NIES	3	Yoshida et al. (2013)
GOSAT	RemoTeC v2.3.8	SRON	5	Butz et al. (2011), Detmers et al. (2017a)
GOSAT	UoL-FP v7.3	UoL	6	Cogan et al (2012) Boesch and Anand (2017)
GOSAT	ACOS v9r	NASA	7	O'Dell et al. (2012), Taylor et al. (2022)
GOSAT	FOCAL v3.0	IUP	8	Noël et al. (2022)
OCO-2	NASA v11.1	NASA	9	Kiel et al. (2019)
OCO-2	FOCAL v10.1	IUP	10	Reuter et al. (2017a,b, 2021)
GOSAT-2	NIES v02.00	NIES	11	Yoshida and Oshio (2020)
GOSAT-2	RemoTeC v2.0.0	SRON	12	Krisna et al. (2021)
GOSAT-2	FOCAL v3.0	IUP	13	Noël et al. (2022)

Table 2: L2 algorithms used in EMMA v4.5 CH4.

Satellite/Instrument	Algorithm	Institution	ID	Reference
SCIAMACHY	WFMD v4.0	IUP	2	Schneising et al. (2018)
GOSAT	FOCAL-FP v3.0	IUP	3	Noël et al. (2022)
GOSAT	FOCAL-PR v3.0	IUP	4	Noël et al. (2022)
GOSAT	NIES v02.9xbc (bias corrected)	NIES	5	Yoshida et al. (2013)
GOSAT	RemoTeC-FP v2.3.8	SRON	7	Butz et al. (2011), Detmers et al. (2017a)
GOSAT	RemoTeC-PR v2.3.9	SRON	8	Butz et al. (2011), Detmers et al. (2017b)
GOSAT	UoL-FP v7.3	UoL	9	Cogan et al (2012) Boesch and Anand (2017)
GOSAT	UoL-PR v9.0	UoL	10	Cogan et al (2012) Boesch and Anand (2017)
GOSAT-2	FOCAL-FP v3.0	IUP	11	Noël et al. (2022)
GOSAT-2	FOCAL-PR v3.0	IUP	12	Noël et al. (2022)
GOSAT-2	RemoTeC-FP v2.0.0	SRON	13	Krisna et al. (2021)
GOSAT-2	RemoTeC-PR v2.0.1	SRON	14	Krisna et al. (2021)
GOSAT-2	NIES v02.00	NIES	15	Yoshida and Oshio (2020)

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 such pairs we compute the standard deviation as an 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.

It shall be noted that we use only those TCCON sites having co-locations with all contributing algorithms within the validation period. This has the advantage to achieve a good comparability among the validation results of the contributing algorithms. However, this limits our analysis to only six TCCON sites, rendering some of the validation results potentially less transferrable to globally valid conclusions. This might particularly be the case for the estimates of systematic biases and drifts.

Additionally, this document shows assessments of temporal and spatial bias patterns based on 10°x10° monthly gridded level 3 data sets (not to be confused with the XCO2_OBS4MIPS and XCH4_OBS4MIPS data sets).

We calculate the fraction of potential outliers according to unrealistically large spatial gradients (>3ppm/10° for XCO_2 and >20ppb/10° for XCH_4), unrealistically large deviations from the climatological model SLIM (Noël et al., 2022) and TCCON (>3ppm for XCO_2 and >20ppb for XCH_4), and larger deviations from EMMA (>2.5ppm for XCO_2 and >10ppb/10° for XCH_4).

We also compare the north/south (N/S) gradient of each month with the SLIM climatological model (Noël et al., 2022) 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 a few grid boxes include TCCON stations.

Additionally, we compare the seasonal (peak-to-peak) amplitude of each grid box with SLIM 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 a few grid boxes with seasonal cycles.

The TCCON-validation of the XCO2_OBS4MIPS and XCH4_OBS4MIPS level 3 products is based on comparisons of monthly mean data and is described in the main PQAR document (D1)

2. Validation Results

2.1 XCO2_EMMA

2.1.1 Validation

XCO_2 retrievals from individual XCO_2 algorithms and EMMA have been validated with TCCON. 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 3. Table 4 shows the validation summary specifically for EMMA v4.5 CO_2, i.e., the XCO2_EMMA product. The results are valid for the periods covered by the individual algorithms but earliest starting in 2009.

The individual algorithms have a single measurement precision in the range of 1.29 ppm (OCO-2 NASA) – 2.91 ppm (GOSAT-2 RemoTeC). EMMA has a single measurement precision of 1.78 ppm. EMMA's combined regional and seasonal biases (0.47 ppm) lie at the lower end of the range of the individual algorithms (0.37 ppm for GOSAT ACOS and GOSAT FOCAL to 0.98 ppm for GOSAT-2 NIES).

Figure 2 (top) 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 Garmisch-Partenkirchen TCCON site and a low bias of similar magnitude at the sites Sodankylä and Lamont. 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 (bottom) shows more or less small negative trends (typically below -0.2 ppm/a) for all satellite retrievals at the Bremen TCCON site. It shall be noted that due to the short validation period, the results for the GOSAT-2 algorithms are less reliable.

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 0.16 ppm/a (OCO-2 NASA) and 0.71 ppm/a (GOSAT RemoTeC). EMMA's year-to-year stability is 0.36 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. An example is shown in Figure 4.

The satellite retrieved seasonal amplitudes are in somewhat better agreement with TCCON than with SLIM (Figure 5, top). The opposite is true for the comparison of the north/south gradient. However, this should not be over interpreted because TCCON contributes only to few grid boxes especially on the southern hemisphere (Figure 5, bottom), even though all TCCON sites are used for this comparison rather than the reduced number of sites used for the L2 validation.

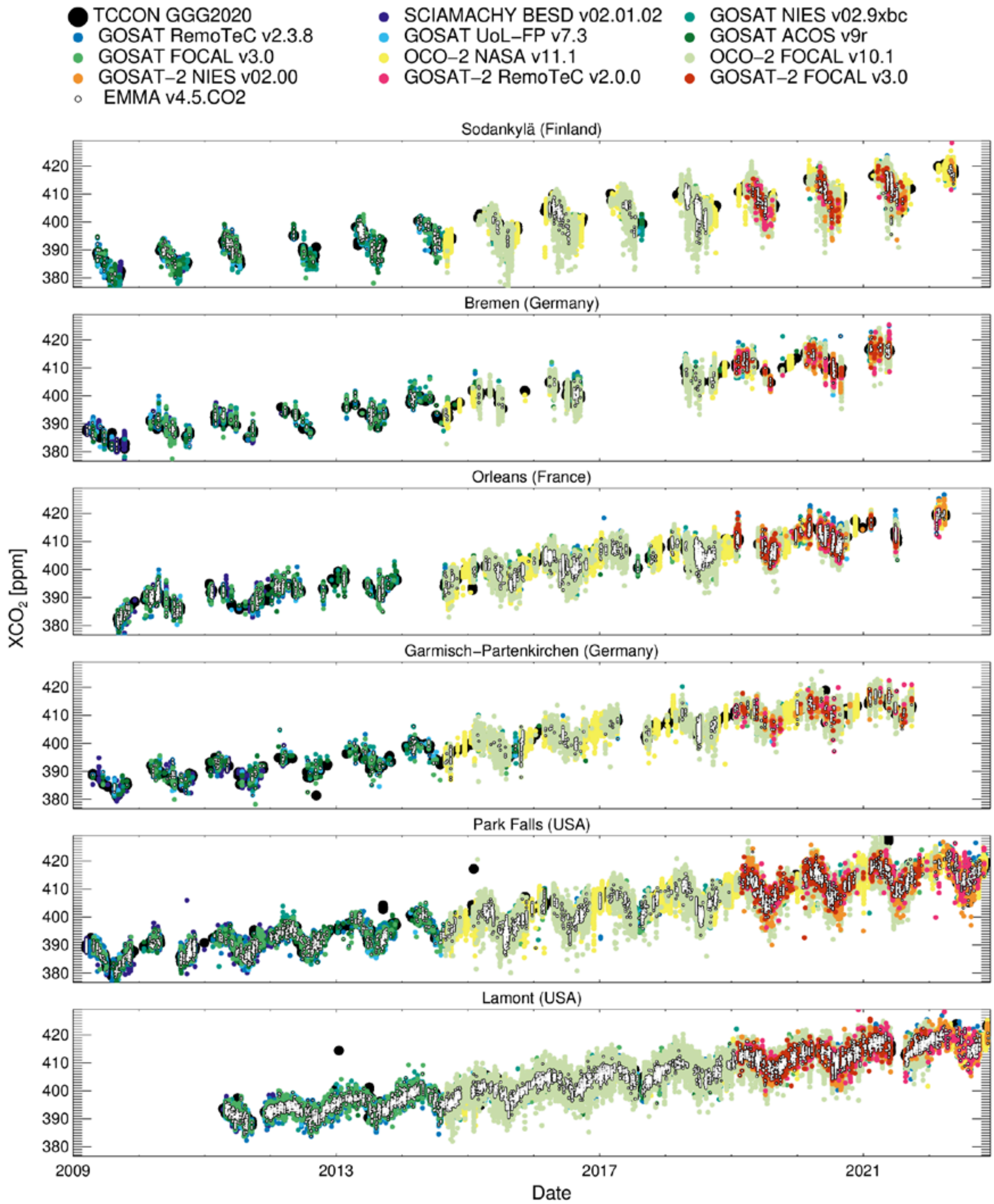


Figure 1: Validation of individual XCO_2 algorithms and EMMA with TCCON.

Table 3: Summarising XCO_2 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. Last 2 columns: '-' for products not generated in this project.

Algorithm	#	Precision [ppm]	Accuracy [ppm]			Stability [ppm/a]		Probability that TR is met [%]	
			Regional	Seasonal		Trend	Year2Year	Accuracy	Stability
SCIAMACHY BESD v02.01.02	6080	1.83	0.25	0.38		0.29±0.27	0.31	-	-
GOSAT NIES v02.9xbc	14088	2.21	0.41	0.31		-0.01±0.02	0.37	-	-
GOSAT RemoTeC v2.3.8	8490	2.26	0.62	0.36		0.06±0.07	0.71	-	-
GOSAT UoL-FP v7.3	10009	1.95	0.45	0.41		-0.02±0.08	0.34	57%	98%
GOSAT ACOS v9r	11239	1.73	0.28	0.25		0.02±0.05	0.29	-	-
GOSAT FOCAL v3.0	9581	2.28	0.28	0.24		-0.07±0.03	0.34	-	-
OCO-2 NASA v11.1	1980981	1.29	0.36	0.24		0.00±0.12	0.16	-	-
OCO-2 FOCAL v10.1	839449	1.76	0.34	0.22		-0.03±0.14	0.23	-	-
GOSAT-2 NIES v02.00	5924	2.52	0.60	0.78		-0.20±0.23	0.28		
GOSAT-2 RemoTeC v2.0.0	2692	2.91	0.63	0.61		-0.30±0.96	0.41	28%	37%
GOSAT-2 FOCAL v3.0	2779	2.09	0.46	0.37		-0.28±0.37	0.26	-	-
EMMA v4.5	27281	1.78	0.41	0.22		0.02±0.09	0.36	71%	98%

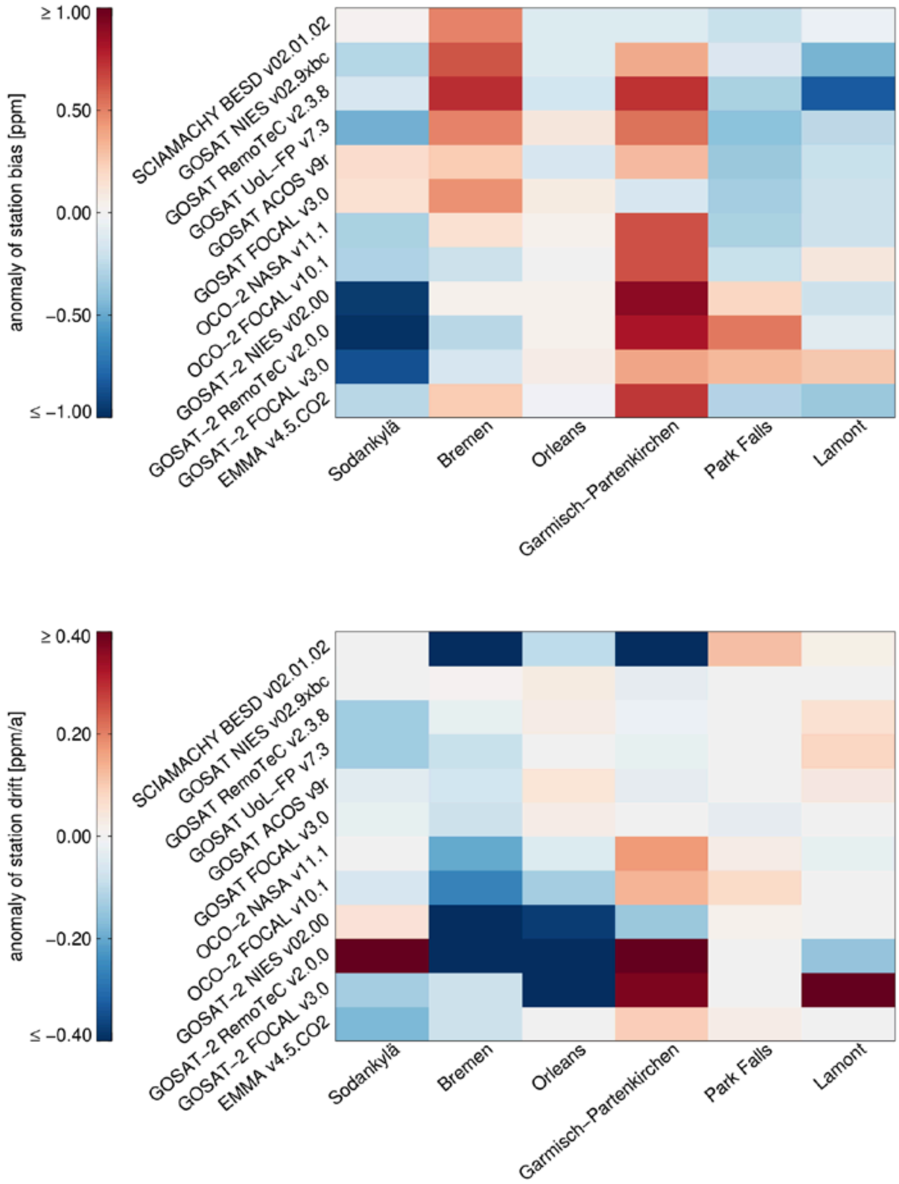


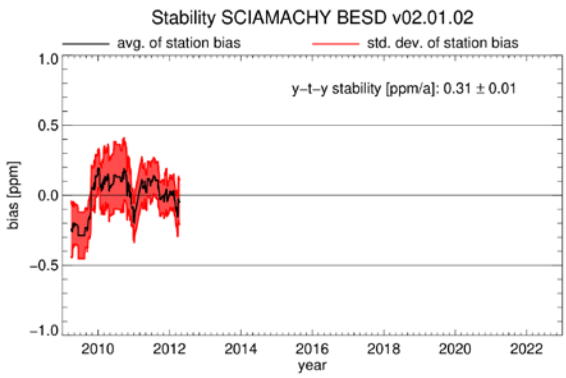
Figure 2: Anomaly of station biases (top) and station drift (bottom).

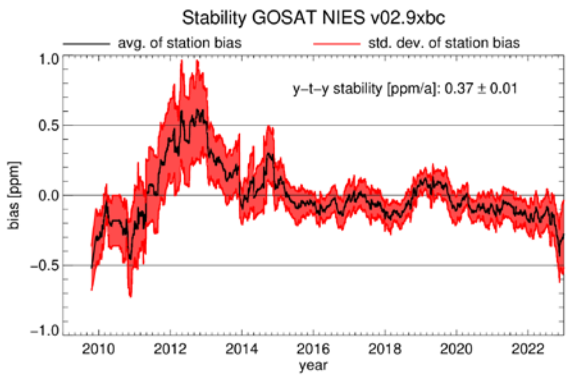
In terms of the frequency of potential outliers and standard deviation of the difference to SLIM and TCCON, EMMA is among the best performing algorithms (Figure 6).

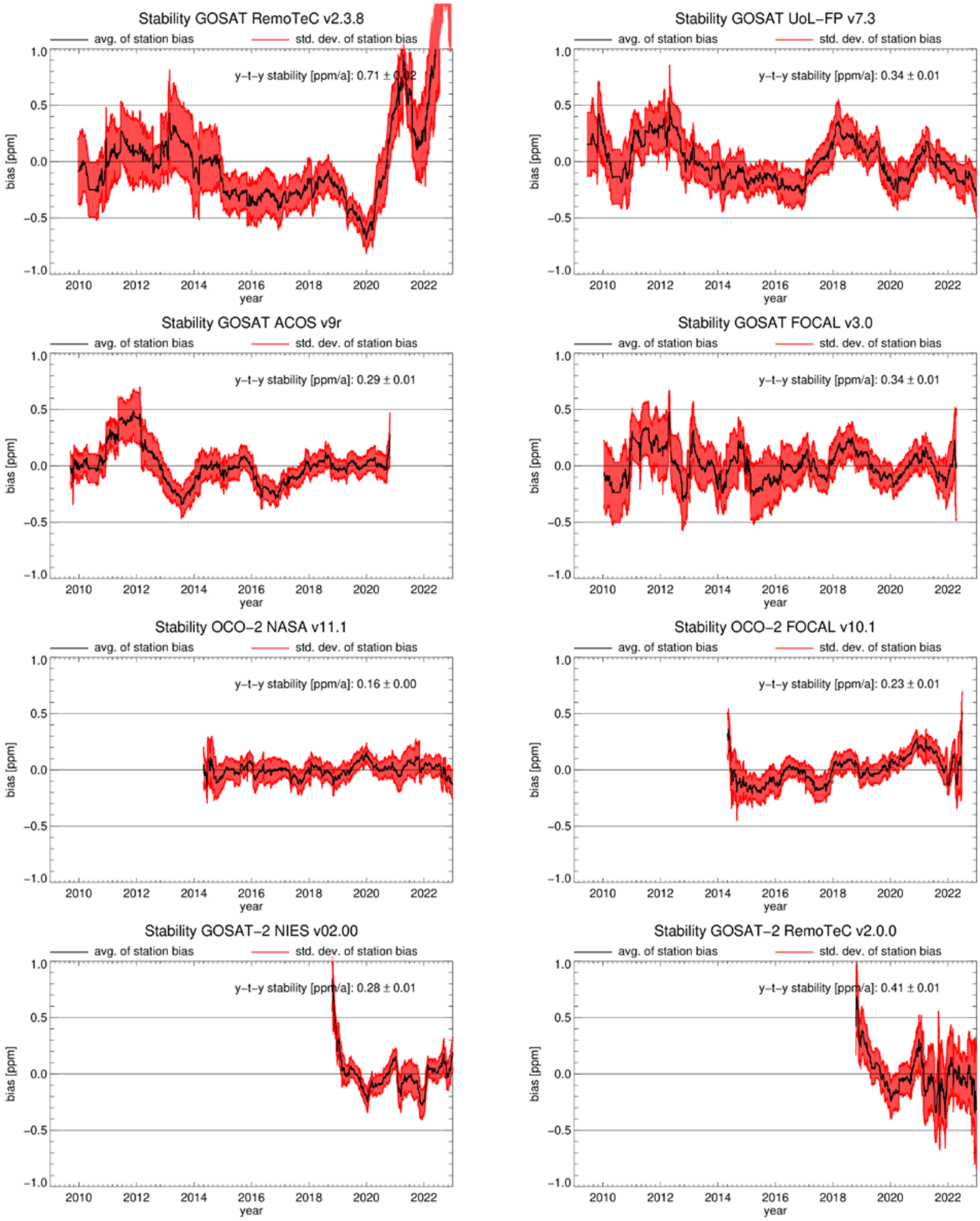
ACOS and FOCAL usually provide the largest part of the relative data weight of the GOSAT algorithms in EMMA (Figure 7). However, the relative data weight of the OCO-2 algorithms starting in 2014 is still much larger. The drop in data weight of the OCO-2 NASA algorithm in 2022 is due to a feature of EMMA which tries to prevent a single algorithm from contributing much more data than the other algorithms. I.e., the OCO-2 NASA algorithm may still be selected as the median as often as in the years before, but its number of soundings per grid box is reduced by random thinning.

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 8, top). 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 8, bottom). The average inter-algorithm spread slightly increases when the GOSAT-2 algorithms start to contribute (Figure 9).

It is interesting to note that the average inter-algorithm spread usually reduced with a new EMMA version (always including the most recent algorithm versions, Figure 10). This means that EMMA observed kind of a convergence among the individual algorithms, even though the number of algorithms and the length of the period increased. 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 and v4.5 do not follow this trend because the average inter-algorithm spread increased slightly. For v3.1 this was most probably caused by adding the not bias corrected operational NIES product and NIES' PPDF-S product. For v4.5, the addition of the GOSAT-2 NIES product and larger year-to-year variations of the GOSAT RemoTeC product in the last years my play a role. Additionally, the EMMA period enhanced so that small drifts in the data sets, which are not accounted for by EMMA's overall offset correction, can contribute to a larger extend. After offset correction, this can, particularly, influence the systematic bias at the start and end of a data set's time series.






Stability GOSAT RemoTeC v2.3.8
avg. of station bias
std. dev. of station bias
y–t–y stability [ppm/a]: 0.71 ± 0.02
Stability GOSAT UoL–FP v7.3
avg. of station bias
std. dev. of station bias
y–t–y stability [ppm/a]: 0.34 ± 0.01
Stability GOSAT ACOS v9r
avg. of station bias
std. dev. of station bias
y–t–y stability [ppm/a]: 0.29 ± 0.01
Stability GOSAT FOCAL v3.0
avg. of station bias
std. dev. of station bias
y–t–y stability [ppm/a]: 0.34 ± 0.01
Stability OCO–2 NASA v11.1
avg. of station bias
std. dev. of station bias
y–t–y stability [ppm/a]: 0.16 ± 0.00
Stability OCO–2 FOCAL v10.1
avg. of station bias
std. dev. of station bias
y–t–y stability [ppm/a]: 0.23 ± 0.01
Stability GOSAT–2 NIES v02.00
avg. of station bias
std. dev. of station bias
y–t–y stability [ppm/a]: 0.28 ± 0.01
Stability GOSAT–2 RemoTeC v2.0.0
avg. of station bias
std. dev. of station bias
y–t–y stability [ppm/a]: 0.41 ± 0.01
bias [ppm]
year
1.0
0.5
0.0
–0.5
–1.0
2010
2012
2014
2016
2018
2020
2022

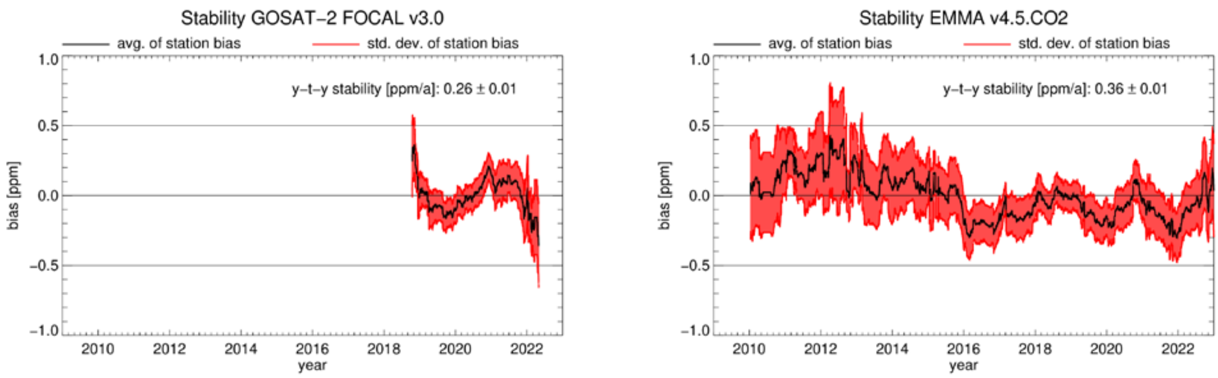


Figure 3: Stability analyses for EMMA and the contributing individual algorithms. The black curve shows the average station bias and the red curves its uncertainty represented by the station-to-station standard deviation and error propagation from single sounding measurement noise.

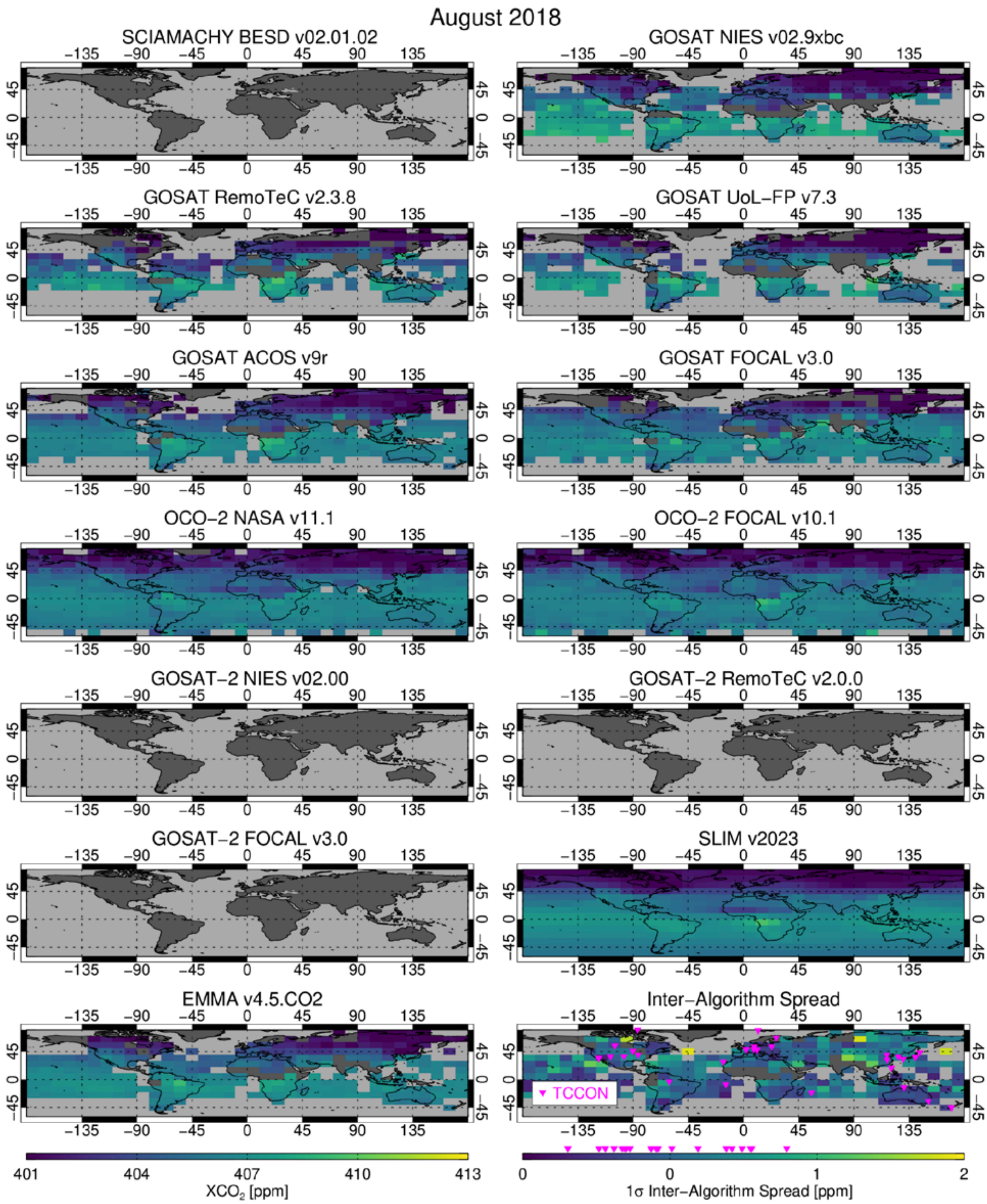
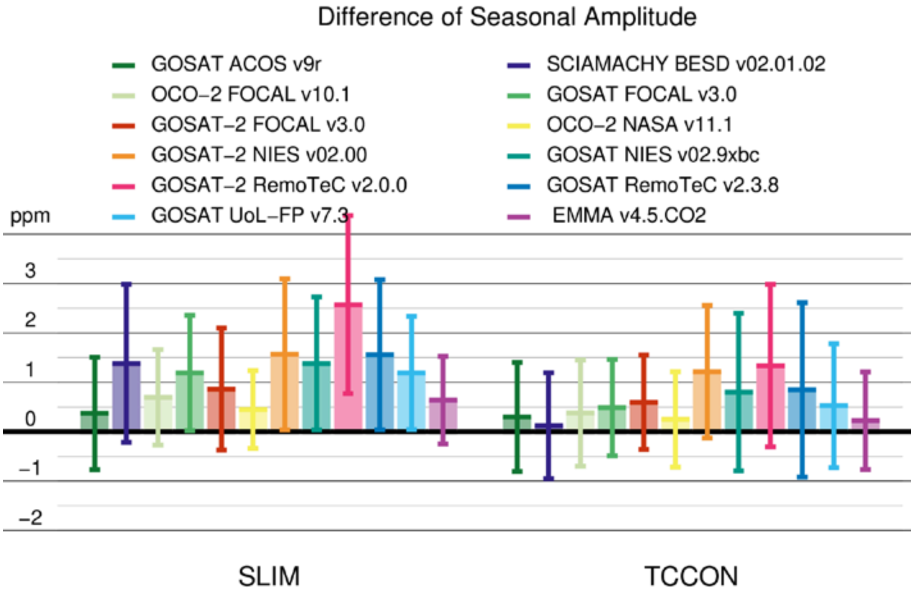


Figure 4: Monthly gridded XCO_2 averages for EMMA and the individual algorithms as well as the inter-algorithm spread at the example of August 2018.



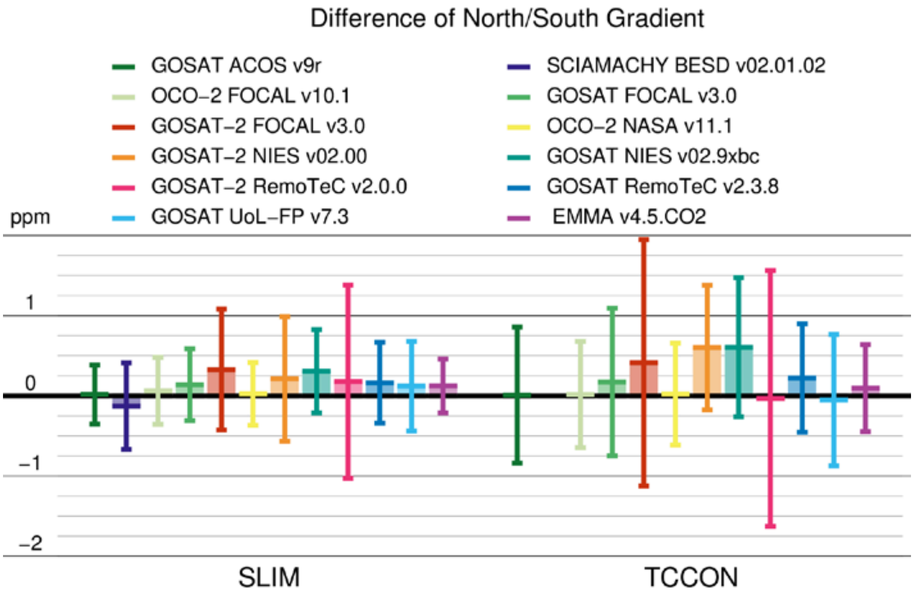


Figure 5: Top: Difference of seasonal cycle amplitude of all individual algorithms as well as EMMA compared with SLIM and TCCON. Bottom: Difference of north/south gradient of all individual algorithms as well as EMMA compared with SLIM and TCCON.

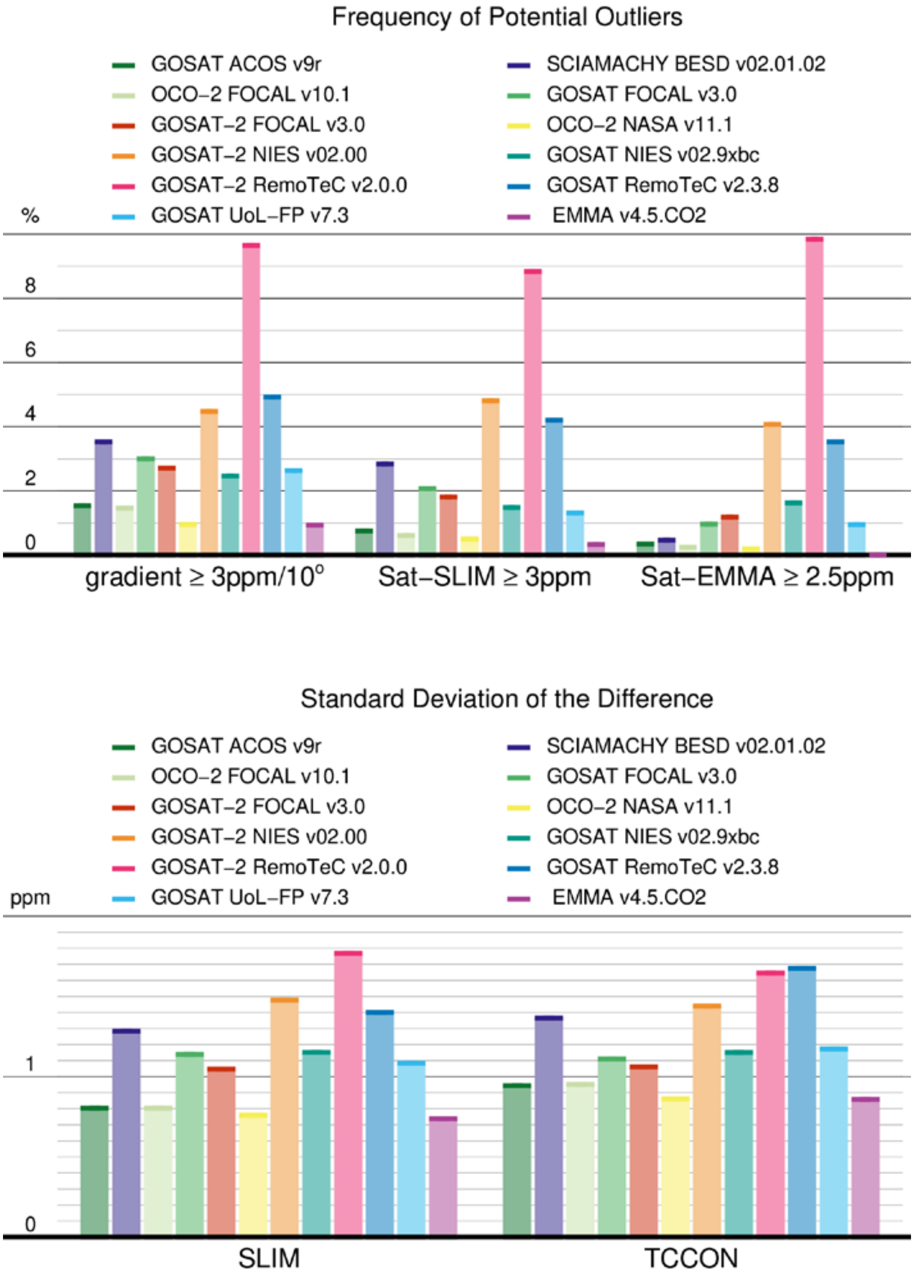


Figure 6: Frequency of potential outliers estimated by large gradients, large differences to SLIM, and large differences to EMMA (top). Standard deviation of the difference of all algorithms and EMMA to SLIM and TCCON (bottom).

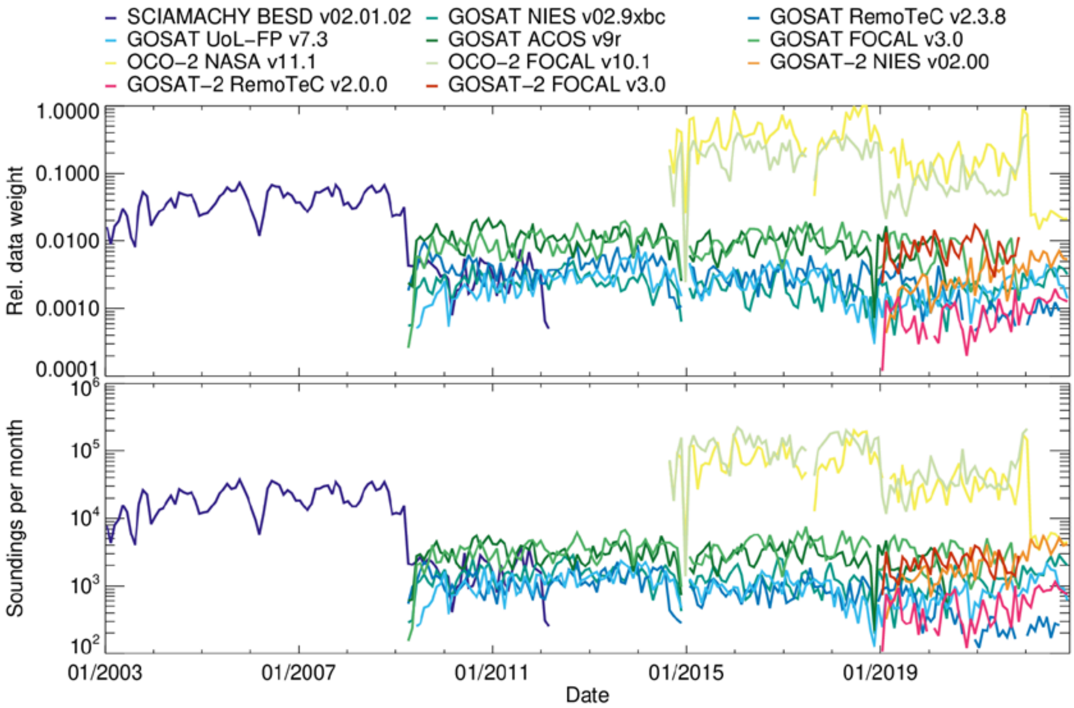


Figure 7: EMMA's normalized relative data weight proportional to $\sum 1/\sigma_i^2$ (top) and number of soundings (bottom) per algorithm and month.

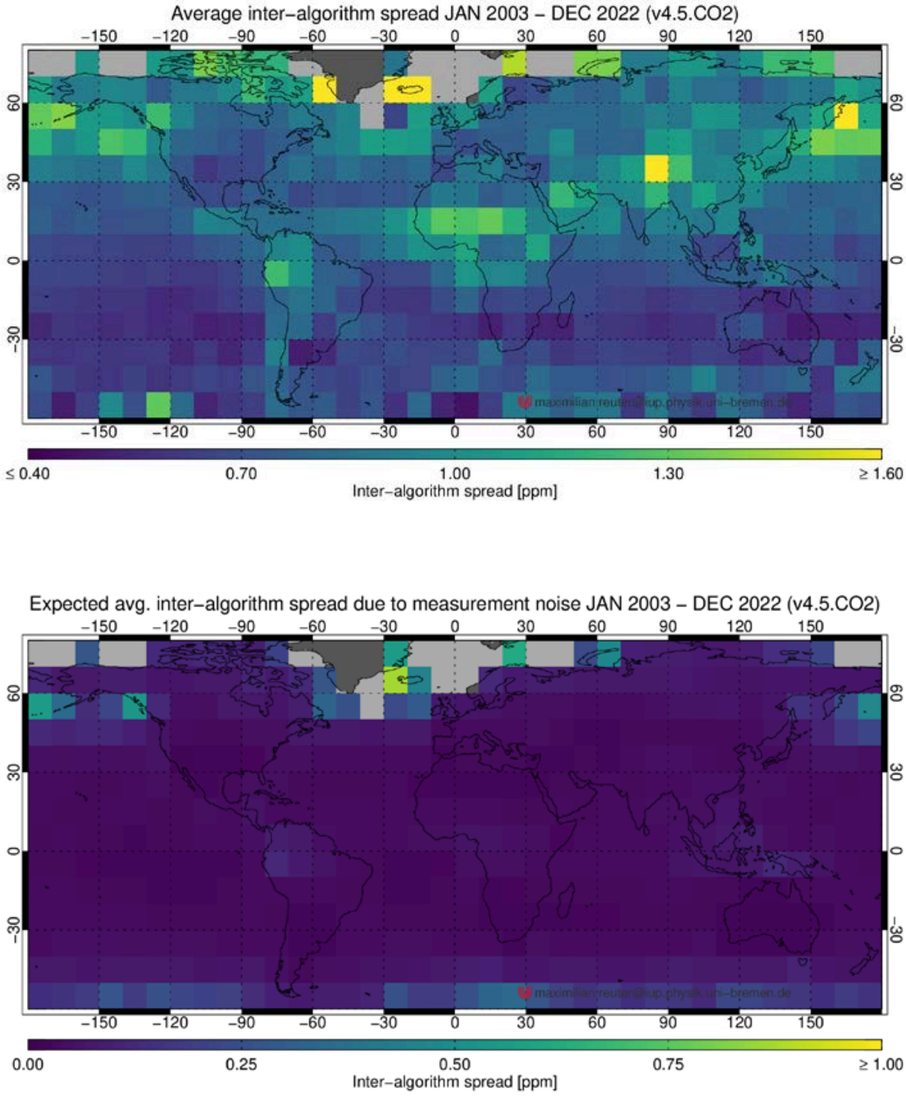


Figure 8: Average inter-algorithm scatter of monthly 10°x10° averages (top) and corresponding expected contribution of measurement noise (bottom).

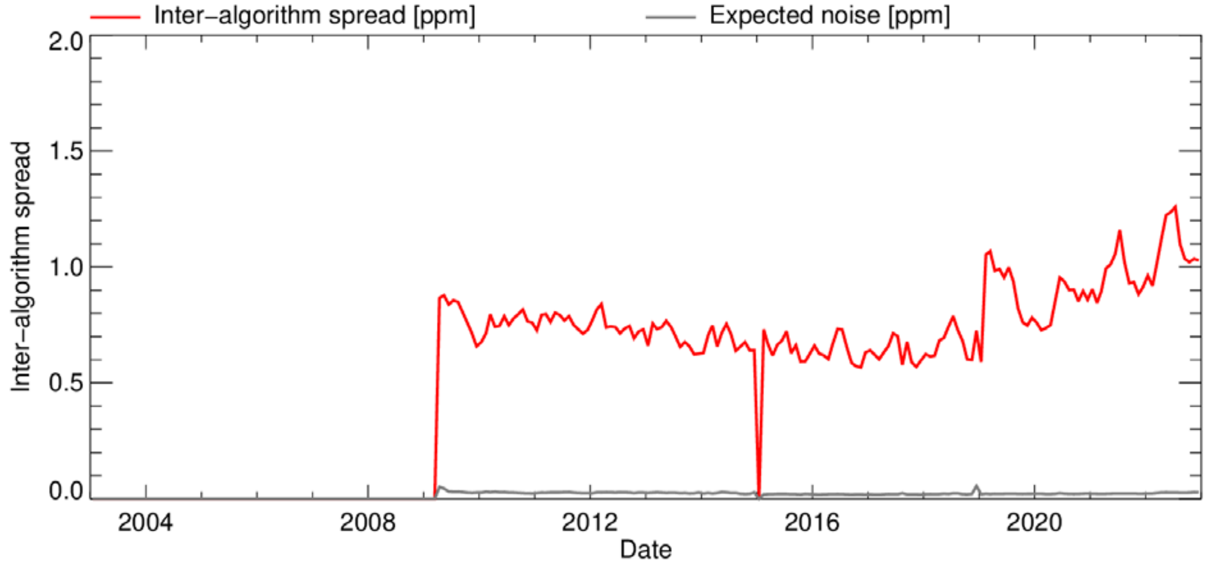


Figure 9: Monthly average of the inter-algorithm scatter and expected contribution of measurement noise.

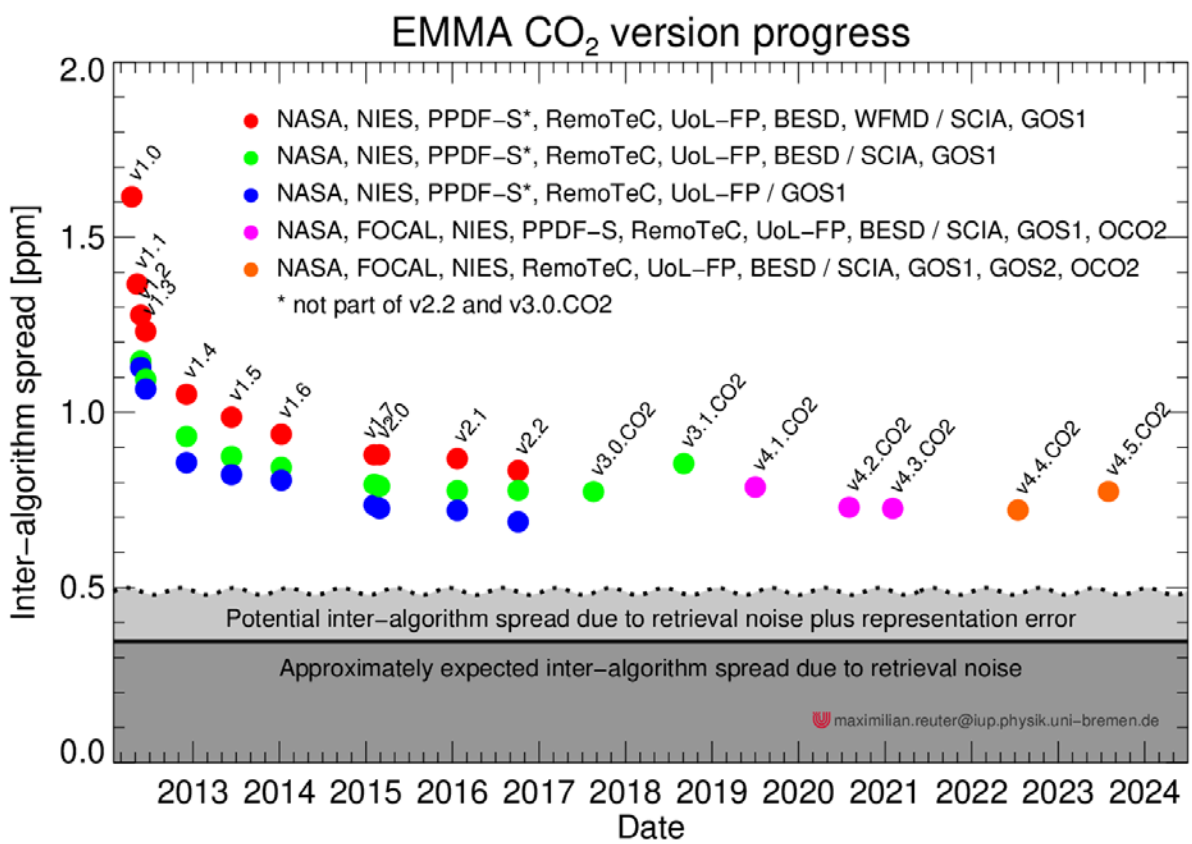


Figure 10: Average inter-algorithm spread of all EMMA versions compared with the approximately expected contribution of retrieval noise and a rough estimate of the representation error.

2.1.2 Summary

The validation results are summarised in Table 4.

Table 4: Product Quality Summary Table for product XCO2_EMMA. The listed requirements are the threshold (T) requirements as given in TRD (D3). For precision (i.e., single observation statistical uncertainty or random error) also the corresponding breakthrough (B) and goal (G) requirements are listed. For the achieved performance of (relative) "Accuracy" two values are listed: The first one is the spatial component of the bias and the second one is the spatio-temporal bias, computed by also considering seasonal biases. The spatio-temporal bias is our estimate of "relative accuracy". TR refers to "target requirement" and reported is the probability that the corresponding TR is met, i.e., the probabilities that accuracy is better than 0.5 ppm and stability is better than 0.5 ppm/year.

<table>
<tr><th colspan="5">Product Quality Summary Table for Product: XCO2_EMMA
Level: 2, Version: v4.5, Time period covered: 01.2003 – 12.2022</th></tr>
<tr><th>Parameter [unit]</th><th>Achieved performance</th><th>Requirement</th><th>TR</th><th>Comments</th></tr>
<tr><td>Single measurement precision (1-sigma) in [ppm]</td><td>1.78</td><td>< 8 (T)
< 3 (B)
< 1 (G)</td><td>-</td><td>-</td></tr>
<tr><td>Uncertainty ratio in [-]: Ratio reported uncertainty to standard deviation of satellite-TCCON difference</td><td>0.97</td><td>-</td><td>-</td><td>No requirement but value close to unity expected for a high quality data product.</td></tr>
<tr><td>Mean bias [ppm]</td><td>0.41</td><td>-</td><td>-</td><td>No requirement but value close to zero expected for a high quality data product.</td></tr>
<tr><td>Accuracy: Relative systematic error [ppm]</td><td>Spatial – spatiotemporal:
0.41 – 0.47</td><td>< 0.5</td><td>Probability that accuracy TR is met: 71%</td><td rowspan="3">Only six TCCON sites fulfilled all validation criteria. Therefore, estimates of these systematic error components are potentially less reliable</td></tr>
<tr><td>Stability: Drift [ppm/year]</td><td>0.02±0.09
(1-sigma)</td><td>< 0.5</td><td>Probability that stability TR is met: 98%</td></tr>
<tr><td>Stability: Year-to-year bias variability [ppm/year]</td><td>0.36
(1-sigma)</td><td>< 0.5</td><td>-</td></tr>
</table>

2.2 XCH4_EMMA

2.2.1 Validation

Figure 11 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 5. Table 6 shows the validation summary specifically for EMMA v4.5 CH_4 i.e., the XCH4_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). 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_4. Similarly, some of the GOSAT-2 validation results are potentially less robust, because of the shorter validation period for these algorithms.

The individual algorithms have a single measurement precision in the range of 11.51 ppb (GOSAT-2 FOCAL-FP) to 17.80 ppb (GOSAT-2 RemoTeC-PR) except for WFMD which has a single measurement precision of 104.27 ppb. EMMA has a single measurement precision of 13.51 ppb which is similar to that of the GOSAT algorithms. EMMA's combined regional and seasonal biases (2.59 ppb) are at the lower end of the range of the individual algorithms (2.00 ppb for GOSAT FOCAL FP to 20.39 ppb for WFMD).

Figure 12 (top) shows the anomaly of station biases of the used algorithms. One can see that all satellite retrievals have a consistent high bias of typically about 2 ppb at the Park Falls TCCON site and a low bias of similar magnitude at Bremen, Orleans, and Lamont. 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 of the GOSAT algorithms in Figure 12 (bottom) shows no consistent drifts for one of the satellite retrievals or at a specific TCCON site. We observer relatively large positive and negative drifts (sometimes exceeding ±3 ppb/a) for WFMD and the GOSAT-2 algorithms. However, these can be considered less reliable due to the short validation period of the corresponding algorithms.

Figure 13 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 1.4 ppb/a (GOSAT-2 RemoTeC PR) to 3.5 ppb/a (GOSAT RemoTeC-PR) for the GOSAT and GOST-2 algorithms. EMMA's year-to-year stability of 1.8 ppb/a is at the lower end of the range of the individual algorithms. The year-to-year stability of WFND could not be reliably determined due to the relatively short time span and few station data within the time period.

Analyses of gridded L3 data show that all algorithms reproduce large scale features well, however, there are still differences of some 10 ppb when looking into the details. An example is shown in Figure 14.

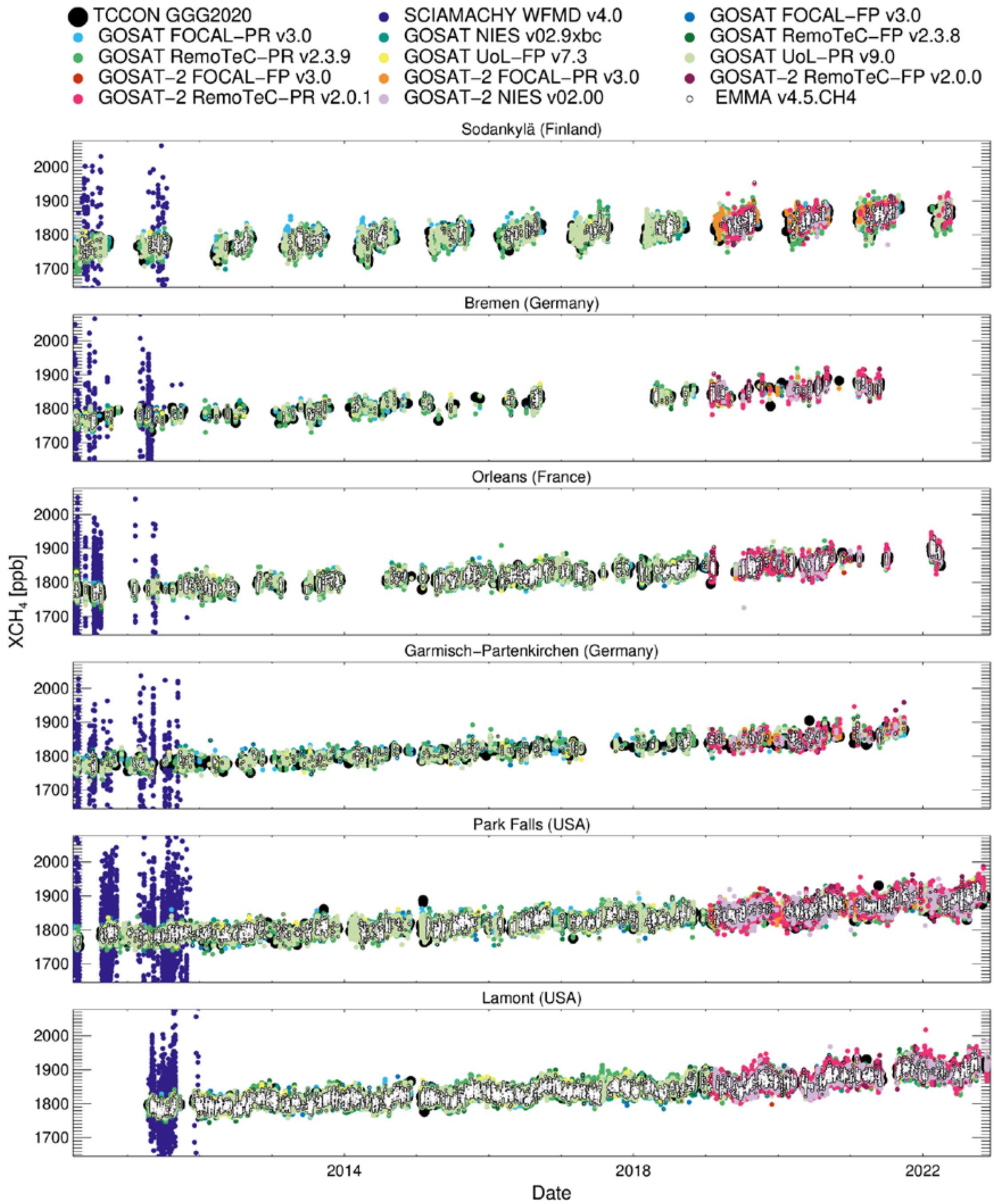
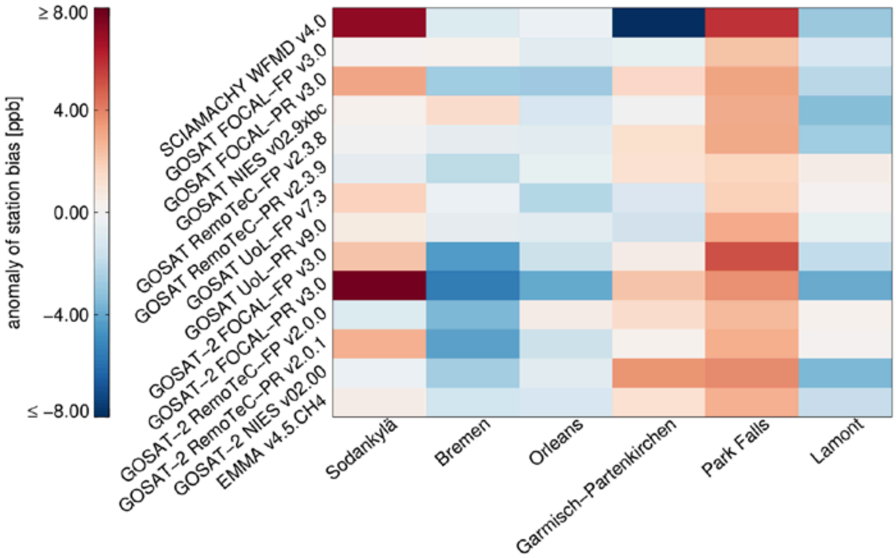


Figure 11: Validation of individual XCH_4 algorithms and EMMA with TCCON.

Table 5: Summarising XCH_4 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. Last 2 columns: '-' for products not generated in this project.

Algorithm	#	Precision [ppb]	Accuracy [ppb]		Stability [ppb/a]		Probability that TR is met [%]	
			Regional	Seasonal	Trend	Year2Year	Accuracy	Stability
SCIAMACHY WFMD v4.0	5425	104.27	5.75	19.56	NA	NA	-	-
GOSAT FOCAL-FP v3.0	11118	12.26	1.27	1.55	-0.87±0.21	1.9	-	-
GOSAT FOCAL-PR v3.0	26346	13.15	2.86	2.97	-0.21±0.26	1.7	-	-
GOSAT NIES v02.9xbc (bias corrected)	13731	13.31	2.15	1.13	-0.23±0.19	1.8	-	-
GOSAT RemoTeC-FP v2.3.8	8138	14.29	1.96	1.57	0.34±0.22	3.5	-	-
GOSAT RemoTeC-PR v2.3.9	32546	15.15	1.36	2.45	0.25±0.34	2.4	-	-
GOSAT UoL-FP v7.3	9343	13.25	1.58	2.98	-0.32±0.14	1.9	93%	100%
GOSAT UoL-PR v9.0	30599	14.31	1.58	2.34	0.34±0.26	1.7	95%	99%
GOSAT-2 FOCAL-FP v3.0	2527	11.51	3.36	2.67	-1.89±3.62	1.5	-	-
GOSAT-2 FOCAL-PR v3.0	6294	12.12	5.29	2.69	-0.70±2.00	1.5	-	-
GOSAT-2 RemoTeC-FP v2.0.0	2692	17.31	2.11	3.81	2.03±4.51	2.4	91%	45%
GOSAT-2 RemoTeC-PR v2.0.0	9415	17.80	2.71	4.01	3.52±1.61	1.4	89%	39%
GOSAT-2 NIES v01.07	5937	13.79	3.07	4.03	-1.59±2.40	1.8	-	-
EMMA v4.5	20659	13.51	1.79	1.87	0.11±0.22	1.8	95%	100%



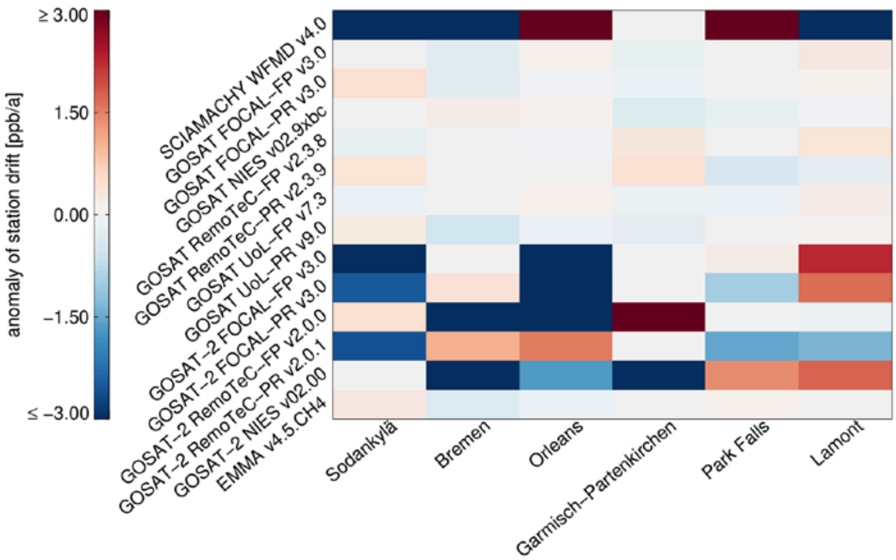


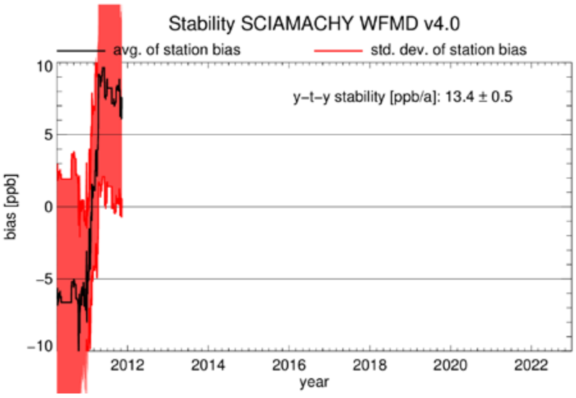
Figure 12: Anomaly of station biases (top) and station drift (bottom).

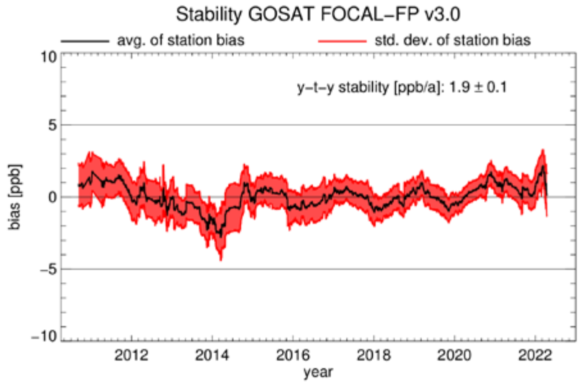
The satellite retrieved seasonal amplitudes are mostly in good agreement with but slightly larger than for TCCON and SLIM (Figure 15, top).

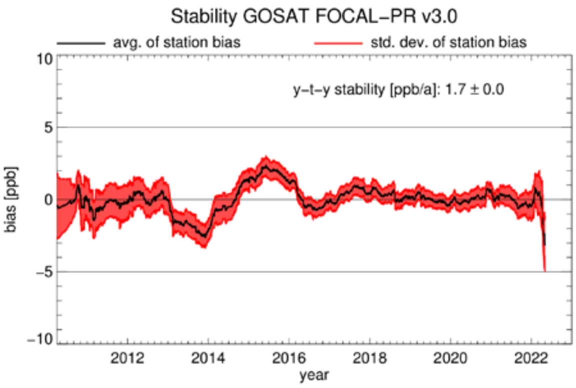
Comparison of the north/south gradient shows a small but relative consistent underestimation relative to SLIM and a small overestimation of similar magnitude relative to TCCON. However, this should not be over interpreted because TCCON contributes only to a few grid boxes especially in the southern hemisphere (Figure 15, bottom).

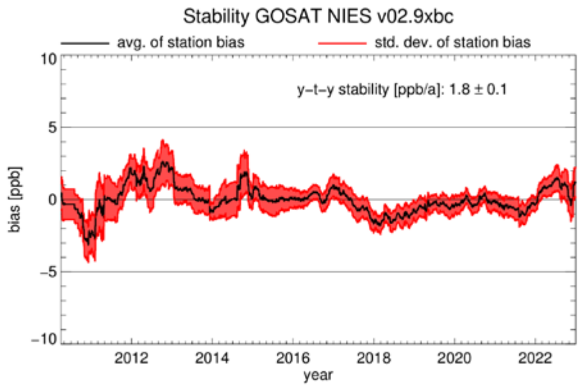
In terms of the frequency of potential outliers and standard deviation of the difference to SLIM and TCCON, EMMA is among the best performing algorithms (Figure 16, bottom).

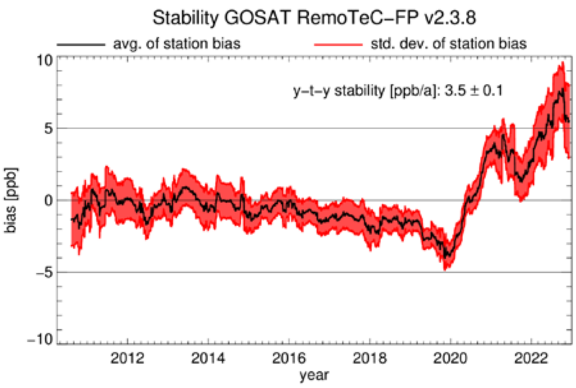
Once the GOSAT algorithms start operating in 2009, the proxy algorithms UoL-PR and FOCAL usually provide the largest part of the relative data weight in EMMA (Figure 17). Note also the drop in the relative data weight of WFMD at the end of 2005 due to instrument degradation.

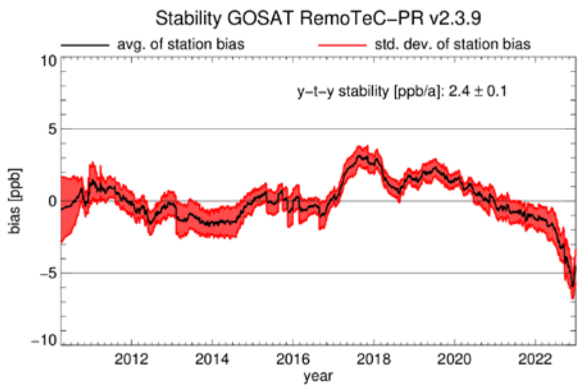

Stability SCIAMACHY WFMD v4.0
avg. of station bias
std. dev. of station bias
y–t–y stability [ppb/a]: 13.4 ± 0.5
bias [ppb]
year

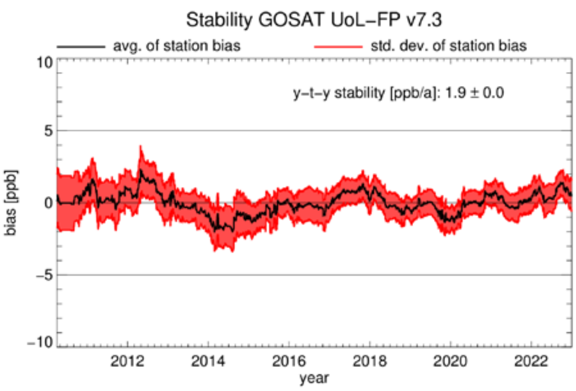

Stability GOSAT FOCAL–FP v3.0
avg. of station bias
std. dev. of station bias
y–t–y stability [ppb/a]: 1.9 ± 0.1
bias [ppb]
year

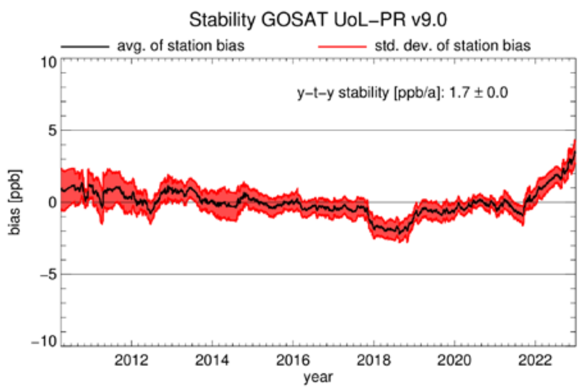

Stability GOSAT FOCAL–PR v3.0
avg. of station bias
std. dev. of station bias
y–t–y stability [ppb/a]: 1.7 ± 0.0
bias [ppb]
year


Stability GOSAT NIES v02.9xbc
avg. of station bias
std. dev. of station bias
y–t–y stability [ppb/a]: 1.8 ± 0.1
bias [ppb]
year


Stability GOSAT RemoTeC–FP v2.3.8
avg. of station bias
std. dev. of station bias
y–t–y stability [ppb/a]: 3.5 ± 0.1
bias [ppb]
year


Stability GOSAT RemoTeC–PR v2.3.9
avg. of station bias
std. dev. of station bias
y–t–y stability [ppb/a]: 2.4 ± 0.1
bias [ppb]
year


Stability GOSAT UoL–FP v7.3
avg. of station bias
std. dev. of station bias
y–t–y stability [ppb/a]: 1.9 ± 0.0
bias [ppb]
year


Stability GOSAT UoL–PR v9.0
avg. of station bias
std. dev. of station bias
y–t–y stability [ppb/a]: 1.7 ± 0.0
bias [ppb]
year

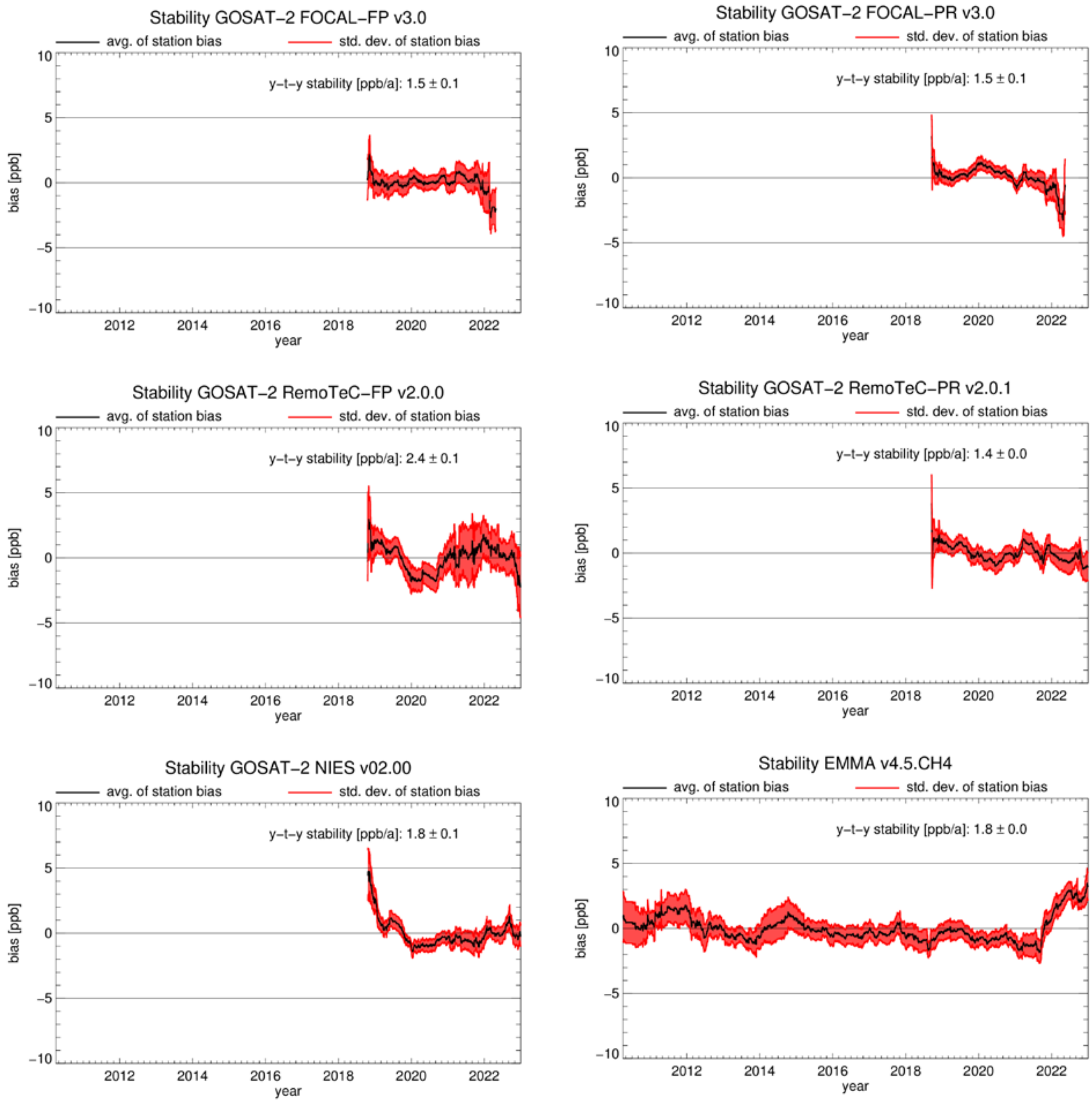


Figure 13: Stability analyses for EMMA and the contributing individual algorithms. The black curve shows the average station bias and the red curves its uncertainty represented by the station-to-station standard deviation and error propagation from single sounding measurement noise.

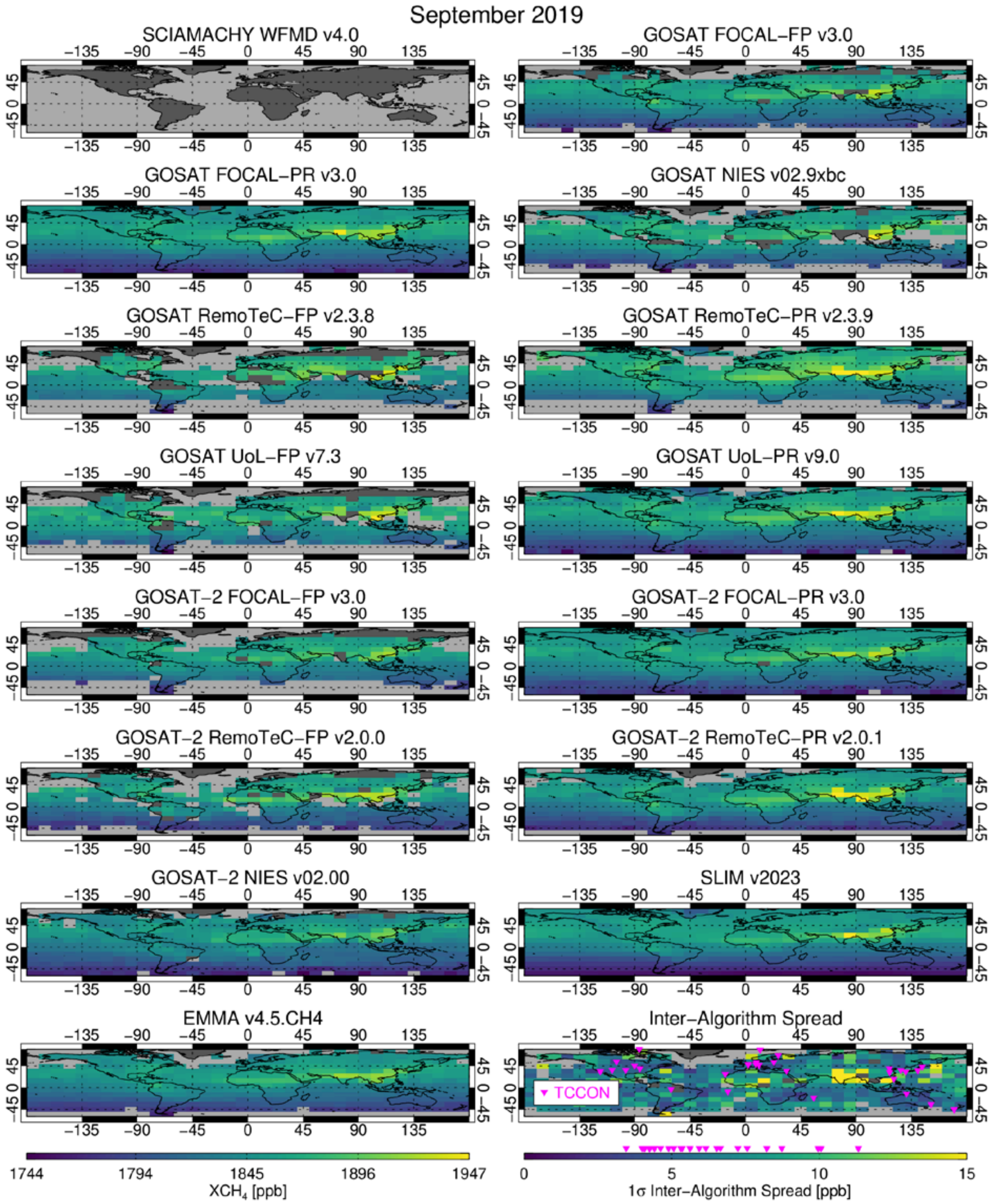
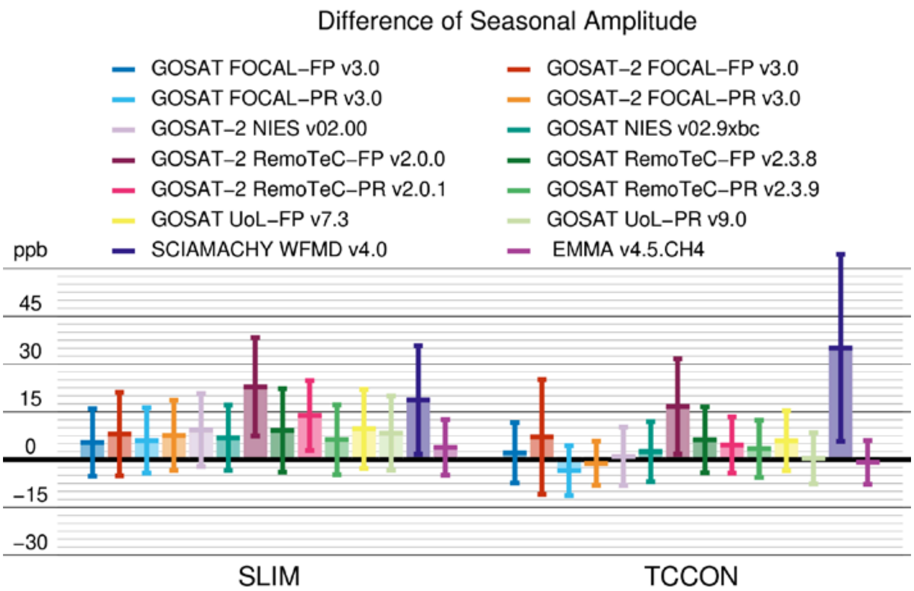


Figure 14: Monthly gridded XCH_4 averages for EMMA and the individual algorithms as well as the inter-algorithm spread at the example of September 2019.



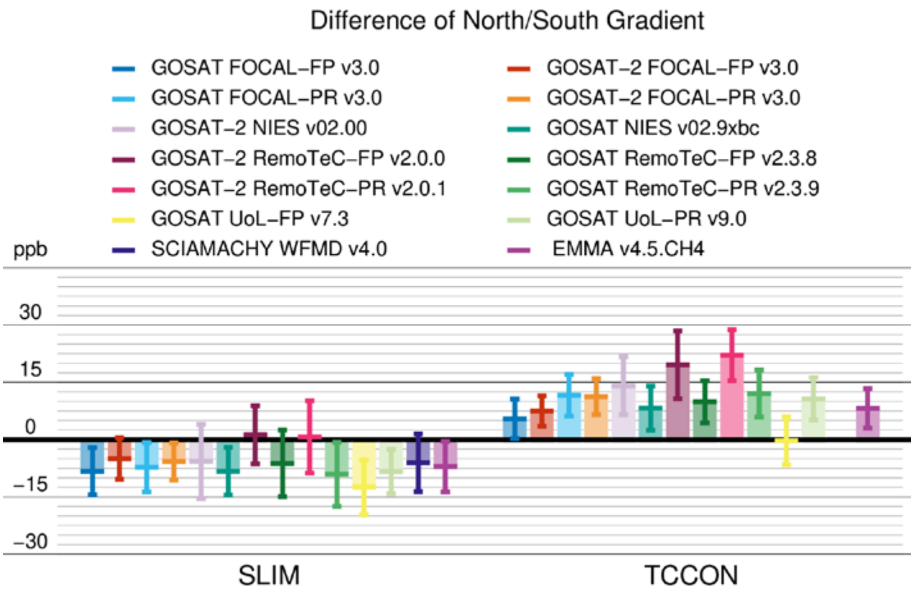


Figure 15: Difference of seasonal cycle amplitude of all individual algorithms as well as EMMA compared with SLIM and TCCON (top). Difference of north/south gradient of all individual algorithms as well as EMMA compared with SLIM and TCCON (bottom).

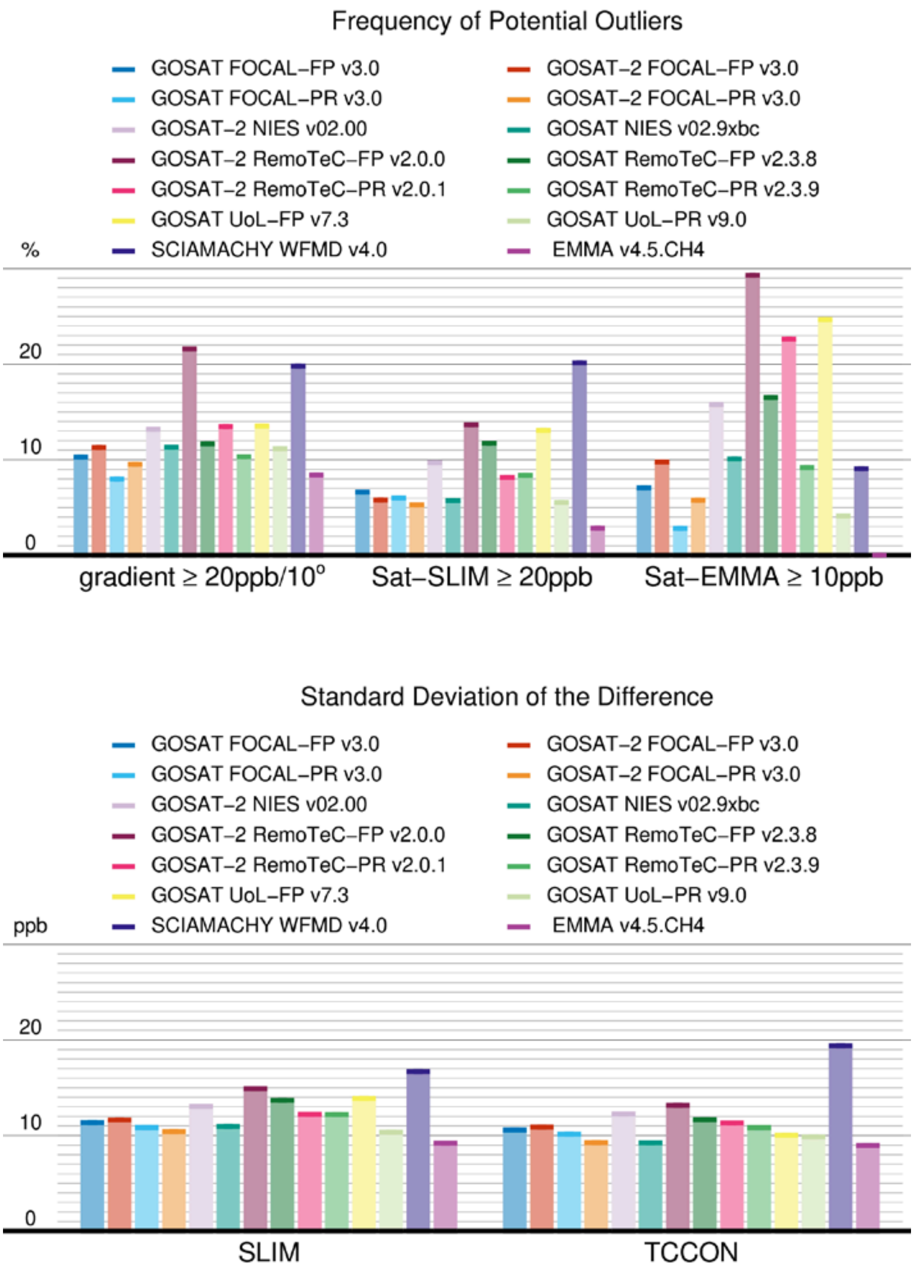


Figure 16: Frequency of potential outliers estimated by large gradients, large differences to SLIM, and large differences to EMMA (top). Standard deviation of the difference of all algorithms and EMMA to SLIM and TCCON (bottom).

The average inter-algorithm spread has values mostly between 2 ppb and 12 ppb and is typically below 9 ppb (Figure 18, top). 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 18, bottom). The average inter-algorithm spread increases when the GOSAT-2 algorithms start to contribute (Figure 19).

The average inter-algorithm spread enhanced till EMMA v3.1 and slightly reduced since then (Figure 20) with the exception of v4.5 showing a slight enhancement again. Part of the enhancement till EMMA v3.1 is most probably caused by adding the non bias corrected operational NIES product and NIES' PPDF-S product in v3.1. Additionally, the EMMA period enhanced so that small drifts in the data sets, which are not accounted for by EMMA's overall offset correction, can contribute to a larger extent. After offset correction, this can, particularly, influence the systematic bias at the start and end of a data set's time series. This may also explain the slight increase of the inter-algorithm spread observed for v4.5.

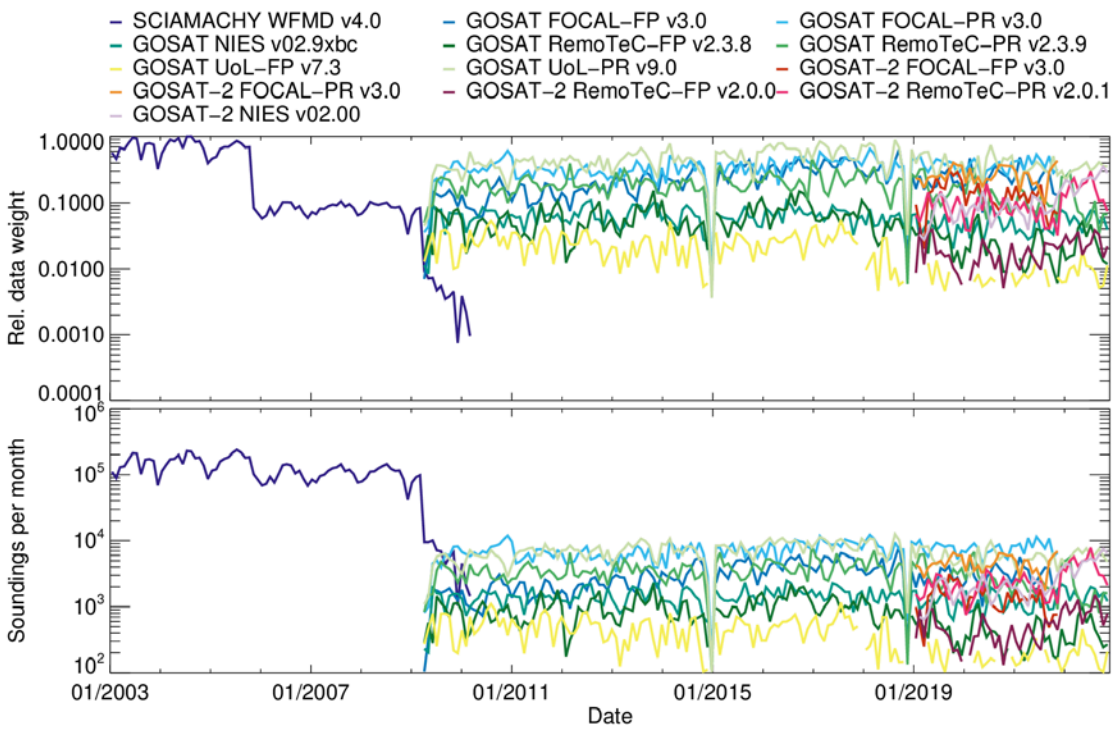
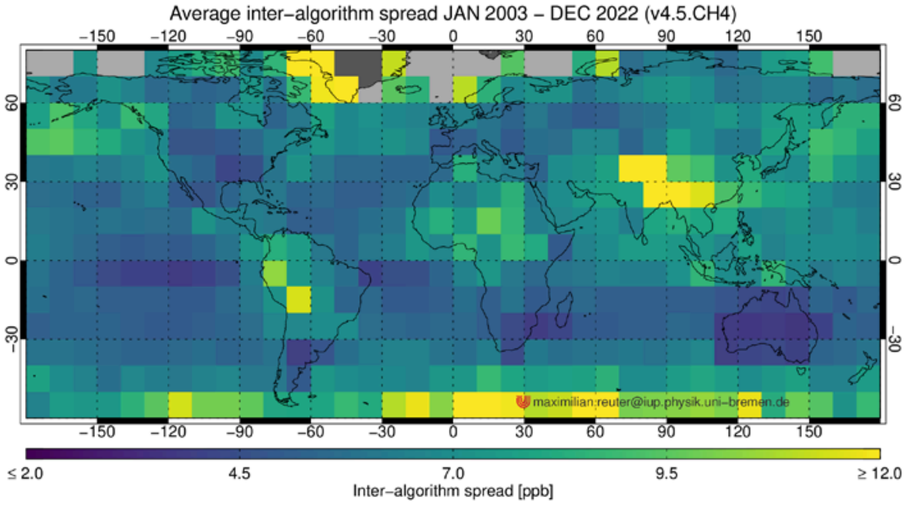


Figure 17: EMMA's normalized relative data weight proportional to $\sum 1/\sigma_i^2$ (top) and number of soundings (bottom) per algorithm and month.



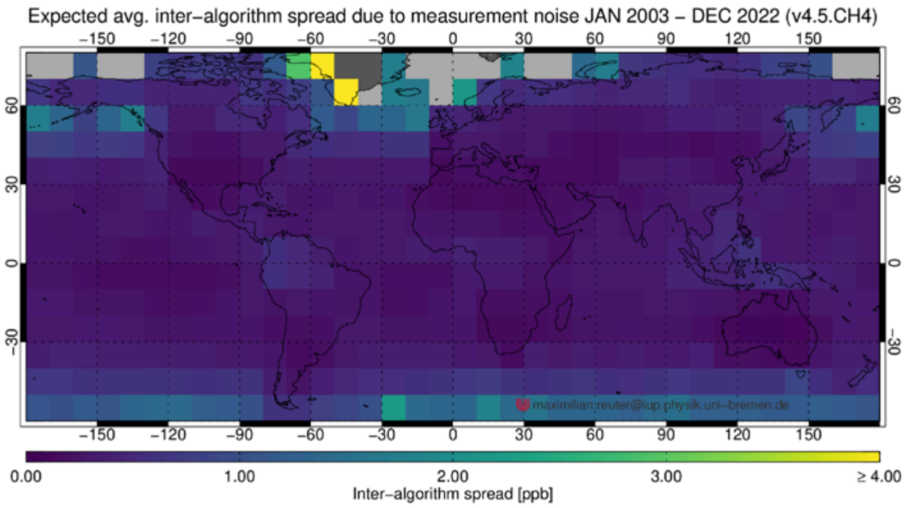


Figure 18: Average inter-algorithm scatter of monthly 10°x10° averages (top) and corresponding expected contribution of measurement noise (bottom).

Figure 19: Monthly average of the inter-algorithm scatter and expected contribution of measurement noise.

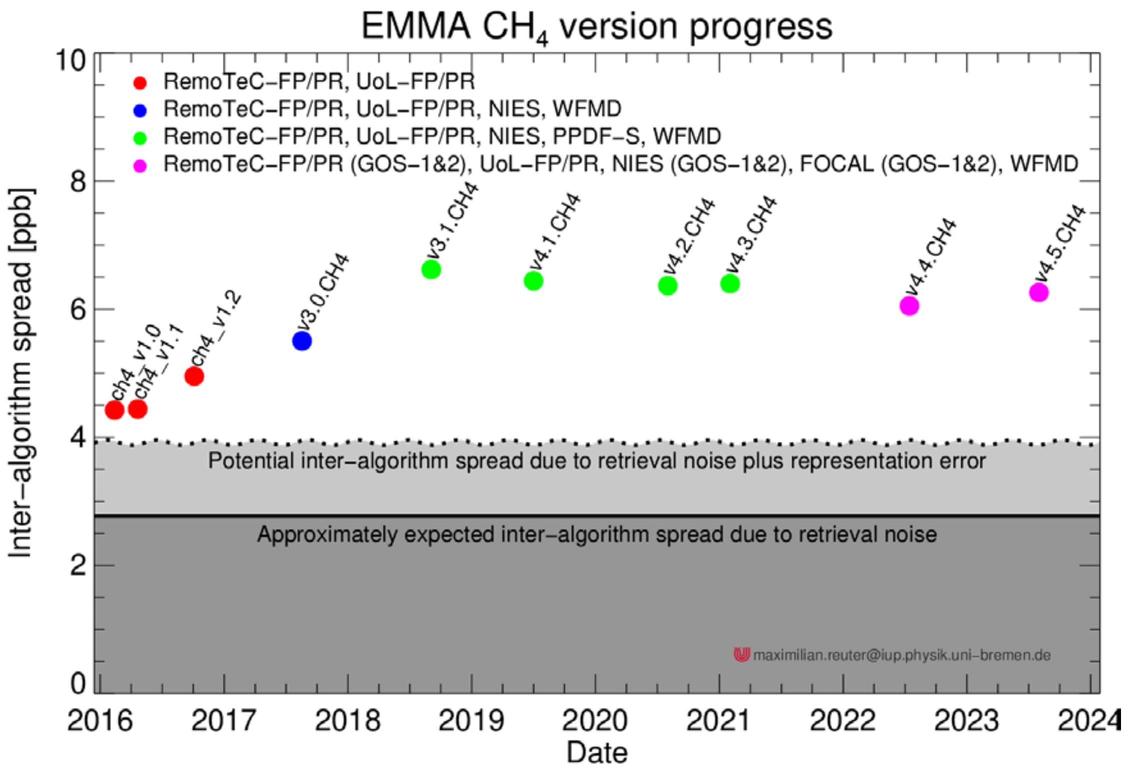


Figure 20: Average inter-algorithm spread of all EMMA versions compared with the approximately expected contribution of retrieval noise and a rough estimate of the representation error.

2.2.2 Summary

The validation results are summarized in Table 6.

Table 6: Product Quality Summary Table for product XCH4_EMMA. The listed requirements are the threshold (T) requirements as given in TRD (D3). For precision (i.e., single observation statistical uncertainty or random error) also the corresponding breakthrough (B) and goal (G) requirements are listed. For the achieved performance of (relative) "Accuracy" two values are listed: The first one is the spatial component of the bias and the second one is the spatio-temporal bias, computed by also considering seasonal biases. The spatio-temporal bias is our estimate of "relative accuracy". TR refers to "target requirement" and reported is the probability that the corresponding TR is met, i.e., the probabilities that accuracy is better than 0.5 ppm and stability is better than 0.5 ppm/year. 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: v4.5, Time period covered: 01.2003 – 12.2022				
Parameter [unit]	Achieved performance	Requirement	TR	Comments
Single measurement precision (1-sigma) in [ppb]	13.51	< 34 (T) < 17 (B) < 9 (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 [ppb]	-0.66	-	-	No requirement but value close to zero expected for a high quality data product.
Accuracy: Relative systematic error [ppb]	Spatial – spatiotemporal: 1.79 – 2.59	< 10	Probability that accuracy TR is met: 95%	Only six TCCON sites fulfilled all validation criteria. Therefore, estimates of these systematic error components are potentially less reliable.
Stability: Linear bias trend [ppb/year]	0.11±0.22 (1-sigma)	< 3	Probability that stability TR is met: 100%	
Stability: Year-to-year bias variability [ppb/year]	1.8 (1-sigma)	< 3	-	

2.3 XCO2_OBS4MIPS

The TCCON-validation of the XCO2_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 XCO2_OBS4MIPS can be summarized as follows:

The overall monthly mean uncertainty is 1 ppm and the mean bias is 0.39 ppm. Relative systematic error, i.e., the spatio-temporal bias, is 0.5±0.6 ppm (1-sigma). The computed linear drift of 0.09±0.23 ppm (1-sigma) is small and not significant. The probability that the 0.5 ppm accuracy requirement is met is 66%. The probability that the 0.5 ppm/year stability requirement is met is 97%. Overall, this product has therefore reasonable accuracy and high stability.

2.4 XCH4_OBS4MIPS

The TCCON-validation of the XCH4_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 XCH4_OBS4MIPS can be summarized as follows:

The overall monthly mean uncertainty is 8.1 ppb and the mean bias is -0.55 ppb. Relative systematic error, i.e., the spatio-temporal bias, is 4.7±6 ppb (1-sigma). The computed linear drift of 0.68±1.1 ppb (1-sigma) is small and not significant. The probability that the 10 ppb accuracy requirement is met is 89%. The probability that the 3 ppb/year stability requirement is met is 98%. Overall, this product has therefore very good accuracy and high stability.

3. Application(s) specific assessments

As also mentioned in the main PQAR (D1), the new data product described and validated in this annex have not yet been used for application specific assessments in terms of peer-reviewed publications or equivalent documentation. However, the main PQAR (D1) describes results from analyses of XCO2 and XCH4 annual mean growth rates using previous data sets and the most recent OBS4MIPS data sets.

In addition, the EMMA algorithm routinely assesses the inter-algorithm spread as described in the EMMA algorithm theoretical basis document (D2). The inter-algorithm spread or ensemble spread is defined as the inter-algorithm standard deviation and can be interpreted as the uncertainty due to potential regional or temporal retrieval biases. EMMA also provides this information to the user on a sounding-by-sounding basis as described in the corresponding product user guide and specification document (D4). Large scale regional or temporal biases can hamper global surface flux inversions, so in particular this application may benefit from making use of EMMA's inter-algorithm spread estimates. Examples of EMMA's inter-algorithm spread and the expected inter-algorithm spread due to measurement noise are shown in D2.

4. Compliance with user requirements

A detailed comparison of validation results from two independent validation experiments with the user requirements is part of the main PQAR (D1) and here we only summarize its most important results.

EMMA's XCO2 single measurement precision is better than the breakthrough requirement of 3 ppm but worse than the goal requirement of 1 ppm. The XCO2_EMMA data set meets the threshold requirement of 0.5 ppm for systematic errors (relative accuracy) with a likelihood of 71%. Its temporal stability is very good and the probability that the threshold XCO2 stability requirement of 0.5 ppm/year is met amounts 98%.

EMMA's XCH4 single measurement precision is close to the breakthrough requirement of 17 ppb. The XCH4_EMMA data set meets the threshold requirement of 5 ppb for systematic errors (relative accuracy) with a likelihood of 95%. Its temporal stability is very good and the probability that the threshold XCH4 stability requirement of 3 ppb/year is met amounts 100%.

References

Boesch and Anand, 2017: H. Boesch and J. Anand, Algorithm Theoretical Basis Document (ATBD) – ANNEX A for products CO2_GOS_OCFP, CH4_GOS_OCFP & CH4_GOS_OCPR, Copernicus Climate Change Service (C3S) project on satellite-derived Essential Climate Variable (ECV) Greenhouse Gases (CO_2 and CH_4) data products (project C3S_312a_Lot6), Version 1 (21/08/2017), 2017.

Buchwitz, M., et al., 2023: Product Quality Assessment Report (PQAR) – Main document for Greenhouse Gas (GHG: CO_2 & CH_4) data set CDR 7 (2003-2022), project C3S2_312a_Lot2_DLR – Atmosphere, v7.0, 2023.

Butz et al., 2011: Butz, A., Guerlet, S., Hasekamp, O., Schepers, D., Galli, A.,Aben, I., Frankenberg, C., Hartmann, J.-M., Tran, H., Kuze,A., Keppel-Aleks, G., Toon, G., Wunch, D., Wennberg, P., Deutscher, N., Griffith, D., Macatangay, R., Messerschmidt, J., Notholt, J., and Warneke, T.: Toward accurate CO2 and CH4 observations from GOSAT, Geophys. Res. Lett., 38, L14812, https://doi.org/10.1029/2011GL047888, 2011

CMUG-RBD, 2010: Climate Modelling User Group Requirements Baseline Document, Deliverable 1.2, Number D1.2, Version 1.3, 2 Nov 2010.

Cogan et al., 2012: 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., Warneke, T., and Wunch, D.: 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., 117, D21301, https://doi.org/10.1029/2012JD018087, 2012

Detmers, 2017a: R. Detmers, Algorithm Theoretical Basis Document (ATBD) – ANNEX B for products CO2_GOS_SRFP & CH4_GOS_SRFP, Copernicus Climate Change Service (C3S) project on satellite-derived Essential Climate Variable (ECV) Greenhouse Gases (CO_2 and CH_4) 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 CH4_GOS_SRPR, Copernicus Climate Change Service (C3S) project on satellite-derived Essential Climate Variable (ECV) Greenhouse Gases (CO_2 and CH_4) data products (project C3S_312a_Lot6), Version 1 (21/08/2017), 2017

JCGM, 2008: JCGM/WG 1, Working Group 1 of the Joint Committee for Guides in Metrology, Evalutation of measurement data – Guide to the expression of uncertainty in measurement, http://www.bipm.org/utils/common/documents/jcgm/JCGM_100_2008_E.pdf, 2008.

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_2 affected by erroneous surface pressure estimates, Atmos. Meas. Tech., 12, 2241–2259, https://doi.org/10.5194/amt-12-2241-2019, 2019

Krisna et al., 2021: Trismono Candra Krisna, Ilse Aben, Lianghai Wu, Otto Hasekamp, Jochen Landgraf: ESA Climate Change Initiative "Plus" (CCI+) Algorithm Theoretical Basis Document (ATBD)

Version 1.3 – For the RemoTeC XCO2 and XCH4 GOSAT-2 SRON Full Physics Products (CO2_GO2_SRFP and CH4_GO2_SRFP) Version 2.0.0 for the Essential Climate Variable (ECV) Greenhouse Gases (GHG), https://www.iup.uni-bremen.de/carbon_ghg/docs/GHG-CCIplus/CRDP7/ATBDv3_GHG-CCI_CO2_CH4_GO2_SRFP_v2p0p0.pdf, 2021

Noël et al., 2022: S. Noël, M. Reuter, M. Buchwitz, J. Borchardt, M. Hilker, O. Schneising, H. Bovensmann, J.P. Burrows, A. Di Noia, R.J. Parker, H. Suto, Y. Yoshida, M. Buschmann, N.M. Deutscher, D.G. Feist, D.W.T. Griffith, F. Hase, R. Kivi, C. Liu, I. Morino, J. Notholt, Y.-S. Oh, H. Ohyama, C. Petri, D.F. Pollard, M. Rettinger, C. Roehl, C. Rousogenous, M.K. Sha, K. Shiomi, K. Strong, R. Sussmann, Y. Té, V.A. Velazco, M. Vrekoussis, and T. Warneke: Retrieval of greenhouse gases from GOSAT and GOSAT-2 using the FOCAL algorithm, Atmos. Meas. Tech., 15, 3401-3437, https://doi.org/10.5194/amt-15-3401-2022, 2022

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., Oyafuso, F., Polonsky, I., Smyth, M., Taylor, T., Toon, G. C.,Wennberg, P. O., andWunch, D.: The ACOS CO2 retrieval algorithm – Part 1: Description and validation against synthetic observations, Atmos. Meas. Tech., 5, 99–121, doi:10.5194/amt-5-99-2012, 2012

Reuter et al., 2010: M. Reuter, M. Buchwitz, O. Schneising, J. Heymann, H. Bovensmann, J. P. Burrows: A method for improved SCIAMACHY CO2 retrieval in the presence of optically thin clouds. Atmospheric Measurement Techniques, 3, 209-232, 2010

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. Velazco, T. Warneke, P. O. Wennberg, D. Wunch: Retrieval of atmospheric CO2 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_2 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., 2016: M. Reuter, H. Bovensmann, M. Buchwitz, J. P. Burrows, J. Heymann, O. Schneising: Algorithm Theoretical Basis Document Version 5 (ATBDv5) - The Bremen Optimal Estimation DOAS (BESD) algorithm for the retrieval of XCO2 for the Essential Climate Variable (ECV) Greenhouse Gases (GHG), 2016

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_2 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_2 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. Boesch, A. Di Noia, J. Anand, R.J. Parker, P. Somkuti, L. Wu, O.P. Hasekamp, I. Aben, A. Kuze, H. Suto, K. Shiomi, Y. Yoshida, I. Morino, D. Crisp, C.W. O'Dell, J. Notholt, C. Petri, T. Warneke, V.A. Velazco, N.M. Deutscher, D.W.T. Griffith, R. Kivi, D.F. Pollard, F. Hase, R. Sussmann, Y.V. Té, K. Strong, S. Roche, M.K. Sha, M. De Mazière, D.G. Feist, L.T. Iraci, C.M. Roehl, C. Retscher, and D. Schepers: Ensemble-based satellite-derived carbon dioxide and methane column-averaged dry-air mole fraction data sets (2003-2018) for carbon and climate applications, Atmos. Meas. Tech., https://www.atmos-meas-tech.net/13/789/2020, 2020.

Reuter et al., 2021: M. Reuter, M. Hilker, S. Noël, M. Buchwitz, O. Schneising, H. Bovensmann, and J. P. Burrows: ESA Climate Change Initiative “Plus” (CCI+) Algorithm Theoretical Basis Document Version 3 (ATBDv3) - Retrieval of XCO2 from the OCO-2 satellite using the Fast Atmospheric Trace Gas Retrieval (FOCAL) for the Essential Climate Variable (ECV) Greenhouse Gases (GHG), http://www.iup.uni-bremen.de/carbon_ghg/docs/GHG-CCIplus/CRDP7/ATBDv3_GHG-CCI_CO2_OC2_FOCA_v10.pdf, 2021.

Reuter, et al., 2023a: Algorithm Theoretical Basis Document (ATBD) – ANNEX D for products XCO2_EMMA, XCH4_EMMA, XCO2_OBS4MIPS, XCH4_OBS4MIPS (v4.5, 01/2003-12/2022), project C3S2_312a_Lot2_DLR – Atmosphere, v7.0, 2023.

Reuter, et al., 2023b: Product User Guide and Specification (PUGS) – ANNEX D for products XCO2_EMMA, XCH4_EMMA, XCO2_OBS4MIPS, XCH4_OBS4MIPS (v4.5, 01/2003-12/2022) C3S_312a_Lot2_DLR – Atmosphere, v7.0, 2023.

Schneising et al., 2018: O. Schneising and the ESA CCI GHG project team: ESA Greenhouse Gases Climate Change Initiative (GHG_cci): Column-averaged CH4 from SCIAMACHY generated with the WFMD algorithm (CH4_SCI_WFMD), version 4.0. Centre for Environmental Data Analysis, date of citation. https://catalogue.ceda.ac.uk/uuid/aa09603e91b44f3cb1573c9dd415e8a8, 2018

Taylor et al., 2022: Taylor, T. E., O’Dell, C. W., Crisp, D., Kuze, A., Lindqvist, H., Wennberg, P. O., Chatterjee, A., Gunson, M., Eldering, A., Fisher, B., Kiel, M., Nelson, R. R., Merrelli, A., Osterman, G., Chevallier, F., Palmer, P. I., Feng, L., Deutscher, N. M., Dubey, M. K., Feist, D. G., García, O. E., Griffith, D. W. T., Hase, F., Iraci, L. T., Kivi, R., Liu, C., De Mazière, M., Morino, I., Notholt, J., Oh, Y.-S., Ohyama, H., Pollard, D. F., Rettinger, M., Schneider, M., Roehl, C. M., Sha, M. K., Shiomi, K., Strong, K., Sussmann, R., Té, Y., Velazco, V. A., Vrekoussis, M., Warneke, T., and Wunch, D.: An 11-year record of XCO2 estimates derived from GOSAT measurements using the NASA ACOS version 9 retrieval algorithm, Earth Syst. Sci. Data, 14, 325–360, https://doi.org/10.5194/essd-14-325-2022, 2022

Wunch et al., 2011: Wunch, D., Toon, G. C., Blavier, J.-F. L., Washenfelder, R. A., Notholt, J., Connor, B. J., Griffith, D. W. T., Sherlock, V., and Wennberg, P. O.: The Total Carbon Column Observing Network (TCCON), Philosophical Transactions of the Royal Society of London, Series A:

Mathematical, Physical and Engineering Sciences, 369, 2087–2112, doi:10.1098/rsta.2010.0240, 2011

Yoshida et al., 2013: Yoshida, Y., Kikuchi, N., Morino, I., Uchino, O., Oshchepkov, S., Bril, A., Saeki, T., Schutgens, N., Toon, G. C., Wunch, D., Roehl, C. M., Wennberg, P. O., Griffith, D. W. T., Deutscher, N. M., Warneke, T., Notholt, J., Robinson, J., Sherlock, V., Connor, B., Rettinger, M., Sussmann, R., Ahonen, P., Heikkinen, P., Kyrö, E., Mendonca, J., Strong, K., Hase, F., Dohe, S., and Yokota, T.: Improvement of the retrieval algorithm for GOSAT SWIR XCO2 and XCH4 and their validation using TCCON data, Atmos. Meas. Tech., 6, 1533–1547, https://doi.org/10.5194/amt-6-1533-2013, 2013

Yoshida and Oshio, 2020: Y. Yoshida and H. Oshio: GOSAT-2 TANSO-FTS-2 SWIR L2 Retrieval Algorithm Theoretical Basis Documen, National Institute for Environmental Studies, GOSAT-2 Project https://prdct.gosat-2.nies.go.jp/documents/pdf/ATBD_FTS-2_L2_SWL2_en_00.pdf, 2020

Copernicus Climate Change Service

ECMWF - Shinfield Park, Reading RG2 9AX, UK

Contact: https://support.ecmwf.int/