ECMWF COPERNICUS REPORT

Copernicus Climate Change Service

Product Quality Assessment Report (PQAR) – ANNEX D for products XCO2_EMMA and XCH4_EMMA

C3S_312a_Lot6_IUP-UB – Greenhouse Gases

Issued by: Maximilian Reuter, University of Bremen,

Institute of Environmental Physics (IUP)

Date: 20/10/2017

Ref: C3S_D312a_Lot6.3.1.7-v1_PQAR_ANNEX-D_v1.1

Official reference number service contract: 2016/C3S_312a_Lot6_IUP-UB/SC1

IMPLEMENTED BY ECMWF

Contributors

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

(IUP)

M. Reuter

M. Buchwitz

O. Schneising-Weigel

Table of Contents

History of modifications

Version	Date	Description of modification	Chapters / Sections
1.0	28-September-2017	New document	All
1.0b	11-October-2017	Page header logo replaced	Page header
1.1	20-October-2017	KPI replaced by TR	All

Related documents

Reference ID	Document
D1	Main PQAR: Buchwitz, M., et al., Product Quality Assessment Report (PQAR) – Main document, C3S project C3S_312a_Lot6_IUP-UB – Greenhouse Gases, v1.1, 2017. *(this document is an ANNEX to the Main PQAR)*
D2	ATBD ANNEX D: Reuter, M., et al., Algorithm Theoretical Basis Document (ATBD) – ANNEX D for products XCO2_EMMA, XCH4_EMMA, C3S project C3S_312a_Lot6_IUP-UB – Greenhouse Gases, 2017.

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 *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 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
XCO2	Column-averaged dry-air mixing ratios (mole fractions) of CO_2
XCH4	Column-averaged dry-air mixing ratios (mole fractions) of CH_4
L1	Level 1 satellite data product: geolocated radiance (spectra)
L2	Level 2 satellite-derived data product: Here: XCO2 and XCH4 information for each ground-pixel
L3	Level 3 satellite-derived data product: Here: Gridded XCO2 and XCH4information, e.g., 5°x5°, monthly
L4	Level 4 satellite-derived data product: Here: Surface fluxes (emission and/or uptake) of CO_2 and CH_4

Scope of document

This document is a Product Quality Assessment Report (PQAR) for the Copernicus Climate Change Service (C3S, https://climate.copernicus.eu/) component as covered by project C3S_312a_Lot6 led by University of Bremen, Germany.

Within project C3S_312a_Lot6 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 C3S_312a_Lot 6 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 / quality assessment of the C3S products XCO2_EMMA and XCH4_EMMA.

These products are merged multi-sensor XCO_2 and XCH_4 Level 2 products generated using algorithms developed at University of Bremen, Germany.

Executive summary

This Product Quality Assessment Report (PQAR) describes the validation of the EMMA v3.0 CO2 and EMMA v3.0 CH4 products (in the following also referred to as XCO2_EMMA and XCH4_EMMA) with TCCON ground based measurements. Originally, the EMMA algorithm (v1.3) was described and validated in the publication of *Reuter et al. (2013)* and their validation is the blueprint for this PQAR. Several of the shown figures are updated versions of figures shown in the publication of *Reuter et al. (2013)*. EMMA is composed of an ensemble of individual SCIAMACHY and GOSAT L2 algorithms and this document also contributes to the inter-comparison of the contributing algorithms.

For XCO2 we find that the individual algorithms have a single measurement precision in the range of 1.70ppm (ACOS) to 2.20ppm (NIES). EMMA has a single measurement precision of 1.89ppm. EMMA's combined regional and seasonal biases (0.56ppm) are on the lower end of the range of the individual algorithms (0.58ppm for BESD to 0.85ppm for NIES). The found linear drifts are small and predominantly not significant, i.e., the trend is usually smaller than twice its uncertainty. The linear drift found for EMMA's XCO2 is -0.04±0.08ppm/a. The year-to-year stability is in the range of 0.28ppm/a (ACOS) and 0.50ppm/a (NIES). EMMA's year-to-year stability of 0.38ppm/a lies within this range.

For XCH4 we find that the individual algorithms have a single measurement precision in the range of 13.0ppb (UoL-PR) to 13.8ppb (RemoTeC-PR) except for WFMD which has a single measurement precision of 90.4ppb. EMMA has a single measurement precision of 43.9ppb which is larger than most of the individual algorithms due to the WFMD contribution till 2010. EMMA's combined regional and seasonal biases (5.52ppb) are within the range of the individual GOSAT algorithms (4.43ppb for RemoTeC-PR to 5.68ppb for NIES) even though WFMD (10.1ppb) is part of EMMA. 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 XCH4 is -0.11±0.89ppb/a. The year-to-year stability is in the range of 1.3ppb/a (UoL-PR) and 2.1ppb/a (NIES) for the GOSAT algorithms and 11.1ppb/a for WFMD. As the first years include only WFMD, EMMA's overall year-to-year stability of 5.9ppb/a is worse than the stability of the GOSAT algorithms.

1. Product validation methodology

As described in *D2*, EMMA v3.0 CO2 and CH4 make use of the following satellite data products: BESD v02.01.02 (*Reuter et al., 2016*), RemoTeC v2.3.8 (*Detmers et al., 2017a*), ACOS v7.3.10a (*O'Dell et al., 2012*), UoL-FP v7.1 (*Boesch and Anand, 2017*), and NIES v02 (*Yoshida et al., 2013*) for XCO2 and WFMD v4.0 (*Schneising et al., 2016*), RemoTeC-FP v2.3.8 (*Detmers et al., 2017a*), RemoTeC-PR v2.3.8 (*Detmers et al., 2017b*), UoL-FP v7.1 (*Boesch and Anand, 2017*), UoL-PR v7.0 (*Boesch and Anand, 2017*), and NIES v02 (*Yoshida et al., 2011*) for XCH4.

The EMMA CO2 and CH4 data products are validated with TCCON (using version GGG2014 R0/1) 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 500km, and a maximal surface elevation difference of 250m.

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 01/2003 to 12/2016.

The station bias is defined as average difference to TCCON and the single measurement precision as standard deviation of the difference to TCCON. Seasonal bias and linear drift are derived by fitting the following trend model:

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

Here, ΔX represents the difference satellite minus TCCON, ε the residual, and a_{0-3} the free fit parameters. Specifically, a_1 represents the linear drift and a_2 the amplitude of the seasonal bias.

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 (standard deviation of the seasonal bias term), and average linear drift. As the linear drift can be assumed to be globally constant, the station-to-station standard deviation of the linear drift is a measure for its uncertainty.

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.

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°x10° monthly gridded level 3 data sets.

We calculate the fraction of potential outliers according to unrealistically large spatial gradients (>3ppm/10° for XCO2 and >20ppb/10° for XCH4), unrealistically large deviations from CarbonTracker CT2016 (>3ppm for XCO2 and >20ppb/10° for XCH4), and larger deviations from EMMA (>2.5ppm for XCO2 and >10ppb/10° for XCH4).

We also compare the north/south (N/S) gradient of each month with CT2016 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 CT2016 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.

2. Validation Results

2.1 XCO2_EMMA

2.1.1 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 v3.0 CO2, i.e., the XCO2_EMMA product. The results are valid for the time periods covered by the individual algorithms or the EMMA time period, respectively.

The individual algorithms have a single measurement precision in the range of 1.70ppm (ACOS) – 2.20ppm (NIES). EMMA has a single measurement precision of 1.89ppm. EMMA's combined regional and seasonal biases (0.56ppm) are on the lower end of the range of the individual algorithms (0.58ppm for BESD to 0.85ppm for NIES).

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 0.3ppm – 1.0ppm at the Sodankylä, Garmisch-Partenkirchen, and Karlsruhe TCCON sites and low bias of similar magnitude at 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 but consistent negative trends (typically below -0.1ppm/a) at the sites Lamont and Wollongong and positive trends of the same magnitude at Orleans and Bialystok. This is a bit surprising because three 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 0.28ppm/a (ACOS) and 0.50ppm/a (NIES). EMMA's year-to-year stability of 0.38ppm/a lies within this range.

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**).

Except for ACOS, all algorithms see consistently larger seasonal amplitudes than CarbonTracker. The satellite retrieved seasonal amplitudes are generally in better agreement with TCCON than with CarbonTracker (**Figure 5**, top left).

Comparison of the north/south gradients show 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).

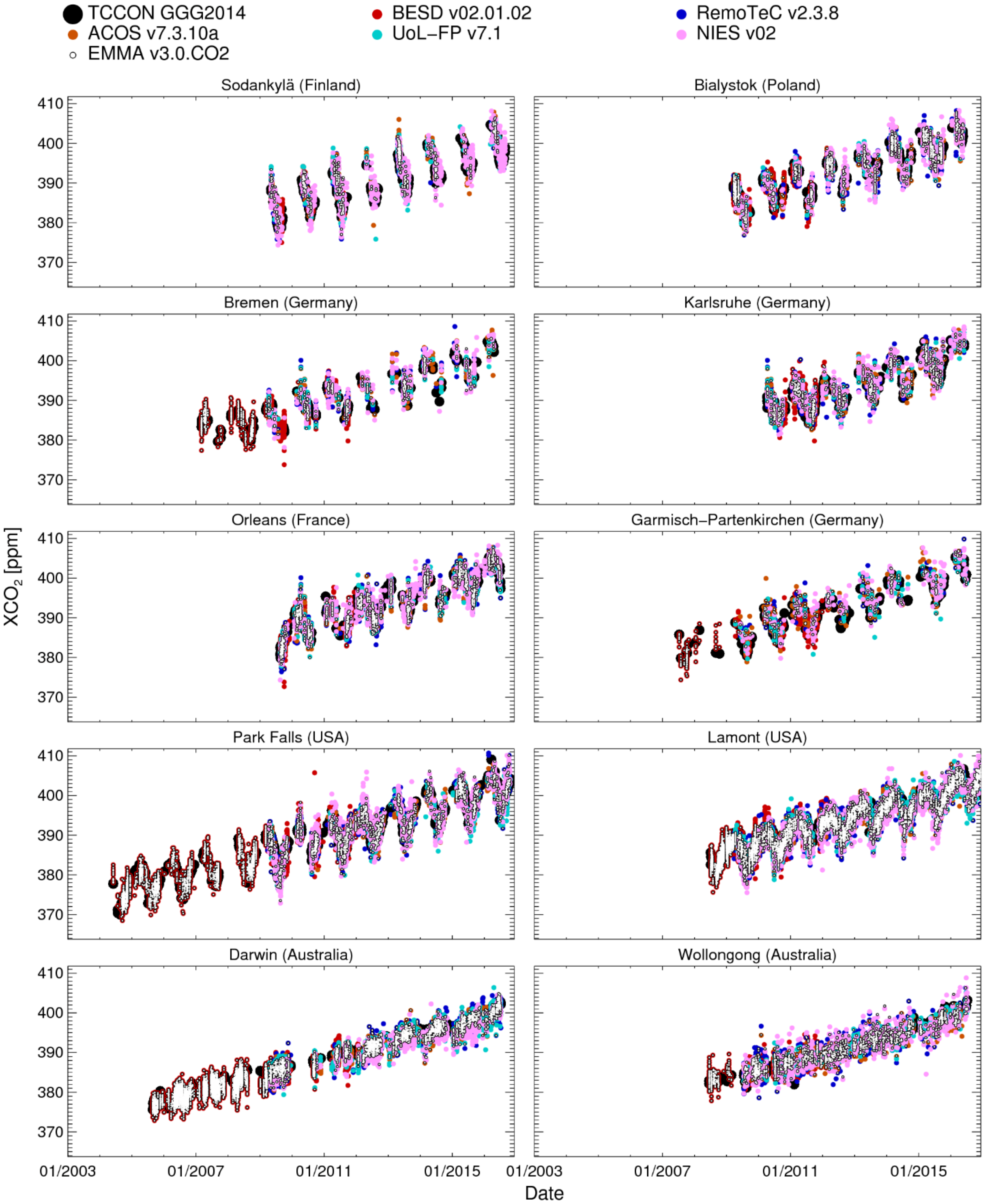


Figure 1: Validation of individual XCO_2 algorithms and EMMA v3.0 CO_2 with TCCON GGG2014.

Table 2: Summarizing 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.

Algorithm	#	Precision [ppm]	Accuracy [ppm]		Stability [ppm/a]		Probability that TR is met [%]	
			Regional	Seasonal	Trend	Year2Year	Accuracy	Stability
BESD v02.01.02	28577	1.91	0.37	0.45	0.15±0.27	0.34	40	82
RemoTeC v2.3.8	11168	2.04	0.53	0.40	-0.06±0.07	0.33	29	98
ACOS v7.3.10a	11988	1.70	0.63	0.41	-0.12±0.07	0.28	-	-
UoL-FP v7.1	9762	1.79	0.43	0.47	-0.13±0.04	0.31	33	96
NIES v02	11484	2.20	0.69	0.49	0.01±0.09	0.50	-	-
EMMA v3.0 CO2	23309	1.89	0.44	0.34	-0.04±0.08	0.38	43	98

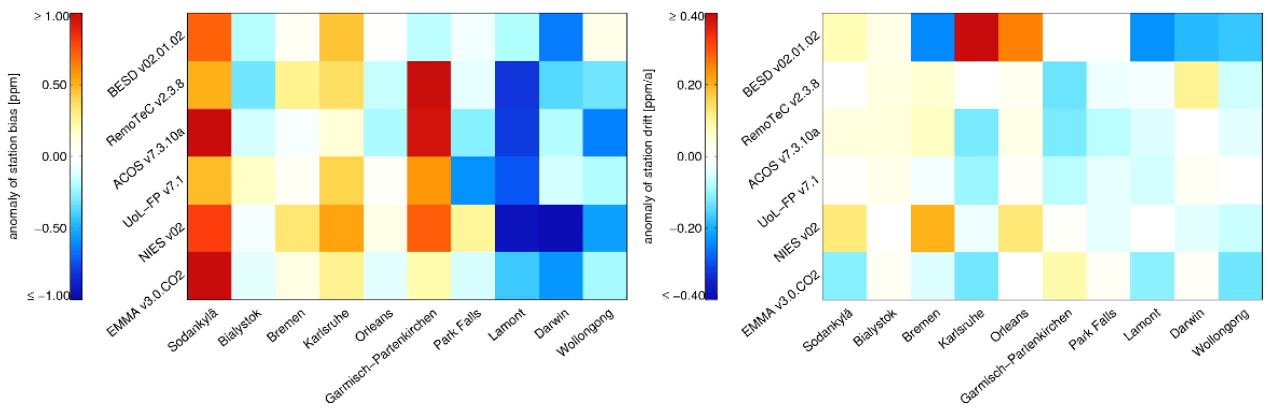


Figure 2: Anomaly of station biases (**left**) and station drift (**right**).

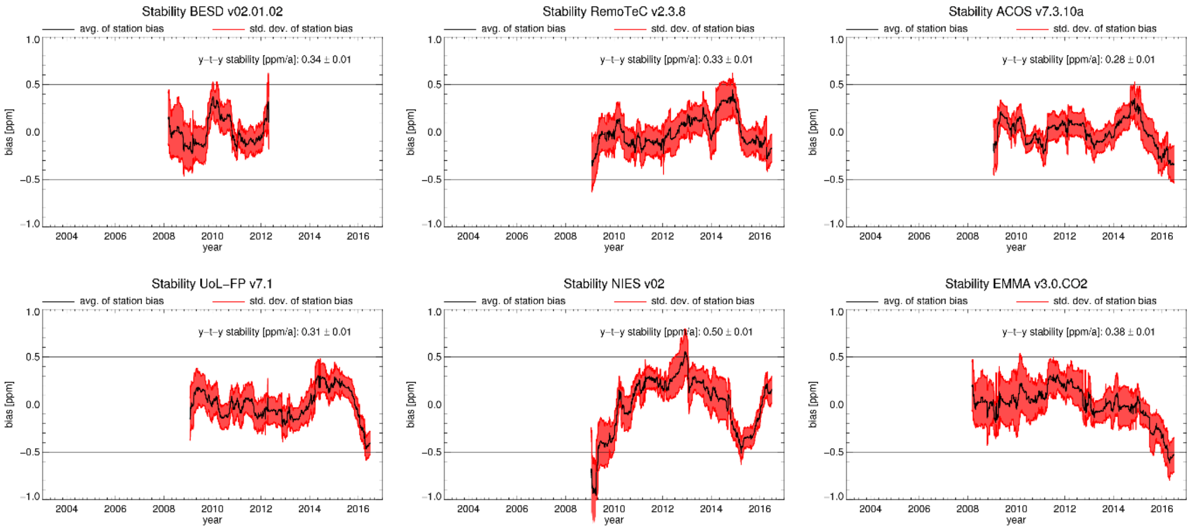


Figure 3: Stability analyses for EMMA and the contributing individual algorithms. The black curve shows the average station bias and the red curves its uncertainty represented by the station-to-station standard deviation and error propagation from single sounding measurement noise.

In terms of the frequency of potential outliers and standard deviation of the difference to CarbonTracker and TCCON, EMMA has a similar performance as the individual algorithms (**Figure 5**, bottom).

Once the GOSAT algorithms kick in in 2009, ACOS and RemoTeC usually provide the largest part of the integrated data weight in EMMA (**Figure 6**).

The average inter-algorithm spread has values between 0.4ppm and1.6 ppm and is typically below 0.9ppm (**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 (till now) the average inter-algorithm spread reduced with every new EMMA version (always including the most recent algorithm versions, **Figure 8**). This means that EMMA observes a kind of convergence among the individual algorithms. It is not entirely clear where this 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).

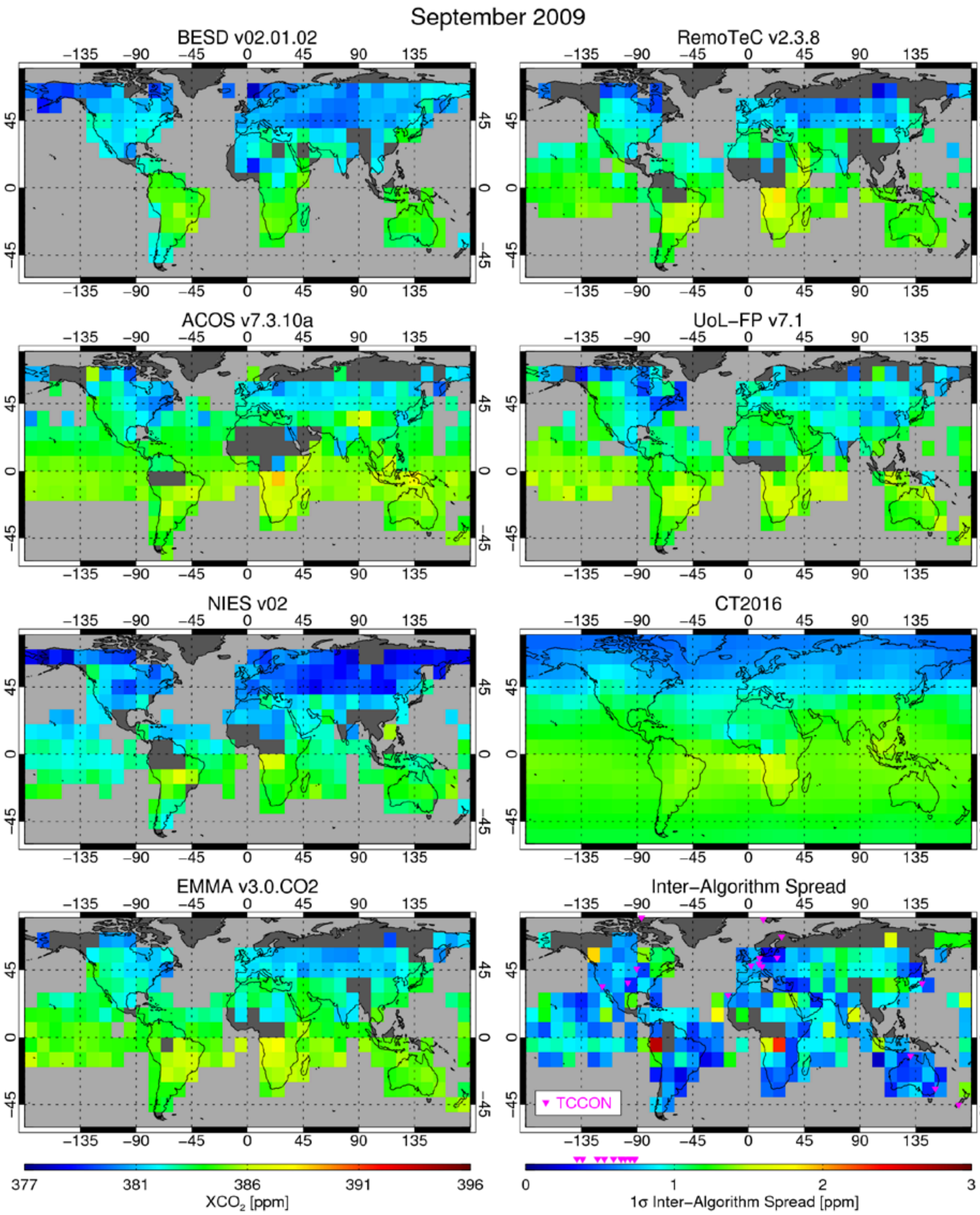


Figure 4: Monthly gridded XCO_2 averages and inter-algorithm spread at the example of September 2009 for EMMA v3.0 CO_2.

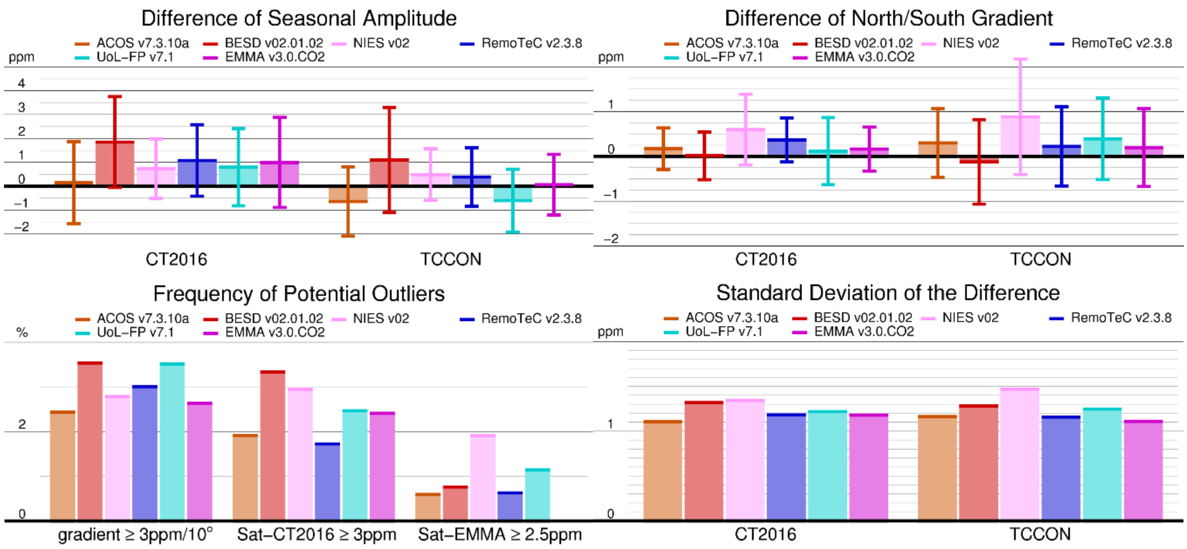


Figure 5: Top left: Difference of seasonal cycle amplitude of all individual algorithms as well as EMMA compared with CarbonTracker v2016 and TCCON GGG2014. **Top right:** Difference of north/south gradient of all individual algorithms as well as EMMA compared with CarbonTracker v2016 and TCCON. **Bottom 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.

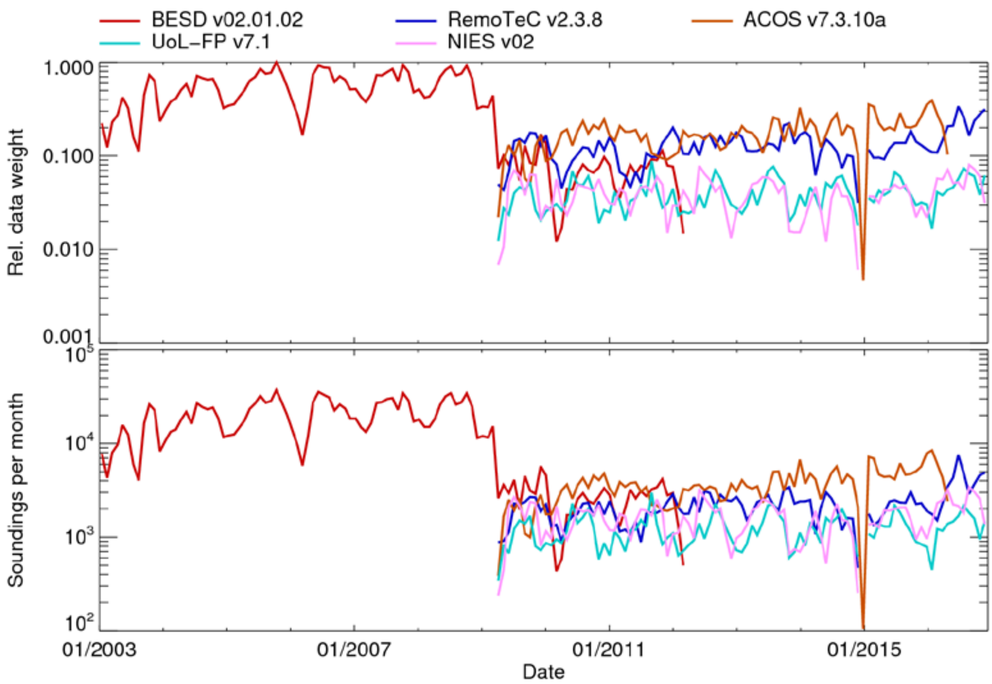


Figure 6: EMMA's normalized relative data weight proportional to $\sum 1/\sigma_i^2$ (**top**) and number of soundings (**bottom**) per algorithm and month.

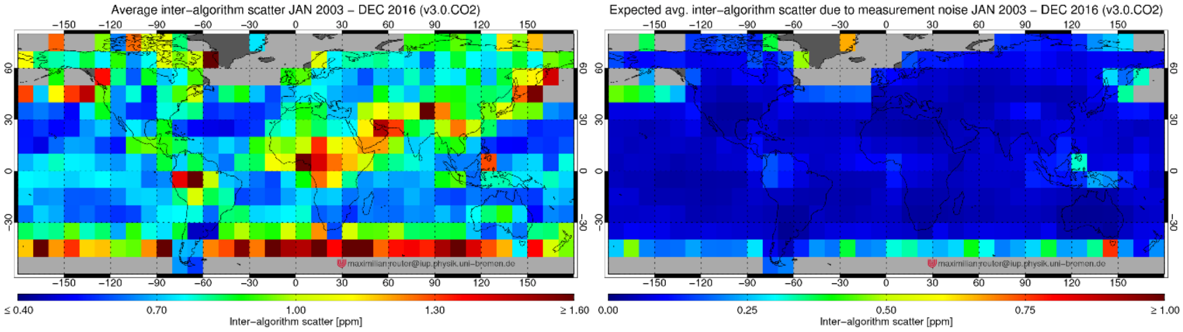


Figure 7: Average inter-algorithm scatter of monthly 10°x10° averages from January 2003 till December 2016 (left) and corresponding expected contribution of measurement noise (right).

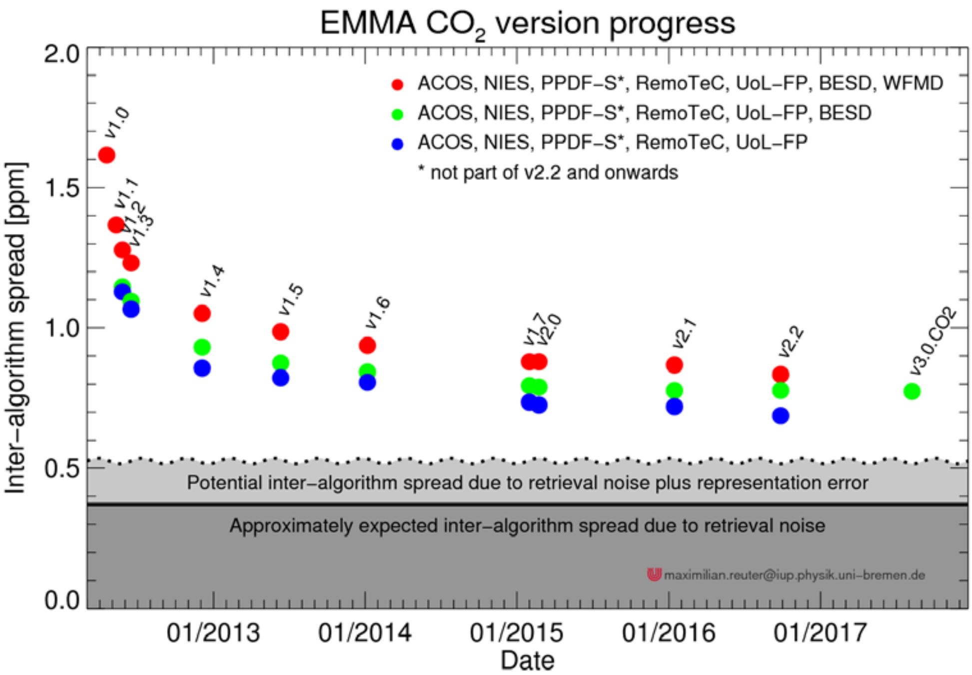


Figure 8: Average inter-algorithm spread of all EMMA versions compared with the approximately expected contribution of retrieval noise and a rough estimate of the representation error.

2.1.2 Summary

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

Table 3: Product Quality Summary Table for product XCO2_EMMA.

Product Quality Summary Table for Product: XCO2_EMMA Level: 2, Version: 3.0, Time period covered: 1.2003 – 12.2016				
Parameter [unit]	Achieved performance	Requirement	TR	Comments
Single measurement precision (1-sigma) in [ppm]	1.89	< 8 (T) < 3 (B) < 1 (G)	-	-
Uncertainty ratio in [-]: Ratio reported uncertainty to standard deviation of satellite-TCCON difference	0.97	-	-	No requirement but value close to unity expected for a high quality data product.
Mean bias [ppm]	0.47	-	-	No requirement but value close to zero expected for a high quality data product.
Accuracy: Relative systematic error [ppm]	Spatial – spatiotemporal: 0.44 – 0.56	< 0.5	Probability that accuracy TR is met: 43%	-
Stability: Drift [ppm/year]	-0.04 +/- 0.08 (1-sigma)	< 0.5	Probability that stability TR is met: 98%	-
Stability: Year-to-year bias variability [ppm/year]	0.38 (1-sigma)	< 0.5	-	-

2.2 XCH4_EMMA

2.2.1 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 v3.0 CH4 i.e., the XCH4_EMMA product. The results are valid for the time periods covered by the individual algorithms or the EMMA time period, respectively.

The individual algorithms have a single measurement precision in the range of 13.0ppb (UoL-PR) – 13.8ppb (RemoTeC-PR) except for WFMD which has a single measurement precision of 90.4ppb. EMMA has a single measurement precision of 43.9ppb which is larger than most of the individual algorithms due to the WFMD contribution till 2010 (**Figure 9**). EMMA's combined regional and seasonal biases (5.52ppb) are within the range of the individual GOSAT algorithms (4.43ppb for RemoTeC-PR to 5.68ppb for NIES) even though WFMD (10.1ppb) is part of EMMA.

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 4ppb at the Sodankylä, Garmisch-Partenkirchen, and Park Falls TCCON sites and low bias of similar magnitude at the southern hemispheric sites Darwin and Wollongong. This feature is similar to the feature observed for XCO2 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 -0.5ppb/a) at the sites Lamont and Wollongong and positive trends of the same magnitude at Bremen and Sodankylä. This is a bit surprising because three of the sites are located in similar latitude bands. This feature is similar to the feature observed for the drift of XCO2.

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 1.3ppb/a (UoL-PR) to 2.1ppb/a (NIES) for the GOSAT algorithms and 11.1ppb/a for WFMD. As the first years include only WFMD, EMMA's overall year-to-year stability of 5.9ppb/a is worse than the stability of the GOSAT algorithms.

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

All algorithms see a consistently larger north/south gradient than TCCON (**Figure 13**, left). Except for UoL-FP, the same is true for the seasonal amplitude (**Figure 13**, left). In terms of the standard deviation of the difference to TCCON, EMMA slightly under-performs compared to the GOSAT algorithms (**Figure 13**, left) which can be explained by the contribution of WFMD in the first years. The same applies to the frequency of potential outliers (**Figure 13**, right).

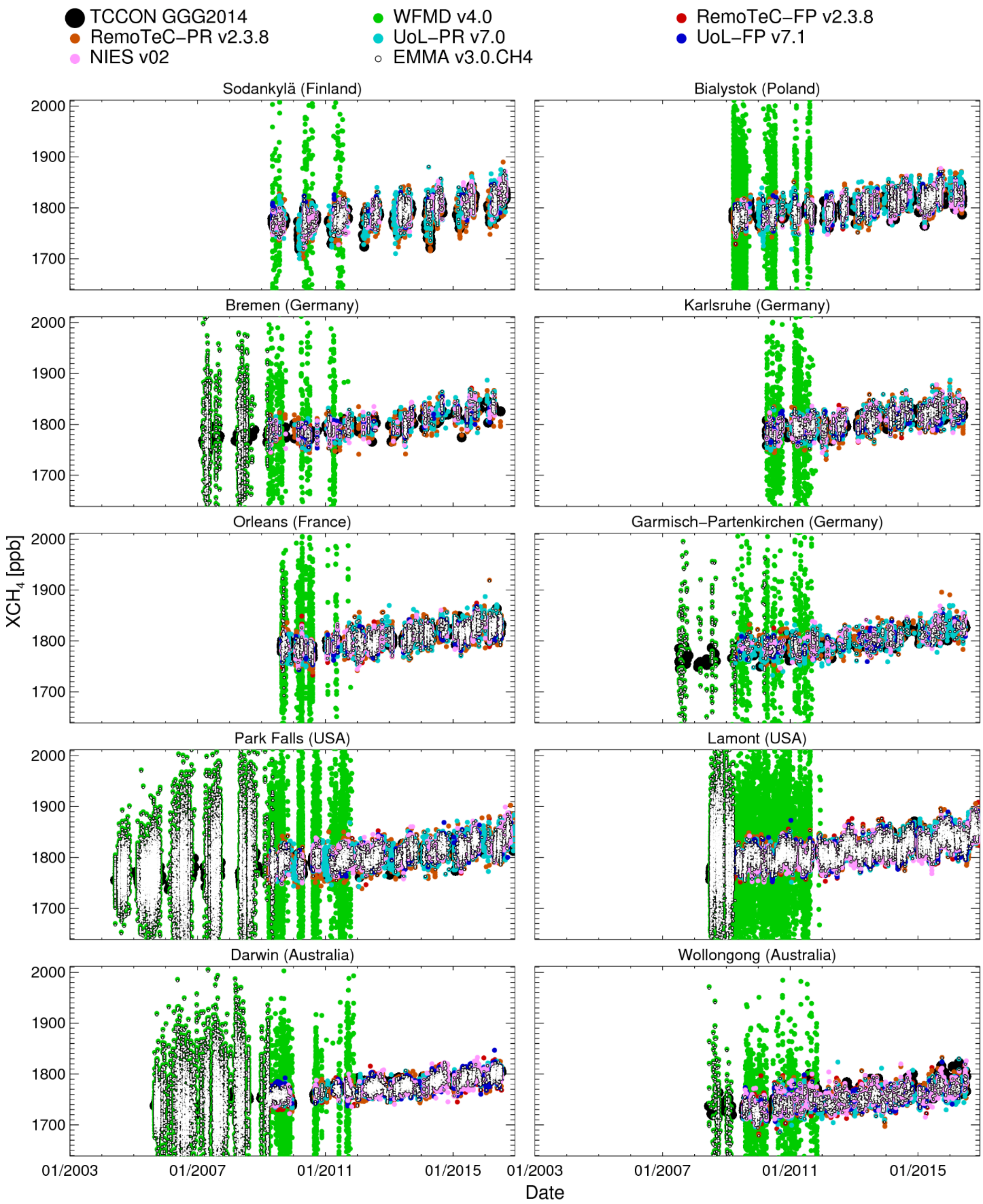


Figure 9: Validation of individual XCH_4 algorithms and EMMA v3.0 CH_4 with TCCON GGG2014.

Table 4: Summarizing 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.

Algorithm	#	Precision [ppb]	Accuracy [ppb]		Stability [ppb/a]		Probability that TR is met [%]	
			Regional	Seasonal	Trend	Year2Year	Accuracy	Stability
WFMD v4.0	54650	90.44	6.88	7.39	1.08±7.31	11.1	49	31
RemoTeC-FP v2.3.8	11168	13.39	3.19	3.11	-0.15±0.76	1.8	100	98
RemoTeC-PR v2.3.8	35293	13.76	2.49	3.66	-0.31±0.60	1.4	100	99
UoL-FP v7.1	28668	13.24	4.28	2.64	0.14±0.78	2.0	100	98
UoL-PR v7.0	9655	12.96	3.49	3.98	-1.01±0.75	1.3	100	94
NIES v02	11485	13.52	3.94	4.09	0.15±0.87	2.1	-	-
EMMA v3.0 CO2	48896	43.94	4.21	3.57	-0.11±0.89	5.9	100	97

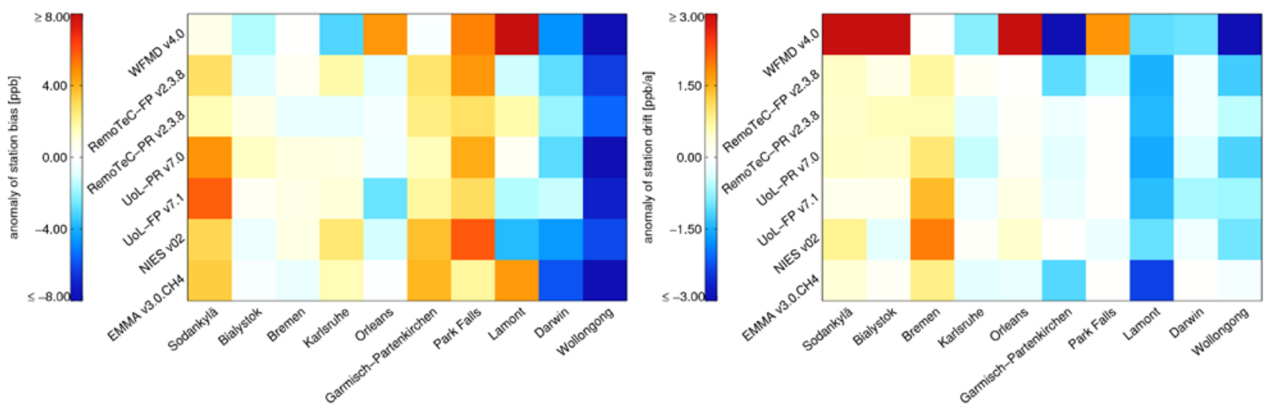
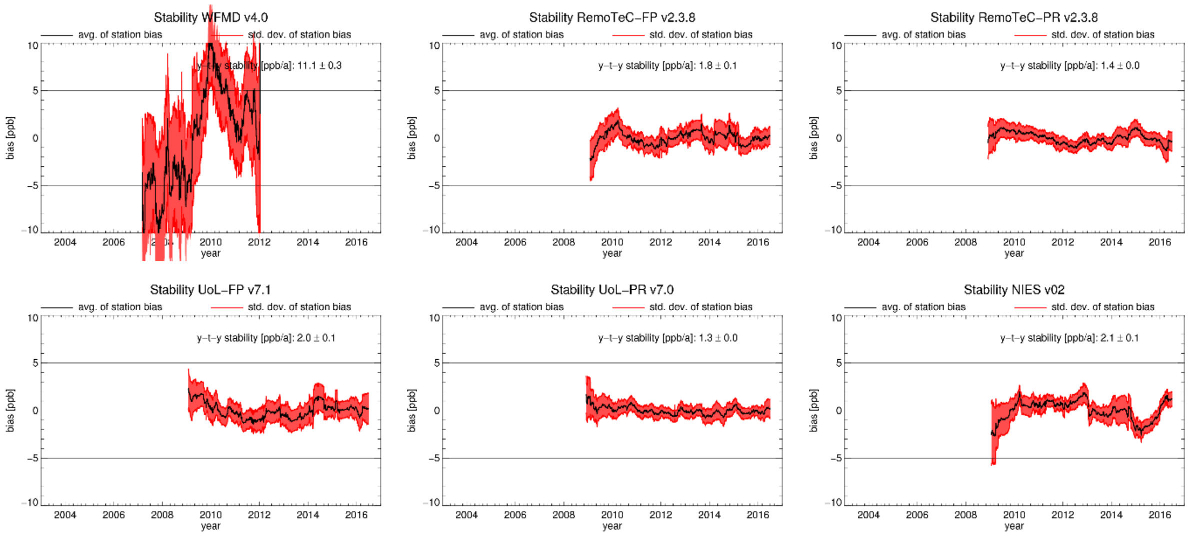


Figure 10: Anomaly of station biases (**left**) and station drift (**right**).



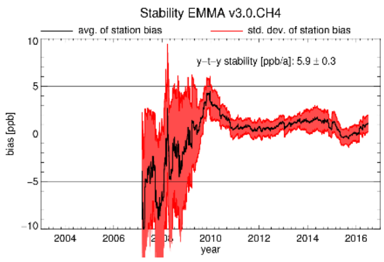


Figure 11: Stability analyses for EMMA and the contributing individual algorithms. The black curve shows the average station bias and the red curves its uncertainty represented by the station-to-station standard deviation and error propagation from single sounding measurement noise.

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

The average inter-algorithm spread has values between 2ppb and 12ppb and is typically below 8ppb (**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 with the last EMMA versions (always including the most recent algorithm versions, **Figure 16**).This can be explained with significant changes in the EMMA CH4 algorithm, and with the addition of the WFMD SCIAMACHY product.

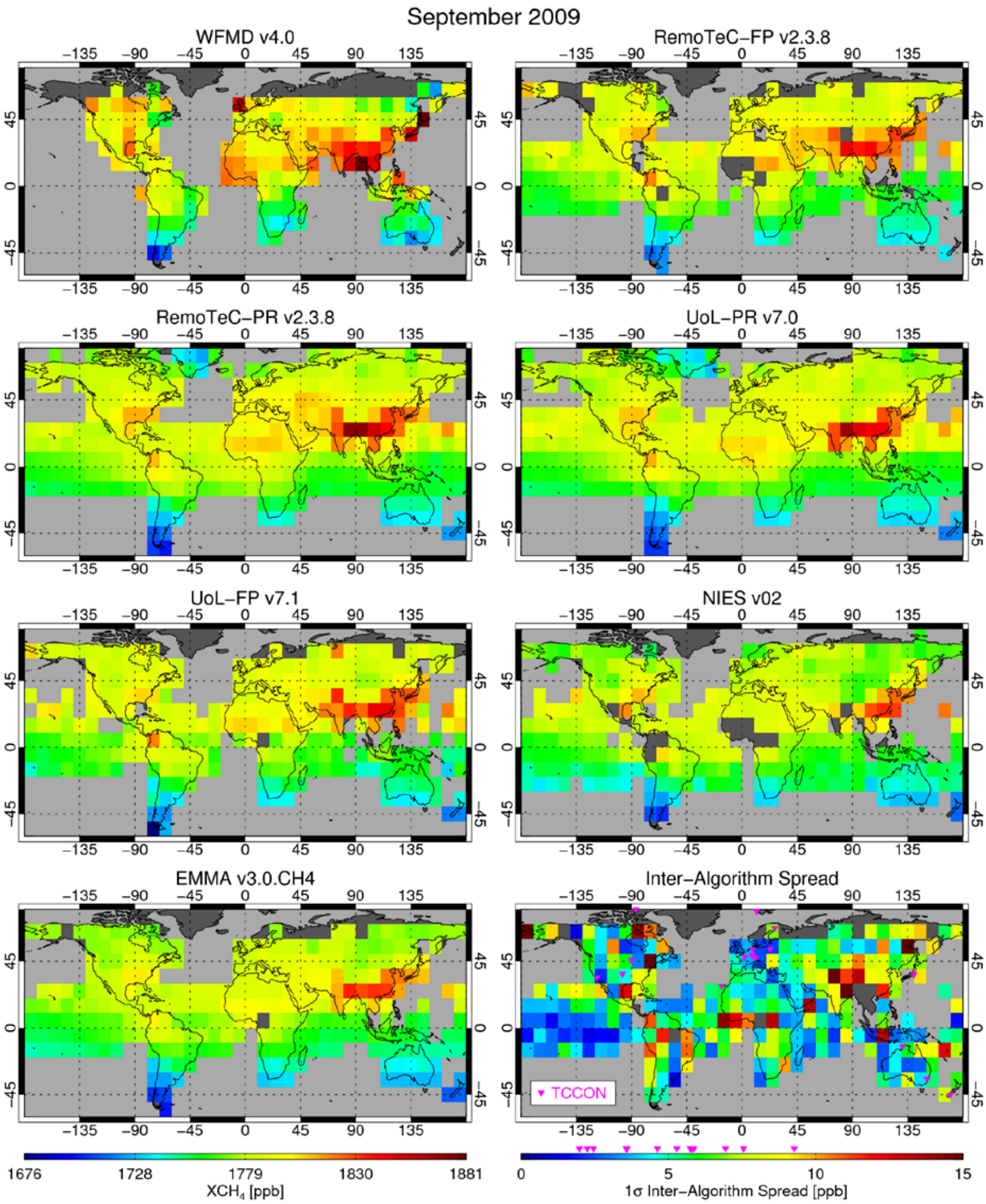


Figure 12: Monthly gridded XCH_4 averages and inter-algorithm spread at the example of September 2009 for EMMA v3.0 CH_4.

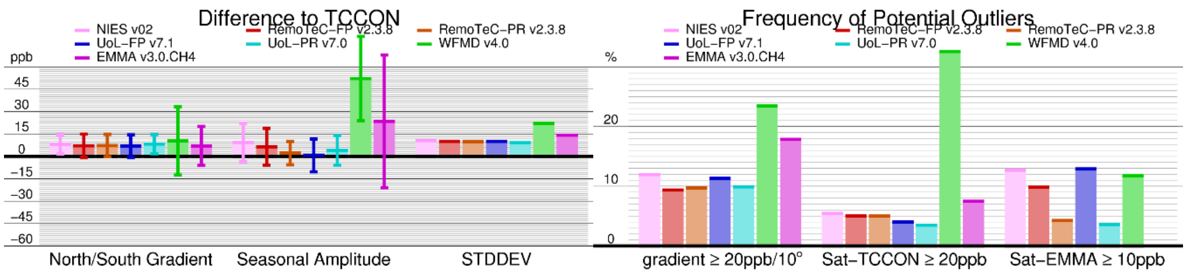


Figure 13: Left: Difference of the north/south gradient, difference of the seasonal cycle amplitude, and standard deviation of the difference in respect to TCCON for all individual algorithms as well as EMMA. **Right:** Frequency of potential outliers estimated by large gradients, large differences to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

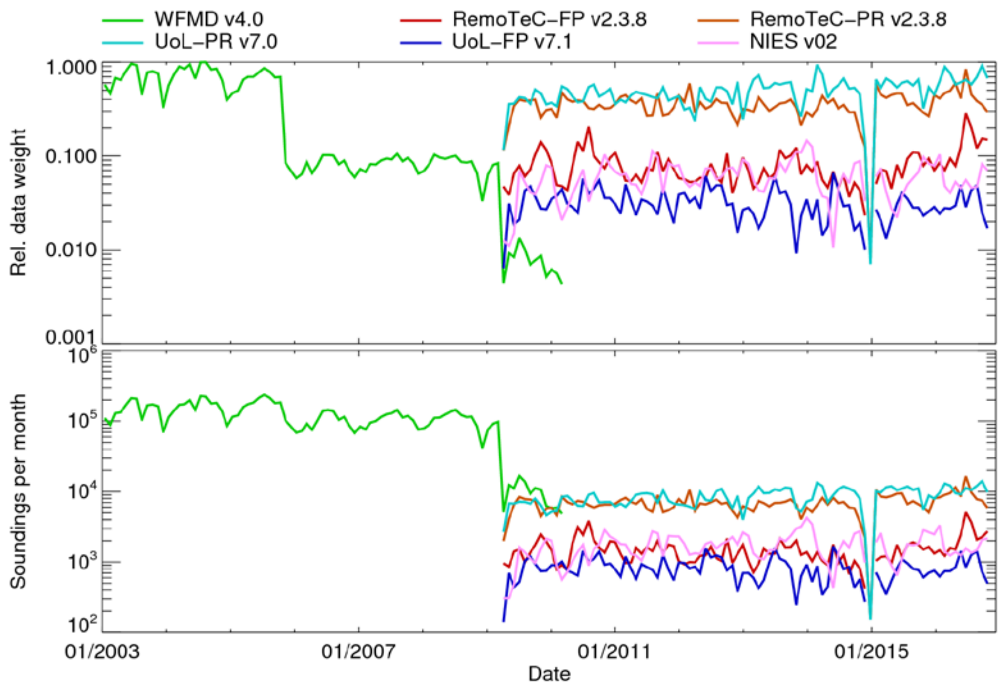


Figure 14: EMMA's normalized relative data weight proportional to $\sum 1/\sigma_i^2$ (**top**) and number of soundings (**bottom**) per algorithm and month.

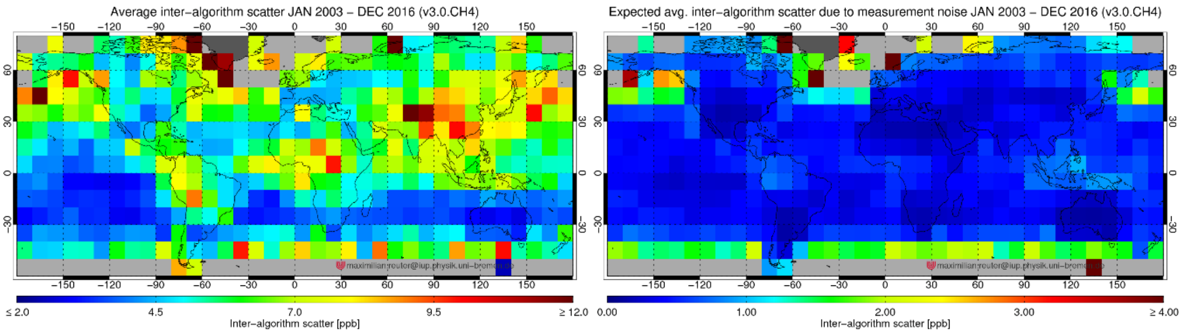


Figure 15: Average inter-algorithm scatter of monthly 10°x10° averages from January 2003 till December 2016 (left) and corresponding expected contribution of measurement noise (right).

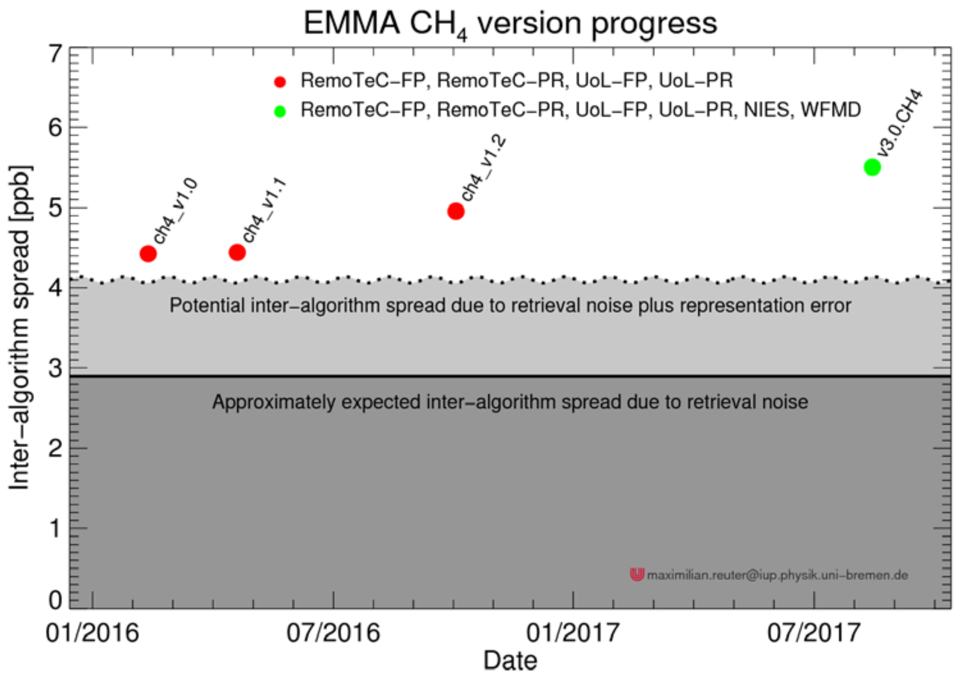


Figure 16: Average inter-algorithm spread of all EMMA versions compared with the approximately expected contribution of retrieval noise and a rough estimate of the representation error.

2.2.2 Summary

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

Table 5: Product Quality Summary Table for product XCH4_EMMA.

Product Quality Summary Table for Product: XCH4_EMMA Level: 2, Version: 3.0, Time period covered: 1.2003 – 12.2016				
Parameter [unit]	Achieved performance	Requirement	TR	Comments
Single measurement precision (1-sigma) in [ppb]	43.9	< 34 (T) < 17 (B) < 9 (G)	-	The achieved performance corresponds to the average over the full time period including SCIAMACHY till 2010. The performance significantly improves afterwards.
Uncertainty ratio) in [-]: Ratio reported uncertainty to standard deviation of satellite-TCCON difference	1.12	-	-	No requirement but value close to unity expected for a high quality data product.
Mean bias [ppb]	-7.4	-	-	No requirement but value close to zero expected for a high quality data product.
Accuracy: Relative systematic error [ppb]	Spatial – spatiotemporal: 4.21 – 5.52	< 10	Probability that accuracy TR is met: 100%	-
Stability: Linear bias trend [ppb/year]	0.11 +/- 0.89 (1-sigma)	< 3	Probability that stability TR is met: 97%	-
Stability: Year-to-year bias variability [ppb/year]	5.9 (1-sigma)	< 3	-	The achieved performance corresponds to the average over the full time period including SCIAMACHY till 2010. The performance significantly improves afterwards.

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

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

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., 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 CO2 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

Yoshida et al., 2011: Yoshida, Y., Ota, Y., Eguchi, N., Kikuchi, N., Nobuta, K., Tran, H., Morino, I., and Yokota, T.: Retrieval algorithm for CO2 and CH4 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

Copernicus Climate Change Service

ECMWF - Shinfield Park, Reading RG2 9AX, UK

Contact: info@copernicus-climate.eu