ESA Climate Change Initiative "Plus" (CCI+)

Product Validation and Intercomparison Report (PVIR) version 3

for the Essential Climate Variable (ECV)

Greenhouse Gases (GHG):

XCO_2 and/or XCH_4 from OCO-2, TanSat, Sentinel-5-Precursor and GOSAT-2

Written by: GHG-CCI+ project team

Lead author: Michael Buchwitz, Institute of Environmental Physics (IUP), University of Bremen, Germany

Co-authors:

Bart Dils (BIRA-IASB, Belgium, Brussels)

Maximilian Reuter, Oliver Schneising, Michael Hilker (IUP, Univ. Bremen, Germany)

Simon Preval, Hartmut Boesch (Univ. Leicester, UK)

Tobias Borsdorff, Jochen Landgraf, Trismono Candra Krisna (SRON, Leiden, The Netherlands)

Change log:

Version Nr.	Date	Status	Reason for change
Version 1.1	13-Mar-2020	Submitted	New document for CRDP5
Version 2.0	10-Feb-2021	Submitted	Update for CRDP6
Version 2.1	19-Mar-2021	Submitted	Content for Sect. 4.2.2 added on request of ESA.
Version 3.0	16-Feb-2022	Submitted	Update for CRDP7

Table of Contents

1 Executive Summary

This document is the Product Validation and Intercomparison Report (PVIR) version 3.0 (v3.0), which is a deliverable of the ESA project GHG-CCI+ (https://climate.esa.int/en/projects/ghgs/). The GHG-CCI+ project, which started in March 2019, is carrying out the research and development (R&D) needed to generate new Greenhouse Gas (GHG) Essential Climate Variable (ECV) satellite-derived CO_2 and CH_4 data products. These products are column-averaged dry-air mole fractions of carbon dioxide (CO_2), denoted XCO_2, and methane (CH_4), denoted XCH_4, from these satellites / satellite sensors using European scientific retrieval algorithms:

- XCO_2 from OCO-2 using the University of Bremen FOCAL algorithm (product **CO2_OC2_FOCA**),
- XCH_4 from Sentinel-5 Precursor (S5P) using University of Bremen's WFM-DOAS (or WFMD) algorithm (product **CH4_S5P_WFMD**),
- XCO_2 from TanSat using University of Leicester UoL-FP (or OCFP) algorithm (product **CO2_TAN_OCFP**), and
- XCO_2 and XCH_4 from GOSAT-2 using SRON's RemoTeC algorithm (products **CO2_GO2_SRFP, CH4_GO2_SRFP, CH4_GO2_SRPR**)

This project aims to generate GHG ECV data products in-line with GCOS (Global Climate Observing System) requirements. GCOS defines the ECV GHG as follows: "Retrievals of greenhouse gases, such as CO_2 and CH_4, of sufficient quality to estimate regional sources and sinks". Within the GHG-CCI+ project satellite-derived XCO_2 (in ppm) and XCH_4 (in ppb) data products are retrieved from satellite radiance observations in the Short-Wave-Infra-Red (SWIR) spectral region. These instruments are used because their measurements are sensitive also to the lowest atmospheric layer and therefore provide information on the regional surface sources and sinks of CO_2 and CH_4. All products are generated with independent retrieval algorithms developed to convert GOSAT-2, OCO-2, TanSat and/or TROPOMI/S5P radiance spectra into Level 2 (L2) XCO_2 and/or XCH_4 data products.

In this document the validation and intercomparison results are presented. The validation is based on comparisons with TCCON (Total Carbon Column Observation Network) ground-based XCO_2 and XCH_4 retrievals. The validation has been carried out by the GHG-CCI+ independent Validation Team (VALT) and by the data provider (DP) of a given product.

For each data product and each assessment method the following validation summary "figures of merit" have been determined and are reported in this document: (i) Single measurement precision, (ii) mean bias (global offset), (iii) relative systematic error (or relative accuracy), (iv) stability (linear bias drift or trend). Furthermore, also the reported XCO_2 and XCH_4 uncertainties have been validated by computing a quantity called "Uncertainty ratio", which is the ratio of the (mean value of the) reported uncertainty and the standard deviation of satellite minus TCCON differences. The results are summarized in **Table 1-1** for the XCO_2 products and **Table 1-2** for the XCH_4 product.

Table 1-1: *Summary of the validation of XCO_2 products CO2_OC2_FOCA and CO2_TAN_OCFP of data set Climate Research Data Package No. 7 (CRDP#7, to be released in March 2022) via comparison with TCCON ground-based XCO_2 retrievals (using version GGG2014). VALT refers to the assessment results of the GHG-CCI+ independent validation team and DP refers to the assessment results of the data provider. (*) Excluding a possible global offset, which is reported separately in this document. The range reported for VALT results in square brackets […] correspond with the upper and lower 95% confidence bound on the parameter. “n.a.” means “not applicable” and “n.e.” means “not evaluated (e.g., because time series is too short).*

Summary validation results GHG-CCI+ CRDP#7 XCO_2 products by comparisons with TCCON (GGG2014)			
Product CO2_OC2_FOCA (v10, global, 9.2014 – 3.2021)			
Parameter	**Achieved**	**Required**	**Comments**
Random error (single obs., 1σ) [ppm]	VALT: 1.60 [1.49,1.67] DP: 1.69	T:<8; B:<3; G:<1	T=threshold; B=breakthrough; G=goal
Systematic error [ppm]	VALT: 0.62 [0.29, 0.91] / 0.83 [0.71, 1.04] DP: 0.57 / 0.62	< 0.5	“Relative accuracy” (*) Spatial / spatio-temp.
Stability: Linear bias trend [ppm/year]	VALT: 0.01 [-0.05, 0.09] DP: -0.01 ± 0.20	< 0.5	1σ uncertainty

Product CO2_TAN_OCFP (v1.2, global land, 3.2017 – 5.2018)			
Parameter	**Achieved**	**Required**	**Comments**
Random error (single obs., 1σ) [ppm]	VALT: 1.51 [1.34, 1.69] DP: 1.78	T:<8; B:<3; G:<1	T=threshold; B=breakthrough; G=goal
Systematic error [ppm]	VALT: 0.50 [-0.07, 0.08] / 0.96 [0.62, 1.20] DP: 0.84 / n.e.	< 0.5	“Relative accuracy” (*) Spatial / spatio-temp.
Stability: Linear bias trend [ppm/year]	VALT: n.e. DP: n.e.	< 0.5	1σ uncertainty Only short time period

Table is continued on the following page …

Table 1-1: *Continued from previous page.*

Product CO2_GO2_SRFP (v02.0.0, global, 2.2019 – 8.2020)			
Parameter	**Achieved**	**Required**	**Comments**
Random error (single obs., 1σ) [ppm]	VALT: 2.04 [1.67, 2.25] DP: 2.26	T:<8; B:<3; G:<1	T=threshold; B=breakthrough; G=goal
Systematic error [ppm]	VALT: 0.57 [-0.19, 0.91] / 1.28 [0.91, 1.79] DP: 1.0 / n.e.	< 0.5	“Relative accuracy” (*) Spatial / spatio-temp.
Stability: Linear bias trend [ppm/year]	VALT: n.e. DP: n.e.	< 0.5	1σ uncertainty Only short time period

Table 1-2: *Summary of the validation of XCH_4 products CH4_S5P_WFMD of data set Climate Research Data Package No. 7 (CRDP#7, to be released in March 2022) via comparison with TCCON ground-based XCH_4 retrievals (using version GGG2014). VALT refers to the assessment results of the GHG-CCI+ independent validation team and DP refers to the assessment results of the data provider. (*) Excluding a possible global offset, which is reported separately in this document. The range reported for VALT results in square brackets […] correspond with the upper and lower 95% confidence bound on the parameter. "n.a." means "not applicable" and "n.e." means "not evaluated (e.g., because time series is too short).*

Summary validation results GHG-CCI+ CRDP#7 XCH_4 products by comparisons with TCCON (GGG2014)			
Product CH4_S5P_WFMD (v1.5, global, 11.2017– 12.2020)			
Parameter	**Achieved**	**Required**	**Comments**
Random error (single obs., 1σ) [ppb]	VALT: 13.8 [12.1, 15.0] DP: 12.9	T:<34; B:<17; G:<9	T=threshold; B=breakthrough; G=goal
Systematic error [ppb]	VALT: 5.0 [1.8, 7.3] / 5.2 [3.6, 6.5] DP: 5.17 / 5.2	< 10	"Relative accuracy" (*) Spatial / spatio-temp.
Stability: Linear bias trend [ppb/year]	VALT: 0.8 [0.4, 2.5] DP: 0.01	< 3	1σ uncertainty Only short time period

Product CH4_GO2_SRFP (v02.0.0, global, 2.2019– 8.2020)			
Parameter	**Achieved**	**Required**	**Comments**
Random error (single obs., 1σ) [ppb]	VALT: 14.7 [13.5, 16.2] DP: 14.4	T:<34; B:<17; G:<9	T=threshold; B=breakthrough; G=goal
Systematic error [ppb]	VALT: 5.8 [2.3, 8.9] / 6.9 [4.4, 9.7] DP: 2.4 / n.e.	< 10	"Relative accuracy" (*) Spatial / spatio-temp.
Stability: Linear bias trend [ppb/year]	VALT: n.e. DP: n.e.	< 3	1σ uncertainty Only short time period

Table is continued on the following page …

Table 1-2: *Continued from previous page.*

Product CH4_GO2_SRPR (v02.0.0, global, 2.2019– 7.2020)			
Parameter	**Achieved**	**Required**	**Comments**
Random error (single obs., 1σ) [ppb]	VALT: 16.7 [15.3, 17.7] DP: 15.5	T:<34; B:<17; G:<9	T=threshold; B=breakthrough; G=goal
Systematic error [ppb]	VALT: 5.0 [1.4, 7.0] / 9.4 [6.6, 12.0] DP: 4.2 / n.a.	< 10	"Relative accuracy" (*) Spatial / spatio-temp.
Stability: Linear bias trend [ppb/year]	VALT: n.a. DP: n.a.	< 3	1σ uncertainty Only short time period

2 Introduction

This document is the Product Validation and Intercomparison Report (PVIR) version 3.0 (v3.0), which is a deliverable of the ESA project GHG-CCI+ (https://climate.esa.int/en/projects/ghgs/).

The GHG-CCI+ project, which started in March 2019, is carrying out the R&D needed to generate new Greenhouse Gas (GHG) Essential Climate Variable (ECV) satellite-derived CO_2 and CH_4 data products.

These products are column-averaged dry-air mole fractions of carbon dioxide (CO_2), denoted XCO_2, and methane (CH_4), denoted XCH_4, from these satellites / satellite sensors using European scientific retrieval algorithms:

- XCO_2 from OCO-2 and TANSAT,
- XCO_2 and XCH_4 from GOSAT-2 and
- XCH_4 from S5P

This project aims to generate GHG ECV data products in-line with GCOS (Global Climate Observing System) requirements **/GCOS-154/ /GCOS-195/ /GCOS-200/**. GCOS defines the ECV GHG as follows: "Retrievals of greenhouse gases, such as CO_2 and CH_4, of sufficient quality to estimate regional sources and sinks".

Once the products are of sufficient quality for a climate service and cover a long enough time period, it is expected that the data will become part of the Copernicus Climate Change Service (C3S, https://climate.copernicus.eu/).

Within GHG-CCI+ satellite-derived XCO_2 (in ppm) and XCH_4 (in ppb) data products are retrieved from satellite radiance observations in the Short-Wave-Infra-Red (SWIR) spectral region. These instruments are used because their measurements are sensitive also to the lowest atmospheric layer and therefore provide information on the regional surface sources and sinks of CO_2 and CH_4.

This document provides validation and intercomparison results for the XCO_2 and XCH_4 datasets as listed in **Table 2-1** for XCO_2 and **Table 2-2** for XCH_4.

All products are generated with independent retrieval algorithms developed to convert GOSAT-2, OCO-2, TANSAT and/or TROPOMI/S5P radiance spectra into Level 2 (L2) XCO_2 and/or XCH_4 data products.

For more information on these products see also **Table 2-3**.

Table 2-1: *Overview GHG-CCI+ algorithms for XCO_2 retrieval.*

XCO_2 Product Identifier	Algorithm (version)	Institute	Technique	Reference
CO2_OC2_FOCA	FOCAL (v10)	IUP, Univ. Bremen, Germany	Optimal Estimation; approximation for an optically thin scattering layer	Reuter et al., 2017a, b
CO2_TAN_OCFP	UoL-FP (v1.2)	Univ. Leicester (UoL), United Kingdom	Optimal Estimation	Boesch et al., 2011
CO2_GO2_SRFP	SRFP or RemoTeC (v2.0.0)	SRON, Netherlands	Phillips-Tikhonov regularization	Butz et al., 2009, 2010

Table 2-2: *Overview GHG-CCI+ algorithms for XCH_4 retrieval.*

XCH_4 Product Identifier	Algorithm (version)	Institute	Technique	Reference
CH4_S5P_WFMD	WFM-DOAS (v1.5)	IUP, Univ. Bremen, Germany	Weighted least squares	Schneising et al., 2019
CH4_GO2_SRPR	SRPR or RemoTeC (v2.0.0)	SRON, Netherlands	Proxy (PR) retrieval method	Frankenberg et al., 2005
CH4_GO2_SRFP	SRFP or RemoTeC (v2.0.0)	SRON, Netherlands	Phillips-Tikhonov regularization; Full Physics (FP) method	Butz et al., 2009, 2010

Table 2-3: *Overview of (other) GHG-CCI+ product related documents. ATBD = Algorithm Theoretical Basis Document, PUG = Product User Guide, E3UB = End-to-End ECV Uncertainty Budget document.*

Product ID	Document	Link
CO2_OC2_FOCA	ATBD	Available from https://www.iup.uni-bremen.de/carbon_ghg/cg_data.html#GHG-CCI and https://climate.esa.int/de/projekte/ghgs/key-documents/
-"-	PUG	- " -
-"-	E3UB	- " -
CH4_S5P_WFMD	ATBD	- " -
-"-	PUG	- " -
-"-	E3UB	- " -
CO2_TAN_OCFP	ATBD	- " -
-"-	PUG	- " -
-"-	E3UB	- " -
CO2_GO2_SRFP	ATBD	- " -
-"-	PUG	- " -
-"-	E3UB	- " -
CH4_GO2_SRFP	ATBD	- " -
-"-	PUG	- " -
-"-	E3UB	- " -
CH4_GO2_SRPR	ATBD	- " -
-"-	PUG	- " -
-"-	E3UB	- " -

3 General description of the processing system

A schematic overview of the GHG-CCI+ processing system is given in **Figure 3-1**.

The processing system consists of the different algorithms (see **Tables 2-1 and 2-2**), running at the different responsible institutes.

The different institutes have their own access to the required input data (satellite data, ECMWF meteo data, model data for priors, spectroscopic databases, etc.), and their own computational facilities in the form of multi CPU Unix/Linux systems.

The Level-2 (L2) output data (XCO_2 and XCH_4) generated by the algorithms at the different institutes are available via the CCI Open Data Portal (*https://climate.esa.int/en/odp/#/dashboard*) and additional information is given at the GHG-CCI+ website (https://climate.esa.int/en/projects/ghgs/).

The different parts of the GHG-CCI+ processing systems running at the different institutes are described in more detail in the System Specification Document (SSD) document **/Aben et al., 2019/**.

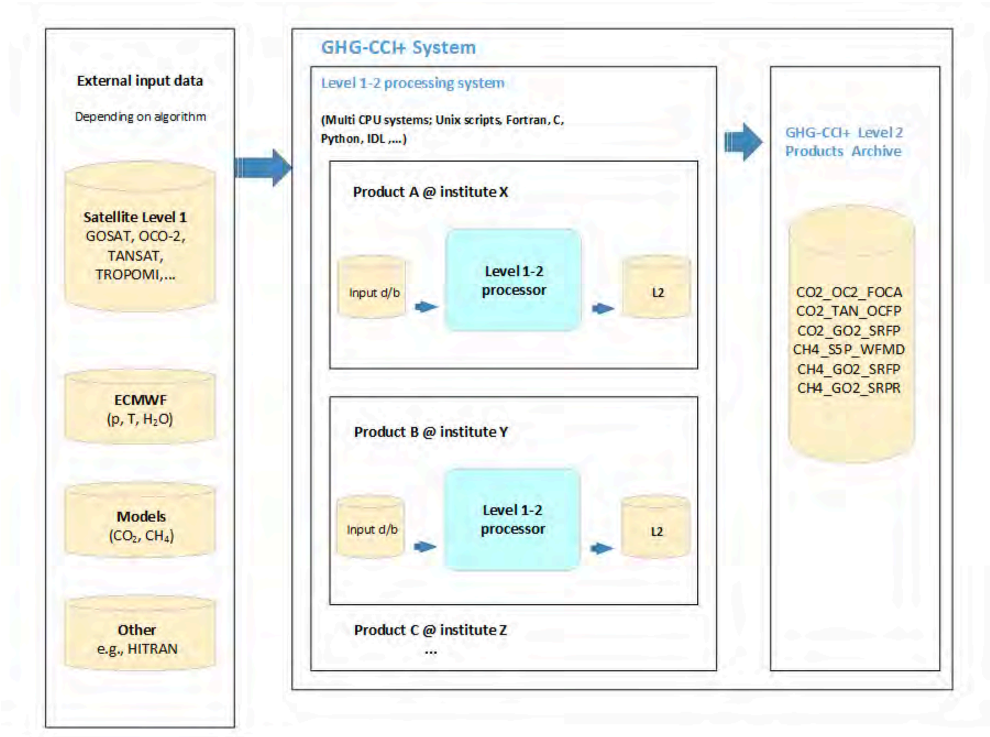


Figure 3-1: *Overview of the GHG-CCI+ processing system. Note that the GHG-CCI+ Level 2 product data archive is the CCI Open Data Portal (https://climate.esa.int/en/odp/#/dashboard).*

4 Independent validation by validation team

This chapter deals with the validation of the GHG-CCI+ retrieval products using ground-based FTIR remote sensing measurements from the Total Carbon Column Observing Network (TCCON) **/Wunch et al.2011/** and, in the case of XCH_4, the Network for the Detection of Atmospheric Composition Change (NDACC) **/De Mazière et al. 2018/**. The latter dataset was not used in our previous PVIR analysis. Take note that NDACC's data protocol is less harmonized as compared to TCCON's. For instance, it allows the use of 2 retrieval algorithms (SFIT4 and PROFFIT9). However, analysis between the two algorithms showed no bias between them **/Hase et al. 2004/**). It also features more stations in what we may call 'challenging environments', that being high altitude sites (Zugspitze, Jungfraujoch, Izaña, Mauna Loa, Reunion (Maido) and Altzomoni), near major urban sites (Toronto, Altzomoni (Mexico City)) and high latitude sites (Eureka, Ny Alesund, Thule, Arrival Heights). It also relies on the surface pressure to derive the dry air mole fraction (see equation 1 in **/Deutscher et al., 2010/**) as it cannot rely on a retrieved CH_4/O_2 ratio to reduce errors in the retrieval process.

TCCON also benefits from an extensive calibration campaign, which results in a calibration factor to reduce its systematic bias **/Wunch et al., 2011/**. TCCON's network accuracy can be determined by the uncertainty on this calibration factor, and amounts to 0.1% for XCO_2, and 0.2% for XCH_4. **/Wunch et al. 2010/.** The random uncertainty of TCCON is about 0.5% for XCH_4 and 0.25% for XCO_2. **/Wunch et al. 2015/.**

For NDACC, the systematic and random uncertainties of CH_4 total columns are estimated to be 3.0% and 1.5%, respectively. The first is mainly coming from the uncertainty of the spectroscopy.

Comparisons between TCCON and NDACC XCH_4 measurements **/Ostler et al., 2014**/ do demonstrated that there is no overall bias between both TCCON and NDACC XCH_4 retrieval methods. Therefore, we feel confident to include NDACC in our analysis, as it may provide some insight into regions that are not sampled by TCCON (Latin America being a prime example). An added benefit of the NDACC data is that it does not use a profile scaling retrieval method, but uses optimal estimation instead, retrieving profiles with ~2.5 degrees of freedom. This should, in principle, reduce the smoothing error, when we apply the satellite averaging kernels as it does not rely on the assumption that the real profile conforms to a pre-determined shape. Nor is the data used in post-retrieval bias-correction methods, that are employed by various satellite algorithms, to reduce the effect of residual systematic error components. While this approach is certainly valid, it also results in retrieval data that is optimized in some sense to the TCCON retrieval sites.

That said, the summary numbers in the tables, are still based on the TCCON analysis only. Mainly due to the much higher prevalence of high altitude/ high latitude sites and higher interstation biases in the NDACC network.

We have used all public TCCON GGG2014 data as available on the TCCON Data Archive (https://tccondata.org/) as well as all publicly available data on the NDACC archive (https://www-air.larc.nasa.gov/missions/ndacc/data.html) on the 1st of January 2022. We also included data from Garmisch, Sodankylä and Porto Velho, which are currently not officially part of NDACC but perform observations and data analysis fully compatible with NDACC guidelines.

Table 4.1: TCCON station coordinates and references.

STATION	Lat	Lon	Alt (km)	Ref
EUREKA	80.05 N	86,42 W	0.61	**/Strong et al., 2019/**
SODANKYLA	67.37N	26.62E	0.19	**/Kivi et al., 2014/**
EASTTROUTLAKE	54.35 N	104.99 W	0.50	**/Wunch et al., 2017/**
BIALYSTOK	53.23 N	23.05 E	0.18	**/Deutscher et al., 2014b/**
BREMEN	53.10 N	8.85 E	0.03	**/Notholt et al., 2019/**
KARLSRUHE	49.10 N	8.44 E	0.12	**/Hase et al., 2015/**
PARIS	48.85 N	2.36 E	0.06	**/Té et al., 2014/**
ORLEANS	47.97 N	2.11 E	0.13	**/Warneke et al., 2019/**
GARMISCH	47.48 N	11.06 E	0.74	**/Sussmann et al., 2018a/**
PARKFALLS	45.95 N	90.27 W	0.44	**/Wennberg et al., 2017/**
RIKUBETSU	43.46 N	143.77 E	0.38	**/Morino et al., 2018/**
LAMONT	36.60 N	97.49 W	0.32	**/Wennberg et al., 2016/**
ANMEYONDO	36.54 N	126.33 E	0.03	**/Goo et al., 2014/**
TSUKUBA	36.05 N	140.12 E	0.03	**/Morino et al., 2018b/**
NICOSIA	35.14 N	33.38 E	0.18	**/Petri et al., 2020/**
EDWARDS	34.96 N	117.88 W	0.70	**/Iraci et al., 2016/**
JPL	34.20 N	118.18 W	0.39	**/Wennberg et al. 2016b/**
PASADENA	34.14 N	118.13 W	0.23	**/Wennberg et al. 2015/**
SAGA	33.24 N	130.29 E	0.01	**/Kawakami et al. 2014/**
HEFEI	31.91 N	117.17 E	0.03	**/Liu et al. 2018/**
IZAÑA	28.30 N	16.50 W	2.37	**/Blumenstock et al., 2017/**
BURGOS	18.53 N	120.65 E	0.04	**/Morino et al., 2018c/**
ASCENSION	7.92 S	14.33 W	0.01	**/Feist et al., 2014/**
DARWIN	12.46 S	130.93 E	0.04	**/Griffith et al., 2014/**
REUNION	20.90 S	55.49 E	0.09	**/De Mazière et al., 2017/**
WOLLONGONG	34.41 S	150.88 E	0.03	**/Griffith et al., 2014b/**
LAUDER	45.04 S	169.68 E	0.37	**/Sherlock et al., 2014/**

Table 4-2: NDACC station coordinates and institutes/references.

STATION	Lat	Lon	Alt (km)	Institutes
EUREKA	80.05 N	86.42 W	0.61	U. of Toronto, /Batchelor et al., 2009/,/Strong 2021/
NY ALESUND	78.92 N	11.93 E	0.01	U. of Bremen, /Notholt et al., 2021a/
THULE	78.90 N	68.77 W	0.02	NCAR /Hannigan et al., 2021/
KIRUNA	67.84 N	20.40 E	0.2	KIT-ASF, IRF Kiruna /Blumenstock et al., 2020/
SODANKYLA	67.37 N	26.65 E	0.18	FMI, BIRA-IASB
HARESTUA	60.20 N	10.80 E	0.60	Chalmers, /Mellqvist et al., 2021/
St. PETERSBURG	59.88 N	29.83 E	0.02	SPbU, /Marakova et al., 2017/
BREMEN	53.11 N	8.85 E	0.03	U. of Bremen, /Notholt et al., 2021b/
GARMISCH	47.48 N	11.06 E	0.74	KIT-IFU
ZUGSPITZE	47.42 N	10.98 E	2.96	KIT-IFU, /Sussmann et al., 2018b/
JUNGFRAUJOCH	46.55 N	7.98 E	3.58	U. of Liège, /Mahieu, 2017/
TORONTO	43.60 N	79.36 W	0.17	U. of Toronto, /Wiacek et al., 2007/
RIKUBETSU	43,46 N	143.77 E	0.38	Nagoya U, NIES
BOULDER	40.04 N	105.24 W	1.61	NCAR, /Ortega et al. 2019/
IZAÑA	28.30 N	16.50 E	2.37	AEMET, KIT-ASF
MAUNA LOA	19.54 N	155.57 W	3.40	NCAR
ALTZOMONI	19.12 N	98.66 W	3.98	UNAM
PARAMARIBO	5.81 S	55.21 W	0.03	U. of Bremen
PORTO VELHO	8.77 S	296.13 W	0.09	BIRA-IASB
REUNION (MAÏDO)	21.08 S	55.38 E	2.16	BIRA-IASB
WOLLONGONG	34.41 S	150.88 E	0.03	U. of Wollongong
LAUDER	45.04 S	169.68 E	0.37	NIWA
ARRIVAL HEIGHTS	77.82 S	166.65	0.20	NIWA

As before, the key concept behind this validation is to apply an as uniform as possible validation strategy for all the involved algorithms. Apart from the already mentioned inclusion of NDACC CH_4 measurements, we uphold the same methodology as in the previous PVIR (see **/PVIR GHG-CCI+ v2.1, 2021**/ for details) analysis apart from 2 minor details. We changed the method in which we extrapolated profiles of higher altitude stations. Previously we used the satellite *a priori*, now we simply extrapolate downward assuming a constant mixing ratio. The first method obviously led to biases when the retrieved sat profile deviated significantly from its prior and in hindsight was the main driver behind the observed difference between smoothed and unsmoothed XCO_2 values. We also no longer ad hoc exclude TCCON mountain stations from our FoM calculations as their parameters no longer stand out among the other stations.

As always, choosing collocation criteria is a balance between minimizing the potential collocation error and still retaining a large enough sample so as to be able to derive adequate statistics. Also of note is that some of the current available algorithms have processed data for a limited time span only, which hampers certain aspects of the analysis.

Concerning the Figures of Merit (FoM), we did not employ any pre-analysis averaging and looked at individual satellite-TCCON pairs. This was done mainly to have statistical parameters that relate to the quality of the original data. Users of the data however should keep in mind that some algorithms opt to have a high density dataset with a larger random error component versus a much stricter quality-flagged low density dataset with a smaller random error component. After averaging (in space or time) the first might outperform the latter.

4.1 Validation method

Each individual satellite measurement is paired, if the criteria are met, with an individual FTS measurement (from TCCON or NDACC). This particular FTS measurement needs to be taken within 2 hours and within 500 km of the satellite measurement. Only for CH4_S5P_WFMD is the collocation criteria tightened to within 100 km and within 1 hour (TCCON) or 2 hours (NDACC) due to its high data density. If more than one FTS measurement fits the above criteria, the FTS measurement that has been measured closest (in time) to the satellite coordinates will be the one paired with said satellite measurement. This creates a collocated dataset with unique individual satellite-FTS pairs.

Prior to the FoM analysis we try to limit the impact of differences in *a priori* and vertical sensitivity between FTS and the satellite product (**/Rodgers, 2000/**). To limit the impact of the former we adjust the satellite dry air mole fraction using the FTS *a priori* as in

$$\hat{c}_{S,adj} = \hat{c_S} + \sum_l pw_l\,(1 - A_l)(x^l_{F,a} - x^l_{S,a})$$

where, $\hat{c_S}$ represents the originally retrieved satellite column-averaged dry air mole fraction, l is the index of the vertical layer, A_l the corresponding column averaging kernel of the satellite algorithm, $\boldsymbol{x}_{S,a}$ and $\boldsymbol{x}_{F,a}$ are the satellite and FTS *a priori* dry air mole fraction profiles respectively. pw_l is the pressure weight associated with level or layer l.

Likewise, to address the latter we apply the satellite averaging kernel onto the FTS data.

Unlike NDACC which directly yields retrieved profiles ($x_{F,r}$), TCCON provides total column dry air mole fractions only. So here we apply this smoothing onto the scaled TCCON *a priori,* where the scaling factor takes into account the actual retrieval (which is based on a scaling an *a priori* profile) as well as the post retrieval correction to bring TCCON in line with in situ measurements. Thus the scaled TCCON profile ($x_{F,r}$) corresponds with

$$x_{F,r} = x_{F,a} \times \hat{c}_{F,r}/\hat{c}_{F,a}$$

where $x_{F,a}$ is the TCCON *a priori* profile. $\hat{c}_{F,r}\ and\ \hat{c}_{F,a}$ are the TCCON retrieved and *a priori* column-averaged dry air mole fractions.

The adjusted FTS dry air mole fraction then corresponds with

$$\hat{c}_{F,adj} = \sum_{l} pw_l \,(x^l_{F,a} + (x^l_{F,r} - x^l_{F,a})A_l)$$

where, pw_l again represents the pressure weight associated with the level or vertical layer with index l and A_l the corresponding column averaging kernel of the satellite algorithm. $\boldsymbol{x}_{F,a}$ and $\boldsymbol{x}_{F,r}$ are the FTS *a priori* and scaled dry air mole fraction profiles respectively.

Prior to these adjustments, the FTS *a priori* needs to be interpolated onto the satellite product vertical grid. This is done using a regridding method that preserves mass (**/Langerock et al., 2015/**) and in case the satellite pixel surface altitude is below that of the FTS site, the regridded FTS profile is extrapolated towards the surface assuming a constant dry air molefraction.

This approach should minimize the differences between satellite and ground-based retrievals, regardless of the algorithm and target species involved.

The bias is defined as the median difference between the individual satellite and FTS pairs

$$\tilde{X}_{bias} = median(\hat{c}_{S,adj} - \hat{c}_{T,adj})$$

This is done for each station after which the overall Bias FoM is defined as the median of all calculated station biases. One could also group all individual measurements, regardless of station, into one sample onto which we calculate the bias, but this would increase the impact of stations where the data density is high. Since having a high data density, does not necessarily correspond with the highest quality data (or best collocation environment), we deem our median of station biases approach more accurate.

The scatter at each station corresponds with the median absolute deviation (mad) scaled by 1.4826 which is a statistically more robust proxy for the standard deviation (std) of said difference as in:

$$scatter = 1.4826 \times median(|X_{bias} - \tilde{X}_{bias}|)$$

where

$$X_{bias} = \hat{c}_{S,adj} - \hat{c}_{F,adj}$$

Again for the overall assessment of the scatter we take the median of all individual station scatter values.

Both parameters, bias and scatter, are presented with their 95% confidence interval in the validation summary tables (see **Tables 4-4, 4-6, 4-8, 4-11, 4-14, 4-17**). These confidence bands have been determined using a bootstrap methodology (**/Lunneborg, 2020/**), where the 95% confidence limits around the median $\tilde{X}$ corresponds with

[$\tilde{X}$ - *(97.5%tile -* $\tilde{X}$*),* $\tilde{X}$ + *(*$\tilde{X}$*- 0.25%tile*)]

Using medians and scaled median absolute deviations instead of means and standard deviations makes for a more robust assessment as it is far less impacted by outliers. These outliers could be haphazard single outliers (in the satellite data as well as for the FTS measurements, due to cloud interference etc.) when calculation the station bias and scatter values, but also caused by far from ideal collocation circumstances, limited data, etc. at various FTS sites when calculating the overall FoMs.

Other FoM are the Relative Accuracy (RA) and Seasonal Relative Accuracy (SRA), which give an indication of the spatial and spatio-temporal accuracy of the algorithm. We define RA as the scaled median absolute deviation on the overall median biases (derived from individual data) obtained at each station. The "Seasonal Relative Accuracy" (SRA), differs from the relative accuracy in that it uses the seasonal bias medians at each station, instead of the overall biases obtained at each station, it is thus the scaled median absolute deviation over all station seasonal median bias results. The seasonal bias results are constructed, for each FTS station, from all data pairs which fall within the months of January till March (JFM), April till June (AMJ), July till September (JAS) or October till December (OND), regardless of the year the measurements are taken. Some stations feature only limited data during certain seasons, which sometimes results in erratic (seasonal) bias results. To avoid the inclusion of these results into the RA and SRA calculation, we do not include those results which are derived from less than 4 individual SAT-FTS pairs. This may seem as a low threshold, but combined with the fact that we draw upon median values, we deem this sufficient.

To verify the stability of the algorithm over time we fit a linear trend and seasonal cycle through the bias timeseries:

$$X = i + s.t + A.\sin(2\pi.(t + ph))$$

Here, *X* represents the satellite minus FTS difference, *i* the intercept, *s* the slope which corresponds with the linear drift, *A* the amplitude of the seasonal cycle and *ph* the phase shift. While the slope yields information on any potential drift, the amplitude in the above fit results gives us information on the potential mismatch between Satellite and FTS seasonal cycles. Ideally there should be no difference between these cycles which would yield a slope and amplitude=0 in the bias timeseries. This is done for all stations provided that the overlapping station satellite timeseries covers a timespan of at least 2 years. The overall long term stability then corresponds with the median slope over all these stations as we expect the linear drift to be consistent for the entire dataset.

Figures 4-6, 4-10, 4-14, 4-22, 4-23, 4-30, 4-31, 4-38 and 4-39 show the monthly medians of all data within certain latitude bands. To determine the seasonal cycle, as with the determination of the long-term stability, a fit as outlined above is performed on the (now monthly median instead of individual) data. For the seasonal cycle representation, we then subtract the linear part from the medians and calculate the mean of all medians for each given month.

Another Figure of Merit is the so-called Uncertainty Ratio, which is defined as the ratio between the algorithm's reported uncertainty and the above mentioned scatter. If the reported uncertainty is correctly assessed, the uncertainty ratio should approach unity. However, this baseline number ignores any aspect of temporal, spatial or FTS variability embedded in the scatter.

We therefore also calculate an improved Uncertainty Ratio, which is the ratio between the reported uncertainty and the uncertainty on the Satellite (σ_{SAT}) as determined from the scatter using the method outlined below. Both are reported in the summary tables of each algorithm (see **Tables 4-4, 4-6, 4-8, 4-11, 4-14, 4-17**), where the improved uncertainty ratio is marked by an *.

Taking into account the variability of the FTS reference data and the collocation error, when assuming independence, the scatter can be written down as:

$$\text{scatter}=\sqrt{(\sigma_{SAT}^2 + \sigma_{FTS}^2 + \sigma_{Collocation}^2)}$$

where σ_{SAT} is the standard deviation due to variability of the satellite product, σ_{FTS} due to variability within the FTS measurements and $\sigma_{Collocation}$ due to variability in time and space. σ_{SAT} as derived from our comparison between the satellite and FTS measurements is thus:

$$\sigma_{SAT} = \sqrt{(scatter^2 - \sigma_{FTS}^2 - \sigma_{Collocation}^2)}$$

The standard deviation on the ground-based FTS measurements can be readily calculated from the average variability of the FTIR measurements within the collocation timeframe (4 hours).

The Collocation uncertainty is harder to define and consists of a spatial and temporal component. The latter can be ignored since it is already embedded in our calculation of the FTS uncertainty (which is based on the actual variability of the FTS measurements in time and thus also contains the temporal natural variability).

Unfortunately, we have no solid information on the spatial collocation uncertainty. Our best, but flawed, estimate of this factor can be derived from fitting a linear equation through the sat-TCCON residuals as a function of distance between the FTS site and the satellite pixel center points (we do this for all satellite FTS pairs drawn from all stations, see **Figure 4-2**). From the obtained slope a, we can then estimate the uncertainty associated with the collocation by simply taking the standard deviation of points along the slope (a×dist(i)), where dist(i) is the distance between the FTS station and satellite centre point for a given sat-FTS pair with index i. Note that we here use the normal standard deviation as, by default, there are no outliers in the points that constitute the slope.

As already mentioned, this is a mere estimate and corresponds more with a lower bound threshold, as station to station bias results can differ profoundly. Most noticeable is to look at bias value differences between sites where the collocation areas overlap to a large degree, such as Pasadena and Edwards (see **Figure 4-1 and Tables 4-3, 4-5, 4-7, 4-9, 4-12, 4-15**).

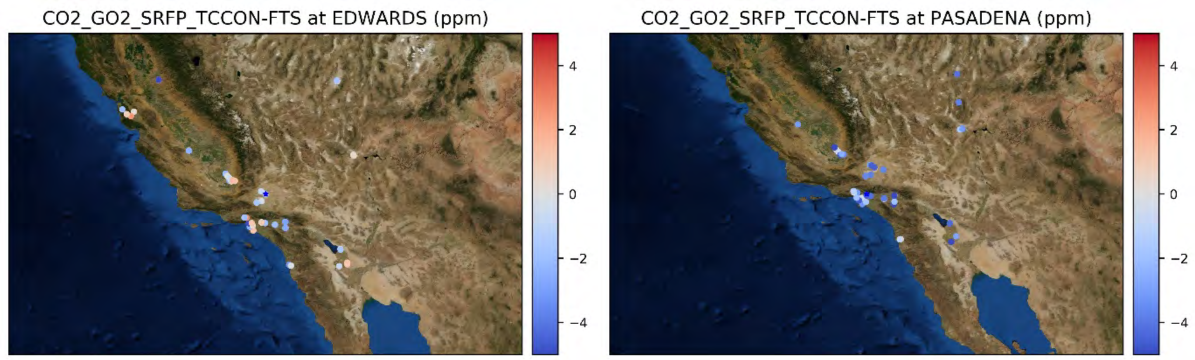


Figure 4-1: Example plot of collocated data (in this case SRFP XCO_2) at Pasadena and Edwards (bottom).

As can be seen in **Figure 4-2**, which shows all the 'bias as a function of distance' plots, the effect is fairly limited. For XCO_2, values range between 0.01 and 0.05 ppm/100 km, for XCH_4 we see values between -0.43 and 0.12 ppb/100km for TCCON and between -3.00 and -2.47 ppb/100km for NDACC. This does not mean that there are only collocation issues within the NDACC constellation, but rather that it does not present itself in TCCON as a general feature over the entire dataset. If we look at the slopes on a per station basis for the two algorithms with the highest datadensity, we find that for CO2_OC2_FOCA these range between -0.34 (Hefei) and 0.41 ppm/100 km (Eureka). For CH4_S5P_WFMD the slopes ranged between -16.4 and +12.4 ppb/100 km (for Izaña and Edwards respectively) in the TCCON dataset and between -42.1 and 41.3 (for Ny Alesund and Mauna Loa respectively) in the NDACC dataset. The Ny Alesund station also delivers TCCON measurements and here the slope equals -10.8 ppb/100 km, which is a large and unexpected difference between the TCCON and NDACC results. Other stations that are joint NDACC-TCCON stations (Eureka, Sodankyla, Bremen, Garmisch, Zugspitze, Izaña, Wollongong and Lauder) feature much smaller differences that typically do not exceed 2 ppb/100 km, thus whatever feature triggered this discrepancy, it is not a general NDACC vs. TCCON feature.

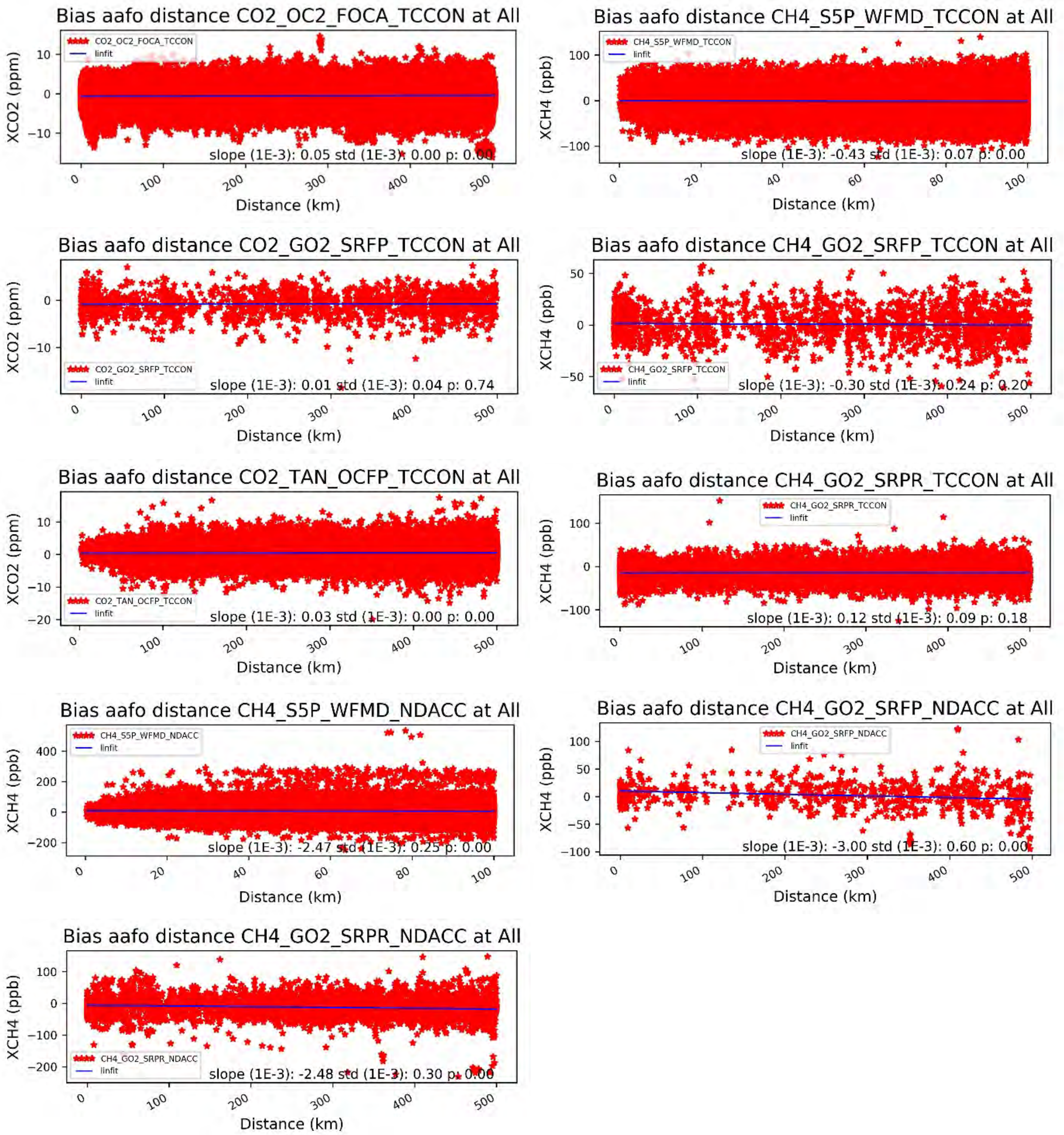


Figure 4-2: Satellite-TCCON or NDACC bias as a function of (aafo) distance between the satellite and TCCON/NDACC sampling point, for all algorithms in this study. Slope in ppm/100 km for XCO_2 and ppb/100 km for XCH_4.

4.2 Validation results

This section lists all validation results for the algorithms presently available in this study. First we show, for each algorithm, a general overview of the collocated data.

This comprises of a Taylor plot and a mosaic overview of the obtained timeseries.

The Taylor plot shows the correlation between the various FTS sites and the retrieval algorithm (straight lines), the standard deviation of the FTS data at each site, relative to the standard deviation of the satellite (normalized to 1) (light grey arches) and the root mean square error of the sat-fts difference (dark grey arches).

After this we discuss the different statistical parameters as obtained on a per station level.

Then the temporal variability is discussed, showing all the station timeseries as well as a more broad 'latitudinal band' based discussion on the long term trend (if any) and seasonality.

After this we discuss the overall FoM, obtained from the analysis of individual data, and their statistical reliability.

Thus in each section, we show:

1) A Taylor and Mosaic overview plot.
2) A table listing all Bias, Scatter, correlation (R), number of collocated data pairs (N) for all stations, and, if the timeseries allows, the slopes and amplitudes of the trend fits.
3) Example timeseries of individual data.
4) Monthly averaged timeseries and seasonal plots for broader latitude bands.
5) A Summary table of the Figures of Merit drawn from the values, drawn from individual measurements, at all stations.

4.2.1 Validation results for product CO2_OC2_FOCA

Below we show the validation results of the XCO_2 concentrations as derived by the CO2_OC2_FOCA v10 algorithm using OCO-2 spectra. Data was available from September 2014 until the end of March 2021. The FOCAL algorithm provides *a priori* and column averaging kernel data on a 5 layer profile. Compared to the last PVIR iteration little has changed in terms of its comparisons with TCCON. There are slight changes in the FoM but never abruptly and always within the previously established confidence bounds.

4.2.1.1 Detailed results

The Taylor diagram below in **Figure 4-3** yields a concise overview of the capabilities of the CO2_OC2_FOCA algorithm. Most TCCON sites cluster around the 0.9 correlation line. Also, the normalized standard deviation of most sites is close to 1, indicating that the variability of both datasets (due to natural variability and random error) is comparable. The normalized standard deviation of the bias (std(sat-fts)/std(sat)) sits (for most sites) around 0.4, which is very encouraging as it indicates that a large fraction of the variability (we can only assume it is the natural variability part) within the TCCON timeseries is also captured by the satellite.

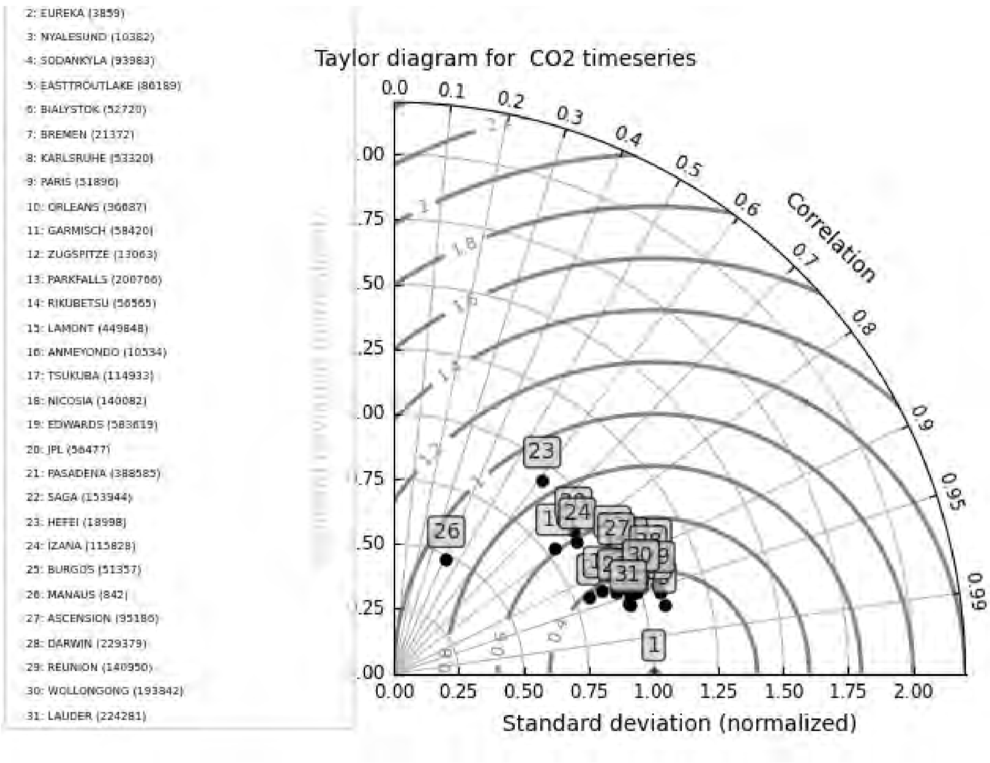


Figure 4-3: Tayor plot of XCO_2 TCCON values relative to CO2_OC2_FOCA. Straight lines correspond with the correlation, light grey lines yield the variability of the TCCON data relative to the satellite variability and the dark grey lines correspond with the variability of the Satellite -TCCON bias relative to the satellite variability.

Notable outliers are Manaus and Hefei with lower correlations (~0.4 and ~0.6) but both datasets only cover a limited fraction of the sampled timeperiod (see **Figure 4-4**)

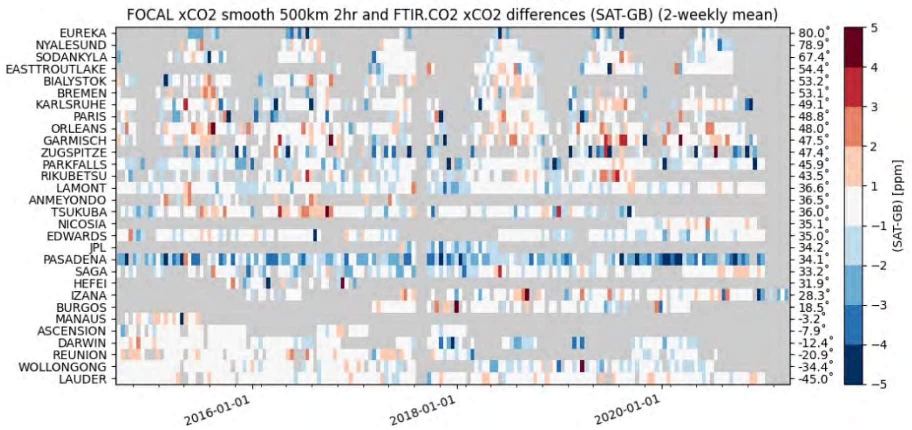


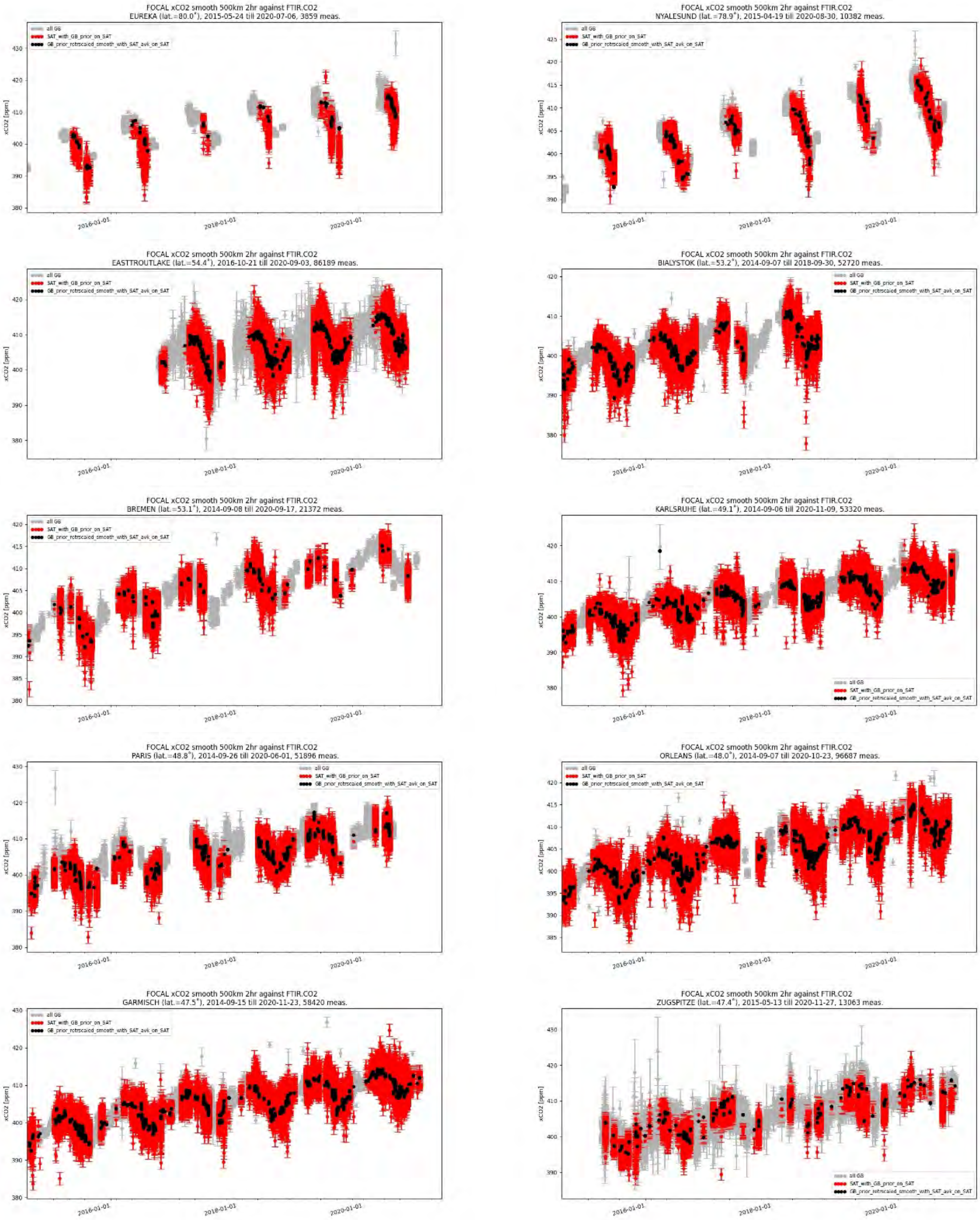
Figure 4-4: Mosaic plot of bi-weekly mean CO2_OC2_FOCA-TCCON XCO_2 biases as a function of time and TCCON station.

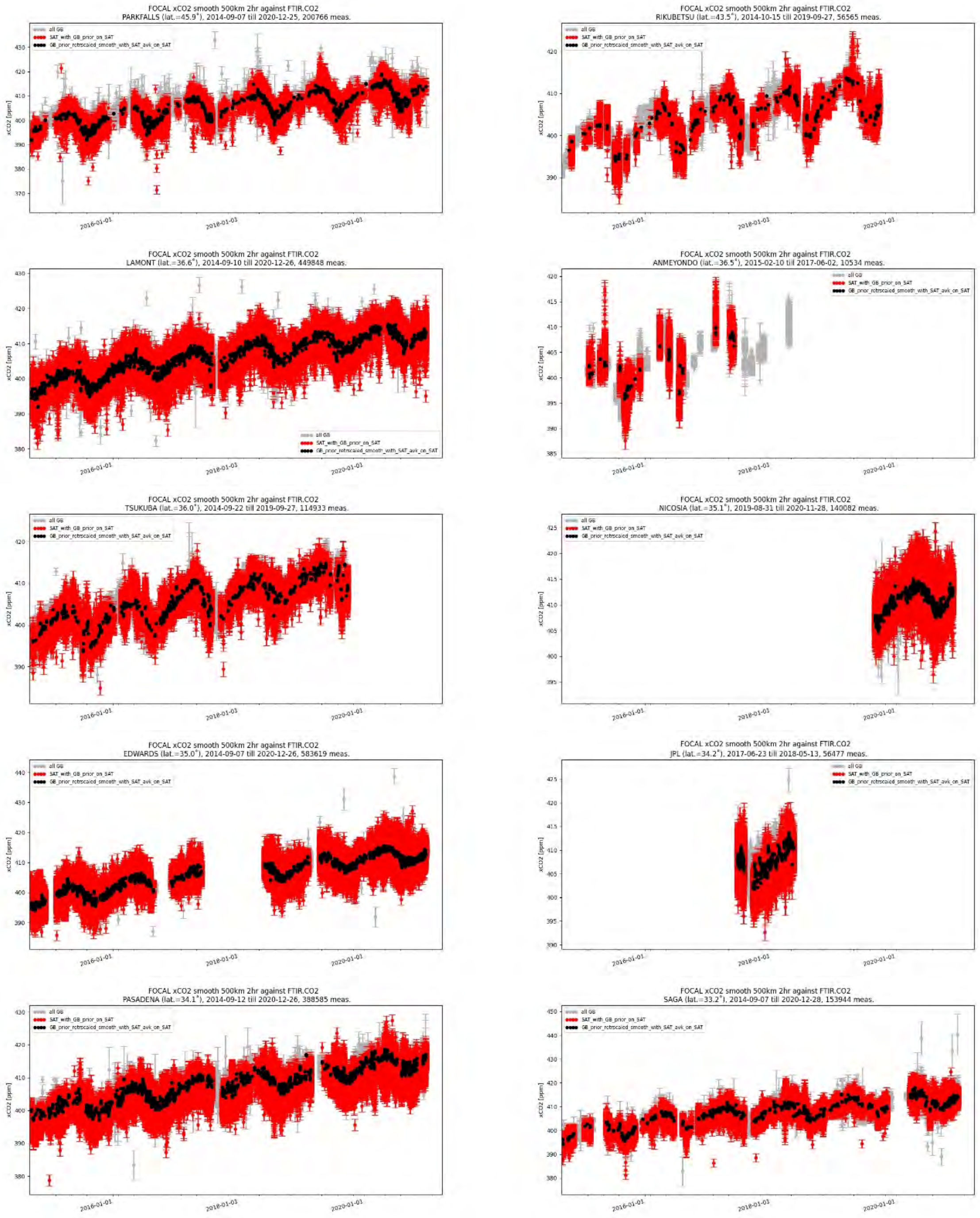
It is hard to discern a pattern in the above mosaic plot (**Figure 4-4**), which shows the mean bi-weekly bias between the satellite and TCCON measurement pairs. One can see the seasonal unavailability of data during winter (not visible for the Southern hemisphere as Lauder (New Zealand) still sits at a modest 45°S. Pasadena has outspoken and consistent negative biases (see also **Table 4-3**). This is not surprising as it is located within the Los Angeles basin and typically measures larger concentrations than what is present outside the basin. The nearby Edwards site which to a large degree has an overlapping collocation area (see **Figure 4-1**) features much different bias values (-0.73 ppm compared to -2.19 ppm at Pasadena). The algorithm produces on average ~90000 data pairs per station. Which roughly corresponds with around 13000 data pairs per station per year. Of the stations, only 7 out of 30 have a correlation coefficient under 0.90 and 3 of those still have a correlation of more than 0.80. The correlation of all data (regardless of station) equals 0.93. The bias ranges between -2.19 ppm (Pasadena) and 1.14 ppm (Manaus) and the scatter between 2.43 ppm (Hefei) and 1.04 ppm (Lauder). Long term trends on the bias (the so-called drift) range between -0.39 ppm/year (Burgos) and 0.29 ppm/year (Saga). Note that we only calculated long-term trends for stations whose collocated dataset spans at least 2 years. The amplitude on the other hand ranges between 1.41 ppm at Karlsruhe and 0.11 ppm at Ascension.

Table 4-3: Number of collocated data pairs (N), Correlation (R), Bias, Scatter, long term trend difference (ltt) and uncertainty thereon (ltt_err), seasonal amplitude difference (A) and uncertainty thereon (A_err) as well as the latitude of the TCCON station. The last row lists the median values over all stations. Product: CO2_OC2_FOCA.

STATION	N	R	Bias	Scat	ltt	ltt_err	A	A_err	lat
EUREKA	3859	0.93	-1.52	2.38	0.29	0.14	1.19	1.76	80.0
NYALESUND	10382	0.97	0.10	1.08	-0.20	0.10	0.35	0.35	78.9
SODANKYLA	93983	0.94	-0.26	1.56	0.01	0.06	1.04	0.34	67.4
EASTTROUTLAKE	86189	0.91	-0.11	1.63	0.07	0.12	0.70	0.35	54.4
BIALYSTOK	52720	0.91	0.10	1.69	0.06	0.11	1.11	0.26	53.2
BREMEN	21372	0.96	0.28	1.72	0.02	0.15	1.23	0.52	53.1
KARLSRUHE	53320	0.93	0.32	1.74	0.02	0.07	1.41	0.21	49.1
PARIS	51896	0.91	-0.56	1.68	0.23	0.09	1.03	0.31	48.8
ORLEANS	96687	0.94	0.44	1.54	-0.03	0.04	0.88	0.15	48.0
GARMISCH	58420	0.92	0.20	1.76	0.10	0.08	0.81	0.21	47.5
ZUGSPITZE	13063	0.92	-1.25	2.06	-0.09	0.21	0.28	0.38	47.4
PARKFALLS	200766	0.96	-0.55	1.49	0.13	0.04	0.83	0.14	45.9
RIKUBETSU	56565	0.95	0.15	1.46	0.26	0.12	0.78	0.20	43.5
LAMONT	449848	0.94	-0.71	1.61	0.00	0.03	0.44	0.09	36.6
ANMEYONDO	10534	0.93	0.09	1.72	0.08	0.38	0.31	0.36	36.5
TSUKUBA	114933	0.93	-0.23	1.54	-0.01	0.11	0.30	0.16	36.0
NICOSIA	140082	0.79	0.35	1.60	-	-	-	-	35.1
EDWARDS	583619	0.95	-0.73	1.79	-0.07	0.04	0.18	0.10	35.0
JPL	56477	0.78	-1.47	2.07	-	-	-	-	34.2
PASADENA	388585	0.90	-2.19	2.07	-0.06	0.06	0.21	0.13	34.1
SAGA	153944	0.94	-0.50	1.60	0.29	0.06	0.17	0.13	33.2
HEFEI	18998	0.61	-0.60	2.43	-	-	-	-	31.9
IZANA	115828	0.81	-0.17	1.47	-0.09	0.13	0.74	0.20	28.3
BURGOS	51357	0.87	0.11	1.13	-0.39	0.11	0.86	0.19	18.5
MANAUS	842	0.42	1.14	1.75	-	-	-	-	-3.2
ASCENSION	95186	0.89	0.17	1.24	0.14	0.14	0.11	0.21	-7.9
DARWIN	229379	0.93	-0.71	1.55	-0.27	0.05	0.34	0.11	-12.4
REUNION	140950	0.95	0.37	1.15	-0.20	0.07	0.77	0.19	-20.9
WOLLONGONG	193842	0.94	-0.44	1.49	-0.10	0.05	0.15	0.11	-34.4
LAUDER	224281	0.96	0.40	1.04	0.09	0.03	0.31	0.08	-45.0
MEDIAN	**90086**	**0.93**	**-0.14**	**1.605**	**0.015**	**0.085**	**0.72**	**0.20**	**36.3**

The timeseries below in **Figure 4-5** show individual satellite and ground-based fts measurements. The capture of the seasonal cycle and long term trend is similar to that of TCCON. Some outliers are still present in the data (for instance in the Hefei, Bialystok or Park Falls plots) but overall most measurements yield good comparison results.




FOCAL xCO2 smooth 500km 2hr against FTIR.CO2
PARKFALLS (lat.=45.9˚), 2014-09-07 till 2020-12-25, 200766 meas.
FOCAL xCO2 smooth 500km 2hr against FTIR.CO2
RIKUBETSU (lat.=43.5˚), 2014-10-15 till 2019-09-27, 56565 meas.
FOCAL xCO2 smooth 500km 2hr against FTIR.CO2
LAMONT (lat.=36.6˚), 2014-09-10 till 2020-12-26, 449848 meas.
FOCAL xCO2 smooth 500km 2hr against FTIR.CO2
ANMEYONDO (lat.=36.5˚), 2015-02-10 till 2017-06-02, 10534 meas.
FOCAL xCO2 smooth 500km 2hr against FTIR.CO2
TSUKUBA (lat.=36.0˚), 2014-09-22 till 2019-09-27, 114933 meas.
FOCAL xCO2 smooth 500km 2hr against FTIR.CO2
NICOSIA (lat.=35.1˚), 2019-08-31 till 2020-11-28, 140082 meas.
FOCAL xCO2 smooth 500km 2hr against FTIR.CO2
EDWARDS (lat.=35.0˚), 2014-09-07 till 2020-12-26, 583619 meas.
FOCAL xCO2 smooth 500km 2hr against FTIR.CO2
JPL (lat.=34.2˚), 2017-06-23 till 2018-05-13, 56477 meas.
FOCAL xCO2 smooth 500km 2hr against FTIR.CO2
PASADENA (lat.=34.1˚), 2014-09-12 till 2020-12-26, 388585 meas.
FOCAL xCO2 smooth 500km 2hr against FTIR.CO2
SAGA (lat.=33.2˚), 2014-09-07 till 2020-12-28, 153944 meas.
all GB
SAT_with_GB_prior_on_SAT
GB_prior_rctrscaled_smooth_with_SAT_avk_on_SAT
xCO2 [ppm]
2016-01-01
2018-01-01
2020-01-01

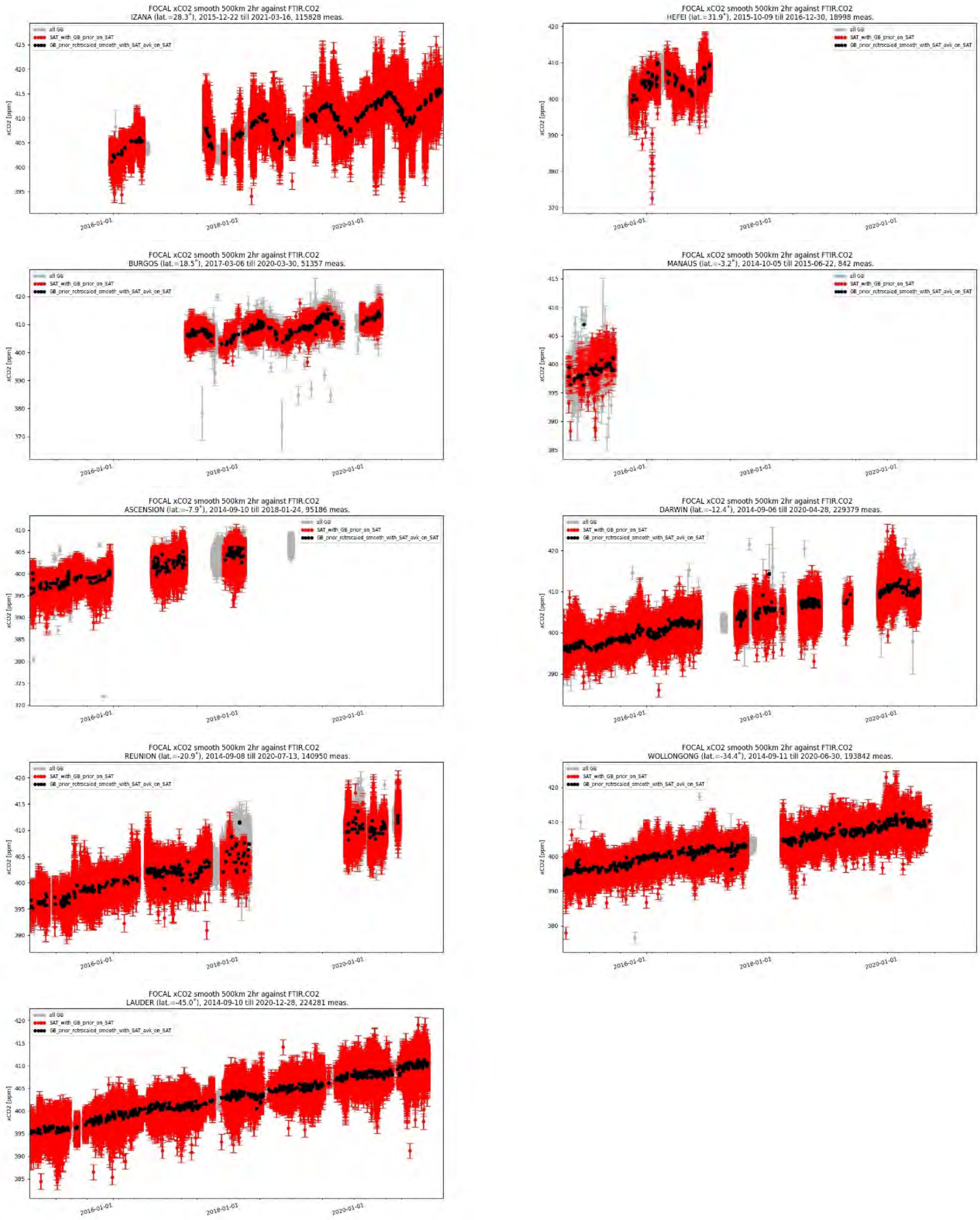


Figure 4-5: XCO_2 timeseries at all TCCON sites (red= CO2_OC2_FOCA data, black is collocated TCCON data and grey are the uncollocated TCCON data).

Figure 4-6 shows monthly median timeseries for TCCON and FOCAL XCO_2 for all data that fall within certain latitude bands, namely all sites north of 40°N latitude (top), all sites between 40°N and the equator (mid) and all sites in the Southern hemisphere (bottom). As can be seen, for all bands, the TCCON and FOCAL data feature long term trends that differ by 0.1 ppm/year only which is well within its uncertainty bounds. On the right hand side of each figure is the detrended monthly median values as a function of month. Again this clearly shows that FOCAL accurately captures the seasonal cycle. The median amplitude derived from seasonal fits through the individual bias data at each station amounts to 0.72 [0.6, 1.12] ppm.

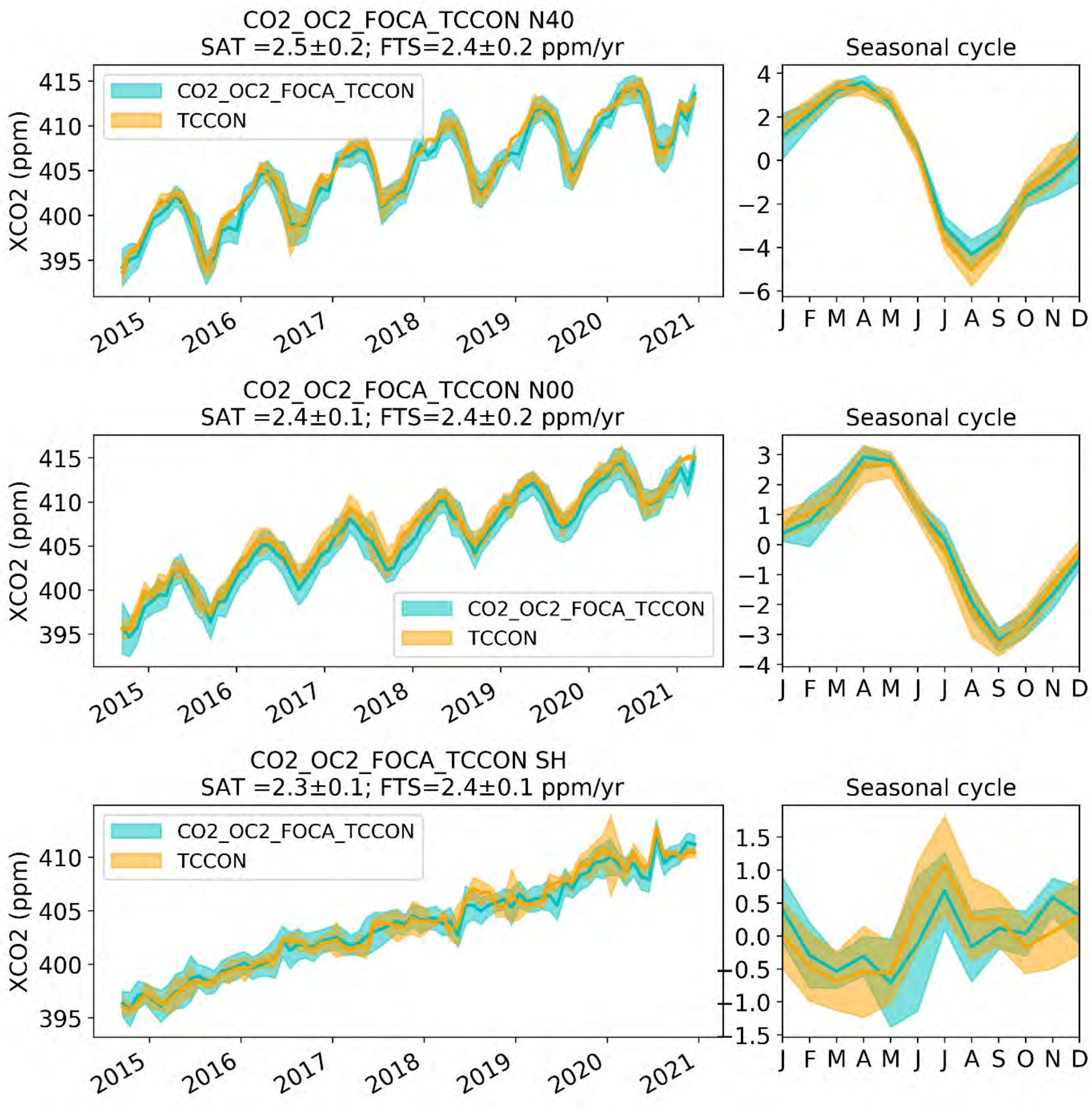


Figure 4-6: Monthly median collocated Sat and TCCON XCO_2 concentrations as a function of time and the detrended monthly medians as a function of season. The shaded areas correspond with the scaled median absolute deviation.

4.2.1.2 Summary

Listed in the table below (**Table 4-4**) are the Figure of Merit parameters as derived from the individual data pairs at the different TCCON stations. Values in square brackets [ ] correspond with the upper and lower 95% confidence bound on the parameter. The uncertainty ratio features 2 numbers as outlined in the validation method

Also important to note is that the results not only pertain to the actual data quality but also contain a collocation error component. For instance, the difference in the observed bias at the relatively close by Pasadena and Edwards stations is 1.46 ppm. The same holds true for Paris and Orleans (1.00 ppm difference).

Overall the CO2_OC2_FOCA product delivers data that matches very well with that of TCCON. This is apparent in the Taylor diagram time series plots as well as the Figures of Merit.

In our previous assessment **/PVIR GHG-CCI+ v2.1, 2021**/ the determined Relative Accuracy (0.41) was lower than the <0.5 ppm accuracy requirements and even the Seasonal Relative Accuracy (SRA at 0.73) was close and had itself confidence bands that overlapped with the target. Currently the estimated Relative Accuracy sits at 0.62 [0.29, 0.91] ppm, while the Seasonal Relative Accuracy equals 0.83 [0.71, 1.04] ppm. These are slightly higher numbers but one has to take into account that the previous analysis was performed using 25 TCCON stations, while this study used 30. The Relative accuracy's confidence bands still overlap with the target value, but those of the Seasonal Relative Accuracy do not. Take note that the accuracy requirements of < 0.5 ppm, assumes the abolishment of any collocation influence, nor any station-to-station differences within the TCCON network (its network accuracy is estimated to be within 0.4 ppm), all of which do contribute to the obtained RA and SRA values.

The reported uncertainty is, when compared to the scatter, very accurate (0.97 or 1.02). The scatter itself (1.60 ppm) has reached the so-called breakthrough levels (< 3 ppm). From the timeseries plots and Taylor diagram we in fact see that the variability closely matches this of TCCON. The overall bias is slightly negative but with confidence bounds that overlap with 0.

And finally the dataset shows no significant long term drift.

Table 4-4: presents an overview of the estimated data quality of CO2_OC2_FOCA, as obtained by the VALT team, from comparisons with TCCON ground-based reference observations. Values in square brackets [ ] correspond with the upper and lower 95% confidence bound on the parameter. The uncertainty ratio features 2 numbers as outlined in the validation method.

Product Quality Summary Table for Product: CO2_OC2_FOCA Level: 2, Version: v10, Time period covered: 9.2014 – 3.2021 Assessment: Validation Team (VALT)			
Parameter [unit]	**Achieved performance**	**Requirement**	**Comments**
Single measurement precision (1-sigma) in [ppm]	1.60 [1.49, 1.67]	< 8 (T) < 3 (B) < 1 (G)	Computed as the median over all station scaled median absolute differences to TCCON
Uncertainty ratio [-]: Ratio reported uncertainty to standard deviation of satellite-TCCON difference	0.97, 1.02*	-	No requirement but value close to unity expected for a high quality data product with reliable reported uncertainty.
Median bias (global offset) [ppm]	-0.14 [-0.41, 0.25]	-	No requirement but value close to zero expected for a high quality data product.
Accuracy: Relative systematic error [ppm]	Spatial: 0.62 [0.29, 0.91] Spatio-temporal: 0.83 [0.71, 1.04]	< 0.5	Spatial: Computed as standard deviation of the biases at the various TCCON sites. Spatio-temporal: As "Spatial" but also considering seasonal biases.
Stability: Drift [ppm/year]	0.01 [-0.05, 0.09]	< 0.5	Linear drift

4.2.2 Validation results for product CO2_TAN_OCFP

Here we present the VALT validation results for the CO2_TAN_OCFP product. The analysis pertains to the global v1.2 dataset, which is essentially still the same algorithm as v1.0. Version 1.1 corrected some minor issues but unfortunately introduced some issues with the quality flag which were uncovered during the course of his validation analysis. These have been resolved in the current v1.2 dataset and we therefore foresee little overall change with respect to the previous analysis. Data was available from March 2017 up to and including May 2018. The OCFP algorithm provides *a priori* and column averaging kernel information on a 20 level profile. Given the very limited time period that is covered by this product, these validation results will be rather preliminary in nature, nor can we make useful statements about long term trends.

4.2.2.1 Detailed results

The Taylor diagram below in **Figure 4-7** shows a short overview of the capabilities of the CO2_TAN_OCFP product. Most TCCON sites cluster around a 0.75 correlation value, but with negative correlation values for Anmeyondo and Darwin (the first due its extremely limited collocated dataset and the latter no doubt due to a combination of the limited seasonal variability in the Southern hemisphere and the short time period covered). Other stations with low correlation values (<0.2) are Ny Alesund, Burgos and Reunion (all featuring very limited temporal overlap). The normalized standard deviation ranges between 0.5 and 1.25 with most sites clustering around the 0.75 mark, indicating that the variability of the TCCON data is (in most cases) smaller. The normalized standard deviation of the bias sits (for most sites) between 1 and 0.6. All this indicates that while OCFP data features a stronger variability (random error and/or seasonal variability) than the TCCON data, the biases still harbors less variability then either of them, an indication of OCFP capturing the natural variability.

There is no real discernible pattern in the mosaic plot (**Figure 4-8**), which shows the mean bi-weekly bias between the satellite and TCCON measurement pairs. August seems to exhibit some more outspoken biases (negative and positive), but since the period covered by the plot is very limited, it is hard to tell if this is indeed a systematic feature or merely coincidence.

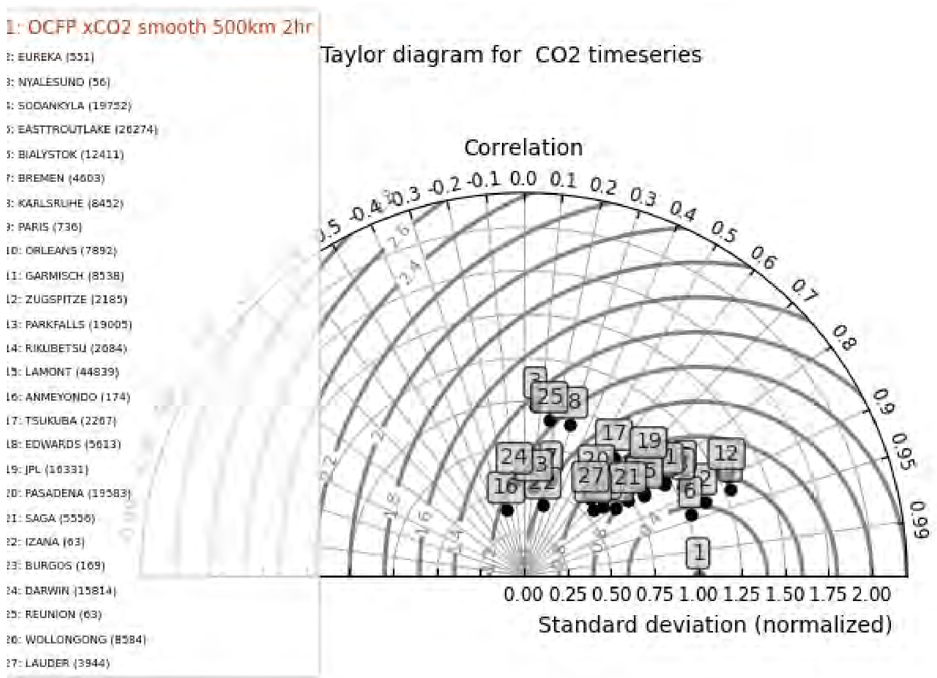


Figure 4-7: Taylor plot of daily averaged XCO_2 TCCON values relative to product CO2_TAN_OCFP. Straight lines correspond with the correlation, light grey lines yield the variability of the TCCON data relative to the satellite variability and the dark grey lines correspond with the variability of the Satellite - TCCON bias relative to the satellite variability.

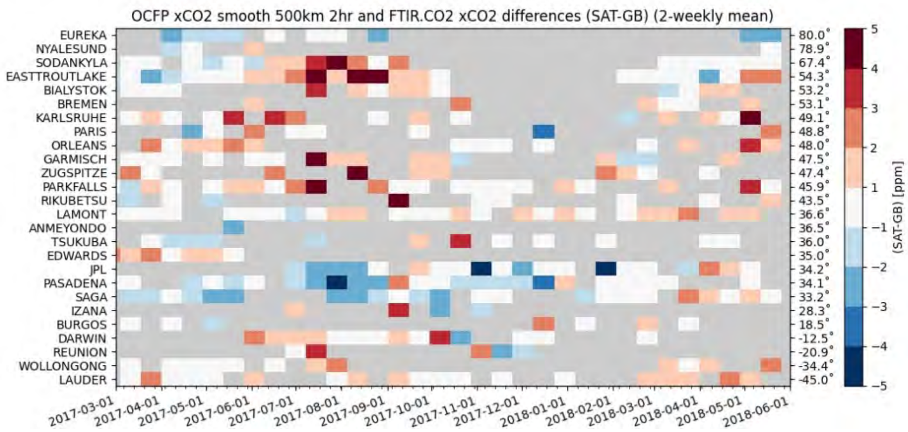


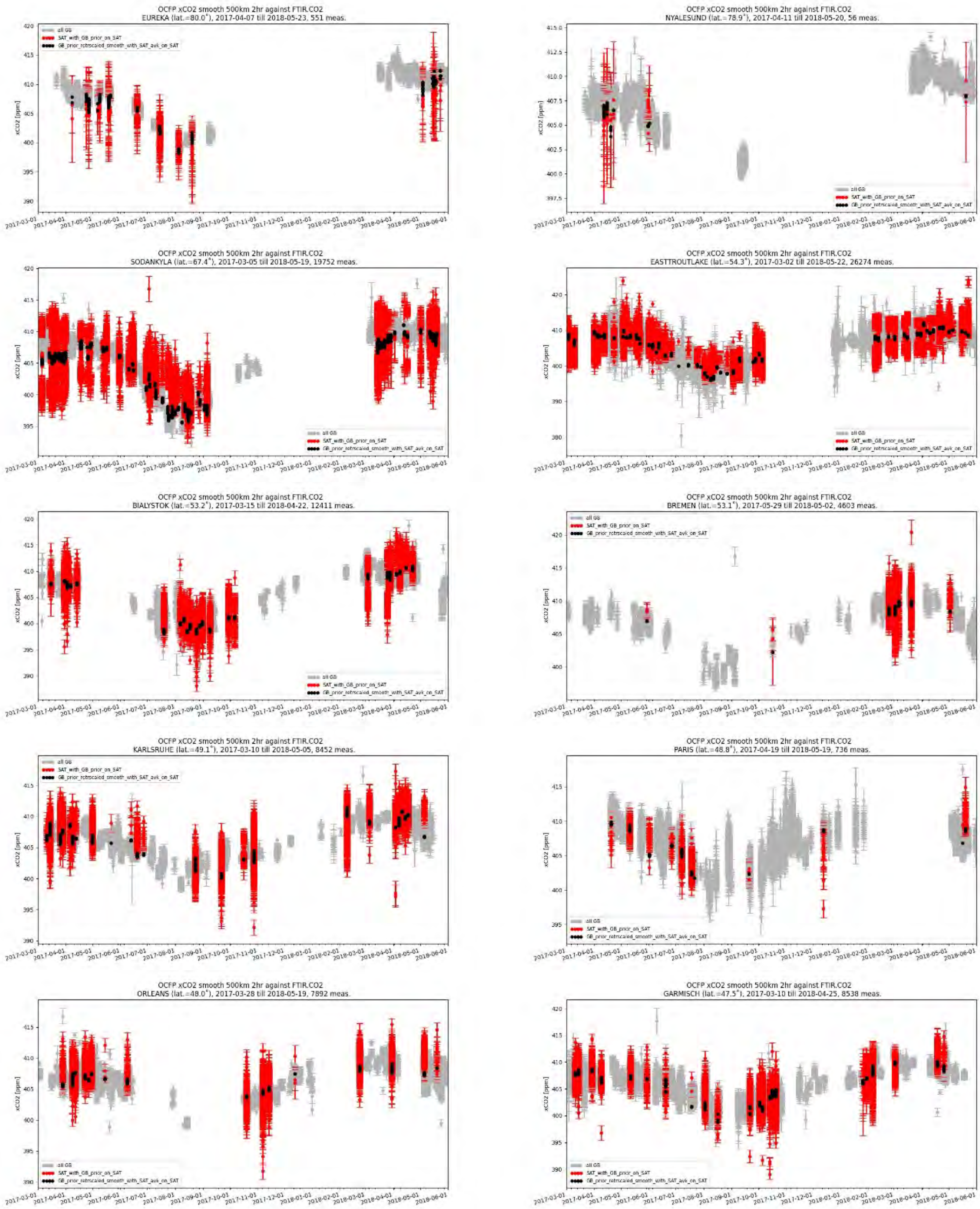
Figure 4-8: Mosaic plot of bi-weekly mean CO2_TAN_OCFP-TCCON XCO_2 biases as a function of time and TCCON station.

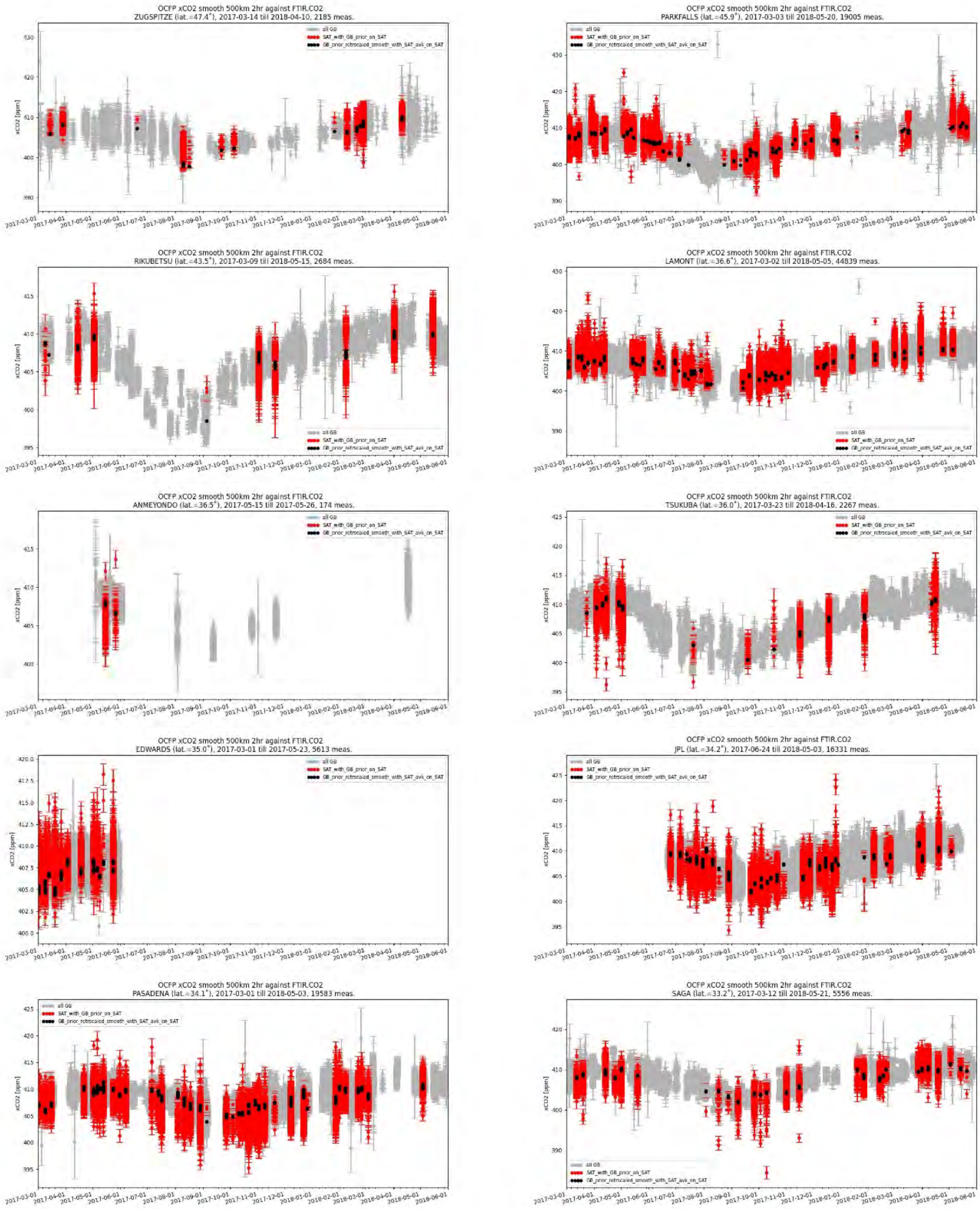
Table 4-5 lists all bias and scatter results derived from individual data pairs at all TCCON stations. The algorithm produces on average ~5600 data pairs per station which corresponds with ~4500 pairs per station per year. The observed median bias ranges between -2.46 (Anmeyondo) and 1.36 ppm(Edwards), while the scatter ranges between 3.23 ppm (Izaña) and 0.98 ppm (Orleans). Note that large bias results are observed at stations that are quite close to one another. One in the Los Angeles basin (Pasadena) and the other just outside on the other side of the San Gabriel Mountain range (Edwards), that separates the basin from the Mojave Desert. Correlation values range between -0.26 (Anmeyondo) and 0.94 (Bialystok), with the median over all stations equal to 0.75. The correlation using all data regardless of station equals 0.82. Given the limited timespan covered by the product, we did not calculate any long term trend. But as can be seen in **Figures 4-9 and 4-10** no clear-cut drift is observable.

Table 4-5: Number of collocated data pairs (N), Correlation (R), Bias, Scatter, long term trend difference (ltt) and uncertainty thereon (ltt_err), seasonal amplitude difference (A) and uncertainty thereon (A_err) as well as the latitude of the TCCON station. The last row lists the median values over all stations. Product: CO2_TAN_OCFP.

STATION	N	R	Bias	Scat	ltt	ltt_err	A	A_err	lat
EUREKA	551	0.93	-0.80	1.31	-	-	-	-	80.0
NYALESUND	56	0.07	-0.49	1.13	-	-	-	-	78.9
SODANKYLA	19752	0.92	0.85	1.55	-	-	-	-	67.4
EASTTROUTLAKE	26274	0.86	0.60	1.85	-	-	-	-	54.3
BIALYSTOK	12411	0.94	0.71	1.40	-	-	-	-	53.2
BREMEN	4603	0.27	0.80	1.04	-	-	-	-	53.1
KARLSRUHE	8452	0.87	0.94	1.63	-	-	-	-	49.1
PARIS	736	0.84	0.55	1.26	-	-	-	-	48.8
ORLEANS	7892	0.80	0.80	0.98	-	-	-	-	48.0
GARMISCH	8538	0.83	0.65	1.74	-	-	-	-	47.5
ZUGSPITZE	2185	0.90	0.99	1.17	-	-	-	-	47.4
PARKFALLS	19005	0.82	0.21	1.70	-	-	-	-	45.9
RIKUBETSU	2684	0.75	-0.51	1.53	-	-	-	-	43.5
LAMONT	44839	0.83	0.53	1.39	-	-	-	-	36.6
ANMEYONDO	174	-0.26	-2.46	1.74	-	-	-	-	36.5
TSUKUBA	2267	0.61	-0.80	1.65	-	-	-	-	36.0
EDWARDS	5613	0.29	1.36	1.48	-	-	-	-	35.0
JPL	16331	0.75	-0.14	1.83	-	-	-	-	34.2
PASADENA	19583	0.62	-1.19	1.84	-	-	-	-	34.1
SAGA	5556	0.81	-0.24	1.71	-	-	-	-	33.2
IZANA	63	0.26	-0.02	3.23	-	-	-	-	28.3
BURGOS	169	0.12	0.92	1.33	-	-	-	-	18.5
DARWIN	15814	-0.10	0.84	1.42	-	-	-	-	-12.5
REUNION	63	0.16	0.78	1.08	-	-	-	-	-20.9
WOLLONGONG	8584	0.72	0.65	1.64	-	-	-	-	-34.4
LAUDER	3944	0.66	1.03	1.24	-	-	-	-	-45.0
MEDIAN	**5584.5**	**0.75**	**0.63**	**1.51**	**-**	**-**	**-**	**-**	**40.1**

The timeseries below in **Figure 4-9** show individual satellite and ground-based fts measurements. As can be seen, and was already apparent from the Taylor diagram, OCFP XCO_2 features a somewhat higher scatter than TCCON, but overall the seasonality is well captured. An occasional outlier is still noticeable (both in the TCCON and OCFP dataset).





OCFP xCO2 smooth 500km 2hr against FTIR.CO2
ZUGSPITZE (lat.=47.4°), 2017-03-14 till 2018-04-10, 2185 meas.
OCFP xCO2 smooth 500km 2hr against FTIR.CO2
PARKFALLS (lat.=45.9°), 2017-03-03 till 2018-05-20, 19005 meas.
OCFP xCO2 smooth 500km 2hr against FTIR.CO2
RIKUBETSU (lat.=43.5°), 2017-03-09 till 2018-05-15, 2684 meas.
OCFP xCO2 smooth 500km 2hr against FTIR.CO2
LAMONT (lat.=36.6°), 2017-03-02 till 2018-05-05, 44839 meas.
OCFP xCO2 smooth 500km 2hr against FTIR.CO2
ANMEYONDO (lat.=36.5°), 2017-05-15 till 2017-05-26, 174 meas.
OCFP xCO2 smooth 500km 2hr against FTIR.CO2
TSUKUBA (lat.=36.0°), 2017-03-23 till 2018-04-16, 2267 meas.
OCFP xCO2 smooth 500km 2hr against FTIR.CO2
EDWARDS (lat.=35.0°), 2017-03-01 till 2017-05-23, 5613 meas.
OCFP xCO2 smooth 500km 2hr against FTIR.CO2
JPL (lat.=34.2°), 2017-06-24 till 2018-05-03, 16331 meas.
OCFP xCO2 smooth 500km 2hr against FTIR.CO2
PASADENA (lat.=34.1°), 2017-03-01 till 2018-05-03, 19583 meas.
OCFP xCO2 smooth 500km 2hr against FTIR.CO2
SAGA (lat.=33.2°), 2017-03-12 till 2018-05-21, 5556 meas.

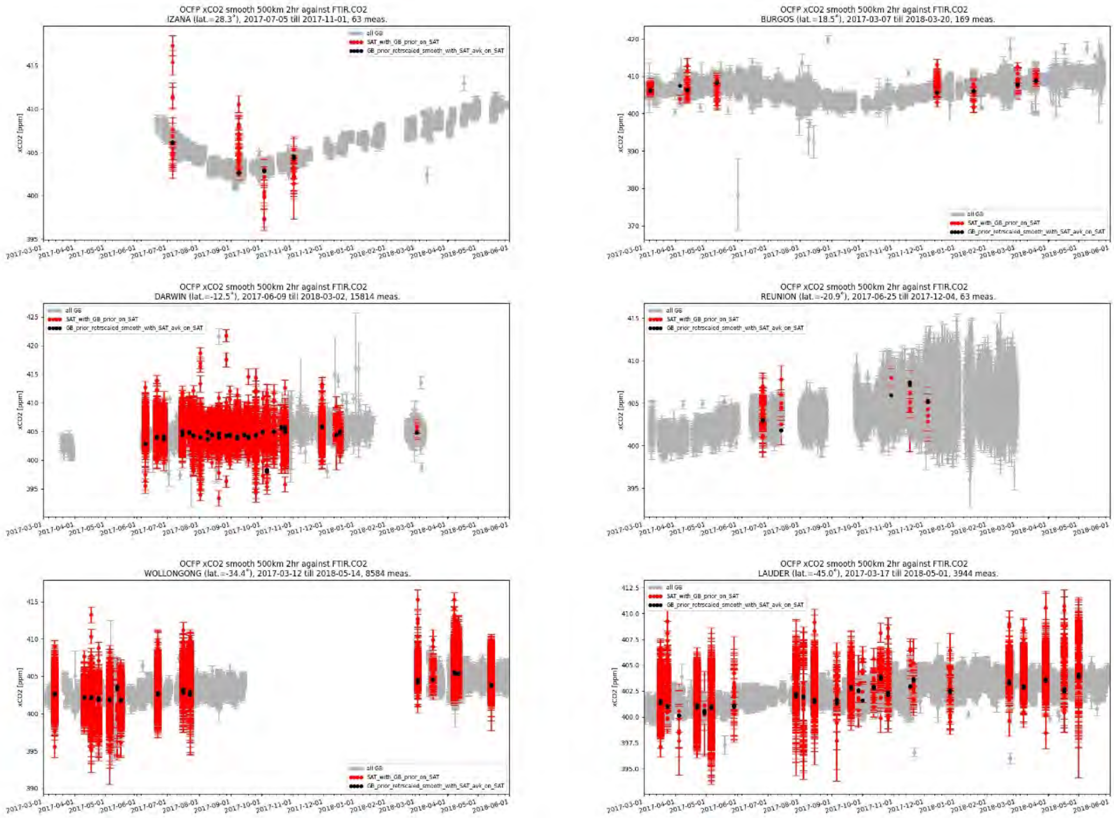


Figure 4-9: XCO_2 timeseries at all TCCON sites (red= CO2_TAN_OCFP data, black is collocated TCCON data and grey are the uncollocated TCCON data).

Figure 4-10 shows monthly median timeseries for TCCON and OCFP XCO_2 for all data that falls within certain latitude bands, namely all sites North of 40°N latitude (top), all sites between 40°N and the equator (mid) and all sites in the Southern hemisphere (bottom). It also features the values for a trend+seasonal fit through both datasets. The obtained long term trends have overlapping standard deviations apart from the Southern hemisphere analysis. Also both FTIR and OCFP XCO_2 seem to follow the same seasonal cycle in the Northern Hemisphere but again not for the Southern hemisphere. However, the observed trend values are, given the short timeframe covered, not robust. Combined with the limited seasonal variability in the Southern hemisphere it is not surprising that we see differences in the fitting parameters.

All in all, we can state that OCFP clearly captures the overall seasonality.

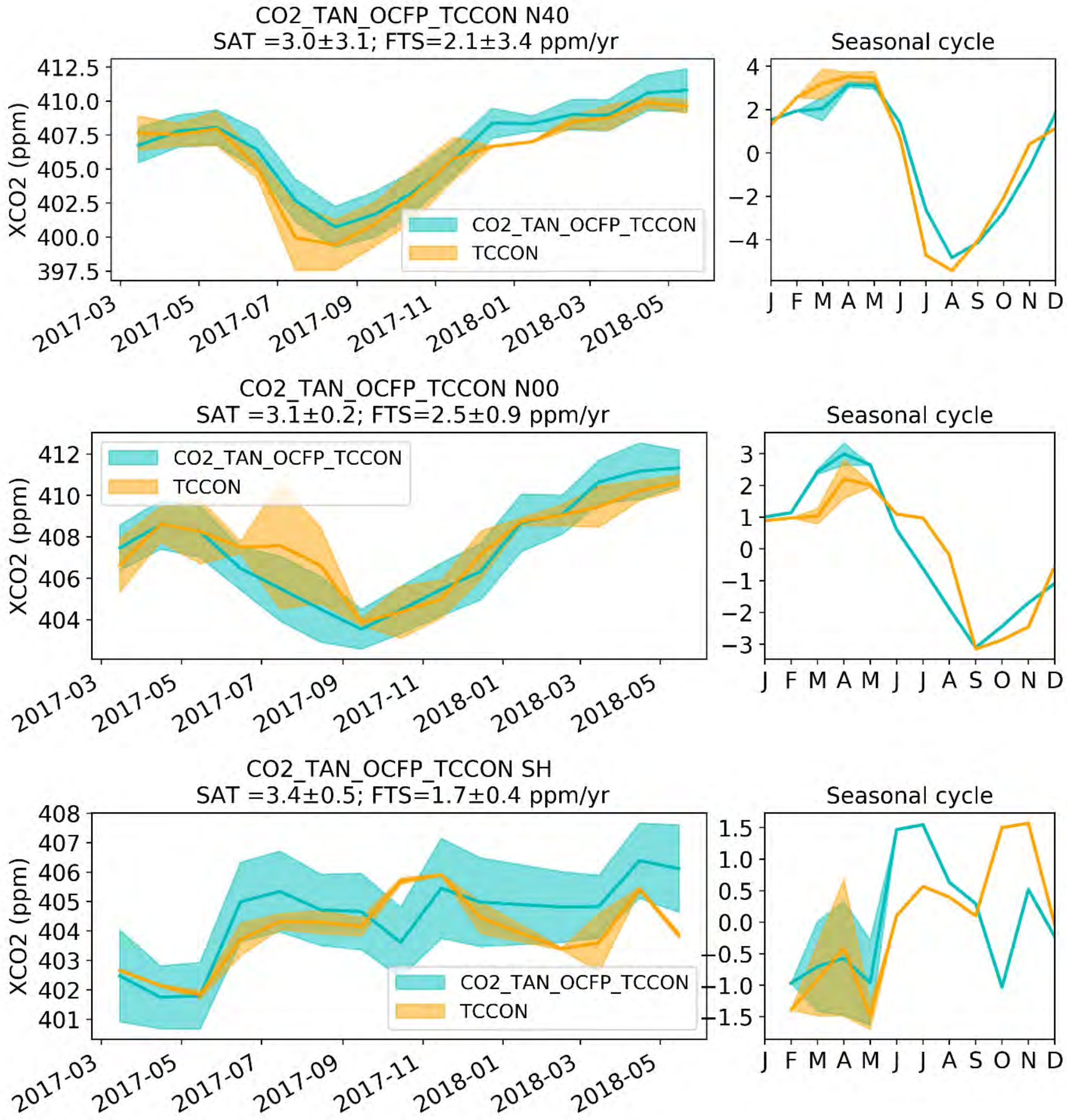


Figure 4-10: Monthly median collocated Sat and TCCON XCO_2 concentrations as a function of time. The shaded areas correspond with the scaled median absolute deviation.

4.2.2.2 Summary

Despite the limited amount of collocated data and the relatively small time period covered, we can already state that we see no obvious defects embedded within the CO2_TAN_OCFP product.

The OCFP reported uncertainty is underestimated by roughly 20% (Uncertainty ratio = 0.77) and the overall bias equals 0.62 ppm and the scatter equals 1.51 ppm. The spatial relative accuracy (RA) has even (just) reached the stated goal requirement (0.5 ppm). The spatio-temporal relative accuracy (SRA), 0.96 ppm, however has not met the stated goal requirement of (>0.5 ppm), nor do its confidence bands overlap. As already mentioned in the analysis of FOCAL XCO2, these numbers ignore TCCON network and collocation errors. Due to the limited temporal coverage, no Stability parameter has been calculated, but we did not see any apparent problems in this area.

Table 4-6 presents an overview of the estimated data quality of CO2_TAN_OCFP, as obtained by the VALT team, from comparisons with TCCON ground-based reference observations. Values in square brackets [ ] correspond with the upper and lower 95% confidence bound on the parameter. The uncertainty ratio features 2 numbers as outlined in the validation method.

Product Quality Summary Table for Product: CO2_TAN_OCFP Level: 2, Version: v01.2.0, Time period covered: 03.2017 – 05.2018 Assessment: Validation Team (VALT)			
Parameter [unit]	**Achieved performance**	**Requirement**	**Comments**
Single measurement precision (1-sigma) in [ppm]	1.51 [1.34, 1.69]	< 8 (T) < 3 (B) < 1 (G)	Computed as the median over all station scaled median absolute differences to TCCON
Uncertainty ratio [-]: Ratio reported uncertainty to standard deviation of satellite-TCCON difference	0.72, 0.77*	-	No requirement but value close to unity expected for a high quality data product with reliable reported uncertainty.
Mean bias (global offset) [ppm]	0.62 [0.45,1.27]	-	No requirement but value close to zero expected for a high quality data product.
Accuracy: Relative systematic error [ppm]	Spatial: 0.50 [-0.07, 0.80] Spatio-temporal: 0.96 [0.62, 1.20]	< 0.5	Spatial: Computed as standard deviation of the biases at the various TCCON sites. Spatio-temporal: As "Spatial" but also considering seasonal biases.
Stability: Drift [ppm/year]	-	< 0.5	Linear drift

4.2.3 Validation results for product CO2_GO2_SRFP

Below we show the validation results of the XCO_2 concentrations as derived by the CO2_GO2_SRFP v2.0.0 algorithm using GOSAT-2 spectra. Data was available from February 2019 up to and including August 2020. The SRFP algorithm provides *a priori* and column averaging kernel information on a 12 layers profile. The covered time period has thus been significantly expanded (end date shifted from October 2019 to August 2020), but has not reached the full 2 years to make an adequate analysis on any long term-trend issues.

4.2.3.1 Detailed results

The Taylor diagram below in **Figure 4-11**shows a short overview of the capabilities of the CO2_GO2_SRFP product. Most TCCON sites cluster around the intercept of the 0.6 correlation line and a normalized standard deviation of ~0.75, with Darwin, Reunion and Bremen, notable exceptions. However, all of these outlier stations have limited collocated data. The normalized standard deviation of most sites range between 0.5 and 1.2, with most being smaller than 1, indicating that on average the variability of the TCCON data is smaller. The normalized standard deviation of the bias sits (for most sites) around 0.8. Notable outliers are again Bremen and Reunion. All this indicates that while SRFP data features a slightly stronger variability (random error and/or seasonal variability) than the TCCON data, the biases still harbors less variability then either of them, an indication of SRFP capturing the natural variability.

There is no real discernible pattern in the mosaic plot (**Figure 4-12**), which shows the mean bi-weekly bias between the satellite and TCCON measurement pairs. However, the period covered by the above plot is limited and there are many gaps in the timeseries, either do due unavailability of TCCON data during winter at high latitudes, interruptions in the measurement cycle or instruments moving to other locations. Sometimes it is merely the result of the sparseness of either data, yielding extremely limiting overlap.

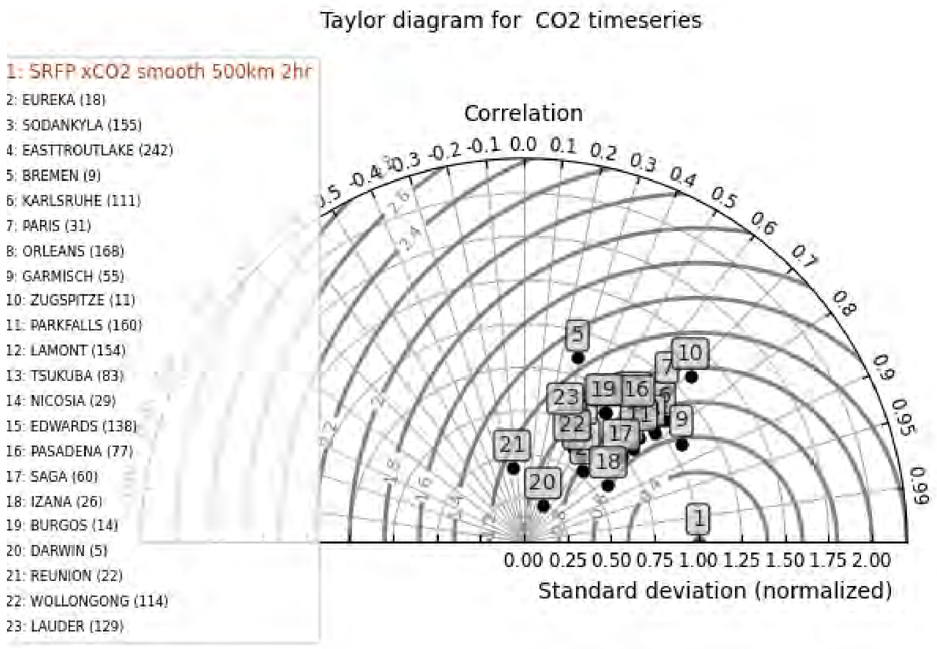


Figure 4-11: Tayor plot of daily averaged XCO_2 TCCON values relative to product CO2_GO2_SRFP. Straight lines correspond with the correlation, light grey lines yield the variability of the TCCON data relative to the satellite variability and the dark grey lines correspond with the variability of the Satellite - TCCON bias relative to the satellite variability.

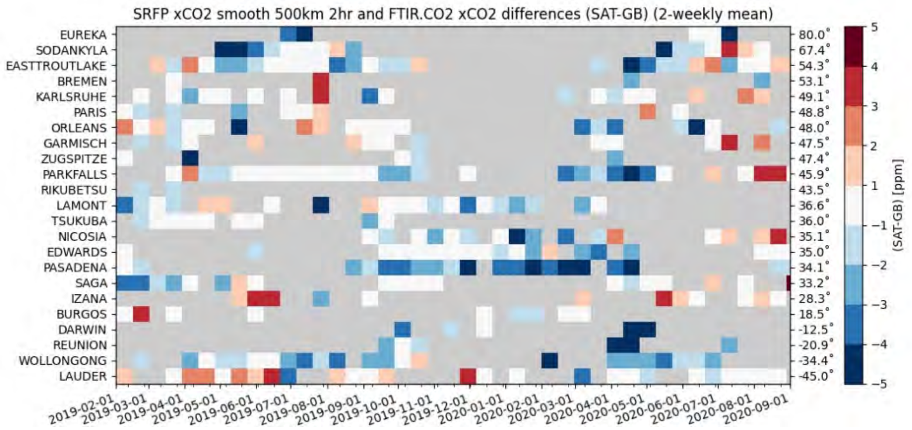


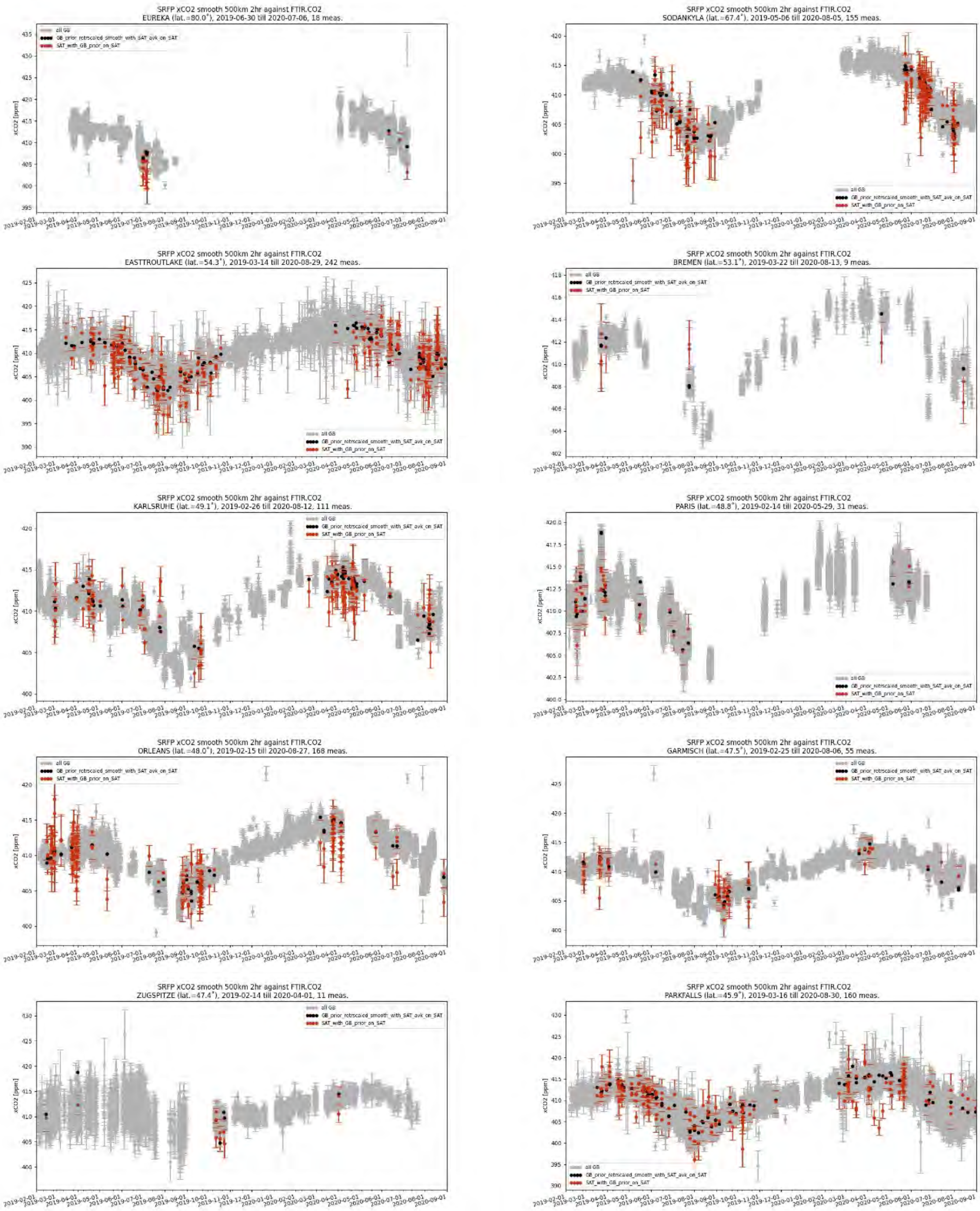
Figure 4-12: Mosaic plot of bi-weekly mean CO2_GO2_SRFP-TCCON XCO_2 biases as a function of time and TCCON station.

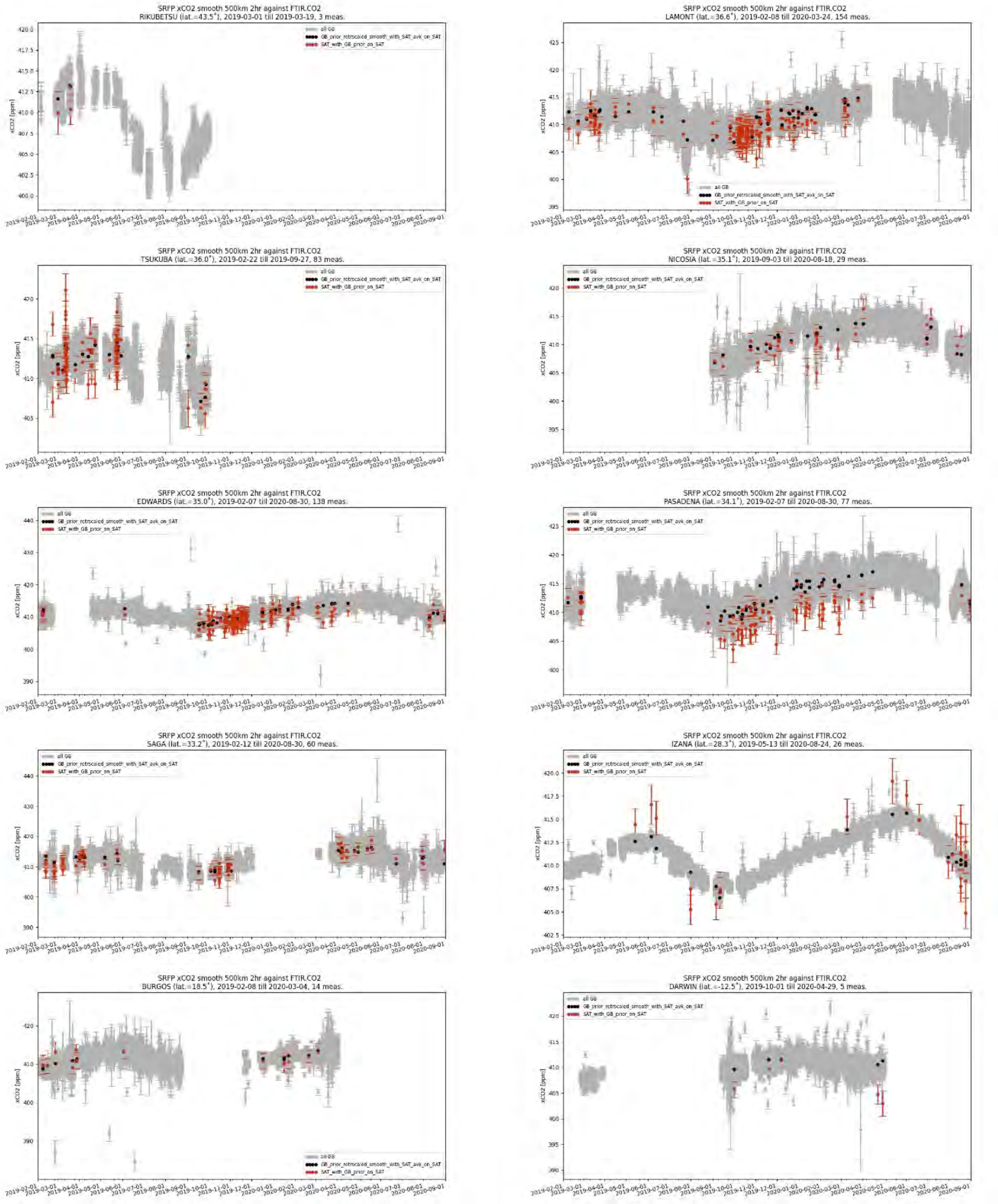
Table 4-7 lists all bias and scatter results derived from individual data pairs at all TCCON stations. The algorithm produces on average ~80 data pairs per station which corresponds with ~50 pairs per station per year. The observed median bias ranges between -4.13 ppm (Eureka) and 0.61 ppm (Lauder and Izaña), while the scatter ranges between 4.06 ppm (Zugspitze) and 1.13 (Burgos). Correlation values range between 0.85 (Garmisch) and -0.16 (Reunion), with most correlation values sitting around 0.6. Of course the limited dataset hampers the correlation values at certain stations. The correlation using all data regardless of station equals 0.73. Given the limited timespan covered by the product, we did not calculate any long term trend. But as can be seen in **Figures 4-13 and 4-14** no clear-cut drift is observable.

Table 4-7: Number of collocated data pairs (N), Correlation (R), Bias, Scatter, long term trend difference (ltt) and uncertainty thereon (ltt_err), seasonal amplitude difference (A) and uncertainty thereon (A_err) as well as the latitude of the TCCON station. The last row lists the median values over all stations. Product: CO2_GO2_SRFP.

STATION	N	R	Bias	Scat	ltt	ltt_err	A	A_err	lat
EUREKA	18	0.63	-4.13	2.93	-	-	-	-	80.0
SODANKYLA	155	0.75	-1.17	2.36	-	-	-	-	67.4
EASTTROUTLAKE	242	0.76	-0.31	2.61	-	-	-	-	54.3
BREMEN	9	0.28	-1.18	2.09	-	-	-	-	53.1
KARLSRUHE	111	0.76	-0.53	1.77	-	-	-	-	49.1
PARIS	31	0.69	-0.54	2.58	-	-	-	-	48.8
ORLEANS	168	0.77	-0.52	1.99	-	-	-	-	48.0
GARMISCH	55	0.85	-0.16	1.58	-	-	-	-	47.5
ZUGSPITZE	11	0.71	-0.72	4.06	-	-	-	-	47.4
PARKFALLS	160	0.74	-0.85	2.47	-	-	-	-	45.9
RIKUBETSU	3	0.66	-1.68	1.58	-	-	-	-	43.5
LAMONT	154	0.64	-0.82	1.65	-	-	-	-	36.6
TSUKUBA	83	0.56	-0.70	1.86	-	-	-	-	36.0
NICOSIA	29	0.44	-0.76	1.84	-	-	-	-	35.1
EDWARDS	138	0.54	-0.57	1.82	-	-	-	-	35.0
PASADENA	77	0.66	-3.10	1.81	-	-	-	-	34.1
SAGA	60	0.76	-0.43	2.22	-	-	-	-	33.2
IZANA	26	0.83	0.61	2.42	-	-	-	-	28.3
BURGOS	14	0.53	-0.34	1.13	-	-	-	-	18.5
DARWIN	5	0.47	-3.95	3.23	-	-	-	-	-12.5
REUNION	22	-0.16	-2.74	2.30	-	-	-	-	-20.9
WOLLONGONG	114	0.46	-1.84	1.86	-	-	-	-	-34.4
LAUDER	129	0.33	0.61	1.72	-	-	-	-	-45.0
MEDIAN	**60**	**0.66**	**-0.72**	**1.99**	**-**	**-**	**-**	**-**	**36.6**

The timeseries below in **Figure 4-13** show individual satellite and ground-based fts measurements. As can be seen, and was already apparent from the Taylor diagram, SRFP XCO_2 features at most stations a somewhat higher scatter than TCCON, but overall the seasonality is well captured.




SRFP xCO2 smooth 500km 2hr against FTIR.CO2
RIKUBETSU (lat.=43.5°), 2019-03-01 till 2019-03-19, 3 meas.
SRFP xCO2 smooth 500km 2hr against FTIR.CO2
LAMONT (lat.=36.6°), 2019-02-08 till 2020-03-24, 154 meas.
SRFP xCO2 smooth 500km 2hr against FTIR.CO2
TSUKUBA (lat.=36.0°), 2019-02-22 till 2019-09-27, 83 meas.
SRFP xCO2 smooth 500km 2hr against FTIR.CO2
NICOSIA (lat.=35.1°), 2019-09-03 till 2020-08-18, 29 meas.
SRFP xCO2 smooth 500km 2hr against FTIR.CO2
EDWARDS (lat.=35.0°), 2019-02-07 till 2020-08-30, 138 meas.
SRFP xCO2 smooth 500km 2hr against FTIR.CO2
PASADENA (lat.=34.1°), 2019-02-07 till 2020-08-30, 77 meas.
SRFP xCO2 smooth 500km 2hr against FTIR.CO2
SAGA (lat.=33.2°), 2019-02-12 till 2020-08-30, 60 meas.
SRFP xCO2 smooth 500km 2hr against FTIR.CO2
IZANA (lat.=28.3°), 2019-05-13 till 2020-08-24, 26 meas.
SRFP xCO2 smooth 500km 2hr against FTIR.CO2
BURGOS (lat.=18.5°), 2019-02-08 till 2020-03-04, 14 meas.
SRFP xCO2 smooth 500km 2hr against FTIR.CO2
DARWIN (lat.=-12.5°), 2019-10-01 till 2020-04-29, 5 meas.

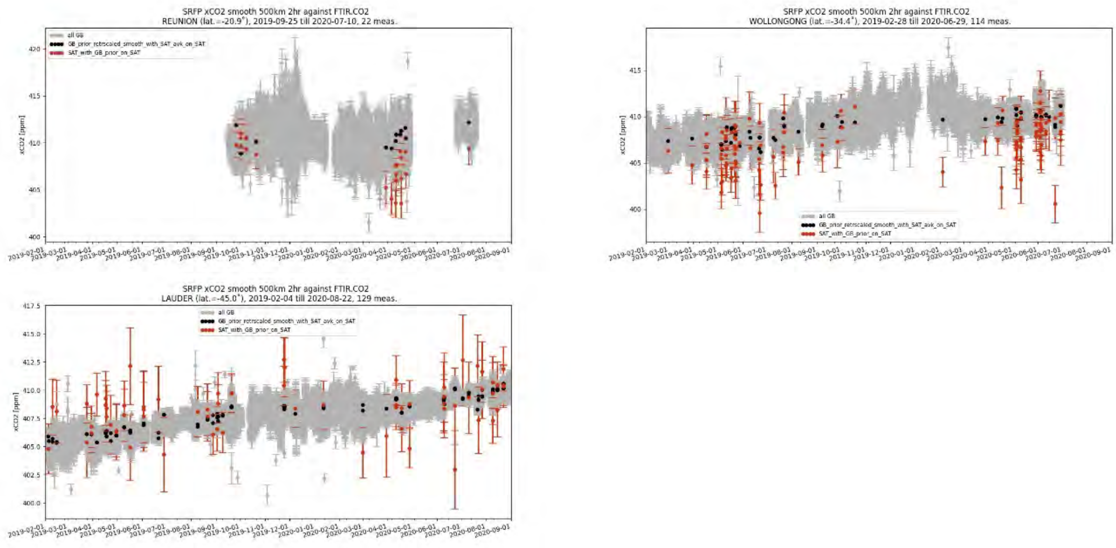


Figure 4-13: XCO_2 timeseries at all TCCON sites (red= CO2_GO2_SRFP data, black is collocated TCCON data and grey are the uncollocated TCCON data).

Figure 4-14 shows monthly median timeseries for TCCON and SRFP XCO_2 for all data that falls within certain latitude bands, namely all sites North of 40°N latitude (top), all sites between 40°N and the equator (mid) and all sites in the Southern hemisphere (bottom). It also features the values for a trend+seasonal fit through both datasets. For both bands in the Northern hemisphere, the obtained long term trends are quasi identical with overlapping standard deviations. For the Southern Hemisphere latitude band a clear difference does exist, but the errors still (just) overlap and if we look at the actual data, a clear outlier is present in February 2020 which can throw off the fit.

All in all, we can state that SRFP clearly captures the overall seasonality.

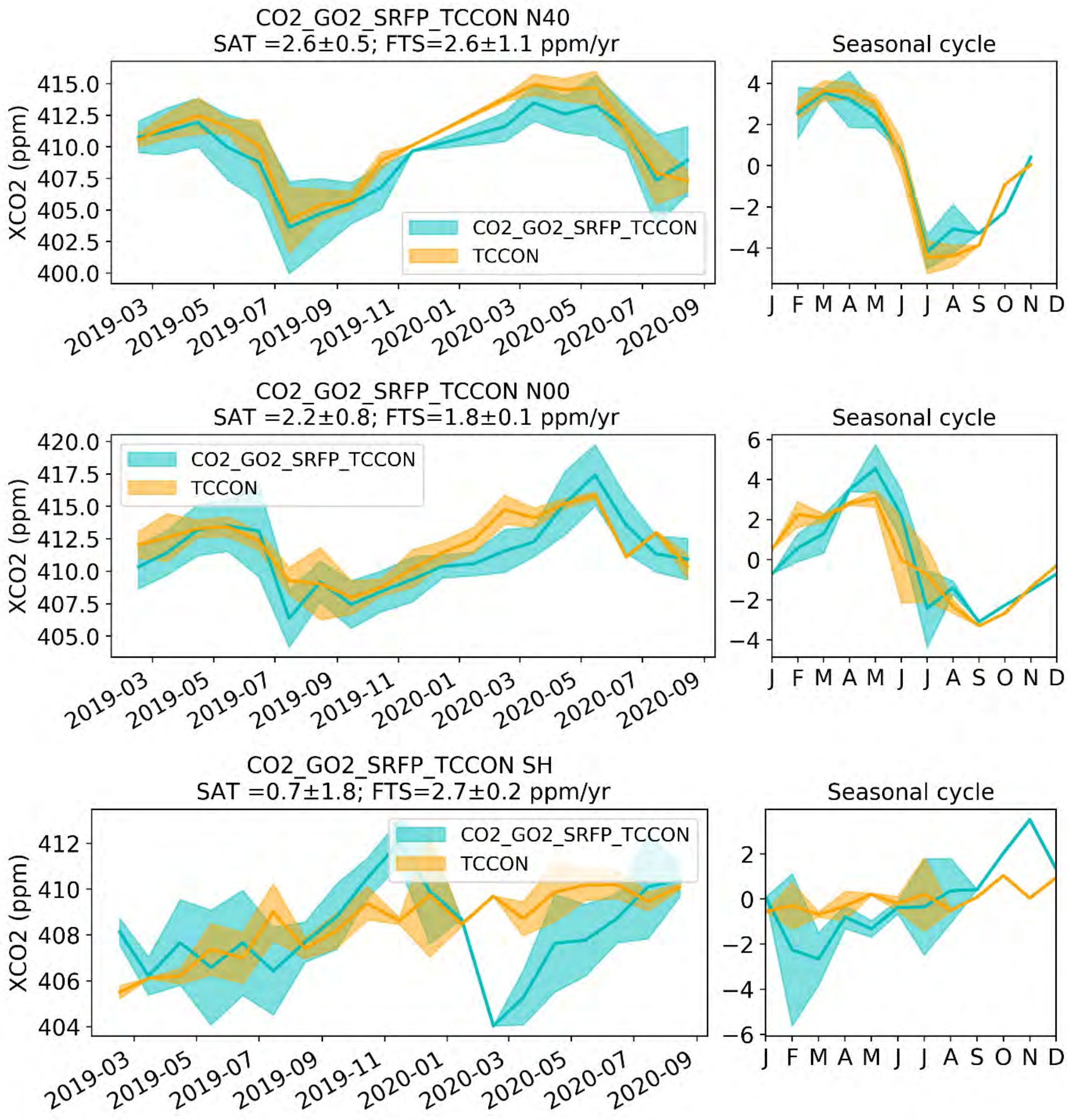


Figure 4-14: Monthly median collocated Sat and TCCON XCO_2 concentrations as a function of time. The shaded areas correspond with the scaled median absolute deviation.

4.2.3.2 Summary

Despite the limited amount of collocated data and the limited time period covered, we can already state that we see no obvious defects embedded within the CO2_GO2_SRFP product apart from a rather consistent negative bias (-0.71 [-0.90 to -0.41] ppm). The SRFP reported uncertainty corresponds closely with our analysis (Uncertainty ratio = 0.93). The spatial (RA) and spatio-temporal relative accuracy (SRA) have not met the stated goal requirement of (>0.5 ppm), but RA's (quite wide) confidence interval overlaps with the target. This is not the case for the SRA. As already mentioned we did not calculate a Stability, due to the limited time period covered but nor did we see any apparent problems in this area.

Table 4-8 presents an overview of the estimated data quality of CO2_GO2_SRFP, as obtained by the VALT team, from comparisons with TCCON ground-based reference observations. Values in square brackets [ ] correspond with the upper and lower 95% confidence bound on the parameter. The uncertainty ratio features 2 numbers as outlined in the validation method.

Product Quality Summary Table for Product: CO2_GO2_SRFP Level: 2, Version: v02.0.0, Time period covered: 2.2019 – 8.2020 Assessment: Validation Team (VALT)			
Parameter [unit]	**Achieved performance**	**Requirement**	**Comments**
Single measurement precision (1-sigma) in [ppm]	2.04 [1.67,2.25]	< 8 (T) < 3 (B) < 1 (G)	Computed as the median over all station scaled median absolute differences to TCCON
Uncertainty ratio [-]: Ratio reported uncertainty to standard deviation of satellite-TCCON difference	0.90, 0.93*	-	No requirement but value close to unity expected for a high quality data product with reliable reported uncertainty.
Mean bias (global offset) [ppm]	-0.71 [-0.90, -0.41]	-	No requirement but value close to zero expected for a high quality data product.
Accuracy: Relative systematic error [ppm]	Spatial: 0.57 [-0.19, 0.91] Spatio-temporal: 1.28 [0.91, 1.79]	< 0.5	Spatial: Computed as standard deviation of the biases at the various TCCON sites. Spatio-temporal: As "Spatial" but also considering seasonal biases.
Stability: Drift [ppm/year]	-	< 0.5	Linear drift

4.2.4 Validation results for product CH4_S5P_WFMD

Below we show the validation results of the XCH_4 concentrations as derived by the CH4_S5P_WFMD v1.5 algorithm using S5P spectra. Data was available from November 2017 up to the end of 2020. The WFMD algorithm provides *a priori* and column averaging kernel data on a 20 layers vertical profile. We have made comparisons with data from both the TCCON and NDACC networks. Note that instead of ‘within 500 km and 2 hour’ collocation criteria, we here have used ‘within 100km and 1 hours’ for TCCON and ‘within 100km and 2 hours’ for NDACC. In the plots and tables below, the TCCON figure/table is always shown first. The obtained Figures of Merit in the summary table (**table 4-11**) pertain to the TCCON analysis only, partly to ensure continuity with previous assessments (even though some aspects of the methodology changed) but also due to the higher systematic uncertainty and high prevalence of high-latitude and mountain sites in the NDACC network which might distort our analysis.

4.2.4.1 Detailed results

The Taylor plot for product CH4_S5P_WFMD is shown in **Figure 4-15**. Most FTIR sites are clustered around the 0.6 correlation line, with the standard deviation of the differences sitting between 0.8 and 1 times the standard deviation of the satellite data itself. The variability on the TCCON data is consistently smaller than that of WFMD apart from the Zugspitze station, with most sites sitting between 50 and 100%. Anmeyondo and Ascension are outliers in this respect, but both also feature a much lower data density.

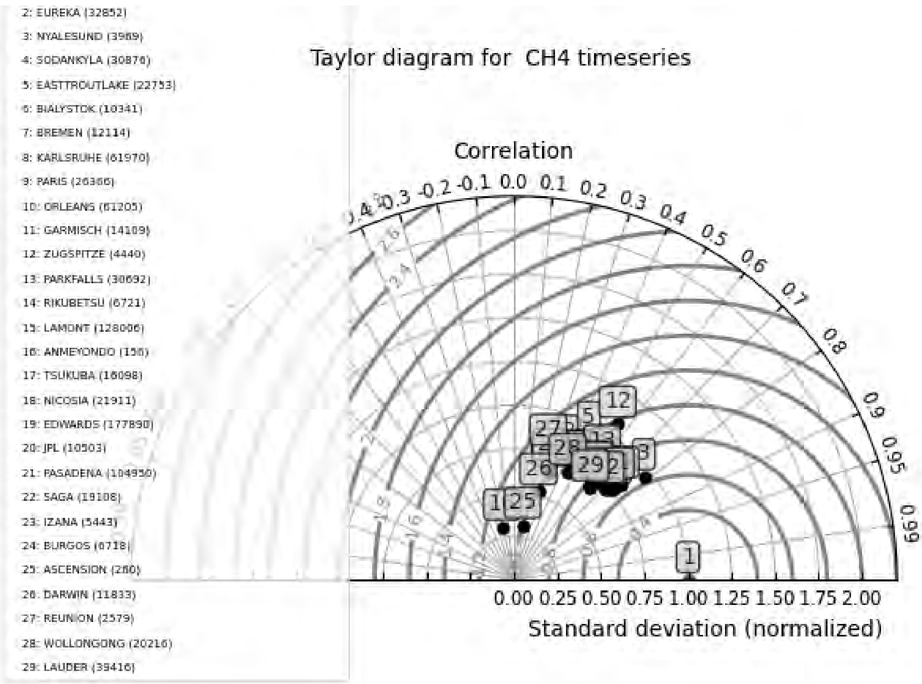


Figure 4-15: Tayor plot of daily averaged XCH4 TCCON values relative to CH4_S5P_WFMD. Straight lines correspond with the correlation, light grey lines yield the variability of the TCCON data relative to

the satellite variability and the dark grey lines correspond with the variability of the Satellite -TCCON bias relative to the satellite variability.

The NDACC Taylorplot shows way more dispersion, indicating either less consistency within the network, less ideal collocation circumstances or a satellite product that is less attuned to the NDACC network. Also the correlation, standard deviation of the difference and standard deviation of the satellite data relative to NDACC yields poorer results. Correlations, on average with a lot of leeway, sits around 0.4, while the standard deviation on the Satellite data has a wide range relative to the NDACC data with some stations showing lower and other higher scatter than NDACC. The scatter on the SAT-NDACC difference, relative to the scatter of the NDACC data itself sits around 1.2. Notable outliers are Toronto and La Reunion Maïdo, with much lower correlation values and Ny Alesund and Mauna Loa with higher correlations than observed at other stations.

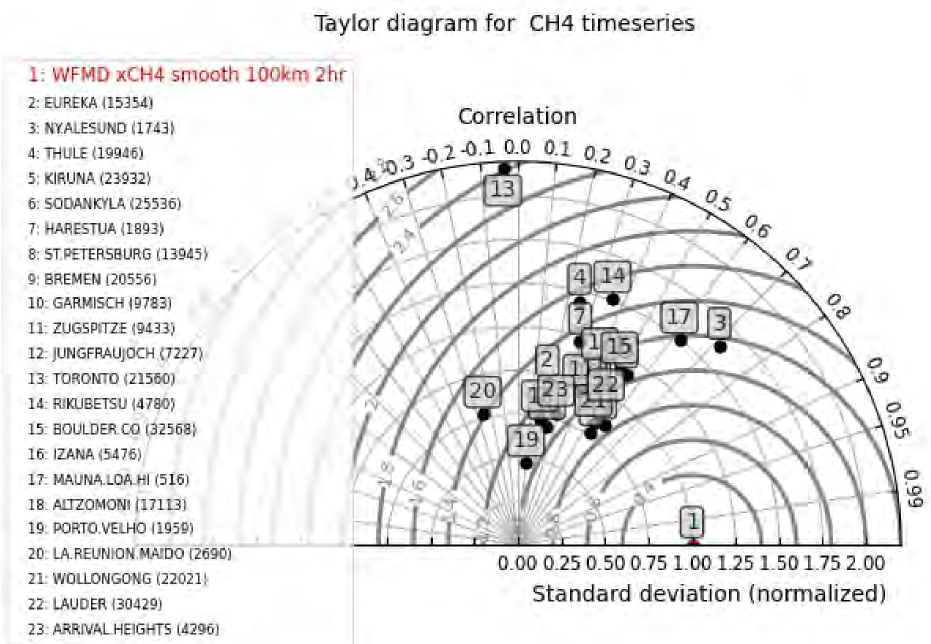


Figure 4-16: Tayor plot of daily averaged XCH4 NDACC values relative to CH4_S5P_WFMD. Straight lines correspond with the correlation, light grey lines yield the variability of the NDACC data relative to the satellite variability and the dark grey lines correspond with the variability of the Satellite -NDACC bias relative to the satellite variability..

The mosaic overview of bi-weekly sat-TCCON biases (**Figure 4-17**) does not reveal any systematic trend over time, nor any as a function of latitude. There are some very pronounced biases (negative in Parkfalls and positive in Zugspitze and Izaña, the latter 2, being high altitude stations).

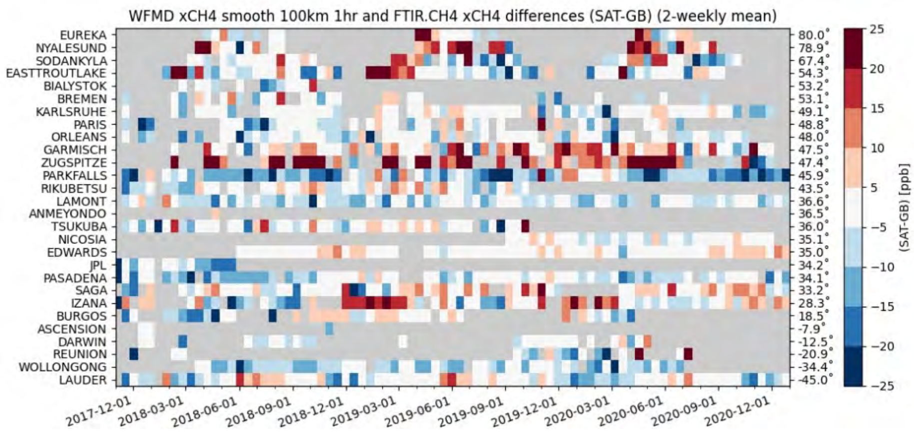


Figure 4-17: Mosaic plot of bi-weekly mean CH4_S5P_WFMD - TCCON XCH_4 biases as a function of time and TCCON station.

For NDACC we see more pronounced differences with strong positive biases at Thule, Altzomoni and Arrival Heights and negative ones at Jungfraujoch, Wollongong and Lauder. For Toronto we even see a shifting bias, with lower values at the start and higher values at the end of the observed timeframe. Paramaribo and Reunion (Maïdo) cover only a tiny fraction of the retrieved timeseries.

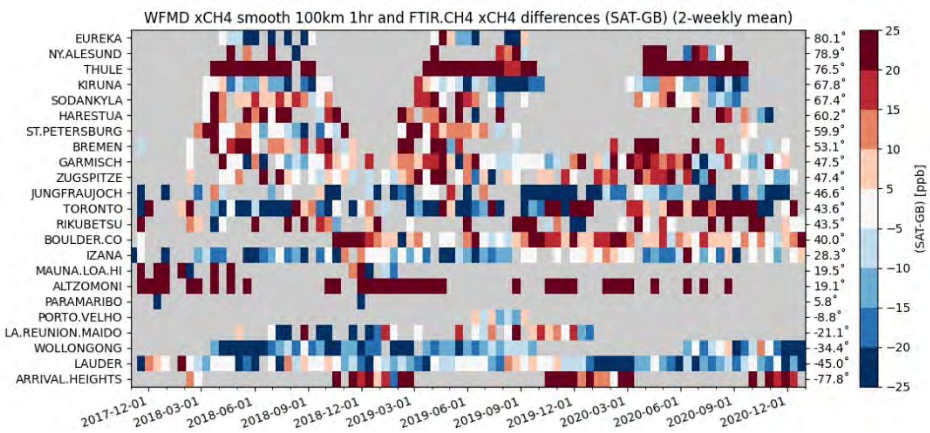


Figure 4-18: Mosaic plot of bi-weekly mean CH4_S5P_WFMD - NDACC XCH_4 biases as a function of time and TCCON station.

Tables 4-9 and 4-10 lists all bias and scatter results derived from individual data pairs at all TCCON and NDACC stations respectively. For TCCON, the algorithm produces on average ~30000 data pairs per station which corresponds with ~10000 pairs per station per year. Also keep in mind that the collocation criteria are substantially stricter. However, the data density suggest they could be even stricter still, to the point we take into account the line-of-sight of the FTIR instrument The observed median bias ranges between -11.87 ppb (Parkfalls) and 23.69 ppb (Zugspitze), while the scatter ranges between 8.35 ppb (Anmeyondon) and 22.56 ppb (Easttroutlake). Correlation values range between -0.22 (Anmeyondo) and 0.78 (Ny Alesund), with most correlation values sitting between 0.6 and 0.75. The correlation of all data, regardless of station, equals 0.84. The long term trend on the bias ranges between -7.8 ppb/year at Zugspitze and 6.8 ppb/year at Eureka. Finally, the seasonal amplitude present in the sat-TCCON bias ranges between 0.6 ppb (Edwards) and 29.1 ppb (Eureka). Of course the latter, being a high latitude station, misses data during autumn and wintertime and cannot capture the full seasonal cycle.

For NDACC (**Table 4-11**), the overall and median correlations are much lower (0.60 and 0.39 respectively). Biases range from a staggering -110.8 ppb (Paramaribo) to 59.3 ppb (Altzomoni). However for the first we only have 3 datapoints, and the latter is a particularly challenging site as it sits in the mountains near Mexico City. It is far from inconceivable that the simple profile extension we employ does not yield satisfying results. However when looking at the collocation area (see **Figure 4-19** left) it looks like the bias extends into the mountainous region so it cannot be the only reason for this discrepancy. Scatter numbers range from 7.73 (Paramaribo again) and 52.6 ppb (Toronto). The collocation area of the latter is also shown in **Figure 4-19** (right). Here we see large differences between datapoints that are sampled over the urban area and those over the countryside. But if we look at the Toronto NDACC data itself (see timeseries in **Figure 4-21**) it is obvious that it is foremost the ground-based data that is showing large variability, no doubt in part due to differing airmasses (urban and rural) being sampled at different times. Long term trends range between -7.0 ppb/year (Harestua) and 29.6 ppb/year (Toronto). The latter is a clear outlier with the next highest positive trends are at 6.3 (Izaña) and 3.9 (St. Petersburg) ppb/year. Looking at the Torono timeseries again we see 2 extreme NDACC outliers at the very end of the timewindow which could heavily impact the long term trend results. However the trend is also visible in the mosaic plot (**Figure 4-18**) so it might point to a more serious issue with the Toronto station.

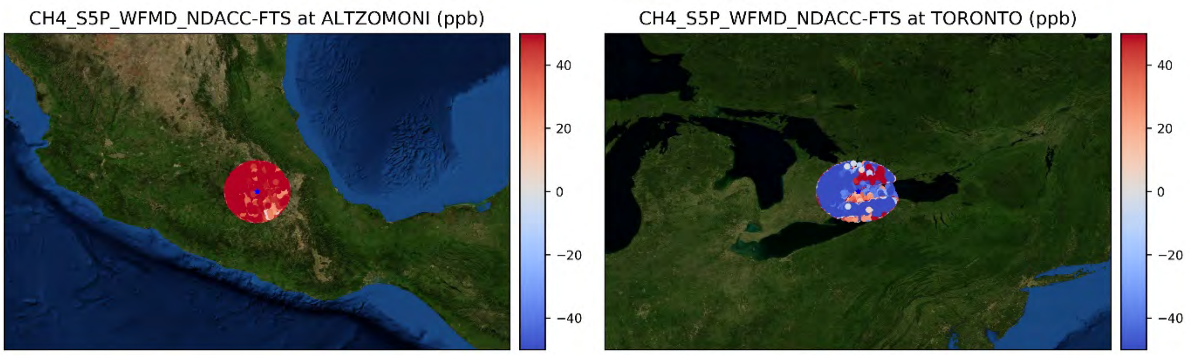


Figure 4-19: Bias between WFMD XCH_4 with respect to several NDACC sites (which showed particular high overall biases).

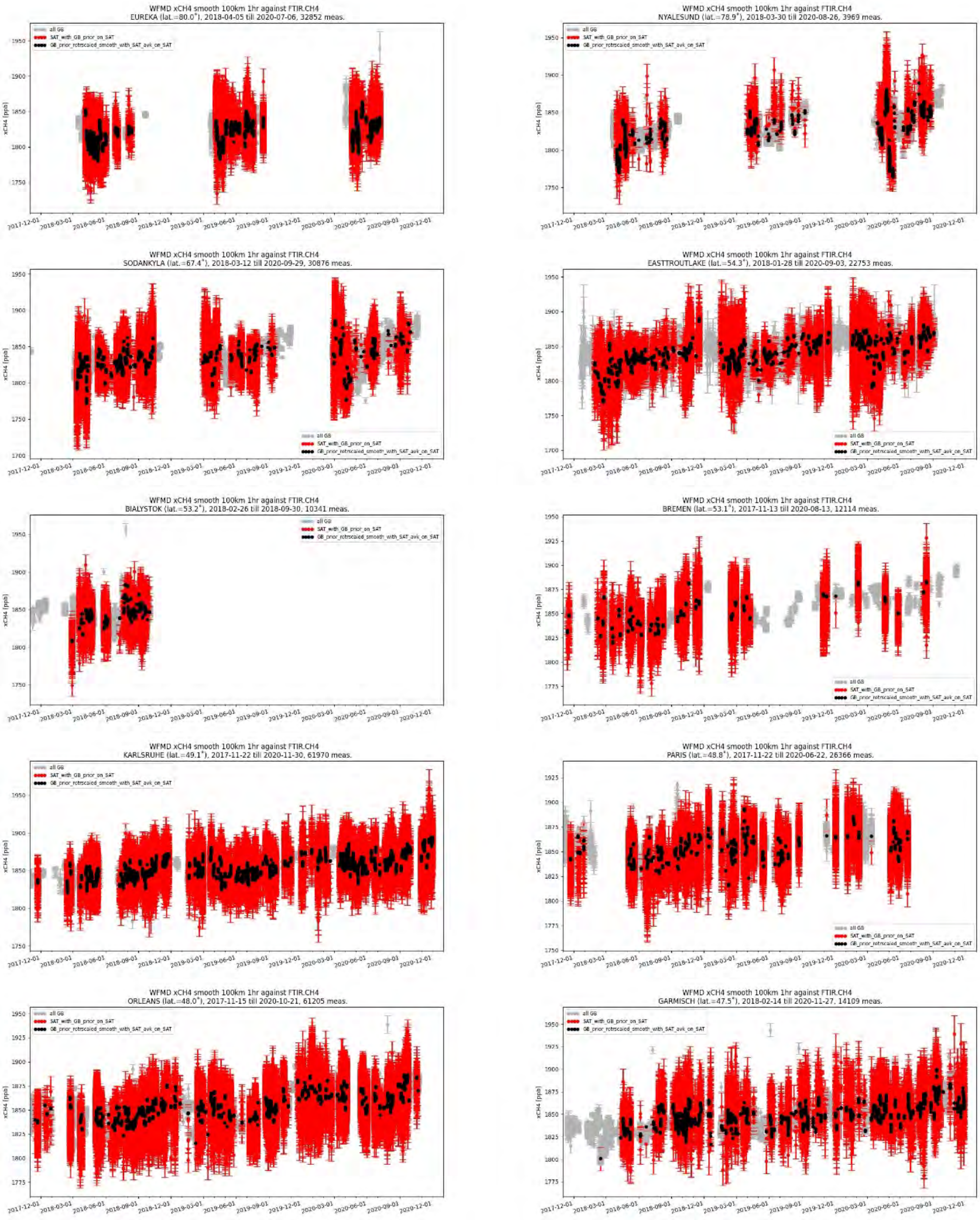
Table 4-9: Number of collocated data pairs (N), Correlation (R), Bias, Scatter, long term trend difference (ltt) and uncertainty thereon (ltt_err), seasonal amplitude difference (A) and uncertainty thereon (A_err) as well as the latitude of the TCCON station. The last row lists the median values over all stations. Product: CH4_S5P_WFMD.

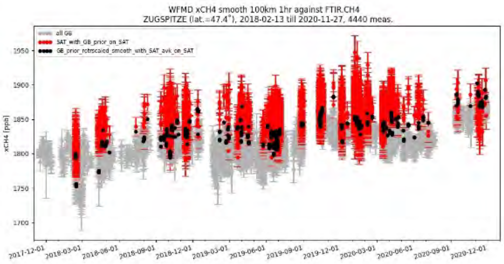
STATION	N	R	Bias	Scat	ltt	ltt_err	A	A_err	lat
EUREKA	32852	0.40	5.57	17.85	6.80	3.60	29.12	7.01	80.0
NYALESUND	3969	0.78	12.77	19.07	2.47	2.40	9.67	5.83	78.9
SODANKYLA	30876	0.68	2.18	16.18	-0.96	2.07	3.63	1.68	67.4
EASTTROUTLAKE	22753	0.47	3.62	22.56	-1.30	1.70	4.53	1.27	54.3
BIALYSTOK	10341	0.45	-1.07	11.79	-	-	-	-	53.2
BREMEN	12114	0.62	0.54	12.97	-3.70	1.24	4.62	2.09	53.1
KARLSRUHE	61970	0.61	1.76	12.50	-0.31	1.02	3.66	0.93	49.1
PARIS	26366	0.54	0.32	12.62	4.21	1.65	2.92	1.36	48.8
ORLEANS	61205	0.67	0.17	11.37	0.77	1.18	1.10	0.88	48.0
GARMISCH	14109	0.46	7.38	14.65	0.49	1.57	3.60	1.54	47.5
ZUGSPITZE	4440	0.56	23.69	18.23	-7.80	2.59	7.78	2.38	47.4
PARKFALLS	30692	0.60	-11.87	14.75	-1.63	1.29	4.91	1.16	45.9
RIKUBETSU	6721	0.74	-0.17	14.88	-	-	-	-	43.5
LAMONT	128006	0.75	-5.83	12.50	-0.63	0.79	1.11	0.75	36.6
ANMEYONDO	156	-0.22	8.79	8.35	-	-	-	-	36.5
TSUKUBA	16098	0.61	1.49	13.29	-	-	-	-	36.0
NICOSIA	21911	0.63	-1.14	12.10	-	-	-	-	35.1
EDWARDS	177890	0.73	0.92	12.43	1.11	0.60	4.82	0.50	35.0
JPL	10503	0.32	-11.36	16.51	-	-	-	-	34.2
PASADENA	104950	0.74	-5.50	13.56	1.10	0.64	4.33	0.57	34.1
SAGA	19108	0.72	6.41	15.75	0.76	1.05	3.61	1.16	33.2
IZANA	5443	0.31	-5.04	17.85	2.06	1.73	8.78	1.50	28.3
BURGOS	6718	0.63	-0.64	13.09	2.89	1.90	5.21	1.32	18.5
ASCENSION	260	0.16	-2.61	8.53	-	-	-	-	-7.9
DARWIN	11833	0.27	-2.65	10.15	2.11	2.00	0.55	1.55	-12.5
REUNION	2579	0.26	-8.90	15.25	-0.95	3.28	16.70	5.19	-20.9
WOLLONGONG	20216	0.45	-5.95	15.60	0.98	1.15	0.62	1.02	-34.4
LAUDER	39416	0.64	-3.16	14.01	-1.27	1.00	6.41	0.78	-45.0
MEDIAN	**17603**	**0.61**	**0**	**13.79**	**0.76**	**1.57**	**4.53**	**1.32**	**36.6**

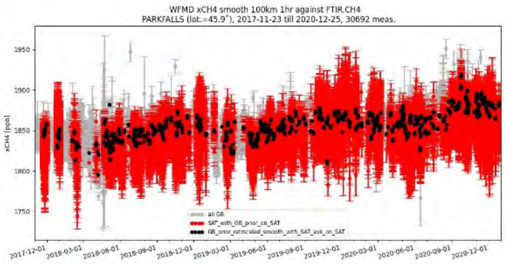
Table 4-10: Number of collocated data pairs (N), Correlation (R), Bias, Scatter, long term trend difference (ltt) and uncertainty thereon (ltt_err), seasonal amplitude difference (A) and uncertainty thereon (A_err) as well as the latitude of the NDACC station. The last row lists the median values over all stations. Product: CH4_S5P_WFMD.

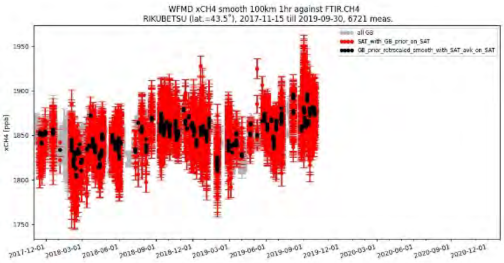
STATION	N	R	Bias	Scat	ltt	ltt_err	A	A_err	lat
EUREKA	15354	0.18	-5.12	22.1	-	-	-	-	80.1
NY.ALESUND	1743	0.71	32.03	31.4	-0.31	4.46	44.76	14.33	78.9
THULE	19946	0.25	38.09	23.4	3.50	3.56	20.87	5.75	76.5
KIRUNA	23932	0.55	-5.13	18.6	1.78	1.81	6.64	1.68	67.8
SODANKYLA	25536	0.59	10.80	17.6	-2.82	1.75	7.15	3.22	67.4
HARESTUA	1893	0.29	21.03	21.8	-6.99	4.06	16.17	5.77	60.2
ST.PETERSBURG	13945	0.48	8.97	17.4	3.92	1.87	10.26	1.90	59.9
BREMEN	20556	0.54	19.34	19.6	2.33	3.62	8.99	2.66	53.1
GARMISCH	9783	0.42	5.09	17.7	1.13	2.00	12.24	1.71	47.5
ZUGSPITZE	9433	0.39	8.27	18.6	-0.05	1.93	8.78	1.99	47.4
JUNGFRAUJOCH	7227	0.42	-16.50	19.7	-4.13	2.82	15.34	2.07	46.6
TORONTO	21560	-0.04	-7.42	52.6	29.60	5.52	6.73	5.21	43.6
RIKUBETSU	4780	0.36	13.17	29.7	-0.61	7.72	36.18	5.64	43.5
BOULDER.CO	32568	0.50	8.31	16.5	-2.53	2.41	3.96	1.83	40
IZANA	5476	0.24	-16.35	16.5	6.28	1.82	1.44	2.09	28.3
MAUNA.LOA.HI	516	0.62	26.10	16.6	-	-	-	-	19.5
ALTZOMONI	17113	0.17	59.26	18.1	3.41	2.35	10.95	3.73	19.1
PARAMARIBO	3	-0.53	-110.79	7.73	-	-	-	-	5.8
PORTO.VELHO	1959	0.10	-4.11	18.5	-	-	-	-	-8.8
LA.REUNION.MAI	2690	-0.26	-7.89	20.7	-	-	-	-	-21.1
WOLLONGONG	22021	0.53	-12.96	18.4	0.49	2.17	9.48	2.16	-34.4
LAUDER	30429	0.54	-10.53	17.8	-6.57	1.23	5.42	1.30	-45.0
ARRIVAL.HEIGHTS	4296	0.28	24.87	20.5	-0.77	2.80	21.85	6.30	-77.8
MEDIAN	**9783**	**0.39**	**8.27**	**18.6**	**0.22**	**2.38**	**9.87**	**2.41**	**43.6**

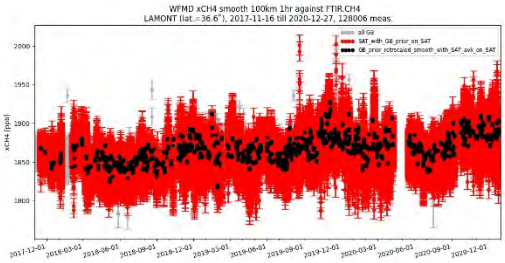
Figure 4-20 shows all collocated WFMD and TCCON data time series. From these figures, it is clear that the variability of WFMD XCH_4 is substantially stronger. Also a fair amount of, particularly negative, outliers is present at many stations.

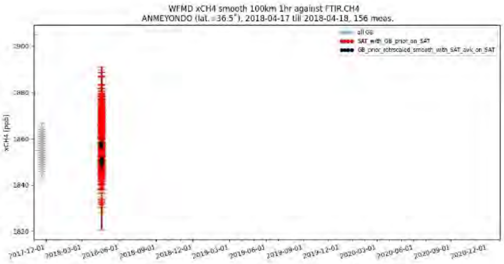

WFMD xCH4 smooth 100km 1hr against FTIR.CH4
EUREKA (lat.=80.0°), 2018-04-05 till 2020-07-06, 32852 meas.
WFMD xCH4 smooth 100km 1hr against FTIR.CH4
NYALESUND (lat.=78.9°), 2018-03-30 till 2020-08-26, 3969 meas.
WFMD xCH4 smooth 100km 1hr against FTIR.CH4
SODANKYLA (lat.=67.4°), 2018-03-12 till 2020-09-29, 30876 meas.
WFMD xCH4 smooth 100km 1hr against FTIR.CH4
EASTTROUTLAKE (lat.=54.3°), 2018-01-28 till 2020-09-03, 22753 meas.
WFMD xCH4 smooth 100km 1hr against FTIR.CH4
BIALYSTOK (lat.=53.2°), 2018-02-26 till 2018-09-30, 10341 meas.
WFMD xCH4 smooth 100km 1hr against FTIR.CH4
BREMEN (lat.=53.1°), 2017-11-13 till 2020-08-13, 12114 meas.
WFMD xCH4 smooth 100km 1hr against FTIR.CH4
KARLSRUHE (lat.=49.1°), 2017-11-22 till 2020-11-30, 61970 meas.
WFMD xCH4 smooth 100km 1hr against FTIR.CH4
PARIS (lat.=48.8°), 2017-11-22 till 2020-06-22, 26366 meas.
WFMD xCH4 smooth 100km 1hr against FTIR.CH4
ORLEANS (lat.=48.0°), 2017-11-15 till 2020-10-21, 61205 meas.
WFMD xCH4 smooth 100km 1hr against FTIR.CH4
GARMISCH (lat.=47.5°), 2018-02-14 till 2020-11-27, 14109 meas.
all GB
SAT_with_GB_prior_on_SAT
GB_prior_retrscaled_smooth_with_SAT_avk_on_SAT
xCH4 [ppb]

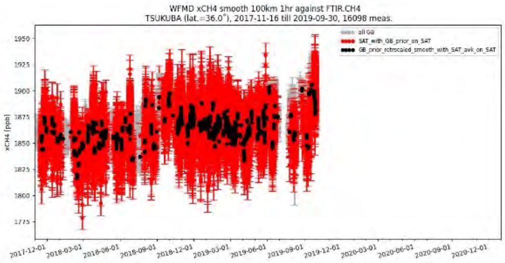

WFMD xCH4 smooth 100km 1hr against FTIR.CH4
ZUGSPITZE (lat.=47.4°), 2018-02-13 till 2020-11-27, 4440 meas.
all GB
SAT_with_GB_prior_on_SAT
GB_prior_rctrscaled_smooth_with_SAT_avk_on_SAT
xCH4 [ppb]

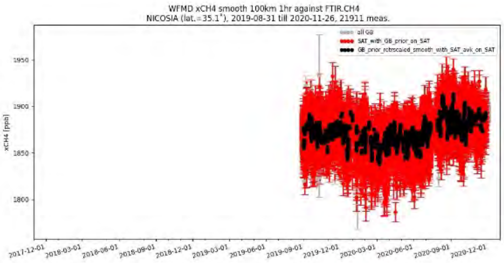

WFMD xCH4 smooth 100km 1hr against FTIR.CH4
PARKFALLS (lat.=45.9°), 2017-11-23 till 2020-12-25, 30692 meas.
xCH4 [ppb]

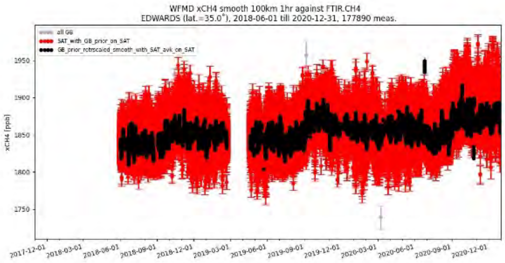

WFMD xCH4 smooth 100km 1hr against FTIR.CH4
RIKUBETSU (lat.=43.5°), 2017-11-15 till 2019-09-30, 6721 meas.
xCH4 [ppb]

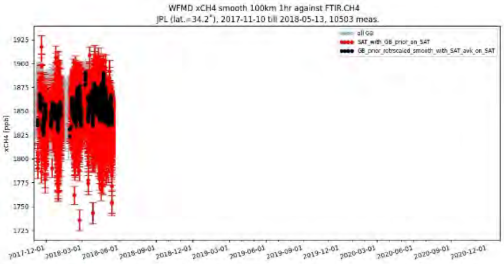

WFMD xCH4 smooth 100km 1hr against FTIR.CH4
LAMONT (lat.=36.6°), 2017-11-16 till 2020-12-27, 128006 meas.
xCH4 [ppb]

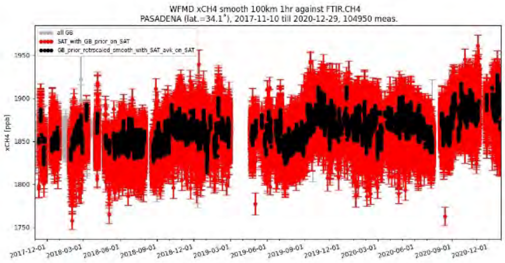

WFMD xCH4 smooth 100km 1hr against FTIR.CH4
ANMEYONDO (lat.=36.5°), 2018-04-17 till 2018-04-18, 156 meas.
xCH4 [ppb]


WFMD xCH4 smooth 100km 1hr against FTIR.CH4
TSUKUBA (lat.=36.0°), 2017-11-16 till 2019-09-30, 16098 meas.
xCH4 [ppb]


WFMD xCH4 smooth 100km 1hr against FTIR.CH4
NICOSIA (lat.=35.1°), 2019-08-31 till 2020-11-26, 21911 meas.
xCH4 [ppb]


WFMD xCH4 smooth 100km 1hr against FTIR.CH4
EDWARDS (lat.=35.0°), 2018-06-01 till 2020-12-31, 177890 meas.
xCH4 [ppb]


WFMD xCH4 smooth 100km 1hr against FTIR.CH4
JPL (lat.=34.2°), 2017-11-10 till 2018-05-13, 10503 meas.
xCH4 [ppb]


WFMD xCH4 smooth 100km 1hr against FTIR.CH4
PASADENA (lat.=34.1°), 2017-11-10 till 2020-12-29, 104950 meas.
xCH4 [ppb]

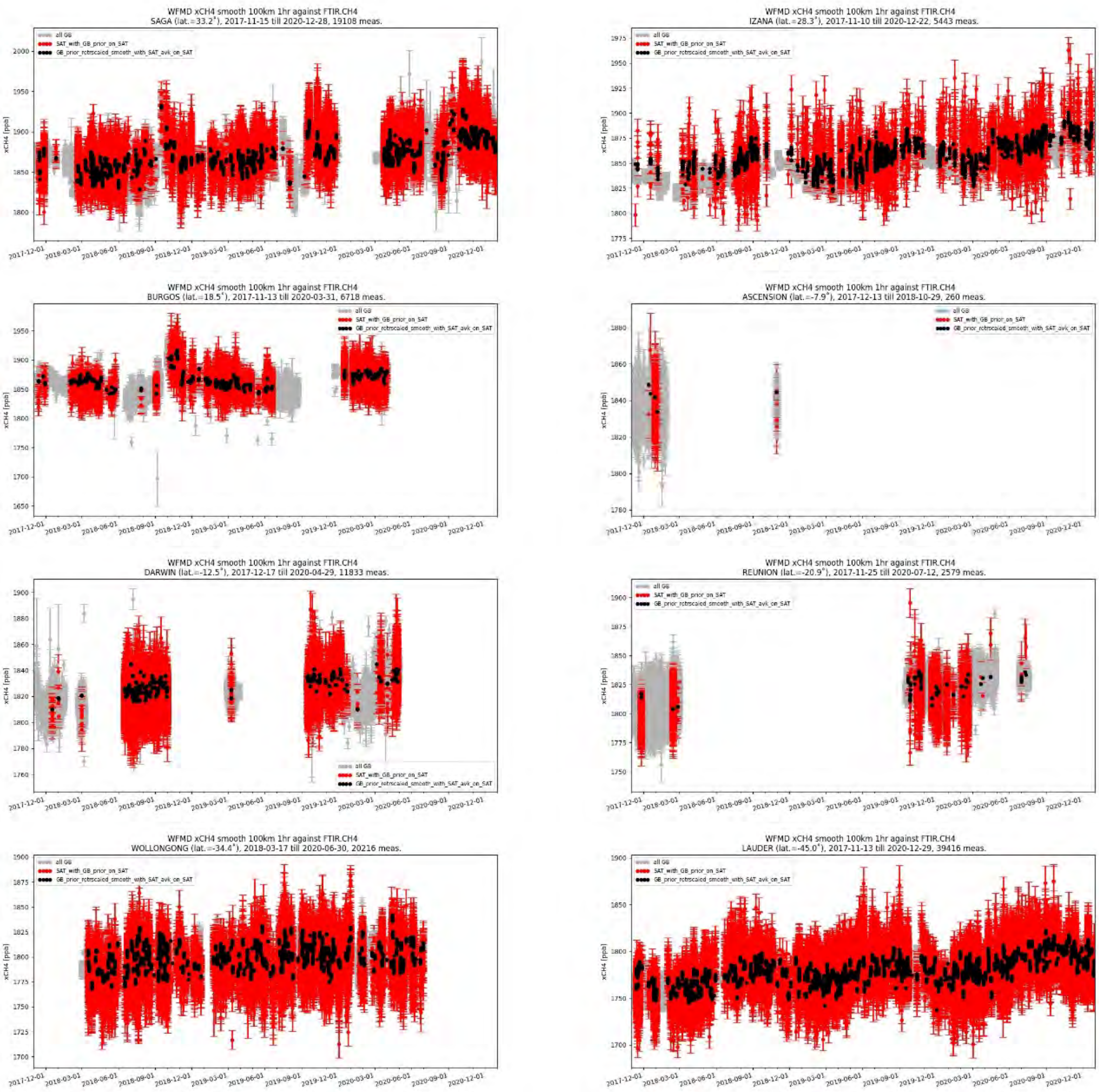
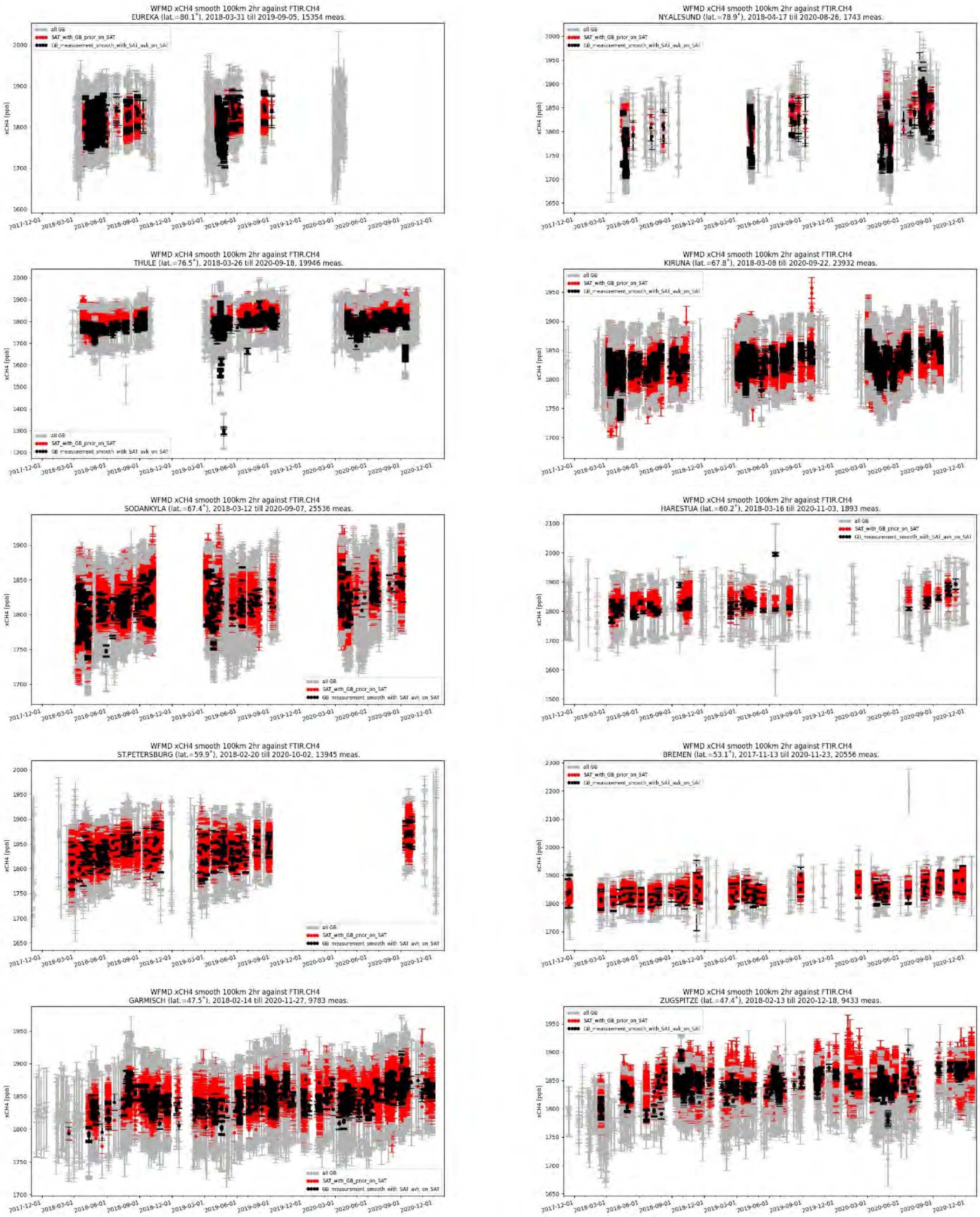
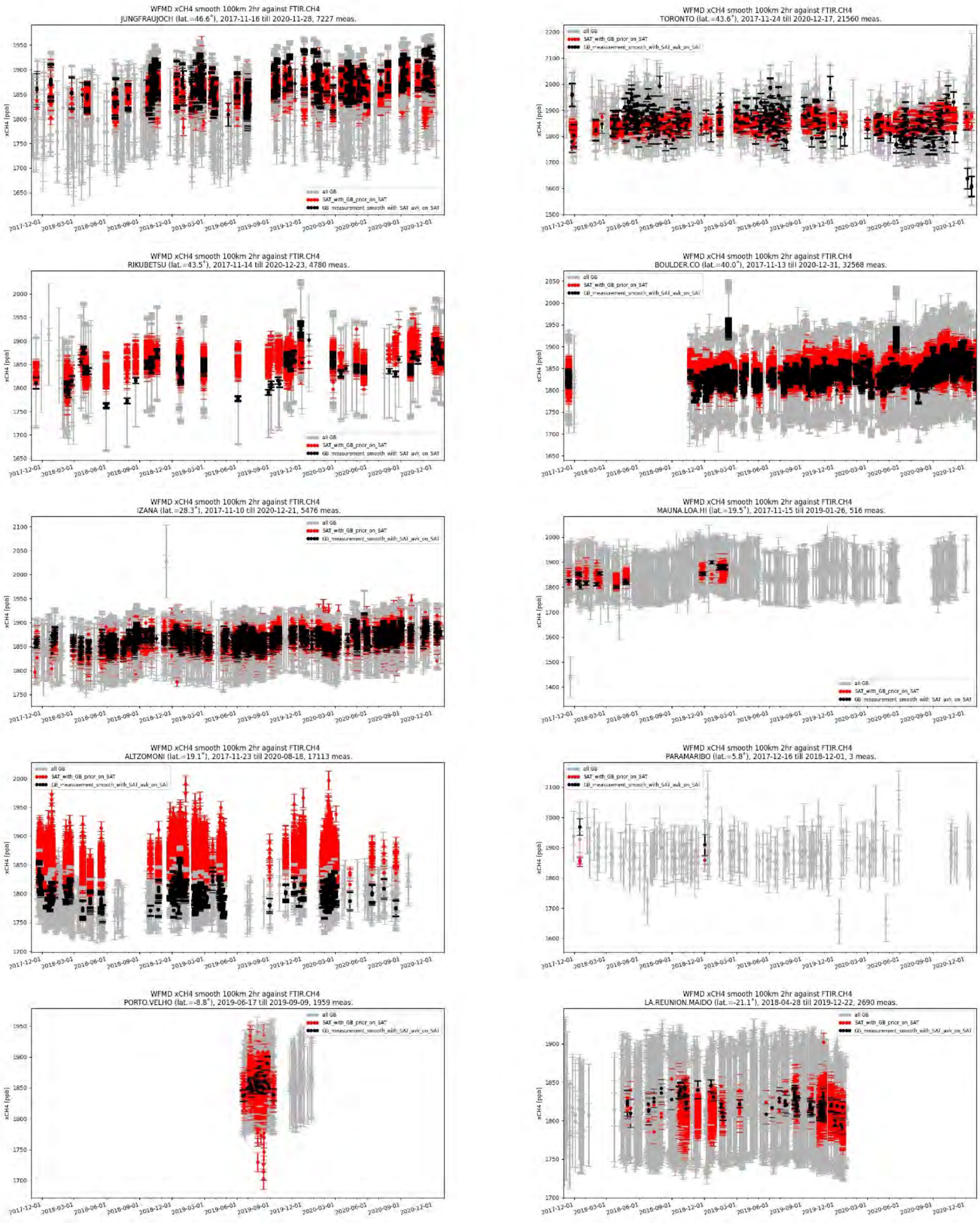


Figure 4-20: Timeseries of XCH_4 TCCON (collocated=black, all=grey) and CH4_S5P_WFMD (red) data at selected TCCON sites.


WFMD xCH4 smooth 100km 2hr against FTIR.CH4
EUREKA (lat.=80.1°), 2018-03-31 till 2019-09-05, 15354 meas.
WFMD xCH4 smooth 100km 2hr against FTIR.CH4
NY.ALESUND (lat.=78.9°), 2018-04-17 till 2020-08-26, 1743 meas.
WFMD xCH4 smooth 100km 2hr against FTIR.CH4
THULE (lat.=76.5°), 2018-03-26 till 2020-09-18, 19946 meas.
WFMD xCH4 smooth 100km 2hr against FTIR.CH4
KIRUNA (lat.=67.8°), 2018-03-08 till 2020-09-22, 23932 meas.
WFMD xCH4 smooth 100km 2hr against FTIR.CH4
SODANKYLA (lat.=67.4°), 2018-03-12 till 2020-09-07, 25536 meas.
WFMD xCH4 smooth 100km 2hr against FTIR.CH4
HARESTUA (lat.=60.2°), 2018-03-16 till 2020-11-03, 1893 meas.
WFMD xCH4 smooth 100km 2hr against FTIR.CH4
ST.PETERSBURG (lat.=59.9°), 2018-02-20 till 2020-10-02, 13945 meas.
WFMD xCH4 smooth 100km 2hr against FTIR.CH4
BREMEN (lat.=53.1°), 2017-11-13 till 2020-11-23, 20556 meas.
WFMD xCH4 smooth 100km 2hr against FTIR.CH4
GARMISCH (lat.=47.5°), 2018-02-14 till 2020-11-27, 9783 meas.
WFMD xCH4 smooth 100km 2hr against FTIR.CH4
ZUGSPITZE (lat.=47.4°), 2018-02-13 till 2020-12-18, 9433 meas.
xCH4 [ppb]
all GB
SAT_with_GB_prior_on_SAT
GB_measurement_smooth_with_SAT_avk_on_SAT



WFMD xCH4 smooth 100km 2hr against FTIR.CH4
JUNGFRAUJOCH (lat.=46.6°), 2017-11-16 till 2020-11-28, 7227 meas.
WFMD xCH4 smooth 100km 2hr against FTIR.CH4
TORONTO (lat.=43.6°), 2017-11-24 till 2020-12-17, 21560 meas.
WFMD xCH4 smooth 100km 2hr against FTIR.CH4
RIKUBETSU (lat.=43.5°), 2017-11-14 till 2020-12-23, 4780 meas.
WFMD xCH4 smooth 100km 2hr against FTIR.CH4
BOULDER.CO (lat.=40.0°), 2017-11-13 till 2020-12-31, 32568 meas.
WFMD xCH4 smooth 100km 2hr against FTIR.CH4
IZANA (lat.=28.3°), 2017-11-10 till 2020-12-21, 5476 meas.
WFMD xCH4 smooth 100km 2hr against FTIR.CH4
MAUNA.LOA.HI (lat.=19.5°), 2017-11-15 till 2019-01-26, 516 meas.
WFMD xCH4 smooth 100km 2hr against FTIR.CH4
ALTZOMONI (lat.=19.1°), 2017-11-23 till 2020-08-18, 17113 meas.
WFMD xCH4 smooth 100km 2hr against FTIR.CH4
PARAMARIBO (lat.=5.8°), 2017-12-16 till 2018-12-01, 3 meas.
WFMD xCH4 smooth 100km 2hr against FTIR.CH4
PORTO.VELHO (lat.=-8.8°), 2019-06-17 till 2019-09-09, 1959 meas.
WFMD xCH4 smooth 100km 2hr against FTIR.CH4
LA.REUNION.MAIDO (lat.=-21.1°), 2018-04-28 till 2019-12-22, 2690 meas.

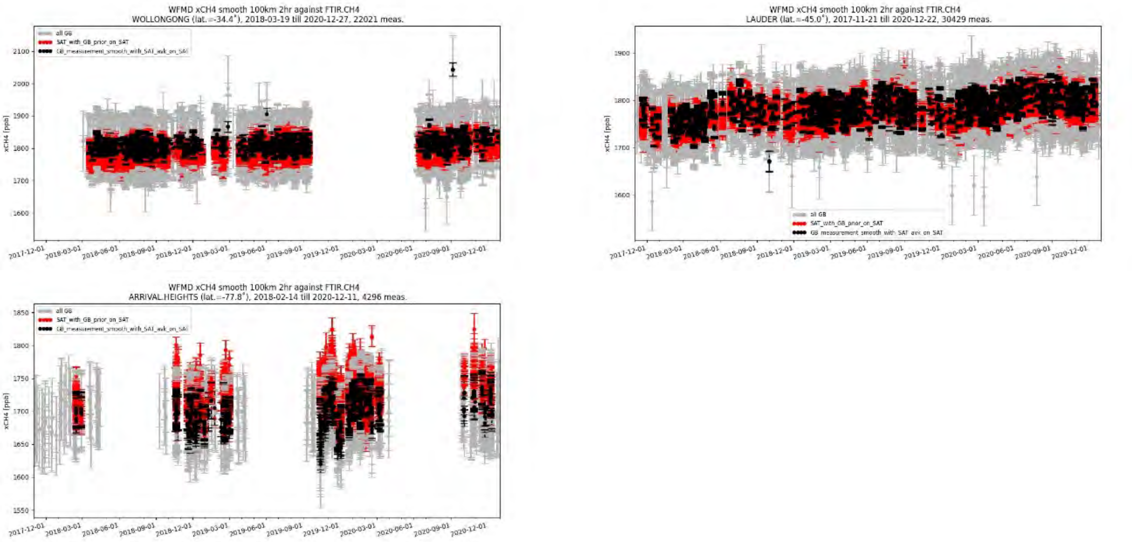


Figure 4-21: Timeseries of XCH_4 NDACC (collocated=black, all=grey) and CH4_S5P_WFMD (red) data at selected TCCON sites.

Figure 4-22 shows monthly median timeseries for TCCON and WFMD XCH_4 for all data that fall within certain latitude bands, namely all sites North of 40°N latitude (top), all sites between 40°N and the equator (mid) and all sites in the Southern hemisphere (bottom). The figures clearly show that WFMD is capable of capturing the larger scale temporal evolution of XCH_4 as well as seasonal variability.

Figure 4-23 shows the same for NDACC. Here we see good agreement for the North of 40° latitude band and the southern hemisphere, but a clear mismatch for the 0-40°N latitude band. Looking at the stations who provide input for this band we have Izaña, Mauna Loa, Altzomoni and Paramaribo (Boulder sits at 40.04°N and contributes to the >40° N latitude band). Of these 4, 3 are high altitude stations and Paramaribo has little to no overlap (3 data pairs). Most data are coming from Altzomoni (17113 data pairs) with smaller contributions from Izaña (5476 data pairs) and Mauna Loa (516 data pairs). As already discussed, Altzomoni is a particularly challenging site. It would therefore be wrong to see the middle figure as an indication of a problem with the satellite product.

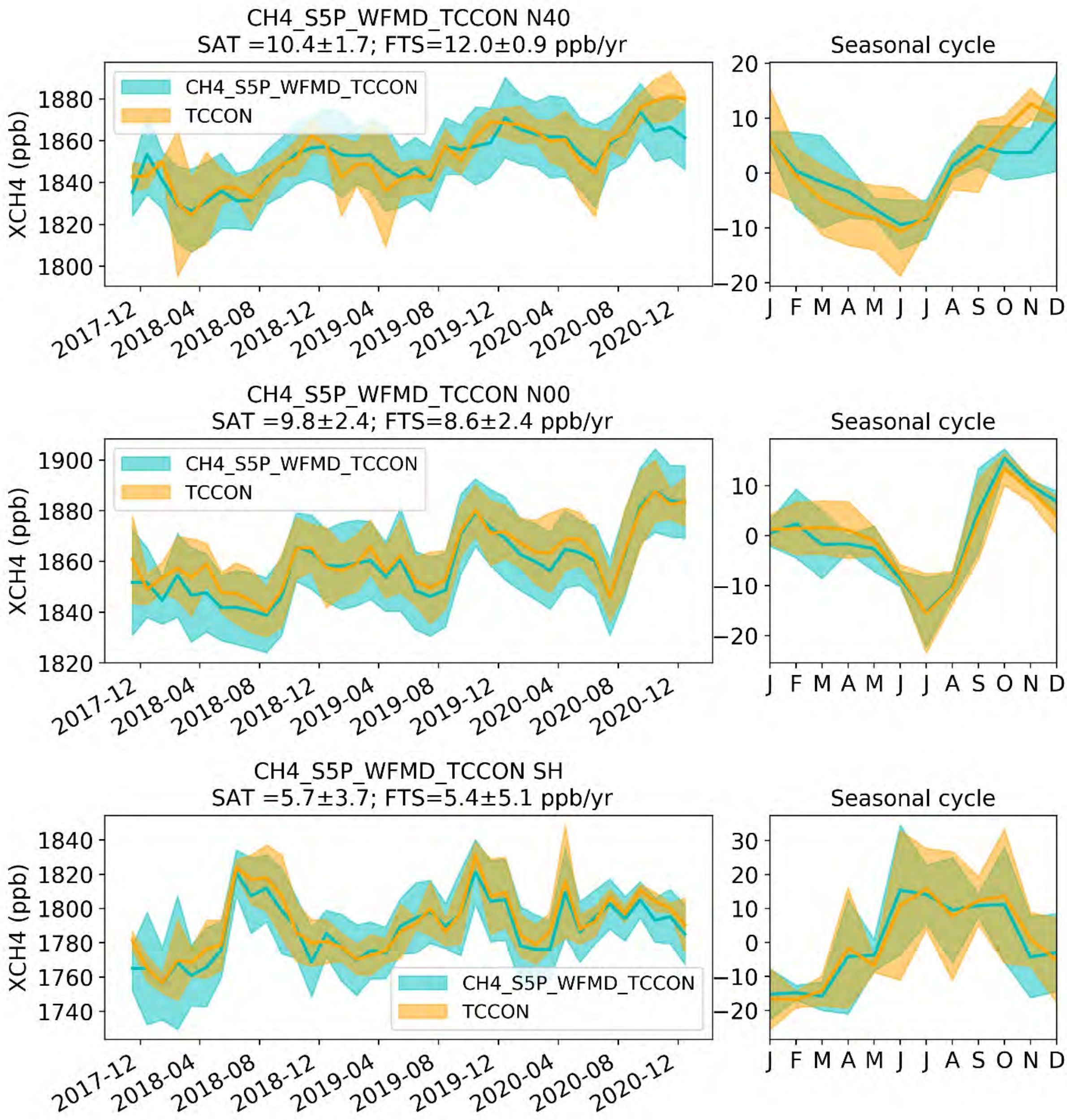


Figure 4-22: Monthly median collocated Sat and TCCON XCH_4 concentrations as a function of time and the detrended monthly medians as a function of season. The shaded areas correspond with the scaled median absolute deviation.

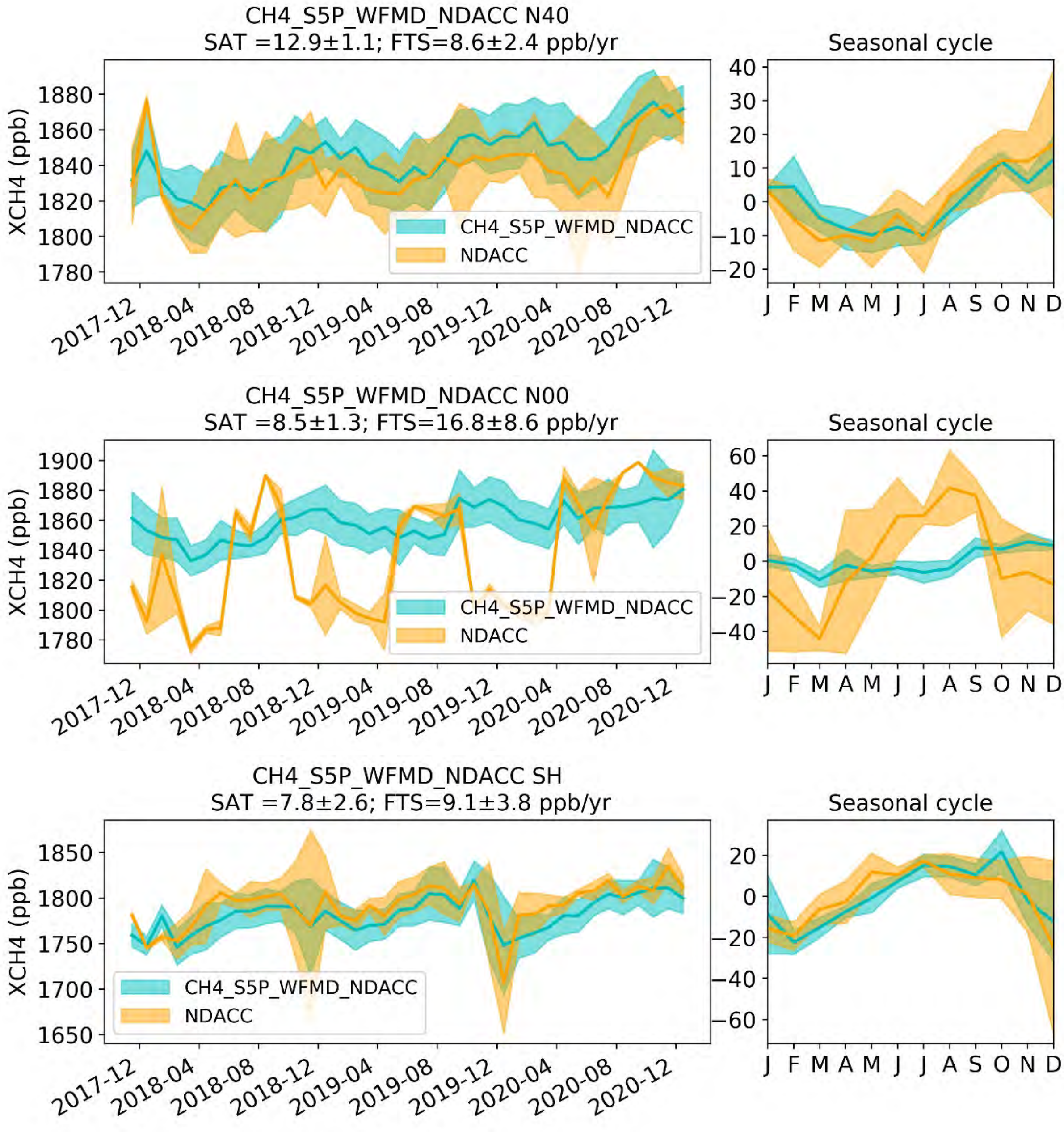


Figure 4-23: Monthly median collocated Sat and NDACC XCH_4 concentrations as a function of time and the detrended monthly medians as a function of season. The shaded areas correspond with the scaled median absolute deviation.

4.2.4.2 Summary

Unlike with our previous assessment the current CH4_S5P_WFMD data contains little noticeable outliers. The seasonal cycles and long term trends seem well captured. The obtained Stability equals 0.8 ppb/year with confidence bands that do not overlap 0. There could therefore be a significant but small trend in the retrieval (still far below the linear drift requirement of <3 ppb/year). The single measurement precision equals 13.8 (again better than the previous 15.1 ppb), thus reaching the breakthrough < 17 ppb target value. The reported uncertainty also improved and now sits at 0.86 times what we find in our analysis. The overall bias sits at 0.

The Relative and Seasonal relative accuracies equal 5.0 and 5.2 ppb respectively, thus reaching the <10 ppb target.

For NDACC we obtain a single measurement precision 18.6 [16.5,19.3] ppb, an overall bias of 8.3 [-1.3,21.7] ppb and relative accuracy values: RA 21.6 [15.4, 32.6] and SRA 18.6 [13.2, 24.7]. The confidence bands for NDACC are significantly wider indicating larger inter-station differences. This naturally also manifests itself in the relative accuracy numbers, where neither RA or SRA overlaps with the <10 ppb target. It is however safe to say that inter-station biases between the NDACC stations contribute to this number.

Table 4-11 presents an overview of the estimated data quality of CH4_S5P_WFMD, as obtained by the VALT team, from comparisons with TCCON ground-based reference observations. Values in square brackets [ ] correspond with the upper and lower 95% confidence bound on the parameter. The uncertainty ratio features 2 numbers as outlined in the validation method.

Product Quality Summary Table for Product: CH4_S5P_WFMD Level: 2, Version: v1.5, Time period covered: 11.2017 – 12.2020 Assessment: Validation Team (VALT)			
Parameter [unit]	**Achieved performance**	**Requirement**	**Comments**
Single measurement precision (1-sigma) in [ppb]	13.8 [12.1,15.0]	< 34 (T) < 17 (B) < 9 (G)	Computed as the median over all station scaled median absolute differences to TCCON
Uncertainty ratio [-]: Ratio reported uncertainty to standard deviation of satellite-TCCON difference	0.86, 0.86*	-	No requirement but value close to unity expected for a high quality data product with reliable reported uncertainty.
Mean bias (global offset) [ppb]	0.0 [-1.6, 2.6]	-	No requirement but value close to zero expected for a high quality data product.
Accuracy: Relative systematic error [ppb]	Spatial: 5.0 [1.8,7.3] Spatio-temporal: 5.2 [3.6,6.5]	< 10	Spatial: Computed as standard deviation of the biases at the various TCCON sites. Spatio-temporal: As "Spatial" but also considering seasonal biases.
Stability: Drift [ppb/year]	0.8 [0.4,2.5]	< 3	Linear drift

4.2.5 Validation results for product CH4_GO2_SRFP

Below we show the validation results of the XCH_4 concentrations as derived by the CH4_GO2_SRFP v2.0.0 algorithm using GOSAT-2 spectra, FP standing for the Full Physics version of the algorithm developed at SRON. Data was available from February 2019 up to and including August 2020. The SRFP algorithm provides *a priori* and column averaging kernel information on a 12 layer profile. Given the very limited time we cannot make useful statements about long term stability as it does not cover 2 full years of data.

4.2.5.1 Detailed results

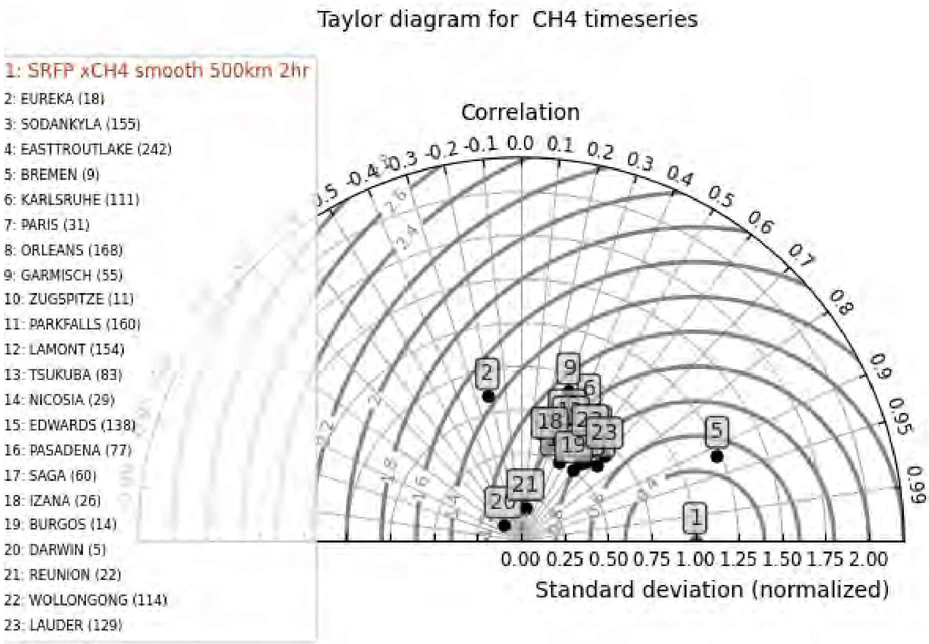


Figure 4-24: Tayor plot of XCH_4 TCCON values relative to CH4_GO2_SRFP. Straight lines correspond with the correlation, light grey lines yield the variability of the TCCON data relative to the satellite variability and the dark grey lines correspond with the variability of the Satellite -TCCON bias relative to the satellite variability.

The Taylor diagram above in **Figure 4-24** yields a concise overview of the capabilities of the CH4_GO2_SRFP algorithm with respect to the TCCON network. Most TCCON sites are nicely clustered apart from Bremen (very high correlation), Eureka, Darwin and Reunion. The last three exhibit a limited seasonal cycle either through data availability (missing autumn winter data for high latitude sites) or naturally (Southern hemisphere). All other sites cluster around the 0.5 correlation line. TCCON yields standard deviations that are 0.5 to 0.75 times that of the algorithm and the relative standard deviation of the bias sits around unity.

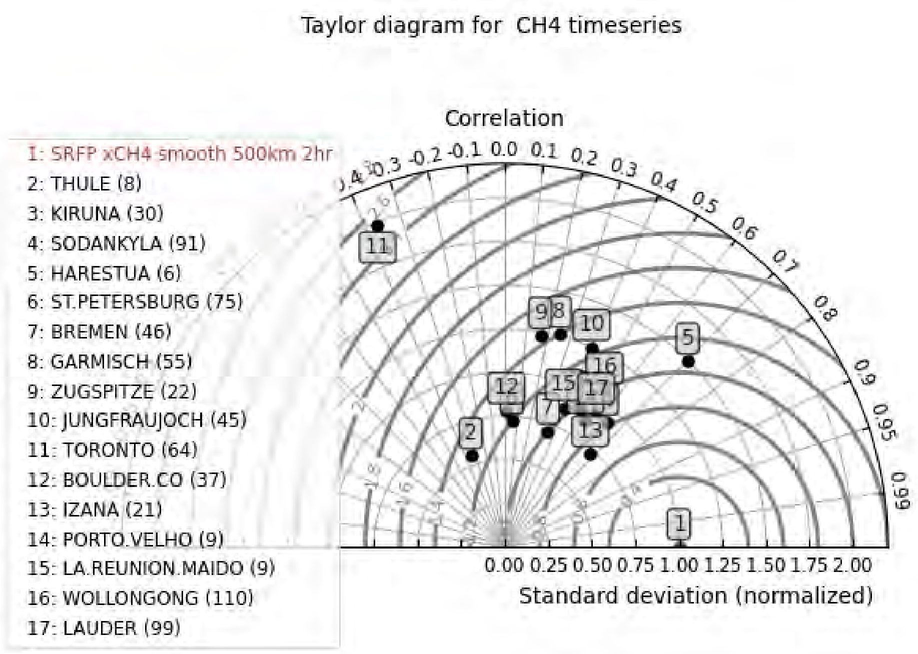


Figure 4-25: Tayor plot of XCH_4 NDACC values relative to CH4_GO2_SRFP. Straight lines correspond with the correlation, light grey lines yield the variability of the TCCON data relative to the satellite variability and the dark grey lines correspond with the variability of the Satellite -TCCON bias relative to the satellite variability.

For NDACC (**Figure 4-25**) we again see much more dispersion with strong outliers at Toronto and Harestua, but also Thule, Boulder and St. Petersburg. Correlations are generally weaker compared to TCCON, whereas its variability relative to the FTIR measurements is lower (indicating higher variability in NDACC)

Again, it is hard to discern a pattern in the mosaic plots which shows the mean bi-weekly bias between the satellite and FTS measurement pairs (**Figure 4-26 and 4-27**), particularly for NDACC which shows substantial data gaps across all latitudes. One of the few stations for which we have a near complete coverage, namely Toronto, again (as with WFMD XCH_4) shows a shift in the bias over time. For TCCON. no station clearly stands out.

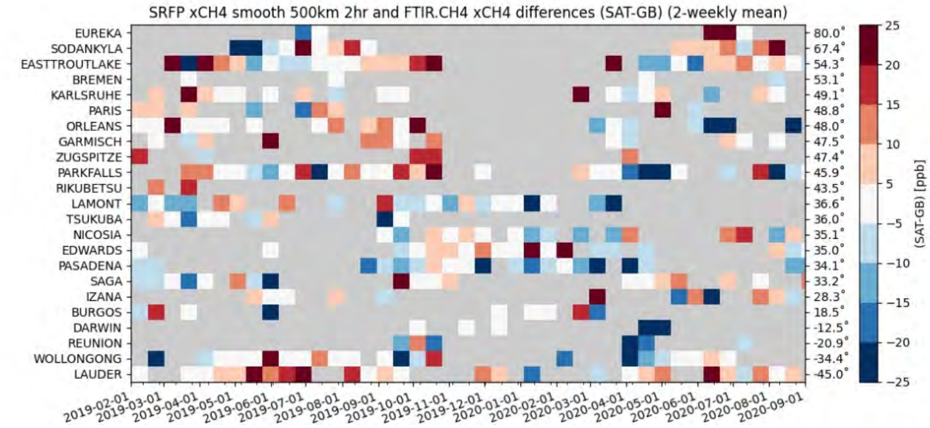


Figure 4-26. Mosaic plot of bi-weekly mean CH4_GO2_SRFP – TCCON XCH_4 biases as a function of time and TCCON station.

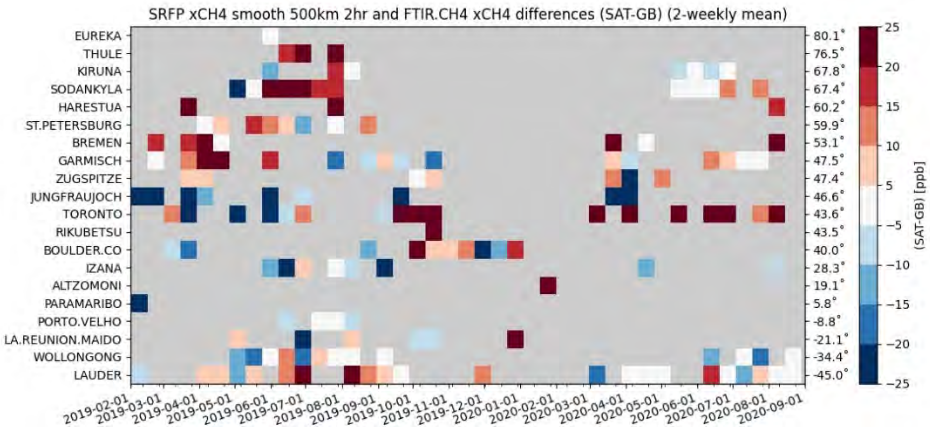


Figure 4-27. Mosaic plot of bi-weekly mean CH4_GO2_SRFP – NDACC XCH_4 biases as a function of time and NDACC station.

Table 4-12 lists all bias and scatter results derived from individual data pairs at all TCCON stations. As with its XCO_2 counterpart, the algorithm produces on average ~80 data pairs per station, which corresponds with ~50 pairs per station per year. Several stations however have far less collocated measurements (Bremen, Rikubetsu and Darwin have less than 10 data pairs) hampering an accurate assessment of the data quality at these sites. The observed median bias ranges between -13.7 ppb (Reunion) and 21.8 ppb (Zugspitze), while the scatter ranges between 1.3 ppb (Rikubetsu and 22.7 ppb (Paris). Ignoring stations with <10 collocated datapoints the lowest scatter is 10.7 ppb (Wollongong). Due to the limited dataset we did not determine long term bias drift numbers. The overall correlation using all collocated data regardless of station equals 0.84.

Table 4-12: Number of collocated data pairs (N), Correlation (R), Bias, Scatter, long term trend difference (ltt) and uncertainty thereon (ltt_err), seasonal amplitude difference (A) and uncertainty thereon (A_err) as well as the latitude of the TCCON station. The last row lists the median values over all stations. Product: CH4_GO2_SRFP.

STATION	N	R	Bias	Scat	ltt	ltt_err	A	A_err	lat
EUREKA	18	-0.23	-6.67	14.47	-	-	-	-	80.0
SODANKYLA	155	0.60	7.03	15.96	-	-	-	-	67.4
EASTTROUTLAKE	242	0.58	1.01	20.64	-	-	-	-	54.3
BREMEN	9	0.92	1.99	4.07	-	-	-	-	53.1
KARLSRUHE	111	0.47	0.18	11.66	-	-	-	-	49.1
PARIS	31	0.45	6.96	22.73	-	-	-	-	48.8
ORLEANS	168	0.28	1.63	15.30	-	-	-	-	48.0
GARMISCH	55	0.30	6.79	14.67	-	-	-	-	47.5
ZUGSPITZE	11	0.38	21.78	18.04	-	-	-	-	47.4
PARKFALLS	160	0.51	-1.31	15.88	-	-	-	-	45.9
RIKUBETSU	3	1.00	14.26	1.34	-	-	-	-	43.5
LAMONT	154	0.43	-6.41	14.49	-	-	-	-	36.6
TSUKUBA	83	0.60	-1.18	13.57	-	-	-	-	36.0
NICOSIA	29	0.70	5.02	12.14	-	-	-	-	35.1
EDWARDS	138	0.44	0.90	15.42	-	-	-	-	35.0
PASADENA	77	0.60	-8.34	17.33	-	-	-	-	34.1
SAGA	60	0.63	3.35	11.88	-	-	-	-	33.2
IZANA	26	0.29	2.82	12.77	-	-	-	-	28.3
BURGOS	14	0.59	-1.26	15.79	-	-	-	-	18.5
DARWIN	5	-0.75	-3.52	7.83	-	-	-	-	-12.5
REUNION	22	0.14	-13.74	17.64	-	-	-	-	-20.9
WOLLONGONG	114	0.58	-2.78	10.72	-	-	-	-	-34.4
LAUDER	129	0.70	7.32	14.73	-	-	-	-	-45.0

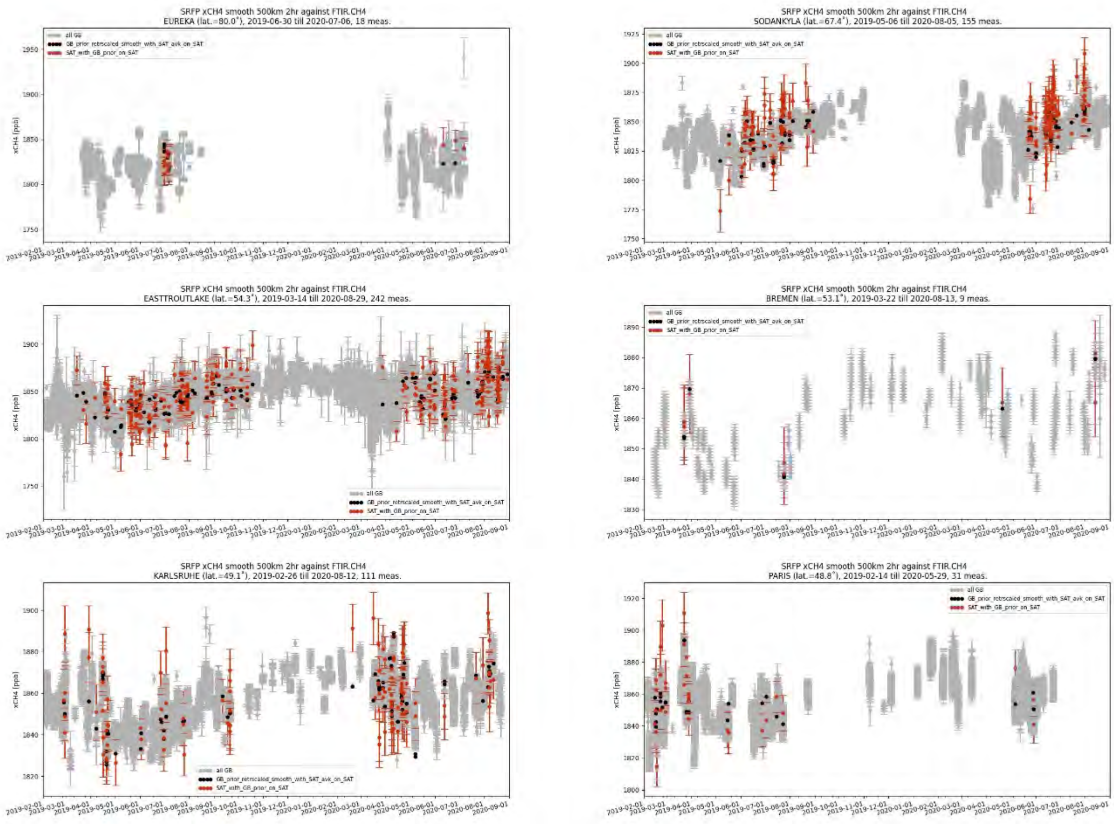
MEDIAN	**60**	**0.51**	**1.01**	**14.67**	**-**	**-**	**-**	**-**	**36.6**

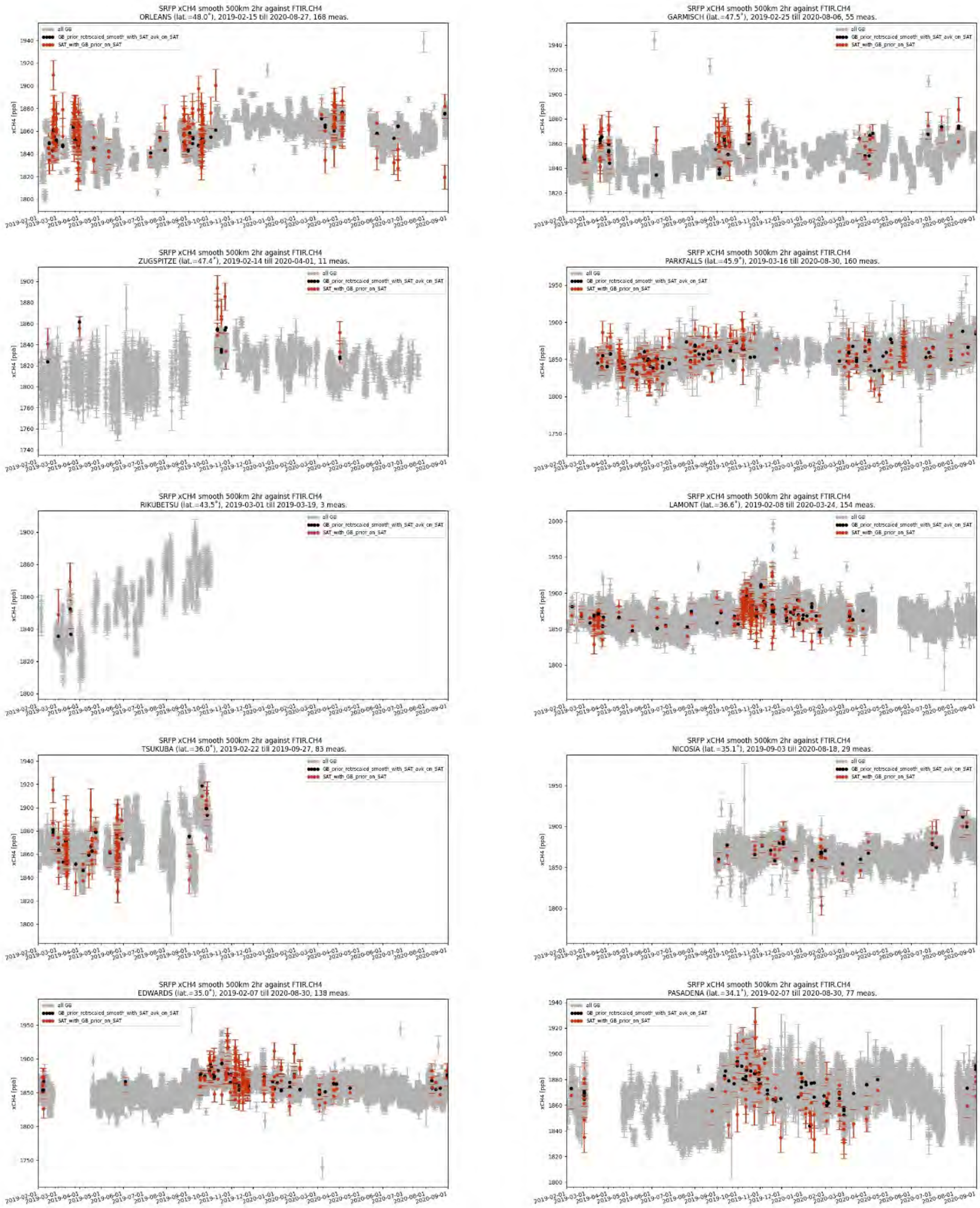
Table 4-13: Number of collocated data pairs (N), Correlation (R), Bias, Scatter, long term trend difference (ltt) and uncertainty thereon (ltt_err), seasonal amplitude difference (A) and uncertainty thereon (A_err) as well as the latitude of the NDACC station. The last row lists the median values over all stations. Product: CH4_GO2_SRFP.

STATION	N	R	Bias	Scat	ltt	ltt_err	A	A_err	lat
EUREKA	1	-	3.91	0.00	-	-	-	-	80.1
THULE	8	-0.35	22.35	21.77	-	-	-	-	76.5
KIRUNA	30	0.48	-6.85	13.85	-	-	-	-	67.8
SODANKYLA	91	0.63	12.22	14.90	-	-	-	-	67.4
HARESTUA	6	0.70	30.39	12.27	-	-	-	-	60.2
ST.PETERSBURG	75	0.06	11.71	16.03	-	-	-	-	59.9
BREMEN	46	0.35	14.16	17.10	-	-	-	-	53.1
GARMISCH	55	0.25	5.30	14.22	-	-	-	-	47.5
ZUGSPITZE	22	0.17	1.69	22.39	-	-	-	-	47.4
JUNGFRAUJOCH	45	0.40	-32.40	20.54	-	-	-	-	46.6
TORONTO	64	-0.37	14.06	52.79	-	-	-	-	43.6
RIKUBETSU	1	-	29.52	0.00	-	-	-	-	43.5
BOULDER.CO	37	0.01	9.72	18.24	-	-	-	-	40.0
IZANA	21	0.68	-10.86	9.73	-	-	-	-	28.3
ALTZOMONI	4	-0.21	64.02	4.02	-	-	-	-	19.1
PARAMARIBO	1	-	-39.12	0.00	-	-	-	-	5.8
PORTO.VELHO	9	0.54	-2.99	15.63	-	-	-	-	-8.8
LA.REUNION.MAIDO	9	0.39	1.32	15.53	-	-	-	-	-21.1
WOLLONGONG	110	0.54	-5.75	16.57	-	-	-	-	-34.4
LAUDER	99	0.56	6.47	15.22	-	-	-	-	-45.0
MEDIAN	**26**	**0.39**	**5.89**	**15.38**	**-**	**-**	**-**	**-**	**45.1**

Table 4-13 lists the same variables but now for the NDACC stations. Here again we sometimes have very little overlap between the ground-based and satellite measurements. Eureka, Rikubetsu and Paramaribo only feature one single collocated measurement. Thule, Harestua, Altzomoni, Porto Velho and Reunion (Maïdo) all feature less than 10 datapoints. Ignoring these low data volume stations, the correlation coefficient ranges between -0.37 (Toronto) and 0.68 (Izaña). The bias ranges between -32.4 ppb (Jungfraujoch) and 14.2 ppb (Bremen) while the scatter ranges between 9.7 ppb (Izaña) and 22.4 ppb (Zugspitze). Note that the scatter at Toronto is 52.8 ppb, but (as can be seen in **Figure 4-29** this is mainly due to the large amount of scatter present in the ground-based FTIR data at this location.

The timeseries below in **Figure 4-28** show individual satellite and ground-based TCCON measurements, while **Figure 4-29** does the same for NDACC. For TCCON we see that SRFP generally manages to capture the seasonal cycle. While the scatter is somewhat higher for SRFP XCH_4, compared to TCCON, it is relatively free of outliers (although some XCH_4 values are clearly suspect such as the low values observed at Wollongong and the high 'outliers' at Karlsruhe). That said, it is clear that the algorithm manages to capture the natural variability of XCH_4.





SRFP xCH4 smooth 500km 2hr against FTIR.CH4
ORLEANS (lat.=48.0°), 2019-02-15 till 2020-08-27, 168 meas.
SRFP xCH4 smooth 500km 2hr against FTIR.CH4
GARMISCH (lat.=47.5°), 2019-02-25 till 2020-08-06, 55 meas.
SRFP xCH4 smooth 500km 2hr against FTIR.CH4
ZUGSPITZE (lat.=47.4°), 2019-02-14 till 2020-04-01, 11 meas.
SRFP xCH4 smooth 500km 2hr against FTIR.CH4
PARKFALLS (lat.=45.9°), 2019-03-16 till 2020-08-30, 160 meas.
SRFP xCH4 smooth 500km 2hr against FTIR.CH4
RIKUBETSU (lat.=43.5°), 2019-03-01 till 2019-03-19, 3 meas.
SRFP xCH4 smooth 500km 2hr against FTIR.CH4
LAMONT (lat.=36.6°), 2019-02-08 till 2020-03-24, 154 meas.
SRFP xCH4 smooth 500km 2hr against FTIR.CH4
TSUKUBA (lat.=36.0°), 2019-02-22 till 2019-09-27, 83 meas.
SRFP xCH4 smooth 500km 2hr against FTIR.CH4
NICOSIA (lat.=35.1°), 2019-09-03 till 2020-08-18, 29 meas.
SRFP xCH4 smooth 500km 2hr against FTIR.CH4
EDWARDS (lat.=35.0°), 2019-02-07 till 2020-08-30, 138 meas.
SRFP xCH4 smooth 500km 2hr against FTIR.CH4
PASADENA (lat.=34.1°), 2019-02-07 till 2020-08-30, 77 meas.
all GB
GB_prior_rctrscaled_smooth_with_SAT_avk_on_SAT
SAT_with_GB_prior_on_SAT
xCH4 [ppb]

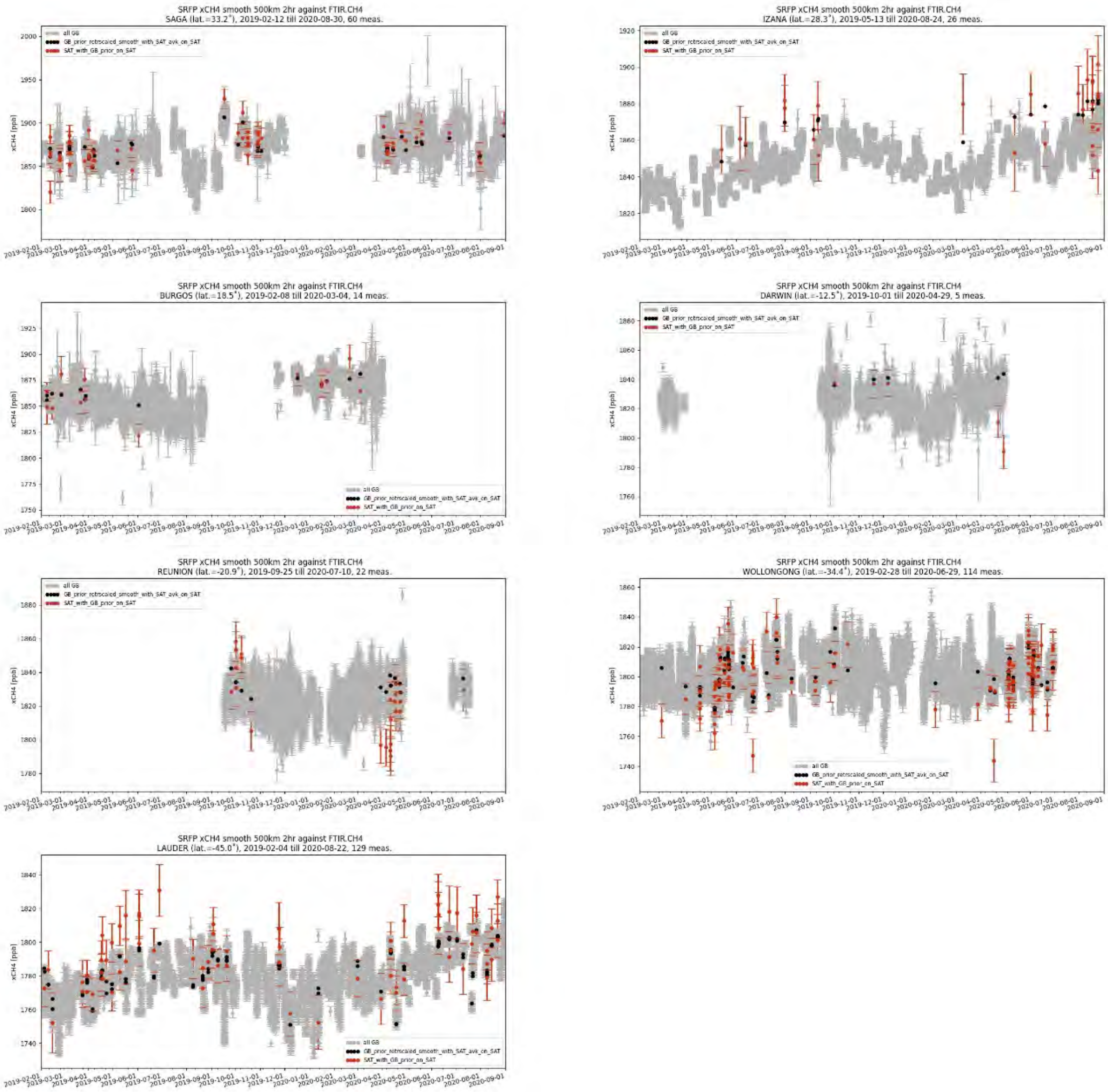
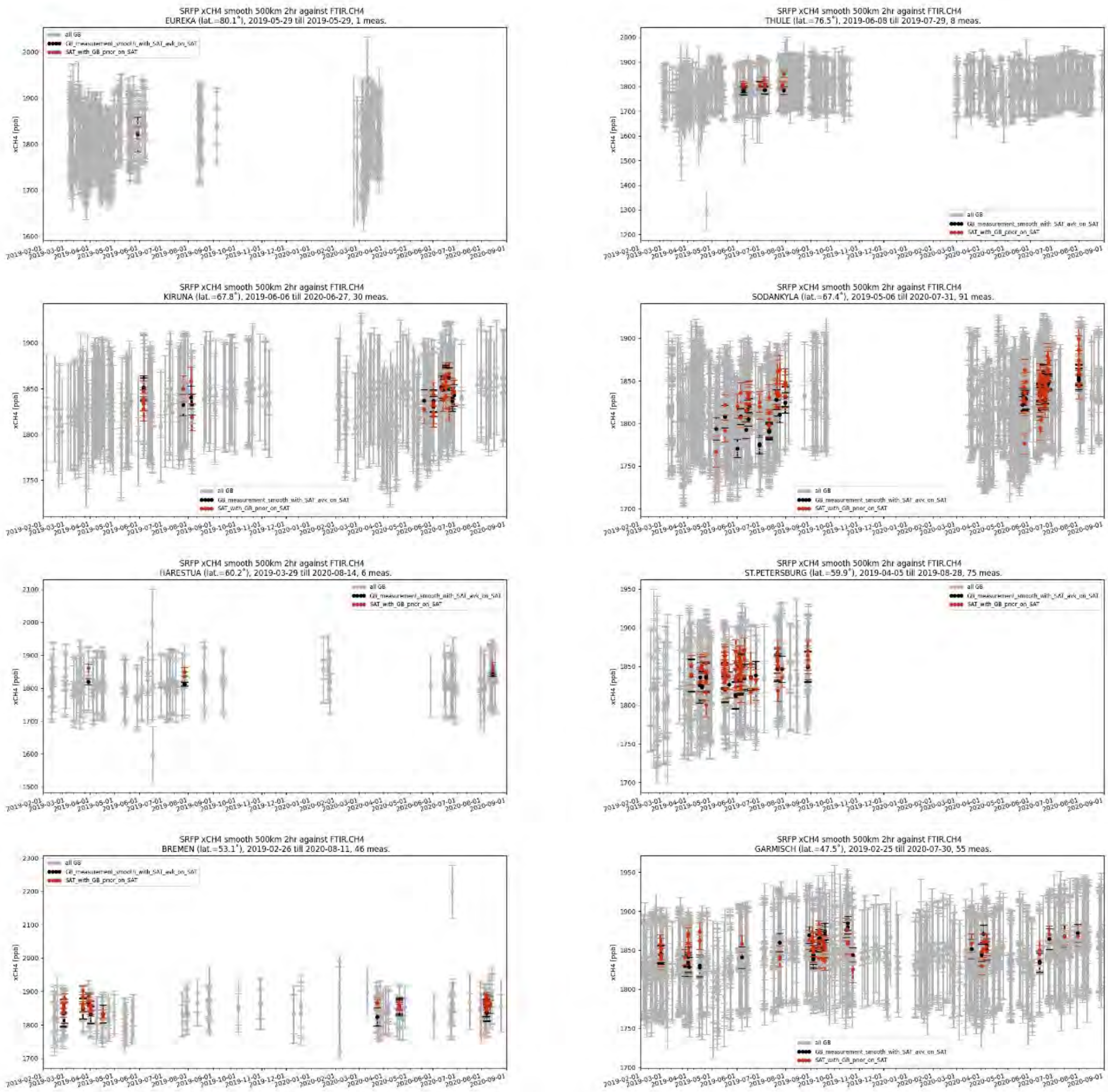
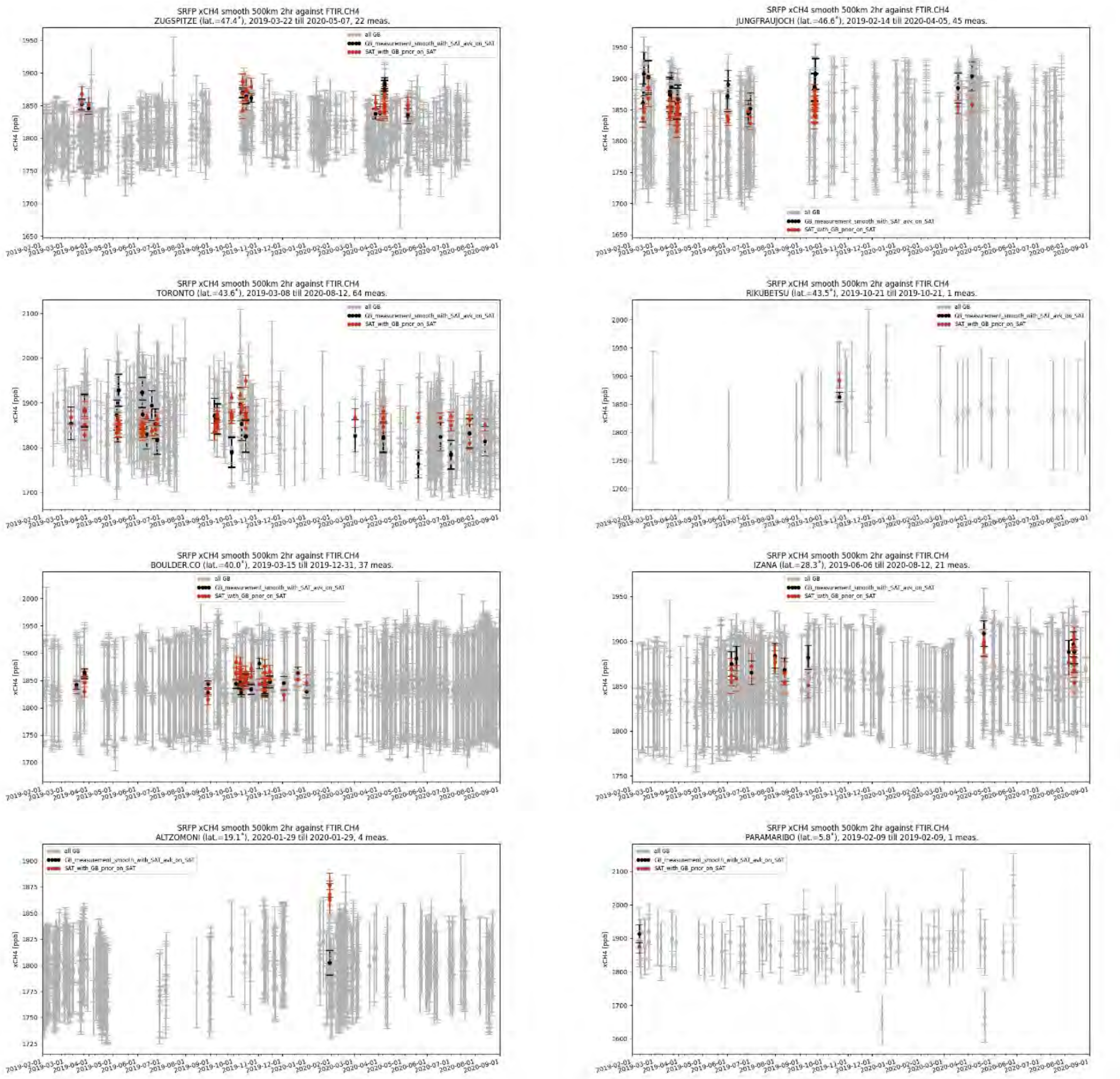


Figure 4-28: XCH_4 timeseries at all TCCON sites (red= CH4_GO2_SRFP data, black is collocated TCCON data and grey are the uncollocated TCCON data).

For NDACC it is clear that SRFP exhibits the same or at some stations even smaller temporal variability than NDACC. Also crealy visible is the sparsness of the dataset, with either little coverage at all, or significant datagaps in the timeseries. For stations where we do have consistent longer sampling, such as Garmisch and Lauder, we see that NDACC and SRFP are in good agreement. For Toronto we clearly see the high variability in the NDACC data, it is therefore not clear whether the strong trend that is observed at this station is real or a sampling issue.




SRFP xCH4 smooth 500km 2hr against FTIR.CH4
ZUGSPITZE (lat.=47.4°), 2019-03-22 till 2020-05-07, 22 meas.
SRFP xCH4 smooth 500km 2hr against FTIR.CH4
JUNGFRAUJOCH (lat.=46.6°), 2019-02-14 till 2020-04-05, 45 meas.
SRFP xCH4 smooth 500km 2hr against FTIR.CH4
TORONTO (lat.=43.6°), 2019-03-08 till 2020-08-12, 64 meas.
SRFP xCH4 smooth 500km 2hr against FTIR.CH4
RIKUBETSU (lat.=43.5°), 2019-10-21 till 2019-10-21, 1 meas.
SRFP xCH4 smooth 500km 2hr against FTIR.CH4
BOULDER.CO (lat.=40.0°), 2019-03-15 till 2019-12-31, 37 meas.
SRFP xCH4 smooth 500km 2hr against FTIR.CH4
IZANA (lat.=28.3°), 2019-06-06 till 2020-08-12, 21 meas.
SRFP xCH4 smooth 500km 2hr against FTIR.CH4
ALTZOMONI (lat.=19.1°), 2020-01-29 till 2020-01-29, 4 meas.
SRFP xCH4 smooth 500km 2hr against FTIR.CH4
PARAMARIBO (lat.=5.8°), 2019-02-09 till 2019-02-09, 1 meas.
all GB
GB_measurement_smooth_with_SAT_avk_on_SAT
SAT_with_GB_prior_on_SAT

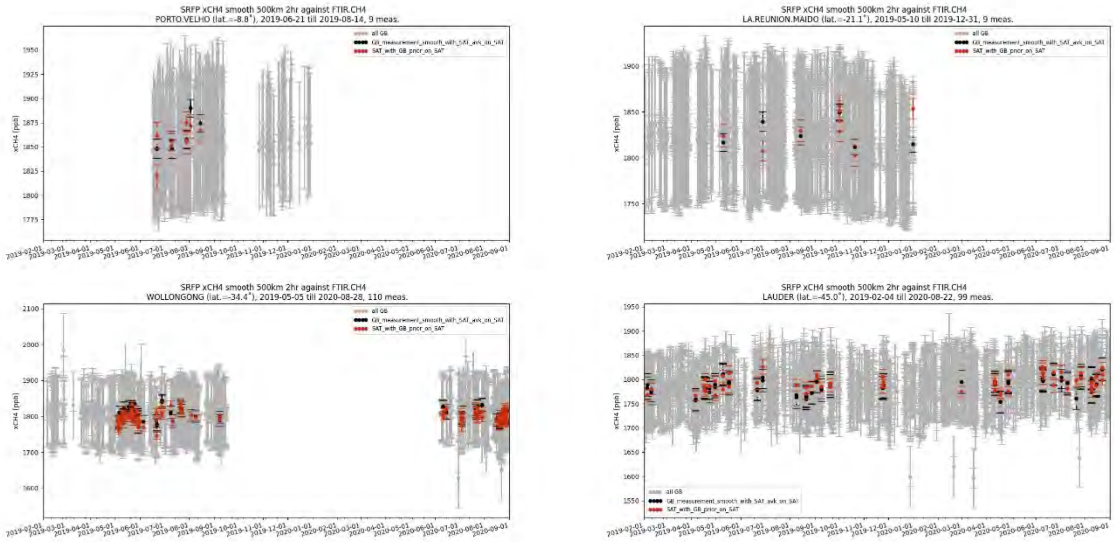


Figure 4-29: Timeseries of XCH_4 NDACC (collocated=black, all=grey) and CH4_GO2_SRFP (red) data at all NDACC sites.

Figure 4-30 shows monthly median timeseries for TCCON and SRFP XCH_4 for all data that fall within certain latitude bands, namely all sites North of 40°N latitude (top), all sites between 40°N and the equator (mid) and all sites in the Southern hemisphere (bottom). The plots also show the trend results of a trend+seasonality fit. The Northen hemisphere plots are very consistent with good overlap between the SRFP and TCCON long term trend and seasonal cycles. For the Southern hemisphere there is a larger difference in the fitted long term trend, but still the errors overlap, nor is there a strong discrepancy between the actual trendlines.

Figure 4-31 shows the same but for NDACC. As already mentioned in our discussion of the WFMD-NDACC results, the 0-40°N latitude band is particularly problematic. Here even to the point where not enough data is available to generate a meaningful seasonal cycle. Ignoring this band we see reasonable agreement (overlapping error range) but with a lot more variability as compared to TCCON (in part, no doubt, due to the lower data density)

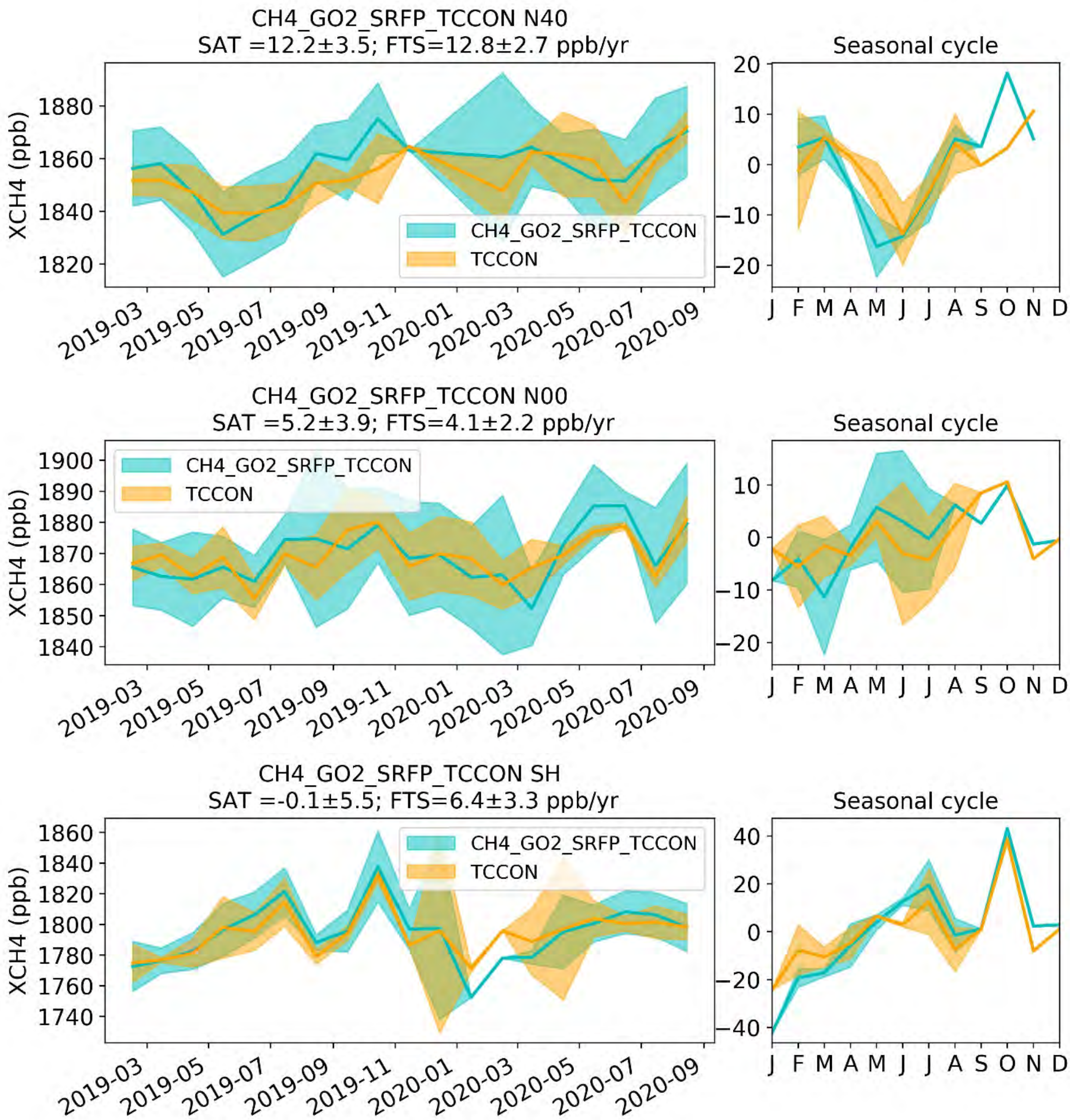


Figure 4-30: Monthly median collocated Sat and TCCON XCH_4 concentrations as a function of time. The shaded areas correspond with the scaled median absolute deviation.

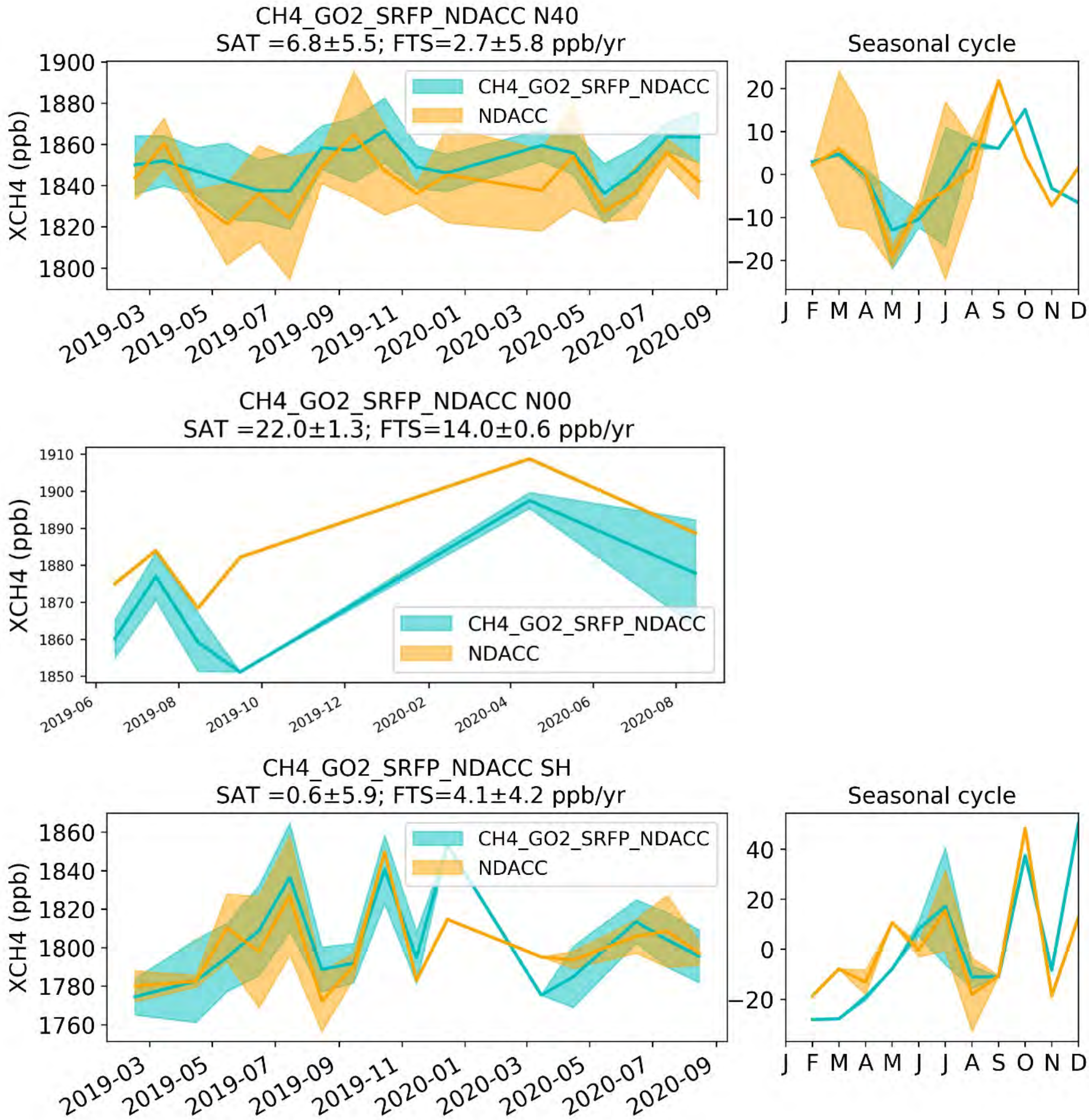


Figure 4-31: Monthly median collocated Sat and NDACC XCH_4 concentrations as a function of time. The shaded areas correspond with the scaled median absolute deviation.

4.2.5.2 Summary

Listed in the table below (**Table 4-14**) are the Figure of Merit parameters as derived from the individual collocated data pairs at each station.

SRFP XCH_4's single measurement precision equals 14.7 ppb, reaching the Breakthrough target of <17 ppb. The error assessment is slightly underestimated with an uncertainty ratio of 0.77. The median bias equals 1.0 ppb and is not significant with confidence bands between -1.4 and 3.2 ppb.

Both the spatial and spatio-temporal relative accuracies reach the <10 ppb target and no meaningful estimate for the drift can be established nor do we see any obvious problems in this regard.

For NDACC, we obtain a single measurement precision of 15.8 [13.3, 16.8] ppb, a positive but not significant median bias of 5.9 [-0.4, 14.8] pp. The median relative accuracy numbers do not meet the target but exhibit very large uncertainty bands (RA 12.2 [3.9, 19.4] ppb, SRA 11;7 [2.2, 16.1] ppb). Given these uncertainties, all obtained data overlap with our TCCON analysis.

Table 4-14 presents an overview of the estimated data quality of CH4_GO2_SRFP, as obtained by the VALT team, from comparisons with TCCON ground-based reference observations. Values in square brackets [ ] correspond with the upper and lower 95% confidence bound on the parameter. The uncertainty ratio features 2 numbers as outlined in the validation method.

Product Quality Summary Table for Product: CH4_GO2_SRFP Level: 2, Version: v02.0.0, Time period covered: 2.2019 – 08.2020 Assessment: Validation Team (VALT)			
Parameter [unit]	**Achieved performance**	**Requirement**	**Comments**
Single measurement precision (1-sigma) in [ppm]	14.7 [13.5,16.2]	< 34 (T) < 17 (B) < 9 (G)	Computed as the median over all station scaled median absolute differences to TCCON
Uncertainty ratio [-]: Ratio reported uncertainty to standard deviation of satellite-TCCON difference	0.76, 0.77*	-	No requirement but value close to unity expected for a high quality data product with reliable reported uncertainty.
Median bias (global offset) [ppm]	1.0 [-1.4,3.2]	-	No requirement but value close to zero expected for a high quality data product.
Accuracy: Relative systematic error [ppm]	Spatial: 5.8 [2.3,8.9] Spatio-temporal: 6.9 [4.4,9.7]	< 10	Spatial: Computed as standard deviation of the biases at the various TCCON sites. Spatio-temporal: As "Spatial" but also considering seasonal biases.
Stability: Drift [ppm/year]	-	< 3	Linear drift

4.2.6 Validation results for product CH4_GO2_SRPR

Below we show the validation results of the XCH_4 concentrations as derived by the CH4_GO2_SRPR v2.0.0 algorithm using GOSAT-2 spectra. 'PR' stands for the proxy version of the algorithm developed at SRON, whereby the retrieved CH_4 concentration is scaled by the modelled CO_2/retrieved CO_2 ratio. Data was available from February 2019 up to and including July 2020. The SRPR algorithm provides *a priori* and column averaging kernel data on a 3 layer vertical profile.

Again, given the limited time period that is covered by this product, we can not make useful statements about long term stability as it does not cover a full 2 years of data.

4.2.6.1 Detailed results

The Taylor diagram below in **Figure 4-32** yields a concise overview of the capabilities of the CH4_GO2_SRPR algorithm. Almost all TCCON sites cluster between the 0.2 and 0.6 correlation line. The TCCON scatter is smaller than that of SRPR while the variability of the bias roughly ranges between 0.8 and 1, relative to the SRPR variability. These results are very similar to the ones obtained from its Full Physics counterpart, but with no station outliers such as Bremen and Eureka in the SRFP-TCCON taylor plot (see **Figure 4-24**).

Figure 4-33, yields the same information but for the NDACC comparisons. Again we see more dispersion as compared to TCCON. Toronto, Wollongong, Boulder and Ny Alesund stand out with much higher scatter in the NDACC data as compared to SRPR. The other stations are clustered between the 0 and 0.4 correlation line, with scatter values of the bias, being 1 to 1.2 times that of SRPR. Compared to SRFP (see **Figure 4-25**), these values seem to be internally more consistent between stations.

When looking at the mosaic plot for TCCON (**Figure 4-34**), we see almost consistent negative biases across all latitudes and times apart from the stations North of 40°N, but up till May 2020 only. With the limited available data it is hard to tell if this apparent bias shift is the result of a long term trend, seasonal mismatch (the same period in 2019 does hint at lower biases but not as outspoken) or something entirely different (and possibly transient in nature). Here the SRPR product does substantially differ from SRFP, where we see a lot less data coverage and more gaps in the timeseries, but no outspoken pattern in the biases, nor were they in general this negative.

Figure 4-44 shows the same but for NDACC. Here we see more data gaps which hampers our ability to draw conclusions. No clear negative bias surge past May 2020 is clearly detectable. Toronto yet again features an outspoken trend.

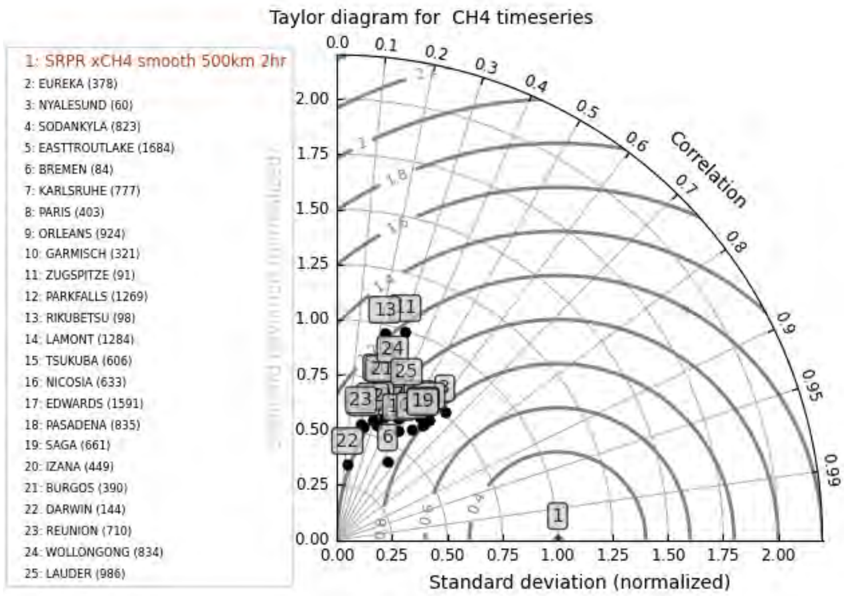


Figure 4-32: Tayor plot of XCH4 TCCON values relative to CH4_GO2_SRPR. Straight lines correspond with the correlation, light grey lines yield the variability of the TCCON data relative to the satellite variability and the dark grey lines correspond with the variability of the Satellite -TCCON bias relative to the satellite variability.

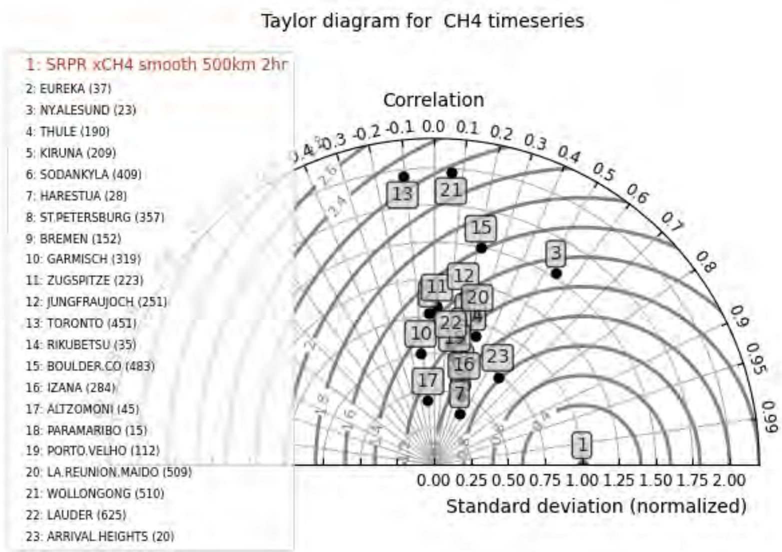


Figure 4-33: Tayor plot of XCH4 NDACC values relative to CH4_GO2_SRPR. Straight lines correspond with the correlation, light grey lines yield the variability of the NDACC data relative to the satellite

variability and the dark grey lines correspond with the variability of the Satellite –NDACC bias relative to the satellite variability.

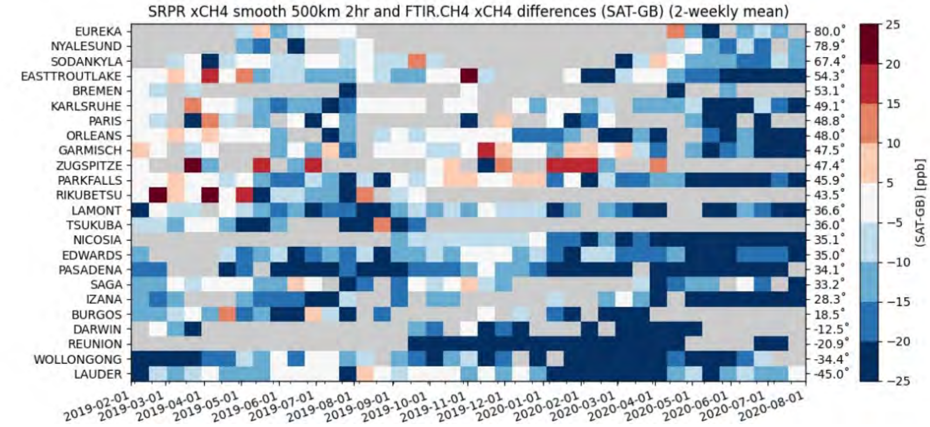


Figure 4-34. Mosaic plot of bi-weekly mean CH4_GO2_SRPR - TCCON XCH_4 biases as a function of time and TCCON station.

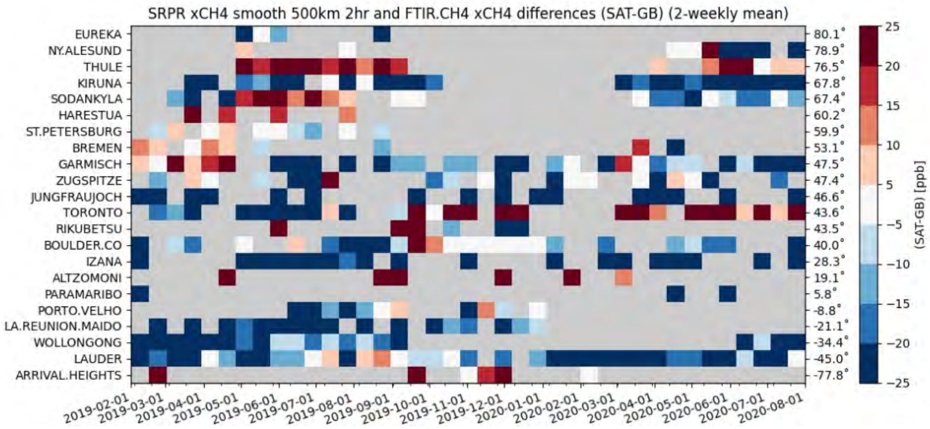


Figure 4-35. Mosaic plot of bi-weekly mean CH4_GO2_SRPR - NDACC XCH_4 biases as a function of time and TCCON station.

Table 4-15 lists all bias and scatter results derived from individual data pairs at all TCCON stations. The Proxy version of the algorithm produces roughly 10 times (note that in the previous PVIR iteration this was little more than 3 times) as many collocated data pairs than its Full Physics counterpart, with on average ~670 data pairs per station, which corresponds with ~500 pairs per station per year. The only stations that feature less than 100 collocated data pairs are Ny Alesund (60), Bremen (84), Zugspitze (91) and Rikubetsu (98). While the data density is higher, the single measurement precision is also somewhat higher (16.7 ppb for SRPR vs. 14.7 ppb for SRFP) with values ranging between 13.67 ppb (Reunion) and 22.05

ppb (Ny Alesund). This in turn impacts the median correlation values (0.51 for SRFP vs. 0.35 for SRPR). Also noticeable is the fact that SRPR features a consistent negative bias (Zugspitze being the only exception and here one could call the effect of profile extension into question. However there is also no correlative data for Zugspitze after May 2020, so this might play a role as well) ranging between -28.42 ppb (Reunion) and 11.04 ppb (Zugspitze), with a median bias of -13.2 ppb. SRFP only features a slight 1.0 ppb positive median bias compared to TCCON. The correlation using all data regardless of station yields 0.80 which is only slightly below SRFP's 0.84.

Table 4-15: Number of collocated data pairs (N), Correlation (R), Bias, Scatter, long term trend difference (ltt) and uncertainty thereon (ltt_err), seasonal amplitude difference (A) and uncertainty thereon (A_err) as well as the latitude of the TCCON station. The last row lists the median values over all stations. Product: CH4_GO2_SRPR.

STATION	N	R	Bias	Scat	ltt	ltt_err	A	A_err	lat
EUREKA	378	0.33	-5.23	17.63	-	-	-	-	80.0
NYALESUND	60	0.64	-8.95	22.05	-	-	-	-	78.9
SODANKYLA	823	0.43	-8.31	18.45	-	-	-	-	67.4
EASTTROUTLAKE	1684	0.30	-11.55	21.76	-	-	-	-	54.3
BREMEN	84	0.54	-10.01	17.91	-	-	-	-	53.1
KARLSRUHE	777	0.45	-13.07	16.40	-	-	-	-	49.1
PARIS	403	0.30	-16.48	17.07	-	-	-	-	48.8
ORLEANS	924	0.36	-9.76	17.91	-	-	-	-	48.0
GARMISCH	321	0.23	-8.71	19.01	-	-	-	-	47.5
ZUGSPITZE	91	0.31	11.04	21.53	-	-	-	-	47.4
PARKFALLS	1269	0.29	-13.96	19.66	-	-	-	-	45.9
RIKUBETSU	98	0.22	-4.70	14.27	-	-	-	-	43.5
LAMONT	1284	0.56	-13.28	15.65	-	-	-	-	36.6
TSUKUBA	606	0.61	-16.34	18.65	-	-	-	-	36.0
NICOSIA	633	0.49	-14.42	15.00	-	-	-	-	35.1
EDWARDS	1591	0.56	-13.43	15.43	-	-	-	-	35.0
PASADENA	835	0.60	-23.67	16.96	-	-	-	-	34.1
SAGA	661	0.60	-9.95	14.68	-	-	-	-	33.2
IZANA	449	0.27	-19.88	15.02	-	-	-	-	28.3
BURGOS	390	0.29	-16.33	15.32	-	-	-	-	18.5
DARWIN	144	0.14	-27.82	15.99	-	-	-	-	-12.5
REUNION	710	0.20	-28.42	13.67	-	-	-	-	-20.9
WOLLONGONG	834	0.32	-18.77	16.23	-	-	-	-	-34.4
LAUDER	986	0.43	-11.19	15.89	-	-	-	-	-45.0
MEDIAN	**647**	**0.35**	**-13.18**	**16.68**	**-**	**-**	**-**	**-**	**40.1**

Table 4-16 shows the same but for NDACC. Consistent with previous NDACC analysis for other algorithms, we see lower data densities (Median at 216), correlation numbers (between -0.12 (Garmisch) and 0.54 (Ny Alesund) and higher scatter (between 12.67 ppb (La Reunion Maïdo) and 79.3ç ppb (Rikubetsu). The median bias equals -11.73 ppb, but with much larger interstation variability (from -62.84 ppb at Paramaribo to 48.94 ppb at Altzomoni). Note that for all three XCH_4 algorithms, Paramaribo and Altzomoni are consistently the stations which feature the highest biases.

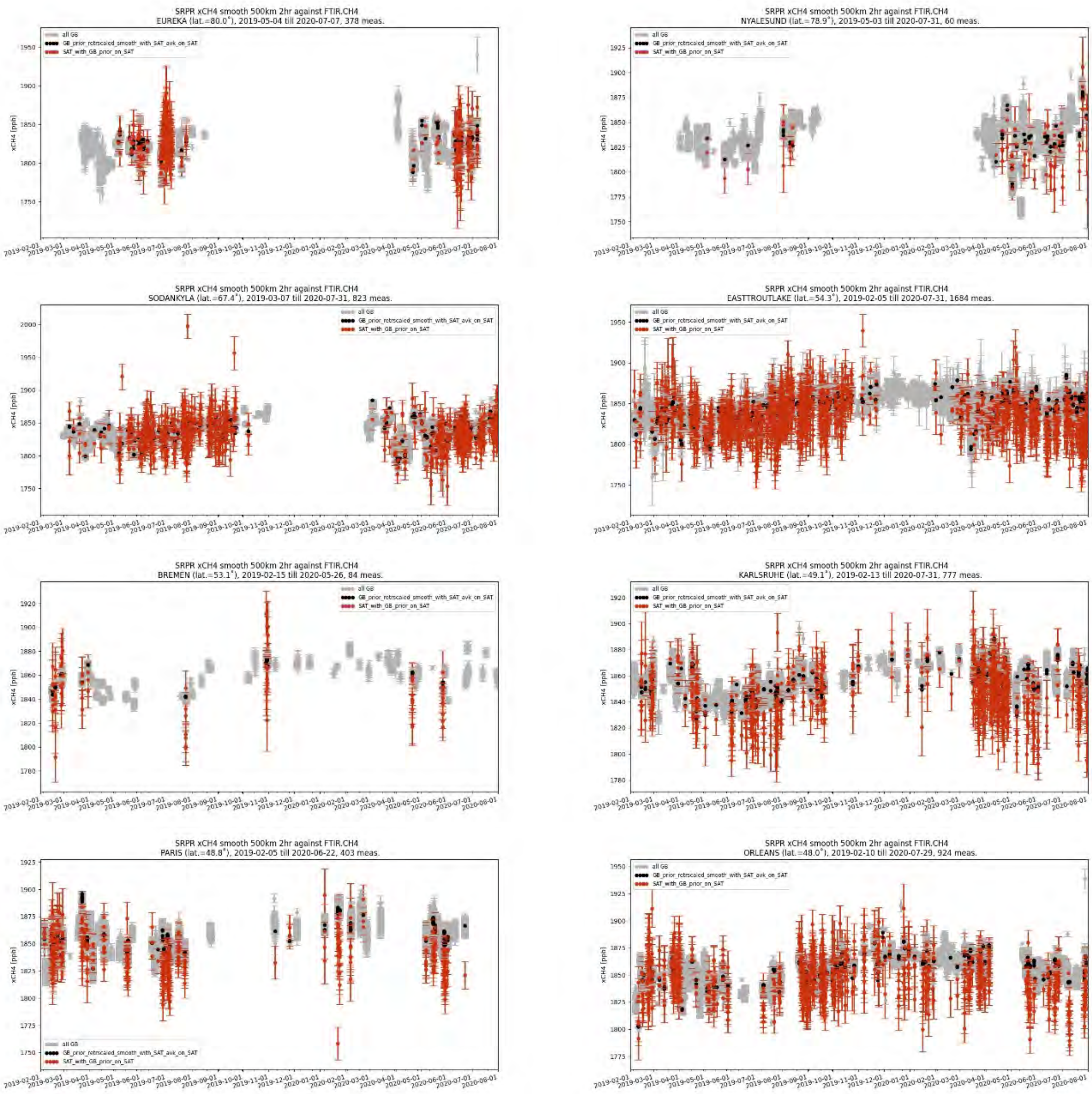
No long term trend analysis was performed since one needs at least 2 full years of data. However the mosaic plots (**Figures 4-34 and 4-35**) and the averaged monthly averaged timeseries do give a good indication (**Figures 4-38 and 4-39**)

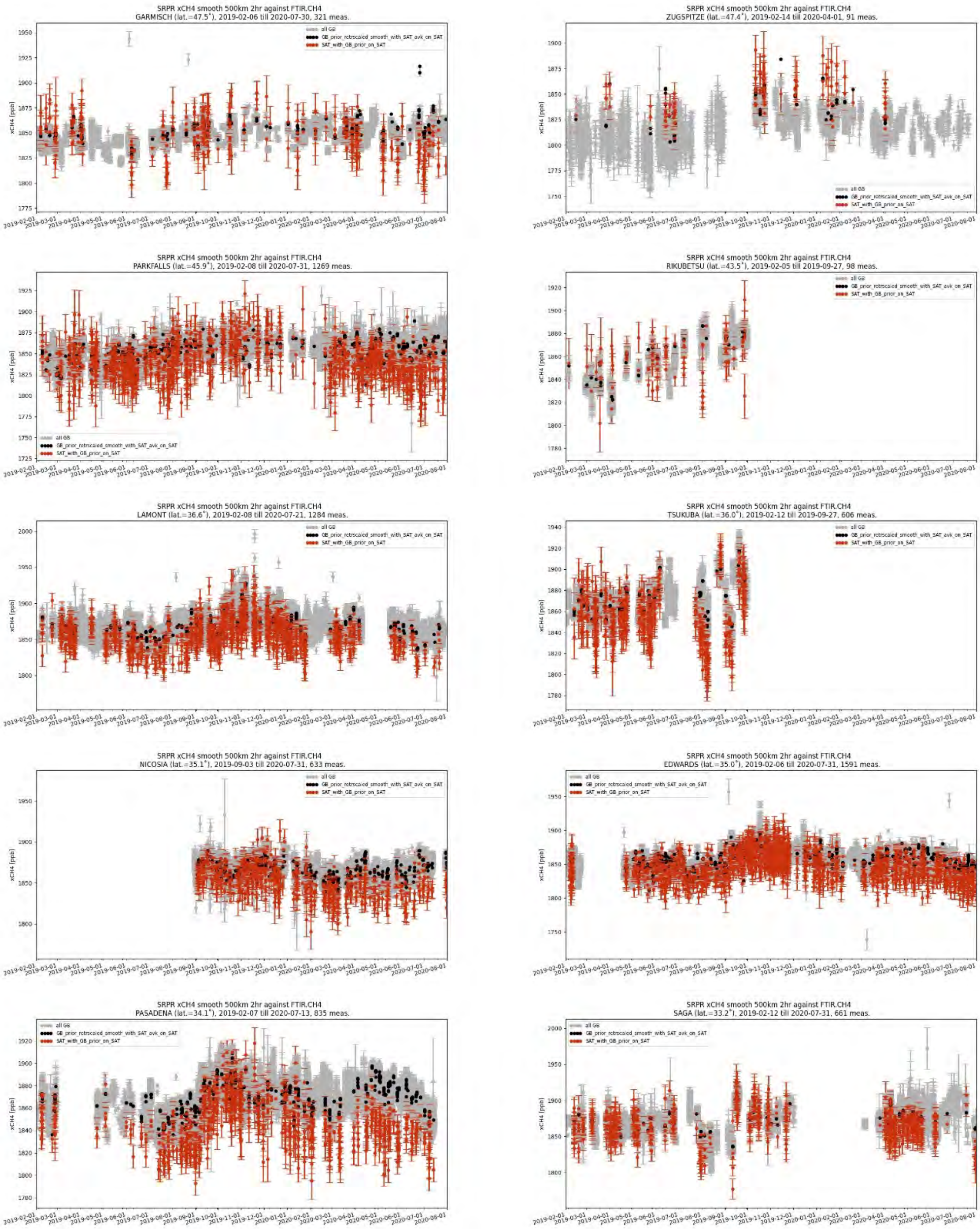
Table 4-16: Number of collocated data pairs (N), Correlation (R), Bias, Scatter, long term trend difference (ltt) and uncertainty thereon (ltt_err), seasonal amplitude difference (A) and uncertainty thereon (A_err) as well as the latitude of the NDACC station. The last row lists the median values over all stations. Product: CH4_GO2_SRPR.

STATION	N	R	Bias	Scat	ltt	ltt_err	A	A_err	lat
EUREKA	37	-0.04	-14.83	24.22	-	-	-	-	80.1
NY.ALESUND	23	0.54	-11.84	30.54	-	-	-	-	78.9
THULE	190	0.31	26.36	20.81	-	-	-	-	76.5
KIRUNA	209	0.36	-27.11	17.82	-	-	-	-	67.8
SODANKYLA	409	0.22	0.19	23.94	-	-	-	-	67.4
HARESTUA	28	0.45	16.78	12.95	-	-	-	-	60.2
ST.PETERSBURG	357	0.33	2.78	17.19	-	-	-	-	59.9
BREMEN	152	0.21	1.22	19.22	-	-	-	-	53.1
GARMISCH	319	-0.12	-14.94	27.42	-	-	-	-	47.5
ZUGSPITZE	223	0.02	-11.62	24.61	-	-	-	-	47.4
JUNGFRAUJOCH	251	0.17	-45.79	27.16	-	-	-	-	46.6
TORONTO	451	-0.11	14.66	41.50	-	-	-	-	43.6
RIKUBETSU	35	0.04	17.90	79.39	-	-	-	-	43.5
BOULDER.CO	483	0.21	-7.97	21.76	-	-	-	-	40.0
IZANA	284	0.36	-37.41	14.98	-	-	-	-	28.3
ALTZOMONI	45	-0.09	48.94	24.63	-	-	-	-	19.1
PARAMARIBO	15	0.53	-62.84	21.25	-	-	-	-	5.8
PORTO.VELHO	112	0.19	-15.25	19.08	-	-	-	-	-8.8
LA.REUNION.MAI	509	0.28	-25.14	12.67	-	-	-	-	-21.1
WOLLONGONG	510	0.06	-20.92	21.40	-	-	-	-	-34.4
LAUDER	625	0.14	-17.22	18.67	-	-	-	-	-45.0
ARRIVAL.HEIGHT	20	0.59	4.84	19.56	-	-	-	-	-77.8
MEDIAN	**216**	**0.21**	**-11.73**	**21.33**	**-**	**-**	**-**	**-**	**45.1**

The timeseries in **Figure 4-36** show individual satellite and ground-based TCCON measurements. While the scatter is even somewhat higher for SRPR XCH_4 with respect to both TCCON and SRFP, it is again relatively free of outliers and manages to capture (in most cases) TCCON's temporal variability.

Figure 4-37 shows the NDACC correlative data timeseries and here again it is obvious that NDACC in itself shows more variability (which affects single measurement precision and correlation numbers). See for instance Toronto and Boulder (at the start of the timeseries).




SRPR xCH4 smooth 500km 2hr against FTIR.CH4
GARMISCH (lat.=47.5°), 2019-02-06 till 2020-07-30, 321 meas.
SRPR xCH4 smooth 500km 2hr against FTIR.CH4
ZUGSPITZE (lat.=47.4°), 2019-02-14 till 2020-04-01, 91 meas.
SRPR xCH4 smooth 500km 2hr against FTIR.CH4
PARKFALLS (lat.=45.9°), 2019-02-08 till 2020-07-31, 1269 meas.
SRPR xCH4 smooth 500km 2hr against FTIR.CH4
RIKUBETSU (lat.=43.5°), 2019-02-05 till 2019-09-27, 98 meas.
SRPR xCH4 smooth 500km 2hr against FTIR.CH4
LAMONT (lat.=36.6°), 2019-02-08 till 2020-07-21, 1284 meas.
SRPR xCH4 smooth 500km 2hr against FTIR.CH4
TSUKUBA (lat.=36.0°), 2019-02-12 till 2019-09-27, 606 meas.
SRPR xCH4 smooth 500km 2hr against FTIR.CH4
NICOSIA (lat.=35.1°), 2019-09-03 till 2020-07-31, 633 meas.
SRPR xCH4 smooth 500km 2hr against FTIR.CH4
EDWARDS (lat.=35.0°), 2019-02-06 till 2020-07-31, 1591 meas.
SRPR xCH4 smooth 500km 2hr against FTIR.CH4
PASADENA (lat.=34.1°), 2019-02-07 till 2020-07-13, 835 meas.
SRPR xCH4 smooth 500km 2hr against FTIR.CH4
SAGA (lat.=33.2°), 2019-02-12 till 2020-07-31, 661 meas.
xCH4 [ppb]

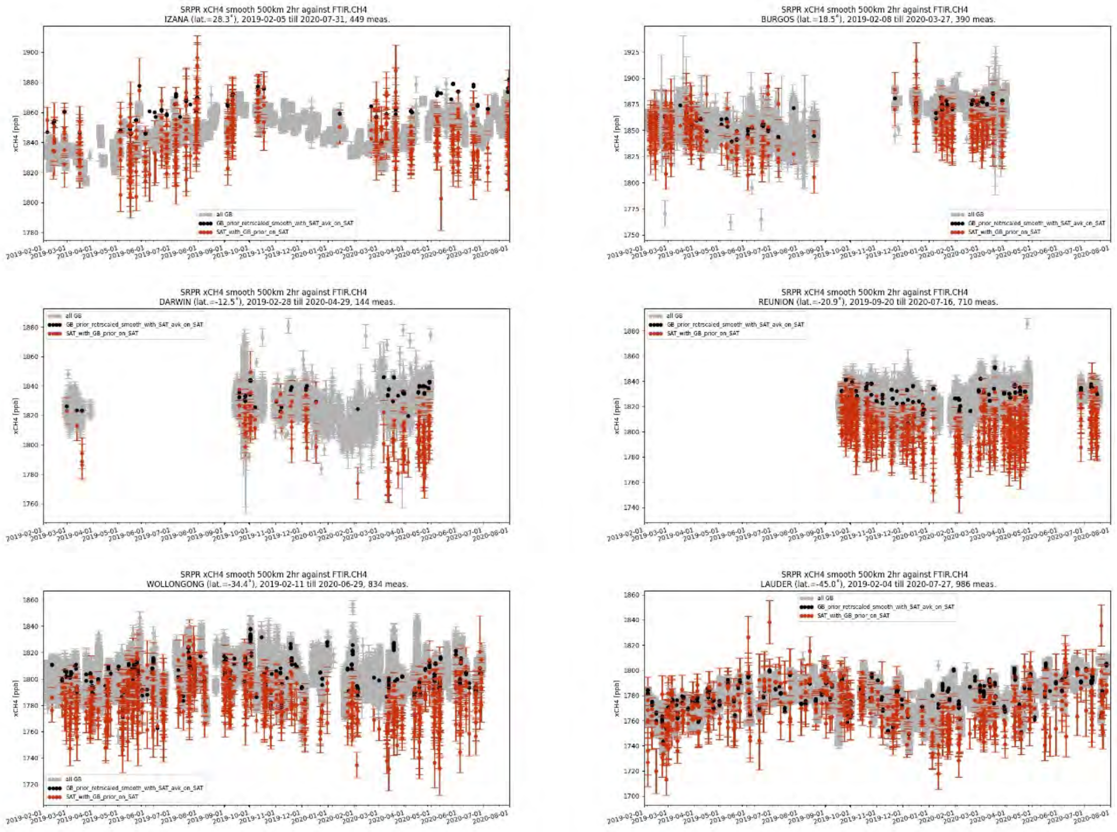
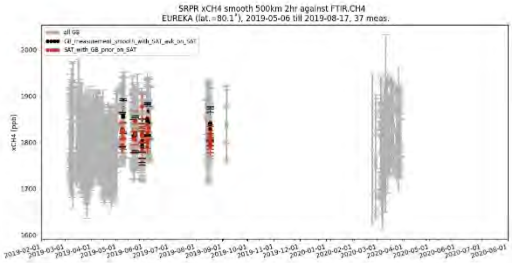
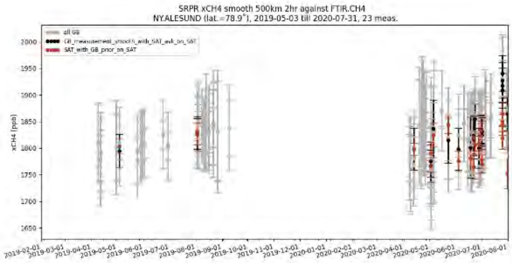
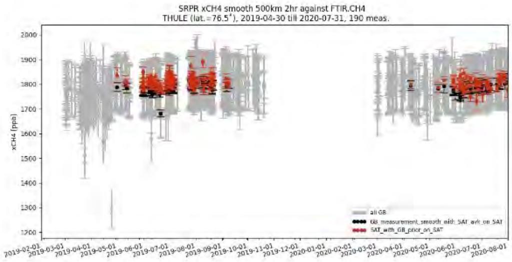
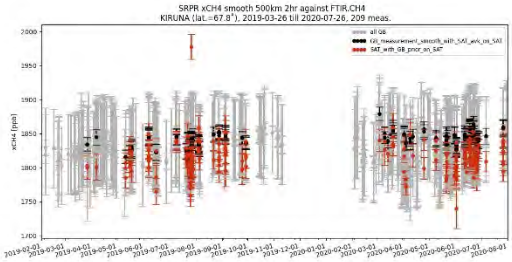


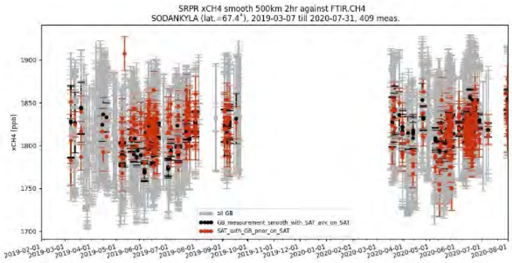
Figure 4-36: Timeseries of XCH_4 TCCON (collocated=black, all=grey) and CH4_GO2_SRPR (red) data at selected TCCON sites.

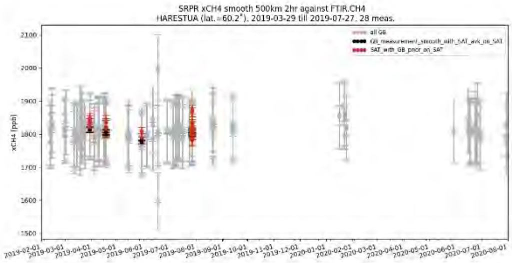

SRPR xCH4 smooth 500km 2hr against FTIR.CH4
EUREKA (lat.=80.1°), 2019-05-06 till 2019-08-17, 37 meas.

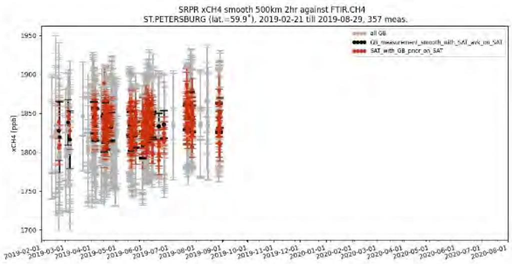

SRPR xCH4 smooth 500km 2hr against FTIR.CH4
NYALESUND (lat.=78.9°), 2019-05-03 till 2020-07-31, 23 meas.

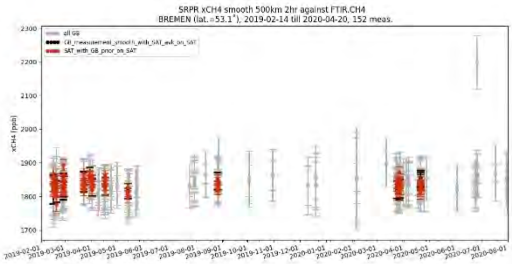

SRPR xCH4 smooth 500km 2hr against FTIR.CH4
THULE (lat.=76.5°), 2019-04-30 till 2020-07-31, 190 meas.

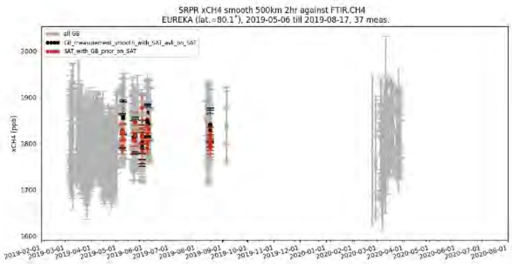

SRPR xCH4 smooth 500km 2hr against FTIR.CH4
KIRUNA (lat.=67.8°), 2019-03-26 till 2020-07-26, 209 meas.

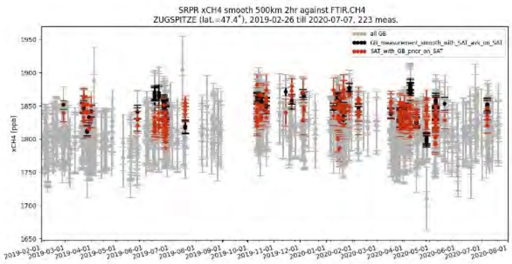

SRPR xCH4 smooth 500km 2hr against FTIR.CH4
SODANKYLA (lat.=67.4°), 2019-03-07 till 2020-07-31, 409 meas.

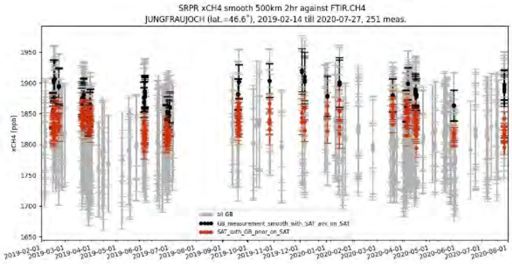

SRPR xCH4 smooth 500km 2hr against FTIR.CH4
HARESTUA (lat.=60.2°), 2019-03-29 till 2019-07-27, 28 meas.

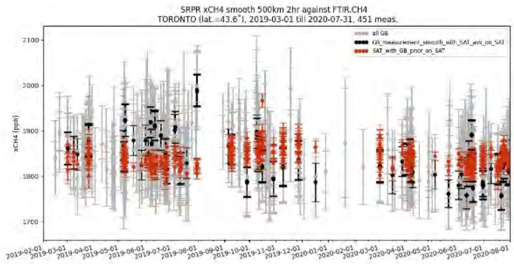

SRPR xCH4 smooth 500km 2hr against FTIR.CH4
ST.PETERSBURG (lat.=59.9°), 2019-02-21 till 2019-08-29, 357 meas.

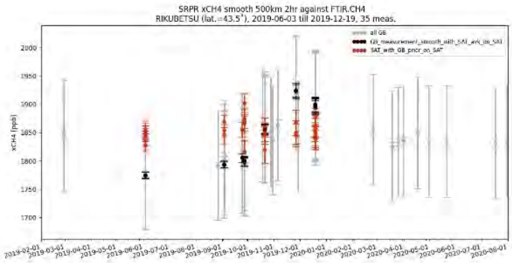

SRPR xCH4 smooth 500km 2hr against FTIR.CH4
BREMEN (lat.=53.1°), 2019-02-14 till 2020-04-20, 152 meas.

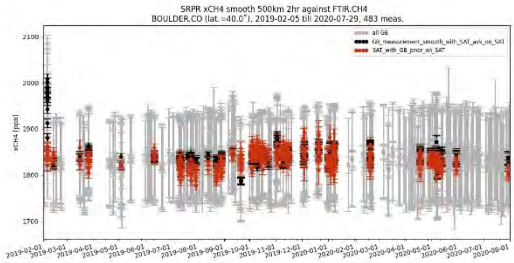

SRPR xCH4 smooth 500km 2hr against FTIR.CH4
EUREKA (lat.=80.1°), 2019-05-06 till 2019-08-17, 37 meas.

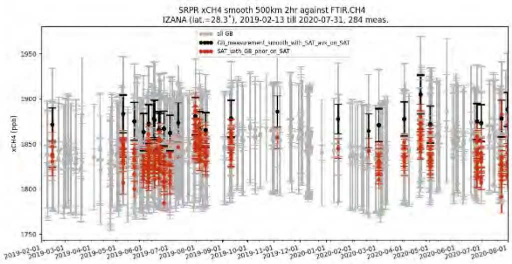

SRPR xCH4 smooth 500km 2hr against FTIR.CH4
ZUGSPITZE (lat.=47.4°), 2019-02-26 till 2020-07-07, 223 meas.

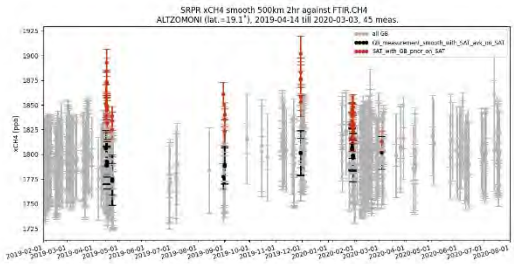

SRPR xCH4 smooth 500km 2hr against FTIR.CH4
JUNGFRAUJOCH (lat.=46.6°), 2019-02-14 till 2020-07-27, 251 meas.
xCH4 [ppb]
all GB
GB_measurement_smooth_with_SAT_avk_on_SAT
SAT_with_GB_prior_on_SAT

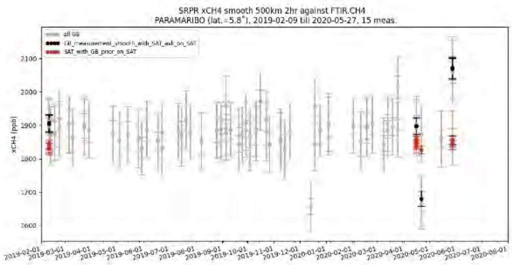

SRPR xCH4 smooth 500km 2hr against FTIR.CH4
TORONTO (lat.=43.6°), 2019-03-01 till 2020-07-31, 451 meas.
xCH4 [ppb]
all GB
GB_measurement_smooth_with_SAT_avk_on_SAT
SAT_with_GB_prior_on_SAT

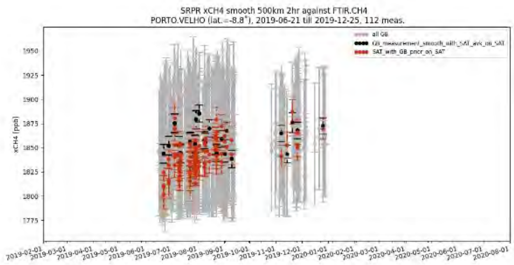

SRPR xCH4 smooth 500km 2hr against FTIR.CH4
RIKUBETSU (lat.=43.5°), 2019-06-03 till 2019-12-19, 35 meas.
xCH4 [ppb]
all GB
GB_measurement_smooth_with_SAT_avk_on_SAT
SAT_with_GB_prior_on_SAT


SRPR xCH4 smooth 500km 2hr against FTIR.CH4
BOULDER.CO (lat.=40.0°), 2019-02-05 till 2020-07-29, 483 meas.
xCH4 [ppb]
all GB
GB_measurement_smooth_with_SAT_avk_on_SAT
SAT_with_GB_prior_on_SAT


SRPR xCH4 smooth 500km 2hr against FTIR.CH4
IZANA (lat.=28.3°), 2019-02-13 till 2020-07-31, 284 meas.
xCH4 [ppb]
all GB
GB_measurement_smooth_with_SAT_avk_on_SAT
SAT_with_GB_prior_on_SAT


SRPR xCH4 smooth 500km 2hr against FTIR.CH4
ALTZOMONI (lat.=19.1°), 2019-04-14 till 2020-03-03, 45 meas.
xCH4 [ppb]
all GB
GB_measurement_smooth_with_SAT_avk_on_SAT
SAT_with_GB_prior_on_SAT


SRPR xCH4 smooth 500km 2hr against FTIR.CH4
PARAMARIBO (lat.=5.8°), 2019-02-09 till 2020-05-27, 15 meas.
xCH4 [ppb]
all GB
GB_measurement_smooth_with_SAT_avk_on_SAT
SAT_with_GB_prior_on_SAT


SRPR xCH4 smooth 500km 2hr against FTIR.CH4
PORTO.VELHO (lat.=-8.8°), 2019-06-21 till 2019-12-25, 112 meas.
xCH4 [ppb]
all GB
GB_measurement_smooth_with_SAT_avk_on_SAT
SAT_with_GB_prior_on_SAT

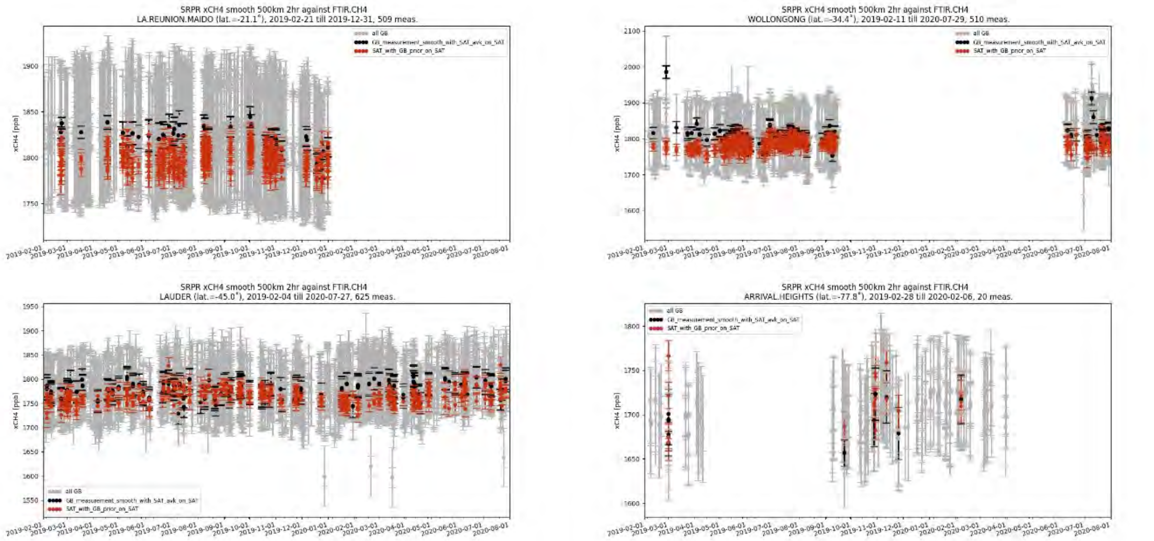


Figure 4-37: Timeseries of XCH_4 NDACC (collocated=black, all=grey) and CH4_GO2_SRPR (red) data at all NDACC sites.

Figure 4-38 shows monthly median timeseries for TCCON and SRPR XCH_4 for all data that fall within certain latitude bands, namely all sites North of 40°N latitude (top), all sites between 40°N and the equator (mid) and all sites in the Southern hemisphere (bottom). Here we see that TCCON exhibits consistently more positive long term trend values as compared to SRPR. Where for SRFP, the TCCON and algorihm trends were for the most part similar and always had overlapping uncertainties, this is not the case for SRPR. For the >40°N band we see a difference of 12.3 ppb/year, for the 0-40°N band we have a 7.8 ppb/ year difference and for the Southern hemisphere we see a 15.4 ppb/year difference. The seasonality however looks in order (note that past August we have no longer an uncertainty attached to the seasonal plot, since we only have August and onward data for a single year).

While these observations are of some concern we once again need to stress that for an accurate assessment of the trends we need more data and therefore these numbers are preliminary in nature.

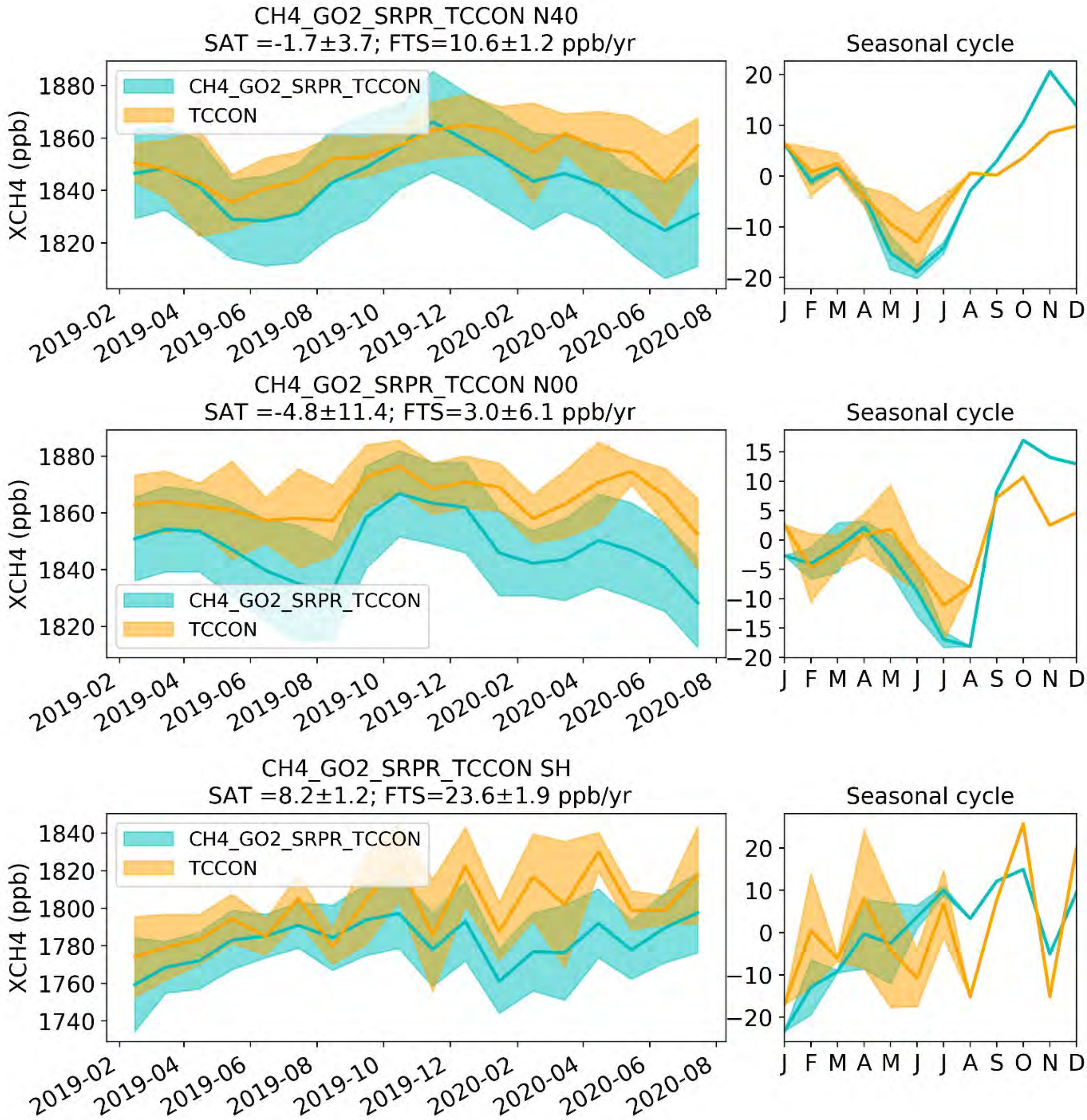


Figure 4-38: Monthly median collocated Sat and TCCON XCH_4 concentrations as a function of time. The shaded areas correspond with the scaled median absolute deviation.

Figure 4-39 shows the same but for NDACC. Ignoring the 0-40°N latitude band (see the analysis of **figure 4-23**), the long term trend difference is not as outspoken with overlapping 1σ-standard deviation intervals.

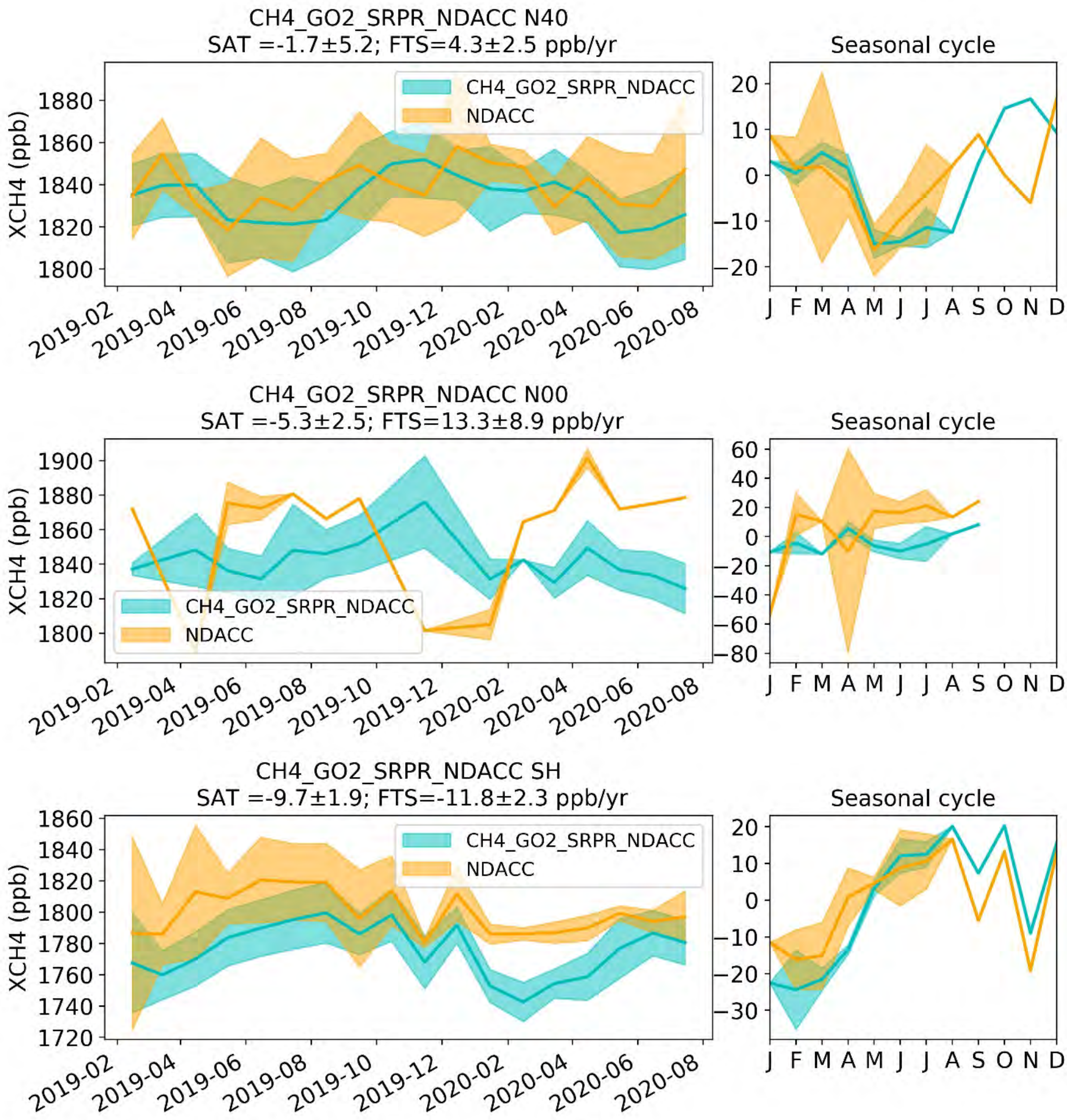


Figure 4-39: Monthly median collocated Sat and NDACC XCH_4 concentrations as a function of time. The shaded areas correspond with the scaled median absolute deviation.

4.2.6.2 Summary

Listed in the table below (**Table 4-17**) are the Figure of Merit parameters as derived from the individual collocated data pairs at each station.

SRPR XCH_4's single measurement precision equals 16.7 ppb, (just) reaching the Breakthrough target of <17 ppb. The error assessment is somewhat underestimated with an uncertainty ratio of 0.69. The median bias is, unlike its SRFP counterpart, significant at -13.3 ppb with confidence bands between -16.4 and -10.0 ppb. Both the spatial and spatio-temporal relative accuracies reach the <10 ppb target (5.0 and 9.4 ppb for the RA and SRA respectively), which is slightly better than SRFP's RA, but slightly worse than its SRA.

When we compare SRPR with SRFP for XCH_4, we observe more data at the cost of higher scatter values for the Proxy version and a significant difference in the overall bias. Of some concern for the proxy algorithm are indications that the bias is not stable in time. However, given the limited temporal coverage these observations are at this point indicative only.

Compared to NDACC we see a single measurement precision of 21.3 [18.2, 23.6] ppb, a likewise negative median bias of -11.7 [-26.2, -6.2] ppb, and relative accuracy values that do not meet the requirements (RA 20.7[5.4, 33.0] ppb, SRA 23.7 [14.4, 28.8] ppb). The latter no doubt in part to the higher inter-station variability within the NDACC network itself.

Table 4-17 presents an overview of the estimated data quality of CH4_GO2_SRPR, as obtained by the VALT team, from comparisons with TCCON ground-based reference observations. Values in square brackets [ ] correspond with the upper and lower 95% confidence bound on the parameter. The uncertainty ratio features 2 numbers as outlined in the validation method.

Product Quality Summary Table for Product: CH4_GO2_SRPR Level: 2, Version: v02.0.0, Time period covered: 2.2019 – 07.2020 Assessment: Validation Team (VALT)			
Parameter [unit]	**Achieved performance**	**Requirement**	**Comments**
Single measurement precision (1-sigma) in [ppm]	16.7 [15.3,17.7]	< 34 (T) < 17 (B) < 9 (G)	Computed as the median over all station scaled median absolute differences to TCCON
Uncertainty ratio [-]: Ratio reported uncertainty to standard deviation of satellite-TCCON difference	0.68,0.69*	-	No requirement but value close to unity expected for a high quality data product with reliable reported uncertainty.
Median bias (global offset) [ppm]	-13.2 [-16.4,-10.0]	-	No requirement but value close to zero expected for a high quality data product.
Accuracy: Relative systematic error [ppm]	Spatial: 5.0 [1.4, 7.0] Spatio-temporal: 9.4 [6.6, 12.0]	< 10	Spatial: Computed as standard deviation of the biases at the various TCCON sites. Spatio-temporal: As "Spatial" but also considering seasonal biases.
Stability: Drift [ppm/year]	-	< 3	Linear drift

5 Validation and intercomparisons results from data provider

5.1 Validation and intercomparison results for product CO2_OC2_FOCA

The validation results shown in this section are valid for version 10 of the OCO-2 XCO_2 retrieval algorithm CO2_OC2_FOCA. The applied methods are similar to those described in BESD's comprehensive error characterization Report **/CECRv3, 2017/** and product validation and inter-comparison reports (e.g., **/PVIRv5, 2017/)** of ESA's GHG CCI project and partly also in the publication of **/Reuter et al., 2020/**. For all comparisons, averaging kernels have been applied and the influence of the smoothing error reduced as described in Section 5.2 of ESA's GHG CCI+ product user guide version 4 (PUGv4) for the FOCAL XCO_2 OCO-2 data product CO2_OC2_FOCA **/PUGv4, 2022/**. The validation results shown in this section are part of ESA's GHG CCI+ end-to-end ECV uncertainty budget version 3 (E3UBv3) for the FOCAL XCO_2 OCO-2 data product CO2_OC2_FOCA **/E3UBv3, 2022/**.

5.1.1 Co-location

FOCAL's XCO_2 has been validated with TCCON GGG2014 measurements **/Wunch et al., 2011/**. The co-location criteria are defined by a maximum time difference of two hours, a maximum spatial distance of 500km, and a maximum surface elevation difference of 250m. Additionally, only TCCON sites with at least 1000 co-locations (4 in the case of daily, weekly, or monthly averages) covering a time period of at least two years are taken into account.

Figure 5.1 shows all 2331159 co-located FOCAL and TCCON XCO_2 retrieval results used for the validation study. One can see that the temporal sampling differs from site to site and that FOCAL captures the year-to-year increase and the seasonal features well.

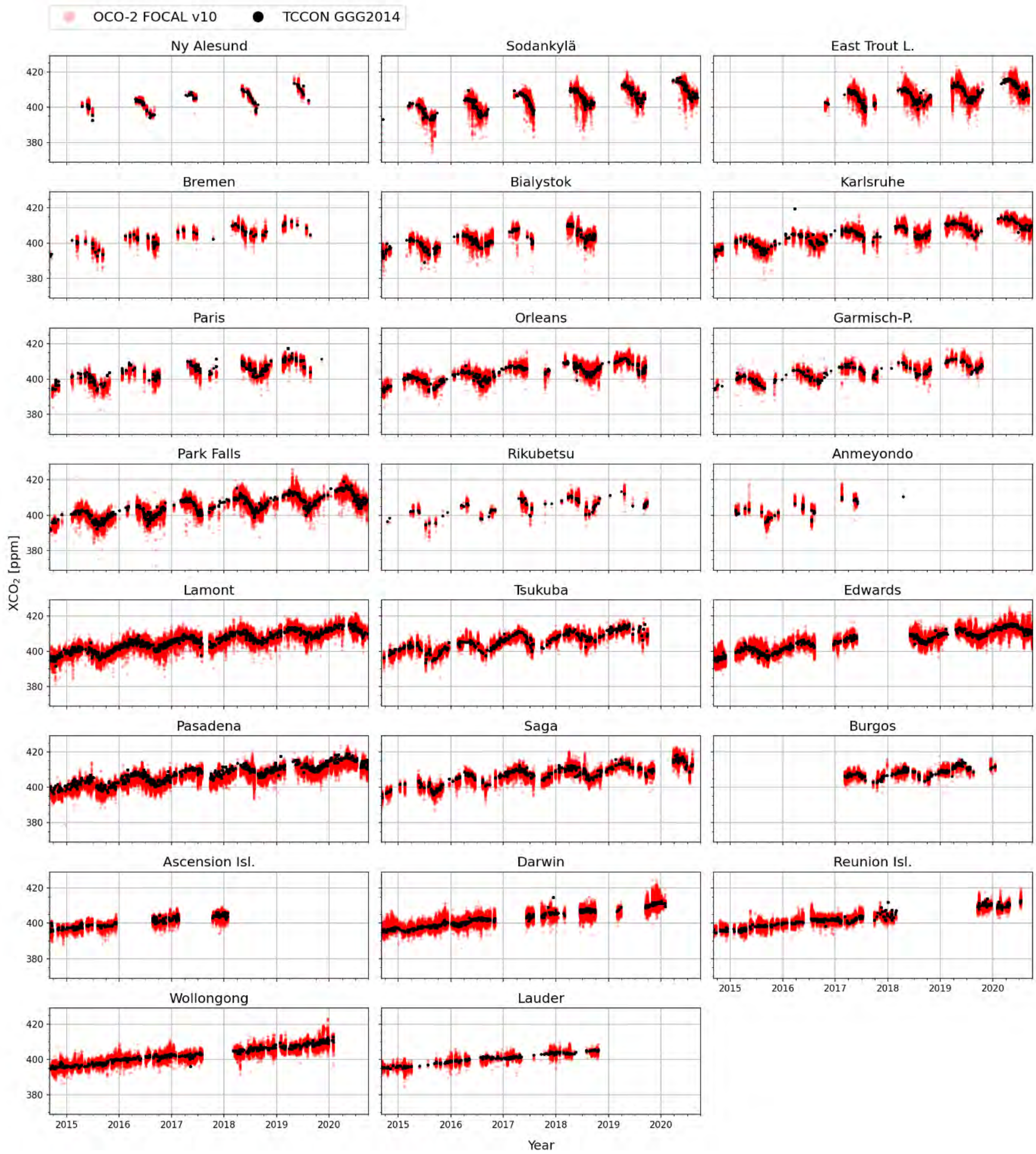


Figure 5.1: *Co-located FOCAL and TCCON XCO_2 retrieval results used for the validation study. The TCCON sites are order from top/left to bottom/right by average latitude of the co-located satellite soundings.*

5.1.2 Daily, weekly, and monthly averages

For some applications, it is expected that FOCAL XCO_2 data will be aggregated to “super soundings” averaging, e.g., all soundings of an orbit in a surrounding of a target. Also FOCAL XCO_2 data might be used to compute L3 (level 3) products, e.g., in the manner of gridded monthly averages. With such application in the mind, we computed daily, weekly, and monthly averages of the FOCAL and TCCON co-locations at each TCCON site. In order to improve the robustness, daily, weekly, and monthly averages are only calculated when averaging at least 10, 30, or 50 individual soundings, respectively. As an example, Figure 5.2 shows the daily, weekly, and monthly FOCAL XCO_2 averages for the Lamont and Darwin TCCON sites. Due to OCO-2’s data density, it is often the case that one overpass generates many co-colocations. This considerably reduces the scatter of the daily averages compared to the individual soundings.

Note that FOCAL reports only on the stochastic uncertainty of the individual soundings. In the case of daily, weekly, and monthly averages we computed the corresponding uncertainties by applying the rules of error propagation under the assumption of uncorrelated errors.

5.1.3 General overview

The overall agreement of the FOCAL data (and its averages) with TCCON data at all sites is illustrated in Figure 5.3. The histograms of the difference (FOCAL – TCCON) show in all cases a near Gaussian distribution with a center between -0.43ppm and -0.29ppm. The standard deviation of the difference reduces from 1.88ppm for individual soundings to 1.14ppm for monthly averages. The FOCAL vs. TCCON heat maps show a pronounced clustering along the one-to-one line for all cases. This is supported by a good agreement of the orthogonal distance regression with the one-to-one line and high Pearson correlation coefficients between 0.94 for individual soundings and 0.97 for monthly averages.

These results provide a first rough overview of FOCAL's agreement with TCCON. However, except for an average bias, they do not allow to separate systematic and stochastic error components.

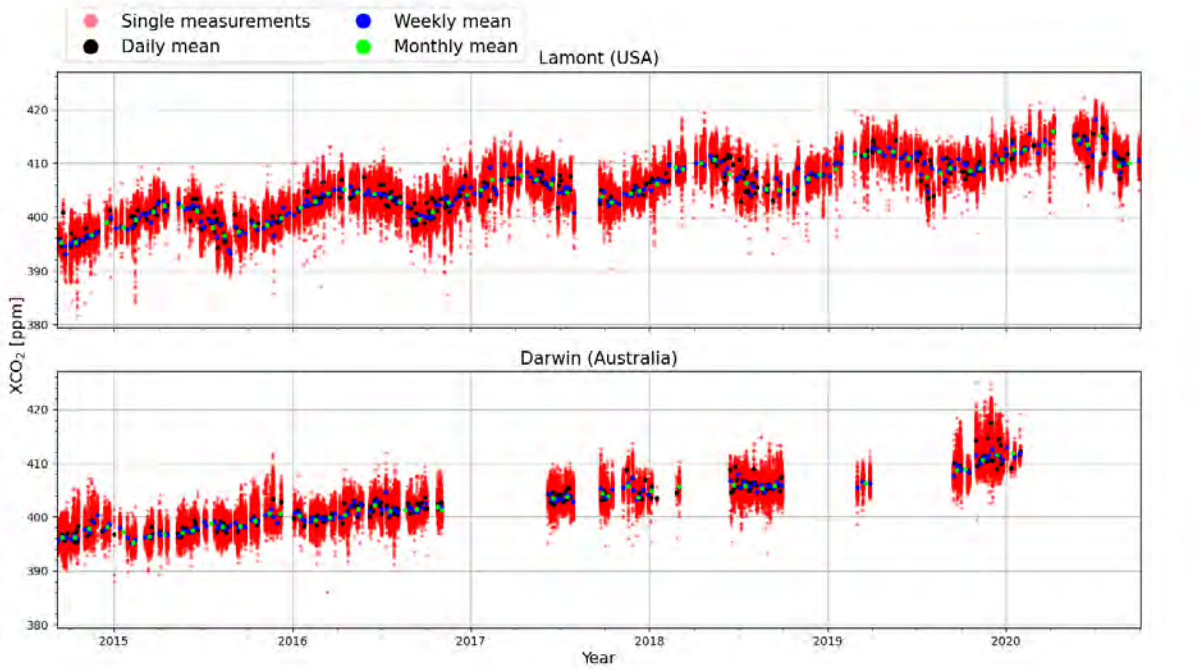


Figure 5.2: *Co-located FOCAL XCO_2 retrieval results and their daily, weekly, and monthly averages at the TCCON sites Lamont (top) and Darwin (bottom) used for the validation study.*

5.1.4 Stochastic and systematic error components

The method described in the following allows us to separate the stochastic errors from potential regional or seasonal biases as well as from a linear drift.

5.1.4.1 Per site performance statistics

For the co-locations of each site, we compute the FOCAL minus TCCON differences ΔX and fit the following bias model:

5-1
$$\Delta X = a_0 + a_1 t + a_2 \sin(2\pi t + a_3) + \varepsilon$$

Here, t is the time of the measurements in fractional years, a_{0-3} the free fit parameters from which we compute the systematic error components, and ε the fit residuum. Figure 5.4 shows at the example of the TCCON sites Lamont and Darwin the fitted bias functions for the individual soundings, daily, weekly, and monthly averages.

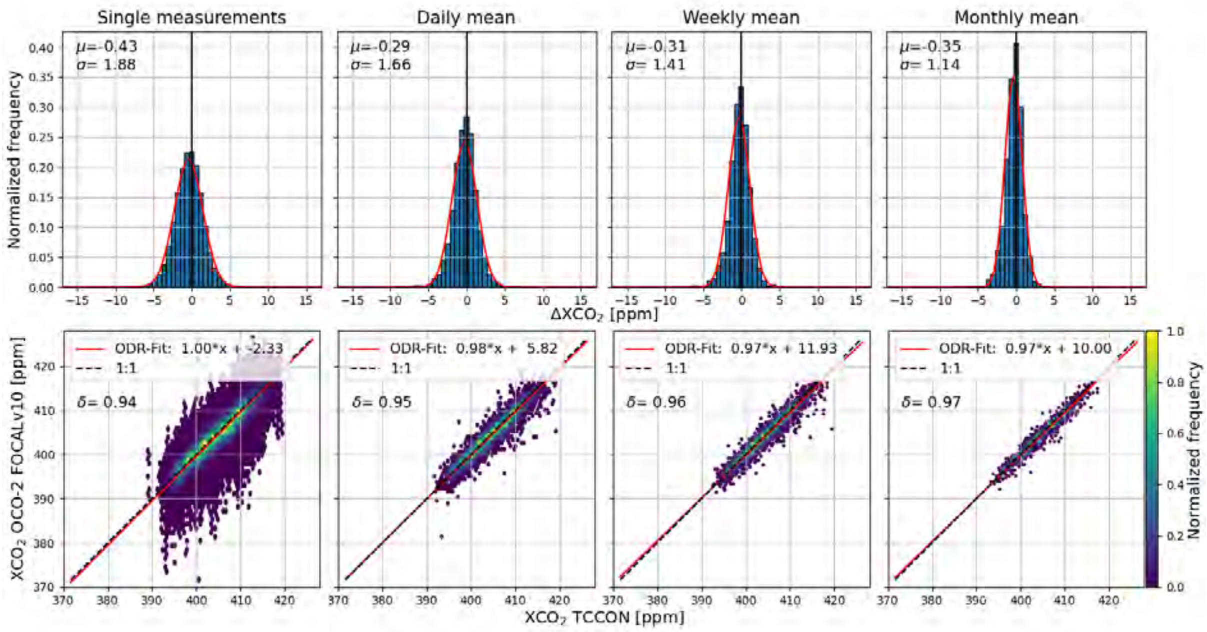


Figure 5.3: *Overall overview on the agreement of the FOCAL data (and its averages) with TCCON data at all sites. Top: Normalized histograms of the difference FOCAL – TCCON. Bottom: Heat maps TCCON vs. FOCAL including one-to-one line, orthogonal distance regression (ODR), and Pearson correlation coefficient **δ**.*

We compute the station or regional bias Δ_{reg} from the average (ave) of the fit values:

5-2
$$\Delta_{reg} = \text{ave}[a_0 + a_1 t + a_2 \sin(2\pi t + a_3)]$$

The seasonal bias Δ_{sea} is computed from the standard deviation (std) of the seasonal component of the fit:

5-3
$$\Delta_{sea} = \text{std}[a_2 \sin(2\pi t + a_3)]$$

It shall be noted that the vector t consists only of the time of the measurements. This means, Δ_{sea} is only computed from those parts of the seasonal cycle actually covered by observations.

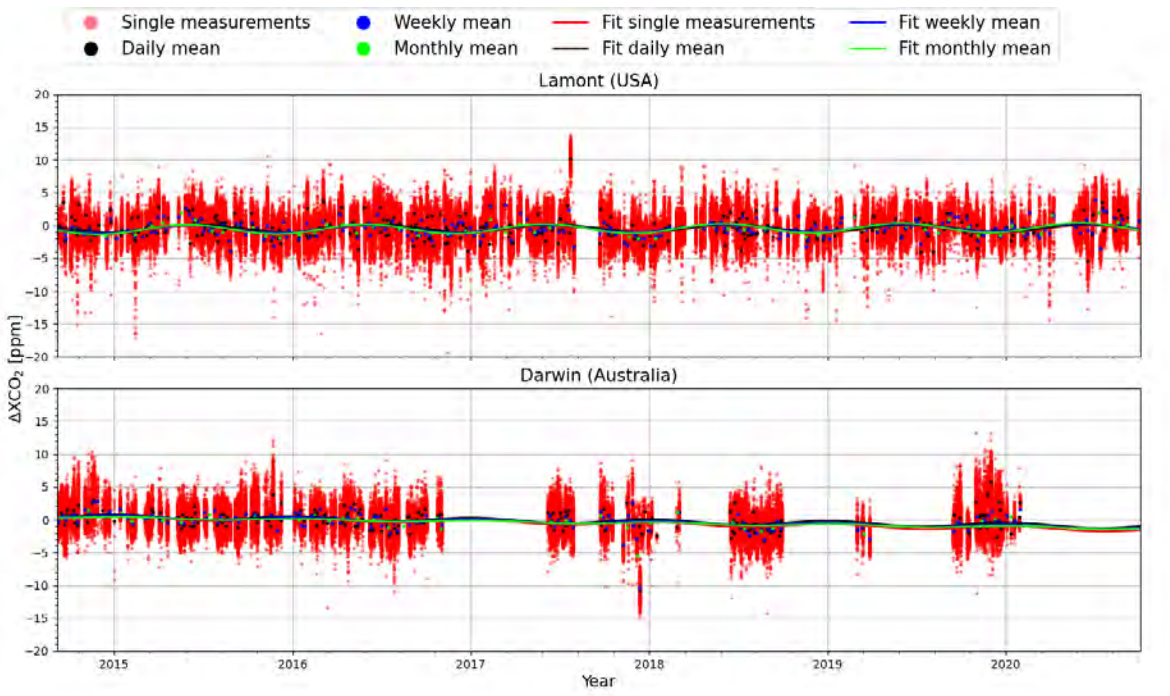


Figure 5.4: *ΔXCO_2 (FOCAL – TCCON) for the co-locations of the single measurements, daily, weekly, and monthly averages at the TCCON sites Lamont (top) and Darwin (bottom). Additionally, the corresponding fits of the bias model (Eq. 5-1) are shown.*

The linear drift corresponds to the fit parameter $\Delta_{dri} = a_1$, and the single sounding precision, i.e., the stochastic retrieval uncertainty σ, is computed from the standard deviation of the residuum.

5-4 $$\sigma = \mathrm{std}[\varepsilon]$$

We define the spatiotemporal bias Δ_{spt} as combination of regional and seasonal bias.

5-5 $$\Delta_{spt} = \sqrt{{\Delta_{reg}}^2 + {\Delta_{sea}}^2}$$

The FOCAL retrieval algorithm reports on the XCO_2 stochastic uncertainty σ'_{rep} for each sounding. From these values, we compute the average reported uncertainty σ_{rep} per station by:

5-6 $$\sigma_{rep} = \sqrt{\text{ave}\left({\sigma'_{rep}}^2\right)}$$

5.1.4.2 Summarizing performance statistics

Based on the per site statistics, the following summarizing performance statistics are calculated.

The average site bias $\overline{\Delta_{reg}}$ and the site-to-site variability is computed from the mean and the standard deviation of the individual site biases:

5-7 $$\overline{\Delta_{reg}} = \text{ave}\left(\Delta_{reg}\right) \pm \text{std}\left(\Delta_{reg}\right)$$

The average seasonal bias $\overline{\Delta_{sea}}$ is computed by:

5-8 $$\overline{\Delta_{sea}} = \text{avg}(\Delta_{sea})$$

The overall spatiotemporal bias $\overline{\Delta_{spt}}$ is computed by:

5-9 $$\overline{\Delta_{spt}} = \sqrt{\overline{\Delta_{reg}}^2 + \overline{\Delta_{sea}}^2}$$

The average drift and the drift uncertainty is computed by:

5-10 $$\overline{\Delta_{dri}} = \text{ave}(\Delta_{dri}) \pm \text{std}(\Delta_{dri})$$

As the linear drift can be assumed to be globally constant, the station-to-station standard deviation of the linear drift can be considered a measure of its uncertainty. The overall single sounding precision and reported uncertainty are computed by:

5-11 $$\overline{\sigma} = \sqrt{\text{ave}(\sigma^2)}$$

5-12 $$\overline{\sigma_{rep}} = \sqrt{\text{ave}\left({\sigma_{rep}}^2\right)}$$

5.1.5 Results

The results of all site performance statistics as well as the summarizing performance statistics for individual soundings, daily, weekly, and monthly averages are illustrated in Figure 5.5. Based on this figure, it can first be noted that averaging does not have a substantial impact on the validation results for the systematic error components. This is especially the case for the summarizing performance statistics which are similar for individual soundings, daily, weekly, and monthly averages. Therefore, it is sufficient that we primarily concentrate on the results for individual soundings from now on and Table 5.1 lists only values of the statistics for individual soundings.

However, the results for the stochastic error component show some important differences. The overall result for the stochastic error of the individual soundings amounts to 1.69ppm which agrees well with the corresponding reported uncertainty of 1.69ppm. This is not the case for the results of the averages. The actual stochastic error reduces for daily (1.39ppm), weekly (1.11ppm), and monthly (0.80ppm) averages, but the reduction is far less pronounced as for the reported uncertainty which has been computed under the assumption of uncorrelated errors. Therefore, it has to be expected that the separation of systematic and stochastic errors by Eq. 5-1 is incomplete at least for the individual soundings. In other words, it can be expected that parts of the residuum ε of Eq. 5-1 for the individual soundings are actually of systematic origin.

For this reason, we grouped the residuum into bins consisting of $n = 1, 2, 3, \cdots$ elements and analyzed its standard deviation as function of the bin size. As the reported retrieval precision is usually relatively constant at one TCCON site, it should be expected that the standard deviation of the binned residuum scales approximately with $1/\sqrt{n}$. We performed this experiment for the TCCON site Lamont because of the large number of co-locations. As shown in Figure 5.6 (top/left), the actual precision (standard deviation of the binned residuum) of the individual soundings does not follow the curve expected for uncorrelated errors. In contrast, the actual precision of daily (Figure 5.6, top/right), weekly (Figure 5.6, bottom/left), and monthly averages (Figure 5.6, bottom/right) agrees well with the expectation for uncorrelated errors. These results may differ in detail from TCCON site to TCCON site, but indicates that the errors of the individual soundings may have additional systematic components not covered by the seasonal component of Eq. 5-1.

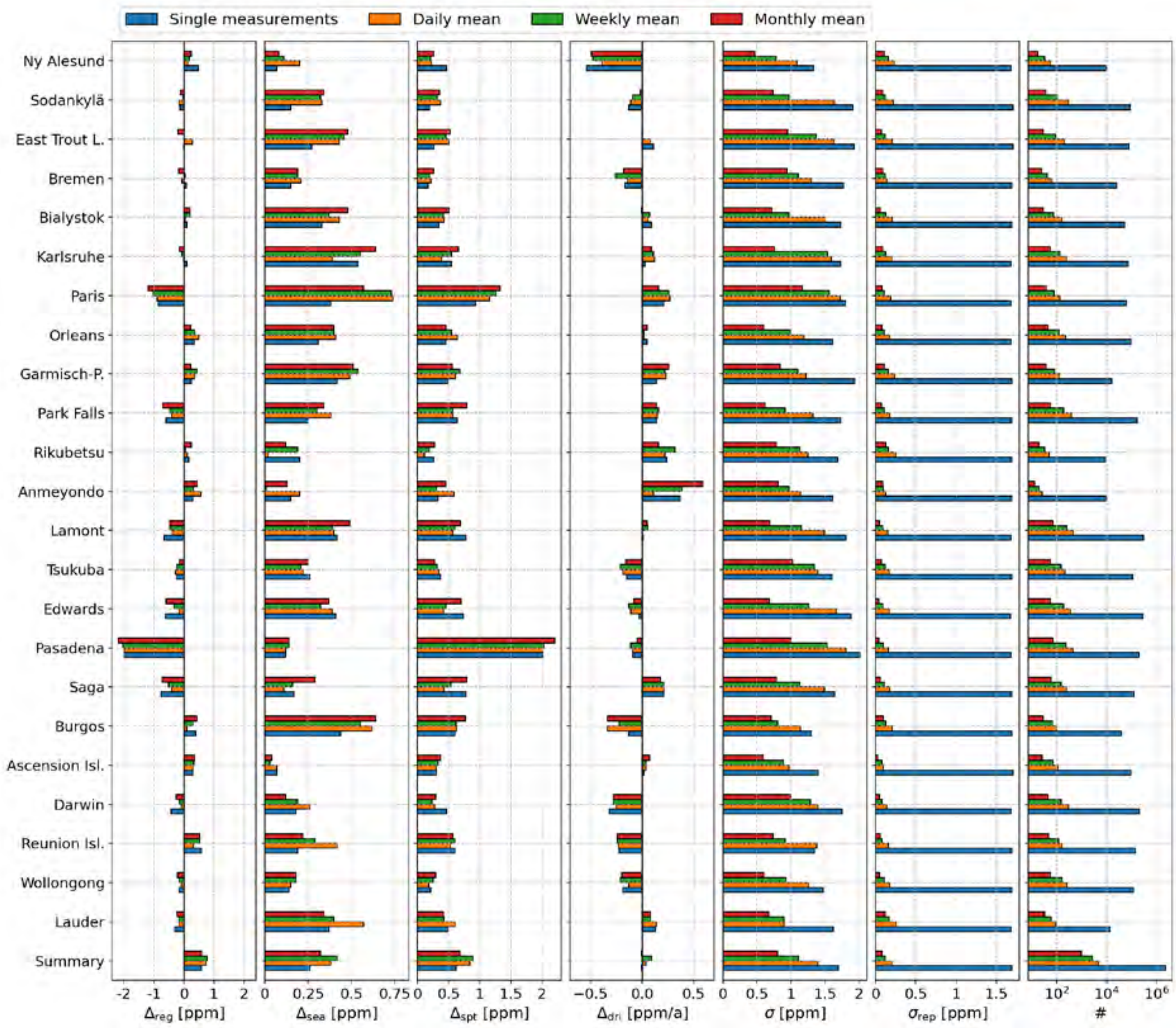


Figure 5.5: *Validation results for FOCAL single measurements, daily, weekly, and monthly averages. From left to right, the figure shows the per site performance statistics (Section 5.1.4.1) regional (Δ_{reg}), seasonal (Δ_{sea}), and spatiotemporal bias (Δ_{spt}), the linear drift (Δ_{dri}), the actual (σ) and reported precision (σ_{rep}), and the number of soundings (#). TCCON sites are order from top to bottom by average latitude of the co-located satellite soundings. The last row includes the summarizing performance statistics as defined in Section 5.1.4.2.*

Table 5.1: *Validation results for FOCAL single measurements. From left to right, the table lists the per site performance statistics (Section 5.1.4.1) regional (Δ_{reg}), seasonal (Δ_{sea}), and spatiotemporal bias (Δ_{spt}), the linear drift (Δ_{dri}), the actual (σ) and reported precision (σ_{rep}), and the number of soundings (#). TCCON sites are order from top to bottom by average latitude of the co-located satellite soundings. The last row includes the summarizing performance statistics as defined in Section 5.1.4.2.*

Station	Δ_{reg} [ppm]	Δ_{sea} [ppm]	Δ_{spt} [ppm]	Δ_{dri} [ppm/a]	σ [ppm]	σ_{rep} [ppm]	#
Ny Alesund	0.47	0.07	0.47	-0.54	1.33	1.68	10223
Sodankylä	-0.13	0.15	0.20	-0.13	1.91	1.70	86093
East Trout Lake	-0.02	0.27	0.27	0.11	1.92	1.70	80390
Bremen	0.08	0.15	0.17	-0.17	1.77	1.69	25681
Bialystok	0.09	0.33	0.35	0.09	1.73	1.69	51219
Karlsruhe	0.09	0.54	0.55	0.03	1.73	1.68	73437
Paris	-0.86	0.38	0.94	0.21	1.80	1.68	63569
Orleans	0.34	0.31	0.46	0.05	1.61	1.68	95043
Garmisch-P.	0.25	0.42	0.49	0.14	1.93	1.69	15863
Park Falls	-0.59	0.25	0.64	0.14	1.73	1.69	167324
Rikubetsu	0.16	0.20	0.26	0.24	1.68	1.69	8879
Anmeyondo	0.29	0.15	0.33	0.37	1.61	1.69	9904
Lamont	-0.66	0.42	0.78	0.01	1.81	1.68	301063
Tsukuba	-0.25	0.26	0.37	-0.15	1.60	1.69	110964
Edwards	-0.61	0.41	0.74	-0.03	1.88	1.67	287748
Pasadena	-2.00	0.12	2.01	-0.09	2.01	1.68	197002
Saga	-0.76	0.17	0.78	0.21	1.64	1.69	127152
Burgos	0.40	0.44	0.60	-0.13	1.30	1.69	39354
Ascension Isl.	0.30	0.07	0.31	0.02	1.40	1.70	95046
Reunion Isl.	-0.44	0.18	0.47	-0.32	1.75	1.68	210700
Darwin	0.57	0.19	0.60	-0.22	1.34	1.69	140479
Wollongong	-0.16	0.14	0.21	-0.19	1.47	1.69	120422
Lauder	-0.32	0.37	0.49	0.13	1.62	1.68	13604
Summary	**-0.16±0.57**	**0.26**	**0.62**	**-0.01±0.20**	**1.69**	**1.69**	**2331159**

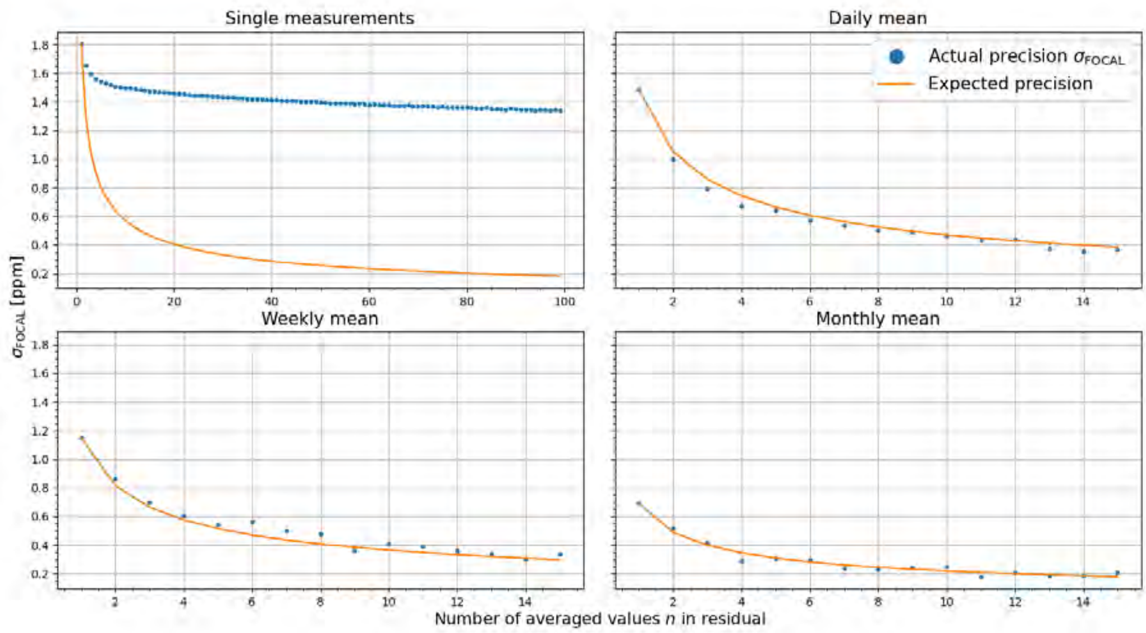


Figure 5.6: *Actual and expected retrieval precision of FOCAL computed from residuals with increasing bin size for the TCCON site Lamont for single measurements (top/left), daily (top/right), weekly (bottom/left), and monthly averages (bottom/right).*

The validation results for the individual soundings (Table 5.1, Figure 5.5) show that there is only a small overall average bias of -0.16ppm. Regional biases estimated from the site-to-site bias variability amount to 0.57ppm and are strongly influenced by the relatively large negative bias of -2ppm at the TCCON site Pasadena. The average seasonal and spatiotemporal bias amounts to 0.26ppm and 0.62ppm, respectively. The overall linear drift of 0.01ppm/a is much smaller than its site-to-site variability of 0.2ppm and, therefore, considered not significant.

Additionally, a measure for the year-to-year stability is computed as follows. For each TCCON site, the residual ε of the bias fit (Eq. 5-1) is smoothed by a running average of 365 days. Only days where more than 10 co-locations contribute to the running average of at least 5 TCCON sites are further considered. At these days, the station-to-station average is calculated (Figure 5.7, black line).

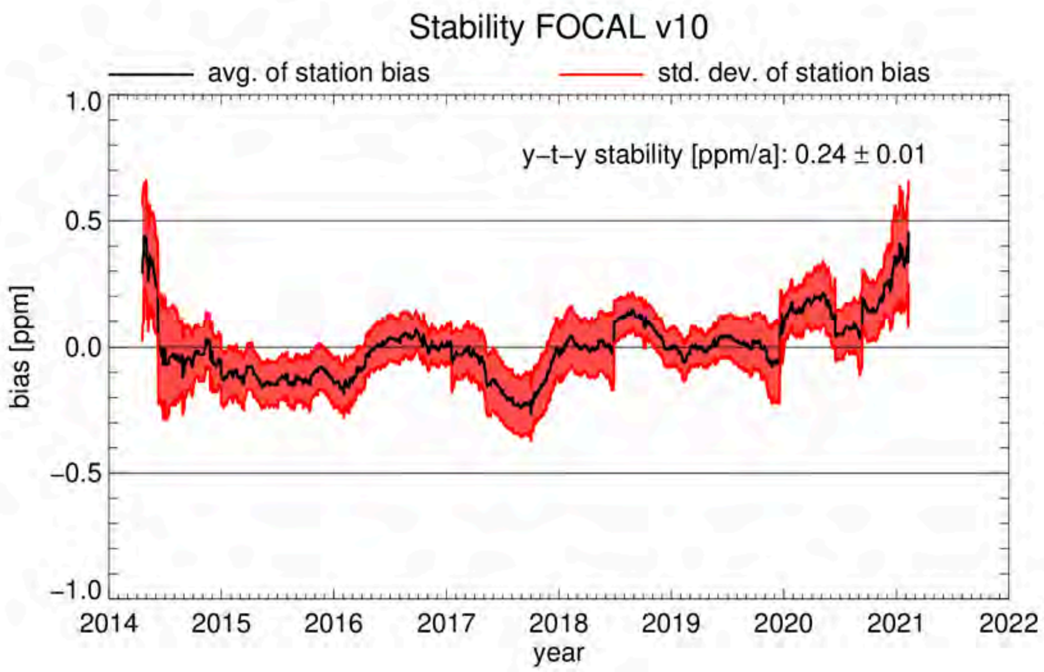


Figure 5.7: S*tability analyses for FOCAL. The black curve shows the average station bias and the red curves its uncertainty represented by the station-to-station standard deviation*

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 (Figure 5.7, red line). For FOCAL, the average is always between about -0.25ppm and 0.45ppm with an uncertainty of typically about 0.15ppm. Most of the time, the average is not significantly different from zero, i.e., its two sigma uncertainty is larger than its absolute value. Due to the relatively large uncertainty, we decided to compute not the maximum minus minimum as a measure for the year-to-year stability because this quantity can be expected to increase with length of the time series simply due to statistics. Therefore, we estimate the year-to-year stability by randomly selecting pairs of dates with a time difference of at least 365 days. For each selection we computed the difference modified by a random component corresponding to the estimated uncertainty. From 1000 of such pairs we compute the standard deviation as estimate for the year-to-year stability. We repeat this experiment 1000 times and compute the average (0.24ppm) and standard deviation (0.01ppm). From this, we conclude that the year-to-year stability is 0.24ppm/a (Figure 5.7).

5.1.6 Summary

We validated the FOCAL v10 XCO_2 data product with TCCON GGG2014 data of the years 2014 – 2021. The validation has been performed for daily, weekly, and monthly averages as well as for single soundings. Analyzing the single soundings without temporal averaging, we find that the overall bias of the FOCAL data amounts to -0.16ppm. Regional biases vary from site to site by 0.57ppm. Seasonal and spatiotemporal biases amount on average to 0.26ppm and 0.62ppm, respectively. We found no significant linear drift (-0.01±0.20ppm). In the context of the systematic error characteristics, it shall be noted that **/Wunch et al., 2010, 2011/** specifies the accuracy (1σ) of TCCON to be about 0.4ppm. This means, e.g., that it cannot be expected to find regional biases considerably less than 0.4ppm using TCCON as reference. We find that the inferred systematic errors, i.e., regional, seasonal, and spatiotemporal biases as well as linear drift, do not critically depend on averaging. The year-to-year stability has been estimated to be 0.24ppm/a. The overall precision of the individual soundings is 1.69ppm which agrees well with the corresponding reported uncertainty of 1.69ppm. The overall precision improves for daily (1.39ppm), weekly (1.11ppm), and monthly (0.80ppm) averages. We find indications that the estimated precision of the individual soundings does actually comprise not only purely stochastic but also residual unknown systematic components. No such indications were found for the daily, weekly, and monthly averages. **Table 5.1-2** presents an overview of the estimated data quality as obtained from comparisons with TCCON ground-based reference observations.

Table 5.1-2: *Summary validation of product CO2_OC2_FOCA.*

Product Quality Summary Table for Product: CO2_OC2_FOCA Level: 2, Version: v10, Time period covered: 9.2014 – 10.2020 Assessment: Data Provider (DP)			
Parameter [unit]	**Achieved performance**	**Requirement**	**Comments**
Single measurement precision (1-sigma) in [ppm]	1.69	< 8 (T) < 3 (B) < 1 (G)	Computed as standard deviation of the difference to TCCON
Uncertainty ratio [-]: 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 with reliable reported uncertainty.
Mean bias (global offset) [ppm]	-0.16	-	No requirement but value close to zero expected for a high quality data product.
Accuracy: Relative systematic error [ppm]	Spatial: 0.57 Spatiotemporal: 0.62	< 0.5	Spatial: Computed as standard deviation of the biases at the various TCCON sites. Spatio-temporal: As “Spatial” but also considering seasonal biases.
Stability: Drift [ppm/year]	-0.01±0.20 (1-sigma)	< 0.5	Linear drift

5.2 Validation and intercomparison results for product CO2_TAN_OCFP

The new TanSat XCO_2 data product CO2_TAN_OCFP v1.2 is part of GHG-CCI project Climate Research Data Package No. 7(CRDP#7), and consists of global, land-based nadir retrievals spanning 1st March 2017 – 23rd May 2018.

CO2_TAN_OCFP v1.2 corrected an issue with the quality-flags identified in CO2_TAN_OCFP v1.1. CO2_TAN_OCFP v1.1 was necessary to correct a mistake in the application of the bias-correction formula presented in **/Yang et al., 2020/** in the production of the CRDP#6 data release. Note that the errors described above do not occur in the CRDP#5 release (CO2_TAN_OCFP v1) and the validation performed with CRDP#5 is applicable to CRDP#7, and hence CO2_TAN_OCFP v1.2. The validation results presented here are therefore valid for v1 (CRDP#5) and v1.2 (CRDP#7).

The UoL core CO_2 ECV product CO2_TAN_OCFP is retrieved from calibrated TanSat SWIR/NIR spectra using the UoL full-physics retrieval algorithm **/Boesch et al., 2011/**. The TanSat L1 spectra are retrieved for all TCCON overpasses for the time period March 2017 to May 2018 and are evaluated against rigorously validated ground based TCCON values.

5.2.1 Detailed results

To assess the quality of CO2_TAN_OCFP observations against TCCON, OCFP (TanSat) soundings are matched to TCCON observations spatially and temporally. OCFP (TanSat) points are co-located with TCCON sites based on a quadrate latitude and longitude region around each TCCON site (in ±3° latitude/longitude box). Matching OCFP soundings with TCCON sites for time is a comparatively simple operation, selecting only those TCCON values whose observation time falls within ±1 hour of each TanSat sounding time. The average is taken of all TCCON points fitting these criteria for each OCFP sounding to provide the TCCON value against which to compare.

The co-location procedure matches 113,120 points for the CO2_TAN_OCFP product. The comparions for each TCCON site is shown in **Figure 5.2-1** and the statics (mean bias, standard deviation and Pearson correlation coefficient R) for each site is given in **Table 5.2-1**. The bias per site varies between -1.40 ppm and 1.57 ppm with a standard deviation of the per-site bias of 0.84 ppm. It is important to highlight that the number of data points and the temporal coverage varies greatly between sites.

The overall correlation between the TanSat and TCCON retrievals is given in **Figure 5.2-2**. We find a small mean overall bias of 0.19 ppm and an all-site Pearson correlation coefficient of 0.82 which details a good match of OCFP and TCCON pairs. The all-site RMSE (mean of the standard deviation per site) of Δ (TCCON- OCFP) is 1.78 ppm.

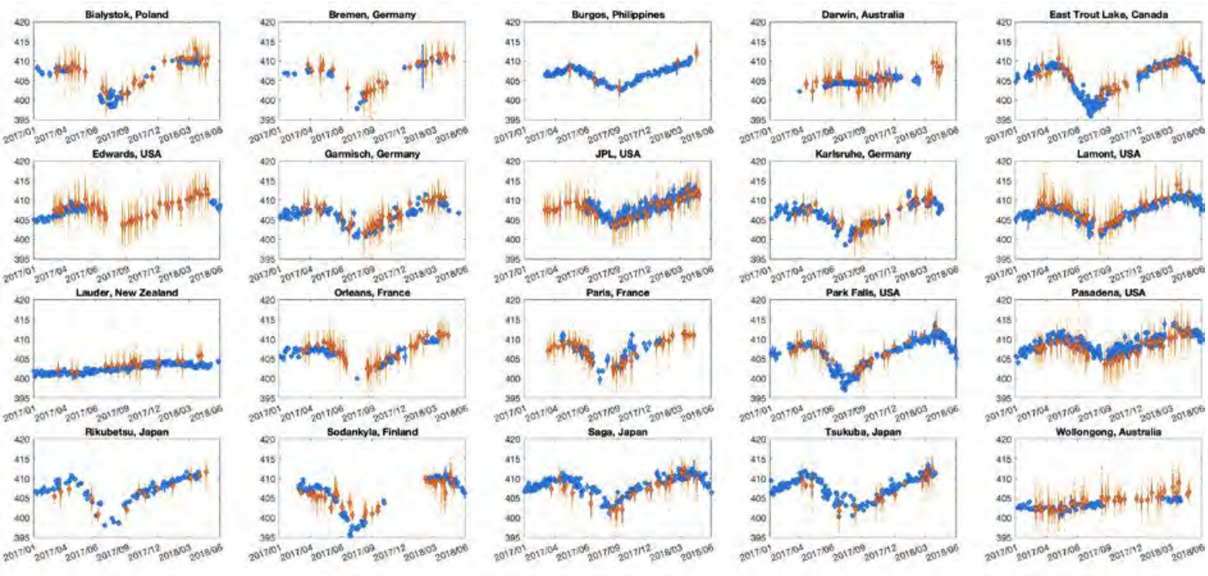


Figure 5.2-1: TanSat XCO_2 observations plotted with their corresponding paired TCCON mean (blue) for the overpass. Overview statistics for each site reference to **Table 5.2-1**.

Table 5-2-1: Overview of the estimated data quality as obtained from comparisons with TCCON ground-based reference observations per site. The bottom row details statistics for all sites, with all co-located points used for calculations. XCO_2 units is in ppm. The overall mean ∆ and σ∆ is calculated by averaging of site values and R is calculated by all individual measurements.

Site	**Mean ∆**	**σ∆**	***R***	**n obs.**
Bialystok, Poland	-0.92	1.68	0.65	3,292
Bremen, Germany	0.25	1.20	0.25	1,610
Burgos, Philippines	-0.08	2.22	0.32	310
Darwin, Australia	-0.64	2.05	-0.33	5,534
East Trout Lake, Canada	-0.17	1.26	0.90	11,923
Edwards, USA	-1.40	1.96	0.55	2,763
Garmisch, Germany	-0.32	1.67	0.67	3,704
JPL, USA	1.17	2.07	0.81	15,209
Karlsruhe, Germany	-0.29	1.62	0.84	3,089
Lamont, USA	-0.35	1.35	0.86	18,274
Lauder, New Zealand	-1.31	1.88	0.72	2,999
Orléans, France	-0.66	1.46	0.18	2,243
Paris, France	-0.08	1.40	0.76	1,503
Park Falls, USA	-0.35	1.45	0.89	13,231
Pasadena, USA	1.57	2.47	0.65	12,807
Rikubetsu, Japan	0.54	1.27	0.84	1,473
Sodankylä Finland	-1.18	2.19	0.93	6,482
Saga, Japan	0.69	1.99	0.77	4,033
Tsukuba, Japan	0.94	2.46	0.79	866
Wollongong, Australia	-1.15	1.93	0.73	1,775
Overall	0.19	1.78	0.82	113,120

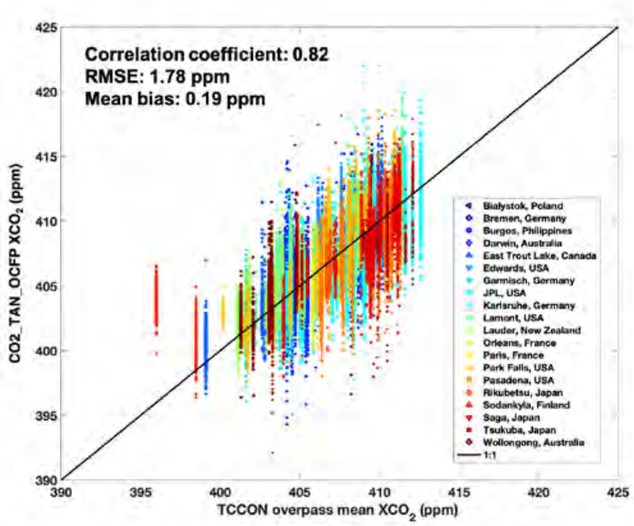


Figure 5.2-2: Correlation plot between all 113,120 co-located CO2_TAN_OCFP and TCCON XCO_2 pairs coloured by site.

The random error is assessed by comparing the overpass-mean reported uncertainty for an overapss over a TCCON site to the standard deviation of the TCCON–OCFP pairs for each overpass. **Figure 5.2-3** shows that the reported uncertainties are between 0.78 ppm (Lamont, U.S.A.) and 4.34 ppm (East Trout Lake, Canada). There is a relatively large spread of the data points with some clear outliers where the observed scatter is largely overestimated. We find that these overestimated errors are correlated with very low surface albedo of the CO_2 band and subsequently low information content for CO_2 so that the retrieved results remain close to the *a priori* values. The slope between the observed scatter between TanSat and TCCON retrievals and the reported uncertainties is 0.96.

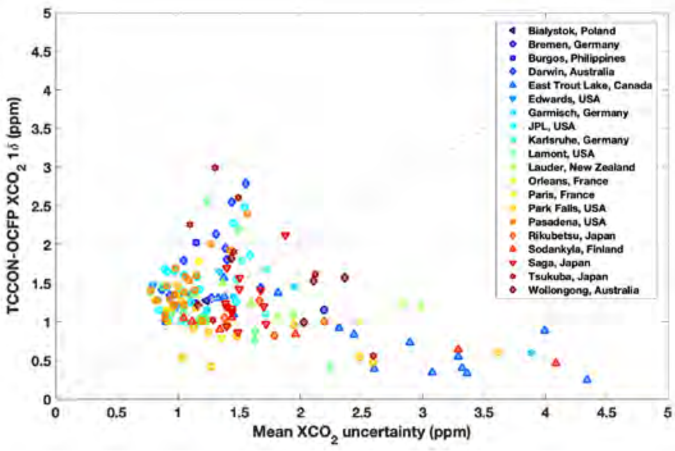


Figure 5.2-3: Correlation plot of the TCCON–OCFP Δ standard deviation per TCCON overpass and the reported overpass-mean a posteriori retrieval error for different TCCON sites.

5.2.2 Summary

The result of the validation of the CO2_TAN_OCFP v1.2 dataset is given in **Table 5.2-2** and compared to the requirement. The mean estimate of the single-measurement precision is 1.78 ppm which exceeds the goal requirement but is within the baseline requirement of 3 ppm. The reported uncertainties agree in average with the observed scatter of the data when compared to TCCON. The mean, global bias of the TanSat XCO_2 retrieval is 0.19 ppm with a relative accuracy of 0.84 ppm which is slightly larger than the requirement of 0.5 ppm. We have not assessed the spatio-temporal bias or the drift due to the short time period covered by the CO2_TAN_OCFP dataset.

Table 5.2-2: Summary validation of product CO2_TAN_OCFP v1.2 by the data provider using TCCON ground-based reference data.

Product Quality Summary Table for Product: CO2_TAN_OCFP Level: 2, Version: v1.2, Time period covered: 3.2017 – 5.2018 Assessment: Data Provider (DP)			
Parameter [unit]	**Achieved performance**	**Requirement**	**Comments**
Single measurement precision (1-sigma) in [ppm]	1.78	< 8 (T) < 3 (B) < 1 (G)	Computed as standard deviation of the difference to TCCON
Uncertainty ratio [-]: Ratio reported uncertainty to standard deviation of satellite-TCCON difference	0.96	-	No requirement but value close to unity expected for a high quality data product with reliable reported uncertainty.
Mean bias (global offset) [ppm]	0.19	-	No requirement but value close to zero expected for a high quality data product.
Accuracy: Relative systematic error [ppm]	Spatial: 0.84 Spatio-temporal: Not evaluated	< 0.5	Spatial: Computed as standard deviation of the biases at the various TCCON sites. Spatio-temporal: As "Spatial" but also considering seasonal biases.
Stability: Drift [ppm/year]	Not evaluated	< 0.5	Linear drift

5.3 Validation and intercomparison results for product CO2_GO2_SRFP

The CO2_GO2_SRFP product is retrieved from GOSAT-2 TANSO-FTS SWIR spectra using the RemoTeC algorithm that has been jointly developed by SRON and KIT **/Butz et al., 2011; Schepers et al., 2012/**. The retrievals are performed globally for the time period between February 2019 and May 2020 and are evaluated against ground based TCCON observations.

5.3.1 Detailed results

To assess the quality of SRFP retrieval XCO_2 observations against TCCON values, SRFP soundings are matched to TCCON observations spatially and temporally. GOSAT-2 observations are co-located with TCCON sites based on a square latitude and longitude region around each TCCON site (in ±2.5º latitude/longitude box). For the temporal co-location we select only the TCCON measurements whose observation time falls within ±2 hour of each GOSAT-2 observation time. The TCCON observations that match these criteria are averaged for each individual GOSAT-2 observation.

We co-located GOSAT-2 and TCCON measurements with a maximum time difference of 2.5h, a maximum distance of 300 km in both longitudinal and latitudinal directions. In cases of multiple TCCON measurements of the same site collocating with a GOSAT-2 sounding, we averaged the TCCON measurements. For 13 TCCON sites used for the validation, in total we found about 1609 collocations during Feb. 2019 - May 2020 for the CO2_GO2_SRFP product.

The comparions for each TCCON site is shown in **Figure 5.3-1**. The statistics (mean bias, standard deviation) for each site are given in **Table 5.3-1**. The spatial accuracy (standard deviation site biases) is 0.9 ppm. The most notable outlier is Lauder, with a remaining bias of 2.10 ppm, which is potentially due to the low number of co-locations. The lauder time-series shows that the TCCON values have an overall offset compared to the GOSAT-2 observations.
The overall correlation between the GOSAT-2 and TCCON retrievals is given in **Figure 5.3-2**. The mean bias (global offset) amounts to 0.01 ppm. The standard deviation of the site biases (spatial accuracy or station-to-station variability) is 1.03 ppm for land and 1.19 ppm for sun-glint observations. The single measurement precision of GOSAT-2 compared to TCCON amounts to 2.26 ppm.

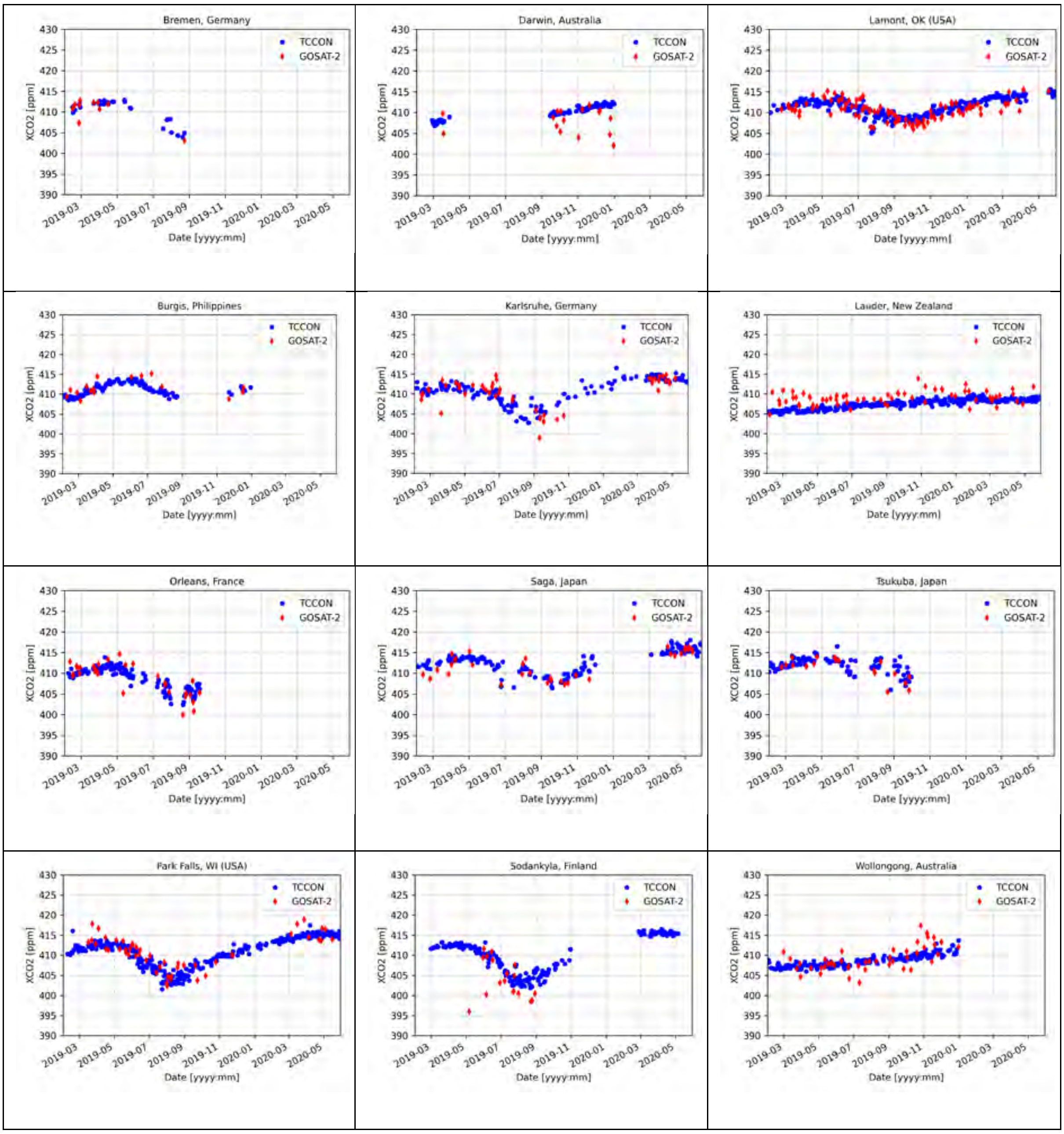


Figure 5-3-1: GOSAT-2 XCO_2 (CO2_GO2_SRFP, red) with co-located TCCON (blue) measurements at 12 TCCON stations used for the validation for the period of February to May 2020.

Table 5.3-1: Overview of the SRFP/RemoTeC XCO_2 validation with TCCON (after bias correction).

TCCON site	Number of co-locations [-]	Mean difference [ppb]	Standard deviation of difference [ppb]
Burgis, Philippines	33	0.774	1.824
Bremen, Germany	41	0.176	2.012
Darwin, Australia	32	-2.029	3.302
Karlsruhe, Germany	171	0.214	2.265
Lamont, USA	475	-0.304	1.753
Lauder, New Zealand	134	1.762	1.946
Orleans, France	153	0.429	2.226
Park Falls, USA	121	0.262	2.457
Saga, Japan	145	-0.467	1.839
Sodankyla, Finland	39	-1.994	3.583
Tsukuba, Japan	139	-0.647	2.133
Wollongong, Australia	96	0.255	2.510
All observations	1579	-0.011	2.264

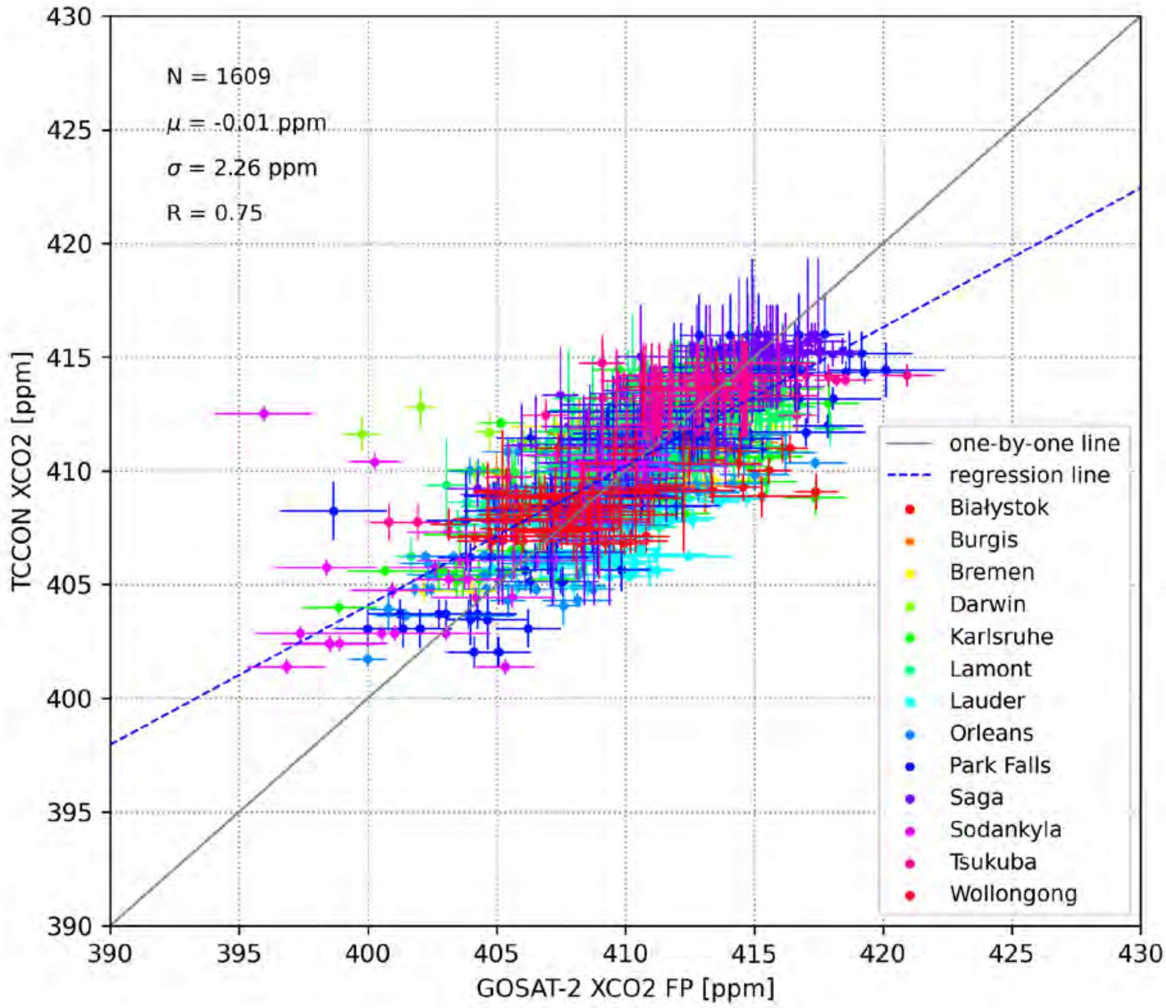


Figure 5-3-2:Validation of single soundings of FP-CO2 with collocated TCCON measurements at all TCCON sites for the period Feb. 2019 - May 2020. Numbers in the figures: μ = bias, i.e., average of the difference; σ = single measurement precision, i.e., standard deviation of the difference; N = number of co-locations; R = Pearson correlation coefficient.

The error that comes out of the RemoTeC retrieval is just a purely statistical error on the radiance that has been propagated through the entire retrieval chain.

In order to more accurately estimate the actual random error on the GOSAT-2 sounding, we applied the following procedure to obtain a scaling factor with which to scale our statistical error. We take the absolute difference of every co-located sounding and divide it by the retrieved statistical error corresponding to that sounding. We then average these values to obtain the average scaling factor by which to scale the retrieved statistical error to obtain a more correct estimate of the random error.

Based on the analysis, we obtain the following scaling factors for the SRFP XCO_2 product, 2.27 for normal mode and 2.05 for sunglint mode and an uncertainty ratio of 0.44 (0.50 sunglint).

5.3.2 Summary

The result of the validation of the CO2_GO2_SRFP dataset is given in **Table 5.3-2** and compared to the requirement. The mean estimate of the single-measurement precision is 2.26 ppm which exceeds the goal requirement but is within the breakthrough requirement of 3 ppm. The uncertainties provided by RemoTeC agree on average with the observed scatter of the data when compared to TCCON. The mean (global bias) of the GOSAT-2 XCO_2 retrieval is 0.01 ppm with a relative accuracy of 0.9 ppm which is larger than the requirement of 0.5 ppm. This can be attributed to the relatively small number of co-locations and short period of comparison. For comparison, we find a value of 0.7 ppm for GOSAT-1 over a multi-year period (2009-2019), while for the same 2019 period it is 1.0 ppm **/E3UBv1.1, 2020/**. We have not assessed the spatio-temporal bias or the drift due to the limited time period covered by the CO2_GO2_SRFP dataset.

Table 5.3-2: Summary validation of product CO2_GO2_SRFP by the data provider using TCCON ground-based reference data.

Product Quality Summary Table for Product: CO2_GO2_SRFP Level: 2, Version: v2, Time period covered: 2.2019 – 8.2020 Assessment: Data Provider (DP)			
Parameter [unit]	**Achieved performance**	**Requirement**	**Comments**
Single measurement precision (1-sigma) in [ppm]	2.26	< 8 (T) < 3 (B) < 1 (G)	Computed as standard deviation of the difference to TCCON
Uncertainty ratio [-]: Ratio reported uncertainty to standard deviation of satellite-TCCON difference	0.44 (0.50 sunglint)	-	No requirement but value close to unity expected for a high quality data product with reliable reported uncertainty.
Mean bias (global offset) [ppm]	-0.01	-	No requirement but value close to zero expected for a high quality data product.
Accuracy: Relative systematic error [ppm]	Spatial: 1.0 Spatio-temporal: Not evaluated	< 0.5	Spatial: Computed as standard deviation of the biases at the various TCCON sites. Spatio-temporal: As “Spatial” but also considering seasonal biases.
Stability: Drift [ppm/year]	Not evaluated	< 0.5	Linear drift

5.4 Validation and intercomparison results for product CH4_S5P_WFMD

Validation results for XCH_4 retrieved from TROPOMI with the WFMDv1.5 algorithm **/Schneising et al., 2019/** are summarised in this section. The validation data set is the GGG2014 collection of the Total Carbon Column Observing Network (TCCON) (available from https://tccondata.org/). To ensure comparability, all TCCON sites use similar instrumentation (Bruker IFS 125HR) and a common retrieval algorithm. The TCCON data are tied to the WMO trace gas scale using airborne in situ measurements applying individual scaling factors for each species. The estimated TCCON accuracy (1σ) is about 3.5 ppb for XCH_4. From the validation with TCCON data at 26 TCCON sites, realistic error estimates of the satellite data are provided.

To compare the satellite data with TCCON quantitatively, it has to be taken into account that the sensitivities of the instruments differ from each other and that individual apriori profiles are used to determine the best estimate of the true atmospheric state, respectively. The first step is to correct for the apriori contribution to the smoothing equation by adjusting the measurements for a common apriori. Here we use the TCCON prior as the common apriori profile for all measurements:

$$\hat{c}_{adj} = \hat{c} + \frac{1}{m_0}\sum_l m_l\,(1 - A_l)(x^l_{a,T} - x^l_a)$$

In this equation, $\hat{c}$ represents the originally retrieved TROPOMI column-averaged dry air mole fraction, l is the index of the vertical layer, A_l the corresponding column averaging kernel of the TROPOMI algorithm, $\boldsymbol{x}_a$ and $\boldsymbol{x}_{a,T}$ the TROPOMI and TCCON apriori dry air mole fraction profiles. m_l is the mass of dry air determined from the dry air pressure difference between the upper and lower boundary of layer l and $m_0 = \sum_l m_l$ is the total mass of dry air. To minimise the smoothing error introduced by the averaging kernels we do not compare $\hat{c}_{adj}$ directly with the retrieved TCCON mole fractions $\hat{c}_T$ but rather with the adjusted expression

$$\hat{c}_{T,adj} = c_{a,T} + \left(\frac{\hat{c}_T}{c_{a,T}} - 1\right)\frac{1}{m_0}\sum_l m_l A_l x^l_{a,T}$$

Thereby, $c_{a,T}$ represents the TCCON apriori column-averaged dry air mole fraction associated with the apriori profile $\boldsymbol{x}_{a,T}$.

5.4.1 Detailed results

For the comparison a set of collocation criteria has been specified. The representativity is maximised by as strict as possible criteria while concurrently ensuring sufficient data for a sound and stable comparison. This trade-off is resolved by the following selection. The spatial collocation criterion requires the satellite measurements to lie within a radius of 100 km around the TCCON site and that the altitude difference is smaller than 250 m. The temporal collocation criterion is set to ±2 hours. For each satellite measurement within the collocation radius, all TCCON data meeting the temporal collocation criterion are averaged to obtain a unique satellite-TCCON data pair. This approach is consistent with the well-established methods used in previous GHG-CCI PVIRs.

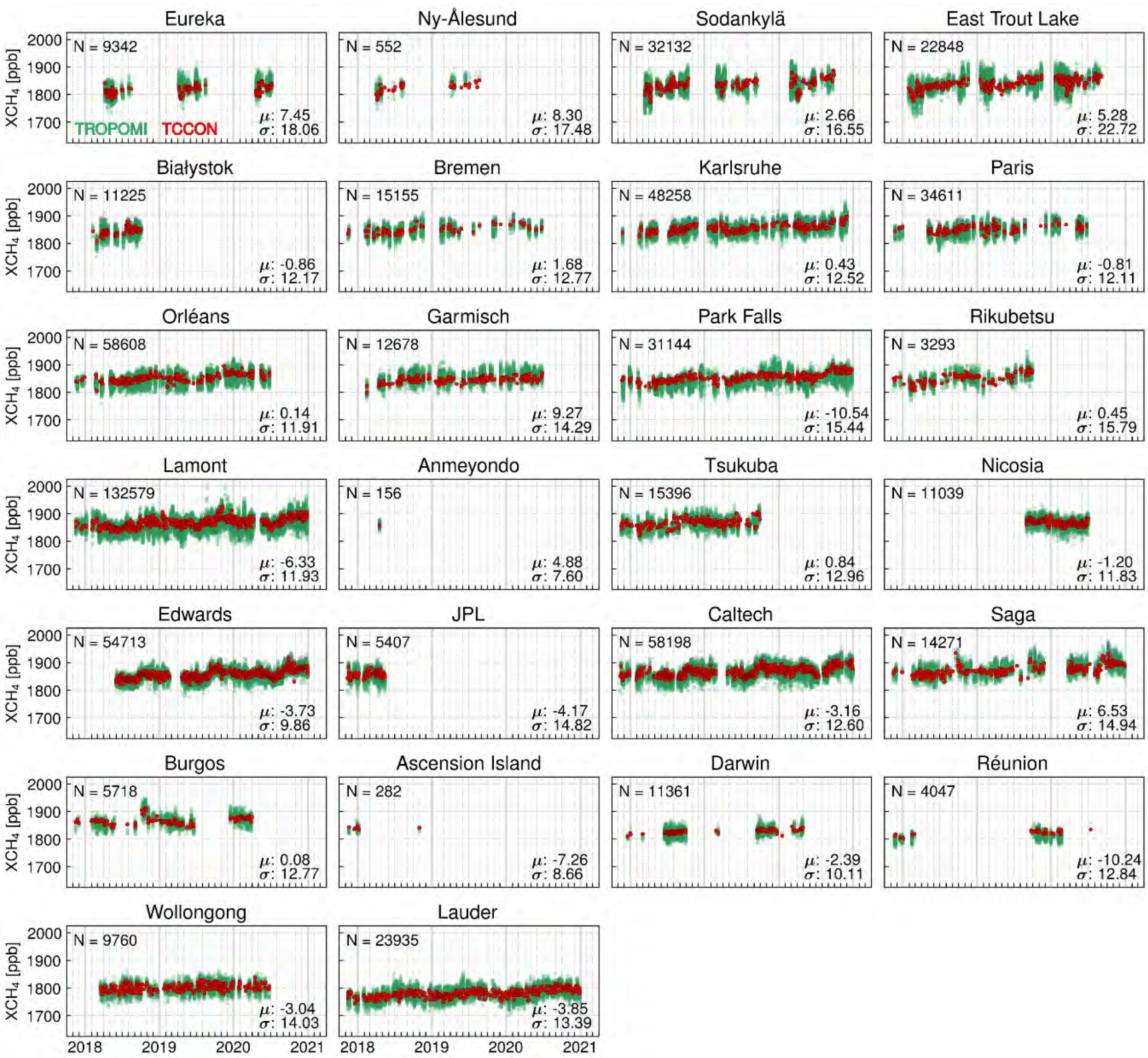


Figure 5.4-1: Comparison of the TROPOMI/WFMD v1.5 XCH_4 time series (green) with ground-based measurements from the TCCON (red). For each site, $\boldsymbol{N}$ is the number of collocations, $\boldsymbol{\mu}$ corresponds to the mean bias and $\boldsymbol{\sigma}$ to the scatter of the satellite data relative to TCCON in ppb.

The validation results are summarised in **Figure 5.4-1** including the mean bias μ and the scatter σ relative to TCCON for each site. As a consequence of the altitude representativity criterion, there are not enough collocations for a robust comparison at the mountain sites Zugspitze and Izaña. The parameter σ is estimated from Huber's Proposal-2 M-estimator, which is a well-established estimator of location and scale being robust against outliers of a normal distribution. This is an appropriate choice and preferred over the standard deviation, because one is interested in the actual single measurement precision without distortion of the results by a few outliers, which are rather attributed to systematic errors, e.g. due to residual clouds. As a consequence, outliers are fully included in the computation of the systematic error but get lower weight in the robust determination of the random error, which is interpreted as a measure of the repeatability of measurements.

It is also checked whether the respective site biases are sensitive to the selection of the spatial collocation radius, which is an indication of sources within the satellite collocation area with only marginal influence on the TCCON measurements itself. A considerable sensitivity was found for XCH_4 at Edwards. The collocation region intersects oil production areas in California's Central Valley (in contrast to Caltech and JPL, see **/Schneising et al., 2019/**) as well as the South Coast Air Basin (SoCAB), which has a well-known methane enhancement. As such nearby sources limit the representativity of affected satellite measurements, the collocation radius is reduced to 50 km for Edwards.

The results for the individual sites are condensed to the following parameters for the overall quality assessment of the satellite data: the global offset is defined as the mean of the local biases at the individual sites, the random error is the global scatter of the differences to TCCON after subtraction of the respective regional biases, and the spatial systematic error is the standard deviation of the local offsets relative to TCCON at the individual sites as a measure of the station-to-station biases. For XCH_4 the global offset amounts to -0.37 ppb, the random error is 12.85 ppb (14.13 ppb when using the standard deviation instead of Huber's Proposal-2 M-estimator), and the spatial systematic error is given by 5.17 ppb. The seasonal systematic error is defined as the standard deviation of the four overall seasonal offsets (using all sites combined after subtraction of the respective local offsets) relative to TCCON and amounts to 0.56 ppb. The spatio-temporal systematic error (defined as the the root-sum-square of the spatial and seasonal systematic errors) amounts to 5.20 ppb, which is on the order of the estimated (station-to-station) accuracy of the TCCON of about 3.5 ppb.

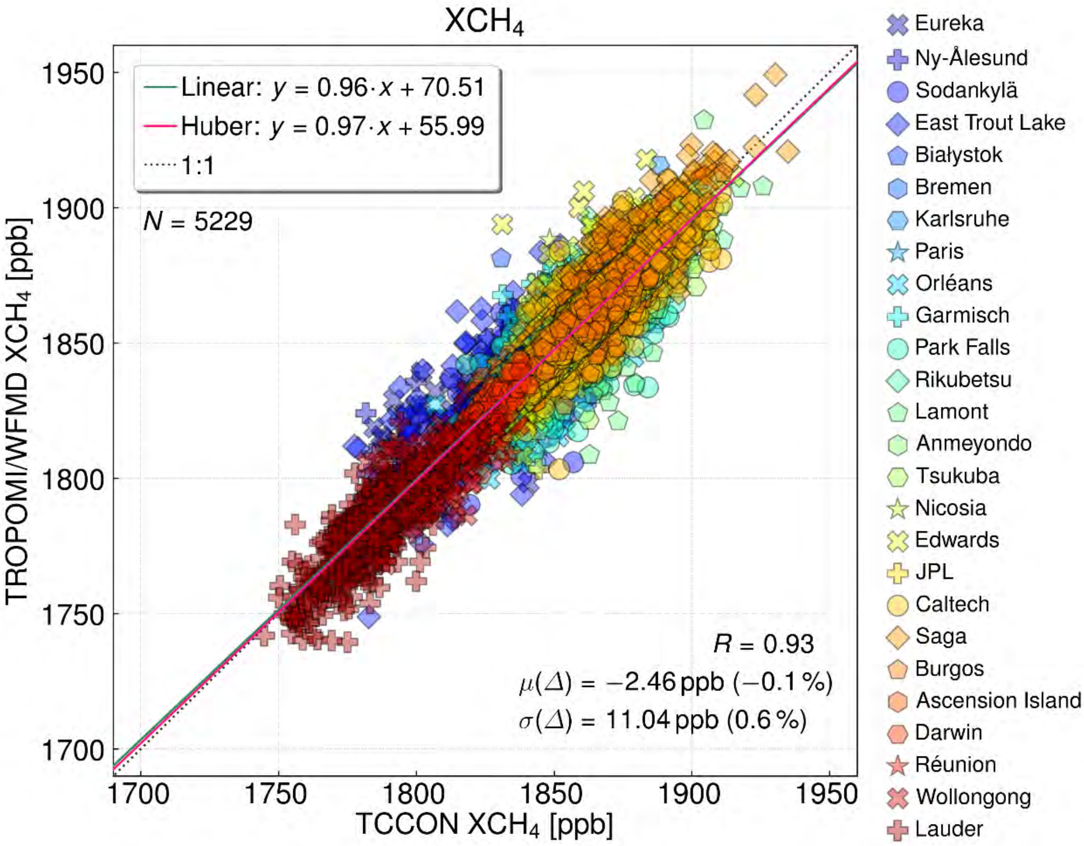


Figure 5.4-2: Comparison of the TROPOMI/WFMD data to the TCCON based on daily means. Specified are the linear regression results and the correlation of the data sets, as well as the mean and standard deviation of the difference. To analyse the impact of outliers, the regression is also performed for the Huber linear regression model, which is robust to outliers.

To further analyse how well the real temporal and spatial variations are captured by the TROPOMI data, **Figure 5.4-2** shows a comparison to TCCON based on daily means for days with more than three collocations. The obvious linear relationship with a high correlation of $R = 0.93$ underlines the typical good agreement of the satellite and validation data.

There are a few outliers where the satellite values are considerably lower than the TCCON values. These occasional instances are not site specific and can probably be ascribed to days with residual or partial cloud cover interfering with the satellite retrievals. Outliers at high latitude sites may be attributable to Arctic polar vortex air potentially causing the following related issues: associated fronts of different air masses may complicate the identification of collocations near the vortex edge and/or the stratospheric part of the methane profiles may be largely affected by the polar vortex leading to a considerable deviation from the assumed apriori profile shapes. It is verified that the impact of outliers on the regression is marginal by repeating the fit with the Huber linear regression model, which is robust to outliers and provides similar results to the standard linear regression here.

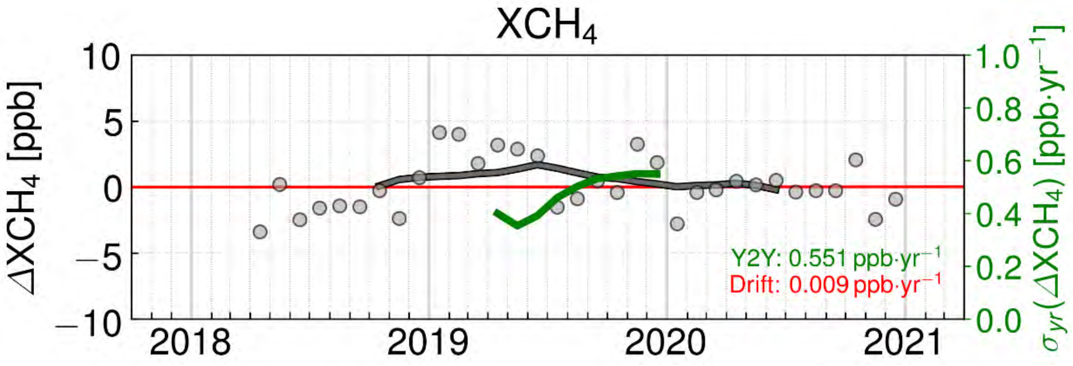


Figure 5.4-3: Long-term drift and year-to-year stability at TCCON sites.

To analyse the stability, we use comparisons with the TCCON since the start of the routine operations phase of Sentinel-5P to have sufficient data coverage. To assess the long-term drift stability, a robust Huber regression of the monthly mean differences relative to the reference (using all data combined after subtraction of the respective regional offsets) with time is used. The resulting stability estimate is 0.01 ppb/year (see red straight line in **Figure 5.4-3**).

The year-to-year stability allowing to detect potential jumps in the time series is defined in the following way: The one-year moving average of the differences relative to the reference (grey curve in **Figure 5.4-3**) is generated. For a given point in time t, let $\sigma_{yr}(t)$ be defined as the standard deviation of this deseasonalised difference within a one-year window around t (green curve in **Figure 5.4-3**). The year-to-year stability is then defined as the maximum of $\sigma_{yr}(t)$ over time, which amounts to 0.55 ppb/year here. Due to the moving average and the one-year moving standard deviation procedure, the green curve loses one year of data at the beginning and end of the time series. A longer time series of satellite data will allow a more sound and stable estimation of the year-to-year stability in the future.

The reported uncertainty of TROPOMI/WFMD v1.5 XCH_4 is validated based on a comparison to the measured scatter relative to the TCCON. After dividing up the reported uncertainties in equal sized bins of about 30000 measurements each, a robust regression provides the results shown in **Figure 5.4-4** (neglecting the random and systematic errors of the TCCON measurements) confirming that the reported estimates are realistic: The uncertainty ratio (reported uncertainty to measured scatter) is 1.00, indicating a reliable estimation of the measurement uncertainties.

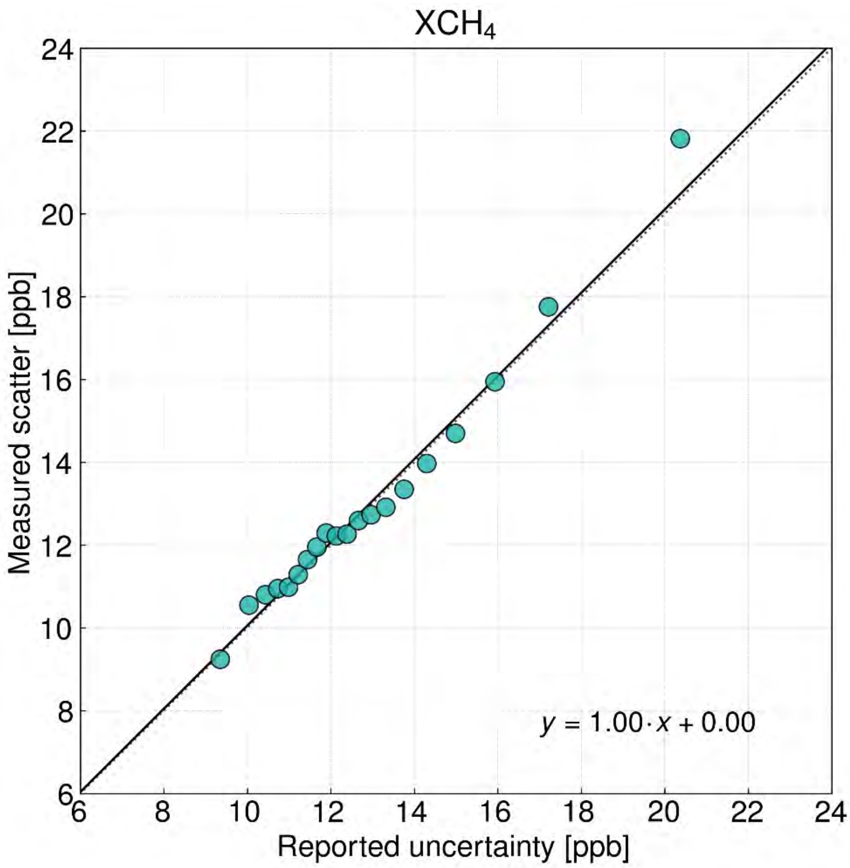


Figure 5.4-4: Comparison of the reported uncertainty of TROPOMI/WFMD v1.5 XCH_4 with the measured scatter relative to the TCCON after dividing up the reported uncertainties in equal sized bins.

5.4.2 Summary

In summary, the natural XCH_4 variations are well captured by the satellite data. We find a single measurement precision of the TROPOMI data of about 0.7%, while the station-to-station accuracy of the satellite data (0.3%) is comparable to the TCCON.

The single measurement precision is below the breakthrough requirement and the uncertainty ratio is close to 1. The accuracy also complies with the requirements and the mean bias is close to zero. The stability is well below the required value. **Table 5.4-1** presents an overview of the estimated data quality as obtained from comparisons with TCCON ground-based reference observations.

Table 5.4-1: Summary validation of product CH4_S5P_WFMD by the data provider using TCCON ground-based reference data.

Product Quality Summary Table for Product: CH4_S5P_WFMD Level: 2, Version: v1.5, Time period covered: 11.2017 – 12.2020 Assessment: Data Provider (DP)			
Parameter [unit]	**Achieved performance**	**Requirement**	**Comments**
Single measurement precision (1-sigma) in [ppb]	12.85	< 34 (T) < 17 (B) < 9 (G)	Computed as standard deviation of the difference to TCCON
Uncertainty ratio [-]: 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 with reliable reported uncertainty.
Mean bias (global offset) [ppb]	-0.37	-	No requirement but value close to zero expected for a high quality data product.
Accuracy: Relative systematic error [ppb]	Spatial: 5.17 Spatio-temporal: 5.20	< 10	Spatial: Computed as standard deviation of the biases at the various TCCON sites. Spatio-temporal: As "Spatial" but also considering seasonal biases.
Stability: Drift [ppb/year]	0.01	< 3	Linear drift

5.5 Validation and intercomparison results for product CH4_GO2_SRFP

The CH4_GO2_SRFP product is retrieved from GOSAT-2 TANSO-FTS SWIR spectra using the RemoTeC algorithm that has been jointly developed by SRON and KIT **/Butz et al., 2011; Schepers et al., 2012/**. The retrievals are performed globally for the time period between February and October 2019 and are evaluated against ground based TCCON observations.

5.5.1 Detailed results

To assess the quality of SRFP retrieval XCH_4 observations against ground based TCCON values, SRFP soundings are matched to TCCON observations spatially and temporally. GOSAT-2 observations are co-located with TCCON sites based on a square latitude and longitude region around each TCCON site (in ±2.5º latitude/longitude box). For the temporal co-location we select only the TCCON measurements whose observation time falls within ±2 hour of each GOSAT-2 observation time. The TCCON observations that match these criteria are averaged for each individual GOSAT-2 observation.

The co-location procedure matches 1587 points for the CH4_GO2_SRFP product. The comparions for each TCCON site is shown in **Figure 5.5-1**. The statistics (mean bias, standard deviation and Pearson correlation coefficient R) for each site are given in **Table 5.5-1**. The bias per site varies between -3.63 ppb at East Trout Lake to 4.18 ppb in Garmisch. The standard deviation of the station-to-station bias is 2.39 ppb, which mostly follows from the large bias found in the few Garmisch and East Trout Lake observations and the overall small number of compared values. Because of the short time period the number of data points and the temporal coverage varies greatly between sites.

The overall correlation between the GOSAT-2 and TCCON retrievals is given in **Figure 5.5-2**. We find a small mean overall bias of 0.09 ppb and an all-site Pearson correlation coefficient of 0.78 which, even for the small number of co-locations, points to a good comparison of GOSAT-2 and TCCON pairs. The all-site RMSE (mean of the standard deviation per site) of Δ (TCCON- GOSAT-2) is 13.03 ppb.

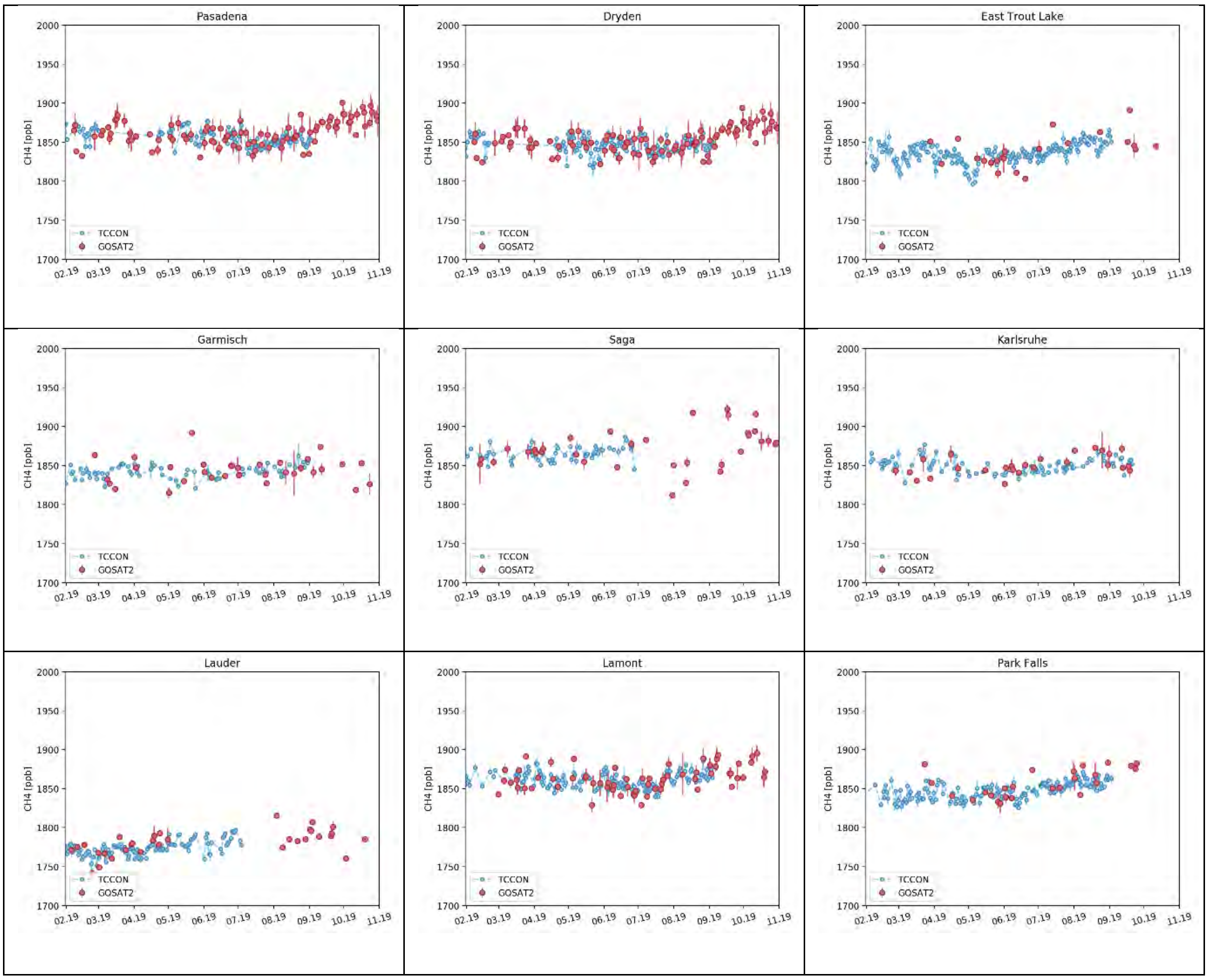


Figure 5.5-1: GOSAT-2 XCH_4 (CH4_GO2_SRFP, red) with co-located TCCON (blue) measurements at nine TCCON stations used for the validation for the period of February to October 2019.

Table 5.5-1: Overview of the estimated data quality as obtained from comparisons with TCCON ground-based reference observations per site. The bottom row details statistics for all sites, with all co-located points used for calculations. XCH_4 units are in ppb. The overall mean ∆ and σ∆ is calculated by averaging of site values and R is calculated by all individual measurements. The mean of site means $\bar{\mu}$ and spatial accuracy $\sigma_{\bar{\mu}}$ are calculated by taking the mean and standard deviation of the site means. The mean standard deviation $\bar{\sigma}$ and standard deviation of the standard deviations $\sigma_{\bar{\sigma}}$ are calculated by taking the mean and the standard deviation of the site standard deviations.

Site	Mean ∆	σ∆	*R*	n obs.
Pasadena, USA	-2.03	15.70	0.47	337
Dryden, USA	0.22	15.36	0.40	448
East Trout Lake, Canada	-3.63	17.31	0.38	62
Garmisch, Germany	4.18	8.41	0.75	28
Saga, Japan	2.18	12.53	0.37	98
Karlsruhe, Germany	-1.31	11.17	0.58	115
Lauder, New Zealand	2.63	10.41	0.74	52
Lamont, USA	1.12	13.18	0.58	344
Park Falls, USA	2.43	13.24	0.64	103
All observations	0.09	14.36	0.78	1587
	$\bar{\mu}$	$\sigma_{\bar{\mu}}$	$\bar{\sigma}$	$\sigma_{\bar{\sigma}}$
Mean of sites	0.64	2.39	13.03	2.64

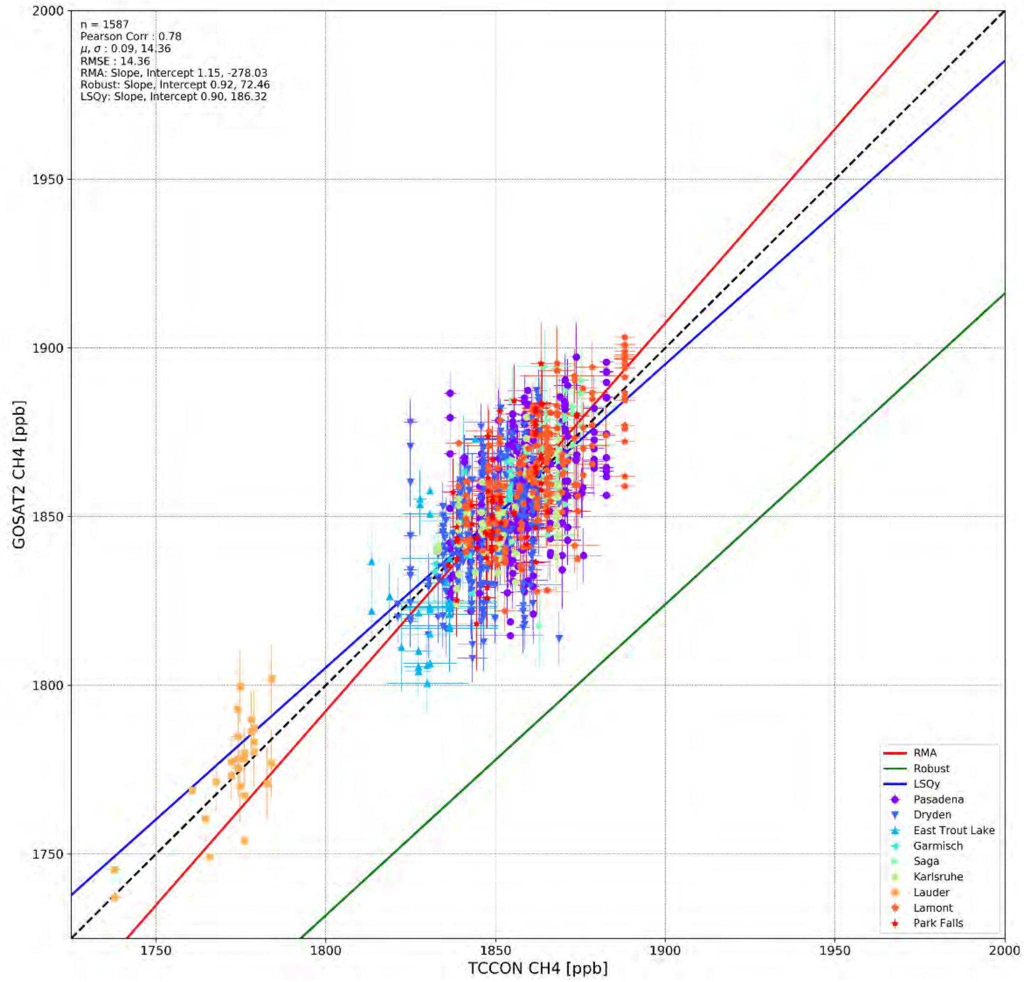


Figure 5.5-2: Correlation plot between all 1587 co-located CH4_GO2_SRFP and TCCON XCH_4 pairs coloured by site.

The error that comes out of the RemoTeC retrieval is just a purely statistical error on the radiance that has been propagated through the entire retrieval chain. In order to more accurately estimate the actual random error on the GOSAT-2 sounding, we applied the following procedure to obtain a scaling factor with which to scale our statistical error. We take the absolute difference of every co-located sounding and divide it by the retrieved statistical error corresponding to that sounding. We then average these values to obtain the average scaling factor by which to scale the retrieved statistical error to obtain a more correct estimate of the random error.

Based on the analysis, we obtain the following scaling factors for the SRFP XCH_4 product, 1.44 for the normal mode and 1.38 for the sunglint mode. Subsequently, we calculate the uncertainty ratio which is defined as the ratio of the mean value of the reported uncertainty and the standard deviation of the difference to TCCON. We obtain uncertainty ratios of 0.69 for the normal mode and 0.72 for the sunglint mode.

5.5.2 Summary

The result of the validation of the CH4_GO2_SRFP dataset is given in **Table 5.5-2** and compared to the requirement. The mean estimate of the single-measurement precision is 14.36 ppb which exceeds the goal requirement but is within the breakthrough requirement of 17 ppb. The uncertainties provided by RemoTeC agree on average with the observed scatter of the data when compared to TCCON. The mean, global bias of the GOSAT-2 XCH_4 retrieval is 0.09 ppb with a relative accuracy of 2.39 ppb which is smaller than the requirement of 10 ppb. We have not assessed the spatio-temporal bias or the drift due to the limited time period covered by the CH4_GO2_SRFP dataset.

Table 5.5-2: Summary validation of product CH4_GO2_SRFP by the data provider using TCCON ground-based reference data.

Product Quality Summary Table for Product: CH4_GO2_SRFP Level: 2, Version: v2, Time period covered: 2.2019 – 8.2020 Assessment: Data Provider (DP)			
Parameter [unit]	**Achieved performance**	**Requirement**	**Comments**
Single measurement precision (1-sigma) in [ppb]	14.36	< 34 (T) < 17 (B) < 9 (G)	Computed as standard deviation of the difference to TCCON
Uncertainty ratio [-]: Ratio reported uncertainty to standard deviation of satellite-TCCON difference	0.69 (0.72 glint)	-	No requirement but value close to unity expected for a high quality data product with reliable reported uncertainty.
Mean bias (global offset) [ppb]	0.09	-	No requirement but value close to zero expected for a high quality data product.
Accuracy: Relative systematic error [ppb]	Spatial: 2.39 Spatio-temporal: Not evaluated	< 10	Spatial: Computed as standard deviation of the biases at the various TCCON sites. Spatio-temporal: As "Spatial" but also considering seasonal biases.
Stability: Drift [ppb/year]	Not evaluated (1-sigma)	< 3	Linear drift

5.6 Validation and intercomparison results for product CH4_GO2_SRPR

The CH4_GO2_SRPR product is retrieved from GOSAT-2 TANSO-FTS SWIR spectra using the RemoTeC algorithm that has been jointly developed by SRON and KIT **/Butz et al., 2011; Schepers et al., 2012/**. The retrievals are performed globally for the time period between February and October 2019 and are evaluated against ground based TCCON observations.

5.6.1 Detailed results

To assess the quality of SRPR retrieval XCH_4 observations against ground based TCCON values, SRPR soundings are matched to TCCON observations spatially and temporally. GOSAT-2 observations are co-located with TCCON sites based on a square latitude and longitude region around each TCCON site (in ±2.5º latitude/longitude box). For the temporal co-location we select only the TCCON measurements whose observation time falls within ±2 hour of each GOSAT-2 observation time. The TCCON observations that match these criteria are averaged for each individual GOSAT-2 observation.

The co-location procedure matches 2642 points for the CH4_GO2_SRPR product. The comparions for each TCCON site is shown in **Figure 5.6-1**. The statistics (mean bias, standard deviation and Pearson correlation coefficient R) for each site is given in **Table 5.6-1**. The bias per site varies between -5.61 ppb for Pasadena to 8.48 ppb in Saga. The standard deviation of the station-to-station bias is 4.24 ppb, which mostly follows from the opposite large bias found in the Pasadena and Saga observations and the overall small number of compared values. Because of the short time period the number of data points and the temporal coverage varies greatly between sites.

The overall correlation between the GOSAT-2 and TCCON retrievals is given in **Figure 5.6-2**. We find a small mean overall bias of 0.10 ppb and an all-site Pearson correlation coefficient of 0.76 which, even for the small number of co-locations, points to a good comparison of GOSAT-2 and TCCON pairs. The all-site RMSE (mean of the standard deviation per site) of Δ (TCCON- GOSAT-2) is 15.32 ppb.

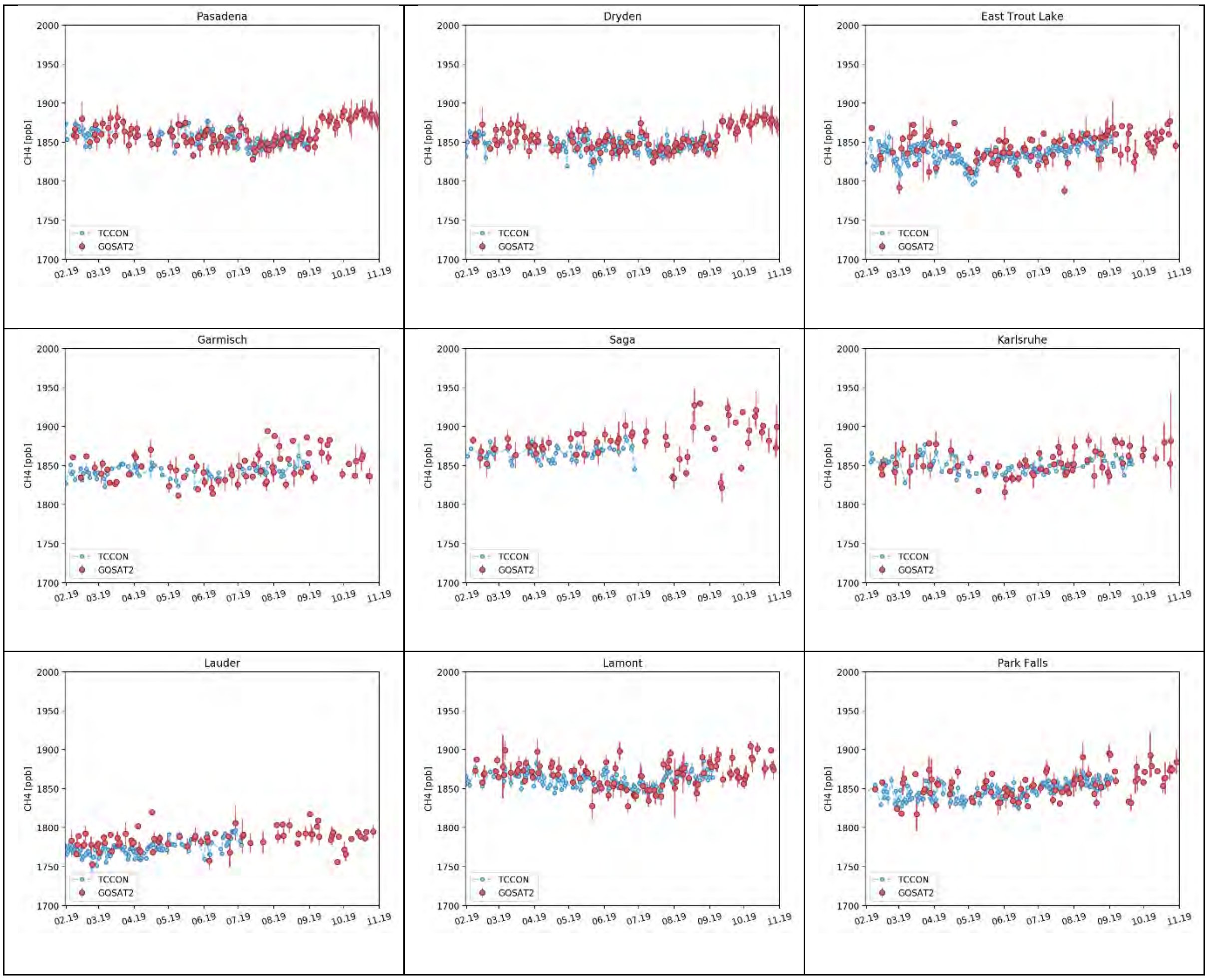


Figure 5.6-1: GOSAT-2 XCH_4 (CH4_GO2_SRPR, red) with co-located TCCON (blue) measurements at nine TCCON stations used for the validation for the period of February to October 2019.

Table 5.6-1: Overview of the estimated data quality as obtained from comparisons with TCCON ground-based reference observations per site. The bottom row details statistics for all sites, with all co-located points used for calculations. XCH_4 units are in ppb. The overall mean Δ and $\sigma\Delta$ is calculated by averaging of site values and R is calculated by all individual measurements. The mean of site means $\bar{\mu}$ and spatial accuracy $\sigma_{\bar{\mu}}$ are calculated by taking the mean and standard deviation of the site means. The mean standard deviation $\bar{\sigma}$ and standard deviation of the standard deviations $\sigma_{\bar{\sigma}}$ are calculated by taking the mean and the standard deviation of the site standard deviations.

Site	Mean Δ	$\sigma\Delta$	*R*	n obs.
Pasadena, USA	-5.61	13.12	0.53	545
Dryden, USA	-0.31	14.59	0.37	732
East Trout Lake, Canada	4.99	17.63	0.53	297
Garmisch, Germany	4.68	17.15	0.45	35
Saga, Japan	8.48	14.53	0.50	184
Karlsruhe, Germany	-2.51	17.12	0.41	115
Lauder, New Zealand	5.70	12.38	0.56	99
Lamont, USA	-0.22	14.70	0.65	400
Park Falls, USA	0.69	16.69	0.36	235
All observations	0.10	15.50	0.76	2642
	$\bar{\mu}$	$\sigma_{\bar{\mu}}$	$\bar{\sigma}$	$\sigma_{\bar{\sigma}}$
Mean of sites	1.76	4.24	15.32	1.79

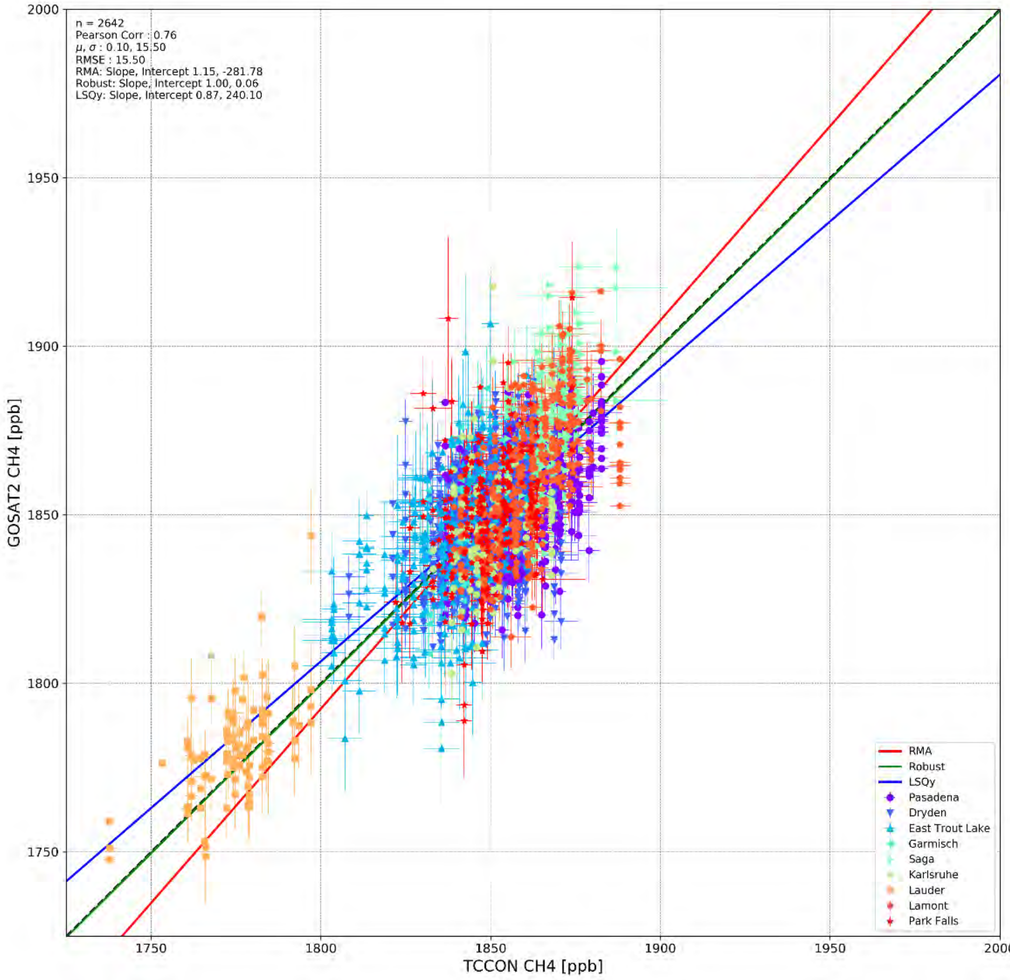


Figure 5.6-2: Correlation plot between all 2642 co-located CH4_GO2_SRPR and TCCON XCH_4 pairs coloured by site.

The error that comes out of the RemoTeC retrieval is just a purely statistical error on the radiance that has been propagated through the entire retrieval chain. In order to more accurately estimate the actual random error on the GOSAT-2 sounding, we applied the following procedure to obtain a scaling factor with which to scale our statistical error. We take the absolute difference of every co-located sounding and divide it by the retrieved statistical error corresponding to that sounding. We then average these values to obtain the average scaling factor by which to scale the retrieved statistical error to obtain a more correct estimate of the random error.

Based on the analysis, we obtain the following scaling factors for the SRPR XCH_4 product, 1.71 for the normal mode and 1.36 for the sunglint mode. Subsequently, we calculate the uncertainty ratio which is defined as the ratio of the mean value of the reported uncertainty and the standard deviation of the difference to TCCON. We obtain uncertainty ratios of 0.58 for the normal mode and 0.74 for the sunglint mode.

5.6.2 Summary

The result of the validation of the CH4_GO2_SRPR dataset is given in **Table 5.6-2** and compared to the requirement. The mean estimate of the single-measurement precision is 15.50 ppb which exceeds the goal requirement but is within the breakthrough requirement of 17 ppb. The uncertainties provided by RemoTeC agree on average with the observed scatter of the data when compared to TCCON. The mean, global bias of the GOSAT-2 XCH_4 retrieval is 0.10 ppb with a relative accuracy of 4.24 ppb which is smaller than the requirement of 10 ppb. We have not assessed the spatio-temporal bias or the drift due to the limited time period covered by the CH4_GO2_SRPR dataset.

Table 5.6-2: Summary validation of product CH4_GO2_SRPR by the data provider using TCCON ground-based reference data.

Product Quality Summary Table for Product: CH4_GO2_SRPR Level: 2, Version: v2, Time period covered: 2.2019 – 8.2020 Assessment: Data Provider (DP)			
Parameter [unit]	**Achieved performance**	**Requirement**	**Comments**
Single measurement precision (1-sigma) in [ppb]	15.50	< 34 (T) < 17 (B) < 9 (G)	Computed as standard deviation of the difference to TCCON
Uncertainty ratio [-]: Ratio reported uncertainty to standard deviation of satellite-TCCON difference	0.58 (0.74 glint)	-	No requirement but value close to unity expected for a high quality data product with reliable reported uncertainty.
Mean bias (global offset) [ppb]	0.10	-	No requirement but value close to zero expected for a high quality data product.
Accuracy: Relative systematic error [ppb]	Spatial: 4.24 Spatio-temporal: Not evaluated	< 10	Spatial: Computed as standard deviation of the biases at the various TCCON sites. Spatio-temporal: As "Spatial" but also considering seasonal biases.
Stability: Drift [ppb/year]	Not evaluated (1-sigma)	< 3	Linear drift

6 References

/Aben et al., 2019/ Aben, I., L. Wu, O. Hasekamp, M. Buchwitz, M. Reuter, O. Schneising, H. Boesch, D. Yang, ESA Climate Change Initiative "Plus" (CCI+) System Specification Document (SSD) for the Essential Climate Variable (ECV) Greenhouse Gases (GHG): XCO_2 and/or XCH_4 from OCO-2, TanSat, Sentinel-5-Precursor and GOSAT-2, Version 1.0, 26-August-2019, pp. 33, 2019.

/Ackermann et al., 2010/ Ackerman, S.; Frey, R.; Strabala, K.; Liu, Y.; Gumley, L.; Baum, B.; Menzel, P. Discriminating Clear-Sky from Cloud with MODIS—Algorithm Theoretical Basis Document (MOD35); Version 6.1; Cooperative Institute for Meteorological Satellite Studies, University ofWisconsin—Madison: Madison, WI, USA, 2010.

/ATBDv1 FOCAL, 2019/ Reuter et al.: Algorithm Theoretical Basis Document Version 1 (ATBDv1) - Retrieval of XCO_2 from the OCO-2 satellite using the Fast Atmospheric Trace Gas Retrieval (FOCAL), ESA GHG-CCI project, 2019.

/ATBDv3, 2021/ M. Reuter, M. Hilker, S. Noël, M. Buchwitz, O. Schneising, H. Bovensmann, and J. P. Burrows: Algorithm Theoretical Basis Document Version 3 (ATBDv3) - Retrieval of XCO2 from the OCO-2 satellite using the Fast Atmospheric Trace Gas Retrieval (FOCAL), ESA Climate Change Initiative "Plus" (CCI+) for the Essential Climate Variable (ECV) Greenhouse Gases (GHG), ESA GHG CCI+, 2021.

/Batchelor at al., 2009/ Batchelor, R. L., Strong, K., Lindenmaier, R., Mittermeier, R. L.,Fast, H., Drummond, J. R., and Fogal, P. F.: A New Bruker IFS 125HR FTIR Spectrometer for the Polar Environment Atmospheric Research Laboratory at Eureka, Nunavut, Canada: Measurements and Comparison with the Existing Bomem DA8 Spectrometer, J. Atmos. Ocean. Tech., 26, 1328–1340,https://doi.org/10.1175/2009JTECHA1215.1, 2009.

/Blumenstock et al., 2020/ Blumenstock T., Raffalski U., and Matsumi Y., NDACC Metadata File for Eureka FTIR and references cited therein, https://www-air.larc.nasa.gov/pub/NDACC/PUBLIC/meta/ftir/Kiruna_ftir_tb_2020.pdf

/Boesch et al., 2011/ Boesch, H., Baker, D., Connor, B., Crisp, D., and Miller, C.: Global characterization of CO_2 column retrievals from shortwave-infrared satellite observations of the Orbiting Carbon Observatory-2 mission, Remote Sens., 3, 270–304, doi:10.3390/rs3020270, 2011.

/Bovensmann et al., 2010/ Bovensmann, H., Buchwitz, M., Burrows, J. P., Reuter, M., Krings, T., Gerilowski, K., Schneising, O., Heymann, J., Tretner, A., and Erzinger, J.: A remote sensing technique for global monitoring of power plant CO2 emissions from space and related applications, Atmospheric Measurement Techniques, 3, 781–811, doi:10.5194/amt-3-781-2010, URL http: //www.atmos-meas-tech.net/3/781/2010/, 2010.

/Butz et al., 2009/ Butz, André, Otto P. Hasekamp, Christian Frankenberg, and Ilse Aben, "Retrievals of atmospheric CO_2 from simulated space-borne measurements of backscattered near-infrared sunlight: accounting for aerosol effects," Appl. Opt. 48, 3322-3336 (2009)

/Butz et al., 2010/ Butz, André, Otto P. Hasekamp, Christian Frankenberg, Jerome Vidot, and Ilse Aben, ",CH4 retrievals from space-based solar backscatter measurements:

Performance evaluation against simulated aerosol and cirrus loaded scenes" , J. Geophys. Res., 115, 24,302, doi:10.1029/2010JD014514, 2010.

/Butz et al., 2011/ Butz, A., S. Guerlet, O. Hasekamp, et al., Toward accurate CO_2 and CH_4 observations from GOSAT, Geophys. Res. Lett., Geophys. Res. Lett., doi:10.1029/2011GL047888, 2011.

/CECRv3 BESD, 2017/ M. Reuter, O. Schneising, M. Buchwitz, J. Heymann: Comprehensive Error Characterisation Report: BESD full-physics retrieval algorithm for XCO_2 for the Essential Climate Variable (ECV) Greenhouse Gases (GHG), ESA GHG-CCI project, 2017.

/CECRv3, 2017/ M. Reuter, O. Schneising, M. Buchwitz, J. Heymann: Comprehensive Error Characterisation Report: BESD full-physics retrieval algorithm for XCO2 for the Essential Climate Variable (ECV) Greenhouse Gases (GHG), ESA GHG-CCI project, 2017.

/De Mazière et al., 2017/ De Mazière, M., Sha, M. K., Desmet, F., Hermans, C., Scolas, F., Kumps, N., Metzger, J.-M., Duflot, V., and Cammas, J.-P.: TCCON data from Réunion Island (RE), Release GGG2014.R1, https://doi.org/10.14291/TCCON.GGG2014.REUNION01.R1, https://data.caltech.edu/records/322, 2017.

/De Mazière et al., 2018/ De Mazière, M., Thompson, A. M., Kurylo, M. J., Wild, J. D., Bernhard, G., Blumenstock, T., Braathen, G. O., Hannigan, J. W., Lambert, J.-C., Leblanc, T., McGee, T. J., Nedoluha, G., Petropavlovskikh, I., Seckmeyer, G., Simon, P. C., Steinbrecht, W., and Strahan, S. E.: The Network for the Detection of Atmospheric Composition Change (NDACC): history, status and perspectives, Atmos. Chem. Phys., 18, 4935–4964, https://doi.org/10.5194/acp-18-4935-2018, 2018.

/Deutcher et al, 2010/ Deutscher, N. M., Griffith, D. W. T., Bryant, G. W., Wennberg, P.O., Toon, G. C., Washenfelder, R. A., Keppel-Aleks, G., Wunch,D., Yavin, Y., Allen, N. T., Blavier, J.-F., Jiménez, R., Daube,B. C., Bright, A. V., Matross, D. M., Wofsy, S. C., and Park,S.: Total column CO2 measurements at Darwin, Australia – site description and calibration against in situ aircraft profiles, Atmos. Meas. Tech., 3, 947–958, https://doi.org/10.5194/amt-3-947-2010, 2010

/Deutscher et al., 2019/ Deutscher, N. M., Notholt, J., Messerschmidt, J., Weinzierl, C., Warneke, T., Petri, C., and Grupe, P.: TCCON data from Bialystok (PL), Release GGG2014.R2, https://doi.org/10.14291/TCCON.GGG2014.BIALYSTOK01.R2, https://data.caltech.edu/records/1300, 2019.

/E3UBv3, 2022/ M.Reuter, M.Hilker, M. Ingendahl, and M.Buchwitz: End-to-End ECV Uncertainty Budget Version 3 (E3UBv3) for the FOCAL XCO_2 OCO-2 Data Product CO2_OC2_FOCA (v10) for the Essential Climate Variable (ECV) Greenhouse Gases (GHG) ESA Climate Change Initiative "Plus" (CCI+), 2022

/Eldering et al., 2015/ Eldering, A.; Pollock, R.; Lee, R.; Rosenberg, R.; Oyafuso, F.; Crisp, D.; Chapsky, L.; Granat, R. Orbiting Carbon Observatory-2 (OCO-2) - LEVEL 1B - Algorithm Theoretical Basis; Version 1.2 Rev 1; National Aeronautics and Space Administration, Jet Propulsion Laboratory, California Institute of Technology: Pasadena, CA, USA, 2015.

/Feist et al.,2014/ Feist, D. G., Arnold, S. G., John, N., and Geibel, M. C.: TCCON data from Ascension Island (SH), Release GGG2014.R0, https://doi.org/10.14291/TCCON.GGG2014.ASCENSION01.R0/1149285, https://data.caltech.edu/records/210, 2014.

/Frankenberg et al., 2005/ Frankenberg, C., Meirink, J.F., van Weele, M., Platt, U., Wagner, T. (2005), Assessing methane emissions from global space-borne obseravtions, Science, ., doi:10.1126/science.1106644

/Frankenberg et al., 2011/ Frankenberg, C., Aben, I., Bergamaschi, P., et al. (2011), Global column-averaged methane mixing ratios from 2003 to 2009 as derived from SCIAMACHY: Trends and variability, J. Geophys. Res., doi:10.1029/2010JD014849.

/GCOS-154/ Global Climate Observing System (GCOS), SYSTEMATIC OBSERVATION REQUIREMENTS FOR SATELLITE-BASED PRODUCTS FOR CLIMATE, Supplemental details to the satellite-based component of the "Implementation Plan for the Global Observing System for Climate in Support of the UNFCCC (2010 update)", Prepared by World Meteorological Organization (WMO), Intergovernmental Oceanographic Commission, United Nations Environment Programme (UNEP), International Council for Science, Doc.: GCOS 154, link: http://cci.esa.int/sites/default/files/gcos-154.pdf, 2011.

/GCOS-195/ Status of the Global Observing System for Climate. GCOS-195. Link: https://library.wmo.int/pmb_ged/gcos_195_en.pdf, 2015.

/GCOS-200/ The Global Observing System for Climate: Implementation Needs, GCOS 2016 Implementation Plan, World Meteorological Organization (WMO), GCOS-200 (GOOS-214), pp. 325, link: https://library.wmo.int/opac/doc_num.php?explnum_id=3417, 2016.

/Goo et al., 2014/ Goo, T.-Y., Oh, Y.-S., & Velazco, V. A. (2014). *TCCON data from Anmeyondo (KR), Release GGG2014.R0* (Version GGG2014.R0) [Data set]. CaltechDATA. https://doi.org/10.14291/TCCON.GGG2014.ANMEYONDO01.R0/1149284, https://data.caltech.edu/records/266

/Griffith et al. 2014a/ Griffith, D. W., Deutscher, N. M., Velazco, V. A., Wennberg, P. O., Yavin, Y., Keppel-Aleks, G., Washenfelder, R. A., Toon, G. C., Blavier, J.-F., Paton-Walsh, C., Jones, N. B., Kettlewell, G. C., Connor, B. J., Macatangay, R. C., Roehl, C., Ryczek, M., Glowacki, J., Culgan, T., and Bryant, G. W.: TCCON data from Darwin (AU), Release GGG2014.R0, https://doi.org/10.14291/TCCON.GGG2014.DARWIN01.R0/1149290, https://data.caltech.edu/records/269, 2014a.

/Griffith et al. 2014b/ Griffith, D. W., Velazco, V. A., Deutscher, N. M., Paton-Walsh, C., Jones, N. B., Wilson, S. R., Macatangay, R. C., Kettlewell, G. C., Buchholz, R. R., and Riggenbach, M. O.: TCCON data from Wollongong (AU), Release GGG2014.R0, https://doi.org/10.14291/TCCON.GGG2014.WOLLONGONG01.R0/1149291, https://data.caltech.edu/records/291, 2014b.

/Hannigan et al., 2021/ Hannigan J. W., Ortega I. and Coffey M.: NDACC Metadata File for Thule FTIR and references cited therein, https://www-air.larc.nasa.gov/pub/NDACC/PUBLIC/meta/ftir/Thule_ftir_jh_2021.pdf

/Hase et al., 2015/ Hase, F., Blumenstock, T., Dohe, S., Groß, J., and Kiel, M.: TCCON data from Karlsruhe (DE), Release GGG2014.R1, https://doi.org/10.14291/TCCON.GGG2014.KARLSRUHE01.R1/1182416, https://data.caltech.edu/records/278, 2015.

/HITRAN2016/ Gordon et al., The HITRAN2016 molecular spectroscopic database, J. Quant. Spectrosc. Radiat. Transfer, 203, 3-69, 2017.

/Iraci et al. 2016/ Iraci, L. T., Podolske, J. R., Hillyard, P. W., Roehl, C., Wennberg, P. O., Blavier, J.-F., Landeros, J., Allen, N., Wunch, D., Zavaleta, J., Quigley, E., Osterman, G. B., Albertson, R., Dunwoody, K., and Boyden, H.: TCCON data from Edwards (US), Release GGG2014.R1, https://doi.org/10.14291/TCCON.GGG2014.EDWARDS01.R1/1255068, https://data.caltech.edu/records/270, 2016.

/Kawakami et al. 2014/ Kawakami, S., Ohyama, H., Arai, K., Okumura, H., Taura, C., Fukamachi, T., and Sakashita, M.: TCCON data from Saga (JP), Release GGG2014.R0, https://doi.org/10.14291/TCCON.GGG2014.SAGA01.R0/1149283, https://data.caltech.edu/records/288, 2014.

/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 XCO2 affected by erroneous surface pressure estimates, Atmos. Meas. Tech., 12, 2241–2259, https://doi.org/10.5194/amt-12-2241-2019, 2019.

/Kivi et al., 2014/ Kivi, R., Heikkinen, P., and Kyrö, E.: TCCON data from Sodankylä (FI), Release GGG2014.R0, https://doi.org/10.14291/TCCON.GGG2014.SODANKYLA01.R0/1149280, https://data.caltech.edu/records/289, 2014.

/Kuruz, 1994/ Kurucz, H.L. The Solar Spectrum: Atlases and Line Identifications, Workshop on Laboratory and Astronomical High Resolution Spectra. Astronomical Society of the Pacific Conference Series. In Proceedings of the ASP Conference, Brussels, Belgium, 29 August–2 September 1994; pp. 17–31.

/Langerock et al., 2015/ Langerock, B., De Mazière, M., Hendrick, F., Vigouroux, C., Desmet, F., Dils, B., and Niemeijer, S.: Description of algorithms for co-locating and comparing gridded model data with remote-sensing observations, Geosci. Model Dev., 8, 911–921, https://doi.org/10.5194/gmd-8-911-2015, 2015.

/Liu et al. 2018/ Liu, C, Wang, W., Sun, Y., TCCON data from Hefei(PRC), Release GGG2014.R0, (Version GGG2014.R0) [Data set]. CaltechDATA. https://doi.org/10.14291/TCCON.GGG2014.HEFEI01.R0, https://data.caltech.edu/records/1092

/Lunneborg, 2020/ Lunneborg, C. E. Data analysis by resampling: Concepts and applications. PacificGrove, CA: Duxbury, 2000.

/Mahieu 2017/ Mahieu, M. NDACC Metadata File for Jungraujoch FTIR and references cited therein https://www-air.larc.nasa.gov/pub/NDACC/PUBLIC/meta/ftir/Jungfraujoch_ftir_em_2017.pdf

/Makarova et al. 2017/, Marakova M.V., Virolainen Y., Polyakov A.: NDACC Metadata File for St. Petersburg FTIR and references cited therein https://www-air.larc.nasa.gov/pub/NDACC/PUBLIC/meta/ftir/StPetersburg_ftir_mm_2017.pdf

/Mellqvist et al., 2021/ Mellqvist J., Persson G., NDACC Metadata File for Harestua FTIR and references cited therein, https://www-air.larc.nasa.gov/pub/NDACC/PUBLIC/meta/ftir/Harestua_ftir_hrt_jm_2021.pdf

/Morino et al. 2018a/ Morino, I., Matsuzaki, T., and Horikawa, M.: TCCON data from Tsukuba (JP), 125HR, Release GGG2014.R2, https://doi.org/10.14291/TCCON.GGG2014.TSUKUBA02.R2, https://data.caltech.edu/records/958, 2018a.

/Morino et al. 2018b/ Morino, I., Yokozeki, N., Matsuzaki, T., and Horikawa, M.: TCCON data from Rikubetsu (JP), Release GGG2014.R2, https://doi.org/10.14291/TCCON.GGG2014.RIKUBETSU01.R2, https://data.caltech.edu/records/957, 2018b.

/Morino et al., 2018c/ Morino, I., Velazco, V. A., Hori, A., Uchino, O., & Griffith, D. W. T. (2018). *TCCON data from Burgos, Ilocos Norte (PH), Release GGG2014.R0* (Version GGG2014.R0) [Data set]. CaltechDATA. https://doi.org/10.14291/TCCON.GGG2014.BURGOS01.R0, https://data.caltech.edu/records/1090

/Notholt et al., 2019/ Notholt, J., Petri, C., Warneke, T., Deutscher, N. M., Palm, M., Buschmann, M., Weinzierl, C., Macatangay, R. C., and Grupe, P.: TCCON data from Bremen (DE), Release GGG2014.R1, https://doi.org/10.14291/TCCON.GGG2014.BREMEN01.R1, https://data.caltech.edu/records/1290, 2019.

/Notholt et al., 2021a/ Notholt, J., Palm, M., NDACC Metadata File for Ny Alesund FTIR and references cited therein https://www-air.larc.nasa.gov/pub/NDACC/PUBLIC/meta/ftir/NyAlesund_ftir_jn_2021.pdf

/Notholt et al., 2021b/ Notholt, J., Palm, M., NDACC Metadata File for Bremen FTIR and references cited therein https://www-air.larc.nasa.gov/pub/NDACC/PUBLIC/meta/ftir/Bremen_ftir_jn_2021.pdf

/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.; et al. The ACOS CO_2 retrieval algorithm - Part 1: Description and validation against synthetic observations. Atmos. Meas. Tech. 2012, 5, 99–121.

/O'Dell et al., 2018/ O'Dell, C. W., Eldering, A., Wennberg, P. O., Crisp, D., Gunson, M. R., Fisher, B., Frankenberg, C., Kiel, M., Lindqvist, H., Mandrake, L., Merrelli, A., Natraj, V., Nelson, R. R., Osterman, G. B., Payne, V. H., Taylor, T. E., Wunch, D., Drouin, B. J., Oyafuso, F., Chang, A., McDuffie, J., Smyth, M., Baker, D. F., Basu, S., Chevallier, F., Crowell, S. M. R., Feng, L., Palmer, P. I., Dubey, M., García, O. E., Griffith, D. W. T., Hase, F., Iraci, L. T., Kivi, R., Morino, I., Notholt, J., Ohyama, H., Petri, C., Roehl, C. M., Sha, M. K., Strong, K., Sussmann, R., Te, Y., Uchino, O., and Velazco, V. A.: Improved retrievals of carbon dioxide from Orbiting Carbon Observatory-2 with the version 8 ACOS algorithm, Atmos. Meas. Tech., 11, 6539–6576, https://doi.org/10.5194/amt-11-6539-2018, 2018.

/Ortega et al., 2019/ Ortega, I., Buchholz, R. R., Hall, E. G., Hurst, D. F., Jordan, A. F.,and Hannigan, J. W.: Tropospheric water vapor profiles obtained with FTIR: comparison with balloon-borne frost point hygrometers and influence on trace gas retrievals, Atmos. Meas. Tech.,12, 873–890, https://doi.org/10.5194/amt-12-873-2019, 2019.

/Ostler et al., 2014/ Ostler, A., Sussmann, R., Rettinger, M., Deutscher, N. M., Dohe, S., Hase, F., Jones, N., and Palm, M.: Multistation intercomparison of column-averaged methane from NDACC and TCCON: Impact of dynamical variability, Atmos. Meas. Tech., 7, 4081–4101, https://doi.org/10.5194/amt-7-4081-2014, 2014.

/Parker et al., 2011/ Parker, R., Boesch, H, Cogan, A., et al. (2011), Methane Observations from the Greenhouse gases Observing SATellite: Comparison to ground-based TCCON data and Model Calculations, Geophys. Res. Lett., doi:10.1029/2011GL047871.

/Petri et al., 2020/ Petri, C., Vrekoussis M., Rousogenous, C., Warneke, T., Sciare, J., Notholt, J., TCCON data from Nicosia, Cyprus (CY), Release GGG2014.R0, https://doi.org/10.14291/TCCON.GGG2014.NICOSIA01.R0/, https://data.caltech.edu/records/1690

/PUGS, 2019/ Buchwitz, M., Aben, I., J., Armante, R., Boesch, H., Crevoisier, C., Di Noia, A., Hasekamp, O. P., Reuter, M., Schneising-Weigel, O., Wu, L.., Product User Guide and Specification (PUGS) – Main document for Greenhouse Gas (GHG: CO_2 & CH_4) data set CDR 3 (2003-2018), C3S project C3S_312b_Lot2_DLR, v3.1, 97 pages, https://www.iup.uni-bremen.de/carbon_ghg/docs/C3S/CDR3_2003-2018/PUGS/C3S_D312b_Lot2.3.2.3-v1.0_PUGS-GHG_MAIN_v3.1.pdf, 2019.

/PUGv4, 2022/ M.Reuter, M.Hilker, and M.Buchwitz: Product User Guide Version 4 (PUGv4) for the FOCAL XCO2 OCO-2 Data Product CO2_OC2_FOCA (v10) for the Essential Climate Variable (ECV) Greenhouse Gases (GHG) ESA Climate Change Initiative "Plus" (CCI+), 2022

/PVIRv5, 2017/ Buchwitz et al., Product Validation and Intercomparison Report (PVIR), version 5, ESA GHG-CCI project, 2017

/PVIR v5, 2017/ Product Validation and Intercomparison Report" (PVIR), ESA Climate Change Initiative (CCI) GHG-CCI project, Version 5, 9 Feb 2017, 2017.

/PVIR GHG-CCI+ v1.1, 2020/ ESA Climate Change Initiative "Plus" (CCI+) Product Validation and Intercomparison Report (PVIR) for the Essential Climate Variable (ECV) Greenhouse Gases (GHG): XCO2 and/or XCH4 from OCO-2, TanSat, Sentinel-5-Precursor and GOSAT-2,13 Mar 2020, version 1.1, pp. 74, 2020.

/Reuter et al., 2011/ Reuter, M., Bovensmann, H., Buchwitz, M., Burrows, J. P., Connor, B. J., Deutscher, N. M., Griffith, D.W. T., Heymann, J., Keppel-Aleks, G., Messerschmidt, J., and et al.: Retrieval of atmospheric CO_2 with enhanced accuracy and precision from SCIAMACHY: Validation with FTS measurements and comparison with model results., Journal of Geophysical Research, 116, doi:10.1029/2010JD015047, URL http://dx.doi.org/10.1029/2010JD015047, 2011.

/Reuter et al., 2012/ M. Reuter, M. Buchwitz, O. Schneising, F. Hase, J. Heymann, S. Guerlet, A.J. Cogan, H. Bovensmann, J.P. Burrows: A simple empirical model estimating atmospheric CO_2 background concentrations, Atmospheric Measurement Techniques, 5, 1349-1357, 2012.

/Reuter et al., 2013/ Reuter, M., H. Boesch, 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, Atmos. Chem. Phys., 13, 1771-1780, 2013.

/Reuter et al., 2017a/ M.Reuter, M.Buchwitz, O.Schneising, S.Noël, V.Rozanov, H.Bovensmann and J.P.Burrows: A Fast Atmospheric Trace Gas Retrieval for Hyperspectral Instruments Approximating Multiple Scattering - Part 1: Radiative Transfer and a Potential OCO-2 XCO_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

/Rodgers, 2000/ Rodgers, C. D.: Inverse Methods for Atmospheric Sounding: Theory and Practice, World Scientific Publishing, 2000.

/Rozanov et al., 2014/ Rozanov, V., Rozanov, A., Kokhanovsky, A., and Burrows, J.: Radiative transfer through terrestrial atmosphere and ocean: Software package SCIATRAN, Journal of Quantitative Spectroscopy and Radiative Transfer, 133, 13 – 71, doi:https://doi.org/10.1016/j.jqsrt.2013.07.004,
URL http://www. sciencedirect.com/science/article/pii/S0022407313002872, 2014

/**Schepers et al., 2012/** Schepers, D., Guerlet, S., Butz, A., et al. (2012), Methane retrievals from Greenhouse Gases Observing Satellite (GOSAT) shortwave infrared measurements: Performance comparison of proxy and physics retrieval algorithms, *J. Geophys. Res.*, 117, D10307, doi:10.1029/2012JD017549.

/Schneising et al., 2011/ Schneising, O., Buchwitz, M., Reuter, M., et al., Long-term analysis of carbon dioxide and methane column-averaged mole fractions retrieved from SCIAMACHY, Atmos. Chem. Phys., 11, 2881-2892, 2011.

/Schneising et al., 2012/ Schneising, O., Bergamaschi, P., Bovensmann, H., et al, Atmospheric greenhouse gases retrieved from SCIAMACHY: comparison to ground-based FTS measurements and model results, Atmos. Chem. Phys., 12, 1527-1540, 2012.

/Schneising et al., 2019/ Schneising, O., Buchwitz, M., Reuter, M., Bovensmann, H., Burrows, J. P., Borsdorff, T., Deutscher, N. M., Feist, D. G., Griffith, D. W. T., Hase, F., Hermans, C., Iraci, L. T., Kivi, R., Landgraf, J., Morino, I., Notholt, J., Petri, C., Pollard, D. F., Roche, S., Shiomi, K., Strong, K., Sussmann, R., Velazco, V. A., Warneke, T., and Wunch, D.: A scientific algorithm to simultaneously retrieve carbon monoxide and methane from TROPOMI onboard Sentinel-5 Precursor, Atmos. Meas. Tech., 12, 6771-6802, https://doi.org/10.5194/amt-12-6771-2019, 2019.

/Sherlock et al., 2014/ Sherlock, V., Connor, B., Robinson, J., Shiona, H., Smale, D., and Pollard, D. F.: TCCON data from Lauder (NZ), 125HR, Release GGG2014.R0, https://doi.org/10.14291/TCCON.GGG2014.LAUDER02.R0/1149298, https://data.caltech.edu/records/281, 2014.

/Stammes, 2002/ Stammes, P. OMI Algorithm Theoretical Basis Document, Volume III, Clouds, Aerosols, and Surface UV Irradiance (ATBD-OMI-03); Royal Netherlands Meteorological Institute (KNMI): De Bilt, The Netherlands, 2002.

/Strong et al., 2019/ Strong, K., Roche, S., Franklin, J. E., Mendonca, J., Lutsch, E., Weaver, D., Fogal, P. F., Drummond, J. R., Batchelor, R., and Lindenmaier, R.: TCCON data from Eureka (CA), Release GGG2014.R3, https://doi.org/10.14291/TCCON.GGG2014.EUREKA01.R3, https://data. caltech.edu/records/1171, 2019.

/Strong, 2021/ Strong K., NDACC Metadata File for Eureka FTIR and references cited therein, https://www-air.larc.nasa.gov/pub/NDACC/PUBLIC/meta/ftir/Eureka_ftir_ks_2021.pdf

/Sussmann et al., 2018a/ Sussmann, R. and Rettinger, M.: TCCON data from Garmisch (DE), Release GGG2014.R2, https://doi.org/10.14291/TCCON.GGG2014.GARMISCH01.R2, https://data.caltech.edu/records/956, 2018.

/Sussmann et al., 2018b/ Sussmann, R. and Rettinger, M.: NDACC Metadata File for Zugspitze FTIR and references cited therein M.: https://www-air.larc.nasa.gov/pub/NDACC/PUBLIC/meta/ftir/Zugspitze_ftir_rs_2018.pdf

/Té et al., 2014/ Té, Y., Jeseck, P., & Janssen, C. (2014). *TCCON data from Paris (FR), Release GGG2014.R0* (Version GGG2014.R0) [Data set]. CaltechDATA. https://doi.org/10.14291/TCCON.GGG2014.PARIS01.R0/1149279, https://data.caltech.edu/records/284

/Thompson et al., 2012/ Thompson, D.R.; Chris Benner, D.; Brown, L.R.; Crisp, D.; Malathy Devi, V.; Jiang, Y.; Natraj, V.; Oyafuso, F.; Sung, K.; Wunch, D.; et al. Atmospheric validation of high accuracy CO_2 absorption coefficients for the OCO-2 mission. J. Quant. Spectrosc. Radiat. Transf. 2012, 113, 2265–2276

/Thompson et al., 2012/ Thompson, D.R.; Chris Benner, D.; Brown, L.R.; Crisp, D.; Malathy Devi, V.; Jiang, Y.; Natraj, V.; Oyafuso, F.; Sung, K.; Wunch, D.; et al. Atmospheric validation of high accuracy CO_2 absorption coefficients for the OCO-2 mission. J. Quant. Spectrosc. Radiat. Transf. 2012, 113, 2265–2276

/URDv2.1/ User Requirements Document (URD), ESA Climate Change Initiative (CCI) GHG-CCI project, Version 2.1, 38 pages, 19-Oct-2016, http://cci.esa.int/sites/default/files/URDv2.1_GHG-CCI_Final.pdf, 2016.

/URDv3.0, 2020/ Chevallier et al., GHG-CCI: User Requirements Document for the GHG-CCI+ project of ESA's Climate Change Initiative, pp. 42, version 3.0, 17 February, 2020. Link: https://www.iup.uni-bremen.de/carbon_ghg/docs/GHG-CCIplus/URD/URDv3.0_GHG-CCIp_Final.pdf.

/Warneke et al., 2019/ Warneke, T., Messerschmidt, J., Notholt, J., Weinzierl, C., Deutscher, N. M., Petri, C., and Grupe, P.: TCCON data from Orléans (FR), Release GGG2014.R1, https://doi.org/10.14291/TCCON.GGG2014.ORLEANS01.R1, https://data.caltech.edu/records/1301, 2019.

/Wennberg et al., 2015/ Wennberg, P. O., Wunch, D., Roehl, C. M., Blavier, J.-F., Toon, G. C., and Allen, N. T.: TCCON data from Caltech (US), Release GGG2014.R1, https://doi.org/10.14291/TCCON.GGG2014.PASADENA01.R1/1182415, https://data.caltech.edu/records/285, 2015.

/Wennberg et al., 2016a/ Wennberg, P. O., Roehl, C. M., Blavier, J.-F.,Wunch, D., and Allen, N. T.: TCCON data from Jet Propulsion Laboratory (US), 2011, Release GGG2014.R1, https://doi.org/10.14291/TCCON.GGG2014.JPL02.R1/1330096, https://data.caltech.edu/records/277, 2016a.

/Wennberg et al., 2016b/ Wennberg, P. O., Wunch, D., Roehl, C. M., Blavier, J.-F., Toon, G. C., and Allen, N. T.: TCCON data from Lamont (US), Release GGG2014.R1, https://doi.org/10.14291/TCCON.GGG2014.LAMONT01.R1/1255070, https://data.caltech.edu/records/279, 2016b.

/Wennberg et al., 2017/ Wennberg, P. O., Roehl, C. M., Wunch, D., Toon, G. C., Blavier, J.-F., Washenfelder, R., Keppel-Aleks, G., Allen, N. T., and Ayers, J.: TCCON data from Park Falls (US), Release GGG2014.R1, https://doi.org/10.14291/TCCON.GGG2014.PARKFALLS01.R1, https://data.caltech.edu/records/295, 2017.

/Wiacek et al., 2007/ Wiacek, A., Taylor, J. R., Strong, K., Saari, R., Kerzenmacher, T. E., Jones, N. B., and Griffith, D. W. T.: Ground-Based Solar Absorption FTIR Spectroscopy: Characterization of Retrievals and First Results from a Novel Optical Design Instrument at a New NDACC Complementary Station, J. Atmos. Ocean. Tech., 24,432–448, https://doi.org/10.1175/JTECH1962.1, 2007

/Wunch et al., 2010/ Wunch, D., Toon, G. C., Wennberg, P. O., Wofsy, S. C., Stephens, B. B., Fischer, M. L., Uchino, O., Abshire, J. B., Bernath, P., Biraud, S. C., Blavier, J.-F. L., Boone, C., Bowman, K. P., Browell, E. V., Campos, T., Connor, B. J., Daube, B. C., Deutscher, N. M., Diao, M., Elkins, J. W., Gerbig, C., Gottlieb, E., Griffith, D. W. T., Hurst, D. F., Jiménez, R., Keppel-Aleks, G., Kort, E. A., Macatangay, R., Machida, T., Matsueda, H., Moore, F., Morino, I., Park, S., Robinson, J., Roehl, C. M., Sawa, Y., Sherlock, V., Sweeney, C., Tanaka, T., and Zondlo, M. A.: Calibration of the Total Carbon Column Observing Network using aircraft profile data, Atmospheric Measurement Techniques, 3, 1351–1362, doi:10.5194/amt-3-1351-2010, URL http://www.atmos-meas-tech.net/3/1351/2010/, 2010.

/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.

/Wunch et al., 2015/ Wunch, D., Toon, G. C., Sherlock, V., Deutscher, N. M., Liu, C., Feist, D. G., & Wennberg, P. O. (2015). Documentation for the 2014 TCCON Data Release. CaltechDATA. https://doi.org/10.14291/TCCON.GGG2014.DOCUMENTATION.R0/1221662, https://data.caltech.edu/records/249

/Wunch et al., 2017/ Wunch, D., Mendonca, J., Colebatch, O., Allen, N. T., Blavier, J.-F., Springett, S., Neufeld, G., Strong, K., Kessler, R., & Worthy, D. (2017). *TCCON data from East Trout Lake, SK (CA), Release GGG2014.R0* (Version GGG2014.R0) [Data set]. CaltechDATA.
https://doi.org/10.14291/TCCON.GGG2014.EASTTROUTLAKE01.R0/1348207, https://data.caltech.edu/records/296

/Yang et al., 2020/ Yang, D., H. Boesch, Y. Liu, P. Somkuti, Z. Cai, X. Chen, A. Di Noia, C. Lin, N. Lu, D. Lyu, R. J. Parker, L. Tian, M. Wang, A. Webb, L. Yao, Z. Yin, Y. Zheng, N. M. Deutscher, D. W. T. Griffith, F. Hase, R. Kivi, I. Morino, J. Notholt, H. Ohyama, D. F. Pollard, K. Shiomi, R. Sussmann, Y. Te, V. A. Velazco, T. Warneke, D. Wunch: Toward high precision XCO_2 retrievals from TanSat observations: Retrieval improvement and validation against TCCON measurements, Journal of Geophysical Research: Atmospheres, 125, e2020JD032794, https://doi.org/10.1029/2020JD032794, 2020.

7 List of Acronyms and Abbreviations

Abbreviation	Meaning
AAI	Absorbing Aerosol Index
ACA	Additional Constraints Algorithm
AOD	Aerosol Optical Depth
AOT	Aerosol Optical Thickness
ATBD	Algorithm Theoretical Basis Document
BIRA-IASB	Royal Belgian Institute for Space Aeronomy
CCI	Climate Change Initiative
CDR	Climate Data Record
CMUG	Climate Modelling User Group (of ESA's CCI)
COD	Cloud Optical Depth
CRG	Climate Research Group
D/B	Data base
DOAS	Differential Optical Absorption Spectroscopy
DPM	Detailed Processing Model
EC	European Commission
ECA	ECV Core Algorithm
ECMWF	European Centre for Medium Range Weather Forecasting
ECV	Essential Climate Variable
EO	Earth Observation
ESA	European Space Agency
ESM	Earth System Model
FCDR	Fundamental Climate Data Record
FOCAL	Fast atmOspheric traCe gAs retrievaL
FoM	Figure of Merit
FP	Full Physics
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
GMES	Global Monitoring for Environment and Security
GOSAT	Greenhouse Gas Observing Satellite
IDL	Interactive Data Language
ITT	Invitation To Tender
IODD	Input Output Data Definition
IPCC	International Panel in Climate Change
IPR	Intellectual Property Right
IUP	Institute of Environmental Physics (IUP) of the University of Bremen, Germany
JCGM	Joint Committee for Guides in Metrology
LMD	Laboratoire de Météorologie Dynamique
LUT	Look-up table
MACC	Monitoring Atmospheric Composition and Climate, EU GMES project
MERIS	Medium Resolution Imaging Spectrometer
MIPAS	Michelson Interferometer for Passive Atmospheric Sounding
MODIS	Moderate Resolution Imaging Spectrometer
N/A	Not applicable
NDACC	Network for the Detection of Atmospheric Composition Change
NASA	National Aeronautics and Space Administration
NIES	National Institute for Environmental Studies
NIWA	National Institute Of Water & Atmospheric Research
NOAA	National Oceanic and Atmospheric Administration
OCO	Orbiting Carbon Observatory
OD	Optical Depth

OE	Optimal Estimation
PBL	Planetary Boundary Layer
PMD	Polarization Measurement Device
PR	Proxy (retrieval method)
PVP	Product Validation Plan
PVR	Product Validation Report
RA	Relative Accuracy
RD	Reference Document
RMS	Root-Mean-Square
RTM	Radiative transfer model
S5P	Sentinel-5 Precursor
SoW	Statement of work
SQWG	SCIAMACHY Quality Working Group
SRA	Seasonal Relative Accuracy
SRD	Software Requirements Document
SRON	Netherlands Institute for Space Research
SUM	Software User Manual
SVR	Software Verification Report
TANSAT	CarbonSat
TANSO	Thermal And Near infrared Sensor for carbon Observation
TBC	To be confirmed
TCCON	Total Carbon Column Observing Network
TBD	To be defined / to be determined
TROPOMI	TROPOspheric Monitoring instrument
UNAM	Universidad Nacional Autónoma de México
WFM-DOAS (or WFMD)	Weighting Function Modified DOAS
WG	Working Group

***** End of Document *****