ICOS Central Analytical Laboratories

Flask and Calibration Laboratory (FCL)

Quality Control Report 2021

Armin Jordan, Marcus Schumacher

ICOS | Central Analytical Laboratory

Jena, March 2022

Table of contents

List of Figures

List of Tables

Summary

The ICOS Central Analytical Laboratories (CALs) play a central role in assuring the accuracy of atmospheric observations within ICOS. This involves the central provision of reference gases to the ICOS atmospheric network and calibrating these standards based on the World Meteorological Organization (WMO) calibration scales. A quality control strategy for the ICOS atmospheric measurements has been described within the Atmospheric Station Specification document [ATC 2020].

In this report the quality control measures are described that are made by the ICOS-CAL Flask and Calibration Laboratory (FCL) to characterize the performance of their calibration of ICOS reference gases. It updates and replaces the QC 2020 report following the same assessment scheme with only minor changes and some few corrections. The results of these activities of the recent years are presented in detail for each of the ICOS core components for in-situ observations (CO_2, CH_4, CO) and N_2O. The results are then assessed and used to substantiate estimates of the measurement uncertainties of the different tracers and to quantify different uncertainty contributions. This involves an evaluation of the uncertainty of the reference values of calibration standard gases ("scale link uncertainty") and the measurement uncertainty related to the respective analyzer's precision or response stability over time.

The resulting overall measurement uncertainty estimates are summarized in the following table.

Table 1: Summary of total estimated measurement uncertainties
Data taken from sections 5.6, 6.6, 7.6 and 8.6 but expressed as expanded uncertainty (95% confidence level), combined uncertainties are calculated as the square root of the sum of squared uncertainty contributions

	CO_2 [ppm]	CH_4 [ppb]	CO [ppb]	N_2O [ppb]
CCL reproducibility[1)]	0.02	1	0.8	0.22
scale propagation to FCL standards	0.026	0.4	1.1[2)]	0.05
scale link uncertainty	**0.032**	**1.1**	**1.4**	**0.23**
instrumental precision	0.028	0.4	0.2	0.03
long-term reproducibility	0.02	0.4	1.0[2)]	0.04
estimated FCL reproducibility	**0.034**	**0.5**	**1.1**	**0.05**
estimated overall uncertainty	**0.047**	**1.2**	**1.4**	**0.23**

[1)] WMO Central Calibration Laboratory (CCL)

[2)] includes uncertainty of CO growth in FCL reference standards

Based on further evidence obtained in 2021 the assumption of a Primary Standard gas set with stable CO_2 appears justified. The scale link uncertainty in CO_2 has been revised with the implementation of the WMO GAW* CO_2 Calibration Scale revision (from X2007 to X2019), that became available in February 2021.

* GAW (Global Atmosphere Watch)

This report is a deliverable (D7) of Annex 2 to the Cooperation Agreement between ICOS ERIC and the Max-Planck-Society that is the host organization for the ICOS Flask and Calibration Laboratory (FCL).

1 Introduction

The mission of ICOS is to run a long-term monitoring network that produces harmonized sets of highly precise and accurate observational data. The data should be of a quality to allow for regularly assessing regional carbon fluxes from atmospheric observations using inversion models, to detect changes in emission patterns and to quantify long-term trends. This requires highly consistent experimental records available over decades. The ICOS strategy to ensure best consistency of the entire atmospheric monitoring network includes the central data processing of the measurement data of all instruments at the monitoring stations (done at the Atmospheric Thematic Center ATC) and a central provision of calibrated reference gases by one of the Central Analytical Laboratories, the Flask and Calibration Laboratory (FCL).

This makes it particularly necessary for the FCL to have a comprehensive QA/QC framework with well-defined analytical procedures in place to assure accurate measurements based on WMO calibration scales. The different components of the FCL quality control system described in this report aim to address all requirements for a comprehensive quality control strategy listed in the ICOS Atmospheric Station Specification Document [ATC_2020]. The results of these quality control activities shall document the achieved accuracy, shall allow an assessment of the uncertainty of the assigned values on reference gases and generate credibility by comparing with various external laboratories, including laboratories that are completely independent from ICOS (as the Bureau International des Poids et Mesures (BIPM) and the WMO-CCL).

The aim of this report is to present the results of the measures undertaken by the FCL that contain information on the data quality of its measurement activities for the ICOS community. In the past years the main function requested from the ICOS FCL was the provision of calibrated reference gases for the routine operation of the station measurements, the recalibration of the station standards and to equip stations that have entered the ICOS labeling process. Hence, this report focusses on the quality control of reference gas measurements performed for the ICOS atmosphere observational network. Mole fraction assignments have been made for the core parameters CO_2, CH_4 and CO as well as for N_2O as recommended parameter and are made with the following instrumentation:

- Picarro G2301 Cavity Ringdown Spectrometer (CO_2 and CH_4)
- Los Gatos CO/N_2O Analyzer EP (CO and N_2O)

2 Measurement Methods

Picarro method brief description (see also Annex I)

CO_2 and CH_4 mole fractions of reference standards that are prepared for the station network in high pressure cylinders are assigned by using a Picarro G2301 Cavity-Ring-Down-Spectroscope. The instrument is operated using the software tool GCwerks that exports averaged one minute Level0 data for further processing. Data is migrated in an automated way into an in-house-developed data base on a daily basis for further processing (quality control, calibration, aggregation), before the data is manually validated and finally forwarded to the ATC's data server. The Level0 data is checked and automatically flagged according to predefined criteria for valid data. This includes instrumental readings (cell pressure, sample flow, sampling frequency), the reproducibility within the one minute averages as well as the scatter of the one minute averages, and noise level (standard deviation of the means and 3-sigma excluded outliers (see also Annex III)).

Each measurement (samples as well as references) takes 20 min of gas injection. To avoid cross contamination of succeeding samples and to flush out the pressure regulator, the first five minutes of data at the beginning of each measurement are ignored and the average of the remaining valid 15min data is further processed.

The instrument is calibrated on a daily basis by a dedicated set of four FCL Secondary Reference Standards. These secondary references are calibrated about quarterly against a set of nine FCL Primary Standards with assignments from the WMO Central Calibration Laboratory (CCL).

Los Gatos method brief description (see also Annex II)

CO and N_2O mole fractions of reference standards that are prepared for the station network in high pressure cylinders are assigned using a Los Gatos CO/N_2O Enhanced Performance Off-Axis Integrated Cavity Output Spectroscopy (OA-ICOS) instrument. The instrument is operated using an in-house build software that controls a multiposition valve for sample provision, collects raw data and delivers averaged 20s Level0 data for further processing. Data are automatically migrated after the termination of the measurement sequence into an in-house-developed data base for further processing (automatic quality control, calibration, aggregation), before the data is manually validated and finally forwarded to the ATC's data server. The Level0 data is checked and flagged automatically according to predefined criteria for valid data. This includes instrumental readings (cell pressure, sample flow, sampling frequency), the repeatability within the one minute averages as well as the scatter of the one minute averages, and noise level (standard deviation of the means and 3-sigma excluded outliers (see also Annex III)).

Each measurement (samples as well as references) involves 20 min of gas injection. To avoid cross contamination of succeeding samples and to flush out the pressure regulator, the first nine minutes of data (27 averages of 20 sec) and the last 20 sec data point of the measurement are ignored and the average of the remaining valid 10 min data is further processed.

Short term drifts of the analyzer are compensated by bracketing every sample analysis by measurements of a working reference standard and normalizing the sample signal to the averaged working standard signal.

The instrument is calibrated by a dedicated set of four FCL Secondary Reference Standards in every series of measurements (at least on a daily basis). These secondary references are calibrated against a set of nine FCL Primary Standards with assignments from the WMO Central Calibration Laboratory (CCL).

GC method description

A gas chromatographic analysis system (GC) has been set up primarily for analysis of flask samples from class1 stations. GC measurements also yield data for the tracers measured by the optical analyzers and thus can be used as an independent check. The GC is equipped with multiple detectors: a Flame Ionization Detector (FID) for CO_2 and CH_4 detection, an Electron Capture Detector (ECD) for N_2O, and a Reduction Gas Detector for CO (HgO Reduction and Hg-UV Detection).

The GC is calibrated for CO_2 and CH_4 by a set of five Secondary Standards dedicated to the GC with currently a bi-weekly to monthly frequency. To calibrate the non-linear detectors for CO and N_2O measurements an extended set of seven Secondary Standards is used. These GC Secondary Reference Gases are calibrated against the set of nine FCL Primary Standards three to four times per year.

3 Calibration gases linking to the WMO Mole Fraction scales

All FCL measurements are traceable to the WMO Mole Fraction Scales. This link is established by a set of standard gases that has been calibrated directly by the WMO Central Calibration Laboratory (CCL). In the WMO/GAW nomenclature these standards are on the level of laboratory tertiary standards (relative to the WMO Mole Fraction scale Primary Standards). However, for the ease of reading they will be referred to throughout this document as FCL Primary Standards. The accuracy of their assignments is an essential prerequisite for the accuracy of the ICOS measurements. Likewise, the knowledge of the stability of the mole fractions of the tracers of interest in these gases is essential for accurate measurements.

Using the set of standards calibrated directly by the CCL as reference (listed in Table 2), additional sets of further working calibration standards (denoted in this document as FCL Secondary Standards) have been derived that are used for daily calibrations of the individual instruments.

All of the FCL Primary Standards have been recalibrated at the CCL twice with the most recent recalibration having been made in 2021. This allows to verify the stability of the respective trace gases or track the rate of change of concentration. In 2021 a change of the WMO scale has been made by the CCL that was implemented at the FCL.

Cylinder ID	Sample ID	Fill date	last CCL calibration	CO_2 (ppm)[1]	CH_4 (ppb)[2]	CO (ppb)[3]	N_2O (ppb)[4]
CB09948	i20140054	07/2013	05/2021	250.13	2933.01	998.13	361.90
CB09944	i20140055	07/2013	05/2021	339.36	1596.74	36.75	317.01
CB09939	i20140056	07/2013	05/2021	365.28	1743.09	84.47	319.93
CB09958	i20140057	07/2013	05/2021	389.77	1896.75	125.03	327.21
CB09983	i20140058	07/2013	05/2021	412.42	2032.96	162.73	330.06
CB09952	i20140059	07/2013	05/2021	433.83	2195.08	203.13	334.57
CB09955	i20140060	07/2013	05/2021	459.17	2343.92	249.93	339.43
CB09957	i20140061	07/2013	05/2021	482.02	2466.23	399.39	343.82
CB09934	i20140062	07/2013	05/2021	515.11	2731.85	697.10	349.14

WMO Mole Fraction scale:
[1] CO_2WMO X2019 (CRDS only)
[2] CH_4 WMO X2004A
[3] CO WMO X2014A
[4] N_2O WMO X2006A

CCL-reproducibility (2 sigma) [reference]:
0.01 ppm [1 sigma, https://gml.noaa.gov/ccl/co2_calsystem.html]
1 ppb (pers. comm., E. Dlugokencky, Feb. 2018)
0.8 ppb [CCL_CO 2017]
0.22 ppb [CCL_N_2O 2011]

Table 2: FCL Primary Standards assignments by the WMO Central Calibration Laboratory (values from recalibration 2021), updated to the currently valid WMO Mole Fraction scales.

4 Quality Control Concept

For all measurements made the general approach is the following:

1. *FCL Primary Standards*: To assure compatibility of ICOS observational data all measurements are linked to the WMO calibration scales. For this the set of FCL Primary Standards covers the atmospheric ranges of the trace gases of interest and has been assigned by the Central Calibration Laboratories (CCL). According to the WMO Experts Group's recommendations these assignments should be re-assessed by regular recalibration by the WMO CCL every third year. In order to always have a sufficient set of Primary Standards at the FCL, sub-groups of each three standards are re-sent to the CCL for recalibration on an annual basis.
2. *FCL Secondary Standards*: All measurements are referenced to daily calibrations using laboratory Secondary Standard gases that have been assigned at the FCL by repeated comparison to the FCL Primary Standards.
3. *Targets*: The performance of daily measurements is characterized by daily analysis of the same gases over long periods of time that are only used for quality assessment (so-called "Target standards")
4. *Inter-Instrument comparisons*: In cases where additional gas chromatographic measurements have been made these results are compared to the spectroscopic data.
5. *External comparisons* have been made routinely with the MPI-BGC GasLab, regularly with a large group of laboratories as part of the "Sausage Intercomparison Program", using the "MENI" (**M**PI-BGC, **E**MPA, **N**OAA and **I**COS) - Intercomparison that includes among others the NOAA-GML as partner laboratory, and occasionally with other laboratories (e.g. WMO Round Robin, BIPM Key Comparison, ATC-Mobile Lab).

All of these steps are evaluated to provide the following information on the FCL data quality (see the respective subsections of chapters 5 to 8 for the respective assessments of the CO_2, CH_4, CO and N_2O measurements):

FCL Primary Standards

- Re-assignments by the CCL provide information on the assignment accuracy or the stability of the specific tracer's mole fraction in the reference gas.
- The observed magnitude of the calibration regression fit residuals contains information on the consistency of the CCL assignments. The persistency of these residuals over time may provide information of the stability of the respective tracers' mole fractions in the Primary Standards.

FCL Secondary Standards

- The consistency of the Secondary Standards' assignments obtained from repeated calibration episodes relative to the FCL Primary Standards is a measure for the uncertainty of the scale transfer and for the stability of the trace gas mole fraction in the reference gases.
- The magnitude of the mean secondary calibration regression fit residuals also contains information on the scale transfer uncertainty.
- The stability of these residuals over time may provide information on the stability of the respective tracers in the Primary Standards.
- The scatter of the daily residuals is an indicator of the reproducibility of the daily calibration.

Targets

- The reproducibility of the daily mean results of the targets shall reflect the long-term reproducibility of measurements that the FCL achieves for ICOS station's standard gases (provided that for the respective targets the tracer mole fractions are constant over time).
- Like the FCL Secondary Standards the targets have received an assignment by calibration directly with FCL Primary Standards. The difference of the daily measurement results (based on the daily secondary calibration) and these assigned values serves as quality control of the actual scale transfer error.

Inter-Instrument comparison

- The agreement of analysis results of the same sample by different detecting techniques is a measure to quantify potential analytical biases.
- The comparison also involves the cross-check of two different sets of laboratory Secondary Standard gases.

External comparison

WMO compatibility goals aim for achieving consistent atmospheric data from different networks with their associated stations and laboratories. Thus, control of this compatibility requires comparison with external partners. Comparison of analytical data from the same sample provides a check for the success of the overall measurement set-ups, including instrumentation, standardization strategy, data processing and the accuracy of the reference material.

5 CO_2

5.1 FCL Primary CO_2 Standards

5.1.1 CCL CO_2 assignments

After initial calibration of all FCL Primary Standards in 2014 the first recalibrations of each three of the standards have been made in 2016, 2017 and 2018, respectively. In 2021 the complete set received a recalibration such that three CCL assignments from different years are now available for each standard. The initial calibration was performed using NDIR technique. NDIR was also applied for the reassignments in 2016 / 2017, when additional measurements with CRDS analysers were also carried out. From 2018 onwards, recalibrations were made at the CCL only by CRDS (PC1). Hence, two CCL measurements with the CRDS technique are available for all of the nine standards. The revision of the WMO/GAW CO_2 X2007 to an updated X2019 Calibration Scale has been disclosed in February 2021.

The CRDS data confirm the temporal stability of the CO_2 mole fractions in each of these standards (Table 3). Earlier ambiguities related to potentially growing CO_2 in many standards probably were result of inferior reproducibility of NDIR X2007 assignments and different isotopic sensitivities between NDIR and CRDS.

Table 3: CO_2 assignments for FCL Primary Standards by CCL based on NDIR (initial data) and CRDS measurements.

Sample ID	Cylinder ID	CCL date 1	CCL date 2	CCL date 3	WMO X2019 NDIRI	WMO X2019 CRDS1	WMO X2019 CRDS2	corresponding drift [ppm/yr]
i20140054	CB09948	Dec-13	Oct-18	Mar-21	250.129	250.116	250.129	**0.005**
i20140055	CB09944	Mar-14	Jul-17	Mar-21	339.327	339.356	339.360	**0.001**
i20140056	CB09939	Dec-13	Oct-18	Mar-21	365.253	365.277	365.281	**0.002**
i20140057	CB09958	Dec-13	Oct-16	Mar-21	389.762	389.753	389.765	**0.003**
i20140058	CB09983	Feb-14	Oct-18	Mar-21	412.381	412.420	412.424	**0.002**
i20140059	CB09952	Jan-14	Sep-16	Mar-21	433.795	433.830	433.832	**0.000**
i20140060	CB09955	May-14	Jun-17	Mar-21	459.121	459.181	459.173	**-0.002**
i20140061	CB09957	Feb-14	Aug-16	Mar-21	481.962	482.010	482.022	**0.003**
i20140062	CB09934	May-14	Jun-17	Mar-21	515.053	515.120	515.113	**-0.002**

5.1.2 Regression fit residuals of FCL Primary CO_2 Standards

The time series of the linear regression fit residuals of CRDS calibrations made with these FCL Primary Standards (based on WMO CO_2 X2019 assignments) is presented in the following Figure 1 for calibration events where the *complete* suite of gases was used. The mean residuals of the individual standards range from -0.019 ppm to +0.015 ppm with a standard deviatin of these means of 0.012 ppm. This is a measure of the consistency of the initial CCL assignments confirming the specifications made by the CCL.

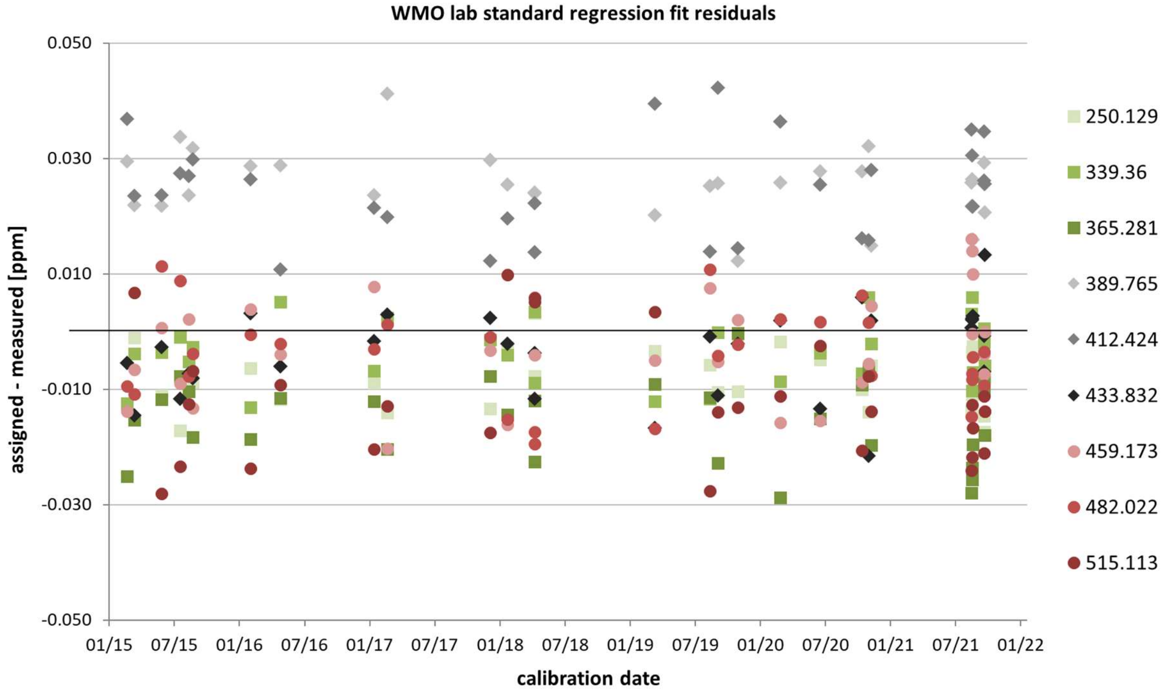


Figure 1: Time series of linear regression fit residuals of the CRDS CO_2 calibration for the individual FCL Primary Standards

The stability of the regression fit residuals over time provides information on possible drifts in individual standard gases. The values of the residuals do not show significant trends for any of the individual standards (within 0.01 ppm). This supports the finding of a set with stable CO_2 mole fractions.

5.2 FCL Secondary CO_2 Standards

5.2.1 Assignment record

The first set of four reference gases that were used as FCL Secondary Standards for the CRDS measurements had been analyzed within 20 (24) valid calibration episodes together with the *complete* set of FCL Primary Standards between Feb 2015 and Nov 2020. During 2020, the first set of FCL Secondary Standards had to be replaced by a new set due to consumption. The replacement was done in two stages, with the replacement of the two standards with higher concentrations in June and the replacement of the two standards with lower concentrations in December. After the replacement of the first two of the initial Secondary Standards in June 2020, updated assigned values were used for the two remaining standards for the period until their replacement from June to December.

The stability of CO_2 values for the new set of Secondary Standards was monitored by repeated measurements against the first set of FCL Secondary Standards, with base numbers for the individual tanks: i20190708 (n=145), i20190803 (n=141), i20190709 (n=115) and i20190438 (n=201). The assigned CO_2 values are based on the records of the CO_2 mole fraction results of the FCL Primary calibration episodes between Aug 2019 and Oct 2021 that are listed for comparison with the first set in Table 4 and Figure 2. The observed CO_2 growth in the second set of Secondary Standards appears significantly larger than in the first set (see Table 4) but with only two years of measurements the uncertainty of this growth rate derived from the red data points in Figure 2 also is higher.

An overall assignment revision was made also for the first set of Secondary Standards. This was based on new recalibration data from the CCL for the 9 Primary Standards, the reassignment arising from the X2019 scale update, and further data from the calibration of the Secondary Standards with the Primary Standards. The resulting time series of calibrations indicates a small growth of CO_2 in three of the four initial Secondary Standards (see Fig 2) i.e. temporal changing assigned values. This then required an update of all former FCL standard gas CO_2 assignments for the ICOS network.

Table 4: CO_2 assignments of CRDS Secondary Standards (as used for data processing in the FCL data base)

Sample ID	Cylinder ID	Assigned Value[1]	Drift/yr[1]	*Date of exchange*	Sample ID	Cylinder ID	Assigned Value[2]	Drift/yr[2]
i20140171	D801336	**359.870 ppm**	+0.003 ppm	*2020-12-08*	i20190708	D761202	**362.751 ppm**	+0.014 ppm
i20140172	D073384	**393.464 ppm**	+0.005 ppm	*2020-12-08*	i20190803	D073381	**402.078 ppm**	+0.010 ppm
i20140173	D073392	**424.724 ppm**	+0.007 ppm	*2020-06-23*	i20190709	D761214	**433.119 ppm**	+0.016 ppm
i20140174	D801331	**454.329 ppm**		*2020-06-23*	i20190438	D073389	**450.779 ppm**	+0.017 ppm

[1] Starting date: 1st January 2015

[2] Starts at 'Date of exchange'

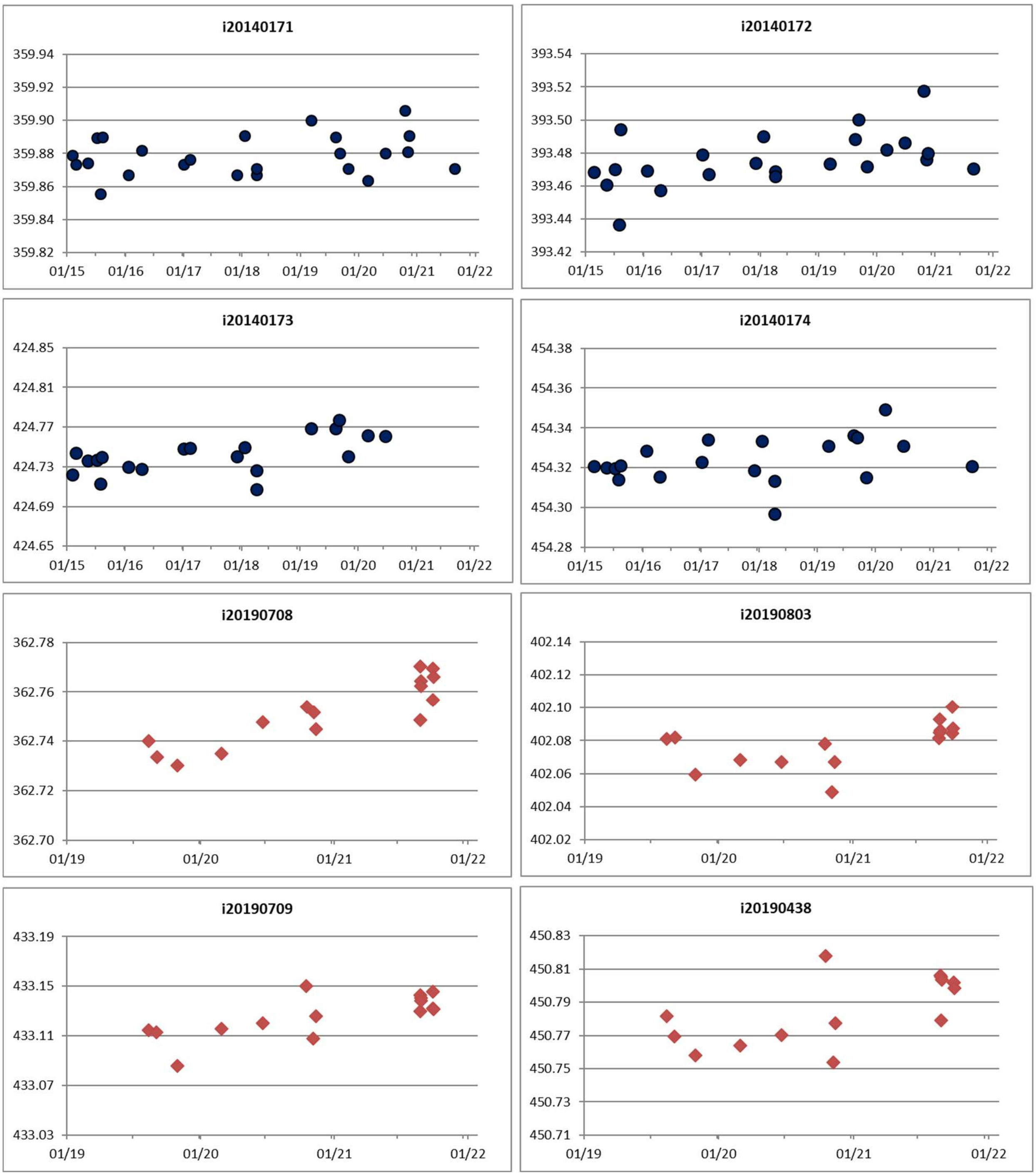


Figure 2: Time series of CRDS CO_2 measurement results for FCL Secondary Standards based on CCL calibrated FCL Primary Standards (all values in [ppm]). The dark dots represent the data from the first FCL Secondary Standards, the red diamonds display results of the four new FCL Secondary Standards.

5.2.2 Residual record

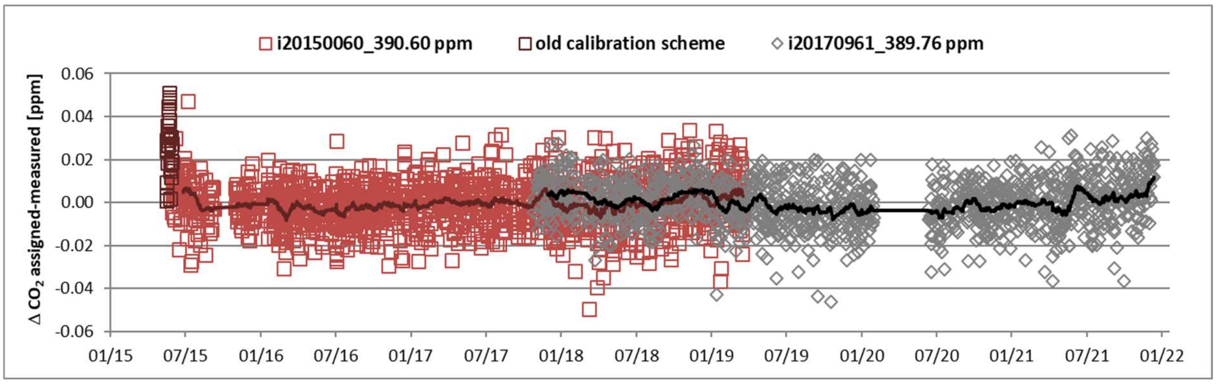
The residuals of the linear regression of the FCL Secondary Standards are given in Figure 3. The mean absolut residuals for the Secondary Standards are on the order of 0.002 ppm and smaller. The standard deviation of the daily residuals for the four individual standard gases in this period amounts to 0.006 ppm. Trends in the residuals over the period 2015-2020 do not exceed 0.005 ppm. This documents the long-term internal consistency of this calibration set over its lifetime. For the new calibration gas set (in use since June / December 2020), the standard deviation of the daily residuals has remained on a similar level with on average 0.007 ppm. The very small values of the mean residuals of all standards provide evidence for a consistent scale transfer to these FCL Secondary Standards.

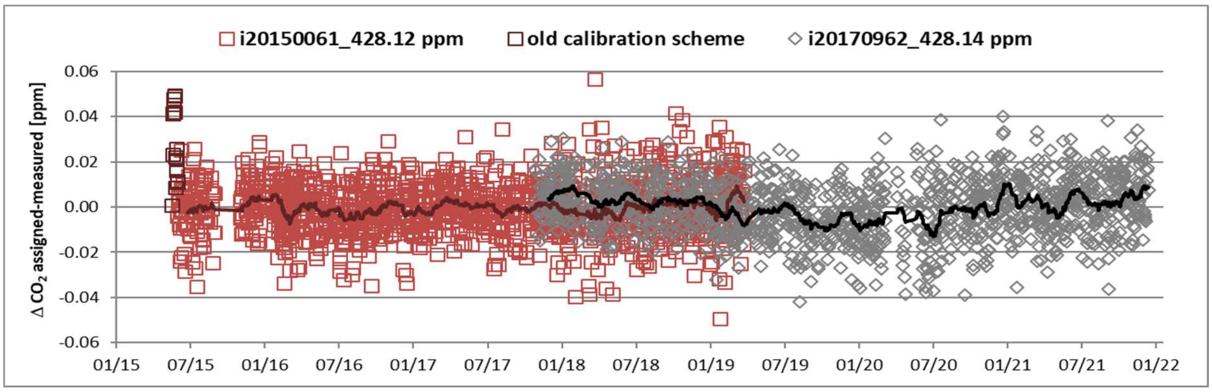
Figure 3: Time series of CO_2 linear regression fit residuals of the FCL Secondary Standards; blue squares first set of FCL Secondary Standards, red diamonds replacement set of FCL Secondary Standards. The dark symbols are indicating the transition phase when only the first part of the standards was replaced.

5.3 CO_2 Targets

In the period from March 2015 to December 2021 two succeeding sets of each three targets have been in use at the CRDS system. On a regular basis two further targets monitor the long-term stability of the instrument around 360 ppm. The Target Standards' mean measurement results are compared to the assigned values based on the Primary Standard calibrations in Figure 4. In this plot, the daily mean results are compared to the trend line in CO_2 observed in multiple calibrations made with the FCL Primary Standards. No bias is observed except for some minor synoptical patterns and variations of the measured results. The standard deviations of the daily target mean results is 0.012 ppm for the respective periods of use excluding the initial period of March-May 2015, when the calibration pattern of the CRDS instrument had not yet been in the same strict routine mode as it has been applied ever since.

Figure 5 displays the time series of CO_2 in the current set of target standards. A closer look suggests that while CO_2 has been relatively stable in the beginning it started to steadily grow since summer 2020. As at that point of time the Secondary Standard set was changed this might suggest that the current assumption of the CO_2 growth rates in the new Secondary Standard set is an overestimate. This will be verified by further calibration episodes.





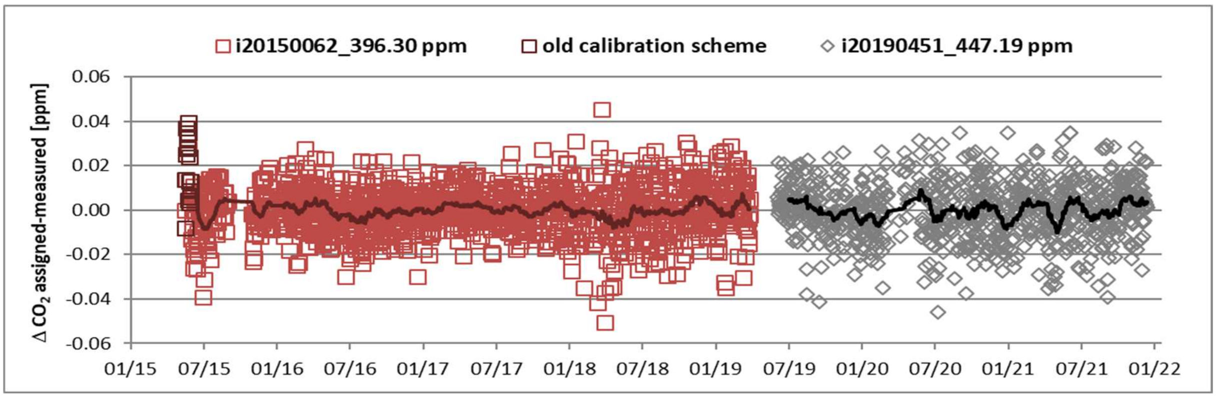
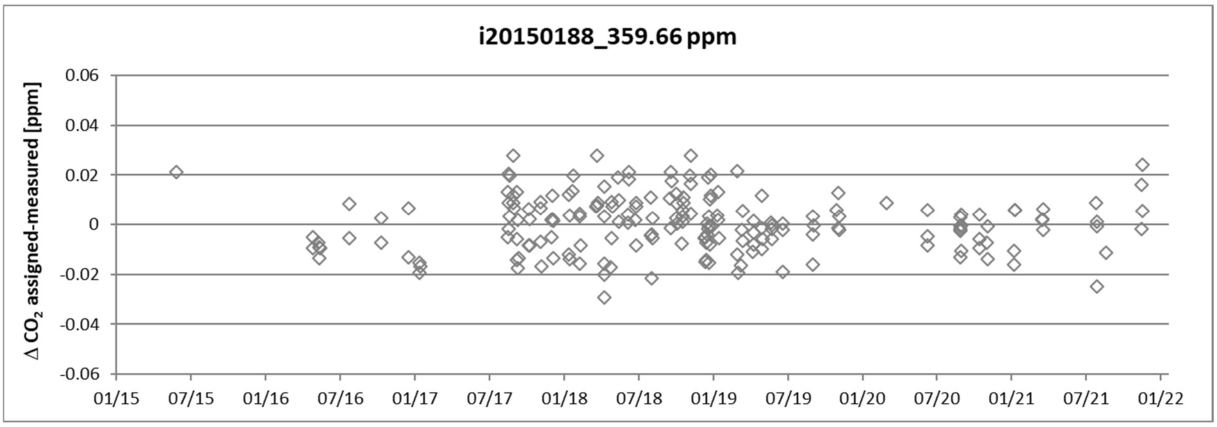


Figure 4: Time series of the offset of CO_2 target measurements to their respective assigned values (detrended). The dark line represents a 30 points-running mean. (Three outliers in January 2018 and January 2019 have been flagged out for i20150060, i20150061 and i20150062 for a more explicit visualisation).



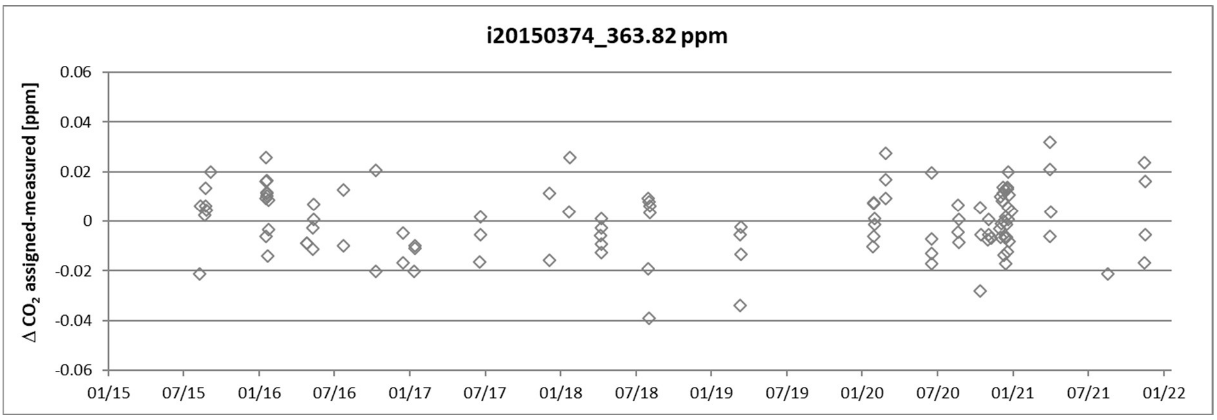
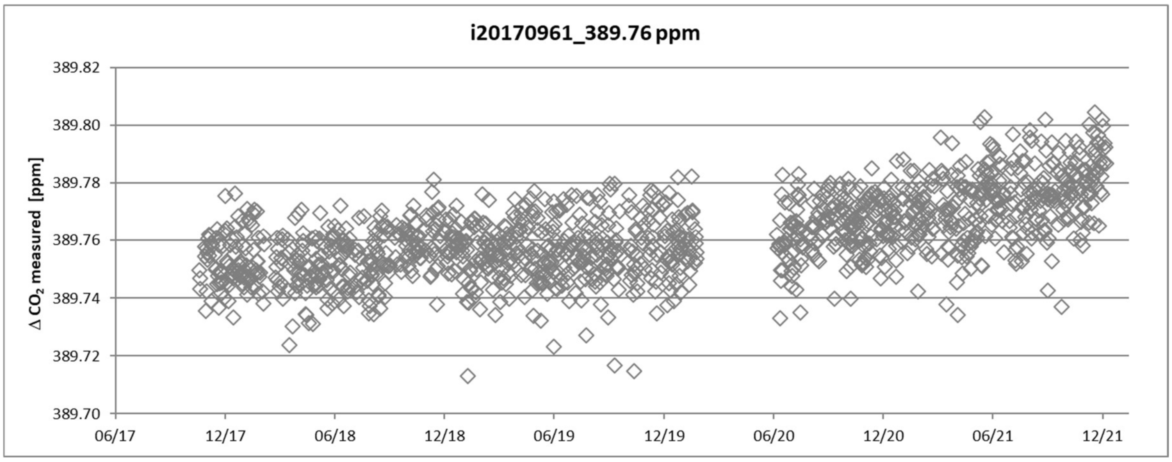
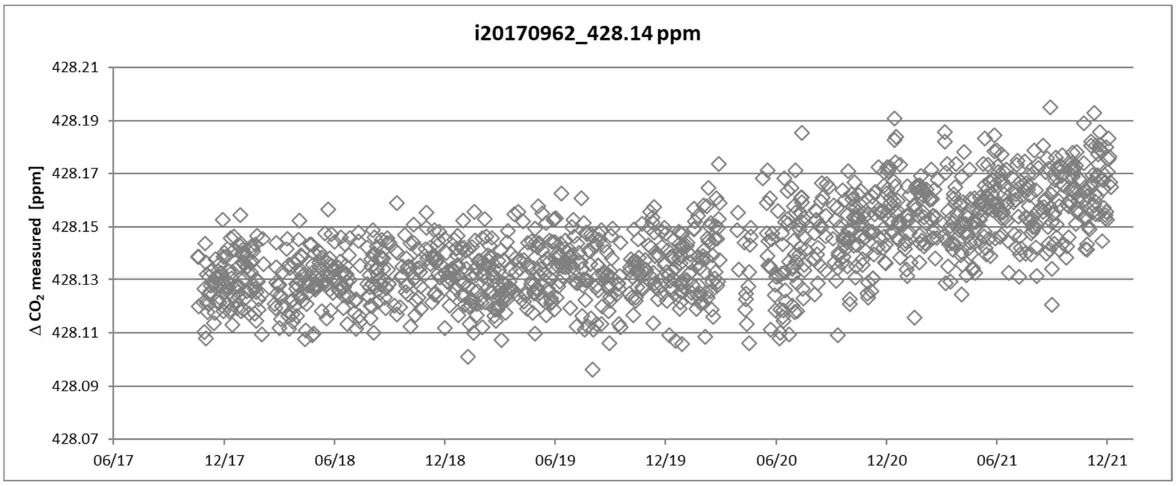


Figure 4: Time series of the offset of CO_2 target measurements to their respective assigned values (detrended). The dark line represents a 30 points-running mean. (Three outliers in January 2018 and January 2019 have been flagged out for i20150060, i20150061 and i20150062 for a more explicit visualisation).





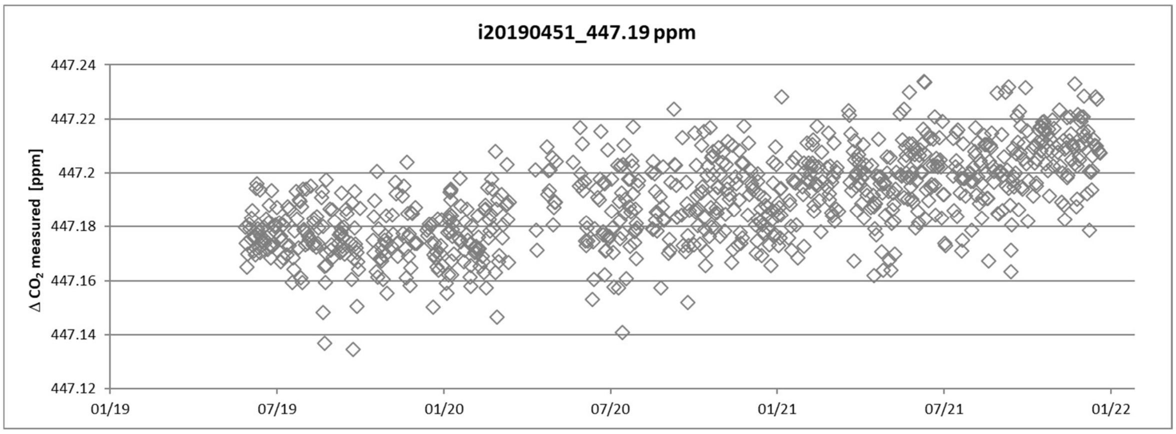
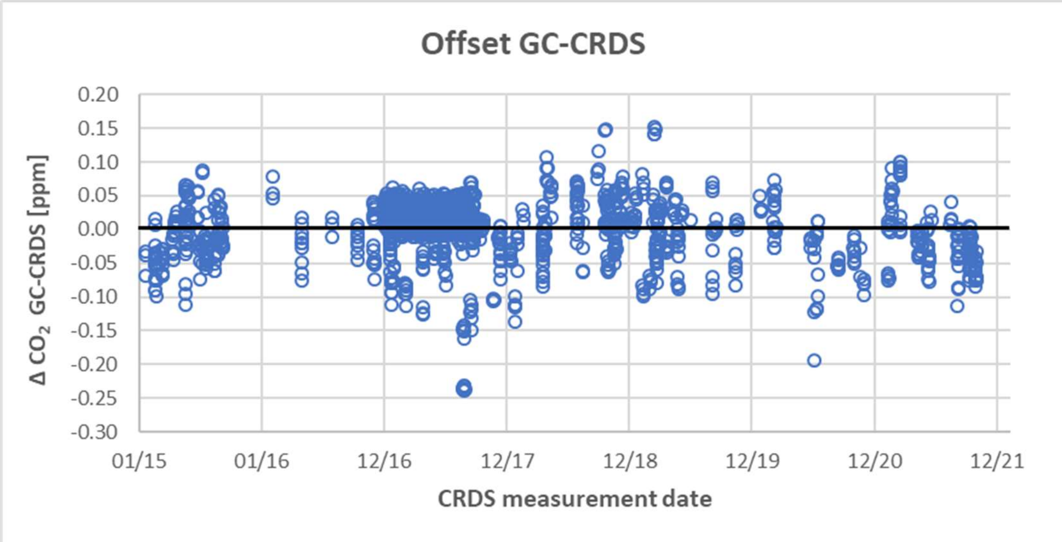


Figure 5: Time series of the measured CO_2 concentrations for the second set of the target tanks.

5.4 Internal CO_2 Comparison: CRDS-GC

Standard gases that are calibrated for CO_2 by CRDS have often also been analyzed by GC. The GC measurements are linked to the same set of FCL Primary Standards but based on a different set of five Secondary Standards. As reproducibility and repeatability of CO_2 measurements using the GC (0.04 ppm and 0.05 ppm, respectively) is in general by a factor of 4-5 worse compared to CRDS (0.01 ppm), only those GC measurements were considered for comparison that have been analyzed on the GC for more than one day with at least 20 injections. The inter-instrumental measurement differences for all standards are depicted in Figure 6 (including only standards within the range defined by the calibration standards). The average offset is 0.01 ppm ± 0.04 ppm.



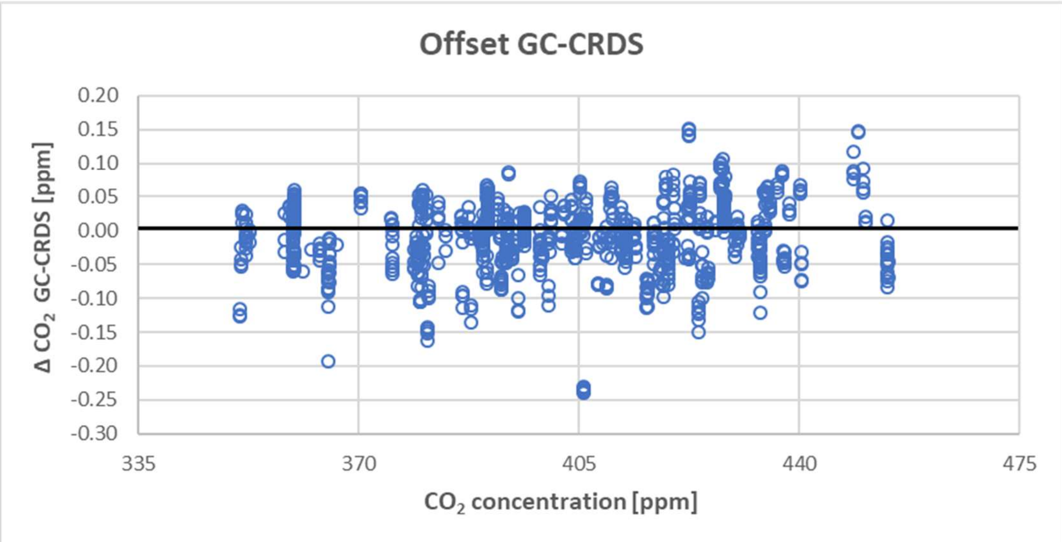


Figure 6: Offsets of daily CRDS CO_2 measurements relative to the average GC result. The black line represents the mean offset.

Note that each data point in Figure 6 represents the difference of one CRDS daily mean result relative to the mean of all GC measurements of the same sample. Some samples have been analyzed much more frequently than others (e.g. target standards) so the occurrence of many data points with -0.03 ppm in 2015 and +0.08 ppm in 2017 are not a sign of a change in the average GC-CRDS agreement in these years but more likely indicate a biased GC measurement of these very samples. The fact that the timing of the period of the GC measurements is different for the different samples limits a more detailed interpretation of the inter-instrument time series. Overall, the comparison with the independent GC measurements does not indicate any significant error in the CRDS measurements that might have been missed.

5.5 External CO_2 Comparisons

5.5.1 CO_2 compatibility ICOS FCL - MPI BGC

The most intensive comparison measurements have been made with the MPI-BGC GasLab. This laboratory is using different instruments (Picarro G1301 through April 2018, G2301 since May 2018) and their measurements are tied to the WMO Mole Fraction scales by an independent set of Lab Primary Standards. These MPI-BGC Primary Standards already have CCL calibration records with multiple measurements in different years for nine individual standards over six to seventeen years.

The MPI-BGC measurements are not relevant for the assignment of the FCL standards and therefore are completely independent.

5.5.1.1 Comparison of Primary CO_2 Standards

Basis for an agreement of FCL and MPI-BGC measurements is the compatibility of the respective sets of Primary Standards. As the FCL Primary Standards have been produced at the MPI-BGC they also were thoroughly analyzed at the MPI-BGC in 2013 and 2014 before being used by the FCL. Before or after the shipment to the CCL for recalibration of sub-sets of this FCL Primary Standard set these standards were also analyzed for a third time. Likewise, MPI-BGC Primary Standards that were simultaneously returned to the CCL for recalibration were also analyzed by the FCL. These data are shown in Figure 7 below.

The results of the MPI-BGC measurements of the complete FCL standard set are on average 0.025 ± 0.027 ppm lower than the CO_2 WMO X2019 PC1 assignments made by the CCL (red symbols). There is an apparent mole fraction dependency of the offset; constraining the compared standards to the four standards in the range of 360-450 ppm results in a mean offset of 0.014 ± 0.018 ppm. The same analysis of FCL measurement results of the MPI-BGC standard set yields a very close match with on average 0.008 ± 0.020 ppm higher values than the CCL PC1 assignments consistent with the mole fraction dependency seen by the MPI measurements (0.000 ± 0.014 ppm within the range of 360-450 ppm) (see Figure 7, blue symbols). Note that the two data sets in Figure 7 are presented on inverse axis because measurements using a set of Primary Standards that are on average carrying too high assignments will detect too little CO_2 in the set of standards that it is analyzing.

Comparison with additional sets of WMO standards could be made with the WMO Lab Standards of FMI (in 2016) and UBA (in 2021). Whereas the observed offset for the UBA set yields the same perfect match there is a consistent larger offset of 0.065 ppm for the FMI set. While this is not fully understood it should be noted that assignments for the FMI set have been made by NDIR which has a less well defined isotopologue selectivity.

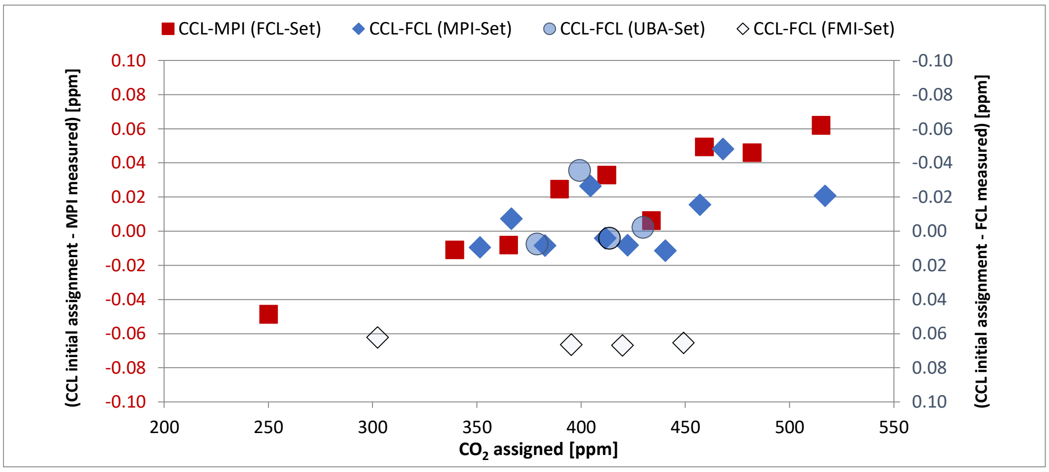


Figure 7: Differences of CCL initial CO_2 assignments to the MPI-BGC analysis results of the FCL Primary Standards (red squares) and to FCL analysis results of MPI and FMI Primary Standards (blue diamonds).
Note that the two data sets are on axis with opposite sign (see text)

5.5.1.2 Comparison of FCL Secondary CO_2 Standards

Three of the four gases from the first set of FCL Secondary Standards have been analyzed at the MPI-BGC in 2013 and 2014.

The results are in perfect agreement for the two lower standards but for the 454 ppm standard the MPI-BGC result is lower than the FCL assignment (ref. Table 5) which exceeds the increasing offset observed in the primary standards sets. The reason for this is not fully understood but this might reflect limitations of the MPI-BGC reproducibility at that time.

Table 5: Offset of MPI-BGC CO_2 analysis results to FCL Secondary CO_2 Standard assignments

FSN	Cylinder	FCL	MPI mean	MPI-FCL
i20140171	D801336	359.870	359.873	0.003
i20140173	D073392	424.724	424.730	0.006
i20140174	D801331	454.329	454.247	-0.082

5.5.1.3 Target standard CO_2 comparison

The three standard gases that were in use as target standards at the FCL also have been analyzed at the MPI-BGC. The differences of FCL measured means and MPI-BGC measured means given in Table 6 below are as small as expected from the agreement of the Secondary Standards.

Table 6: Comparison of MPI-BGC CO_2 analysis results of FCL target cylinders
$FCL_{measured}$ data are based on the daily calibration with FCL lab Secondary Standards and refer to results from Dec 2016 and July 2015 (i20150062) close to the periods when the measurements were made at MPI-BGC

Sample Number (FSN)	Cylinder Code	$FCL_{measured}$	$MPI_{measured}$	ΔMPI-FCL
i20150060	D073381	390.60	390.61	0.01
i20150061	D073389	428.11	428.12	0.01
i20150062	D073391	396.29	396.28	-0.01

5.5.1.4 Sample CO_2 comparison

High pressure standards have been regularly exchanged between MPI-BGC and FCL in earlier years and analyzed in both laboratories. The difference in results for about 280 comparisons is presented in Figure 8 below for all gases that have been analyzed within one year (only samples with CO_2 mole fractions within the calibrated ranges have been considered). The average offset of FCL - MPI of 0.009 ppm ± 0.029 ppm is again very small but also shows a mole fraction dependency with FCL results being larger compared to MPI results at higher concentrations and smaller at lower concentrations. Note that these differences include the measurement uncertainties of both laboratories. The MPI-BGC precision has been inferior up to May 2018 when a Picarro 1301 analyzer was replaced by a 2301 analyzer. The current MPI-BGC reproducibility is estimated as 0.02 ppm.

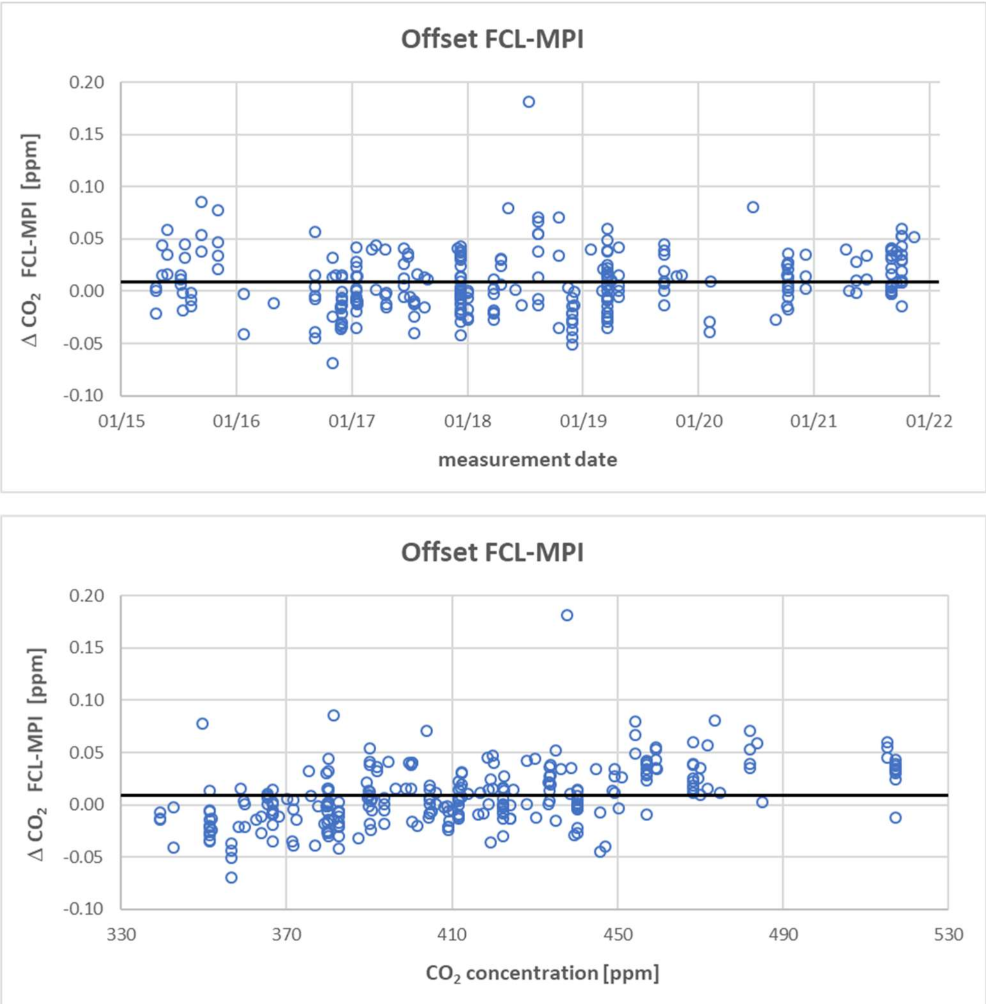


Figure 8: Differences of results for samples that have been analyzed for CO_2 at FCL and MPI

5.5.2 CO_2 compatibility ICOS FCL - NOAA

Comparison with the NOAA-GML laboratory (and other laboratories) is routinely made in two independent exercises, the Sausage Flask Intercomparison Program and the MENI (**M**PI – **E**MPA – **N**OAA -**I**COS) high pressure cylinder round robin program.

In the Sausage intercomparison, samples for comparison are prepared by connecting sets of flasks in line and filling them with dry air from a high-pressure cylinder at the FCL. The FCL generally analyzes the composition of the filling air using the normal instrumentation for calibrating standards. Therefore, the results of the flask measurements provided by NOAA can be compared to these high-pressure cylinder measurements. The respective data are compiled in Figure 9. The average agreement of NOAA mean flask results compared to FCL-CRDS filling gas data is very good (filled black circles, NOAA - FCL = - 0.01 ± 0.05 ppm) without any clear concentration dependency. Some larger scatter at lower concentrations may relate to less homogeneous CO_2 isotopic composition for air depleted in CO_2 affecting the isotope sensitive NDIR analysis.

The MENI round robin between NOAA-GML (as WMO-CCL), EMPA (as WMO-WCC), MPI-BGC, FCL and FMI-ATC (ICOS Mobile Lab) has been established to check the ICOS scale link to the WMO mole fraction in a regular manner. In this program a set of three cylinders is prepared and maintained by the FCL. One of these cylinders constitutes a blind sample and is modified in its composition after every completed loop. In Figure 10 results of the first four iterations are shown for FCL, MPI and NOAA. A general small trend in the CO_2 mole fractions has been observed by all labs. To account for the different times of analysis of the comparison samples this trend is defined by the measurements of the CCL as the reference. The mean offset between CCL and FCL is 0.006 ppm (trend revised) and an overall scatter of 0.012 ppm for the period 2017-2020, similar to the agreement established by measurement of other CCL assigned standard gases.

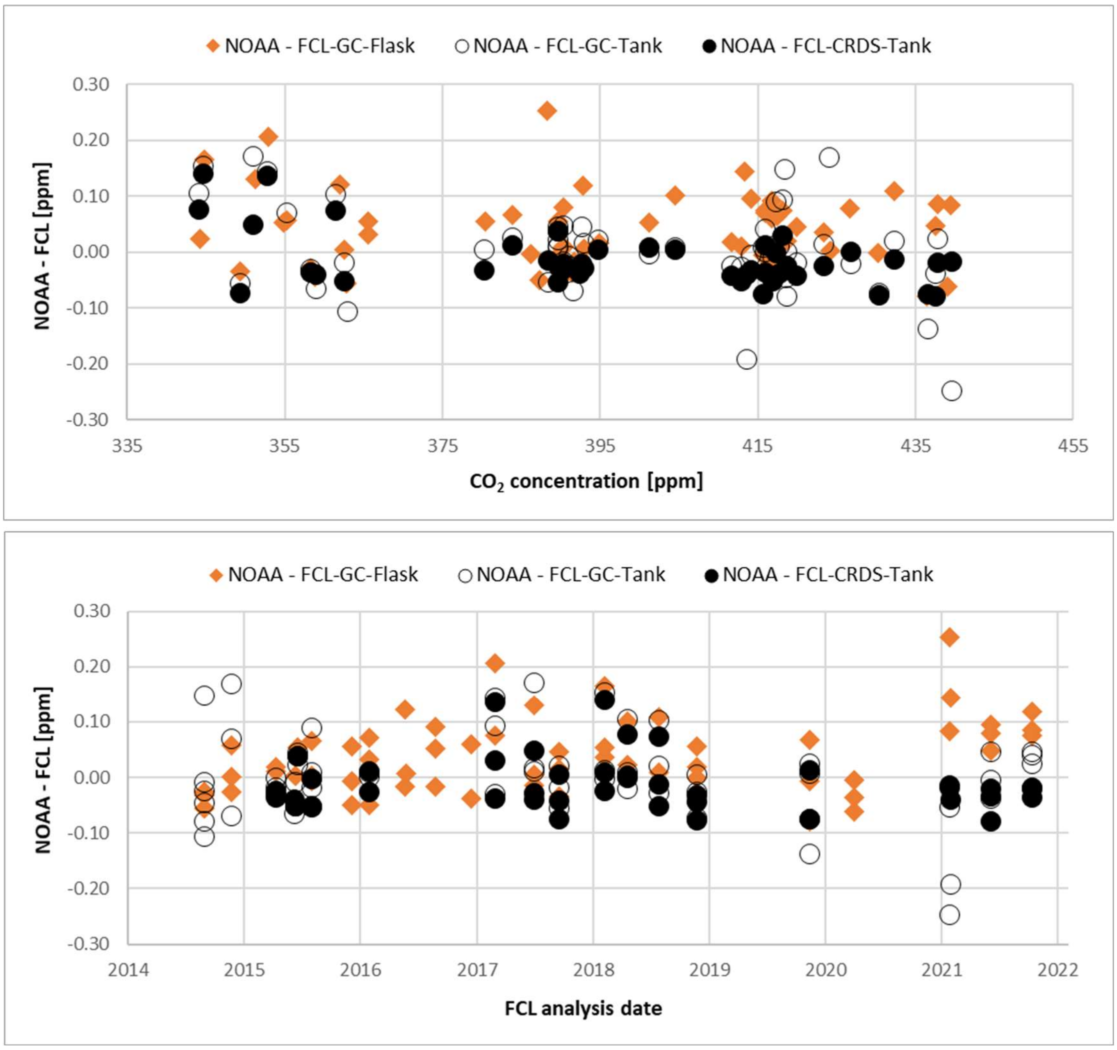
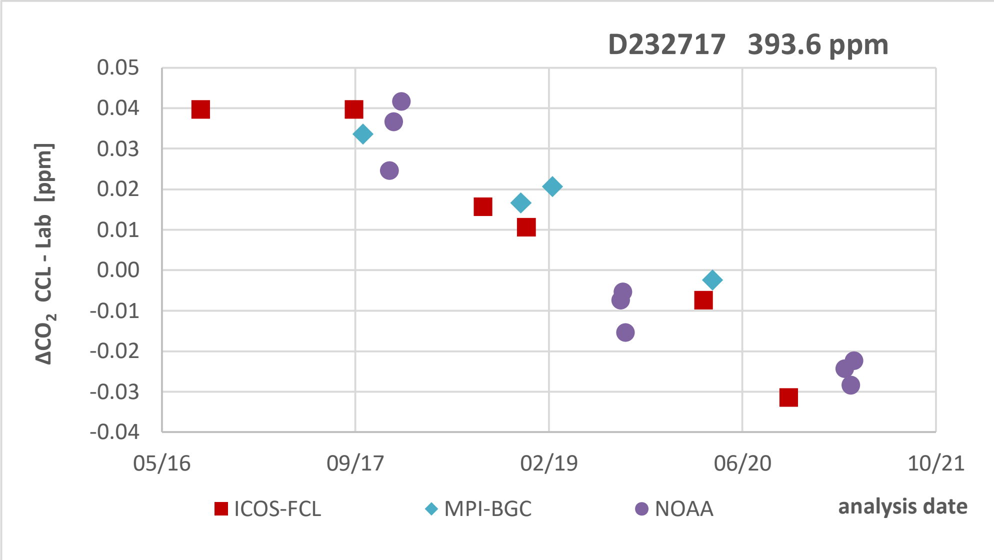


Figure 9: CO_2 offset between FCL and NOAA-GML based on analyses of flask samples (by NOAA), analyses of flask samples by FCL (GC-FID) and their respective source gases from high pressure cylinders (by FCL using CRDS)



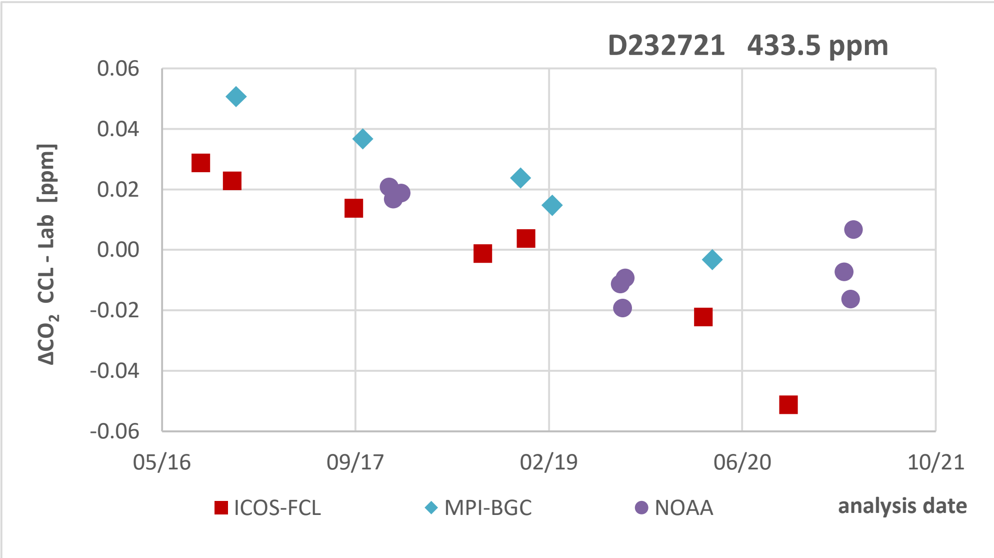


Figure 10: CO_2 offset between NOAA-GML, FCL and MPI relative to NOAA mean of MENI high pressure cylinders.

5.6 CO_2 uncertainty evaluation

The WMO Expert Group recommendations request investigators to report uncertainty estimates for their data that include all potential sources of error [WMO 2018]. A scheme for a comprehensive uncertainty discussion has been suggested by Andrews et al. 2014. Adapting this scheme, we have made such an overall measurement uncertainty estimate based on a performance assessment of the CRDS system. In this assessment we have considered the following uncertainty contributions and checked them using the quality control data of this report.

5.6.1 FCL Primary CO_2 Standards

The X2019 scale revision has solved earlier ambigiues on the stability of the CO_2 mole fractions in the Primary FCL Standards. There are no signs indicating a significant drift in any of the nine standard gases any more. Regression fit residuals of 0.01 ppm confirm the consistency of the standard assignments.

The comparison of FCL measurement results of WMO tertiary standards of other groups (MPI, UBA) also confirm the specified scale transfer uncertainty at the CCL of 0.01 ppm at atmospheric concentration levels. At CO_2 mole fractions beyond a range of 360 - 450 ppm small offsets appear.

The excellent agreement in CO_2 measurements with the NOAA laboratory is ascertained by the results of the ongoing round robin.

5.6.2 CO_2 scale transfer uncertainty

The statistics of repeated calibrations of the FCL Secondary Standards by the FCL Primary Standards provide a measure for the uncertainty of their assignments. The average reproducibility of these assignments is 0.013 ppm for the first set of Secondary Standards, 0.016 ppm for the send set of Secondary Standards. With a total of 20 and 24 calibration episodes, respectively, for each of the Secondary Standards of the first set the uncertainty of their assignments is expected to be below 0.005 ppm relative to Primary Standard set. This is consistent with an average daily calibration standard error of 0.006 ppm. The new set of Secondary standards has been calibrated within 10 calibration episodes, but over a shorter period. This increases the uncertainty in the size of the CO_2 growth within the cylinders. The time series of the target standard CO_2 results suggest a potential overestimate of this drift.

A small negative offset in the Secondary Standard assignments is consistent with the slightly larger difference compared to CCL assignments observed in FCL CO_2 measurement results of the MPI Primary Standards relative to the difference of MPI measurement results compared to CCL assignments of the FCL Primary Standards.

Mean results from the inter-instrument comparison as well as all external comparisons with MPI and NOAA result in average agreements within at maximum 0.015 ppm which also provides evidence for a very small FCL internal scale-transfer uncertainty.

5.6.3 CO_2 long-term reproducibility

The reproducibility as derived from the target standard measurement record is <0.01 ppm from 2015-2020 and may be 0.02 ppm at present (which is expected to decrease once Secondary Standard drift rates can be better constrained). Within the scatter of this time series there are systematic shifts of mean results occasionally observed over periods of many days to weeks to months that are not cancelled out by the standardization scheme. These typically do not exceed 0.005 ppm and point to small system changes over time that are not always understood.

5.6.4 CO_2 measurement uncertainty estimate

Based on these evaluations the following combined standard uncertainty (k=1) is calculated as the square root of the sum of the individual uncertainty squares:

1. Scale link uncertainty = 0.018 ppm:

- The uncertainty from the reproducibility of the CO_2 WMO X2019 CCL CRDS assignments on calibration standards is specified as 0.01 ppm (k=1) [Hall 2021]. This is consistent with the consistency of the regression fit residuals of the FCL Primary Standards. Results from various external comparisons also support that the assignments of the FCL Primary Standard set are well within this uncertainty.
- The uncertainty of the FCL internal scale transfer to the Secondary Standard assignments of the first set deduced from the record of the secondary measurement results from FCL primary calibration events results (uncertainty of the regression of the trend line of CO_2 over time) is estimated as 0.013 ppm (2015 to June 2020). For the second set the change in the growth rates in the target standards' CO_2 time series is indicating a currently larger uncertainty of up to 0.02 ppm.

2. Measurement uncertainty of daily means = 0.014 ppm:

- mean uncertainty of the daily calibration regression fit = 0.01 ppm
- typical uncertainty of unaccounted detector response drift throughout the validity of the daily calibration = 0.009 ppm
- uncertainty from the repeatability of the daily sample measurements =0.0025 ppm (for 15 min means)

3. Additional long-term variability = 0.01 ppm

The reproducibility derived from the target standard record is consistent with the uncertainty estimate for measurement of daily means.

The accuracy with respect to the WMO scale arises from the root of the sum of squares of the scale link uncertainty, the measurement uncertainty and additional long-term variability amounting to 0.02 ppm for calibration data before July 2020 and 0.03 ppm for calibration data since July 2020.

6 CH_4

6.1 FCL Primary CH_4 Standards

6.1.1 CCL CH_4 Assignments

After initial calibration of all FCL Primary Standards in 2014 the first recalibrations of each three of the standards have been made in 2016 and 2017, respectively. In 2018 the final sub-set received its recalibration and in 2021 the complete set was recalibrated again, such that three CCL assignments from different years are available for each standard. In 2017 the CCL has changed instrumentation now using CRDS instead of GC-FID. With the recalibration in 2021, two CRDS measurements are now available for the entire set. For the tanks, the difference in concentrations between the CRDS and the initial values measured with GC-FID lies within the range of the standard deviations specified by the CCL for the individual measurements (range of CRDS-GC-FID difference is -0.28 to 0.57 ppb; see Table 7).

Table 7: CH_4 assignments for FCL Primary Standards by the CCL (WMO X2004A scale)

Sample ID	Cylinder ID	CCL date 1	CCL date 2	CCL date 3	WMO X2004A GC data	WMO X2004A CRDS data*	difference CRDS - GC [ppb]
i20140055	CB09944	Dec-13	May-17	Mar-21	1596.637	1596.682	**0.045**
i20140056	CB09939	Feb-14	Oct-18	Mar-21	1743.128	1743.114	**0.015**
i20140057	CB09958	Dec-13	Aug-16	Mar-21	1896.817	***1896.898***	**0.081**
i20140058	CB09983	Dec-13	Oct-18	Mar-21	2032.923	2032.925	**0.002**
i20140059	CB09952	Feb-14	Aug-16	Mar-21	2195.338	***2195.063***	**-0.275**
i20140060	CB09955	Dec-13	Jun-17	Mar-21	2344.051	2343.898	**-0.154**
i20140061	CB09957	Dec-13	Aug-16	Mar-21	2466.717	***2466.600***	**-0.117**
i20140062	CB09934	Jan-14	Jun-17	Mar-21	2731.277	2731.841	**0.570**
i20140054	CB09948	Jan-14	Oct-18	Mar-21	2932.819	2933.037	**0.218**

* values in bold and italics are indicating results from measurements performed only at date 3

Thus, the recalibrations by the CCL have not changed the assignments significantly and the signs of the update terms for the various standards are such that they largely compensate in sum. Therefore, there is no need for an update of the assigned values used.

6.1.2 Regression fit residuals of FCL Primary CH_4 Standards

The time series of the linear regression fit residuals is presented in Figure 11 for calibration events where the *complete* FCL Primary Standard gas suite was used.

CH_4 mole fractions are known to be generally very stable in aluminium high pressure cylinders. Accordingly, the regression fit residuals do not show significant trends over time for any of the individual standards (generally within 0.2 ppb), which is supporting the assumption of a stable set.

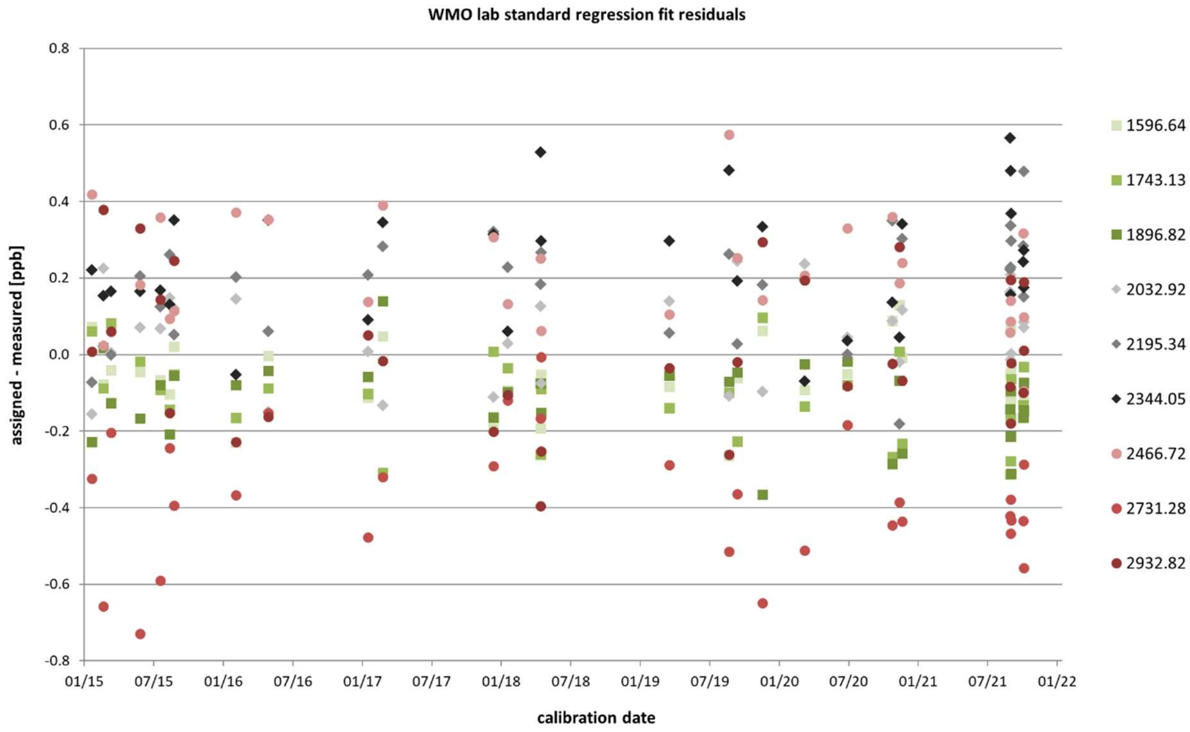


Figure 11: Time series of linear regression fit residuals of CRDS CH_4 calibrations for the individual FCL Primary Standards

6.2 FCL Secondary CH_4 Standards

6.2.1 Assignment record

The four reference gases that were used as initial set of FCL Secondary Standards for the CRDS measurements have been analyzed within 20-24 valid calibration episodes together with the FCL Primary Standards between Feb 2015 and July2020/Sep2021. During 2020, the first set of FCL Secondary Standards had to be replaced by a new set due to consumption. The replacement was done in two stages, with the replacement of the two standard gases with higher mole fractions in June and the replacement of the two standard gases with lower mole fractions in December.

The assigned values for the new standards were also determined by repeated measurements against the FCL Primary Standards (n=10). The assigned values are listed for comparison with the first set in Table 8.

The record of the CH_4 mole fraction results of these FCL primary calibration episodes is displayed in the graphs below. The measured values for the first set of FCL Secondary Standards are shown with dark dots, those for the subsequent second set, which is currently in use, with red diamonds. The reproducibility for both sets shows no difference and is within the order of magnitude of the standard deviations specified by the CCL for the calibration measurements of the primary standards.

For the initial set of Secondary Standards the initial assigned values have not yet been replaced by the mean of the complete set of calibrations given the marginal difference. However, after the replacement of the first two of the initial Secondary Standards in June 2020, updated assigned values were used for the two remaining Secondary Standards of the remaining two initial set for the period until their replacement from June to December 2020.

Table 8: Initial CH_4 assignments of CRDS Secondary Standards

Sample ID	Cylinder ID	Assigned Value	Avg. Ω*	*Date of exchange*	Sample ID	Cylinder ID	Assigned Value
i20140171	D801336	**1795.46 ppb**	1795.51 ppb	*2020-12-08*	i20190708	D761202	**1799.01 ppb**
i20140172	D073384	**1960.24 ppb**	1960.43 ppb	*2020-12-08*	i20190803	D073381	**1949.24 ppb**
i20140173	D073392	**2288.57 ppb**	2288.54 ppb	*2020-06-23*	i20190709	D761214	**2296.10 ppb**
i20140174	D801331	**2092.46 ppb**	2092.41 ppb	*2020-06-23*	i20190438	D073389	**2098.41 ppb**

* Average of all calibration episodes between Feb 2015 and Nov 2020

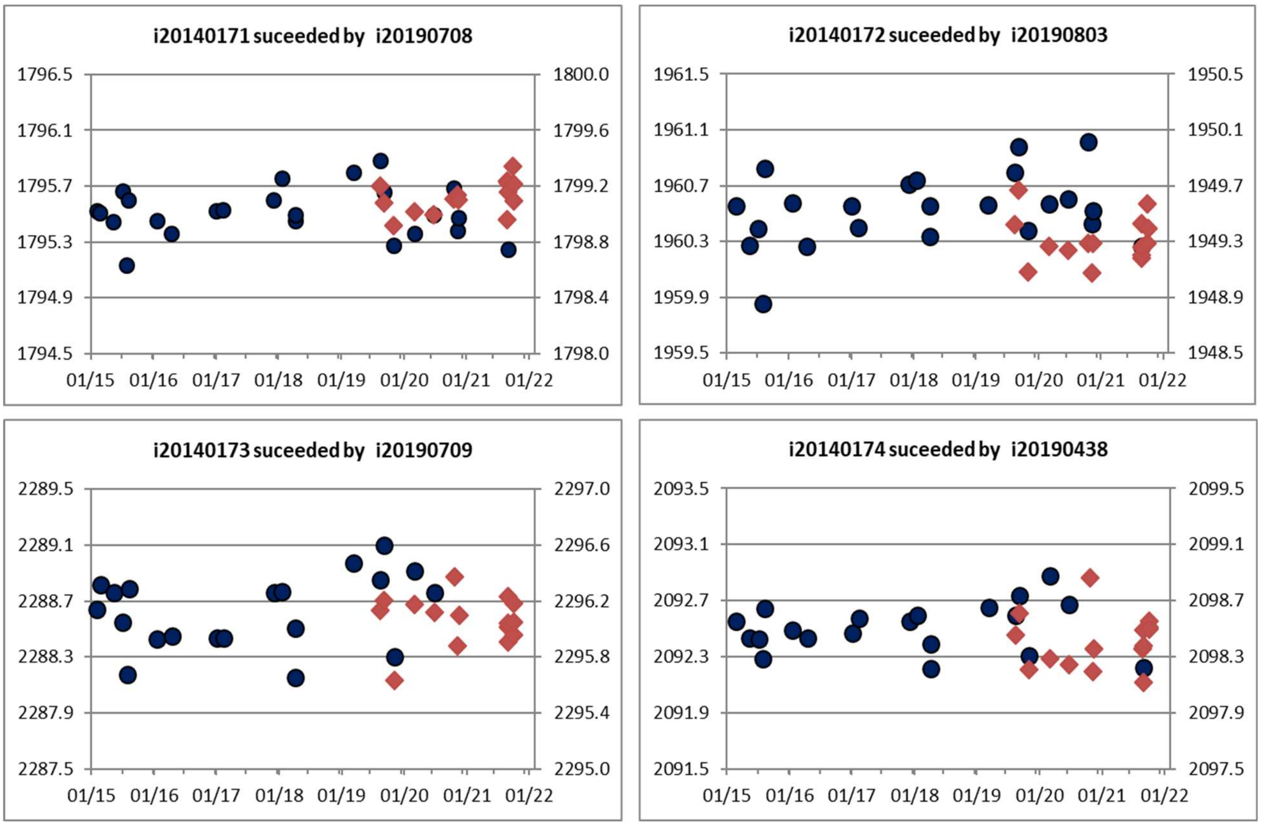


Figure 12: Time series of CRDS CH_4 measurement results for Secondary Standards based on FCL Primary Standards (all values in [ppb]). Dark blue dots represent the assignments for the first set of FCL Secondary Standards, the red diamonds display the four new FCL Secondary Standards.

6.2.2 Residual record

The time series of the residuals of the linear regression fit of the Secondary Standard calibrations are given in Figure 13. The scatter of the residual time series for the individual standards until June 2020 is on average 0.07 ppb without any trend in the residuals being apparent. This documents the long-term internal consistency of the calibration set over time. The replacement of two of the Secondary Standards has resulted in changing residuals for the remaining two standards from the first set by ~0.1 - 0.2 ppb reflecting the small bias in the initial assignments. The internal consistency of the new FCL Secondary Standard set expressed as the standard deviation of the mean residuals is ~0.03 ppb compared to 0.08 ppb for the first FCL Secondary Standard set indicating very little scale transfer uncertainty.

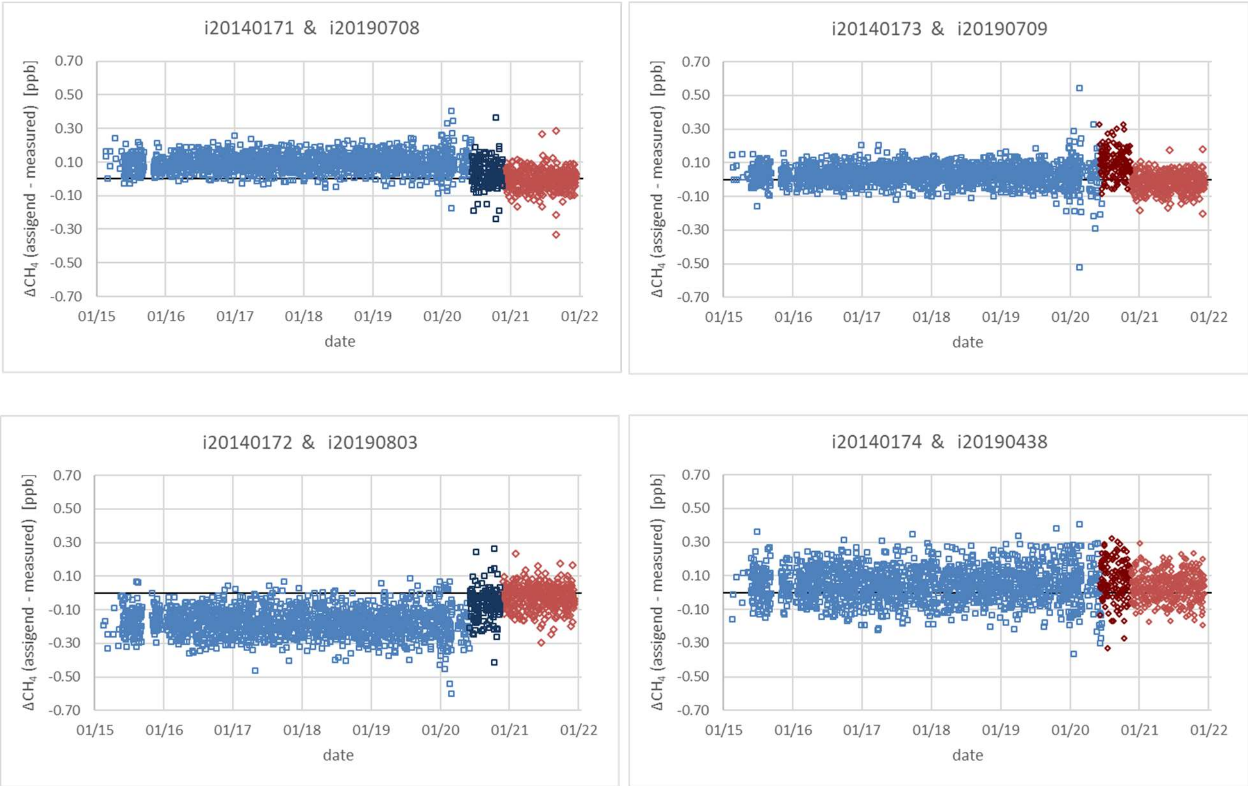


Figure 13: Time series for CH_4 of linear regression fit residuals of the CRDS daily calibration with FCL Secondary Standards The dark symbols are indicating the transition phase when only the first part of the standards was replaced

6.3 CH_4 Targets

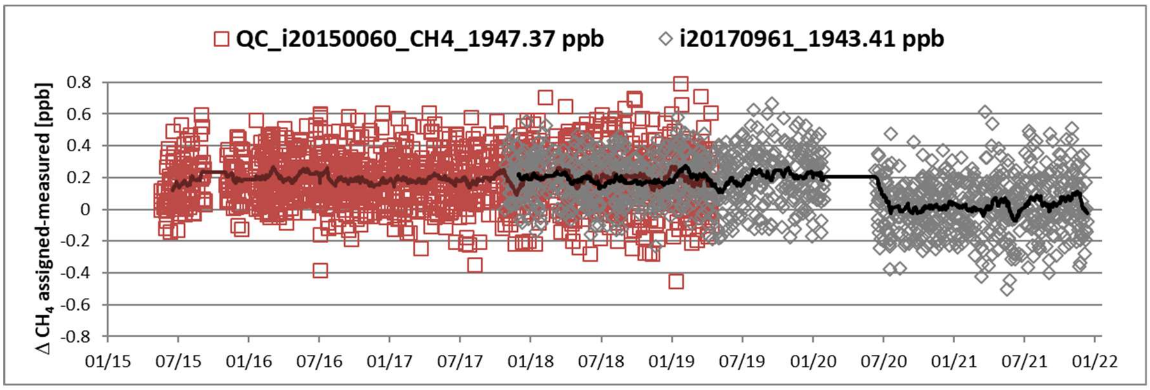
In the period from March 2015 to December 2021 two succeeding sets of each three target standards have been in use for the CRDS system. The replacement was made in 2019 after having analyzed them in parallel for more than 200 analysis days to characterize the transition. On a regular basis two additional Targets monitor the long-term stability of the instrument for concentrations below the calibrated range (1600 ppb and 1700 ppb). The consistent step of 0.2 ppb after changing the Secondary Standard calibration sets apparent in Fig. 14 comply with the offset of the initial assignment of the Secondary Standards made in 2015 relative to results of all primary calibration events of these standards (see section above). Since the exchange of the secondary set the offset has decreased to <0.07 ppb for all targets. This confirms that very little uncertainty contribution results from the scale propagation.

Table 9: Target standards for the CRDS CH_4 analyses

sample ID	tank ID	measured CH_4 [ppb]*	std.dev. [ppb]*	n [days]*	Primary Calibration CH_4 mean [ppb]	std.dev. Calibration mean [ppb]	n calibration values
i20150062	D073391	1914.71	0.17	1303	1914.94	0.16	29
i20150061	D073389	2043.05	0.19	1288	2043.30	0.11	32
i20150060	D073381	1947.18	0.17	1352	1947.37	0.13	82
i20170961	D761211	**1943.21** *1943.39*	**0.12** *0.12*	**739** *551*	1943.41	0.12	31
i20170962	D801332	**2032.71** *2032.90*	**0.12** *0.12*	**787** *537*	2032.94	0.14	25
i20190451	D073391	**2085.97** *2086.17*	**0.12** *0.13*	**300** *547*	2086.27	0.12	19
i20150188	D073398	**1595.55** *1595.70*	**0.10** *0.10*	**160** *30*	1595.72	0.10	26
i20150374	CA05755	**1703.32** *1703.50*	**0.10** *0.11*	**61** *47*	1703.53	0.11	31

* for i20170961, i20170962, i20190451, i20150188 and i20150374 the values are given for the period from start until the change of the FCL Secondary Standards on 23rd June 2020 in bold, after that in italics

The time series of the residuals does not show any significant drift of the targets (Figure 14). But it reveals rare synoptical patterns that introduce variations on the time scale of weeks to months of mostly not more than 0.1 ppb.



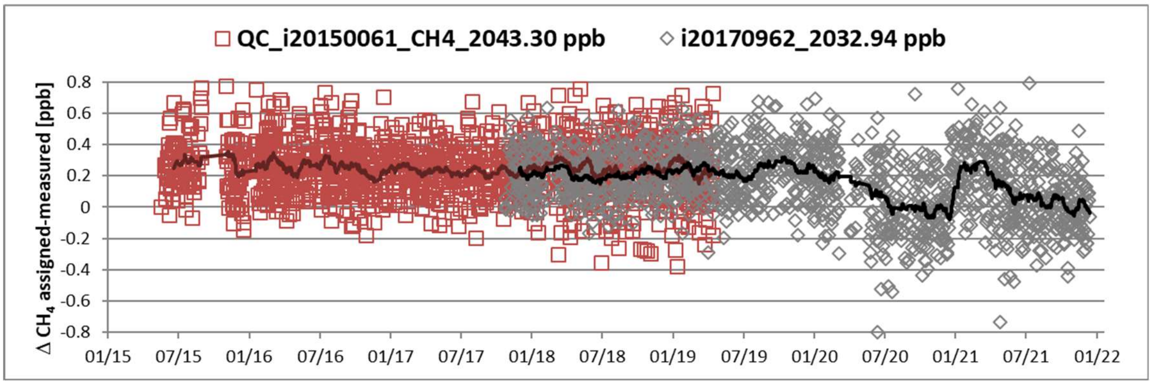
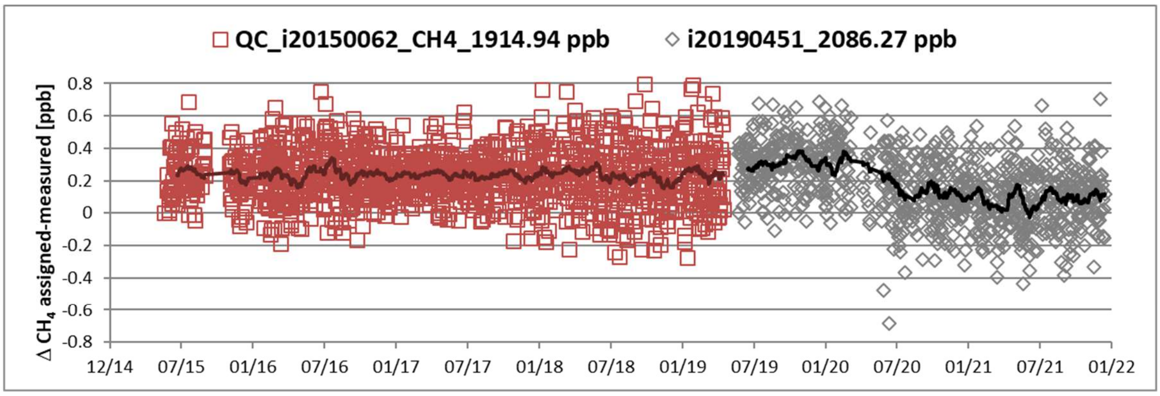
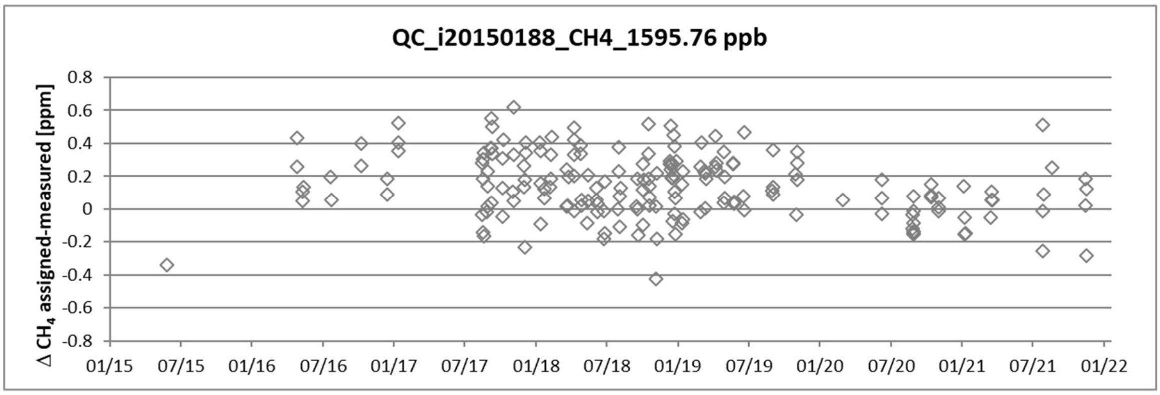


Figure 14: Time series of the CH_4 offset of target measurements to their respective assigned values. The dark line represents a 30 points-running mean.





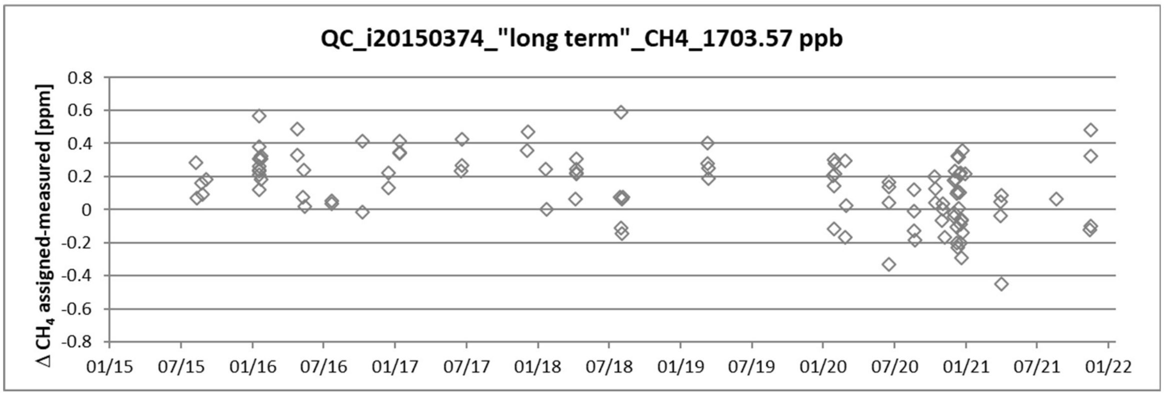
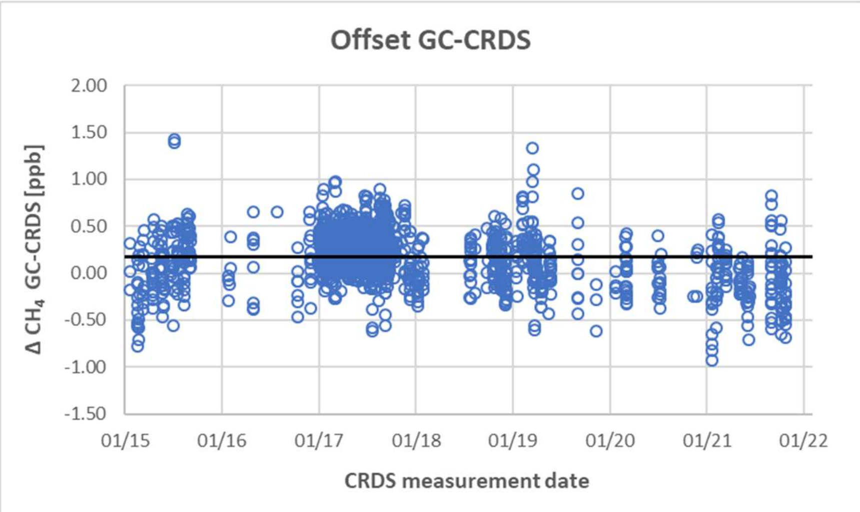


Figure 14: Time series of the CH_4 offset of target measurements to their respective assigned values. The dark line represents a 30 points-running mean.

6.4 Internal CH_4 Comparison: CRDS-GC

Standard gases that are calibrated for CH_4 using CRDS have often also been analyzed by GC-FID. The GC measurements are linked to the same set of Primary FCL Reference Standards but based on a different set of five Secondary Standards. As the reproducibility and typical repeatability of the GC-FID (0.4 ppb and 0.8 ppb, respectively) is approximately by a factor of 3-5 worse than that of the CRDS instrument, only GC measurements have been considered that have been analyzed on the GC on more than one day with at least 20 injections. The inter-instrumental measurement differences for all samples are depicted in Figure 15 (only standards within the range defined by the calibration standards were considered). The average offset is 0.20 ppb ± 0.25 ppb for the initial phase until the change of the FCL Secondary Standards on 23rd June 2020, from that date onwards about -0.08 ppb ± 0.30 ppb, which again is of the same magnitude as the small bias of the initial CRDS Secondary Standard assignments.



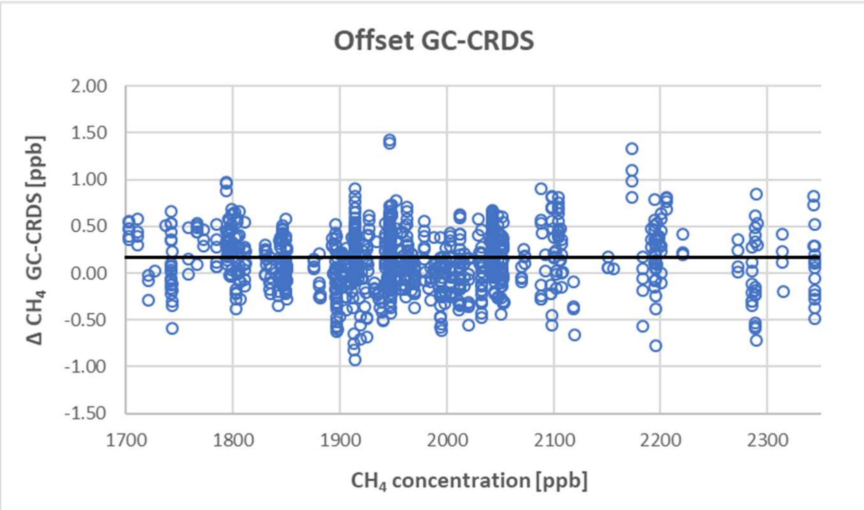


Figure 15: Offset of CRDS daily mean CH_4 results relative to the GC average result of the same sample. The black line represents the mean offset.

Note that each data point in Figure 15 represents the difference of one CRDS daily mean result relative to the annual means of all GC measurements of the same sample. Some samples have been analyzed much more frequently than others (e.g. target standards) which explain the occurrence of many clustered data points in the Figure. Overall, the comparison with the independent GC measurements does not indicate any significant error in the CRDS measurements that might have been missed.

6.5 External CH_4 Comparisons

6.5.1 CH_4 compatibility ICOS FCL - MPI-BGC

The most intensive external comparison measurements have been made with the MPI-BGC GasLab. This laboratory is using different instrumentation (Picarro G1301 through April 2018, G2301 since May 2018) and their measurements are tied to the WMO Mole Fraction scales by an independent set of Primary Standards. These MPI-BGC Primary Standards already have CCL calibration records with multiple measurements in different years for nine individual standards over six to seventeen years.

The MPI-BGC measurements are not relevant for the assignment of the FCL standards and are therefore completely independent.

6.5.1.1 Comparison of CH_4 Primary Standards

Basis for an agreement of FCL and MPI-BGC measurements is the compatibility of the respective sets of Primary Standards. As the FCL Primary Standards have been produced at the MPI-BGC they also were thoroughly analyzed at the MPI-BGC in 2013 and 2014 before being used by the FCL. Before or after the shipment to the CCL for recalibration of sub-sets of this FCL Primary Standard gas suite these standards were also analyzed for a third time. Likewise, MPI-BGC Primary Standards that were simultaneously returned to the CCL for recalibration were also analyzed by the FCL. These data are shown in Figure 16.

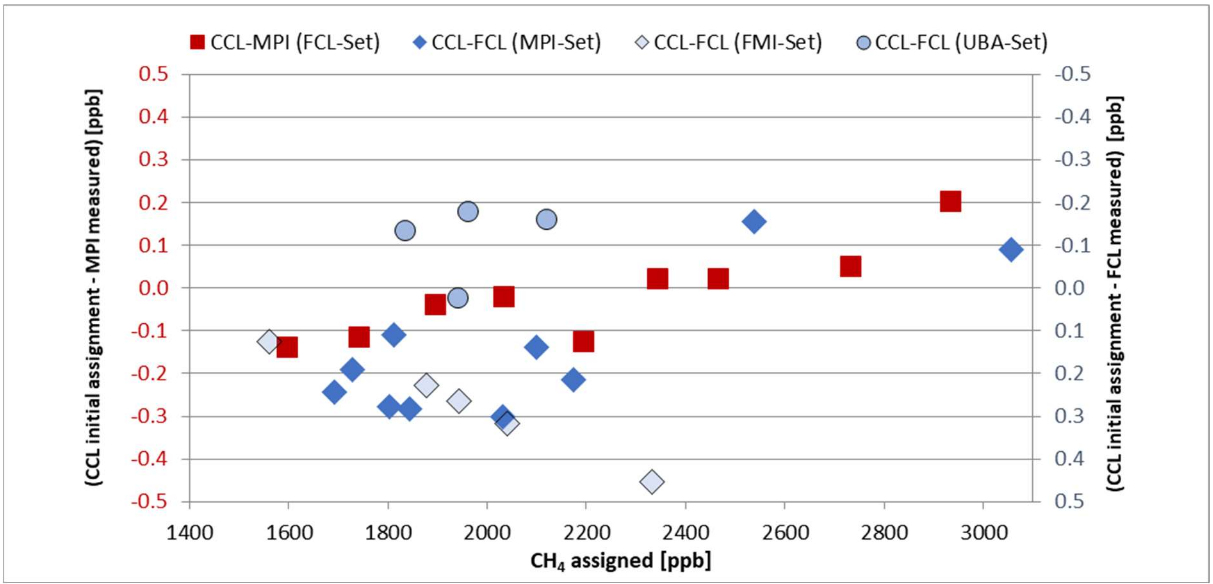


Figure 16: Differences of measured CH_4 results to CCL assigned values for FCL Primary CH_4 Standards.

The mean difference of the MPI-BGC measurement results relative to the CCL assignments is <0.02 ppb and the differences for the individual standards closely follow the regression fit residuals observed with the FCL Picarro CRDS instrument (see section above). The mean difference of the FCL measurement results relative to the CCL assignments has been 0.19 ppb before June 2020 (using MPI-Set and FMI-Set) and 0.11 ppb in 2021 (using the UBA set). This is fully consistent with the findings in the previous sections and confirms the excellent accuracy of the CH_4 CCL assignments.

6.5.1.2 Comparison of FCL Secondary CH_4 Standards and Target standards

Three of the four gases from the first set of FCL Secondary Standards have been analyzed at the MPI as well as three of the target standards. The differences between MPI-BGC measurement results and FCL assignments (Figure 17, blue symbols) are very consistent to the difference of FCL measurement results of the MPI-BGC Primary CH_4 Standards.

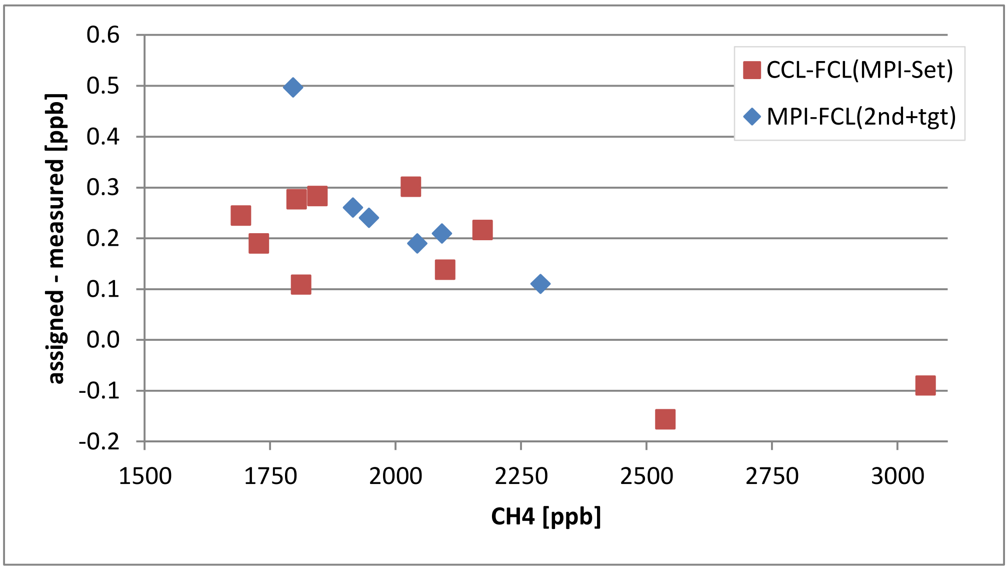


Figure 17: Differences of MPI-BGC measured results to FCL Secondary Standard assigned CH_4 values (blue diamonds) compared to the differences of FCL measured results relative to CCL CH_4 assignments of MPI-BGC Primary Standards (red squares)

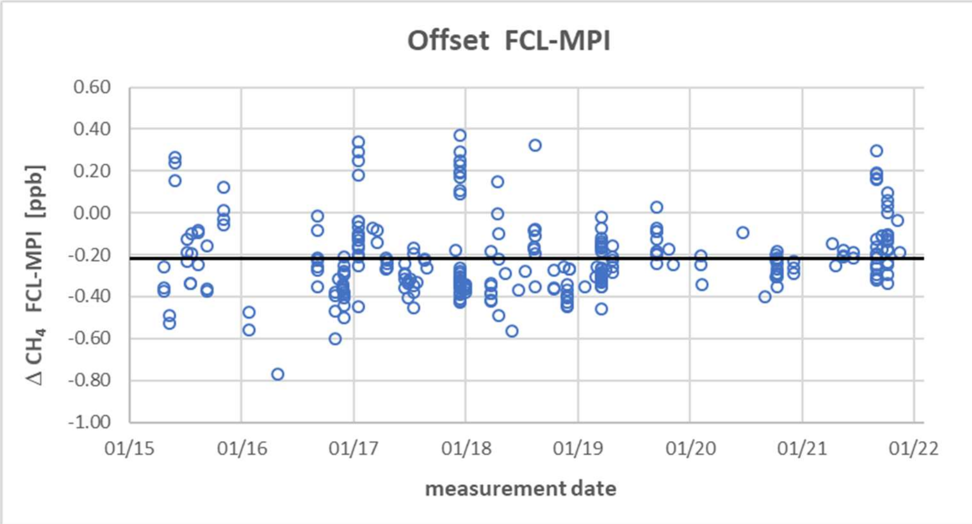
The mean differences of FCL-assigned values (based on the initial calibrations with the FCL Primary Standards for the secondaries but accounting for all calibrations of the targets), the FCL measured means and the MPI-BGC measured means are given in Table 10. As seen in Fig. 17 MPI-BGC measurement results show a difference on average 0.2 ppb to the assigned values of the Secondary Standards and the measured values of the targets. However, they do not show any difference to the FCL assignments (0.0 ± 0.1 ppb).

Table 10: Comparison of MPI-BGC analysis results and FCL for target cylinders

FSN	Cylinder	$FCL_{assigned}$	$FCL_{measured}$	$MPI_{measured}$	$MPI\text{-}FCL_{assigned}$	$MPI\text{-}FCL_{measured}$
i20140171	D801336	1795.46		1795.93	0.47	
i20140173	D073392	2288.57		2288.72	0.15	
i20140174	D801331	2092.46		2092.70	0.24	
i20150060	D073381	1947.37	1947.18	1947.42	0.05	0.24
i20150061	D78910	2043.30	2043.05	2043.24	-0.06	0.19
i20150062	D073391	1914.94	1914.71	1914.97	-0.03	0.26

6.5.1.3 Sample CH_4 comparison

High pressure standards have been regularly exchanged between MPI-BGC and FCL and analyzed in both laboratories. The difference in results for about 280 comparisons is presented in Figure 18. The average offset of all MPI-FCL sample comparisons amounts to 0.22 ppb ± 0.18 ppb. This difference confirms the observed offset in the Secondary Standard assignments.



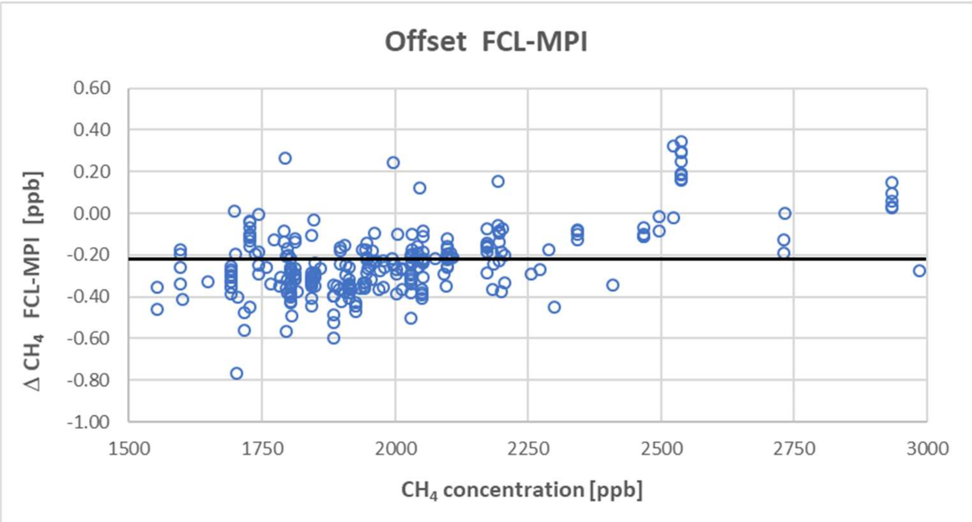


Figure 18: Differences of CH_4 measurement results for samples that have been analyzed at FCL and MPI-BGC

6.5.2 CH_4 compatibility ICOS FCL - NOAA

Comparison with the NOAA-GML laboratory (and other laboratories) is routinely made in two independent exercises, the Sausage Flask Intercomparison Program and the MENI high pressure cylinder program. In the Sausage intercomparison samples for comparison are prepared by connecting sets of flasks in line and filling them with dry air from a high-pressure cylinder at the FCL. The FCL generally analyzes the composition of the filling air using the normal instrumentation for calibrating standards. Therefore, the results of the flask measurements provided by NOAA can be compared to these high-pressure cylinder measurements. The respective data are compiled in the following figures. The CH_4 offset of all samples is NOAA - FCL = 0.42 ppb ± 0.53 ppb. In 2019 NOAA has changed the instrumentation for flask analysis to a CRDS system; constraining the comparisons to data since 2021 (after the change in the FCL Secondary Standards) results in a CH_4 offset of NOAA - FCL = 0.11 ppb ± 0.23 ppb.

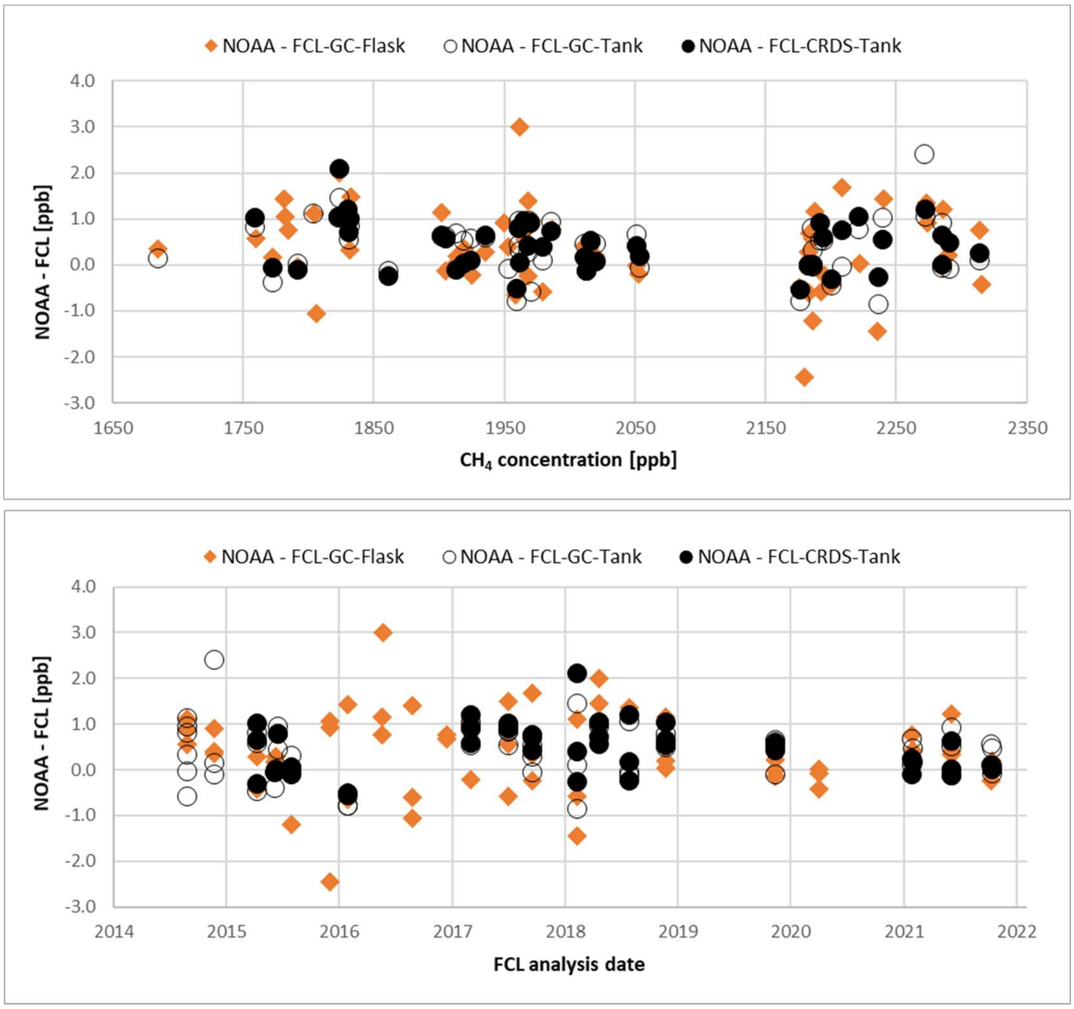


Figure 19: CH_4 offset between FCL and NOAA-GML based on analyses of flask samples (by NOAA), analyses of flask samples by FCL (GC-FID) and their respective source gases from high pressure cylinders by FCL (CRDS)

The MENI round robin between NOAA-GML (as WMO-CCL), EMPA (as WMO-WCC), MPI-BGC, FCL and FMI-ATC (ICOS Mobile Lab) has been established to check the ICOS scale link to the WMO mole fraction in a regular manner. In this program a set of three cylinders is prepared and maintained by the FCL. One of these cylinders constitutes a blind sample and is modified in its composition after every loop completed. In Figure 20 results of the first four circulations are shown. The observed offset between CCL and FCL is 0.15 ± 0.10 ppb.

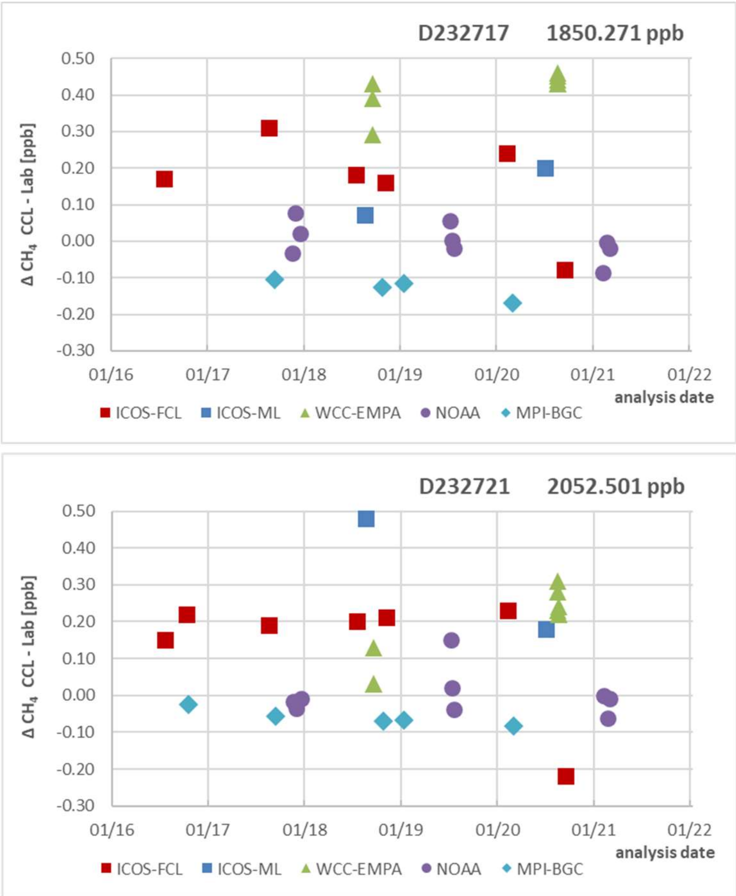


Figure 20: CH_4 offset between NOAA-GML, FCL and other labs' analyses results of MENI high pressure cylinders

6.6 CH_4 uncertainty evaluation

The WMO Expert Group recommendations request investigators to report uncertainty estimates for their data that include all potential sources of error [WMO 2018]. A scheme for a comprehensive uncertainty discussion has been suggested by Andrews et al. 2014. Adapting this scheme we have derived an overall measurement uncertainty based on a performance assessment of the CRDS system. In this assessment we have considered the following uncertainty contributions and checked them using the quality control data of this report.

6.6.1 FCL Primary CH_4 Standards

According to all available evidence CH_4 mole fractions within the FCL Primary Standards are accurately assigned and stable with all metrics (re-calibration by the CCL, repeated analysis by the MPI-BGC, consistency of regression fit residuals) pointing to a consistency of 0.2 ppb. For this evaluation, however, we consider the uncertainty specification of the scale propagation to individual standard gases at the CCL as 0.5 ppb (k=1), pers. comm. E. Dlugokencky, Feb. 2018).

6.6.2 CH_4 scale transfer uncertainty

The statistics of repeated calibrations of the FCL Secondary Standards by the FCL Primary Standards provide a measure for the uncertainty of their assignments. The reproducibility of these assignments is 0.2 ppb. However, the initial assignments in 2015 have been based on a set of calibration events that turned out to be all lower by 0.1 - 0.3 ppb than the mean results from all calibration episodes. This finding of such a marginal offset in the FCL Secondary Standards' CH_4 mole fractions is quantitatively confirmed by the comparison FCL measurement results of standard sets assigned by the CCL for other laboratories (MPI-BGC and FMI). It is also consistent with the offsets observed up to 2020 in various comparisons including the MENI intercomparison with NOAA. With the replacement of the FCL Secondary Standard set when the first set from 2014 was exhausted this offset has been remedied. It could be considered to update the initial assignments as all gases of this Secondary Standard set have received their final calibration in 2021. However, the update terms are very minor. This will be discussed with ICOS ATC and ICOS Atmosphere MSA.

6.6.3 CH_4 long-term reproducibility

The reproducibility derived from the target standard measurement record is 0.2 ppb (standard deviation of daily means). Within the scatter of this time series there are occasional systematic shifts of mean results observed over periods of many days to weeks to months that are not cancelled out by the standardization scheme. These typically do not exceed about 0.2 ppb and point to additional uncertainty arising from small system sensitivity changes that are not always understood.

6.6.4 CH_4 measurement uncertainty estimate

Based on these evaluations the following combined standard uncertainty (k=1) is calculated as the square root of the sum of the individual squared uncertainty contributions:

1. Scale link uncertainty = 0.54 ppb

- uncertainty of the FCL Primary Standards set based on CCL assignments =0.5 ppb
- uncertainty of the FCL internal scale transfer to FCL Secondary Standard assignments =0.2 ppb (both for first and second Secondary Standard sets (n=22 and n=10 calibration episodes, resp.).

2. Measurement uncertainty of daily means = 0.18 ppb

- mean uncertainty of the daily calibration regression fit = 0.1 ppb
- uncertainty of the detector response drift throughout the validity of a daily calibration ≈ 0.15 ppb
- uncertainty from the repeatability of the daily sample measurements = 0.03 ppb (for 15 min means)

3. Additional long-term variability = 0.2 ppb

The accuracy with respect to the WMO Mole Fraction scale arises from the root of the sum of squares of the scale link uncertainty, the measurement uncertainty and additional long-term variability amounting to 0.6 ppb.

The internal reproducibility is estimated to be 0.3 ppb which is consistent with the results from the target standard record.

7 CO

7.1 FCL Primary CO Standards

7.1.1 CCL CO Assignments

After initial calibration of all FCL Primary Standard gases in 2014 the first recalibrations of each three of the standards have been made in 2016, 2017, and 2018. In December 2020 the complete set of FCL Primary Standards was sent to the CCL for the third calibration, performed in March 2021.

The CCL calibration record of the FCL Primary Standards is summarized in Table 11 indicating that the increase in CO exhibits a concentration dependency with standards with lower CO mole fraction having a larger increase in CO. While all initial CCL assignments have been made based on measurements with the LGR2 instrument, not all of the first recalibration measurements were made using this instrument but one third was re-assigned using the V3 Aerolaser VURF analyzer only. For the last recalibration both instruments were used for all standard gases. Figure 21 shows that systematically larger CO values result from the LGR2 measurements for all standards with CO below 400 ppb compared to VURF results. The effect is under investigation at the CCL [CCL_CO 2018].

Growth of CO in high pressure aluminium cylinders is a known limitation for accurate CO measurements that has to be accounted for. To account for the increasing CO concentrations in FCL Primary Standards we applied a linear interpolation between the initial and the second calibration data point for every standard where the increase exceeded the analytical uncertainty of the CCL calibrations. This includes all standards with CO below 210 ppb. Based on these new, drifting assigned values, the calibration results of the FCL Secondary Standards were recalculated. Lab-internally, all data sets have been updated to take the CO growth in the reference standards into account. Data presented here are therefore also compensated for this drift.

A further refinement of the CO growth in the Primary Standards would be possible with the third CCL assignment. However, this has not been performed because a re-assignment would not only entail a re-processing of all FCL-CO calibration measurements but also require a re-computation of all ICOS atmospheric CO data. At this point this effort appears not justified as Fig. 21 shows that the FCL measurement results of the

Table 11: CO assignments for FCL Primary Standard gases by the CCL (WMO X2014A scale)

Sample ID	Cylinder ID	CCL date 1	CCL date 2	CCL date 3	WMO X2014A	WMO X2014A	WMO X2014A	corresponding drift [ppb/yr]
i20140055	CB09944	Dec-13	Jun-17	Mar-21	31.310	*34.414*	36.754	**0.750**
i20140056	CB09939	Feb-14	Oct-18	Mar-21	80.656	82.872	84.465	**0.538**
i20140057	CB09958	Dec-13	Sep-16	Mar-21	120.693	122.362	125.030	**0.598**
i20140058	CB09983	Dec-13	Oct-18	Mar-21	158.920	161.614	162.726	**0.525**
i20140059	CB09952	Feb-14	Sep-16	Mar-21	199.466	200.773	203.125	**0.516**
i20140060	CB09955	Dec-13	Jul-17	Mar-21	247.136	*247.877*	249.934	**0.386**
i20140061	CB09957	Dec-13	Sep-16	Mar-21	397.062	396.194	399.389	**0.321**
i20140062	CB09934	Jan-14	Jun-17	Mar-21	697.558	*697.724*	697.099	**-0.064**
i20140054	CB09948	Jan-14	Oct-18	Mar-21	998.629	1001.751	998.128	**-0.070**

Drift calculated on period CCLdate3 - CCLdate1
Values without V3 if possible (if only V3 available, these are in italics)

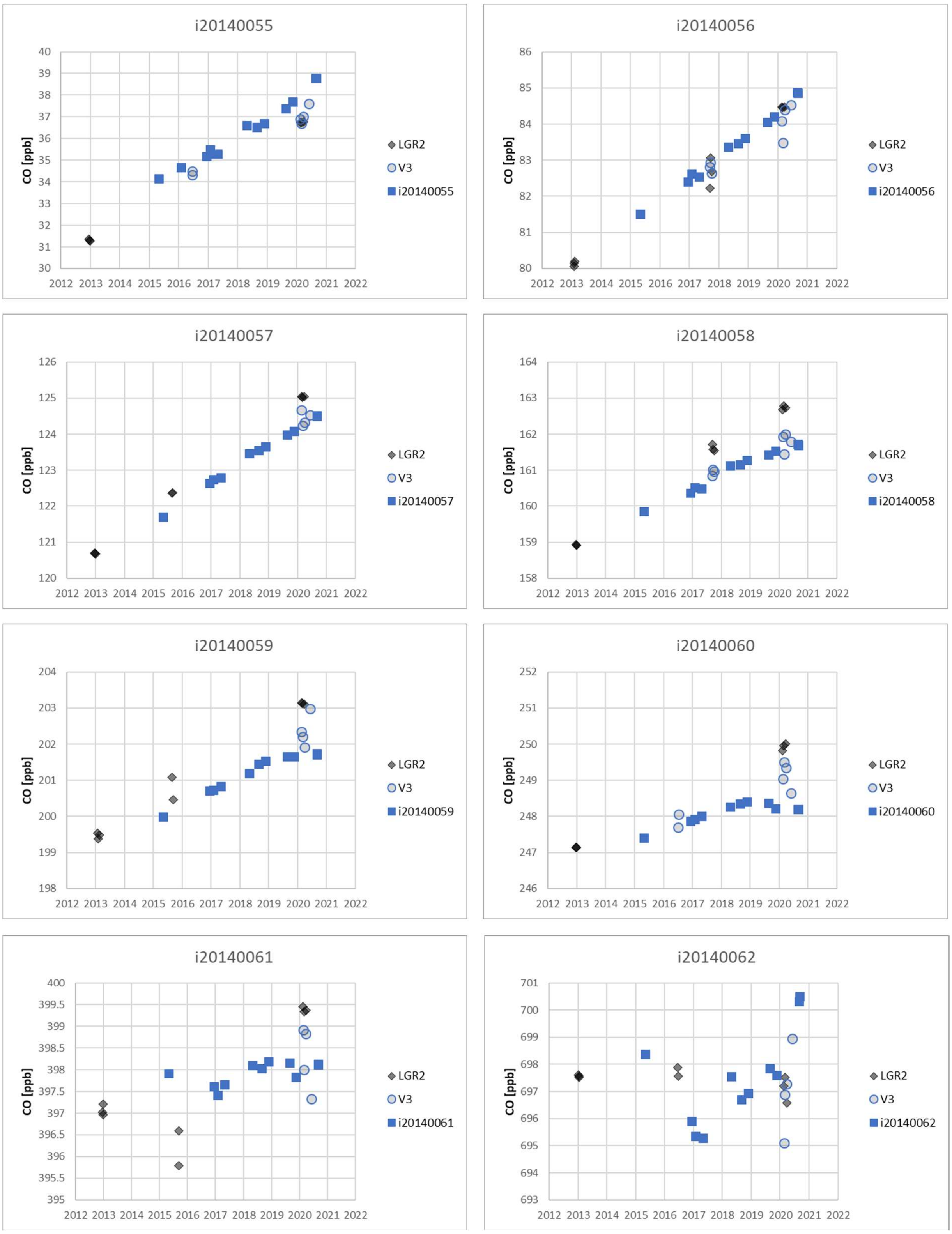


Figure 21: CO Primary Standards, measured at the FCL (blue squares) and at the CCL with the two different analysers LGR (grey diamonds) and Aerolaser (light blue circles)

Primary Standards (blue squares) are generally well in line with the trend arising from the second and third CCL calibration with differences not exceeding the offset between the results from different instruments used at the CCL (LGR (grey diamonds) and Aerolaser (light blue circles)).

7.1.2 Regression fit residuals of FCL Primary CO Standards

The time series of the regression fit residuals displayed in Figure 22 shows very consistent results but with trends on the order of 0.05 – 0.2 ppb/yr for the individual standard gases. This reflects the limited accuracy of the applied trend functions. This is partly result from only two CCL calibration results having been available to assign the trend lines. On the one hand, information has been presented by the CCL that the WMO Mole Fraction scale Primary Standards drift at a different rate as the values used for calibrating tertiary standards (https://www.esrl.noaa.gov/gmd/ccl/co_scale_update.html). Whereas this drift may be overcorrected at low CO mole fractions, it is below the CCL defined significance level for the high CO standards.

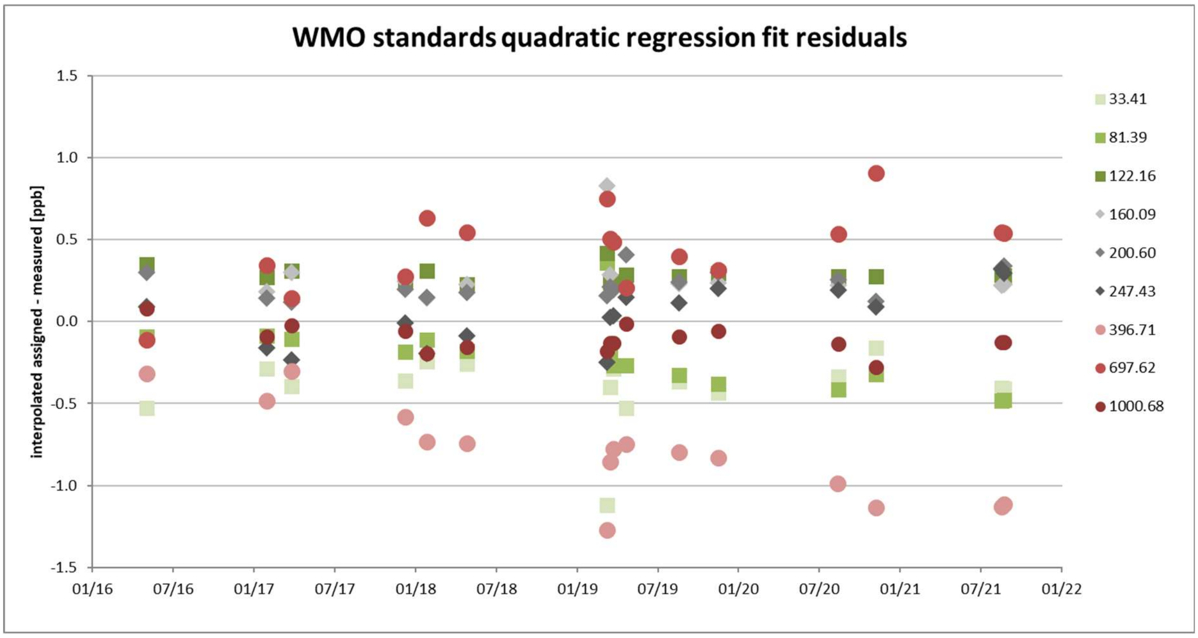


Figure 22: Time series for CO regression fit residuals for CO of FCL Primary Standard calibrations of the LGR instrument

7.2 FCL Secondary CO Standards

7.2.1 Assignment record

The four reference gases that are used as FCL Secondary Standards for the daily LGR calibration have been analyzed within 24 to 25 calibration episodes together with the complete set of FCL Primary Standards between May 2016 and Sep 2021. In May 2018 one of the Secondary Standard gases was exhausted (i20150251; five calibration episodes together with FCL Primary Standards) and was succeeded by a new standard with a similar CO content (i20170889; 21 calibration episodes together with FCL Primary Standards). The three remaining tanks were replaced when they were exhausted in July 2021. For the three replacements, the assigned values and drift rates were determined by measurements against the old set of FCL Secondary Standards (the number of comparison measurements is: i20201308 (n=114), i20201254 (n=114), i20201255 (n=117)). Table 12 summarizes the initial assignment values and the CO growth rates for all tanks.

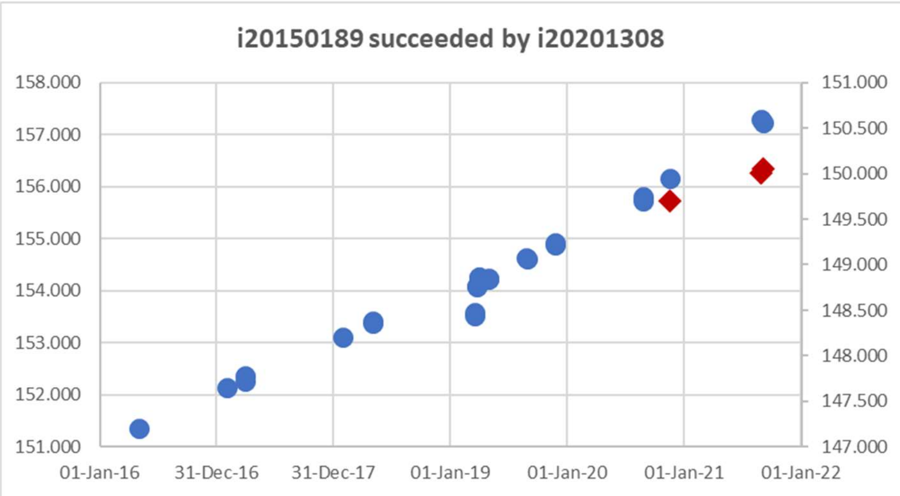
Table 12: CO assignments for FCL Secondary Standard gases (WMO X2014A scale)

Sample ID	Cylinder ID	Assigned Value[1]	drift/yr[1]	*Date of exchange*	Sample ID	Cylinder ID	Assigned Value[2]	drift/yr[2]
i20150251	CA05640	**78.52 ppb**	**+0.76 ppb**	*2018-05-03*	i20170889	D557226	**79.63 ppb**	**+1.28 ppb**
i20150189	D073397	**150.79 ppb**	**+0.97 ppb**	*2021-07-26*	i20201308	D753834	**149.53 ppb**	**+0.25 ppb**
i20150191	D073395	**305.80 ppb**	**+0.72 ppb**	*2021-07-26*	i20201254	D753835	**293.54 ppb**	**-0.16 ppb**
i20150544	D073396	**433.52 ppb**	**+0.51 ppb**	*2021-07-26*	i20201255	D753836	**423.90 ppb**	**-0.07 ppb**

[1] Starting date: 1st November 2015

[2] Starts for i20170889 on 1st October 2017, for the other three tanks on 26th July 2021

In Figure 23 the record of the Secondary Standards' CO mole fraction results of these FCL primary calibration episodes is presented based on time dependent assigned values for the FCL Primary CO Standard gases. It appears that the two high standards are relatively stable in CO in the currently used set of Secondary Standards.



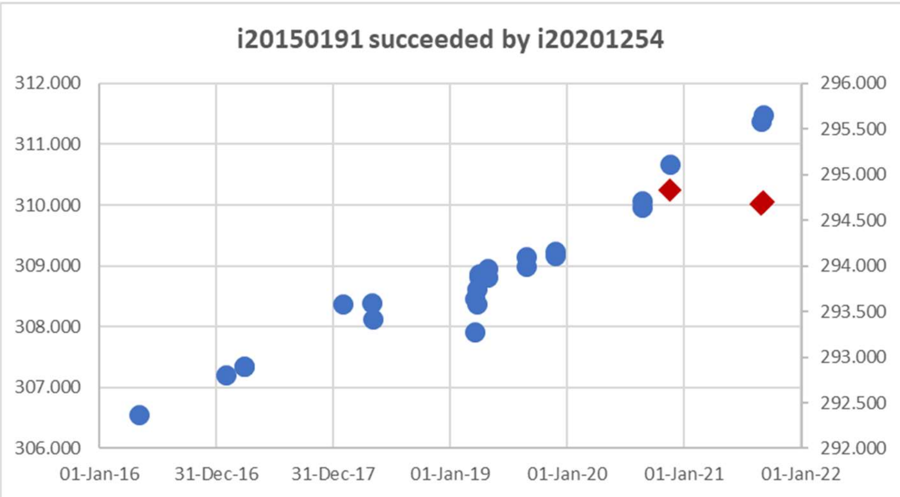
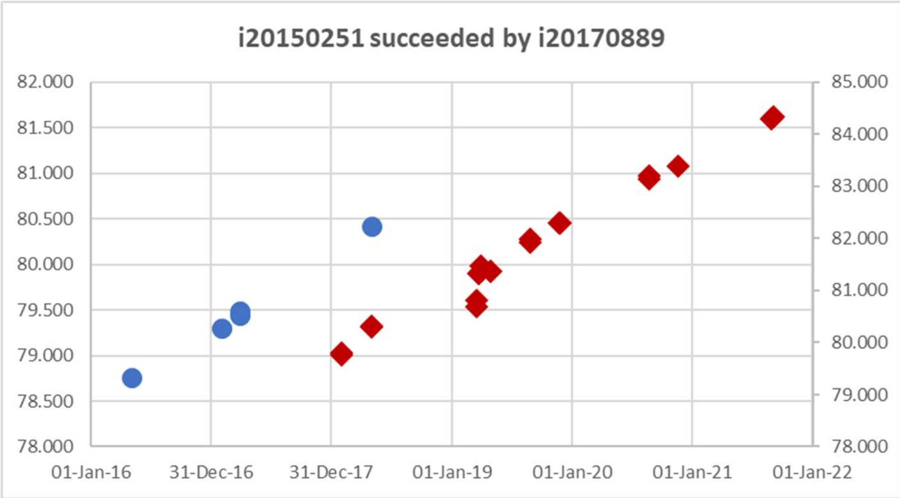


Figure 23: Secondary CO standard assignment record (values in [ppb]).



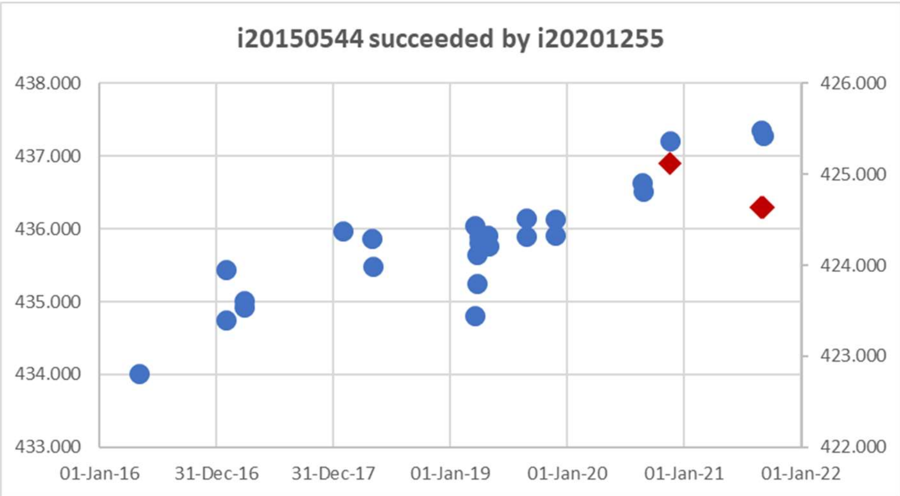
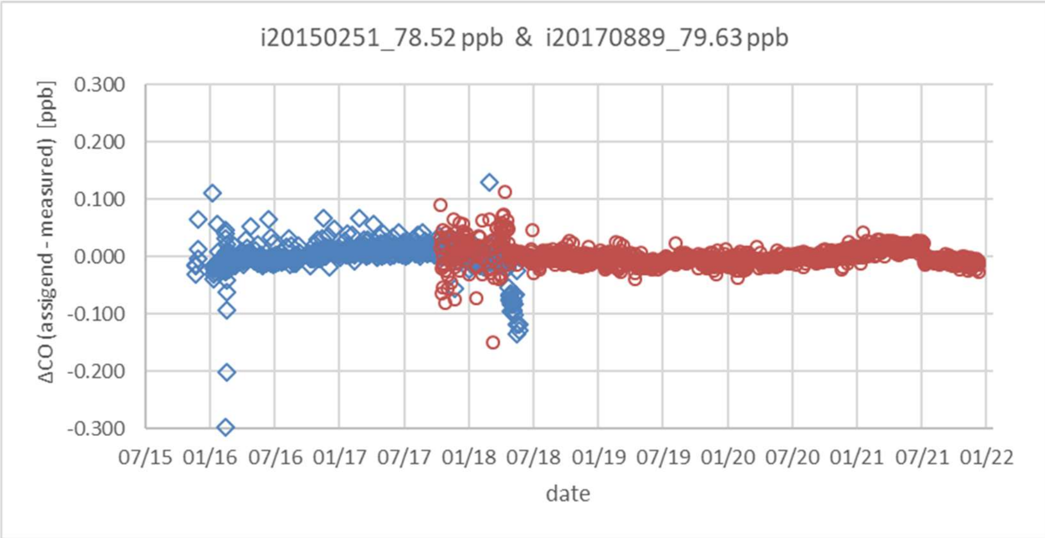
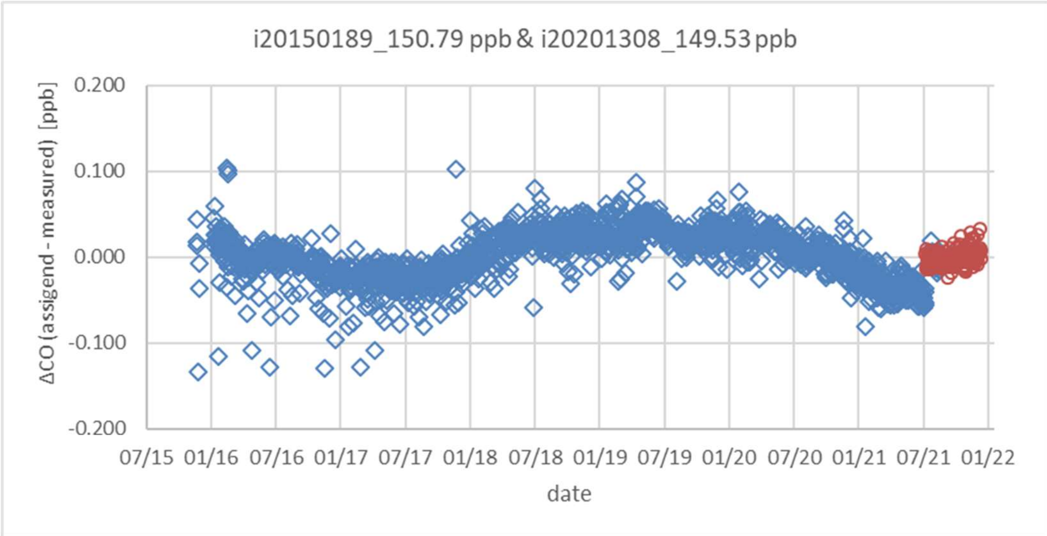


Figure 23: Secondary CO standard assignment record (values in [ppb]).

7.2.2 Residual record

The residuals of the quadratic regression of the FCL Secondary Standards are displayed in Figure 24. These residuals document an excellent consistency of this reference gas set. The time series also illustrates that the reproducibility is not at all limited by the instrumental precision capability. Note, though, that the changes in these plots only reflect the relative changes between the FCL Secondary Standards and do not allow deducing any absolute trends.





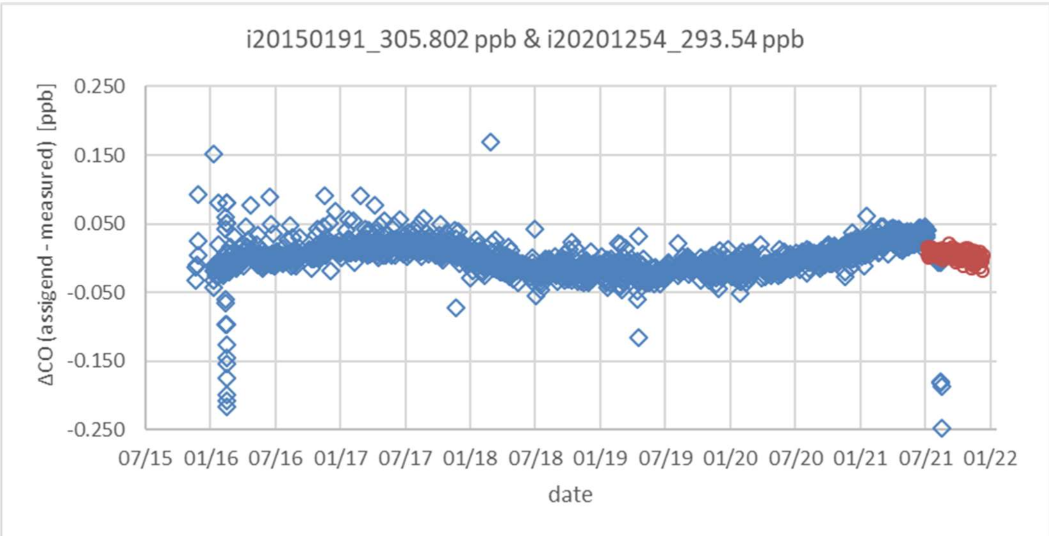


Figure 24: Time series of quadratic regression fit residuals of the LGR CO calibration using FCL Secondary Standards with drifting assignments based on the revised Primary Standard assignments

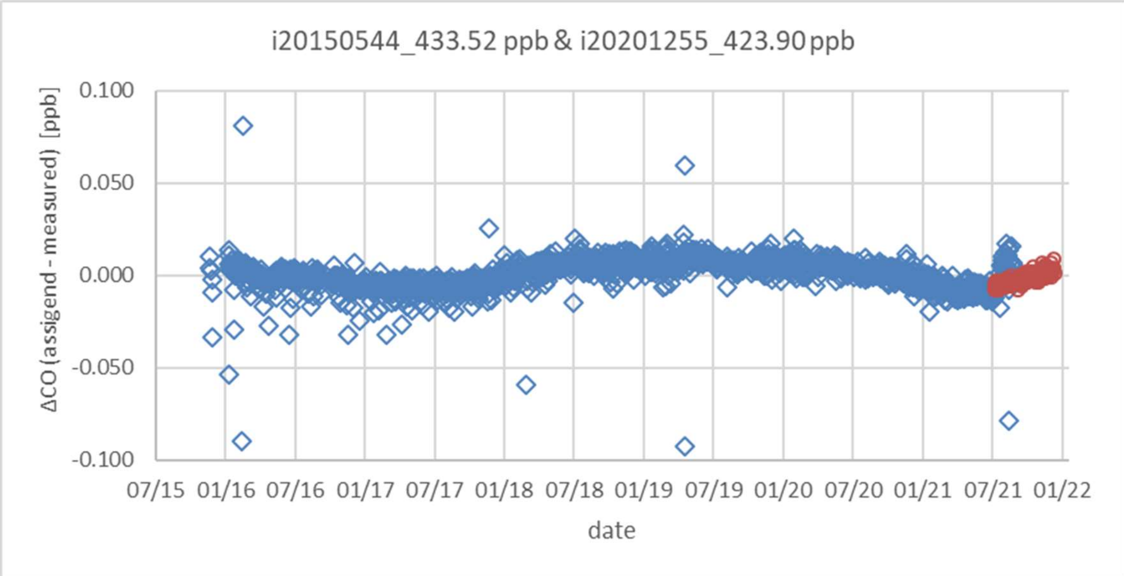


Figure 24: Time series of quadratic regression fit residuals of the LGR CO calibration using FCL Secondary Standards with drifting assignments based on the revised Primary Standard assignments

7.3 CO Targets

In the period from Feb 2016 to December 2021 two targets have been in use on the LGR system. They are complemented by three Long Term Targets (also Archive Quality Control). These are measured less frequent after an initial daily analysis frequency to maintain a long-term link of succeeding targets in future. The record of daily measurement results is presented in in Figure 25.

As for the reference standards also the target standards exhibit a steady increase of CO for those gases with the concentration below 300 ppb. This trend appears mostly linear but the systematic curvatures of the detrended grey symbols might indicate that the initial growth rate is slightly higher.

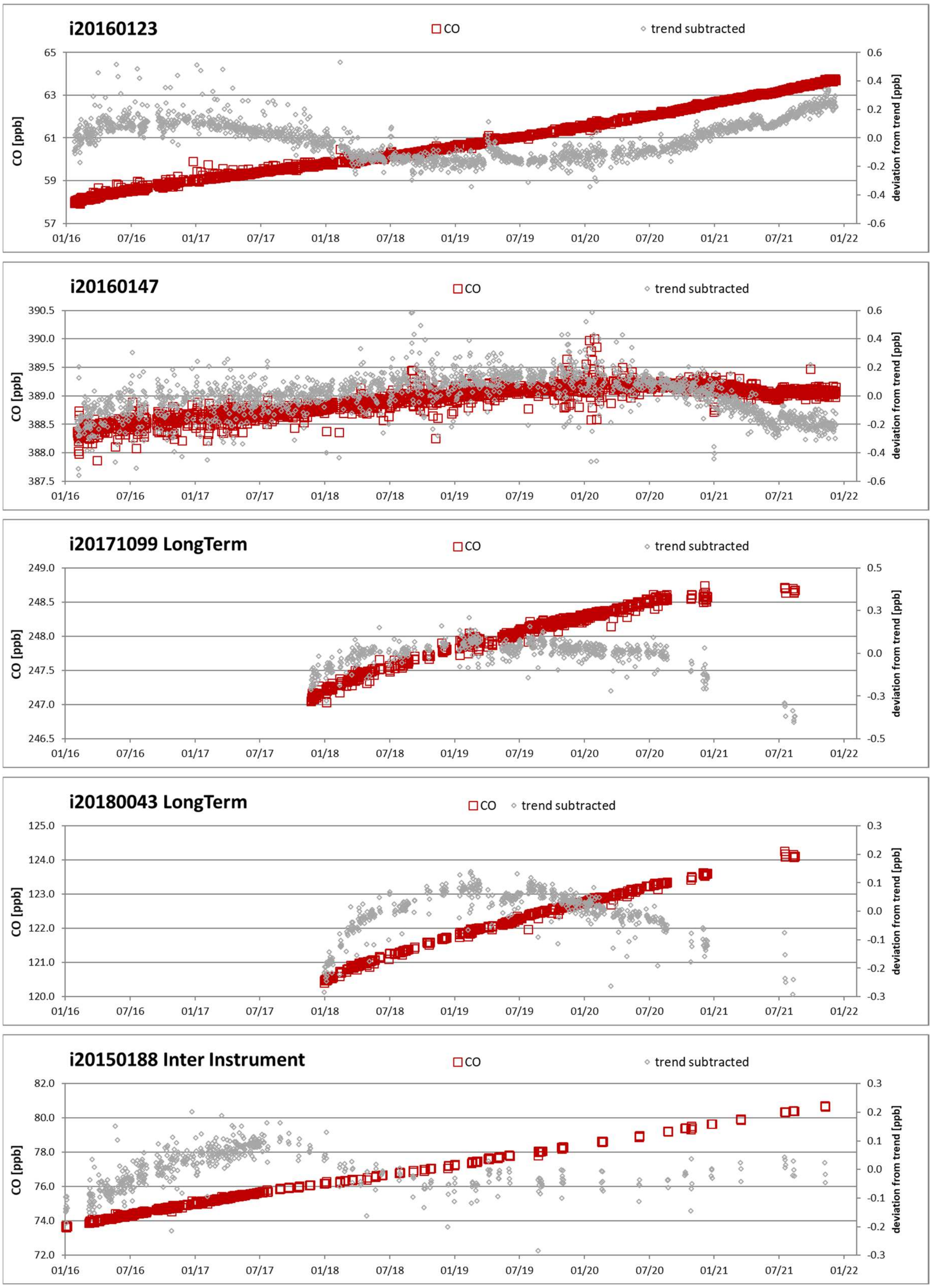
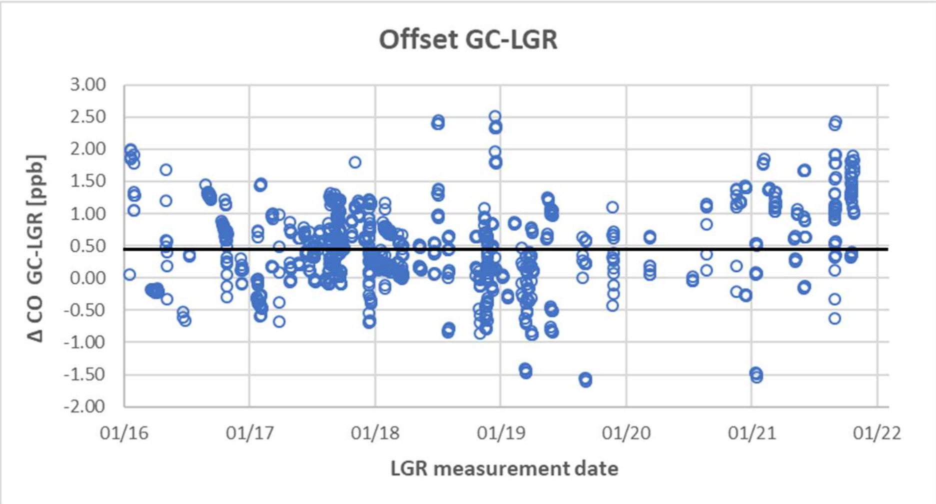


Figure 25: Time series of daily CO data of target standards analyzed on the LGR instrument based on revised assignments (red squares)

7.4 Internal CO Comparison: LGR-GC

Standard gases that are calibrated for CO by the LGR have often also been analyzed by GC. The GC measurements are linked to the same set of Primary FCL Standards but are based on a different set of seven Secondary Standards. The inter-instrumental measurement differences for all standards that have been analyzed within the same month (in order to avoid any overlaying CO growth in the lag period) are depicted in Figure 26 (only standards within the range defined by the calibration standards were considered).

GC results for the intercomparison samples are on average slightly higher (LGR-GC = -0.4 ± 0.6 ppb). There is a small concentration dependency in the offset between the instruments. It has changed only a little over time but offsets in 2021 have increased a bit. Note, that the GC-RGA precision in general is by a factor of 10 worse than the LGR, and the scatter and most likely mean biases of the data can primarily be attributed to the GC analysis.



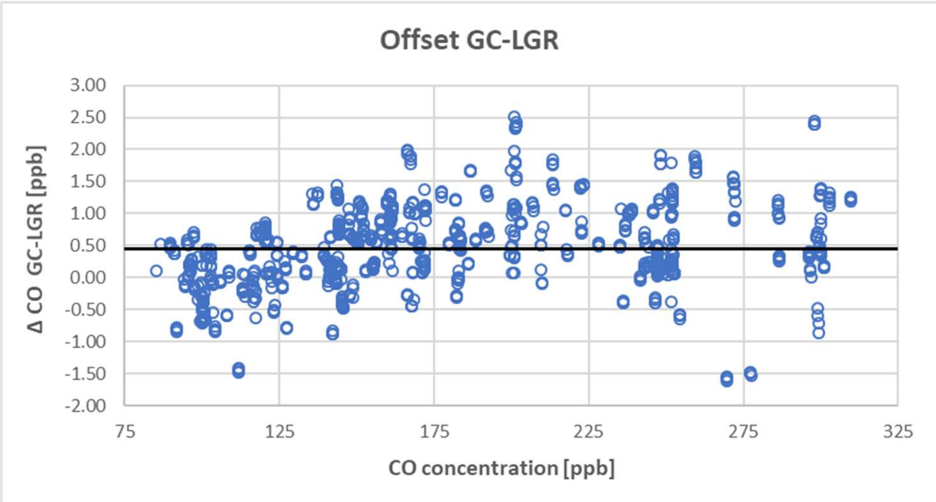


Figure 26: Time series of LGR-GC CO differences of measurement results of the same samples

7.5 External CO Comparisons

7.5.1 CO compatibility ICOS FCL - MPI-BGC

The most intensive comparison measurements have been made with the MPI-BGC GasLab. This laboratory is using different analytical technology (Aerolaser AL5002) and their measurements are tied to the WMO Mole Fraction scales by an independent set of Primary Standards. These MPI-BGC Primary Standards already have CCL calibration records with multiple measurements in different years for nine individual standards over six to seventeen years, partly with established drift rates and partly with apparently stable composition. In contrast to the other trace gases covered by this report calibrations made by the CCL before 2011 are not tied to the same WMO primary standards. The comparability of these old calibrations to calibrations since 2012 is therefore inferior. The assessment of the drift of MPI-BGC CO standards based on the old calibrations therefore may be not as accurate as the assessment of the drift of FCL Standards.

The MPI-BGC measurements are not relevant for the assignment of the FCL standards and therefore are completely independent.

7.5.1.1 Comparison of CO Primary Standards

Basis for an agreement of FCL and MPI-BGC measurements is the compatibility of the respective sets of Primary Standards. Before or after the shipment to the CCL for recalibration of sub-sets of the respective Lab Primary Standard suites these standards were mutually exchanged between MPI-BGC and FCL and analyzed. This allows a direct comparison with the CCL.

The comparison data of the measurement results relative to the CCL assignments are shown in the Figure 27 also including the set of FMI standards that had been calibrated by the CCL. Note that the two data series in the plot are on inverted y-axes. FCL CO data for MPI-BGC Primary Standards within the calibrated range of the FCL measurements are on average 0.5 ± 0.5 ppb lower than the CCL assignments, the offset of MPI-BGC results relative to CCL assignments is 0.2 ± 0.8 ppb,

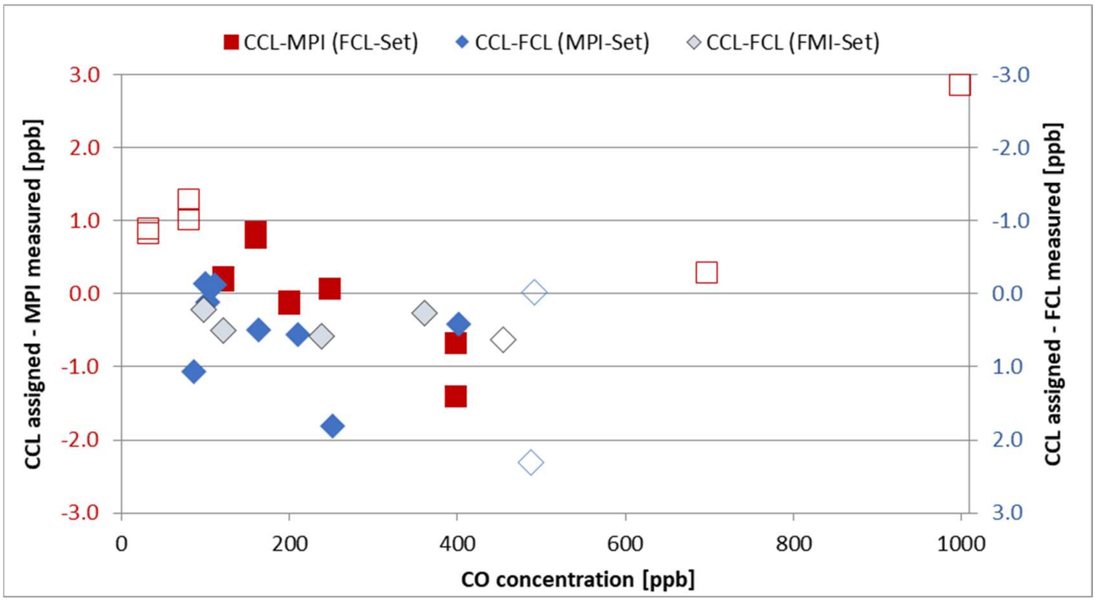


Figure 27: Difference of measured CO values to CCL assignment of the WMO standards of the partner laboratory (unfilled symbols for concentrations beyond the calibrated range)

7.5.1.2 Sample CO comparison

High pressure standards have been regularly exchanged between MPI-BGC and FCL in earlier years and analyzed in both laboratories. To make sure that the comparison is not affected by growing CO in the comparison standards only comparisons are taken into consideration where the analysis was done within six months. The difference in results for about 120 comparisons is presented in Figure 28 and Figure 29. The average offset of MPI-FCL amounts to 0.2 ppb ± 0.6 ppb (red squares) based on the revised assignments. This difference exhibits a slight concentration dependence that is in accordance with the different patterns of concentration dependent offsets to the CCL shown in Figure 27.

Plotting the inter-laboratory differences against the analysis date at the FCL reveals a trend in the offset. This trend is explainable by an overestimate of the CO increase in the FCL references [Crotwell 2019] or an underestimate of a CO growth in MPI-BGC reference standards or a combination of both.

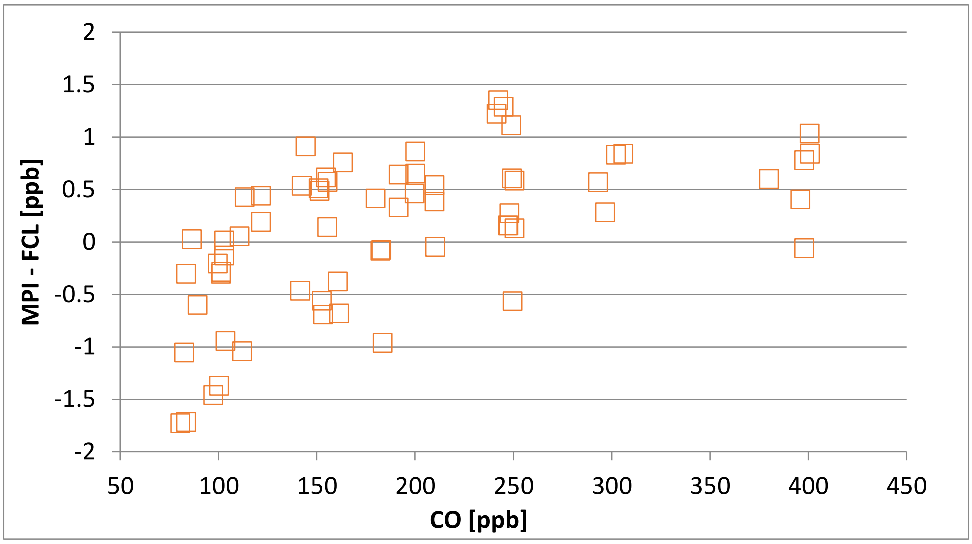


Figure 28: CO offset between FCL and MPI-BGC derived from measurements of the same high-pressure samples

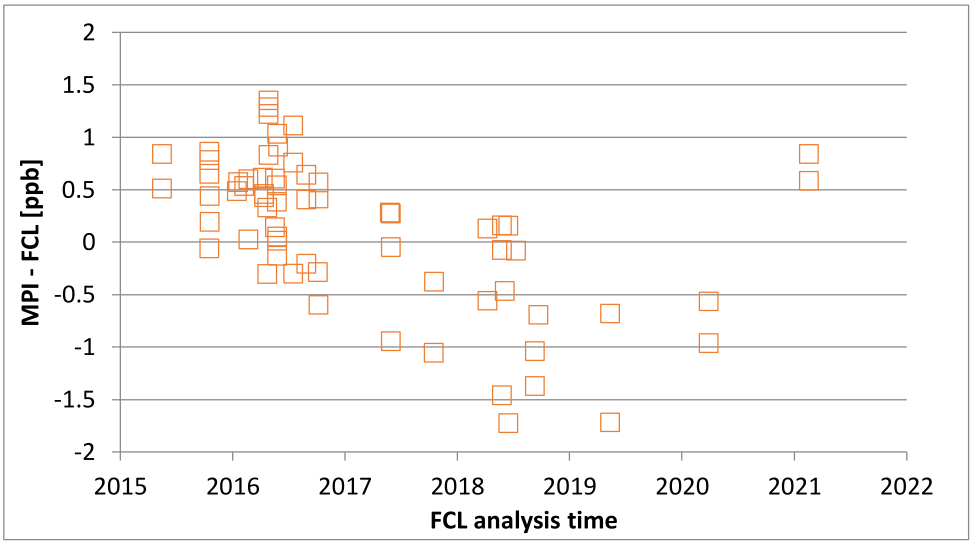


Figure 29: Time series of CO offset between FCL and MPI-BGC derived from measurements of the same high-pressure samples

7.5.2 CO compatibility ICOS FCL - NOAA

A comparison with the NOAA-GML laboratory (and other laboratories) is routinely made in two independent exercises, the Sausage Flask Intercomparison Program and the MENI high pressure cylinder program. In the Sausage intercomparison samples for comparison are prepared by connecting sets of flasks in line and filling them with dry air from a high-pressure cylinder at the FCL. The FCL is generally analyzing the composition of the filling air using the normal instrumentation for calibrating standards. Therefore, the results of the flask measurements provided by NOAA can be compared with these high-pressure cylinder measurements. The respective data are compiled in the following figures. The difference between FCL and NOAA increases with increasing CO, the mean CO offset for of all tank samples (black symbols) is FCL-NOAA = -1.1 ± 0.9 ppb.

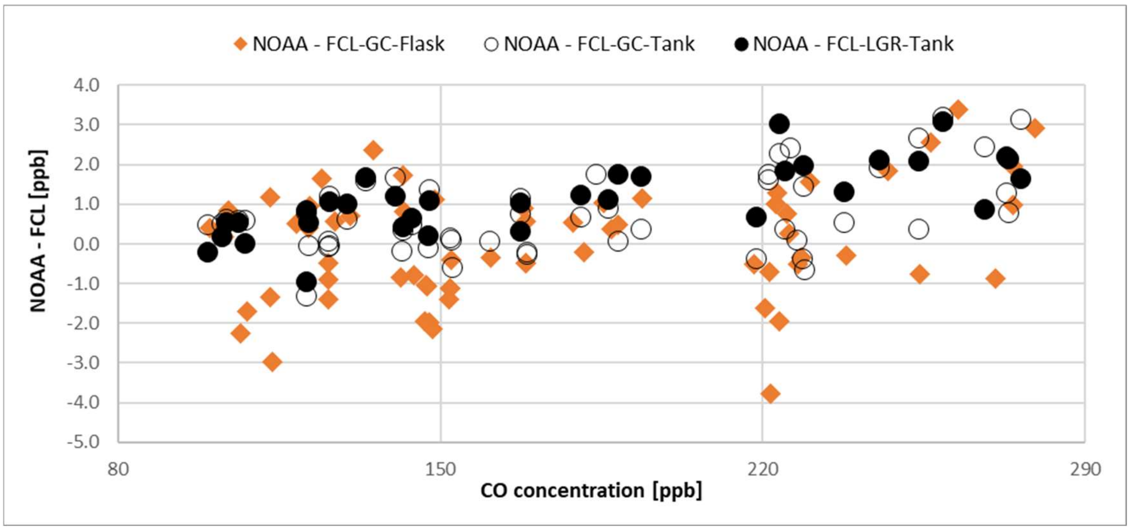


Figure 30: CO offset between NOAA Sausage flask data and FCL data. Black dots represent FCL's analysis of the Sausage fill gas (filled symbol: LGR measurement; unfilled symbol: GC); orange diamonds represent GC flask sample analysis

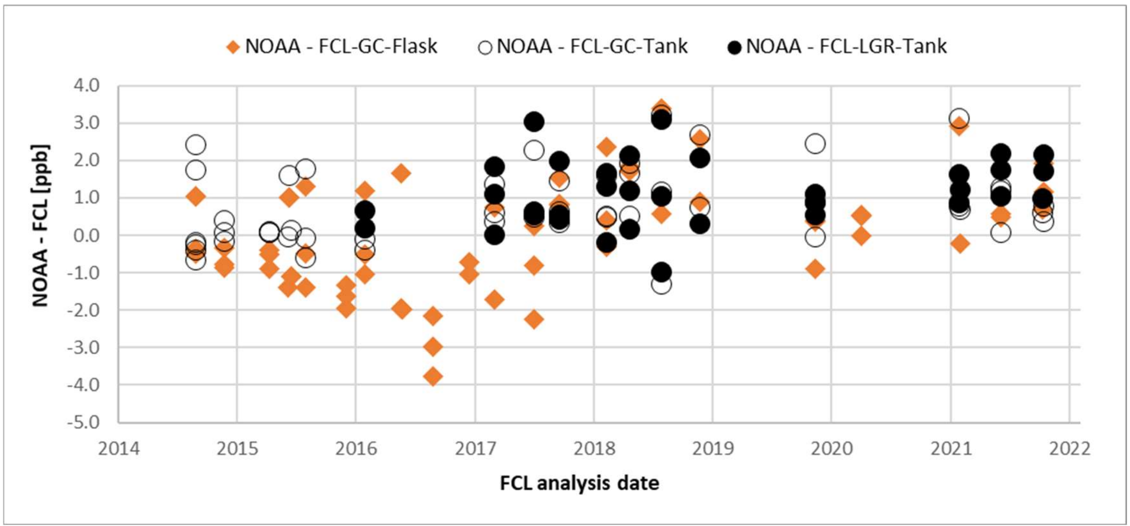


Figure 31: Record of CO offset of NOAA Sausage flask data compared with FCL data of the source gas for these samples (black symbols, see Fig. 31) and of the FCL flasks data (orange diamonds)

A round robin test between NOAA-GML (as WMO-CCL), EMPA (as WMO-WCC), MPI-BGC, FCL and FMI-ATC (ICOS Mobile Lab) has been established to check the FCL CO calibration facility's WMO Mole Fraction scale link directly. In this program a set of three cylinders is prepared and maintained by the FCL. One of these cylinders constitutes as blind sample and is modified in its composition after every completed loop. In Figure 32 results of the first iterations are shown.

It turned out that the CO mole fractions in the cylinders were growing over the time of the experiment. This needs to be taken into consideration when comparing data from measurements made at different points of time. Therefore, in Figure 32 the CO grow rate is assessed based on the CCL measurement records. The displayed offset takes the trendline of the CCL measurements as reference. On average there is no offset (FCL - CCL = - 0.1 ± 0.7 ppb) but there may be a drift of 0.5 ppb/yr in the offset of the high comparison sample.

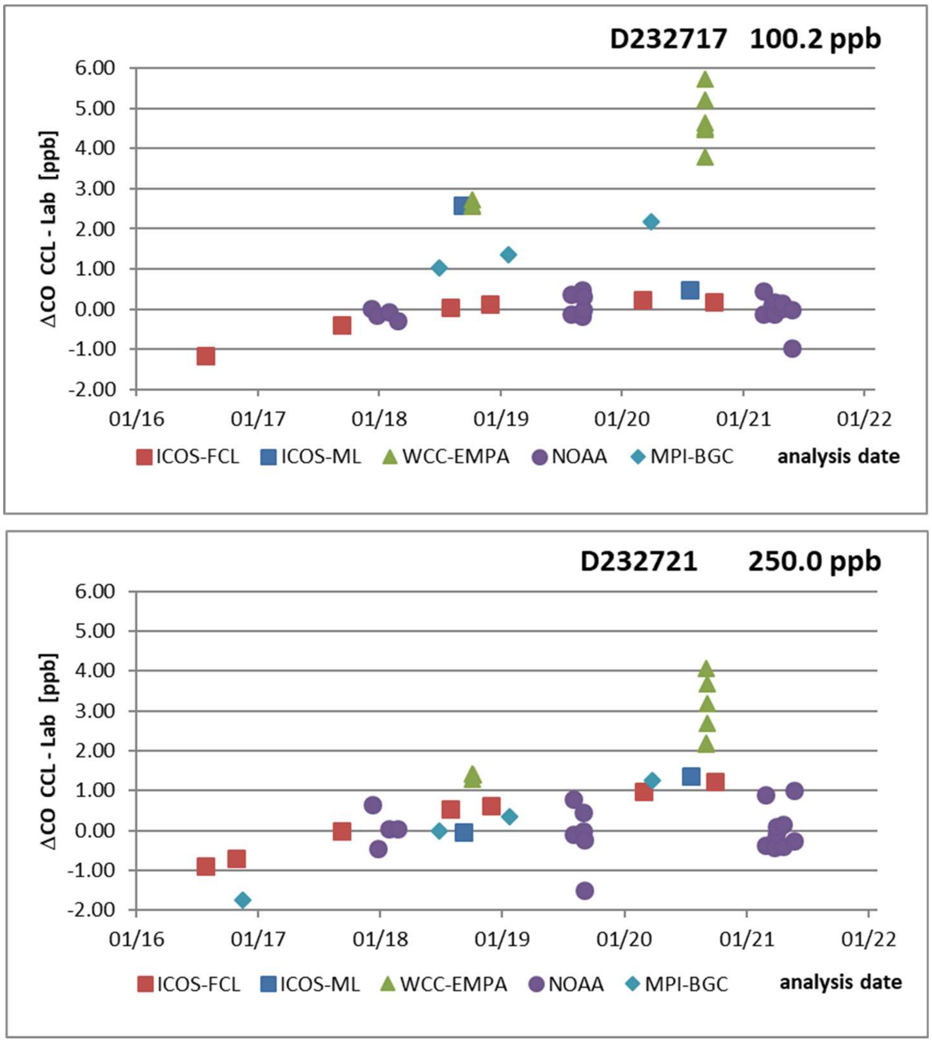


Figure 32: CO offset between FCL and NOAA-GML, based on the comparison of MENI measurements (adjusted for CO growth)

7.6 CO uncertainty evaluation

The WMO Expert Group recommendations request investigators to report uncertainty estimates for their data that include all potential sources of error [WMO 2018]. A scheme for a comprehensive uncertainty discussion has been suggested by Andrews et al. 2014. Adapting this scheme we have tried to derive such an overall measurement uncertainty based on a performance assessment of the LGR system and an evaluation of the consistency of CO assignments in the reference gases. The latter is as well the dominant source of uncertainty and at the same time the most difficult to quantify reliably.

In this assessment we have considered the following uncertainty contributions and checked them using the quality control data of this report.

7.6.1 FCL Primary CO Standards

The CCL specifies a scale transfer uncertainty of 0.4 ppb (k=1) in the range up to 400 ppb increasing (in particular for LGR assignments) to 2 ppb at 700 ppb and 4 ppb at 1000 ppb. The CCL has pointed to systematic differences they have observed between the OA-ICOS (LGR) and VURF measurement data that causes a mole fraction dependent bias in results between the analytical techniques of 0.5 – 1.5 ppb (LGR-VURF). All initial assignments had been made using the LGR instrument whereas recalibrations in 2017 were all made using the VURF instrument. While this may suggest a larger uncertainty than specified above, the quadratic regression fit residuals of the calibrations using the FCL Primary Standards are consistent with the above quote.

The growth of CO in most FCL Primary Standards is clearly documented by results from the recalibration of these standard gases by the CCL. Its results suggest a concentration dependent CO increase: standards with low mole fractions exhibit a large drift and standards with high CO mole fractions a minor to no drift. The trend function for the CO assigned values had been defined in 2019 by the first two CCL calibration events only and is currently being extrapolated beyond the time of the recalibration. This also contributes to the uncertainty. The average difference in the CO growth rates of the FCL Primary Standards when considering the 2021 recalibration compared to when ignoring it is -0.06 ppb/year for all standards < 200 ppb and +0.3 ppb/yr for the two standards at 250 ppb and 400 ppb. Based on this an additional extrapolation uncertainty of 0.5 ppb at standards < 200 ppb is assumed and an uncertainty of 1.5 ppb at larger CO values. A slowly degrading consistency of the Primary Standard set is also indicated by the steadily growing regression fit uncertainty (rising from 0.3 ppb to 0.6 ppb from 2016 to 2021).

7.6.2 CO scale transfer uncertainty

Knowledge of the CO mole fractions in the individual FCL Secondary Standards is based on the record of repeated calibrations using the FCL Primary Standards. The additional uncertainty arising from the FCL internal scale transfer measurements is expressed by the scatter of the individual calibration episode results relative to the trend line of increasing CO. The mean absolute residuals of the up to 15 assignment periods are mole fraction dependent between 0.2 and 0.3 ppb for the low and high FCL Secondary Standard, respectively.

7.6.3 CO long-term reproducibility

Any long-term reproducibility limitations beyond the random errors in daily measurements could be derived from discontinuities in the target standard measurement record (Figure 27). While also the target standards exhibit growing CO the daily means scatter on average by less than 0.2 ppb around the respective trend line.

7.6.4 CO measurement uncertainty estimate

Based on the above considerations the following combined standard uncertainty (k=1) is calculated as the square root of the sum of the individual uncertainty squares:

1. Scale link uncertainty = 0.7 ppb (standards with CO < 200 ppb)

= 1.6 ppb (standards with CO > 200 ppb)

- The scale link uncertainty estimate is derived from the specified CCL assignment uncertainty (0.4 ppb below 400 ppb) and the CO growth function uncertainty (0.5 ppb below 200 ppb).
- Uncertainty of the FCL internal scale transfer to FCL Secondary Standard assignments (0.2 ppb).

2. Measurement uncertainty of daily means = 0.1 ppb

- mean uncertainty of the daily calibration regression (0.07 ppb)
- uncertainty of the detector response drift throughout the validity of a daily calibration (0.01 ppb)
- uncertainty from insufficient sample flushing and instrumental repeatability of the daily sample measurements (0.05 ppb, for 10 min means)

3. Additional long-term variability = 0.2 ppb

The accuracy with respect to the WMO Mole Fraction scale is limited by the uncertain knowledge of the current assigned values in the drifting reference standards. It calculates as the root of the sum of squares of the scale link uncertainty and the measurement uncertainty and amounts to 0.7 ppb for samples with CO < 200 ppb and 1.6 ppb for samples with larger CO. The internal reproducibility is estimated to be 0.5 ppb. This is consistent with observed external comparison results.

The CCL by definition provides the link to the WMO Mole Fraction scale but it has announced that the way the growth of CO in the WMO Scale Primary References was prescribed likely overcompensated this drift for low concentrated standards. The evaluation of the scale is ongoing at the CCL. All uncertainty estimates made here refer to the uncertainty of the measurements and assignments relative to the current scale and do not include a term for any potential concentration dependent scale error.

8 N_2O

8.1 FCL Primary N_2O Standards

8.1.1 CCL N_2O assignments

After initial calibration of all FCL Primary Standard gases in 2014 the first recalibrations of each three of the standards have been made in 2016 and 2017, and 2018. In December 2020 the complete set of FCL Primary Standards was sent to the CCL for the third calibration, performed in May 2021.The reassignments by the CCL have generally been within the uncertainty of the initial assignment and have not caused a systematic shift of the entire set. (Δ = -0.01 ±0.06 ppb) see Table 13. However, there is a slight concentration dependent difference for the FCL Primary Standards < 320 ppb having probably been determined too low initially.

Table 13: N_2O assignments for FCL Primary Standards by the CCL (WMO X2006A scale)

Sample ID	Cylinder ID	CCL date 1	CCL date 2	CCL date 3	WMO X2006A	WMO X2006A	WMO X2006A	Δ initial – mean [ppb]
i20140055	CB09944	Jan-14	Jul-17	May-21	316.77	316.90	317.01	**-0.12**
i20140056	CB09939	Jan-14	Feb-19	May-21	319.86	319.97	319.93	**-0.06**
i20140057	CB09958	Jan-14	Oct-16	May-21	327.12	327.02	327.21	**0.00**
i20140058	CB09983	Jan-14	Jan-19	May-21	329.92	329.89	330.06	**-0.04**
i20140059	CB09952	Apr-14	Nov-16	May-21	334.60	334.52	334.57	**0.04**
i20140060	CB09955	Jan-14	Jul-17	May-21	339.48	339.52	339.43	**0.00**
i20140061	CB09957	Jan-14	Nov-16	May-21	343.95	343.88	343.82	**0.07**
i20140062	CB09934	Mar-14	Jun-17	May-21	349.13	349.18	349.14	**-0.02**
i20140054	CB09948	Jan-14	Jan-19	May-21	362.13	362.12	361.90	**0.08**

8.1.2 Regression fit residuals of FCL Primary N_2O Standards

The time series of the quadratic regression fit residuals is presented in Figure 33 for calibration events where the complete FCL Primary Standard suite was used.

N_2O mole fractions are known to be generally stable in aluminium high pressure cylinders. The assumption of a stable standard set is supported by the fact that the regression fit residuals do not show significant trends for any of the individual standards (within 0.03 ppb).

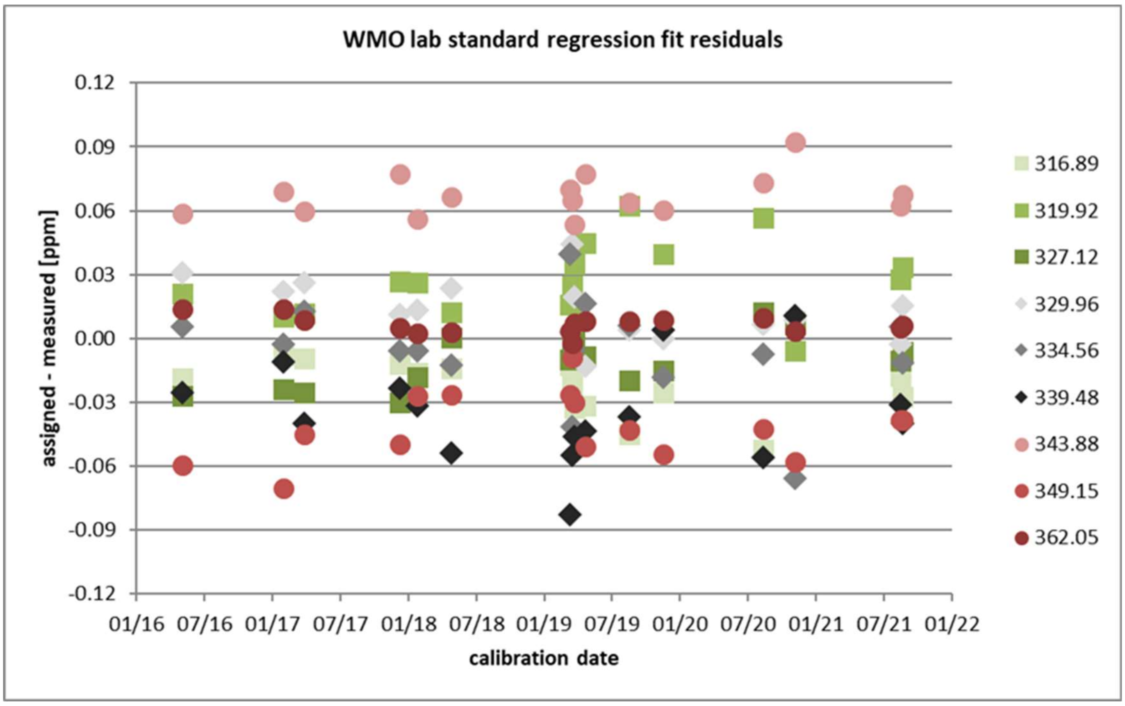


Figure 33: Time series of N_2O regression fit residuals for LGR calibration events using FCL Primary Standards

8.2 FCL Secondary N_2O Standards

8.2.1 Assignment record

The four reference gases that have been used as FCL Secondary Standards for the daily LGR calibration have been analyzed within 24 to 25 calibration episodes together with the complete set of FCL Primary Standards between May 2016 and Sep 2021. In May 2018 one of the Secondary Standard gases was exhausted (i20150251; five calibration episodes together with FCL Primary Standards) and was succeeded by another standard with a similar N_2O content (i20170889; 18 calibration episodes together with FCL Primary Standards). The three remaining tanks were replaced when they were exhausted in July 2021. For the three replacements, the assigned values were determined by measurements against the old set of FCL Secondary Standards (the number of comparison measurements is: i20201308 (n=114), i20201254 (n=114), i20201255 (n=117)). As explained in the section above the initial CCL assignments for standards with N_2O < 320 ppb were lower than for both recalibration results. Therefore, for the assignment of all new Secondary Standards (including the standard i20170889) the assigned values of the initial Secondary Standard set have been adjusted and are based on the mean of all three CCL calibrations (not only the initial CCL assignment). This results in a new assigned value for i20170889 that is about 0.10 ppb higher than the one used between 3rd May 2018 and 26th July 2021. As indicated in the previous section the update is very minor for ambient atmospheric N_2O concentrations (< 0.03 ppb).

Table 14: N_2O assignments for FCL Secondary Standards (WMO X2006A scale)

Sample ID	Cylinder ID	Assigned Value	Avg. Ω*	*Date of exchange*	Sample ID	Cylinder ID	Assigned Value
i20150251	CA05640	**316.923 ppb**	316.922 ppb	*2018-05-03*	i20170889	D557226	**315.682 ppb**
i20150189	D073397	**324.506 ppb**	324.514 ppb	*2021-07-26*	i20201308	D753834	**324.395 ppb**
i20150544	D073396	**334.201 ppb**	334.194 ppb	*2021-07-26*	i20201254	D753835	**339.360 ppb**
i20150191	D073395	**344.970 ppb**	344.972 ppb	*2021-07-26*	i20201255	D753836	**348.730 ppb**

* Average of all calibration episodes between Nov 2015 and Jul 2021; for i20170889 the assigned value was 315.580 ppb from 3rd May 2018 to 26th July 2021

8.2.2 Residual record

The residuals of the quadratic regression fit of the FCL Secondary Standard daily calibration are given in Figure 34. The absolute values are all extremely small, the average scatter of the individual standard's residual time series is 0.008 ppb, containing random noise but also systematic variations that last for several weeks to months. No steady trend is apparent in the residuals. This is good supporting evidence for the assumption that all FCL Secondary Standards are stable in their N_2O mole fractions.

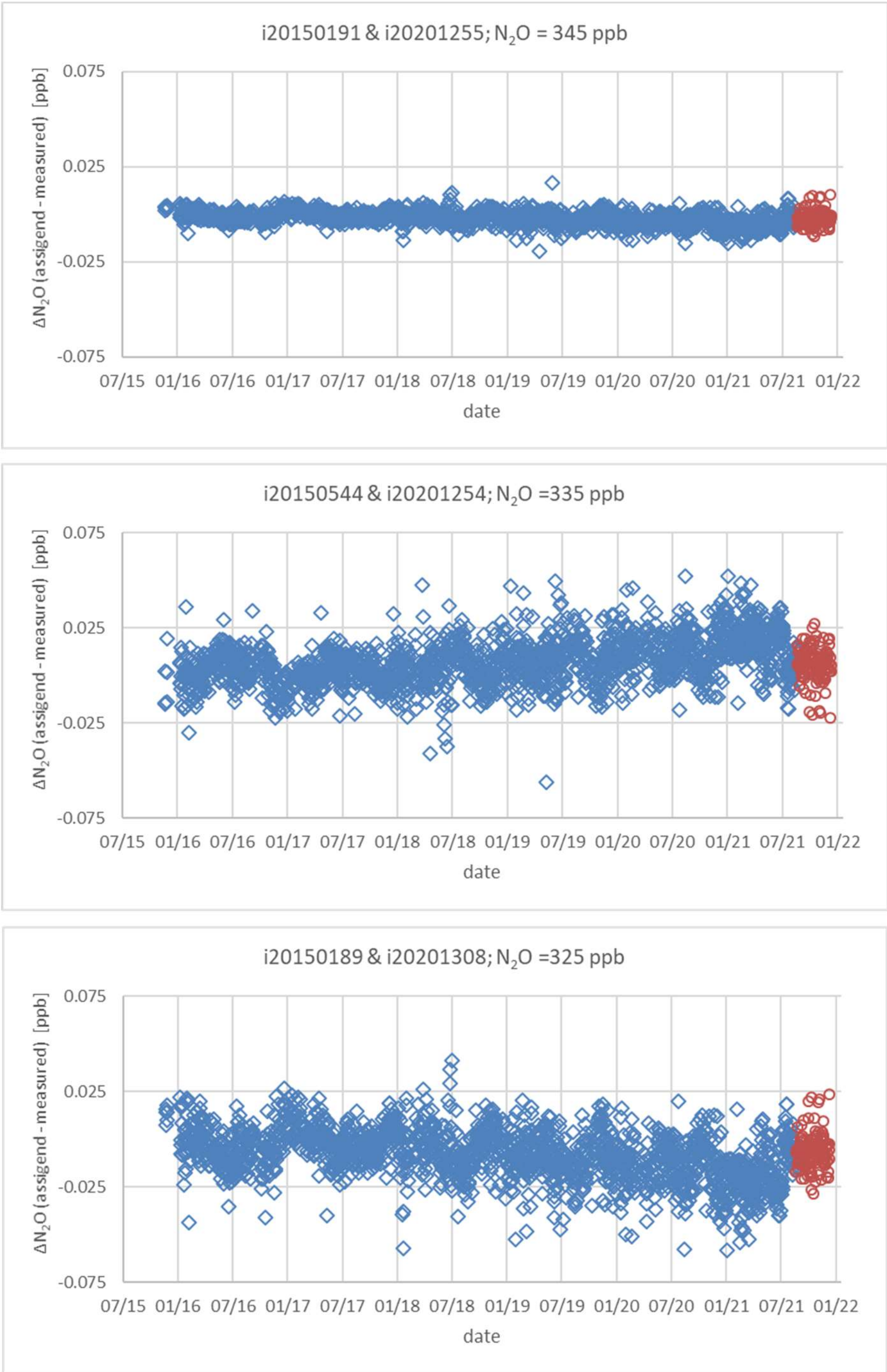


Figure 34: Quadratic regression fit residuals of the daily LGR N_2O calibration with FCL Secondary Standards

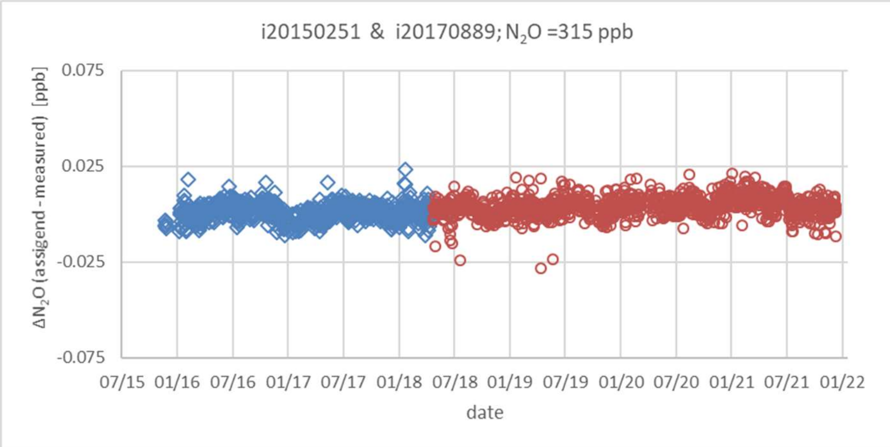


Figure 34: Quadratic regression fit residuals of the daily LGR N_2O calibration with FCL Secondary Standards

8.3 N_2O Targets

In the period from March 2016 to December 2020 two targets have been constantly in use for the LGR system. They are complemented by additional Long Term Targets (also Archive Quality Control). This shall maintain a long-term link of succeeding (short term) targets. After initial daily analysis they have been assessed on a regular, less frequent basis since 2020 to extend their lifetime.

In Table 15 the daily mean target results based on the LGR daily calibration are compared to the assignment by calibration directly with FCL Primary Standards. The adjustment of the assigned value for i20170889 in the course of the replacement of the three emptied tanks from the first set of FCL Secondary Standards, the daily mean target results were split into two tables, for the period before the change (26th July 2021) and after.

One standard is slightly above currently ambient atmospheric mole fractions (336 ppb) while the other two targets contain 308 ppb and 313 ppb N_2O, respectively. The reproducibility over time gets worse when the mole fractions are below the calibrated range and also the offset between the Primary Standard calibration results and the regular (Secondary Standard based) analysis results increases the further the N_2O mole fraction is away from the calibrated range. The low target therefore rather serves as an early indicator for problems with the system and is not considered as a measure for the performance of the measurements.

Table 15: Target reference standards for N_2O

sample ID	tank ID	measured N_2O mean [ppb]	std.dev. of mean [ppb]	n [days]	Primary Calibration N_2O mean [ppb]	std.dev. Calibration mean [ppb]	n calibration events
i20160123	D073388	308.732	0.044	1731	308.857	0.052	16
I20150188	D073398	313.460	0.023	558	313.532	0.037	11
I20180043	D557243	324.163	0.018	580	324.165	0.013	10
I20171099	D557242	333.719	0.022	596	333.718	0.027	11
i20160147	D801333	335.769	0.017	1689	335.784	0.034	15

Table 15: Target reference standards for N_2O

sample ID	tank ID	measured N_2O mean [ppb]	std.dev. of mean [ppb]	n [days]	Primary Calibration N_2O mean [ppb]	std.dev. Calibration mean [ppb]	n calibration events
i20160123	D073388	308.884	0.046	173	308.857	0.052	16
I20150188	D073398	313.570	0.013	10	313.532	0.037	11
I20180043	D557243	324.209	0.014	8	324.165	0.013	10
I20171099	D557242	333.726	0.012	8	333.718	0.027	11
i20160147	D801333	335.773	0.027	173	335.784	0.034	15

The upper table contains the values up to the adjustment of the scale link on 26th July 2021, the lower table the subsequent period.

The N_2O time series of the target standard measurements is depicted in Figure 35 (the adjustment of the assigned value for i20170889 is accounted for in the graphs). For mole fractions within the calibrated range the agreement between assigned and mean measured value is very good (0.02 ppb). The record of the low standard reveals different periods where the mean results are stable for weeks to months on different levels that are different by up to 0.1 ppb. This provides some estimate for the uncertainty of measurements beyond the calibrated range.

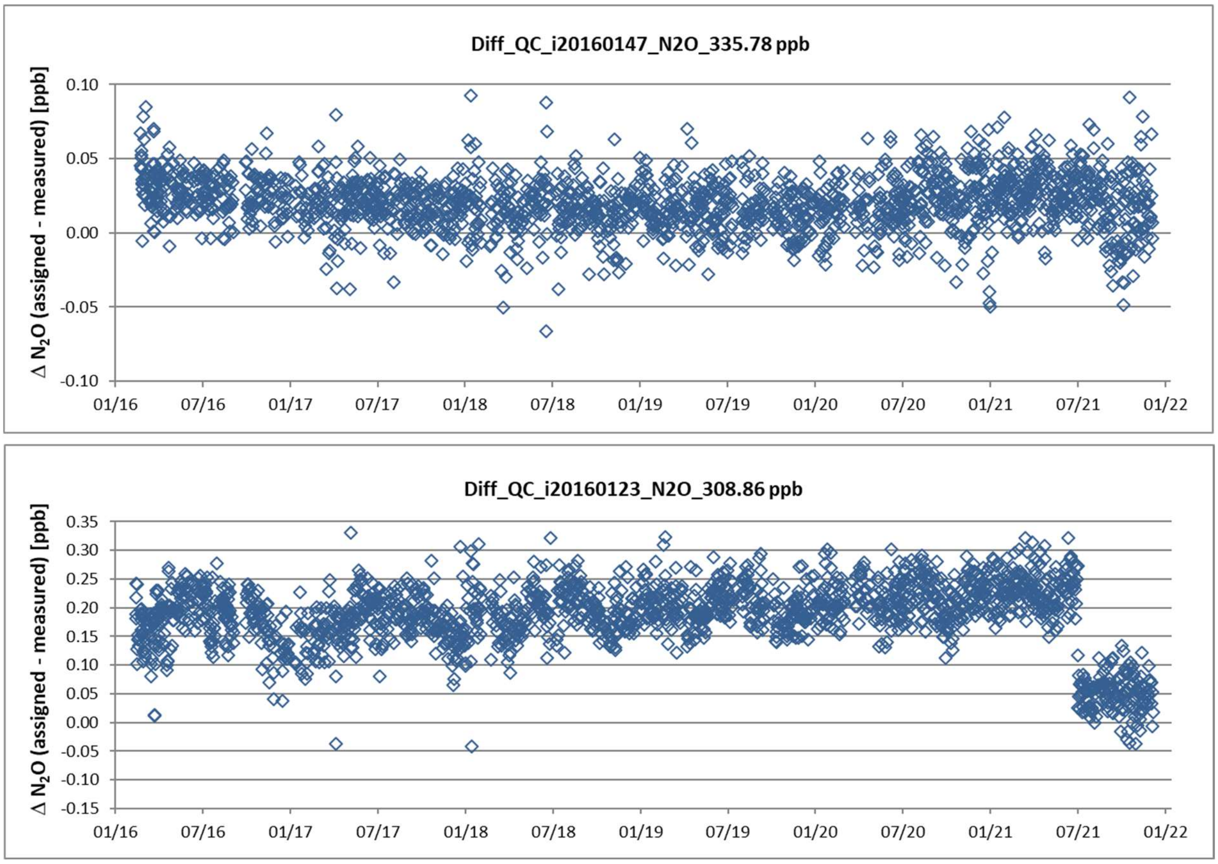


Figure 35: Time series of the offset of N_2O target measurements to their respective assigned values.

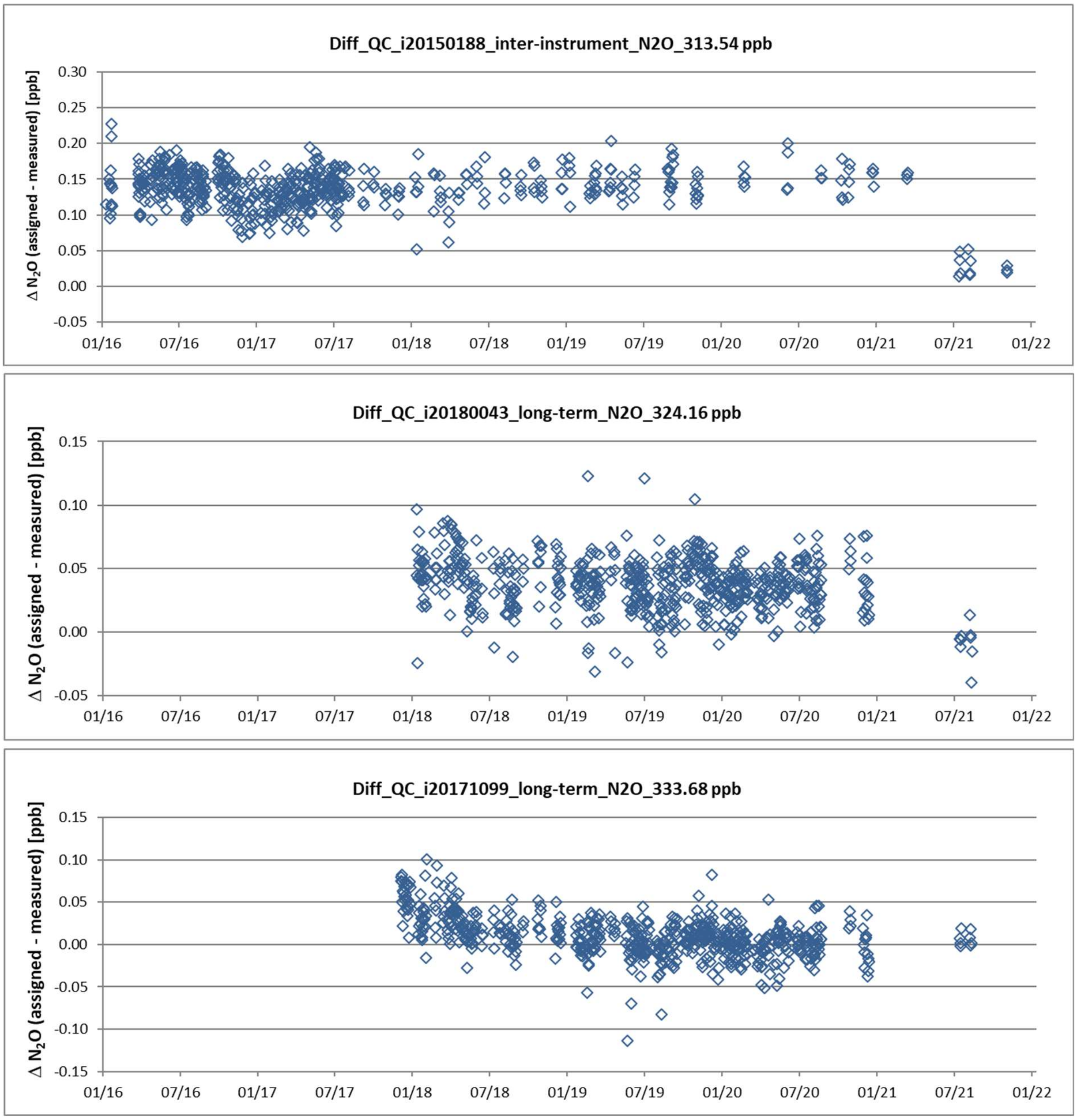
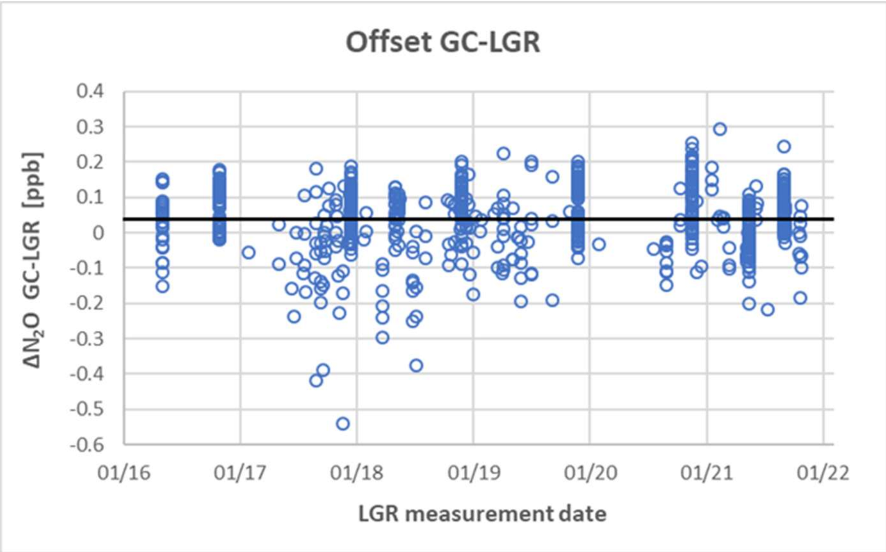


Figure 35: Time series of the offset of N_2O target measurements to their respective assigned values.

8.4 Internal N_2O Comparison: LGR-GC

Standard gases that are calibrated for N_2O by the LGR instrument have often also been analyzed by GC. The GC measurements are linked to the same set of FCL Primary WMO Standards but are based on a different set of six Secondary Standards. The GC detection of N_2O by an Electron Capture Detector (ECD) can be influenced by SF_6 mole fractions if they differ significantly from the atmospheric air abundance. Therefore, only samples have been included in the comparison that contain 8-30 ppt SF_6 at ambient N_2O values of 319-350 ppb. As the reproducibility and repeatability of the GC-ECD (0.1 ppb and 0.14 ppb, respectively) are in general by a factor of 7 inferior to that of the LGR, only GC measurements have been considered that have been analyzed on the GC on more than one day with at least 10 injections. The averaged inter-instrumental measurement difference for all comparison samples is -0.02 ppb ± 0.12 ppb (see Figure 36). This includes the marginal bias of the initial LGR Secondary Standard assignments for measurements before July 2021 and does not provide indication for any other bias in either of the instruments.



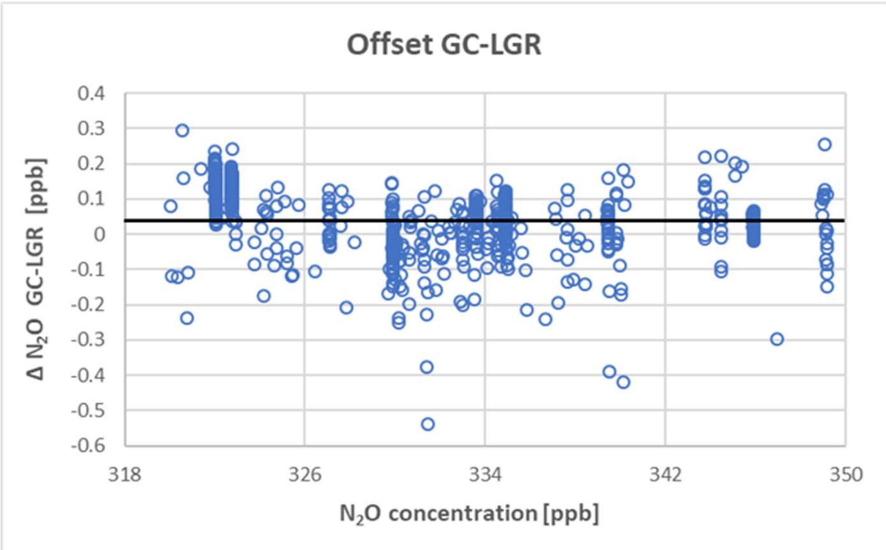


Figure 36: Offsets of daily LGR N_2O measurements relative to the annual mean of GC results

8.5 External N_2O Comparisons

8.5.1 N_2O compatibility ICOS FCL - MPI BGC

The most intensive comparison measurements have been made with the MPI-BGC GasLab. This laboratory is using different instrumentation (Agilent 6890 GC-ECD) and their measurements are tied to the WMO Mole Fraction scales by an independent set of Primary Standards. These MPI-BGC Primary Standards already have CCL calibration records with multiple measurements in different years for fifteen individual standards assigned over 14 years.

The MPI-BGC measurements are not relevant for the assignment of the FCL standards and therefore are completely independent.

8.5.1.1 Comparison of N_2O calibration standards

Basis for an agreement of FCL and MPI-BGC measurements is the compatibility of the respective sets of calibration standards. As the FCL Primary Standards have been produced at the MPI-BGC they also were thoroughly analyzed at the MPI-BGC in 2013 and 2014 before being used by the FCL. In addition, these standards were also analyzed for a third time before or after the shipment to the CCL for the first recalibration of subsets of this FCL Primary Standard suite. Likewise, MPI-BGC Primary Standards that were simultaneously returned to the CCL for recalibration were also analyzed by the FCL. Measurements at MPI-BGC have started 15 years earlier and thus the mole fraction range of the Primary Standards is about 15 ppb lower compared to the FCL Primary Standards. Therefore, the high FCL standard and low MPI-BGC standard are far beyond the calibrated ranges of the other lab and the bias for these standards is largely due to an extrapolation error. For the remaining standards a small, consistent offset between CCL-assignments and MPI-BGC measurements of the FCL Primary Standard set of -0.06 ± 0.04 ppb is observed whereas an offset between CCL-assignments and FCL measurements of the MPI-BGC Primary Standard set of 0.17 ± 0.07 ppb is apparent which also shows up with 0.15 ± 0.05 ppb in the CCL-FCL difference for the FMI set in that range. However, these offsets are reduced to 0.12 ppb when changing the basis of the N_2O assignments of the FCL Secondary Standards from the initial CCL assignment values of the FCL Primary Standards to the average value of all CCL calibration results up to 2021 (see section 8.2.1). Including the measurements of the UBA standard set in October 2021 an overall offset of 0.08 ppb ± 0.08 ppb is observed for all standards in the range relevant for atmospheric measurements (325 - 350 ppb). This is consistent with the standard assignment uncertainty of 0.11 ppb specified by the CCL and very similar to the corresponding CCL-MPI offset for the FCL Primary Standards as shown in Figure 37.

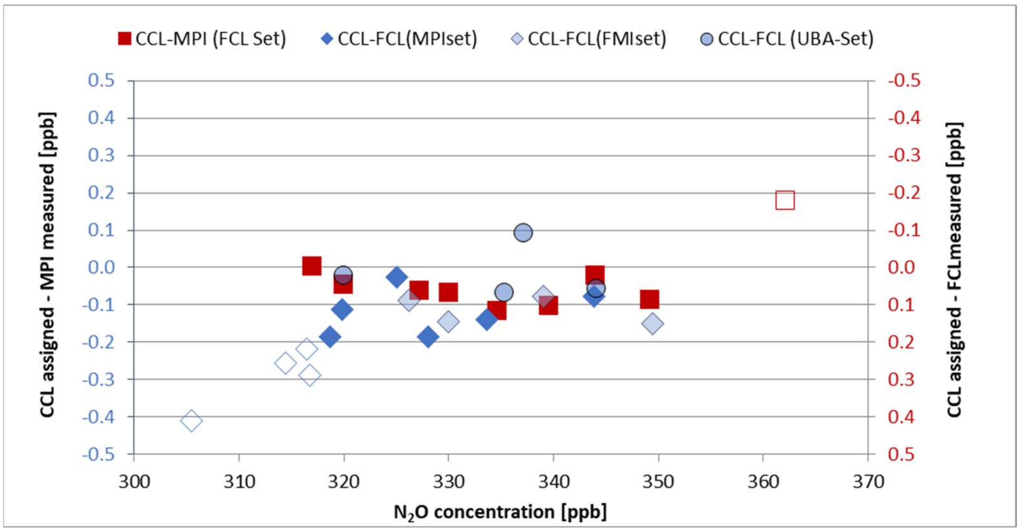


Figure 37: Differences of Primary Standard measured N_2O results to CCL assigned values MPI-BGC measurements of FCL Primary set (red squares- right y axis) and FCL measurements of MPI-BGC (blue diamonds, open symbols represent values extrapolated beyond the calibrated range set by Secondary Standards), ATC-MobileLab Primary set (grey diamonds) and UBA Schneefernerhaus (bluish dots) (*Note that the two axes have opposite signs*)

8.5.1.2 Sample N_2O comparison FCL

High pressure standards have been regularly exchanged between MPI-BGC and FCL and analyzed in both laboratories. The resulting differences for about 180 comparisons (for FCL LGR values only) are presented in Fig 38. The average offset of MPI-FCL within the Secondary Standards' range amounts to 0.09 ppb ± 0.17 ppb. This corresponds to the offset established in the preceding section and confirms the mole fraction dependence.

8.5.2 N_2O compatibility ICOS FCL - NOAA

Comparison with the NOAA-GML laboratory (and other laboratories) is routinely made in two independent exercises, using the Sausage Flask Intercomparison Program and MENI high pressure cylinder program. In the Sausage Program, samples for comparison are prepared by connecting sets of flasks in line and filling them with dry air from a high-pressure cylinder at the FCL. The FCL is generally analyzing the composition of the filling air using the normal instrumentation for calibrating standards. Therefore, the results of the flask measurements provided by NOAA can be compared with these high-pressure cylinder measurements (see Figure 39). The agreement of all valid samples (defined by a flask pair agreement within 0.3 ppb) yields a difference of NOAA - FCL = 0.07 ppb ± 0.09 ppb. In summer 2019 the NOAA laboratory has changed instrumentation which did not change the offset but the standard deviation of the difference has been reduced to 0.08 ppb.

The MENI round robin test between NOAA-GML (as WMO-CCL), EMPA (as WMO-WCC), MPI-BGC, FCL and -ATC (ICOS Mobile Lab) has been established to check the ICOS WMO mole fraction scale link in a regular manner. In this program a set of three cylinders is prepared and maintained by the FCL. One of these cylinders constitutes a blind sample and is modified in its composition after every completed loop. Results of the first iterations are shown in Figure 40. The observed offset between CCL and FCL is 0.04 ± 0.05 ppb. This small offset is in line what has been revealed by the comparison of other standards assigned by the CCL and the Sausage Program.

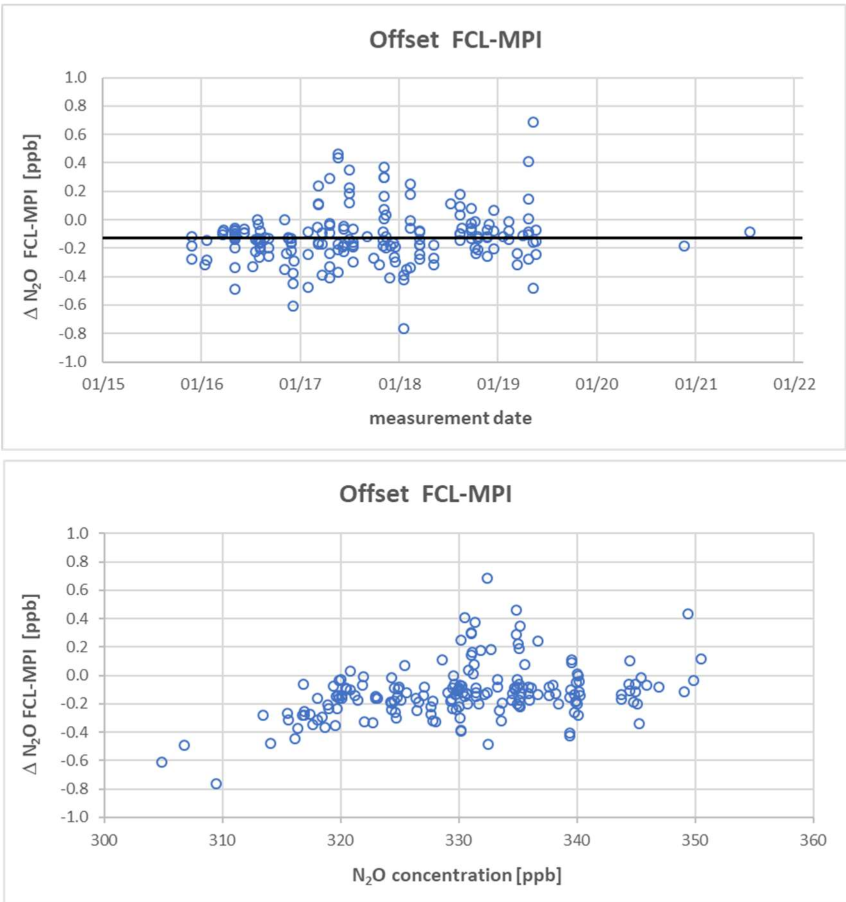


Figure 38: Difference of N_2O measurement results for samples that have been analyzed at FCL and MPI-BGC
All MPI-BGC GC measurements since 2015 with minimum 6 injections within the calibrated range are considered in aggregated means (blue symbols)

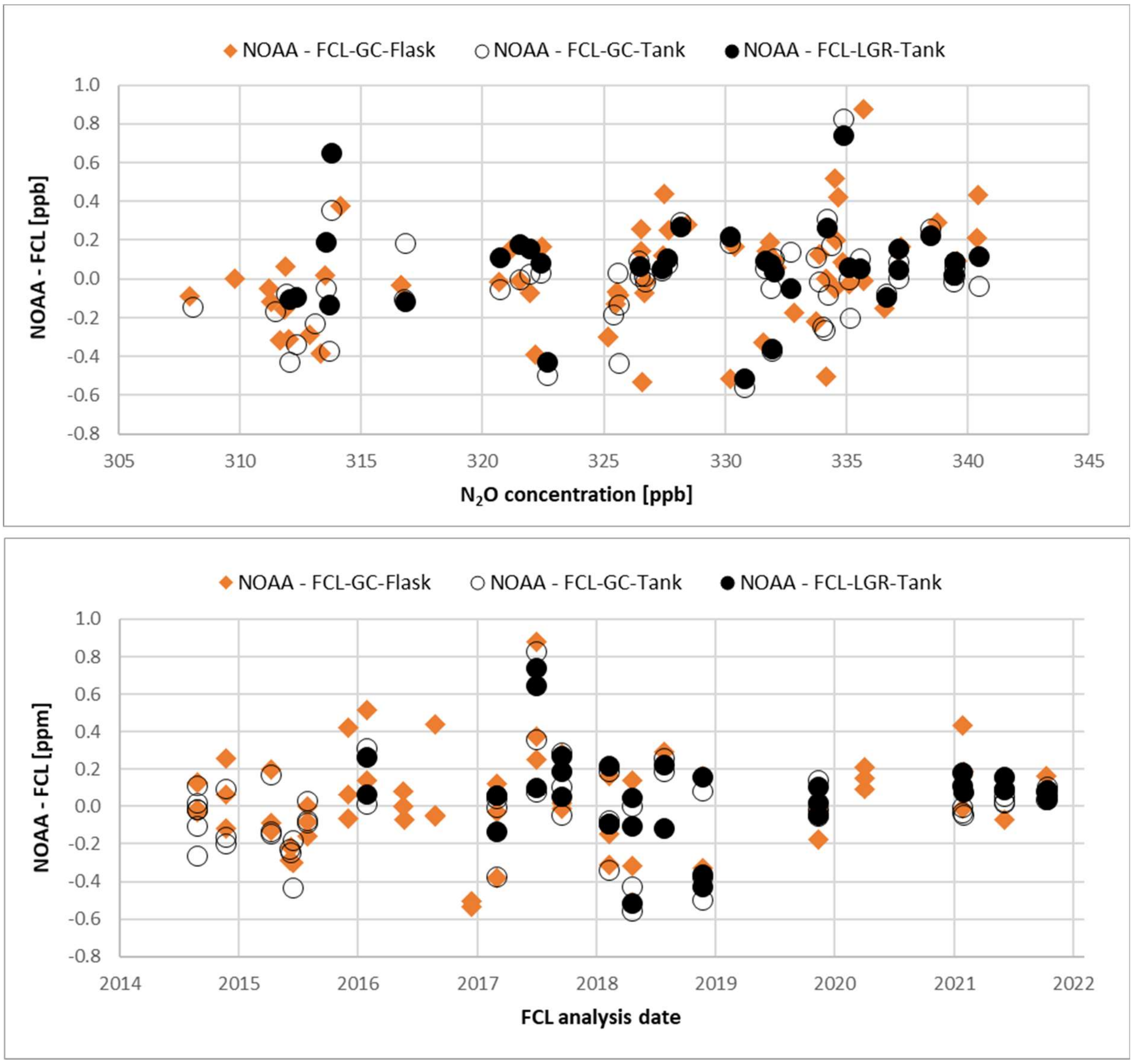


Figure 39: N_2O offset between NOAA-GML and FCL based on flask samples (by NOAA-GML)

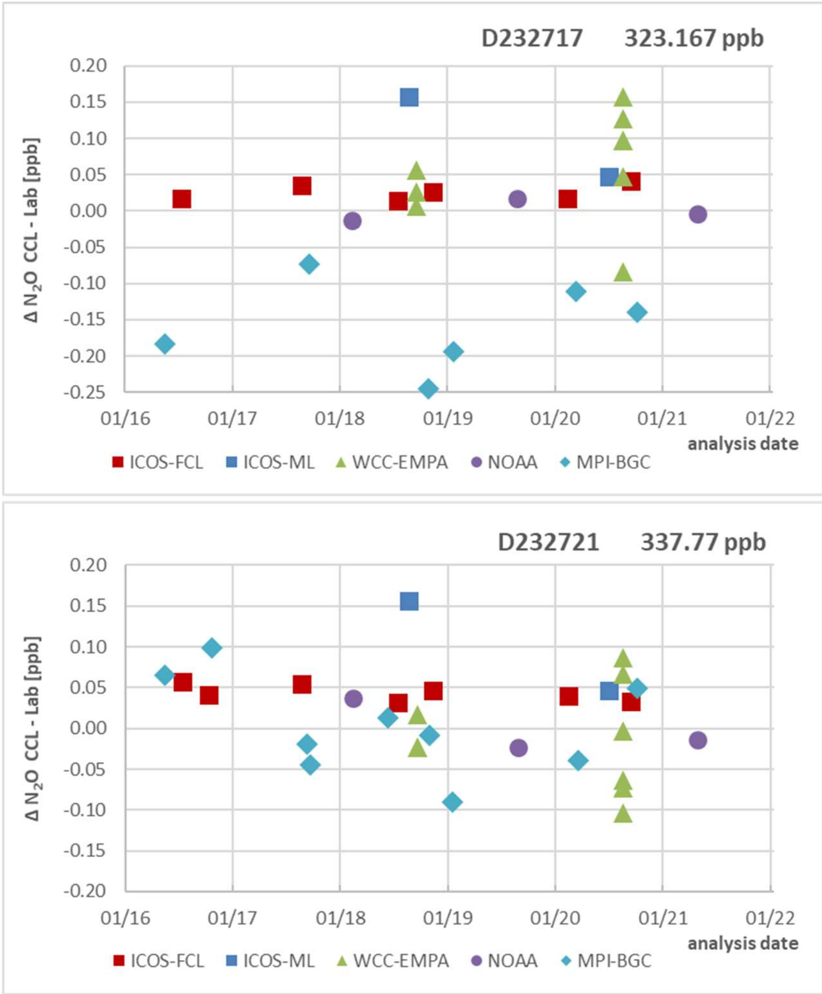


Figure 40: N_2O offset between FCL and NOAA-GML based on analyses of high-pressure cylinder samples

8.6 N_2O uncertainty evaluation

According to the WMO Expert Group recommendations investigators must report uncertainty estimates for their data that include all potential sources of error. A scheme for a comprehensive uncertainty discussion has been suggested by Andrews et al. 2014. Following this scheme we have derived an overall N_2O measurement uncertainty based on a performance assessment of the LGR system. In this assessment we have considered the following uncertainty contributions:

8.6.1 FCL Primary N_2O Standards

The CCL specifies reproducibility for N_2O calibrations of 0.11 ppb (68% confidence level). This CCL uncertainty quote is in line with the assessment of the FCL Primary Standard set. The initial absolute residuals of the FCL Primary Standard set are on average 0.05 ppb and the differences from the succeeding CCL recalibrations relative to the first were also within this range with the exception of a systematic increase of the assignments of the N2O assignments in the two standard gases with the lowest N_2O. The compatibility of the FCL Primary Standard set with other CCL calibrated standards (held by MPI-BGC, FMI, UBA) yield a systematic offset of 0.08 ppb on average lower than the CCL assignments of the respective standards for gases with $N_2O \geq 320$ ppb. The reverse assessment of FCL Primary Standards by the MPI-BGC laboratory results in a very similar mean offset of 0.06 ppb. This offset includes the assignment uncertainties of each calibration gas set.

8.6.2 N_2O scale transfer uncertainty

The assigned values of the first set of LGR Secondary Standards was defined by the first calibration using the FCL Primary Standards. Repeated calibrations of the FCL Secondary Standards using these FCL Primary Standards indicated that these initial assignments were slightly too low in the atmospheric concentration range (≈ 0.03 ppb in the range 330 - 340 ppb) with a larger offset for the low Secondary Standard (0.11 ppb). As the LGR instrument is characterized by a quadratic curve this introduces a systematic extrapolation error for samples with N_2O outside the calibrated range. The reproducibility of these assignments is ranging between 0.018 ppb and 0.036 ppb. The absolute values of the regression fit residuals of the daily calibration with the former Second Standard set were on average < 0.008 ppb for all individual standards and are even smaller for the current set. This suggests very small uncertainties for the FCL internal scale transfer based on 18 - 25 calibration points using the FCL Primary Standard Set (initial Secondary Standard set) and more than 110 measurement days to assign the current Secondary Standard set using the initial FCL Secondary set.

8.6.3 N_2O long-term reproducibility

The time series of the target standard and the calibration fit residuals, respectively, indicate periods where the result stabilizes on varying levels within a very minor range. While the reason for this variability is not understood it is used to deduce an additional uncertainty of 0.02 ppb for long-term system changes that are not cancelled out by the standardization scheme.

8.6.4 N_2O measurement uncertainty estimate

Based on the above considerations the following combined standard uncertainty (k=1) is calculated as the square root of the sum of the individual uncertainty squares:

1. Scale link uncertainty = 0.11 ppb

- uncertainty of the CCL assignments for individual FCL Primary Standards (0.11 ppb)
- uncertainty of the FCL internal scale transfer to Secondary Standard assignments (0.026 ppb)

2. Measurement uncertainty of daily means = 0.015 ppb

- mean uncertainty of the daily calibration regression fit (0.011 ppb)
- uncertainty of the detector response drift throughout the validity of a daily calibration (0.01 ppb)
- uncertainty from the repeatability of the daily sample measurements (0.004 ppb)

3. Additional long-term variability = 0.02 ppb

In sum the accuracy with respect to the WMO Mole Fraction scale arises from the root of the sum of squares of the scale link uncertainty, the long-term reproducibility and the measurement uncertainty which amounts to 0.12 ppb (k=1). The FCL reproducibility is estimated to be 0.03 ppb.

The reproducibility derived from the target standard record is consistent with the uncertainty estimate for measurement of daily means.

The analytical precisions of many instruments that are involved in comparison activities are considerably inferior to the FCL LGR system. Therefore, the time series of these comparisons are mostly dominated by this scatter and contain little information on the LGR's reproducibility but the consistently small mean offsets support the uncertainty estimate. The mean offset relative to NOAA-GML based on measurement results for CCL assigned standards from partner labs and the MENI comparison samples are compatible with this uncertainty estimate.

References

[Andrews et al. 2014] Andrews, A. E., Kofler, J. D., Trudeau, M. E., Williams, J. C., Neff, D. H., Masarie, K. A., Chao, D. Y., Kitzis, D. R., Novelli, P. C., Zhao, C. L., Dlugokencky, E. J., Lang, P. M., Crotwell, M. J., Fischer, M. L., Parker, M. J., Lee, J. T., Baumann, D. D., Desai, A. R., Stanier, C. O., De Wekker, S. F. J., Wolfe, D. E., Munger, J. W., and Tans, P. P.: CO_2, CO, and CH_4 measurements from tall towers in the NOAA Earth System Research Laboratory's Global Greenhouse Gas Reference Network: instrumentation, uncertainty analysis, and recommendations for future high-accuracy greenhouse gas monitoring efforts, Atmos. Meas. Tech., 7, 647-687, https://doi.org/10.5194/amt-7-647-2014, 2014

[ATC 2020] ICOS Atmospheric Station specifications Version 2.0 - September 2020.pdf; section 4; pp 41ff (accessed March 2022)

[CCL_CO 2017] Information published at CCL's webpage: https://gml.noaa.gov/ccl/co_scale.html (accessed March 2022)

[CCL_CO_2 2017] Information published at CCL's webpage: https://gml.noaa.gov/ccl/co2_scale.html (accessed March 2022)

[CCL_N_2O 2011] Information published at CCL's webpage: https://gml.noaa.gov/ccl/n2o_scale.html (accessed March 2022)

[Crotwell 2017] Presentation by A. Crotwell at the WMO meeting GGMT2017: https://wmoomm.sharepoint.com/sites/wmocpdb/eve_activityarea/Forms/AllItems.aspx?id=%2Fsites%2Fwmocpdb%2Feve%5Factivityarea%2FGlobal%20Atmosphere%20Watch%20Programme%20%28GAW%29%5F63452102%2D7575%2De911%2Da98e%2D000d3a44bd9c%2FGGMT2017%5FT05%5FCrotwell%2Epdf&parent=%2Fsites%2Fwmocpdb%2Feve%5Factivityarea%2FGlobal%20Atmosphere%20Watch%20Programme%20%28GAW%29%5F63452102%2D7575%2De911%2Da98e%2D000d3a44bd9c&p=true (accessed March 2022)

[Crotwell 2019] Presentation by A. Crotwell at the WMO meeting GGMT2019: https://wmoomm.sharepoint.com/sites/wmocpdb/eve_activityarea/Forms/AllItems.aspx?id=%2Fsites%2Fwmocpdb%2Feve%5Factivityarea%2FGlobal%20Atmosphere%20Watch%20Programme%20%28GAW%29%5F63452102%2D7575%2De911%2Da98e%2D000d3a44bd9c%2FMeetings%2F2019%2FGGMT2019%2FOral%2FT5%5Fcrotwell%5Fggmt%5F2019%2Epdf&parent=%2Fsites%2Fwmocpdb%2Feve%5Factivityarea%2FGlobal%20Atmosphere%20Watch%20Programme%20%28GAW%29%5F63452102%2D7575%2De911%2Da98e%2D000d3a44bd9c%2FMeetings%2F2019%2FGGMT2019%2FOral&p=true (accessed march 2022)

[Hall 2017] Presentation by B. Hall at the WMO meeting GGMT2017: https://wmoomm.sharepoint.com/sites/wmocpdb/eve_activityarea/Forms/AllItems.aspx?id=%2Fsites%2Fwmocpdb%2Feve%5Factivityarea%2FGlobal%20Atmosphere%20Watch%20Programme%20%28GAW%29%5F63452102%2D7575%2De911%2Da98e%2D000d3a44bd9c%2FGGMT2017%5FT04%5FHall%2Epdf&parent=%2Fsites%2Fwmocpdb%2Feve%5Factivityarea%2FGlobal%20Atmosphere%20Watch%20Programme%20%28GAW%29%5F63452102%2D7575%2De911%2Da98e%2D000d3a44bd9c&p=true (accessed March 2022)

[Hall 2021] Hall, B. D., Crotwell, A. M., Kitzis, D. R., Mefford, T., Miller, B. R., Schibig, M. F., and Tans, P. P.: Revision of the WMO/GAW CO_2 Calibration Scale, https://doi.org/10.5194/amt-2020-408

[WMO 2018] Crotwell, A., & Steinbacher, M. (Eds.). (2018). *19th WMO/IAEA meeting on carbon dioxide, other greenhouse gases and related measurement techniques (GGMT-2017)*. *GAW Report: Vol. 242*. *19th WMO/IAEA meeting on carbon dioxide, other greenhouse gases, and related measurement techniques (GGMT-2017)*. Feb. 2019 retrieved from https://library.wmo.int/index.php?lvl=notice_display&id=20698#.W_5gRvJgifE; Section 1.5 a General recommendations for the quality control of atmospheric trace gas measurements, p. 12

[CCQM-P188] International comparison-Report of the Pilot study CCQM-P188 (in parallel with CCQM-K120.a and b): Carbon dioxide in Air (380-800) µmol/mol; Draft B; Edgar Flores*1, Joële Viallon1, Tiphaine Choteau1, Philippe Moussay1, Faraz Idrees1, Robert I. Wielgosz1, C. Meyer2, Daniel Rzesanke 3.; 2018

Annex I

Analysis of CO_2 and CH_4 mole fractions in reference standard mixtures at near-ambient concentration levels

Instrumentation:

Analysis of dried atmospheric air samples, pressurized in high pressure cylinders is performed by a Picarro Inc. G2301 CO_2/CH_4 Cavity Ringdown Spectroscopy (CRDS) Analyzer. The instrument (S/N CFADS2193) retrieves concentrations by analyzing the characteristic absorption of light of infrared-active molecules (near-IR spectroscopy).

Procedure:

Sample flow and cell pressure are controlled in an automated way and protocolled by the instrument.

The sample is provided via an external multi-position valve (VICI Valco, EMT2C16UWE; MPV) to the instrument's inlet. Commonly up to 16 high pressure cylinder air samples are analyzed within a sequence.

For data collection and synchronization of the MPV position and detector data an additional, external PC supervises the setup (see Fig.1). The resulting data and .log files are compiled by this PC and provided to the lab internal data management and data storage system. Following the automated migration of the raw data into the central data base quality checks and calibration of the instrument are self-controlled performed.

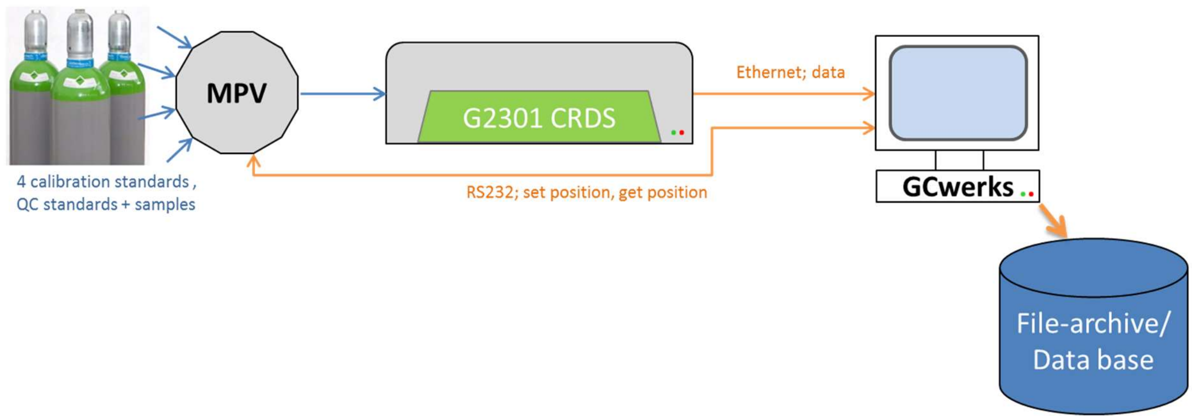


Figure 1: Schematics of the instrumental setup, blue lines= sample, orange lines=data/commands

Mode of Operation:

The operator defines the sequence of analysis using the GCwerks software at the supervising PC. Required information is shown in Table 1 and includes the date and time of initial connection, a MPV port number, sample identifier and meta information like the specific regulator mounted or the type of the sample. This information is stored in a *ports.log*-file, that supplies identifiers for the GCwerks-internal data base and sequencing as well as meta information for later summary purposes.

In a second step, the operator sets up the sequencing of the sampling ports stored as **.sequence*-file (as shown in Table 2). This list contains the port to be addressed and the residence time at this position as well as the runfile, that specifies the parameter set for this sample analysis. In the subsequent results file, both input files are merged with the raw data to automatically link the data collected during a specific port position to the respective sample identifier.

To keep the optical cell dry and maintain the instrument in a defined state the default sequence terminates with continuous analysis of a purge gas. The analyzed sample gas is discarded.

Date	Time	# port	Sample	Regulator	Type
170711	1200	6	i20150060	Tes1-021	qc
170711	1440	4	i20150062	Sco2-005	qc
170711	1440	3	i20150062	Sco2-005	qc
170711	1440	1	i20140171	Tes1-009	cal
170711	1440	5	i20140172	Sco2-001	cal
170711	1440	9	i20140173	Sco2-002	cal
170711	1440	13	i20140174	Tes1-004	cal
170711	1440	12	i20170017	Tes1-007	tank
170711	1440	14	i20170205	Sco3-008	tank

Table 1: Exemplary *ports.log* meta look-up table

While preparing the schedule, the operator has to make sure, that every sequence contains at least one Quality Control Standard (Target) and that for each calendar day the four FCL Secondary Calibration Standards have to be analyzed once at least.

Table 2 shows an exemplary analysis sequence containing 3 target samples (qc), 4 calibration gases (cal), a sample (tank) and the closing purge gas (for 60 minutes).

Duration [min]	Procedure	Type	# port
20	picarro.runfile	qc	6
20	picarro.runfile	qc	3
20	picarro.runfile	qc	4
20	picarro.runfile	cal	1
20	picarro.runfile	cal	5
20	picarro.runfile	cal	9
20	picarro.runfile	cal	13
20	picarro.runfile	tank	12
60	picarro.runfile	tank	14

Table 2: Exemplary sampling sequence

Every sample is fed to the analyzer for 20 minutes. During the initial 5 minutes the results are discarded with respect to running-in effects, like purging of the tube and allows for equilibration in pressure regulators, thermal equilibration and settling of the regulating loops. The instrument itself runs up to 0.2 Hz analysis frequency but raw data is aggregated in 60 s integration intervals to reduce the data volume. This leaves the opportunity to observe the sampling time series for subsequent flagging and averaging.

The optical cell is evacuated to 140 Torr, so the sample has not necessarily to be provided at over pressure. Pressure regulators (either and most common Tescom 64 series regulators or Scott Specialties 14C series brass regulators) are mounted on the cylinders at least the day before the analysis, flushed and stored pressurized with closed cylinder head valves. Before analysis this pressure is released and a slight overpressure of about 100mbar is generally adjusted to purge the regulators. This purging step, with pressurization followed by pressure release at closed cylinder head valve is performed three times.

The instrument is calibrated on a daily basis. The operator has to ensure that an analysis of the FCL Secondary Calibration Standards occurs within each calendar day. If it is more frequent the raw results of these standards are averaged for a daily mean. During data processing the daily mean calibration standard data are fitted by a regression function to their assigned concentration values to calculate the calibration coefficients of this day. For calibration of CO_2 and CH_4 a linear equation is applied.

Five dedicated samples, called Targets are regularly analyzed for quality control of the instrument's performance including the daily calibrations. Two of them are included within every sequence. They have CO_2 and CH_4 concentrations close to the boundaries set by the range of the calibration gases to give a conservative assessment that is meaningful for all concentrations. The two additional targets are analyzed less frequently (at least four times a year) as "long term targets" to assess long term variability and potential drifts of the instrument's calibration suite. A fifth QC standard is shared between different instruments in the laboratory to assess the link of their respective results on a regular basis.

Instrument calibration:

Measurement data are calibrated relative to the current WMO Mole Fraction Scales for all reported gas mole fractions. The current scales that are maintained by NOAA ESRL as Central Calibration Laboratory (CCL) are: WMO CO_2 X2007 and WMO CH_4 X2004A. For spring 2021 a revision of the WMO/GAW CO_2 Calibration Scale is announced (X2019).

The traceability to these scales is realized by a suite of high pressure standard gases calibrated by the WMO Central Calibration Laboratory. The link is actively maintained by regular re-calibrations of subgroups of these laboratory calibration standards. The respective standard cylinders are listed in Table 3 with the currently used assigned values. These values may change due to scale revisions by the Central Calibration Laboratory or additional measurements done by the CCL. The updated data is available in the internet (http://www.esrl.noaa.gov/gmd/ccl/refgas.html).

With this CCL-calibrated reference suite (FCL Primary Calibration Standards) the daily used FCL Secondary Calibration Standards were calibrated/assigned at the same instrument.

Cylinder ID	Fill date	CO_2 [ppm]	CH_4 [ppb]	Last CCL analysis
CB09944	07/2013	339.24	1596.64	06/2017
CB09939	07/2013	365.12	1743.13	10/2018
CB09958	07/2013	389.53	1896.82	09/2016
CB09983	07/2013	412.21	2032.92	10/2018
CB09952	07/2013	433.58	2195.34	11/2016
CB09955	07/2013	458.92	2344.05	07/2017
CB09957	07/2013	481.75	2466.72	09/2016
CB09934	07/2013	514.79	2731.28	06/2017
CB09948	07/2013	250.08	2932.82	10/2018

Table 3: Calibration standards assigned by the WMO Central Calibration Laboratory

Data evaluation:

The detector response function and the mole fractions of the various trace species in the FCL Secondary Standard are determined by analysis of a suite of laboratory standard gases measured by the WMO Central Calibration Laboratory (see Table 3). Measurements of these highest level laboratory calibration standards are generally repeated four times a year to capture small changes in the composition of the FCL Secondary Standards or in cases where quality control measurements suggest sudden changes.

To evaluate the validity of the analytical results the following is regularly checked:

- Instrumental parameters during analysis (sampling frequency; cell temperature as well as pressure level and variability).
- Baseline drift and noise when the instrument is purged with a constant sample;
- The measurement results of the target standards relative to their known composition
- The regression fit residuals of the associated daily calibration and their time series.

Measurements are flagged invalid in cases where instrumental variables indicate a system malfunctioning or if the sample flow points to insufficient supply.

Annex II

Analysis of CO and N_2O mole fractions in reference standard mixtures at near-ambient concentration levels

Instrumentation:

Analysis of dried atmospheric air samples, pressurized in high pressure cylinders is performed by a Los Gatos Research Inc. CO/N_2O-analyzer Enhanced Performance (LGR). The instrument (S/N 15-0140) retrieves concentrations by analyzing the characteristic absorption of light of infrared-active molecules (near-IR spectroscopy). The instrument's uses the technical principle of Off-Axis Integrated Cavity Output Spectroscopy (OA-ICOS) is implemented (see Fig.1).

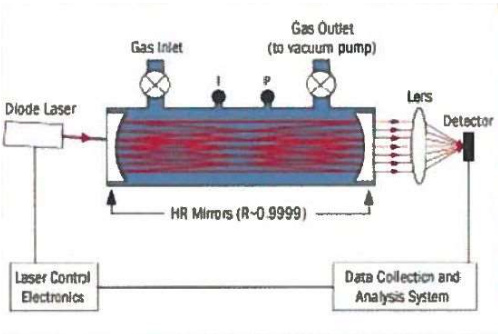


Figure 1: Schematic diagram of an OA-ICOS analyzer

Data retrieval is performed with tunable-laser absorption-spectroscopy (TDL) by scanning a narrowband wavelength across the absorption band of a target species to record the loss in the emitted light (ref. Fig.2). Under knowledge of the gas temperature, pressure in the cell, effective path length and known line strength the concentration can be calculated from the integrated loss-signal following Lambert-Beer's-Law.

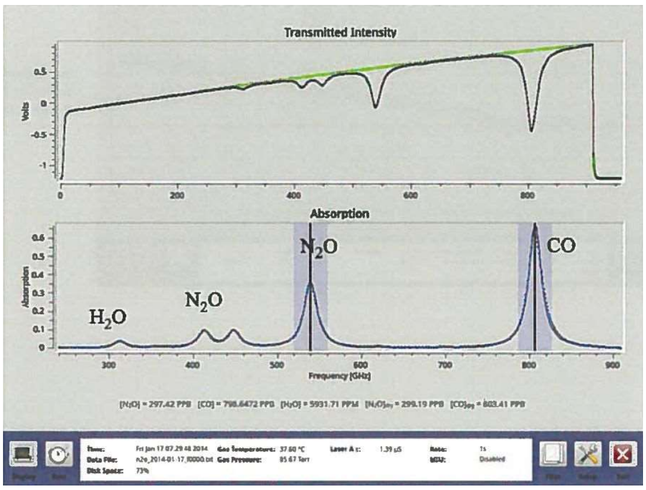


Figure 2: Screen shot of spectrum display, upper panel shows photo detector voltage, lower panel shows optical absorption of species of interest.

Instrumental Setup:

Sample flow and cell pressure are controlled and protocolled automatically. Figure 3 gives an overlook of the sample flow and meta information retrieval within the instrument.

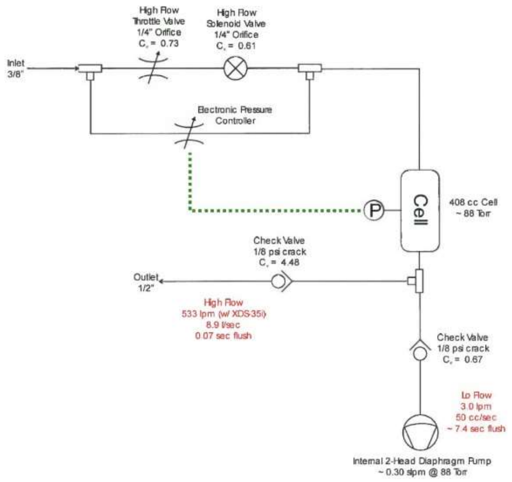


Figure 3: Internal flow schematics of the LGR instrument

The sample is provided via an external multi-position valve (MPV; VICI Valco EMT2C16UWE) to the instrument's inlet. Commonly up to 16 high pressure cylinder air samples are analyzed within a sequence. The analyzed sample gas is discarded.

For data collection, synchronization of the MPV and merging of position and detector data an additional, external PC supervises the setup (see Fig.4). The resulting data and .log files are compiled by this PC and provided to the lab internal data management and data storage system. Following the automated parsing process to migrate the raw data into the central data base the data processing includes a short term stability correction, automated quality checks and automated calibration of the instrument.

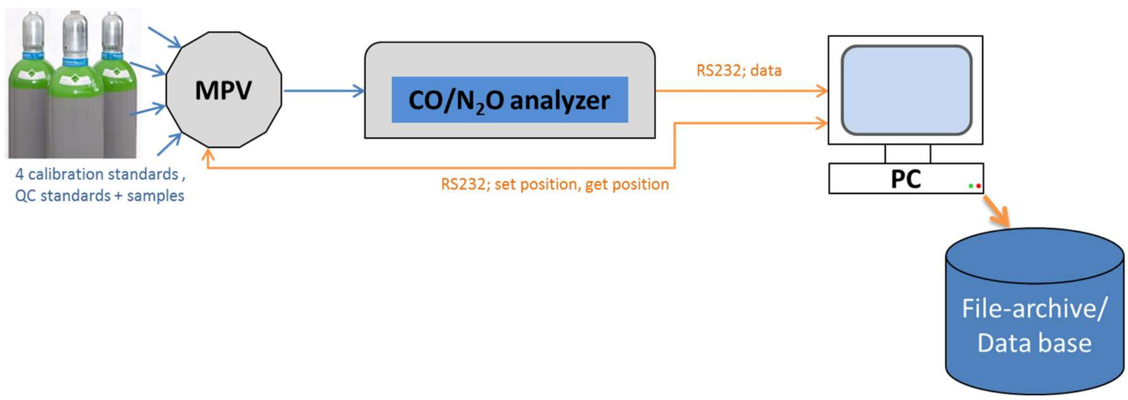


Figure 4: Schematics of analysis station, blue lines= sample, orange lines=data/commands

Mode of Operation:

The operator defines the analysis sequence using an in-house programmed software at the supervising PC. Required information to be entered is shown in Table 1 and includes the sample identifier, measurement duration, and the port number of the multi position valve.

The mandatory structure of the sample sequence scheme is:

1. *Every* sample analysis has to be bracketed by analysis of the Working Standard (WT) that is used for short term drift correction.

2. The first samples in the sequence have to be the calibration gases for the automated data processing.

3. Every sequence has to include the analysis of minimum one Target Standard that is analyzed for quality control purposes.

4. Samples described as "purging" are ignored and not transferred to the database.

Table 1 shows an exemplary analysis sequence containing 4 Calibration Standards, a sample, 3 Target Standards (QC) and the periodic Working Standard (WT). To keep the optical cell dry and maintain the instrument in a defined state the default sequence terminates with continuous analysis of a dried purge gas.

Duration [min]	# port	Sample	Description
20	10	i20160515	WT
20	2	i20150251	Cal1_CA05640
20	10	i20160515	WT
20	3	i20150189	Cal2_D073397
20	10	i20160515	WT
20	6	i20150544	Cal3_D073396
20	10	i20160515	WT
20	5	i20150191	Cal4_D073395
20	10	i20160515	WT
20	1	i20160123	QClow_D073388
20	10	i20160515	WT
20	11	i20170274	sample_D073386
20	10	i20160515	WT
20	13	i20150188	QCinter_D073398
20	10	i20160515	WT
20	4	i20160147	QChigh_D801333
20	10	i20160515	WT
720	15	i20170299	purging

Table 1: Exemplary sampling sequence

Every sample is fed to the analyzer for 20 minutes. During the initial 10 minutes the results are discarded due to running-in effects like sample purging of and equilibration in pressure regulators, thermal equilibration and settling of the regulating loops. The instrument itself runs up to 1 Hz analysis frequency but raw data is aggregated in 20 s integration intervals to reduce the data volume. This leaves the opportunity to observe the sampling time series for later flagging and averaging. This 20 s averaging interval set by the LGR instrument is not synchronized with the valve switch schedule set by the controlling software such that there is the possibility that the last data point combines the signals of two subsequent samples. Therefore, the very last data point is generally discarded. The remaining 20s-data points are the raw reading of this analysis.

The optical cell is evacuated to 85 Torr, so the sample has not necessarily to be pressurized. Pressure regulators (either and most common Tescom 64 series regulators or Scott Specialties 14C series brass regulators) are mounted on the cylinders at least the day before the analysis, flushed and stored pressurized with closed cylinder head valves. Before analysis this pressure is released and a slight overpressure of about 100mbar is generally adjusted at the inlet to purge the regulators.

Every sample analysis (including the Calibration Standards) is bracketed by analysis of the Working Standard (WT_{prior}, WT_{after}). Thus short term drifts of the analyzer are accounted for by normalization to the Working Standard's raw signal in the same way for unknown samples as for Calibration Standards:

$$C_{corr} = 2\frac{C_{raw}}{\left(\frac{WT_{prior}}{WT_{ref}}+\frac{WT_{after}}{WT_{ref}}\right)},$$

with C_{raw} – raw signal of sample, C_{corr} – the normalized sample and WT_{ref} - assigned value of the Working Standard Tank.

Every sequence has to be started by the set of the four FCL Secondary Calibration Standards. If all 16 available ports are occupied with bracketing by the WT and sampling time of 20min, an analysis takes no longer than 11 hours. Therefore, the instrument is practically calibrated on a daily basis.

During data processing the normalized calibration standard data are fitted by a regression function to their assigned concentration values to calculate the calibration coefficients of this run. For calibration of CO a linear and for N_2O a quadratic equation is applied.

Three dedicated standards, called Targets are regularly analyzed for quality control the instrument's performance including the daily calibrations. Two of them are included within every sequence. They have CO and N_2O concentrations close to the boundaries set by the range of the calibration gases to allow a conservative assessment that is meaningful for all concentrations. The third QC is shared between different instruments in the laboratory to assess the link of their respective results on a regular basis. It serves as "long term target" to assess long term variability and potential drifts of the calibration suite.

Instrument calibration:

Measurement data are calibrated relative to the current WMO Mole Fraction Scales for all reported gas mole fractions. The current scales that are maintained by NOAA ESRL as Central Calibration Laboratory (CCL) are: WMO CO X2014A and WMO N_2O X2006A.

The traceability to these scales is realized by a suite of nine high pressure standard gases calibrated by the WMO Central Calibration Laboratory. The link is actively maintained by regular (annual) re-calibrations of subgroups of these FCL Primary Calibration Standards by the CCL. The respective standard cylinders are listed in Table 2 with the currently used assigned values. These values may change due to scale revisions by the Central Calibration Laboratory or additional measurements done by the CCL. The updated data is available in the internet (http://www.esrl.noaa.gov/gmd/ccl/refgas.html).

With this CCL assessed reference suite the FCL Secondary Calibration Standards (used on a daily basis) are calibrated at the same instrument. Measurements of the FCL Primary Calibration Standards are generally repeated four times a year.

Cylinder ID	Fill date	CO [ppb]	N2O [ppb]	Last CCL analysis
CB09944	07/2013	31.31	316.77	06/2017
CB09939	07/2013	80.14	319.86	02/2019
CB09958	07/2013	120.69	327.12	09/2016
CB09983	07/2013	158.92	329.92	01/2019
CB09952	07/2013	199.47	334.60	11/2016
CB09955	07/2013	247.14	339.48	07/2017
CB09957	07/2013	397.06	343.95	09/2016
CB09934	07/2013	697.56	349.13	06/2017
CB09948	07/2013	998.63	362.13	01/2019

Table 2: Calibration standards assigned by the WMO Central Calibration Laboratory

Data evaluation:

A regular analysis sequence consists of alternate measurements of the Working Standard and Target Standards that are used for quality control assessment. Raw data of any sample measurement is normalized to the raw data of the Working Standard to cancel out instrumental drifts within hours (triggered e.g. by variations in atmospheric pressure or other laboratory environment variations). The detector response function and the mole fractions of the various trace species in the Working Standard are determined by analysis of the FCL Secondary Calibration Standards.

To evaluate the validity of the analytical results the following is regularly checked:

- Instrumental parameters during analysis (sampling frequency; cell temperature as well as pressure level and variability),
- Baseline drift and noise when the instrument is purged with a constant sample;
- Every Working Standard raw data 10 min mean relative to the means of the preceding and subsequent Working Standard measurements,
- Measurement results of the Target Standards relative to their known composition,
- Regression fit residuals of the associated daily calibration.

Measurements are flagged invalid in cases where instrumental variables indicate a system malfunctioning or if the sample flow points to insufficient supply. In cases of an invalid Working Standard measurement it is checked if this individual reference point can be replaced by the next Working Standard measurement result. However, it is also checked if this failed standard measurement indicates a problem that makes the sample measurement also invalid and has to be flagged accordingly.

Annex III

Overview of flagging parameters for measurements performed with Picarro and Los Gatos analyzers

Picarro:

Flag	description
NSigma	NSIGMA with a sigma factor of 3
RI	RUNNINGIN with a running in duration of 300 seconds
P	PCavity with range 139.99 ... 140.01
MISS CO2	Missing value in CO2 related measurements
MISS CH4	Missing value in CH4 related measurements
OPV	OutletProportinalValve Flag 28800 … 34000
SDMinRaw CO2	Standard deviation of MinRaw data, range 0 … 0.035
SDMinRaw CH4	Standard deviation of MinRaw data Flag, range 0 … 0.3
INMinRaw CO2	Insufficient number (of MinRaw values)
INMinRaw CH4	Insufficient number (of MinRaw values)
RC	Insufficient number of calibration gases, <4
CO2_DYN_sd	static dynamic upper bound: 0.070 threshold: 450
CH4_DYN_sd	static dynamic upper bound: 0.6 threshold: 2300

Los Gatos:

Flag	description
NSigma	NSIGMA with a sigma factor of 3
RI	RUNNINGIN with a running in duration of 540 seconds (28 measurement points)
RO	RUNNINGOUT with a running out duration of 5 seconds (1 measurement point)
Gas pressure	with range 85.17 ... 85.28
Gas pressure sd	with range 0 ... 0.006
MISS CO	Missing value in CO related measurements
MISS N2O	Missing value in N2O related measurements
H2O	leakage on the basis of water signal
H2O sd	leakage on the basis of water signal stdev
CO sd	Standard deviation of CO out of range, -1.0 … 0.00014
N2O sd	Standard deviation of NO2 out of range, -1.0 … 0.0001
RC	Insufficient number of calibration gases, <4
N2O_DYN_sd	Dynamic upper bound, Minimum: 4.0e-04 dyn_poly: 6.357375e-04
CO_DYN_sd	Dynamic upper bound, Minimum: 2.2e-04 dyn_poly: 1.074092e-03