ICOS Central Analytical Laboratories

Flask and Calibration Laboratory (FCL)

Quality Control Report 2024

Armin Jordan, Fadwa Damak

ICOS Central Analytical Laboratory

Jena, May 2025

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 2023 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. For CO_2 and CH_4, the first analyzer named Picarro2 replaced Picarro1 in February 2024 (refer to section 2 and Annex I), thus the two different uncertainty terms values.

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, combined uncertainties are calculated as the square root of the sum of squared uncertainty contributions

	CO_2 [ppm] (Picarro1/Picarro2)	*CH_4 [ppb] (Picarro1/Picarro2)*	*CO [ppb]*	*N_2O [ppb]*
CCL reproducibility[1)]	*0.01*	*0.5*	*0.4*	*0.11*
scale propagation to FCL standards	*0.01[2)]*	*0.05[3)]*	*2[4)]*	*0.006[5)]*
scale link uncertainty	***0.012***	***0.5***	***2[6)]***	***0.11***
instrumental precision	*0.015/0.011*	*0.24/0.08*	*0.05*	*0.024*
long-term reproducibility	*0.006*	*0.07*	*NA[7)]*	*0.02*
estimated FCL reproducibility	***0.019/0.016***	***0.3/0.14***	***2***	***0.03***
estimated overall uncertainty	***0.022/0.020***	***0.58/0.50***	***2***	***0.11***

[1)] WMO Central Calibration Laboratory (CCL)

[2)] This does not include a bias resulting from an incorrect accounting of the CO_2 stable isotopic composition of the FCL Primary Standards (see Annex IV).

[3)] Refers to uncertainty to the FCL standards in use since December 2020. N.B. Provisional scale propagation uncertainty until 2020 is 0.2 ppb (refer to section 6.6.4).

[4)] The term includes the uncertainty of CO growth in FCL Secondary Standards.

[5)] After correction of the initial assignment bias of the first Secondary Standard set.

[6)] For CO mole fractions at atmospheric background levels.

[7)] Refer to section 7.6.3.

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 of 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. This report focuses 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 at the end of each daily sub-sequence 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 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)).

On February 6th 2024, the Picarro used since the early days of FCL in 2015 was finally taken out of service after a period of accelerating deterioration of performance related to laser ageing. From February 8th to 27th, a replacement Picarro was set up in place and the measurement method (outlet valve value, flushing and measurement times) was tested and optimized.From then onwards, E]each measurement (samples as well as references) takes thirty minutes of gas injection instead of the previous twenty minutes. To avoid cross contamination of succeeding samples, the new method measures the same flushing gas for five minutes between each samle (samples and references) . To flush out the pressure regulator, the first ten minutes of data at the beginning of each sample measurement are ignored and the average of the remaining valid twenty minutes data is further processed.

The instrument is calibrated on a daily basis by a dedicated set of four FCL Secondary 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 built 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 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 calibrated at the CCL four times with the most recent recalibration having been made in 2024. This shall allow to verify the stability of the respective trace gases or track the rate of change of their mole fraction. Some tracers have been analyzed using different measurement techniques at the CCL and for CH4 and CO2 not all calibration results are considered (see sections 5.1 and 6.1). The table below shows the results of the last CCL calibration for each gas.

Cylinder ID	Sample ID	Fill date	last CCL calibration date	CO_2 (ppm)[1]	CH_4 (ppb)[2]	CO (ppb)[3]	N_2O (ppb)[4]
CB09948	i20140054	07/2013	11/2024	250.11	2932.82	1000.81	362.06
CB09944	i20140055	07/2013	11/2024	339.34	1596.57	38.87	316.74
CB09939	i20140056	07/2013	11/2024	365.28	1743.20	86.60	319.88
CB09958	i20140057	07/2013	11/2024	389.75	1896.87	126.93	327.16
CB09983	i20140058	07/2013	11/2024	412.41	2032.83	164.25	329.86
CB09952	i20140059	07/2013	11/2024	433.84	2195.21	204.54	334.44
CB09955	i20140060	07/2013	11/2024	459.17	2343.89	250.33	339.37
CB09957	i20140061	07/2013	11/2024	482.01	2466.79	398.86	343.69
CB09934	i20140062	07/2013	11/2024	515.11	2731.85	696.45	348.95

WMO Mole Fraction scale:	**CCL-reproducibility** (1 sigma) [reference]:
[1] CO_2WMO X2019 (CRDS only)	0.01 ppm [CCL_CO_2 2021]
[2] CH_4 WMO X2004A	0.5 ppb (pers. comm., E. Dlugokencky, Feb. 2018)
[3] CO WMO X2014A	0.4 ppb for CO < 400 ppb [CCL_CO 2018]
[4] N_2O WMO X2006A	0.11 ppb [CCL_N_2O 2011]

TABLE 2: *FCL PRIMARY STANDARDS ASSIGNMENTS BY THE WMO CENTRAL CALIBRATION LABORATORY*

4 QA/QC 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 for Greenhouse Gases 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 have been re-sent to the CCL for recalibration on an annual basis for the first three years. A next batch of re-calibrations is planned for 2024.
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. The FCL Secondary Standard assignments are made a certain point in time and in general kept fixed despite the comparisons to the FCL Primary Standards are being continued. A re-evaluation of these Secondary Standard assignments is commonly not made before they are fully exhausted and thus the record of Primary Standard calibrations has been completed.
3. *Targets*: The performance of daily measurements is characterized by daily analysis of the same gases in high-pressure cylinders 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* are made routinely. Initially an intensive exchange of samples analyzed at the FCL and the MPI-BGC GasLab was made which is still ongoing with lower frequency. International comparisons with a large group of laboratories are performed in the "Sausage Intercomparison Program" (using flask samples), and within the “MENI” (**M**PI-BGC, **E**MPA, **N**OAA and **I**COS) - Intercomparison that includes among others the NOAA-GML as partner laboratory. Additional such activities that FCL is involved are of more sporadic nature (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 uncertainty (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 used Secondary Standards' assigned values with the results obtained from repeated further 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 Secondary Standards.
- The scatter of the daily residuals is an indicator for 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 another quality control of the actual scale transfer uncertainty.

Inter-Instrument comparison

- The agreement of analysis results of the same sample by different detecting techniques provides the chance to identify and quantify potential analytical biases related to either of the techniques.
- 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, the accuracy of the reference material, the standardization strategy and data processing.

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 and again in 2024, the complete set received additional recalibrations such that four CCL assignments from different years are now available for each standard. The initial calibration was performed using only the NDIR (L9) 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, three 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. The standard approach for X2019 assignments is based on CRDS measurements in combination with the determination of the CO_2 stable isotope composition of the respective standard gas. Therefore, CCL information based on NDIR measurements without consideration of the CO_2 isotopic composition are not further considered any more. Atmospheric observations of CO_2 are performed within ICOS almost exclusively using CRDS instrumentation that is selective for the $^{12}C^{16}O_2$ isotopologue only. The FCL Primary Standards are modified, dried real air. The modification involves addition of pure CO_2 to achieve the wanted composition resulting in standard gases with a CO_2 stable isotope composition that is similar to but not perfectly matching the range of naturally observed atmospheric CO_2. To account for this, the assigned values of the individual standards are adjusted for the offset resulting from the isotopic deviation between standard and atmosphere. The values specified in the second column on the right-hand side of Table 3 are those that are currently in use. It has recently been discovered that they are 0.02 µmol/mol too high (see details of the adjustment procedure as described in Annex IV.). The last column shows the corrected isotope-adjusted values based on the the total CRDS measurements, but those values will not be used until the CO_2 scale links update is implemented jointly between FCL and the ATC which is planned to happen in 2025.

Table 3: CO_2 *X2019 ASSIGNMENTS FOR FCL PRIMARY STANDARDS BY CCL [PPM].*

Sample ID	Cylinder ID	CCL date 1	CCL date 2	CCL date 3	CCL date 4	NDIR Date 1	CCL-CRDS date 2	CCL-CRDS Date 3	CCL-CRDS Date 4	adjusted CRDS used*	adjusted CRDS**
i20140054	CB09948	Dec-13	Oct-18	Mar-21	Aug-24	250.129	250.116	250.129	250.113	**250.144**	250.131
i20140055	CB09944	Mar-14	Jul-17	Mar-21	Aug-24	339.327	339.356	339.360	339.342	**339.387**	339.369
i20140056	CB09939	Dec-13	Oct-18	Mar-21	Aug-24	365.253	365.277	365.281	365.278	**365.306**	365.308
i20140057	CB09958	Dec-13	Oct-16	Mar-21	Aug-24	389.762	389.753	389.765	389.750	**389.781**	389.755
i20140058	CB09983	Feb-14	Oct-18	Mar-21	Aug-24	412.381	412.420	412.424	412.407	**412.447**	412.424
i20140059	CB09952	Jan-14	Sep-16	Mar-21	Aug-24	433.815	433.833	433.832	433.839	**433.853**	433.850
i20140060	CB09955	May-14	Jun-17	Mar-21	Aug-24	459.121	459.181	459.173	459.169	**459.224**	459.197
i20140061	CB09957	Feb-14	Aug-16	Mar-21	Aug-24	481.962	482.014	482.022	482.008	**482.068**	482.041
i20140062	CB09934	May-14	Jun-17	Mar-21	Aug-24	515.053	515.120	515.113	515.107	**515.183**	515.138

*adjustment: see Annex IV, based on CRDS date 2 results; ** adjustment based on total CRDS until 2024 (not in use yet)

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 all calibration events with the *complete* set of the primary standards. The mean residuals of the individual standards range from -0.019 ppm to +0.012 ppm with a standard deviation of these means of 0.011 ppm. This is a measure of the consistency of the initial CCL assignments confirming the specifications made by the CCL.

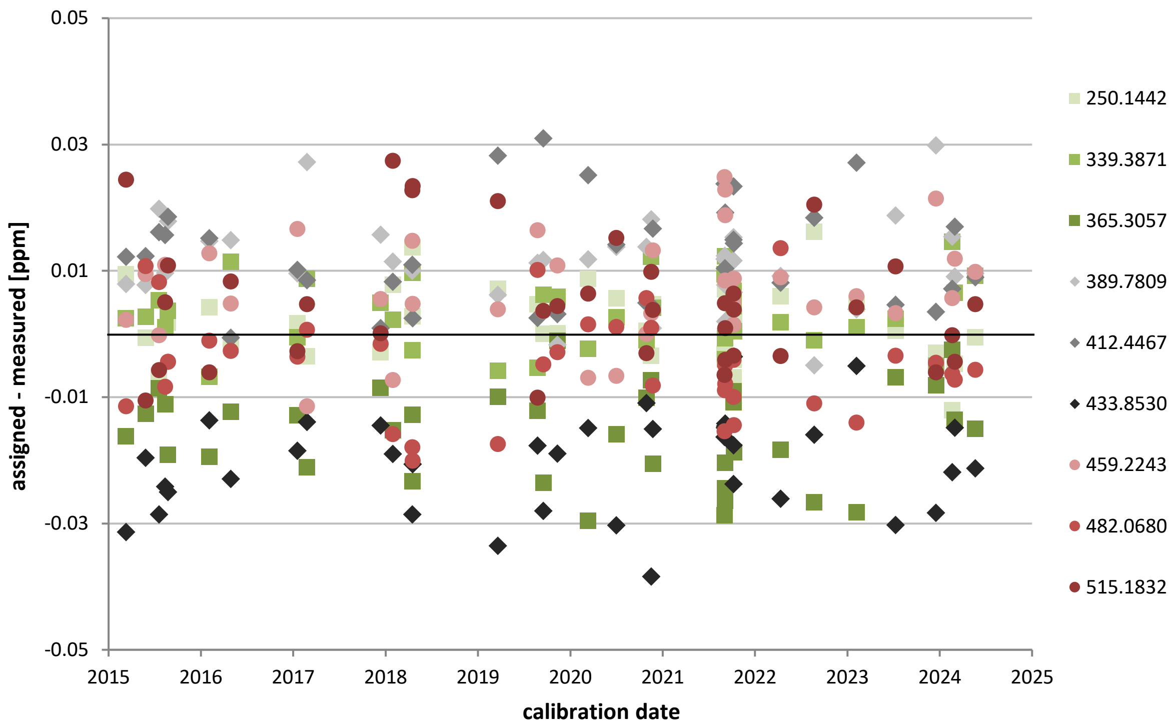


FIGURE 1: *TIME SERIES OF LINEAR REGRESSION FIT RESIDUALS OF THE CRDS CO_2 CALIBRATION FOR FCL PRIMARY STANDARDS*

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 25 to 29 valid calibration episodes together with the *complete* set of FCL Primary Standards between Feb 2015 and either July 2020 or September 2021. During 2020, the first set of FCL Secondary Standards had to be replaced by a new set because they were consumed. The replacement was done in two steps, with the replacement of the two standards with higher mole fractions made in June and the replacement of the two standards with lower mole fractions made in December.

The stability of CO_2 values for the second set of Secondary Standards had been monitored by repeated measurements against the first set of FCL Secondary Standards for an extended period. The assigned CO_2 X2019 values were based on the records of the CO_2 mole fraction results of the FCL primary calibration episodes between Aug 2019 and Oct 2021. With a limited number of calibration episodes there seemed to be an annual CO_2 growth of 0.01 ppm and more in all standards of the second set. This impression changed with further calibrations made in 2022. Calibrated results of target standard measurements also showed inconsistent behaviour that pointed to an overestimation of the CO_2 drift. Thus, assigned values of the second set were reassessed and none of the standards is currently assumed to grow CO_2 any more. The FCL CO_2 measurement results from June 2020 to April 2022 are still affected by this preliminary assignment error with maximum biases at the end of this period of 0.02 to 0.03 µmol/mol. While a correction at FCL internally would be a minor effort, it is a larger computational work load to reprocess all continuous CO_2 measurements in the ICOS network based on standards assigned by FCL during that time. This requires that the correction needs to be done in collaboration with the ATC in due course. Therefore, these will be rectified latest when this set of Secondary Standards will be replaced at the end of its lifetime. At that point of time the assignment history based on the FCL Primary Standards will be completed. Such a final assignment revision had been made already for the first set of Secondary Standards (see Table 4).

TABLE 4: *CO_2 ASSIGNMENTS OF FCL SECONDARY STANDARDS [PPM]*

Sample ID	Cylinder ID	Assigned Value[4]	Drift/yr[1]	*Date of change*	Re-assigned Value[5]	Re-assigned Drift/yr[2]
i20140171[1]	D801336	**359.870**	+0.003	*2020-12-08*		
i20140172[1]	D073384	**393.464**	+0.005	*2020-12-08*		
i20140173[1]	D073392	**424.724**	+0.007	*2020-06-23*		
i20140174[1]	D801331	**454.329**		*2020-06-23*		
i20190708[2]	D761202	**362.751**	+0.014	*2022-04-29*	**362.751**	
i20190803[2]	D073381	**402.078**	+0.010	*2022-04-29*	**402.077**	
i20190709[3]	D761214	**433.119**	+0.016	*2022-04-29*	**433.124**	
i20190438[3]	D073389	**450.779**	+0.017	*2022-04-29*	**450.784**	

Starting dates: [1]1st January 2015; [2]8th December 2020; [3]23rd June 2020

[4]Assigned value at start date; [5]Re-assigned value since date of change

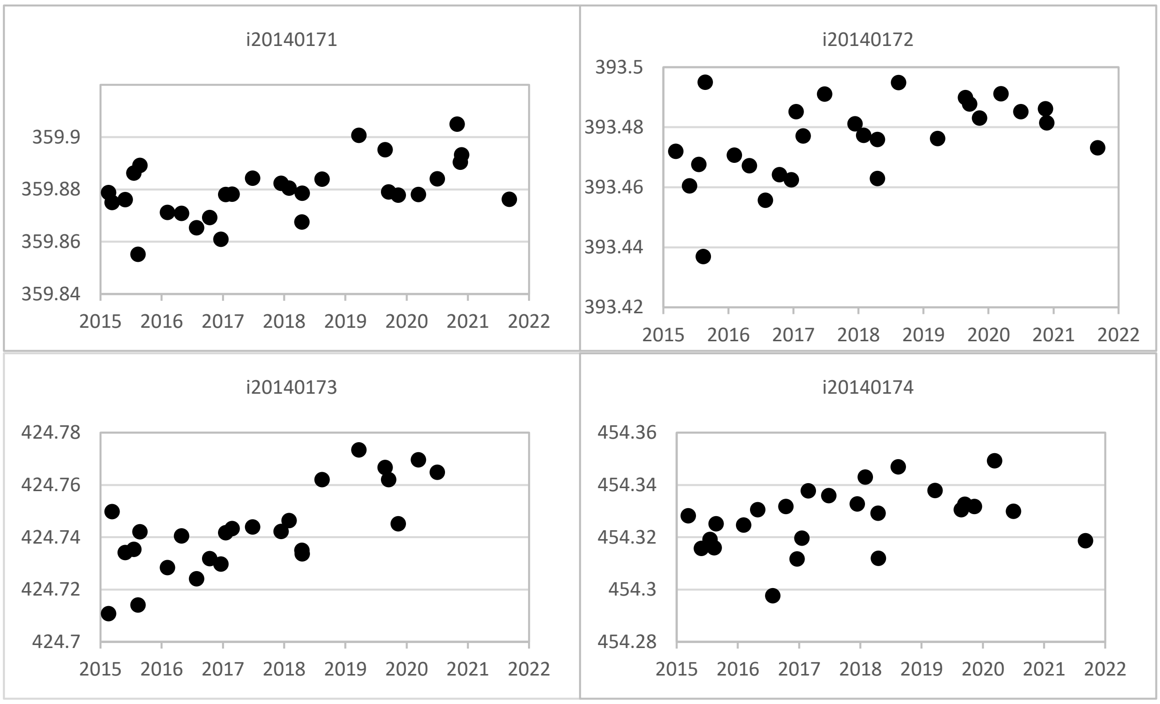


FIGURE 2: *: FCL SECONDARY STANDARDS CO2 ASSIGNMENT TIME SERIES, DATA OF THE FIRST SET (VALUES IN [PPM]).*

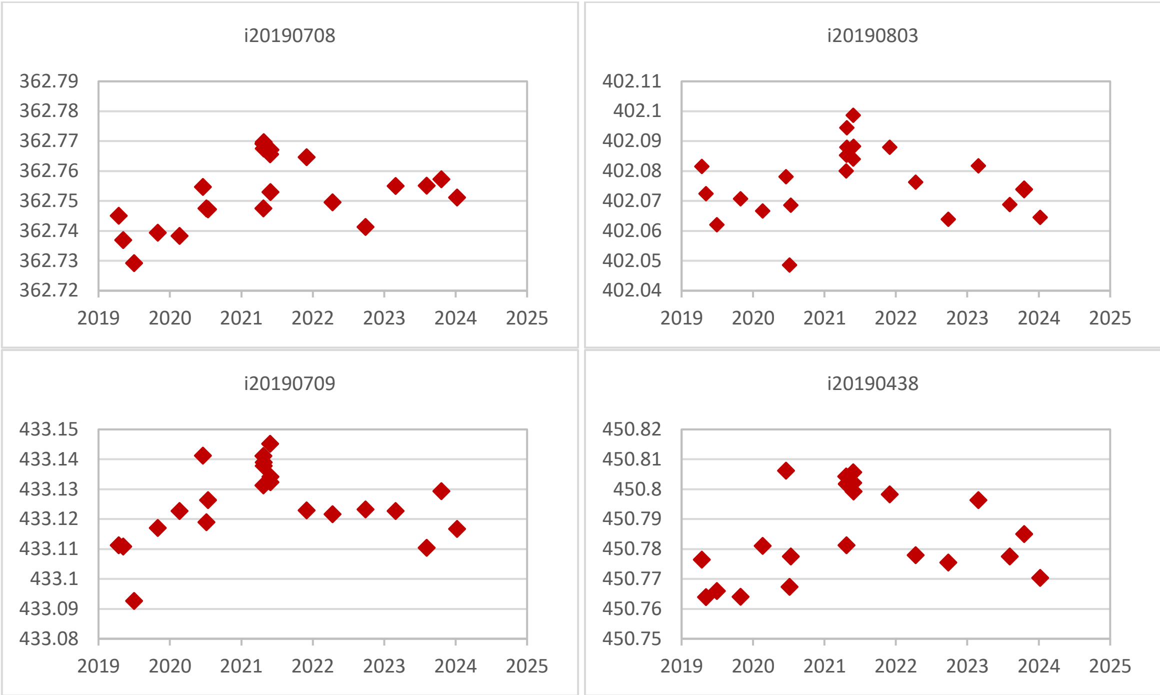


FIGURE 2: *FCL SECONDARY STANDARDS CO_2 ASSIGNMENT TIME SERIES, DATA OF THE SECOND SET (VALUES IN [PPM]).*

5.2.2 Residual record

The residuals of the linear regression of the FCL Secondary Standards are given in Figure 3. The mean absolute residuals for the Secondary Standards are on the order of 0.001 ppm and smaller. The standard deviation of the daily residuals for the four individual standard gases for the entire period amounts to maximum 0.005 ppm. These very small values of the mean residuals of all standards provide evidence for a consistent scale transfer to these FCL Secondary Standards. Trends in the residuals over the periods of the respective Secondary Standard sets do not exceed 0.01 ppm. This documents the long-term internal consistency of the calibration sets throughout their lifetime.

FIGURE 3: *TIME SERIES OF CO_2 LINEAR REGRESSION FIT RESIDUALS OF THE FCL SECONDARY STANDARDS. Blue symbols represent the first set of FCL Secondary Standards, red symbols the second set of FCL Secondary Standards. The dark symbols are indicating the transition phase when only two of the standards were replaced.*

5.3 CO_2 Targets

In the period from March 2015 to December 2024 two succeeding sets of each three Target Standards have been in use at the CRDS system. Two of those targets were succeeded by a third set of two respective standards from mid of 2022, with the second high CO_2 standard still continued to be measured daily until July 2024. Whereas the third target (i20170962) of the second set switched to a long-term standard that is measured on a monthly basis from end of August 2022. On a regular basis two further targets monitor the long-term stability of the instrument around 360 ppm. The Target Standards' daily 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 records of daily target mean residuals are below 0.017 ppm. There are two exceptional periods: firstly, the initial period until-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. Secondly, Target results are higher by up to 0.02 ppm in the period between June 2020 and April 2022. During this period only preliminary assignment information for the Secondary Standards were available and resulting in incorrect CO_2 growth estimates. While the diverging data in that second exceptional period will represent similar deviations of FCL assignments on ICOS standards, this bias will be corrected at a later point in time (see section 5.2.1).

FIGURE 4: *TIME SERIES OF THE OFFSET OF CO_2 TARGET MEASUREMENTS TO THEIR RESPECTIVE ASSIGNED VALUES. The dark lines represent 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 visualization).*

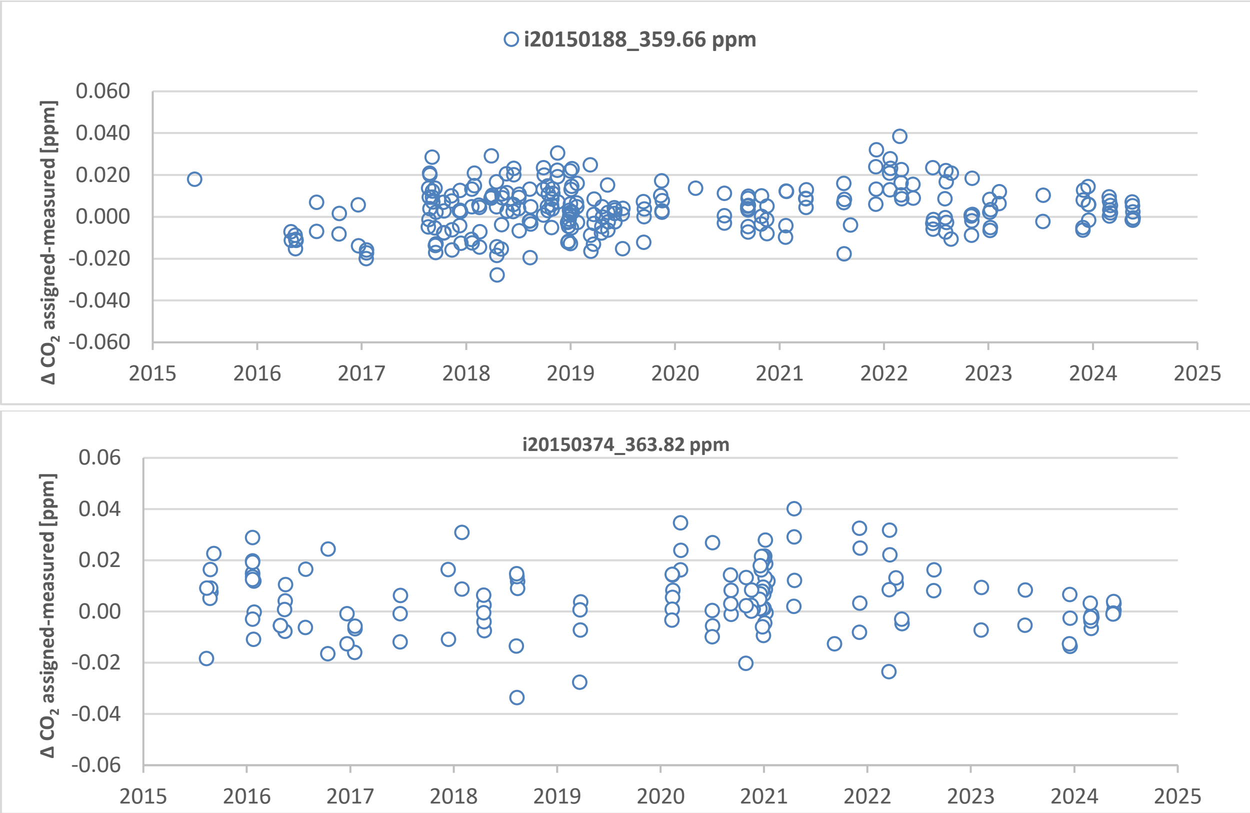
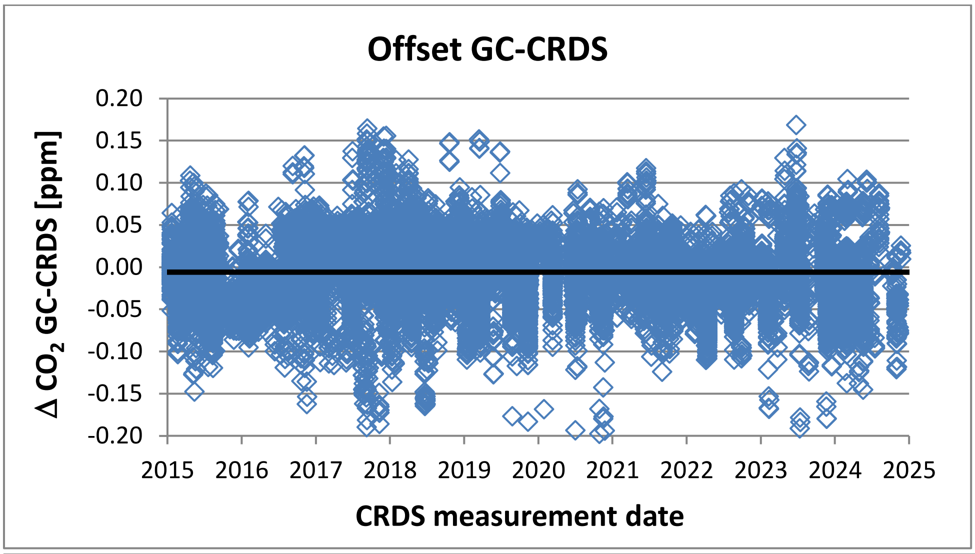


FIGURE 4: *TIME SERIES OF THE OFFSET OF CO_2 TARGET MEASUREMENTS TO THEIR RESPECTIVE ASSIGNED VALUES*

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 with at least ten injections. The inter-instrumental measurement differences for all standards analyzed within one year are depicted in Figure 5 (including only standards within the range defined by the calibration standards). On average there is no offset (-0.006 ppm ± 0.040 ppm), neither any evidence for a trend in time nor a systematic mole fraction dependency of the agreement.

Note that each data point in Figure 5 represents the difference of one CRDS daily mean result relative to the means of GC measurements of the same sample averaged over one calibration episode. Some samples have been analyzed much more frequently on the CRDS system than on the GC giving these latter measurements more weight in the figures which are based on 219 individual samples in total.



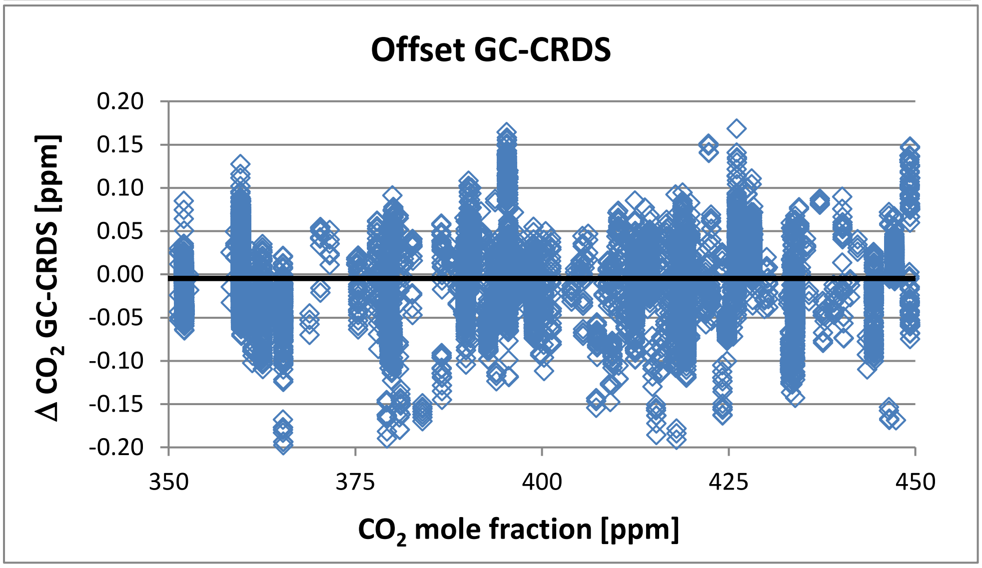


FIGURE 5: *OFFSETS OF DAILY CRDS CO_2 MEASUREMENTS RELATIVE TO AVERAGE GC RESULTS. Only analyses results made within one year are considered. The black line represents the mean offset.*

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 only serve as independent quality control check.

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 the standards were also analyzed for another time at MPI-BGC. 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 6 below.

The results of the MPI-BGC measurements of the complete FCL standard set are on average 0.001 ± 0.012 ppm lower than the CO_2 WMO X2019 PC1 assignments made by the CCL (red symbols) when considering the isotopic composition of CO_2 in the standards (see Annex IV). There is an apparent mole fraction dependency of the offset. The same analysis of FCL measurement results of the MPI-BGC standard set yields a very close match with on average 0.001 ± 0.015 ppm lower values than the CCL PC1 assignments (see Figure 6, blue symbols). Note that the two data sets in Figure 6 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 by FCL with the WMO Lab Standards of FMI (in 2016), UBA Zugspitze (in 2021) and DLR (in 2022). Whereas the agreement with the UBA and DLR sets is very good there is a small consistent offset for the FMI set (FCL CO_2 results ca. 0.04 ppm lower than CCL assignments).

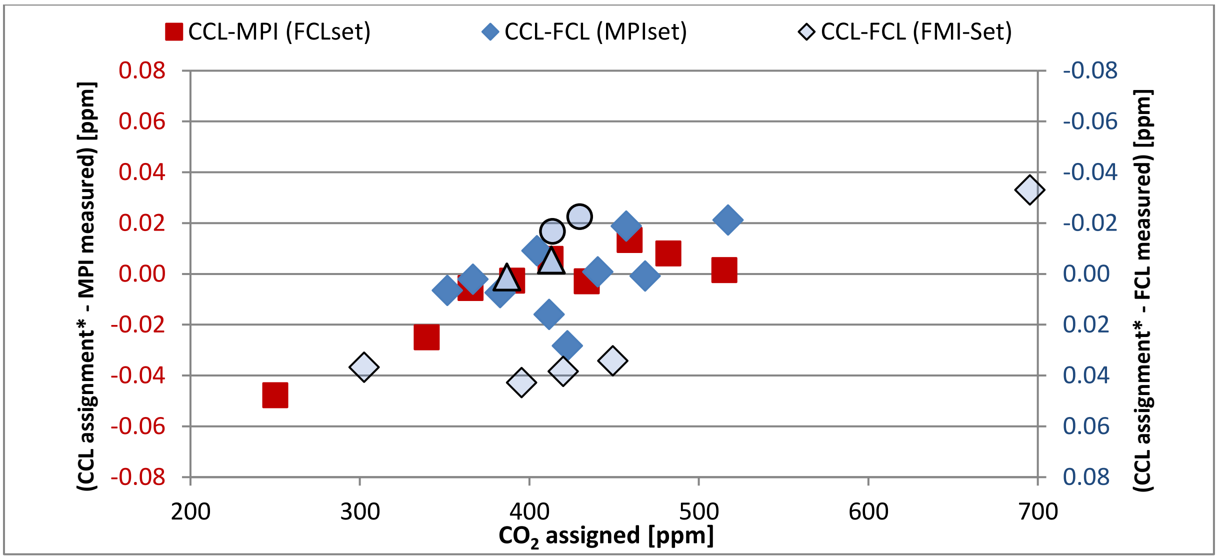
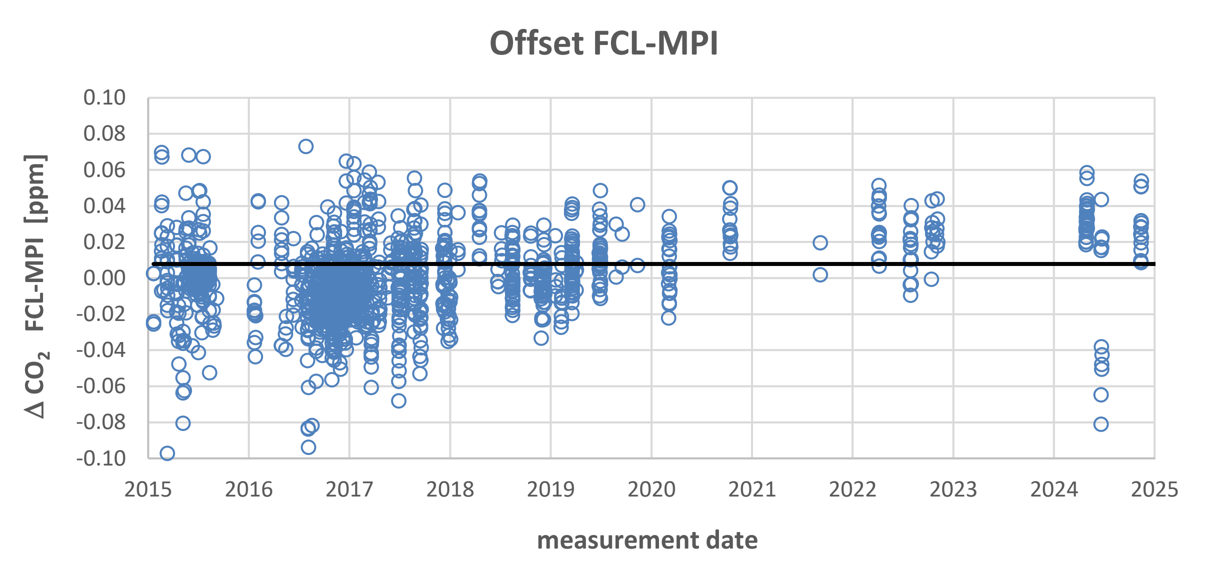


FIGURE 6: *DIFFERENCES OF FCL ANALYSIS RESULTS OF EXTERNAL WMO TERTIARY STANDARDS TO CCL CO_2 ASSIGNMENTS (BLUE DIAMONDS) AND OF MPI-BGC ANALYSIS RESULTS OF THE FCL PRIMARY STANDARDS TO CCL ASSIGNMENTS (RED SQUARES). Note that the data sets with different colours are on axis with opposite sign (see text) and that the assigned values have been adjusted for the isotopic composition of CO_2 in the respective standards (see Annex IV). *only CCL PC1 data considered*

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, however not in 2023. The difference in the results of the two labs for about 1402 daily mean results (involving 102 samples) is presented in Figure 7 below. These comprise all gases that have been analyzed within one year (only samples with CO_2 mole fractions within the calibrated ranges have been considered). There is no mean offset between FCL and MPI for the entire period Mar 2015 through Nov 2024 (-0.00 ppm ± 0.02 ppm) but a very minor mole fraction dependent difference with FCL results being larger compared to MPI results at higher mole fractions and smaller at lower mole fractions. Between end of 2020 through Nov 2024 the offset has been 0.022 ppm ± 0.024 ppm which is again very similar to the one established in the previous sections. Note that these differences include the measurement uncertainties of both laboratories and for some samples with growing CO_2 part of the difference will be result of the analysis time delay. As explained in section 5.3 measurements up to May 2015 were not yet made using the same strict procedure that has been adopted since resulting in more noise in the offset. 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.



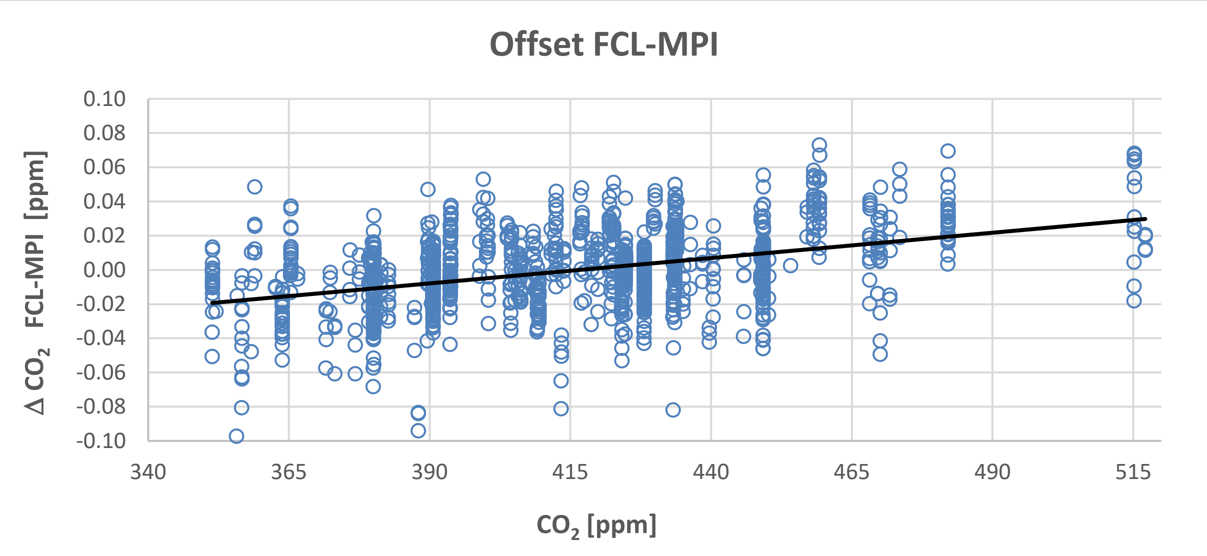


FIGURE 7: *CO2 OFFSET BETWEEN FCL AND MPI IN STANDARD MEASUREMENTS. Note that there are time lags between the analysis time in both laboratories that can cause biases for gases that are not stable in their CO_2 mole fraction over time in this graph.*

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 8. The average agreement of NOAA mean flask results compared to FCL-

CRDS filling gas data is NOAA - FCL = - 0.03 ± 0.05 ppm (filled black circles) without any clear mole fraction dependency. Some larger scatter at lower mole fractions in earlier years 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 (as WMO-CCL), EMPA (as WMO-WCC), MPI-BGC, FCL and FMI-ATC (ICOS Mobile Lab) is made on an annual basis to check the ICOS scale link to the WMO mole fraction. In this program a set of three cylinders is prepared and maintained by the FCL. One of these cylinders (D232733) constitutes a blind sample and is modified in its composition after every completed loop. A small trend of increasing 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 NOAA data record as the reference for the two comparison samples that have been used over several years. The "blind" sample is analysed at different points of time only at the FCL, therefore the CO_2 growth is determined by these measurements and the FCL trend serves as reference. In Figure 9 results of the first four iterations are shown as difference relative to the respective CO_2 trend function. The mean FCL-NOAA offsets relative to the reference trends for the low (D232717) and the blind (D232733) comparison standards have remained stable within 0.02 ppm whereas a growing offset is observed for the high comparison standard (D232721) up to 0.06 ppm since 2022. FCL data suggests an accelerated CO_2 growth caused by the low pressure in the cylinder where the NOAA trendline might not be valid any more. Yet, it might also point to a real bias. The last data points of the high and blind comparison data have not yet measurements by any lab at different points in time so a trend adjustment might have to be applied by hindsight.

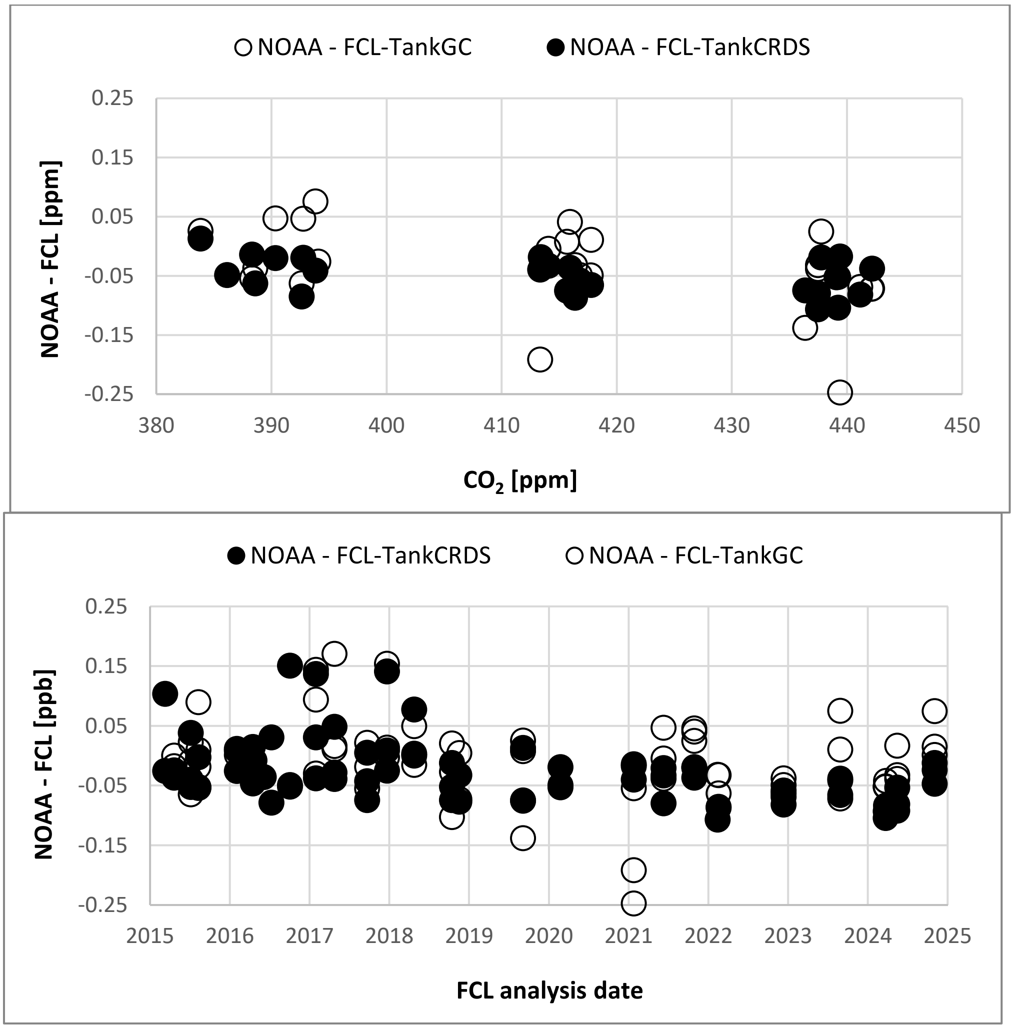


FIGURE 8: *CO2 OFFSET BETWEEN NOAA SAUSAGE FLASK DATA AND FCL DATA.*
Black dots represent fcl's analysis of the sausage fill gas (filled symbol: CRDS measurement; unfilled symbol: GC); comparisons are only considered if the flask pair agreement is < 0.3 ppm. The upper plot is based on data from 2019-2024 only.

FIGURE 9: *CO_2 OFFSET IN MENI ICP BETWEEN FCL, MPI, ICOS MOBILELAB AND WCC RELATIVE TO NOAA*

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 2020]. 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 CCL assignment record using CRDS instrumentation does not indicate a significant drift in any of the nine standard gases. The standard deviation of the mean regression fit residuals of the Primary Standards set of 0.01 ppm confirm the consistency of the used assignments.

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 from between 25 and 29 calibration episodes of the first set and of 23 calibration episodes for the second set is 0.013 ppm. The uncertainty of the assignments of the Secondary Standards is expected to be below 0.01 ppm relative to the Primary Standard set reflecting the ambiguity of the assignement record on whether CO_2 is stable or slightly growing. This is consistent with an average daily calibration standard error of 0.006 ppm.

The comparison of FCL measurement results of WMO tertiary standards of other groups (MPI, FMI, UBA, DLR) results in a mean offset of 0.005 +/- 0.02 ppm to the CCL assignments at atmospheric mole fractions. This is in agreement with the above uncertainty estimate, although an arithmetic error that has been made in the calculation to account for the differences in the isotopic calculation has not yet been corrected (see Annex IV). A preliminary assignment of the second set of Secondary Standards based on a limited number of Primary Standard calibrations had suggested a growth of CO_2 in the standards that were not confirmed by further Primary Standard calibrations. While the assignments were adjusted end of April 2022 they have not been rectified for the period before (June 2020 - April 2022). As a result CO_2 results are currently slightly too high for that period with a maximum offset in April 2022 of 0.02 - 0.03 ppm. The small offset in the Secondary Standard assignments in that period shows consistently up also in the target residuals records, as well as in all external comparisons with MPI and NOAA. An adjustment to correct for the erroneous trend assignments will be made by hindsight.

5.6.3 CO_2 long-term reproducibility

The reproducibility of CO_2 measurements as derived from the standard deviations of the monthly averaged measurement residuals of the target standards is estimated to be equal to 0.006 ppm from 2015-2024. Within the scatter of this time series there are minor 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 0.006 ppm (except for the period between June 2020 and April 2022, see section 5.6.2) 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 = 0.012 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 et al. 2021]. With three PC1 calibrations indicating a stable standard composition, the uncertainty is assumed to be equal to 0.007 ppm. This is in agreement with the consistency of the regression fit residuals of the FCL Primary Standards.
- The uncertainty of the FCL internal scale transfer to the Secondary Standards is estimated as 0.01 ppm (uncertainty of the mean mole fraction or the trend function of CO_2 over time).

2. Measurement uncertainty of daily means = 0.015 ppm for Picarro1 and 0.011 ppm for Picarro2:

- mean uncertainty of the daily calibration regression fit = 0.006 ppm
- typical uncertainty of unaccounted detector response drift throughout the validity of the daily calibration = 0.013 ppm before the change of Picarro method and less than 0.009 ppm after the change.
- approximated uncertainty for insufficient sample flushing time = 0.005 ppm before the change of Picarro method and less than 0.002 ppm after the change.
- Uncertainty from the repeatability of the daily sample measurements = 0.003 ppm before the change of Picarro method and 0.002 ppm after the change.

This uncertainty of daily means estimate is similar to the mean observed standard deviation of multiple daily means within one calibration period for individual samples (0.013±0.004 ppb, n=1160).

3. Additional long-term variability = 0.006 ppm

The long-term variability estimation is based on the reproducibility of the monthly-averaged residulas of the targets measurements on the FCL Secondary Standards against their Primary Standards calibrated estimations.

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 Picarro1 and 0.017 ppm for Picarro2.

The internal reproducibility is estimated to be 0.019 ppm for Picarro 1 and 0.016 ppm for Picarro 2 which is consistent with the results from the target standard record.

The assignment error made when accounting for the isotopic composition of CO_2 is on average 0.03 ppm in the range of 390 to 460 ppm. This term is a systematic offset so is not counted as an uncertainty and will be corrected for with the next scale link update.

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, 2017, and 2018, respectively. In 2021 and then in 2024, the complete set was recalibrated again, such that four CCL assignments from different years are available for each standard. In 2017 the CCL has changed instrumentation now using CRDS instead of GC-FID. For the tanks, the difference in mole fractions 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 -0.20 to 0.54 ppb).

TABLE 5: *CH_4 X2004A ASSIGNMENTS FOR FCL PRIMARY STANDARDS BY THE CCL [PPB]*

Sample ID	Cylinder ID	CCL date 1	CCL date 2	CCL date 3	CCL date 4	mean GC data	mean CRDS data	Assignment used *
i20140055	CB09944	Dec-13	May-17	Mar-21	Aug-24	1596.76	1596.58	**1596.64**
i20140056	CB09939	Feb-14	Oct-18	Mar-21	Aug-24	1743.13	1743.14	**1743.13**
i20140057	CB09958	Dec-13	Aug-16	Mar-21	Aug-24	1896.80	1896.88	**1896.82**
i20140058	CB09983	Dec-13	Oct-18	Mar-21	Aug-24	2032.92	2032.89	**2032.92**
i20140059	CB09952	Feb-14	Aug-16	Mar-21	Aug-24	2195.27	2195.14	**2195.34**
i20140060	CB09955	Dec-13	Jun-17	Mar-21	Aug-24	2344.02	2343.89	**2344.05**
i20140061	CB09957	Dec-13	Aug-16	Mar-21	Aug-24	2466.60	2466.69	**2466.72**
i20140062	CB09934	Jan-14	Jun-17	Mar-21	Aug-24	2731.47	2731.82	**2731.28**
i20140054	CB09948	Jan-14	Oct-18	Mar-21	Aug-24	2932.82	2932.97	**2932.82**

* Corresponds to initial CCL GC assignment

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 was no need for an update of the assigned values and the initial assignment is still used (last column in Table 5).

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 10 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.

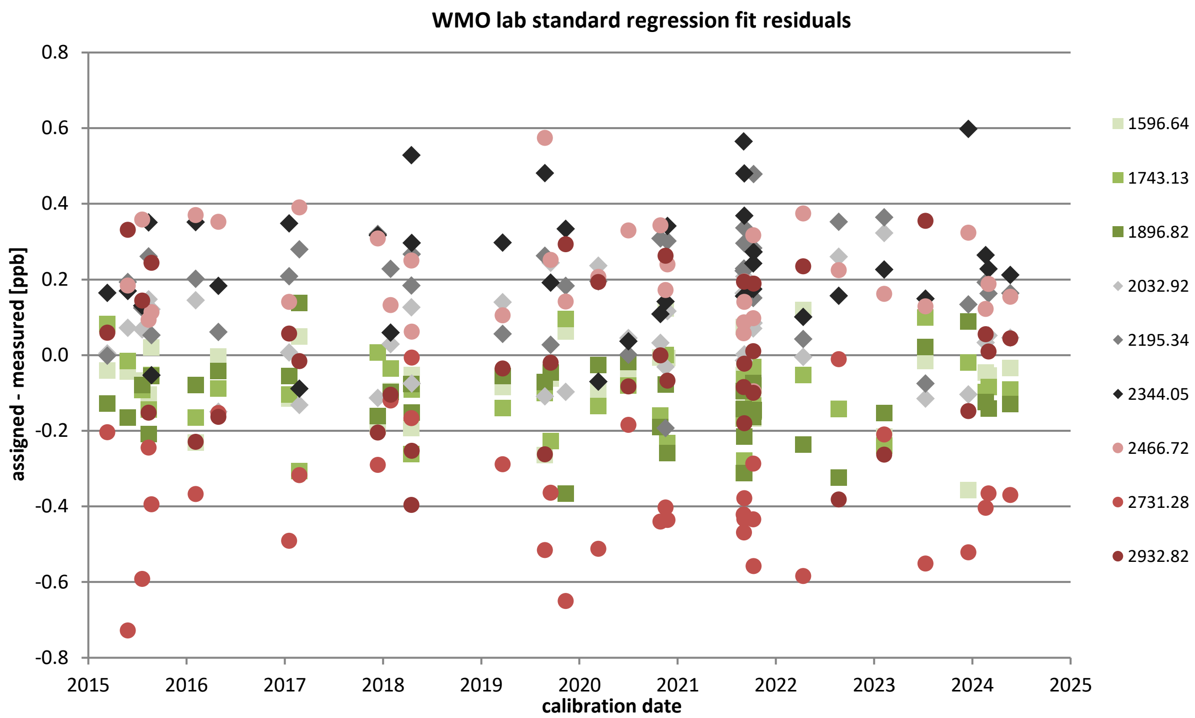


FIGURE 10: *TIME SERIES OF LINEAR REGRESSION FIT RESIDUALS OF CRDS CH_4 CALIBRATIONS FOR FCL PRIMARY STANDARDS*

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 either July2020 or Sep2021, respectively. 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 steps, 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 determined by repeated measurements against the FCL Primary Standards (n=8). The assigned values are listed for both sets in Table 6.

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.

For the initial set of Secondary Standards used until June 2020, 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 initial set for the period until their replacement in December 2020.

TABLE 6: *CH_4 ASSIGNMENTS OF FCL SECONDARY STANDARDS [PPB]*

Sample ID	Cylinder ID	Assigned Value	Re-assigned*	*Date of exchange*	Sample ID	Cylinder ID	Assigned Value
i20140171	D801336	**1795.46 ppb**	1795.56 ppb	*2020-12-08*	i20190708	D761202	**1799.01 ppb**
i20140172	D073384	**1960.24 ppb**	1960.54 ppb	*2020-12-08*	i20190803	D073381	**1949.24 ppb**
i20140173	D073392	**2288.57 ppb**		*2020-06-23*	i20190709	D761214	**2296.10 ppb**
i20140174	D801331	**2092.46 ppb**		*2020-06-23*	i20190438	D073389	**2098.41 ppb**

** Re-assignments used from 2020-06-23 to 2020-12-07*

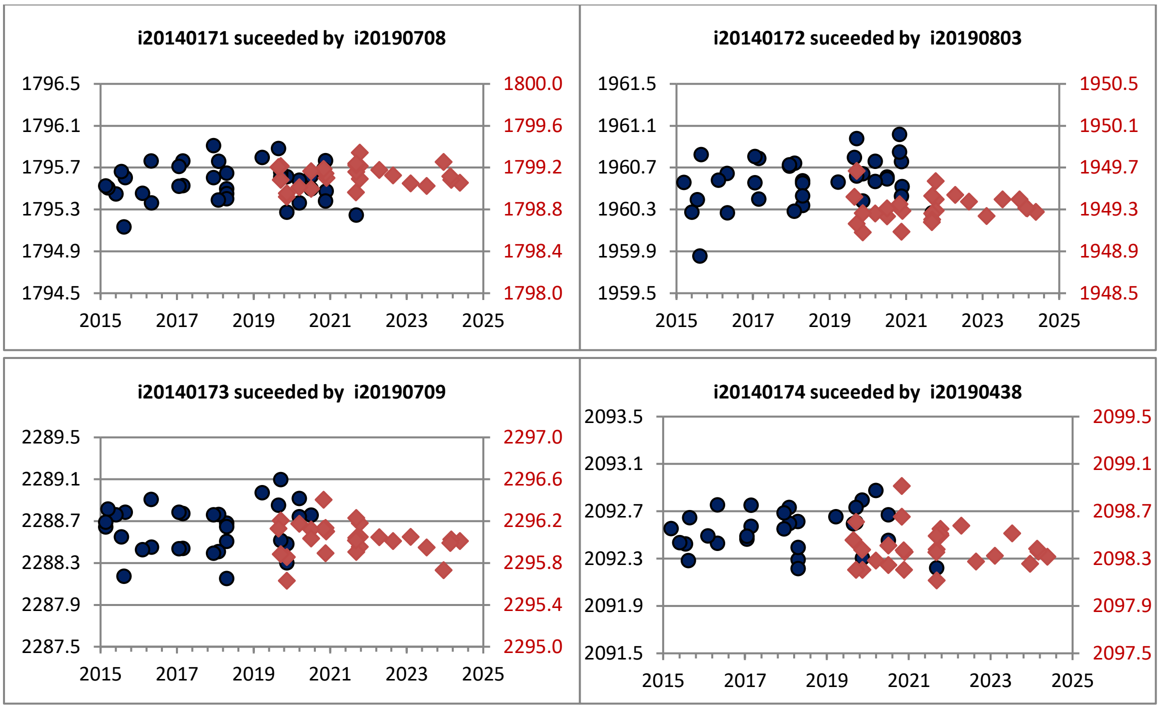


FIGURE 11: *FCL SECONDARY STANDARDS CH_4 ASSIGNMENT TIME SERIES (ALL VALUES IN [PPB])*
Dark blue dots represent the assignments for the first set of FCL Secondary Standards, the red diamonds display the four new FCL Secondary Standards.

6.2.2 Residual record

The record of the residuals of the linear regression fit of the Secondary Standard calibrations are given in Figure 12. The scatter of the residual time series for the individual standards is mostly < 0.1 ppb without any trend in the residuals being apparent. Note that the scatter of the residuals became narrower < 0.05 ppb after the exchange of the CRDS analyzer in February 2024. This documents the long-term internal consistency of the calibration set over time. The internal consistency of the new FCL Secondary Standard set expressed as the standard deviation of the mean residuals is ~0.02 ppb compared to 0.11 ppb for the first FCL Secondary

Standard set. This reflects the small bias in the initial assignments of the first set of standards but still indicates the overall very little scale transfer uncertainty.

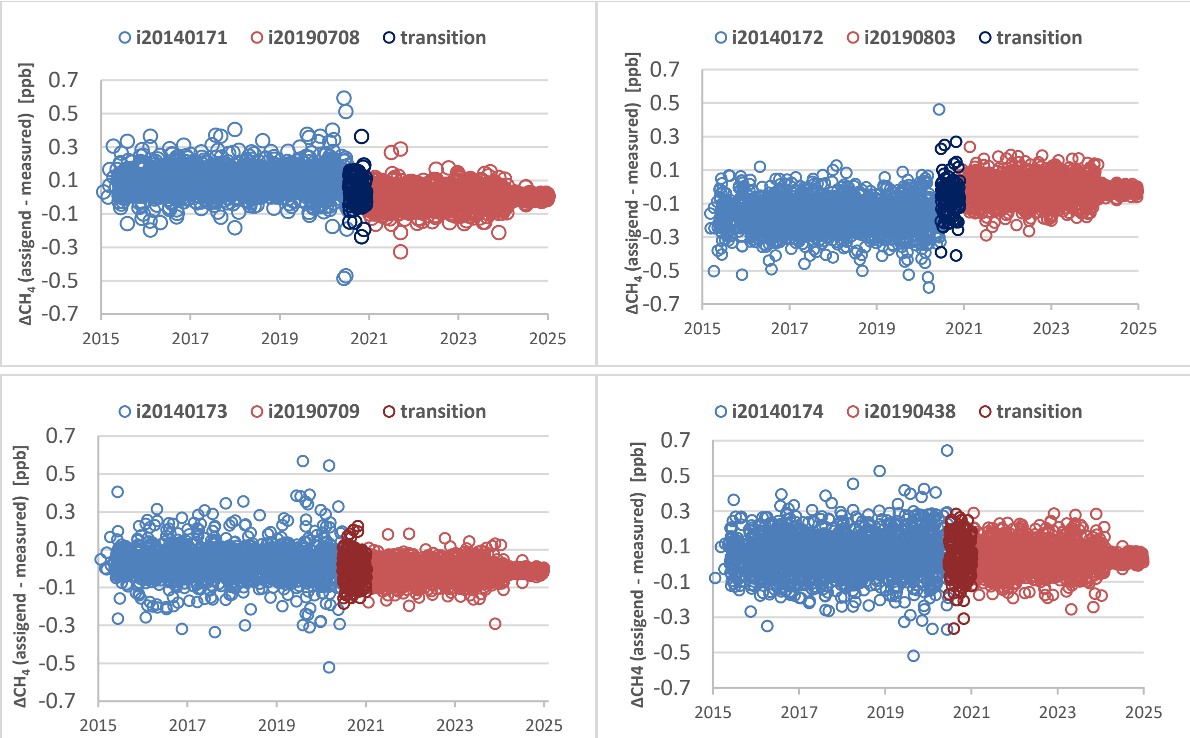


FIGURE 12: *TIME SERIES FOR CH_4 LINEAR REGRESSION FIT RESIDUALS OF THE FCL SECONDARY STANDARDS. Dark symbols indicate the transition phase when only the first part of the standards was replaced.*

6.3 CH_4 Targets

The performance of the measurements is controlled on a daily basis by analysis of two short term target standards and at lower frequency by additional long term targets. Table 6 lists the periods of use and mole fractions of those standards. Figure 13 shows the time series of the residuals of the measurement results relative to the average mole fraction based on the calibrations using all FCL Primary Standards. The consistent step of 0.2 ppb after changing the Secondary Standard calibration sets apparent in this figure complies with the small initial assignment bias of the Secondary Standards made in 2015 (see section above). Since the exchange of the Secondary Standard set, the mean residuals have decreased to within 0.06 ppb for all targets. This confirms that very little uncertainty contribution results from the scale propagation. Similarly to the secondaries, the scatter of the targets residuals became less than 0.1 ppb after the exchange of the CRDS analyzer in February 2024.

TABLE 7: *TARGET STANDARDS FOR THE CRDS CH_4 ANALYSES*

sample ID	tank ID	measured CH_4 [ppb]*	std.dev. [ppb]*	Primary Calibration CH_4 mean [ppb]	std.dev. Calibration mean [ppb]	n calibration values	Period of use
i20150062	D073391	1914.71	0.17	1914.92	0.19	21	05.2015-04.2019
i20150061	D073389	2043.05	0.19	2043.25	0.18	24	05.2015-04.2019
i20150060	D073381	1947.18	0.18	1947.36	0.18	63	05.2015-04.2019
i20170961	D761211	**1943.21** *1943.40*	**0.15** *0.19*	1943.43	0.14	49	11.2017-04.2023
i20170962	D801332	**2032.71** *2032.92*	**0.17** *0.19*	2032.94	0.17	34	11.2017-present
i20190451	D073391	**2085.97** *2086.19*	**0.19** *0.19*	2086.25	0.18	28	06.2019-07.2024
i20150188	D073398	**1595.54** *1595.73*	**0.19** *0.16*	1595.76	0.14	28	05.2015-present
i20150374	CA05755	**1703.32** *1703.52*	**0.16** *0.16*	1703.57	0.16	32	08.2015-present
i20222329	D994882	1897.37	0.14	1897.40	0.09	20	09.2022-present
i20222170	D487652	2067.68	0.14	2067.71	0.01	6	10.2022-present

**For Targets i20170961, i20170962, i20190451, i20150188 and i20150374 the mean values for the period from start until 23.06.2020 (change of the FCL Secondary Standards) are displayed in bold, for the period since then in italics*

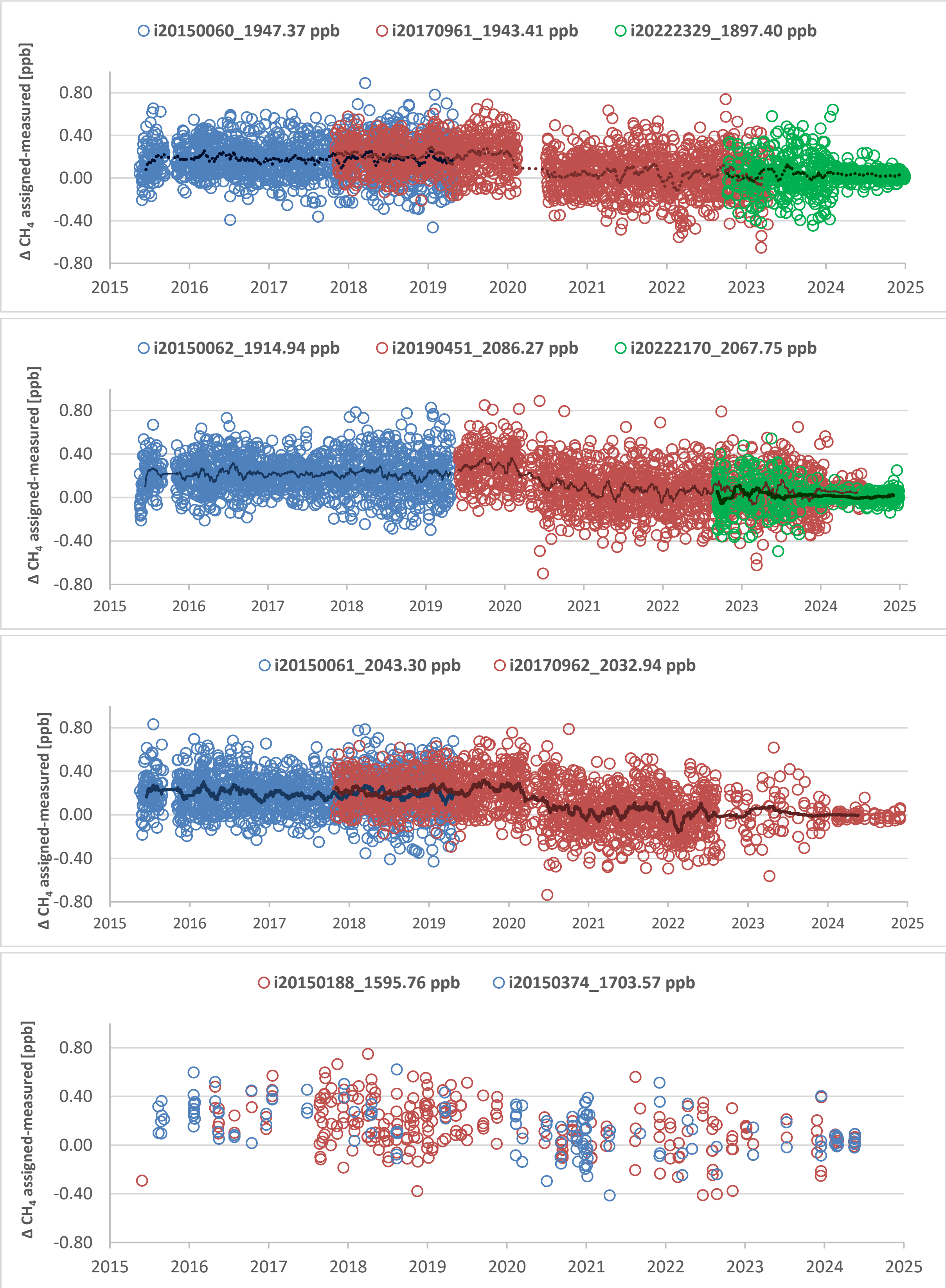


FIGURE 13: *TIME SERIES OF THE CH_4 OFFSET OF TARGET MEASUREMENTS TO THEIR RESPECTIVE ASSIGNED VALUES. The dark line represents a 30 points-running mean.*

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 FCL Primary 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 ten injections. The inter-instrumental measurement differences for all samples are depicted in Figure 14 (only standards within the range defined by the calibration standards were considered, n=260). The average offset is 0.13 ppb ± 0.30 ppb for the initial phase until the change of the FCL Secondary Standards on 23rd June 2020, from that date onwards about -0.06 ppb ± 0.28 ppb, which again reflects the small bias of the initial CRDS Secondary Standard assignments.

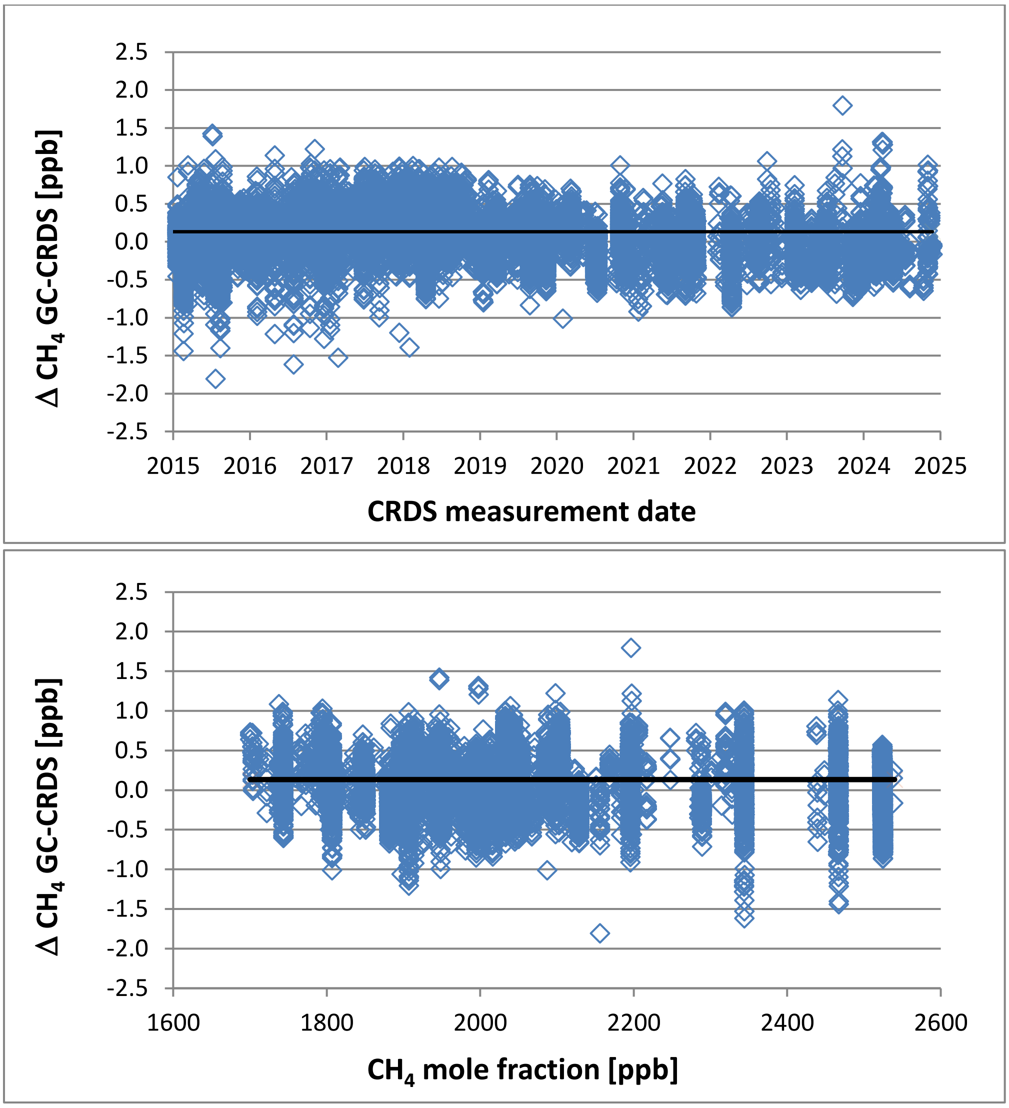


FIGURE 14: *OFFSET OF CRDS DAILY MEAN CH_4 RESULTS RELATIVE TO GC AVERAGE RESULTS OF THE SAME SAMPLE*

Note that each data point in Figure 14 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. Comparison with additional sets of WMO tertiary standards could be made by FCL with the WMO standard sets of FMI (in 2016), UBA Zugspitze (in 2021) and DLR (in 2022). These data are shown in Figure 15.

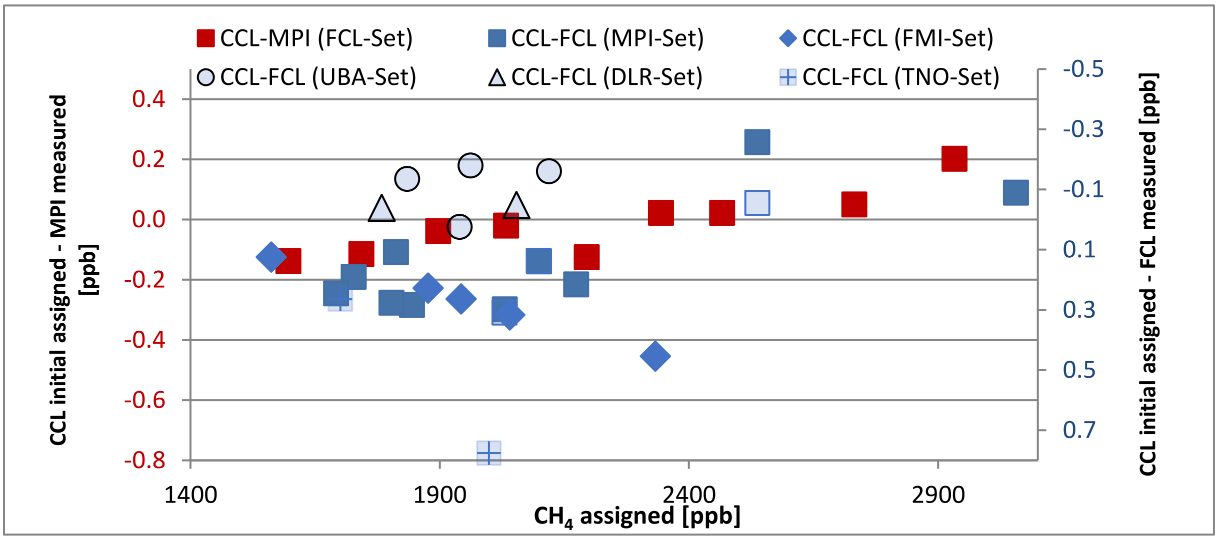


FIGURE 15: *DIFFERENCES OF FCL ANALYSIS RESULTS OF EXTERNAL WMO TERTIARY STANDARDS TO CCL CH_4 ASSIGNMENTS (BLUE DIAMONDS) AND OF MPI-BGC ANALYSIS RESULTS OF THE FCL PRIMARY STANDARDS TO CCL ASSIGNMENTS (RED SQUARES). Note that the data sets with different colours are on axis with opposite sign (see text)*

The mean difference of the measurement results CCL - FCL has been 0.2 ppb before June 2020 (dark blue symbols in Fig. 15) and 0.1 ppb in 2021-2024 (light blue symbols in Fig. 15). Similarly, the mean CCL - MPI-BGC difference of is <0.02 ppb. The differences for the individual standards closely follow the regression fit residuals observed (see section above). 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 16, blue symbols) are very consistent to the difference of FCL measurement results of the MPI-BGC Primary CH_4 Standards.

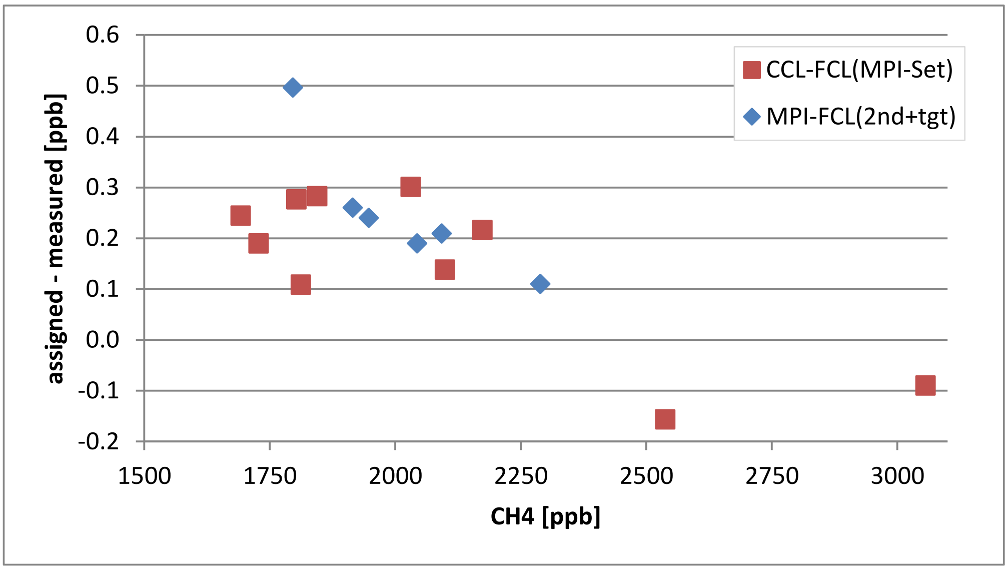


FIGURE 16: *DIFFERENCES OF MPI-BGC MEASURED RESULTS TO FCL SECONDARY STANDARD ASSIGNED CH_4 VALUES (BLUE DIAMONDS) compared to the differences of FCL measured results relative to CCL CH_4 assignments of MPI-BGC Primary Standards (red squares)*

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 8. As seen in Fig. 16 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 8: *CH_4 COMPARISON OF MPI-BGC ANALYSIS RESULTS AND FCL FOR TARGET STANDARDS*

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 by CRDS in both laboratories in former years (no comparison was added in 2023). The difference in results for about 100 compared samples is presented in Figure 17. The average offset of all MPI-FCL sample comparisons that were measured using the first Secondary Standard set at FCL amounted to 0.25 ppb ± 0.20 ppb, since the time the second Secondary Standard set is in use the mean offset is 0.04 ppb ± 0.21 ppb.

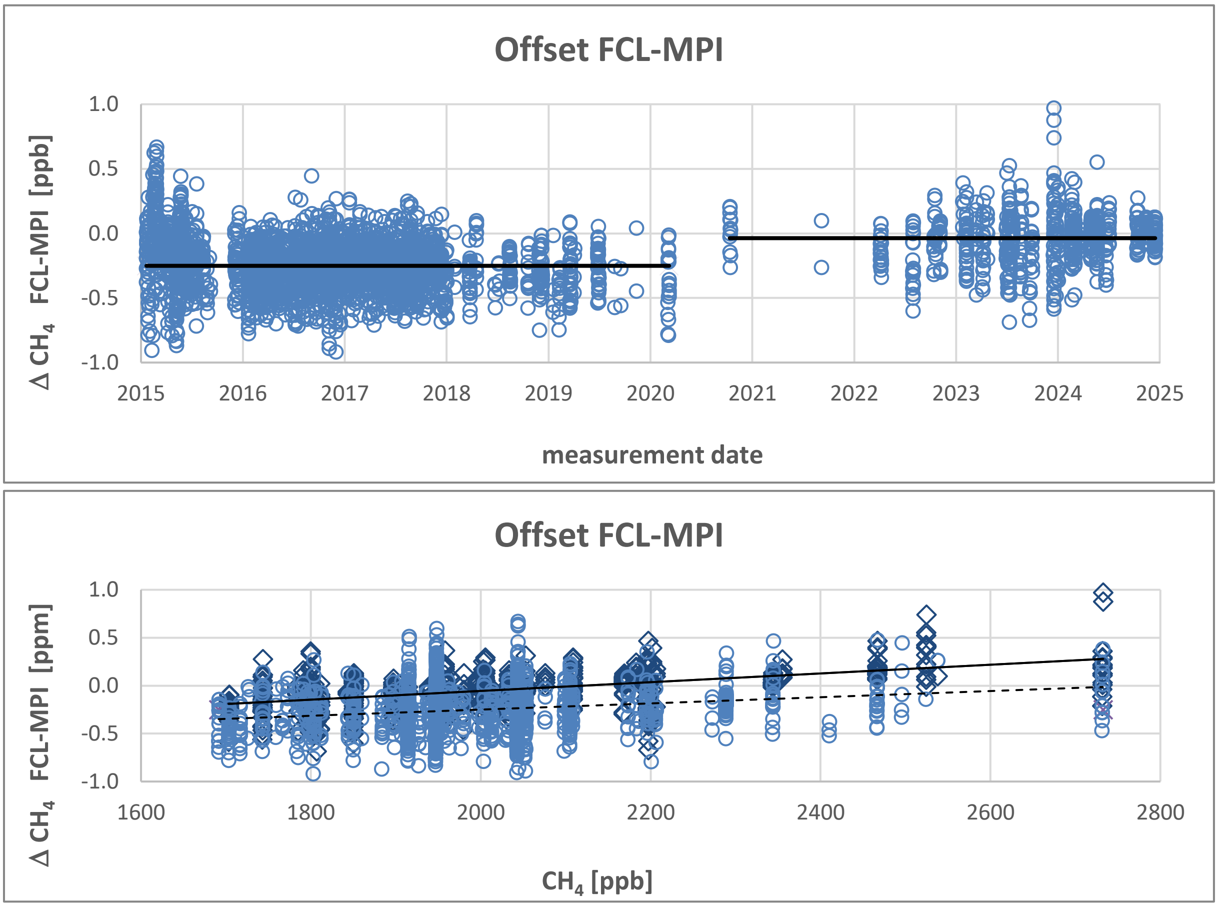


FIGURE 17: *CH_4 OFFSET BETWEEN FCL AND MPI IN STANDARD MEASUREMENTS (DARK BLUE DIAMONDS REPRESENT COMPARISON RESULTS BASED ON THE SECOND SECONDARY STANDARDS SET; LIGHT BLUE DIAMONDS REPRESENT COMPARISON RESULTS BASED ON THE FIRST SECONDARY STANDARDS SET)*

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 = 0.4 ppb ± 0.5 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 = 0.1 ppb ± 0.2 ppb.

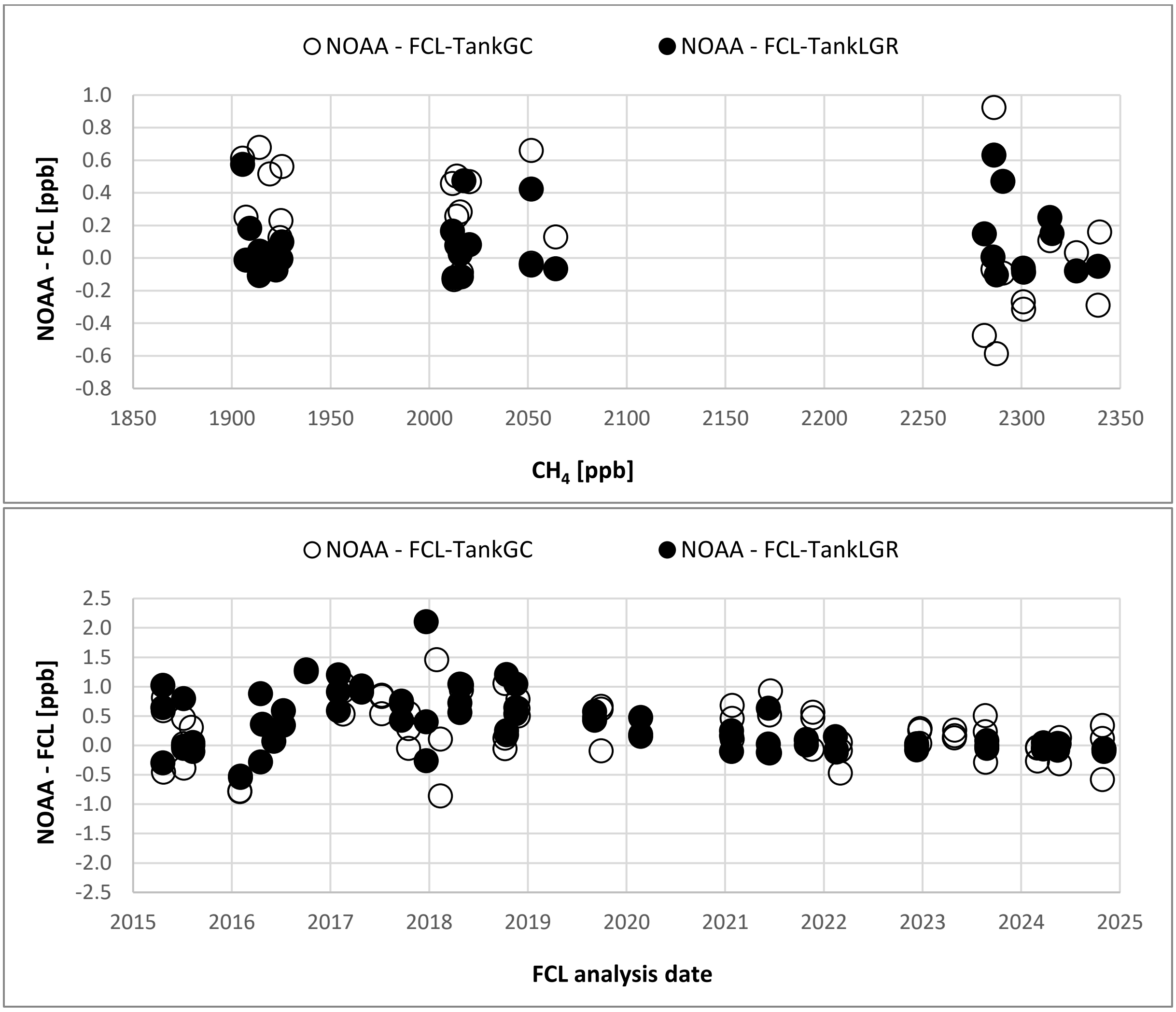


FIGURE 18: *CH_4 OFFSET BETWEEN NOAA SAUSAGE FLASK DATA AND FCL DATA. Black dots represent FCL's analysis of the sausage fill gas (filled symbol: CRDS measurement; unfilled symbol: GC); the upper plot is based on data from 2019-2024 only.*

The MENI round robin between NOAA (as WMO-CCL), EMPA (as WMO-WCC), MPI-BGC, FCL and FMI-ATC (ICOS Mobile Lab) is made on an annual basis to check the ICOS scale link to the WMO mole fraction. 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 19 results of the first four circulations are shown. The total observed offset between FCL-CCL is Δ = - 0.1 ± 0.2 ppb, for the FCL analysis period up to June 2020 (using the first set of Secondary Standards) Δ = - 0.2 ppb, since then no mean offset has remained (Δ = ± 0.0 ppb).

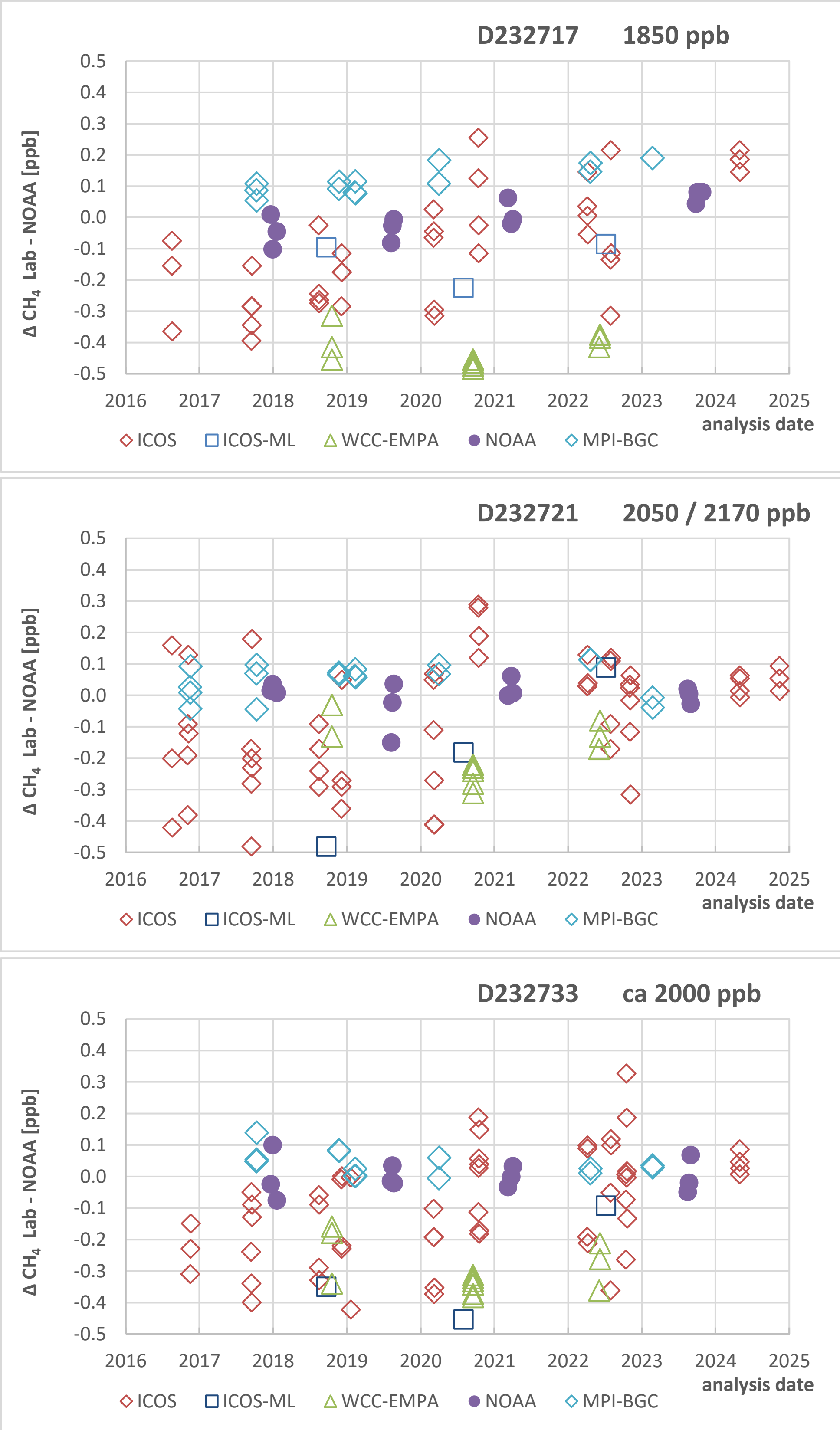


FIGURE 19: *CH_4 OFFSET IN MENI ICP BETWEEN FCL, MPI, ICOS MOBILELAB AND WCC RELATIVE TO NOAA*

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 available evidence with all metrics (re-calibration by the CCL, repeated analysis by the MPI-BGC, consistency of regression fit residuals) CH_4 mole fractions within the FCL Primary Standards are accurately assigned and stable 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 standard deviation of these assignments of individual Secondary Standards is approximately 0.2 ppb. The uncertainty of the scale transfer depends on the number of calibration events. The initial assignments in 2015 have been based on only four 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. The assigned values of the current secondary standards set are based on minimum 11 calibration events. We consider to update the initial assignments of the first set in 2025, as all gases of this Secondary Standard set have received their final calibration.

6.6.3 CH_4 long-term reproducibility

Within the scatter of the 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. The related uncertainty is approximated by the standard deviations of monthly averaged CH4 measurement residuals of the target standards resulting in 0.07 ppb from 2015-2024.

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 = 0.5 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 is 0.05 ppb (second Secondary Standard set). The assignments of the first set of secondary standards are based on four calibration events within the period of February to August 2015. The reproducibility of the four assignments within this period suggests an uncertainty of 0.09 ppb. As stated in the above section 6.6.2, a larger

bias of 0.2 ppb was established by various quality control measures. The reason has not been fully unerstood.

2. Measurement uncertainty of daily means = 0.24 ppb (based on Picarro1 data) and 0.08 ppb for Picarro2.

- mean uncertainty of the daily calibration regression fit = 0.13 ppb for Picarro1 and 0.03 ppb for Picarro2.
- uncertainty of the detector response drift throughout the validity of a daily calibration = 0.2 ppb for Picarro1 and 0.07 ppb for Picarro2.
- uncertainty from the repeatability of the daily sample measurements = 0.03 ppb (for 15 min means) and 0.02 ppb (for 20 min means) after the change of the Picarro (in Feb. 2024).

This uncertainty estimate of daily means is similar to the mean observed standard deviation of multiple daily means within one calibration period for individual samples (0.17±0.06 ppb, n=1195).

3. Additional long-term variability = 0.07 ppb

The long-term variability estimation is based on the reproducibility of the monthly-averaged residulas of the targets measurements on the FCL Secondary Standards against their Primary Standards calibrated estimations.

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 for Picarro1 and 0.5 ppb for Picarro2.

The internal reproducibility is estimated to be 0.3 ppb for Picarro 1 and 0.14 ppb for Picarro 2 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 2021 and 2024, the *complete* set of FCL Primary Standards received the third and the fourth calibrations by the CCL, respectively.

The CCL calibration record of the FCL Primary Standards is summarized in Table 9 indicating that the increase in CO exhibits a mole fraction dependency with standards with lower CO mole fraction having a larger increase in CO. Further investigations at the CCL have revealed that the CO growth assigned to the WMO primary standards was inaccurate and required an update [Crotwell 2024]. This will cancel out the mole fraction dependency and reduce the magnitude of the assigned CO growth in FCL Primary Standards [Crotwell 2025]. 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 the V3 and for the first time the AR3 Aerodyne QC-TILDAS instruments were used for all standard gases. Figure 20 shows that mostly 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 obtain consistency with the WMO X2014A scale a linear interpolation between the initial and the second calibration data points was applied for every standard where the increase exceeded the analytical uncertainty of the CCL calibrations to account for the increasing CO mole fractions in FCL Primary Standards. This includes all standards with CO below 250 ppb.

A further refinement of the CO growth in the Primary Standards would be possible with the further CCL assignments. 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. It had appeared that this effort was not justified before obtaining the results of the announced WMO scale revision. As Fig. 20 shows, the FCL measurement results of the Primary Standards (grey squares) were 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 (blue dots) and Aerolaser (red crosses). However, after 2021 the FCL measurements of the Primary Standards started to deviate from the aforementioned trend, and show an ever growing offset with the trend line defined by all the four CCL calibration episodes. The link of the FCL CO scale implementation, will use all available CCL data, once the WMO CO scale revision has been published by the CCL, which is scheduled for 2025.

TABLE 9: *CO X2014A ASSIGNMENTS FOR FCL PRIMARY STANDARDS BY THE CCL [PPB]*

Sample ID	Cylinder ID	CCL date 1	CCL date 2	CCL date 3	CCL date 4	CO-date1	CO-date 2	CO-date 3	CO-date 4	drift [ppb/yr.]*	Assignment used**
i20140055	CB09944	Dec-13	Jun-17	Jun-21	Sep-24	31.31	34.41	36.92	39.13	**0.768**	32.20
i20140056	CB09939	Feb-14	Oct-18	Jun-21	Sep-24	80.14	82.73	84.27	86.43	**0.582**	80.62
i20140057	CB09958	Dec-13	Sep-16	Mar-21	Oct-24	120.69	122.36	124.70	126.91	**0.547**	121.31
i20140058	CB09983	Dec-13	Oct-18	Jun-21	Sep-24	158.92	161.28	162.19	164.24	**0.450**	159.47
i20140059	CB09952	Feb-14	Sep-16	Jun-21	Oct-24	199.47	200.77	202.69	204.63	**0.448**	199.92
i20140060	CB09955	Dec-13	Jul-17	Jun-21	Sep-24	247.14	247.88	249.48	250.77	**0.331**	247.37
i20140061	CB09957	Dec-13	Sep-16	Jun-21	Oct-24	397.06	396.19	398.75	399.47	**0**	397.90
i20140062	CB09934	Jan-14	Jun-17	Jun-21	Oct-24	697.56	697.72	697.07	696.95	**0**	697.30
i20140054	CB09948	Jan-14	Oct-18	Jun-21	Sep-24	998.63	1002.38	999.21	1000.76	**0**	999.05

* Drift calculated based on period CCL date3 – CCL date1; ** On 1/1/2015, calculated based on the corresponding drift

FIGURE 21: *CO PRIMARY STANDARDS, MEASURED AT THE FCL (GREY SQUARES) AND AT THE CCL. CCL analysers: LGR (blue dots), Aerolaser (red crosses) and Aerodyne (green diamonds)*

7.1.2 Regression fit residuals of FCL Primary CO Standards

The time series of the regression fit residuals displayed in Figure 21 shows 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 due to the fact that only two CCL calibration results have been applied to assign the trend lines. On the one hand, as mentioned before the WMO Mole Fraction scale Primary Standards the assigned X2014A trend functions were inaccurate and will be revised in 2025 (https://www.esrl.noaa.gov/gmd/ccl/co_scale_update.html). As a result CO growth is currently overcorrected at low CO mole fractions, and is underestimated for the high CO standards.

FIGURE 22: *TIME SERIES FOR CO QUADRATIC REGRESSION FIT RESIDUALS OF LGR FCL PRIMARY STANDARD CALIBRATIONS*

All data presented in the above figure refers to measurements by LGR 1, with exception to data of 08.2022 and one of the two points per standard displayed in 03.2023 (refer to section 8.4.1)

7.2 FCL Secondary CO Standards

7.2.1 Assignment record

The reference gases that are used as FCL Secondary Standards for the daily LGR calibration have been analyzed together with the complete set of FCL Primary Standards between May 2016 and December 2024 for up to 27 times. In May 2018 one of the Secondary Standard gases was exhausted (i20150251) and was succeeded by a new standard with a similar CO mole fraction (i20170889). 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. In January 2024, standard i20170889 was replaced by another standard i20201317 with a lower CO concentration (to extend the calibrated range at the lower end). The latter was assigned on the basis of its calibrations against the nine FCL Primary Standards (n=6).

At the upcoming WMO scale revision the number of primary calibration events will be sufficient to base the assignments for the current set of Secondary Standards, on those measurements. Table 10 summarizes the initial assignment values and the CO growth rates currently used for all tanks.

In Figure 22 the record of the Secondary Standards' CO mole fraction results of these FCL primary calibration episodes is presented. It appears that the two high standards are relatively stable in CO in the currently used set of Secondary Standards on the X2014A scale.

Table 10: *CO ASSIGNMENTS FOR FCL SECONDARY STANDARDS [PPB]*

Role in the scale link	Sample ID	Cylinder ID	Date of start of use	Date of end of use	Assigned Value (ppb)	drift/yr *	Basis of the assignment	n
CAL1	i20150251	CA05640	25-11-2015	2018-05-03	78.52	+0.76	Primary Stds.	4
CAL2	i20150189	D073397	25-11-2015	26-07-2021	150.79	+0.97	Primary Stds.	6
CAL3	i20150544	D073396	25-11-2015	26-07-2021	305.80	+0.72	Primary Stds.	6
CAL4	i20150191	D073395	25-11-2015	26-07-2021	433.52	+0.51	Primary Stds.	6
CAL1	i20170889	D557226	2017-10-01	11-01-2024	79.63	+1.28	Prim. and Second. Stds.	2 and 95, respectively
CAL2	i20201308	D753834	26-07-2021	NA	149.53	+0.25	Secondary Stds.	109
CAL3	i20201254	D753835	26-07-2021	NA	293.54	-0.16	Secondary Stds.	109
CAL4	i20201255	D753836	26-07-2021	NA	423.90	-0.07	Secondary Stds.	111
CAL1	i20201317	D753833	20-12-2023	NA	55.02	+0.86	Primary Stds.	6

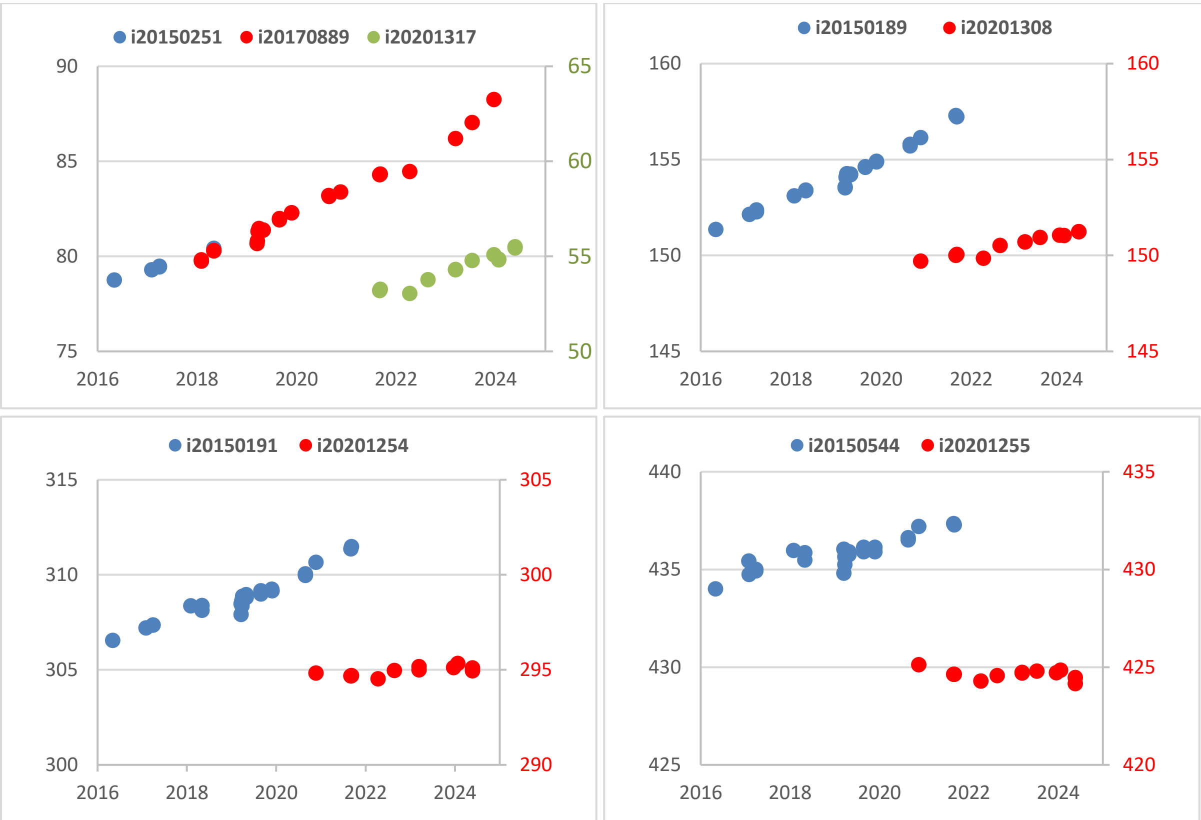


FIGURE 23: *SECONDARY CO STANDARD ASSIGNMENT RECORD (VALUES IN [PPB]).* Blue points represent the assignments for the first set of FCL secondary standards, the red points represent those for the second set of FCL secondary standards and the green points represent the current lower CO standard (replaced since 11.01.2024)

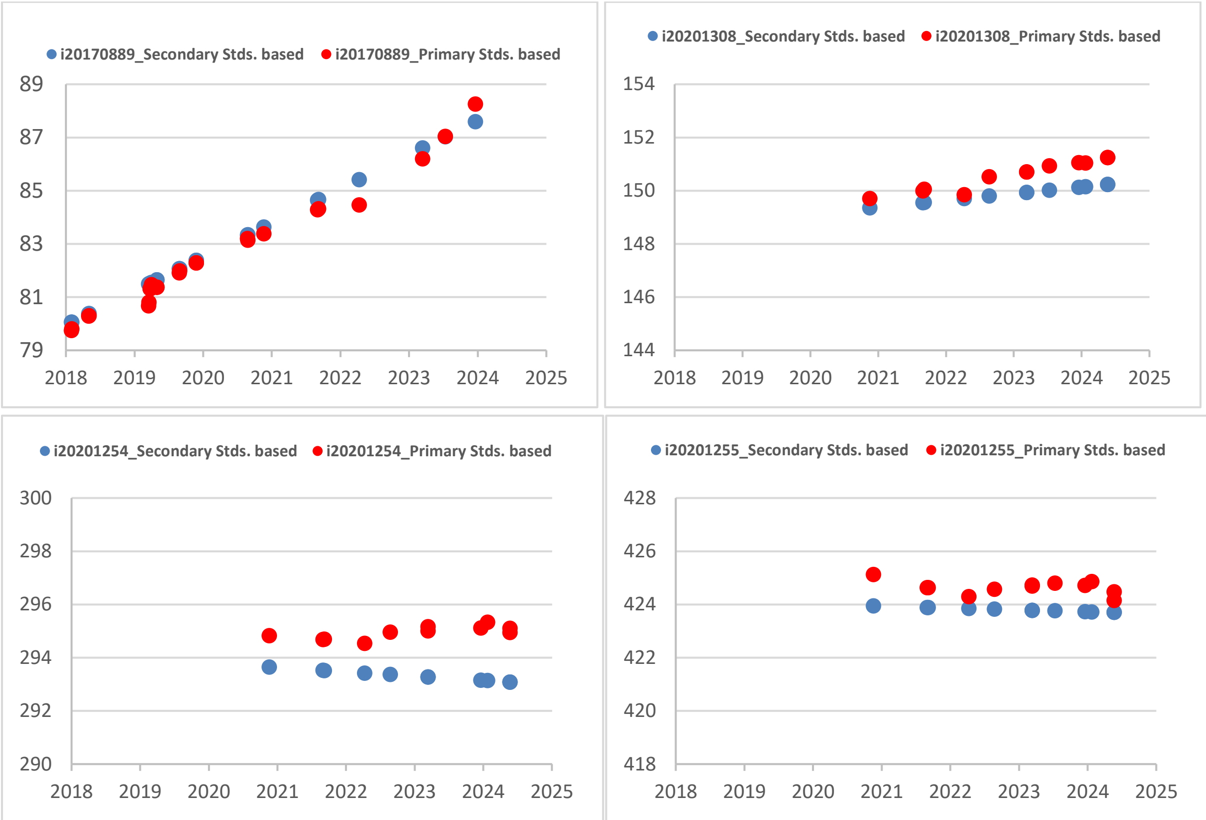


FIGURE 24: *SECOND SET OF SECONDARY CO STANDARDS: ASSIGNMENT RECORD AGAINST FCL'S PRIMARIES (RED DOTS) AND CO MOLE FRACTIONS BASED ON ASSIGNED TREND FUNCTIONS RELATING TO FCL'S FIRST SET OF SECONDARIES (BLUE DOTS) (VALUES IN [PPB]).*

Figure 23 shows the records of the second set of the FCL Seconday Standards' CO mole fraction results from the primary calibration episodes on the one hand (red dots, same values as in Figure 22), and their currently used drifting assignments as defined when they were put into use on the other hand (blue dots, based on growth rates listed in Table 10). It appears that results from these two calculation methods agree only for the low CO mole fraction Secondary Standard, but show a persisting offset, either constant (for the high CO tank) or divergent for the remaining two. This can be explained by the combined effect of the limited accuracy of the initial assignments of the second Secondary Standards set, and on the limited accuracy of the applied trend functions for CO in the Primaries on the other hand (refer back to 7.1.1).

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 up to end of 2022. 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.

From beginning of 2023, the residuals started to drift significantly resulting in values of up to 0.3 ppb. In January 2024, the Seconday Standard with the lowest CO mole fraction needed to be replaced because of its low pressure.

Since the trend in the residuals has disappeared for all of the Secondary Standards since then, it is most likely that the assigned growth rate for the replaced Secondary Standard had not been correctly characterized in its past year of operation.

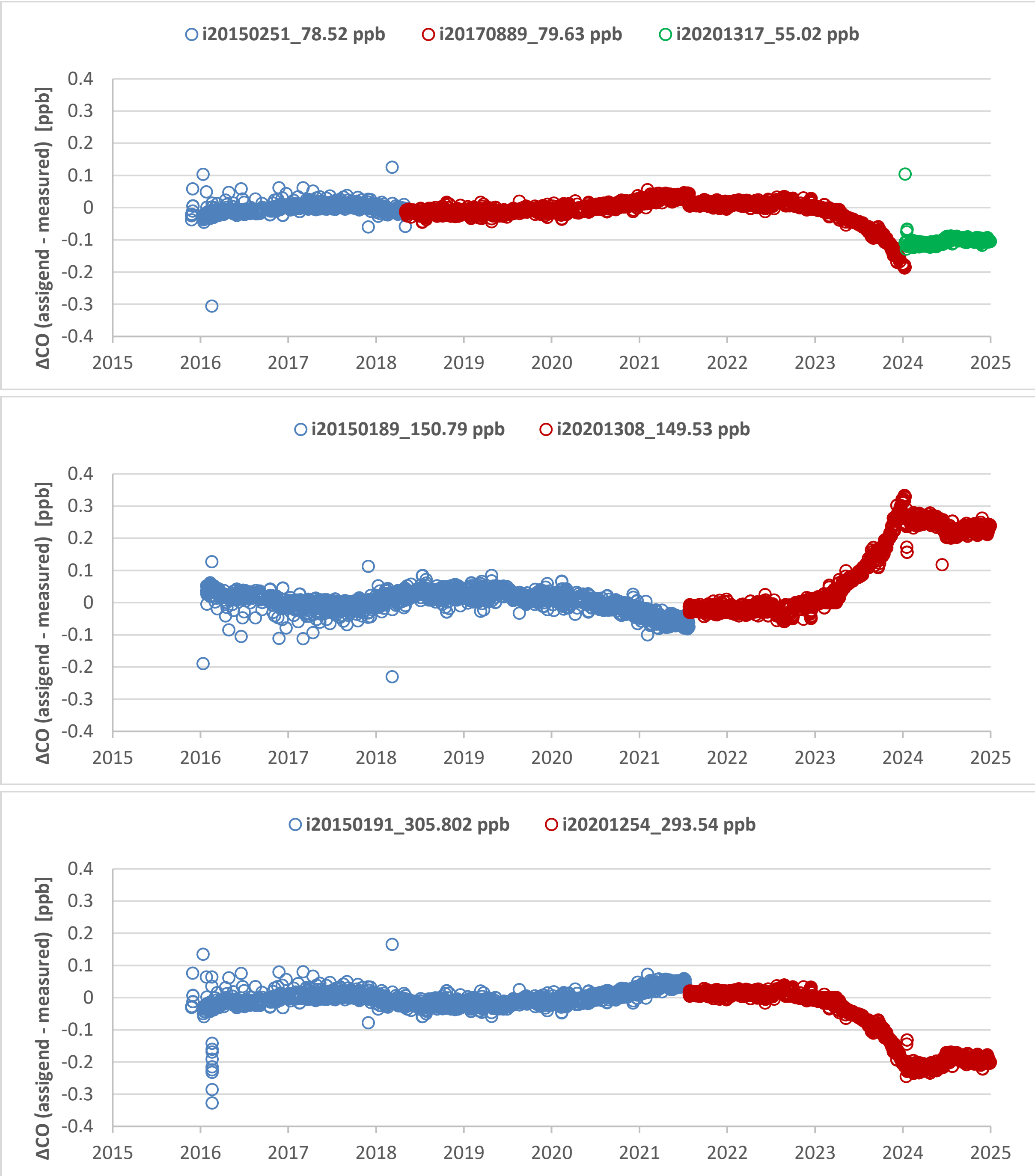


FIGURE 25: *TIME SERIES OF QUADRATIC REGRESSION FIT RESIDUALS OF LGR CO FCL SECONDARY STANDARDS CALIBRATIONS BASED ON DRIFTING ASSIGNMENTS RESULTING FROM FCL PRIMARY STANDARD ASSIGNMENTS*

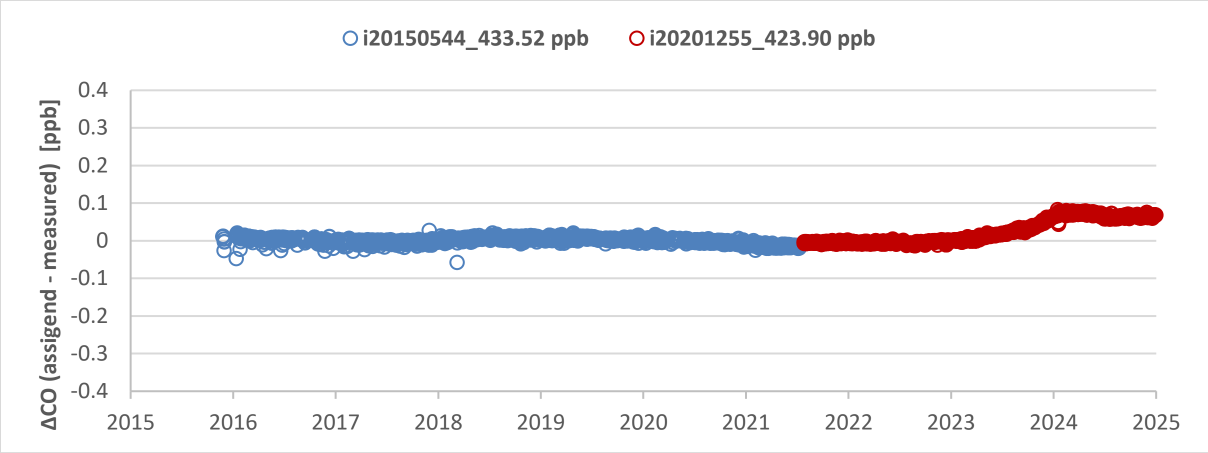


FIGURE 24: *TIME SERIES OF QUADRATIC REGRESSION FIT RESIDUALS OF LGR CO FCL SECONDARY STANDARDS. CALIBRATIONS BASED ON DRIFTING ASSIGNMENTS RESULTING FROM FCL PRIMARY STANDARD ASSIGNMENTS*

7.3 CO Targets

In the period from Feb 2016 to December 2024 two Short Term Targets have been in use on the LGR system. When the old targets got empty, they were succeeded by two new targets. The Short Term Targets are complemented by three Long Term Targets. These are currently being measured less frequently after an initial phase of daily analysis frequency to maintain a long-term link of succeeding targets in future. The time series of CO in the target standards exhibits a noticeable change, more remarkable from 2023. The record of the residuals of daily measurement results relative to the assigned trend based on Primary Standard calibrations is presented in Figure 25. There are trends and offsets apparent in the residuals of most of these targets of up to 2 ppb. These are probably result of the limited accuracy of the drift assignment for the standards of the FCL Primary and Secondary Standard sets involved as explained above.

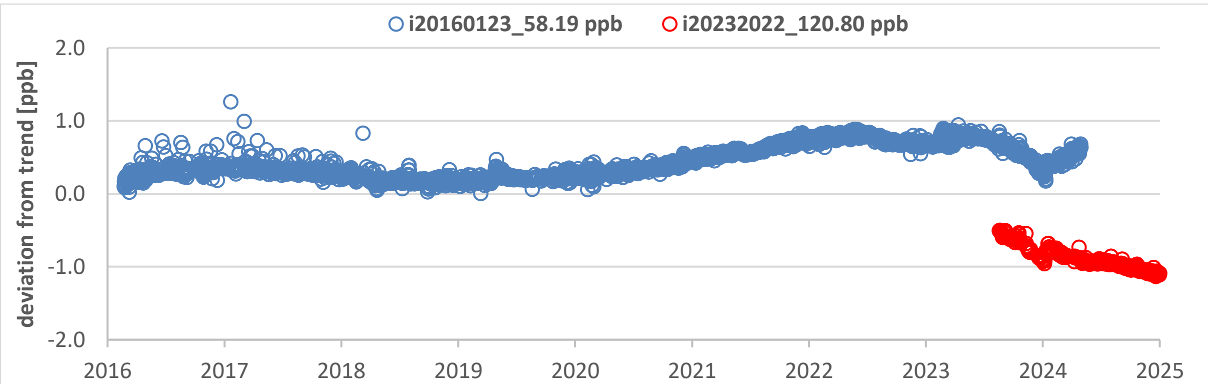


FIGURE 26: *TIME SERIES OF DAILY CO RESIDUALS DATA OF TARGET STANDARDS ANALYZED ON THE LGR INSTRUMENT*

FIGURE 25: *TIME SERIES OF DAILY CO RESIDUALS DATA OF TARGET STANDARDS ANALYZED ON THE LGR INSTRUMENT*

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 = -0.5 ± 0.8 ppb). There is a small mole fraction dependency in the offset between the instruments. It has changed only a little over time but offsets in 2024 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.

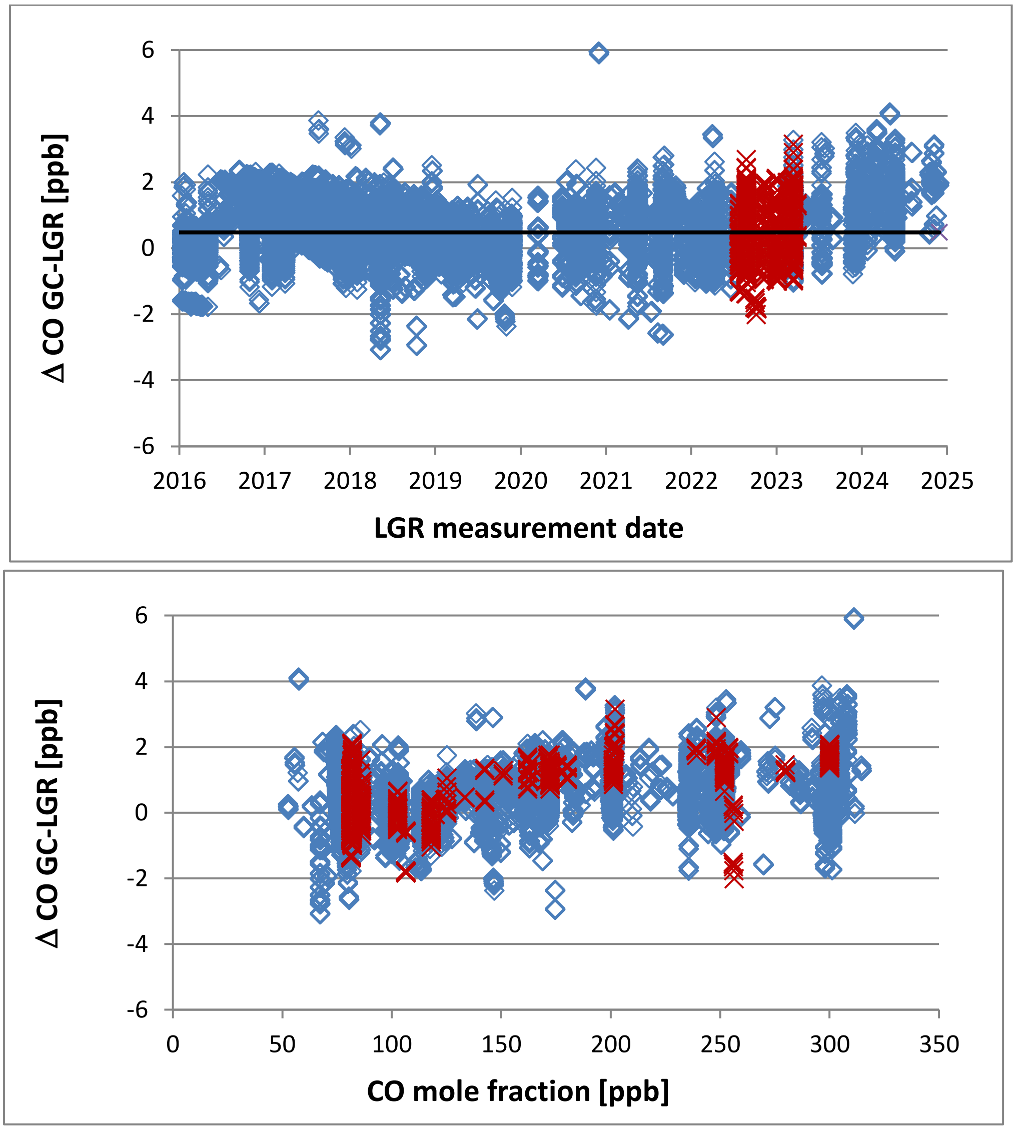


FIGURE 27: *TIME SERIES OF LGR-GC CO DIFFERENCES OF MEASUREMENT RESULTS OF THE SAME SAMPLES. RED CROSSES SHOW DATA FROM LGR 2 (REFER TO SECTION 8.4.1)*

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 X2014A Mole Fraction scale 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.

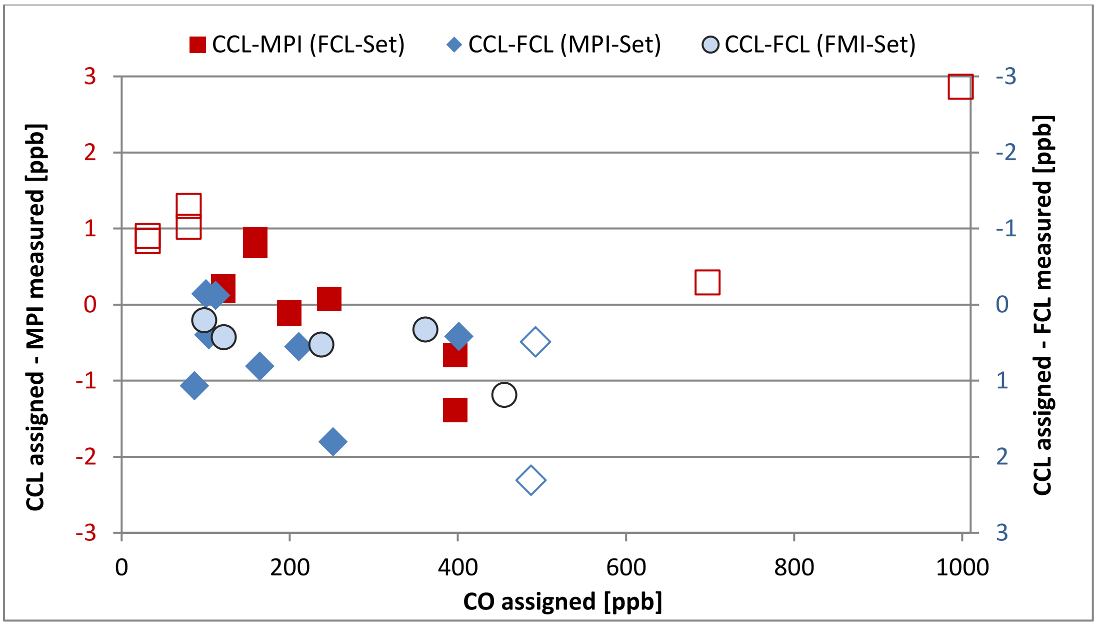


FIGURE 28: *DIFFERENCE OF MEASURED CO DATA TO CCL ASSIGNMENT OF THE WMO STANDARDS of partner labs, Standards are considered only where multiple CCL calibrations allow to characterize the CO growth rate or where the FCL and CCL measurements were performed within 6 months; unfilled symbols indicate mole fractions 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 based on 68 sample measurements using the VURF instrument is presented in Figure 28 and Figure 29. There is no average offset between FCL and MPI with Δ = -0.0 ppb ± 0.7 ppb. The difference exhibits a clear mole fraction dependence. This is in accordance with the different patterns of mole fraction dependent offsets to the CCL shown in Figure 27. It is also result of the different calibration approaches: for the LGR a multi-point quadratic fit follows the primary scale more closely than the one-point calibration of the linear VURF instrument.

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.

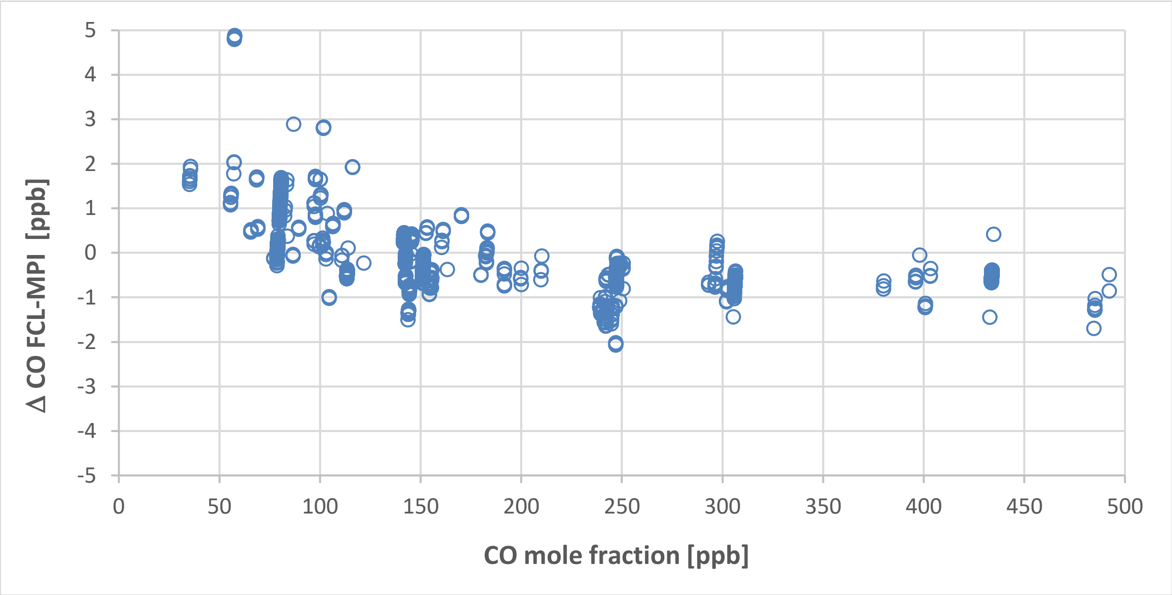


FIGURE 29: *MOLE FRACTION DEPENDENCE OF CO OFFSETS FOR SAMPLES ANALYZED AT FCL AND MPI-BGC*

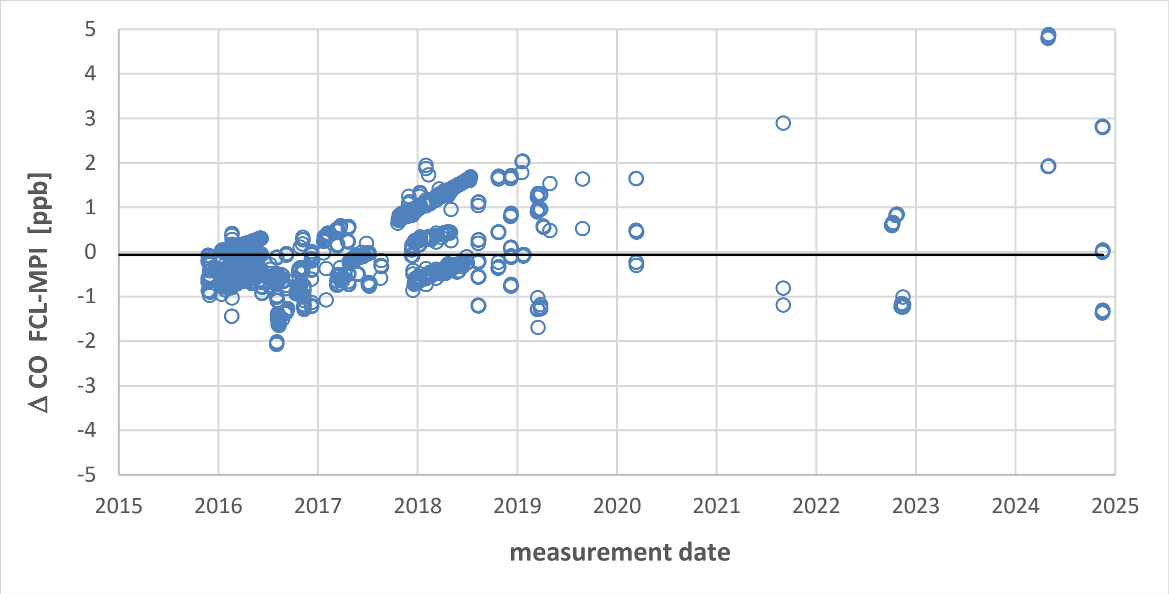
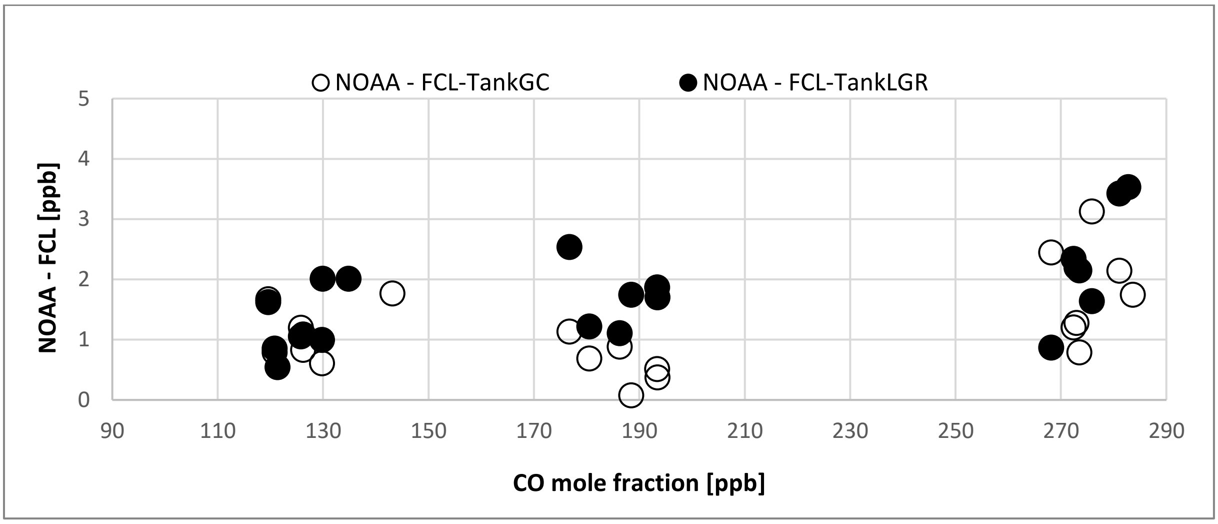


FIGURE 29: *CO OFFSET BETWEEN FCL AND MPI IN STANDARD MEASUREMENTS*

7.5.2 CO compatibility ICOS FCL - NOAA

A comparison with the NOAA-GML laboratory (and other laboratories) is routinely made in the Sausage Flask Intercomparison program. In this 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. The respective data are compiled in Fig. 30. The difference between FCL and NOAA increases with increasing CO, the mean CO offset for of all tank samples (black symbols) is NOAA-FCL = 1.5 ± 1.0 ppb.



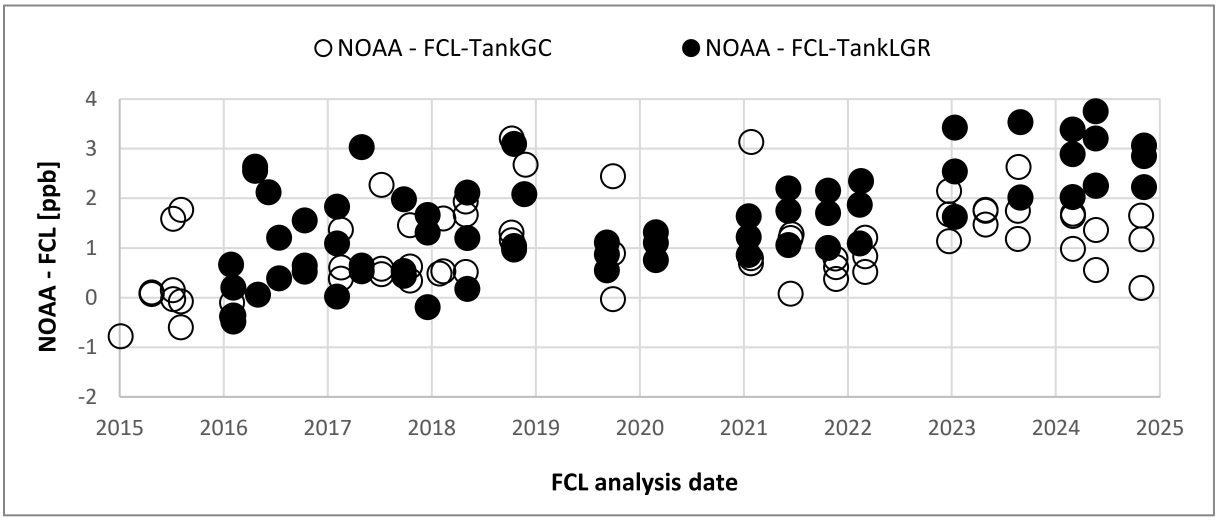


FIGURE 30: *CO OFFSET BETWEEN NOAA SAUSAGE FLASK DATA AND FCL DATA. Black dots represent FCL's analysis of the Sausage fill gas (filled symbol: LGR measurement; unfilled symbol: GC); comparisons are only considered if the flask pair agreement is < 6 ppb. The upper plot is based on data from 2019-2024 only.*

A complementary round robin test between NOAA (as WMO-CCL), EMPA (as WMO-WCC), MPI-BGC, FCL and FMI-ATC (ICOS Mobile Lab) (called "MENI" program) is made on an annual basis. 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 31 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, the CO growth rate is assessed based on the CCL measurement records. As the "blind" sample is analysed at different points of time only at the FCL, the CO growth is determined by these measurements and the FCL trend serves as reference. The offsets displayed in Fig. 31 are based on the respective reference trendlines. On average the offset between FCL and CCL is within 1 ppb but there may be a drift in the offset of the high comparison sample.

FIGURE 31: *CO OFFSET IN MENI ICP BETWEEN FCL, MPI, ICOS ML AND EMPA RELATIVE TO NOAA (LOWEST PLOT: RELATED TO ICOS)*

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 later years were made using various instruments. 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 mole fraction 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 2018 recalibration. This also contributes to the uncertainty. The difference in CO assigned values of the FCL Primary Standards in 2024 based on this extrapolation compared to when considering the complete CCL calibration record is on average 0.5 ppb for mole fractions below 200 ppb or above 700 ppb. For the two standards at 400 ppb and 700 ppb, an offset of 3 ppb appears. 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 0.8 ppb from 2016 to 2024).

7.6.2 CO scale transfer uncertainty

Attribution of the CO mole fraction trends in the individual FCL Secondary Standards is based on repeated calibrations using the FCL Primary Standards up to April 2019. 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 12 assignment periods are mole fraction dependent up to 0.2 ppb.

Comparison with results from calibration measurements against the FCL Primary Standards since 2019 shows a bias that has grown by 2024 to between 1 and 2 ppb between 150 and 420 ppb.

The assigned CO trend functions of the second set of FCL Secondary Standards are mostly based on daily measurement results relative to the first set of FCL Secondary Standards. For the first time since 2023 there have been trends in the residuals of Secondary Standards that have grown up to 0.3 ppb. These residuals have stabilized at this high level since 2024 with the replacement of the CAL1 of the Secondary Standards. This clearly points to limitations in the accuracy of the actually assigned CO trend of the FCL Secondary Standards. Once the new WMO scale will be available, those limitations are expected to be solved.

7.6.3 CO long-term reproducibility

To derive long-term reproducibility limitations beyond the random errors in daily measurements, discontinuities in the Target Standard measurement record have been used for CH_4, CO_2 and N_2O. The comparison of daily Target Standards measurements to the trends established by the measurements calibrated by the FCL Primary Standard set indicates such limitations also for CO. The dominant uncertainty from assignment inaccuracies result in residual trends of up to 2 ppb consistent with the external comparison offset trends. This limitation does not enable a sound quantitative assessment of a reproducibility term, though.

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 = 2 ppb (standards with CO < 400 ppb)

- The scale link uncertainty estimate is derived from the specified CCL assignment uncertainty (0.4 ppb below 400 ppb).
- Uncertainty of the FCL internal scale transfer to FCL Secondary Standard assignments (0.1 ppb).
- Uncertainty in the CO growth rates of the FCL Primary and Secondary Standards (2 ppb)

2. Measurement uncertainty of daily means = 0.05 ppb

- mean uncertainty of the daily calibration regression (0.04 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 (0.025 ppb, for 10 min means)

This uncertainty of daily means estimate is similar to the mean observed standard deviation of multiple daily means within one calibration period for individual samples (0.04±0.07 ppb, n=1076).

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. The scale transfer uncertainty and the measurement uncertainty do not contribute significantly and the overall uncertainty and internal reproducibility is assumed to be 2 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 WMO X2014A scale and do not include a term for any potential mole fraction 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. The complete set of FCL Primary Standards was sent to the CCL for the third then the fourth calibrations, performed in May 2021 and in November 2024, respectively. The reassignments by the CCL have generally been within the uncertainty of the initial assignment (? = -0.02 ±0.07 ppb) (see Table 11). However, there is a slight mole fraction dependent difference withFCL Primary Standards < 320 ppb having been determined too low initially.

TABLE 11: *N_2O X2006A ASSIGNMENTS FOR FCL PRIMARY STANDARDS BY THE CCL [PPB]*

Sample ID	Cylinder ID	CCL date 1	CCL date 2	CCL date 3	CCL date 4	N_2O date 1	N_2O date 2	N_2O date 3	N_2O date 4*	Assignment**
i20140055	CB09944	Jan-14	Jul-17	May-21	Nov-24	316.77	316.90	317.03	316.74	**316.90**
i20140056	CB09939	Jan-14	Feb-19	May-21	Nov-24	319.86	319.97	319.94	319.88	**319.92**
i20140057	CB09958	Jan-14	Oct-16	May-21	Nov-24	327.12	327.02	327.22	327.16	**327.12**
i20140058	CB09983	Jan-14	Jan-19	May-21	Nov-24	329.92	329.89	330.03	329.86	**329.95**
i20140059	CB09952	Apr-14	Nov-16	May-21	Nov-24	334.60	334.52	334.58	334.44	**334.57**
i20140060	CB09955	Jan-14	Jul-17	May-21	Nov-24	339.48	339.52	339.44	339.37	**339.48**
i20140061	CB09957	Jan-14	Nov-16	May-21	Nov-24	343.95	343.88	343.80	343.69	**343.88**
i20140062	CB09934	Mar-14	Jun-17	May-21	Nov-24	349.13	349.18	349.13	348.95	**349.15**
i20140054	CB09948	Jan-14	Jan-19	May-21	Nov-24	362.13	362.12	361.90	362.06	**362.05**

** based solely on GC-ECD data; ** Represents the mean of WMO X2006A date 1-date 3*

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 32 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).

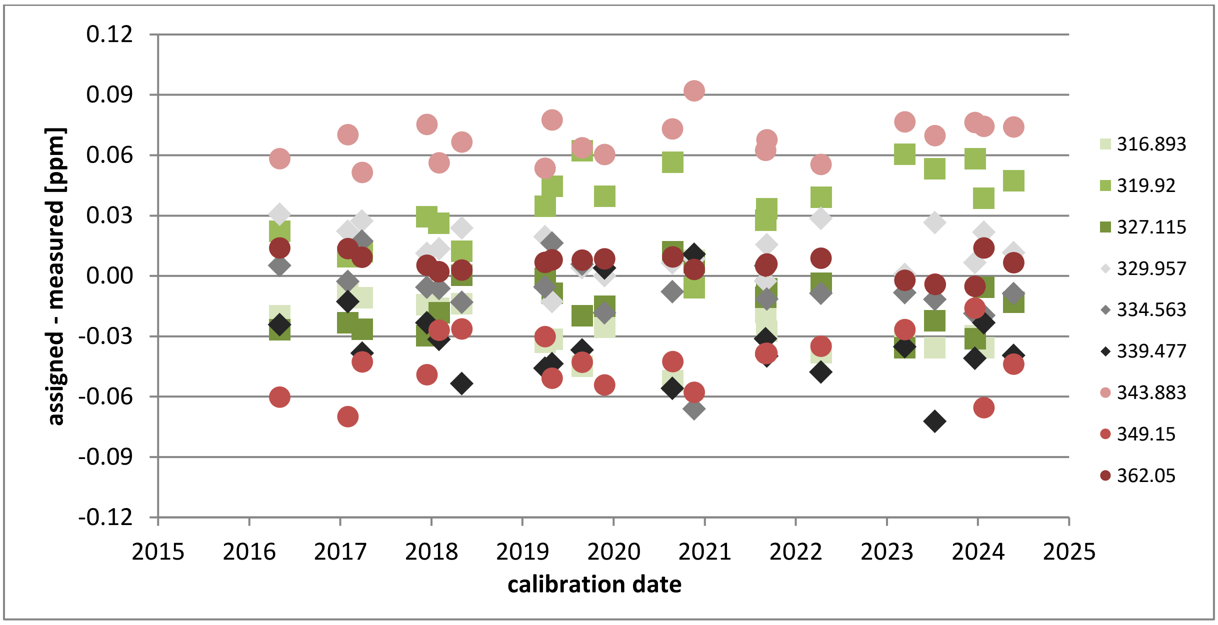


FIGURE 32: *TIME SERIES OF N_2O QUADRATIC REGRESSION FIT RESIDUALS OF LGR FCL PRIMARY STANDARD CALIBRATIONS. All data presented in the above figure refers to measurements by LGR 1. Data of 08.2022 and one of the two measurements per standard done in 03.2023 with LGR 2 are neither displayed nor used to assign FCL's secondary standards anymore (refer to section 8.4.1)*

8.2 FCL Secondary N_2O Standards

8.2.1 Assignment record

In January 2024 the basis of the assignments of FCL Secondary Standards changed to the averages of the first three CCL calibration episodes instead of the initial assignment. The four reference gases that had been used as initial FCL Secondary Standards for the daily LGR calibration have been analyzed within 6 to 15 calibration episodes (i.e. between 9 and 24 individual measurements) together with the complete set of FCL Primary Standards between May 2016 and September 2021. In May 2018 one of the Secondary Standard gases was exhausted and was succeeded by another standard with a similar N_2O content. The three remaining tanks were replaced when they were exhausted in July 2021. For the three replacement Secondary Standards, their initial assigned values in 2021 were determined by measurements against the first set of FCL Secondary Standards (see Table 12 for details).

TABLE 12: *N_2O ASSIGNMENTS FOR FCL SECONDARY STANDARDS [PPB]*

Sample ID	Cylinder ID	Date of start of use	Date of end of use	Assigned Value* (ppb)	Assigned Value** (ppb)	Previous basis of the assignment	n	Actual basis of the assignment	n
i20150251	CA05640	25-11-2015	03-05-2018	316.923	317.033	Primary Stds.	1	Primary Stds.	9
i20150189	D073397	25-11-2015	26-07-2021	324.506	324.562	Primary Stds.	1	Primary Stds.	24
i20150544	D073396	25-11-2015	26-07-2021	334.201	334.215	Primary Stds.	1	Primary Stds.	25
i20150191	D073395	25-11-2015	26-07-2021	344.970	345.023	Primary Stds.	1	Primary Stds.	24
i20170889	D557226	03-05-2018 / 26-07-2021	11-01-2024	315.58 / 315.682	315.677	Primary Stds.	1	Primary Stds.	21
i20201308	D753834	26-07-2021	NA	324.395	324.385	Secondary Stds.	114	Primary Stds.	7
i20201254	D753835	26-07-2021	NA	339.360	339.383	Secondary Stds.	114	Primary Stds.	6
i20201255	D753836	26-07-2021	NA	348.730	348.803	Secondary Stds.	117	Primary Stds.	7
i20201317	D753833	11-01-2024	NA	315.324	315.324	NA	NA	Primary Stds.	7

** Pre Sep 2024; ** Post Sep 2024*

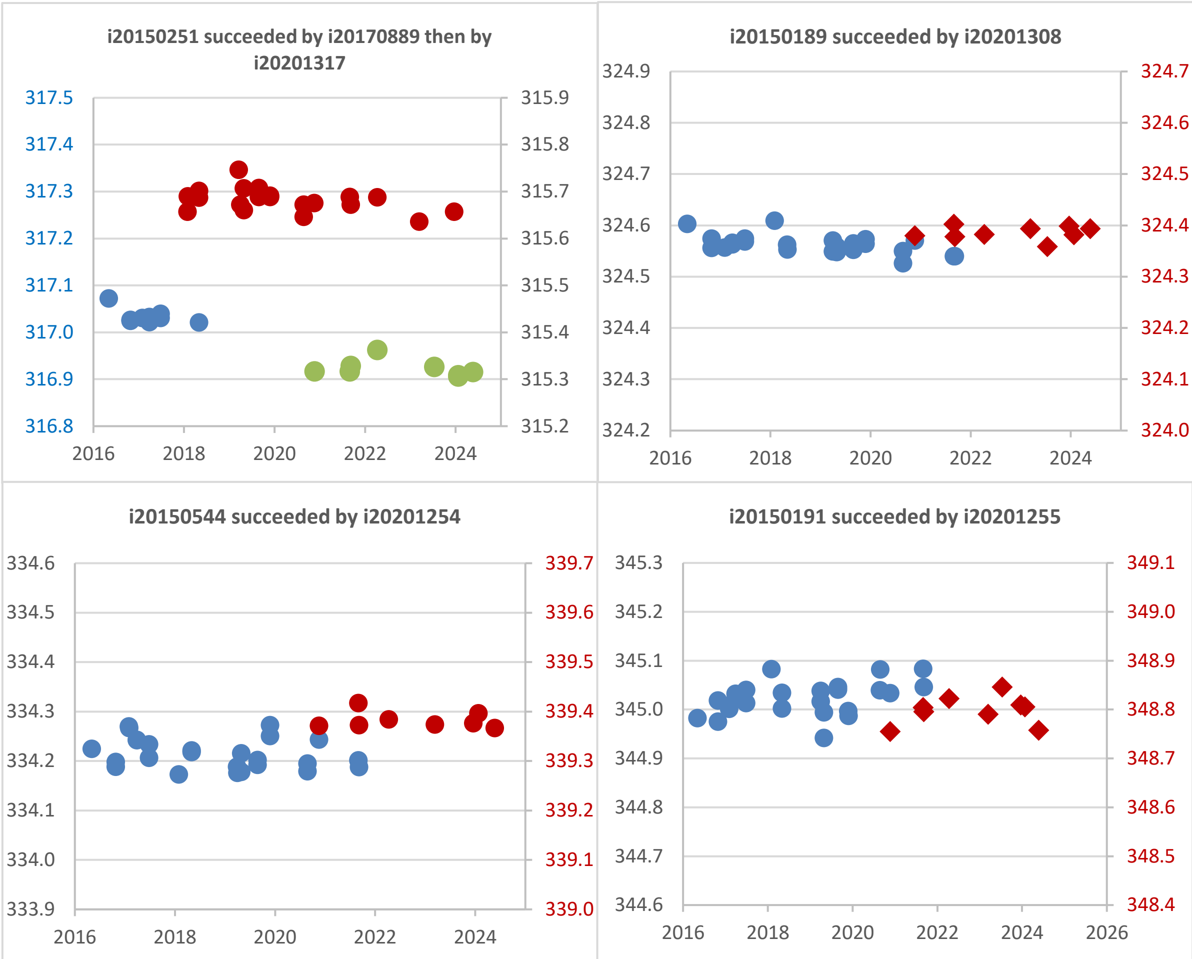


FIGURE 33: *SECONDARY N_2O STANDARD ASSIGNMENT RECORD (VALUES IN [PPB]).* Blue points represent the assignments for the first set of fcl secondary standards, the red points display the four new fcl secondary standards

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 generally smaller than 0.01 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.

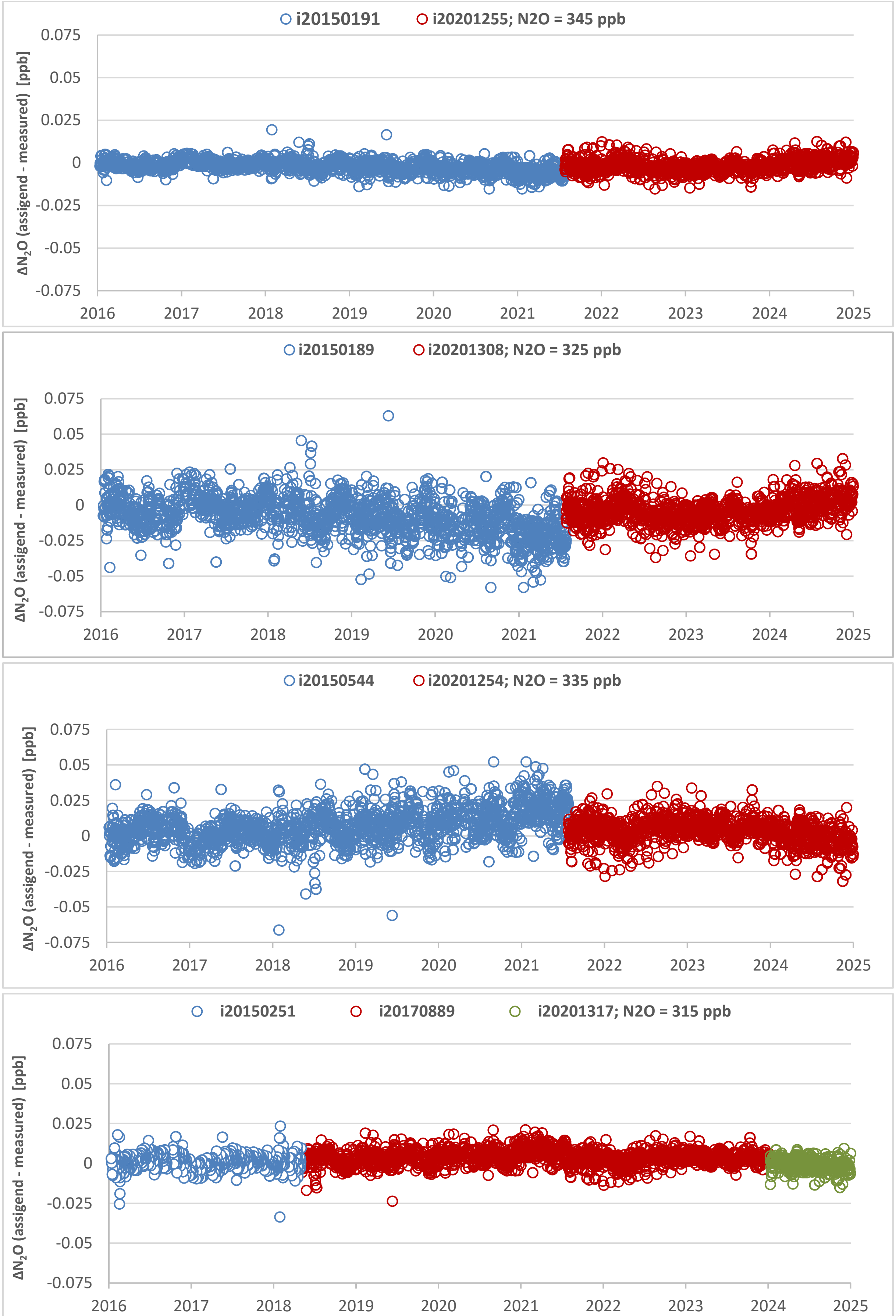


FIGURE 34: *QUADRATIC REGRESSION FIT RESIDUALS OF THE DAILY LGR N_2O CALIBRATION WITH FCL SECONDARY STANDARDS*

8.3 N_2O Targets

In the period from March 2016 to December 2024, two Short Term Targets have been constantly in use for the LGR system. They are complemented by additional Long Term Targets. This shall maintain a long-term link of succeeding (short term) targets. After an initial phase of daily analysis, they have been assessed on a regular, less frequent basis since 2020 to extend their lifetime.

The time series of the Target Standard N_2O measurement residuals is depicted in Figure 35. For mole fractions within the calibrated range, the agreement between assigned and mean measured value is generally very good (mean residuals ≤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 and at the end of its lifetime by up to 0.2 ppb. This provides some estimate for the uncertainty of measurements beyond the calibrated range. A failure of the LGR analyzer in July 2022 required an exchange of the instrument. The replacement analyzer (further on named LGR2) was operated until March2023 and data points in Figure 35 originating from measurements with the LGR2 analyser are displayed differently. Persistent small biases between the two analysers were confirmed by parallel analysis of the same samples in Feb-Apr 2023. In consequence, a decision was made together with the MSA during the meeting of November 2023 to reject all N_2O assignments of the standards prepared for ICOS atmospheric stations network that had been made with LGR2.

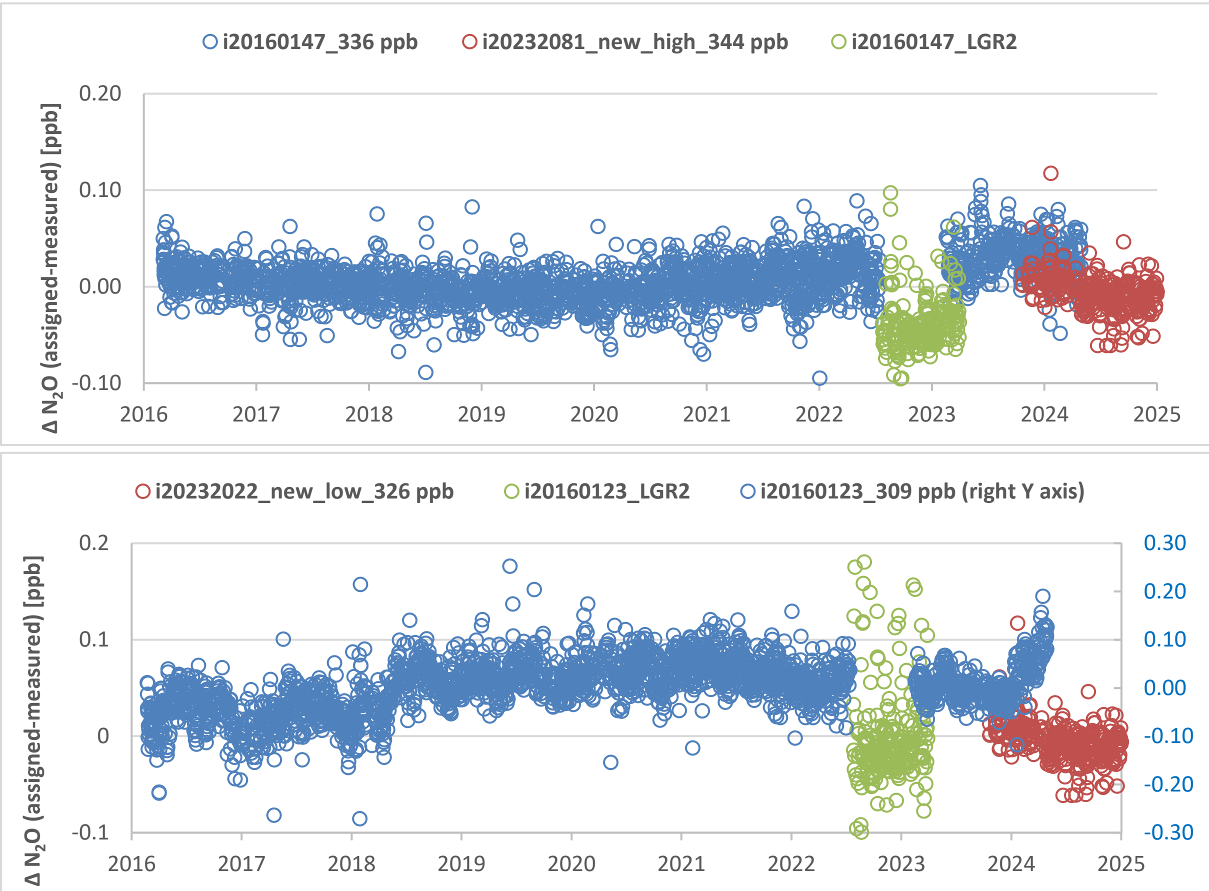


FIGURE 35: *TIME SERIES OF THE OFFSET OF N_2O TARGET MEASUREMENTS TO THEIR RESPECTIVE ASSIGNED VALUES***.**

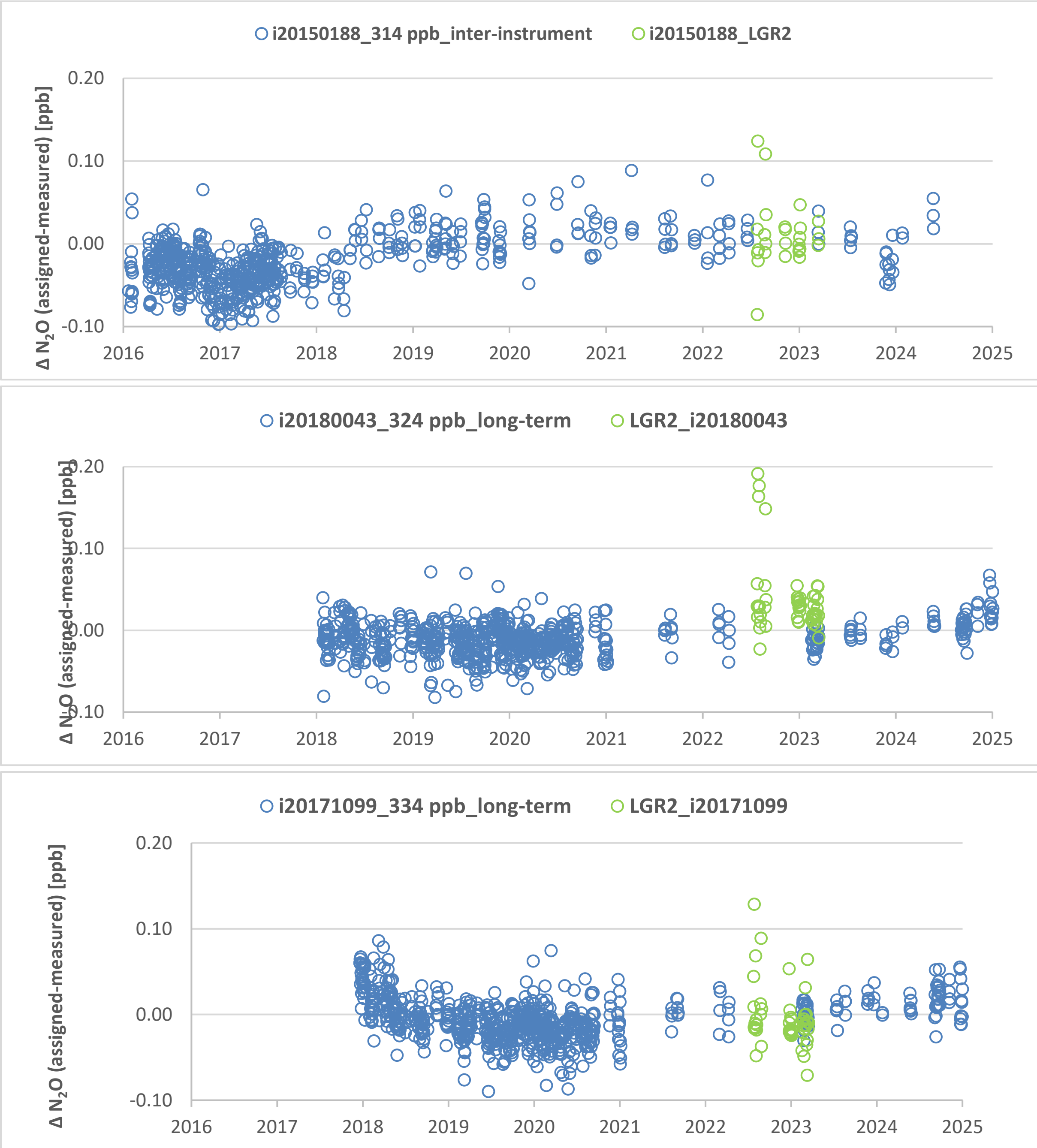


FIGURE 35: *TIME SERIES OF THE OFFSET OF N_2O TARGET MEASUREMENTS TO THEIR RESPECTIVE ASSIGNED VALUES*.

8.4 Internal N_2O Comparison

8.4.1 N_2O comparison of two LGR instruments

A failure of the LGR analyzer in July 2022 required an exchange of the instrument. The replacement analyzer (further on named LGR2) was operated until March 2023. The original instrument (further on named LGR1) was repaired on February 16th, 2023 In the following weeks, several standards had been analyzed on both instruments simultaneously. Figure 36 shows the offsets of the mean results of the comparison of both instruments, including also other standards that have been analysed with time lags of up to 20 months. Results show a total average offset of 0.05 ± 0.06 ppb for all samples. These biasses remain constant over time without any underlying cause yet having been identified.

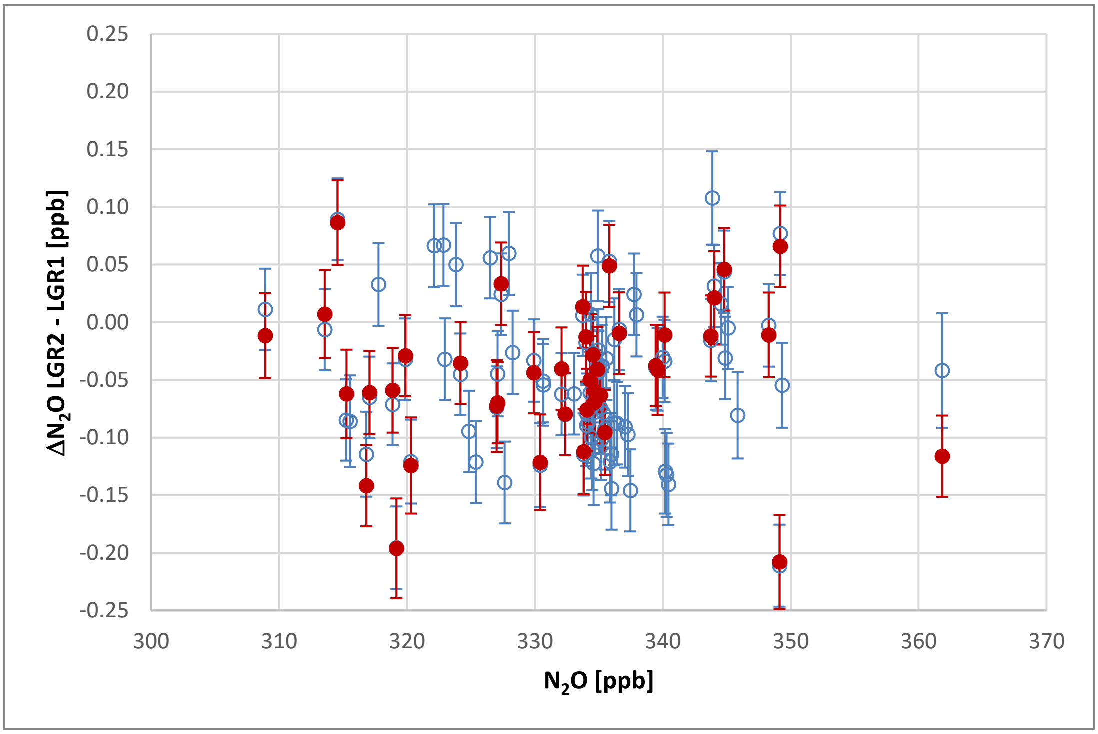
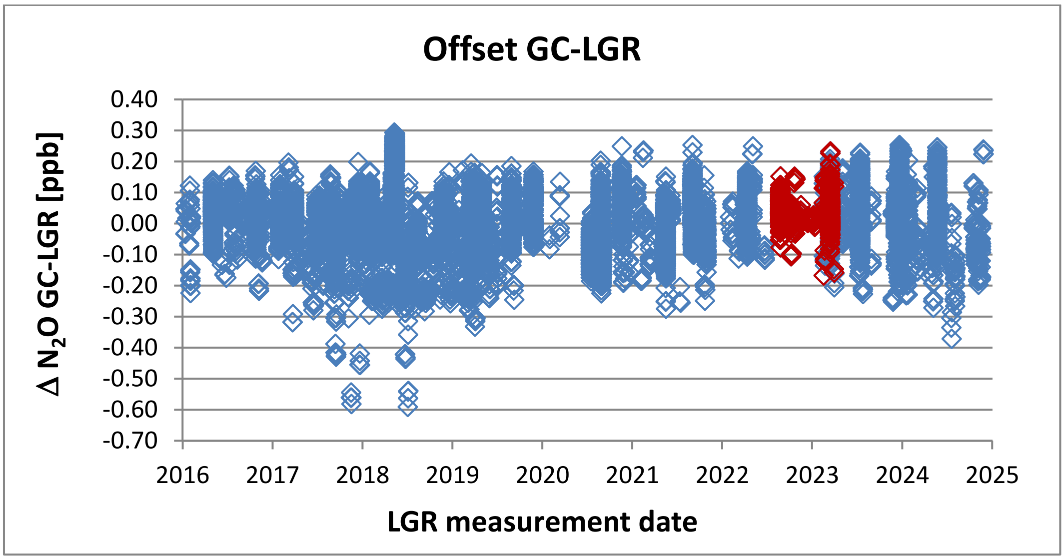


FIGURE 35: *N_2O OFFSET OF TWO LGR INSTRUMENTS FOR THE SAME SAMPLES.*
Blue symbols represent offsets of total means of all measurement days, red symbols represent data analysed on the same day. Samples within 308 - 362 ppb are considered, both analyzers are operated in the same way and based on the same set of FCL Secondary Standards; error bars are combied uncertainties assessed in section 8.6.4 2. and 3.

8.4.2 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 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 mole fractions 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 -0.01 ppb ± 0.10 ppb for LGR 1 (based on 228 samples) and +0.04 ± 0.05 ppb for LGR 2 (based on 25 samples) (see Figure 37).



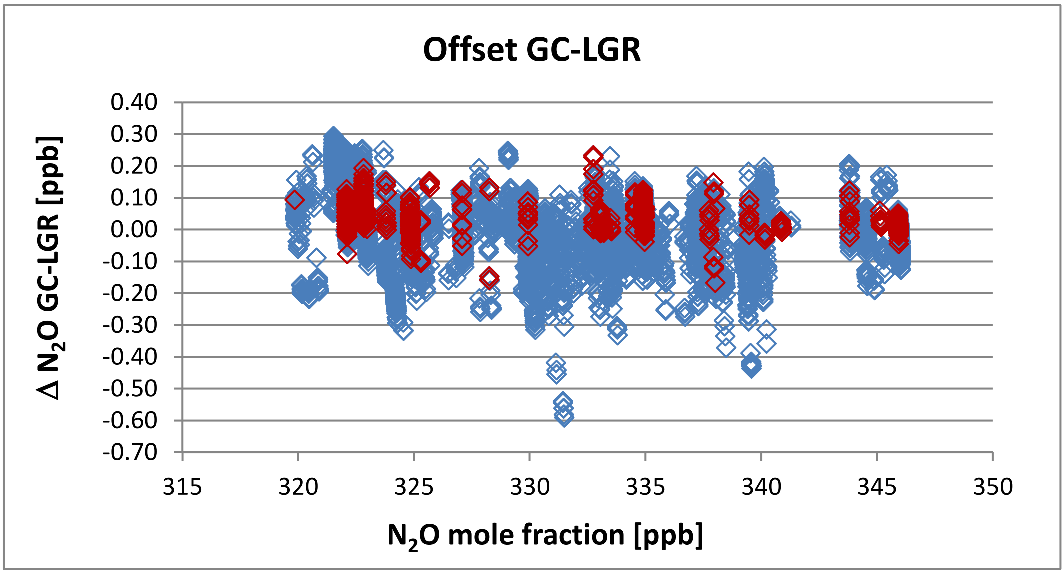


FIGURE 36: *OFFSETS OF DAILY LGR N_2O MEASUREMENTS RELATIVE TO THE ANNUAL MEAN OF GC RESULTS. Blue diamond symbols refer to GC-LGR 1 comparison results and red diamond symbols refer to GC-LGR 2 comparison results.*

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 CCL assignments - MPI BGC measurements of the FCL Primary Standard set of -0.09 ± 0.04 ppb is observed whereas an offset of CCL assignments - FCL measurements of the MPI BGC Primary Standard set of 0.12 ± 0.07 ppb is apparent which also shows up with 0.13 ± 0.04 ppb in the CCL-FCL difference for the FMI set in that range. Including also the measurements of UBA and TNO standard sets, an overall offset of 0.07 ppb ± 0.08 ppb is observed for all standards in the range relevant for atmospheric measurements (325 - 350 ppb); one such comparison was made using LGR2 (see section 8.4.1) in 2022 (DLR set) resulting in a very similar offset of CCL-FCL = 0.10 ppb. This is consistent with the standard assignment uncertainty of 0.11 ppb specified by the CCL and a corresponding offset CCL-MPI = -0.09 ppb for the FCL Primary Standards as shown in Figure 38.

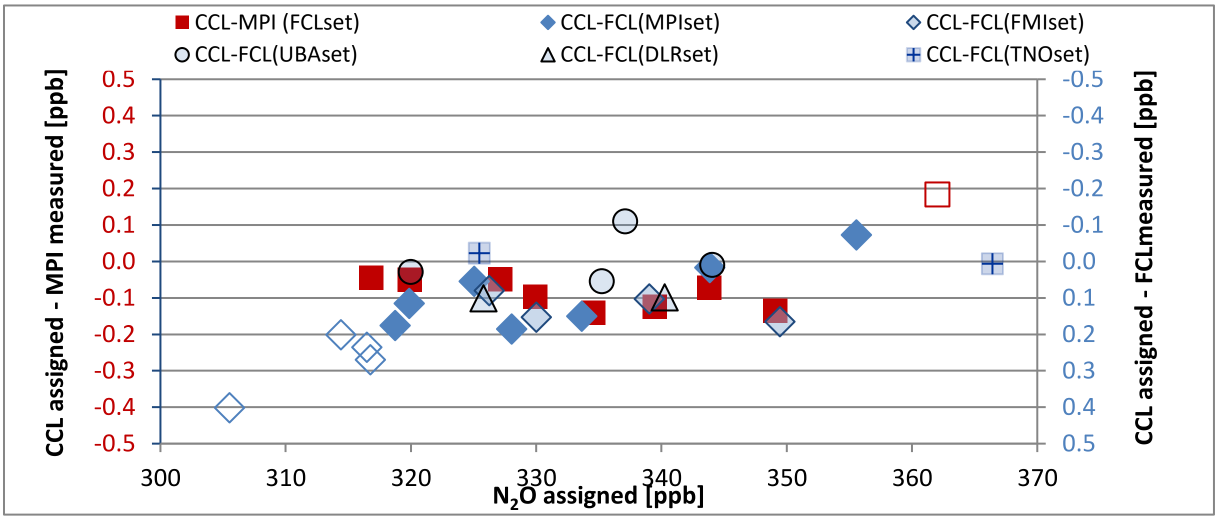


FIGURE 37: *DIFFERENCES OF PRIMARY STANDARD MEASURED N_2O RESULTS TO CCL ASSIGNED VALUES MPI-BGC measurements of FCL Primary set (red squares- right y axis) and FCL measurements of MPI-BGC (blue diamonds, open symbols represent values extrapolated beyond the calibrated range set by Secondary Standards), 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 193 comparisons (for FCL LGR values only) are presented in Figure 39. The average offset of MPI-FCL within the Secondary Standards' range amounts to 0.09 ppb ± 0.16 ppb. This corresponds to the offset established in the preceding section and confirms the mole fraction dependence.

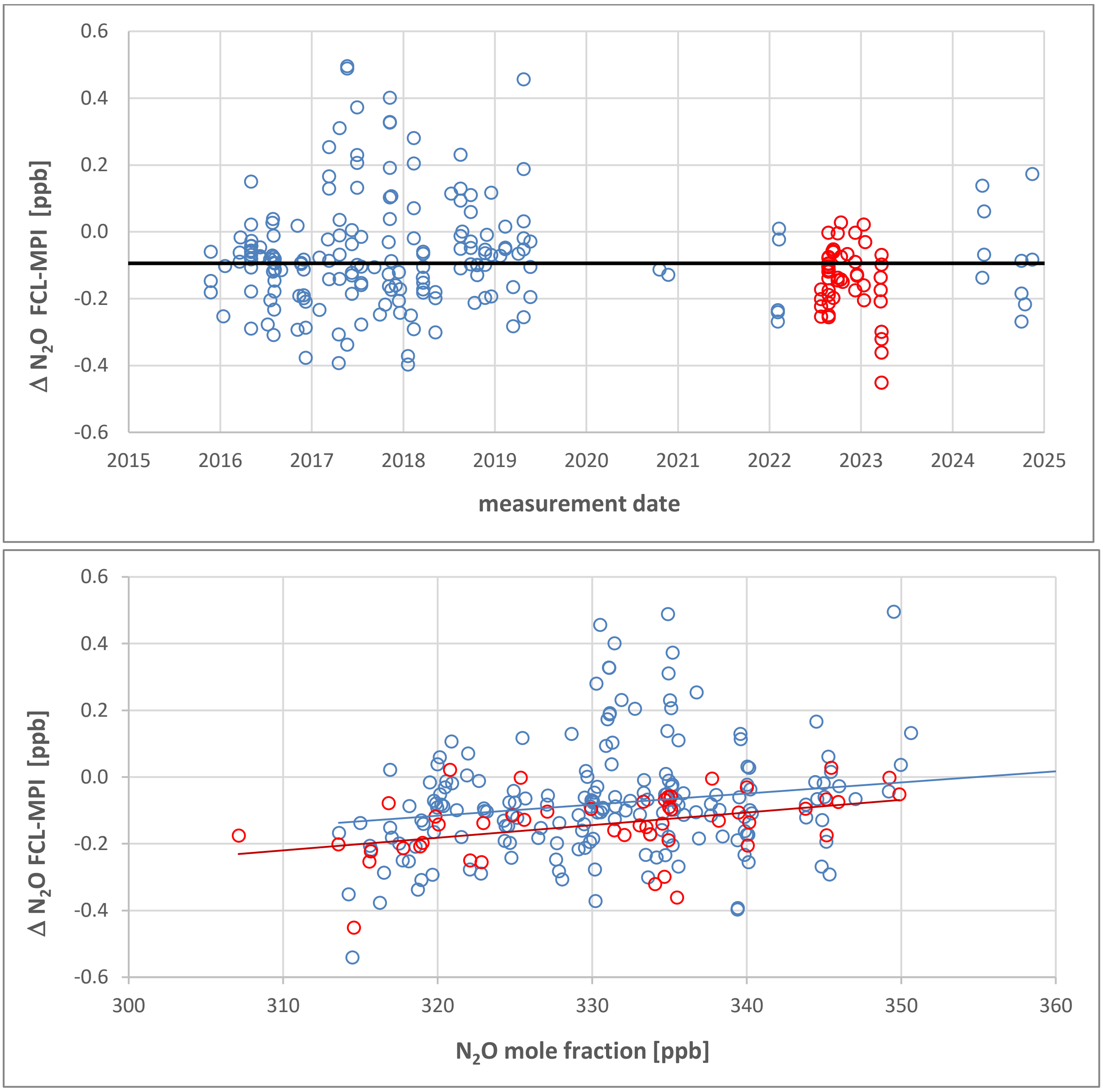
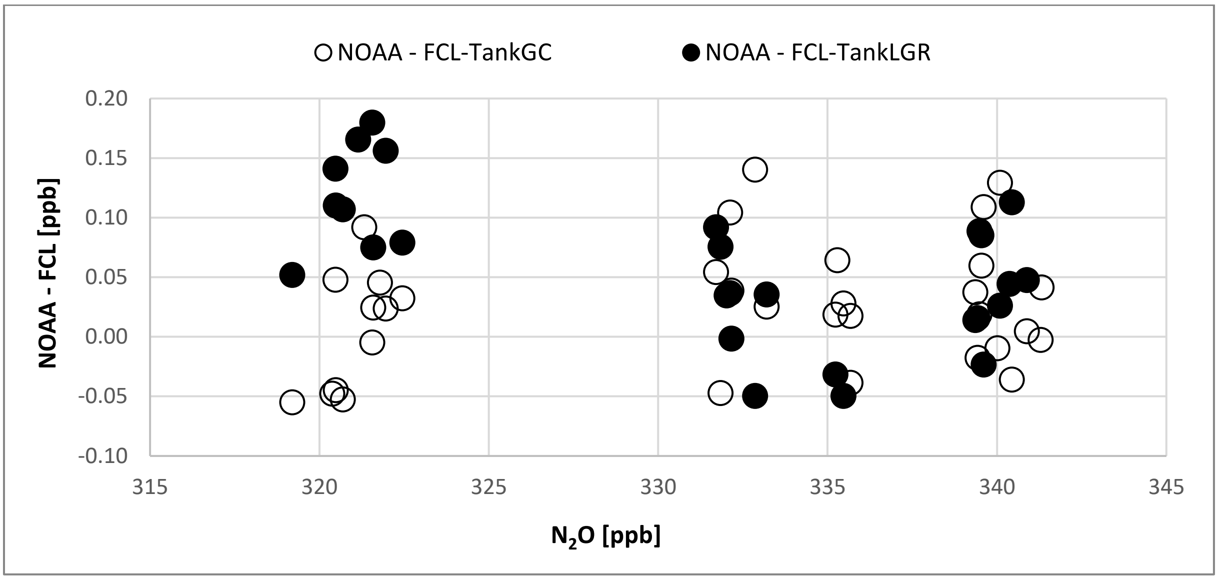


FIGURE 38: : *N_2O OFFSET BETWEEN FCL AND MPI IN STANDARD MEASUREMENTS. All MPI-BGC GC measurements since 2015 with minimum 6 injections within the range of 313 - 350 ppb are considered in aggregated means.*

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 40). The agreement of all valid samples (defined by a flask pair agreement within 0.7 ppb) yields a difference of NOAA-FCL = 0.06 ppb ± 0.2 ppb. In summer 2019 the NOAA laboratory has changed instrumentation resulting in a similar but more stable agreement of NOAA-FCL = 0.06 ppb ± 0.06 ppb.



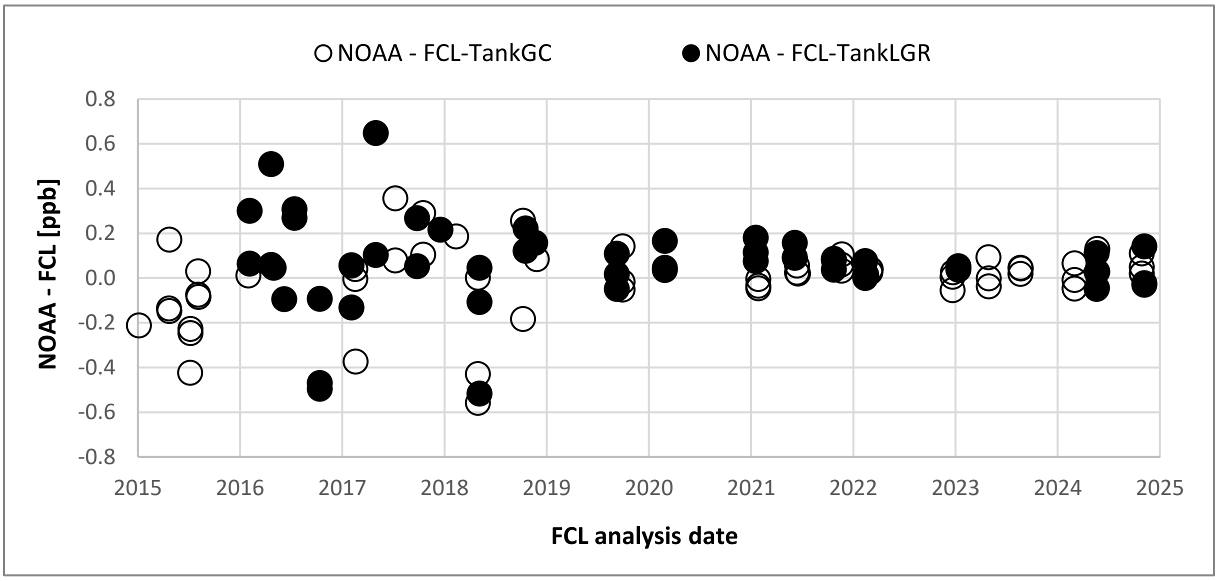


FIGURE 39*: N_2O OFFSET BETWEEN NOAA SAUSAGE FLASK DATA AND FCL DATA.* Black dots represent FCL's analysis of the sausage fill gas (filled symbol: LGR measurement; unfilled symbol: GC) comparisons are only considered if the flask pair agreement is < 0.5 ppb. The upper plot is based on data from 2019-2024 only.

The MENI round robin test between NOAA (as WMO-CCL), EMPA (as WMO-WCC), MPI-BGC, FCL and -ATC (ICOS Mobile Lab) is made on an annual basis to check the ICOS scale link to the WMO mole fraction. 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 are shown in Figure 41. The observed offset FCL - CCL is -0.01 ± 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. The last data points in 2022 all are results from the LGR2 (displayed as crosses, not considered for the overall mean).

FIGURE 40: *N_2O OFFSET IN MENI ICP BETWEEN FCL, MPI, ICOS MOBILELAB AND WCC RELATIVE TO NOAA*

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. This is consistent with the absolute residuals of the FCL Primary Standard set being on average 0.03 ppb and with the standard deviations of the four CCL assignments being on average 0.09 ppb. The compatibility of the FCL Primary Standard set with other CCL calibrated standards (held by MPI-BGC, FMI, UBA, DLR) yield a systematic offset of 0.07 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 0.09 ppb. This offset includes the assignment uncertainties of each calibration gas set.

8.6.2 N_2O 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 standard deviation of these assignments of individual Secondary Standards is approximately 0.02 ppb. The uncertainty of the scale transfer depends on the number of calibration events that range from 8 to 25 and is on average estimated to be 0.006 ppb.

The absolute mean values of the regression fit residuals of the daily calibration using the Secondary Standards are on average < 0.004 ppb for all individual standards. This is consistent with the above estimate and confirms very small uncertainties for the FCL internal scale transfer.

The overall small difference of 0.01 ppb between GC measurements and LGR ones of the same samples also confirms small internal scale transfer uncertainty.

A comparison of the FCL Primary Standard set with other CCL calibrated standards (held by MPI-BGC, FMI, UBA) was made. On average a systematic offset of FCL - CCL of - 0.07 ± 0.08 ppb for gases with $N_2O \geq 320$ ppb was established. The reverse assessment of FCL Primary Standards by the MPI-BGC laboratory results in a similar mean offset of 0.09 ppb.

This systematic small offset is consistent with results from ongoing comparison activities with NOAA (refer to section 8.5.2).

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 without the reason being always understood. The related uncertainty is approximated by the standard deviations of monthly averaged N_2O measurement residuals of the target standards resulting in 0.02 ppb from 2015-2024.

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 Secondary Set assignments (0.006 ppb)

2. Measurement uncertainty of daily means = 0.024 ppb

- mean uncertainty of the daily calibration regression fit (0.013 ppb)
- uncertainty of the detector response drift throughout the validity of a daily calibration (0.02 ppb)
- uncertainty from the repeatability of the daily sample measurements (0.004 ppb)
  This uncertainty of daily means estimate is similar to the mean observed standard deviation of multiple daily means within one calibration period for individual samples (0.028±0.006 ppb, n=1438).

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 0.11 ppb (k=1). The FCL reproducibility is estimated to be 0.031 ppb.

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 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 2018] Information published at CCL's webpage: https://gml.noaa.gov/ccl/co_scale.html (accessed May 2025)

[CCL_CO_2 2021] Information published at CCL's webpage: https://gml.noaa.gov/ccl/co_2_scale.html (accessed May 2025)

[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)

[Crotwell 2024] Presentation by A. Crotwell at the WMO meeting GGMT2024:

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%2F2024%2FGGMT%202024%2Fpresentations%2FGGMT%5F2024%5FPdf%2FGGMT%5F2024%5FCrotwell%5F0782024%2Epdf&parent=%2Fsites%2Fwmocpdb%2Feve%5Factivityarea%2FGlobal%20Atmosphere%20Watch%20Programme%20%28GAW%29%5F63452102%2D7575%2De911%2Da98e%2D000d3a44bd9c%2FMeetings%2F2024%2FGGMT%202024%2Fpresentations%2FGGMT%5F2024%5FPdf&p=true&ga=1 (accessed May 2025)

[Crotwell 2025] Personal communication in April 2025

[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 et al. 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 mole fraction 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 retrieves mole fractions by analyzing the characteristic absorption of light of infrared-active molecules (near-IR spectroscopy). In February 2024, the Picarro analyzer in use since the start of FCL operations in 2015 (S/N CFADS2193), was taken out of service and replaced by a new Picarro G2301 CO_2/CH_4 CRDS analyzer (S/N 2696-CFADS2461). The method of analysis was also optimized (refer to Mode of Operation section).

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.A1.1). 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 database, quality checks and calibration of the instrument are performed in a self-controlled manner.

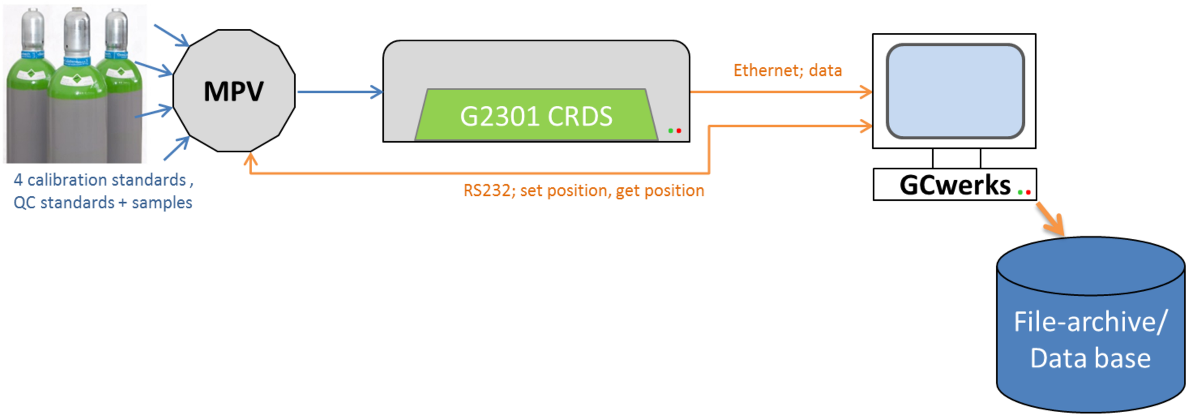


FIGURE A1.1: *SCHEMATICS OF THE INSTRUMENTAL SETUP, BLUE LINES= SAMPLE, ORANGE LINES= DATA/COMMANDS*

Mode of Operation:

The operator defines the sequence of analysis using the GCwerks software at the supervising PC. Required information is shown in Table A1.1 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 database 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 A1.2). 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.

TABLE A1.1: *EXEMPLARY PORTS.LOG META LOOK-UP TABLE*

Date	Time	# port	Sample	Regulator	Type
240712	1430	6	i20222329	Tes1-021	qc
240712	1430	3	i20222170	Sco2-005	qc
240712	1430	1	i20190708	Tes1-009	cal
240712	1430	5	i20190803	Sco2-001	cal
240712	1430	9	i20190709	Sco2-002	cal
240712	1430	13	i20190438	Tes1-004	cal
240712	1430	12	i20242122	Tes1-007	tank
240712	1430	15	i20242100	Tes1-008	tank
240712	1430	14	i20242384	Sco3-008	tank

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 A1.2 shows an exemplary analysis sequence containing 2 target samples (qc), 4 calibration gases (cal), two samples (tank) and the closing purge gas (for 60 minutes).

TABLE A1.2: *EXEMPLARY SAMPLING SEQUENCE*

Duration [min]	Procedure	Type	# port
5	picarro.runfile	tank	14
30	picarro.runfile	cal	1
5	picarro.runfile	tank	14
30	picarro.runfile	cal	5
5	picarro.runfile	tank	14
30	picarro.runfile	cal	9
5	picarro.runfile	tank	14
30	picarro.runfile	cal	13
5	picarro.runfile	tank	14
30	picarro.runfile	tank	12
5	picarro.runfile	tank	14
30	picarro.runfile	qc	3
5	picarro.runfile	tank	14
30	picarro.runfile	tank	15
5	picarro.runfile	tank	14
30	picarro.runfile	qc	6
60	picarro.runfile	tank	14

The Picarro method changed from 27th Februry 2024, with a longer measurement time of thirty instead of the previous twenty minutes. In addition, all samples measurements are preceded by five minutes of cell flushing with purge gas. During the initial ten minutes of sample measurement (previously five minutes), the results are discarded with respect to running-in effects, like purging of the cavity, and allows for equilibration in pressure regulators, thermal equilibration and settling of the regulating loops. These changes aim to further reduce the carry-over effect from the last sample in the Picarro cavity to exclude a bias towards the mole fraction of the succeeding sample and such to improve the overall long-term reproducibility.

The instrument itself runs at 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 200 mbar 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 mole fraction values to calculate the calibration coefficients of this day. For calibration of CO_2 and CH_4 a linear equation is applied.

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 mole fractions close to the boundaries set by the range of the calibration gases to give a conservative assessment that is meaningful for all mole fractions. The 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. One of those is shared between different instruments in the laboratory to assess the link of their respective results on a regular basis.

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 A1.3). 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; outlet valve value and variability).
- The measurement results of the target standards relative to their known composition
- The regression fit coefficients and 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 mole fraction 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 mole fractions by analyzing the characteristic absorption of light of infrared-active molecules (near-IR spectroscopy). The instrument uses the technical principle of Off-Axis Integrated Cavity Output Spectroscopy (OA-ICOS) (see Fig.A2.1).

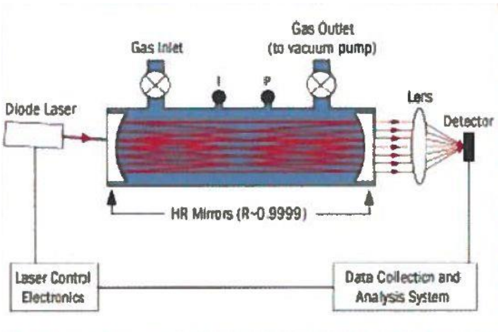


FIGURE A2.1: *SCHEMATIC DIAGRAM OF AN OA-ICOS ANALYZER*

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.A2.2). Under knowledge of the gas temperature, pressure in the cell, effective path length and known line strength the mole fraction can be calculated from the integrated loss-signal following Lambert-Beer's-Law.

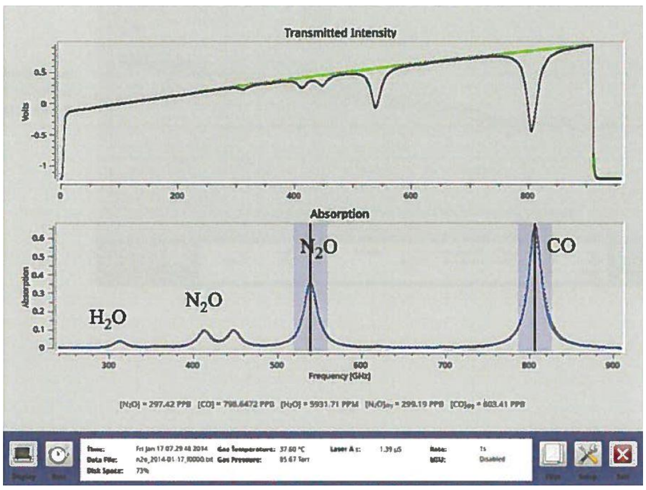


FIGURE A2.2: *SCREEN SHOT OF SPECTRUM DISPLAY, UPPER PANEL SHOWS PHOTO DETECTOR VOLTAGE, LOWER PANEL SHOWS OPTICAL ABSORPTION OF SPECIES OF INTEREST.*

Instrumental Setup:

Sample flow and cell pressure are controlled and protocolled automatically. Figure A2.3 gives an overview of the sample flow and meta information retrieval within the instrument.

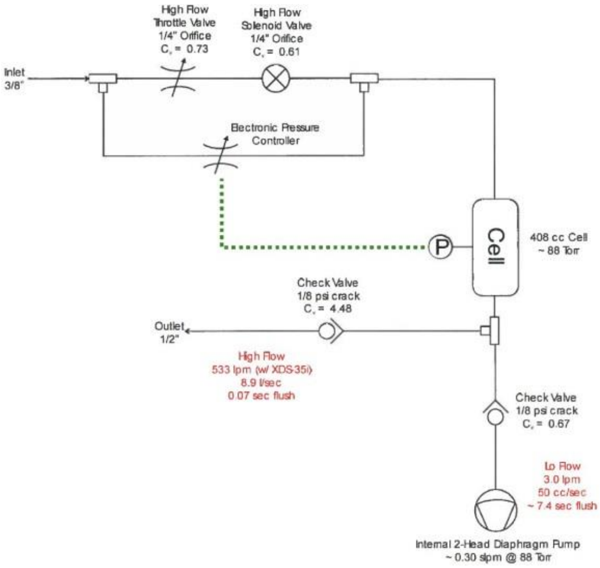


FIGURE A2.3: *INTERNAL FLOW SCHEMATICS OF THE LGR INSTRUMENT*

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. A2.4). 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 database, the data processing includes a short term stability correction, automated quality checks and automated calibration of the instrument.

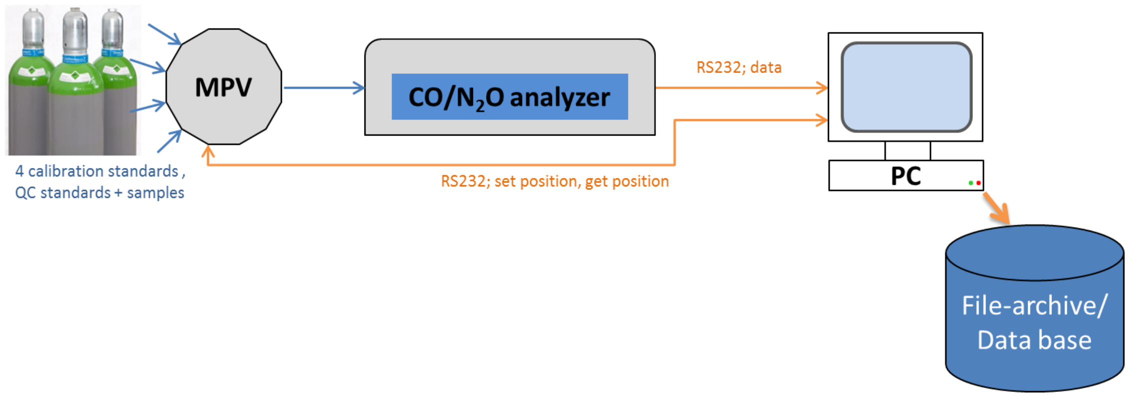


Figure A2.4: Schematics of analysis station, blue lines = sample, orange lines = data/commands

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 A2.1 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 A2.1 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.

TABLE A2.1: *EXEMPLARY SAMPLING SEQUENCE*

Duration [min]	# port	Sample	Description
20	10	i20240065	WT
20	2	i20201317	Cal1_D753833
20	10	i20240065	WT
20	3	i20201308	Cal2_D753834
20	10	i20240065	WT
20	6	i20201254	Cal3_D753835
20	10	i20240065	WT
20	5	i20201255	Cal4_D753836
20	10	i20240065	WT
20	1	i20242104	sample_D337364
20	10	i20240065	WT
20	4	i20232022	QClow_D761211
20	10	i20240065	WT
20	11	i20170274	sample_D073386
20	10	i20240065	WT
20	4	i20232081	QChigh_D073383
20	10	i20240065	WT
720	15	i20241823	purging

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 and equilibration in pressure regulators, thermal equilibration and settling of the regulating loops. The instrument itself runs at 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 100 mbar 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 20 min, 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 mole fraction values to calculate the calibration coefficients of this run. For calibration of CO and N_2O, quadratic equations are applied.

Dedicated standards, called Targets are regularly analyzed to quality control the instrument's performance including the daily calibrations. Two of them are included within every sequence. They have CO and N_2O mole fractions close to the boundaries set by the range of the calibration gases to allow a conservative assessment that is meaningful for all mole fractions. Another Target 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.

Data evaluation:

A regular analysis sequence consists of alternate measurements of the Working Standard and samples (including Targets that are used for quality control assessment). 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),
- Measurement results of the Target Standards relative to their known composition,
- Regression fit coefficients and 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 III

Overview of parameters for automized flagging of 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 600 seconds
P	PCavity with range 139.99 ... 140.01
MISS CO_2	Missing value in CO_2 related measurements
MISS CH_4	Missing value in CH_4 related measurements
OPV	OutletProportinalValve Flag 29999 … 35000
SDMinRaw CO_2	Standard deviation of MinRaw data, static upper bound: 0.035
SDMinRaw CH_4	Standard deviation of MinRaw data Flag, static upper bound: 0.3
INMinRaw CO_2	Insufficient number (of MinRaw values)
INMinRaw CH_4	Insufficient number (of MinRaw values)
RC	Insufficient number of calibration gases, <4
CO_2_DYN_sd	static dynamic upper bound: 0.07 threshold: 450
CH_4_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 N_2O	Missing value in N_2O related measurements
H_2O	leakage on the basis of water signal
H_2O sd	leakage on the basis of water signal stdev
CO sd	Standard deviation of CO out of range, minimum: 0.00022
N_2O sd	Standard deviation of N_2O out of range, range: 0 … 0.0002
RC	Insufficient number of calibration gases, <4
N_2O_DYN_sd	Dynamic upper bound, Minimum: 4.0e-04 dyn_poly: [0.0006357375, 9.565958e-05]
CO_DYN_sd	Dynamic upper bound, Minimum: 2.2e-04 dyn_poly: [0.001074092, 6.098591e-05]
CO_wt_diff	Absolute difference of the series means of two neighboring WT CO measurements, Maximum: 0.0005
N_2O_wt_diff	Absolute difference of the series means of two neighboring WT N_2O measurements, Maximum: 0.0002

Annex IV

CO_2 mole fraction measurement calibrations using an isotopolgue selective analyzer

The analyzer for CO_2 calibration used in the FCL as well as in the ICOS observational network is applying the CRDS technique. This method is selective only for the $^{12}C^{16}O_2$ isotopologue. However, the standard gases to calibrate the analyzers have CO_2 mole fraction assignments from the WMO-CCL for total CO_2 that account for the complete suite of all CO_2 isotopologues. So in principle, this calibration approach is working without bias only if the fractional abundance of the main CO_2 isotopologue of the standard gases is similar to the one observed in the atmosphere. Figure A4.1 displays the relationship between the CO_2 mole fraction and the fractional abundance of its main isotopologue derived from the $\delta^{13}X$ and $\delta^{18}O$ data for the FCL Primary Standard gases and background atmosphere, respectively. The atmospheric values represent data points from flask sample data of the ICOS Jungfraujoch background station using MPI-BGC flask data from 2007-2024 [Heimann et al. 2021].

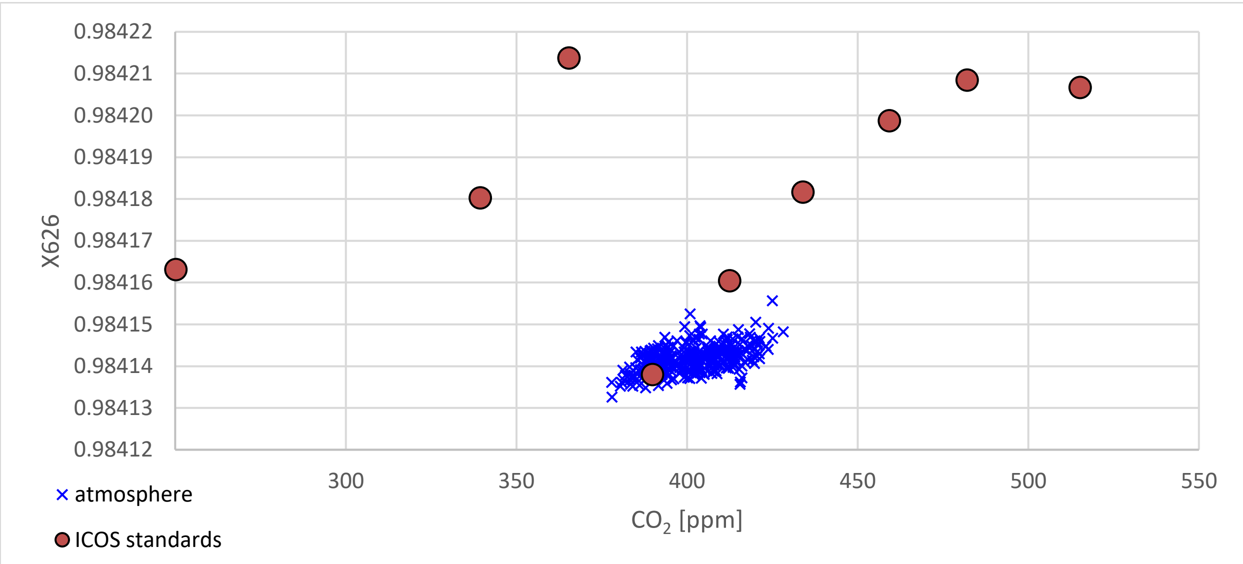


FIGURE A4.1: *CO_2 ISOTOPE VS. CO_2 MOLE FRACTION RELATIONSHIP IN STANDARD GASES AND ATMOSPHERIC SAMPLES*

Modification of CO_2 isotopic composition resulting from preparation of standard gases

Standard gases are prepared at FCL on the basis of compressed, dried real air collected at the roof tops of either the MPI-BGC building at the South-Western edge of Jena city or the FCL building close to Jena city centre. To prepare standard gases with sub-atmospheric mole fractions of CO_2 and other trace gases (CH_4, CO, N_2O; in the case of the FCL primaries also SF_6) the CO_2 is partly taken out using molecular sieve as scrubber which is mostly followed by an addition of pure CO_2 to achieve the wanted composition. Other standards only undergo the spiking step. For this spiking there are two pure CO_2 gases available with 13C either depleted or enriched relative to atmospheric CO_2 ($\delta^{13}C$ = -2 ‰ and -38 ‰, respectively). The spiking is generally made such that the selected relative amounts of each of the two CO_2 that are added result in a $\delta\ ^{13}C$-CO_2 value that is expected to match the range of typically observed $\delta\ ^{13}C$-CO_2 in atmospheric CO_2. In contrast, the δ18O value of both CO_2 spike gases is more negative in either spike gas (-15 ‰ and -30 ‰, respectively) than in the atmosphere (0 … -2 ‰) causing spiked standards to exhibit a higher fraction of the $^{12}C^{16}O_2$ isotopologue as displayed in Fig. A4.1.. This is similar to the situation described by the WMO-CCL [Tans et al. 2017].

Mole fraction adjustments accounting for standard - atmosphere isotope mismatch

Table A4.1 lists the CO_2 mole fractions of the FCL Primary Standards and their measured isotope delta values relative to the VPDB-CO_2 scale. For each standard gas the isotope amount-fraction ($X^{12}C^{16}O_2$) of the main isotopologue $^{12}C^{16}O_2$ relative to total CO_2 is calculated. This calculation is based on the $\delta^{13}C$- and $\delta^{18}O$- CO_2 measurement results by the CCL, $\delta^{17}O$- CO_2 data that are deduced from a $\delta^{17}O$ to $\delta^{18}O$ relationship of 0.5281 [Assonov and Brenninkmeijer 2003] and the isotope-amount fractions for the VPDB reference as compiled by Tans et al 2017:

13xVPDB = 0.010564 (eq. 4a [Tans et al. 2017]), 17xVPDB-CO_2 = 0.0003941 (eq. 4b [Tans et al. 2017]), 18xVPDB-CO_2 = 0.0020832 (eq. 4c [Tans et al. 2017]).

The resulting X12C16O2 std of the standard gas is then compared to the X12C16O2 atm that is expected to be observed in the atmosphere at the respective mole fraction based on the trend line through the data points presented in Figures 1a and 1b. The ratio of xstd/xatm indicates if a larger fraction of CO2 is detectable by the analyzer in either the standard gas or the atmosphere and therefore serves as adjustment factor assigned values for total CO2 by the WMO-CCL. The correction term is insignificant for the FCL Primary Standards at current atmospheric background CO2 mole fractions (<0.01 µmol/mol) but systemetically biased for low and high CO2 standards by up to 0.03 µmol/mol. Depending on their isotopologue composition other standard gases assigned by the WMO-CCL will have different adjustment factors. In the WMO tertiary set held by the MPI-BGC GasLab adjustments of up to 0.05 µmol/mol were required.

Table A4.1: *FRACTIONAL ABUNDANCE OF THE $^{12}C^{16}O_2$ ISOTOPOLOGUE IN STANDARD GASES AND ATMOSPHERIC AIR AND DERIVED CO_2 ASSIGNMENT UPDATE FOR STANDARD GASES*

FSN	UCN	CO_2 [ppm]	$\delta^{13}C$ [‰]	$\delta^{18}O$ [‰]	$X^{12}C^{16}O_{2\,std}$	$X^{12}C^{16}O_{2\,atm}$	adj.factor	CO_2 [ppm] $_{iso\_adjusted}$
i20140054	CB09948	250.119	-8.5	-5.4	0.984171	0.984124	1.000047	250.131
i20140055	CB09944	339.352	-8.4	-9.2	0.984186	0.984137	1.000050	339.369
i20140056	CB09939	365.279	-9.4	-13.9	0.984218	0.984140	1.000079	365.308
i20140057	CB09958	389.756	-8.3	0.4	0.984142	0.984144	0.999998	389.755
i20140058	CB09983	412.417	-9.7	-1.0	0.984164	0.984147	1.000018	412.424
i20140059	CB09952	433.834	-9.7	-5.9	0.984185	0.984150	1.000036	433.850
i20140060	CB09955	459.174	-11.7	-4.8	0.984202	0.984154	1.000049	459.197
i20140061	CB09957	482.015	-12.4	-4.8	0.984211	0.984157	1.000055	482.041
i20140062	CB09934	515.113	-12.2	-5.0	0.984209	0.984161	1.000049	515.138

In order to avoid any such measurement bias the assigned values by the CCL should be adjusted to the value specified in the last column of Table A4.1.

Erroneous initial X2019 mole fraction assignments

The CO_2 mole fractions listed in Table A4.1 in the last column are 0.02 µmol/mol lower than those listed in the QC-Report Table 3. This is a result from an arithmetic error made when initially calculating the $X^{12}C^{16}O_2$ amount fraction that was discovered. While an update of this error internally in the CAL database is a moderate work effort it is a big computational work load to reprocess all continuous CO_2 measurements in the ICOS network. This requires that the correction needs to be done in collaboration with the ATC in due course.