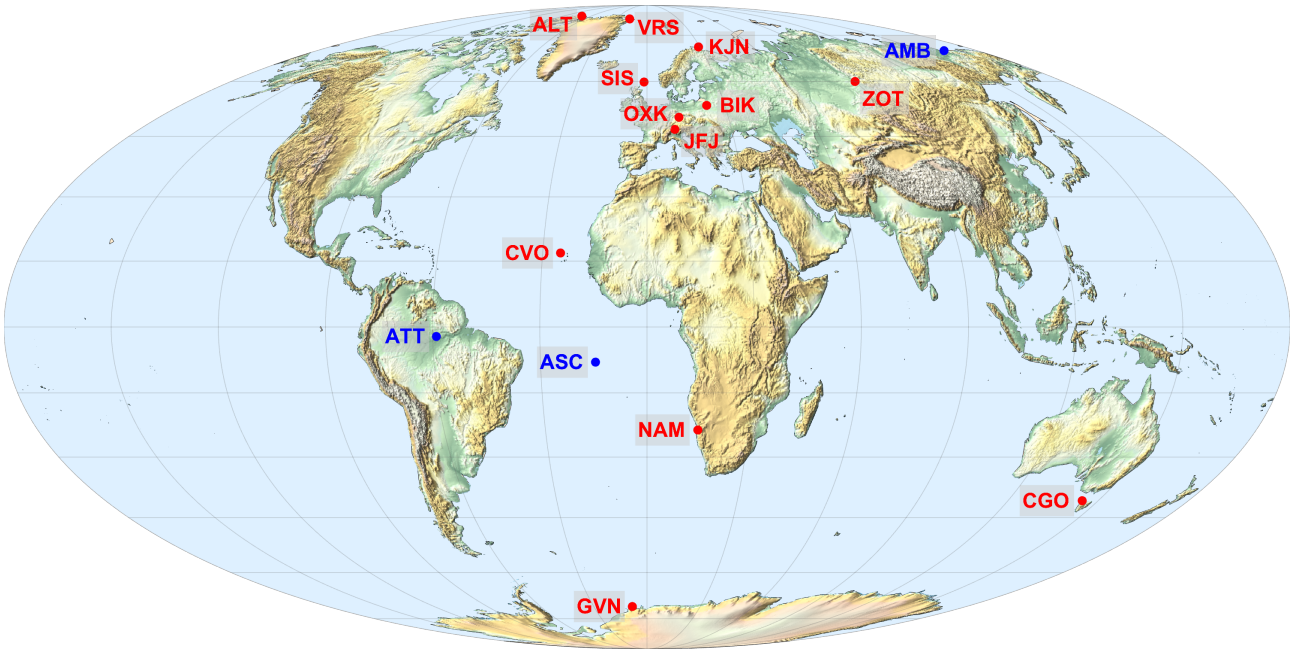
The atmospheric flask sampling program of MPI-BGC
Version 13, 2022

Martin Heimann, Armin Jordan,
Willi Brand, Jost Lavric, Heiko Moossen, Michael Rothe



(MPI-BGC atmospheric greenhouse gas measurement network. Red: Stations with flask sampling.)

Max Planck Institute for Biogeochemistry, Hans-Knöll-Str. 10, Jena, Germany
https://dx.doi.org/10.17617/3.8r January 11, 2022

Contents

1 Introduction

Many of the main biogeochemical cycles are reflected in the atmosphere by important greenhouse gases (GHG) or other gases. Among these are CO_2, CH_4, N_2O, H_2, CO and O_2. Additional GHGs include also anthropogenically produced species, such as SF_6. The spatio-temporal distribution of these species in the atmosphere provides valuable information on location and temporal evolution of their sources and sinks. In addition, the isotopic composition of the gases bears valuable information on their origin because of differences in their source or sink isotopic signatures. To assess this information requires the establishment of an observation system for measurements of these atmospheric gases and their isotopic composition. Since most of these species have relatively long atmospheric lifetimes, their spatio-temporal concentration gradients are very small and measurements have to be made with very high accuracy and, for comparability among different networks, according to internationally documented calibration scales (Crotwell et al., 2019).

As part of its scientific objectives, the MPI for Biogeochemistry (MPI-BGC) since its foundation in 1997 has been developing the capacity to contribute to the regional (primarily Europe) and the global atmospheric observation network for biogeochemically relevant atmospheric species. By 2021, MPI-BGC has established regular atmospheric measurements at 15 sites (see Table 1). Measurements include *in situ* continuous measurements at tall towers (ATT, OXK, BIK, ZOT) and remote coastal sites (CVO, NAM, AMB), as well as flask sampling sites (ALT, VRS, KJN, SIS, JFJ, CGO, GVN) and a site for remote sensing of the vertical column composition at ASC.

The atmospheric flask sampling program of MPI-BGC has three main objectives:

- Flask samples analysed in house at MPI-BGC provide an independent quality control on *in situ* measurements at the remote stations.
- In house analyses of the flask samples allow the measurement of additional species, which provide additional information on atmospheric transport and/or on sources and sinks, such as isotopic composition as well as gases that are not easy to measure continuously at remote sites due to technical and/or logistical limitations.
- Taking flask samples at sites operated by partner institutions provides an important possibility to intercompare the measurement programmes of the different laboratories. Intercomparisons of the MPI-BGC measurements with other laboratories are possible at Alert (ALT), Jungfraujoch (JFJ), Ochsenkopf (OXK) and Cape Grim Observatory (CGO).

After the laboratory facilities were built up by the year 2002, the MPI-BGC atmospheric flask program was gradually expanded to presently 12 sites with regular sampling. Typically, at each site every week flask triplets are filled with atmospheric air and subsequently shipped to the institute for measurement analysis. Routinely, on all flask samples measurements by gas chromatography (GC) are performed of the mixing ratio of CO_2, CO, CH_4, N_2O, H_2, and SF_6. In addition, measurements of the isotopic composition of CO_2 ($^{13}C/^{12}C$ and $^{18}O/^{16}O$), as well as O_2/N_2 and Ar/N_2 are performed by mass spectrometry (MS). Since 2012 also the isotopic composition of methane ($^{13}C/^{12}C(CH_4)$ and $D/H(CH_4)$) is routinely determined.

This report documents the procedures adopted for the atmospheric flask sampling program of MPI-BGC. In the following we describe the characteristics of the measurement station network followed by the procedures for the flask sampling. We then document the in house analysis methods and the employed calibration scales for the different species. This is followed by the description of the post processing and flagging, the Ar-correction for the O_2/N_2 measurements and a brief assessment of the overall quality of the measurement program. Graphical displays of the measurements from some of the stations are shown in Appendix B.

This report describes version 13.4 of the full data release containing measurements up to October 2021. With a continuing flask sampling program future updated releases will become available on a yearly basis.

2 MPI-BGC atmospheric GHG measurement station network

As of the year 2021, BGC atmospheric GHG measurements are or have been performed regularly at 15 stations, of which 12 include a flask sampling program. Figure 1 shows a global map of the station locations, while a summary of the station characteristics are given in Table 1. For more information see Appendix C.

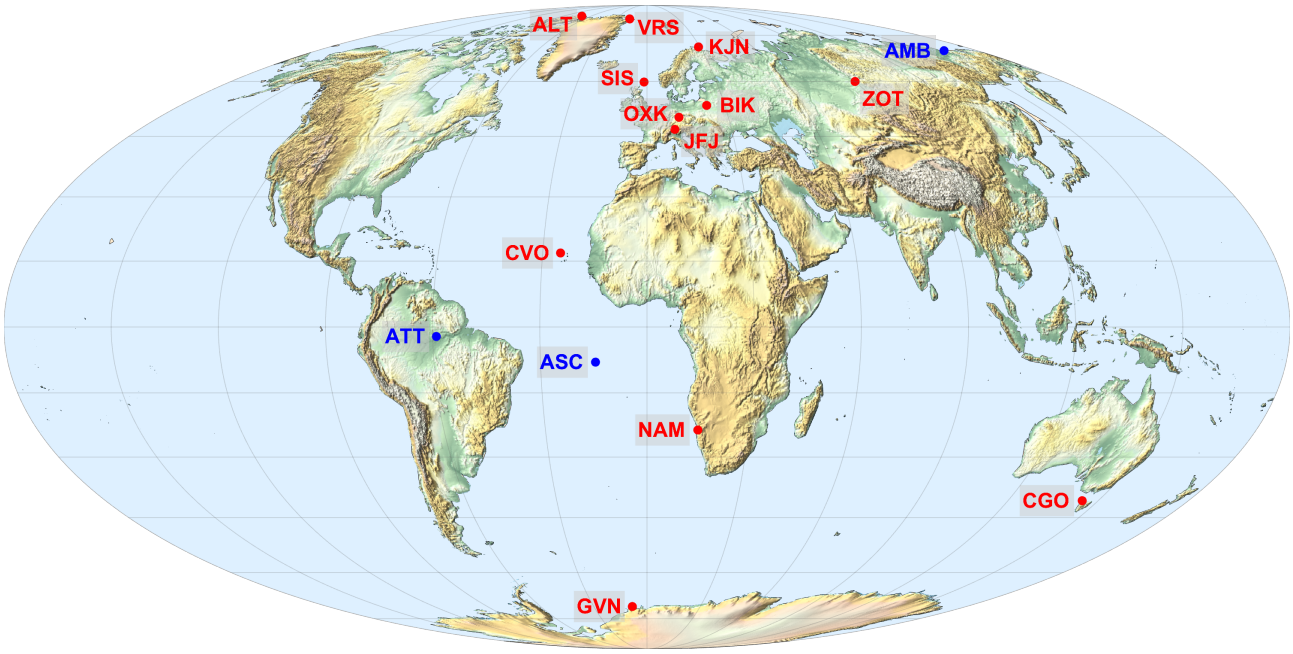


Figure 1: Map of the MPI-BGC atmospheric GHG measurement network. Red labels: stations with regular flask sampling, blue labels: stations with greenhouse gas measurements but no flask sampling.

Station-code	Station name and country	Latitude	Longitude	Sampling altitude*	Type**	Begin***	Station characteristics	Partner group
ALT	Alert, Canada	82.45N	62.51W	5m (210m asl)	F + I_2	2004	Arctic coastal background station	Environment Canada
VRS	Villum Research Station (Station Nord), Greenland	81.58N	16.64W	50m (10m asl)	F + I_2	2019	Arctic coastal background station	Aarhus University (Denmark)
KJN	Kjølnes, Norway	70.85N	29.23E	10m (5m asl)	F + I_2	2014	Arctic coastal background station	University of Exeter, UK
ZOT	ZOTTO, Russian Federation	60.80N	89.35E	301m (114m asl)	F + I_1	2006	Zotino Tall Tower Observatory (300m tall measurement mast)	Institute of Forest, Krasnojarsk, Russian Federation
SIS	Shetland Islands, UK	59.85N	1.27W	5m (20m asl)	F	2003	Coastal marine boundary layer site	R. Robertson, Sumburgh Head, Shetlands
BIK	Bialystok, Poland	53.23N	23.03 E	300m (183m asl)	F + I_1	2005	300m tall communication tower	Aerometeo Bialystok, Poland
OXK	Ochsenkopf, Germany	50.03N	11.80E	163m (1022m asl)	F + I_1 F + I_2	2006	TV tower	MPI-BGC, DWD/ICOS (since 2016), Germany
JFJ	Jungfraujoch, Switzerland	46.82N	9.86E	(3570m asl)	F + I_2	2008	High Altitude Alpine site	University of Bern, Switzerland
CVO	Cape Verde Atmospheric Observatory	16.86N	24.87W	30m (10m asl)	F + I_1	2007	Coastal marine boundary layer background site	INMG, Cape Verde, University of Exeter, UK
NAM	Gobabeb, Namibia	23.56S	15.05E	21m (408m asl)	F + I_1	2013	Desert/coastal site	Gobabeb Namib Research Institute, Namibia
CGO	Cape Grim, Australia	40.68S	144.69E	70m (94m asl)	F + I_2	2002[x]	Coastal site	CSIRO, Australia
GVN	Neumayer station, Antarctica	70.67S	8.28W	4m (40m asl)	F + I_2	2017	Antarctic coastal background station	Alfred Wegener Institute and University of Heidelberg, Germany
AMB	Ambarchik, Russian Federation	69.63N	162.39E	25m (10m asl)	I_1	2014	Arctic coastal site	NESS, Russian Federation
ATT	ATTO, Brazil	2.15S	59W	79m, 325m (120m asl)	I_1	2012	Amazon Tall Tower Observatory (325m tall measurement mast erected in 2016)	INPA et al., Brazil
ASC	Ascension Island, UK	7.97S	14.5W	–	R	2012	FTIR Instrument	ESA, DLR Germany (since 2020)

Table 1: Summary of main MPI-BGC atmospheric GHG measurement sites. *Altitude above local ground level; numbers in parentheses show altitude of local ground level above sea level. **Measurement type: F: Flask sampling, I: *in situ* measurements (I_1: *in situ* measurements operated by MPI-BGC, I_2: *in situ* measurements operated by partner group), R: remote sensing instrument. ***Year of begin of regular measurements by MPI-BGC. [x]Discontinued after 2018.

3 Flask sampling procedures

Air sampling is performed with 1l glass flasks equipped with PCTFE seals which exhibit a better performance compared to other sealing O-ring polymers such as Viton, PTFE or PFA (Sturm et al. (2004), Rothe et al. (2005)), see Figure 2. As only exception to this, samples from CGO are collected in 1.6 L stainless steel flasks as part of the SF_6 observation programme by the University of Heidelberg (Fraser et al. (2004)).

Flasks are conditioned at MPI-BGC prior to shipping to the sampling stations. Thereby all flasks undergo an initial pre-treatment at MPI-BGC that is repeated whenever a flask is repaired or exposed to undried air. They are evacuated for 3 days to 0.5 mbar at a temperature of 60 °C to remove residual humidity from the surface. This has been shown to be crucial to maintain the $\delta^{18}O - CO_2$ value of CO_2 of an air sample over time (Ghosh et al. (2005)). Flasks are also conditioned each time prior to shipping to the sampling stations by flushing them with 90 L of dry standard air (Rothe et al. (2005)) at the same pressure that they are normally filled with. This serves to maintain the dryness of the glass surface and to minimize pressure changes during the sample collection. It also helps to identify flasks that where not successfully sampled if the sample composition is found to be identical as the one of the conditioning gas.

At the stations generally three flasks are filled with local ambient air using a portable flask sampler (Figure 3), except for Bialystok where an automated flask sampling system from NIWA has been installed (Popa et al. (2010)) filling two flasks simultaneously and except for Cape Grim where the CSIRO sampling equipment is used to fill just one sample (Fraser et al. (2004)). It includes an air pump, a drying agent and gauges for monitoring the line and flask filling pressure. The air drying is achieved either using a cartridge with anhydrous magnesiumperchlorate or a cryodryer. Table 2 lists the drying method in place at the different sampling stations. The flasks to be sampled (in general triplets) are connected in series to the sampler and are filled by flushing for at least 15 min at their sampling pressure. Originally the filling pressure was approximately 2 bar; this was later reduced to about 1.6 bar in order to reduce fractionation effects caused by gas diffusion through microleaks in the PCTFE seals (particularly noticeable in O_2/N_2).

The standard instructions for the flask sampler operators at the stations are listed in Appendix D. The filled flasks are afterwards shipped back to MPI-BGC for subsequent analysis. Figure 4 shows the flask processing flowchart.

Station	Drying method
ALT	m, c (since 2016)
BIK	c (-90C)
CGO	m
CVO	m, c (since 05/2008)
JFJ	m
KJN	c (-80C)
NAM	c (-80C)
GVN	c (-50C)
OXK	m, c (-80C since 04/2014)
SIS, SIS0	m
VRS	c (-45C)
ZOT	m, c (since 05/2009)

Table 2: Air drying methods employed at the different sampling stations. m: Magnesiumperchlorate, c: cryodrier.

Figure 2: Sampling flask with PCTFE seal

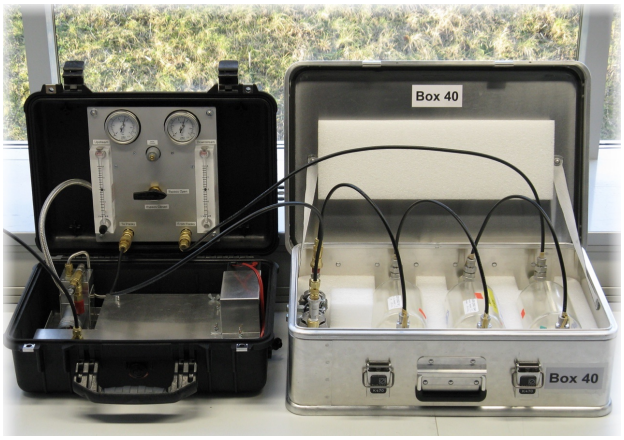


Figure 3: Flask sampler

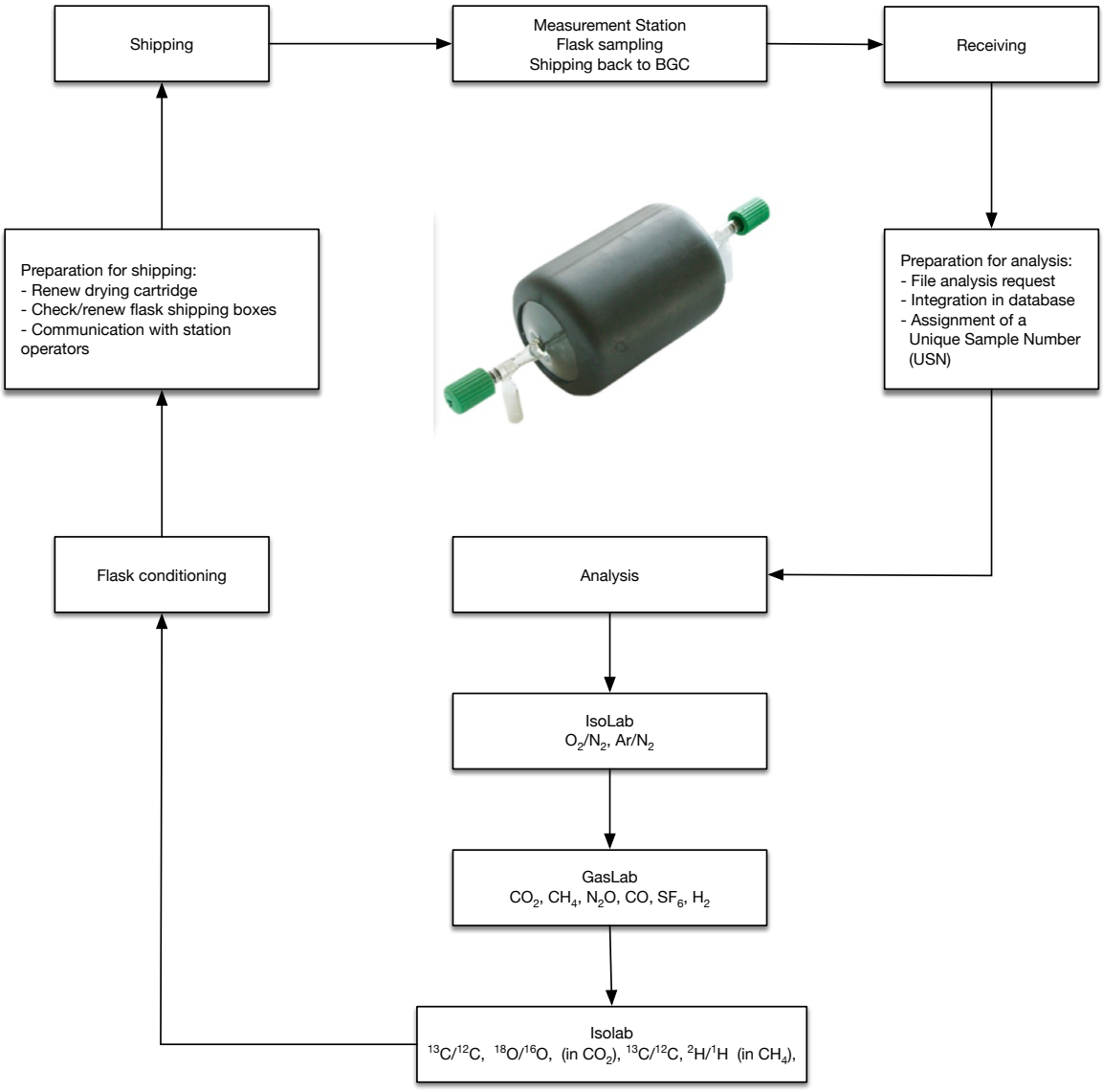


Figure 4: Flowchart of the flask sampling, shipping and analysis cycle.

4 Analytical methods

4.1 Summary

Table 3 lists the trace species that are routinely measured on the flask samples.

Trace species	Method	Analytical precision[a]	WMO target[b]	Unit
CO_2	GC	0.07	0.1 (0.05 in SH)	ppm
CO	GC	0.8	2	ppb
CH_4	GC	1.4	2	ppb
N_2O	GC	0.17	0.1	ppb
H_2	GC	0.7/2.5[c]	2	ppb
SF_6	GC	0.03	0.02	ppt
$\delta^{13}C - CO_2$	MS	0.02	0.01	‰
$\delta^{18}O - CO_2$	MS	0.02	0.05	‰
O_2/N_2	MS	4	2	permeg
Ar/N_2	MS	8	—	permeg
$\delta^{13}C - CH_4$	MS	0.1	0.02	‰
$\delta^{2}H - CH_4$	MS	1	1	‰

Table 3: Summary of trace species measured routinely on flask samples at MPI-BGC. [a]Long-term reproducibility based on target gas QC measurements. [b]Network compatibility goal recommended by the WMO-GAW expert group (Crotwell et al., 2019). [c]Precision of PDD resp. RGA analysers.

4.2 GC measurements

4.2.1 Instrumentation

Gas chromatographic (GC) analysis of air samples in glass flasks is performed by a GC system that combines two Agilent 6890 gas chromatographs with several separation and detecting units. Gas Analysing Unit 1 (GAU1) is equipped with an Electron Capture Detector (anode-purged style, product number G1533A) for N_2O and a Flame Ionisation Detector (FID) for CH_4 and CO_2 analysis (see Figure 5). The second Gas Analysing Unit (GAU2) comprises an HgO-Reduction Gas Analyser (RGA3, Trace Analytical) for CO and H_2 analysis, an additional ECD for SF_6 analysis, and, since July 2012, a Pulsed Discharge Helium Ionisation Detector (PDD) (D-3-I-HP, Valco Instruments, Schenkon, Switzerland) for H_2 analysis (see Figure 6). A list of the specific detector and chromatographic settings is given in Table 4.

4.2.2 Chromatographic Method

For all of the individual chromatographic sub-units the course of an analysis is similar. The sample gas is first flushed through the sample loops attached to the coupled injection valves of each unit (ten-port, 2-position valves, ET6C10UWE, Valco Instuments) for 48 s at a flow of 70 ccm per minute set by a mass flow controller (MFC)(1179A, MKS Instruments) and then vented to the room (displayed in Figure 5 for GAU1). After sufficient time for the sample to equilibrate with atmospheric pressure (40 s) and GC oven temperature the sample is directed onto the first of two columns by switching the injection valves (depicted in Figure 6 for GAU2). Once the analytes of interest have quantitatively been transferred to the second column (main-column) these valves are switched back such that the flow through the first chromatographic column (pre-column) is reversed (back-flushed). This prevents that unwanted compounds accumulate on the column and enhance the detector baseline noise.

By this back-switching of the injection valves the sample loops are removed from the carrier gas stream back to the initial fill position. To avoid a possible contamination of the subsequent FID measurement by the methane containing carrier gas used for the ECD chromatography another 2-position valve is placed between the two injection valves of GAU1 ("loop separation valve" in Figure 5). Such the

Table 4: Configuration details of the Gas Analysing Units

Parameter	CH_4	CO_2	N_2O	CO	SF_6	H_2	H_2
Detection							
Detector	FID	Ni_{cat}-FID	ECD	HgO-R-UV	ECD	HgO-R-UV	PDD
Det.temp. [°C]	180	180 (Ni_{cat}: 360)	385	272	385	272	100
Det. auxiliary gas	fuel gas H_2; FID air	fuel gas H_2; FID air					discharge gas He 6.0
Flow rate [mL/min]	35 ; 220	35 ; 220					32
Bypass period [s]	-	0 - 174	0 - 138	-	0 - 144	-	-
Peak quantity	height	area	area	height	height	height	height
Response function	linear	quadratic	quadratic	cubic[1)]	quadratic	quadr. exp.	quadratic
Chromatography							
Sample loop [mL]	2.2	2.2	5	1	15	1	3
Pre-column	Hayesep Q 6ft x 1/8" 60-80 mesh	Hayesep Q 6ft x 1/8" 60-80 mesh	Hayesep Q 6ft x 1/8" 60-80 mesh	Unibeads 1S 2.5ft x 1/8" 60-80 mesh	Hayesep Q 6ft x 3/16" 80-100 mesh	Unibeads 1S 2.5ft x 1/8" 60-80 mesh	Hayesep Q 9ft x 1/8" 100-120 mesh
Main column	Porapak Q 12ft x 1/8" 50-80 mesh	Porapak Q 12ft x 1/8" 50-80 mesh	Porapak Q 6ft x 1/8" 50-80 mesh	Mole sieve 5A 2.5ft x 1/8" 60-80 mesh	Hayesep Q 12ft x 3/16" 80-100 mesh	Mole sieve 5A 2.5ft x 1/8" 60-80 mesh	Hayesep Q 9ft x 1/8" 80-100 mesh
Oven temp. [°C]	75	75	50	105	72	105	40
Carrier gas	He 5.0	He 5.0	Ar(95)-CH_4(5)	synthetic air	Ar(95)-CH_4(5)	synthetic air	He 6.0
Gas Purifier [2)]	HP2	HP2	EL	-	HCP	-	HP2
Flow rate [mL/min]	30	30	26	25	100	25	20
Flow control[3)]	EPC	EPC	EPC / MFC	EPC	EPC / MFC	EPC	EPC / MFC
Backflush switch [s]	234	234	138	78	163	78	131
Retention time [s]	155	230	335	98	330	46	210
Run time [s]	390	390	390	420	420	420	420

1) A cubic fit function is applied for the RGA response in the range of 80 - 500 ppb CO (covering the CO mole fractions common in background atmospheric air of the Northern Hemisphere). The detector non-linearity is particularly strong at low mole fractions, such that this function is not valid for lower levels. Samples with CO levels in the range of 20 - 110 ppb (mainly from the Southern Hemisphere or stratospheric air) are quantified applying a five point quadratic curve response based on secondary standards calibrated at the MPI-BGC Gaslab using a linear optical analyzer (Vacuum Ultraviolet Resonance Flourescence AL5002, Aerolaser, Garmisch-Partenkirchen, Germany).

2) HP2: Helium Purifier HP2 (Valco Instruments); Eliminator: ; HCP: High Capacity Purifier (Supelco-Sigma-Aldrich, Bellefonte, PA, USA); Eliminator E0040-CA-VR4-I-XL (NuPure, Ottawa, Canada);

3) EPC: Electronic Pressure Control; MFC: Mass Flow Control

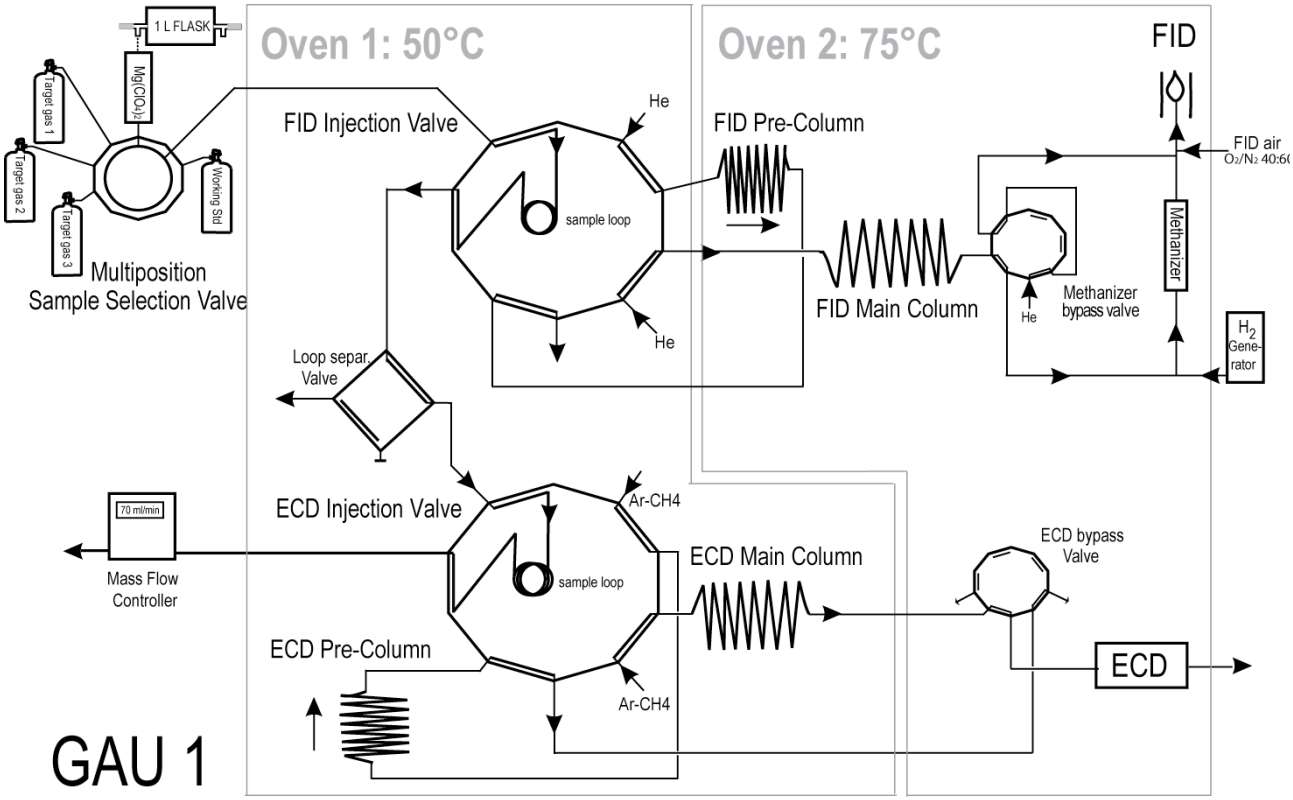


Figure 5: Gas flow schematics of Gas Analysing Unit 1(GAU1) in the fill = backflush position.

two sample loops can be disconnected after they have been filled and reconnected after the back-flush switch of the ECD injection valve just before the back-flush switch of the FID injection valve. This way, the pressurized argon-methane in the ECD sample loop is not expanding to the FID injection valve, and the over-pressurized helium from the FID sample loop blows out the argon-methane to the vent.

In order to make CO_2 detectable by the FID it is converted to methane on a heated Nickel catalyst after the chromatographic separation. To protect the hot catalyst from damage by oxygen another 2-position valve serves to bypass the major air compounds from this methanizer and leads them directly into the detector. The methanizer is put in line after elution of methane. Likewise, the sensitivity of the ECDs has turned out to slowly degrade when regularly being exposed to oxygen. Therefore, additional four-port 2-position valves are installed between the ECDs and the outlets of the respective main columns to bypass the main air peak from the detector to the vent. Once oxygen has eluted the bypass valves are switched and the column effluent is connected to the detectors. Although only four ports are used for these bypass valves 10-port valves are used, because the longer switching times of 4-port valves causes a larger disturbance of the stability of the FID flame and the ECD baseline, respectively.

Two independent chromatographic lines have been set up for N_2O (GAU1) and SF_6 (GAU2) to allow for individual chromatographic parameter optimization. Due to the very small atmospheric mole fraction of a few picomoles per mole the precision limit for the SF_6 analysis is mainly defined by the signal-to-noise ratio of the small chromatographic peak. A large sample loop volume and a quick separation resulting in a sharp and high peak hence improve analytical repeatability. These optimized conditions do not ensure a baseline separation of CO_2 from N_2O. Although CO_2 has a very weak (negative) response on the ECD it does have a cross-sensitivity to the N_2O detection. Parameters for the N_2O chromatography are therefore set such that there is no interference of CO_2 on the N_2O detection.

After the implementation of the PDD in routine operation in July 2012 H_2 data are generally generated using this detector due to its superior repeatability and linearity characteristics as compared to the

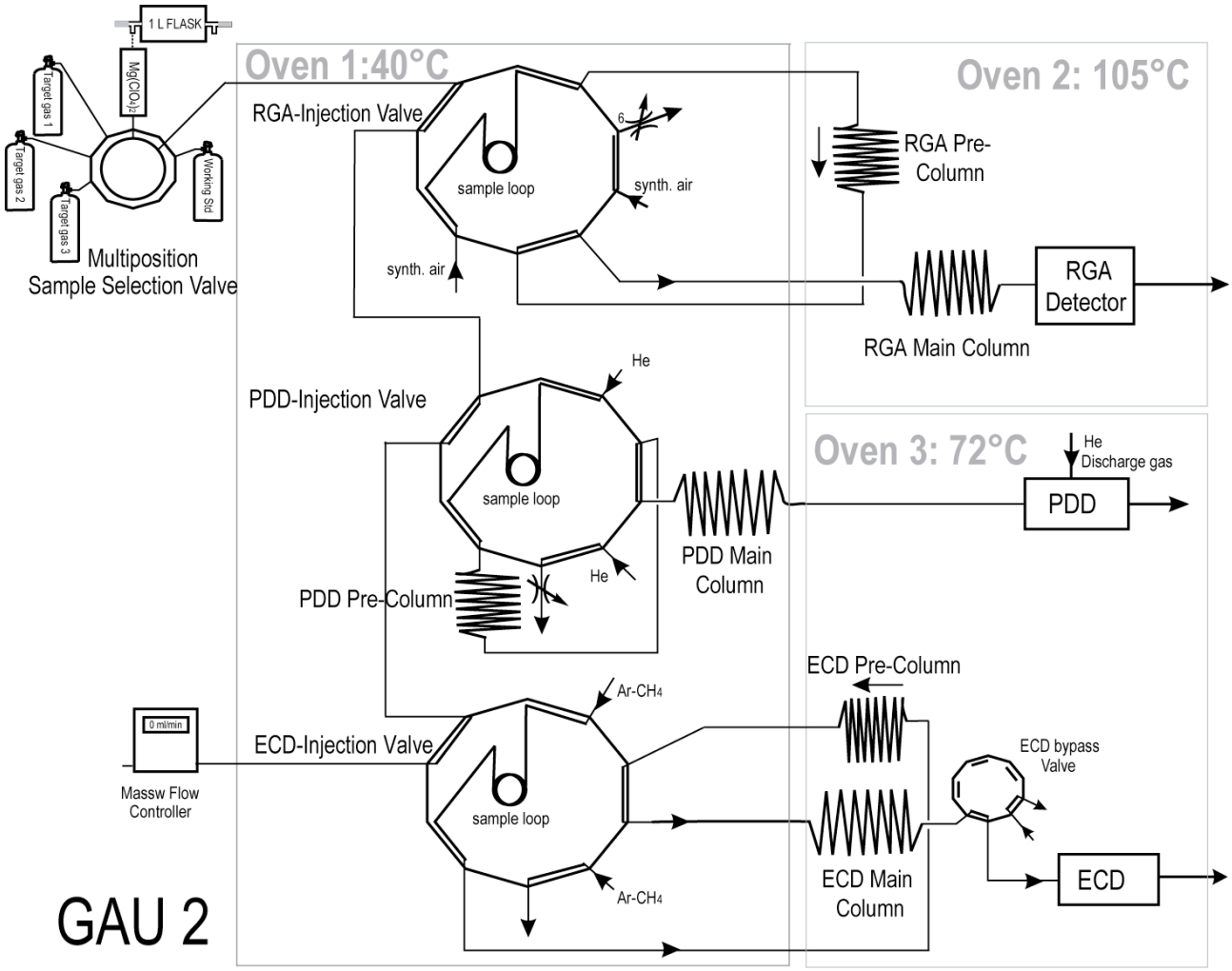


Figure 6: Gas flow schematics of Gas Analysing Unit 2(GAU2) in the inject position.

HgO-R-UV detector. During the measurement period of July 2019 to August 2020 the PDD was out of order and H_2 was again analysed using the HgO-R-UV detector.

Integration of all detector signals is done using the Chemstation software (Agilent Technologies, Santa Clara, CA, USA). Depending on the chromatographic peak shape, peak height and baseline noise for different compounds a quantification using either the peak area or peak height have proven to be more robust. The information on the peak dimension used for the respective tracers is specified in Table 4. Analog voltages from the RGA and the MFC are imported using an Analog-Digital-Converter (35900E, Agilent Technologies).

Sets of up to 15 flasks each can be connected to two flask selection valves (Valco EMT2C16UWE multi-position valves, Valco Instruments) using Ultratorr-connectors. The 16th valve position is attached to a target gas. Each of these flask selection valves is linked through a 2-position valve to the sample selection valves of either of the two Agilent 6890 GCs. A small glass tube (10 cm x 1/4") filled with anhydrous magnesium perchlorate (ca. 800mg) is placed before the GC valve to remove humidity from occasionally poorly dried samples.

One of the eight inlet ports of the GC sample selection valves is connected to the flask selection valves. The other ports are linked to natural air standards in high pressure cylinders equipped with high-purity, two-stage gas regulators (most common Tescom 64 series regulators; Scott Specialties are also used for H_2, CO and SF_6). One of these standards serves as working standard gas that is used to correct the measurements for short term drifts, three as target standards for quality control on different concentration levels. Three ports are not permanently tied to specific gases but used to mount either calibration gases or pressurized gas samples to be analyzed.

The two sets of 15 flasks each are measured in parallel on the two systems (GAU1 and GAU2). A regular analysis sequence consists of alternate measurements of the working standard and the flask samples. Atmospheric air samples are usually analyzed twice (with a time delay of about 4 hours). Once the first unit has finished analysis of the first set of flask samples it is paused until the second unit has completed its measurements of the other set of samples. Then the two-way-valve connecting the two flask selection valves to the respective GC units is switched to exchange the combination of the flask set and instrument and the analysis of the other set of samples is started on both units. The measurement sequences are completed by (generally triplicate) analysis of multiple target standard gases likewise bracketed by working standard analysis.

4.2.3 Instrument Calibration

Measurement data are calibrated relative to the current WMO Mole Fraction Scales for all reported gas mole fractions. The current scales that are maintained by NOAA GML as Central Calibration Laboratory (CCL) are: WMO CO_2 X2019, WMO CH_4 X2004A, WMO CO X2014A, WMO N_2O X2006A, WMO SF_6 X2014. The traceability to these scales is realized by suites of high pressure standard gases calibrated by the WMO Central Calibration Laboratory and is actively maintained by regular re-calibrations of subgroups of these laboratory calibration standard sets. The respective standard cylinders are given in Table 5 with the currently used assigned values. These values may change due to scale revisions by the Central Calibration Laboratory or additional measurements done by the CCL. The updated data are available in the internet online.

The WMO H_2 X2009 Mole Fraction Scale is maintained by MPI-BGC on the same instrument that is used to analyze flask samples such that these measurements are calibrated directly relative to the primary scale.

4.2.4 Data evaluation

A regular analysis sequence consists of alternate measurements of the working standard and either flask samples or target standards that are used for quality control assessment. Raw data of any sample measurement is normalized to the raw data of the working standard to cancel out drifts due to variations in atmospheric pressure or other laboratory environment variations. The detector response function

CylinderID	Fill Date	CO2 [ppm] X2019[1]	CO [ppb] X2014A[2]	CH4 [ppb] X2004A[1]	N2O [ppb] X2006A[3]	SF6 [ppt] X2014[3]	Last CCL Analysis
CA01601	12.1999	366.54	163.7	1727.4	314.46	4.32	2016
CA01650	12.1999	411.58	86.3	2099.1	325.12	3.81	2021
CA01675	12.1999	457.01	245.9	2537.6	333.64	4.31	2021
CA01680	12.1999	517.12	492.7	3055.9	343.90	4.31	2017
CA01794	09.2011		358.7				2011
CA04605	03.2001	351.39	111.7	1692.01)	305.42	4.57	2018
CA04611	03.2001	382.47	210.4	1844.4	316.78	4.91	2017
CA04639	03.2001	422.28	401.0	2030.61)	316.49	5.02	2021
CA05281	11.2005				339.22	6.82	2007
CA05290	11.2005				321.80	7.32	2007
CA05295	11.2005				305.15	16.46	2007
CA05435	06.2007				328.68	6.40	2007
CA05863	11.2003				315.77	8.15	2007
CA06720	11.2005	440.32	103.2	1803.2	318.72	5.83	2017
CC121969	11.2005	404.45	100.0	1812.0	319.86	5.84	2018
CC339513	01.2011	468.14	486.8	2173.5	328.03	7.26	2019

Table 5: Calibration standards assigned by the WMO Central Calibration Laboratory. Mole fraction data on the specified WMO scale taken from NOAA-GML, as of June 2021.
[1]CO2 and CH4 results assignments by PC1 instrument only, if available.
[2]last assigned CO value (most standards exhibit growth of CO).
[3]2001 CCL assignments with high uncertainty are not considered.

and the mole fractions of the various trace species in the working standard are determined by analysis of a suite of laboratory standard gases assigned by the WMO Central Calibration Laboratory (see Table 5). Measurements of these highest level laboratory calibration standards are generally repeated every second month to capture small changes of detector responses or in cases where quality control measurements suggest such sudden changes.

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

- chromatographic parameters (retention time; baseline level, drift, and noise; raw peak magnitude) of the sample and the working standard chromatograms,
- the sample flushing volume registered by the MFC (imported by the Chemstation software),
- the repeatability of the related working standard raw data relative to those of the preceding and subsequent working standard measurements,
- the measurement results of the target standards relative to their known composition.

Measurements are flagged invalid in cases where chromatographic variables indicate a system malfunctioning or if the MFC output points to insufficient flushing of the sample loop. In cases of an invalid working standard measurement it is checked if this individual reference point can be replaced by the next working standard measurement result or if this failed standard measurement indicates that the sample measurement also has to be flagged invalid.

In cases of invalid flask sample measurements or if the pair analysis agreement is worse than expected from the typical instrument performance a third analysis is usually made (provided the remaining sample pressure is sufficiently high).

The typical long-term reproducibilities obtained from target gas measurement results are listed in Table 3.

4.3 MS measurements

Stable isotope ratios of CO_2 ($\delta^{13}C-CO_2$ and $\delta^{18}O-CO_2$) and of CH_4 ($\delta^{13}C-CH_4$ and $\delta^{2}H$) in air samples as well as the O_2/N_2 and Ar/N_2 ratios are routinely analyzed by mass spectrometry. Automated sample preparation systems are coupled to commercially available isotope ratio mass spectrometers. All systems were built and assembled at MPI-BGC. The technique for CO_2 isotopes is in operation

since October 2000. The analysis of O_2/N_2 and Ar/N_2 ratios was added in 2005. The rather complex methane system started routine flask analysis in 2012.

4.3.1 Instruments

4.3.1.1 CO_2-in-air isotopic analysis For air-CO_2 isotopic analysis, the MPI-BGC Stable Isotope Laboratory ("BGC IsoLab") has developed a dedicated, fully automated cryogenic extraction line ("BGC Airtrap"), coupled to the dual inlet system of a Finnigan MAT 252 stable isotope ratio mass spectrometer (IRMS) available from Thermo-Fisher Scientific, Bremen, Germany (Werner et al., 2001). In order to cope with the number of sample analysis requests, a second BGC Airtrap with MAT 252 system was assembled and taken into routine operation in early 2002.

The systems can perform 18 complete analysis runs each in a daily sequence. Four of the extracted samples are working reference air from a high pressure cylinder, and one is a QA reference air. The remaining 13 analyses are made from the standard BGC 1-L flask samples. The isotopic analysis is the last in line, following the O_2/N_2 and trace gas quantification measurements. At the end of these analyses, the pressure in the flasks is usually down to about one atmosphere. For a single analysis, 600 Bar mL of this remaining air amount is consumed during extraction.

A number of corrections are applied to the raw mass spectrometric data in order to arrive at the final calibrated delta-values on the international JRAS-06 (Ghosh et al., 2005; Brand et al., 2009; Wendeberg et al., 2013) scale. These include correction for the isobaric interference from N_2O in the air samples (Ghosh and Brand, 2004), the correction for the ^{17}O contribution to the m/z 45 and 46 ion currents (Assonov and Brenninkmeijer, 2003a,b; Brand et al., 2010), and corrections for the capillary crimp mismatch and working reference gas drift (Werner et al., 2001). Most importantly, the mass spectrometers are operated at reduced electron emissions and, thus, sensitivity, resulting in a negligible crosstalk of the gases upon gas exchange ("eta-effect"; (Ghosh et al., 2005; Meijer et al., 2000; Verkouteren et al., 2003a,b)).

The daily QA measurements allow monitoring the overall analytical performance over time. The two mass spectrometric systems are administered with the same working reference (WR, isotopic composition near ambient air) and quality assurance (QA, isotopic composition offset from ambient) gases. Quality assurance measurements are conducted in each sample analysis sequence and now cover a period of 21 years. The average long-term measurement precision is 0.02‰ for both, $\delta^{13}C - CO_2$ and $\delta^{18}O - CO_2$ measurements. While in principle the reproducibility is closely similar for both, one instrument is measuring soil air samples frequently, which often contain organic compounds like ethanol that interfere with the measurements. Without this effect, $\delta^{13}C - CO_2$ has been analyzed with a reproducibility of 0.014‰. Without adjusting one instrument to the other, both instruments measure the same isotopic differences (within a few permeg) between the WR and QA gases.

4.3.1.2 Methane stable isotope analysis A dedicated analysis system called iSAAC ("Integrated System for Analyzing Air Constituents") has been developed for measuring the two stable isotope ratios of methane simultaneously (plus those of CO_2) in flask samples (Brand et al., 2016). The system is comprised of

- a sample carousel with 16 positions for flasks or cylinders,
- a 1st dual-loop 10-port valve (VICI AG, Schenkon, CH) for administering 1-mL sample aliquots to a GC-IRMS (with sub-ambient Poraplot Q column) for measuring the CO_2 isotopes
- a 2nd dual loop 10-port valve for pre-concentrating methane at -130°C (generated using a PCC "Cryotiger" compressor cooler from Brooks Automation Chelmsford, MA, USA) from the sample streams for $\delta^{13}C$ analysis on the same IRMS
- a 3rd dual-loop 10-port valve for pre-concentrating methane at -130°C from the sample streams for $\delta^{2}H$ analysis on a second IRMS dedicated to hydrogen measurements

The analytical conduits from the 2nd and 3rd valve include one cryogenic focus-trap, each, and a GC column (Shincarbon, 0.5 mm id, Restek, Bellfonte, PA, USA) operated isothermally at 90°C inside a Trace-GC (Thermo-Fisher, Bremen). CH_4-peaks are heart cut automatically and reacted on-line to CO_2 and H_2, respectively, using established irm-GCMS techniques (Brand et al., 1996; Hilkert et al., 1999). Following conversion, the GC effluents are subjected to a last GC separation, thereby preventing co-elution of Krypton co-trapped from air (Bock et al., 2014; Schupbach et al., 2009; Schmitt et al., 2013) or residual CH_4 (Merritt et al., 1995). The transfer to the respective mass spectrometer is made with an open split (Merritt et al., 1994), modified for efficient exclusion of eluting air components other than the selected peaks. The open split for CO_2 transfer can receive effluent helium from either the first valve (CO_2-line) or from the 2nd valve (methane δ^{13}C-line). The system design includes full calibration for every sample by alternating acquisition from the sample and from a calibrated reference cylinder. A complete analysis including measurement of CO_2 ($\delta^{13}C - CO_2$ and $\delta^{18}O - CO_2$) and CH_4 ($\delta^{13}C - CH_4$ and δ^2H) from sample and reference gas takes about 90 minutes. An entire sequence of 14 samples and 5 reference extractions is completed after about 28 hours. Within every sequence, one QA reference material is analysed. Since the start of the measurements the QA records show a reproducibility of 0.05 and 0.058‰ for $\delta^{13}C - CO_2$ and $\delta^{18}O - CO_2$, respectively, and 0.1 and 1.2‰ for $\delta^{13}C - CH_4$ and δ^2H, respectively. Although these are fully calibrated values, including the daily mean reference analysis, the $\delta^{13}C - CH_4$ values for methane still need to be improved.

4.3.1.3 O_2/N_2 and Ar/N_2 measurement system The quantification of O_2/N_2 and Ar/N_2 ratios in air samples is made on a dedicated mass spectrometer (DeltaPLUS XL, Thermo-Fisher, Bremen, Germany) (Brand, 2005; Werner and Brand, 2001). Inlet of air samples in flasks or in high-pressure cylinders is made from a home-built auto sampler system with 16 ports. Some of the ports are permanently connected to reference air cylinders; the bulk is available for flasks. Gas transfer to the mass spectrometer is made via an open split, consisting of a small i.d. glass tube closed on one side. The open side has a 10μ/60 cm fused silica conduit to the mass spectrometer inlet, which permanently admits ambient gas to the ion source. The flow results in an operating pressure of $\sim 2 \cdot 10^{-7}$ mbar. In order to cope with this deliberately high load of air, the ion source emission is reduced to 0.2 mA electron current. Two fused silica capillaries transport the sample and reference air, respectively, to the sniffing point of the open split tube; either one is inside while the other is retracted. The air flow through these capillaries is ∼2 mL/min, which is sufficient to keep ambient air out and to exchange the gas volume within seconds. Gas change in the mass spectrometer takes ∼6 seconds, governed mainly by the transfer time through the flux-limiting capillary to the mass spectrometer. Although no dual inlet system is needed, the mass spectrometric measurement is made in a dual-inlet fashion, a standard technique for stable isotopes which is capable of very high precision comparisons of gas mixtures as well, as long as these are very close in composition. This is the case for air samples, where O_2/N_2 and Ar/N_2 ratios vary much less than 0.1% over a decade or from pole to pole.

Special attention has been given to the overall temperature sensitivity of any O_2/N_2 and Ar/N_2 quantification system for air samples. This sensitivity arises from the requirement for measurement reproducibility in the 10^{-6} region, the results are communicated in "permeg" units (Keeling and Shertz, 1992). At this level, fractionation effects arising from thermal diffusion can already produce a difference of 6 permeg owing to a small temperature difference of 0.1 °C for instance across a high pressure cylinder (Keeling et al., 2007). Therefore, the long-term temperature stability of the entire laboratory was improved to $\sim \pm 0.4$ °C, with the larger variation in the diurnal cycle. To further dampen the influence of the remaining temperature variability the reference cylinders are encased in a wooden box and after mounting to the carousel the entire set of sample flasks are enclosed by a large foam cylinder before and during analysis. Cylinders for analysis as well as cylinders defining the local calibration scales (see Section 4.3.2) are positioned horizontally on a shelf with encasement as well.

4.3.2 Calibration scales

4.3.2.1 CO_2 isotopes The calibration of the CO_2 isotopes in air has been described in detail previously (Brand et al., 2009; Ghosh et al., 2005; Wendeberg et al., 2011, 2013). In short: the scale

anchor is generated by mixing CO_2 from reference carbonates (NBS 19, MAR-J1) into CO_2-free air. The $\delta^{13}C - CO_2$ reproducibility of the procedure has been established at the 0.01‰ level, that of $\delta^{18}O - CO_2$ is at 0.02‰. The scale has been termed "JRAS-06" (for "Jena Reference Air Set"); it has been recognised as the reference scale for such measurement by the WMO, designating the BGC IsoLab as the Central Calibration Laboratory for CO_2-in-air isotopic measurements. Transfer of the scale is achieved by sending 5-L glass flasks filled with calibrated air to interested laboratories. Moreover, calibration of cylinders with dry air is available. Scale offsets over a wider range of delta values between different laboratories have been studied over time and made public (Wendeberg et al., 2013).

4.3.2.2 Methane isotopes The methane isotope scales have initially been anchored at BGC IsoLab by measurements of two air cylinders made at the University of Utrecht. This laboratory is well "inter-calibrated" with other laboratories involved in such measurements. However, there is a well-studied $\delta^{13}C - CH_4$ scale discrepancy, as revealed by comparing measurements from the same or nearby locations (Levin et al., 2012), which amounts up to about 0.3‰ between laboratories. This is of the same order of magnitude as the differences to be studied (e.g. the inter-hemispheric difference and its development over time (Kai et al., 2011). Similar inter-laboratory offsets due to isotope scaling variations have been found for almost all laboratories involved with measurements of stable isotopes of methane (Umezawa et al., 2018). At BGC-IsoLab, we therefore have investigated the relation between different methane gases in high-pressure cylinders and the primary isotope scales, VPDB and VSMOW (Sperlich et al., 2016, 2021). In short, comparison for $\delta^{13}C - CH_4$ of methane has been made by reacting carbonates (IAEA-CO-9) and the CH_4 gases in the same oxidation device (EA-ConFlo-IRMS (Werner et al., 1999)) and thus establishing a "true" $\delta^{13}C$ value for the gases. In a second step, mixtures of pure CH_4-free air and all gases were made using the same system as that for mixing CO_2. These gases were analyzed on the iSAAC system, resulting in a scale comparison between IMAU and the newly established scale anchor. In a similar fashion, water samples precisely calibrated on the VSMOW scale were used on a TC/EA-ConFlo-IRMS to compare their δ^2H with those from the co-injected CH_4 gases. The results were also compared with those from the iSAAC system. Thus, firm offsets were found between the IMAU and the new scale anchor for δ^2H. The established offsets were +0.066‰ for $\delta^{13}C - CH_4$ and +3.97‰ for δ^2H (with IMAU positive versus the new scale anchors). Thus methane isotope measurements are reported on the JRAS-M16 (Methane) scale that has the above stated offsets to the IMAU scale.

4.3.2.3 O_2/N_2 and Ar/N_2 ratios The O_2/N_2 scale at BGC Isolab is currently still a local implementation of the Scripps O_2/N_2 "S2" scale (Keeling et al., 2007). The basis for the local implementation is a set of three 30-L high pressure cylinders obtained from Scripps in 2007. Prior to 2007 the scale has been covered by a set of two tanks, also obtained from SIO. These Scripps Tanks are used to anchor working standard tanks and have periodically been analysed as samples (nearly 200 times since 2007) for QA puproses. These measurements have yielded an average standard deviation of 2 permeg. The transition to the new SIO2017 O_2/N_2 scale is currently in preparation and is expected to be completed by the end of the year 2022.

For Ar/N_2 measurements, a common scale does not exist (yet). We have gotten informative values for three tanks from Scripps. Based on these, the local BGC "scale", as realised by the same working standards that are used for O_2/N_2 measurements, has an offset of +126.3 permeg with a reproducibility of 5.6 permeg. No further attempts have been made to unify Ar/N_2 scales, neither between Scripps and BGC nor between other laboratories. This is probably due to the low annual variability ($\sim\pm5$ permeg) of this ratio as well as the larger sensitivity towards mass related fractionation, which is about 2.5 times larger than that of O_2/N_2. We mainly use the measured Ar/N_2 values for correcting the raw O_2/N_2 data for fractionation effects during sampling and storage (see Appendix A).

5 Postprocessing and automatic flagging

Postprocessing of the results from the flask measurements is performed in three steps. In step 1 the flask analysis results reported from GASLAB and ISOLAB are merged, a series of consistency checks performed, station codes unified and derived quantities are computed. In step 2 each measurement undergoes a series of tests yielding a numerical quality flag (0-9). In step 3 flask replicate measurements are averaged and, together with valid single flask measurements are reported for each station and each trace species.

The flowchart of this semi-automatic processing is schematically shown in Figure 7. The three processing steps are described in more detail in the following subsections.

5.1 Step 1: Merging of measurement databases and initial consistency checks

5.1.1 Merging

GASLAB and ISOLAB measurements are reported in separate excel tables on the institute server. In step 1 these tables are merged based on the unique sample number (USN) that is assigned to each flask sample when the analysis request is registered in the ISOLAB database.

5.1.2 Initial consistency checks

Subsequently, a series of initial checks are performed:

- USN: each distinct USN should appear only once in each of the two measurement tables.
- FlaskID: for any USN the corresponding flaskID should be the same in both tables.
- Sampling time: for any USN the recorded sampling time should be the same in both tables.
- Sampling location: for any USN the recorded sampling location should be the same in both tables.
- Sampling height: for any USN the recorded sampling height should be the same in both tables.

5.1.3 Assignment of station codes

Since the GASLAB and ISOLAB tables contain a different column layout of the measurement results including also intermediate values, a merged and cleaned master table is constructed by extracting for each USN the correct values out of the individual tables for each trace species.

Over time different flask operators have used slightly different sampling location descriptions. Therefore a unified station code has been defined for each sampling location and added to each valid USN record. The mapping of the location descriptions to the unified station codes is given in in Table 6. Some of the site descriptions refer to flask measurement tests which are retained in the database; these are indicated with a station code XXXn and not further processed. There exist also flask measurements from mobile sampling campaigns, such as from the flight program at Bialystok (BIKF) or the flask sampling on the Meteor cruise near the Namibia station (NAMS). Flask measurements from additional short term sampling campaigns exist from Amam. KJA denotes flasks prepared for the calibration of the in situ measurement system in Dikson, Russia. SIS0 indicates early measurements on flasks sampled at the original Shetland-Burra site, before the sampling location was moved to the Sumburgh Head location 26km further south (see section C.5).

5.1.4 APO calculation

If both CO_2 and O_2/N_2 measurements are available for a particular flask, the derived species APO ("Atmospheric Potential Oxygen", (Stephens et al., 1998)) is calculated according to the formula:

$$\mathrm{APO} = \delta \mathrm{O_2/N_2} + 1.1(\mathrm{CO_2} - 350\mathrm{ppm})/0.20964 \quad (1)$$

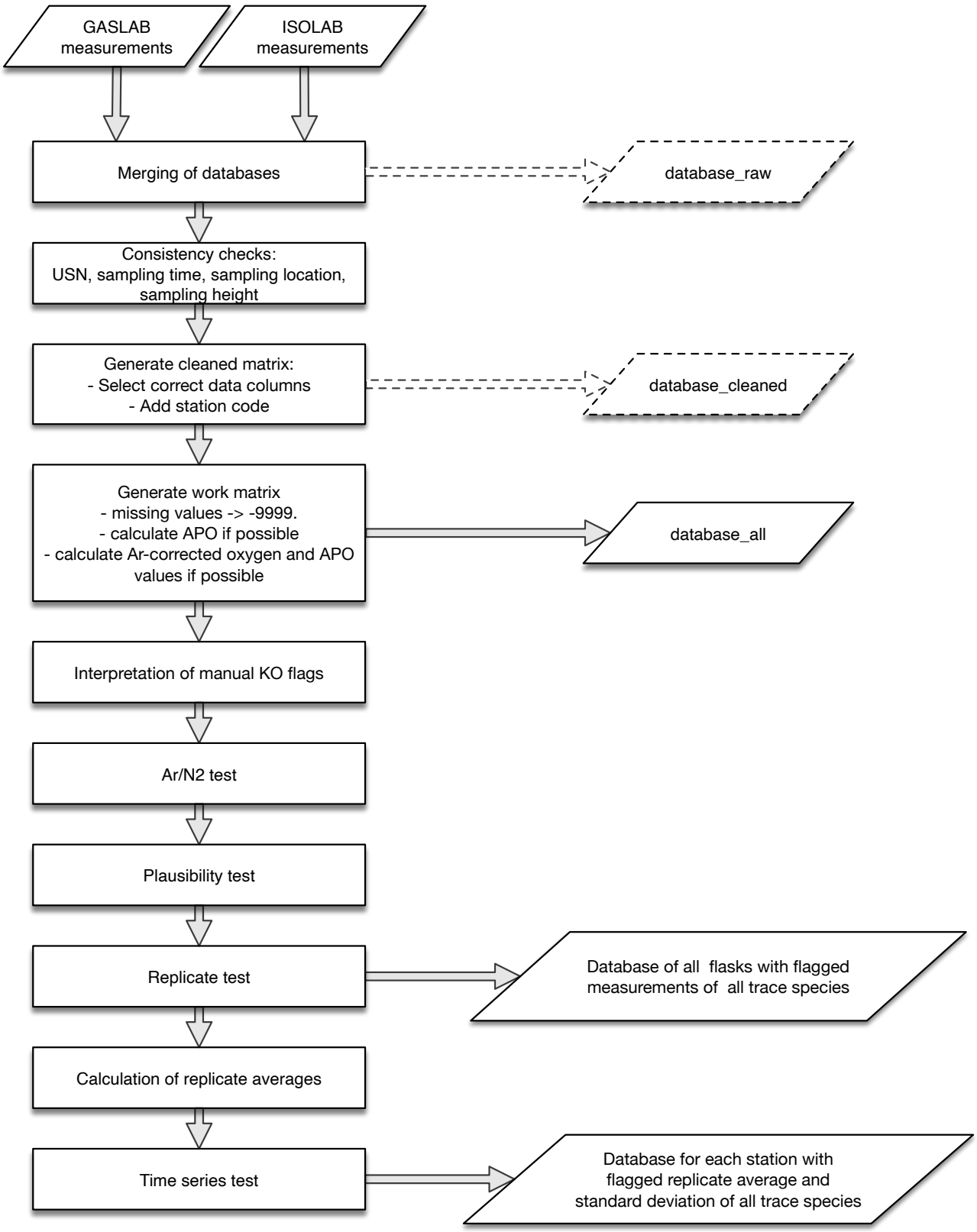


Figure 7: Postprocessing flowchart. Left column: processing steps, right column: output database files. Dashed items indicate intermediate files generated for checking purposes.

Site description in flask sampling data sheet	Stationcode	No of flasks
Alert	ALT	1310
Amam	AMA	26
Bialystok	BIK	715
Bialystok (Base)	BIK0	75
Bialystok flight	BIKF	9
Bialystok flight area	BIKF	213
Bialystok Tall Tower	BIK	77
cape grim	CGO	12
Cape Grim	CGO	155
Cape Verde	CVO	1772
Jungfraujoch	JFJ	652
KJA	KJA	12
Kjölnes	KJN	86
Kjölnes Norway	KJN	276
Namibia	NAM	549
Namibia Meteor	NAMS	197
Neumayer Station	GVN	224
Ochsenkopf	OXK	2201
Shetland-Burra	SIS0	33
Shetland Isles	SIS	1356
Shetland Isles Station	SIS	1124
Sphynx, Air sampling point	JFJ	234
Station North, Greenland	VRS	87
Zotino Tall Tower Observatory (ZOTTO)	ZOT	631
Zotto	ZOT	789
ZOTTO	ZOT	659

Table 6: Assignment of site description to unified station codes.

This yields APO in units of permeg. Thereby we neglect the small contribution from CH_4 and CO in the original APO definition of Stephens et al. (1998).

5.1.5 Ar-correction of O_2/N_2 and APO measurements

In the case of O_2/N_2, the spread of flask replicate measurements was found to be substantially larger than the measurement repeatability, and the correlation between replicate differences and the corresponding Ar/N_2 ratio is highly significant (see Appendix A). Because of the sometimes long storage times of the pressurised flasks ($\sim$ 1.6 bar) prior to analysis at MPI-BGC, effusion of air from the flask through micro-leaks at the stopcocks would fractionate the trace species in the remaining air sample. Indeed, the observed slope is 0.4 corresponding to the ratio of the theoretical fractionation factor for O_2/N_2 and Ar/N_2 by Knudsen diffusion (Langenfelds et al., 2005).

In this case the measured Ar/N_2 ratio (δAr) can be used to derive an Ar-corrected O_2/N_2 ratio (δO^c):

$$\delta O_2^c = \delta O_2 - g(\delta Ar - \delta Ar_0) \quad (2)$$

where the constant $g = 0.4$ is the theoretical fractionation factor mentioned above.

For the constant δAr_0 we use a spatially and temporally constant value of 140 permeg, corresponding to the median of all flask measurements from all station except GVN. The assumption of an approximately uniform constant atmospheric Ar/N_2 ratio is based on the fact, that natural Ar/N_2 variations are only driven by changes in ocean heat content. The observed seasonal cycle is smaller than 20 permeg, latitudinal gradients are smaller than 20 permeg (Battle et al. (2003), Keeling et al. (2004)) and the expected long-term trend due to global ocean warming is on the order of 0.3 permeg/yr. These natural variations are smaller than our measurement precision.

A special case was found at the flasks from the Neumayer station (GVN) in Antarctica: At this station the median Ar/N_2 value is -100 permeg, substantially lower than at the other stations. Most likely this indicates a fractionation effect during sampling, possibly caused by the relatively low local temperatures (see also Battle et al. (2003)). Hence for the GVN measurements, the base value for the Ar-correction was chosen as -100 permeg.

The Ar-corrected O_2/N_2 ratio is denoted by $O_2/N_2{}^c$. Using this value also an Ar-corrected APO value (denoted by APO^c) is calculated (using equation 1). Ar-corrected $O_2/N_2{}^c$ and APO^c values of flask replicates exhibit substantially smaller spread compared to the uncorrected values (Appendix A). The Ar-corrected $O_2/N_2{}^c$ and APO^c values are retained as separate trace species in the subsequent postprocessing and flagging tests.

5.2 Step 2: Flagging of individual flask measurements

Prior to the flagging tests missing values are indicated with the value -9999. and indicated with the numerical flag = 9.

5.2.1 Incorporation of manual KO flags

Based on information on the sampling logsheets and/or abnormal values on some of the trace species, sample collection problems or irretrievable sample storage effects, artefacts in some or all trace species measurements of a flask can be identified. These are indicated as manual k.o. flags in the comment field of the GASLAB data sheet with the text string "KO=xxx;". Thereby the letters in the string xxx denote one or several of the trace species whose measurements are judged as suspicious and not valid. Table 7 lists the letters and the corresponding trace species.

If one or several manual KO flag letters are identified in the comment field, the corresponding trace species values are marked with the numerical flag = 8.

KO flag letter	Suspect trace species measurement
A	Ar/N_2
B	O_2/N_2
C	CH_4
D	CO_2
E	N_2O
F	H_2
G	CO
H	SF_6
I	$\delta^{13}C - CO_2$
J	$\delta^{18}O - CO_2$
K	$\delta^{13}C - CH_4$
L	$\delta^{2}H - CH_4$
Y	all trace species measurements suspect

Table 7: Manual KO flag single letter indicators.

5.2.2 Ar/N_2 test

Similar to the correlation between O_2/N_2 and Ar/N_2 of differences of replicate flask measurements also the outlier differences of all other trace species except CO were found to be linearly related to the corresponding Ar/N_2 measurements. The empirically found slopes for all species except CO agree very well with the theoretical relationships assuming Knudsen diffusion (Langenfelds et al., 2005). Because the linear relationship holds only for outliers, it cannot be used for a proper correction procedure but it provides a means to diagnose outliers based on the corresponding Ar/N_2 measurement (see Appendix A).

For each flask for which a valid Ar/N_2 measurement is available (i.e. for which the Ar/N_2 measurement lies within the plausibility range, see below), each of the trace species measurements are flagged with the numerical flag = 7 if the Ar/N_2 measurement is higher resp. lower a base value ± the corresponding Ar-criterion listed in Table 8.

For all stations except GVN we choose a nominal base value of 140 permeg, which is very close to the median of all Ar/N_2 flask measurements in the entire flask database. For flasks from the GVN station we choose a base value of -100 permeg (see above).

5.2.3 Plausibility test

All flask measurements are furthermore scanned for non-plausible concentration values. These flask measurement entries are flagged with the numerical value = 6. The plausible concentration ranges employed as criterion for each trace species are listed in Table 8.

5.2.4 Replicate test

Replicate samples are identified as flasks sampled at the same location concurrently or sequentially within less than two hours. For each replicate group the median c_{med} of all samples of the group is computed. Then for each flask sample of the group (c_i), the positive difference from the median $d_i = \|c_i - c_{med}\|$ is analysed: if d_i is larger than a given species specific criterium d_{crit}, the flask sample is flagged as not having passed the replicate criterion. If there is only one flask measurement of the group with $d_i < d_{crit}$ remaining, it is also flagged as not passing the replicate criterion. (The latter case may occur if e.g. three flask values are spaced wider than d_{crit}. In this case the middle value representing the median would remain as single value with $d_i = 0 < d_{crit}$.)

The criteria values d_{crit} for each of the 15 trace species are listed in Table 8.

5.2.5 Numerical values of the criteria employed in the tests

As basis for the criteria used in the different tests we define a "BGC base precision" by selecting for each trace species the higher value of analytical reproducibility and the WMO target as listed in Table 3. This base precision is used to calculate the respective Ar-criterion as described in Appendix A. For the replicate test we use as value of d_{crit} three times the BGC base precision. Table 8 lists the base precision and the numerical criteria selected for the different tests and trace species.

Trace species	Units	Base precision	Ar-criterion (permeg)	d_{crit}	Plausibility range	
CO_2	ppm	0.1	220	0.3	340	1500
CO	ppb	2	10000	6	10	3000
CH_4	ppb	2	520	6	1700	4000
N_2O	ppb	0.17	470	0.5	315	360
H_2	ppb	2	240	6	100	1500
SF_6	ppt	0.03	1500	0.1	4.5	25
$\delta^{13}C - CO_2$	permil	0.02	360	0.06	-20	-5
$\delta^{18}O - CO_2$	permil	0.05	460	0.15	-15	5
O_2/N_2	permeg	4	10000	12	-5000	0
Ar/N_2	permeg	8	10000	24	-1000	5000
$O_2/N_2{}^c$	permeg	4	10000	12	-5000	0
APO	permeg	4	10000	12	-2000	0
APO^c	permeg	4	10000	12	-2000	0
$\delta^{13}C - CH_4$	permil	0.1	410	0.3	-60	-30
$\delta^{2}H - CH_4$	permil	1	4100	3	-150	20

Table 8: "BGC base precision" and the criteria selected for the Ar/N_2, the replicate and the plausibility test.

5.2.6 Summary of numerical flags assigned to trace species values of individual flasks

Table 9 lists the numerical flags assigned to each trace species measurement of an individual flask. These flags are included in the complete flagged flask database (see Appendix B).

Numerical Flag	Description
0	good flask measurement, passed all tests
1	single flask measurement - passed all tests
2	flask measurement failed replicate test but passed Ar/N_2 and plausibility test
6	flask measurement failed plausibility test
7	flask measurement failed Ar/N_2 test
8	flask measurement judged suspect as indicated by manual KO flag
9	no measurement performed

Table 9: Summary description of the numerical flags

5.3 Step 3: Calculation of flask averages and time series test

In the third step possible outliers indicating possible local effects are identified using an iterative data fitting procedure performed for all species on all flask measurements from a single sampling station.

1. Selection of flask replicate measurements with flags 0 or 1.
2. Computation of mean and standard deviation for each replicate group. In case of only one available measurement, an uncertainty value of three times the base precision criterion is assigned.

3. A smooth curve consisting of a seasonal cycle component superimposed on a longer-term trend is fit to the replicate means. The seasonal cycle consists of a series of harmonics to the base period of 1 year; the number of harmonics depends on the location of the flask sampling station and the species under consideration. The trend is represented with a spline curve with specified 50% attenuation temporal frequency (typically 1/1.5 – 1/2 years^{-1}) (Enting, 1987).
4. Calculation of the root mean square (RMS) of the residuals from the fit.
5. If the residual of a single flask replicate mean is larger than 3 times the RMS, it is considered an outlier and discarded in the next fitting iteration.
6. Steps 3–5 are repeated iteratively until no further outlier is identified. In practice this takes typically 4–6 iterations.
7. Finally, all replicate means for the particular trace species and station are sorted in time.

Based on this time series test, the flask replicate averages are assigned flags 0 or 1:

Flag	Description
0	good flask replicate average
1	good flask replicate average but failed time series test, possibly not representative for regional scale background

The resulting data files for all station are available as described in Appendix B.

6 Quality assessment

6.1 Overall flask measurement statistics

Table 10 contains an overview of the number of measurements of the different species on individual flasks and associated flags after step 1. Table 11 shows a similar summary for the number of flask replicate averages after the time series test.

Flag:	0	1	2	6	7	8	9	0+1	Total	F (%)
CO_2	10730	635	692	0	419	485	572	11365	12961	88.
CO	9859	829	293	4	0	1899	649	10688	12884	83.
CH_4	11458	579	238	0	217	466	575	12037	12958	93.
N_2O	9301	753	59	0	206	2602	612	10054	12921	78.
H_2	9082	738	159	0	343	2607	604	9820	12929	76.
SF_6	11631	551	182	0	46	401	722	12182	12811	95.
$\delta^{13}C - CO_2$	9584	456	552	26	269	419	2227	10040	11306	89.
$\delta^{18}O - CO_2$	8952	445	1224	42	225	419	2226	9397	11307	83.
O_2/N_2	8468	388	1359	116	0	445	2757	8856	10776	82.
Ar/N_2	8194	364	1718	31	0	470	2756	8558	10777	79.
$O_2/N_2{}^c$	9377	358	550	7	0	470	2771	9735	10762	90.
APO	8380	419	1111	171	0	525	2927	8799	10606	83.
APO^c	9289	376	415	1	0	525	2927	9665	10606	91.
$\delta^{13}C - CH_4$	4257	727	213	0	113	652	7571	4984	5962	84.
$\delta^{2}H - CH_4$	4708	721	133	2	2	394	7573	5429	5960	91.

Table 10: Statistics of all flask measurements for the respective species and associated flags. Columns 0–9 show the numbers of measurement entries with the respective flag. Column 0+1 lists the "valid" measurements (i.d. measurements with flag 0 or 1). Column Total lists the total number of available measurements. Column F lists the percentage of "valid" measurements with respect to the total number of measurements of each species.

Table 10 shows that for most species over 80 % of all valid flask measurements contain flag 0, indicating that they passed the tests described in Section 5. The oxygen-based species (O_2/N_2 and APO) improve significantly after the applied Ar-correction as seen in the statistics for $O_2/N_2{}^c$ and APO^c. The somewhat lower percentage in valid N_2O and H_2 flasks is due to specific contamination problems that occurred for extended periods at some flask sampling stations. E.g. there is a known N_2O problem in air drying with a magnesiumperchlorate product that had been used for several years (2011-2014), and there is a problem in H_2 if the inlet system exhibits corrosion as identified at the CVO und SIS stations.

6.2 Flask turnover times

Figure 8 shows the statistics of the time difference between flask sampling and flask analysis for the different stations. In the case of logistically more challenging stations (NAM, CVO, ZOT, GVN), a substantial fraction of flasks are analysed later than 180 days after sampling. In particular at ZOT a substantial fraction of flasks has delay times exceeding one year, which was caused by a temporary flask export ban by Russian authorities.

6.3 Flask replicate residuals

Having flask replicates provides an important quality control assessment opportunity. For this all possible replicate flask pairs were selected and for each pair the differences in all available trace species computed. Ideally for each trace species these residual differences should be distributed normally with a mean of zero and a standard deviation of $\sqrt{2}\sigma_{ana}$, where σ_{ana} is the measurement analytical repeatability.

	0	1
CO_2	4390	135
CO	4273	210
CH_4	4465	147
N_2O	3937	104
H_2	3819	127
SF_6	4408	207
$\delta^{13}C - CO_2$	3817	156
$\delta^{18}O - CO_2$	3694	102
O_2/N_2	3445	159
Ar/N_2	3322	184
$O_2/N_2{}^c$	3736	132
APO	3417	201
APO^c	3684	184
$\delta^{13}C - CH_4$	2308	48
$\delta^2H - CH_4$	2456	43

Table 11: Total number of flask replicate averages for the each species with associated flags (0 or 1).

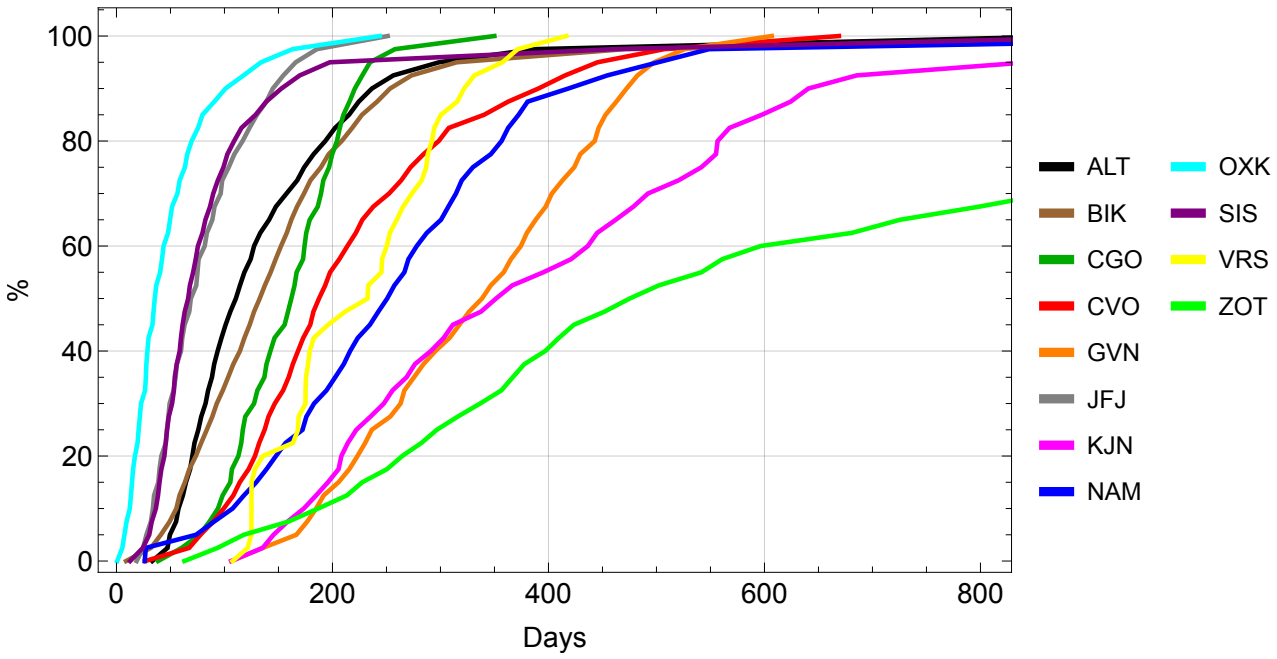


Figure 8: Cumulative distribution of the time difference between flask sampling and flask analysis for the different stations.

The statistics of the flask replicate pair differences for all species and the entire database of replicates are shown in Table 12 in terms of mean and standard deviation. Only flasks measurements are included in this statistic which are flagged = 0, i.e. which have passed all the quality tests. Evidently, this provides only an internal consistency test, but contrasting the standard deviation of replicate pair differences with the BGC base precision multiplied by $\sqrt{2}$ (last column) indicates that this uncertainty assumption is realistic.

	Mean	Standard deviation	(Base precision) $\cdot\sqrt{2}$
CO_2	0	0.12	0.14
CO	-0.03	2.1	2.83
CH_4	0.01	1.88	2.83
N_2O	-0.002	0.169	0.24
H_2	0.05	1.89	2.83
SF_6	0	0.042	0.042
$\delta^{13}C - CO_2$	0	0.024	0.028
$\delta^{18}O - CO_2$	-0.001	0.071	0.071
O_2/N_2	0.08	5.49	5.66
Ar/N_2	-0.14	13.27	11.31
$O_2/N_2{}^c$	0.05	4.74	5.66
APO	0.09	5.42	5.66
APO^c	0.06	4.67	5.66
$\delta^{13}C - CH_4$	-0.005	0.163	0.141
$\delta^2H - CH_4$	0.02	1.46	1.41

Table 12: Mean and standard deviation of all flask replicate pair differences; selected with flag 0. The last column shows an expected standard deviation based on the assumed BGC base precision multiplied by $\sqrt{2}$.

An alternative view on the statistics of the flask replicates is shown in Figure 9. It displays the cumulative distribution of the standard deviations assigned to the flask replicate means for each station and each trace species.

For most species and at most stations more than 80% of the standard deviation values of the replicates lie within the BGC base precision. Somewhat conspicuous are the O_2/N_2 and APO measurements at GVN and to a minor extent also at BIK. At GVN this is likely a consequence of the low sampling temperatures in Antarctica inducing fractionation effects at the air inlet as indicated also in the Ar/N_2 measurements. At BIK the statistics improve after applying the Ar-correction as indicated in the graphs of $O_2/N_2{}^c$ and APO^c, which may point to a leakage problem of the installed automated flask sampler. A somewhat larger scatter is also seen at OXK in the GC measurements of most trace gases, most likely caused by the rather large local short-term variability experienced at this site (see below).

6.4 Comparison with other measurement programs

MPI-BGC participates in an intercomparison program with partner institutions at several sites. Thereby flasks are regularly sampled almost concurrently (within about 2 hours) for the different institutions. This provides in principle an end-to-end comparison between the different measurement systems (sampling, flask shipping and storage, handling, laboratory analysis), however short-term atmospheric variations in the trace species concentrations (if present) will add additional variability on top of the noise from the measurements. Here we show a preliminary comparison of the measurements at Alert and at Ochsenkopf.

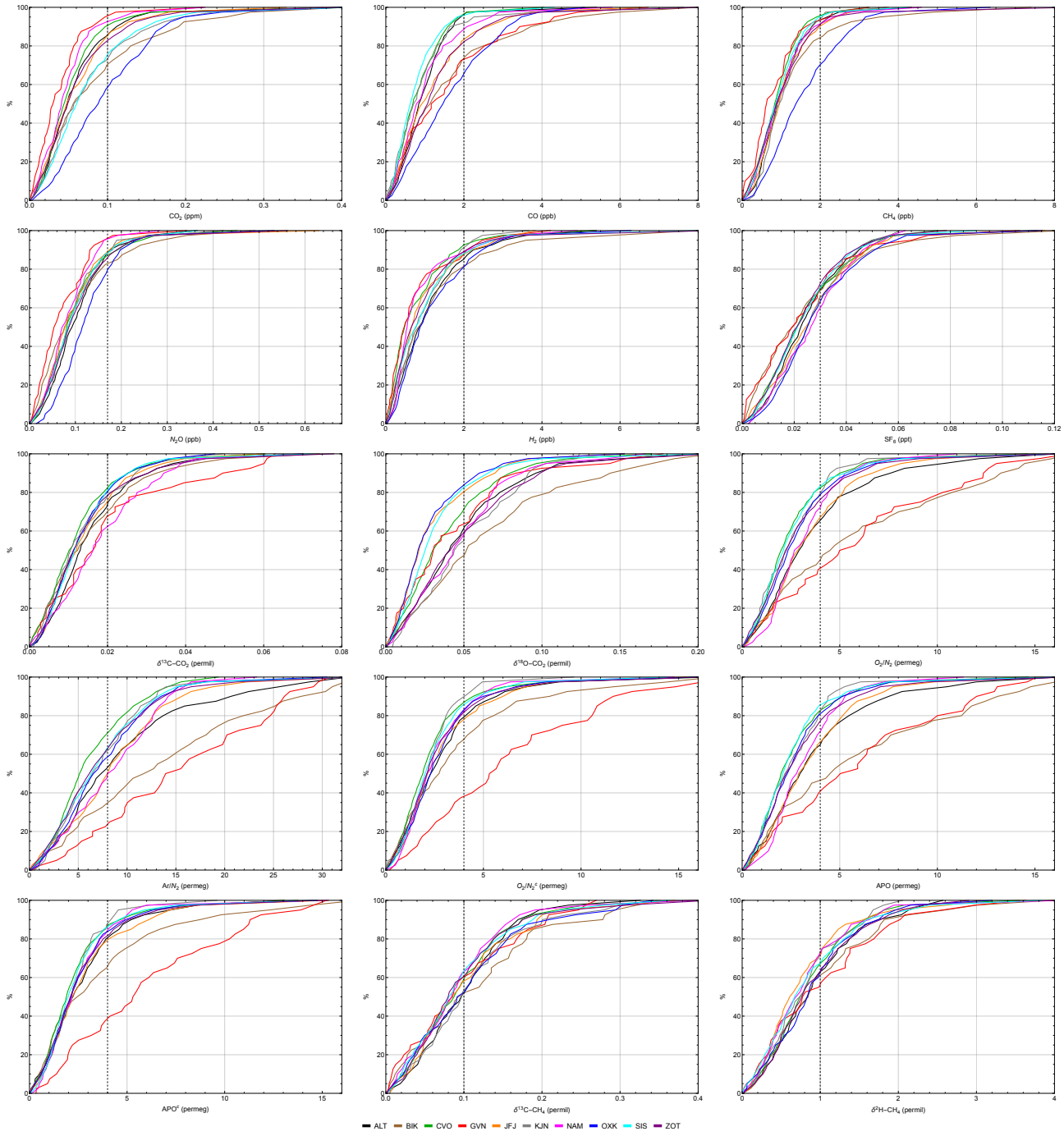


Figure 9: Cumulative distribution of the standard deviation of the flask replicate means for the different trace species and for each station. The vertical dashed line indicates the base BGC base precision.

6.4.1 Alert

Figure 10 shows the differences between concurrently sampled flasks at Alert (ALT) between measurements by MPI-BGC minus measurements by NOAA-GML (gas measurements), NOAA-INSTAAR (isotope measurements) and the Scripps Institution of Oceanography (O_2/N_2 and Ar/N_2 measurements).

For CO_2, CH_4 and SF_6 the intercomparison is excellent both in bias and variability, meeting well the WMO target. The small differences in the CO flask measurements may be caused by the relatively long storage time of the flasks, which is known to induce spurious concentration increases with time. The agreement in N_2O between the labs has been determined by limited instrumental repeatability for most of the time with an average offset of 0.2 ppb. After an instrumentation upgrade in August 2019 the NOAA measurement uncertainty has improved significantly. While a small bias of 0.1 ppb has remained this offset is very consistent since then. A large number of MPI-BGC samples between 2011 and 2014 were contaminated in N_2O caused by a change in the supplier of the drying agent (anhydrous magnesium perchlorate) and had to be flagged. A subtle influence on the remaining unflagged N_2O data before 2014 cannot be fully excluded.

For the MS-measurements $\delta^{13}C - CH_4$ exhibits a known offset between the labs. Efforts to establish a calibration standard for this species are currently underway. The slight trend in time of the lab difference still needs to be assessed.

There is an improvement in the differences of the O_2/N_2 measurements if the Ar-correction is applied (O_2/N_2 vs $O_2/N_2{}^c$) in that the standard deviation of the differences is slightly reduced from 6.75 permeg to 5.25 permeg. Until the year 2015 the agreement of the oxygen measurements between the two labs is excellent (in particular if for BGC the Ar-corrected values are considered). Between 2015 and 2019 the SIO oxygen values are about 20 permeg lower than the BGC measurements. At this point the reasons for this discrepancy is not known. For Ar/N_2 there is a constant offset of 116 permeg, close to the difference of the calibration standards of 126.3 permeg as noted in Section 4.3.2. Over the 15 years of intercomparison no systematic trend is observed.

6.4.2 Ochsenkopf

At Ochsenkopf normally two sets of flasks are filled subsequently using the same inlet line but different flasks and flask samplers (minimum delay ca 15 min, in ca 50% of the cases the set sampling times are within 30 min, in 75% of the cases within 60 min). One set of samples is collected using the MPI-BGC flask sampler described in Section 3, the other set using a manual portable sampling unit and the type of flasks that is generally employed in the NOAA Global Monitoring Laboratory's Cooperative Air Sampling Network. Figure 11, left column, shows the gas measurement differences between flask replicate averages sampled within 2 hr at Ochsenkopf (OXK) between MPI-BGC (both flask sample sets) minus NOAA-GML (NOAA flask replicate averages only).

This continental site exhibits a much larger atmospheric short-term variability than at Alert in most species, hence the plotted differences show a substantially larger noise compared to the expected differences due to the measurement systems alone. In order to remove the part of the atmospheric variability MPI-BGC and NOAA-GML results from the NOAA flask samples only are also compared. The results are shown in the right hand column in Figure 11. Obviously, for most species the scatter of the measurements is much smaller and similar to the results shown above for Alert. The mean difference between the measurements of the two labs lies within the base BGC precision for all species.

The corresponding differences of the MS measurements between MPI-BGC minus those from NOAA-INSTAAR (only data prior to 2015 are available) is shown in Figure 12. For this intercomparison only the separately (within 2 hr) sampled flask measurements from the two labs are available.

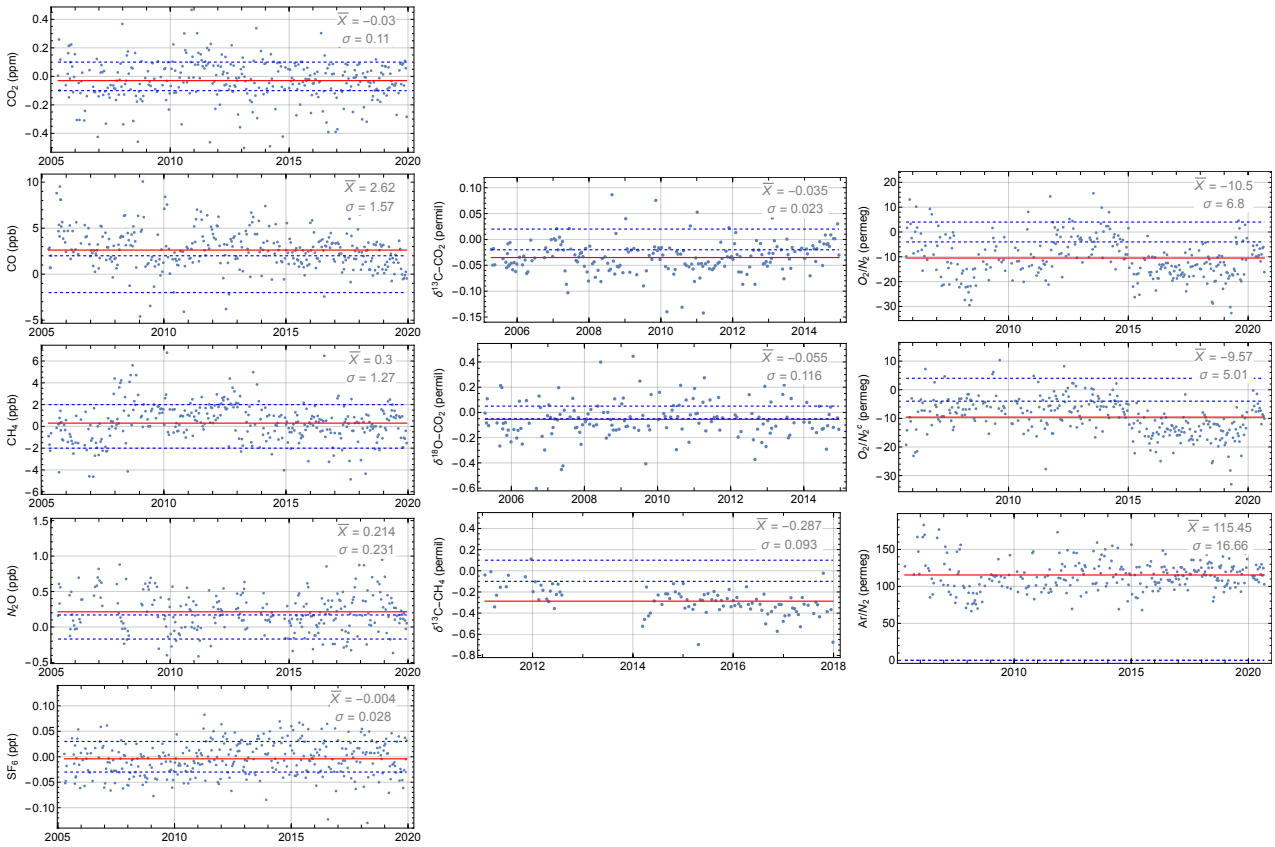


Figure 10: Concurrently sampled valid flask replicate means from Alert measured at MPI-BGC minus flask replicate means measured by NOAA-GML (GC-measurements, left), NOAA-INSTAAR (MS-measurements, center) and SIO (right) as a function of time. The red lines show the mean difference (trimmed by 5%) of all concurrent measurements. The blue dashed lines indicate the range around zero spanned by the base BGC precision.

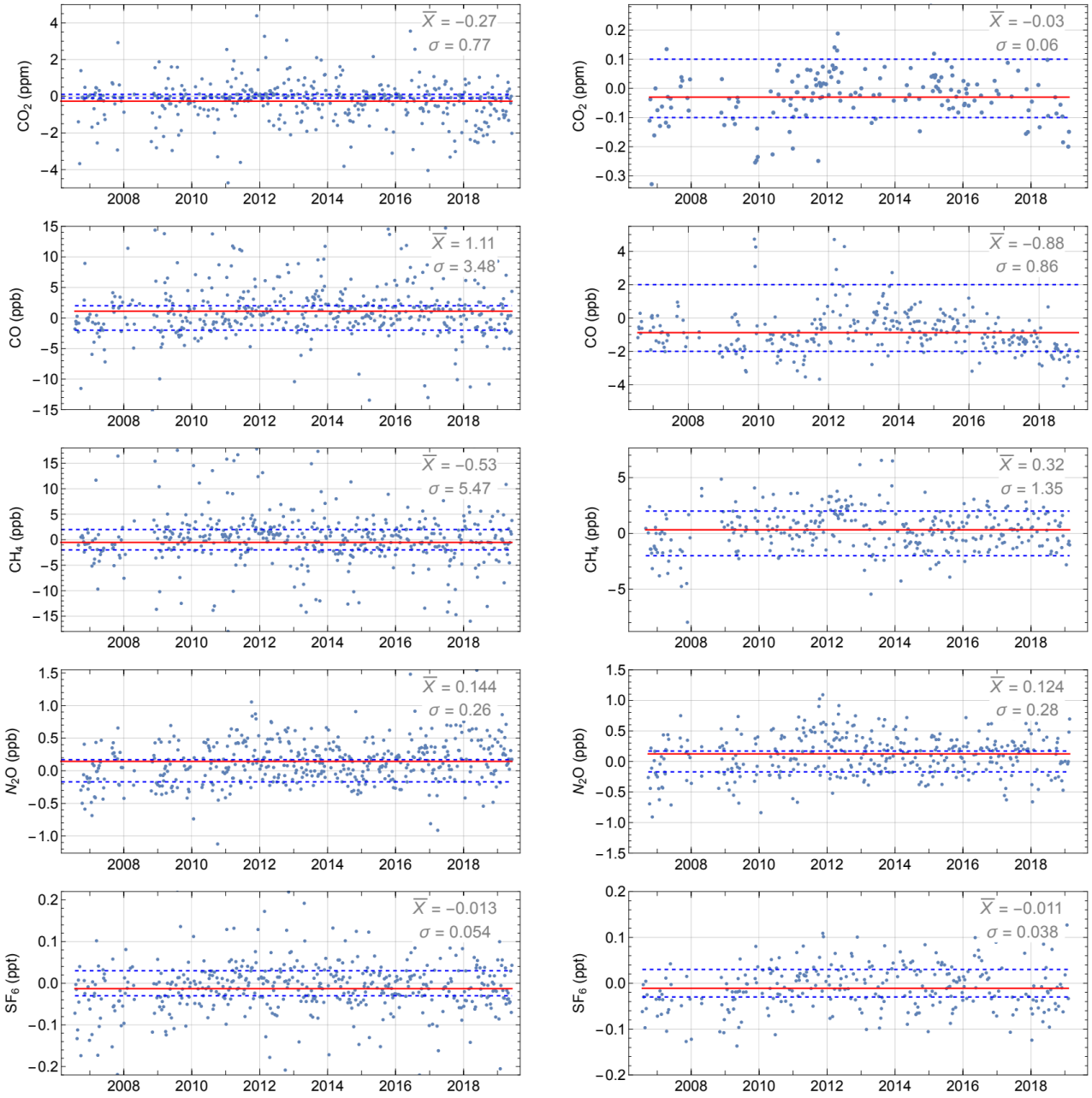


Figure 11: Left column: Flask replicate means from Ochsenkopf measured at MPI-BGC minus replicate means measured by NOAA-GML; separately sampled flasks within 2 hours. Right column: Differences between flasks from Ochsenkopf filled at the same time in series on the same air line (MPI-BGC measurement minus NOAA-GML measurement). Notice the different scales on the vertical axis. The red lines show the mean difference (trimmed by 5%) of all measurements. The horizontal blue dashed lines indicate the range around zero spanned by the base BGC precision.

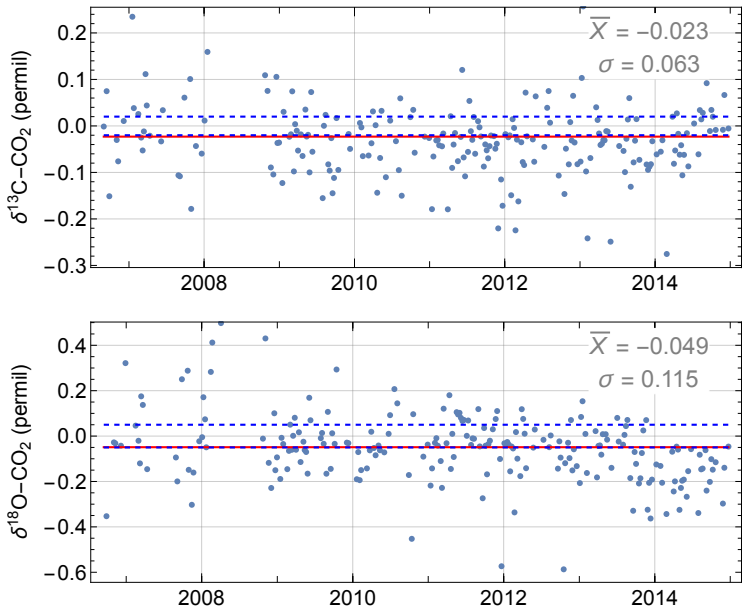


Figure 12: Concurrently sampled flask replicate means from Ochsenkopf measured at MPI-BGC minus flask replicate means measured by NOAA-INSTAAR.

7 Acknowledgements

The flask sampling program of MPI-BGC would not have been possible without diligent personnel performing the laboratory analyses over the many years of the program. In particular we thank Jürgen Richter, Isabell von Rein, Steffen Knabe and Bert Steinberg. We are also grateful to Uwe Schultz and Michael Hielscher for building the flask samplers which have been deployed at the stations. We also thank the many operators of the different institutions at the stations for their meticulous regular sampling of the flasks; in particular Ronnie Robertson, from Virkie, Shetlands (UK). The support by our intercomparison partners in providing valuable information through sharing their data is highly appreciated. The flask program was initially set up by Andrew Manning and the sampling initiated at the stations by Elena Kozlova (ZOTTO, Cape Verde), Elena Popa (Bialystok) and Rona Thompson (Ochsenkopf, Gobabeb). Funding for the MPI-BGC flask sampling program is provided by the Max-Planck-Society.

A QC and Ar-correction of O_2/N_2 and APO using Ar/N_2 measurements

A.1 Motivation for correcting O_2/N_2 with Ar/N_2 measurements

Analysis of the O_2/N_2 measurements of flask replicates yielded a relatively large spread; substantially larger than the analysis repeatability. Since the flasks are overpressured ($\sim$ 1.6 bar) and due to logistical reasons are often stored for a substantial time period before shipping and analysis at MPI BGC (see Section 6.2), it was suspected that potential micro-leaks might cause fractionation effects.

If fractionation effects occur, they would affect the Ar/N_2 ratio in a similar way as O_2/N_2. Theoretically, the ratio of the fractionation effects for the two ratios should be approximately $1/2.5 = 0.4$, caused by molecular effusion (Knudsen diffusion effect) (Langenfelds et al., 2005) (see also subsection below). Indeed, a graph of the individual flask replicate residuals of O_2/N_2 vs Ar/N_2 shows a clear linear relationship with a slope (0.44) close to the theoretical value (see Figure 13).

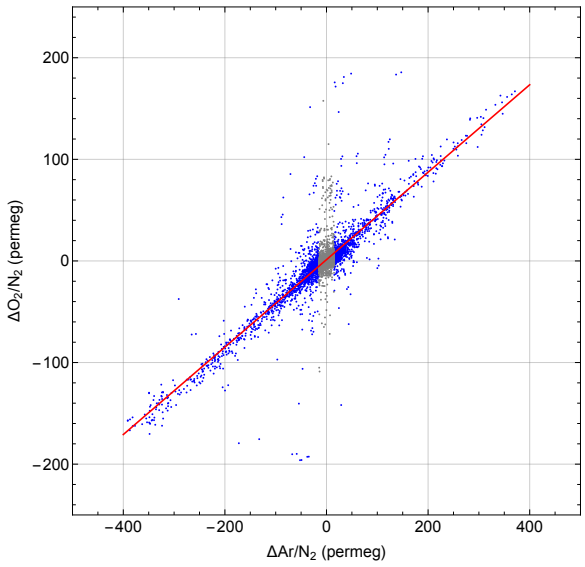


Figure 13: Observed replicate flask residuals of O_2/N_2 vs concurrent flask residuals of Ar/N_2. The red linear regression line computed from residuals lying outside of the d_{crit} criterion for Ar/N_2 ($\pm$15 permeg, blue points) has a slope of 0.44, close to the theoretical slope of 0.4 expected from molecular effusion through micro-leaks.

Natural variations in atmospheric Ar/N_2 are at least one order of magnitude smaller than in O_2/N_2, driven only by changes in ocean heat content causing out- and ingassing of these gases in slightly different proportions due to different solubilities as a function of water temperature. The observed amplitude of the seasonal cycle of Ar/N_2 is smaller than 20 permeg, latitudinal gradients are smaller than 20 permeg (Battle et al. (2003), Keeling et al. (2004)) and the expected long-term trend due to global ocean warming is on the order of 0.3 permeg/yr. These natural variations are smaller than our measurement reproducibility.

If we tentatively attribute the entire measured variability of Ar/N_2 to fractionation effects assuming a constant uniform atmospheric Ar/N_2 ratio, we can define a "corrected" O_2/N_2 ratio as:

$$\delta O_2^c = \delta O_2 - g(\delta Ar - \delta Ar_0) \tag{3}$$

The reference Ar/N_2 ratio, δAr_0, was set at a nominal value of 140 permeg. The effect of varying g on the spread of the flask measurements of individual replicates is shown in Figure 14. The solid line shows the square root of the trimmed variance of all replicate residuals as a function of g. There is a clear minimum close to the theoretical value of g=0.4.

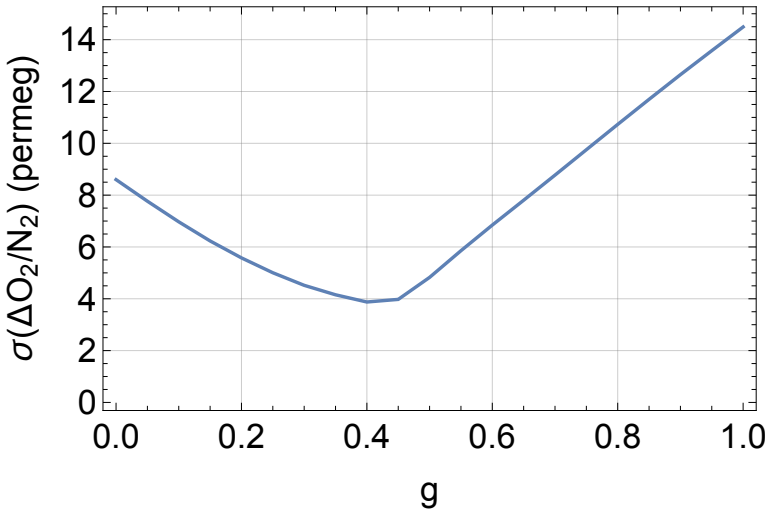


Figure 14: Standard deviation of O_2/N_2 replicate residuals as a function of the fractionation parameter g.

Evidently, the Ar-correction of the O_2/N_2 measurements reduces considerably the scatter of individual replicate residuals, as also seen in Table 12.

A.2 Theory

Knudsen diffusion through micro-leaks fractionates to first order gas species with respect to bulk air according to Graham's law inverse proportionally to the square root of the molecular mass of the gas (m_g) and bulk air ($m_{air} = 28.97$):

$$\kappa_g = \sqrt{\frac{m_{air}}{m_g}} \tag{4}$$

The effect of this process on the remaining gas in the flask can be derived by considering the change in air mass:

$$M_1 = M_0 - \Delta M \tag{5}$$

where M_0 and M_1 are the amount of air in the flask at sampling and at measurement in the lab. For the trace gas the fractionation has to be taken into account:

$$\chi_1 M_1 = \chi_0 M_0 - \kappa_g \chi_0 \Delta M \tag{6}$$

where χ denotes the mixing ratio of the trace gas. Dividing by M_1 and inserting (5) yields:

$$\chi_1 = \chi_0 \frac{M_0 - \kappa_g \Delta M}{M_0 - \Delta M} \tag{7}$$

which can be written as

$$\chi_1 = \chi_0 \frac{1 - \kappa_g x}{1 - x} \tag{8}$$

where $x = \frac{\Delta M}{M_0}$ denotes the relative loss of air mass in the flask.

Thus the change in mixing ratio of gas g due to fractionated air loss becomes

$$\Delta\chi_g = \chi_1 - \chi_0 = \chi_0 \left(\frac{1 - \kappa_g x}{1 - x} - 1 \right) = \chi_0 \cdot \frac{x}{1 - x} \cdot (1 - \kappa_g) \tag{9}$$

Measurements of two gas species permit the elimination of the term $\frac{x}{1-x}$ in equation (9) and defines the slope of measurements of $\Delta\chi_{g1}$ vs $\Delta\chi_{g2}$:

$$\frac{\Delta\chi_{g1}}{\Delta\chi_{g2}} = \frac{1 - \kappa_{g1}}{1 - \kappa_{g2}} \left(\frac{\chi_{0,1}}{\chi_{0,2}} \right) \tag{10}$$

The right hand side term in parentheses denotes the initial mixing ratio of the two gases in the flask. It can be approximated with the ratio of typical reference mixing ratios of the two gases.

Using Ar/N_2 ratios as an indicator of mixing ratio changes caused by micro leaks requires an expansion of the formalism, since Ar/N_2 is a gas concentration ratio and not a mixing ratio. Using the "permeg" definition in which Ar/N_2 ratios are reported:

$$^{Ar}\delta = \left(\frac{\frac{\chi_{Ar}}{\chi_{N_2}}}{(\frac{\chi_{Ar}}{\chi_{N_2}})_{ref}} - 1 \right) \cdot 10^6 \tag{11}$$

deviations of this quantity can be expressed as

$$\Delta^{Ar}\delta = \Delta \left(\frac{\chi_{Ar}}{\chi_{N_2}} \right) \cdot \frac{1}{(\frac{\chi_{Ar}}{\chi_{N_2}})_{ref}} \cdot 10^6 \tag{12}$$

Assuming small deviations around the reference ratio values, the deviation of this quantity can be related to relative changes in the Ar and N_2 content

$$\Delta^{Ar}\delta = \left(\frac{\Delta\chi_{Ar}}{\chi_{Ar_{ref}}} - \frac{\Delta\chi_{N_2}}{\chi_{N_{2,ref}}} \right) \cdot 10^6 \tag{13}$$

With equation (9) setting $\chi_{Ar,0} \approx \chi_{Ar_{ref}}$ and $\chi_{N_2,0} \approx \chi_{ref}$ this becomes

$$\Delta^{Ar}\delta = x(\kappa_{N_2} - \kappa_{Ar}) \cdot 10^6 \tag{14}$$

The theoretical slope of mixing ratio deviations caused by micro leaks of a gas g vs the corresponding deviations in the Ar/N_2 ratio is thus:

$$\frac{\Delta\chi_g}{\Delta^{Ar}\delta} = \frac{(1-\kappa_g)\chi_g}{(\kappa_{N_2} - \kappa_{Ar}) \cdot 10^6} \tag{15}$$

If not a gas but an isotopic ratio is concerned, a similar derivation leads in the case of $^{13}C/^{12}C$ to

$$\frac{\Delta^{13}\delta}{\Delta^{Ar}\delta} = \frac{(\kappa_{^{12}CO_2} - \kappa_{^{13}CO_2}) \cdot 10^3}{(\kappa_{N_2} - \kappa_{Ar}) \cdot 10^6} \tag{16}$$

and analogously for $^{18}O/^{16}O$, $^{13}C/^{12}C(CH_4)$ and $D/H(CH_4)$.

A.3 Observations

Outliers of flask replicate residuals, as defined in Section 6.3, with concurrent Ar/N_2 residuals indeed exhibit the theoretical relationships derived in the previous section. Figure 15 shows the result for the different species. In each case flask replicate residuals of Ar/N_2 larger than 200 permeg are indicated with red dots. The red regression line, computed from these outlier residuals, indicates the observed and the black dotted line the expected theoretical relation computed using the equations in the previous section (using standard values of the molecular weights of the different molecules). Table 13 show for the different species the theoretical value of the slopes and the observed values from the computed regression, including the statistical error and the p-value of the t-test.

Obviously, not all concentration or isotope ratio deviations are caused by micro leaks. Other sources of error dominate the flask measurement close to the flask median. It appears, however, that a large fraction of the real outliers are indeed caused by diffusion through micro leaks.

The computed statistics, computed from the outliers (defined as Abs ($\Delta Ar/N_2$)> 200 permeg) are for most species highly significant and the observed slopes with its errors bracket well the theoretical slopes. Exceptions are CO (which has a molecular weight very close to that of atmospheric air), and H_2, for which diffusion through micro-leaks appears to be more complicated than assumed by simple Knudsen diffusion.

The very high correlation of the O_2/N_2 deviations reflects what was discussed in Appendix A.1.

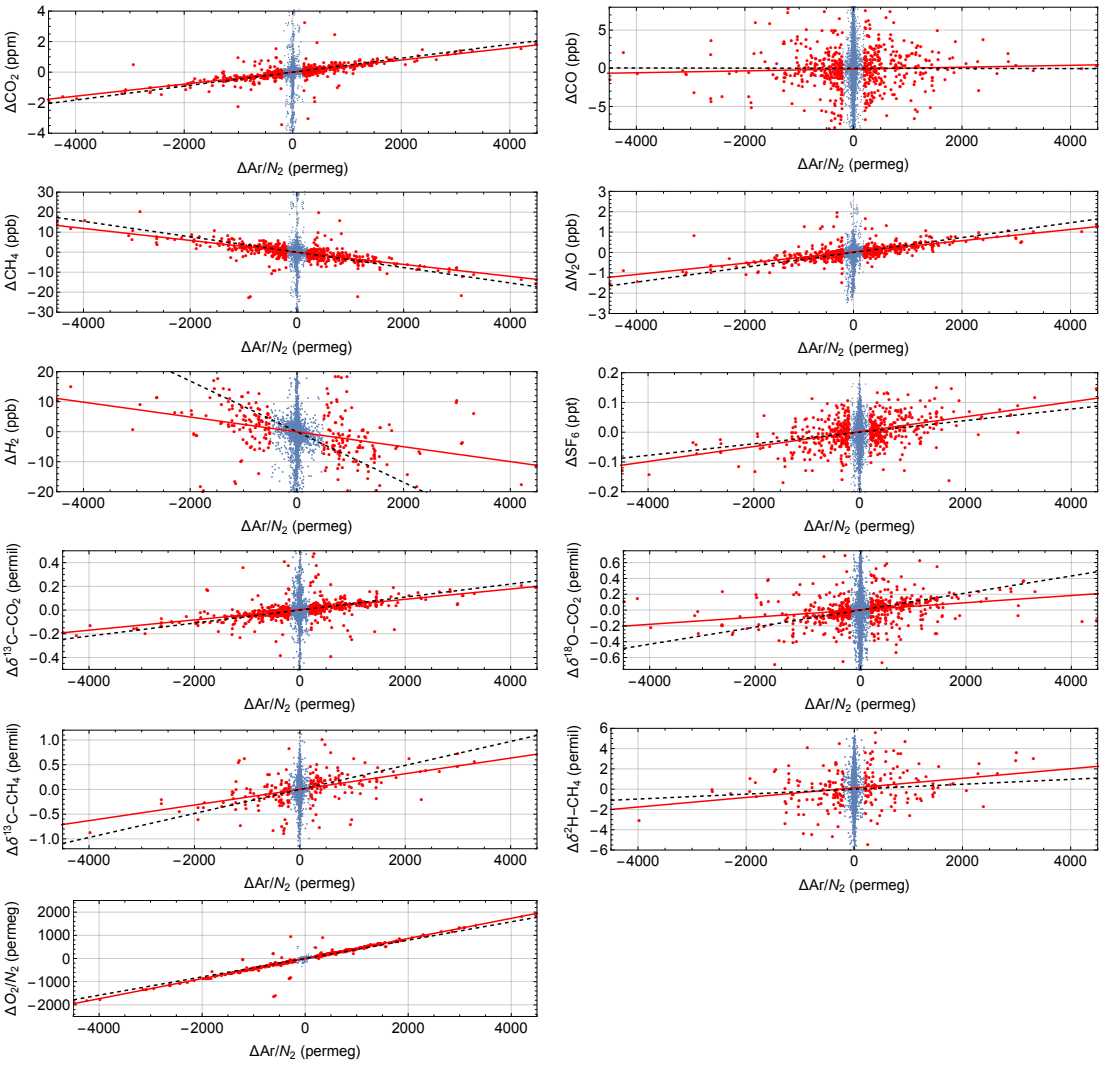


Figure 15: Correlations of flask replicate residuals vs Ar/N_2 residuals. Outliers with $Abs(\Delta Ar/N_2)>200$ permeg (except for H_2: $Abs(\Delta Ar/N_2)>500$ permeg) are indicated with red dots. Black dashed lines show the theoretical relationship; red lines the regression line taking all the outliers.

Species	Units	Theory ($\times 10^3$)	Observed ($\times 10^3$)	σ(Observed) ($\times 10^3$)	p	N
CO_2	ppm/permeg	0.456	0.394	0.017	2.1×10^{-86}	570
CO	ppb/permeg	-0.0103	0.121	0.12	0.29	529
CH_4	ppb/permeg	-3.85	-2.99	0.14	1.4×10^{-78}	569
N_2O	ppb/permeg	0.365	0.277	0.012	1.1×10^{-80}	580
H_2	ppb/permeg	-8.44	-2.48	0.37	1.1×10^{-10}	240
SF_6	ppt/permeg	0.0195	0.0251	0.0022	5.7×10^{-28}	531
$\delta^{13}C-CO_2$	permil/permeg	0.0548	0.0434	0.0031	7.5×10^{-39}	578
$\delta^{18}O-CO_2$	permil/permeg	0.108	0.0455	0.0072	4.4×10^{-10}	555
$\delta^{13}C-CH_4$	permil/permil	0.243	0.158	0.02	2.5×10^{-13}	207
$\delta^{2}H-CH_4$	permil/permil	0.243	0.472	0.12	0.00017	205
O_2/N_2	permeg/permeg	396.	433.	4.9	0.	592

Table 13: Numerical values of the slopes between flask replicate residuals of trace species and Ar/N_2.

A.4 Ar-test

Unlike for O_2/N_2 the relationship is for the other trace species much less robust and hence not useful for a correction procedure. However it can be used for defining an Ar/N_2 based outlier-criterion, beyond which the species flask measurement is flagged as suspicious.

As discussed for O_2/N_2 (Appendix A.1) we assume for Ar/N_2 a globally uniform, constant reference value of 140 permeg and define for each trace species an outlier criterion based on the measured Ar/N_2 value: if it is outside of the reference value $\pm$ the tracer specific outlier criterion this tracer value gets flagged as potentially suspicious. Thereby the compatibility goals (1-sigma) defined by the WMO are mapped to the corresponding bounds of Ar/N_2 using the theoretical relationship derived in Appendix A.2. Table 8 lists the resulting Ar/N_2 criteria for the different trace species.

B Data and graphics for all stations

The flagged replicate means for all stations are available together with graphics in subdirectories of a collection in the Edmond repository of the Max Planck Digital Library (`https://dx.doi.org/10.17617/3.8r`, direct link).

Excel files for each station containing all species data in separate sheets are in subdirectory "data_xls"; plain CSV ASCII textfiles for each station are in subdirectory "data_csv" for each species.

Graphs for each species measured at all stations are available in the subfolder "graphics". As an example, Figure 16 shows the complete record of GC measurements and Figure 17 of MS measurements at the station Alert. Blue dots indicate flask replicate averages with flag=0, red dots with flag=1. The gray shaded area shows the calculated fit (seasonal cycle plus trend) to the replicate averages with flag=0; the width of the shading indicates the standard deviation. The dashed line shows the trend.

The complete database of all flasks with flags for each species measurement is available together with the present report draft in the main Edmond collection:

`database_BGC_flagged_v13.4_2022-01-03.xls`

`flask_report_v13_2022-01-11a.pdf`

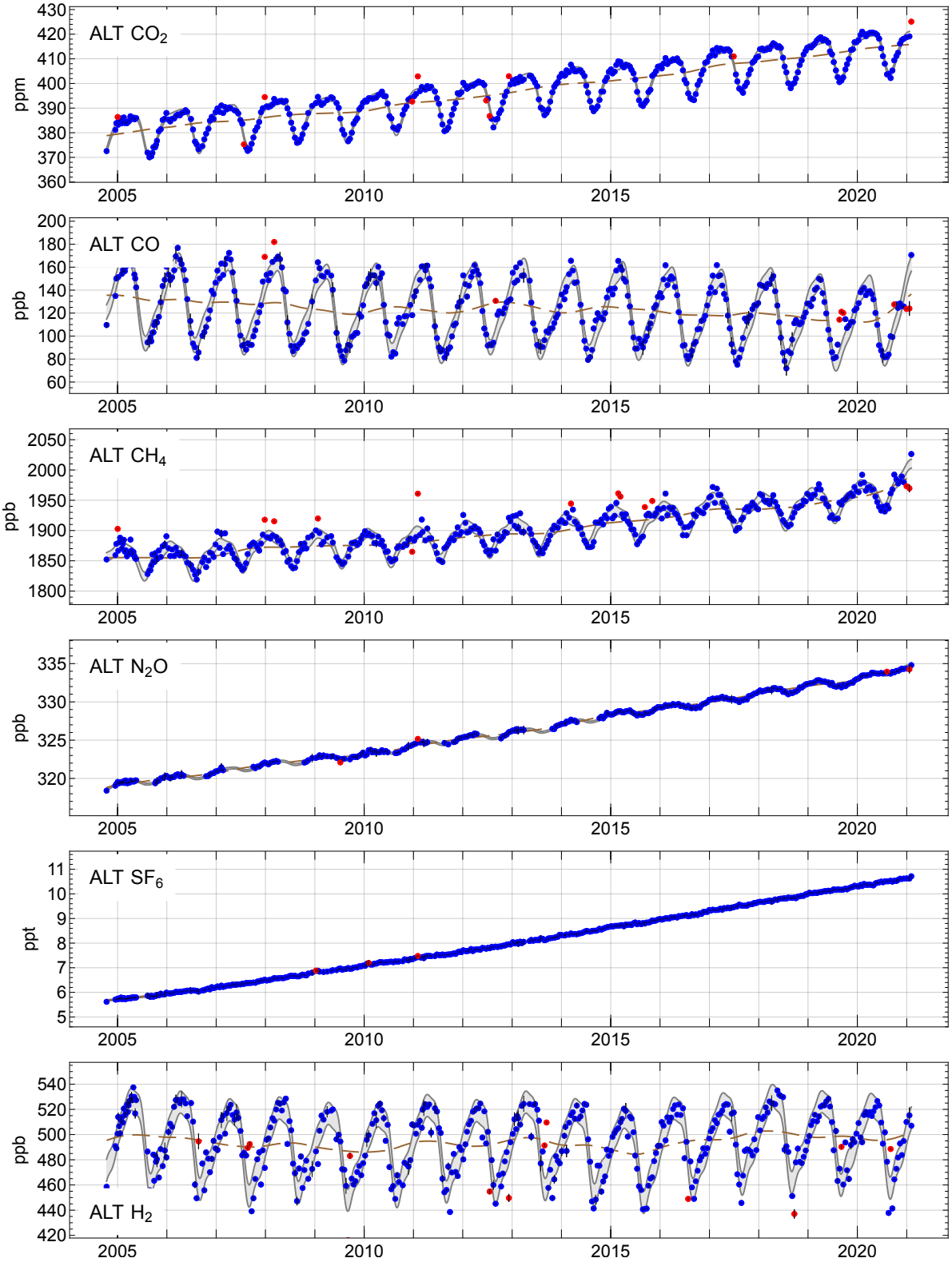


Figure 16: GC measurements (flask replicate averages) from the station Alert (ALT). Blue: flag=0, red: flag=1. Gray shaded area: calculated fit (seasonal cycle plus trend) to the replicate averages with flag=0; width of the shading indicates the standard deviation. The dashed line shows the trend.

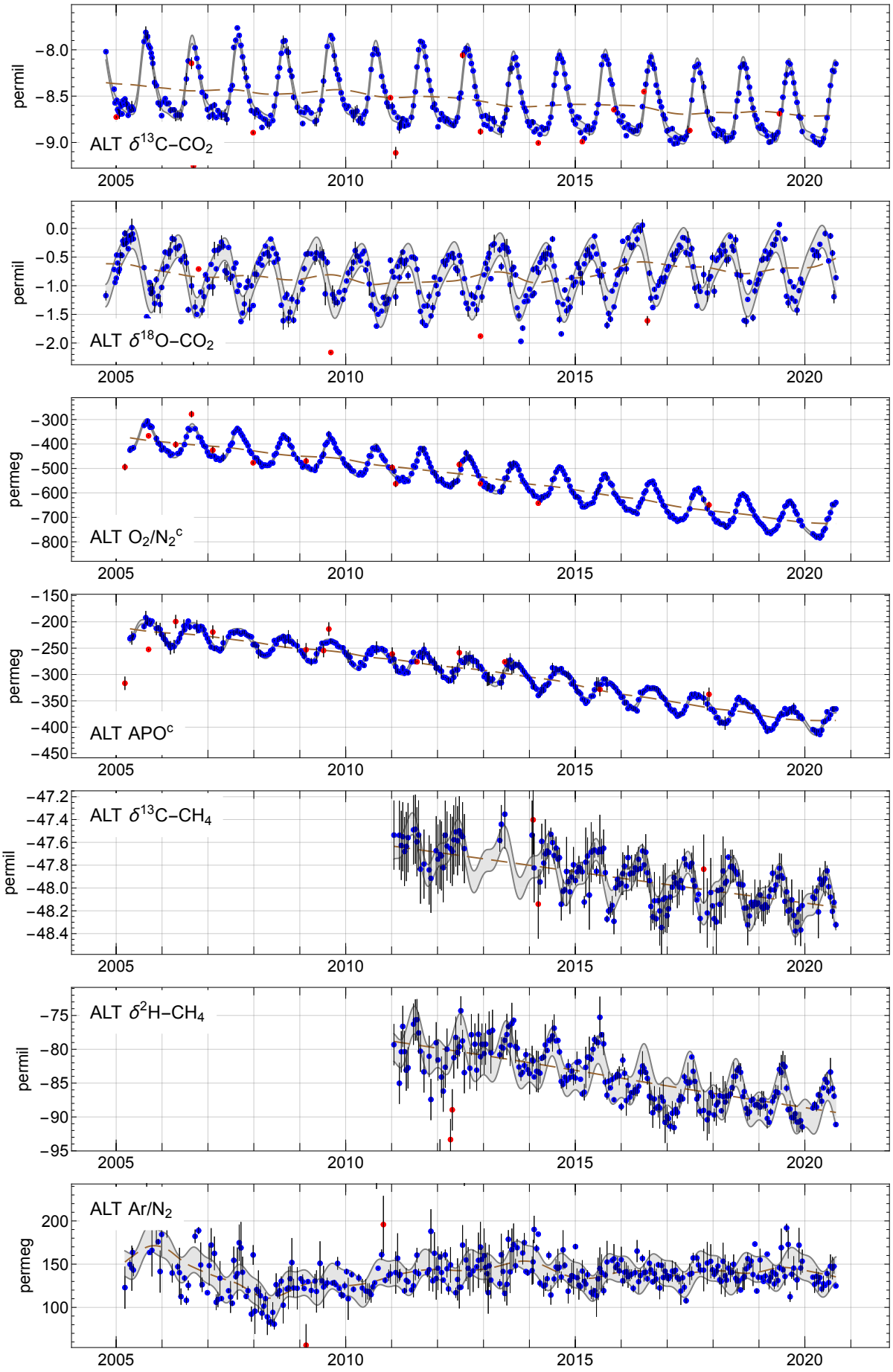


Figure 17: MS measurements (flask replicate averages) from the station Alert (ALT). Blue: flag=0, red: flag=1. Gray shaded area: calculated fit (seasonal cycle plus trend) to the replicate averages with flag=0; width of the shading indicates the standard deviation. The dashed line shows the trend.

C Station information

C.1 ALT, Alert, Canada

Canadian Forces Station Alert (CFS Alert) lies at the northeastern tip of Ellesmere Island in the Canadian arctic (82.4508N,62.5072W). The greenhouse gas measurement program is operated by Environment and Climate Change Canada. The station serves several networks, hence provides an opportunity for comparisons between the measurement programs of the different agencies.

Figure 18: Canadian Forces Station Alert, Canada. (Photo: K. Rawlings, Wikimedia)

C.2 VRS, Villum Research Station, Station Nord, Greenland

Villum Research Station, Station Nord, lies at the northeastern tip of Greenland (81.5812N,16.6402W). The research station is operated by Aarhus University, Denmark, while the logistics is provided by the Danish military. See the research station website.

Figure 19: Air Monitoring House at Villum Research Station, Station Nord, Greenland (Photo: Hendrik Skov)

C.3 KJN, Kjølnes, Norway

Air sampling is performed at the Kjølnes Lighthouse near Berlevåg at the Barents Sea coast in northern Norway (70.852227N, 29.232191E). Air inlets are located on the roof of one of the houses close to the lighthouse which also contains instruments for continuous measurements. The measurement program and the flask sampling is operated by the Exeter Atmospheric and Ocean Science research group of Exeter University, UK.

Figure 20: Kjølnes Lighthouse at Barents Sea coast, Norway

C.4 ZOT, Zotino Tall Tower Observatory (ZOTTO)

The Zotino Tall Tower Observatory (ZOTTO) is located in the taiga in central Siberia, about 20km west of the village Zotino at the river Yenisey. The station includes a 304m tall mast from which air is pumped down through pipes and sampled at the tower base in a laboratory bunker. Flasks are filled with air from the 301m level.

ZOTTO has been established jointly between the MPI-BGC, the Max Planck Institute for Chemistry, Mainz, Germany, TROPOS, Leipzig, Germany and the Institute of Forest, Krasnojarsk, Russian Federation (IFOR). It is operated by staff from the Institute of Forest. Details on the measurement program at ZOTTO can be found in Heimann et al. (2014).

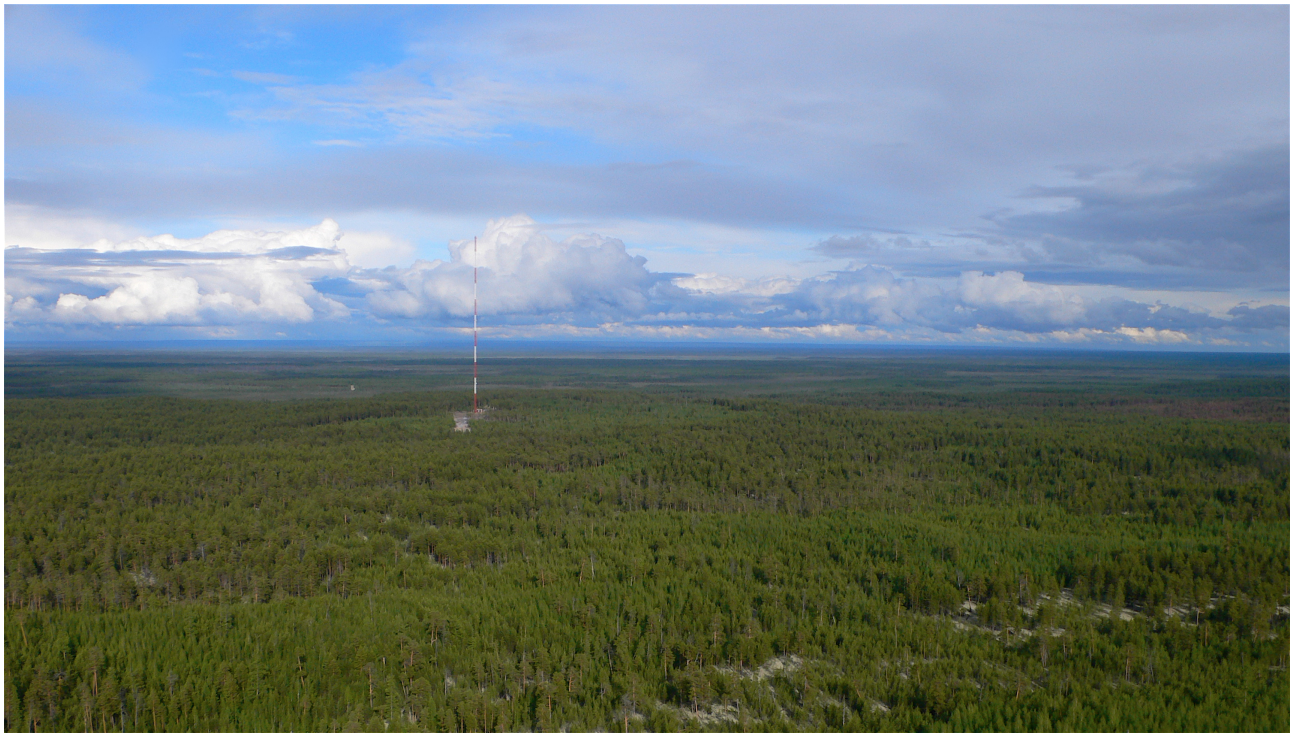


Figure 21: Zotino Tall Tower Observatory (ZOTTO), Russian Federation.

C.5 SIS, Shetland Islands, UK

In 1993 CSIRO (Australia) as part of their atmospheric baseline measurement program began with air sampling on Shetland at an eastern (60.1403N, 1.1792W) and western location (60.0889N, 1.2553W) of the main island ("Shetland-Burra") depending on wind direction (Francey et al., 1998). In 2003 the flask sampling program was transferred to the MPI-BGC. A new remote location more suitable for most wind directions was found at Sumburgh Head on the southern tip of the main island, about 26 km to the south (59.854833N, 1.274298W). Early samples from the Shetland-Burra site are indicated with station code SIS0.

Figure 22: Sumburgh Head, Shetland, UK.

C.6 BIK, Bialystok, Poland

Air sampling is performed on the 300m tall television tower mast at Krynice near Bialystok in eastern Poland (53.231517N, 23.026803E). The measurement system was installed by the Max Planck Institute for Biogeochemistry as part of the CHIOTTO EU-funded project (EVK2-CT-2002-00163). Details of the online measurement system are described in Popa et al. (2010). It is expected that the station will be integrated in a future Polish component of the European research infrastructure "Integrated Carbon Observation System" (ICOS).

Figure 23: 300m Krynice television mast near Bialystok, Poland

C.7 OXK, Ochsenkopf, Germany

Measurement and flask sampling systems were installed by the MPI-BGC in 2003 on a telecommunication tower on the top of the Ochsenkopf mountain in the German Fichtelgebirge. Information on the site and the installed measurement system can be found in Thompson et al. (2009). In 2019 the station was officially integrated into the atmospheric national network of the German component of ICOS, operated by the German national weather service (DWD).

Figure 24: Ochsenkopf communication tower, Germany

C.8 JFJ, Jungfraujoch, Switzerland

Jungfraujoch is a high altitude research station located on the ridge between the Mönch and Jungfrau mountains in Switzerland. Air sampling is performed on the Sphinx platform on the top of the observatory (46.547481N, 7.985129E, 3578masl). The greenhouse gas measurement program is operated by the Climate and Environmental Physics institute of the University of Bern. Continuous greenhouse gas measurements at Jungfraujoch are now also being performed as part of the ICOS atmospheric observations by the Swiss Materials Science and Technology Institute (EMPA).

Figure 25: High Altitude Station Jungfraujoch, Switzerland.

C.9 CVO, Cape Verde Atmospheric Observatory (CVAO)

The Cape Verde Atmospheric Observatory (CVAO) is located on the north-north-eastern coast of the Sao Vicente island of Cape Verde (16.864015N, 24.867496W). The station faces most of the time the trade wind from the NNE direction. Flask sampling is performed on air pumped down from a 30m tall tower.

The station has been established as a joint project with MPI-BGC, TROPOS (Leipzig, Germany), the University of York (UK), the Exeter Atmospheric and Ocean Science research group of Exeter University (UK) and the meteorological service of Cape Verde (INMG). CVAO complements from the atmospheric side the corresponding Cape Verde Ocean Observatory (CVOO) operated by GEOMAR (Kiel, Germany) and the Ocean Science Center Mindelo in Cape Verde. Maintenance of the local measurements and the flask sampling is performed by INMG. The site and the comprehensive measurement program are described in Carpenter et al. (2010). Details on CVAO can be found on the NCAS website.

Figure 26: Cape Verde Atmospheric Observatory.

C.10 NAM, Gobabeb station, Namibia

The Gobabeb Namib Research Institute is a desert research facility in the Namib desert of Namibia at the ephemeral Kuiseb river, about 55km eastward from the coast of the Atlantic Ocean. The Namib Desert Atmospheric Observatory (NDAO) was established by MPI-BGC in 2012. The site (23.561742S, 15.047052E) and the *in situ* station measurements are described in Morgan et al. (2015) and Morgan et al. (2019).

Figure 27: Air measurement and sampling station at Gobabeb, Namibia.

C.11 CGO, Cape Grim Observatory, Australia

The Cape Grim Observatory is a long term atmospheric "Baseline Air Pollution Station" operated by the Bureau of Meteorology and Commonwealth Scientific and Industrial Research Organisation (CSIRO) of Australia. The station is located at the north-western coast of Tasmania (40.683S, 144.689E), sampling primarily the maritime air from the southern Indian ocean.

Flask sampling for MPI-BGC began in 2002 when CSIRO had been instrumental to build up the analytical facilities of the institute. Flask sampling for MPI-BGC ended after 2018.

Figure 28: Cape Grim Observatory, Tasmania, Australia

C.12 GVN, Neumayer Station, Antarctica

The German Neumayer Station (currently III) is located in Antarctica on an ice shelf close to the southern ocean. The facility is operated by the German Alfred-Wegener-Institute (Bremerhaven). Flask measurements are taken at the air chemistry observatory close to the station (70.6666324S, 8.2688599W). The air sampling program at GVN was established by the University of Heidelberg. Since 2017 air flasks are analysed at MPI-BGC for greenhouse gas concentrations.

Figure 29: Air chemistry observatory at Neumayer Antarctic station. (Photo: AWI)

D Flask sampler operating instructions

Flask Sampling Instructions

Max Planck Institute for Biogeochemistry

1. System installation

1.1 Place compressor and sampling suitcase next to each other.

1.2 Open the lids of both suitcases (pump & flask module) and fix them in vertical position.

1.3 Connect the fixed air inlet tubing (e.g. from a tower) tubing from the mast to the quick connector labelled ***INLET***.

1.4 Take three flasks from the shipping box and place them in the place holders of the sampling suitcase. Flask connectors must be facing up.

1.5 Record flask codes, site code as well as sampling date and name of the operator on the **flask sampling data sheet**.

1.6 Connect flasks, drying cartridge and compressor module in the following described order (***see Figure 1 & 2***).

Note: Flasks have one connector at each end. It is recommended to use the connector which is the nearest from you as the flask inlet.

1. Connect the pump module port labelled ***TO FLASKS*** with the inlet of the drying trap.
2. Connect the outlet of the drying trap to the inlet connector of the first, leftmost flask (closest to drying trap).
3. Connect the outlet of the first flask to the inlet of the second, middle flask.
4. Connect the outlet of the second flask to the inlet of the third, rightmost flask.
5. Connect the outlet of the third flask to compressor suitcase at the port labelled ***FROM FLASKS***.

Note: All Quick Connectors should easily snap into place. A leak-tight sealing of the glass/ metal connectors (Ultra Torr Fittings) are essential for successful air sampling. Thus, make sure that all connections feel tight.

Fig. 1 The Compressor Module

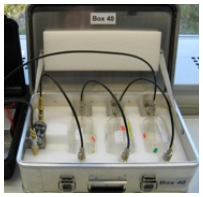


Fig. 2 The Sampling Suitcase

If you have questions concerning the sampling procedure, or if spare parts or supplies are needed please contact us directly.

Jošt V. Lavrič
email: jlavric@bgc-jena.mpg.de
phone: +49 3641576368
fax: +49 3641 577300

Flask Sampling Instructions

Max Planck Institute for Biogeochemistry

2. Collecting air samples

2.1 Put the bypass valve (in the center of panel) in ***BYPASS OPEN*** position.

2.2 Switch on pump to start flushing the system.

2.3 Check upstream and downstream flow rate as well as pressures indicated on the corresponding flow meters ***(~2.0 L/min)*** and pressure gauges ***(~1.0 bar)*** in the compressor suitcase.

Note: If necessary, the sample flow rate can be adjusted by using the valve at the flow meter labelled UPSTREAM.

2.4 Proceed to opening the flask valves, starting with the inlet valve of the first flask followed by the outlet valve of the first flask. Finish with the outlet valve of the third flask.

Note: Opening flask valves in that order is crucial to prevent large pressure gradients in flasks.

2.5 Close bypass valve (valve position: ***BYPASS CLOSE***) to force the air sample stream through the flasks.

Note: If flow rate indicated on the upstream flow meter suddenly drops to zero, immediately reopen the bypass valve. In this case, one or more flask valves are still closed or the drying cartridge is not properly connected to the system.

2.6 Record time on the sampling sheet in row ***"Start flushing time"***. Mark the corresponding check box concerning time format.

Note: Coordinated Universal Time (UTC) time is preferred. Daylight saving time (DST) in summer should be avoided.

2.7 Flush flasks for **15 minutes**.

2.8 After flushing is accomplished record on the sampling sheet flow rates and pressures indicated on the corresponding flow meters and pressure gages.

2.9 Reopen bypass valve (valve position: ***BYPASS OPEN***)

Note: Flushing and pressurizing takes place simultaneously. After the bypass valve is reopened sample air from the flushing procedure remains in the flask with an overpressure of 1.0 bar.

2.10 Close all six flask valves in the following described order.

1. **Inlet valve** of the **first flask**.
2. **Outlet valve** of the **third flask**.
3. Close all open flask valves.

2.11 Record time on the sampling sheet in row ***"Stop sampling time"***.

If you have questions concerning the sampling procedure, or if spare parts or supplies are needed please contact us directly.

Jošt V. Lavrič
email: jlavric@bgc-jena.mpg.de
phone: +49 3641576368
fax: +49 3641 577300

Flask Sampling Instructions

Max Planck Institute for Biogeochemistry

2.12 Fill in the sample sheet with available meteorological data. Additionally, note down any observations concerning sampling conditions that might affected sample quality during collection.

2.13 Turn off pump and disconnect all tubes.

2.14 Place filled flasks and sampling sheet back in the shipment box and return them back to MPI-BGC, Jena.

3. Battery Charging

3.1 Connect the battery charger with the 230 V electricity mains. The green LED labelled ***"Mains"*** lights up.

3.1 Set the slide switch of the battery charger to 12 V.

3.2 Connect the battery charger to the socket labelled **BATTERY CHARGER** at the compressor suitcase.

3.3 The device starts the charging procedure automatically which is indicated by the LED labelled ***"Charging"***.

3.4 The battery is charged and can be disconnected from the battery charger when the LED labelled **"Refresh"** lights up. Now the compressor suitcase is again ready for use.

If you have questions concerning the sampling procedure, or if spare parts or supplies are needed please contact us directly.

Jošt V. Lavrič
email: jlavric@bgc-jena.mpg.de
phone: +49 3641576368
fax: +49 3641 577300

E References, List of Figures and List of Tables

References

Assonov, S. and Brenninkmeijer, C. (2003a). On the ^{17}O correction for CO_2 mass spectrometric isotopic analysis. *Rapid Commun Mass Spectrom*, 17(10):1007–1016. 15

Assonov, S. and Brenninkmeijer, C. (2003b). A redetermination of absolute values for 17RVPDB-CO2 and 17RVSMOW. *Rapid Commun Mass Spectrom*, 17(10):1017–1029. 15

Battle, M., Bender, M., Hendricks, M., Ho, D., Mika, R., McKinley, G., Fan, S., Blaine, T. A., and Keeling, R. F. (2003). Measurements and models of the atmospheric Ar/N_2 ratio. *Geophys Res Lett*, 30(15). 21, 34

Bock, M., Schmitt, J., Beck, J., Schneider, R., and Fischer, H. (2014). Improving accuracy and precision of ice core δD(CH4) analyses using methane pre-pyrolysis and hydrogen post-pyrolysis trapping and subsequent chromatographic separation. *Atmospheric Measurement Techniques*, 7:1999–2012. 16

Brand, W. A. (2005). O_2/N_2 *Storage Aspects and Open Split Mass Spectrometric Determination*, pages 146–151. GAW Report 161, World Meteorological Organization. 16

Brand, W. A., Assonov, S. S., and Coplen, T. B. (2010). Correction for the ^{17}O interference in $\delta^{13}C$ measurements when analyzing CO_2 with stable isotope mass spectrometry (IUPAC technical report). *Pure and Applied Chemistry*, 82(8):1719–1733. 15

Brand, W. A., Avak, H., Seedorf, R., Hofmann, D., and Conradi, T. (1996). New methods for fully automated isotope ratio determination from hydrogen at the natural abundance level. *Isotopes in Environmental and Health Studies*, 32(2-3):263–273. 16

Brand, W. A., Huang, L., Mukai, H., Chivulescu, A., Richter, J. M., and Rothe, M. (2009). How well do we know VPDB? variability of $\delta^{13}C$ and $\delta^{18}O$ in CO_2 generated from NBS19-calcite. *Rapid Communications in Mass Spectrometry*, 23(6):915–926. 15, 16

Brand, W. A., Rothe, M., Sperlich, P., Strube, M., and Wendeberg, M. (2016). Automated simultaneous measurement of the $\delta^{13}C$ and $\delta^{2}H$ values of methane and the $\delta^{13}C$ and $\delta^{18}O$ values of carbon dioxide in flask air samples using a new multi cryotrap/gas chromatography/isotope ratio mass spectrometry system. *Rapid Communications in Mass Spectrometry*, 30(13):1523–1539. 15

Carpenter, L. J., Fleming, Z. L., Read, K. A., Lee, J. D., Moller, S. J., Hopkins, J. R., Purvis, R. M., Lewis, A. C., Mueller, K., Heinold, B., Herrmann, H., Fomba, K. W., van, Pinxteren, D., Mueller, C., Tegen, I., Wiedensohler, A., Mueller, T., Niedermeier, N., Achterberg, E. P., Patey, M. D., Kozlova, E. A., Heimann, M., Heard, D. E., Plane, J. M. C., Mahajan, A., Oetjen, H., Ingham, T., Stone, D., Whalley, L. K., Evans, M. J., Pilling, M. J., Leigh, R. J., Monks, P. S., Karunaharan, A., Vaughan, S., Arnold, S. R., Tschritter, J., Poehler, D., Friess, U., Holla, R., Mendes, L. M., Lopez, H., Faria, B. V. E., Manning, A. J., and Wallace, D. W. R. (2010). Seasonal characteristics of tropical marine boundary layer air measured at the Cape Verde Atmospheric Observatory. *J Atmos Chem*, 67:87–140. 50

Crotwell, A., Lee, H., and Steinbacher, M., editors (2019). *20th WMO/IAEA Meeting on Carbon Dioxide, Other Greenhouse Gases and Related Measurement Techniques (GGMT-2019)*. GAW Report 255, World Meteorological Organisation. 4, 9, 61

Enting, I. G. (1987). On the use of smoothing splines to filter CO_2 data. *J. Geophys. Res*, 92(D9):10977–10984. 24

Francey, R., Steele, L., Langenfelds, R., Allison, C., Cooper, L., Dunse, B., Bell, B., Murray, T., Tait, H., Thompson, L., and Masarie, K. (1998). Atmospheric carbon dioxide and its stable isotope ratios, methane, carbon monoxide, nitrous oxide and hydrogen from Shetland Isles. *Atmospheric Environment*, 32(19):3331–3338. 46

Fraser, P. J., Porter, L. W., Baly, L. W., Krummel, P. B., Dunse, P. B., Steele, L. P.and Derek, N., Langenfelds, R. L., Levin, I., Oram, D. E., Elkins, J. W., Vollmer, M. K., and Weiss, R. F. (2004). *Sulfur hexafluoride at Cape Grim: long term trends and regional emissions*, page 18-23. Bureau of Meteorology and CSIRO Atmospheric Research, Melbourne, Australia. 7

Ghosh, P. and Brand, W. A. (2004). The effect of N_2O on the isotopic composition of air-CO_2 samples. *Rapid Communications in Mass Spectrometry*, 18(16):1830–1838. 15

Ghosh, P., Patecki, M., Rothe, M., and Brand, W. A. (2005). Calcite-CO_2 mixed into CO_2-free air: a new CO_2-in-air stable isotope reference material for the VPDB scale. *Rapid Communications in Mass Spectrometry*, 19(8):1097–1119. 7, 15, 16

Heimann, M., Schulze, E., Winderlich, J., Andreae, M. O., Chi, X., Gerbig, C., Kolle, O., Kübler, K., Lavric, J., Mikhailov, E., Panov, A., Park, S., Rödenbeck, C., and Skorochod, A. (2014). The Zotino Tall Tower Observatory (ZOTTO): Quantifying large scale biogeochemical changes in central Siberia. *Nova Acta Leopoldina NF*, 117(399):51–64. 45

Hilkert, A. W., Douthitt, C. B., Schlüter, H. J., and Brand, W. A. (1999). Isotope ratio monitoring gas chromatography/mass spectrometry of D/H by high temperature conversion isotope ratio mass spectrometry. *Rapid Communications in Mass Spectrometry*, 13:1226–1230. 16

Kai, F. M., Tyler, S. C., Randerson, J. T., and Blake, D. R. (2011). Reduced methane growth rate explained by decreased northern hemisphere microbial sources. *Nature*, 476(7359):194–197. 17

Keeling, R. F., Blaine, T. A., Paplawsky, B., Katz, L., Atwood, C., and Brockwell, T. (2004). Measurement of changes in atmospheric Ar/N_2 ratio using a rapid-switching, single-capillary mass spectrometer system. *Tellus B*, 56(4):322–338. 21, 34

Keeling, R. F., Manning, A. C., Paplawsky, W. J., and Cox, A. C. (2007). On the long-term stability of reference gases for atmospheric O_2/N_2 and CO_2 measurements. *Tellus B*, 59(1):3–14. 16, 17

Keeling, R. F. and Shertz, S. R. (1992). Seasonal and interannual variations in atmospheric oxygen and implications for the global carbon cycle. *Nature*, 358(6389):723–727. 16

Langenfelds, R. L., van der Schoot, M. V., Francey, R. J., Steele, L. P., Schmidt, M., and Mukai, H. (2005). Modification of air standard composition by diffusive and surface processes. *J Geophys Res-Atmos*, 110(D13307). 21, 22, 34

Levin, I., Veidt, C., Vaughn, B. H., Brailsford, G., Bromley, T., Heinz, R., Lowe, D., Miller, J. B., Poss, C., and White, J. W. C. (2012). No inter-hemispheric $\delta^{13}CH_4$ trend observed. *Nature*, 486(7404):E3–E4. 17

Meijer, H. A. J., Neubert, R. E. M., and Visser, G. H. (2000). Cross contamination in dual inlet isotope ratio mass spectrometers. *International Journal of Mass Spectrometry*, 198(1):45–61. 15

Merritt, D. A., Brand, W. A., and Hayes, J. M. (1994). Compound-specific isotope analysis in biogeochemistry and petroleum research isotope-ratio-monitoring gas chromatography-mass spectrometry: methods for isotopic calibration. *Organic Geochemistry*, 21(6):573–583. 16

Merritt, D. A., Freeman, K. H., Ricci, M. P., Studley, S. A., and Hayes, J. M. (1995). Performance and optimization of a combustion interface for isotope ratio monitoring gas-chromatography mass-spectrometry. *Analytical Chemistry*, 67(14):2461–2473. 16

Morgan, E. J., Lavric, J. V., Arévalo-Martínez, D. L., Bange, H. W., Steinhoff, T., Seifert, T., and Heimann, M. (2019). Air–sea fluxes of greenhouse gases and oxygen in the northern Benguela Current region during upwelling events. *Biogeosciences*, 16(20):4065–4084. 51

Morgan, E. J., Lavrič, J. V., Seifert, T., Chicoine, T., Day, A., Gomez, J., Logan, R., Sack, J., Shuuya, T., Uushona, E. G., Vincent, K., Schultz, U., Brunke, E.-G., Labuschagne, C., Thompson, R. L., Schmidt, S., Manning, A. C., and Heimann, M. (2015). Continuous measurements of greenhouse

gases and atmospheric oxygen at the Namib Desert Atmospheric Observatory. *Atmos. Meas. Tech.*, 8(6):2233–2250. 51

Popa, M. E., Gloor, E., Manning, A. C., Jordan, A., Schultz, U., Haensel, F., Seifert, T., and Heimann, M. (2010). Measurements of greenhouse gases and related tracers at Bialystok tall tower station in Poland. *Atmospheric Measurement Techniques*, 3(2):407–427. 7, 47

Rothe, M., Jordan, A., and Brand, W. A. (2005). *Trace gases, ^{13}C and ^{18}O of CO_2-in-air samples: Storage in glass flasks using PCTFE seals and other effects*, pages 64–70. GAW Report 161, World Meteorological Organization. 7

Schmitt, J., Seth, B., Bock, M., van der Veen, C., Muller, L., Sapart, C. J., Prokopiou, M., Sowers, T., Röckmann, T., and Fischer, H. (2013). On the interference of Kr during carbon isotope analysis of methane using continuous-flow combustion, isotope ratio mass spectrometry. *Atmos. Meas. Tech.*, 6(5):1425–1445. 16

Schupbach, S., Federer, U., Kaufmann, P. R., Hutterli, M. A., Buiron, D., Blunier, T., Fischer, H., and Stocker, T. F. (2009). A new method for high-resolution methane measurements on polar ice cores using continuous flow analysis. *Environmental Science & Technology*, 43(14):5371–5376. 16

Sperlich, P., Moossen, H., Geilmann, H., Bury, S. J., Brown, J. C., Moss, R. C., Brailsford, G. W., and Brand, W. A. (2021). A robust method for direct calibration of isotope ratios in gases against liquid/solid reference materials, including a laboratory comparison for δ^{13}C-CH_4. *Rapid Communications in Mass Spectrometry*, 35:e8944. 17

Sperlich, P., Uitslag, N. A. M., Richter, J. M., Rothe, M., Geilmann, H., van der Veen, C., Röckmann, T., Blunier, T., and Brand, W. A. (2016). Development and evaluation of a suite of isotope reference gases for methane in air. *Atmos. Meas. Tech.*, 9(8):3717–3737. 17

Stephens, B. B., Keeling, R. F., Heimann, M., Six, K. D., Murnane, R., and Caldeira, K. (1998). Testing global ocean carbon cycle models using measurements of atmospheric O_2 and CO_2 concentration. *Global Biogeochem. Cycles*, 12(2):213–230. 18, 21

Sturm, P., Leuenberger, M., Sirignano, C., Neubert, R. E. M., Meijer, H. A. J., Langenfelds, R. L., Brand, W. A., and Tohjima, Y. (2004). Permeation of atmospheric gases through polymer O-rings used in flasks for air sampling. *J. Geophys. Res.*, 109(D4):D04309. 7

Thompson, R. L., Manning, A. C., Gloor, E., Schultz, U., Seifert, T., Haensel, F., Jordan, A., and Heimann, M. (2009). In-situ measurements of oxygen, carbon monoxide and greenhouse gases from Ochsenkopf tall tower in Germany. *Atmospheric Measurement Techniques*, 2(2):573–591. 48

Umezawa, T., Brenninkmeijer, C. A. M., Rockmann, T., van der Veen, C., Tyler, S. C., Fujita, R., Morimoto, S., Aoki, S., Sowers, T., Schmitt, J., Bock, M., Beck, J., Fischer, H., Michel, S. E., Vaughn, B. H., Miller, J. B., White, J. C., Brailsford, G., Schaefer, H., Sperlich, P., Brand, W. A., Rothe, M., Blunier, T., Lowry, D., Fisher, R. E., Nisbet, E. G., Rice, A. L., Bergamaschi, P., Veidt, C., and Levin, I. (2018). Interlaboratory comparison of δ^{13}C and δ^{2}H measurements of atmospheric CH_4 for combined use of data sets from different laboratories. *Atmospheric Measurement Techniques*, 11(2):1207–1231. 17

Verkouteren, R. M., Allison, C. E., Studley, S. A., and Leckrone, K. J. (2003a). Isotopic metrology of carbon dioxide. I. Interlaboratory comparison and empirical modeling of inlet equilibration time, inlet pressure, and ion source conductance. *Rapid Communications in Mass Spectrometry*, 17(8):771–776. 15

Verkouteren, R. M., Assonov, S., Klinedinst, D. B., and Brand, W. A. (2003b). Isotopic metrology of carbon dioxide. ii. effects of ion source materials, conductance, emission, and accelerating voltage on dual-inlet cross contamination. *Rapid Communications in Mass Spectrometry*, 17(8):777–782. 15

Wendeberg, M., Richter, J. M., Rothe, M., and Brand, W. A. (2011). ^{18}O anchoring to VPDB: calcite

digestion with ^{18}O-adjusted ortho-phosphoric acid. *Rapid Communications in Mass Spectrometry*, 25(7):851–860. 16

Wendeberg, M., Richter, J. M., Rothe, M., and Brand, W. A. (2013). Jena reference air set (JRAS): a multi-point scale anchor for isotope measurements of CO_2 in air. *Atmos. Meas. Tech.*, 6(3):817–822. 15, 16, 17

Werner, R. A. and Brand, W. A. (2001). Referencing strategies and techniques in stable isotope ratio analysis. *Rapid Communications in Mass Spectrometry*, 15(7):501–519. 16

Werner, R. A., Bruch, B. A., and Brand, W. A. (1999). ConFlo III, an interface for high precision δ^{13}C and δ^{15}N analysis with an extended dynamic range. *Rapid Communications in Mass Spectrometry*, 13:1237–1241. 17

Werner, R. A., Rothe, M., and Brand, W. A. (2001). Extraction of CO_2 from air samples for isotopic analysis and limits to ultra high precision δ^{18}O determination in CO_2 gas. *Rapid Communications in Mass Spectrometry*, 15(22):2152–2167. 15

List of Figures

List of Tables