InGOS - Integrated non-CO_2 Observing System
FP7 Integrating Activity
Grant agreement no 284274

Deliverable D3.3			
Title	First ICP report on compatibility of CH4, N2O, SF6 and H2 measurements within the InGOS Network		
Delivery date Annex 1	Month 24		
Actual delivery date	Month 30	09/05/2014	
Lead participant	Work package	Nature	Dissemination level
UEA	**NA3**	**report**	**Public**

Deliverable description:
A detailed report on the current state of compatibility of the InGOS atmospheric network of greenhouse gas measurements, including summary graphs and tables of results received, both during the InGOS period as well as in the years pre-dating InGOS when the Cucumbers programme was running.

Compatibility of atmospheric greenhouse gas measurements in the InGOS network as assessed by the 'Cucumbers' intercomparison programme

<u>InGOS Deliverable 3.3:</u> First ICP report on compatibility of CH_4, N_2O, SF_6 and H_2 measurements within the InGOS network

A. C. Manning[1], M. Schmidt[2,3], A. Jordan[4], I. Levin[3], M. Hewitt[1], A. Etchells[1], P. A. Pickers[1], S. Hammer[3], M. Delmotte[2], A. Vermeulen[5], M. Alemanno[6], C. Allison[7], F. Apadula[8], J. Arduini[9], Z. Barcza[10], W. Brand[4], H. Chen[11], M. Ernst[12], B. D. Hall[13], L. Haszpra[14], J. Hatakka[15], L. Huang[12], K. Katsumata[16], V. Kazan[2], P. Keronen[17], R. Langenfelds[7], J. Lavric[4], M. C. Leuenberger[18], M. Lopez[2], T. Machida[16], F. Meinhardt[19], J. Moncrieff[20], J. Necki[21], S. O'Doherty[22], S. Piacentino[23], J. Richter[4], L. Ries[19], M. Rothe[4], H. A. Scheeren[11], M. F. Schibig[18], L. P. Steele[7], I. Szilagyi[24], B. Wastine[2], P. A. Wilson[1], D. E. Worthy[12], C. Yver[2]

[1] Centre for Ocean and Atmospheric Sciences, School of Environmental Sciences, University of East Anglia, United Kingdom (UEA/ENV)
[2] Laboratoire des Sciences du Climat et de l'Environnement, France (LSCE)
[3] Institut für Umwelphysik, Heidelberg University, Germany (UHEI)
[4] Max Planck Institute for Biogeochemistry, Germany (MPI-BGC)
[5] Dept. Air Quality, Energy Research Centre of the Netherlands, The Netherlands (ECN)
[6] Italian Air Force Meteorological Service, Italy (IAFMS)
[7] Centre for Australian Weather and Climate Research / CSIRO Marine and Atmospheric Research, Australia (CMAR)
[8] Ricerca sul Sistema Energetico, Italy
[9] Department of Basic Sciences, Università degli Studi di Urbino, Italy
[10] Eotvos Lorand University, Hungary
[11] Centre for Isotope Research, University of Groningen, The Netherlands (CIO-RUG)
[12] Climate Research Division, Environment Canada, Canada (EnCan)
[13] National Oceanic and Atmospheric Administration, USA (NOAA)
[14] Hungarian Meteorological Service, Hungary and Hungarian Academy of Sciences (MTA) Research Centre for Astronomy and Earth Sciences, Hungary
[15] Finnish Meteorological Institute, Finland (FMI)
[16] National Institute for Environmental Studies, Japan (NIES)
[17] University of Helsinki, Department of Physics, Division of Atmospheric Sciences, Finland (UHEL)
[18] Physics Institute and Oeschger Centre for Climate Change Research, University of Bern, Switzerland
[19] Umweltbundesamt, Germany (UBA)
[20] School of GeoSciences, University of Edinburgh, United Kingdom
[21] AGH - University of Science and Technology, Poland (AGH-UST)
[22] School of Chemistry, University of Bristol, United Kingdom (UBRIS)
[23] National Agency for New Technologies, Energy and Sustainable Economic Development, Italy (ENEA)
[24] Hungarian Academy of Sciences (MTA) Research Centre for Natural Sciences, Hungary

Published: 09 May 2014

Disclaimer:

All data presented in this report are preliminary and subject to change, especially as central laboratories and field stations revise their internal calibration scales based on new information, for example, adjusted values received for their primary calibration standards from the World Meteorological Organization Central Calibration Laboratory (WMO CCL).

Description and aims of the programme

The 'Cucumbers' intercomparison programme was originally initiated as part of the European Union (EU) 'CarboEurope' project in 2005, and was continued in the EU 'Infrastructure for Measurements of the European Carbon Cycle' (IMECC) project from 2008 – 2011, and in the EU 'Integrated Non-CO_2 Greenhouse gas Observing System' (InGOS) project since the end of 2011. Although there were several other intercomparison programmes included in the EU CarboEurope project, these other programmes were to some extent limited in that they only compared central laboratories, and did not compare the many atmospheric CO_2 monitoring field stations within Europe. The Cucumbers intercomparison programme, however, comprised of 21 field stations within Europe (initially) in addition to nine central laboratories (including four non-European international partners), thus addressing this limitation.

The overarching objective of atmospheric measurement intercomparison programmes is to assess and quantify potential offsets in calibration scales, and consequently in ambient air measurements, between different laboratories and field stations. Specifically, the three main aims of such programmes are to:

1) document the state of compatibility so that laboratories and field stations become more aware of where problems lie, and thus are able to address their analytical procedures with the expectation of reducing calibration scale offsets progressively;
2) contribute to an estimation of uncertainties in the field station data sets; and
3) achieve improved scientific interpretation of data sets by taking into account the quantified calibration scale offsets.

An example of the latter objective is where the quantified scale offsets are incorporated in inverse atmospheric transport modelling exercises to determine fluxes from mole fraction measurements. Peters et al. [2010] carried out such an exercise, examining the sensitivity of CO_2 fluxes derived from inverse modelling of CO_2 measurements at European field stations, when applying relative offsets to the measurements as defined by the Cucumbers programme. A conceptually more simplistic example is to consider atmospheric measurements of a given species made from two locations, say Mace Head, on the west coast of Ireland, and Białystok, in eastern Poland. If the measurements at Białystok are elevated with respect to Mace Head, then given the prevailing westerly winds, this would suggest a net source for that gas species over western Europe. If, however, owing to analytical artefacts, Białystok measurements are biased high with respect to Mace Head, then the size of the western European source would be overestimated. By quantifying such possible biases (offsets) and incorporating them into our use of data sets, the Cucumbers programme allows us to achieve more accurate scientific interpretation of our measurements, in this case a more accurate estimate of source fluxes over western Europe.

It is very important to note, however, that results from intercomparison exercises such as the Cucumbers programme should not be used to 'correct' a field station's data set, nor to merge together data sets from different field stations by applying such 'corrections'.

A 'Cucumber' is a 20 L Luxfer aluminium cylinder filled initially to high pressure (approximately 200 bar) with dry, real air. The Cucumbers programme has consisted of seven loops, each of which comprises three Cucumbers (see Fig. 1; station name codes are defined in Appendix 1). Of these seven loops, there are five 'Euro' loops that each comprise of a European central laboratory and several field stations. The remaining two loops, known as 'Inter' loops, comprise only of central laboratories, where Inter-1 is a faster loop with only European laboratories, and Inter-2 is a slower loop that also includes participants from USA, Canada, Australia and Japan. This design of five 'Euro' loops linked together by the two 'Inter' loops enables laboratories and field stations to be compared over much shorter timescales than would otherwise be possible, and than is possible through the WMO/GAW 'Round Robin' intercomparison programme, for example.

In the unusual case of the Inter-2 loop, Cucumbers are 29.5 L Luxfer aluminium cylinders, initially filled with real air to a pressure of about 130 bar. These larger cylinders are manufactured in the USA, whereas all the 20 L cylinders used in the other loops are manufactured in the UK. The vast majority of cylinders manufactured in Europe are not 'DOT-certified', and therefore can not be shipped to the USA, thus requiring the use of different cylinders for the Inter-2 loop.

The Cucumbers were initially prepared and filled at MPI-BGC, and where feasible, mole fractions were adjusted so that each trio of Cucumbers within a loop spanned a range in mole fraction similar to that observed at background atmospheric field stations in the northern hemisphere. Cucumbers were not calibrated, but initial measurements were made at MPI-BGC to ensure that all mole fractions were at acceptable values. For each trio of Cucumbers in a given loop, further initial measurements were made by the relevant central laboratory for that loop (for many loops this was MPI-BGC again), so as to establish initial mole fraction values for each species in each Cucumber. Following these initial measurements at the central laboratories, Cucumbers began their rotation amongst the field station participants in each loop.

Within each loop, the Cucumbers cycle perpetually between field stations and laboratories, and are analysed at each location according to the procedures detailed in Appendix 2. Measurement results are then sent to UEA, where they are incorporated into graphs of local calibration scale offsets that are displayed on a dedicated website at http://www.cucumbers.uea.ac.uk, viewable by all participants in the Cucumbers programme. This online dissemination of results allows field stations and laboratories to identify potential problems with their measurement systems, providing a vital, independent and rigorous method of quality control of the *in situ* records collected at the field stations.

It is important to realise a key limitation in this quality control methodology, however, namely that the air from the Cucumbers does not pass through the inlet tubing, sample pumps and drying systems present at the field station. This means that if any contamination or mole fraction artefacts are introduced by these parts of the system, then they will not be picked up by the Cucumbers programme. Thus the Cucumbers programme quantifies calibration scale offsets, which are informative about offsets in ambient air measurements, but may not incorporate all of the latter offsets.

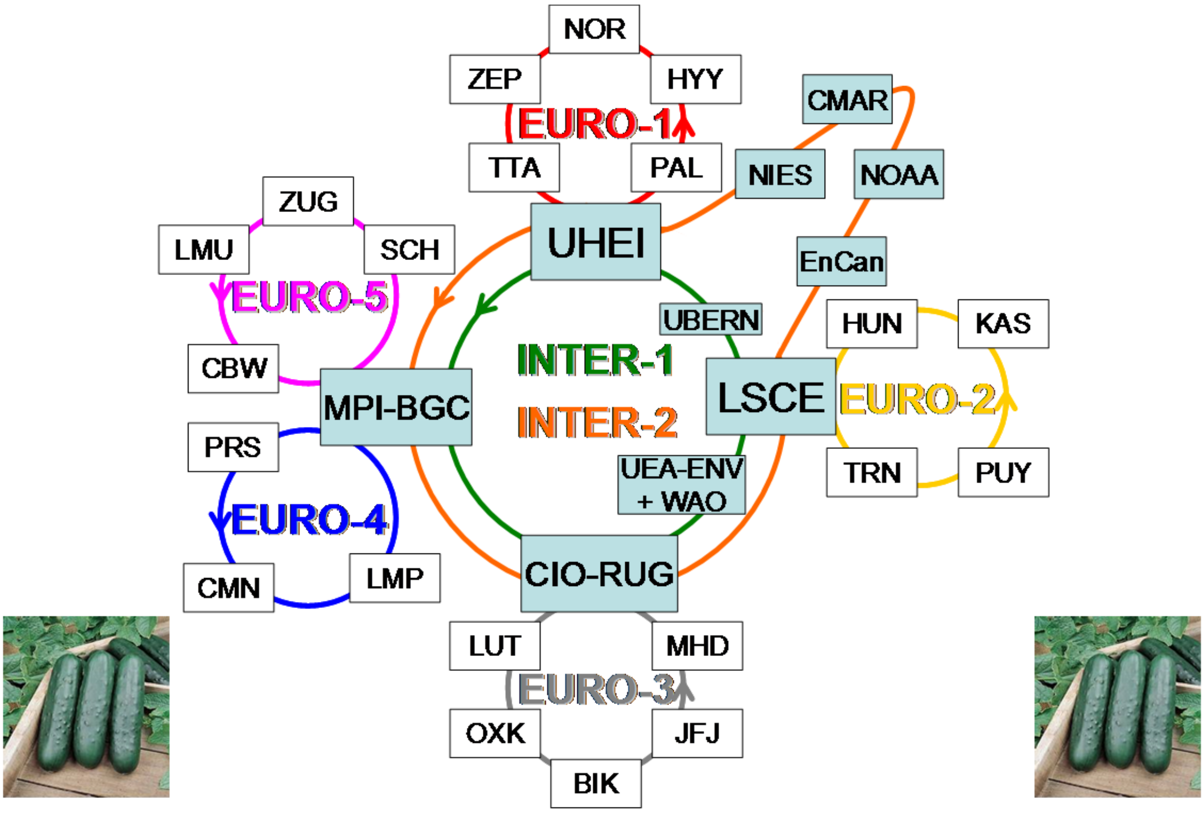


Fig. 1. Schematic of the original seven Cucumber loops. Station name codes are defined in Appendix 1. Note that although CIO-RUG are responsible for the Euro-3 loop, these data are reported as offsets from MPI-BGC, not CIO-RUG. MPI-BGC were added to the Euro-3 loop (not shown) because the CIO-RUG laboratory does not have the capability to measure all the species reported from some of the field stations (e.g. N_2O).

The Cucumbers programme was originally established in 2005 specifically for the purpose of CO_2 mole fraction intercomparison. It was later expanded to include the comparison of eight additional atmospheric species: CH_4, CO, N_2O, SF_6, H_2, O_2/N_2, $\delta^{13}C$-CO_2, $\delta^{18}O$-CO_2, with stations and laboratories analysing these additional species as their analytical capabilities and time allow. Because of the original focus on CO_2, however, this meant that the trios of Cucumbers within a given loop do not typically span an appropriate range in mole fraction (or isotope ratio) in all of the species, and in a few cases, mole fractions were well outside of typical ranges experienced in background northern hemisphere air. Key examples are a few Cucumbers with N_2O mole fractions in the range 295-310 ppb (ambient is ~325 ppb) and SF_6 mole fractions in the range 10-15 ppt (ambient is ~8 ppt). The unfortunate consequence of this is that for some species, a few Cucumbers are outside of the calibration ranges of the stations and laboratories within the loop, thus unfairly influencing compatibility results for those stations and laboratories. To address this problem, we have ignored all data reported from a few Cucumbers for given species, when we consider the mole fraction (or isotopic ratio) of that species to be outside of typical calibration ranges. The species ignored in specific Cucumbers are clarified in the Results section below.

The design of the loops was decided primarily based on the geographic location of the field stations, in an attempt to minimise shipping costs, but it also partially takes into account which species are measured at each field station and laboratory. For example, the Euro-3 loop was specifically designed to include all field stations that make O_2/N_2 measurements. One O_2/N_2 field station, WAO, which was added at a later date, was unable to be included in Euro-3 without making the loop too large, thus it was included in the Inter-1 loop, which also reports O_2/N_2 results. The results from all nine atmospheric species are displayed graphically on the Cucumbers website.

<u>Progress – Historical</u>

Circulation of all the Cucumbers initially began in October 2005, with the following deadlines in place: Cucumbers are to be analysed within 4 weeks of receipt and then shipped to the next participant, and data are to be reported within a subsequent 4 weeks. Shipping times are not included in the deadlines. In general, analysis deadlines have been reasonably well adhered to, but unfortunately reporting deadlines have been widely disregarded. This situation improved following the launch of the website in early 2008, but then deteriorated significantly again at the conclusion of the IMECC project and beginning of the InGOS project. This problem remains the most serious detriment to the success of the programme, and is discussed further in Apendix 1, section D).

There were several unfortunate logistical problems that caused disruption to the programme in 2005 and 2006, such as repeated damage to very expensive regulators, cylinder damage during shipping necessitating a replacement to be prepared, changes to international shipping regulations regarding the custom-made wooden shipping boxes that we used, and finally, concerns over regulations regarding the use of US-manufactured high pressure cylinders in Europe. This last issue resulted in the programme being suspended in August 2006, and 21 new European replacement cylinders had to be purchased, prepared, filled and calibrated. This disruption caused a major delay to the programme, which was re-instated only in January 2008. The international loop, Inter-2, was not suspended, however, and now consists of a continuous dataset spanning almost nine years between eight laboratories.

Since resuming in 2008, the programme has run much more smoothly, and still continues, in spite of the fact that the original CarboEurope project and the successor IMECC project have concluded.

<u>Progress – Upgrades in InGOS:</u>

A number of upgrades have been carried out in the Cucumbers programme since the beginning of InGOS in Oct2011. The most time-consuming aspect of running the programme has been dealing with errors and inconsistencies in the logsheets submitted to UEA. Between 50-75% of logsheets contain errors, ranging from:

- incorrect participant ID;
- incorrect cylinder ID;
- missing data for example no analysis dates or no cylinder pressures;

  - logsheets submitted with incorrect data formats, for example using American-style dates;
  - submission of corrupt logsheets, for example containing many logsheets on separate worksheet tabs;
  - populating rows and columns with data for the wrong gas species;
  - confusing which cylinder ID is which when reporting data; and
  - inconsistent analysis dates when submitting revisions of previously submitted data.

We have effected several procedural changes to try to reduce the workload in processing these logsheets, including:

- Emails to all participants stating the most common data entry errors to be wary of;
- Correcting logsheets at UEA, then sending them back to each participant, so they have an example of a correct logsheet;
- Adding detailed automated error checking in our logsheet processing software;
- Adding software to produce detailed automated report logs of the errors found in the logsheets, reducing the need for human checking;
- Developing new software to process an ASCII version of the Excel logsheet, in the hope that participants would then automatically generate the logsheets, thus reducing human error at the point of creating the logsheets.

These steps mitigated the problem, but unfortunately it still continued to be a large time sink: emails and versions of the correct logsheets appeared to be ignored, and the ASCII version of the logsheet was declared unfit for purpose. Thus, very recently, two new steps have been taken which we optimistically anticipate will greatly improve the problem:

- All processed, latest revision logsheets are now available to all participants on the website. This removes the need for participants to be organised with their logsheets, as UEA maintains the logsheet database, and makes it available to all.
- New software is being developed to process a CSV input file instead of Excel logsheets. The CSV input file has been designed from the perspective of InGOS and ICOS file formats, thus will be robust to future changes. The CSV file contains different information from the Excel logsheets, thus it requires some complex programming to 'massage' the data into the existing database design.

It is anticipated that in a future ICOS-Cucumbers programme, all of the above problems will disappear, since all data submission from ICOS field stations will be fully automated. It must be recognised, however, that the custom programming required to accept, store, and present these data in ICOS will be complex. Furthermore, presenting the data both graphically and in tabular format is the second biggest current time demand in the Cucumbers programme. The programming code has been frequently upgraded to cope with changes in cylinders within a loop, changes in participants within a loop, decisions to ignore certain cylinder results for certain gas species, and even seemingly trivial matters such as a field station changing their designated three-letter code identifying their station.

Other upgrades implemented under InGOS include providing on the website plots of cylinder pressure for each loop, providing summary data files of all processed results, and upgrading the field station summary plots by removing the error bars (the plots were becoming too busy), and adding new plots showing separately standard deviation of the average offsets.

Results

The main results from the field stations since the programme re-started in 2008 are shown in Figures 2 to 8 below. Each of these figures consists of two plots: a) shows the mean difference or offset (of all three Cucumbers in a given loop) of the field station analyses from the initial central laboratory analyses; and b) shows the standard deviation of the mean difference of the three Cucumber analyses. Note that a field station may make several analyses of a single Cucumber, but if so, only the mean results are reported to the programme, and while there would also be a standard deviation associated with these repeated measurements from one Cucumber, we do not show these in the figures in this report (other figures available on the website do display these standard deviations). Rather, we show the standard deviation of the mean of the three differences from the central laboratory (one difference for each Cucumber). 'Zero' in part a) of these figures is defined as the initial analyses (of all three Cucumbers) made at the central laboratory. Subsequent analyses at the central laboratory are also shown, so that possible drifts in mole fraction of one or more of the Cucumbers can be identified. Subsequent central laboratory analyses can also highlight potential calibration scale issues at the central laboratory itself, as well as providing information on reproducibility at the central laboratory. Lines connect the data points from a single field station (or laboratory), and show the history of analyses by that station in the same loop. The central laboratory data are denoted by red symbols and lines for all loops. The World Meteorological Organization (WMO) inter-laboratory compatibility goals are illustrated with horizontal dashed lines (also summarised in Appendix 3).

As an example of how to interpret the figures, in Fig. 2, the BIK station (Białystok. Poland; cyan squares) is in the Euro-3 loop with MPI-BGC as central laboratory, and in Sep2007 measured an average offset of 1.7 ppb from the initial MPI-BGC measurements, with a standard deviation of the average offset of ±0.4 ppb. Y-axis ranges in Figures 2 to 8 have been drawn consistently in that all average offset figures are symmetrical about zero, and all standard deviation figures exhibit exactly half the y-axis range of the accompanying average offset figure.

It is important to note that it is not possible to directly compare all of the field stations with each other from these figures. For example, Fig. 3 shows that average offsets for TRN (Trainou, France; green asterisks) are relative to LSCE as central laboratory, whereas BIK average offsets are relative to MPI-BGC. The BIK and TRN offsets can only be compared by also considering the offset of LSCE relative to MPI-BGC. Such comparisons are possible indirectly, however, owing to the existence of the Inter-1 and Inter-2 loops, where offsets between the participating central laboratories are established. A compilation of both direct and indirect comparisons of all field station offsets relative to both their own central laboratory and to MPI-BGC is shown later in this report in Tables 1 and 3.

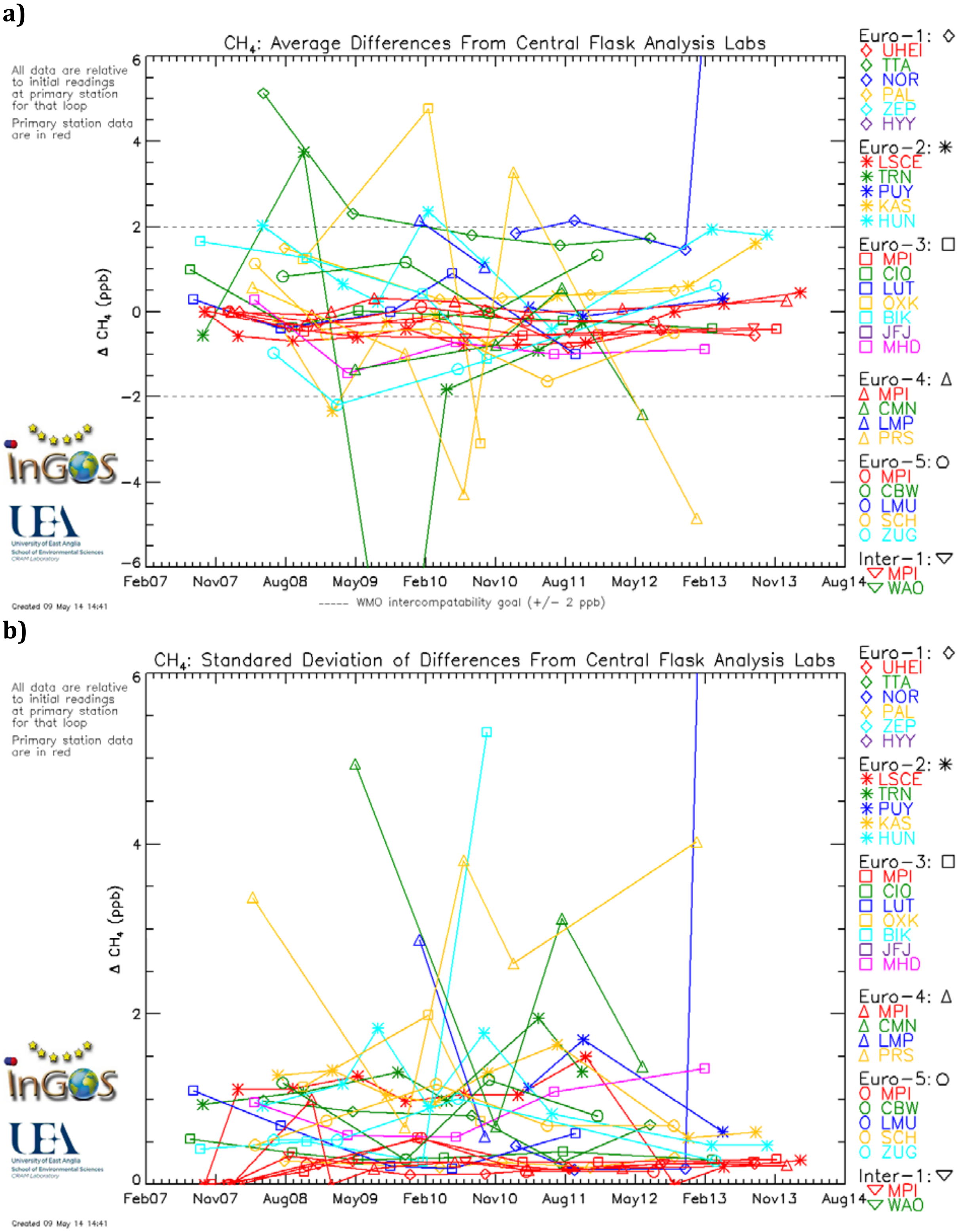


Fig. 2. CH$_4$ field station average differences (a) and standard deviation of differences (b) from central laboratories. Most of the field station analyses and all of the central laboratory analyses are within the WMO compatibility goal and the standard deviation data are also good with respect to this goal. There is no evidence of instability of CH$_4$ mole fraction in any of the cylinders over time. TTA had known instrumental problems during their initial analyses in 2008, which were later resolved.

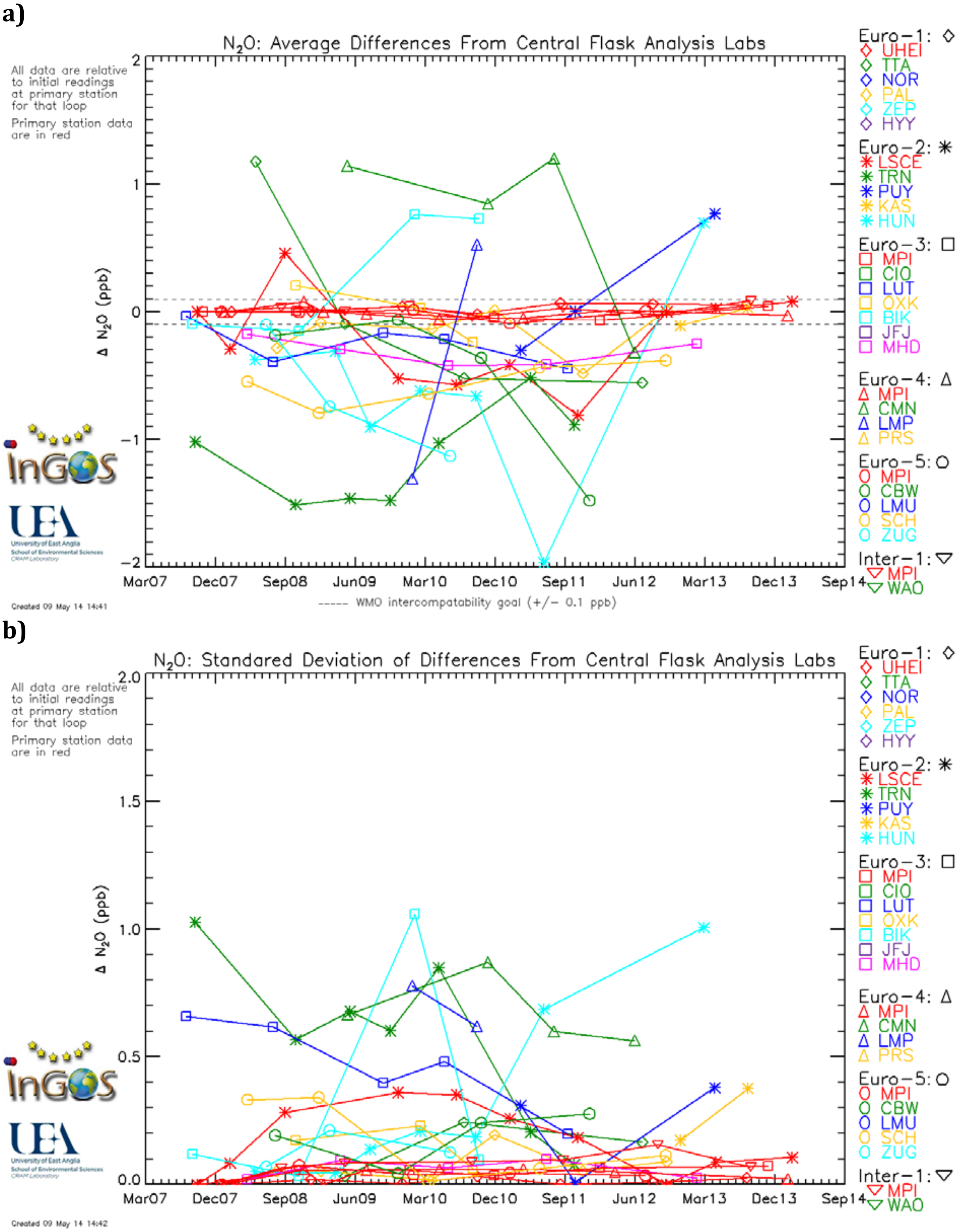
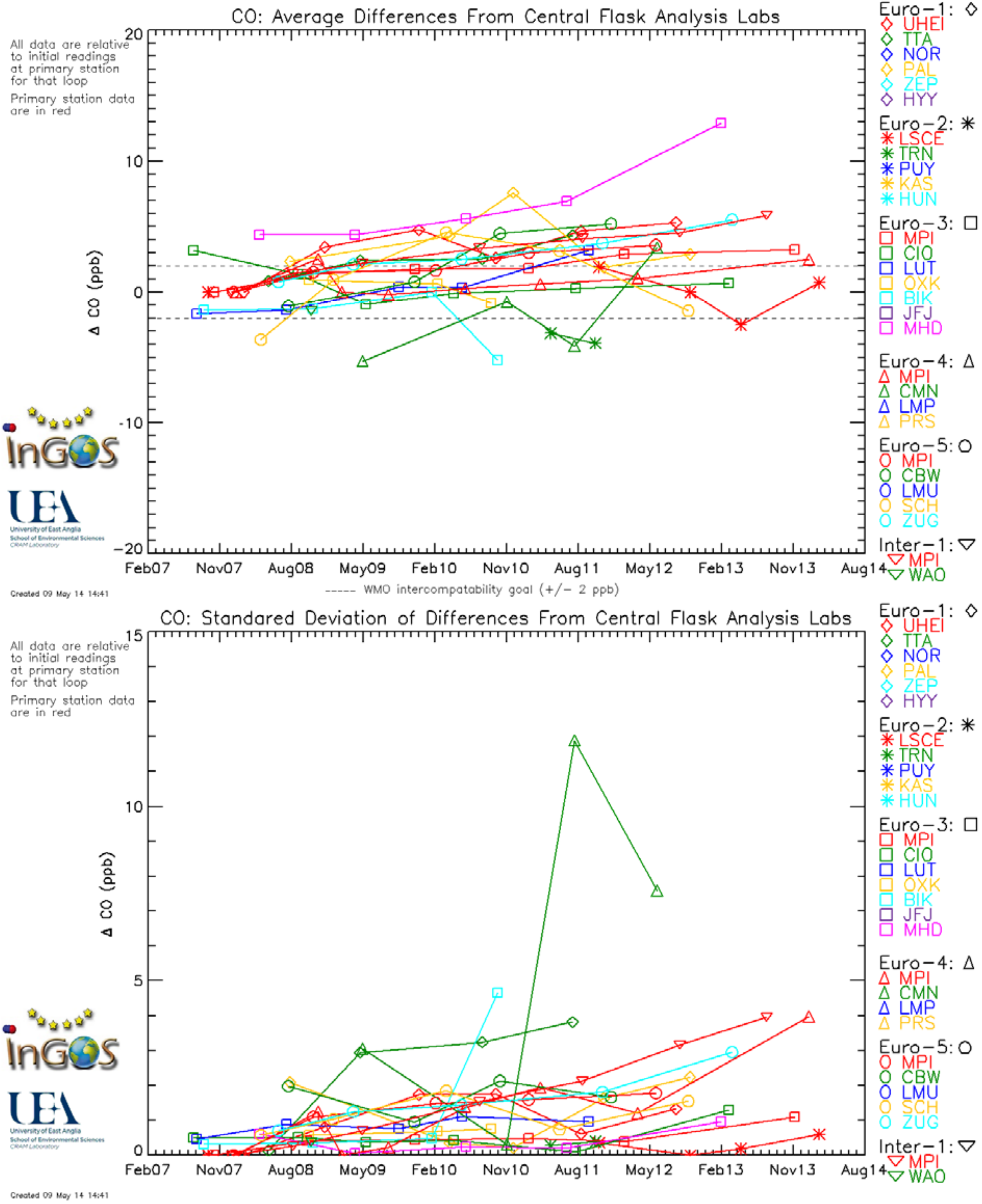


Fig. 3. N_2O field station average differences (a) and standard deviation of differences (b) from central laboratories. All but one of the central laboratories (LSCE) show results that consistently lie within the WMO compatibility goal, however, not many of the field stations are able to meet this goal. The standard deviation data of the field stations is also very poor with respect to the WMO goal. The data do not show any evidence of instability of N_2O mole fraction in any of the cylinders over time. Results from only two cylinders are shown in each of Euro-1 and Euro-5 loops, since the third cylinder in these loops had an N_2O mole fraction significantly

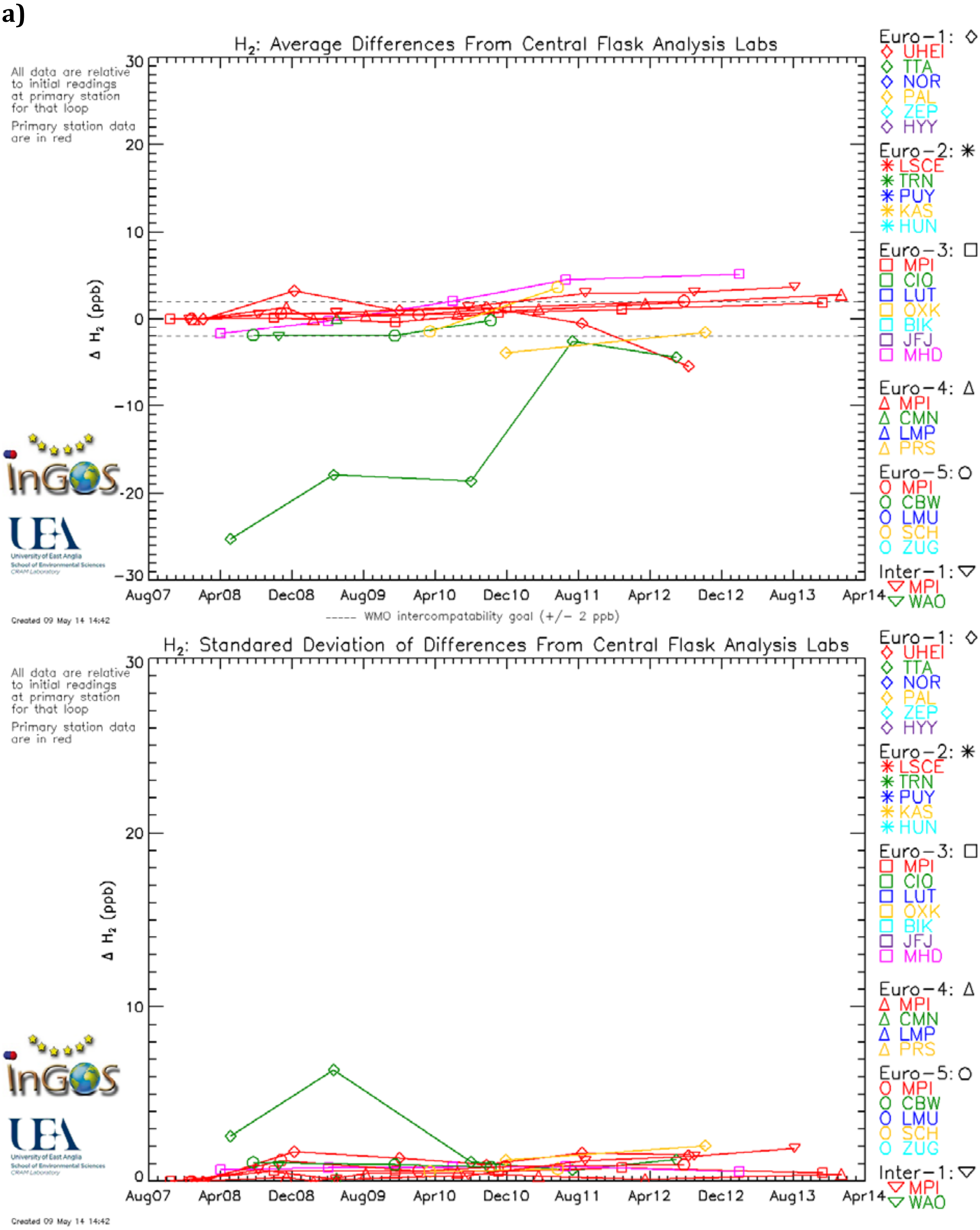
different from ranges experienced in ambient air. As with CH_4, TTA had known instrumental problems during their initial analyses in 2008, which were later resolved.

a)



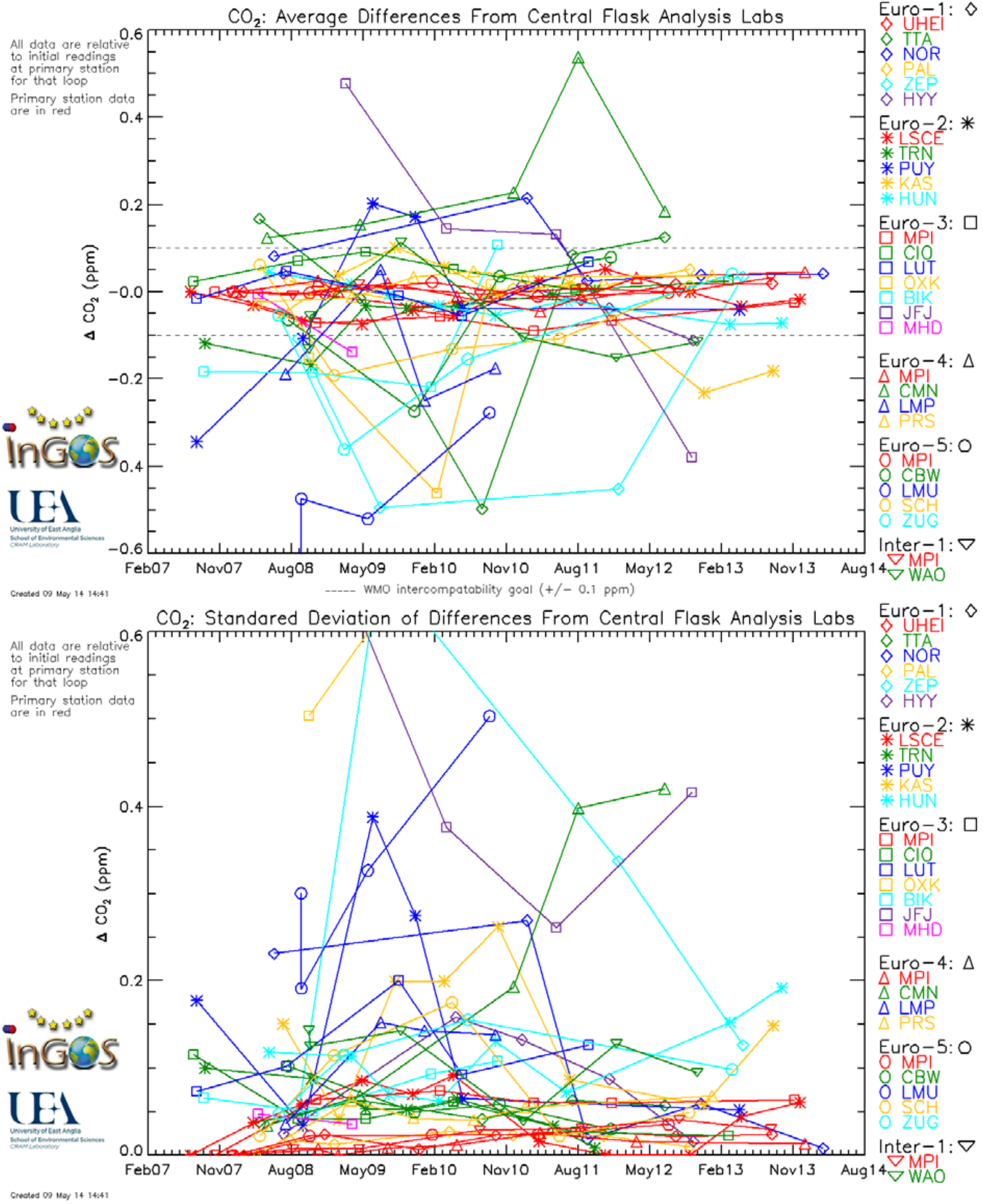
b) Fig. 4. CO field station average differences (a) and standard deviation of differences (b) from central laboratories. Figure (a) demonstrates a consistent upwards drift in the CO offset of the UHEI and MPI-BGC central laboratories over time, suggesting instability in either the Cucumber mole fractions or the laboratory calibration scales. Consequently, we are unable to determine if any of the central laboratories or field stations meet the WMO compatibility goal for CO. The fact that there also appears to be an increasing trend in standard deviation values at MPI-BGC and possibly UHEI (Fig. 4b), suggests that the problem is not with laboratory scales but rather with instability in the CO mole fractions in one or more of the cylinders. This is because if CO were unstable in the cylinders, we would only expect constant standard deviations if the CO mole fraction in all three cylinders changed by exactly the same absolute amount over time, which is unlikely (although not impossible). We would expect varying standard deviations over time if one or more of the three cylinders behaved differently from the others. As a specific example, if one cylinder exhibited a continually increasing CO mole fraction while the other two exhibited approximately constant CO mole fraction, then we would observe an incrementally increasing standard deviation over time. PAL results (Euro-1) are negatively influenced because one Cucumber (with a CO mole fraction of 350 ppb) is outside of their calibration range.

Continued on next page…



b) Fig. 5. H_2 field station average differences (a) and standard deviation of differences (b) from central laboratories. In general, the central laboratories are all performing quite well with respect to the WMO compatibility goal, as are some of the field stations. Recent analyses show a possible upwards drift in the UHEI and MPI-BGC analyses; this drift is small, however, and so might not be indicative of any cylinder instability. Further analyses and monitoring will provide greater insight into this feature in the data. In the Euro-4 loop, prior to Jan2009, results from only two Cucumbers are shown, since the third cylinder appeared to be unstable for H_2 mole fraction (and also CO_2 mole fraction). As with CH_4, TTA had known instrumental problems during their initial analyses in 2008, which were later resolved.

a)



b) Fig. 6. CO$_2$ field station average differences (a) and standard deviation of differences (b) from central laboratories. All the central laboratories and the majority of the field stations are performing well, with results within the WMO compatibility goal. However, even though the majority of field stations meet the WMO compatibility goal, when one examines their average offset from all three Cucumbers, many of them exhibit a large standard deviation of the average, often several times larger than the WMO compatibility goal. Further investigation is needed on a case by case basis, for example, to determine if there are mole fraction-dependent offsets. In the Euro-4 loop, prior to Jan2009, results from only two Cucumbers are shown for, since the third cylinder appeared to be unstable for CO$_2$ mole fraction (and also H$_2$ mole fraction). Some of the data shown as JFJ and ZEP were in fact from analyses made in the University of Bern and Stockholm University laboratories respectively.

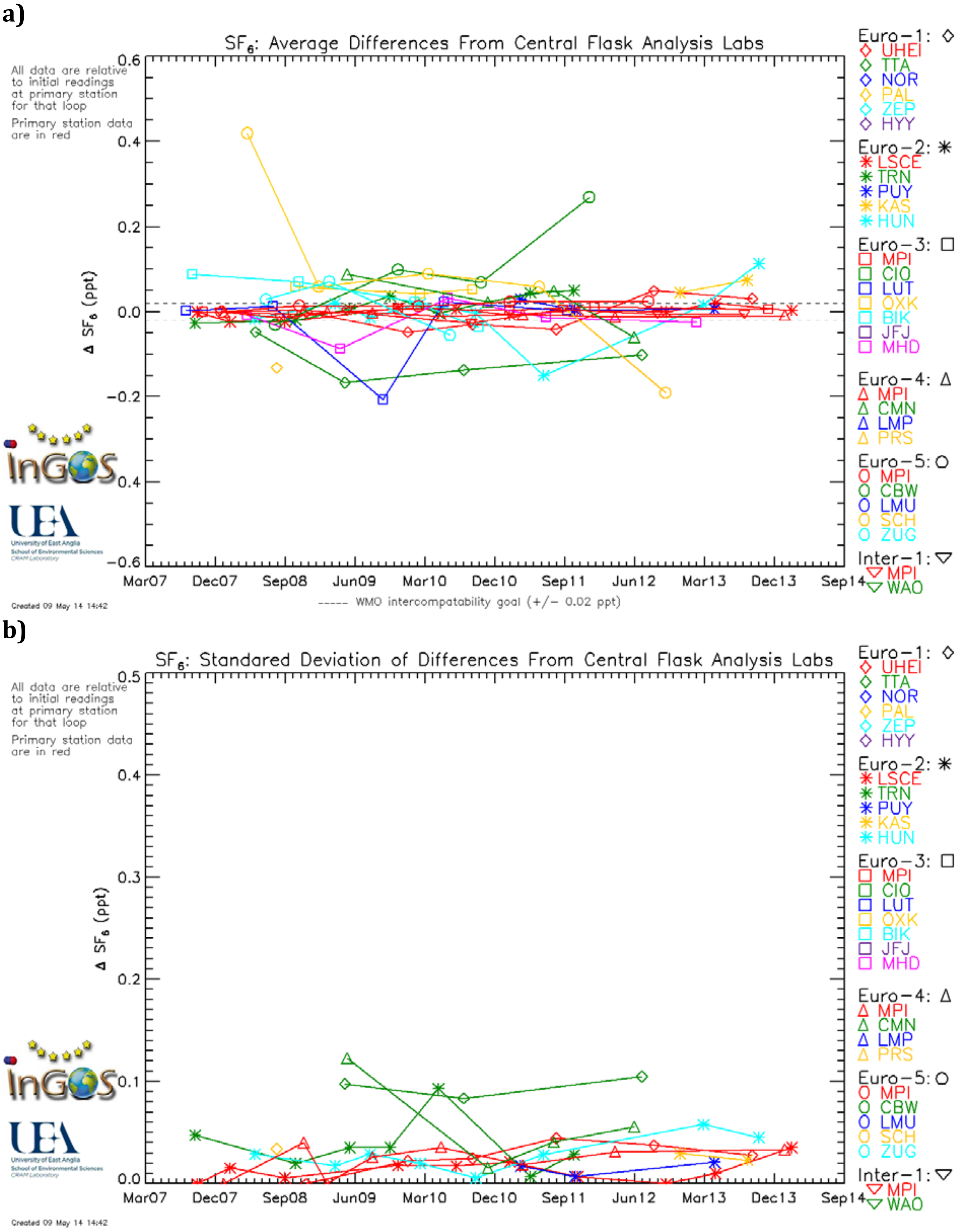
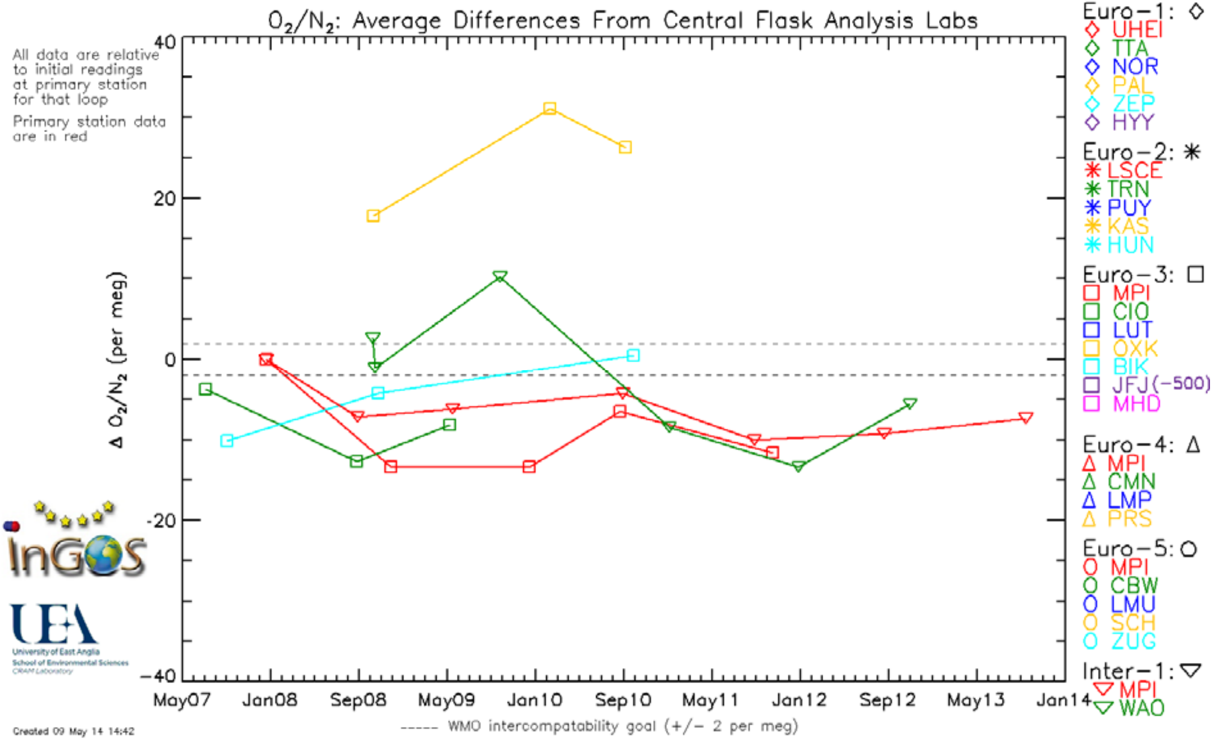


Fig. 7. SF$_6$ field station average differences (a) and standard deviation of the differences (b) from central laboratories. The central laboratory results are generally within or close to the WMO compatibility goal, however, few of the field station results meet this goal. There have been some improvements during the period of the programme, however; for example, SCH data were off-scale in 2008, but have since improved significantly, and their latest results are close to the WMO compatibility goal, although admittedly with a standard deviation about ten times

larger than this goal. BIK and KAS have also improved over time. There is no evidence of drifting SF_6 mole fraction within the Cucumbers. Results are shown from only two cylinders in the Euro-1 loop and one cylinder in each of Euro-3 and Euro-5 loops, since the other cylinders in these loops had SF_6 mole fractions significantly different from ranges experienced in ambient air.

a)



b)

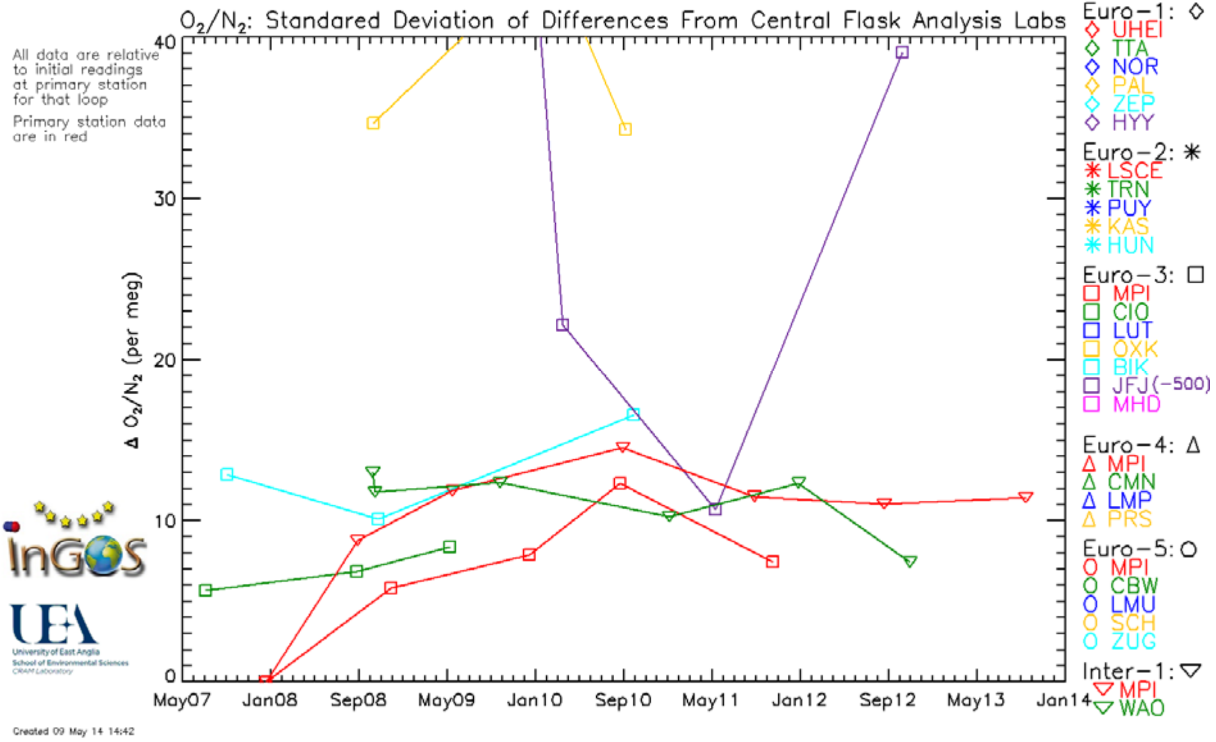


Fig. 8. O_2/N_2 field station average differences (a) and standard deviation of differences (b) from central laboratories. Initially, it was thought that the central laboratory scales and/or Cucumber O_2/N_2 ratios might be unstable, but further analyses have shown that this is likely not the case. None of the central laboratories or field stations have results that are consistently

within the WMO compatibility goal. The data points shown as JFJ were in fact from analyses made in the University of Bern laboratory.

Ideally, the three Cucumbers of each loop span a range of mole fraction for each species, in which case one can also look for possible mole fraction-dependent offsets. These cannot be seen in the summary plots of Figures 2-8, but additional figures available on the Cucumbers website demonstrate such offsets if they exist. For example, Fig. 9 shows the individual CH_4 analyses made at NOAA (National Oceanic and Atmospheric Administration) in the Inter-2 loop. All three sets of analyses show evidence of a small mole fraction-dependent offset of NOAA data relative to MPI-BGC. In Fig. 9, and all similar figures on the website, the Cucumbers are colour-coded based on mole fraction, with the sequence red, green, blue indicating decreasing mole fraction. In the specific case of Fig. 9, the CH_4 mole fractions are nominally 1981 ppb, 1894 ppb and 1883 ppb. The relative differences in mole fraction between these three Cucumbers agree very well with the offsets observed between NOAA – MPI-BGC shown in Fig. 9 and the hypothesis that these offsets are mole fraction dependent. . This example highlights just one other type of additional information that can potentially be gleaned from the Cucumbers programme. This example also highlights the importance of each trio of Cucumbers spanning a range in mole fraction, but unfortunately this has not always been possible for every species in every loop.

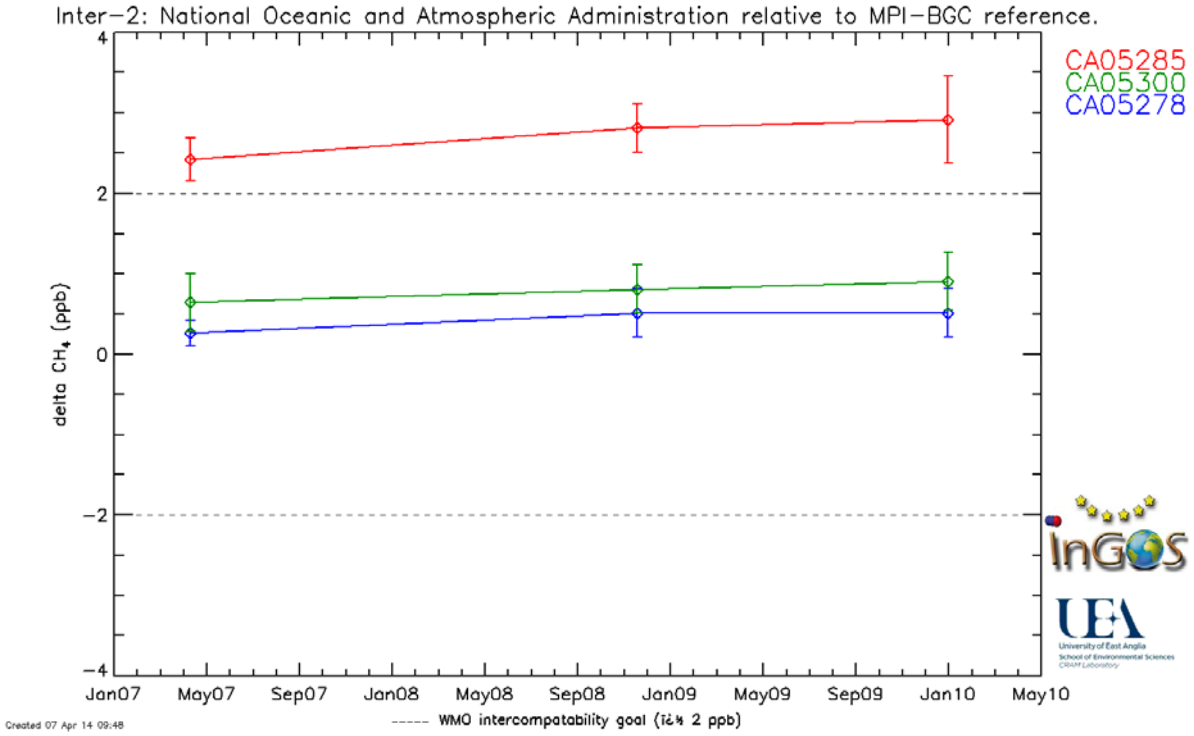


Fig. 9. Difference of individual Cucumber analyses for CH_4 from NOAA (Inter-2 loop) relative to the initial MPI-BGC analysis in Sep2005.

Figures 2 to 8 shown above present the offsets found at European field stations relative to the central laboratory of each loop. In contrast, figures 10 to 17 below present the inter-laboratory offsets, including those from the four non-European partners, over more than eight years from the combined results of the Inter-1 and Inter-2 loops. As with Figures 2 to 8, Figures 10 to 17 have y-axis ranges that are symmetrical around zero.

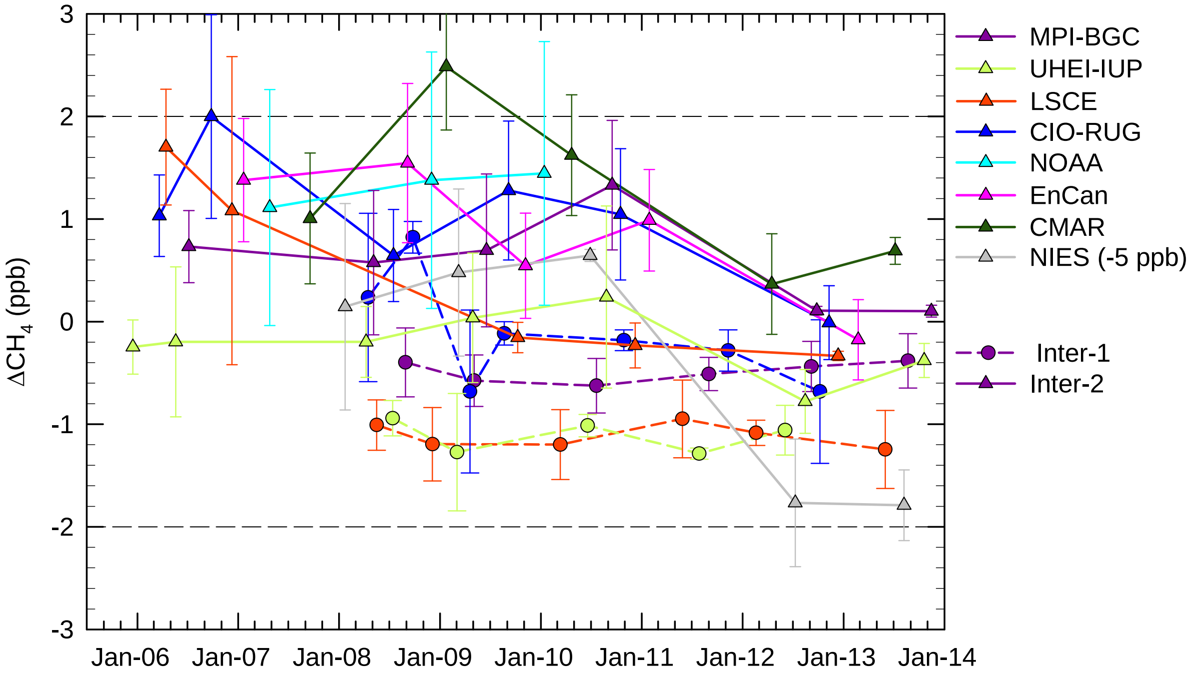


Fig. 10. Central laboratory CH_4 differences from MPI-BGC in the Inter-1 and Inter-2 loops. Dashed lines are Inter-1; solid lines are Inter-2. Horizontal black dashed lines indicate the WMO compatibility goal. Almost all of the central laboratory analyses are within the WMO compatibility goal. All data are on the NOAA04 scale, with the exception of NIES, who are on their own gravimetric scale. Although the Japanese gravimetric scale has a large offset from the NOAA scale, the offset has proven to be very stable over time, as demonstrated previously by the 'Sausages' intercomparison programme [*Levin et al.*, 2004]. One of the cylinders was replaced in Jan2012 in the Inter-2 loop, owing to an accidental venting of this cylinder.

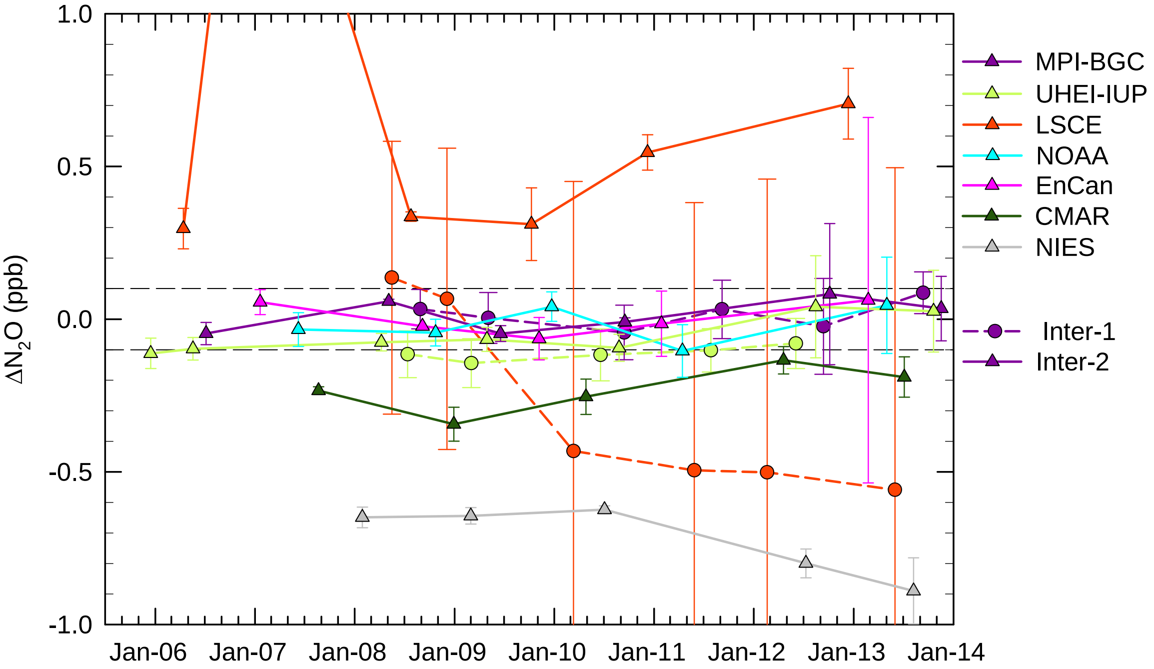


Fig. 11. Central laboratory N_2O differences from MPI-BGC in the Inter-1 and Inter-2 loops. Dashed lines are Inter-1; solid lines are Inter-2. Horizontal black dashed lines indicate the WMO compatibility goal. Some of the laboratories are performing within the WMO compatibility goal, however, some laboratories are not. Results from only two cylinders are shown in the Inter-2 loop, since the third cylinder had an N_2O mole fraction significantly different from ranges experienced in ambient air. As for CH_4, NIES results are measured on an independent gravimetric scale and are therefore offset from the other laboratories. UHEI data shown are not presently on the NOAA scale owing to incorrect assignment of their primary calibration standards on the NOAA website. One of the cylinders was replaced in Jan2012 in the Inter-2 loop, owing to an accidental venting of this cylinder.

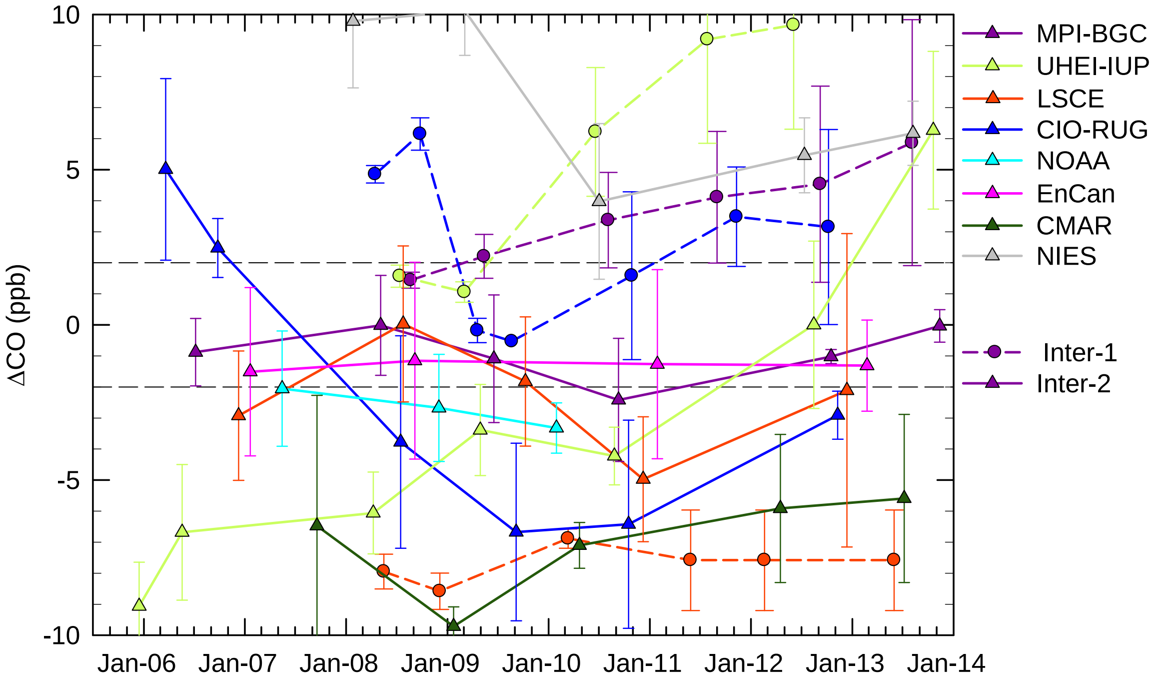


Fig. 12. Central laboratory CO differences from MPI-BGC in the Inter-1 and Inter-2 loops. Dashed lines are Inter-1; solid lines are Inter-2. Horizontal black dashed lines indicate the WMO compatibility goal. Prior to Jan2012, results from only two cylinders are shown in the Inter-2 loop, since the third cylinder had a CO mole fraction significantly different from ranges experienced in ambient air. Following an accidental venting of this cylinder, the replacement was prepared with ambient levels of CO and so could be used in the results shown. As for CH_4 and N_2O, NIES results are measured on an independent gravimetric scale.

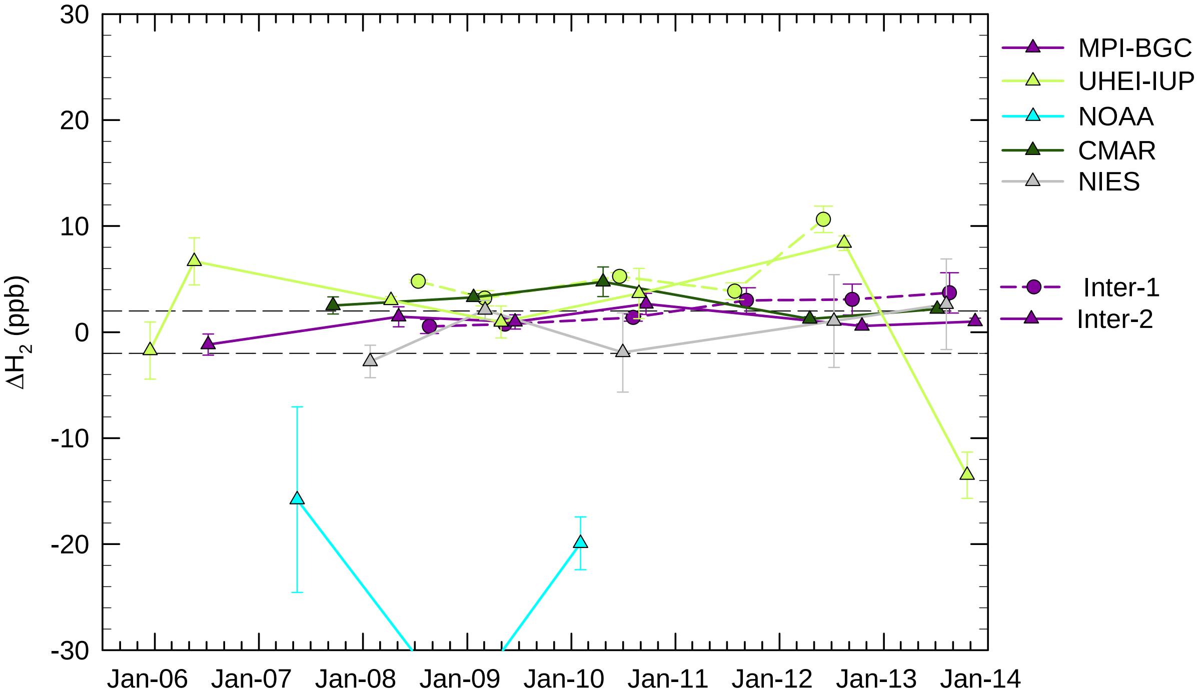


Fig. 13. Central laboratory H_2 differences from MPI-BGC in the Inter-1 and Inter-2 loops. Dashed lines are Inter-1; solid lines are Inter-2. Horizontal black dashed lines indicate the WMO compatibility goal. Prior to Jan2012, results from only two cylinders are shown in the Inter-2 loop, since the third cylinder had an H_2 mole fraction significantly different from ranges experienced in ambient air. Following an accidental venting of this cylinder, the replacement was prepared with ambient levels of H_2 and so could be used in the results shown.

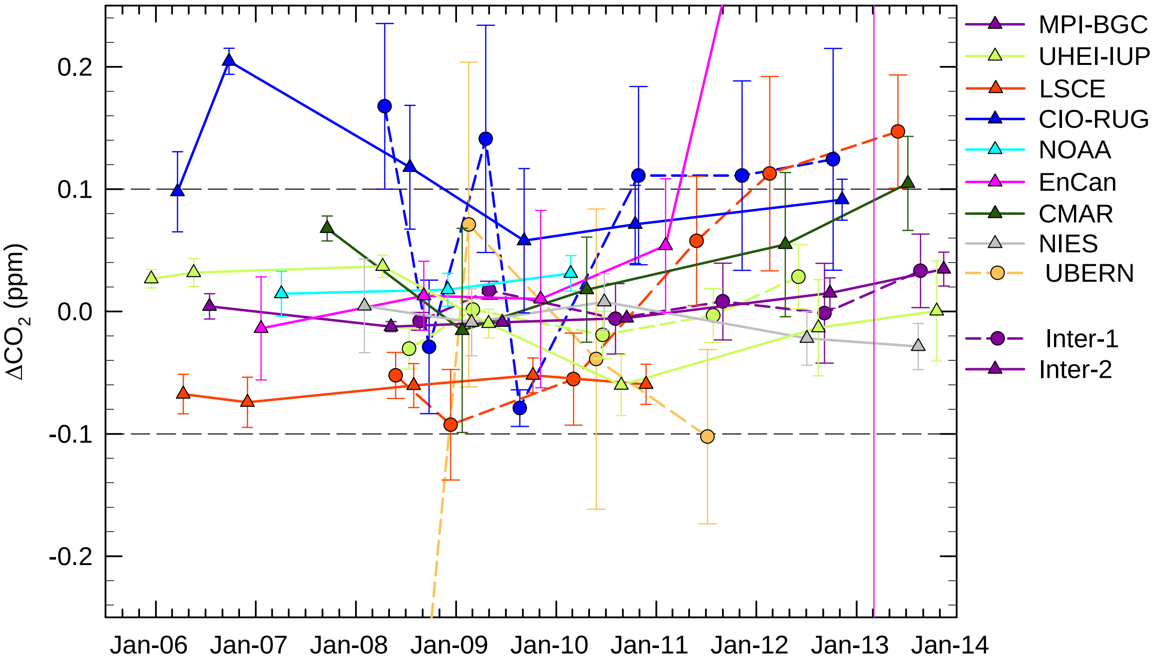


Fig. 14. Central laboratory CO_2 differences from MPI-BGC in the Inter-1 and Inter-2 loops. Dashed lines are Inter-1; solid lines are Inter-2. Horizontal black dashed lines indicate the WMO compatibility goal. Almost all of the analyses are within the WMO compatibility goal for CO_2. As for CH_4, N_2O and CO, NIES results are measured on an independent gravimetric scale, although for CO_2, there appears to be effectively no offset. One of the cylinders was replaced in Jan2012 in the Inter-2 loop, owing to an accidental venting of this cylinder. EnCan found an error in one of their 2013 analyses at the last minute that could not be corrected before publication.

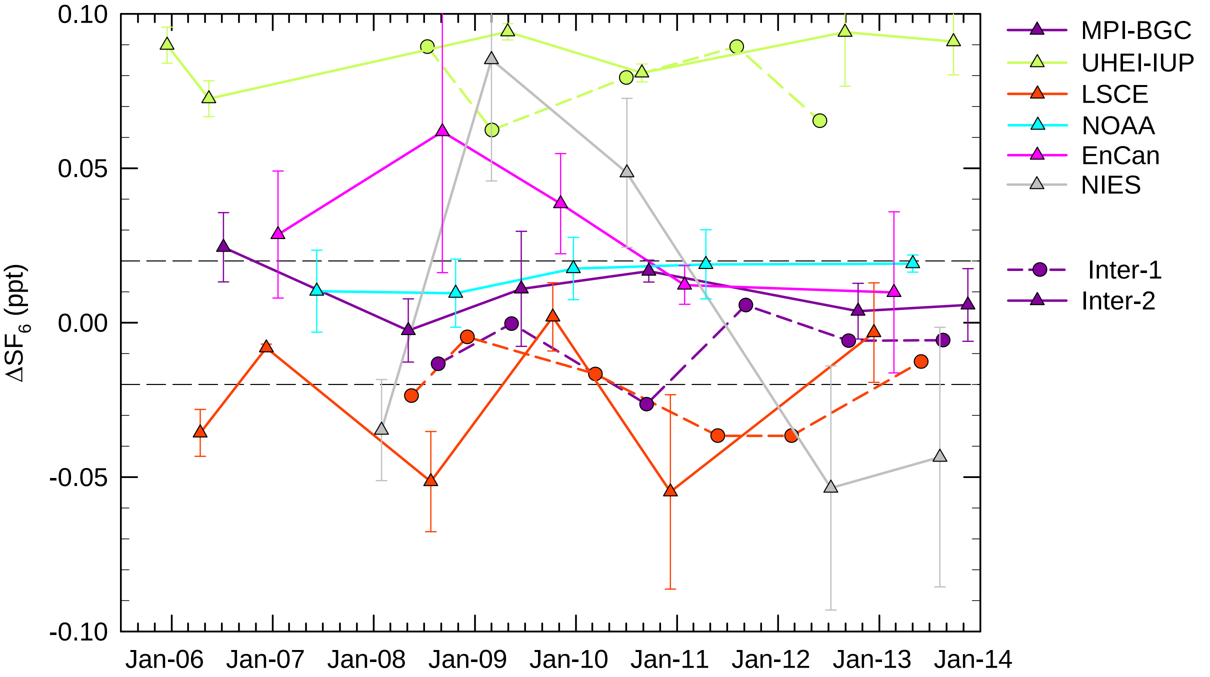


Fig. 15. Central laboratory SF_6 differences from MPI-BGC in the Inter-1 and Inter-2 loops. Dashed lines are Inter-1; solid lines are Inter-2. Horizontal black dashed lines indicate the WMO compatibility goal. Although few measurements are within the ±0.02 ppt WMO compatibility goal, this is with respect to the initial MPI-BGC measurements made in 2005. When one examines the MPI-BGC measurements over time, one can see that NOAA, and much of the time also EnCan and LSCE, are compatible with the nearest in time MPI-BGC measurements. This is one example illustrating that defining compatibility of any two laboratories (or field stations) is not a straightforward task. Note that UHEI data are reported on their own internal gravimetric scale, which has a known offset from the NOAA scale of about 0.1 ppt; if one takes into account this offset, then UHEI measurements would also be compatible with MPI-BGC within the WMO goals for the full time period shown. Results from only one cylinder are shown in the Inter-1 loop, since the other two cylinders had SF_6 mole fractions significantly different from ranges experienced in ambient air (therefore these measurements show no standard deviations). One of the cylinders was replaced in Jan2012 in the Inter-2 loop, owing to an accidental venting of this cylinder.

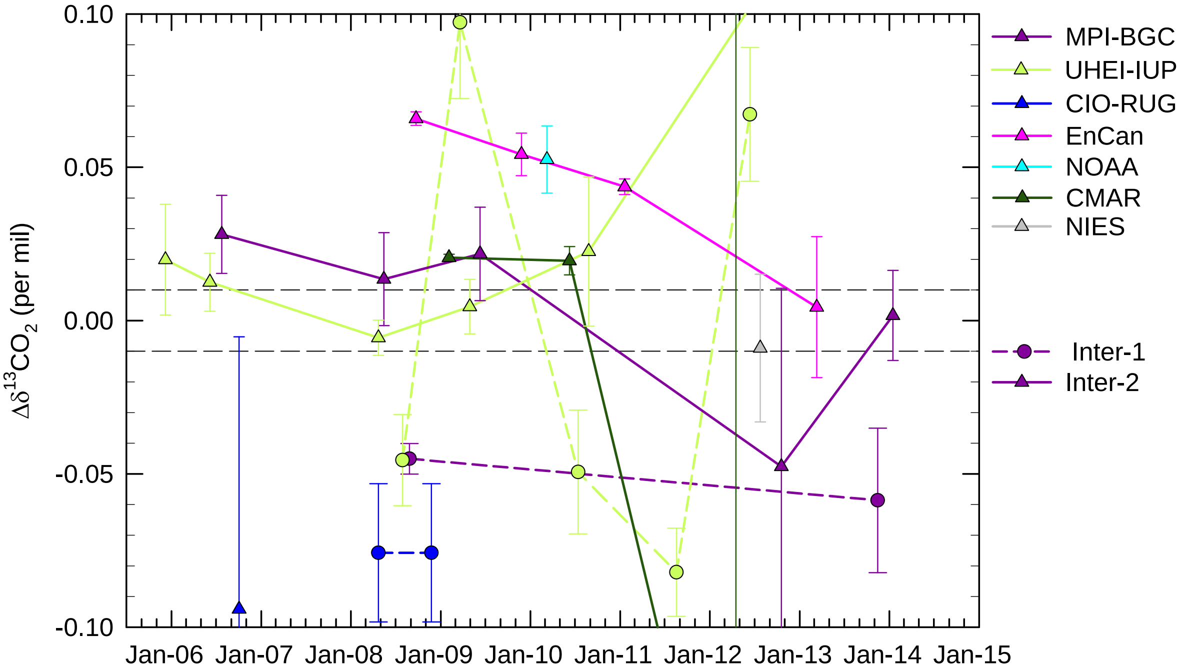


Fig. 16. Central laboratory δ^{13}C-CO$_2$ differences from MPI-BGC in the Inter-1 and Inter-2 loops. Dashed lines are Inter-1; solid lines are Inter-2. Horizontal black dashed lines indicate the WMO compatibility goal. Most of the results are not within the WMO compatibility goal, however, it is difficult to obtain good reproducibility from high pressure cylinders for isotopic measurements. In other words, the Cucumbers programme may not be the appropriate method to assess compatibility in isotopic measurement. One of the cylinders was replaced in Jan2012 in the Inter-2 loop, owing to an accidental venting of this cylinder. CMAR found an error in their 2012 analyses at the last minute that could not be corrected before publication. UHEI revised their data at the last minute giving improved performance; not shown in this figure.

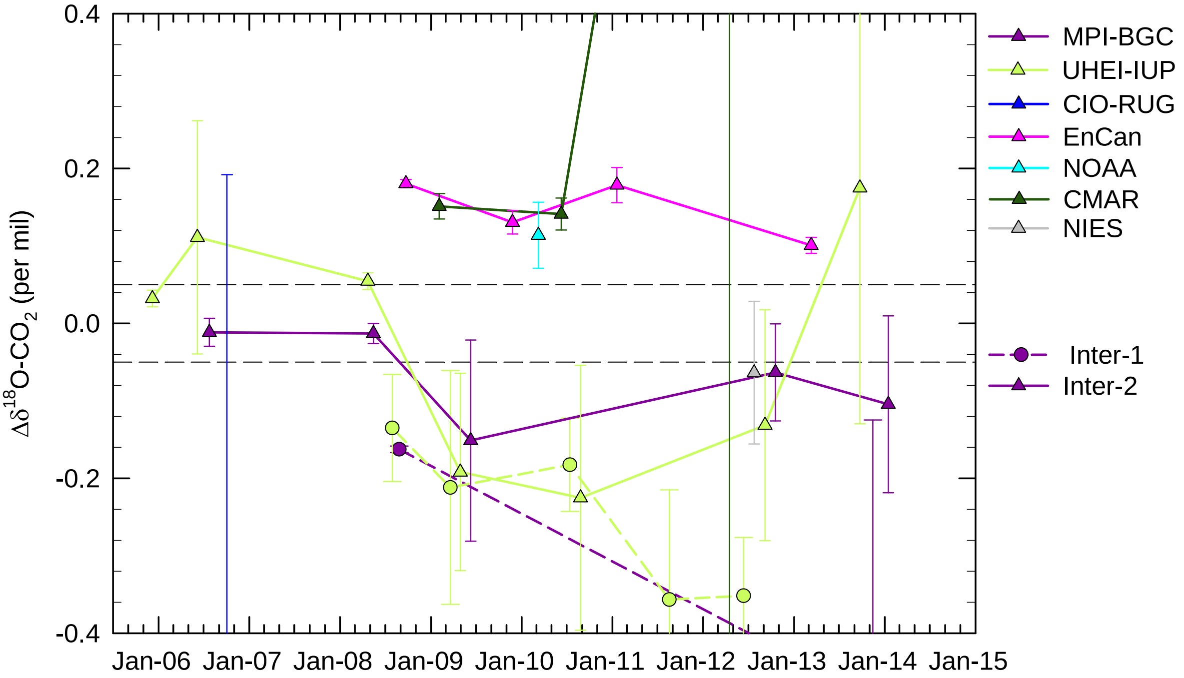


Fig. 17. Central laboratory δ^{18}O-CO$_2$ differences from MPI-BGC in the Inter-1 and Inter-2 loops. Dashed lines are Inter-1; solid lines are Inter-2. Horizontal black dashed lines indicate the WMO compatibility goal. Most of the results are not within the WMO compatibility goal, but as mentioned in the Fig. 16 caption above, we are unsure of the validity and usefulness of the Cucumbers programme for assessing compatibility in isotopic measurements. One of the cylinders was replaced in Jan2012 in the Inter-2 loop, owing to an accidental venting of this cylinder. CMAR found an error in their 2012 analyses at the last minute that could not be corrected before publication. UHEI revised their data at the last minute giving improved performance; not shown in this figure.

Discussion and conclusions

The Cucumbers field station offsets for all species are summarised in Table 1 below. Table 1 presents the average offsets of each station over the entire Cucumbers programme period. Offsets for different field stations are reported against different central laboratories, depending on which loop the field station belongs to, shown in the last row of the table. Note that the values used to determine the 'average offset' are themselves averages, that is, they represent the average offset of the three different Cucumbers measured at a given point in time. When we refer to 'average offset', however, unless stated otherwise, we mean the average in time over the entire duration of the Cucumbers programme. The standard deviation values given in the table are calculated as follows: we calculate two measures of uncertainty; first, we calculate the standard deviation of the average offset reported in the table for a given station or laboratory. In other words, this is the standard deviation of the average of the number of values indicated in the 'no. analyses' row. Second, each time the Cucumbers are analysed at a location, both an offset and a standard deviation are calculated. The offset is the average offset of the three Cucumbers measured and the standard deviation is the standard deviation of this average. We then calculate the average of these individual standard deviations from all analyses of the trio of Cucumbers made at the station for the duration of the programme. Finally, we report in Table 1 the larger of these two measures. In this way, the more conservative measure of uncertainty over the time period of all the analyses is entered into the tables.

As stated at the beginning of this report, one aim of intercomparison programmes is to document compatibility so that laboratories and field stations become more aware of where problems may lie in their analytical procedures, providing information on what could possibly be improved. As such, if the Cucumbers programme is being implemented effectively, we would hope that there would be a reduction in the average offset values in this report compared to those of the previous February 2009 CarboEurope report [*Manning et al.*, 2009]. To examine for such improvements, we have colour coded the data in Table 1: blue indicates an improvement in the average offset or standard deviation compared to the 2009 report, whereas red indicates a deterioration in the values reported. Black values indicate that a station did not make measurements for that gas species at the time the 2009 report was produced.

The results are mixed. SF_6 shows improvement at every station, both the average offsets and the standard deviations (for all ten stations that reported measurements both in 2009 and more recently). However, this comparison is likely not valid, since in Table 1 we have ignored results for a number of cylinders with SF_6 values outside of typical ambient mole fractions, which we did not do in the 2009 report. In a future analysis, we will produce a revised data set over the same time period used in the 2009 report, ignoring the same cylinders, so as to make this comparison more valid.

In the case of CH_4, only three stations out of twelve have improved their average offset. It could be argued that the standard deviation is a more important diagnostic of a station's compatibility than the average offset, because if the standard deviation is relatively low, then one has a higher degree of confidence in making use of the average offset that has been quantified here, and thus one can more robust scientific conclusions using data sets from a diverse combination of stations. Following this argument and considering CH_4, we observe that four stations have improved their standard deviation; interestingly, all four of these stations have reported a greater average offset than was reported in 2009. In the case of N_2O, six of nine stations have improved their average offset, but only four of nine improved their standard deviation. Three stations improved both their average offset and their standard deviation. Three of the six with improved average offset, however, had a cylinder with non-ambient N_2O mole fraction removed from the more recent analyses. For CO_2, 14 of 21 stations improved their average offset, but only four of 21 improved their standard deviation. All four of the latter stations also improved their average offset. One caveat to mention in these comparisons of standard deviation, is that at the time of the 2009 report, many stations had only carried out one analysis of the Cucumbers, therefore the standard deviation reported had to be the standard deviation of the average offset of the three Cucumbers. Whereas as explained above, more recent results report the higher of two different standard deviation measures.

Table 2 shows the average offsets of the central laboratories, including the non-European partners, from the initial MPI-BGC measurements, over the entire Cucumbers programme period. Data are a compilation of both Inter-1 and Inter-2 results, for those laboratories in both loops, and both average offsets and standard deviations are calculated the same as for Table 1. Average offsets and standard deviations are calculated in the same manner as described above for Table 1. Colour coding also indicates the same information. Note that it is possible for a laboratory to have a value in red, indicating a worse value compared to 2009, but to still be well within the WMO compatibility goal. This could be true for laboratories or stations in either Table 1 or 2, but occurs more frequently in Table 2. CH_4 in Table 2 is the most prominent example, where six out of eight laboratories exhibit a worse average offset, but five of the eight are still well within the WMO compatibility goal of ±2 ppb. The sixth laboratory, NIES, does not report data on the WMO NOAA calibration scale, but rather their own internal gravimetric scale. So although the NIES CH_4 offset is large, at about 5.2 ppb, it is a reproducible offset, as has been indicated by a number of measures, including the standard deviation reported here, of only 0.6 ppb.

For N_2O, four of seven laboratories show improved average offset, and five of seven show improved standard deviation, but once again, these results have been biased favourably owing to the exclusion in this report's table of a cylinder with non-ambient N_2O mole fraction (in the Inter-2 loop). Four of the seven laboratories show average offsets within the WMO compatibility goal of ±0.1 ppb (including MPI-BGC itself), which is an encouraging result considering the great analytical challenges in measuring N_2O to high precision. One of these four laboratories, however, has a large standard deviation of the average offset of ±0.20 ppb.

In order to directly compare all field stations with each other, we have constructed Table 3, which shows all the field station average offsets relative to MPI-BGC. For each station, we generate these average offsets by adding the average offset in Table 1 to the the average offset in Table 2 for the relevant central laboratory for that field station. So that all field stations are treated equally, we also combine average offsets for those stations for which MPI-BGC is the central laboratory, using data in the last column of Table 2. The validity of combining offsets in such a manner is assessed by the standard deviation values of the central laboratories shown in Table 2 (shown in blue and yellow highlighting, indicating valid and invalid, respectively). This assessment was made based on whether the standard deviation is above (yellow highlight) or below (blue highlight) the WMO compatibility goals. Thus, all data shown in blue in Table 3 indicate what we believe to be a valid calculation of the station's average offset relative to the initial MPI-BGC measurements. Data shown in red indicate those stations for which we do not have confidence in the calculations, owing to poor results in Table 2. These poor central laboratory results could be either because one or more cylinders exhibits drift for the gas species in question (as we believe is the case for some cylinders for CO mole fraction, for example), or because of analytical challenges faced by the central laboratory.

Some blue data in Table 3 show the importance of having constructed this table, for example, TTA and PAL SF_6 average offsets are much better in Table 3 than Table 1, since the UHEI SF_6 average offset from MPI-BGC is taken into account in Table 3 but not in Table 1. In the case of red data in Table 3, one should not draw conclusions about how 'good' or 'bad' a field station's results are. Instead, the best one can do is to refer to Table 1 for these results; but Table 1 may also have issues, depending on the underlying reasons for the difficulty in achieving good results. For example, since CO mole fractions appear to be unstable in many of the Cucumbers, this implies that the Cucumbers programme is not an appropriate mechanism for determining compatibility of CO calibration scales at field stations or laboratories.

In conclusion, with the design of the Cucumbers programme and with the calculation methodologies presented in this report, we believe that CH_4, CO_2 and SF_6 can be assessed at all field stations as to whether their results are compatible with a single laboratory, namely MPI-BGC. CO and O_2/N_2 can not be assessed in such a manner for any field stations, whereas N_2O and H_2 can be for some, but not all stations. In the latter case for N_2O and H_2, the limitation appears to be due to the analytical challenges faced by specific central laboratories.

Continued on next page…

Table 1. Compilation of all field station average offsets relative to the initial central laboratory analyses.

	Station:	TTA	NOR	PAL	ZEP	HYY	TRN	PUY	KAS	HUN	LUT	OXK	BIK	MHD	PRS	CMN	LMP	SCH	ZUG	LMU	CBW	WAO
CH_4	avg. offset	3.08		0.61			-4.90	0.11	6.54	1.22	-0.04	0.96	0.56	-0.80	-1.26	0.81	1.60	-0.39	-1.58		2.03	
	std. dev.	1.80		0.51			3.91	1.15	19.51	1.05	0.72	3.94	1.62	0.94	3.40	4.02	1.72	0.99	0.86		0.76	
	no. analyses	3		5			4	3	8	8	5	3	4	5	5	4	2	5	2		2	
N_2O	avg. offset	0.30		-0.04			-1.32	0.16	-0.19	-0.59	-0.25	0.00	0.31	-0.30		0.68	-0.39	-0.61	-0.42		-0.26	
	std. dev.	0.88		0.19			0.54	0.55	1.31	0.79	0.47	0.22	0.50	0.10		0.72	1.30	0.33	0.45		0.43	
	no. analyses	3		6			2	3	2	7	5	3	4	5		4	2	5	2		2	
CO	avg. offset	1.9		3.8			-0.1				0.2	0.2	-2.0	6.8		-1.4		0.7	1.4		-6.7	-1.5
	std. dev.	2.1		2.3			1.2				1.9	0.9	2.3	3.6		5.8		3.3	1.0		1.4	0.4
	no. analyses	3		5			1				5	3	4	5		4		5	2		2	1
H_2	avg. offset	-20.6		-2.7										-9.9		-26.7		1.1			-7.1	-19.2
	std. dev.	4.0		1.7										2.9		0.6		3.6			7.3	0.9
	no. analyses	3		3										4		1		2			2	1
CO_2	avg. offset	-0.12	0.08	0.01	-0.30	-0.03	-0.05	-0.02	0.04	-0.01	0.01	-0.20	-0.12	-0.07	0.03	0.28	-0.12	-0.07	-0.21	-0.54	0.08	-0.07
	std. dev.	0.34	0.23	0.03	0.36	0.08	0.06	0.22	0.19	0.09	0.12	0.61	0.15	0.09	0.05	0.27	0.16	0.10	0.22	0.33	0.15	0.11
	no. analyses	3	1	6	2	5	5	5	7	6	5	3	4	2	5	4	3	5	2	4	2	6
SF_6	avg. offset	-0.12		-0.13			0.03	0.01	0.25	0.00	-0.03	0.05	0.04	-0.02		0.01		0.09	0.05		-0.03	
	std. dev.	0.09		0.03			0.05	0.02	0.44	0.07	0.10	0.01	0.06	0.03		0.06		0.22	0.03		0.11	
	no. analyses	3		1			3	3	2	8	5	3	4	5		4		5	2		2	
O_2/N_2	avg. offset											25.1	-4.6									-2.6
	std. dev.											38.0	13.2									11.2
	no. analyses											4	3									6
Central lab:		UHEI	UHEI	UHEI	UHEI	UHEI	LSCE	LSCE	LSCE	LSCE	MPI	MPI	MPI	MPI	MPI	MPI	MPI	MPI	MPI	MPI	MPI	MPI

- Data are a compilation of all analyses from a given field station, with the exception that the following results are ignored because the relevant mole fractions are outside of typical calibration scale ranges: Euro-1, cylinders D88486 for N_2O and D88473 for SF_6; Euro-3, cylinders D88472 and D88476 for SF_6; Euro-5, cylinders D88480 for N_2O and SF_6 and D88484 for SF_6. In Euro-4, analyses of cylinder D88483 were also ignored for CO_2 and H_2 as this cylinder proved to be very unstable for these two species, for unknown reasons. The cylinder was replaced in Jan2009.

- 'Avg. offset' is the average offset calculated over the entire period of the Cucumbers programme of the difference of a field station's measurements of Cucumbers from the initial central laboratory measurements. 'Std. dev.' is defined in detail in the text; briefly, it is the larger value of either the standard deviation of the average offset reported, or the average of the standard deviations determined each time a trio of Cucumbers is analysed at a station. 'No. analyses' reports the number of times the trio of Cucumbers has visited the field station and being analysed; that is, the number of analyses used to calculate the average offset. The last row of the table indicates who is the central laboratory for each field station.
- Units are: CO_2 in ppm; CH_4, CO, N_2O, H_2 in ppb; SF_6 in ppt; O_2/N_2 in per meg.
- Colour coding: blue indicates an improvement in that value compared to the equivalent value reported in the 2009 CarboEurope report [*Manning et al.*, 2009]; red indicates a deterioration in the value compared to the earlier report; black indicates that the station did not measure the given species at the time the 2009 report was produced.
- TTA analysers had known problems with their initial measurements in Apr2008, which were later resolved. Some ZEP analyses were in fact made at the Stockholm University laboratory. KAS found errors in their 2012 and 2013 analyses at the last minute that could not be corrected before publication, significantly affecting their results in this table. UHEI corrected their Euro-1 SF_6 data at the last minute that could not be included here, affecting offsets given for TTA and PAL stations.

Continued on next page…

Table 2. Compilation of all central laboratory offsets relative to the initial MPI-BGC analyses.

	Station:	UHEI	LSCE	CIO	UBERN	EnCan	NOAA	CMAR	NIES	MPI-BGC
CH_4	avg. offset	-0.59	-0.42	0.39		0.86	1.31	1.23	5.16	0.05
	std. dev.	0.53	0.99	0.82		0.69	1.23	0.84	0.57	0.65
	no. analyses	12	11	12		5	3	5	5	12
N_2O	avg. offset	-0.08	0.20			0.00	-0.02	-0.23	-0.72	0.01
	std. dev.	0.07	0.73			0.20	0.08	0.08	0.12	0.08
	no. analyses	12	12			5	5	5	5	12
CO	avg. offset	0.4	-5.3	0.5		-1.3	-2.7	-7.0	7.1	1.3
	std. dev.	6.4	3.0	4.3		2.6	1.5	2.1	2.7	2.7
	no. analyses	12	11	13		4	3	5	5	12
H_2	avg. offset	2.9					-23.8	2.8	0.2	1.5
	std. dev.	6.1					10.5	1.3	3.0	1.4
	no. analyses	12					3	5	5	12
CO_2	avg. offset	0.00	-0.02	0.09	-0.14	0.17	0.02	0.05	-0.01	0.01
	std. dev.	0.03	0.08	0.08	0.25	0.34	0.02	0.05	0.03	0.02
	no. analyses	12	11	12	4	5	3	5	5	12
SF_6	avg. offset	0.08	-0.02			0.03	0.02		0.00	0.00
	std. dev.	0.01	0.02			0.02	0.01		0.06	0.01
	no. analyses	11	12			5	5		5	12
O_2/N_2	avg. offset			-2.5						-7.4
	std. dev.			8.3						11.5
	no. analyses			2						6
^{13}C-CO_2	avg. offset	0.02		-0.08		0.04	0.05	-0.06	-0.01	-0.01
	std. dev.	0.07		0.04		0.03	0.01	0.13	0.02	0.04
	no. analyses	12		3		4	1	3	1	7
^{18}O-CO_2	avg. offset	-0.12		-0.80		0.15	0.11	0.57	-0.06	-0.16
	std. dev.	0.17		0.61		0.04	0.04	0.74	0.09	0.19
	no. analyses	12		2		4	1	3	1	5

- Data are a compilation of all results from Inter-1 and Inter-2 analyses, with the exception that the following results are ignored because the relevant mole fractions are outside of typical calibration scale ranges: Inter-1, cylinders D88471 and D88474 for SF_6; Inter-2, cylinder CA05300 for N_2O and cylinder CA05285 for CO and H_2.
- ‘Avg. offset’ is calculated as for Table 1, where offsets are relative to the initial MPI-BGC measurements, made in Sep2005 for Inter-2 and Jan2008 for Inter-1. ‘Std. dev.’ and ‘no. analyses’ are also calculated as for Table 1. The final column shows MPI-BGC’s own reanalyses, also as offsets against the same initial measurements made by themselves.
- Units are: CO_2 in ppm; CH_4, CO, N_2O, H_2 in ppb; SF_6 in ppt; O_2/N_2 in per meg; $\delta^{13}C$-CO_2 and $\delta^{18}O$-CO_2 in per mil.
- Blue, red and black colour coding is as for Table 1. Blue and yellow highlighting, applied only to standard deviation values for UHEI, LSCE and MPI-BGC, indicates

where we believe results are good enough to calculate a field station’s offsets relative to MPI-BGC, as given in Table 3. See text for more details.
- NIES data are on independent gravimetric scales for CO_2, CH_4, CO and N_2O.
- UHEI data are on an independent gravimetric scale for SF_6.
- EnCan (CO_2, 2013 analyses) and CMAR (isotopes, 2012 analyses) found errors at the last minute that could not be corrected before publication, which significantly affected the results shown in this table.

Continued on next page…

Table 3. Compilation of all field station average offsets relative to MPI-BGC.

	Station:	TTA	NOR	PAL	ZEP	HYY	TRN	PUY	KAS	HUN	LUT	OXK	BIK	MHD	PRS	CMN	LMP	SCH	ZUG	LMU	CBW	WAO
CH_4	avg. offset	2.49		0.02			-5.32	-0.31	6.12	0.80	0.01	1.01	0.61	-0.75	-1.20	0.86	1.65	-0.34	-1.53		2.08	
	std. dev.	1.87		0.73			4.03	1.52	19.54	1.44	0.96	3.99	1.75	1.14	3.46	4.07	1.84	1.18	1.08		0.99	
	no. analyses	3		5			4	3	8	8	5	3	4	5	5	4	2	5	2		2	
N_2O	avg. offset	0.2		-0.1			-1.1	0.4	0.0	-0.4	-0.2	0.0	0.3	-0.3		0.7	-0.4	-0.6	-0.4		-0.2	
	std. dev.	0.9		0.2			0.9	0.9	1.5	1.1	0.5	0.2	0.5	0.1		0.7	1.3	0.3	0.5		0.4	
	no. analyses	3		6			2	3	2	7	5	3	4	5		4	2	5	2		2	
CO	avg. offset	2.3		4.2			-5.4				1.5	1.6	-0.6	8.2		-0.1		2.1	2.8		-5.3	-0.1
	std. dev.	6.7		6.8			3.3				3.3	2.8	3.5	4.4		6.4		4.3	2.8		3.0	2.7
	no. analyses	3		5			1				5	3	4	5		4		5	2		2	1
H_2	avg. offset	-17.7		0.2										-8.4		-25.2		2.6			-5.6	-17.7
	std. dev.	7.3		6.3										3.2		1.5		3.9			7.4	1.6
	no. analyses	3		3										4		1		2			2	1
CO_2	avg. offset	-0.12	0.08	0.01	-0.30	-0.03	-0.06	-0.04	0.02	-0.03	0.01	-0.19	-0.11	-0.07	0.03	0.28	-0.12	-0.07	-0.20	-0.53	0.08	-0.07
	std. dev.	0.35	0.23	0.04	0.36	0.09	0.10	0.24	0.20	0.13	0.12	0.61	0.15	0.10	0.06	0.27	0.16	0.11	0.22	0.33	0.15	0.11
	no. analyses	3	1	6	2	5	5	5	7	6	5	3	4	2	5	4	3	5	2	4	2	6
SF_6	avg. offset	-0.03		-0.05			0.00	-0.01	0.23	-0.03	-0.03	0.05	0.04	-0.02		0.01		0.09	0.05		-0.03	
	std. dev.	0.09		0.04			0.05	0.02	0.44	0.08	0.10	0.02	0.06	0.03		0.06		0.22	0.03		0.11	
	no. analyses	3		1			3	3	2	8	5	3	4	5		4		5	2		2	
O_2/N_2	avg. offset											9.4	-12.0									-10.0
	std. dev.											32.7	17.5									16.1
	no. analyses											4	3									6

- As for Table 1, data are a compilation of all analyses from a given field station, with the same results for some cylinders ignored as given in Table 1.
- ‘Avg. offset’ is the sum of the average offset given in Table 1 and the relevant average offset in Table 2. For example, the PAL CH_4 average offset is the value from Table 1 (0.61 ppb) added to the UHEI average offset from Table 2 (-0.59 ppb), since UHEI is the central laboratory for PAL measurements, giving a result as shown in Table 3 of 0.02 ppb. ‘Std. dev.’ is defined as the quadrature sum of errors of the relevant standard deviations reported in Tables 1 and 2. Using PAL CH_4 as example again, the standard

deviation given is the square root of $0.51^2 + 0.53^2 = 0.73$ ppb. Where 0.51 ppb and 0.53 ppb are the reported standard deviations for PAL and UHEI, respectively, given in Tables 1 and 2. 'No. analyses' is identical to Table 1.

- Units are: CO_2 in ppm; CH_4, CO, N_2O, H_2 in ppb; SF_6 in ppt; O_2/N_2 in per meg.
- Colour coding: blue indicates data for which we believe a valid average offset to the initial MPI-BGC measurements can be made, based on Table 2 central laboratory standard deviation information (see text for more details), whereas red indicates data for which we believe the average offset shown does not represent a valid calculation. In the case of red data, therefore, one should not draw conclusions about how 'good' or 'bad' the field station's results are. The best one can do in such cases, is to refer to Table 1 data rather than Table 3.
- TTA analysers had known problems with their initial measurements in Apr2008, which were later resolved. Some ZEP analyses were in fact made at the Stockholm University laboratory. KAS found errors in their 2012 and 2013 analyses at the last minute that could not be corrected before publication, significantly affecting their results in this table.

Continued on next page…

References

Levin, I., C. Facklam, M. Schmidt, M. Ramonet, P. Ciais, I. Xueref, R. Langenfelds, C. Allison, R. Francey, A. Jordan, M. Rothe, W. A. Brand, R. E. M. Neubert, H. A. J. Meijer, T. Machida, and H. Mukai, Results of intercomparison programme for analysis of "sausage" flask air samples, in *EU TACOS-Infrastructure project*, 10 pp., 2004.

Manning, A. C., A. Jordan, I. Levin, M. Schmidt, R. E. M. Neubert, A. Etchells, B. Steinberg, P. Ciais, T. Aalto, F. Apadula, W. A. Brand, M. Delmotte, A. Giorgio di Sarra, B. Hall, L. Haszpra, L. Huang, D. Kitzis, S. van der Laan, R. L. Langenfelds, M. Leuenberger, A. Lindroth, T. Machida, F. Meinhardt, J. Moncrieff, J. A. Morguí, J. Necki, M. Patecki, E. Popa, L. Ries, K. Rozanski, R. Santaguida, L. P. Steele, J. Strom, Y. Tohjima, R. L. Thompson, A. T. Vermeulen, F. Vogel, and D. E. Worthy, Final report on CarboEurope 'Cucumbers' intercomparison programme, 2009.

Peters, W., M. C. Krol, G. R. van der Werf, S. Houweling, C. D. Jones, J. Hughes, K. Schaefer, K. Masarie, A. Jacobson, J. B. Miller, C. H. Cho, M. Ramonet, M. Schmidt, L. Ciattaglia, F. Apadula, D. Heltai, F. Meinhardt, A. G. di Sarra, S. Piacentino, D. Sferlazzo, T. Aalto, J. Hatakka, J. Ström, L. Haszpra, H. A. J. Meijer, S. van der Laan, R. E. M. Neubert, A. Jordan, X. Rodo, J.-A. Morguí, A. T. Vermeulen, E. Popa, K. Rozanski, M. Zimnoch, A. C. Manning, M. Leuenberger, C. Uglietti, A. J. Dolman, P. Ciais, M. Heimann, and P. P. Tans, Seven years of recent European net terrestrial carbon dioxide exchange constrained by atmospheric observations, *Global Change Biology*, *16*, 1317-1337, doi:10.1111/j.1365-2486.2009.02078.x, 2010.

Appendix 1. List of responsible scientists for each loop and participants

Responsible scientists

Euro-1 = Ingeborg Levin (Ingeborg.Levin@iup.uni-heidelberg.de)
Euro-2 = Martina Schmidt (Martina.Schmidt@cea.fr) / Marc Delmotte (marc.delmotte@lsce.ipsl.fr)
Euro-3 = Rolf Neubert (R.E.M.Neubert@rug.nl) / Armin Jordan (ajordan@bgc-jena.mpg.de)
Euro-4 = Armin Jordan (ajordan@bgc-jena.mpg.de)
Euro-5 = Armin Jordan (ajordan@bgc-jena.mpg.de)
Inter-1 = Andrew Manning (Andrew.UEA@gmail.com)
Inter-2 = Andrew Manning (Andrew.UEA@gmail.com)

Euro-1

UHEI	University of Heidelberg, GERMANY, Ingeborg Levin
TTA	Tall Tower, Angus, UNITED KINGDOM, John Moncrieff
NOR	Norunda, SWEDEN, Anders Lindroth
HYY	Hyytiälä, FINLAND, Tuula Aalto
PAL	Pallas, FINLAND, Tuula Aalto
ZEP	Zeppelin, NORWAY, Johan Strom

Euro-2

LSCE	Lab. des Sciences du Climat et de l'Environ., FRANCE, Marc Delmotte (Martina Schmidt prior to 2014)
TRN	Trainou Tower, FRANCE, Martina Schmidt
PUY	Puy de Dome, FRANCE, Martina Schmidt
KAS	Kasprowy Wierch, POLAND, Kazimierz Rozanski
HUN	Hegyhatsal, HUNGARY, Laszlo Haszpra

Euro-3

CIO-RUG	University of Groningen, THE NETHERLANDS, Rolf Neubert
LUT	Lutjewad, THE NETHERLANDS, Rolf Neubert
OXK	Ochsenkopf, GERMANY, Uwe Schultz
BIK	Bialystok, POLAND, Falk Haensel
JFJ	Jungfraujoch, SWITZERLAND, Markus Leuenberger
MHD	Mace Head, IRELAND, Martina Schmidt

Euro-4

MPI-BGC	Max Planck Institute for Biogeochemistry, GERMANY, Armin Jordan
PRS	Plateau Rosa, ITALY, Francesco Apadula
CMN	Monte Cimone, ITALY, Riccardo Santaguida
LMP	Lampedusa, ITALY, Alcide Giorgio di Sarra

Euro-5

MPI-BGC	Max Planck Institute for Biogeochemistry, GERMANY, Armin Jordan
SCH	Schauinsland, GERMANY, Frank Meinhardt
ZUG	Zugspitze, GERMANY, Ludwig Ries
LMU	La Muela, SPAIN, Josep Anton Morguí
CBW	Cabauw, THE NETHERLANDS, Alex Vermeulen

Inter-1

UHEI	University of Heidelberg, GERMANY, Ingeborg Levin
MPI-BGC	Max Planck Institute for Biogeochemistry, GERMANY, Armin Jordan
CIO-RUG	University of Groningen, THE NETHERLANDS, Rolf Neubert
UEA-ENV	University of East Anglia, UNITED KINGDOM, Andrew Manning
WAO	Weybourne, UNITED KINGDOM, Michael Patecki
LSCE	Lab. des Sciences du Climat et de l'Environ., FRANCE, Marc Delmotte (Martina Schmidt prior to 2014)
UBERN	University of Bern, SWITZERLAND, Markus Leuenberger

Inter-2

UHEI	University of Heidelberg, GERMANY, Ingeborg Levin
MPI-BGC	Max Planck Institute for Biogeochemistry, GERMANY, Armin Jordan
CIO-RUG	University of Groningen, THE NETHERLANDS, Rolf Neubert
LSCE	Lab. des Sciences du Climat et de l'Environ., FRANCE, Marc Delmotte (Martina Schmidt prior to 2014)
EnCan	Environment Canada, CANADA, Doug Worthy
NOAA	National Oceanic and Atmospheric Admin., USA, Duane Kitzis / Brad Hall
CMAR	CSIRO Marine Atmospheric Research, AUSTRALIA, Paul Steele
NIES	National Institute for Environ. Studies, JAPAN, Toshinobu Machida

Appendix 2. Measurement and data reporting protocols

A) Cucumber cylinder analysis protocols

1. Immediately email your "Responsible Scientist" to notify receipt of shipment (Responsible Scientists are listed at the end of this document). You now have **4 weeks** to complete all analyses and to ship the cylinders to the next participant.
2. Position the cylinders in your laboratory, ensuring that there are no significant heat sources near the cylinders (for example, cryo-chillers, air conditioners, heaters).
3. Remove cap nuts from the cylinders (with 27 mm wrench). Note: these caps have been installed both to prevent and test for leaks. A leak could be present if you have severe difficulty removing the cap. In such a case, please inform your Responsible Scientist immediately.
4. Check the white plastic gasket (PCTFE) on the regulator DIN14 connectors for scratches or damage. If it is necessary, replace the gasket using a new gasket which you can find underneath the foam cube (below the regulator) in the high mole fraction Cucumber box. Please also inform your Responsible Scientist, if you replace any of the gaskets. If there are no gaskets left, contact Armin Jordan.
5. Install the 3 regulators supplied, and open the cylinders to the high pressure side of the regulators.
6. Record the pressure in each cylinder, noting it in the Cucumber Logsheet Excel file.
7. Make a leak check of the DIN14 regulator connections to the cylinders. This should be done with the regulators at high pressure, with the main cylinder valves closed, and the regulator outlets closed (turn black knobs fully anti-clockwise, with nothing connected to the female Quick Connect on the outlets). Leave this for **at least half a day**, then check for any pressure drop on the high pressure side of the regulators. If

the pressure has dropped, then all pressure should be released from the regulator (use the male Quick Connect which was supplied to your lab), and the connection carefully re-tightened (overwinding of the connection will destroy the PCTFE gasket). Then repeat the at-least-half-a-day leak check.

8. If you continue to have difficulty making a leak-tight seal, please do not over tighten. The PCTFE gasket should be examined, and replaced if necessary, for example, if it has become too "flat" (see 4. above). If problems continue, please contact your Responsible Scientist.
9. Now do a pressure flush of the regulators 4 times. This means filling the regulators to the cylinder pressure, closing the cylinders, emptying all air from the regulators (use the male Quick Connect which was supplied to your lab), and then repeating, for a total of 4 times for each cylinder.
10. After installing regulators, leak checking, and pressure flushing, **wait a minimum of 2 days with the cylinders in the same lab where the analyses will take place, and with the high pressure side of the regulator at cylinder pressure, but with the main cylinder valve closed**, before the first analysis from the cylinders. After 2 days, check again for leaks by looking for any pressure drop on the high pressure side of the regulators. (If your measurements are at a remote field site and do not allow a 2-day wait, please discuss with your Responsible Scientist).
11. Each laboratory should then follow their normal procedures for analysis from unknown high pressure cylinders. Our recommendations are:
    a) Flush between 5 and 10 Litres of air out of each cylinder and regulator (at low flow rate, that is, less than 500 mL/min) before you start your proper analyses.
    b) Perform as high level on-site calibration as possible immediately before or after you analyse the Cucumbers.
12. Each laboratory may conduct more than one analysis from the cylinders, but we ask that no laboratory remove more than a total of 40 Litres of air from each cylinder.
13. If a laboratory conducts more than one analysis, we encourage, if possible, to do them on different days, but ensuring to keep the 4 week deadline.
14. After final analysis, remove the regulators, first noting again the cylinder pressures and writing them on the Logsheet, reinstall the cylinder DIN14 cap nuts, and place cylinders back in their boxes.
15. Ship the cylinders to the next participant in the round robin (if you are not sure who this is, ask your Responsible Scientist – usually it will be as indicated in the schematic below).
16. Email your Responsible Scientist when the cylinders are shipped out to the next participant.

** Please note, for O_2/N_2 analysis, there are a few additional requirements. Contact Andrew Manning if you are not sure about these.
** Inter-2 loop is still with the old CGA-type cylinders, and so protocols should be slightly different where DIN14 specifications are given above.

B) Cucumber cylinder data reporting protocols

1. After analysing a set of Cucumber cylinders, you should report your results to both your Responsible Scientist and to Andrew Manning. There is a **4 week reporting deadline** for these results. We understand that some labs require 6 months or more before they can report final mole fraction numbers, however, at least provisional

numbers are required 4 weeks after analysis.

2. Please use the Excel template provided for reporting your results.
3. Any change or update in mole fraction data for these analyses should be reported in a new copy of the same Excel template, and sent to your Responsible Scientist and Andrew Manning.
4. CO_2 results are required to be reported in the programme. Reporting (and analysis) of other species is optional.
5. Your analysis results, combined with results from other participants, may be presented at international meetings and workshops. If you have any concerns about this, you should contact Andrew Manning.

Appendix 3. WMO inter-laboratory compatibility goals

The WMO inter-laboratory compatibility goals are:

CH_4	± 2 ppb
N_2O	± 0.1 ppb
CO	± 2 ppb
H_2	± 2 ppb
CO_2	± 0.10 ppm
SF_6	± 0.02 ppt
O_2/N_2	± 1 per meg
$\delta^{13}C\text{-}CO_2$	± 0.01 ‰
$\delta^{18}O\text{-}CO_2$	± 0.05 ‰