RINGO | Readiness of ICOS

Readiness of ICOS for Necessities of integrated Global Observations

D1.3

An ICOS flask sampling protocol based on historical time series and high-resolution footprint modelling

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 730944

RINGO (GA no 730944)

Public Report

RINGO | Readiness of ICOS

Deliverable: D1.3

Author(s): I. Levin, U. Karstens, F. Maier, M. Eritt, D. Rzesan, S. Conil, D. Kubistin, M. Lindauer, S. Hammer

Date: 2. 4. 2019
Activity: Task 1.3
Lead Partner: UHEI
Document Issue: Report
Dissemination Level: Public
Contact: Ingeborg.Levin@iup.uni-heidelberg.de

	Name	Partner	Date
From	Ingeborg Levin	UHEI	2. 4. 2019
Reviewed by			
Approved by			

Version	Date	Comments/Changes	Author/Partner
V2	14.5.2019	1 addition (flow rate, P. 7)	M. Leuenberger (CH)

Deliverable Review Checklist

A list of checkpoints has been created to be ticked off by the Task Leader before finalizing the deliverable. These checkpoints are incorporated into the deliverable template where the Task Leader must tick off the list. √

- Appearance is generally appealing and according to the RINGO template. Cover page has been updated according to the Deliverable details. ☐
- The executive summary is provided giving a short and to the point description of the deliverable. ☐
- All abbreviations are explained in a separate list. ☐
- All references are listed in a concise list. ☐
- The deliverable clearly identifies all contributions from partners and justifies the resources used. ☐
- A full spell check has been executed and is completed. ☐

ABSTRACT

Footprint model runs with the regional transport model STILT conducted by the ICOS Carbon Portal for 4 labelled ICOS class 1 stations have been analysed together with the first in situ CO_2 and CO measurements at the stations to develop an optimal strategy for flask sampling with an automated sampler. Flask sampling in the ICOS atmospheric station network has three different purposes: 1) Provide an independent quality control for in situ observations, 2) provide representative information on components currently not monitored in situ at the stations, 3) collect samples that are significantly influenced by fossil fuel CO_2 ($ffCO_2$) emission areas for $^{14}CO_2$ analysis. Based on the existing data and experimental results obtained at the Heidelberg pilot station with a prototype flask sampler, we suggest that single flask samples should be collected regularly every 3rd day around noon/afternoon from the highest level of a tower station and analysed for all accessible components in the ICOS Flask and Calibration Laboratory (FCL). Flasks shall be collected integrated over one hour with temporally decreasing flow rate to obtain a true hourly mean. At all stations studied, more than 50% of these mid-day flasks will likely be sampled during low ambient variability and, based on experimental test, suitable to detect biases larger than 0.1 ppm CO_2 between flask and in situ observations within one month of weekly observations. In order to have a maximum chance to also sample $ffCO_2$ emission areas, additional flasks need to be collected on all other days in the afternoon. Using the continuous in situ CO observations, the CO offset ΔCO compared to an estimated background value must be determined the day after each flask sampling and, depending on this offset a decision must be made if a flask shall be retained for $^{14}CO_2$ analysis. It turned out that, based on existing data of the last three years with a preliminary threshold of $\Delta CO > 40$ ppb that would translate into a fossil fuel CO_2 concentration of about 4-5 ppm, summer events will be very rare. During the other seasons events could be collected more frequently. A lower threshold for ΔCO would increase the uncertainty of the $ffCO_2$ determination to > 30%. The strategy developed in this task now needs to be approved by the ICOS Atmospheric Monitoring Station Assembly (MSA) and the procedures be implemented at the stations with support from the Atmospheric Thematic Center (ATC).

TABLE OF CONTENTS

1 INTRODUCTION: Aims of the flask sample analyses as ICOS stations

Flask sampling shall be conducted at all ICOS class-1 stations with subsequent analysis for greenhouse and other trace gases as well as for isotopic analysis of CO_2 in the Central Analytical Laboratories (CAL). There are three main aims for regular flask sampling:

1. Use flask value for comparison with trace gas components measured in situ at the station (CO_2, CH_4, CO, (N_2O)). This comparison shall provide an ongoing quality control of the in situ measurements.
2. Obtain data of components not measured continuously at the station, such as SF_6 or H_2, but also stable isotopes of CO_2 or O_2/N_2: Here we aim at monitoring their large-scale representative concentration levels to allow estimating their continental fluxes with help of inverse modelling.
3. Analysis of ^{14}C in CO_2 to allow determining the atmospheric fossil fuel CO_2 component ($ffCO_2$) and with help of these observations and inverse modelling to estimate the continental fossil fuel CO_2 source strength.

To meet aims 1 and 2 requires flask sampling at well-mixed meteorological conditions and sampled footprints not dominated by particular hot spot source areas. For aim 2, in addition, we aim at covering the entire footprint of the station. In contrast, meeting aim 3, due to the expected generally small fossil fuel signals at ICOS stations (which are, if possible, located far away from anthropogenic source areas), requires some “hot spot targeted” sampling to maximize the fossil fuel CO_2 signal in the samples. Note that the detection limit (or measurement uncertainty) of the fossil fuel CO_2 ($ffCO_2$) component is of order 1 - 1.5 ppm.

In order that flask sample results are useful for flux estimates with current inversion models, at least at the standard ICOS stations, which are tall towers, flasks should be collected during mid-day or early afternoon. During this time of the day atmospheric mixing is strong and model transport errors are small. A different sampling strategy needs to be developed for mountain stations or coastal sites, to take into account particular circulation patterns such as land-sea brises or upslope/downslope winds. For all sampling wind speeds should be larger than about 2 m/s, so that the sampled footprint is well defined. The strategy outlined below has been developed for tall tower sites that are located some ten kilometers away from the coast.

The sampling strategies have been investigated based on footprint model simulations with the regional transport model STILT [Lin et a., 2003] implemented at the ICOS Carbon Portal (https://www.icos-cp.eu/about-stilt) and real continuous CO_2 and CO observations available in the first release of ICOS atmospheric data [ICOS RI, 2018] available at ICOS Carbon Portal (https://meta.icos-cp.eu/collections/4H3RS8YtXIt_WTcjrSSkaQ-O). The STILT model is coupled to the EDGAR v4.3 emission inventory [Janssens-Meanhout et al., 2017] and the biosphere model VPRM [Mahadevan et al., 2008] to simulate atmospheric CO_2 and CO concentrations.

2 Requirements and technical constraints for flask sampling and analysis

1. Flask duplicates: Flask sampling is performed with an automated flask sampler designed and constructed by the Flask and Calibration Laboratory (FCL) of the CAL and Max Planck Institute for Biogeochemistry (MPI-BGC) in Jena. The sampler can hold 24 flasks (4 drawers with 6 flasks each). Up to three flasks can principally be sampled in parallel (i.e. duplicates or triplicates). However, it turned out that with parallel sampling the two or three flasks are in practice flushed with different flow rates if only one flow controller is used (as in the current design of the sampler). Flasks, thus, currently do not sample exactly the same air when flushed as pairs or triples. Therefore, the Atmospheric Monitoring Station Assembly (MSA) decided to sample only single flasks (unless additional flow controllers are installed in the sampler that would guarantee identical flow through parallel sampled flasks).

2. Intake line: There will be only one line from the tower running to the flask sampler, therefore only continuous observations from this height can be quality-controlled with parallel sampled flasks (aim 1). As modelers prefer using data from the highest level of the tower (largest footprint, most representative, etc.), all flasks should be sampled from that highest level (as specified in the ICOS Atmospheric Station Specification Document, https://www.icos-ri.eu/sites/default/files/cmis/ICOS%20Atmospheric%20Station%20specifications%20Version%201.3%20-%20November%202017.pdf).

3. Sample integration and timing: Flasks should be sampled as integrals, i.e. the collected sample should represent a real mean of ambient air (e.g., a 1-hour mean, comparable to current model resolution). Synchronizing in situ continuous observations and integrated flask sampling is similarly important for quality control issues. Further, for comparison reasons, continuous observations should be kept at the flask sampling height during the entire flask sampling period (i.e. no switching of *in situ* intake heights during flask sampling, no profile information available).

4. Flask size: The original choice of flask size for sole analysis in the FCL (aims 1 and 2) was 2 liters. Precise $^{14}CO_2$ analysis does however, require/prefer larger air volumes, i.e. sampling into 3-liter flasks. Both flask types can be sampled in the automated flask sampler, however, the four drawers in the sampler must be filled with only one flask size (either 6 x 2-liter or 6 x 3-liter flasks). Similarly, the shipment boxes (holding 12 flasks each) have been adjusted to the flask size and are different for 2-liter and 3-liter flasks. It was suggested/decided by the CAL that the standard flask size should be 3-liter, but existing 2-liter flasks can still be used (see below).

5. Average number of flask analyses per station: While the capacity for flask analysis at the FCL has been designed for a total of ca. 100 flask analyses per station per year, the capacity for $^{14}CO_2$ analyses in the Central Radiocarbon Laboratory (CRL), which are performed AFTER analysis of all other components at the FCL, is only about ¼, i.e., on average, 25 samples per station per year. Consequently, all flasks will be shipped from the station to the FCL and after analysis a subset, preferably 3-liter flasks, will be shipped for further analysis to the CRL. The way back to the stations after analysis will be reverse: all flasks (also those, which were analyzed at the CRL) will be leak-tested and conditioned at the FCL before dispatch to the stations.

6. Flask handling at the station: Flasks need to be installed and removed manually from the sampler. Remote stations are normally visited about once per month by a technician who would do this work. Also, at least the flasks sampled to meet aim 1 should be shipped to the FCL not much later than one month after sampling, so that a potential bias between in situ and flask analyses is detected without too much delay. Analysis of (3-liter) flasks in the CRL is less urgent, therefore shipment of one box every 4-6 months would be acceptable.

3 Sampling strategy to serve aims 1 and 2

In order to best monitor the compatibility of better than 0.1 ppm for CO_2 between a 1-hour mean continuous in situ measurement and a flask sampled in parallel integrated over one hour, mid-day situations with low ambient concentration variability should be chosen. These situations occur most frequently when the actual footprint is stable and/or has no large point sources or hot-spot areas. At coastal stations one would then sample only the marine sector to support aim 1, but this would introduce a bias in the representativeness, i.e. no continental footprints would be sampled. Also at the continental ICOS tower stations targeting for stable footprints without point sources could introduce a sampling bias. When aiming at representative data of components that are only measured in flasks (aim 2), such footprints may not cover the entire influence area of the station. To illustrate this potential sampling bias, Fig. 1 compares footprints calculated from 10-day back trajectories for the ICOS station Gartow (Germany) for July 2017 at the highest level (344 m) at mid-day. The left panel shows a composite of all 31 footprints at 12 h UTC (13 h LT), while the right panel displays only those situations where the ambient CO_2 variability from 11 to 15 h was less than 1 ppm. Under this stability restriction only the North Sea sector was sampled. Collecting samples randomly every 3rd day to best meet aim 2 would, however, rather cover almost the entire footprint of that month (Fig. 1 middle panel). Figure 2 shows a similar evaluation for the station OPE (France) for the same month (July 2017) and the maximum sampling height of 120m. Here sampling of low variability situations only would also bias the footprint to the western sector.

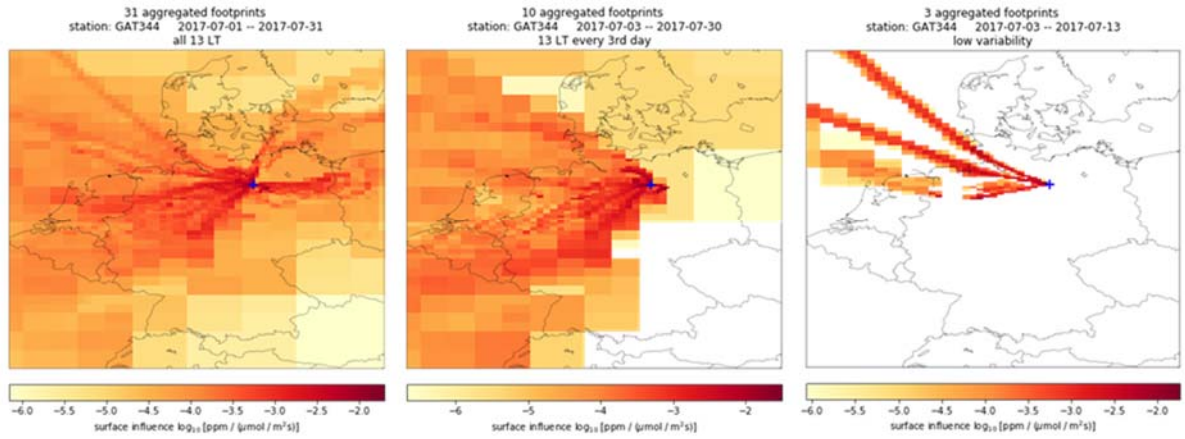

Figure 1: Gartow aggregated footprints at 344m for July 2017. The left panel shows the footprints for all 31 days at 13h LT, the middle panel the same footprints, but sampled only every 3rd day and the right panel shows the three days with low variability (<1ppm) between 11h and 15h (note different color code scaling).

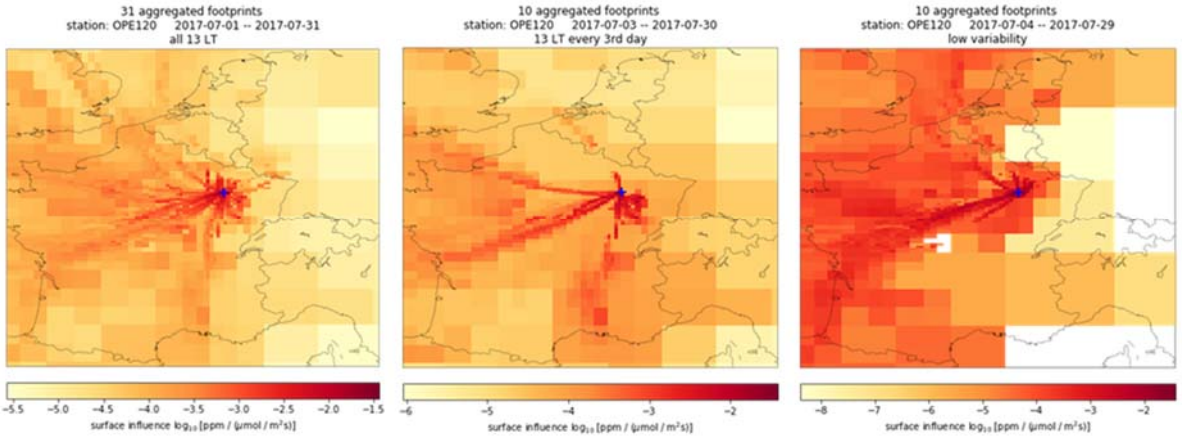

Figure 2: Aggregated footprints for OPE at 120m and 13h LT in July 2017. The left panel shows all 31 footprints, the middle panel only every 3rd day and the right panel shows only situations with low variability (<1ppm) between 11h and 15h (note different color code scaling).

The question now is, would we collect with random sampling every 3rd day enough flasks to significantly detect a bias of 0.1 ppm as is suggested by WMO as compatibility goal for monitoring stations in the Northern Hemisphere [WMO 2018]? On one hand, this depends on the frequency of situations with low ambient variability and on the other hand it depends on the measurement precision of the flask samples and the exact synchronization of their collection with the in situ measurement. To answer the question of acceptable ambient air variability thresholds for reliable bias detection, we have conducted two types of experiments at the Heidelberg pilot station.

For the first test, ambient air was collected with the Heidelberg ICOS prototype flask sampler from the 30 m intake line from the roof of the Institute's building. Flasks were collected during a period of about three months from June to August 2018. Single flasks were flushed with constant flow rate of 1.2 – 1.4 l min^{-2} at an over-pressure of 1.6 bar for one hour and were subsequently analyzed at the FCL in Jena. Individual CO_2 results were then compared with the parallel conducted continuous in situ measurements. The in situ measurements had to be weighted with an exponential function to account for the exponential weighting of the sampled air in the flask (see Fig. 3, left panel), i.e. consider that most of the air in the flask stems from the last 5-10 minutes.

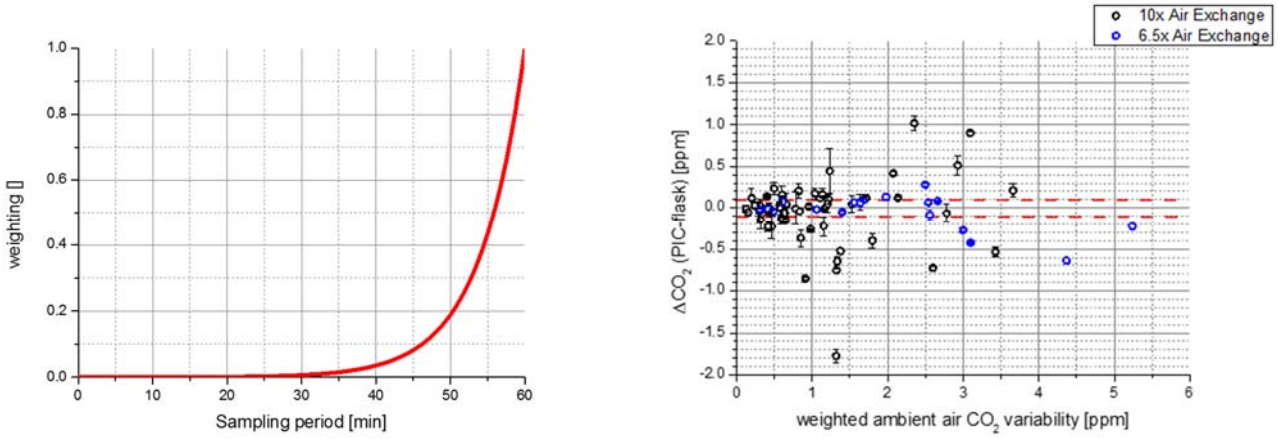

Figure 3: left: Exponential weighting of in situ measurements to allow comparison with results from flasks sampled with constant air exchange over one hour. Right: Difference between continuous ambient CO_2 observations in Heidelberg (PIC) with parallel sampled flask results. The continuous concentrations have been weighted with the exponential function displayed in the left panel. The dashed red lines indicate the 0.1 ppm compatibility goal.

During this summer testing period, ambient air variability was up to ±5.2 ppm. Results of the comparison between in situ and flask are displayed in the right panel of Fig. 3. The mean of all ΔCO_2 values for this period was -0.05 ppm with a standard deviation of 0.37 ppm. ΔCO_2 stayed within ±0.3 ppm if ambient variability was below about 0.5 ppm. For these data the mean was -0.02 ppm with a standard deviation of ±0.11 ppm. From these results we are confident that a bias between in situ and flask observations of about 0.1 ppm can be detected by averaging a few comparison events, e.g. after one month when collecting one flask per week at low ambient variability.

A second test was conducted in Heidelberg during the period of September 2018 to February 2019, where we mounted into the flask sampler an additional flow controller with a larger range to be able to collect flasks with decreasing flushing rate and then obtain a real hourly mean ambient air sample. This makes the flask results representative and comparable to model results; it also simplifies comparison with in situ measurements. The respective results from 1/t flask sampling in Heidelberg are displayed in Fig. 4. During these experiments we obtained three outliers, where flask CO_2 results have been up to more than 3 ppm higher than the in situ measurements. CH_4 and CO (not shown) did, however, compare very well between flask and in situ. Although the additional mass flow controller we used in these experiments had some problems to exactly regulate the flow rate, we did not find obvious reasons for malfunction of the sampling system. The only explanation for the outliers may be contamination of the flask with room air, which is elevated in CO_2 but not in CH_4 or CO compared to outside air.

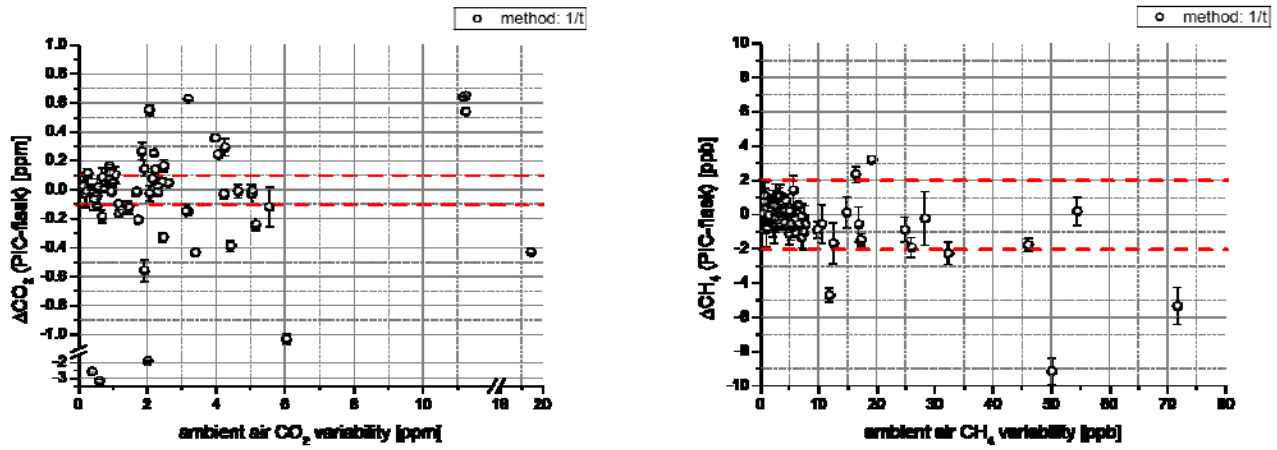

Figure 4: Left: In situ – flask CO_2 results obtained with the with a 1/t flask flushing method to sample a real hourly mean sample. Right: Same as left panel but for CH_4. The red dashed lines indicate the compatibility goals for these two components.

If we disregard the two CO_2 outliers, the 1/t sampling test results in Heidelberg give confidence that flask samples collected over one hour at low ambient CO_2 (and CH_4) variability seem to be well suited to meet our aim 1 of ongoing quality control at class 1 stations. The mean differences between in situ and flask measurements for CO_2 have been 0.0 ppm with a standard deviation of ±0.05 ppm (n=8) for CO_2 and for a variability <10 ppb for CH_4 the mean difference was -0.07 ppb with a standard deviation of 0.7 ppb (n = 70). More such tests and practical experience at ICOS stations are needed to support these findings.

Table 1: Number of low variability hourly means (σ_{CO2} < 0.5 ppm for 1-minute values) at 13h LT and potential fossil fuel CO_2 events based on a ΔCO threshold of > 40 ppb from ambient measurements in 2016-2018

	Gartow	Gartow	OPE	OPE	Kresin	Kresin	Hyltemossa	Hyltemossa
	σ_{CO2} < 0.5 ppm	ΔCO > 40 ppb	σ_{CO2} < 0.5 ppm	ΔCO > 40 ppb	σ_{CO2} < 0.5 ppm	ΔCO > 40 ppb	σ_{CO2} < 0.5 ppm	ΔCO > 40 ppb
Jan	26	11	26	11	20	5	28	9
Feb	22	14	23	12	17	19	21	9
Mar	24	14	21	5	19	18	28	11
Apr	24	7	24	0	22	4/5	23	1
May	23	4/1	17	2	10	3	24	0
Jun	13	1	21	2	13	0	23	0
Jul	10	2	29	1	12	0	18	0
Aug	19	4	27	2	(4)	2	21	3
Sept	17	10	20	3/4	13	1	24	0
Oct	20	9	24	12	21	1	23	2
Nov	15	11	21	13	23	7	22	6
Dec	20	6	22	21	21	5	25	12

To evaluate how frequent afternoon events with less than 0.5 ppm variability occur at typical ICOS stations, we evaluated real in situ CO_2 data from the stations Gartow (DE), OPE (FR), Kresin (CZ) and Hyltemossa (SE). In the years 2016 to 2018 (at least 1 year of data at each station) at all four stations the number of situations with hourly standard deviations smaller than 0.5 ppm was at least 10 hours per month and on average over the one year of existing data between 56% and 78% of all 13h LT (Table 1). The number of events was slightly lower during early summer than during the other seasons. For an entire year, more than half of all midday hours had thus standard deviations below 0.5 ppm CO_2.

4 Realization of flask sampling to serve aim 3 (catching events with high regional $ffCO_2$)

While ^{14}C analyses on integrated CO_2 samples at ICOS stations aim at monitoring the *average* fossil fuel CO_2 level at the sites, $^{14}CO_2$ analysis on flask samples focusses on monitoring areas in the footprint of the sites with particularly large fossil fuel CO_2 emissions. The fossil CO_2 contribution in a sample should be large enough to determine it with an uncertainty of less than 30% (i.e., larger than 4-5 ppm). Further, as sample preparation for ^{14}C analysis is very laborious, one should know beforehand, if a sample potentially contains a significant regional fossil fuel CO_2 component. This could either be found out with Near Real Time transport model simulations or using the in situ observations at the station. A good indicator of significant fossil fuel CO_2 at a station is the ambient CO concentration [Levin and Karstens, 2007], a trace gas that is monitored continuously at all ICOS sites. A preliminary threshold of 40 ppb CO above local background was chosen to identify episodes with a significant $ffCO_2$ component.

An important question is, if at the ICOS stations we can at all sample regional fossil fuel CO_2 signals > 4 ppm, and, if yes, how often do we experience such situations and which emission areas are then sampled. Note that, in order for the flask results to be used in transport model investigations, similar to all other flask samples, also $^{14}CO_2$ flasks should be collected during early afternoon when atmospheric mixing can be modelled with good confidence. As in the previous section, we investigated these questions for the stations Gartow and OPE, first theoretically with STILT model simulations for 2017 transporting EDGARv4.3 emissions to the two sites. As a second step, we investigated real continuous CO_2 and CO observations from both sites in 2017.

4.1 Investigation of afternoon fossil fuel CO_2 events at Gartow

Figure 5 shows in the upper three panels STILT-simulated CO_2 and CO mole fractions (13h LT values highlighted) as well as the different regional CO_2 components at 13h LT relative to the European model background (CO_2 concentration at the border of the STILT model area) at Gartow 344 m in July 2017. During three afternoons, i.e. on July 1, 7 and 27, modelled $ffCO_2$ was larger than 4 ppm (indicated by red dots in the upper and the third panels). At the same time, the modelled CO offset was elevated, but not reaching 40 ppb. If the ΔCO threshold would have been reached, this would be indicated in the CO_2 plot by a magenta cross. The CO offset was estimated relative to the minimum modelled CO concentration of the last 7 days (red line in second panel). Comparing modelled concentrations with observations (fourth panel), the agreement is very good, particularly during afternoon hours. Deviations of the model simulations from observations are larger during night and in summer, when the model seems to underestimate the measured concentration pile-up. This model deficiency is the reason why we want to collect the flask samples at midday or in the afternoon, making sure the data can best be used in inversion estimates of fluxes.

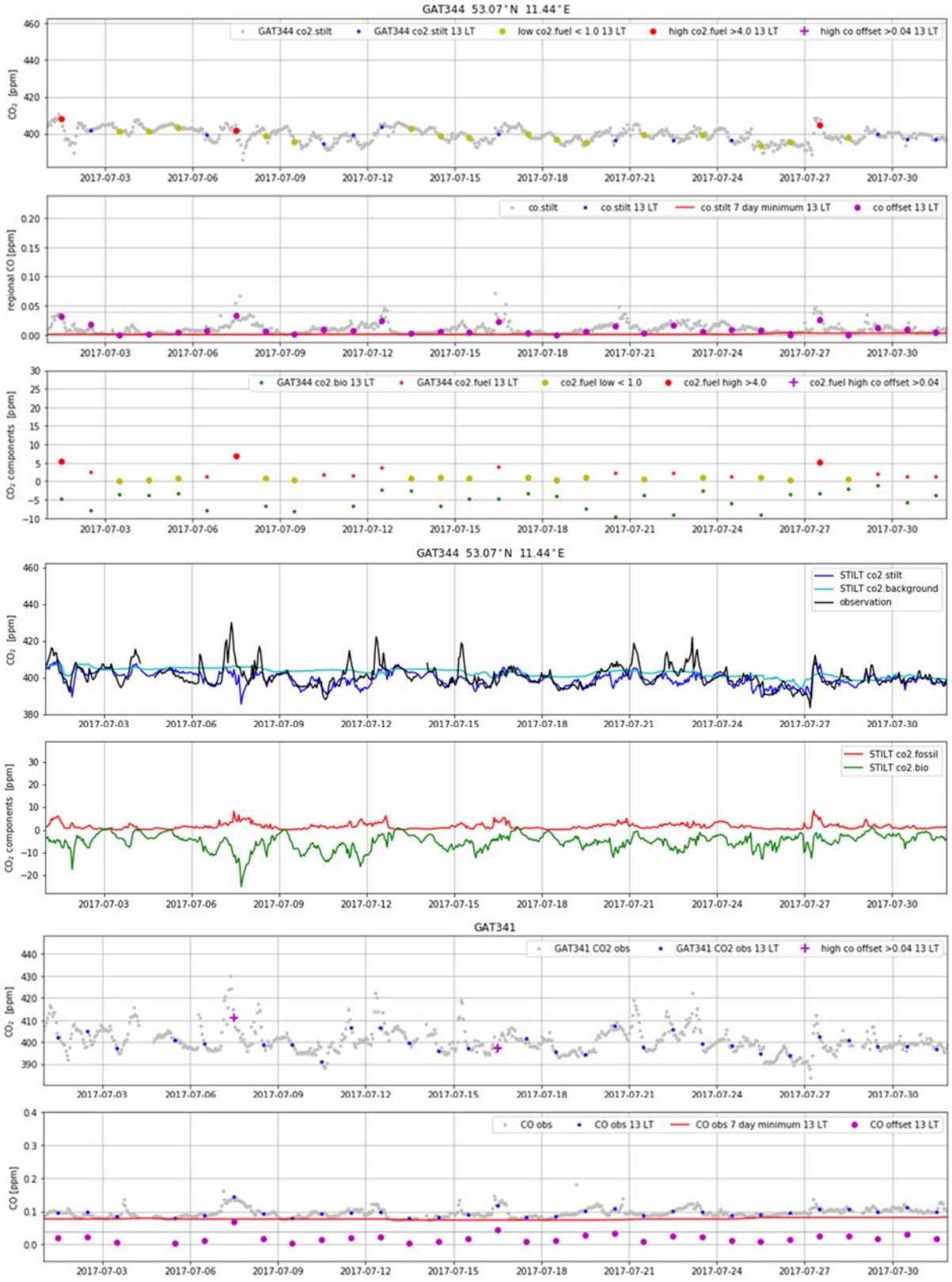

Figure 5: Variability of modelled (upper five panels) and measured (lower panels) CO_2 and CO concentrations at Gartow at the 344/341 m level in July 2017. Afternoon values are highlighted with colored symbols, situations with elevated $ffCO_2$ based on modelled or measured CO would be marked with a magenta cross in the CO_2 records.

The fifth panel in Fig. 5 shows the modelled CO_2 components and indicates the generally moderate fossil fuel CO_2 signal at Gartow in July. Indeed, based on the real CO observations shown in the lowest panel of Figure 5, summer situations with potentially high $ffCO_2$ concentrations are rare (1-5 cases). However, starting at the end of August,

such events become more frequent, and CO can well be used as a criterion for $ffCO_2$ in collected flasks. In October 2017 the measured CO offsets rather frequently exceeded the threshold of 40 ppb, in particular in the observations (see Fig. 6).

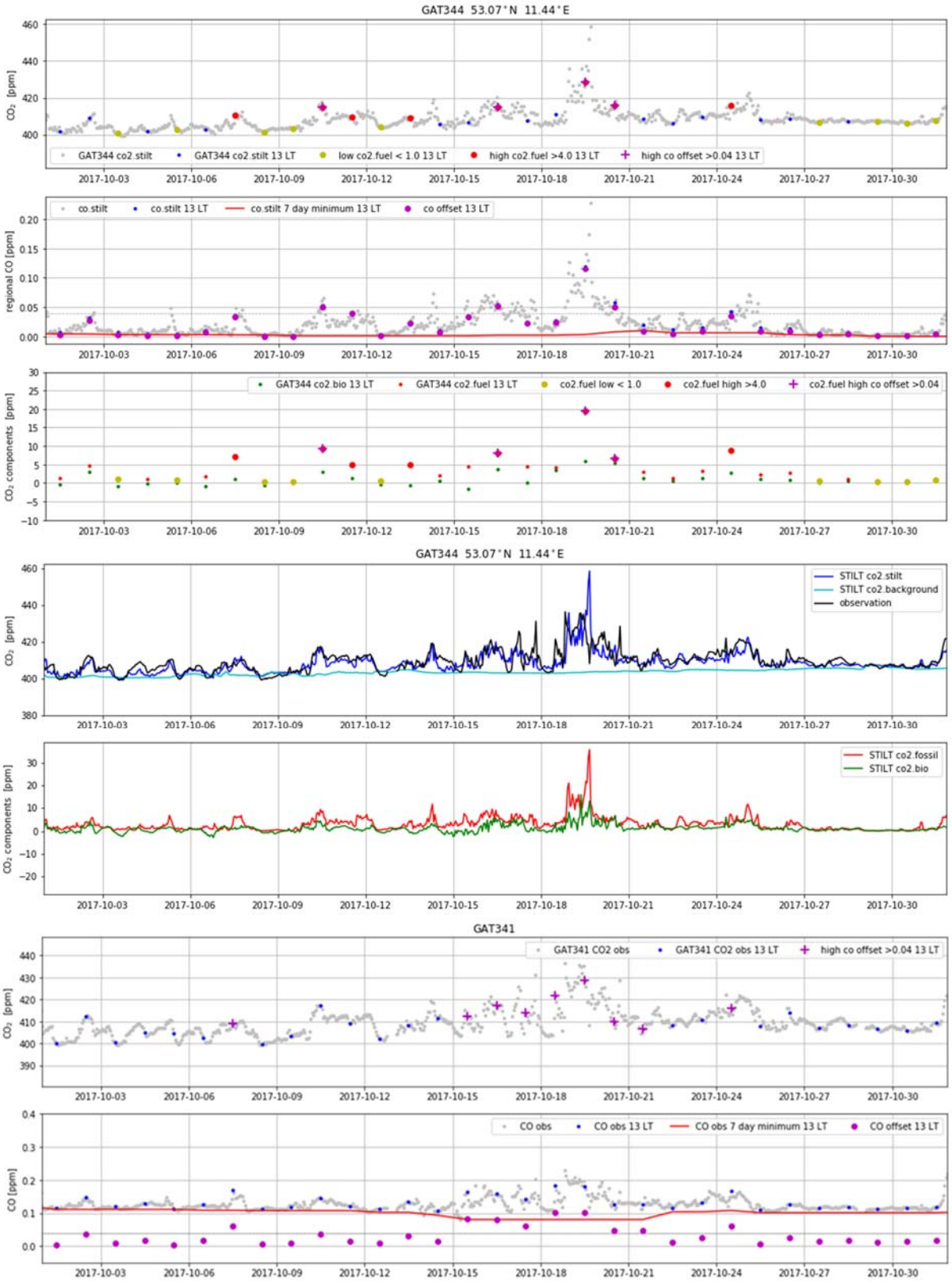

Figure 6: Same as Figure 5, but for October 2017.

The aggregated footprints of the three afternoon situations with STILT-modeled $ffCO_2$ > 4 ppm in July 2017 are displayed in Figure 7 (upper panels). They show south-westerly trajectories and a dominating surface influence from the highly populated German Ruhr area, but also some influences from large emitters (e.g., power plants) in North western Germany and at the Netherland's North Sea coast. The main influence area with high $ffCO_2$ emissions in October 2017 (Figure 7, lower panels), now show also Berlin as a significant emitter and some "hot spots" close to the German borders in the south-east.

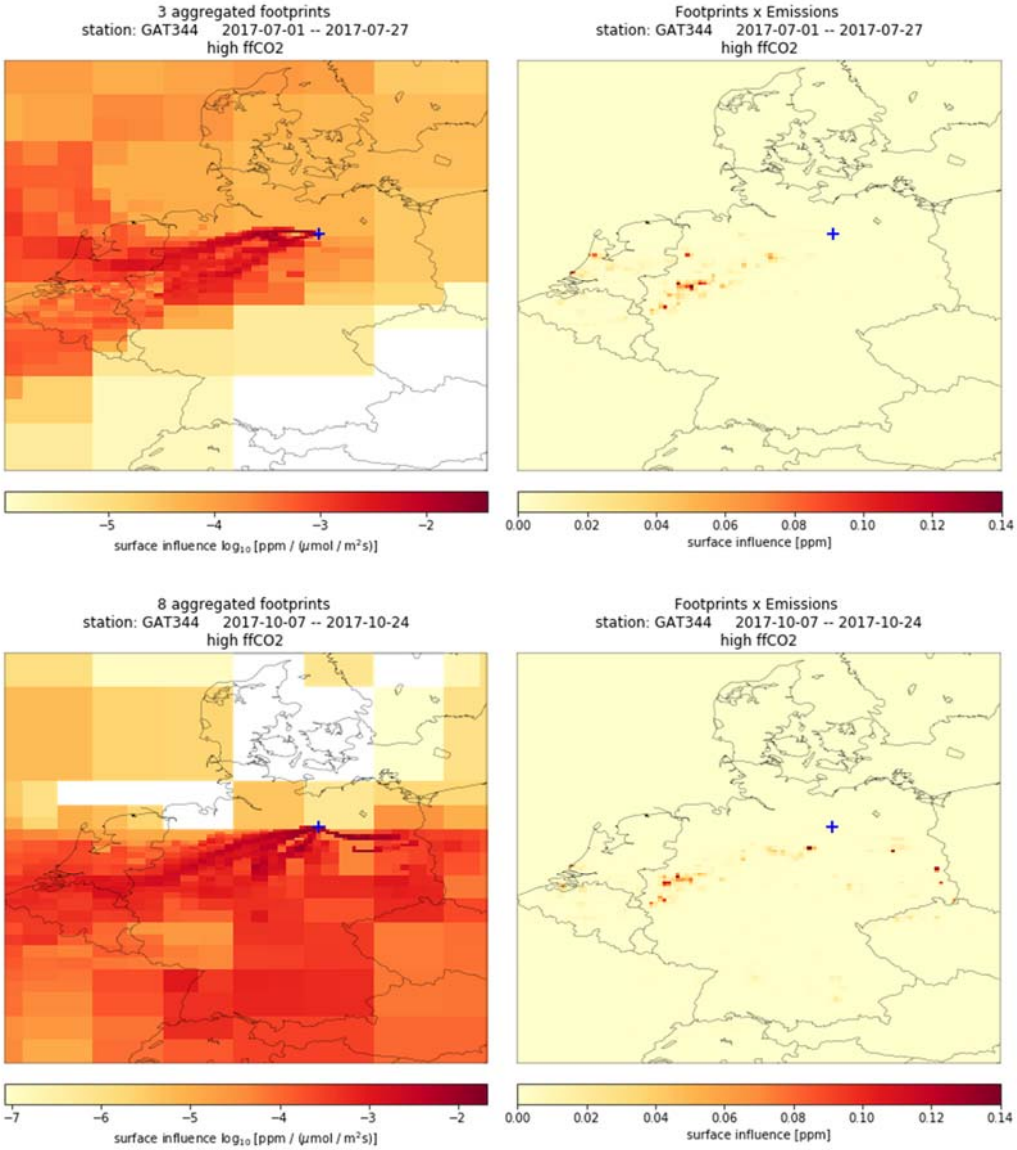

Figure 7: Aggregated footprints with elevated $ffCO_2$ (left panels) and the corresponding surface influences (right panels) for Gartow in July 2017 (upper panels) and October 2017 (lower panels)

4.2 Investigation of afternoon fossil fuel CO_2 events at OPE

To be able to make a direct comparison of the sizes of fossil fuel CO_2 signals at Gartow (341 m) and OPE (120 m), we display modelled and measured data from OPE in October 2017 in Figure 8 (note that in July 2017, no significant

$ffCO_2$ signals are measured or simulated for OPE). On October 31st, $ffCO_2$ at OPE was simulated higher than 4 ppm and a CO offset larger than 40 ppb. The respective trajectory on that day was from the North, a source area close to the border to Luxemburg (Fig. 9).

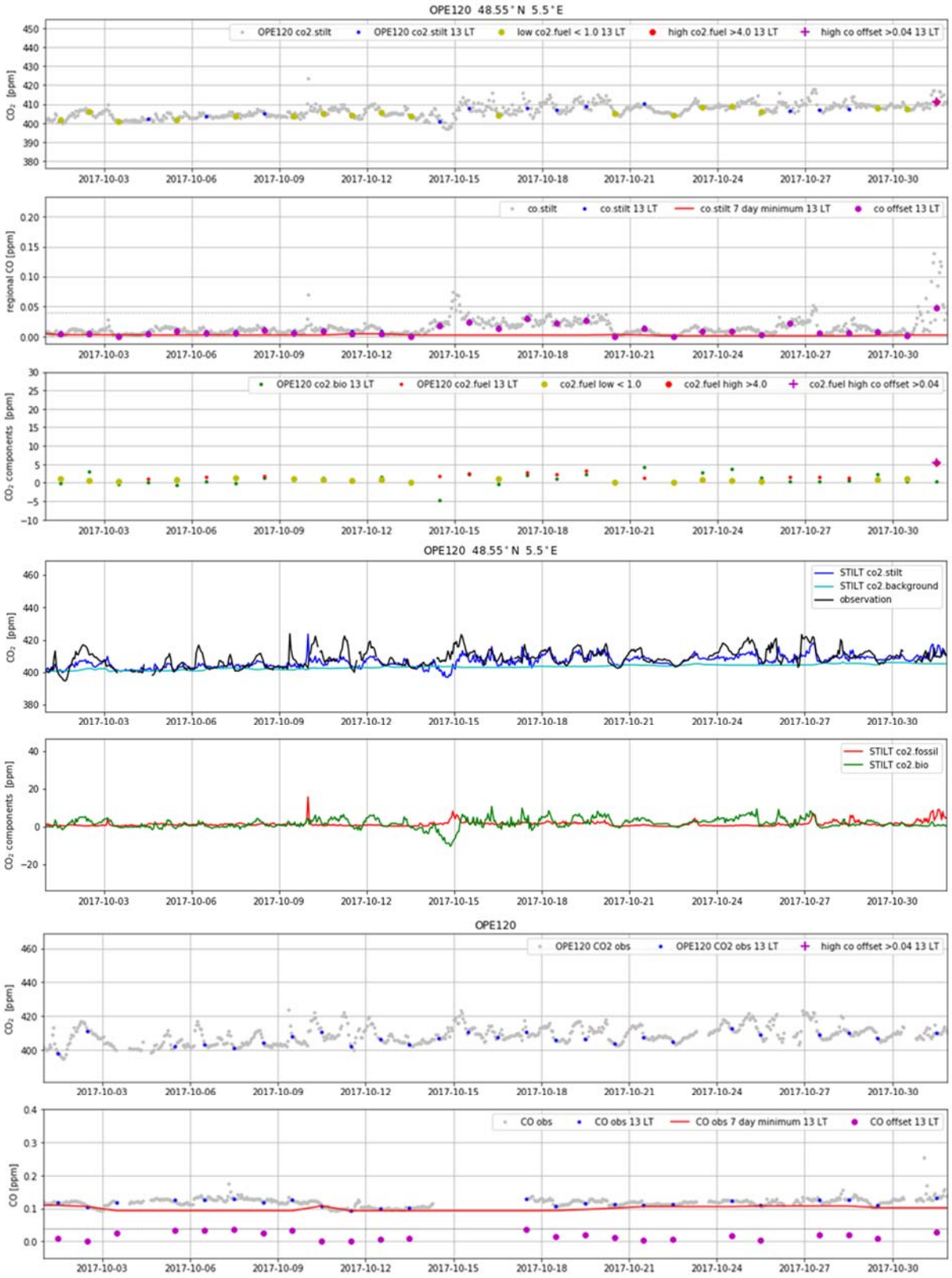

Figure 8: Variability of modelled (upper five panels) and measured (lower panels) CO_2 and CO concentrations at OPE at the 120 m level in October 2017. Afternoon values are highlighted with colored symbols, situations with elevated $ffCO_2$ based on modelled or measured CO are marked with a magenta cross in the CO_2 records.

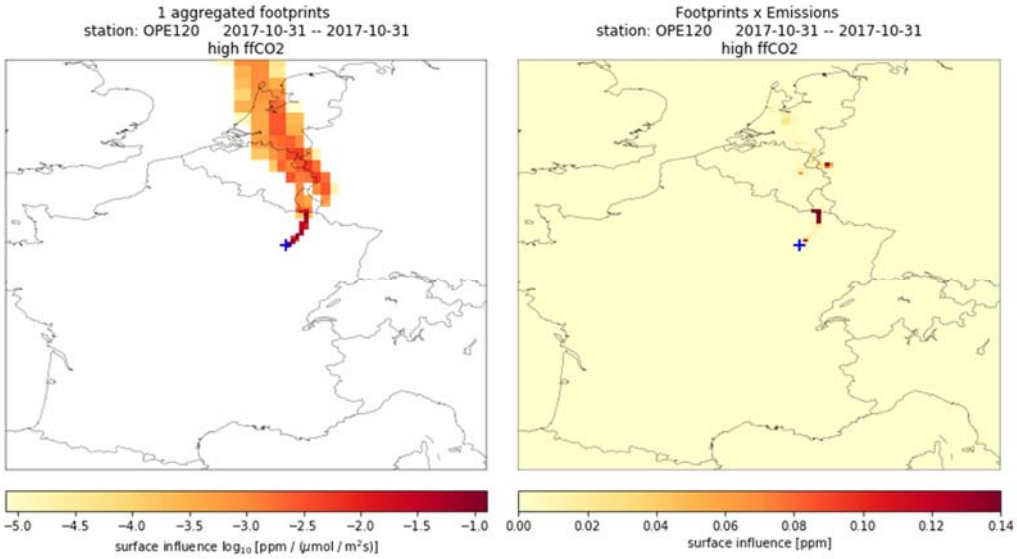

Figure 9: Single footprint with modelled elevated $ffCO_2$ (left panel) and the corresponding surface influence (right panel) for OPE in October 2017

Measured CO offsets were, however, never equal or larger than 40 ppb in October 2017. The threshold to retain a flask for ^{14}C analysis will thus have to be slightly lower here than at Gartow, at least in summer. During spring, autumn and winter a few $ffCO_2$ events were identified in the OPE observations. The emission areas were then very often located north of OPE, but in some cases also Paris was seen. All-in-all, situations with high $ffCO_2$ were less frequent at OPE than at Gartow, also in winter.

Similar analyses of potential (and measured) $ffCO_2$ signal sizes as discussed here for Gartow and OPE have been made for Kresin and Hyltemossa, but can be extended to all ICOS stations with a Jupyter notebook developed by the Carbon Portal (for a static example of the Jupyter notebook see https://www.icos-cp.eu/jupyter/examples). The number of potential fossil fuel CO_2 events based on measured data and using the same estimate for the CO background are listed in Table 1 for all months were observations exist at the four currently studied sites in 2016 – 2018. It is obvious that at these ICOS stations during summer very few high $ffCO_2$ events can be expected with the current threshold. Lower potential thresholds could be used for summer, taking into account larger uncertainties of the $ffCO_2$ component. An alternative could be to concentrate ^{14}C analysis on flasks to autumn, winter and spring.

5 Suggestion for a simple flask sampling strategy and requirements for its practical realization

In general, the ICOS flask sampling strategy will evolve with time, experience and modelling capabilities. At the moment we foresee a tiered approach, which will be implemented in the next years:

Tier 1: Simple time triggered flask sampling strategy (as outlined below)

Tier 2: Direct, station internal, coupling of in-situ analyzer and flask sampler to allow for in-situ triggered sampling of ^{14}C flasks to meet aim 3

Tier 3: Triggering of the flask sampler via prognostic GHG data e.g. based on CAMS forecast.

Analyses of STILT-simulated and in situ measured CO_2 and CO concentration data at Gartow (341/344m a.g.l.) and at OPE (120 m a.g.l.), combined with preliminary experiments of parallel afternoon flask and in situ measurements at the Heidelberg pilot station lead to the following conclusions:

1. As long as noon/early afternoon ambient hourly standard deviations of CO_2 are smaller than 0.5 ppm, flask sampling with the constant air exchange method in comparison to exponential weighting of in situ measurements allows detection of biases larger than 0.1 ppm between both measurements on the basis of e.g. weekly comparisons within a one month period.

2. In order to sample representative hourly mean ambient conditions, flasks should, however, be sampled with 1/t decreasing flow rate. This requires modification of the standard flask sampler by adding/substituting the flow controller by one with much higher flow rate. Tests of such a modified system in Heidelberg yielded similarly good results as the tests with constant air exchange. Modification of the existing ICOS flask samplers is, thus, recommended.

3. At the four stations Gartow, OPE, Kresin and Hyltemossa, which were included in this study, random flask sampling every 3 days around noon, independent of ambient concentration variability, should provide a sufficient number of flasks to detect a bias > 0.1 ppm between flask and in situ measurement within one month sampling period. At the same time, this “random” flask sampling provides a simple and reliable strategy to monitor the entire catchment area of a station.

4. Afternoon concentrations with regional $ffCO_2$ signals >4 ppm (resp. ΔCO > 40 ppb) are rare at the yet studied four ICOS stations. In order to have a maximum (i.e. 66%) chance to catch such an event, 3-liter flasks for potential ^{14}C analysis should be sampled at all afternoons, where no (2-liter) flasks for aims 1 and 2 are collected. Few days after sampling, the in situ measurements at the station need to be (automatically) evaluated, either based on the in situ CO measurement or, in the future, on NRT modelling to check if a 3-liter flask with a potentially high $ffCO_2$ signal has been collected. If this has been the case, the flask should be retained for analysis. If not, the respective flask will be set to be available for re-filling in the flaks sampler controlling software.

5. In practice this strategy can be realized as follows:

- If not all flasks at a station are 3-liter flasks, fill 2 drawers of the flask sampler with 2-liter flasks and 2 drawers with 3-liter flasks
- Program the sampler to collect one 2-liter flask every 3rd day; the 12 flasks will then be filled after slightly more than one month and need to be replaced by new ones. The box with filled flasks is shipped immediately to the FCL for analysis.
- Program the sampler to fill the 3-liter flasks always on the two days between sampling of 2-liter flasks.
- Check every day, based on the continuous CO record, if a $ffCO_2$ event has occurred on the day before (use a CO “background” estimate e.g. similar to the one we used in the current study, to calculate the CO offset, ΔCO). Retain all flasks with a ΔCO above station threshold for ^{14}C analysis. If this happens to be a 2-liter flask, mark it for the FCL so that it is analyzed there only for the most “CO_2-relevant” components (needs to be discussed, which these are). Make all flasks with a ΔCO below threshold available for re-filling.
- If there are consecutive $ffCO_2$ events (ΔCO above threshold) and no 3-liter flasks are left in the sampler, free the flask of a “pair” with the lower ΔCO. Repeat this procedure to free one of two consecutive flasks for refill. If there are no consecutive 3-liter flasks left, then we should have collected a maximum of one $ffCO_2$ flask every 3rd day, i.e. less than 12 flasks per month, and a technician should come or have been at the station for replacement of all flasks.
- Once per month, when 2-liter flasks are exchanged, also the 3-liter flasks that were identified for ^{14}C analysis shall be put into the shipping box and new flasks be put into the sampler.
- Ship full 3-liter flask boxes to the FCL for concentration analysis, when 12 flasks have been retained.

6 Next steps to test and implement this strategy

As soon as possible, refurbish all ICOS flask samplers with flow controllers that allow 1/t sampling (FCL).

Further test reliability of 1/t sampling and applicability of flask sampling for QC at existing stations (KIT, HPB, (CRL))

Implement an automated warning procedure if flask-in situ difference exceeds compatibility goal (ATC & CRL)

Develop strategies and tools for trouble shooting if compatibility goal was missed (ATC & MSA)

Program ffCO_2 event sampling strategy (ATC & FCL)

Implement a background estimation routine to allow estimating ΔCO (ATC)

Identify stations (e.g. mountain stations, coastal stations), which require a different flask sampling strategy (MSA)

7 CONCLUSIONS

Developing a flask sampling strategy for ICOS is an innovative undertaking. It aims at optimizing technical requirements/efforts at the (remote) ICOS stations as well as the analytical capacities and capabilities of the ICOS Central Analytical Laboratories. The strategy was designed to meet, on one hand, the requirements for quality control, making sure ICOS data are of highest precision and accuracy and at the same time be most useful for current inverse modelling tasks to estimate continental fluxes, not only of core ICOS components such as CO_2 and CH_4, but also of trace substances, which are not monitored continuously, such as molecular hydrogen or oxygen/nitrogen ratios. Monitoring also fossil fuel CO_2 emission hot spots from ICOS stations that were - on purpose - located far away from the direct influence of such regions is a particular challenge. Experience of the coming years will show if our current strategy is useful to meet the aims. In any case, it will most probably be subject to future changes and further improvement.

8 REFERENCES

Colomb, A., Conil, S., Delmotte, M., Heliasz, M., Hermannsen, O., Holst, J., Keronen, P., Komínková, K., Kubistin, D., Laurent, O., Lehner, I., Levula, J., Lindauer, M., Lunder, C., Lund Myhre, C., Marek, M., Marklund, P., Mölder, M., Ottosson Löfvenius, M., Pichon, J.-M., Plaß-Dûlmer, C., Ramonet, M., Schumacher, M., Steinbacher, M., Vítková, G., Weyrauch, D., and Yver-Kwok, C., 2018. ICOS Atmospheric Greenhouse Gas Mole Fractions of CO_2, CH_4, CO, $^{14}CO_2$ and Meteorological Observations 2016-2018, final quality controlled Level 2 data, doi:10.18160/rhkc-vp22.

Janssens-Maenhout, G., Crippa, M., Guizzardi, D., Muntean, M., Schaaf, E., Dentener, F., Bergamaschi, P., Pagliari, V., Olivier, J. G. J., Peters, J. A. H. W., van Aardenne, J. A., Monni, S., Doering, U., and Petrescu, A. M. R., 2017. Global Atlas of the three major Greenhouse Gas Emissions for the period 1970-2012, Earth Syst. Sci. Data Discuss., https://doi.org/10.5194/essd-2017-79.

Levin, I. and Karstens, U., 2007. Inferring high-resolution fossil fuel CO_2 records at continental sites from combined (CO_2)-C-14 and CO observations. Tellus B. 59, 245–250.

Lin, J. C., Gerbig, C., Wofsy, S. C., Andrews, A. E., Daube, B. C., Davis, K. J., and Grainger, C. A., 2003. A near-field tool for simulating the upstream influence of atmospheric observations: The Stochastic Time-Inverted Lagrangian Transport (STILT) model, J. Geophys. Res., 108, 4493, doi:10.1029/2002JD003161.

Mahadevan, P., Wofsy, S. C., Matross, D. M., Xiao, X., Dunn, A. L., Lin, J. C., Gerbig, C., Munger, J. W., Chow, V. Y., and Gottlieb, E. W., 2008. A satellite-based biosphere parameterization for net ecosystem CO_2 exchange: Vegetation Photosynthesis and Respiration Model (VPRM), Global Biogeochem. Cy., 22, GB2005, doi:10.1029/2006GB002735.

WMO, 2018. GAW Report No. 242. 19th WMO/IAEA Meeting on Carbon Dioxide, Other Greenhouse Gases and Related Tracers Measurement Techniques (GGMT-2017). Crotwell Andrew; Steinbacher M. World Meteorological Organization, Geneva, Switzerland.

In this study we used the following in situ datasets from the ICOS 2018 release:

ICOS RI, 2018. ICOS ATC CO2 Release, Gartow (341.0 m), 2017-04-11-2018-05-31, https://doi.org/11676/vcqlKCuljtd3G-Fak1tt3qD8

ICOS RI, 2018. ICOS ATC CO Release, Gartow (341.0 m), 2017-04-11-2018-05-31, https://doi.org/11676/3lmUvoUYAXpzI7kVYdYVLUWl

ICOS RI, 2018. ICOS ATC CO2 Release, Observatoire Pérenne de l'Environnement (120.0 m), 2016-08-18-2017-11-23, https://doi.org/11676/i2asBCpa_1FnrKUwI-TMLT6r

ICOS RI, 2018. ICOS ATC CO Release, Observatoire Pérenne de l'Environnement (120.0 m), 2016-08-18-2017-12-07, https://doi.org/11676/uvHlewSsNaMwF7Cp5aezuwxS

ICOS RI, 2018. ICOS ATC CO2 Release, Křešín u Pacova (250.0 m), 2017-04-12-2018-05-31, https://doi.org/11676/Td5c-8z6gUHtWfk6A8-nF7-C

ICOS RI, 2018. ICOS ATC CO Release, Křešín u Pacova (250.0 m), 2017-04-12-2018-05-31, https://doi.org/11676/4-6BhuUTjVIv8s24f63zsF72

ICOS RI, 2018. ICOS ATC CO2 Release, Hyltemossa (150.0 m), 2017-04-17-2018-05-03, https://doi.org/11676/Z0LhpFBypn1SEhKlVPryBfZX

ICOS RI, 2018. ICOS ATC CO Release, Hyltemossa (150.0 m), 2017-04-17-2018-05-03, https://doi.org/11676/qNshckvwoUdRD9iQ0l1SR4Wo

9 DEFINITIONS, ACRONYMS AND ABBREVIATIONS

ATC:	Atmospheric Thematic Center of ICOS
CAL:	Central Analytical Laboratories of ICOS
CAMS:	Copernicus Atmospheric Monitoring System
CRL:	Central Radiocarbon Laboratory of ICOS
EDGAR:	Emission Database for Global Atmospheric Research
FCL:	Flask and Calibration Laboratory of ICOS
$ffCO_2$	Fossil fuel CO_2
GAT:	Gartow (ICOS station in Germany)
GHG:	Greenhouse Gas
HTM:	Hyltemossa (ICOS station in Sweden)
ICOS:	Integrated Carbon Observation System
KIT:	Karlsruhe Institute of Technology (ICOS station in Germany)
KRE:	Kresin (Czech station of ICOS)
MPI-BGC:	Max Planck Institute for Biogeochemistry (Jena, DE)
MSA:	Monitoring Station Assembly
OPE:	Observatoire Pérenne de l'Environnement (ICOS station in France)
QC:	Quality Control
STILT:	Stochastic Time Inverted Lagrangian Transport model
UHEI:	Heidelberg University
VPRM:	Vegetation Photosynthesis and Respiration Model