Author comment on "Evaluation and optimization of ICOS atmospheric station data as part of the labeling process" by Camille Yver-Kwok et al.

Anonymous Referee #3
Received and published: 9 September 2020

We thank the reviewer for their positive review and their constructive and helpful comments. We answer them below, highlighted in bold and italic.

This work describes the specific workflow and quality assurance processes within the ICOS network. The authors did a wonderful job composing a well written, compre hensive paper about the difficulties and challenges faced in the world of high-quality greenhouse gas observations. I think this work is a substantial contribution to the knowledge base and scientifically very important. Everyone in the field relies on intercomparable high-quality long-term observations but is rare to see papers that clearly describe and outline the hard work that lies behind them. I especially appreciate that the examples given in the paper include times where problems were found and solved, as well the realistic description of timeframes when setting up a new site. It is the lived reality in the field and will hopefully be educational to both users of the ICOS data as well as groups interested in setting up long-term observation stations. While the quality of the paper is generally high, there are some inconsistencies in the use of language mainly the interchangeable use of site names and abbreviations. Some of the parameters mentioned in the text and/or in figures have no clear explanation of how they were derived. The figures contain a lot of vocabulary and abbreviations that are probably useful within ICOS but can be confusing for the casual reader (tank and instrument numbers) without additional explanation. A lot of the figures also use very small fonts that can only be read by zooming and will not be legible in print outs at all. The colour in the figures is generally not suitable for colour-blind individuals (a lot of red and green with the same saturation levels). The descriptions of the figures in both the text and figure legends are very perfunctory even for some of the more complex figures

Specific Comments:
While the introduction gives a brief overview of greenhouse gas measurements and observation networks what the data is used for but lacks information about why we need such high-quality greenhouse gas observations and the benefits of the labelling process for the end-user of the data product.
ICOS was first designed to serve as a backbone network to monitor fluxes away from main anthropogenic sources. There the concentration gradients between European sites is typically of only a few ppm on seasonal time scales. It is this signal that is used to make atmospheric inversion where from atmospheric gradients, using atmospheric transport models, one can deduce surface emission fluxes. To correctly capture this signal of a few ppms, a high precision and integrated network is needed. As a precise example, Ramonet et al. 2020, show that a strong drought in Europe like the one seen in summer 2018 produces an atmospheric signal of only 1 to 2 ppm.

The labeling process is very useful for new stations coming into the network to ensure proper setting; good measurement practice and in the end be able to reach the precision and stability requirement of ICOS. The whole process is lead by the ICOS ATC that offers expert support to the PIs of the stations. For the end-user, the labeling process is a guarantee of high quality measurement with proper metadata description and associated traceability along the data processing.

In 2.4 general requirements and table 1, the different parameters mandatory and recommended for the different classes of station are given but not much information about why they were deemed mandatory or recommended. Two parameters in table 1, the mandatory atmospheric pressure observation at ground level and the recommended eddy covariance flux for CO2 are not mentioned in the text at all.

We will add a few sentences about why these parameters were chosen.

ICOS atmospheric network aims to provide high precision measurements of greenhouse gases, and in priority CO_2 and CH_4 which represent the main anthropogenic GHG. In-situ measurements of N_2O, the third contributor to the additional radiative forcing, was not required in the initial phase of ICOS due to the difficulty to find at this time reliable instruments enable to provide the expected precision (Lebegue et al. 2016). This gas will be added in the new phase of ICOS. Flask sampling is required at class 1 stations for quality control of in-situ measurements, and to provide additional trace gases measurements like N_2O, H_2, CO_2 isotopes (Levin et al., 2020). Other parameters are required in order to support the interpretation of the GHG variabilities, like CO as a tracer of combustions, and meteorological parameters to characterize the local winds, vertical stability along tall towers and weather conditions (P, T, RH). The eddy covariance flux have been selected in this list with the idea to characterize the local surface fluxes either from biogenic and/or anthropogenic activities and to monitor possible long term changes around the ICOS sites. So far this parameter is not required for the labeling process due to logistical difficulties to install such measurements at several atmospheric sites.

In 2.2 greenhouse gas calibration requirement: Add an example or range for the automatically filtered data.

We will add an example of such range. For example, for the cavity pressure, the range for the CRDS instruments is usually 139..8 to 140.2 Torr. These ranges are also given in Hazan et al (2016). Of course, those parameters are depending on the type of analyzer.

2.3.3 Intake line and water vapour correction tests: The text is a bit unclear, it mentions a shelter test every 6 months and testing the outside lines every year, then later in the text it mentions testing of the whole line is recommended at sites where lines are older than 10 years. Is the yearly testing of the outside line just to tests from the base of the tower to the shelter or is the yearly testing of the whole outside line only recommended for lines older than 10 years? I assume that since ICOS focusses on tall tower sites, yearly full line tests would be very expensive (hiring climbers to climb the tower twice for each test to connect and disconnect the lines).

We will clarify this section. Every 6 months, there is a shelter test and then every year the intake line test. This test can be either by injecting a target gas from the top of the lines (that we recommend for older lines only as it it complicated to implement) or by maintaining a depression in the line. For almost all towers, PIs can climb themselves so it is not a problem.

2.4.4 Calibration drift and optimization: In this section and in figure 6 the word calibration drift is used for the instrument drift that is then corrected and optimized with the calibration, this is very confusing as there is also such a thing as a drift in the concentration of calibration gases.

We will clarify the text and reformulate to talk about instrument drift instead.

2.4.7 Diagnostic parameters, Page 10, line 7: A low flow rate within the line could also be indicative of an obstruction in the line (damage to sampling line or blockage).

Thanks for that addition included in the manuscript.

3. Presentation of the 23 labelled stations Page 11 line33: The paper generally describes all the site setups in great detail, but there is no description of the 4 sites with buffer volumes. I appreciate that

everyone in the community has strong opinions about the usage of buffer volumes, but regardless of their merits or lack thereof, a more detailed description of the buffer volume setups (integration volume, flow rate and in integration time) would be helpful.
We will add some details about the buffer volumes. At the Swedish sites, buffer volumes of 8L are used with an integration time from 3.8min to 4.9min and a flushing rate between 1.6 to 2.1M/min. At SMR, buffer volumes of 5.6L are flushed at 0.325 L/min which gives an integration time of 16.9min.

4.1 Calibrations: Is the drift described this chapter and in table 5 the same instrument response drift described in 2.4.4 and figure 6? The naming is ambiguous as it implies it is the calibration that is drifting not the instrument response. There is also no description of how the drift was calculated.
Yes, we are still talking about instrument drift. As above, we will clarify and explain how we calculated the drift. When the drift looked linear enough, we used the difference between the first and last data points and divided by the number of months.

4.3 Uncertainties: Figures 16-18 contain a lot of information, while some of it is explained here in detail, other artefacts are left unmentioned or are mentioned later in the text and then not referenced. For example in Figure 16: For JFJ, The continuous instrument repeatability (CMR) is good but the long-term repeatability (LTR) is high, later in paragraph 4.6 it is explained that this is was due to a polytube Nafion, but the text does not reference figure 16. Then there is OPE that seems to have had a bias in the target value for both CH_4 and CO_2 for a while. On page13, line 21 it is mentioned that two sites show values outside of the WMO targets for CO, but it is not specified which of the subplots this refers to. 3 additional sites show larger biases after the initial test period (IPR, SMR and NOR) but they are not mentioned in this section. Later in the troubleshooting section, there is a mention of the issue with CO that was related with the use of heated inlet cups but figure 18 is not mentioned in that section and the sites are referred to with their full name.
We will comment in more details these figures and reference them in a better way in the text. In the case of OPE, it is most probably due to the manual QC that has not been done in time. It may disappear as all data have been carefully quality control for the recent ICOS release in September 2020. For CO, on Figure 18, bias to the assigned values, the black lines shows the WMO compatibility goals. In the text, we talk about GAT and HTM (next sentences). We see that HTM, IPR, NOR and SMR are slightly below the threshold over the last year while during the test period, GAT was exceeding the goals and HTM showed very noisy data. The data from the intake cups were QC invalid, moreover they were not concerning the target data that are used to produce Figure 18 only the air data.

Figure 11: This figure is not easy to read and within the text, it is just casually referred too, does it add any value? How does it help evaluate the influence of different sources (is it because it shows the different inlet heights?) whatever information it is supposed to convey is lost in the sheer amount of data (1 year 4 heights 3 compounds, plus quality assurance subplots). I could see the value of a plot like this online where you can zoom in to it. The short-term long-term target stability on the right-hand side is interesting but is not even mentioned in the text or legend of the figure and the short and long-term targets are also shown in figure 1.
We will comment this figure in more details. This figure is available on the ATC web site in daily, 10 days, monthly and yearly versions (updated every day). Indeed, on the yearly figure, we mostly look for patterns in the targets, data gaps, outliers, whereas the ambient air signals are much more visible on shorter period's figures. ...

We will clarify or correct as proposed in the technical comments. If needed, we add a more detailed answer to the comments.

Technical comments:

Page 3, line 11: First mention of WMO compatibility goals but no information what they are for the gases discussed in the paper.

Page 5, line 1: No reference for the ICOS specification document.

Page 5, line 25 mentions that Table 2. Contains the raw minute and cycle but then Table 2 contains the minute, injection, and cycle data. Is an injection not the same as a cycle? The words are interchangeably used throughout the paper for example in figure 7.

We will clarify: an injection is one sampling of 30 minutes, a cycle is the suite of the injection of each calibration cylinders (so 4 injections of different mixing ratios), the calibration is then a suite of 3 to 4 cycles.

Page 7 line 22: What is meant with the intrinsic bias of the instrument?

What is called "intrinsic bias" in the Mlab initial test report is the bias observed in the final test when the tested instrument with dry air is compared to reference instrument with dry air. We will simplify in the manuscript.

Page 7, line 31: Rephrase to clarify that the onsite water test needs to be performed if the last instrument test at MLab was longer than a year ago.

Page 10, line 20-21: Clarify that the instrument flow rate can be used to estimate the lifetime of cylinders.

Page 11, line 1: LTR, is defined as the Long-Term Repeatability which is I understand to be the 3 day average of the standard deviation of the short-term target measurement.

The text does not specify that it is based on the short-term target although it must be as the long-term target would not be measured often enough for this. The naming is confusing as the short-term target makes up the long-term repeatability but then there is also a long-term target.

Sorry about the confusion, we will clarify in the text but cannot change the term. The LTR is in opposition with the short-term repeatability calculated over 10 injections within a few hours (see Yver et al, 2014).

Page 13, line 15: Figure 12, not 13 describes the bias calculation, the text would be easier to follow it the calculation was described within the text of the chapter and or the relevant figure description.

We will correct and add some details.

Page 13, line 18 change to: The red dot.

Page 13, line 29-Page 14, line 4: Restructure this part to make it easier to read. The text references that the description of figure 19 contains how the bias was calculated but figure 19 contains no calculations. Later in the paragraph in page 14, line 4 it is clarified that the bias is (measured-reference) I assume what is described as the bias is just the difference between the concentration measured in the line vs. directly in the tank but the roundabout description makes it harder to understand than necessary.

Maybe just rephrase the second sentence to make it clear it is measured – reference as currently, it is the other way around.

Page 14, line 13: Reference water droplet test protocol used.

This test is shortly described in 2.3.3. The protocol used is not available publicly.

Page 14, line 30: If available add part numbers and manufacturer for both the recommended Nafion membrane and cryogenic water trap.

The recommended Nafion membrane is the model MD from Perma Pure. For the cryogenic water trap, no model is recommended except one that goes at least at -50°C.

Table 1: is missing a reference to the ICOS specification document.
Table 4: 5069% in KRE?.
We will correct this typo, the right value is 69%.

Figure 2: No labelling and description of parts and existing text font size too small.
Figure 11: The colour choice is not ideal (yellow on white background, and red and green lines and circles for both the target values). The numbers above the target plots are not explained anywhere, one I assume is the targets assigned concentration but Ptp is not explained anywhere. For the H2O % the target numbers are also present but empty (0.00 for all) please remove.
This figure is produced automatically (once a week) and stand as an example of what is provided in the reports. It was taken from the report for the site Torfhaus (TOH). We will add comments about the different numbers.

Figure 15-19: It is hard to visually align the legend at the bottom with the two figures above, especially in figure 16-18 where there are 2 box and whisker plots for each instrument. Adding the legend to the other 2 plots or some kind of shading or line might help.
Figure 16: Has some random signs at below the figure legend # # \.
This is due to writing the paper in R+Latex, hopefully it will disappear once processed by the editor.