Comment on amt-2022-50
Anonymous Referee #2


Summary:

Azevedo et al. 2022 present a laboratory study on lower-cost commercial CO2 sensors from SenseAir AB (type: K30). Building on previous work they extend the range of environmental variables to find a calibration strategy that allows the use of such sensor on small unmanned aerial vehicle (UAV) systems in quantitative studies. The paper is clearly structured, easy to follow and well-written. The topic itself is of great interest as UAV monitoring systems for greenhouse gases continue to increase in their importance for local and facility scale monitoring.

However, there are serious concerns about the reliability of the results. Should the authors be able to address these major comments the paper could suitable for publication, but the amount of additional work necessary might be too challenging to be achievable in a few weeks.

General comments:

The first major concern is the lack of a calibrated and reliable reference dataset for ambient air CO2 dry air mole fractions (or as referred to in manuscript: the CO2 concentration). The authors LiCOR LI-840 and LI-820 systems as a reference. However, it is obvious that neither of the two instruments was appropriately calibrated. An indoor air
concentration of 200-300ppm CO₂ as reported in Figure 3 is completely unrealistic. Ambient clean air data from NOAA can be found here: Global Monitoring Laboratory - Carbon Cycle Greenhouse Gases (noaa.gov) demonstrating that 200-300 ppm is not possible unless in artificial gas mixtures or environments. Given that the LiCORs are not calibrated it is also unlikely that they were properly tested for cross-sensitivities, offsets or non-linearity.

The LI-820 and LI-40 have known temperature dependent drifts. According to the LI-840 manual this the calibration drift is <0.4ppm per degree C 840A_Manual_10690.pdf | Powered by Box (boxenterprise.net)

How can the reader be sure that there isn’t a residual drift in the reference data?

2.) The range for the temperature calibration is too small and only a single test (at only one RH level) was conducted. Atmospheric temperatures outside the tropics frequently reach values below 10°C, which seems to have been the lowest temperature setting tested in section 3. Also, the chamber experiment holds the temperature stable for multiple hours. Is this really a realistic temperature profile for a drone flight?

3.) It is unclear how/if the lab bench setup described in figure 4 was able to provide a homogeneously heated air-stream to all instruments. It would be necessary have many more temperature sensors placed around the 2xK30 and the Li-COR to be sure they measured the same (temperature) air. Furthermore, the lab bench tests measured a response to a short-term temperature change within a few minutes, while the chamber test duration was over 6 hours with 2 hours time for instruments to equilibrate. How can those to experiments be compared? The low correlation seen in Figure 6 could well be related to the change in time scale of the experiment.
Specific comments:

L1: Suggestion to mention that this study focuses on (lower-cost) NDIR sensors.

L14: Please clarify: what does “mentioned measurement systems” refer to. Also please add a citation of studies that demonstrated the claim that no suitable measurement systems existed for local and regional scale work. Since the 2010s cavity ring down spectroscopy (CRDS) and integrated cavity output spectroscopy (ICOS) systems have been in regular use for atmospheric CO2 measurements and have allowed high-resolution and accurate measurements, even on mobile platforms (e.g. Chen et al. 2010, AMT - High-accuracy continuous airborne measurements of greenhouse gases (CO2 and CH4) using the cavity ring-down spectroscopy (CRDS) technique (copernicus.org)).

L85: You mention the need for instrument specific correction coefficients, but only decided to measure 2 instruments. How representative are two units? Martin et al. 2017 (AMT - Evaluation and environmental correction of ambient CO2 measurements from a low-cost NDIR sensor (copernicus.org)) tested at least 6 units of the K30 series.

Figure 2. It is very difficult to distinguish the time series of the different instruments.

L144, Figure 8: A linear fit does not seem appropriate for the left panels. Did you consider a non-linear instrument response?

L170: The temperature experiment, especially Figure4 clearly show that the concentration inside and outside the chamber can (and do) differ. Why do you consider the LI-840 on the outside as a reliable reference here, especially after the potential ‘unknown external interference’?

L214: This is a major concern: Can the results reported here be useful to other researchers, if the time delayed response to pressure changes is specific to the inlet and housing design?

L240: The accuracy of the instruments has not been investigated at all. No gas standards from NOAA or NIST was used here, neither were calibrated reference instruments.
The authors did demonstrate that they can reproduce measurements within 2.5ppm for same air sampling of another optical instruments under certain conditions.