

This paper evaluates the use of low-cost non-dispersive infrared (NDIR) CO₂ sensors for long-term atmospheric monitoring in rural agricultural environments. The authors developed a data-processing framework that corrects errors caused by automatic baseline correction and environmental influences, particularly temperature and pressure, improving the agreement with reference-grade instruments. The study is important as it applies sensor correction and QA/QC procedures to non-urban applications. However, it is unclear how some of the corrections were derived, as well as which sensor data are presented, that make this manuscript not publishable in its current form. I recommend review once the points below are addressed.

Major comments

- 1) The derivation of temperature and pressure corrections are one of (if not *the*) main results of this work. However these are just presented in Section 2.4.2 as Equations (3) to (6). No detail is provided on how the coefficients in these equations were determined, how the authors came to decision of applying linear corrections in some cases, exponential corrections in others. It should be made clear to the reader how these were derived. This section should be significantly expanded to allow readers and sensor users to perform the same corrections.

Line 255: similar to the point above, no detail is provided on how the ABC correction is derived and applied to the data. The authors do the describe what ABC is and the issues associated with it, but I could not find anywhere in the text describing how the authors correct for obvious step changes in the raw sensor data introduced by ABC.

Line 277: similarly, this linear scaling is not described anywhere in the text.

Figure 6: similarly, the baseline drift correction is not described.

Line 500: similar comment as line 255, the reader is not told how to correct for step changes.

- 2) The discussion of the pressure correction appears, at the same time, too detailed and overly simplistic. I do not think Eq. (1) and (2) are needed for the reader to appreciate how ambient pressure affects number density. A reference to previous work on the subject (e.g., Gaynullin et al., <https://www.mdpi.com/1424-8220/23/5/2872>, Müller et al. <https://amt.copernicus.org/articles/13/3815/2020/amt-13-3815-2020.pdf>) should suffice.

While the authors correctly identify the effect of pressure on number density, there is also an effect of ambient pressure of line broadening that is not considered (see the Gaynullin paper above). Differentiating the magnitude of the

two separate pressure effects is probably beyond the scope of the manuscript, but the additional effect of line broadening should be mentioned.

On a related note, the pressure correction derived in this work should also be compared to previous studies. Muller et al.

(<https://amt.copernicus.org/articles/13/3815/2020/amt-13-3815-2020.pdf>)

found the effect of pressure on the sensor measurements to be rather small ($\sim 0.1 \text{ ppm hPa}^{-1}$). It would be useful (for future comparisons) if the authors could quantify the pressure and temperature effects from their work in a similar manner.

- 3) The main figures depicting timeseries (and text therein) are way too small to be readable or interpretable. Also it is never clear which sensors the plotted data originated from. The study used two distinct sensor systems (Magic Box and Air Gradient). The authors should revise the captions for every figure and state clearly where the data plotted came from. Currently it only says “LCS CO₂ data”, which is not particularly useful.
- 4) Similarly to point 3), the conclusions only apply to some of the sensors tested (Air Gradient) and not to others (Magic Box). The authors should provide more details on what conclusions apply to which sensor system, and also comment on the limitations they have identified.
- 5) There is a lack of quantification to some of the statements in the text:

Lines 410-411: can the authors add some stats to back up the “excellent agreement”?

Lines 427-440: phrases such as “daytime errors (...) were substantially reduced”, “short duration enhancement”, “stronger influence” are of little use if not backed up by quantitative figures.
- 6) Section 3.5: why are the internal RH readings from the two Magic Boxes in Fig 10 very different from the reference ambient RH, whilst they seem to agree much better with the reference value in Fig 11? Is this just an effect of temperature? If so, it would help to show the air temperature too. The reference ambient RH at the two sites are very similar (slightly higher at the rural site, but similar range), so it is difficult to understand why the agreement with the internal readings is so different.
- 7) There should be a discussion section in which the results of this paper are compared to previous studies (e.g., the two references in point 2 above).

Minor comments

Line 1: Surely the title should read “Evaluating low-cost NDIR CO₂ **sensors** for atmospheric observations **s** in rural settings”?

Line 16 (first occurrence) and throughout the manuscript (e.g., lines 92, 152, 165): the authors use “LCS NDIR CO₂ sensors” and “LCS CO₂ sensors” repeatedly. However the “S” in LCS already contains the word “sensor”, so the formula does not quite read correctly. I suggest the authors revise this throughout to “low-cost CO₂ sensors” or, more simply, “CO₂ sensors”.

Line 24: long sentence. Add comma after “summer”: “... in summer, were obtained when compared to a...”

Lines 27/28: move “15 ppm” to later in the sentence to improve flow: “... which is more than three times the typical change in the urban environment (**15 ppm**).”

Section 2.1: the information about the two sites is very minimal. Can the authors provide geographical coordinates for the rural site, as well as information such as distance from the nearest urban centre? It would also be useful for readers not familiar with either site to add information on the dominant CO₂ sources and processes at both locations.

Line 98: it should read “Figure **1b**”.

Lines 115-118: it would be useful to add the distances of each airport from the two sites, to better put these locations into context.

Line 119: the section title should read “2.3 Low-Cost CO₂ Sensor Nodes **and reference instrument**”.

Line 125: do the authors mean Section 2.4.1?

Line 127: long sentence that needs breaking up. “...for a period of 12 days at the urban site (7 to 18 December 2022). **The** nodes were then deployed at the rural site ...”.

Line 129: I think this strictly classifies as an adjustment or correction, not a calibration. Change “calibrated” to either “adjusted” or “corrected”.

Line 148: Add “as the”: “... to reduce power demand in the sensor system, **as the** S8 average current is...”

Lines 153-161: Please describe how the Licor was calibrated. What standards were used, how often, etc.

Line 160: change “instrumentations” to “instrumentation” (singular).

Line 161: change “platform” and “type” to “platform**s**” and “type**s**” (plural).

Table 1: add an extra column with the number or sensor nodes used in each deployment.

Line 177: the authors should either provide a reference or a brief description of how the detrending should be performed.

Line 178: the way “section 3.3” appears makes it look like the detrending is described in section 3.3, but instead the authors mean that the measurement framework is presented in that section. I would suggest changing it as follows: “Users should therefore characterise this drift by detrending the data prior to applying the environmental correction framework presented in this work (**Section 3.3**)”.

Line 188: there is an extra space between the words “impact” and “sensor”.

Line 195: P_{lab} is defined but never used in the rest of the text

Line 219: change to “...represents mixing ratios for...”

Line 225-226: move “IR-high and IR-low” to later in the sentence for better flow: “An NDIR sensor measures CO₂ concentrations by comparing two infrared (IR) signals from the same IR source at two distinct wavelengths (**IR-high and IR-low**)”

Line 253: It would be useful to point the reader to where exactly in Figure 2 they should be looking at. Is it around 12:00 on Sept 20th in panel a?

Lines 278-282: this section was very confusing. The author should state that which metrics being mentioned here are from the application of the full correction protocol, and which ones from the linear scaling.

Figures 3, 4, 5, 7, 8, S1 and S2: are data from a single sensor system? Do similar plots for every sensor system deployed exist? It would be useful to show data from all sensors in each time series, like the authors have done in Fig 9. The same applies to Taylor’s diagrams: are these metrics from just one sensor system? Or an average of the performance of the many systems tested and/or deployed?

Lines 466 and 469: surely these two values of RH (80 % and 75 %) are meant to be the same?

Supplementary Material

Line 2: Surely the title should read “Evaluating low-cost NDIR CO₂ **sensors** for atmospheric observations in rural settings”?