

Atmos. Meas. Tech. Discuss., referee comment RC1
<https://doi.org/10.5194/amt-2021-372-RC1>, 2022
© Author(s) 2022. This work is distributed under
the Creative Commons Attribution 4.0 License.

Comment on amt-2021-372

Anonymous Referee #1

Referee comment on "Performance characterization of low-cost air quality sensors for off-grid deployment in rural Malawi" by Ashley S. Bittner et al., Atmos. Meas. Tech. Discuss., <https://doi.org/10.5194/amt-2021-372-RC1>, 2022

This article discusses the calibration via collocation experiments and five different calibration models, deployment, stability, and relocation/post deployment check of low-cost air quality sensors (specifically ARISENSE measuring PM, O₃, NO_x, and CO) in Malawi. Low cost sensors are increasing in use, specifically in areas with less infrastructure or resources for reference air quality monitoring stations, and long-term deployments are useful to see how they perform. Comparing sensor calibration robustness over time is also useful. I recommend eventual publication. However, there are a few large issues with this paper that must be addressed first. General comments first, then specific comments.

General Comments

This paper feels like an add-on/afterthought to another paper (on the air quality findings by the low-cost sensors in Malawi, which is mentioned by the authors as a separate publication in prep). That's an efficient data use strategy, but means that this study was not designed to optimize its stated main goal of sensor calibration scheme comparison and robustness. There are no reference monitors available for comparison during deployment, so the ability of the scientists to really understand how their calibrations performed over time and how performance changed when exposed to different conditions, beyond knowing when the sensors are returning non-physical data (e.g., negative values), is unclear. A comparison of sensor data to satellite data is done; however, the authors themselves acknowledge past work that shows that, in Africa, comparison of satellite to ground data is non-ideal. There are mentions of 'known emissions' near each site, but no clear definition and numbering of these emissions or discussion of their distance/location from the site. If the goal of the study was really to compare long-term performances of these models, this study was not ideal for that. Perhaps this could still be done in part—could you also pretend the initial collocation was a full experiment, .

I acknowledge that this is hard in this particular region; however, another major issue is that there is no one from the region involved as an author on this study. Involving local scientists would have greatly increased the scientific merit of this paper in a few ways. 1. Better understanding of conditions on the ground and local context 2. Better data capture by regular maintenance 3. Better understanding of how these sensors might be used 4. Better understanding of how well correction factors on data might be applied. (e.g., are the correction factors easy to apply with limited computing power and limited software?). Understanding the actual use, by people in the region, of these sensors could have made up for some of the lack of reference data comparisons by discussing another essential facet of low-cost sensor use (the people using them).

Specific comments

Table 1: define QR in table caption

How reasonable is developing individual models for each sensor and each gas component?

60 minute averaged data: will miss 'events' potentially like cooking/agricultural and trash burning—can you speak to the significance of these types of events?

Figure 1: I found this confusing. And hard to tell if the different models worked better or worse on the different sensors, or there was consistent agreement from looking. Are they the same graph just color-coded differently? If so you could make sure no overlaps in color on the two graphs. I think you are trying to convey a bit too much information on one graph for ease of interpretation. A table, or clearer labeling with different colors, or an extra panel, might work better to convey your message, which is a bit lost right now. I think a table (like those in the supplementary nfo) or adding some numbers of relative performance in your text could make the determination of which model worked best clearer for the reader.

For the atmospheric pressure versus sensor performance issues: are there any studies you could quickly compare your work to? Maybe some done in Boulder, for extrapolation? If not, I don't think that limits this work at all, as the difference in elevation is not huge.

Figure 2: These plots are somewhat confusing and hard to interpret, in my opinion. I would at least add a color scale bar. I think there's likely an easier way to display this info. Also I searched your text and I don't see V1,V2, Uni defined anywhere? I assume it is village 1 village 2 university, but I would repeat that info in the graph caption and maybe

in the text when introducing the sites define that V1=village 1 so readers just browsing the graphs can figure out the meaning more quickly. f

Line 279 : would recommend you to define more what you mean by 'well controlled'

Line 288 can you expound a bit on the 'difference in ozone precursor regimes?' I know a separate study is likely coming out about the measurements themselves, and this will be expounded on, but this line as it is just is hanging there begging for a bit more info.

Line 295 discussions: is calculating a diurnal trend really a way to see if the models are transferable, if you don't have any nearby ground-based air quality data? How did you get your local knowledge of air quality? (anything to cite?)

Line 319: sensor has less RH/T interference at higher concentration—can you provide a reference for this? It makes some logical sense but would be nice to expand this thought. Also, are there any chemical species co-emitted with CO that, at higher levels of CO, would also influence CO? How did you characterize the known CO sources?

The grid averaged for CO surface concentration is 12,000 km², right?. How many different CO sources are within an area of that size that would cause heterogenous CO measurements?

Figure 4: Label individual village sites, be consistent with naming of your sites.

Line 651: I assume the 'relatively short' and '100 hours' is 100 hours for the whole year? Or is that per biomass burning episode?

Also, are you certain that O₃, NO, and NO₂ came from 'fresh' biomass burning emissions, and not other sources (e.g., NO_x from diesel trucks with poor emission controls)? O₃ from biomass burning not so straightforward and a result of aging of biomass burning emissions.

Would be helpful to have a map of the stations & a figure showing when sensors were collocated with and when the instruments weren't collecting data. There are some maps in the supplement, but a nice and simplified summary graphic in the main paper would be good. If possible, a map with direction of known emitters could help (would assist with the discussion in change in sensor performance with wind direction). As it is, emission sources are mentioned but are unclear where and what they are, and how information that they

even existed was obtained.

The conclusions could be tightened up. What are the absolute main points your study, in particular, found? The novel work here, as stated at least in the introduction and abstract, is the comparison of different correction models over time. This gets a bit lost in the weeds.