Atmos. Meas. Tech. Discuss.,
doi:10.5194/amt-2020-129-RC2, 2020

Atmospheric Measurement Techniques Discussions

Interactive comment on "Robust statistical calibration and characterization of portable low-cost air quality monitoring sensors to quantify real-time O_3 and NO_2 concentrations in diverse environments" *by* Ravi Sahu et al.

Anonymous Referee #3

Received and published: 31 July 2020

General Comments:

This paper provides a comprehensive investigation of several factors affecting the calibration of low-cost sensors across different seasons and environments. While some aspects of the presentations of the results could be improved to enhance clarity, the basic approach appears sound. Most of my suggestions are related to specific changes that should be made, as outlined under the "specific comments" section below. In addition to these, I have some general comments, which should be taken more as suggestions of possible topics for further investigation or discussion rather than recom-

Printer-friendly version

Discussion paper

mendations of anything that must be improved in this paper.

It was mentioned in Section 3.2 that 52% of timestamps were invalid. Does this mean that more than half of the collected data was corrupted in some way, or simply that not all sensor were operating at all times? If the former, this would be concerning, as it means that less than half of the data during a long-term deployment might be useable, which would throw into question the validity of, for example, long-term average measurements. While your investigation of the effects of "small" datasets indicates this might not be too detrimental to overall performance, if such missing data were clustered in time rather than truly random, this might introduce biases into the dataset.

With respect to the "Aggregated Datasets", while it is possible to identify the effects of seasonally diverse datasets and of seasonally and spatially diverse data, investigating the effects of spatially diverse data only is not possible with the current setup (since the datasets which are spatially diverse feature data from different seasons as well). Have you investigated the creation of a dataset which is diverse in terms of the location but not in season (e.g. combining DD1(Jun) and MM5(Jun))? Of course, this would introduce the effects of sensor-to-sensor variability into the results, but that could also be investigated by comparing how, for example, a calibration on DD1 could apply to DD3. While you observe that parametric and non-parametric methods both suffer from poorer performance when generalizing beyond their calibration site and season, it would be interesting to see if parametric and non-parametric methods suffered to different degrees. Based on my experience, parametric methods may perform more poorly that non-parametric methods on data drawn from the same distribution as the training data, but generalize better than non-parametric methods when the underlying distribution in changed, i.e. by moving to a new site or climate.

Specific Comments:

Line 4: I don't know what "commodity" refers to here.

Lines 4-5: I would rather say that low-cost sensors can possess high precision (i.e.

high consistency between measurements), but require calibration to attain accuracy (i.e. similarity of low-cost measurements to those of regulatory instruments).

Lines 14-15: The “2” in “R2” is not appearing as a superscript (this seems to be true throughout the document).

Line 15: “upto” should be “up to”.

Line 52: Again, I am uncertain about what “commodity” means in this context.

Line 94: The word “include” appears redundant here.

Lines 94-96: The “2.5” of “PM2.5” should be subscripted. This occurs several times throughout the document.

Line 116: Again, I am uncertain about what “commodity” means in this context.

Line 118: The “2.5” of “PM2.5” should be subscripted.

Line 146: “sensor” is repeated.

Figure 3: It is mentioned elsewhere that four sensors are swapped between the two deployment locations, yet only three are listed in the figure.

Line 182: It is mentioned that there are seven sensors, but only six are listed.

Lines 184-186: Again, only three of the mentioned four sensors are listed.

Line 188: I believe “commodity” means “commercially available” or “commercially purchased” in this context. I would assume all other references to “commodity” have similar meanings as well.

Line 196: It should be specified if the T and RH values come from the internal sensors within the LCAQ sensors, or if these are from the reference instrument.

Figure 6: This figure could be clarified significantly. First, for the upper row, the pollutant in question should be specified on each plot. Also, since these plots refer to val-

ues measured by the reference instrument, specifying the LCAQ sensor data set from which the measures are plotted is unnecessary and potentially misleading. Instead, for example, sub-figure a could be titled “NO2 at site D”, and the two colors could indicate simply “Jun” or “Oct”. For the second row, I would also recommend simplifying the labels and switching data around so that each figure compares the two sites for the same season; as it is currently presented, differences are shown for both season and site, which makes it harder to separate these two effects. So, for example, sub-figure e could present “no2diff in Jun” with the two colors representing sites “D” and “M”. If sensor-to-sensor differences are a concern, you could use the average or median of values across all sensors deployed to a common site. However, I think an even easier approach would be to instead continue to use the reference data; since this figure is mainly serving to show how concentrations vary by site and by season, using only reference data (combined in different ways depending on what comparisons are being made) would be a valid approach. If you still want to include a comparison of the raw signal differences, I would recommend moving that to the supplemental information; I think trying to do that in one figure which properly accounts for all three sources of variability (season-to-season differences, site-to-site differences, and sensor-to-sensor differences) would be too complicated for a clear main-paper result. Finally, the dotted curves in the figures are not described. They appear to represent Gaussian distributions fit to the data, but since that is not an appropriate distribution for these data (e.g. they are strictly non-negative for the top row), I would recommend omitting these as they can be potentially misleading.

Line 385: What does “statistically distributed” mean?

Line 391: The figure does not appear to have 3rd or 4th rows.

Figure 7: I recommend moving this to the supplemental information; Violin plots are sufficiently common that they do not need to be described in such detail in the main text. However, in the case of the split violin plot, you should specify whether the depicted median and ranges refer to what is plotted on the left or right halves of the plot (or, are

two such depictions provided?).

Figure 9: Whether the figure refers to NO2 or O3 should be clearly denoted in the figures themselves, rather than just in the caption. This will minimize potential confusion and misunderstandings.

Section 6.4: This section appears incomplete. It does not refer to any specific results or figures. It is possible that sections 6.5 through 6.8 were meant to be sub-sections of this section.

Section 6.7: Although it has been mentioned elsewhere, I would recommend restating in the main body of the text that you observe that the effect is more severe for NO2 data than for O3 data.

Interactive comment on Atmos. Meas. Tech. Discuss., doi:10.5194/amt-2020-129, 2020.

Printer-friendly version

Discussion paper