Atmospheric Measurement Techniques Discussions

Atmos. Meas. Tech. Discuss., referee comment RC2
https://doi.org/10.5194/amt-2020-489-RC2, 2021

Comment on amt-2020-489

Anonymous Referee #4

Referee comment on "Unravelling a black box: an open-source methodology for the field calibration of small air quality sensors" by Seán Schmitz et al., Atmos. Meas. Tech. Discuss., https://doi.org/10.5194/amt-2020-489-RC2, 2021

General comments. The manuscript describes an open source, systematic methodology to calibrate low-cost sensors (LCSs). The Authors propose a 7-step statistical method based on: 1) preliminary analysis of raw data; 2) data cleaning; 3) flag data; 4) selection of the model by using both multiple linear regression and random forest and several statistical parameters; 5) model validation; 6) export of the experimental data as concentrations; 7) error predictions. Finally, the Authors tested the proposed model with an example during a field campaign in urban environment.

The manuscript shows a very interesting and systematic methodology to calibrate LCSs, suggesting to employ a univocal and standardized method to let comparable the LCSs measurements, considering the more and more frequently use of this technology. Despite this, the manuscript requires revisions before to be accepted for final publication. Following suggestions and specific comments.

Specific comments.

For this calibration procedure, reference instruments are needed. Trends due to specific events (i.e., burning etc…) could be not properly described by the sensors, if not calibrated in the same conditions?

Moreover, did the Authors try to do a calibration procedure by using chemical standard to produce a calibration curve at different concentrations and conditions in laboratory experiment? If yes, could the Authors discuss difference between the two approaches?

In case of LCSs time drift, did the proposed methodologies take into account of it (or allow to)? Is the proposed frequency (2 weeks every 2-3 months) enough to take into account of seasonal variations and, eventually, time drift of the sensors?

Lines 430 and following. LCSs could have a T and RH dependency. Is it appropriate to apply the suggested corrections by the manufacturer for T and RH to the raw data before to apply the models in the proposed methodology? Could the Authors discuss this aspect and if the models relationships with RH and T are in line with what, eventually, suggested by the manufacturer?

Other species could interfere with measurements: if the concentration of those compounds changes (season, night-day, etc…), this could affect the sensors response. How the Author suggest to deal with this eventuality?

The Authors showed their experiment results only for NO2 and O3 MOS sensors. Is this methodology applicable to other compounds (i.e., VOCs) or technologies (i.e., PID, electrochemical) with the same characteristics proposed in the manuscript? This information should be included in the manuscript.

The Authors refer to Ammonia and Reducing gas sensor in the manuscript (see Table 1), but results regarding these sensors are not present. Is this due to lacking of reference instrument?

Could the Author describe in more details how the Step 4 is performed? How the experiment and co-location data have been used in this step? The Authors describe that in Step 5 and 6 the co-locations data were used, but information about data used in Step 4 seem missed.

Did the Authors intercompare between them similar sensors, i.e. two Zephyrs, before and after the calibration to check the response of same sensors in same conditions?

About the data cleaning, how the Authors correct data for possible bias effect? Line 190: the duration of the moving window chosen to remove the outliers avoid to exclude from the dataset some specific and real events with short duration?

Lines 218-220. To identify which model better describe the measurements in term of over or under estimation, could the Authors consider to include also a statistical parameter such as the Fractional Bias?

Lines 365-369. Co-location 3 was at the end of the summer campaign (i.e., October). Anyway data for experiment 2 are not available. It seems from Figure 4 there is a seasonal impact. Did the Author use this co-location for their calibration for Experiment 1? Did the different season affect the calibration procedure? Are the 2 weeks every 2-3 months enough to take into account of it?

Lines 390-396. Is the GSM the only way to transfer data to a database? The warm up time was provided by the manufacturer?

Lines 421-423. Since the 3rd co-location is in October, could this be indication that closer and more frequent co-location are needed? See also the following Section 3.5 (line 539) and Figures 14-15.

Figures 14-15. Could the Authors add the 1:1 lines and indicate the R2 in the plots? How the Authors can explain the constant thresholds in the plots of panels 15e and 15f? Looking at Figure 11, the models using internal T and RH seem to give lower O3 and NO2 compare to the ones that use the ambient T and RH. In figure 11 this is less evident: could the Authors explain it and the reasons/meaning of the slopes (typically lower than the unit) and intercepts?

Supplementary. Why for the winter campaign, the Authors use co-location 1 and 2 instead of 4 and 5, which are closer to Experiment 3? Comparing Table 9 and Table S4, the models identified for O3 are different (and similarly for NO2): how the Authors could explain this?

Technical comments.

Line 88. See "host".

Line 200. Do the Authors refer to Section 3.5?

Line 291. Decent and good agreement should have to be quantitative and not qualitative information.

Lines 306-307 and 317-318. Information about the date of the campaigns are confused and should be coherent. The information could be furnished only once clearly and I would suggest to add the dates in Figure 2, as well.

Line 328-329. This sentence should be clarified. Zephyr s71 and s72 were located as in Figure 3 or with reference in an office on the 6th floor? In the former case, this information is redundant and could be included in previous paragraphs, when describing the setting (line 311 and following). In the second: how air masses have been sampled?

Line 359. When the Authors refer to "combined" co-location, this means an average of co-location 1 and 2?

Lines 419-423. This section is not well described. Could the Authors explain in more details the criteria to be used to flag the data?

Lines 430. Could the Authors specify in this or previous paragraph the units of the input data?

Lines 436-438. The Author report that relationship between Oxa and O3 was determined be inverse; but, since the predictive accuracy for no transformation is similar, they selected the latter. Anyway, in Table 3 there is not inverse relationship and a log dependency between O3 and Oxa was selected (also in Table 5 there is not inverse transformation). Could the Author explain this discrepancy or illustrate better this paragraph?

Lines 490-498. A comparison with the reference O3 and NO2 data should be included here (and in Figure 11).

Line 595. See "this is should be".