Atmos. Meas. Tech. Discuss.,
doi:10.5194/amt-2019-299-RC1, 2019

Atmospheric Measurement Techniques Discussions

Interactive comment on "Field Calibration of Low-Cost Air Pollution Sensors" *by* Andres Gonzalez et al.

Anonymous Referee #1

Received and published: 6 September 2019

The authors have constructed a relatively compact, 'low cost', transportable air pollution monitoring system (which they call a Mobile Autonomous Air Quality Sensor box – MAAQSbox) that contains a suite of commercially-available AlphaSense B4 electrochemical sensors for CO, NO, NO2 and O3, and an AlphaSense OPC-N2 optical particle monitor for PM2.5 (and other PM size fractions). The MAAQSbox also includes thermal conditioning of the inlet sampling stream, which incorporates internal T and RH sensing, and wireless data transmission.

The purpose of this paper is to report the results of a field calibration of the above-listed air pollutant sensors in the MAAQSbox as derived from co-location of the box at a 'reference' Air Monitoring Station in a 'near-road' location in Minneapolis, USA. (The paper states that the MAAQSbox also contains sensors for SO2, CO2 and VOC, but

Printer-friendly version

Discussion paper

only calibration for the five pollutant sensors mentioned above is reported here.)

There is considerable motivation for developing lower-cost air pollution measurement instrumentation; namely, the large adverse health impacts caused by poor air quality and the consequent need to collect more spatially spread measurements in order to better characterise human exposure, to develop and evaluate models and, subsequently, to evaluate effectiveness of mitigation. Description of some of the above motivation is appropriately covered by the authors in the Introduction to their paper.

The authors have clearly given good thought to construction of their MAAQSbox, particularly the efforts to provide some T and RH conditioning of the incoming air stream in order to negate dependencies of sensor response to these ambient variables.

In contrast, however, the calibration work described here is extremely limited, and, in my view, not of sufficient level to support full publication of this work in AMT. In both the Introduction (Line 80) and the Methods (Line 180) it is stated that the aim of this work is to 'evaluate' the performance of the sensors in their MAAQSbox, yet the design of their study and the data presented only reports a calibration (and not even the full calibration data), not an evaluation.

The co-location dataset comprises a single 6.5-day (154 hour) co-location at one time of year at one site. This dataset is then used to derive a multivariate linear regression calibration for each sensor against its relevant 'reference instrument' value using sensor signals and the internal airstream T and RH values as dependent variables. For calibration of the NO2 and O3 sensors, sensor signals from the other species sensor were also included to allow for potential cross-species interference. However, the authors do not present the actual calibration equation coefficient values and their p values (they only state which variables are included in each sensor calibration equation). Nor do they present visualisations and/or statistics for the raw comparisons of sensor values against respective reference concentrations. Consequently, in the absence of such information the reader is not able to gauge how well or not each sensor performs

prior to the multivariate regression fits, i.e. to gauge how much modification to raw sensor output is being made by the derived multiple regression calibration equation. In other words the reader does not get a sense of how much the sensor signal needs to be corrected for the contribution of other variables to the signal, particularly the extent to which there has to be correction for cross-interference between the NO2 and O3 sensors. Such information would tell the reader how important other variables are.

A more fundamental flaw, however, is that there is no independent evaluation of the calibration: the same data is used both to derive a calibration equation and then to justify the goodness of the calibration once applied to that data. If one derives a predictor equation from a dataset and then applies the predictor equation to exactly the same dataset then of course the predictions (and their 'evaluation' statistics) are likely to be very good. At the very least, there needs to be sufficient co-location data to (randomly) split into 'training' and 'test' sub-datasets in order to provide some (quasi)independent statistical evaluation of a derived calibration. More usefully still, what potential users of this MAAQSbox need to know is how well does a calibration equation hold in time and at different locations. Is there evidence of any long-term drift in sensor performance/calibration? If the sensors in the MAAQbox is calibrated at one location, does the same calibration hold at another location and/or at another time? If the MAAQbox is calibrated prior to a mobile deployment and is then used as intended on a mobile platform how well does its calibration hold up when the MAAQBox is co-located back at the reference monitoring station? All that the data presented in this paper show is that an underlying relationship for sensor performance self-consistently holds within a single 154 hour period.

Some additional comments:

The regression equation written on the panel of Figure 6c does not seem correct. The intercept appears to be much larger than 0.29 ppb, and eyeballing this panel suggests that the plotted regression line is giving much higher values for estimated NO2 than the stated regression line would predict; for example, for a reference value of 15 ppb

the regression equation predicts a sensor value of 12.14 ppb but the plotted line shows higher estimated NO2 than this. Also, this panel should include the origin of the scatter plot.

The scatter plot in panel 6a should also include the origin of the plot, and why does the regression equation for this panel not have an intercept coefficient? Even if the coefficient is not statistically significant its value should be included to indicate that the regression included an intercept in the fit.

The increased scatter in the calibration scatter plot for NO2 (Figure 6c) is noted but there is no discussion of this. Given that NO2 is a key pollutant in the urban environment, for which quantification by instruments such as MAAQSbox is most keenly sought, there needs to be further comment on what is underlying this poorer performance for NO2 measurement. As indicated above, we are not given the magnitudes or p values of coefficients in the calibration equations: which one of the variables is having the most influence on the NO2 response during this deployment?

In places the paper contains too much description of material that the authors can assume readers of this paper will be aware of. For example, it is not necessary to provide so much description of what statistical metrics such as R2, RMSE and p values mean. There is also quite a lot of generic description of operation of an electrochemical gas sensor.

Interactive comment on Atmos. Meas. Tech. Discuss., doi:10.5194/amt-2019-299, 2019.

Printer-friendly version

Discussion paper