Atmospheric Measurement Techniques Discussions

Atmos. Meas. Tech. Discuss., referee comment RC1
https://doi.org/10.5194/amt-2021-282-RC1, 2021

Comment on amt-2021-282

Anonymous Referee #1

Referee comment on "Machine learning techniques to improve the field performance of low-cost air quality sensors" by Tony Bush et al., Atmos. Meas. Tech. Discuss., https://doi.org/10.5194/amt-2021-282-RC1, 2021

General Comments:

In this work, the authors developed a machine learning calibration process that combines a 4-stage baseline offset correction and Random Forest Regression Modelling (RF). They adjusted the RF model by identifying readily available training features and optimizing the number of leaf nodes and trees. This work compared the performance of the RF correction model against values from a reference monitor, the raw sensor value, and baseline-corrected sensor values over a time span of ~7 months. This baseline + RF model improved the performance of low-cost NO_2, PM_{10}, and $PM_{2.5}$ sensors relative to the raw and baseline-corrected values. This machine learning technique is a reasonable method to improve data quality from low-cost air sensors and is suitable for publication after minor revisions.

Major:

Alphasense NO2-A43F electrochemical NO2 sensors (and Alphasense NO2-B43F) have a known cross-sensitivity to ozone (Spinelle et al.). Although the Praxis Urban sensor system and the St Ebbe's monitoring site do not appear to measure ozone, the study fails to mention/address this concern. While inclusion of this variable into feature training could restrict the spread of this model to other networks, it could greatly enhance the performance of the NO_2 model. Spinelle et al. also found that sensors from the same manufacturer can behave differently in the same environmental conditions. This manuscript would greatly benefit from applying your model to more than one sensor to demonstrate its capability to nullify discrepancies from sensor to sensor. (Spinelle, L.; Gerboles, M.; Kotsev, A.; Signorini, M. Evaluation of Low-Cost Sensors for Air Pollution Monitoring: Effect of Gaseous Interfering Compounds and Meteorlogical Conditions; Publications Office of the European Union:Luxemborg, 2017. https://doi.org/10.2760/548327)

It is unclear how this model could be applied to sensors throughout a network. Would each sensor need to spend *x* number of months at a reference site to develop the model prior to deployment? How well would a baseline established at the reference site transfer to the deployment site?

Line 163: "The filtering criteria presented in Table 1 were identified empirically from an analysis of typical sensor performance from the sensor network and from similar parameters logged at the St Ebbe's AURN station" It is not fully clear how these criteria were chosen. Was this based on limits set by the sensor manufacturer? Please clarify. It would also be useful to state the sample population percentage that was removed based on these criteria, as you did on line 188.

Minor:

Line 69: "multiple linear regression (MLR) models have been successfully used with variable results" Conflicting statement, please clarify.

Line 136: Please provide more information regarding the location of the sensor relative to the reference instrumentation.

Table 4 & Table 5: Please re-format the column headers as it is currently difficult to differentiate between them.

Line 319: "The performance of each component of the correction method is presented in Table 3" Should read Table 4 I believe. All table references after this point in the manuscript need to be shifted +1 up to Table7.

Line 392: "December 2020 saw the occurrence of several pollution events in the particle sensor time series (as also noted above). Although these events were observed throughout Oxford in multiple particle sensor time series, they were not reciprocated in reference measurements, nor in NO_2 data" It seems that around 12/25 in Figs 12-14 all corrected sensor values for NO_2, PM_{10}, & $PM_{2.5}$ experience an increase relative to the reference value. Therefore, it does seem like some event affected all three pollutant models. Have you investigated these anomalies further to locate a common factor?