

Interactive comment on “Evaluating and Improving the Reliability of Gas-Phase Sensor System Calibrations Across New Locations for Ambient Measurements and Personal Exposure Monitoring” by Sharad Vikram et al.

Christine Braban (Referee)

chri2@ceh.ac.uk

Received and published: 17 March 2019

General overview:

This paper is a well thought through experiment and has some exciting ideas about the building of sensor networks and data processing to improve or understand error and bias in the systems. It uses a coherent approach, and develops a new statistical method which is mostly well described and accessible to the atmospheric scientist reader. There are a few major areas for improvements which are suggested below.

Printer-friendly version

Discussion paper



Major comments: 1) Presentation of measured data: Measurement data: despite doing a great job in static analysis of sensor data, this paper lacks a figure with the EPA and low cost sensor measurement data, ideally 1 panel with initial error envelope and 1 with final error envelope. For example NO₂ sensor measurement (+MAE+95th percentile MAE) (-MAE-95th percentile MAE) (i.e. from Figure 9 level2) - the authors may have better suggestions but if the paper and sensor systems data are to be used by scientists, citizens and community groups this is the information which needs to be presented. 2) Use of reference station data Generally the authors seem to treat the data from the reference station as not existing outside of the experiment and repeatedly make comments about the results in this paper showing new information about pollutant variation at the test sites. This shows a lack of forethought and analysis of existing data. The chemical climate for NO₂ and ozone at the sites is measured by the reference station, presumably for some years. The information to calculate the variability therefore is already collected and reported somewhere (e.g. EPA or San Diego authority reports). The authors should use existing knowledge to inform their data analysis and interpretation rather than present the data in a knowledge vacuum. For example the speculation about the chemical climatology (my phrase not used in manuscript) of Shafer shows a clear disregard for existing evidence outside their experiment which is really avoidable.

3) Terminology: The error terminology is not applied consistently through the paper. "Error" is used a lot without it being clear whether it is overall error, expanded uncertainty, a replacement for Mean Absolute Error. Terminology is only introduced on p14 in the results section and then it is only sparsely used after that with terms such as centred error and bias error being added in (p15 line 6) This lack of clarity in the terminology leads to more questions than are answered in the manuscript e.g. "How do the authors assess the total error with the set of benchmarks" "When the difference in error is discussed (e.g. p14 line 11), which difference is being discussed?" "In the discussion the authors do not specify which error they are referring to e.g. p16 line 6 onwards. "Difference in error is discussed in qualitative terms despite there being

[Printer-friendly version](#)[Discussion paper](#)

quantitative data in the work. Significance of differences are not discussed.

Overall the authors should revise with a consistent terminology and perhaps a table glossary. Also only MAE is shown in the main paper (Figure 6), please could the authors add the equivalent plots in the same figure for each of the benchmark errors. It is hard to assimilate the many tables in the Appendix with the results. The tables potentially should be moved into the main manuscript.

4) Discussion of cost Interestingly I think there are 2 points from this paper: the new N-N splitting method offers improvements for incrementally built networks and the Simple linear algorithmic gives the best understanding of changes. The former involves large amounts of expensive model development and potentially lack of clarity as to why the model works to improve sensor data (to a non-scientist sensor user). The cost implications for what is now not low cost at all (enclosures, powered fans, telemetry, expert algorithm development and maintenance) seem to be not openly explored, rather a nebulous “small cost increment for new nodes” offered as a positive: Could the authors perhaps discuss whether there is unconscious bias in their cost increment assessment?

Minor comments and corrections:

Abstract Could do with quantitative analysis of results in abstract e.g. how much does N-N improve model?

Introduction

Well written and readable. P4 lines 20-35: the list of the results does not fit in the introduction

Methods Section 2.1 sampling sites: the authors describe the sampling sites “expected profile” in descriptive terms. Given that the sites are regulatory local environmental and emission metadata for the site is probably available in the EPA station records. No references for the regulatory station information or data is provided or links to EPA reports using site data.

P6 line 13: “over the air” what does that mean?

P6 line 34: the noise on the signals and SD are discussed of the sensor data. It would be useful to have the statistics of the raw (level 0) data and the “cleaned” or level 1 data for each sensor deployment as per epa site format used in Appendix B- or in a table in the main paper as it is critical for understanding the data processing effects

P7 line 6 is the sensiron accurate to 0.05% rh or is it just the data resolution on readout. Please could the accuracy/precision be stated. Resolution is not really useful.

Section 2.2.1 and 2.3 : the passive electrochemical samplers are placed in an actively ventilated housing “for this study”. If truly for only this study and the sensors would be deployed differently under a normal operation, how are the results relevant to different setups?

2.3 Data is stored in the cloud: are they available for the public? What is the archive for data (and data identifier)?

P9 line 9: was the data from the reference station provisional or final ratified, i.e. regulatory automatic network data is ratified on a cycle. What was the date of data provision and data capture statistics for those periods? It is not enough to just say they came from the EPA. Data should be referenced properly.

Results: p14 line 15 is the difference between level 2 error vs level 1 statistically significant? It looks quite small on the Figure.

P9: averaging of minute data: arithmetic mean, time weighted average? How are data gaps in a minute treated?

Data filtering: filtering for “the realm of reasonable values” probably needs explaining more completely. Please list the QA filter steps in the appendix. Just to note, short lived plumes do not give reasonable values when you are normally used to looking at average values, but they may be real and relevant. What does +5V represent for each parameter? Probably the filtering is fine, but from the paper I cannot tell that.

[Printer-friendly version](#)[Discussion paper](#)

P10 lines 1-5: you can tell that that would be overlap from the reference site data. Not necessary to confirm it with sensor measurements.

P15: It is interesting so see the change in bias between level 1 and 2 and I feel it should warrant further discussion in the manuscript. How many extra levels would be need to achieve an acceptable bias?

Figure 7: this figure needs a more complete caption to describe the graphs

P16 line 10: which error metric?

Figure 8: no x-axis label

Figure 9: please match scales on the two NO₂ graphs and the two O₃ graphs for ease of comparison. Also putting zero or an integer at the origin would be good practice.

P20 line 16. The authors mention using the 5s data to get more information and improve data quality despite the response time of the system likely to be not 5s (not quantified in this paper) and no reference given to show that this would a likely significant improvement rather than addition of more noise. It would be useful if the authors expanded on why they are optimistic about this.

Fig A1-A3: Figure captions could be more explanatory. What is the line vs the bars?

Appendix B too many decimal places in the tables!

Interactive comment on Atmos. Meas. Tech. Discuss., doi:10.5194/amt-2019-30, 2019.

[Printer-friendly version](#)[Discussion paper](#)