Atmos. Meas. Tech. Discuss.,
doi:10.5194/amt-2019-30-AC1, 2019

Atmospheric Measurement Techniques
Discussions
Open Access

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.

Sharad Vikram et al.

ashley.collier@colorado.edu

Received and published: 18 May 2019

Comment: "This manuscript describes the assessment of several approaches that could be used to improve both the performance and the transferability of low cost gas phase sensor system calibrations. This is a crucial step in the enabling of these technologies for use air pollution monitoring, and this work is a valuable contribution to the growing body of literature on this major remaining challenge for these technologies. Previous work has demonstrated that although successful calibrations can be derived for low cost sensors through co-location with reference grade instruments, these cali-

Printer-friendly version

Discussion paper

brations do not hold if the sensors are moved to a new location, or even at the same location under significantly different chemical or meteorological conditions, and are prone to model over-fitting. The lack of a robust and transferrable calibration strategy is most likely due to variations in the multiple environmental parameters, both chemical and physical, that effect sensor signals. The authors of this work propose that by using the data from multiple low cost sensors systems co-located with reference instruments in different locations the resultant calibration will be more generalized. This approach has been suggested previously, however, to this reviewers knowledge this is the most extensive investigation of this approach for gas phase electrochemical sensors to date. The authors also propose a novel two-stage "split-NN" approach to address the challenge of sensor to sensor variability when creating a global calibration. The analysis presented in this manuscript is thorough and well written, and although the generalized calibration models developed still maintain large sensor errors the methods do show promise. I therefore recommend publication after the following minor comments have been addressed."

Response: We thank reviewer 1 for their thoughtful and detailed comments. We believe we have completely addressed the reviewer's comments through revisions, as discussed below. We are grateful to reviewer 1 for their help in markedly improving the paper.

Minor Comments

Comment: "Sect. 2.3 pg 9 lines 13-15: It would be useful to the reader to know how much data was removed during the preprocessing steps."

Response: Thanks for noting this omission. We have added details about how much data was filtered to Section 2.4 Preprocessing, in particular, that 2.4% of the 5-second data was filtered.

Comment: "Sect. 2.5: The split-NN is a novel approach for correcting for sensor-to-sensor variability in sensor signal and response to target compound concentrations. If

I am not mistaken however, the environmental variables such as temperature are only used in the second stage of the process. As individual sensors are known to have different responses to their target compound it is more than likely that they will also differ in their responses to interfering compounds and environmental factors (this has been shown previously e.g. Smith et al. 2017). Would the authors not therefore get an improved result if the environmental parameters were included in both stages of the split-NN procedure? The authors should provide further justification of the variables chosen for each step in the split-NN."

Response: The environmental parameters contribute to training both stages through the process of backpropagation, as the sensor-specific model and the generic model are trained together. The concept is that, during training, the model-in-development generates a prediction, which determines an error with respect to the ground truth. Backpropagation pushes this error back into the network and the weights in the neurons in each layer are adjusted. As a consequence, even though the environmental parameters are injected downstream, their effects are felt upstream. We hesitated to provide this kind of detail in the article, leaving that for the cited literature on machine learning. However, if a more detailed account is desired, we'll be happy to accommodate.

Comment: "Fig. 6: Needs units on y-axis."

Response: Units have been added to the figure and the units for all metrics have been clarified in the text both where the metrics are introduced and in Appendix C.

Comment: "Fig. 7: Needs units on plot axes and the time averaging used for the data points needs to be stated in the fig. Caption."

Response: Units have been added to the figure. The details on minute-averaging have been elaborated in Section 2.4 Preprocessing, and apply to all the analyses, so we feel it is better centralized here. However, we believe that your comment was also directed at what each point in the target plot represents, so we have clarified that in the caption, saying that each point in the plot corresponds to a different individual benchmark (i.e., a unique round, location, and board).

Comment: “Sect. 3.1 pg 15 lines 7-8: The sentence “The increase in bias is more pronounced in the higher capacity models” does not seem to be strongly supported by the data presented in Fig. 7. This statement needs supporting quantitatively or removing.”

Response: True, nor do the results of the paper depend on this sentence. It has been removed.

Comment: “Sect. 3.2: It would be interesting to see the performance improvements from each stage of the split-NN approach. The addition of error plots similar to Fig. 7 for a single sensor after both stages of the process would help visualize the power of the approach.”

Response: Unfortunately, a Split-NN provides ppb predictions only in the final stage. The earlier stage provides a set of latent variables that are learned from the input variables.

Comment: “Fig. 9: Needs units on y-axis.”

Response: Units have been added to the figure.

Comment: “Discussion: The authors are open about the limited success of the transferable calibration approaches investigated. It would, however, be beneficial to the field if the authors were to expand further on possible reasons for this and potential ways to improve the methods moving forward.”

Response: In the Conclusion we mention one direction for future work, using the higher resolution data of our sensor. We have now added two others (improvements to split-NN and use of infrastructure data). In the Discussion section, we do discuss some of these issues, including taking a closer look at bias error (3rd paragraph). But as of now, we’re thinking of this as a pretty strong tradeoff between transferability and accu-

Printer-friendly version

Discussion paper

racy that can only be addressed through more diverse measurement (4th paragraph of Discussion). We still hold out some hope for split-NN as an economical approach to gaining more diverse measurement (5th paragraph).

References: Smith K. R., Edwards P. M., Evans M. J., Lee J. D., Shaw M. D., Squires F., Wilde S. and Lewis A. C.: Clustering approaches to improve the performance of low cost air pollution sensors. Faraday Discuss, 15, 1-15, 2017.

Response: Thank you. This work and several others were added in a related work paragraph in Section 2.6. Although the methods were trained similarly, it is important to note that in the intended use case, the MetaSense sensors would be generating a prediction about their current location, which would be wherever the end user happened to carry the sensor. It would be unlikely that they would be near other sensors.

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