Response to Reviewers

Dear Reviewers,

We are grateful for your comments and suggestions, which have helped us improve the manuscript. The necessary revisions have been implemented, which can be found in the attached file (highlighted in yellow). Below, we provide responses to your comments and suggestions, along with corresponding changes made in the revised manuscript, where applicable.

Sincerely,

On behalf of all authors,
Ayah Abu-Hani

Response to Reviewer #1

<u>General Comments</u>

The manuscript presents research work looking at the application of machine learning (ML) techniques in developing a transferable calibration methodology for a network of low-cost sensor units (SU) focusing on NO and NO2 pollutants. This study assessed the performance for collocated and non-collated networks of low-cost sensors in different urban environments in Switzerland and Italy. They incorporated several commonly used variables (raw sensor signals, RH, temperature) in their algorithm but emphasised the key role ozone plays as an input in their model, concluding that the best results were obtained in studies involving co-located networks which have ozone as part of the input variable.

Thank you for taking the time to review our manuscript. We do appreciate your feedback and have carefully considered your comments. We have improved the revised manuscript in response to your suggestions.

<u>Specific comments</u>

The authors have used low-cost SU that have a pair of electrochemical sensors for the two species of interest (NO, NO2). The reviewer found it odd that the pair of signals were used in the model setup for each species. For instance in modelling the corrected NO2, both the 'NO2_A' and 'NO2_B' are used but I would expect these two signals to be very correlated as summarised in Table 2. I would have thought one of the pairs should be sufficient, particularly the one with the best R value in Table 3.

The study by Bigi et al. (2018) demonstrated that the feature combinations with a pair of electrochemical sensors leads to better performance in calibration models compared to a single sensor. Given that we utilized the same dataset in our study, we maintained the same features for benchmarking purposes. Furthermore, Smith et al. (2019) reported the effectiveness of employing an array of sensors rather than a single sensor. They utilized the instantaneous median signal from six identical electrochemical sensors for NO_2 and O_3, resulting in minimized randomized drifts and inter-sensor differences, thus, addressing some limitations of individual sensors.

We have updated the corresponding statements in our revised manuscript (P.16, L.310) and (P.18, L.331).

While the reviewer agree with the authors on the inclusion of O3 for as input variable for the NO2 calibration (there are ample literature evidence for this), there are very little evidence for the same for NO, thus questioning if this could lead to overfitting/training dependence for NO on O3 and potentially result in additional error in transferability of this method to regions where O3 is high but low NOX.

Thank you for pointing this out. Yes, we agree that limited evidence for O_3's inclusion in NO calibration, so, we made sure to clarify this in our revised manuscript (P.4, L.125).

In this work, we've evaluated our method's transferability across diverse environmental conditions, including regions with high O_3 but low NO_x levels. For example, training a calibration model on LAU data and transfer to ZUE showed improved performance (as depicted in Fig. 9 Case B), where ZUE is characterized with higher O_3 and lower NO_x compared to LAU. This is also supported by the feature importance (Fig. 10), shows that in ZUE O_3 is among the most significant attributes.

Moreover, we've employed robust validation techniques to address concerns about overfitting and training dependence.

<u>Technical corrections</u>

Figures 1 & S2, the caption describes the central line of box plots to mean the median but the median are not shown in these figures.

The figures are now updated. We also included the mean values (indicated by "*") in the modified manuscript and Supplement. (Fig.1 & Fig.S2).

P.11, line 232 & 234, units missing for the RMSE values. Autor need to correct instance of this in the whole manuscript

Thank you. The manuscript has been modified accordingly (P.11, L.335 & L.337).

1. ***17, line 311-312: the statement "Moreover, this study advocated enhancing global calibration models by incorporating O_3 measurements from available nearby monitoring stations." This statement is too generic, it implies that O_3 needs to be considered for all low-cost SU network calibration like CO, PM, CO_2 and NO (see reviewers' general comment about this species above) etc.***

   The statement is now modified in the updated manuscript, and it is made specific for NO_2 and NO (P.17, L.318).

2. ***18 line 333, the statement "The utilization of multiple electrochemical cells within each SU targeting the same pollutant to enhance data reliability" needs to be revised in context of the reviewer's general comments above for: 1) NO species and 2) overfitting by incorporating pair sensor of reading for same species.***

   Corresponding statements have been modified to address these comments in the updated manuscript (P.18, L.331) and (P.4, L.125).

References

Bigi, A., Mueller, M., Grange, S. K., Ghermandi, G., and Hueglin, C.: Performance of NO, NO 2 low cost sensors and three calibration approaches within a real world application, Atmospheric Measurement Techniques, 11, 3717–3735, https://doi.org/10.5194/amt-11-3717- 2018, 2018.

Smith, K. R., Edwards, P. M., Ivatt, P. D., Lee, J. D., Squires, F., Dai, C., Peltier, R. E., Evans, M. J., Sun, Y., and Lewis, A. C.: An improved low-power measurement of ambient NO2 and O3 combining electrochemical sensor clusters and machine learning, Atmos. Meas. Tech., 12, 1325–1336, https://doi.org/10.5194/amt12-1325-2019, 2019.