

Reply to the comments of reviewer #3 on the manuscript „Design of an ozone and nitrogen dioxide sensor unit and its long-term operation within a sensor network in the city of Zurich“

Michael Müller¹, Jonas Meyer² and Christoph Hüglin¹

¹Empa, Swiss Federal Institute of Materials Science and Technology, Dübendorf, Switzerland.

²Decentlab GmbH, Dübendorf, Switzerland.

We thank the reviewer for his effort in evaluating this manuscript and for his valuable suggestions for improvements. All points made by the reviewer are addressed on the following pages.

The text isn't very clear however about how data from the redundant sensors is used in the paper's subsequent analysis. Is the information from all three NO₂ sensors included in the statistics on for example RMSE (mean, median?), or is one single sensor chosen, with the other two NO₂ sensors truly redundant spares in case of failure? The later part of the paper shows how the identical sensors in each box compare to one another and this is very interesting, but there needs to be better clarification of how each contributes to the datasets that are the main conclusions of the study. Page 10 refers to the mean of sensors in each box, but it is not clear if this approach is used through the paper, or just in this part.

Response: We agree with the reviewer that the reader should easily understand which part of the manuscript refers to measurements from individual sensors and in which parts the mean of all the sensors in a sensor unit has been used. This was not always clear.

Modification: We therefore added a paragraph at the end of section 3.1 that explains in which section individual sensors are discussed and in which sections the discussion refers to the mean. This information is repeated in section 5.1 as in section 5 discussions refer to both individual sensor measurements and the mean of the individual sensor measurements.

The paper includes some lab testing of sensor interferences, and this shows some similar results to other studies. Are the various mixtures of co-pollutants that are tested presented to the sensors in 'real' air (e.g. zero air, or synthetic air) or blended in pure N₂?

Response: NO, NO₂, CO, CO₂ were added in specific concentrations to conditioned zero air.

Modification: We added this information in the text.

Figure 2 is really far too small to be read clearly for so many different chemicals, so a better way of showing this data is needed.

Modification: We optimized Figure 2 in order to present our findings in a clearer way.

There is a lot of information in Figure, and the text refers to this figure as showing for example the impacts and temperature RH correction compared between models. It is quite hard for the reader to find this in a many different models, so could the figure or text more directly identify those models that show these differences?

Response: We assume the reviewer is referring to Figure 3. The figures are linked with equations 2 and 3 to 6 in terms of the variable naming. The discussion about the model performance is also based on these equations. We are aware that there is a lot of information in Figure 3, nevertheless Figure 3 is complete and provides all the information needed for the assessment of the dependence of the model performance on the selection of the different terms in the sensor models.

Modification: We suggest keeping Figure 3 as it is.

The comparisons between sensors in Figure 6 are interesting (associated text is on page 10). There is reference to drift in RMSE and R2 over time, but its quite hard to see this in the data. I presume this is inferred because towards the end of the period more sensors have RMSE value above 4 ppb, than at the start? Is there perhaps another way this could be shown graphically since this is an important point.

Response: The large differences between the mean differences and RMSE values between Figure 6 (a) and Figure 6 (b) can clearly be identified. The term “drift” is not entirely adequate. We agree with the reviewer.

Modification: We changed a sentence in section 5.2 in order to focus on the main message of Figure 6. In addition, we limited the time axis of the figure to August 2, 2016 as this is the end of the data period used in the sensor calibration.

The comparison of sensors in homogeneous air is an interesting approach. Can the figure and text be made a little clearer about which bits of data in Figure 8 are from Sensor Units, and which are data being reported from standard reference instruments. I have assumed that the sites in normal font, eg SWD, SCH, STA are only showing reference instrument data, for those selected periods, whilst WIN, STV, ETH for example are only showing sensor data?

Response: It is correct the site names in normal font refer to measurements from reference instruments and the site names in bold refer to sensor measurements.

Modification: We changed the caption of Figure 8 in order to improve the explanation of the figure content. We think it is much clearer now.

Figure 11 seemed to be a little surplus to requirements in the paper. It is good that the Sensor Units captured a plausible diurnal trend, but since the two sensor locations and two AQ reference stations are not co-located, there isn't much to infer from comparing the two types of data.

Response: Traffic causes most NO₂ emissions in Zurich. Diurnal variation of traffic volume in the city centre of Zurich is not homogeneous but comparable. Therefore, we think that locations impacted by similar traffic volume can qualitatively be compared. This plot is valuable in order to give the reader an impression of the obtained data quality which is encouraging (as indicated in this plot) but not yet what is required for long-term use (comparison with diffusion tubes).

Modifications: We therefore suggest keeping this Figure in the manuscript and not to remove it from the paper.

Minor editorial changes

Page 2 'metal oxide'

→ Corrected.

Page 2 Line 27, ' : : were operated at these locations until August 2 : : .

→ Corrected.

Page 5. Line 11. Presumably this should read something like: ' : : leading to the omission of a few measurements where there were small variations in measurement frequency'?

→ Changed.

Page 6 Line 4. It is temperature and humidity that interfere with the sensors, not the other way around.

→ Corrected.

Page 7. Line 7. This isn't clear, but I have assumed this to mean the sensors were operated some way away from the reference site. The explanation of the PAR/REM approach to calibration needs a slightly expanded and better description here. It becomes clearer the more you read on in latter pages.

→ We reworded section 4.1 in order to better explain the calibration concept.

P8 line 14 – obviously

→ Corrected.

Figure 6 (and elsewhere). Can the captioning use the same abbreviations as the text, eg DS1, DS3. There is some interchange between dataset 2 - DS 2 etc.

→ We made several changes in the text to be more coherent concerning naming of the data sets.