Anonymous Referee #2, 13 Mar 2023

AC: We thank Referee #2 for her/is time and effort to go over our manuscript and provide useful comments. Below are our responses to all the points s/he raised. The associated modifications in the revised manuscript are noted with red font colour.

RC2: You mention that you use the reference thermometers at the station to compute the different correction factors. Do you think that there can be a difference between the temperature recorded by the reference thermometers and the real temperature at which the sensor box is exposed? Can it impact the correction factors calculations or the final concentrations? Are you including temperature sensors in the sensor box?

AC: A temperature/RH sensor was included in the two AQ monitors, and there was a big difference between the measurements of these sensors compared to the respective reference ones, especially during high temperature days. The reason of using the ambient temperature (rather than the temperatures recorded by the sensors inside the air quality monitor) was that the sensing area of the sensors was exposed to the ambient environment. In fact, we explored that using the temperatures recorded by the sensors located inside the AQ monitors for converting the LCS signals to concentrations, but that resulted to much higher deviations from the reference measurements.

To address and clarify this point, we have updated the following lines from:

"*The temperatures used as input for the calculation of the correction factors n_T and k_T are the ones measured by the reference thermometers at the station.*"

to

"*While a temperature/RH sensor was included inside the two AQ monitors, the temperatures used as input for the calculation of the correction factors n_T and k_T are the ones measured by the reference thermometers at the station. We should note here that the temperature and RH sensor used in the two AQ monitors, recorded measurements that deviated substantially from those reported by the reference sensors, and thus were not used for determining the concentrations from the LCSs.*" (lines 173-176 in the updated version of the manuscript).

RC2: Table S1 shows some sort of compilation errors "Error! Reference source not found".

AC: We thank the reviewer for pointing this out. The error has now been corrected. This was literature citation software issue.

RC2: Many recent works use a machine learning-based in-situ calibration, comparing the LCS measurements with the reference instrument. Do you think that some sort of sensor recalibration where the sensor calibration model is trained with data with large temperatures may reduce the worsening of the sensor?

AC: Indeed, this is the case. In this work we focused only on the performance evaluation of low-cost electrochemical gas sensors at extreme temperature and RH conditions, using the calibration models provided by the manufacturer. In fact, we are currently exploring the effectiveness of machine learning algorithms for calibrating measurements from LCSs, including identifying the best calibration schemes one can follow in doing so. We deliberately did not include any results from that work in this manuscript, as it would deviate from its main focus. We plan to submit a follow-up paper on the results from the work we did on the use of machine learning algorithms very soon.

RC2: As noticed in Table I, different sensors have different numbers of useful samples (N), the difference can be up to 2000. Is it due to missing data, sensor issues?

AC: This point is well taken. The difference in the number of hourly averaged data points used for deriving the performance of each LCS is attributed to the following two reasons: (a) missing data points due to data acquisition systems malfunctions and/or overheating of the sensors electronics, and (b) omission of negative concentrations (i.e., measured by the LCSs). We should note that in the case of the Winsen LCSs, less measurements were recorded during summer in comparison to the Alphasense LCSs, because their operation was interrupted earlier than the Alphasense due to electronics overheating.

We believe that these points are clear in the manuscript. For instance, in Lines 304-306 that read:

"*It should be noted that even though the sensors were not operating during the summer months, they were exposed to ambient temperatures (up to 45 °C), while their main body containing the electrolyte was at even higher values (i.e., > 50 °C) during midday, as a result of heat built up within the cases of the AQ monitors*."

and in lines 352-355:

"*This was feasible for the datasets of the Alphasense sensors but not so easy for the Winsen, as those were not always functional after the summer period, yielding significantly less measurements during that period as indicated in Table 1 and reflected by the high temperature differences in the two periods analysed here (cf. Table 4).*"

RC2: In Figure 3, it can be seen how the O3 estimation for the Alphasense sensor completely overestimates the reference values. Does it make sense? As far as I understand this period corresponds to beginning of the testing and the temperature is not large is this case.

AC: Even though LCSs are proven to be highly influenced by high temperature and low RH conditions, the performance of Alphasense O_3 sensor is poor from the beginning of the measuring period. This is possibly due to the cross-sensitivity of O_3 sensor to NO_2, as mentioned in section 3.1 (lines 272-277 in the original manuscript; now lines 270 – 275 in the revised manuscript) that read:

"The high errors of the Alphasense O_3 measurement can be associated to interferences with NO_2. This is in fact taken into account, as the signal from the NO_2 LCS (mV) is subtracted from the O_3 sensor signal following the guidelines by the manufacturer (cf. first term in the nominator of Eq. (1)). Although this correction is used for the performance analysis of the Alphasense O_3 LCS, it yields a measurement error that is higher compared to the case when the correction is not applied (cf. Fig. S6), most likely due to error propagation from the signal of the NO_2 sensors to the reported concentrations of O_3"

We believe that point is addressed, and thus no further action is taken.

RC2: Have you evaluated/shown the possible joint effect of temperature and relative humidity? It would be interesting to show a result by diving the data into, low T+low RH, low T + high RH, etc.

AC: As stated in the manuscript, during the period of our study the low temperatures are typically accompanied with high RH values and vice versa (cf. Fig. S5 in the supplement and Figure 1 below). As shown by measurements (Figure 1 below), there are not many cases (data points) with low temperature and low RH (bottom left corner in the plot), or high temperature and high RH (top right corner) to proceed with the comparative analysis suggested by the reviewer. Analysis of the other

cases is already provided in the manuscript, so no further action was taken with respect to this point.

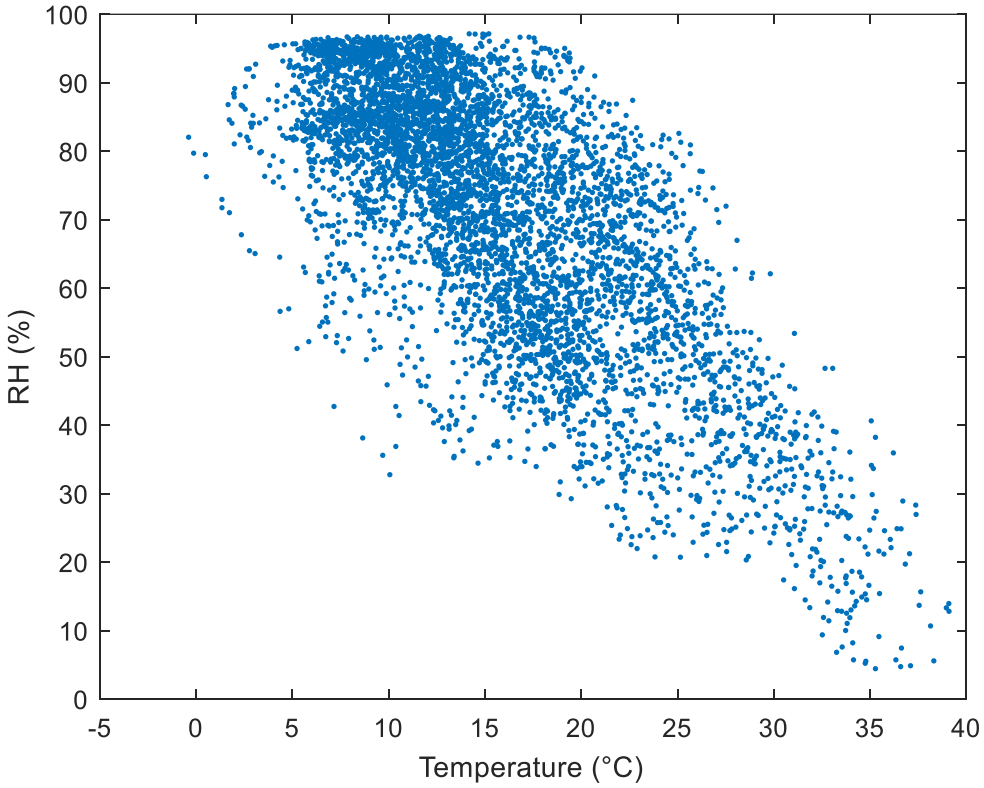

Figure 1: Correlation between the temperature and RH measurements recorded by the reference monitoring station for the whole measurement period.