Atmos. Meas. Tech. Discuss.,
doi:10.5194/amt-2017-43-RC2, 2017

Atmospheric Measurement Techniques Discussions

Interactive comment on "Practical field calibration of electrochemical NO_2 sensors for urban air quality applications" *by* Bas Mijling et al.

Anonymous Referee #2

Received and published: 24 June 2017

The paper attempts to describe a way to use low-cost electrochemical NO2 sensors (Alphasense) to extract meaningful information about ambient levels of NO2 in an urban area. Data were collected from several identical NO2 sensors (and 2 from a previous generation of Alphasense NO2 sensors) using a co-locate/measure/co-locate. Decoupling the interference between NO2 and O3 with Alphasense sensors is a difficult task, as highlighted throughout the literature. However, after reading through the manuscript several times, it does not appear the author's goal was accomplished based on their thesis: to describe a "practical method for in-field calibration and regression modeling" of electrochemical NO2 sensors. Several major concerns including the use of a reference instrument (ozone) as an independent variable within the model and lack of rigorous validation data must be addressed.

The best-performing model includes data from a reference ozone monitor which does not constitute a "practical method" for using low-cost NO2 sensors, and the regression modeling nearly completely describes how well these sensors performed in the past, without properly withholding validation data to describe how they will hold up in the future (predictive versus descriptive modeling). The modeling approach (multivariate linear regression using WE and AE) is not novel in the literature concerning Alphasense electrochemical sensors, especially when considering species other than NO2 (see Lewis 2015[1]) as an example that uses both linear regression and other statistical models).

In addition to a few major corrections, many minor corrections should be addressed as well (outlined below). Therefore, publication of this manuscript in AMT should only be considered after the comments below have been addressed.

Major Comments

P. 6, line 24: Including a reference ozone measurement as an independent variable in the linear model is inappropriate for low-cost sensing. If the goal is to describe a method by which you can use low-cost NO2 sensors to obtain a decent NO2 concentration, then including data from a $5000+ instrument in the analysis simply cannot be included. I understand that there is a strong cross-sensitivity to ozone, but claiming even a poor ozone measurement would improve results without any evidence to support the claim is invalid. This should be removed completely from the analysis.

To show the model is predictive (rather than descriptive), previously withheld validation data should be used to evaluate the model. Currently, this work only shows that these sensors can reasonably describe what has been measured in the past, but provides no insight into well they will hold up in the future.

All fit parameters in the tables (and throughout the paper) should have error estimates/confidence intervals.

A focus on the absolute RMSE, rather than just the bias-corrected RMSE should be highlighted in the abstract

Minor Comments

There are many English language errors (mostly grammar) that need to be worked out

P. 2, line 6: These sensors are commercial, not experimental, despite their quality. Stating otherwise supports the idea that they are not currently on the market, which they are.

P. 4, lines 3-4: Rather than just throwing away data based on arbitrary filters, a digital filter could be used. Throwing away data that is not within 10% of the mean is probably not the best methodology; one gives up the ability to measure higher concentrations if a local source were to emerge!

If the analysis is going to be based on the "more linear" regime of these sensors (dropping all data > 30C), it should be more pronounced in the abstract and introduction (page 4, lines 5-6). This is a huge limitation and one of the most important research topics for electrochemical sensors (as used for ambient measurements).

P. 6 line 10: If the DHT22 sensor does not need to be individually calibrated, the authors should explain why they observed such large variance between DHT22 sensors and how this affects their model results

P. 6, line 15: Comments suggest the sensor loses sensitivity at higher temperatures. This seems counterintuitive given that diffusion across the membrane should be faster at higher T. What is the explanation for this effect?

P. 11, lines 5-6: Diverging results for two different models of Alphasense NO2 sensors are discussed; Alphasense explains why the newer version of the sensor obtains better selectivity towards NO2 and has a reference (Hossain 2016 [2]) that should be examined/discussed.

Equations 6, 7: A time-based interpolation for back-calculation of NO2 is used without sufficient evidence the decay in sensitivity/accuracy is linear in time.

P. 9, line 22: Is there a reason the authors decide to use r2 rather than adjusted-r2 for comparing to adjusted-r2?

The median value throughout the campaign is 15 ugm-3 and the stated 95% CI is 14 ugm-3 (2*RMSE); what is signal and what is noise?

What makes a measurement "good enough" (page 10, line 15)?

Claiming the calibration period should be "as long as possible" isn't very helpful. Eventually, the sensitivity of the sensor would begin to decay and one would lose valuable time to move the device and measure other places! Is there a quantitative way to phrase "as long as possible"?

The description of the in-field co-location (when an NO2 sensor is compared to a close-by reference sensor) is quite confusing. It took several read-throughs to really understand when and where everything was taking place. This could be greatly simplified by adapting the map figure with notes.

P. 9, lines 6-14: The authors claim an in-field co-located NO2 sensor stays calibrated at another site, but the error bars on those measurements are the same as the absolute value of those measurements. How can one be sure they are not just looking at noise?

Figure Comments

Each figure should be able to stand alone and tell a story; many of the figures do not contribute substantially to the paper and could be omitted. Specific comments include:

Figure 1 needs labels for the co-location stations (text) to make it easier to understand what was taking place

Figure 3 demonstrates a large absolute error on some of the RH and T measurements (15 C swing on Temp and 20% on RH). Why? Should counts be converted to volts to

ease comparisons with existing literature? What is going on with the clear outlier?

Figure 5 should have more descriptive axis labels – using just the title to describe the plot makes it hard for the reader to understand what is going on. Many of these plots are not needed ([row 2, col 2], [row 3, col 1], [row 3, col 2], [row 4, col 3]). The authors claim ozone is correlated with AE response, but clearly, that is just a temperature effect. Otherwise, the authors need to describe how ozone can diffuse across the analyte and undergo a redox reaction at the AE surface.

Figure 6 is not needed. It does not add anything to the paper and is well known through basic photochemistry.

Figure 7a should not include the model with ozone in the regression (row 2, col 2)

Figure 8a does not do a good job at conveying the point (that transient temperature spikes affect the signal) since temperature is not shown anywhere.

Figure 8b is not needed. These details are in the technical spec sheet and previous literature – just cite those.

Figure 9 is not needed – simply describing the start-up/warm-up period in the methods section along with other filtering methodology was sufficient.

Figure 10a and 10b do not seem to convey what you are trying to convey – plotting a distribution of the residuals during the two co-location periods would be much more helpful.

Figure 11 was already described in a Table – no need for a plot as well. They are very confusing and don't add anything in terms of advancing the story. It just makes it seem like the linear model is not very robust or repeatable. It also appears to suggest the is a drift in the y-intercept of nearly 1000 ugm3 in some instances!

Figure 12b could also be plotted as a distribution of residuals – one would then be able to see clear overlap (or not) if there is/isn't bias.

[1]: http://xlink.rsc.org/?DOI=C5FD00201J [2]: http://pubs.acs.org/doi/full/10.1021/acssensor

Interactive comment on Atmos. Meas. Tech. Discuss., doi:10.5194/amt-2017-43, 2017.