Atmos. Meas. Tech. Discuss.,
doi:10.5194/amt-2017-183-AC4, 2017

Atmospheric Measurement Techniques Discussions

Interactive comment on "Intra-urban spatial variability of surface ozone and carbon dioxide in Riverside, CA: viability and validation of low-cost sensors" *by* Kira Sadighi et al.

Kira Sadighi et al.

kira.sadighi@colorado.edu

Received and published: 10 December 2017

Anonymous Referee #3

In Sadighi et al., the authors describe the calibration and deployment of the UC Boulder U-Pod for measuring ozone and CO2 in the Riverside and San Bernardino counties for a period of three weeks. The dataset is meaningful as it discusses not just validation of the low-cost sensor but also an actual deployment.

GENERAL COMMENTS I believe the contribution of the authors is valuable and should be published in AMT if major revisions are provided. Overall, I found the structure of

Printer-friendly version

Discussion paper

the manuscript to be confusing at times to follow – the authors make specific references to various units as D0, DA, DB etc without many reminders to the significance of the deployment locations. Sensor performance should be put in the context of the expected nearby sources (e.g., it makes sense that the sensor near Highway 91 would be different). I also think the manuscript is lacking in synthesis of the findings – many of the findings are stated without much interpretation. I also think both the calibration approach, pre- and post-treatment of the data and the calibration models should be discussed in significantly more detail. You mention the significance of a time variable in the calibration, but the magnitude and direction of this coefficient isn't discussed – would this not be critical for other uses to decide on an appropriate sensor? I am also a little concerned on how data was screened to be included in the manuscript. I don't think it's very clear what data were omitted – you make some reference to only looking at data where the deployment exactly matched the calibration range. Why did you do this? Why would you expect your model could not extrapolate? I think given the structure of the mode extrapolation should be discussed. Lastly, I would suggest you comment on how this study might translate to other areas – Riverside has some of the highest ozone in the U.S.; I would imagine that the sensors would have the least difficulty accurately measuring ozone in this area. What about other areas of the US that have occasionally observe high ozone concentrations but where it is more erratic (e.g., Pittsburgh can be in non-attainment, but average ozone is generally Âń 30 ppb.) Some generalizations on the findings would strengthen the paper. Lastly, I echo other reviewers when they say the CO2 discussion as written does not seem to add much value to the paper. I am also not clear on the value of the results or the impact of this auto-calibration which is underdiscussed in the manuscript. Specific comments follow.

* After much consideration, we have decided to remove the CO2 analysis portion from the manuscript. The availability of the CO2 data, due to technical issues with our certain set of sensors, has made evaluation of that dataset difficult. We feel that it makes the analysis of the ozone sensors less strong. We have also added a few figures to make the ozone analysis stronger. Figure 5 shows information about the

coefficients of the calibrations and Figure 7 and the surrounding discussion shed more light on the topic of drift through time. We have also moved some of the figures in the SI to the main body of the paper, to flesh out the discussion on the calibration itself. More than one reviewer was confused about the time periods discussed in the paper, which we have made an effort to clarify this everywhere.

* The purpose of filtering the deployment ozone data by environmental characteristic was to reduce the error in those values. We didn't train the model to use values outside of the calibration range. The purpose of this paper was not to test whether the model could extrapolate, but analyze spatial differences on the best data we could. We have conceded that LA is an easier place to measure high levels of ozone, but that this study does not speak to those people living in areas with lower levels of ozone, but periods of non-attainment. In the conclusion, we propose this as an important area of future research. Some of the general comments are more thoroughly answered in addressing the specific comments below, marked with asterisks.

SPECIFIC COMMENTS P3 Line 6-7: I am not sure why you mention the CyberSEES project or what it is. Delete? * This research was conducted, in part, under the CyberSEES project which we mention in the introduction. We feel it is important to recognize the project under which our research is conducted.

P3 Line 17: What is this site C? *It is a third site in that study, clarified in the text. er P3 Lines 13-16: In this paragraph you list several findings, but there is not much synthesis of the results or general conclusions. The narrative should be improved. * We have added more synthesis in the discussion section of the paper and framed our results with potential explanations for such trends including proximity to roadways, weekend vs weekday activities as well as meteorological changes. We focused our analysis to show variability existed and that low cost sensors have the capacity to quantify these and less forming hypotheses on what is causing these differences which has been a well-researched topic especially in the LA area.

P3, Line 33: Is it really an ideal test bed? Wouldn't this be the best case scenario where you a) have lots of ozone and b) and lots of sources? I feel like to truly test this question, you would need to try the sensors somewhere more representative of federal average ozone concentrations. * We have edited this paragraph to add clarity to this statement. This region is a good test bed for answering our research question due to the reasons we mention in the text. There is increased interest from the regional governmental regulatory body in sensors and due to the potential health consequences from ozone exposure in this area, which makes it a good candidate. Testing our research question (detect variability knowing our uncertainty) in this region is the first step and then expanding to other locations is the next step.

P5, Section 2.1: I don't think enough is said here about the long term drift or crosssensitivity of the sensors. What is the expected life span? You should be explicit about the tradeoffs. * A more detailed discussion and explanation of the sensor characteristics are described in the text body here. Specific life spans are not mentioned by manufacturers, but long term drift and calibration frequency are discussed further as addressed from Reviewer 1.

P6 Line 6: What do you mean by relationships? Linear regressions? You should be specific early on. * We did mean linear regressions. The term "relationships" and "regressions" have been clearly defined and identified in the text. For reference here, multivariate linear regressions (with interactions) were used.

P6 Lines 20-28: What are the coefficients? What are their directions? Can you discuss the physical meaning behind the calibrations? How significant was the drift over three weeks (i.e., what is the size of p5?). If it is large, then are these monitors really suitable to replace EPA reference monitors? These sorts of details are critical to this paper. *These are great questions and the sensor community is especially interested in these questions (hence previously addressed these comments from Reviewer 1). Therefore, more analysis of the coefficients and dispersion was added to shed light on how these sensors perform over time and variable space. The validation set offers the

most information regarding long term accuracy and integrity. It is important to note that we are not suggesting to replace AQMS but rather employing these sensor networks to supplement AQMS as needed to answer specific research questions. Figure 5 and the surrounding discussion hopefully addresses some of these questions. The median value of p5 is quite small and close to 0. We have also included Figure 7, which shows the difference between the validation from week 1 to week 9.

P7, Lines 10-20: Again, I think some of these model calibrations are critically important and should not be in the SI. The coefficients and their interpretation should be front and center in the paper. * We have moved some of those figures into the main paper and decided to focus the attention on the best performing model (4T, Eq.1) and thoroughly explained it in detail.

P8, Lines 1-9: I am confused by what this means? Did you only use data where there was exact overlap between calibration and deployment for T and RH? i.e., you did not try to extrapolate from the model? It's not clear to me what you mean in this paragraph. Also how many RH sensors failed? A number or percentage would be helpful to assess whether deploying these low cost sensors is a feasible alternative to EPA monitoring. * There has been some confusion here. This comment has been addressed in responses to Reviewer 1. To add more here: Four sensors had unrecoverable data, all of them experienced at least one week of missing data, and 2 failed outright.

P9 Line 1-2: How did you deal with the data from the two ozone sensors? Did you average? Or choose one? This should be briefly mentioned. * This has been clarified in the text. After finding there is very little difference in the estimates from these duplicate sensors, we used the best performing (higher R2 and lower RMSE) sensor in the calibration period in the deployment analysis. Also, see Reviewer 1 comments for more information

P9 Line 5-10: What is this auto-calibration setting? I think more detail is needed here

about what the auto-calibration does, and how you corrected for it. *Analysis of CO2 has been omitted to further investigate at a later date.

P9 Line 19: “We thought that” is informal. Why do you think that? Describe, and also use more formal scientific language. *The language here has been adjusted in the text.

P9 Line 26: Not sure what “Values that are less than around 5 ppb different” means. * Changed the language to be more clear

P9 Table 2: I think you need to do a better job describing how to interpret the results of Table 2. I am not clear on the takeaway. *Table 2 shows the validation statistics from the interactive technique described in the text. Uncertainty from the validation (RMSE from D0 and D7, D5 was omitted) is the uncertainty of the U-Pod measurements.

P13 Lines 11-14: You should provide some rationale for these differences and similarities by time of day. Is the 9AM difference due to rush hour? Why not at 8AM? Please expand. *Additional context and discussion were added to help the reader know what activities are happening around these time frames. We have commented that differences in ozone could be related to rush hour times, but since our U-Pods were not equipped with sensors to detect gasses that would indicate more or less traffic, we cannot be more specific about when these times occur. This comment was also addressed from Reviewer 1.

P15, Lines 1-8: So is the statement here that the disagreement between the R2 metric of spatial variability and the absolute differences is real? Does it not then follow that lower R2 between two sites is not a good predictor of spatial variability? You should be clear on the implications of the findings. From my perspective, if two sites are different based on absolute concentrations, that is the better marker of spatially variability. Especially because R2 can be disproportionately affected by high leverage points in a linear regression. * We expounded on these finding and their differences in the text, specifically explaining the benefits and drawbacks of each comparison technique. We agree in part, absolute differences do present a more telling indicator of spatial vari-

ability, and correlations are more susceptible to misinterpretations or leveraging from outliers.

P16, Lines 7-10: Are you sure this is really spatial variability and not some confounding factor with the sensor? I am not convinced the data you have collected here is sufficient to claim that the interesting features of D0, DA, DB and DE are related to spatial variability. * This is a reasonable hesitation to have. When the scatter/dispersion of the measurements is compared to the expected uncertainty we found during the validation, we are certainty seeing variability. Now for the claw shape, we agree, and have changed the narrative to actually suggest some sort of confounder acting temporarily for a few hours at most. Please see the added analysis and discussion around Figure 7.

P16, Line 24: Proximity to a major highway would be a critical factor in differences between two sensors, even if they are <2 km apart. There is a very near-road effect of NO+O3 forming NO2 – is the O3 at site DC consistently lower? The direction of these differences should be included. * Details regarding the exact explanations of the higher and lower values observed are limited to hypotheses but we do mention proximity to highways and the influence of NOx and VOCs as potential reasons why some sites/measurements are lower or higher.

P18 Line 11-12: Please rephrase the question as a statement – inserting a question like this seems gimmicky. It is up to you as the researcher to provide a hypothesis for the observation and inform the reader succinctly what you observed and whether this matches expectations. * We have rephrased these lines.

P18: Lines 15-17: Again, I am not convinced this isn't just an artifact of the sensor. You should be clear on the potential uncertainty. *The uncertainty of the ozone sensors is well referenced using the validation from D0 and D7. Analysis of CO2 has been omitted to further investigate at a later date.

P21, Line 11: Some hypotheses should be provided. The discussion should be more

than statements of observations but also include some scientific assessment of what was observed. What kind of future investigation is warranted? What would that future study look like? * Discussion and synthesis of the observations and trends is boosted but our main goal is to show that the sensors are viable for detecting variability on the scales (temporal and spatial) mentioned. Analysis of CO2 has been omitted to further investigate at a later date.

P22, Line 1: This is the first time, to my knowledge, that you directly state that the MOx sensors cannot replace EPA monitors. Some space should be devoted to discussing the suitability of the U-Pods for monitoring and what is a reasonable expectation from the units. * A clear discussion of the limits of the sensors and their desired applications as we see it has been added here and a clear message has been added to the beginning of the paper that the intent is not to replace the AQMS but rather to supplement them for more information.

P22 Line 3-6: Can you expand on the computation time, people demand, etc to give a sense of the rigor involved in the deployment as a guide to others? *Additional information on time resources of using the sensors was added.

P22 Line 19: What is “frequently”? *As addressed in Reviewer 1 and 2, additional discussion has been added on this point, including figure 7, which show the difference between the validation of week 1 and week 9.

TECHNICAL CORRECTIONS P1, Line 11: I am not sure it’s grammatically correct to say the tool provides low-cost sensors. This just didn’t make sense to me. * We changed “provides” to “contains”. The U-pod (the tool) contains sensors. I would replace every instance of “Author and coworkers” with “Author et al.” – e.g. on P2 Line 29, and elsewhere in the manuscript. * We have taken this suggestion and made changes in the text.

P3 Line 4 and 5 “the Riverside-San Bernardino counties” *Changed

P12 Line 6: Some word missing here in “Figure 4 gives some context temporal variability” * This was changed.

P22-23, Line 31/Line 1: You should just delete the sentence about it being a lot of work. *This was removed. paper P23, Line 3: “undergo” vs “undergone” * This was changed.

P24: Why is there an appendix that is separate from the Supporting Information? I would just put in the SI? * This has been addressed by a previous comment. It has all been combined into the SI.

Interactive comment on Atmos. Meas. Tech. Discuss., doi:10.5194/amt-2017-183, 2017.