Review of "Impact of Sampling Frequency on Low-Cost PM Sensor Performance"

General Comments

This manuscript investigates the impact of sampling frequency on lower-cost PM sensor precision, accuracy, and plume detection capability in a polluted megacity, Delhi. The dataset itself could be a useful training/exploration dataset with the right application in mind. It spans a dynamic polluted month with variable meteorological conditions while collocated with a reference instrument. However, the paper is largely unfocused on a specific application - reading as an unhappy hybrid of a measurement report (listing off communication protocols and daily readings), and letters paper (brief punchy analysis). As a result, the findings are muddled with distracting information more suitable to an Supplement. Furthermore, the analysis itself lacks impact. The findings on LCS sensor accuracy as a function of averaging time has been investigated previously in literature - in North India no less (doi.org/10.5194/amt-11-4823-2018) - however, here it lacks a nuanced discussion of aerosol optical characteristics and the influence of different sensor technology on the underlying quantification of plumes. It is unclear how these findings could be directly applied in a novel way for health studies, regulation, source apportionment, etc.

In addition to the underlying structural problems of the manuscript, the organization of the paper is weak. Nearly every figure requires regeneration or rethinking to meet the standards laid out by the publisher. Furthermore, literature review largely sticks to well cited more general work (e.g., Zimmerman et al) while not citing similar work from the region (e.g., doi.org/10.3390/s20051347). A more through analysis of work from Delhi and similar high background signal environments is sorely needed to better contextualize the results.

With these criticisms in mind, as well as my more specific comments below, I recommend the paper be **rejected** as unsuitable for publication in AR at this time. I empathize with the authors as I fully understand the difficult and often grueling nature of field work and scientific writing. I believe the underlying data is useful. If the authors develop a new hypothesis focusing on a specific application (perhaps in coordination with the other instruments/historical data at IIT), the work could eventually lead to a useful reference for the broader community. I wish all the authors good luck.

Specific Comments

- The title is too vague. I recommend a hypothesis centered title, something like "Lower Cost PM Sensor Accurately Quantify Short-term Plumes in High Background Polluted

Environments." This is more informative to readers and easier to find in search engines.

- Citation formatting is incorrect. In-text citations should be in chronological order, and the references section should be organized alphabetically by last name. This may seem overly picky, however cross checking citations is key to the review process, and missing this step unnecessarily added time to my review. See other recent preprints (e.g., ar.copernicus.org/preprints/ar-2024-35/) for examples.
- In the introduction, please define what affordability means in the context of "low-cost" sensors. Are you using a benchmark for "low-cost" (e.g., compared to other non-regulatory optical monitors)? Furthermore, it maybe worth noting these are low capital cost sensors - clearly the work in deployment, and calibrating amounts to a very high cost.
- In the second paragraph of the introduction, you write LCS are compared to OPC. However, recent work (doi.org/10.1080/02786826.2023.2285935) has conclusively demonstrated LCS such as the Plantower cited this work are OPCs, albeit low efficiency OPCs. Please briefly clarify the differences in operating methodology between a reference OPC and an LCS OPC. Additionally, highlight how this difference could lead to mischaracterization of short term peaks. Especially since OPCs infer mass concentration from particle counts rather than directly measuring mass.
- The third paragraph references opinions or beliefs from the scientific community using language like "widely assumed" or "lack of understanding" but offers no literature or evidence. Please either support these statements with some theoretical backing in supplement, and/or cite (preferably recent) review work on the current understanding of high resolution monitoring.
- Furthermore, in the third paragraph of the introduction it would help to be specific when using terms like "high frequency" and "low frequency." Techniques such as eddy covariance would consider 1 Hz or higher standard, while many epidemiological studies use annual or decadal data as standard. Only "high frequency" is defined earlier in the manuscript.
- Again, in paragraph 4 of the introduction it is worth discussing what the LCS actually measures. For example, most PM LCS don't accurately identify resuspended dust from vehicles which can make up a significant portion of supramicron (dp > 1 micron) PM2.5 nor fresh vehicle emissions (dp < 0.1 microns). While the effective size range of most PM LCS is still useful, it largely lies in the accumulation mode, which includes a big share of the regional aerosol burden rather than the hyperlocal burden. Please rework this paragraph to discuss which short-term PM signals LCS are useful to quantify in the light of previous epidemiological work and aerosol fundamentals.
- The last paragraph of the introduction needs to be more specific on why this month was useful, why the location is valuable, and why the Sensiron sensor rather than the Plantower, or Honeywell monitor. Citing other work contrasting the Sensirion with other sensor models, especially in Delhi or similar environments, is useful.

- There is a key strength of this manuscript which is excluded from the introduction and should be better described. This work focuses on characterizing plumes in highly polluted ambient environments. Much of the previous work in this area focus on comparing plumes in low to moderately polluted ambient environments in which the signal to background is high so analysis is straightforward. Despite the largely regional signal in Delhi, for which the design of many LCS is optimized, this work demonstrates its usefulness in source identification.
- Figure 1 should go to the Supplement, as should probably all of Section 2.2. These details are ubiquitous to most LCS packages, and are distracting in the main manuscript.
- I suggest adding 2 or at most 3 sentences in Section 2.3 contrasting how the BAM technique assess particle mass concentrations differently than an optical technique and why it matters for source identification.
- Lines 85-90 would be more efficiently communicated as a table with the 25th, 50th, and 75th percentiles for the relevant parameters.
- Figure 2 should be more clearly illustrated. For PM2.5 and RH, please ensure the y-axis minimum is **exactly** 0 bounded. For all plots the y-axis minimums and maximums should be clearly marked. The x-axis date labels should be more clearly marked, Oct 29 and Nov 1 overlap in a confusing to read way. Consider excluding the year after the first date marker. Adding grid lines would also improve interpretation.
- Figure 4 needs to be entirely regenerated. The top panel (a), should not contain both reasonably 0 bounded elements (i.e., PM2.5 and RH) on the same axis as temperature. Additionally, please label the beginning and end points on x-axis. Furthermore, when comparing different colored lines on the same plot, it is best practice to use color-blind friendly color schemes (see EGU publishing guidelines for more details) and/or different line styles. In this case, I'd recommend 3 different subplots for each variable. Panel (b), as in Figure 2 should have clear bounds, and be 0 bound on the y-axis minimum. Also, the red text in panel b should be better spaced to avoid overlap with the highlighted boxes or other text. Panel (c) looks to me like a screenshot from Plotly. I strongly recommend against Plotly and similar interactive tools for publication since they are difficult to reproduce. Also the y-axis label in Panel c is insufficient (no units!). Please remove Panel (c), or replace with a clearly labeled static boxplot, or use a table to describe the key percentiles of the monthly distribution.
- Figure 5 needs to be entirely regenerated. All scatterplots should be "square" - or equal ranges on both the x- and y-axes. The long lists of metrics would be easier to read on an accompanying table, with the panel labels in the upper left corner of the plot. Additionally, a legend denoting what each line and scatter point corresponds to is compulsory. To my eye, it appears as though there is some red shading around the regression line. Please explain how this is derived and how it is relevant. If not relevant, remove. I would guess its a confidence interval based on the sample size and therefore irrelevant and distracting.

- Figure 6 needs to be entirely rethought. As it stands the figure is too low resolution to read normally, I had to zoom in to read it and it rendered blurry. The "clock plots" don't follow best practices as highlighted in the previous figure comments, and don't add anything not clearly visible from the time series plots. Focusing on a few episodes can be helpful, especially as these demonstrate clear contrast, but the takeaway seems uninteresting.
- Figure 7 is also too low resolution (i.e., blurry) and should use colorblind friendly perceptually uniform colormaps in accordance with EGU publishing guidelines. Furthermore, there should be some discussion of aerosol properties which explain the "swings" in LCS hourly and sub-hourly data. Clearly the LCS is also systemically **underestimating** in addition to overestimating (probably due to RH?). Therefore, the magnitude of the plume spikes maybe incorrectly quantified relative to the BAM. Some baseline-spike decomposition algorithm could more holistically address these limitations. There are many in literature, especially in mobile monitoring.
- Overall the plume discussion doesn't focus on investigating aerosol characteristics. Simply finding a spike does not necessarily offer useful or actionable information. Therefore topics of interest to the community including background-plume decomposition analysis - which is key to both epidemiology and source characterization - are not clearly discussed or analyzed in the conclusion.

Technical Comments

- Line 65 - Needs a citation (maybe more than 1) when you assert its "among the best"
- Line 66 - It is customary to refer to Supplement items as Table S1, Figure S1, or Section S1. Refering to the sensors as S1, S2, etc. could be confusing for readers in publication.
- Lines 87-88 - This is not PM2.5 exposure, its simply the mean mass concentration.
- Line 101 - While higher temperatures *can* result in lower PM2.5 concentrations, I think this is merely an indicator of increased ventilation due to higher planetary boundary layer height during midday rather than a direct temperature effect - please clarify.
- Line 112 - This looks like the standard way to calculate SD, please either simply cite a statistics textbook/manual or move to supplement.
- Line 118 - Since you're using US-EPA standards, please state the US-EPA guideline for CV (about 0.1 or 10%, although 0.2 or 20% is a commonly used more lax requirement for LCS).
- Line 125 - Please explain or cite why this is the case