Atmospheric Measurement Techniques Discussions

Atmos. Meas. Tech. Discuss., referee comment RC2
https://doi.org/10.5194/amt-2021-5-RC2, 2021

Comment on amt-2021-5

Anonymous Referee #2

Referee comment on "SIBaR: a new method for background quantification and removal from mobile air pollution measurements" by Blake Actkinson et al., Atmos. Meas. Tech. Discuss., https://doi.org/10.5194/amt-2021-5-RC2, 2021

General comments:

This paper presents a novel method to quantify and remove background signals from air pollution data, relevant to the processing and interpretation of mobile monitoring measurements. This has the potential to be a significant methodological advance relevant to a broad body of mobile monitoring studies. However, I find that the paper in its current state does not clearly demonstrate that the novel method improves upon prior techniques, nor does it fully justify the additional statistical complexity inherent in this new method. I would suggest significant additions to this analysis to more carefully interpret the output of the HMM, consider the physical/mechanistic interpretation of the signal classification, and evaluate how the conclusions drawn from mobile monitoring data may be altered by the removal of background signal.

I find two significant flaws in the current analysis. The first is simply that the prior methodology by Brantley et al. (2019) seems to outperform the proposed SIBaR in representing background concentrations and also seems to produce very similar spatial results, leaving the reader questioning the purpose of this more conceptually complex and computationally intensive method. SIBaR has the attractive quality that it provides data-driven, variable time windows in its signal classification scheme, a method that performs well compared with manually classified data (Section 3.1) potentially making it a method that can be adapted to other data series and settings. However, that data-driven quality is somewhat undercut by its sensitivity to the initial smoothing of the input time series. It would be useful for the reader if the authors could point to use cases where the Brantley method may result in significant misinterpretation or misclassification of the data while the SIBaR method performs more favorably. I understand that the comparison against the elevated fixed site monitor may have been an attempt to do so, but I believe there is a fairly reasonable hypothetical explanation for why this failed, which I address in the following paragraph. Are there other ways that SIBaR is more replicable, portable, or robust that provide it an advantage over the Brantley method?

The second flaw with this analysis is conceptual and relates to the division of mobile data into two distinct modes of background (defined in the introduction as "measured air pollution independent of local source influences") and non-background/source. I agree with the concept that within mobile measurements there is a hypothetical pollution signal that is time-variant but spatially invariant, and this signal should match concentrations at a semi-remote background monitoring site. However, following the framework described by Shairsingh et al. (2018) (included in this manuscript's references), a mobile measurement represents the superposition of several time- and spatially-variant patterns including (a) this spatially-invariant hourly/daily background, (b) spatially variant/neighborhood elevations, and (c) isolated spikes caused by localized and/or transient emission plumes. Visually, the data classified by SIBaR as "background" appear more similar to (b) type signals than (a), further underscored by how much higher the SIBaR signal appears compared to the Brantley signal in Figure 4. This is not necessarily a failing of the SIBaR method! It is useful to distinguish between (b) and (c) signals. However, an evaluation of SIBaR results against methods for isolating the (a) signal may sell the method short. There may be a rich array of conclusions that could be drawn by looking at characteristics of the "background" vs. source signals, grouping by neighborhood and considering how parameters of the distributions may vary. I was also curious to know whether any useful information was captured by the time covariate included in the HMM (Line 97-98).

I would suggest that the authors consider the conceptual reasons why the Brantley background removal method outperformed SIBaR in the current evaluation frameworks and whether SIBaR provides value if it is evaluated in a different framework.

Specific comments:

- This is a nit-pick but applying a temporal smoothing of 30 seconds on data collected in-motion results in a spatial smoothing effect that means that the measurements do not quite match the nominal resolution of 50-meter road points (as per discussion in Chambliss et al. 2020, cited in this work). This isn't a crucial problem, but it does mean that the road segment observations presented in Fig 8 and related analysis aren't independent data points. I don't think it warrants restructuring the analysis, but it may be worth mentioning.
- It would be nice to see some additional information on the time parameter described on Line 97.
- RE: Line 98, "we assume that the probability distribution governing yt are log normal" – is this assumption justified for the background signal? I can understand why a time series with plume-related peaks would be log-normally distributed but why would we assume a long tail for background measurements?
- Figure 2: Is there a reason that these figures present the transformed data and not concentrations? For ease of interpretation, it would be useful to show these in units of ppb.
- Lines 244-245: The authors mention source-dominated hot spots (presumably other than roads) but these are not obvious to the reader, absent local context for interpreting the maps. It would be useful to include annotations on map figures if possible.

Technical corrections:

- Line 53: "a way to determine whether measurements were taken in locations representative of background versus locations subject to local influences." Precisely speaking, the method applies to a time series and not a set of locations, so it determines whether measurements were taken during periods representative of background patterns vs. periods of transient plumes or localized elevations.
- Line 125: "designate **points** as background or source. State assigned **points**"—to me, "points" suggests a location and here you are referring to observations in a time series, so I would prefer the word "observations"
- Line 151-154 vs. 228-230: The authors contradict themselves in describing the fixed site as background and then walking that description back when interpreting the results. If you believe that it is influenced by transportation emissions patterns it would be appropriate to include that information in the original description.
- Figure 8: "Best Fit" description is the same in both panels