

Atmos. Meas. Tech. Discuss., referee comment RC1
<https://doi.org/10.5194/amt-2021-5-RC1>, 2021
© Author(s) 2021. This work is distributed under
the Creative Commons Attribution 4.0 License.

Comment on amt-2021-5

Anonymous Referee #1

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-RC1>, 2021

The authors present a new technique for estimating and removing background concentrations of air pollutants from mobile measurement data. A major benefit of this method is that it does not depend on a static time window in estimating background concentrations and instead uses a time-varying approach. The authors compare the results using their developed SIBaR method to two previously-published methods as well as datasets from a stationary site and data collected in a prior mobile measurement campaign. The authors conclude that the SIBaR method results in background estimates similar to those obtained when a previously published technique is used on the same dataset.

Overall my suggestion is to reconsider the paper after major edits have been applied. My two specific points of concern are in regard to the susceptibility of the SIBaR method to outliers and to the high modeled background values shown in the time series in Figure 4. First, the strong dependence of the SIBaR technique on the smoothing time interval is concerning. Looking at Figure S1, the background designations vary wildly and do not make much sense when the data is not smoothed or when 10 second smoothing is applied. Figure S1 also shows that the background designations (in ppm) are much higher than source designations, which is worrisome. If this is not the case then the technique and figure have not been sufficiently explained. Overall the current technique appears too susceptible to time series outliers. Secondly, the SIBaR-modeled background concentrations in the example time series shown in Figure 4 are much too high compared to the other 2 techniques applied to same dataset. The SIBaR-calculated backgrounds in Figure 4 are higher than most of the total ambient measurements, which doesn't make any sense, and the calculated background values do not show enough temporal variability in the background over the course of the day. This Figure is meant to present a single day as an example, however it shows that the SIBaR technique may not be very accurate in modeling background concentrations.

In addition to these concerns, the authors did not sufficiently address the observations that are an essential part of this research. The paper needs to provide more details on the observations such as the accuracy, precision, limit of detection, etc. of each

instrument or measurement technique to provide the reader with some idea of the uncertainty associated with measurements of each pollutant. The reader also does not know the dates or seasons of the measurements, which are vital parts to list for any measurement. This is important because air pollutants such as CO₂ which exhibit a seasonal cycle and therefore observations in January cannot simply be compared to those in July. For Table S2, a third column in the table listing the measurement techniques for each of the analyzers would be great.

Below are general comments regarding the figures.

Figures 1, 2, 3, 4, 6, 8: The figures do not "stand alone" if read without the rest of the paper. Even so, some of the figures are not easy to understand even with the text. The captions are very short and need more information such as the dates, locations, and temporal resolution of the measurements, a brief description of the modeling techniques used, etc. Some other suggestions are to take the titles out of the figures and put the information into the figure captions instead. For figures with multiple subplots, designate one as "a" and the other as "b". For the x-axis, specify the time zones (local time, UTC, etc) on your plots. If a legend is provided, make sure it includes units.

Figure 3 specifically: Please check your units for CO. Thousands of ppm is VERY high. I'm not sure if these measurements were taken over a single day or if they were averaged at each time-of-day across multiple days of measurements.

Figure 4 specifically: If I'm understanding this figure correctly, then it appears that SIBaR overestimates the background concentrations by a lot, and this is rather worrisome. The background concentrations estimated by Brantley and Apte appear much more reasonable, although the Apte technique may have too high of a temporal variability.

Figure 8 specifically: Check your y-intercept and slope values in the two plots. They are the same values but the bottom best-fit line should be much different. Please explain the subplot within the figures in the caption. The circles are rather large and hide some of the data behind them. Perhaps consider using small points as markers instead.

Below are specific technical comments to consider for the manuscript.

Specify "Houston, TX" the first time it is mentioned on line 17 (don't need to repeat "TX" on line 19).

Line 31: "Air pollutant concentrations" in place of "pollution concentrations"

Line 32: "ambient background levels" in place of "background"

Line 36: "to determine the background" instead of "to determine background"

Line 47: "background concentrations" instead of just "background"

Line 47: State which traffic-related air pollutants you're looking at here.

Lines 66-67: Need to state the temporal resolution of the data you collected (hourly, minutely, secondly, etc). Also provide the year and months of the campaign. This is especially important for emissions that vary seasonally, such as CO₂.

Line 72: "neighborhoods" and "highways"

Lines 75-76: Only need to list the pollutants you did use in your study, not ones that weren't used.

Line 81: "for each GSV" car

Line 130: Please define what variable "k" represents.

Line 130: "day-by-day"

Line 134: "NO_x concentrations"

You mention that the RMSE equation is provided by equation (10), but you need to state this equation after the first mention of it here.

Line 139: Include the units for the RMSE of 0.01 (ppb?)

Line 142: "Mobile monitoring dataset?"

Line 143: "Collecting measurements"

Line 143: Define CO at the first mention.

Line 147: State which measurements you're comparing. Carbon monoxide?

Line 151: Be consistent with writing out numbers. In the abstract you wrote "70 m". Either replace "seventy meters" here with "70 m", or replace "70 m" in the abstract.

Line 152: Replace "it" with something like "this site". Also, specify if you mean "Houston background NOx concentrations" or emissions.

Lines 154 and 155: Again, be consistent with spelling out fifth or 5th (similar to Table 1) and tenth.

Line 159: Time series of what pollutants?

Line 168: "...the mean of measurements collected as the GSV car drove past a road segment..."

Lines 169-172: Confusing statement. Please revise.

Line 188: Change "data set" to "dataset".

Lines 189-190: Unclear with what you mean by this sentence.

Line 195: Specify the dates of your comparison.

Line 198: Don't need to redefine RMSE.

Table 2: Need to include more detail in the table caption, including the dates of the measurement periods, the locations, etc. Are the RMSE and MAE values for the entire dataset or an average or median of daily-calculated RMSE and MAE values? Need to state this.

Line 218: "outperforms" in place of "out performs"

Lines 218-219: This is the first time that the reader is introduced to how many days were considered in the analysis. Please be sure to include this information in the methodology with details on the exact dates of measurements.

Figure 4: Place Figure 4 after the paragraph that first mentions it (after line 230).

Lines 223-225: Confusing as written.

Lines 227-228: Please rewrite the second "local pollution influences" with another term to not repeat the same wording twice in the sentence.

Figure 5: Move to the end of line 238 where Figure 5 is first mentioned. Good caption!

Line 267: Please specify which pollutants.

Line 272: "highways" in place of "highway"

Line 283: Add "regression" after OLS

Figure 9: Move to below line 306, after the first mention of the figure.