Atmos. Meas. Tech. Discuss.,
doi:10.5194/amt-2017-316-RC2, 2017
© Author(s) 2017. This work is distributed under
the Creative Commons Attribution 4.0 License.

Atmospheric Measurement Techniques Discussions

Interactive comment on "Adaptive Baseline Finder, a statistical data selection strategy to identify atmospheric CO_2 baseline levels and its application to European elevated mountain stations" *by* Ye Yuan et al.

J. Kim (Referee)

jjkim@ucsd.edu

Received and published: 8 November 2017

This work presents a new statistical algorithm, named ABF, for identifying "baseline" levels from CO2 measurements. The title of the work refers to elevated mountain sites as its application focus, but the work also includes some analysis of non-mountain sites as well. While there are some issues that I would like to see the authors address, overall I do feel the authors have done a good job of presenting a unique algorithm and comparing it to other frequently used methods in the measurement community, and as such I suggest that the manuscript be published with some revisions.

Printer-friendly version

Discussion paper

Before I proceed with my comments on the paper, I would like to comment on the term "baseline" itself. My concern is that the definition of "baseline" is very subjective open to interpretation. For example the authors mention that ABF in this study was used specifically to identify periods of free troposphere concentrations in the high elevation sites, and that is certainly one valid definition of "baseline". With this definition, however, sites that may have statistically stable concentrations at certain times of the day but do not necessarily measure the free troposphere will by definition have no "baseline". If the definition of "baseline" was "typical concentrations you would probably measure at a certain location at a certain time" with the goal of creating a global spatial map of average concentrations, I suppose you would end up with something close to the trend and seasonal components in the STL analysis, which you may (or may not) be able to find through statistical methods such as ABF. On the other end of the spectrum, for a regional modeler, the useful definition of "baseline" would be whatever concentrations enter the modeling domain and not necessarily any clean/stable condition, and if the air was polluted coming into the grid box then the model needs to know about it. I've seen attempts to distinguish between "baseline" and "background" to try to navigate through the subtle (and sometimes not-so-subtle) differences in definitions, but in my view all attempts at defining "baseline" is inherently subjective and the best practice is to be specific about what the particular definition for the study is, and that definition should encompass the specific intended use of this definition. All this to say, I feel the name Adaptive BASELINE Finder, while sounding nice, can be misleading. I would suggest that the authors consider another name, but will leave the decision to the authors.

[General comments, questions]

- P5, ln 15: Why the window of 6 hours? I suppose this assumes that baseline conditions occur for longer than 6 hours? Have you tried shorter windows and found you come to the same conclusion? I almost wonder whether it would be more beneficial as a general algorithm to have as short a window as possible, such that the window never exceeds the actual window of a baseline occurrence?

CC BY

- P10, ln 15: The increase in the mean annual growth rates is within the noise, I'm not sure that much can be made of this.

- Figure 2: I have a hard time understanding this figure. First off, the figure seems to represent data from the full data set (spanning years), and yet the method describes that the baseline "window" is adaptive, potentially changing each day and by season. What criteria was used to derive a representative window for the whole period?

- P10, ln 27: Regarding "active vegetation", wouldn't signals from respiration also explain these results, and wouldn't that also be one form of active vegetation? I think this possibility can't be ignored since the authors suggest that the lower VAL values in summer are likely due to vegetation. Are the anthropogenic emission activities in this region such that you would expect emissions only in winter, or are they small enough to be masked by the summer drawdown? I do think that the authors' interpretations on the findings are likely to be correct, however I do think that a much deeper analysis of the data (perhaps beyond the scope of this paper) may be needed to conclusively determine the source of these discrepancies.

- One discussion I think is missing is regarding the "adaptiveness" of the algorithm, in other words do the results show baseline windows changing with season. The authors state this as a strength of the ABF (P4 ln 29), so I had expected this to be one of the early points of discussion.

[Minor comments]

- Page 4, ln 10: "At last", change to "Finally"?

- P4, ln 27: "No upwind air masses with depleted CO2 levels by photosynthesis of vegetation like in summer are recorded." -> "Unlike summer, no upwind air masses with depleted CO2 levels by photosynthesis of vegetation are recorded."

- P5, ln 12: "but preserves of the diurnal pattern." -> "while preserving the diurnal pattern."

CC BY

- P6, ln 10: “Step 3” is not actually a step, but a general description of Step 5 and 6. Perhaps it makes more logical sense to include it in “Step 2”, presenting it as an “If/Else” step.

- P9, ln 5, Table 2: Can the authors clarify whether the percentages are based on just the time windows considered in the algorithm or the complete dataset?

Interactive comment on Atmos. Meas. Tech. Discuss., doi:10.5194/amt-2017-316, 2017.

Printer-friendly version

Discussion paper

CC BY