

Atmos. Meas. Tech. Discuss., referee comment RC1
<https://doi.org/10.5194/amt-2021-16-RC1>, 2021
© Author(s) 2021. This work is distributed under
the Creative Commons Attribution 4.0 License.

Comment on amt-2021-16

Anonymous Referee #2

Referee comment on "An algorithm to detect non-background signals in greenhouse gas time series from European tall tower and mountain stations" by Alex Resovsky et al., Atmos. Meas. Tech. Discuss., <https://doi.org/10.5194/amt-2021-16-RC1>, 2021

The manuscript presents an algorithm for detecting anomalous signals in European atmospheric CO₂ and CH₄ data by decomposing the time series into background (long-term and seasonal), non-background (synoptic/regional) elements, and very localised emissions spikes. The method relies on the CCGCRV decomposition algorithm of Thoning et al. 1989 with additional LOESS smoothing applied with spans of 30 and 90 days to isolate anomalies on synoptic and seasonal timescales respectively.

I think the manuscript warrants more clarity in a few places as mentioned in the specific comments below. In particular, clearer definitions of some of the terms (seasonal, synoptic, etc.) and demonstration of how they are treated by the framework, as these terms are susceptible to subjective interpretation. I would also recommend that in the results section of the manuscript, the interpretation of the events should either be strengthened with evidence that precludes a dominant role of atmospheric transport influences, or that the language here should be softened as I find a few of the interpretations unsupported by the observational evidence in its current form.

Nevertheless, this manuscript is generally of high quality and will be of value to the atmospheric greenhouse gas community; it therefore justifies publication in AMT, subject to some alterations as suggested below.

Specific comments

I think some additional practical information about the algorithm would be very helpful in the methods. For example, is the algorithm designed only for use with daily data? Can the smoothing spans be easily altered by the user to other values if required? What is the minimum amount of data required to make reasonable fits (e.g. 1 year, 3 years, etc)? Is the algorithm sensitive to end effects?

Top of page 4: it seems a bit limiting and perhaps a bit subjective to assume that synoptic events predominate only in the winter and seasonal perturbations occur only during the growing season. I can think of instances where this might not be the case, such as uncommonly warm winters, and periods of summer storminess/heavy rainfall. Would it not be better to apply both smoothing windows to the whole year and allow the data to determine what kind of event occurs and when, instead of this being pre-defined by the user? It is surely also possible in reality for both types of event (seasonal and synoptic) to occur at the same time, superimposed on each other. How would the algorithm treat such cases?

Top of page 5: is the time window user settable? I think so, but this could be made more explicit. Although afternoon periods are favoured in many analyses, there might be some instances where it is more appropriate to use the full daily period. How does this pre-selection of only 5 hours of data per day (at non-mountain sites) influence the detection of events?

Figure 3: I would recommend adding plots of the deseasonalised anomalies, and somewhere I would also like to see plots of the trends, since the way CCGCRV assigns 'seasonal' and 'trend' variability depends on the settings and these things are not independent of each other (more variability in the seasonal component usually results in smoother trends, and vice versa). In general, I think the manuscript would be stronger for clearer definition and demonstration of the terms "long-term trend", "seasonal cycle", "synoptic", "localized", etc., perhaps including some discussion (or by utilizing the figures) to explain the limitations/assumptions associated with each term. E.g. Here, we have assigned variability longer than XX days to the trend component of the decomposition procedure, thus excluding it from the seasonal cycle, etc.

Page 9: again, I am a bit concerned about the assumptions made regarding the length of events, when they occur and what they are caused by. It is also not clear to me why the user would only want to identify BLO type events in the winter and not also NAO events (2nd paragraph), or only seasonal events in summer, not synoptic events. If it is the case that these sorts of events in specific seasons have been chosen simply as illustrative examples to demonstrate the scope of algorithm, then I think this needs to be made much clearer to the reader. If this is not the case, I am finding it difficult to see the rationale behind this approach, so I think either more justification is required, or a broader/looser demonstration of the algorithm is needed.

Page 10: is the distinction of localized fluctuations vs SSAs part of the algorithm/framework, or was this done manually afterwards?

Pages 16-17: I would urge caution with the interpretation here. Although the signals observed at some sites appear to agree with the expected signals associated with droughts, heatwaves, etc. a causal link has not been established in this manuscript. It is entirely possible that the signals are predominantly caused by variability in atmospheric transport at the sites and that this just happens to show what you are expecting. Without some sort of wind sector/back trajectory analysis demonstrating that this is not the case,

it is very difficult to be conclusive about these signals and I do not think the influence of atmospheric transport should be underestimated, nor neglected from the interpretation of the results. I would recommend some additional analysis and an extra section showing how the impact of variability in atmospheric transport might manifest in the frequency and magnitude of anomalies. If this is not feasible, I think the interpretation on fluxes needs to be significantly scaled back and the potential influence of atmospheric transport brought much more to the fore in this section.

Page 17, lines 304-305. It is difficult to see a trend in the spikes from Figure 9. I think that at some stations the record is not long enough to determine a trend, and others, such as OPE, there seems to be a reduction in all anomalies, not just negative ones. In addition, if I understand the procedure correctly, you have used a mean representation of the seasonal cycle for each station, but at most CO₂ sites there is evidence that the seasonal cycle amplitude is slowly increasing, especially at northern sites. Any trend in anomalies, especially in the longer records, might therefore be partially explained by the exclusion of this increasing trend in amplitude from the seasonal cycle used in the algorithm. I would actually not expect you to see a reliable trend in the anomalies unless interannual variability in atmospheric transport processes has been accounted for. In my experience, atmospheric transport easily masks trends in emissions (natural and anthropogenic) at most atmospheric measurement sites, unless these trends are extremely significant, which is not the case here.

Discussion section: In general, I think that time series decomposition and anomaly detection are analysis techniques that are fraught with complication and I commend the authors at highlighting many of the difficulties inherent in the process. It seems to me that the method has been designed for use in the ICOS network, or an ICOS-type network. I would suggest that this network/multi-station requirement should be mentioned earlier on in the manuscript, perhaps on page 3, 'We present here...'.

Conclusion: The first sentence needs to be softened considerably unless additional analyses are done to exclude the possibility that the pattern of anomalies predominantly reflects atmospheric transport variability, as per my previous comment.

Technical corrections

Lines 16-17: I would recommend re-wording this sentence as it is not clear whether you mean the stations are uninfluenced or the regional signals are uninfluenced by the local processes.

Abstract and elsewhere: the terms 'weather' and 'atmospheric transport' are used

somewhat synonymously through the manuscript, but these do mean different things. I would recommend the authors pick one and only use that, except in specific instances where a different meaning is necessary. The same might apply to 'meteorological' and 'climatic'. Also 'swath'/'span'.

Line 44 and elsewhere: the usual notation I think is $\pm 2\sigma$, $\pm 3\sigma$, etc.

Line 54: 'A Gaussian was then defined using...' is this a typo? I am not familiar with the word Gaussian as a noun. Should it be Gaussian curve/fit/function? Same with 'Gamma'.

Line 78: at multiple [European] sites?

Line 210: no '-' after NAO?

Figure 5: it looks to me like the positive SSA in CH₄ at HPB at the end of Jan 2019 is caused by a gap, followed immediately by a very transitory positive spike in the data and is therefore perhaps not an SSA at all? I wonder if a similar thing happens also in CO₂ in Fig 4 at GAT in later Oct 2018, only this time the spike (negative) comes first and the data gap is directly after. At GAT it looks like there might be a straight turquoise line interpolating this data gap when perhaps there should not be?

Line 308-309: 'CCGCRV is ill-suited to handle such gaps...' I think this requires a citation.