Review of revised manuscript by Yuan et al.
"Adaptive selection of Diurnal Minimum Variation: a statistical strategy to obtain representative atmospheric CO_2 data and its application to European elevated mountain stations"

The manuscript has improved in this revision, the authors have especially clarified parts of the text. The replies to my comments are also helpful, and clarify the method and some issues. However, there are a number of questions that have been answered only in the reply, and the clarifications were not included in the main text. I highlight these questions/answers below and think this should still be included in the manuscript. Besides these suggestions, I also give some comments below on the revised version.

The authors write that the manuscript was proofread for English, however, the level of English in this new version still does not seem adequate and would need to be improved.

All in all, as I wrote in the review of the first version of the manuscript, I think the work done by the authors is interesting enough to be published, after taking into account the following comments.

=======

(Some comments from the first review are repeated here in gray. Author replies are included in blue, and new comments are given in black.)

Some comments on the revised manuscript or replies to the first review:

I agree with the second reviewer that replacing the terminology "baseline" in the name of the presented method is a good idea. However, I think that the new name "Adaptive Diurnal Minimum Variation (ADMV)" is not clear and possibly not correct English, and specifically I do not understand what the authors mean by "variation". The previous choice for "finder" gave an idea of the goal of the new technique, this is now missing. In the title it is more clear, as it includes 'selection'. Please reconsider the name again, possibly by including 'selection technique' or a similar term in the full name.

Page 8 line 10: Why hourly? How did you define hourly values? As the average of the whole hour? Or just last part? Is the hour defined at the beginning of the averaging interval or at the end? This is important information and should be included in methods.
Hourly values are used because of the availability of hourly averages as the highest time resolution in the World Data Center for Greenhouse Gases (WDCGG). Therefore in order to keep the format of input data constant for ADMV method, we calculated the average of the whole hour for all data sets. The time stamp for the hourly average was defined as the beginning of the averaging interval.

Moreover, originally ADMV was developed based on 30-min time resolution at the station ZSF. Therefore ADMV method can also handle data with higher time resolution than one hour.
We added in Section 2.1, "In addition, the time stamp was defined at the beginning of the averaging interval."
The argument to use hourly data because that is what is available on WDCGG does not make sense, as for each measurement site there is a co-author included in the manuscript, suggesting that they contributed to the research. Since continuous data is available, it would have been more solid if the authors had analyzed the influence of using hourly averages, versus original data. The reply that the method works on half hourly averages in that sense does not add much information, it would be more informative to compare the outcomes, and conclude what would be the optimal time resolution of the observations. Can the authors add some information on that?

Page 11 line 10: is this referring to the global growth rate? Why would you compare to an average over the last 10 years, and not to the mean growth rate during the same time period as your data sets? You could use e.g. the annual global growth rates from https://www.esrl.noaa.gov/gmd/ccgg/trends/global.html.

Page 13 line 16 'increasing percentage of data': explain that this is the amount of data left after selection and that it is supposed to represent background conditions. In that way it is more clear to read the conclusions without the rest of the text.

Page 13 line 19 and conclusions: 'ADMV is the most restrictive'. This should be put in perspective? Is it good or bad to have a restrictive filter? Would ADMV result in better representative conditions than the other methods?

Page 10 line 27: if VAL is all validated data it can never over- or underestimate CO_2 levels, as they are the actual observations!
We apologize for the misleading wording.
VAL data are validated correct measurements, adjusted to the international standard reference scales and following the Global Atmosphere Watch quality objectives. Nevertheless, due to the different time scales of transportation effects VAL data may contain values from a time period where the well mixing assumption is violated (short time events). Since we referred VAL as validated unselected measurements, the CO_2 levels mentioned here refer to the background level of CO_2 which are supposed to take place at the measurement sites.
If all validated data are used, this would result in an overestimation of the atmospheric CO_2, due to the dominance of anthropogenic activities and no active vegetation in winter. Thus, it indicates that the VAL data are not representative.
Part of the answer here shows that my question was misunderstood: 'If all validated data are used, this would result in an overestimation of the atmospheric CO_2'. This is not true, as validated data are the measurements of atmospheric CO_2 concentrations themselves. Yes, this includes local effects etc. and are not background levels, but it is the actual CO_2 concentration at that

location. Change on page 11 line 21-22 to e.g.: 'indicating that the CO_2 concentrations estimated by VAL are above the background levels'.

=======

In a number of places, I had raised some questions with the intention to also clarify these issues in the manuscript itself. In some cases the authors have just answered these questions in the reply, but have not updated the text accordingly. I would recommend to revise the text especially to include parts of the following replies in the text:

Page 8 line 15: Does it make sense to have different windows at the different levels?
The different start time windows at the different levels result automatically from the ADMV method. It always searches for the optimal start time window based on specific data sets. In our opinion, these are very interesting and valuable results, which reflect to some extent the different characteristics of different measurement sites and also different levels. In this respect, the different time windows at the different sampling levels are results of differences in the dynamics of atmospheric transport.

Page 8 line 19: The results ARE not fully comparable. Does it even make sense to analyze such a short record which does not even give a complete annual cycle?
We agree that the data were not fully comparable because the time period was too short in contrast to the other stations. However, the results showed that for time periods shorter than a full year, the ADMV method was still applicable to the data from the tower measurements, which highlights the flexibility of the approach.

Page 9 line 2: It would make sense to look at the differences by season, as the diurnal cycle is not the same throughout the year. Also, the data sets all cover different time periods, so it is difficult to compare.
We agree that there are differences in diurnal patterns among seasons. We also applied the ADMV method separated by season, i.e., data sets were processed and selected by the ADMV method only during a specific season over the whole time period. However, we found that the start time windows didn't differ significantly (see Supplement S1.1).
Regarding different time periods of the sites we also included data of 2015 for SSL. Now except for HPB, all the measurement sites cover the same time period.

Page 12 line 21: how applicable is the method to other stations?
This would be one of our next research questions and would be tested in the near future.