Ye Yuan et al., Adaptive Baseline Finder, a statistical data selection strategy to identify atmospheric CO$_2$ baseline levels and its application to European elevated mountain stations

Answers to Anonymous Referee #1

The referee's comments are in black, answers are in blue.

Short notice:

According to the suggestion from J. Kim (Referee #2), we changed the name of our method "Adaptive Baseline Finder (ABF)" into "Adaptive Diurnal Minimum Variation (ADMV)". All the names and abbreviations of this method have been adjusted throughout the answer.

Yuan et al. present a data selection method for records of atmospheric CO$_2$ mole fraction observations from mountain locations. Their method, the adaptive baseline finder (ABF), is an interesting one and in that sense worth publishing. However, unfortunately the manuscript in its current form remains very descriptive and does not include clear conclusions on how the community would benefit from using this method in comparison to the other methods to which ABF is compared.

Another main point is that the English should be checked by a language editor, as in several places the manuscript is not written in correct English (e.g. articles are often omitted and commas are used incorrectly).

All in all, I think the authors have done a substantial amount of interesting work, and could be worth publishing after taking into account the specific comments below and especially focus on placing their work in larger context and making more explicit what the use of ABF could contribute to the field.

We would like to thank the referee for the very detailed and constructive comments. Besides answering all the specific comments respectively, we also added more explanations and arguments mainly in the Results and discussion, as well as the Conclusion section.

Specific comments:

Page 1 line 21: 'measuring sites' should be replaced by 'measurement sites', throughout the manuscript.

This was corrected.

Page 1 line 22: Why would this lead to a bias when comparing different stations? Only when the data of these different stations has been selected with different methods.

We apologize for the misleading wording. We replaced "bias" by "reduced compatibility."

Also, it needs to be noted that different stations do have different methods for data selection and data processing due to different station characteristics and measurement conditions (including instrument, etc.). In a way, it is also true that data

compatibility needs to be improved within GAW network. This methodological approach focuses on common structural features of measurement data from mountain stations with the aim of finding a more general solution for the selection of representative measurement data. The aim is to improve the compatibility of the data and to facilitate the conclusion from a point measurement to a larger area.

5

Page 1 line 23: pattern -> patterns

This was corrected.

Page 1 line 24: 'measuring records' -> records of atmospheric CO2 observations

10 This was corrected.

Page 1 line 27: implemented -> included/applied

This was corrected.

15 Page 1 line 27: Among the studied methods, our ABF method …

This was corrected, and the full name was added in the text above (because of the abbreviation used here).

Page 1 line 27: This is very descriptive: lower percentage of selected data; is this 'better'? What does it imply to have less or more data selected?

20 We added by the end of this sentence, "…, which can be understood as a better representation of the lower free troposphere."

Page 1 line 30: STL is not explained

We rewrote the sentence, "The measured time series were analyzed for long-term trend and seasonality by seasonal-

25 trend decomposition technique."

And STL would be explained later in detail in Section 2.4.

Page 2 line 13: what do you mean by correction for interference from other GHG?

We added an example for the potential interference, "such as water vapor."

30

Page 2 line 24: here it would be good to elaborate on the work of Uglietti et al. 2011 (ACP), which is referred to on the same page.

Since the method for calculating the background corrected CO_2 record was the same as Satar et al. (2016), we included the citation of Uglietti et al. (2011) at the end of Section 2.3 for the method description.

Regarding the use of the transport model, we elaborated on the reference to the work of Uglietti et al. (2011) in the same paragraph but some lines further below when the modeling techniques were mentioned. The added sentence is, "Uglietti et al. (2011) focused on the origins of atmospheric CO_2 at Jungfraujoch (Switzerland) by the FLEXible PARTicle dispersion model (FLEXPART)."

5

Page 2 line 29: explain why afternoon values should be excluded.

We added, "…due to the influence of convective upward transport."

Page 3 line 3: MHD flasks are only sampled during restricted base line conditions, so no filtering is applied.

10   We apologize for the misleading wording. It is correct that MHD flasks were only sampled during Restricted Baseline Conditions (RBC), i.e. during periods with specific wind directions and wind velocities requirements.

When citing this paper, we wanted to stress the approaches for aiming on a regional analysis instead of filtering technique. More details can be seen in Section 3 (Data selection) of Sirignano et al. (2010).

For this sentence, we rephrased as, "Threshold limits of 300 ppb for CO and 2000 ppb for CH_4 were defined by

15 Sirignano et al. (2010) to perform a regional analysis of CO_2 data at Lutjewad (the Netherlands) and Mace Head (Ireland)."

Page 3 line 6: Hawaii, USA. Also add Switzerland for JFJ.

These were added.

20 Page 3 line 18: what is REBS?

We rephrased the sentence as, "Ruckstuhl et al. (2012) developed a method based on robust local regression, called Robust Extraction of Baseline Signal, to estimate…"

Page 3 line 23: why do the authors choose to focus on mountain sites only? This should be made more clear in the

25 manuscript.

This work concentrates on finding common properties for the lower free troposphere in ground-based measurements. As the first approach data have been taken from mountain stations due to their remote location with limited anthropogenic influence and increased representativeness. We also focus on mountain sites only because mountain stations have a diurnal variation which results in a daily time window for well-mixed air with a better representation of the lower free

30 troposphere. The explanation in Section 2.2 was improved considerably, which can be seen in the answers below.

Also, we added at the beginning of Section 2.1, "The data have been taken from mountain stations due to their remote location with limited anthropogenic influence and increased representativeness."

Page 4 line 7-9: what do these classifications mean? E.g. "weakly influenced, constant deposition" is not very clear.

We added the explanation at the beginning of this sentence, "Henne et al. (2010) presented a method of categorizing site representativeness based on the influence and variability of population and deposition by the surface fluxes."

Page 4 line 13: did you use hourly data or higher time resolution? This is not clear from this section.

Both time resolutions hourly and half-hourly are available. We used hourly data throughout the work except for the evaluation of the influences of different time resolutions (see Supplement S1.3).

Therefore, we added, "For this study, hourly data were used consistently, unless otherwise indicated."

Page 4 line 15-16: specify which reference is for which station.

We added the station names in the sentence, "Schmidt et al. (2003) for *SSL*, Gilge et al. (2010) for *HPB* and *SNB*, Gomez-Pelaez et al. (2010) for *IZO*, Risius et al. (2015) for *ZSF* and Schibig et al. (2015) for *JFJ*."

Page 4 line 19-20: This sentence is very vague, make more clear what the motivation of this research is.

We rephrased this sentence as, "ADMV is a tool for automatic and systematic analysis of diurnal CO_2 cycles at elevated mountain stations in order to select consecutive time sequences with minimum variation, which can be regarded as representing well-mixed air conditions."

Page 4 line 21-22: This sentence is not clear: what traffic activities are relevant to the mountain sites? And why is vegetation active in the afternoon only? How about respiration?

We rephrased this sentence as, "For example, at ZSF, these can be characterized by anthropogenic CO_2 sources, detectable especially in winter during the day, whereas in summer the convective upwind transport results in a strong impact of air masses with depleted CO_2 concentrations due to photosynthesis at lower altitudes. Plant respiration activities, which may contribute small amounts, are primarily not visible in the convective upwind air masses (which arrive at mountain sites predominantly in the afternoon). Although high elevated mountain stations do not have vegetation in their surroundings, mountain stations at lower altitudes but still in the vegetation zone may be influenced by plant respiration, especially at night."

Page 4 line 22-24: this sentence is not clear. What do you mean by 'which in turn in an effective tool'? What tool?

We apologize for the misleading wording. We rephrased "is an effective tool for data selection" as, "can be used for selecting representative data."

Page 4 line 25: explain PBL and explain the changing degree of entrainment.

PBL was introduced in Section 2.1 (Page 4 line 1), as "planetary boundary layer."

And hereby we rephrased "(e.g., due to changing degree of entrainment of PBL air)" as "because of variations in the dynamics of transport to the site (e.g., Birmili et al., 2009; Herrmann et al., 2015)."

Page 4 line 27-31: The level of English needs to be assessed particularly in these sentences.

We rephrased this part as, "…whereas in winter, significantly longer stable periods occur. In winter, no upwind air masses with depleted CO_2 levels due to photosynthesis of vegetation are recorded. To receive as much representative data as possible, it is desirable to select the time window dynamically. ADMV is constructed to select a subset from the measured data, being best representative for baseline conditions with an adaptive selected time window specific for every day."

Page 5 line 7: What is the time resolution of the data sets?

The time resolution is hourly. Therefore, we added "hourly" in the sentence.

Page 5 line 11-20, and page 6 line 6-22: Revise English especially here, including use of complete sentences including articles ('the') and correct use of commas.

An English proofreading has been done throughout the paper.

This part is shown in the following.

"**Step 1**: Detrending is done by subtracting a 3-day average for each day, including the neighboring two days. It is the shortest possible time window to remove sudden changes in the time series related to the previous and posterior days while preserving the diurnal pattern.

Step 2: The overall mean diurnal variation, $\bar{d}_i$ (i = 0 to 23 h), is calculated from the complete set of detrended data.

Step 3: The standard deviations s_{Δ_j} from the overall mean diurnal variation $\bar{d}_i$ are calculated on a moving window Δ_j ($j = 6\ h$). To be able to place a full set of 24 moving time windows over the overall mean diurnal variation, time windows across midnight (e.g., 6 h from 11 p.m. to 4 a.m. LT) are also included, that is, its first j hours are appended to the end of the 24 h in the overall mean diurnal variation. The time window with the smallest standard deviation is selected as the *start time window*.

Result: The *start time window* $[i_{start}, ..., i_{end}]$."

Page 6 line 26: This is a vague sentence, data only exists on a single day, so why talking about selecting it in 'any day'?

We rephrased the sentence as, "We always label the data as "selected" once it has been selected by ADMV."

Page 7 line 15: photosynthesis starts long before 11 a.m.

We rephrased "potential influences of local photosynthesis" as "transported air influenced by photosynthesis."

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."

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 9 line 4-10: Revise English.

The English proofreading has been done throughout the paper.

And we rephrased this paragraph as,

"With the determined *start time windows*, ADMV selected the data for all stations (see Fig. 3). In addition, we calculated the percentages of ADMV selected data values among all values of the complete datasets for all stations, which are listed in the first column of Table 2. The higher the selection percentage is the more well-mixed air is measured at the station, which is assumed to be a representation of lower free tropospheric conditions. This holds especially for *IZO*. Because of this the greatest amount of accepted data points with 36.2% was found at this station. The sites with intermediate percentages are *JFJ* (22.1%), *SNB* (19.3%), and *ZSF* (14.8%). For the three sampling heights at *HPB*, only 3.2% (50 m), 4.8% (93 m), and 6.2% (131 m) of the data were selected by ADMV. Finally, a similar percentage was found for *SSL* (4.0%), probably due to its higher data variability."

Page 9 line 6-9: But what do these percentages actually mean? This is too descriptive and needs more analysis and perspective.

We added as mentioned above, "The higher the selection percentage is the more well-mixed air is measured at the station, which is assumed to be a representation of lower free tropospheric conditions. This holds especially for *IZO*. Because of this the greatest amount of accepted data points with 36.2% was found at this station."

Page 9 line 10: what is 'major step' and what do the percentages by each step mean?

These different definitions of selection percentages were explained in the Supplement S3.1.

Besides, we rephrased the sentence as, "we additionally calculated selection percentages after completing both the *starting selection* and *adaptive selection* steps mentioned in Section 2.2 (see Supplement S3.1)."

Page 10 line 3: This previous section remains very descriptive. What do the differences between all methods mean, and what is more useful for what type of analysis? This needs more work.

We explained the differences of all methods and the potential reasons in the paragraph above (please see page 9 line 24–33). For ADMV method, the detailed stepwise results of selection percentages were made in Supplement S3.1. For SI and THO methods the major difference is the requirement of consecutive hours. As for MA method, the selection criteria would become too strict for stations with very small data variability (e.g., *IZO*).

On the other hand, this paper focuses more on the mechanism and results of ADMV method. Our intention is to give a clear and detailed instruction on this data selection method and provide options to the users. The advantage of ADMV method can be seen in the Conclusion section. To compare more thoroughly of different data selection methods and present a clear strategy of applying different methods require further researches and are beyond the focus of this paper.

Page 10 line 10: What is the use of comparing growth rates for different time periods? Growth rates are very variable from year to year, so choosing a different period gives different growth rates.

The STL technique has been re-run. The underlying time period is 2010 to 2015 for all sites except for SNB, for which data of 2010 to 2011 are missing. We changed the color for SNB to gray in Table 3 and Table 4 for differentiation.

By comparing growth rates, we aim at showing there are no significant influences on the trend components by these data selection methods.

5

Page 10 line 11: A positive trend in what? In the CO2 concentrations in general?

We rephrased as, "Based on the 95% confidence interval for the slope, a positive trend i.e. increasing CO_2 concentrations are observed."

10   Page 10 line 12: Explain VAL

VAL is validated data and thus delivered to the GAW data bases (Level 2 data). It has been explained in Section 1 (Page 2 line 12).

Page 10 line 12: what differences?

15   We rephrased as, "differences in the mean annual growth rates…"

Page 10 line 15: What do you mean by tendency?

What we can observe from the resulted mean annual growth rates is that, the growth rates resulted after data selection are mostly higher than the ones of validated datasets and always approaching toward the values at station *IZO*.

20   However, this is not statistically significant based on the confidence intervals. That is the reason we meant by tendency.

We rephrased as, "Moreover, the following fact is observed for all sites except for *SSL*."

Page 10 line 18: 2015 had a much higher growth rate compared to the years before, so that also influences the results at SSL. Why not including 2015? It is publically available through ObsPack.

25   Thank you very much for your suggestion. Now the CO_2 data of 2015 for *SSL* are included and all respective analyses have been re-run.

Page 10 section 3.3.1: I do not understand the added value of this paragraph. It should include more details on what was exactly studied and more conclusive remarks instead of only descriptive statements.

30   This paragraph showed the results of the trend components of all data sets at all stations. The comparison of trends based on datasets with and without data selection methods clearly indicated that mean annual growth rates were not significantly different.

To include more details, we added the following arguments in the text.

8

"Compared to unselected data (VAL), the mean annual growth rates based on selected data sets are systematically higher approaching the growth rates at *IZO*. *IZO* can be considered as better representing the lower free tropospheric conditions and agrees well with the mean annual absolute increase during last 10 years (2.21 ppm yr^{-1}) reported by WMO (2017a). The exception at SSL is probably is caused by stronger local influences as a result of its lower elevation. Besides, the confidence intervals of the mean annual growth rates are always smaller after data selection, which improves the precision of trends."

Besides, we also added a general description at the beginning of the section, "The following sections discuss the resulting components obtained by STL, namely the trend component over the observation period, the seasonal component and finally the remainder component."

Page 10 line 20: this is not clearly described (the difference between val and selected data).

As mentioned above, VAL was validated data and thus delivered to the GAW data bases, which was taken as data input for our study. Selected data were resulted data after different data selection methods.

We rephrased, "systematic differences which calculated for validated (VAL) and selected data sets…"

Page 10 line 24: this percentage is given too much precision.

We changed it to "18.9%."

Page 10 line 27: if VAL is all validated data it can never over- or underestimate CO2 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.

Page 11 line 1-7: Very descriptive, add more details and analyses and perspective.

We rewrote this paragraph partly and added more content as the following.

"The magnitude of these delays may be related to mixing features in the lower free troposphere. Rapid changes are usually observed close to sources and sinks, e.g., from anthropogenic and biogenic activities. Thus, the higher the station is above the boundary layer, the later the maxima during the winter can be observed, because of the late response due to inhibited mixing conditions. However, this delay does not occur for the minima during the summer because of the very

effective upward transport and more favorable mixing conditions at that time of year. Consequently, no changes in the seasonal minima are observed at all measurement sites, which is taken as an indicator of enhanced thickness of the mixing layer as good mixing conditions. Taking *ZSF* as an example, Birmili et al. (2009) observed low concentrations of particle number in winter and found it representative for the free tropospheric air by analyzing the annual and diurnal cycles. From

5   spring on, the warmer it gets the higher the PBL goes. The intense vertical atmospheric exchange during summer months results in a daily air mass transport from the boundary layer to reach *ZSF* due to thermal convection (Reiter et al., 1986; Birmili et al., 2009). Thus there are optimal transportation and mixing conditions."

Page 11 line 6: explain in more detail 'thickness of the mixing layer'.

10   Please see the last answer above.

Page 11 line 14: what does 'least standard deviations' mean?

This refers to the variability in the remainder components.

Thus, we rephrased it as, "the smallest standard deviations in the remainder components."

15

Page 11 line 14: we already knew that IZO is least influenced.

IZO was taken as reference station, according to our assumption at the beginning. Therefore, as the results showed the remainder component at *IZO* with the least standard deviation, it supported our assumption that it was the least influenced station among all.

20

Page 11 line 15: what are intermediate results?

We rephrased as, "The three alpine measuring stations (*ZSF*, *SNB* and *JFJ*) exhibit intermediate variability."

Page 11 line 19: could, but why not done?

25   We deleted this sentence as it was beyond the focus of this paper.

Page 11 line 25, figure 6: why is red included in the color scale as those values do not occur? Also the caption of figure 6 contains a lot of information that should be included in the main text as well (pearson corr. matrix etc).

We changed the color scale in Figure 6 from white to blue only.

30   And we added the explanation of the figure in the main text, "The trend and seasonal components of all VAL and selected datasets were firstly compiled, and then Pearson's correlation coefficients were calculated assuming normal distribution of data examined by Anderson Darling test (P < 0.05). The correlation matrices are shown for each type of data sets individually. Data used for correlation were chosen only when available at all stations (2012–2015)".

Page 12 line 1: what does this mean/imply?

We rephrased the sentence as, "The number of insignificant correlations between the station pairings is the greatest for ADMV."

And we also added at the end of the paragraph, "This means that by ADMV, the combination of trend and seasonal components correlate best and the remaining unselected data have the lowest correlation among the methods. If these two criteria are used to separate the representative part of the data from the unrepresentative part, the ADMV method produces the best results".

Page 12 conclusions: should be especially checked for level of English.

An English proofreading has been done throughout the paper.

Page 12 line 7: not all 6 cover this period.

We deleted "from 2010 to 2016."

Page 12 line 7: rewrite, ABF does not select..

We replaced by, "The ADMV selection resulted in…"

Page 12 line 8: growing elevation?

We deleted "growing."

Page 12 line 10: but what does it mean/imply that is the most restrictive? When would you recommend the ABF method?

ADMV is the most restrictive in terms of selection percentage, which selects the least data representative for the lower free troposphere. However, additional indicators should be defined for the selection quality, such as STL and correlation analysis mentioned in the content.

Regarding the results of correlation analysis, we would recommend ADMV for selection of representative well-mixed lower free tropospheric air for the elevated mountain stations mentioned in this study.

Page 12 line 14: what do reduced and delayed mean here?

We rephrased "reduced and delayed influences of CO_2 sources and sinks" as, "with smaller seasonal amplitudes and delayed occurrences of seasonal maxima."

Page 12 line 18-19: what do you mean?

We rewrote this paragraph as the following.

"The presented method ADMV is useful for data selection of atmospheric CO_2 data representative of the lower free troposphere. It requires only data from a single measurement site. It is easily adjustable to the local conditions and it runs automatically. The method can also be applied on historical datasets. The results provide evidence that the proposed ADMV method confers the possibility of selecting data that are representative of CO_2 concentrations of a larger area of the lower free troposphere. This is an elementary prerequisite for application of the method to a large number of different stations and an essential step toward generalization. It directly supports the objective of GAW to extrapolate from a set of point measurements from single stations to a larger representative area or region in the lower free troposphere (WMO, 2017b). In future, there is a need to test whether such results could be used for additional tasks, such as ground calibration of satellite measurements".

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.

Figure 1: add larger map to know which region of the world this is.

This was corrected.

Reference

Birmili, W., Ries, L., Sohmer, R., Anastou, A., Sonntag, A., König, K., and Levin, I.: Feine und ultrafeine Aerosolpartikeln an der GAW-Station Schneefernerhaus/Zugspitze, Gefahrst. Reinhalt. L, 69, 1/2, 31–35, 2009.

Henne, S., Brunner, D., Folini, D., Solberg, S., Klausen, J., and Buchmann, B.: Assessment of parameters describing representativeness of air quality in-situ measurement sites, Atmos. Chem. Phys., 10, 3561–3581, doi:10.5194/acp-10-3561-2010, 2010.

Herrmann, E., Weingartner, E., Henne, S., Vuilleumier, L., Bukowiecki, N., Steinbacher, Coen, F., Collaud Conen, M., Hammer, E., Jurányi, Z., Baltensperger, U., and Gysel, M.: Analysis of long-term aerosol size distribution data from Jungfraujoch with emphasis on free tropospheric conditions, cloud influence, and air mass transport, J. Geophys. Res., 120, doi 10.1002/2015JD023660, 2015.

Reiter, R., Sladkovic, R., Kanter, H.-J., 1986. Concentration of trace gases in the lower troposphere, simultaneously recorded at neighboring mountain stations. Meteorl. Atmos. Phys. 35 (4), 187–200. 10.1007/BF01041811.

Satar, E., Berhanu, T. A., Brunner, D., Henne, S., and Leuenberger, M.: Continuous $CO_2/CH_4/CO$ measurements (2012–2014) at Beromünster tall tower station in Switzerland, Biogeosciences, 13, 2623–2635, doi:10.5194/bg-13-2623-2016, 2016.

5   Sirignano, C., Neubert, R. E., Rödenbeck, C., and J. Meijer, H. A.: Atmospheric oxygen and carbon dioxide observations from two European coastal stations 2000-2005: continental influence, trend changes and APO climatology, Atmospheric Chemistry and Physics, 10, 1599–1615, 2010.

Uglietti, C., Leuenberger, M., and Brunner, D.: European source and sink areas of CO_2 retrieved from Lagrangian transport
10   model interpretation of combined O_2 and CO_2 measurements at the high alpine research station Jungfraujoch, Atmospheric Chemistry and Physics, 11, 8017–8036, doi:10.5194/acp-11-8017-2011, 2011.

WMO: WMO Greenhouse Gas Bulletin, No. 13, ISSN 2078-0796, 30 October, 2017a.

15   WMO: WMO Global Atmosphere Watch (GAW) Implementation Plan: 2016-2023, 2017b.

Ye Yuan et al., Adaptive Baseline Finder, a statistical data selection strategy to identify atmospheric CO₂ baseline levels and its application to European elevated mountain stations

Answers to J. Kim (Referee)

The referee's comments are in black, answers are in blue.

5

Short notice:

According to the suggestion from J. Kim (Referee #2), we changed the name of our method "Adaptive Baseline Finder (ABF)" into "Adaptive Diurnal Minimum Variation (ADMV)". All the names and abbreviations of this method have been adjusted throughout the answer.

10

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

15 measurement community, and as such I suggest that the manuscript be published with some revisions.

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

20 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

25 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

30 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.

Thank you very much for your concerns and elaboration. We fully agree with the elaborate remark of the reviewer. We would like to recall that our study largely focuses on elevated stations, but we define the term "baseline" in the beginning of the introduction (see Page 2 line 8-10) in a broader sense referring to "well-mixed air masses with minimized short-term

14

external influences." We also agree that the name "Baseline Finder" can be misleading, thus after discussion with all co-authors, we decided to change the name into Adaptive Diurnal Minimum Variation (abbreviated: ADMV). Besides, we also added more arguments mainly in the Results and discussion, as well as the Conclusion section.

5   **[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
10   window of a baseline occurrence?

The window of 6 hours and the standard deviation threshold were choices based on empirical visual inspection of the available datasets and on literature review. For this study, we specify that the 6 hours in the *start time window* have to meet two constraints: the standard deviation of measured values less than 0.3 ppm and the missing data in the 6 hours less than 50%. If the requirements are fulfilled, then the data selection will start with the *start time window* for that day. Otherwise, all
15   values in that day with the *start time window* are labelled as "unselected."

The length of 6 hours was considered as a reasonable time length to determine whether the measured air masses are well-mixed and thus most representative, largely following the approach of Pales and Keeling (1965) (as mentioned in the introduction, page 3 line 15). More detailed information can also be found in Levin et al. (1995) and Brailsford et al. (2012). Shorter time windows will lead to less robust statistics and thus more variable standard deviations. Thus the selection
20   procedure might rely on less representative data and risk of accepting the wrong *start time window* increases.

However it is worth noting that 6 hours were only chosen for this study. It can be variably adjusted by users according to their measurement sites.

- 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.
25   We agree that, the tendency is not statistically significant based on the confidence intervals.
We rephrased as, "Moreover, the following fact is observed for all sites except for *SSL*."

- 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
30   season. What criteria was used to derive a representative window for the whole period?

We apologize for the misleading wording. There are different terms regarding time window in our ADMV method: *start time window* and *selected time window*.

The *start time window* is different based on different running frequency of ADMV. It is the result from the first part of ADMV data selection – *starting selection* (see Section 2.2.1). For this study we applied overall frequency, indicating the

start time window for the full data set (spanning years) is the same. Figure 2 shows the *start time windows* at each measurement site. Theoretically we can also apply ADMV by yearly, seasonally or monthly depending on the requirements.

And for calculation, the *start time window* is derived from the diurnal cycles which are the mean over the detrended data. The criterion for selecting such window is the least variable time period (6 hours) during the night time (6 p. m. to 5 a. m. LT), due to the focus on mountain stations for this study. More details can be found in Section 2.2.1 (Page 5 line 6 onwards).

On the other hand, the *selected time window* represents the selected data sets from ADMV data selection. It is the result from the second part of ADMV data selection – *adaptive selection* (see Section 2.2.2). After both *forward* and *backward adaptive selection*, the *selected time window* result is different for each individual day.

- 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.

Based on our results, it is very likely that the lower free troposphere will respond in a delayed manner to CO_2 concentration changes by effective sources and sinks on the ground. The lower free troposphere acts like an atmospheric "memory" with delayed reaction.

Regarding anthropogenic emissions in summer, we agree that they are certainly small enough to be masked in the drawdown. One example can be found in Oney et al. (2017) for a comparison of biospheric and anthropogenic contribution from CO_2 observations at a tall tower station on the Swiss Plateau, which is the most populated and most industrialized region in Switzerland. Both Fig. 2 and Fig 3 in Oney et al. (2017) show the difference in anthropogenic and biospheric signals, especially in summer time. The magnitudes of anthropogenic contributions are much smaller than the biospheric ones.

- 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.

The adaptiveness of the algorithm is indeed the ability to select values in different time windows for every individual day. It is the ability to adapt the selection on a daily basis in order to receive a maximum amount of representative data (For more details please see Section 2.2.1 and 2.2.2). One of the results is shown in Section 3.1 for the different *start time* windows among stations.

16

Moreover, the ADMV data selection can also be processed for each season individually (with individual settings manually). A comparison of the resulting *start time windows* between overall and seasonal running frequencies can be found in Supplement S1.1.

5 **[Minor comments]**

- Page 4, ln 10: "At last", change to "Finally"?
  This was corrected.

10 - 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."
  This was corrected.

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

- 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.
  We combined Step 3 into Step 2 with the "If/Else" step, and changed the following step numbers accordingly.

20
- 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?
  The percentages refer to the ratio of the selected data values in the data values of the complete data sets. The selection percentage regarding the selected time windows and the selected days can be found in Supplement S3.1 in detail.
25 Thus for clarification, we rephrased "data in all data for all stations" as, "data values in all values of the complete data sets".

Reference

30 Brailsford, G. W.; Stephens, B. B.; Gomez, A. J.; Riedel, K.; Mikaloff Fletcher, S. E.; Nichol, S. E.; Manning, M. R. (2012): Long-term continuous atmospheric CO_2 measurements at Baring Head, New Zealand. In Atmospheric Measurement Techniques 5 (12), pp. 3109–3117. DOI: 10.5194/amt-5-3109-2012.

Levin, I.; Graul, R.; Trivett, N. B. A. (1995): Long-term observations of atmospheric CO_2 and carbon isotopes at continental sites in Germany. In Tellus 47B, pp. 23–34.

Oney, B., Gruber, N., Henne, S., Leuenberger, M., & Brunner, D. (2017). A CO-based method to determine the regional biospheric signal in atmospheric CO_2. Tellus B: Chemical and Physical Meteorology, 69(1), 1353388.

Pales, J. C. and Keeling, C. D.: The Concentration of Atmospheric Carbon Dioxide in Hawaii, J. Geophys. Res., 70, 6053–6076, doi:10.1029/JZ070i024p06053, 1965.

5

Adaptive selection of Diurnal Minimum Variation: ~~Baseline Finder,~~ a statistical ~~data selection~~ strategy to obtain representative ~~identify~~ atmospheric CO_2 data ~~baseline levels~~ and its application to European elevated mountain stations

Ye Yuan[1], Ludwig Ries[2], Hannes Petermeier[3], Martin Steinbacher[4], Angel J. Gómez-Peláez[5], Markus C. Leuenberger[6], Marcus Schumacher[7], Thomas Trickl[8], Cedric Couret[2], Frank Meinhardt[9], Annette Menzel[1,10]

[1]Department of Ecology and Ecosystem Management, Technische Universität München, Freising, Germany
[2]German Environment Agency (UBA), Zugspitze, Germany
[3]Department of Mathematics, Technische Universität München, Freising, Germany
[4]Empa, Laboratory for Air Pollution/Environmental Technology, Dübendorf, Switzerland
[5]Izaña Atmospheric Research Center, Meteorological State Agency of Spain (AEMET), Santa Cruz de Tenerife, Spain
[6]Climate and Environmental Physics Division, Physics Institute and Oeschger Centre for Climate Change Research, University of Bern, Bern, Switzerland
[7]Meteorological Observatory Hohenpeissenberg, Deutscher Wetterdienst (DWD), Hohenpeissenberg, Germany
[8]Institute of Meteorology and Climate Research, Atmospheric Environmental Research (IMK-IFU), Karlsruhe Institute of Technology ~~(KIT)~~, Garmisch-Partenkirchen, Germany
[9]German Environment Agency (UBA), Schauinsland, Germany
[10]Institute for Advanced Study, Technische Universität München, Garching, Germany

Correspondence to: Ye Yuan (yuan@wzw.tum.de)

Abstract. Critical data selection is essential for determining representative baseline levels of atmospheric trace gas<u>es</u> ~~measurements~~ even at remote <u>measurement sites</u> ~~measuring sites~~. Different data selection techniques have been used around the world<u>,</u> which could potentially lead to <u>reduced compatibility</u> ~~bias~~ when comparing data from different stations. This paper presents a novel statistical data selection method <u>named Adaptive Diurnal Minimum Variation (ADMV)</u> based on CO_2 diurnal <u>patterns</u> ~~pattern~~ occurring typically at high elevated mountain stations. Its capability and applicability <u>were</u> ~~was~~ studied <u>on records of atmospheric CO_2 observations</u> ~~for atmospheric measuring records of CO_2 from 2010 to 2016~~ at six Global Atmosphere Watch ~~(GAW)~~ stations in Europe, namely<u>,</u> Zugspitze-Schneefernerhaus (Germany), Sonnblick (Austria), Jungfraujoch (Switzerland), Izaña (Spain), Schauinsland (Germany)<u>,</u> and Hohenpeissenberg (Germany). Three other frequently applied statistical data selection methods were <u>included</u> ~~implemented~~ for comparison. <u>Among the studied methods, our ADMV method</u> ~~Among all selection routines, the new method named Adaptive Baseline Finder (ABF)~~ resulted in lower selection percentages with lower maxima during winter and higher minima during summer in the selected data<u>, which can be understood as a better representation of the lower free troposphere. The measured time series were analyzed for long-term trends and seasonality by seasonal-trend decomposition technique.</u> ~~To investigate long-term trend and seasonality, seasonal decomposition technique STL was applied.~~ Compared with ~~the~~ unselected data, mean annual growth rates of all selected data-sets were not significantly different<u>,</u> except for <u>the data recorded for</u> Schauinsland. However, clear differences

were found in the annual amplitudes as well as ~~for~~ the seasonal time structure. Based on correlation analysis, results obtained by ADMV ~~ABF~~ selection showed a better representation of the lower free tropospheric conditions.

1 Introduction

Continuous in situ measurements of greenhouse gases (GHGs) at remote locations have been established since 1958
5 (Keeling, 1960). Knowledge of background atmospheric GHG concentrations is ~~a~~ key to ~~the~~ understanding ~~of~~ the global carbon cycle and its effect on climate, as well as the GHG responses to a changing climate. A crucial issue when using data from remote stations remains the identification of time periods that are representative of ~~for~~ larger spatial areas and their differentiation ~~separation~~ from periods influenced by local and regional pollution. If these two regimes are well disaggregated, the available data-sets can represent more reliable information about long-term changes of undisturbed
10 atmospheric GHG levels, or be used to investigate local and regional GHG sources and sinks when specifically analyzing the deviations from the baseline conditions. In this study, the baseline conditions refer to a selected subset of data from the validated data-set, representing well-mixed air masses with minimized short-term external influences (Elliott, 1989; Calvert, 1990; Balzani Lööv et al., 2008; Chambers et al., 2016).

Measurement results depend on sampling methods, analytical instrumentation, and data processing. Validated data (labeled
15 ~~labelled~~ as VAL in this study to differentiate from the selected data) are usually obtained after signal correction, for example, due to interferences from other GHGs such as water vapor, calibration accounting for sensitivity changes of the analyzer, and validation based on plausibility checks. Data selection starts with validated data and identifies in subsequent steps a final subset of the validated data-set based on pre-defined criteria for ~~a~~ specific qualities ~~quality~~ such as representativeness. With a particular focus on CO_2 in this study, it is commonly accepted that data selection methods can be categorized into
20 meteorological, tracer, and statistical selection methods (Ruckstuhl et al., 2012; Fang et al., 2015).

Meteorological data selection makes use of the meteorological information at the measurement sites ~~measuring sites~~, which provides valuable information about the surrounding environment as well as air mass transport (Carnuth and Trickl, 2000; Carnuth et al., 2002). Forrer et al. (2000), Zellweger et al. (2003), and Kaiser et al. (2007) ~~studied~~ intensively studied the relationship between measured trace gases (such as O_3, CO, and NO_x) and meteorological processes at Zugspitze,
25 Jungfraujoch, Sonnblick, and Hohenpeissenberg. For CO_2, the most common parameters applied in the literature are wind speed and wind direction. They can provide information on critical variations at stations with sources and sinks in their vicinity, while these parameters are less suited at stations in largely pristine environments. For example, Lowe et al. (1979) performed a pre-selection on the CO_2 record at Baring Head (New Zealand) during the southerly wind period only (clean marine air). In addition, Massen and Beck (2011) found that the CO_2 versus wind speed plot can be valuable for baseline
30 CO_2 estimation without a local influence of continental measurements. Besides, fixed time window selection has been widely used, by selecting data in a certain time interval of the day based on local and mesoscale mechanisms of air mass transportation. For selecting well-mixed air at elevated mountain sites, night-time ~~night time~~ is usually chosen with a special

focus on the exclusion of afternoon periods due to the influence of convective upward transport (Bacastow et al., 1985). Brooks et al. (2012) also limited their mountaintop CO_2 results in the Rocky Mountains (USA) by "time-of-day" from 0 a.m. till 4 a.m. local time (LT) ~~for a more likely~~ to increase the likelihood of sampling the free tropospheric environment ~~sampled~~ at the station. Apart from this, modeling techniques such as backward trajectories are very helpful for analyzing in detail the

5   origins and transport processes of air masses arriving at the station (Cui et al., 2011~~; Uglietti et al., 2011~~). Uglietti et al. (2011) focused on the origins of atmospheric CO_2 at Jungfraujoch (Switzerland) by the FLEXible PARTicle dispersion model (FLEXPART). Using tracers, data selection can be performed by investigating the correlations between the air components of interest. Many tracers have been applied and compared with CO_2. Threshold limits of 300 ppb for CO and 2000 ppb for CH_4 were defined by Sirignano et al. (2010) to perform a regional analysis of ~~filter the analyzed~~ CO_2 data at

10   Lutjewad (the Netherlands) and Mace Head (Ireland) ~~by Sirignano et al. (2010)~~. Similar approaches with black carbon and CH_4 were performed by Fang et al. (2015) at Lin'an (China). Moreover, Chambers et al. (2016) developed and applied a data selection technique to identify baseline air masses using atmospheric radon measurements at the stations Cape Grim (Australia), Mauna Loa (Hawaii, USA), and Jungfraujoch (Switzerland).

Unlike most of the methods mentioned above, which require additional data or advanced transport modeling ~~modelling~~,

15   statistical data selection only relies on the time series of interest and typically investigates the variability of signal. It is usually assumed that the most representative CO_2 data are found during well-mixed conditions revealing small variations in time (Peterson et al., 1982) and in space (Sepúlveda et al., 2014). For continuous measurements, it is possible to investigate within-hour and hour-to-hour variability in the data-sets. The within-hour variability is often expressed as the standard deviation of the measured data within 1 h ~~one hour~~. The hour-to-hour variability compares the differences between hourly

20   averaged concentrations either during a certain time period, or from one hour to the next. Pales and Keeling (1965) marked ambient data as "variable" when the within-hour variability for the air sample is significantly larger than the within-hour variability for the reference gas. Consequently, they only selected CO_2 data in "steady" conditions for 6 h ~~hours~~ or more. Besides, Peterson et al. (1982) also rejected sampled CO_2 data values for adjacent hours when the hour-to-hour variability exceeded 0.25 ppm. Thoning et al. (1989) combined these two strategies using an iterative approach by selecting data

25   according to deviations of daily averages from a spline curve fit. Ruckstuhl et al. (2012) developed a method based on robust local regression, called Robust Extraction of Baseline Signal, ~~chose a different statistical method based on robust local regression (REBS)~~ to estimate the baseline curves generalized for atmospheric compounds, which is available in the R package IDPmisc (Locher and Ruckstuhl, 2012).

The present study focuses on the comparison of results from statistical data selection methods and the Adaptive Diurnal

30   Minimum Variation (ADMV) ~~Baseline Finder (ABF)~~. ~~The ABF~~ The ADMV is seen as a possible alternative to already known data selection methods as discussed above. By applying ADMV ~~ABF~~ to the atmospheric CO_2 records from six European elevated mountain stations, the selection results are compared with those derived from three other statistical data selection methods. To investigate the potential influences of the data selection method on trend and seasonality, further

analyses focus on the decomposition of validated and selected data~~-~~sets in trend and seasonal components. Finally, differences between ADMV ~~ABF~~ and other data selection methods were assessed by correlation analysis.

2 Methods

2.1 CO$_2$ measurements at elevated European sites

5   CO$_2$ measurements from six selected European mountain stations (see Fig. 1) within the Global Atmosphere Watch (GAW) network were used to test the data selection algorithms. The data have been taken from mountain stations due to their remote location, which results in them being subjected to limited anthropogenic influence and having increased representativeness. Three ~~high alpine~~ high-alpine measurement sites ~~measuring sites~~ were included: Zugspitze-Schneefernerhaus (*ZSF*, DE, 47°~~25'~~25′ N, 10°~~59'~~59′ E, 2670 m a.s.l.), Jungfraujoch (*JFJ*, CH, 46°~~33'~~33′ N, 7°~~59'~~59′ E, 3580 m a.s.l.), and Sonnblick

10   (*SNB*, AT, 47°~~03'~~03′ N, 12°~~57'~~57′ E, 3106 m a.s.l.). They are often above the planetary boundary layer (PBL)~~,~~ and thus exposed to free and assumed clean lower tropospheric air masses, but are periodically influenced by regional emissions from lower altitudes. Additionally, to test data selection for a less remote environment, CO$_2$ measurements from Schauinsland (*SSL*, DE, 47°~~55'~~55′ N, 7°~~55'~~55′ E, 1205 m a.s.l.) at a clearly lower elevation in the mid-range Black Forest were investigated. Data selection was also applied to three recently started CO$_2$ time series from different sampling heights above

15   ground at a tall tower at the Hohenpeissenberg observatory (*HPB*, DE, 47°~~63'~~63′ N, 11°~~01'~~01′ E, 934 m a.s.l.), located in the northern ~~Northern~~ foothills of the Alps. Henne et al. (2010) presented a method of categorizing site representativeness based on the influence and variability of population and deposition by the surface fluxes. ~~Based on the station categorization by Henne et al. (2010),~~ *JFJ* and *SNB* are classified as "mostly remote~~",~~," while *ZSF* is considered as "weakly influenced, constant deposition~~",~~," and *SSL* and *HPB* are considered as "rural" (Henne et al., 2010). Finally, ~~At last,~~ Izaña station on

20   Tenerife Island in the North Atlantic (*IZO*, ES, 28°~~19'~~19′ N, 16°~~30'~~30′ W, 2373 m a.s.l.) was chosen as a reference for comparison due to its location above the subtropical temperature inversion layer, which means that ~~makes~~ the station is ~~to be~~ rarely affected by any local or regional CO$_2$ sources and sinks (Gomez-Pelaez et al., 2013).

For this study, hourly data were used consistently, unless otherwise indicated. The validated CO$_2$ hourly averages from all stations were downloaded from ~~are available at~~ the World Data Centre for Greenhouse Gases (~~WDCGG;~~

25   http://ds.data.jma.go.jp/gmd/wdcgg/). Data with higher time resolution required for this study were provided directly by the station investigators. In addition, the time stamp was defined as the beginning of the averaging interval. Descriptions of the sampling elevation and time period of available data are given in Table 1. Further information on each station can be found in the report by Schmidt et al. (2003) for *SSL*, Gilge et al. (2010) for *HPB* and *SNB*, Gomez-Pelaez et al. (2010) for *IZO*, Risius et al. (2015) for *ZSF*, and Schibig et al. (2015) for *JFJ*. Practical data selections and analyses in this study were ~~paper~~

30   ~~have been~~ performed in the R Statistical Environment (R Core Team, 2017).

2.2 ADMV ~~Adaptive Baseline Finder (ABF)~~

ADMV is a tool for automated and systematic analysis of diurnal CO_2 cycles at elevated mountain stations in order to select consecutive time sequences with minimum variation, which can be regarded as representing well-mixed air conditions. ~~ABF is motivated by the investigation of CO_2 diurnal cycles at elevated mountain stations that are in the free troposphere or~~

5   ~~representing a regional background during certain periods.~~ Even though such measurement sites ~~measuring sites~~ are remotely located, the CO_2 levels are still influenced by local sources and sinks. For example, at ZSF, these can be characterized by anthropogenic CO_2 sources, detectable especially in winter during the day, whereas in summer the convective upwind transport results in a strong impact of air masses with depleted CO_2 concentrations due to photosynthesis at lower altitudes. Plant respiration activities, which may contribute small amounts, are primarily not visible in the convective upwind air

10 masses (which arrive at mountain sites predominantly in the afternoon). Although high elevated mountain stations do not have vegetation in their surroundings, mountain stations at lower altitudes but still in the vegetation zone may be influenced by plant respiration, especially at night. ~~For instance, both traffic activities starting in the morning hours, and vegetation activities in the afternoon hours contribute to the measured CO_2 signal at *ZSF* on a diurnal basis.~~ This points out the importance of finding a certain diurnal time window representing the most stable and representative CO_2 level, which in turn

15 can be used for selecting representative data. ~~is an effective tool for data selection.~~ However, the duration of this time window during the day varies with the season ~~seasons~~ and from day to day because of variations in the dynamics of transport to the site (e.g., Birmili et al., 2009; Herrmann et al., 2015). ~~(e.g., due to changing degree of entrainment of PBL air).~~ In summer, larger variabilities in the CO_2 signal are observed due to more prevalent convective boundary layer air mass injections influencing the diurnal pattern, resulting in shorter periods of stable ~~condition.~~ conditions, whereas in ~~In~~ winter,

20 significantly ~~much~~ longer stable periods occur ~~can be found.~~ In winter, no upwind air masses with depleted CO_2 levels due to photosynthesis by vegetation are recorded. ~~No upwind air masses with depleted CO_2 levels by photosynthesis of vegetation like in summer are recorded.~~ To receive ~~capture~~ as much ~~many~~ representative data as possible, it is desirable to select the time window dynamically. ADMV ~~ABF~~ is constructed to ~~fulfil these requirements for an automated data selection method that~~ selects a subset from ~~of~~ the measured ~~validated~~ data, being best representative for baseline conditions ~~at the measuring~~

25 ~~station~~ with an adaptive selected time window specific for every day.

The algorithm is based on two basic assumptions. First, ~~Firstly,~~ air masses measured at altitude stations contain well-mixed air, closest to baseline levels, within a certain time window of several hours during the day. For the elevated mountain stations discussed in this paper, this time interval is around midnight. Different diurnal patterns are apparent at each station, so ~~that~~ the selected time window should be adjusted accordingly. Second, ~~Secondly,~~ it is assumed that the real baseline

30 conditions are not subject to local influences and thus represent air masses originating ~~originated~~ only from the uninfluenced lower free troposphere. This ~~It~~ indicates that the variability of the measured CO_2 signal should be minimal within this selected time window. The methodological steps of ADMV ~~ABF~~ are introduced in detail below in the two sections *starting selection* and *adaptive selection*.

2.2.1 *Starting selection*

For a given validated <u>hourly</u> data~~-~~set, <u>ADMV</u> ~~ABF~~ starts data selection by finding a *start time window* for all days. The standardized selection procedure for the *start time window* results from site-specific parameters. This time interval is set as the most stable period from the diurnal variation. The step is referred to as *starting selection*. It begins by analyzing the mean diurnal cycle of the data input.

Step 1: Detrending is done by subtracting a 3-day average for each day, including the neighboring two days. It is the shortest possible time window to remove sudden changes in the time series related to the previous and posterior days <u>while preserving</u>~~, but preserves of~~ the diurnal pattern.

Step 2: The overall mean diurnal variation, $\bar{d}_i$ ($i = 0\ to\ 23\ h$), <u>is</u> ~~are~~ calculated from the complete set of detrended data.

Step 3: The standard deviations s_{Δ_j} from the overall mean diurnal variation $\bar{d}_i$ are calculated on a moving window Δ_j ($j = 6\ h$). To be able to place a full set of 24 moving time windows over the overall mean diurnal variation, time windows across midnight (e.g.<u>,</u> 6 <u>h</u> ~~hours~~ from 11 p.m. to 4 a.m. LT) are also included, <u>that is,</u> ~~i.e.~~ its first j hours are appended to the end of the 24 <u>h</u> ~~hours~~ in the overall mean diurnal variation. The time window with the smallest standard deviation is selected as the *start time window*.

Result: The *start time window* $[i_{start}, \ldots, i_{end}]$.

With the focus on elevated mountain stations, *starting selection* is purposely designed with the moving window Δ_j of 6 <u>h</u> ~~hours~~, and the starting hour i_{start} to be between 6 p.m. and 5 a.m. LT for this study. For other stations with possibly different diurnal patterns, *starting selection* can be adjusted accordingly. For instance, at urban stations or stations completely within the continental PBL, the *start time window* can be chosen based on the best mixing conditions<u>, which</u> ~~that~~ often occur in the afternoon with a shorter moving window, when the PBL reaches its maximum depth after "ingesting" free<u>-</u>tropospheric air during its growth. Being aware that calculating the *start time window* from overall data could differ from the *start time windows* calculated by <u>season</u> ~~seasons~~, the overall generated *start time windows* have been compared with seasonal generated *start time windows* for <u>highly</u> ~~high~~ elevated mountain stations (see Supplement S1.1). Because these differences are mostly minimal to moderate and this work aims <u>at</u> ~~to~~ a methodical comparison under identical conditions, ~~a~~ constant generation of *start time window* from overall data has been chosen.

2.2.2 *Adaptive selection*

The second part *adaptive selection* is designed to determine the most suitable time window for each day, based on the data variability. The length of the *start time window* is adapted (expanding only) in both directions in time. *Adaptive selection* is <u>performed</u> ~~done~~ on a daily basis, starting with the first day of the given data~~-~~set. The following steps only describe the *forward adaptive selection*. <u>ADMV</u> ~~ABF~~ runs the *backward adaptive selection* in an analogous manner <u>but going back in time</u> ~~reverse in time~~.

Step 1: The mean molar ~~mole~~ fraction $\bar{x}_i$, standard deviation s_i, and the proportion of missing values $\pi_{missing}$ are calculated from data in the *start time window* $[i_{start}, \dots, i_{end}]$.

Step 2: If $s_i \leq 0.3$ ppm (CO_2) and $\pi_{missing} \leq 0.5$, ADMV ~~ABF~~ continues to advance in time to examine whether the next data point x_f can be included in the selected time window W with $f = i_{end} + 1$. Otherwise, it is considered that the *start time window* does not fulfill the assumptions. No data are ~~is~~ selected for this day. One should proceed on ~~Go~~ to **Next Day**.

~~**Step 3**: ABF advances in time to examine whether the next data point x_f can be included in the selected time window W with $f = i_{end} + 1$.~~

Step ~~3~~ 4: The absolute difference between x_f and $\bar{x}_i$ is calculated, and the following threshold criterion is applied: $\left| x_f - \bar{x}_i \right| \leq \kappa \cdot s_i$, where κ is the threshold parameter. If this criterion holds, x_f is included in W and ADMV ~~ABF~~ continues. Otherwise, ADMV ~~ABF~~ is finished only with the *start time window* for this day and one should proceed on ~~go~~ to **Next Day**.

Step ~~4~~ 5: Mean $\bar{x}_W$ and standard deviation s_W for the new selected time window W are calculated. If $s_W \leq 0.3$ ppm (CO_2), ADMV ~~ABF~~ continues with the next data point x_f while $f = f + 1$. Otherwise, ADMV ~~ABF~~ is finished with the previous selected time window and one should proceed on ~~go~~ to **Next Day**.

Step ~~5~~ 6: The new absolute difference between x_f and $\bar{x}_W$ is calculated, as well as the new threshold criteria. If condition $\left| x_f - \bar{x}_W \right| \leq \kappa \cdot s_W$ holds, x_f is included in W and ADMV ~~ABF~~ goes back to **Step ~~4~~ 5**. Otherwise, ADMV ~~ABF~~ is finished for this day and one should proceed on ~~go~~ to **Next Day**.

When selection for all days is finished, ADMV ~~ABF~~ continues with *backward adaptive selection*. Afterwards, one should proceed on ~~go~~ to **Result**.

Result: ~~The~~ This is the final selected time window, which is a combination of $W_{forward}$ and $W_{backward}$ for the referring day.

The following limitations of the forward and backward expansions of the time window should be considered. ADMV ~~ABF~~ always runs for no longer than 24 h ~~hours~~ including the *start time window*, namely, ~~i.e.~~ $f \leq 24 \cdot tr$, where tr is the time resolution in data points per hour of the input data. ~~Sometimes this~~ This sometimes results in an overlap of "selected" and "unselected" data for two consecutive days. We always label ~~consider~~ the data as "selected" once it has been selected by ADMV ~~ABF in any day~~. The threshold parameter κ is the controlling factor of ADMV ~~ABF~~ for the length of the selected time window. As κ increases, the length of the selected time window becomes larger. The value of 2 was chosen heuristically for this study as a compromise between selecting as many data points as possible and achieving the least data variability. Similar values of sensitivity-controlling parameters in other data selection methods can be found (Thoning et al., 1989; Sirignano et al., 2010; Uglietti et al., 2011; Satar et al., 2016). In **Step 2**, values of 0.3 ppm and 0.5 indicate the threshold ~~value~~ values for s_i and $\pi_{missing}$. We denote them as $s_{i,threshold}$ and $\pi_{missing,threshold}$. It has been shown that less remote stations at lower altitudes require a larger value than 0.3 ppm because of different mixing conditions. When performing ADMV ~~ABF~~ data selection at lower sites such as *HPB* and *SSL*, we recommend a higher $s_{i,threshold}$, such as 1.0 ppm instead of 0.3 ppm. However, throughout this study, we used the described parameter setting (0.3 ppm) for a methodical

inter-comparison of selection methods at all stations. Potential influences of these parameter sizes ($s_{i,threshold}$ and tr) are discussed in ~~Supplement~~ Supplements S1.2 and S1.3.

2.3 Other statistical data selection methods for comparison

We compared ADMV ~~ABF~~ with three statistical data selection methods. The first method named SI is based on "steady
5 intervals" (Lowe et al., 1979; Stephens et al., 2013). Steady intervals, which are considered as baseline conditions, are defined by a standard deviation being lower than or equal ~~than~~ to 0.3 ppm for 6 or more consecutive hours.

Second~~ly~~, we ~~took~~ adopted a method applied by NOAA ESRL, which originated from Thoning et al. (1989). This selection routine has been applied specifically for measurements of background CO_2 levels at Mauna Loa. This method (~~labelled~~ referred to as THO) was applied as described ~~in~~ on the website
10 (http://www.esrl.noaa.gov/gmd/ccgg/about/co2_measurements.html~~)~~. The first step of THO examines the within-hour variability by selecting hours with hourly standard deviation less than 0.3 ppm. Second, ~~ly~~ it computes hourly averages, and checks the hour-to-hour variability by retaining any two consecutive hourly values where the hour-to-hour difference is less than 0.25 ppm. The last step is based on the diurnal pattern (similar to ADMV ~~ABF~~), by excluding data from 11 a.m. to 7 p.m. LT due to transported air influenced by photosynthesis ~~potential influences of local photosynthesis~~.
15 The last method compared is a moving average technique (MA). A moving time window of 30 days and a threshold criterion of two standard deviations from the moving averages were applied to discard outliers. Afterwards, new moving averages and new threshold criteria were calculated for data exclusion. This step is repeated until no more outliers can be found. A more detailed description can be found in Uglietti et al. (2011) and Satar et al. (2016).

2.4 Seasonal-trend decomposition STL

20 To analyze and compare the selected results from different data selection methods as well as the original validated data-sets, we applied the seasonal-trend decomposition technique based on locally weighted regression smoothing (Loess), named STL (Cleveland, 1979; Cleveland et al., 1990). STL has been widely used on measurements of atmospheric CO_2 and other trace gases ~~measurements~~ (Cleveland et al., 1983; Carslaw, 2005; Brailsford et al., 2012; Hernández-Paniagua et al., 2015; Pickers and Manning, 2015). It decomposes a time series of interest into a trend component T, a seasonal component S, and a
25 remainder component R, which allows separate~~ly~~ detailed analyses and comparisons of trend and seasonality. Two recursive procedures are included in the STL technique: an inner loop where seasonal and trend smoothing based on Loess are performed and updated in each pass, and an outer loop ~~which~~ that computes the robustness weights to reduce the influences of extreme values for the next run of the inner loop (Cleveland et al., 1990).

For this study, we used the implemented function `stl` in R (R Core Team, 2017). ~~Due~~ Owing to limitations of function
30 `stl`, full time coverage of monthly data is needed in order to reduce the risk of large time gaps or unequal spacing (Pickers and Manning, 2015). All data results were first~~ly~~ aggregated to monthly averages. Then, missing data were substituted by

linear interpolation, using R function `na.approx` (Zeileis and Grothendieck, 2005). For the application of STL, two parameters need to be specified, which are the seasonal smoothing parameter $n_{(s)}$ (*s.window* in function `stl`) and the trend smoothing parameter $n_{(t)}$ (*t.window* in function `stl`). As $n_{(s)}$ and $n_{(t)}$ increase, the seasonal and trend components get smoother (Cleveland et al., 1990). For ~~an optimum~~ optimal compatibility in this study, the same parameters were chosen for

5   all stations as $n_{(s)} = 7$ and $n_{(t)} = 23$, based on the recommendation of Cleveland et al. (1990). Another parameter ~~parameters~~ combination of $n_{(s)} = 5$ and $n_{(t)} = 25$ was also tested according to Pickers and Manning (2015), but with no significant differences in results.

3 Results and discussion

3.1 *Start time window*

10   ADMV ~~ABF~~ was applied to the validated hourly averages from all six stations with the parameter settings as described above. The detrended mean diurnal cycles were obtained with the *start time window* for each station by *starting selection* (see Fig. 2, for conventional mean diurnal plots see Supplement S2). The observed differences in the *start time windows*, as well as in the widths of the confidence intervals (gray ~~grey~~ shades in Fig. 2), can be explained by different site environments and thus differing data variabilities. The first subplot column (*HPB50*, *HPB93*, and *HPB131*), ~~represents~~ representing the

15   three sampling heights at *HPB*, shows similar detrended diurnal patterns with similar *start time windows*. The decreasing amplitude with increasing sampling height indicates that the higher the sampling inlet is above the ground, the less it is affected ~~it is~~ by the local surface fluxes. The three *start time windows* suggest that the most stable period at *HPB* occurs during the last few hours of a day, ~~and~~ also including midnight. However, in contrast to all other stations covering at least a full year, *HPB* data are only from September of 2015 to June of 2016. The results may ~~be~~ not be fully comparable.

20   Regarding the second subplot column (*SSL*, *SNB*, and *IZO*), the *start time windows* can be found from midnight on or later in the morning. The *start time window* for *SSL* encompasses its diurnal maximum, indicating that data variability is considerably smaller in the early morning than in the afternoon because of its vicinity to the Black Forest region, which has strong influence due to local photosynthetic activities (Schmidt et al., 2003). A similar diurnal pattern can be found at *SNB*. The influence of CO_2 sources is not as prominent as the effect of distant CO_2 sinks, since it is situated at the single summit

25   peak of Hoher Sonnblick only surrounded by mountains and glaciers, with a ~~negligible~~ negligibly small number of tourists, thus anthropogenic activities are minimal. *IZO* is a special case, since it is located on a remote mountain plateau on the Island of Tenerife above the strong subtropical temperature inversion layer. Even though the *start time window* is limited to 6 h ~~six hours~~, *IZO* presents ~~a very~~ an ideal mean diurnal cycle for data selection from a potentially much longer time window.

30   ~~On~~ In the right subplot column, both *ZSF* and *JFJ* find their *start time windows* around midnight (with more hours after midnight). *ZSF* shows higher diurnal CO_2 amplitude ~~compared with~~ than *JFJ*, but ~~both~~ the two sites show similar diurnal

patterns. For the choice of the *start time window* from the mean diurnal variation, relatively close or even local anthropogenic sources may influence the CO_2 at these two stations, possibly due to touristic influences.

3.2 Selection percentage

With the determined *start time windows*, ADMV ~~ABF~~ selected the data for all stations (see Fig. 3). ~~And~~ In addition, we calculated the percentages of ADMV ~~ABF~~ selected data ~~in all data~~ values among all values of the complete datasets for all stations, which are listed in the first column of Table 2. The higher the selection percentage is the more well-mixed air is measured at the station, which is assumed to be a representation of lower free tropospheric conditions. This holds especially for *IZO*. Because of this the greatest amount of accepted data points with 36.2% was found at this station. ~~Among all stations, the highest percentage of data, found by ABF data selection belongs to IZO, with 36.2%.~~ The ~~following~~ sites with intermediate percentages are ~~stations~~ *JFJ* (22.1%), *SNB* (19.3%), and *ZSF* (14.8%). For the three sampling heights at *HPB*, only 3.2% (50 m), 4.8% (93 m), and 6.2% (131 m) of the data were selected by ADMV ~~ABF~~. ~~At last,~~ Finally, a similar percentage ~~is~~ was found for *SSL* (4.0% ~~3.8%~~), probably due to its higher data variability. ~~To examine the characteristic growth of ABF selection percentages during the selection process, we additionally calculated selection percentages after every major step. The detailed percentage results were listed in the Supplement S3.1. All the results of percentages show similar order of stations as above, and the selection percentages increase steadily step by step for all stations.~~

Since the stations were listed according to their altitudes, it was clear ~~visible~~ that all four selection percentages increase with ~~altitudes,~~ altitude, indicating ~~which indicated~~ that measurements at higher altitudes could capture progressively well-mixed and hence representative air. Therefore, linear least-squares regression was applied between the absolute altitude and the selection percentage for continental stations. *IZO* was on a remote island and therefore not comparable. ~~As a result,~~ This approach revealed a significant positive linear trend ~~was observed~~ (see coefficient in Table 2). The related figure of linear regression can be found in Supplement S3.2.

To examine the characteristic growth of ADMV selection percentages during the selection process, we additionally calculated selection percentages after completing both the *starting selection* and *adaptive selection* steps mentioned in Section 2.2 (see Supplement S3.1). All of the results of percentages show an order of stations similar to that above, and the selection percentages increase steadily step by step for all stations.

The selection percentages of ADMV ~~ABF~~ were ~~again~~ compared again with those of the already mentioned statistical data selection methods SI, THO, and MA in Table 2, with the corresponding figure shown in Supplement S3.3. Since the selection percentages indicate not only the amount of data declared as representative but also show the characteristics of the selection methods, this criterion was used for further assessment. All other methods except for MA resulted in higher selection percentages for ~~higher~~ more highly elevated stations (*IZO*, *ZSF*, *SNB*, and *JFJ*) than ~~lower~~ for less elevated ones (*HPB* and *SSL*). ADMV ~~ABF~~ always performs the strictest in all cases. Based on the stepwise study of the selection percentages (see Supplement S3.1), ~~the reason for~~ such low percentages ~~is~~ are due to the precise definition of the *start time window*. With *adaptive selection*, the selection percentages have grown but are still ~~maintain~~ lower than those of the ~~lowest~~

~~compared with~~ other methods. SI and THO, on the other hand, show differences between stations at high and low elevations ~~elevated stations~~. Compared with SI, THO is higher at stations at lower elevations ~~in low elevated stations~~, but lower at ~~in~~ high ones ~~elevated stations~~. A major limitation of SI seems to be the requirement for ~~of~~ consecutive hours, in our case of ~~six hours~~ 6 h with 0.3 ppm standard deviation threshold, which might be too restrictive for ~~low elevated~~ stations at lower

5   elevations. However, this criterion results in a fairly large percentage for ~~high elevated~~ stations at high elevations. At *ZSF*, *SNB*, and *JFJ*, it results in the second largest, and even the largest in the case of *IZO*.

The highest selection percentages of approximately 80% ~~are~~ were obtained with MA. ~~But due~~ However, owing to the minimal data variability of CO_2 measurements at *IZO*, the selection interval in MA becomes so small that the selection percentage becomes considerably smaller than at ~~compared with~~ all other stations. However, *IZO* obtains ~~gets~~ the largest

10 selection percentages from all other selection methods. Thus, we conclude that MA does not work properly in the case of very well-mixed air (*IZO*). At all other stations, it is possible that MA declares too much ~~many~~ data as representative. Therefore, MA was excluded from further analyses.

3.3 STL decomposed results

STL was applied to the validated data-sets before and after selection with SI, THO, and ADMV ~~ABF~~, except for *HPB* due to

15 its limited length of ~~in~~ time (less than one year). ~~, which is less than one year.~~ Depending on data availability, STL was performed on CO_2 data from ~~2010 to 2014 at *SSL* and from~~ 2012 to 2015 at *SNB*, while data inputs at *SSL*, *IZO*, *ZSF*, and *JFJ* cover the ~~complete~~ whole period from 2010 to 2015. Figure 4 gives an overview of the decomposition in each component by STL. The following sections discuss the resulting components obtained by STL, namely the trend component over the observation period, the seasonal component and finally the remainder component.

20 ### 3.3.1 Trend component

From the trend component, the mean annual growth rate is estimated by linear regression (see Table 3). Based on the 95% confidence interval for the slope, a positive trend i.e. increasing CO_2 concentrations are ~~can be~~ observed. ~~Due~~ Owing to the overlapping ~~ping~~ of the confidence intervals, differences in the mean annual growth rates among VAL and ~~all~~ selected data-sets at the same station are all in good agreement ~~negligible~~. This indicates that the trend component is not influenced by the

25 statistical data selection method, which agrees well with the finding of Parrish et al. (2012) ~~for the~~ from a study of baseline ozone concentrations that there were no significant differences of the long-term changes ~~were found~~ between the baseline and unfiltered data-sets. ~~However, there is a tendency observed for all stations except for *SSL*.~~ Moreover, the following fact is observed for all sites except for *SSL*. Compared to unselected data (VAL), the mean annual growth rates based on ~~from~~ selected data-sets are systematically higher approaching the growth rates at *IZO*. ~~, in direction to the concentration levels at~~

30 ~~*IZO*, which here is considered as~~ *IZO* can be considered as better representing the ~~reference for better approximation of~~ lower free tropospheric conditions and agrees well with the mean annual absolute increase during last 10 years (2.21 ppm yr[-1]) reported by WMO (2017a). The exception ~~deviation from this tendency~~ at *SSL* is probably ~~is~~ caused by stronger local

influences as a result of its lower elevation. Besides, the confidence intervals of the mean annual growth rates are always smaller after data selection, which improves the precision of trends.

3.3.2 Seasonal component

The resulting seasonal components show systematic differences between VAL and selected data~~-~~sets. The mean monthly
5   variations were calculated on a monthly scale over the entire period from the analyzed data. Figure 5 (a) and (b) present the results at stations *ZSF* and *IZO*. At most stations (except for *IZO*), the seasonal amplitudes have been substantially reduced compared to VAL (see also Fig. 4). At *ZSF*, the averaged peak-to-peak seasonal amplitude, defined as mean seasonal maximum minus seasonal minimum, drops the most by 18.9% ~~18.88%~~ from VAL with the ADMV ~~ABF~~ selected data~~-~~set. An explanation of this reduction is CO_2 signal exclusion from local sources and sinks by data selection. When taking a closer
10   look ~~on~~ at the monthly averages, lower CO_2 values are found in the selected data-sets in the winter months from October to April, indicating that the CO_2 level is overestimated by VAL because of more dominant anthropogenic activities and ~~almost~~ no active vegetation. Higher values in the summer months from May to September explain ~~an~~ underestimation of VAL due to intensified upward transport of photosynthetic signatures resulting from vegetation ~~signals~~. Similar patterns can be found at stations *SSL*, *SNB*, and *JFJ* (see Supplement S4). *IZO*, as expected by its location, ~~shows~~ always shows the smallest
15   seasonal amplitude and nearly uninfluenced monthly results between VAL and selected data~~-~~sets. Based on this consideration, it is very likely that the lower free troposphere will react with a delay to CO_2 concentration changes of effective sources and sinks on the ground, acting like an atmospheric memory.

A time delay of one month in the mean seasonal maximum is ~~found~~ shown in Fig. 5 (a) at *ZSF* with selected data-sets by SI and ADMV ~~ABF~~ (March), compared with the maximum from the validated data (February). In addition, a similar time shift
20   can also be found by other selection methods at stations *SSL* (one-month delay from February to March by SI and ADMV ~~ABF~~) and *JFJ* (two-months delay from February to April by SI, THO, and ADMV ~~ABF~~). As for ~~stations~~ station *IZO* (April) in Fig. 5 (b) and station *SNB* (March), the seasonal maxima stay the same. The magnitude of these delays may be related to mixing features in the lower free troposphere. Rapid changes are usually observed close to sources and sinks, e.g., from anthropogenic and biogenic activities. Thus, the higher the station is above the boundary layer, the later the maxima during
25   the winter can be observed, because of the late response due to inhibited mixing conditions. However, this delay does not occur for the minima during the summer because of the very effective upward transport and more favorable mixing conditions at that time of year. ~~Regarding the seasonal minima, no changes~~ Consequently, no changes in the seasonal minima are observed at all measurement sites, which is taken as an indicator ~~of~~ for enhanced thickness of the mixing layer as good mixing conditions. Taking *ZSF* as an example, Birmili et al. (2009) observed low concentrations of particle number in
30   winter and found it representative for the free tropospheric air by analyzing the annual and diurnal cycles. From spring on, the warmer it gets the higher the PBL goes. The intense vertical atmospheric exchange during summer months results in a daily air mass transport from the boundary layer to reach *ZSF* due to thermal convection (Reiter et al., 1986; Birmili et al., 2009). Thus there are optimal transportation and mixing conditions. Therefore after data selection, the timing of seasonal

peaks better corresponds among the stations~~, a better agreement of the seasonal peaks has been reached after data selection among all stations. This better represents the seasonal cycles of the baseline conditions.~~

3.3.3 Remainder component

The remainder component contains data with external and random influences. It has characteristics of random noise, being basically different from site to site and statistically uncorrelated with the general signal of CO_2 concentrations in the lower free troposphere (Thoning et al., 1989). The standard deviation of the remainder component is taken here as a measure for external influences (see Fig. 4). Table 4 shows the calculated standard deviations from the remainder components at each station. Comparable results are derived from all selected data~~-~~sets. *SSL*, as the lowest elevated station, exhibits the most variation. *IZO* with the ~~least~~ smallest standard deviations in the remainder components proves to be the station least influenced by its surrounding environment. The three alpine measuring stations (*ZSF*, *SNB* and *JFJ*) exhibit intermediate variability. ~~show similar intermediate results.~~ From this perspective, STL performs well to show the site characteristics. Consequently, the noise of the remainder components, given in Table 4, decreases with increasing altitude of the continental mountain stations, which is in inverse relation to the selection percentage (Table 2). *IZO* was excluded in both regressions against altitude because of its maritime character. ~~Further analyses with the remainder components could also yield the local influences and characteristics at each station in more detail.~~

3.4 Correlation analysis

As mentioned above, data selection is defined here as an approach of extracting a group of data to be the best representative for the lower free troposphere. Consequently, the selected CO_2 data~~-~~sets from all stations should theoretically agree better among themselves. For validation of this, we took the combination of the trend and seasonal components from STL and examined the correlations between each pair of stations in a Pearson correlation matrix (see the upper panel of Fig. 6). The trend and seasonal components of all VAL and selected datasets were first compiled, and then Pearson's correlation coefficients were calculated assuming normal distribution of data examined by the Anderson Darling test (P < 0.05). The correlation matrices are shown for each type of dataset individually. Data used for correlation were chosen only when available at all stations (2012–2015). In general, most pairs show higher correlation coefficients with selected data from the different selection methods, especially between the three Alpine stations (*ZSF*, *SNB*, and *JFJ*). This evaluation hence shows a similar result to the method presented by Sepúlveda et al. (2014) for identifying baseline conditions based on the correlation between distant measuring stations. Pairs including *IZO* after data selection by ADMV ~~ABF~~ show a ~~noticeable~~ notable increase in the correlation coefficients, meaning a better coherence between the reference station *IZO* and the others. On the other hand, when selecting representative data more effectively, the results should contain less local and regional influences. Therefore, we compared the remainder components derived from STL pairwise to check whether the Pearson correlation coefficients decreased after data selection, as shown in the lower panel of Fig. 6. The number of insignificant correlations between the station pairings is the greatest for ADMV. ~~From all selection types compared with VAL, ABF~~

31

yields the maximum number of pairs of insignificant correlations. For the only two coefficients significant at the 0.05 confidence level (*ZSF-SNB* and *ZSF-JFJ*), they it drops largely from 0.75 to 0.48, and from 0.75 to 0.40, respectively, which cannot be observed by the other selection methods. This means that by ADMV the combination of trend and seasonal components correlate best and the remaining unselected data have the lowest correlation among the methods. If these two

5 criteria are used to separate the representative part of the data from the unrepresentative part, the ADMV method produces the best results.

4. Conclusions and outlook

We presented a novel statistical data selection method, the ADMV Adaptive Baseline Finder (ABF), for CO_2 measurements at elevated GAW mountain stations. For validation and assessment of the data selection procedure, we applied the method to

10 six CO_2 data-sets from 2010 to 2016, measured at GAW mountain stations in the European Alps. For mountain stations in the European Alps, The ADMV ABF method resulted in selected an increasing percentage of data with growing altitude elevation which is reasonable due to the underlying atmospheric dynamics. Comparing ADMV ABF with three other well-known statistical data selection methods, all methods yielded rather consistent characteristics across different stations. Nevertheless, among all the methods, ADMV ABF is the most restrictive in terms of the number of selected data in the

15 overall data-sets.

In addition, we applied the time series decomposition tool STL to all validated and selected data-sets. All statistical data selection methods resulted in the same annual trend in terms of the 95% confidence interval from the validated data-sets, while the seasonal signal varied substantially with smaller seasonal amplitudes and delayed occurrences of seasonal maxima due to the reduced and delayed influences of CO_2 sources and sinks. We also presented an additional assessment of the

20 proposed new ABF method compared with the other statistical data selection methods based on correlation analysis. Both, higher correlation coefficients of the trend and seasonal components by STL and inversely lower coefficients of the remainder indicate a better performance of ADMV ABF than the other methods SI and THO.

The presented method ADMV is useful for data selection of atmospheric CO_2 data representative of the lower free troposphere. It requires only data from a single measurement site. It is easily adjustable to the local conditions and it runs

25 automatically. The method can also be applied to historical datasets. The results provide evidence that the proposed ADMV method confers the possibility of selecting data that are representative of CO_2 concentrations of a larger area of the lower free troposphere. In all, this paper showed evidence that data selection based on the herein presented statistical properties enables practically feasible adjustments to individual conditions at GAW altitude measurement stations in Central Europe. This is an elementary a basic prerequisite for the methodical application of the method to a larger number of different

30 stations and an essential step toward towards generalization. It directly supports the objective of GAW to extrapolate from a set of point measurements from single stations to a larger representative area or region in the lower free troposphere (WMO,

2017b). ~~ABF, as an automated method, was shown to be a good option for mountain stations.~~ In future, there is a need to test whether such results could be used for additional tasks, such as ground calibration of satellite measurements. ~~Hence for future research~~ Finally, it would be extremely interesting ~~is of substantial interest~~ to test~~,~~ whether this presented concept also holds in other regions and on other continents. ~~Thus future investigations will target on including further~~

5 ~~altitude stations and on the question, how ABF will work with other air components. Also~~ Moreover, the ~~question~~ issue of whether ~~and~~ how to include coastal ~~seaside~~ stations in ~~into~~ a systematic and practically generalizable approach for selecting representative data at GAW stations will be a particular ~~of essential~~ concern.

Table 1: Information of measured CO_2 data-sets at six GAW mountain stations.

Station (GAW ID)	Sampling elevation (a.s.l.)	Time period (yyyy.mm)	Data provider
Hohenpeissenberg (*HPB*)	984/1027/1065 m	2015.09–2016.06	DWD
Schauinsland (*SSL*)	1210 m	2010.01–2015.12 ~~2014.12~~	UBA-De
Izaña (*IZO*)	2403 m	2010.01–2015.12	AEMET
Zugspitze-Schneefernerhaus (*ZSF*)	2670 m	2010.01–2015.12	UBA-De
Sonnblick (*SNB*)	3111 m	2010.01–2015.12	UBA-At
Jungfraujoch (*JFJ*)	3580 m	2010.01–2015.12	Empa

Table 2: Selection percentages of selected data ~~in~~ among all data by different data selection methods. The bottom shows the linear regression coefficients of stations' (*HPB* is represented by *HPB50*; *IZO* is excluded) altitudes and the selection percentages at a significance level of 0.05 (***).

Station ID	ADMV ~~ABF~~	SI	THO	MA
HPB50	3.2	13.9	21.7	79.8
HPB93	4.8	18.5	25.0	79.4
HPB131	6.2	21.3	27.3	79.8
SSL	4.0 ~~3.8~~	17.9 ~~17.3~~	25.4 ~~25.3~~	83.2 ~~81.4~~
IZO	36.2	82.2	56.0	60.5
ZSF	14.8	47.1	40.8	79.0
SNB	19.3	58.7	44.2	76.9
JFJ	22.1	62.1	46.3	77.6
Linear regression coefficient (γ^2)	0.996***	0.992 ~~0.991~~***	0.985 ~~0.986~~***	0.645 ~~0.751~~

Table 3: Mean annual growth rates (ppm y $^{-1}$ ~~-1~~) with 95% confidence intervals from linear regression applied on the trend components by STL over 2010 to 2015, except for *SNB*. Data at *SNB* were decomposed over 2012 to 2015 due to missing data from 2010 to 2011 and thus shown in gray.

Station ID	VAL	SI	THO	ADMV ~~ABF~~
SSL	2.04 ± 0.09 ~~2.12 ± 0.14~~	1.89 ± 0.06 ~~1.72 ± 0.05~~	2.04 ± 0.06 ~~2.05 ± 0.10~~	2.03 ± 0.09 ~~1.85 ± 0.10~~
IZO	2.24 ± 0.03	2.26 ± 0.02	2.25 ± 0.02	2.25 ± 0.02
ZSF	2.13 ± 0.08	2.16 ± 0.05	2.17 ± 0.06	2.19 ± 0.06
SNB	2.02 ± 0.07	2.06 ± 0.06	2.06 ± 0.06	2.08 ± 0.04
JFJ	2.13 ± 0.03	2.15 ± 0.02	2.14 ± 0.02	2.14 ± 0.02

Table 4: Standard deviation of the remainder components by STL over 2010 to 2015, except for *SNB*. Data at *SNB* were decomposed over 2012 to 2015 due to missing data from 2010 to 2011 and thus shown in gray.

Station ID	VAL	SI	THO	ADMV ABF
SSL	1.61 ~~1.58~~	1.16 ~~1.27~~	1.26 ~~1.27~~	1.99 ~~2.03~~
IZO	0.34	0.33	0.30	0.30
ZSF	0.89	0.75	0.72	0.73
SNB	0.66	0.56	0.55	0.70
JFJ	0.56	0.45	0.48	0.47
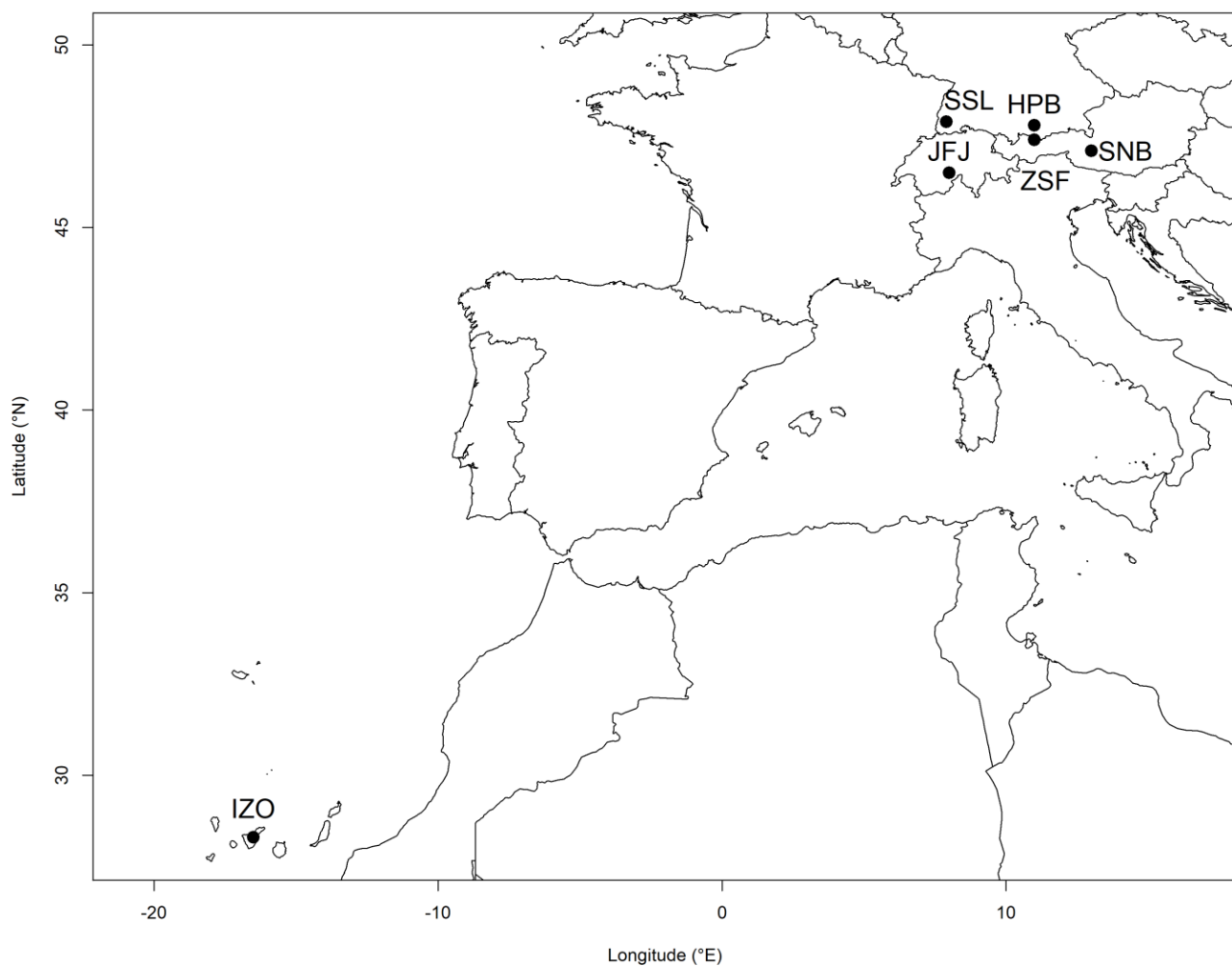
Figure 1: Locations of six European elevated mountain stations ~~included in this study~~. Symbols from left to right stand for: IZO – Izaña, Spain; SSL – Schauinsland, Germany; JFJ – Jungfraujoch, Switzerland; HPB – Hohenpeissenberg, Germany; ZSF – Schneefernerhaus-Zugspitze, Germany; SNB – Sonnblick, Austria. ~~Codes in brackets stand for countries (DE – Germany; AT – Austria; CH – Switzerland; ES – Spain).~~
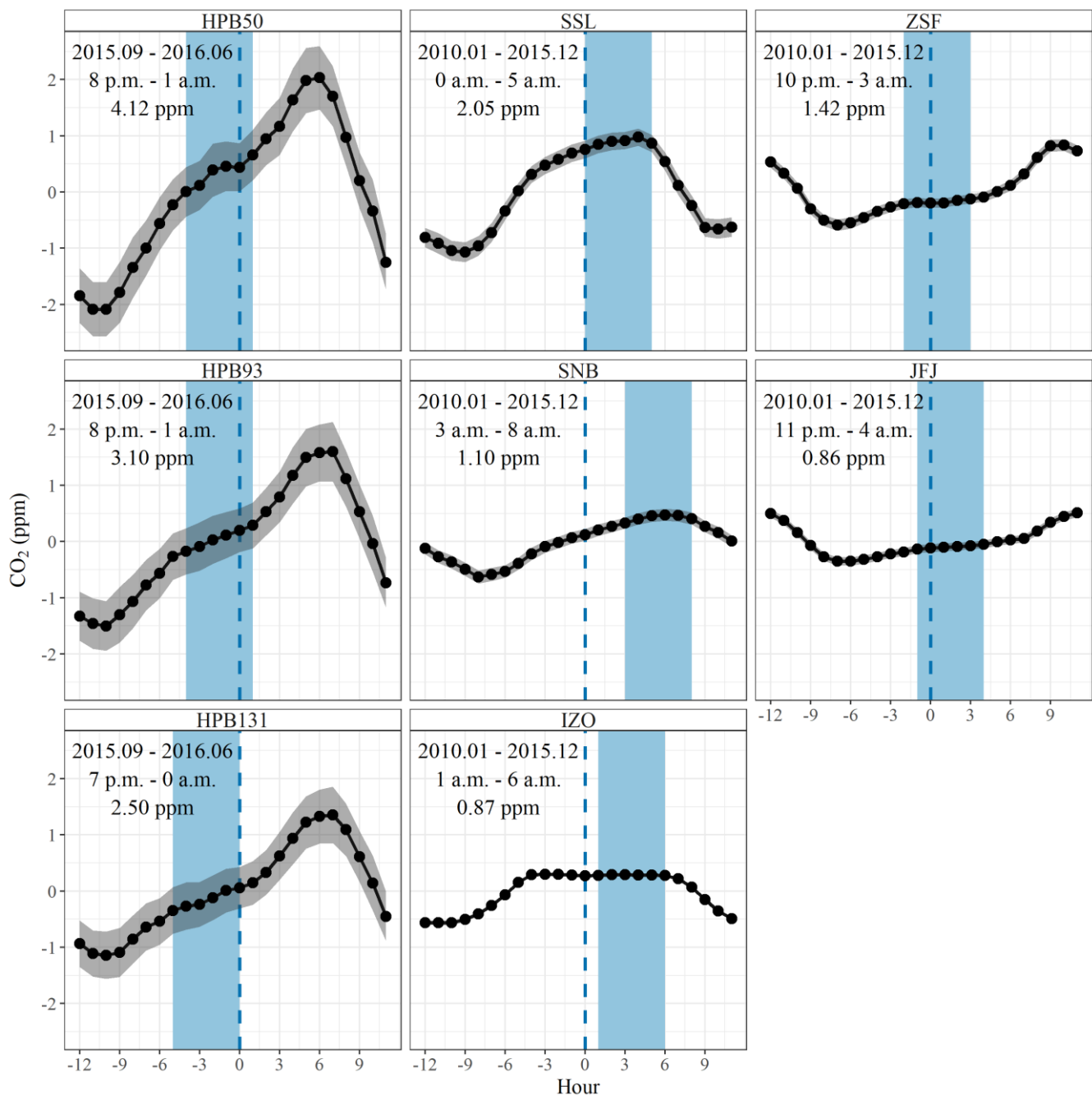
Figure 2: Detrended mean diurnal cycles of validated CO₂ data-sets (black) with 95% confidence intervals (~~grey~~ gray) from six GAW stations (hours in LT). Measurements at *HPB* are differentiated by the sampling heights (e.g., *HPB50* for 50 m a.g.l.). The covered time periods (top text), resulting *start time windows* (middle text, also in light blue shades), and mean diurnal amplitudes (bottom text) are shown in each subplot.
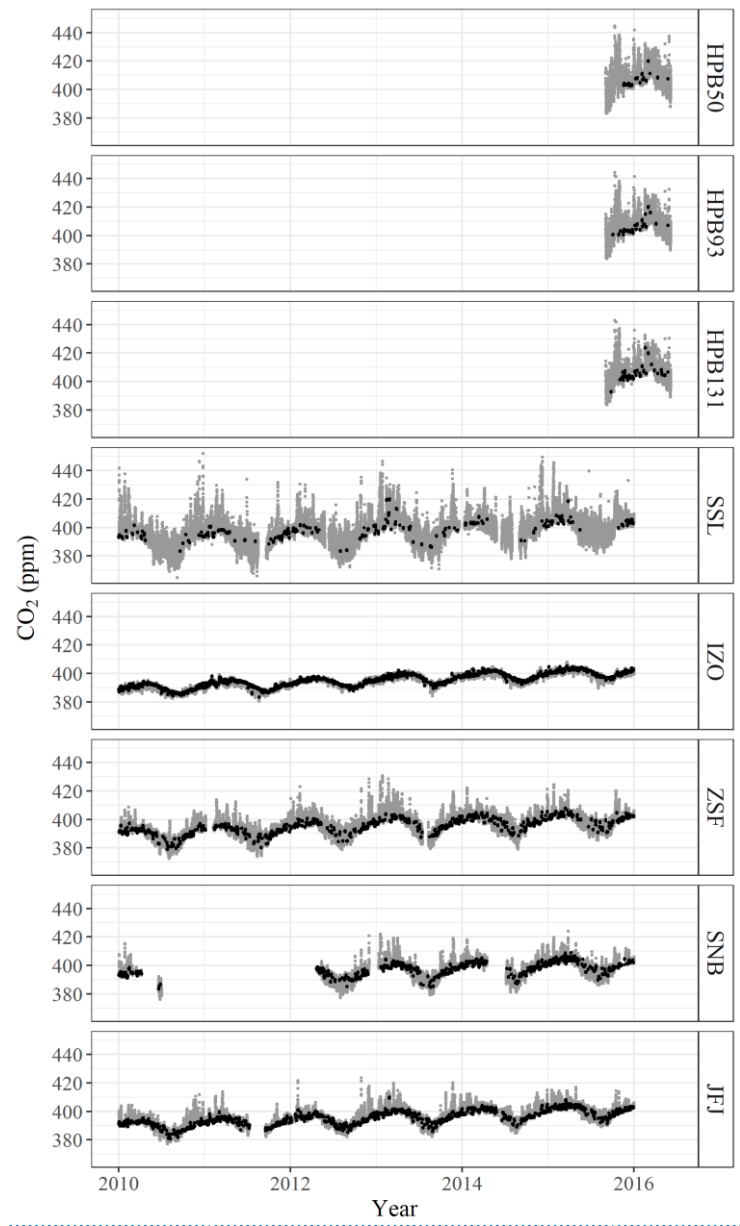
Figure 3: Time series plots of validated CO₂ data-sets (~~grey~~ gray), and selected data-sets by ADMV ~~ABF~~ (black) at six GAW stations.
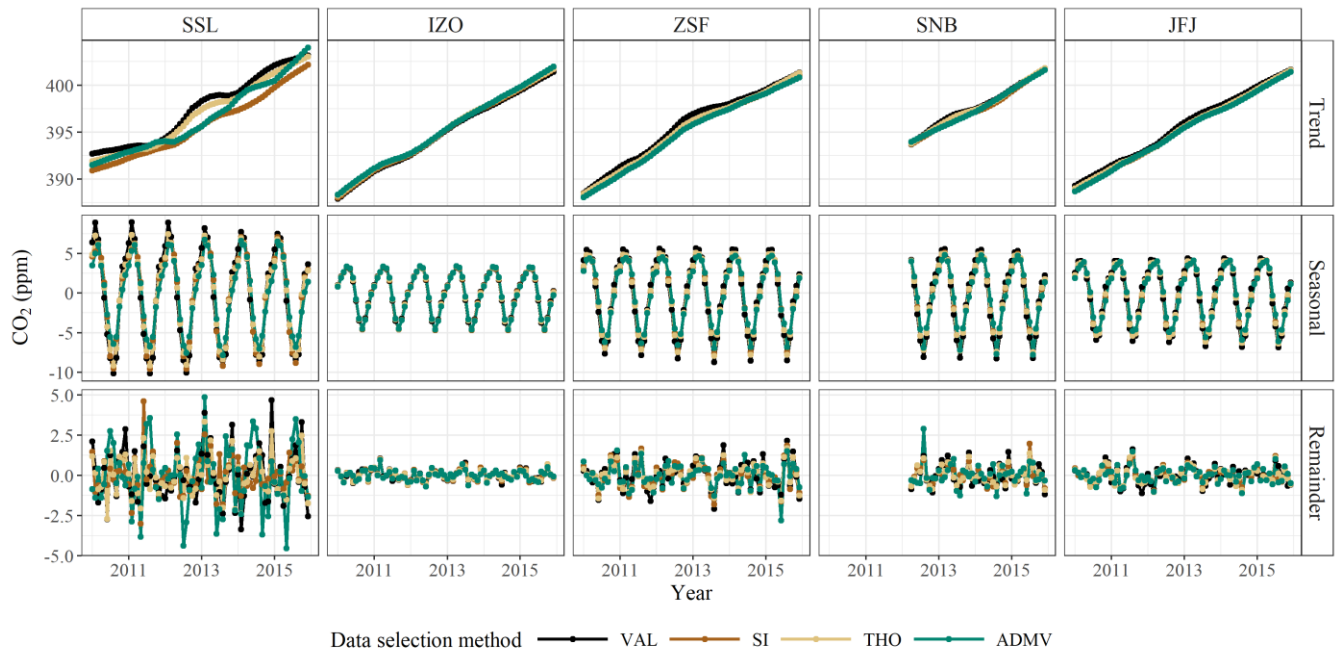
Figure 4: STL decomposition results from VAL (black), SI-selected (brown), THO-selected (yellow), and ADMV ABF-selected (green) data sets at five GAW stations.
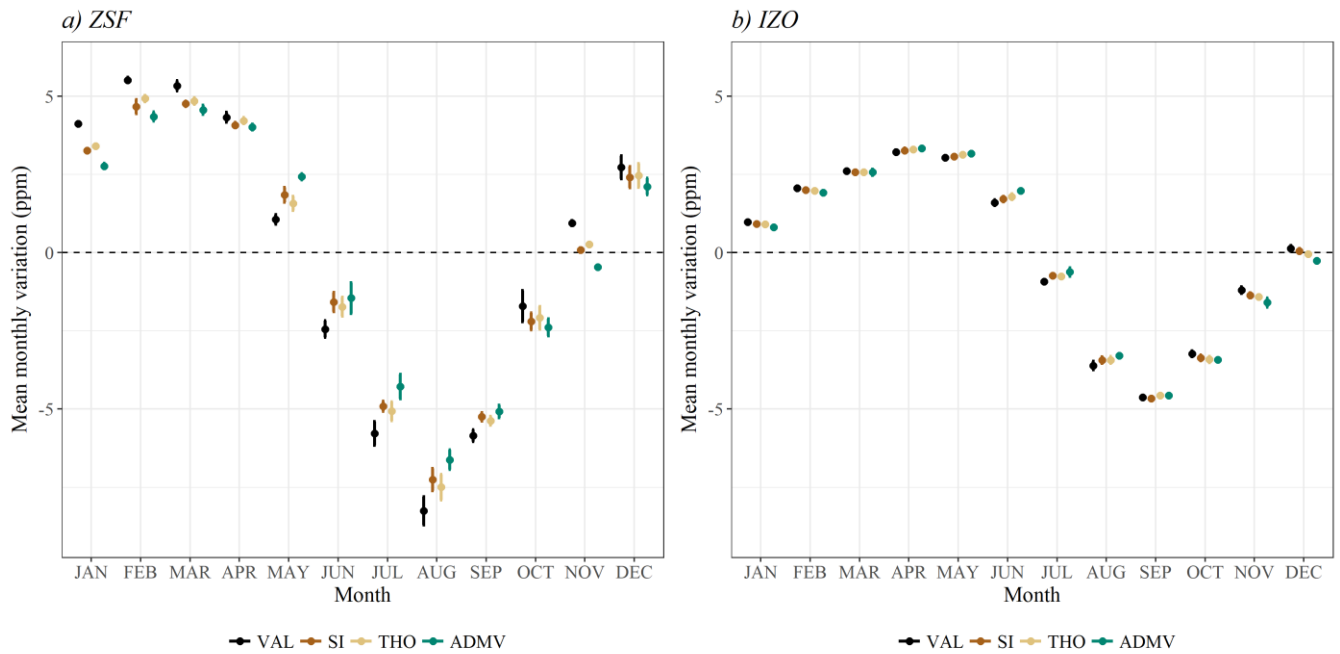
Figure 5: Mean monthly variation of the seasonal component decomposed by STL at a) *ZSF* and b) *IZO* over the whole period. For better ~~visualizing~~ visualization of the results of selection methods, dots have been separated horizontally ~~equidistant~~ equidistantly. The 95% confidence intervals are shown as error bars.
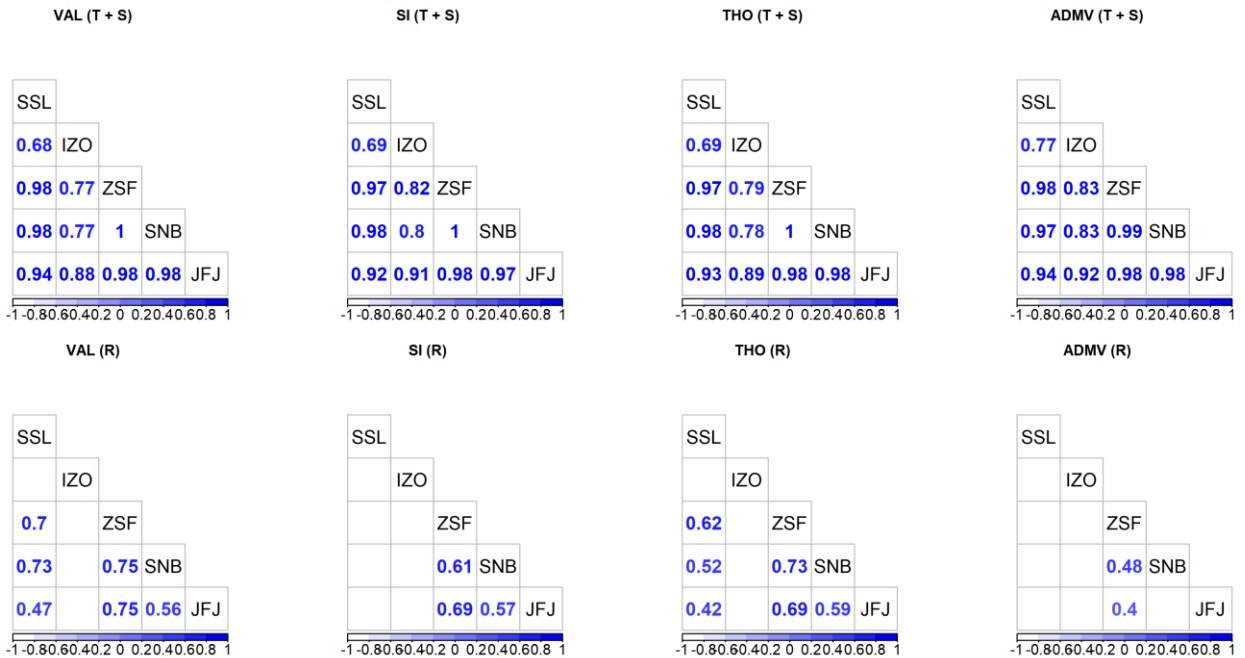
5

Figure 6: ~~Pearson~~ Pearson's correlation ~~matrices~~ ~~matrix~~ of combinations of trend (T) and seasonal (S) components (upper panel), and only remainder (R) components (lower panel) at stations *SSL*, *IZO*, *ZSF*, *SNB*, and *JFJ* by different selection methods, as indicated on the top. ~~Pearson correlation was applied due to normal distribution of data examined by Anderson-Darling test. Data used for correlation were chosen when available at all stations (2012.04 – 2014.12).~~ The ~~blue~~ color scale reflects the strength of ~~positive~~ correlation. Correlations with no significant coefficient at the 0.05 confidence level were left blank.

43

References

Bacastow, R. B., Keeling, C. D., and Whorf, T. P.: Seasonal Amplitude Increase in Atmospheric CO_2 Concentration at Mauna Loa, Hawaii, 1959-1982, J. Geophys. Res. Journal of Geophysical Research, 90, 10529–10540, doi:10.1029/JD090iD06p10529, 1985.

5 Balzani Lööv, J. M., Henne, S., Legreid, G., Staehelin, J., Reimann, S., Prévôt, A. S. H., Steinbacher, M., and Vollmer, M. K.: Estimation of background concentrations of trace gases at the Swiss Alpine site Jungfraujoch (3580 m asl), J. Geophys. Res., 113, doi:10.1029/2007JD009751, 2008.

Birmili, W., Ries, L., Sohmer, R., Anastou, A., Sonntag, A., König, K., and Levin, I.: Feine und ultrafeine Aerosolpartikeln an der GAW-Station Schneefernerhaus/Zugspitze, Gefahrst. Reinhalt. L, 69, 1/2, 31–35, 2009.

10 Brailsford, G. W., Stephens, B. B., Gomez, A. J., Riedel, K., Mikaloff Fletcher, S. E., Nichol, S. E., and Manning, M. R.: Long-term continuous atmospheric CO_2 measurements at Baring Head, New Zealand, Atmospheric Measurement Techniques, 5, 3109–3117, doi:10.5194/amt-5-3109-2012, 2012.

Brooks, B. G. J., Desai, A. R., Stephens, B. B., Bowling, D. R., Burns, S. P., Watt, A. S., Heck, S. L., and Sweeney, C.: Assessing filtering of mountaintop CO_2 mole fractions for application to inverse models of biosphere-atmosphere carbon

15 exchange, Atmospheric Chemistry and Physics, 12, 2099–2115, doi:10.5194/acp-12-2099-2012, 2012.

Calvert, J. G.: Glossary of atmospheric chemistry terms, Pure and Applied Chemistry, 62, 2167–2219, doi:10.1351/pac199062112167, 1990.

Carnuth, W., Kempfer, U., and Trickl, T.: Highlights of the tropospheric lidar studies at IFU within the TOR project, Tellus B, 54, 163–185, doi:10.1034/j.1600-0889.2002.00245.x, 2002.

20 Carnuth, W. and Trickl, T.: Transport studies with the IFU three-wavelength aerosol lidar during the VOTALP Mesolcina experiment, Atmospheric Environment, 34, 1425–1434, doi:10.1016/S1352-2310(99)00423-9, 2000.

Carslaw, D. C.: On the changing seasonal cycles and trends of ozone at Mace Head, Ireland, Atmos. Chem. Phys., 5, 3441–3450, doi:10.5194/acp-5-3441-2005, 2005.

Chambers, S. D., Williams, A. G., Conen, F., Griffiths, A. D., Reimann, S., Steinbacher, M., Krummel, P. B., Steele, L. P.,

25 van der Schoot, M. V., Galbally, I. E., Molloy, S. B., and Barnes, J. E.: Towards a Universal "Baseline" Characterisation of Air Masses for High- and Low-Altitude Observing Stations Using Radon-222, Aerosol Air Qual. Res., 16, 885–899, doi:10.4209/aaqr.2015.06.0391, 2016.

Cleveland, R. B., Cleveland, W. S., McRae, J. E., and Terpenning, I.: STL: A seasonal-trend decomposition procedure based on Loess, 6, 3–73, 1990.

30 Cleveland, W. S.: Robust locally weighted regression and smoothing scatterplots, Journal of the American Statistical Association, 74, 829–836, doi:10.1080/01621459.1979.10481038, 1979.

Cleveland, W. S., Freeny, A. E., and Graedel, T. E.: The Seasonal Component of Atmospheric CO_2: Information From New Approaches to the Decomposition of Seasonal Time Series, J. Geophys. Res. ~~Journal of Geophysical Research~~, 88, 10934–10946, doi:10.1029/JC088iC15p10934, 1983.

Cui, J., Pandey Deolal, S., Sprenger, M., Henne, S., Staehelin, J., Steinbacher, M., and Nédélec, P.: Free tropospheric ozone changes over Europe as observed at Jungfraujoch (1990–2008): An analysis based on backward trajectories, J. Geophys. Res., 116, doi:10.1029/2010JD015154, 2011.

Elliott, W. P. (Ed.): The Statistical treatment of CO_2 data records, NOAA Technical Memorandum ERL ARL, 173, U.S. Dept. of Commerce, National Oceanic and Atmospheric Administration, Environmental Research Laboratories, Silver Spring, Md., ix, 131, 1989.

Fang, S. -X., Tans, P. P., Steinbacher, M., Zhou, L. -X., and Luan, T.: Comparison of the regional CO_2 mole fraction filtering approaches at a WMO/GAW regional station in China, Atmospheric Measurement Techniques, 8, 5301–5313, doi:10.5194/amt-8-5301-2015, 2015.

Forrer, J., Rüttimann, R., Schneiter, D., Fischer, A., Buchmann, B., and Hofer, P.: Variability of trace gases at the high-Alpine site Jungfraujoch caused by meteorological transport processes, J. Geophys. Res., 105, 12241–12251, doi:10.1029/1999JD901178, 2000.

Gilge, S., Plass-Duelmer, C., Fricke, W., Kaiser, A., Ries, L., Buchmann, B., and Steinbacher, M.: Ozone, carbon monoxide and nitrogen oxides time series at four alpine GAW mountain stations in central Europe, Atmos. Chem. Phys., 10, 12295–12316, doi:10.5194/acp-10-12295-2010, 2010.

Gomez-Pelaez, A. J., Ramos, R., Cuevas, E., and Gomez-Trueba, V.: 25 years of continuous CO_2 and CH_4 measurements at Izaña Global GAW mountain station: annual cycles and interannual trends, in: Proceedings of the Symposium on Atmospheric Chemistry and Physics at Mountain Sites (ACP Symposium 2010), Interlaken, Switzerland, June 8-10, 2010, 157–159, 2010.

Gomez-Pelaez, A. J., Ramos, R., Gomez-Trueba, V., Novelli, P. C., and Campo-Hernandez, R.: A statistical approach to quantify uncertainty in carbon monoxide measurements at the Izaña global GAW station: 2008–2011, Atmos. Meas. Tech., 6, 787–799, doi:10.5194/amt-6-787-2013, 2013.

Henne, S., Brunner, D., Folini, D., Solberg, S., Klausen, J., and Buchmann, B.: Assessment of parameters describing representativeness of air quality in-situ measurement sites, Atmos. Chem. Phys., 10, 3561–3581, doi:10.5194/acp-10-3561-2010, 2010.

Herrmann, E., Weingartner, E., Henne, S., Vuilleumier, L., Bukowiecki, N., Steinbacher, Coen, F., Collaud Conen, M., Hammer, E., Jurányi, Z., Baltensperger, U., and Gysel, M.: Analysis of long-term aerosol size distribution data from Jungfraujoch with emphasis on free tropospheric conditions, cloud influence, and air mass transport, J. Geophys. Res., 120, doi:10.1002/2015JD023660, 2015.

Hernández-Paniagua, I. Y., Lowry, D., Clemitshaw, K. C., Fisher, R. E., France, J. L., Lanoisellé, M., Ramonet, M., and Nisbet, E. G.: Diurnal, seasonal, and annual trends in atmospheric CO_2 at southwest London during 2000–2012: Wind

45

sector analysis and comparison with Mace Head, Ireland, Atmospheric Environment, 105, 138–147, doi:10.1016/j.atmosenv.2015.01.021, 2015.

Kaiser, A., Scheifinger, H., Spangl, W., Weiss, A., Gilge, S., Fricke, W., Ries, L., Cemas, D., and Jesenovec, B.: Transport of nitrogen oxides, carbon monoxide and ozone to the Alpine Global Atmosphere Watch stations Jungfraujoch (Switzerland), Zugspitze and Hohenpeissenberg (Germany), Sonnblick (Austria) and Mt. Krvavec (Slovenia), Atmospheric Environment, 41, 9273–9287, doi:10.1016/j.atmosenv.2007.09.027, 2007.

Keeling, C. D.: The Concentration and Isotopic Abundances of Carbon Dioxide in the Atmosphere, Tellus, 12, 200–203, doi:10.1111/j.2153-3490.1960.tb01300.x, 1960.

Locher, R. and Ruckstuhl, A.: IDPmisc: Utilities of Institute of Data Analyses and Process Design (www.idp.zhaw.ch): https://CRAN.R-project.org/package=IDPmisc, 2012.

Lowe, D. C., Guenther, P. R., and Keeling, C. D.: The concentration of atmospheric carbon dioxide at Baring Head, New Zealand, Tellus, 31, 58–67, doi:10.1111/j.2153-3490.1979.tb00882.x, 1979.

Massen, F. and Beck, E. -G.: Accurate Estimation of CO_2 Background Level from Near Ground Measurements at Non-Mixed Environments, in: The Economic, Social and Political Elements of Climate Change, Leal Filho, W. (Ed.), Climate Change Management, Springer Berlin Heidelberg, Berlin, Heidelberg, 509–522, 2011.

Pales, J. C. and Keeling, C. D.: The Concentration of Atmospheric Carbon Dioxide in Hawaii, J. Geophys. Res. ~~Journal of Geophysical Research~~, 70, 6053–6076, doi:10.1029/JZ070i024p06053, 1965.

Parrish, D. D., Law, K. S., Staehelin, J., Derwent, R., Cooper, O. R., Tanimoto, H., Volz-Thomas, A., Gilge, S., Scheel, H. -E., Steinbacher, M., and Chan, E.: Long-term changes in lower tropospheric baseline ozone concentrations at northern mid-latitudes, Atmos. Chem. Phys., 12, 11485–11504, doi:10.5194/acp-12-11485-2012, 2012.

Peterson, J. T., Komhyr, W. D., Harris, T. B., and Waterman, L. S.: Atmospheric carbon dioxide measurements at Barrow, Alaska, 1973-1979, Tellus, 34, 166–175, doi:10.1111/j.2153-3490.1982.tb01804.x, 1982.

Pickers, P. A. and Manning, A. C.: Investigating bias in the application of curve fitting programs to atmospheric time series, Atmos. Meas. Tech., 8, 1469–1489, doi:10.5194/amt-8-1469-2015, 2015.

R Core Team: R: A Language and Environment for Statistical Computing, Vienna, Austria: https://www.R-project.org/, 2017.

Reiter, R., Sladkovic, R., Kanter, H.-J., 1986. Concentration of trace gases in the lower troposphere, simultaneously recorded at neighboring mountain stations. Meteorl. Atmos. Phys. 35 (4), 187–200. 10.1007/BF01041811.

Risius, S., Xu, H., Di Lorenzo, F., Xi, H., Siebert, H., Shaw, R. A., and Bodenschatz, E.: Schneefernerhaus as a mountain research station for clouds and turbulence, Atmos. Meas. Tech., 8, 3209–3218, doi:10.5194/amt-8-3209-2015, 2015.

Ruckstuhl, A. F., Henne, S., Reimann, S., Steinbacher, M., Vollmer, M. K., O'Doherty, S., Buchmann, B., and Hueglin, C.: Robust extraction of baseline signal of atmospheric trace species using local regression, Atmos. Meas. Tech., 5, 2613–2624, doi:10.5194/amt-5-2613-2012, 2012.

Satar, E., Berhanu, T. A., Brunner, D., Henne, S., and Leuenberger, M.: Continuous CO_2/CH_4/CO measurements (2012–2014) at Beromünster tall tower station in Switzerland, Biogeosciences, 13, 2623–2635, doi:10.5194/bg-13-2623-2016, 2016.

Schibig, M. F., Steinbacher, M., Buchmann, B., van der Laan-Luijkx, I. T., van der Laan, S., Ranjan, S., and Leuenberger, M. C.: Comparison of continuous in situ CO_2 observations at Jungfraujoch using two different measurement techniques, Atmos. Meas. Tech., 8, 57–68, doi:10.5194/amt-8-57-2015, 2015.

Schmidt, M., Graul, R., Sartorius, H., and Levin, I.: The Schauinsland CO_2 record: 30 years of continental observations and their implications for the variability of the European CO_2 budget, Journal of Geophysical Research D: Atmospheres, 108, ACL 14-1 - 14-7, 2003.

Sepúlveda, E., Schneider, M., Hase, F., Barthlott, S., Dubravica, D., García, O. E., Gomez-Pelaez, A., González, Y., Guerra, J. C., Gisi, M., Kohlhepp, R., Dohe, S., Blumenstock, T., Strong, K., Weaver, D., Palm, M., Sadeghi, A., Deutscher, N. M., Warneke, T., Notholt, J., Jones, N., Griffith, D. W. T., Smale, D., Brailsford, G. W., Robinson, J., Meinhardt, F., Steinbacher, M., Aalto, T., and Worthy, D.: Tropospheric CH_4 signals as observed by NDACC FTIR at globally distributed sites and comparison to GAW surface in situ measurements, Atmos. Meas. Tech., 7, 2337–2360, doi:10.5194/amt-7-2337-2014, 2014.

Sirignano, C., Neubert, R. E., Rödenbeck, C., and J. Meijer, H. A.: Atmospheric oxygen and carbon dioxide observations from two European coastal stations 2000-2005: continental influence, trend changes and APO climatology, Atmospheric Chemistry and Physics, 10, 1599–1615, 2010.

Stephens, B. B., Brailsford, G. W., Gomez, A. J., Riedel, K., Mikaloff Fletcher, S. E., Nichol, S., and Manning, M.: Analysis of a 39-year continuous atmospheric CO_2 record from Baring Head, New Zealand, Biogeosciences, 10, 2683–2697, doi:10.5194/bg-10-2683-2013, 2013.

Thoning, K. W., Tans, P. P., and Komhyr, W. D.: Atmospheric Carbon Dioxide at Mauna Loa Observatory: 2. Analysis of the NOAA GMCC Data, 1974-1985, J. Geophys. Res. ~~Journal of Geophysical Research~~, 94, 8549–8565, doi:10.1029/JD094iD06p08549, 1989.

Uglietti, C., Leuenberger, M., and Brunner, D.: European source and sink areas of CO_2 retrieved from Lagrangian transport model interpretation of combined O_2 and CO_2 measurements at the high alpine research station Jungfraujoch, Atmospheric Chemistry and Physics, 11, 8017–8036, doi:10.5194/acp-11-8017-2011, 2011.

WMO: WMO Greenhouse Gas Bulletin, No. 13, ISSN 2078-0796, 30 October, 2017a.

WMO: WMO Global Atmosphere Watch (GAW) Implementation Plan: 2016-2023, 2017b.

Zeileis, A. and Grothendieck, G.: zoo: S3 Infrastructure for Regular and Irregular Time Series, J. Stat. Soft., 14, doi:10.18637/jss.v014.i06, 2005.

Zellweger, C., Forrer, J., Hofer, P., Nyeki, S., Schwarzenbach, B., Weingartner, E., Ammann, M., and Baltensperger, U.: Partitioning of reactive nitrogen (NO_y) and dependence on meteorological conditions in the lower free troposphere, Atmos. Chem. Phys., 3, 779–796, doi:10.5194/acp-3-779-2003, 2003.