Response to Reviews

May 23, 2017

Referee #1 (Thomas Bauska):

This paper presents continuous histories of greenhouse gas variability and radiative forcing over the past 150,000 years based on a carefully selected compilation of previously published ice core records and simple, generally well-described time-series smoothing techniques. The authors are experts in the field of ice core gas measurements and well qualified to undertake this data compilation study (albeit they are not always the originators of all the data used here). The atmospheric histories presented here will be of interest to a number of modelling studies (particularly transient simulations) currently underway and to a wide-range of paleoclimate researchers seeking the best ice core gas data.

I have a number of comments on the paper that are individually minor and can probably be addressed with further clarifications in the text. However, in aggregate they probably amount to a major revision. I will outline a few broader points, then digress with a suggestion, and end with some line-by-line comments.

Edge effects and stitching the splines together. The use of splines to smooth atmospheric gas records (particularly CO_2) is relatively wide-spread and justified here. However, there are a number of pitfalls that may be encountered and need to be addressed with further clarification in the text. My main concern is that splines can sometimes overfit the data if they are weighted too heavily to minimizing the data fit while allowing the curvature to vary greatly. In practice this can cause the spline to produce overshoots in between data points. The authors state that ideally, a smoothing spline removes all high frequencies (sic) sine functions. I would suggest that its also possible for splines to produce erroneous high-frequency variability if the Pcutoff is too small. Though I believe this is generally not the case (but see comments on N_2O and CH_4 in the Holocene), how the analysis has tested for these possible errors needs to be discussed.

It was unclear to me about how the final, continuous spline is constructed from a number of different, discrete time intervals. Are the splines calculated with overlapping data and then truncated at both ends? Or are the splines simply spliced end-to-end? The tables provided show tstart and tstop intervals that are almost always separated by intervals that are much greater than dt (in other words there appear to be gaps between the splines). I suspect some crucial information on the method is missing here. My worry is that there can be significant edge effects with splines and therefore the truncation method could induce some artificial features in the final spline. By the eye this doesnt appear to be a problem but a little more information is needed.

Our reply: Some clarification on the splines seem to be necessary:

1. Overfit: As general guideline we first investigate the data coverage of a GHG (Δt, Figs 1B, 3B, 5B), which then gives us a low limit of the cutoff frequency. We here rely more on the 11-pts running mean than on the individual data, but keep Pcutoff well above Δt. Second, we take into account if abrupt changes in the GHG are expected in a time interval. For, example, during the Holocene only little climate related changes in CO_2 are expected, leading to a chosen cutoff period of 3000 years. Certainly, the final choice of the cutoff period is to some extend subjective. To add some further details here we will in a revised version calculate splines, in which the cutoff periods are systematically varied from the control values ($\pm$ 50%) and will plot resulting splines against the underlying data to detect any overfitting as function of cutoff period. The question of overfitting for the Holocene and CH_4 and N_2O is addressed further below

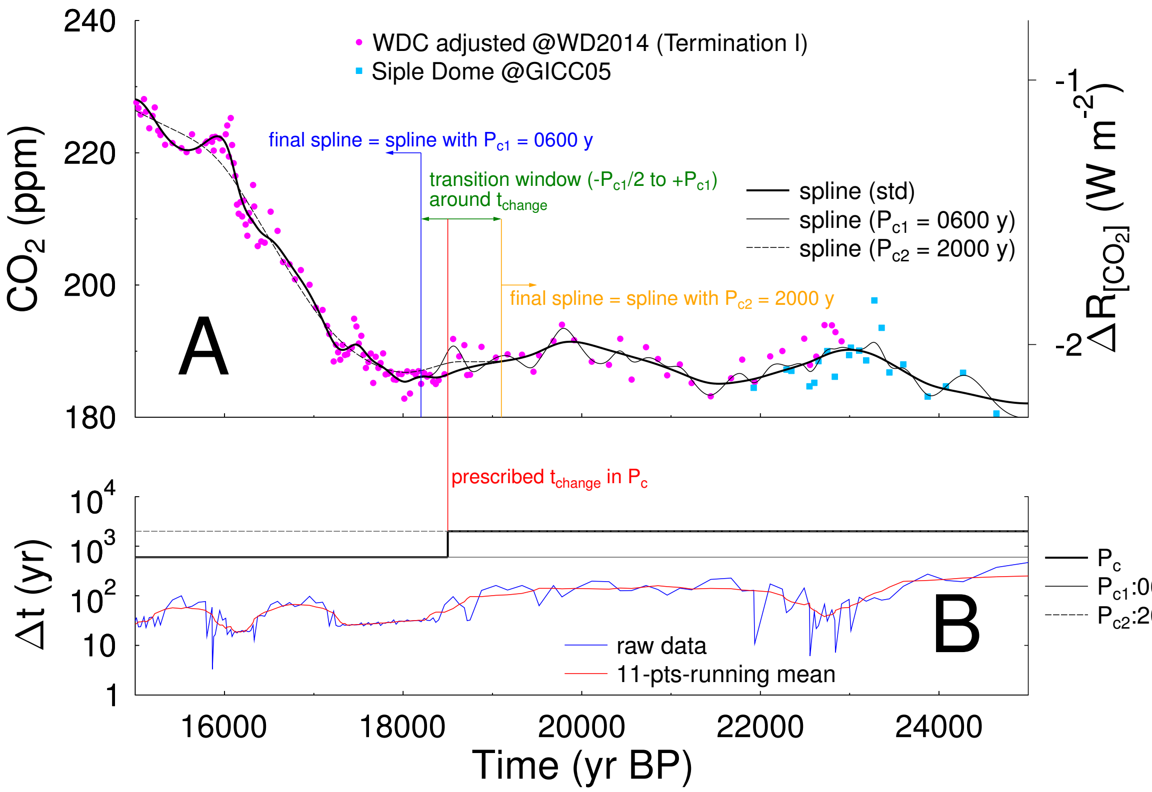
(when these specific points are brought up again), but as you will see we there is no overfitting contained in the splines so far.

2. Continuous spline: Only one continuous spline is calculated over the whole time window of interest, so there are no shorter splines which need to get spliced together. The different cutoff periods for the different time windows are taken into account by the calculation of $\lambda' = \lambda \cdot s^2$ with s being the scaling factor which is defined by the cutoff period as given by Eq 4 in the text. In practise, based on λ' an intermediate product (time series of y_i and v'_i) is calculated, in which the revised uncertainties v'_i related to each data point i are calculated following the relation of s and λ as given by Eq. 4 in the text.

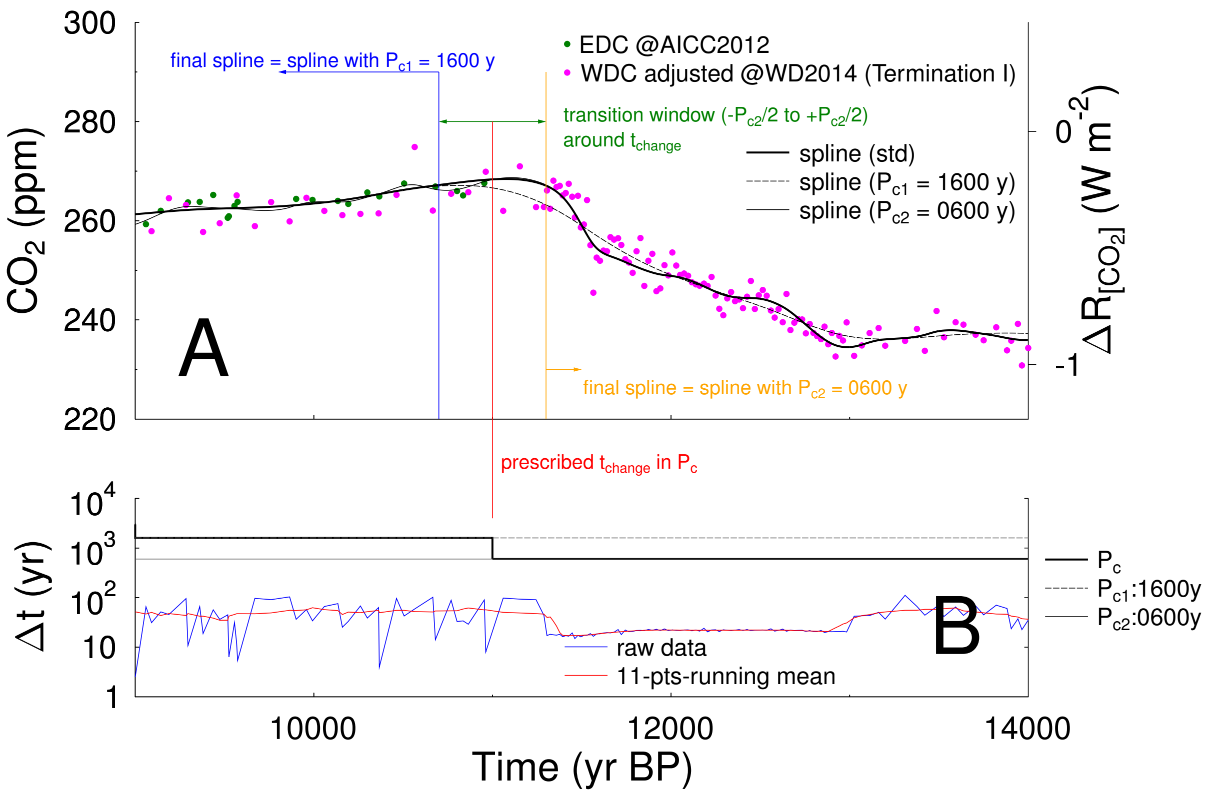
3. Data gaps: The tables with details on the splines (Tables 3, 6, 8) give the exact date of the first and the last data point in an interval. There are no gaps between any two intervals with data points. In other words, if interval 4 stops at 9425.6 yr BP and interval 5 starts at 9517.2 yr BP, then there are no data points between 9425.6 yr BP and 9517.2 yr BP (existing example from N_2O).

4. Change in cutoff period: When changing the cutoff period from P_{c1} to P_{c2} there is a time window of transition around the prescribed time of change in the cutoff period t_{change}, starting around $t_{\text{change}} - P_c$ to around $t_{\text{change}} + P_c$, where P_c is the smaller of both P_{c1} and P_{c2}, in which the transition of the cutoff period can be identified in the final spline. Before that transition window the spline is identical to a spline with P_{c1}, after that transition window the spline is identical to a spline with P_{c2}. This change can be seen in the two plots below, in which two different cutoff periods and a transition between both have been applied in two examples. Note, that in the examples the transition window can be defined in more detail $[(t_{\text{change}} - P_{c1}/2$ to $t_{\text{change}} + P_{c2})$, or $(t_{\text{change}} - P_{c2}/2$ to $t_{\text{change}} + P_{c2}/2)]$ but these are special cases, and cannot by generalised. This is also the reason why the tables which contain the statistics of the spline include a column $\overline{P_{\text{cutoff}}}$ with the realised cutoff period, which is due to this transition window slightly smaller than the prescribed P_{cutoff}.

We will extend the description of the spline on these details.



The figure above illustrates the transition effect, when changing the cutoff period from a low P_{c1} to a high P_{c2}. In sub-panel B not only the age distance of the data points, but also the cutoff periods are plotted.



The figure above illustrates another example of the transition effect, when changing the cutoff period from a high P_{c1} to a low P_{c2}. In sub-panel B not only the age distance of the data points, but also the cutoff periods are plotted.

A thorough description of the smoothing induced by the splines including modification of the data files. At the moment, the description of the spline fitting routine is fairly broad, relies a little too heavily on other references, is not exactly precise, and contains some grammatical errors. See line-by-line comments for more details, but broadly I would like to more emphasis on how the Pcutoff value relates to the smoothing of the data.

The authors state the code used to construct the splines is based on Enting, 1987. As there are also a number of software packages capable of spline fitting (Matlab, IGOR, even Excel, etc.) which other researches may apply to the data for comparison it would be useful to have little information about the actual code used. Is there a specific source for this code (I believe the Enting used to be available online) or will the new code for constructing the splines be made available? I didnt see anything in the ESSD requirements regarding code availability, so it doesnt appear necessary for publication, but perhaps it should be considered by the authors in the spirit of open-access/open-source. Also, there are a wide variety of smoothing splines techniques, so more information is needed on what specific type of spline technique has been used.

The authors have a done an excellent job at producing well referenced and easily understood data files. I am particularly glad to see a complete list of references for the original data with tags to individual data points. However, one important aspect of the study has not been transferred to the data files is the smoothing induced by the splines. In the spline files, I would suggest including the Pcutoff value so a user of the data can readily assess the degree to which the data has been smoothed.

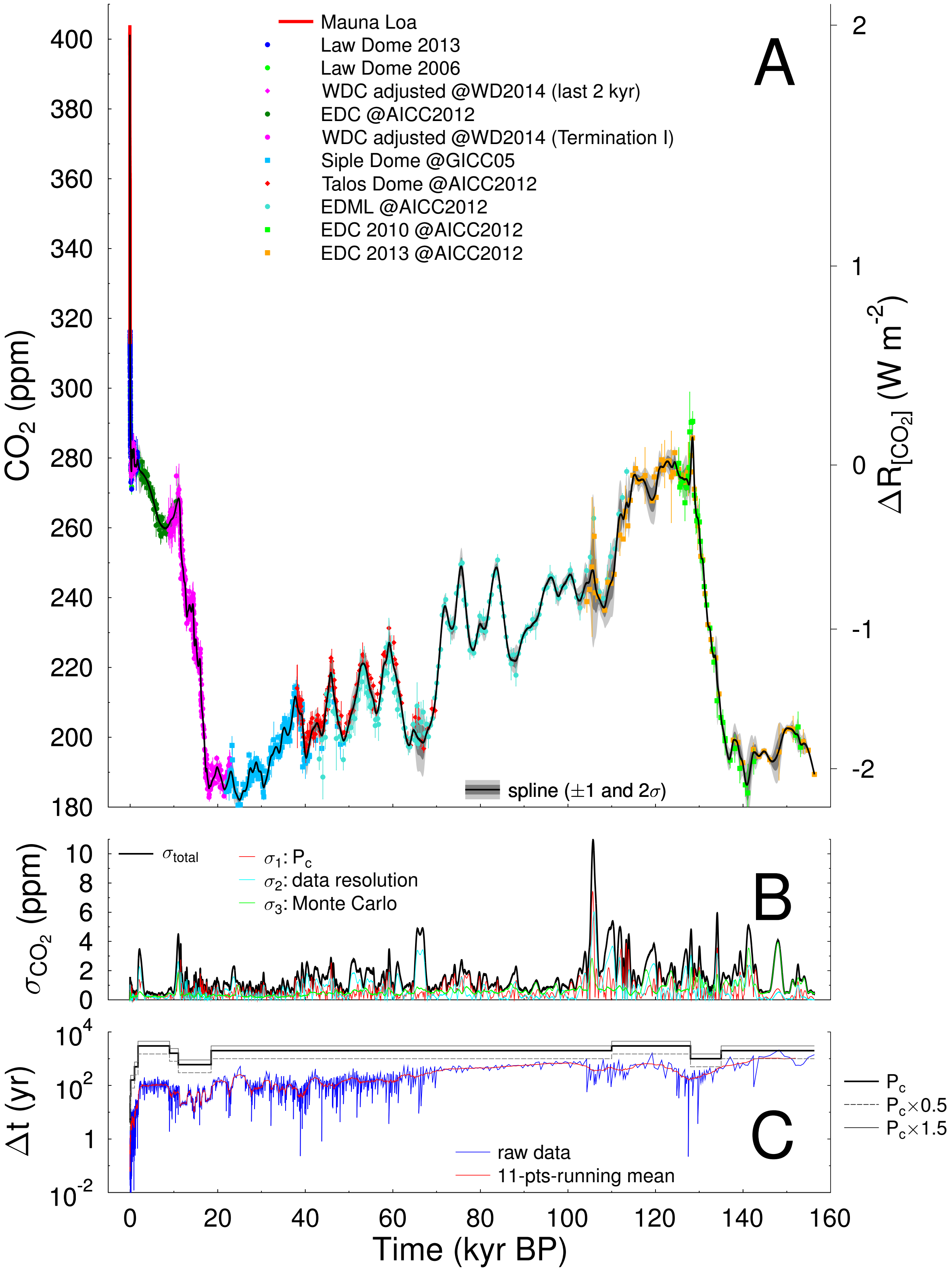
I was somewhat surprised to not see any uncertainty analysis carried out in this study. Varying the Pcutoff values and the data within their measurement uncertainty in a Monte Carlo analysis is relatively standard. Perhaps because the error in the final data product is the radiative forcing dominated by other uncertainties, this was deemed unnecessary? However, I don't see any error analysis carried over into the radiative forcing histories either (see comments below on radiative forcing).

Our reply: The splines are calculated with Fortran routines provided by Fortunat Joos, University of Bern. He is no coauthor of our paper, but agreed that we use his routines here. We are therefore not in the position to put them online, since we would then claim authorship for something, which has in the first place been developed by somebody else. Since the guide to authors for ESSD does not provide that we put all applied tools online we believe refusal of putting the routines online is acceptable (although maybe not desirable).

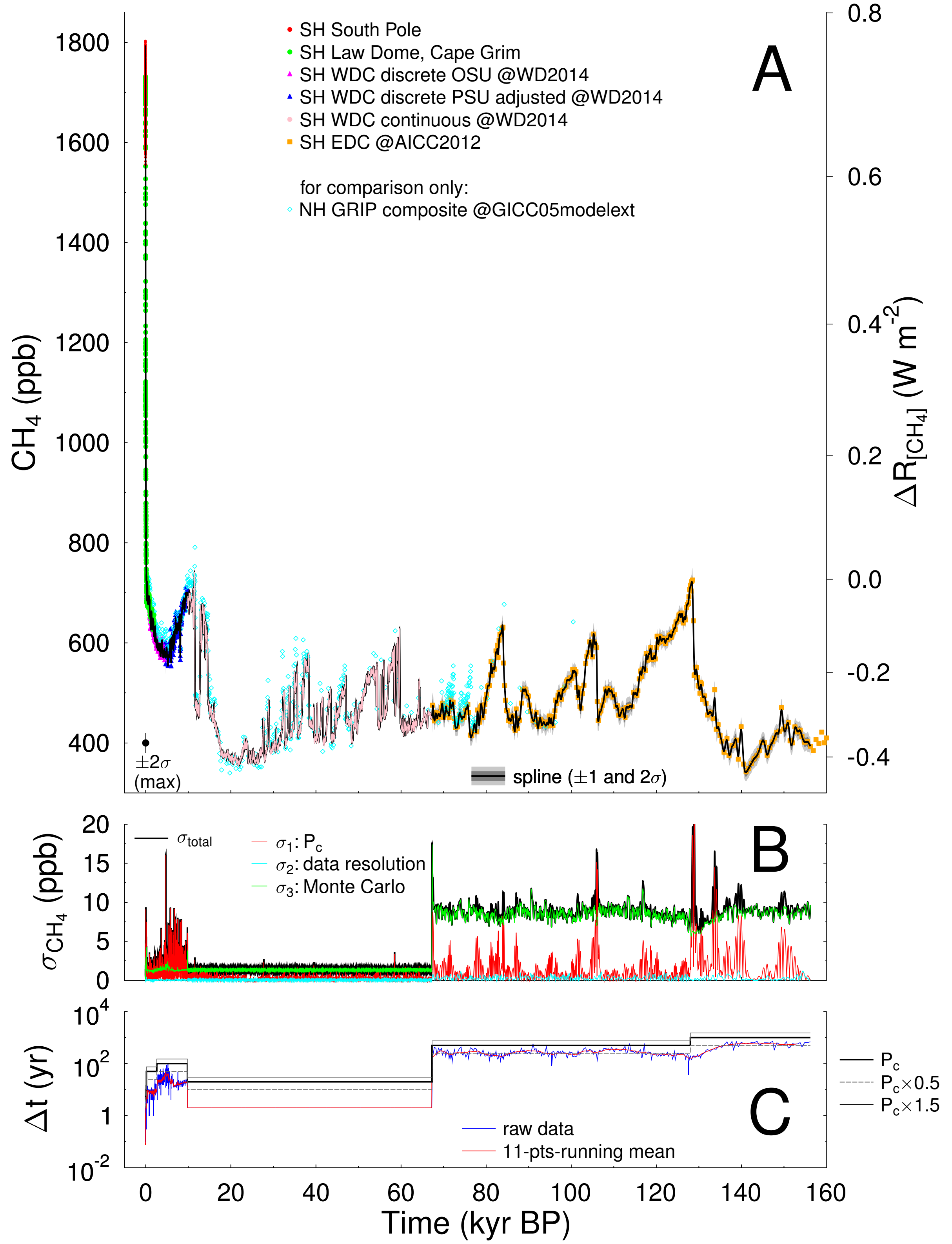
We will now contain an uncertainty analysis and will introduce a now sub-panel in Figs 1, 3, 5 with details on the spline errors. There are three different sources of error: (1) The cutoff periods P_{cutoff} are partially subjectively chosen. We will therefore calculate splines, in which P_{cutoff} is varied by $\pm 50\%$ and will calculate the mean difference from these two additional splines from our standard spline as σ_1 related to P_{cutoff}. (2) Monte Carlo statistics will be applied in the revision. However, please note that for the spline calculations not only the the mean values y_i of each data point i, but also their (1-sigma) uncertainties v_i are used. If we now vary the selected values supporting a particular spine within the uncertainty range, as typically done in Monte Carlo statistics, the uncertainty can then no longer be used for the spline calculations because the allowable range of a data point would then shift also. Take, for example, the data point i with $y_i = 1$, $v_i = 0.1$. In the standard case, the 1-sigma data range of i would be 0.9–1.1. If in Monte Carlo we randomly select a value of 0.95, the 1-sigma data range seen by the spline would then shift to 0.85–1.05, which is clearly different from 0.9–1.1. We will circumvent this issue in the following way: We will in a Monte Carlo approach (with $n = 500$ repetitions (n has been tested to lead to converging results)) calculate values y_i by chance taken out of normal-distributed data using the given uncertainty range v_i. However, when calculating the spline the uncertainty of all points v_i will be largely reduced (artificially set to 0.01). This will then give us a second uncertainty σ_2 caused by setting v_i to a small value, and a third uncertainty σ_3 from the Monte Carlo statistics. The overall, or total, uncertainty will then be calculated from the square root of the sum of squares of the individual errors ($\sigma = \sqrt{\sigma_1^2 + \sigma_2^2 + \sigma_3^2}$).

To visualise the data uncertainty we will in a revision include the error bars of the data in the figure, whenever possible, or include typically size of the data uncertainties .

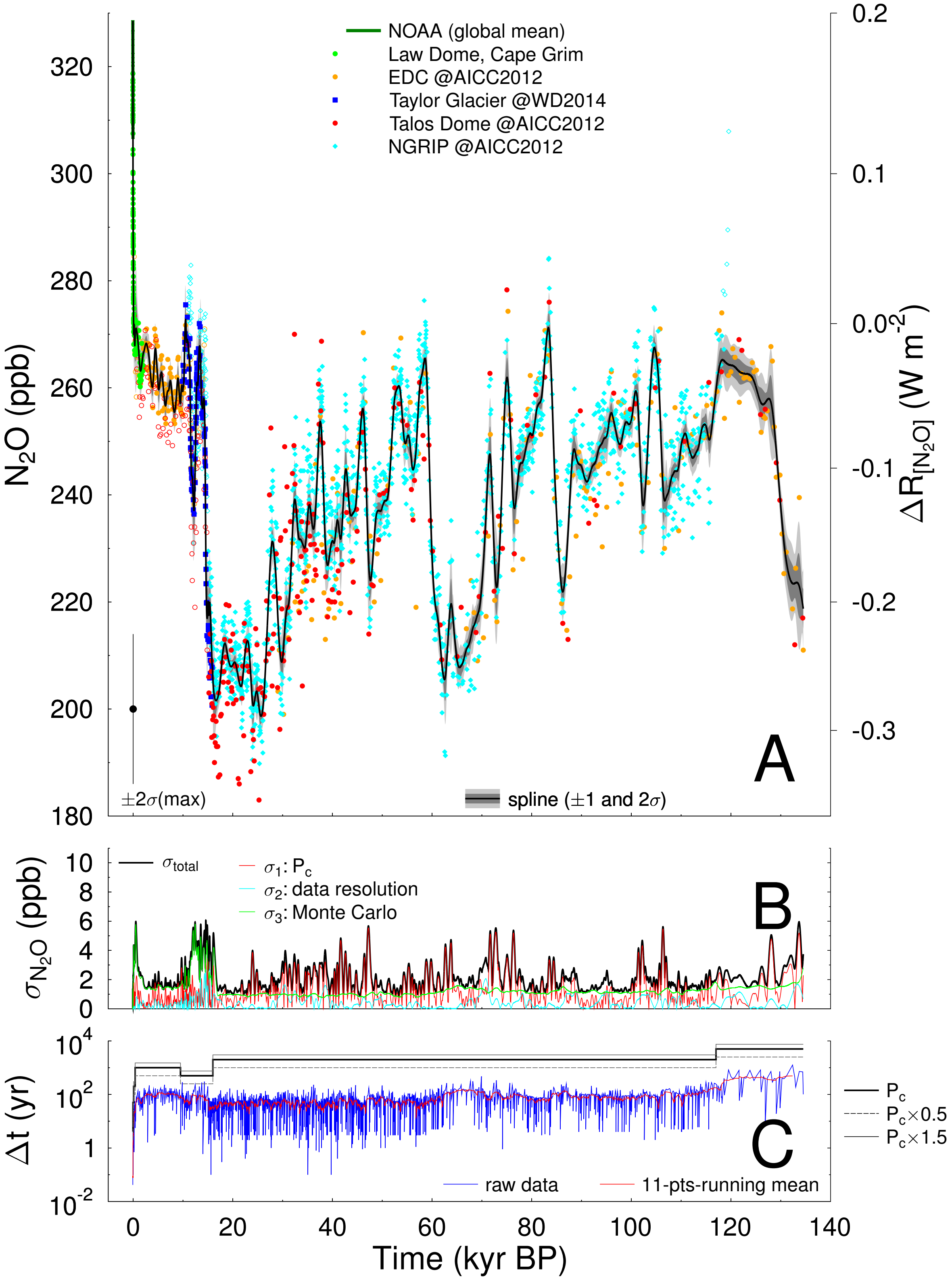
Uncertainty analysis for radiative forcing is of minor importance here, since full GCMs will use GHG time series as forcing and will calculate the corresponding radiative forcing internally. Thus, our calculated ΔR are only some first order simplified expressions, of interest to compare different GHG data sets and to be used in rather simple approaches. We will nevertheless use the uncertainties of the spline calculations also when calculating the radiative forcing, using the assumptions on uncertainties in ΔR as taken in Köhler et al. 2010 (QSR).



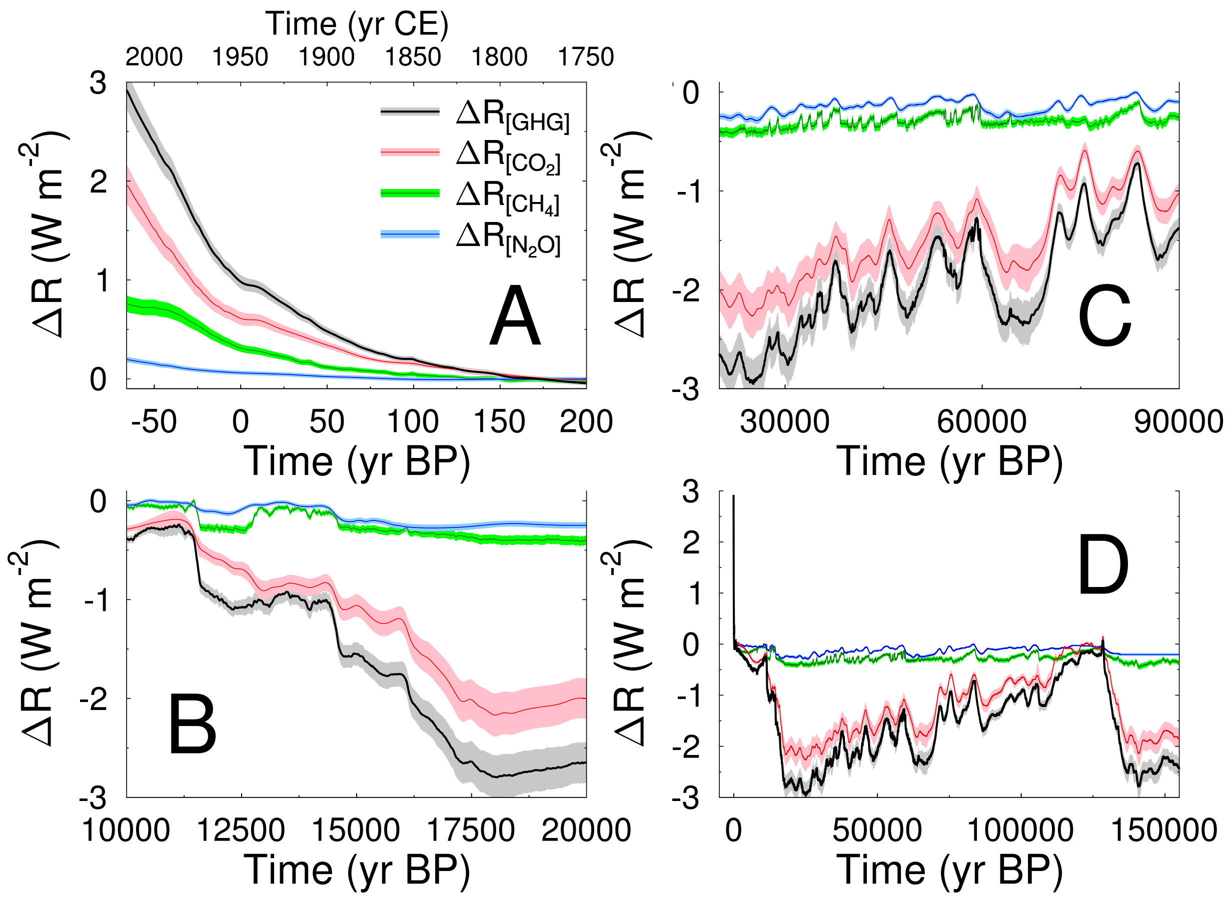
Revised Fig 1 with cutoff period ($P_c \pm 50\%$) used in the splines, also in comparison to the data spacing Δt. Error bars in (A) are $\pm 2\sigma$. An error calculation for the spline is now included (panel B). From panel (C) one can learn that even when we reduce the cutoff frequency by 50% we are well above the distance of the data points (apart from a short interval around 150 kyr BP). This implies that overfitting of the data with our spline can be excluded in the case of CO_2.



Revised Fig 3 with cutoff period ($P_c \pm 50\%$) used in the splines, also in comparison to the data spacing Δt. An error calculation for the spline is now included (panel B). From panel (C) one can learn that the cutoff frequency is at the limit of what the CH_4 data allow and a further reduction might indeed lead to overfitting. Note, this figure contains a revised calculation of the spline based on the continuous WDC CH_4 data, and changed cutoff periods.



Revised Fig 5 with cutoff period ($P_c \pm 50\%$) used in the splines, also in comparison to the data spacing Δt. An error calculation for the spline is now included (panel B). From panel (C) one can learn that even with a reduced cutoff period we are technically well above the distance of the data points and therefore technically no overfitting is occuring.



Revised Fig 6 including uncertainty estimates $\pm 1\sigma$. Note, that when calculating the radiative forcing out of the GHG concentration by the simplified equations adds another 10% relative uncertainty to ΔR (IPCC 2007).

Calculating radiative forcing. The radiative impact of the greenhouse gases are calculated based on Myhre et al., 1998. First, in all three equations presented in the text (5,6, and 7) all the terms need to be defined. Second, how does this approach differ from previous calculations of past radiative forcing (Joos and Spahni, 2008; Schilt et al., 2010)? I was under the impression that is it standard to include the interacting effects. A naive question: is there a set protocol for calculating radiative forcing from a body like the IPCC, CMIP or the PALMOD project that this study is focused on?

Our reply: The equations for the calculation of the radiative forcing due to GHGs are simplified expressions to be used in simple approaches or simple models. Full GCMs calculate radiative forcing of GHGs internally, and therefore slightly differ from each other and these simplified equations. Therefore, there is no protocol for these calculations. Our calculations are identical to those of Köhler et al. (2010). Joos and Spahni (2008) used the identical equations as in the present study. While Schilt et al (2010) used the same equations for CH_4 and N_2O, they did apply a different equation for CO_2. Note, that in the IPCC 2001 three alternative equations to calculate the radiative forcing of CO_2 have been summarised. In the present study and Joos and Spahni 2008 one of the other two equations stated in the IPCC 2001 is used. This detail, however, is not stated in the Schilt paper (only a reference to the IPCC is given, but not which equation has been used), and only known to the first author due to correspondence with A. Schilt in 2011 on this issue. Both Joos and Spahni (2008) and Schilt et al (2010) include the interacting effect of CH_4 and N_2O (which we ignore), but neglect the 40% increase in CH_4 radiative forcing due to indirect effects of CH_4 on stratospheric H_2 and tropospheric O_3. The difference between the equations used in Schilt and here (when ignoring the 40% interactive rise in CH_4) leads to offsets in the radiative forcing of less than 0.1 W m^{-2}, if based on the the same GHG data set. We agree with the reviewer, that it is more or less standard to include the interacting effects of CH_4 and N_2O, however since it is so small we prefer to ignore it in order to calculate the radiative forcing of the change in a single GHG without having the need to know details in the other GHGs. This choice is also motivated by the fact, that full GCMs (climate models) which might use our final splines to force transient simulations, will calculate their own radiative forcing internally. We will slightly change the wording in the text to clarify the here used equations. They contain only one free variable (the GHG of interest), constant values and standard mathematical functions together with units.

Comparison with Schilt et al., 2010. In a similar vein, I would strongly suggest adding a brief comparison with the results of Schilt et al., 2010, which produced a very similar data product of radiative forcing,. What new data has been added? What improvements had been made in chronology? Has the method for calculating radiative forcing changed?

Our reply: We disagree on the importance of Schilt et al (2010). With respect to earlier studies in which radiative forcing of GHG have been calculated the more complete previous studies were in our view either Joos and Spahni (2008) or Köhler et al (2010). Joos and Spahni (2008) was restricted to the last 20kyr while Köhler et al (2010) covered the last 800 kyr. Both approaches use the same equations to calculate the radiative forcing of GHG as applied here, but different GHG input data. As already mentioned above, full GCMs will calculate their own radiative forcing by using our GHG splines as input data. The calculation of radiative forcing performed in Schilt et al (2010) is problematic because of its vague documentation (see reply to last comment). Furthermore, for all GHG a reference state has to be selected, which is not per se defined and which is also not clearly stated in Schilt et al (2010). It is said, that changes with respect to year 1750 CE have been calculated, but not which GHGs values for that year have been assigned. Since the same equation for $\Delta R_{[CO_2]}$ are used here and in Joos and Spahni (2008) as well as in Köhler et al (2010) differences in the radiative forcing are direct consequences of the differences in the underlying data. The size of the interacting effect of CH_4 and N_2O is already estimated in the paper. Most important, every individual GCM will calculate own radiative forcing. Again, these calculations are only simplified equations of interest for some analytical calculations, but not for transient climate simulations using full GCMs.

One possible suggestion. As argued in this paper, data compilations like this are very useful for getting the best data into the hands of non-specialists. However, I've also found that a compilation can act as a stand-in for a substantial body of work and obscure the original research publications within the scientific literature. Whether because of convenience or unrealistic limitations on references in some journals, a compilation is often cited in place of the original work. It also seems to me that even when a new ice record supplants older data as the new reference record, that new work owes a great deal to everything that came before. I wonder if the paper would benefit from briefly mentioning a wider body of ice core gas research, perhaps as a brief introduction to each GHG section.

Within the relatively narrow scope of the paper (and journal) this doesnt seem completely necessary. But perhaps it is necessary in the sense that the reader cannot evaluate if the data being presented is truly the best by reading deeper into the literature. Since the reference list is nearly complete already, I think this would only involve mentioning a few of the very early studies. Again, just a suggestion to discuss.

Our reply: We agree with the reviewer opinion, but we believe this will difficult to achieve. It would shift our effort from a best guess data compilation to a full review paper. Furthermore, to give full credit to earlier studies we would need to cite more or less every ice core related GHG paper ever published. This was not in the focus of our study and we do not feel in the position to do so.

"ESSD Living Data". Are there plans to update the compilation as new data emerge?

Our reply: This is a good suggestion. Right now, no plans in that direction exist, but our understanding of this concept is, that an "ESSD Living Data" starts to be created in the moment the first revision (so the 2nd submission) of a data set is uploaded at ESSD. We will keep an eye on this option, but we are not in the position to make already now a clear statement, if this will happen in the future.

Line-by-line comments:

- P.1 L.22 "within sight" sounds a little too informal
  Our reply: We will change "within sight" into "achievable in the near future".
- P.2. L. 13 "extend" = extent
  Our reply: Corrected.
- P.3. L.10 last "or" = and
  Our reply: Changed.
- P.3. L.20 semi-colon is out of place, probably best to use and
  Our reply: The semi-colon is automatically generated with BibTeX, but we will change the sentence manually as suggested.
- P.3 L.26 Why isn't the cost function shown? Also, plural problem with "cost functions". The phrase "spline/second derivative" is very confusing, is it the curvature of the spline or is it the curvature of second derivative that is being minimized?
  Our reply: We do not show the cost function, that is minimised when calculating the

spline, because our focus is not on the method, but rather on the final spline through the data. Such technical details can be found in the cited papers, e.g. Enting (1987) or Enting et al. (2006). Plural in "cost functions" is set to singular. The second derivative which represents the curvature of the spline is minimised. We will change the text for clarity.

- P.3 L.27. Starting off this paragraph with a general description of Fourier analysis leads to quite a bit of confusion, as the spline fitting and Fourier analysis are very different techniques. I believe the authors were attempting an analogy with Fourier techniques but I found the results quite muddled. Also, stating "ideally, a smoothing spline removes all high frequencies (sic) sine functions" sounds strange given that the spline fit is not specifically designed to remove sine functions (as is Fourier analysis) but rather reduce the curvature of spline and the data-spline fit (as stated in the text). I would suggest refocusing the paragraph on the Pcutoff values and its relation to attenuating the amplitude of variability (at various frequencies) and then turn to an analogy with other techniques. I also would like to see a clear demonstration that the spline does indeed reduce amplitude by 50% at the Pcutoff frequency or have a stronger reference attached to this statement so the reader knows where to look. Visually, it would be useful to have a figure where the effect of the spline fitting is displayed on artificial time-series (for instance a set of sine curves possibly including some with red noise), but this is just a suggestion.
  Our reply: We do not understand the comment on Fourier analysis here. Since the spline smoothing is performed in the frequency space with the target to dampen higher frequencies and therefore to act as a low-pass filter, Fourier analysis is part of it. Since the underlying technical details of the spline methods have been described various times in other papers, we refrain from going to more details here and follow also not the suggestion of showing a figure in which the effect of the spline routines is visualised. The 50% amplitude attenuation comes from the definition of the cutoff frequency, and can, for example, in more detail be found in Enting et al (2006). Nevertheless, this paragraphs will be revised for clarity.
- P.5. L.15 "goes back in time until" = extends to
  Our reply: Changed.
- P.6 L.12 Simply quantifying the offset between WDC and EDC would be sufficient here. The sentence is a bit convoluted at the moment and needs to be rewritten.
  Our reply: We will split the sentence in two and simplify for clarity.
- P.7 L.5 Need to explain all the terms in this equation
  Our reply: This comment refers to Eq 5, which is an expression of a rough estimate of the radiative forcing of changes in CO_2:
  $\Delta R_{[\mathrm{CO_2}]} = 5.35 \cdot ln(\mathrm{CO_2}/(278 \text{ ppm}))$ W m^{-2}.
  The right-hand-side contains only one variable (CO_2), all else are either constants (5.35, 278), mathematical standard functions (ln) or units (ppm, W m^{-2}), which have been derived in the corresponding underlying paper (Myhre et al., 1998). So, we can not see what we need to explain here. However, we will revise the sentence around this equation to clarify its meaning.
- P.7 L.22 Need a better reference to the uncertainty than "the NOAA website". At least a URL.
  Our reply: On https://www.esrl.noaa.gov/gmd/ccgg/about/co2_measurements.html one can read, that the accuracy of the Mauna Loa CO_2 data has been estimated to be in general (but not necessarily always) better than 0.2 ppm. To be on the save side, we

have chosen to use a slightly higher uncertainty of 0.3 ppm for those data. However, in comparison to the uncertainty in the ice cores data used here, this data uncertainty of the instrumental record is very low, and the data coverage high (monthly means), and a cutoff period of 4 years assigned to this time interval. A revision in assumed uncertainty for the Manua Loa CO_2 data to the stated 0.2 ppm would change our final smoothing spline only very slightly. The text will be extended on these details.

- P.7 L.29 "the knowledge of abrupt changes" why not something like "to preserve the rapid variability described in Marcott et al.,"
  Our reply: We will revise the sentence for clarity.

- P.7 L. 30 use of "could be" makes it sound like this is simply a possibility and not what was actually done.
  Our reply: We will change "could be" into "has been".

- P.9 L.14. "The interhemispheric gradient of the NH...." Something is missing in this and the following sentence. Also, please state exactly how the 0.1 W m^{-2} calculated even if it is straightforward. Are models sensitive to the spatial pattern of radiative forcing from GHGs? The earth must be sensitive to the spatial pattern of (the much larger) changes in insolation otherwise we wouldn't have glacial-interglacial cycles so why not GHGs?
  Our reply: The sentences around the interhemipsheric gradient will be revised and extended for clarity. We will in detail explain, how this radiative forcing difference between N and S of 0.1 W m^{-2} has been calculated. Models are indeed sensitive to spatial pattern of GHG concentrations. However, one needs to be aware, that the radiative forcing of the GHG, which we here calculate by simplified expressions, is in most cases and in most models calculated internally in their radiative schemes. So, our approximations of ΔR are only some first steps, to be used for comparisons and simply applications.

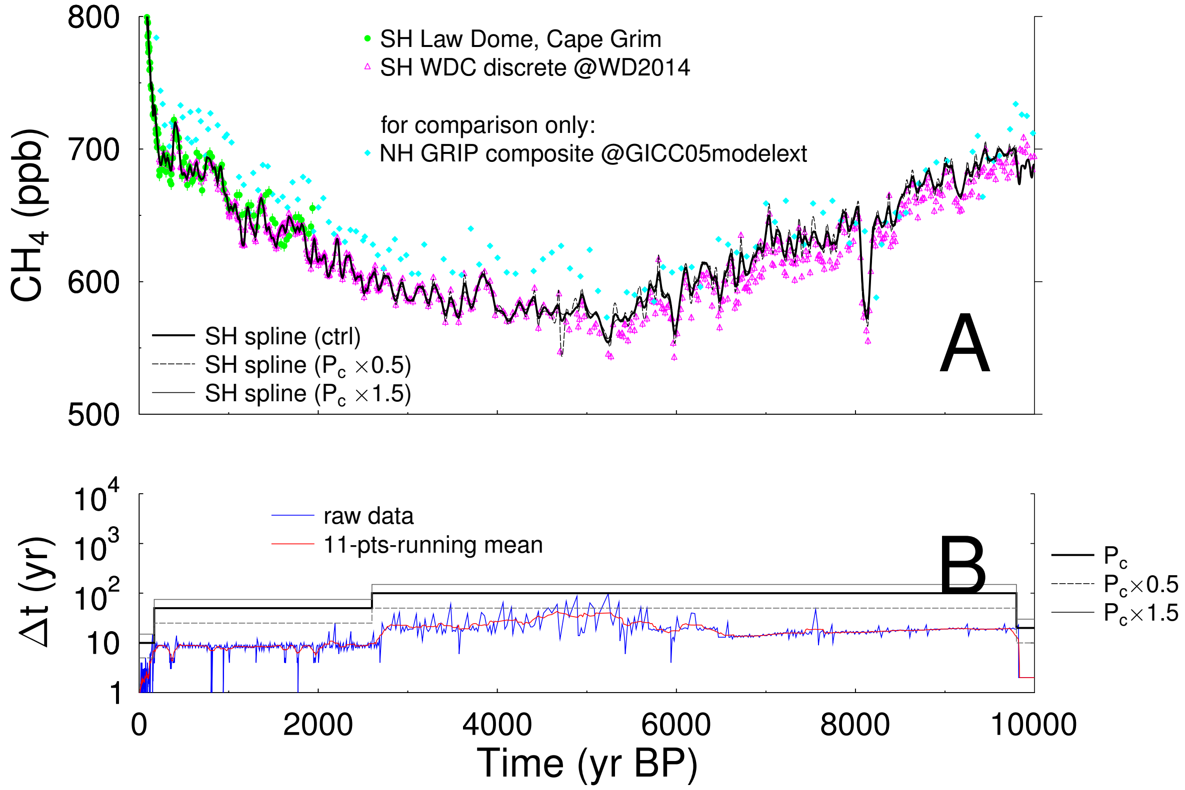
- P.9 L.22 It seems to me there are some offsets between WDC and Law Dome from 0-500CE but this could be due to age model problems.
  Our reply: We will briefly mention this offset between WDC and Law Dome, but no adjustment is necessary, since WDC is higher resolved than Law Dome and solely taken in support for the spline.

- P.9 L.24 I have already suggested to the authors in a correspondence that they should consider using the WDC continuous CH_4 data (Rhodes et al., 2015) as their primary CH_4 record from the early Holocene to the last glacial period. Not only is the highest-resolution data available, but it has already been processed by the original authors to be near continuous (2 year resolution). The authors may be interested in this recently published paper on the reproducibly and nature of centennial-scale features in the continuous dataset (Rhodes et al., 2017).
  Our reply: We thank for this suggestion of the continuous WDC CH_4 data. This has now been taken in the revised calculation of the CH_4 spline.

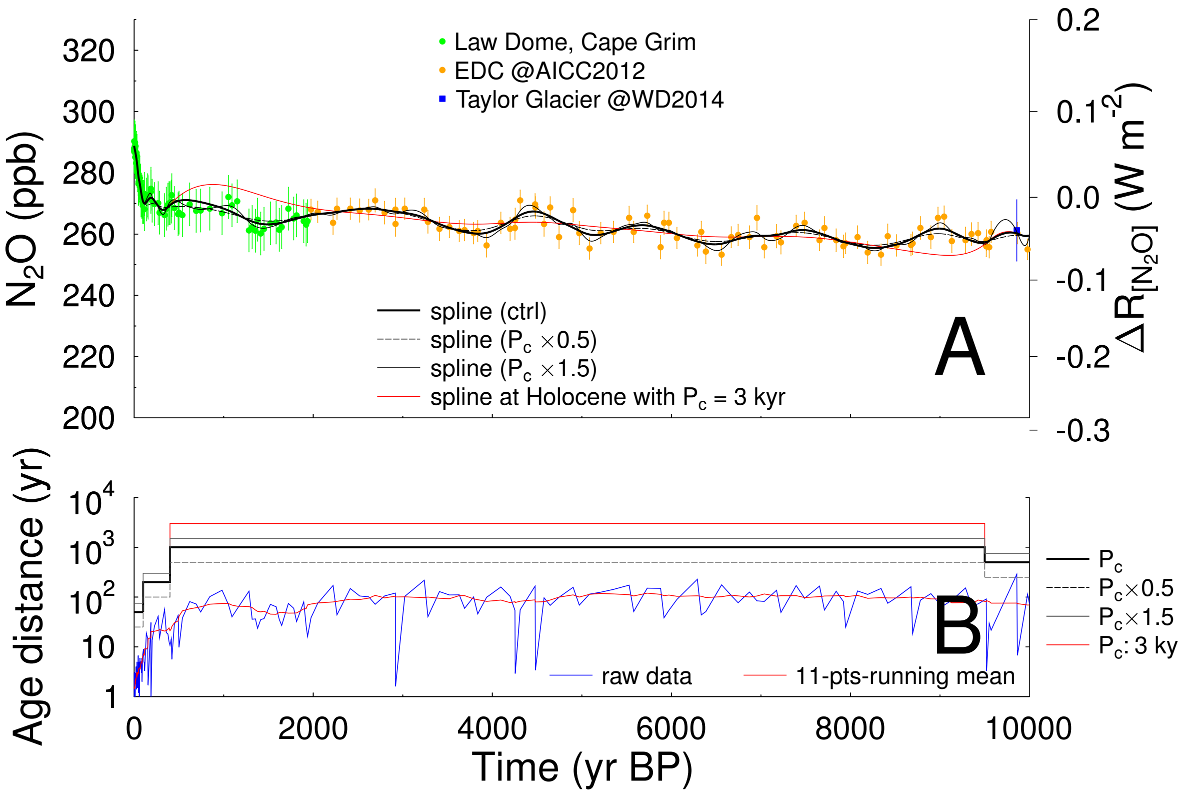
- The other thing to note regarding the WDC CH_4 is that the complete record is actually the combination of two different labs: Oregon State (late-Holocene, deglaciation) and Penn State (early-Holocene and last glacial period). I would suggest the authors correspond with these groups to make sure the datasets provided online have been harmonized for any small differences in procedural corrections.
  Our reply: Thanks for the suggestion. Since we now use the continuous CH_4 spline between 9.8 ka and 67 ka, the discrete CH_4 data from WDC are only used in the

Holocene. We have taken our WDC CH_4 data from the online available combined data set from both labs Oregon State University (OSU) and Penn State University (PSU) (https://www1.ncdc.noaa.gov/pub/data/paleo/icecore/ antarctica/wais2015ch4osupsu.txt). After checking for offsets between data from both labs in this data set we realised that no harmonisation of the data have been performed so far. We therefore adjusted the PSU data by +9.9 ppb, which is the stated, (but yet unexplained) offset between both labs, found in the SI to Rhodes et al (2015) Science, which describes to continuous WDC CH_4 data.

- Finally, it is not clear whether the spline fit to the CH_4 data in the early-Holocene have been overfit from the figure provided. In fact it looks like there is a decrease in the magnitude of variability around 4ka, which is right when the Oregon State data (Mitchell et al., 2013) starts. A clearer figure spanning the entire Holocene, a comparison with the continuous data, or references that described this high-frequency variability in the early Holocene are needed to convince the reader that the spline has faithfully captured real variability.
  Our reply: In the process of recalculating the CH_4 spline by implementing the continuous CH_4 data we slightly revised the cutoff periods in the Holocene to avoid overfitting. Below you find a clearer picture of the Holocene CH_4 data. The continuous WDC CH_4 data the reviewer is referring to start only at 9.8 kyr BP, so they can not be used for comparison in the Holocene. Also note that the cutoff period of our spline is now 100 years (has been 50 years in initial submission) between 2.6 and 6.5 kyr BP, but 50 years (has been 20 yr) in younger times (see Table 6 for further details). Our understanding is that these are now centennial-scale changes in CH_4, which have not been visual before due to measurement uncertainty and which are understood to be of climatic origin (e.g. Mitchell et al. 2013). In this plot below we see no sign of overfitting. Note, that the continuous CH_4 data have a temporal resolution of 2 years and are already a spline. We nevertheless used a cutoff period of 20 years here (9.8-67 ka), which is outstanding small (compared to the other ice core GHG data) to reduce still existing noise in the data.



- P. 10 L. 21 Regarding the reference to Ivanovic et al., 2016, my understanding is that those experiments are transient and thus need continuous data. I think a brief but closer comparison with the data in Ivanovic is warranted given that they basically used WDC CO_2 with EDC CH_4 (despite the obviously higher-quality of the WDC CH_4).
  Our reply: Ivanovic et al. (2016) has chosen to base transient simulations on the Bereiter et al. (2008) CO_2 compilation, Loulergue et al. (2008) CH_4 data (EDC) and the Schilt et al. (2010) N_2O compilation. This is just a different selection of GHG data sets. We will briefly mention that differences exits, but this additional information will be included in the revised paper at the point, when Ivanovic et al. (2016) is first cited in the introduction. We like to highlight that the paleo runs within PMIP4 specifically allow to use other pre-processed data sets, if well motivated.
- P.11 L.1 Please reword some of the informal terms (e.g. "get...right" = obtain accurate....)
  Our reply: Done.
- P11. L.3 Remove "again".
  Our reply: Done.
- I was surprised to see the spline fits capture high-frequency variability in N_2O during the Holocene. The Pcutoff in this study is set to 1000 years, yet the original data publication (Flueckiger et al., 2002) use a cutoff of 1500 years (only very tentatively) and 3000 (deemed the robust features). It seems this paper have overfit the data here. This and possibly some of the early-Holocene CH_4 data are the only places I could see the potential for improper use of the spline fitting procedure.
  Our reply: Our choices of Pcutoff was motivated by the data coverage, or age distance of neighbouring points, and if higher frequency changes are expected for a data set in a specific time window. This leaves some spaces for subjectivity. We have overlooked that Flueckiger et al. (2002) already made an expert judgement, on which cutoff frequency might produce a spline which might contain robust features. As you can see in the figure below, the data coverage of the Holocene clearly justifies our chosen cutoff period of 1000 years in the Holocene. Furthermore, if we would use a cutoff period of 3000 yr during most of the Holocene, we would not only average over some of the variability, which is contained in the data, but we would also produce a spline with some unsatisfactory edge effects during the transitions in and out of the Holocene. Such edge efforts might occur during the transition window, but need to be avoided by a careful selection of cutoff periods and the times of cutoff transitions, which we have done here.



The figure above contains a zoom-in on the N_2O during the Holocene on the question of overfitting the data by a spline with cutoff period of 1000 yr.

- Overall, I was less convinced by the (difficult) choices the authors made regarding the N_2O reconstruction. For all intents and purposes, the authors have averaged a number of different ice cores that appear to each their own unique biases. Given that Schilt et al., 2010 focused on reconstructing N_2O and subsequently produced a very similar set of radiative forcing histories, I would like to see a brief comparison with further justification for the inclusions of the NGRIP data. It seems clear that quite a bit of work on reconstructing N_2O is still needed.
  Our reply: Our N_2O compilation is not so much different from the one compiled in Schilt et al. (2010). We added additional NGRIP data points, since new NGRIP N_2O data have been published in 2013 (Schilt et al., 2013), and decided then to rely Greenland N_2O data solely on NGRIP, while in Schilt also data from GRIP and GISP2 are plotted. We furthermore have chosen to use from the Antarctic ice cores only data from Talos Dome and EDC, while in Schilt also EDML N_2O data have been plotted. However, since EDML N_2O largely agree with the EDC N_2O data no additional knowledge is added by them. Some Greenland data (here NGRIP) are necessary, since data gaps are to large, when the N_2O data compilation is solely based on Antarctic ice cores.

- When discussing the differences between NGRIP and the Antarctic cores, it should be noted that such a large interhemispheric gradient is impossible to maintain given the residence time of N_2O in the atmosphere ($\sim$100 years).
  Our reply: We are not sure one what this comment is based on. The difference between NGRIP and Talos Dome for individual times is around 10 ppb. However, the individual data points uncertainty is up to 7 ppb (1σ), so such N-S offsets would within the uncertainties still not point to a large interhemispheric gradient. Furthermore, at the beginning of this section, when explaining the instrumental N_2O data from NOAA we already point to the fact that due to the large atmospheric lifetime of N_2O no interhemispheric gradient is expected in this data set.

- P. 12 L.24 remove "also"
  Our reply: Done.
- P.15 L.10 Baggenstos, D., reference can updated to a paper in discussion in Climate of the Past
  Our reply: Since the discussion phase of this paper is already finished and the reviews suggest only minor changes we agree that the reference to the PhD thesis of Baggenstos should indeed be revised to this paper submitted to Climate of the Past. Hopefully, the full citation of the discussion paper might be available when we correct our final page proofs.
- P.15 L. 23 CO_2 subscript
  Our reply: Corrected.
- P.15 L. 32 "&ndash" is an incorrect
  Our reply: Corrected.
- P.17 L.10 Subscript problem
  Our reply: Corrected.
- P. 17 L.19 Is Meinhausen now accepted?
  Our reply: Accepted in final version, but not yet available. We will revise this reference in our final version.
- P. 18 L.19: N_2O subscript
  Our reply: Corrected.
- Table 2: In the reference to Monnin et al, 2004 for Holocene CO_2 data, it says only EDC data was used. This paper also presented EDML and Taylor Dome CO_2, so I just wanted to check that these data were excluded from the Holocene CO_2 record.
  Our reply: Correct, only the EDC CO_2 data from the Monnin et al, 2004 paper are considered here, not those from EDML or Taylor Glacier.
- Table 4/6/8: Pcutoff (the realised cutoff period) should be defined in the text.
  Our reply: We will extend the description of the spline and will also explain the realised cutoff period in more detail.
- Figures mix "yr BP" and "kyr BP". They should be made consistent, ideally with all the conventions in the manuscript. Given the use of both CE and BP (and the more confusing B2K used in some ice core publications), a definition of BP is needed in the main text. This can sometimes be confusing for researchers outside the paleoclimate field, who will likely be interested in this paper.
  Our reply: We will include a definition of BP and CE as suggested. However, we will stick to our usage of 'either "yr BP" or "kyr BP", since this depends on the time window of data shown and the space available for the labels of the x-axis.
- Data files on Pangaea: Methane is misspelled in the data labels as Methan
  Our reply: Error will be corrected in revision.
  To do: To be corrected.