Reply to the referee comments of Sapart et al., CH_4 isotopic studies from firn air at 11 sites

We thank both reviewers for the relevant and constructive comments, which we used to improve the quality of the paper. Below is our point-by-point reply, but first a general comment.

Our paper has two major goals: 1) To investigate whether a consistent $\delta^{13}CH_4$ history over the last 50 years can be constructed by combining firn air measurements from multiple sites in Greenland and Antarctica. For this purpose, we use the only existing firn air model to date that allows reconstructing atmospheric time trends from isotopic data at a large number of firn drilling sites ("multi-site" inversions). 2) To highlight the uncertainties in isotope reconstructions from firn air when the atmospheric signal is of the same order as the firn fractionation effects, as in the case of $\delta^{13}CH_4$. Moreover, recent investigations in several laboratories including IMAU show that Krypton ions interfer with $\delta^{13}CH_4$ and this may likely be an additional cause of the discrepancies observed between the different firn datasets.

As the referee comments highlighted, several processes cannot be quantified within the state of the art on firn modeling (e.g. dispersive mixing and 3D transport). It is clearly beyond the scope of our manuscript to improve the physics included in existing firn models. The model used here has performed well in many previous studies (e.g. Wang et al., 2012, Buizert et al., 2012 and Witrant et al., 2012), and was compared to other state of the art models in (Buizert et al., 2012).

Considering the comments of the referees we clarified our main goal in the revised version of the paper and we propose a new title.

Reply to anonymous referee 2:

The manuscript describes the most elaborate study of d13C of methane from firn air to date. The authors combine records from nine different firn air sites, and use these to derive a record of d13C for both hemispheres. The topic is of importance to both our understanding of the global methane budget, as well as firn air transport processes. The manuscript is well written, and manages to clearly explain the complexities of firn fractionation and how it impacts their ability to reconstruct atmospheric d13C.

12) The rate of d13C change the authors derive for the recent atmosphere differs considerably (>50%) from earlier estimates. The authors do discuss the discrepancy with earlier estimates from firn and atmospheric measurements, but do not show which rate is the correct one. The authors also convincingly show that the large discrepancy between their sites is related to the firn fractionation, and uncertainties in the reconstructed diffusivity profiles. These two things combined leave the reader unnecessarily confused, and in doubt that atmospheric d13C can be reconstructed from firn air in the first place.

***Answer 12*:**

In the work described in this manuscript, we attempt for the first time to derive a consistent atmospheric $\delta^{13}CH_4$ history from all available firn air data. Indeed the reconstructed trends are different from earlier publications, but the deviations are not in all cases very large. Table 2 shows the $\delta^{13}CH_4$ rates of change from our study in comparison to values reported in the literature. The largest discrepancy is with the reconstruction of Sowers et al., 2005 which shows a trend three times larger than our reconstruction for the period 1995-2001. The differences in the agreement with previous studies also translate to inconsistencies in our own isotope reconstruction between individual sites, which leads to rather large error bars in the multisite inversion. Based on this analysis, we then discuss the possible factors that cause uncertainties in the $\delta^{13}CH_4$ reconstructions from firn air measurements. Given these uncertainties, we refrain from interpreting our $\delta^{13}CH_4$ trends in terms of rate of source and sink changes. For example, the uncertainty of the methane mixing ratio scenario used already induces a large uncertainty in the rate of change of $\delta^{13}CH_4$ (see Supplementary Section 2).

Reference period	***1995-2001***	***1984-1999***	***1978-1996***
Francey et a., 1999			***0.6 ‰***
Braunlich et al., 2001		***0.6 ‰***	
Sowers et al., 2005	***0.35 ‰***		

This study, from firn data	***0.11 ‰***	***0.37 ‰***	***0.46 ‰***

Table 2: Comparison of $\delta^{13}CH_4$ trends from our study with published values from atmospheric data (Francey et al., 1999) and firn air (Bräunlich et al., 2001; Sowers et al., 2005) in the Southern Hemisphere. The values from this study derived from firn air data were calculated from the scenario shown in red on Figure 5d of our manuscript.

13) Unfortunately I cannot agree with the main conclusions the authors derive from their data and modeling efforts. I believe there is a fundamental error in the methodology, and given all the presented data and model output, I come to a different conclusion. I believe the work should be published in ACP eventually, given the importance of the topic and the amount of data and model work compiled in this study. I am certain this work will make a valuable contribution to our understanding of firn transport, and limitations therein. First a major revision of the work is needed, along the lines detailed below.

My main concern is the implicit assumption in the work that by averaging results at different sites one would obtain a more correct estimate of the true atmospheric variations. This would be correct if the histories reconstructed for the individual sites were consistent with each other – in which case averaging would make the final result more robust against measurement and sampling errors at the individual sites. In this study the authors find huge discrepancies between the single-site reconstructions that some- times exceed the estimated measurement uncertainties by a factor of 10 or more. This hints at a problem with the reconstruction method, or perhaps calibration issues between the datasets from different labs. Given that the firn fractionation is so problematic, I think the right way to proceed would be to try and understand the firn fractionation better, and to see whether consistent site histories could be obtained using different assumptions about the firn gas transport. We know that the individual reconstructions are unreliable, because they give inconsistent results. By taking the average of two unreliable histories, one does not obtain a reliable history. An erroneous firn correction would probably bias all the individual site reconstructions in a similar manner which cannot be corrected for by averaging between sites. The available atmospheric records (NOAA-ESRL, Cape Grim archive) are inconsistent with the obtained reconstruction in both hemispheres: in the NH the monitoring data show no downward trend for the last 10 years, and in the SH the isotopic trend in the reconstruction has only half the slope of the direct observations. The atmospheric monitoring data are not subject to the uncertainty of the firn fractionation, and should be considered more reliable.

__Answer 13__:

We clearly state in our paper that the discrepancy to other datasets (monitoring and ice core air) is a concern. We also discuss in much detail that uncertainties in the firn fractionation used in firn models could contribute to these differences. It is clear that future work on firn physics effects as suggested by both referees is necessary to ultimately resolve this issue.

As an immediate reaction to this referee comment that could be carried out with adequate effort, we reconstructed a "best scenario" using DEO8, Cape Grim archive and atmospheric monitoring data in the revised version of the Supplementary Information (see answer to comment 17). This was then used as input in the forward model to see how well it fits all firn sites (see revised supplementary material). As expected, and consistent with our findings, this approach does not lead to a good fit of the firn data, especially at two early drill sites (Fig. 2). This may be due to the lack of constraints in the deep firn, because the sampling resolution of the DEO8 firn is low compared to other sites. Unfortunately, the existence of several important sources of uncertainty of different nature makes it difficult to understand the $\delta^{13}CH_4$ trends better than indicated by our large error bars: These uncertainties arise from the scenario used for methane (Supplementary Section 2), from uncertainties on the firn diffusivity (e.g. manuscript Section 5.2) and from experimental uncertainties (see below).

14) The authors discuss the differences between their reconstruction and previous ones, but do not make any statements about which reconstruction is more reliable. For this reason I think the work as it stands adds more confusion than it resolves.

__Answer 14__:

It is one of the results of our study that there are still unresolved issues regarding the reconstruction of $\delta^{13}CH_4$ records from firn air. Therefore, we cannot yet identify the most reliable record, but make some first steps in this direction. For example, we suggest that the most reliable sites are the sites with diffusivity profiles constrained with the largest number of species. One paragraph has been added to the discussion to

clarify this aspect and one section about uncertainties has been added to the revised supplementary material.

As an additional complication (and maybe part of the solution), during the review process of our paper, it has been discovered, with contributions from the IMAU research group, that $\delta^{13}CH_4$ measurements in several laboratories worldwide are affected by a Krypton (Kr) interference (Schmidt et al, submitted to AMT, 2013). This interference affects measurements of samples with varying CH_4/Kr ratio, for example firn air samples where the CH_4 mixing ratio strongly varies over the last century, and the Kr mixing ratio stays constant. The corresponding correction for the Kr interference is system-specific, and the samples measured at IMAU have been corrected for the Kr interference in the revised version of our manuscript. Based on NEEM-09, NEEM-08 and NGRIP measurements analyzed in the IMAU lab, a new NH reconstruction, corrected for the Kr interference (added in the revised supplementary material) has been build. However, the analytical system at LGGE is no longer in use and it is therefore not possible to correct the older firn air record for the Kr interference. Although the details of this interference will be published separately (Schmidt et al., manuscript in preparation, 2013), an explanation about the Kr interference has been added in the discussion section of the revised manuscript.

15) I would also like to see a more thorough discussion of calibration issues between the different datasets. The authors are trying to resolve small atmospheric trends, and the sites cover different time intervals. Therefore small calibration issues will influence the observed trends. The information given now on calibration issues (Page 9591 lines 21-27) is limited and incomplete. For example, how does the CIC and NOAA-ESRL data relate to IMAU? E.g. in Figure 6a NOAA-ESRL data is introduced, which appears to be isotopically lighter than all the firn data. Most puzzling to me was the following inconsistency: on page 9591 it is claimed that CSIRO and LGGE have no systematic differences, and that IMAU is corrected to be consistent with LGGE (implying IMAU is consistent with CSIRO); yet on page 9602 there is an 0.28 permil offset between IMAU and PSU, while PSU is consistent with CSIRO Cape grim and law dome records (implying IMAU is 0.28 permil offset with CSIRO). It becomes very difficult to see how

proper intercalibration is guaranteed. In many places the authors use words such as "possibly" (P9604) and "indirect evidence" (P9603) when discussing the calibration scales. Would it be possible to provide more clarity? Do the SH zero depth firn air samples (i.e. atmospheric measurements) match the reconstructed d13C history and monitoring data?

<u>Answer 15</u>:

No large-scale intercalibration exercise has been carried out between LGGE, IMAU, CSIRO, CIC and NOAA and since many of the firn measurements have been carried out long ago, such an exercise is not possible. However, intercalibration exercises have been performed in the past between CSIRO and LGGE and more recently between CIC and IMAU. Moreover, IMAU re-measured some firn bottles previously measured by the LGEE to evaluate the possible scale offset. For clarity on the intercalibration issue, a table showing the differences between the different measurement systems and the uncertainties associated has been added in the revised Supplementary Information. As stated in the previous answer, besides systematic scale shifts, the Kr interference likely induces additional discrepancies dependent on the depth in firn for which differences between labs have not yet been evaluated.

16) The technique the authors use is ultimately a statistical one: data from different sites are weighed by their uncertainty, and averaged to constrain the problem. I think some statistical tests would therefore be in order to assess the robustness of the final product; in particular bootstrapping and jackknife tests. Looking at the NH reconstruction (Fig. 6a) it appears to me that the NGR dataset is isotopically heavy. How does the reconstruction respond if it is left out? I suspect the downward trend after 2000 would disappear (consistent with the direct NOAA-ESRL data).

<u>Answer 16</u>:

The suggested statistical tests: bootstrapping and jackknifing are re-sampling techniques. An additional re-sampling method (cross validation) was recently implemented in our model, together with a more rigorous isotopic inversion method based on a forward model written in "δ" unit (Witrant and Martinerie, in press, <u>http://www.gipsa-lab.grenoble-inp.fr/~e.witrant/papers/13_delta_sssc.pdf</u>).

Our main results are not significantly modified. The slight (well within error bars) downward trend after 2000 in the North Hemisphere is also present on the single-site scenarios based on NEEM-08 data (from both IMAU and CIC) but not on the smoother scenario based on NEEM-09 data (see Figure 2a). It disappears from both single-site and multi-site scenarios when krypton-corrected data are used (see answer to comment 15). This illustrates the fact that detailed features well within the error bars are difficult to interpret as multiple sources of uncertainties co-exist.

17) I would like to ask the authors to consider, and comment on, the following alternative conclusion from their dataset. I suspect they may have thought of this themselves; I would be very interested in their response.

There are four lines of evidence in the paper that DE08 is the most reliable of all the sites, because the firn fractionation is smallest there: 1. DE08 has very high accumulation, which gives the gases little time to fractionate diffusively, 2. The LGGE-GIPSA and CSIRO models give comparable values for the DE08 firn fractionation (Fig 7), 3. Changes to the diffusivity profile do not influence the calculated fractionation much (Section 5.3), and 4. The DE08 single site reconstruction agrees well with the Cape Grim air archive (Fig 6b). For these reasons I have more confidence in the combined Cape Grim/DE08 firn and ice reconstruction than in the multi-site reconstruction presented in the MS. Would it be possible to use the DE08/cape grim d13C record as a constraint in the diffusivity reconstruction? The authors mention this technique was applied successfully for DML (page 9598, lines 20-28).

Answer 17:

Following this comment, an atmospheric scenario was built using atmospheric and DE08 data (Figure 2, left panel) and used as input to the LGGE-GIPSA forward model of gas transport in firn. Note that NIWA data was preferred to NOAA data here because inter-calibration tests showed no significant discrepancies between LGGE, CSIRO and NIWA data (Aballain, 2002, page 82), whereas no direct comparison with LGGE data is available for NOAA data (a comparison of NIWA and NOAA atmospheric data series suggest a ~0.2 ‰ shift). When applied to other firn sites, this scenario shows important discrepancies with firn air data especially at the two earliest drilled sites besides DE08: South Pole 1995 and Vostok (also sampled in 1995).We agree with the referee on the fact that using one or a few high accumulation rate site(s) to build a reference $\delta^{13}CH_4$ atmospheric trend (at least two sites, one from each

hemisphere are needed) and better constrain the diffusivity at other sites is a promising line of research, as mentioned p9599 l3-8 and p9605 l14-19. The DE08 record could be improved by performing a high depth resolution firn sampling (especially in the lock-in zone) and acquiring krypton corrected $\delta^{13}CH_4$ data from this site.

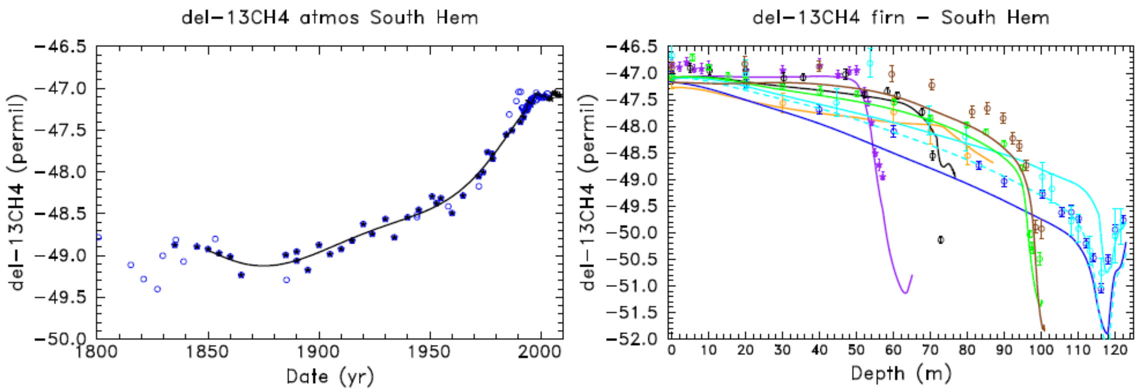

Figure 2: Left panel: atmospheric scenario (black line) based on NIWA atmospheric data at Arrival Heights (Antarctica, Lowe et al., 1991), Cape Grim air archive (Francey et al., 1999), and DE08 firn and ice data (Ferretti et al., 2005). Black stars show the dataset used to build the scenario, blue circles show the whole Ferretti et al. (2005) dataset also including the lower accumulation rate DSS site. Right panel: results obtained by using the left panel scenario as input to the LGGE-GIPSA forward model of gas transport in firn, color code for firn drilling sites: see e.g. Fig.5 of our manuscript. The light blue dashed line shows results for South Pole 2001 obtained with the South Pole 1995 diffusivity.

This figure has been added in the revised version of the supplementary material.

18) The discussion section (section 7) of the MS is basically a summary of the previous pages, and not really a discussion. I suggest merging this section with the previous ones, or with the conclusions.

<u>Answer 18</u>:

The discussion section has been replaced with a section focused on uncertainties (on model and data). The former conclusion has been turned into a conclusion and perspectives section.

Other corrections:

Title: I think 11 polar sites is misleading. DI is not used in the reconstruction, and NM and SPO are sampled twice (Should NM 2008 and 2009 be considered different sites?).

polar sites would be more appropriate.

To avoid any confusion and to clarify the aim of the paper, we propose to change the title to: Can the carbon isotopic composition of methane be reconstructed from multi-site firn air measurements?

P9589 L14-15: State your actual conclusions in the abstract. What trend do you get, how does the reconstruction compare to other records, etc.

Due to the difficulty to reconcile all datasets and the multiple possible causes of the inconsistencies, we prefer not to discuss the trends and growth rates in detail (see answer 12). However, one sentence has been added stating our actual conclusion in the abstract of the revised manuscript.

P9592 L3: measurements usually lead to the conclusions that. . .. Was this the case for this study?

This sentence has been slightly modified for clarity in the revised manuscript.

L17-18: (Martinerie et al 2009): I think Clark et al. JGR 2007 is the more appropriate citation for DI.

The citation Clark et al., 2007 has been added in the revised manuscript.

L24-25: A possible trend. . ..our dataset: This is a very important statement. Do the authors mean that the reconstruction after 1993 is unreliable? This has implications for the interpretation of the results. Please elaborate.

This sentence means that the surface individual data points of each site taken together do not show a clear atmospheric trend, but of course this does not include possible surface perturbations, seasonality, etc… To avoid misleading the reader, this sentence has been removed.

P9593 L10-12: The connections between the individual pores and the bubble closure (on the microscopic scale) do determine the transport properties to a great degree. I would say both micro and macro-scale features are important.

This sentence refers to the second paragraph of Section 4 in Fabre et al. (2000). As firn micro-structure is not the main subject of our manuscript, this sentence was removed.

L26: Martinerie 2012 is not in the reference list. If the work is still in preparation, remove the citation.

The citation has been removed of the revised manuscript.

P9594 L1-12: mention in this first paragraph that you use a smoothness requirement for the solution (with link to supplement).

One sentence referring to the SI has been added to the revised manuscript.

L19: is horizontal diffusion ever negligible? I suspect it occurs at all depths, the model simply does not require/capture it.

This has been changed in the revised manuscript. "is no longer negligible" has been replaced by "is important".

P9597 L10 and throughout the MS: Please specify you are referring to Fig 4b (instead of Fig 4). All figures have several panels, so refer to the specific panel.

This has been specified in the revised manuscript.

P9598 L5-8: is it possible that the problem is mathematically under-constrained, and that the differences between NM-08 and NM-09 are in part due to that?

This comments allowed us to find an error in Section 5.3 and caption of Fig.4 of our manuscript: the results discussed are not inverse scenario reconstructions but firn fractionation calculations (as in Section 5.2 and Fig.3). Firn fractionation is calculated with the forward model of gas transport in firn hence the solution is not mathematically under-constrained.

L10: high accumulation sites do not have a thinner diffusive zone. Compare Law Dome DE08 and DSSW20K, where the former high acc site has a longer diffusive zone. Also densification models predict a longer firn column at higher accumulation. From Fig 4 it is clear that the firn fractionation is mostly sensitive to the diffusivity variations in the lock-in zone. In this zone advection is the dominant transport mechanism; advection does not fractionate, and hence the small sensitivity at DE08.

We have removed "due to a thinner diffusive column height and". This was a mistake also noticed by Referee 1.

L26: Replace Fig 4 with Fig 4b

This has been corrected in the revised manuscript.

P9599 L20-21: Give a justification for doing this. I suspect using SPO-01 gives better results, but this seems a completely ad-hoc adjustment.

The justification is provided at p9597 l1-3: the SPO-95 diffusivity is constrained with 6

species whereas only 3 species are available for SPO-01. We have modified this sentence (P9599 L20-21) in the revised manuscript.

P9600 L10-12: Which reconstruction (green or red) is the better one, and for what reasons? Which one is used in the remainder of the study?

The two reconstructions using equal weight (red) or weighed differently per site according to the uncertainties of the data from the different sites (green) show relatively similar results, well within error bars of one another. They would have led to different results if the $\delta^{13}CH_4$ data were undergoing very different uncertainties (e.g. a factor of 10) from site to site. The fact that the green and red reconstructions are similar supports the idea that it is not the case here. Due to the difficulty to precisely evaluate overall uncertainties, the equal-weight approach has been used in the final multisite reconstruction.

L14: what does "supposedly" mean here?

The word "supposedly" has been removed in the revised manuscript.

P9601 L24-25: This is too bad. An IPG estimate would be one of the most interesting things coming out of a reconstruction on both hemispheres. Could it be placed in the supplement (with large uncertainty bars)?

We agree that this would be very interesting information, but with the uncertainties presented in our manuscript, it does not seem achievable at the moment. Determining the IPG based on $\delta^{13}CH_4$ firn air data is very uncertain, because the IPG is of the same order of magnitude as the error bars of the reconstruction, therefore it is not possible to reconstruct the temporal changes of the IPG over the last 50 years with firn data. In relation with several comments above, especially comment 7 from Referee 1, the NH+SH simulations were removed.

P9602 L5: Discuss the implications of the constant IPG reconstruction, or leave it out.

A reliable reconstruction of the IPG is not possible with our scenarios therefore this paragraph has been left out in the revised manuscript.

L16: Leave out "in review" references

This article is now published, the reference has been modified.

L17-19: I disagree. The reconstruction shows a downward trend in the NH which is not there in the atmospheric data. Also, there appears to be an offset of around 0.2 permil.

Perhaps it would make sense to show de-seasonalized atmospheric data (moving 1 year average filter) because the firn reconstruction has no seasonality either.

We wrote that it agrees well, because the atmospheric data are within the error bars. When compared to the width of the scenario uncertainty, which reflects the mismatches between firn data and modeled firn concentrations, the downward trend in the NH is insignificant (see also answer 16). An offset due to calibration issues is possible, but unfortunately, no international scale exists for $\delta^{13}CH_4$ and we don't have a way to calibrate all firn data with respect to atmospheric data. A section concerning intercalibration has been added to the revised Supplementary Information. Moreover, problems due to Kr interference (as explained above) may affect atmospheric and firn air samples and this could contribute to the observed offset.

L20-27: It should be mentioned here that the cape grim archive does not suffer from firn fractionation, and is therefore probably more reliable. See my comment above on the calibration scale offset. I am confused, because on it is stated that IMAU and CSIRO are consistent.

We realized that the brief discussion about intercalibration is confusing, so we added a table about intercalibration in the revised version of the Supplementary Information and discussed further the uncertainties associated to intercalibration issues in section 7 of the revised manuscript.

P9603 L9: Also mention here that the difference in slope could be due to the firn fractionation correction that is applied to the data.

One sentence has been added to state that uncertainty in the firn fractionation may be as well a cause of the observed differences.

L24: The fact that both models agree well on DE08 also shows that the firn fractionation can be reliably calculated at this site due to the high accumulation rate.

Due to the uncertainties related to the krypton correction and the low depth resolution of the DE08 firn data, we are not fully confident in the estimated firn fractionation at DE08 (see also answer 13). However, we agree with the referee on the fact that high accumulation rate sites such as DE08 offer the best perspective to improve our understanding of $\delta^{13}CH_4$ in firn.

P9604 L2: what does "possibly" mean here? Please be precise with the calibration scales!

This sentence has been modified in the revised manuscript.

P9604 L15-18: This is an important and interesting new observation, which is completely unexpected. Can you explain this effect? Please elaborate.

The comparison between the SPO-95 and SPO-01, and between the NM-EU-08, NM-US-08 and NM-09 reconstruction led to this statement. However, the same kind of behavior occurs at Devon Island and DML (strong atmospheric gradients predicted in single site reconstructions, see Fig. 2), which is likely a diffusivity-related problem, so not due to the number of reference gases used but to the presence of melt layers for Devon (=> complex shape of the diffusivity profile not fully represented) and to the low depth resolution in the bottom firn (where steep concentration gradients occur) for DML.

This section has been modified (e.g. in answer to comment 18) in the revised manuscript.

P9605 L9-11: I would disagree with this statement. As I mentioned earlier I consider DE08 to be a more reliable site for d13C(CH4) reconstruction, because it is not affected by the uncertainty in the firn fractionation modeling. More effort should be made to improve the consistency of the single-site reconstructions, e.g. by trying different firn air transport parameterizations, or by using different firn air models.

See answer 13)

This sentence has been removed from the revised version and the uncertainties associated to our reconstruction have been discussed further in this section.

L16-19: This is an important and interesting conclusion. However, the cape grim record could be used to constrain firn models back to the 1970s.

"This sentence has been suppressed (see also answer 17).

P9606 L4: 8 sites, or 10 if different sampling campaigns at the same site are counted double.

In this study we consider measurements from 11 boreholes from 11 sampling campaigns occurring at different dates, and different locations, even though the locations were sometime very close to each other. For clarification, we changed this term to " 11 firn sampling campaigns" in the revised manuscript. As Devon Island is

excluded in the multi-site simulations, we use 10 firn sampling campaigns, not 11.

P9613 and other figures: would it be possible to supply the figures as vector graphics? This should be a simple operation with most graphical editing software, and would considerably improve the readability of the figures.

Our figures are vector graphics, but pasted in Word, their quality decreases. Figures will be delivered in the good format for the final publication.

P9614: Why are the SPO-01 error bars so large suddenly? In Fig 1b they seem comparable to other study sites.

A mistake occured in Fig.1 where the error bars of SPO-01 should be larger. This has been corrected in the revised version.

P9618: This is one thing I don't understand. How is it possible that for 1960-1980 the "best-estimate" reconstruction on both hemispheres is considerably below all the data points? This does not make sense if one tries to minimize the RMS. Is the regularization term of the solution set too high?

Fig. 6 plots firn data corrected for firn-fractionation versus the mean age of each datapoint (Green function sense) and this is compared to the inverse model atmospheric trend. The mean age of firn data is a biased indicator of atmospheric age because age distributions in firn are not symmetric, and the firn data corrected for firn-fractionation are a biased indicator of atmospheric concentrations e.g. because of diffusional mixing in firn. Thus the corrected firn data are only roughly consistent with the inverse model scenario and that is the reason why we need an inverse model. The consistency between firn data and model results can be better appreciated on the right panels of Fig. 2 and Fig.5. However, plotting firn fractionation corrected glaciological data versus age is the usual method for interpreting ice core data. In this respect, the fact that nearly all firn data on Fig. 6 fall within the error bars of the inverse model scenario suggests that the method used for ice core data is not strongly biased.

Supplement Section 1: the misfit for DI is really remarkable (Fig S1e). Could you comment?

The misfit is due to the fact that near-surface $\delta^{13}CH_4$ in Devon Island firn is distinctly lower than at other sites. The seasonality test indicates that it is not due to a seasonality

effect. Diffusivity at Devon Island is smaller than at all other sites, and stronger CH_4 trends are observed in the upper firn than at other sites (Witrant et al., 2012). In relation with the melt layers, firn fractionation may already operate and be under-estimated by the model in the upper Devon Island firn.

Moreover, as discussed in the revised section 7 of the manuscript, intercalibration offsets or analytical scale differences (Kr interference) could also play a role and explain at least partly this misfit.

Supplement Section 3: what are the units of the kˆ2 term? Is it somehow related to a timescale on which the solution is smooth?

While k^2 directly affects the smoothness of the solution, it acts as a weight on the rugosity of the atmospheric scenario (L term in Rommelaere et al., 1997, section 5.1) and it is relative to the firn measurements error weighted by their uncertainty (Is term, Eq. 34 of Rommelaere et al.,1997). Thus, k has units of squared time per unit of concentration (to be consistent with the fact that Is is dimensionless and rugosity is a second order time derivative). Here we use $\delta^{13}CH_4$ concentrations in ppb and the unit of k^2 is $(year^2/ppb)^2$.

A better "normalized" weight would be k x σ (units of squared time) and taking the squared root of k x σ gives a timescale of the order of the decade. Nevertheless such a criterion is not sufficient to provide satisfying inversion results for the different sites/species configurations and a better defined optimal value of k^2 has to be used.

More advanced statistical methods would be necessary in this case as discussed in Answer 16 and Witrant and Martinerie, 2013.

Supplement Section 4: Two methods are described, but which one was used for this study? This is not motivated or even discussed.

The first method is used in the main paper, and the second method is used on Supplementary Fig. S4. This is now more clearly explained in the revised supplementary information.

P9589 L10: firn air transport models (insert “air” or “gas”)

This has been corrected in the revised manuscript.

L18: Its atmospheric mixing ratio has rapidly. . . (make ratios singular, insert “has”)

This has been corrected in the revised manuscript.

L22: list a few source mechanisms as well (sinks are explicitly mentioned)

One sentence has been added on the major methane sources.

P9590 L26-27: replace “results” with “data”

This has been corrected in the revised manuscript.

P9593 L5: replace “scenarios” with “histories”

This has been corrected in the revised manuscript.

P9595 L20: replace “changes” with “signals”. Add a comma after column

This has been corrected in the revised manuscript.

L21: add “a” before d13C(CH4) trend

This has been corrected in the revised manuscript.

P9596 L3: add “d13C” before “atmospheric history”

This has been corrected in the revised manuscript.

P9599 L16: remove “sometime”

This has been corrected in the revised manuscript.

P9600 L27: replace “convert. . . into temporal isotope values” with “place isotope measurements on a time scale”

This has been corrected in the revised manuscript.

P9601 L3: “note” is used twice in sentence.

This has been corrected in the revised manuscript.

P9601 L15: add “strongly” in front of “constrained”

This has been corrected in the revised manuscript.

P9603 L22: typo, “LGEE” should be “LGGE”

This has been corrected in the revised manuscript.

P9604 L2: remove “ice core”

This has been removed in the revised manuscript.

P9604 L18: replace “less” by “fewer”

This sentence and the discussion section have been modified.

P9606 L5: add comma after “inversion”, replace “estimate of” with “calculated” or “modeled”

This has been corrected in the revised manuscript.

L6: replace “good” with “well”

This has been corrected in the revised manuscript.

P9612: is z_last in the caption the same variable as z_lowest in the table?

Both are the same. This has been corrected in the revised manuscript.

Supplement P3 L18: make plural: two different wayS of excluding. . .

This has been corrected in the revised SI.

References not cited in the ACPD paper:

Anderson, J.: Fundamentals of Aerodynamics, McGraw-Hill Companies, 1991.

Battle et al., Nature, 383(6597), 231-235, 1996.

Battle et al., Atmos. Chem. Phys., 11, 18633-18675, 2011.

Butler et al., Nature, 399(6738), 749-755, 1999.

Fujita et al., J. Geophys. Res., 114, F03023, 2009.

Francey et al., Tellus, 51B(2), 170-193, 1999b.

Freitag et al., J Glaciol, 50(169), 243-250, 2004.

Kai et al., Nature, 476, 194-197, 2011.

Levin et al., Nature, 486, E3–E4, 2012.

Lomonaco et al., J. Glaciol. 57(204), 755-762, 2011.

Martinerie et al., Earth Planet. Sci. Lett.,112, 1-13, 1992.

Sapart et al., Nature, 490, 87-89, 2012.

Severinghaus and Battle, Earth Planet. Sci. Lett., 244, 474-500, 2006.

Severinghaus et al., Earth Planet. Sci. Lett., 293, 359-367, 2010.

Spaudling et al., J. Glaciol, 57(205), 796-810, 2011.

Sturges et al., Atmos. Chem. Phys., 12, 3653-3658, 2012.

Sturrock et al., J. Gepohys. Res., 107(D24), 4765, 2002.

Trudinger et al., Atmos. Chem. Phys. Discuss., 12, 17773–17834, 2012.

Witrant et al., Atmos. Chem. Phys., 12, 11465-11483, 2012

E. Witrant and P. Martinerie, "Input Estimation from Sparse Measurements in LPV Systems and Isotopic Ratios in Polar Firns",Proc. of the IFAC Joint Symposium on SSSC, TDS and FDA, Grenoble, France, Feb. 4-6, 2013.