

## ***Interactive comment on “Intercomparisons of $\delta^{13}\text{C}$ and $\delta\text{D}$ measurements of atmospheric $\text{CH}_4$ for combined use of datasets from different laboratories” by Taku Umezawa et al.***

**Anonymous Referee #2**

Received and published: 26 October 2017

The authors take on the difficult and important task of summarizing comparisons of measurements of  $\text{CH}_4$  stable isotopes in air at  $\text{CH}_4$  levels appropriate for the current atmosphere and air extracted from ice cores. Done properly, this would give data users correction factors to combine data from different laboratories and give them an understanding of the major issues involved so they fully understand the limitations of the combined data sets. This would allow more  $\text{CH}_4$  isotope data to be used in studies of the global  $\text{CH}_4$  budget. Unfortunately, the paper seems to be written for isotope measurement experts, like the authors, rather than for data users who may be very interested in more  $\text{CH}_4$  isotopic composition measurements in their studies. It contains too much jargon and too little explanation of the major issues that prevent labs

C1

from preparing a combined data set for  $\text{CH}_4$  isotopes with meaningful temporal and spatial gradients. These issues seem to be inherent in the the community's measurement approach. But what are those issues? Are they associated with deficiencies in measurement techniques themselves and how instruments are calibrated? Is there a lack of appropriate isotope standards for  $\text{CH}_4$ ? Is the issue with propagating standard scales from carbonates ( $^{13}\text{C}$ ) or water (D/H) to  $\text{CH}_4$  in air? The fundamental hierarchy of standards used within the  $\text{CH}_4$  isotope community should be described; much of what is described seems to violate good metrological practice. Are sample collection or processing methods causing differing amounts of fractionation among labs? These are just some possibilities from a non-expert. I suggest that a brief, systematic description of the important issues involved in making measurements of atmospheric  $\text{CH}_4$  isotopic composition and ultimately preparing a consistent data set across measurement labs is given in Section 2. This would be especially useful to data users and also help the authors focus on how to move forward.

This study is supposed to help scientists utilize more  $\text{CH}_4$  isotopic data by providing correction factors to make data from different labs more compatible, but are the data sets available? A quick look at the World Data Center shows only NIWA and NOAA data have been updated within the past couple years for  $^{13}\text{CH}_4$ , and the only other data set, quite outdated, is from Tohoku University. Are the data from ice core and firn air available? Do other labs make data available through their institutes' web sites? That issue aside, using the comparison information in the paper is complicated by Kr interference; it is not always clear where a correction has been applied and, as a result, what data sets the information in this paper is appropriate for. It gives the impression that the paper was written before the isotope measurement community is ready for such an effort to be more generally useful to the  $\text{CH}_4$  research community.

It was surprising that  $\text{XCH}_4$  (i.e.,  $\text{CH}_4$  mole fractions) was not reported in the paper, at least for the cylinders that were circulated as part of this study. It is important that  $\text{CH}_4$  measurements agree among labs measuring air from ice cores and the modern

C2

atmosphere to get radiative forcing correct, and this is a good forum to show that level of agreement.

The manuscript is poorly written. It is far too wordy. While the first author is not a native English speaker, there are at least 10 authors who are. Shame on all of you for not improving the English. Language is often vague. Scientific terms are misused.

Overall, I think the work described in the paper is important, despite that the community is not yet capable of preparing useful combined data sets. I recommend that the paper is re-evaluated for acceptance after the authors respond to the comments in this review to the satisfaction of the journal editor.

General comments: 1. Use appropriate metrological terms. "Precision" is a qualitative term, yet it is used as a quantitative measure of the noise or uncertainty in a measurement system. Is it short-term noise (repeatability) or does it represent long-term variations of a measurement system (reproducibility)? When the proper terms are used, how are they quantified? More appropriately, uncertainty should be stated with its confidence interval.

2. Calibration: paraphrased, it links the measured response of an analyzer to the known values of standards. In the text, it seems to be confused with standard, and its use is often unclear. Given that, terms like "calibration offsets" are vague. Are the standards different? Is the issue with propagation of the standards? Could the offset result from fractionation during sample processing?

3. Differences between labs are often given with standard errors; I think standard deviation would be a better metric. In cases where n is large, e.g., for ongoing comparisons that happen over years, standard error exaggerates how well the difference is known.

4. Remove unnecessary words: assessed to be... comparison exercises (delete "exercises") evaluated to be X L in volume (delete "in volume") considered to be "in the time period" and "the years" offset value (delete "value")

C3

5. Each participant in measurement of CH<sub>4</sub> isotopes from the circulated cylinder should report XCH<sub>4</sub>.

6. "Concentration" is misused. In most cases, it can be deleted, because the unit provided (ppb) defines the measured quantity.

7. Why were other labs measuring CH<sub>4</sub> isotopic composition not included (e.g., Oregon State University)?

8. What are the main sources of differences among labs? You imply it is differences in standard scales, but why? Don't the working standard scales propagate back to a primary scale, e.g., VPDB for <sup>13</sup>C?

9. What is the path forward, beyond what was mentioned regarding newly-developed standards? Many deviations from good metrological practice, especially regarding propagation of isotope standard scales, have occurred within this community. The new standards, although developed using an approach that may not be defensible in a pure metrological sense, seems practical given the limited resources of the measurement community. Is that alone sufficient? What else needs to be done to make existing data more compatible? How could new laser-based spectroscopic methods help this measurement community and the science? What else could improve the quality and compatibility of measurements of CH<sub>4</sub> isotopic composition across the labs involved here and beyond to others? As mentioned, data availability is not considered.

Specific comments:

P1L32: suggest ..from an inter-laboratory comparison of measurements.... (Also for title.)

P1L32: ..among worldwide...

P2L3: What does "the data" refer to? The differences among labs?

P2L4-5: As presented, it is not clear how this will help combine data sets. It could

C4

be more clear it a table was given of available data sets and if (i.e., with respect to Kr interference), and how, the offsets in the paper apply.

P2L8-12: The description of how CH<sub>4</sub> isotopes are used to constrain the CH<sub>4</sub> budget could be stated more clearly. I suggest something like "The mass-weighted average delta-13C of emissions from all sources will equal the delta-13C of atmospheric CH<sub>4</sub>, after correction for fraction by removal processes." While you give some references for studies that use isotopes of CH<sub>4</sub> (some are poor examples), you don't reference early literature that identified their usefulness (e.g., Stevens).

P2L21: Is not this ratio more generally rare/common isotope, i.e., more abundant isotope in denominator?

P3L3: Condensable? At what temperature? How about CO? Why not describe the method directly?

P3L17: "types of" is vague. Be more specific or delete it.

P3L24: delete "datasets".

P4L1: "reliable calibrations"? Do you mean being able to reliably characterize the response of your instrument with standards, or do you refer to the standards themselves? The following discussions of "calibrations" is vague.

P4L4: "primary calibrations"? Do you mean calibration of primary standards? What defines the primary standard scale for CH<sub>4</sub> isotopes?

P4L9+: What is the Kr interference? I assume it is something with same mass/charge as the "CH<sub>4</sub>-derived peak"?

P4L11: What is the "CH<sub>4</sub>-derived peak"?

P4L26: You are summarizing the analytical methods used by each laboratory, not reviewing the labs.

## C5

P5L6: In place of standard errors, standard deviations (with "n" given) should be used.

P5L9+: I think the general discussion of techniques (DI-IRMS, GC-IRMS, and optical), calibration, propagation of standards and their traceability to fundamental SI quantities, limitations of current methods, interferences, memory effects (scale compression?), etc. would be useful to isotope data users and fit better here (section 2) rather than in the introduction. This would be a good place to define how metrics like repeatability and reproducibility are quantified and other terms that might be unclear to non-specialists.

P5L15: replace "the early years" with a year or range of years.

P6L19-20: How can calibrations of an instrument in one lab be the basis of calibrations in another lab? Do you mean the standards developed at MPIC were propagated to IMAU?

P6L25: unclear what "calibrations were made against ..." means. What do these abbreviations mean? Was water from these standards injected directly into the spectrometer, or were intermediate standards traceable to them injected? This is used other places, too.

P7L5: what is "a working standard air"?

P7L21: delete "because of"

P9L19-20: Kr interference is significant.

P9L21: Correction of the data for Kr interference ...

P9L23: "Kr removal"? do you mean only these data were corrected for Kr interference?

P9L25: ..RHUL) measured atmospheric ... using....

P10L9: How can one lab share its "calibration" with another. They shared standards?

P10L16: "calibration is made against gas bottles"? What is in the "gas bottles"? What standard is it traceable to?

## C6

P10L21-22: "anchored to the INSTAAR calibration"?

P10L23: "overall measurement precision"? How is that different from "precision", repeatability, or reproducibility?

P10L25: What is "an Antarctic bottled air"?

P11L7: You can not transfer a standard scale by measurement of a single sample or even multiple samples. Comparisons of measurements can not replace propagation and maintenance of a standard scale.

P11L18: Dates (or ranges) should be given for each comparison.

P12L6-7: TU filled four ....., two with dry ambient air (give dew point) and two with ... (what is "synthetic standard air"? How is it different from real air?).. with CH4 (delete "concentrations" - the units make it clear what the quantity is (which is not concentration, anyway)).

P12L12: .. after transport to... (throughout)

P12L14-15: "Calibration offsets" is too vague.

P12L21: "scale compression" should be defined.

P12L22: what are the differences in matrix?

P13L12: This title implies something else. Ice cores are not part of the round robin. It is a comparison among labs that measure delta-13CH4 from air extracted from ice cores.

P13L20" "elemental"? Do you mean measurements of XCH4 (i.e., CH4 mole fraction)?

L13L24-26: A change after 9 years could mean a change at PSU, not necessarily drift in the isotopic composition of CH4 in the cylinders. What happened to XCH4? Did it drift? If not, could CH4 isotopic composition change without a change in XCH4? If so, how?

C7

P14L5: The level of agreement depends on your perspective.

P14L14: suite of cylinders...

P15L10: "RHUL-INSTAAR offset"? Are they both different from NIWA, which offsets are calculated from, by the same amount? I suggest "difference" rather than "offset".

P15L12-13: With such large "n", the uncertainty on the difference is deceptively small. SD would be more representative of true uncertainty in the difference, since their could be drifts over time in one measurement vs another.

P15L14: data were corrected...

P15L20: among laboratories

P15L21: delete "It has also happened that"

P15L27-P16L1: Fig. 2 combines....

P16L11: delete "that" or rewrite as "stability ... 1992 to 2007 continues until 2011."

P16L17-18: was the GC-IRMS with the post separation column used to define the empirical correction in L15 or the DI-IRMS?

P16L22,L25: rather than say "high" or "middle" cylinder, why not ...for CH4 at ~X ppb to make it clear?

P17L2: it can not be the "calibration" that is propagated, but rather a standard or standard scale.

P17L3: Since standard scales were never defined, I'm not sure what "primary calibration" refers to? It should be calibration of the instrument at MPI with primary standards, what ever they are.

P17L14: it is the scales, not the calibrations that are important.

P17L14-15: what does "their" refer to? Since Sperlich's affiliation is given as NIWA,

C8

then is this difference between IMAU and NIWA?

P17L22: delete "has".

P18L7: "instrument circumstance" is too vague.

P18L18: both use the same scale.

P18L26: update of the standard scale or the method used to calibrate the response of the instrument with that scale?

P19L12: "intercomparison in this study"? The round robin?

P19L16: ...comparison described by Nisbet?

P20L5: This is the best comparison after correction for Kr interference, so why is excluded from Fig.2a?

P20L18: internal standard?

P21L2: measurements of air in cylinders exchanged between... How many cylinders?

P21L3: applied an offset correction (delete "to").. to all data.. (delete "the")

P22L1: replace "shows the offset to be" with "gives"

P23L10: what makes these "programs"?

P23L11: The results are about measurement offsets; they do not address differences in standard scales directly.

P23L15: among labs...

P24L3: atmospheric CH<sub>4</sub> level - when? modern?

P24L26: ...similar to...

---

Interactive comment on Atmos. Meas. Tech. Discuss., doi:10.5194/amt-2017-281, 2017.