

Response to reviewer 1

The authors thank the reviewer for the positive response and thorough comments that have improved the manuscript. The point-by-point response is below. The reviewers' comments are numbered and are in black font, the authors responses are also numbered and in blue font to ease readability.

Comments from Reviewer

The manuscript by Worton et al. describes the production and evaluation of several multicomponent gaseous primary reference materials (PRM) with the aim to allow PTR-MS users to better constrain the transmission curves of their respective instruments. The manuscript discusses the challenges of including low-volatility compounds, but manages well to reproduce their concentrations in the different PRMs with uncertainties typically below 10%. The quantitative addition of such high-mass low-volatility molecules to gas-standards will be of great use for the PTR-MS user community in general and is also important for atmospheric measurements. Challenges in the quantification of several compounds were overcome by using a combination of GC-MS and GC-FID and in some cases Cryo-GC-FID. The PRMs described were prepared following standardized procedures and evaluation results are presented in great detail and with particular dedication to precision and uncertainty analysis in a well-structured manner. However, the manuscript was in parts difficult to understand on the first read, because abbreviations were used excessively and partly confusing. The paper gives insight into the reproducibility and stability of gas standards that will be very valuable to end users. Transmission-curve constraints with one single reference material can simplify the lives of many PTR-MS users around the world and enhance data comparability and quantification. I therefore suggest that the manuscript is published in AMT after some minor comments have been addressed.

R1.1 The GC-MS method used is not fully described. Please specify especially the type of ionization (electron impact, chemical ionization ...) (lines 94 ff.)

Added the following text 'electron ionisation (70 eV)' before mass spectrometer on line 94.

R1.2 How was the separation of the 3-carene and 1,2,4-TMB peak treated? Was a multipeakfit performed and the data corrected accordingly or do the two compounds influence the other's signal? (lines 96 ff.) In line 144, the baseline separation issue is mentioned for Acetone and DMS as well, but it is missing in lines 96 ff.

While the 3-carene and 1,2,4-TMB peaks and the acetone and DMS peaks are not baseline resolved the separations were observed to be very consistent throughout all the chromatograms collected and this is described already in the manuscript between lines 114 and 118. A multipeakfit deconvolution was not used.

R1.3 For the Cryo-GC-FID, the volume of the loop, and the trapping / heating cycle are not described (l. 105 ff.)

These details have now been added to the text.

R1.4 Why are the FID data on acetonitrile so much more noisy than the acetonitrile GC-MS data and also so much more noisy than the methanol FID data, although the latter should give a smaller FID signal? Consider adding a sentence or two on this matter.

It was not clear as to what the main reason for the noisier GC-FID data. It is possible that a co-eluting impurity that was present at different amount fractions in the different PRMs could be an explanation but we do not have an conclusive evidence to support this suggestion. We have added some text eluding to this in the main manuscript.

R1.5 Consistent use of abbreviations:

Line 57: NPL is not defined (don't let the reader search for the hint in Dave Worton's email address...)

Line 154: Is PTRMS NPL PRMs a subgroup of NPL PRMs? If yes, please define, if no, please remove the "PTRMS"

Line 152: Mentioning of the certified primary reference material without the short-form "CRM" used in the supplement.

Apologies for the confusion here. We have rechecked the manuscript and have refined the abbreviations to be consistent throughout and to be simpler so as to convey only the most pertinent information (i.e., PRM, CRM).

R1.6 Also, the supplement about the certified reference material with fig. S3 is not mentioned in the main text and so the difference between certified reference material and PRM remains unclear in the main text. It appears, as if the uncertainty for CRMs is larger than for PRMs due to the simpler preparation method. What does that mean for the PFTBA validation?

To clarify this we have added the anew section 3.5 CRM Validation and have added the following text to the main manuscript citing the discussion of the CRM in the supplement as well as explaining the difference between a CRM and PRM: **"To enable a more cost effective and timely delivery of the PTR-MS transfer standard an NPL certified reference material (CRM) was also developed. In contrast to the NPL PRMs the CRMs are not prepared by gravimetry and are produced by the direct addition of multicomponent mixtures derived from the original pure liquids. Further details of the preparation method are given in the Supporting Information (Supplementary text: Preparation and validation of certified reference materials). The amount fractions for the components in the CRMs were assigned through analytical comparisons between each CRM and an NPL PRM. In this way, preparation is quicker and more cost effective while maintaining the integrity of the values and their traceability. Initially with the developed CRM preparation method it was possible to produce mixtures that had blend tolerances of 20 – 30 % (see Figure S3, Supporting Information), which are suitable for end users but work is continuing to improve this with the aim of achieving better than 10 % blend tolerances in the near future. The blend tolerances are just an indication of the repeatability of the preparation process, the amount fractions of each CRM are analytical assigned by comparison to one or more NPL PRMs. The uncertainties in the amount fractions for the CRMs are independent on the blend tolerances and where between 2 – 10 %, compound dependent, and these uncertainties were predominantly drive by the observed differences from the preparation of the PRMs. "**

The CRMs have a larger uncertainty than the PRMs because in addition to the uncertainties in the gravimetric preparation data and the validation the uncertainty from the analytical value assignment of the CRMs from the PRMS also needs to be included. However, for the PFTBA, the validations were done using the gravimetric data of the preparations. As such this was developed using a primary method (the gravimetric method). This was done by addition of PFTBA by two separate weighed loop additions to two independent cylinders, which were then validated through analytical

comparisons to the values expected from the gravimetric data as such it was validated in line with the PRMs.

R1.7 In this context, there is also missing information on the BTEX NPL PRM in the main text (especially its age, and if it is produced also gravimetrically)

The following information on the BTEX PRM has been added to the main text at line XXX: "Toluene was validated by comparison against an existing NPL PRM containing BTEX (benzene, toluene, *m*-xylene, *p*-xylene and *o*-xylene) components **that was prepared gravimetrically in December 2018 and had been independently validated against other NPL PRMs that were internationally compared as part of key comparison CCQM-K10.2018 (Cecelski et al., 2022). These BTEX PRMs are known to be stable for more than 5 years and at the time of the comparisons the BTEX PRM was less than 3 years old.**"

R1.8 Text and Fig. S3 in the supplement: Decide on CRM, PTRMS CRM or NPL CRM to support the reader, except you want to address differences, that then should be clearly defined.

Apologies for the confusion. These have all been changed to NPL CRM throughout.

R1.9 The names of the different reference materials are not clearly motivated and therefore reduce the readability. I suggest changing them so they e.g. contain the month and year of production. To guide the reader for a faster understanding of Table 2 consider colouring the Cylinder IDs in table 2 as you do in the figures.

The names of the different reference materials came originally from the unique identifiers of the actual cylinders used but we acknowledge that these are not meaningful to the reader so we have changed them using the month and year suggested by the reviewer. Thus A574, A578, A638, A643, 2819, D961492 become 0917a, 0917b, 1218, 0119, 0819, 0821, the colour coding suggestion has also been adopted in Table 2.

R1.10 line 158 ff: - did you really measure "A" always twice directly after each other or were the repeats performed on different dates / which longer breaks in between? If so, add this. Otherwise, consider correcting to (AB)nA.

What is meant by „bracketing“ between the nearest neighbors? Consider using a more common word that is clear also for non-english natives.

Add to Eq. 1 that it is calculating the relative difference in %

We have corrected the repeating pattern to (AB)nA. We have removed the reference to nearest neighbours and added the following text to clarify what is meant here: "The ratio in response was determined by **dividing B by the average response of the A's immediately before and after each analysis of B.**", We have added the % sign into Eq. 1.

R1.11 l. 198 „A638“ is mentioned twice

This was a typo and should read A643 (now 0119 in the new naming convention, see R1.9 above).

R1.12 The footnotes in table 3 (a,b,c,...) do not correspond to the markers within table 3 (1,2,3,...)

This has been corrected.

R1.13 it is not clear to me, why the reference would be a PRM that is produced from many old parent mixtures. Thus I wonder, what is the advantage of 2819 compared to using D961492? (line 149) In general, only little information has been given on the stability of the parent mixtures.

The reason that 2819 is used as the reference is because many of the validations between the PRMs took place before D961492 had been produced so comparing to 2819 provides a consistent link between all of the PRM mixtures. The parents have been produced and validated over many years and form part of the National Standards, maintained by NPL. While some have been produced a number of years ago they have been measured on numerous occasions against newly prepared mixtures that were destined for third parties as part of our measurement service offerings. These validations and those presented in this work demonstrate that the parents are stable at least on the timeframe of the PRM produced in this work and probably much longer. Typically, we observed stabilities that are longer for higher amount fraction mixtures compared to those at lower amount fractions. The following text has been added to section 3.4 (PRM validation) to clarify the choice of 2819 as the reference PRM: **“In the majority of cases PRM 0819 is used as the reference to which all others are compared. It was chosen as such because at the time it was the newest PRM to be produced and was used to benchmark all the others that had already been made. Thus, PRM 0821 was also referenced to PRM 0819 to provide a link between all six PRMs.”**

R1.14 In fig S2, D4-siloxane and D5-siloxane, biased data do not appear corrected again (as they do e.g. in the plot for 1,2,4-TMB). Please add a short comment regarding this to the figure caption.

Noted. Additional text has been added in the caption and is shown in bold here: **“The open symbols show the original data before being corrected for biases in the parent mixtures (for methanol, MVK and 1,2,4-TMB) or which has been excluded from the regression analysis (for D4-siloxane and D5-siloxane) as discussed in the text.”**

R1.15 The PTR-MS transmission curve constraints would be performed with the CRMs appearing only in the supplement fig. S3 in the future, as I understood. The preparation repeatability is reduced to 20-30%. Is this still sufficient to determine a transmission curve, which usually varies within a factor 1.5-2 or will the CRMs be cross-evaluated against the PRMs? Please clarify.

This is correct. The intention is to disseminate the CRMs to end users and for the PRMs to be maintained for the purposes of value assignment of the CRMs. The preparation repeatability or blend tolerance of the CRMs is just an indication of the repeatability of the process. The amount fractions of the CRMs are based on analytical values determined by comparison to an NPL PRM. See response to R1.5 above.

R1.16 Another technique to determine the transmission curve is the addition of single species to the sample air in step-wise increasing concentrations and observe the change in primary ion signal and the ion signal of the sample molecule and repeating this for multiple compounds with different masses. This technique is time-consuming for the end-user but gives generally good results. By having one gas standard that combines many compounds spanning the whole mass range, the technique of comparing with the primary ion signal is certainly not possible anymore, which requires the CRM to really be precise and contain species that are detected at the kinetic limit. I am looking forward to your discussion of this point. A short “how-to” for the usage of your CRMs for transmission curve correction, potential error sources and pre-requirements would add great value to your manuscript, especially for new users.

Besides being time consuming, the compound-by-compound method has the disadvantage that the high product ion concentration interfere with the transmission of the primary ions. Especially in

new-generation instruments counter intuitive behaviour has been observed (e.g., increasing primary ion signal as the compound concentration is increased). So, while this method produces acceptable results in older instruments, this is not an option for instruments featuring more complex ion optics. For the 'how-to' question we refer to Holzinger et al. (2019) where the use of the standard is discussed in detail. We have added the following text to the manuscript at the end of the introduction to direct the reader to Holzinger et al., 2019 for details on how to use the CRMs for transmission curve correction: **“For details on how to use the CRMs to constrain the PTR-MS transmission curve the reader is directed to H”**