

Author's Response

Reviewers' comment: Blue

Author's Response: Black and white; with added text in italic font

Review 1

General comments:

Regarding the determining of the effect of changes in N₂O concentration, I am rather surprised at the scale and direction of these changes in comparison to the calculated values. It stands in contrast to the levels previously measured by CRDS-III (as shown in table 4), which while not correct are on the same order of the calculated offset. What is your best hypothesis for why the three instruments behave so differently, and have discussed potential explanations with the manufacturer? Because unless there has been relevant software updates for CRDS-III since Harris, 2020, I am under the assumption that the built-in post-correction for non-linearity should be close to the same, and if not it would appear as if the updated correction performs worse?

We agree this contrast in N₂O non-linearity effects on delta values, which are reported by the analyser, both between instruments and with time is striking and deserves explicit discussion. We would like to note, however, that the analyser-specific nature of the N₂O non-linearity correction and its temporal variability, raised by reviewer 1, are already discussed in the manuscript (Sections 3.1 and 4.1.2, Table 4). In addition, we extended discussions in section 4.1.2 (see tracked changes line 585-590).

To directly address the reviewer's question regarding the underlying mechanism: the observed slopes represent the residual N₂O non-linearity after Picarro's built-in post-correction, with experimentally determined instrument-specific correction factors. The parametrization of this data post-correction is optimized before instrument delivery and can be updated by the company upon request later. Since the internal algorithm and parameter settings are not disclosed, we cannot confirm / exclude whether inter-instrument differences and the divergence for CRDS-III from Harris et al. (2020) reflect differences in factory parametrisation, software updates / different correction functions, or gradual changes in instrument optical properties over time. All three possibilities or an interplay might contribute to the observed changes, and we cannot distinguish between them with the available data.

In addition, instrumental effects such as temperature fluctuations in the cavity, minor pressure instabilities, or gradual increase in background absorption (due to e.g. mirror contamination) may contribute to the observed differences between instruments, and over time. We added an explicit

discussion of this reasoning to Section 4.1.2 to reinforce our recommendation that correction terms must be empirically re-determined routinely for each individual instrument, and that previously published coefficients, even for the same physical instrument, cannot be assumed transferable: *The observed slopes represent the residual non-linearity after the Picarro Inc. G5131-i analyser built-in post-correction, whose parametrisation is instrument-specific, optimised before delivery, and can be updated by the manufacturer upon request. Since the internal algorithm is not disclosed, we cannot evaluate whether inter-instrument differences and the divergence of CRDS-III corrections from previous work (Harris et al., 2020) reflect differences in factory parametrisation, software updates, gradual changes in optical properties, or an interplay of these factors. Instrumental effects such as thermal fluctuations, minor pressure instabilities, or gradual mirror contamination may further contribute. Likewise, the Consultative Committee for Amount of Substance: Metrology in Chemistry and Biology (CCQM) Gas Analysis Working Group (GAWG) and Isotope Ratio Working Group (IRWG) suggests instrument developers to provide access to isotopologue raw data to e.g. resolve instrumental artifacts from shortcomings in post-correction.*

A final question in relation to this topic is; whether you also tried diluting below the reported detection limit, because the non-linear behavior of the calculated values first appear pronounced at $1/N_2O > 0.003$.

We did not perform dilution experiments to concentrations below the specified measurement as concentrations < 300 ppb N_2O are of minor importance for most monitoring applications. It might well be that the instrument performance degrades outside the measurement range as speculated by reviewer 1, but this is beyond the focus of this study.

While access to MATLAB is common within the field, it would have been great and more inclusive if the developed MATLAB code functions were available in a free software such as R or python. I understand that it is quite the endeavor to change programming language, but I would be hopeful if it is something that you would consider within outlook and future work. Because as is, I unfortunately cannot test the code as I would otherwise have liked to do.

We thank the reviewer for this suggestion. We agree that an R or Python implementation would improve accessibility, but rewriting and validating the code in another language is beyond the scope of the present study. The present code was implemented in MATLAB because this is the environment in which the authors were most confident and efficient. The MATLAB source code is openly available under an open license and can be viewed and adapted without MATLAB using any standard text editor. In future work, we intend to apply the code for other isotope systems (e.g. $^{13}C/^{12}C$ in CH_4) and convert it to Python.

The uncertainty calculations in the appendix appear to be correct, so I believe it would improve the impact of the paper if you could provide an example of the scale of error one could expect. A suggestion would be correcting for the matrix effect of Cal 2.1_{330ppb}, where the value and error could nicely be compared to the known value?

We agree with the referee, and we have added a new paragraph to the manuscript providing an exemplary uncertainty calculation (Supplementary Information, S.3.1). Since no measurements of Cal 2.1_{330 ppb} gas at different CH₄ and/or CO₂ concentrations were performed, the uncertainty due to matrix effects (namely, the CH₄ spectral interference correction) was evaluated using measurements of Cal 1_{90 ppm}, diluted to 330 ppb N₂O, performed on 05.02.2024 with CRDS-I:

To provide an exemplary uncertainty estimate, we evaluated the uncertainty components for experiment 2 (sec. 3.2) performed at 330 ppb N₂O (Cal-1) with different additions in CH₄ (up to 10 ppm). In detail, measurements on 05.02.2024 with CRDS-I were evaluated. The following input parameters for the uncertainty calculations were considered; the uncertainties in the N₂O non-linearity correction slope (Δm_{N_2O}), the CH₄ correction slope (Δm_{CH_4}), and the isotopic composition of Cal_{1.1} and Cal_{2.1} reference gases, provided in Tables 4, 5, and 1, respectively. The uncertainty contribution by the N₂O concentration correction (Eq. A20) is negligible for all delta values (< 0.01 ‰), since sample and reference gases are measured at similar N₂O concentration (330 ppb). The uncertainty due to "imperfect" CH₄ spectral interference correction (Eq. A21) increases with increasing CH₄ concentration, up to 0.23 ‰ in $\delta^{15}N^\alpha$, 0.13 ‰ in $\delta^{15}N^\beta$ and 0.05 ‰ in $\delta^{18}O$ for the highest CH₄ sample (10 ppm). The propagated uncertainty due to the calibration correction (Eq. A23-A24) is approximately 0.31 ‰ in $\delta^{15}N^\alpha$, 0.11 ‰ in $\delta^{15}N^\beta$, and 0.35 ‰ in $\delta^{18}O$. The uncertainty due to poorly understood effects, estimated by the repeatability of target gas measurements, is 0.5 ‰ for all isotopologues. The total uncertainties (Eq. A19) for the N₂O sample gas with 10 ppm CH₄ sample are 0.63 ‰, 0.53 ‰, 0.62 ‰ for $\delta^{15}N^\alpha$, $\delta^{15}N^\beta$ and $\delta^{18}O$, respectively.

Supplement comments:

In the tables S8-10 you have calculated the mixing ratio of oxygen, though for table S9 and S10 I believe that you have incorrectly copied the concentration from CO₂ from the previous set of tables. A further correction to the correct oxygen values of table S8, is that the matrix of the Cal1 and Cal2 include oxygen, which means that you have a base concentration of oxygen at about 770 ppm rather than 0 for N₂O 0.33 ppm and double and triple that for 0.66 and 0.99 ppm N₂O. I understand that those can be considered negligible levels for the relevant purpose, but I would still prefer the correct values.

Thank you for catching this! The reviewer is correct on both these points. The target oxygen concentrations in Tables S9 and S10 were incorrectly copied from the CO₂ concentration data of the preceding tables, and the base oxygen concentration in Table S8-10 should reflect the matrix contribution from Cal1 and Cal2 rather than zero. We have corrected all three tables accordingly, with the Cal1 and Cal2 oxygen values now starting from the matrix baseline of approximately 770

ppm rather than 0 ppm for the 0.33 ppm N₂O standard, and scaling proportionally for the 0.66 and 0.99 ppm standards. For clarity, we have added this to the experimental description in Supplementary section S.1.4, stating that target oxygen mixing ratios were set at 12, 14, 16, 20, and 20.7%, and that the Cal 1 and Cal 2 reference gases contribute an additional matrix baseline of approximately 770, 1544, and 2316 ppm O₂ at N₂O concentrations of 330, 660, and 990 ppb respectively, such that target values do not start from zero as they do for N₂O.

I appreciate the inclusion of experiment 6 and find it to be a worthwhile aspect to investigate, but I find the description of it difficult to follow. In part because the text refers to gas mixes not shown in the figure S6 or described in table S13, leaving me unsure how they are introduced and diluted. Additionally the many gasses with very similar names led me to mix up which gas was being discussed. I would suggest perhaps expanding on table S13 to show the dilution scheme for all discussed dilutions. And maybe update the figure to better indicate channel 1 and 2.

- 1) We agree that the description was difficult to follow, and thank the reviewer for spotting this. We have revised the text to improve clarity in several ways; 1) We have added an explicit clarification of the key distinction between the gases used (Supplementary Information, tracked changes lines 97–100):

"The key distinction is that RM-1-1Pure is an undiluted pure N₂O standard, while RM-1-1Diluted-1 and RM-1-1Diluted-2 are gravimetrically prepared ambient-concentration dilutions of the same gas (~327 ppb; Table 1) (Mohn et al., 2014). The dynamic dilution approach prepares the same gas at ambient concentration in real time using a sequence of dilution steps."

- 2) In addition, Figure S6 has been updated to explicitly include RM-1-1Diluted-1 and RM-1-1Diluted-2 as input gas streams, making their introduction into the measurement sequence visually traceable. To avoid further confusion regarding the concentration of these gases, a note has been added to the legend of Table S13 stating:

"RM-1-1Diluted-1 and RM-1-1Diluted-2 have ambient N₂O concentrations and were therefore introduced without dilution."

And the legend of Figure S6:

"RM-1-1Diluted-1 and RM-1-1Diluted-2 have ambient N₂O concentrations and were therefore routed directly through the system without the use of a dilution gas. Cal 1 and Cal 2 were diluted in a single step in mixing chamber 1, analogous to the general experimental setup, with flow rates as given for example in Table S2. RM-1-1Pure was diluted dynamically to ambient concentrations via the three-step dilution scheme described in Table S13."

We believe these changes together resolve the ambiguity regarding which gases were diluted and how they entered the system.

Minor comments:

Line 165: You have written $^{14}\text{N}^{16}\text{O}^{17}\text{O}$ and $^{14}\text{N}^{16}\text{O}^{18}\text{O}$, where I believe you meant to write $^{14}\text{N}^{14}\text{N}^{17}\text{O}$ and $^{14}\text{N}^{14}\text{N}^{18}\text{O}$

We thank the reviewer for pointing out this error. We have changed it in the manuscript as suggested.

Line 462: There is a misplaced 6 following the (990 ppb)

We thank the reviewer for pointing out this typographical error. We have changed it in the manuscript as suggested.

Line 523: I harbor suspicions of the non-linearity measurement of the 10/4-2025, as it shows a significantly different behavior for both CRDS II & III.

Thank you for this comment. We share this suspicion and agree that the N_2O non-linearity measurement of 10.04.2025 shows a notably different behaviour for both CRDS-II and CRDS-III. As discussed in the manuscript (lines 523–524), the larger time interval between the validation experiment (20.03.2025) and the N_2O non-linearity characterisation (10.04.2025) compared to the later dataset is a plausible contributing factor, consistent with our broader observation that the N_2O non-linearity correction drifts substantially over time (Table 4). However, we recognise that we cannot fully exclude that this particular measurement was anomalous, as no repetition was conducted for comparison. Nevertheless, this pattern further reinforces our recommendation to characterise the correction function as close in time as possible to the actual sample measurements.

Line 574: Did you conduct validation measurements for the inclusion of CO_2 removal to compare the found values to your validation experiment?

This is a good point. However, no dedicated validation experiment for CO_2 removal was conducted in the study. CO_2 spectral interference on delta values seem negligible when CO_2 is the sole interferant (Exp. 3, Section 3.3, Table 6), which is also consistent with findings in Harris et al. (2020). Our recommendation to remove CO_2 is to eliminate the source of the non-additive CH_4 - CO_2 interaction documented in Section 3.5. A direct experimental validation of CO_2 removal as a practical strategy is acknowledged as a valuable future test but was outside the scope of the current study.

Minor supplement comments:

Line 17: Missing lower case 4 for methane.

Thank you for your comment. We have carefully checked the manuscript and confirmed that methane is consistently written as CH₄, with the “4” correctly formatted as a subscript throughout.

Table S1: You have in the accompanying text written that the change was in the range of 0.33-1.20 ppm, but the lowest concentration I see in the figure is 0.35 ppm.

During the initial experiments, a target N₂O concentration of 335 ppb was applied instead of the 330 ppb used in subsequent experiments. This minor discrepancy of 5 ppb arises from a slight difference in the dilution constant applied and has no meaningful impact on the results, as the derived correction functions are normalised to the reference concentration of Cal 1 in the post-processing. The concentration range in the text of Experiment S1 has been corrected accordingly to 0.35 to 1.200 ppm (Supplementary Information, tracked changes).

Review 2

L20, gives the impression that 3 brands will be tested, but they are all Picarro.

We see why the sentence can be read like that, although we did write in parenthesis (G5131-*i*, Picarro Inc.). We have changed the wording to be more clear. The new sentence reads:

*This study presents a comprehensive and detailed correction and calibration protocol to post-process N₂O isotopic data, exemplified by data obtained from three commercial G5131-*i* cavity ring-down spectroscopy (CRDS) analysers manufactured by Picarro Inc.*

L24, if only Picarro's were tested, is it fair to say “any” instrument must be routinely characterized? Would another instrument brand be more reliable?

It is true that we do not know this for other instruments, although we suspect this is the case. We have added “G5131-*i*” as exemplified below:

“Consequently, any G5131-i instrument must be...” (See tracked changed)

L25-26, The earlier Stephen Harris paper seems to suggest that some instruments have additive effects and that these corrections are possible. Are these non-additive effects specific to Picarro?

Based on our observations in this study, the behaviour described here includes non-additive components that cannot be fully corrected using purely additive approaches. Our earlier work in Harris et al. did not test additive effects but speculates on additivity of corrections. At present, we lack sufficient comparative data across different instrument types to determine whether these non-additive effects are specific to Picarro CRDS instruments or represent a more general characteristic of this detection scheme of wavelength region applied for N₂O isotope measurements. In section 4.1.1 we speculate, that effects might be related to an erroneous CO₂ interference post-correction. However, this would require retrieval and analysis of analyser raw data in agreement with recommendations by CCQM GAWG-IRWG. However, this is beyond the scope of the present study, and our data cannot add much substance to draw broader conclusions at present.

L28, the MATLAB code is certainly a nice addition. Of course not necessary, but it would be wonderful if there was an R version available so that everyone can access it.

Both reviewers have flagged this, and in principle, we agree that an R or Python implementation would improve accessibility and inclusion, but rewriting and validating the code in another language is beyond the scope of the present study. The present code was implemented in MATLAB because this is the environment in which the authors were most confident and efficient. The MATLAB source code is openly available under an open license and can be viewed and adapted without MATLAB using any standard text editor. This is a worthwhile direction for future work (see our reply to RC1).

L31, consider reversing the words to state “case study”.

We have revised the wording in L31 as suggested by the reviewer.

L39, Consider using “N₂O” consistently.

We thank the reviewer for pointing this out. We have made this change in the manuscript.

L465-466, when it is stated that the data fits well with a linear model, how is that assessed?

We used a first-order polynomial (linear) regression model in MATLAB to determine the slope and intercept of the data. The adjusted R^2 value was used to estimate the goodness of fit. The adjusted R^2 provides a measure of how well the data conform to the linear model while accounting for the number of fitted parameters. The details of the linear fit model and the formulae used for the computation of adjusted R^2 is presented in L250.

In our analysis, an R^2 value greater than 0.60 is used as a pragmatic threshold to indicate an adequate representation of the data by a linear model. For our dataset ($n = 21$, $p = 1$), this corresponds to an adjusted R^2 value of approximately 0.58. We therefore use this as a threshold to assess the goodness of the fit. We have now added this information in L265:

“In this study, values of $R^2 > 0.60$ (adjusted $R^2 \sim 0.58$ for our dataset) are used as a pragmatic threshold to indicate an adequate representation of the data by a linear model.”

Fig 1. Please include more tick labels in the X and Y axis of figure 1. Unclear what is meant by “sliced” data and it would be helpful to include that in the figure caption.

We have added additional tick labels to both the X and Y axes in Fig. 1 to improve the visual presentation of data. For the Allan variance analysis shown in Figure 1, pressurised air was used as the sample. The detailed description of raw data, sliced data and drift-corrected data is presented in L115-130. We have now added the term ‘sliced data’ in the Fig. 1 caption for clarity.

Fig 5. There might be value in adding the R^2 of the relationships to better understand the linear relationships (see above comment).

Thank you for the very helpful suggestion. The adjusted R^2 values corresponding to the data presented in Figure 5 are listed in Table 4. Since the figure already contains multiple data sets, fitted line plots and their 95% confidence bounds, we chose to present the slope, intercept and the adjusted R^2 values in a separate table to keep the figures clean and understandable. We have now revised the figure caption to explicitly state that the slope, intercept, and adjusted R^2 values for each linear fit are provided in Table 4.

Fig 10, the caption refers to solid gray lines, but I can't find gray lines in the figure other than the dashed ones.

Thanks for spotting this. Given the high information density of Figure 9 and 10, we have decided to remove both the grey and black solid lines from the figure and retain only the symbols. This should reduce visual clutter and improve readability. The captions have been updated accordingly.

Harris, S. J., Liisberg, J., Xia, L., Wei, J., Zeyer, K., Yu, L., Barthel, M., Wolf, B., Kelly, B. F., and Cendón, D. I.: N₂O isotopocule measurements using laser spectroscopy: analyzer characterization and intercomparison, *Atmos. Meas. Tech.*, 13, 2797–2831, <https://doi.org/10.5194/amt-13-2797-2020>, 2020.

Mohn, J., Wolf, B., Toyoda, S., Lin, C. T., Liang, M. C., Brüggemann, N., Wissel, H., Steiker, A. E., Dyckmans, J., and Szewc, L.: Interlaboratory assessment of nitrous oxide isotopomer analysis by isotope ratio mass spectrometry and laser spectroscopy: current status and perspectives, *Rapid communications in mass spectrometry*, 28, 1995–2007, 2014.