Response to RC2

We sincerely thank the reviewer for carefully evaluating our manuscript and providing constructive and insightful comments. Line numbers in our responses refer to revised manuscript.

General Comment

This manuscript presents methodologies of vial treatment and continuous-flow IRMS measurement for δ^{13}C-CO_2. They are fairly developed for ecosystem measurement applications and worth being shared in the isotope measurement community; therefore, this study is well within the scope of the journal. However, I do not recommend publication of the current manuscript until the following issues are addressed. In addition, Referee#1 already provided a long list of constructive comments, which I totally support and express sincere respect to the thorough evaluation.

We thank the reviewer for the feedback and for highlighting the relevance of our work to the isotope measurement community.

Given that not a few papers on continuous-flow δ^{13}C-CO_2 measurement in similar principles are already available over the last 20 years (those both cited and not cited in this study), I do not find the method novel, in contrast to what the authors write. They improved sampling and laboratory treatment methodologies for a reduced sample amount, and these careful descriptions are of value. However, from my point of view, many places in the manuscript look exaggerated in particular for novelty and outlook. I strongly recommend toning down with focus on actual application fields only.

We thank the reviewer for this remark. We agree that the analytical principle is not new and have therefore toned-down statements about novelty throughout the manuscript. Following comments from all reviewers, we now clearly emphasise that the contribution of our work lies in its practical application to volume-limited systems (e.g., growth chambers, microcosms, respiration closed systems chambers, etc.) and in providing a simple, low-cost, and rapid method based on standard equipment for reliably analysing very small air volumes. The revised manuscript reflects this clarified and more moderate positioning (lines 17, 74, 340).

One simple way to minimize storage effects is to employ a glass flask with a stopcock valve with a rubber O-ring, which has been long used in the atmospheric monitoring community (e.g., Worthy et al. 2023). The flasks are in many cases much larger (up to litres), but stability of sample air has been ensured for months to a year as the flaks are transported worldwide including Antarctica. In this regard, one might wonder why the authors take riskier septum, despite of availability of the safer method. I surmise there are several reasons such as practice in field deployment, familiarity and cost, but I did not find them well explained in the manuscript.

We thank the reviewer for raising this point. We agree that the systems mentioned—glass flasks with stopcock valves—are highly reliable and well known to ensure long-term isotopic stability. However, due to their large internal volume, they are not compatible with our experimental constraints. We acknowledge that the original manuscript did not clearly explain the rationale behind our choice of septum-sealed vials. Our approach was specifically guided by the need for simplicity, low cost, rapid handling, and the use of standard commercially available materials, while enabling the analysis of very small air volumes (1 mL).

We have now made these motivations and constraints explicit throughout the revised manuscript, including in the introduction, materials and methods, and discussion.

Throughout the manuscript, precision, which the authors consider important, is not consistently evaluated and described. In this type of experiments, I suppose that efforts are made to minimize measurement uncertainty, which includes uncertainties associated with every measurement process such as pre-treatment, storage, and isotope analysis. These processes affect both variability (precision) and systematic offset (accuracy). I believe that the authors would like to ensure minimized systematic offsets with acceptable measurement variability. Description and discussion on the former are incomplete in the current manuscript, thereby it is difficult to consider that the present method was evaluated sufficiently. Regarding the latter, the authors should define "precision" explicitly throughout the manuscript. I would use "repeatability" or "reproducibility" depending on the context.

We thank the reviewer for this remark. In the revised manuscript, we now explicitly define what we mean by precision (line 83). We chose precision because, although the procedure and instrument remain constant, the operator may vary between analytical sequences, making precision the most appropriate term for our workflow (Belouafa et al. 2017; Squara et al., 2020). These definitions and adjustments are now consistently applied throughout the manuscript.

To address accuracy, we (i) added the uncertainties of our secondary standards to Figures 3–7, (ii) performed one-sample t-tests comparing each batch to the certified reference value, and (iii) discuss these results in the relevant sections. These additions show that most apparent offsets were overestimated in the original version and that accuracy is mainly improved by storage temperature, not by septum type or dual sealing. Altogether, the revised manuscript now evaluates both precision and accuracy more completely and transparently.

The authors presumably consider that "precision" of 0.1 per mil is a criterion to evaluate quality of the measurement, but it is not clear why this value is appropriate. If the authors assume ecosystem measurement applications, one would see relatively large variability through a series of samples (e.g., some ‰), therefore relaxed "precision" might be acceptable.

We thank the reviewer for this important comment. We chose a precision target of ±0.1 ‰ for δ^{13}C–CO$_2$ based on both practical and methodological considerations. As mentioned application targeted, typically biologically meaningful processes, exhibits δ^{13}C variations of 0.3–1 ‰ (Pataki et al., 2003; Tu et al., 2001; Joos et al., 2008; Breecker et al., 2014; Leitner et al., 2023). Achieving ± 0.1 ‰ precision allows these changes to be resolved reliably.

Our objective was also to reach comparable analytical performance with standard IRMS setups (i.e. precision <0.2 ‰ with 10–12 mL samples) (Giammanco et al., 2017; van Geldern et al., 2014), using only 1 mL of air, balancing analytical resolution, operational simplicity, and constraints of small-volume experimental systems.
This rationale and justification for the 0.1 ‰ target have now been clarified in the revised manuscript (line 90).

I do not think that Table 1 is useful for readers, because the laboratory settings could change and some part of information is clear from Figure 3. Alternatively, a schematic of the whole measurement system including a sample vial connection (GasBench), ConFlo, and IRMS might be presented. The schematic along with the current Figure 3 should be presented in section 2.3.

We thank the reviewer for this helpful suggestion. Following this recommendation, we removed Table 1 and added a schematic representation of the complete measurement system adapted from Gas Bench II operating manual (Thermo Fisher Scientific., 2018)—including the sample vial connection to the Gas Bench, the ConFlo interface, and the IRMS (now Fig. 2A)—alongside the chromatographic output (now Fig. 2B previous Fig. 3). Both elements are now presented together in Section 2.3 to provide a clearer and more coherent description of the analytical workflow.

Below my specific concerns are also detailed, but I think the authors might need corrections which span the entire manuscript, as several comments apply to several relevant places.

Specific Comment

P1 L10: The "stable" carbon isotopic "ratio", if only δ^{13}C of CO_2 is referred. One might include ^{14}C if the term "composition" is preferred. Modified (line 11).

P1 L11: …, yet its extended application remains limited due to analytical and sampling restrictions. Modified (line 12).

P1 L14: "Terostat" here and other first place where it appears, the authors should describe what it is, not the trade name, e.g., sealant/adhesive tape? A brief description has been added (line 137), and the term "Terostat" has been replaced throughout the manuscript with butyl-rubber sealing compound.

P1 L19: "carbon stable" at another place, "stable carbon" was used. Harmonised throughout the manuscript.

P1 L23: "composition" to "ratio". The bracket "(δ^{13}C)" should come just after "ratio", i.e., "The stable carbon isotopic ratio (δ^{13}C) of atmospheric CO_2…" Corrected (line 26)

P2 L37: It is weird that the authors mention to the analyzer of Picarro as emerging instrument, but no reference paper is cited. The authors later cite Sperlich et al. (2022), but their instrument was not Picarro. As Picarro is not the only availability, these sentences might be reformulated in a more balance way. Sentence reformulated and reference added (line 39).

P2 L40: Picarro could measure sample air with 1/10 atmospheric CO_2 concentration (e.g., ~40 ppm), but it would make the precision worse. According to their data sheet, detection limit is not clearly defined, and also in principle it would not make sense. I would think that it is matter of whether the worse precision at low CO2 concentration is acceptable. Moreover, dilution measurement with laser-based instrument would cause different problem; change in matrix could cause measurement offset. The sentence has been revised to more accurately reflect the performance and limitations of laser-based analysers at low CO_2 concentrations (line 42).

P2 L45: …during "mixture between atmospheric and ecosystem reservoirs." Corrected (line 48).

P2 L58: At other places, the term "concentration" is used. Is "mixing ratio" used with a different meaning? Corrected, use only "concentration" in the revised manuscript.

P3 L86: "0.5 bar" is this absolute pressure or above ambient? Above ambient, this has been specified in the text (line 107).

P4 Figure 1: I wondered that there may be a close valve at the vial side of the 3-way valve, otherwise the vials cannot be evacuated with N2 or exhaust connected. The 3-way valve (SS-43GXS4) serves three functions: closing the manifold, opening the N_2 line, or opening the exhaust, eliminating the need for an additional shut-off valve. This is clarified in the manuscript (lines 106, 108 and 113).

P4 L107: Explain about "Terostat" here, so that readers without prior knowledge about the product can follow. And how was it used at where of the vial? Describe explicitly ("apply" it, as in the manuscript, says almost nothing). Corrected (line 137).

P4 L109: As in my earlier comment, a schematic figure including the whole measurement system might help readers. The figure in the reference (Fiebig et al. 2005) represent only part of the system and the authors should present the system more in detail as they improved from the earlier one. Figure 2 in

Section 2.3 presents a schematic of the complete measurement system, together with a representative chromatogram.

P4 L117: What is the cryofocus unit? Is it a tubing or a capillary of which material and size? A cryogenic trap is a U shaped 1/16 stainless steel tube fixed to a pneumatically operated lifting unit and placed above a 2 litre dewar of liquid nitrogen, these details were added in the revised manuscript (line 143).

P5 L122: "the cryofocus was raised" to "the cryofocus column (*or tubing as appropriate*) was lifted out from liquid nitrogen" Corrected.

P5 L129: This sentence does not describe traceability clearly. The authors should mention to the international reference material (RM) to which the reported values are traceable eventually. It is unclear if the authors determined δ^{13}C value of the working standard against an RM or if they have a suit of different gases whose δ^{13}C values were determined against an RM. We have added a detailed description of the calibration hierarchy and traceability in the revised manuscript. Specifically, we now clarify that the three ambient air working standards (WS1–3) used for δ^{13}C–CO_2 calibration were calibrated against secondary reference gases, which are themselves traceable to NOAA primary standards on the RAMCES platform (LSCE, France) line 177).

P5 L133: "…, and analytical precision consistently reached ±0.1 ‰ for δ^{13}C" here and other places, it is unclear if "reached" means whether the value is larger than 0.1 ‰ or not. I would avoid ambiguous verbs or adjectives when large or small matters. For instance, here the sentence could be rephrased like: …analytical precision of δ^{13}C was consistently <0.1 ‰. Note my comment on "precision" and consider it consistently throughout the manuscript. Ambiguous terms such as "reached" have been replaced with precise formulations indicating whether the precision is above or below the target threshold.

P5 Table 1: See my earlier comment. Removed.

P6 L147: In section 2.1, the authors described (O_2-free) N_2 was used to prevent possible N_2O production at ion source, but here they use air containing both N_2 and O_2. Reformulated.

P6 L155: It is unclear if the δ^{13}C value is a nominal one from the gas company (Air Liquid) or that determined by the authors' laboratory so as to be traceable to an RM. This is important because it would be no wonder if the values have independent traceabilities. We have clarified in Section 2.3 that the three working standards (WS1–3) used for δ^{13}C–CO_2 calibration are secondary standards prepared in-house and calibrated against reference gases that are themselves traceable to NOAA primary standards (RAMCES platform, LSCE, France) (line 177).

P6 L158: See my earlier comment. If the term "precision" refers to standard deviation, define so at early place of the manuscript. Now defined at the end of the introduction (line 84).

P7 L169: This section and Figure 3 could be merged into section 2.3. We have merged the figures as suggested: the content of the original Figure 3 is now incorporated as Figure 2A and 2B within Section 2.3. However, we have retained Section 3.2 as a separate section because it presents a result that is meaningful only after demonstrating the effect of flushing with dry CO_2-free air.

P8 L186: "50 µL" this amount depends on the CO_2 concentration of sample air. How much concentration corresponds to the sample amount 50 µL? The CO_2 concentration in the sample corresponding to the 50 µL aliquot ranged between 10,000 and 20,000 ppm.

P8 L194: Same comment as P5 L133. Does "exceed" mean smaller or larger than 0.1 ‰? Corrected (line 252).

P9 L210: "reduce" delete "s"; the risk of gas leakage, diffusion "and associated" isotopic drift. Corrected (line 271).

P10 L233: Same comment as P6 L155. See details added on line 177.

P11 L238: "…improved measurement precision" it seems to me more important that the low temperature storage resulted in the values in agreement to the nominal value, than the magnitudes of the error bars. In the revised manuscript, we now emphasise accuracy more clearly. Each figure includes the nominal value of the working standard together with its ± 0.1 ‰ uncertainty range, enabling direct visual comparison. We also added one-sample statistical tests comparing measured means to the certified value. These analyses show that low-temperature storage does not merely improve precision, but—more importantly—produces δ^{13}C-CO_2 values statistically indistinguishable from the nominal standard.

P11 L257: Same comment as P6 L155. See details added on line 177.

P12 L274: “composition” to “ratio” Corrected.

P13 L297–L301: For integrity, the authors should discuss total uncertainty, not only “precision.” In the revised version, the paragraph now explicitly addresses total uncertainty by discussing both precision and accuracy relative to the working standard, and by discussing the physical mechanisms (diffusion, permeation, adsorption–desorption) that influence systematic offsets during storage (line 377).

P13 L302–305: Appropriate reference should be mentioned. Otherwise, readers cannot understand whether the description is an established knowledge or the authors’ speculation. At least I do not find this discussion convincing with the current style. Ghiara et al. (2025) provide experimental evidence that CO_2 diffusion and permeation through elastomeric materials decrease sharply at cryogenic temperatures, supporting our interpretation of reduced leakage and isotopic drift at –80 °C. Similarly, Sreenath and Sam (2023) demonstrate that low-temperature conditions suppress gas transport in polymer–membrane systems used for CO_2 capture. These two studies have now been cited to the relevant section (Section 4.2).

P14 L320: As the authors explain at P14 L323, I agree that diffusion is plausibly the dominant process that caused the result with δ^{13}C offset and larger variability. The authors might come up with adsorption or desorption, but I could not find any signal that support these processes occurring. This paragraph could be reformulated. We have reformulated the paragraph to clarify that while diffusion is likely the dominant process behind the δ^{13}C-CO_2 offset and variability, minor contributions from CO_2 adsorption/desorption on vial surfaces or elastomeric seals cannot be excluded, especially in small-volume samples. We support this point with recent literature (Aoki et al., 2022; Schukraft et al., 2022; Ghiara et al., 2025; Sreenath and Sam, 2023). See sections 4.2 and 4.3.

P14 Section 4.4: I think discussion of this section is too general and thereby reads exaggerated. For instance, isotope equilibrium with water vapor during storage largely matters in δ^{18}O of CO_2 measurement, thereby it would go much less smoothly. The authors additionally explain prior treatment of vials, but storage over days, weeks and months are totally different, and I would not be optimistic as written currently. If the authors have a specific application plan deemed feasible only with the result presented in this study, they could focus on it in this section, but otherwise I think it is difficult to keep this section in a style convincing to readers. Modifications have been made to refine the adaptation for δ^{18}O–CO_2, limit extrapolation regarding storage, and clearly highlight the applications under testing or in realistic conditions (section 4.4).

P15 Conclusion: The conclusion (as well as abstract) section should be reformulated after all issues are addressed. Done.

P15 L360: I do not think that the exact methods in this study can help reduce cost of long-term monitoring network, because they analyse δ^{13}C of CO_2 using sub aliquot of original flask samples (use of extra vials would increase cost). We removed the statement regarding long-term monitoring networks. Instead, we highlight more appropriate applications, such as long-term laboratory incubations and specialised microcosm experiments, as well as inter-laboratory sample exchange, where small-volume δ^{13}C–CO_2 sampling provides a clear advantage.

P15 L362: A week is too short for international collaborations. Extra efforts to cool samples would also complicate logistics. The following sentence also reads like overstatement. Removed.