Response to RC3

We thank the Reviewer for this constructive review. Below we provide a point-by-point response to all comments. Line numbers refer to the revised manuscript.

Sauze et al. (2025) developed an analytical workflow to measure the carbon isotopic composition of CO_2 in small atmospheric air samples (1 mL). Their work demonstrates that all the successive steps described in the paper are necessary to achieve a good precision (0.1‰). The ability to measure isotopic composition in such small quantities opens new perspectives in ecological and paleoenvironmental research. For this reason, the study fits well within the scope of the journal.
I find this study interesting, particularly in the way it tests different parameters. However, as already noted by the other two reviewers, several points need to be addressed before publication.

We thank the reviewer for the positive evaluation of our study and for highlighting the relevance of developing an analytical workflow capable of measuring the carbon isotopic composition of CO_2 in very small air samples. All comments have been addressed point-by-point in the detailed responses below, and corresponding modifications have been implemented in the revised version of the manuscript.

I also believe that the authors should explain in more detail why achieving a precision of 0.1‰ is novel. This could be related to Berryman's 2011 study, where a precision of 0.3‰ was reported.

We thank the reviewer for raising this point. We agree that the relevance of a 0.1 ‰ precision target need to be clearly justified and we have clarified this rationale in the revised manuscript (line 90).
In atmospheric and ecosystem applications, $\delta^{13}C$–CO_2 signals typically vary by only a few tenths of a per mil, and 0.1 ‰ is widely considered the precision required to resolve meaningful biological or ecosystem-level processes.

Current IRMS setups (including Gas Bench II systems) generally deliver precision <0.2‰ using 10–12 mL of air. Our aim was to reach comparable performance with 1 mL of sample—an order-of-magnitude reduction—while maintaining analytical robustness and compatibility with small chamber volumes and high sampling frequency.

Quality of the figures

Although this study focuses on optimizing precision in $\delta^{13}C$ measurements, it is difficult—based solely on the figures—to assess the difference or similarity between the measured precision and the accepted values reported in the literature.

For example, in all figures the mean value is shown in red, but what are the corresponding precisions and accepted standard deviations? How far are the measured values from the reference ones?

We thank the reviewer for this comment. In the revised manuscript, figures 3-7 have been updated to display the reference $\delta^{13}C$ value (red dashed line) and its associated uncertainty range (± 0.1‰) as a shaded band. This allows direct visual evaluation of how each treatment, condition, or protocol compares with both the reference value and its accepted precision.

Need for more quantitative information

Since this study is based on the analysis of very small amounts of CO_2, more quantitative detail and precision are needed. For example:

- L93: "A control set of vials (i.e., blanks) was analyzed and showed no detectable CO_2 signal."
- L249: "These blank vials showed detectable CO_2 signals, confirming the integrity of the sealing protocol and the absence of background contamination during storage under ultra-low-temperature conditions."

The authors mention possible contamination, but without quantifying the blanks and background, how can one assess the effect on the results? If the $\delta^{13}C$ of atmospheric CO_2 is around –8‰, could this explain the higher $\delta^{13}C$ values measured compared to the reference value of the working standard? A possible mixing process could be discussed.

We thank the reviewer for this constructive comment. To quantitatively assess blank integrity, we reprocessed all blank chromatograms using higher sensitive detection settings (minimum peak height 5 mV; start/stop thresholds 1 mV s^{-1}), well below standard analytical thresholds used here. Under these conditions, small CO_2 peaks (10–15 mV) were detected in approximately one-third of the blanks (line 116).

We now also discuss the theoretical possibility of mixing with atmospheric CO_2 ($\delta^{13}C \approx -8‰$). Given that blanks remain below detection limits, any such contribution would represent a negligible fraction of the CO_2 in a sample and would fall well within analytical uncertainty (line 361).

Accuracy

Although this study focuses on improving $\delta^{13}C$ precision, it does not address accuracy. There are often systematic offsets between the mean measured values and the reference values, but these are not discussed. For example, why does the single-septum test (l. 189, 3.3) show an offset from the reference value that is not observed in the other tests?

We thank the reviewer for this comment. We agree that the apparent discrepancies between target and measured values required further clarification. Accuracy is now explicitly addressed. In the revised figures, we display the working-standard reference value together with its ± 0.1‰ uncertainty envelope, enabling direct comparison. The apparent offsets are considerably reduced, showing that the method is more accurate than initially assumed.

We also performed statistical tests comparing the certified reference value (treated as fixed) with the mean $\delta^{13}C–CO_2$ values obtained for each batch, and the results of these tests are now discussed in the relevant sections.

Finally, we provide an overall interpretation indicating that storage temperature had a clear positive effect on accuracy, whereas septum type and dual sealing did not yield any convincing improvements (section 4.2).

Even if measurements were performed using a continuous-flow setup, it would be useful to test different IRMS settings. What were the instrumental parameters and operating conditions used during the study? Providing this information could help other researchers reproduce the method.

The analytical setup for $\delta^{13}C–CO_2$ is now shown in Figure 2A. All operating details—including carrier gas flow, GC temperature, peak detection thresholds, number of reference injections, autosampler settings, and others—are described in sections 2.1 and 2.2 of the revised manuscript.

Additional aspects that could be discussed include:

1. the impact of instrumental drift during analysis,
2. the memory effect, and
3. the background signal

Data accuracy is strongly linked to background levels, with increasing background often leading to less accurate and less precise results. What are the quantitative values associated with these three parameters, and how were the data corrected for them?

(1) Instrumental drift was monitored through systematic injections of the working standard at regular intervals during each run (at both the beginning and end of each daily sequence). Drift remained below 0.05 ‰ over the longest sequences (4–5 h), which is well within the analytical uncertainty. No time-dependent trend was detected, and this information is now reported in the revised manuscript (line 185).

(2) Memory effects were assessed during the experiment described by Siegwart et al. (2023). Highly contrasting samples were analysed in a random order, resulting in successive injections of CO_2 with very high concentrations (10,000–20,000 ppm) and highly enriched isotopic signatures (up to 3.5 AT%), followed by samples with atmospheric CO_2 concentrations (~420 ppm) and depleted signatures (~ –10‰). No memory effect was detected under these conditions, indicating that carry-over is negligible and does not require additional correction or mitigation in our analytical setup. This is now specified in the revised manuscript (line 171).

(3) It is not entirely clear what is expected regarding the background signal. Instrumental background is already corrected by the software through the blanking procedure (line 167). Background originating from the vials is now discussed in the revised manuscript: under highly sensitive detection settings, small CO_2 background peaks (10–15 mV) were detected in approximately one-third of the blanks. These correspond to very small amounts of CO_2 and remain well below standard analytical detection thresholds (50 mV) (line 361).

Finally, the authors mention diffusion as a plausible cause of some effects. It would be useful to provide references or a brief explanation to support this hypothesis

Ghiara et al. (2025) provide experimental evidence that CO_2 diffusion and permeation through elastomeric materials decline sharply at cryogenic temperatures, supporting our interpretation that leakage and isotopic drift are minimized at –80 °C. Likewise, Sreenath and Sam (2023) show that low temperatures substantially suppress gas transport in polymer–membrane systems designed for CO_2 capture. These findings are now incorporated into the discussion of diffusion processes in Section 4.2.