

Dear Editor,

Thank you very much for reviewing the revised version of our manuscript.

- Concerning the use of the monitoring data, following your recommendation, we refer to the paper of Friedlingstein et al. (2023) as an example instead of the Global Carbon Project website. This paper describes the methodology for the calculation of the atmospheric CO<sub>2</sub> budget from the concentration measurements (and other data).

- You are right, the low average bias does not necessarily mean that the measurements continuously met the WMO compatibility goal, therefore, so this statement has been deleted. I think details of the flask – in situ comparison would not fit into the main text, therefore, the new figures have been added to the Supplement. One of the figures shows all available data, while the other shows the average bias and the scatter of the data aggregated into 5-year bins. In addition to the statistical analysis, not much can be said about the data. In the first figure (Fig. S2 upper panel), it can be seen that the scatter of the data and the number of extreme outliers were higher when the Li-Cor 6251 analyzer was used until 2007. Both the change to the LI-7000 analyzer and a more trustworthy technician might contribute to the lower scatter and a lower number of extreme outliers after 2007 but we do not know anything about the technical changes at NOAA. The reasons for the recent increase in the scatter and the appearance of a few outliers are not clear yet. Most of the extreme outliers do not correlate with the variability of the concentration in the time window used for the comparison. This suggests that the extreme outliers are likely caused by sampling (e.g. contamination, improper flushing, misdating of the sample, etc.) or analytical errors, and the omission of these data is justified.

The second figure (Fig. S2 lower panel) aggregates the bias and the scatter into 5-year bins. Not considering the first period, the 5-year average biases are always within the WMO extended compatibility goal of 0.2 μmol mol<sup>-1</sup>. This indicates at least that the measurements at Hegyhátsál are on the NOAA scale and that the quality of the measurements is fairly consistent throughout the monitoring period.

- The reference to Barcza et al. (2020) has been rephrased and extended.