Response to reviewer comments

We thank reviewers for the critical comments and helpful suggestions. We have taken all these comments and suggestions into account, and they have improved our manuscript considerably. A point-by-point response to reviewers' comments please found as below.

Responses to Referee 2:

Wu et al. (2023) provide a valuable analysis focusing on the impact of continental site inclusion when calculating global CO2 growth rates. Employing the CTE model, the authors conduct synthetic tests to ascertain the accuracy of various growth rate estimate methods. The study is a valuable contribution to our understanding of the sampling error in the growth rate of atmospheric CO2 and is generally well-conceived. Nonetheless, the paper would benefit from clarifications and adjustments to enhance its readability and coherence.

Main Comments:

Presentation quality: The primary analysis of the paper focuses on the impact of including the continental sites for calculating the global CO2 growth rate. The study compares growth rate estimates from three sets of observations using, in essence, the NOAA's growth rate method:

1. NOAA: MBL sites only
2. WDCGG: MBL and some continental sites
3. CTE: MBL and a more extensive inclusion of continental sites

Given the many tests conducted and the slight variations between them, I recommend presenting this information in a table. Please specify in the table what is being compared to what is in each test to enhance the clarity of the methodology.

Response: We appreciate the reviewer for suggesting this.

The three observation networks are clarified in Table 1 (in paper or two pages below) and Fig. 1 caption (Lines 106-111): "Three observation networks are employed to assess the impact of continental site inclusion when calculating global CO_2 mole fraction and its growth rates. The NOAA network (43 sites, yellow stars) comprises MBL sites only. The selected GAW global network (139 sites, red dots) includes both MBL sites and continental sites, for example from the Advanced Global Atmospheric Gases Experiment (AGAGE) and European ICOS contribution network. The CTE network serves as the global network for the CTE model evaluations (230 sites, blue dots), comprises MBL sites and a more extensive inclusion of continental sites."

We have created four pair-wise comparison heatmaps, as shown in the table below, to enhance the clarity of comparisons among various networks, methodologies, periods (depending on data availability), and temporal resolutions. For example, the Figure below (Fig. 5 in the paper) displays a monthly comparison for the period 2001-2020.

	2000-2020	1980-2020
Monthly	Fig 5 (paper)	Fig S1 (supplementary)

Annual	Fig S2 (supplementary)	Fig S3 (supplementary)

We have revised our manuscript to make this point clearer (Lines 216-219): “The statistical metrics assessing the agreement of these monthly comparisons are available in Fig. 5 (for 2001-2020) and Fig. S1 (for 1980-2020). The statistical metrics for the annual comparisons can be found in Fig. S2 (for 2001-2020) and Fig. S3 (for 1980-2020).”

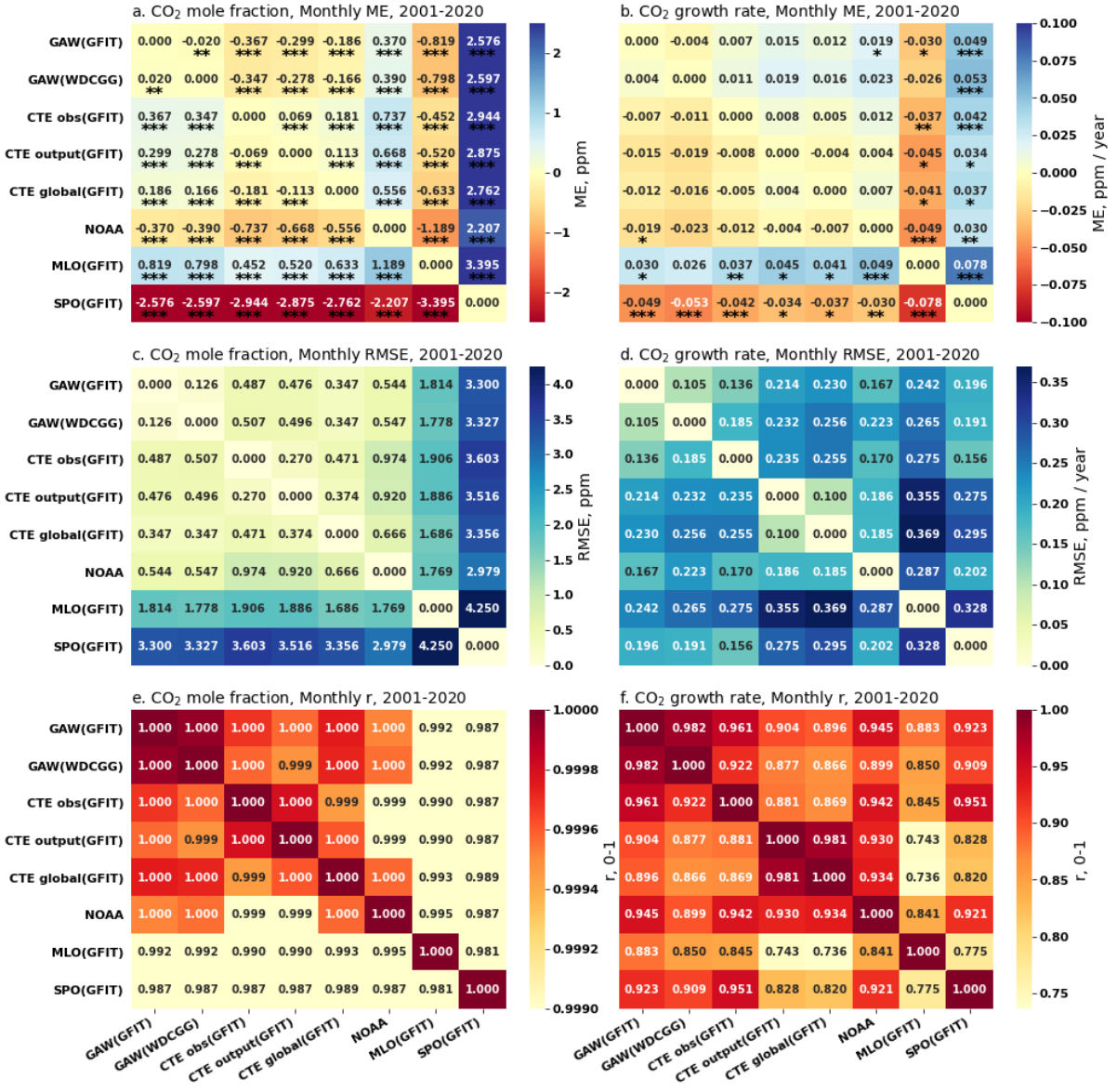

Figure 5. Pair-wise statistical metrics assess the agreement of monthly global and local CO_2 mole fraction (ppm) and its G_{ATM} (ppm yr^{-1}) across various networks and methodologies (see Table 1 and Fig. 4) for the period 2001-2020. Panel (a) presents the Mean Error (ME) quantifying the difference for each pair, focusing on CO_2 mole fraction, while panel (b) does the same for G_{ATM}. The significant levels of paired t-test for ME are indicated as follows: * p<0.1, ** p<0.05, *** p<0.01. Panel (c) and (d) present the Root Mean Squared Error (RMSE) for CO_2 mole fraction and G_{ATM}, respectively. Panel (e) and (f) present the Pearson Correlation Coefficient (r) for CO_2 mole fraction and G_{ATM}, respectively.

"The semi-NOAA method": The authors introduce a method called "semi-NOAA," adding unnecessary complexity to the presentation. The approach is not new, mainly the NOAA

approach on an observation set including continental sites. Referring to all the filtering and fitting procedures as components of the original NOAA method would be more effective. Subsequently, the authors could delineate any variations they are implementing compared to the standard NOAA and WDCGG methods.

Response: To avoid confusion, the method right now is named GFIT, which stands for "global fit procedure". We acknowledge that the GFIT method is not a novel approach; instead, it represents a hybrid method derived from both the standard NOAA and WDCGG methods. In the method section, we have already referred the station selection and CO_2 averaging method to the WDCGG approach (Text S1).

However, when it comes to the filtering, fitting, and growth rate calculation steps, we have chosen to describe and illustrate them in the method section. This decision is deliberate, as we believe it aids readers in understanding the GFIT method, particularly those who may not be familiar with the NOAA method. Without this clarity, readers might perceive these aspects as a 'black box.' Thus, our preference is to provide a detailed description and illustration of the filtering, fitting, and growth rate calculation within the method section, thereby enhancing understanding of the GFIT method.

Minor Comments:

1. I suggest modifying the abstract to clearly state the study's purpose: to evaluate the impact of using continental sites in CO2 growth rate calculations. It drifts off by introducing the "GFIT" method, which I do not think is the main point of this work.

   Response: We appreciate the reviewer's suggestion. We have improved the abstract.

2. It needs to be clarified how CTE is precisely used. CTE is sometimes a network, a growth rate, and a transport/inversion model run. Please use more clear terminology to differentiate. State this information in a table.

   Response: We appreciate the reviewer's suggestion. In response, we created a table (Table 1 in the paper) to differentiate and clarify the observation network and its analysis method. This table is placed in front of the results section. Specifically, CTE alone stands for CarbonTracker Europe model.

Table 1. Description of the three observation networks and their analysis methods.

Terminology	Description
NOAA network	NOAA network comprises MBL sites only (43 sites).
GAW network	The selected GAW global network (139 sites) includes both MBL sites and some continental sites.
CTE network	The CTE network serves as the global network for the CTE model evaluations (230 sites), comprises MBL sites and a more extensive inclusion of continental sites.
GAW (GFIT)	GAW network observations analyzed using the GFIT method
GAW (WDCGG)	GAW network observations analyzed using the WDCGG method without extrapolation

GAW (WDCGG+)	GAW network observations analyzed using the WDCGG method with extrapolation
CTE_obs (GFIT)	CTE network observations analyzed using the GFIT method. The observations come from the ObsPack data product (Schuldt et al., 2022)
CTE_output (GFIT)	CTE model output at the 230 sites (sampled at the same location, altitude and time) analyzed using the GFIT method
CTE_global (GFIT)	CTE model output for full global grids (averaged over the first three levels, 0 to 0.35 km Alt.) analyzed using the GFIT method
MLO (GFIT)	Mauna Loa (MLO) observations analyzed using the GFIT method
SPO (GFIT)	South Pole (SPO) observations analyzed using the GFIT method

3. The study mainly addresses monthly and multi-decadal scales. I suggest adding an analysis on annual growth rates, which have been the scales that NOAA and WDCGG report the growth rates.

   Response: In this study, we conducted analyses at both monthly and annual temporal resolutions, as we mentioned earlier. To facilitate comparisons, we created four pair-wise comparison heatmaps, as described above (also see Fig. 5, S1, S2 and S3).

4. Throughout the manuscript, excessive use of parentheses interrupts the reading flow. Consider using tables to present some of the information the reader can refer to easily.

   Response: We appreciate the reviewer's suggestion. We have incorporated the suggestions by increasing the number of tables and heatmaps to enhance the presentation, meanwhile reducing the parentheses.

5. Many sentences are unnecessarily long and could be divided into shorter, more readable sentences.

   Response: We appreciate the reviewer for the suggestion. We have reduced the number of long sentences and improved the flow and clarity.

Technical Corrections:

1. Line 82: The term "biased" seems unfair when referring to NOAA's estimate.

   Response: We have revised the sentence (Lines 84-86): "The NOAA estimate of global surface annual mean CO_2 mole fraction is expected to be lower (e.g. ~0.35 ppm lower than the WDCGG estimate, Tsutsumi et al., 2009) compared to a full global surface average because areas with large sources are not represented."

2. Lines 184-189: These lines could be made clearer to understand.

   Response: We appreciate the reviewer's comment. We have removed the mentioned lines, and instead we created a table (Table 1, as previously presented) to provide clarity regarding the observation network and its analysis methods.

We have also revised the sentence (Lines 192-194): “Global averaged surface CO_2 and its G_{ATM} are calculated using the WDCGG method and our GFIT method based on the data from the GAW and CTE networks (Fig. 1). The different observation networks and their analysis methods are listed in Table 1.”

3. Line 328: Explain the acronym IVA.

   Response: We appreciate the reviewer for pointing this out. It was a typo; it should be the interannual variability (IAV). We have now corrected it.