Response to Reviewer #2

This study evaluates the performance of various carbon flux products against eddy covariance measurements at three forest sites in France. The authors investigate the monthly, seasonal, and inter-annual variability of NEE, GPP, and RECO to assess different global products and explore their relationships with meteorological variables. While the manuscript is generally written clearly, the analysis lacks sufficient depth and significance for the scientific community. This isn't to say that evaluating existing products isn't valuable, but the limited number of eddy covariance sites and the selection of only four global products raise concerns about the comprehensiveness of the study.

We would like to thanks Reviewer 2 for their constructive feedback, which will greatly contributed to the improvement of the manuscript. In response to both reviewers' comments, we have decided to enhance the study by including approximately 17 additional ICOS sites, incorporating more models (specifically the TRENDY ensemble and FLUXCOM-X), and focusing exclusively on the monthly timescale. This adjustment will enable a more detailed exploration of the influence of climate on the temporal variability of CO_2 fluxes.

In the initial version of the manuscript, this influence was analyzed by considering all months together, which yielded results that closely resembled those at the annual timescale (e.g., see Figs. 6 and 8). In the revised version, we will examine the influence of climate on CO_2 fluxes for each month individually, as we anticipate that this influence will vary throughout the annual cycle. The inclusion of additional sites will also allow us to analyze the zonal and meridional variations in this influence across Europe.

1.Choice of Models (LPJ-GUESS and FLUXCOM v1): Why were LPJ-GUESS and FLUXCOM v1 selected when many other land surface models or upscaled products are available for evaluation? Including more models and products could improve relevance, especially since ensemble models are commonly used in land carbon sink studies.

The paper does not aim to evaluate all existing CO_2 flux models. One objective is to assess the strengths and limitations of the CarbonSpace data-driven model in comparison to several widely used models. The CarbonSpace model is distinctive due to its very high spatial resolution and its ability to differentiate between tree species.

We acknowledge that the use of a single land surface model (LPJ-GUESS) was not ideal, given the large spread across land surface models. To address this, we will incorporate the TRENDY v12 ensemble (S3 simulations, which include time-varying CO_2, climate, and land use) to better account for uncertainties in process-based models.

The choice of FLUXCOM remains appropriate, as it is frequently referenced in the literature. However, we will consider adding the FLUXCOM-X product, which provides global monthly CO_2 data at an unprecedented spatial resolution of 0.05°.

2.Spatial Resolution Mismatch: The 50 km spatial resolution of LPJ-GUESS and FLUXCOM products may not align with the footprint of the eddy covariance sites, and this mismatch is not addressed in the manuscript.

Comparing coarse-resolution models with local measurements can indeed result in significant discrepancies that may not necessarily reflect model errors. While we do not anticipate a one-to-one match, we expect that the main observed patterns to be captured by both data-driven and process-based models. It is also important to note that all the data-driven models used in this study are trained using FLUXNET2015 local measurements, making comparisons between CarbonSpace/FLUXCOM estimates and local measurements meaningful. We will include a brief discussion on the impact of spatial resolution mismatches.

3.Insufficient Number of Sites: Only three forest sites are used in the analysis, despite the availability of hundreds of FLUXNET sites globally. Using only these sites may not provide a robust basis for summarizing product performance.

The initial aim of this paper was to evaluate data-driven and process-based models to capture the annual cycle, interannual variability, and trends of CO_2 fluxes in temperate deciduous broadleaf (DBF) and evergreen needleleaf (ENF) forests of Western Europe. We recognize that the number of sites analyzed was insufficient. To address this, we will expand the analysis to encompass all of Europe. This expansion will allow us to include 17 additional ICOS sites that were not part of the initial version, resulting in a total of 20 ICOS sites under study. Each of the 20 ICOS sites provides at least five years of data (Figure 1).

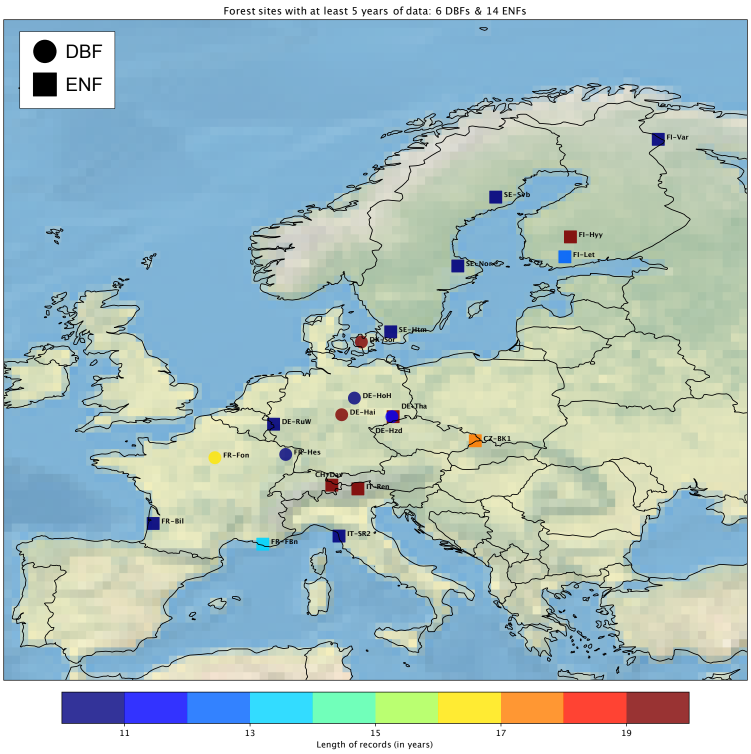

Figure 1: ICOS sites providing CO_2 fluxes for at least 5 years.

4.Correlation Analysis: The correlation analysis in Figure 6 lacks a logical basis, as some variables (e.g., VPD and RECO) do not have clear biogeochemical or biophysical relationships. Also, the analysis does not account for multicollinearity among variables, which affects the validity of the results.

While we recognize that the physical influence of VPD on GPP and NEE is more evident than on RECO, it is important to note that autotrophic respiration is closely correlated with GPP. As a result, VPD indirectly influences RECO through its impact on GPP.

The other meteorological variables (2 m temperature, soil moisture, total precipitation, and real and potential evapotranspiration) exert a significant, physically grounded influence on NEE, GPP and RECO, justifying their inclusion in the analysis.

While multicollinearity poses a strong issue when analysing the combined influence of climate variables on CO_2 fluxes (e.g., in multiple linear regression), this study computes correlations variable by variable, which is methodologically sound and unaffected by multicollinearity.

5.Focus on Temperature: The authors only consider temperature when analyzing carbon fluxes in Figure 6 and do not include other important variables, such as soil moisture, which were emphasized in the introduction. Given this, the use of polynomial regressions without considering other factors raises questions about interactive effects of multivariate factors of the carbon fluxes.

The polynomial regressions shown in Figure 7 are valuable for discussing temperature-induced threshold effects, known to affect both GPP and RECO. However, with the inclusion of new models and additional sites, Figure 7 would occupy too much space. Therefore, it will be omitted from the revised manuscript.

Specific Points:

Inconsistent Visualization (Figures 5 and 13): Figures 5 and 13 present similar data for annual and monthly scales, but the visualizations need to be consistent to enable direct comparison.

Figure 13 will be removed in the revised manuscript to maintain a focus solely on the monthly timescale and further investigate the climate – CO_2 flux relationship along the annual cycle.

Climate Anomalies Definition: The authors should clarify how climate anomalies are defined, as the methods section only explains CO2 flux anomalies. Also, the choice of the -0.5/+0.5 thresholds for carbon flux anomalies seems arbitrary and needs further justification.

Climate anomalies are computed in the same manner as CO_2 flux anomalies, resulting in variations in the reference period depending on the site. With regard to the composite analysis, higher thresholds (-1/+1 and -1.5/+1.5) significantly limit the sample size due to the short duration of the data. Therefore, we have opted for -0.5/+0.5 thresholds as a compromise. In the revised manuscript, we will provide a more detailed explanation of how climate anomalies are computed and justify our choice of thresholds in Section 2.4 (Methodology).