Response Letter to Reviewer #2

CEDAR-GPP: spatiotemporally upscaled estimates of gross primary productivity incorporating CO_2 fertilization

Dear Dr. Wang,

Thank you very much for your thorough review of and positive comments on our manuscript. We greatly appreciate your constructive feedback, which has significantly improved our paper. Below, we outline our responses (in blue) to your comments (in black) and detail the corresponding revisions made to the manuscript (in red). We included line numbers from the "track-change" version of the revised manuscript for your reference.

Thank you once again for the time and efforts you dedicated to reviewing our manuscript. We hope that our revisions have adequately addressed your comments and concerns. We eagerly look forward to any further feedback and suggestions that you may have!

Sincerely,

Yanghui Kang
On behalf of all co-authors

General comments:

This manuscript presents a new global GPP dataset by using the machine learning method and many various datasets. They mainly considered the direct CO2 fertilization effect on GPP through both the machine learning and the theoretical approaches. They then modelled the short and long-term global GPP by using different periods of earth observations and climate variables. They also validated their new GPP data by comparing to flux sites GPP data and other products.

This study fits well for ESSD and I appreciate the efforts that the authors made. Several minor concerns from my side are listed below, which I think may be helpful for revising the manuscript.

Specific comments:

Line 20: Maybe not the "first", since you also said that the revised EC-LUE GPP have considered this.

Response: Thank you for pointing this out. We have revised this sentence to elaborate our product was the first dataset upscaled from eddy covariance measurements using machine learning models that incorporated the direct CO2 fertilization effect.
Revisions: Line 20 – 21 (Abstract): *Here, we introduce CEDAR-GPP, the first global machine-learning-upscaled GPP product that incorporates the direct CO2 fertilization effect on photosynthesis.*

Line 114: Some flux sites have the data before 2001? Why did include them into the model? And, this means that the information before 2001 from flux sites are not included into your GPP products, right?

Response: Thank you for the comment! It is correct that eddy covariance data before 2001 were not included in our models. While we recognize that sites with data back to the 1990s are valuable for evaluating long-term changes, such sites are very few (e.g. only four sites have data before 1996) and could induce model biases towards a limited range of environmental conditions in these early years. This is particularly critical for modeling the relationship between GPP and CO2. Therefore, we were concerned that applying such models globally could result in unreliable predictions. Moreover, the short-term models were based on MODIS data that was only available after 2000, and thus did not require training data beyond the time range. This approach also ensures the consistency between our short-term (ST) and long-term (LT) models.
Revisions: Line 140 – 143 (Section 2.1): *Note that we did not include eddy covariance data before 2001, since it was limited to only a few sites. This scarcity might introduce biases in the machine learning models, particularly in the relationship between GPP and CO2, leading to unreliable extrapolations across space and time*.

Line 214: Both the MODIS and CISF data are used in the model. Will them from the same observations, i.e., the CSIF is also be trained from MODIS data right. Will the multi-collinearity issue exist in this?

Response: This is a great point. Indeed, CSIF, along with other variables derived from MODIS, are highly correlated. We expect the XGBoost model to effectively manage such multi-collinearity in the feature space, as it does not rely on assumptions about the statistical distributions of the predictor variables. While multi-collinearity may affect model interpretability, repetitive information shared among the different MODIS-derived variables can actually enhance the machine learning model's predictive ability and reduce noise.
Revisions: Line 319 – 321 (Section 2.3.3): *Without relying on prior assumptions about the functional forms or statistical distributions, the model is also robust to multi-collinearity between the predictors in our dataset, particularly for the variables derived from MODIS data*.

Line 252-256: When adding the CO2 into the machine-learning model, the output of your model is GPP, but LUE. That means, the model could not distinguish the direct CO2 fertilization on LUE and the indirect CO2 impact on FAPAR? A simple way to check this is to check the CO2 impacts on GPP, LUE and FAPAR at some FACE sites, and to check that if your model is correct.

Response: Thank you for the feedback and suggestion! We agree that disentangling the direct CO2 effect on LUE and the indirect effects on FAPAR is a major challenge within machine learning frameworks, which primarily focus on optimizing the overall predictive accuracy of GPP.

Indeed, FACE experiments offer unique opportunities to quantify the direct CO2 effect on LUE and indirect effects on LAI, as thoroughly analyzed by previous studies analysis (e.g. Ainsworth et al., 2005). However, applying our machine learning models directly to such experimental conditions is challenging. These experiments involved elevated CO2 levels (e.g., 542 ppm in Duke, 547 ppm in Oak Ridge) significantly above current ambient concentration (420 ppm as of 2020). Given that the machine learning models do not extrapolate over unseen conditions, they cannot reliably produce predictions under the FACE experiment scenarios.

Our objective was to provide comprehensive quantifications of the overall GPP trends, encompassing both the direct and indirect CFE, to address omissions of the direct effects in previous upscaled datasets. By incorporating both CO2 and remote sensing proxies indicating canopy structure (LAI, fAPAR) as predictors, our machine learning models (i.e. CFE-ML) aimed to capture the overall CO2 effects. The enhanced performance of the CFE-ML models compared to the baseline models, in capturing the significant increase of GPP from eddy covariance data, suggests that the CFE-ML model effectively captured the direct CFE absent in the Baseline models, providing an improved representation of the overall trend.

In light of your valuable feedback, we have expanded our discussion (Section 4.2) to emphasize the importance and challenges of isolating the direct and indirect CFE with the machine learning frameworks and highlight future opportunities through advanced machine learning and causal inference techniques. Moreover, we have extended our model validation against eddy covariance data across different plant functional types and climate zones (Section 3.1.3, Figure 5, Figure S3), and included further analysis at sites undergoing long-term GPP changes (Figure S4). These additional analyses aimed to demonstrate our models' capacity to capture the overall GPP trends including both the direct and indirect CFE.

Revisions: Line 818 – 823 (Section 4.2): *Moreover, disentangling the direct CO2 effects on LUE, water use efficiency, and its indirect effects on fAPAR remains challenging with machine learning models due to the correlations and interactions between CO2 and other climatic or environmental factors. Future work may exploit explainable machine learning and causal inference to unravel the complex mechanisms and distinct pathways of CO2 effects on vegetation carbon uptake.*

Line 467 – 485 (Section 3.1.3): *Aggregated eddy covariance GPP experienced increasing trends of varied magnitudes across different climate zones and plant functional types (Figure 5b,c; Figure S3b,c). While the machine learning models generally did not fully capture the enhancement in GPP for most categories, the CFE-ML and/or CFE-hybrid models consistently outperformed the Baseline models in both ST and LT setups. The CFE-ML setup predicted a higher trend than CFE-hybrid in most cases, suggesting that the data-driven approach captured more dynamics not represented in the theoretical model, which was based on conservative assumptions regarding the CO2 sensitivity of photosynthesis (see Sect. 2.3.3 and Appendix A). The choice of remote sensing data (ST vs. LT configurations) did not lead to substantial differences in the predicted GPP trend. Most long-term flux sites (at least 10 years of records) with a significant trend experienced an increase in GPP, and the CFE-ML and/or CFE-hybrid models again aligned closer to eddy covariance data than the Baseline models (Figure S4). Additionally, we found a considerably higher trend in eddy covariance GPP measurements derived from the day-time versus night-time partitioning approach, potentially associated with uncertainties in GPP partitioning methods (Figure S4). Yet, machine learning model predicted trends were not strongly affected by GPP partitioning methods (Figure S3, S4).*

Line 261-262: A constant χ value for global and long-term analysis seems will introduce large uncertainties. Since we know that χ will change with different climate conditions, and also will change largely with the CO2 concentration. Although we know that the modelling of global χ is very hard, but some discussions and limitations on this aspect is at least needed.

Response: Thank you for the feedback! We totally concur that applying a constant χ value in global and long-term analysis can introduce uncertainties. We have expanded our discussion to underscore the dynamics of χ in response to meteorological conditions and associated uncertainties, and provided insights for potential improvements in future work.

Revisions: Line 792 – 799 (Section 4.2): *For simplicity, we assumed a fixed ratio of leaf internal to ambient CO2 (χ) representing an average long-term value typical for C3 plants in the* *theoretical CO_2 sensitivity function. However, χ varies by environmental conditions, including temperature and vapor pressure deficit, and robustly modeling these dependencies remains challenging (Wang et al., 2017). Future work could incorporate more comprehensive representations of the χ and evaluate how the associated uncertainties affect the quantification of GPP and its temporal variations.*