

Response Letter

CEDAR-GPP: spatiotemporally upscaled estimates of gross primary productivity incorporating CO₂ fertilization

Referee #1

Dear reviewer,

We are grateful for your thorough and constructive feedback, which has helped us improve our manuscript. We have carefully considered each comment and made the following key changes to address your suggestions:

- 1) We conducted additional site-level evaluation using 11 independent sites in tropical and arid areas, which were previously not involved in model development.
- 2) We expanded the intercomparison analysis for global and regional GPP trends using three process-based models forced with remote sensing data.
- 3) We developed further analysis of the direct CO₂ effect on GPP trends
- 4) We added new illustrations and discussion on uncertainties associated with the estimated GPP trends.

Below we provide our point-to-point response to reviewers' comments, with revisions highlighted in red. For revisions, we provide line numbers from the track-change revised manuscript.

As I pointed out, the trend of GPP is the most important result for model effectiveness, so validation is very important. Besides, I have provided several sites (comment #8) for model data validation in the last round of revision, however, the new version of the manuscript has nothing new about it. Because the site validation is the most effective way for the direct validation for GPP trend, and the ESSD also highlight the robustness for dataset especially in the model validation. So the validation has to be done.

So I think four possible ways to prove the results are robustness especially the trend of GPP

[Reviewer Comment 1] 1、 Site validation in the EBF at tropics.

The authors should use at least 6 or more sites to validate other results at EBF at tropics, because the sites in the EBF tropics contributes a very high amount of global GPP and now the sites in this study cannot reproduce the spatial representativeness. So please follow the [comment#8] from the last round of revision and validate it.

Response: Thank you for your suggestion of enhancing site validation in EBF and tropics. In response, we expanded our validation to include the new OzFlux FluxNet dataset

<https://data.ozflux.org.au/portal/pub/viewColDetails.aspx?collection.id=1882723&collecti>

[on.owner.id=450&viewType=anonymous](#)), which contains 11 Australian sites, previously not involved in our model training. This dataset contains three tropical sites (savannas) and two non-tropical EBF sites (**Table S3**). Notably, it includes two sites you previously suggested—AU-Lit (Litchfield) and AU-ASM (Alice Springs). Regarding the other five suggested sites, we note that long-term high-quality data necessary for trend evaluation were not available. A detailed explanation of each site’s data limitation are provided at the end of the response.

In the validation based on the OzFlux FluxNet dataset, we evaluated annual trends in seven sites with at least four years of good-quality data (**Figure S9**). Of these, only two sites - AU-Cpr (Tropical) and AU-Stp (Arid) exhibited a trend with p-value < 0.3. The CEDAR-GPP model estimates align closely with the observed trend. The other sites showed strong interannual variabilities, which CEDAR-GPP captured effectively, particularly in AU-ASM (savanna) and AU-Cum (EBF) (**Figure S9, Figure S10, Table S4**).

Within OzFlux and our existing data from FLUXNEXT2015, ICOS, and AmeriFlux datasets, we have a total of 5 tropical and 5 EBF sites (**Figure R1**). However, high-quality data spanning over five years are available only in 1 tropic and 3 EBF sites. This still limits our ability to assess aggregated trends for tropical areas and the effect of CO₂ fertilization.

Lastly, we carefully assessed the sites suggested by the reviewer. AU-Rob, GF-Guy, and BR-Sa1 were part of FLUXNEXT2015, thus already included in our analysis. Unfortunately, AU-Rob had only one year of high-quality data in FLUXNET2015. BR-Sa1 had frequent missing values for at least three months per year from 2002 to 2011. GF-Guy was located in a coastal area and did not contain GPP estimates from the night-time partitioning method. Moreover, while both the Xishuangbanna and CongoFlux sites are valuable for enhancing representation in Africa and Asia, Xishuangbanna only has five years of data with large interannual variation, and CongoFlux data collection started after 2020, thus lacking sufficient long-term observations for trend evaluation.

Nevertheless, the most robust assessment of GPP trends and CO₂ fertilization between observations and models remains the comparison of aggregated trends, for which we have included the maximum amount of data that met analysis criteria and results indicate that models which incorporate direct CO₂ effects improve the match with observed GPP trends compared to models that do not include direct CO₂ effects

Table S3. Sites from the OzFlux FluxNet dataset used for independent validation.

Site ID	IGBP	Koppen zone	Data range	No. of site-months
AU-ASM	SAV	Arid	2010-2019	111
AU-Adr	SAV	Tropical	2007-2009	19
AU-Boy	SAV	Temperate	2017-2019	24
AU-Cpr	SAV	Arid	2011-2019	104
AU-Cum	EBF	Temperate	2014-2019	71
AU-Dry	WSA	Tropical	2010-2019	90
AU-GWW	SAV	Arid	2013-2019	83
AU-Lit	SAV	Tropical	2015-2019	53
AU-Rgf	CRO	Temperate	2016-2019	39
AU-Stp	GRA	Arid	2009-2019	114
AU-War	EBF	Temperate	2013-2019	53

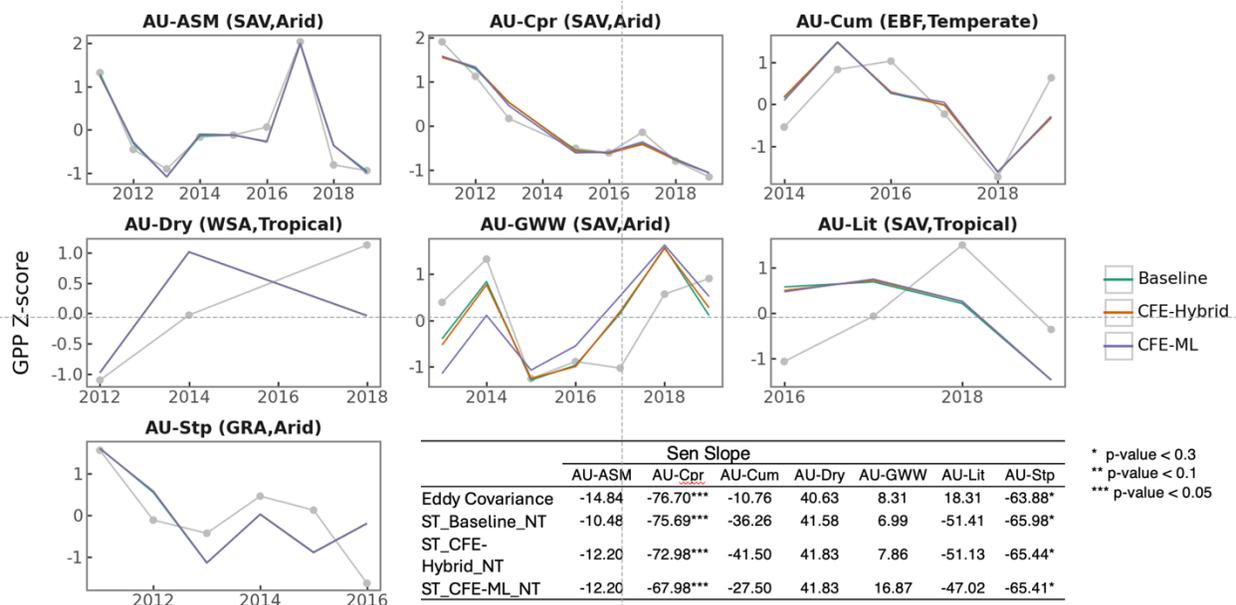


Figure S9. Standardized annual GPP anomalies from eddy covariance data and estimated by CEDAR-GPP for seven independent sites from the OzFlux FluxNet dataset. The results compare three CEDAR-GPP model setups – ST_Baseline_NT, ST_CFE-Hybrid_NT, and ST_CFE-ML_NT. Eddy covariance GPP was partitioned using the Night-time (NT) approach. The bottom right inset table lists the annual GPP trends based on Sen’s slopes and the Mann-Kendall test.

Table S4. CEDAR-GPP model performance based on independent data from the OzFlux FluxNet dataset

Model Setup	R ²				RMSE			
	Overall	MS C	Anomalies	Cross-site	Overall	MS C	Anomalies	Cross-site
ST_Baseline_NT	0.75	0.77	0.33	0.77	1.27	1.23	0.77	1.00
ST_CFE-Hybrid_NT	0.75	0.77	0.33	0.77	1.27	1.23	0.77	0.99
ST_CFE-ML_NT	0.75	0.77	0.33	0.76	1.27	1.24	0.77	1.01
LT_Baseline_NT	0.74	0.80	0.26	0.77	1.29	1.15	0.81	0.98
LT_CFE-Hybrid_NT	0.74	0.79	0.26	0.76	1.28	1.16	0.81	1.00
ST_Baseline_DT	0.73	0.74	0.50	0.69	1.40	1.40	0.67	1.17
ST_CFE-Hybrid_DT	0.73	0.74	0.50	0.69	1.39	1.40	0.67	1.16
ST_CFE-ML_DT	0.74	0.74	0.50	0.69	1.38	1.39	0.67	1.16
LT_Baseline_DT	0.74	0.78	0.43	0.72	1.37	1.29	0.71	1.11
LT_CFE-Hybrid_DT	0.74	0.77	0.43	0.71	1.38	1.30	0.71	1.13

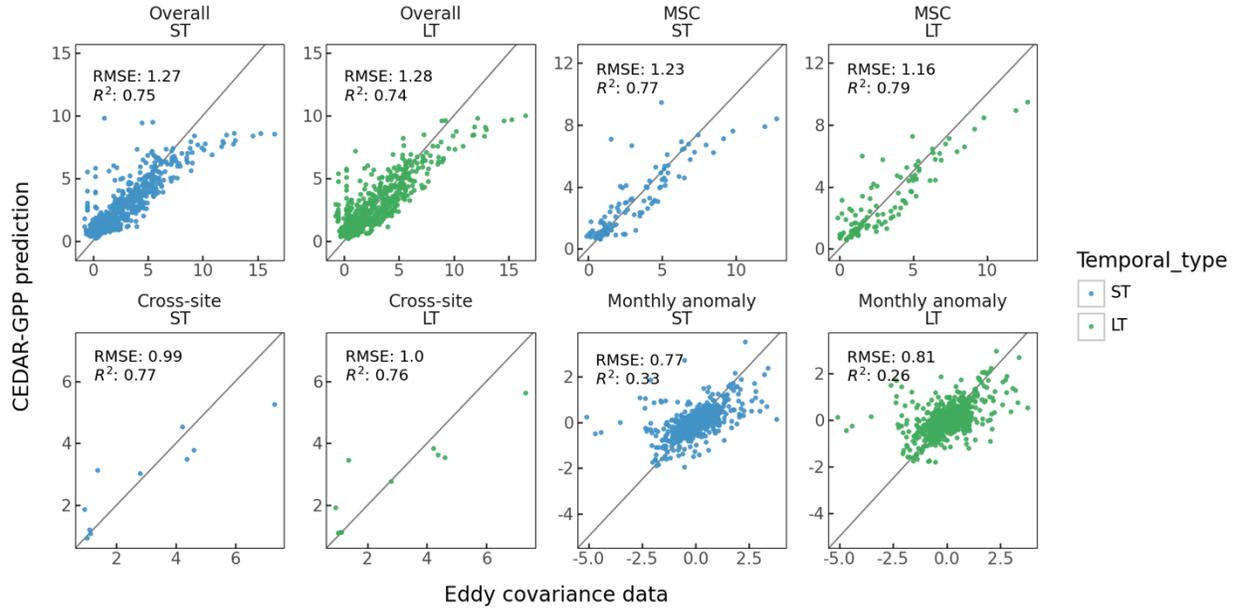


Figure S10. CEDAR-GPP performance in estimating monthly GPP, mean seasonal cycle, monthly anomalies, and spatial variations in the OzFlux FluxNet dataset.

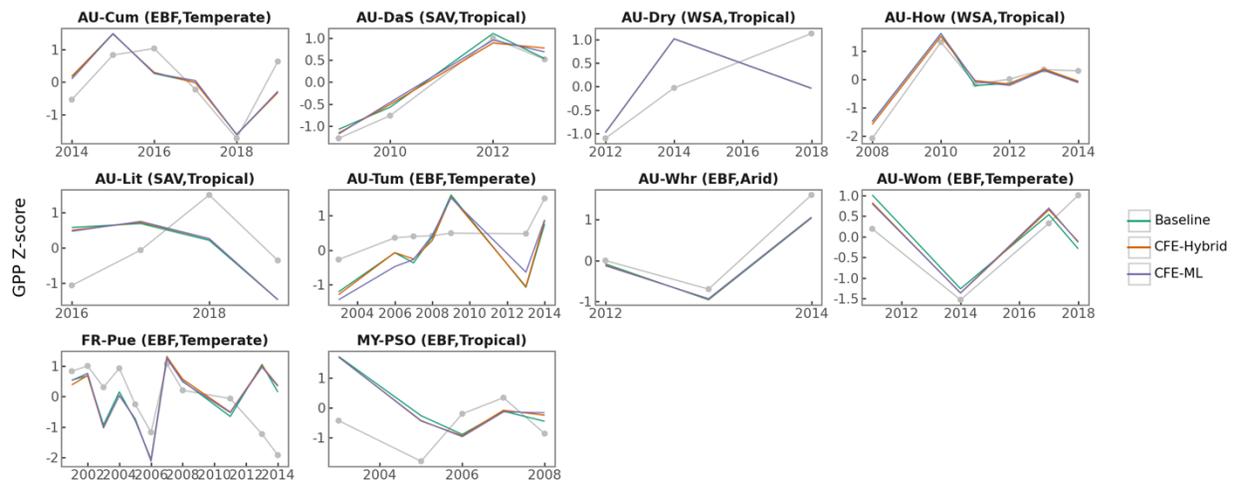


Figure R1. Standardized annual GPP anomalies of tropical and EBF sites.

Revisions: Table S3, S4; Figure S9, S10

Line 517 – 523: Finally, we evaluated CEDAR-GPP using independent eddy covariance data (11 sites, Table S3) that was not involved in model training and obtained from the OzFlux FluxNet dataset (Ozflux, 2024). Among these sites, only two - AU-Cpr (Tropical) and AU-Stp (Arid) - with more than five years of records exhibit a GPP trend with p-value less than 0.3. CEDAR-GPP shows strong consistency with the observed trend (Figure S9). Additionally, CEDAR-GPP achieves reasonable accuracy in predicting monthly GPP ($R^2 \sim 0.73 - 0.75$), mean seasonal cycle ($R^2 \sim 0.74 - 0.78$), and monthly anomalies ($R^2 \sim 0.26 - 0.50$) (Table S4, Figure S10), closely aligning with the cross-validation results.

[Reviewer Comment 1] 2、 Provide the uncertainties or more model validation on the global trend of GPP.

The authors could compare their GPP trend with other existing long-term GPP products (e.g., FLUXCOM, FLUXSAT, MODIS). However, all of them did not include the direct CO₂ fertilization effect. Some new datasets, such as BESS (Li et al. 2023), BEPS (Leng et al. 2024), PML-v2 (Zhang et al. 2019), have the direct CO₂ fertilization effect (CFE), but with a different model structure. The authors should compare their trend of GPP with these datasets that include the direct CO₂ fertilization effect.

Leng, J., Chen, J. M., Li, W., Luo, X., Xu, M., Liu, J., ... & Yan, Y. (2024). Global datasets of hourly carbon and water fluxes simulated using a satellite-based process model with dynamic parameterizations. *Earth System Science Data*, 16(3), 1283-1300.

Li, B., Ryu, Y., Jiang, C., Dechant, B., Liu, J., Yan, Y., & Li, X. (2023). BESSv2. 0: A satellite-based and coupled-process model for quantifying long-term global land-atmosphere fluxes. *Remote Sensing of Environment*, 295, 113696.

Zhang, Y., Kong, D., Gan, R., Chiew, F. H., McVicar, T. R., Zhang, Q., & Yang, Y. (2019). Coupled estimation of 500 m and 8-day resolution global evapotranspiration and gross primary production in 2002–2017. *Remote sensing of environment*, 222, 165-182.

Response: Thank you for the suggestion. It is beneficial to compare our GPP trend estimates with other models that incorporate the direct CO₂ fertilization effect. In the revised manuscript, we have followed the reviewer's suggestion and expanded our evaluation to include BEPS, BESS, and PLM-V2. We found that CEDAR-GPP aligned well with BESS and BEPS in global and regional GPP trends, while patterns in PML V2 showed considerable differences, especially in Africa and tropics (**Figure S17, S18**). We have incorporated these results in the manuscript as detailed below. Additionally, we note that we previously incorporated the intercomparison with rEC-LUE which also considered the CO₂ effects.

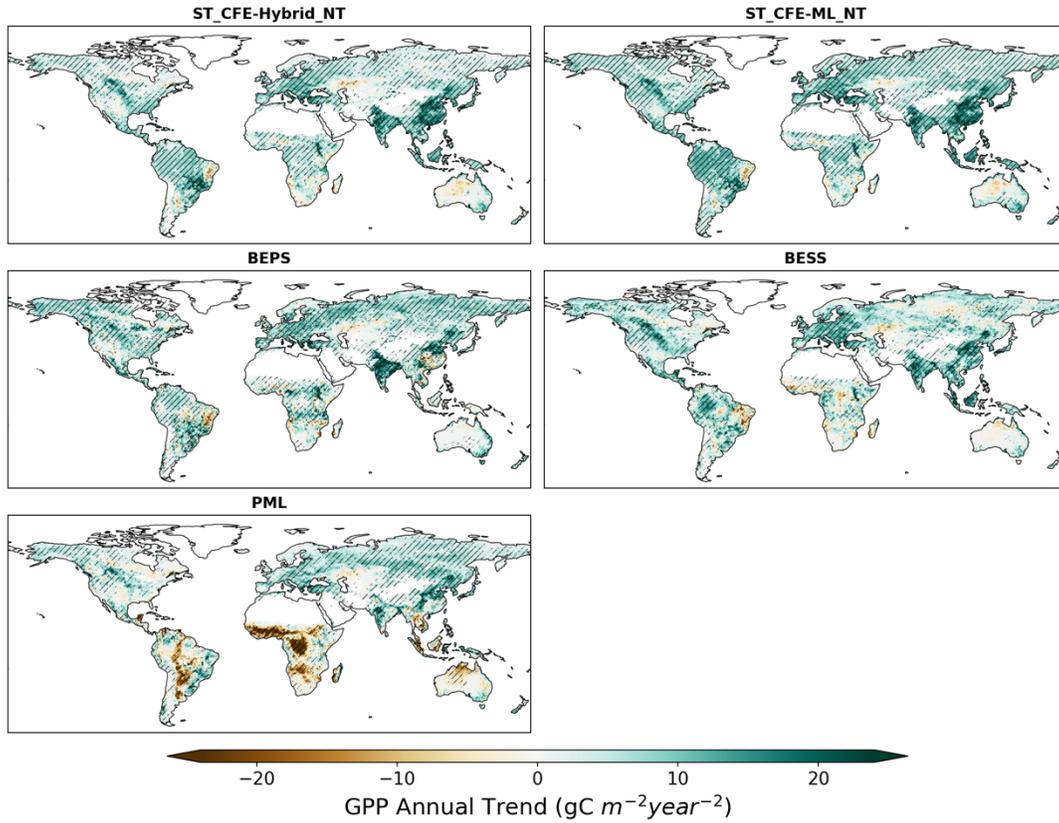


Figure S17. Annual GPP trends over 2001 – 2002 from short-term night-time CEDAR GPP datasets, BEPS, BESS v2, and PML V2. Hatched areas indicate the GPP trend that is statistically significant at $p < 0.05$ level under the Mann-Kendal test.

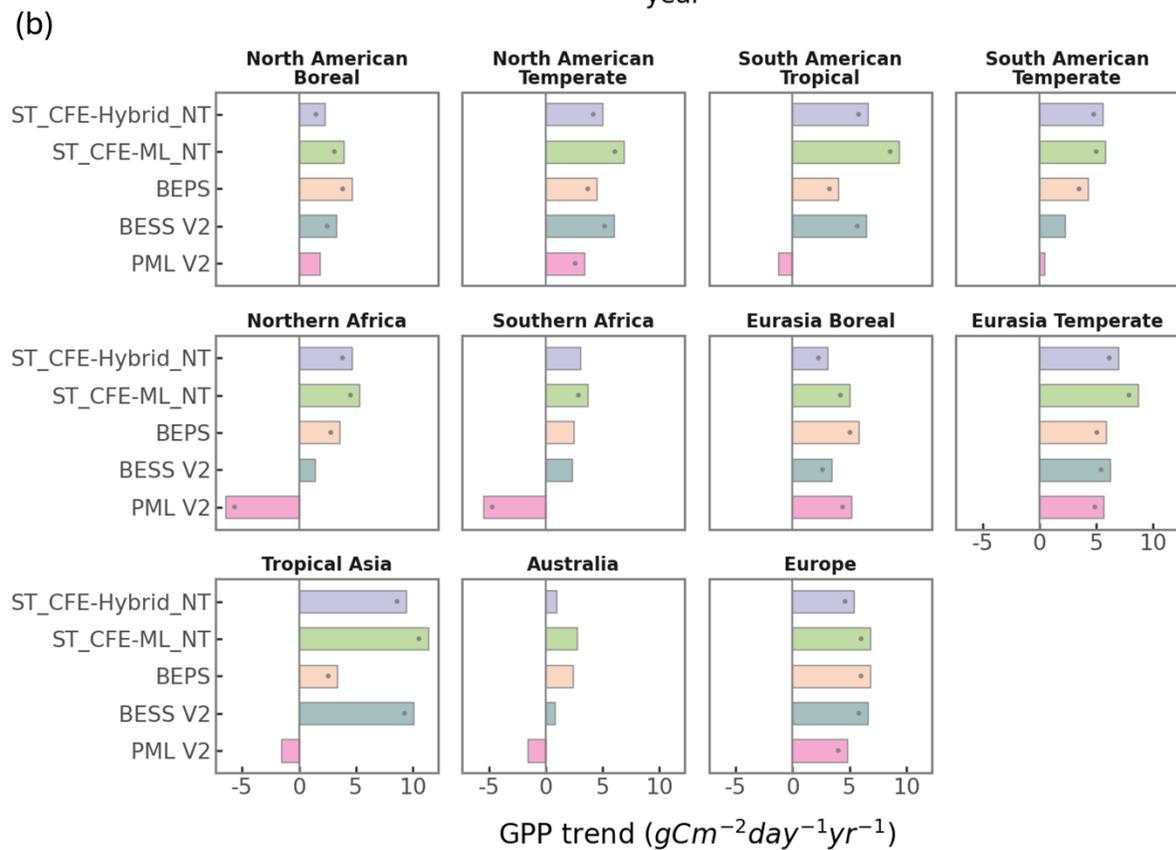
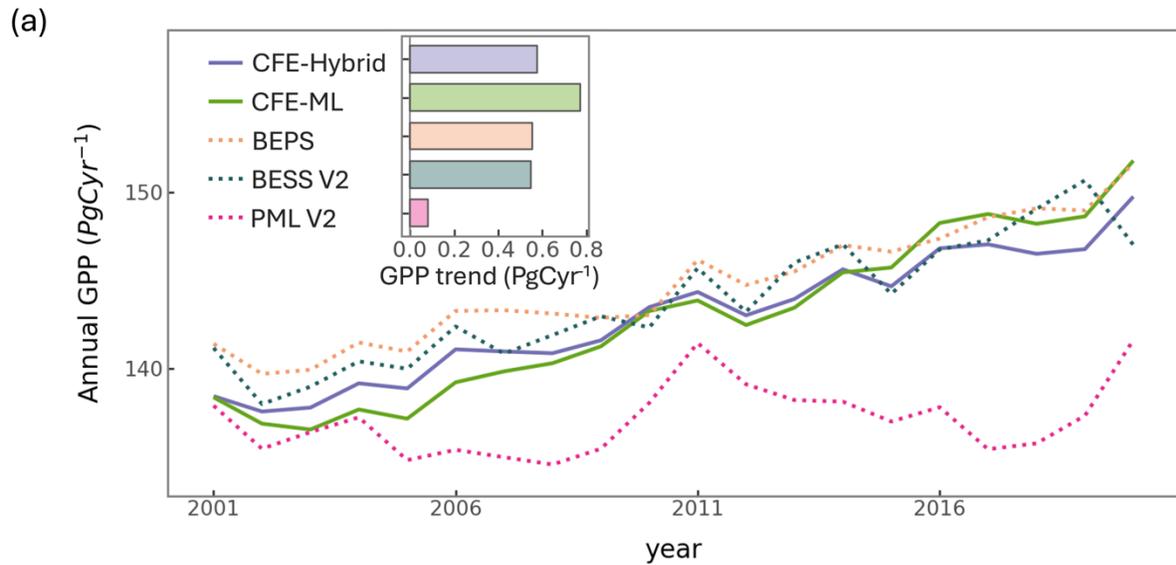


Figure S18. Global and regional GPP trends from 2001 to 2020 short-term night-time CEDAR GPP datasets, BEPS, BESS v2, and PML V2. a) Global annual GPP over time, with an inset showing GPP trends. b) GPP trend in 11 TRANCOSM regions. Bars marked with a grey dot represent statistically significant trends at $p < 0.05$ level under the Mann-Kendal test. Figure S16 shows a map of 11 TRANCOSM regions.

Revision

Revisions: Figure S17, S18

Line 386 – 390: Additionally, to evaluate GPP trends, we compared our product against three process-based models forced by remote sensing data – BEPS (Leng et al., 2024), BESSv2 (Li et al., 2023a), and PML V2 (Zhang et al., 2019). These products estimate GPP by scaling leaf-level biochemical photosynthesis models to the canopy level, using satellite-derived vegetation structural variables such as LAI. All three products incorporate the direct CO₂ effects within their biochemical photosynthesis models.

Line 660 – 668: Among the process-based models forced by remote sensing data, CEDAR-GPP aligns closely with BESS v2 and BEPS, showing widespread positive GPP trends in the boreal and temperate regions of the Northern Hemisphere, as well as across tropical areas (Figure S17, S18). In contrast, PML V2 presents minimal GPP changes in tropics and substantial reduction in Africa. The CEDAR-GPP CFE-ML model exhibits higher trends in tropical and northern hemisphere temperate regions, while BEPS shows more pronounced trends in boreal regions. None of the products indicates significant GPP trends in Australia. Globally, trends from BEPS and BESS V2 are around 0.55 PgC year⁻¹, consistent with CEDAR's ST_CFE-Hybrid_NT model at 0.58 PgC year⁻¹, while PML V2 displays a neutral trend of 0.08 PgC year⁻¹.

Line 875 – 880: CEDAR CFE-ML and CFE-Hybrid models align well with two process-based models forced with remote sensing data which consider direct CO₂ effects (BESS and BEPS). Nevertheless, considerable differences between CEDAR-GPP and other remote sensing products that include direct CO₂ effects (rEC-LUE and PML V2) warrant more in-depth investigations into long-term GPP responses to changes in atmospheric CO₂ and climate patterns.

[Reviewer Comment 3] 3、Evaluating the sensitivity of CFE to GPP trend at the site level

As the direct CFE is important, the authors could conduct a new analysis on the sensitive of CFE to GPP trend. This could provide a new point to demonstrate how the direct CFE affects GPP trends at site levels. The analytical framework could follow Chen et al. 2022. These results could be added to the supplementary data.

Chen, C., Riley, W. J., Prentice, I. C., & Keenan, T. F. (2022). CO₂ fertilization of terrestrial photosynthesis inferred from site to global scales. *Proceedings of the National Academy of Sciences*, 119(10), e2115627119.

Response: We agree with the reviewer that evaluating the contribution of the direct CFE on GPP trends offers valuable insights. As the only distinction between the CFE and Baseline models is the incorporation of the direct CFE in the former, the observed differences in GPP trends between these two types of models reflect the impact of direct CFE on the GPP trends (see Fig. 5).

While Chen et al. utilize partial derivatives ($\Delta\text{GPP}/\Delta\text{CO}_2$) from their analytical evolutionary optimality model to quantify the direct CO₂ effect, such an approach is not applicable to machine learning models due to the lack of analytical solutions. Therefore, in the revised manuscript, we used two explainable machine learning approaches, Accumulated Local Effect (ALE) and SHAP, to isolate the direct CO₂ effect on GPP. ALE

is a global explaining method, similar to partial dependence plot. It evaluates how a feature (e.g. CO₂) influences the prediction of a machine learning model on average, across the dataset. SHAP is a local effect model, describing the prediction of a single instance (i.e. a site-month) as the sum of individual feature effects.

In our revisions, we present the CO₂ dependence plots (GPP responses to CO₂) using ALE and SHAP (**Figure S8**). Both analyses demonstrated an increase of GPP in response to CO₂, as expected from theory (Chen et al., 2022).

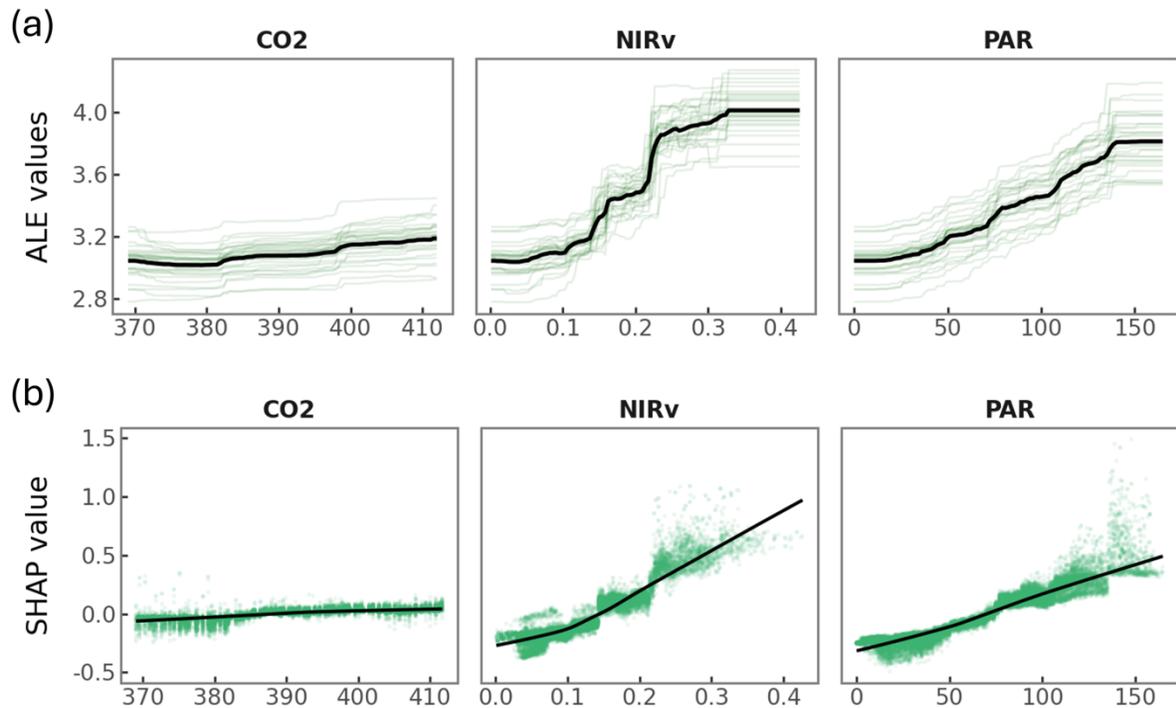


Figure S8. GPP responses to CO₂, NIRv, and PAR from the ST_CFE-ML_NT model evaluated with (a) the Accumulated Local Effects (ALE) and (b) the SHAP (SHapley Additive exPlanations) explaining approaches. Light green lines in (a) represent ALE response curves of 30 model ensembles, and the thick black presents the ensemble mean curve. Green dots in (b) correspond to SHAP values of individual samples (i.e. GPP observation from one site-month).

Revisions: Figure S8

Line 350-357: To further analyze GPP responses to CO₂ in the CFE-ML models, we leveraged two explainable machine learning approaches: ALE (Accumulated Local Effects) and SHAP (SHapley Additive exPlanations). SHAP is a model interpretation method derived from game theory, providing a value for each feature's contribution to a prediction, elucidating how each feature impacts the model's output in a specific instance. Conversely, ALE quantifies the average effect of a feature across the data, isolating its impact by aggregating local effects and avoiding the biases associated with correlated features.

Line 510-516: The differences in estimated GPP trends between the Baseline and CFE models underscore the significant long-term GPP changes driven by the direct CO₂ effect. Using explainable machine learning approaches – ALE and SHAP, we further assessed the CFE-ML models for quantifying the direct CO₂ effect. Both approaches revealed a consistently positive influence of CO₂ on GPP, aligning with biophysical theories (Figure S8). Compared to the effects from light (PAR) and vegetation structures (e.g. NIRv), the impacts of CO₂ are considerably smaller, which explains the minimal differences in overall model accuracy between the Baseline and CFE models.

[Reviewer Comment 4] 4、 Provide a new result for the uncertainties in the GPP trend. I appreciate that the authors provide the new results in Fig.12. But I think it is not enough. The authors could plot these results according to this rule. First, provide the GPP trend spatially in CEDAR (subplot a, unit in gC m⁻² yr⁻¹ in GPP trend); second, the minimum GPP trend (could be the minimum trend of 30 models or the original GPP trend minus the std of the trend) spatially in CEDAR (subplot b, unit in gC m⁻² yr⁻¹ in GPP trend); third, the maximum GPP trend spatially in CEDAR (subplot c, unit in gC m⁻² yr⁻¹ in GPP trend). With these results, the user could easily understand the trend of GPP in CEDAR and its uncertainty globally.

Response: Thank you for your feedback. We have included a new figure according to your suggestion for our six CFE models (**Figure S20**). Additionally, we note that we have previously provided the error range in the global GPP trends for our models (Figure 11), which showed the 25% and 75% percentiles as well as minimum and maximum trends for all CEDAR-GPP models. In the revisions, we have further detailed the standard deviations of the trends within the text.

Revisions: Figure S20

Line 625 – 630: The ST_Baseline_NT model predicts a GPP growth rate of 0.35 (± 0.02) Pg C per year, aligning with FLUXCOM-RS, but lower than FLUXSAT (0.51 Pg C yr⁻²) and MODIS (0.39 Pg C yr⁻²) (Figure 11b). The CFE-hybrid models estimate a notably faster GPP growth at 0.58 (± 0.03) Pg C yr⁻². The CFE-ML models predict the highest trends, up to 0.76 (± 0.15) Pg C yr⁻² from the ST_CFE-ML_NT model and 0.59 (± 0.13) Pg C yr⁻² from the ST_CFE-ML_DT model.

Line 639 – 642: The LT_Baseline_NT and LT_Baseline_DT models show a trend of 0.13 (± 0.02) and 0.15 (± 0.02) Pg C yr⁻² respectively, while the LT_CFE-Hybrid_NT and LT_CFE-Hybrid_DT models double these rates with 0.33 (± 0.02) and 0.31 (± 0.03) Pg C yr⁻² respectively. rEC-LUE shows a two-phased pattern with a strong increase in GPP from 1982 to 2000 (0.54 Pg C yr⁻²), followed by a decreasing trend after 2001 (-0.20 Pg C yr⁻²) (Figure S16).

Line 682 – 685: Figure S20 further illustrates the trend uncertainties with the ensemble mean error range based on one standard deviation. Both the CFE-ML models show large discrepancies between the upper and lower uncertainty ranges particularly within

the tropics. Additionally, the long-term models also show a higher uncertainty compared to the short-term models.

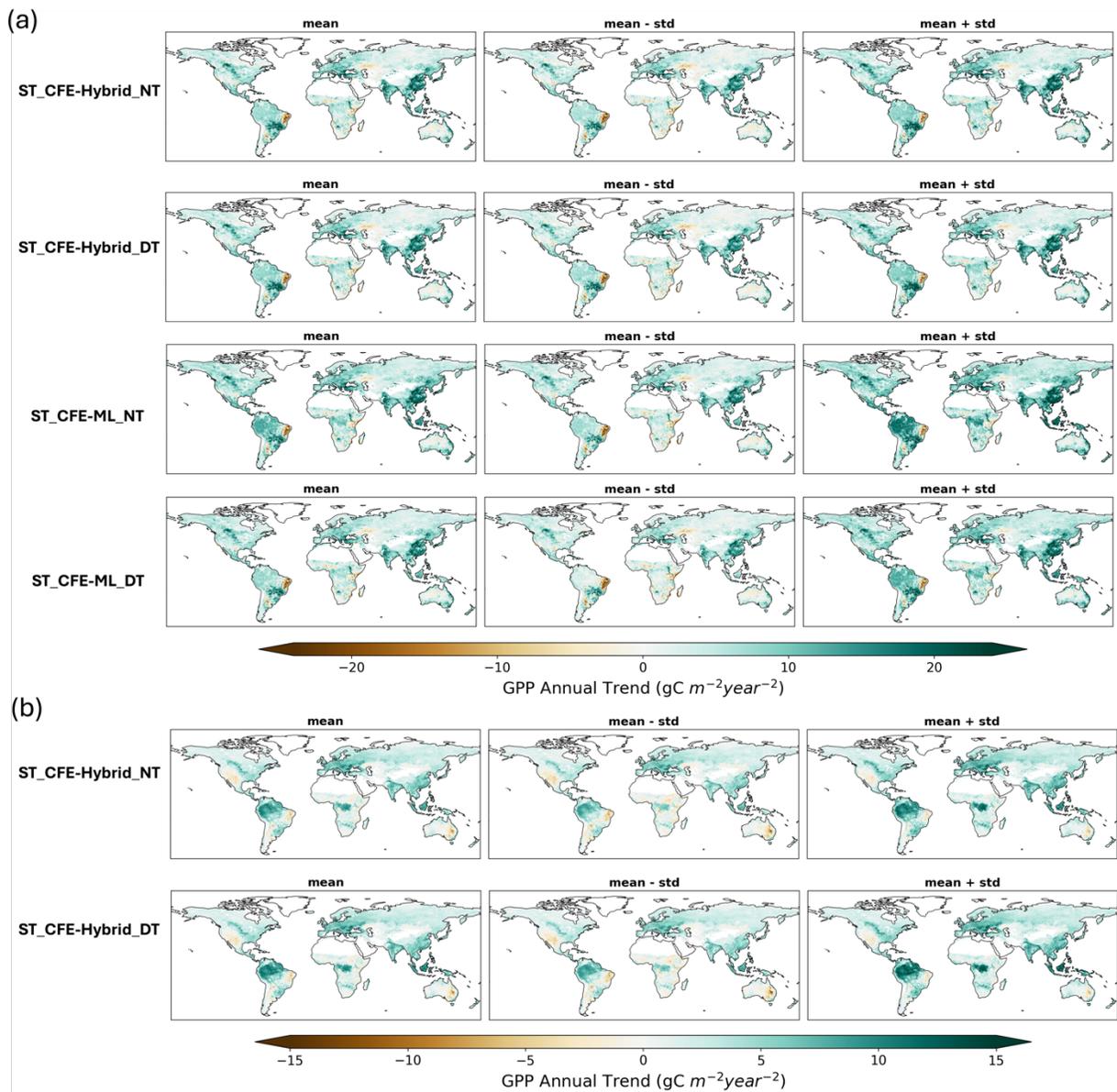


Figure S20. Maps of GPP trends and uncertainty range for CEDAR-GPP CFE datasets. The first column presents ensemble mean trends, the second column shows trends from the mean minus one standard deviation, and third column indicates the trend from the mean plus one standard deviation. (a) Trends from the short-term (ST) datasets evaluated from 2001 to 2020. (b) Trends from the long-term (LT) datasets evaluated from 1982 to 2020.

Referee #4

Thank you for the extensive review, my only suggestion at this point would be to change the color scale in Fig. 9, as it is very difficult to differentiate (even though there are redundant labels).

Dear reviewer, thank you for your positive feedback. We have revised color scheme in Fig. 9 to enhance readability.