We are grateful to the editor and referees for their time and energy in providing helpful comments and guidance that have improved the manuscript. In this document, we describe how we have addressed the reviewers' comments. Please note that the quantified results have slightly changed due to adjustments in the temporal range (1982–2011) and the datasets (specifically, the removal of GOSIF GPP). Referee comments are shown in black italics and author responses are shown in blue regular text. A manuscript with tracking changes is attached at the end.

Reviewer #2:

General comments

The authors utilized existing GPP observations and multi-model simulations to examine latitudinal differences in temperature responses (positive in boreal, negative in tropics) and vegetation-type specific sensitivities. While the topic is within the scope of Biogeosciences, I have several major concerns regarding the novelty and methodology of this work.

- Thank you for your evaluations. We made substantial revisions following your comments. We hope this version of paper have answered your concerns.

Numerous previous studies have already investigated GPP responses to temperature, precipitation, and drought across different regions and vegetation types, as well as global GPP responses to specific meteorological factors, i.e., tropical region (Piao et al., 2013, Lomax et al. 2024), boreal forest to drought (Lindroth et al. 2020, Martínez-García et al., 2024). The general relationships between GPP and meteorological variables described in the abstract could essentially be obtained through literature review alone. Therefore, I am not fully convinced by the motivation, novelty, and critical insights in re-examining these well-documented responses.

- Thank you for your insightful comment. We acknowledge that previous studies have extensively investigated the responses of GPP to climatic drivers. However, we provided several new insights as listed below:
  (1) This study systematically explored the responses of GPP to temperature, precipitation, and drought across various vegetation types. Most previous studies have focused on either a single vegetation type (e.g., boreal forests in Lindroth et al., 2020) or a single climatic factor (e.g., precipitation in Lomax et al., 2024).
  (2) We validated the performance of state-of-the-art vegetation models, which provide long-term carbon simulations for the annual Global Carbon Budget

report, against available benchmark datasets. This updated assessment (e.g., compared with Piao et al., 2013) offers an up-to-date understanding of the credibility of the land carbon budget predicted by these models.

(3) Our analyses showed that state-of-the-art models tend to overestimate the positive effects of precipitation on GPP, particularly in tropical regions. Such biases may lead to an overestimation of carbon sink losses during drought years. Furthermore, structural differences among models, such as the inclusion of carbon-nitrogen coupling and the separation of direct and diffuse radiation, can substantially influence the simulated GPP sensitivity to drought, offering critical insights for model improvement.

In this revised paper, we have incorporated site-level validations of benchmark GPP (Fig. 2), analyzed the simulated GPP sensitivity to climatic factors across different model structures (Fig. 9), and substantially expanded the discussion to examine the drivers of GPP responses and the sources of biases in model simulations. Detailed revisions will be shown in the following responses. With these updates, along with the original analyses, we aim to provide a robust foundation for understanding multi-model ensemble predictions, particularly for interpreting long-term trends and interannual fluctuations of terrestrial carbon sinks.

The authors acknowledged the potential nonlinear relationship between GPP and meteorological variables in the introduction, yet the analysis relied entirely on linear regression/correlation methods. This approach is inadequate for capturing non-linear GPP responses to extremes (e.g., drought thresholds, temperature optima).

➢ Thank you for your insightful comment. During the revision process, we attempted to perform quadratic regressions to derive the optimal thresholds as suggested. However, these regressions yielded low R2 values for most plant functional types (not shown), indicating that the nonlinear relationships between GPP and meteorological variables are highly complex and cannot be adequately captured by a simple threshold. Moreover, for some PFTs, such as evergreen trees in boreal regions, environmental temperatures are generally too low to reach their optimal levels, making it difficult to determine such thresholds.

In this revision, we have expanded the discussion to acknowledge the importance of the nonlinear responses of GPP to climatic factors:

"Some PFTs show limited GPP responses to climatic variations (Fig. 5), such as DBF and shrubs to temperature, and ENF and EBF to precipitation. This is likely because their GPP are not primarily constrained by these factors under typical environmental conditions. In relatively stable climates, such as temperate forests or humid tropical regions, temperature and water availability often remain near optimal levels for photosynthesis, so additional warming or increased precipitation brings little further benefit and may even reduce GPP(Reichstein et al., 2006)." (Lines 507-512)

We also explicitly discussed the limitations of ignoring nonlinear responses:

"Finally, the nonlinear effects of climatic factors on GPP responses were not considered in the analyses. We employed the linear regression to estimate GPP sensitivity to the climatic changes. In reality, GPP often exhibits threshold-like responses to climate drivers, such as declining photosynthetic efficiency under heat stress (Doughty et al., 2023) or diminishing GPP with increasing precipitation once water availability is no longer limiting (Li et al., 2022). Ignoring these nonlinear dynamics may lead to underestimation or overestimation of GPP sensitivities, particularly in regions where climatic conditions approach environmental limits for plant growth. Future studies could incorporate nonlinear analyzing approaches, such as quadratic or piecewise regressions, to better capture the full range of GPP responses to climate variability." (Lines 577-586)

Additionally, the observational datasets (GLASS, GOSIF, JUNG) cover different periods (1982 – 2017 vs. 2001 – 2018 vs. 1982 – 2011). The non-overlapping timeframes may introduce biases in trend and sensitivity analyses, particularly given accelerated climate change after the 2000s.

- Thank you for your insightful comment. We have revised the analysis by removing the GOSIF dataset due to its relatively shorter and misaligned temporal coverage. All analyses now rely on the overlapping period (1982 – 2011) between the GLASS and JUNG datasets. The updated text is as follows:

  "In this study, we use the overlapping period of 1982-2011 from both GLASS and JUNG as the reference period." (Lines 127-128)

The 17 models included in this study vary substantially in resolution, carbon-nitrogen coupling, and radiation schemes as illustrated in Table 1. The authors should address *how structural differences contribute to inter-model variability through sensitivity analyses or other methods.*

- Thank you for your valuable suggestion. In the revised paper, we conducted a sensitivity analysis to assess the impact of model structure on the simulated GPP responses to climatic drivers. We added a new section 3.5 and Fig. 9 as follows:

  "3.5 Impact of model structure on drought sensitivity

  The TRENDY models use different structures and parameterizations for carbon cycle simulations (Table 1), leading to varied GPP responses to climatic variables (Fig. 4). To assess the impact of model structure on simulated GPP sensitivity to drought, we grouped the models based on whether they include carbon-nitrogen coupling and/or whether they use separate schemes for diffuse and direct radiation. For most PFTs, incorporating the nitrogen cycle generally increased GPP sensitivity to scPDSI compared to models without carbon-nitrogen coupling, leading to a further overestimation of drought responses relative to the benchmark datasets (Fig. 9a). Meanwhile, models that distinguish between direct and diffuse radiation showed higher GPP sensitivity to scPDSI than those without such a radiation scheme on the global scale (Fig. 9b). This feature is most pronounced for non-tree PFTs (e.g., C3G, C4G, and crop). In contrast to the effect of carbon-nitrogen coupling, which consistently magnified the simulated drought responses, the two-leaf radiation scheme reduced GPP sensitivity to scPDSI for EBF and shrubs, thereby improving the simulated drought responses for these two PFTs." (Lines 449-462)

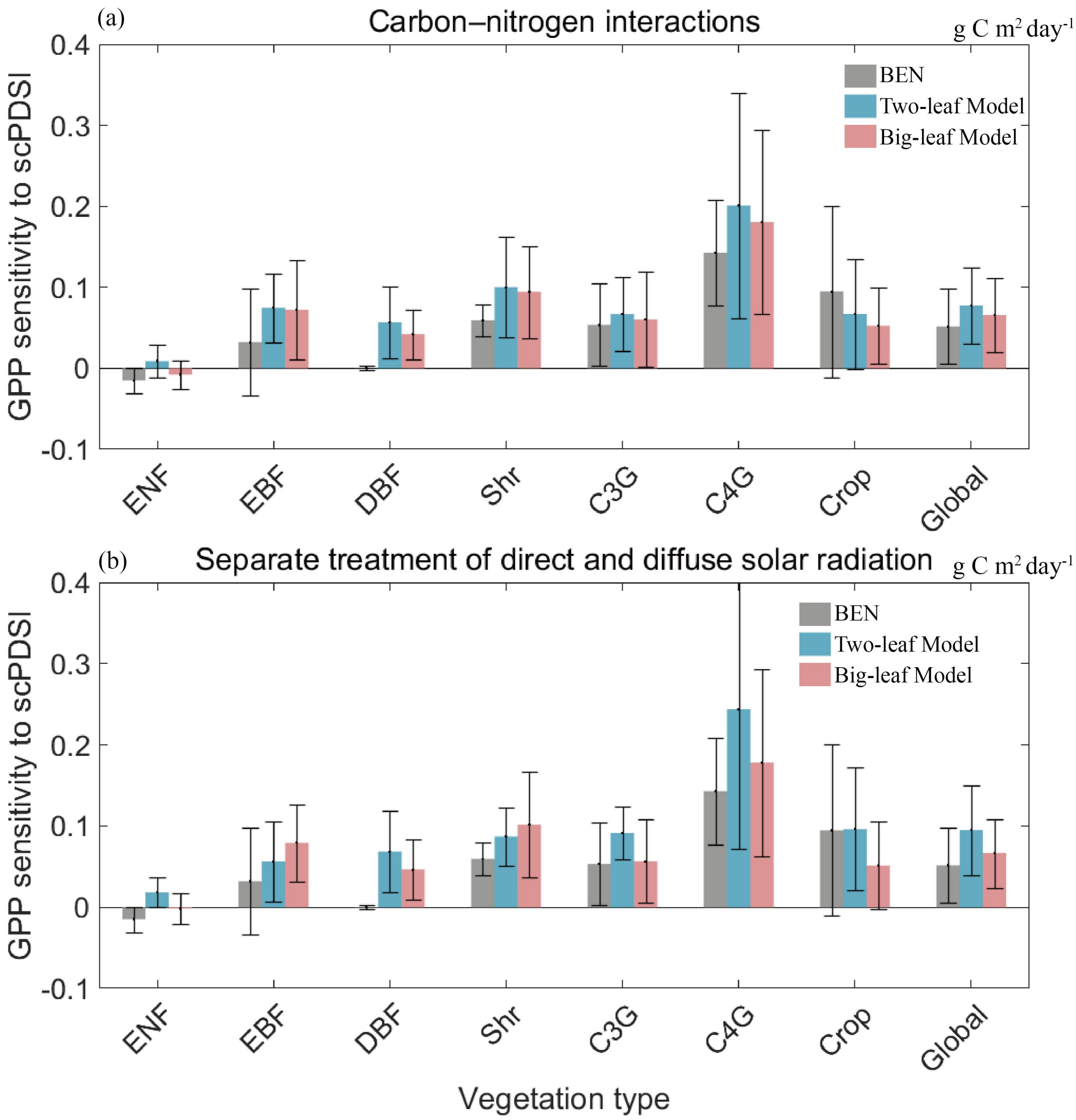

Fig. 9. Comparison of GPP sensitivity to scPDSI across model configurations. Model groups are categorized based on the presence or absence of carbon-nitrogen (C-N) coupling and the separate treatment of diffuse and direct radiation (two-leaf). For each structural feature, ensemble means are calculated from benchmark datasets (grey; including GLASS and JUNG), models with the feature (blue), and models without the feature (pink). The errorbars indicate one standard deviation across benchmark datasets or model simulations.

We also expanded discussion for the possible causes and implications:

"Our analyses showed that differences in model structures partly contribute to the variations in simulated GPP sensitivities (Fig. 9). For instance, the representations of carbon-nitrogen (C-N) coupling can substantially influence the simulated response to water stress (Luo et al., 2008; Liu et al., 2025). By emphasizing nitrogen limitation on photosynthesis, C-N coupled models may amplify the interactions between water availability and carbon cycling in nitrogen-limited ecosystems such as ENF, potentially leading to deviations from observed sensitivity patterns (Yang et al., 2025). Furthermore, differences in canopy radiative scheme affect simulated GPP responses to drought. In ecosystems with complex canopy structures, such as EBF and shrubs, models incorporating a two-leaf radiation transfer scheme show better agreement with observed GPP sensitivity to scPDSI. This likely reflects a more realistic representation of light distribution and leaf functional traits in multi-layered canopies, enhancing diffuse radiation use efficiency and mitigating the limiting effect of water stress on photosynthesis (De Pury and Farquhar, 2008). In contrast, in grassland ecosystems with simplified canopy structure (both C_3 and C_4), models that separate direct and diffuse radiation tend to overestimate GPP sensitivity to scPDSI, likely due to overestimated diffuse fertilization effects (Kanniah et al., 2013)." (Lines 544-558)

Specific comments

Please clarify all the units, Pg C or Pg CO2, throughout the manuscript.

- We clarified that the units of "Pg C" was used throughout the manuscript. All relevant figures, tables, and text have been updated accordingly.

Only the GLASS GPP product is shown in Figure 1a, while the other two observation datasets are not included. Please explain the rationale for this selective presentation.

- In the revised version, we have focused on GLASS and JUNG datasets as the observational benchmarks. The time series for JUNG GPP has now been added to Fig. 1a to ensure a complete and accurate representation of the data used in our analysis.

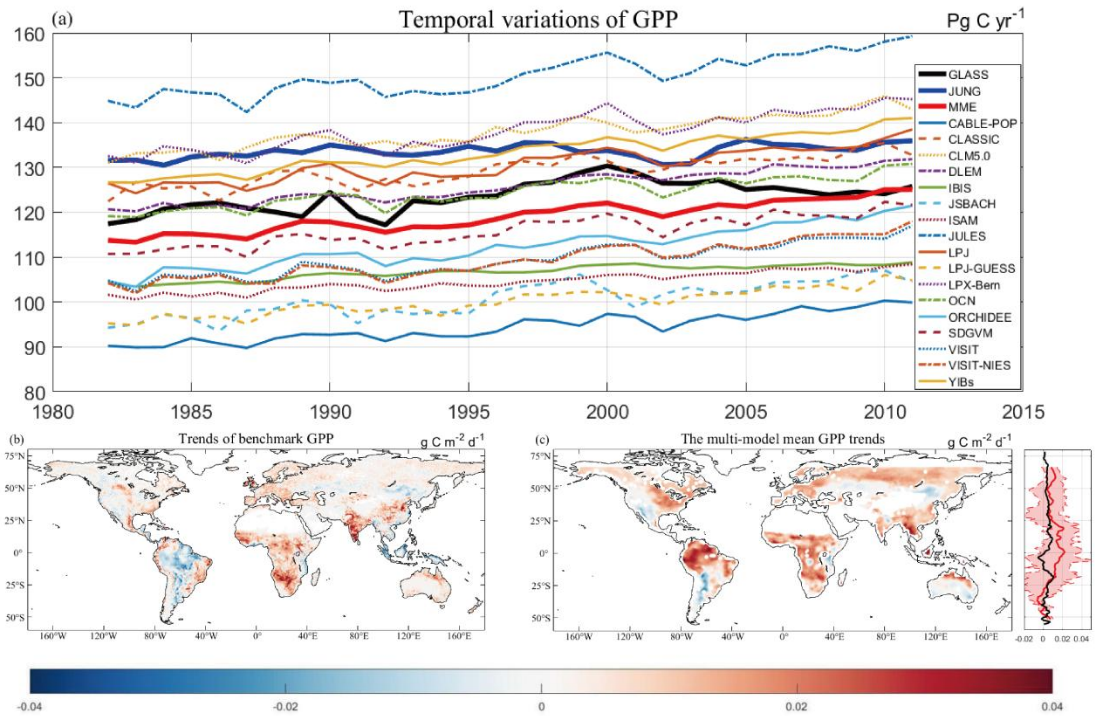

Fig. 1. Comparison of spatiotemporal variations in Gross Primary Productivity (GPP) between benchmark and model simulations. The (a) temporal variations in simulated GPP from individual models and the multi-model ensemble (MME) mean (thick red line) are shown alongside benchmark data from GLASS (thick black line) and JUNG (thick blue line). Spatial patterns of GPP trends from (b) benchmark, represented as the mean of two products (GLASS and JUNG), are compared with (c) the MME of the simulations. Latitudinal variations in GPP trends are also shown, with benchmark data represented in black and model simulations in red; shading indicates one standard deviation.

Line 380-383: The divergence between observations and models is already evident in the GPP trends shown in Figure 1. This discrepancy should be explained before analyzing the GPP responses to meteorological factors.

➢ Section 4.1 focuses on the drivers of GPP sensitivity to climatic factors based on benchmark products, while Section 4.2 addresses the performance of model simulations. Accordingly, we explain the divergence between benchmark data and model results at the beginning of Section 4.2 as follows:

“For this study, we used two benchmark datasets to assess the simulated GPP responses to climatic variables from 17 state-of-the-art vegetation models. The model in general captured the increasing trends in benchmark GPP but with large inter-model variability (Fig. 1). Meanwhile, the MME failed to reproduce the decline in GPP over in Amazon. These biases likely originate from the inadequate representation of key processes in the models, such as moisture stress, phenological responses, and anthropogenic disturbances (e.g., deforestation and forest degradation) in tropical ecosystems (Gu et al., 2002; Koch et al., 2021). Many models also overestimate the benefits of CO_2 fertilization under moisture-constrained conditions, while failing to adequately represent the constraints imposed by combined heat and drought stress on photosynthetic activity, such as photoinhibition and stomatal closure (Green et al., 2020). Furthermore, insufficient representation of thermal adaptation and resilience mechanisms in tropical vegetation may lead to an overestimation of carbon sink capacity under prolonged climatic stress (Doughty et al., 2023)." (Lines 521-532)

Line 385-388: references are needed for this. Is this from your analysis or previous studies? Please clarify and provide appropriate citations.

- This statement is based on our analyses, particularly the results shown in Fig. 5. In the revised manuscript, we have clarified it as follows:

  "However, our analyses showed that models overestimated the impact of precipitation on GPP, particularly in vegetation types such as C3G, DBF, ENF, and EBF (Fig. 5)." (Lines 534-536)

References

De Pury, D. G. G. and Farquhar, G. D.: Simple scaling of photosynthesis from leaves to canopies without the errors of big-leaf models, Plant, Cell & Environment, 20, 537-557, https://doi.org/10.1111/j.1365-3040.1997.00094.x, 2008.

Doughty, C. E., Keany, J. M., Wiebe, B. C., Rey-Sanchez, C., Carter, K. R., Middleby, K. B., Cheesman, A. W., Goulden, M. L., da Rocha, H. R., Miller, S. D., Malhi, Y., Fauset, S., Gloor, E., Slot, M., Oliveras Menor, I., Crous, K. Y., Goldsmith, G. R., and Fisher, J. B.: Tropical forests are approaching critical temperature thresholds, Nature, 621, 105-111, https://doi.org/10.1038/s41586-023-06391-z, 2023.

Green, J. K., Berry, J., Ciais, P., Zhang, Y., and Gentine, P.: Amazon rainforest photosynthesis increases in response to atmospheric dryness, Sci Adv, 6, https://doi.org/10.1126/sciadv.abb7232, 2020.

Gu, L., Baldocchi, D., Verma, S. B., Black, T. A., Vesala, T., Falge, E. M., and Dowty, P. R.: Advantages of diffuse radiation for terrestrial ecosystem productivity, Journal of Geophysical Research: Atmospheres, 107, https://doi.org/10.1029/2001jd001242, 2002.

Kanniah, K. D., Beringer, J., and Hutley, L.: Exploring the link between clouds, radiation, and canopy productivity of tropical savannas, Agricultural and Forest Meteorology, 182-183, 304-313, https://doi.org/10.1016/j.agrformet.2013.06.010, 2013.

Koch, A., Hubau, W., and Lewis, S. L.: Earth System Models Are Not Capturing Present-Day Tropical Forest Carbon Dynamics, Earth's Future, 9, https://doi.org/10.1029/2020ef001874, 2021.

Li, W., Migliavacca, M., Forkel, M., Denissen, J. M. C., Reichstein, M., Yang, H., Duveiller, G., Weber, U., and Orth, R.: Widespread increasing vegetation sensitivity to soil moisture, Nature Communications, 13, https://doi.org/10.1038/s41467-022-31667-9, 2022.

Liu, H., Gao, X., Fan, W., and Fu, X.: Optimizing carbon and nitrogen metabolism in plants: From fundamental principles to practical applications, J Integr Plant Biol, 67, 1447-1466, https://doi.org/10.1111/jipb.13919, 2025.

Luo, Y., Gerten, D., Le Maire, G., Parton, W. J., Weng, E., Zhou, X., Keough, C., Beier, C., Ciais, P., Cramer, W., Dukes, J. S., Emmett, B., Hanson, P. J., Knapp, A., Linder, S., Nepstad, D. A. N., and Rustad, L.: Modeled interactive effects of precipitation, temperature, and [CO2] on ecosystem carbon and water dynamics in different climatic zones, Global Change Biology, 14, 1986-1999, https://doi.org/10.1111/j.1365-2486.2008.01629.x, 2008.

Reichstein, M., Ciais, P., Papale, D., Valentini, R., Running, S., Viovy, N., Cramer, W., Granier, A., OgÉE, J., Allard, V., Aubinet, M., Bernhofer, C., Buchmann, N., Carrara, A., GrÜNwald, T., Heimann, M., Heinesch, B., Knohl, A., Kutsch, W., Loustau, D., Manca, G., Matteucci, G., Miglietta, F., Ourcival, J. M., Pilegaard, K., Pumpanen, J., Rambal, S., Schaphoff, S., Seufert, G., Soussana, J. F., Sanz, M. J., Vesala, T., and Zhao, M.: Reduction of ecosystem productivity and respiration during the European summer 2003 climate anomaly: a joint flux tower, remote sensing and modelling analysis, Global Change Biology, 13, 634-651, https://doi.org/10.1111/j.1365-2486.2006.01224.x, 2006.

Yang, H., Zhang, P., Wang, G., Wang, Q., Wang, D., Wang, R., Zhang, X., and Yin, H.: The synergistic strategy of leaf nitrogen conservation and root nitrogen acquisition in an alpine coniferous forest along an elevation gradient, Plant and Soil, https://doi.org/10.1007/s11104-025-07514-3, 2025.