We thank the reviewers for their suggestions to improve our manuscript. Our response to specific point are given below *in italics*, with revised text in blue.

Review 1

The statement, “Thus, our estimates of the effect of different drivers to changes in GPP during the MH are not expected to be influenced by structural biases or sensitivity to parameters.” may not be entirely accurate, as the authors themselves acknowledge the use of an empirical soil water stress function to scale GPP (Lines 173–175). This indicates that model sensitivity to parameterization may still influence the results.
We do use an empirical function for soil water stress in applying the P model. Stocker et al. (2020) tested this function by comparing the performance of the P model before and after its application against eddy-covariance flux measurement sites under different levels of aridity and shown using this function improves model performance. The model with this empirical function provides a good prediction of the observations on multiple time scales. The impact of water stress on photosynthesis will not be affected by running the model under MH conditions. CO_2 *levels which would affect photosynthesis through water use efficiency are explicitly and mechanistically accounted for in the P model. Overall, the EEO framework only uses a very small number of empirical parameters, and these are well quantified from global observations. Stocker et al. (2020) also showed that the sensitivity to these parameters was small. Therefore, our statement that the effect of different drivers on GPP will not be influenced by structural biases or sensitivity to empirical parameters seems justifiable.*
In response to comments by the second reviewer, however, we have substantially expanded the description of the EEO models in general, and the application of the soil-moisture stress function in particular, to make it clearer how these are applied in the current context (please see revised text in response to the second reviewer’s comment below).

In their response, the authors emphasize that the novelty of this work lies in applying a parameter-sparse model (EEO) to investigate the relatively understudied mid-Holocene (MH) period. While this approach may offer advantages, the manuscript would benefit from a clearer explanation of how the EEO model has been validated. Specifically, the authors should provide evidence of model validation using independent proxies such as pollen reconstructions or other observational datasets. Since the conclusions draw upon the effects of CO_2 concentrations, precipitation, radiation, and C3/C4 vegetation fractions, the validation of the model's ability to simulate these drivers should be explicitly addressed. This could be presented as a table summarizing existing validations or as a standalone section detailing independent assessments.
We provided information about the validation of each component of the model in the revised text, and have further expanded on this in the revised of Section 2.1 required by the second reviewer. This included validation of the impact of CO_2 *under both enhanced and lower* CO_2 *levels. The performance of the models with respect to variations in climate parameters was also discussed in that Section. Since we have cited the appropriate evaluations in the text, the value of adding a table is unclear – although we could include the table below if this would make things clearer.*
Our ability to validate the mid-Holocene and LGM simulations is limited. There is no global data source that provides information on GPP or C_3/C_4 *vegetation fraction. As we stated in the Discussion, pollen data do not provide a way to separate* C_3 *and* C_4 *grasses. It could be used to estimate tree cover fraction, although this involves models such as REVEALS with their own methodological uncertainties and therefore would not provide a strong validation*

of our results. There are some local or regional studies which provide compound-specific $\delta^{13}C$ analyses of leaf wax biomarkers. As we state in the Discussion, these regional studies are consistent with our simulations of C_3/C_4, both in regions where C_4 was more abundant at the LGM and the more limited regions where it was less abundant - this provides strong support for our results. We have revised the text in the Discussion to make the palaeodata limitations for global validation clearer and to point out the implications of the consistency with the limited amount of regional data that exists as follows:

The modelled abundance of C_4 plants was nearly double at the LGM compared to the pre-industrial era (40% versus 23% of the vegetation fraction) and C_4 vegetation was responsible for 56% of the total modelled GPP at that time. These changes are broadly consistent with pollen records, indicating a substantial reduction in tree taxa at the LGM (Prentice et al., 2000; Williams, 2003: Pickett et al., 2004; Cordova et al., 2009; Marchant et al., 2009; Williams et al., 2011). It is difficult to estimate the magnitude of this reduction because existing regional reconstructions have not been applied to the LGM (e.g. Zanon et al., 2018; Serge et al., 2023) and furthermore employ techniques that are based on modern calibrations and therefore do not account for the impact of CO_2 on tree cover (Prentice et al., 2022). However, while pollen data can be used discriminate between trees (virtually all C_3) and grasses, it cannot be used to infer changes in the importance of C_3 and C_4 grasses. Compound-specific $\delta^{13}C$ analyses of leaf wax biomarkers provide evidence of the relative contribution of C_3 and C_4 plants (Eglinton & Eglinton, 2008; Diefendorf et al., 2010) and have shown that C_4 plants were more abundant at the LGM than during the Holocene in many regions (e.g. in southern Africa: Rommerskirchen et al., 2006; Vogts et al., 2012; eastern Africa: Sinninghe Damsté et al., 2011; Himalayan Basin: Galy et al., 2008; southern China: Jiang et al., 2019; south-western North America: Cotton et al., 2016; northern South America: Makou et al., 2007), consistent with our simulations. There are a few regions where C_4 plants were less abundant at the LGM than during the Holocene, including the Chinese Loess Plateau and the Great Plains of North America (Cotton et al., 2016). Both of these regions are identified as characterised by reduced C_4 abundance in our simulations. The consistency of the signs of the regional changes in the observed relative abundance of C_3 to C_4 plants to our simulated changes provides strong support for the model predictions.

Additional references

Cordova, C.E**.,** Harrison, S.P., Mudie, P.J., Riehl, S, Leroy, S.A.G., and Ortiz, N.: Pollen, plant macrofossil and charcoal records for palaeovegetation reconstruction in the Mediterranean-Black Sea Corridor since the Last Glacial Maximum, Quat. Int. 197, 12-26, 2009.

Prentice, I.C., Villegas-Diaz, R., and Harrison, S.P.: Accounting for atmospheric carbon dioxide variations in pollen-based reconstruction of past hydroclimates, Glob. Planet. Change, 211,103790, https://doi.org/10.1016/j.gloplacha.2022.103790, 2022.

Serge, M., Mazier, F., Fyfe, R., Gaillard, M.-J., Klein, T., Lagnoux, A., Galop, D., Githumbi, E., Mindrescu, M., Nielsen, A., Trondman, A.-K., Poska, A., Sugita, S., Woodbridge, J., Abel-Schaad, D., Åkesson, C., Alenius, T., Ammann, B., Andersen, S., Anderson, R., Andrič, M., Balakauskas, L., Barnekow, L., Batalova, V., Bergman, J., Birks, H., Björkman, L., Bjune, A., Borisova, O., Broothaerts, N., Carrion, J., Caseldine, C., Christiansen, J., Cui, Q., Currás, A., Czerwiński, S., David, R., Davies, A., De Jong, R., Di Rita, F., Dietre, B., Dörfler, W., Doyen, E., Edwards, K., Ejarque, A., Endtmann, E., Etienne, D., Faure, E., Feeser, I., Feurdean, A., Fischer, E., Fletcher, W., Franco-Múgica, F., Fredh, E., Froyd, C., Garcés-Pastor, S., García-Moreiras, I., Gauthier, E., Gil-Romera, G., González-Sampériz, P., Grant, M., Grindean, R., Haas, J., Hannon, G., Heather, A.-J., Heikkilä, M., Hjelle, K., Jahns, S., Jasiunas, N., Jiménez-Moreno, G., Jouffroy-Bapicot, I., Kabailienė, M., Kamerling, I., Kangur, M., Karpińska-Kołaczek, M., Kasianova, A., Kołaczek, P., Lagerås, P., Latalowa, M., Lechterbeck, J., Leroyer, C., Leydet, M., Lindbladh, M., Lisitsyna, O., López-Sáez, J.-A., Lowe, J., Luelmo-Lautenschlaeger, R., Lukanina, E., Macijauskaitė, L., Magri, D., Marguerie, D., Marquer, L., Martinez- Cortizas, A., Mehl, I., Mesa-Fernández, J., Mighall, T., Miola, A., Miras, Y., Morales-Molino, C., et al.: Testing the effect of relative pollen productivity on the REVEALS model: A validated reconstruction of Europe-wide Holocene vegetation, Land, 12, 986, https://doi.org/10.3390/land12050986, 2023.

Williams, J. W., Tarasov, P., Brewer, S., and Notaro, M.: Late Quaternary variations in tree cover at the northern forest-tundra ecotone, J. Geophys. Res., 116, G01017, doi:10.1029/2010JG001458, 2011.

Zanon, M., Davis, B. A. S., Marquer, L., Brewer, S., and Kaplan, J. O.: European forest cover during the past 12,000 years: A palynological reconstruction based on modern analogs and remote sensing, Front. Plant Sci., 9, 1–25, https://doi.org/10.3389/fpls.2018.00253, 2018.

Additionally, my original concern regarding the large inter-model spread in Last Glacial Maximum (LGM) GPP estimates was not fully addressed. It remains unclear how the newly derived estimate fits within the broader context of previous estimates, and what implications this has for reducing uncertainty or reconciling existing model disagreements. A more in-depth discussion of how the new estimate compares to the existing range—and whether it provides any constraints or insights—would significantly strengthen the manuscript.

We reiterate the point that we made last time about the difficulties of diagnosing the cause of the large range in previous estimates of GPP at the LGM. The forcing data used for these simulations (e.g. ice sheet configuration) varied, the climate models themselves differ structurally and in terms of specific parameter values used, and the vegetation component of these models are also different. Moreover, the vegetation components of the models are very complex being plant-functional type (PFT) based, and this involves specifying many (poorly constrained) parameters for each PFT. Evaluation of the performance of these models tends to focus on outcomes (e.g. vegetation distribution) rather than the individual processes (e.g. photosynthesis, evapotrainspiration) that gave rise to these outcomes. Our EEO-based models are parameter sparse, have been evaluated at process level using a wide range of different types of observation (controlled experiments, FACE experiments, field studies and geographic patterns of key traits in response to specific environmental gradients) and therefore provides robust simulations of key vegetation properties. The EEO-based models have repeatedly been shown to out-perform the models that are routinely used to make projections in simulating the modern GPP. Although we are forced to use a climate model to provide LGM climate inputs, we have chosen the best-performing model as independently evaluated using quantitative climate reconstructions. Thus, we believe that our LGM simulations are likely to be more reliable than the previous estimates, although we have acknowledged in the Discussion that it would have been useful to run simulations with other CMIP6/PMIP4 models – but unfortunately other models did not archive all of the required variables. Since we recognise that we have perhaps been somewhat cautious about claiming our model is better tested and likely to provide more robust estimates of past GPP, we will re-organise and modify the Discussion section on the comparison with previous estimates, as follows:

We have shown that the LGM was characterised by a large reduction in modelled GPP, while the mid-Holocene was characterised by a small increase in GPP compared to the pre-industrial state. Estimated GPP at the LGM was ca 84 PgC yr^{-1} compared to ca 110 PgC yr^{-1} in the PI. The simulated reduction at the LGM is consistent with previous model-based estimates (e.g. François et al., 1998; Prentice et al., 2011; Hoogakker et al., 2016), including those from the latest phase of the Couple Model Intercomparison project (CMIP6/PMIP4: Supplementary Table 3). However, previous estimates of GPP span a considerable range, from 40-110 PgC yr^{-1}). This reflects differences in the boundary conditions used, differences in the vegetation models used and their sensitivity to changes in CO2, and differences in the structure and parameterisations of the climate models overall. Diagnosing the specific causes of this large range is therefore extremely difficult. The parameter sparse nature of our EEO-based modelling approach, and the fact that the individual processes that give rise to the simulated GPP have been independently validated, suggest that our estimate of ca 84 PgC yr^{-1} is more likely to be realistic than previous estimates. A limited number of studies have estimated GPP at the LGM by constraining model estimates using oxygen isotope records from ice core (Landais et al., 2007; Ciais et al., 2011; Yang et al., 2022). The still large range in simulated GPP (40-110 PgC yr^{-1}) reflects, in part, uncertainties associated with estimating ocean productivity and respiration fractionation rates. Thus, although there is a consensus that GPP was considerably lower at the LGM than during pre-industrial times, and this is consistent with pollen evidence for a very large reduction in tree cover over much of the world (Prentice et al., 2000; Williams, 2003; Pickett et al., 2004; Marchant et al., 2009), the absolute magnitude of this change is uncertain. Nevertheless, since the climate simulated by the MPI ESM has been shown to reproduce pollen-based climate reconstructions better than most other CMIP6/PMIP4 models (Kageyama et al., 2021) and we use robust EEO-models to estimate the overall change in GPP, the partitioning of the impacts of different factors in the simulated reduction of GPP is likely to be robust.

We have also added a concluding sentence to the final paragraph of the Discussion to reiterate the advantages of using EEO-based models for the simulation of past vegetation, as follows:

…. more accurately than the land-surface models used to predict the terrestrial carbon cycle (Cai et al., 2025; Zhou et al., 2025). Given their simplicity, the fact that the very few parameters required are well constrained from observations, and the demonstrated quality of their performance, EEO-modelling holds considerable promise for understanding past vegetation changes and their impact on climate.

Review 2

This is a very interesting study. Definitely, eco-evolutionary optimization modeling is an effective approach in estimating global GPP in different geological periods.
We thank the reviewer for their positive evaluation of the worth of this approach

The main issue for me is the lack of clarity regarding how the flowchart in Figure 1 was conducted in the simulations. While I understand the general intent, these steps raise important questions that I could not figure out and also could not find the answers in this manuscript and related references:
In Step 1: The use of the P-model to estimate potential GPP for C_3 and C_4 plants by assuming fAPAR = 1.0 is straightforward and aligns with standard applications of the P-model. This

part is clear and no problem for me.

In Step 2: The integration of global tree cover data and a C_3/C_4 model is where the confusion begins. It is unclear how tree cover data from the three specified periods (LGM, MH, and PI) were obtained, and more importantly, how these data were used in the C_3/C_4 model. Given that most C_4 species are grasses, the relationship between tree cover and the selection of C_3 versus C_4 GPP is not intuitive. Moreover, the referenced preprint describing the C_3/C_4 model lacks sufficient detail to explain it in this context. What mechanism or criteria does the model use to distinguish between C_3 and C_4 dominance? Without this, the reader cannot assess the validity of this step.

In Step 3: The introduction of the LAI model is also vague. No detailed descriptions are provided regarding the model's structure or rationale. The paragraph between lines 155–176 (Page 5) appears to describe this model, but it remains opaque after multiple times of reading. A clear explanation of how LAI and fAPAR are calculated and linked to the earlier steps is needed.

Step 4: The reapplication of the P-model and C_3/C_4 model using the updated fAPAR from Step 3 raises the question of whether more iterations are needed. Additionally, the use of the SPLASH model is introduced somewhat abruptly. It is said to provide a soil moisture correction to GPP (Lines 174–176), but the connection between cloud cover, soil moisture limitation, and GPP is not explained in sufficient detail.

Overall, the simulation system employed in this study appears to draw upon several existing models—P-model, C_3/C_4 model, LAI model, and SPLASH—some of which include optimization principles. However, the manuscript offers only general descriptions of these models and lacks a coherent and specific explanation of how they were implemented and integrated in this research and how the "optimal" was solved. Key model assumptions, parameterizations, and linkages between steps are either not provided or are difficult to follow. As a result, it is unclear how the full system operates or how optimization principles are realized within it.

I recommend that the authors substantially revise the methods section to clearly explain each step in the simulation workflow, including detailed roles of each model, data inputs, and how optimization is achieved. A more explicit and self-contained description would significantly improve the transparency and reproducibility of the study.

We have expanded the model description to give a general overview of the simulation system we are using and the different steps involved, as shown in Figure 1. For each component model, we have provided a fuller description of the model, including the key assumptions and parameterizations. We have also revised the text to make it clear what EEO principles we are invoking in each of the models. We have also made it clear when we are using empirical relationships or empirically derived parameter values. The SPLASH model, which is not an EEO model, was only used to calculate soil moisture based on the climate model inputs, and this is then used to determine when to apply the soil moisture correction that is used in drier regions when running the P model. We have therefore expanded the description of this empirical correction and made it clear what the purpose of SPLASH is. The reapplication of the P-model and C_3/C_4 *model in Step 4 using the updated fAPAR from Step 3 is in order to translate from the leaf level to the canopy level. This is therefore not an iterative process. We hope that the revised text (Section 2.1) explains what we have done more clearly. The revised text is as follows:*

We simulated vegetation changes at the LGM and the MH compared to the pre-industrial (PI) state using a sequence of linked models that predict GPP, leaf area index (LAI) and C_3/C_4 competition based on EEO theory (Fig. 1). We first simulate potential GPP (equivalent to leaf-level photosynthesis) for C_3 and C_4 plants independently. These estimates are used to derive total potential GPP allowing for competition between C_3 and C_4 plants. Potential GPP is converted to actual GPP using a model that simulates the seasonal cycle of leaf area index (LAI), which is converted to the fraction of absorbed photosynthetically active radiation (fAPAR) using using Beer's law. Finally, we use a soil water balance model to calculate soil moisture, then take account of the impact of low soil mosture on GPP using an empirical correction.

The P model (Wang et al., 2017, Stocker et al., 2020) is a light-use efficiency model that simulates GPP. It uses the Farquhar-von Caemmerer-Berry photosynthesis model (Farquhar et al., 1980) for instantaneous biochemical processes combined with two EEO hypotheses describing photosynthetic acclimation, the 'coordination' and 'least- cost' hypotheses (Prentice et al., 2014, Wang et al., 2017), to account for the spatial and temporal acclimation of carboxylation and stomatal conductance to environmental variations at weekly to monthly time scales. Although the P model simulates both C_3 and C_4 photosynthesis, it does not need to make any other distinctions between plant functional types. The required inputs to the model (Fig. 1) are air temperature (°C), vapour pressure deficit (VPD, Pa) derived from relative humidity, air pressure (Pa) (to account for the effect of elevation on photosynthesis, incident photosynthetic photon flux density (PPFD, μmol m^{-2} s^{-1}) estimated from shortwave (solar) radiation, and ambient CO_2 concentration. The P model has been extensively validated and shown to predict the geographic patterns and seasonal cycles of GPP under modern conditions successfully (Wang et al., 2017; Stocker et al., 2020). Furthermore, it correctly predicts related physiological characteristics, including the global pattern of the maximum carboxylation (V_{cmax}) rate in relation to gradients in PPFD, temperature and VPD (Smith et al., 2019), the seasonal variation of V_{cmax} in different biomes (Jiang et al., 2020), its response to atmospheric CO_2 (Smith and Keenan, 2020), and the variation of photosynthetic traits along elevational gradients (Peng et al., 2020). The responses of photosynthetic properties to enhanced CO_2 as simulated by the P model have been validated against both Free Air Carbon dioxide Enrichment (FACE) experiments (Wang et al., 2017) and controlled-environment experiments (Smith and Keenan, 2020). Moreover, the model's implied response of photosynthetic capacity to CO_2 has been validated by measurements on plants experimentally grown at low (160 ppm) CO_2 (Harrison et al., 2021).

The P model first simulates potential GPP for C_3 and C_4 plants separately, without consideration of competition between them (Figure 1). These estimates are fed into a simple model of C_3/C_4 competition based on the P model (Lavergne et al., 2024). The principle of the C_3/C_4 model is as follows. C_4 plants (mainly grasses, some shrubs) have a higher rate of photosynthesis under hot and/or dry conditions, especially under low CO_2, which reduces C_3 photosynthesis. On the other hand, C_4 plants can only become dominant or co-dominant if tree cover is limited, because (C_3) trees have the advantage in competition for light: C_3 trees can outcompete C_4 grasses through shading even where the C_4 pathway would yield higher rates of photosynthesis. The relative photosynthetic advantage of C_4 plants is estimated in the model as the difference between the monthly potential GPP for C_3 and C_4 plants, summed over the year. The C_4 share of total GPP is then estimated based on a logistic relationship between model-estimated C_4 relative advantage and observed C_4 abundance. An additional function relates the proportion of GPP from trees to total potential GPP based on the relationship between annual mean percentage tree cover and the simulated annual GPP of C_3 plants. Thus, tree cover is an additional required input to the competition model (Figure 1). The competition model also enforces a minimum temperature threshold to define conditions under which C_4 plants cannot grow, where this limit is set to a minimum temperature of the coldest month of –

24° based on experimental data. The competition model has been shown to reproduce global patterns in the relative abundance of C_3/C_4 plants as well as the observed rate of $\Delta^{13}C$ in recent decades, as shown by independent atmospheric estimates (Lavergne et al., 2020).

To convert potential GPP to actual GPP, we use an LAI model (Figure 1) that predicts the seasonal cycle of LAI based on environmental conditions and the P-model estimates of potential GPP, i.e. the GPP predicted when fAPAR is set to 1 (Zhou et al., 2025). This model is based on the EEO hypothesis that seasonal variations in LAI are coordinated with variations in potental GPP because leaves are displayed at (or close to) the time when they are able to be most productive. A seasonal maximum LAI model was embedded in this model to provide an upper limit to the seasonal LAI predictions (Zhu et al., 2022; Cai et al., 2025). The calculation of seasonal maximum LAI incorporates an EEO-based water-carbon trade-off and is defined as the lesser of an energy-limited and a water-limited estimate (Zhu et al., 2022; Cai et al., 2025). The model assumes that, under energy limitation, plants allocate carbon to leaves to maximize GPP after accounting for the costs of leaf construction and maintenance, including the costs of obtaining water and nutrients. This leads to a clear optimum because investing in leaf tissue yields diminishing returns due to mutual leaf shading. Under water limitation, the model assumes that plants adjust their rooting behaviour to extract a portion of annual precipitation from the soil, irrespective of its distribution throughout the year, and allocate carbon to leaves so that all this water is transpired, thereby maximizing GPP. There are inherent delays between the steady-state LAI and the real-time dynamic LAI because photosynthetic and phenological processes do not respond instantaneously to weather fluctuations: the allocation of photosynthate to leaves can take days to months. Thus, the seasonal dynamics of LAI were calculated using a moving average to represent the time lag between allocation to leaves and modelled steady-state LAI (Zhou et al., 2025). The seasonal LAI model has been shown to capture observed LAI dynamics across all biomes at different temporal scales (weekly, seasonal, annual and interannual variability) both at individual eddy-covariance flux measurement sites and when compared to satellite-derived LAI (Zhou et al., 2025). Furthermore, it predicts both the multi-year average LAI and the annual trends in LAI better than any of the biosphere models used in the Trends and Drivers of Terrestrial Sources and Sinks of Carbon Dioxide (TRENDY) project (Zhou et al., 2025). The seasonal cycle of fAPAR is calculated from the seasonal cycle of LAI using Beer's law (Swinehart, 1962) and this is then used to calculate seasonal changes in actual GPP using the P model.

Finally, we apply an empirical soil moisture correction ($\beta(\theta)$: Stocker et al., 2020) to account for the additional impact of soil moisture stress on GPP. This has the form of a quadratic expression applied when soil moisture is below a given threshold value, where the sensitivity of this relationship varies depending on aridity such that the decline in $\beta(\theta)$ with drying soils is steep in dry climates and less steep in wetter climates. The soil moisture stress function was developed by comparing simulations of GPP with flux-tower data at sites across a range of climatic aridity (Stocker et al., 2020). Soil moisture was calculated using the Simple Process-Led Algorithms for Simulating Habitats (SPLASH) model (Davis et al., 2017), which is a generic soil water accounting model in which daily losses depend on potential evaporation, reduced in proportion to relative soil water content.

We realised that we never actually specified that the tree cover information came from the MPI simulations, and have now added this information in section 2.2 – which we have renamed accordingly:

2.2. Derivation of LGM, MH and PI climate and tree cover inputs

The revised text to specify tree cover is:

…… and uniquely has archived all the necessary climate and vegetation outputs needed to run the EEO-based models

The climate and tree cover outputs necessary to run the EEO-based models …….