Dear Reviewer 2 and Topical Editor:

On behalf of my co-authors, we thank you very much for reviewing our manuscript and giving us rigorous, thought-provoking, and highly constructive evaluation. We deeply appreciate your feedback on our manuscript entitled "*Sensitivity and Uncertainty Analysis of China's Terrestrial Carbon-Water Cycle Using a Dynamic Global Vegetation Model*" (egusphere-2025-6076).

This document contains our detailed, point-by-point responses to all the comments raised by **Reviewer 2**. In these responses, we have addressed your concerns and clearly outlined the comprehensive modifications we plan to implement in the formal revised manuscript.

We firmly believe that your insights will greatly improve the biogeoscientific relevance and overall quality of our study. Looking forward to hearing from you.

Best regards,

Fulai Feng

School of Geography and Tourism, Shaanxi Normal University

No. 620, West Chang'an Street, Chang'an District, Xi'an 710119, China

Response to the Reviewers #2

RC2: Comment on egusphere-2025-6076

Fulai Feng and coauthors present a manuscript that assesses the sensitivity of selected model outputs from the Dynamic Global Vegetation Model LPJ-GUESS (version 4.1) to a set of 39 parameters for ecosystems in China. To do so, the authors simulate vegetation dynamics for 13 sites in China (selected to cover main ecosystem types in China) and assess the response of carbon pools, ecosystem properties and carbon and water fluxes to variations in these parameters.

The manuscript has generally a very technical approach to the modelling and sensitivity analysis, which results in an assessment that is very specific for simulations of the 13 specific sites made with this specific model. The wide range of LPJ-GUESS parameters that are addressed here is not further explained with physical meanings or equations. In addition, the sensitivity indicators used in the study are highly dependent on the allowed range of parameter variations, which makes the outcomes rather subjective. A more carefully constrained set of parameter ranges would help to provide more ecological meaning to the outcomes.

The strong focus on the technical aspects of the model and limited focus on the (biogeo-)scientific aspects makes the manuscript relevant for a small group of users and developers of the LPJ-GUESS model that are aware of the model structure and meaning of the parameters, but it will be of limited value for the broader community. To turn this into a contribution that would be of relevance to the broader community, the study should provide more detail on the actual model, and should attempt to discuss the implications of these findings beyond the applied model and study area.

In its current form, I therefore do not consider the manuscript suitable for Biogeosciences. I provide my main concerns below and hope that these can help the authors when revising the manuscript.

Response: We sincerely thank the reviewer for their rigorous, thought-provoking, and highly constructive evaluation of our manuscript. We deeply appreciate your candid feedback regarding the manuscript's overly technical focus and the need to broaden its biogeoscientific relevance. We completely agree with your assessment: a pure technical diagnostic of 39 parameters across 13 sites is of limited value to the broader Biogeosciences community unless these technical findings are explicitly translated into fundamental ecological mechanisms and broader modeling implications.

In response to your critical feedback, we have planned a major revision of the manuscript to fundamentally shift the narrative from "model-specific technicalities" to "general biogeochemical insights." Specifically, we will:

Bridge the mechanistic gap by significantly expanding the physical, mathematical, and structural explanations of all parameters. We will achieve this through a comprehensively overhauled, module-based parameter table that clearly links each parameter to its core biogeochemical equation and model function.

Clarify the exploratory nature of our parameter ranges, addressing the critical issue of ecological plausibility versus technical bounds (e.g., for physically constrained parameters like *FRADPAR*).

Quantify the ecological context by formally defining "resource-rich" and "resource-limited" ecosystems using fundamental climatological drivers (e.g., Mean Annual Precipitation, MAP). We will introduce new quantitative analyses to visually demonstrate how parameter dominance mechanistically shifts along these hydro-thermal gradients.

Elevate the discussion to emphasize how our findings—such as the hidden dominance of hard-coded physiological traits and the critical trade-off between relative and absolute uncertainty—provide universal lessons for the broader Dynamic Global Vegetation Model (DGVM) community (e.g., models participating in TRENDY or CMIP6).

We firmly believe that addressing your profound concerns will transform this manuscript into a much more valuable and widely applicable contribution. Our detailed, point-by-point responses to your specific concerns are provided below.

[Major Comment 1] In order to understand the meaning of model parameters, basic information on their meaning (possibly in relation to the equations in which they are used) should be provided, and information on their units should be given. The current descriptions in Table 1 are too concise for a reader without prior in-depth knowledge of LPJ-GUESS to understand how these parameters are used in the model (e.g., descriptions like "colimitation (shape) parameter", "Constant in tree allometric growth equation" are not very informative).

Response: We fully agree with this criticism. The original Table 1 was indeed too concise and failed to provide the necessary physical, mathematical, and structural context for readers outside the LPJ-GUESS developer community. To resolve this, we will undertake a comprehensive overhaul of Table 1 in the revised manuscript to bridge the gap between technical parameter names and actual biogeochemical processes.

Specifically, we will expand Table 1 by adding the following detailed columns:

Units: Explicitly stating the physical dimension of each parameter.

Associated Module: Categorizing each parameter into its core functional module within the LPJ-GUESS architecture (e.g., Photosynthesis, Carbon Allocation, Allometry, Soil Biogeochemistry, Fire Dynamics). This will immediately clarify where the parameter impacts the model's logic.

Physical Meaning & Key Equations: Providing a clearer explanation of how the parameter functions mathematically. For example, instead of merely stating "colimitation (shape) parameter" for *THETA*, we will categorize it under the "Photosynthesis Module" and describe its role in the Farquhar photosynthesis equation, explaining how it governs the transition between light-limited and Rubisco-limited assimilation rates.

We believe this expanded, module-based parameter table will provide the broader ecological community with a highly transparent understanding of our sensitivity analysis inputs.

[Major Comment 2] The outcomes of the GSA... is highly dependent on the range of parameter values used... As a result of this, the conclusions are sometimes rather arbitrary. E.g., the study attributes major sensitivity to FRADPAR (the fraction of SW radiation that is in the photosynthetically active part of the spectrum, L317), which is a physical property of the light that can vary across sky conditions but is overall fairly well constrained. Fractions are allowed to range within their technically permissible bounds (L177) while these may not be their reasonable range...

Response: This is an exceptionally sharp and valid criticism, and it echoes a similar concern raised by Reviewer 1. While applying standardized relative perturbations (e.g., ±25%) is a widely established exploratory precedent for initial sensitivity screening in complex vegetation and crop models (e.g., Li et al., 2022; Miller, 1974), we completely agree that for physically well-constrained constants like *FRADPAR*, this artificially inflates their apparent sensitivity and compromises the ecological interpretation. In the revised manuscript, we will address this in two concrete ways:

Reframing the *FRADPAR* result: We will explicitly acknowledge that the high sensitivity of *FRADPAR* is an artifact of its artificially wide sampling range. However, we will pivot this into a meaningful insight: the fact that the model responds so violently to changes in *FRADPAR* mathematically highlights the ecosystem's extreme vulnerability to light availability, underscoring the critical necessity of using highly accurate radiation forcing data in dryland modeling.

Addressing the General Limitation: We will add a dedicated subsection in the Discussion (Section 4.3) acknowledging that our exploratory technical boundaries do not equate to ecological boundaries, and subsequent regional calibration must strictly constrain these parameters using empirical trait databases (e.g., TRY database).

References:

Li, Y., Wang, Y., Sun, Y., and Li, J.: Global sensitivity analysis of the LPJ model for Larix olgensis Henry forests NPP in Jilin Province, China, Forests, 13, 874, https://doi.org/10.3390/f13060874, 2022.

Miller, D. R.: Sensitivity analysis and validation of simulation models, J. Theor. Biol., 48, 345–360, 1974.

[Major Comment 3] The number of sites applied to address assess the parameter sensitivity is overall very low. This makes that the assessment concludes that sensitivities vary between sites, but it does not provide useful insights in similarities and differences between the sites. The authors have an interesting and potentially promising approach by distinguishing "resource-rich" and "resource-limited" ecosystems, but the authors would need to define these and find a way to quantify how this can be used as prior information about sensitivity.

Response: We appreciate your constructive suggestion. We acknowledge that 13 sites are numerically few for a country as vast as China. However, these sites were not chosen randomly. As detailed in Text S1 of our original Supplement ("Detailed Methodology for Study Site Selection"), these sites were rigorously identified through a multi-step screening of 23-year (2001-2023) MODIS land cover dynamics, strictly confined within national nature reserves to eliminate anthropogenic disturbances, and finalized via a spatial equilibrium algorithm. Thus, they serve as highly representative "archetypes" capturing the extreme gradients of China's climatic and ecological space.

To make our spatial conclusions more robust and address your excellent point about "resource-rich vs. resource-limited" ecosystems, we have quantitatively defined these classifications and introduced a new mechanistic analysis:

Defining the Environmental Gradient: We utilized Mean Annual Precipitation (MAP) to formally locate each of the 13 sites along a quantitative hydro-climatic gradient.

Quantitative Spatial Analysis (New Figure): We have generated a new analysis (presented below for this discussion, and to be formally added as Figure 9 in the revised manuscript) plotting the Total Sensitivity Indices (S_{T_i}) of key representative parameters directly against the MAP gradient. This figure mathematically demonstrates how the dominant model control mechanism smoothly transitions from "photosynthetic capacity" (e.g., *ALPHA_C3*, which dominates in humid, resource-rich regimes) to "carbon allocation" (e.g., *common_reprfrac*, which dominates in arid, resource-limited regimes).

By anchoring our spatial narrative to quantifiable climatic gradients rather than mere geographical coordinates, we believe the insights drawn from these 13 representative sites now provide highly generalized and scientifically predictive value for ecosystems beyond the specific study areas.

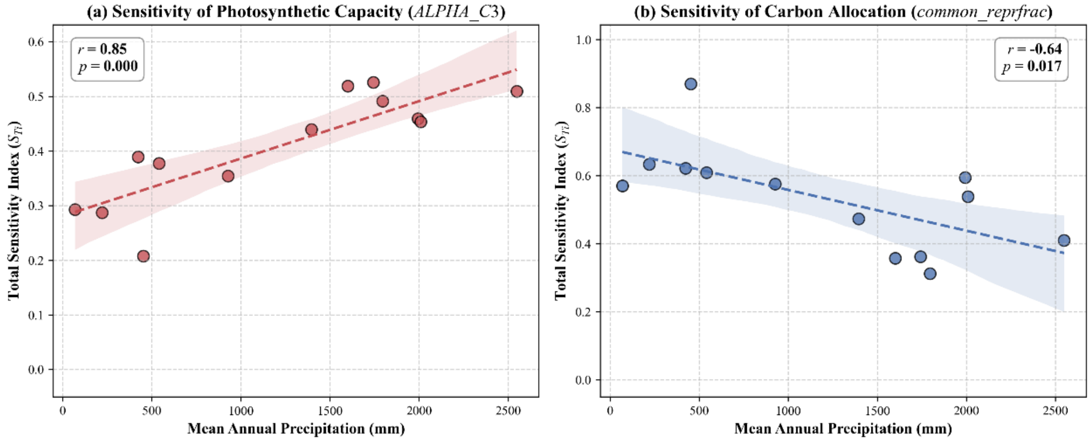

Fig 9. Mechanistic shift in parameter dominance along the precipitation gradient. Scatter plots showing the Total Sensitivity Index (S_{T_i}) of GPP to (a) photosynthetic capacity (*ALPHA_C3*) and (b) carbon allocation (*common_reprfrac*) across the 13 archetypal sites, plotted against Mean Annual Precipitation (MAP). The shaded areas represent the 95% confidence intervals of the linear regression fits. The strong positive ($r = 0.85$, $p < 0.001$) and negative ($r = -0.64$, $p = 0.017$) correlations statistically demonstrate the model's transition from a photosynthesis-limited regime in resource-rich (humid) ecosystems to an allocation-limited regime in resource-limited (arid) ecosystems.

[Major Comment 4] For the PFT-specific parameters (second half of Table 1), it is unclear how the PFT aspect is taken into account here. Are the provided ranges generic across PFTs? Why would a parameter that is PFT-specific be allowed to vary across the same range for all PFTs? Some of the parameters appear to be specific for a subset of PFTs, whereas others exist for all PFTs. Moreover, with the limited number of sites, some of the PFTs will be represented very sparsely in the set of outputs, whereas others may play a more important role, which the study should address.

Response: We sincerely thank the reviewer for their careful examination of Table 1 and for raising this important methodological question. We apologize for the lack of clarity in our initial manuscript regarding how PFT-specific ranges were constructed.

To clarify, our sampling ranges were not assigned arbitrarily, nor were they uniformly treated. Instead, the parameter boundaries were strictly guided by the availability of empirical data in the existing literature, which resulted in two distinct treatments for PFT-specific parameters:

Literature-Constrained PFT-Specific Ranges: For traits with well-documented biological variations across plant functional types, we assigned distinct, PFT-specific ranges. These bounds are explicitly supported by the respective citations provided in the final column of Table 1.

Generic Ranges for Uncertain PFT Traits: For certain parameters—particularly those inherited from earlier iterations of the model (e.g., the original LPJ-DGVM) where high-resolution, PFT-differentiated trait bounds are still absent in the literature—we conservatively assigned a uniform plausible range across all relevant PFTs. It is important to note that while the boundaries of uncertainty are assumed to be consistent across PFTs due to the lack of differentiated empirical evidence, the sampling engine still allows the parameters to take on different actual values for different PFTs during the simulations. This is a standard and scientifically justifiable approach in global sensitivity analyses when specific prior trait distributions are missing. In the revised Section 2.3, we will explicitly clarify this literature-driven, hybrid boundary-setting strategy to prevent any future misunderstanding.

Regarding your excellent point on the sparse representation of certain PFTs across the 13 sites: this is an acute observation. In an aggregated site-level analysis, the sensitivity signal of a parameter specific to a dominant PFT (e.g., evergreen broadleaf trees in southern sites) will naturally drown out the signal of a sparse PFT. In the revised manuscript, we will add a paragraph in the Discussion explicitly acknowledging this limitation. We will clarify that our aggregated sensitivity indices inherently reflect the "dominant vegetation dynamics" of the chosen archetypal sites. Consequently, evaluating the parametric sensitivity of rare or sparsely distributed PFTs falls outside the scope of this macroscopic assessment and would require targeted, single-biome experimental designs in future studies.

[Major Comment 5] There is a clear logic in the modelling of carbon cycle processes that is not used sufficiently in the analysis: Changes to the productivity affect vegetation carbon pools, but changing these will have subsequent impacts further down the line (litterfall, soil C, soil respiration). It would be nice to bring this logic into the analysis...

It would be nice if the overview of the parameters (e.g. in Table 1 or in a figure) could provide more information on the place where they impact processes in the model.

Response: We completely agree. Failing to explicitly map the statistical sensitivities back onto the chronological "cascade" of the ecological carbon cycle was a missed opportunity to provide deeper biogeoscientific insights.

To directly address this and bring the carbon cycle logic to the forefront of our analysis, we will implement the following two changes:

Integrating Process Mapping into Table 1: As suggested by your comment ("e.g. in Table 1"), and detailed in our response to [Major Comment 1], we have opted to comprehensively overhaul Table 1 rather than introducing a redundant standalone figure. By adding the "Associated Module" column to Table 1, we will explicitly categorize each parameter into its chronological functional step (e.g., Photosynthesis, Carbon Allocation, Soil Biogeochemistry). This will provide readers with a clear overview of exactly where each parameter impacts the model's logic.

Rewriting Section 3.4 & Updating Figure 5: We will fundamentally restructure the narrative of Section 3.4 to explicitly describe the "Uncertainty Propagation Cascade" along the carbon cycle. Furthermore, as part of our rigorous statistical overhaul across the manuscript (detailed in our response to Reviewer 1), the completely revised Figure 5 now explicitly displays the true, unstandardized Total Sensitivity Index (S_{T_i}). This allows us to trace the actual absolute fraction of variance as it logically propagates down the carbon cascade—demonstrating how leaf-level quantum efficiency (ALPHA_C3) dictates the initial carbon influx (GPP), how allocation strategies (common_reprfrac) then partition this into vegetation pools (Veg C), and finally how turnover rates (tree_turnover_root) propagate this uncertainty into litterfall and soil residence times (Litter C and Soil C).

By linking the module definitions in Table 1 with the sequential variance propagation shown in Figure 5, we believe the logical flow of the carbon cycle will be deeply integrated into the sensitivity analysis narrative.

[Major Comment 6] The description of the three applied methods to assess model sensitivity should be extended and homogenised. In Eqs. 2-5, the symbols need to be described (in a similar way as done for Eq. 1), and it would be good to avoid using different symbols for the same properties (e.g., I presume that Y in Eq. 2-4 is the same as f() in Eq. 1, and X_i in eq. 5 the same as x_i in eq. 1).

Response: We thank you for pointing out these inconsistencies in our mathematical notation. You are absolutely correct; the mixing of notations (Y vs. f(X), X_i vs. x_i) across the descriptions of Morris, eFAST, and Sobol' methods causes unnecessary confusion.

In the revised manuscript, we will thoroughly overhaul Section 2.4 (Sensitivity Analysis Methods). We will establish a unified mathematical nomenclature at the beginning of the section (e.g., universally defining the model as $Y = f(X_1, X_2, ..., X_n)$) and strictly adhere to these symbols across all equations (Eqs. 1-6). We will also ensure that every symbol in Eqs. 2-5 is explicitly defined immediately following the equation.

[Major Comment 7] The authors make a very pronounced distinction between "hard-coded" generic parameters and parameters describing specific PFT traits. It is unclear to me whether there is a fundamental difference between these parameters: whether or not parameters are hard-coded is a merely technical distinction that does not alter their significance for the model. Moreover, the PFT-specific parameters seem to be equally treated as generic... I would urge the authors to elaborate on this distinction and use the aspect of "generic" vs. "PFT-specific" more clearly in their analysis.

Response: This is a profound observation. From a pure mathematical or biogeochemical perspective, you are entirely correct: a parameter governing a process is simply a parameter, regardless of where it resides in the software architecture.

However, our strong emphasis on the "hard-coded" vs. "PFT-specific (user-facing)" distinction is driven by a very real operational and technical barrier within the DGVM user community. Historically, sensitivity analyses and calibration efforts have overwhelmingly focused on adjusting the easily accessible PFT parameter text files (e.g., .ins files), treating the core physiological constants hidden deep within the C++ source code as immutable.

This technical bias was explicitly acknowledged in a recent comprehensive LPJ-GUESS sensitivity analysis by Oberpriller et al. (2022). In their limitation section, they noted that highly sensitive physiological parameters like *ALPHA_C3* were *"omitted from our analysis"* specifically because they are *"not accessible in the parameter input file. Instead, they are hard coded in the model's source code."* While they acknowledged that it *"would certainly be useful (although very complicated) to explore these uncertainties together with the factors considered here in a joint analysis,"* they ultimately restricted their scope to user-accessible parameters.

Our study directly tackles this complicated challenge. By modifying the C++ source code to expose these hard-coded physiological traits, our GSA results empirically demonstrate that these "hidden" generic parameters (like *ALPHA_C3* or *THETA*) actually co-dominate or even override the user-facing PFT traits in driving model uncertainty.

To make this intention clearer and better contextualized in the revision, we will:

Contextualize in the Introduction: Explicitly state that this distinction reflects current operational modeling practices rather than biological reality. We will cite Oberpriller et al. (2022) to highlight the prevalent technical bias of excluding hard-coded parameters, framing our joint analysis as a necessary step forward.

Elevate in the Discussion: Use this distinction to formulate a strong recommendation for the modeling community: robust model calibration cannot treat DGVMs as black boxes by only tuning user-interface files; it must critically evaluate the generic physiological equations hard-coded at the core of the model.

References:

Oberpriller, J., Herschlein, C., Anthoni, P., Arneth, A., Krause, A., Rammig, A., Lindeskog, M., Olin, S., and Hartig, F.: Climate and parameter sensitivity and induced uncertainties in carbon stock projections for European forests (using LPJ-GUESS 4.0), Geosci. Model Dev., 15, 6495–6519, https://doi.org/10.5194/gmd-15-6495-2022, 2022.

[Major Comment 8] In general, table captions as well as some figure captions (e.g., Fig. 7, 8) should be extended to provide sufficient information for understanding and interpreting the table/figure content.

Response: We accept this feedback. In the revised manuscript, we will thoroughly review and expand all table and figure captions. We will ensure that every caption contains sufficient detail—including descriptions of axes, color scales, abbreviations, and the statistical methods used to generate the plots—so that the figures and tables can be fully understood independently from the main text.

[**Major Comment 9**] L65: In the introduction, it would be nice to get a brief characterization of different methods for GSA (helping the reader to see pros and cons of these different methods). Also, when discussing the existing literature in the subsequent paragraph, it would be nice to highlight which of these methods have been used.

Response: We thank the reviewer for this excellent suggestion, which will greatly improve the logical flow and educational value of the Introduction. In the revised manuscript, we will add a concise summary of the three GSA methods in the Introduction (Section 1), outlining their pros and cons: the Morris method (computationally efficient for screening but qualitative regarding interactions), the eFAST method (variance-based, captures interactions with moderate cost), and the Sobol' method (the rigorous gold standard for variance decomposition but computationally expensive). Furthermore, we will explicitly mention which GSA methods were utilized in the cited literature to better contextualize our multi-method approach.

[**Major Comment 10**] L86: This sentence is a repetition from L66. Consider removing one of the instances.

Response: Thank you for catching this oversight. We will remove the redundant sentence in Line 86 in the revised manuscript to keep the text concise.

[**Major Comment 11**] L137: Clarify what is meant with "upper soil moisture content". Soil moisture content of the upper layer? Which depth does this represent?

Response: We apologize for the ambiguity. In the LPJ-GUESS model structure, the soil column is typically divided into an upper and a lower layer. The “upper soil layer” corresponds to the top 0–50 cm of the soil profile. This interpretation is consistent with the model source code: in the multilayer soil hydrology scheme, water is explicitly distributed within the "upper 50 cm" of the soil column, and the soil-layer settings indicate that the upper soil domain typically consists of five 100 mm layers, i.e., 500 mm in total. Therefore, in our manuscript, "upper soil moisture content" refers to the soil moisture content of the top 0–50 cm. We will explicitly state this depth range (0–50 cm) in the revised text to ensure absolute clarity regarding the fire probability function.

[**Major Comment 12**] Section 2.2.3 The model setup is described for the spin-up period. It would be nice to elaborate in a similar way about the settings used for the period after spin-up.

Response: We agree that the description of the simulation protocol was incomplete. In the revised manuscript, we will expand Section 2.2.3 to include a description of the transient (historical) run. We will clarify that following the 500-year spin-up, the model was driven by transient historical meteorological forcing (temperature, precipitation, radiation), varying annual atmospheric CO_2 concentrations, and nitrogen deposition data for the period from 1980 to 2023 to generate the final outputs used for the sensitivity analysis.

[**Major Comment 13**] L265: unclear what "Python process pools" are. Process tools?

Response: We apologize for the technical jargon. "Python process pools" refers to the multiprocessing.Pool module in the Python programming language, which we used to run multiple LPJ-GUESS model simulations simultaneously across multiple CPU cores (parallel computing) to reduce the total computation time. To avoid confusion for readers unfamiliar with computer science terminology, we will rephrase this in the revised manuscript to: "parallel computing utilizing Python's multiprocessing module."

[**Major Comment 14**] L388 and elsewhere: I would avoid using the word "drivers" for the parameters. In a modelling context, the drivers or driving data are typically considered to be the model forcing, not the equation parameters.

Response: This is a very precise and important terminological correction. We completely agree that mixing "drivers" (which should strictly refer to external climatic forcing like temperature and precipitation) with "parameters" (internal equation constants or traits) creates confusion. We will conduct a global search throughout the revised manuscript to ensure the word "drivers" is strictly reserved for meteorological forcing data, and we will replace it with "parameters" wherever we are discussing the sensitivity of internal model variables.

We appreciate your warm work earnestly, and hope that the correction will meet with approval. We tried our best to improve the manuscript and made some changes in the manuscript. These changes will not influence the content and framework of the paper.

Once again, thank you very much for your comments and suggestions.