Response to Reviewer 2

The comments of Reviewer 2 are accessible here:
https://doi.org/10.5194/egusphere-2025-1260-RC2

Below, we provide a point-by-point reply to the review. Text by the reviewer is indented and in blue font. Our reply is in black font and not indented. For ease of referencing our replies, we numbered them.

General comments:

> The authors presented a Simulating Optimal FUNctioning {rsofun} R package, aiming to lower the bar of entry to ecosystem modelling and model-data integration for scientists. This R package providing the potential of efficiently model parameterization and estimation of uncertainty. The authors tested the {rsofun} R package at the site level by applying parameterization to the P-model and presenting a corresponding case study. The results demonstrate that this tool can be used to calibrate photosynthesis-related parameters of the P-model at site scale and to simulate GPP accordingly. The package shows potential to advance research in ecosystem modeling. However, further development is still required. The case study presentation lacks sufficient detail regarding key aspects of the tool, such as how model–data integration is implemented. Moreover, the core functionalities of the tool appear to rely heavily on other existing R packages. The authors are encouraged to clearly identify the unique contributions and core features developed specifically within this package—particularly whether it provides a generalized interface to invoke different ecosystem models. Overall, the {rsofun} R package is a promising tool to support ecosystem modeling studies, but it remains incomplete in its current form. Its potential for application at regional scales is still uncertain. Therefore, I recommend substantial revisions before the manuscript can be considered for acceptance.

[r2.1] We appreciate the critical and constructive feedback by the reviewer. In their comment above, the reviewer raises (in our reading) three main points.

1. Lacking detail about the implementation of model-data integration.

We appreciate that this point was not sufficiently covered by our initially submitted manuscript. As stated also in our reply to reviewer 1 [r1.1], we will provide a detailed description of likelihood functions - the core of the model-data integration - for the

different calibration setups that will be employed. A description of the planned calibration setups is given in our reply to reviewer 1 [r1.8]. Added details about the implementation will also be provided by the planned addition of a brief overview of the architecture and implementation of the model code. This will describe how diverse data types (flux time series, static traits) serve as calibration targets via separate sets of simulations (a simulation set for the set of flux measurement sites and a simulation set of the set of locations at which traits were measured).

2. Unique contributions and features of the *rsofun* package.

The *rsofun* package serves two main purposes: First, it provides a comprehensively tested and documented, open-access implementation of two vegetation models, the P-model and BiomeE, of which the former is described in our study. The models themselves have been presented elsewhere (Stocker et al., 2020; Weng et al., 2019). Second, it "provides off-the-shelf methods for Bayesian (probabilistic) parameter and prediction uncertainty estimation" (p. 6, l. 22). The implementation of Bayesian methods is provided by the *BayeisanTools* R package and described elsewhere (Hartig et al., 2023). The combination of the two general functionalities into one package is a key contribution by our work and "acknowledges that the data is an integral part of the modelling process (Dietze et al., 2013)" (p. 2, l. 59). Our work thus makes workflows of model-data integration transparent, easily implementable, and readily reproducible. In other words, "*rsofun* lowers the bar of entry to ecosystem modelling and model-data integration and serves as an open-access resource for model development and dissemination." (p. 1, l. 15). We note also that the implementation of a model for simulating vegetation and land surface processes as an open source package and its distribution via a rigorously tested software library repository (here CRAN) is well beyond the current practice in vegetation modelling. Some state-of-the-art models are open access (e.g., the Community Land Model https://www.cesm.ucar.edu/models/clm, or Noah-MP https://ral.ucar.edu/model/noah-multiparameterization-land-surface-model-noah-mp-lsm). However, it is common that such models are provided as Fortran source code, without forcing data, making it challenging for inexperienced users to compile and run the model, let alone to implement a systematic model calibration. *rsofun* addresses this challenge by providing R wrappers that connect to two specific models (P-model, BiomeE) and calibration routines (BayesianTools and GenSa), while model code itself remains implemented in computationally efficient Fortran. We consider that this is a critical step towards making vegetation modelling accessible to novices in the field (e.g., graduate students) and thereby advancing research through a practical

innovation. In revising the manuscript, we will more clearly state (and re-state) these points in the Discussion section.

3. The potential application of the model across regions (to global).

We agree that the spatial generalisability of the model was not tested in our initially submitted manuscript version (pre-print). In the revised version, we will address this limitation by including data from multiple sites for model calibration and including data from a disjunct set of (out-of-sample) sites for model evaluation (see also [r1.8]). This will yield a model calibration that is informed by observations from a wide range of climates and will yield a model evaluation that addresses spatial generalisability within the data covered by the networks of ecosystem flux measurements. We note however, that the measuring of spatial generalisability and upscaling for global model applications remains a challenge which goes beyond the scope of our study (Ludwig et al., 2023; Meyer and Pebesma, 2021). Ultimately, such investigations are limited by how flux measurement sites cover the (global) space of environmental conditions, and how transferrable models are across plant species and communities. Our revised manuscript will discuss these limitations and present results of the evaluation with a view to spatial generalisability of the model. However, we consider a rigorous test of model generalisability across large environmental gradients to not be a central aspect of this study and we will thus limit our scope to using the set of sites described in our reply to reviewer 1 [r1.8].

Specific comments:

> [line number]: Please using continuous line numbers.

[r2.2] We apologise for this inconvenience. The display of re-initiated line numbers by page was unintentional and will be avoided in a re-submitted manuscript.

> [Page 1, Line 11]: Whether site-scale simulations or parameterization need parallelizable computation？How about applying this package at a larger scale？

The package is designed for site-scale simulations, but allows for a distribution of any number of simulations across multiple “sites”, representing for example gridcells for a global simulation. This distribution is parallelisable. The rationale for the “time-first” organisation of the model forcing data structure and model integration are given by the following points. First, due to temporal memory in the model (e.g., soil water balance),

full forcing time series must be provided to the model, unless information is kept in memory for multiple (spatially arranged) gridcells. This is restrictive for the memory usage if the model simulation covers a large number of gridcells. Second, because there is no lateral dependency of gridcells through mechanisms simulated in the model, gridcell-wise (site-scale) simulations can be run in isolation, i.e., they can be easily parallelised, with a very limited penalty from the overhead created by distributing jobs across compute cores. We have made use of this aspect and have implemented code for parallelised global simulations using *rsofun* in the *grsofun* repository (https://github.com/geco-bern/grsofun). However, since these functionalities are not contained within the *rsofun* package and since *grsofun* is designed not for CRAN and a wide distribution, we excluded this application from the scope of our present study.

> [Page 2, Line 58]: There is no table before this table, so please rename it as Tab. 1.

[r2.3] This is a typo which we will correct.

> [Page 4, Line 08]: Was net radiation an output in P-model？and it is not suitable to list soil temperature in ecosystem water balance.

[r2.4] We will rename ‘Ecosystem water balance’ to ‘Ecosystem states’ and delete net radiation from the list of model outputs (here, it is used as a model input - this was an error).

> [Page 4, Line 14]: There is no information about spin-up period or corresponding years in Table 1, please move this sentence to corresponding section.

[r2.5] All relevant details for reproducing simulations (in addition to code provided already along with our initial submission) will be included (see also our reply to Reviewer 1 [r1.1]).

> [Page 7, Line 41]: Why the author selected this site, and as rsofun package so computationally efficient, why not test at more sites that can represent different conditions. What is the depth of the root zone soil at this site?

[r2.6] We will use a larger set of sites for model calibration and evaluation in our revised manuscript. See also our reply [r1.8].

> [Page 7, Line 60]: If the authors mentioned analysis, please give the result in manuscript or supplementary materials.

[r2.7] We will use a set of three alternative calibration setups in our revised manuscript. See also our reply [r1.8].

> [Page 7, Line 62]: When all parameters are used for model calibration, the multidimensional parameter space increases exponentially, potentially leading to model non-convergence. This also indicates that different values of c*, beta, tau and b0 can influence the parameterization outcomes. Meanwhile, the fixed parameters should be explicitly justified in terms of how they were determined.

[r2.8] With our proposed changes (alternative calibration setups and corresponding sets of calibrated vs. fixed parameters), this point will be addressed. See also our reply [r1.8].

> [Page 8, Line 71]: Given that c* is the third most sensitive parameter, why was it not included in the calibration, while β0, which ranks second to last in sensitivity, was? The authors are requested to further clarify the rationale behind the parameter selection.

[r2.9] With our proposed changes (alternative calibration setups and corresponding sets of calibrated vs. fixed parameters), this point will be addressed. See also our reply [r1.8].

> [Page 10, Line 90]: It is hard to recognise the light green band, and please explain more about the model uncertainty and parameter uncertainty.

[r2.10] We will add a brief discussion about the magnitude of different error magnitudes in the revised manuscript.

> [Page 10, Line 99]: It's dark orange or red line, the authors described as red line before.

[r2.11] We will revise our selection of colors to avoid confusion. What the reviewer refers to as ‘orange’ is called “tomato” in R.

References

Dietze, M. C., Lebauer, D. S., and Kooper, R.: On improving the communication between models and data, Plant, Cell & Environment, 36, 1575–1585, https://doi.org/10.1111/pce.12043, 2013.

Hartig, F., Minunno, F., Paul, S., Cameron, D., Ott, T., and Pichler, M.: BayesianTools: General-Purpose MCMC and SMC Samplers and Tools for Bayesian Statistics, 2023.

Ludwig, M., Moreno-Martinez, A., Hölzel, N., Pebesma, E., and Meyer, H.: Assessing and improving the transferability of current global spatial prediction models, Global Ecology and Biogeography, 32, 356–368, https://doi.org/10.1111/geb.13635, 2023.

Meyer, H. and Pebesma, E.: Predicting into unknown space? Estimating the area of applicability of spatial prediction models, Methods Ecol Evol, 12, 1620–1633, https://doi.org/10.1111/2041-210X.13650, 2021.

Stocker, B. D., Wang, H., Smith, N. G., Harrison, S. P., Keenan, T. F., Sandoval, D., Davis, T., and Prentice, I. C.: P-model v1.0: an optimality-based light use efficiency model for simulating ecosystem gross primary production, Geoscientific Model Development, 13, 1545–1581, https://doi.org/10.5194/gmd-13-1545-2020, 2020.

Weng, E., Dybzinski, R., Farrior, C. E., and Pacala, S. W.: Competition alters predicted forest carbon cycle responses to nitrogen availability and elevated CO₂: simulations using an explicitly competitive, game-theoretic vegetation demographic model, Biogeosciences, 16, 4577–4599, https://doi.org/10.5194/bg-16-4577-2019, 2019.