Anonymous Referee #2

The authors of the manuscript 'A differentiable ecosystem modeling framework for large-scale inverse problems: demonstration with photosynthesis simulations' describe the application of the 'differentiable parameter learning'(dPL) framework to the photosynthesis module of FATES model. The framework, and concept, overcomes extrapolation limitations from site-by-site calibration approaches and allows leveraging information content in large-scale datasets towards a global parameterization of photosynthesis models. Neither the concept (Tsai et al., Nature Communications, https://www.nature.com/articles/s41467-021-26107-z, 2021; Bao et al., Authorea, https://www.authorea.com/doi/full/10.1002/essoar.10512186.3, 2022) nor the dPL framework (Tsai et al., 2021; Feng et al., 2022ab) are new. However, the framework is used in the FATES model for the first time and the results would be of interest for further model development, but also to the scientific community at large.

At this point, the experiment focuses on inverting two parameters, V_{cmax25} and B, resulting in that the accuracy of the simulated net photosynthesis rate being slightly improved. The main concerns at this stage relate to apparently incorrect formulations of some key equations, to issues about the validation strategy, to the fact that the forcing data and the experiments are not described sufficiently, challenging the acceptance of the study, while hampering any reproducibility efforts. Please see below for details.

Thank you for your comments! We are preparing a fuller response to your comment, but here are some rapid response to two of your comments.

We were indeed following our previous differentiable parameter learning paradigm which were applied in hydrology (Tsai et al., 2021; Feng et al., 2022), as noted in the manuscript, but this is a novel use in ecosystem modeling. We could not have noticed Bao et al., 2022 as it went online after ours did and seems to be undergoing review. Upon some examination, we believe the basic modules are very different. They are using a light-use-efficiency approach and predicted GPP, while our paper focused on photosynthesis using a Farquhar-type model. Hence we don't think there is much overlap between the two.

Major comments:

1. Two key equations are incorrect in the paper:

1) line 140: equation 5, $C_i=C_a-A_n*P_{atm}*(1.4g_s+1.6g_b)/(g_s+g_b)$;

2) line 505: equation A1, $A_c=V_{cmax}*(C_i-\Gamma_*)/(C_i+K_c*(1+K_o/O_i))$.

According to the user guide of the FATES model (https://fates-users-guide.readthedocs.io/projects/tech-doc/en/latest/fates_tech_note.html#fundamental photosynthetic-physiology-theory), the equations should be:

1) $C_i=C_a-A_n*P_{atm}*(1.4g_s+1.6g_b)/(g_s*g_b)$;

2) $A_c=V_{cmax}*(C_i-\Gamma_*)/(C_i+K_c*(1+O_i/K_o))$.

Since the FATES model is reimplemented in Julia and PyTorch by the authors, the codes might be also wrong. If so, the unit of C_i will be incorrect, leading to errors in the inversion of

V_{cmax25} and B. The wrong computation of the effective Michaelis-Menten coefficient ($=K_c*(1+O_i/K_o)$) might only have a slight effect if the temperature is close to 25°C, but should be concerned if the temperature is too low or high (and I do see some points with low leaf temperature in the 'Lin15' database). Thus, I have doubts about the current results and relevant analysis.

Regarding the equations --- we were cautious to adhere to the original FATES equations before implementing it on PyTorch or Julia. Unfortunately, we **realized there were some typos in the manuscript** in line 140 and line 505 in the paper which will then be modified. However, we used the correct equations in our differentiable model as the following:
$C_i=C_a-A_n*P_{atm}*(1.4g_s+1.6g_b)/(g_s*g_b)$;
$A_c=V_{cmax}*(C_i-\Gamma_*)/(C_i+K_c*(1+O_i/K_o))$.

No results need to be changed. The code was correct as we compared carefully against the Fortran code in these subroutines as we developed the differentiable versions of the code. We will be publishing the code as the paper gets closer to acceptance so this can be examined in the code. Again, we apologize for the errors in the manuscript.

As all the results are validated only once using the temporal holdout data or the random holdout data, the generalizability of the dPL (or NN_B+NN_v) is not clear. If the N-fold or leave-one-out cross-validation can be adopted, the statistical metrics can be more justifiable to reflect the model performance.

Thanks for being rigorous. We believe the randomly-selected points were representative but we will conduct a cross validation (CV) and show the results. However we would also like to note that a temporal test is actually more stringent than a random-holdout cross validation. The temporal test examines if the model can run forward correctly for a period of time not included in the training whereas a random selection will always include some points from most recent times. The metrics also show that random holdout is better than temporal holdout. We noticed this issue through careful benchmarking in previous papers [Feng et al., 2021], but the uncertainty associated with temporal projection was often not paid attention to. We expect the full CV results to be better than temporal test results. We believe a spatial test, though, would belong to a different paper as the paper is already getting long. There are many techniques to improve spatial generalization and larger dataset from remote sensing which, if combined with the present content, would just be too much for a first paper. We plan to clarify this point in the paper.

The forcing variables and parameters are not clearly differentiated in the paper. For example, is the leaf layer boundary conductance, g_b, a constant parameter across sites or a temporally changing variable? If it is a forcing variable for FATES, where is g_b from? is θ_{ice} a forcing variable or a parameter correlated with temperature and θ_{liq}? Is the C_a a constant value or variable? The model would be different if the spatial and temporal variability of all these factors are considered. If all these are parameters (i.e., scalars), what are the values?

The Lin15 dataset included different forcing variables that we used in our model including:

RH	Relative humidity
T_{air}	Air temperature
T_{leaf}	Leaf temperature
P_{atm}	Atmospheric pressure
PAR (φ)	Photosynthetic active radiation
g_b	Boundary layer conductance

Concerning (g_b, θ_{ice} and C_a), here are details about how they were considered in the model:

- g_b, the boundary layer conductance values were already available in Lin15 dataset. However, it has some missing values which were computed using the (GetCanopyGasParameters) subroutine available in Fates model. https://github.com/NGEET/fates/blob/main/biogeophys/FatesPlantRespPhotosynthMod.F90
- θ_{ice}, the volumetric ice content values were ignored (considered as zero) since both the air and leaf temperatures in our dataset were above the freezing temperature (0 °C or 273.15 K) by at least 5 degrees.
- C_a, the CO_2 partial pressure near the leaf surface values were variable spatially and temporally and they were taken as 0.039% of the atmospheric pressure

Line 216-218: the reason for replacing saturated soil matric potential (Ψsat) with soil matric potential for closed stomata (Ψc) is not explained. Equation 10 shows that the Ψsat is replaced with soil matric potential for open stomata(Ψo), not Ψc. Furthermore, the Ψi was still calculated using Ψsat in Appendix A (equations A16-A18). I'm confused about which variable was used to calculate Ψi.

Line (216 – 218), we stated the actual equations that we used in for computing ψ_i (in which ψ_{sat} was replaced with ψ_o).

In Appendix A, we kept all the original equations the same whether those related to FATES or to computing the soil water stress function (β_t).

Actual equation used in this study (Line 216 – 218)	Original equation (Appendix A)
$\Psi_i = \Psi_o \times S_i^{-B_i} \geq \Psi_c$	$\Psi_i = \Psi_{sat,i} \times S_i^{-B_i} \geq \Psi_c$

Reasons for this replacement:

In the original CLM4.5 equations, ψ_{sat} is based on empirical functions, percentage of sand *(%sand)*, and fraction of organic matter (F_{om}) (Equations A17 – A18). Using the original Equation 7 for computing ψ_i results in a plant wilting factor w_i equals to one for more than 90% of the data points across different soil layers.

To give the model more flexibility in the computation of ψ_i and thus allow more variability in w_i values, ψ_{sat} was replaced with ψ_o. However, to ensure having w_i values less than or equal 1 as in the original equation 9, we tried to create equation 10 in a way that satisfies this condition using ψ_o. For parameter B (outputted from NN_B), it was restricted to be within the range 0 and 1 to satisfy the same condition as well. Applying those changes, we were able to get ψ_i values within the range of ψ_o and ψ_c while showing more variability in the computed w_i.

Also, we plan to add this paragraph to the **(Model changes)** section for clarification:

"*These changes were implemented to give more flexibility in the computation of the soil matric potential ψ_i. Using the original Equation 7 for computing ψ_i results in a plant wilting factor w_i equals to one for more than 90% of the datapoints across different soil layers. Thus, changing Equation 7 to the form shown in Equation 10 helped to express more variability in w_i and eventually in the computed soil water stress function (β_t).*"

Here, the point is to calculate photosynthesis. We can see clearly the modified model works very well for photosynthesis. The differentiable modeling approach was specifically designed to enable inspection of various modules and assumptions in the model to update the formula.

As indicated, we are preparing a fuller response to you and will soon post it here. The description above only represents a quick reply. Thank you for your comments again!

References.

Shen, CP. et al., Differentiable modeling in Geosciences to unify machine learning and physical models. https://arxiv.org/abs/2301.04027

Wen-Ping Tsai, Dapeng Feng, Ming Pan, Hylke Beck, Yuan Yang, Kathryn Lawson, Jiangtao Liu, and Chaopeng Shen. From calibration to parameter learning: Harnessing the scaling effects of big data in geoscientific modeling. Nature Communications, (2021). doi: 10.1038/s41467-021-26107-z.

Feng, DP., K. Lawson and CP. Shen, Mitigating prediction error of deep learning streamflow models in large data-sparse regions with ensemble modeling and soft data, Geophysical Research Letters, doi: 10.1029/2021GL092999 (2021)