Dear Biogeosciences editor and reviewers,

Thank you for handling and thanks to the reviewers for their constructive suggestions. It seems most comments were asking for clarifications and further exploration of data. Reviewers have commented that they would like to see a revised version. While responding to reviewers, we have already run cross-validation experiments and have provided additional visualizations including plots for different PFTs as requested by the reviewers. We would also stress that as a first paper with a differentiable model for a Farquhar formulation, this paper needs to have a strong focus. We did not claim the model can well simulate seasonality at a site, and will further clarify it in the revised manuscript that this point is to be studied in future work. Because the backbone of the model is Farquhar, its seasonal behavior should be comparable to what we expect out of the other Farquhar models, because in this paper we only estimated static parameters.

In the following, we add more complete answers to a few detailed questions that were asked before the interactive system closes so we didn't have the time to address.

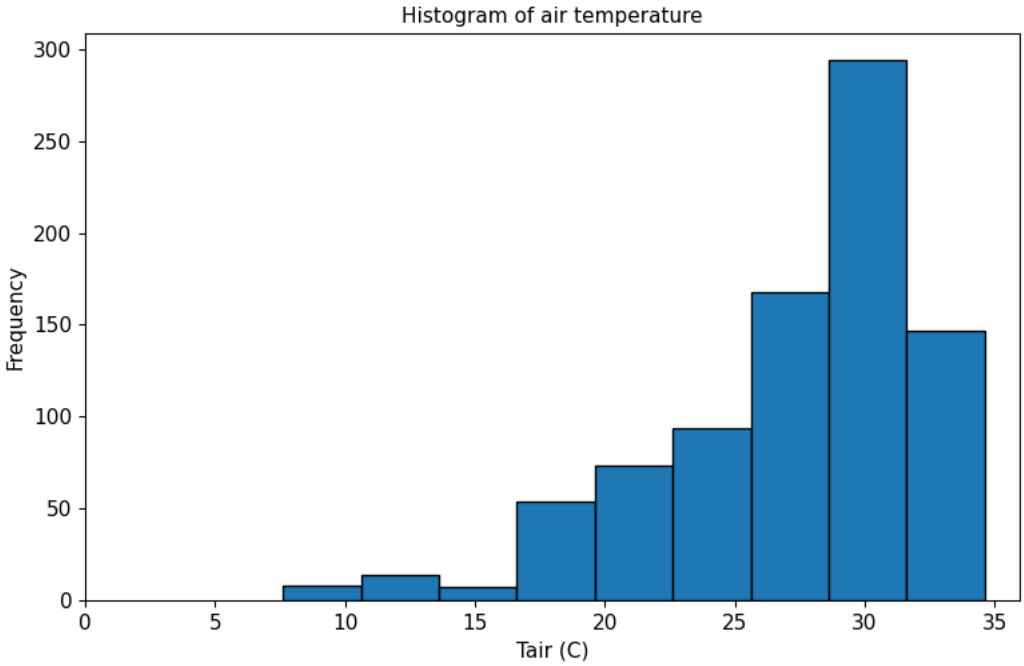
---

1. Line 218-220: is NN_B used to predict B_i or Ψ_i? B depends on only %clay and F_{om} according to equations A22-A23, while the authors add %sand, which is related to Ψ_{sat} and, therefore, Ψ_i. I didn't find a direct relationship between B_i and %sand according to the original equations in the FATES model. If NN_B is used to predict Ψ_i, I think the equation can be $\Psi_i=\theta_{liq}*NN_B$(%sand,%clay,PFT,F_{om},T), where T represents the factors controlling θ_{ice}, e.g., temperature.

It was suggested in this comment to include the temperature in NN_B to represent θ_{ice} as the following:
$\Psi_i=\theta_{liq}*NN_B$(%sand,%clay,PFT,F_{om},T), need to rescale between Ψ_c and Ψ_o

However, we claim that including T in NN_B to represent θ_{ice} might not very effective as justified afterwards.

The histogram of air temperature (shown in the figure below) indicates that we don't have in our dataset air temperatures below 5 °C and that clarifies why θ_{ice} was ignored in our calculations. Thus, there is low probability that the temperature would have a great effect being included in NN_B.

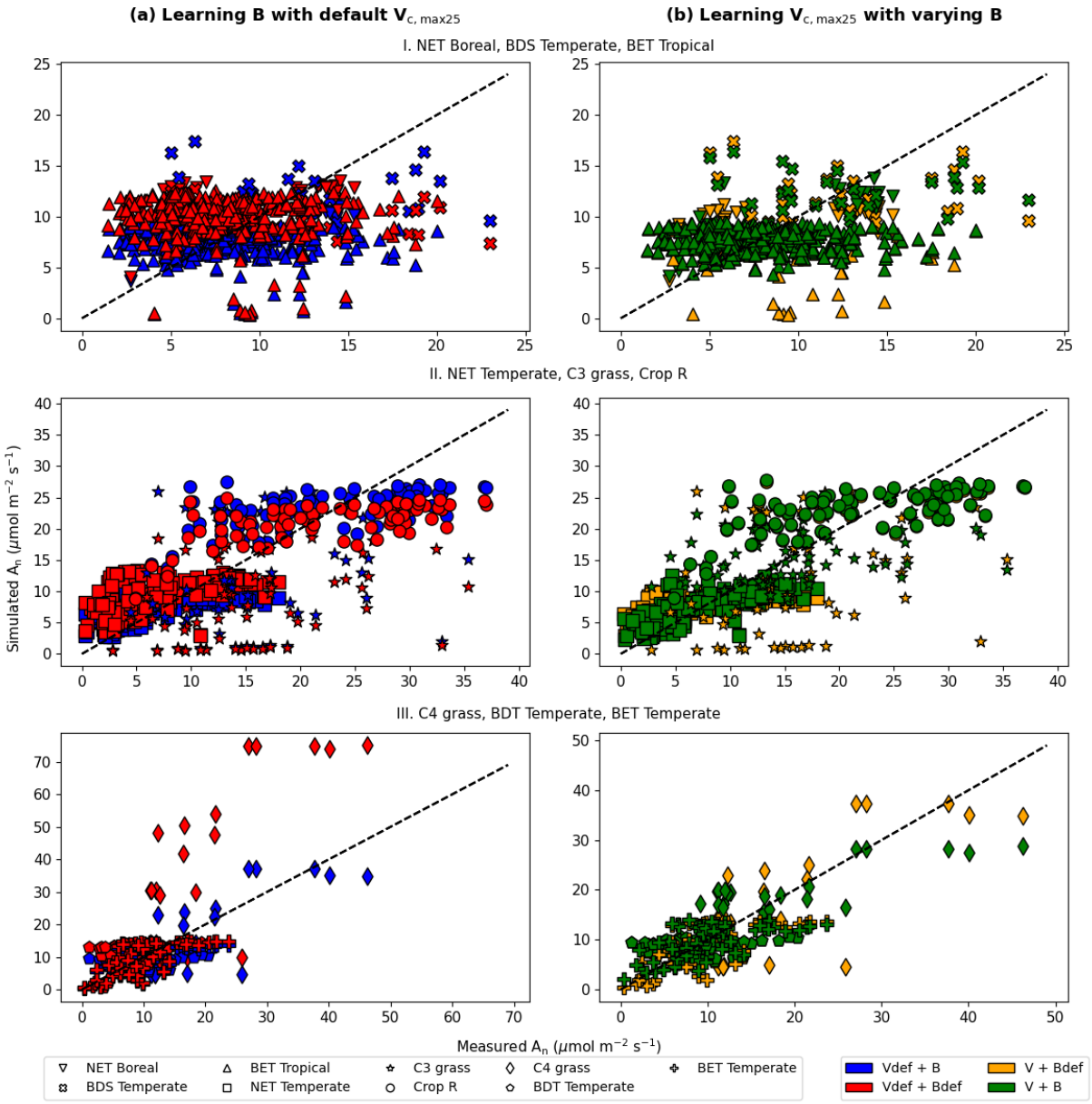


2. Line 431-432: I cannot identify the C3 grass at the lower left corner of figure 5b. Maybe a violin plot per PFT can be helpful to show the difference between optimizing B or not for a specific plant type. The figures in the paper only show the net photosynthesis rate across all sites. However, the site-level comparison might be more meaningful to assess the four parameterization strategies: Vdef+B, Vdef+Bdef, V+Bdef, and V+B.

   The violin plots showed the net photosynthesis per PFT, but I think readers would be more interested in how different is the simulated net photosynthesis from the measured net photosynthesis. Maybe the fourth violin plot (measured values) can be added on the right side and the NSE can be displayed at the top. Moreover, I think only the test dataset (or better cross-validated dataset) should be compared with the measured values (e.g., Fig. 5) and used to make the violin plots.

   For better clarification of different PFTs, we split figure 5 into 3 rows representing 3 PFTs (each of 2 subplots) as shown below. We used cross validated dataset for making the plot to avoid confusion concerning which dataset was used for the plot (train or test). Splitting figure 5 this way helps present the same information as the violin plot (which we can then exclude in the new version to avoid redundancy).

   Regarding computing the NSE per PFT, as expected, this leads to lowering down the performance (lower NSE values) as shown in the attached table for each PFT, especially BET Tropical and BDT temperate, and C3 grass. We will comment on it in the revised manuscript.



PFT	NSE_train	NSE_test	No. of datapoints per PFT	No. of species per PFT	No. of locations per PFT
Crop R	0.3687	0.3648	85	5	4
NET Boreal	0.6604	0.6178	32	3	3
BET Tropical	-0.1400	-0.1419	367	185	11
NET Temperate	0.5859	0.5804	153	10	9
BET Temperate	0.3880	0.3401	69	16	9
BDT Temperate	-0.0193	-0.0329	61	12	7
C3 grass	-0.0180	-0.0564	58	17	2

BDS Temperate	-0.0246	-0.2257	28	12	4
C4 grass	0.5299	0.5021	21	6	1

3. There are limited site-level temporal data, thus the seasonality of net photosynthesis cannot be assessed.

We did not claim the model can simulate seasonality very well at a site. Currently our differentiable model follows the same structure of the photosynthesis module in the process-based model "FATES". We didn't make significant changes to the model. Because the backbone of the model is Farquhar, its seasonal behavior should be comparable to what we expect out of the other Farquhar models, because in this paper we only estimated static parameters. On the other hand, the nature of our dataset doesn't enable us to test the seasonality and we didn't mention this in the manuscript. This might be our scope in the future work.

To avoid confusion, we would add limitations section in the new version of the manuscript including the following paragraph:

"*Although applying the dPL framework improved the parameters to an extent, the model still has similar structural limitations as other Farquhar-type models. We didn't test the model's ability to capture the seasonality of the net photosynthetic rate due the limited site level temporal data. The seasonal behavior of the model is expected to be similar to other Farquhar models.*"

4. The authors clarified that the Vcmax25 is predicted per PFT but did not mention Bi. Is Bi predicted per PFT or per site? How is the predicted Bi compared with values from CLM?

B_i differs between different sites and for one site it differs for different PFTs; $B_i = NN_B$(%sand, %clay, F_{om}, PFT). Contrary to $V_{c,max25}$, there are no default values for B because of two reasons:

a. B in the default CLM4.5 equations come from empirical equations based on %clay and F_{om}
b. We changed equation 7 to equation 10 (as shown below). Thus, parameter B in equation 7 has a completely different range from the one in equation 10 which ranges between 0 and

Equations 7 and 10

$$\Psi_i = \Psi_{sat,i} \times S_i^{-B_i} \geq \Psi_c \qquad \text{(7) default}$$

$$\Psi_i = \Psi_o \times S_i^{-B_i} \geq \Psi_c \qquad \text{(10) New}$$

Default B equations in CLM4.5

$$B_i = (1 - F_{om,i}) \times B_{min,i} + F_{om,i} \times B_{om} \qquad \text{(A22)}$$

$$B_{min,i} = 2.91 + 0.159 \times (\%clay)_i \qquad \text{(A23)}$$

For better clarification, this text can be modified in the new version in which we state that we used the two symbols Ψ_i (PFT) and C_i (in place of Ψ_i and B_i respectively) to avoid confusion with the terms in the original equations.

[*In the original water limitation function in CLM4.5, the stomata response to soil water potential is based on a linear function between the water potential for stomata openness and closeness. In view that plant could respond to soil water potential differently dependent on plant hydraulic traits (Christoffersen et al. 2016), in this study, we modified the soil water limitation for PFTs so that they could have different shapes. Specifically, we designed a PFT-dependent soil water stress* $[\Psi_i\,(PFT)$, *ranging from* Ψ_c *and* $\Psi_o]$ *depending on the soil water content, which is calculated as follows:*

$$\Psi_i\,(\mathrm{PFT}) = \Psi_o \times S_i^{-C_i(soil,PFT)} \geq \Psi_c \qquad (10)$$

Where C_i *is a PFT- and soil-texture-dependent shape parameter (between 0 and 1) estimated as:*

$$C_i = NN_c(\%\mathrm{sand}, \%\mathrm{clay}, F_{om}, \mathrm{PFT}) \qquad (11)$$

The PFT-dependent soil water stress $\Psi_i\,(PFT)$*, is then feed into the plant wiling equation (equation 9) as the following:*

$$w_i = \frac{\Psi_c - \Psi_i\,(\mathrm{PFT})}{\Psi_c - \Psi_o} = \frac{\Psi_c - \max\left(\Psi_c, \Psi_o \times S_i^{-C_i(soil,PFT)}\right)}{\Psi_c - \Psi_o} \qquad (12)$$

]

5. Since the site-level comparison and the site-average An comparison are not possible, the generalizability cannot be evaluated. However, the model performance across sites can be compared to other papers using the Farquhar model (e.g., Fig 1B of Chen at al., PNAS, https://doi.org/10.1073/pnas.2115627119, 2022).

Concerning, the spatial generalization or the site-level comparison, as mentioned in the previous responses, spatial test is not the scope of this paper. we are working on further improving the spatial generalization with some error mitigation approaches. This will add lots of content and should be for the scope of another paper. So, we can add the following paragraph to the limitations section and the future work:

"This study doesn't cover the spatial generalization of the dPL model since we don't present results for spatial tests or based on site-level comparison. To improve spatial generalization may require further changes in the model, dynamical parameters, or using other error mitigation approaches. This is not our scope for this study; however, it would be conserved for future work."

Things being asked for and will be added or modified in the new manuscript version:

1. Explanation for temporal and random holdout tests >> (will be clarified in the new version through paragraphs added)

2. f1 and f2 equations clear explanation in the manuscript body >> (proposed figure 2 will be added with explanation for the terms in the equations)

3. Details on NNs hyperparameters and hyperparameters tuning >> (will be clarified in the new version through paragraphs added)

4. Inquiries about Lin15 dataset >> (Number and full name of PFTs, forcing variables, atmospheric CO_2 (C_a), leaf layer boundary conductance (g_b)) >> (will be clarified in the new version through paragraphs added)

5. Reasons of replacing Ψ_{sat} by Ψ_o in equations 7 and 10 >> (will be clarified in the new version through paragraphs added)

6. CLM4.5 contribution to the study >> (will be clarified in the new version through paragraphs added)

7. Inquiries about B calculations across soil layers >> (will be clarified in the new version through paragraphs added for synthetic and real case experiments)

8. Cross validation tests >> (were performed and results will be added in the new version)

9. Model performance is impacted by certain set of model equations and forcings >> (will be clarified in the new version through paragraphs added)

10. Modify typos in model equations >> (will be modified in the new version)

11. NN_B and NN_v constraints on outputs and output range >> (will be clarified in the new version through paragraphs added)

12. More complex NN for the real case than synthetic case >> (already done for NN_B but not applicable for NN_v)

13. Loss function clarification >> (will be better clarified in the new version + Figure 1)

14. Timeseries of observations >> (can't be provided due to the site limited temporal data)

15. Spatial variability of the parameters not fully captured by dPL >> (spatial test is not the scope of this paper, we are working on further improving the spatial generalization with some error mitigation approaches. This will add lots of content and should be for the scope of another paper.

16. Soil organic carbon content unit conversion >> (will be clarified in the new version)

17. $V_{c,max25}$ correlation literature values >> (the proposed figure 8 shown in previous response will be added showing the correlation between learnt and reference $V_{c,max25}$ values)

18. Split plots per PFT >> (Figure 5 will be spitted into 3 rows each with only 3 PFTs)

19. Figure 5b plotted using both training and test datasets >> (will be replotted using the cross-validation test results in the new version)

References

Christoffersen, B. O., M. Gloor, S. Fauset, N. M. Fyllas, D. R. Galbraith, T. R. Baker, B. Kruijt, L. Rowland, R. A. Fisher, O. J. Binks, S. Sevanto, C. Xu, S. Jansen, B. Choat, M. Mencuccini, N. G. McDowell, and P. Meir. 2016. "Linking Hydraulic Traits to Tropical Forest Function in a Size-Structured and Trait-Driven Model (TFS v.1-Hydro)." Geosci. Model Dev. 9(11):4227–55. doi: 10.5194/gmd-9-4227-2016.