https://doi.org/10.5194/bg-2022-211
Preprint. Discussion started: 2 November 2022

Biogeosciences Discussions

A differentiable ecosystem modeling framework for large-scale inverse problems: demonstration with photosynthesis simulations

Doaa Aboelyazeed[1], Chonggang Xu[2], Forrest M. Hoffman[3,4], Alex W. Jones[5], Chris Rackauckas[6], Kathryn Lawson[1], Chaopeng Shen[1]

[1] Civil and Environmental Engineering, The Pennsylvania State University, University Park, PA 16802, USA
[2] Earth and Environmental Sciences Division, Los Alamos National Laboratory, Los Alamos, NM, 87544, USA
[3] Computational Sciences & Engineering Division and the Climate Change Science Institute, Oak Ridge National Laboratory, Oak Ridge, Tennessee, USA
[4] Department of Civil and Environmental Engineering, University of Tennessee, Knoxville, Tennessee, USA
[5] SciML Open Source Software Organization, https://sciml.ai
[6] Computer Science and Artificial Intelligence Laboratory (CSAIL), Massachusetts Institute of Technology, Cambridge, MA, 02139, USA

Correspondence to: Chaopeng Shen (cshen@engr.psu.edu)

Abstract. Photosynthesis plays an important role in carbon, nitrogen, and water cycles. Ecosystem models for photosynthesis are characterized by many parameters that are obtained from limited in-situ measurements and applied to the same plant types. Previous site-by-site calibration approaches could not leverage big data and faced issues like overfitting or parameter non-uniqueness. Here we developed a programmatically differentiable (meaning gradients of outputs to variables used in the model can be obtained efficiently and accurately) version of the photosynthesis process representation within the Functionally Assembled Terrestrial Ecosystem Simulator (FATES) model. This model is coupled to neural networks that learn parameterization from observations of photosynthesis rates. We first demonstrated that the framework was able to recover multiple assumed parameter values concurrently using synthetic training data. Then, using a real-world dataset consisting of many different plant functional types, we learned parameters that performed substantially better and dramatically reduced biases compared to literature values. Further, the framework allowed us to gain insights at a large scale. Our results showed that the carboxylation rate at 25°C ($V_{c,max25}$), was more impactful than a factor representing water limitation, although tuning both was helpful in addressing biases with the default values. This framework could potentially enable a substantial improvement in our capability to learn parameters and reduce biases for ecosystem modeling at large scales.

Short Summary. Photosynthesis is critical for life and is affected by a changing climate. Many parameters come into play when modeling, but traditional calibration approaches have faced many issues. Our framework trains coupled neural networks to provide parameters to a photosynthesis model. Using big data, we independently found parameter values that were correlated with those in the literature while giving higher correlation and reduced biases in photosynthesis rates.

1 Introduction

Plant photosynthesis is critically important for regulating the global carbon and nutrient cycles, and thus the future climate. Understanding future climate trajectories requires the understanding of photosynthetic responses to changes in environmental factors including atmospheric CO_2 concentrations, radiation, temperature, humidity, and nutrient and water availability

https://doi.org/10.5194/bg-2022-211
Preprint. Discussion started: 2 November 2022

(Kirschbaum, 2004). Photosynthesis is influenced by many factors such as higher CO_2 levels, reduced productivity of vegetation (i.e., nutrient concentration) (Thompson et al., 2017), intensified droughts (Urban et al., 2017; Xu et al., 2019) and rising temperatures (Dusenge et al., 2019) under a changing climate. To comprehensively evaluate the impacts of these changing processes and vegetation feedbacks to the atmosphere, we need accurate representations of photosynthesis in models.

For global assessments of the carbon cycle, vegetation models were developed to simulate terrestrial ecosystem processes and the distributions of vegetation, both vertically in the soil-plant system and horizontally across the landscape. Substantial efforts over the last few decades have improved the representation of vegetation and its responses and feedbacks to climate change (Fisher et al., 2018). A typical framework structure for a vegetation model is to keep track of changes in carbon and optionally nutrient states, driven by climatic variables and modulated by soil properties, with feedback to the climate, e.g., CO_2 releases, radiation, and vegetation composition and structure. A core component of the vegetation module is photosynthesis (Quillet et al., 2010).

Present ecosystem models for photosynthesis are based primarily on mechanistic descriptions of plant photosynthesis pathways, but this theoretically-sound modeling paradigm faces many challenges, with parametric uncertainty being a major one. Photosynthesis models may describe limitations of carboxylation rates, light availability, and plant-specific factors like enzyme efficiencies for C_3 and C_4 plants differently (Farquhar et al., 1980; Farquhar and Caemmerer, 1982; Meyer, 1983; Von Caemmerer, 2003, 2013; Yin and Struik, 2009). They contain many parameters that quantify these efficiencies and limitations. In the past, these parameters have been estimated from different approaches: 1) obtained from a limited set of *in-situ* sites and scaled based on climate and environmental factors (Verheijen et al., 2013); 2) calibrated on observational data site by site or for a few sites for a plant functional type (PFT) (Mäkelä et al., 2019; Wang et al., 2014); or 3) optimized based on environmental conditions (Ali et al., 2016). However, these estimated values may not be optimal at the global scale. Site-by-site calibrations using genetic or similar algorithms are highly expensive and are limited in their spatial coverage and generalizability to different PFTs and species. Furthermore, such calibration faces the issue of nonuniqueness (which some call equifinality (Beven and Freer, 2001)), where different parameter sets produce the same outcome. As a result, calibration can easily lead to poorly-generalizable parameter values. This problem exists for many domains with diverse parameters, including ecosystem modeling (Tang and Zhuang, 2008). It is similarly found in hydrologic modeling and has troubled scientists there for decades (Beven, 2006). More recently, some parameters can be fitted directly from large datasets with directly measured parameter values (Luo et al., 2021), which is highly valuable but is limited to those parameters with extensive observations, e.g., soil water retention and hydraulic properties. An efficient way to permit large-scale inverse modeling is needed.

https://doi.org/10.5194/bg-2022-211
Preprint. Discussion started: 2 November 2022

There has been substantial progress in utilizing modern machine learning (ML) for geosciences. Purely data-driven deep learning models (LeCun et al., 2015; Reichstein et al., 2019; Shen, 2018; Shen et al., 2018) directly learn from data so they tend to be fairly accurate and many have outperformed traditional models for a large number of applications such as hydrological (Feng et al., 2020, 2021; Rahmani et al., 2020, 2021), agricultural (ElSaadani et al., 2021; Hossain et al., 2019; Liu et al., 2022a; Saleem et al., 2019), cryosphere (Leong and Horgan, 2020; Zhang et al., 2019), water quality (Hrnjica et al., 2021; Zhi et al., 2021), and ecosystem modeling (Zhang et al., 2020, 2021). Unfortunately, deep learning models also lack interpretability and process clarity, and can only output trained variables with extensive observations. This need for data is often not satisfied for ecological processes.

To aid geoscientific models in general, Tsai et al. (2021) presented an efficient framework known as differentiable parameter learning (dPL), in an effort to leverage recent progress in ML to mitigate the issues listed above for parameter inversion. This framework turns parameter estimation into a large-scale ML problem. It is mainly composed of a parameter estimation module based on a neural network (NN), combined with a process-based model (or its surrogate). The whole framework must be "programmatically differentiable" (Baydin et al., 2018; Innes et al., 2019), which refers to a programming paradigm where we can efficiently and accurately obtain the gradients of the outputs with respect to any of the variables used in the model. Once we have programmatic differentiability, dPL can efficiently learn unknown functions from big data to serve as either a parameterization or process representation. Tsai et al. (2021) found that this framework scales well with more data, produces spatially and temporally well-generalized parameter sets, extends well to uncalibrated variables, and saves orders of magnitude in computational time. Feng et al. (2022a) further showed that an adopted, differentiable process-based hydrologic model with dPL could approach the performance of a purely data-driven ML model, and potentially outperform ML in data-sparse regions (Feng et al., 2022b). These successes can be conveniently migrated to the ecosystem modeling domain.

Here, we apply the dPL framework to the photosynthesis module of the Functionally Assembled Terrestrial Ecosystem Simulator (FATES) model. FATES is an ecosystem model that describes co-existence and competition in plant functional types (PFTs) (Koven et al., 2020). FATES can be used as an ecosystem module in the Community Land Model (CLM) (Oleson et al., 2013; Lawrence et al., 2019) to represent the ecosystem demography (Fisher et al., 2015). The photosynthesis module is based on the Farquar photosynthesis model. To apply the dPL framework in our study, we first reimplemented the photosynthesis module in FATES so that it became programmatically differentiable. Second, we connected this model to neural networks for parameter estimation. With this tool, we aim to answer the following questions: *(1) What is the achievable model performance, in terms of predicting photosynthesis rates in space and in time, by tuning the parameters for the classical photosynthesis module without changing the model structure? (2) Are parameters like $V_{c,max25}$ and soil water limitation factor simultaneously identifiable? (3) Are parameters learned from a large global dataset similar to the values we used in our current models?* In the following, we first described the photosynthesis model with different parameter

https://doi.org/10.5194/bg-2022-211
Preprint. Discussion started: 2 November 2022

estimation experiments and target datasets. We then discuss the parameters chosen to be estimated and their significance. Afterward, we presented the results from synthetic experiments and experiments based on real datasets from sites around the globe. Finally, we compared the learned parameters to values from the literature, and provided some suggestions for future work.

2 Methods and datasets

2.1. General overview

Our work focused on the photosynthesis module in FATES. Ignoring the impacts of vegetation states, this module, just as many others, can be formulated as a system of nonlinear equations with many parameters:

$$f_1(\mathrm{x};\ \theta, \theta_c, \mathrm{F}) = 0; y = f_2(\mathrm{x}, \theta, \theta_c, \mathrm{F}) \quad (1)$$

where f_1 represents a system of nonlinear equations, x represents the unknowns of the equations (in this case the internal leaf CO_2 partial pressure [pa]), y is an observable variable (in this case photosynthetic rate [μmol m^{-2} s^{-1}]) that is dependent on x via nonlinear equations f_2, F represents some forcing variables such as radiation and air temperature, θ represents a list of tunable physical parameters, and θ_c represents untuned constant attributes. Given a set of θ with known θ_c and F, we need to solve for x from f_1 and send the solution into f_2 to further compute y: $y = f_2(f_1^{-1}(\theta, \theta_c, F), \theta, \theta_c, F)$. This whole workflow can be lumped into one model:

$$y = \delta_{psn}(\theta, \theta_c, F) \quad (2)$$

where δ_{psn} represents the overall model. Some of the tunable parameters are typically formulated as being PFT-dependent (e.g., the maximum carboxylation rate), or related to soil water availability (e.g., the soil water stress). We posit that there exists a parameterization scheme, $\theta = g^W(R)$, which is a mapping relationship from some underlying attributes R to the physical parameters represented by a neural network with W as the learnable weights. Thus, we can learn W so the simulated variable y matches the observations y^*:

$$W = \mathrm{argmin}(L(\delta_{psn}(\theta, \theta_c, F), y^*)) = \mathrm{argmin}_W(L(\delta_{psn}(g^W(R), \theta_c, F), y^*)) \quad (3)$$

where L is the loss function. For the purpose of solving the inverse problem and training the neural network g^W in an "online" mode using gradient descent (the only practically-employed algorithm for neural network training), we reimplemented the photosynthesis module in FATES onto two differentiable platforms: Julia and PyTorch (discussed in more detail below).

In order to test the learning capability of our framework and the identifiability of the parameters, we first ran synthetic experiments to verify if the model would be able to retrieve assumed values for the physical parameters. Second, using a

https://doi.org/10.5194/bg-2022-211
Preprint. Discussion started: 2 November 2022

dataset with thousands of photosynthesis rate measurements, we trained the differentiable model to obtain estimated parameters at the global scale, and compared them to the literature.

2.2. The Farquhar photosynthesis model

The FATES photosynthesis module is based on the classical Farquhar model for C_3 plants (Farquhar et al., 1980), which calculates the photosynthetic rate based on carbon fluxes under different limitations. For C_4 plants, it uses the Collatz model (Collatz et al., 1992). Both models assume that the gross photosynthetic rate is affected by the maximum rate of carboxylation and is limited by RuBP carboxylase (Rubisco) (A_c), light and electron transport (A_j), and PEP carboxylase enzyme in C_4 plants (A_p). The final gross photosynthetic rate "A" is calculated using the empirical curvature parameters (θ_{cj} and θ_{ip}), while the net photosynthetic rate A_n is the same as the gross rate (A) after the plant respiration rate (R_d) is subtracted. The system can be described succinctly as the following, with Equations 5 and 4 playing the roles of f_1 and f_2 in Equation 1, respectively, and the whole set of associated equations in Appendix A.

$$\mathrm{A_n} = \mathrm{A(C_i)} - \mathrm{R_d} \tag{4}$$

$$\mathrm{C_i} = \mathrm{C_a} - \mathrm{A_n P_{atm}} \frac{(1.4\mathrm{g_s} + 1.6\mathrm{g_b})}{(\mathrm{g_s} + \mathrm{g_b})} \tag{5}$$

Equation 5 is a single-variable nonlinear equation, with the intercellular leaf CO_2 pressure (C_i) as the unknown term to be solved, (serving as the x term in Equation 1). C_i is influenced by the CO_2 partial pressure at the leaf surface (C_a), the net photosynthetic rate (A_n), the atmospheric pressure (p_{atm}), the leaf stomatal conductance (g_s), and the leaf boundary layer conductance (g_b). Upon solving for C_i, we can further calculate A_n, which is the y term in equation 1. In the original FATES and CLM, the system of nonlinear equations were solved iteratively using fixed-point iteration (Oleson et al., 2013).

In order to train the physical equations and neural networks together using gradient descent, the above equations were implemented on differentiable platforms to support backpropagation. We developed two alternative implementations: PyTorch (Paszke et al., 2019) and Julia (Bezanson et al., 2012). The PyTorch version solves the coupled nonlinear equations using our own parallel implementation of Newton iteration, while the Julia version is implemented via a symbolic computer algebra system (CAS) and is compatible with a wide variety of nonlinear solvers (Gowda et al., 2022). In contrast to the previous fixed-point iteration used by FATES, our PyTorch Newton iteration solver can run on a graphical processing unit (GPU) in parallel for many sites. Newton's iteration features second-order convergence compared to the slower convergence of fixed-point iteration, while GPU parallelism represents orders of magnitude in computational savings compared to the original algorithm in FATES. The photosynthesis problem studied here has only one unknown (C_i) even though there are many other supporting equations, while we have tested other nonlinear systems with <10 unknowns. For higher-dimensional systems, alternative treatment for the adjoint equations may be needed to speed up the gradient calculation since the Jacobian needs to be inverted which can impact the efficiency. Altogether, we can train this model with the coupled neural networks

https://doi.org/10.5194/bg-2022-211
Preprint. Discussion started: 2 November 2022

for hundreds of data points in under 10 minutes (typically in 600 iterations) and could also train the model on 10,000 data points. For the Julia implementation, the symbolic toolbox ModelingToolKit.jl (Gowda et al., 2022; Ma et al., 2021) was employed to automatically generate the solution scheme as Julia code, and along with solvers from NonlinearSolve.jl, solve the system of equations in the forward problem. Presently, we have implemented the Julia version in serial mode only. Results presented in this paper were produced using the PyTorch version, although the computational results were the same with the Julia version.

2.3. The parameterization pipeline and model changes

We used multilayer perceptron (MLP) neural networks as the parameterization module g in Equation 3. The purpose of the MLPs is to estimate parameters θ, which are then fed into the photosynthesis module to obtain the net photosynthetic rate (A_n) (Appendix A). The MLPs were trained based on the loss function between the solved and observed values of A_n. As described in Equation 2, the whole workflow is hereafter referred to as the δ_{psn} model ("delta-photosynthesis model") (the greek letter δ is selected because the model is programmatically differentiable and δ is often associated with gradients). There may be multiple MLPs to estimate different parameters in θ, each with different inputs of either continuous or categorical data, and they can all be trained together. Figure 1a below shows the framework for different parameter estimation experiments. We carried out both single-parameter and dual-parameter (learning two parameters simultaneously) experiments for both synthetic and real case datasets.

We chose to estimate one or both of two specific parameters in our experiments. The first one is the plant maximum carboxylation rate at 25°C ($V_{c,max25}$), which is normally formulated as a PFT-dependent parameter. Although $V_{c,max25}$ is hypothesized to be PFT-dependent, recent studies have shown that the parameter can vary in space and time and by different species in the PFT as well (Ali et al., 2015; Chen et al., 2022; Qian et al., 2019). Estimating $V_{c,max25}$ is not a trivial matter due to its high variability and sensitivity to different factors such as drought, leading some studies to suggest a substitute for it. For example, Croft et al. (2017) suggested using the leaf chlorophyll content as a direct proxy for $V_{c,max25}$. Nevertheless, considering this is an initial study applying dPL, we followed the convention and parameterized it based on PFT:

$$V_{c,max25} = NN_V(PFT) \quad (6)$$

where PFT is the plant functional type category (in one-hot encoding format) and the neural network used for parameterization of $V_{c,max25}$ is referred to as NN_V hereafter.

The second parameterization is $B = NN_B(R)$ (R includes %sand, %clay, fraction of organic matter, and PFT; see equation 11), and the neural network used for parameterization of B is referred to as NN_B. B is the parameter defined by Clapp & Hornberger (1978) and it influences the soil water stress function (β_t, where the subscript t indicates transpiration). β_t is called "btran" in the CLM code and it reflects the impacts of soil wetness on stomatal conductance and ranges from zero

https://doi.org/10.5194/bg-2022-211
Preprint. Discussion started: 2 November 2022

(extreme dry conditions causing stomata closure) to one (wet conditions with stomata fully open). B is purely a function of soil properties whereas β_t is affected by soil water potential and plant root distribution across different soil layers and is a PFT-dependent feature. B comes into play when calculating the soil water potential ψ_i using a power-law formulation:

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

where ψ_{sat} is the saturated soil matric potential and S is the soil wetness, both defined for a specific soil layer. Different soil attributes such as percentages of sand (*%sand*) and clay (*%clay*), fraction of organic matter (F_{om}), and soil moisture (θ_{liq}) are used in computing ψ_{sat}, S, and B (Appendix A). These equations will later be replaced by our NN-based parameterization scheme (NN_B, see equation 11) because they were originally empirical and may not be optimal at the global scale.

β_t can be calculated by aggregating the plant wilting factor (w) and plant root distribution (r) across different soil different layers based on the PFT (see equation 8) (Oleson et al., 2013).

$$\beta_t = \sum_i w_i r_i \quad (8)$$

The plant wilting factor (w_i), is mainly dependent on the soil water potential ψ_i and other PFT-dependent parameters such as the soil matric potentials for closed stomata ψ_c and open stomata ψ_o, which represent the soil water potentials when stomata are fully closed and fully open, respectively, as in equation (9). The factor w_i is also dependent on other factors like the temperature of the soil layer (T_i) relative to the freezing temperature (T_f), the volumetric liquid water (θ_{liq}) and ice (θ_{ice}) contents, and the volumetric water content at saturation (θ_{sat}).

$$w_i = \begin{cases} \frac{\Psi_c - \Psi_i}{\Psi_c - \Psi_o}\left[\frac{\theta_{sat,i} - \theta_{ice\,i}}{\theta_{sat,i}}\right] \leq 1 \; ; \; T_i > T_f - 2 \text{ and } \theta_{liq,i} > 0 \\ 0 \qquad\qquad\qquad\qquad\qquad\qquad ; \; T_i \leq T_f - 2 \text{ and } \theta_{liq,i} \leq 0 \end{cases} \quad (9)$$

https://doi.org/10.5194/bg-2022-211
Preprint. Discussion started: 2 November 2022

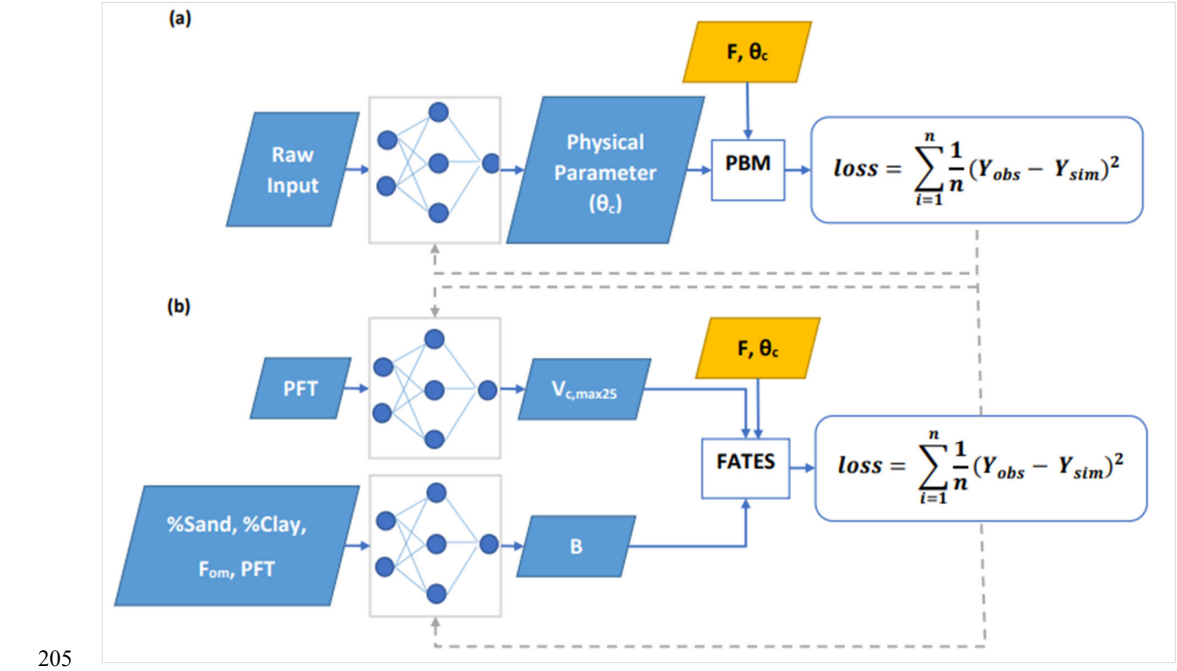


Figure 1. Diagram showing the differentiable parameter learning (dPL) framework which is a hybrid of neural networks and the photosynthesis module in the FATES ecosystem model written on a differentiable platform. (a) The generic workflow: Some raw information is mapped into physical parameters via a neural network. These parameters are sent into a process-based model, which then outputs variable Y that is compared with observations. Direct supervision for the physical parameters is not required -- we do not need ground truth for these parameters. The loss function is "global" in that it involves all training data points, rather than being computed site-by-site as done in traditional calibration. (b) The workflow for the computational example described in this work. We estimate either $V_{c,max25}$ or the parameter B using neural networks, or both of them at the same time. When they were not estimated from data, default values from the literature were used.

2.3.1 Model changes

Two simple changes were applied in computing the soil water stress function (β_t). In place of equation 7, the soil matric potential is calculated using the soil matric potential for closed stomata as:

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

where B is estimated using NN_B as:

$$B = NN_B(\%sand, \%clay, PFT, F_{om}) \quad (11)$$

The default equations in the Community Land model V4.5 (CLM4.5) for computations of B (Appendix A) show that the parameter *B* depends on two attributes, *%clay* and F_{om}, which is why they were used in NN_B. To account for the dependence

https://doi.org/10.5194/bg-2022-211
Preprint. Discussion started: 2 November 2022

of ψ_{sat} on *%sand* (Appendix A) and its replacement by ψ_o (see equations 7 and 10), *%sand* was also added to NN_B. Since in NN_B, we use quantitative inputs (*%sand*, *%clay*, F_{om}) along with categorical inputs (PFT), we used an one-hot embedding layer in PyTorch, which translates each category to a vector of quantitative variables. This categorical data can then easily be combined with other quantitative inputs we provide to our neural network.

2.3.2 Case 1: Synthetic data

In our synthetic experiments, we assumed values for some parameters to generate synthetic photosynthesis rates as the training data. Then, we estimated those parameters with NNs while keeping other components unmodified. These experiments were intended to verify the plausibility and efficiency of the differentiable learning framework, and the identifiability of parameters.

In the first synthetic case, "vcmax-only", the δ_{psn} framework was tested for retrieving a single PFT-dependent parameter, $V_{c,max25}$, using NN_V. We used the suggested values for $V_{c,max25}$ from CLM4.5 for different PFTs to calculate the synthetic net photosynthetic rates.

In the second synthetic case, "$V_{c,max}$ - B", we tested retrieving both $V_{c,max25}$ and B, the latter of which varies spatially and temporally. To generate the synthetic data, we assumed B = 0.1 * F_{om} + 0.45 * (%sand + %clay), and then the soil matric potential (ψ_i) was calculated using equation 10. The plant wilting factor (w_i) and the soil water stress function (β_t) were calculated using the default equations 9 and 8 respectively. For simplicity, the computations of B, ψ_i, w_i, β_t were performed for the topsoil layer only. To retrieve B, we used NN_B (see equation 11) but excluded the PFT term.

For both synthetic runs "vcmax-only" and "vcmax-B", the MLP models were trained concurrently for all PFTs with several data points for each PFT. Moreover, white noise was added to the synthetic values of A_n with a standard deviation of 5% of the mean value.

2.3.3. Case 2: Real data

The model passing the test of the synthetic case was then applied to a real dataset (Lin et al., 2015) using observation data. This tested whether the model, learning from this dataset for many of the PFTs, could find parameters to better describe photosynthesis data than the literature values. We do not know the ground truth in this case, so we tested multiple formulations to understand the impacts of allowing more or less flexibility in the estimation and the role of each parameter.

We tested several formulations to estimate either one ($V_{c,max25}$) or two parameters ($V_{c,max25}$ and B) at a time. In essence, we compared allowing either one or two of the parameters to be estimated vs. using the default formulation or values from the

https://doi.org/10.5194/bg-2022-211
Preprint. Discussion started: 2 November 2022

original model. For $V_{c,max25}$, the default values were those defined in CLM4.5, while for β_t, the default equations (Appendix A) were used to obtain its values. Altogether, we trained the following models:

V_{def} + B_{def} : in this case, $V_{c,max25}$ took the default values from CLM4.5 and B was calculated using the default equations (Appendix A). This was used as a reference case.

V_{def} + B : in this formulation, the default $V_{c,max25}$ values from CLM4.5 were used while B was estimated using NN_B.

V + B_{def} : in this formulation, $V_{c,max25}$ was estimated using NN_V, while B was calculated using the default equations (Appendix A).

V + B : in this formulation, we employed both NN_V and NN_B. They were trained concurrently to see if this interfered with parameter retrieval.

Just as in the synthetic case, the MLPs were shared between all sites. All sites were used to calculate one loss function as in typical machine learning tasks, with the hope of ensuring the wide applicability of the MLPs and leveraging the synergy between all sites (Fang et al., 2022). In this way, we also hope to identify parameters that can generalize well in space and be applicable at large scales.

We then compared the values of $V_{c,max25}$ learned by the V+B model, trained on all data points, against values of $V_{c,max25}$ in other data sources (Kattge et al., 2020; Rogers, 2014), which highlights the variability of these parameters. The TRY database (Kattge et al., 2020) has $V_{c,max25}$ values defined for several species which can be aggregated to get unique values for each PFT (Table 3). Moreover, we compared our $V_{c,max25}$ values to the ones used in different earth system models (Rogers, 2014) for various PFTs. The comparison can first check if the inversely determined values are on the same order of magnitude as previously-employed values and are physically possible. We also expect our values for different PFTs to be at least partially correlated with the ones used in the literature, as they were meant to represent the same physical quantity. A complete disagreement or a different order of magnitude would suggest that our values may be not physical. However, partial discrepancies would also highlight the knowledge gaps.

2.4. Statistical metrics

In order to evaluate different experiments and see the sensitivity of the results to changing different parameters, we chose four different metrics as shown in table 1, below. The four metrics were root-mean-square error (RMSE), bias, Pearson's correlation coefficient (COR), and Nash-Sutcliffe Efficiency (NSE). Both RMSE and bias measure how far the model simulations are from the observations; however, RMSE is the standard deviation of all errors while bias is calculated as the

https://doi.org/10.5194/bg-2022-211
Preprint. Discussion started: 2 November 2022

285 average. COR measures the linear relationship between both the simulations and the observations, ranging between -1 and 1. NSE measures the relative magnitude of the residual variance relative to the observed data variance (Nash and Sutcliffe, 1970), and has a perfect score of 1. Table 1 below shows the formulations of the four metrics and their possible ranges.

Table 1. Performance metrics used for evaluation and their possible ranges

Metric	Formula	Range
COR	$\frac{\sum_{i=1}^{n}(OBS - \underline{OBS})(SIM_i - \underline{SIM})}{\sigma_{OBS}\sigma_{SIM}}$	[-1 , 1]
RMSE	$\sqrt{\frac{\sum_{i=1}^{n}(SIM_i - OBS_i)^2}{n}}$	[0, ∞]
BIAS	$\frac{\sum_{i=1}^{n}(SIM_i - OBS_i)^2}{n}$	[-∞,∞]
NSE	$1 - \frac{\sum_{i=1}^{n}\lvert SIM_i - OBS_i\rvert}{\sum_{i=1}^{n}\lvert OBS_i - \underline{OBS}\rvert}$	[-∞,1]

290

2.5. Input and observation datasets

2.5.1 Forcing and Photosynthesis rates:

We used the ERA5 Reanalysis dataset (Copernicus Climate Change Service (C3S), 2017), which provides hourly estimates of soil moisture at different soil levels. The soil moisture contributes to computing β_t (see Appendix A), where the soil
295 wetness S depends on both the soil moisture and the saturated soil moisture.

We used data from the leaf gas exchange database (Knauer et al., 2018; Lin et al., 2015) which is a global database of stomatal conductance measurements. It incorporates data from several sites around the world in Australia, Europe, USA, and Asia (Figure 2). 43 sites were chosen because they had complete data (with dates of measurements) available. We refer to
300 this dataset as Lin15 throughout the rest of this work. Lin15 also contained meteorological forcing variables, including air temperature, atmospheric pressure, relative humidity, and radiation. Moreover, we used ERA5 to fill in for any missing forcing variables in Lin15.

https://doi.org/10.5194/bg-2022-211
Preprint. Discussion started: 2 November 2022

2.5.2 Static attributes:

For β_t calculations, we used data from Hengl & Wheeler (2018) for the soil organic carbon content at different soil depths, while data for sand and clay percent was obtained from Hengl (2018). Both are global datasets available at 250 m resolution at 6 different soil depths (0, 10, 30, 60, 100, and 200 cm).

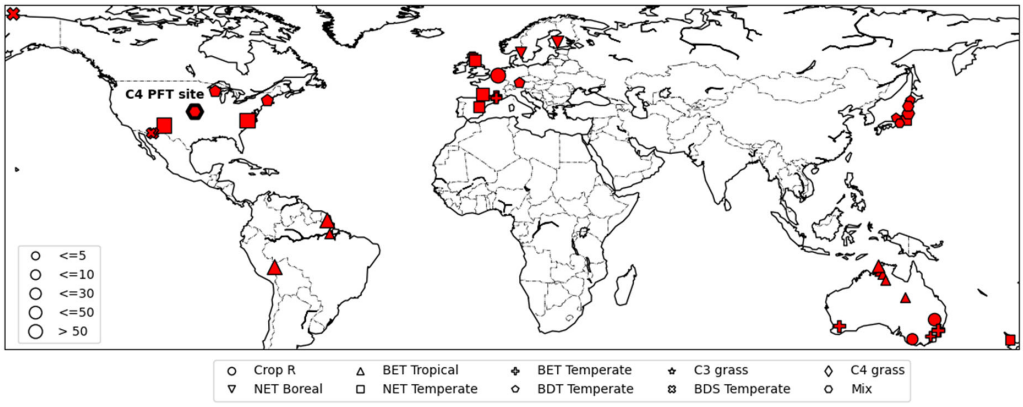


Figure 2. Map of sites available from the leaf gas exchange database (Lin et al., 2015). Different symbols represent different plant functional types. The C4 site is highlighted by a thick-bordered hexagon. The marker sizes represent the quantity of data available for each site. (map based on matplotlib basemap, Jeffrey Whitaker)

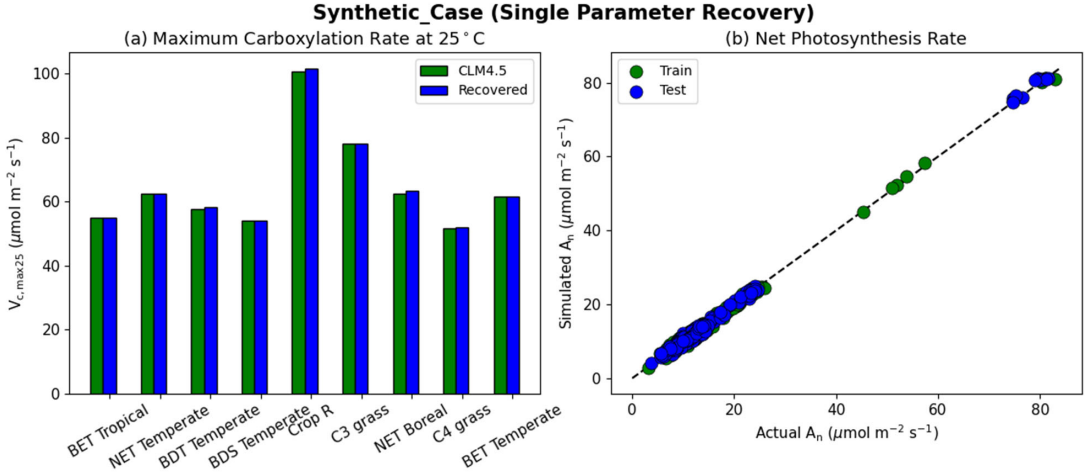
3 Results

3.1. Results for synthetic case

The results of the synthetic experiments showed that our workflow successfully recovered the parameters in either the one-parameter (vcmax-only) (Figure 3) or two-parameter ("vcmax-B") cases (Figure 4). In the one-parameter "vcmax-only" case, the recovered parameters agreed with the assumed values almost completely for each PFT (Figure 3a). The model was able to capture the variability in the values of $V_{c,max25}$ for different PFTs, where the values ranged from 100.7 μmol m^{-1} s^{-1} for the rainfed crop (defined as Crop R in CLM4.5) to around 50 μmol m^{-1} s^{-1} for C_4 grass (Figure 3a). Moreover, we found nearly complete agreement between the synthetic and recovered net photosynthesis rates (A_n) (Figure 3b). This case demonstrated that the dPL framework and the posited formulation $V_{c,max25} = NN_V(PFT)$ were functional, but could not show the effects of parameter interactions. Furthermore, we found a similarly near-complete recovery for $V_{c,max25}$ with the dual-parameter case (Figure 4a) and a near-complete reproduction of simulated photosynthesis (Figure 4d). However, we noticed a negligible amount of scattering with β_t (Figure 4c), and to a larger extent, with B (Figure 4b). For all experiments, we verified that the training and test periods were highly consistent (between green and blue points in the scattered plots).

https://doi.org/10.5194/bg-2022-211
Preprint. Discussion started: 2 November 2022

325 The results indicate that the problem formulation allows for sufficient sensitivity of the net photosynthesis rate with respect to PFT-specific $V_{c,max25}$ and the soil water constraint. In addition, $V_{c,max25}$ and B influence the photosynthesis rate in different ways so that, along with a large dataset with different combinations of moisture conditions and PFTs, they can be identified simultaneously. This forms the basis of the next stage of the work. The soil moisture parameter identifiability was slightly weakened compared to $V_{c,max25}$ because there were more equations involved between B and A_n, and some of them had
330 parameters in the exponential operators, e.g., $\psi_i = \psi_o * S^{-B}$. Mathematically, this curve can be flat and the gradients can be small in some ranges of S. Mechanistically, A_n can have reduced sensitivity to B under some conditions. Therefore, we do not expect soil properties to be fully identifiable from photosynthesis data, but the general pattern may still be learnable.



335 **Figure 3. Single parameter recovery for synthetic data. (a) Comparison of modeled parameter values to literature values by plant functional type (PFT). (b) Actual and modeled net photosynthesis rates for training and testing periods (dashed line indicates the ideal 1:1 relationship).**

https://doi.org/10.5194/bg-2022-211
Preprint. Discussion started: 2 November 2022

340

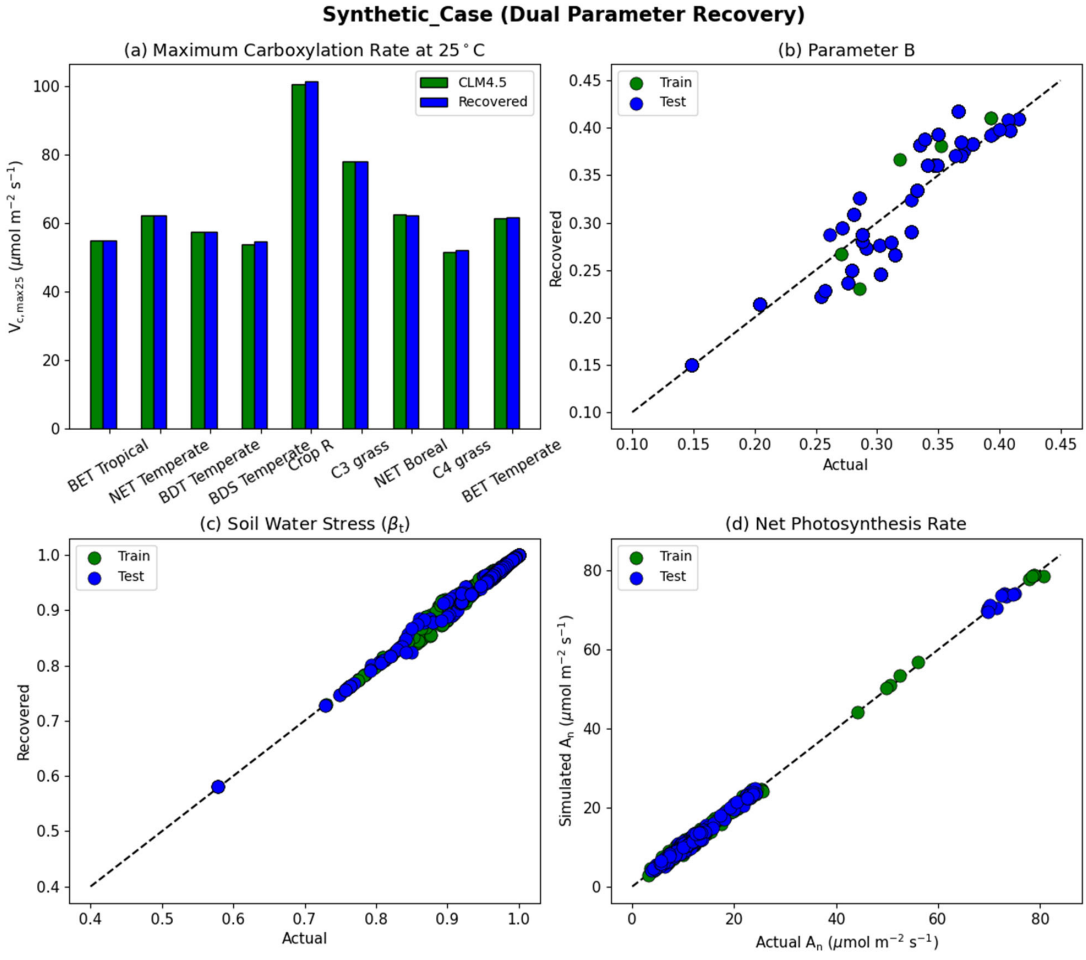


Figure 4. Dual parameter recovery for synthetic data. (a) Comparison of modeled parameter values to literature values by plant functional type (PFT) estimated using NN_V. (b) Actual and modeled parameter values for B, estimated using NN_B (dashed line indicates the ideal 1:1 relationship). (c) Actual and modeled parameter values for βt, calculated using equations 8, 9, and 10 for the topsoil layer. (d) Actual and modeled net photosynthesis rates for training and testing periods.

345

3.2. Real dataset

3.2.1. Comparisons between candidate formulations

The V+B model exhibited obvious advantages over the default FATES model, as well as the models that learned only one of the parameters (Table 2). The default CLM4.5 parameters (V_{def}+B_{def}) led to a low correlation (0.565), a large bias (1.476

https://doi.org/10.5194/bg-2022-211
Preprint. Discussion started: 2 November 2022

μmol m^{-2} s^{-1}) and nearly zero NSE (0.041) (Table 2a). In particular, the default values appeared to cause a large over-estimation of the net photosynthetic rate (A_n) for the high-photosynthesis data points (especially C_4), which are visible in Figure 5a. After allowing B to be learned (V_{def}+B), the correlation for testing remained the same, while the bias remained high (1.754 μmol m^{-2} s^{-1}). It seems that the treatment of water stress alone did not address the bias. On the other hand, when we only allowed $V_{c,max25}$ to be estimated (V+B_{def}), the bias was slightly reduced and the test NSE increased to (0.229). Finally, if we allowed both parameters to be estimated (V+B), a decent correlation was obtained (0.748), the bias was the smallest value yet (0.347 μmol m^{-2} s^{-1}) and the test NSE was 0.532, which means the model explained about half of the variance in the observed photosynthesis rate. The remaining error might be attributable to other untuned parameters as well structural deficiencies of the current model, which can be further improved in the future using the differentiable modeling paradigm.

A similar behavior in the performance metrics was observed for the random holdout test (Table 2b). In this test, instead of splitting the dataset in chronological order with earlier dates set for training and the rest reserved for testing, we randomly chose the data points for training and testing. Based on the results shown in table 2b, the random splitting decreased to a great extent the disparity in the metrics' values between training and testing. However, contrary to the temporal holdout test, we found a slight improvement in COR (0.619) and NSE (0.171) when B was learned (V_{def}+B), while a much higher boost was found in metrics when $V_{c,max25}$ was learned (V+B_{def}). This shows the higher impact of learning $V_{c,max25}$ on the simulation of A_n, where the COR and NSE increased to 0.695 and 0.442, respectively, while the bias decreased to -0.374. Similar to the temporal holdout test, the V+B model showed the best metrics in comparison to other models with the lowest RMSE (4.480) and bias (0.177) values, and the highest COR (0.758) and NSE (0.566) values.

Consistent with the observations of the synthetic experiments, $V_{c,max25}$ and B impacted A_n in different ways. When $V_{c,max25}$ was not adjusted, the photosynthesis rates simulated for a number of sites in the high-A_n range (most of them C_4 plants) had some substantial overestimation, regardless of whether B had learned or default values (Figure 5a). It was only after we also learned $V_{c,max25}$ that these high biases were reduced (Figure 5b). Hence, apparently, the learning reduced the $V_{c,max25}$ for these sites compared to the default values. In contrast, learning B mainly corrected the low bias for low-A_n data points (both C_3 and C_4 plants) (Figure 5b). A group of sites with A_n <2 μmol $m^{-2}s^{-1}$ have been corrected upward (from yellow to green, bottom points in Figure 5b), due to a correction in the soil parameter B. Apparently, the original parameters overestimated the water stress for these sites. Our results suggest the adjustments to both parameters improved the results, but $V_{c,max25}$ was more impactful, especially in addressing the bias.

Our test showed that the framework is moderately impacted by long-term nonstationarity, as the temporal test had worse metrics than the random test (comparing Table 2b with 2a). The absolute value of the bias increased from 0.177 in the random test to 0.347 in the temporal test. This suggests the current model (and perhaps training data) still has some

https://doi.org/10.5194/bg-2022-211
Preprint. Discussion started: 2 November 2022

limitations with representing long-term changes. Possible reasons may include CO2 fertilization and its impact on water use efficiency or differences in the state of plants, as this factor is not included in our present parameterization. In the future, these issues could be addressed by assembling a more long-term training dataset (the Lin15 dataset has data ranging from 1991 to 2013), as well as improving the parameterization and physics of the model.

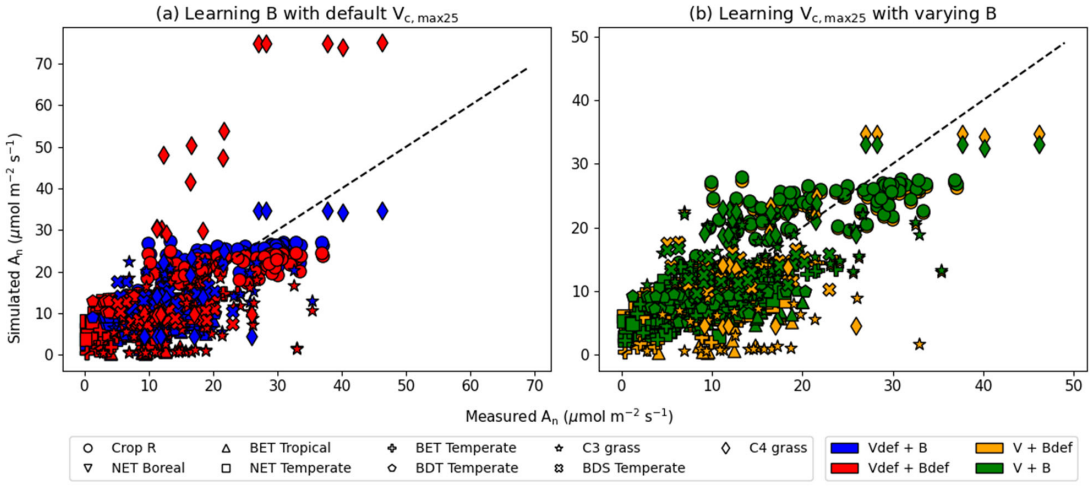


Figure 5. Comparisons of photosynthesis model calibration: mean estimated value of default parameters vs. mean estimated value of best learned parameters vs. observed value for different candidate models. (a) Impact of learning B with default $V_{c,max25}$. (b) Impact of learning $V_{c,max25}$ with varying B. The colors represent the results from the four different models, the shapes indicate the plant functional type (PFT) groups, and the dotted line in each panel indicates the ideal 1:1 relationship.

https://doi.org/10.5194/bg-2022-211
Preprint. Discussion started: 2 November 2022

395 **Table 2. Performance metrics for the candidate models for the Lin15 dataset. In all the following, subscript $_{def}$ indicates the default parameter value from CLM4.5 was used, while a parameter lacking $_{def}$ means the parameter was estimated as an output from a neural network (in all cases, V indicates that $V_{c,max25}$ was estimated as a function of PFT using NN_V and B indicates estimation using NN_B). Panel (a) is for the temporal holdout test where 70% of data points were used for training and 30% were reserved for testing; panel (b) is for the test where training and testing data groups were randomly selected.**

(a) Temporal holdout test for the following system

Runs	Corr		RMSE (µmol m^{-2} s^{-1})		Bias (µmol m^{-2} s^{-1})		NSE	
	Train	Test	Train	Test	Train	Test	Train	Test
$V_{def}+B_{def}$	0.565		6.780		1.476		0.041	
$V_{def}+B_{def}$ **	0.592		5.488		1.034		0.318	
$V_{def}+B$	0.678	0.547	5.887	6.730	1.353	1.754	0.321	-0.084
$V+B_{def}$	0.769	0.593	4.595	5.677	-0.129	-1.368	0.587	0.229
V+B	0.800	0.748	4.299	4.421	0.037	0.347	0.638	0.532
V+B **	0.774	0.768	4.269	4.198	0.056	0.092	0.597	0.581

400 ** *refers to using C3_only plants in dataset*

(b) Random holdout test for the following system

Runs	Corr		RMSE (µmol m^{-2} s^{-1})		Bias (µmol m^{-2} s^{-1})		NSE	
	Train	Test	Train	Test	Train	Test	Train	Test
$v_{def}+B_{def}$	0.565		6.780		1.476		0.041	
$v_{def}+B$	0.644	0.619	6.156	6.185	1.349	1.424	0.220	0.171
$v+B_{def}$	0.722	0.695	4.928	5.073	-0.389	-0.374	0.500	0.442
v+B	0.792	0.758	4.262	4.480	0.070	0.177	0.626	0.566

https://doi.org/10.5194/bg-2022-211
Preprint. Discussion started: 2 November 2022

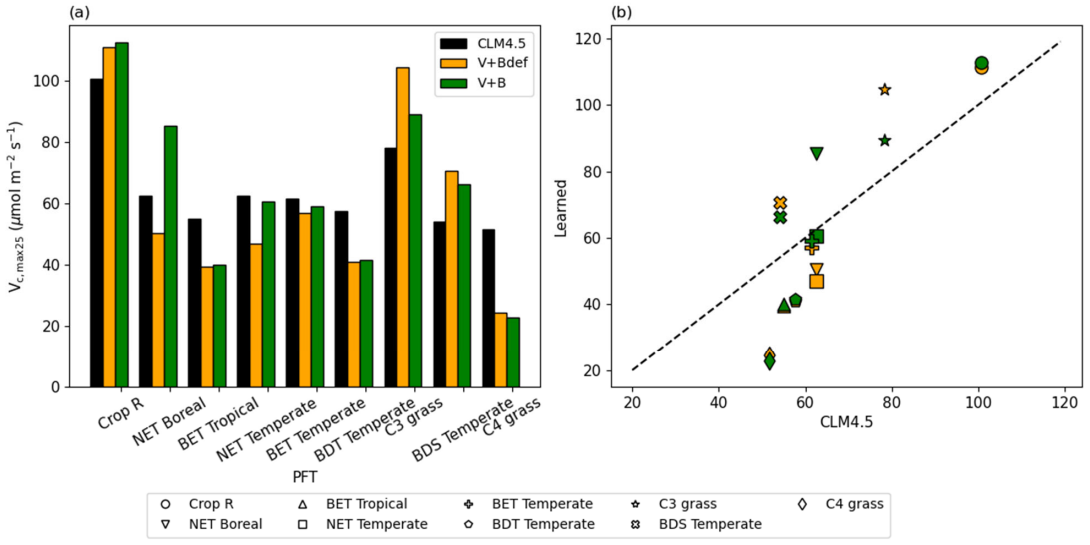


Figure 6. Parameter recovery for real data. (a) Comparison of modeled parameter values to literature values by plant functional type (PFT). (b) Actual and modeled $V_{c,max25}$ values plotted by PFT (dashed line indicates the 1:1 ideal relationship). In this figure, both "V+Bdef" and "V+B" models were trained using the whole dataset.

3.2.2. Recovered parameters

Even though we did not prescribe the values of $V_{c,max25}$, the training on the dataset converged to parameter values that were well correlated with, yet still substantially different from, the literature values (Figure 6). The default $V_{c,max25}$ values came from in-situ measurements at a limited number of sites, while our values came from learning from a large dataset (essentially an inversion process limited to the model structure). The fact that they agreed with each other in the main pattern suggests both have merit, and that the learning process captured fundamental physics. The upper half of Figure 6b saw a high correlation, but $V_{c,max25}$ values for the V+B model were uniformly higher than the CLM4.5 defaults, especially for the NET Boreal PFT. The correlation was lower toward the lower half of Figure 6b (where $V_{c,max25}$ from CLM4.5 was lower than 65 μmol m^{-2} s^{-1}) -- the learned $V_{c,max25}$ had a larger variability. In particular, the learned $V_{c,max25}$ (V+B) for C_4 grass is much lower than the default, which could be attributed to species-level variability and the fact that the dataset contains very limited sites with C_4 plants. Hence, we do not argue that the values learned here would be applicable globally to other C_4 grasses. For these cases, the influence of B or B_{def} were mostly small ($V_{c,max25}$ from V+B and V+B_{def} models were mostly similar) and thus parameter interaction from soil water stress was not significant. It seems the inter-PFT variability in $V_{c,max25}$ was previously under-represented by the CLM4.5 default parameter values (C_4 grass, BET tropical, BDS temperate, BET temperature), and the learning process used here expanded the variability. The overall results showcase the ability of the algorithm to adapt to data.

https://doi.org/10.5194/bg-2022-211
Preprint. Discussion started: 2 November 2022

425

In our interpretation, the learned values represent a more "precisely tuned" version of the literature $V_{c,max25}$ values, with the interference from soil water stress disentangled. The magnitude and ranking for PFTs remained similar to the literature values, but the results were improved in different ways for different PFTs. The V+B model obtained lower $V_{c,max25}$ for C_4 grasses, addressing the significant overestimation bias for these sites, which we noted in Figure 5a. Due to their different
430 photosynthesis pathway, C_4 plants have the lowest learned $V_{c,max25}$, but overall the highest net photosynthesis rates, which were not heavily influenced by the choice of the B parameter. For C_3 grasses, V+B only slightly increased $V_{c,max25}$ compared to the default CLM4.5 values, which addressed the low bias noticeable in Figure 5b. The default soil parameterization for C_3 grass sites seemed somewhat deficient as soil water stress accounted for the other parts of variance in net photosynthesis, as demonstrated by the comparison between V+B and V+B_{def} models in Figures 5b and 6b for C_3 grass.

435

We compared our learned $V_{c,max25}$ values (Table 3) with values from other earth system models (ESMs) and with some observatory values in the TRY database (Kattge et al., 2020; Rogers, 2014). The learned $V_{c,max25}$ values are higher than those of the TRY database for most PFT classes except for BDS Temperate and BDT Temperate; however, they are within the range of values used in other ESMs except for relatively higher estimations for NET Boreal and C_3 grasses. On the scale of
440 ESMs, several values for $V_{c,max25}$ are adopted by those models. For instance $V_{c,max25}$ for C_4 grass is taken as 25 and 20 (μmol m^{-2} s^{-1}) in AVIM and BETHY models, respectively (Table 3). These values agree with the learned $V_{c,max25}$ by the V+B model of 22.86 (μmol m^{-2} s^{-1}), whereas much higher values were found to be adopted for C_4 grasses with 60 (μmol m^{-2} s^{-1}) in the Biogeochemical cycles model "BiomeBGC" as reported in Rogers (2014), and 51.6 (μmol m^{-2} s^{-1}) in CLM4.5. $V_{c,max25}$ from the V+B model and TRY database are similar for BET Tropical and BDT Temperate. For BDS Temperate, the learned
445 $V_{c,max25}$ is lower than that in TRY by ~20 (μmol m^{-2} s^{-1}), but similar values were used by BETHY and lower values were used by AVIM. For NET Boreal, BET Temperate, Crop R, and NET Temperate, the learned $V_{c,max25}$ values were all ~20 (μmol m^{-2} s^{-1}) higher than those of the TRY database, but (except for NET Boreal) similar values have been used by AVIM or BETHY. Both the learned (V+B) and the observed (TRY database) $V_{c,max25}$ values show a similar pattern with the lowest $V_{c,max25}$ for BET Tropical and a high value assigned for Crop R.

450

https://doi.org/10.5194/bg-2022-211
Preprint. Discussion started: 2 November 2022

Biogeosciences Discussions Open Access EGU

Table 3. $V_{c,max25}$ simulated by V+B model versus observed values from the TRY database (with partial overlap in species with the Lin15 dataset -- the percentage of overlap is provided in the table), and used in different earth system models such as CLM4.5, Atmosphere-Vegetation Interaction Model "AVIM", and the Biosphere Energy Transfer Hydrology scheme "BETHY".

PFT	CLM4.5	AVIM	BETHY	V+B (ours)	TRY (mean / % species overlap)	TRY (std)
BET Temperate	61.5	68	58	59.04	39.54 / 31.3%	4.05
BET Tropical	55	64	28/36	40.07	33.14 / 86.5%	14.09
BDT Temperate	57.7	60	54	41.63	50.27 / 50.0%	21.62
BDS Temperate	54	52	65	66.22	87.61/ 58.3%	11.77
NET Temperate	62.5	60	58	60.64	44.33 / 50.0%	7.13
NET Boreal	62.6	58	58	85.30	62.90 / 100.0%	22.53
C_3 grass	78.2	55/40	71	89.26	-	-
C_4 grass	51.6	25	20	22.86 (limited data points)	-	-
Crop R	100.7	55	90	112.61	84.20 / 60.0%	2.19

https://doi.org/10.5194/bg-2022-211
Preprint. Discussion started: 2 November 2022

4 Discussion

As an initial exploration of the potential of the emerging differentiable computing paradigm for application to ecosystem modeling, our work showed promise but also had many limitations, as the goal was not to produce the best-performing photosynthetic model. We restricted our parameter sets to be dependent on PFT, whereas it is known that within-PFT variation can be significant and parameters could also be determined on the trait level as well as by multiple environmental factors. Our model did not consider the effects of memory, e.g., rainfall in previous days, and the states of the vegetation, e.g., carbon stored in the canopy or carbon: nitrogen ratios in the canopy. The soil moisture comes from the ERA5 dataset, which, based on comparisons to in-situ data, would be outperformed by ML-based soil moisture predictions (Fang et al., 2017; Liu et al., 2022a, b), but we used it due to its global seamless coverage and availability for multiple soil depths. This work also only modified the parameterization scheme and did not learn model structures. Recently, development in differentiable hydrologic models allows learning parts of the model using neural networks (Feng et al., 2022a, b). In summary, we believe there is still lots of room for improving model quality, but at some point we may run into the limits of measurements (aleatoric uncertainty) or data availability (epistemic uncertainty) (Hüllermeier and Waegeman, 2021). Future effort can harness deep networks to establish reference levels as a measure of the data uncertainty (Feng et al., 2022a).

This work appears to be the first evaluation of the Lin15 dataset, and, as such, it establishes a reference level to which future studies can compare. The current dataset may still have limitations in that the number of sites for C_4 plants is small and does not allow ample testing. Some geoscientific domains have well-known benchmark datasets, e.g., the CAMELS dataset in hydrology (Feng et al., 2020). Having such a common (and hopefully large) benchmark dataset allows better model structures to be rapidly discovered and is highly beneficial to the growth of the community (Shen et al., 2018). Related to the limits of measurement errors discussed above, multiple deep-learning-based studies have explored the approximate limit of data error (or best achievable model) of CAMELS and that knowledge has been appreciated by the community (Feng et al., 2021). Moreover, deep learning methods benefit from data synergy effects (Fang et al., 2022), where more sites and more diverse data lead to a more robust model and better performance for each site.

5 Conclusions

In this study, we proposed a novel differentiable ecosystem modeling framework that uses neural networks as a parameterization scheme to support a process-based model (FATES). Training coupled neural networks was not previously possible without differentiable programming, and it allows us to approximate complex, *a priori* unknown mapping relationships between plant functional types, landscape characteristics, and physical parameters. The photosynthesis module was treated as a system of nonlinear equations, and, like other such systems, could be solved efficiently and in a massively

https://doi.org/10.5194/bg-2022-211
Preprint. Discussion started: 2 November 2022

parallel fashion on graphical processing units (GPUs) by our differentiable framework. $V_{c,max25}$ and a soil water parameter could be simultaneously identified in our synthetic experiments, because they played different roles in the model.

Compared to purely data-driven machine learning approaches, the differentiable programming framework provides physically meaningful variables and can be used to learn relationships from big data. Via training on a global dataset, we found $V_{c,max25}$ values for global sites that correlate with the values in the literature, but produce more accurate net photosynthesis rates. It is noteworthy that these values were identified without any supervision from experts other than the preparation of the training dataset and the model. We conclude that $V_{c,max25}$ has a larger impact on photosynthesis than the soil water stress parameter, but both can be useful in tuning model responses and their default values were not optimal. Not only is this method able to improve simulation quality and provide model parameterization, it can allow us to modify model structure and ask questions regarding unclear parts of the model in the future. There is significant room for this framework to improve and expand to other ecosystem modeling application.

https://doi.org/10.5194/bg-2022-211
Preprint. Discussion started: 2 November 2022

6 Appendices

Appendix A

The System of nonlinear equations

The FATES photosynthesis module is based on the classical Farquhar model for C_3 plants (Farquhar et al., 1980), which calculates the photosynthetic rate based on carbon fluxes under different limitations. For C_4 plants, it uses the Collatz model (Collatz et al., 1992). Both models assume that the gross photosynthetic rate is affected by the maximum rate of carboxylation and is limited by RuBP carboxylase (Rubisco) (A_c, equation 1), light and electron transport (A_j, equation 2), and PEP carboxylase enzyme in C_4 plants (A_p, equation 3). A_c, A_j, and A_p are calculated as:

$$A_c = \begin{cases} \dfrac{V_{c,max} \times (C_i - \Gamma_*)}{C_i + K_c\left(1 + \dfrac{K_o}{O_i}\right)}, & \text{for C3 plants} \\ V_{cmax}, & \text{for C4 plants} \end{cases} \quad \text{(A1)}$$

$$A_j = \begin{cases} \dfrac{J_x \times (C_i - \Gamma_*)}{4C_i + 8\Gamma_*}, & \text{for C3 plants} \\ \alpha(4.6\varphi), & \text{for C4 plants} \end{cases} \quad \text{(A2)}$$

$$A_P = \left\{ K_p \frac{C_i}{P_{atm}}, \quad \text{for C4 plants} \right\} \quad \text{(A3)}$$

where $V_{c,max}$ is the maximum carboxylation rate, C_i is the intercellular leaf CO_2 pressure, Γ^* is the CO_2 compensation point, K_c and K_o are the Michaelis-Menten constants, O_i is the O_2 partial pressure, J_x is the electron transport rate, α is the quantum efficiency (0.05 mol CO_2 mol^{-1} photon), φ is the photosynthetically active radiation, K_p is the initial slope of C_4 CO_2 response curve, and p_{atm} is the atmospheric pressure.

The gross photosynthetic rate (A) is then calculated by solving for the minimum root of the quadratic equations:

$$\theta_{cj}A_i^2 - (A_c + A_j)A_i + A_cA_j = 0 \quad \text{(A4)}$$

$$\theta_{ip}A^2 - (A_i + A_p)A + A_iA_p = 0 \quad \text{(A5)}$$

where A_i is an intermediate co-limited photosynthetic rate calculated using the empirical curvature parameter (θ_{cj}). Using A_i *and* A_p, and the empirical curvature parameter (θ_{ip}), the gross rate is given by the smaller root of equation 5. To get the net photosynthetic rate (A_n), the plant respiration (R_d) is subtracted from the gross rate (A) as the following:

https://doi.org/10.5194/bg-2022-211
Preprint. Discussion started: 2 November 2022

$$A_n = A - R_d \quad \text{(A6)}$$

Then using A_n, the CO_2 partial pressure at the leaf surface (C_s) is calculated as:

$$C_s = \max\left(1.0e-6, C_a - \frac{1.4\,P_{atm}A_n}{g_b}\right) \quad \text{(A7)}$$

where C_a is CO_2 partial pressure near the leaf surface and g_b is the leaf boundary layer conductance. The stomatal conductance (g_s) is then given by the maximum root of the quadratic equation:

$$ag_s^2 + bg_s + c = 0 \quad \text{(A8)}$$

where a, b, and c are functions in some PFT-dependent parameters (the intercept and the slope of medlyn stomatal conductance model (Medlyn et al., 2011)), A_n, C_s, p_{atm}, g_b, and the vapor pressure deficit (D). Finally, C_i, is related to An using C_a, p_{atm}, g_s, and g_b as the following:

$$C_i = C_a - A_n P_{atm}\frac{(1.4g_s + 1.6g_b)}{(g_s + g_b)} \quad \text{(A9)}$$

Both $V_{c,max25}$ and β_t affect the plant-specific photosynthetic fluxes, despite not directly appearing in equations (1-9). They, however, indirectly affect some parameters where the maximum electron transport rate (J_{max25}), the initial slope for C_4 CO_2 response curve (K_{p25}), and the plant respiration rate (R_{d25}) at 25 °C depend on $V_{c,max25}$ as:

$$J_{max25}, K_{p25}, R_{d25} = \emptyset_1(V_{c,max25}) \quad \text{(A10)}$$

Then, the biophysical rates $V_{c,max}$, J_{max}, and K_p, at the vegetation temperature (T_v) are calculated using $V_{c,max25}$, J_{max25}, K_{p25}, β_t, and φ as:

$$V_{c,max}, J_{max}, K_p = \emptyset_2(V_{c,max25}, J_{max25}, K_{p25}, \beta_t, \varphi, T_v) \quad \text{(A11)}$$

Similarly, the plant respiration (R_d) at T_v is calculated using R_{d25} as:

$$R_d = \emptyset_3(R_{d25}, T_v) \quad \text{(A12)}$$

$V_{c,max}$, K_p, and R_d directly appear in equations (1-9), while J_{max}, doesn't directly affect those equations but the J_x term. J_x is given by the minimum root of the following quadratic equation:

$$\theta_{PSII}J_x^2 - (I_{PSII} + J_{max})J_x + I_{PSII}J_{max} = 0 \quad \text{(A13)}$$

where θ_{PSII} is an empirical curvature for the electron transport rate and I_{PSII} is the light utilized in electron transport.

Computations of btran (β_t) in CLM4.5

β_t is calculated by aggregating the plant wilting factor (w) and plant root distribution (r) across different soil different layers.

https://doi.org/10.5194/bg-2022-211
Preprint. Discussion started: 2 November 2022

$$\beta_t = \sum_i w_i r_i \quad \text{(A14)}$$

The plant wilting factor (w_i) is mainly dependent on the soil water potential ψ_i and other PFT-dependent parameters such as the soil matric potentials for closed stomata ψ_c and open stomata ψ_o, which represent the soil water potentials when stomata are fully closed and fully open, respectively, as in equation (15). The factor w_i is also dependent on other factors like the temperature of the soil layer (T_i) relative to the freezing temperature (T_f), the volumetric liquid water (θ_{liq}) and ice (θ_{ice}) contents, and the volumetric water content at saturation (θ_{sat}).

$$w_i = \begin{cases} \frac{\Psi_c - \Psi_i}{\Psi_c - \Psi_o}\left[\frac{\theta_{sat,i} - \theta_{ice\,i}}{\theta_{sat,i}}\right] \le 1 \; ; \; T_i > T_f - 2 \text{ and } \theta_{liq,i} > 0 \\ 0 \qquad\qquad\qquad\qquad\qquad\qquad ; \; T_i \le T_f - 2 \text{ and } \theta_{liq,i} \le 0 \end{cases} \quad \text{(A15)}$$

The soil matric potential ψ_i is calculated using a power-law formulation:

$$\Psi_i = \Psi_{sat,i} \times S_i^{-B_i} \ge \Psi_c \quad \text{(A16)}$$

where ψ_{sat} is the saturated soil matric potential, S is the soil wetness, and B is the Clapp and Hornberger parameter, all defined for a specific soil layer (i). Different soil attributes such as percentages of sand (*%sand*) and clay (*%clay*), fraction of organic matter (F_{om}), and soil moisture (θ_{liq}) are used in computing ψ_{sat}, S, and B. ψ_{sat} is calculated as:

$$\Psi_{sat,i} = \left(1 - F_{om,i}\right) \times \Psi_{sat,min,i} + F_{om,i} \times \Psi_{sat,om} \quad \text{(A17)}$$

where $F_{om,i}$ is the fraction of organic matter, $\psi_{sat,om}$ is the saturated organic matter matric potential, and $\psi_{sat,min,i}$ is the saturated mineral soil matric potential calculated using (*%sand*) as:

$$\Psi_{sat,min,i} = -10.0 - 10^{1.88 - 0.0131 \times (\%sand)_i} \quad \text{(A18)}$$

The soil wetness (S) is calculated using the volumetric contents θ_{liq}, θ_{ice} , and θ_{sat} as:

$$S_i = \frac{\theta_{liq,i}}{\theta_{sat,i} - \theta_{ice,i}}, 0.01 \le S \le 1 \quad \text{(A19)}$$

where θ_{sat} for a soil layer is:

$$\theta_{sat,i} = \left(1 - F_{om,i}\right) \times \theta_{sat,min,i} + F_{om,i} \times \theta_{sat,om} \quad \text{(A20)}$$

$\theta_{sat,om}$ is the porosity of the organic matter, while the porosity of the mineral soil ($\theta_{sat,min}$) using (*%sand*) is:

$$\theta_{sat,min,i} = 0.489 - 0.00126 \times (\%sand)_i \quad \text{(A21)}$$

Similar to ψ_{sat} and θ_{sat} (see equation 17 and 20), the parameter B is calculated as:

https://doi.org/10.5194/bg-2022-211
Preprint. Discussion started: 2 November 2022

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

where the B_{om} is the exponent for organic matter while $B_{min,i}$ the exponent for mineral soil is:

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

7 Code Availability

A code release with the example training dataset will be provided upon acceptance of the paper.

8 Data Availability

Datasets used in the model are publicly available from the sources cited in this paper.

9 Author Contribution

DA implemented the numerical models, ran experiments, and produced the figures. DA and CS completed the initial manuscript. CS and CX conceived the study. CX, FH, KL and CS edited the manuscript. CS implemented the parallel Newton solver for nonlinear system on PyTorch, while DA implemented the Julia version with assistance from AJ and CR.

10 Competing Interests

KL and CS have financial interests in HydroSapient, Inc., a company which could potentially benefit from the results of this research. This interest has been reviewed by the University in accordance with its Individual Conflict of Interest policy, for the purpose of maintaining the objectivity and the integrity of research at The Pennsylvania State University.

11 Acknowledgements

This work was supported by the US. Department of Energy, Office of Science under award DE-SC0021979.

https://doi.org/10.5194/bg-2022-211
Preprint. Discussion started: 2 November 2022

References

Ali, A. A., Xu, C., Rogers, A., McDowell, N. G., Medlyn, B. E., Fisher, R. A., Wullschleger, S. D., Reich, P. B., Vrugt, J. A., Bauerle, W. L., Santiago, L. S., and Wilson, C. J.: Global-scale environmental control of plant photosynthetic capacity., Ecol Appl, 25, 2349–2365, https://doi.org/10.1890/14-2111.1, 2015.

Ali, A. A., Xu, C., Rogers, A., Fisher, R. A., Wullschleger, S. D., Massoud, E. C., Vrugt, J. A., Muss, J. D., McDowell, N. G., Fisher, J. B., Reich, P. B., and Wilson, C. J.: A global scale mechanistic model of photosynthetic capacity (LUNA V1.0), Geoscientific Model Development, 9, 587–606, https://doi.org/10/f8qhr3, 2016.

Baydin, A. G., Pearlmutter, B. A., Radul, A. A., and Siskind, J. M.: Automatic differentiation in machine learning: A survey, Journal of Machine Learning Research, 18, 1–43, 2018.

Beven, K.: A manifesto for the equifinality thesis, Journal of Hydrology, 320, 18–36, https://doi.org/10/ccx2ks, 2006.

Beven, K. and Freer, J.: Equifinality, data assimilation, and uncertainty estimation in mechanistic modelling of complex environmental systems using the GLUE methodology, Journal of Hydrology, 249, 11–29, https://doi.org/10/fgmngv, 2001.

Bezanson, J., Karpinski, S., Shah, V. B., and Edelman, A.: Julia: A Fast Dynamic Language for Technical Computing, ArXiv, abs/1209.5145, 2012.

Chen, J. M., Wang, R., Liu, Y., He, L., Croft, H., Luo, X., Wang, H., Smith, N. G., Keenan, T. F., Prentice, I. C., Zhang, Y., Ju, W., and Dong, N.: Global Datasets of Leaf Photosynthetic Capacity for Ecological and Earth System Research, Earth System Science Data Discussions, 2022, 1–26, https://doi.org/10.5194/essd-2022-136, 2022.

Clapp, R. B. and Hornberger, G. M.: Empirical equations for some soil hydraulic properties, Water Resources Research, 14, 601–604, https://doi.org/10.1029/wr014i004p00601, 1978.

Collatz, G., Ribas-Carbo, M., and Berry, J.: Coupled Photosynthesis-Stomatal Conductance Model for Leaves of C 4 Plants, Australian Journal of Plant Physiology, 19, 519, https://doi.org/10/cw8rtn, 1992.

Copernicus Climate Change Service (C3S): ERA5: Fifth generation of ECMWF atmospheric reanalyses of the global climate. Copernicus Climate Change Service Climate Data Store (CDS), 2017.

Croft, H., Chen, J. M., Luo, X., Bartlett, P., Chen, B., and Staebler, R. M.: Leaf chlorophyll content as a proxy for leaf photosynthetic capacity, Global Change Biology, 23, 3513–3524, https://doi.org/10.1111/gcb.13599, 2017.

Dusenge, M. E., Duarte, A. G., and Way, D. A.: Plant carbon metabolism and climate change: elevated CO2 and temperature impacts on photosynthesis, photorespiration and respiration, New Phytologist, 221, 32–49, https://doi.org/10.1111/nph.15283, 2019.

ElSaadani, M., Habib, E., Abdelhameed, A. M., and Bayoumi, M.: Assessment of a Spatiotemporal Deep Learning Approach for Soil Moisture Prediction and Filling the Gaps in Between Soil Moisture Observations, Frontiers in Artificial Intelligence, 4, https://doi.org/10.3389/frai.2021.636234, 2021.

Fang, K., Shen, C., Kifer, D., and Yang, X.: Prolongation of SMAP to spatiotemporally seamless coverage of continental U.S. using a deep learning neural network, Geophys. Res. Lett., 44, 11,030-11,039, https://doi.org/10/gcr7mq, 2017.

Fang, K., Kifer, D., Lawson, K., Feng, D., and Shen, C.: The data synergy effects of time-series deep learning models in hydrology, Water Resources Research, 58, e2021WR029583, https://doi.org/10.1029/2021WR029583, 2022.

Farquhar, G. D. and Caemmerer, S. von: Modelling of Photosynthetic Response to Environmental Conditions, https://doi.org/10.1007/978-3-642-68150-9_17, 1982.

Farquhar, G. D., von Caemmerer, S., and Berry, J. A.: A biochemical model of photosynthetic CO2 assimilation in leaves of C 3 species., Planta, 149, 78–90, https://doi.org/10/fs9dpz, 1980.

Feng, D., Fang, K., and Shen, C.: Enhancing streamflow forecast and extracting insights using long-short term memory networks with data integration at continental scales, Water Resources Research, 56, e2019WR026793, https://doi.org/10.1029/2019WR026793, 2020.

Feng, D., Lawson, K., and Shen, C.: Mitigating prediction error of deep learning streamflow models in large data-sparse regions with ensemble modeling and soft data, Geophysical Research Letters, 48, e2021GL092999, https://doi.org/10.1029/2021GL092999, 2021.

Feng, D., Liu, J., Lawson, K., and Shen, C.: Differentiable, learnable, regionalized process-based models with multiphysical outputs can approach state-of-the-art hydrologic prediction accuracy, Water Resources Research, e2022WR032404, https://doi.org/10.1029/2022WR032404, 2022a.

Feng, D., Beck, H., Lawson, K., and Shen, C.: The suitability of differentiable, learnable hydrologic models for ungauged regions and climate change impact assessment, Hydrology and Earth System Sciences Discussions, 1–28, https://doi.org/10.5194/hess-2022-245, 2022b.

Fisher, R. A., Muszala, S., Verteinstein, M., Lawrence, P., Xu, C., McDowell, N. G., Knox, R. G., Koven, C., Holm, J., Rogers, B. M., Spessa, A., Lawrence, D., and Bonan, G.: Taking off the training wheels: the properties of a dynamic vegetation model without climate envelopes, CLM4.5(ED), Geoscientific Model Development, 8, 3593–3619, https://doi.org/10.5194/gmd-8-3593-2015, 2015.

Fisher, R. A., Koven, C. D., Anderegg, W. R. L., Christoffersen, B. O., Dietze, M. C., Farrior, C. E., Holm, J. A., Hurtt, G. C., Knox, R. G., Lawrence, P. J., Lichstein, J. W., Longo, M., Matheny, A. M., Medvigy, D., Muller-Landau, H. C., Powell, T. L., Serbin, S. P., Sato, H., Shuman, J. K., Smith, B., Trugman, A. T., Viskari, T., Verbeeck, H., Weng, E., Xu, C., Xu, X., Zhang, T., and Moorcroft, P. R.: Vegetation demographics in Earth System Models: A review of progress and priorities, Global Change Biology, 24, 35–54, https://doi.org/10.1111/gcb.13910, 2018.

Gowda, S., Ma, Y., Cheli, A., Gwóźdź, M., Shah, V. B., Edelman, A., and Rackauckas, C.: High-Performance Symbolic-Numerics via Multiple Dispatch, ACM Commun. Comput. Algebra, 55, 92–96, https://doi.org/10.1145/3511528.3511535, 2022.

Hengl, T.: Sand content in % (kg / kg) at 6 standard depths (0, 10, 30, 60, 100 and 200 cm) at 250 m resolution (v0.2), https://doi.org/10.5281/ZENODO.2525662, 2018.

Hengl, T. and Ichsani Wheeler: Soil Organic Carbon Content In X 5 G / Kg At 6 Standard Depths (0, 10, 30, 60, 100 And 200 Cm) At 250 M Resolution, https://doi.org/10.5281/ZENODO.1475458, 2018.

https://doi.org/10.5194/bg-2022-211
Preprint. Discussion started: 2 November 2022

Hossain, M. S., Al-Hammadi, M., and Muhammad, G.: Automatic Fruit Classification Using Deep Learning for Industrial Applications, IEEE Transactions on Industrial Informatics, 15, 1027–1034, https://doi.org/10.1109/TII.2018.2875149, 2019.

Hrnjica, B., Mehr, A. D., Jakupović, E., Crnkić, A., and Hasanagić, R.: Application of Deep Learning Neural Networks for Nitrate Prediction in the Klokot River, Bosnia and Herzegovina, in: 2021 7th International Conference on Control, Instrumentation and Automation (ICCIA), 2021 7th International Conference on Control, Instrumentation and Automation (ICCIA), 1–6, https://doi.org/10.1109/ICCIA52082.2021.9403565, 2021.

Hüllermeier, E. and Waegeman, W.: Aleatoric and epistemic uncertainty in machine learning: an introduction to concepts and methods, Machine Learning, 110, 457–506, https://doi.org/10.1007/s10994-021-05946-3, 2021.

Innes, M., Edelman, A., Fischer, K., Rackauckas, C., Saba, E., Shah, V. B., and Tebbutt, W.: A Differentiable Programming System to Bridge Machine Learning and Scientific Computing, http://arxiv.org/abs/1907.07587, 18 July 2019.

Kattge, J., Bönisch, G., Díaz, S., Lavorel, S., Prentice, I. C., Leadley, P., Tautenhahn, S., Werner, G. D. A., Aakala, T., Abedi, M., Acosta, A. T. R., Adamidis, G. C., Adamson, K., Aiba, M., Albert, C. H., Alcántara, J. M., Alcázar C, C., Aleixo, I., Ali, H., Amiaud, B., Ammer, C., Amoroso, M. M., Anand, M., Anderson, C., Anten, N., Antos, J., Apgaua, D. M. G., Ashman, T.-L., Asmara, D. H., Asner, G. P., Aspinwall, M., Atkin, O., Aubin, I., Baastrup-Spohr, L., Bahalkeh, K., Bahn, M., Baker, T., Baker, W. J., Bakker, J. P., Baldocchi, D., Baltzer, J., Banerjee, A., Baranger, A., Barlow, J., Barneche, D. R., Baruch, Z., Bastianelli, D., Battles, J., Bauerle, W., Bauters, M., Bazzato, E., Beckmann, M., Beeckman, H., Beierkuhnlein, C., Bekker, R., Belfry, G., Belluau, M., Beloiu, M., Benavides, R., Benomar, L., Berdugo-Lattke, M. L., Berenguer, E., Bergamin, R., Bergmann, J., Bergmann Carlucci, M., Berner, L., Bernhardt-Römermann, M., Bigler, C., Bjorkman, A. D., Blackman, C., Blanco, C., Blonder, B., Blumenthal, D., Bocanegra-González, K. T., Boeckx, P., Bohlman, S., Böhning-Gaese, K., Boisvert-Marsh, L., Bond, W., Bond-Lamberty, B., Boom, A., Boonman, C. C. F., Bordin, K., Boughton, E. H., Boukili, V., Bowman, D. M. J. S., Bravo, S., Brendel, M. R., Broadley, M. R., Brown, K. A., Bruelheide, H., Brumnich, F., Bruun, H. H., Bruy, D., Buchanan, S. W., Bucher, S. F., Buchmann, N., Buitenwerf, R., Bunker, D. E., et al.: TRY plant trait database – enhanced coverage and open access, Global Change Biology, 26, 119–188, https://doi.org/10.1111/gcb.14904, 2020.

Kirschbaum, M. U. F.: Direct and indirect climate change effects on photosynthesis and transpiration., Plant biology, 6 3, 242–53, https://doi.org/10.1055/s-2004-820883, 2004.

Knauer, J., Zaehle, S., Medlyn, B. E., Reichstein, M., Williams, C. A., Migliavacca, M., De Kauwe, M. G., Werner, C., Keitel, C., Kolari, P., Limousin, J.-M., and Linderson, M.-L.: Towards physiologically meaningful water-use efficiency estimates from eddy covariance data, Global Change Biology, 24, 694–710, https://doi.org/10.1111/gcb.13893, 2018.

Koven, C. D., Knox, R. G., Fisher, R. A., Chambers, J. Q., Christoffersen, B. O., Davies, S. J., Detto, M., Dietze, M. C., Faybishenko, B., Holm, J., Huang, M., Kovenock, M., Kueppers, L. M., Lemieux, G., Massoud, E., McDowell, N. G., Muller-Landau, H. C., Needham, J. F., Norby, R. J., Powell, T., Rogers, A., Serbin, S. P., Shuman, J. K., Swann, A. L. S., Varadharajan, C., Walker, A. P., Wright, S. J., and Xu, C.: Benchmarking and parameter sensitivity of physiological and

https://doi.org/10.5194/bg-2022-211
Preprint. Discussion started: 2 November 2022

vegetation dynamics using the Functionally Assembled Terrestrial Ecosystem Simulator (FATES) at Barro Colorado Island, Panama, Biogeosciences, 17, 3017–3044, https://doi.org/10.5194/bg-17-3017-2020, 2020.

Lawrence, D. M., Fisher, R. A., Koven, C. D., Oleson, K. W., Swenson, S. C., Bonan, G., Collier, N., Ghimire, B., van Kampenhout, L., Kennedy, D., Kluzek, E., Lawrence, P. J., Li, F., Li, H., Lombardozzi, D., Riley, W. J., Sacks, W. J., Shi, M., Vertenstein, M., Wieder, W. R., Xu, C., Ali, A. A., Badger, A. M., Bisht, G., van den Broeke, M., Brunke, M. A., Burns, S. P., Buzan, J., Clark, M., Craig, A., Dahlin, K., Drewniak, B., Fisher, J. B., Flanner, M., Fox, A. M., Gentine, P., Hoffman, F., Keppel-Aleks, G., Knox, R., Kumar, S., Lenaerts, J., Leung, L. R., Lipscomb, W. H., Lu, Y., Pandey, A., Pelletier, J. D., Perket, J., Randerson, J. T., Ricciuto, D. M., Sanderson, B. M., Slater, A., Subin, Z. M., Tang, J., Thomas, R. Q., Val Martin, M., and Zeng, X.: The Community Land Model Version 5: Description of New Features, Benchmarking, and Impact of Forcing Uncertainty, Journal of Advances in Modeling Earth Systems, 11, 4245–4287, https://doi.org/10.1029/2018ms001583, 2019.

LeCun, Y., Bengio, Y., and Hinton, G.: Deep Learning, Nature, 521, 436–444, https://doi.org/10/bmqp, 2015.

Leong, W. J. and Horgan, H. J.: DeepBedMap: a deep neural network for resolving the bed topography of Antarctica, The Cryosphere, 14, 3687–3705, https://doi.org/10.5194/tc-14-3687-2020, 2020.

Lin, Y.-S., Medlyn, B. E., Duursma, R. A., Prentice, I. C., Wang, H., Baig, S., Eamus, D., de Dios, V. R., Mitchell, P., Ellsworth, D. S., de Beeck, M. O., Wallin, G., Uddling, J., Tarvainen, L., Linderson, M.-L., Cernusak, L. A., Nippert, J. B., Ocheltree, T. W., Tissue, D. T., Martin-StPaul, N. K., Rogers, A., Warren, J. M., De Angelis, P., Hikosaka, K., Han, Q., Onoda, Y., Gimeno, T. E., Barton, C. V. M., Bennie, J., Bonal, D., Bosc, A., Löw, M., Macinins-Ng, C., Rey, A., Rowland, L., Setterfield, S. A., Tausz-Posch, S., Zaragoza-Castells, J., Broadmeadow, M. S. J., Drake, J. E., Freeman, M., Ghannoum, O., Hutley, L. B., Kelly, J. W., Kikuzawa, K., Kolari, P., Koyama, K., Limousin, J.-M., Meir, P., Lola da Costa, A. C., Mikkelsen, T. N., Salinas, N., Sun, W., and Wingate, L.: Optimal stomatal behaviour around the world, Nature Climate Change, 5, 459–464, https://doi.org/10.1038/nclimate2550, 2015.

Liu, J., Rahmani, F., Lawson, K., and Shen, C.: A multiscale deep learning model for soil moisture integrating satellite and in situ data, Geophysical Research Letters, 49, e2021GL096847, https://doi.org/10.1029/2021GL096847, 2022a.

Liu, J., Hughes, D., Rahmani, F., Lawson, K., and Shen, C.: Evaluating a Global Soil Moisture dataset from a Multitask Model (GSM3 v1.0) for current and emerging threats to crops, Geoscientific Model Development Discussions, 2022, 1–23, https://doi.org/10.5194/gmd-2022-211, 2022b.

Luo, X., Keenan, T., Chen, J., Croft, H., Prentice, I., Smith, N., Walker, A., Wang, H., Wang, R., Xu, C., and Zhang, Y.: Global variation in the fraction of leaf nitrogen allocated to photosynthesis, Nature Communications, 12, https://doi.org/10.1038/s41467-021-25163-9, 2021.

Ma, Y., Gowda, S., Anantharaman, R., Laughman, C., Shah, V. B., and Rackauckas, C.: ModelingToolkit: A Composable Graph Transformation System For Equation-Based Modeling, CoRR, abs/2103.05244, 2021.

Mäkelä, J., Knauer, J., Aurela, M., Black, A., Heimann, M., Kobayashi, H., Lohila, A., Mammarella, I., Margolis, H., Markkanen, T., Susiluoto, J., Thum, T., Viskari, T., Zaehle, S., and Aalto, T.: Parameter calibration and stomatal

conductance formulation comparison for boreal forests with adaptive population importance sampler in the land surface model JSBACH, Geoscientific Model Development, 12, 4075–4098, https://doi.org/10.5194/gmd-12-4075-2019, 2019.

Medlyn, B. E., Duursma, R. A., Eamus, D., Ellsworth, D. S., Prentice, I. C., Barton, C. V. M., Crous, K. Y., de Angelis, P., Freeman, M., and Wingate, L.: Reconciling the optimal and empirical approaches to modelling stomatal conductance, Global Change Biology, 17, 2134–2144, https://doi.org/10.1111/j.1365-2486.2010.02375.x, 2011.

Meyer, F. H.: Encyclopedia of Plant Physiology, New Series. Editors: Pirson, A.; Zimmermann, M.H., Vol. 12, Part A (in 4 parts) Physiological Plant Ecology I. Responses to the Physical Environment, Editors: Lange, O.L.; Nobel, P.S.; Osmond, C.B.; Ziegler, H., Springer-Verlag, Berlin–Heidelberg–New York, 1981, 110 figs. XV, 625 pages. Cloth DM 239,–, Zeitschrift für Pflanzenernährung und Bodenkunde, 146, 543–544, https://doi.org/10.1002/jpln.19831460417, 1983.

Nash, J. E. and Sutcliffe, J. V.: River flow forecasting through conceptual models part I — A discussion of principles, Journal of Hydrology, 10, 282–290, https://doi.org/10/fbg9tm, 1970.

Oleson, K. W., Lawrence, D. M., Bonan, G. B., Drewniak, B., Huang, M., Charles, D., Levis, S., Li, F., Riley, W. J., Zachary, M., Swenson, S. C., Thornton, P. E., Bozbiyik, A., Fisher, R., Heald, C. L., Kluzek, E., Lamarque, F., Lawrence, P. J., Leung, L. R., Muszala, S., Ricciuto, D. M., and Sacks, W.: NCAR/TN-503+STR NCAR Technical Note, 2013.

Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., Killeen, T., Lin, Z., Gimelshein, N., Antiga, L., Desmaison, A., Kopf, A., Yang, E., DeVito, Z., Raison, M., Tejani, A., Chilamkurthy, S., Steiner, B., Fang, L., Bai, J., and Chintala, S.: PyTorch: An imperative style, high-performance deep learning library, in: Advances in Neural Information Processing Systems 32, 8024–8035, 2019.

Qian, X., Liu, L., Croft, H., and Chen, J.: C_3 plants converge on a universal relationship between leaf maximum carboxylation rate and chlorophyll content, Biogeosciences Discussions, 2019, 1–18, https://doi.org/10.5194/bg-2019-228, 2019.

Quillet, A., Peng, C., and Garneau, M.: Toward dynamic global vegetation models for simulating vegetation–climate interactions and feedbacks: recent developments, limitations, and future challenges, Environmental Reviews, 18, 333–353, https://doi.org/10.1139/A10-016, 2010.

Rahmani, F., Oliver, S., Ouyang, W., Appling, A., Lawson, K., and Shen, C.: Developing and testing a long short-term memory stream temperature model in daily and continental scale, AGU 2020 Fall Meeting, publisher: Earth and Space Science Open ArchiveDOI: 10.1002/essoar.10505077.1, https://doi.org/10.1002/essoar.10505077.1, 2020.

Rahmani, F., Shen, C., Oliver, S., Lawson, K., and Appling, A.: Deep learning approaches for improving prediction of daily stream temperature in data-scarce, unmonitored, and dammed basins, Hydrological Processes, 35, e14400, https://doi.org/10.1002/hyp.14400, 2021.

Reichstein, M., Camps-Valls, G., Stevens, B., Jung, M., Denzler, J., Carvalhais, N., and Prabhat: Deep learning and process understanding for data-driven Earth system science, Nature, 566, 195–204, https://doi.org/10/gfvhxk, 2019.

Rogers, A.: The use and misuse of V(c,max) in Earth System Models., Photosynth Res, 119, 15–29, https://doi.org/10.1007/s11120-013-9818-1, 2014.

https://doi.org/10.5194/bg-2022-211
Preprint. Discussion started: 2 November 2022

Saleem, M. H., Potgieter, J., and Arif, K. M.: Plant Disease Detection and Classification by Deep Learning, Plants, 8, https://doi.org/10.3390/plants8110468, 2019.

Shen, C.: Deep learning: A next-generation big-data approach for hydrology, Eos, 99, 2018.

Shen, C., Laloy, E., Elshorbagy, A., Albert, A., Bales, J., Chang, F.-J., Ganguly, S., Hsu, K.-L., Kifer, D., Fang, Z., Fang, K., Li, D., Li, X., and Tsai, W.-P.: HESS Opinions: Incubating deep-learning-powered hydrologic science advances as a community, Hydrology and Earth System Sciences, 22, 5639–5656, https://doi.org/10.5194/hess-22-5639-2018, 2018.

Tang, J. and Zhuang, Q.: Equifinality in parameterization of process-based biogeochemistry models: A significant uncertainty source to the estimation of regional carbon dynamics, Journal of Geophysical Research, 113, G04010, https://doi.org/10/btw8qf, 2008.

Thompson, M., Gamage, D., Hirotsu, N., Martin, A., and Seneweera, S.: Effects of elevated carbon dioxide on photosynthesis and carbon partitioning: A perspective on root sugar sensing and hormonal crosstalk, Frontiers in Physiology, 8, https://doi.org/10.3389/fphys.2017.00578, 2017.

Tsai, W.-P., Feng, D., Pan, M., Beck, H., Lawson, K., Yang, Y., Liu, J., and Shen, C.: From calibration to parameter learning: Harnessing the scaling effects of big data in geoscientific modeling, Nat Commun, 12, 5988, https://doi.org/10.1038/s41467-021-26107-z, 2021.

Urban, L., Aarrouf, J., and Bidel, L.: Assessing the effects of water deficit on photosynthesis using parameters derived from measurements of leaf gas exchange and of chlorophyll A fluorescence, Frontiers in Plant Science, 8, https://doi.org/10.3389/fpls.2017.02068, 2017.

Verheijen, L. M., Brovkin, V., Aerts, R., Bönisch, G., Cornelissen, J. H. C., Kattge, J., Reich, P. B., Wright, I. J., and van Bodegom, P. M.: Impacts of trait variation through observed trait–climate relationships on performance of an Earth system model: a conceptual analysis, Biogeosciences, 10, 5497–5515, https://doi.org/10.5194/bg-10-5497-2013, 2013.

Von Caemmerer, S.: C4 photosynthesis in a single C3 cell is theoretically inefficient but may ameliorate internal CO2 diffusion limitations of C3 leaves, Plant, Cell & Environment, 26, 1191–1197, https://doi.org/10.1046/j.0016-8025.2003.01061.x, 2003.

Von Caemmerer, S.: Steady-state models of photosynthesis, Plant, Cell & Environment, 36, 1617–1630, https://doi.org/10.1111/pce.12098, 2013.

Wang, H. B., Ma, M. G., Xie, Y. M., Wang, X. F., and Wang, J.: Parameter inversion estimation in photosynthetic models: Impact of different simulation methods, Photosynthetica, 52, 233–246, https://doi.org/10.1007/s11099-014-0027-8, 2014.

Xu, C., McDowell, N. G., Fisher, R. A., Wei, L., Sevanto, S., Christoffersen, B. O., Weng, E., and Middleton, R. S.: Increasing impacts of extreme droughts on vegetation productivity under climate change, Nature Climate Change, 9, 948–953, https://doi.org/10.1038/s41558-019-0630-6, 2019.

Yin, X. and Struik, P. C.: C3 and C4 photosynthesis models: An overview from the perspective of crop modelling, NJAS: Wageningen Journal of Life Sciences, 57, 27–38, https://doi.org/10.1016/j.njas.2009.07.001, 2009.

https://doi.org/10.5194/bg-2022-211
Preprint. Discussion started: 2 November 2022

Zhang, E., Liu, L., and Huang, L.: Automatically delineating the calving front of Jakobshavn Isbræ from multitemporal TerraSAR-X images: a deep learning approach, The Cryosphere, 13, 1729–1741, https://doi.org/10.5194/tc-13-1729-2019, 2019.

Zhang, X.-Y., Huang, Z., Su, X., Siu, A., Song, Y., Zhang, D., and Fang, Q.: Machine learning models for net photosynthetic rate prediction using poplar leaf phenotype data., PLoS One, 15, e0228645, https://doi.org/10.1371/journal.pone.0228645, 2020.

Zhang, Z., Xin, Q., and Li, W.: Machine Learning-Based Modeling of Vegetation Leaf Area Index and Gross Primary Productivity Across North America and Comparison With a Process-Based Model, Journal of Advances in Modeling Earth Systems, 13, e2021MS002802, https://doi.org/10.1029/2021MS002802, 2021.

Zhi, W., Feng, D., Tsai, W.-P., Sterle, G., Harpold, A., Shen, C., and Li, L.: From hydrometeorology to river water quality: Can a deep learning model predict dissolved oxygen at the continental scale?, Environ. Sci. Technol., 55, 2357–2368, https://doi.org/10.1021/acs.est.0c06783, 2021.