https://doi.org/10.5194/egusphere-2026-2241
Preprint. Discussion started: 8 May 2026

Constraining boreal carbon allocation and turnover by assimilating forest growth dynamics in a differentiable framework

Jincheng Wu[1], Philippe Ciais[2], Yuanyuan Huang[3,4], Jianing Fang[5], Pierre Gentine[5], Yidi Xu[2], Daniel S. Goll[2], Fayong Liu[3,6], Xiaomeng Du[1], Rui Ma[2], Nan Meng[2], Mengjie Han[1], Jinlong Zang[1], Runda Jiang[7], Wei Li[1]*

[1] Department of Earth System Science, Ministry of Education Key Laboratory for Earth System Modeling, Institute for Global Change Studies, Tsinghua University, Beijing 100084, China
[2] Laboratoire des Sciences du Climat et de l'Environnement, CEA CNRS UVSQ, Université Paris-Saclay, Gif-sur-Yvette 91191, France
[3] Key Laboratory of Ecosystem Network Observation and Modeling, Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing 100101, China
[4] State Key Laboratory of Resources and Environmental Information System, Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing 100101, China
[5] Department of Earth and Environmental Engineering, Columbia University, New York, NY 10027, USA
[6] University of Chinese Academy of Sciences, Beijing 100049, China
[7] School of Earth System Science, Tianjin University, Tianjin 300072, China
* Correspondence to Wei Li: wli2019@tsinghua.edu.cn

Abstract

Terrestrial biosphere models (TBMs) exhibit substantial uncertainty in simulating the land carbon sink in particular at interannual to decadal time scales, partly because parameters that govern carbon allocation and biomass turnover rates are weakly constrained. Typical model-data fusion approaches, which rely heavily on high-frequency (minutes to days) observations related to fluxes (e.g., gross primary production and leaf area index), often struggle to constrain the slow turnover processes that govern long-term biomass accumulation over multiple years. Here, we employ DifferLand, a JAX-based differentiable TBM, to jointly assimilate satellite-derived boreal forest biomass growth trajectories and high-frequency observations. Calibrations using only high-frequency fluxes reproduce short-term dynamics but yield large biases in mature forest biomass (RMSE = 138.7 Mg ha^{-1}), while incorporating a single-year biomass stock constraint only partially reduces the error (RMSE = 87.5 Mg ha^{-1}) and provides limited constraints on the allocation patterns. In contrast, incorporating our biomass growth curves reduce the biomass RMSE to 11.3 Mg ha^{-1} (a 91.9% reduction) without degrading fits to high-frequency fluxes. Our findings reveal that while a single-year biomass stock data provides some constraints on biomass residence times, the full growth trajectory is essential to simultaneously constrain carbon allocation and turnover. The retrieved parameters indicate that boreal forests sustain biomass primarily through longer wood carbon residence times (74.1% lower wood turnover rates) rather than higher allocation to wood, compared to calibrations using only high-frequency flux observations. Attribution analyses further show that climate conditions are the dominant driver of wood turnover, with a sharp increase when the temperature of the coldest month exceeds -20°C. Our study demonstrates the importance of assimilating slow ecological trajectories to improve long-term predictions of carbon storage and highlights the potential acceleration of the sensitivity of the boreal carbon sink to warming under future climate change.

https://doi.org/10.5194/egusphere-2026-2241
Preprint. Discussion started: 8 May 2026

1. Introduction

Terrestrial ecosystems play a pivotal role in regulating the global carbon cycle by storing a substantial fraction of anthropogenic CO_2 emissions (Schimel et al., 2015). As a core component of Earth System Models (ESMs), terrestrial biosphere models (TBMs) provide a fundamental tool to simulate terrestrial carbon and water fluxes and assess how climate change influences land ecosystem functioning and the terrestrial carbon sink (Zaehle, 2013). However, despite increasing complexity in TBMs, their projections of future changes in the terrestrial carbon sinks still exhibit substantial uncertainties (Famiglietti et al., 2021). These uncertainties primarily stem from differences in how models represent key biogeochemical and biophysical processes, such as photosynthesis and respiration, carbon allocation, vegetation dynamics, disturbance, and soil carbon turnover, and in how these processes respond to climate and CO_2 (Canadell et al., 2021). The magnitude of resulting uncertainty in carbon sink projections arising from these process-level differences is comparable to the uncertainty in physical climate sensitivity (e.g., equilibrium climate sensitivity, ECS), and together these uncertainties represent a major barrier to reliable future climate projections (Kaufhold et al., 2025).

A primary uncertainty in TBMs lies in the post-photosynthetic processes, especially the parameterization of carbon allocation and biomass turnover (i.e., carbon residence time) (Bloom et al., 2016). Uncertainty in carbon residence time is a major contributor to the spread of TBM projections under changing climate and rising atmospheric CO_2, as reflected by large inter-model differences in baseline forest biomass turnover times (Friend et al., 2014; Pugh et al., 2020). Such divergence arises from contrasting assumptions about how photosynthates are allocated among tissues and how quickly carbon is transferred and lost from biomass (Pugh et al., 2020). TBMs typically rely on Plant Functional Types (PFTs) to prescribe these parameters, but this PFT-based classification exhibits a limited ability to account for the observed spatial heterogeneity in carbon allocation and residence times (Bloom et al., 2016).

Using model-data fusion methods to constrain parameters faces the challenge of equifinality, where different parameter combinations can fit the observational data equally well (Famiglietti et al., 2021). Most model-data fusion studies rely on high-frequency observational data at hourly, daily, or monthly scales from satellites or eddy-covariance flux towers, such as gross primary productivity (GPP), ecosystem respiration (RECO), and leaf area index (LAI) to constrain critical parameters of the carbon cycle (Bloom et al., 2016; Kuppel et al., 2014; Meroni et al., 2019; Wang et al., 2019; Zheng et al., 2020). These high-frequency observations capture diurnal and seasonal dynamics of canopy photosynthesis and topsoil respiration, but they are insufficient for constraining model predictions at annual or longer timescales (Keenan et al., 2012; Peylin et al., 2016). Indeed, signals originating from large, slow-turnover carbon pools (such as woody biomass) are difficult to detect in high-frequency flux data (Braswell et al., 2005). Consequently, even if a model near perfectly fits short-term GPP and RECO, its internal allocation and long-term carbon turnover parameters may still be largely unconstrained, undermining the reliability of predictions at decadal to centennial time scales, such as for forest biomass growth and long-term carbon storage (Bloom et al., 2016; Keenan et al., 2012; Pugh et al., 2020).

Forest biomass growth curves, which describe the trajectory of biomass accumulation with stand age following major disturbances such as fires, represent a potentially powerful constraint for constraining carbon allocation and turnover parameters in TBMs. While a single-year biomass stock data from lidar or vegetation optical depth can be assimilated to characterize current carbon stocks (Dubayah et al., 2020; Fan et al., 2019), such single-point constraints fail to effectively resolve the rate of biomass change or the underlying turnover times (Shchepashchenko & Ciais, 2026; Zhou et al., 2015), both of which are critical for making reliable long-term projections. Unlike high-frequency flux data and biomass snapshots, biomass growth curves capture the integrated outcomes of long-term carbon allocation and turnover

https://doi.org/10.5194/egusphere-2026-2241
Preprint. Discussion started: 8 May 2026

processes over decades (Xu et al., 2026; Zhou et al., 2015). A key challenge in TBM parameter calibration is equifinality, where divergent parameter sets, such as high carbon allocation paired with rapid turnover versus low allocation with slow turnover, produce identical biomass stocks. Single-year biomass stock constraints cannot distinguish between these pathways because they provide no information on the actual rate of carbon accumulation. By assimilating age-related biomass stock data derived from sources such as forest inventories (Thum et al., 2017), chronosequences (Zhou et al., 2015), or remote sensing observations (Neigh et al., 2025), model parameters governing the accumulation of slow-turnover carbon pools can be more directly constrained, thereby reducing equifinality in post-photosynthetic parameter estimates. Such slow-process constraint is particularly valuable in boreal forests, where carbon storage is dominated by long-lived pools and long-term biomass trajectories can be strongly influenced by both disturbance (e.g., fire) regimes and climate warming.

Boreal forests store approximately 30% of global terrestrial carbon (Hansen et al., 2023), most of which resides in slow-turnover carbon pools such as woody biomass and carbon-rich permafrost soils (Schuur, 2019). Carbon dynamics in this region are largely driven by disturbance-induced, non-equilibrium biomass accumulation following stand-replacing events (Jonsson & Wardle, 2009), making it an ideal testbed for evaluating the effectiveness of using forest biomass growth curves to constrain model parameters. Accordingly, this study focuses on boreal forests.

To achieve this, we employed DifferLand, an advanced, JAX-based differentiable TBM (Fang & Gentine, 2024). Using the differentiable DALEC implementation provided within DifferLand, automatic differentiation computes gradients of the multi-source mismatch with respect to model parameters, enabling efficient gradient-based calibration against heterogeneous constraints that combine fast fluxes with slow carbon stocks/trajectories (Fang & Gentine, 2024; Lin, 2024; Shi et al., 2024).

The objectives of this study are threefold: (1) evaluate the efficacy of assimilating forest biomass growth curves to constrain critical carbon allocation and turnover parameters; (2) characterize the spatial patterns of these parameters across the entire boreal biome (Thurner et al., 2016); and (3) identify the key environmental drivers (e.g., climate, soil properties, and vegetation types) governing these spatial patterns.

2. Materials and Methods

2.1 Calibration Framework

We used DifferLand, a JAX-based differentiable data assimilation framework (Bradbury et al., 2018; Fang & Gentine, 2024), to calibrate ecosystem carbon-cycle parameters against multi-source observational data streams. JAX is a high-performance numerical computing library that combines NumPy-like syntax with automatic differentiation and XLA-accelerated execution (Bradbury et al., 2018). By leveraging automatic differentiation to compute exact gradients with respect to model parameters and initial conditions through the entire simulation period, DifferLand enables the accurate joint calibration of the parameters, along with data assimilation (using observation operators between model variables and observations). Calibrating ecological processes across long timescales is particularly challenging, as error signals from slow-turnover pools are often difficult to attribute to the fast physiological drivers that govern them. DifferLand addresses this difficulty by allowing error signals from slow processes (e.g., biomass turnover) to be backpropagated to constrain fast processes (e.g., photosynthesis), effectively bridging the timescale gap. This integrated approach resolves the trade-offs between diverse datasets within a physics-constrained framework, thereby more effectively constraining carbon allocation and potentially carbon residence times.

https://doi.org/10.5194/egusphere-2026-2241
Preprint. Discussion started: 8 May 2026

The physical core of the framework is the Data Assimilation Linked Ecosystem Carbon (DALEC) model (Bloom & Williams, 2015). We used the "β-JS" (Jarvis-Stewart) configuration, which couples the process-based carbon balance model with a semi-empirical soil moisture stress function (Fang & Gentine, 2024). The model represents the ecosystem state using eight pools: six carbon pools (labile, foliar, wood, fine roots, litter, and soil organic matter) and two water pools (plant available and unavailable water).

Carbon enters the ecosystem through gross primary productivity (GPP), which is simulated by the Aggregate Canopy Model (ACM) based on leaf area index (LAI), meteorological drivers, and atmospheric CO_2 concentration (Williams et al., 1997). In this configuration ("β-JS"), GPP is further regulated by a water stress factor (β) derived from the plant available water using a piecewise-linear Jarvis-Stewart function. Coupled with the carbon cycle, the model simultaneously simulates evapotranspiration (ET) as a function of meteorological drivers and canopy state, which determines the depletion of the plant available water pool. Net primary productivity (NPP) is calculated as the fraction of GPP remaining after accounting for autotrophic respiration. The resulting NPP is partitioned into the four well-mixed live biomass pools based on time-invariant allocation fractions, assuming no vegetation demography (i.e., tree-size or age classes are not explicitly resolved). Leaf phenology follows the Combined Deciduous-Evergreen Analytical (CEDA) scheme, while the turnover of live biomass feeds into litter and soil organic matter pools (Bloom et al., 2016). In total, the model comprises 26 tunable physical parameters (Table S1) governing these ecohydrological processes.

For the subsequent analysis, the "allocation to leaf" is calculated as the sum of allocation fractions to the labile and foliar pools since they both support canopy development. Additionally, the "leaf lifespan" reported in this study represents the effective mean residence time derived from the integrated annual turnover rate in the CEDA phenology scheme, rather than the raw shape parameter used in the model equations (see Text S1, Fig. S1 for details).

2.2 Data Source and Preprocessing

2.2.1 Meteorological Forcing Data

We used the CRU-JRA v2.2 reanalysis dataset (Harris, 2021) as the climate forcing to drive the DALEC model. This dataset provides meteorological fields at a spatial resolution of 0.5° × 0.5° and a 6-hourly temporal resolution. We aggregated the 6-hourly 2 m air temperature, precipitation, and downward shortwave solar radiation to a daily time step to match the model simulation step. Daily maximum and minimum temperatures (T_{max}, T_{min}) were extracted from the 6-hourly records. Atmospheric CO_2 concentrations were prescribed using annual global mean values from the TRENDY protocol (Friedlingstein et al., 2025), assuming a uniform distribution across the globe within each year. Since the vapor pressure deficit (VPD) was not directly provided in the CRU-JRA dataset, we derived it from air temperature and specific humidity (Table S2).

2.2.2 Observational Constraints for Fast Processes

We assimilated multiple earth observation data to constrain the fast-turnover carbon and water fluxes. For GPP, we used the FluxSat v2.0 daily GPP product (Joiner et al., 2018; Joiner & Yoshida, 2020, 2021). This dataset upscales eddy covariance measurements from FLUXNET 2015 using Moderate-Resolution Imaging Spectroradiometer (MODIS) nadir bidirectional reflectance distribution function adjusted reflectance (NBAR) data and a neural network approach, explicitly accounting for light use efficiency and meteorological drivers. Daily evapotranspiration (ET) estimates were obtained from the Global Land Evaporation Amsterdam Model (GLEAM) dataset (Miralles et al., 2025). GLEAM4

https://doi.org/10.5194/egusphere-2026-2241
Preprint. Discussion started: 8 May 2026

employs a hybrid modelling framework that combines physical principles (Penman-Monteith equation) with deep learning-based evaporative stress derived from eddy-covariance and sap flow data, while assimilating satellite soil moisture and vegetation optical depth (VOD) (Miralles et al., 2025), as a proxy for vegetation water stress. For ecosystem respiration (RECO), we used a data-driven upscaled product provided as 10-day means (Zeng et al., 2020). This product was generated using a random forest algorithm trained on FLUXNET 2015 observations. Vegetation state was constrained using the Global Land Surface Satellite (GLASS) 8-day leaf area index (LAI) v6 product (Ma & Liang, 2022). This version uses a bidirectional long short-term memory (Bi-LSTM) deep learning model to generate spatiotemporally continuous LAI profiles from MODIS surface reflectance.

All observational datasets were aggregated to 0.5° × 0.5° resolution. To derive constraints representative of mature forests, we utilized the ESA CCI dataset of forest aboveground biomass (AGB) for the year 2020 (Santoro & Cartus, 2023), which was first resampled to a spatial resolution of 0.05°. Using this map, we identified mature forest pixels within each 0.5° grid cell as those with a AGB density exceeding both the 90th percentile of the grid cell and a minimum threshold of 50 Mg ha^{-1}. The observational data for GPP, ET, RECO, and LAI were then aggregated to the 0.5° resolution by averaging only the values corresponding to these identified mature forest pixels.

2.2.3 Forest Biomass Constraints for Slow Processes

To constrain the long-term carbon turnover and allocation, we used a recently developed dataset of post-fire forest biomass growth curves specifically for boreal forests (Xu et al., 2026). These growth curves were derived using a space-for-time substitution approach, combining 36-year Landsat-based fire history records (Long et al., 2019) with a high-resolution (100 m) woody biomass map (Santoro & Cartus, 2023). For each 0.5° grid cell, the relationship between AGB and stand age (i.e., years after the previous fire) was fitted using the Richard-Chapman growth function. These spatially explicit growth trajectories provide a critical constraint on the biomass accumulation potential (AGB_{max}) and growth rate across the boreal biome (Fig. S2). As a reference for comparison, we also used the ESA CCI dataset of forest AGB for the year 2020 (Santoro & Cartus, 2023) to provide a static constraint for mature forest states.

While the satellite-derived curves and the single-year AGB represent only the aboveground component, the DALEC model simulates carbon dynamics through distinct pools where the wood pool aggregates both stems and belowground coarse roots. To ensure structural consistency between the observational constraint and the model's state variables, we converted both the satellite AGB trajectories and the 2020 AGB into total biomass curves (sum of above- and belowground biomass). This conversion was performed using spatially explicit Root-to-Shoot (R:S) ratios derived from a global map of root biomass, which was generated by a random forest model trained on extensive field measurements (Huang et al., 2021). The resulting total biomass curves provide a comprehensive constraint for the sum of all living carbon pools in the model.

2.2.4 Data for Parameter Priors

We incorporated trait-based datasets to define priors and ecological constraints for specific model parameters. Spatial reference maps for leaf carbon mass per area (LCMA) were derived from a global specific leaf area (SLA) product (Butler et al., 2017), converted to carbon mass units assuming a carbon fraction of 0.5. Additionally, to represent the known physiological trade-off between leaf lifespan and construction cost (Wright et al., 2004), we fitted a functional relationship between leaf lifespan and LCMA ($R^2 = 0.42$, $p < 0.001$, Fig. S3). This relationship was derived using a compiled global leaf trait dataset (Wang et al., 2023), which integrates the Glopnet database with the China Plant Trait

https://doi.org/10.5194/egusphere-2026-2241
Preprint. Discussion started: 8 May 2026

Database, providing a robust global representation of trait coordination. Furthermore, we utilized a global reference map of ecosystem carbon use efficiency (CUE), which was generated using a random forest model trained on CUE derived from eddy covariance observations (Luo et al., 2025).

2.2.5 Ancillary Data for Spatial Analysis

We incorporated additional time-invariant datasets to interpret the spatial heterogeneity of the calibrated parameters. Data on soil physical and chemical properties, specifically bulk density, clay content, and soil H_2O pH, were obtained from the Global Soil Dataset for Earth System Modeling (GSDE) at an original spatial resolution of 30 arc-seconds (Shangguan et al., 2014). Additionally, the soil carbon-to-nitrogen ratio was derived from the organic carbon and total nitrogen layers within the same dataset. The spatial distribution of vegetation composition was characterized using the plant functional type (PFT) maps based on the CCI land cover products, provided at 0.25° resolution (Li et al., 2018). We specifically extracted the 1985–2020 mean areal fractions of boreal needleleaf evergreen (ENF), boreal broadleaf summergreen (DBF), and boreal needleleaf deciduous (DNF). All ancillary datasets were subsequently aggregated to the 0.5° model grid.

2.2.6 Independent Validation Data

To independently validate the calibrated carbon allocation parameters, we utilized a dataset of forest biomass allocation derived from field measurements (Xia et al., 2019). We aggregated the plot-level measurements to the 0.5° grid resolution to match our model simulations. Further details on the processing of this dataset are provided in Supplementary Text S3.

2.3 Experimental Design and Model Calibration

2.3.1 Simulation Protocol

In the framework of DifferLand, the initial sizes of the eight ecological pools (labile, foliar, wood, root, litter, soil organic matter, plant available water, plant unavailable water) are treated as trainable parameters, retrieved simultaneously with the 26 physical parameters from the observational data. In other words, within the differentiable framework, initial conditions can be treated as parameters and we can compute the gradients/sensitivity of the model to those initial conditions.

To capture the long-term legacy effects of carbon turnover and accumulation, we designed a 107-year continuous simulation experiment (growth-curve-constrained experiment) for each grid cell. The meteorological drivers from 1985 to 2020 (36 years) were cycled three times, with the final year (2020) excluded from the last cycle to align the simulation end date (2019) with the extent of the observational record. This protocol simulates the forest growth trajectory from a stand-replacing disturbance event with a biomass of nearly zero to a mature state.

To quantify the specific impact of assimilating forest biomass growth dynamics, we conducted two additional experiments, a baseline experiment and a single-year biomass-constrained experiment. Both simulations spanned only 19 years (2001–2019), driven directly by the corresponding meteorological forcing without cycling. The baseline setup closely follows the original DifferLand configuration (Fang & Gentine, 2024), where the model is constrained only by high-frequency data (GPP, ET, RECO, LAI) without the information provided by the long-term forest biomass growth curves. While the baseline experiment excludes long-term biomass constraints entirely, the single-year biomass-

https://doi.org/10.5194/egusphere-2026-2241
Preprint. Discussion started: 8 May 2026

constrained experiment additionally assimilates the 2020 biomass to constrain the biomass at the end of the simulation and evaluate the efficacy of this constraint in reducing parameter equifinality.

For all experimental setups, the calibration process utilized observational data from the period 2001–2012 (corresponding to the simulation years 89–100 for the 107-year run, and years 1–12 for the 19-year run), while data from 2013–2019 were reserved for independent validation. By aligning the calibration period with simulation years 89–100, we assume that after approximately 90 years of recovery following the stand-replacing disturbance, the simulated ecosystem has reached a mature state consistent with the present-day observational constraints (GPP, RECO, ET and LAI), which were specifically filtered to represent mature forest pixels. To ensure consistency between model outputs and observations, daily model fluxes were temporally processed to match the assimilation datasets: daily outputs were averaged to match the 10-day means of RECO, while state variables were sampled to match the 8-day interval of LAI.

During the calibration of all three experimental setups, the initial values of the physical parameters and ecological pools for each grid cell were subject to five independent random initializations. Initiating multiple independent optimization sequences from overdispersed starting values is a widely recognized standard in iterative simulation and parameter estimation to ensure robust posterior inference (Gelman & Rubin, 1992). For the subsequent analysis of parameter spatial patterns, the representative estimate for each parameter was determined by taking the median across this 5-member ensemble. To ensure the reliability of the retrieved parameters, we assessed the robustness of the 5-member ensemble for each grid cell using a robust coefficient of variation (Fig. S4), and excluded parameters within a grid cell from subsequent spatial and attribution analyses if their robust coefficient of variation exceeded a threshold of 0.3 to minimize the influence of non-convergent or unstable optimizations (Leys et al., 2013; Rousseeuw & Croux, 1993).

Finally, to evaluate the impact of the retrieved optimized parameters on long-term carbon accumulation, we performed a 100-year forward simulation using the initial state derived from the growth-curve-constrained experiment. By applying the physical parameters retrieved from the baseline and the single-year biomass-constrained experiments to this same initial state from the growth-curve-constrained experiment and comparing the result against the growth-curve-constrained experiment, we isolated the divergence in 100-year biomass accumulations caused solely by the differences in physical parameters. The detailed workflow of these experimental designs is illustrated in Fig. 1.

To rigorously propagate parameter uncertainty into the prognostic outputs, these forward simulations under all three experimental setups were executed individually for each of the five ensemble members, rather than running a single simulation forced by the aggregated median parameters. Specifically, by applying the physical parameters retrieved from the baseline and single-year biomass-constrained experiments to the corresponding initial states derived from the growth-curve-constrained experiment, we simulated a 5-member ensemble of 100-year biomass accumulation trajectories. We then calculated the median of these prognostic trajectories and compared it against the ensemble median of the growth-curve-constrained experiment. This approach allowed us to robustly isolate the divergence in long-term biomass accumulations caused solely by the differences in physical parameters.

https://doi.org/10.5194/egusphere-2026-2241
Preprint. Discussion started: 8 May 2026

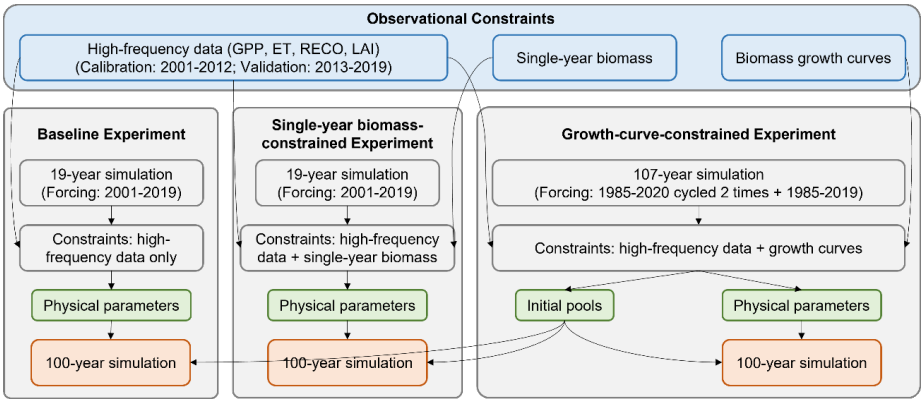

Fig. 1 Schematic illustration of the experimental design. The framework consists of two calibration experiments: a baseline experiment constrained solely by high-frequency observations (GPP, ET, RECO, and LAI) over a 19-year period; a single-year biomass-constrained experiment that additionally incorporates a single-year mature forest biomass over the same 19-year period; and a growth-curve-constrained experiment that additionally assimilates long-term forest biomass growth curves over a 107-year simulation. To isolate the impact of physical parameters on biomass storage, the 100-year forward simulations for all three experiments use the initial state retrieved from the growth-curve-constrained experiment.

2.3.2 Pre-calibration of Canopy Efficiency

Prior to the model calibration, we performed an independent pre-calibration step to determine the prior value for the canopy efficiency (ce) parameter. This parameter governs the maximum light use efficiency in the ACM module and represents the intrinsic photosynthetic capacity of the vegetation under non-water-stressed conditions. Simultaneous optimization of ce and water stress parameters often leads to an equifinality problem, where an underestimated capacity is compensated by an underestimated stress. To isolate the optimized physiological capacity from soil moisture limitations, we identified "well-watered" periods using the evaporative fraction (EF). Theoretically, EF is defined as the ratio of latent heat flux (LE) to the available energy (net radiation R_n minus ground heat flux G). In our study, this ratio is approximated by the ratio of actual evaporation to potential evaporation (E / E_p) (Miralles et al., 2011; Seneviratne et al., 2010). We calculated this proxy using daily E and E_p estimates from the GLEAM v4.2a dataset. For each grid cell, we selected days where the EF exceeded the 80th percentile of its temporal distribution. During these periods, we assumed that soil water stress was negligible (i.e., the water stress factor $\beta \approx 1$). We then optimized the *ce* parameter by minimizing the mean squared error (MSE) between the ACM-simulated GPP (driven by observed LAI and meteorology) and the observed FluxSat GPP. The retrieved optimal value (ce_{opt}) was subsequently incorporated into the full model calibration as a prior constraint.

2.3.3 Optimization and Cost Function

To retrieve the 26 physical parameters and initial states of the model, simultaneously, we employed the Adam optimizer within the JAX framework to minimize a composite cost function (J). This function integrates observational errors (J_{obs}), parameter priors (J_{prior}), and ecological constraints (J_{EDC}) into a single scalar metric. The observational loss (J_{obs}) quantifies the mismatch between model simulations and data using the negative Normalized Nash-Sutcliffe Efficiency (NNSE). The NNSE transforms the standard error metric onto a bounded range of (0, 1], preventing variables with

https://doi.org/10.5194/egusphere-2026-2241
Preprint. Discussion started: 8 May 2026

large absolute magnitudes from dominating the optimization gradient (Text S2). Accordingly, J_{obs} is calculated as the negative sum of NNSE values for the long-term total biomass trajectory and the high-frequency variables (GPP, ET, RECO, and LAI) during the calibration period.

To constrain the parameter search space and prevent ecologically unrealistic solutions, we included prior cost terms that penalize deviations of the calibrated parameters from independent estimates. These priors specifically target the canopy efficiency (ce), LCMA and the fraction of GPP allocated to autotrophic respiration (f_{auto}). The model is penalized for deviations from the pre-calibrated canopy efficiency (ce_{opt}), the satellite-derived LCMA reference maps, and the CUE-inferred f_{auto} (1 - CUE) maps, ensuring that the retrieved parameters remain anchored to observational evidence.

Finally, we imposed ecological and dynamical constraints (EDCs) to ensure biologically realistic process rates. We adopted five standard EDCs (EDC1–5) from the CARDAMOM framework (Bloom et al., 2016; Fang & Gentine, 2024) to enforce hierarchical relationships in carbon pool turnover and allocation. EDC1: Turnover rate of the litter pool should be greater than the turnover rate of soil organic matter; EDC2: The rate of decomposition of litter into soil organic matter should be faster than the release of carbon from the soil organic matter pool into the atmosphere; EDC3: Foliar carbon turnover should be faster than the wood carbon turnover; EDC4: Root turnover faster than soil organic matter turnover at mean temperature; and EDC5: Allocation to canopy is comparable to allocation to roots. Beyond the five standard constraints, we introduced specific modifications for this study. A modified constraint (EDC6) was applied to enforce a quasi-steady-state assumption for solely the soil organic matter pool, allowing vegetation carbon pools to accumulate mass over time to simulate forest biomass growth. Furthermore, we incorporated the leaf economics spectrum as a distinct ecological constraint (EDC7). This constraint enforces the fundamental physiological trade-off between leaf longevity and construction cost by penalizing parameter combinations that deviate from the global log-linear relationship between leaf lifespan and LCMA fitted in Section 2.2.4 (Fig. S3) (Wang et al., 2023; Wright et al., 2004).

2.4 Attribution Analysis using Machine Learning

To attribute the spatial heterogeneity of the calibrated parameters to environmental drivers, we employed an attribution analysis framework based on a gradient boosting regression algorithm (Friedman, 2001; Pedregosa et al., 2011). We developed separate machine learning models to predict the calibrated values of key process parameters, specifically leaf lifespan, carbon turnover rates for wood and fine roots, and NPP allocation fractions to leaf, wood, and fine roots. The predictor variables for these gradient boosting models include climatic variables derived from the CRU-JRA dataset (e.g., mean annual temperature, precipitation seasonality, and solar radiation), alongside with the soil properties and vegetation fractions (Table S4).

To evaluate the model performance in predicting the spatial variation of the parameters, we conducted a five-fold cross validation, where the results dataset of each parameter was randomly divided into five subsets, and the gradient boosting models were trained and tested iteratively. The goodness of fit was quantified using the coefficient of determination (R^2) and root mean square error (RMSE) obtained from these validation sets. We then interpreted their internal decision-making using the SHAP (SHapley Additive exPlanations) tool. SHAP quantifies the relative importance of each predictor and estimates its marginal effect on the output parameter, allowing us to disentangle the contributions of climate, soil, and vegetation structure in driving the observed spatial distribution of carbon allocation and turnover strategies across the boreal forest biome.

https://doi.org/10.5194/egusphere-2026-2241
Preprint. Discussion started: 8 May 2026

3. Results

3.1 Reducing Long-term Biomass Bias with Slow-process Constraints

The baseline experiment, which constrained the DALEC model solely using high-frequency flux and LAI observations, failed to reproduce the magnitude and spatial variability of carbon accumulation in boreal forests (Fig. 2). Despite successfully fitting the daily dynamics of GPP and RECO (Fig. S5), the baseline experiment exhibited a general underestimation simulating mature forest biomass stocks throughout the boreal biome (Fig. 2a). The model tended to underestimate biomass in high-latitude regions and overestimate biomass in the southern ecotones of the boreal biome. Consequently, the baseline simulation yielded a poor fit with the 100-year biomass derived from the forest growth curves (R^2 = 0.02), with a large RMSE of 138.7 Mg ha^{-1}. These discrepancies highlight that accurate short-term flux simulations do not guarantee reliable long-term carbon storage estimates, because turnover processes remain weakly constrained and allow compensating parameter combinations.

The single-year biomass-constrained experiment partially mitigated these biases but exhibited a distinct spatial pattern of underestimating biomass in high-latitude regions and overestimating it in the southern ecotones (Fig. 2b). While it improved the overall fit (R^2 = 0.35 and RMSE = 87.5 Mg ha^{-1}), significant systematic errors remained. Notably, this improvement was achieved primarily through a 64.8% average reduction in wood turnover rates compared to the baseline (Fig. 2g), while the fraction of NPP allocated to wood remained nearly unchanged (Fig. 2d). This suggests that a static biomass snapshot can constrain the residence time to some extent but lacks the information necessary to adjust internal carbon allocation patterns.

Assimilating the forest biomass growth curves effectively resolved these spatial biases. In the experiment with growth curve constraints, the simulated 100-year biomass aligned closely with the observational constraints along the 1:1 line (Fig. 2c). This constraint largely improved the spatial performance of the model simulation, increasing the R^2 to 0.99 and reducing the RMSE to 11.3 Mg ha^{-1}. Notably, this improvement in simulating slow carbon pools was achieved without compromising the accuracy for fast processes (Fig. S5). The spatial patterns and magnitudes of the NNSE for GPP, RECO, and ET remained similar across all experimental setups (Fig. S5), demonstrating the capacity of the framework to reconcile multi-timescale observational constraints.

The elimination of the biomass bias was primarily driven by substantial changes in carbon allocation and turnover parameters. To support higher biomass accumulation in mature stands, the optimization algorithm adjusted the parameter towards longer residence times. Specifically, the wood carbon turnover rate was reduced by an average of 74.1% compared to the baseline (Fig. 2i), implying more than a doubling of the woody biomass residence time. This reduction in turnover allowed for biomass accumulation despite a 19.9% decrease in the fraction of NPP allocated to wood (Fig. 2f). The model allocated more carbon to fine roots (Fig. S6i), potentially reflecting the high nutrient acquisition demands in boreal soils. Meanwhile, allocation to leaves decreased moderately (Fig. S6c), compensated by a 114.6% extension in leaf lifespan (Fig. S6f) to maintain the LAI required to match satellite observations, a shift consistent with the dominance of evergreen conifers in these high-latitude ecosystems.

Comparisons with field measurements of NPP allocation (Xia et al., 2019) provided independent evidence that these parameter changes moved the model towards a more ecologically realistic state (Text S3, Fig. S7). While the baseline and single-year biomass-constrained estimates for wood allocation showed poor agreement with field data ($r < 0.01$ and bias > 25%; Fig. S7b, e), the growth-curve-constrained experiment showed a marked improvement in the

https://doi.org/10.5194/egusphere-2026-2241
Preprint. Discussion started: 8 May 2026

consistency between the retrieved optimized wood allocation and the field observations, increasing the correlation to r = 0.57 and a reducing the bias to 5.86% (Fig. S7h). This suggests that only by assimilating forest biomass growth trajectories can the model effectively constrain internal carbon flows to be more consistent with observed ecological allocation patterns.

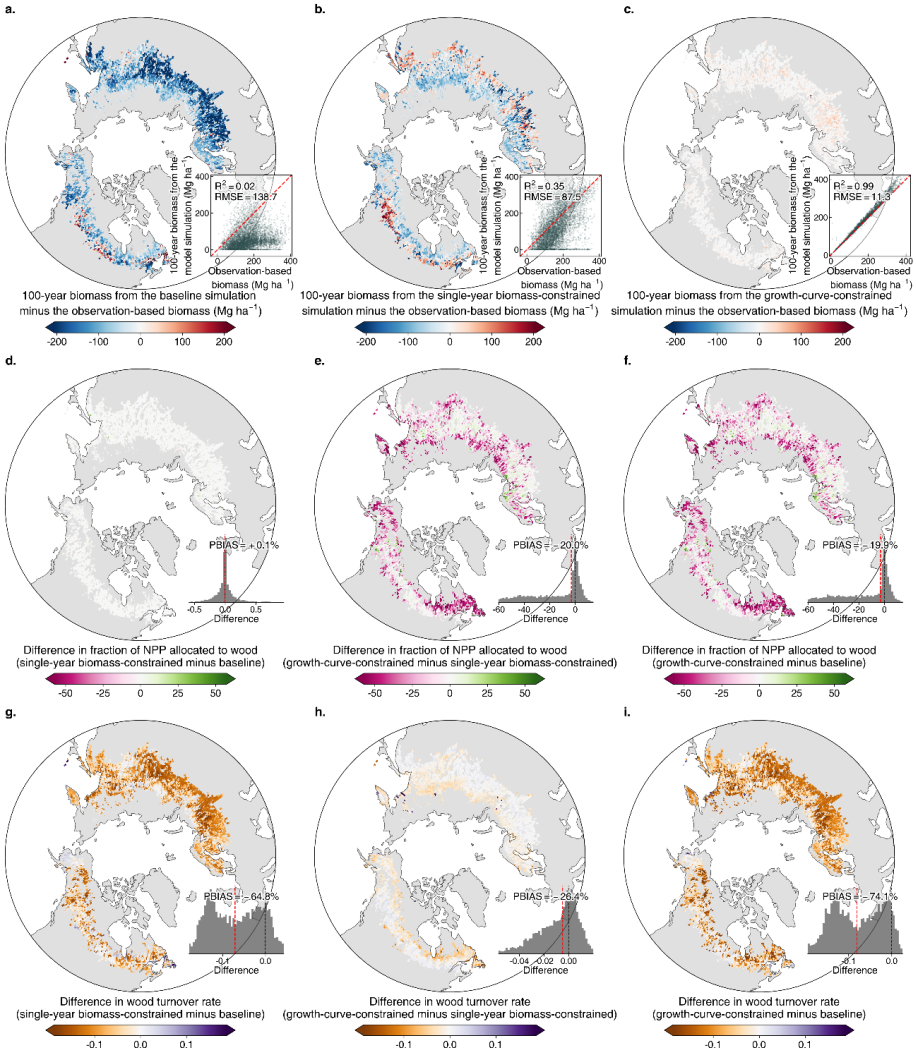

Fig. 2. Comparison of model simulated long-term biomass with that from the reference forest biomass growth curves (observation-based), and changes in wood carbon turnover and allocation. (**a-c**) Spatial distribution of the difference in the 100-year total biomass between model simulations and the observation-based estimates (**a**, baseline experiment; **b**, single-year biomass-constrained experiment; **c**, growth-curve-constrained experiment). The baseline experiment is constrained by GPP, RECO, LAI, and ET. The single-year biomass-constrained experiment additionally assimilates present-day biomass stock, whereas the growth-curve-constrained experiment assimilates forest biomass growth curves. (**d–f**) Spatial difference maps showing the changes in the fraction of NPP allocated to wood across the three experiments (**d**, single-year biomass-constrained minus baseline; **e**, growth-curve-constrained minus single-year biomass-constrained; **f**, growth-curve-constrained minus baseline). **(g–i)** Spatial difference maps showing the changes

https://doi.org/10.5194/egusphere-2026-2241
Preprint. Discussion started: 8 May 2026

in wood turnover rate (**g**, single-year biomass-constrained minus baseline; **h**, growth-curve-constrained minus single-year biomass-constrained; **i**, growth-curve-constrained minus baseline). PBIAS in the insets denotes the percent bias, indicating the average relative difference between the corresponding experiments.

3.2 Environmental Drivers Controlling Carbon Turnover and Allocation

To disentangle the complex environmental controls on the calibrated carbon turnover and allocation parameters, we employed a gradient boosting regression framework. The gradient boosting models captured ecologically consistent signals from environmental drivers, explaining a moderate proportion of the spatial variance with R^2 values ranging from 0.28 to 0.76 (Fig. S8). Predictive performance was highest for canopy-related traits, specifically the fraction of NPP allocated to leaves (R^2 = 0.60) and leaf lifespan (R^2 = 0.76). In contrast, parameters associated with slower or belowground processes, such as wood and fine root dynamics, showed lower explainability (R^2 between 0.26 and 0.34). This remaining unexplained variance likely reflects unobserved local heterogeneity (e.g., micro-site soil conditions) and stochastic non-climatic mortality agents (e.g., windthrow, insect outbreaks) that are not fully captured by the coarse-resolution environmental predictors.

3.2.1 Environmental Drivers Controlling Carbon Turnover

For the spatial variability of our optimized wood turnover rate (R^2 = 0.44), climatic variables were the most important predictors, accounting for 67.2% of the relative importance (Fig. 3b). Solar radiation was found to be the most important predictor and exhibited a strong positive correlation with the wood turnover rate (Fig. 3c). This suggests that forests in light-limited environments tend to adopt a strategy of extended biomass residence times to maintain standing stocks. Additionally, the response to the temperature of the coldest month revealed a non-linear threshold (Fig. 3e). Wood turnover rates remained stable in regions with extreme winter cold (i.e., the mean temperature of the coldest month below -25°C) but increased sharply as the temperature of the coldest month rose above -20°C. This implies that extreme cold may currently act as a stabilizing mechanism for woody biomass, probably by limiting biotic disturbance agents. Annual mean temperature also exhibited a positive relationship with turnover rate (Fig. 3h). Furthermore, wood turnover rates showed a sharp decline in more alkaline soils (Fig. 3j), likely reflecting the impacts of soil nutrient availability and pH-mediated stress on tree longevity.

In terms of leaf lifespan, the fraction of ENF was found to be the most important predictor, positively correlated with leaf lifespan (Fig. S9b, c). The fraction of DBF also emerged as an important driver, exhibiting a sharp negative relationship with leaf lifespan (Fig. S9e). This relationship confirms that the data assimilation framework successfully retrieved the distinct phenological traits of needleleaf and broadleaf species (reflecting multi-year leaf retention versus annual turnover) directly from the constraints of satellite observations.

https://doi.org/10.5194/egusphere-2026-2241
Preprint. Discussion started: 8 May 2026

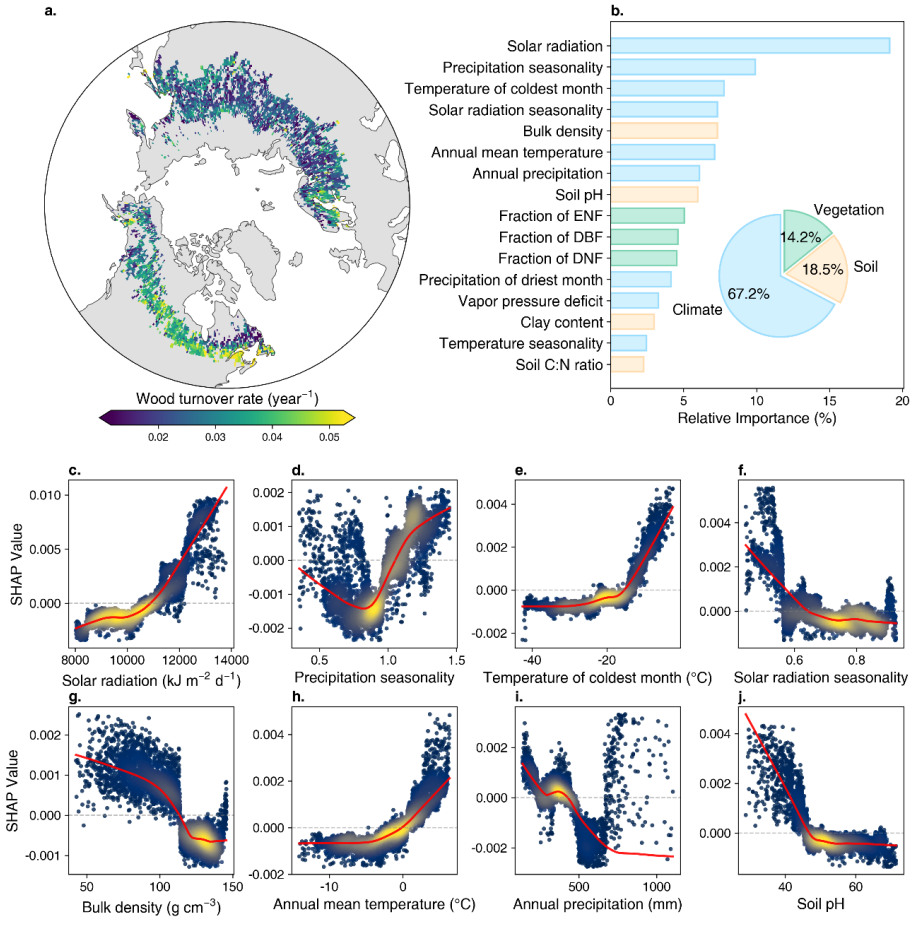

Fig. 3 Spatial patterns and environmental drivers of wood turnover rates. (**a**) Spatial distribution of the calibrated wood turnover rates across the boreal biome. (**b**) Relative importance of environmental predictors derived from the gradient boosting regression model. (**c–j**) SHAP dependence plots for the top eight most important predictors, illustrating the marginal effect of each variable on the wood turnover rate. Red solid lines indicate the locally weighted scatterplot smoothing (LOWESS) fit to highlight the trend. Note that ENF, DBF, DNF and C:N represent boreal needleleaf evergreen, boreal broadleaf summer green, boreal needleleaf deciduous, and carbon:nitrogen, respectively.

3.2.2 Environmental Drivers Controlling Carbon Allocation

For the fraction of NPP allocated to wood, VPD emerged as the most important predictor, followed by annual mean temperature and solar radiation (Fig. 4b). Climatic variables collectively accounted for 71.5% of the relative importance. Wood allocation exhibited a distinct non-linear threshold response to water stress, where investment initially increased with VPD but dropped abruptly when it exceeded about 0.35 kPa (Fig. 4c). This implies that vegetation in drier environments allocates a smaller fraction of NPP to wood. Furthermore, annual mean temperature acted as a fundamental constraint on structural investment, with wood allocation increasing as temperatures rose (Fig. 4d). In these cold environments, the reduction in wood investment was accompanied by a shift towards fine roots (Fig. S11e). This pattern suggests that vegetation in cold-limited biomes tends to prioritize belowground resource acquisition over

https://doi.org/10.5194/egusphere-2026-2241
Preprint. Discussion started: 8 May 2026

aboveground structural development (Reich et al., 2014). Environmental seasonality also remains a key driver, with wood investment generally declining under conditions of high temperature or solar radiation seasonality (Fig. 4g, j).

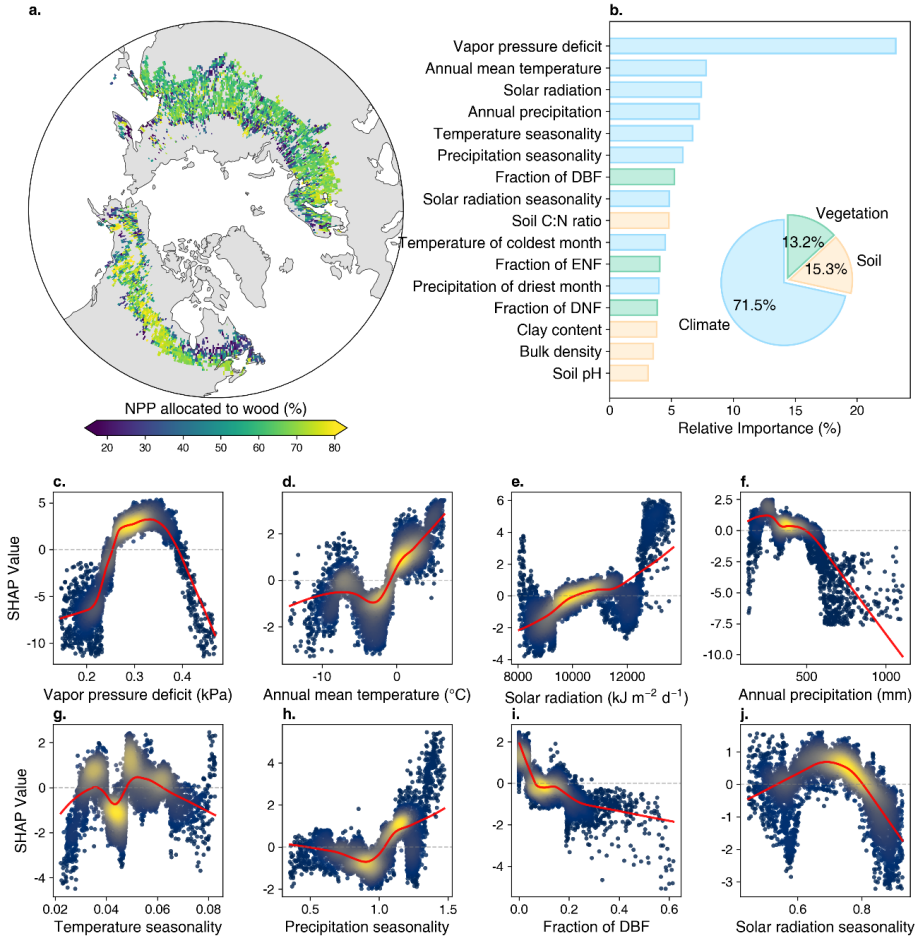

Fig. 4 Spatial patterns and environmental drivers of the fraction of net primary productivity (NPP) allocated to wood. (**a**) Spatial distribution of the fraction of NPP allocated to wood across the boreal biome. (**b**) Relative importance of environmental predictors derived from the gradient boosting regression model. (**c–j**) SHAP dependence plots for the top eight most important predictors, illustrating the marginal effect of each variable on the fraction of NPP allocated to wood. Red solid lines indicate the locally weighted scatterplot smoothing (LOWESS) fit to highlight the trend. Note that ENF, DBF, DNF and C:N represent boreal needleleaf evergreen, boreal broadleaf summer green, boreal needleleaf deciduous, and carbon:nitrogen, respectively.

For the fraction of NPP allocated to leaf, the fraction of deciduous needleleaf forests emerged as the most important predictor (Fig. 5b). Leaf investment showed a rapid increase as the fraction of deciduous needleleaf forests rose, before stabilizing at higher fraction (Fig. 5c). Atmospheric water demand also played an important role. It initially increased as VPD rose from 0.1 to 0.35 kPa (Fig. 5d), likely because a moderate increase in evaporative demand enhances the transpiration required for nutrient and water transport to the canopy, providing the necessary resources to support the construction and maintenance of a larger canopy (López et al., 2021; Massmann et al., 2019; Matimati et al., 2014).

https://doi.org/10.5194/egusphere-2026-2241
Preprint. Discussion started: 8 May 2026

However, beyond the threshold of about 0.35 kPa, NPP allocated to leaf declined (Fig. 5d), consistent with a shift toward stomatal closure to prevent hydraulic failure, which in turn limits carbon uptake and restricts further canopy expansion (Koehler et al., 2023). Moreover, leaf investment showed a strong positive correlation with bulk density (Fig. 5g). Leaf allocation also rose as the annual mean temperature increased from about -10°C to 0°C but plateaued thereafter (Fig. 5i), suggesting that extreme cold limits canopy expansion (Way & Oren, 2010).

For the fraction of NPP allocated to fine roots, it was upregulated under both high VPD conditions (indicative of hydraulic stress) and low VPD conditions typical of cold, humid areas (Fig. S11c). The NPP allocated to fine roots also increased when the soil C: N (carbon: nitrogen) ratio exceeded 25 (Fig. S11f). This to some extent confirms that the suppression of vegetation growth in high-latitude regions is driven by a functional shift towards nutrient acquisition in nitrogen-limited soils.

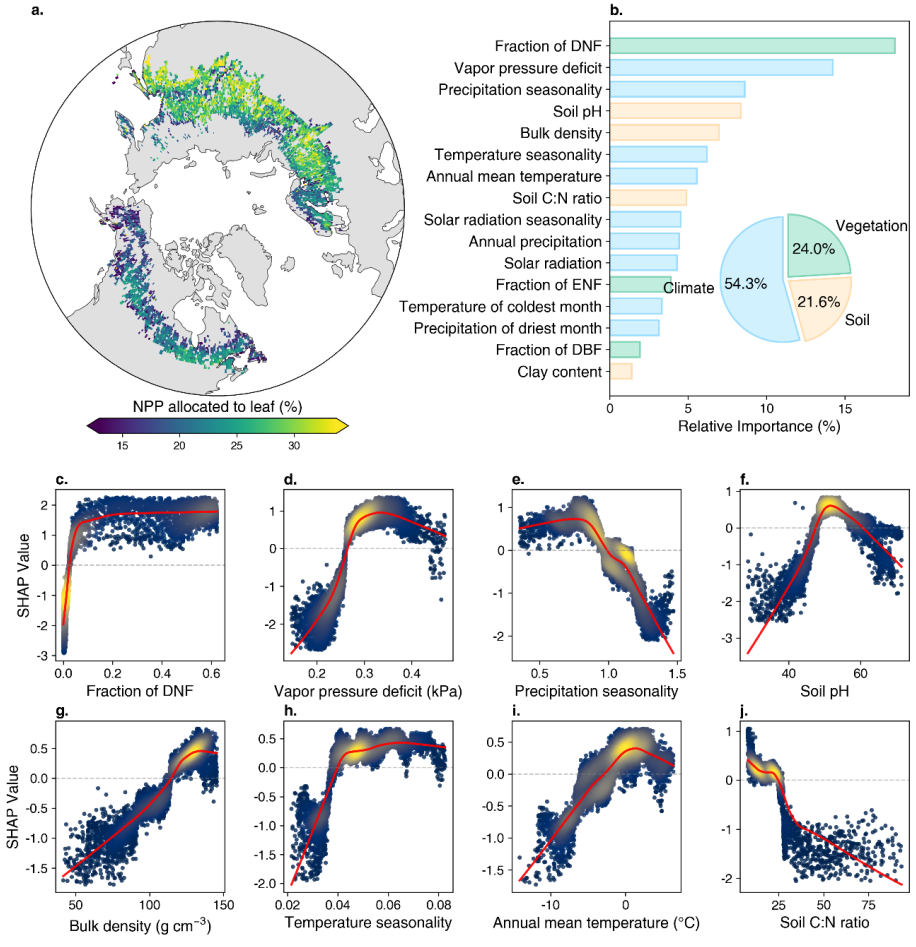

Fig. 5 Spatial patterns and environmental drivers of the fraction of net primary productivity (NPP) allocated to leaf. (**a**) Spatial distribution of the fraction of NPP allocated to leaf across the boreal biome. (**b**) Relative importance of environmental predictors derived from the gradient boosting regression model. (**c–j**) SHAP dependence plots for the top eight most important predictors, illustrating the marginal effect of each variable on the fraction of NPP allocated

https://doi.org/10.5194/egusphere-2026-2241
Preprint. Discussion started: 8 May 2026

to leaf. Red solid lines indicate the locally weighted scatterplot smoothing (LOWESS) fit to highlight the trend. Note that ENF, DBF, DNF and C:N represent boreal needleleaf evergreen, boreal broadleaf summer green, boreal needleleaf deciduous, and carbon:nitrogen, respectively.

3.3 Benchmarking against Global Carbon Cycle Retrievals

We compared our calibrated parameters against the CARDAMOM product, a spatially explicit global carbon cycle reanalysis derived by assimilating satellite observations into the DALEC model (Bloom et al., 2016). Although both frameworks perform grid-level parameter retrieval, they differ critically in the observational constraints used for boreal ecosystems. CARDAMOM relies primarily on LAI constraints in the boreal zone, whereas our framework assimilates forest biomass growth curves. This methodological difference resulted in substantial variations in the estimated carbon turnover and allocation parameters (Fig. 6 and Fig. S12).

For the wood carbon pool, our results indicate a more conservative turnover strategy compared to the CARDAMOM benchmark. We found that the fraction of NPP allocated to wood was generally lower than CARDAMOM estimates across the boreal biome, showing a mean percent bias (PBIAS) of -7.9% (Fig. 6a). The wood turnover rates were also retrieved on average to be 30.8% lower than CARDAMOM estimates (Fig. 6b). These results suggest that when constrained by long-term biomass growth trajectories rather than LAI alone, the model identifies extended biomass residence times as the primary mechanism for maintaining boreal biomass carbon stocks, rather than high allocation inputs.

Substantial differences also emerged in canopy parameters. Our framework retrieved longer leaf lifespan across the boreal biome, higher than CARDAMOM estimates by an average of 123.1% (0.82 years vs 0.37 years; Fig. S12b). This retrieval is consistent with the physiological traits of the evergreen conifers that dominate the region. In contrast to previous estimates, our results show a slightly higher fraction of NPP allocated to the canopy, which was 3.8% higher than the CARDAMOM baseline (Fig. S12a).

The reduction in allocation to wood and leaf in our retrievals was balanced by a higher allocation to fine roots compared to the CARDAMOM baseline, pointing to a larger investment in nutrient acquisition relative to structural growth than the benchmark (Fig. S12c). Our estimates of fine root turnover rates are also higher than those of CARDAMOM (Fig. S12d), although this parameter remains the less constrained by the available observation dataset. These comparisons indicate that assimilating forest biomass growth dynamics results in a functional shift in the model parameters towards a longer-lived wood and leaves, supported by a higher proportional investment in fine root allocation.

https://doi.org/10.5194/egusphere-2026-2241
Preprint. Discussion started: 8 May 2026

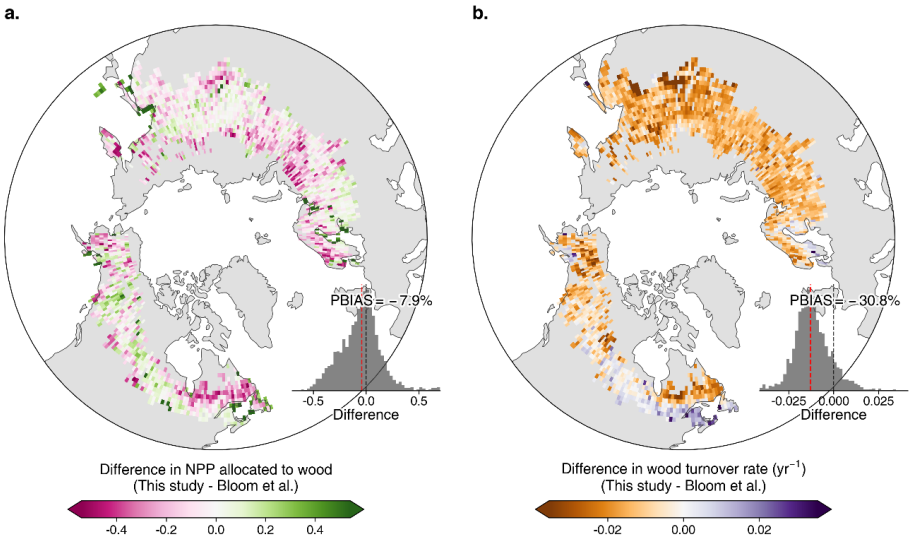

Fig. 6 Spatial differences in the calibrated (a) NPP allocated to wood and (b) the wood turnover rate between this study and the estimates from the CARDAMOM product (Bloom et al., 2016)**.** PBIAS in the insets denotes the percent bias, indicating the average relative difference between the estimates in this study and the CARDAMOM product.

4. Discussion

A major uncertainty in terrestrial biosphere modeling arises from equifinality, where different parameter combinations yield similar fits to observational data (Bloom et al., 2016; Famiglietti et al., 2021). In our baseline experiment, constrained solely by high-frequency observations (GPP, RECO, ET, and LAI), the model reproduced seasonal flux dynamics well but showed large biases in long-term biomass stocks. It is mainly because high-frequency data primarily constrain fast physiological processes, such as photosynthesis, while leaving the processes and parameters governing long-term turnover and allocation less constrained (Friend et al., 2014; Pugh et al., 2020). Specifically, without explicit constraints of accumulated biomass, the optimization algorithm tends to exploit the equifinality between carbon allocation and turnover. In our baseline experiment, the model compensated for the lack of long-term constraints by converging toward a fast-turnover strategy. This allowed the model to maintain appropriate LAI and GPP levels to match satellite observations but prevented the accumulation of woody biomass, resulting in the systematic underestimation of carbon stocks in boreal regions. Our results further demonstrate that while incorporating a snapshot of mature biomass information (single-year biomass-constrained experiment) provides a valuable anchor for carbon residence times, it remains insufficient to resolve the equifinality of internal allocation patterns. Furthermore, potential inconsistencies between independent GPP and RECO products could introduce biases in the net carbon flux (NEE). In the absence of stock constraints, such possible flux biases might propagate into long-term biomass estimates. The assimilation of forest biomass growth dynamics can effectively mitigate this uncertainty by providing an independent constraint on the net cumulative carbon.

To constrain these slower processes, we assimilated forest biomass growth dynamics derived from high-resolution satellite observations (Xu et al., 2026). Unlike instantaneous flux measurements or static biomass snapshots, these biomass growth curves reflect the cumulative outcome of carbon allocation and turnover over decades (Zhou et al.,

https://doi.org/10.5194/egusphere-2026-2241
Preprint. Discussion started: 8 May 2026

2015). Incorporating this constraint significantly reduced the error in the simulated biomass stocks, with the RMSE decreasing from 138.7 to 11.3 Mg ha^{-1}. Notably, this improvement in carbon stock estimation was achieved while maintaining similar performance for high-frequency fluxes. These findings indicate that robust model predictions require integrating data across multiple timescales. Jointly constraining fast fluxes and slow state variables helps reduce equifinality by ruling out compensating combinations of carbon input and turnover that can reproduce short-term dynamics while producing biased long-term carbon storage.

Our assimilation of forest biomass growth curves relies on the 0.5° grid-level space-for-time substitution approach, which assumes that spatial biomass gradients across different stand ages accurately represent the temporal growth trajectory of a single stand. However, in the context of rapid climate change, the conditions under which mature stands established decades ago may differ from those influencing regeneration today (Likens, 1989; Walker et al., 2010). Historical biomass growth curves might not fully capture future growth trajectories under climate change and increasing atmospheric CO_2. Furthermore, converting satellite-derived aboveground biomass to total biomass may introduce additional uncertainty. Although we employed spatially explicit root-to-shoot ratios (Huang et al., 2021), these estimates may not fully capture stand-age-dependent changes in belowground allocation. Accordingly, the belowground processes, especially the fine root turnover rates, still have great uncertainties due to the lack of effective constraints from direct observations. Moreover, the accuracy of our parameter retrieval depends on the quality of the assimilated data products. Systematic biases in satellite-derived GPP or biomass estimates would inevitably propagate into the calibrated parameters (Keenan et al., 2011). Continuous validation against independent field measurements and the integration of next-generation biomass products are therefore essential to further refine these constraints.

While the DALEC configuration assumes time-invariant allocation fractions, our assimilation of 107-year biomass growth curves effectively captures the integrated outcome of these potentially dynamic processes. In real-world boreal successions, allocation patterns may shift as stands mature (Gundersen et al., 2021; Xia et al., 2019). However, by optimizing parameters against the full accumulation trajectory, DifferLand retrieves effective mean parameters that robustly represent the long-term carbon balance of the ecosystem. This approach offers a significant advancement over baseline calibrations constrained only by high-frequency fluxes, as it ensures that the retrieved traits are consistent with observed decadal biomass stocks. Future iterations could incorporate age-dependent functions to further refine these internal dynamics without compromising the parsimony of the current differentiable framework.

Additionally, the representation of non-structural carbohydrates (NSC) in DALEC is simplified into a single labile pool that exclusively supplies foliar development. In contrast, real-world NSCs are distributed throughout various plant tissues and serve multiple functions, including maintenance metabolism, osmoregulation, and xylem embolism repair (Dietze et al., 2014; Hartmann & Trumbore, 2016; McDowell et al., 2008; Zhou et al., 2025). This structural simplification may introduce uncertainties in the retrieved parameters, as the model may compensate for the absence of a more complex NSC storage and remobilization system by adjusting carbon allocation or turnover rates to match the observed biomass trajectories.

Beyond forest biomass growth curves, additional long-term constraints could further reduce equifinality, particularly for slow soil carbon pools which were not directly constrained in this study. For example, observational constraints on soil organic carbon (SOC) stocks with chronosequence-based SOC trajectories (Walker et al., 2010) or repeated soil resampling from long-term monitoring networks (Bellamy et al., 2005; He et al., 2016) could be used to constrain litter and soil turnover parameters that are otherwise weakly informed by high-frequency fluxes. Where available, isotope-

https://doi.org/10.5194/egusphere-2026-2241
Preprint. Discussion started: 8 May 2026

or radiocarbon-informed metrics of SOC mean age and residence time can provide complementary information on long-term decomposition dynamics (He et al., 2016; Trumbore, 2000), helping to separate uncertainties in biomass turnover from those in soil carbon processes. Integrating such slow-pool constraints would enable a more complete constraint on long-term ecosystem carbon storage, because biomass and SOC jointly determine multi-decadal to centennial carbon accumulation (Schimel et al., 2015).

Although this study focuses on the boreal biome, the method of jointly constraining fast fluxes with long-term trajectory information should be transferable to other regions. In temperate forests, long-term forest inventories and management records can provide stand-age-dependent biomass trajectories similar to growth-curve constraints (Pan et al., 2011; Pugh et al., 2019). In the tropics, repeated biomass mapping (e.g., lidar- or radar-informed products) combined with disturbance chronosequences may offer similar long-term constraints (Heinrich et al., 2021; Poorter et al., 2016), although high species diversity and mixed disturbance types (e.g., selective logging, windthrow) can complicate interpretation. In semi-arid and savanna systems, long-term constraints on woody cover and carbon stocks may be particularly valuable because short-term fluxes are strongly driven by episodic rainfall and can be highly equifinal (Brandt et al., 2016; Poulter et al., 2014). Extending this framework globally will therefore require region-specific long-term datasets and disturbance histories and may motivate developing growth-curve constraints for multiple disturbance agents rather than fire alone.

Our study emphasizes that boreal forest biomass stocks are sustained by slower carbon turnover rates than those inferred from high-frequency flux constraints alone, highlighting the potential vulnerability of this carbon sink to disturbances that shorten carbon residence time. Current warming trends pose multiple threats to this stabilizing effect of long residence times. The increasing fire frequency (Magney & Pierrat, 2025) and drought-induced mortality (Peng et al., 2011) are already reducing carbon residence times. Our analysis identifies a specific non-linear threshold for turnover stability when the temperature of the coldest month rises above -20°C. This finding aligns with thermal constraints on several disturbance agents. For example, winter minimum temperatures ranging from -16°C to -20°C act as a lethal lower limit for major boreal pests, such as the southern pine beetle (*Dendroctonus frontalis*) (Lesk et al., 2017; Ungerer et al., 1999). Crossing this thermal threshold effectively lifts the climatic barrier to pest survival and northward expansion (De Grandpré et al., 2018). Furthermore, warm winters can disrupt tree cold-hardening processes, increasing the risk of frost damage (Hänninen, 2006) or winter embolism (Mayr et al., 2020). This implies that future ESM projections relying on fixed turnover rates may underestimate the risk of a rapid release of stored carbon as these bioclimatic barriers are removed. Together, these results suggest that representing climate-sensitive turnover processes with effective constraints from long-term biomass information may be critical for projecting the persistence of boreal carbon storage under future climate change.

5. Conclusion

This study demonstrates that long-term forest biomass growth trajectories provide an effective constraint on carbon allocation and turnover parameters in terrestrial biosphere models. Using a differentiable DALEC framework, we showed that calibrations based only on high-frequency observations reproduced short-term GPP, RECO, ET, and LAI dynamics but generated large biases in long-term biomass stocks. Adding a single-year biomass constraint partially reduced these biases, whereas assimilating the full biomass growth curves reduced the biomass RMSE from 138.7 to 11.3 Mg ha^{-1} without degrading the simulation of fast processes. These results show that slow ecological trajectories provide information that is not captured by flux observations or static biomass snapshots alone.

https://doi.org/10.5194/egusphere-2026-2241
Preprint. Discussion started: 8 May 2026

The growth-curve-constrained retrievals indicate that boreal forest biomass stocks are maintained primarily by long woody carbon residence times rather than by high allocation to wood. Compared with the baseline calibration, the optimized wood turnover rate decreased by 74.1%, while the fraction of NPP allocated to wood also declined. Environmental attribution further suggests that boreal carbon residence times are sensitive to climate, with wood turnover increasing sharply when the temperature of the coldest month exceeds approximately -20°C. More broadly, our results highlight the need to assimilate observations that match the timescales of the processes being constrained. Future work should integrate additional long-term constraints, such as soil carbon trajectories, repeated biomass observations, and disturbance-specific recovery curves, to improve projections of ecosystem carbon residence time and the stability of the land carbon sink under climate change.

Code and Data Availability

The code and data that support the findings of this study are openly available in Zenodo at https://zenodo.org/records/19135427 (Wu, 2026).

Author Contributions

Conceptualization: W.L. and P.C. Methodology: J.W. Investigation: J.W. Visualization: J.W. Writing—original draft: J.W. Writing—review & editing: All authors.

Acknowledgements

This study was funded by the National Natural Science Foundation of China (grant number: 42571109). This study was supported by the CALIPSO (Carbon Loss in Plant Soils and Oceans) project and the CLARiTy (Combining LAnd-use, modeling and Remote-sensing to Transform carbon budgets) project funded through the generosity of Schmidt Science. This study is supported by Center of High-performance computing, Tsinghua University.

Conflicts of Interest

The authors declare no conflicts of interest.

Reference

Bellamy, P. H., Loveland, P. J., Bradley, R. I., Lark, R. M., & Kirk, G. J. D. (2005). Carbon losses from all soils across England and Wales 1978–2003. *Nature*, *437*(7056), 245–248. https://doi.org/10.1038/nature04038

Bloom, A. A., Exbrayat, J.-F., van der Velde, I. R., Feng, L., & Williams, M. (2016). The decadal state of the terrestrial carbon cycle: Global retrievals of terrestrial carbon allocation, pools, and residence times. *Proceedings of the National Academy of Sciences*, *113*(5), 1285–1290. https://doi.org/10.1073/pnas.1515160113

Bloom, A. A., & Williams, M. (2015). Constraining ecosystem carbon dynamics in a data-limited world: Integrating ecological "common sense" in a model–data fusion framework. *Biogeosciences*, *12*(5), 1299–1315. https://doi.org/10.5194/bg-12-1299-2015

Bradbury, J., Frostig, R., Hawkins, P., Johnson, M. J., Leary, C., Maclaurin, D., Necula, G., Paszke, A., VanderPlas, J., Wanderman-Milne, S., & others. (2018). *JAX: composable transformations of Python+ NumPy programs*.

https://doi.org/10.5194/egusphere-2026-2241
Preprint. Discussion started: 8 May 2026

622 Brandt, M., Hiernaux, P., Tagesson, T., Verger, A., Rasmussen, K., Diouf, A. A., Mbow, C., Mougin, E., & Fensholt, R.
623 (2016). Woody plant cover estimation in drylands from Earth Observation based seasonal metrics. *Remote*
624 *Sensing of Environment*, *172*, 28–38. https://doi.org/10.1016/j.rse.2015.10.036

625 Braswell, B. H., Sacks, W. J., Linder, E., & Schimel, D. S. (2005). Estimating diurnal to annual ecosystem parameters by
626 synthesis of a carbon flux model with eddy covariance net ecosystem exchange observations. *Global Change*
627 *Biology*, *11*(2), 335–355. https://doi.org/10.1111/j.1365-2486.2005.00897.x

628 Butler, E. E., Datta, A., Flores-Moreno, H., Chen, M., Wythers, K. R., Fazayeli, F., Banerjee, A., Atkin, O. K., Kattge, J.,
629 Amiaud, B., Blonder, B., Boenisch, G., Bond-Lamberty, B., Brown, K. A., Byun, C., Campetella, G., Cerabolini,
630 B. E. L., Cornelissen, J. H. C., Craine, J. M., … Reich, P. B. (2017). Mapping local and global variability in plant
631 trait distributions. *Proceedings of the National Academy of Sciences*, *114*(51), E10937–E10946. (world).
632 https://doi.org/10.1073/pnas.1708984114

633 Canadell, J. G., Monteiro, P. M. S., Costa, M. H., Cotrim da Cunha, L., Cox, P. M., Eliseev, A. V., Henson, S., Ishii, M.,
634 Jaccard, S., Koven, C., Lohila, A., Patra, P. K., Piao, S., Rogelj, J., Syampungani, S., Zaehle, S., & Zickfeld, K.
635 (2021). Global Carbon and Other Biogeochemical Cycles and Feedbacks. In V. Masson-Delmotte, P. Zhai, A.
636 Pirani, S. L. Connors, C. Péan, S. Berger, N. Caud, Y. Chen, L. Goldfarb, M. I. Gomis, M. Huang, K. Leitzell, E.
637 Lonnoy, J. B. R. Matthews, T. K. Maycock, T. Waterfield, O. Yelekçi, R. Yu, & B. Zhou (Eds.), *Climate Change*
638 *2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the*
639 *Intergovernmental Panel on Climate Change* (pp. 673–816). Cambridge University Press.
640 https://doi.org/10.1017/9781009157896.007

641 De Grandpré, L., Pureswaran, D., Bouchard, M., & Kneeshaw, D. (2018). Climate-induced range shifts in boreal forest
642 pests: Ecological, economic, and social consequences. *Canadian Journal of Forest Research*, *48*(3), v–vi.
643 https://doi.org/10.1139/cjfr-2018-0058

644 Dietze, M. C., Sala, A., Carbone, M. S., Czimczik, C. I., Mantooth, J. A., Richardson, A. D., & Vargas, R. (2014).
645 Nonstructural Carbon in Woody Plants. *Annual Review of Plant Biology*, *65*(Volume 65, 2014), 667–687.
646 https://doi.org/10.1146/annurev-arplant-050213-040054

647 Dubayah, R., Blair, J. B., Goetz, S., Fatoyinbo, L., Hansen, M., Healey, S., Hofton, M., Hurtt, G., Kellner, J., Luthcke, S.,
648 Armston, J., Tang, H., Duncanson, L., Hancock, S., Jantz, P., Marselis, S., Patterson, P. L., Qi, W., & Silva, C.
649 (2020). The Global Ecosystem Dynamics Investigation: High-resolution laser ranging of the Earth's forests and
650 topography. *Science of Remote Sensing*, *1*, 100002. https://doi.org/10.1016/j.srs.2020.100002

651 Famiglietti, C. A., Smallman, T. L., Levine, P. A., Flack-Prain, S., Quetin, G. R., Meyer, V., Parazoo, N. C., Stettz, S. G.,
652 Yang, Y., Bonal, D., Bloom, A. A., Williams, M., & Konings, A. G. (2021). Optimal model complexity for
653 terrestrial carbon cycle prediction. *Biogeosciences*, *18*(8), 2727–2754. https://doi.org/10.5194/bg-18-2727-2021

654 Fan, L., Wigneron, J.-P., Ciais, P., Chave, J., Brandt, M., Fensholt, R., Saatchi, S. S., Bastos, A., Al-Yaari, A., Hufkens, K.,
655 Qin, Y., Xiao, X., Chen, C., Myneni, R. B., Fernandez-Moran, R., Mialon, A., Rodriguez-Fernandez, N. J., Kerr,
656 Y., Tian, F., & Peñuelas, J. (2019). Satellite-observed pantropical carbon dynamics. *Nature Plants*, *5*(9), 944–951.
657 https://doi.org/10.1038/s41477-019-0478-9

658 Fang, J., & Gentine, P. (2024). Exploring Optimal Complexity for Water Stress Representation in Terrestrial Carbon
659 Models: A Hybrid-Machine Learning Model Approach. *Journal of Advances in Modeling Earth Systems*, *16*(12),
660 e2024MS004308. https://doi.org/10.1029/2024MS004308

https://doi.org/10.5194/egusphere-2026-2241
Preprint. Discussion started: 8 May 2026

661 Friedlingstein, P., O'Sullivan, M., Jones, M. W., Andrew, R. M., Hauck, J., Landschützer, P., Le Quéré, C., Li, H., Luijkx,
662 I. T., Olsen, A., Peters, G. P., Peters, W., Pongratz, J., Schwingshackl, C., Sitch, S., Canadell, J. G., Ciais, P.,
663 Jackson, R. B., Alin, S. R., … Zeng, J. (2025). Global Carbon Budget 2024. *Earth System Science Data*, *17*(3),
664 965–1039. https://doi.org/10.5194/essd-17-965-2025

665 Friedman, J. H. (2001). Greedy Function Approximation: A Gradient Boosting Machine. *The Annals of Statistics*, *29*(5),
666 1189–1232.

667 Friend, A. D., Lucht, W., Rademacher, T. T., Keribin, R., Betts, R., Cadule, P., Ciais, P., Clark, D. B., Dankers, R.,
668 Falloon, P. D., Ito, A., Kahana, R., Kleidon, A., Lomas, M. R., Nishina, K., Ostberg, S., Pavlick, R., Peylin, P.,
669 Schaphoff, S., … Woodward, F. I. (2014). Carbon residence time dominates uncertainty in terrestrial vegetation
670 responses to future climate and atmospheric CO2. *Proceedings of the National Academy of Sciences*, *111*(9),
671 3280–3285. (world). https://doi.org/10.1073/pnas.1222477110

672 Gelman, A., & Rubin, D. B. (1992). Inference from Iterative Simulation Using Multiple Sequences. *Statistical Science*,
673 *7*(4), 457–472. https://doi.org/10.1214/ss/1177011136

674 Gundersen, P., Thybring, E. E., Nord-Larsen, T., Vesterdal, L., Nadelhoffer, K. J., & Johannsen, V. K. (2021). Old-growth
675 forest carbon sinks overestimated. *Nature*, *591*(7851), E21–E23. https://doi.org/10.1038/s41586-021-03266-z

676 Hänninen, H. (2006). Climate warming and the risk of frost damage to boreal forest trees: Identification of critical
677 ecophysiological traits. *Tree Physiology*, *26*(7), 889–898. https://doi.org/10.1093/treephys/26.7.889

678 Hansen, W. D., Foster, A., Gaglioti, B., Seidl, R., & Rammer, W. (2023). The Permafrost and Organic LayEr module for
679 Forest Models (POLE-FM) 1.0. *Geoscientific Model Development*, *16*(7), 2011–2036.
680 https://doi.org/10.5194/gmd-16-2011-2023

681 Harris, I. C. (2021). *CRU JRA v2. 2: A forcings dataset of gridded land surface blend of Climatic Research Unit (CRU)*
682 *and Japanese reanalysis (JRA) data; Jan. 1901–Dec. 2020*. University of East Anglia Climatic Research Unit.

683 Hartmann, H., & Trumbore, S. (2016). Understanding the roles of nonstructural carbohydrates in forest trees – from what
684 we can measure to what we want to know. *New Phytologist*, *211*(2), 386–403. https://doi.org/10.1111/nph.13955

685 He, Y., Trumbore, S. E., Torn, M. S., Harden, J. W., Vaughn, L. J. S., Allison, S. D., & Randerson, J. T. (2016).
686 Radiocarbon constraints imply reduced carbon uptake by soils during the 21st century. *Science*, *353*(6306), 1419–
687 1424. https://doi.org/10.1126/science.aad4273

688 Heinrich, V. H. A., Dalagnol, R., Cassol, H. L. G., Rosan, T. M., de Almeida, C. T., Silva Junior, C. H. L., Campanharo,
689 W. A., House, J. I., Sitch, S., Hales, T. C., Adami, M., Anderson, L. O., & Aragão, L. E. O. C. (2021). Large
690 carbon sink potential of secondary forests in the Brazilian Amazon to mitigate climate change. *Nature*
691 *Communications*, *12*(1), 1785. https://doi.org/10.1038/s41467-021-22050-1

692 Huang, Y., Ciais, P., Santoro, M., Makowski, D., Chave, J., Schepaschenko, D., Abramoff, R. Z., Goll, D. S., Yang, H.,
693 Chen, Y., Wei, W., & Piao, S. (2021). A global map of root biomass across the world's forests. *Earth System*
694 *Science Data*, *13*(9), 4263–4274. https://doi.org/10.5194/essd-13-4263-2021

695 Joiner, J., & Yoshida, Y. (2020). Satellite-based reflectances capture large fraction of variability in global gross primary
696 production (GPP) at weekly time scales. *Agricultural and Forest Meteorology*, *291*, 108092.
697 https://doi.org/10.1016/j.agrformet.2020.108092

https://doi.org/10.5194/egusphere-2026-2241
Preprint. Discussion started: 8 May 2026

Joiner, J., & Yoshida, Y. (2021). Global MODIS and FLUXNET-derived Daily Gross Primary Production, V2, ORNL DAAC, Oak Ridge, Tennessee, USA. *Earth Data [Data Set]*, *10*.

Joiner, J., Yoshida, Y., Zhang, Y., Duveiller, G., Jung, M., Lyapustin, A., Wang, Y., & Tucker, C. J. (2018). Estimation of Terrestrial Global Gross Primary Production (GPP) with Satellite Data-Driven Models and Eddy Covariance Flux Data. *Remote Sensing*, *10*(9), 1346. https://doi.org/10.3390/rs10091346

Jonsson, M., & Wardle, D. A. (2009). Structural equation modelling reveals plant-community drivers of carbon storage in boreal forest ecosystems. *Biology Letters*, *6*(1), 116–119. https://doi.org/10.1098/rsbl.2009.0613

Kaufhold, C., Willeit, M., Talento, S., Ganopolski, A., & Rockström, J. (2025). Interplay between climate and carbon cycle feedbacks could substantially enhance future warming. *Environmental Research Letters*, *20*(4), 044027. https://doi.org/10.1088/1748-9326/adb6be

Keenan, T. F., Carbone, M. S., Reichstein, M., & Richardson, A. D. (2011). The model–data fusion pitfall: Assuming certainty in an uncertain world. *Oecologia*, *167*(3), 587–597. https://doi.org/10.1007/s00442-011-2106-x

Keenan, T. F., Davidson, E., Moffat, A. M., Munger, W., & Richardson, A. D. (2012). Using model-data fusion to interpret past trends, and quantify uncertainties in future projections, of terrestrial ecosystem carbon cycling. *Global Change Biology*, *18*(8), 2555–2569. https://doi.org/10.1111/j.1365-2486.2012.02684.x

Koehler, T., Wankmüller, F. J. P., Sadok, W., & Carminati, A. (2023). Transpiration response to soil drying versus increasing vapor pressure deficit in crops: Physical and physiological mechanisms and key plant traits. *Journal of Experimental Botany*, *74*(16), 4789–4807. https://doi.org/10.1093/jxb/erad221

Kuppel, S., Peylin, P., Maignan, F., Chevallier, F., Kiely, G., Montagnani, L., & Cescatti, A. (2014). Model–data fusion across ecosystems: From multisite optimizations to global simulations. *Geoscientific Model Development*, *7*(6), 2581–2597. https://doi.org/10.5194/gmd-7-2581-2014

Lesk, C., Coffel, E., D'Amato, A. W., Dodds, K., & Horton, R. (2017). Threats to North American forests from southern pine beetle with warming winters. *Nature Climate Change*, *7*(10), 713–717. https://doi.org/10.1038/nclimate3375

Leys, C., Ley, C., Klein, O., Bernard, P., & Licata, L. (2013). Detecting outliers: Do not use standard deviation around the mean, use absolute deviation around the median. *Journal of Experimental Social Psychology*, *49*(4), 764–766. https://doi.org/10.1016/j.jesp.2013.03.013

Li, W., MacBean, N., Ciais, P., Defourny, P., Lamarche, C., Bontemps, S., Houghton, R. A., & Peng, S. (2018). Gross and net land cover changes in the main plant functional types derived from the annual ESA CCI land cover maps (1992–2015). *Earth System Science Data*, *10*(1), 219–234. https://doi.org/10.5194/essd-10-219-2018

Likens, G. E. (Ed.). (1989). *Long-Term Studies in Ecology*. Springer. https://doi.org/10.1007/978-1-4615-7358-6

Lin, M. (2024). Automatic Functional Differentiation in JAX. In B. Kim, Y. Yue, S. Chaudhuri, K. Fragkiadaki, M. Khan, & Y. Sun (Eds.), *International Conference on Representation Learning* (Vol. 2024, pp. 8139–8159). https://proceedings.iclr.cc/paper_files/paper/2024/file/20d50ff26f7ab9c8c2578a3eec69c922-Paper-Conference.pdf

Long, T., Zhang, Z., He, G., Jiao, W., Tang, C., Wu, B., Zhang, X., Wang, G., & Yin, R. (2019). 30 m Resolution Global Annual Burned Area Mapping Based on Landsat Images and Google Earth Engine. *Remote Sensing*, *11*(5), 489. https://doi.org/10.3390/rs11050489

https://doi.org/10.5194/egusphere-2026-2241
Preprint. Discussion started: 8 May 2026

López, J., Way, D. A., & Sadok, W. (2021). Systemic effects of rising atmospheric vapor pressure deficit on plant physiology and productivity. *Global Change Biology*, *27*(9), 1704–1720. https://doi.org/10.1111/gcb.15548

Luo, X., Zhao, R., Chu, H., Collalti, A., Fatichi, S., Keenan, T. F., Lu, X., Nguyen, N., Prentice, I. C., Sun, W., Yu, K., & Yu, L. (2025). Global variation in vegetation carbon use efficiency inferred from eddy covariance observations. *Nature Ecology & Evolution*, *9*(8), 1414–1425. https://doi.org/10.1038/s41559-025-02753-0

Ma, H., & Liang, S. (2022). Development of the GLASS 250-m leaf area index product (version 6) from MODIS data using the bidirectional LSTM deep learning model. *Remote Sensing of Environment*, *273*, 112985. https://doi.org/10.1016/j.rse.2022.112985

Magney, T. S., & Pierrat, Z. A. (2025). Shifting equilibria in a warming boreal forest. *Proceedings of the National Academy of Sciences*, *122*(3), e2424669122. https://doi.org/10.1073/pnas.2424669122

Massmann, A., Gentine, P., & Lin, C. (2019). When Does Vapor Pressure Deficit Drive or Reduce Evapotranspiration? *Journal of Advances in Modeling Earth Systems*, *11*(10), 3305–3320. https://doi.org/10.1029/2019MS001790

Matimati, I., Verboom, G. A., & Cramer, M. D. (2014). Nitrogen regulation of transpiration controls mass-flow acquisition of nutrients. *Journal of Experimental Botany*, *65*(1), 159–168. https://doi.org/10.1093/jxb/ert367

Mayr, S., Schmid, P., Beikircher, B., Feng, F., & Badel, E. (2020). Die hard: Timberline conifers survive annual winter embolism. *New Phytologist*, *226*(1), 13–20. https://doi.org/10.1111/nph.16304

McDowell, N., Pockman, W. T., Allen, C. D., Breshears, D. D., Cobb, N., Kolb, T., Plaut, J., Sperry, J., West, A., Williams, D. G., & Yepez, E. A. (2008). Mechanisms of plant survival and mortality during drought: Why do some plants survive while others succumb to drought? *New Phytologist*, *178*(4), 719–739. https://doi.org/10.1111/j.1469-8137.2008.02436.x

Meroni, M., Fasbender, D., Lopez-Lozano, R., & Migliavacca, M. (2019). Assimilation of Earth Observation Data Over Cropland and Grassland Sites into a Simple GPP Model. *Remote Sensing*, *11*(7), 749. https://doi.org/10.3390/rs11070749

Miralles, D. G., Bonte, O., Koppa, A., Baez-Villanueva, O. M., Tronquo, E., Zhong, F., Beck, H. E., Hulsman, P., Dorigo, W., Verhoest, N. E. C., & Haghdoost, S. (2025). GLEAM4: Global land evaporation and soil moisture dataset at 0.1° resolution from 1980 to near present. *Scientific Data*, *12*(1), 416. https://doi.org/10.1038/s41597-025-04610-y

Miralles, D. G., De Jeu, R. a. M., Gash, J. H., Holmes, T. R. H., & Dolman, A. J. (2011). Magnitude and variability of land evaporation and its components at the global scale. *Hydrology and Earth System Sciences*, *15*(3), 967–981. https://doi.org/10.5194/hess-15-967-2011

Neigh, C. S. R., Montesano, P. M., Sexton, J. O., Wooten, M., Wagner, W., Feng, M., Carvalhais, N., Calle, L., & Carroll, M. L. (2025). Russian forests show strong potential for young forest growth. *Communications Earth & Environment*, *6*(1), 71. https://doi.org/10.1038/s43247-025-02006-9

Pan, Y., Birdsey, R. A., Fang, J., Houghton, R., Kauppi, P. E., Kurz, W. A., Phillips, O. L., Shvidenko, A., Lewis, S. L., Canadell, J. G., Ciais, P., Jackson, R. B., Pacala, S. W., McGuire, A. D., Piao, S., Rautiainen, A., Sitch, S., & Hayes, D. (2011). A Large and Persistent Carbon Sink in the World's Forests. *Science*, *333*(6045), 988–993. https://doi.org/10.1126/science.1201609

https://doi.org/10.5194/egusphere-2026-2241
Preprint. Discussion started: 8 May 2026

771 Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R.,
772 Dubourg, V., & others. (2011). Scikit-learn: Machine learning in Python. *The Journal of Machine Learning*
773 *Research*, *12*, 2825–2830.

774 Peng, C., Ma, Z., Lei, X., Zhu, Q., Chen, H., Wang, W., Liu, S., Li, W., Fang, X., & Zhou, X. (2011). A drought-induced
775 pervasive increase in tree mortality across Canada's boreal forests. *Nature Climate Change*, *1*(9), 467–471.
776 https://doi.org/10.1038/nclimate1293

777 Peylin, P., Bacour, C., MacBean, N., Leonard, S., Rayner, P., Kuppel, S., Koffi, E., Kane, A., Maignan, F., Chevallier, F.,
778 Ciais, P., & Prunet, P. (2016). A new stepwise carbon cycle data assimilation system using multiple data streams
779 to constrain the simulated land surface carbon cycle. *Geoscientific Model Development*, *9*(9), 3321–3346.
780 https://doi.org/10.5194/gmd-9-3321-2016

781 Poorter, L., Bongers, F., Aide, T. M., Almeyda Zambrano, A. M., Balvanera, P., Becknell, J. M., Boukili, V., Brancalion,
782 P. H. S., Broadbent, E. N., Chazdon, R. L., Craven, D., de Almeida-Cortez, J. S., Cabral, G. A. L., de Jong, B. H.
783 J., Denslow, J. S., Dent, D. H., DeWalt, S. J., Dupuy, J. M., Durán, S. M., … Rozendaal, D. M. A. (2016).
784 Biomass resilience of Neotropical secondary forests. *Nature*, *530*(7589), 211–214.
785 https://doi.org/10.1038/nature16512

786 Poulter, B., Frank, D., Ciais, P., Myneni, R. B., Andela, N., Bi, J., Broquet, G., Canadell, J. G., Chevallier, F., Liu, Y. Y.,
787 Running, S. W., Sitch, S., & van der Werf, G. R. (2014). Contribution of semi-arid ecosystems to interannual
788 variability of the global carbon cycle. *Nature*, *509*(7502), 600–603. https://doi.org/10.1038/nature13376

789 Pugh, T. A. M., Lindeskog, M., Smith, B., Poulter, B., Arneth, A., Haverd, V., & Calle, L. (2019). Role of forest regrowth
790 in global carbon sink dynamics. *Proceedings of the National Academy of Sciences*, *116*(10), 4382–4387.
791 https://doi.org/10.1073/pnas.1810512116

792 Pugh, T. A. M., Rademacher, T., Shafer, S. L., Steinkamp, J., Barichivich, J., Beckage, B., Haverd, V., Harper, A., Heinke,
793 J., Nishina, K., Rammig, A., Sato, H., Arneth, A., Hantson, S., Hickler, T., Kautz, M., Quesada, B., Smith, B., &
794 Thonicke, K. (2020). Understanding the uncertainty in global forest carbon turnover. *Biogeosciences*, *17*(15),
795 3961–3989. https://doi.org/10.5194/bg-17-3961-2020

796 Reich, P. B., Luo, Y., Bradford, J. B., Poorter, H., Perry, C. H., & Oleksyn, J. (2014). Temperature drives global patterns in
797 forest biomass distribution in leaves, stems, and roots. *Proceedings of the National Academy of Sciences*, *111*(38),
798 13721–13726. https://doi.org/10.1073/pnas.1216053111

799 Rousseeuw, P. J., & Croux, C. (1993). Alternatives to the Median Absolute Deviation. *Journal of the American Statistical*
800 *Association*, *88*(424), 1273–1283. https://doi.org/10.2307/2291267

801 Santoro, M., & Cartus, O. (2023). *ESA Biomass Climate Change Initiative (Biomass_cci): Global datasets of forest above-*
802 *ground biomass for the years 2010, 2017, 2018, 2019 and 2020, v4*. NERC EDS Centre for Environmental Data
803 Analysis. https://doi.org/10.5285/AF60720C1E404A9E9D2C145D2B2EAD4E

804 Schimel, D., Stephens, B. B., & Fisher, J. B. (2015). Effect of increasing CO2 on the terrestrial carbon cycle. *Proceedings*
805 *of the National Academy of Sciences*, *112*(2), 436–441. https://doi.org/10.1073/pnas.1407302112

806 Schuur, T. (2019). Permafrost and the Global Carbon Cycle. In J. Richter-Menge, M. L. Druckenmiller, & M. Jeffries
807 (Eds.), *Arctic Report Card 2019* (pp. 58–65). NOAA. https://www.arctic.noaa.gov/Report-Card

https://doi.org/10.5194/egusphere-2026-2241
Preprint. Discussion started: 8 May 2026

Seneviratne, S. I., Corti, T., Davin, E. L., Hirschi, M., Jaeger, E. B., Lehner, I., Orlowsky, B., & Teuling, A. J. (2010). Investigating soil moisture–climate interactions in a changing climate: A review. *Earth-Science Reviews*, *99*(3), 125–161. https://doi.org/10.1016/j.earscirev.2010.02.004

Shangguan, W., Dai, Y., Duan, Q., Liu, B., & Yuan, H. (2014). A global soil data set for earth system modeling. *Journal of Advances in Modeling Earth Systems*, *6*(1), 249–263. https://doi.org/10.1002/2013MS000293

Shchepashchenko, D., & Ciais, P. (2026). *Updated report on existing biomass and biomass change datasets*.

Shi, Z., Hu, Z., Lin, M., & Kawaguchi, K. (2024). Stochastic Taylor Derivative Estimator: Efficient amortization for arbitrary differential operators. In A. Globerson, L. Mackey, D. Belgrave, A. Fan, U. Paquet, J. Tomczak, & C. Zhang (Eds.), *Advances in Neural Information Processing Systems* (Vol. 37, pp. 122316–122353). Curran Associates, Inc. https://doi.org/10.52202/079017-3887

Thum, T., MacBean, N., Peylin, P., Bacour, C., Santaren, D., Longdoz, B., Loustau, D., & Ciais, P. (2017). The potential benefit of using forest biomass data in addition to carbon and water flux measurements to constrain ecosystem model parameters: Case studies at two temperate forest sites. *Agricultural and Forest Meteorology*, *234–235*, 48–65. https://doi.org/10.1016/j.agrformet.2016.12.004

Thurner, M., Beer, C., Carvalhais, N., Forkel, M., Santoro, M., Tum, M., & Schmullius, C. (2016). Large-scale variation in boreal and temperate forest carbon turnover rate related to climate. *Geophysical Research Letters*, *43*(9), 4576–4585. https://doi.org/10.1002/2016GL068794

Trumbore, S. (2000). Age of Soil Organic Matter and Soil Respiration: Radiocarbon Constraints on Belowground C Dynamics. *Ecological Applications*, *10*(2), 399–411. https://doi.org/10.1890/1051-0761(2000)010%255B0399:AOSOMA%255D2.0.CO;2

Ungerer, M. J., Ayres, M. P., & Lombardero, M. J. (1999). Climate and the northern distribution limits of Dendroctonus frontalis Zimmerman (Coleoptera: Scolytidae). *Journal of Biogeography 26:1133-1145*. https://research.fs.usda.gov/treesearch/1327

Walker, L. R., Wardle, D. A., Bardgett, R. D., & Clarkson, B. D. (2010). The use of chronosequences in studies of ecological succession and soil development. *Journal of Ecology*, *98*(4), 725–736. https://doi.org/10.1111/j.1365-2745.2010.01664.x

Wang, H., Colin Prentice, I., J. Wright, I., I. Warton, D., Qiao, S., Xu, X., Zhou, J., Kikuzawa, K., & Chr. Stenseth, N. (2023). Leaf economics fundamentals explained by optimality principles. *Science Advances*. (world). https://doi.org/10.1126/sciadv.add5667

Wang, H., Li, X., Ma, M., & Geng, L. (2019). Improving Estimation of Gross Primary Production in Dryland Ecosystems by a Model-Data Fusion Approach. *Remote Sensing*, *11*(3), 225. https://doi.org/10.3390/rs11030225

Way, D. A., & Oren, R. (2010). Differential responses to changes in growth temperature between trees from different functional groups and biomes: A review and synthesis of data. *Tree Physiology*, *30*(6), 669–688. https://doi.org/10.1093/treephys/tpq015

Williams, M., Rastetter, E. B., Fernandes, D. N., Goulden, M. L., Shaver, G. R., & Johnson, L. C. (1997). Predicting Gross Primary Productivity in Terrestrial Ecosystems. *Ecological Applications*, *7*(3), 882–894. https://doi.org/10.1890/1051-0761(1997)007%255B0882:PGPPIT%255D2.0.CO;2

https://doi.org/10.5194/egusphere-2026-2241
Preprint. Discussion started: 8 May 2026

Wright, I. J., Reich, P. B., Westoby, M., Ackerly, D. D., Baruch, Z., Bongers, F., Cavender-Bares, J., Chapin, T., Cornelissen, J. H. C., Diemer, M., Flexas, J., Garnier, E., Groom, P. K., Gulias, J., Hikosaka, K., Lamont, B. B., Lee, T., Lee, W., Lusk, C., … Villar, R. (2004). The worldwide leaf economics spectrum. *Nature*, *428*(6985), 821–827. https://doi.org/10.1038/nature02403

Wu, J. (2026). *Supplementary data and code for "Constraining boreal carbon allocation and turnover by assimilating forest growth dynamics in a differentiable framework"* [Dataset]. Zenodo. https://doi.org/10.5281/zenodo.19135427

Xia, J., Yuan, W., Lienert, S., Joos, F., Ciais, P., Viovy, N., Wang, Y., Wang, X., Zhang, H., Chen, Y., & Tian, X. (2019). Global Patterns in Net Primary Production Allocation Regulated by Environmental Conditions and Forest Stand Age: A Model-Data Comparison. *Journal of Geophysical Research: Biogeosciences*, *124*(7), 2039–2059. https://doi.org/10.1029/2018JG004777

Xu, Y., Ciais, P., & Li, W. (2026). *Post-fire biomass recovery in boreal forest during the last three decades*.

Zaehle, S. (2013). Terrestrial nitrogen–carbon cycle interactions at the global scale. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *368*(1621), 20130125. https://doi.org/10.1098/rstb.2013.0125

Zeng, J., Matsunaga, T., Tan, Z.-H., Saigusa, N., Shirai, T., Tang, Y., Peng, S., & Fukuda, Y. (2020). Global terrestrial carbon fluxes of 1999–2019 estimated by upscaling eddy covariance data with a random forest. *Scientific Data*, *7*(1), 313. https://doi.org/10.1038/s41597-020-00653-5

Zheng, Y., Shen, R., Wang, Y., Li, X., Liu, S., Liang, S., Chen, J. M., Ju, W., Zhang, L., & Yuan, W. (2020). Improved estimate of global gross primary production for reproducing its long-term variation, 1982–2017. *Earth System Science Data*, *12*(4), 2725–2746. https://doi.org/10.5194/essd-12-2725-2020

Zhou, T., Shi, P., Jia, G., Dai, Y., Zhao, X., Shangguan, W., Du, L., Wu, H., & Luo, Y. (2015). Age-dependent forest carbon sink: Estimation via inverse modeling. *Journal of Geophysical Research: Biogeosciences*, *120*(12), 2473–2492. https://doi.org/10.1002/2015JG002943

Zhou, W., Zhang, J., Yin, B., Peng, L., Wang, L., Zhou, X., Zhou, Y., Di, Y., Zheng, H., & Zhang, Y. (2025). Mapping spatial heterogeneity of non-structural carbohydrates in *Haloxylon ammodendron* using remote sensing in extreme desert environments. *Plant Stress*, *15*, 100790. https://doi.org/10.1016/j.stress.2025.100790