Supplementary information

This file includes:

Supplementary Text

Text S1 Derivation of Derived Ecological Metrics from Model Parameters

Text S1.1 Mathematical formulation of leaf phenology and realized leaf lifespan

In the DALEC model configurations used in this study (the "$\beta - \mathrm{JS}$" setup), leaf phenology is simulated by the Combined Deciduous-Evergreen Analytical (CEDA) scheme. According to the model formulation, the instantaneous leaf fall factor $\Phi_{\mathrm{fall}}(t)$, which defines the daily fraction of the foliar carbon pool turned over at day t, is calculated as:

$$\Phi_{\mathrm{fall}}(t) = \frac{\sqrt{2}}{\sqrt{\pi} \times c_{\mathrm{rfall}}} \left(-\ln\left(1 - \frac{1}{\tau_{\mathrm{foliar}}}\right)\right) \times e^{-\left(\frac{365.25 \times \sqrt{2}}{\pi \times c_{\mathrm{rfall}}} \times \sin\left(\frac{\pi(t - d_{\mathrm{fall}} + \Psi)}{365.25}\right)\right)^2} \quad \text{(S1)}$$

where c_{rfall} controls the duration of the leaf fall period, τ_{foliar} controls the amplitude of the function, d_{fall} determines the day of year corresponding to the peak leaf fall rate, Ψ is a phase offset function dependent on c_{rfall} and τ_{foliar} (Fig. S1).

The true annual turnover fraction (Φ_{annual}) is then derived by integrating the instantaneous rate over an annual cycle:

$$\Phi_{\mathrm{annual}} = \int_0^{365.25} \Phi_{\mathrm{fall}}(t)\, \mathrm{dt} \quad \text{(S2)}$$

and the realized leaf lifespan (τ_{real}) is defined as the inverse of this integrated turnover:

$$\tau_{\mathrm{real}} = \frac{1}{\Phi_{\mathrm{annual}}} \quad \text{(S3)}$$

In this study, all reported "leaf lifespan" refer to τ_{real}, providing a more accurate representation the mean residence time of foliar carbon than the raw parameter τ_{foliar}.

Text S1.2 Net Primary Productivity (NPP) Allocation Fractions

The DALEC configuration used in this study employs a hierarchical allocation scheme to partition NPP into the four live pools: foliar, labile, fine roots, and wood. This approach ensures that the sum of allocation fractions is always 100%. We derived the allocation percentages for the canopy (A_{leaf}), fine roots (A_{root}), and wood (A_{wood}) from the raw model parameters (f_{lab}, f_{fol}, f_{root}) as follows:

$$A_{\mathrm{leaf}} = \left[f_{\mathrm{fol}} + f_{\mathrm{lab}} \times \left(1 - f_{\mathrm{fol}}\right) \right] \times 100\% \quad \text{(S4)}$$

$$A_{\mathrm{root}} = \left(100\% - A_{\mathrm{leaf}}\right) \times f_{\mathrm{root}} \quad \text{(S5)}$$

$$A_{\mathrm{wood}} = 100\% - A_{\mathrm{leaf}} - A_{\mathrm{root}} \quad \text{(S6)}$$

Text S2 Formulation of the Normalized Nash-Sutcliffe Efficiency

To balance the contribution of multiple data output with different units and magnitudes, we utilized the Normalized Nash-Sutcliffe Efficiency (NNSE) as the objective function for all observational constraints.

The standard Nash-Sutcliffe Efficiency (NSE) is defined as (Nash & Sutcliffe, 1970):

$$NSE = 1 - \frac{\sum_{t=1}^{N} \left(O_t - M_t\right)^2}{\sum_{t=1}^{N} \left(O_t - \overline{O}\right)^2} \quad \text{(S7)}$$

where O_t is the observed value, M_t is the modeled value, and $\bar{O}$ is the mean of observations. Since NSE ranges from $-\infty$ to 1, extremely poor model performance (large negative NSE) can cause unstable gradients during optimization. The NNSE transforms the NSE into a bounded range of (0, 1]:

$$NNSE = \frac{1}{2 - NSE} \tag{S8}$$

In our loss function, we maximize NNSE (or minimize –NNSE). This formulation ensures that the loss for each data output is bounded between -1 (perfect fit) and 0 (infinitely poor fit), preventing any single data output from dominating the joint optimization process solely due to scaling issues.

Text S3 Independent Validation using Field-based Allocation Data

To independently validate the calibrated parameters, we utilized a global dataset of forest biomass allocation derived from field measurements (Xia et al., 2019). Since the field data are point-based, we aggregated all plots falling within each 0.5° × 0.5° grid cell to match the spatial resolution of our model outputs.

Establishing a consistent benchmark for these allocation data required reference forest growth trajectories. As the biomass growth curves used in this study (Xu et al., 2026) are specific to the boreal zone, we incorporated additional datasets to characterize growth potentials for validation sites located in other regions. We utilized the forest recovery curves from (Ritter et al., 2025) for sites within the non-boreal regions of Europe and the spatially explicit growth curves from (Robinson et al., 2025) for other non-boreal regions. Because these external datasets differ in spatial resolution (18 km and 1 km, respectively) from our model grid, we applied an aggregation procedure to generate compatible constraints. For each 0.5° grid cell containing validation points, we calculated biomass accumulation trajectories (0–100 years) for all underlying high-resolution pixels and averaged them to produce a single, representative biomass growth curve.

Supplementary Figures

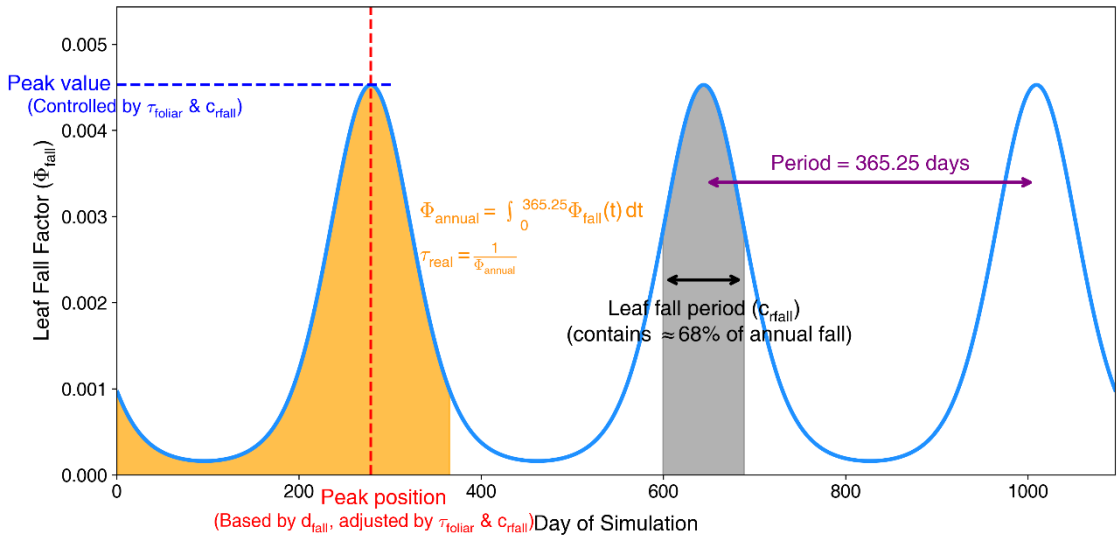

Fig. S1 Schematic illustration of the seasonal leaf fall dynamics and the derivation of the realized leaf lifespan. The blue curve represents the instantaneous leaf fall factor (Φ_{fall}) simulated by the CEDA scheme over time. The shape of the curve is governed by key model parameters: the peak timing is determined by d_{fall}, the duration of the fall period is controlled by c_{rfall} (gray shaded region), and the amplitude is adjusted by τ_{foliar}. The true annual turnover fraction (Φ_{annual}, orange shaded area) is obtained by integrating Φ_{fall} over a full annual cycle (365.25 days). The realized leaf lifespan (τ_{real}) is then defined as the inverse of this integrated turnover (Φ_{annual}).

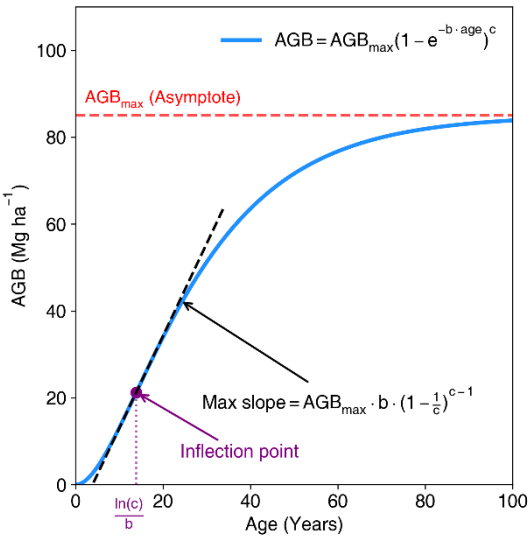

Fig. S2 Visualization of key parameters in the Richards-Chapman growth function. The plot illustrates the geometric correspondences of the maximum asymptotic capacity (AGB_{max}), growth rate parameter (b), and shape parameter (c) on the growth curve.

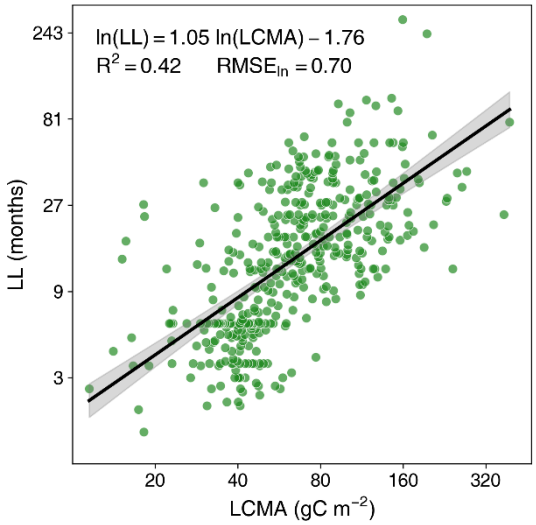

Fig. S3 The log-linear relationship between leaf lifespan (LL) and leaf carbon mass per area (LCMA). Dots represent the compiled measurements from the global leaf trait dataset (Wang et al., 2023). The red line indicates the fitted linear regression on log-transformed data, with the gray shaded area representing the 95% confidence interval. The goodness of fit is indicated by an R^2 of 0.42 and an RMSE of 0.70 (on log-transformed data). This functional relationship serves as a prior constraint in the model calibration to enforce the physiological trade-off between leaf longevity and construction cost.

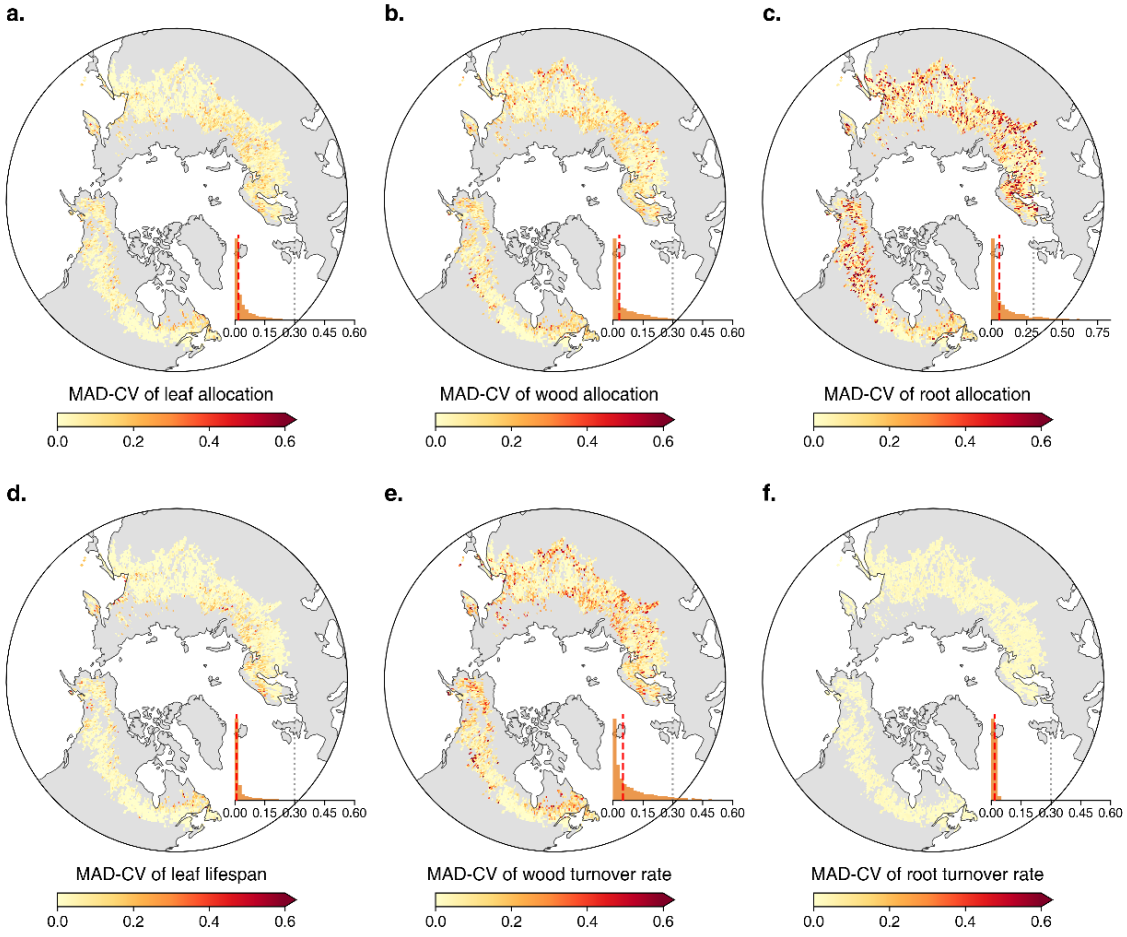

Fig. S4 Spatial distributions of the robust coefficient of variation (MAD-CV) for six key parameters under the growth-curve-constrained scenario. MAD-CV = 1.4826×MAD / Median, where MAD is the median absolute deviation.

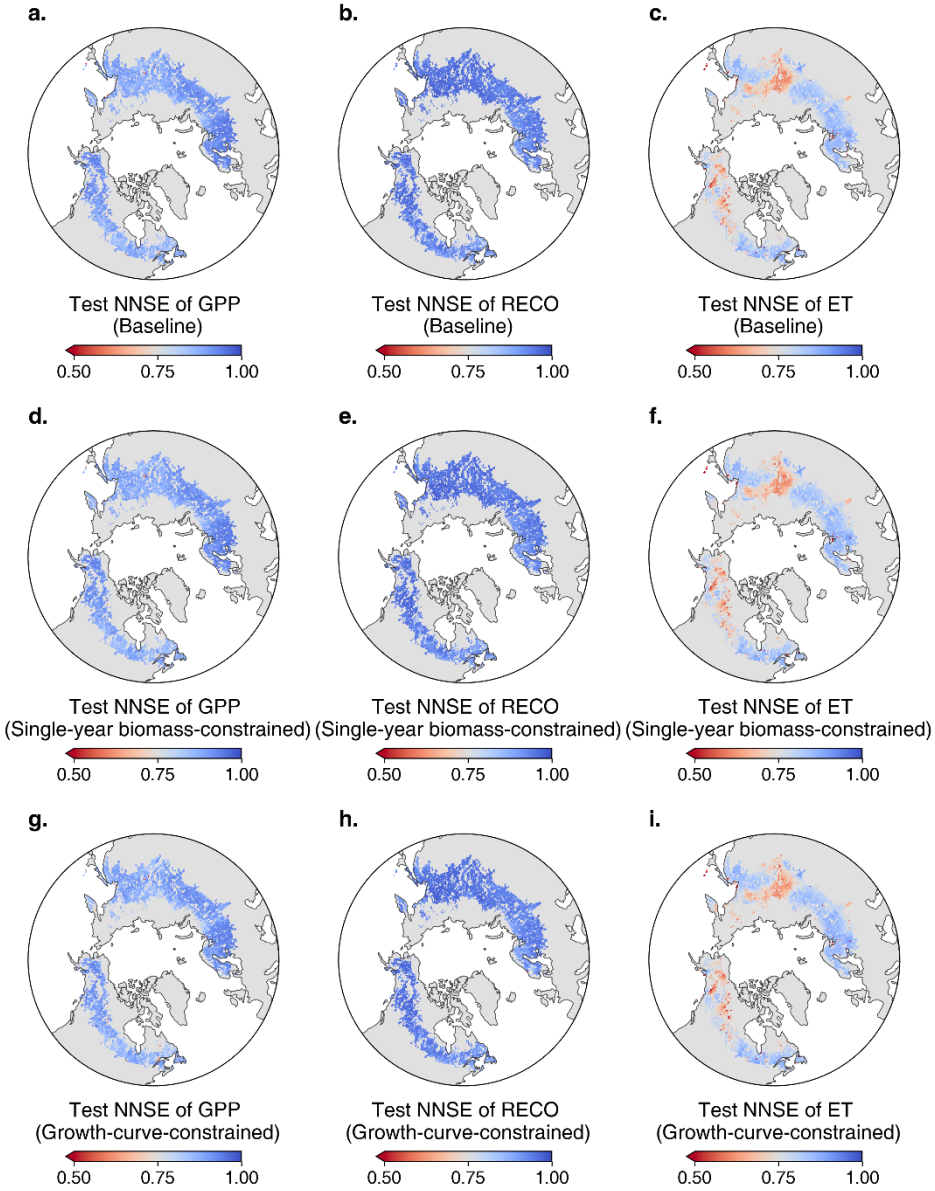

Fig. S5 Evaluation of model performance for model fast processes outputs. Spatial distribution of the Normalized Nash-Sutcliffe Efficiency (NNSE) on the independent test set (2013–2019) for GPP, RECO, and ET. (**a–c**) Results from the baseline experiment. (**d–f**) Results from the single-year biomass-constrained experiment. (**g–i**) Results from the growth-curve-constrained experiment. High NNSE values (approaching 1, shown in blue) indicate high model accuracy. The similarity across all three rows demonstrates that assimilating slow-process constraints (single-year biomass or forest recovery curves) does not compromise the ability of the model to reproduce fast-process dynamics.

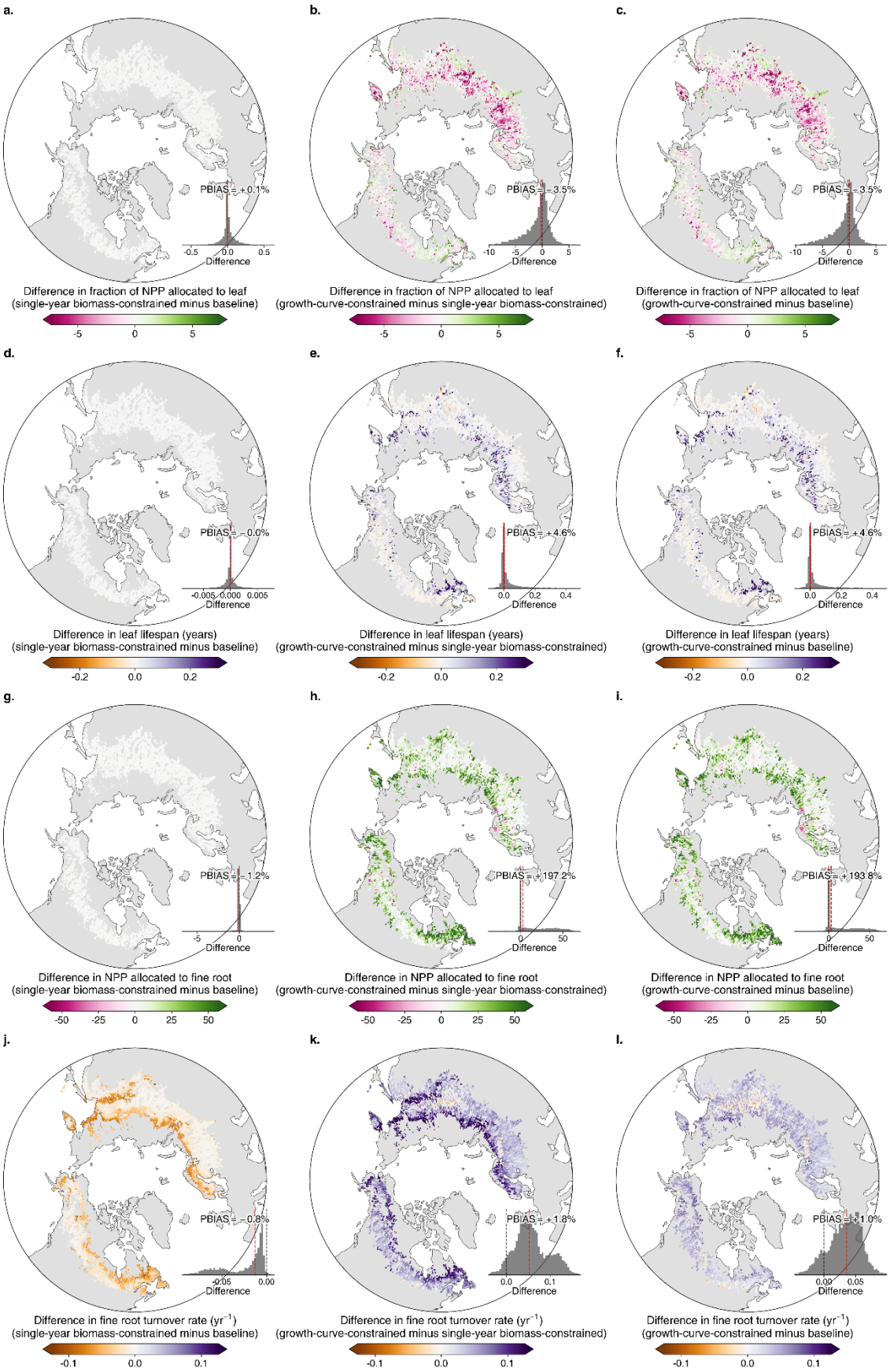

Fig. S6 Changes in leaf and fine root carbon turnover allocation across different biomass constraints. Spatial difference maps showing the changes in calibrated parameters among the baseline, single-year biomass-constrained, and growth-curve-constrained experiments. Difference in (**a–c**) the fraction of NPP allocated to leaf, (**d–f**) leaf lifespan, (**g–i**) the fraction of NPP allocated to fine roots, and (**j–l**) fine root turnover rate.

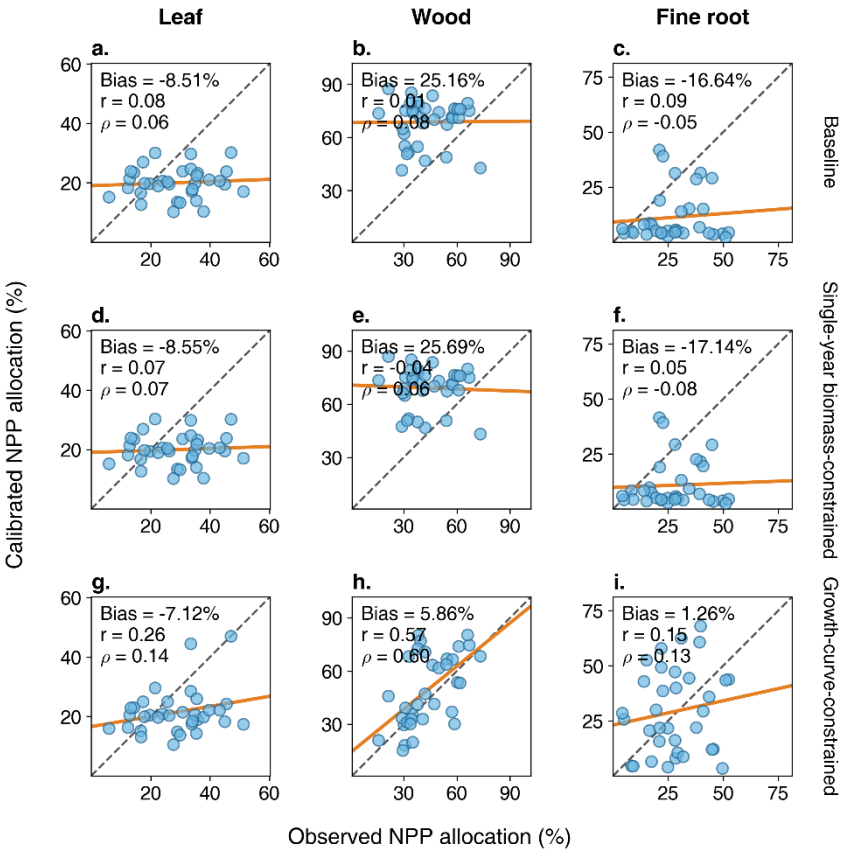

Fig. S7 Independent validation of calibrated NPP allocation fractions against field measurements. Comparison of the calibrated NPP allocation fractions against a global dataset of field measurements (Xia et al., 2019) for (**a, d, g**) leaf, (**b, e, h**) wood, and (**c, f, i**) fine root. (**a–c**) Results from the baseline experiment. (**d–f**) Results from the single-year biomass-constrained experiment. (**g-i**) Results from the growth-curve-constrained experiment. Performance is quantified using mean bias (average difference between calibrated and observed value), Pearson correlation coefficient (r), and Spearman's rank correlation coefficient (ρ).

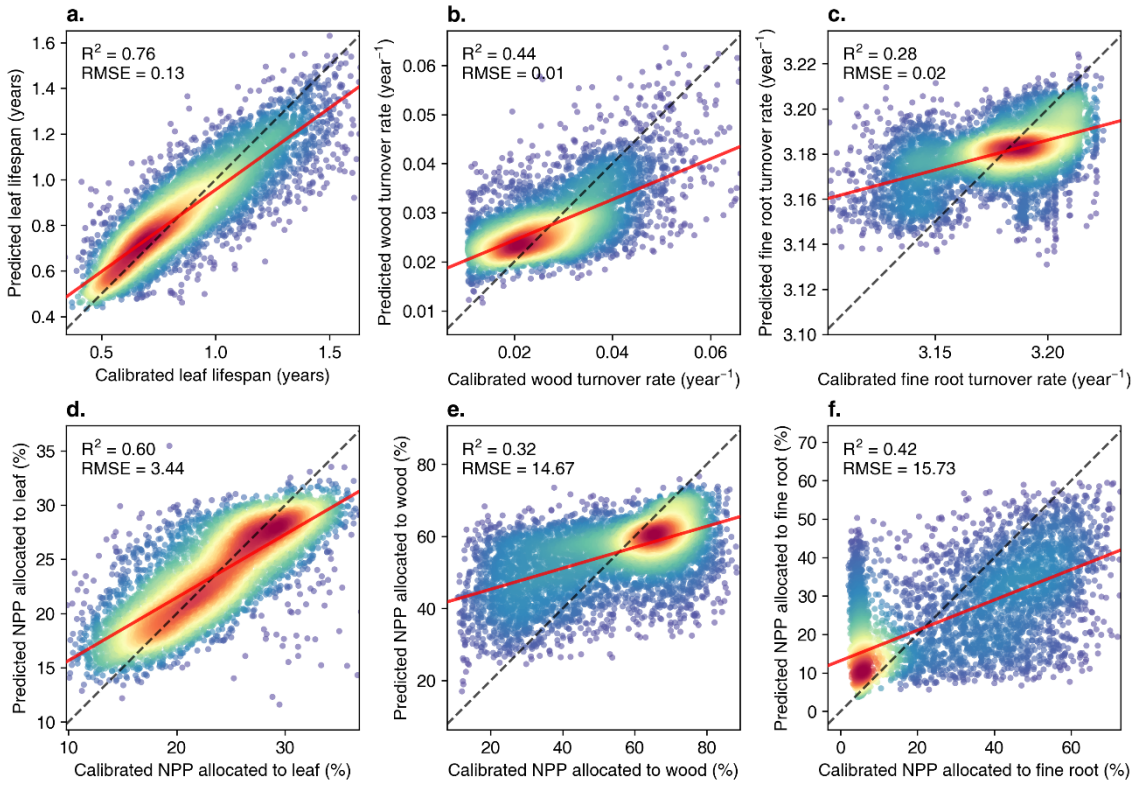

Fig. S8 Predictive performance of machine learning models for the calibrated ecological parameters. Scatter plots comparing the calibrated results (x-axis, derived from the DifferLand assimilation) against the predictions made by the gradient boosting regressor models (y-axis) based on climate, soil, and vegetation drivers. The comparisons include (**a**) leaf lifespan, (**b**) wood turnover rate, (**c**) fine root turnover rate, and the fraction of NPP allocated to (**d**) leaf, (**e**) wood, and (**f**) fine root.

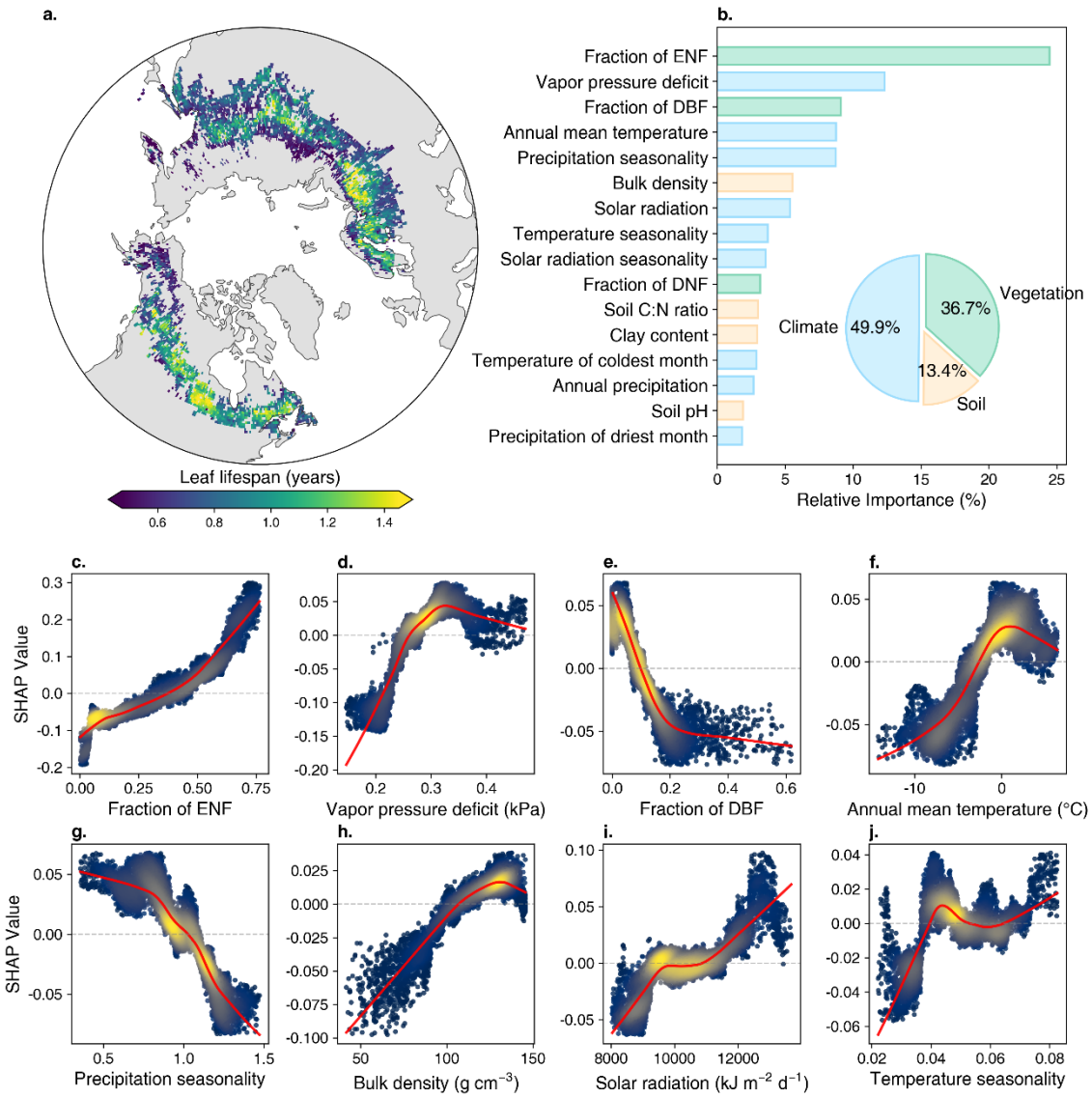

Fig. S9 Same as Fig. 3 but for the leaf lifespan.

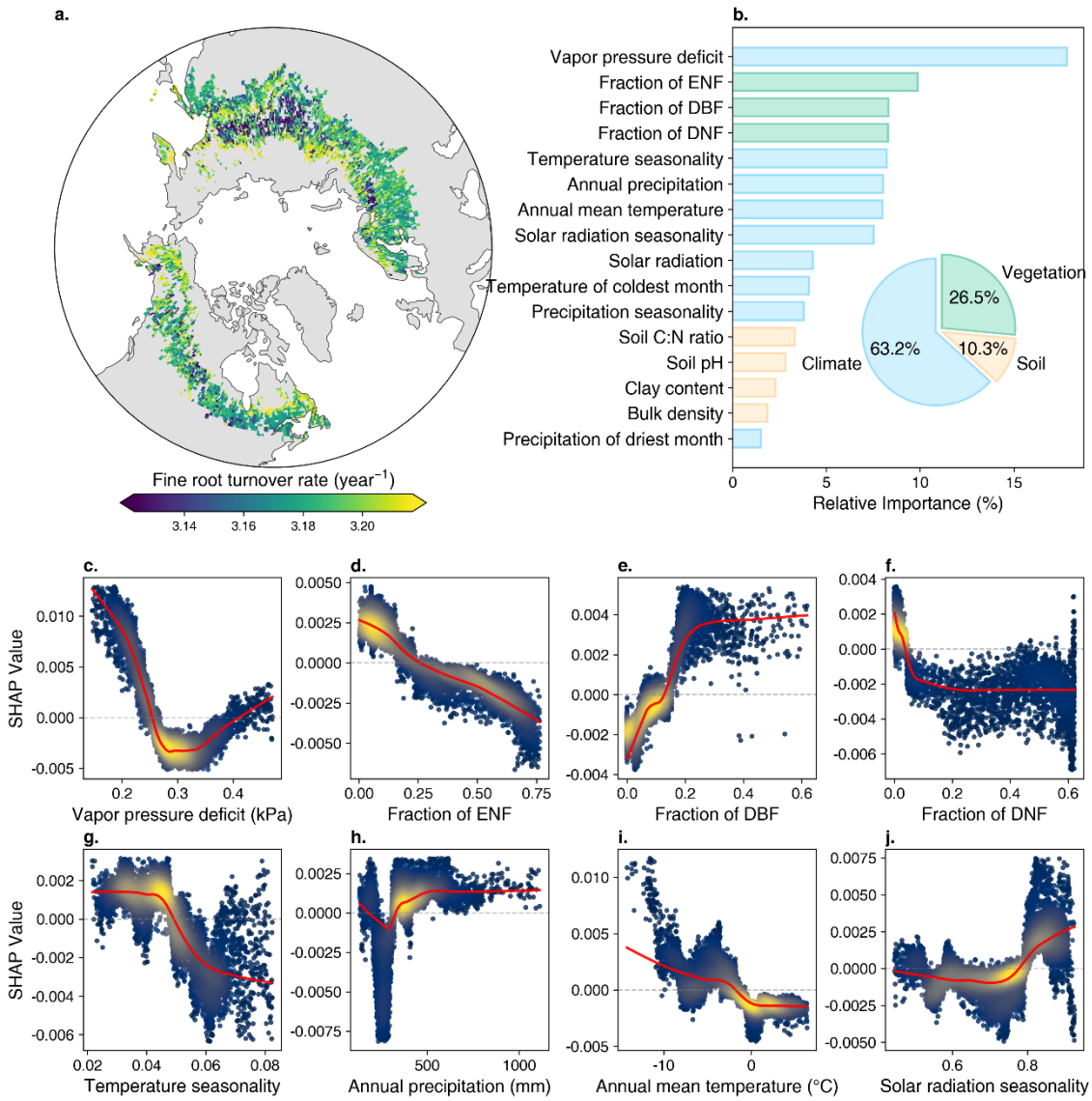

Fig. S10 Same as Fig. 3 but for the root turnover rate.

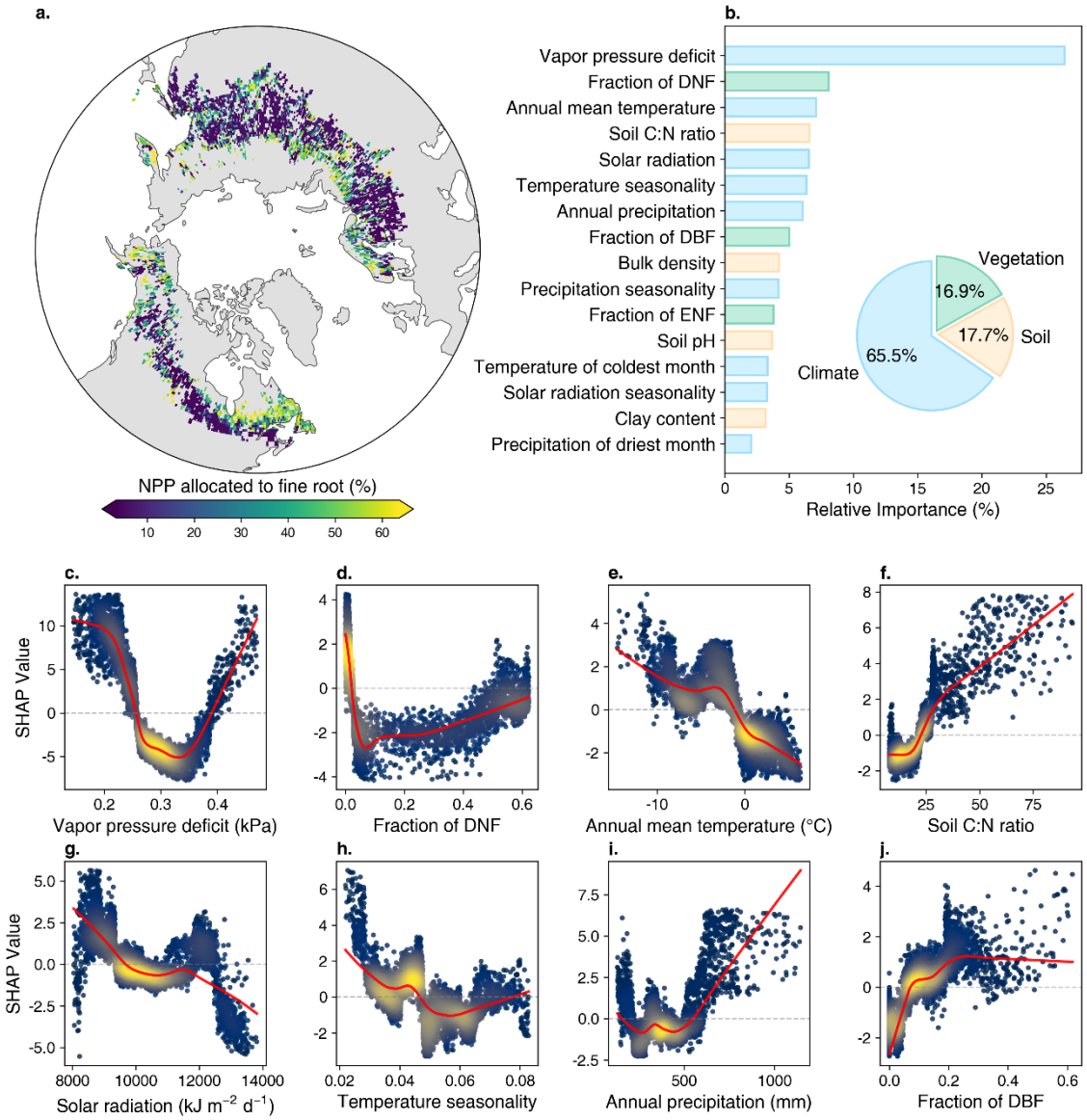

Fig. S11 Same as Fig. 3 but for the fraction of NPP allocated to fine root.

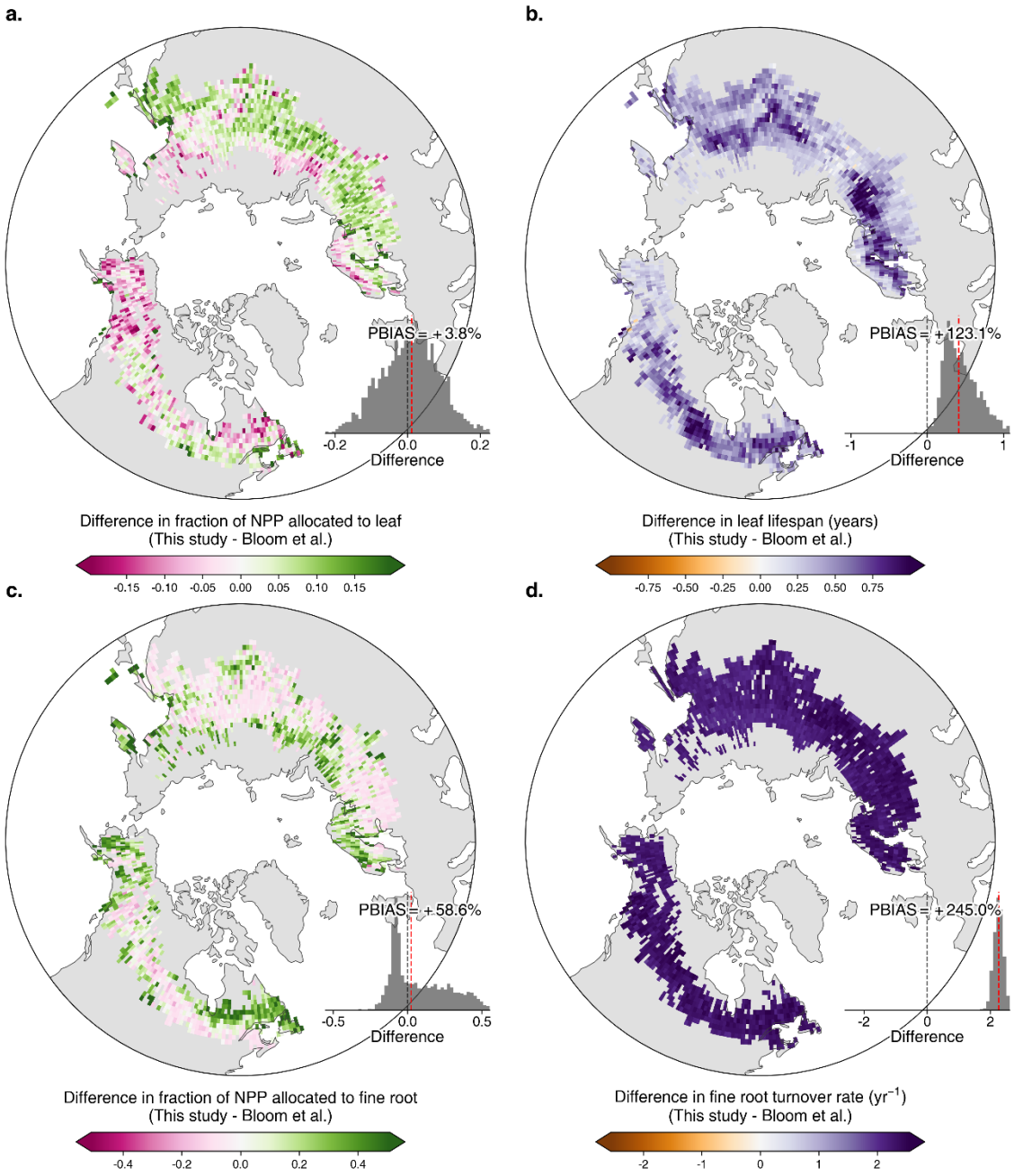

Fig. S12 Same as Fig. 6 but for (a) NPP allocation fraction to leaf, (b) leaf lifespan, (c) NPP allocation fraction to fine root, (d) fine root turnover rate.

Supplementary Tables

Table S1 Physical parameters in DALEC2. There are a total of 34 parameters, including 26 physical parameters and 8 initial pools. PAW and PUW is the abbreviation of plant available water and plant unavailable water pool. The model version we used does not use parameters related to PUW (in italic).

Parameter	Description	Unit	Min	Max	Note
f_{auto}	Fraction of GPP allocated to the autotrophic respiration	unitless	0.2	0.8	1
f_{lab}	Fraction of NPP allocated to the labile pool	unitless	0.01	0.5	1
f_{fol}	Fraction of (NPP – labile production) allocated to the foliar pool	unitless	0.01	0.5	1
f_{root}	Fraction of (NPP – labile production – foliar production) allocated to the root pool	unitless	0.01	1	1
τ_{lab}	Residence time of C in the labile pool	year	1.001	8	1
τ_{foliar}	Residence time of C in the leaf pool	year	1.001	8	1
k_{wood}	C turnover rate of the wood pool	day^{-1}	2.5×10^{-7}	0.001	1
k_{root}	C turnover rate of the fine root pool	day^{-1}	0.0001	0.01	1
k_{litter}	C turnover rate of the litter pool	day^{-1}	0.0001	0.01	1
k_{som}	C turnover rate of the soil organic matter	day^{-1}	10^{-7}	0.001	1
Q10	Temperature sensitivity of litter decomposition and heterotrophic respiration	unitless	0.018	0.14	1
ce	Nitrogen-limited canopy efficiency of the Aggregate Canopy Model	unitless	5	50	1
d_{onset}	Leaf onset day, on DOY = d_{onset} % 365.25 labile C release rate reaches the maximum	unitless	365.25	1461	1
c_{ronset}	Labile C release period	day	30.4375	100	1
d_{fall}	Leaf fall day, on DOY = d_{fall} % 365.25 leaf fall rate reaches the maximum	unitless	365.25	1461	1
c_{rfall}	Leaf fall period	day	30.4375	150	1
LCMA	Leaf carbon mass per area	$gC\ m^{-2}$	5	200	1
uWUE	Underlying water use efficiency	$gC\ kg\ H_2O^{-1}\ hPa^{-0.5}$	0.5	30	1
$Boese_r$	Solar shortwave downward radiation adjustment factor to evapotranspiration	$kg\ H_2O\ MJ^{-1}$	0.01	0.3	1
$Qmax_{PAW}$	Runoff focal point of the plant available water pool	$kg\ H_2O\ m^{-2}$	1	10^5	1
PAW_{fc}	Field capacity for the PAW pool	$kg\ H_2O\ m^{-2}$	1	10000	1

f_{wp}	Wilting point fraction as of field capacity	unitless	0.01	0.5	1
θ_p	Precipitation sensitivity of litter decomposition and heterotrophic respiration	unitless	0.01	1	1
k_{decomp}	Litter decomposition rate	day^{-1}	10^{-5}	0.01	1
X_{H2O}	*Fraction of water leaving the PAW pool being transferred to the PUW pool*	*unitless*	*0.01*	*1*	*1*
$Qmax_{PUW}$	*Runoff focal point of the PUW pool*	*kg H_2O m^{-2}*	*1*	*10^5*	*1*
$C_{lab, i}$	Initial C stock in the labile pool	gC m^{-2}	1	2000	1
$C_{fol, i}$	Initial C stock in the foliar pool	gC m^{-2}	1	2000	1
$C_{wood, i}$	Initial C stock in the wood pool	gC m^{-2}	1	10^5	1
$C_{root, i}$	Initial C stock in the fine root pool	gC m^{-2}	1	2000	1
$C_{litter, i}$	Initial C stock in the litter pool	gC m^{-2}	1	2000	1
$C_{som, i}$	Initial C stock in the soil organic matter pool	gC m^{-2}	1	2000	1
PAW_i	Initial water storage in the PAW pool	kg H_2O m^{-2}	1	10000	1
PUW_i	*Initial water storage in the PUW pool*	*kg H_2O m^{-2}*	*1*	*10000*	*1*

Table S2 Information of the drivers used in DALEC2.

Driver	Description	Unit	Data source
T_{max}	Daily maximum temperature	°C	CRU-JRA v2.2
T_{min}	Daily minimum temperature	°C	CRU-JRA v2.2
SSRD	Downward shortwave solar radiation	MJ m^{-2} day^{-1}	CRU-JRA v2.2
CO_2	Atmospheric CO_2 concentrations	ppm	TRENDY GCB2024
VPD	Vapor pressure deficit	hPa	Derived from CRU-JRA v2.2 *
DOY	Day of the year	\	\
PREC	Precipitation	mm day^{-1}	CRU-JRA v2.2
Lat	Latitude	\	\

* We calculated it as $\mathrm{VPD} = 0.6108\times e^{\frac{17.27\times T_{air}}{T_{air}+237.3}} - \frac{q_{air}\times P_{surf}}{0.378\times q_{air}+0.622}$, in which T_{air} is air temperature (°C), q_{air} is specific humidity (g g^{-1}), P_{surf} is surface pressure (kPa), and VPD is vapor pressure deficit (kPa) (Monteith & Unsworth, 2013; Tetens, 1930).

Table S3 Information of the constraining data.

Constraint for	Time frequency	Data source
Biomass	Yearly	Derived from forest recovery curve (Huang et al., 2021; Xu et al., 2026)
GPP	Daily	Global MODIS and FLUXNET-derived Daily Gross Primary Production, V2 (Joiner & Yoshida, 2021)
ET	Daily	GLEAM4 (Miralles et al., 2025)
LAI	8 days	GLASS LAI V6 (Ma & Liang, 2022)
RECO	10-day means	A Data-driven Upscale Product of Global Gross Primary Production, Net Ecosystem Exchange and Ecosystem Respiration (Zeng et al., 2020)

Table S4 Predictor variables for the machine learning models.

Category	Variable	Unit	Source
Climate	Annual mean temperature	°C	CRU-JRA v2.2
	Temperature of coldest month	°C	
	Temperature seasonality (Variation coefficient)	-	
	Annual precipitation	mm	
	Precipitation of driest month	mm	
	Precipitation seasonality (Variation coefficient)	-	
	Solar radiation	kJ m^{-2} day^{-1}	
	Solar radiation seasonality (Variation coefficient)	-	
	Vapor pressure deficit	kPa	
Soil	Bulk density	%	The Global Soil Dataset for Earth System Modeling
	Clay content	%	
	Soil pH	-	
	Soil C:N ratio	%	
Vegetation	Fraction of Evergreen Needleleaf Forest	%	ORCHIDEE
	Fraction of Deciduous Broadleaved Forest	%	
	Fraction of Deciduous Needleleaf Forest	%	

Reference

Huang, Y., Ciais, P., Santoro, M., Makowski, D., Chave, J., Schepaschenko, D., et al. (2021). A global map of root biomass across the world's forests. *Earth System Science Data*, *13*(9), 4263–4274. https://doi.org/10.5194/essd-13-4263-2021

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*.

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

Miralles, D. G., Bonte, O., Koppa, A., Baez-Villanueva, O. M., Tronquo, E., Zhong, F., et al. (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

Monteith, J. L., & Unsworth, M. H. (2013). Chapter 2 - Properties of Gases and Liquids. In J. L. Monteith & M. H. Unsworth (Eds.), *Principles of Environmental Physics (Fourth Edition)* (Fourth Edition, pp. 5–23). Boston: Academic Press. https://doi.org/10.1016/B978-0-12-386910-4.00002-0

Nash, J. E., & Sutcliffe, J. V. (1970). River flow forecasting through conceptual models part I — A discussion of principles. *Journal of Hydrology*, *10*(3), 282–290. https://doi.org/10.1016/0022-1694(70)90255-6

Ritter, F., Ciais, P., Senf, C., Santoro, M., Xu, Y., Pelissier-Tanon, A., et al. (2025, April 16). Alarming decline in the carbon sink of European forests driven by disturbances. Research Square. https://doi.org/10.21203/rs.3.rs-3671432/v1

Robinson, N., Drever, C. R., Gibbs, D. A., Lister, K., Esquivel-Muelbert, A., Heinrich, V., et al. (2025). Protect young secondary forests for optimum carbon removal. *Nature Climate Change*, *15*(7), 793–800. https://doi.org/10.1038/s41558-025-02355-5

Tetens, O. (1930). Uber einige meteorologische Begriffe. *Z. Geophys*, *6*, 297–309.

Wang, H., Colin Prentice, I., J. Wright, I., I. Warton, D., Qiao, S., Xu, X., et al. (2023). Leaf economics fundamentals explained by optimality principles. *Science Advances*. https://doi.org/10.1126/sciadv.add5667

Xia, J., Yuan, W., Lienert, S., Joos, F., Ciais, P., Viovy, N., et al. (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.

Zeng, J., Matsunaga, T., Tan, Z.-H., Saigusa, N., Shirai, T., Tang, Y., et al. (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