Stepwise Calibration of Age-Dependent Biomass in the Integrated Biosphere Simulator (IBIS) Model

Rui Ma, Philippe Ciais, Yidi Xu, Daniel Goll, Yuan Zhang, Jingfeng Xiao, Shunlin Liang

JAMES | Journal of Advances in Modeling Earth Systems

RESEARCH ARTICLE
10.1029/2023MS004048

Stepwise Calibration of Age-Dependent Biomass in the Integrated Biosphere Simulator (IBIS) Model

Rui Ma[1,2], Yuan Zhang[3], Philippe Ciais[2], Jingfeng Xiao[4], Yidi Xu[2], Daniel Goll[2], and Shunlin Liang[5]

[1]School of Remote Sensing and Information Engineering, Wuhan University, Wuhan, China, [2]Le Laboratoire des Sciences du Climat et de l'Environnement, IPSL-LSCE CEA/CNRS/UVSQ, Gif sur Yvette, France, [3]Key Laboratory of Alpine Ecology, Institute of Tibetan Plateau Research, Chinese Academy of Sciences, Beijing, China, [4]Earth Systems Research Center, Institute for the Study of Earth, Oceans, and Space, University of New Hampshire, Durham, NH, USA, [5]Jockey Club STEM Laboratory of Quantitative Remote Sensing, Department of Geography, University of Hong Kong, Hong Kong, China

Key Points:

- We presented a stepwise calibration framework for better integrating age-dependent biomass into the integrated biosphere simulator model
- We utilized a machine learning-based model as an alternative to the physical model, accelerating the calibration process
- The calibration noticeably improved the simulation of gross primary productivity, leaf area index, and biomass

Supporting Information:

Supporting Information may be found in the online version of this article.

Correspondence to:

S. Liang and Y. Zhang,
shunlin@hku.hk;
yuan.zhang@itpcas.ac.cn

Citation:

Ma, R., Zhang, Y., Ciais, P., Xiao, J., Xu, Y., Goll, D., & Liang, S. (2024). Stepwise calibration of age-dependent biomass in the integrated biosphere simulator (IBIS) model. *Journal of Advances in Modeling Earth Systems*, *16*, e2023MS004048. https://doi.org/10.1029/2023MS004048

Received 14 OCT 2023
Accepted 6 APR 2024

Abstract Many land surface models (LSMs) assume a steady-state assumption (SS) for forest growth, leading to an overestimation of biomass in young forests. Parameters inversion under SS will potentially result in biased carbon fluxes and stocks in a transient simulation. Incorporating age-dependent biomass into LSMs can simulate real disequilibrium states, enabling the model to simulate forest growth from planting to its current age, and improving the biased post-calibration parameters. In this study, we developed a stepwise optimization framework that first calibrates "fast" light-controlled CO_2 fluxes (gross primary productivity, GPP), then leaf area index (LAI), and finally "slow" growth-controlled biomass using the Global LAnd Surface Satellite (GLASS) GPP and LAI products, and age-dependent biomass curves for the 25 forests. To reduce the computation time, we used a machine learning-based model to surrogate the complex integrated biosphere simulator LSM during calibration. Our calibrated model led to an error reduction in GPP, LAI, and biomass by 28.5%, 35.3% and 74.6%, respectively. When compared with net biome productivity (NBP) using no-age-calibrated parameters, our age-calibrated parameters increased NBP by an average of 50 gC m^{-2} yr^{-1} across all forests, especially in the boreal needleleaf evergreen forests, the NBP increased by 118 gC m^{-2} yr^{-1} on average, increasing the estimate of the carbon sink in young forests. Our work highlights the importance of including forest age in LSMs, and provides a novel framework for better calibrating LSMs using constraints from multiple satellite products at a global scale.

Plain Language Summary Physical and biological process-based models always overestimate the biomass of young forests, with an assumption that they usually hold maximum carbon stocks like old-growth stands. Such an assumption can lead to biased carbon fluxes and stocks in further simulation. Considering stand age in LSMs improves their ability to simulate real forest growth. In this study, we develop a stepwise method to account for stand age effects in model simulations by assimilating remotely sensed information on vegetation productivity, leaf area, biomass, and age. To reduce the computational cost of the complex original code, we use a substitute model constructed using a machine learning method for calculations. The improved model successfully reproduces the changes in ecosystem biomass and fluxes as forest age varies. Our research provides a novel approach to improving other land surface models for predicting age-dependent ecosystem properties.

1. Introduction

Land surface models (LSMs) are an important tool for exploring the dynamics of energy budgets (Alton, 2013; Li et al., 2016), water and carbon cycles (Lafont et al., 2012), and plant growth (Liu et al., 2016) due to their capability of representing the complex interplay between biophysical and biogeochemical processes. However, carbon fluxes predicted by LSMs exhibit substantial variability in spatiotemporal patterns (Krause et al., 2022; Seiler et al., 2022). This variability can be due to uncertainties in input variables (Barman et al., 2014), over-simplified model structures and assumptions like a simple big leaf canopy model for vegetation (Chaney et al., 2016), or poorly specified parameters resulting from a lack of information (Bastrikov et al., 2018; Xiao et al., 2014). To increase the realism of LSMs, the modeling community has been adding more and more detailed processes into LSMs, which may have decreased model structural uncertainties, but inevitably increased model complexity and the number of unknown parameters (Famiglietti et al., 2021).

To reduce the number of unconstrained parameters, observations on ecosystem properties from various temporal and spatial scales are being used for model calibration (MacBean et al., 2016, 2022; Xiao et al., 2019). Current calibration of carbon-related parameters mainly focuses on using data that are intensively observed by field measurements or space sensors, such as gross primary productivity (GPP), leaf area index (LAI), soil moisture, and the fraction of absorbed photosynthetically active radiation (FAPAR), as independent or joint constraints (Bacour et al., 2015; Forkel et al., 2019; He et al., 2021; Kumar et al., 2019; Ma et al., 2022a). These land surface variables provide information on fast carbon and water fluxes exchanging, but poorly indicate the ecosystem carbon stock evolution at longer timescales, leading to uncertainties in biomass, respiration, and net biome productivity (NBP) predictions (Santaren et al., 2014; Thum et al., 2017).

Adding carbon stock observations into parameter optimization schemes makes it possible to constrain slower processes like carbon allocation and tissue turnover, but will be challenged by the lack of representation of individual land age cohorts in models (Carvalhais et al., 2010). Initialization of LSMs usually runs repeating climate conditions until equilibrium, or steady state (SS), when net carbon exchanges approach zero at the annual scale (Ge et al., 2018; Pietsch & Hasenauer, 2006). This will result in an overestimation of the simulated biomass for young forests, as the SS assumption would imply that they always hold maximum carbon stocks (Ciais et al., 2008). However, in reality, young forests are considered to be strong carbon sinks irrespective of environmental changes with GPP exceeding carbon losses from respiration (Heinrich et al., 2021, 2023). It will impact the assessment of regional or global carbon budget, since young forests make up a significant proportion of the global forested areas (Besnard et al., 2021; Pugh, Lindeskog, et al., 2019). Moreover, the SS assumption could also introduce bias in parameter inversion within a model-data fusion framework. For example, several ecological parameters, like the decay rate of recalcitrant pools, turnover time and allocation have been indicated to exhibit poor performances under the SS assumption (Ge et al., 2018; Wutzler & Reichstein, 2007). Such biases in the initialization of pools and retrieved parameters could lead to incorrect assessments of forest carbon storage capacity.

New frameworks for model calibration are required which consider the non-equilibrated state of ecosystems. For example, Carvalhais et al. (2008) introduced a factor to initialize the model running from an imbalance state (carbon sink or source) and demonstrated that the differences between modeled carbon pools and observations could be decreased with such relaxed steady-state assumption. Considering that the growth process of a forest can be reflected by the relationship between biomass and age (Amiro et al., 2010; Bukoski et al., 2022; Liu et al., 2014; Repo et al., 2021; Thum et al., 2017) proposed a method that removed all biomass pools in the model when the stand age was zero, followed by a simulation up to the current age. This approach enables the modeled biomass to match the real age of the forest, and facilitates a more accurate representation of forest growth dynamics. Such age-dependent biomass data assimilation has been performed at a site or plot scale but rarely at the regional or global scale (Smallman et al., 2017; Thum et al., 2017). The lack of long-term carbon stock data and spatial age maps and the high computational cost of LSMs have been major bottlenecks for global applications (MacBean et al., 2016; Thum et al., 2017).

The emergence of novel large-scale biomass and age maps now provides the opportunity to explore the growth trajectory of forests and the data basis for calibration of age-dependent biomass dynamics in global models (Besnard et al., 2021; Poulter et al., 2019; Xu et al., 2021). In addition, computationally efficient machine learning (ML) emulators of computationally intensive LSMs are becoming available for LSMs calibration (Reichstein et al., 2019). ML emulators are typically constructed with parameter samples and the corresponding model outputs or errors, and then substitute the LSMs for the calibration. Many studies have shown that with a well-trained ML emulator, the calibration can be dramatically accelerated while maintaining sufficient model precision (Fer et al., 2018; Hawkins et al., 2019; Ma et al., 2022a; Xu et al., 2018).

In this study, we developed a stepwise calibration framework to assimilate age-dependent biomass, GPP, and LAI into a LSM (the integrated biosphere simulator, IBIS). We sequentially calibrated the sensitive parameters in three steps by taking GPP, LAI, and biomass observations as constraints separately. Real stand age information was added to the IBIS by artificially removing all the biomass pools in IBIS for a simulated forest. To reduce the computing time, we applied a modified adaptive surrogate modeling optimization method (MASM) with a ML emulator in our calibration and tested the framework on several representative forests across the world. The objectives of this study are to present a new framework for improving both the fast and

slow processes of IBIS with multiple remotely sensed products and to explore the possibility of global application.

2. Model and Materials

2.1. Model Description

The IBIS is a process-based LSM representing energy balance, hydrological, soil, and vegetation dynamics in different terrestrial ecosystems like forests, grass, and shrubs (Foley et al., 1996). The carbon cycle in IBIS depends on processes including photosynthesis, respiration, allocation, and phenology. The canopy level photosynthesis, that is, GPP, of vegetation follows Farquhar's scheme (Farquhar et al., 1980), and is calculated as a function of absorbed light, temperature, CO_2, and Rubisco enzyme capacity (Kucharik et al., 2000; Winter et al., 2009). For forests, autotrophic respiration (Ra) is deduced from GPP to calculate net primary productivity, which is then allocated into three biomass pools including leaf (Cbiol), wood (Cbiow), and fine root (Cbior) (Equation 1) (Sierra et al., 2022). These pools are updated annually with constant allocation ratios (*aleaf, awood, aroot*) and turnover times (*tauleaf, tauwood, tauroot*) as follows:

$$\frac{dBio_i}{dt} = a_i(GPP - Ra) - \frac{Bio_i}{tau_i} \qquad \sum a_i = 1 \tag{1}$$

where *Bio* represents live biomass including leaf, wood, and root biomass (Cbiol, Cbiow, Cbior), *a* is the allocation ratio of different carbon pool *i* (*i* = leaf, wood, or fine root), and *tau* defines the turnover time (years) of each pool.

The annual updated Cbiol is used to calculate daily LAI using specific leaf area (*specla*) and phenology function in each plant functional type (PFT) as follows:

$$\mathrm{LAI} = plai * phen = Cbiol * specla * phen \tag{2}$$

where "*plai*" is the potential LAI (yearly maximum LAI), "*specla*" represents the specific leaf area depending on forest types, and "*phen*" is the function of temperature, representing the seasonal variation of leaves in each PFT. The value of *phen* is one during the peak growing season, while equals zero when all leaves are shed. The climate triggers of leaf flushing and shedding are different for deciduous and evergreen forests (Botta & Foley, 2002; Kucharik et al., 2000).

The detailed description and parameterization strategy of the IBIS model can be found in Foley et al. (1996) and Kucharik et al. (2000).

2.2. Data Sources

2.2.1. Model Forcing

The meteorological forcings including air temperature (°C), pressure (Pa), wind speed (m s^{-1}), specific humidity (%), precipitation (mm), and downward solar radiation flux (W m^{-2}) are from a gridded historical climate data set CRUJRA v2.2 (Harris, 2021), which were reconstructed using the Japanese Reanalysis data (JRA) and the Climate Research Unit data set from 1901 to 2020 with a resolution of 0.5° × 0.5° (Harris et al., 2014, 2020; Kobayashi et al., 2015) (Table 1). The original 6 hourly data was aggregated into daily values to drive the IBIS model. Soil sand and clay contents (%) are derived from a gridded Global Soil Data set for use in Earth System Models (GSDE) with a resolution of 0.083° × 0.083° (Shangguan et al., 2014). The CO_2 concentration used in our simulations was reconstructed based on the ice-core records before 1958, while from the Mauna Loa Observatory measurements after 1958 (Thoning et al., 1989).

2.2.2. Satellite Products for Model Calibration: GPP, LAI, Biomass, and Age

We used eight-day gridded GPP (2000–2019) and LAI (2000–2020) maps from the Global Land Surface Satellite (GLASS) products suite (Liang et al., 2021) with a spatial resolution of 0.05°. The GPP product was derived from a revised light use efficiency model that was calibrated using eddy covariance measurements and driven by meteorological data (Zheng et al., 2020). The LAI data was generated by combining several existing global LAI

Table 1
Description of the Data sets Used in This Study

Variables	Products	Time span	Spatial resolution	Temporal resolution	References
Model forcing					
Climate variables*	CRUJRA	1901–2019	0.5°	6-hourly	Harris (2021)
CO_2	CO_2	1700–2019	–	Yearly	Thoning et al. (1989)
Soil sand, soil clay	Soil textures	–	~10 km	–	Shangguan et al. (2014)
Land cover	MCD12Q1	2010–2020	500 m	Yearly	Homer et al. (2020)
Tree cover and Lossyear	Hansen_2022		30 m		Hansen et al. (2013)
Remotely sensed products					
LAI	GLASS LAI	2000–2020	0.05°	8-day	Ma and Liang (2022)
GPP	GLASS GPP	2000–2019	0.05°	8-day	Zheng et al. (2020)
Biomass	Biomass	2000–2019	0.1°	yearly	Xu et al. (2021)
Age	Age	2010	1 km	yearly	Besnard et al. (2021)

Note. Climate variables include daily air temperature (°C), the minimum temperature (°C), the maximum temperature (°C), precipitation (mm), pressure (Pa), specific humidity (%), wind speed (m s^{-1}), and downward solar radiation flux (W m^{-2}). Lossyear: year of tree cover loss.

products and the Terra and Aqua combined Moderate Resolution Imaging Spectroradiometer (MODIS) reflectance information using a deep learning algorithm (Ma & Liang, 2022).

The biomass data used in this paper comes from a global annual biomass carbon density data set from 2000 to 2019 with a spatial resolution of 0.1° (Xu et al., 2021). It was developed spatially using a self-improving ML model trained with extensive biomass and forest height measurements (more than 100,000 plots) from ground, air, and space. For the age of forests, we used a global ML product trained with more than 40,000 plots of global forest inventories, biomass, and climate records (Besnard et al., 2021). It provides the age information of each 1 km pixel circa 2010, with forest age values capped at 300 years. We resampled the age data into 0.1° to match the biomass map by averaging all the 1 km forested pixels within every 0.1-degree cell.

To ensure a relatively pure and unchanged forest type for our calibration, initially, we screened the data based on two criteria: (a) vegetation type remained unchanged from 2010 to 2020, according to the MODIS annual land cover product (MCD12Q1) (Homer et al., 2020), which was resampled to 0.1° × 0.1° to align with the satellite products; (b) grid cells did not experience forest loss from 2010 to 2020 according to the forest cover and forest loss data extended from Hansen et al. (2013) (data available at https://glad.earthengine.app/view/global-forest-change). Following this, we randomly select 25 forests (one 0.1° × 0.1° pixel for each) from the remaining data, with 3–4 in each PFT (Table 2).

3. Parameter Calibration Scheme

The objective of the calibration is to improve the simulation of biomass, which is regulated by "fast" (photosynthesis, phenology, respiration, etc.) and "slow" processes (long term carbon accumulation, etc.). We employed a step-by-step calibration approach to sequentially calibrate various processes operating on distinct timescales. This method allows for a more targeted and iterative calibration, ensuring that each data stream exclusively constrains the most relevant model processes based on their relative contributions (Peylin et al., 2016). First, we use the observed 8-day LAI as forcing and 8-day GPP as a reference to calibrate the light-dependent carbon parameters. Second, the observed LAI (8-day) was used as a reference to calibrate the leaf carbon allocation and turnover (thus indirectly the allocation to wood) as well as phenology parameters. Third, we used the annual biomass and forest age data to calibrate the slow processes affecting wood and fine root biomass.

3.1. Parameter Selection

We used a two-step parameter selection approach. First, based on model processes and parameter selection of previous model calibration studies, we select 20 key parameters that obviously affect photosynthesis, carbon allocation, and phenology simulation as the candidate parameters (Text S1 in Supporting Information S1). Then

Table 2
Information on all the Selected Forests (See Spatial Locations in the Supplement Information, Figure S1 in Supporting Information S1)

Num	ID	Latitude	Longitude	PFT_{IBIS}	Age_{2010}	$Year_{ini}$
1	BoND1	52.87	120.95	Boreal needleleaf cold-deciduous tree (BoND)	74	1936
2	BoND2	60.47	116.05	Boreal needleleaf cold-deciduous tree (BoND)	115	1895
3	BoND3	61.68	70.76	Boreal needleleaf cold-deciduous tree (BoND)	93	1917
4	BoBD1	56.17	−115.85	Boreal broadleaf cold-deciduous tree (BoBD)	94	1916
5	BoBD2	58.07	59.55	Boreal broadleaf cold-deciduous tree (BoBD)	100	1910
6	BoBD3	39.07	−79.60	Boreal broadleaf cold-deciduous tree (BoBD)	78	1932
7	TeNE1	42.87	−122.75	Temperate needleleaf evergreen tree (TeNE)	169	1841
8	TeNE2	−33.97	150.08	Temperate needleleaf evergreen tree (TeNE)	89	1921
9	TeNE3	28.14	96.44	Temperate needleleaf evergreen tree (TeNE)	100	1910
10	TeBD1	35.37	−83.85	Temperate broadleaf cold-deciduous tree (TeBD)	74	1936
11	TeBD2	48.47	23.45	Temperate broadleaf cold-deciduous tree (TeBD)	70	1940
12	TeBD3	−43.19	−72.39	Temperate broadleaf cold-deciduous tree (TeBD)	113	1897
13	BoNE1	63.07	44.35	Boreal needleleaf evergreen tree (BoNE)	87	1923
14	BoNE2	44.28	−65.60	Boreal needleleaf evergreen tree (BoNE)	54	1956
15	BoNE3	59.94	12.81	Boreal needleleaf evergreen tree (BoNE)	57	1953
16	TroBD1	21.57	96.75	Tropical broadleaf drought-deciduous tree (TroBD)	47	1963
17	TroBD2	−23.13	−64.15	Tropical broadleaf drought-deciduous tree (TroBD)	36	1974
18	TroBD3	−8.78	29.42	Tropical broadleaf drought-deciduous tree (TroBD)	36	1974
19	TroBE1	−1.63	23.75	Tropical broadleaf evergreen tree (TroBE)	300	-
20	TroBE2	0.47	114.85	Tropical broadleaf evergreen tree (TroBE)	300	-
21	TroBE3	−17.31	−64.86	Tropical broadleaf evergreen tree (TroBE)	60	1950
22	TroBE4	−24.49	−48.00	Tropical broadleaf evergreen tree (TroBE)	34	1976
23	TeBE1	−30.03	152.55	Warm-temperate broadleaf evergreen tree (TeBE)	71	1939
24	TeBE2	20.16	103.97	Warm-temperate broadleaf evergreen tree (TeBE)	33	1977
25	TeBE3	52.04	−8.04	Warm-temperate broadleaf evergreen tree (TeBE)	112	1898

Note. PFT_{IBIS} stands for plant functional type in IBIS. Age_{2010} refers to age values from the age product, $Year_{ini}$ represents the year when the age is equal to zero according to the age map.

we performed sensitivity analyses among the candidate parameters to identify the ones affecting GPP, LAI, and biomass most for the final calibration. The bounds of each parameter used in sensitivity analysis and calibration were derived from the literature (Cunha et al., 2013; Kucharik et al., 2000; Varejão et al., 2013) or set to ±40% of the default values (Table 3). Specifically, for phenology-related parameters, temperature-related thresholds (*tempu_tth, tempu_gth, tempu_gdd*) were applied in the boreal and temperate deciduous forests to optimize the flushing and shedding of the leaves. While for tropical deciduous forests, the leaves begin to drop with the trigger by the 10-day carbon photosynthesis rates via the regulation of the parameter *tempu_drop*.

When selecting the parameters to be optimized, we first identified the parameters that are most relevant to the simulation processes of LAI, GPP, and biomass based on the modules they affect. Then we performed a sensitivity analysis based on the Sobol method (Sobol, 1993) within the ranges of parameters, with GPP, LAI, biomass, Cbiol, Cbiow, and Cbior as the target variables. All the parameters were set to a uniform distribution and sampled more than 8,000 combinations from their prior spaces randomly. Finally, we selected the 11 most influential parameters (Figure 1), including three GPP-sensitive parameters (*alpha3, theta3, vmax* for step 1), four LAI-sensitive parameters (*aleaf, tauleaf, specla, tempu_gth* for step 2) and four biomass-sensitive parameters (*awood, tauwood0, tauroot, rgrowth* for step 3).

Journal of Advances in Modeling Earth Systems
10.1029/2023MS004048

Table 3
Description, Prior Values and Ranges of the 20 Key Parameters Used in Sensitivity Analysis

Parameters	Description (units)	TroBE	TroBD	TeBE	TeNE	TeBD	BoNE	BoBD	BoND	Range
alpha3	intrinsic quantum efficiency for C3 plants				0.060					[0.03, 0.125]
theta3	coupling coefficient of C3 plants				0.970					[0.6, 0.996]
rgrowth	growth respiration coefficient				0.300					[0.1, 0.6]
beta1	parameter for the lower fine root				0.950					[0.75, 0.99]
beta2	parameter for the upper fine root				0.975					[0.75, 0.99]
stressfac	moisture stress factor				−5					[−6.5, −3.5]
gamma	leaf respiration coefficient	0.015	0.015	0.015	0.015	0.015	0.015	0.015	0.015	[0.0075, 0.0225]
coefm	coefficient for stomatal conductance	10	10	10	6	10	6	10	10	[5, 15]
coefb	coefficient for stomatal conductance	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	[0.005, 0.015]
v_{max}	the maximum carboxylase capacity of Rubisco (molCO2 m^{-2} s^{-1}) ($\times 10^{-5}$)	6.50	6.50	4.00	3.00	3.00	2.50	3.00	3.00	[2.00, 10.00]
specla	specific leaf area (m^2 kg^{-1})	25	25	25	12.5	25	12.5	25	25	[4, 50]
tauleaf	foliar biomass turnover time (yr)	1	1	1	2	1	2.5	1	1	[0.05, 3]
tauwood0	stem biomass turnover time (yr)	25	25	25	50	50	100	100	100	[5, 200]
tauroot	fine root biomass turnover time (yr)	1	1	1	1	1	1	1	1	[0.05, 5]
aleaf	carbon allocation fraction to leaves	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	[0.05,1]
awood	carbon allocation fraction to woods	0.5	0.5	0.5	0.3	0.5	0.3	0.5	0.5	[0.05, 1]
tempu_tth	temperature threshold for leaf activity (°C)	–	–	–	–	0	–	0	0	[−10, 10]
tempu_gdd	a growing degree day (gdd) threshold	–	–	–	–	100	–	100	100	[60, 140]
tempu_gth	temperature threshold of gdd for leaf budburst (°C)	–	–	–	–	5	–	5	5	[0, 10]
tempu_drop	days that affect leaf shedding of TroBD	–	15	–	–	–	–	–	–	[9, 21]

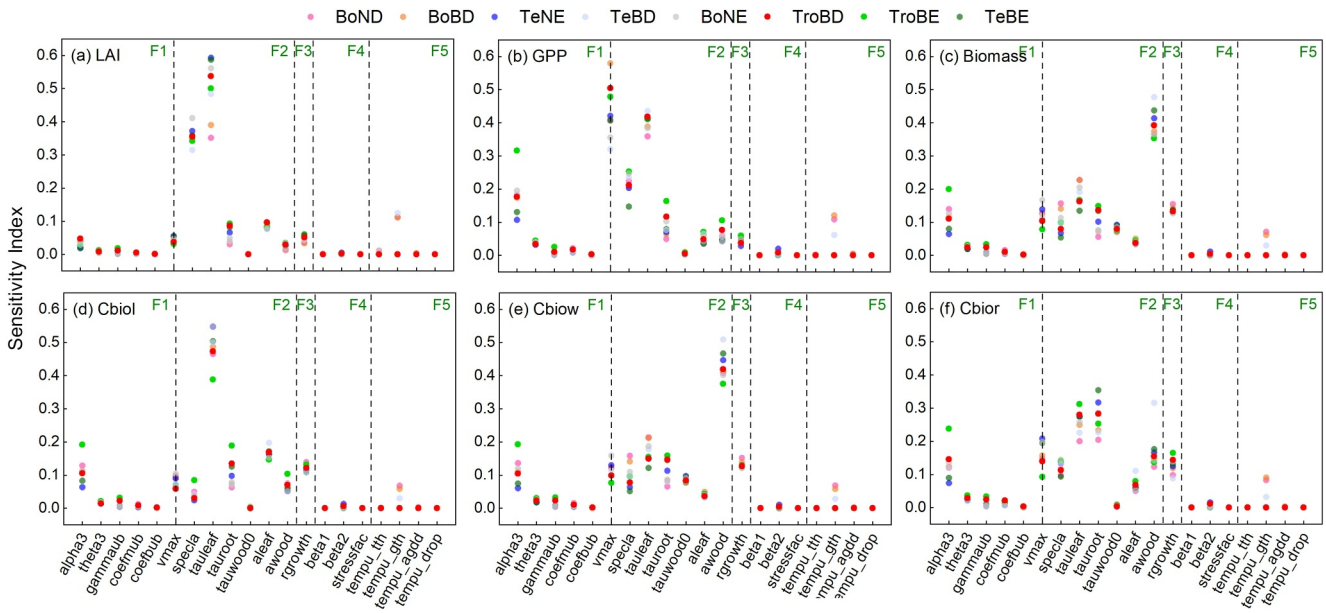

Figure 1. Sensitivity indexes of the 20 key parameters in integrated biosphere simulator at 25 forests, represented by different symbols. The dashed lines are employed to group parameters that serve distinct functions (F1: Photosynthesis; F2: Allocation and turnover; F3: Respiration; F4: Stress; F5: Phenology).

It should be noted that we did not include *specla* and *tauleaf* when calibrating GPP, although they are more influential than *alpha3, theta3* for GPP. This is because we used the observed LAI as forcing when calibrating GPP. The *specla* and *tauleaf* are directly associated with leaf biomass and LAI, but cannot affect GPP. The large sensitivities of GPP to the two parameters (Figure 1) imply a strong impact of LAI on GPP in the model, highlighting the necessity of using more accurate LAI for GPP calibration. Regarding biomass, the influential parameters for GPP and LAI also show a strong impact. However, they have been calibrated in the first two steps (See Section 3.3.3). Therefore, we only calibrated *awood, tauwood0, tauroot, rgrowth* in the last step.

3.2. Modified Adaptive Surrogate Modeling Optimization

The Markov Chain Monte Carlo method (MCMC) is a commonly used algorithm for model parameter calibration. It iteratively samples parameter values within predefined priors, evaluating each set through model runs. Subsequently, the cost function is computed to determine the acceptance or rejection of this set of parameters based on the Metropolis–Hastings criterion. All accepted parameters are then used to generate posterior distributions, with a mode approximating a normal distribution indicating well-constrained posterior parameters. To assimilate biomass data, simulation from planting (or the most recent disturbance) up to its current age is needed, which may markedly increase computing time, especially when using Metropolis algorithms. Therefore, running the LSMs using all the random parameter sets is computationally intensive. Here we applied a Modified Adaptive Surrogating Modeling optimization method (MASM), which trains an emulator of the IBIS model using a ML technique and uses the computationally cheap emulator to examine all the parameter sets (Gong & Duan, 2017; Ma et al., 2022a).

In parameter distribution estimation, the target distribution is usually the posterior probability distribution generated through the framework of Bayes' rule, and is proportional to the likelihood and prior probability density under the observation constraints:

$$P\left(\theta_i \middle| y_t^O\right) \propto P\left(y_t^O \middle| \theta_i\right) P(\theta) = L\left(\theta_i \middle| y_t^O\right) P(\theta_i) \qquad i = 1,2,3,\cdots,n \tag{3}$$

where y_t^O is the referenced (observed) data, $P\left(\theta_i \middle| y_t^O\right)$ denotes the posterior probability density functions of target parameters given by y, $P\left(y_t^O \middle| \theta_i\right)$ stands for likelihood. $P(\theta_i)$ represents prior distribution (set as uniform in this

study), and n is the number of sensitive parameters. Based on this, the MASM method includes the following steps (Figure S2 in Supporting Information S1):

1. Randomly sample parameter sets ns times ($pars^{ns}$) within the prior range of each parameter (ns = 18*number of parameters), and run the IBIS model to generate the time series of target variables.

$$y_M^{ns} = IBIS\left(pars^{ns}, mf_t\right) + \varepsilon_t \quad t = 1,2,\ldots,n \tag{4}$$

where $pars^{ns}$, mf_t and ε_t are parameters, model forcings, and model error at time t. The definition of the likelihood function (L) and cost function CF^{ns} is as follows:

$$CF^{ns} = -2\log\{L\} = -2\,log\left\{(2\pi\sigma)^{-n/2}\prod_t exp\left[-\frac{\left(y_o - y_M^{ns}\right)^2}{2\sigma^2}\right]\right\}, \tag{5}$$

where σ^2 should be the error variance of each observation. We assumed that the errors between the model and the observation ($y_o - y_M^{ns}$) are independent (i.e., the covariance is zero), and therefore σ^2 here is expressed by the variance of each observation product (Yuan et al., 2012).

2. Fit a surrogate model using the Gaussian Processed Regression (GPR) model between the sampled parameter sets ($pars^{ns}$) and the corresponding model errors CF^{ns}, and apply ASMO-PODE (Adaptive Surrogate Modeling-based Optimization-Parameter Optimization and Distribution Estimation, Gong and Duan (2017)) to the GPR model to decrease the number of model evaluations. Update the GPR model through an iterative loop ($iloop$ = 120) to improve its accuracy. In each iteration, three representative points from the posterior distribution, along with the corresponding CF values obtained from the IBIS model, are added to the sample for re-training the GPR model. The IBIS model may find a parameter set with lower model error within a limited number of runs, but achieving convergence is challenging (Figure S3 in Supporting Information S1, gray lines). The predicted CF values from GPR model gradually approached the minimum CF value of the IBIS model during 120 loops (Figure S3 in Supporting Information S1, red lines).
3. After finishing the iterative loop, we carried out Metropolis optimization on the trained GPR model. Once the loop ends and the Markov Chain converges (GR < 1.2, Equation 6) (Gelman & Rubin, 1992), the final established GPR model can be obtained. By applying the Metropolis algorithm to the final GPR model (200,000 samplings × 4 chains), the best CF value can be quickly approached as the number of runs increases (Figure S3 in Supporting Information S1, blue lines). Additionally, the posterior probability distribution of each parameter can be determined using the last 100,000 values.

$$GR = \sqrt{\frac{N_2 - 1}{N_2} + \frac{N_{chain} + 1}{N_{chain} \times N_2} \times \frac{B}{W}}, \tag{6}$$

where N_2 = 200,000, N_{chain} = 4, B represents the variation among the averaged values of each chain, and W is the average of the variances within chains.

A detailed explanation of the MASM and its corresponding code is available in Ma et al. (2022a, 2022b) and Gong and Duan (2017). Unlike traditional parameter calibration, we only need to run the IBIS model a few hundred times (ns+3*$iloop$) with the MASM method instead of millions of times, which greatly reduces the time cost of the model running without excessive loss of accuracy.

3.3. Optimization Strategy

3.3.1. Biomass-Age Curves Fitting

To define long-term age-dependent biomass relationships for calibration of the IBIS parameters, we first fitted relationships between biomass and stand age using the widely employed Chapman-Richard growth function (Richards, 1959), where the biomass increases with age until it reaches the asymptote:

$$Bio_t = A(1 - \exp(-k \cdot t))^c, k \text{ and } c > 0 \tag{7}$$

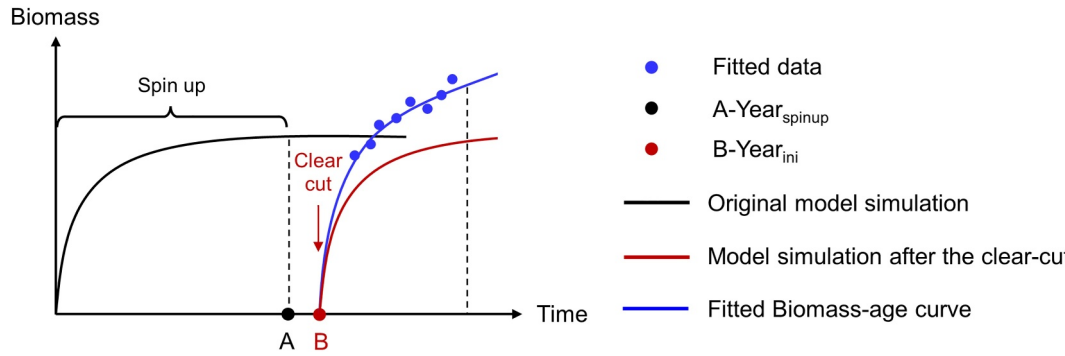

Figure 2. The flow of biomass assimilation in the integrated biosphere simulator model. The blue points correspond to the data used for fitting the curve. Point A indicates the time after the model spin-up, whereas point B refers to the time when the forest age is zero (B = 2020–age). The black solid curve represents the original model trajectory, the red curve depicts the model trajectory after a clear cut, and the blue curve corresponds to the biomass-age curve utilized as a reference during our calibration process.

where Bio_t is the (observed) biomass stock in MgC ha^{-1} at (observed) age t; A stands for the asymptote limit of the equation where biomass saturates; k refers to the growth rate coefficient of Bio; c determines the shape of the fitted curve.

For younger forests with an age of less than 300 years old in Besnard et al. (2021), we collected all the gridded biomass and age data of the corresponding vegetation type within a 5° × 5° window and grouped them into 5-year age bins. We then randomly selected one point in each age bin and fitted a curve using these points weighted by the size of each group. This step was repeated 100 times. The median of the 100 curves was taken as the final fitting curve, and the upper and lower quartiles (25 and 75th percentiles) were used to quantify the fitting uncertainty. For the forest which was set to 300 years old circa 2010 in the age map (regarded as an old forest in Besnard et al., 2021), it was difficult for us to build reasonable growth curves if its surrounding forests are all old forests, due to the lack of data at earlier growth stages (TroBE1 and TroBE2 in Table 2).

3.3.2. Real-Age Forest Simulation

For forests younger than 300 years, we calculated the year when the forest age is 0 based on the observation-based age map (as shown at point B in Figure 2) and removed all the biomass pools (including leaves, woods, and roots) in the year when the forest age was zero, and continued the simulation from that year with varying climate forcings (transient simulation) until 2020 (Figure 2). The experiment protocol used during the calibration can be found in Table S1 in Supporting Information S1.

3.3.3. Stepwise Calibration Approach

The calibration was performed step by step using the MASM algorithm described above in Section 3.2. The best value was defined as the mean value of the parameter corresponding to the last 100 minimum cost functions. The parameters calibrated in each step were kept in the following steps. The calibration was conducted individually for each chosen forest. The calibration made use of the GLASS GPP and LAI products from 2000 to 2009, while the years following 2009 were set aside for independent validation. To establish a biomass-age curve reference, 80% of the data was randomly selected for calibration, and the remaining points were used for independent validation. Posterior parameters were obtained by the following steps:

Step 1. We performed the IBIS model simulation with prescribed GLASS LAI to calibrate the key physiological parameters regulating photosynthesis, including *alpha3, theta3* and *vmax,* to minimize the mismatch between the modeled and GLASS GPP (2000–2009). Since LAI is prescribed to the model instead of predicted, the parameters affecting LAI will not be included in this step.

Step 2. We calibrated key LAI parameters using GLASS LAI as a reference. Three parameters (*tauleaf, aleaf, specla*) related to the leaf carbon allocation, turnover, and structure were calibrated in this step. Besides, parameters regulating leaf phenology were also tuned in each forest type (*tempu_gth* for TeBD, BoBD, and BoND; *tempu_drop* for TroBD). Note that calibrating leaf carbon allocation indirectly constrained allocation to wood and fine root, which can affect biomass calibration in the next step.

Step 3. For forests younger than 300 years old, the modeled biomass at different stand ages from simulations was compared against the fitted biomass-age curve from the initial year (age = 0) to 2020. For older forests, we directly compared the modeled biomass against the reference during 2000–2019 without a clear-cut. Parameters including *tauroot, awood, tauwood, and rgrowth* were calibrated at this step to fit the observed biomass-age curves.

We used a comprehensive index called the distance between indices of simulation and observation (DISO) as our accuracy index to evaluate model simulations (Hu et al., 2019). DISO is a combined statistical value that integrates three widely used indices: Pearson's correlation coefficient (R), root mean square error, and average error (AE). A low value of DISO indicates that the simulation is closer to the observation, while a high value indicates a greater discrepancy between modeled and observed data.

$$\mathrm{R} = \frac{\sum_{i=1}^{n}\left(S_i - \overline{S}\right)\left(O_i - \overline{O}\right)}{\sqrt{\sum_{i=1}^{n}\left(S_i - \overline{S}\right)^2}\sqrt{\sum_{i=1}^{n}\left(O_i - \overline{O}\right)^2}} \tag{8}$$

$$\mathrm{AE} = \frac{1}{n}\sum_{i=1}^{n}\left(S_i - O_i\right) \tag{9}$$

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(S_i - O_i\right)^2} \tag{10}$$

$$\mathrm{DISO} = \sqrt{(R-1)^2 + \left(\frac{AE}{\overline{O}}\right)^2 + \left(\frac{RMSE}{\overline{O}}\right)^2}, \tag{11}$$

where n represents the total number of variables, S_i and O_i denote the simulated and observed variables, and $\overline{S}$ and $\overline{O}$ are averaged values.

4. Results

4.1. Improvement to the Simulation of GPP and LAI

Figures 3 and 4 reveal the temporal dynamics of 8-day GPP and LAI in three forests of different PFTs (TeNE1, TeBD2, TroBE1). The other forests showing similar results can be found in the supplement (Figures S4–S5 in Supporting Information S1). The posterior best parameters with their uncertainties are listed in Table S2 in Supporting Information S1. In forests with distinct GPP seasonality (Figures 3a and 3b), both the prior and posterior estimates are capable of capturing the general seasonal patterns exhibited in the GLASS GPP. Specifically, optimized parameters result in reduced biases of GPP across all the forests compared to the original parameters. For all temperate and boreal forests, the optimized GPP during the growing season has shown obvious improvements. Notably, for all TeNE types and BoNE2 (Figure 3a and Figure S4 in Supporting Information S1), the peak GPP has increased by almost twofold. This is due to a noticeable increase in optimized and vmax, indicating considerable improvements in the maximal photosynthetic rate of leaves (Table S2 in Supporting Information S1). For tropical forests, the optimized vmax has decreased, while alpha3 has increased by nearly 50%. For TroBE (Figure 3c; Figure S4 in Supporting Information S1), which is sensitive to alpha3, the improvement in the intrinsic quantum efficiency has a more pronounced effect on increasing carbon fixation. For TroBD (Figures S4–S15, S17, S18 in Supporting Information S1), which is more sensitive to vmax, the reduction in the maximum photosynthetic rate precisely reduces the overestimation of GPP by default parameters (Figure S4 in Supporting Information S1).

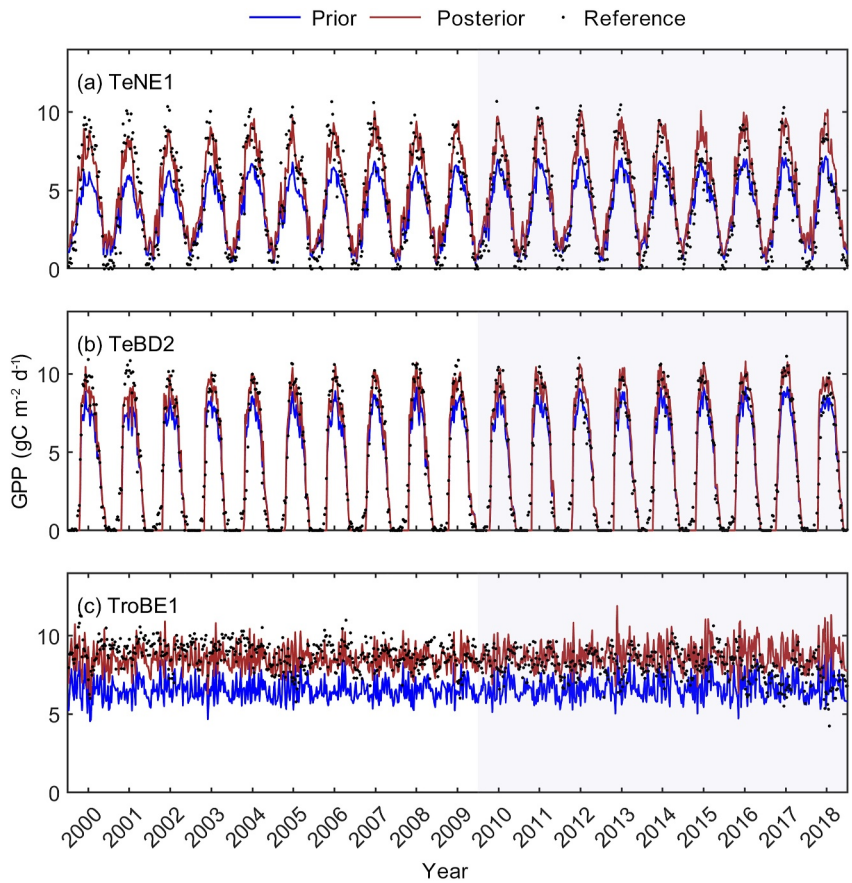

Figure 3. Time series of modeled 8-day GPP before (blue line) and after (red line) calibration, and GLASS GPP (black points) from 2000 to 2018 in the three selected forests. Data from 2000 to 2009 is used for training and the remaining years (shaded) for validation. The calibrated GPP was derived using the best posterior parameters.

In the second step, the LAI is simulated using the model, and the GLASS LAI is used as a reference for calibrating the influential parameters of LAI (*aleaf, tauleaf, specla, phen_gth*). Our findings reveal that the IBIS model performs reasonably well in simulating the timing of leaf flushing and shedding in temperate and boreal deciduous forests, as seen in Figure 4b and Figure S5 in Supporting Information S1. However, the model has difficulty in capturing the peaks and valleys of the GLASS LAI due to the oversimplified phenology parameterization that cannot describe the variation of LAI within growing seasons. For evergreen forests, we find that GLASS LAI displays noticeable seasonal changes, while IBIS shows a constant LAI value. The reason for this is that IBIS calculates the maximum LAI potential for the entire year using the leaf biomass from the previous year and then adjusts LAI values throughout the seasons based on phenological changes. However, IBIS assumes that the leaves in evergreen forests do not undergo phenological changes and remain consistent throughout the year, resulting in a constant LAI value. To ensure the utmost accuracy of the LAI simulation model, it is imperative to rely not only on high-quality observations but also on more refined phenology models that encompass detailed descriptions of leaf dynamics.

4.2. Improvement to the Simulation of Biomass

4.2.1. Construction of Biomass-Age Curves in Forests

We construct the relationship between forest biomass and age for each forest using the Chapman-Richard function, as shown in Figure 5 and Figure S6 in Supporting Information S1. The biomass for young forests accumulates rapidly at the early stages of growth; this increase slows down as trees age, and eventually biomass reaches a steady state, typically at around 150–200 years. In contrast, old forests (TroBE1 and TroBE2) have

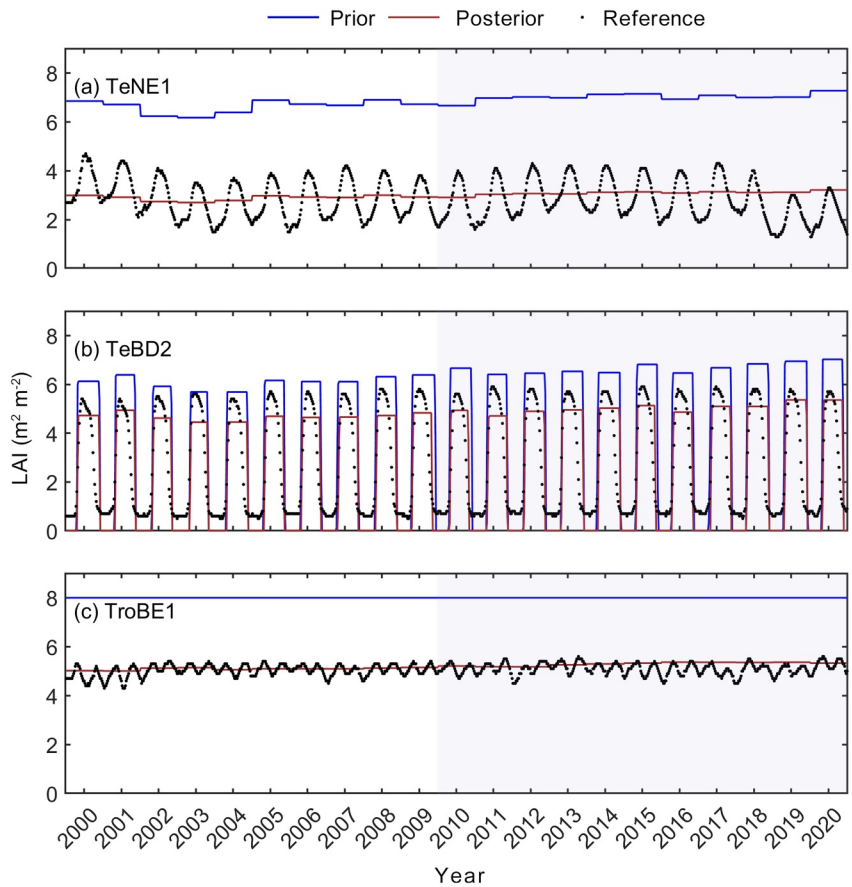

Figure 4. Time series of modeled 8-day leaf area index (LAI) before (blue line) and after (red line) calibration, and GLASS LAI (black points) from 2000 to 2020 in the three selected forests. Data from 2000 to 2009 is used for training and the remaining years (shaded) for validation. The calibrated LAI was derived using the best posterior parameters.

already reached maturity and maximum biomass. Therefore, we only display the annual biomass in 2000–2019 in Figure 5, as it is not possible to establish a reliable growth curve due to the absence of data at the early stage of growth (Figure 5c, Figure S6 in Supporting Information S1). Of all the forests considered, the maximum carbon storage varies among deciduous forests, with boreal forests exhibiting a slightly lower ability to sequester carbon.

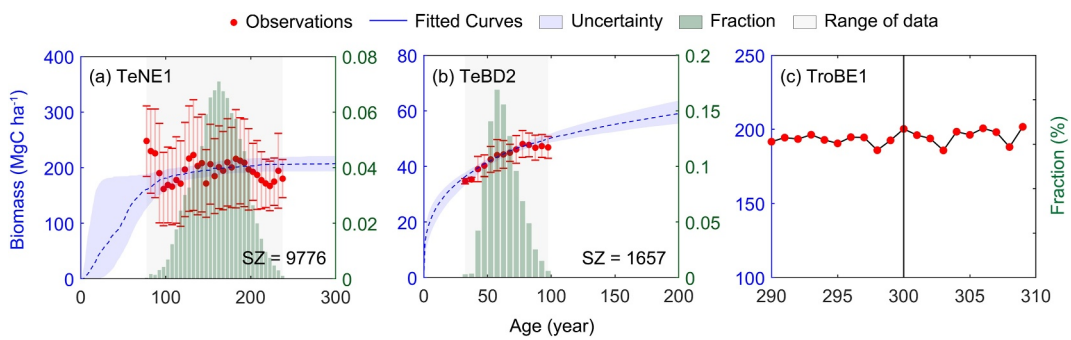

Figure 5. Annually observed biomass with observed forest age in three selected forests. All the curves are fitted based on the parameterized Chapman-Richards growth model. SZ refers to the size of samples when fitting the curves. The blue broken lines refer to the median values of the 100 curves fitted in Section 3.3.2. The blue shaded area represents the 75th and 25th percentiles of the 100 curves. The red points with the error bar are the median of the biomass records corresponding to each age bin (5-year interval). The right *y*-axis represents the fraction of the size of each age bin to the total and is used as a weighting factor for fitting curves (Section 3.3.2).

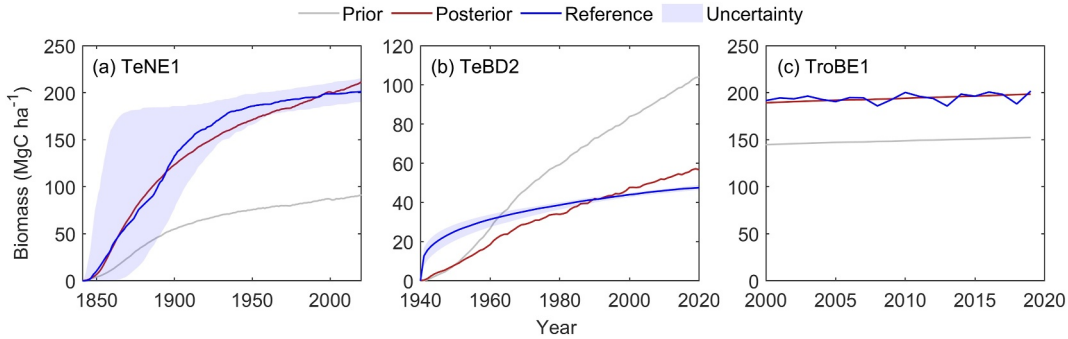

Figure 6. Validation of modeled biomass with default (Prior) and optimized parameters (Posterior) from the initial year (age = 0) to 2020. Blue lines used as reference refer to the fitted biomass-age curves in Section 3.3.2. The blue shaded area represents uncertainties of corresponding curves. The calibrated biomass was derived using the best posterior parameters.

For evergreen forests, biomass accumulates more rapidly and has a higher maximum stock for high-productivity forests. When the forest reaches a steady state, TeNE1 exhibits approximately 2.5 times higher biomass compared to deciduous forests such as BoBD and BoND. We also find a large uncertainty in the curves at the early growth stage, primarily attributable to the absence of biomass data for the entire period of forest growth. This uncertainty varies across different types of forests. For example, the uncertainty during the early growth for TeNE1 can reach up to 170 MgC ha^{-1}, while it shows lower values for TeBD2.

4.2.2. Model Performance in Simulating Biomass With Calibrated and Prior Parameters

We use the fitted biomass-age curves as a reference to calibrate the IBIS parameters affecting biomass (*awood, aroot, tauwood0, rgrowth*). Basically, all optimized biomass-age growth curves (red) can better capture the accumulation of biomass (blue), compared with the curves from the default parameters (gray) (Figure 6, Figure S7 in Supporting Information S1). In the BoBDs, the untuned model shows saturation of biomass at older ages with a much higher maximum carbon stock than the reference. In needleleaf forests, the annual carbon accumulation is seriously underestimated in the original estimates due to the lower accumulation rate at the early stages (Figure 6a, Figure S7 in Supporting Information S1). The posterior parameters result in curves that improved the carbon stocks in forests over 50 years old, but still underestimate carbon accumulation rates during early growth stages. For two mature forests older than 300 years, the biomass at the stable state fluctuates around 200 MgC ha^{-1} with slight interannual variations. These biomass values are severely underestimated by the original model but are much better simulated after the calibration. For instance, in TeBD2, the age-calibration decreased the wood turnover time (*tauwood0*), indicating that the overestimation of original biomass is mainly due to slower turnover. In contrast, the biomass calibration corrected the underestimation of biomass in TeNE1 by increasing *awood* and *tauwood0*. During the early phases of forest growth, greater wood allocation means more resources being allocated to wood growth, resulting in an accelerated rate of biomass accumulation. However, the original model parameters underestimated the carbon allocation to wood and overestimated the turnover speed of wood carbon pools in TeNE1.

We also found that our stepwise calibrated parameters showed large differences in posterior values (e.g., *tauwood0*) even within the same PFT (Table S2 in Supporting Information S1). This is partly because ecosystem characteristics, density, or disturbance history can vary substantially within a PFT across the region. Some studies have shown that the carbon turnover time exhibits noticeably spatial differences and is closely related to temperature and precipitation (Carvalhais et al., 2014), which indicates that each specific environment can influence model parameters when performing calibration. This is also the reason why we emphasize the importance of conducting calibration at the pixel scale rather than a parameter per PFT, and this study has provided the potential for pixel-scale calibration with our surrogate-based calibration framework.

4.3. Back-Compatibility Check for the Stepwise Calibration Approach

The consistency of the stepwise strategy has been evaluated with a back-compatibility check after the third calibration. Figure 7 and Figure S8 in Supporting Information S1 summarize the prior and posterior performance

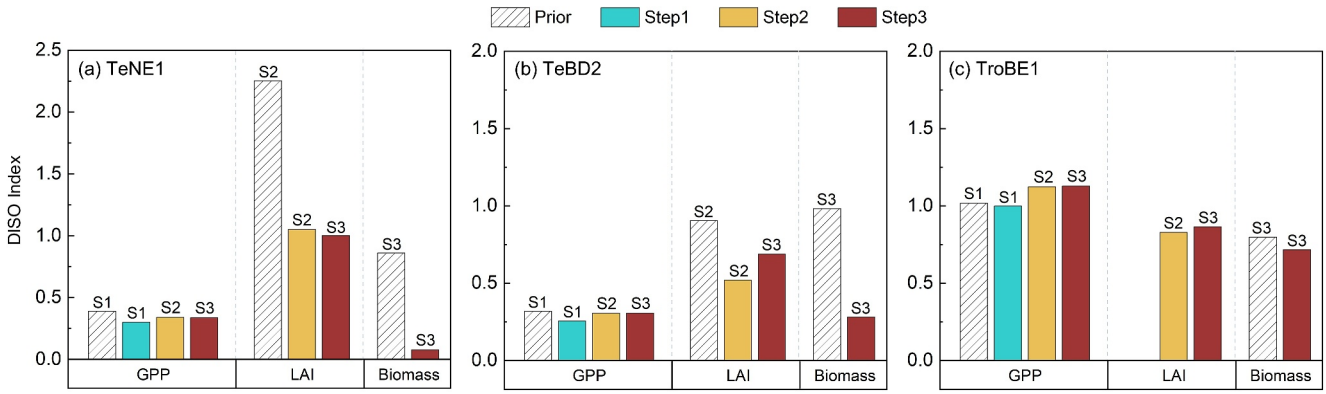

Figure 7. Accuracy index between model outputs and references of gross primary productivity (2000–2009), leaf area index (LAI) (2000–2009), and biomass (Bio) (from the initial year to 2020) in three steps. The bars with a light color are results obtained using prior parameters, while those with a dark color are optimized indexes. For TroBE1 with missing values for prior LAI, the modeled LAI with default parameters is a fixed value (8 m^2/m^2) from year to year, making it unable to calculate the correlation.

of the model-data fit of each step, with an obvious DISO reduction in GPP, LAI, and biomass when compared to the prior model. The optimization with the final set of parameters does not appreciably degrade the fit to GLASS GPP that is used as a reference in the first steps (only minor changes of the DISO indexes occurred in our results), even though LAI and biomass have been improved greatly. In the 25 forests (Figure 7 and Figure S8 in Supporting Information S1), only three forests (TroBE1, BoNE1, and TeBE1) exhibited worse GPP and three forests (BoND1, TeNE2, and TeBD3) exhibited worse LAI after the step 3 calibration. This may due to the impact of changed biomass on respiration, which affects the carbon allocated in leaves, consequently resulting in changes in GPP and LAI. Nevertheless, this effect is small in most of the forests. We also make a final comparison between prior and posterior values of GPP, LAI, and biomass from 2000 to 2020 (Figure S9 in Supporting Information S1). The optimization results in a substantial reduction of errors in LAI (−35.3%), GPP (−28.5%), and biomass (−74.6%). For biomass, the DISO for most forests has been reduced by over 70%. The improvements of LAI and GPP are also remarkable, although some errors are derived from the imperfect model structure. Nevertheless, there remain considerable errors in LAI not reduced even when phenological parameters have been taken into account during the calibration, which is mainly due to the oversimplified phenology parameterization in IBIS. In general, the stepwise calibration strategy used in this study enables a better constraint of both "fast" and "slow" processes in IBIS, with the application of remotely sensed GPP, LAI, and biomass.

5. Discussion

5.1. Importance of Stand Age in Improving Ecosystem Carbon Estimation

Forest age offers a starting point for LSMs to correctly simulate forest recovery from the last disturbance (Pan et al., 2011; Pugh, Arneth, et al., 2019; Zhu et al., 2019). Previous LSMs often omitted stand age, and start simulating carbon pools from an equilibrium state (i.e., model spin-up). This will result in a systematic overestimation of carbon pools since young forests are not yet equilibrated. Similar findings were reported by Thum et al. (2017), wherein the authors demonstrated an overestimation of modeled biomass when forest age was not considered.

Attempting to calibrate the model under such a steady state will lead to biases in parameters. Using these biased parameters, LSMs will not be able to correctly simulate the regrowth of forests, even when real age is considered in the simulation. For example, we repeated the step 3 calibration on each forest but without considering real age (Figure S10 in Supporting Information S1). The calibrated biomass derived in such a way ($Post_{NAge}$) exhibited larger differences compared to the observations than our previous results ($Post_{Age}$). In this study, parameter inversion under the SS mostly overestimated wood turnover time (Figure S11 in Supporting Information S1, *tauwood0*) but underestimated wood allocation rates (except TroBE types). This result agrees with Ge et al. (2018), who demonstrated that biased inverted parameters under the SS overestimated mean turnover times in young-aged ecosystems, which in turn led to obvious deviations in simulating carbon effluxes (e.g., litterfall and respiration). With proper consideration of carbon pool initialization and parameter

retrieval, our calibration framework yields a better model capacity for simulating carbon stocks (Figure S10 in Supporting Information S1).

The biomass error due to the omission of age will cause further biases in ecosystem carbon budgets (Amiro et al., 2010; Bellassen et al., 2010). The biomass of young forests is much smaller than equilibrated old forests, which leads to much smaller Ra. Meanwhile, young forests have smaller LAI, resulting in smaller GPP. The coupling effects from Ra and GPP can drive changes in carbon balances (Amiro et al., 2010; Williams et al., 2014). Biased parameters retrieved from disequilibrium initialization further propagate into changes in NBP in a transient simulation. We reveal that the parameters retrieved from noage-calibration (SS) can result in underestimation of the mean annual NBP during 2000–2020 (Figure S12 in Supporting Information S1). Our age-calibrated parameters increase NBP by an average of 50 gC m^{-2} yr^{-1} across all forests, especially in the BoNE type, the NBP increased by 118 gC m^{-2} yr^{-1} on average, increasing the estimate of the carbon sink in young forests. Such underestimation of NBP under SS assumption is also supported by Ge et al. (2018) and Yu et al. (2014). It should be noted that we did not calibrate the parameters affecting heterotrophic respiration, thus the NBP may be biased. Nevertheless, our results still highlight the importance of considering age in calibrating LSMs for correct carbon flux simulations.

5.2. Uncertainties for Age-Dependent Biomass Assimilation

The accuracy of the biomass-age curve is crucial for model calibration. Therefore, the uncertainties in the biomass and age maps will have a big impact on the carbon stock estimates. Both of these two maps are the products of machine learning models trained on observations from the ground and space. Errors in training data and models will directly affect the spatial accuracy of these products, and then ultimately propagate to our fitted curves (Raczka et al., 2021).

An additional factor contributing to uncertainty is the imperfect fitting of biomass-age curves. First, we averaged age pixels (1 km) within 0.1° to match the biomass product, which may result in a loss of variability in ages. Second, to ensure sufficient data for curve fitting, we collected data within a 5° × 5° window instead of temporal features (Section 3.3.1). This "space for time" approach may ignore the spatial variation in growth rates. Although our framework provides a good method for assimilating age-dependent biomass in LSMs, high-resolution and long-time-series data are still needed to provide a more robust and reliable relationship between biomass and age. In this study, the fitting curves have large uncertainty at young forest ages (Figure 5). This is because our data does not cover the entire period of forest growth and the predictions at such ages rely on extrapolation of the observations using the parameterized equation (Equation 7). Nevertheless, when we use the fitting curve within the observed forest age range as the reference in the calibration (the gray shaded area in Figure 5 and Figure S6 in Supporting Information S1), the model performance does not deteriorate obviously. By comparing the fitted curves with field survey data, we found obvious mismatches in this comparison (results not shown), which are likely due to scale differences. For example, small plots are associated with edge effects that may cause discrepancies in tree representation between remote sensing footprints and field plots (Nesha et al., 2022). In large-scale regional studies, more field data needs to be collected for exploring a more convincing relationship than that obtained solely through remote sensing products. This could be achieved by constraining the fitted curves at different scale through field observations (Shang et al., 2023), or upscaling tree/plot level observations to a regional scale with the support of spatial remote sensing data. Meanwhile, calibrating multiple parameters of complex DGVMs at high resolution is challenging due to computational cost. We still need to develop a faster calibration approach that can facilitate regional-scale parameter optimization by integrating remote sensing and field observations.

Another source of uncertainty is the oversimplified allocation parametrization in the IBIS model, which assumes constant allocations of carbon among plant tissues (Merganičová et al., 2019). However, in reality, plants may alter their allocation strategies under the influence of climate changes (Collalti et al., 2020; Xu et al., 2012), and during different periods. The static allocation assumption has been reported to considerably affect the simulated carbon pools (Pappas et al., 2013; Xia et al., 2017, 2019). Such uncertainties in model structure will be transmitted to the calibrations and simulations, leading to biases in both posterior parameters and modeled biomass. To improve the simulation of carbon dynamics, adaptive carbon allocation strategies that account for varying climate changes are needed, as suggested by Xia et al. (2019).

5.3. Stepwise Versus Simultaneous Calibration Approach

Adding extra data streams helped to constrain unresolved sub-spaces of the total parameter space. In multiple constraints, simultaneous and stepwise calibration approaches are both options for integrating several observations with the LSMs.

For multi-objective calibration, simultaneous optimization is mathematically advantageous, which stems from its preservation of strong parameter linkages across diverse processes, coupled with its more rigorous handling of error correlations within the same inversion (MacBean et al., 2016). When dealing with disparate numbers of observations, like GPP and biomass in our study, the different constraint effects provided by such unbalanced data could bring a performance conflict for each variable (Oberpriller et al., 2021). That is to say, the optimized results may match well with one set of data and deviate from others (Figure S13 in Supporting Information S1, Opt_Sim2). Assigning appropriate weights to different observations within the same objective function is crucial but challenging for optimization, as it directly influences the constraints imposed by each reference, and may result in an overestimation of posterior errors for parameters constrained by the abundant data streams (Thum et al., 2017; Wutzler & Carvalhais, 2014). We performed three simultaneous calibrations with different weighting strategies (Table S3 in Supporting Information S1) in Figure S13 in Supporting Information S1, and found different performances in improving GPP, LAI, and Biomass. Simultaneous calibration may also provide alternative sets of posterior parameters to achieve a similar reduction in the model-data inconsistency, a phenomenon referred to as parameter equifinality, which requires more model evaluations to find the optimal sets.

Given the complexities involved in designing the cost function and the computational costs, we tend to prefer adopting a step-wise calibration framework, which ensures that each data stream only constrains the most relevant processes of the model based on their relative contributions. This way can move beyond the model equifinality of multi-objective in calibration, and avoid weight problems that need to be considered in simultaneous calibration, and also bring more information at different spatial and temporal scales (Peylin et al., 2016). Another reason for us to use a stepwise approach is that we can separate our calibration into "fast" and "slow" processes calibration. In the "fast" steps (step 1 and step 2), we only need a short simulation period for GPP and LAI to rapidly reach a steady state, while in the "slow" step (step 3), we need a longer spin-up to optimize biomass.

Nonetheless, stepwise optimization still comes with its challenges, especially the propagation of parameter error matrices among each step. Because the parameters calibrated from previous steps are assumed constant in later steps, its error is thus taken by the parameters in later steps. Although we cannot avoid the error propagation, our stepwise calibration still results in better GPP, LAI, and biomass of most forests compared with default parameters after the final step. This is because we use each data constraint only for its most relevant parts of the model, which seems more feasible and straightforward. So, for stepwise calibration, it is still essential to perform a back-compatibility check to ensure that subsequent calibrations have an acceptable impact on other variables.

Furthermore, considering that the parameters (aleaf, tauleaf, specla, and tempu_gth) associated with LAI in Step 2, also have a notable effect on biomass (Figure 1), we conduct a test to see if we could optimize the same 11 parameters only limited by GPP and biomass (Table S4 in Supporting Information S1). In this scenario (GPP-Bio), we keep Step 1 unchanged, while calibrating 8 parameters using biomass in Step 2. According to the results presented in Table S4 in Supporting Information S1, the modeled biomass does not show an obvious difference with our three-step calibration (GPP-LAI-Bio). However, it is observed that the accuracy of the modeled LAI noticeably decreases. For evergreen forests, the decrease in LAI accuracy does not greatly impact GPP. However, for deciduous forests, there is a 45% increase in DISO (i.e., worse GPP). In our three-step calibration experiment, we find that GPP, LAI, and biomass all have good performance in each step, and subsequent calibrated parameters have minimal impact on the variables of the first two steps (Figure 7 and Figure S8 in Supporting Information S1). This result indicates our three-step calibration framework can better improve carbon cycle simulation in IBIS.

5.4. Challenges in Large-Scale Assimilation of GPP, LAI and Biomass

This study conducted a stepwise calibration across 25 forests, enhancing the ability of the IBIS model to simulate GPP, LAI, and biomass. It provided an initial framework and theoretical foundation for establishing a GPP-LAI-Biomass calibration system. Although these calibrations are performed at individual forests, they can be potentially implemented at a regional or even global scale. The emergence of machine learning methods provides an opportunity to scale up the parameters obtained in this study. Compared with previously used parameters that

are fixed for each PFT, the parameters derived from scaling up site-level calibrations can better represent the heterogeneity of vegetation within each PFT, thus resulting in much better simulations of carbon dynamics. There have been successful implementations of such methods to reduce model uncertainties of GPP and LAI over North America (Ma et al., 2022a) and also soil organic carbon over the Conterminous United States (Tao et al., 2020).

Despite the potential, there remain several challenges to accurately simulate biomass evolution at the global scale. First, most grid cells include different PFTs and cohorts with different ages which most models cannot deal with. To fully consider the sub-grid heterogeneity requires explicitly modeling efforts to represent various age categories present within each PFT (Bellassen et al., 2010; Zaehle et al., 2006). Second, the current forest age data used in our study is only available in 2010 (Besnard et al., 2021). The forest age distribution before and after that year still needs to be reconstructed. Finally, we can only assume temporal consistency of parameters for our calibration. Although this assumption is accepted by most of the LSMs, it may still fail due to structural and biogeophysical differences between young and old forests. With the advent of new forest age maps, and if those maps were accurate enough, we could nevertheless use a space-for-time approach where LAI and GPP high-resolution values in the present period are matched to certain forest age, allowing us to derive GPP-age and LAI-age relationships that could be used for calibration.

6. Conclusions

Forest age is an effective indicator of the duration since the last disturbance and drives forest NBP. In LSMs, forests are often assumed undisturbed, leading to unrealistic forest growth patterns and inaccurate estimations of carbon stocks. Joint assimilation of remotely sensed GPP, LAI, and age-dependent biomass has been proven feasible in 25 forests across the world, and provided great constraints on parameters related to photosynthesis, allocation, phenology, and respiration processes. With consideration of stand age, modeled biomass was greatly improved in each forest, and also revealed a marked impact of age on NBP simulation. Our three-step calibration framework offers a reliable and robust approach for optimizing both "fast" and "slow" processes in IBIS, and has the potential to be applied worldwide using machine learning-based surrogate modeling. The quality of referenced data used in our framework is of utmost importance for calibration, and the performance of our fitted biomass-age curves exerts a major influence on carbon allocation and turnover rates. We expect that the availability of reliable observations or remotely sensed products will offer additional insights for optimizing future carbon flux simulations.

Data Availability Statement

The scripts, and instructions of the MASM method can be archived through Zenodo at https://doi.org/10.5281/zenodo.6953354 (Ma et al., 2022b). The version of IBIS model and codes for fitting biomass-age curves used in this paper are available at https://doi.org/10.5281/zenodo.8376379 (Ma, 2023).

Acknowledgments
This study was partially supported by the Open Research Program of the International Research Center of Big Data for Sustainable Development Goals, Grant CBAS2022ORP01, and the National Natural Science Foundation of China (No. 42090011), and the China Scholarship Council No. 202106270126 to R. Ma, and the European Space Agency Climate Change Initiative (ESA CCI) RECCAP2 project (No. ESRIN/4000123002/18/I-NB). Y.Z. is supported by the European Commission, Horizon 2020 Framework Programme 4C (No. 821003), the Science and Technology Major Project of Tibetan Autonomous Region of China (XZ202201ZD0005G01). We gratefully acknowledge the data support from National Earth System Science Data Center, National Science & Technology Infrastructure of China (http://www.geodata.cn), and thank Dr. Wenping Yuan for providing the code of the IBIS model which was used in this paper.

References

Alton, P. B. (2013). From site-level to global simulation: Reconciling carbon, water and energy fluxes over different spatial scales using a process-based ecophysiological land-surface model. *Agricultural and Forest Meteorology*, *176*, 111–124. https://doi.org/10.1016/j.agrformet.2013.03.010

Amiro, B. D., Barr, A. G., Barr, J. G., Black, T. A., Bracho, R., Brown, M., et al. (2010). Ecosystem carbon dioxide fluxes after disturbance in forests of North America. *Journal of Geophysical Research*, *115*(G4). https://doi.org/10.1029/2010JG001390

Bacour, P. P., Peylin, P., MacBean, N., Rayner, P. J., Delage, F., Chevallier, F., et al. (2015). Joint assimilation of eddy covariance flux measurements and FAPAR products over temperate forests within a process-oriented biosphere model. *Journal of Geophysical Research: Biogeosciences*, *120*(9), 1839–1857. https://doi.org/10.1002/2015JG002966

Barman, R., Jain, A. K., & Liang, M. L. (2014). Climate-driven uncertainties in modelling terrestrial gross primary production: A site level to global-scale analysis. *Global Change Biology*, *20*(5), 1394–1411. https://doi.org/10.1111/gcb.12474

Bastrikov, V., MacBean, N., Bacour, C., Santaren, D., Kuppel, S., & Peylin, P. (2018). Land surface model parameter optimization using in situ flux data: Comparison of gradient-based versus random search algorithms (a case study using ORCHIDEE v1.9.5.2). *Geoscientific Model Development*, *11*(12), 4739–4754. https://doi.org/10.5194/gmd-11-4739-2018

Bellassen, V., Le Maire, G., Dhôte, J. F., Ciais, P., & Viovy, N. (2010). Modeling forest management within a global vegetation model-Part 1: Model structure and general behavior. *Ecological Modeling*, *221*(20), 2458–2474. https://doi.org/10.1016/j.ecolmodel.2010.07.008

Besnard, S., Koirala, S., Santoro, M., Weber, U., Nelson, J., Gütter, J., et al. (2021). Mapping global forest age from forest inventories, biomass and climate data. *Earth System Science Data*, *13*(10), 4881–4896. https://doi.org/10.5194/essd-13-4881-2021

Botta, A., & Foley, J. A. (2002). Effects of climate variability and disturbances on the Amazonian terrestrial ecosystems dynamics. *Global Biogeochemical Cycles*, *16*(4), 1070. https://doi.org/10.1029/2000GB001338

Bukoski, J. J., Cook-Patton, S. C., Melikov, C., Ban, H., Chen, J. L., Goldman, E. D., et al. (2022). Rates and drivers of aboveground carbon accumulation in global monoculture plantation forests. *Nature Communications*, *13*(1), 4206. https://doi.org/10.5194/bg-15-5801-2018

Carvalhais, N., Forkel, M., Khomik, M., Bellarby, J., Jung, M., Migliavacca, M., et al. (2014). Global covariation of carbon turnover times with climate in terrestrial ecosystems. *Nature*, *514*(7521), 213–217. https://doi.org/10.1038/nature13731

Carvalhais, N., Reichstein, M., Ciais, P., Collatz, G. J., Mahecha, M. D., Montagnani, L., et al. (2010). Identification of vegetation and soil carbon pools out of equilibrium in a process model via eddy covariance and biometric constraints. *Global Change Biology*, *16*(10), 2813–2829. https://doi.org/10.1111/j.1365-2486.2010.02173.x

Carvalhais, N., Reichstein, M., Seixas, J., Collatz, G. J., Pereira, J. S., Berbigier, P., et al. (2008). Implication of the carbon cycle steady state assumption for biogeochemical modelling performance and inverse parameter retrieval. *Global Biogeochemical Cycles*, *22*(2), GB2007. https://doi.org/10.1029/2007GB003033

Chaney, N. W., Herman, J. D., Ek, M. B., & Wood, E. F. (2016). Deriving global parameter estimates for the Noah land surface model using FLUXNET and machine learning. *Journal of Geophysical Research: Atmosphere*, *121*(22), 13218–13235. https://doi.org/10.1002/2016JD024821

Ciais, P., Schelhaas, M. J., Zaehle, S., Piao, S. L., Cescatti, A., Liski, J., et al. (2008). Carbon accumulation in European forests. *Nature Geoscience*, *1*(7), 425–429. https://doi.org/10.1038/ngeo233

Collalti, A., Ibrom, A., Stockmarr, A., Cescatti, A., Alkama, R., Fernández-Martínez, M., et al. (2020). Forest production efficiency increases with growth temperature. *Nature Communications*, *11*(1), 5322. https://doi.org/10.1038/s41467-020-19187-w

Cunha, A. P. M. A., Alvalá, R. C. S., Sampaio, G., Shimizu, M. H., & Costa, M. H. (2013). Calibration and validation of the integrated biosphere simulator (IBIS) for a Brazilian semiarid region. *Journal of Applied Meteorology and Climatology*, *52*(12), 2753–2770. https://doi.org/10.1175/JAMC-D-12-0190.1

Famiglietti, C. A., Smallman, T. L., Levine, P. A., Flack-Prain, S., Quetin, G. R., Meyer, V., et al. (2021). Optimal model complexity for terrestrial carbon cycle prediction. *Biogeosciences*, *18*(8), 2727–2754. https://doi.org/10.5194/bg-18-2727-2021

Farquhar, G. D., von Caemmerer, S., & Berry, J. A. (1980). A biochemical model of photosynthetic CO_2 assimilation in leaves of C_3 species. *Planta*, *149*(1), 78–90. https://doi.org/10.1007/BF00386231

Fer, I., Kelly, R., Moorcroft, P. R., Richardson, A. D., Cowdery, E. M., & Dietze, M. C. (2018). Linking big models to big data: Efficient ecosystem model calibration through Bayesian model emulation. *Biogeosciences Discussions*, *15*, 5801–5830. https://doi.org/10.5194/bg-15-5801-2018

Foley, J. A., Prentice, C., Ramankutty, N., Levis, S., Pollard, D., Sitch, S., & Haxeltine, A. (1996). An integrated biosphere model of land surface processes, terrestrial carbon balance, and vegetation dynamics. *Global Biogeochemical Cycles*, *10*(4), 603–628. https://doi.org/10.1029/96GB02692

Forkel, M., Druke, M., Thurner, M., Dorigo, W., Schaphoff, S., Thonicke, K., et al. (2019). Constraining modeled global vegetation dynamics and carbon turnover using multiple satellite observations. *Scientific Reports*, *9*(1), 18757. https://doi.org/10.1038/s41598-019-55187-7

Ge, R., He, H., Ren, X., Zhang, L., Yu, G., Smallman, T. L., et al. (2018). Underestimated ecosystem carbon turnover time and sequestration under the steady state assumption: A perspective from long-term data assimilation. *Global Change Biology*, *25*(3), 938–953. https://doi.org/10.1111/gcb.14547

Gelman, A., & Rubin, D. B. (1992). Inference from iterative simulation using multiple sequences. *Statistical Science*, *7*(4), 457–472. https://doi.org/10.1214/ss/1177011136

Gong, W., & Duan, Q. Y. (2017). An adaptive surrogate modeling-based sampling strategy for parameter optimization and distribution estimation (ASMO-PODE). *Environmental Modelling & Software*, *95*, 61–75. https://doi.org/10.1016/j.envsoft.2017.05.005

Hansen, M. C., Potapov, P. V., Moore, R., Hancher, M., Turubanova, S. A., Tyukavina, A., et al. (2013). High-resolution global maps of 21st-century forest cover change. *Science*, *342*(6160), 850–853. https://doi.org/10.1126/science.1244693

Harris, I. (2021). CRU JRA v2.2: A forcings dataset of gridded land surface blend of Climatic Research Unit (CRU) and Japanese reanalysis (JRA) data [Dataset]. *Centre for Environmental Data Analysis, University of East Anglia Climatic Research Unit*. Retrieved from https://catalogue.ceda.ac.uk/uuid/4bdf41fc10af4caaa489b14745c665a6

Harris, I., Jones, P. D., Osborn, T. J., & Lister, D. H. (2014). Updated high-resolution grids of monthly climatic observations – The CRU TS3.10 Dataset. *International Journal of Climatology*, *34*(3), 623–642. https://doi.org/10.1002/joc.3711

Harris, I., Osborn, T. J., Jones, P., & Lister, D. (2020). Version 4 of the CRU TS monthly high-resolution gridded multivariate climate dataset. *Scientific Data*, *7*(1). https://doi.org/10.1038/s41597-020-0453-3

Hawkins, L. R., Rupp, D. E., NcNeall, D. J., Li, S., Betts, R. A., Mote, P. W., et al. (2019). Parametric sensitivity of vegetation dynamics in the TRIFFID model and the associated uncertainty in projected climate change impacts on western U.S. Forests. *Journal of Advances in Modeling Earth Systems*, *11*(8), 2787–2813. https://doi.org/10.1029/2018MS001577

He, X., Xu, T., Bateni, S. M., Ki, S. J., Xiao, J., Liu, S., et al. (2021). Estimation of turbulent heat fluxes and gross primary productivity by assimilating land surface temperature and leaf area index. *Water Resources Research*, *57*(11), e2020WR028224. https://doi.org/10.1029/2020WR028224

Heinrich, V. H. A., Dalagnol, R., Cassol, H. L. G., Rosan, T. M., de Almeida, C. T., Silva Junior, C. H. L., et al. (2021). Large carbon sink potential of secondary forests in the Brazilian Amazon to mitigate climate change. *Nature Communications*, *12*(1), 1785. https://doi.org/10.1038/s41467-021-22050-1

Heinrich, V. H. A., Vancutsem, C., Dalagnol, R., Rosan, T. M., Fawcett, D., Silva-Junior, C. H. L., et al. (2023). The carbon sink of secondary and degraded humid tropical forests. *Nature*, *615*(7952), 436–442. https://doi.org/10.1038/s41586-022-05679-w

Homer, C., Dewitz, J., Jin, S., Xian, G., Costello, C., Danielson, P., et al. (2020). Conterminous United States land cover change patterns 2001–2016 from the 2016 national land cover database. *ISPRS Journal of Photogrammetry and Remote Sensing*, *162*, 184–199. https://doi.org/10.1016/j.isprsjprs.2020.02.019

Hu, Z., Chen, X., Zhou, Q., Chen, D., & Li, J. (2019). DISO: A rethink of Taylor diagram. *International Journal of Climatology*, *39*(5), 2825–2832. https://doi.org/10.1002/joc.5972

Kobayashi, S., Ota, Y., Harada, Y., Ebita, A., Moriya, M., Onoda, H., et al. (2015). The JRA-55 reanalysis: General specifications and basic characteristics. *Journal of the Meteorological Society of Japan. Ser. II*, *93*(1), 5–48. https://doi.org/10.2151/jmsj.2015-001

Krause, A., Papastefanou, P., Gregor, K., Layritz, L. S., Zang, C. S., Buras, A., et al. (2022). Quantifying the impacts of land cover change on gross primary productivity globally. *Scientific Reports*, *12*(1), 18398. https://doi.org/10.1038/s41598-022-23120-0

Kucharik, C. J., Foley, J. A., Delire, C., Fisher, V. A., Coe, M. T., Lenters, J. D., et al. (2000). Testing the performance of a dynamic global ecosystem model: Water balance, carbon balance, and vegetation structure. *Global Biogeochemical Cycles*, *14*(3), 795–825. https://doi.org/10.1029/1999GB001138

Kumar, S. V., Mocko, D. M., Wang, S., Peters-Lidard, C. D., & Borak, J. (2019). Assimilation of remotely sensed leaf area index into the Noah-MP land surface model: Impacts on water and carbon fluxes and states over the continental United States. *Journal of Hydrometeorology*, *20*(7), 1359–1377. https://doi.org/10.1175/JHM-D-18-0237.1

Lafont, S., Zhao, Y., Calvet, J. C., Peylin, P., Ciais, P., Maignan, F., & Weiss, M. (2012). Modeling LAI, surface water and carbon fluxes at high-resolution over France: Comparison of ISBA-A-gs and ORCHIDEE. *Biogeosciences*, *9*(1), 439–456. https://doi.org/10.5194/bg-9-439-2012

Li, J., Wang, Y.-P., Duan, Q., Lu, X., Pak, B., Wiltshire, A., et al. (2016). Quantification and attribution of errors in the simulated annual gross primary production and latent heat fluxes by two global land surface models. *Journal of Advances in Modeling Earth Systems*, *8*(3), 1270–1288. https://doi.org/10.1002/2015MS000583

Liang, S. L., Cheng, J., Jia, K., Jiang, B., Liu, Q., Xiao, Z. Q., et al. (2021). The global land surface satellite (GLASS) product suite. *Bulletin of the American Meteorological Society*, *102*(2), E323–E337. https://doi.org/10.1175/BAMS-D-18-0341.1

Liu, X., Chen, F., Barlage, M., Zhou, G., & Niyogi, D. (2016). Noah-MP-Crop: Introducing dynamic crop growth in the Noah-MP land surface model. *Journal of Geophysical Research: Atmospheres*, *121*(23), 139553–213972. https://doi.org/10.1002/2016JD025597

Liu, Y., Yu, G., Wang, Q., & Zhang, Y. (2014). How temperature, precipitation and stand age control the biomass carbon density of global mature forests. *Global Ecology and Biogeography*, *23*(3), 323–333. https://doi.org/10.1111/geb.12113

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

Ma, R. (2023). Integrated biosphere simulator model (IBIS) (v1.0) [Software]. *Zenodo*. https://doi.org/10.5281/zenodo.8376379

Ma, R., Xiao, J. F., Liang, S. L., Ma, H., He, T., Guo, D., et al. (2022a). Pixel-level parameter optimization of a terrestrial biosphere model for improving estimation of carbon fluxes with an efficient model–data fusion method and satellite-derived LAI and GPP data. *Geoscientific Model Development*, *15*(17), 6637–6657. https://doi.org/10.5194/gmd-15-6637-2022

Ma, R., Xiao, J. F., Liang, S. L., Ma, H., He, T., Guo, D., et al. (2022b). Modified adaptive surrogate modeling (MASM) (v1.0) [Software]. *Zenodo*. https://doi.org/10.5281/zenodo.6953354

MacBean, N., Bacour, C., Raoult, N., Bastrikov, V., Koffi, E. N., Kuppel, S., et al. (2022). Quantifying and reducing uncertainty in global carbon cycle predictions: Lessons and perspectives from 15 years of data assimilation studies with the ORCHIDEE terrestrial biosphere model. *Global Biogeochemical Cycles*, *36*(7), e2021GB007177. https://doi.org/10.1029/2021GB007177

MacBean, N., Peylin, P., Chevallier, F., Scholze, M., & Schürmann, G. (2016). Consistent assimilation of multiple data streams in a carbon cycle data assimilation system. *Geoscientific Model Development*, *9*(10), 3569–3588. https://doi.org/10.5194/gmd-9-3569-2016

Merganičová, K., Merganič, J., Lehtonen, A., Vacchiano, G., Sever, M. Z. O., Augustynczik, A. L. D., et al. (2019). Forest carbon allocation modelling under climate change. *Tree Physiology*, *39*(12), 1937–1960. https://doi.org/10.1093/treephys/tpz105

Nesha, K., Herold, M., Sy, V. D., Bruin, S., Araza, A., Málaga, N., et al. (2022). Exploring characteristics of national forest inventories for integration with global space-based forest biomass data. *The Science of the Total Environment*, *850*(4), 157788. https://doi.org/10.1016/j.scitotenv.2022.157788

Oberpriller, J., Cameron, D. R., Dietze, M. C., & Hartig, F. (2021). Towards robust statistical inference for complex computer models. *Ecology Letters*, *24*(6), 1251–1261. https://doi.org/10.1111/ele.13728

Pan, Y., Chen, J. M., Birdsey, R. B., McCullough, K., He, L., & Deng, F. (2011). Age structure and disturbance legacy of North American forests. *Biogeosciences*, *8*(3), 715–732. https://doi.org/10.5194/bg-8-715-2011

Pappas, C., Fatichi, S., Leuzinger, S., Wolf, A., & Burlando, P. (2013). Sensitivity analysis of a process-based ecosystem model: Pinpointing parameterization and structural issues. *Journal of Geophysical Research: Biogeosciences*, *118*(2), 505–528. https://doi.org/10.1002/jgrg.20035

Peylin, P., Bacour, C., MacBean, N., Leonard, S., Rayner, P. J., Kuppel, S., et al. (2016). A new step-wise carbon cycle data assimilation system 2 using multiple data streams to constrain the simulated land 3 surface carbon cycle. *Geoscientific Model Development*, *9*(9), 3321–3346. https://doi.org/10.5194/gmd-9-3321-2016

Pietsch, S. A., & Hasenauer, H. (2006). Evaluating the self-initialization procedure for large-scale ecosystem models. *Global Change Biology*, *12*(9), 1658–1669. https://doi.org/10.1111/j.1365-2486.2006.01211.x

Poulter, B., Aragão, L., Andela, N., Bellassen, V., Ciais, P., Kato, T., et al. (2019). *The global forest age dataset and its uncertainties (GFADv1.1)*. NASA National Aeronautics and Space Administration, PANGAEA. https://doi.org/10.1594/PANGAEA.897392

Pugh, T. A. M., Arneth, A., Kautz, M., Poulter, B., & Smith, B. (2019). Important role of forest disturbances in the global biomass turnover and carbon sinks. *Nature Geoscience*, *12*(9), 730–735. https://doi.org/10.1038/s41561-019-0427-2

Pugh, T. A. M., Lindeskog, M., Smith, B., Poulter, B., Arneth, A., Haverd, V., & Calle, L. (2019). Role of forest regrowth in global carbon sink dynamics. *Proceedings of the National Academy of Sciences*, *116*(10), 4382–4387. https://doi.org/10.1073/pnas.181051211

Raczka, B., Hoar, T. J., Duarte, H. F., Fox, A. M., Anderson, J. L., Bowling, D. R., & Lin, J. C. (2021). Improving CLM5.0 biomass and carbon exchange across the western United States using a data assimilation system. *Journal of Advances in Modeling Earth Systems*, *13*(7), e2020MS002421. https://doi.org/10.1029/2020MS002421

Reichstein, M., Camps-Valls, G., Stevens, B., Jung, M., Denzler, J., Carvalhais, N., & Prabhat (2019). Deep learning and process understanding for data-driven Earth system science. *Nature*, *566*(7743), 195–204. https://doi.org/10.1038/s41586-019-0912-1

Repo, A., Rajala, T., Henttonen, H. M., Lehtonen, A., Peltoniemi, M., & Heikkinen, J. (2021). Age-dependence of stand biomass in managed boreal forests based on the Finnish National Forest Inventory data. *Forest Ecology and Management*, *498*, 119507. https://doi.org/10.1016/j.foreco.2021.119507

Richards, F. J. (1959). A flexible growth function for empirical use. *Journal of Experimental Botany*, *10*(2), 290–301. https://doi.org/10.1093/jxb/10.2.290

Santaren, D., Peylin, P., Bacour, C., Ciais, P., & Longdoz, B. (2014). Ecosystem model optimization using in situ flux observations: Benefit of Monte Carlo versus variational schemes and analyses of the year-to-year model performances. *Biogeosciences*, *11*(24), 7137–7158. https://doi.org/10.5194/bg-11-7137-2014

Seiler, C., Melton, J. R., Arora, V. K., Sitch, S., Friedlingstein, P., Anthoni, P., et al. (2022). Are terrestrial biosphere models fit for simulating the global land carbon sink? *Journal of Advances in Modeling Earth Systems*, *14*(5), e2021MS002946. https://doi.org/10.1029/2021MS002946

Shang, R., Chen, J. M., Xu, M. Z., Lin, X. D., Li, P., Yu, G., et al. (2023). China's current forest age structure will lead to weakened carbon sinks in the near future. *The Innovation*, *4*(6), 100515. https://doi.org/10.1016/j.xinn.2023.100515

Shangguan, W., Dai, Y. J., Duan, Q. Y., Liu, B. Y., & 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

Sierra, C. A., Ceballos-Núñez, V., Hartmann, H., Herrera-Ramírez, D., & Metzler, H. (2022). Ideas and perspectives: Allocation of carbon from net primary production in models is inconsistent with observations of the age of respired carbon. *Biogeosciences*, *19*(16), 3727–3738. https://doi.org/10.5194/bg-19-3727-2022

Smallman, T. L., Exbrayat, J. F., Mencuccini, M., Bloom, A. A., & Williams, M. (2017). Assimilation of repeated woody biomass observations constrains decadal ecosystem carbon cycle uncertainty in aggrading forests. *Journal of Geophysical Research: Biogeosciences*, *122*(3), 528–545. https://doi.org/10.1002/2016JG003520

Sobol, I. M. (1993). Sensitivity estimates for nonlinear mathematical models. *Mathematical Modelling in Civil Engineering*, *4*, 407–414.

Tao, F., Zhou, Z. H., Huang, Y. Y., Li, Q. Y., Lu, X. J., Ma, S., et al. (2020). Deep learning optimizes data-driven representation of soil organic carbon in earth system model over the conterminous United States. *Frontiers in Big Data*, *3*, 1–15. https://doi.org/10.3389/fdata.2020.00017

Thoning, K. W., Tans, P. P., & Komhyr, W. D. (1989). Atmospheric carbon dioxide at Mauna Loa observatory: 2. *Analysis of the NOAA GMCC data, 1974–1985*, *94*(D6), 8549–8565. https://doi.org/10.1029/JD094iD06p08549

Thum, T., MacBean, N., Peylin, P., Bacour, C., Santaren, D., Longdoz, B., et al. (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

Varejão, C. G., Costa, M. H., & Camargos, C. C. S. (2013). A multi-objective hierarchical calibration procedure for land surface/ecosystem models. *Inverse Problems in Science and Engineering*, *21*(3), 357–386. https://doi.org/10.1080/17415977.2011.639453

Williams, C. A., Collatz, G. J., Masek, J., Huang, C. Q., & Goward, S. N. (2014). Impacts of disturbance history on forest carbon stocks and fluxes: Merging satellite disturbance mapping with forest inventory data in a carbon cycle model framework. *Remote Sensing of Environment*, *15*, 57–71. https://doi.org/10.1016/j.rse.2013.10.034

Winter, J. M., Pal, J. S., & Eltahir, E. A. B. (2009). Coupling of integrated biosphere simulator to regional climate model version 3. *Journal of Climate*, *22*(10), 2743–2757. https://doi.org/10.1175/2008JCLI2541.1

Wutzler, T., & Carvalhais, N. (2014). Balancing multiple constraints in model-data integration: Weights and the parameter block approach. *Journal of Geophysical Research: Biogeosciences*, *119*(11), 2112–2129. https://doi.org/10.1002/2014JG002650

Wutzler, T., & Reichstein, M. (2007). Soils apart from equilibrium? Consequences for soil carbon balance modelling. *Biogeosciences*, *4*(1), 125–136. https://doi.org/10.5194/bg-4-125-2007

Xia, J. Z., Yuan, W. P., 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

Xia, J. Z., Yuan, W. P., Wang, Y. P., & Zhang, Q. G. (2017). Adaptive carbon allocation by plants enhances the terrestrial carbon sink. *Scientific Reports*, *7*(1), 3341. https://doi.org/10.1038/s41598-017-03574-3

Xiao, J. F., Chevallier, F., Gomez, C., Guanter, L., Hicke, J. A., Huete, A. R., et al. (2019). Remote sensing of the terrestrial carbon cycle: A review of advances over 50 years. *Remote Sensing of Environment*, *233*, 111383. https://doi.org/10.1016/j.rse.2019.111383

Xiao, J. F., Davis, K. J., Urban, N. M., & Keller, K. (2014). Uncertainty in model parameters and regional carbon fluxes: A model-data fusion approach. *Agricultural and Forest Meteorology*, *189–190*, 175–186. https://doi.org/10.1016/j.agrformet.2014.01.022

Xu, C. Y., Turnbull, M. H., Tissue, D. T., Lewis, J. D., Carson, R., Schuster, W. S. F., et al. (2012). Age-related decline of stand biomass accumulation is primarily due to mortality and not to reduction in NPP associated with individual tree physiology, tree growth or stand structure in a Quercus-dominated forest. *Journal of Ecology*, *100*(2), 428–440. https://doi.org/10.1111/j.1365-2745.2011.01933.x

Xu, H. Y., Zhang, T., Luo, Y. Q., Huang, X., & Xue, W. (2018). Parameter calibration in global soil carbon models using surrogate-based optimization. *Geoscientific Model Development*, *11*(7), 3027–3044. https://doi.org/10.5194/gmd-11-3027-2018

Xu, L., Saatchi, S. S., Yang, Y., Yu, Y. F., Pongratz, J., Bloom, A. A., et al. (2021). Changes in global terrestrial live biomass over the 21st century. *Science Advances*, *7*(27), eabe9829. https://doi.org/10.1126/sciadv.abe9829

Yu, G. R., Chen, Z., Piao, S. L., Peng, C. H., Ciais, P., Wang, Q. F., et al. (2014). High carbon dioxide uptake by subtropical forest ecosystems in the East Asian monsoon region. *Proceedings of the National Academy of Sciences*, *111*(13), 4910–4915. https://doi.org/10.1073/pnas.1317065111

Yuan, W. P., Liang, S. L., Liu, S. G., Weng, E. S., Luo, Y. Q., Hollinger, D., & Zhang, H. (2012). Improving model parameter estimation using coupling relationships between vegetation production and ecosystem respiration. *Ecological Modelling*, *240*, 29–40. https://doi.org/10.1016/j.ecolmodel.2012.04.027

Zaehle, S., Sitch, S., Prentice, I. C., Liski, J., Cramer, W., Erhard, M., et al. (2006). The importance of age-related decline in forest NPP for modelling regional carbon balances. *Ecological Applications*, *16*(4), 1555–1574. https://doi.org/10.1890/1051-0761(2006)016[1555:tioadi]2.0.co;2

Zheng, Y., Shen, R., Wang, Y., Li, X., Liu, S., Liang, S., et al. (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

Zhu, K., Song, Y. L., & Qin, C. (2019). Forest age improves understanding of the global carbon sink. *Proceedings of the National Academy of Sciences*, *116*(10), 3962–3964. https://doi.org/10.1073/pnas.1900797116