Biogeochemical model Biome-BGCMuSo v6.2 provides plausible and accurate simulations of carbon cycle in Central European beech forests

Katarína Merganičová[1,2], Ján Merganič[3], Laura Dobor[1], Roland Hollós[4,5], Zoltán Barcza[1,4], Dóra Hidy[9], Zuzana Sitková[6], Pavel Pavlenda[6], Hrvoje Marjanovic[8], Daniel Kurjak[3], Michal Bošeľa[3], Doroteja Bitunjac[8], Masa Zorana Ostrogovic Sever[8], Jiří Novák[7], Peter Fleischer[3], Tomáš Hlásny[1]

[1]Czech University of Life Sciences Prague, Faculty of Forestry and Wood Sciences, Kamýcká 129, 16500 Praha 6 – Suchdol, Czech Republic

[2]Department of Biodiversity of Ecosystems and Landscape, Institute of Landscape Ecology, Slovak Academy of Sciences, 949 01 Nitra, Slovak Republic

[3]Technical University in Zvolen, Faculty of Forestry, T. G. Masaryka 24, 960 53 01 Zvolen, Slovak Republic

[4]ELTE Eötvös Loránd University, Department of Meteorology, Pázmány P. sétány 1/A, H-1117 Budapest, Hungary

[5]HUN-REN Centre for Agricultural Research, Agricultural Institute,Brunszvik u. 2, H-2462 Martonvásár, Hungary

[6]National Forest Centre -, Forest Research Institute Zvolen, T. G. Masaryka 2175/22, SK – 960 92 01 Zvolen, Slovak Republic

[7]Forestry and Game Management Research Institute, Research Station at Opočno, Na Olive 550, 517 73 Opočno, Czech Republic

[8]Croatian Forest Research Institute, Division for Forest Management and Forestry Economics, Cvjetno naselje 41, 10450 Jastrebarsko, Croatia

[9]HUN-REN-MATE Agroecology Research Group, 2100 Gödöllő, Páter K. u. 1.

Correspondence to: Katarína Merganičová (k.merganicova@forim.sk)

Abstract. Process-based ecosystem models are increasingly important for predicting forest dynamics under future environmental conditions, which may encompass non-analogous climate coupled with unprecedented disturbance regimes. However, challenges persist due to the extensive number of model parameters, scarce calibration data, and the trade-offs between the local precision and the ~~model wide-ranging~~ applicability of the model over a wide range of environmental conditions. In this paper, we describe a protocol that allows a modeller to collect transferable ecosystem properties based on ecosystem characteristic criteria and to compile the parameters that need to be described in the field.

We applied the procedure to develop a new parameterization for the European beech (*Fagus sylvatica* L.) for the Biome-BGCMuSo model, the most advanced member of the Biome-BGC family. For model calibration and testing, we utilized multiyear forest carbon data from 87 plots distributed across five European countries. The initial values of 48 new ecophysiological parameters were defined based on the literature review. The final values of 6 calibrated parameters were optimised for single sites as well as for multiple sites using the Generalised Likelihood Uncertainty Estimation and model output conditioning that ensured plausible simulations based on user-defined ranges of carbon stock output variables (carbon stock in aboveground wood biomass, soil, and litter) and finding the intersections of site-specific plausible parameter hyperspaces. To support the model use, we tested the model performance in simulating tree aboveground wood, soil, and litter carbon across a large geographical gradient of Central Europe and evaluated the trade-offs between parameters tailored to single plots and parameters estimated using multiple sites.

Our findings indicated that parameter sets derived from single sites provided an improved local accuracy of simulations of aboveground wood, soil, and litter carbons stocks by 35 %, 55 %, and 11 % in comparison to the a priori parameter set. However, their broader applicability was very limited. A multi-site optimised parameter set, on the other hand, performed satisfactorily across the entire geographical domain studied here, including sites not involved in the parameter estimation, but the errors were on average by 26 %, 35 % and 9 % greater for the aboveground wood, soil, and litter carbons stocks than those obtained with the site-specific parameter sets. Importantly, model simulations demonstrated plausible responses across large-scale environmental gradients, featuring a clear production optimum of beech that aligns with empirical studies. These findings suggest that the model is capable of accurately simulating the dynamics of the European beech across its range and can be used for more comprehensive experimentations.

1   Introduction

~~Forests are dynamic long-lived ecosystems, with the life span of single trees potentially exceeding several centuries (Di Filippo et al., 2015). Forests longevity and ecological complexity make studies of their dynamics challenging, mainly under unstable environmental conditions, changing ecological interactions and management (Mouquet et al., 2015), due to limited options of~~

55 ~~research repeatability in empirical studies (Schnitzer and Carson, 2016). Computer models have become indispensable in overcoming these problems since they account for complex interactions, feedback loops, and dependencies between ecosystem components (Bugmann and Seidl, 2022; Geary et al., 2020). The development of models was considerably accelerated by advances in computer science and the progress in ecophysiological and ecological research, which helped define different model structures and parameters (Minunno et al., 2019; Augustynczik et al., 2017). Hence, currently a great number of different~~

60 ~~models using varying methodological approaches, different temporal and spatial scales, and focusing on answering different primary questions is available (for an overview see e.g. Fabrika et al., 2019; Bosela et al., 2022).~~

~~Over time, the objectives of vegetation dynamics modelling have shifted from supporting practical decision-making towards providing novel insights into ecological patterns and processes, ecosystem transformations driven by climate change, and complex scenario and effect attribution studies (Geary et al., 2020). An urgency of informing ecosystem management and~~

65 ~~policies about the impacts of climate change and effects of management actions has become a prominent use of ecosystem models (Pecl et al., 2017). Rapid environmental changes have increasingly favoured process-based ecosystem models (PBM) over the empirical ones (Adams et al., 2013), because they incorporate a mechanistic description of physical and biological processes and interactions between plants and the environment (Waring and Running, 1998). Hence, they should possess the capacity to better address novel and non-analogous conditions, which could occur in the future but were not observed in the~~

70 ~~past (Adams et al., 2013).~~

Complex process-based vegetation dynamics models (PBMs) ~~simulating the dynamics of vegetation~~ typically contain many parameters specifying physiology, biochemistry, phenology, and allocation patterns of different vegetation types or species (Cameron et al., 2013; van Oijen, 2017). Parameter values are estimated based on different field or laboratory measurements, trial-and-error parameter adjustments or probabilistic methods (Forrester et al., 2021). A comprehensive review of calibration

75 methods can be found in Hollós et al. (2022). Thereby, each measurable parameter has its ~~intrinsic~~ own variability that emerges from environmental conditions, sampling and measurement errors. Such a value range can be interpreted as a parameter probability distribution or parameter uncertainty (van Oijen, 2017). Calibration is often applied to narrow the initial parameter ranges and capture regional or local peculiarities.

The challenges of model calibration include Aa selection of most influential variables to be calibrated using a sensitivity

80 analysis (SA) and coping with~~coping wit~~ ~~calibration~~ equifinality, i.e. a situation when various combinations of parameter values produce the same results (Beven, 2006)~~, can challenge calibration efforts~~. To this end, different calibration approaches, such as trial-and-error parameter adjustments or probabilistic methods (Forrester et al., 2021; Hollós et al., 2022) including Bayesian methods (Fer et al., 2018; van Oijen, 2017) or the Generalized Likelihood Uncertainty Estimation (GLUE) (Beven

and Binley, 2014) have been proposed. The calibration can focus on one or several variables simultaneously. Multivariate approaches are ~~generally~~ preferred as they aim to identify parameter values that minimise the differences between simulated and observed values of multiple variables (Kamali et al., 2022; Wöhling et al., 2013). The spatial aspect may be dealt with in a similar manner. The model calibrated for single sites provides ~~model~~ outputs with a high local accuracy, ~~but~~ whereas it loses the ability to generalize outside the calibration data (Blyth et al., 2011; Kramer et al., 2002; Levins, 1966). Therefore, introducing advanced calibration designs that are supposed to maintain~~ing~~ a balance between the local accuracy and ~~generality~~wide applicability, ~~and~~ provid~~eing~~ good performance across multiple variables, is needed~~, have been developed (Keenan et al., 2011; Wallach et al., 2021)~~.

In this study, we ~~focused on~~aimed to ~~creating~~formulate a calibration workflow that offers improvements beyond the broadly used methods detailed in ~~(~~Keenan et al.~~,~~ (2011), and~~;~~ Wallach et al.~~,~~ (2021). ~~For this prursposе w~~We used~~addressed~~ the process-based model Biome-BGCMuSo (BBGCMuSo, Hidy et al., 2016, 2022, 2012), which is the most rapidly developing member of the Biome-BGC model family (Thornton, 1998). It simulates the storage~~s~~ and fluxes of carbon, nitrogen and water~~these elements~~ within and between the pools. The model has been extensively used in the research of forest ecosystems concerning their productivity (Kimball et al., 1997; Sever et al., 2017), carbon (Churkina et al., 2003; Ostrogović Sever et al., 2021; Yan et al., 2016), water (Pietsch et al., 2003), and nitrogen dynamics (Merganičová et al., 2005; Pietsch et al., 2003), including effects of climate change (Churkina and Running, 2000; Hlásny et al., 2011; Jager et al., 2000; White et al., 1999). The recent developments of BBGCMuSo included a multilayer soil profile; complex water cycling between soil, vegetation, and atmosphere; intra-annual phenology; and complex management operations (Hidy et al., 2012, 2016, 2022). However, robust testing of ecological plausibility and model performance in forest ecosystems as well as ~~the~~ regionally calibrated species-specific parameter sets are still lacking. These tasks are challenging given the substantial increase in model structural complexity and the number of parameters, which limits the use of former parameter sets (Pietsch et al., 2005).

The aim of this study is to develop a multi-objective calibration procedure of model parameters that considers balancing the trade-off between the local precision of model outputs and a broad applicability of parameters, and to perform a comprehensive model benchmarking of the ecological plausibility of model results across large environmental gradients. We hypothesise that parameter estimates optimized for single sites are not sufficiently robust to be applied across large geographical space, while the multi-site optimisation reduces the ability of model ~~output~~ to capture the phenotypic~~al~~ plasticity of vegetation that causes alterations in plant properties, e.g. allocation ratios between different plant organs in response to environmental conditions, thereby reducing the local accuracy (Gratani, 2014). The proposed method was applied to calibrate ecophysiological parameters of BBGCMuSo v6.2 for the European beech (*Fagus sylvatica* L.), the most widespread deciduous tree species in Europe ~~to facilitate a large scale utilization of the model in addressing research questions related to alterations in species productivity and carbon cycle under different climatic and management scenarios~~. The~~se~~ findings ~~should~~ may serve as a reference for calibrating other tree species and/or different models.

2. Data and Methods

2.1. Model

Biome-BGCMuso (Hidy et al., 2012, 2016, 2022) is a descendant of Biome-BGC model (Thornton, 1998). It is a biogeochemical model that simulates cycling of carbon, water, nitrogen, and energy in terrestrial ecosystems at a daily time step (Thornton, 1998). Biome-BGC was one of the earliest biogeochemical models that included explicit carbon, water and nutrient cycles. The represented processes include photosynthesis, evapotranspiration, allocation, respiration, litterfall, and decomposition (Thornton et al., 2002). These processes are defined for a unit ground area that is considered homogeneous. A so-called „two-leaf" ~~representation~~model that represents stand foliage with one sunlit and one shaded leaf is used to simulate radiation interception, evapotranspiration, and gross primary production for the sunlit and shaded canopy fractions (Thornton and Rosenbloom 2005). The modelled ecosystem consists of several components representing different plant parts (leaf, stem, roots), litter, soil, and coarse woody debris. The ecosystem status and dynamics are represented by carbon (C), nitrogen (N) and water (W) pools and fluxes between the pools. The main pools represent leaf (C, W, N), aboveground wood (C, N), coarse root (C, N), fine root (C, N), coarse woody debris (C, N), litter (C, N), soil (C, W, N), yield (C, N), standing-dead-biomass (C, N) cut-down biomass (C, N). BBGCMuSo contains a number of new features including a multilayer soil representation that allows simulating ~~a~~ more realistic dynamics of water, carbon and nitrogen across the soil profile; a possibility of using ~~a~~ dynamic annual mortality rates; adjustable intra-annual allocation driven by phenology; improved representation of transpiration, soil evaporation, and inorganic nitrogen; and flexible simulation of management operations, including forest thinning and harvesting (Hidy et al., 2012, 2016, 2022). It can also simulate acclimation to temperature and short-term temperature dependence of maintenance respiration, drought legacy effects through reduction of non-structural carbohydrate storage pools, and a CO_2 concentration-dependent stomatal conductance (Hidy et al., 2016). The model version 6.2 uses 53 soil-related and 105 ecophysiological parameters (Hidy et al., 2021), of which some are site-specific (e.g. soil depth, soil texture), while others, such as C:N ratio in different tree compartments, are species-specific. The number of parameters in BBGCMuSo has been tripled in comparison to the original model Biome-BGC, although~~mostly~~ 17 parameters ~~thanks to the fact that the model was also made suitable~~ were specified to allow~~for simulations of~~concern ~~the~~ crop ~~functioning~~s and are not relevant for forest ecosystems. To perform simulations, a species or a plant functional type needs to be defined. In addition, site, soil and daily climate data are required inputs. In the case of forest ecosystems, stand age and past forest management are also necessary input data. The model, including its source code, is available at https://nimbus.elte.hu/bbgc/.

2.2. Data

The dendrometric and environmental data represented 87 forest sites distributed across Central Europe within the distributional range of European beech (Figure 1). The dataset includes sites from the International Co-operative Programme on Assessment and Monitoring of Air Pollution Effects on Forests (ICP Forests, Michel et al., 2018) long-term forest research plots from thinning trials supervised by different national institutions, and highly-instrumented intensively

monitored plots equipped with weather stations, dendrometers, instruments measuring sap flow, soil water content, etc.
(Table S2). The plots are located in Croatia, Hungary, Slovakia, the Czech Republic, and Poland along an elevation gradient
150   from 20 to 1,325 m a.s.l. Their mean annual precipitation totals vary from 419 to 1,883 mm, mean annual temperatures range
from 3.5 to 13.3°C, and soil depths vary between 0.4 and 2 m (
Table **1**). Most plots were of a circular shape with a size from 0.09 to 1.05 ha. Forest stands were established between 1787 and
1984, i.e., their age in the year 2022 varied from 38 to 235 years. Both managed and unmanaged forest stands ~~that originated~~
originating either from natural or artificial regeneration or a combination ~~of both~~thereof, are represented. Time series lengths
155   and the number of observations differed between the sites depending on the year of plot establishment, and the frequency of
re-measurements of tree dimensions. The maximum time series length was 60 years, and the maximum number of observations
in a single ~~time~~ series was 30.

(a)   (b)

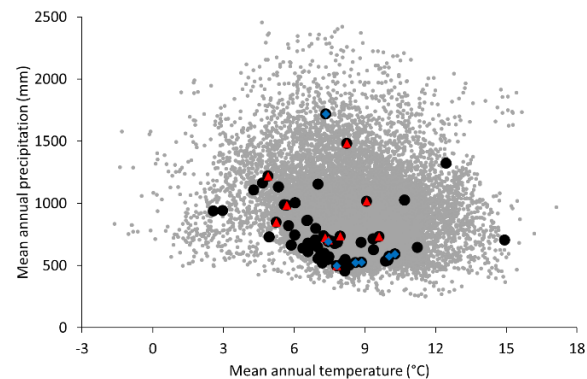
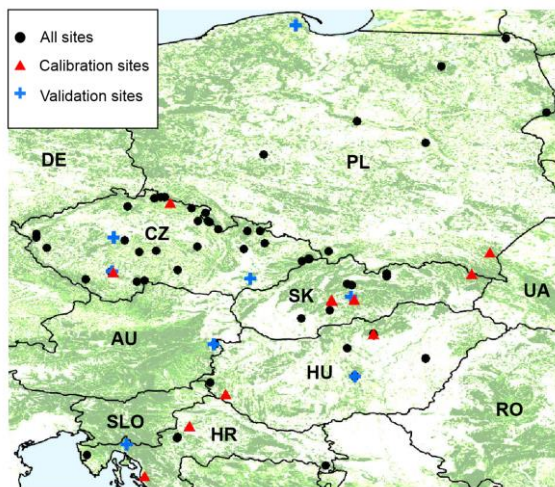



**Figure 1 a) ~~Distribution~~ Distribution of used forest sites across Europe with ~~the~~ forest cover displayed in the background (CORINE
Land Cover, 2023), ~~a~~b)~~, and in t~~The coverage of the entire climatic space of European beech in Europe (represented by grey dots
160   using the, data from Caudullo et al. 2017) by the used forest sites. ~~derived from Caudullo et al. (2017) and indicated by grey dots, b).~~
Black dots indicate 87 ~~research~~ sites used for testing model plausibility ~~(realism),~~. ~~red~~ Red triangles indicate eleven data-rich sites
used for the calibration of the BBGCMuSo model. ~~for simulating European beech forests, and b~~Blue crosses and diamonds represent
eight validation sites.**

Out of the whole data set, 11 beech-dominated sites with the most comprehensive data and the balanced coverage of the
165   geographical and environmental space~~coverage ensuring that each country was represented by at least one site,~~ were used for
model calibration (hereafter referred to as calibration sites~~, see~~ Figure 1,

Table **1**). The northern part of the selected region is underrepresented due to the insufficient data for model calibration at northern sites. Eight beech-dominated sites with repeated stand measurements covering the full range of environmental conditions represented in our database (

170   Table **1**) representing all included countries were selected to perform used for an independent model validation (hereafter referred to as validation sites). All 87 sites were used for testing the plausibility (realism) of simulated model output using the calibrated parameter set across an extended geographical gradient. The comparison of beech natural distribution ranges (Pagan, 1996) with the covered ranges of latitude, elevation, climatic, and soil characteristics suggested that the selected sites should be representative of the Central European beech population.

175

Table 1 Summary of site and forest stand characteristics for the whole data set (All), calibration and validation sets. Climate data represent a period from 1950 to 2018.

Data set	All	Calibration	Validation
Number of sites	87	11	8
Dominant tree species	*Fagus sylvatica, Picea abies, Pinus sylvestris, Quercus spp.*	*Fagus sylvatica*	*Fagus sylvatica*
Site variables	mean (min - max)		
Years with observations	(1990 - 2018)	(1949 - 2018)	(1990 - 2018)
Latitude (°)	(N 45.4814 - N 54.559)	(N 44.8164 - N 50.7349)	(N 45.4814 - N 54.5592)
Longitude (°)	(E 14.2736 - E 23.72)	(E 14.3000 - E 22.4917)	(E 14.2736 - E 19.4701)
Elevation (m a.s.l.)	544.2 (20 - 1325)	686.8 (240 - 1325)	502.5 (120 - 1180)
Annual precipitation (mm)	789 (419 – 1884)	1018 (663 – 1884)	784 (458 - 1592)
Mean temperature (°C)	7.75 (3.52 - 13.28)	7.34 (4.86 - 10.3)	8.37 (6.41 - 11.11)
Soil depth (m)	0.99 (0.4 - 2)	0.92 (0.45 - 2)	0.99 (0.5 - 1.8)
Plot size (ha)	0.31 (0.09 - 1.05)	0.45 (0.25 - 1.05)	0.28 (0.09 - 0.84)
Stand age (yr)	84 (35 - 232)	118 (64 - 232)	79 (41 - 122)
Aboveground wood carbon (kgC m^{-2})	8.62 (0.01 - 36.87)	16.92 (0.03 - 35.87)	16.89 (4.30 - 33.74)
Soil carbon (kgC m^{-2})	13.31 (1.13 - 16.24)	13.21 (1.13 - 16.24)	NA
Litter carbon (kgC m^{-2})	0.256 (0.086 - 0.344)	0.256 (0.086 - 0.344)	NA

The plot dataset contains information on site, forest stand structure and development, soil, climate, nitrogen deposition, and
180   physiological processes (Table S2). Site description data comprise plot coordinates (latitude, longitude), elevation, aspect, and slope. Forest stand data comprise information on tree species composition, stand age, stand structure, data on individual tree dimensions or mean stand characteristics (e.g. mean diameter at breast height (DBH), mean tree height (H)), mortality, the applied management and damages. Soil dataset contains soil depth, texture, pH, nutrient stocks, and indicators of water regime. If soil information was not available, it was obtained from the Harmonized World Soil Database (HWSD v 1.21, FAO, 2012)
185   that provides soil attributes, such as soil depth, soil texture, pH, etc., at a grid cell size of approximately 1km.

Climate data include daily values of minimum and maximum temperature, solar radiation, precipitation, and vapour pressure deficit (VPD). These data were compiled from different sources, including observations at individual sites, nearby meteorological stations, or if no local data were available, the E-OBS gridded dataset providing daily minimum, maximum temperature, and precipitation with 0.1 deg. resolution was used (Cornes et al., 2018). The climate data covered a period from 1950 to 2018. To cope with limited availability of local data, we combined the on-site measurements with the E-OBS data. The MTClim model (Hungerford et al., 1989) was used to extrapolate the climate time series from the nearest E-OBS grid cell to account for the elevation difference between the source cell and the target site, and to calculate daylight values of mean temperature, VPD, solar radiation, and daylength at individual sites required by BBGCMuSo.

Annual CO_2 data were taken from Mauna Loa observations (Keeling et al., 1976). Since nitrogen deposition data were directly available only for ICP Forests plots, they were taken from Tian et al. (2018) for the remaining sites.

Data about site, soil texture, pH, stand age, management, nitrogen deposition, CO_2 concentration and daily climate data were used as model inputs to run site-specific simulations. Model calibration and validation was based on carbon stocks in aboveground wood, litter and soil (AbgwC, SoilC, LitterC). AbgwC was derived from the dendrometric characteristics of individual trees by calculating first the total aboveground wood volume using species-specific two-parameter regressions (Petráš and Pajtík 1991). The produced values were subsequently converted to carbon stock using the species-specific basic wood density (Merganič et al. 2017, 570 kg m^{-3} for beech) and 50 % carbon content in biomass (IPCC, 2003).

2.3 Simulation design

The simulations were performed in three steps: (1) spin-up run, (2) transient run, and (3) normal run (Hidy et al., 2021). The spin-up was performed with a constant CO_2 concentration and nitrogen deposition equal to the pre-industrial values of 277.15 ppm and 0.002 kgN m^{-2} year^{-1}, respectively. During the transient run, CO_2 concentration and nitrogen deposition increased annually from their pre-industrial to the current values. The transient run started in 1850 and lasted until the year preceding the establishment of the current stand. Hence, the length of the transient run varied between sites, and it was used only for the simulations of the stands established after 1850, while the maximum length of the transient run was 134 years. Both spin-up and transient runs were performed with no management, no fire-induced mortality, and constant natural (background) mortality.

The normal run was driven by the temporally varying CO_2 and N-deposition, and included management reconstructed based on forest inventory records or yield tables, if no site-specific information on management interventions was available. The simulations started at stand age 0 (i.e., the year of stand establishment) and continued to the present day; i.e. the simulation length equalled the actual stand age. In simulation year 1, a clearcut was applied followed by the removal of 90 % of the aboveground woody biomass accumulated during the spin-up and transient runs, with all non-woody biomass, i.e. the foliage, remaining in the stand. Natural mortality was changing annually (Fig. S1). In the first 30 years after the stand

establishment, natural mortality rates followed a decreasing exponential function ~~with~~ reaching the highest annual mortality rate one year after the stand establishment and its subsequent gradual reduction over time, resembling the survival rates of forest regeneration from experimental studies focusing on beech (Barna et al., 2011; Hülsmann et al., 2018). After 30 years, we used a constant annual mortality rate of 0.9 % (Pajtík et al. 2018; Vanoni et al. 2019) ~~to simulate~~in managed ~~forests~~stands, while unmanaged stands~~sites~~ were simulated using the dynamic natural mortality rates that fluctuated between 0.76 and 4.1 % during a cycle of 300 years. ~~and followed~~ following the elliptical function (Merganičová and Merganič, 2014).

2.4. Parameter estimation

The estimation of model parameters (i.e. model calibration) consisted of several phases (Figure 2) described below.

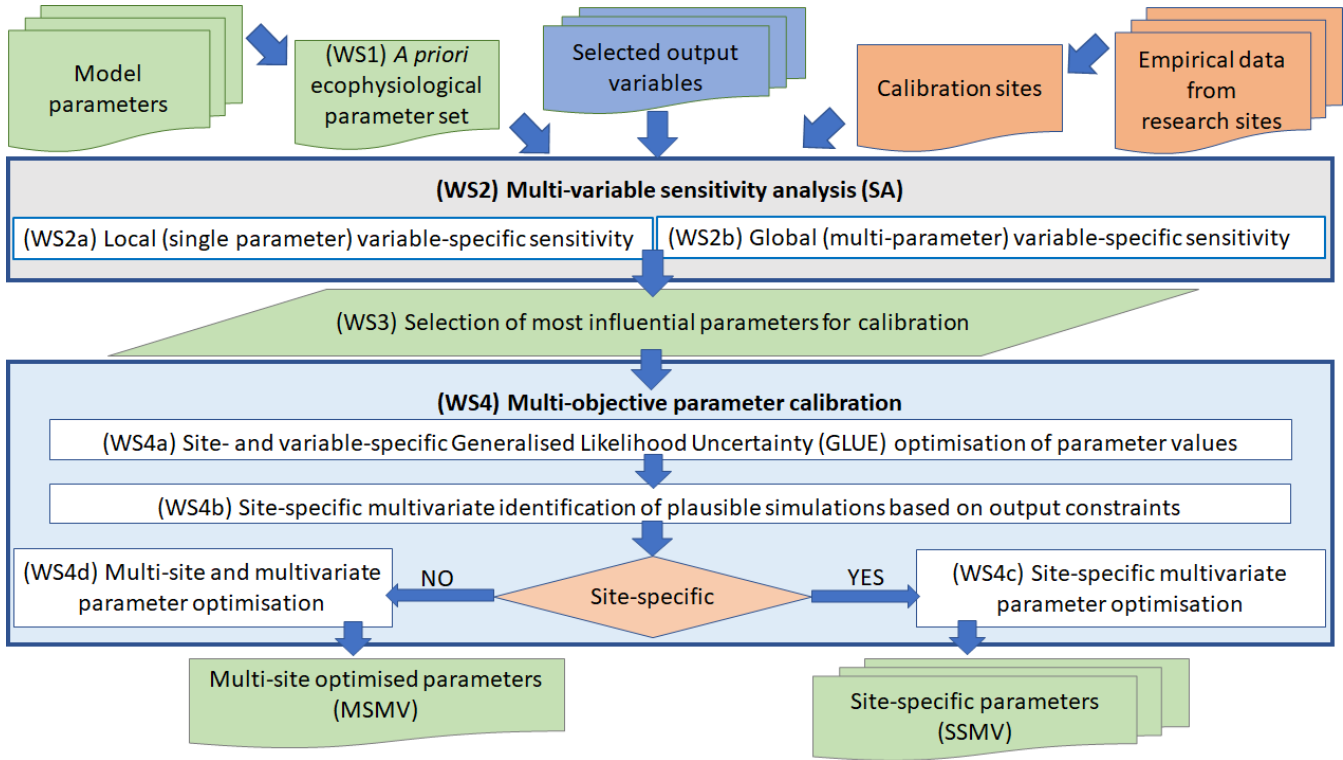
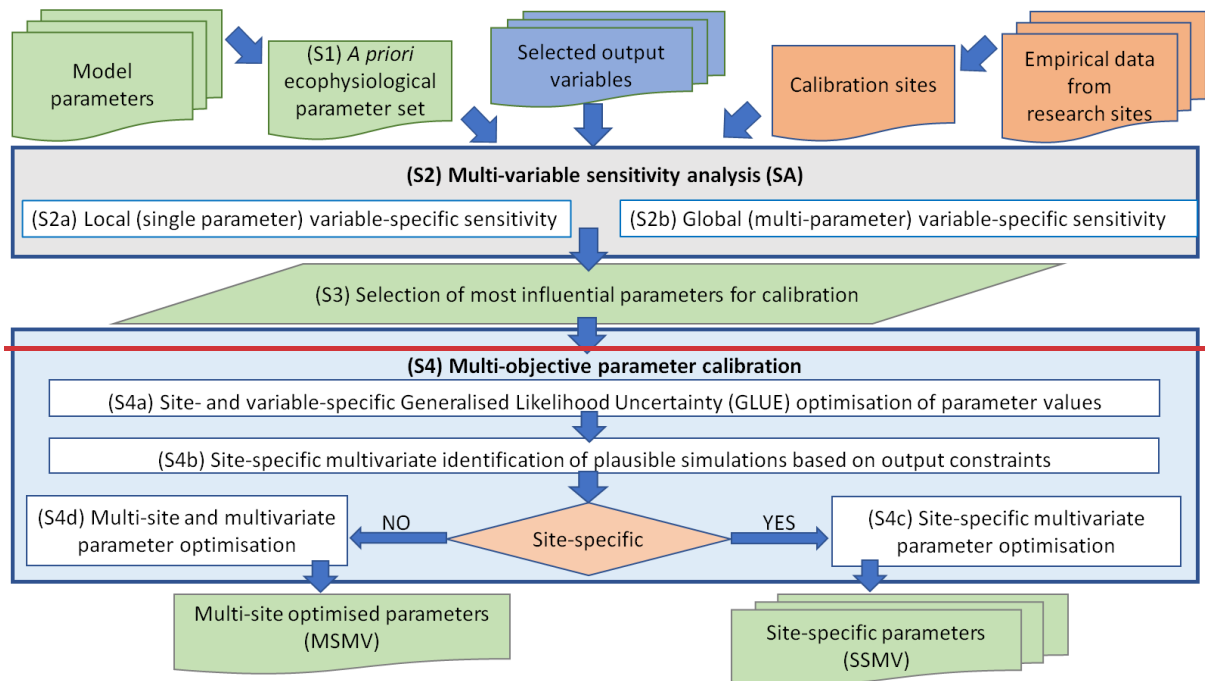



Figure 2 A general workflow used for ~~the~~ optimis~~ing values~~ation of ecophysiological parameters of BBGCMuSo for the European beech using the multi-objective calibration approach.

230

2.4.1. A priori parameter setting

Prior to the calibration of model parameters, the default or a priori parameter values were set (~~the W~~S1 phase in Figure 2). The initially defined set (~~called~~ column "First" in Table 2) was defined based on values of 40 ecophysiological parameters included in previous model versions taken from Pietsch et al. (2005)~~. and t~~The values of new 48 parameters, which were introduced in

235 the latest model version, were set based on the literature review and the TRY database (Kattge et al., 2020). This set was used for ~~the~~ first simulations of all 87 sites to examine the success~~fulness~~ of simulated development. ~~Under~~By a successful simulation we understand a simulation from spin-up until the end of ~~the~~ normal run, during which the ecosystem ~~existence~~ stability was maintained, i.e.~~ during the spin-up run,~~ carbon stock~~s~~ in vegetation and soil w~~ere~~as accumulated during the spin-up until they reached a balanced state, and ~~the development during the normal run simulation maintained the ecosystem~~ the

240 vegetation existence, confirmed by non-zero values of vegetation (AbgwC, foliage, roots) carbon pools, was maintained ~~in the model~~ throughout the entire normal run simulation. Based on the results of unsuccessful simulations~~ at specific sites~~, we identified ~~potential~~ parameters, the values of which may have caused~~s of~~ the problems, such as insufficient water or nitrogen supply ~~during simulations~~ resulting in the cessation of vegetation existence~~,~~. ~~and changed respective parameters using~~Subsequently, we used ~~the~~ a trial-and-error approach to alter the respective parameters until we ~~obtained~~ reached a 100

245 % success~~fulness~~ rate of the simulations. Those parameter values represented ~~and~~ an a priori parameter set (~~called~~ column "A priori" in Table 2).

Table 2 List of parameters ~~that were~~ tested and modified during the calibration of Biome-BGCMuSo for *Fagus sylvatica* L. Columns Min and Max represent the ~~= m~~Minimum and ~~Max =~~ maximum values ~~represent parameter ranges~~ used to define the parameter ranges for the sensitivity analysis and optimisation~~,~~. Local Sensitivity contains~~presents~~ the values of the sensitivity index per output variable (carbon stock in aboveground wood, soil, and litter labelled as AbgwC, SoilC, and LitterC) calculated following Hoffman and Gardner (1983, Eq. 1)~~,~~. ~~the~~ The column ~~named~~ "First" is the ~~very~~ first ~~setting~~ parameter approximation~~for beech that we used~~ identified based on the species-specific values derived for Biome-BGC by Pietsch et al. (2005) (in italics) and the complementary literature review~~,~~. The column ~~while~~ "A priori" ~~is the setting that~~contains parameter values that enabled~~ to~~ the successful simulations of all 87 sites ~~in the database before parameter optimisation~~ (result of the WS1 phase in Figure 2)~~,~~. The column "MSMV ~~optimised~~" ~~is~~ contains the multi-site and multivariate calibrated parameter set (result of the WS4d phase in Figure 2) based on 11 calibration sites and 3 output variables (AbgwC, SoilC, and LitterC). The ~~Calibrated~~ calibrated parameters are indicated with the grey background.

Parameter			Tested range		Local sensitivity			Parameter set		
Name	Abbreviation	Unit	Min	Max	AbgwC	SoilC	LitterC	First	A priori	MSMV
Transfer growth period as fraction of growing season	GP	dim	0.05	0.3	0.005	0.073	0.003	0.2	0.2	0.2
Litterfall as fraction of growing season	LP	dim	0.2	0.6	0.002	0.027	0.001	0.2	0.2	0.2
Base temperature	T_base	°C	0	7	0	0	0	5	5	5
Annual live wood turnover fraction	WTF	dim	0.5	1	0.00002	0.00026	0.00001	0.7	0.7	0.7
Annual fire mortality fraction	FM	dim	0	1	0.078	0.716	0.008	0	0	0
Whole-plant mortality fraction in vegetation period	WPM	dim	0	0.1	0.996	0.493	0.006	0.005	0.005	0.005
C:N of leaves	CN_lv	kgC kgN^{-1}	16.5	40	0.017	0.183	0.004	26.9	26.9	26.9
C:N of leaf litter, after retranslocation	CN_li	kgC kgN^{-1}	10	114	0.0002	0.015	0.178	44	44	44
C:N of fine roots	CN_ro	kgC kgN^{-1}	10	75.8	0.005	0.054	0.007	47.6	47.6	47.6
C:N of live wood	CN_lw	kgC kgN^{-1}	17	100	0.004	0.041	0.001	50	50	50
C:N of dead wood	CN_dw	kgC kgN^{-1}	300	819	0.00002	0.0002	0.005	550	550	550
Leaf litter labile proportion	LLaP	dim	0.1	0.6	0.00002	0.111	0.004	0.124	0.124	0.124
Leaf litter cellulose proportion	LCeP	dim	0.1	0.7	0.00002	0.157	0.004	0.561	0.561	0.561
Fine root labile proportion	RLaP	dim	0.1	0.6	0.00002	0.067	0.003	0.34	0.34	0.34
Fine root cellulose proportion	RCeP	dim	0.1	0.6	0.00003	0.080	0.002	0.44	0.44	0.44
Dead wood cellulose proportion	WCeP	dim	0.5	0.9	0.00004	0.361	0.376	0.77	0.77	0.77
Canopy water interception coefficient	CWIC	mm LAI^{-1} d^{-1}	0.01	0.063	0.003	0.017	0.011	0.034	0.034	0.034
Canopy light extinction coefficient	CLEC	dim	0.3	0.7	0.012	0.136	0.004	0.6	0.6	0.6616
All-sided to projected leaf area ratio	SLA:PA	dim	1.5	2.5	0.0007	0.001	0.00002	2	2	2
Ratio of shaded SLA:sunlit SLA	shSLA:suSLA	dim	0.2	5	0.073	0.568	0.004	2	2	2
Fraction of leaf N in Rubisco	FLNR	dim	0.1	0.3	0.008	0.087	0.003	0.162	0.162	0.1383
Maximum stomatal conductance (projected area basis)	MSC	m s^{-1}	0.001	0.009	0.088	0.925	0.020	0.006	0.005	0.0051
Cuticular conductance (projected area basis)	CC	m s^{-1}	0.00001	0.0001	0.001	0.011	0.000	0.00006	0.00006	0.00006
Boundary layer conductance (projected area basis)	BLC	m s^{-1}	0.01	0.09	0.005	0.076	0.002	0.01	0.01	0.01
Maximum depth of rooting zone	MRD	m	0.2	4.1	0.001	0.021	0.003	2	2	2
Root distribution parameter	rootDistr	dim	0.5	4	0.001	0.088	0.123	3.67	1.5	1.5
Growth resp per unit of C grown	GRC	dim	0.1	0.5	0.020	0.216	0.016	0.3	0.3	0.3
Maintenance respiration in kgC/day per kg of tissue N	MRperN	kgC kgN^{-1} d^{-1}	0.1	0.4	0.091	0.902	0.023	0.218	0.218	0.218
Theoretical maximum proportion of non-structural and structural carbohydrates	NSC:SCmax	dim	0.05	0.3	0.0001	0.0004	0.00001	0.1	0.1	0.1
Proportion of non-structural carbohydrates available for maintanance respiration	NSC2MR	dim	0.1	0.5	0.00002	0.00011	0.000008	0.3	0.3	0.3
Symbiotic+asymbiotic fixation of N	Nfix	kgN m^{-2} yr^{-1}	0.0001	0.01	0.046	0.494	0.991	0.0005	0.01	0.0091
Time delay for temperature in photosynthesis acclimation	tau	day	0	50	0.0003	0.0035	0.0001	10	10	10
Volumetric water content ratio to calculate soil moisture limit 1	VWCratio_lim1	dim	0.1	0.9	0.004	0.067	0.002	0.99	0.1	0.1
Volumetric water content ratio to calculate soil moisture limit 2	VWCratio_limit2	dim	0.5	1	0.0001	0.0049	0.0021	0.99	0.99	0.99
Minimum of soil moisture limit2 multiplicator (full anoxic stress value)	min_soilstress2	dim	0	1	0	0	0	0.4	0.4	0.4
Vapor pressure deficit: start of conductance reduction	VPDS	Pa	500	1500	0.005	0.079	0.003	600	600	600
Vapor pressure deficit: complete conductance reduction	VPDC	Pa	1500	3500	0.080	0.783	0.009	3000	3000	2910
Maximum senescence mortality coefficient of aboveground plant material	SMCA	dim	0	0.01	0	0	0	0.001	0	0
Maximum senescence mortality coefficient of belowground plant material	SMCB	dim	0	0.01	0.0001	0.001	0.00002	0.001	0	0
Maximum senescence mortality coefficient of non-structured plant material	SMCL	dim	0	0.01	0.003	0.079	0.004	0.0001	0	0
Lower limit extreme high temperature effect on senescence mortality	SNSC_ext1	°C	30	40	0	0	0	30	30	30
Upper limit extreme high temperature effect on senescence mortality	SNSC_ext2	°C	30	50	0	0	0	40	40	40
Turnover rate of wilted standing biomass to litter	TRWB	dim	0.01	0.1	0	0	0	0.01	0.01	0.01
Turnover rate of non-woody cut-down biomass to litter	TRCN	dim	0.01	0.1	0	0	0	0.05	0.01	0.01
Turnover rate of woody cut-down biomass to litter	TRCW	dim	0.0001	0.1	0	0	0	0.01	0.0009	0.0009
Drought tolerance parameter	DSWScirt	nday	0	100	0	0	0	30	100	100
Effect of soilstress factor on photosynthesis	Sseff	dim	0	0.4	0.006	0.066	0.002	0	0	0
Leaf carbon allocation proportion		dim	-	-	-	-	-	0.173	0.173	0.173
Fine root carbon allocation proportion		dim	-	-	-	-	-	0.094	0.094	0.094
Live woody stem carbon allocation proportion		dim	-	-	-	-	-	0.101	0.101	0.101
Dead woody stem carbon allocation proportion		dim	-	-	-	-	-	0.556	0.556	0.556
Live coarse root carbon allocation proportion		dim	-	-	-	-	-	0.012	0.012	0.012
Dead coarse root carbon allocation proportion		dim	-	-	-	-	-	0.064	0.064	0.064
Canopy average specific leaf area (projected area basis)		m2 kgC^{-1}	-	-	-	-	-	48	34.5	34.5
Canopy growth proportion		dim	-	-	-	-	-	0.5	0.5	0.5

2.4.2. Sensitivity analysis

260 The sensitivity analysis (SA, the WS2 phase in Figure 2) was performed to identify the ~~impact~~ effect of model parameters on ~~the~~ simulated carbon stocks in ~~the~~ aboveground wood, soil, and litter (AbgwC, SoilC, LitterC) ~~using BBGCMuSo simulations of~~ at the calibration sites (

Table **1**). The three carbon stock variables were selected instead of typically used carbon fluxes to cover a wider range of environmental conditions, since fluxes are usually measured only at a limited number of ~~research~~ sites. Moreover, ~~It has been~~
265 ~~documented that~~ BBGCMuso ~~can be well fitted~~has already been found to effectively simulate C fluxes (Hidy et al., 2016; Maselli et al., 2009), while its ability to simulate C stocks is much less documented (Ostrogović Sever et al., 2021). Moreover, focusing on C stocks rather than fluxes enables us to evaluate model performance over the multidecadal time span of up to 69 years represented by our data.

First, we performed a local, i.e. single parameter, SA (WS2a in Figure 2) using regular sampling of parameter values from
270 their pre-defined ranges based on the literature review. The sensitivity of variable i to parameter P was quantified ~~with~~using the sensitivity index (SI) ~~using the equation of~~proposed by Hoffman and Gardner (1983):

$$SI_{Pi} = \frac{Vmax_{Pi} - Vmin_{Pi}}{Vmax_i} \qquad (1)$$

where $Vmax_{Pi}$ and $Vmin_{Pi}$ are the maximum and minimum values of the simulated output variable i when testing the effect of parameter P, and $Vmax_i$ is the absolute maximum value of the output variable obtained from the tests of all parameters.

275 Afterwards, a global, i.e. multi-parameter, SA (WS2b in Figure 2), ~~which~~ assesses the sensitivity of all selected parameters across the entire parameter space simultaneously, was performed ~~following~~ using the least squares linearisation (LSL) approach ~~by~~ (Verbeeck et al., 2006). A commonly used global SA method is the variance-based Sobol analysis, which performs Monte-Carlo simulations on the parameter-space. The method estimates the Sobol sensitivity index ~~,~~that distribut~~ing~~es the overall variability of model outputs to the contributions from each model input (Saltelli et al., 2004). As the parameter space
280 expands, an increasing number of Monte-Carlo simulations is required to accurately estimate the sensitivity. To simplify this process and enhance the accuracy with a fewer number of simulations, surrogate models are employed. The simplest surrogate models are multivariable linear models (Verbeeck et al. 2006), including LSL. This approach utilises the ~~widely applied~~ ordinary least squares method to approximate the process-based model with a surrogate multidimensional linear model. The coefficients derived from the fitted model are then used to calculate the relative Sobol sensitivity indices.

285 The procedure first ~~The~~ simultaneously sampl~~ing of~~es values of all~~the~~ selected model parameters from their predefined ranges ~~was done~~ with Monte Carlo simulations, while assuming a multivariate uniform distribution. The~~n, the~~ simulated output~~s was~~ are examined with regard to parameter deviations from the mean using the LSL~~least square linearisation method~~, which splits the overall output uncertainty into its individual sources. This allowed us to estimate the contribution of each tested parameter to the ~~analysed~~ model output uncertainty and identify the parameters that affected AbgwC, SoilC, and LitterC most. The results

290   of SA were used to select the parameters to be calibrated (WS3 phase in Figure 2). The SA was performed using the musoMonte and musoSensi functions implemented in the RBBGCMuSo package (Hollós et al., 2023) available at https://github.com/hollorol/RBBGCMuso.

2.4.3. Parameter calibration

295   Within the parameter optimisation we fFirst, we performed a site- and variable-specific calibration (WS4a phase in Figure 2) using the GLUE method (Beven and Binley, 2014) implemented in the calibMuso function of the RBBGCMuSo package (Hollós et al., 2023). With this procedure, a selected parameter set was optimised using the least squares likelihood function based on the comparison of the simulated values of the selected output variable with its the corresponding observations in the pre-defined parameter space:

$$L = e^{-\sqrt{\frac{(Vobs-Vsim)^2}{n}}} \tag{2}$$

300   where L is the estimated likelihood, *Vobs* and *Vsim* are the observed and the simulated values of the output variable, and n is the number of observations.

For every site, we performed 100,000 Monte Carlo simulations, each with a unique combination of parameter values randomly generated sampled from the predefined parameter ranges (Table 2). The results of all simulations were used in the subsequent
305 optimisation steps.

FirstThen, we evaluated the plausibility (i.e. the realism) of simulated AbgwC, LitterC, and SoilC (WS4b phase in Figure 2) at the end of the spin-up and in individual years of the normal run simulations. The plausibility check is a form of output data streams conditioning similar to the CIRM method proposed by Hollós et al. (2022). We applied the following constraints derived from the literature to evaluate the plausibility of simulated values: carbon stock in aboveground wood below 70 kgC
310 m^{-2} (Barna et al., 2011; Georgi et al., 2018; Standovár and Kenderes, 2003; Trotsiuk et al., 2012), soil carbon in the whole soil profile between 5 and 25 kgC m^{-2}, and litter carbon amount between 0.1 to and 4 kgC m^{-2} (De Vos and Cools, 2011; Pavlenda and Pajtík, 2010; Wellbrock et al., 2016; Wellbrock and Bolte, 2019). A simulation was identified as plausible if all three examined output variables fulfilled the above listed constraints at the end of the spin up and in all simulated years of the normal runwere within the given ranges.
315 The site-specific multivariate optimised parameter values (SSMV parameter sets resulting from the WS4c phase in Figure 2) were derived from the subsets of plausible simulations selected for each calibration site as those minimising the estimation errors of AbgwC, SoilC, and LitterC, and maximizing the joint-likelihood function (WS4b in Figure 2). We appliedused the normal likelihood function to each of the three output variables and afterwards calculated the sum of loglikelihood values for each timestamp (year,) assuming the that independence of the estimation errors weare independent from each other.

320 Next, we performed a n parameter optimisation of parameters across the studied geographical domain (WS4d phase in Figure 2) by processing the plausible subsets of outputssimulations from all calibration sites (selected in the WS4b phase in Figure 2) together. First, wWe identified feasible ranges for each parameter and site based on intersections of variable-specific parameter ranges (Figure 3a). by overlapping the results of the three output variables (a). Then, we derived the multi-site feasible parameter ranges by as an intersection intersecting of the site-specific feasible parameter ranges (Figure 3b) in the whole tested 6D

325 parameter space defined by the calibrated parameters. Afterwards, we divided the multi-site feasible ranges of each parameter into five equally wide sub-intervals that defineing 5 discrete steps on each dimension (Figure 3b) of the multi-dimensional parameter space. This categorisation was neededrequired because the applied parameter values in site-specific Monte Carlo simulations differed between the sites. In the hyperspace, we identified ten cells with the highest number of allocated sites and simulations and the smallest arithmetic mean errors of AbgwC, SoilC, and LitterC, . Finally, and we calculated mean parameter

330 values of parameters per for each cell. The final multi-objective, i.e. multi-site multivariate (MSMV) optimal parameter set was the one that led to successful simulations of all calibration sites and generated the smallest errors.

(a)

(b)

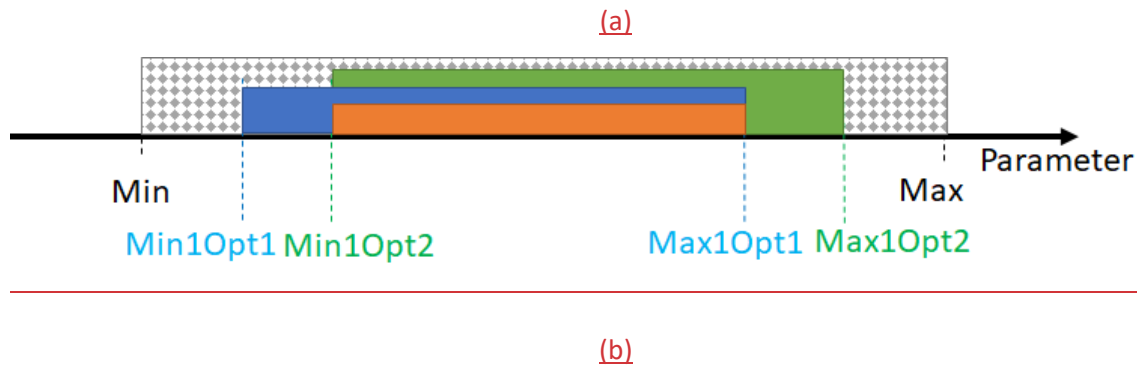
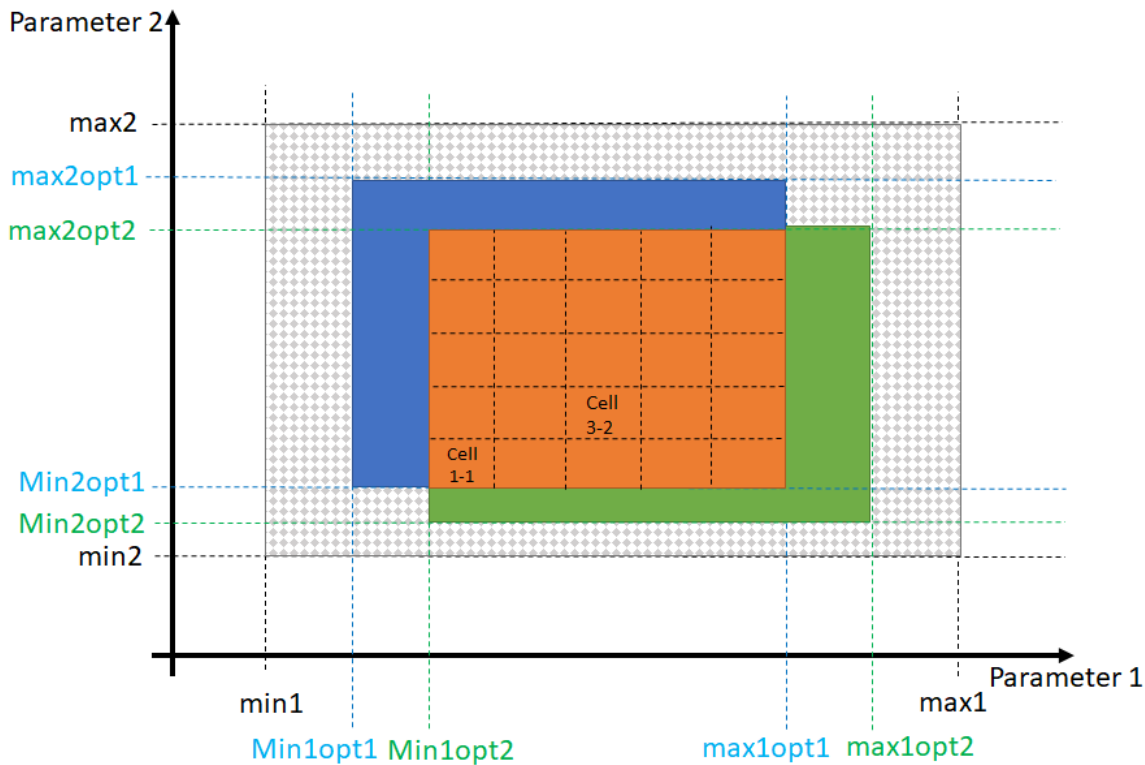



Figure 3 Identification of site-specific feasible parameter ranges ~~per site~~ using a multivariate approach (a); and identification of multi-site and multivariate optimised (MSMV) parameter values and their parameter ranges (b). The approach in (a) is demonstrated ~~by~~ using an example for a single site, a single parameter and two variables. The tested parameter interval is indicated by the grey dotted pattern, feasible parameter ranges for individual variables are indicated with blue and green colours, and the multivariate feasible parameter range if shown in orange. The approach (b) is demonstrated using a ~~on an example of~~hypothetical 2D parameter space and ~~the~~ site-specific optimised parameter ranges ~~of~~for two sites. The grey dotted space indicates the tested 2D parameter space, ~~T~~the blue and green ~~C~~colours refer to 2D parameter space for specific sites, and the orange space refers to multi-site and multivariate feasible parameter ranges. Min = minimum, Max = maximum, Opt = optimised values of tested parameters.~~the .B~~

2.5. Robustness, validation and plausibility tests

2.5.1. Robustness of calibrated parameter values

The robustness of the site-specific (SSMV) and multi-site (MSMV) optimised values of parameters was tested by simulating all calibration sites with all derived parameter sets and calculating the root mean square errors (RMSE) of the simulated output variables. This test allowed us to examine the ~~problem of~~possible overfitting ~~and~~ as well as the applicability of site-specific parameter sets outside the ~~specific~~actual site conditions. ~~The results were presented with heatmaps produced using ggplot2 R package (Wickham, 2016).~~

Next, we applied ~~the machine learning method of~~ decision trees (DT) to identify the problems ~~due~~arising from~~to~~ the trade-off between variables and/or sites based on the evaluation of the entropy. For each site and output variable w~~With the DT method~~

350 ~~w~~We determined parameter ranges, within which the plausible simulations can be expected with the highest probability. ~~This was done by~~ by ~~deriving~~ constructing ~~ing classification decision trees~~DT ~~from~~for each site and output variable using the outputs of ~~all~~ 100,000 Monte Carlo simulations per calibration site. The simulations ~~that~~ were split into ~~2 groups of~~ plausible and implausible ones based on the constraints specified for the three tested variables (done under ~~WS4a and~~the WS4b phase in Figure 2, ~~see~~ Sect. 2.4.3.). In total we derived 33 decision trees (11 sites x 3 variables) using the rpart R package (Therneau et

355 al., 2023). ~~From the DT results, w~~We ~~obtained~~ The plausible ranges for each parameter, site and variable were determined based on the selection of the most probable leaf node. The ranges~~,~~ may differ from those obtained under the WS4b phase, where each parameter was analysed separately (Figure 3a), because DT evaluate all parameters at the same time (Figure 4). Then, we searched for the intersections of these ranges to obtain the final parameter-wise plausible ranges for a specific site and all variables together and for multi-site ranges for all calibration sites and variables together (Figure 4). The SSMV and

360 MSMV optimised parameter values were considered robust if they occurred within the respective DT parameter ranges.

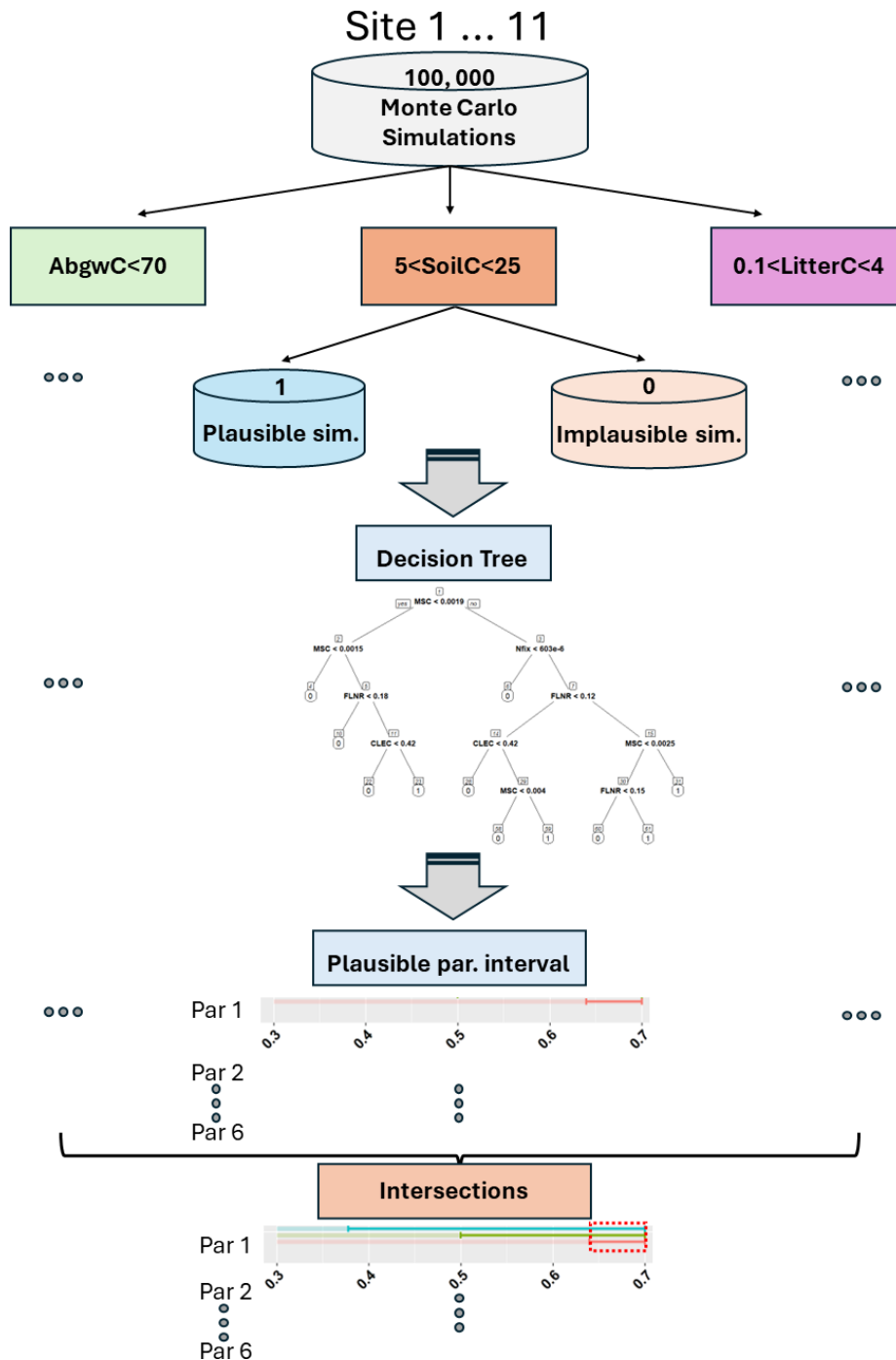



Figure 4 Determination of plausible ranges of parameters using ~~the machine learning method of~~ decision trees across all sites and output variables. Par = parameter, background transparent coloured horizontal lines represent the tested parameter range, while non-transparent coloured lines show the optimised parameter ranges based on decision trees performed for carbon stock in aboveground wood (AbgwC), soil (SoilC), and litter (LitterC). The dashed red rectangle represents the plausible range of ~~one~~ a single parameter based on the intersection of its plausible ranges derived for the three output variables.

The next test was aimed at analysing if SSMV parameter values followed any trends along specific gradients. Specifically, we examined the interdependencies between the pairs of parameters, and the trends of their site-specific values along stand and site gradients, namely age, elevation, latitude, climate and soil characteristics. Any significant trend may indicate that the respective parameter should not be handled as a constant but should vary with specific conditions. The analyses were based on Spearman correlations, and linear regressions performed in R environment (R Core Team, 2018). Subsequently, Wwe examined the physiological meaningfulness of revealed trends and relationships using the empirical evidence collected from the scientific literature.

Following the identification of the significant relationships and their biological plausibility, we derived ten multiple linear regression models explaining the variation in site-specific values of the canopy light extinction coefficient (CLEC) of calibration sites using different combinations of environmental predictors and analysed the explanatory power and performance of the models with the following statistical characteristics: R-squared, adjusted R-squared, AKAIKE information criterion, Bayesian information criterion, Mallows' statistic, and the residual standard deviation. This was performed in R environment using the car (Fox et al., 2023) and lmSubsets (Hofmann et al., 2021) R packages. Then we applied the derived functions to all 87 sites, for which we calculated site-specific values of CLEC. and simulated each site with the respective CLEC value. Afterwards, we examined the robustness and plausibility of the simulated output with varying CLEC across the whole geographical domain.

2.5.2. Model performance with parameter values optimised for multiple sites

The MSMV parameter set of BBGCMuSo derived in the WS4d calibration phase (Figure 2) was validated using an independent dataset from 8 European beech--dominated sites located within the studied geographical domain (

Table 1, Figure 1), each represented by at least two repeated observations of the aboveground wood carbon. The simulations consisted of spin-up, transient run and normal run, as described in Sect. 2.3above. The validation was based on the comparison of modelled and observed carbon stocks in aboveground wood at specific time points., for which wWe calculated the bias defined as an arithmetic mean of differences between modelled and observed values of the respective variable, mean absolute error (MAE), mean percentage error (MPE), and root mean square error (RMSE). SoilC and LitterC were compared to the plausible ranges derived from the literature (De Vos and Cools, 2011; Pavlenda and Pajtík, 2010; Wellbrock et al., 2016; Wellbrock and Bolte, 2019), since no observed data on soil and litter carbon were available for the validation sites.

2.5.3. Robustness and plausibility of simulated output at a large scale

To evaluate the broad applicability of the derived multi-objective parameter set across the whole studied geographical region, we simulated forest development atacross 87 sites, encompassing the main climatic and soil gradients in the study area (Figure 1). We specifically assessed the plausibility of absolute values of AbgwC, LitterC, and SoilC by comparing them with values documented in the literature (e.g. Barna et al., 2011; Pavlenda and Pajtík, 2010). We also examined the responses of simulated carbon stocks to environmental conditions (e.g. latitude, longitude, elevation, annual precipitation, mean temperature, proportion of sand, silt, and clay fractions, soil depth) and compared their observed shapes to the patterns published infrom empirical studies. We analysed the responses of simulated outputs using Spearman correlations, linear and quadratic regressions, and generalised additive models (GAM). When explaining the patterns of the main three variables along environmental gradients we examined also other stocks, such as carbon stock in roots, leaves, as well as some carbon fluxes, particularly heterotrophic respiration, to reveal the mechanisms driving the model responses. All tests were performed in the R environment (R Core Team, 2018) and. the results were visualised using ggplot2 R package (Wickham, 2016).

3. Results

3.1 Parameter sensitivity analysis

The local (single parameter) sensitivity analysis (WS2a in Figure 2) focusing on evaluating the effects of individual parameters revealed that the aboveground wood carbon stock was most affected by the whole plant mortality rate (WPM). Soil and litter carbon stocks were most sensitive to the maximum stomatal conductance (MSC) and nitrogen fixation (Nfix), respectively (Table 2). The analysis of trends in variable changes due to modifications of parameter values clarified how the increase of the parameter value affected the values of the respective output variable, e.g. the increase of MSC caused an increase of all tested output variables (AbgwC, SoilC, LitterC) in the whole parameter range (Fig. S4). The impact of other parameters was more complex, as we revealed both positive and negative trends, while in the case of the increase of e.g. Nfix the positive ones prevailed in AbgwC and SoilC, and negative ones in LitterC. A more detailed analysis identified the changes of output variables along the parameter range, e.g. the increase of Nfix caused an initial increase of LitterC, which was followed by its gradual reduction as Nfix was increasing (Fig. S2a).

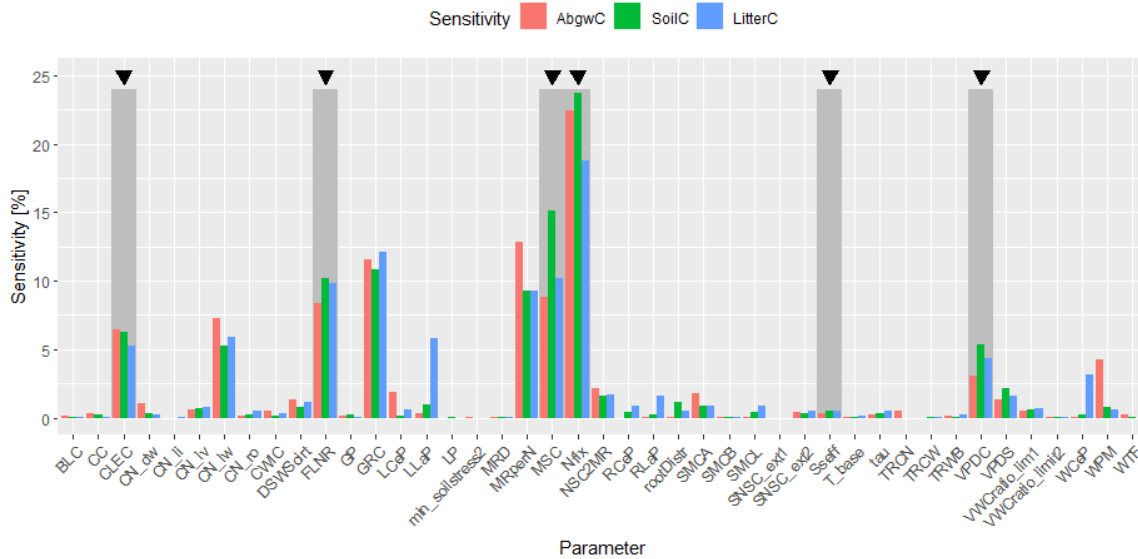
The global (multi-parameter) sensitivity analysis (WS2b in Figure 2) showed that Nfix had the highest impact on all three analysed carbon pools (Figure 5). The subsequent parameters were MSC, growth respiration per unit of carbon allocation (GRC), maintenance respiration in kgC day^{-1} per kg of tissue nitrogen (MRperN), fraction of leaf nitrogen in Rubisco (FLNR). However, the ranking of parameters differed between the individual output variables (Figure 5).

425

Figure 5 Results of global (multi-parameter) sensitivity analysis (WS2b in Figure 2) of simulated carbon stock in aboveground wood, soil and litter (AbgwC, SoilC, LitterC) to ecophysiological parameters of BBGCMuSo. Parameter abbreviations are given in Table 2. Dark grey rectangles with black arrows above indicate parameters selected for the calibration procedure: canopy light extinction coefficient (CLEC), fraction of leaf nitrogen in rubisco (FLNR), maximum stomatal conductance (MSC), nitrogen fixation (Nfix),
430 **effect of soil stress factor on photosynthesis (Sseff), and vapour pressure deficit for complete conductance reduction (VPDC).**

Based on the results of ~~the sensitivity analysis~~SA, we selected six parameters to be calibrated: canopy light extinction coefficient (CLEC), FLNR, MSC, Nfix, vapour pressure deficit for complete conductance reduction (VPDC), and effect of soil stress factor on photosynthesis (Sseff), as they had a substantial effect on carbon stock in aboveground wood, soil, and litter (Figure 5). Other parameters with a high influence on the simulated C pools (such as C:N ratio in leaves, Fig. S2b) were
435 not to selected for calibration due to a strong support of their actual values from the literature (e.g. Fig. S3).

3.2 Parameter estimation

The site-specific multivariate (SSMV) optimised values of all 6 calibrated parameters for 11 calibration sites differed from the a priori ones, varied within the whole tested ranges (Table 2) and strongly differed between the sites (Figure 6). Two parameters (FLNR, VPDC) showed a gradual change of SSMV values across the whole tested range, while the SSMV values of others,
440 especially CLEC and Sseff, were clustered. In comparison to a priori values, SSMV values of individual parameters changed by 41 % on average, while Nfix was modified most substantially (67 % on average).

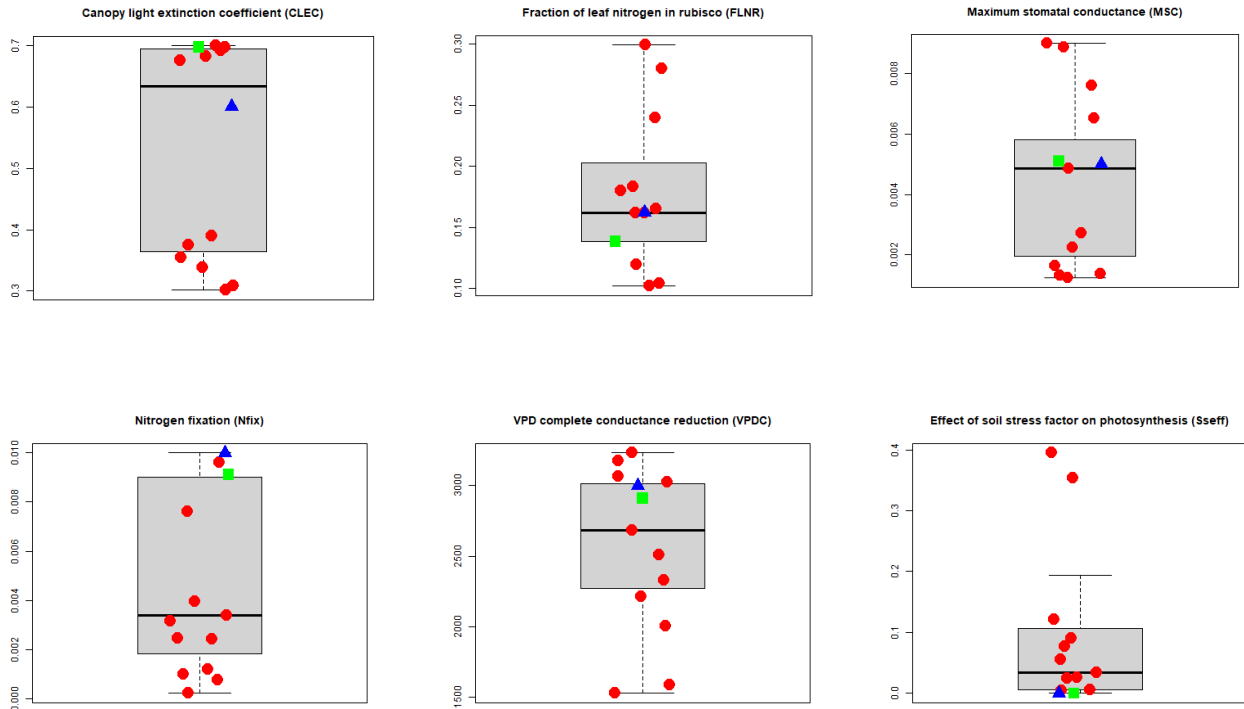



Figure 6 A priori (blue triangles), site-specific (SSMV, red dots) and multi-site (MSMV, green squares) optimised values of calibrated ecophysiological parameters selected based on the multi-objective sensitivity analysis (S4 WS2 in Figure 2). The thick horizontal lines represent medians, boxes represent the interquartile ranges (IQR), and the whiskers represent ±1.5IQR.

445   The median reduction of model errors at calibration sites simulated with SSMV parameter sets in comparison to the a priori set was 35 %, 55 %, and 11 % for AbgwC, SoilC, and LitterC, respectively, and 26 %, 35 % and 9 % in comparison to the simulation output obtained with the MSMV parameter set (Figure 7).

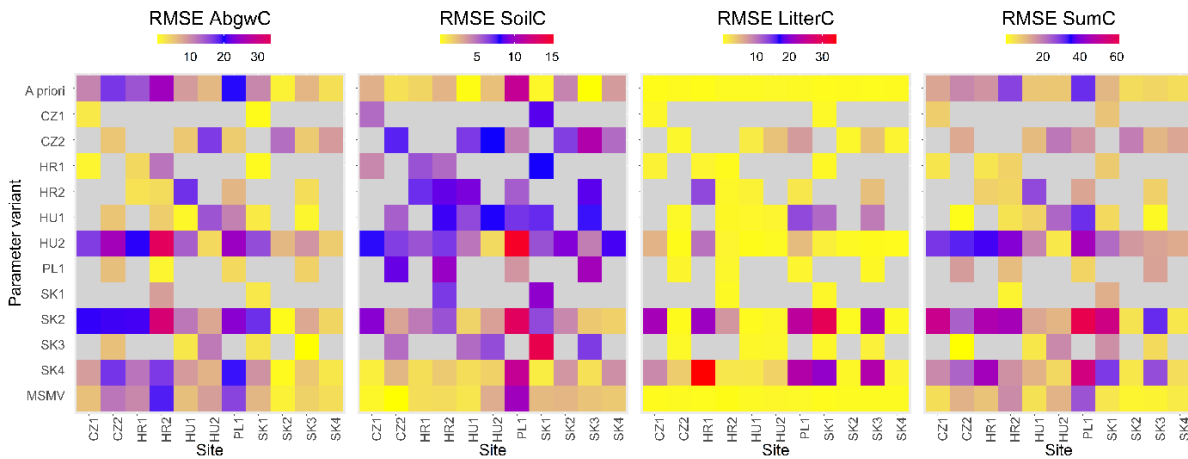



Figure 7 Heatmaps of root mean square errors (RMSEs) for carbon stocks in aboveground wood, soil, and litter (AbgwC, SoilC, LitterC) and their sum (SumC) for individual calibration sites (~~presented on x~~X axis) and 13 ~~different~~ variants of parameter sets (~~presented on y~~Y axis): one a priori set, 11 site-specific (SSMV) optimised ~~ones~~ sets identified by site abbreviations, and one multi-objective (multi-site and multivariate, MSMV) optimised set. The ~~red colour indicates the largest RMSEs, while the yellow one indicates the smallest RMSEs, the~~ grey colour indicates unsuccessful simulations, which ended with zero or close to zero values of carbon state variables.

Parameter values obtained from the multi-site optimisation (~~W~~S4d in Figure 2) were changed by 2 to 15 % ~~of~~ relative to their a priori estimates except for the Sseff, which remained unchanged (Figure 6, Table 2). The MSMV values substantially differed from the SSMV values of most sites, although the differences were parameter specific. The lowest variation was observed for Sseff, while the largest differences between MSMV and SSMV values were found for Nfix (Figure 6). The ~~Simulation~~ simulation results obtained using the MSMV parameter set showed reduced mean errors ~~for~~of all three ~~tested~~ output variables, i.e. AbgwC, SoilC and LitterC, in comparison to the results obtained with the a priori set, by 10, 26, and 5 %, respectively (Table 3, Figure 7).

Table 3 Evaluation of model performance ~~based on simulations of beech ecosystems~~ at 11 calibration sites using the a priori parameter set and the ~~optimised~~ parameter sets optimised with respect to ~~the measured~~ observed ~~sets of~~ carbon stocks in aboveground wood, soil, and litter (AbgwC, SoilC, and LitterC in kgC m^{-2}). "Site-specific" refers to parameter sets derived for single sites using a multivariate approach (SSMV, WS4c phase in Figure 2). "Multi-site" refers to multi-site and multivariate (MSMV) parameter values derived collectively for all calibration sites (WS4d in Figure 2). RMSE = root mean square error, Bias = arithmetic mean of differences between modelled and observed values of the respective variable, MAE = mean absolute error, MPE = mean percentage error, MinDif and MaxDif are minimum and maximum differences between the modelled and observed values of the respective variable.

Parameter set	Variable	N	RMSE	Bias	MAE	MPE	MinDif	MaxDif
A priori		96	9.7727	7.3911	7.8366	29.502	-8.1569	27.007
Site-specific	AbgwC	96	2.0698	-0.2887	1.4873	0.012	-5.4830	5.7422
Multi-site		96	6.8718	4.1734	5.4798	19.246	-9.5527	27.756
A priori		72	3.5242	2.2276	2.2630	58.329	-0.2268	11.729
Site-specific	SoilC	72	6.8068	-2.7876	6.3782	-0.408	-9.7400	4.7630
Multi-site		72	3.0803	0.1167	2.6423	38.068	-2.4427	10.243
A priori		10	0.2424	0.2117	0.2117	121.774	0.0611	0.4651
Site-specific	LitterC	10	0.2035	0.0030	0.1653	42.005	-0.2894	0.4382
Multi-site		10	0.1849	0.1417	0.1419	91.327	-0.0007	0.4185

The simulations performed with the site-specific and multi-site optimised parameter sets produced more accurate estimates of carbon stocks in aboveground wood, soil and litter than the a priori set. The non-parametric Wilcoxon signed rank test (Wilcoxon, 1945) with the continuity correction data confirmed insignificant differences between the observed and modelled AbgwC ~~carbon stock in aboveground wood~~ and LitterC simulated with SSMV parameter sets (V = 1807, p = 0.14 for AbgwC, and V = 26, p = 0.92 for LitterC), while the differences for soil carbon were significant (V = 528, p = 1.04e-05, see Sect. 4.2 for explanation). The use of the MSMV parameter set resulted in insignificant differences of SoilC (V = 1396, p = 0.64), while the estimates of AbgwC and LitterC were significantly different from the observations (V = 4094, p = 1.10e-10, and V = 54, p = 0.004, respectively). Nevertheless, the magnitudes of their mean errors calculated for the whole set of calibration sites as well as for most of the individual sites were substantially reduced (Table 3, Figure 7).

3.3 Robustness, validation, and plausibility tests

3.3.1 Robustness of calibrated parameter values

When simulating the development of calibration sites with the site-specific parameter sets (SSMV) optimised for other sites, we revealed a high variation in modelled outputs per site (Figure 7 and Fig. S5). In 47 % of cases we encountered unsuccessful simulations, during which the modelled forests ~~did not survive~~ceased to exist. Only three SSMV parameter sets led to

successful simulations ~~of~~ in all calibration sites, but their errors exceeded those obtained with the a priori, respective SSMV or MSMV optimised sets (MPE of AbgwC, SoilC, and LitterC were 39 %, 296 %, and 570 %, respectively).

The robustness test of SSMV and MSMV optimised parameter values using the ~~method of~~ decision trees (DT) revealed that in 78 % of cases site-specific and muti-site parameter values occurred within the parameters ranges derived from decision trees for individual output variables and sites (Table 4, Fig. S8). The discrepancies between the optimised parameter values and DT ranges occurred for the variables and sites with lower proportions of plausible simulations (Fig. S16~~Fig. A15~~) or in the cases when DT parameter ranges derived for individual output variables did not overlap (e.g. the ranges derived for FLNR and site CR2015, Fig. S8).

Table 4 Robustness of site-specific (SSMV) and multi-site (MSMV) parameter values based on the analysis ~~if~~ of the optimised parameter value occurred ~~inside~~within the plausible ~~the~~ parameter ranges derived for the individual calibration sites and carbon stock~~s~~ ~~variables~~ (aboveground wood, soil and litter carbon labelled as AbgwC, SoilC, LitterC) using ~~the method of~~ decision trees (DT, Sect. 2.5.1). The values for SSMV represent the mean ~~from~~ of 11 calibration sites. The value 1 indicates that the MSMV value or all SSMV optimised parameter values occurred within the DT ranges. ~~The abbreviations of parameters are~~ CLEC = canopy light extinction coefficient, FLNR = fraction of leaf nitrogen in rubisco, MSC = maximum stomatal conductance, Nfix = nitrogen fixation, VPDC = vapour pressure deficit for complete conductance reduction, Sseff = effect of soil stress factor on photosynthesis.

| | | Proportion of optimised parameter values inside parameter ranges derived using decision tree method | | | | | |
| | Output variable | AbgwC | | SoilC | | LitterC | |
	Parameter set	SSMV	MSMV	SSMV	MSMV	SSMV	MSMV
Parameter	CLEC	1.00	1	0.91	1	0.91	1
	FLNR	0.73	1	0.64	0	0.64	0
	MSC	1.00	1	0.55	1	0.36	0
	Nfix	0.91	0	1.00	0	0.36	1
	VPDC	1.00	1	1.00	1	1.00	1
	Sseff	1.00	1	1.00	1	1.00	1
	Mean of All	0.94	0.83	0.85	0.67	0.71	0.67

The analysis of interdependencies between the site-specific optimised parameter values revealed the only significant Spearman correlation between CLEC and MSC, which were negatively correlated at 95 % significance level (r=-0.6, p=0.04, Figure 8a). The highest, although non-significant, correlation was found between Nfix and FLNR (r=0.7, p=0.37) suggesting that if Nfix increases FLNR should also increase.

Spearman correlations between parameters and site characteristics revealed that the site-specific optimised values of two calibrated parameters (CLEC, VPDC) were significantly related to elevation (r=0.5 and -0.7, p=0.04, respectively). Significant relationships were also found between CLEC and several climatic variables (Figure 8). The highest positive correlation (r=0.9, p=0.01) of CLEC was with the long-term mean annual precipitation total (AMPRCP), and the highest negative correlation (r=-0.8, p=0.01) with the long-term mean annual vapour pressure deficit (AMVPD). The increasing AMVPD was positively

significantly related to the values of MSC (r=0.7, p=0.05). For other parameters we did not reveal any significant relationships with climate conditions, nor could we confirm any significant correlations of parameter values to soil conditions (Fig. S9),

525 although positive or negative trends with several environmental characteristics were ~~visible~~ identified for some parameters, mainly for MSC and VPDC (Figure 8b, Fig. S9).

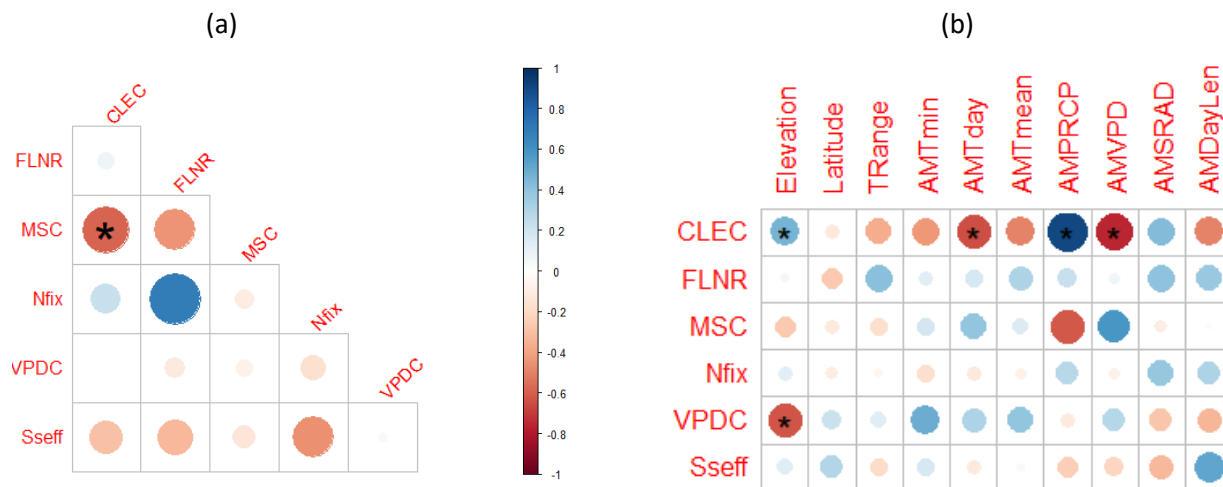



Figure 8 Spearman correlations between site-specific optimised values of calibrated parameters (a) and between site-specific optimised parameter values and site characteristics (b). The colour and the size of the circles indicate the value of the correlation
530 **coefficient~~,~~ ~~. and t~~The stars indicate the significance of the correlation (1 star for 95 % significance level, 2 stars for 99 % significance level). ~~The abbreviations of parameters are:~~ CLEC = canopy light extinction coefficient, FLNR = fraction of leaf nitrogen in Rubisco, MSC = maximum stomatal conductance, Nfix = nitrogen fixation, VPDC = vapour pressure deficit for complete conductance reduction, and Sseff = effect of soil stress factor on photosynthesis. Climate characteristics represent long-term annual averages: TRange = Temperature range, AMTmin = Minimum temperature, AMTday = Daylight temperature, AMTmean = Mean**
535 **temperature, AMPRPCP= Precipitation total, AMVPD = Vapour pressure deficit, AMSRAD = Daily solar radiation, AMDayLen = Daylength.**

3.3.2 Model performance with parameter values optimised for multiple sites

The mean absolute error (MAE) between the simulated and observed aboveground wood carbon of 8 validation sites was 0.26

540 kgC m^{-2} with a 95 % confidence interval from -0.025 to 0.56 kgC m^{-2}, while the individual absolute differences varied between -2.06 and 5.11 kgC m^{-2} (Figure 9). The root mean square error was 1.22 kgC m^{-2}. The non-parametric Wilcoxon signed rank test with continuity correction indicated non-significant differences between simulations and observations of aboveground wood carbon (V = 1385, p-value = 0.1962). The mean percentage error of AbgwC was 1.25 % of the observed carbon stock in aboveground wood. Hence, both absolute and relative differences were of negligible magnitudes.

545 Since no observed data on soil and litter carbon were available for validation sites, the simulated SoilC and LitterC were tested against their ranges reported in the literature. The results showed that both variables occurred within the plausible ranges (Fig. S12).

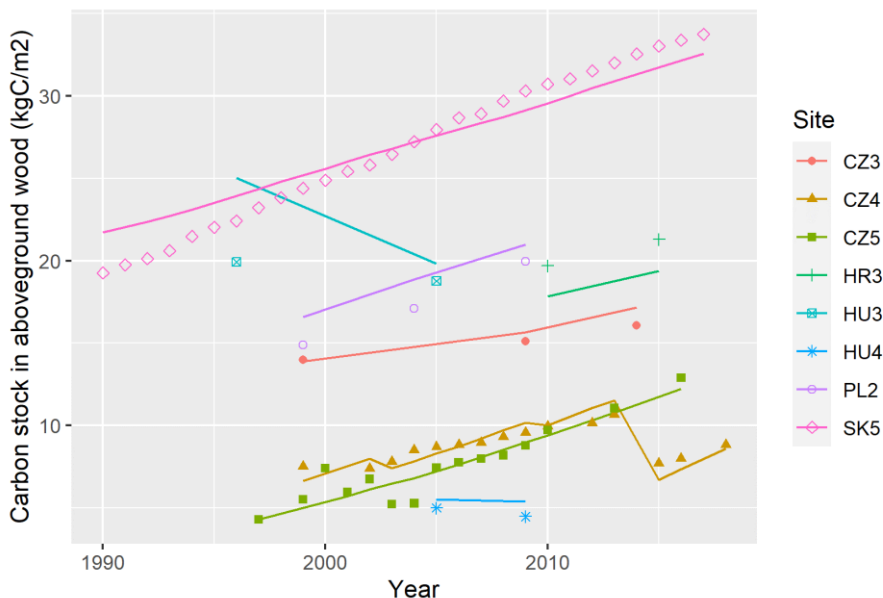



Figure 9 Comparison of the temporal development of observed (points) carbon stock in aboveground wood (AbgwC) and simulated AbgwC (lines) using the multi-site multivariate optimised (MSMV) parameter set at 8 validation research sites.

550

3.3.3 Robustness and plausibility of simulated output at a large scale

The simulations of all 87 research sites using the multi-site optimised (MSMV) parameter values were successful and the simulated values of the three output variables (i.e. AbgwC, SoilC, and LitterC) at the end of the spin-up run were well-aligned with the plausible ranges indicated in the literature (Fig. S13).

555 The simulated values of output variables varied across the studied geographical space (Figure 10) and were significantly correlated with multiple several site characteristics (Figure 11, Fig. S17, Fig. S18Fig. A17). The modelled AbgwCaboveground wood carbon exhibited distinct unimodal responses along the gradients of elevation, long-term mean air temperature, and vapour pressure deficit (VPD). The results manifested a production optimum of beech in Central Europe at elevations of 500 - 600 m a.s.l., mean annual air temperature of 9°C, and VPD of 530 Pa. The SoilC and LitterC demonstrated an increasing

560 trend along the elevation gradient and decreasing trends along the climatic gradients. The responses of carbon stocks to soil properties generally followed a linear pattern, while SoilC and LitterC significantly decreased with the increasing clay content, and were positively correlatedincreased with the increasing sand content in soil. Aboveground wood carbon was found to be significantly correlated with the proportion of clay in the first and second soil layers, and with the proportion of silt in the fifth

layer (Fig. S18~~Fig. A17~~). The highest levels of simulated AbgwC~~aboveground wood carbon~~ were observed on loamy or sandy-

565   loamy soils (Fig. S20).

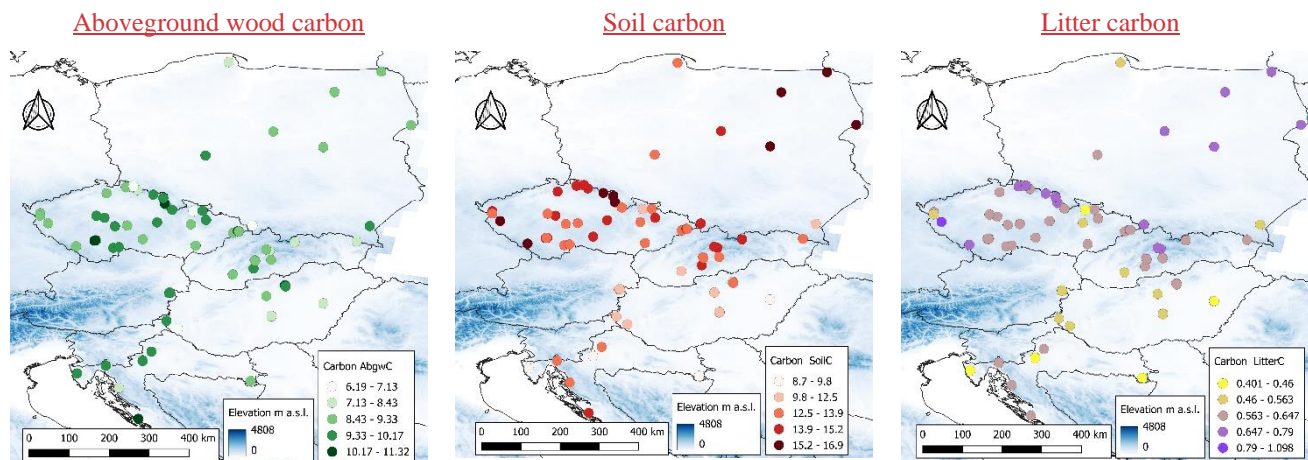
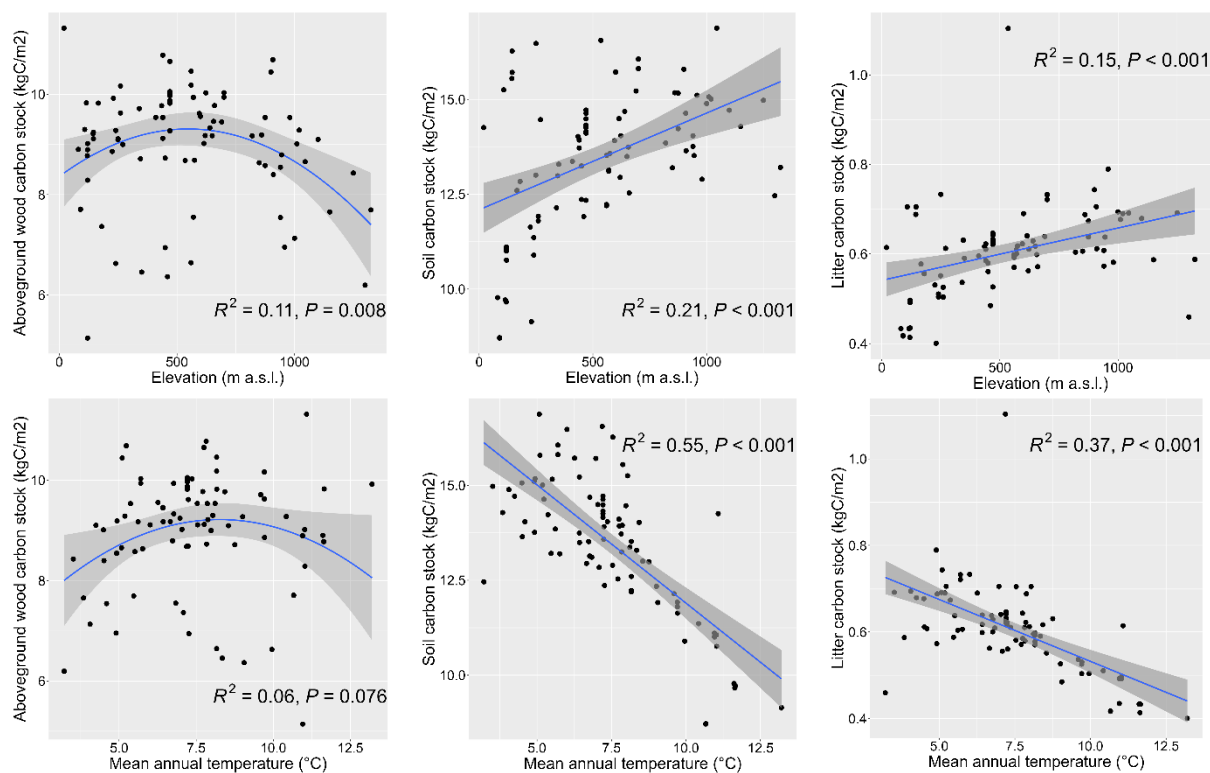
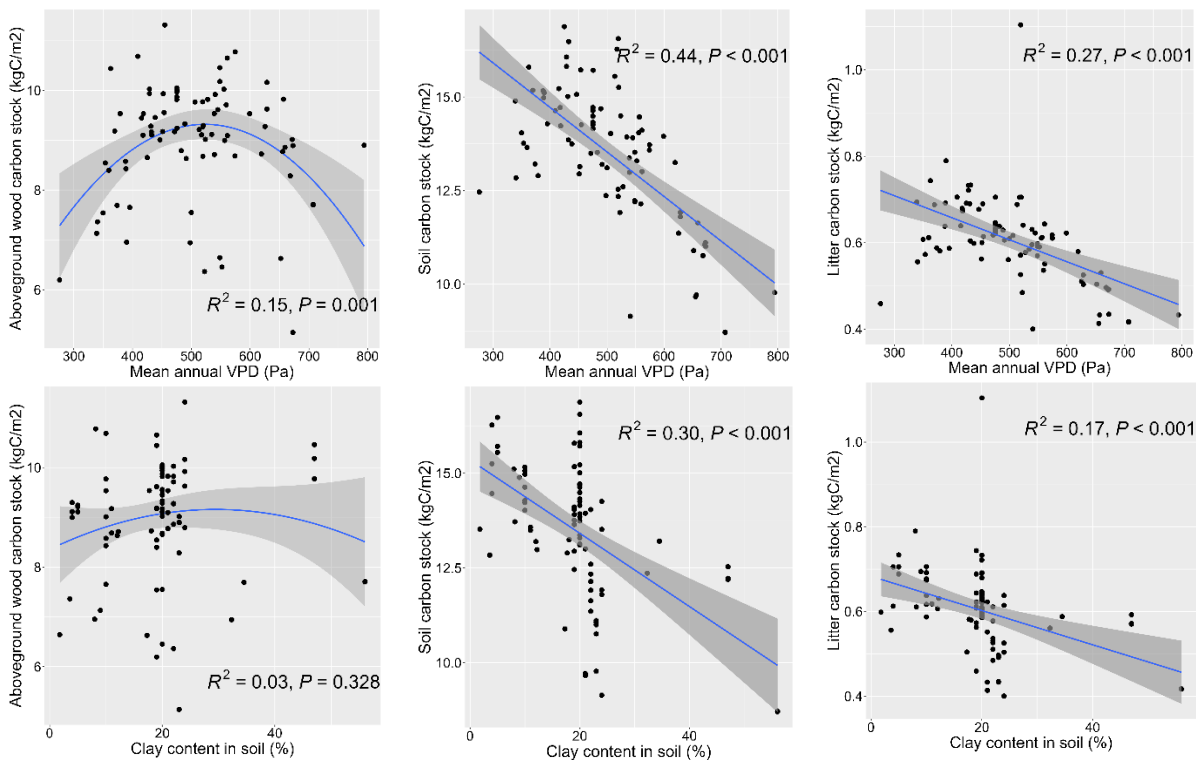



Figure 10 **Simulated values of carbon stocks in aboveground wood, soil and litter (kgC m^{-2}) in beech ecosystems at the standardised stand age of 35 years across 87 sites. The blue background indicates the elevation gradient taken from (****Hengl et al.****, (2020)****.**

570 **Figure 11 Responses of modelled carbon stocks in aboveground wood at the standardised stand age of 35 years (left), soil (middle) and litter (right) carbon stocks to selected environmental variables. The simulations ~~with BBGCMuSo~~ were conducted for 87 sites distributed across central Europe (~~see~~ Figure 1). Linear or quadratic regressions were fitted to the data. R² represents the squared regression coefficient, ~~and~~ P is the p-value for the F-test of the fitted model.**

4. Discussion

575 ### 4. 1. Selection of parameters for calibration

~~The sensitivity analysis identified the parameters with the highest impact on the selected model output (Table 2, Figure 5).~~ The global SA showed small differences in the parameter impact between the three examined output variables, because they are all a part of the carbon cycle, which makes them naturally interconnected ~~within forest ecosystem~~s. The similarity in parameters affecting these output variables underscores the importance of considering a forest holistically, since due to the interactions

580 between different components of the ecosystem changes in one part have cascading effects throughout the system. Therefore, a comprehensive approach that accounts for these interdependencies is crucial for accurate modelling and understanding of forest carbon dynamics.

Several highly influential parameters identified by SA were excluded from the calibration for different reasons. Growth respiration per unit of carbon allocation (GRC), and maintenance respiration per kg of tissue nitrogen (MRperN)~~,~~ were not

585 calibrated because we could not support the modification of their values~~y were already~~ used in previous model versions

(Thornton et al., 2005) ~~, and we could not support the modification of their values~~ by observations. Such data would be needed since the empirical evidence suggests that respiratory parameters may substantially differ between sites (Lavigne and Ryan, 1997). Other parameters were not included in calibration because the data from literature supported their current values and/or because of the adverse impact on the variables of interest. For example, C:N ratio in leaves was found to have a substantial

590 effect on all examined carbon pools (Figure 5), but to obtain a desired reduction ~~of~~ in ~~carbon stock in the aboveground wood~~AbgwC, we would need to increase this parameter from its a priori value (Fig. S2). However, the analysis of the values obtained from site-specific measurements performed at~~on~~ some ~~of our~~ sites and the data from the TRY database (Kattge et al., 2020) ~~revealed~~ suggested that increasing the value of C:N ratio in leaves would cause a~~its~~ significant deviation from ~~its~~the mean or median of experimental observations (Fig. S3). Similarly, the parameter representing the natural whole plant mortality

595 was not calibrated since ~~instead of the temporarily constant mortality~~ we applied dynamic mortality rates during the normal run simulations instead of the temporarily constant mortality. The rates ~~that~~ depended on the applied ~~forest~~ management and were derived from published field observations (Barna et al., 2011; Hülsmann et al., 2018; Pajtík et al., 2018; Vanoni et al., 2019). The possibility of using dynamic mortality rates in BBGCMuSo is a major improvement over~~in comparison to~~ the original Biome-BGC model, as mortality has been found to be a driving process of vegetation dynamics in forest growth

600 models (Bugmann et al., 2019; Hlásny et al., 2014).

The identification of influential parameters is crucial not only for model calibration, but also for future ~~local~~ studies as it provides valuable information about the key parameters, the values of which should be collected in the field, since applying site-specific values obtained from experimental data may substantially reduce the uncertainty of model simulations. As~~lready~~ ~~(~~Thornton et al. ~~(~~, 2002) already presented~~, that~~ some parameters, such as C:N ratio in leaves, should be treated as site-specific.

605 In the case the site-specific values of parameters are available, they need to be set prior to the calibration due to the covariance of parameters and need to be excluded from the sensitivity analysis. In our case we aimed at a generic parameter set applicable across the whole studied region, ~~However, such data were not available for most of our sites,~~ and thus, we used a generalised value of this parameter.

4.2. Parameter estimation

610 The mean error metrics (root mean square errors, mean absolute and percentage errors) showed the increased accuracy of modelled output obtained using single-site (SSMV) and multi-site (MSMV) optimised parameter sets in comparison with those obtained with the a priori parameter set (Table 3). Greater errors of soil carbon for SSMV parameter sets resulted from site-specific very tight positive relationships between the simulated AbgwC and SoilC revealed when analysing Monte Carlo simulations. Due to this, even the high number of performed simulations (100,000 simulations per site) generally covered only

615 a small portion of the 2D space defined by AbgwC and SoilC (Fig. S6). Hence, in some cases improving the results for one of these two carbon pools caused the increase of the error of the other pool within the tested space of six calibrated parameters (Fig. S7). Although some empirical studies reported the positive correlation between aboveground wood and soil carbon stock~~,~~

~~particularly~~ in~~of~~ the top soil, the relationship is not ~~less~~ strong (R^2=0.24, ~~(~~Woollen et al., 2012), and frequently insignificant (Osei et al., 2022), as SoilC primarily depends on climate, topography, soil mineralogy and soil texture, especially the content

620 of clay (Powers and Schlesinger, 2002) or sand (Devi, 2021). Our results indicated that different or additional parameters may need to be included in the calibration that may increase the variability of model output and thus loosen the current high correlations between AbgwC and SoilC. ~~or the model to increase the accuracy of model estimates of both variables and/or the spatial coverage defined by observations.~~

The application of site-specific calibrated parameter sets outside the respective sites pointed out at the contradiction between

625 their generality and the local accuracy of model estimations. SSMV optimised parameters were not generally applicable, as 47 % of simulations of calibration sites with SSMV parameter sets optimised for different sites collapsed (Figure 7). Calibrating models for individual sites may often result in model overfitting due to the small amount of available data (Tsai et al., 2021), which may lead to completely different parameter sets in the case of a recalibration, and hence a high variance in calibrated values, reducing thus the reusability of calibrated parameter values.

630 The parameter values optimised for single and multiple sites frequently substantially differed (Figure 6), which indicates the existence of the calibration equifinality, i.e. that many different parameter sets may produce similar output predictions (Beven, 2019). This issue was not apparent at the multi-site level, but occurred at the level of individual sites, at which it can be partially solved using the conditional interval reduction method (CIRM, Hollós et al., 2022). The CIRM approach is based on the iterative narrowing of plausible intervals of parameters using the constraints on the model output. It is an efficient way of

635 dealing with the equifinality unless the contradictions between different output variables occur.

As expected based on similar calibration works of forest growth models that used comprehensive data sources (Forrester et al., 2021; Minunno et al., 2019), the multisite and multivariate calibration increased the generality and robustness of the model application by finding a parameter set that worked across all calibration sites (Figure 7), and the validation set (Figure 9). Nevertheless, differences in data quality and availability across space can substantially influence the calibration results.

640 ~~Although t~~This problem can be mitigated with~~using~~ ~~Bayesian calibration~~ calibration techniques that utilis~~ze~~ ~~and by choosing the~~a more -appropriate likelihood function (i.e. formal likelihood; Hollós et al., 2022). ~~in frequentist calibrations,~~ In this way, the issue of the significant spatial autocorrelation ~~data quality at individual sites is not always independent from each other, since there is often a significant spatial autocorrelation. The extent of this interdependence is rarely quantifiable, therefore the~~can be resolved using the inclusion of an appropriate error-covariance matrix ~~used for the likelihood functions is usually~~

645 ~~incorrect, which can mislead the calibration process.~~in the construction of the likelihood function (Tarantola, 2005).

In addition, the multi-site calibration can also reduce spatial heterogeneity of model outputs ~~due to the "averaging out effect"~~ causing that the average model performance will be good, but at specific sites it can be completely wrong. This could be overcome by applying a hybrid approach that combines the site-specific and multisite calibration. With such an approach, site-specific values obtained from local measurements, well-established relationships derived from large databases or calibrated

650 for specific sites will be used for the parameters that are known to vary in space, while generic values will be used for the other

parameters. Thus, the overall correctness can be ensured by multisite calibration, while spatial heterogeneity by site-specific calibration.

Another aspect affecting the calibration is data availability. Although we tried to select calibration sites to cover the whole geographical and environmental space, the central part of the region was overrepresented, while northern parts were not covered due to the insufficient data required for calibration. Nevertheless, the ranges of environmental conditions (

Table 1, Figure 1) covered by our data included also extreme sites and seem to represent the natural distribution of European beech (Pagan, 1996) well. To ensure more robust calibration results, more balanced geographical coverage, more long-term data of multiple variables of interests at individual sites, and a combination of the information about stocks and fluxes at same sites would be required.

4.3. Trends in parameters

4.3.1 Covariance between parameters

The covariance analysis between parameters found correlations of different magnitudes indicating that in most cases the parameters were not independent. The revealed significant negative linear relationship between the site-specific values of the canopy light extinction coefficient (CLEC) and the maximum stomatal conductance (MSC, Figure 8) suggested that low values of CLEC should be coupled with high values of MSC and vice versa. However, we have not found any empirical evidence in the literature to confirm or refute the revealed relationship between CLEC and MSC, and therefore it is not clear if the revealed pattern is biologically realistic, or if it is only a side effect of the calibration procedure. Due to this, we did not incorporate this relationship into the calibration procedure. Another strong (R=0.7) though a non-significant relationship was revealed between nitrogen fixation (Nfix) and fraction of leaf nitrogen in Rubisco (FLNR, Figure 8). Examining this relationship in more detail revealed a non-linear pattern between the two parameters resembling a parabolic curve reaching a maximum of FLNR in the middle of Nfix range (Fig. S10). Similarly as for the previous relationship, we have not found any empirical research dealing with the presented issue although the study by (Tang et al., (2019) analysing different species in subtropical ecosystems suggests that nitrogen fixing trees allocate lower fractions of N to Rubisco than non-nitrogen fixing trees. The FLNR values of different *Eucalyptus* species published by (Warren and Adams, (2004) do not show any significant trend with the increasing nitrogen amount. The global study by (Luo, (2021) showed that FLNR is considerably affected by climate and soil factors, including light, atmospheric dryness, soil pH, and sand content. Based on these results, nitrogen fixation does not seem to be directly related to FLNR. Nevertheless, the environmental conditions that affect nitrogen availability can indirectly influence how nitrogen is allocated within the leaf, including Rubisco, suggesting a complex relationship between them. Still, the pattern of the relationship between FLNR and Nfix across one tree species in temperate ecosystems remains unclear.

Our results indicate the necessity of analysing the covariance between parameters during a model calibration as it not only enlightens the model behaviour and interdependencies between specific parameters but can also increase the efficiency of the calibration procedure by excluding one of the correlated parameters from the calibrated parameter set and estimating its value

only subsequently. In addition, such information may also help to identify the gaps in the available empirical evidence and the direction of future empirical research.

685

4.3.2 Parameter correlations with site characteristics

The analysis of the variation of the optimised site-specific values of parameters across environmental gradients revealed some significant trends (see Figure 8, Fig. S9). This may indicate that a specific parameter should not be kept as a constant but rather

690 as a characteristic that changes depending on driving conditions. An example of such a parameter is the canopy light extinction coefficient (CLEC) that specifies the proportion of solar radiation intercepted in the canopy. A number of process-based models that use this parameter set its value around 0.5, while some models differentiate values between different species (Zhang et al., 2014). Based on our multi-objective optimisation for beech ecosystems we set the MSMV value of CLEC to 0.66, which is by 10 % higher than the value used by Pietsch et al. (2005) in the original Biome-BGC model for beech (0.6), but it is within the

695 range for broadleaved forests reported by Zhang et al. (2014).

Although most models keep this parameter constant across sites and throughout their simulations (Liu et al., 2021; Zhang et al., 2014), in reality its value changes during a day as well as during a year as it depends on the solar zenith angle, leaf area, leaf inclined angle, and leaf clumping (Parker, 2020; Wang et al., 2004; Zhang et al., 2014). It also changes with stand age, while it reaches its maximum in young stands (Brown and Parker, 1994). A constant value of CLEC causes intra-annual errors

700 in estimations of plant transpiration and soil evaporation during a year (Tahiri et al., 2006). Due to this, a variable CLEC seems to be a more appropriate option. Our analysis revealed significant trends in the SSMV optimised values of CLEC with multiple environmental characteristics, while the trends with elevation and precipitation were positive, and with temperature and VPD negative (Figure 8). Such patterns were not observed in other studies analysing measured data of CLEC, but the number of tested observations was low (Zhang et al., 2014) as it was also in our case. Nevertheless, the empirical data showed that CLEC

705 increases with ~~the~~ decreasing plant density (Timlin et al., 2014) and LAI (Zhang et al., 2014). Since our model operates at a stand level, stand biomass can be considered as a proxy of stand density. Our results showed the decreasing trend of CLEC with the increasing aboveground wood carbon stock (Fig. S11~~Fig. A10~~), which is in agreement with ~~(~~Timlin et al.~~,~~ (2014). Moreover, the simulated AbgwC of calibration sites decreased along the elevational gradient (Fig. S11~~Fig. A10~~) explaining the positive correlation of CLEC to elevation, which can be considered as a side effect of stand density that is lower at high

710 elevations.

Tahiri et al. (2006) successfully applied a simple empirical approach using a linear regression with leaf area. Parker (2020) calculated CLEC as a ratio between the effective LAI and the total LAI. CLEC is usually calculated following the simplified Beer Lambert law as a function of above- and below-canopy solar radiation and leaf area (Zhang et al., 2014). A more sophisticated approach includes the solar zenith angle and the clumping index. Some models also use the inclination angle of

715 leaves, while most commonly the spherical distribution of leaves is assumed (Liu et al., 2021), although Pisek et al. (2013)

found that tree species in temperate and boreal regions are usually characterised by planophile or plagiophile leaf angle distribution.

Although we derived several multiple linear regressions explaining the variation in site-specific CLEC values using different combinations of environmental predictors (Fig. S15), the simulations with varying CLEC based on environmental conditions did not produce satisfactory results (33 % and 27 % of all and calibration sites collapsed if simulated with CLEC derived from its regression to annual precipitation, respectively). The possible reason is the existence of interdependencies between parameters discussed in Sect. 4.3.1.

~~We revealed a significant negative linear relationship between site-specific CLEC and the maximum stomatal conductance (MSC, Figure 6), suggesting that low values of CLEC should be coupled with high values of MSC and vice versa. Hence, changing the value of one parameter should be coupled with the modification of the value of another parameter. However, we have not found any empirical evidence in the literature to confirm or refute the revealed relationship between CLEC and MSC, and therefore it is not clear if the revealed pattern is realistic, or if it is only a side effect of the calibration procedure.~~

Moreover, we found that MSC was also significantly related to environmental conditions, namely VPD (Figure 8). MSC specifies the highest possible rate at which stomata can open and allow the exchange of gases between the plant and the environment under present-day CO_2 concentration and optimal environmental conditions, i.e. maximum radiation, and unlimited water availability, when VPD is zero, and there is no soil water stress. Such conditions rarely occur in the field, and hence, the observed maximum conductance, which represents the highest conductance on fully expanded leaves that was measured during the summer growing season (Murray et al., 2019), does not usually reach the theoretical maximum (McElwain et al., 2016). The theoretical MSC can be derived from leaf anatomy, namely stomatal density, maximum stomatal pore area and stomatal pore depth (McElwain et al., 2016; Murray et al., 2020). The SSMV optimised values of MSC were found to be positively related to the long-term mean VPD (Figure 8), which decreases with elevation, whereas the stomatal conductance as well as stomatal characteristics specifying MSC usually increase with elevation (Bresson et al., 2011; Petrik et al., 2022). In line with this, the studies from temperate European ecosystems reported an inverse relationship between the stomatal conductance and VPD (Körner, 1995; Urban et al., 2017). Similarly, the site-specific values of another ecophysiological parameter representing the stomata closure (VPDC, i.e. the vapour pressure deficit causing the complete conductance reduction) were found to be significantly negatively related to elevation (Figure 8), while the empirical studies did not reveal any differences in the onset of stomatal closure along an elevational gradient (Körner and Cochrane, 1985). Hence, we assume that the revealed correlations in our data are by-products of the site-specific calibrations.

Due to the above above-stated inconsistencies, and the lack of data and supporting information, we decided not to apply the dynamically changing CLEC, MSC and/or VPDC along environmental gradients. However, this approach may be considered as a potential way forward in a future model development when more scientific knowledge becomes available. Nonetheless, our results pointed out that for simulations at a local level, some parameters may need site-specific values. Such a hybrid approach of using a combination of general and site-specific parameters, which was already applied by e.g. ~~(~~Thornton et al.~~,~~

(2002), may be beneficial to reduce the uncertainty of local predictions. Since the values of many of the parameters are usually

750 not measured at research plots, global trait databases, such as TRY (Kattge et al., 2020) or the ones by Liu et al. (2023); Maire et al. (2015); Lin et al. (2015), might be useful to estimate the local values for a specific site and species considering site-specific environmental conditions. Naturally, the best solution for any local study is to obtain measurements of required parameters from specific sites, which is however not always feasible due to time and financial restrictions.

4.4. Robustness and plausibility tests of simulated outputs at a large scale

4.4.1 Carbon stock in soil

Soil carbon represents a large storage of terrestrial carbon (Amundson, 2001) accounting for approximately a half of total forest ecosystem carbon (Domke et al., 2017; Jobbágy and Jackson, 2000). A similar proportion was also revealed in the output of our simulations (median= 47.7 %, mean=51.5 %, 1st Qu.= 32.6%, 3rd Qu.=69.9 %). The absolute values of simulated

760 carbon stock in soil per unit area occurred within the range of soil organic carbon (SOC) reported by empirical studies (De Vos and Cools, 2011; 2010; Wellbrock et al., 2016; Wellbrock and Bolte, 2019), although the variability of simulated values was lower (Fig. S13). The mean value of the simulated SoilC (Min.=7.9, 1st Qu.=12.1, Median = 13.3, Mean = 13.1, 3rd Qu. = 14.4, Max. =17.0 kgC m^{-2}) was similar to the mean values observed in European beech forest stands (e.g. Meier and Leuschner, 2010; Mund, 2004).

765 The site-specific simulated SoilC significantly decreased with the increasing air temperature (Figure 11), which is consistent with the observed patterns in soil carbon stocks from soil profile data along temperature gradients (Hartley et al., 2021; Jobbágy and Jackson, 2000; Post et al., 1982; Sun et al., 2019; Wang et al., 2013). The impact of temperature was also apparent in the relationships between the simulated SoilC and elevation or latitude, both of which were significantly positive (Fig. S17). The increasing trend of SoilC with latitude matched the trend of soil carbon stock found in temperate regions of the northern

770 hemisphere (Minasny et al., 2014). These trends result from the faster decomposition (Wang et al., 2013) and hence the microbial soil respiration as the temperature increases (Cao et al., 2019; Rodeghiero and Cescatti, 2005; Sun et al., 2019), the pattern that was found significant (r=0.33, t = 3.15, p-value = 0.002, 95 % CI = 0.12, 0.51) also in our simulated output (Fig. A18Fig. S19a).

Unlike the increasing trend of SOC with the increasing precipitation reported in the literature (Jobbágy and Jackson, 2000;

775 Post et al., 1982), the simulated SoilC was not significantly related to the precipitation amount (Fig. S19Fig. A18b). In general, the impact of precipitation on SOC changes depending on whether the examined ecosystems are water-limited (Wiesmeier et al., 2019). The small-scale study of SOC in beech forests in Germany revealed its significant correlation to precipitation (Meier and Leuschner, 2010), while at high-latitude ecosystems precipitation has only a minor impact on SOC stock (Devos et al., 2022). Our study includes a much wider variety of environmental conditions including different temperature ranges, soil

780 depths, and soil textures than the study of Meier and Leuschner (2010), which may mask the relationship between SoilC and

precipitation. Unfortunately, we could not derive the regional relationships between measured SOC and environmental characteristics from our dataset due to the lack of data on soil carbon stock at all plots. Hence, we performed only plausibility tests with modelled values and compared the revealed trends with those reported in published papers from elsewhere.

When we checked the relationship of simulated SoilC to soil characteristics, we found the opposite trend of SoilC with the
785 increasing clay content (Figure 11) to the one reported in the literature based on soil measurements (Hartley et al., 2021; Jobbágy and Jackson, 2000). The fine mineral fraction composed of medium to fine silt particles and clay is known to have a stabilisation effect on SOC (Hartley et al., 2021), due to which it is often used as an indicator for SOC storage (Wiesmeier et al., 2019). However, our model results showed that SoilC decreased with the increasing content of fine particles (clay or silt) and increased as sand fraction dominated (Figure 11). Under real conditions, higher clay content supports the formation of soil
790 aggregates that can save organic matter from decomposers and sequester SOC (Angst et al., 2018; Schmidt et al., 2011). In BBGCMuSo this mechanism is not accounted for, as SOC formation is driven solely by temperature and SWC, and the litter input (Hidy et al., 2022). Moreover, the data used for our model simulations did not include the full range of clay content, since the maximum in our database was 56 %, and most site-specific values did not exceed 30 % (median=20 %, mean=19.6 %, 3rd quartile = 22 %). When we experimentally increased the clay content in soils of some sites to the maximum value (i.e.
795 clay content =100 %), we could see the reversed pattern in the relationship (~~Fig. A18~~Fig. S19c).

Soil acidity enhances the storage of SOC by reducing soil microbial activities driving the decomposition of soil organic matter (Funakawa et al., 2014). The new BBGCMuSo model includes soil pH as a factor affecting the process of nitrification in soil layers (Hidy et al., 2022). The observed decreasing trend in the simulated output of SoilC with the increasing pH (~~Fig. A18~~Fig. S19d) is consistent with the experimental results (Funakawa et al., 2014) confirming the correct implementation of pH impact
800 on soil processes in the model.

Soil carbon is the result of carbon inputs from vegetation followed by decomposition processes. In the model, decomposition is driven by soil temperature and soil water content, which is dependent on the precipitation amount, and water infiltration driven by soil texture (Hidy et al., 2022). In nature, around 60 % of the variability in soil respiration is explained by soil temperature and precipitation (Čater and Ogrinc, 2011). Our simulated output also showed a significantly increasing trend of
805 heterotrophic soil respiration with the increasing temperature (Fig. S19~~Fig. A18~~g), but the trends with the soil water content or _precipitation were insignificant (~~Fig. A18~~Fig. S19h,i). Nevertheless, the increase of the simulated soil water content with the increasing fraction of clay and the decreasing fraction of sand in soil (Fig. S20) was consistent with the general knowledge about the impact of soil texture on soil moisture (Kaufmann and Cleveland, 2008). However, the simulated soil microbial respiration was found to have an increasing though insignificant trend with the increasing clay proportion (~~Fig. A18~~Fig. S19e),
810 and a significant decreasing trend with the sand proportion (~~Fig. A18~~Fig. S19f). Although these results explain the negative correlation between SoilC and clay content, they contradict our expectations based on the evidence from empirical studies that suggest that decomposition should be faster in coarse-sized soils (Hartley et al., 2021). In addition, soil respiration is strongly driven by root biomass (Čater and Ogrinc, 2011), which was also detected in our simulations (~~Fig. A18~~Fig. S19j, significant

Pearson's product-moment correlation between the carbon stock in fine roots and heterotrophic respiration with r=0.53, t = 5.74, p-value = 1.45e-07, 95 % CI = (0.36, 0.67)). These findings suggest that while the impact of temperature and vegetation on decomposition is captured well in the model well, the influence of soil water seems to be insufficient. Without a thorough data-based analysis it is however not possible to state if the reason lies in the missing process description in the model or in the values of decomposition-related parameters. Nevertheless, the last methodological paper presenting Biome-BGCMuSo (Hidy et al., 2022) also identified decomposition as a process requiring further development.

Similarly to the reported positive relationship of soil carbon to organic carbon input (Cao et al., 2019; Jobbágy and Jackson, 2000), our outputs showed that SoilC increased with the increasing vegetation carbon stock (Fig. S19Fig. A18k), although the correlation was not significant (Pearson's product-moment correlation r=0.15, t = 1.43, p-value = 0.16, 95 % CI= (-0.06, 0.36)). In the model direct carbon inputs into soil storage come from the litter (Hidy et al., 2022). The significant positive correlation (r=0.89, t=16.79, p-value < 2.2e-16, 95 % CI= (0.82, 0.92)) between the simulated litter and soil carbon stocks (Fig. S19Fig. A18l), confirmed that the model captured the carbon flow from vegetation to the soil according to expectations based on published field data (Hilli et al., 2010). In the model, litter is formed by leaf fall, fine root mortality, and defragmentation of coarse woody debris (CWD, Hidy et al., 2022). Surprisingly, the relationships of simulated SoilC to the annual amount of carbon in leaves or fine roots were insignificant (Fig. S19Fig. A18m,n), while the correlation of SoilC with CWD was significantly positive (Fig. S19Fig. A18o). These results were caused by a much higher amount of accumulated CWD in the simulated ecosystem in comparison to the input from leaves or fine roots. Hence CWD represents the main source of carbon for soil. In the current model version, the actual amount of CWD results from the accumulation over the whole simulation that cannot be reduced by a user, although the usual practice in managed forests has been to remove dead trees during logging operations for sanitary reasons (Kirby et al., 1998; Paletto et al., 2012). Hence, the actual amount of CWD found in managed forests normally represents only a small fraction of the CWD stock of what can be found in nature reserves (Christensen et al., 2005). To make the ecosystem development under the human influence more realistic, the future model version will include the possibility to simulate the extraction of CWD or its part from the system at any time during the normal run simulation.

4.4.2 Carbon stock in litter

The absolute values of simulated carbon stock in litter (Min.= 0.33, 1st Qu.= 0.53, Median= 0.59, Mean = 0.61, 3rd Qu. = 0.68, Max. = 1.07 kgC m^{-2}) were consistent with the litter carbon stock reported from beech forests in Europe (Meier and Leuschner, 2010; Mund, 2004; Vesterdal et al., 2008). The litter amount represented approximately 2.5 % of the total ecosystem carbon (Min.= 0.6, 1st Qu.= 1.4, Median = 2.2, Mean= 2.5, 3rd Qu.= 3.7, Max.= 9.8 %), which is lower than the relative mean litter C stock reported globally (5 % based on Pan et al. (2011)) or for the U.S.A. (7 %, Domke et al., 2016). Such a relatively small amount of organic litter is typical for temperate hardwood forests on fertile soils (BMELF, 1997).

We revealed similar trends of LitterC along environmental gradients as for SoilC, e.g. the simulated carbon stock in litter significantly decreased with the increasing temperature since the heterotrophic respiration also increases along the temperature

gradient (Sun et al., 2019). Similar trends were also found with elevation, and latitude, as well as with soil characteristics (Fig. S17 and Fig. S18~~Fig. A17~~). The decreasing trends of simulated LitterC with the increasing pH, clay and silt proportion in soil and with the decreasing content of coarse sand were consistent with the trends derived from field measurements (Vesterdal

850 and Raulund-Rasmussen, 1998). Based on the empirical evidence by Meier and Leuschner (2010), fine root biomass is the major factor affecting carbon stock in litter. However, our analysis did not reveal a significant relationship between carbon stock in fine roots and litter (Fig. S19~~Fig. A18~~p) probably due to the differences in the perception of the term "litter" in the model and in empirical studies. Vesterdal and Raulund-Rasmussen (1998) also found significant correlations of LitterC to the soil content of other chemical elements (P, Ca, K, Mg), which are not included in our model and in most available models of

855 vegetation dynamics (Merganičová et al., 2019). Nevertheless, due to the ongoing climate change including the dynamics of other nutrients in models may become more important especially if they represent limiting factors for ecosystems (Zaehle, 2013).

4.4.3 Carbon stock in aboveground wood biomass

860 The accumulated carbon stock in wood biomass strongly depends on the forest age or the forest developmental phase. Due to this we first compared the absolute values of AbgwC at the end of spinup simulations to the stock observed in over-aged and old-growth beech forests (Barna et al., 2011). The absolute values occurred within the reported range, although the variability of simulated values was substantially lower than in the observed ones (Fig. S13). On average, around 40 % of total ecosystem carbon was fixed in simulated AbgwC (1st Qu.=21.6, Median=43.5, Mean=39.9, 3rd Qu.=57.9, Max.=78.7 %), similarly as

865 reported from temperate European forest ecosystems (Wellbrock et al., 2017).

Simulated values of AbgwC exhibited a parabolic relationship to elevation and temperature, with maximum values at elevations of around 500 - 600 m a.s.l., and a mean annual air temperature of approximately 9°C (Figure 11). These results coincide with optimum growth conditions for beech reported in the literature, based on which beech growth optimum occurs between 450 and 900 m a.s.l. (Schieber et al., 2013), and at sites with a mean annual temperature of 7 to 10°C (Czajkowski et

870 al., 2006; Pagan, 1996; Paule, 1995). The literature also suggests the optimum annual precipitation total for beech from 700 to 1,000 mm (Czajkowski et al., 2006; Paule, 1995), but the relationship of the simulated AbgwC with precipitation explained only 5 % of AbgwC variability. Nevertheless, the unimodal relationship of AbgwC with VPD revealed the maximum AbgwC at around 530 Pa of VPD (Figure 11), which falls within the optimum VPD range for plant growth (500 to 1,200 Pa~~,~~) (Noh and Lee, 2022). Although empirical studies reported an inverse relationship of beech production to VPD (Lendzion and

875 Leuschner, 2008; Roibu et al., 2022; Tumajer et al., 2022), they focus on short-term changes, whereas in our analysis we used a long-term mean VPD characterising overall site conditions. Already Leuschner (2002) showed in his experiment that the prevailing VPD during the plant development determines the growth potential of plants under the conditions of Central Europe. The laboratory experiment by Lihavainen et al. (2016) revealed that the effect of VPD changes in time. While the initial

reduction of VPD to low values caused an acceleration in the growth rate of silver birch, the effect diminished in time due to nitrogen limitation (Lihavainen et al., 2016). Since VPD seems to have a more profound effect on an intra-annual growth of broadleaved tree species than temperature (Tumajer et al., 2022), more research is required to clarify the impacts at different temporal levels.

The simulated AbgwC trend along the soil gradient was consistent with the empirical knowledge about optimum soil conditions for beech. Beech prefers well-drained soils and does not tolerate wet clay soils. It frequently occurs on loamy or sandy-loamy soils (Packham et al., 2012), on which the simulated AbgwC was the highest (Fig. S20). Loams are the most productive soils because of their moderate soil texture due to which they are able to hold a large amount of water available for plants (Kolb, 2022). Soil texture also affects fine root production, e.g. Weemstra et al. (2017) observed significantly higher fine root biomass on sand than on clay. We did not reveal such a trend in our simulation outputs (Fig. S19Fig. A18q, Pearson's product-moment correlation r=0.11, t = 1.0, df = 82, p-value = 0.32, 95 % CI= (-0.11, 0.32)) because C allocation in the model is fixed and does not depend on soil texture. Although beech forests grow on soils with a large range of pH from 3.5 to 8.5 (Packham et al., 2012), the optimum values at which the maximum biomass production is achieved fluctuate between 5.5. and 6 (Pagan, 1996). The pH of the plots in our database varied between 3.69 and 7.5 (1st Qu.= 5.1, Median = 5.2, Mean = 5.6, 3rd Qu. = 6.4), but we did not reveal a significant trend in AbgwC with pH in the simulated output (Fig. S19Fig. A18r).

4.5. Future model development

Model structural uncertainty and parameter uncertainty are not distinguishable. Inevitable structural uncertainties exist in Biome-BGCMuSo and essentially in all other process-based models, which means that the processes are simplified, and some internal processes can compensate each other. We typically call this phenomenon as getting good results for wrong reasons.

The variability in output data along the gradients of individual characteristics indicates the complex nature of the model and the combined impact of multiple environmental and ecosystem conditions on the final state of the system. In general, we can say that the model output behaves according to well-known natural rules along environmental gradients. The revealed discrepancies are of lesser importance and point out at the issues in the model that should be dealt with in the future model development. Water seems to play a minor role in modifying the simulated carbon-related output, but this requires more thorough tests using data capturing the water cycle that were not used in the current study. Similarly, Tthe impact of soil texture might needs to be examined in more detail to drive the conclusion. Moreover, there are environmental characteristics which are not accounted for in the current model but may explain the differences between the observed and modelled trends in soil carbon stock, e.g. parent rock material (Wiesmeier et al., 2019), or the proportion of coarse rock fragments in soil, which that may substantially influence soil properties, such as water holding capacity and movement, plant growth and decomposition processes (Poesen and Lavee, 1994). In the current study we have not addressed these issues due to the lack of field observations.

Carbon cycling in simulated forest ecosystems is primarily driven by allocation, respiration, mortality, and decomposition.

The relationships between individual output variables representing the carbon cycle are in general consistent with the empirically based knowledge. Including the possibility of CWD removal from an ecosystem due to management interventions will enable more realistic simulations of managed forests and should also result in better capturing the relationships between SoilC and carbon in foliage or fine roots.

Although in the model, a forest is represented by two leaves, one sunlit and one shaded one, this has implications only on the calculations of photosynthesis, while the other processes are not separated between the two parts. Due to this, the parameter called the ratio of shaded specific leaf area (SLA) to sunlit SLA did not currently have a substantial effect on the examined carbon stocks, especially in aboveground wood (

Table 1). Future model development could account for the differences in respiration and allocation proportions and mortality of overstorey and understorey. This would enhance the model applicability to simulate the development of two-storeyed forests and should also increase the variability of model output due to the differences in the growth efficiency between the forest storeys.

Another limitation is the fixed C allocation over the whole simulated period driven by species-specific C allocation parameters. This approach is the simplest one (Merganičová et al., 2019) and was found sufficient when simulating short-term dynamics of ecosystems. However, for multi-site simulations covering long-term dynamics, a fixed C allocation may lead to bias in model output at certain sites or during certain developmental phases of forests, which may require site-specific or phase-specific parameter values.

Other structural improvements needed in Biome-BGCMuSo include improved N cycle and consideration of additional SOC decomposition mass flows including root exudates, priming and litter decomposition to avoid the bias in the estimated parameters. It is a major challenge to the community, and it is not foreseen that the parameter estimation will ever be free from errors.

5. Conclusion

The is work presents presented a novel multi-objective calibration approach that uses the generalised likelihood uncertainty estimation method, plausibility checks of output variables, and intersection principles. The proposed multi-objective calibration approach solves the problems of model overfitting, calibration efficiency, spatial heterogeneity, and the quantity amount and quality of available data. The sensitivity results highlighted the need for the multivariate approach, as the impact rates of parameters and the trends of changes differed between the selected output variables. The integration of the plausibility checks of model outputs ensures the realism of the simulated ecosystemsdynamics. The most important advantage of the presented method is the fact that it considers the environmental dependency of model ecophysiological parameters in a spatial context. Moreover, the approach is also usable to select site (environment) invariant parameters, that are globally applicable.

~~Another advantage compared to traditional Bayesian or frequentist methods is the plausibility check of the optimised parameters and their ranges, for~~where global databases on plant traits play a crucial role.~~ The solution improves the reliability of the optimisation and may be generally applicable to any process-based model~~s~~ of ecosystem dynamics. ~~Such models heavily rely on how they characterize the underlying ecosystem. The process is crucial for every modelling application, and it is considered as one of the main challenges, because of its vast computational, and data need.~~

~~The method includes the multi-dimensional sensitivity analysis used as a basis for the selection of parameters for calibration. The sensitivity results highlighted the need for the multivariate approach, as the impact rates of parameters and the trends of changes differed between the selected output variables. In addition, the analysis also documented the importance of plausibility range checks for parameters, for which global databases on plant traits, such as TRY, play a crucial role.~~

The disadvantage of the optimisation method is the possible bias of optimised model parameters that can occur because the parameters are forced to~~wards~~ specific values during the optimisation process to match the simulated output with observations. This can be partially avoided by including multiple data into calibration, which represent diverse parts of an ecosystem, such as vegetation, litter, and soil, as it was presented in this study, simulated nutrients (in our case it would be carbon, nitrogen, and water), and processes. To identify the bias in parameter values, on-site measurements of parameters would be needed. Hence, ~~from this point of view,~~ it is worth considering obtaining the information on some plant characteristics, e.g. C:N ratios in different ecosystem compartments, FLNR, etc., routinely from research plots.

The presented optimisation method can be further enhanced. In this study the likelihood function did not include the uncertainty of the observations, which means a lack of weighting of errors ~~due~~with respect to their magnitudes. Thus, the likelihood can be reformulated to include the observation uncertainty. Moreover, the method does not currently account for the covariance between output variables. This could be done by ~~creating~~constructing a covariance matrix representing the relationships between the output variables from the model simulations and incorporating it into a multivariate likelihood function. Such an approach ~~sh~~could provide a more accurate and realistic estimation of the uncertainties associated with ~~the~~ model parameters. On the other hand, including covariance w~~ill~~ould further increase the computational complexity of the method, which is already ~~characterized by its significant demands~~high.

The calibration of the model performed at individual sites (SSMV) and multiple sites (MSMV) revealed pros and cons of both approaches. Site specific parameter values improved the accuracy of the simulated outputs of interest~~aboveground wood, soil, and litter carbons stocks~~ for the specific sites ~~by 35 %, 55 %, and 11 % in comparison to the a priori parameter set. However, when using them for different locations, we observed the collapse of the simulations in 47 % of the cases. When applying the MSMV parameter values obtained from the multi-site calibration, all simulations were successful, but the errors were on average by 26 %, 35 % and 9 % greater for the aboveground wood, soil, and litter carbons stocks than those obtained with the SSMV parameter sets. T~~ and t~~hus,~~ are more suitable for local simulation studies~~, site-specific values of the parameters may be more suitable~~ than the generalised parameter set, which is more appropriate for studies covering a larger spatial scale. ~~On the other hand, the multi-site calibration reduces the spatial heterogeneity due to the "averaging out effect" causing that the average~~

~~model performance will be good, but at specific sites it can be completely wrong. This could be overcome by applying a hybrid approach that combines the site-specific and multisite calibration.~~

The independent validation, robustness and plausibility tests confirmed the robustness of the multi-site and multivariate calibrated set of ecophysiological parameters for the European beech at a regional level. ~~In general, the examined model output followed well-known natural rules along environmental gradients. The revealed discrepancies were of lesser importance and pointed out at the issues in the model structure that can be improved in the future model development~~. The study ~~also~~ highlighted the gaps in the empirical data and knowledge explaining the relationships between parameters, or between parameters and environmental conditions, which should be addressed by future research. For future applications, additional parameters that were not considered in this study, such as parameters specifying drought-induced mortality, may need to be calibrated with additional empirical data since the occurrence of extreme events and disturbances has been increasing due to the climate change.

Code and data availability. The current version of Biome-BGCMuSo, together with sample input files and a detailed user guide, is available from the website of the model at http://nimbus.elte.hu/bbgc/download.html under the GPL-2 license. Biome-BGCMuSo v6 is also available at GitHub: https://github.com/bpbond/Biome-BGC/tree/Biome-BGCMuSo_v6. The exact version of the model (v6.2 alpha) used to produce the results in this paper is archived on Zenodo (https://doi.org/10.5281/zenodo.5761202; Hidy and Barcza, 2021). The RBBGCMuSo package (Hollós et al., 2023) is available at https://github.com/hollorol/RBBGCMuso. Experimental data used in the study are available from ICP Forests (http://icp-forests.net/) and from authors that provided the data upon request. The code for optimisation of parameters is in the supplement.

Supplement. The supplement related to this article is available online.

Author contributions. KM and TH conceived and designed the study and KM carried the experiments with the assistance from JM, LD, and RH. DH and ZB developed Biome-BGCMuSo, and maintained the source code. RH and ZB developed RBBGCMuSo. KM and DK performed the literature review. KM executed simulations, and KM and JM performed the analyses. LD and JM contributed with model benchmarking. ZS, PP, HM, JN, MB, DB, MZOS, and PF contributed with experimental data. KM prepared the paper and the Supplement with contributions from all co-authors. KM, JM, TH, ZB, RH, MZOS, HM, and DB revised the paper based on the reviews of two anonymous reviewers. All authors reviewed and approved the present article and the Supplement.

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements

The study was supported by COST PROCLIAS CA19139, grant "EVA4.0", No. CZ.02.1.01/0.0/0.0/16_019/0000803 financed by OP RDEEVA4.0, and the Slovak Research and Development Agency, projects No. APVV-21-0412 (KM, PF), APVV-18-0305, APVV-22-0001 (JM), APVV-19-0183 (MB). HM, MZOS, and DB were supported by the Croatian Science Foundation National MODFLUX (HRZZ IP-2019-04-6325). ZB, RH and DH were supported by the National Multidisciplinary Laboratory for Climate Change, RRF-2.3.1-21-2022-00014 project. PP was supported by „TreeAdapt" funded on the base of contract

between the National Forest Centre and Ministry of Agriculture and Rural Development of the Slovak Republic. We thank two anonymous reviewers for their valuable comments that substantially improved the manuscript.

References

1015   Amundson, R.: The Carbon Budget in Soils, Annu. Rev. Earth Planet. Sci., 29, 535–562, https://doi.org/10.1146/annurev.earth.29.1.535, 2001.

Angst, G., Messinger, J., Greiner, M., Häusler, W., Hertel, D., Kirfel, K., Kögel-Knabner, I., Leuschner, C., Rethemeyer, J., and Mueller, C. W.: Soil organic carbon stocks in topsoil and subsoil controlled by parent material, carbon input in the rhizosphere, and microbial-derived compounds, Soil Biol. Biochem., 122, 19–30,
1020   https://doi.org/10.1016/j.soilbio.2018.03.026, 2018.

Barna, M., Kulfan, J., and Bublinec, E.: Beech and Beech Ecosystems of Slovakia / Buk a bukové ekosystémy Slovenska, Veda, Bratislava, 636 pp., 2011.

Beven, K.: A manifesto for the equifinality thesis, J. Hydrol., 320, 18–36, https://doi.org/10.1016/j.jhydrol.2005.07.007, 2006.

1025   Beven, K.: Validation and Equifinality, in: Computer Simulation Validation: Fundamental Concepts, Methodological Frameworks, and Philosophical Perspectives, edited by: Beisbart, C. and Saam, N. J., Springer International Publishing, Cham, 791–809, https://doi.org/10.1007/978-3-319-70766-2_32, 2019.

Beven, K. and Binley, A.: GLUE: 20 years on, Hydrol. Process., 28, 5897–5918, https://doi.org/10.1002/hyp.10082, 2014.

Blyth, E., Clark, D. B., Ellis, R., Huntingford, C., Los, S., Pryor, M., Best, M., and Sitch, S.: A comprehensive set of
1030   benchmark tests for a land surface model of simultaneous fluxes of water and carbon at both the global and seasonal scale, Geosci. Model Dev., 4, 255–269, https://doi.org/10.5194/gmd-4-255-2011, 2011.

BMELF: Deutscher Waldbodenbericht 1996, Bundesministerium für Ernährung, Landwirtschaft und Forsten, Bonn, 1997.

Bresson, C. C., Vitasse, Y., Kremer, A., and Delzon, S.: To what extent is altitudinal variation of functional traits driven by genetic adaptation in European oak and beech?, Tree Physiol., 31, 1164–1174, https://doi.org/10.1093/treephys/tpr084, 2011.

1035   Brown, M. J. and Parker, G. G.: Canopy light transmittance in a chronosequence of mixed-species deciduous forests, Can. J. For. Res., 24, 1694–1703, https://doi.org/10.1139/x94-219, 1994.

Bugmann, H., Seidl, R., Hartig, F., Bohn, F., Brůna, J., Cailleret, M., François, L., Heinke, J., Henrot, A.-J., Hickler, T., Hülsmann, L., Huth, A., Jacquemin, I., Kollas, C., Lasch-Born, P., Lexer, M. J., Merganič, J., Merganičová, K., Mette, T., Miranda, B. R., Nadal-Sala, D., Rammer, W., Rammig, A., Reineking, B., Roedig, E., Sabaté, S., Steinkamp, J., Suckow, F.,
1040   Vacchiano, G., Wild, J., Xu, C., and Reyer, C. P. O.: Tree mortality submodels drive simulated long-term forest dynamics: assessing 15 models from the stand to global scale, Ecosphere, 10, e02616, https://doi.org/10.1002/ecs2.2616, 2019.

Cameron, D. R., Van Oijen, M., Werner, C., Butterbach-Bahl, K., Grote, R., Haas, E., Heuvelink, G. B. M., Kiese, R., Kros, J., Kuhnert, M., Leip, A., Reinds, G. J., Reuter, H. I., Schelhaas, M. J., De Vries, W., and Yeluripati, J.: Environmental change impacts on the C- and N-cycle of European forests: a model comparison study, Biogeosciences, 10, 1751–1773,
1045   https://doi.org/10.5194/bg-10-1751-2013, 2013.

Cao, B., Domke, G. M., Russell, M. B., and Walters, B. F.: Spatial modeling of litter and soil carbon stocks on forest land in the conterminous United States, Sci. Total Environ., 654, 94–106, https://doi.org/10.1016/j.scitotenv.2018.10.359, 2019.

Čater, M. and Ogrinc, N.: Soil respiration rates and in natural beech forest (Fagus sylvatica L.) in relation to stand structure, Isotopes Environ. Health Stud., 47, 221–237, https://doi.org/10.1080/10256016.2011.578214, 2011.

1050   Caudullo, G., Welk, E., and San-Miguel-Ayanz, J.: Chorological maps for the main European woody species, Data Brief, 12, 662–666, https://doi.org/10.1016/j.dib.2017.05.007, 2017.

Christensen, M., Hahn, K., Mountford, E. P., Ódor, P., Standovár, T., Rozenbergar, D., Diaci, J., Wijdeven, S., Meyer, P., Winter, S., and Vrska, T.: Dead wood in European beech (Fagus sylvatica) forest reserves, For. Ecol. Manag., 210, 267–282, https://doi.org/10.1016/j.foreco.2005.02.032, 2005.

1055   Churkina, G. and Running, S.: Investigating the balance between timber harvest and productivity of global coniferous forests under global change, Clim. Change, 47, 167–191, https://doi.org/10.1023/a:1005620808273, 2000.

Churkina, G., Tenhunen, J., Thornton, P., Falge, E. M., Elbers, J. A., Erhard, M., Grünwald, T., Kowalski, A. S., Rannik, Ü., and Sprinz, D.: Analyzing the ecosystem carbon dynamics of four European coniferous forests using a biogeochemistry model, Ecosystems, 6, 168–184, https://doi.org/10.1007/s10021-002-0197-2, 2003.

1060   Cornes, R. C., van der Schrier, G., van den Besselaar, E. J. M., and Jones, P. D.: An Ensemble Version of the E-OBS Temperature and Precipitation Data Sets, J. Geophys. Res. Atmospheres, 123, 9391–9409, https://doi.org/10.1029/2017JD028200, 2018.

Czajkowski, T., Kompa, T., and Bolte, A.: Zur Verbreitungsgrenze der Buche (Fagus sylvatica L.) im nordöstlichen Mitteleuropa (The distribution boundary of European beech (Fagus sylvatica L.) in north-eastern Europe), Forstarchiv, 77,
1065   203–216, 2006.

De Vos, B. and Cools, N.: Second European Forest Soil Condition Report, Research Institute for Nature and Forest, Geraardsbergen, 2011.

Devi, A. S.: Influence of trees and associated variables on soil organic carbon: a review, J. Ecol. Environ., 45, 5, https://doi.org/10.1186/s41610-021-00180-3, 2021.

1070   Devos, C. C., Ohlson, M., Næsset, E., and Bollandsås, O. M.: Soil carbon stocks in forest-tundra ecotones along a 500 km latitudinal gradient in northern Norway, Sci. Rep., 12, 13358, https://doi.org/10.1038/s41598-022-17409-3, 2022.

Domke, G. M., Perry, C. H., Walters, B. F., Woodall, C. W., Russell, M. B., and Smith, J. E.: Estimating litter carbon stocks on forest land in the United States, Sci. Total Environ., 557–558, 469–478, https://doi.org/10.1016/j.scitotenv.2016.03.090, 2016.

1075   Domke, G. M., Perry, C. H., Walters, B. F., Nave, L. E., Woodall, C. W., and Swanston, C. W.: Toward inventory-based estimates of soil organic carbon in forests of the United States, Ecol. Appl., 27, 1223–1235, https://doi.org/10.1002/eap.1516, 2017.

CORINE Land Cover: https://land.copernicus.eu/en/products/corine-land-cover, last access: 17 November 2023.

FAO, - Food and Agriculture Organization of the United Nations: Harmonized World Soil Database v 1.2, 2012.

1080   Fer, I., Kelly, R., Moorcroft, P. R., Richardson, A. D., Cowdery, E. M., and Dietze, M. C.: Linking big models to big data: Efficient ecosystem model calibration through Bayesian model emulation, Biogeosciences, 15, 5801–5830, https://doi.org/10.5194/bg-15-5801-2018, 2018.

Forrester, D. I., Hobi, M. L., Mathys, A. S., Stadelmann, G., and Trotsiuk, V.: Calibration of the process-based model 3-PG for major central European tree species, Eur. J. For. Res., 140, 847–868, https://doi.org/10.1007/s10342-021-01370-3, 2021.

1085   Fox, J., Weisberg, S., Price, B., Adler, D., Bates, D., Baud-Bovy, G., Bolker, B., Ellison, S., Firth, D., Friendly, M., Gorjanc, G., Graves, S., Heiberger, R., Krivitsky, P., Laboissiere, R., Maechler, M., Monette, G., Murdoch, D., Nilsson, H., Ogle, D., Ripley, B., Short, T., Venables, W., Walker, S., Winsemius, D., Zeileis, A., and R-Core: car: Companion to Applied Regression, 2023.

Funakawa, S., Fujii, K., Kadono, A., Watanabe, T., and Kosaki, T.: Could Soil Acidity Enhance Sequestration of Organic
1090   Carbon in Soils?, in: Soil Carbon, edited by: Hartemink, A. E. and McSweeney, K., Springer International Publishing, Cham, 209–216, https://doi.org/10.1007/978-3-319-04084-4_22, 2014.

Georgi, L., Kunz, M., Fichtner, A., Härdtle, W., Reich, K. F., Sturm, K., Welle, T., and Oheimb, G. von: Long-Term Abandonment of Forest Management Has a Strong Impact on Tree Morphology and Wood Volume Allocation Pattern of European Beech (Fagus sylvatica L.), Forests, 9, 704, https://doi.org/10.3390/f9110704, 2018.

1095   Gratani, L.: Plant Phenotypic Plasticity in Response to Environmental Factors, Adv. Bot., 2014, 208747, https://doi.org/10.1155/2014/208747, 2014.

Hartley, I. P., Hill, T. C., Chadburn, S. E., and Hugelius, G.: Temperature effects on carbon storage are controlled by soil stabilisation capacities, Nat. Commun., 12, 6713, https://doi.org/10.1038/s41467-021-27101-1, 2021.

Hengl, T., Leal Parente, L., Krizan, J., and Bonannella, C.: Continental Europe Digital Terrain Model at 30 m resolution
1100   based on GEDI, ICESat-2, AW3D, GLO-30, EUDEM, MERIT DEM and background layers (v0.3), https://doi.org/10.5281/zenodo.4724549, 2020.

Hidy, D., Barcza, Z., Haszpra, L., Churkina, G., Pintér, K., and Nagy, Z.: Development of the Biome-BGC model for simulation of managed herbaceous ecosystems, Ecol. Model., 226, 99–119, https://doi.org/10.1016/j.ecolmodel.2011.11.008, 2012.

1105   Hidy, D., Barcza, Z., Marjanović, H., Ostrogović Sever, M. Z., Dobor, L., Gelybó, Gy., Fodor, N., Pintér, K., Churkina, G., Running, S. W., Thornton, P. E., Bellocchi, G., Haszpr, L., Horváth, F., Suyker, A., Verma, S., and Nagy, Z.: Terrestrial Ecosystem Process Model Biome-BGC MuSo: Summary of improvements and new modeling possibilities, Geosci. Model Dev., 9, 4405–4437, https://doi.org/10.5194/gmd-2016-93, 2016.

Hidy, D., Barcza, Z., Hollós, R., Thornton, P. E., Running, S. W., and Fodor, N.: User's Guide for Biome-BGCMuSo 6.2,
1110   2021.

Hidy, D., Barcza, Z., Hollós, R., Dobor, L., Ács, T., Zacháry, D., Filep, T., Pásztor, L., Incze, D., Dencső, M., Tóth, E., Merganičová, K., Thornton, P., Running, S., and Fodor, N.: Soil-related developments of the Biome-BGCMuSo v6.2 terrestrial ecosystem model, Geosci. Model Dev., 15, 2157–2181, https://doi.org/10.5194/gmd-15-2157-2022, 2022.

Hilli, S., Stark, S., and Derome, J.: Litter decomposition rates in relation to litter stocks in boreal coniferous forests along
1115   climatic and soil fertility gradients, Appl. Soil Ecol., 46, 200–208, https://doi.org/10.1016/j.apsoil.2010.08.012, 2010.

Hlásny, T., Barcza, Z., Fabrika, M., Balázs, B., Churkina, G., Pajtík, J., Sedmák, R., and Turcáni, M.: Climate change impacts on growth and carbon -balance of forests in Central Europe, Clim. Res., 47, 219–236, https://doi.org/10.3354/cr01024, 2011.

Hlásny, T., Barcza, Z., Barka, I., Merganičová, K., Sedmák, R., Kern, A., Pajtík, J., Balázs, B., Fabrika, M., and Churkina, G.: Future carbon cycle in mountain spruce forests of Central Europe: Modelling framework and ecological inferences, For. Ecol. Manag., 328, 55–68, https://doi.org/10.1016/j.foreco.2014.04.038, 2014.

Hoffman, E. and Gardner, R.: Evaluation of Uncertainties in Environmental Radiological Assessment Models., in: Radiological Assessments: a Textbook on Environmental Dose Analysis., Nuclear Regulatory Commission, Washington, D.C, USA, 11.1-11.51, 1983.

Hofmann, M., Gatu, C., Kontoghiorghes, E. J., Colubi, A., and Zeileis, A.: lmSubsets: Exact Variable-Subset Selection in Linear Regression, J Stat Softw, 1–21, 2021.

Hollós, R., Fodor, N., Merganičová, K., Hidy, D., Árendás, T., Grünwald, T., and Barcza, Z.: Conditional interval reduction method: A possible new direction for the optimization of process based models, Environ. Model. Softw., 158, 105556, https://doi.org/10.1016/j.envsoft.2022.105556, 2022.

Hollós, R., Kristóf, E., Fodor, N., Hidy, D., Horváth, F., and Barcza, Z.: RBBGCMuso: an R package to support the application of the Biome-BGCMuSo biogeochemical mode., 2023.

Hülsmann, L., Bugmann, H., Meyer, P., and Brang, P.: Natürliche Baummortalität in Mitteleuropa: Mortalitätsraten und -muster im Vergleich, Schweiz. Z. Forstwes., 169, 166–174, https://doi.org/10.3188/szf.2018.0166, 2018.

Hungerford, R. D., Nemani, R. R., Running, S. W., and Coughlan, J. C.: MTCLIM: a mountain microclimate simulation model, U.S. Department of Agriculture, Forest Service, Intermountain Forest and Range Experiment Station, Ogden, UT, https://doi.org/10.2737/INT-RP-414, 1989.

ICPP: Good Practice Guidance for Land Use, Land-Use Change and Forestry, 2003.

Jager, H. I., Hargrove, W. W., Brandt, C. C., King, A. W., Olson, R. J., Scurlock, J. M. O., and Rose, K. A.: Constructive contrasts between modeled and measured climate responses over a regional scale, Ecosystems, 3, 396–411, https://doi.org/10.1007/s100210000035, 2000.

Jarčuška, B. and Barna, M.: Plasticity in above-ground biomass allocation in Fagus sylvatica L. saplings in response to light availability, Ann. For. Res., 54, 151–160, https://doi.org/10.15287/afr.2011.86, 2011.

Jobbágy, E. G. and Jackson, R. B.: The Vertical Distribution of Soil Organic Carbon and Its Relation to Climate and Vegetation, Ecol. Appl., 10, 423–436, https://doi.org/10.1890/1051-0761(2000)010[0423:TVDOSO]2.0.CO;2, 2000.

Kamali, B., Stella, T., Berg-Mohnicke, M., Pickert, J., Groh, J., and Nendel, C.: Improving the simulation of permanent grasslands across Germany by using multi-objective uncertainty-based calibration of plant-water dynamics, Eur. J. Agron., 134, 126464, https://doi.org/10.1016/j.eja.2022.126464, 2022.

Kattge, J., Bönisch, G., Díaz, S., Lavorel, S., Prentice, I. C., Leadley, P., Tautenhahn, S., Werner, G. D. A., Aakala, T., Abedi, M., Acosta, A. T. R., Adamidis, G. C., Adamson, K., Aiba, M., Albert, C. H., Alcántara, J. M., Alcázar C, C., Aleixo, I., Ali, H., Amiaud, B., Ammer, C., Amoroso, M. M., Anand, M., Anderson, C., Anten, N., Antos, J., Apgaua, D. M. G., Ashman, T. L., Asmara, D. H., Asner, G. P., Aspinwall, M., Atkin, O., Aubin, I., Baastrup-Spohr, L., Bahalkeh, K., Bahn,

M., Baker, T., Baker, W. J., Bakker, J. P., Baldocchi, D., Baltzer, J., Banerjee, A., Baranger, A., Barlow, J., Barneche, D. R., Baruch, Z., Bastianelli, D., Battles, J., Bauerle, W., Bauters, M., Bazzato, E., Beckmann, M., Beeckman, H., Beierkuhnlein, C., Bekker, R., Belfry, G., Belluau, M., Beloiu, M., Benavides, R., Benomar, L., Berdugo-Lattke, M. L., Berenguer, E., Bergamin, R., Bergmann, J., Bergmann Carlucci, M., Berner, L., Bernhardt-Römermann, M., Bigler, C., Bjorkman, A. D., Blackman, C., Blanco, C., Blonder, B., Blumenthal, D., Bocanegra-González, K. T., Boeckx, P., Bohlman, S., Böhning-Gaese, K., Boisvert-Marsh, L., Bond, W., Bond-Lamberty, B., Boom, A., Boonman, C. C. F., Bordin, K., Boughton, E. H., Boukili, V., Bowman, D. M. J. S., Bravo, S., Brendel, M. R., Broadley, M. R., Brown, K. A., Bruelheide, H., Brumnich, F., Bruun, H. H., Bruy, D., Buchanan, S. W., Bucher, S. F., Buchmann, N., Buitenwerf, R., Bunker, D. E., et al.: TRY plant trait database – enhanced coverage and open access, Glob. Change Biol., 26, 119–188, https://doi.org/10.1111/gcb.14904, 2020.

Kaufmann, R. K. and Cleveland, C. J.: Environmental Science, McGraw-Hill Higher Education, 596 pp., 2008.

Keeling, C. D., Bacastow, R. B., Bainbridge, A. E., Ekdahl Jr., C. A., Guenther, P. R., Waterman, L. S., and Chin, J. F. S.: Atmospheric carbon dioxide variations at Mauna Loa Observatory, Hawaii, Tellus, 28, 538–551, https://doi.org/10.1111/j.2153-3490.1976.tb00701.x, 1976.

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

Kimball, J. S., White, M. A., and Running, S. W.: BIOME-BGC simulations of stand hydrologic processes for BOREAS, J. Geophys. Res., 102, 29043–29051, https://doi.org/10.1029/97JD02235, 1997.

Kirby, K., Reid, C., Thomas, R., and Goldsmith, F.: Preliminary estimates of fallen dead wood and standing dead trees in managed and unmanaged forests in Britain, J. Appl. Ecol., 35, 148–155, 1998.

Soils and Water Availability – Climate, Forests and Woodlands: https://climate-woodlands.extension.org/soils-and-water-availability/, last access: 8 December 2022.

Körner, Ch.: Leaf Diffusive Conductances in the Major Vegetation Types of the Globe, in: Ecophysiology of Photosynthesis, edited by: Schulze, E.-D. and Caldwell, M. M., Springer, Berlin, Heidelberg, 463–490, https://doi.org/10.1007/978-3-642-79354-7_22, 1995.

Körner, Ch. and Cochrane, P. M.: Stomatal responses and water relations of Eucalyptus pauciflora in summer along an elevational gradient, Oecologia, 66, 443–455, https://doi.org/10.1007/BF00378313, 1985.

Kramer, K., Leinonen, I., Bartelink, H. H., Berbigier, P., Borghetti, M., Bernhofer, C., Cienciala, E., Dolman, A. J., Froer, O., Gracia, C. A., Granier, A., Grünwald, T., Hari, P., Jans, W., Kellomäki, S., Loustau, D., Magnani, F., Markkanen, T., Matteucci, G., Mohren, G. M. J., Moors, E., Nissinen, A., Peltola, H., Sabaté, S., Sanchez, A., Sontag, M., Valentini, R., and Vesala, T.: Evaluation of six process-based forest growth models using eddy-covariance measurements of CO2 and H2O fluxes at six forest sites in Europe, Glob. Change Biol., 8, 213–230, https://doi.org/10.1046/j.1365-2486.2002.00471.x, 2002.

Lavigne, M. and Ryan, M.: Growth and maintenance respiration rates of aspen, black spruce and jack pine stems at northern and southern BOREAS sites, Tree Physiol., 17, 543–551, https://doi.org/10.1093/treephys/17.8-9.543, 1997.

Lendzion, J. and Leuschner, C.: Growth of European beech (Fagus sylvatica L.) saplings is limited by elevated atmospheric vapour pressure deficits, For. Ecol. Manag., 256, 648–655, https://doi.org/10.1016/j.foreco.2008.05.008, 2008.

Leuschner, C.: Air humidity as an ecological factor for woodland herbs: leaf water status, nutrient uptake, leaf anatomy, and productivity of eight species grown at low or high vpd levels, Flora - Morphol. Distrib. Funct. Ecol. Plants, 197, 262–274, https://doi.org/10.1078/0367-2530-00040, 2002.

1190 Levins, R.: The strategy of model building in population biology arises, Am. Sci., 54, 421–431, 1966.

Lihavainen, J., Ahonen, V., Keski-Saari, S., Kontunen-Soppela, S., Oksanen, E., and Keinänen, M.: Low vapour pressure deficit affects nitrogen nutrition and foliar metabolites in silver birch, J. Exp. Bot., 67, 4353–4365, https://doi.org/10.1093/jxb/erw218, 2016.

Lin, Y.-S., Medlyn, B. E., Duursma, R. A., Prentice, I. C., Wang, H., Baig, S., Eamus, D., de Dios, V. R., Mitchell, P.,
1195 Ellsworth, D. S., de Beeck, M. O., Wallin, G., Uddling, J., Tarvainen, L., Linderson, M.-L., Cernusak, L. A., Nippert, J. B., Ocheltree, T. W., Tissue, D. T., Martin-StPaul, N. K., Rogers, A., Warren, J. M., De Angelis, P., Hikosaka, K., Han, Q., Onoda, Y., Gimeno, T. E., Barton, C. V. M., Bennie, J., Bonal, D., Bosc, A., Löw, M., Macinins-Ng, C., Rey, A., Rowland, L., Setterfield, S. A., Tausz-Posch, S., Zaragoza-Castells, J., Broadmeadow, M. S. J., Drake, J. E., Freeman, M., Ghannoum, O., Hutley, L. B., Kelly, J. W., Kikuzawa, K., Kolari, P., Koyama, K., Limousin, J.-M., Meir, P., Lola da Costa, A. C.,
1200 Mikkelsen, T. N., Salinas, N., Sun, W., and Wingate, L.: Optimal stomatal behaviour around the world, Nat. Clim. Change, 5, 459–464, https://doi.org/10.1038/nclimate2550, 2015.

Liu, C., Sack, L., Li, Y., Zhang, J., Yu, K., Zhang, Q., He, N., and Yu, G.: Relationships of stomatal morphology to the environment across plant communities, Nat. Commun., 14, 6629, https://doi.org/10.1038/s41467-023-42136-2, 2023.

Liu, S., Baret, F., Abichou, M., Manceau, L., Andrieu, B., Weiss, M., and Martre, P.: Importance of the description of light
1205 interception in crop growth models, Plant Physiol., 186, 977–997, https://doi.org/10.1093/plphys/kiab113, 2021.

Luo, X.: Global variation in the fraction of leaf nitrogen allocated to photosynthesis, https://doi.org/10.5281/zenodo.5090497, 2021.

Maire, V., Wright, I. J., Prentice, I. C., Batjes, N. H., Bhaskar, R., van Bodegom, P. M., Cornwell, W. K., Ellsworth, D., Niinemets, Ü., Ordonez, A., Reich, P. B., and Santiago, L. S.: Global effects of soil and climate on leaf photosynthetic traits
1210 and rates, Glob. Ecol. Biogeogr., 24, 706–717, https://doi.org/10.1111/geb.12296, 2015.

Maselli, F., Papale, D., Puletti, N., Chirici, G., and Corona, P.: Combining remote sensing and ancillary data to monitor the gross productivity of water-limited forest ecosystems, Remote Sens. Environ., 113, 657–667, https://doi.org/10.1016/j.rse.2008.11.008, 2009.

McElwain, J. C., Yiotis, C., and Lawson, T.: Using modern plant trait relationships between observed and theoretical
1215 maximum stomatal conductance and vein density to examine patterns of plant macroevolution, New Phytol., 209, 94–103, https://doi.org/10.1111/nph.13579, 2016.

Meier, I. C. and Leuschner, C.: Variation of soil and biomass carbon pools in beech forests across a precipitation gradient, Glob. Change Biol., 16, 1035–1045, https://doi.org/10.1111/j.1365-2486.2009.02074.x, 2010.

Merganič, J., Merganičová, K., Konôpka, B., and Kučera, M.: Country and regional carbon stock in forest cover – estimates
1220 based on the first cycle of the Czech National Forest Inventory data (2001–2004), Cent. Eur. For. J., 63, 113–125, https://doi.org/10.1515/forj-2017-0018, 2017.

Merganičová, K. and Merganič, J.: The Effect of Dynamic Mortality Incorporated in BIOME-BGC on Modelling the Development of Natural Forests, J. Environ. Inform., 24, 24–31, 2014.

Merganičová, K., Pietsch, S. A., and Hasenauer, H.: Testing mechanistic modeling to assess impacts of biomass removal, For. Ecol. Manag., 207, 37–57, https://doi.org/10.1016/j.foreco.2004.10.017, 2005.

Merganičová, K., Merganič, J., Lehtonen, A., Vacchiano, G., Sever, M. Z. O., Augustynczik, A. L. D., Grote, R., Kyselová, I., Mäkelä, A., Yousefpour, R., Krejza, J., Collalti, A., and Reyer, C. P. O.: Forest carbon allocation modelling under climate change, Tree Physiol., 39, 1937–1960, https://doi.org/10.1093/treephys/tpz105, 2019.

Michel, K., Prescher, A. K., Seidling, W., and Ferretti, M.: A policy-relevant infrastructure for long-term, large-scale assessment and monitoring of forest ecosystems, Thünen Institute of Forest Ecosystems, Eberswalde, Germany, 2018.

Minasny, B., McBratney, A. B., Malone, B. P., Lacoste, M., and Walter, C.: Quantitatively Predicting Soil Carbon Across Landscapes, in: Soil Carbon, edited by: Hartemink, A. E. and McSweeney, K., Springer International Publishing, Cham, 45–57, https://doi.org/10.1007/978-3-319-04084-4_5, 2014.

Minunno, F., Peltoniemi, M., Härkönen, S., Kalliokoski, T., Makinen, H., and Mäkelä, A.: Bayesian calibration of a carbon balance model PREBAS using data from permanent growth experiments and national forest inventory, For. Ecol. Manag., 440, 208–257, https://doi.org/10.1016/j.foreco.2019.02.041, 2019.

Mund, M.: Carbon pools of European beech forests (Fagus sylvatica) under different silvicultural management, Dissertation thesis., Georg-August-Universität Göttingen, Göttingen, 256 pp., 2004.

Murray, M., Soh, W. K., Yiotis, C., Batke, S., Parnell, A. C., Spicer, R. A., Lawson, T., Caballero, R., Wright, I. J., Purcell, C., and McElwain, J. C.: Convergence in Maximum Stomatal Conductance of C3 Woody Angiosperms in Natural Ecosystems Across Bioclimatic Zones, Front. Plant Sci., 10, 2019.

Murray, M., Soh, W. K., Yiotis, C., Spicer, R. A., Lawson, T., and McElwain, J. C.: Consistent Relationship between Field-Measured Stomatal Conductance and Theoretical Maximum Stomatal Conductance in C3 Woody Angiosperms in Four Major Biomes, Int. J. Plant Sci., 181, 142–154, https://doi.org/10.1086/706260, 2020.

Noh, H. and Lee, J.: The Effect of Vapor Pressure Deficit Regulation on the Growth of Tomato Plants Grown in Different Planting Environments, Appl. Sci., 12, 3667, https://doi.org/10.3390/app12073667, 2022.

van Oijen, M.: Bayesian Methods for Quantifying and Reducing Uncertainty and Error in Forest Models, Curr. For. Rep., 3, 269–280, https://doi.org/10.1007/s40725-017-0069-9, 2017.

Osei, R., del Río, M., Ruiz-Peinado, R., Titeux, H., Bielak, K., Bravo, F., Collet, C., Cools, C., Cornelis, J.-T., Drössler, L., Heym, M., Korboulewsky, N., Löf, M., Muys, B., Najib, Y., Nothdurft, A., Pretzsch, H., Skrzyszewski, J., and Ponette, Q.: The distribution of carbon stocks between tree woody biomass and soil differs between Scots pine and broadleaved species (beech, oak) in European forests, Eur. J. For. Res., 141, 467–480, https://doi.org/10.1007/s10342-022-01453-9, 2022.

Ostrogović Sever, M. Z., Barcza, Z., Hidy, D., Kern, A., Dimoski, D., Miko, S., Hasan, O., Grahovac, B., and Marjanović, H.: Evaluation of the Terrestrial Ecosystem Model Biome-BGCMuSo for Modelling Soil Organic Carbon under Different Land Uses, Land, 10, 968, https://doi.org/10.3390/land10090968, 2021.

Packham, J. R., Thomas, P. A., Atkinson, M. D., and Degen, T.: Biological Flora of the British Isles: Fagus sylvatica, J. Ecol., 100, 1557–1608, https://doi.org/10.1111/j.1365-2745.2012.02017.x, 2012.

Pagan, J.: Lesnícka dendrológia., 2. vyd., Technická univerzita Zvolen, Zvolen, 1996.

Pajtík, J., Čihák, T., Konôpka, B., Merganičová, K., and Fabiánek, P.: Annual tree mortality and felling rates in the Czech Republic and Slovakia over three decades, Cent. Eur. For. J., 64, 238–248, https://doi.org/10.1515/forj-2017-0048, 2018.

Paletto, A., Ferretti, F., Cantiani, P., and Meo, I. D.: Multi-functional approach in forest landscape management planning: an application in Southern Italy, For. Syst., 21, 68–80, https://doi.org/10.5424/fs/2112211-11066, 2012.

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

Parker, G. G.: Tamm review: Leaf Area Index (LAI) is both a determinant and a consequence of important processes in vegetation canopies, For. Ecol. Manag., 477, 118496, https://doi.org/10.1016/j.foreco.2020.118496, 2020.

Paule, L.: Gene conservation in European beech (Fagus Sylvatica L.), For. Genet., 2, 160–170, 1995.

Pavlenda, P. and Pajtík, J.: Monitoring lesov Slovenska, LVÚ Zvolen, Zvolen, NLC, 2010.

Petráš, R. and Pajtík, J.: Sústava československých objemových tabuliek drevín, Lesn. Časopis, 37, 49–56, 1991.

Petrik, P., Petek-Petrik, A., Kurjak, D., Mukarram, M., Klein, T., Gömöry, D., Střelcová, K., Frýdl, J., and Konôpková, A.: Interannual adjustments in stomatal and leaf morphological traits of European beech (Fagus sylvatica L.) demonstrate its climate change acclimation potential, Plant Biol., 24, 1287–1296, https://doi.org/10.1111/plb.13401, 2022.

Pietsch, S. A., Hasenauer, H., Kučera, J., and Čermák, J.: Modeling effects of hydrological changes on the carbon and nitrogen balance of oak in floodplains, Tree Physiol., 23, 735–746, https://doi.org/10.1093/treephys/23.11.735, 2003.

Pietsch, S. A., Hasenauer, H., and Thornton, P. E.: BGC-model parameters for tree species growing in central European forests, For. Ecol. Manag., 211, 264–295, https://doi.org/10.1016/j.foreco.2005.02.046, 2005.

Pisek, J., Sonnentag, O., Richardson, A. D., and Mõttus, M.: Is the spherical leaf inclination angle distribution a valid assumption for temperate and boreal broadleaf tree species?, Agric. For. Meteorol., 169, 186–194, https://doi.org/10.1016/j.agrformet.2012.10.011, 2013.

Poesen, J. and Lavee, H.: Rock fragments in top soils: significance and processes, CATENA, 23, 1–28, https://doi.org/10.1016/0341-8162(94)90050-7, 1994.

Post, W. M., Emanuel, W. R., Zinke, P. J., and Stangenberger, A. G.: Soil carbon pools and world life zones, Nature, 298, 156–159, https://doi.org/10.1038/298156a0, 1982.

Powers, J. S. and Schlesinger, W. H.: Relationships among soil carbon distributions and biophysical factors at nested spatial scales in rain forests of northeastern Costa Rica, Geoderma, 109, 165–190, https://doi.org/10.1016/S0016-7061(02)00147-7, 2002.

R Core Team: R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, 2018.

Rodeghiero, M. and Cescatti, A.: Main determinants of forest soil respiration along an elevation/temperature gradient in the Italian Alps, Glob. Change Biol., 11, 1024–1041, https://doi.org/10.1111/j.1365-2486.2005.00963.x, 2005.

Roibu, C.-C., Palaghianu, C., Nagavciuc, V., Ionita, M., Sfecla, V., Mursa, A., Crivellaro, A., Stirbu, M.-I., Cotos, M.-G., Popa, A., Sfecla, I., and Popa, I.: The Response of Beech (Fagus sylvatica L.) Populations to Climate in the Easternmost Sites of Its European Distribution, Plants, 11, 3310, https://doi.org/10.3390/plants11233310, 2022.

1295 Saltelli, A., Tarantola, S., Campolongo, F., and Ratto, M. (Eds.): Sensitivity analysis in practice: a guide to assessing scientific models, John Wiley & Sons Ltd, Hoboken, NJ, 219 pp., 2004.

Schieber, B., Janík, R., and Snopková, Z.: Phenology of common beech (Fagus sylvatica L.) along the altitudinal gradient in Slovak Republic (Inner Western Carpathians), 9, 2013.

Schmidt, M. W. I., Torn, M. S., Abiven, S., Dittmar, T., Guggenberger, G., Janssens, I. A., Kleber, M., Kögel-Knabner, I.,
1300 Lehmann, J., Manning, D. A. C., Nannipieri, P., Rasse, D. P., Weiner, S., and Trumbore, S. E.: Persistence of soil organic matter as an ecosystem property, Nature, 478, 49–56, https://doi.org/10.1038/nature10386, 2011.

Sever, M. Z. O., Paladinić, E., Barcza, Z., Hidy, D., Kern, A., Anić, M., and Marjanović, H.: Biogeochemical modelling vs. tree-ring measurements - Comparison of growth dynamic estimates at two distinct oak forests in Croatia, South-East Eur. For., 8, 71–84, https://doi.org/10.15177/seefor.17-17, 2017.

1305 Standovár, T. and Kenderes, K.: A review on natural stand dynamics in Beechwoods of East Central Europe, Appl. Ecol. Environ. Res., 1, https://doi.org/10.15666/aeer/01019046, 2003.

Sun, X., Tang, Z., Ryan, M. G., You, Y., and Sun, O. J.: Changes in soil organic carbon contents and fractionations of forests along a climatic gradient in China, For. Ecosyst., 6, 1, https://doi.org/10.1186/s40663-019-0161-7, 2019.

Tahiri, A. Z., Anyoji, H., and Yasuda, H.: Fixed and variable light extinction coefficients for estimating plant transpiration
1310 and soil evaporation under irrigated maize, Agric. Water Manag., 84, 186–192, https://doi.org/10.1016/j.agwat.2006.02.002, 2006.

Tang, J., Sun, B., Cheng, R., Shi, Z., Luo, D., Liu, S., and Centritto, M.: Seedling leaves allocate lower fractions of nitrogen to photosynthetic apparatus in nitrogen fixing trees than in non-nitrogen fixing trees in subtropical China, PLOS ONE, 14, e0208971, https://doi.org/10.1371/journal.pone.0208971, 2019.

1315 Tarantola, A.: Inverse Problem Theory and Methods for Model Parameter Estimation, Society for Industrial and Applied Mathematics, 348 pp., https://doi.org/10.1137/1.9780898717921, 2005.

Therneau, T., Atkinson, B., and Ripley, B.: Recursive Partitioning and Regression Trees - Package 'rpart,' 2023.

Thornton, P., Running, S. W., and Hunt, E. R.: Biome-BGC: Terrestrial Ecosystem Process Model, Version 4.1.1, https://doi.org/10.3334/ornldaac/805, 2005.

1320 Thornton, P. E.: Regional Ecosystem Simulation: Combining Surface- and Satellite-Based Observations to Study Linkages between Terrestrial Energy and Mass Budgets, College of Forestry, The University of Montana, Montana, 288 pp., 1998.

Thornton, P. E. and Rosenbloom, N. A.: Ecosystem model spin-up: estimating steady state conditions in a coupled terrestrial carbon and nitrogen cycle model, Ecol. Model., 189, 25–48, https://doi.org/10.1016/j.ecolmodel.2005.04.008, 2005.

Thornton, P. E., Law, B. E., Gholz, H. L., Clark, K. L., Falge, E., Ellsworth, D. S., Goldstein, A. H., Monson, R. K.,
1325 Hollinger, D., Falk, M., Chen, J., and Sparks, J. P.: Modeling and measuring the effects of disturbance history and climate on

carbon and water budgets in evergreen needleleaf forests, Agric. For. Meteorol., 113, 185–222, https://doi.org/10.1016/S0168-1923(02)00108-9, 2002.

1330 Tian, H., Yang, J., Lu, C., Xu, R., Canadell, J. G., Jackson, R. B., Arneth, A., Chang, J., Chen, G., Ciais, P., Gerber, S., Ito, A., Huang, Y., Joos, F., Lienert, S., Messina, P., Olin, S., Pan, S., Peng, C., Saikawa, E., Thompson, R. L., Vuichard, N., Winiwarter, W., Zaehle, S., Zhang, B., Zhang, K., and Zhu, Q.: The Global N2O Model Intercomparison Project, Bull. Am. Meteorol. Soc., 99, 1231–1251, https://doi.org/10.1175/BAMS-D-17-0212.1, 2018.

Timlin, D. J., Fleisher, D. H., Kemanian, A. R., and Reddy, V. R.: Plant Density and Leaf Area Index Effects on the Distribution of Light Transmittance to the Soil Surface in Maize, Agron. J., 106, 1828–1837, https://doi.org/10.2134/agronj14.0160, 2014.

1335 Trotsiuk, V., Hobi, M. L., and Commarmot, B.: Age structure and disturbance dynamics of the relic virgin beech forest Uholka (Ukrainian Carpathians), For. Ecol. Manag., 265, 181–190, https://doi.org/10.1016/j.foreco.2011.10.042, 2012.

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

1340 Tumajer, J., Scharnweber, T., Smiljanic, M., and Wilmking, M.: Limitation by vapour pressure deficit shapes different intra-annual growth patterns of diffuse- and ring-porous temperate broadleaves, New Phytol., 233, 2429–2441, https://doi.org/10.1111/nph.17952, 2022.

Urban, J., Ingwers, M., McGuire, M. A., and Teskey, R. O.: Stomatal conductance increases with rising temperature, Plant Signal. Behav., 12, e1356534, https://doi.org/10.1080/15592324.2017.1356534, 2017.

1345 Vanoni, M., Cailleret, M., Hülsmann, L., Bugmann, H., and Bigler, C.: How do tree mortality models from combined tree-ring and inventory data affect projections of forest succession?, For. Ecol. Manag., 433, 606–617, https://doi.org/10.1016/j.foreco.2018.11.042, 2019.

Verbeeck, H., Samson, R., Verdonck, F., and Lemeur, R.: Parameter sensitivity and uncertainty of the forest carbon flux model FORUG: a Monte Carlo analysis, Tree Physiol., 26, 807–817, https://doi.org/10.1093/treephys/26.6.807, 2006.

1350 Vesterdal, L. and Raulund-Rasmussen, K.: Forest floor chemistry under seven tree species along a soil fertility gradient, Can. J. For. Res., 28, 1636–1647, https://doi.org/10.1139/x98-140, 1998.

Vesterdal, L., Schmidt, I. K., Callesen, I., Nilsson, L. O., and Gundersen, P.: Carbon and nitrogen in forest floor and mineral soil under six common European tree species, For. Ecol. Manag., 255, 35–48, https://doi.org/10.1016/j.foreco.2007.08.015, 2008.

1355 Wallach, D., Palosuo, T., Thorburn, P., Hochman, Z., Gourdain, E., Andrianasolo, F., Asseng, S., Basso, B., Buis, S., Crout, N., Dibari, C., Dumont, B., Ferrise, R., Gaiser, T., Garcia, C., Gayler, S., Ghahramani, A., Hiremath, S., Hoek, S., Horan, H., Hoogenboom, G., Huang, M., Jabloun, M., Jansson, P. E., Jing, Q., Justes, E., Kersebaum, K. C., Klosterhalfen, A., Launay, M., Lewan, E., Luo, Q., Maestrini, B., Mielenz, H., Moriondo, M., Nariman Zadeh, H., Padovan, G., Olesen, J. E., Poyda, A., Priesack, E., Pullens, J. W. M., Qian, B., Schütze, N., Shelia, V., Souissi, A., Specka, X., Srivastava, A. K., Stella, T., 1360 Streck, T., Trombi, G., Wallor, E., Wang, J., Weber, T. K. D., Weihermüller, L., de Wit, A., Wöhling, T., Xiao, L., Zhao, C., Zhu, Y., and Seidel, S. J.: The chaos in calibrating crop models: Lessons learned from a multi-model calibration exercise, Environ. Model. Softw., 145, 105206, https://doi.org/10.1016/j.envsoft.2021.105206, 2021.

Wang, G., Zhou, Y., Xu, X., Ruan, H., and Wang, J.: Temperature Sensitivity of Soil Organic Carbon Mineralization along an Elevation Gradient in the Wuyi Mountains, China, PLoS ONE, 8, e53914, https://doi.org/10.1371/journal.pone.0053914, 2013.

Wang, Q., Tenhunen, J., Granier, A., Reichstein, M., Bouriaud, O., Nguyen, D., and Breda, N.: Long-term variations in leaf area index and light extinction in a Fagus sylvatica stand as estimated from global radiation profiles, Theor. Appl. Climatol., 79, 225–238, https://doi.org/10.1007/s00704-004-0074-3, 2004.

Warren, C. R. and Adams, M. A.: What determines rates of photosynthesis per unit nitrogen in Eucalyptus seedlings?, Funct. Plant Biol., 31, 1169–1178, https://doi.org/10.1071/FP04115, 2004.

Weemstra, M., Sterck, F. J., Visser, E. J. W., Kuyper, T. W., Goudzwaard, L., and Mommer, L.: Fine-root trait plasticity of beech (Fagus sylvatica) and spruce (Picea abies) forests on two contrasting soils, Plant Soil, 415, 175–188, https://doi.org/10.1007/s11104-016-3148-y, 2017.

Wellbrock, N. and Bolte, A. (Eds.): Status and Dynamics of Forests in Germany: Results of the National Forest Monitoring, Springer Nature, https://doi.org/10.1007/978-3-030-15734-0, 2019.

Wellbrock, N., Bolte, A., and Flessa, H.: Dynamik und räumliche Muster forstlicher Standorte in Deutschland - Ergebnisse der Bodenzustandserhebung im Wald 2006 bis 2008, in: Thünen Report 43, Johann Heinrich von Thünen-Institut, Braunschweig, 550, 2016.

Wellbrock, N., Grüneberg, E., Riedel, T., and Polley, H.: Carbon stocks in tree biomass and soils of German forests, Cent. Eur. For. J., 63, 105–112, https://doi.org/10.1515/forj-2017-13, 2017.

White, M. A., Running, S. W., and Thornton, P. E.: The impact of growing-season length variability on carbon assimilation and evapotranspiration over 88 years in the eastern US deciduous forest, Int. J. Biometeorol., 42, 139–145, 1999.

Wickham, H.: ggplot2: Elegant Graphics for Data Analysis., Springer-Verlag New York, 2016.

Wiesmeier, M., Urbanski, L., Hobley, E., Lang, B., von Lützow, M., Marin-Spiotta, E., van Wesemael, B., Rabot, E., Ließ, M., Garcia-Franco, N., Wollschläger, U., Vogel, H.-J., and Kögel-Knabner, I.: Soil organic carbon storage as a key function of soils - A review of drivers and indicators at various scales, Geoderma, 333, 149–162, https://doi.org/10.1016/j.geoderma.2018.07.026, 2019.

Wilcoxon, F.: Individual Comparisons by Ranking Methods, Biom. Bull., 1, 80–83, https://doi.org/10.2307/3001968, 1945.

Wöhling, T., Gayler, S., Priesack, E., Ingwersen, J., Wizemann, H. D., Högy, P., Cuntz, M., Attinger, S., Wulfmeyer, V., and Streck, T.: Multiresponse, multiobjective calibration as a diagnostic tool to compare accuracy and structural limitations of five coupled soil-plant models and CLM3.5, Water Resour. Res., 49, 8200–8221, https://doi.org/10.1002/2013WR014536, 2013.

Woollen, E., Ryan, C. M., and Williams, M.: Carbon Stocks in an African Woodland Landscape: Spatial Distributions and Scales of Variation, Ecosystems, 15, 804–818, https://doi.org/10.1007/s10021-012-9547-x, 2012.

Yan, M., Tian, X., Li, Z., Chen, E., Wang, X., Han, Z., and Sun, H.: Simulation of forest carbon fluxes using model incorporation and data assimilation, Remote Sens., 8, https://doi.org/10.3390/rs8070567, 2016.

Zaehle, S.: Terrestrial nitrogen-carbon cycle interactions at the global scale, Philos. Trans. R. Soc. B Biol. Sci., 368, 20130125–20130125, https://doi.org/10.1098/rstb.2013.0125, 2013.

Zhang, L., Hu, Z., Fan, J., Zhou, D., and Tang, F.: A meta-analysis of the canopy light extinction coefficient in terrestrial ecosystems, Front. Earth Sci., 8, 599–609, https://doi.org/10.1007/s11707-014-0446-7, 2014.

1400