Supplemental information

Target growth curve derivation

The growth curves for the three FPFTs were derived from observations and LPJmL predictions for the corresponding natural PFTs using Bayesian non-linear regression. We assume that observed stemwood C is given by:

$$C_{\text{SWC}}^{\text{obs}}(t) = C_{\text{SWC}}^{\text{pred}}(t) + \varepsilon_t = C_{\text{SWC,max}}\left(1 - e^{-k\,t}\right)^p + \varepsilon_t,$$

where ε_t is the model-data residual, assumed to be stochastic. We used normal priors for the parameters of the growth curve. For $C_{\text{SWC,max}}$ strong priors were used, based on maximum stemwood C for the corresponding natural PFTs predicted by LPJmL (see section 2.3.1 in the main text; $N(5.43, 1.88)$, $N(14.0, 2.40)$, and $N(6.16, 1.24)$ for the temperate, tropical and boreal FPFT, respectively). For k and p weaker priors were used ($N(0.05, 1)$ and $N(2, 20)$, respectively, for all three FPFTs), allowing them to be more strongly influenced by the observations. Model-data residuals can be expected to scale with stemwood C, since both predictions and observations cannot be lower than zero. This is exacerbated by the fact that we incorporate a strong prior for $C_{\text{SWC,max}}$, which may to some extent be inconsistent with the data, leading to a stronger misfit for higher plantation ages. Ignoring this heteroscedasticity may result in biased parameter estimates. Therefore we assumed that the variance of ε_t scales linearly with predicted stemwood C:

$$\varepsilon_t \sim N\left(0, \sigma_s C_{\text{SWC}}^{\text{pred}}(t)\right)$$

where σ_s is a scaling factor, simultaneously determined in the fitting, using a weak prior ($N(1, 10)$).

The fitting was carried out with Markov Chain Monte Carlo sampling, implemented in the Stan programming language (Carpenter et al., 2017), linked to Matlab (Stan Development Team, 2017). Specifically, we used the No-U-Turn-Sampler (NUTS; Hoffman and Gelman, 2014) with four chains of 10,000 samples each, and a warmup period of 10,000 iterations per chain. Figure S1 depicts the marginal posterior distributions of the four estimated parameters for the three FPFTs, together with the prior distributions. The parameter set with highest posterior density, together with variance over the complete sample, was used for the target growth curve in the LPJmL calibration (see table S2).
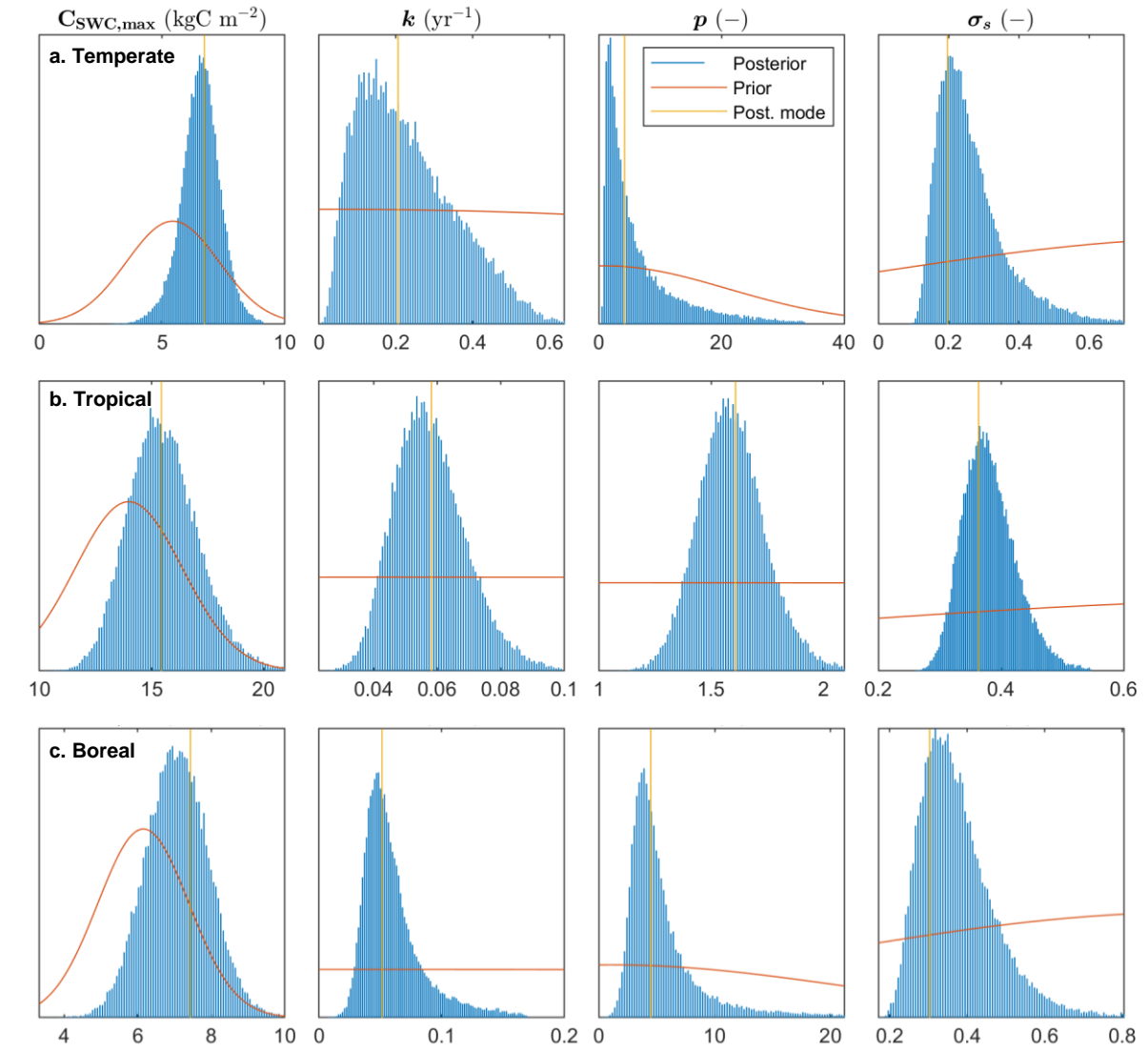
Figure S1. Results of the Monte Carlo sampling to determine the parameters of the target growth curves for temperate (a), tropical (b), and boreal (c) forest plantations. Note that the multivariate posterior mode, determined as the sample with highest posterior density (yellow lines in the graphs), deviates from the marginal modes (peak of the histograms), due to non-linear parameter correlations.

LPJmL calibration

The optimization of the FPFT parameters in this study comprised minimization of a Bayesian cost-function, essentially negative log posterior density. The cost-function is defined as he sum of a log-likelihood term and log prior density term. Prior distributions on the parameters were defined by gaussian probability density functions, defined on logit scale, for α_a, and log-scale for all other parameters. Similarly, the observations were log- or logit-transformed as well. For a given parameter set θ the overall cost $C(\theta)$ is given by:

$$C(\theta) = \sum_{i=1}^{I} \frac{(f_i(\theta_i) - \mu_{\theta_i})^2}{\sigma_{\theta_i}^2} + \sum_{j=1}^{J} \frac{(f_j(\varphi_j) - \mu_{\varphi_j})^2}{\sigma_{\varphi_j}^2}, \qquad (1)$$

 The first term on the right-hand size of (3) represents the prior in which: I is the number of optimized parameters (15), θ_i and f_i are the value and the transformation function (logit or log) of parameter i, and μ_{θ_i} and $\sigma_{\theta_i}^2$ are the mean and variance of the prior for parameter i on transformed scale. Additionally, bounds were enforced on the parameters. The lower bound was at 0 and the upper bound at the 99% quantile of the prior distributions, except for α_a, for which it was set at 1. Table S1 shows μ_{θ_i}, $\sigma_{\theta_i}^2$, and the upper bound of all parameters for the three FPFTs. The second term on the right-hand size of (3) represents the likelihood, in which: J is the number of observations (5), φ_j and f_j are the LPJmL output (based on parameter set θ) and transformation function (logit or log) of observation variable j, and μ_{φ_j} and $\sigma_{\varphi_j}^2$ are the mean and variance of observation variable j on transformed scale. Table S2 shows the μ_{φ_i}, $\sigma_{\varphi_i}^2$ for all observations for the three FPFTs.

Table S1. Mean (μ), variance (σ^2), and upper bound (u.b.) of the prior distributions on the parameters used in the LPJmL calibration. Note that μ and σ^2 are defined on transformed (log or logit) scale, while the upper bound is defined on untransformed scale.

Parameter (i)	Transformation (f_i)	Temperate			Tropical			Boreal		
		μ_{θ_i}	$\sigma_{\theta_i}^2$	u.b.	μ_{θ_i}	$\sigma_{\theta_i}^2$	u.b.	μ_{θ_i}	$\sigma_{\theta_i}^2$	u.b.
α_a	logit	0.00	0.700	1.00	0.00	0.700	1.00	0.00	0.700	1.00
g_{min}	log	-0.693	0.300	2.41	-0.693	0.300	2.41	-1.20	0.300	1.45
E_{max}	log	1.61	0.300	24.1	1.95	0.300	33.8	1.61	0.300	24.1
r	log	0.182	0.300	5.79	-1.61	0.300	0.965	0.182	0.300	5.79
k_{allom1}	log	4.61	0.300	481	4.61	0.300	4.61	4.70	0.300	530
k_{allom2}	log	3.69	0.300	193	3.69	0.300	193	3.69	0.300	193
k_{allom3}	log	-0.400	0.300	3.23	-0.40	0.300	3.23	-0.40	0.300	3.23
lr_{max}	log	0.00	0.300	4.83	0.00	0.300	4.83	0.00	0.300	4.83
$C_{sapwood;sapl}$	log	0.182	0.300	5.79	0.182	0.300	5.79	0.182	0.300	5.79
LAI_{sapl}	log	0.405	0.300	7.24	0.405	0.300	7.24	0.405	0.300	7.24
α_{leaf}	log	-0.693	0.300	2.41	0.693	0.300	9.65	1.39	0.300	19.3
$\tau_{sapwood}$	log	3.00	0.300	96.5	3.00	0.300	96.5	3.00	0.300	96.5
$\tau_{leaf,root}$	log	0.00	0.300	4.83	0.693	0.300	9.65	1.39	0.300	19.3
P_{init}	log	-1.90	0.300	0.724	-1.90	0.300	0.724	-1.90	0.300	0.724
k_{mort1}	log	-3.51	0.300	0.145	-3.51	0.300	0.145	-3.51	0.300	0.145

Table S2. Mean (μ), variance (σ^2), of the likelihood functions for the observations used in the LPJmL calibration. Note that μ and σ^2 are defined on transformed (log or logit) scale.

Obser-vation (j)	Transfor-mation (f_j)	Temperate		Tropical		Boreal	
		μ_{φ_j}	$\sigma^2_{\varphi_j}$	μ_{φ_j}	$\sigma^2_{\varphi_j}$	μ_{φ_j}	$\sigma^2_{\varphi_j}$
$C_{SWC,max}$	log	8.82	0.0163	9.66	0.00989	8.92	0.0150
k	log	-1.63	0.431	-2.87	0.0389	-3.00	0.300
p	log	1.36	0.865	0.465	0.00889	1.45	0.164
CUE	logit	-0.488	0.200	-0.169	0.0262	-0.159	0.123
τ_{vegC}	log	2.82	0.0500	3.09	0.0500	3.10	0.05

2 The parameter set θ at which the cost is minimal was obtained using the genoud (GENetic Optimization

3 Using Derivatives) algorithm. Genoud is a hybrid optimization scheme that combines a genetic

4 algorithm with gradient-search approach (Sekhon and Mebane, 1998). Genetic algorithms are schemes

5 that search for solutions to optimization problems with complex, non-smooth functions. These

6 problems are often difficult to solve using gradient-search algorithms (e.g. Newton-Raphson), which

7 rely on derivatives in order to determine successive steps. Genetic algorithms use an approach inspired

8 by the process of natural selection, working with a population of candidate solutions (parameter sets),

9 which are evolved in successive generations. This is done by means of "genetic operators", which

10 combine and mutate individuals in the population. While genetic algorithms are highly suitable for

11 irregular functions, they are often slower than gradient-search algorithms in the neighborhood of the

12 global solution, where the surface is usually smooth and concave. Therefore, genoud employs a

13 gradient-search algorithm on the best candidate solution in the population after each generation. The

14 parameter value resulting from this gradient-search step is used in next generation. The algorithm is

15 implemented in R language and available as the rgenoud package (Mebane Jr. and Sekhon, 2011). In

16 our calibrations we used a population size of 1000 individuals and a maximum number of 20

17 generations. For the gradient-search algorithm the BGFS scheme was used.

18
19

20

21
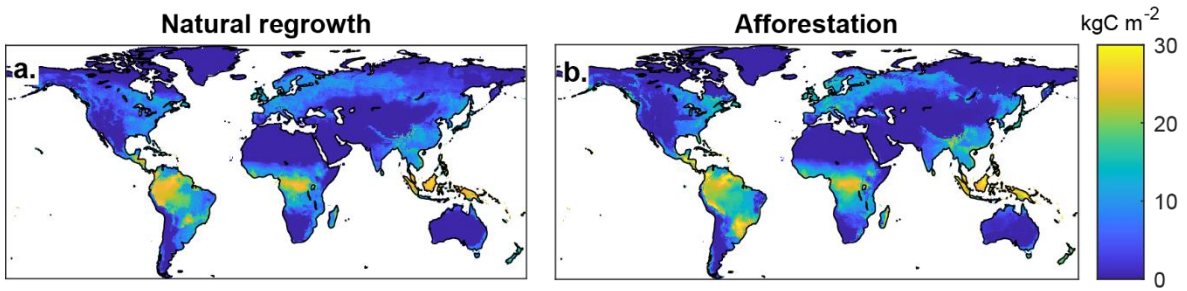
Figure S2. Total ecosystem C after 300 years for a global simulation with only natural vegetation and only forest plantations.
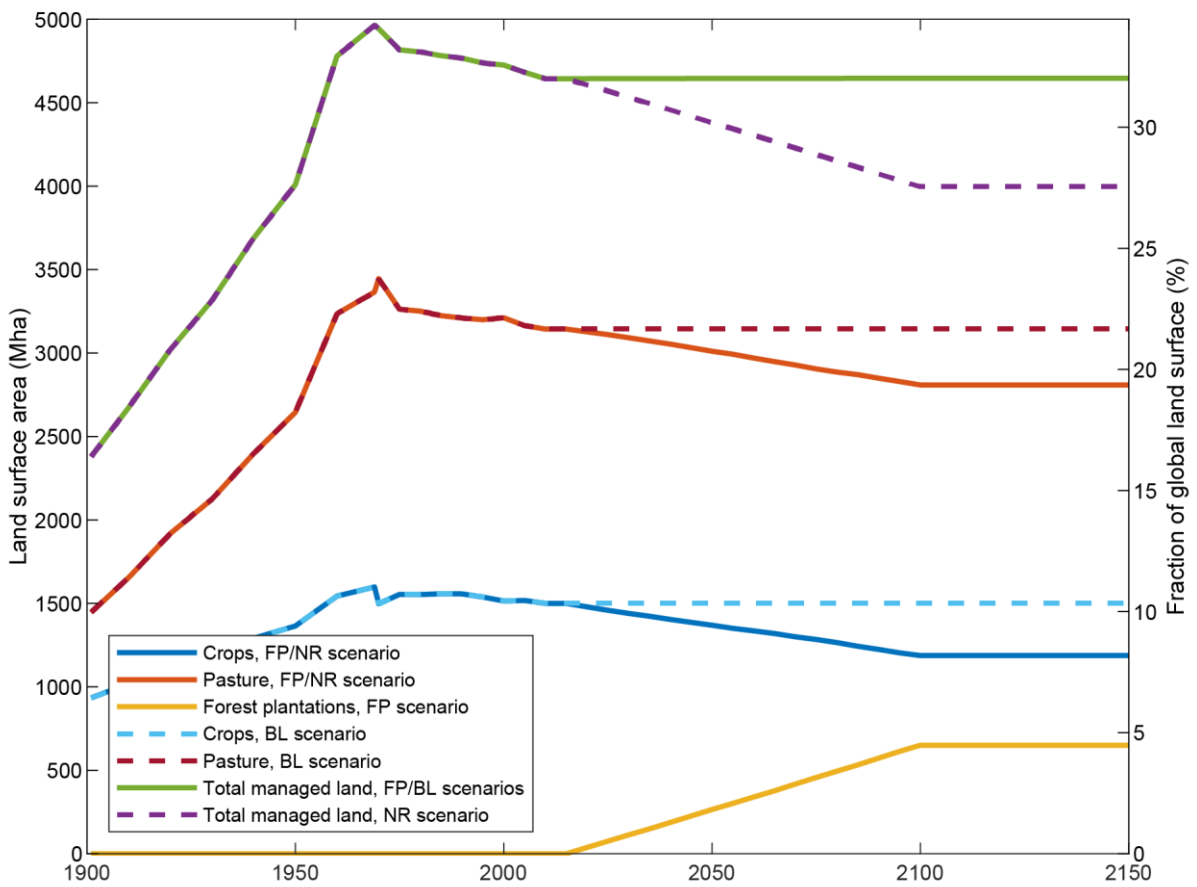
Figure S3. Development of land-use fractions for the transient afforestation scenario (FP) and accompanying scenarios for natural regrowth (NR) and the baseline (BL) with constant land-use after 2015. "Total managed land" refers to the sum of croplands, pasture, and forest plantations.
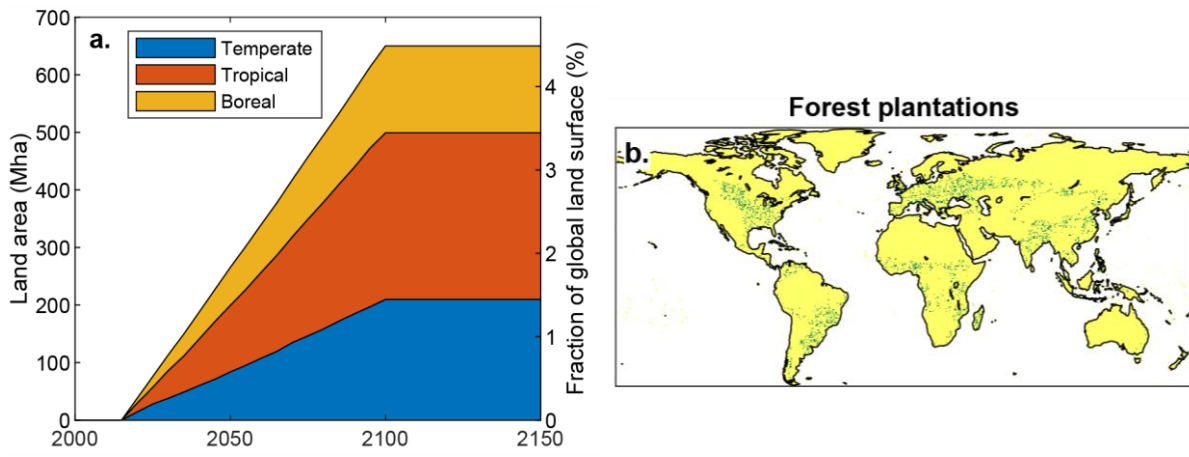
Figure S4. Forest plantation land use fraction for the transient afforestation scenario. (a) global total afforestation area versus time; (b) afforested grid cells from 2100.