Supplementary Information for

Forestation at the right time with the right species can generate persistent carbon benefits in China

Hao Xu[1], Chao Yue[2]*, Yao Zhang[1], Dan Liu[3], Shilong Piao[1,3]*

*Corresponding author: Chao Yue (chaoyue@ms.iswc.ac.cn), Shilong Piao (slpiao@pku.edu.cn)

This file includes:

Supplementary Method 1-3, Supplementary Table 1-8,

Supplementary Figure 1-10, and References

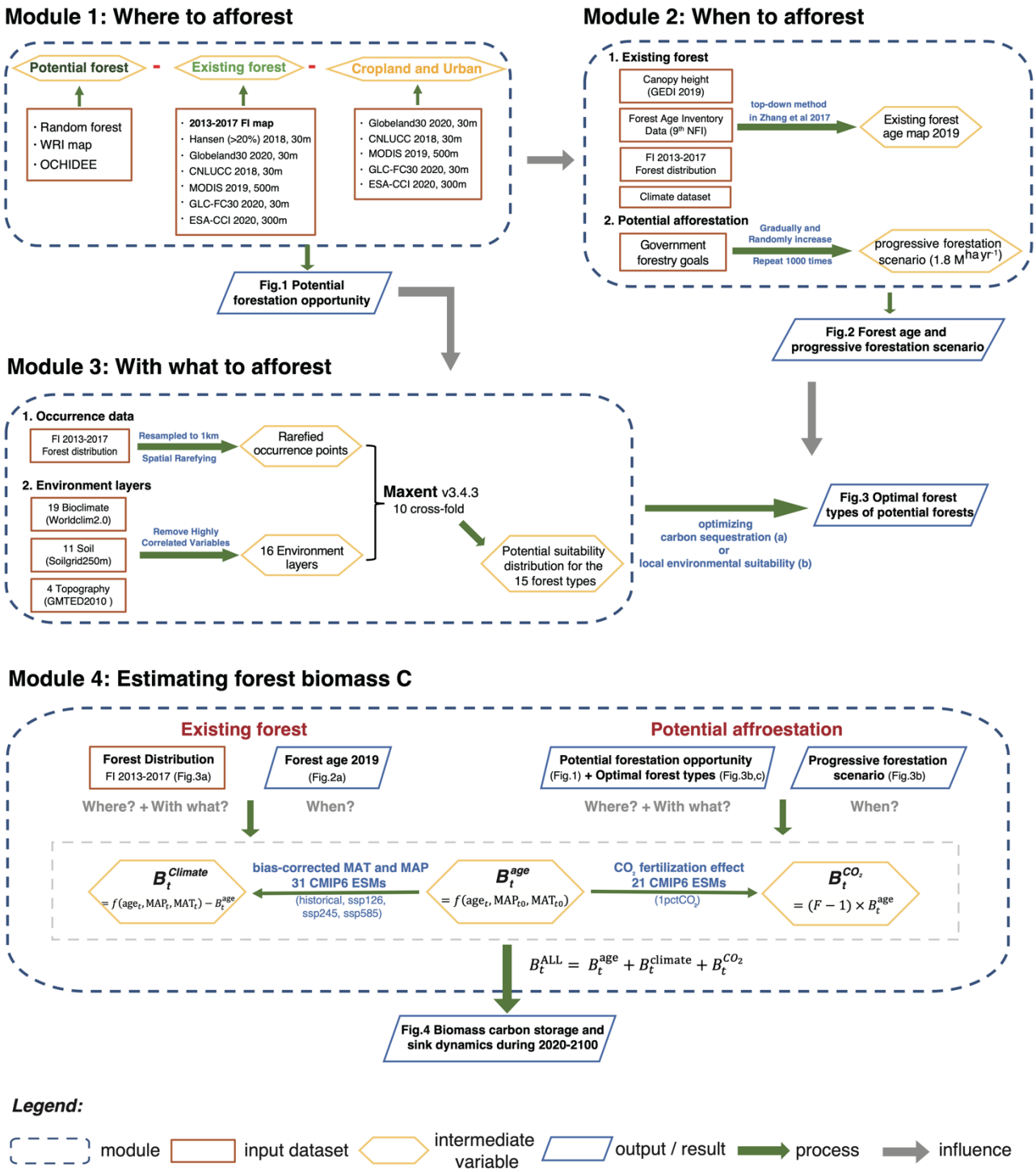

Figure S1. Research framework and workflow of our study.

Supplementary Method 1： Potential forest distribution

1) Random forest

We obtained the potential tree cover distribution in China by using random-forest modeling following an approach similar to Bastin et al. (1). In that study, 78,774 measurements of tree cover (0–100%) in protected regions across the globe were made by direct photo interpretation and were used as tree cover observations under various natural environment conditions with minimal human influence. We fitted a random forest model (R^2 = 0.81, number of trees: 20) using their observations in China (3840 measurements, coefficient of determination and various environmental drivers (5 climatic, 3 soil and 2 topographic variables, Table S2). The climatic, soil and topographic data were extracted from Worldclim2.1 (2), Soilgrid250m (3) and GMTED2010 (4), respectively (Table S1). All covariate layers were resampled to a unified World Geodetic System 1984 Coordinate System (WGS-1984) projection, at 30 arc-seconds resolution (≈1km at the equator). Layers with a higher original pixel resolution were re-gridded using an area-weighted mean aggregation method. The potential tree cover (0-100%) was then spatially extrapolated outside protected areas for each pixel using model coefficients combining the values of 10 selected variables across China at a spatial resolution of 30 arcsec (~ 1km). Grid cells at a 1km resolution with a tree cover ≥ 20% were defined as "forest", consistent with the definition used in the Chinese National Forest Continuous Inventory (5).

2) Potential forest coverage map from WRI

As part of an effort made by the World Resources Institute Global Restoration Initiative to map forest and landscape restoration opportunities (6), the map of potential forests (1-km resolution) was generated as an estimate of where forests would grow under current climate conditions and without human influence. The global distribution of terrestrial ecoregions (7) was used as the main data source to define potential forest extent in this map. More specifically, each ecoregion was classified as one of four categories: dense forest (canopy cover>45%), open forest (canopy cover 25-45%), woodland (canopy cover < 25%), or non-forest. The classification was determined according to the nature of the ecoregion concerned. In this study, the three classes: dense forest, open forest and woodland were combined and considered as 'forest' when deriving the potential forest distribution.

3) Dynamic global vegetation model simulation of ORCHIDEE

We employed the process-based dynamic global vegetation model ORCHIDEE (Organizing Carbon and Hydrology in Dynamic Ecosystems) to simulate the potential forest distributions in China. We used the version ORC-HL-NVD which has improved parameterization for vegetation dynamics in northern temperate and boreal regions, including the processes of establishment, mortality, and species competition for different tree functional types (8). Current climate variables from the 3-hourly China Meteorological Forcing Dataset (CMFD) (9) at a 0.5° resolution (resampled from the original 0.1° resolution data) were used to force the model. These data consisted of 2-meter air temperature, surface pressure, specific humidity, 10-meter wind speed, downward shortwave and longwave radiation, and

precipitation (Table S1).

The simulation started from a state of 100% ground coverage (i.e., bare soil everywhere) and lasted for 200 years, with cycling 2006-2015 CMFD climate forcing and constant contemporary CO_2 concentration (405 ppm). As ecosystems were simulated to succeed one another over time, the proportions of ground area covered by forest and grass ecosystems gradually increased, and finally reached equilibrium in the later stages of succession (Fig. S2). The average proportion of PFTs after model equilibration (the last 20 years) was the potential forest distribution we obtained (0.5°). Aiming to integrate with the other two 1-km distribution maps, we converted the potential forest coverage (%) as simulated by the ORCHIDEE model at a 0.5° spatial resolution to forest distribution at a 1-km resolution. This was done by first ranking the underlying 1-km pixels within each 0.5º grid cell by their tree cover as simulated by the random-forest modeling approach, and then by choosing an appropriate threshold such that the total area of the 1-km pixels with a predicted forest cover above the threshold was equal to the area of forest coverage predicted by the ORCHIDEE model for the 0.5º grid cell concerned.

Supplementary Method 2: Species distribution model

We constructed species distribution models to predict the species suitability distribution (i.e., probability of occurrence) across China for 15 dominant forest types. This method consists of the following four steps: (1) preparing the species occurrence observations, (2) filtering the environment variables, (3) running the MaxEnt model, (4) model performance evaluation.

Firstly, we prepared the species occurrence observations by resampling from the national field survey data, and when necessary, further spatially rarefying the observations. We obtained the initial presence data for each forest type by resampling the 1:1000000 vector map of the FI 2013-2017 (10) into a matrix of dots spaced 1-km apart. These initial species presence points have a strong spatial auto-correlation, which often causes model over-fitting and inflated model performance. Thus, we used the "*spatially rarefy occurrence data*" tool of SDMtoolbox2.0 (11) developed by Jason L. Brown to extract statistically viable observations from the complete set of species presence points by removing spatially and climatically redundant observations. The key function of the SDMtoolbox2.0 tool is to reduce the extent of spatial clustering of presence observations by filtering presence data through user-defined tolerance distances and/or a map of climate heterogeneity index. We first performed principal component analysis for all bioclimate variables across China and then used the first 3 principal components to calculate the climate heterogeneity index. The extent of climate heterogeneity was classified into three classes (high, medium, low) using the "Natural Breaks" method in SDMtoolbox2.0. Different distance thresholds were applied for the three classes of climate heterogeneity, with a distance threshold of 25 km used for the low-heterogeneity areas, 10 km for the medium-heterogeneity areas and 5 km for the high-heterogeneity areas. Spatially and climatically superfluous observations were removed until only a single point remained within the specified distance thresholds. The remaining presence data after redundancy

removal were used as observations in the construction of species distribution models for each forest type.

Secondly, we filtered the environment variables which were used as dependent variables in the species distribution models. Initially 19 bioclimatic variables from 1970-2000 WorldClim2.1 (2), 11 soil variables from the SoilGrid250m (3) dataset, and 4 topographic variables from GMTED2010 (4) were included as candidate variables, with all the variables resampled to a consistent resolution (1-km). To reduce the influence of multicollinearity among environmental variables, correlation analysis was carried out among all candidate variables. Variables that were highly correlated (with a Pearson's correlation coefficient $|r| > 0.7$) were selectively removed based on expertise and their respective importance from the initial MaxEnt model with all variables input. Finally, 17 variables were retained: 6 bioclimatic, 8 soil and 3 topographic variables (Table S2).

The statistically viable species presence observations and environmental variables were then fed into the MaxEnt software (Version 3.4.3) (12) to fit species distribution models, using a maximum iteration number of 5000 and 10-fold cross validations. The occurrence and environmental datasets were divided randomly into 10 equal-sized groups, and models were built using k-1 bins for calibration in each iteration (training set), with the left-out bin used for evaluation (test set).

Model performance was assessed using the receiver operating characteristic (ROC) curves, and the area under the curve (AUC) scores from the 10-fold cross-validations were used to evaluate the accuracy of the resulting model (12). The AUC score varies from 0 to 1, with larger values indicating better model performance and more reliable predictions. The obtained MaxEnt models of the 15 forest types all showed high accuracy (AUC $\geq$ 0.83, Table S4), which confirmed the feasibility of using a species distribution model to guide forestation planning and forest type selection.

Supplementary Method 3: Bias corrections for future climate

Since outputs of climate simulations from CMIP6 models are generally biased compared to observations and these outputs are provided at very coarse resolutions (0.8-2.8 degrees), both bias correction and spatial downscaling were required. We applied a trend-preserving bias correction method (13) to the projected temperature and precipitation data based on the MAT and MAP ‘‘observations" from the Peng_China_1km (14) dataset for the same historical period (T_0). Simultaneously, spatial downscaling can also be accomplished through spatial interpolation in cooperation with this bias correction process (13).

The MAT and MAP outputs from the CMIP6 models were first spatially interpolated to the higher spatial resolution (i.e., 1-km) using a nearest-neighbor approach. Next, we selected the 30-yr period from 1985 to 2014 as the historical reference period (T_0). The monthly temperature data were corrected by an additive factor C_m for each month m (m=1, 2…, 12), which represents the long-term differences between

the simulated ($tas^{\mathrm{sim}}_{T_0,m}$) and observed ($tas^{\mathrm{obs}}_{T_0,m}$) monthly mean data during the historical period (Eq. S1), i.e.

$$\widetilde{tas}^{\mathrm{sim}}_{t,m} = tas^{\mathrm{sim}}_{t,m} + C_m = tas^{\mathrm{sim}}_{t,m} + \left(tas^{\mathrm{obs}}_{T_0,m} - tas^{\mathrm{sim}}_{T_0,m}\right) \quad (S1)$$

where $\widetilde{tas}^{\mathrm{sim}}_{t,m}$ and $tas^{\mathrm{sim}}_{t,m}$ are the corrected and uncorrected monthly mean temperatures for period *t* at the 1-km resolution.

Given the positivity constraint on precipitation data, a similar additive approach is not appropriate. Instead, we correct the monthly mean precipitation values using a multiplicative factor c_m for each month *m* (*m*=1, 2…, 12), representing the long-term differences between the simulated ($pr^{\mathrm{sim}}_{T_0,m}$) and observed ($pr^{\mathrm{obs}}_{T_0,m}$) monthly mean data during the historical period (Eq.S2), i.e.

$$\widetilde{pr}^{\mathrm{sim}}_{t,m} = pr^{\mathrm{sim}}_{t,m} \times c_m = pr^{\mathrm{sim}}_{t,m} \times \left(pr^{\mathrm{obs}}_{T_0,m} \,/\, pr^{\mathrm{sim}}_{T_0,m}\right) \quad (S2)$$

where $\widetilde{pr}^{\mathrm{sim}}_{t,m}$ and $pr^{\mathrm{sim}}_{t,m}$ are the corrected and uncorrected monthly mean temperature for period *t*. We impose an upper bound of 10 on *c*, to avoid unrealistically high precipitation values. The above algorithms accomplished not only the bias correction to observation, but also the spatial downscaling, and thus generated high-resolution (i.e., 1-km) and bias-corrected MAP and MAT data, which were appropriate for use in estimating climate change-induced changes in total forest biomass C storage.

Table S1. Environment variables used in our study.

Datasets	Description and Variables	Available Period	Resolution		Reference
			spatial	temporal	
	Climatic				
Worldclim2.1	WorldClim, the most cited climatic dataset provides climatic indicators at very high spatial resolution(~1km). **Variables: minimum, mean, and maximum temperature, precipitation, solar radiation, wind speed, water vapor pressure and 19 standard bioclimatic variables**	1970-2000	1km	climatological average / monthly	(2)
Peng_china_1km	Spatially downscaled from CRU-TS-4.03, with the climatology dataset of WorldClim using delta spatial downscaling and evaluated using observations collected in 1951–2016 by 496 weather stations across China. **Variables: monthly precipitation and temperature**	1901-2020	1km	monthly	(14)
CMFD	The China Meteorological Forcing Dataset (CMFD) was produced through fusion of remote sensing products, reanalysis datasets and in-situ station data. **Variables: 2-meter air temperature, surface pressure, specific humidity, 10-meter wind speed, downward shortwave and longwave radiation, and precipitation rate**	1979-2018	0.1° × 0.1°	3-hourly	(9)
	Soil				
Soilgrid250m	Provides global predictions of standard numeric soil **properties: organic carbon, bulk density, Cation Exchange Capacity (CEC), pH, soil texture fractions and coarse fragments** at seven standard depths (0, 5, 15, 30, 60, 100 and	~	250m	~	(3)

	200 cm), in addition to predictions of **depth to bedrock** and distribution of soil classes. The soil samples from the first five layers were weighed and averaged as the attribute values of the soil in the corresponding layer.				
	Topographic				
GMTED2010	Global Multi-resolution Terrain Elevation Data 2010 is a notably enhanced global **elevation** model that replaces GTOPO30 as the elevation dataset of choice for global and continental scale applications. Other topographic properties (i.e., **slope, aspect, hillshade**) were calculated in ArcGIS 10.8	~	1km, 500m, 250m	~	(4)

Table S2. Candidate explanatory variables used for the Random Forest model and MaxEnt.

Variables name		Description	Type/Resource	Original Spatial Resolution
		Random Forest		
Bio1	✓	Annual Mean Temperature	*Climatic;* Worldclim 2.1	1km
Bio8	✓	Mean Temperature of Wettest Quarter		
Bio12	✓	Annual Precipitation		
Bio15	✓	Precipitation Seasonality (Coefficient of Variation)		
Bio17	✓	Precipitation of Driest Quarter		
elevation hillshade	-		*Topographic:* GMTED2010	250m
soc	✓	Soil Organic Carbon Stock from 0.00m-0.05m	*Soil:* Soilgrid250m	250m
sand	✓	Proportion of sand particles (>0.05mm) in the fine earth fraction at 0.05m		
depth	✓	Depth to Bedrock		
		MaxEnt v3.4.3		
Bio1	✓	Annual Mean Temperature	*Climatic;* Worldclim 2.1	1km
Bio3	✓	Isothermality (Mean Diurnal Range / Temperature Annual Range) (× 100)		
Bio4	✓	Temperature Seasonality (standard deviation ×100)		
Bio12	✓	Annual Precipitation		
Bio15	✓	Precipitation Seasonality (Coefficient of Variation)		
hillshade aspect slope	-		*Topographic:* GMTED2010	250m
ocs	✓	Organic carbon stock	*Soil:* Soilgrid250m	250m
bdod	✓	Bulk density of the fine earth fraction		
clay	✓	Proportion of clay particles (<0.002mm) in the fine earth fraction		
silt	✓	Proportion of silt particles (≥0.002mm and ≤0.05mm) in the fine earth fraction		
cfvo	✓	Volumetric fraction of coarse fragments (>2mm)		
cec	✓	Cation exchange capacity of the soil		
nitrogen	✓	Total nitrogen (N)		
soc	✓	Soil organic carbon content in the fine earth fraction		

Table S3. Different land cover datasets and "forest" and "cropland" definition used in our study.

Datasets	Description	"Forest" Definition	"cropland" definition	Spatial resolution	Reference
FI 2013-2017	Forest distribution map from the 2013-2017 China Forest Vegetation Survey project (2013FY111600) conducted by the Chinese Academy of Forestry Sciences. This map contains specific forest type and species information.	areas spanning more than 0.0667 ha, canopy coverage >20 %, arbor forest	-	1:1000000, Vector map	(10)
Hansen 2019	Results from time-series analysis of Landsat images in characterizing global forest extent and change from 2000 through 2019	tree cover>20%	-	30m	(15)
CNLUCC 2018	Regularly updated (every 5 years) from the late 1980s to 2018 using information sourced from high-resolution satellite remote-sensing images	(**a**) canopy coverage>10% (**b**) canopy coverage>30%	cropland including economic tree species	30m	(16)
MODIS 2019	The MODIS Land Cover Type Product (MCD12Q1 C6) supplies global maps of land cover at annual time steps and 500-m spatial resolution for 2001-present	canopy coverage>30% , tree height>2 m (including 5 forests and woody savannas 30-60% tree cover)	at least 40 % cultivated cropland	500m	(17)
GLC-FCS30 2020	Global land-cover product with fine classification system at 30 m using time-series Landsat imagery	canopy coverage>15%	(*) rainfed and irrigated cropland (*add) rainfed and cropland, including herbaceous	30m	(18)

			and orchard		
Globeland 30 2020	The GlobeLand30 data sets are freely available and comprise ten types of land cover, including forests, artificial surfaces and wetlands. They were extracted from more than 20,000 Landsat and Chinese HJ-1 satellite images.	canopy coverage>10%	cropland, including fruits and economic tree species	30m	(19)
ESA-CCI 2020	The European Space Agency (ESA) initiated a new program - namely the Climate Change Initiative (CCI) to develop a global monitoring dataset. The CCI-LC project delivers consistent global LC maps at 300 m spatial resolution on an annual basis from 1992 to 2020.	canopy coverage>15%	cropland	300m	(20)

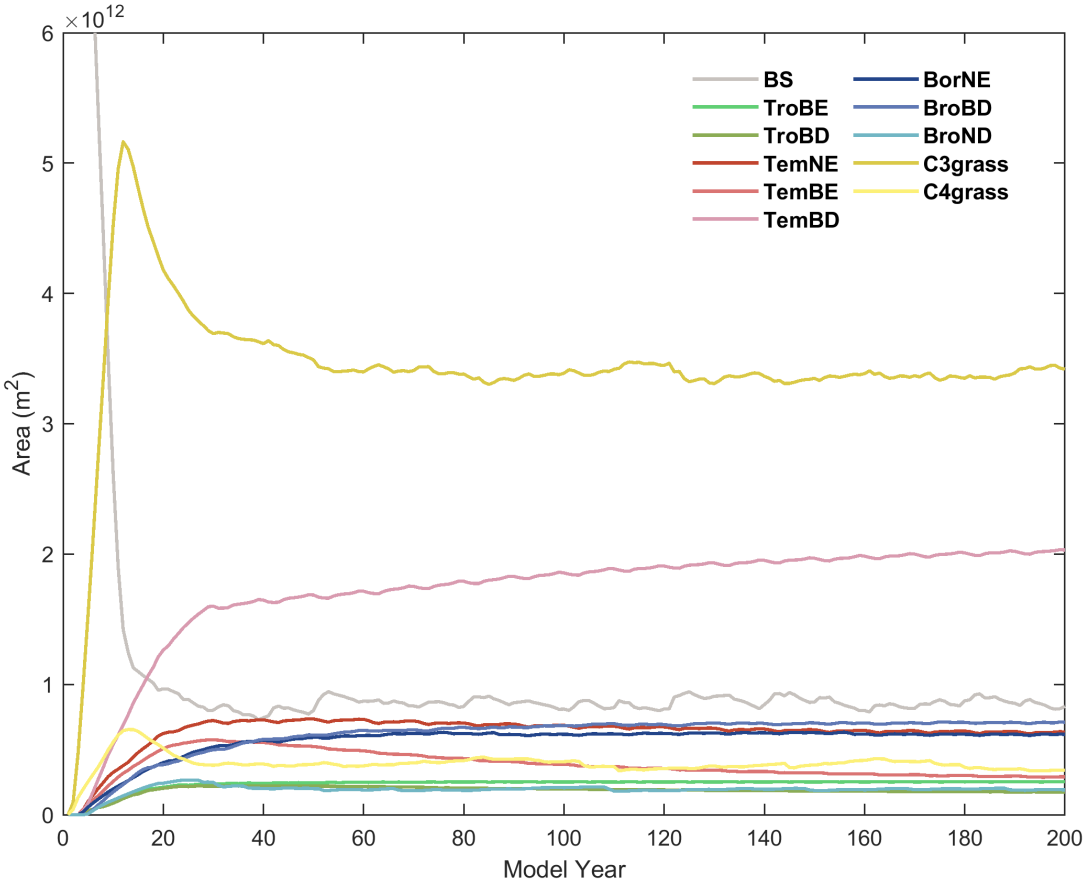

Figure S2. The area dynamics of bare soil, forests and grass in China over the 200 model years in ORCHIDEE. With the introduction of new plant functional types (PFTs), competition and succession between PFTs, proportions of ground area covered by forest (TroBE: Tropical Broadleaf Evergreen, TroBD: Tropical Broadleaf Deciduous, TemNE: Temperate Needleleaf Evergreen, TemBE: Temperate Broadleaf Evergreen. TemBD: Temperate Broadleaf Deciduous, BorNE: Boreal Needleleaf Evergreen, BroBD: Boreal Broadleaf Deciduous, BroND: Boreal Needleleaf Deciduous) and grass (C3grass, C4grass) ecosystems gradually increased, and finally reached equilibrium in the late stage of succession. The improved model version: ORC-HL-NVD was from Zhu et al. (8).

Figure S3. The current land type of the potential forestation opportunity in Fig.1*A*. Panel A corresponds to the high-confidence forestation areas in the main Fig. 1*A*, while B, C correspond to the medium- and low-confidence areas, respectively.

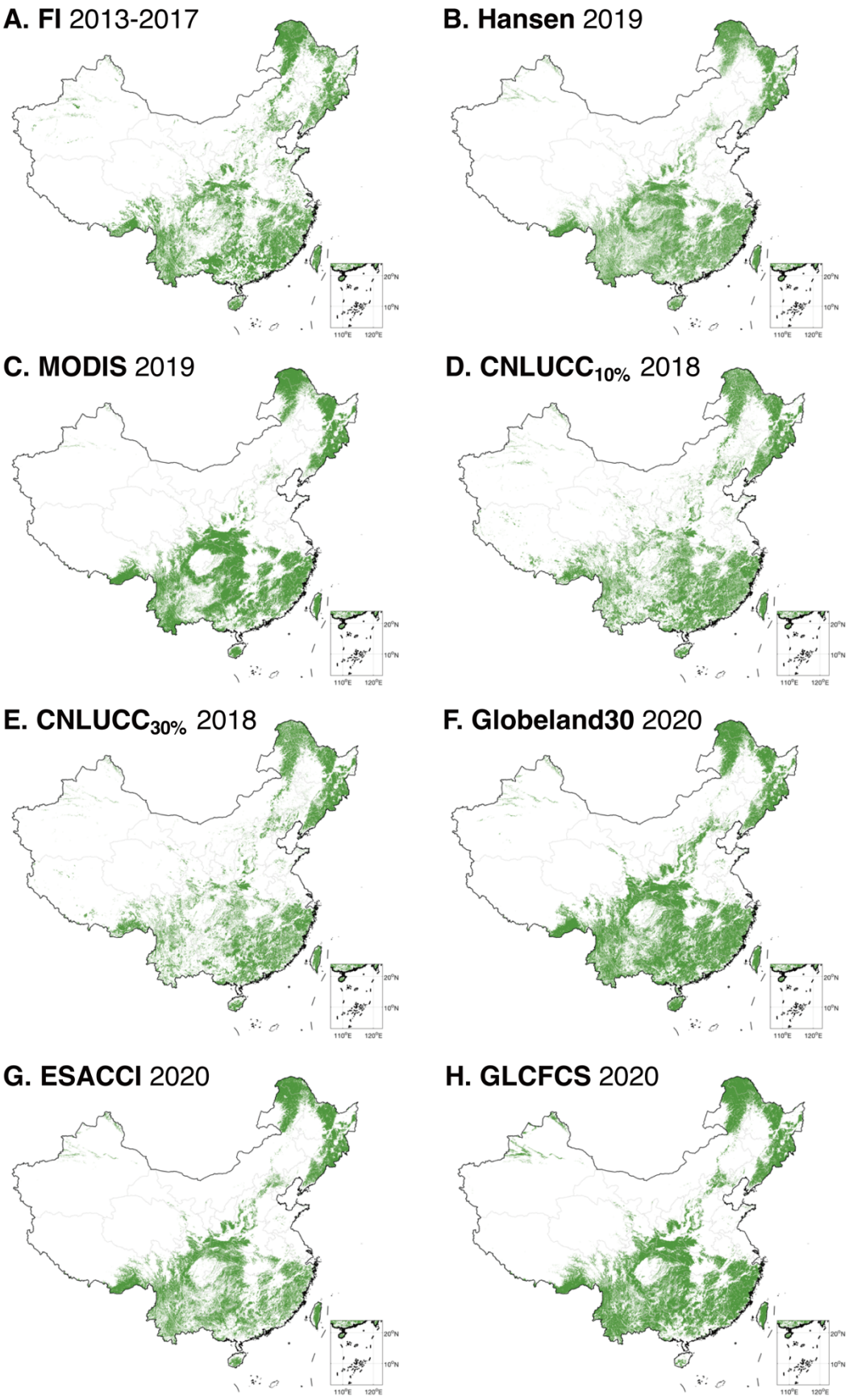

Figure S4. The distribution of existing forests in the 8 different landcover datasets in Fig. 1. Details of these datasets are given in Table S3. Each forest gridcell is shown in green.

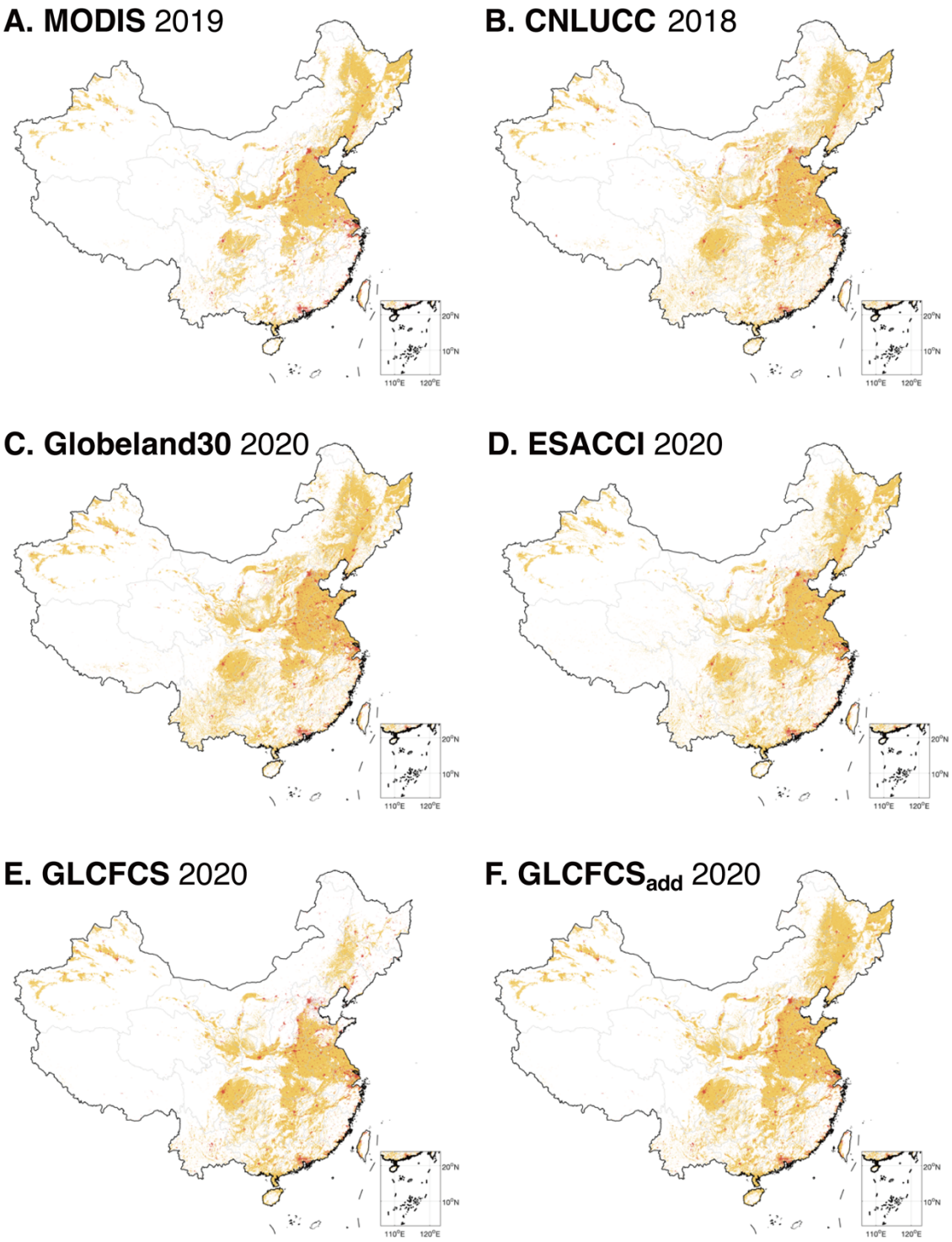

Figure S5. The distribution of cropland and urban land from the 6 different landcover datasets in Fig. 1. Details of these datasets are given in Table S3. Each cropland gridcell is orange, while the urban gridcells are red.

Table S4. The MaxEnt model accuracy (AUC value) of the 15 forest groups. The optimized classification thresholds were calculated as the average of the 'maximum test sensitivity plus specificity threshold' in the MaxEnt model.

No	Forest type	Training data AUC	Test data AUC	Optimized classification threshold
1	*Picea-Abies*	0.944	0.942	0.211
2	*Larix* spp.	0.836	0.835	0.338
3	*Pinus tabuliformis*	0.955	0.951	0.176
4	Other temperate *Pinaceae* spp.	0.923	0.920	0.229
5	*Cunninghamia lanceolata*	0.944	0.942	0.181
6	*Pinus massoniana*	0.935	0.933	0.285
7	Other warm *Pinaceae* spp.	0.844	0.843	0.341
8	Typical deciduous broadleaf forest	0.960	0.953	0.191
9	*Betula-Populus*	0.876	0.871	0.345
10	Other deciduous broadleaf forest	0.955	0.947	0.142
11	Typical evergreen broadleaf forest	0.954	0.952	0.223
12	Evergreen fast-growning forest	0.978	0.976	0.173
13	Other evergreen broadleaf forest	0.930	0.928	0.167
14	Temperate mixed forest	0.964	0.960	0.195
15	Subtropical mixed forest	0.985	0.982	0.129

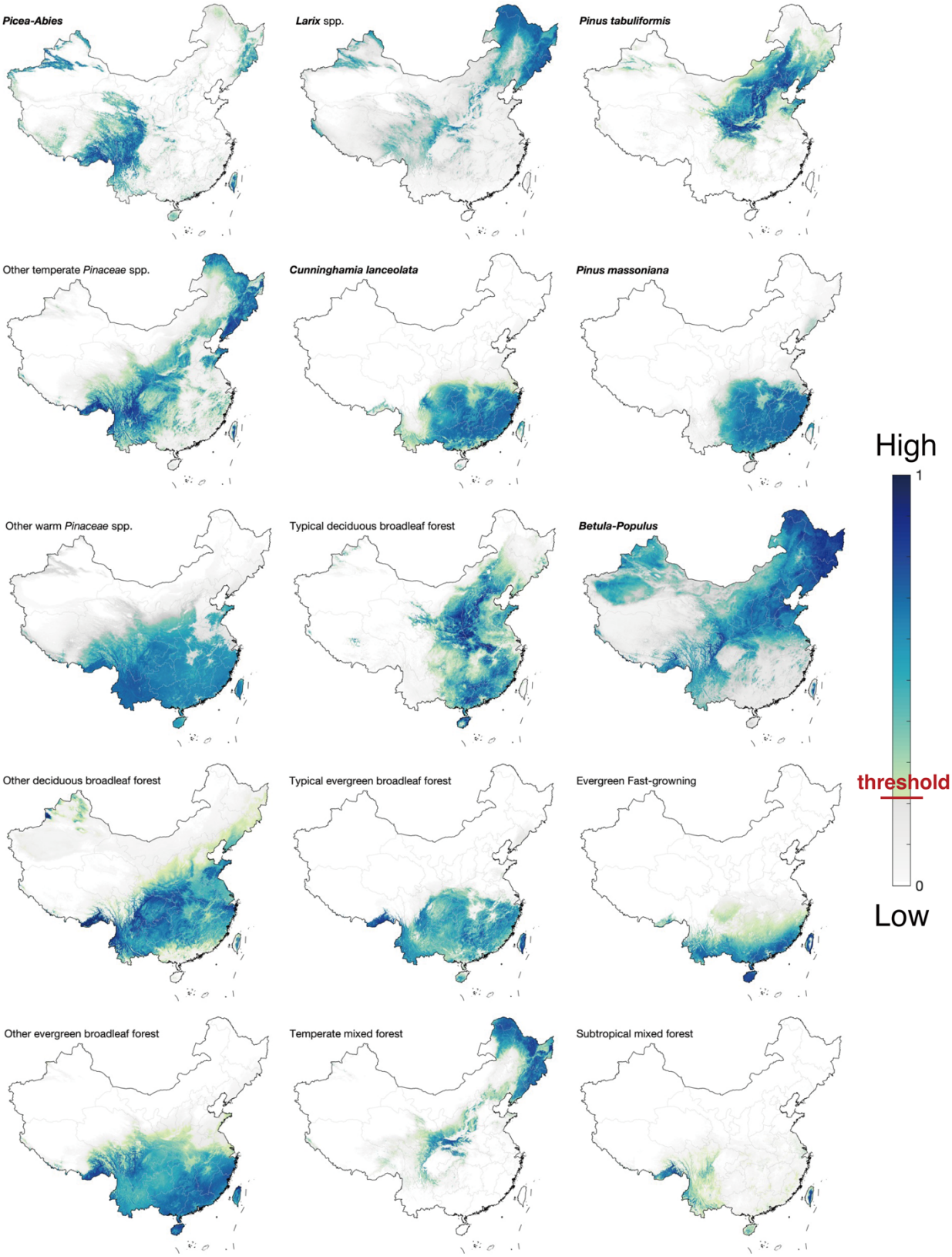

Figure S6. The potential suitability distribution for the 15 forest types. The gained value from the MaxEnt models represents the distribution probability, ranging between 0 and 1. With the gained value increasing, the probability of the presence of the forest type also increases. The threshold for each forest type is shown in Table S4.

Table S5. Updated parameters for the height-age equation $H = aA^{bT+cP+d}$ for each forest type. *A* is forest stand age, *H* is mean forest height, *T* is mean annual temperature, *P* is mean annual precipitation, and *a*, *b*, *c*, and *d* are parameters.

Forest Type*	a	b	c	d
Pinus tabulaeformis	0.20	0.043	3.6×10^{-4}	0.51
Pinus massoniana	0.20	0.023	-6.5×10^{-5}	0.92
Larix spp.	0.20	-0.017	6.5×10^{-4}	0.68
Picea spp.	0.20	-0.071	4.5×10^{-4}	0.84
Abies spp.	0.20	0.024	8.0×10^{-5}	0.68
Other Pinaceae spp.[a]	0.23	0.0050	-7.2×10^{-5}	1.11
Cunninghamia lanceolata	4.0	0.012	-6.8×10^{-5}	0.31
Other Taxodiaceae spp.[b]	0.44	0.099	-6.6×10^{-4}	0.80
Cupressaceae spp.	0.20	0.28	1.7×10^{-4}	0.43
Quercus spp.	3.8	-0.0096	1.9×10^{-4}	0.27
Betula spp.	0.85	-0.021	0.0010	0.13
Other hardwood broadleaf spp.[c]	0.59	2.1×10^{-4}	1.5×10^{-4}	0.68
Softwood broadleaf spp.	0.20	-0.025	0.0010	0.67
Broadleaf mixed forests	0.57	0.016	5.1×10^{-4}	0.58
Needle-leaf and broadleaf mixed forests	0.39	0.017	-3.7×10^{-4}	1.2

* We used the average forest height within each 1-km gridcell rather than the maximum tree height to avoid unreasonable overestimation.

** The classification of forest types here is from Zhang et al. (21), which is different from the dominant 15 forest type in the main text.

a. Pinaceae spp. except *Pinus tabulaeformis*, *Pinus massoniana*, *Larix* spp., *Picea* spp., and *Abies* spp.

b. Taxodiaceae spp. except *Cunninghamia lanceolata*.

c. Hardwood broadleaf spp. excluding *Quercus* spp. and *Betula* spp.

Table S6. The forestry goals of the Chinese government as in Fig. 3*B*. The bold font indicates targets for total forest area or total forest stock volume that were originally mentioned in official documents.

	Year	Total forest* (Mha)	Arbor forest (Mha, This study)	Forest stock volume (10^8 m^3)	ratio = Arbor / Total
Historical change	5th, 1994-1998	158.94	132.41		0.833
	6th, 1999-2003	174.91	142.78		0.816
	7th, 2004-2008	195.45	157.54		0.806
	8th, 2009-2013	207.69	166.55		0.802
Baseline year (9th)	2016 (2014-2018)	220.45	179.89	175.60	0.816
Targe 1 (a)	2025	**231.36**	188.79		
Targe 2 (b)	2030	247.58**	202.03	**197.21**	
Targe 3 (c)	2035	**249.60**	203.68		
Targe 4 (d)	2050	**294.72**	240.50		
100% Potential - high confidence	2062		257.79		

* "Total Forest area" reported in the national report including arbor forest, bamboo forest and national special shrubbery.

** "Total forest area" of Targe 2 = "Forest stock volume" of Targe 2 / "Forest stock volume" of Baseline × "Total forest area" of Baseline

a. The "Fourteenth Five-Year" Plan Outline for the Protection and Development of Forestry and Grassland: 24.1% of forest cover, by 2025

b. Building on Past Achievements and Launching a New Journey for Global Climate Actions (Xi Jinping, at the Climate Ambition Summit, 2020): increase the forest stock volume by 6 billion cubic meters from the 2005 level, by 2030

c. National Major Project for the Protection and Restoration of Important Ecosystems Master Plan (2021-2035): 26% of forest cover by 2035.

d. According to the document of the 2018 National Forestry Department Directors' Conference, the national forest cover may reach the world average of 30.7% in 2050.

The total forest area = target forest cover × 960 M ha (i.e., total land area of China)

Table S7. Nonlinear statistical forest biomass growth models used for 15 forest types (N = 1105). The bold terms in the equations were replaced by linear combinations of climate factors as (**α**×MAT+**β**×MAP+**γ**). MAT, mean annual temperature (°C); MAP, mean annual precipitation (mm). The fitted functions were obtained according to the method of Yao et al. (22), which was based on *in situ* field measurement data from Luo et al. (23).

No.	Forest Type	Number of Samples	Fitted function	α	β	γ
1	*Picea-Abies*	36	$y = \frac{\boldsymbol{a}}{1 + b \times age^{-c}}$	15.96	0.13	198.45
2	*Larix* spp.	62	$y = \frac{\boldsymbol{a}}{(1 + b \times age^{-d})^{c}}$	-1.47	0.15	67.83
3	*Pinus tabuliformis*	87	$y = \frac{\boldsymbol{a}}{1 + b \times e^{-c \times age}}$	-4.68	0.40	16.24
4	Other temperate *Pinaceae* spp.	41	$y = \frac{\boldsymbol{a}}{1 + b \times age^{-c}}$	-9.25×10^7	1.77×10^6	1.07×10^9
5	*Cunninghamia lanceolata*	268	$y = \frac{\boldsymbol{a}}{1 + b \times e^{-c \times age}}$	0.47	-0.01	506.92
6	*Pinus massoniana*	92	$y = \frac{\boldsymbol{a}}{1 + b \times e^{-c \times age}}$	-2.32	0.03	228.11
7	Other warm *Pinaceae* spp.	109	$y = \boldsymbol{a} \times e^{\frac{-b}{age}}$	12.92	0.09	-125.67
8	Typical deciduous broadleaf forest	55	$y = \frac{a}{1 + \boldsymbol{b} \times age^{-c}}$	5.26	-0.20	249.50

9	*Betula-Populus*	41	$y = \frac{\boldsymbol{a}}{(1 + b \times age^{-d})^{c}}$	3.71	-0.04	87.96
10	Other deciduous broadleaf forest	25	$y = \frac{a}{1 + \boldsymbol{b} \times age^{-c}}$	15.79	-0.51	628.13
11	Typical evergreen broadleaf forest	49	$y = \frac{\boldsymbol{a}}{(1 + b \times age^{-d})^{c}}$	64.32	0.62	-34.23
12	Evergreen fast-growing forest	95	$y = \frac{\boldsymbol{a}}{1 + b \times age^{-c}} + d$	0.88	0.03	-417.65
13	Other evergreen broadleaf forest	30	$y = \frac{\boldsymbol{a}}{(1 + b \times age^{-d})^{c}}$	-62.09	0.20	1416.68
14	Temperate mixed forest	45	$y = \frac{\boldsymbol{a}}{(1 + b \times age^{-d})^{c}}$	-10.51	1.50	-362.32
15	Subtropical mixed forest	70	$y = \frac{\boldsymbol{a}}{(1 + b \times age^{-d})^{c}}$	19.89	0.04	-201.15

Table S8. List of CMIP6 models used in this study. Detailed information on the CMIP6 experiments and these models can be obtained from O'Neill et al. (24) and from the official website https://pcmdi.llnl.gov/CMIP6/.

historical, SSP1-2.6, SSP2-4.5, SSP5-8.5					
No.	Models	No.	Models	No.	Models
1	ACCESS-CM2	12	EC-Earth3	23	KACE-1-0-G
2	ACCESS-ESM1-5	13	EC-Earth3-Veg	24	MIROC6
3	AWI-CM-1-1-MR	14	EC-Earth3-Veg-LR	25	MPI-ESM1-2-HR
4	BCC-CSM2-MR	15	FGOALS-f3-L	26	MPI-ESM1-2-LR
5	CAMS-CSM1-0	16	FGOALS-g3	27	MRI-ESM2-0
6	CAS-ESM2-0	17	FIO-ESM-2-0	28	NESM3
7	CESM2-WACCM	18	GFDL-ESM4	29	NorESM2-LM
8	CIESM	19	IITM-ESM	30	NorESM2-MM
9	CMCC-CM2-SR5	20	INM-CM4-8	31	TaiESM1
10	CMCC-ESM2	21	INM-CM5-0		
11	CanESM5	22	IPSL-CM6A-LR		
1%CO2 experiments: cVeg					
No.	Models	No.	Models	No.	Models
1	ACCESS-ESM1-5	9	EC-Earth3-Veg	17	MPI-ESM1-2-LR
2	BCC-CSM2-MR	10	GFDL-ESM4	18	NorCPM1
3	BCC-ESM1	11	INM-CM4-8	19	NorESM2-LM
4	CESM2	12	INM-CM5-0	20	NorESM2-MM
5	CESM2-WACCM	13	IPSL-CM5A2-INCA	21	SAM0-UNICON
6	CMCC-CM2-SR5	14	IPSL-CM6A-LR	22	TaiESM1
7	CMCC-ESM2	15	KIOST-ESM		
8	CanESM5	16	MPI-ESM-1-2-HAM		

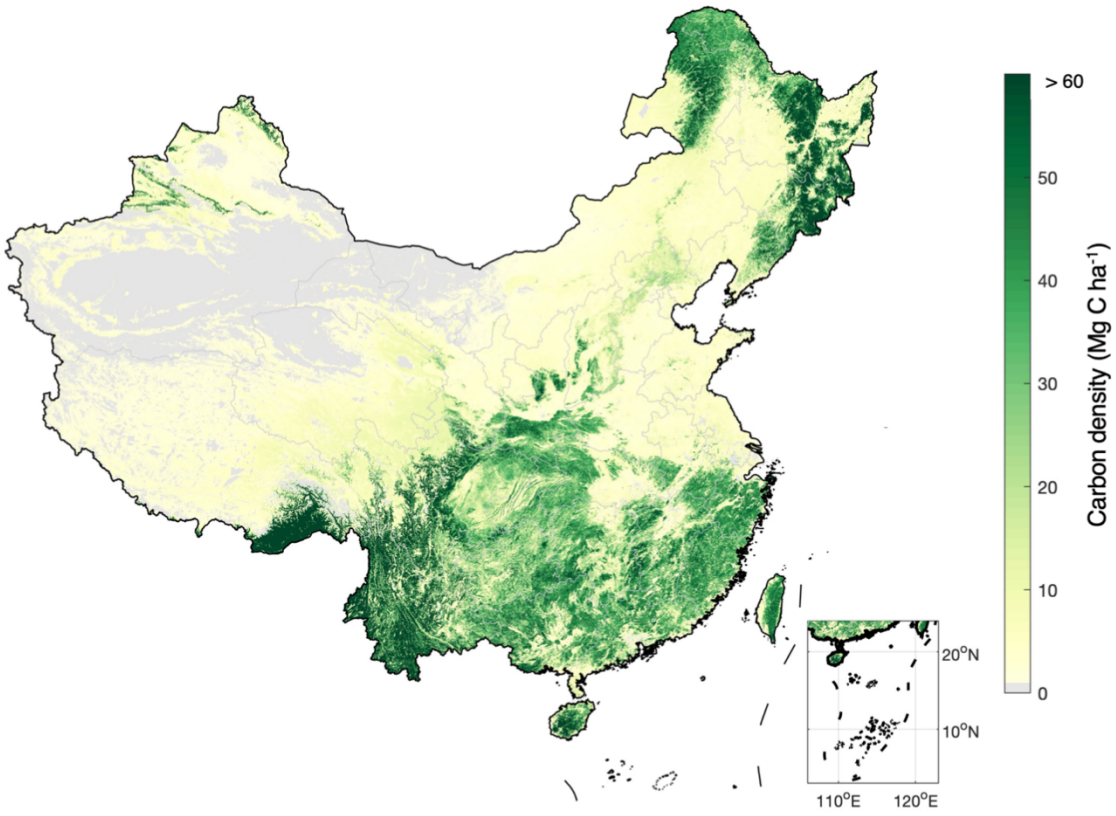

Figure S7. Living biomass carbon densities in 2010. Both aboveground and belowground biomass carbon are included. The original biomass map is from Spawn et al. (25).

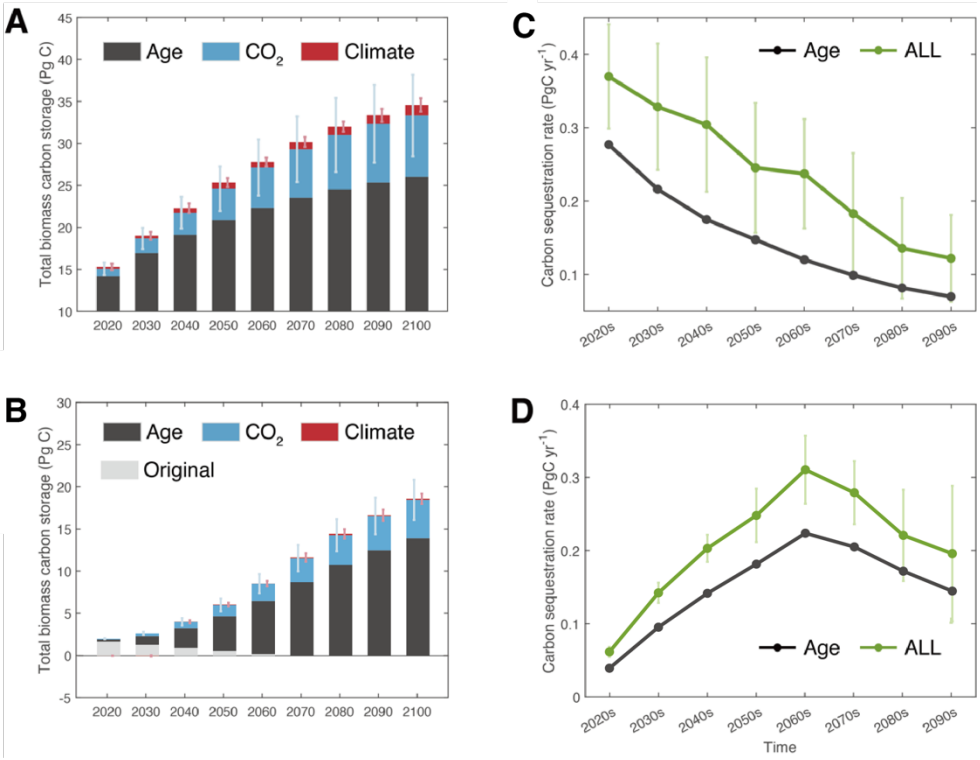

Figure S8. Biomass carbon storage and carbon sink dynamics of existing forests and newly established forests under the progressive forestation scenario for 2020–2100 under the SSP2-4.5 scenario. Same as Fig. 4 in the main text, but the climate change and CO_2 effects are based on the SSP2-4.5 scenario.

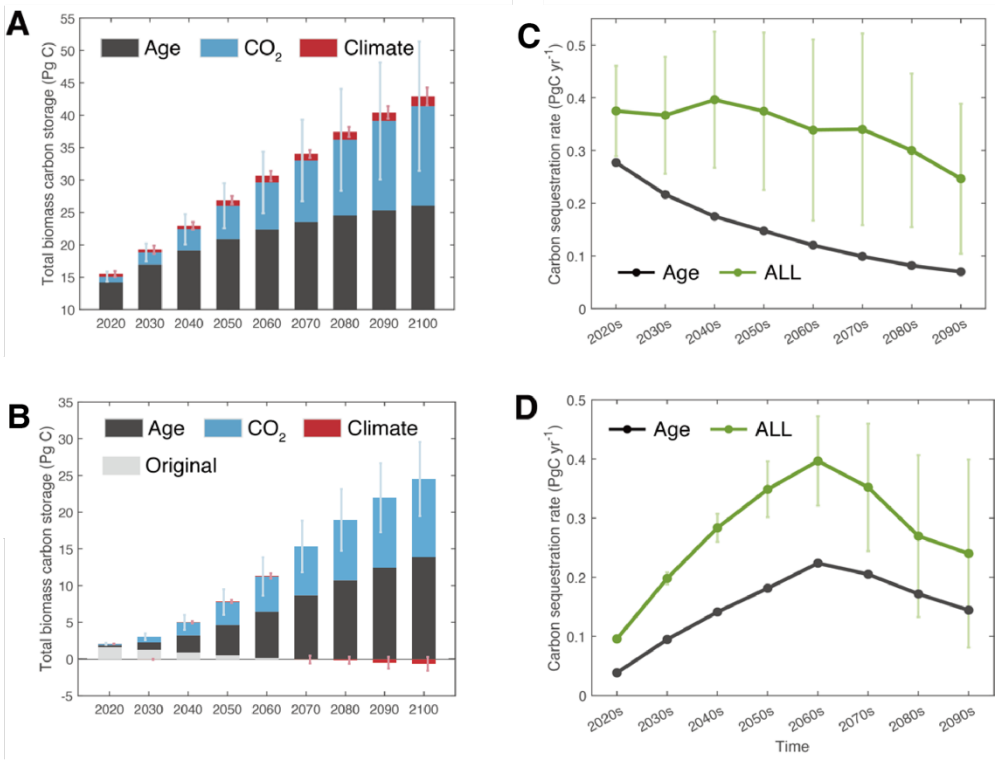

Figure S9. Biomass carbon storage and carbon sink dynamics of existing forests and newly established forests under the progressive forestation scenario for 2020–2100 under the

SSP5-8.5 scenario. Same as Fig. 4 in the main text, but the climate change and CO_2 effects are based on the SSP5-8.5 scenario.

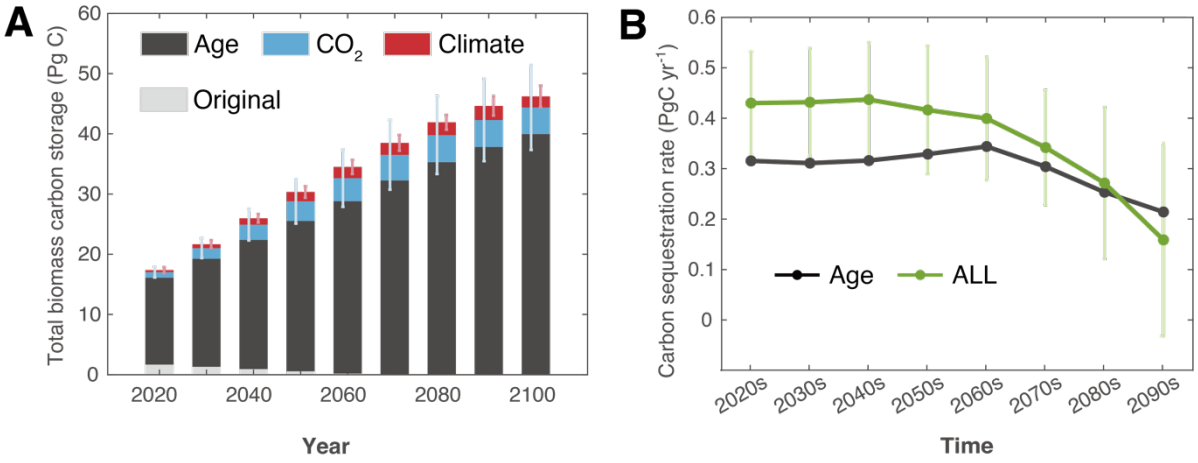

Figure S10. Biomass carbon storage and carbon sink dynamics of the entire forest ecosystems under the progressive forestation scenario for 2020–2100 under the SSP1-2.6 scenario. The results for entire forest ecosystems are the sum of the results for existing forests and newly established forests presented in Fig. 4.

Reference:

1. J.-F. Bastin, *et al.*, The global tree restoration potential. *Science* **365**, 76–79 (2019).
2. S. E. Fick, R. J. Hijmans, WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. *Int. J. Climatol.* **37**, 4302–4315 (2017).
3. L. Poggio *et al.*, Data from "SoilGrids250m 2.0." ISRIC. Available at https://files.isric.org/soilgrids/latest/. Deposited 31 January 2022.
4. M. Friedl, D. Sulla-Menashe, Data from "MCD12Q1 MODIS/Terra+Aqua Land Cover Type Yearly L3 Global 500m SIN Grid V006." NASA EOSDIS Land Processes DAAC. Available at https://doi.org/10.5067/MODIS/MCD12Q1.006. Accessed 6 January 2021.
5. National Forestry and Grassland Administration, *Forest Resource Report of China (2014–2018)* (Forestry Publishing House, 2019).
6. World Resources Institute, Data from "Atlas of Forest and Landscape Restoration Opportunities". https://www.wri.org/applications/maps/flr-atlas/#. Accessed 6 January 2021.
7. D. M. Olson, *et al.*, Terrestrial Ecoregions of the World: A New Map of Life on Earth. *BioScience* **51**, 933 (2001).
8. D. Zhu, *et al.*, Improving the dynamics of Northern Hemisphere high-latitude vegetation in the ORCHIDEE ecosystem model. *Geosci. Model Dev.* **8**, 2263–2283 (2015).
9. J. He, *et al.*, The first high-resolution meteorological forcing dataset for land process studies over China. *Sci. Data* **7**, 25 (2020).
10. S. Liu, Y. Chen, R. Zang, Data from "1:1000000 forest distribution map of China for 2013-2017." Geodata. Available at http://www.doi.org/10.12041/geodata.43370179401687.ver1.db. Deposited 19 May 2020.
11. J. L. Brown, J. R. Bennett, C. M. French, SDMtoolbox 2.0: the next generation Python-based GIS toolkit for landscape genetic, biogeographic and species distribution model analyses. *PeerJ* **5**, e4095 (2017).
12. S. J. Phillips, R. P. Anderson, R. E. Schapire, Maximum entropy modeling of species geographic distributions. *Ecol. Model.* **190**, 231–259 (2006).
13. S. Hempel, K. Frieler, L. Warszawski, J. Schewe, F. Piontek, A trend-preserving bias correction – the ISI-MIP approach. *Earth Syst. Dynam.* **4**, 219–236 (2013).
14. S. Peng, Y. Ding, W. Liu, Z. Li, 1 km monthly temperature and precipitation dataset for China from 1901 to 2017. *Earth Syst. Sci. Data* **11**, 1931–1946 (2019).
15. M. C. Hansen, *et al.*, High-Resolution Global Maps of 21st-Century Forest Cover Change. *Science* **342**, 850–853 (2013).
16. J. Liu, *et al.*, Spatiotemporal characteristics, patterns, and causes of land-use changes in China since the late 1980s. *J. Geogr. Sci.* **24**, 195–210 (2014).
17. M. Friedl, D. Sulla-Menashe, Data from "MCD12Q1 MODIS/Terra+Aqua Land Cover Type Yearly L3 Global 500m SIN Grid V006." NASA EOSDIS Land Processes DAAC. Available at https://doi.org/10.5067/MODIS/MCD12Q1.006. Accessed 6 January 2021.
18. X. Zhang, *et al.*, GLC_FCS30: global land-cover product with fine classification system at 30 m using time-series Landsat imagery. *Earth Syst. Sci. Data* **13**, 2753–2776 (2021).
19. C. Jun, Y. Ban, S. Li, Open access to Earth land-cover map. *Nature* **514**, 434–434 (2014).
20. European Space Agency, Data from "Land cover classification gridded maps from 1992 to present derived from satellite observations." C3S Climate Data Store (CDS). Available at https://doi.org/10.24381/cds.006f2c9a. Accessed 20 October 2020.

21. Y. Zhang, Y. Yao, X. Wang, Y. Liu, S. Piao, Mapping spatial distribution of forest age in China. *Earth Space Sci.* **4**, 108–116 (2017).
22. Y. Yao, S. Piao, T. Wang, Future biomass carbon sequestration capacity of Chinese forests. *Sci. Bull.* **63**, 1108–1117 (2018).
23. Y. Luo, X. Zhang, X. Wang, F. Lu, Biomass and its allocation of Chinese forest ecosystems. *Ecology* **95**, 2026–2026 (2014).
24. B. C. O'Neill, *et al.*, The Scenario Model Intercomparison Project (ScenarioMIP) for CMIP6. *Geosci. Model Dev.* **9**, 3461–3482 (2016).
25. S. A. Spawn, C. C. Sullivan, T. J. Lark, H. K. Gibbs, Harmonized global maps of above and belowground biomass carbon density in the year 2010. *Sci. Data* **7**, 112 (2020).