

Spatial database of planted forests in East Asia

Akane O. Abbasi, Xiaolu Tang, Nancy L. Harris, Elizabeth D. Goldman, Javier G. P. Gamarra, Martin Herold, Hyun Seok Kim, Weixue Luo, Carlos Alberto Silva, Nadezhda M. Tchepakova, Ankita Mitra, Yelena Finegold, Mohammad Reza Jahanshahi, César Iván Álvarez, Tae Kyung Kim, Daun Ryu, Jingjing Liang

Angaben zur Veröffentlichung / Publication details:

Abbasi, Akane O., Xiaolu Tang, Nancy L. Harris, Elizabeth D. Goldman, Javier G. P. Gamarra, Martin Herold, Hyun Seok Kim, et al. 2023. "Spatial database of planted forests in East Asia." *Scientific Data* 10 (1): 480.
<https://doi.org/10.1038/s41597-023-02383-w>.

Nutzungsbedingungen / Terms of use:

CC BY 4.0

Dieses Dokument wird unter folgenden Bedingungen zur Verfügung gestellt: / This document is made available under these conditions:

CC-BY 4.0: Creative Commons: Namensnennung

Weitere Informationen finden Sie unter: / For more information see:

<https://creativecommons.org/licenses/by/4.0/deed.de>





OPEN

Spatial database of planted forests in East Asia

DATA DESCRIPTOR

Akane O. Abbasi^{1,18}, Xiaolu Tang^{2,3,18}, Nancy L. Harris⁴, Elizabeth D. Goldman⁴, Javier G. P. Gamarra⁵, Martin Herold⁶, Hyun Seok Kim^{7,8,9,10}, Weixue Luo¹¹, Carlos Alberto Silva¹², Nadezhda M. Tchebakova¹³, Ankita Mitra¹, Yelena Finegold⁵, Mohammad Reza Jahanshahi^{14,15}, Cesar Ivan Alvarez¹⁶, Tae Kyung Kim⁷, Daun Ryu^{9,17} & Jingjing Liang¹✉

Planted forests are critical to climate change mitigation and constitute a major supplier of timber/non-timber products and other ecosystem services. Globally, approximately 36% of planted forest area is located in East Asia. However, reliable records of the geographic distribution and tree species composition of these planted forests remain very limited. Here, based on extensive *in situ* and remote sensing data, as well as an ensemble modeling approach, we present the first spatial database of planted forests for East Asia, which consists of maps of the geographic distribution of planted forests and associated dominant tree genera. Of the predicted planted forest areas in East Asia (948,863 km²), China contributed 87%, most of which is located in the lowland tropical/subtropical regions, and Sichuan Basin. With 95% accuracy and an F1 score of 0.77, our spatially-continuous maps of planted forests enable accurate quantification of the role of planted forests in climate change mitigation. Our findings inform effective decision-making in forest conservation, management, and global restoration projects.

Background & Summary

Planted forests are forest ecosystems established by artificial tree planting or seeding for the provision of income and goods, as well as for climate change mitigation and the restoration of ecosystem services and processes^{1,2}. According to the Food and Agriculture Organization of the United Nations (FAO)², planted forests globally increased by 41,000 km² per year between 2000 and 2020 and currently amounted to approximately 2,930,000 km². Today, FAO estimates that 36% of the world's planted forests are distributed in East Asian

¹Forest Advanced Computing and Artificial Intelligence (FACAI) Lab, Department of Forestry and Natural Resources, Purdue University, 715 W State St., West Lafayette, IN, 47907, USA. ²State Key Laboratory of Geohazard Prevention and Geoenvironment Protection, Chengdu University of Technology, Dongsanlu, Erxianqiao, Chengdu, 610059, Sichuan, P.R. China. ³College of Ecology and Environment, Chengdu University of Technology, Dongsanlu, Erxianqiao, Chengdu, 610059, Sichuan, P.R. China. ⁴World Resources Institute, 10 G Street N.E., Washington, DC, 20002, USA. ⁵National Forest Monitoring (NFM) Team, Forestry Division, Food and Agriculture Organization of the United Nations, Viale delle Terme di Caracalla, 00153, Rome, Italy. ⁶Helmholtz Center Potsdam GFZ German Research Centre for Geosciences, Section 1.4 Remote Sensing and Geoinformatics, Telegrafenberg, Potsdam, 14473, Germany. ⁷Department of Agriculture, Forestry, and Bioresources, Seoul National University, Seoul, 08826, Republic of Korea. ⁸Research Institute of Agriculture and Life Sciences, Seoul National University Seoul, Seoul, Republic of Korea. ⁹Interdisciplinary Program in Agricultural and Forest Meteorology, Seoul National University, Seoul, 08826, Republic of Korea. ¹⁰National Center for AgroMeteorology, Seoul, Republic of Korea. ¹¹Research Center of Forest Management Engineering of State Forestry and Grassland Administration, Beijing Forestry University, Beijing, 100107, P.R. China. ¹²Forest Biometrics, Remote Sensing and Artificial Intelligence Lab (Silvalab), School of Forest, Fisheries, and Geomatics Sciences, University of Florida, 342 Newins-Ziegler Hall, Gainesville, FL, 32611, USA. ¹³Sukachev Forest Institute, FRC KNC, Siberian Branch, Russian Academy of Sciences, Akademgorodok, 50/28, Krasnoyarsk, 660036, Russia. ¹⁴Lyles School of Civil Engineering, Purdue University, 550 Stadium Mall Drive, West Lafayette, IN, 47907, USA. ¹⁵Elmore Family School of Electrical and Computer Engineering, Purdue University, 465 Northwestern Ave, West Lafayette, IN, 47907, USA. ¹⁶Environmental Research Group for Sustainable Development (GIADES), Salesian Polytechnic University, Rumichaca y Moran Valverde, Quito, EC, 170702, Ecuador. ¹⁷Urban Forests Division, National Institute of Forest Science, Seoul, 02455, Republic of Korea. ¹⁸These authors contributed equally: Akane O. Abbasi, Xiaolu Tang. ✉e-mail: albeca.liang@gmail.com

countries, namely China, Japan, the Republic of Korea (ROK), and the Democratic People's Republic of Korea (DPRK)². In East Asia, a large proportion of forest area is planted forests (39% in China, 41% in Japan, 36% in ROK, and 16% in DPRK in 2020, according to FAO²), while other regions in the world remain well below 20% (19% in Africa, 7% in Europe, and 9% in the United States). Unlike Western countries, where planting was traditionally conducted for silvicultural practices, East Asian countries planted trees for varying purposes with local species and unique history^{3–11}.

East Asian countries have implemented a variety of tree-planting policies at different spatial and temporal scales. China leads all countries worldwide with the largest estimated plantation area of about 840,000 km². Since the end of the 1970s, China has established several afforestation projects, including the Three-North Forest Shelterbelt Program⁴, the Natural Forest Conservation Program (also known as Natural Forest Protection Program), and the Grain to Green Program (GGP; also known as the Sloping Land Conversion Program)^{5,6}. Currently, China has committed to preserving and expanding forest cover, aiming at mitigating soil erosion, air pollution, and climate change in the coming decades⁷. Although hundreds of tree species have been used for plantation establishment in China, a few species dominate the planted forests across the country, such as Chinese fir (*Cunninghamia lanceolata*) and eucalyptus (*Eucalyptus* spp.)⁸. In Japan, most planted forests were established after World War II to meet the growing demand for timber and other wood products. Thus, fast-growing and highly productive species, such as Japanese cedar (*Cryptomeria japonica*) and Hinoki cypress (*Chamaecyparis obtusa*), were extensively planted⁹. ROK underwent severe deforestation and forest degradation during World War II and the Korean War (1950–1953), followed by active conversion of forests to agricultural lands due to post-war poverty¹⁰. In response, the government implemented five National Forest Development Plans from 1973 to 2017. A variety of fast-growing species were planted during this period, and the successful recovery of healthy forests and ROK's sustainable management strategies are internationally recognized^{10,11}.

With active tree planting being implemented throughout the world for climate change mitigation, forest restoration, and biological conservation, it has become urgent to establish cost-effective guidelines for all ongoing and upcoming tree-planting projects. Assessment of the costs and benefits of planted forests, the key to the development of such cost-effective guidelines, is contingent on knowing where the existing planted forests are distributed^{12–16} and which tree species are planted¹⁷. The geospatial distribution of planted forests in East Asia still remains unclear due to a scarcity of complete, transparent, and publicly accessible data records. National governments have published some planted forest maps based on site visits, forest inventory, and satellite data. Yet, the spatial coverage is incomplete for Japan¹², and the map produced by the Chinese Forest Inventory remains unverified and largely inconsistent with independent studies^{13,14}. The existing large-scale maps of planted forests are based on inconsistent data sources with varying reliability and scale¹³ or solely based on satellite images¹⁴. Because of these differences in spatial extent, underlying data sources, and methods in existing datasets, a database that provides complete, consistent, and ground-truth-based records of the geographic distribution of planted forests and associated dominant tree species for East Asia constitutes a consistent and harmonized product.

Here, we produced the spatial database of planted forests in East Asia at a 1-km resolution and identify dominant tree species in these planted forests to the genus level. Our planted forest map encompasses forests of all ages planted for various purposes, including forest restoration, commercial plantation, and disaster prevention. These mapping products are based on ensemble machine learning models, data fusion, and multi-source data of planted forests. Our multi-source data comprised ~7,000 ground-truth inventory plots in China, five independent digitized maps across the study region, as well as 57 auxiliary datasets and layers, including satellite data such as the Global Ecosystem Dynamics Investigation (GEDI)¹⁸ and Moderate Resolution Imaging Spectroradiometer (MODIS) data to account for potential differences in forest structure and vegetation characteristics between planted and natural forests. In addition to the main products, we also estimated the upper and lower bounds of potential planted forest extent to account for the uncertainty associated with the varied quality of multi-source training data. With previous records of planted forests being inconsistent in resolution, quality, and accessibility, our map provides a complete, consistent, and *in situ* data-based estimation of the extent and species distribution of planted forests in East Asia.

Methods

To estimate the spatial distribution of planted forests over East Asia, we integrated multi-source planted-natural forest data from multiple *in situ* inventories and digitized data sources in a high-level data fusion algorithm (Fig. 1). For each observation, we first created a response variable explicitly labeled as either “planted” or “natural” forests. We then obtained data on 57 potential predictor variables encompassing forest structure, vegetation characteristics, bioclimate, topography, anthropogenic information, and soil characteristics, and merged these layers with the response variable layer based on spatial coordinates. The training dataset was then masked to the forested area in 2020 and separated into three biomes based on the Nature Conservancy Terrestrial Ecoregions map¹⁹. For each biome, we selected the optimal machine learning classification model and fine-tuned hyperparameters. Finally, we mapped planted forest distribution and the distribution of the dominant tree species in these forests to the genus level. Our study area covers China, Japan, ROK, and DPRK.

Data fusion. We collected and integrated *in situ* and digitized planted-natural forest data from multiple independent sources using a high-level data fusion algorithm (Fig. 1). Observations from China came from published literature^{20–265} (Fig. 2a), which included 2,542 and 4,394 *in situ* records of confirmed locations of planted and natural forests, respectively. The *in situ* planted forest observations include the plantation of commercial species, such as pine (*Pinus* spp.) and eucalyptus (*Eucalyptus* spp.), and forests planted for restoration purposes. We also obtained the national planted forest map of China (Fig. 2b)¹⁵, which depicts the distribution of planted forests in 2000. Data specific to Japan was obtained from the national vegetation map created based on site visits and

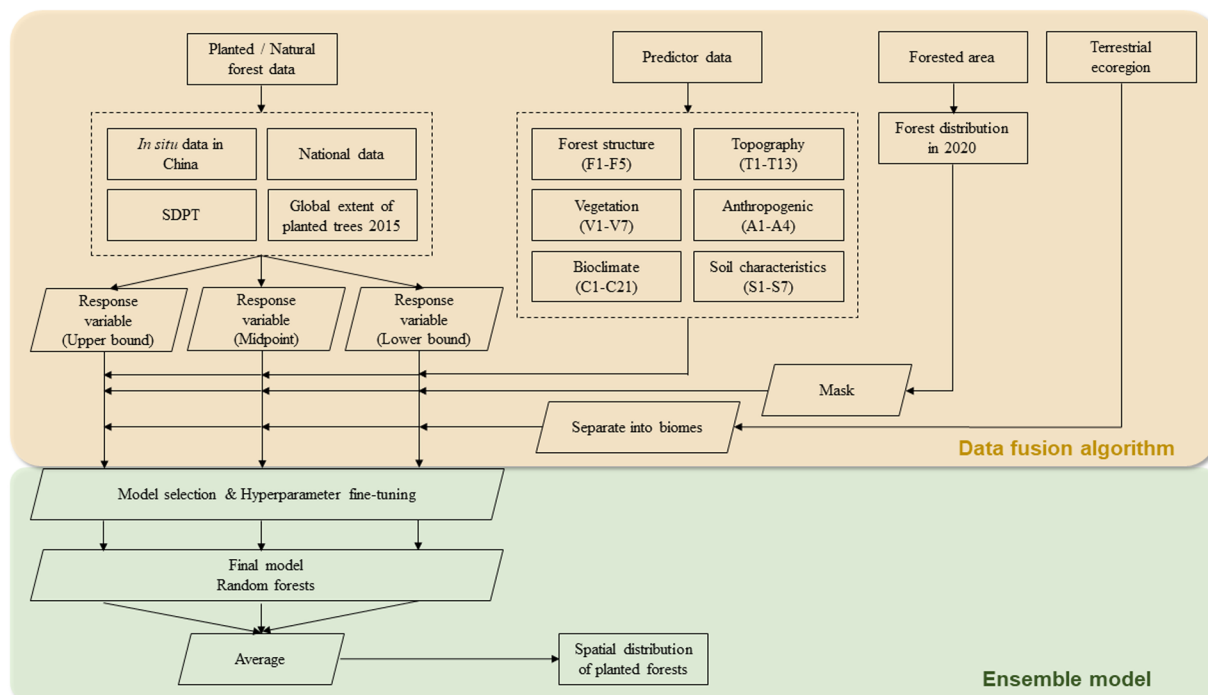


Fig. 1 Workflow for developing the spatial database of planted forests. The top section (yellow) represents the data fusion algorithm we used to integrate multi-source data into coherent training datasets. The bottom section (green) represents the ensemble model we developed to predict the spatial patterns of planted forests.

satellite images, where “planted forest” was one of the attributes of vegetation types (Fig. 2b)¹². This “planted forest” attribute includes restoration-oriented forests composed of broadleaf species, commercial forests dominated by productive species like Japanese cedar (*Cryptomeria japonica*) and Hinoki cypress (*Chamaecyparis obtusa*), and disaster prevention planting, such as Japanese black pine (*Pinus thunbergii*) from coastal erosion and tropical species (e.g., *Acacia confusa*) as windbreaks. The national vegetation map has been gradually developed and improved since 2005. Finally, data specific to ROK was a polygon map of planted and natural forests from the national forest cover map (Fig. 2b)¹⁶. The ROK maps depict the distribution of planted and natural forests from 2009 to 2013, depending on the province. In addition to the country-specific data, we obtained the Spatial Database of Planted Trees covering China, Japan, and ROK (SDPT version 1.0; Fig. 2c)¹³ and a global extent of planted trees 2015¹⁴, which includes the land use classes of planted forest, woody plantations, and agroforestry of the global forest management map²⁶⁶ (Fig. 2d). There is no data specific to DPRK used in this study due to the lack of available data.

To prepare a training dataset for machine learning classification models, we prepared a 0.009° by 0.009° grid (approximately 1 km²) for the study region in East Asia. National planted forest maps of China¹⁵, Japan¹², and ROK¹⁶, as well as SDPT¹³ and the Global Extent of Planted Trees¹⁴ were extracted to the centroid of each grid cell using the “sf” or “raster” packages in R^{267,268}. China’s *in situ* observations were associated with each grid cell by taking the majority vote of *in situ* points within each grid cell to determine whether that cell is a planted or natural forest. Grid cells with a 50/50 vote were removed from the training dataset. We then derived the response variable – a label of “planted” or “natural” forest – based on these underlying datasets following the Quality-Oriented Data Integration (QODI).

Quality-oriented data integration (QODI). Since the underlying datasets differed in data sources and estimation methods, we developed a quality-oriented data integration approach in which the response variable was defined in three different levels of integration (Fig. 3). For each level of integration, we trained a separate set of machine learning models, so that we can quantify the potential range in estimated planted forest areas.

The first level of integration took the most conservative approach in deriving the lower bound of our estimation. Since China’s *in situ* observations^{20–265}, Japan’s national vegetation map¹², and ROK’s national planted forest map¹⁶ were largely based on *in situ* observations, we labeled a unit forest area (i.e., grid cell) as planted if and only if the grid cell was identified as a planted forest by either of these *in situ*-based datasets or identified by at least three other datasets as a planted forest.

The second level of integration took a midway approach in which, in addition to planted forests identified in the first level of integration, a given grid cell was also labeled as a planted forest if two out of the national planted forest maps of China¹⁵, SDPT¹³, and Global Extent of Planted Trees¹⁴ datasets agreed so.

The third level of integration took the most liberal approach in deriving the upper bound of our estimation, in which we assumed all underlying data sources were equally reliable and labeled a given grid cell as planted forest if it was identified as a planted forest by either of these datasets.

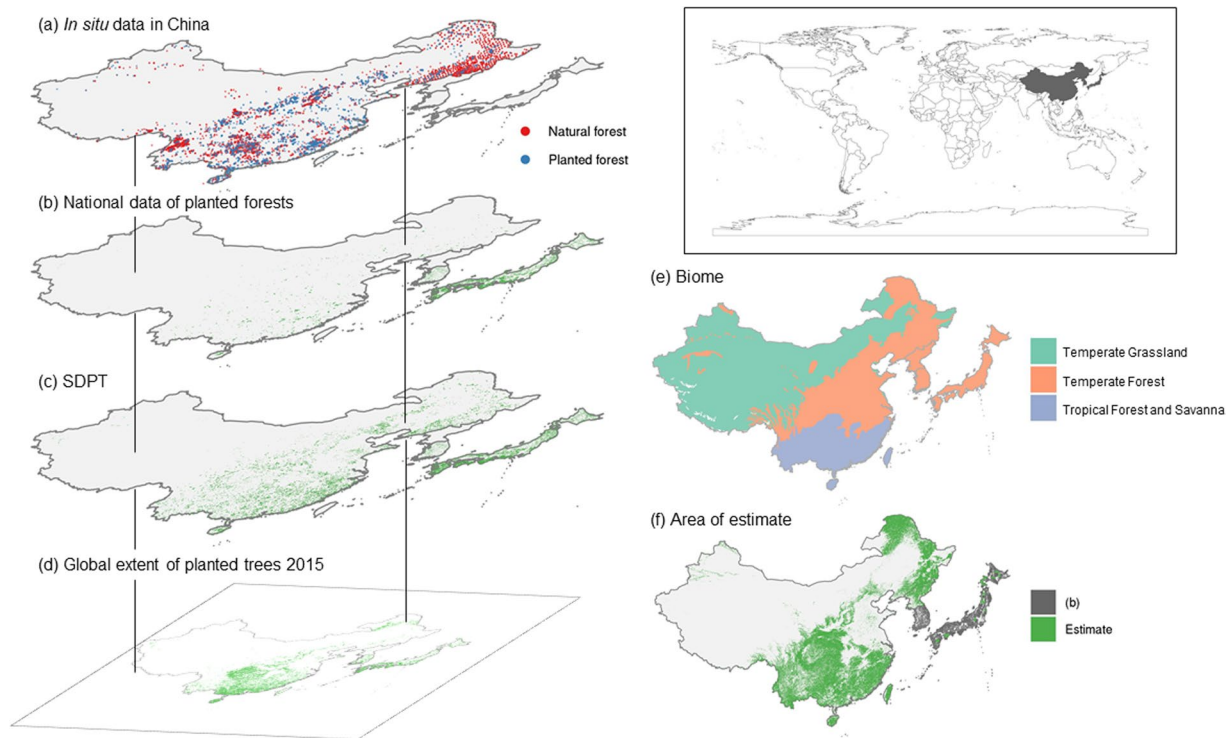


Fig. 2 Training data consists of a series of *in situ* and digital maps of planted-natural forest data from multiple independent sources. **(a)** The *in situ* data in China encompass 2,542 and 4,394 ground observations^{20–265}, which represent confirmed locations of planted and natural forests, respectively, by previously published articles. **(b)** National maps of planted forests were obtained for China¹⁵, Japan¹², and ROK¹⁶. **(c)** The Spatial Database of Planted Trees (SDPT version 1.0)¹³. **(d)** An estimated Global Extent of Planted Trees 2015¹⁴. **(e)** Distribution of the three biomes in our study area. We developed a machine learning classification model for each biome to predict planted forests. Note that forests are distributed in the Temperate Grassland according to the FAO’s definition of forest (≥ 5 m tree height)^{2,270} although the area is limited. **(f)** Distribution of planted forests was estimated mainly for China, DPRK, and small areas in Japan. For the ROK and the majority of Japan, the national planted forest maps^{12,16} **(b)** were used as a final label.

We also compiled 57 predictor variables for the supervised learning of the classification models (Fig. 1, Supplementary Table S1). The predictor variables consisted of five forest structure attributes^{269,270}, seven MODIS-derived vegetation characteristics, 21 bioclimatic attributes^{271–274}, 13 topographic attributes²⁷⁵, four anthropogenic attributes^{276–279}, and seven soil attributes²⁸⁰. We obtained four forest structure attributes from the most recent Global Ecosystem Dynamics Investigation (GEDI) dataset, namely canopy height (rh100), plant area index (pai), foliage height diversity (fhd_normal), and total canopy cover (cover) (see Supplementary Table S1)^{18,269}. We downloaded the raw footprint-level GEDI data (L2B), among which only full-power lasers were used in this study to ensure the accuracy of the measurement. GEDI data was processed using the “rGEDI” package in R²⁸¹. Another forest structure attribute, tree height²⁷⁰, represents the 90th or 95th percentile of energy return height relative to the ground.

We extracted predictor variables to the centroid of each grid cell using the “sf” or “raster” packages in R^{267,268}. GEDI footprint-level data was associated with each grid cell by taking the mean value of each attribute. We kept only grid cells with a minimum of 5 m tree height in accordance with FAO’s definition of “forest”^{2,270}. Our final training dataset encompassed more than 1.5 million grid cells for the upper bound dataset, 1.0 million grid cells for the midpoint dataset, and 0.9 million grid cells for the lower bound dataset, consisting of one response variable labeled as either “planted” or “natural” and 57 predictor variables. Finally, to account for the differences in terrestrial ecoregions, we divided the overall training dataset into three biomes (Fig. 2e). Based on the global terrestrial biome map¹⁹, Temperate Grassland/Savanna and Montane and Montane Flooded Grassland were grouped into “Temperate Grassland”. Temperate Broadleaf and Mixed and Temperate Conifer were grouped into “Temperate Forest”, and Tropical Moist, Tropical Dry, and Tropical Grassland/Savanna were grouped into “Tropical Forest and Savanna.” The three biomes remained separated for the upper bound dataset, but Temperate Grassland and Temperate Forest were merged for the midpoint and lower bound datasets to form the “Temperate Forest and Grassland” biome due to low sample size in Temperate Grassland.

For mapping purposes, we prepared another 0.009° by 0.009° grid (approximately 1 km²), covering forested area (≥ 5 m tree height)^{2,270} in the study region with all predictor variables (new data; Fig. 2f). We chose the resolution 0.009° to align with most of the predictor variables (Supplementary Table S1). After a machine learning classification model was trained, estimation was made for each grid cell of this new data. For ROK and a

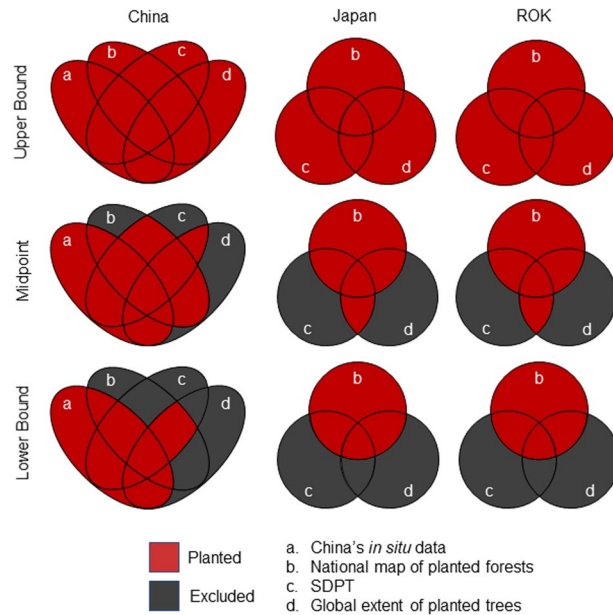


Fig. 3 The response variable (“planted” or “natural” forest) was defined in a quality-oriented data integration approach based on multiple underlying data sources. Underlying datasets a–d correspond to Fig. 2a–d. Upper and lower bound models represent the most liberal and conservative approaches in labeling planted forest, respectively. The grey area was removed from the respective training dataset. All areas outside of the Venn diagrams were labeled natural forest. DPRK is not included in this figure due to the absence of training data associated with the country.

majority of areas in Japan, however, we utilized the existing planted forest maps, namely the national forest cover map of ROK¹⁶ and the national vegetation map of Japan (Fig. 2b)¹², respectively, to label the grid cells. Since reliable planted forest data already exist for these areas, we used our estimation only for the remaining areas in China, DPRK, and a small portion of Japan (Fig. 2f). Nevertheless, the existing data for ROK and a majority of areas in Japan were converted to the 0.009° resolution within the forested area for consistency. For the areas where our estimation is used, we imputed missing values in predictor variables of the new data using the “Hmisc” package in R²⁸² to provide a spatially continuous map. For the GEDI attributes (Supplementary Table S1), however, we imputed missing values by training random forest (RF) models (see below for details of RF) with seven MODIS attributes due to a large number of missing values (22%, 34%, and 44% of the sample size for the upper bound, midpoint, and lower bound dataset, respectively). For the midpoint and lower bound datasets, we used the average predicted values from 10 repetitions of random forest models using 200,000 data points to minimize computational time (Table 1). To assess the performance of the RF model in imputing missing values in GEDI attributes, we performed cross-validation using bootstrapping. For the upper bound dataset, we randomly sampled the dataset into the training (90%) and testing (10%) sets with replacement. For the midpoint and lower bound datasets, we randomly sampled 200,000 data points for the training sets with replacement, and the remaining was used as the testing dataset (Table 1). Based on 20 random iterations, we calculated the 95% confidence interval (CI) of the root mean square error (RMSE) and R-squared (R^2). We calculated a 95% CI using the $t_{0.975}$ value with 19 degrees of freedom.

Ensemble machine learning model. We developed an ensemble model to estimate the spatial distribution of planted forests, with three candidate machine learning models: RF, support vector machines (SVM), and XGBoost. RF is a non-parametric ensemble learning approach²⁸³, which combines a variant of decision trees and an additional level of randomness by bootstrapping sub-data and different sets of predictor variables to mitigate potential multicollinearity issues often encountered in multidimensional machine learning models²⁸⁴. We used the “randomForest” package in R²⁸⁵. SVM is a supervised learning model which constructs a hyperplane or set of hyperplanes in a high- or infinite-dimensional space to help data analysis²⁸⁶. We used the “e1071” package in R²⁸⁷. XGBoost is a gradient-boosted decision tree machine learning, designed to accommodate large data at high speed. We used the “xgboost” package in R²⁸⁸. The three candidate models are frequently used in ecological and biological research with satisfactory performance^{266,289}. Other potential candidate models include artificial neural networks, k-nearest neighbor, Naïve Bayes, *etc.*, which are not necessarily superior²⁹⁰. All modeling processes were conducted in R²⁹¹.

To assess the performance of the three candidate models in estimating planted forests, we conducted cross-validation using bootstrapping. Due to data size, we randomly sampled 50,000 points (25,000 for each class) for the upper bound and midpoint datasets and 80% of the sample points for the lower bound dataset for each of the ten repetitions to create the training set and the rest composed the testing set (Table 1). Default hyperparameter values were used for the three candidate models. Based on 10 iterations, we calculated the

Model objective	Machine learning	Level of integration	Response	Predictors	Training size	Testing size	Iterations
To impute GEDI attributes in new data for mapping purposes (final model)	Random forests (RF)	Upper Bound	Each GEDI attribute	Seven MODIS attributes	All	NA	1
		Midpoint			200,000 data points	NA	10
		Lower Bound			200,000 data points	NA	10
To impute GEDI attributes in new data for mapping purposes (cross-validation)	RF	Upper Bound	Each GEDI attribute	Seven MODIS attributes	90%	10%	20
		Midpoint			200,000 data points	Remaining points	20
		Lower Bound			200,000 data points	Remaining points	20
To determine the best machine learning model to predict planted forest	RF, support vector machines (SVM), and XGBoost	Upper Bound	Planted or natural	All listed in Supplementary Table S1	50,000	Remaining points	10
		Midpoint			50,000	Remaining points	10
		Lower Bound			80%	20%	10
To fine-tune hyperparameters of the final RF model to predict planted forest	RF	Upper Bound	Planted or natural	All listed in Supplementary Table S1	50,000	Remaining points	10
		Midpoint			50,000	Remaining points	10
		Lower Bound			80%	20%	10
To predict planted forest (final model)	RF	Upper Bound	Planted or natural	All listed in Supplementary Table S1	All	NA	20
		Midpoint			All	NA	20
		Lower Bound			All	NA	20
To predict dominant tree genus (final model)	RF	Upper Bound	Genus	All listed in Supplementary Table S1 except for roadless areas and GEDI	All	NA	1
		Midpoint			All	NA	1
		Lower Bound			All	NA	1
To predict dominant tree genus (cross-validation)	RF	Upper Bound	Genus	All listed in Supplementary Table S1 except for roadless areas and GEDI	90%	10%	100
		Midpoint			90%	10%	100
		Lower Bound			90%	10%	100

Table 1. Summary of tasks conducted in this study.

95% CI of classification accuracy and F1 score. We calculated a 95% CI using the $t_{0.975}$ value with 9 degrees of freedom. Classification accuracy shows the proportion of overall correct prediction. While accuracy is the most widely used and intuitive evaluation metric of a classification problem, it overestimates the performance of imbalanced data. F1 score is an equal measure of precision and recall and is more appropriate for imbalanced data²⁹². Precision represents the correct prediction of the positive class (*i.e.*, planted) among all positive predictions, and recall represents the correct prediction of the positive class among all actual positive cases²⁹³. Since precision and recall are in an inverse relationship, the combined metric, F1 score, provides a better evaluation perspective of incorrectly predicted cases. Using both accuracy and F1 score, we present a suite of evaluation metrics of our candidate models for both correct and incorrect predictions of an imbalanced dataset. Other potential evaluation metrics include Cohen's Kappa. However, we did not use it in our study due to the controversy of its use²⁹⁴. Compared with SVM and XGBoost, the RF model was 0.7–8.1% more accurate in terms of overall classification accuracy and 1.4–4.5% more reliable in terms of F1 score (Fig. 4). Thus, we chose RF as the final model.

To improve the performance of the model while minimizing the time it takes to compute, we adjusted two hyperparameters of the RF algorithm: the number of decision trees and the number of predictor variables. Similar to the cross-validation described above, we randomly sampled 50,000 points (25,000 for each class) for the upper bound and midpoint models and 80% of the sample points for the lower bound model for each of the ten repetitions to assess RF performance using different hyperparameter values (Table 1). Specifically, we calculated the classification accuracy and F1 score for different hyperparameter values. Based on 10 iterations, we chose the number of 100 decision trees for the upper bound and midpoint models and 200 for the lower bound model where both accuracy and F1 score converged (Fig. 5). We used the default number of predictor variables (seven) for all biomes for the upper bound model. We chose 26 and 42 for Temperate Forest and Grassland and Tropical Forest and Savanna, respectively, for the midpoint model (Fig. 6). We chose 20 and 40 for Temperate Forest and Grassland and Tropical Forest and Savanna, respectively, for the lower bound model (Fig. 6).

For the final RF model, we ensured that the training set had an equal number of points for each class (*i.e.*, 50% planted forest and 50% natural forest) by randomly under-sampling the dominant class. The prediction of our classification model was the percent planted forest based on how many decision trees returned the “planted” prediction. We built 20 models to derive the mean percentage for each biome and model (upper bound, midpoint, and lower bound) (Table 1). Finally, we calculated the mean percentage of the three models as a final value, while upper and lower bounds serve as a potential range (Fig. 7). Grid cells with a predicted percentage $\geq 50\%$ are considered planted forest (Fig. 8). Using the spatially continuous dataset of 57 predictor variables (see *Data fusion*), we created a map covering the entire forested area in East Asia using model prediction.

Mapping dominant tree species of the planted forests. Over the planted forest expanse in East Asia identified by the final RF classification model, we predicted the dominant tree species (to the genus level) of the planted forest for each criterion (Fig. 9). For the training set, we combined 2,481 *in situ* records in China^{20–265} with the tree-level records of Japan²⁹⁵ and ROK²⁹⁶ National Forest Inventories (NFI). Specifically, we calculated importance value for each species for each NFI plot within the predicted planted forest expanse and identified the species

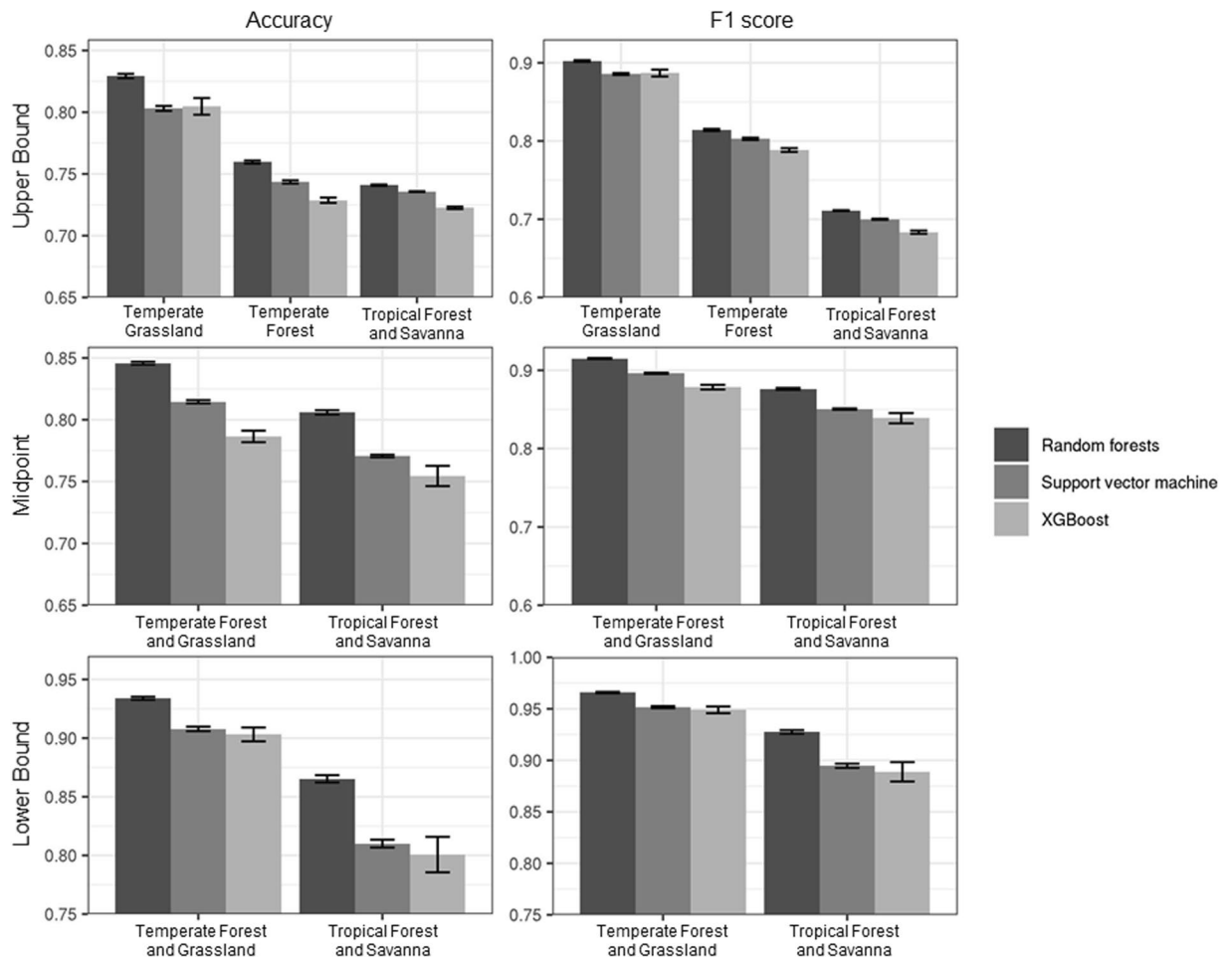


Fig. 4 Performance of three candidate machine learning models to map planted forests. Classification accuracy and F1 score of random forest (RF), support vector machine (SVM), and XGBoost imputation models are shown. Mean values from 10 repetitions and 95% confidence intervals are shown for each biome. RF outperformed SVM and XGBoost in all cases, and thus RF was used to model planted forests in our study.

with the highest importance value as the dominant species for the given plot. Importance value is the sum of the percent basal area and the percent number of individuals of each species and represents the overall dominance of the species^{297,298}. After identifying the dominant species for each NFI plot, we aggregated the plots into the 0.009° by 0.009° grid cells by taking the majority vote of the dominant species. We retained the genus names of the dominant species, and only genera with 60 or more samples were included to ensure a sufficient size of training data.

We trained an RF classification model using the same package in R, with the default hyperparameter setting and an identical set of predictor variables, except for roadless areas and GEDI attributes due to a substantial number of missing values (86% and 34% of the sample size, respectively). We ensured that the training set had an equal number of points for each class (*i.e.*, genus) by combining random under-sampling and oversampling using the “UBL” package in R²⁹⁹. To assess the performance of the RF model in mapping dominant genera across the planted forest expanse in East Asia, we performed a 90/10 cross-validation using bootstrapping. In each iteration, we used stratified sampling to split the entire training dataset into the training (90%) and testing (10%) sets using the “caret” package in R³⁰⁰ and conducted a combination of under-sampling and oversampling of the training set to address the class imbalance (Table 1). Based on 100 random iterations, we calculated the 95% CI of overall classification accuracy and precision, recall, and F1 score for each class.

Data Records

The spatial database of planted forests consists of maps of estimated planted forest distribution (Figs. 7, 8) and dominant tree species (Fig. 9) of East Asia, available at <https://doi.org/10.6084/m9.figshare.21774725.v3>³⁰¹. The database is in shapefiles where each polygon is 0.009° by 0.009° in size within the forested area of 2020 (≥ 5 m tree height) based on the FAO’s definition of “forest”^{2,270}. Each polygon contains the following attributes:

- ID:** Polygon ID
- Biome:** Biome classes used in the study
- Country:** Country

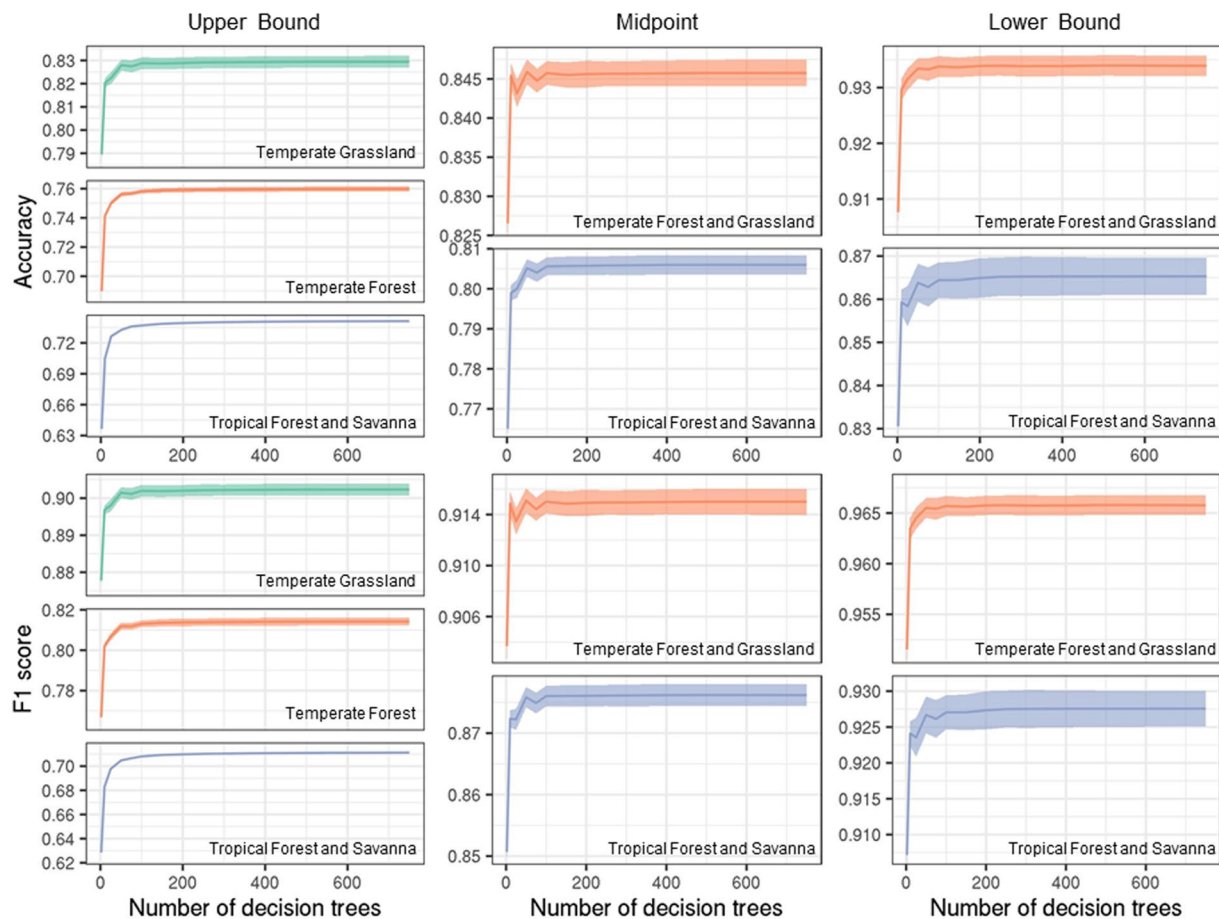


Fig. 5 Performance of random forest models in terms of classification accuracy and F1 score with different numbers of decision trees. For each biome (Fig. 2c), we tested a different number of decision trees in the random forest ranging from 2 to 750. The solid lines represent the mean of 10 repetitions, and the bands represent the standard deviation. The number of trees = 100 for the upper bound and midpoint models and 200 for the lower bound model were chosen to maximize the model performance while minimizing computational time.

Prc_Pln: Percent planted forest. The values represented the average of the three models (upper bound, midpoint, and lower bound). NA for ROK and a majority of areas in Japan, where national planted forest maps^{12,16} were used as the final planted/natural label (Fig. 2f).

Prc_P_U: Percent planted forest predicted by the upper bound model. NA for ROK and a majority of areas in Japan, where national planted forest maps^{12,16} were used as the final planted/natural label (Fig. 2f). Note that values are not always higher than Prc_Pln.

Prc_P_L: Percent planted forest predicted by the lower bound model. NA for ROK and a majority of areas in Japan, where national planted forest maps^{12,16} were used as the final planted/natural label (Fig. 2f). Note that values are not always lower than Prc_Pln.

Type: “Planted” or “Natural” forests based on the main result (*i.e.*, the average of the three models). For our predicted percent planted forest, “Planted” if $\text{Prc_Pln} \geq 0.5$ and “Natural” if $\text{Prc_Pln} < 0.5$. For $\text{Prc_Pln} = \text{NA}$, national planted forest maps^{12,16} were used to determine if the given polygon is a planted forest, and if not, “Natural.”

Typ_Upp: “Planted” or “Natural” forests based on the upper-bound model.

Typ_Lwr: “Planted” or “Natural” forests based on the lower-bound model.

Genus: For Type = “Planted”, this attribute indicates the predicted dominant genus. NA for Type = “Natural”.

Gns_Upp: For Typ_Upp = “Planted”, this attribute indicates the predicted dominant genus. NA for Typ_Upp = “Natural”.

Gns_Lwr: For Typ_Lwr = “Planted”, this attribute indicates the predicted dominant genus. NA for Typ_Lwr = “Natural”.

Besnard_Yr: Estimated planted year based on forest age³⁰² (<https://doi.org/10.17871/ForestAgeBGI.2021>). See Usage Notes.

Du_Yr: Estimated planted year based on the map of planting year of plantations^{303,304} (<https://doi.org/10.6084/m9.figshare.19070084.v2>). A value of 1981 indicates the planting year was before 1982, and values from 1982 to 2019 correspond to the planting years. See Usage Notes.

Area_m2: Area of the planted forest polygons in square meters.

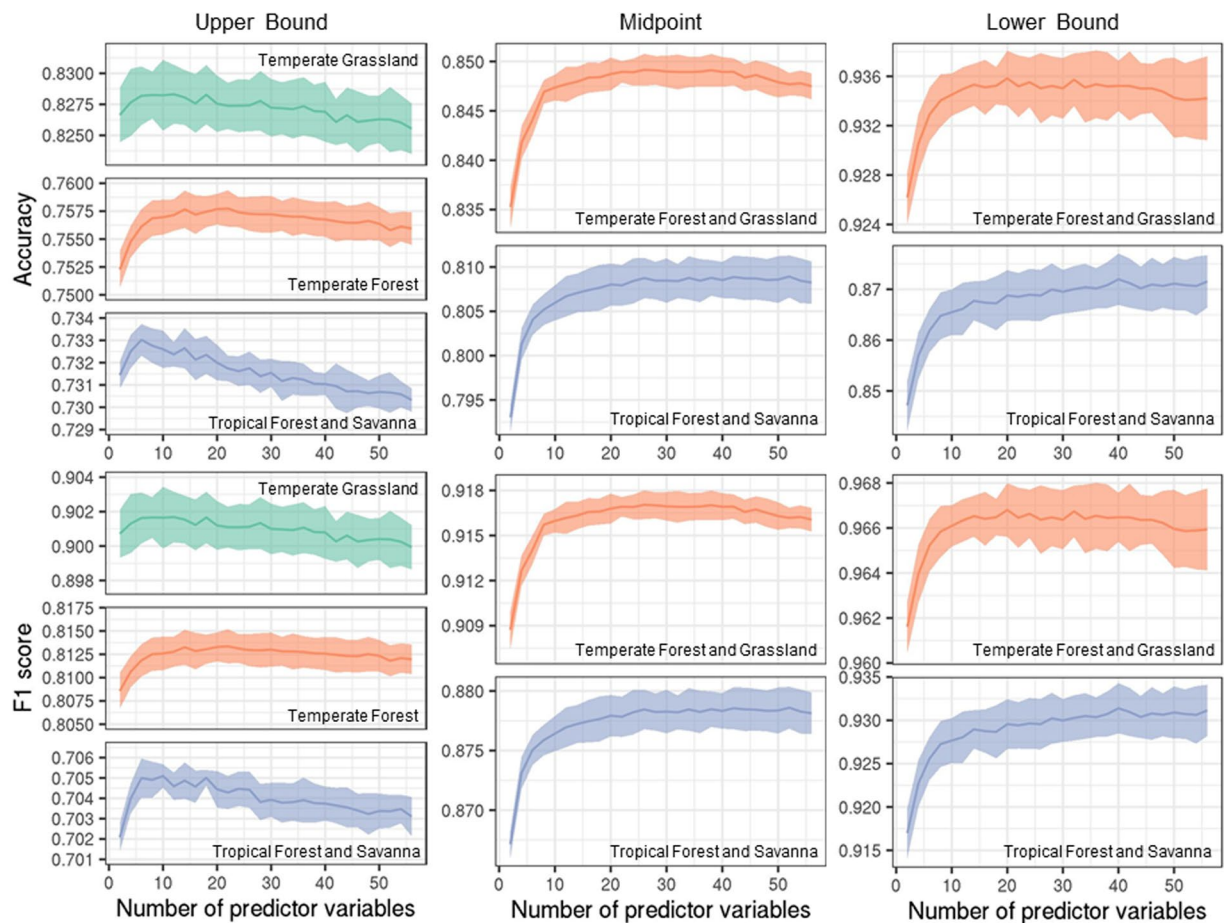


Fig. 6 Performance of random forest models in terms of classification accuracy and F1 score with different numbers of predictor variables. For each biome (Fig. 2e), we tested a different number of predictor variables in the random forest ranging from 2 to 56. The solid lines represent the mean of 10 repetitions, and the bands represent the standard deviation. We used the default number of predictor variables (seven) for all biomes for the upper bound model. We chose 26 and 42 for Temperate Forest and Grassland and Tropical Forest and Savanna for the midpoint model. We chose 20 and 40 for Temperate Forest and Grassland and Tropical Forest and Savanna for the lower bound model.

Raster layers are also available for percent planted forest, type (planted or natural forest), and dominant genus, at <https://doi.org/10.6084/m9.figshare.21774725.v3>³⁰¹.

Based on our prediction, the total area of planted forests in East Asia was 948,863 km², ranging between 600,529 and 1,277,549 km². China shared 87% of the planted forest area in East Asia, most of which is in the lowland subtropical and tropical regions, and Sichuan Basin (Fig. 8). More than half of China's planted forest area was dominated by *Cunninghamia* (Table 2) in the subtropical region and Sichuan Basin (Fig. 9). Larch (*Larix* spp.), black locust (*Robinia* spp.), and pine (*Pinus* spp.) were widely observed in northern and central China, and eucalyptus dominated planted forests in tropical regions.

In Japan and ROK, planted forests were uniformly distributed across the country (Fig. 8). More than half of Japan's total planted forest area was *Chamaecyparis*- or *Cryptomeria*-dominant (Table 2), while other coniferous genera (e.g., *Abies* and *Pinus*) covered northern planted forests (Fig. 9). ROK's planted forests were characterized by diverse genera; more than half of planted forest areas were dominated by pine, followed by deciduous trees including oak (*Quercus* spp.) and chestnut (*Castanea* spp.). DPRK's planted forests were mainly distributed in the south, largely composed of oak, larch, and pine.

The input training data, including the response variable and predictor variables, used in this study are available at <https://doi.org/10.6084/m9.figshare.21774812.v2>³⁰⁵. Underlying data included *in situ* and digitized planted-natural forest data:

The *in situ* observational data of China^{20–265}

The Japan Vegetation Map¹² (<http://gis.biodic.go.jp/webgis/sc-025.html?kind=v67>)

The national planted forest map of China¹⁵

The national planted forest map of ROK¹⁶

SDPT version 1.0¹³ (<https://www.wri.org/research/spatial-database-planted-trees-sdpt-version-10>)

Global planted trees extent 2015¹⁴ (<https://doi.org/10.5281/zenodo.3931930>)

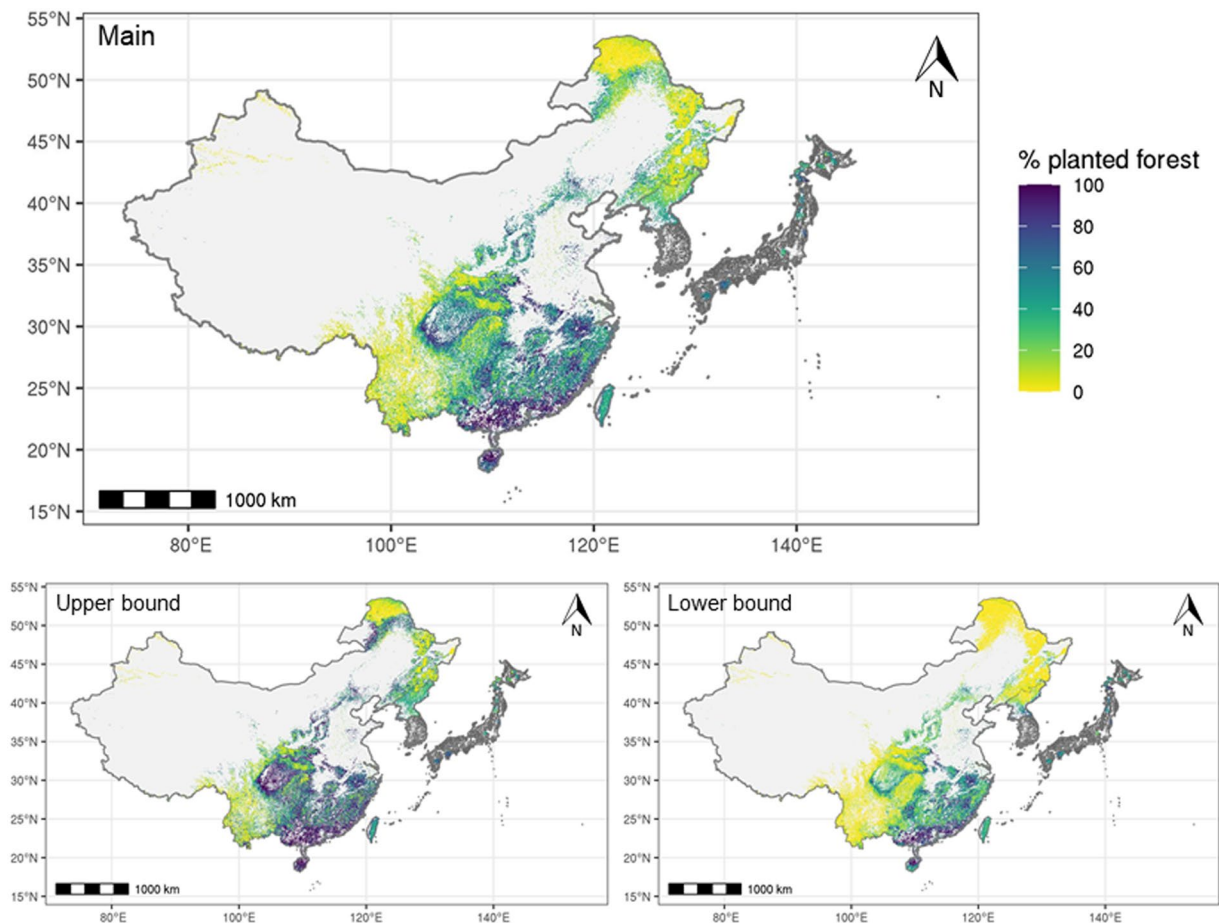


Fig. 7 Spatial distribution of percent planted forest in East Asia. Our main prediction was the mean percent planted forest from the three models (upper bound, midpoint, and lower bound), while upper and lower bounds present potential ranges. Prediction was made for China, DPRK, and small portions of Japan. National planted forest maps of Japan¹² and ROK¹⁶ were used for the remaining areas in ROK and the majority of areas in Japan, indicated in gray. The data is in a vector format with each polygon representing a 0.0090° by 0.0090° (approximately 1 km) grid in the WGS84 datum.

Japan National Forest Inventory²⁹⁵ (<http://forestbio.jp/datafile/datafile.html>)

ROK National Forest Inventory²⁹⁶

The predictor variables used in this study are all available through open sources as follows:

GEDI L2B²⁶⁹ (<https://search.earthdata.nasa.gov/search>)

Tree height (<https://glad.umd.edu/dataset/GLCLUC2020>)

MODIS (<https://modis.gsfc.nasa.gov/>)

Corrected precipitation: PBCOR²⁷¹ (<http://www.gloh2o.org/pbcor/>)

Bioclimate data: CHELSA^{272,273} (<https://chelsa-climate.org/bioclim/>)

Global aridity index and potential evapotranspiration: CGIAR-CSI v.2²⁷⁴ (<https://doi.org/10.6084/m9.figshare.7504448.v3>)

Topography: EarthEnv²⁷⁵ (<http://www.earthenv.org/topography>)

Global cattle distribution²⁷⁶ (<https://doi.org/10.7910/DVN/GIVQ75>)

Roadless area²⁷⁷ (<https://doi.org/10.1126/science.aaf7166>)

Protected area: UNEP-WCMC²⁷⁸ (<https://www.protectedplanet.net/en>)

Human footprint²⁷⁹ (<https://doi.org/10.5061/dryad.052q5>)

Soil characteristics: WISE30sec v1.0²⁸⁰ (<https://www.isric.org/explore/wise-databases>)

Other data used in this study include:

The Nature Conservancy (TNC) Terrestrial Ecoregions map¹⁹ (<https://geospatial.tnc.org/datasets/b1636d640ede4d6ca8f5e369f2dc368b/about>)

All the data listed above are open access, except the national planted forest map of China¹⁵, the national planted forest map of ROK¹⁶, and the ROK National Forest Inventory²⁹⁶. The sensitive information in these datasets will be available upon request via Science-i (<https://science-i.org/>) and approval from data contributors.

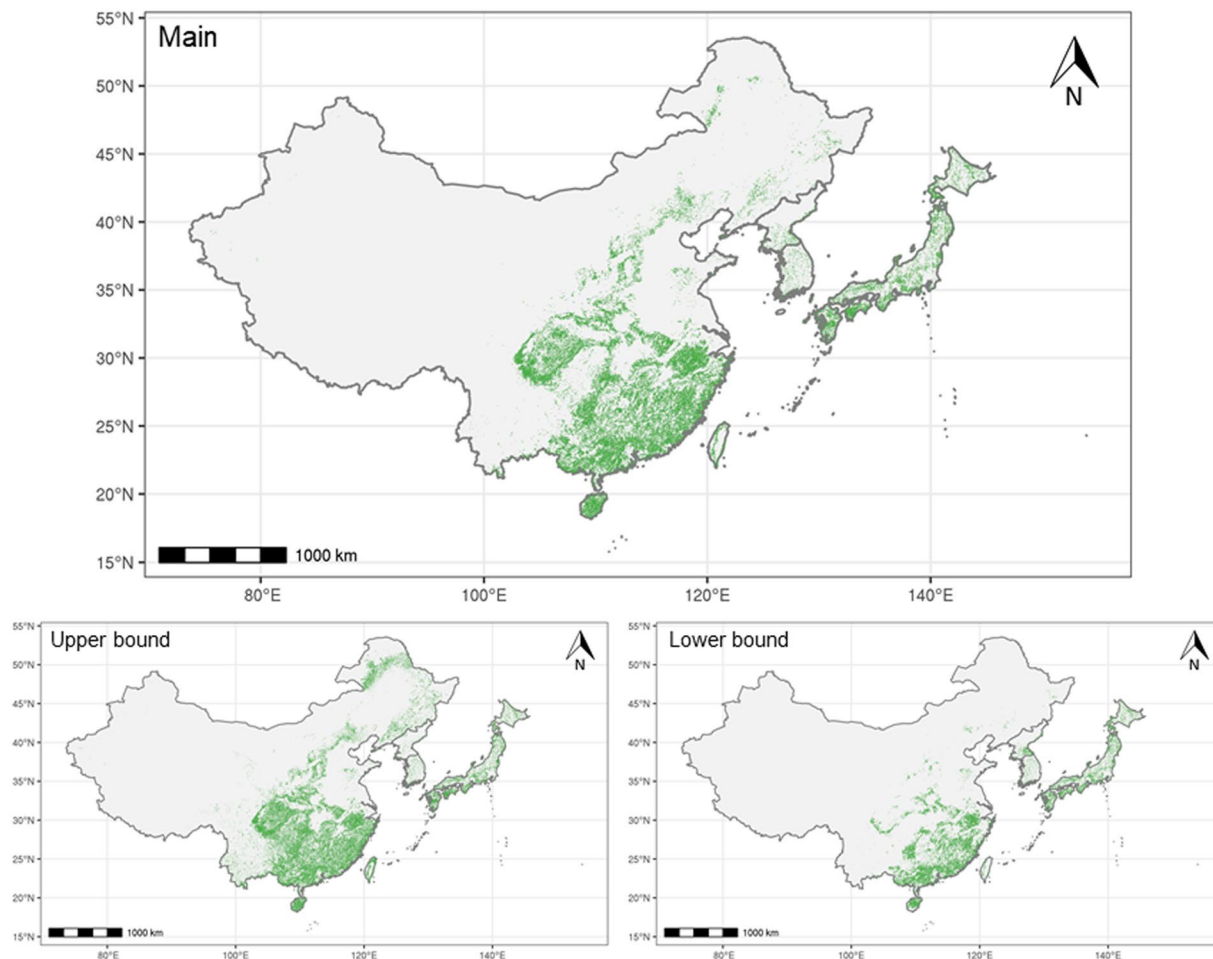


Fig. 8 Spatial distribution of planted forests in East Asia. The map shows the estimated areas where the percent planted forest is greater than 50%. For ROK and most areas in Japan, national planted forest maps^{12,16} were used to determine the distribution of planted forest. The data is in a vector format with each polygon representing a 0.0090° by 0.0090° (approximately 1 km) grid in the WGS84 datum.

Technical Validation

Model validation in imputing GEDI missing values. We conducted cross-validation with bootstrapping to evaluate the model in imputing the missing values in GEDI attributes for the high-latitude areas (Supplementary Table S1; see *Quality-Oriented Data Integration (QODI)* in Methods). R^2 was within the range of 31% and 42% for all the GEDI attributes in Temperate Grassland and Temperate Forest (Table 3). For Tropical Forest and Savanna, canopy height showed R^2 of 22%, and the rest of the attributes showed R^2 of almost 30%. Foliage height diversity showed the highest R^2 and total canopy cover showed the lowest root mean square error (RMSE) among all GEDI attributes in all groups (Table 3).

Model validation in estimating planted forests. To evaluate the performance of our mapping product of East Asia, we compared our main prediction (Fig. 8) with the planted/natural labels of the midpoint dataset for China. We calculated classification accuracy, precision, recall, F1 score, and four elements of confusion matrices in percentage (true positive, false positive, false negative, and true negative, where positive class represented planted, and negative class represented natural forest). Our prediction is characterized by a high recall (0.99), indicating that 99% of the observed planted forests were correctly predicted as planted forest (Table 4). Our precision was 0.63, which indicates that approximately two out of three positive predictions are actually planted forests. This level of accuracy is similar to those of other large-scale forest mapping studies (0.60–0.80)^{306–308}.

While precision is often negatively associated with recall, the F1 score, 0.77, indicates that our model is well-balanced between precision and recall. The low precision is attributable to the imbalanced distribution of positive and negative classes in the validation set (the midpoint dataset for China). The number of samples for natural forests was almost 10 times greater than that of planted forests in our validation set (Table 4). While we maximized the predictive performance by balancing the training data, high accuracy and low precision are inevitable due to the imbalanced validation set.

To further validate the quality of our prediction, we also compared our estimated total area of planted forests against the reported values from the FAO Global Forest Resources Assessment (FRA)² and the National

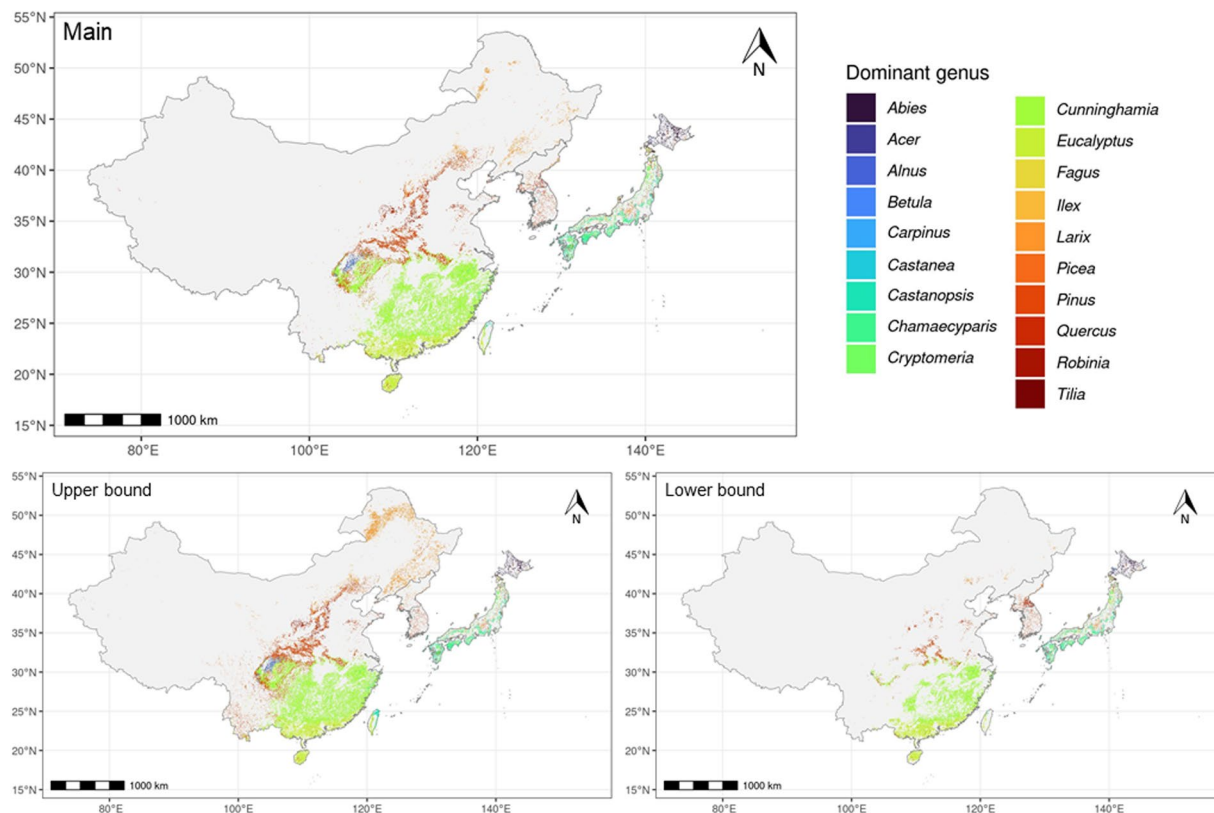


Fig. 9 Spatial distribution of dominant tree species to the genus level across the planted forest range in East Asia (Fig. 8).

Genus	Predicated area in km ² (%)				
	East Asia	China	Japan	ROK	DPRK
<i>Cunninghamia</i>	480,913 (50.7)	480,901 (58.2)	12 (0.0)	0 (0.0)	0 (0.0)
<i>Pinus</i>	151,749 (16.0)	142,031 (17.2)	2,334 (2.2)	5,692 (60.0)	1,691 (21.2)
<i>Eucalyptus</i>	137,318 (14.5)	137,318 (16.6)	0 (0.0)	0 (0.0)	0 (0.0)
<i>Larix</i>	40,487 (4.3)	32,400 (3.9)	5,383 (5.1)	525 (5.5)	2,179 (27.3)
<i>Chamaecyparis</i>	35,267 (3.7)	3,117 (0.4)	32,115 (30.4)	35 (0.4)	0 (0.0)
<i>Cryptomeria</i>	22,772 (2.4)	798 (0.1)	21,973 (20.8)	1 (0.0)	0 (0.0)
<i>Robinia</i>	14,046 (1.5)	13,858 (1.7)	87 (0.1)	100 (1.1)	1 (0.0)
<i>Abies</i>	11,695 (1.2)	211 (0.0)	11,455 (10.8)	0 (0.0)	29 (0.4)
<i>Alnus</i>	11,140 (1.2)	10,438 (1.3)	697 (0.7)	4 (0.0)	0 (0.0)
<i>Quercus</i>	10,793 (1.1)	21 (0.0)	4,510 (4.3)	2,402 (25.3)	3,860 (48.3)
<i>Castanopsis</i>	10,367 (1.1)	3,032 (0.4)	7,333 (6.9)	2 (0.0)	0 (0.0)
<i>Fagus</i>	5,892 (0.6)	4 (0.0)	5,887 (5.6)	0 (0.0)	0 (0.0)
<i>Castanea</i>	3,669 (0.4)	0 (0.0)	3,089 (2.9)	561 (5.9)	19 (0.2)
<i>Carpinus</i>	3,530 (0.4)	437 (0.1)	2,748 (2.6)	143 (1.5)	203 (2.5)
<i>Ilex</i>	3,515 (0.4)	8 (0.0)	3,507 (3.3)	0 (0.0)	0 (0.0)
<i>Acer</i>	2,640 (0.3)	1 (0.0)	2,637 (2.5)	1 (0.0)	2 (0.0)
<i>Betula</i>	1,979 (0.2)	885 (0.1)	1,092 (1.0)	2 (0.0)	0 (0.0)
<i>Picea</i>	681 (0.1)	289 (0.0)	392 (0.4)	0 (0.0)	0 (0.0)
<i>Tilia</i>	410 (0.0)	0 (0.0)	381 (0.4)	26 (0.3)	3 (0.0)
Total	948,863 (100.0)	825,751 (100.0)	105,633 (100.0)	9,493 (100.0)	7,986 (100.0)

Table 2. Predicted area of planted forest for each dominant genus based on the main model. East Asia is the sum of all four countries. The numbers in parenthesis represent percent area in each country or region.

Forest Inventory dataset from China³⁰⁹ (Table 5). Our total predicted area of planted forests in East Asia was 948,863 km² with a range between 600,529 and 1,277,549 km², which is consistent with the FRA estimate (981,390 km²). The predicted area of China's planted forests was 825,751 km² (475,566–1,159,009 km²), while

Biome	Response variable (unit)	RMSE (mean±95%CI)	R ² (mean±95%CI)
Temperate Grassland	Canopy height (rh100) (cm)	560.80 ± 1.57	0.35308 ± 0.00226
	Plant area index (pai) (-)	0.73181 ± 0.00093	0.38590 ± 0.00158
	Foliage height diversity (fhd_normal) (-)	0.29674 ± 0.00058	0.42148 ± 0.00167
	Total canopy cover (cover) (%)	0.15642 ± 0.00018	0.39395 ± 0.00166
Temperate Forest	Canopy height (rh100) (cm)	614.98 ± 0.21	0.31465 ± 0.00021
	Plant area index (pai) (-)	0.88935 ± 0.00016	0.39003 ± 0.00013
	Foliage height diversity (fhd_normal) (-)	0.29609 ± 0.00006	0.41618 ± 0.00012
	Total canopy cover (cover) (%)	0.16938 ± 0.00002	0.39714 ± 0.00012
Tropical Forest and Savanna	Canopy height (rh100) (cm)	716.28 ± 0.38	0.22118 ± 0.00025
	Plant area index (pai) (-)	1.01696 ± 0.00032	0.28806 ± 0.00017
	Foliage height diversity (fhd_normal) (-)	0.30687 ± 0.00009	0.29766 ± 0.00024
	Total canopy cover (cover) (%)	0.16035 ± 0.00004	0.29317 ± 0.00018

Table 3. Evaluation in imputing missing data of GEDI attributes for mapping purposes. We conducted cross-validation with bootstrapping. The mean and 95% confidence interval (CI) from 20 iterations are shown for root mean square error (RMSE) and R-squared (R²).

Evaluation metrics	Value
Accuracy	0.945
Precision	0.633
Recall	0.990
F1 score	0.772
True positive	0.093
False positive	0.054
False negative	0.001
True negative	0.853

Table 4. Evaluation metrics and elements of confusion matrices of the main prediction of planted forest distribution. Our final prediction was evaluated against the planted/natural labels of the midpoint dataset in China. The elements of confusion matrices are represented in percentages. The positive class represents planted, and the negative class represents natural forests. Accuracy shows the proportion of overall correct prediction, precision represents the correct prediction of the positive class (*i.e.*, planted) among all positive predictions, recall represents the correct prediction of the positive class among all actual positive cases, and F1 score represents a balanced score of precision and recall.

Country or region	Predicted area main (km ²)	Predicted area upper bound model (km ²)	Predicted area lower bound model (km ²)	FAO's FRA 2020 (km ²) ²	The Ninth National Forest Inventory of China (km ²) ³⁰⁹
East Asia	948,863	1,277,549	600,529	981,390	NA
China	825,751	1,159,009	475,566	846,960	795,428
Japan	105,633	103,447	103,823	101,840	NA
DPRK	7,986	5,601	11,648	9,870	NA

Table 5. Predicted area of planted forest for each country and the entire region and estimated area by other sources. ROK is not shown here as the national planted forest map¹⁶ was used in the final map. East Asia includes China, Japan, ROK, and DPRK.

the FRA reports 846,960 km² and the Ninth National Forest Inventory of China reports 795,428 km². For Japan, the range of estimated areas of planted forests was between 103,447 and 105,633 km², while the FRA reported value is 101,840 km². Our estimated area of planted forests in DPRK was 7,986 km² (5,601–11,648 km²), while the FRA reported value is 9,870 km². Overall, our estimate was consistent with those reported by the FRA and the National Forest Inventory of China.

Model validation in estimating dominant tree species. Our 90/10 bootstrapping cross-validation in estimating the dominant tree species across planted forests showed an overall classification accuracy of 0.396 (±0.003 95% CI). Among all the planted tree species, *Cunninghamia* and *Eucalyptus* had the highest F1 score (0.745 and 0.733, respectively), with high recall (0.893 and 0.802, respectively) and satisfactory precision (0.644 and 0.403, respectively) (Table 6). Meanwhile, *Carpinus* and *Castanea* showed the lowest F1 score (0.124 and 0.136, respectively), which likely resulted from a small sample size compared to other genera. *Acer*, *Alnus*, *Betula*,

Genus	Precision (mean \pm 95%CI)	Recall (mean \pm 95%CI)	F1 score (mean \pm 95%CI)
<i>Abies</i>	0.393 \pm 0.007	0.679 \pm 0.013	0.497 \pm 0.008
<i>Acer</i>	0.203 \pm 0.018	0.125 \pm 0.012	0.154 \pm 0.013
<i>Alnus</i>	0.039 \pm 0.028	0.011 \pm 0.007	0.182 \pm 0.005
<i>Betula</i>	0.288 \pm 0.020	0.187 \pm 0.014	0.223 \pm 0.015
<i>Carpinus</i>	0.086 \pm 0.013	0.129 \pm 0.018	0.124 \pm 0.011
<i>Castanea</i>	0.078 \pm 0.012	0.152 \pm 0.023	0.136 \pm 0.012
<i>Castanopsis</i>	0.163 \pm 0.008	0.522 \pm 0.027	0.246 \pm 0.011
<i>Chamaecyparis</i>	0.404 \pm 0.007	0.525 \pm 0.012	0.456 \pm 0.008
<i>Cryptomeria</i>	0.541 \pm 0.008	0.352 \pm 0.008	0.425 \pm 0.007
<i>Cunninghamia</i>	0.644 \pm 0.018	0.893 \pm 0.014	0.745 \pm 0.014
<i>Eucalyptus</i>	0.694 \pm 0.030	0.802 \pm 0.030	0.733 \pm 0.024
<i>Fagus</i>	0.403 \pm 0.008	0.771 \pm 0.015	0.527 \pm 0.009
<i>Ilex</i>	0.131 \pm 0.011	0.436 \pm 0.034	0.200 \pm 0.016
<i>Larix</i>	0.437 \pm 0.011	0.502 \pm 0.014	0.465 \pm 0.011
<i>Picea</i>	0.265 \pm 0.043	0.190 \pm 0.031	0.282 \pm 0.024
<i>Pinus</i>	0.584 \pm 0.010	0.445 \pm 0.010	0.504 \pm 0.009
<i>Quercus</i>	0.389 \pm 0.013	0.151 \pm 0.006	0.216 \pm 0.008
<i>Robinia</i>	0.460 \pm 0.035	0.527 \pm 0.037	0.479 \pm 0.029
<i>Tilia</i>	0.057 \pm 0.024	0.031 \pm 0.012	0.168 \pm 0.010

Table 6. Evaluation of the random forest classification model in mapping the dominant tree species across the planted forest expanse in East Asia. We conducted a rigorous 90/10 bootstrapping cross-validation. The mean and 95% confidence interval (CI) are shown for the precision, recall, and F1 score of each class (*i.e.*, genus) based on 100 random iterations.

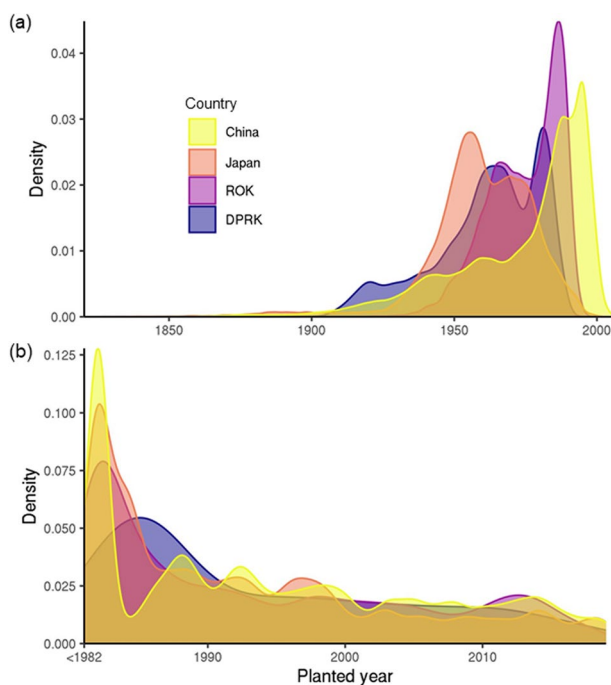


Fig. 10 Density plot showing the concentration of estimated planted forests in the range of planted year for each country. (a) planted year was estimated based on forest age³⁰² with a maximum value of 2010. (b) planted year was estimated based on the map of planting year of plantations^{303,304}.

Cryptomeria, *Picea*, *Pinus*, *Quercus*, and *Tilia* showed low recall compared to precision, indicating that true labels for these genera tended to be classified as other genera. *Abies*, *Carpinus*, *Castanea*, *Castanopsis*, *Chamaecyparis*, *Cunninghamia*, *Eucalyptus*, *Fagus*, *Ilex*, *Larix*, and *Robinia* had lower precision than recall due to the overprediction of these genera (Table 6).

Uncertainties. While this study advances the current understanding of planted forests in East Asia based on multi-source data consisting of *in situ*, digitized, and modeled datasets, uncertainties arose from two main sources. First, limited *in situ* data, especially from Japan, ROK, and DPRK constitute one of the largest sources of uncertainties. The limited *in situ* data from these countries could lead to lower accuracy in our planted forests prediction. Nevertheless, to mitigate this uncertainty, we integrated different data sources for modeling (e.g., SDPT¹³ and the Global Planted Trees Extent 2015¹⁴), and the final map product for these countries relied on external sources^{12,16}.

Secondly, our map of planted tree species depicts the spatial distribution of the dominant tree species to the genus level across the range of planted forests. However, it is beyond the scope of this study to identify the spatial distribution of monoculture planted forests versus mixed-species planted forests, the latter of which are common in certain regions³¹⁰. This uncertainty in tree species richness can be mitigated by integrating the mapping products presented here with recent global high-resolution maps of local tree species richness and co-limitation²⁸⁹. Furthermore, some genera predicted in our study had low F1 scores, which can be mitigated by increasing the sample size for these species. Nevertheless, it is not realistic to achieve perfectly balanced data, and differences in predictive performance among genera are inevitable.

Usage Notes

Our final maps of planted forest range (Fig. 8) for Japan and ROK consist of data directly obtained from the national planted forest maps of Japan¹² and ROK¹⁶. Users of these particular maps should cite these sources accordingly.

Planted forests in this study include forests of all ages that have been planted for ecological restoration, commercial plantation, and other purposes, such as landscape and disaster prevention.

Since the underlying training datasets differ by planting years, we were only able to quantify a roughly estimated range of underlying years. Specifically, we overlaid our final map with two existing map layers with estimated fore

st age³⁰² and planted year^{303,304} values. Based on these two sources, some planted forests were planted more than 100 years ago, while other planted forests are less than five years in age (Fig. 10). Estimation based on forest age³⁰² presents consistency with planting history in each country; the majority of planted forests were established post-war in Japan, followed by efforts in the Korean peninsula, while planted forests in China come from more recent planting (Fig. 10a). We included planted year information in our map product (see Data Records).

Code availability

The R code, saved RF models, and training datasets to reproduce the results of this study are available at <https://doi.org/10.6084/m9.figshare.21774812.v2>³⁰⁵.

Received: 5 February 2023; Accepted: 12 July 2023;

Published online: 22 July 2023

References

1. Brancalion, P. H. S. & Holl, K. D. Guidance for successful tree planting initiatives. *J. Appl. Ecol.* **57**, 2349–2361 (2020).
2. FAO. *Global Forest Resources Assessment 2020: Main Report*. (Food and Agriculture Organization of the United Nations, 2020).
3. Evans, J. *Planted forests: Uses, impacts and sustainability* (Food and Agricultural Organisation of the United Nations and CAB International, 2009).
4. Zhang, D., Zuo, X. & Zang, C. Assessment of future potential carbon sequestration and water consumption in the construction area of the Three-North Shelterbelt Programme in China. *Agric. For. Meteorol.* **303**, 108377 (2021).
5. Liu, J., Li, S., Ouyang, Z., Tam, C. & Chen, X. Ecological and socioeconomic effects of China's policies for ecosystem services. *Proc. Natl. Acad. Sci.* **105**, 9477–9482 (2008).
6. Xu, J., Yin, R., Li, Z. & Liu, C. China's ecological rehabilitation: Unprecedented efforts, dramatic impacts, and requisite policies. *Ecol. Econ.* **57**, 595–607 (2006).
7. Chen, C. *et al.* China and India lead in greening of the world through land-use management. *Nat. Sustain.* **2**, 122–129 (2019).
8. Farooq, T. *et al.* Perspectives of plantation forests in the sustainable forest development of China. *Iforest.* **14**, 166–174 (2021).
9. Yamaura, Y., Oka, H., Taki, H., Ozaki, K. & Tanaka, H. Sustainable management of planted landscapes: lessons from Japan. *Biodivers. Conserv.* **21**, 3107–3129 (2012).
10. Park, M. S. & Youn, Y.-C. Reforestation policy integration by the multiple sectors toward forest transition in the Republic of Korea. *For. Policy Econ.* **76**, 45–55 (2017).
11. Lee, D. & Lee, Y. Roles of Saemaul Undong in reforestation and NGO activities for sustainable forest management in Korea. *J. Sustain. For.* **20**, 1–16 (2005).
12. Biodiversity Center of Japan. *Vegetation Survey* <http://gis.biodic.go.jp/webgis/sc-025.html?kind=vg67> (2021).
13. Harris, N., Goldman, E. D. & Gibbes, S. *Spatial Database of Planted Trees (SDPT Version 1.0)*. Technical Note. <https://www.wri.org/publication/spatialdatabase-planted-trees> (World Resources Institute, 2019).
14. Lesiv, M. *et al.* Global Planted Trees Extent 2015. *Zenodo* <https://doi.org/10.5281/zenodo.3931930> (2020).
15. Yu, Z. *et al.* Mapping forest type and age in China's plantations. *Sci. Total Environ.* **744**, 140790 (2020).
16. Kim, K.-M., Kim, C.-M. & Jun, E. J. Study on the standard for 1:25,000 scale digital forest type map production in Korea. *J. Korean Assoc. Geograp. Infor. Stud.* **12**, 143–151 (2009).
17. Paquette, A. & Messier, C. The role of plantations in managing the world's forests in the Anthropocene. *Front. Ecol. Environ.* **8**, 27–34 (2010).
18. Dubayah, R. *et al.* The Global Ecosystem Dynamics Investigation: High-resolution laser ranging of the Earth's forests and topography. *Sci. Remote Sens.* **1**, 100002 (2020).
19. Olson, D. *et al.* Terrestrial ecoregions of the world: A new map of life on Earth: A new global map of terrestrial ecoregions provides an innovative tool for conserving biodiversity. *Bioscience* **51**, 933–938 (2001).
20. Ai, X. R. & Zhou, G. L. Study on the biomass of Chinese fir plantation's ecosystem in the north edge of median sub-tropic. *J. Hubei For. Technol.* **2**, 17–20 (1996).
21. An, H. P., Jin, X. Q. & Yang, C. H. A study on biomass growth and change law of major vegetation types in the pre-governance stage of Banqiaohe small watershed. *Guizhou For. Sci. Technol.* **19**, 20–34 (1991).

22. Cao, J. *et al.* Study of biomass and carbon storage in mixed stands of locust and catalpa in karst area of Guizhou. *J. Cent. South Univ. For. Technol.* **31**, 145–139 (2011).
23. Chen, D. *et al.* Study on biomass and net primary productivity of *Podocarpus imbricatus* plantation in Jianfengling, Hainan Island. *Forest Research—Chinese Academy of Forestry* **17**, 604–608 (2004).
24. Chen, D. & Zhao, W. Studies on the biomass of *Castanopsis orthacantha* community. *J. Yunnan Univ. Nat. Sci.* **3** (1994).
25. Chen, C. Y., Zhang, J. W., Zhou, C. L. & Zheng, H. Y. Researches on improving the quality of forest land and the productivity of artificial *Cunninghamia lanceolata* stands. *J. Appl. Ecol.* **1**, 97–106 (1990).
26. Chen, H. W. *et al.* Study of biomass of six kinds of plantations in Xishuangbanna. *Yunnan For. Sci. Technol.* **3**, 19–22 (2002).
27. Chen, H. Y. & Xu, Y. B. Study on productivity of mixed forest consisting of *Pinus massoniana* and *Castanopsis fissa*. *J. South China Agric. Univ.* **14**, 144–148 (1993).
28. Chen, Z. H., Zhang, H. D. & Wang, B. S. Studies on biomass and production of the lower subtropical evergreen broad-leaved forest in Heishiding Natural Reserve—biomass increment and net primary production. *Acta Ecol. Sin.* **12**, 377–386 (1992).
29. Chen, Z. H., Zhang, H. D., Wang, B. S. & Zhang, Z. Q. Studies on biomass and its allocation of the evergreen broadleaved forest in Heishiding, Guangdong. *Acta Phytocol. Geobot. Sin.* **17**, 289–298 (1993).
30. Cheng, T. R. *et al.* Research on forest biomass in Xiaolong Mountains, Gansu Province. *J. Beijing For. Univ.* **29**, 31–36 (2007).
31. Cheng, P. F. *et al.* Carbon storage and density of four main trees in Shangrila based on plot data. *For. Inventory Plan.* **36**, 12–15 (2011).
32. Dang, C. & Wu, Z. Studies on the biomass of *Castanopsis orthacantha* community. *J. Yunnan Univ. Nat. Sci.* **16**, 195–199 (1994).
33. Dang, C. L. & Wu, Z. L. Studies on the biomass for *Castanopsis echidnocarpa* community of monsoon evergreen broad-leaved forest. *J. Yunnan Univ.* **14**, 95–107 (1992).
34. Dang, C. L. & Wu, Z. L. Studies on the biomass of *Castanopsis orthacantha* community. *J. Yunnan Univ.* **16**, 195–199 (1994).
35. Dang, C. L. & Wu, Z. L. Studies on the biomass of *Pinus yunnanensis* forest. *Acta Bot. Yunnanica* **13**, 59–64 (1991).
36. Dang, C. L. & Wu, Z. L. Studies on the net primary production for *Castanopsis echidnocarpa* community of monsoon evergreen broad-leaved forest. *J. Yunnan Univ.* **14**, 108–117 (1992).
37. Dang, C. L. & Wu, Z. L. Studies on the net primary production of *Castanopsis orthacantha* community. *J. Yunnan Univ.* **16**, 200–204 (1994).
38. Dang, C. L. & Wu, Z. L. Studies on the net primary production of *Pinus yunnanensis* forest. *Acta Bot. Yunnanica* **13**, 161–166 (1991).
39. Dang, C. L., Wu, Z. L., Wang, C. Y. & He, Z. R. Studies on biomass and net primary production of *Abies georgei* Community. *J. Yunnan Univ.* **16**, 214–219 (1994).
40. Dang, C. L., Wu, Z. L. & Zhang, Z. Studies on the biomass of *Cyclobalanopsis delavayi* community. *J. Yunnan Univ.* **16**, 205–209 (1994).
41. Deng, K. M., Shi, P. L. & Yang, Z. L. Biomass allocation and net primary productivities and treeline ecotone on the Changbai Mountains, Northeast China. *J. Nat. Resour.* **21**, 942–948 (2006).
42. Deng, Q. X., Zhao, Y., Wu, L. J., Guo, W. F. & Lu, L. H. Biomass productivity of a 12-years-old *Cunninghamia lanceolata* plantation in Daqingshan of Guangxi. *Guangxi For. Sci.* **37**, 187–190 (2008).
43. Deng, S. J., Liao, L. P., Wang, S. L., Gao, H. & Lin, B. Bioproductivity of *Castanopsis hysrix*-*Cyclobalanopsis glauca*-*Machilus pauhoi* community in Huitong, Hunan. *Chin. J. Appl. Ecol.* **11**, 651–654 (2000).
44. Deng, S. J., Wang, K. P. & Gao, H. Biological productivity and nutrient distribution in over-mature plantation of *Cunninghamia lanceolata*. *Chin. J. Ecol.* **7**, 13–18 (1988).
45. Deng, T. X. & Liu, G. F. A preliminary study on alder and cypress mixed forest. *Acta Phytocol. Geobot. Sin.* **11**, 59–66 (1987).
46. Ding, G. Study on biomass and productivity of Masson Pine planting stand. *J. Fujian Coll. For.* **23**, 34–38 (2003).
47. Ding, G. J. Study on biomass and productivity of Masson Pine planting stand I. Biomass and density effect of different planting density. *J. Fujian Coll. For.* **23**, 34–38 (2003).
48. Ding, G. J. Study on change laws of biomass and productivity of Masson Pine forest plantation III. Biomass and productivity of different sites. *J. Mt. Agric. Biol.* **19**, 411–417 (2000).
49. Ding, G. J. & Wang, P. C. Study on change laws of biomass and productivity of Masson Pine forest plantation II. Biomass and productivity of stand at different ages. *For. Res.* **15**, 54–60 (2001).
50. Fan, J. Y., Tan, D. W., Jiang, L. & Qin, W. M. Research on the tree biomass and productivity of *Cryptomeria fortunei* Hooibrenk plantation. *J. Fujian For. Sci. Technol.* **38**, 1–5 (2011).
51. Fang, J. Study on biomass and productivity of *Picea likiangensis* var. *linzhiensis* forest in Nanyigou of Tibet. *For. Res.* **25**, 582–589 (2012).
52. Fang, H. B., Tian, D. L. & Kang, W. X. Biomass dynamics of a thinned Chinese fir plantation ecosystem. *J. Cent. South For. Univ.* **19**, 16–19 (1999).
53. Fang, Q. Effects of strengthening soil and cover plants management on energy utilization and nutrient cycle of ecosystem biomass in *Cunninghamia lanceolata*. *Sci. Silvae Sin.* **26**, 201–208 (1990).
54. Fang, X., Tian, D. L. & Xu, C. H. Productivity and carbon dynamics of Masson Pine plantation. *J. Cent. South For. Univ.* **23**, 11–15 (2003).
55. Fang, Y. T. & Mo, J. M. Study on carbon distribution and storage of a pine forest ecosystem in Dinghushan Biosphere Reserve. *Guihaia* **22**, 305–310 (2002).
56. Fang, Y. T., Mo, J. M., Huang, Z. L. & Ouyang, X. J. Carbon accumulation and distribution *Pinus massoniana* and *Schima superba* mixed forest ecosystem in Dinghushan Biosphere Reserve. *J. Trop. Subtrop. Bot.* **11**, 47–52 (2003).
57. Feng, Z. *et al.* The biological productivity on Chinese fir stands at different zones. *Acta Phytocol. Geobot. Sin.* **8**, 93–100 (1984).
58. Feng, Z. L., Gan, J. M., Zheng, Z. & Feng, Y. L. A comparative study on *Amomum villosum* cultivation under tropical wet seasonal rainforest and secondary forest at Xishuangbanna. *Chin. J. Appl. Ecol.* **15**, 1318–1322 (2004).
59. Feng, Z. L. *et al.* Biomass dynamics of the pioneer *Trema orientalis* community in the early stages of secondary succession of tropical forest in Xishuangbanna. *Chin. J. Ecol.* **18**, 1–6 (1999).
60. Feng, Z. L., Zheng, Z., Tang, J. W., Song, Q. S. & Zhang, J. H. Biomass of tropical secondary *Mallotus paniculatus* forest in Xishuangbanna. *Chin. J. Ecol.* **24**, 238–242 (2005).
61. Feng, Z. L. *et al.* Biomass and its allocation of a tropical wet seasonal rain forest in Xishuangbanna. *Acta Phytocol. Sin.* **22**, 481–488 (1998).
62. Feng, Z. W., Chen, C. Y., Zhang, J. W., Wang, K. P. & Zhao, J. L. Biological productivity of two forest communities in Huitong County of Hunan province. *Acta Phytocol. Geobot. Sin.* **6**, 257–267 (1982).
63. Feng, Z. W. *et al.* Determination of biomass of *Pinus massoniana* stand in Huitong County, Hunan province. *Sci. Silvae Sin.* **18**, 127–134 (1982).
64. Fu, Y. S., Guo, W. Y. & Ai, S. Studies on the community composition and biomass of pine plantation in Dawu County, Hubei province. *J. Hubei Univ. Nat. Sci.* **14**, 178–182 (1992).
65. Guo, L. D. *et al.* Evaluation of carbon sequestration function and its economic value of the *Eucalyptus urophylla* plantation in western Pearl River Basin. *J. Guangdong For. Technol.* **25**, 8–13 (2009).
66. Guo, Y. Q. *et al.* Biomass and productivity of *Jatropha curcas* plantation in Honghe River area. *J. Northwest For. Univ.* **25**, 1–4 (2010).

67. He, B. *et al.* Changing regularity of biomass and productivity of *Taiwania flousiana* plantation. *J. Northeast For. Univ.* **36**, 17–18 (2008).
68. He, B. *et al.* Vertical distribution of biomass and carbon storage in *Acacia melanoxylon* plantation ecosystem. *J. Northeast For. Univ.* **40**, 48–50 (2012).
69. He, H., Qiao, Y. K., Liu, Q., Wu, Y. & Lin, B. Dynamics of biomass and stem volume of *Picea asperata* stands in artificial restoration process of subalpine coniferous forest. *Chin. J. Appl. Ecol.* **15**, 748–752 (2004).
70. Hong, J. A study on the biomass and production of *Picea purpurea* forest communities. *Acta Phytoecol. Sin.* **2**, 8 (1986).
71. Hu, Q., Wang, S. L., Chen, L. C., Zhang, W. D. & Gong, C. Biomass and carbon density of main forest ecosystems in Hubei province. *Chin. J. Ecol.* **31**, 1626–1632 (2012).
72. Huang, T., Zhong, Q. & Peng, X. Study on biomass and productivity of *Liriodendron* Chinese Plantation. *For. Sci. Technol.* **9**, 5 (2000).
73. Huang, C. Y., Zhuang, X. P., Zhou, X. Y., Liu, Y. T. & Xie, D. M. Preliminary investigation of thin bark masson pine. *J. Cent.-South For. Coll.* **11**, 79–88 (1991).
74. Huang, L. J. Study on biomass and carbon density of *Acacia mangium* plantation. *Guangxi Trop. Agric.* **2**, 8–12 (2009).
75. Huang, S. D., Wu, Q. B., Liao, K. B., Mo, D. X. & Qin, J. Carbon storage and its allocation in an artificial *Tsoongiodendron odorum* ecosystem in southern subtropical region of China. *Chin. J. Ecol.* **30**, 2400–2404 (2011).
76. Huang, T., Zhong, Q. P. & Peng, X. Y. Study on biomass and productive forces of artificial *Liriodendron chinense*. *For. Sci. Technol.* **9**, 12–15 (2000).
77. Huang, T., Zhong, Q. P. & Peng, X. Y. Study on biomass and productive forces of artificial *Liriodendron chinense*. *Jiangxi For. Sci. Technol.* **5**, 4–9 (2000).
78. Jiang, P., Ye, J. & Wu, G. Woody species composition and biomass of main tree species in a 25hm² plot of broad-leaved and Korean pine mixed forests of Changbai Mountain, northeast China. *J. Beijing For. Univ.* **27**, 112–115 (2005).
79. Jiang, H. A primary study about the relationship between natural stand production of *Picea asperata* and ecology environment condition. *Acta Bot. Sin.* **28**, 538–548 (1986).
80. Jiang, H. A study on the biomass and production of *Picea purpurea* forest communities. *Acta Phytoecol. Geobot. Sin.* **10**, 146–152 (1986).
81. Kang, B. *et al.* Carbon accumulation and distribution in *Pinus massoniana* and *Cunninghamia lanceolata* mixed forest ecosystem in Daqingshan, Guangxi of China. *Acta Ecol. Sin.* **26**, 1321–1329 (2006).
82. Kang, W. X., Tian, D. L., Yan, W. D. & Fang, X. Solar energy storage and distribution in bole stage of Chinese fir plantation. *Sci. Silvae Sin.* **40**, 205–209 (2004).
83. Lan, Z. J. *et al.* Biomass distribution of major plant communities in Jiuzhaigou valley, Sichuan. *Chin. J. Appl. Environ. Biol.* **10**, 299–306 (2004).
84. Li, J. *et al.* Studies on the biomass and productivity of *Manglietia glauca*. *J. Fujian For. Sci. Technol.* **1**, 3 (2011).
85. Li, Q. *et al.* Study on biomass and productivity of *Populus* plantation in the Huanghuaihai Plain. *Henan Sci.* **26**, 434–437 (2008).
86. Li, G. X. *et al.* Characteristics of *Alnus cremastogyne* plantation community and its biomass in central Yuannan Plateau. *J. Zhejiang For. Coll.* **23**, 362–366 (2006).
87. Li, J. *et al.* Dynamics of biomass and productivity for young and middle-aged plantations of *Pinus kesiya* var. *langbianensis*. *J. Northeast For. Univ.* **38**, 36–38 (2010).
88. Li, J. Z. *et al.* Studies on the biomass and productivity of *Manglietia glauca*. *J. Fujian For. Sci. Technol.* **38**, 1–5 (2011).
89. Li, S. C., Liu, X. L., Liu, X. Q. & Shi, Y. H. Studies on biomass and forest hydrological effect in *Pinus massoniana* plantation. *Cent. South For. Inventory Plan.* **3**, 26–29 (1995).
90. Li, W. B., Bao, W. K., He, B. H., Wu, W. Y. & Li, F. L. Biomass compositions of *Pinus tabulaeformis* plantation and their relationships in the Dagou valley of the upper Minjiang River. *J. Mt. Sci.* **25**, 236–244 (2007).
91. Li, Z. H., Chen, S. X., Xie, Y. J., Zhang, B. & Zhu, B. L. Effects of stand density upon the biomass and productivity of *Eucalyptus urograndis*. *J. Cent. South Univ. For. Technol.* **28**, 49–54 (2008).
92. Liang, Y. *et al.* A study on biomass and productivity of *Michelia macclurei* Dandy plantation in southeast of Guangxi. *China For. Sci. Technol.* **5** (2010).
93. Liang, N., Wang, W. B., Ni, J. B. & Tian, K. A study on biomass in sapling stage of pure *Betula alnoides* forest and *Betula alnoides* and *Cinnamomum cassia* mixed forest. *J. West China For. Sci.* **36**, 44–49 (2007).
94. Liang, N., Wang, W. B. & Tian, K. Biomass distribution characteristics of 4 and 13 years old *Betula alnoides* plantations. *J. West China For. Sci.* **35**, 118–222 (2006).
95. Liang, Y. X. *et al.* The growth regularity, biomasses and productivity of *Acacia melanoxylon* in the southwest of Guangxi. *J. Fujian For. Sci. Technol.* **37**, 1–5 (2010).
96. Liang, Y. X., Wei, Z. M., Yu, G. C., Qin, J. & Chen, J. C. A study on biomass and productivity of *Michelia macclurei* Dandy plantation in southeast of Guangxi. *China For. Sci. Technol.* **24**, 45–49 (2010).
97. Liao, B., Zheng, D. & Zheng, S. Biomass research on the *Sonneratia caseolaris*. *For. Res.* **3**, 47–54 (1990).
98. Liao, B. & Zheng, D. Study on the forest biomass and productivity of Olive wood. *For. Res.* **4**, 22–29 (1991).
99. Lin, J. Study on productivity of mixed forest of Chinese Fir and *Schima*. *Sci. Technol. Qinghai Agric. For.* **4**, 2–31 (2004).
100. Lin, P. *et al.* The biomass and productivity of the mangrove forest. *J. Xiamen Univ. Nat. Sci.* **29**, 209–213 (1990).
101. Lin, P., Hu, H. Y., Zheng, W. J., Li, Z. J. & Lin, Y. M. A study on the biomass and energy of mangrove communities in Shenzhen bay. *Sci. Silvae Sin.* **34**, 18–24 (1998).
102. Lin, W. H., Chen, K. M. & Liu, Z. G. Biomass and nutrient element content of *Eucalyptus* plantations of dry-hot valley in southwest Sichuan. *Mt. Res.* **12**, 251–255 (1995).
103. Liqiang, M. *et al.* Study on the tree layer biomass of *Picea koraiensis* artificial forests. *Zhiwu Yanjiu* **15**, 551–557 (1994).
104. Liu, X. L. *et al.* Studies on root biomass and productivity in dominant plantation populations in the mountainous land in western Sichuan. *Acta Ecol. Sin.* **26**, 542–551 (2006).
105. Liu, X. *et al.* Biomass study of the plantation of *Alnus cremastogyne* burkill at different stages of age. *J. Cent. South Univ. For. Technol.* **2**, 17 (2007).
106. Liu, Y., Wu, Z. & Guo, Z. Henan day treasure net nature reserve of *Quercus variabilis* forest biomass and productivity research. The progress of science and technology and social and economic development of the 21st century (part ii) (1999).
107. Liu, J. H., Tang, D. Y. & Yang, Q. P. Study on production in *Corylus chinensis* plantation. *Chin. Horticu. Abstr.* **11**, 26–28 (2009).
108. Liu, Q. Studies of the biomass and productivity of different age-group *Pinus massoniana* plantation. *J. Cent. South For. Univ.* **16**, 47–51 (1996).
109. Liu, W. Y. Studies in biomass and productivity of *Acacia dealbata* plantation in the protected district of water sources in north Kunming. *Guihaia* **15**, 327–334 (1995).
110. Liu, X. C., Wen, S. Z., Feng, H. H. & Yang, L. L. Biomass study of the plantation of *Alnus cremastogyne* Burkill at different stages of age. *J. Cent. South Univ. For. Technol.* **27**, 83–86 (2007).
111. Liu, Y. G. *et al.* Organic carbon storage in four ecosystem types in the karst region of southwestern China. *PLoS ONE* **8**, e106876 (2013).
112. Liu, Y. L. & Xue, J. H. Quantitative properties of degraded karst forest communities in Maolan mountain area of Guizhou. *J. Nanjing For. Univ. Nat. Sci. Ed.* **29**, 23–27 (2005).

113. Luo, T. *et al.* Root biomass along subtropical to alpine gradients: global implication from Tibetan transect studies. *For. Ecol. Manag.* **206**, 349–363 (2005).
114. Luo, T., Li, W. & Zhu, H. Estimated biomass and productivity of natural vegetation on the Tibetan Plateau. *Ecol. Appl.* **12**, 980–997 (2002).
115. Luo, Y. *et al.* Root: shoot ratios across China's forests: forest type and climatic effects. *For. Ecol. Manag.* **269**, 19–25 (2012).
116. Luo, J., Tian, Y. X., Yang, N. & Yao, M. Study on biomass of different types for protection forest system area around Dongting Lake. *Hunan For. Sci. Technol.* **38**, 27–29 (2011).
117. Luo, J., Yang, Z. & Yang, Q. W. A study on the biomass and productivity of forest on the Gongga Mountain. *Acta Phytoecol. Sin.* **24**, 191–196 (2000).
118. Luo, W., Zhang, C., Zhao, X. & Liang, J. Understanding patterns and potential drivers of forest diversity in northeastern China using machine-learning algorithms. *J. Veg. Sci.* **32**, e13022 (2021).
119. Luo, Y. *et al.* A review of biomass equations for China's tree species. *Earth Syst. Sci. Data* **12**, 21–40 (2018).
120. Lv, X. T. *et al.* Biomass and its allocation in tropical seasonal rain forest in Xishuangbanna. *China. J. Plant Ecol.* **31**, 11–22 (2007).
121. Ma, M. D., Jiang, H. & Liu, Y. J. Biomass, carbon content, carbon storage and their vertical distribution of *Phoebe bournei* artificial stand. *Sci. Silvae Sin.* **44**, 34–39 (2008).
122. Ma, M. D., Jiang, H., Luo, C. D. & Liu, Y. J. Preliminary study of carbon density, net production and carbon stock in natural spruce forests of northwest subalpine Sichuan. *China. J. Plant Ecol.* **31**, 305–312 (2007).
123. Ma, M. D., Jiang, H. & Yang, J. Y. study on biomass of *Phoebe zhennan* plantation in the western hilly regions of the Sichuan basin. *J. Sichuan For. Sci. Technol.* **10**, 6–14 (1989).
124. Miao, S. Y. *et al.* Biomasses and distributive patterns of mangrove populations in Zhanjiang Nature Reserves, Guangdong, China. *Guihaia* **18**, 19–23 (1998).
125. Ming, A. G. *et al.* Biomass and its allocation in a 28-year-old *Mytilaria laosensis* plantation in southwest Guangxi. *Chin. J. Ecol.* **31**, 1050–1056 (2012).
126. Ming, A. G. *et al.* The study of the growth and biomass of *Mytilaria laosensis* plantation in different slope positions. *Chin. Agric. Sci. Bull.* **27**, 90–93 (2011).
127. Mo, D., Guan, D. S., Liu, S. W., Huang, K. Y. & Liang, Y. D. Biomass, foliar dust and species diversity of forests in ecological safety islands of urban Guangzhou. *Acta Sci. Circumst.* **31**, 666–679 (2011).
128. Mo, D. X., Liao, K. B., Wu, Q. B. & Qin, J. The carbon storage amount and spatial distribution characteristics of *Paramichelia bailonii* plantations. *J. Anhui Agric. Sci.* **39**, 14072–14075 (2011).
129. Mo, J. M., Peng, S. L., Brown, S., Kong, G. H. & Fang, Y. T. Response of biomass production to human impacts in a pine forest in subtropical China. *Acta Ecol. Sin.* **24**, 193–200 (2004).
130. Ning, C. *et al.* Biomass and productivity of *Betula luminifera* and Italian poplar mixed-forests ecosystem in karst city. *J. Cent. South Univ. For. Technol.* **31**, 161–166 (2011).
131. Niu, D., Wang, S. L. & Ouyang, Z. Y. Comparisons of carbon storages in *Cunninghamia lanceolata* and *Michelia macclurei* plantations during a 22-year period in southern China. *J. Environ. Sci.* **21**, 801–805 (2009).
132. Pan, K. & Liu, G. Biomass research on the ecological artificial community. *Chin. J. Appl. Environ. Biol.* **5**, 121–130 (1999).
133. Pan, P., Li, R. W., Qing, Z. G., Quan, M. J. & Cao, J. A study on biomass and productivity of *Eucommia ulmoides* plantation. *Resour. Environ. Yangtze Basin* **9**, 71–77 (2000).
134. Pan, P., Li, R. W., Xiang, C. H., Zhu, Z. F. & Yin, X. M. Biomass and productivity of *Cupressus lusitanica*. *Resour. Environ. Yangtze Basin* **11**, 133–136 (2002).
135. Pan, P. *et al.* Research on biomass and productivity of *Pinus radiata* plantations. *J. Sichuan For. Sci. Technol.* **26**, 21–27 (2005).
136. Peng, S., Yu, Z. & Zhang, W. Coenological analysis of five man-made forests on down-land in Heshan, Guangdong. *Acta Phytoecol. Geobot. Sin.* **16**, 1–10 (1992).
137. Peng, P. H., Peng, J. S., Wang, C. S. & Wang, J. X. Studies on the biomass and production of *Populus schneideri* var. *tibetica* plantation. *For. Sci. Technol.* **28**, 14–18 (2003).
138. Peng, S. L. & Fang, W. Features of biomass and productivity dynamics in successional process of low subtropical forest. *J. Ecol. Sci.* **2**, 1–9 (1995).
139. Qi, J. F. & Tang, J. Biomass and its allocation pattern of monsoon rain forest over limestone in Xishuangbanna of Southwest China. *Chin. J. Ecol.* **27**, 167–177 (2008).
140. Qi, L. H. *et al.* Species diversity and biomass allocation of vegetation restoration communities on degraded lands. *Chin. J. Ecol.* **26**, 1697–1702 (2007).
141. Qi, Z. Y., Ma, J. X. & Li, S. M. Studies on the biomass and productivity of artificial *Michelia macclurei* var. *sublanaea* Forest. *Chin. J. Ecol.* **30**, 17–19 (1985).
142. Qin, W., He, B. & Qin, S. Biomass and productivity in *Acacia crassicaarpa* plantation. *J. Northwest For. Univ.* **23**, 17 (2008).
143. Qin, J., Meng, H. S., Qin, W. M., Yu, J. M. & Qin, D. W. Studies on the biomass and productivity of *Tsoongiodendron odorum* plantation. *China For. Sci. Technol.* **25**, 65–68 (2011).
144. Qin, L. *et al.* Allocation pattern of biomass and productivity for three plantations of *Castanopsis hystrix*, *Pinus massoniana* and their mixture in south subtropical area of Guangxi, China. *Sci. Silvae Sin.* **47**, 17–21 (2011).
145. Qin, W. M., He, B., Yu, H. G. & Chen, W. J. Biomass productivity of *Acacia mangium* plantations of different age classes. *J. Northeast For. Univ.* **35**, 22–24 (2007).
146. Qin, W. M., Qiu, B. F., Qin, J., Xu, T. B. & Qin, D. W. Study on the biomass and growth law of *Paramichelia bailonii* plantation. *J. Fujian Coll. For.* **31**, 110–114 (2011).
147. Qiu, F. *et al.* Studies on biomass of evergreen broad-leaved forest of Jinpenshan forests in Jiangxi Province. *Jiangxi For. Sci. Technol.* **4** (2011).
148. Qiu, F. Y. *et al.* Studies on biomass of evergreen broad-leaved forest of Jinpenshan forests in Jiangxi province. *Jiangxi For. Sci. Technol.* **4**, 1–5 (2011).
149. Qiu, X. Z., Xie, S. C. & Jin, G. F. A preliminary study on biomass of *Lithocarpus xylocarpus* forest in Xujiaba region, Ailao mts., Yunnan. *Acta Bot. Yunnanica* **6**, 85–92 (1984).
150. Rong, Y. *et al.* Biomass productivity of the second generation *Cunninghamia lanceolata* plantation in northwest of Guangxi. *J. Guangxi Agric. Biol. Sci.* **27**, 451–455 (2008).
151. Sang, W., Su, H. & Chen, L. Coupling biomass and energy in warm temperate deciduous broad-leaved oak (*Quercus liaotungensis*) forest ecosystem. *Acta Phytoecol. Sin.* **26**, 88–92 (2002).
152. Shao, P. & Zhu, Z. Biomass and primary productivity of dominant species *Aporosa yunnanensis* and *Blastus cochinchinensis* of forest vegetation on Dinghu Mountain. *Chin. J. Appl. Ecol.* **3**, 1 (1992).
153. Shen, Y. *et al.* Biomass and productivity of natural secondary *Sassafras tzumu* and *Liquidambar formosana* mixed forest. *J. Cent. South Univ. For. Technol.* **31**, 26–30 (2011).
154. Sheng, C. Y., Liu, L. H. & Liu, W. Y. Biomass and dynamics of soil environment during the early stage of vegetation restoration in a degraded dry-hot mountain area of Nanjian, Yunnan. *Acta Phytoecol. Sin.* **24**, 575–580 (2000).
155. Shi, P., Zhong, Z. & Li, X. A study on the biomass of alder and cypress artificial mixed forest in Sichuan. *Chin. J. Plant Ecol.* **20**, 524–533 (1996).

156. Shi, S. *et al.* Biomass and carbon storage of the secondary forest (*Populus davidiana*) at different stand growing stages in southern Daxinganling temperature zone. *Ecol. Environ. Sci.* **21**, 428–433 (2012).
157. Su, Y., Wu, Q. B., Shi, F. J., Liang, J. & Duan, W. W. Storage and distribution pattern of carbon in the ecological system of man-made forest of *Parashorea chinensis*. *J. Anhui Agric. Sci.* **39**, 5271–5273 (2011).
158. Su, Y. M., Liu, X. L. & Xiang, C. H. studies on biomass and production of *Abies fabri* plantation. *J. Sichuan For. Sci. Technol.* **21**, 31–35 (2000).
159. Su, Y. M. *et al.* Research on biomass of *Quercus liaotungensis* natural secondary stand. *J. Nanjing For. Univ. Nat. Sci. Ed.* **27**, 107–109 (2003).
160. Tang, J. W. *et al.* preliminary study on the biomass of secondary tropical forest in Xishuangbanna. *Acta Phytoecol. Sin.* **22**, 489–498 (1998).
161. Tang, J. W. *et al.* A preliminary study on the net primary production of the secondary tropical forest in Xishuangbanna. *Acta Phytoecol. Sin.* **27**, 756–763 (2003).
162. Tang, J. W. *et al.* Biomass and net primary productivity of artificial tropical rainforest in Xishuangbanna. *Chin. J. Appl. Ecol.* **14**, 1–6 (2003).
163. Tian, Q., Zhou, R. & Zhang, J. Biomass research on the *Quercus variabilis* plantation. *J. Beijing For. Univ.* **2** (1997).
164. Tian, D. L., Pan, H. H., Kang, W. X. & Fang, H. B. A study of the biomass of a second generation Chinese fir plantation. *J. Cent. South For. Univ.* **18**, 11–16 (1998).
165. Tian, D. L., Xiang, W. H. & Kang, W. X. Study on biological cycling of microelements in *Pinus massoniana* plantations. *Sci. Silvae Sin.* **39**, 1–8 (2003).
166. Tian, D. L., Xiang, W. H., Yan, W. D. & Kang, W. X. Effect of successive-rotation on productivity and biomass of Chinese fir plantation at fast growing stage. *Sci. Silvae Sin.* **38**, 14–18 (2002).
167. Tian, D. L., Zhang, C. J., Luo, Z. P. & Yuan, W. X. Biomass and distribution of nutrient elements in natural *Sassafras* mixed forest II. Spatial distribution of nutrient elements. *J. Cent.-South For. Coll.* **11**, 35–43 (1991).
168. Tian, D. L., Zhang, C. J., Luo, Z. P. & Yuan, W. X. Biomass and distribution of nutrient elements in natural *Sassafras* mixed forests. *J. Cent.-South For. Coll.* **10**, 121–128 (1990).
169. Wang, B. & Yang, X. S. Comparison of biomass and species diversity of four typical zonal vegetations. *J. Fujian Coll. For.* **29**, 345–350 (2009).
170. Wang, M. & Li, H. Quantitative study on the soil water dynamics of various forest plantations in the loess plateau region in northwestern Shanxi. *Acta Ecol. Sin.* **15** (1995).
171. Wang, X. Y., Hu, D. & He, J. S. Biomass research of *Fagus Engleriana* and *Quercus aliena* var. *acuteserrata* forest in Shennongjia Forest District. *J. Capit. Nor. Univ. Nat. Sci. Ed.* **28**, 62–67 (2007).
172. Wang, D. *et al.* Study of biomass and production of the forest *Quercus mongolica* in Wuling Mountain. *Chin. J. Ecol.* **1**, 9–15 (1998).
173. Wang, X., Fang, J. & Zhu, B. Forest biomass and root–shoot allocation in northeast China. *For. Ecol. Manag.* **255**, 4007–4020 (2008).
174. Wang, B. & Yang, X. Q. Comparison of carbon content and carbon density of four typical zonal forest ecosystems. *J. Hum. Agric. Univ. Nat. Sci.* **36**, 464–470 (2010).
175. Wang, J. Investigation on community characteristics and biomass of *Alnus cremastogyne* and *Cupressus funebris* mixed young forest. *J. Sichuan For. Sci. Technol.* **14**, 66–69 (1993).
176. Wang, W. B. A study on community characteristics of *Betula alnoides* plantation. *J. West China For. Sci.* **35**, 8–13 (2006).
177. Wang, X. K. *et al.* Structure of ecosystem biomass in karst urban poplar plantations. *J. Cent. South Univ. For. Technol.* **31**, 13–18 (2011).
178. Wang, Y. Q., Sun, B. P., Zhao, Y., Zhou, X. S. & Zhong, X. J. Comparison of vegetative biomass in different configuration modes for reforestation of cultivated land in Liping County of Guizhou province. *Hunan Agric. Sci.* **11**, 120–122 (2011).
179. Wei, J. Study on the biomass structure of *Fokienia hodginsii* plantations in South Fuji. *J. Fujian For. Sci. Technol.* **3** (2001).
180. Wen, D. Z., Wei, P., Kong, G. H., Zhang, Q. M. & Huang, Z. L. Biomass study of the community of *Castanopsis chinensis* + *Cryptocarya concinna* + *Schima superba* in a southern China reserve. *Acta Ecol. Sin.* **17**, 497–504 (1997).
181. Wen, S. Z., Tian, D. L., Yang, L. L. & Fang, X. Carbon density, carbon stock and carbon sequestration in *Alnus cremastogyne* plantation. *Sci. Silvae Sin.* **46**, 15–21 (2010).
182. Wen, Y. G. *et al.* Relationship between species diversity and biomass of *Eucalyptus* plantation in Guangxi. *Sci. Silvae Sin.* **44**, 14–19 (2008).
183. Wen, Y. G. & Huang, C. B. Litter production in Chinese fir plantation in Liluo. *For. Sci. Technol.* **7**, 12–16 (1986).
184. Wen, Z. L., Zheng, W. & Chen, R. M. Biomass and nutrient distribution of *Acacia mangium* plantations. *J. Anhui Agric. Sci.* **38**, 15040–15045 (2010).
185. Wu, P., Ding, F. J., Cui, Y. C., Zhu, J. & Li, C. R. Study on biomass and productivity of young *Pinus massoniana* forests of Pearl River Shelterbelt Construction Project in Qiannan Prefecture. *Guizhou Agric. Sci.* **40**, 169–172 (2012).
186. Wu, X. *et al.* CPSPDv0: a forest stand structure database for plantation forests in China. *Big Earth Data* **7**, 212–230 (2022).
187. Wu, X. S. & Huang, C. D. Carbon density, storage and distribution in birth forest ecosystem on the forestland converted from farmland. *Chin. J. Ecol.* **26**, 323–326 (2007).
188. Wu, Z. L. & Dang, C. L. A preliminary study on biomass and net primary productivity of *Quercus senescens* forest near Kunming. *J. Yunnan Univ.* **16**, 235–240 (1994).
189. Wu, Z. L. & Dang, C. L. The biomass and net primary productivity of *Pinus kesiya* var. *langbianensis* stands in Chang-Nin district, Yunnan. *J. Yunnan Univ.* **14**, 137–145 (1992).
190. Wu, Z. L. & Dang, C. L. The biomass of *Pinus kesiya* var. *langbianensis* stands in Pu'Er district, Yunnan. *J. Yunnan Univ.* **14**, 119–127 (1992).
191. Wu, Z. L. & Dang, C. L. The net primary productivity of *Pinus kesiya* var. *langbianensis* stands in Pu'Er district, Yunnan. *J. Yunnan Univ.* **14**, 128–136 (1992).
192. Wu, Z. L., Dang, C. L., He, Z. R. & Wang, C. Y. A preliminary study on biomass and net primary productivity of *Quercus pannosa* forest in northwest Yunnan province, China. *J. Yunnan Univ.* **16**, 245–249 (1994).
193. Wu, Z. L., Dang, C. L., Wang, C. Y. & He, Z. R. A preliminary study on biomass of *Picea brachytyla* var. *complanata* forests in northwest Yunnan province, China. *J. Yunnan Univ.* **16**, 230–234 (1994).
194. Wu, Z. L., Dang, C. L., Wang, C. Y. & He, Z. R. A preliminary study on biomass of *Pinus densata* forests in northwest Yunnan province, China. *J. Yunnan Univ.* **16**, 220–224 (1994).
195. Wu, Z. L., Dang, C. L., Wang, C. Y. & He, Z. R. A preliminary study on net primary productivity of *Picea brachytyla* var. *complanata* forest in northwest Yunnan province, China. *J. Yunnan Univ.* **16**, 225–229 (1994).
196. Wu, Z. L., Dang, C. L., Wang, C. Y. & He, Z. R. A preliminary study on net primary productivity of *Pinus densata* forest in northwest Yunnan province, China. *J. Yunnan Univ.* **16**, 225–229 (1994).
197. Xia, H. B. Biomass and net primary production in different successional stages of karst vegetation in Maolan, SW China. *Guizhou For. Sci. Technol.* **38**, 1–7 (2010).
198. Xiao, Y. Biomass and productivity by natural *Pinus henryi* forests. *Chin. J. Plant Ecol.* **16**, 227–232 (1992).
199. Xiao, Y. Quantitative evaluation on the influence of environmental factors on timber production of Mukden pine. *J. Appl. Ecol.* **3** (1990).

200. Xiao, X. C. *et al.* Effects of stand density on biomass and productivity of *Pinus elliottii*. *J. Cent. South Univ. For. Technol.* **31**, 123–129 (2011).
201. Xie, S., Liu, W. & Li, S. Preliminary studies on the biomass of middle mountain moist evergreen broadleaved forests in Ailao mountain, Yunnan [China]. *Acta Phytoecol. Sin.* **20** (1996).
202. Xie, W. An investigation of biomass and its structure in Masson pine forests in mountainous areas of the north of Fujian province. *East China For. Manag.* **21**, 21–23 (2007).
203. Xie, S. C., Liu, W. Y., Li, S. C. & Yang, G. P. Preliminary studies on the biomass of middle-mountain moist evergreen broadleaved forests in Ailao Mountain, Yunnan. *Acta Phytoecol. Sin.* **20**, 167–176 (1996).
204. Xie, W. D., Ye, S. M., Yang, M. & Zhao, L. J. Biomass and Distribution pattern of *Pinus massoniana* plantation in southeast area of Guangxi. *J. Beihua Univ. Nat. Sci.* **10**, 68–71 (2009).
205. Xiong, D. G. *et al.* A study of annual growth and biomass of *Alnus formosana* in Yuanba District of Guangyuan City. *J. Sichuan For. Technol.* **27**, 55–58 (2006).
206. Xiu, Y. & Liu, X. Biomass and productivity study on Emei fir artificial forest. *J. Sichuan For. Sci. Technol.* **21**, 31–35 (2000).
207. Xu, Z. *et al.* Study on biodiversity and biomass in *Larix principis-rupprechtii* Mayr. plantation community in Yanshan mountainous region. *For. Res. Manag.* **0**, 43–49 (2010).
208. Xu, D. P., Yang, Z. J. & He, Q. X. Above ground biomass production and nutrient cycling of middle-age plantation of *Acacia mangium*. *For. Res.* **11**, 592–598 (1998).
209. Xue, P. P. *et al.* A study of carbon storage distribution patterns of typical *Pinus massoniana* forest ecosystem in the Three Gorges Reservoir area. *J. Sichuan For. Sci. Technol.* **32**, 62–67 (2011).
210. Yan, W. & Shi, Z. Biomass and net productivity of *Picea schrenkiana* var. *tianshanica* forest. *Chin. J. Appl. Ecol.* **4** (1994).
211. Yan, Z. Biomass and its allocation in a 28-year-old *Castanopsis kawakamii* plantation. *J. Fujian Coll. For.* **2** (1996).
212. Yan, H. Z., Lin, J. P., Wang, J., Su, D. & Miao, Y. DEM model space analysis on regional biomass based on ARCGIS. *J. Anhui Agric. Sci.* **39**, 852–855 (2011).
213. Yang, D. & Yang, X. Q. Studies on biomass and productivity of *Pinus tabulaeformis* plantation in the Wufengshang of Wudu, Gansu Province. *J. Northwest Nor. Univ. Nat. Sci.* **1**, 70–73 (2004).
214. Yang, D. *et al.* Comparison of biomass of *Betula alnoides* plantations and natural secondary forests in Xishuangbanna. *J. West China For. Sci.* **1** (2009).
215. Yang, Q. *et al.* Dynamics of biomass and net primary productivity in succession of south subtropical forests in southwest Guangdong. *J. Appl. Ecol.* **14**, 2136 (2003).
216. Yang, C. *et al.* Biomass distribution and net primary productivity of a 14-year-old stand of successive rotations of Chinese fir plantation. *J. Cent. South Univ. For. Technol.* **31**, 1–6 (2011).
217. Yang, D. J., Zhang, J. F., Qiu, Q. & Wang, W. B. Comparison of biomass of *Betula alnoides* plantations and natural secondary forests in Xishuangbanna. *J. West China For. Sci.* **38**, 77–81 (2009).
218. Yang, H. K. & Cheng, S. Z. Study on biomass of the karst forest community in Maolan, Guizhou province. *Acta Ecol. Sin.* **11**, 307–312 (1991).
219. Yang, K. & Guan, D. S. Forest biomass and its dynamics in Pearl River Delta. *Chin. J. Appl. Ecol.* **18**, 705–712 (2007).
220. Yang, L. L., Wen, S. Z., Wang, Z. Z. & He, G. X. Comparison between biomass and productivity of young *Alnus cremastogyne* Burkhill plantation under different site conditions. *J. Cent. South Univ. For. Technol.* **28**, 122–126 (2008).
221. Yang, Q. P., Li, M. G. & Li, R. W. Studies on the dynamic succession of *Pinus massoniana* community in Heishiding Natural Reserve. *Guihaia* **21**, 295–300 (2001).
222. Yang, Q. P., Li, M. G., Wang, B. S., Li, R. W. & Wang, C. W. Dynamics of biomass and net primary productivity in succession of south subtropical forests in southwest Guangdong. *Chin. J. Appl. Ecol.* **14**, 2136–2140 (2003).
223. Yang, Z., Zhang, J. P., Wan, D. J. & Zhang, X. B. Preliminary study on the biomass of artificial *Eucalyptus camaldulensis* Dehnl forests in arid-hot valleys, Yuanmou. *J. Mt. Sci.* **19**, 503–510 (2001).
224. Yao, Y. J., Kang, W. X. & Tian, D. L. Study of the biomass and productivity of *Cinnamomum camphora* plantation. *J. Cent. South For. Univ.* **23**, 1–5 (2003).
225. Ye, S. M., Wen, Y. G., Yang, M., Liang, H. W. & Lan, J. X. Correlation analysis on productivity and plant diversity of *Eucalyptus* plantations under successive rotation. *Acta Bot. Boreal.-Occident. Sin.* **30**, 1458–1467 (2010).
226. Ye, S. M. *et al.* Biomass and productivity of stratified mixed stands of *Eucalyptus urophylla* and *Acacia mangium*. *J. Beijing For. Univ.* **30**, 37–43 (2008).
227. Yi, Y. *et al.* On the biomass of secondary *Scikima superba* forest in Hangzhou. *J. Zhejiang For. Coll.* **2** (1993).
228. Yi, A. Y. Study on biomass and production of *Cryptomeria japonica* and *Cunninghamia lanceolata* mixed plantation. *Sichuan For. Exp. Des.* **3**, 50–52 (1998).
229. Yi, L. P., Wen, S. Z., Wang, Z. Z. & Yang, L. L. Biomass and productivity of *Liquidambar formosana* Hance. *J. Cent. South Univ. For. Technol.* **28**, 50–53 (2008).
230. Yi, W. M., Zhang, Z. P., Ding, M. M. & Wang, B. S. Biomass and efficiency of radiation utilization in *Erythrophleum fordii* community. *Acta Ecol. Sin.* **20**, 397–404 (2000).
231. Yin, G. Q., Tian, D. L., Fang, X., Xiang, W. H. & Deng, X. W. Studies on biomass of different young forests converted from farm land in Huitong, Hunan Province. *J. Cent. South Univ. For. Technol.* **30**, 9–14 (2010).
232. Yin, G. T. *et al.* Study on the above-ground biomass production and nutrition accumulation in various mixed plantations of forest trees and rattan. *For. Res.* **6**, 358–367 (1993).
233. You, W. *et al.* Biomass and net primary productivity of *Larix olgensis* plantation in Bingla Mountains, Northeast China. *J. Shenyang Agric. Univ.* **5** (2011).
234. Yuan, C. *et al.* The structural feature and biomass of soil and water conservative plantation of Armand Pine in the upper reach of the Yangtze River. *J. Northeast For. Univ.* **30**, 5–7 (2002).
235. Yuan, C. M. *et al.* The structural feature and biomass of soil and water conservation plantation of Armand Pine in the upper reach of the Yangtze River. *J. Northeast For. Univ.* **30**, 5–7 (2002).
236. Yue, F. & Yang, B. Effect of different control methods on carbon sink function of artificial *Pinus kesiya* var. *langbianensis* plantation. *J. Anhui Agric. Sci.* **39**, 3433–3435 (2011).
237. Yue, J. W. *et al.* Biomass characteristics of the deteriorated *Pinus massoniana* forests in red soil low hilly. *China For. Sci. Technol.* **24**, 19–22 (2010).
238. Zan, Q. J., Wang, Y. J. & Liao, W. B. Biomass and net productivity of *Sonneratia carseolaria*-*Sonneratia apetala* mangrove forest. *J. Wuhan Bot. Res.* **15**, 391–397 (2001).
239. Zeng, S. C. *et al.* The biomass and water-holding capacities of some forest communities of Baiyunshan scenic spot, Guangzhou. *J. South China Agric. Univ. Nat. Sci. Ed.* **23**, 41–44 (2001).
240. Zhang, B. & Chen, C. The red star ChangWuXian locust plantation forest biomass and production. *Shaanxi For. Sci. Technol.* **3**, 13–17 (1992).
241. Zhang, F. & Shang, G. Close the emperor mountain north China larch forest community characteristics and biomass. *J. Shanxi Agric. Univ. Nat. Sci. Ed.* **15**, 72–77 (1992).
242. Zhang, J. & Yuan, Y. A Study on the biomass and production of *Pinus fenzieliana* forest in Hainan. *Guizhou For. Sci. Technol.* **3** (1987).

243. Zhang, L. *et al.* Biomass and net primary productivity of secondary evergreen broadleaved forest in Huangmian Forest Farm, Guangxi. *J. Appl. Ecol.* **15**, 2029–2033 (2004).
244. Zhang, Q. *et al.* A study on biomass and productivity of *Populus x euramericana* cv. 'San Martino' (I - 72/58) plantation on beach land of Yangtze River. *For. Res.* **21**, 542–547 (2008).
245. Zhang, W. Studies on the biomass and productivity of Slash Pine plantation. *For. Prospect Des.* **2**, 11–16 (2010).
246. Zhang, X. & Shang, Z. The bio-cycle patterns of nutrient elements and stand biomass in forest communities in Hilly Loess Regions. *Acta Ecol. Sin.* **3**, 527–537 (2005).
247. Zhang, X. *et al.* Distribution pattern of nutrient elements in masson pine plantation of Anhui province. *Chin. J. Appl. Ecol.* **4**, 7–11 (1993).
248. Zhang, J. X. & Yuan, Y. Z. A study on the biomass and productivity of *Pinus fenzeliana* forest in Hainan. *Acta Phytoecol. Geobot. Sin.* **12**, 63–69 (1988).
249. Zhang, L., Huang, Y., Luo, T. X., Dai, Q. & Deng, K. M. Age effects on stand biomass allocations to different components: A case study in forests of *Cunninghamia lanceolata* and *Pinus massoniana*. *J. Grad. Sch. Chin. Acad. Sci.* **22**, 170–178 (2005).
250. Zhang, Z. J. *et al.* Study on the biomass structure and distribution of natural secondary forest of *Pinus massoniana*. *J. Agric. Univ. Hebei* **29**, 37–43 (2006).
251. Zhang, Z. P. & Ding, M. M. Biomass and efficiency of radiation utilization in monsoon evergreen broadleaved forest in Dinghushan Biosphere Reserve. *Acta Ecol. Sin.* **16**, 525–534 (1996).
252. Zhang, Z. P., Ding, M. M., Zai, G. L., Yi, W. M. & Huang, Y. J. Biomass and net primary production of *Cryptocarya concinna* community in Dinghushan. *J. Ecol. Sci.* **1**, 8–11 (1991).
253. Zhang, Z. P., He, D. Q. & Ao, H. X. The biomass and solar energy utilization efficiency in *Zenia insignis* forest. *Acta Phytoecol. Sin.* **20**, 502–509 (1996).
254. Zhao, K. & Tian, D. L. Study of the biomass and productivity of mature Chinese fir stand in Huitong County. *J. Cent. South For. Univ.* **20**, 7–13 (2000).
255. Zheng, Z. *et al.* Forest structure and biomass of a tropical seasonal rain forest in Xishuangbanna, southwest China. *Biotropica* **38**, 318–327 (2006).
256. Zheng, Z. *et al.* A study on biomass of the primary tropical seasonal rain forest in Xishuangbanna. *Guihaia* **19**, 309–314 (1999).
257. Zheng, Z., Feng, Z. L., Cao, M., Liu, H. M. & Liu, L. H. Biomass and net primary production of primary tropical wet seasonal rainforest in Xishuangbanna. *Acta Phytoecol. Sin.* **24**, 197–203 (2000).
258. Zheng, Z., Feng, Z. L., Liu, H. M., Meng, Y. & Gan, J. M. Influence of planting *Amomum villosum* in tropical wet seasonal rainforest on biomass of the rainforest in Xishuangbanna. *J. Mt. Sci.* **19**, 237–242 (2001).
259. Zheng, Z., Liu, H. M. & Feng, Z. L. Biomass of tropical montane rain forest in Xishuangbanna of southwest China. *Chin. J. Ecol.* **25**, 347–353 (2006).
260. Zheng, Z., Liu, L. H., Feng, Z. L., Liu, H. M. & Cao, M. The net primary production of the tropical seasonal rain forest in Xishuangbanna. *J. Mt. Sci.* **17**, 212–217 (1999).
261. Zhong, X. H. & Luo, J. Study of characteristics of ecological function of the natural and degraded ecosystems in the mountain dark coniferous forest zone in Mt. Gonggashan. *J. Mt. Sci.* **19**, 201–206 (2001).
262. Zhou, Q. Y. *et al.* Biomass-and energy allocation in *Eucalyptus urophylla* x *Eucalyptus tereticornis* plantations at different stand ages. *J. Appl. Ecol.* **21**, 16–22 (2010).
263. Zhou, S. Q. & Huang, J. Y. A study on biomass and productivity of *Larix mastersiana* plantation in Sichuan. *Acta Phytoecol. Geobot. Sin.* **15**, 9–16 (1991).
264. Zhou, Z. Z., Zheng, H. S., Yang, Z. J. & Yin, G. T. Research on spatial pattern of biomass and nutrient element in plantation of rubber intercropped with *Amomum longiligulare*. *Acta Ecol. Sin.* **17**, 225–233 (1997).
265. Zhu, X., Shi, Q. & Li, Y. A preliminary study on the qinghai's treasure house of the forest biomass and shrubs. *Sci. Technol. Qinghai Agric. For.* (1993).
266. Lesiv, M. *et al.* Global forest management data for 2015 at a 100 m resolution. *Sci. Data* **9**, 199 (2022).
267. Hijmans, R. & van Etten, J. *raster: Geographic Data Analysis and Modeling*. R Package Version 3.6-11. <https://cran.r-project.org/web/packages/raster/index.html> (2022).
268. Pebesma, E. Simple Features for R: standardized support for spatial vector data. *The R Journal* **10**, 439–446 (2018).
269. Dubayah, R. *et al.* *GEDI L2B Canopy Cover and Vertical Profile Metrics Data Global Footprint Level V001*. https://doi.org/10.5067/GEDI/GEDI02_B.001 (2020).
270. Potapov, P. *et al.* Mapping and monitoring global forest canopy height through integration of GEDI and Landsat data. *Remote Sens. Environ.* **253**, 112165 (2020).
271. Beck, H. E. *et al.* Bias correction of global high-resolution precipitation climatologies using streamflow observations from 9372 catchments. *J. Clim.* **33**, 1299–1315 (2020).
272. Karger, D. N. *et al.* Climatologies at high resolution for the earth's land surface areas. *Sci. Data* **4**, 170122 (2017).
273. Karger, D. N. *et al.* Climatologies at high resolution for the earth's land surface areas. <https://doi.org/10.5061/dryad.kd1d4> (2022).
274. Trabucco, A. & Zomer, R. Global Aridity Index and Potential Evapotranspiration (ET0) Climate Database v2. [figshare https://doi.org/10.6084/m9.figshare.7504448.v3](https://doi.org/10.6084/m9.figshare.7504448.v3) (2019).
275. Amatulli, G. *et al.* A suite of global, cross-scale topographic variables for environmental and biodiversity modeling. *Sci. Data* **5**, 180040 (2018).
276. Gilbert, M. *et al.* Global cattle distribution in 2010 (5 minutes of arc). *Harvard Dataverse* <https://doi.org/10.7910/DVN/GIVQ75> (2018).
277. Ibisch, P. L. *et al.* A global map of roadless areas and their conservation status. *Science* **354**, 1423–1427 (2016).
278. UNEP-WCMC. *User Manual for the World Database on Protected Areas and world database on other effective area-based conservation measures: 1.6*. 79 pp. UNEP-WCMC: Cambridge, UK (2019).
279. Venter, O. *et al.* Sixteen years of change in the global terrestrial human footprint and implications for biodiversity conservation. *Nat. Commun.* **7**, 12558 (2016).
280. Batjes, N. H. Harmonized soil property values for broad-scale modelling (WISE30sec) with estimates of global soil carbon stocks. *Geoderma* **269**, 61–68 (2016).
281. Silva, C. A. *et al.* *rGEDI: NASA's Global Ecosystem Dynamics Investigation (GEDI) Data Visualization and Processing*. R Package Version 0.1.11. <https://CRAN.R-project.org/package=rGEDI> (2021).
282. Harrell, F. E. Jr *Hmisc: Harrell Miscellaneous*. R package version 4.7-1. <https://CRAN.R-project.org/package=Hmisc> (2022).
283. Breiman, L. Random forests. *Mach. Learn.* **45**, 5–32 (2001).
284. James, G., Witten, D., Hastie, T. & Tibshirani, R. *An Introduction to Statistical Learning with Applications in R*. (Springer, 2013).
285. Liaw, A. & Wiener, M. Classification and regression by randomForest. *R News* **2/3**, 18–22 (2002).
286. Cortes, C. & Vapnik, V. Support-vector networks. *Mach. Learn.* **20**, 273–297 (1995).
287. Meyer, D., Dimitriadou, E., Hornik, K., Weingessel, A. & Leisch, F. *e1071: Misc Functions of the Department of Statistics, Probability and Theory Group (Formerly: E1071), TU Wien*. R package version 1.7–12. <https://CRAN.R-project.org/package=e1071> (2022).
288. Chen, T. *et al.* *xgboost: extreme gradient boosting*. R package version 1.6.0.1. <https://CRAN.R-project.org/package=xgboost> (2022).
289. Liang, J. *et al.* Co-limitation towards lower latitudes shapes global forest diversity gradients. *Nat. Ecol. Evol.* **6**, 1423–1437 (2022).

290. Kampichler, C., Wieland, R., Calmé, S., Weissenberger, H. & Arriaga-Weiss, S. Classification in conservation biology: a comparison of five machine-learning methods. *Ecol. Inform.* **5**, 441–450 (2010).
291. R Core Team. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing <https://www.R-project.org/> (2021).
292. Kulkarni, A., Chong, D. & Batarseh, F. A. in *Data Democracy* 1st Edition Ch. 5 (Academic Press, 2020).
293. Hossin, M. & Sulaiman, M. N. A review on evaluation metrics for data classification evaluations. *Int. J. Data Min. Knowl. Manag. Process* **5**, 1–11 (2015).
294. Delgado, R. & Tibau, X. A. Why Cohen's Kappa should be avoided as performance measure in classification. *PLoS one* **14**, e0222916 (2019).
295. Forestry Agency, Ministry of Agriculture, Forestry and Fisheries of Japan. *The National Forest Inventory of Japan* <http://forestbio.jp/datafile/datafile.html> (2023).
296. Korea Forest Research Institute. *The 5th National Forest Inventory - Field manual* (Korea Forest Research Institute, 2008).
297. Dyer, J. M. Revisiting the deciduous forests of eastern North America. *BioScience* **56**, 341–352 (2006).
298. Iverson, L. R., Peters, M. P., Prasad, A. M. & Matthews, S. N. Analysis of climate change impacts on tree species of the eastern US: Results of DISTRIB-II modeling. *Forests* **10**, 302 (2019).
299. Branco, P., Ribeiro, R. P. & Torgo, L. UBL: an R package for Utility-based Learning. <http://arxiv.org/abs/1604.08079> (2016).
300. Kuhn, M. Building Predictive Models in R Using the caret Package. *J. Stat. Softw.* **28**, 26 (2008).
301. Abbasi *et al.* Spatial database of planted forests in East Asia using machine learning (final products). *figshare* <https://doi.org/10.6084/m9.figshare.21774725.v3> (2023).
302. Besnard, S. *et al.* Mapping global forest age from forest inventories, biomass and climate data. *Earth Syst. Sci. Data* **13**, 4881–4896 (2021).
303. Du, Z. *et al.* A global map of planting years of plantations. *Sci. Data* **9**, 141 (2022).
304. Du, Z. *et al.* A global map of planting years of plantations v2. *figshare* <https://doi.org/10.6084/m9.figshare.19070084.v2> (2022).
305. Abbasi *et al.* Spatial distribution of planted forests in East Asia (code, training datasets, and models). *figshare* <https://doi.org/10.6084/m9.figshare.21774812.v2> (2023).
306. Potapov, P. *et al.* The Global 2000–2020 Land Cover and Land Use Change Dataset Derived From the Landsat Archive: First Results. *Fron. Remote Sens.* **3** (2022).
307. Hansen, M. C. *et al.* High-resolution global maps of 21st-century forest cover change. *Science* **342**, 850–853 (2013).
308. Liang, J. & Gamarra, J. G. P. The importance of sharing global forest data in a world of crises. *Sci. Data* **7**, 1–5 (2020).
309. National Forestry and Grassland Administration. *China Forest Resources Report (2014–2018)*. (China Forestry Publishing House, 2019).
310. Xiang, Y. *et al.* Mixed plantations enhance more soil organic carbon stocks than monocultures across China: Implication for optimizing afforestation/reforestation strategies. *Sci. Total Environ.* **821**, 153449 (2022).

Acknowledgements

This study is supported, in parts, by the World Resources Institute (WRI) project “Mapping planted forests in China,” and by the Department of Forestry and Natural Resources, Purdue University. We thank the Global Forest Biodiversity Initiative (GFBI) and Science-I (Project#EA-2023-001) for facilitating data sharing and international research collaboration. We also thank the Institute for a Sustainable Future (ISF) at Purdue University for facilitating Purdue-wide research collaboration. A.O.A. benefited from Takenaka Scholarship Foundation. X.T. benefited from the National Natural Science Foundation of China (32271856). M.H. was supported by the CGIAR MITIGATE+ project.

Author contributions

J.L., N.L.H. and E.D.G. conceived the study. X.T. and A.O.A. coordinated data compilation. A.O.A. conducted data analysis. A.O.A. and X.T. wrote the initial draft. J.L. supervised the manuscript development. Everyone contributed to the writing, revision, and editing.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41597-023-02383-w>.

Correspondence and requests for materials should be addressed to J.L.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023