



RESEARCH PAPER

Open Access



Combining Weibull distribution and k -nearest neighbor imputation method to predict wall-to-wall tree lists for the entire forest region of Northeast China

Yuanyuan Fu^{1,2,3} , Hong S. He^{4*} , Shaoqiang Wang^{1,3} and Lunche Wang^{1,3}

Abstract

Key message: We propose a coupled framework to combine the strengths of the Weibull function in modeling diameter distributions and the ability of the k -nearest neighbor (k NN) method to impute spatially continuous forest stand attributes for the prediction of wall-to-wall tree lists (lists of stems per hectare by species and diameter at breast height (DBH)) at regional scales. The tree lists of entire Northeast China's forests predicted by the above framework reasonably reflect the species-specific tree density and diameter distributions.

Context: Detailed tree lists provide information about forest stocks disaggregated by species and size classes, which are crucial for forest managers to accurately characterize the current forest stand state to formulate targeted forest management strategies. However, regional tree list information is still lacking due to limited forest inventory.

Aims: We aimed to develop a coupled framework to enable the prediction of wall-to-wall tree lists for the entire forest region of Northeast China, then analyze the species-specific diameter distributions and reveal the spatial patterns of tree density by species.

Methods: A two-parameter Weibull function was used to model the species-specific diameter distributions in the sample plots, and a maximum likelihood estimation (MLE) was used to predict the parameters of the Weibull distributions. The goodness-of-fit of the predicted species-specific Weibull diameter distributions in each plot was evaluated by Kolmogorov-Smirnov (KS) test and an error index. The k NN model was used to impute the pixel-level stand mean DBH.

Results: Weibull distribution accurately described the species-specific diameter distributions. The imputed stand mean DBH from the k NN model showed comparable accuracy with earlier studies. No difference was detected between predicted and observed tree lists, with a small error index (0.24–0.58) of diameter distributions by species. The fitted species-specific diameter distributions generally showed a right-skewed unimodal or reverse J-shaped pattern.

Handling editor: Andreas Bolte

*Correspondence: heh@missouri.edu

⁴ School of Natural Resources, University of Missouri, Columbia, MO 65211, USA

Full list of author information is available at the end of the article



© The Author(s) 2022. **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 licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence 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 licence, visit <http://creativecommons.org/licenses/by/4.0/>.

Conclusion: Overall, the coupled framework developed in this study was well-suited for predicting the tree lists of large forested areas. Our results evidenced the spatial patterns and abundance of tree species in Northeast China and captured the forest regions affected by disturbances such as fire.

Keywords: Weibull function, *k*NN imputation, Coupled framework, Diameter distribution, Tree density

1 Introduction

As knowledge of the total stand volume, aboveground biomass, or basal area is insufficient, forest managers cannot accurately characterize the current stand state, project the future growth and yield, formulate the harvest schedules, and prioritize the stands for treatments (e.g., thinning or regeneration harvest) (Mauro et al. 2019). Accordingly, detailed information about forest stocks disaggregated by species and size class is of primary importance for effective forest resource management (Brosofske et al. 2014). Notably, detailed tree lists [lists of stems per hectare by species and diameter at breast height (DBH)] (Temesgen et al. 2003) can be used to satisfy the above requirements. Therefore, the prediction of spatially continuous tree lists is urgently needed for these evaluations and is an important output of a complete forest inventory.

Tree lists can be used as inputs for tree and stand projection models (growth and yield models) to predict tree growth (Lamb et al. 2018) and assess stand-level structures (Temesgen et al. 2003). In particular, regional tree lists provide key information for the initialization of forest landscape models (e.g., LANDIS PRO) to simulate the long-term effects of succession, climate change, and disturbances (e.g., fire and pests) on forest composition over large areas (Duan et al. 2022; Huang et al. 2021). Therefore, spatially explicit tree lists are required to develop tactical forest management in response to various disturbances. The generated maps of wall-to-wall tree lists can be used as valuable indicators by forest managers to intuitively understand the species-specific density distribution at different DBH classes, evaluate the necessity of thinning, and plan the location and intensity of the next selective cuttings (Valbuena et al. 2014).

The collection of detailed tree list information contiguously across large forest regions is often not practical as it is time-consuming, labor-intensive, and spatially constrained (Zhang et al. 2019). In Northeast China, forests cover approximately 45 million hectares, and most forested lands are located in mountains with few roads. Due to the large forested land base and difficult access, detailed tree lists are only available for limited sample stands or plots. Stand-level inventory information, such as stand DBH, stand height, and species composition, collected from the national forest inventory (Zhang and Liang 2014), is available for thousands of stands in each

forestry bureau. This information can be imputed to large areas by integrating remotely sensed data with complete spatial coverage (Fu et al. 2019; Zhang et al. 2018a). Therefore, an emerging need exists to develop a cost-effective methodology that integrates forest inventory data, remotely sensed data, and other ancillary data sets to predict wall-to-wall tree lists for forest regions based on limited sample plot data.

Tree lists can be approximated using species-specific diameter distributions for mixed forest stands (Brosofske et al. 2014). Therefore, many studies have used stand- or plot-level information measured on the ground to generate tree lists using a parametric diameter distribution modeling approach (Bankston et al. 2021; Olofsson and Olsson 2018). This approach usually uses a probability density function (PDF) (e.g., gamma, beta, exponential, log-normal, Johnson's SB, Weibull) to fit the diameter distribution for each stand and then predicts the PDF parameters using stand attributes (Liu et al. 2009; Temesgen et al. 2003). The two-parameter Weibull function has been reported to be the simplest and most accurate PDF for diameter distribution modeling for tree species worldwide, owing to its flexibility in describing different shapes of diameter distributions and fewer parameters (Diamantopoulou et al. 2015; Schmidt et al. 2020; Schütz and Rosset 2020). Schmidt et al. (2020) modeled the diameter distributions of clonal eucalyptus stands in Brazil using a Weibull function and re-measured information from 56 permanent sample plots as predictor variables. Schütz and Rosset (2020) compared different methods of predicting Weibull distribution parameters to better describe the diameter distributions of Norway spruce (*Picea abies* (L.) Karst.) and European beech (*Fagus sylvatica* L.) in monospecific, regular temperate European forests. However, the diameter distribution modeling approach is only suitable for simple stands with few species and small contiguous areas (Temesgen et al. 2003). For large regions, the application of the Weibull diameter distribution is limited by the availability of the stand attributes that are assumed to have strong correlations with the PDF parameters, which can be predicted by nonparametric imputation methods (Fu et al. 2019).

The nonparametric *k*-nearest neighbor (*k*NN) imputation model has been widely applied to predict wall-to-wall forest stand attributes across large regions as

it can simultaneously predict a suite of response variables, such as species-specific volume (Breidenbach et al. 2010), species-level biomass (Fu et al. 2019), basal area (Wilson et al. 2012), or stand mean DBH (Zhang et al. 2018b). k NN imputation is also a reliable strategy for the attribute prediction of complex forest stands as it can maintain both the spatial and attribute covariance structures of the stand attributes (Temesgen et al. 2003). The process of k NN imputation for generating wall-to-wall stand attributes (response variables) involves computing the statistical distance between the unsampled target pixel and the reference samples (neighbors) and then assigning the mean value of the k NNs' attributes to the target pixel (Brosofske et al. 2014). In the development of a k NN imputation model, in addition to the suitable selection of distance metrics and k values, the selection of predictor variables can be more crucial. The predictor variables should be functionally or statistically related to the response variables (Eskelson et al. 2009). The predictor variables used in the k NN model built for stand attributes imputation usually include soil, climate, topographic variables, and vegetation indices derived from optical remote sensing imagery, such as Landsat (Gjertsen 2007) and Moderate Resolution Imaging Spectroradiometer (MODIS) (Fu et al. 2019; Zhang et al. 2018a). The use of airborne light detection and ranging (LiDAR) in the k NN model for the prediction of forest attributes has been rapidly increasing and is promising for application in specific forest management (Hudak et al. 2008; Mauro et al. 2019). For example, the accurate prediction of fine-scale (e.g., tree-level or stand-level) attributes might require predictor variables with three-dimensional and high-resolution information. However, LiDAR lacks imaging capabilities and cannot provide spatially continuous maps (Chi et al. 2015). Therefore, optical imagery is more suitable for use in the k NN model to generate the wall-to-wall stand attributes across large regions.

The objective of this study was to develop a coupled framework combining the strengths of the Weibull function in modeling diameter distributions and the ability of the k NN method to impute spatially continuous forest stand attributes for the prediction of wall-to-wall tree lists at regional scales based on limited sample plot data via the integration of national forest inventory, remotely sensed, and other ancillary data. Further, the coupled framework was tested and assessed, and the wall-to-wall tree lists at a 250-m spatial resolution were predicted for the entire forest region of Northeast China. Finally, the species-specific diameter distributions were analyzed and the spatial patterns of tree density by species were revealed.

2 Material and methods

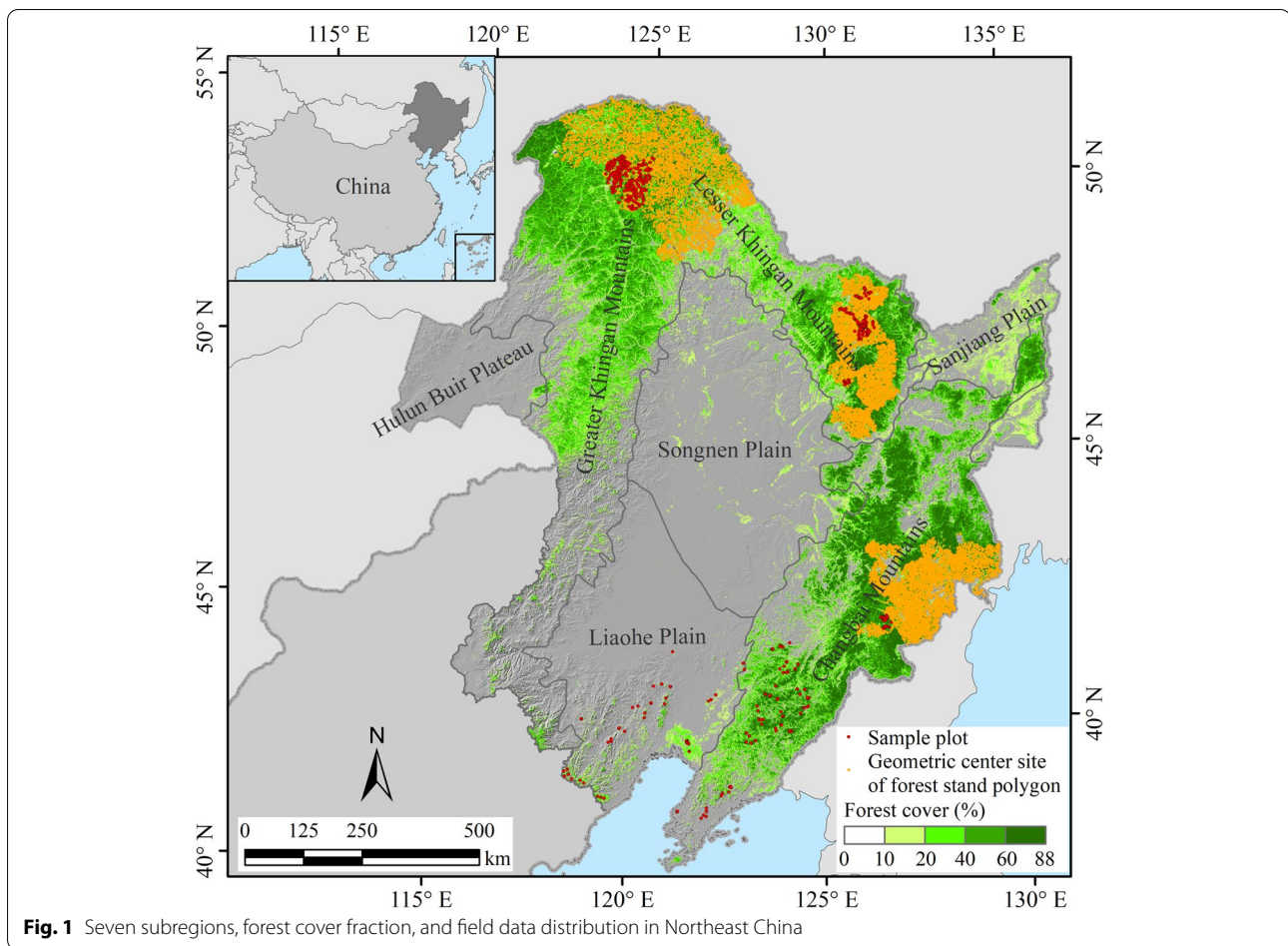
2.1 Study area

The study area is located in northeastern China (38° 42'–53° 35' N, 115° 32'–135° 09' E) and covers 124 million hectares, encompassing three provinces (Heilongjiang, Jilin, and Liaoning) and the eastern part of the Inner Mongolia Autonomous Region. The area is divided into seven major subregions that reflect diverse climates, terrains, soils, and vegetation (Fu et al. 2019): the Greater Khingan Mountains (GKM), Lesser Khingan Mountains (LKM), Changbai Mountains (CBM), Sanjiang Plain (SJP), Songnen Plain (SNP), Liaohe Plain (LHP), and Hulun Buir Plateau (HBP) (Fig. 1). Approximately 40% of the region is forested (Pan et al. 2011), with forests mainly distributed over the Greater Khingan, Lesser Khingan, and Changbai Mountains. The forests in Northeast China are mixed, spanning various forest types, such as cold-temperate conifer mixed forests and temperate coniferous and broad-leaved mixed forests. Cold-temperate conifer mixed forests are distributed in the northern part of the Greater Khingan Mountains and are dominated by larch (*Larix gmelinii* (Rupr.) Kuzen), white birch (*Betula platyphylla* Suk.), aspen (*Populus davidiana* Dode), spruce (*Picea koraiensis* Nakai), Asian black birch (*Betula davurica* Pall., hereafter black birch), Mongolian oak (*Quercus mongolica* Fisch. ex Ledeb.), Mongolian Scotch pine (*Pinus sylvestris* var. *mongolica* Litv., hereafter Scotch pine), and willow (*Chosenia arbutifolia* (Pall.) A. Skv.). Temperate coniferous and broad-leaved mixed forests are distributed in the Lesser Khingan and Changbai Mountains, mainly composed of Korean pine (*Pinus koraiensis* Sieb. et Zucc.), larch, fir (*Abies nephrolepis* (Trautv.) Maxim.), spruce, Mongolian oak, white birch, aspen, ribbed birch (*Betula costata* Trautv.), basswood (*Tilia amurensis* Rupr.), mono maple (*Acer mono* Maxim.), elm (*Ulmus pumila* L.), Manchurian walnut (*Juglans mandshurica* Maxim., hereafter walnut), Manchurian ash (*Fraxinus mandschurica* Rupr., hereafter ash), and Amur corktree (*Phellodendron amurense* Rupr.).

2.2 Field data

Two types of field data (plot and stand) were collected for this study. The limited sample plot data with detailed tree lists were used to build species-specific Weibull distribution parameter prediction models (WDPPM_s), and the stand inventory data without tree lists were used as response variables in the k NN imputation model to predict the wall-to-wall stand attributes.

The detailed tree lists of 641 rectangular (20 m × 30 m) sample plots were measured from 2009 to 2013 in the study area. The sample plots were mainly distributed over



the Greater Khingan Mountains, Lesser Khingan Mountains, and Changbai Mountains (Fig. 1) and contained 17 major tree species in Northeast China. The plot number, GPS coordinates, elevation, slope, forest type, and canopy density for each plot were recorded. Within each plot, tree species, diameter at breast height (DBH), and tree height for every tree with a DBH larger than 5 cm were measured. The arithmetic mean DBH by species in each plot was calculated by dividing the sum of tree DBHs by the number of corresponding tree species.

The stand inventory data were derived from the National Secondary Forest Resources Continuous Inventory in China and organized as stand polygons. A total of 25,000 forest stand polygons (average polygon size: 20.6 ha) from Northeast China surveyed in the early 2000s were collected from the National Forestry and Grassland Data Center (<http://www.cfsdc.org/>). Each stand polygon was a contiguous area ranging from a few hectares to tens of hectares with relatively homogeneous forest attributes, such as stand mean DBH, stand height, stand age, stand volume density, tree species composition (volume proportion by species), and dominant species.

The stand mean DBH is the DBH corresponding to the average basal area of the stand, which differs from the arithmetic mean DBH. In this study, owing to the lack of precise conversion relationships between stand mean DBH and mean arithmetic DBH of major tree species in Northeast China, the stand mean DBH was used instead of the mean arithmetic DBH by species to predict the wall-to-wall Weibull diameter distribution parameters by species. The geometric center sites of 25,000 stand polygons are shown in Fig. 1.

Polygon boundaries were generated by interpreting aerial photos in accordance with the technical regulations for forest management planning and design inventory in China. Within each polygon, stand attributes were estimated using angle-count sampling. Each plot was more than 50-m away from the stand boundary, and the distance between the two angle-count plots was at least 100 m (Zhang et al. 2018a). Trees with a DBH larger than 5 cm were counted in each angle gauge plot (basal area factor = 1), and the DBH of each tree was converted to volume according to the species-specific DBH-volume relationships summarized by the National Forestry and

Grassland Data Center (<http://www.cfsdc.org/>). The volume density of each angle gauge plot was derived by aggregating all single-tree volume estimates counted by species. Further, the stand volume density was estimated by averaging the volume density of all angle gauge plots in each polygon (Zhang et al. 2018a). The species-level biomass of each stand polygon was transformed from the species-level volume based on species-specific biomass-volume relationships (Fang et al. 1998).

2.3 MODIS and environmental data

The selection of predictor variables is crucial when the k NN imputation method is employed to predict wall-to-wall stand attributes (e.g., stand mean DBH or species-level biomass). Forest stand attributes can be characterized by rich spectral information from multi-spectral remote sensing images and are influenced by environmental factors. In this study, seven Moderate Resolution Imaging Spectroradiometer (MODIS) surface reflectance bands from MOD09Q1 (b1–b2; 250 m) and MOD09A1 (b3–b7; 500 m) from 2000 were selected as basic data to derive the predictor variables for further stand attribute imputation. By averaging all reflectance values of each month in 2000, the above seven reflectance bands were processed into monthly data and resampled to 250 m using nearest-neighbor interpolation. Several spectral indices that correlate with the stand characteristics (Fu et al. 2019) were calculated from the monthly reflectance bands (Additional file 1: Table S1). As the surface reflectance in Northeast China is largely affected by snow or frost cover from January to April and November to December (Zhang et al. 2013), only the monthly MODIS reflectance bands and spectral indices from May to October were used as predictor variables. To avoid imputing stand attributes to non-forest areas, forest areas that were defined as tree cover greater than 10% (Schmitt et al. 2009) were extracted using the MODIS Vegetation Continuous Fields (VCF) product MOD44B (250 m) for the year 2000.

Environmental data (e.g., topographic, climate, and soil data) that correlate with the stand attributes were selected as auxiliary predictor variables (Additional file 1: Table S1) to reduce prediction uncertainties due to environmental heterogeneity. Topographic data (elevation, slope, and cosine of aspect) were derived from the Shuttle Radar Topography Mission (SRTM) digital elevation model (DEM), with a 90-m spatial resolution (<http://srtm.csi.cgiar.org/>). Climate data (mean annual temperature and cumulative precipitation during 1982–2015; 250 m) were interpolated from the 103 meteorology stations in Northeast China (<https://data.cma.cn/>) using ANUSPLIN 4.3 software, in which a thin-plate spline function with the resampled SRTM DEM (250

m) as the covariate was used (Fu et al. 2019). Other climate data, such as mean annual radiation and potential evapotranspiration from 1982 to 2015 with a spatial resolution of $0.1^\circ \times 0.1^\circ$, were derived from the China Meteorological Forcing Dataset (<http://westdc.westgis.ac.cn>). Soil data with a 1-km spatial resolution were derived from the Harmonized World Soil Database v1.2 (Fao/Iiasa/Isric/Isscas/Jrc 2012). To match the spatial resolution of the MODIS data, the topographic data and other environmental data were resampled to a 250-m spatial resolution using bilinear interpolation and nearest-neighbor interpolation, respectively.

2.4 Overall framework for wall-to-wall tree list prediction

The framework for wall-to-wall tree list prediction included the following three steps (Fig. 2): (1) building Weibull distribution parameter prediction models (WDPPM_S) with limited tree list data collected from sample plots to predict species-specific diameter distributions in each pixel; (2) mapping spatially continuous stand attributes (e.g., species-level biomass and stand mean DBH) using the k NN imputation models based on stand inventory data, MODIS, and environmental data, and providing the independent variables for WDPPMs in each pixel; and (3) predicting the wall-to-wall tree lists over Northeast China by combining the Weibull diameter distributions predicted by the WDPPM_S and forest attributes imputed by the k NN models.

2.4.1 Building Weibull distribution parameter prediction models (WDPPM_S)

The two-parameter Weibull function was used to model the diameter distribution of the 641 plots by species; the function is expressed as follows:

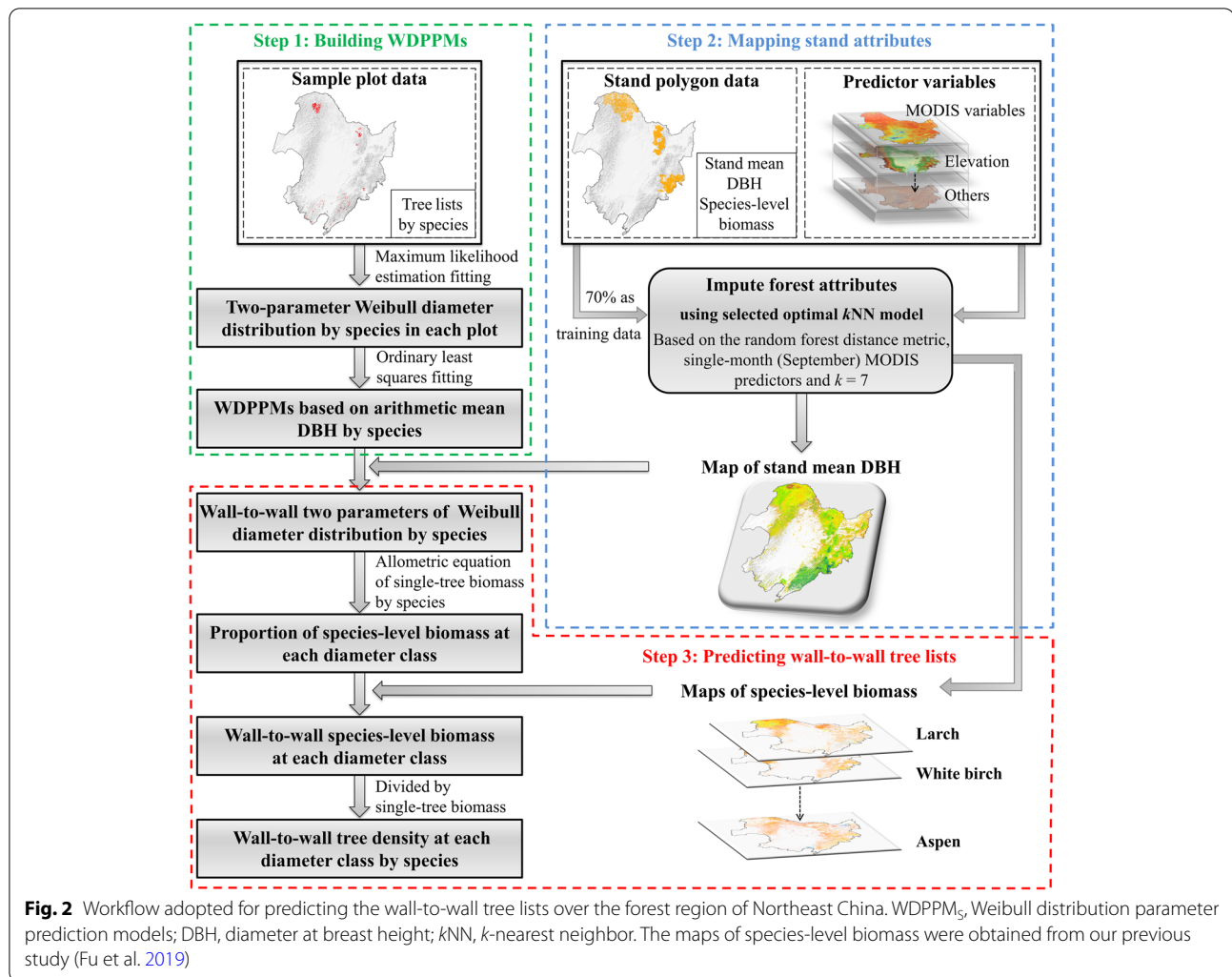
$$f(x) = \frac{b}{c} \left(\frac{x}{c}\right)^{b-1} \exp \left[-\left(\frac{x}{c}\right)^b \right] \quad (1)$$

where $f(x)$ represents the probability density function for x , x represents the DBH, b and c are the shape and scale parameters of the Weibull function, respectively.

The maximum likelihood estimation (MLE) was used to predict the parameters of the two-parameter Weibull function for each species in each plot. Owing to its good statistical characteristics and applicability, MLE is generally considered to be the most frequently used method for predicting distribution parameters (Diamantopoulou et al. 2015). The MLE equation is as follows:

$$L(x_1, x_2, \dots, x_n; \theta) = \prod_{i=1}^n f(x_i; \theta) \quad (2)$$

where x_1, x_2, \dots, x_n are the sample observations, $L(x_1, x_2, \dots, x_n; \theta)$ is the MLE function, and $f(x_i; \theta)$ is the



probability density function based on the unknown parameter θ . The maximum probability of the sample observations (x_1, x_2, \dots, x_n) is necessary when the MLE method is used to predict the value of θ . If the MLE function (L) is differentiable, the value of θ must satisfy the following formula:

$$\frac{dL}{d\theta} = 0 \quad (3)$$

The unknown parameter θ can be predicted according to (3). The MLE method was implemented using the R package, MASS. Once the shape (b) and scale (c) parameters of the Weibull diameter distributions for each species in each plot were obtained, the relationship between these two parameters and stand variables could be established. In this study, the arithmetic mean DBH by species was selected as the independent variable to

build WDPPM_s for each species using the ordinary least squares (OLS) method.

2.4.2 Mapping stand attributes

The wall-to-wall stand attributes, including species-level biomass and stand mean DBH, were mapped by integrating the forest attributes extracted from stand polygons (as response variables) with spatially continuous predictor variables (e.g., MODIS and environmental data), as described in Section 2.3, using the k NN imputation model. By comparing the prediction accuracy of 630 k NN models in imputing stand attributes, our previous study mapped the wall-to-wall species-level biomass in 2000 for forests in Northeast China using the optimal k NN model based on the random forest distance metric, single-month (September) MODIS predictor variables, and $k = 7$ (Fu et al. 2019). In this study, the k NN imputation model was

used to map the stand mean DBH in 2000 for the forested region of Northeast China. The imputation process was implemented using the `yaImpute` package in R.

The k NN imputation method involves the identification of the k -nearest reference samples in a feature space defined by the predictor variables for each target pixel. The values of each response variable within these k -nearest samples are averaged and assigned to the target pixel. Formally, the nearest neighbors prediction, \tilde{y}_i , for the i th target pixel is calculated as follows (McRoberts 2012):

$$\tilde{y}_i = \sum_{j=1}^k w_{ij} y_j^i \quad (4)$$

where $\{y_j^i, j=1, 2, \dots, k\}$ is the set of response variable observations for the k reference pixels that are nearest to the i th target pixel in feature space with respect to a given distance metric, and w_{ij} is the weight assigned to the j th nearest neighbor with $\sum_{j=1}^k w_{ij} = 1$. The w_{ij} is defined as follows:

$$w_{ij} = \frac{1/(1+d_{ij})}{\sum_{j=1}^k [1/(1+d_{ij})]} \quad (5)$$

where d_{ij} is the distance between the j th nearest-neighbor reference pixel and i th target pixel.

To execute the k NN imputation models, a forest stand polygon (average polygon size of 20.6 ha) was used as the unit of observation. The MODIS and environmental predictor variables of each stand polygon were calculated as the mean values of the raster cells (250-m spatial resolution) with more than 50% of the pixel area covered by the stand polygon. The 25,000 stand polygons containing both response variables (stand attributes) and predictor variables (MODIS and environmental variables) were split into training and testing datasets using a 7:3 split ratio, 70% of which were used to build the k NN imputation model to predict wall-to-wall stand attributes (Fig. 2); the remaining 30% of the stand polygons were used to evaluate the prediction accuracy.

2.4.3 Predicting wall-to-wall tree lists

Maps of species-level biomass for the 17 major species in the study area in 2000 were derived from our previous study, where the prediction accuracy of the species-level biomass (Additional file 1: Fig. S1) was provided (Fu et al. 2019). In this step, maps of species-level biomass were used to predict the tree lists in each pixel by combining the Weibull diameter distribution by species. Once the parameters of the Weibull diameter distribution by species for each pixel were predicted by WDPPM_S, the proportion of tree density at each

diameter class by species (PTDS) could be calculated for each pixel. Thereafter, the proportion of the species-level biomass at each diameter class (PSB) was calculated by combining PTDS and single-tree biomass by species at each diameter class (Dong et al. 2014; Dong et al. 2015). Finally, the wall-to-wall species-level biomass was divided into each diameter class according to the PSB and converted to the tree density at each diameter class (wall-to-wall tree lists) by dividing them by the single-tree biomass at each diameter class (Fig. 2).

2.5 Accuracy assessment

To guarantee the reliable accuracy of our coupled framework, each step of the framework was assessed using various accuracy metrics.

In step 1, the Kolmogorov-Smirnov (KS) test (Riemann et al. 2010) and the error index proposed by Packalén and Maltamo (2008) were applied to assess the goodness-of-fit of the two-parameter Weibull function for predicting the diameter distribution by species in each plot from the MLE. The KS test does not make any assumption about the distribution of data and quantifies the consistency between the distributions of the two data sets (e.g., observed and predicted samples) based on the maximum difference (D) in their empirical distribution functions [e.g., $F(x)$ and $G(x)$], which is defined as follows:

$$D = \max |F(x) - G(x)| \quad (6)$$

The larger the P value of the KS test, the higher the prediction accuracy of the DBH Weibull distribution parameters of each tree species obtained by MLE. For each tree species, the plots that passed the KS test were selected as training data to establish the relationship between the stand mean DBH and the two parameters of the Weibull distribution and further calculate the error index of the DBH distribution of each species in each plot. The error index for each plot was calculated as follows:

$$e = \sum_{i=1}^k 0.5 \left| \frac{f_i}{N} - \frac{\hat{f}_i}{\hat{N}} \right| \quad (7)$$

where f_i and \hat{f}_i are the observed and predicted stem numbers of class i by species, respectively, k is the number of diameter classes, and N and \hat{N} are the observed and predicted stem numbers of all diameter classes by species, respectively. The class interval was 2 cm. The value of the error index ranges from 0 to 1, with 0 indicating a perfect fit and 1 indicating no overlap of the distributions (Packalén and Maltamo 2008).

The WDPPM_S was evaluated using R^2 , the root mean square error (RMSE), and bias. The three metrics were

calculated based on the parameters of the Weibull diameter distribution function using leave-one-out cross-validation (LOOCV):

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y}_i)^2} \quad (8)$$

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n}} \quad (9)$$

$$\text{Bias} = \frac{\sum_{i=1}^n (y_i - \hat{y}_i)}{n} \quad (10)$$

where n is the number of samples, y_i is the observed value for sample i , \hat{y}_i is the predicted value for sample i , and \bar{y}_i is the mean value of all observed values. To assess the WDPPM_S accuracy, n is the number of plots used to fit WDPPM_S for each tree species, y_i is the predicted shape (b) or scale parameter (c) value of plot i based on MLE, and \hat{y}_i is the predicted shape (b) or scale parameter (c) value of plot i based on WDPPM_S.

In step 2, the accuracy of the predicted wall-to-wall stand mean DBH over the forest region of Northeast China was evaluated at two scales (stand and ecotype). The study area was classified into 239 ecotypes based on seven subregions, nine landforms (Additional file 1: Fig. S2), and five forest types (Additional file 1: Fig. S3). The seven subregions reflect differences in temperature and moisture conditions. Nine landforms (ridge, upper slope, sunny slope, semi-sunny slope, semi-shady slope, shady slope, flat slope, lower slope, and valley) were classified from DEM using a Topographic Position Index (Zhang et al. 2009). Five forest types (evergreen coniferous, evergreen broad-leaved, deciduous coniferous, deciduous broad-leaved, and mixed forest) were extracted from the resampled MODIS land cover product MCD12Q1 (250 m) in 2000.

At the stand scale, the predicted stand mean DBH of the remaining 30% of stand polygons was calculated by averaging the stand mean DBH values of all pixels with over 50% of their area located in each forest stand polygon. Thereafter, both the observed and predicted stand mean DBH of the testing dataset were averaged to 65 ecotypes at the ecotype scale. The accuracy of the predicted wall-to-wall stand mean DBH over the forest region of Northeast China at stand and ecotype scales was evaluated by calculating the R^2 , RMSE, bias, empirical cumulative distribution functions (ECDFs), and KS statistic (Eq. (6)) between the predicted and observed stand mean DBH. The R^2 , RMSE, and bias provide an overall assessment of the prediction accuracy. The ECDFs and KS statistics can

quantify the discrepancy in the distributions of the predicted and observed stand mean DBH.

In step 3, to assess the accuracy of the predicted wall-to-wall tree lists in 2000, the tree lists of the 641 sample plots in 2000 were first obtained as the validation data by transforming the tree lists of 641 sample plots from 2009 to 2013 by subtracting the DBH growth of the 17 species in the past 9–13 years using the age-DBH relationships summarized in previous studies (Xu et al. 2020; Zhang et al. 2018b). The central points of the 641 sample plots were used to extract the predicted wall-to-wall tree lists at the corresponding pixel locations. Thereafter, the observed and predicted tree lists of the 641 sample plots in 2000 included the tree density (trees ha⁻¹) of each diameter class by species was used to calculate the error index (Eq. (7)), which was used to evaluate the prediction accuracy of the wall-to-wall tree lists over Northeast China. We separately calculated the error indices by species for each plot and for all plots as a whole.

3 Results

3.1 Weibull distribution parameter prediction models (WDPPM_S)

Based on the KS test, the diameter distributions of 17 major tree species in the forest region of Northeast China could be well fitted by the two-parameter Weibull functions predicted by the MLE method, especially for the white birch (*B. platyphylla* Suk.), basswood (*T. amurensis* Rupr.), black birch (*B. davurica* Pall.), mono maple (*A. mono* Maxim.), elm (*U. pumila* L.), walnut (*J. mandshurica* Maxim.), Korean pine (*P. koraiensis* Sieb. et Zucc.), ribbed birch (*B. costata* Trautv.), and fir (*A. nephrolepis* (Trautv.) Maxim.) (Table 1; agreement > 90%).

The calculated error indices revealed that the diameter distributions by species in each plot were well-fitted, with the mean error index value ranging from 0.24 to 0.48 and standard deviation of the error index ranging from 0.09 to 0.2. When the number of trees in each diameter class of all plots as a whole by species is summed to calculate the error index, the diameter distributions of all species were well represented with a lower value of error index (0.1–0.38) (Table 2).

The accuracy assessment of WDPPMs by LOOCV demonstrated that the scale parameter had a strong correlation with the arithmetic mean DBH for all 17 tree species, with high R^2 (close to 1) and low RMSE (0.06–0.52) (Table 3). Although the relationship between the shape parameter and arithmetic mean DBH of each tree species was relatively weak (low R^2 and large RMSE), the shape parameter prediction models for all species passed the significance test ($P < 0.05$), with bias = 0 (Table 3).

Table 1 Kolmogorov-Smirnov (KS) test ($P > 0.05$) of the DBH distribution based on the two-parameter Weibull function on 17 tree species in the forest region of Northeast China

Species	Number of samples passing the KS test	Total number of samples	Agreement (%)
White birch (<i>B. platyphylla</i> Suk.)	306	329	93.01
Larch (<i>L. gmelinii</i> (Rupr.) Kuzen)	270	319	84.64
Mongolian oak (<i>Q. mongolica</i> Fisch. ex Ledeb.)	91	112	81.25
Aspen (<i>P. davidiana</i> Dode)	94	106	88.68
Basswood (<i>T. amurensis</i> Rupr.)	160	167	95.81
Black birch (<i>B. davurica</i> Pall.)	30	32	93.75
Scotch pine (<i>P. sylvestris</i> var. <i>mongolica</i> Litv.)	61	73	83.56
Mono maple (<i>A. mono</i> Maxim.)	112	116	96.55
Elm (<i>U. pumila</i> L.)	97	102	95.10
Walnut (<i>J. mandshurica</i> Maxim.)	52	55	94.55
Spruce (<i>P. koraiensis</i> Nakai)	95	112	84.82
Korean pine (<i>P. koraiensis</i> Sieb.et Zucc.)	103	110	93.64
Ash (<i>F. mandschurica</i> Rupr.)	54	60	90.00
Ribbed birch (<i>B. costata</i> Trautv.)	81	85	95.29
Fir (<i>A. nephrolepis</i> (Trautv.) Maxim.)	112	115	97.39
Amur corktree (<i>P. amurense</i> Rupr.)	60	68	88.24
Willow (<i>C. arbutifolia</i> (Pall.) A. Skv.)	25	30	83.33

Agreement was defined as the percentage of samples that passed the KS test ($P > 0.05$)

Table 2 Error index of the DBH Weibull distribution for 17 tree species in all plots

Species	Mean value of error index	Minimum value of error index	Maximum value of error index	Standard deviation of error index	Error index of summing all plots
White birch (<i>B. platyphylla</i> Suk.)	0.31	0.11	0.75	0.14	0.19
Larch (<i>L. gmelinii</i> (Rupr.) Kuzen)	0.34	0.07	0.80	0.17	0.20
Mongolian oak (<i>Q. mongolica</i> Fisch. ex Ledeb.)	0.34	0.15	0.52	0.09	0.14
Aspen (<i>P. davidiana</i> Dode)	0.24	0.08	0.67	0.12	0.10
Basswood (<i>T. amurensis</i> Rupr.)	0.37	0.13	0.86	0.17	0.20
Black birch (<i>B. davurica</i> Pall.)	0.42	0.14	0.90	0.18	0.13
Scotch pine (<i>P. sylvestris</i> var. <i>mongolica</i> Litv.)	0.32	0.13	0.86	0.14	0.24
Mono maple (<i>A. mono</i> Maxim.)	0.46	0.17	0.92	0.12	0.38
Elm (<i>U. pumila</i> L.)	0.42	0.12	0.78	0.12	0.19
Walnut (<i>J. mandshurica</i> Maxim.)	0.45	0.11	0.89	0.17	0.19
Spruce (<i>P. koraiensis</i> Nakai)	0.30	0.13	0.60	0.10	0.23
Korean pine (<i>P. koraiensis</i> Sieb.et Zucc.)	0.36	0.12	0.92	0.18	0.16
Ash (<i>F. mandschurica</i> Rupr.)	0.43	0.17	0.72	0.13	0.27
Ribbed birch (<i>B. costata</i> Trautv.)	0.43	0.15	0.83	0.20	0.20
Fir (<i>A. nephrolepis</i> (Trautv.) Maxim.)	0.43	0.14	0.86	0.17	0.18
Amur corktree (<i>P. amurense</i> Rupr.)	0.48	0.19	0.71	0.13	0.28
Willow (<i>C. arbutifolia</i> (Pall.) A. Skv.)	0.44	0.22	0.68	0.12	0.32

Error index of summing all plots is calculated by combining all plots as a whole

3.2 Map of stand mean DBH

The average value of the stand mean DBH over the forest region of Northeast China was 13.22 cm. The predicted stand mean DBH of the severely burned

area in 1987 was the lowest, with an average value of 8.86 cm (Fig. 3). Subregion CBM, especially the CBM Nature Reserve, had the highest prediction of stand mean DBH. In fact, the average stand mean DBH of

Table 3 Equations of shape parameter (b) and scale parameter (c) of Weibull distribution based on arithmetic mean DBH (d , unit: cm) for 17 tree species and accuracy assessment using LOOCV

Species	Equations	R^2	RMSE	Bias	P value
White birch (<i>B. platyphylla</i> Suk.)	$b = -0.182d + 5.078$ $c = 1.154d - 0.347$	0.26 1.00	0.76 0.16	0.00 0.00	0.00 0.00
Larch (<i>L. gmelinii</i> (Rupr.) Kuzen)	$b = 0.168d + 1.623$ $c = 1.068d - 0.509$	0.23 1.00	0.81 0.20	0.00 0.00	0.00 0.00
Mongolian oak (<i>Q. mongolica</i> Fisch. ex Ledeb.)	$b = 0.068d + 1.245$ $c = 1.141d - 0.212$	0.18 1.00	1.11 0.52	0.00 0.00	0.05 0.00
Aspen (<i>P. davidiana</i> Dode)	$b = -1.169d + 4.960$ $c = 1.154d - 0.359$	0.34 1.00	0.49 0.06	0.00 0.00	0.00 0.00
Basswood (<i>T. amurensis</i> Rupr.)	$b = 0.101d + 1.230$ $c = 1.091d + 0.445$	0.12 1.00	1.14 0.27	0.00 0.00	0.00 0.00
Black birch (<i>B. davurica</i> Pall.)	$b = 0.281d + 0.231$ $c = 1.039d + 0.896$	0.24 0.99	1.56 0.33	0.00 0.00	0.01 0.00
Scotch pine (<i>P. sylvestris</i> var. <i>mongolica</i> Litv.)	$b = 0.228d + 0.089$ $c = 1.081d + 0.513$	0.21 0.99	0.97 0.19	0.00 0.00	0.01 0.00
Mono maple (<i>A. mono</i> Maxim.)	$b = -0.236d + 8.473$ $c = 1.142d - 0.659$	0.18 0.99	1.83 0.35	0.00 0.00	0.00 0.00
Elm (<i>U. pumila</i> L.)	$b = 0.131d + 0.882$ $c = 1.082d + 0.614$	0.19 1.00	0.92 0.22	0.00 0.00	0.00 0.00
Walnut (<i>J. mandshurica</i> Maxim.)	$b = 0.163d + 0.732$ $c = 1.064d + 0.816$	0.28 1.00	1.45 0.41	0.00 0.00	0.00 0.00
Spruce (<i>P. koraiensis</i> Nakai)	$b = -0.210d + 4.703$ $c = 1.165d - 0.357$	0.29 1.00	0.57 0.08	0.00 0.00	0.00 0.00
Korean pine (<i>P. koraiensis</i> Sieb. et Zucc.)	$b = 0.122d + 1.578$ $c = 1.068d + 0.628$	0.33 1.00	1.53 0.46	0.00 0.00	0.00 0.00
Ash (<i>F. mandshurica</i> Rupr.)	$b = -0.108d + 3.489$ $c = 1.122d + 0.067$	0.16 1.00	0.59 0.13	0.00 0.00	0.00 0.00
Ribbed birch (<i>B. costata</i> Trautv.)	$b = 0.251d - 0.135$ $c = 1.038d + 1.061$	0.49 1.00	1.66 0.39	0.00 0.00	0.00 0.00
Fir (<i>A. nephrolepis</i> (Trautv.) Maxim.)	$b = 0.097d + 1.515$ $c = 1.087d + 0.460$	0.13 1.00	1.34 0.33	0.00 0.00	0.00 0.00
Amur corktree (<i>P. amurense</i> Rupr.)	$b = -0.170d + 4.351$ $c = 1.134d - 0.063$	0.16 1.00	0.80 0.12	0.00 0.00	0.00 0.00
Willow (<i>C. arbutifolia</i> (Pall.) A. Skv.)	$b = 0.156d + 1.0219$ $c = 1.095d + 0.329$	0.16 0.99	0.77 0.17	0.00 0.00	0.05 0.00

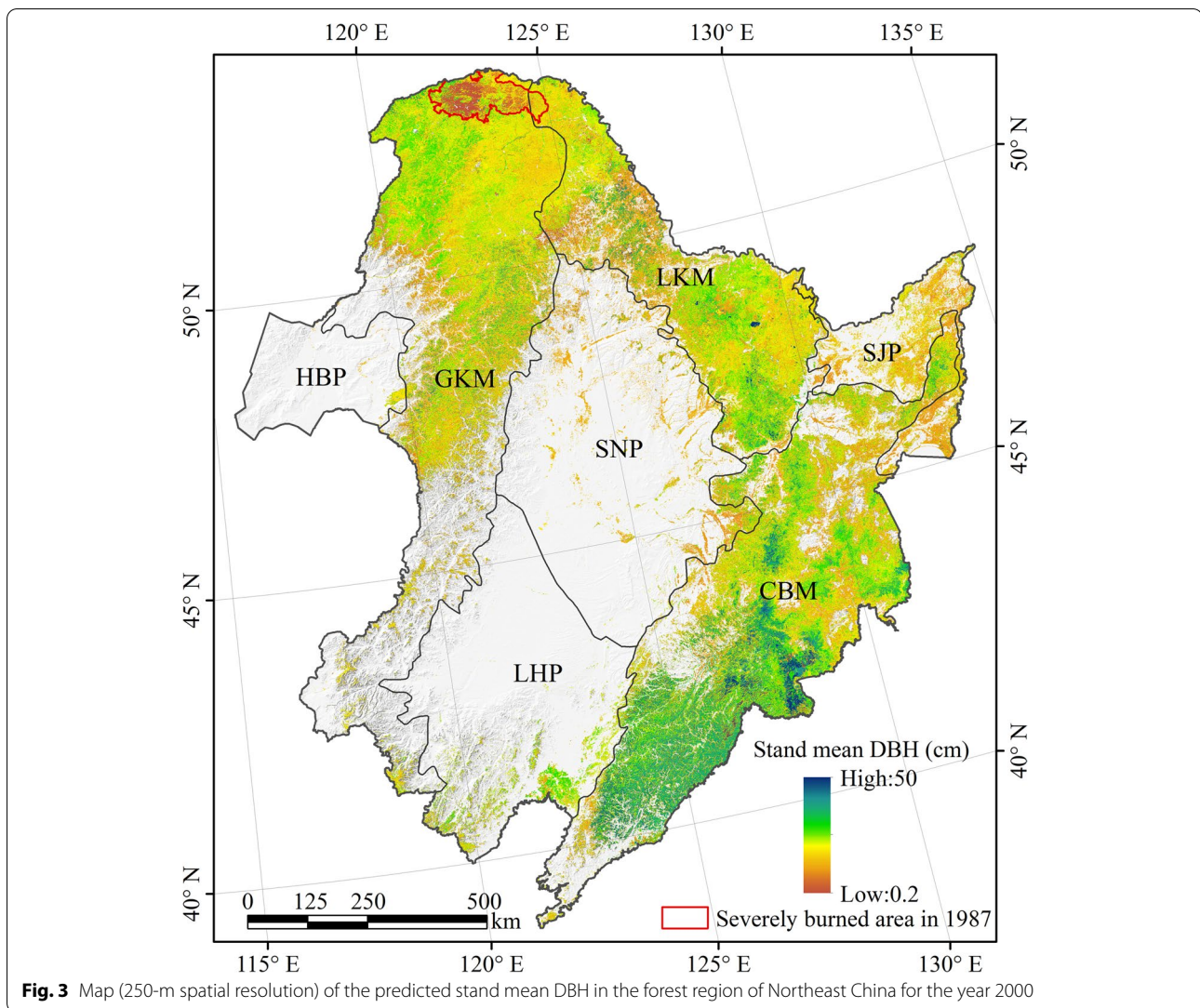
subregion CBM was 15.20 cm. The predicted stand mean DBH of the other two major forest regions (subregion LKM and GKM) was lower than that of subregion CBM, with an average value of 12.67 cm and 12.25 cm, respectively.

The accuracy of the predicted stand mean DBH improved substantially from the stand scale to the ecotype scale, the R^2 increased from 0.56 to 0.88, the RMSE decreased from 3.33 to 1.08 cm, and the bias decreased from 0.13 cm to 0.01 cm (Fig. 4a, c). Although the KS distance ($D = 0.09$) was the same at the ecotype and the stand scale, a higher P value (0.94 vs. 0) for KS distance at the ecotype scale revealed that the predicted

and observed stand mean DBH ECDFs became more similar (Fig. 4b, d).

3.3 Prediction of the tree lists

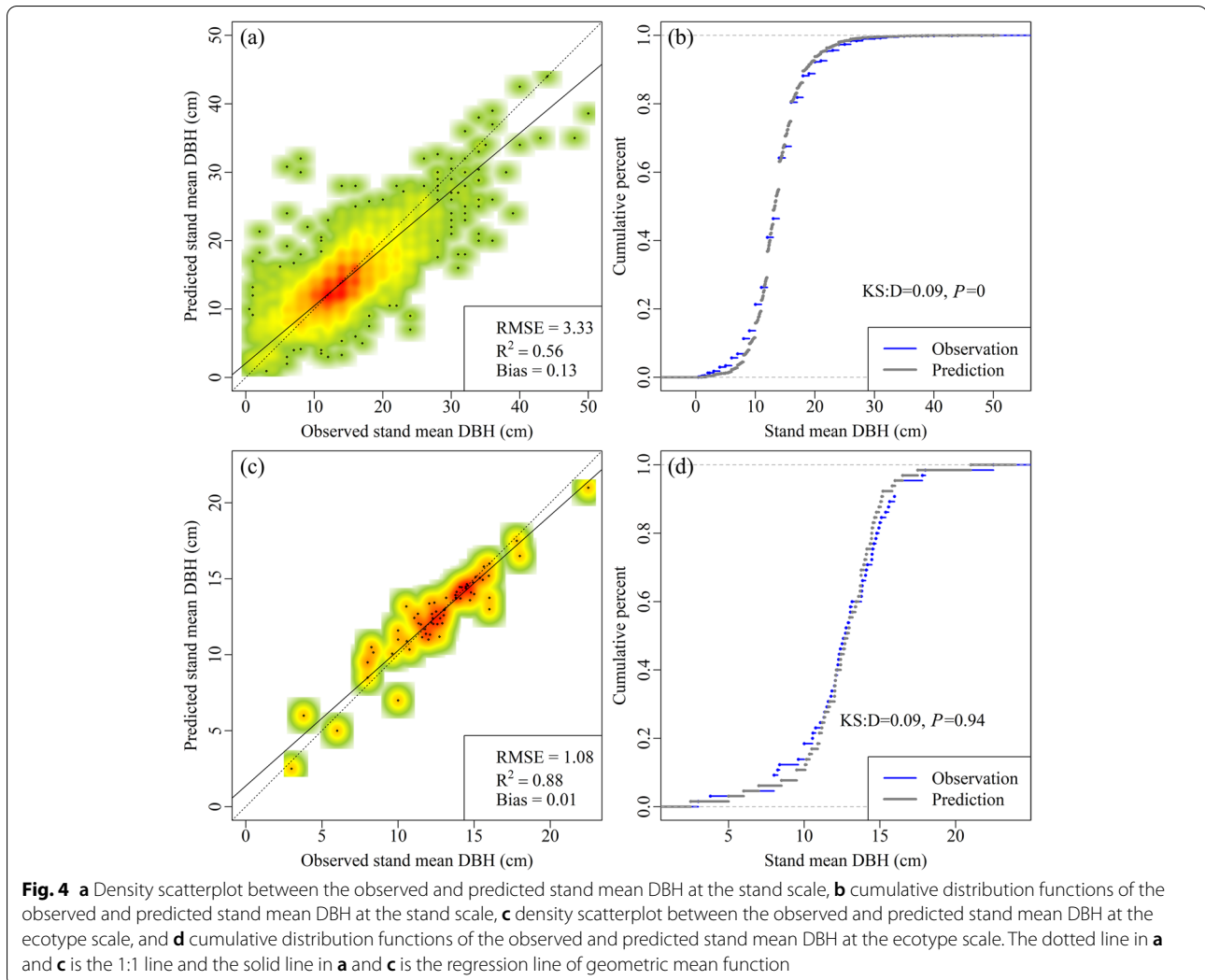
The predicted wall-to-wall tree lists revealed the DBH distributions of the 17 major tree species in the forest region of Northeast China (Fig. 5). The predicted DBH of most tree species was concentrated in the range of 6–38 cm, except for spruce (*P. koraiensis* Nakai), ash (*F. mandshurica* Rupr.), and Amur corktree (*P. amurense* Rupr.), which was distributed at a DBH greater than 40 cm. The DBH distributions of aspen (*P. davidiana* Dode), Scotch pine (*P. sylvestris* var. *mongolica* Litv.), spruce (*P. koraiensis* Nakai), ash (*F.*



mandschurica Rupr.), Amur corktree (*P. amurense* Rupr.), and willow (*C. arbutifolia* (Pall.) generally showed a reverse J-shape, and the density of the above tree species was the highest at a DBH of 6 cm. The DBH of white birch (*B. platyphylla* Suk.), larch (*L. gmelinii* (Rupr.) Kuzen), Mongolian oak (*Q. mongolica* Fisch. ex Ledeb.), basswood (*T. amurensis* Rupr.), black birch (*B. davurica* Pall.), mono maple (*A. mono* Maxim.), elm (*U. pumila* L.), walnut (*J. mandshurica* Maxim.), Korean pine (*P. koraiensis* Sieb. et Zucc.), ribbed birch (*B. costata* Trautv.), and fir (*A. nephrolepis* (Trautv.) Maxim.) displayed a right-skewed unimodal distribution, and the tree density of these tree species reached its peak at a DBH of 12, 14, or 16 cm.

The predicted wall-to-wall tree lists in 2000 had a comparable accuracy to the predicted tree lists of 641 plots from 2009 to 2013 (Table 2, Figs. 6 and 7), with a mean value for the error index of DBH distribution of

each species between 0.24 and 0.58 (Fig. 6). When all plots were combined as a whole, the error index of the DBH distribution of each tree species was lower than the mean value of the error index calculated based on all plots, indicating improved accuracy of the predicted tree lists for most species (Fig. 7). Except for aspen (*P. davidiana* Dode), walnut (*J. mandshurica* Maxim.), and Amur corktree (*P. amurense* Rupr.), the predicted density of the other tree species in the smallest DBH class (6 cm) was lower than the observed density. Notably, the predicted and observed DBH distribution curves for black birch (*B. davurica* Pall.) was the most consistent, showing a left-skewed unimodal distribution, and the error index was the smallest (0.13). Therefore, the prediction accuracy of the DBH distribution of black birch (*B. davurica* Pall.) was deemed to be the highest (Fig. 7).



3.4 Maps of species-specific tree density from tree lists

For the maps of species-specific tree density in the forest region of Northeast China in 2000 (Fig. 8), pixels with a tree density greater than 0 indicated the location of tree species growth. The total tree density of the 17 species mapped by summing the species-level tree density within each pixel mainly ranged from 700 trees ha⁻¹ to 3000 trees ha⁻¹, and the average total tree density for the forest region of Northeast China was approximately 1462 trees ha⁻¹ (Fig. 8). Among the three major forest distribution areas of Northeast China (subregions GKM, LKM, and CBM), the average total tree density in subregion LKM (1510 trees ha⁻¹) was higher than that in subregions GKM (1506 trees ha⁻¹) and CM (1393 trees ha⁻¹). In the northeastern part of ecoregion SJP, owing to the large number of newly planted Scotch pine (*P. sylvestris* var. *mongolica* Litv.) forests, the total tree density in this region was high (Fig. 8).

By counting the mean value of pixels with tree density greater than 0 for each tree species, the average tree density of larch (*L. gmelinii* (Rupr.) Kuzen) (1009 trees ha⁻¹) in its existing area was found to be the highest. In terms of spatial distribution, the tree density of larch (*L. gmelinii* (Rupr.) Kuzen) in each pixel was significantly higher than that of the other tree species (Fig. 8). White birch (*B. platyphylla* Suk.) had an opposing trend to larch (*L. gmelinii* (Rupr.) Kuzen) in the spatial distribution of tree density, especially in the severely burned area of the GKM subregion. Although the distribution range of willow (*C. arbutifolia* (Pall.) A. Skv.) was the smallest among the 17 tree species, with a high average tree density (766 trees ha⁻¹) in its existing area, second only to that of larch (*L. gmelinii* (Rupr.) Kuzen), indicating that the distribution of willow (*C. arbutifolia* (Pall.) A. Skv.) was relatively dense. The tree density of Scotch pine (*P. sylvestris* var.

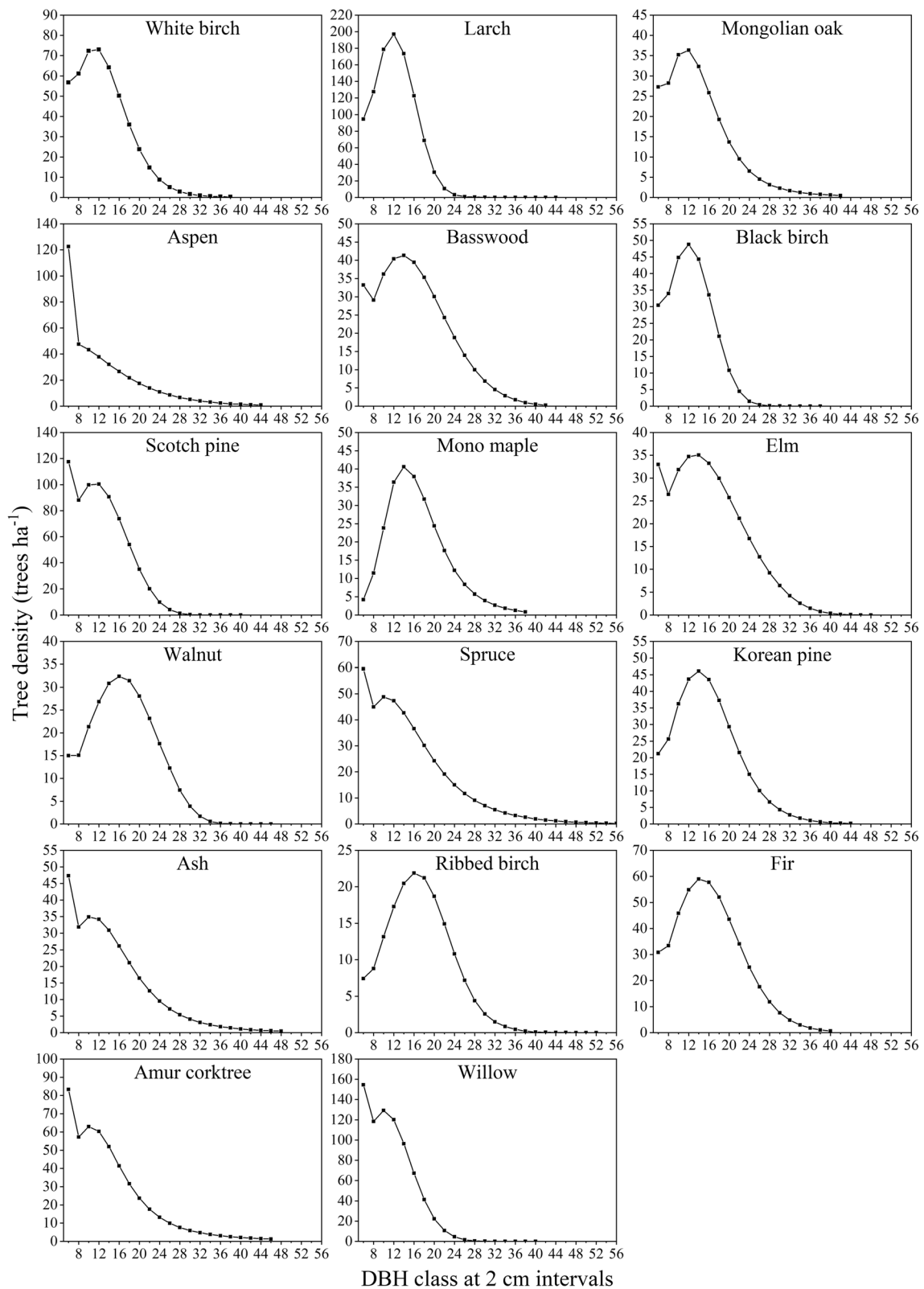


Fig. 5 Predicted tree lists by species over the forest region of Northeast China in 2000

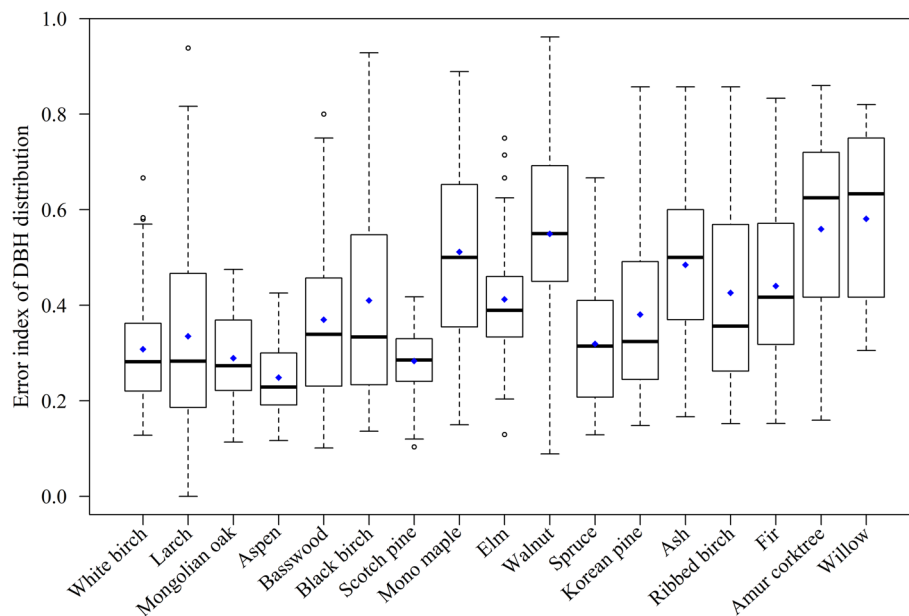


Fig. 6 Error index calculated based on the predicted wall-to-wall and transformed inventory tree lists in 2000 by species for each plot. The blue diamond represents the mean value of the error index calculated based on all plots

mongolica Litv.) in the eastern part of the LKM subregion and the northeastern part of the SJP subregion was significantly higher than that in the other regions. The tree density of spruce (*P. koraiensis* Nakai) and fir (*A. nephrolepis* (Trautv.) Maxim.) in subregion LKM was higher than that in subregion CM. For other tree species, the predicted wall-to-wall tree density for most pixels was less than 700 trees ha⁻¹. For the Mongolian oak (*Q. mongolica* Fisch. ex Ledeb.), black birch (*B. davurica* Pall.), mono maple (*A. mono* Maxim.), walnut (*J. mandshurica* Maxim.), ash (*F. mandshurica* Rupr.), and ribbed birch (*B. costata* Trautv.), the predicted wall-to-wall tree density for most pixels was less than 300 trees ha⁻¹.

4 Discussion

4.1 Selection of the coupled framework for predicting wall-to-wall tree lists

With the increasing need for detailed forest attributes (e.g., tree lists) to answer regional questions, including understanding the current structure and succession status of a forest, formulating targeted forest management strategies for different forest regions, or predicting tree species distribution shifts under climate change (Ferraz et al. 2020; Wang et al. 2019), the development of efficient and accurate methods to obtain such regional tree list information has attracted significant research interest. Although previous studies have applied both parametric diameter distribution modeling (e.g., Weibull function) and nonparametric imputation (e.g., *k*NN

model) to predict tree lists (Packalén and Maltamo 2008; Schütz and Rosset 2020; Shang et al. 2017), the integration of these two methods proposed in our study is novel for generating wall-to-wall tree lists over large areas with limited forest inventory data.

The Weibull function, especially the two-parameter Weibull function, is extensively used to describe the diameter distributions of single- and mixed-species stands (Hao et al. 2022; Liu et al. 2014; Schmidt et al. 2020; Zhang et al. 2018b). Although Liu et al. (2014) found that a finite mixture model (FMM) of Weibull functions was more flexible for describing irregular and highly skewed diameter distributions for the whole plot in the mixed-species forest stands of the Greater Khingan Mountains in Northeast China than a single Weibull function to fit the whole plot only, a single Weibull function to fit each species component separately was more accurate than the FMM. FMM is often difficult to apply to large regions owing to the many unknown parameters that require good fitting (Zasada and Cieszewski 2005).

A prerequisite for applying a single Weibull function to fit each species component separately to large regions is the available information on wall-to-wall species composition, which is difficult to obtain through traditional forest resource inventory alone (Liu et al. 2014). Our coupled framework can use a Weibull function to fit each species component separately and provide information regarding the species composition across large regions using *k*NN imputation models based on MODIS data.

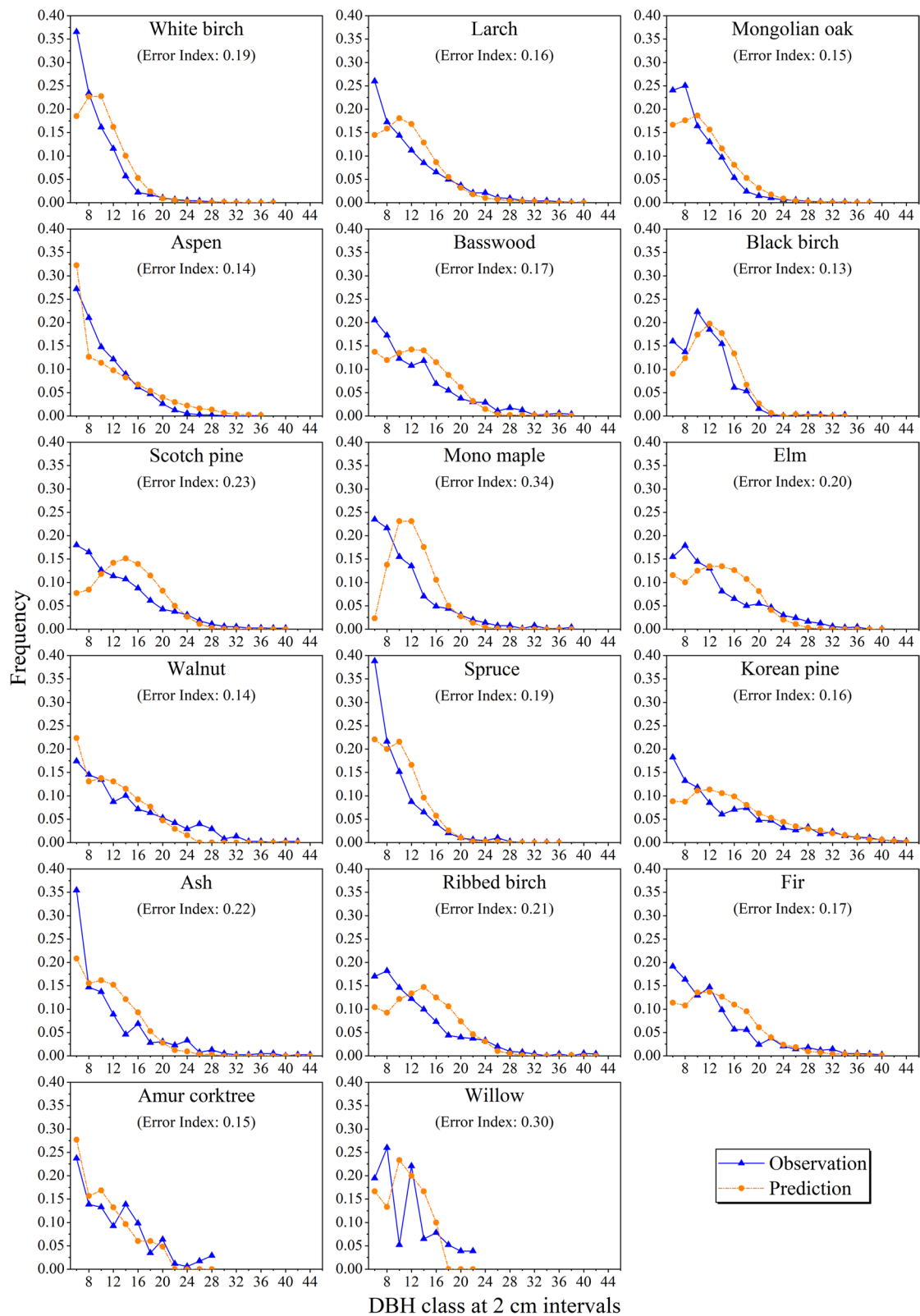
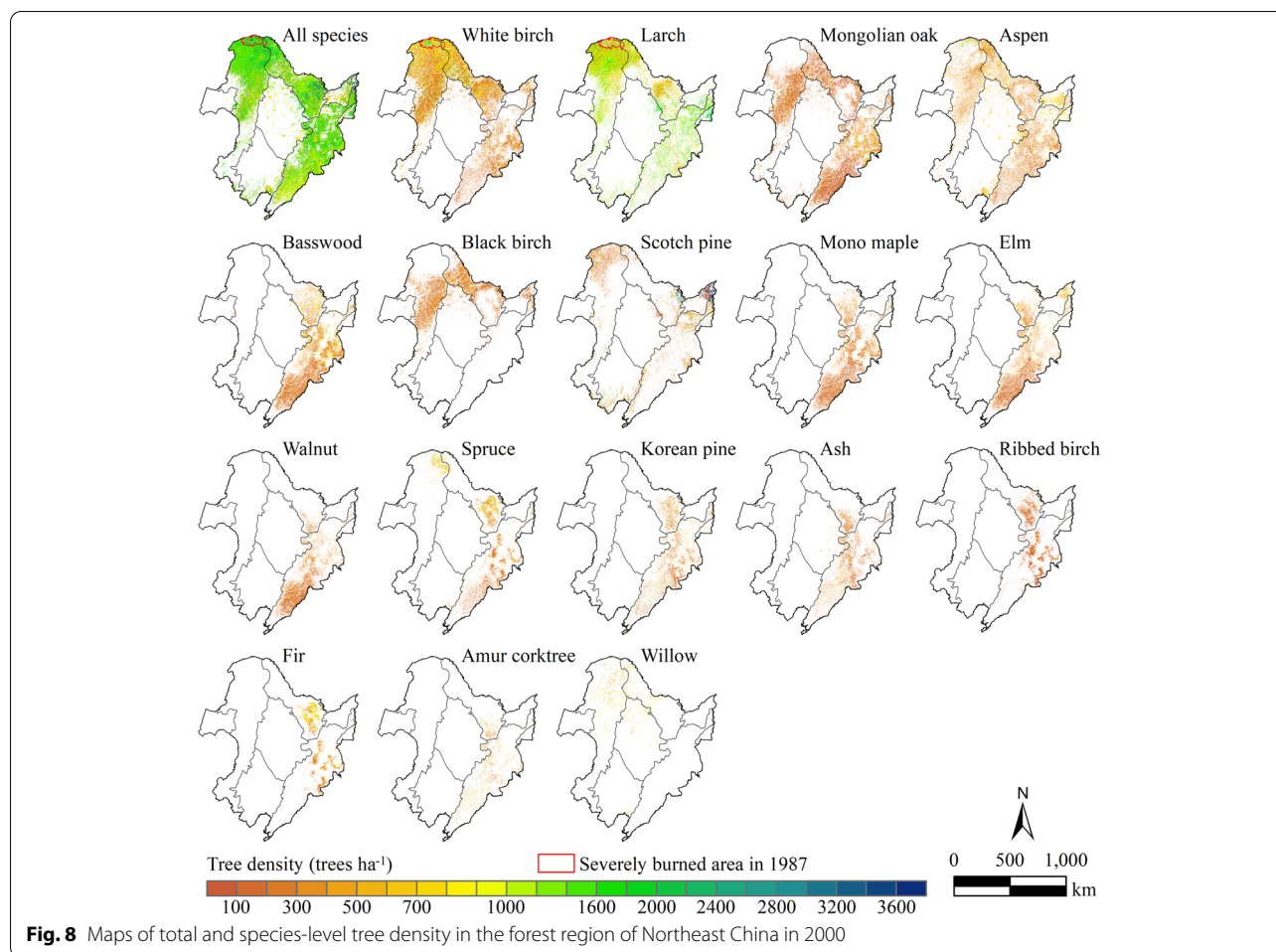


Fig. 7 DBH distribution frequency by species of predicted wall-to-wall tree lists versus observed tree lists in 2000 based on the transformed sample plot data. The error index was calculated by combining all transformed plots as a whole



MODIS data have broader spatial coverage and higher temporal resolution than other high spatial resolution multispectral satellite imagery (e.g., Landsat), provide abundant and near real-time spectral information, and have been widely used to map spatially explicit forest attributes over large regions (Beaudoin et al. 2018; Hayes et al. 2008; Zhang et al. 2018a; Zhang et al. 2013). Therefore, our coupled framework is feasible and effective for predicting species-specific diameter distributions and tree density on a regional scale.

4.2 Accuracy and reliability of the coupled framework's predictions

Accuracy is an important issue in regional forest attributes predictions. The accuracy of the wall-to-wall tree lists predicted in this study was mainly determined by the WDPPM_S and *k*NN imputation models. For a more general assessment of our coupled framework, we compared the accuracies of the WDPPM_S and *k*NN models with those of similar studies. The WDPPM_S used in our study produced an accuracy comparable to that

of previous studies. For example, the scale parameter was usually strongly correlated with the stand attributes (Zhang et al. 2018b), consistent with previous studies, and the prediction accuracy of the scale parameter of the 17 tree species in this study was better than the results ($R^2 = 0.96\text{--}0.99$) obtained in the natural stands of the Crimean Juniper in the southern and southwestern (Mediterranean) region of Turkey (Diamantopoulou et al. 2015). Furthermore, our studies led to a conclusion similar to that previously reported by Hao et al. (2022), Wang et al. (2006), and Zhang et al. (2018b), who found that the shape parameter is often more difficult to accurately predict than the scale parameter. This finding might be due to the remarkable complexity of the shape parameter in the mixture distribution to allow characterization by the stand mean DBH (Wang et al. 2006). The average error index (0.32–0.34) of the predicted diameter distribution by tree species in a typical managed boreal forest area of Finland (Packalén and Maltamo 2008) was close to that obtained herein for the dominant tree species

in Northeast China. The above comparisons indicate that the WDPPM_S developed in this study can be used at least as accurately as previous studies to predict species-specific diameter distributions.

The accuracy assessment at different scales for our predicted stand attributes (e.g., species-level biomass and stand mean DBH) could provide useful information for different applications. For example, stand projection models (growth and yield models) require detailed stand attributes at the stand scale to set the initial conditions (Scolforo et al. 2019), whereas forest ecosystem process research may require forest attributes at the ecotype scale (Dijak et al. 2017). In this study, the stand mean DBH predicted using the *k*NN imputation model over Northeast China at the stand scale had a higher accuracy than that of similar studies conducted in Chinese boreal forests (Zhang et al. 2018b). Owing to the relatively homogeneous temperature, humidity, topography, forest type, and species composition of each ecotype, the prediction accuracy of stand mean DBH at the ecotype scale was significantly improved, indicating that the map of stand mean DBH generated in this study is more suitable for application to forest ecosystem models (e.g., LINKAGES) operated at the ecotype scale. The areas with lower predictions of stand mean DBH in our study were consistent with the burned areas in 1987, further indicating that our predicted stand attributes were realistic and captured the significant disturbance processes. The predicted mean values of total and species-level tree density derived from the tree lists over the Greater Khingan Mountains, Lesser Khingan Mountains, and Changbai Mountains were similar to those of previous studies conducted in Northeast China (Chen et al. 2003; Dai et al. 2013; Liu et al. 2020), suggesting that our predicted tree lists were reasonable.

4.3 Ecological implications of the prediction results

The DBH distribution of tree species exhibits variable patterns owing to site quality, regional climate, and disturbances (Chhin et al. 2008; Lin et al. 2016; Rodrigo et al. 2022). Generally, consistent with the results of previous studies fitting the similar DBH distribution of the corresponding tree species in Northeast China (Hao et al. 2022; Liu et al. 2014; Xu and Jin 2012), the overall DBH distribution of most tree species predicted in this study displayed a right-skewed unimodal pattern (Fig. 5), suggesting that at the regional scale, a continuous regeneration and relatively stable community structure occurred in most tree species in Northeast China (Fang et al. 2012). Further, the regional DBH distribution of a few tree species, such as aspen (*P. davidiana* Dode), Scotch pine (*P. sylvestris* var. *mongolica* Litv.), spruce (*P. koraiensis* Nakai), ash (*F. manschurica*

Rupr.), Amur corktree (*P. amurense* Rupr.), and willow (*C. arbutifolia* (Pall.)), which were fitted in this study, followed a reverse J-shape (Fig. 5), which is a general pattern observed in virgin or near-natural forests. Such findings indicate that the stands of the above tree species were mature and reached an equilibrium state in Northeast China (Kral et al. 2010; Westphal et al. 2006). Our predicted species-specific diameter distributions could provide a reference for species conservation and strategic forest management in the entire distribution range of tree species.

Our spatial predictions of tree density by species extracted from tree lists can capture the dominant ecological patterns and processes. In particular, for the high-intensity burned areas of the Greater Khingan Mountains in 1987, the predicted tree density of white birch (*B. platyphylla* Suk.) was higher than that of larch (*L. gmelinii* (Rupr.) Kuzen) in 2000, which was lower than that of larch (*L. gmelinii* (Rupr.) Kuzen) in low-intensity burned areas (Fig. 8). This spatial pattern was attributed to the strong seed dispersal and establishment abilities of white birch (*L. gmelinii* (Rupr.) Kuzen) and the strong fire resistance of larch (*L. gmelinii* (Rupr.) Kuzen) (Chang et al. 2007; Wang et al. 2017). Therefore, the high-resolution maps of total and species-level tree density for the entire Northeast China generated in our study could reveal the tree species composition and regeneration patterns of forest regions affected by large disturbance events (Crowther et al. 2015). These maps would also be useful for modeling broad-scale biological and biogeochemical processes (Slik et al. 2010).

4.4 Limitations of the current study and future directions

In this study, we proposed a coupled framework for using MODIS data as predictor variables to impute wall-to-wall stand attributes and predict species-specific diameter distributions using the unimodal Weibull model for continuous forest regions. This framework provides an innovative means of generating detailed wall-to-wall tree lists across large areas based on limited field data. However, the coupled framework of the current study has limitations, and some potential improvements must be addressed in future work. For instance, the limited and unevenly distributed forest inventory data over the study area may introduce error sources to forest stand attribute predictions. Advances in LiDAR techniques have demonstrated great potential for providing accurate estimates of tree metrics, such as tree height (Su et al. 2016). Moreover, as LiDAR footprints are distributed evenly and widely (Popescu et al. 2011), the derived tree-level variables from these footprints can be input to the *k*NN imputation models to compensate for the lack of field data. Thus, future work

should consider the synergy of LiDAR and MODIS data to develop links between inherent forest structural attributes and their spectral information to enhance the reliability of predicting wall-to-wall stand attributes.

Owing to the lack of accurate conversions between stand mean DBH and mean arithmetic DBH of major tree species in Northeast China, the stand mean DBH was used instead of mean arithmetic DBH by species to predict pixel-level Weibull distribution parameters. This replacement will lead to great uncertainties when applied to forests with complex stand structures, especially for uneven-aged natural mixed forests. Some studies revealed that a unimodal pattern may not adequately describe the diameter distribution modality of irregular, uneven-aged stands (Hao et al. 2022; Thomas et al. 2008; Wang et al. 2006). Therefore, in future work, on the premise of collecting sufficient forest ground inventory and LiDAR data, the accurate conversion relationships between stand mean DBH and mean arithmetic DBH of each species must first be established. Thereafter, an appropriate modeling approach must be applied to fit the multimodal diameter distributions of mixed species or uneven-aged forest stands at the regional scale.

5 Conclusion

This study revealed the appropriateness of a novel framework using limited forest inventory data to predict wall-to-wall tree lists across large regions by integrating the two-parameter Weibull function and *k*NN imputation method. The Weibull function displayed a strong ability to describe the diameter distribution of tree species in Northeast China. The Weibull distribution parameter prediction models (WDPPM_S) of 17 species for predicting the parameters of Weibull functions, and the random forest-based *k*NN model for imputing stand mean DBH built produced considerable or higher accuracy than those of previous studies. Generally, the fitted DBH distribution of 17 tree species across Northeast China in 2000 displayed a right-skewed unimodal or reverse J-shaped pattern, suggesting that most tree species exist in a state of continuous regeneration and relatively stable structure at the regional scale. Additionally, the total and species-level tree density maps derived from the tree lists captured the forest regions influenced by fire disturbance and are prerequisites for ecological modeling and assessment of the entirety of Northeast China. However, this study presented regional tree lists across Northeast China for only 1 year (2000). To better analyze the natural succession of forests and assess the influence of disturbances and climate change on tree species composition, the temporal and spatial variation of tree lists over Northeast China should be further studied.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s13595-022-01161-9>.

Additional file 1: Table S1. Candidate predictor variables used in the *k*NN imputation models (Fu et al. 2019). The growing season was defined as May to October. **Fig. S1.** Maps of total AFB and 17 species-level AFB in Northeast China generated by Fu et al. (2019). AFB: Aboveground forest biomass. **Fig. S2** Nine landforms in Northeast China classified from the DEM using Topographic Position Index (Zhang et al. 2009). **Fig. S3** Five forest types in Northeast China extracted from the resampled MODIS land cover product MCD12Q1 (250 m) in 2000. R codes for generating wall-to-wall species-level biomass. R codes for generating wall-to-wall stand mean DBH. R codes for generating wall-to-wall tree lists.

Acknowledgements

We thank Qinglong Zhang from the Shandong University of Technology for the technical support.

Authors' contributions

Conceptualization: Yuanyuan Fu and Hong S. He; Methodology: Yuanyuan Fu; Data curation: Yuanyuan Fu; Writing—original draft preparation: Yuanyuan Fu; Writing—review and editing: Hong S. He, Shaoqiang Wang, and Lunche Wang; Supervision: Hong S. He. The authors read and approved the final manuscript.

Funding

This research was supported by the National Natural Science Foundation of China (Grant No. 32101327), and the Fundamental Research Funds for the Central Universities, China University of Geosciences (Wuhan) (Grant No. 111-G1323522043).

Availability of data and materials

The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

Declarations

Ethics approval and consent to participate

The authors declare that they follow the rules of good scientific practice.

Consent for publication

All authors gave their informed consent to this publication and its content.

Competing interests

The authors declare that they have no competing interests.

Author details

¹Key Laboratory of Regional Ecology and Environmental Change, School of Geography and Information Engineering, China University of Geosciences, Wuhan 430074, China. ²Key Laboratory of Geographical Processes and Ecological Security in Changbai Mountains, Ministry of Education, School of Geographical Sciences, Northeast Normal University, Changchun 130024, China. ³Hunan Key Laboratory of Remote Sensing of Ecological Environment in Dongting Lake Area, School of Geography and Information Engineering, China University of Geosciences, Wuhan 430074, China. ⁴School of Natural Resources, University of Missouri, Columbia, MO 65211, USA.

Received: 29 March 2022 Accepted: 19 September 2022

Published online: 04 October 2022

References

- Bankston JB, Sabatia CO, Poudel KP (2021) Effects of sample plot size and prediction models on diameter distribution recovery. *For Sci* 67(3):245–255. <https://doi.org/10.1093/forsci/xxaa055>
- Beaudoin A, Bernier P, Villemaire P, Guindon L, Guo XJ (2018) Tracking forest attributes across Canada between 2001 and 2011 using a *k* nearest

- neighbors mapping approach applied to MODIS imagery. *Can J For Res* 48(1):85–93. <https://doi.org/10.1139/cjfr-2017-0184>
- Breidenbach J, Næsset E, Lien V, Gobakken T, Solberg S (2010) Prediction of species specific forest inventory attributes using a nonparametric semi-individual tree crown approach based on fused airborne laser scanning and multispectral data. *Remote Sens Environ* 114(4):911–924. <https://doi.org/10.1016/j.rse.2009.12.004>
- Brosföske KD, Froese RE, Falkowski MJ, Banskota A (2014) A review of methods for mapping and prediction of inventory attributes for operational forest management. *For Sci* 60(4):733–756. <https://doi.org/10.5849/forsci.12-134>
- Chang Y, He HS, Bishop I, Hu Y, Bu R, Xu C, Li X (2007) Long-term forest landscape responses to fire exclusion in the Great Xing'an Mountains, China. *Int J Wildland Fire* 16(1):34–44. <https://doi.org/10.1071/WF05093>
- Chen X, Li B-L, Lin Z-S (2003) The acceleration of succession for the restoration of the mixed-broadleaved Korean pine forests in Northeast China. *For Ecol Manage* 177(1-3):503–514. [https://doi.org/10.1016/S0378-1127\(02\)00455-3](https://doi.org/10.1016/S0378-1127(02)00455-3)
- Chhin S, Hogg ET, Loeffers VJ, Huang S (2008) Potential effects of climate change on the growth of lodgepole pine across diameter size classes and ecological regions. *For Ecol Manage* 256(10):1692–1703. <https://doi.org/10.1016/j.foreco.2008.02.046>
- Chi H, Sun G, Huang J, Guo Z, Ni W, Fu A (2015) National forest aboveground biomass mapping from ICESat/GLAS data and MODIS imagery in China. *Remote Sens (Basel)* 7(5):5534–5564. <https://doi.org/10.3390/rs70505534>
- Crowther TW, Glick HB, Covey KR, Bettigole C, Maynard DS, Thomas SM, Smith JR, Hintler G, Duguid MC, Amatulli G (2015) Mapping tree density at a global scale. *Nature* 525(7568):201–205. <https://doi.org/10.1038/nature14967>
- Dai L, Jia J, Yu D, Lewis BJ, Zhou L, Zhou W, Zhao W, Jiang L (2013) Effects of climate change on biomass carbon sequestration in old-growth forest ecosystems on Changbai Mountain in Northeast China. *For Ecol Manage* 300:106–116. <https://doi.org/10.1016/j.foreco.2012.06.046>
- Diamantopoulou MJ, Özçelik R, Crecente-Campo F, Eler Ü (2015) Estimation of Weibull function parameters for modelling tree diameter distribution using least squares and artificial neural networks methods. *Biosyst Eng* 133:33–45. <https://doi.org/10.1016/j.biosystemseng.2015.02.013>
- Dijk WD, Hanberry BB, Fraser JS, He HS, Wang WJ, Thompson FR (2017) Revision and application of the LINKAGES model to simulate forest growth in central hardwood landscapes in response to climate change. *Landsc Ecol* 32(7):1365–1384. <https://doi.org/10.1007/s10980-016-0473-8>
- Dong L, Zhang L, Li F (2014) A compatible system of biomass equations for three conifer species in Northeast, China. *For Ecol Manage* 329:306–317. <https://doi.org/10.1016/j.foreco.2014.05.050>
- Dong L, Zhang L, Li F (2015) Developing additive systems of biomass equations for nine hardwood species in Northeast China. *Trees* 29(4):1149–1163. <https://doi.org/10.1007/s00468-015-1196-1>
- Duan S, He HS, Spetich MA, Wang WJ, Fraser JS, Xu W (2022) Long-term effects of succession, climate change and insect disturbance on oak-pine forest composition in the US Central Hardwood Region. *Eur J For Res* 141(1):153–164. <https://doi.org/10.1007/s10342-021-01428-2>
- Eskelson BN, Temesgen H, Lemay V, Barrett TM, Crookston NL, Hudak AT (2009) The roles of nearest neighbor methods in imputing missing data in forest inventory and monitoring databases. *Scand J For Res* 24(3):235–246. <https://doi.org/10.1080/02827580902870490>
- Fang J, Wang GG, Liu G, Xu S (1998) Forest biomass of China: an estimate based on the biomass–volume relationship. *Ecol Appl* 8(4):1084–1091. [https://doi.org/10.1890/1051-0761\(1998\)008\[1084:FBOCAE\]2.0.CO;2](https://doi.org/10.1890/1051-0761(1998)008[1084:FBOCAE]2.0.CO;2)
- Fang J, Wang X, Liu Y, Tang Z, White PS, Sanders NJ (2012) Multi-scale patterns of forest structure and species composition in relation to climate in northeast China. *Ecography* 35(12):1072–1082. <https://doi.org/10.1111/j.1600-0587.2012.00086.x>
- Fao/liasa/lisric/lsscscas/Jrc (2012) Harmonized world soil database (version 1.2). FAO, Rome and IIASA, Laxenburg. <http://web.archive.iiasa.ac.at/Research/LUC/External-World-soil-database/HTML/>
- Ferraz A, Saatchi SS, Longo M, Clark DB (2020) Tropical tree size–frequency distributions from airborne lidar. *Ecol Appl* 30(7):e02154. <https://doi.org/10.1002/eap.2154>
- Fu Y, He HS, Hawbaker TJ, Henne PD, Zhu Z, Larsen DR (2019) Evaluating k-nearest neighbor (kNN) imputation models for species-level aboveground forest biomass mapping in Northeast China. *Remote Sens (Basel)* 11(17):2005. <https://doi.org/10.3390/rs11172005>
- Gjertsen AK (2007) Accuracy of forest mapping based on Landsat TM data and a kNN-based method. *Remote Sens Environ* 110(4):420–430. <https://doi.org/10.1016/j.rse.2006.08.018>
- Hao Y, Widagdo FRA, Liu X, Quan Y, Liu Z, Dong L, Li F (2022) Estimation and calibration of stem diameter distribution using UAV laser scanning data: A case study for larch (*Larix olgensis*) forests in Northeast China. *Remote Sens Environ* 268:112769. <https://doi.org/10.1016/j.rse.2021.112769>
- Hayes DJ, Cohen WB, Sader SA, Irwin DE (2008) Estimating proportional change in forest cover as a continuous variable from multi-year MODIS data. *Remote Sens Environ* 112(3):735–749. <https://doi.org/10.1016/j.rse.2007.06.003>
- Huang C, He HS, Liang Y, Hawbaker TJ, Henne PD, Xu W, Gong P, Zhu Z (2021) The changes in species composition mediate direct effects of climate change on future fire regimes of boreal forests in northeastern China. *J Appl Ecol* 58(6):1336–1345. <https://doi.org/10.1111/1365-2664.13876>
- Hudak AT, Crookston NL, Evans JS, Hall DE, Falkowski MJ (2008) Nearest neighbor imputation of species-level, plot-scale forest structure attributes from LiDAR data. *Remote Sens Environ* 112(5):2232–2245. <https://doi.org/10.1016/j.rse.2007.10.009>
- Král K, Vrška T, Hort L, Adam D, Šamonil P (2010) Developmental phases in a temperate natural spruce–fir–beech forest: determination by a supervised classification method. *Eur. J For Res* 129(3):339–351. <https://doi.org/10.1007/s10342-009-0340-0>
- Lamb SM, MacLean DA, Hennigar CR, Pitt DG (2018) Forecasting forest inventory using imputed tree lists for LiDAR grid cells and a tree-list growth model. *Forests* 9(4):167. <https://doi.org/10.3390/f9040167>
- Lin C, Tsogt K, Zandraabal T (2016) A decompositional stand structure analysis for exploring stand dynamics of multiple attributes of a mixed-species forest. *For Ecol Manage* 378:111–121. <https://doi.org/10.1016/j.foreco.2016.07.022>
- Liu C, Beaulieu J, Prgent G, Zhang S (2009) Applications and comparison of six methods for predicting parameters of the Weibull function in unthinned *Picea glauca* plantations. *Scand J For Res* 24(1):67–75. <https://doi.org/10.1080/02827580802644599>
- Liu F, Li F, Zhang L, Jin X (2014) Modeling diameter distributions of mixed-species forest stands. *Scand J For Res* 29(7):653–663. <https://doi.org/10.1080/02827581.2014.960891>
- Liu Y, Yue C, Wei X, Blanco JA, Trancoso R (2020) Tree profile equations are significantly improved when adding tree age and stocking degree: an example for *Larix gmelinii* in the Greater Khingan Mountains of Inner Mongolia, northeast China. *Eur J For Res* 139(3):443–458. <https://doi.org/10.1007/s10342-020-01261-z>
- Mauro F, Frank B, Monleon VJ, Temesgen H, Ford KR (2019) Prediction of diameter distributions and tree-lists in southwestern Oregon using LiDAR and stand-level auxiliary information. *Can J For Res* 49(7):775–787. <https://doi.org/10.1139/cjfr-2018-0332>
- McRoberts RE (2012) Estimating forest attribute parameters for small areas using nearest neighbors techniques. *For Ecol Manage* 272:3–12. <https://doi.org/10.1016/j.foreco.2011.06.039>
- Olofsson K, Olsson H (2018) Estimating tree stem density and diameter distribution in single-scan terrestrial laser measurements of field plots: a simulation study. *Scand J For Res* 33(4):365–377. <https://doi.org/10.1080/02827581.2017.1368698>
- Packalén P, Maltamo M (2008) Estimation of species-specific diameter distributions using airborne laser scanning and aerial photographs. *Can J For Res* 38(7):1750–1760. <https://doi.org/10.1139/X08-037>
- Pan Y, Birdsey RA, Fang J, Houghton R, Kauppi PE, Kurz WA, Phillips OL, Shvidenko A, Lewis SL, Canadell JG (2011) A large and persistent carbon sink in the world's forests. *Science* 333(6045):988–993. <https://doi.org/10.1126/science.1201609>
- Popescu SC, Zhao K, Neuenschwander A, Lin C (2011) Satellite lidar vs. small footprint airborne lidar: Comparing the accuracy of aboveground biomass estimates and forest structure metrics at footprint level. *Remote Sens Environ* 115(11):2786–2797. <https://doi.org/10.1016/j.rse.2011.01.026>
- Riemann R, Wilson BT, Lister A, Parks S (2010) An effective assessment protocol for continuous geospatial datasets of forest characteristics using USFS Forest Inventory and Analysis (FIA) data. *Remote Sens Environ* 114(10):2337–2352. <https://doi.org/10.1016/j.rse.2010.05.010>

- Rodrigo R, Pettit JL, Matula R, Kozák D, Bače R, Pavlin J, Janda P, Mikoláš M, Nagel TA, Schurman J (2022) Historical mixed-severity disturbances shape current diameter distributions of primary temperate Norway spruce mountain forests in Europe. *For Ecol Manage* 503:119772. <https://doi.org/10.1016/j.foreco.2021.119772>
- Schmidt LN, Sanquetta MNI, McTague JP, da Silva GF, Fraga Filho CV, Sanquetta CR, Soares Scolforo JR (2020) On the use of the Weibull distribution in modeling and describing diameter distributions of clonal eucalypt stands. *Can J For Res* 50(10):1050–1063. <https://doi.org/10.1139/cjfr-2020-0051>
- Schmitt CB, Burgess ND, Coad L, Belokurov A, Besançon C, Boisrobert L, Campbell A, Fish L, Gliddon D, Humphries K (2009) Global analysis of the protection status of the world's forests. *Biol Conserv* 142(10):2122–2130. <https://doi.org/10.1016/j.biocon.2009.04.012>
- Schütz JP, Rosset C (2020) Performances of different methods of estimating the diameter distribution based on simple stand structure variables in monospecific regular temperate European forests. *Ann For Sci* 77:1–11. <https://doi.org/10.1007/s13595-020-00951-3>
- Scolforo HF, McTague JP, Burkhart H, Roise J, Campoe O, Stape JL (2019) Eucalyptus growth and yield system: linking individual-tree and stand-level growth models in clonal Eucalypt plantations in Brazil. *For Ecol Manage* 432:1–16. <https://doi.org/10.1016/j.foreco.2018.08.045>
- Shang C, Treitz P, Caspersen J, Jones T (2017) Estimating stem diameter distributions in a management context for a tolerant hardwood forest using ALS height and intensity data. *Can J Remote Sens* 43(1):79–94. <https://doi.org/10.1080/07038992.2017.1263152>
- Slik J, Aiba SI, Brearley FQ, Cannon CH, Forshed O, Kitayama K, Nagamasu H, Nilus R, Payne J, Paoli G (2010) Environmental correlates of tree biomass, basal area, wood specific gravity and stem density gradients in Borneo's tropical forests. *Glob Ecol Biogeogr* 19(1):50–60. <https://doi.org/10.1111/j.1466-8238.2009.00489.x>
- Su Y, Guo Q, Xue B, Hu T, Alvarez O, Tao S, Fang J (2016) Spatial distribution of forest aboveground biomass in China: estimation through combination of spaceborne lidar, optical imagery, and forest inventory data. *Remote Sens Environ* 173:187–199. <https://doi.org/10.1016/j.rse.2015.12.002>
- Temesgen H, LeMay V, Froese K, Marshall P (2003) Imputing tree-lists from aerial attributes for complex stands of south-eastern British Columbia. *For Ecol Manage* 177(1-3):277–285. [https://doi.org/10.1016/S0378-1127\(02\)00321-3](https://doi.org/10.1016/S0378-1127(02)00321-3)
- Thomas V, Oliver R, Lim K, Woods M (2008) LiDAR and Weibull modeling of diameter and basal area. *For Chron* 84(6):866–875. <https://doi.org/10.5558/tfc84866-6>
- Valbuena R, Vauhkonen J, Packalen P, Pitkänen J, Maltamo M (2014) Comparison of airborne laser scanning methods for estimating forest structure indicators based on Lorenz curves. *ISPRS J of Photogramm Remote Sens* 95:23–33. <https://doi.org/10.1016/j.isprsjprs.2014.06.002>
- Wang J, Wang C, Zang S (2017) Assessing re-composition of Xing'an larch in boreal forests after the 1987 fire, Northeast China. *Remote Sens* 9(5):504. <https://doi.org/10.3390/rs9050504>
- Wang S, Dai L, Liu G, Yuan J, Zhang H, Wang Q (2006) Modeling diameter distribution of the broadleaved-Korean pine mixed forest on Changbai Mountains of China. *Sci China Ser E: Technol Sci* 49(1):177–188. <https://doi.org/10.1007/s11431-006-8119-8>
- Wang WJ, Thompson FR III, He HS, Fraser JS, Dijk WD, Jones-Farrand T (2019) Climate change and tree harvest interact to affect future tree species distribution changes. *J Ecol* 107(4):1901–1917. <https://doi.org/10.1111/1365-2745.13144>
- Westphal C, Tremer N, von Oheimb G, Hansen J, von Gadow K, Härdtle W (2006) Is the reverse J-shaped diameter distribution universally applicable in European virgin beech forests? *For Ecol Manage* 223(1-3):75–83. <https://doi.org/10.1016/j.foreco.2005.10.057>
- Wilson BT, Lister AJ, Riemann RI (2012) A nearest-neighbor imputation approach to mapping tree species over large areas using forest inventory plots and moderate resolution raster data. *For Ecol Manage* 271:182–198. <https://doi.org/10.1016/j.foreco.2012.02.002>
- Xu L, Jin G (2012) Species composition and community structure of a typical mixed broadleaved-Korean pine (*Pinus koraiensis*) forest plot in Liangshui Nature Reserve, Northeast China. *Biodivers Sci* 20(4):470 (in Chinese). <https://doi.org/10.3724/SPJ.1003.2012.12233>
- Xu W, He HS, Hawbaker TJ, Zhu Z, Henne PD (2020) Estimating burn severity and carbon emissions from a historic megafire in boreal forests of China. *Sci Total Environ* 716:136534. <https://doi.org/10.1016/j.scitotenv.2020.136534>
- Zasada M, Cieszewski CJ (2005) A finite mixture distribution approach for characterizing tree diameter distributions by natural social class in pure even-aged Scots pine stands in Poland. *For Ecol Manage* 204(2-3):145–158. <https://doi.org/10.1016/j.foreco.2003.12.023>
- Zhang Q, He HS, Liang Y, Hawbaker TJ, Henne PD, Liu J, Huang S, Wu Z, Huang C (2018a) Integrating forest inventory data and MODIS data to map species-level biomass in Chinese boreal forests. *Can J For Res* 48(5):461–479. <https://doi.org/10.1139/cjfr-2017-0346>
- Zhang Q, Liang Y, He HS (2018b) Tree-lists estimation for Chinese boreal forests by integrating Weibull diameter distributions with MODIS-based forest attributes from kNN imputation. *Forests* 9(12):758. <https://doi.org/10.3390/f9120758>
- Zhang Y, He HS, Dijk WD, Yang J, Shifley SR, Palik BJ (2009) Integration of satellite imagery and forest inventory in mapping dominant and associated species at a regional scale. *Environ Manag* 44(2):312–323. <https://doi.org/10.1007/s00267-009-9307-7>
- Zhang Y, Liang S (2014) Changes in forest biomass and linkage to climate and forest disturbances over Northeastern China. *Glob Chang Biol* 20(8):2596–2606. <https://doi.org/10.1111/gcb.12588>
- Zhang Y, Liang S, Sun G (2013) Forest biomass mapping of northeastern China using GLAS and MODIS data. *IEEE J Sel Top Appl Earth Obs Remote Sens* 7(1):140–152. <https://doi.org/10.1109/JSTARS.2013.2256883>
- Zhang Z, Cao L, Mulverhill C, Liu H, Pang Y, Li Z (2019) Prediction of diameter distributions with multimodal models using LiDAR data in subtropical planted forests. *Forests* 10(2):125. <https://doi.org/10.3390/f10020125>

Publisher's note

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

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

