Received: 2 December 2023 | Revised: 10 February 2024 | Accepted: 12 February 2024

DOI: 10.1111/gcb.17224

Global Change Biology WILEY

RESEARCH ARTICLE

Global patterns of tree wood density

Hui Yang[1] | Siyuan Wang[1,2] | Rackhun Son[1,3] | Hoontaek Lee[1,2] | Vitus Benson[1,4] | Weijie Zhang[1] | Yahai Zhang[5] | Yuzhen Zhang[1] | Jens Kattge[1,6] | Gerhard Boenisch[1] | Dmitry Schepaschenko[7] | Zbigniew Karaszewski[8] | Krzysztof Stereńczak[9] | Álvaro Moreno-Martínez[10] | Cristina Nabais[11] | Philippe Birnbaum[12,13] | Ghislain Vieilledent[12] | Ulrich Weber[1] | Nuno Carvalhais[1,4,14]

[1]Max Planck Institute for Biogeochemistry, Jena, Germany

[2]Institute of Photogrammetry and Remote Sensing, Technische Universität Dresden, Dresden, Germany

[3]Department of Environmental Atmospheric Sciences, Pukyong National University, Busan, South Korea

[4]ELLIS Unit Jena, Jena, Germany

[5]State Key Laboratory of Earth Surface Processes and Resource Ecology, Faculty of Geographical Science, Beijing Normal University, Beijing, China

[6]German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig, Leipzig, Germany

[7]International Institute for Applied Systems Analysis (IIASA), Laxenburg, Austria

[8]Research Group of Chemical Technology and Environmental Protection, Łukasiewicz Research Network Poznań Institute of Technology Center of Sustainable Economy, Poznań, Poland

[9]Department of Geomatics, Forest Research Institute, Raszyn, Poland

[10]Image Processing Laboratory (IPL), Universitat de València, València, Spain

[11]Centre for Functional Ecology, Associate Laboratory TERRA, Department of Life Sciences, University of Coimbra, Coimbra, Portugal

[12]AMAP, Univ Montpellier, CIRAD, CNRS, INRAE, IRD, Montpellier, France

[13]Institut Agronomique néo-Calédonien (IAC), Nouméa, New Caledonia

[14]Departamento de Ciências e Engenharia do Ambiente, DCEA, Faculdade de Ciências e Tecnologia, FCT, Universidade Nova de Lisboa, Caparica, Portugal

Correspondence
Hui Yang and Nuno Carvalhais, Max Planck Institute for Biogeochemistry, Jena, Germany.
Email: huiyang@bgc-jena.mpg.de and ncarvalhais@bgc-jena.mpg.de

Funding information
GlobBiomass DUE Project, Grant/Award Number: 4000113100/14/I-NB; German Federal Ministry for Economic Affairs and Climate Action, Grant/Award Number: 50EE1904; ESM2025; H2020 European Research Council, Grant/Award Number: 855187; International Max Planck Research School for Biogeochemical Cycles; ESA IFBN project, Grant/Award Number: 4000114425/15/NL/FF/gp; ESA FRM4BIOMASS, Grant/Award Number: 4000142684/23/I-EF-bgh; Poland National Centre for

Abstract

Wood density is a fundamental property related to tree biomechanics and hydraulic function while playing a crucial role in assessing vegetation carbon stocks by linking volumetric retrieval and a mass estimate. This study provides a high-resolution map of the global distribution of tree wood density at the 0.01° (~1 km) spatial resolution, derived from four decision trees machine learning models using a global database of 28,822 tree-level wood density measurements. An ensemble of four top-performing models combined with eight cross-validation strategies shows great consistency, providing wood density patterns with pronounced spatial heterogeneity. The global pattern shows lower wood density values in northern and northwestern Europe, Canadian forest regions and slightly higher values in Siberia forests, western United States, and southern China. In contrast, tropical regions, especially wet tropical areas, exhibit high wood density. Climatic predictors explain 49%–63% of spatial variations, followed by

Research and Development REMBIOFOR project, Grant/Award Number: BIOSTRATEG1/267755/4/NCBR/2015

vegetation characteristics (25%–31%) and edaphic properties (11%–16%). Notably, leaf type (evergreen vs. deciduous) and leaf habit type (broadleaved vs. needleleaved) are the most dominant individual features among all selected predictive covariates. Wood density tends to be higher for angiosperm broadleaf trees compared to gymnosperm needleleaf trees, particularly for evergreen species. The distributions of wood density categorized by leaf types and leaf habit types have good agreement with the features observed in wood density measurements. This global map quantifying wood density distribution can help improve accurate predictions of forest carbon stocks, providing deeper insights into ecosystem functioning and carbon cycling such as forest vulnerability to hydraulic and thermal stresses in the context of future climate change.

KEYWORDS
carbon stocks, climate stresses, machine learning, plant traits, tree physiology, vegetation resilience

1 | INTRODUCTION

Forests, occupying one-third of the global land areas, play an important role in absorbing (through photosynthesis) approximately 15.6 billion metric tons of carbon dioxide from the atmosphere each year (Harris et al., 2021). About half of this assimilated carbon is used for forest growth and subsequently stored in terrestrial ecosystems for a long period (Pan et al., 2011), a process known as carbon sequestration. However, our understanding of the magnitude of carbon sequestration within terrestrial ecosystems, along with its inherent variability, remains inadequately characterized, despite the availability of remote sensing products that have recently provided extensive information related to forest coverage (Song et al., 2018), forest structure (Burt et al., 2021) and canopy height (Potapov et al., 2021). This is because remote sensing mainly provides volumetric proxies, which necessitate the incorporation of a density estimate to obtain the mass of carbon stocks. One major source of uncertainty concerning the spatial distribution of carbon stocks within forests lies in wood density (Chave et al., 2019), which is defined as the ratio of dry mass of wood to its green volume and cannot be directly observable from space.

Beyond its role in computing carbon sequestration, wood density has obtained growing attention as a plant functional trait, particularly for its links to the biomechanical support of trees and hydraulic safety (Serra-Maluquer et al., 2022). Wood density has direct or indirect influence on a range of ecological processes, including tree growth, tree resistance, resilience and recovery to disturbances, and tree mortality primarily through two mechanisms (King et al., 2005; O'Brien et al., 2017; Roderick & Berry, 2001). First, studies have reported a tight relationship between wood density and tree mortality rate or ecosystem carbon turnover, representing the overall carbon loss rate over several years or even decades. For example, the distinct variations in tree mortality rates observed in eastern and western Amazon forests have been attributed to differences in wood density, where lower density is associated with higher mortality rates (Chao et al., 2008). Additionally, plot-level synthetic studies have revealed the predictive capacity of wood density in characterizing forest dynamics in response to abrupt natural and anthropogenic disturbances such as extreme weather conditions, fires, and biotic attacks. Anderegg et al. (2016) demonstrated that wood density helped explain the observed variability in tree mortality rates across angiosperms. Nevertheless, the role of wood density in explaining the response of trees to water stress remains a topic of ongoing debate. Liang et al. (2021) and Greenwood et al. (2017), drawing on global synthesis of seedlings and adult trees' measurements, reported positive relationship between wood density and the drought resistance. However, this relationship may take a opposite direction in certain regional studies, such as Hoffmann et al. (2011) reported in temperate forest that the species with high wood density, owing to their insensitivity to water stress, exhibit limited capacity to regulate plant water potential. Thus, high wood density species show lower resistance and higher mortality rate under extremely severe drought conditions.

Up to now, despite several forest inventory data have been used for shaping regional wood density distribution, as demonstrated by Chave et al. (2009), Poyatos et al. (2018), and Oliveira et al. (2021), a reliable global-scale spatially explicit product of tree wood density is lacking. Here, we conducted a comprehensive collection of tree-level wood density measurements spanning all global climate biomes. Our primary objective is to generate a global high-resolution map of tree wood density and thoroughly assess its robustness. To achieve this, we developed a comprehensive global dataset of georeferenced species-based wood density observations together with basic leaf attributes of the respective species. We then trained four distinct machine learning models using forest inventory data, combined with data for climate, edaphic, and vegetation characteristics derived from global observation products (Figure S1). We then used eight cross-validation methods to assess the predictive capacity of

these models in estimating variations in wood density. Finally, we provided an exposition of both the global and regional patterns of the wood density product, including an analysis of the associated uncertainties, and investigated the predictive factors influencing the observed spatial variations in wood density.

2 | MATERIALS AND METHODS

2.1 | Wood density measurements

We requested tree-level measurements for "wood density" (Trait ID 4 and 3064) from the TRY v6 database (Kattge et al., 2020; https://www.try-db.org/TryWeb/Home.php). These measurements are mainly sourced from the Biomass And Allometry Database (BAAD), RAINFOR trait database, Neotropical trait database, Tropical plant trait database, Netherland trait database, Chinese trait database, Panama trait database, and Spanish trait database (detailed information see Table S1). It is important to note that we retain only those entries pertaining to trees with precise coordinate information and documented tree species. In total, 26,709 measurements obtained from the TRY database are used in our analysis. Moreover, we incorporate 8743 tree-level wood density measurements in the Eurasia region from Schepaschenko et al. (2017). This dataset is derived from a compilation of experiments undertaken by the authors, along with data extracted from scientific publications. Additionally, we include two unpublished databases, one from the Poland Forest Research Institute and the other from the UMR AMAP wood density database. The former comprises over 48,000 wood density samples, which were measured in the year 2018 for 2971 trees within Poland. The latter includes a total of 2967 wood density measurements per tree originating from New Caledonia, an archipelago situated in the southwest Pacific Ocean. To reduce the disparities in sample size among regions and achieve a more balanced spatial distribution of measurements, we perform a random selection process, retaining only 10% of the data from the Poland and UMR AMAP databases for our analysis. In summary, we construct a global database consisting of 36,046 tree-level wood density measurements spanning every continent where forests are present (Figure 1a). This dataset covers a wide range of climate space, although it has only a few data in the desert and tundra regions due to their low forest coverage (Figure 1b).

2.2 | Covariates variables

For each species for which we gathered wood density measurement, we identify its specific leaf type (broadleaf or needleleaf) and leaf habit type (evergreen or deciduous) from TRY database (Kattge et al., 2020). The attributes of leaf type and leaf habit type serve as predictive variables in our wood density prediction model. We decided to use leaf type and leaf habit type instead of tree species, because a global map of tree species abundance is currently unavailable. Besides, to enhance our capacity to predict global wood density variations, based on the coordinate information, we extract a range of proxies for vegetation characteristics, including land cover type, tree cover fraction, vegetation carbon productivity, canopy greenness, and vegetation water content, which also act as predictive variables (details on Table S2). Furthermore, we extract all the climatic variables from the high-resolution database of WorldClim version 2 (Fick & Hijmans, 2017), and extracted all relevant soil-related variables from the global gridded soil information, SoilGrids database (https://soilgrids.org/; Batjes et al., 2020).

When confronted with missing values in the training dataset, we addressed them by approximating average values based on the spatial proximity of data collection sites. The proportion of missing data for climate- and soil-related covariates is less than 1%; however, several satellite-based vegetation covariates have a high fraction of missing data (Figure S5a). Specifically, we utilized data from the five nearest sites within a geospatial radius of 500 kilometers

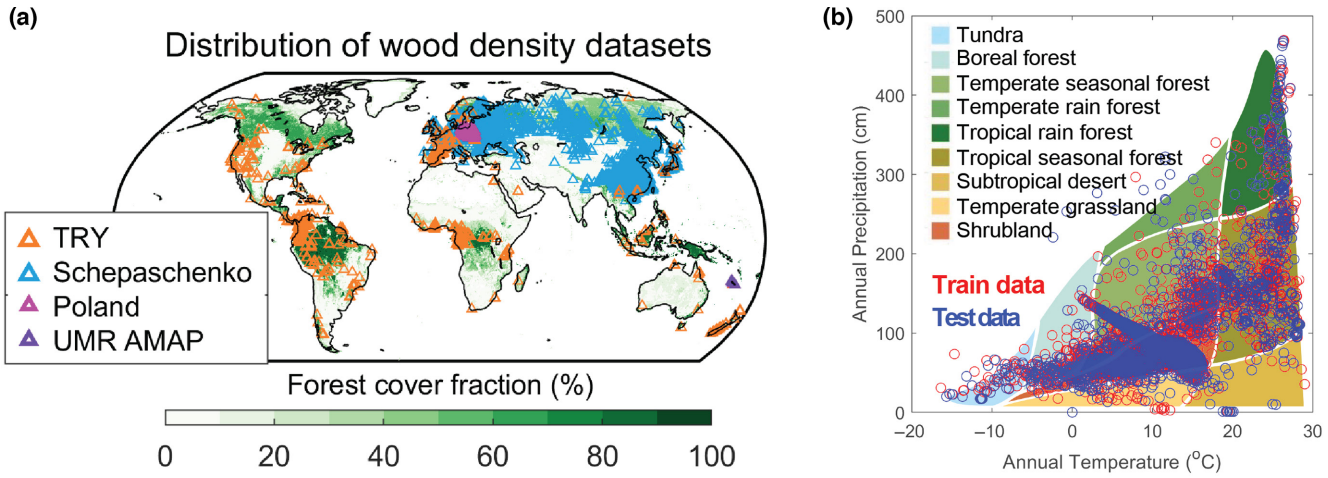

FIGURE 1 (a) Spatial distribution of the forest plots originating from four datasets used for the wood density maps. The background colors indicate the fraction of forest cover from Climate Change Initiative land cover maps. (b) Distribution of the forest plots in a climate space defined by annual temperature and total annual precipitation. Red dots represent training data, blue dots testing data.

to estimate missing values for each data point. In instances where the search radius encompassed fewer than five available sites, the number of sites used for averaging was determined by the actual count of data available points, if the available sites numbered more than three, otherwise we opted to compute the average using the data from the three closest sites, with the weights inversely proportional to their geographical distance. The coefficients of variation of some vegetation covariate values from the nearest sites exhibit a notable degree of variability, which has the potential to result in biases; however, these vegetation covariates in our machine learning models have low importance (Figure S5b; Figure 4). In addition, the target value is the pixel-level averaged wood density, calculated as the average of all tree-level wood density measurements for each leaf type and leaf habit type within the identical geographic grid cell (0.01°×0.01°), since their predictive variables are identical. This calculation is based on the assumption that sample distributions are consistent with population distributions. It is important to acknowledge that the pixel-level wood density estimations derived from tree-level measurements may be subject to bias when the selection of sampling distribution is skewed, particularly in regions with high species richness.

2.2.1 | Machine learning models

Our machine learning models are based on decision trees, which are inherently interpretable and have been shown to be superior compared to neural networks in handling tabular data (Grinsztajn et al., 2022). Specifically, we utilize two popular training schemes for decision tree ensembles: Bagging (Random Forest) and Boosting (Gradient Boosting). Furthermore, for each of these schemes, we leverage two current state-of-the-art software package implementations (Table 1). First, regarding random forest schemes, we use two models that differ in the treatment of categorical variables: the Scikit-Learn Random Forest (Pedregosa et al., 2011) cannot handle categorical data explicitly, but instead treats them as continuous or using one-hot encoding. In contrast, the LightGBM Regularized Random Forest uses categorical information for leaf splitting and uses a training routine optimized for speed. Second, we use two models based on the gradient boosting schemes with different training routines. Extreme gradient boosting tree (XGBoost, Chen & Guestrin, 2016), which is known for maximum predictive performance, adopts a level-wise tree growth strategy, navigating through gradient values and partially aggregating them to assess the quality of all possible splits. To mitigate overfitting issues associated with increasing tree depth, XGBoost introduces complexity as a regularization term, seamlessly incorporated into its cost functions. Unlike XGBoost's level-wise tree growth strategy, LightGBM adopts a leaf-wise strategy, efficiently increasing the complexity of the tree structure to determine branch points through histogram-based methods (Ke et al., 2017). LightGBM is centered around parallelization and is characterized by two innovative techniques: gradient-based one-side sampling, which reduces the use of low-gradient data during

TABLE 1 A list of four machine learning models used for prediction.

Model	LightGBM	LightGBM-RF	Random forest	XGBoost
Python package	lightgbm.LGBMRegressor()	lightgbm.LGBMRegressor()	sklearn.ensemble.RandomForestRegressor()	xgboost.XGBRegressor()
Training scheme	Boosting	Bagging	Bagging	Boosting
Training algorithm	Histogram-based approach	Histogram-based approach	Pre-sorting approach	Pre-sorting approach
Categorical feature treatment	Node splitting	Node splitting	Ordinal encoder	None
Prevents overfitting	Yes	Yes	No	Yes
Computational cost	(relatively) Low	Low	High	Medium
Advantage	Fast training	Fast training	Fewer parameters	Parallel processing
Disadvantage	Higher overfitting risk	Higher overfitting risk	Larger memory req.	Complex parameter tuning

the training process, and exclusive feature bundling, which groups features together. These techniques provide advantages, including the acceleration of the training process with high accuracy and reduction of memory usage. In this analysis, we randomly partition our wood density database into training and testing subsets, allocating 80% of the measurements to the training set and reserving the remaining 20% for testing. Both the training and testing datasets cover the identical climate space (Figure 1b).

2.3 | Cross-validation strategies

To evaluate the performance of these machine learning models, we use a leave-one-cluster-out cross-validation method. The clusters for cross-validation are defined using eight different methods, that is, random fivefold, spatial blocked 10-fold (where the global domain was divided into spatial blocks of 5° and randomly assigned to one of 10 folds), and classifications based on two sets of Köppen climate maps, European Space Agency Climate Change Initiative (ESA CCI) land cover types, FAO ecozones, and latitudinal and longitude binning. These cross-validation methods (except for random 5-fold method) are chosen to mitigate potential spatial auto-correlation and to avoid optimistically evaluating model performance (Ludwig et al., 2023; Ploton et al., 2020). For the assessment of predictive capabilities concerning wood density, we used two metrics, namely the root-mean-square error (RMSE) and the goodness of fit (R^2).

2.4 | Upscaling and generating global maps

A total of 79 selected vegetation, climatic, and edaphic variables (listed in Table S2) are used to train the four machine learning models, to conduct cross-validation analysis, and to upscale the global maps. To upscale wood density, we prepare global maps of all predictive vegetation, climatic, and edaphic variables. The initial spatial resolution of climatic, edaphic, and some satellite-based vegetation covariates is higher than the target resolution of 1 km for the resulting product. These input drivers were subjected to spatial aggregation, through the computation of spatial averages within 1-km spatial windows. The upscaling procedure is done in two steps. First, we use our trained machine learning models to predict wood density for distinct leaf type and leaf habit types. Then, we use a global map of the fractions of plant function type from the ESA CCI as weighting factors for calculating average wood density values at the 1 km pixel level. In total, we generate an ensemble of 32 gridded tree wood density maps, all characterized by a spatial resolution of 1 km. These maps are obtained using the four different machine learning models, generated as the average of maps derived from eight cross-validation methods (shown in Figure 2). Furthermore, the overall ensemble average of these four maps is calculated (shown in Figure 3).

3 | RESULTS AND DISCUSSION

3.1 | Machine learning model evaluation

Overall model performance is high with an R^2 value of 0.55 ± 0.03 (mean ± SD across machine learning models and cross-validation splits) and an RMSE of 0.11 ± 0.01 g/cm^3 (Figure S2a) on the test data. Between machine learning models as well as across cross-validation splits, there is little variability. The cross-validation analyses are performed to assess the extrapolation capability of machine learning models and provide a more comprehensive understanding of the driving factors (Sweet et al., 2023). When using the random fivefold cross-validation method, the cross-validation predictions from the four machine learning models show high R^2 and low RMSE (R^2 = .50–.56, RMSE = 0.05–0.12 g/cm^3), while the predictive capacity exhibits a reduction when using the other cross-validation methods to reduce the influence of spatial auto-correlation, indicating the limited extrapolation capability (Figure S2b). Such reduction could be due to the unique features inherent to particular climatic biome or ecology zone, so excluding one cluster from a specific climatic biome or ecology zone during the training process may result in biases in cross-validation predictions, ultimately leading to a decrease in the model predictive performance. Thus, it highlights the importance of using the data covering all the climatic biomes during the training process to mitigate extrapolation when generating a global map.

Our models suffer from a small generalization gap, mainly visible in the decrease in R^2 value from training data to testing dataset. Performance on the training set displays an R^2 value of 0.68 ± 0.15 and an RMSE of 0.10 ± 0.15 (g/cm^3) (Figure S2a). More specifically, however, the drop in R^2 value is strong for the Random Forest and particularly within the tropical region (30° S–15° N), as well as mid-latitude areas (45° N–60° N) (Figure S2c,d). This discrepancy in the tropical region may be due to high species richness and the limited number of wood density measurements in this region, which result in the biases found in wood density predictions. In highly diverse tropical forests, high phylogenetic endemism among tree species is not only affected by current environmental conditions but also characterized by long-term climate stability (Guo et al., 2023). For instance, even within the same location and forest type, several species exhibit varying wood densities. Ideally, incorporating species as a categorical variable to refine the analysis would be valuable, but a comprehensive global species distribution map is unavailable, precluding the inclusion of species as a predictor in our machine learning models.

3.2 | Global and regional patterns

The global maps of wood density generated by the four machine learning models show consistent patterns (Figure 2), characterized by pronounced spatial heterogeneity across latitudes and from dry to wet regions (notably shown in the zoomed regional maps in

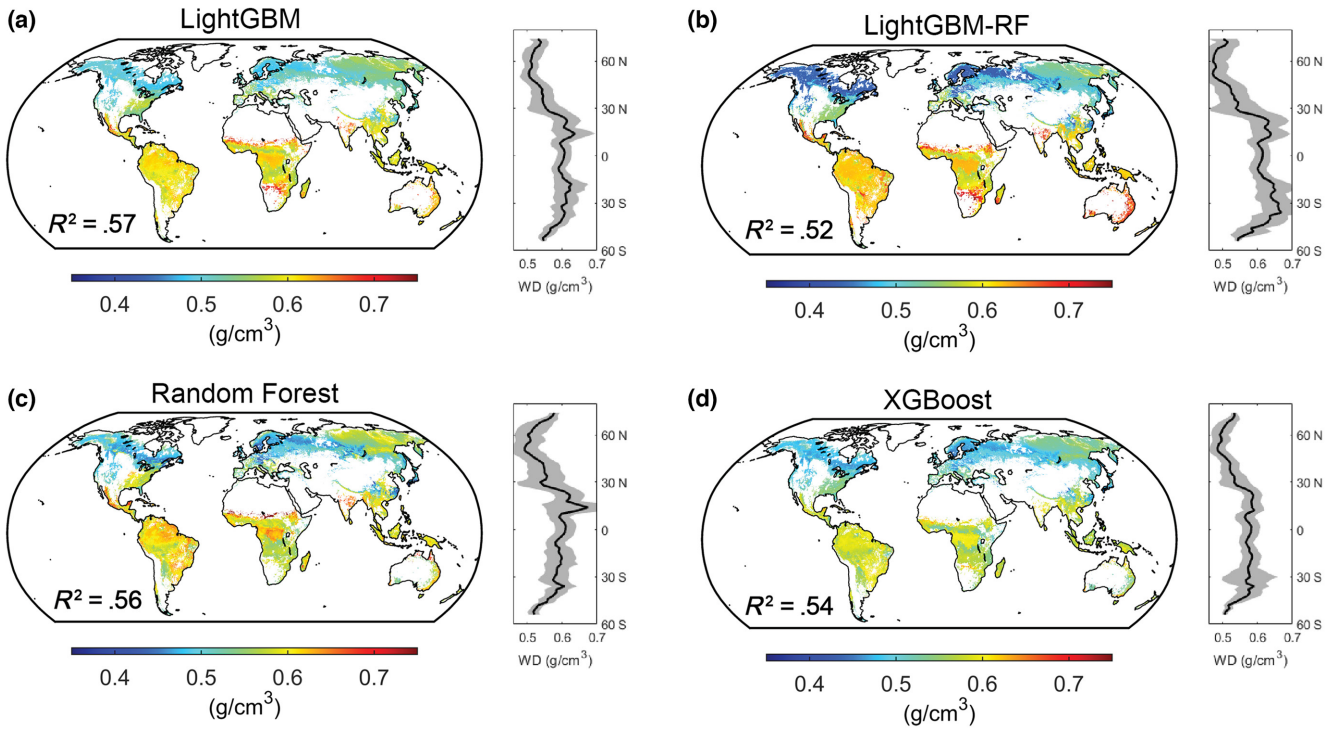

FIGURE 2 Global 1 km mean wood density maps from four distinct machine learning models: LightGBM (a), LightGBM-RF (b), Random Forest (c), and XGBoost (d). Patterns represent the average wood density values derived from models using eight distinct cross-validation strategies. These wood density values are presented exclusively within regions where the forest fraction exceeds 10%. The R^2 values for each machine learning model's prediction of test data are written on the maps. The inset on the right-hand side shows the latitudinal averages (black line) and standard deviations (grey shading) of wood density.

Figure 3; Figure S6). The main features are as follows: In the northern high latitudes, wood density values are notably lower in northern and northwestern Europe, as well as in the Canadian forest regions, where evergreen needleleaf trees predominate. In contrast, Siberia exhibits higher wood density values (~0.55 g/cm^3), particularly in its eastern regions. This disparity can be attributed to the predominance of deciduous needleleaf trees, primarily the *Larix* genus, within the arid landscapes of eastern Siberia (Sato et al., 2016). Then, a slightly higher wood density is observed in temperate regions such as western United States and southern China. Moreover, wood density in the tropical regions generally holds higher values compared to the extra-tropical regions. Within the tropics, wood density in wet tropical areas tends to be relatively higher than that in dry tropical regions.

We further compare our spatial patterns of wood density maps with other independent regional wood density products (Chave et al., 2009; Oliveira et al., 2021). Chave et al. (2009) used the 16,468 wood density entries from Dryad data repository, which have not been used in our analysis due to the unavailability of geographic coordinate information, to generate a wood density map for North and South America (hereafter JC09). Overall, JC09 product and our product both exhibit consistent spatial distributions, that is, lower wood density in high latitudes of North America and areas near the Andes Mountain while higher wood density is primarily observed in the wet and transitional tropical regions of South America. But our wood density product depicts relatively lower absolute values of wood density in the wet tropical region compared to the JC09 product. Furthermore, differences are also apparent in the dry tropics and semiarid areas of South America. The JC09 product shows higher wood density in these regions, compared to the wet tropical areas, while our dataset shows the opposite, particularly in eastern Brazil and around Paraguay region. These discrepancies could be explained either by the lack of measurements used in our product or the simplicity of the multiple regression method utilized in JC09 product. Additionally, a recent analysis by Oliveira et al. (2021), which generated a wood density pattern for eastern Brazil through data collection and kriging extrapolation approach, exhibits a consistent range of wood density values and similar spatial distributions when compared to our maps.

3.3 | Uncertainty analysis

We explore the uncertainty stemming from the selection of machine learning models and the application of various cross-validation methods (Figure S3). First of all, when using the same machine learning model, the utilization of different cross-validation methods for model training tends to generate relatively consistent global wood density maps. That is, the standard deviation in wood density across eight

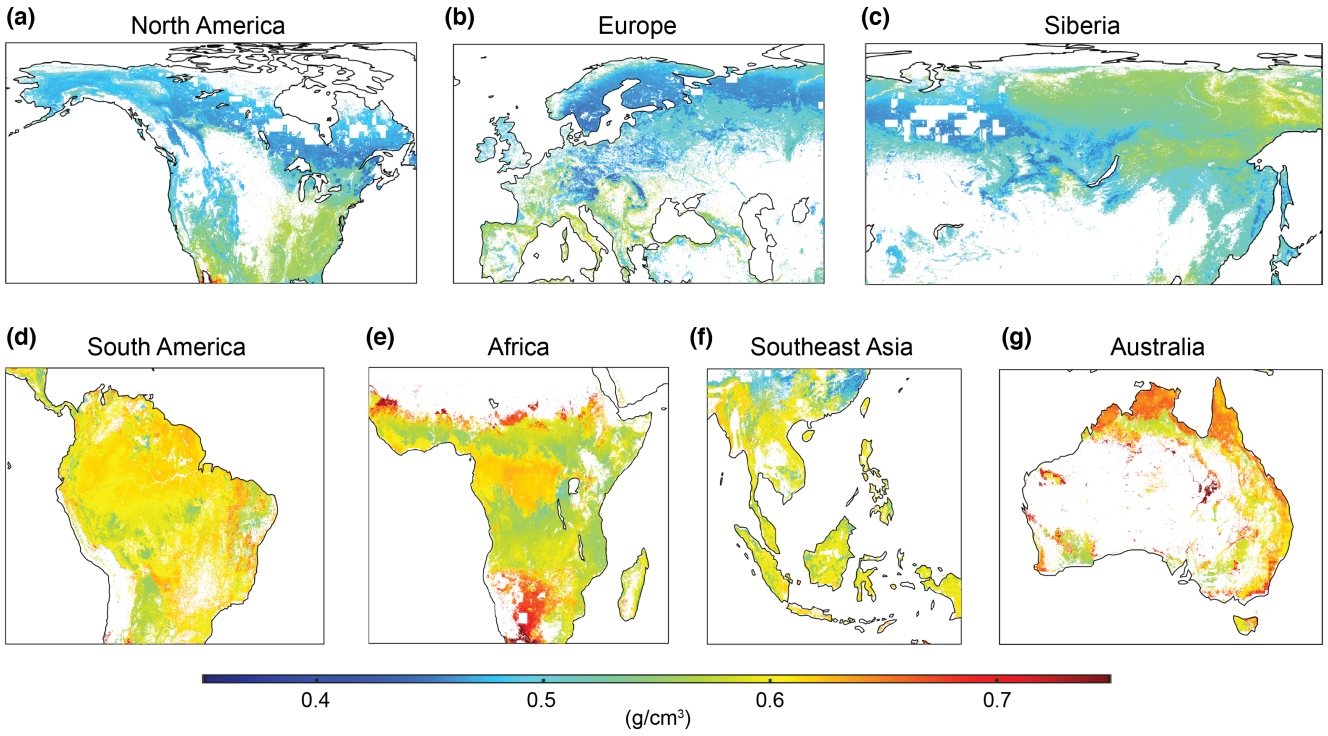

FIGURE 3 Zoomed-In 1 km wood density map over: North America (a), Europe (b), Siberia (c), South America (d), Africa (e), Southeast Asia (f), and Australia (g). Shown is the average wood density prediction from four machine learning models for grid cells with a forest fraction above 10%.

distinct cross-validation strategies is consistently below 0.01 g/cm^3, albeit some tropical savannas show a relatively high standard deviation (~0.03 g/cm^3) (Figure S3a,c). Conversely, the uncertainty associated with the choice of machine learning models is relatively higher, but the standard deviation in wood density remains below 0.05 g/cm^3, which is less than 8% of the mean values (Figure S3b,d). Regions with high uncertainty across machine learning models include northern Canada, northern Europe, southern Siberia, southwestern China as well as part of tropical savanna regions. Furthermore, we implement an analysis of the number of wood density measurements collected for 44 subregions based on the IPCC subregion reference map (Iturbide et al., 2020). Our analysis reveals a significant negative relationship between the standard deviation in wood density predictions across machine learning models and the number of observations (Figure S3e,f) This suggests that the regions with limited wood density measurements exhibit higher uncertainty in the areas due to the well-known low generalizability of machine learning models for small and unbalanced sample classes (Jung et al., 2020).

3.4 | Factors influencing spatial variations

We further assess the feature importance of the four machine learning models, aiming to elucidate which and how the factors influence spatial variations in wood density (Figure 4). First of all, the category of climatic condition plays a critical role, accounting for a substantial proportion of spatial variations, with contributions ranging from 49% to 63%. Following in significance is the category of vegetation characteristics, explaining 25%–31% of the variations, while edaphic properties have a discernible and relatively smaller influence, contributing to 11%–16% of the variations. More precisely, the importance of vegetation characteristics category mainly arises from attributes associated with leaf type and leaf habit type, which is also the most dominant individual feature among all 79 selected ones. In contrast, within the category of climatic condition, several factors such as temperature seasonality, total annual precipitation, and cloud cover have high importance in explaining wood density variation.

The importance of species in shaping wood density has been highlighted in various prior studies (e.g., Nabais et al., 2018; Ogle et al., 2014; Thurner et al., 2014). In our study, leaf type and leaf habit types, which are identified based on the species, also serve as key dominants in predicting wood density across all four machine learning models. It is worth noting that forest age and plant leaf traits (with the exception of leaf type and leaf habit type) have not been incorporated as predictive variables in our analysis, despite their potential relevance in explaining wood density, as reported by previous studies (Bouriaud et al., 2004, 2005; Chave et al., 2009). The primary rationale for the omission is the lack of accurate, high-resolution, independent, and global maps for these variables. However, we recognize the potential advancements offered by recently published global maps of forest age and plant traits, generated through machine learning techniques (Besnard et al., 2021; Huang et al., 2021; Moreno-Martínez et al., 2018). First, we acknowledge

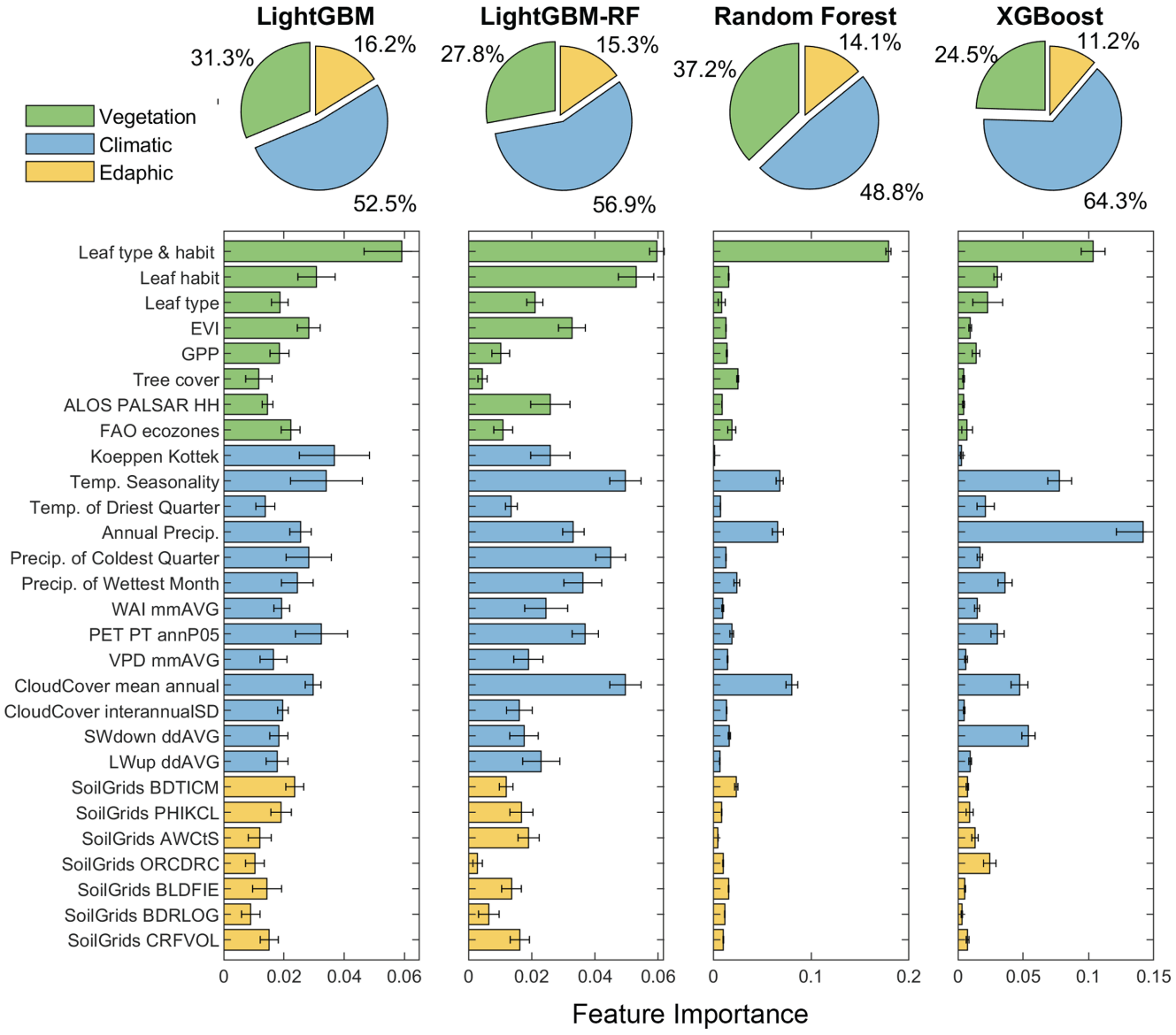

FIGURE 4 Relative feature importances across machine learning models. The pie plots show the sum of importance of all variables in each group (vegetation, climatic, and edaphic). The bar plots show the importance of each predictor. The error bars denote the s.d. of importance across cross-validation strategies. The details of predictor covariates are listed in Table S2.

that the forest age map from Besnard et al. (2021) is incompatible with our analysis due to the interdependence with forest biomass from GlobBiomass—a critical covariate in predicting forest age. The algorithm for biomass in GlobBiomass relies on estimated wood density, creating a circular dependency that precludes the use of Besnard et al.'s map in our framework. Second, we conducted additional assessments of model performance, comparing models with and without plant leaf traits and forest age as predictive variable. The results show no discernible differences in model performance (Figure S7). This suggests that the role of plant leaf traits and forest age may be effectively replaced by other vegetation and/or climate factors with similar spatial distributions in our predictive models. Another possible explanation for the lack of new information provided by leaf traits is that their maps were generated using the same machine learning approaches and shared many covariates, such as climate data from WorldClim and vegetation indexes (Moreno-Martínez et al., 2018).

The variation in wood density derived from the four machine learning models, categorized by leaf types, leaf habit types, as well as climatic biomes, aligns with the features observed in wood density measurements (Figure 5a). Nonetheless, our machine learning models have limited capability in predicting the variations of wood density within leaf (habit) types and climatic biomes. This limitation arises from the fact that wood density measurements are tree-level estimations, which exhibit more significant variations within the category compared to the pixel-level predictions generated by the models, where the inter-tree variations within a pixel are not considered. In the context of different leaf types and leaf habit types, wood density tends to be higher for broadleaf trees compared to needleleaf trees. The

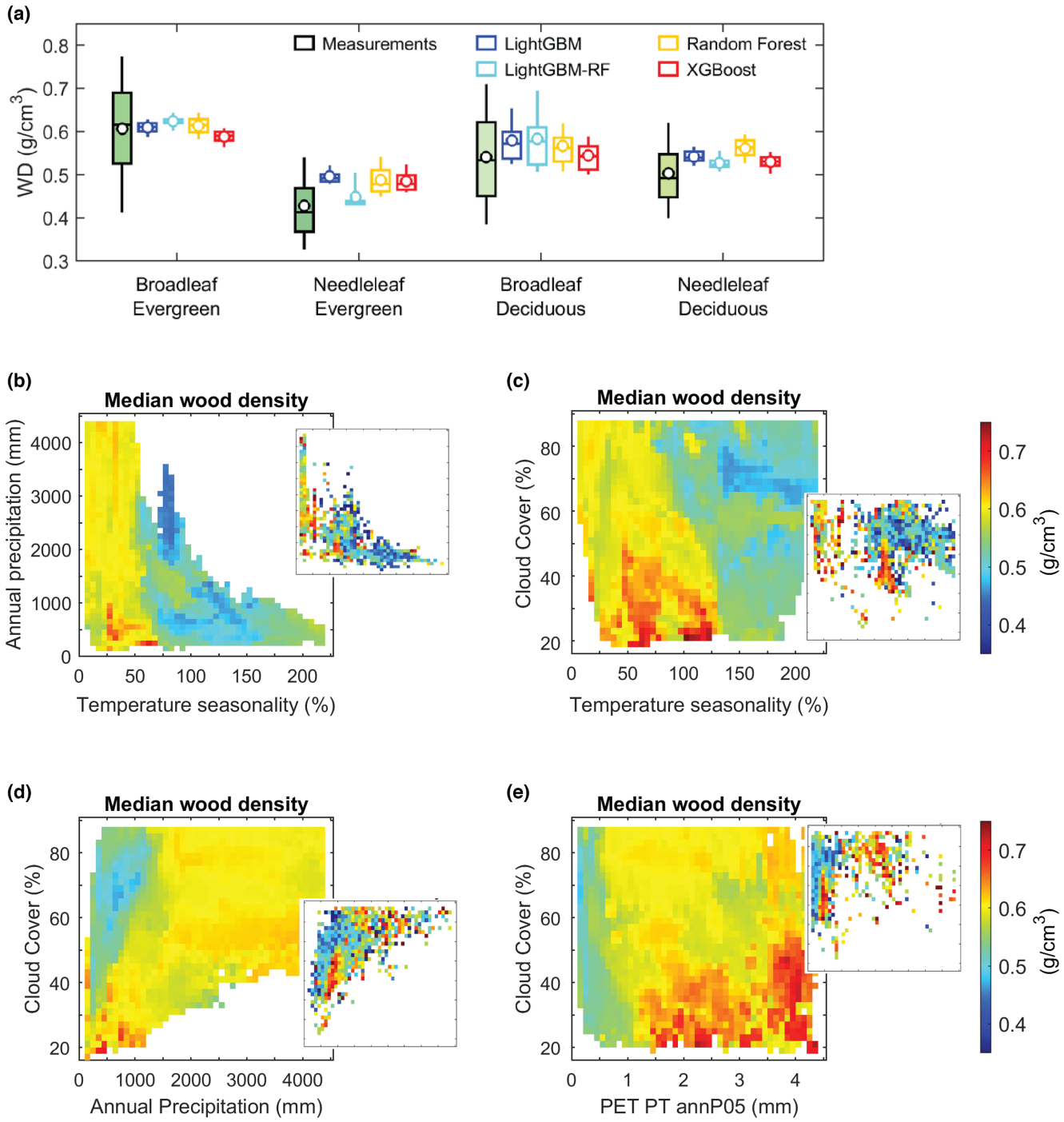

FIGURE 5 (a) The boxplots show the distribution of wood density for different categories of leaf types and leaf habit types. Both wood density measurement (filled boxes) and our estimates derived from four machine learning models (transparent boxes) are shown. In the plots, the white dot represents the mean value, and the lines outside and inside the boxes represent, from top to bottom, 90th, 75th, 50th, 25th, and 10th percentiles. (b–e) The distributions of predicted wood density by machine learning models with the climate, hydrological, and radiation spaces defined by temperature seasonality, total annual precipitation, precipitation of coldest quarter, potential evapotranspiration, and cloud coverage. The predicted wood density map at 1 km pixels used here corresponds to a forest fraction threshold of 10%. The insets in the corners show the values of wood density measurements over the corresponding climate spaces.

leaf types can be used to broadly categorize trees into hardwoods and softwoods, representing the difference in wood structure, specifically characterized by the presence of vessels and fibers in hardwoods, and predominantly tracheids in softwoods (Barnett & Jeronimidis, 2003; Pallardy, 2010). Among broadleaf trees, deciduous species exhibit slightly lower wood density than evergreen species. Conversely, among needleleaf trees, deciduous trees of gymnosperm tend to have higher wood density than evergreen

species. Regarding different climatic biomes (Figure S4), trees within tropical biomes exhibit higher wood density values compared to those in temperate biomes. In turn, trees in temperate biomes generally have higher wood density than those in boreal and polar biomes. The high wood density in tropical regions reflects the predominance of shade-tolerant species characterized by a slow growth rate replacing the fast-growing, gap-dependent pioneer species (Finegan, 1996; Martínez-Cabrera et al., 2012; Poorter et al., 2008).

The influences of climate variables such as temperature seasonality, annual precipitation, cloud cover, and potential evapotranspiration (PET) on wood density in our machine learning models are further explored (Figure 5b,e). These climate variables hold high importance in our machine learning models and exhibit relatively low collinearity among them. First of all, the key characteristics of the relationship between wood density and these climate covariates in the predicted map exhibit a good agreement with that in the observed wood density measurements. More specifically, we observed that temperature seasonality, defined as the coefficient of variation in monthly mean temperatures, serves to differentiate regions with low and high wood density, but has no monotonic relationship with wood density (Figure 5b,c). Furthermore, we found that wood density spatially varies with aridity degree, including total annual precipitation and PET. Specifically, we observed that under drier conditions, characterized by lower precipitation and higher PET, trees tend to have higher wood density values. This coincides with our previous results showing the prevalence of high wood density in forests situated within tropical dry and arid zones (Figure 3). Additionally, in the arid regions (annual precipitation <1000 mm), the influence of cloud cover on wood density is pronounced, that is, lower cloud coverage fraction, which may represent higher shortwave radiation availability, is associated with forests with higher wood density (Figures 3e and 5d).

Regarding the climate factors influencing wood density, previous studies have predominantly focused on the impacts of temperature along with one of precipitation, soil moisture or PET. But the impact of cloud coverage or radiation on wood density, which plays an important role in our machine learning models, has been largely neglected in earlier research. Furthermore, earlier studies have reported the dominant climate factor influencing wood density changes over region, for example, Wiemann and Williamson (2002) suggested that, within tropical regions, precipitation was the primary driver of wood density variations, while in extratropical regions, temperature played a key role. However, their investigations were limited to linear relationships between wood density and climate variables over very large regions such as the whole North and South Americas, or mid and high northern latitudes (Chave et al., 2009; Ogle et al., 2014; Roderick & Berry, 2001). In contrast, in our results, the climate covariates and wood density show complex nonlinear relationships over the global scale. Our findings do not exhibit clear, monotonous linear relationships within the observed and predicted patterns. It is important to note that our analysis is based on pixel-level wood density values in which the characteristics of both gymnosperms and angiosperms are combined. This differs from earlier studies that reported the significant climate role specifically for gymnosperms (Clough et al., 2017). In summary, our study sheds light on the multifaceted and nonlinear relationships between wood density and climate covariates, challenging previous linear assumptions.

4 | CONCLUSIONS

We provide the first spatially continuous map of wood density at a global scale at the 0.01° (~1 km) spatial resolution, using a wide set of climate, soils, topography, and vegetation properties applied to four machine learning models. The four machine learning models compared in our analyses all show good capacity to predict tree-level wood density estimates ($R^2 = 0.55 \pm 0.03$ and RMSE $= 0.11 \pm 0.01$ g/cm^3). Nevertheless, it is worth noting that variations and biases in model predictions tend to be more pronounced in regions where wood density measurements are scarce. To increase the reliability of our findings, we use eight cross-validation strategies, which not only confirm the robustness of our results but also underscore the limitations of machine learning models when it comes to extrapolation. This highlights the critical need for widespread wood density measurements and emphasizes the importance of leveraging measurements with global coverage. The spatial patterns of wood density generated by machine learning models exhibit a remarkable alignment with the leaf type and leaf habit types and climate conditions, which reflects the wood structure and growing rates of trees. In the northern high latitudes where deciduous and evergreen needleleaf trees dominate, wood density values are generally lower compared to the evergreen trees in tropical regions. Within the tropics, wood density in arid areas tends to be relatively lower than in wet regions. It is imperative to acknowledge the models' limited capacity to predict tree-level wood density variations in the wet tropical regions. Due to the absence of species maps, we used leaf types and leaf habit types as predictors, offering partial insights into the wood density variations related to species. Nevertheless, this method may limit the models' capacity to accurately predict wood density variations in regions characterized by high biodiversity. To overcome this limitation, there is a compelling need for the development of high-resolution global species abundance maps in the future, by taking advantage of available open-data resources, including complex networks of plot data (e.g., sPlot data; Sabatini et al., 2021), national forest inventories data, and species occurrences dataset (e.g., GBIF). Nevertheless, our newly global spatial explicit dataset will allow us to provide more accurate estimations of vegetation carbon stocks and give valuable insights into how forests resist and recover from future environmental challenges. This provides a valuable opportunity to obtain a better understanding of ecosystem function and services in the face of future climate change.

AUTHOR CONTRIBUTIONS

Hui Yang: Data curation; formal analysis; methodology; visualization; writing – original draft; writing – review and editing. **Siyuan Wang:** Formal analysis; software; writing – review and editing. **Rackhun Son:** Data curation; formal analysis; methodology; writing – review and editing. **Hoontaek Lee:** Formal analysis; writing – original draft; writing – review and editing. **Vitus Benson:** Formal analysis; methodology; software; writing – review and editing. **Weijie Zhang:** Formal analysis; writing – review and editing. **Yahai Zhang:** Formal analysis. **Yuzhen Zhang:** Formal analysis. **Jens Kattge:** Data curation; writing – review and editing. **Gerhard Boenisch:** Data curation. **Dmitry Schepaschenko:** Data curation. **Zbigniew Karaszewski:** Data curation. **Krzysztof Stereńczak:** Data curation. **Álvaro Moreno-Martínez:** Data curation; writing – review and editing. **Cristina Nabais:** Data curation; writing – review and editing. **Philippe Birnbaum:** Data curation; writing – review and editing. **Ghislain Vieilledent:** Data curation; writing – review and editing. **Ulrich Weber:** Data curation; software. **Nuno Carvalhais:** Conceptualization; formal analysis; methodology; project administration; supervision; writing – review and editing.

ACKNOWLEDGMENTS

We acknowledge Thomas Ibanez for his invaluable assistance in collecting wood density measurements. H.Y. is supported by the Project Office BIOMASS (grant number 50EE1904) funded by the German Federal Ministry for Economic Affairs and Climate Action. R.S. was supported by ESM2025. D.S. was supported by the ESA IFBN (ESA Contract No. 4000114425/15/NL/FF/gp) and ESA FRM4BIOMASS (ESA Contract No. 4000142684/23/I-EF-bgh) projects. Á.M.-M. was supported by the European Research Council under the ERC-SyG-2019 USMILE project (grant agreement 855187). S.W. and W.Z. acknowledge support from the International Max Planck Research School for Biogeochemical Cycles (IMPRS-gBGC). N.C. acknowledges contribution by the GlobBiomass DUE Project (ESA Contract No. 4000113100/14/I-NB). The data from Poland used for analysis were collected under REMBIOFOR project entitled "Remote sensing-based assessment of woody biomass and carbon storage in forests", which was financially supported by the National Centre for Research and Development (Poland), under the BIOSTRATEG programme (Agreement No. BIOSTRATEG1/267755/4/NCBR/2015). Open Access funding enabled and organized by Projekt DEAL.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflicts of interest.

DATA AVAILABILITY STATEMENT

The 1 km global maps of wood density are made accessible to the public downloading: https://doi.org/10.5281/zenodo.10692059. Code availability: The code used for training machine learning models and generating global maps is shared and made available on GitLab: https://gitlab.gwdg.de/siyuan.wang/global-wood-density.git.

ORCID

Hui Yang https://orcid.org/0000-0001-6454-8954
Vitus Benson https://orcid.org/0000-0003-4760-5501
Jens Kattge https://orcid.org/0000-0002-1022-8469
Dmitry Schepaschenko https://orcid.org/0000-0002-7814-4990
Ghislain Vieilledent https://orcid.org/0000-0002-1685-4997
Nuno Carvalhais https://orcid.org/0000-0003-0465-1436

REFERENCES

Anderegg, W. R., Klein, T., Bartlett, M., Sack, L., Pellegrini, A. F., Choat, B., & Jansen, S. (2016). Meta-analysis reveals that hydraulic traits explain cross-species patterns of drought-induced tree mortality across the globe. *Proceedings of the National Academy of Sciences of the United States of America*, *113*(18), 5024–5029.

Barnett, J. R., & Jeronimidis, G. (Eds.). (2003). *Wood quality and its biological basis*. CRC Press.

Batjes, N. H., Ribeiro, E., & Van Oostrum, A. (2020). Standardised soil profile data to support global mapping and modelling (WoSIS snapshot 2019). *Earth System Science Data*, *12*(1), 299–320.

Besnard, S., Koirala, S., Santoro, M., Weber, U., Nelson, J., Gütter, J., Herault, B., Kassi, J., N'Guessan, A., Neigh, C., Poulter, B., Zhang, T., & Carvalhais, N. (2021). Mapping global forest age from forest inventories, biomass and climate data. *Earth System Science Data*, *13*(10), 4881–4896.

Bouriaud, O., Bréda, N., Le Moguedec, G., & Nepveu, G. (2004). Modelling variability of wood density in beech as affected by ring age, radial growth and climate. *Trees*, *18*(3), 264–276.

Bouriaud, O., Leban, J. M., Bert, D., & Deleuze, C. (2005). Intra-annual variations in climate influence growth and wood density of Norway spruce. *Tree Physiology*, *25*(6), 651–660.

Burt, A., Boni Vicari, M., Da Costa, A. C., Coughlin, I., Meir, P., Rowland, L., & Disney, M. (2021). New insights into large tropical tree mass and structure from direct harvest and terrestrial lidar. *Royal Society Open Science*, *8*(2), 201458.

Chao, K. J., Phillips, O. L., Gloor, E., Monteagudo, A., Torres-Lezama, A., & Martínez, R. V. (2008). Growth and wood density predict tree mortality in Amazon forests. *Journal of Ecology*, *96*(2), 281–292.

Chave, J., Coomes, D., Jansen, S., Lewis, S. L., Swenson, N. G., & Zanne, A. E. (2009). Towards a worldwide wood economics spectrum. *Ecology Letters*, *12*(4), 351–366.

Chave, J., Davies, S. J., Phillips, O. L., Lewis, S. L., Sist, P., Schepaschenko, D., Armston, J., Baker, T. R., Coomes, D., Disney, M., Duncanson, L., Hérault, B., Labrière, N., Meyer, V., Réjou-Méchain, M., Scipal, K., & Saatchi, S. (2019). Ground data are essential for biomass remote sensing missions. *Surveys in Geophysics*, *40*, 863–880.

Chen, T., & Guestrin, C. (2016). Xgboost: A scalable tree boosting system. In *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining* (pp. 785–794). Association for Computing Machinery.

Clough, B. J., Curzon, M. T., Domke, G. M., Russell, M. B., & Woodall, C. W. (2017). Climate-driven trends in stem wood density of tree species in the eastern United States: Ecological impact and implications for national forest carbon assessments. *Global Ecology and Biogeography*, *26*(10), 1153–1164.

Fick, S. E., & Hijmans, R. J. (2017). WorldClim 2: New 1-km spatial resolution climate surfaces for global land areas. *International Journal of Climatology*, *37*(12), 4302–4315.

Finegan, B. (1996). Pattern and process in neotropical secondary rain forests: The first 100 years of succession. *Trends in Ecology & Evolution*, *11*(3), 119–124.

Greenwood, S., Ruiz-Benito, P., Martínez-Vilalta, J., Lloret, F., Kitzberger, T., Allen, C. D., Fensham, R., Laughlin, D. C., Kattge, J., Bönisch, G., Kraft, N. J., & Jump, A. S. (2017). Tree mortality across biomes is

promoted by drought intensity, lower wood density and higher specific leaf area. *Ecology Letters*, *20*(4), 539–553.

Grinsztajn, L., Oyallon, E., & Varoquaux, G. (2022). Why do tree-based models still outperform deep learning on typical tabular data? *Advances in Neural Information Processing Systems*, *35*, 507–520.

Guo, W. Y., Serra-Diaz, J. M., Eiserhardt, W. L., Serra-Diaz, J. M., Eiserhardt, W. L., Maitner, B. S., Merow, C., Violle, C., Pound, M. J., Sun, M., Slik, F., Blach-Overgaard, A., Enquist, B. J., & Svenning, J. C. (2023). Climate change and land use threaten global hotspots of phylogenetic endemism for trees. *Nature Communications*, *14*, 6950. https://doi.org/10.1038/s41467-023-42671-y

Harris, N. L., Gibbs, D. A., Baccini, A., Birdsey, R. A., de Bruin, S., Farina, M., Fatoyinbo, L., Hansen, M. C., Herold, M., Houghton, R. A., Potapov, P. V., Suarez, D. R., Roman-Cuesta, R. M., Saatchi, S. S., Slay, C. M., Turubanova, S. A., & Tyukavina, A. (2021). Global maps of twenty-first century forest carbon fluxes. *Nature Climate Change*, *11*, 234–240.

Hoffmann, W. A., Marchin, R. M., Abit, P., & Lau, O. L. (2011). Hydraulic failure and tree dieback are associated with high wood density in a temperate forest under extreme drought. *Global Change Biology*, *17*(8), 2731–2742.

Huang, Y., Ciais, P., Santoro, M., Makowski, D., Chave, J., Schepaschenko, D., Abramoff, R. Z., Goll, D. S., Yang, H., Chen, Y., Wei, W., & Piao, S. (2021). A global map of root biomass across the world's forests. *Earth System Science Data*, *13*(9), 4263–4274.

Iturbide, M., Gutiérrez, J. M., Alves, L. M., Bedia, J., Cimadevilla, E., Cofiño, A. S., Di Luca, A., Faris, S. H., Gorodetskaya, I. V., Hauser, M., Herrera, S., Hennessy, K., Hewitt, H. T., Jones, R. G., Krakovska, S., Manzanas, R., Martínez-Castro, D., Narisma, G. T., … Vera, C. S. (2020). An update of IPCC climate reference regions for subcontinental analysis of climate model data: Definition and aggregated datasets. *Earth System Science Data*, *12*, 2959–2970.

Jung, M., Schwalm, C., Migliavacca, M., Walther, S., Camps-Valls, G., Koirala, S., Anthoni, P., Besnard, S., Bodesheim, P., Carvalhais, N., Chevallier, F., Gans, F., Goll, D. S., Haverd, V., Köhler, P., Ichii, K., Jain, A. K., Liu, J., Lombardozzi, D., … Reichstein, M. (2020). Scaling carbon fluxes from eddy covariance sites to globe: Synthesis and evaluation of the FLUXCOM approach. *Biogeosciences*, *17*(5), 1343–1365.

Kattge, J., Bönisch, G., Díaz, S., Lavorel, S., Prentice, I. C., Leadley, P., Tautenhahn, S., Werner, G. D. A., Aakala, T., Abedi, M., Acosta, A. T. R., Adamidis, G. C., Adamson, K., Aiba, M., Albert, C. H., Alcántara, J. M., Alcázar, C. C., Aleixo, I., Ali, H., … Wirth, C. (2020). TRY plant trait database–enhanced coverage and open access. *Global Change Biology*, *26*(1), 119–188.

Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., Ye, Q., & Liu, T. Y. (2017). Lightgbm: A highly efficient gradient boosting decision tree. In I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, & R. Garnett (Eds.), *Advances in neural information processing systems* (pp. 3146–3154). Curran Associates.

King, D. A., Davies, S. J., Supardi, M. N., & Tan, S. (2005). Tree growth is related to light interception and wood density in two mixed dipterocarp forests of Malaysia. *Functional Ecology*, *19*(3), 445–453.

Liang, X., Ye, Q., Liu, H., & Brodribb, T. J. (2021). Wood density predicts mortality threshold for diverse trees. *The New Phytologist*, *229*(6), 3053–3057.

Ludwig, M., Moreno-Martinez, A., Hölzel, N., Pebesma, E., & Meyer, H. (2023). Assessing and improving the transferability of current global spatial prediction models. *Global Ecology and Biogeography*, *32*(3), 356–368.

Martínez-Cabrera, H. I., Estrada-Ruiz, E., Castañeda-Posadas, C., & Woodcock, D. (2012). Wood specific gravity estimation based on wood anatomical traits: Inference of key ecological characteristics in fossil assemblages. *Review of Palaeobotany and Palynology*, *187*, 1–10.

Moreno-Martínez, Á., Camps-Valls, G., Kattge, J., Robinson, N., Reichstein, M., van Bodegom, P., Kramer, K., Cornelissen, J. H. C., Reich, P., Bahn, M., Niinemets, Ü., Peñuelas, J., Craine, J. M., Cerabolini, B. E. L., Minden, V., Laughlin, D. C., Sack, L., Allred, B., Baraloto, C., … Running, S. W. (2018). A methodology to derive global maps of leaf traits using remote sensing and climate data. *Remote Sensing of Environment*, *218*, 69–88.

Nabais, C., Hansen, J. K., David-Schwartz, R., Klisz, M., López, R., & Rozenberg, P. (2018). The effect of climate on wood density: What provenance trials tell us? *Forest Ecology and Management*, *408*, 148–156.

O'Brien, M. J., Engelbrecht, B. M., Joswig, J., Pereyra, G., Schuldt, B., Jansen, S., Kattge, J., Landhäusser, S. M., Levick, S. R., Preisler, Y., Väänänen, P., & Macinnis-Ng, C. (2017). A synthesis of tree functional traits related to drought-induced mortality in forests across climatic zones. *Journal of Applied Ecology*, *54*(6), 1669–1686.

Ogle, K., Pathikonda, S., Sartor, K., Lichstein, J. W., Osnas, J. L., & Pacala, S. W. (2014). A model-based meta-analysis for estimating species-specific wood density and identifying potential sources of variation. *Journal of Ecology*, *102*(1), 194–208.

Oliveira, G. M. V., de Mello, J. M., de Mello, C. R., Scolforo, J. R. S., Miguel, E. P., & Monteiro, T. C. (2021). Behavior of wood basic density according to environmental variables. *Journal of Forest Research*, *33*, 497–505.

Pallardy, S. G. (2010). *Physiology of woody plants*. Academic Press.

Pan, Y., Birdsey, R. A., Fang, J., Houghton, R., Kauppi, P. E., Kurz, W. A., Phillips, O. L., Shvidenko, A., Lewis, S. L., Canadell, J. G., Ciais, P., Jackson, R. B., Pacala, S. W., McGuire, A., Piao, S., Rautiainen, A., Sitch, S., & Hayes, D. (2011). A large and persistent carbon sink in the world's forests. *Science*, *333*(6045), 988–993.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., … Duchesnay, É. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, *12*, 2825–2830.

Ploton, P., Mortier, F., Réjou-Méchain, M., Barbier, N., Picard, N., Rossi, V., Dormann, C., Cornu, G., Viennois, G., Bayol, N., Lyapustin, A., Gourlet-Fleury, S., & Pélissier, R. (2020). Spatial validation reveals poor predictive performance of large-scale ecological mapping models. *Nature Communications*, *11*(1), 4540.

Poorter, L., Wright, S. J., Paz, H., Ackerly, D. D., Condit, R., Ibarra-Manríquez, G., Harms, K. E., Licona, J. C., Martínez-Ramos, M., Mazer, S. J., Muller-Landau, H. C., Peña-Claros, M., Webb, C. O., & Wright, I. J. (2008). Are functional traits good predictors of demographic rates? Evidence from five neotropical forests. *Ecology*, *89*(7), 1908–1920.

Potapov, P., Li, X., Hernandez-Serna, A., Tyukavina, A., Hansen, M. C., Kommareddy, A., Pickens, A., Turubanova, S., Tang, H., Silva, C. E., Armston, J., Dubayah, R., Blair, J. B., & Hofton, M. (2021). Mapping global forest canopy height through integration of GEDI and Landsat data. *Remote Sensing of Environment*, *253*, 112165.

Poyatos, R., Sus, O., Badiella, L., Mencuccini, M., & Martínez-Vilalta, J. (2018). Gap-filling a spatially explicit plant trait database: Comparing imputation methods and different levels of environmental information. *Biogeosciences*, *15*(9), 2601–2617.

Roderick, M. L., & Berry, S. L. (2001). Linking wood density with tree growth and environment: A theoretical analysis based on the motion of water. *The New Phytologist*, *149*, 473–485.

Sabatini, F. M., Lenoir, J., Hattab, T., Arnst, E. A., Chytrý, M., Dengler, J., De Ruffray, P., Hennekens, S. M., Jandt, U., Jansen, F., Jiménez-Alfaro, B., Kattge, J., Levesley, A., Pillar, V. D., Purschke, O., Sandel, B., Sultana, F., Aavik, T., Ačić, S., … Bruelheide, H. (2021). sPlotOpen—An environmentally balanced, open-access, global dataset of vegetation plots. *Global Ecology and Biogeography*, *30*(9), 1740–1764.

Sato, H., Kobayashi, H., Iwahana, G., & Ohta, T. (2016). Endurance of larch forest ecosystems in eastern Siberia under warming trends. *Ecology and Evolution*, *6*(16), 5690–5704.

Schepaschenko, D., Shvidenko, A., Usoltsev, V., Lakyda, P., Luo, Y., Vasylyshyn, R., Lakyda, I., Myklush, Y., See, L., McCallum, I., Fritz, S.,

Kraxner, F., & Obersteiner, M. (2017). A dataset of forest biomass structure for Eurasia. *Scientific Data*, *4*, 170070.

Serra-Maluquer, X., Gazol, A., Anderegg, W. R., Martínez-Vilalta, J., Mencuccini, M., & Camarero, J. J. (2022). Wood density and hydraulic traits influence species' growth response to drought across biomes. *Global Change Biology*, *28*(12), 3871–3882.

Song, X. P., Hansen, M. C., Stehman, S. V., Potapov, P. V., Tyukavina, A., Vermote, E. F., & Townshend, J. R. (2018). Global land change from 1982 to 2016. *Nature*, *560*(7720), 639–643.

Sweet, L. B., Müller, C., Anand, M., & Zscheischler, J. (2023). Cross-validation strategy impacts the performance and interpretation of machine learning models. *Artificial Intelligence for the Earth Systems*, *2*(4), e230026.

Thurner, M., Beer, C., Santoro, M., Carvalhais, N., Wutzler, T., Schepaschenko, D., Shvidenko, A., Kompter, E., Ahrens, B., Levick, S. R., & Schmullius, C. (2014). Carbon stock and density of northern boreal and temperate forests. *Global Ecology and Biogeography*, *23*(3), 297–310.

Wiemann, M. C., & Williamson, G. B. (2002). Geographic variation in wood specific gravity: Effects of latitude, temperature, and precipitation. *Wood and Fiber Science*, *34*(1), 96–107.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Yang, H., Wang, S., Son, R., Lee, H., Benson, V., Zhang, W., Zhang, Y., Zhang, Y., Kattge, J., Boenisch, G., Schepaschenko, D., Karaszewski, Z., Stereńczak, K., Moreno-Martínez, Á., Nabais, C., Birnbaum, P., Vieilledent, G., Weber, U., & Carvalhais, N. (2024). Global patterns of tree wood density. *Global Change Biology*, *30*, e17224. https://doi.org/10.1111/gcb.17224