### nature plants

Article

# Boreal tree species diversity increases with global warming but is reversed by extremes

Received: 19 January 2024

Accepted: 21 August 2024

Published online: 11 September 2024

Check for updates

Yanbiao Xi<sup>1,2</sup>, Wenmin Zhang  $\mathbb{O}^2 \boxtimes$ , Fangli Wei<sup>3</sup>, Zhongxiang Fang<sup>4</sup> & Rasmus Fensholt  $\mathbb{O}^2$ 

Tree species diversity is essential to sustaining stable forest ecosystem functioning. However, it remains unclear how boreal tree species diversity has changed in response to climate change and how it is associated with productivity and the temporal stability of boreal forest ecosystems. By combining 5,312 field observations and 55,560 Landsat images, here we develop a framework to estimate boreal tree species diversity (represented by the Shannon diversity index, H') for the years 2000, 2010 and 2020. We document an average increase in H' by 12% from 2000 to 2020 across the boreal forests. This increase accounts for 53% of all boreal forest areas and mainly occurs in the eastern forest-boreal transition region, the Okhotsk-Manchurian taiga and the Scandinavian-Russian taiga. Tree species diversity responds positively to increasing temperatures, but the relationship is weakened for higher temperature changes, and in areas of extreme warming (>0.065  $^{\circ}$ C yr<sup>-1</sup>), a negative impact on tree species diversity is found. We further show that the observed spatiotemporal increase in diversity is significantly associated with increased productivity and temporal stability of boreal forest biomass. Our results highlight climate-warming-driven increases in boreal tree species diversity that positively affect boreal ecosystem functioning but are countered in areas of extreme warming.

Global forests have been documented to represent a consistent carbon sink<sup>1</sup> that is essential to mitigating climate warming<sup>2</sup>. The diversity of tree species in forests has been found to play a pivotal role in maintaining the functionality of forest ecosystems as a vital carbon sink on the basis of plot observations<sup>3</sup>, further sustaining ecosystem stability<sup>4.5</sup>. The global forest extent has experienced gains and losses over the past three decades<sup>6</sup>, owing primarily to forestry, agriculture and wildfires<sup>7</sup>, while climate change and rising CO<sub>2</sub> have caused continental-specific trade-offs in forest dynamics (for example, growth or mortality)<sup>8</sup> and have been reported to cause shifts in trees species at the plot level<sup>9.10</sup>. Yet, the response of tree species diversity to changes in forests over the past few decades remains unknown at a large spatial scale. Furthermore, the responses of forest diversity to ecosystem productivity and stability have been documented on the basis of plot measurements<sup>4,11</sup>. These relationships established at the plot level may vary depending on regional climate and soil conditions<sup>12,13</sup>, and therefore cannot easily be used to generalize across a range of diverse environmental conditions. The lack of large-scale spatial continuous mapping of tree species diversity, and the temporal dynamics thereof, impedes further exploration of how tree species diversity contributes to various aspects of forest ecosystem functioning.

Existing studies have attempted to estimate global forest tree species diversity<sup>14</sup> and vascular plant diversity on the basis of statistical relationships between species diversity metrics and environmental

<sup>1</sup>State Key Laboratory of Black Soils Conservation and Utilization, Northeast Institute of Geography and Agroecology, Chinese Academy of Sciences, Changchun, China. <sup>2</sup>Department of Geosciences and Natural Resource Management, University of Copenhagen, Copenhagen, Denmark. <sup>3</sup>Key Laboratory of Ecosystem Network Observation and Modeling, Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing, China. <sup>4</sup>Institute of Geography, Augsburg University, Augsburg, Germany. ⊠e-mail: wenminzhg@gmail.com





variables<sup>15</sup>, but these diversity maps represent static maps provided in coarse spatial resolution defined by the environmental input data. Satellite-based tree species diversity estimates have been conducted on the basis of spectral heterogeneity<sup>16–19</sup> and have produced reasonable predictions of tree species diversity, thereby offering improved spatial resolution and adding a temporal dimension to the mapping. Long-term satellite-based species diversity estimation would further make it possible to identify temporal changes in diversity across large spatial scales, which currently represents a major unknown. The combined impacts of climate change and human appropriation have profoundly transformed global forest ecosystems in recent decades<sup>20</sup>, highlighting the urgent need for continuous, time-varying observations.

Boreal forests represent 30% of the Earth's total forested area<sup>21</sup>, enduring harsh winters with freezing temperatures for approximately six to eight months and snow cover persisting for several months<sup>10,22</sup>. Boreal forests are further characterized by a much lower diversity of tree species than tropical forests<sup>21,23</sup>, and even minor positive or negative changes in species diversity may have substantial effects on the boreal ecosystems' carbon uptake and stability<sup>24</sup>. For example, Larix spp. (the dominant species across Eurasian boreal forests) are approaching a thermal tipping point at their southern margin, which is expected to cause an abrupt ecological collapse of ecosystem functioning (for example, the capacity of carbon uptake) of *Larix* spp. under continued climate warming<sup>25</sup>. These unique characteristics of boreal forest ecosystems make them particularly susceptible to climate change, and boreal forests have therefore been identified as a critical 'tipping element' of the Earth's climate system<sup>26</sup>. Profound climate warming and increasing frequency of fire activity have been observed in boreal ecosystems<sup>27,28</sup>, and the associated droughts have been documented to trigger widespread increases in tree mortality and decreases in the biomass carbon sink<sup>29-31</sup>. Yet, the extent to which boreal forest tree species diversity has undergone changes and the potential impact on ecosystem functioning remain largely unexplored to date. This knowledge would make a notable contribution to the assessment of carbon sequestration sustainability in boreal forests<sup>32</sup> and boreal forest ecosystem transitions<sup>9,33</sup>.

In this study, we aim to quantify changes in boreal tree species diversity with climate change over recent decades and the impacts of species diversity change on the boreal ecosystem carbon cycle. We developed a framework to produce spatially continuous representations of forest



Project Phase 6 (ref. 87). The shaded areas represent 95% confidence intervals. **b**, Tree species diversity, represented by H', estimated using Landsat satellite data in 2000, 2010 and 2020. The boxes and whiskers show the 5th, 25th, median, 75th and 95th percentiles of the H' values, while the black dots show the mean of the H' values for each year (n = 128, 715, 043).

tree species diversity quantified by a Shannon diversity index  $(H')^{34,35}$ accounting for both species richness and evenness across the boreal forest ecosystems (Supplementary Fig. 1). This is achieved by deploying machine learning based on a combination of field and Landsat satellite observations for the years 2000, 2010 and 2020. We used 5,312 field observations including observations of 190,516 trees and a deep learning approach based on the InceptionTime architecture to train a predictive model. Subsequently, 55,560 Landsat scenes were subjected to the model to upscale tree species diversity across boreal forest ecosystems. Next, we analysed the spatiotemporal dynamics of boreal tree species diversity in relation to a comprehensive set of environmental factors potentially impacting diversity (including climate, population density, fire activity and soil conditions). Finally, we analysed the spatiotemporal changes in boreal tree species diversity associated with the productivity and temporal stability of boreal forest biomass. This analysis is particularly timely in the boreal areas given the observed distinct climate warming and increase in wildfire activity in this region. and it further contributes to generalizing our current understanding of the associations between tree diversity and ecosystem functions.

#### Results

#### Spatiotemporal changes in tree species diversity

The boreal biome has experienced a dramatic increase in air temperature over the past few decades, compared with other regions across the globe, with continuous warming by 2100 being projected in future climate scenarios (Fig. 1a and Supplementary Fig. 2). This warming is expected to profoundly alter the structure and functioning of boreal forest ecosystems. We used Landsat satellite data, trained by plot-based measurements of tree species and a deep learning approach based on the Inception Time architecture (Methods, Extended Data Fig. 1 and Supplementary Figs. 3 and 4), to predict tree species diversity as described by  $H'^{34,35}$ , with a high accuracy (coefficient of determination ( $R^2$ ), 0.77; root mean square error, 0.12) for the entire boreal forest area (Extended Data Fig. 2). We observed a 12% increase in boreal tree species diversity by H' values from 2000 to 2020, with average H' values increasing from  $0.41 \pm 0.14$  (mean  $\pm$  standard deviation) to  $0.46 \pm 0.16$  (Fig. 1b), representing a 5%  $\pm$  2% increase during 2000–2010 and a 7%  $\pm$  3% increase during 2010-2020.

Large spatial variations in tree species diversity were observed across boreal forest ecosystems (Fig. 2a-c; zoomed-in figures are provided in Extended Data Fig. 3 for improved visual interpretation).

#### Article



**Fig. 2** | **Spatiotemporal changes in boreal tree species diversity. a**-**c**, Spatial patterns of tree species diversity by *H'* values across boreal forests in 2000 (**a**), 2010 (**b**) and 2020 (**c**). **d**, Changes in *H'* between 2000 and 2020 superimposed on mean seasonal air temperature. *H'* is aggregated at a spatial resolution of  $1.5^{\circ} \times 1.5^{\circ}$  and is shown by dots, with larger dots indicating higher diversity gains (blue) or losses (purple). Basemap in **a**-**d** from Natural Earth

(https://www.naturalearthdata.com/). **e**, Frequency plot of boreal tree species diversity in 2000, 2010 and 2020; the vertical lines denote the mean *H'* values of 0.41 in 2000, 0.43 in 2010 and 0.46 in 2020. **f**, Frequency plot of gains and losses in tree species diversity during 2000–2020; the vertical dashed line represents no change in the *H'* value.

In Eurasia, relatively high tree species diversity was observed over the Okhotsk–Manchurian taiga and Scandinavian–Russian taiga, located in northern Scandinavia as well as the northwestern and central regions of Russia, with average *H'* values of  $0.56 \pm 0.16$  and  $0.55 \pm 0.18$ , respectively, whereas a lower diversity ( $0.28 \pm 0.07$ ) was observed in the Northeast Siberian taiga located at the northern edge of Eurasia (Fig. 2a–c and Supplementary Table 1). In North America, the eastern forest–boreal transition region, characterized by boreal and temperate tree species, had the highest diversity ( $0.77 \pm 0.20$ ), followed by Central Rockies forests ( $0.49 \pm 0.16$ ) and Mid-Continental Canadian forests ( $0.44 \pm 0.12$ ). The Alaska–Yukon lowland taiga and Northern Canadian Shield taiga, dominated by spruce species, had the lowest diversity, with average *H'* values of  $0.33 \pm 0.09$  and  $0.31 \pm 0.08$ , respectively.

The extent of tree species diversity gains (calculated from the differences in *H'* values between 2020 and 2001; Methods) accounts for 53% of boreal forest areas (approximately 8,165,000 km<sup>2</sup>), and losses accounted for 17% (approximately 2,684,000 km<sup>2</sup>), while areas of no distinct change (determined by changes in *H'* values ranging from –0.01 to 0.01) were observed for 20% of all boreal forest areas (Supplementary Table 2). Diversity gains in boreal forest were primarily observed in the Scandinavian–Russian taiga (Fig. 2d and Supplementary Table 3), with *H'* values increasing by 35% ± 16% (0.18 ± 0.08); the Okhotsk–Manchurian taiga (33% ± 17%, 0.17 ± 0.09); and the eastern forest–boreal transition region (27% ± 11%, 0.19 ± 0.08). Diversity losses occurred mainly in the Kamchatka Mountain forest and the West Siberian taiga, with *H'* values decreasing by 25% ± 16% (-0.13 ± 0.08) and 26% ± 10%

 $(-0.13 \pm 0.05)$ , respectively. Areas of gains in diversity were mainly observed in warmer regions, while losses tended to occur in colder regions (Fig. 2d). Moreover, a clear reduction in the proportions of relatively low diversity values (approximately H' < 0.4) was observed, accompanied by an increase in the proportions of relatively high diversity values (H' > 0.4) at the pixel level for the three epochs from 2000 to 2020 (Fig. 2e); also, a higher frequency of pixels of diversity gains than losses was observed (a gain-to-loss ratio of 3:1) (Fig. 2f).

We further assessed the relative contributions of changes in species richness and evenness to diversity dynamics using a multiple linear regression model, on the basis of forest inventory data with repeated measurements (n = 648; most of the plots are located in North America) (Extended Data Fig. 4). The results show that in the eastern forest-boreal transition region and Canadian Shield forests, species richness and evenness contribute almost equally to the changes in tree species diversity (H' values), while the contribution of evenness ( $\beta = 0.47$ ;  $\beta$  represents the sensitivity of change in diversity against changes in the explanatory variable) significantly exceeds that of richness ( $\beta = 0.21$ ) in the Northern Canadian Shield taiga and the Central Rockies forests (0.56 versus 0.28).

#### Determinants of tree species diversity changes

We quantified the spatial variability in boreal tree species diversity for the three epochs driven by a range of potential environmental variables accounting for geographic variations in climate, soil properties and disturbances (fire activity and human population density) (Supplementary



**Fig. 3** | **Environmental determinants of observed spatial variability in boreal tree species diversity. a**, The relative importance of predictor variables controlling the spatial variability of boreal tree species diversity, determined by a BRT model. The model was run ten times to avoid stochastic errors. The coloured circles represent the relative importance of each predictor variable for each run, the bars are the mean relative importance and the error bars are one standard deviation of the mean. TMP, mean seasonal temperature; PR, mean seasonal precipitation; POP, human population density; OC, topsoil organic carbon;

SAND, topsoil sand fraction; CEC, cation exchange capacity; FIRE, fire activity frequency; DEM, digital elevation model. **b**–**e**, Partial dependency plots of the top four variables explaining variability in boreal tree species diversity for 2020. The blue lines are smoothed representations of the responses, with fitted values (model predictions based on the original data), and the shaded areas represent 95% confidence intervals. The distributions of the predictor observations are indicated by the density of the vertical grey lines above the *x* axis.

Fig. 5) using a boosted regression tree (BRT) algorithm (Methods). We show that BRT can explain, on average, 62% of the spatial variability in boreal tree species diversity (Fig. 3a). We found mean seasonal temperature to be the most important predictor (53%), followed by mean seasonal precipitation (21%), while each of the remaining individual variables contributed less than 10% to the variability in diversity, including elevation (9%), human population density (6%), fire activity (4%), cation exchange capacity (3%), topsoil organic carbon (2%) and topsoil sand fraction (2%). When investigating the diversity response to an individual variable independent of other variables (partial dependency plots; Methods), we found that temperature shows a strong positive impact on diversity when mean seasonal temperature is below 12 °C (Supplementary Fig. 6), while the impact appears to reach a plateau for values above this threshold (Fig. 3b). Precipitation shows a clear positive effect with mean seasonal precipitation above 100 mm that tends to saturate after reaching 400 mm (Fig. 3c). Elevation generally shows a weak positive impact on diversity until approximately 700 m, beyond which elevation shows a weak negative impact (Fig. 3d). Human population density shows a weak positive effect on diversity with population density increasing to 1.5 people per km<sup>2</sup>, from where population density exerts a slightly negative effect on diversity (Fig. 3e). Similar patterns of the partial responses were observed for each individual period (Extended Data Fig. 5 and Supplementary Figs. 7-9).

Mean seasonal temperature and precipitation exerted strong positive effects on the spatial variation of boreal tree species diversity, whereas temporal changes in diversity were observed to show a negative response to increasing temperature over the past 20 years ( $\rho = -0.51$ , P < 0.01) (Fig. 4a,d). While lower rates of increasing temperature are associated with increasing trends in diversity, this relationship

gradually changes towards higher rates of increasing temperature  $(exceeding 0.065 \,^{\circ}C \,yr^{-1})$  associated with decreasing trends in diversity. These areas of decreasing trends are primarily observed to be located in northeastern Siberia (Fig. 4a). Similarly, the trends in diversity show a negative response to increasing fire activity frequency (Fig. 4b.e). The diversity trend generally shows a positive response to precipitation in cases of minor positive and negative precipitation trends, whereas for more extreme precipitation trends, the positive diversity trends approach zero (Fig. 4c, f). Furthermore, a negative relationship between diversity and stand age was observed across boreal forests (Extended Data Fig. 6), with higher gains in diversity in young forests than in mature forests. Accounting for these variables together, our analysis reveals that temperature trends and stand age exert a greater relative influence on regulating changes in diversity than precipitation trends and the frequency of fire activity (Supplementary Table 4). When the effect of stand age is removed, temperature trends play a predominant role in controlling changes in diversity (Supplementary Table 4).

#### Association with carbon fluxes, stocks and stability

We quantified the associations between boreal tree species diversity (and spatiotemporal changes therein) and six independent indicators (and spatiotemporal changes therein) characterizing forest carbon, including carbon fluxes (net primary production (NPP), kernel normalized difference vegetation index (kNDVI) and vegetation optical depth climate archive Ku-band (VOD Ku-band)), carbon stocks (aboveground-biomass-based LiDAR and optical satellite data, and L-band passive microwave data (AGB\_1 and AGB\_2)) and the temporal stability of boreal forest biomass (Methods). Our results, based on a multiple linear regression, show significantly (*P* < 0.05) positive associations





dots indicate one standard deviation of tree species diversity trends at a spatial resolution of  $1.5^{\circ} \times 1.5^{\circ}$ , with larger dots indicating higher variations of species diversity. The solid lines are fitted by a linear or quadratic model, and the shaded ribbons indicate the 95% confidence intervals of the fitted lines. The rugs on the *x* axis show the distribution of dots, while the grey contour lines show the 50%, 75% and 95% quantiles of the occurrence probability of the dots.  $\rho$  indicates the Spearman correlation coefficient between detected variables based on the raw data. The two-sided Student's *t*-test is used for statistical testing, and *P* values are indicated.

between diversity and forest carbon fluxes and stocks across spatial and temporal scales (Fig. 5 and Supplementary Figs. 10–12).

The spatial variability of all carbon indicators was significantly associated with tree species diversity and stand age, as well as with climate, disturbances, soil properties and topography; climate variables had the highest impact on forest carbon, followed by diversity, disturbances and stand age (Fig. 5 and Supplementary Fig. 10). In the spatial domain, diversity generally showed a positive relationship with forest carbon stock and fluxes, while stand age showed a negative relationship with carbon fluxes and a positive relationship with carbon stock (Fig. 5). Trends in NPP, kNDVI and AGB\_2 were significantly positively correlated with diversity trends, while stand age showed significant negative relationships with trends for most carbon indicators (Fig. 5 and Supplementary Fig. 11). Temperature and precipitation changes showed significant positive relationships with trends in NPP and AGB\_1, whereas temperature and precipitation had varying relationships with other carbon indicators (Fig. 5 and Supplementary Fig. 11). Disturbances were generally negatively correlated with most carbon indicators, with the impact of fire activity being larger than that of human population density. Elevation showed a positive correlation with most carbon indicator trends, whereas soil characteristics were both positively and negatively correlated (Fig. 5 and Supplementary Fig. 12). The temporal stability of boreal forest biomass was also found to be significantly associated with diversity at both spatial and temporal scales, while being co-regulated by climate, fire activity and topography.

#### Discussion

We developed a data-driven tree species diversity assessment framework that utilized remote sensing data in combination with in situ observations to generate spatially continuous boreal tree species diversity maps with a high level of spatial detail ( $30 \text{ m} \times 30 \text{ m}$ ), but also with a temporal dimension covering three different epochs around 2000, 2010 and 2020. This approach thus provides distinct advantages over other global mapping methods for diversity assessment<sup>15,23</sup> and offers an unparalleled evaluation of the nature of spatiotemporal changes in boreal tree species diversity in response to global environmental changes at different scales of time and space. The success of the satellite-remote-sensing-based estimates of tree species diversity



Fig. 5 | Spatiotemporal changes in boreal tree species diversity associated with the forest carbon cycle. The colour scale indicates the strength of the relationship (standardized coefficients  $\beta$ ) between each predictor (bottom) and each response variable (left) (NPP, kNDVI, AGB and temporal stability of AGB) in a multiple linear regression model. The black dots indicate significant impacts at a 95% confidence level, and the hashed areas indicate no statistical relationships.

was mainly attributed to spectral heterogeneity (for example, plant chemical properties of the tissue related to photosynthetic pigments and water, branching structure, leaf size and colour, leaf clumping and leaf angle distribution) being sufficiently characterized by information from the visible, near-infrared and short-wave infrared regions<sup>16,36</sup>, while the InceptionTime deep learning approach used here captured well the complex relationships under the varying phases. Uncertainties in the prediction of diversity may still exist owing to data quality, environmental conditions (for example, lighting conditions and shadows) and the reflectance influenced by understory vegetation (that is, shrubs and grasses) in sparse forests<sup>18,37</sup>. However, our approach, including the use of segmentation and multiple vegetation indices in the model building, has largely reduced the direct effects of changes in greenness and tree cover on changes in H' (indicated by the smaller  $R^2$ ; Extended Data Fig. 7). Future work could nonetheless consider techniques such as spectral unmixing, radiative transfer model inversion and data fusion to further reduce these uncertainties<sup>38</sup>.

Not surprisingly, the spatial patterns of forest tree species diversity show signs of latitudinal dependency, with decreasing species diversity towards the tundra biome, largely related to the temperature gradient from higher to lower temperatures<sup>23</sup>. However, supported by the high spatial resolution of the satellite data, the *H'* maps (accounting for tree species richness and evenness) present distinctly spatial variations in diversity across areas at the same latitude, unlike previous studies of tree species richness generally displaying monotonic and homogeneous changes in diversity across boreal forest ecosystems<sup>15,23</sup>. This satellite-based spatial pattern of diversity may thus better reflect natural diversity changes in boreal forests.

We documented an overall increase in boreal tree species diversity during the period of analysis linked to climate warming, which is consistent with predictions and observations of other studies<sup>9,10,39</sup>. Warming of the climate fosters conditions conducive to the expansion Diversity, climate and disturbances have mean values and temporal trends, with  $\Delta$  indicating temporal trends of variables from 2000 to 2020. MODIS NPP, NPP from the MODIS MODI7A3HGF v.6.1 product; AGB\_1, L-VOD-based AGB from the Soil Moisture and Ocean Salinity (SMOS) satellite; AGB\_2, AGB from ref. 64; Stability, temporal stability of boreal forest biomass (calculated from AGB\_2 from 2000 to 2019); AGE, stand age.

of boreal forests and to the proliferation of species, achieved through mechanisms such as earlier snowmelt providing more time for seed germination<sup>40</sup>, sapling growth<sup>41</sup>, altered disturbance regimes<sup>42</sup> and the augmentation of soil nutrient availability<sup>9,10</sup>. Particularly, moderate disturbances could catalyse tree community responses to climate change<sup>42</sup>, potentially shifting forest composition (both species richness and evenness) towards warm-adapted species (Extended Data Fig. 4)-for example, a transition from coniferous species to mixed species (temperate and boreal), as documented in southern boreal forest areas<sup>43,44</sup>. However, we observe that the diversity increase is negatively responding to increasing temperatures, suggesting that warming only to a certain extent can promote boreal tree diversity. A rapid increase in temperature is possibly detrimental to boreal tree diversity, because these species cannot adapt to such abrupt changes in temperature<sup>8</sup>, whereas pioneer species may adapt to the changes more guickly, thereby encroaching on the habitat of other species and potentially limiting their space and resources<sup>43,45</sup>. Moreover, extreme warming is likely to surpass the thermal tolerance of trees and triggers wildfires, resulting in tree mortality and thereby a reduction in tree diversity<sup>46,47</sup>. Such effects could particularly be occurring in northern regions with lower tree species diversity and scarce environmental resources (for example, low soil nutrients and seed availability)<sup>12,40</sup>, hindering the recovery of tree species. Here we observe a warming rate exceeding 0.065 °C yr<sup>-1</sup> to have a negative impact on temporal changes in diversity. Such negative effects can also be enhanced due to the increased frequency of fire occurrences induced by rising temperatures<sup>48,49</sup>, as well as the co-occurrence with warming droughts<sup>50</sup> and an increasing vapour pressure deficit exerting a higher demand for water availability<sup>51</sup>.

Increasing boreal tree species diversity is further found to be associated with high carbon stocks and the stability of boreal ecosystems, and thus there is no potential conflict of interest between preserving boreal tree species diversity and emphasising the role of boreal forest ecosystems as an essential global source of carbon sequestration<sup>3</sup>. Additionally, changes in tree species diversity and the stability of forest ecosystems can also be largely regulated by tree stand age. Our results suggest that young forests have higher increases in diversity and carbon cycling dynamics than mature forests, which is probably because young forests are undergoing rapid compositional/structural transitions. Finally, we note that some uncertainty may exist regarding the observed relationships introduced by the different data sources on carbon fluxes, carbon stocks and diversity, despite the efforts made (for example, the use of different data sources) to reduce such potential bias or uncertainty.

Our results indicate that increasing temperatures across the boreal zone overall corresponded to an increase in tree species diversity over the past two decades, but these increasing trends of diversity were reversed in areas of extreme warming, providing new insights into the long-term temporal changes in diversity in response to climate change across boreal forest ecosystems. However, here we did not account for the response of possible changes in dominant tree species due to climate warming, which may contribute to the change in productivity and stability of boreal ecosystems<sup>52,53</sup>. Further studies are also needed on the responses of forest structural diversity<sup>54</sup> and functional trait diversity<sup>55</sup> to climate change and their associated impacts on ecosystem functions. Incorporating these variables would enhance our understanding of the underlying mechanisms governing the associations between forest tree diversity, productivity and the temporal stability of ecosystems. Additionally, increasing diversity is probably associated with the emergence of a biome shift52.56, with trees and shrubs expanding towards tundra regions<sup>56,57</sup>, reshaping the structure and functioning of boreal ecosystems. Taken together, such an expanded knowledge base would provide a stronger foundation for promoting solutions for sustainable ecosystem functioning and mitigating the risk of destabilizing the terrestrial carbon sink under climate warming.

#### Methods

#### National forest inventory data

In this study, we collected boreal tree species diversity datasets from six countries (Supplementary Table 5). The datasets include 5,312 field observations of a total of 190,516 trees divided into 254 tree species. Most of the datasets come from national forest inventory databases, including Canada and China, while the remaining data come from publicly available datasets covering the United States (Alaska)<sup>58</sup>, northeastern Siberia<sup>59</sup> and Northern Europe<sup>3,60</sup>. *H'* was calculated on the basis of the number and species of trees in each sample plot<sup>35,61</sup>. Considering differences in plot size across datasets, we normalized forest tree species diversity to a common basis of 900 m<sup>2</sup> in area and 10 cm in threshold diameter at breast height using plot area and diameter at breast height as predictor variables, following the approach adopted by ref. 23.

#### Landsat data

The Landsat surface reflectance data, with a spatial resolution of  $30 \text{ m} \times 30 \text{ m}$ , from the US Geological Survey Earth Resources Observation and Science archive were used to upscale boreal tree species diversity. We obtained three sets of Landsat data covering the growing season (May to October) for three time periods: 1999–2001, 2009-2011 and 2019-2021. We applied a three-year time interval for data compositions to increase the number of clear-sky satellite observations of the different time epochs. These data were obtained from Landsat-5 Thematic Mapper, Landsat-7 Enhanced Thematic Mapper Plus and Landsat-8 Operational Land Imager, and they were synthesized into monthly composites for the best-quality collection (that is, minimal cloud, fog and snow cover) of each period using published cross-calibration coefficients for surface reflectance (Supplementary Fig. 3).

#### carbon indicator data

Multiple and independent long-term datasets related to carbon flux and stock were obtained to assess the association of tree species diversity with boreal forest carbon indicators. Two datasets were used to calculate the mean and changes of boreal forest carbon fluxes: the latest optimized annual NPP from the MODIS MOD17A3HGF v.6.1 product and the kNDVI introduced by ref. 62. The MOD17A3HGF is generated on the basis of the radiation use efficiency model that takes photosynthetically active radiation, leaf area index, climate factor and biome parameter as input. The kNDVI has high correlations with plot-level measurements of primary productivity and satellite retrievals of solar-induced fluorescence, and has thus been proposed as a robust proxy for terrestrial carbon sink dynamics<sup>51</sup>. We calculated the kNDVI for boreal forests from 2000 to 2020 using MODIS reflectance bands, on the basis of a method proposed by ref. 62 (equation (1)). The formula is as follows:

kNDVI = 
$$tanh\left(\left(\frac{NIR - red}{2\sigma}\right)^2\right)$$
 (1)

where  $\sigma$  represents the sensitivity of the index to sparsely/densely vegetated regions; in this study,  $\sigma = 0.5$  (NIR + red).

In addition to carbon flux indicators, three products measuring carbon stocks were obtained: VOD Ku-band from the Vegetation Optical Depth Climate Archive63; AGB\_1, driven by the L-band microwave radiometer of SMOS missions; and AGB\_2, from integrated ground and airborne measurements and MODIS and PALSAR observations by ref. 64. The VOD Ku-band data (period 1987-2017) were generated with combinations of multiple sensors (SSM/I, TMI, AMSR-E, WindSat and AMSR2) using the Land Parameter Retrieval Model and are closely related to the density, biomass and water content of vegetation<sup>63</sup>. AGB 1 was derived from the SMOSL-VOD (vegetation optical depth of L-band microwave missions) ascending product in the IC v.1.05 (ref. 65). SMOS L-VOD was converted to carbon density using the previously published biomass map<sup>66</sup> as a reference by a linear regression with mean L-VOD. Xu et al.<sup>64</sup> mapped live biomass carbon stocks (used in this study as AGB 2) over all woody vegetation globally from 2000 to 2019 by using a large number of ground inventory plots, in combination with LiDAR data (ICESat) and optical and microwave satellite data (MODIS and PALSAR).

Finally, using time series of AGB\_2 data, we quantified the temporal stability of boreal forest biomass for the three epochs as the ratio of mean AGB to its temporal standard deviation over a five-year period (2000–2004, 2008–2012 and 2014–2019), as similarly done in several other studies<sup>67,68</sup>.

#### **Environmental data**

We collected a set of environmental variables to explore the underlying mechanism of the variations of boreal tree species diversity. These variables were grouped into four categories: climate (that is, precipitation and temperature), soil properties (that is, topsoil sand fraction, topsoil organic carbon and cation exchange capacity), disturbances (that is, fire activity and population density) and topography (that is, elevation) (Supplementary Table 6). Incoming solar radiation can also strongly impact diversity but was excluded due to its high collinearity with temperature for the boreal biome. Seasonal precipitation and temperature from 2000 to 2020 obtained from the ECMWF Reanalysis v.5 were aggregated for the vegetation growing seasons from May to October. Topsoil sand fraction, organic carbon and cation exchange capacity of the clay fraction were obtained from the Regridded Harmonized World Soil Database v.1.2. We quantified the fire frequency using the monthly MODIS burned area product (MCD64A1) by summing the number of fire occurrences for each 500 m pixel from 2001 to 2020. The global human population density, provided in 30-arcsecond (approximately 1 km) grid cells, was used here as an indicator of human disturbance impact on forest resources. These data were derived from the Gridded Population of the World Version 4 Revision 11, which holds the estimates

of population density for 2000, 2010 and 2020. The global stand age map was generated from ref. 69 using forest inventories and biomass and climate data, which were divided into several age classes (0-150+with a decadal interval) at a spatial resolution of 1 km. A digital elevation model with 30 m spatial resolution was acquired from the ASTER Global Elevation Model. These variables were resampled to a 500 m × 500 m spatial resolution using the bilinear interpolation method.

#### Shannon diversity index

In this study, we applied H' (equation (2))<sup>34,35,61</sup>, which represents alpha diversity, to characterize tree species diversity within each defined spatial unit in boreal forests. H' considers both tree species richness and evenness to provide a comprehensive measure of alpha diversity:

$$H' = -\sum_{i=1}^{S} p_i \ln p_i \tag{2}$$

where S is the total number of tree species in a plot, and  $p_i$  is the proportional abundance of species *i* relative to the total abundance of all species S in a plot.

#### Mapping boreal tree species diversity

Satellite-remote-sensing-based diversity estimates are based on the spectral variability hypothesis<sup>18,19</sup>, which relates the spectral heterogeneity determined by plant biochemical and morphological differences (including photosynthetic pigments, branching structure, leaf clumping and leaf angle distribution) to environmental heterogeneity. A higher spectral heterogeneity is thus expected to be associated with higher environmental heterogeneity that sustains more species, providing a proxy for species diversity<sup>36,70</sup>. To reduce the potential impact of other land-cover types on spectral heterogeneity, we masked out all non-forested areas where shrubs and herbaceous vegetation dominate (tree cover less than 30%) across the boreal forests.

Accordingly, we developed a workflow for mapping boreal tree species diversity using satellite remote sensing imagery, consisting of the following five steps: (1) object segmentation, (2) spectral metrics, (3) matching to in situ *H'*, (4) prediction of *H'* and (5) post-processing (Extended Data Fig. 1).

Object segmentation. To improve the categorization of morphologically similar species, we used object clustering, based on a simple non-iterative clustering algorithm. By grouping pixels on the basis of their spectral characteristics, shape, texture and spatial relationship with the surrounding pixels to accurately estimate spatial/spectral metrics, we better captured the spectral heterogeneity compared with an individual-pixel-level-based spectral characterization. The simple non-iterative clustering algorithm was performed by initializing cluster centres (called seeds) at regular grid points throughout the images of Landsat-based spectral bands and vegetation indices. Each pixel in the image was then assigned to the nearest cluster centre on the basis of both spatial distance and feature similarity. Various spacing distances (in pixels) between seeds (that is, 5, 10, 15, 20, 25, 30, 40 and 50) were tested to derive the optimal seed spacing based on the boundary recall<sup>71</sup>. After initial assignment, the cluster centres were updated to the mean position of all the pixels assigned to that cluster, so that the cluster centres better represent the pixels belonging to that cluster. The clusters thus represent forest segments (the smallest unit of a forest community), and the calculations of spectral metrics were performed within each segment, with representative spectral metrics obtained by averaging pixel-wise metrics per segment. This analysis was implemented in Google Earth Engine.

**Spectral metrics.** We obtained three classes of commonly used satellite-based spectral metrics<sup>17–19</sup>. First, we calculated the spectral heterogeneity metrics defined as the degree of spatial variations in

spectral reflectance-that is, the coefficient of variation, spectral angle mapper (spectral dilation and spectral gradient) and texture features (dissimilarity and entropy) (Supplementary Table 7). These metrics have been proposed on the basis of different mathematical principles (that is, variance, distance, angle and volume) and have proved effective in capturing the spectral heterogeneity of a given area. Second, we calculated spectral/temporal metrics and derived five statistical metrics: median, minimum, maximum, standard deviation and mean for each spectral band and vegetation index. Third, we used the original spectral bands of the Landsat imagery and calculated six vegetation indices for each month during the growing season (May to October). The spectral bands, vegetation indices and temporal metrics help differentiate between trees, shrubs and herbaceous vegetation, thereby reducing the impact of changes in tree cover on the prediction of H'. We thus derived a total of 217 spectral metrics including 85 spectral heterogeneity metrics, 60 spectral/temporal metrics (related to temporal variations), 36 spectral bands and 36 vegetation indices, which were ultimately calculated per segment and used as predictors for the modelling (Supplementary Fig. 3 and Supplementary Table 8).

**Matching to in situ** *H'*. We established the spatial matching between segment-based spectral metrics and in situ *H'* included in national forest inventory records across 5,312 sites to derive the training samples. To improve the model robustness, we also implemented data augmentation to increase the size of training samples. The augmentation was applied by calculating the satellite-based spectral metrics averaged over window sizes of  $1 \times 1$ ,  $3 \times 3$  and  $5 \times 5$  pixels centred on each plot location accounting for varying sizes in the segmented patches. This increases variations in the training data and allows control over the number of training samples, thereby improving the generalization performance of modelling. By removing missing values and outliers induced by clouds and shadows, we finally derived 20,100 samples (each with 217 metrics) paired with in situ *H'*, of which 70% were used for training the model, 10% were used for validation and 20% were used as the test dataset.

**Prediction of** *H'*. We applied the InceptionTime architecture deep learning approach to establish a predictive model with in situ *H'* as the response variable and 217 spectral metrics as predictors. InceptionTime has been extensively used for classification and regression<sup>72</sup>, because it allows the model to simultaneously analyse patterns exhibited at different convolutional scales and cope with time series with complex patterns and varying temporal frequencies. We fine-tuned InceptionTime, including deleting the maximum pooling layer, reducing inception modules, adding dropout layers and modifying residual connections to achieve the best architecture for the training data of this study (Supplementary Fig. 4). A grid search method combined with cross-validation was used to determine the optimal hyperparameters such as filters, kernel sizes, batch size and learning rate.

The performance of the predictive model was evaluated using a tenfold cross-validation method, which ensures that the validation set is independent and spans the entire range of the data. Mean values of  $R^2$  and the root mean square error over the ten iterations were computed to quantify the model performance. Finally, we established the best model for predicting H' using the optimal hyperparameters and the selected predictors. The resultant model was used to generate H' maps for the entire study area in 2000, 2010 and 2020.

**Post-processing.** The distortion of spectral reflectance caused by tree shadowing, cloud cover, topography, diverse understory vegetation and other factors may lead to larger variability or dispersion of data records around the mean. We thus used standard errors based on the inventory data to minimize the uncertainty of predicted *H'* (Supplementary Fig. 13). We calculated the standard deviation of *H'* using the forest inventory data with repeated measurements of *H'* during different

years in the same plot (n = 648 plots) and derived the 95th quantile value as the maximum standard deviation corresponding to 0.20 as the threshold applied to obtain H' with the lowest potential uncertainty (Supplementary Fig. 14). About 10% of the study area (approximately 1,531,000 km<sup>2</sup>) was marked as uncertain, probably due to noise and randomness in the data, and was excluded from further analysis.

#### Statistical analysis

We quantified the changes of tree species diversity in boreal forests by calculating the per-pixel difference in H' values between the 2000s and the 2020s. We defined distinct changes larger than 0.01 as a diversity gain and less than -0.01 as a diversity loss, while a change ranging between -0.01 and 0.01 was defined as no distinct change, according to the minimal units of changes in H' values and the dependency of the two parameters (that is, species richness and evenness)<sup>73,74</sup>. We used a BRT analysis to assess the relative impacts of the explanatory variables, including temperature, rainfall and soil properties, on the spatial distribution of boreal tree diversity. This method has been used extensively in ecological studies to study response variables<sup>75,76</sup>. BRT is an advanced machine learning algorithm that iteratively fits and combines multiple regression tree models to improve predictive performance<sup>76,77</sup>. We randomly selected 92,416 samples (pixels) with a minimum distance of 50 km between neighbouring samples to establish the relationships between the explanatory variables and the response variable (H') with the BRT model. The model was iterated ten times to avoid stochastic errors. Partial dependency plots resulting from the BRT analysis were derived to describe how the tree species diversity responds to change in each predictor independent of the other predictors.

We used a multiple linear regression model to explore the responses of spatiotemporal changes in carbon indicators to diversity as well as changing environmental variables. The carbon indicators were used as dependent variables, while *H'*, stand age and the other environmental variables were used as explanatory variables. To avoid bias introduced by spatial autocorrelation between neighbouring samples, we calculated Moran's index for each variable and used the maximum distance of 50 km for the random selections of sampling (Supplementary Table 9). All explanatory variables were standardized (with an average of 0 and a standard deviation of 1) to obtain standardized coefficients  $\beta$ . Significance tests were set at a 95% confidence level (*P* < 0.05). The analyses and graphs were performed using R v.4.2.0 (ref. 78) and with the following packages: caret<sup>79</sup>, gbm<sup>80</sup>, tidyverse<sup>81</sup>, Ime4 (ref. 82) and ggplot2 (ref. 83).

#### **Reporting summary**

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

#### Data availability

All data used to support the findings of this study are publicly available. The Landsat surface reflectance data used in this study are freely available and can be obtained from https://earthengine.google.com/. The ECMWF Reanalysis v.5 climatic data are available from https://www. ecmwf.int/en/forecasts/dataset/ecmwf-reanalysis-v5. The Regridded Harmonized World Soil Database v.1.2 is available from https://www. fao.org/soils-portal/data-hub. The Terra and Aqua combined MCD64A1 Version 6 Burned Area data are available from https://lpdaac.usgs.gov/ products/mcd64a1v006/. The Gridded Population of the World Version 4 data are available from https://sedac.ciesin.columbia.edu/data/set/ gpw-v4-population-count-rev11. The ASTER Global Elevation Model is available from https://asterweb.jpl.nasa.gov/gdem.asp. The AGB\_1 data are available from https://doi.org/10.3390/rs9050457, and the AGB\_2 data are available from https://doi.org/10.1126/sciadv.abe9829. The MODIS MOD17A3HGF v.6.1 product is available from https://lpdaac. usgs.gov/products/mod17a3hgfv061/. The VOD Ku-band data are available via Zenodo at https://doi.org/10.5281/zenodo.2575599 (ref. 84).

#### **Code availability**

All code used in the analyses is available via GitHub at https://github. com/surxuxu/Boreal-forest-diversity.

#### References

- 1. Pan, Y. D. et al. A large and persistent carbon sink in the world's forests. *Science* **333**, 988–993 (2011).
- 2. Bonan, G. B. Forests and climate change: forcings, feedbacks, and the climate benefits of forests. *Science* **320**, 1444–1449 (2008).
- 3. Liang, J. et al. Positive biodiversity–productivity relationship predominant in global forests. *Science* **354**, aaf8957 (2016).
- 4. Schnabel, F. et al. Species richness stabilizes productivity via asynchrony and drought-tolerance diversity in a large-scale tree biodiversity experiment. *Sci. Adv.* **7**, eabk1643 (2021).
- Liu, D., Wang, T., Peñuelas, J. & Piao, S. Drought resistance enhanced by tree species diversity in global forests. *Nat. Geosci.* 15, 800–804 (2022).
- 6. Hansen, M. C. et al. High-resolution global maps of 21st-century forest cover change. *Science* **342**, 850–853 (2013).
- Curtis, P. G., Slay, C. M., Harris, N. L., Tyukavina, A. & Hansen, M. C. Classifying drivers of global forest loss. *Science* **361**, 1108–1111 (2018).
- 8. Kharuk, V. I. et al. Climate-driven conifer mortality in Siberia. *Glob. Ecol. Biogeogr.* **30**, 543–556 (2021).
- 9. Reich, P. B. et al. Even modest climate change may lead to major transitions in boreal forests. *Nature* **608**, 540–545 (2022).
- Dial, R. J., Maher, C. T., Hewitt, R. E. & Sullivan, P. F. Sufficient conditions for rapid range expansion of a boreal conifer. *Nature* 608, 546–551 (2022).
- Huang, Y. et al. Impacts of species richness on productivity in a large-scale subtropical forest experiment. *Science* 362, 80–83 (2018).
- 12. Massey, R. et al. Forest composition change and biophysical climate feedbacks across boreal North America. *Nat. Clim. Change* **13**, 1368–1375 (2023).
- Joswig, J. S. et al. Climatic and soil factors explain the two-dimensional spectrum of global plant trait variation. *Nat. Ecol. Evol.* 6, 36–50 (2022).
- Keil, P. & Chase, J. M. Global patterns and drivers of tree diversity integrated across a continuum of spatial grains. *Nat. Ecol. Evol.* 3, 390–399 (2019).
- Cai, L. et al. Global models and predictions of plant diversity based on advanced machine learning techniques. *N. Phytol.* 237, 1432–1445 (2023).
- Schweiger, A. K. et al. Plant spectral diversity integrates functional and phylogenetic components of biodiversity and predicts ecosystem function. *Nat. Ecol. Evol.* 2, 976–982 (2018).
- Xi, Y., Zhang, W., Brandt, M., Tian, Q. & Fensholt, R. Mapping tree species diversity of temperate forests using multi-temporal Sentinel-1 and -2 imagery. Sci. Remote Sens. https://doi.org/ 10.1016/j.srs.2023.100094 (2023).
- Liu, X., Frey, J., Munteanu, C., Still, N. & Koch, B. Mapping tree species diversity in temperate montane forests using Sentinel-1 and Sentinel-2 imagery and topography data. *Remote Sens. Environ.* 292, 113576 (2023).
- 19. Wang, R. & Gamon, J. A. Remote sensing of terrestrial plant biodiversity. *Remote Sens. Environ.* **231**, 111218 (2019).
- 20. McDowell, N. G. et al. Pervasive shifts in forest dynamics in a changing world. Science **368**, eaaz9463 (2020).

- Gauthier, S., Bernier, P., Kuuluvainen, T., Shvidenko, A. Z. & Schepaschenko, D. G. Boreal forest health and global change. Science 349, 819–822 (2015).
- 22. Jarvis, P. & Linder, S. Constraints to growth of boreal forests. Nature **405**, 904–905 (2000).
- 23. Liang, J. et al. Co-limitation towards lower latitudes shapes global forest diversity gradients. *Nat. Ecol. Evol.* **6**, 1423–1437 (2022).
- 24. Pearson, R. G. et al. Shifts in Arctic vegetation and associated feedbacks under climate change. *Nat. Clim. Change* **3**, 673–677 (2013).
- 25. Rao, M. P. et al. Approaching a thermal tipping point in the Eurasian boreal forest at its southern margin. *Commun. Earth Environ.* **4**, 247 (2023).
- 26. Lenton, T. M. et al. Tipping elements in the Earth's climate system. Proc. Natl Acad. Sci. USA **105**, 1786–1793 (2008).
- 27. Tamarin-Brodsky, T., Hodges, K., Hoskins, B. J. & Shepherd, T. G. Changes in Northern Hemisphere temperature variability shaped by regional warming patterns. *Nat. Geosci.* **13**, 414–421 (2020).
- Mack, M. C. et al. Carbon loss from boreal forest wildfires offset by increased dominance of deciduous trees. *Science* **372**, 280–283 (2021).
- 29. Liu, Q. Y. et al. Drought-induced increase in tree mortality and corresponding decrease in the carbon sink capacity of Canada's boreal forests from 1970 to 2020. *Glob. Change Biol.* **29**, 2274–2285 (2023).
- Ma, Z. H. et al. Regional drought-induced reduction in the biomass carbon sink of Canada's boreal forests. Proc. Natl Acad. Sci. USA 109, 2423–2427 (2012).
- Peng, C. H. et al. A drought-induced pervasive increase in tree mortality across Canada's boreal forests. *Nat. Clim. Change* 1, 467–471 (2011).
- Wang, J. A., Baccini, A., Farina, M., Randerson, J. T. & Friedl, M. A. Disturbance suppresses the aboveground carbon sink in North American boreal forests. *Nat. Clim. Change* **11**, 435–441 (2021).
- Scheffer, M., Hirota, M., Holmgren, M., Van Nes, E. H. & Chapin, F. S.3rd Thresholds for boreal biome transitions. *Proc. Natl Acad. Sci.* USA 109, 21384–21389 (2012).
- Shannon, C. E. A mathematical theory of communication. Bell Syst. Tech. J. 27, 379–423 (1948).
- Peet, R. K. The measurement of species diversity. Annu. Rev. Ecol. Evol. Syst. 5, 285–307 (1974).
- Williams, L. J. et al. Remote spectral detection of biodiversity effects on forest biomass. *Nat. Ecol. Evol.* 5, 46–54 (2021).
- Zhao, Y. et al. Forest species diversity mapping using airborne LiDAR and hyperspectral data in a subtropical forest in China. *Remote Sens. Environ.* 213, 104–114 (2018).
- Hauser, L. T. et al. Explaining discrepancies between spectral and in-situ plant diversity in multispectral satellite earth observation. *Remote Sens. Environ.* 265, 112684 (2021).
- 39. Chen, X. et al. Tree diversity increases decadal forest soil carbon and nitrogen accrual. *Nature* **618**, 94–101 (2023).
- Wang, X. Y. et al. Disentangling the mechanisms behind winter snow impact on vegetation activity in northern ecosystems. *Glob. Change Biol.* 24, 1651–1662 (2018).
- 41. Dial, R. J. et al. Arctic sea ice retreat fuels boreal forest advance. Science **383**, 877–884 (2024).
- 42. Brice, M. H. et al. Moderate disturbances accelerate forest transition dynamics under climate change in the temperate-boreal ecotone of eastern North America. *Glob. Change Biol.* **26**, 4418–4435 (2020).
- Brice, M. H., Cazelles, K., Legendre, P. & Fortin, M. J. Disturbances amplify tree community responses to climate change in the temperate-boreal ecotone. *Glob. Ecol. Biogeogr.* 28, 1668–1681 (2019).

- 44. Fisichelli, N. A., Frelich, L. E. & Reich, P. B. Temperate tree expansion into adjacent boreal forest patches facilitated by warmer temperatures. *Ecography* **37**, 152–161 (2014).
- Boisvert-Marsh, L., Périé, C. & de Blois, S. Divergent responses to climate change and disturbance drive recruitment patterns underlying latitudinal shifts of tree species. *J. Ecol.* **107**, 1956–1969 (2019).
- 46. Wu, X. et al. Exposures to temperature beyond threshold disproportionately reduce vegetation growth in the northern hemisphere. *Natl Sci. Rev.* **6**, 786–795 (2019).
- 47. Adams, H. D. et al. Temperature sensitivity of drought-induced tree mortality portends increased regional die-off under global-change-type drought. *Proc. Natl Acad. Sci. USA* **106**, 7063–7066 (2009).
- 48. Randerson, J. T. et al. The impact of boreal forest fire on climate warming. Science **314**, 1130–1132 (2006).
- 49. Descals, A. et al. Unprecedented fire activity above the Arctic Circle linked to rising temperatures. *Science* **378**, 532–537 (2022).
- Gampe, D. et al. Increasing impact of warm droughts on northern ecosystem productivity over recent decades. *Nat. Clim. Change* 11, 772–779 (2021).
- 51. Li, F. et al. Global water use efficiency saturation due to increased vapor pressure deficit. *Science* **381**, 672–677 (2023).
- Berner, L. T. & Goetz, S. J. Satellite observations document trends consistent with a boreal forest biome shift. *Glob. Change Biol.* 28, 3275–3292 (2022).
- Juday, G. P., Alix, C. & Grant, T. A. Spatial coherence and change of opposite white spruce temperature sensitivities on floodplains in Alaska confirms early-stage boreal biome shift. *For. Ecol. Manage.* 350, 46–61 (2015).
- 54. Li, W. et al. Human fingerprint on structural density of forests globally. *Nat. Sustain.* **6**, 368–379 (2023).
- 55. Gross, N. et al. Functional trait diversity maximizes ecosystem multifunctionality. *Nat. Ecol. Evol.* **1**, 132 (2017).
- 56. Beck, P. S. A. et al. Changes in forest productivity across Alaska consistent with biome shift. *Ecol. Lett.* **14**, 373–379 (2011).
- 57. Wang, J. A. et al. Extensive land cover change across Arctic-Boreal Northwestern North America from disturbance and climate forcing. *Glob. Change Biol.* **26**, 807–822 (2020).
- Malone, T., Liang, J. & Packee, E. C. Cooperative Alaska Forest Inventory Vol. 785 (US Department of Agriculture, Forest Service, Pacific Northwest Research Station, 2009).
- 59. van Geffen, F. et al. SiDroForest: a comprehensive forest inventory of Siberian boreal forest investigations including drone-based point clouds, individually labeled trees, synthetically generated tree crowns, and Sentinel-2 labeled image patches. *Earth Syst. Sci. Data* **14**, 4967–4994 (2022).
- 60. Mauri, A., Strona, G. & San-Miguel-Ayanz, J. EU-Forest, a high-resolution tree occurrence dataset for Europe. *Sci. Data* **4**, 160123 (2017).
- 61. Nolan, K. A. & Callahan, J. E. Beachcomber biology: the Shannon-Weiner species diversity index. In Proc. the 27th Workshop/Conference of the Association for Biology Laboratory Education (ABLE) 334–338 (ABLE, 2006).
- 62. Camps-Valls, G. et al. A unified vegetation index for quantifying the terrestrial biosphere. Sci. Adv. **7**, eabc7447 (2021).
- 63. Moesinger, L. et al. The global long-term microwave Vegetation Optical Depth Climate Archive (VODCA). *Earth Syst. Sci. Data* **12**, 177–196 (2020).
- 64. Xu, L. et al. Changes in global terrestrial live biomass over the 21st century. *Sci. Adv.* **7**, eabe9829 (2021).
- 65. Fernandez-Moran, R. et al. SMOS-IC: an alternative SMOS soil moisture and vegetation optical depth product. *Remote Sens.* **9**, 457 (2017).

- Baccini, A. et al. Estimated carbon dioxide emissions from tropical deforestation improved by carbon-density maps. *Nat. Clim. Change* 2, 182–185 (2012).
- 67. Ma, Z. et al. Climate warming reduces the temporal stability of plant community biomass production. *Nat. Commun.* **8**, 15378 (2017).
- 68. Garcia-Palacios, P., Gross, N., Gaitan, J. & Maestre, F. T. Climate mediates the biodiversity–ecosystem stability relationship globally. *Proc. Natl Acad. Sci. USA* **115**, 8400–8405 (2018).
- 69. Besnard, S. et al. Mapping global forest age from forest inventories, biomass and climate data. *Earth Syst. Sci. Data* **13**, 4881–4896 (2021).
- Madonsela, S., Cho, M. A., Ramoelo, A. & Mutanga, O. Remote sensing of species diversity using Landsat 8 spectral variables. *ISPRS J. Photogramm. Remote Sens.* 133, 116–127 (2017).
- Soares, A., Körting, T., Fonseca, L. & Bendini, H. Simple nonlinear iterative temporal clustering. *IEEE Trans. Geosci. Remote Sens.* 59, 7669–7679 (2020).
- 72. Ismail Fawaz, H. et al. InceptionTime: finding AlexNet for time series classification. *Data Min. Knowl. Discov.* **34**, 1936–1962 (2020).
- Beisel, J. N. & Moreteau, J. C. A simple formula for calculating the lower limit of Shannon's diversity index. *Ecol. Model.* 99, 289–292 (1997).
- 74. Dušek, R. & Popelková, R. Theoretical view of the Shannon index in the evaluation of landscape diversity. *AUC Geogr.* **47**, 5–13 (2017).
- Elith, J. & Leathwick, J. Boosted Regression Trees for ecological modeling. CRAN http://download.nust.na/pub3/cran/web/ packages/dismo/vignettes/brt.pdf (2017).
- 76. Elith, J., Leathwick, J. R. & Hastie, T. A working guide to boosted regression trees. J. Anim. Ecol. **77**, 802–813 (2008).
- Zhang, W. et al. Socio-economic and climatic changes lead to contrasting global urban vegetation trends. *Glob. Environ. Change* **71**, 102385 (2021).
- 78. R Core Team R: A Language and Environment for Statistical Computing (R Foundation for Statistical Computing, 2013); http://www.R-project.org/
- Kuhn, M. Building predictive models in R using the caret package. J. Stat. Softw. 28, 1–26 (2008).
- Ridgeway, G. Generalized Boosted Models: A Guide to the gbm Package (CRAN, 2007).
- Wickham, H. et al. Welcome to the Tidyverse. J. Open Source Softw. 4, 1686 (2019).
- 82. Bates, D. et al. Package 'lme4'. Convergence 12, 2 (2015).
- Wickham, H. et al. ggplot2: Create elegant data visualisations using the grammar of graphics. R package version 2.0.0. *GitHub* https://github.com/tidyverse/ggplot2/releases/tag/v2.0.0 (2015).
- Moesinger, L. et al. The Global Long-Term Microwave Vegetation Optical Depth Climate Archive VODCA (1.0). *Zenodo:* https://doi.org/10.5281/zenodo.2575599 (2019).
- 85. Xi, Y. Boreal tree species diversity increases with global warming but is reversed by extremes. *figshare* https://doi.org/10.6084/m9.figshare.25034342 (2024).
- Harris, I., Osborn, T. J., Jones, P. & Lister, D. Version 4 of the CRU TS monthly high-resolution gridded multivariate climate dataset. *Sci. Data* 7, 109 (2020).
- Eyring, V. et al. Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization. *Geosci. Model Dev.* 9, 1937–1958 (2016).

#### Acknowledgements

Y.X. is supported by the Science & Technology Fundamental Resources Investigation Program (grant no. 2022FY101902), the National Natural Science Foundation of China (grant no. 42171367) and the China Scholarship Council (grant no. 202106190083). W.Z. acknowledges funding support from the Independent Research Fund Denmark (CLISA, grant no. 2032-00026B) and the ERC project TOFDRY (grant no. 947757). R.F. acknowledges support by the Villum Synergy project DeReEco.

#### **Author contributions**

W.Z. and Y.X. conceived the study and wrote the first draft of the paper. Y.X. and Z.F. performed the data analyses. R.F., Z.F. and F.W. contributed to the discussions of the results and to the text.

#### **Competing interests**

The authors declare no competing interests.

#### **Additional information**

Extended data is available for this paper at https://doi.org/10.1038/s41477-024-01794-w.

**Supplementary information** The online version contains supplementary material available at https://doi.org/10.1038/s41477-024-01794-w.

**Correspondence and requests for materials** should be addressed to Wenmin Zhang.

**Peer review information** *Nature Plants* thanks Colin Maher and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

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

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

**Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, 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 you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. 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-nc-nd/4.0/.

© The Author(s) 2024



Extended Data Fig. 1|See next page for caption.

#### Article

Extended Data Fig. 1 | Workflow of the tree species diversity mapping incorporating input data, object segmentation, spectral metrics calculation, model training, testing and map prediction. SBs: spectral

bands, VIs: vegetation indices, STMs: spectral temporal metrics, SHMs: spectral heterogeneity. A detailed description of the spectral metrics can be found in Supplementary Table 6.



**Extended Data Fig. 2** | **Feature importance and accuracy assessment of boreal tree species diversity estimation using Landsat satellite imagery.** Left, importance ranking of results from four group of predictors Phenology: monthly spectral bands and vegetation indices from May to October; Band + VIs: spectral bands and vegetation indices; SHMs: spectral heterogeneity metrics (that is,

coefficient of variation (CV), spectral dilation (SD), spectral gradient (SG), texture dissimilarity (TD), and entropy (TE)); STMs: Spectral-Temporal-Metrics. Right, scatterplots between the *in situ* H' values of 10 independent validation sample subsets (n = 10, 776) obtained by a ten-fold cross-validation method and the predicted H' values from the best model. The red dotted line is 1:1 line.



Extended Data Fig. 3 | Zoom-in examples of true color Landsat-8 images (RGB = bands 4, 3, 2) and tree diversity predictions at different spatial scales. a-d, 200 × 200 km<sup>2</sup>; e-h, 50 × 50 km<sup>2</sup>; i-l, 5 × 50 km<sup>2</sup>. The columns are true color Landsat-8 images (a, e, i) and diversity for 2000 (b, f, j), and 2010 (c, g, k), and 2020 (d, h, l).



**Extended Data Fig. 4** | **The relative contributions of changes in species richness and evenness to the dynamics of tree species diversity.** (*H'*) (following a Gaussian distribution), based on a multiple linear regression model with a response variable of changes in *H'* and the explanatory variables of changes in species richness and evenness. The explanatory variables richness and evenness were standardized. The plots with repeated measurements of *H*' derived from forest inventory dataset were used for the analysis (n = 648). EF: Eastern forest-boreal transition; CS: Canadian Shield forests; CR: Central Rockies forests; MC: Mid-Continental Canadian forests; NC: Northern Canadian Shield taiga.



**Extended Data Fig. 5** | **Pairwise correlations between different environmental variables and tree species diversity. for a**, 2000, **b**, 2010, and **c**, 2020. The color gradient of the legends indicates Pearson's correlation coefficient, with more positive/negative values (dark blue/red) indicating stronger positive/ negative correlations. Asterisks '\*' denote the significance levels of the Pearson's correlation coefficients based on two-sided t-tests: \*\*\* p < 0.001, \*\* p < 0.01, \* p < 0.05. D\_2000: H' values in 2000, D\_2010: H' values in 2010, D\_2020: H' values in 2020, TMP\_2000: mean seasonal temperature in 2000; PR\_2000: mean seasonal precipitation in 2000; Fire\_2000: fire activity frequency in 2000; POP\_2000: human population density in 2000; TMP\_2010: mean seasonal temperature in 2010; PR\_2010: mean seasonal precipitation in 2010; Fire\_2010: fire activity frequency in 2010; POP\_2010: human population density in 2010; TMP\_2020: mean seasonal temperature in 2020; PR\_2020: mean seasonal precipitation in 2020; Fire\_2020: fire activity frequency in 2020; POP\_2020: human population density in 2020; S\_OC: topsoil organic carbon; S\_SAND: topsoil sand fraction; S\_CEC: cation exchange capacity; DEM: Digital Elevation Model.



**Extended Data Fig. 6** | **Temporal changes in boreal tree species diversity in relation to stand age.** Spatial distribution of trends in diversity associated with stand age (**a**). *H'* is aggregated at a spatial resolution of  $1.5^{\circ} \times 1.5^{\circ}$  and is shown by dots, with larger dots indicating higher diversity gains (green) or losses (purple). **b**, Response of age to diversity trend. Diversity trends were binned according to

age at intervals of 10 years. The black squares indicate average diversity trends within each bin, the grey lines represent the standard error of the mean of the diversity trend, and the color bars indicate the number of grid cells in each bin. Age: stand age.



Extended Data Fig. 7 | Relationship between changes in NDVI/tree cover and H' based on a multiple linear regression model. Seasonal mean (May– October) NDVI for 2000 (NDVI\_2000) and 2020 (NDVI\_2020) obtained from the Landsat-7 and Landsat-8 satellites. The tree cover data for 2000 (TC\_2000) and 2020 (TC\_2020) were obtained from Hansen et al.<sup>6</sup> and the Copernicus Global Land Service (Buchhorn et al. 2019). Changes in greenness and tree cover were calculated from the differences between NDVI\_2020 and NDVI\_2000, TC\_2020 and TC\_2000, respectively, and were used as explanatory variables. The  $\beta$  and R<sup>2</sup> marked in black were calculated using the *in situ* H' as the response variable, while those in grey were calculated using the predicted H' as the response variable. Both NDVI and tree cover were standardised prior to the linear regression analysis. The plots with repeated measurements of H' derived from the forest inventory dataset were used for this analysis (n = 648).

# nature portfolio

Corresponding author(s): Wenmin Zhang

Last updated by author(s): Aug 8, 2024

# **Reporting Summary**

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our <u>Editorial Policies</u> and the <u>Editorial Policy Checklist</u>.

#### Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.						
n/a	Cor	firmed				
	$\boxtimes$	The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement				
	$\boxtimes$	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly				
	$\boxtimes$	The statistical test(s) used AND whether they are one- or two-sided Only common tests should be described solely by name; describe more complex techniques in the Methods section.				
	$\boxtimes$	A description of all covariates tested				
	$\boxtimes$	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons				
	$\boxtimes$	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)				
	$\boxtimes$	For null hypothesis testing, the test statistic (e.g. F, t, r) with confidence intervals, effect sizes, degrees of freedom and P value noted Give P values as exact values whenever suitable.				
$\boxtimes$		For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings				
$\boxtimes$		For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes				
	$\boxtimes$	Estimates of effect sizes (e.g. Cohen's d, Pearson's r), indicating how they were calculated				
	•	Our web collection on <u>statistics for biologists</u> contains articles on many of the points above.				

#### Software and code

Policy information about availability of computer code						
Data collection	Satellite remote sensing data were accessed from GEE and no softwares were used to access field measurements of data used in this study.					
Data analysis	Analyses were performed using R version 4.2.0 (2022) and with the following packages: caret, gbm, tidyverse, Ime4, and ggplot2 as well as GEE. Image segmentation was processed with the SNIC algorithm while the the InceptionTime architecture deep learning approach was used for prediction. The relevant codes for data processing can be accessed from: https://github.com/surxuxu/Boreal-forest-diversity					

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio guidelines for submitting code & software for further information.

#### Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

Boreal tree species diversity data used to conduct analyses for the manuscript are available at https:// figshare.com/s/a7cac743013cf536caaf. The Landsat surface reflectance data used in this study are freely available and can be obtained from https://earthengine.google.com/. ECMWF Reanalysis v5 (ERA5)

Inature portfolio | reporting summary

climatic data are available from https://www.ecmwf.int/en/forecasts/dataset/ecmwf-reanalysis-v5. Regridded Harmonized World Soil Database v.1.2 are available from https://www.fao.org/soils-portal/data-hub. Terra and Aqua combined MCD64A1 Version 6 Burned Area data are available from https://lpdaac.usgs.gov/ products/mcd64a1v006/. Gridded Population of World Version 4 (GPWv4) are available from https://sedac.ciesin.columbia.edu/data/set/gpw-v4-population-countrev11. ASTER Global Elevation Model are available from https://asterweb.jpl.nasa.gov/gdem.asp. The AGB\_1(aboveground biomass) data are available from https:// doi.org/10.3390/rs9050457, and the AGB\_2 data are available from https://doi.org/10.1126/sciadv.abe9829. MODIS MOD17A3HGF Version 6.1 product are available from https://lpdaac.usgs.gov/products/mod17a3hgfv061/. VOD Ku-band data are available from https://doi.org/10.5281/zenodo.2575599.

#### Research involving human participants, their data, or biological material

Policy information about studies with human participants or human data. See also policy information about sex, gender (identity/presentation), and sexual orientation and race, ethnicity and racism.

Reporting on sex and gender	N/A
Reporting on race, ethnicity, or other socially relevant groupings	N/A
Population characteristics	(N/A
Recruitment	(N/A
Ethics oversight	N/A

Note that full information on the approval of the study protocol must also be provided in the manuscript.

## Field-specific reporting

Please select the one be	elow that is the best fit for your research.	. If you are not sure, read the appropriate sections before making your selection					
Life sciences	Behavioural & social sciences	Ecological, evolutionary & environmental sciences					
For a reference copy of the document with all sections, see <u>nature.com/documents/nr-reporting-summary-flat.pdf</u>							

# Ecological, evolutionary & environmental sciences study design

All studies must disclose on these points even when the disclosure is negative.

Study description	This study developed a novel approach to map continuous spatial distribution of tree species diversity and explored the associated impact on boreal ecosystem productivity and stability					
Research sample	We collected samples including 5.127 field observations of a total of 188.639 trees divided into 254 tree species rom from six					
Research sumple	countries (Supplementary Table 5) and satellited-based gridded datasets.					
Sampling strategy	NA					
Data collection	Plot based data are available referring to Supplementary Table 5 and the satellite data are open access from Google Earth Engine.					
liming and spatial scale	Data covering the boreal forest ecosystem spaning 2000-2020 are used.					
Data ovelusions	Satallita pixals with page quality					
Data exclusions	Satellite pixels with pool quality.					
Poproducibility	All results are being processed many time and the data and code are available from https://github.com/survuyu/Boreal-forest-					
Reproducibility	diversity					
Randomization	NA					
Blinding	(NA					
Did the study involve field work? 🛛 Yes 🕅 No						

## Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

# nature portfolio | reporting summary

#### Materials & experimental systems

n/a	Involved in the study
$\boxtimes$	Antibodies
$\boxtimes$	Eukaryotic cell lines
$\boxtimes$	Palaeontology and archaeology
$\boxtimes$	Animals and other organisms
$\boxtimes$	Clinical data
$\boxtimes$	Dual use research of concern
$\boxtimes$	Plants

#### Methods

n/a Involved in the study

ChIP-seq

- Flow cytometry
- MRI-based neuroimaging