communications
earth & environment

ARTICLE

https://doi.org/10.1038/s43247-023-00758-w OPEN

Spaceborne LiDAR reveals the effectiveness of European Protected Areas in conserving forest height and vertical structure

Guido Ceccherini [1,3✉], Marco Girardello [1,3✉], Pieter S. A. Beck [1], Mirco Migliavacca [1], Gregory Duveiller [2], Grégoire Dubois [1], Valerio Avitabile [1], Luca Battistella [1], José I. Barredo [1] & Alessandro Cescatti [1]

The effectiveness of Protected Areas in conserving forest ecosystems has been examined at the continental scale using area-based habitat parameters, but knowledge of the three-dimensional structure of forest habitats is still lacking. Here, we assess the effectiveness of European Protected Areas in conserving the vertical structure of forests by analysing more than 30 million records from the Global Ecosystem Dynamics Investigation (GEDI), a spaceborne LiDAR (Light Detection And Ranging) mission. We compare a suite of indicators of the vertical structure of forests inside and outside nearly 10,000 protected areas. We find that European forests are on average 2 m taller and vertically more complex in protected areas than in nearby unprotected areas, albeit with some regional differences. At the same time, forests outside protected areas show greater variations in canopy height than inside, probably as a result of past and current forest management operations. Our findings highlight the positive imprint of environmental policies on forest structure across Europe and underscore how spaceborne LiDAR enables the large-scale monitoring of forest vertical structural attributes that are key to conservation and restoration policies.

[1] European Commission, Joint Research Centre, Ispra, Italy. [2] Max Planck Institute for Biogeochemistry, Jena, Germany. [3] These authors contributed equally: Guido Ceccherini, Marco Girardello. ✉email: guido.ceccherini@ec.europa.eu; marco.girardello@ec.europa.eu

Protected Areas (PA) are a cornerstone of environmental policies worldwide and play a crucial role in conserving biodiversity[1,2], maintaining genetic resources[3], sustaining important ecosystem services and related functions as well as cultural landscapes[4]. The most common ecosystem types in European terrestrial PAs are forests[5].

On the policy side, the assessment of the forest ecosystem structure within PAs is important for the post-2020 global biodiversity targets[6] and the EU's restoration policies[7]. Spatially exhaustive, consistent and robust records on the physical complexity of forests, including the three-dimensional structure of their canopy are required to carry out biodiversity assessments and provide baseline information for ecosystem restoration initiatives[8].

Forest structural complexity is a multifaceted concept related to the number of structural attributes and their relative abundance in forests[9]. Despite the growing body of literature on this topic, at the moment there is no unique definition of forest structural complexity and there is a lack of consensus on the most appropriate means to measure it. Forest structural complexity can be broadly classified into horizontal and vertical complexities: vertical complexity describes the layering of the vertical profiles while horizontal complexity describes the variability across a horizontal space (e.g.[10–13]).

Among forest structural indicators, canopy height is a robust proxy for aboveground biomass, ecosystem structure, and biodiversity[14–16]. Canopy height and its spatial variation might reflect stand age and often also the intensity of disturbances or management[17]. Taller forests are known to provide more niches and promote species' coexistence[18,19]. Moreover, they often provide a greater diversity of microhabitats, harbouring species characterised by divergent evolutionary histories and functional traits, which translates into high phylogenetic and functional diversity[20,21]. Canopy height influences maximum productivity and water use efficiency[22], by moderating air temperature extremes and water stress. Furthermore, height - and its local spatial variation - leads to a more thorough mixing of the atmospheric boundary layer, enabling greater heat exchange[23].

Assessing the structure of forests over large areas remains difficult. Whilst it can be done manually through ground-based surveys, it is generally prohibitively expensive. Stereogrammetry, i.e. the science of using overlapping imagery acquired from different angles, allows forest height estimation. However, applications of stereogrammetry have been generally limited to small study areas where both trees and the ground are visible to the sensor[24,25]. As an alternative, laser scanning, i.e. LiDAR (Light Detection And Ranging) is an active remote sensing technique, where tree height is derived from 3D data measured through the echoes of the pulses emitted by a laser instrument mounted on a location rotating around its axes (i.e. Terrestrial Laser Scanning) or mounted on an aircraft and directed to the ground (i.e. Airborne Laser Scanning). Airborne LiDAR is commonly used for national-level surveys[26]. More recently, estimates of tree height from satellite retrievals have become available[27,28]. However, data generated from these missions are provided at very coarse resolutions, relying on instruments optimised for variables other than vegetation parameters. For these reasons, the study of the forest's structural complexity has so far been limited to relatively small areas (e.g.[29,30]).

The Global Ecosystem Dynamics Investigation (GEDI) instrument, placed on board the International Space Station (ISS) in 2018, is now revolutionising the assessment of forest structural complexity[31] and biomass[32]. GEDI retrieves regular and accurate information on forest height and structure through LiDAR shots that yield a vertical energy profile (i.e. full-waveform) covering a footprint of 25 m diameter on the ground. Canopy top height and other vertical structural parameters associated with the structural complexity are derived from these shots. The GEDI mission is designed as a sampling mission and collects sparse samples of the Earth's surface. It overcomes the shortcomings of in situ field measurements and airborne laser scanning, which have limited and non-uniform spatial and temporal coverage. Moreover, GEDI provides denser observational coverage than the other spaceborne LiDAR missions (i.e. ICESat and IceSat2) mainly designed to monitor changes in polar glaciers and sea ice. Recent studies have shown that GEDI structural measurements are strong predictors of tree species richness over natural forests[33]. GEDI measurements can greatly advance our ability to quantify forest biodiversity and habitat quality[34] as well as biomass and productivity[35], all directly relevant to forest management and conservation. A recent analysis of protected areas at global scale assessed[36] the impact of protection measures on the growing vegetation. However, this was only focused on the top canopy top height and did not consider the vertical structure below and within the canopy.

In this study, we use more than 30 million GEDI observations to investigate the effectiveness of European PAs in conserving forest vertical structural complexity compared to non-protected forests. While vertical structure metrics can be derived from single shots, horizontal structure metrics require the analysis of neighbouring shots, which is limited by the density of GEDI data (see GEDI section in Methods). We calculated a suite of forest structural variables from GEDI (version 2) data, including canopy top height, Foliage Height Diversity (FHD), and Relative Height (RH) metrics (i.e., RH50 and RH25) (for further details, see Fig. S1 in Supplementary Information). RH metrics[31] represent the height above ground elevation below which a given percentage of LiDAR signal has been returned. RH50, for example, is the height relative to the ground elevation below which 50% of the LiDAR signal has been returned. FHD is a canopy structural index describing the vertical heterogeneity of the foliage profile[37]. For the present analysis, we only considered forest areas where the top canopy height has a value greater or equal to 5 m - to be compliant with the Food and Agriculture Organization (FAO) forest definition[38]. In addition to the abovementioned GEDI metrics, we computed two indices of forest structural complexity (theoretical framework described in Methods). The first index is a measure of the entropy (i.e. Shannon Diversity Index), while the second one is a measure of dissimilarity (i.e. a mean dissimilarity Index) of structural complexity. Higher values for individual GEDI structural indicators and the alpha diversity index indicate an increase in height or structural complexity within a given PA. On the other hand, higher values for the coefficient of variation or the dissimilarity index indicate an increase in the spatial variability of structural indicators among samples of a given PA.

We assessed the effectiveness of PAs included in the Natura 2000 network[39]. Natura 2000 is the *de facto* European reference network of areas, designated as "Special Areas of Conservation" and "Special Protection Areas", implementing the EU's Habitats Directive and the Birds Directive, respectively. PAs can vary considerably in their protection levels and silvicultural activities might be allowed. To have a baseline containing forest habitats where human impact is minimised, we also considered PAs classified within the World Database on Protected Areas (WDPA) (UNEP-WCMC, IUCN (2021)[40]) where protection level is reported. The WDPA database (November 2021 version) is the most comprehensive global spatial dataset on terrestrial, marine, and coastal protected areas and includes all nationally designated (e.g. national parks) and internationally recognized (e.g. RAMSAR sites) protected areas. Within WDPA, we selected only terrestrial sites reported with a greater degree of protection (i.e.

Table 1 Description of International Union for Conservation of Nature (IUCN) Protected Area Management Category.

Category	Definition
Category Ia	Strict Nature Reserve.
Category Ib	Wilderness Area.
Category II	National Park.
Category III	Natural Monument or Feature.
Category IV	Habitat/Species Management Area.
Category V	Protected Landscape/Seascape.
Category VI	Protected area with sustainable use of natural resources.

Strict Nature Reserves, Wilderness Areas and National Parks, for further details, see Table 1). The majority of these areas are the most protected sections of Natura 2000 sites.

Here, we addressed the following research questions: (Q1) are the structural characteristics of forests different between Protected Areas and their unprotected surroundings (henceforth referred to as Unprotected Surroundings)? (Q2) If so, in which regions and to what degree? (Q3) Are canopy heights more spatially homogeneous or heterogeneous in Protected Areas compared to Unprotected Surroundings? (Q4) What is the influence of the year of establishment of protected areas on forests' vertical structural complexity?

Results

Differences in vertical structural complexity between Protected Areas and Unprotected Surroundings. We provided the first EU-scale assessment (up to 51.6° North) of the forest vertical structural differences between Protected Areas and unprotected areas in their surroundings. To compare changes inside and outside PAs, we considered a 10 km unprotected buffer zone around each PA (i.e. referred to here as Unprotected Surroundings). For 7831 of the 9898 Natura 2000 sites assessed, we retrieved information on their protection level from the WDPA database. Only 391 sites have a high degree of protection - i.e. "Strict Nature Reserve", "Wilderness Area" and "National Park" (hereafter WDPA strictly PAs, for further details, see Table 1). Our results clearly showed that the mean values of the structural variables are higher when forests are in PAs than in Unprotected Surroundings. On average, the EU-wide mean difference in canopy top height between Natura 2000 PAs and Unprotected Surroundings is 2 ± 0.13 m (Fig. 1a and density plots in Fig. 2). Particularly large differences were found in Eastern Europe, where trees can be up to 7 m taller in PAs. In a similar manner to canopy top height, we observed a positive difference for FHD (0.10 ± 0.06), Shannon Diversity Index (0.14 ± 0.02), RH50 (1.7 ± 0.10 m), and RH25 (1.1 ± 0.07 m) (See Fig. 1 and Fig. S1). For clarity, we show and discuss mostly the results for canopy top height, FHD and Shannon Diversity Index, since they are the most important structural features - or synthetic indices of structural complexity - of trees. Results for the other forest structural variables showed identical effects in terms of direction (i.e. increase in PAs) and are reported in the Supplementary Information figure (Fig. S1). A close inspection of the inter-correlation among the GEDI structural variables (scatterplot in Fig. S2) showed that these are highly correlated.

PAs that are strictly protected according to WDPA showed similar patterns to those only in Natura 2000 (Fig.1b for canopy top height and Fig.S1 for the other variables), with a mean difference in canopy top height of 2.15 ± 0.36 m, in FHD of 0.10 ± 0.15 and in Shannon Diversity Index of 0.03 ± 0.05 across the EU. Differences between PAs and Unprotected Surroundings were statistically significant for a high proportion of the area considered (i.e. 80% and 64% of the hexagons in Fig.1 for Natura 2000 and WDPA strictly PAs, respectively) according to the t-test analysis ($p < 0.05$ with Bonferroni adjustment - shown as dots in Fig.1). These patterns were consistently observed for WDPA strictly PAs sites for both deciduous and evergreen forests (See Fig. S3), while for Natura 2000 evergreen forests positive differences are less marked (i.e. mean difference in canopy top height of 0.49 ± 0.11 m).

Unprotected Surroundings generally show greater spatial variation in canopy top height than nearby PAs (quantified as Coefficient of variation of canopy top height, Fig. 1c and d), thus suggesting that this variable may be able to capture the fingerprint left by forest management. This result is further supported by the Mean Dissimilarity (Fig. S1) where higher values are found in Unprotected Surroundings.

Regional differences in vertical structural complexity. We compared differences in canopy top height and FHD between Protected Areas and Unprotected Surroundings for different biogeographical regions (Fig. 3). A 2 m EU-wide difference was found both for Natura 2000 sites, and for the strictly protected WDPA sites. For Natura 2000, the Pannonian and Continental regions showed the largest differences in Canopy Top Height (3.9 m and 2.2 m, respectively), while the Mediterranean and Alpine regions showed the smallest ones (1.6 m and 1.1 m, respectively). When comparing Natura 2000 with WDPA strictly PAs (Fig. 2), used here as a reference for forests where human impact is limited, the analysis suggests that the Alpine and the Mediterranean regions exhibited a large difference, while the Atlantic region had a small one, even if the sample size is very different. As for Fig. 1, FHD and the other metrics related to vertical structural complexity showed similar results (see Fig. S5). The relative difference in Canopy Top Height between Protected Areas and Unprotected Surroundings is ~9% for Natura2000 and ~8% for strictly protected WDPA sites.

The observed differences among regions were strongly supported by Generalized Linear Mixed Models (Fig. 4), fitted within a Bayesian framework that allowed us to 1) account for spatial autocorrelation in the data and 2) fully estimate uncertainties around the estimated differences.

For both Canopy Top Height and FHD, WDPA results show larger width for the credibility intervals, probably due to the small sample size of the areas with a high degree of protection. Except for the Atlantic biogeographical region, there were marked differences between Protected Areas and Unprotected Surroundings. Conversely, summaries of the posterior distributions of the models fitted to the Natura2000 data show little overlap.

We quantified differences in GEDI structural variables and indices between PAs and Unprotected Surroundings according to the protection level (Fig. S6). PAs in the WDPA dataset are classified according to different categories based on their management objectives and are labelled using seven classes, ranging from strict protection (Class Ia) to multiple uses (Class VI). For further details, see the Protected Areas section and Table S1 in Supplementary Information. Generally, for strict nature reserves (i.e. Class Ia) the differences in Canopy Top Height, FHD, RHs and Shannon Diversity Index values between PAs and Unprotected Surroundings are more evident than for Wilderness Areas (i.e. Class Ib). Classes III to VI show an opposite pattern, with higher trees in the Unprotected Surroundings, probably due to the active management in those PAs. However, results are still very dependent on the small number of sample sites, and they should be taken with caution.

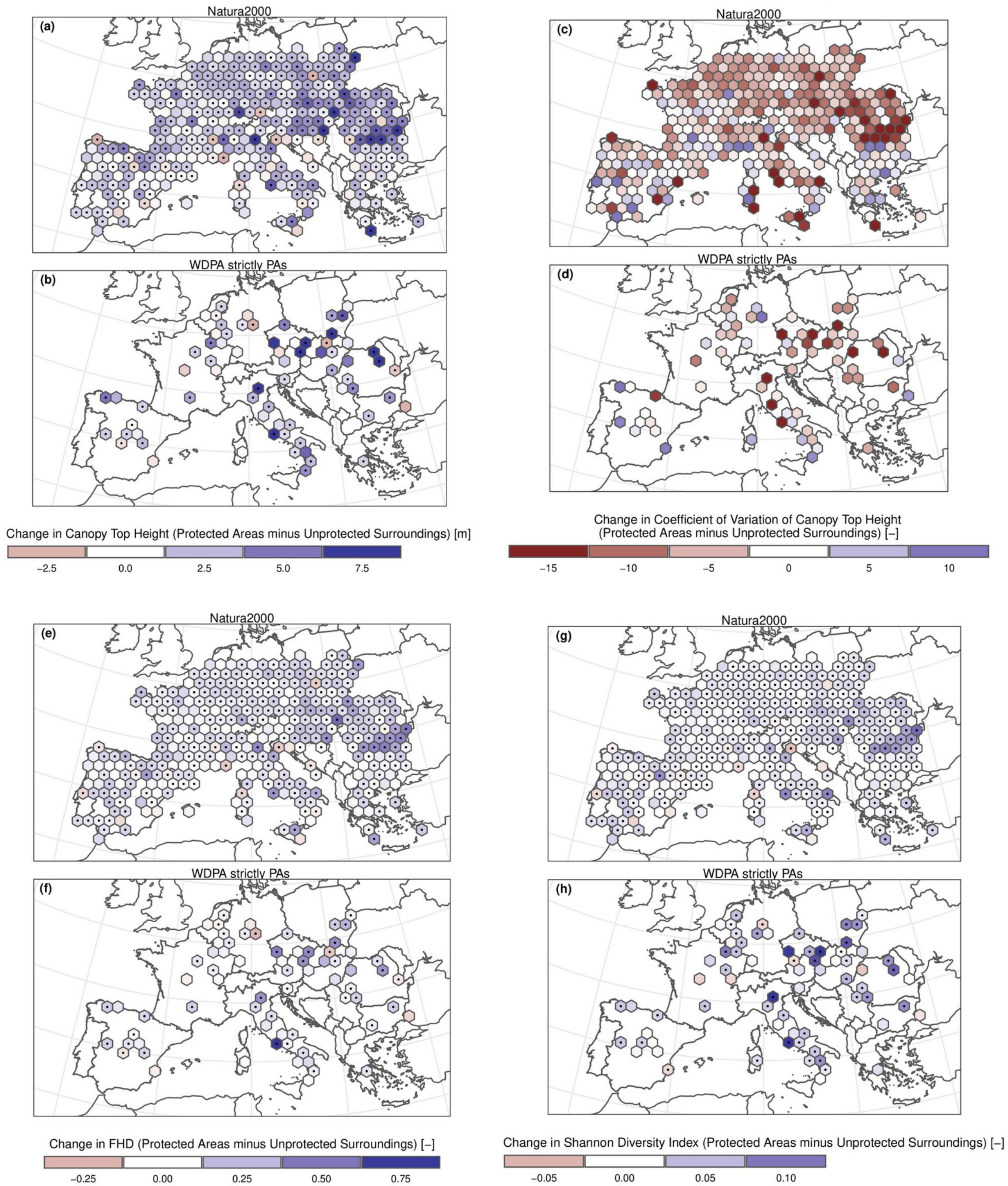

Fig. 1 Differences between Protected Areas and Unprotected Surroundings. Difference in Canopy Top Height between Protected Areas and Unprotected Surroundings for (**a**) Natura 2000 sites and (**b**) strictly protected sites according to the World Database on Protected Areas (WDPA). Difference in coefficient of variation of canopy top height (**c**, **d**), Foliage Height Diversity (**e**, **f**) and Shannon Diversity Index (**g**, **h**). Within each hexagon grid cell (side 100 km), we computed the mean values - and the coefficient of variation - of canopy top height in Protected Areas and Unprotected Surroundings and then their difference. Grey-filled colour identifies EU countries where concurrent PAs and GEDI information is available. Note that Croatia does not appear in the WDPA because the protection classification is missing for its PAs. Dots in the hexagons indicate when differences are statistically significant (i.e. $p < 0.05$) using a t-test with Bonferroni correction (not performed for the coefficient of variation of canopy top height). Density plots of the differences (**i–l**) for Natura 2000 sites (blue filled) and strictly protected sites according to the World Database on Protected Areas (yellow filled - WDPA). The vertical solid line represents the mean value, while the vertical dot-dashed lines the standard deviation.

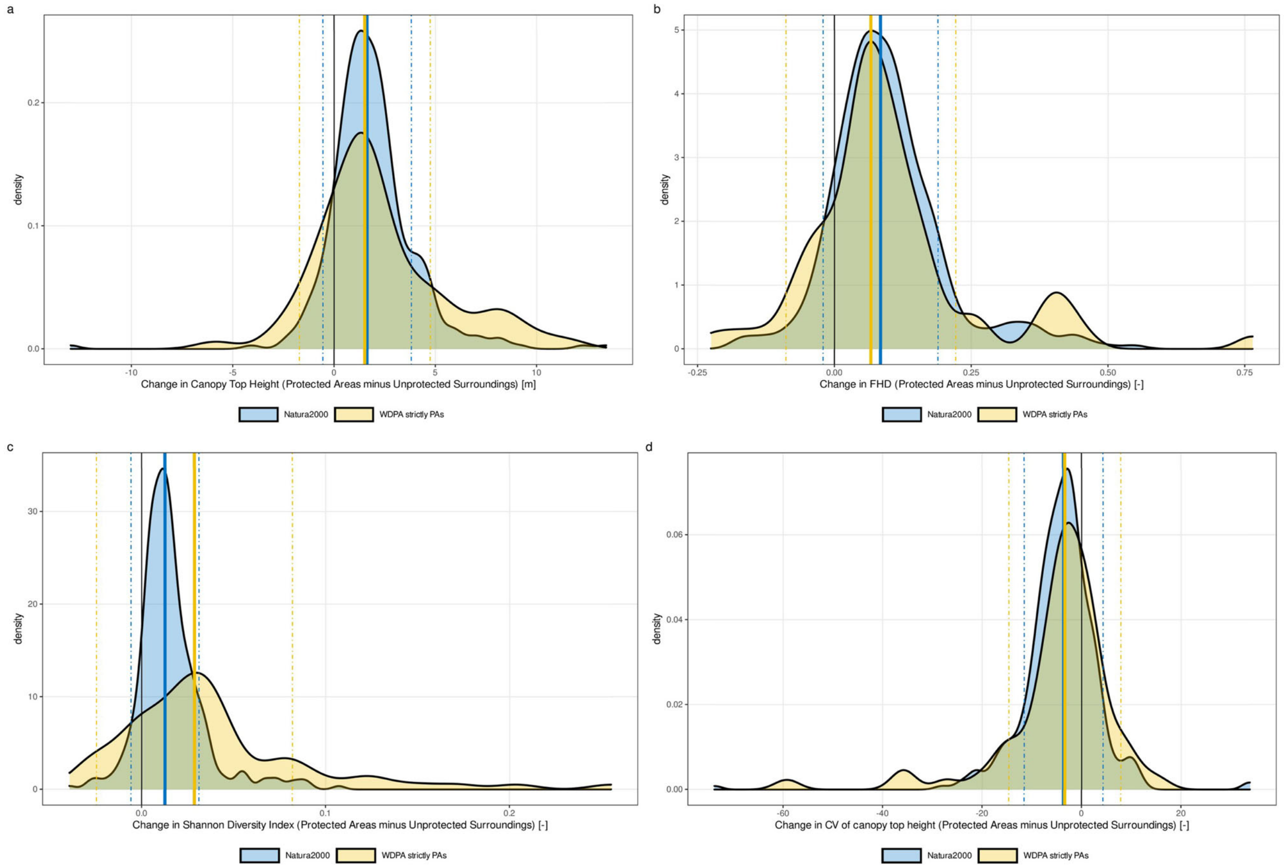

Fig. 2 Density plots of the differences between Protected Areas and Unprotected Surroundings. Density plots of the differences(**a**–**d**) for Natura 2000 sites (blue filled) and strictly protected sites according to the World Database on Protected Areas (yellow filled - WDPA). The vertical solid line represents the mean value, while the vertical dot-dashed lines the standard deviation.

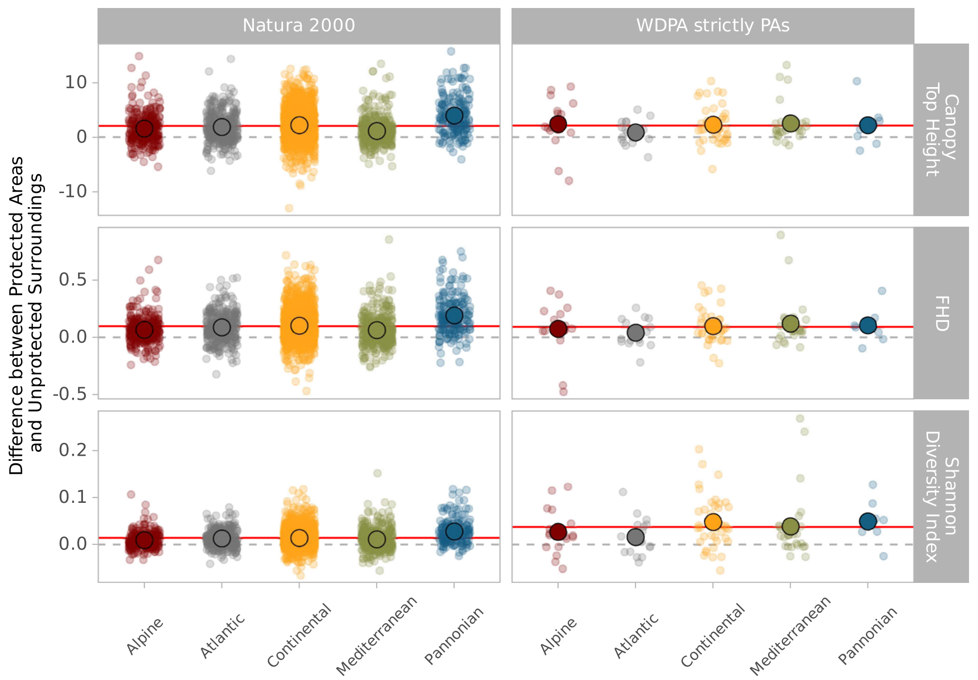

Fig. 3 Differences between Protected Areas and Unprotected Surroundings for different biogeographical regions. Canopy Top Height, Foliage Height Diversity and Shannon Diversity Index differences between forests within Protected Areas and their respective Unprotected Surroundings, separated according to the designation of the PA: Natura 2000 sites and strictly protected sites according to the World Database on Protected Areas (WDPA). Different columns refer to biogeographical regions[96] shown in Fig.S4. The horizontal red line is the European Union (EU) average difference, the big circles represent the regional averages, and the small dots are individual PAs.

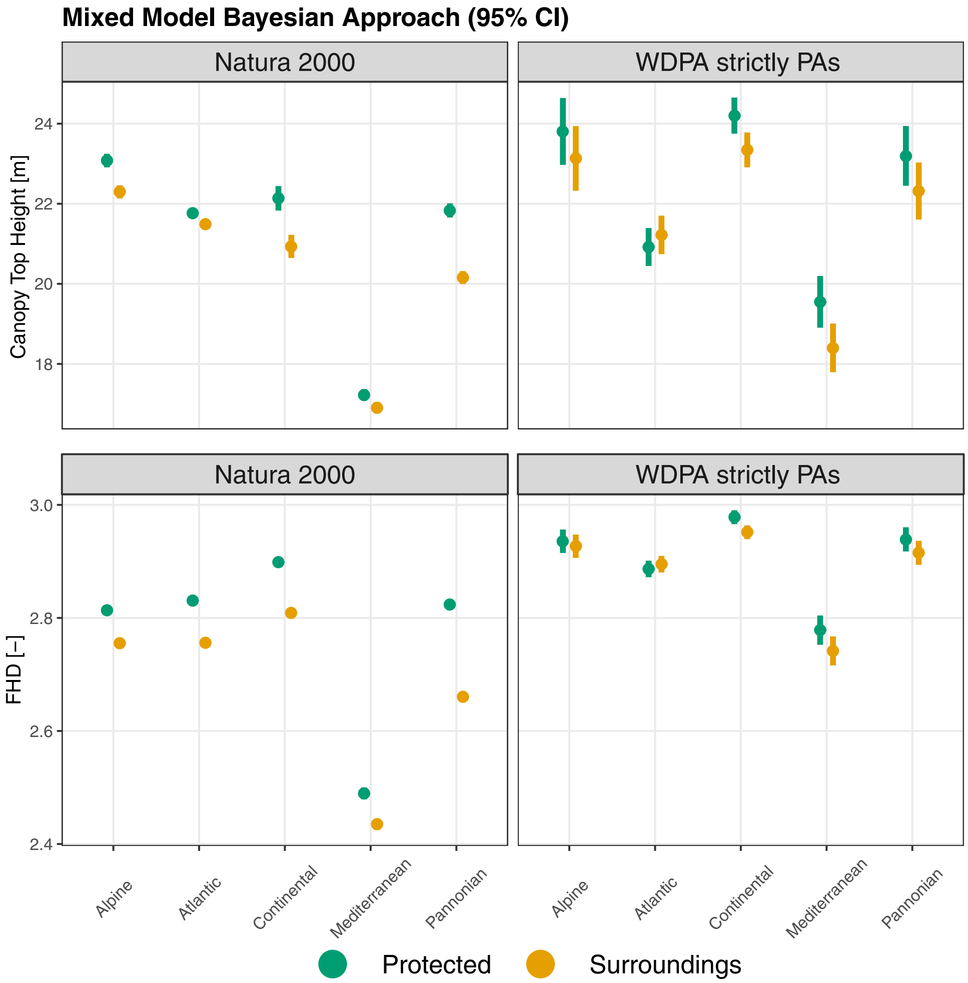

Fig. 4 Generalized Linear Mixed Model analysis of differences in vertical structural attributes. Results of the generalized linear mixed model showing the Canopy Top Height and Foliage Height Diversity for Protected Areas and Unprotected Surroundings in Natura 2000 and in the strictly protected sites according to the World Database on Protected Areas (WDPA) for different biogeographical regions.

Structural complexity in relation to the designation year of PAs. The analysis of the correlation between the structural variables and the PAs's designation year showed that the latter could be important in increasing the structural complexity of the forests within PAs in the Alpine, Continental, and Mediterranean biogeographical regions (Fig. 5). We did not examine the relationship between establishment year and Canopy Top Height and FHD differences for the Pannonian and Atlantic regions as the sample size was too low ($n = 5$ and $n = 17$, respectively).

For the Mediterranean region, the median value of the slope is strictly positive (i.e. 0.08 m / year) and the 2.5th and 97.5th percentiles of its posterior distribution do not overlap with 0 (i.e. 0.01 and 0.16, respectively), thus suggesting that age has a moderate impact on Canopy Top Height. For the other regions, while the slope is positive, the 2.5th and 97.5th percentiles of the posterior distribution are encompassing 0 (−0.06 to 0.16 and −0.03 to 0.19 for Alpine and Continental regions, respectively), providing little evidence for an effect of establishment year.

For FHD and RHs we observed similar results as for Canopy Top Height, indicating an increase in the different metrics in older protected areas (see Fig. S7 in Supplementary Information where the relationship is shown at EU scale for different levels of protection). One notable difference in the Canopy Top Height and FHD results, was for the continental region, for which age had a higher explanatory capacity for FHD.

Discussion

Effectiveness of protected areas in Europe. Our results suggest that European PAs effectively protect forest structural complexity. These findings are based on robust statistical inference, combined with novel spaceborne LiDAR retrievals, comprising more than 30 million observation records. Forests in PAs are higher than those in their immediate surroundings - with an average difference of ~2 m - and are more structurally complex (see the increase of RHs as reported in Fig.S1). The results were consistent for all Natura 2000 areas and a subset of strictly protected areas, thus providing strong support for the effectiveness of PAs, including those where active land management is allowed.

The effectiveness of PAs has already been examined in several studies (e.g.[41–43]). However, previous global or continental-scale assessments have either focused on direct measurements of biodiversity, such as species richness, or area-based habitat proxies (e.g. land cover). Forest structural complexity and, in general, the vertical dimension of the canopy had not been considered before[44]. Structural variables are known to be excellent indicators of biodiversity and ecosystem functioning and strongly correlate with habitat quality, primary production, nutrient cycling, and light capture[45–47]). Additionally, more structurally complex forests are known to buffer the effect of biotic disturbances[48,49]. Recently[50] found that PAs are on average cooler, probably due to their higher structural complexity and

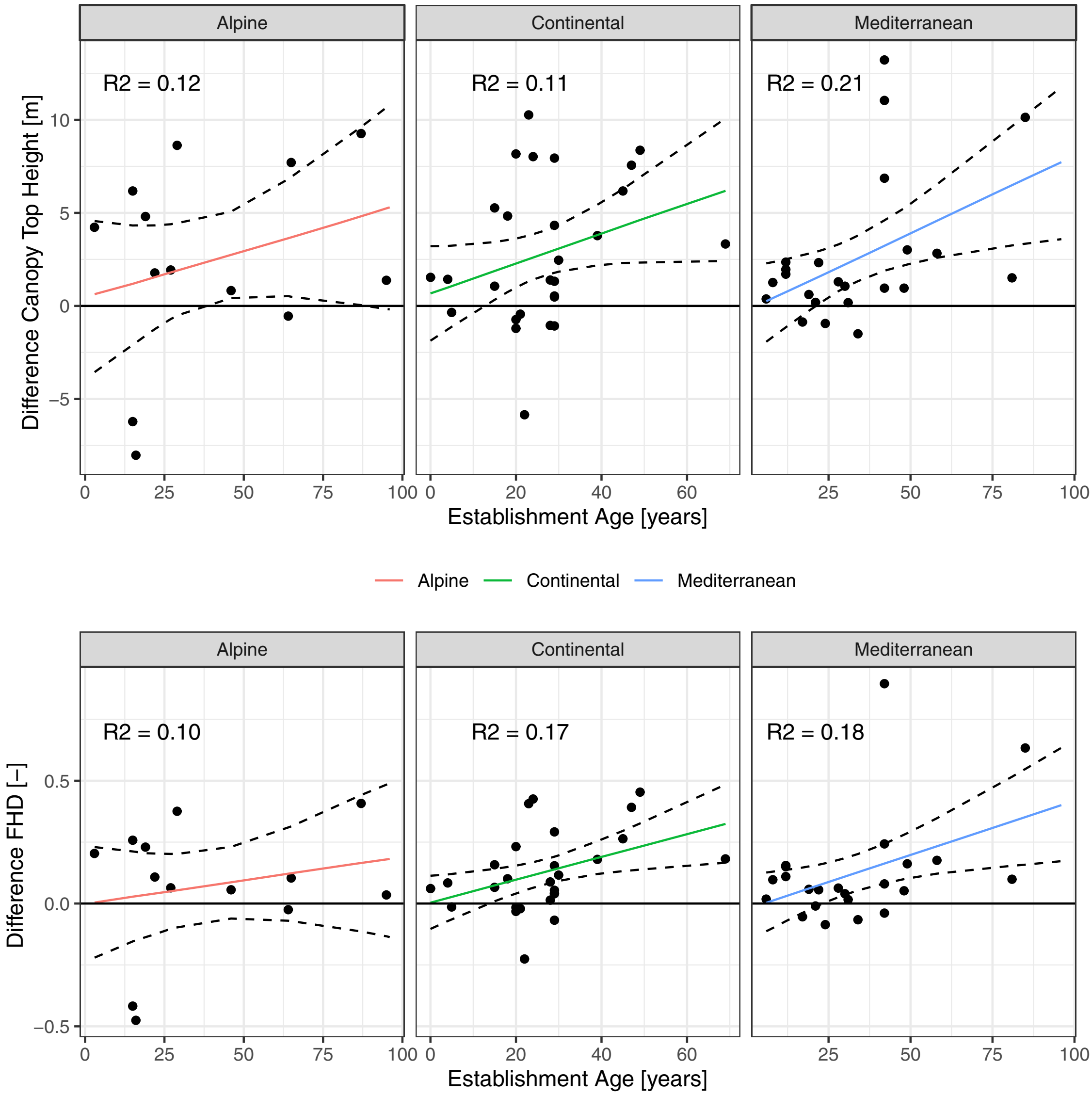

Fig. 5 Establishment age of PAs and differences in Canopy Top Height and Foliage Height Diversity between Protected Areas and Unprotected Surroundings. Relationship between Establishment age and difference in Canopy Top Height and Foliage Height Diversity between Protected Areas and Unprotected Surroundings based on biogeographical regions for the WDPA dataset. Dashed lines 95% show credibility intervals calculated from the posterior distributions of three linear models fitted within a Bayesian framework.

heterogeneity and lower disturbances. By focusing on the vertical component of forest habitats, as measured by LiDAR, our study overcomes the limitations of previous assessments, based on metrics describing habitat extent.

While, with the current data, it is not possible to identify all the mechanisms underlying geographical variation in the differences of structural variables between PAs and unprotected surroundings, we identify two main causes: current management practices and past land management.

First, current management practices within and outside PAs contribute to the observed differences in canopy height and other forests' vertical structural complexity variables. Management can have profound effects on the spatial distribution and connectivity of forest habitats and trees with important implications for forest demography and species composition. Across Europe, there is a large variation in management strategies, ranging from planted forests with different rotation cycles to naturally regenerating forests that are actively managed[51]. Landscapes around PAs are generally heterogeneous, made of several forests with active management of different species and ages[52,53]. This is supported by our findings, (i.e. Fig1 c, d), showing that Canopy Top Height varies more spatially in Unprotected Surroundings than in PAs (i.e. Fig1 c, d). This indicates that forest management intensity and turnover rates are higher in unprotected areas. On the contrary, PAs mainly relate to forests with fewer disturbances, resulting in a higher homogeneity across sites. Management objectives in PAs are set around conservation targets, ranging from sites where the main goal is the strict protection of habitats and biodiversity to sites where the sustainable use of natural resources and other economic activities are allowed. As shown by our results, the mean value of individual structural variables and complexity indices was higher in PAs when compared to their unprotected surroundings. This finding, somehow expected for WDPA, is noteworthy, as it suggests less restrictive conservation measures within Natura 2000 sites are still effective at preserving the structural integrity of forests.

Second, observed differences are likely to be the result of site history and past land management[54]. While a small percentage of European PAs host primary forests that have remained stable over centuries, the vast majority of PAs were established in economically marginal lands or residual areas, that minimise conflicts with other land uses[55]. These areas comprise forests at various successional stages, following land abandonment or the cessation of management activities. Forests are particularly likely to exhibit land use legacies because they are persistent elements in landscapes due to the long lifespan of trees[56]. These, somehow idiosyncratic, differences can be seen when comparing larger areas which share common management history. The observed regional differences in our results are likely to be a reflection of the environmental and socio-economic constraints that have shaped forest ecosystems over long-time scales, and that may still have an impact on current management practices.

In addition to the abovementioned causes, we acknowledge that other factors could be driving the observed differences, mainly natural forest disturbances and large-scale ecological factors influencing species composition. However, we note our analyses are local comparisons, where each PA was compared to its unprotected surroundings. Our subsampling sampling scheme was designed to minimise confounding factors operating at these scales. To this end, we compared forests that were similar in terms of local-scale environmental conditions.

Despite the broad support for the effectiveness of PAs, there were few cases in which forests were more structurally complex outside than inside PAs. These differences may be due to several reasons. For instance, specific national PA designation policies may have favoured the establishment of PAs in areas with low soil fertility or on steeper slopes (e.g. in the Alpine region[57]). Additionally, before their establishment, forest structure within certain PAs may have been affected by its existing legislation and how intended management translates into practical reinforcement[43]. Whilst it is currently not possible to disentangle the role of many ecological and socio-economic drivers, the number of areas (Fig. 1) where PAs are effective at conserving structurally complex forests is overwhelmingly high compared to those that are not effective, suggesting a certain degree of generality in our results.

Our findings show relatively weak but positive evidence of the importance of the time of the establishment of European PAs in preserving forest structural complexity. Whilst, to some extent, the results do support the hypothesis that older PAs may harbour more structurally complex forests (i.e. in the Mediterranean region), there may be a number of unaccounted factors that could have contributed to our results. Notably, among-PA variability in the age of forests, at the time of PA establishment, is likely to have translated into uneven trajectories in the ecological succession of forest communities within PAs.

Note that this study does only evaluate “global” statistics, i.e. the spatial variance of the canopy top height and other GEDI-derived metrics within large regions (i.e. within PAs and their Unprotected Surroundings). Local statistics within smaller neighbourhoods might be more meaningful to studying the horizontal structure of forest ecosystems instead. Such analyses would be possible with GEDI in higher latitudes, where the point density is higher, but restricted at lower latitudes, where the point density is strongly reduced. Additionally, recent efforts aiming at increasing the spatial resolution of GEDI data through data fusion (see e.g.[36,58]) might circumvent this issue.

Implications for forest restoration policies. The EU’s Biodiversity Strategy to 2030[59] aims to ensure that by 2050 all terrestrial and marine ecosystems are restored, resilient and adequately protected, in line with the 2030 Agenda for Sustainable Development and the Paris Agreement on Climate Change. One of the legal instruments for achieving these objectives is the proposed Regulation on Nature Restoration[60], which sets legally binding targets for the restoration of degraded ecosystems, habitats and species. By 2030, restoration measures will cover 20% of the EU’s land and sea areas and, by 2050, all ecosystems in need of restoration.

A robust methodology for assessing and monitoring forest ecosystem conditions (i.e[61].) and multi-temporal spatially-explicit indicators representing forest traits, in particular structural aspects, is key to the success of the Regulation on Nature Restoration. The GEDI-derived indicators used in this study could represent key tools, contributing to assessing forest structural state characteristics, whereas the potential for temporal monitoring is limited. GEDI can help identify and map valuable forests for which the European Green Deal[62] and the European Biodiversity Strategy for 2030 propose increasing financial and legal support for their effective conservation.

Improvements in the framework developed here would also enable the identification of forest areas in poor or degraded ecological conditions, which in turn can be subject to restoration measures. Moreover, operational LiDAR data could provide harmonised indicators for assessing the effects of restoration measures on forest structural characteristics and conditions.

Furthermore, GEDI offers great potential in higher latitudes, where the point density is highest, for understanding the response of horizontal and vertical structural diversity to disturbance across scales. This can make an important contribution to predicting at different scales - from sample to landscape scales - how diversity might respond to ongoing changes in natural disturbance regimes[63].

Another interesting aspect to analyse in the future is the use of the skewness of the distribution of canopy height to create an indicator that might help in assessing the impacts of forest management.

Opportunities for Monitoring Forest Vertical Structure across large areas through rapid, scalable, and accurate remote sensing. In recent years, LiDAR data have played a pivotal role in the study of biodiversity at regional and national scales. This includes studies aimed at mapping functional diversity[29], correlating forest structural parameters to ground-based species diversity and abundance data[64–66].

The use of LiDAR data has overcome the traditional limitations imposed by passive optical sensors, which are not able to capture specific forest structural parameters such as Canopy Top Height and FHD. Furthermore, the availability of spaceborne LiDAR sensors such as GEDI is revolutionising our understanding of the role of habitat structure at unprecedented spatial scales. Continental-level assessments of forest vertical structural complexity such as the one presented here are now feasible. These can potentially be repeated at regular intervals and they can be used for monitoring and informing conservation policies. In agreement with others[67], we envisage that spaceborne LiDAR data will play an essential role in developing a suite of Essential Biodiversity Variables from remote sensing data.

Furthermore, the use of LiDAR data is not only relevant to assessing the effectiveness of Protected Areas, but also to the expansion of the global network of PAs in light of the upcoming post-2020 biodiversity targets[68]. For instance, GEDI could be employed as an indicator of tree species richness[33,69]. However, for future global ecosystem monitoring with spaceborne LiDAR, the scientific community will need to find alternatives to GEDI, which has a lifetime limited to about 2019 – 2022 and a latitudinal coverage limited to ± 51.6°, to achieve a global, continuous spaceborne LiDAR coverage (e.g.[70]). Given the unprecedented information provided by GEDI, there are efforts to extend its lifetime ([71]) but space agencies should soon consider a long-term, operational LiDAR mission dedicated to forest monitoring to continue the path opened by GEDI. Recent research efforts are addressing these crucial issues of restricted geographical coverage and limited capabilities for temporal monitoring by fusing GEDI data with other satellite sensor data such as Landsat and Sentinel-2(see e.g.[36,58]).

We hope that future studies will address more questions related to the fitness-for-purpose of GEDI for monitoring forest structure at different spatial scales. These should include methodological aspects, related to the adequacy of footprint size and the spatial frequency of the soundings. A combination of GEDI and airborne LiDAR and/or ground-based data is likely to increase our understanding of GEDI's capabilities, especially in terms of its ability to discriminate the differences in structural features of highly biodiverse sites from those subject to intensive management.

Conclusions

In conclusion we show that forests within PAs are on average 2 m taller and have more vertical structural complexity than those in unprotected areas. Furthermore, spatial variability in height is larger in Unprotected Surroundings than in PAs, showing the footprint of management practices in forests outside PAs. Importantly, our study demonstrates the utility of spaceborne LiDAR, beyond the study of forest vertical structure per se, showing that it can be used as a powerful tool for evaluating the condition of forest habitats within PAs.

Our study constitutes a notable advancement revealing the possibility of developing new avenues for monitoring the effectiveness of PAs, paving the way for a demonstrative framework for monitoring changes in forest habitat quality and identifying areas that may be the subject of restoration measures. As major international policy initiatives, such as the post-2020 global biodiversity framework and the EU Biodiversity Strategy, set ambitious targets aimed at halting biodiversity loss and restoring ecosystems by 2030, information on the state forest habitats is needed more than ever. We anticipate that LiDAR data will play a fundamental role in providing the required evidence base for meeting these targets and we hope that the collaboration between remote sensing scientists, ecologists and biodiversity policy experts will help to find effective solutions[72,73] needed to tackle some of the grand environmental challenges of our times.

Methods

GEDI. The NASA GEDI is a spaceborne LiDAR instrument operating onboard the International Space Station from April 2019 until at least January 2023. One of its main scientific objectives is to map forest structural properties and understand the effects of vegetation structure on biodiversity[31]. It provides sparse measurements of vegetation structure, including forest canopy height, over an area defined by a sampling footprint of ~25 m wide. These footprint data are geolocated with a mean positional error of 10.3 m. GEDI is mounted on the International Space Station (ISS) and is measuring along its ground tracks, collecting data globally between 51.6° N and 51.6° S latitudes - i.e. the inclination of the International Space Station orbit. Data is not equally distributed in space, as more samples are available in the northern part of the GEDI data range due to the convergence of International Space Station ground tracks.

For our analysis, we used GEDI version 2 which consists of 1) GEDI L2A[74] Elevation and Height Metrics; and 2) GEDI L2B[75] Canopy Cover and Vertical Profile Metrics. The data represent waveform return metrics for each 25 m diameter GEDI footprint. The footprint data are geolocated and have expected positional errors of 10–20 m (i.e. horizontal geolocation accuracy, see[31] section 4.3.1). For each footprint, we extracted a set of relative height (RH) metrics and the Foliage Height Diversity (FHD). RH metrics represent the height (in meters) at which a percentile of the laser's energy is returned relative to the ground. For example, RH50 = 20 m means that 50% of the laser's energy was returned by objects up to 20 m above the ground. RH metrics are saved at 1% intervals, so each shot contains 101 values representing RH at 0%–100%. RH98 corresponds to the canopy top height (hereafter "canopy top height") as this is a more stable height metric than RH100[76]. The vertical accuracy is on the order of 50 cm[31]. The foliage height diversity (FHD) is a canopy structural index describing the vertical heterogeneity of the foliage profile[37]. FHD is calculated from the vertical profile and is a measure of the complexity of the vertical canopy structure. Canopy Top Height and FHD are key morphological traits. Additionally, FHD is describing ecosystem structure, an essential biodiversity variable and a strong predictor of species richness at local to global scales[69]. In this study, we considered target variables RH98, RH50, RH25, and FHD.

In addition to the individual GEDI structural metrics, we quantified structural complexity using two synthetic indices following the methods described in[77]. These were calculated using all the abovementioned GEDI variables (i.e. RH98, RH50, RH25, and FHD) and include an index that measures alpha diversity (or within-sample structural diversity) and an index that measures beta diversity (or among-sample structural diversity). Alpha diversity quantifies the within-sample structural diversity for a given PA or its Unprotected Surroundings. Beta diversity characterizes the differences in structural composition among the samples of PAs or their Unprotected Surroundings.

Alpha diversity was quantified using the Shannon Diversity Index[78] (hereinafter referred to as Shannon Diversity Index). Specifically, we quantified alpha diversity for each GEDI sample. The alpha diversity values for a given area (a PA or its unprotected surroundings) were subsequently averaged. Beta diversity was calculated by taking the mean value of the pairwise dissimilarities between all samples for a given PA (hereinafter referred to as Dissimilarity). A Euclidean distance matrix was used to obtain pairwise dissimilarities. All the variables were normalized before the calculations. All calculations were done using the "vegan" and "stats" packages in the R software for statistical computing (R Core Team[79]).

Canopy heights and cover metrics for the first three years of GEDI observations (April 2019 – August 2021) are currently processed and available for analysis. We downloaded the GEDI L2A and L2B version 2 data that intersects our study regions through NASA's Earthdata Search website. GEDI data come with a series of flags and properties to help the user filter for data of quality appropriate for the specific application. Only GEDI footprints labelled as good waveform quality (i.e. quality flag equals 1) without the potential for degraded geolocation under suboptimal operating conditions (i.e. degrade flag equals 0) were used. We assumed that on-orbit waveforms with erroneous RH metrics due to low signal-to-noise ratio would be flagged in L2A. Also, we discarded footprints with a cover less than 0.1 (i.e. GEDI cover product is defined as the percentage of ground covered by the vertical projection of canopy material such as leaves, branches, and stems), winter and autumn acquisitions (i.e. shots acquired during January to March and October to December) where broadleaf forests might encounter issues due to leaf-off. Finally, we filtered GEDI samples when the top canopy height was less than 5 m to be compliant with the FAO forest definition[38]. There were a total of 49.364.247 GEDI samples over the entire EU after this first quality check.

The GEDI metric, filtered when the canopy top height (i.e. RH98) is greater than 5 m, still does not discriminate between the height of trees and man-made objects (buildings, power lines). Also, GEDI performances within steep slopes are poor. For these very reasons, we decided to cross GEDI data with satellite optical data and other ancillary datasets. Each GEDI sample was represented by a vector feature describing the spatial delineation of the LiDAR footprint, i.e. a square with a 50 m side. We created this squared footprint to accommodate the geolocation errors of the GEDI sampler. Each GEDI sample was uploaded on Google Earth Engine. For each GEDI sample footprint, we extract other data such as NDVI from Sentinel 2, slope, and aspect and we filter only samples belonging to Protected Areas or their relative 10 km buffers (i.e. Unprotected Surroundings).

Only GEDI footprints that fell inside the PAs or the 10 km buffer (see section Protected Areas) were used. This resulted in a total of 26.9 Million and 39.7 Million footprints that were considered in this study within PAs of Natura 2000 and WDPA - including all the different protection categories - respectively.

The data with a total volume of 19 Terabytes was stored on the Big Data platform (BDAP[80]) of the Joint Research Centre.

Sentinel 2. GEDI retrievals include man-made objects (e.g. buildings) and need to be filtered based on vegetation presence. For this reason, we used the Sentinel 2 optical satellites that are part of the European Copernicus program (European Commission funding, European Space Agency implementation). Sentinel 2 is a high-resolution Multi-Spectral Instrument with 13 spectral bands spanning from the visible and the near-infrared to the short-wave infrared. The spatial resolution varies from 10 m to 60 m, depending on the spectral band.

Over the study area, we created a mosaic of Sentinel 2 images for the year 2020 from May to September, using Level 1 C (i.e. Top of Atmosphere reflectance). Since images are contaminated with clouds and aerosols which need to be masked out before further processing, we applied cloud masking using the quality band (i.e. band QA60 and the Sentinel-2: Cloud Probability[81]). Then, we computed the Normalized Difference Vegetation Index (NDVI) and we extracted for each GEDI footprint the mean value of NDVI. If this value is less than 0.4 - thus indicating that the vegetation over the footprint is almost absent - then the GEDI sample was discarded; otherwise the GEDI sample is kept for analysis.

Shuttle radar topography mission (SRTM). To be conservative, and following[32], we filtered GEDI samples that lie in slopes above 10° as they likely lead to increased errors in canopy height. For this reason, we used slope information from the Shuttle Radar Topography Mission (SRTM)[82]. SRTM is a digital elevation dataset that provides consistent, high-quality elevation data. As done for NDVI, for each GEDI sample, we retained only measurements where the SRTM-derived slope is less than 10°. Note that the GEDI dataset already comes with terrain height retrieval, with better accuracy than the SRTM product[83]. However, the slope information - necessary to filter out samples that lie on steep slopes - is missing.

High-resolution layer forest type. The High-Resolution Layer Forests Type product[84] provides land cover classification with 3 thematic classes: 1) all non-tree covered areas, 2) broadleaved and 3) coniferous at 10 m spatial resolution at European level. The production of the High-Resolution Forest layers was coordinated by the European Environment Agency (EEA) in the frame of the EU Copernicus programme. As previously done for NDVI and slope, for each GEDI footprint, we extracted the mode of the High-Resolution Layer Forests Type, thus indicating the dominant leaf type. Information from High-Resolution Layer Forests Type has been only used to quantify differences in canopy top height between Protected Areas and Unprotected Surroundings for broadleaved and coniferous forests in Fig. S3.

Protected areas and 10 km buffer zones - unprotected surroundings. We took into account two different PAs classification systems and databases: (i) the World Database on Protected Areas (WDPA) and (ii) Natura 2000.

Natura 2000[39] is the European database of protected areas that collects "Special Areas of Conservation" and "Special Protection Areas" designated under the Habitats Directive and the Birds Directive, respectively. Natura 2000 is not classified by management objectives, therefore PAs can vary considerably in their protection levels. Traditional silvicultural activities are allowed to continue as before if they have no negative impact on the species or habitat types for which the site has been designated.

With regards to the WDPA, it is the most comprehensive global database on protected areas, representing *de facto* the global standard for PAs. It is a joint project between the United Nations Environment Programme and the International Union for Conservation of Nature (IUCN). The WDPA is compiled in collaboration with over 600 data providers ranging from governments to individuals. PAs are classified according to different categories - called IUCN classes - based on their management objectives and are labelled using seven classes, ranging from strict protection (Class Ia) to multiple uses (Class VI) (Table 1). Due to the inherent variability of data submitted by a wide range of data providers with different capacities and resources to digitise protected area boundaries, issues with the accuracy of the WDPA data should be expected. Throughout this study, we mainly used WDPA Protected Areas belonging to classes Ia, Ib and II to have a baseline for forests where human pressure is limited. The other IUCN categories (>= III) may have active management oriented to different conservation objectives but still can affect forest structure[85].

It is currently not possible to accurately identify overlapping areas between Natura 2000 and WDPA over Europe since Natura 2000 sites generally overlap with several protected areas classified within WDPA[86]. We, therefore, used two separate databases, the Natura 2000, which comprises the WDPA sites and the WDPA one which includes information on the protection levels, according to the IUCN criteria.

To compare changes within and around the PAs, we considered, around each PA, a 10 km unprotected buffer zone. This zone covered the land around each PA that did not overlap with any other PA, hereafter referred to as the Unprotected Surroundings. This buffer distance accounts for both positive and negative spillover effects[87] and was recently used over PAs in Tanzania along with GEDI shots[88].

In this study, only PAs of 1 km^2 or larger were considered. Also, we excluded the following PAs. First, all PAs with forest area less than 50% using the Hansen map of Global Forest Change[89] (version 1.8). Second, we excluded all marine PAs. Third, we excluded PAs with a "proposed" or "not reported" status in the WDPA, in line with common practice for global PA analyses (e.g.[90]). Using these filters, we removed Croatia which was not reporting IUCN classes and for this reason, Croatia is not included in the WDPA analysis. Applying all these selection criteria produced a final dataset of 9890 and 7831 PAs for Natura 2000 and WDPA, respectively (Fig.S8).

Since many analyses on WDPA took into consideration only PAs with high degree of protection - with IUCN Classes equal to Ia, Ib, and II, the number of WDPA is reduced to 391 PAs sites, thus rendering difficult a robust statistical analysis. Yet, the presence of WDPA is crucial for our analysis because of the IUCN classification that allows us to have a benchmark for PAs degree of protection.

Sampling of GEDI data. Comparing changes within and around PAs using an approach solely based on a fixed buffer zone might include environmental characteristics substantially different from the PA when increasing the distance to the edges of the PA[91]. For instance, using only a fixed buffer the environmental conditions might be different within and around a given PA. Therefore, to ensure a set of comparable observations of forest structure metrics and hence, allow a comparison of like with like, we used a sampling approach. Specifically, we downsampled GEDI points of the Unprotected Surroundings based on the i) elevation and ii) the distance to the centroid of the Protected Area of forest structure metrics samples. This was done because there are generally more GEDI samples in the 10 km buffer (nB) than in the PA (nP). To avoid the negative effects of an unbalanced dataset, we applied a hybrid downsampling approach so that all classes (protected areas and Unprotected Surroundings) have always the same frequency as the minority class. To avoid issues linked to canopy heights measured at different altitudes, for each site we considered valid in the Unprotected Surroundings only GEDI samples whose elevation is lying within the 10th and 90th percentiles of the elevation of GEDI samples within the Protected Area. The reason for this "percentile preprocessing" is that there is the possibility of altitudinal gradients: e.g. PAs might be on the top of mountains where the temperature is much lower than in its 10-km Unprotected Surroundings. This in turn might influence the canopy top heights, with trees taller in Unprotected Surroundings only for temperature reasons.

We considered only sites where there are at least 30 GEDI observations inside the Protected Area and 30 observations in the Unprotected Surroundings. We consider 30 observations sufficient to have a robust estimate of the vegetation structure. If the area of the site is greater than 100 km^2, the minimum number of observations is 100 instead of 30. By doing so, we try to ensure a denser sampling.

Specifically, we downsampled GEDI points of the Unprotected Surroundings based on the elevation and the distance to the centroid of the protected area.

First, each GEDI sample in the Unprotected Surroundings was ranked based on the distance to the centroid of the protected area. If the elevation range (i.e. the difference between the highest and the lowest elevation of GEDI samples in the PA) is less than 150 m, then we keep the nP closest GEDI points of the Unprotected Surroundings. If the elevation range is greater than 150 m, then we divide the dataset into 100-m elevation bins (from i to N) and for each, we keep the closest nP_i GEDI points of the Unprotected Surroundings. There is also the (less likely) case when there are more samples in the Protected Areas than in the Unprotected Surroundings. In this case, we ranked the GEDI samples inside the Protected Area according to the distance from the centroid and we retained only the first nB samples. We employed a 150 m elevation threshold to minimize the effect of the elevation on the canopy top height as already done with the percentile preprocessing.

Statistical analysis. We used a Generalized Linear Mixed Modelling (GLMM) framework to quantify differences in forest structure between Protected Areas and Unprotected Surroundings. Such models are useful when the data are clustered, e.g. GEDI shots nested in different sites (see Methods section). GLMMs were used to account for spatial dependencies between observations within the same site i.e. a given protected area and the Unprotected Surroundings (i.e. the 10 km buffer around it). The site was therefore inserted as a random effect in our models. Given that our response variables are strictly positive and continuous, an appropriate statistical distribution for the residual errors needs to be chosen. We carried out an exploratory analysis, by which we fitted gamma, lognormal and Weibull distributions to GEDI's measurements of canopy top height and FHD, to explore which distribution is better at describing the structural variables. We used the Aikake Information Criterion (AIC) to compare which statistical distribution fitted best our dataset. We found that Weibull was the best fit for both canopy top height and FHD (Fig. S9 where only the Pannonian biogeographical region is shown for the sake of simplicity).

The model takes the general form:

$$Forstruct_{ij} \sim Weibull(\alpha, \beta) \quad (1)$$

$$\mu_{ij} = \beta\Gamma(1 + 1/\alpha) \quad (2)$$

$$E(Forstruct_{ij}) = \mu_{ij} \quad (3)$$

$$\mathrm{var}(Forstruct_{ij}) = \beta^2 \left[\Gamma(1 + 2/\alpha) - [\Gamma(1 + 1/\alpha)]^2\right] \quad (4)$$

$$\mu_{ij}^{-1} = \eta_{ij} \quad (5)$$

$$\eta_{ij} = \alpha + X_{ijy} \quad (6)$$

$$a_i N(0, \sigma^2_{Site}) \quad (7)$$

where:

$Forstruct_{ij}$ is the jth observation of site i for a given structural variable

X_{ijy} is a two-level categorical variable describing whether a given observation falls within a Protected Area or the Unprotected Surroundings.

a_i is a random intercept following a normal distribution with zero mean and variance.

Models were fitted within a Bayesian framework. We fitted the models using the programming language Stan via the *brms* package in the R software for statistical computing (R Core Team[79]). In brief, Stan fits models using a stochastic Markov chain Monte Carlo (MCMC) algorithm No-U-Turn sampler, using a variant of the Hamiltonian Monte Carlo (HMC) sampling algorithm. The algorithm is programmed to take a guided random walk through parameter space[92]. Models were run using 4 chains, each with 4,000 iterations, with a warm-up of 1,000. We used the *brms* default priors for our models. These are set to be weakly informative and have been designed to provide moderate model regularisation and help stabilise computation[92,93].

Convergence was visually assessed using trace plots and using the *Rhat* values (i.e. the ratio of the effective sample size to the overall number of iterations, with values close to one indicating convergence). We checked model residuals for evidence of spatial autocorrelation using a variogram analysis[94]. Specifically, we compared variograms derived from residuals of a simple GLM without random effects *versus* the residuals of the GLMM. The results confirmed that spatial autocorrelation was much reduced when using the GLMM framework (Fig. S10 where only the Pannonian biogeographical region is shown for the sake of simplicity).

Cloud-computing platform: Google Earth Engine. Google Earth Engine[95] is a cloud-based infrastructure that enables "access to high-performance computing resources for processing very large geospatial datasets". It consists of "a multi-petabyte analysis-ready data catalogue co-located with a high-performance, intrinsically parallel computation service". The data catalogue hosts a repository of geospatial datasets, including the Sentinel 2 optical satellites and the Shuttle Radar Topography Mission (SRTM) map. Also, Google Earth Engine allows the upload of data in the form of a raster or vector. In this study, we uploaded all the GEDI footprints as vectors. All data extraction for this study was performed in Google Earth Engine, which provides the ability to compute GEDI footprint statistics and analyze the entire data records with high computational efficiency.

Data availability
To ensure full reproducibility and transparency of our research, we provide all of the data analysed during the current study. The data are permanently and publicly available on the Figshare repository https://doi.org/10.6084/m9.figshare.22128503.v1.

Code availability
To ensure full reproducibility and transparency of our research, we provide all of the scripts used in our analysis. Codes used for this study (i.e. R and javascript scripts) are permanently and publicly available on the Figshare repository https://doi.org/10.6084/m9.figshare.22128503.v1.

Received: 1 September 2022; Accepted: 15 March 2023;
Published online: 29 March 2023

References

1. Jung, M. et al. The global exposure of species ranges and protected areas to forest management. *Divers. Distrib.* **28**, 1487–1496 (2022).
2. Guo, W.-Y. et al. High exposure of global tree diversity to human pressure. *Proc. Natl. Acad. Sci.* **119**, e2026733119 (2022).
3. McNeely, J. A. Protected areas for the 21st century: working to provide benefits to society. *Biodivers. Conserv.* **3**, 390–405 (1994).
4. secretariatGlobal Assessment Report on Biodiversity and Ecosystem Services. IPBES secretariat https://ipbes.net/global-assessment (2019).
5. UNSDG. https://unstats.un.org/sdgs/dataportal.
6. Xu, H. et al. Ensuring effective implementation of the post-2020 global biodiversity targets. *Nat. Ecol. Evol.* **5**, 411–418 (2021).
7. Pioneering proposals to restore Europe's nature by 2050. *European Commission - European Commission* https://ec.europa.eu/commission/presscorner/detail/en/ip_22_3746.
8. Bergen, K. M. et al. Remote sensing of vegetation 3-D structure for biodiversity and habitat: Review and implications for lidar and radar spaceborne missions. *J. Geophys. Res.* **114**, G00E06 (2009).
9. McElhinny, C., Gibbons, P., Brack, C. & Bauhus, J. Forest and woodland stand structural complexity: Its definition and measurement. *For. Ecol. Manag.* **218**, 1–24 (2005).
10. Valbuena, R., Maltamo, M., Mehtätalo, L. & Packalen, P. Key structural features of Boreal forests may be detected directly using L-moments from airborne lidar data. *Remote Sens. Environ.* **194**, 437–446 (2017).
11. Moran, C. J., Rowell, E. M. & Seielstad, C. A. A data-driven framework to identify and compare forest structure classes using LiDAR. *Remote Sens. Environ.* **211**, 154–166 (2018).
12. Crespo-Peremarch, P., Fournier, R. A., Nguyen, V.-T., van Lier, O. R. & Ruiz, L. Á. A comparative assessment of the vertical distribution of forest components using full-waveform airborne, discrete airborne and discrete terrestrial laser scanning data. *For. Ecol. Manag.* **473**, 118268 (2020).
13. Hagar, J. C., Yost, A. & Haggerty, P. K. Incorporating LiDAR metrics into a structure-based habitat model for a canopy-dwelling species. *Remote Sens. Environ.* **236**, 111499 (2020).
14. Feng, G., Zhang, J., Girardello, M., Pellissier, V. & Svenning, J.-C. Forest canopy height co-determines taxonomic and functional richness, but not functional dispersion of mammals and birds globally. *Glob. Ecol. Biogeogr.* **29**, 1350–1359 (2020).
15. Jucker, T. et al. Canopy structure and topography jointly constrain the microclimate of human-modified tropical landscapes. *Glob. Change Biol.* **24**, 5243–5258 (2018).
16. Skidmore, A. K. et al. Environmental science: Agree on biodiversity metrics to track from space. *Nature* **523**, 403–405 (2015).
17. Reis, C. R. et al. Forest disturbance and growth processes are reflected in the geographical distribution of large canopy gaps across the Brazilian Amazon. *J. Ecol.* **110**, 2971–2983 (2022).
18. MacArthur, R. H. & MacArthur, J. W. On Bird Species Diversity. *Ecology* **42**, 594–598 (1961).
19. Roll, U., Geffen, E. & Yom-Tov, Y. Linking vertebrate species richness to tree canopy height on a global scale. *Glob. Ecol. Biogeogr.* **24**, 814–825 (2015).
20. Aguirre-Gutiérrez, J. et al. Butterflies show different functional and species diversity in relationship to vegetation structure and land use. *Glob. Ecol. Biogeogr.* **26**, 1126–1137 (2017).
21. Bae, S. et al. Taxonomic, functional, and phylogenetic diversity of bird assemblages are oppositely associated to productivity and heterogeneity in temperate forests. *Remote Sens. Environ.* **215**, 145–156 (2018).
22. Migliavacca, M. et al. The three major axes of terrestrial ecosystem function. *Nature* **598**, 468–472 (2021).
23. Bonan, G. B. Forests and Climate Change: Forcings, Feedbacks, and the Climate Benefits of Forests. *Science* **320**, 1444–1449 (2008).
24. Neigh, C. S. R. et al. Deciphering the Precision of Stereo IKONOS Canopy Height Models for US Forests with G-LiHT Airborne LiDAR. *Remote Sens* **6**, 1762–1782 (2014).
25. Montesano, P. M. et al. The use of sun elevation angle for stereogrammetric boreal forest height in open canopies. *Remote Sens. Environ.* **196**, 76–88 (2017).
26. European Commission, Joint Research CentreFlorio, P., Kakoulaki, G. & Martinez, A. *Non-commercial Light Detection and Ranging (LiDAR) data in Europe*. (Publications Office). https://doi.org/10.2760/212427 (2021).
27. Lefsky, M. A. A global forest canopy height map from the Moderate Resolution Imaging Spectroradiometer and the Geoscience Laser Altimeter System. *Geophys. Res. Lett.* **37**, L15401 (2010).
28. Simard, M., Pinto, N., Fisher, J. B. & Baccini, A. Mapping forest canopy height globally with spaceborne lidar. *J. Geophys. Res.* **116**, G04021 (2011).
29. Asner, G. P. et al. Airborne laser-guided imaging spectroscopy to map forest trait diversity and guide conservation. *Science* **355**, 385–389 (2017).
30. Fricker, G. A., Wolf, J. A., Saatchi, S. S. & Gillespie, T. W. Predicting spatial variations of tree species richness in tropical forests from high-resolution remote sensing. *Ecol. Appl.* **25**, 1776–1789 (2015).
31. Dubayah, R. et al. The Global Ecosystem Dynamics Investigation: High-resolution laser ranging of the Earth's forests and topography. *Sci. Remote Sens.* **1**, 100002 (2020).
32. Dubayah, R. et al. GEDI launches a new era of biomass inference from space. *Environ. Res. Lett.* https://doi.org/10.1088/1748-9326/ac8694 (2022).
33. Marselis, S. M., Keil, P., Chase, J. M. & Dubayah, R. The use of GEDI canopy structure for explaining variation in tree species richness in natural forests. *Environ. Res. Lett.* **17**, 045003 (2022).
34. Fisher, C. et al. Creating an Earth Archive. *Proc. Natl. Acad. Sci.* **119**, e2115485119 (2022).
35. Wieczynski, D. J. et al. Improving landscape-scale productivity estimates by integrating trait-based models and remotely-sensed foliar-trait and canopy-structural data. *Ecography* **8**, e06078 (2022).
36. A high-resolution canopy height model of the Earth. https://langnico.github.io/globalcanopyheight/.
37. MacArthur, R. H. & Horn, H. S. Foliage Profile by Vertical Measurements. *Ecology* **50**, 802–804 (1969).
38. Global Forest Resources Assessment 2020.pdf.
39. European Commission, Directorate-General for Environment & Sundseth, K. Natura 2000: protecting Europe's biodiversity. (European Commission). https://doi.org/10.2779/45963 (2008).
40. Home. *Protected Planet* https://www.protectedplanet.net/en.
41. Starnes, T. et al. The extent and effectiveness of protected areas in the UK. *Glob. Ecol. Conserv.* **30**, e01745 (2021).
42. Geldmann, J., Manica, A., Burgess, N. D., Coad, L. & Balmford, A. A global-level assessment of the effectiveness of protected areas at resisting anthropogenic pressures. *Proc. Natl. Acad. Sci.* **116**, 23209–23215 (2019).
43. Rodrigues, A. S. L. & Cazalis, V. The multifaceted challenge of evaluating protected area effectiveness. *Nat. Commun.* **11**, 5147 (2020).
44. Lines, E. R., Fischer, F. J., Owen, H. J. F. & Jucker, T. The shape of trees: Reimagining forest ecology in three dimensions with remote sensing. *J. Ecol.* **110**, 1730–1745 (2022).
45. Gough, C. M., Atkins, J. W., Fahey, R. T., Hardiman, B. S. & LaRue, E. A. Community and structural constraints on the complexity of eastern North American forests. *Glob. Ecol. Biogeogr.* **29**, 2107–2118 (2020).
46. Forrester, J. A., Mladenoff, D. J., Gower, S. T. & Stoffel, J. L. Interactions of temperature and moisture with respiration from coarse woody debris in experimental forest canopy gaps. *For. Ecol. Manag.* **265**, 124–132 (2012).

47. Ehbrecht, M. et al. Global patterns and climatic controls of forest structural complexity. *Nat. Commun.* **12**, 519 (2021).
48. Forzieri, G. et al. Emergent vulnerability to climate-driven disturbances in European forests. *Nat. Commun.* **12**, 1081 (2021).
49. Jactel, H. & Brockerhoff, E. G. Tree diversity reduces herbivory by forest insects. *Ecol. Lett.* **10**, 835–848 (2007).
50. Xu, X., Huang, A., Belle, E., De Frenne, P. & Jia, G. Protected areas provide thermal buffer against climate change. *Sci. Adv*. **8**, eabo0119 (2022).
51. Lesiv, M. et al. Global forest management data for 2015 at a 100 m resolution. *Sci. Data* **9**, 199 (2022).
52. Hardiman, B. S. et al. Spatial Variation in Canopy Structure across Forest Landscapes. *Forests* **9**, 474 (2018).
53. Wulder, M. A., White, J. C., Andrew, M. E., Seitz, N. E. & Coops, N. C. Forest fragmentation, structure, and age characteristics as a legacy of forest management. *For. Ecol. Manag.* **258**, 1938–1949 (2009).
54. Fahey, R. T. et al. Defining a spectrum of integrative trait-based vegetation canopy structural types. *Ecol. Lett.* **22**, 2049–2059 (2019).
55. Venter, O. et al. Bias in protected-area location and its effects on long-term aspirations of biodiversity conventions. *Conserv. Biol.* **32**, 127–134 (2018).
56. Munteanu, C. et al. Legacies of 19th century land use shape contemporary forest cover. *Glob. Environ. Change* **34**, 83–94 (2015).
57. Boyle, J. R. FOREST SOILS. in Encyclopedia of Soils in the Environment (ed. Hillel, D.) 73–79 (Elsevier). https://doi.org/10.1016/B0-12-348530-4/00033-3 (2005).
58. Potapov, P. et al. Mapping global forest canopy height through integration of GEDI and Landsat data. *Remote Sens. Environ.* **253**, 112165 (2021).
59. EU. COMMUNICATION FROM THE COMMISSION TO THE EUROPEAN PARLIAMENT, THE COUNCIL, THE EUROPEAN ECONOMIC AND SOCIAL COMMITTEE AND THE COMMITTEE OF THE REGIONS Stepping up Europe's 2030 climate ambition Investing in a climate-neutral future for the benefit of our people. (2020).
60. Nature restoration law. https://environment.ec.europa.eu/publications/nature-restoration-law_en.
61. Ecosystem Accounting | System of Environmental Economic Accounting. https://seea.un.org/ecosystem-accounting.
62. EU. COMMUNICATION FROM THE COMMISSION TO THE EUROPEAN PARLIAMENT, THE EUROPEAN COUNCIL, THE COUNCIL, THE EUROPEAN ECONOMIC AND SOCIAL COMMITTEE AND THE COMMITTEE OF THE REGIONS The European Green Deal COM/2019/640 final. (2019).
63. Seidl, R. et al. Forest disturbances under climate change. *Nat. Clim. Change* **7**, 395–402 (2017).
64. Müller, J. et al. LiDAR-derived canopy structure supports the more-individuals hypothesis for arthropod diversity in temperate forests. *Oikos* **127**, 814–824 (2018).
65. Müller, J. & Brandl, R. Assessing biodiversity by remote sensing in mountainous terrain: the potential of LiDAR to predict forest beetle assemblages. *J. Appl. Ecol* **46**, 897–905 (2009).
66. Bakx, T. R. M., Koma, Z., Seijmonsbergen, A. C. & Kissling, W. D. Use and categorization of Light Detection and Ranging vegetation metrics in avian diversity and species distribution research. *Divers. Distrib.* **25**, 1045–1059 (2019).
67. Skidmore, A. K. et al. Priority list of biodiversity metrics to observe from space. *Nat. Ecol. Evol.* **5**, 896–906 (2021).
68. Visconti, P. et al. Protected area targets post-2020. *Science* **364**, 239–241 (2019).
69. Marselis, S. M. et al. Evaluating the potential of full-waveform lidar for mapping pan-tropical tree species richness. *Glob. Ecol. Biogeogr.* **29**, 1799–1816 (2020).
70. Hancock, S., McGrath, C., Lowe, C., Davenport, I. & Woodhouse, I. Requirements for a global lidar system: spaceborne lidar with wall-to-wall coverage. *R. Soc. Open Sci*. **8**, 211166.
71. Greenfield, P. Extend life of key climate sensor that maps world's forests, Nasa told. *Guardian*, https://www.theguardian.com/environment/2022/mar/20/nasa-urged-to-extend-life-of-key-climate-sensor-that-maps-worlds-forests-gedi-aoe (2022).
72. Díaz, S. et al. Set ambitious goals for biodiversity and sustainability. *Science* **370**, 411–413 (2020).
73. Leclère, D. et al. Bending the curve of terrestrial biodiversity needs an integrated strategy. *Nature* **585**, 551–556 (2020).
74. GEDI L2A Vector Canopy Top Height (Version 2). https://doi.org/10.5067/GEDI/GEDI02_A.002.
75. GEDI L2B Vector Canopy Cover Vertical Profile Metrics (Version 2). https://doi.org/10.5067/GEDI/GEDI02_B.002.
76. Blair, J. B. & Hofton, M. A. Modeling laser altimeter return waveforms over complex vegetation using high-resolution elevation data. *Geophys. Res. Lett.* **26**, 2509–2512 (1999).
77. Senf, C., Mori, A. S., Müller, J. & Seidl, R. The response of canopy height diversity to natural disturbances in two temperate forest landscapes. *Landsc. Ecol.* **35**, 2101–2112 (2020).
78. A Mathematical Theory of Communication - Shannon - 1948 - Bell System Technical Journal - Wiley Online Library. https://onlinelibrary.wiley.com/doi/10.1002/j.1538-7305.1948.tb01338.x.
79. R Core Team. R: A Language and Environment for Statistical Computing. (2022).
80. Soille, P. et al. A versatile data-intensive computing platform for information retrieval from big geospatial data. *Future Gener. Comput. Syst.* **81**, 30–40 (2018).
81. Skakun, S. et al. Cloud Mask Intercomparison eXercise (CMIX): An evaluation of cloud masking algorithms for Landsat 8 and Sentinel-2. *Remote Sens. Environ.* **274**, 112990 (2022).
82. Jarvis, A., Reuter, H. I., Nelson, A. & Guevara, E. Hole-Filled SRTM for the Globe Version 4.CGIAR-CSI SRTM 90 m Database.
83. Liu, A., Cheng, X. & Chen, Z. Performance evaluation of GEDI and ICESat-2 laser altimeter data for terrain and canopy height retrievals. *Remote Sens. Environ.* **264**, 112571 (2021).
84. Forests — Copernicus Land Monitoring Service. https://land.copernicus.eu/pan-european/high-resolution-layers/forests.
85. *Fourth ministerial conference on the protection of forests in europe. Vienna resolution 4 conserving and enhancing forest biological diversity in Europe*. (2003).
86. Lawrence, A., Friedrich, F. & Beierkuhnlein, C. Landscape fragmentation of the Natura 2000 network and its surrounding areas. *PLoS ONE* **16**, e0258615 (2021).
87. Joppa, L. N. & Pfaff, A. Global protected area impacts. *Proc. R. Soc. B Biol. Sci.* **278**, 1633–1638 (2011).
88. Liang, M. et al. Assessing protected area's carbon stocks and ecological structure at regional-scale using GEDI lidar. *Glob. Environ. Change* **78**, 102621 (2023).
89. Hansen, M. C. et al. High-Resolution Global Maps of 21st-Century Forest Cover Change. *Science* **342**, 850–853 (2013).
90. Saura, S. et al. Protected area connectivity: Shortfalls in global targets and country-level priorities. *Biol. Conserv.* **219**, 53–67 (2018).
91. Gao, Y. et al. Control selection for the assessment of protected areas in the Hengduan Mountains: A case study in Yunlong Tianchi National Nature Reserve, China. *Glob. Ecol. Conserv.* **23**, e01170 (2020).
92. Gelman, A., Hill, J. & Vehtari, A. *Regression and Other Stories*. (Cambridge University Press). https://doi.org/10.1017/9781139161879 (2020).
93. Gelman, A. Prior distributions for variance parameters in hierarchical models (comment on article by Browne and Draper). *Bayesian Anal* **1**, 515–534 (2006).
94. Cressie, N. *Statistics for Spatial Data, Revised Edition*. (Wiley, 2015).
95. Gorelick, N. et al. Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote Sens. Environ.* **202**, 18–27 (2017).
96. Biogeographical regions, Europe 2016, version 1. *EEA geospatial data catalogue* https://sdi.eea.europa.eu/catalogue/srv/api/records/c6d27566-e699-4d58-a132-bbe3fe01491b.

Acknowledgements

The authors would like to thank the JRC Big Data Platform (BDAP)[80] for the support provided.

Author contributions

G.C. and M.G. conceived the idea and designed the methodology; G.C. and M.G. analyzed the data and wrote the Google Earth Engine and R scripts; G.C., M.G., G.Duv., M.M., V.A. and A.C. wrote the manuscript with contributions from P.B., J.B., and G.Dub. and L.B. All authors contributed critically to the interpretation of the results and gave final approval for publication.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s43247-023-00758-w.

Correspondence and requests for materials should be addressed to Guido Ceccherini or Marco Girardello.

Peer review information *Communications Earth & Environment* thanks the anonymous reviewers for their contribution to the peer review of this work. Primary Handling Editor: Aliénor Lavergne. Peer reviewer reports are available.