Forest Ecology and Management 589 (2025) 122763

Contents lists available at ScienceDirect

Forest Ecology and Management

journal homepage: www.elsevier.com/locate/foreco

# Predicting woody species assemblages using ecophylogenetics and Earth observation data

Alexander Wellenbeck [a,e,*], Nils Hein [b,c], David Tarkhnishvili [b,d], Bernhard Misof [a,e], Sebastian Schmidtlein [f], Zurab Janiashvili [g], Lasha Dzadzamia [d,h], Hannes Feilhauer [i,j,k]

[a] Chair of Systematic Zoology, University Bonn, Bonn, Germany
[b] Caucasus Leibniz Biodiversity Center (LBiC), Ilia State University, Tbilisi, Georgia
[c] Institute of Geography Education, University of Cologne, Cologne, Germany
[d] Department of Evolutionary Biology and Ecology, Ilia State University, Tbilisi, Georgia
[e] Leibniz Institute for the Analysis of Biodiversity Change (LIB), Museum Koenig Bonn, Germany
[f] Institute of Geography and Geoecology, Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany
[g] Department of Biodiversity and Forestry, Ministry of Environmental Protection and Agriculture of Georgia, Tbilisi, Georgia
[h] Deutsche Gesellschaft Für Internationale Zusammenarbeit (GIZ) GmbH, Tbilisi, Georgia
[i] Institute for Earth System Science and Remote Sensing, Leipzig University, Leipzig, Germany
[j] German Center for Integrative Biodiversity Research (iDiv), Halle-Jena-Leipzig, Germany
[k] Helmholtz-Center for Environmental Research (UFZ), Leipzig, Germany

ARTICLE INFO

*Keywords:*
Biodiversity monitoring
Species diversity
Forest community classification
Community assembly
Phylogenetic diversity
Random forests
Forest inventory analysis
Remote-sensing ecology

ABSTRACT

Organizing species assemblages based on compositional characteristics enables the identification of ecologically meaningful patterns in biodiversity and supports forest diversity monitoring, conservation, and management. In this context, ecophylogenetics offers powerful opportunities by exploring how evolutionary relationships between species reflect community distributions within ecological space. Using national forest inventory data of Georgia (Sakartvelo), we classify woody species assemblages based on interspecies phylogenetic dissimilarity and evaluated whether cluster membership could be predicted from multivariate Earth observation data describing site-specific environmental conditions. Principal components of 30 explanatory variables were used to model class membership across three sample groups with increasing disturbance levels. Prediction accuracy reached 53.6 % (OOB error 46.4 %) for undisturbed samples, 67.5 % for disturbed (OOB 32.5 %), and 45.7 % for disturbed samples with neophytes (OOB 54.3 %), based on 12, 6, and 5 clusters, respectively. The decline in classification accuracy with increasing disturbance reflects compositional homogenization and a weakened alignment of the phylogenetic signal with environmental gradients. Our findings demonstrate that incorporating phylogenetic variability in the classification of woody species assemblages enables coherent clustering and effectively captures distributions along environmental gradients particularly under low-disturbance conditions. This approach offers a solid framework to improve forest community classification and to support sustainable forest and conservation management.

## 1. Introduction

Forest ecosystems support terrestrial biodiversity by providing habitat and sustaining essential ecological processes. Reflecting this importance, recent approaches have advanced forest biodiversity monitoring through the integration of ecological indicators into national forest inventories (Gillerot et al., 2021; Heym et al., 2021). Large scale, repetitive forest inventories support comprehensive biodiversity monitoring systems by providing systematically assessed data on community composition, stand structure and characteristics such as dead wood, tree habitats and other ecological components outside the scope of conventional forestry objectives (Newton and Kapos, 2002; Chirici et al., 2012; Godoy and Rueda, 2016; Reise et al., 2019; Ette et al., 2023). Data generated by these assessments enable the quantification of various aspects of community diversity and temporal biodiversity changes in forest ecosystems and are available worldwide (Winter et al.,

* Corresponding author at: Chair of Systematic Zoology, University Bonn, Bonn, Germany.
*E-mail address:* alex.wellenbeck@uni-bonn.de (A. Wellenbeck).

https://doi.org/10.1016/j.foreco.2025.122763
Received 5 March 2025; Received in revised form 24 April 2025; Accepted 26 April 2025
Available online 8 May 2025

2008; Corona et al., 2011; Traub and Wüest, 2020; Heym et al., 2021). Woody species are the main contributors to biomass in forest ecosystems and are a key component of forest biodiversity (Zhou et al., 2021). Forest inventories usually record all woody species observations occurring within a standardized sample plot area, according to predefined inclusion thresholds (Henttonen and Kangas, 2015). Thus recorded woody species assemblages represent subsets of the regional species pool filtered through multiple abiotic and biotic mechanisms, like species' habitat and dispersal limitations, and local dynamics, such as disturbances or species competing for space and resources (Pavoine and Bonsall, 2011; Swenson, 2011b). Organizing woody species assemblages according to shared compositional or other characteristics supports conservation and forest management and increases our understanding of community distributions across spatial scales (Hao et al., 2021). The field of ecophylogenetics extends the scope of compositional characteristics to the phylogenetic structure of species assemblages providing means to infer the assembly mechanisms that shape community distribution patterns from the phylogenetic relationships of their constituents in a multidimensional ecological space (Davies, 2021).

The non-stochastic distribution patterns of woody species assemblages are shaped by assembly rules defined by scale-dependent environmental drivers. At finer spatial scales (e.g., on sample plot or stand level), co-occurring species are often distantly related, reflecting limiting similarity and competitive exclusion (Cavender-Bares and Wilczek, 2003). At coarser spatial scales however (e.g., across regions), habitat heterogeneity becomes more pronounced and closely related species tend to cluster due to niche conservatism, as environmental filtering selects similar species suited to particular conditions (Webb, 2000; Webb et al., 2002; Holt, 2009; Ascanio et al., 2024). Thus, species' phylogenetic dissimilarity reflects niche differences as a result of abiotic constraints which interact with biotic factors that shape diversity patterns (Gerhold et al., 2015; Kraft et al., 2015; Cadotte, 2017; Cadotte and Tucker, 2017) and determine which species co-inhabit sites with similar conditions, forming habitat specific communities (Weiher and Keddy, 2001; Webb et al., 2006; Norberg et al., 2019). The interrelation between increasing spatial distance and species or community dissimilarity is well established, and consistent with Tobler's first law of geography (Tobler, 1970). This relationship is attributed to the decrease in environmental similarity along contemporary environmental gradients or dispersal limitations and niche width differences among taxa, mediated by evolutionary diversification (Bosch et al., 2021; Kusumoto et al., 2021). Consequently, the phylogenetic structure of species assemblages is spatially dependent and provides a proxy metric that characterizes community responses to environmental conditions. In ecology, this approach is used to disentangle the relative roles of environmental filtering, competitive exclusion and biogeographical processes that shape community structure (Webb et al., 2002; Emerson and Gillespie, 2008; Xu et al., 2019; Davies, 2021). Since phylogenetic dissimilarity reflects species' functional traits, this approach offers potential for distinguishing species compositions by organizing communities in a way that incorporates natural evolutionary processes shaped by environmental filtering mechanisms. The notion that the inherent genetic signal of species assemblages reflects abiotic site conditions suggests that species ecological demands can be inferred from environmental variables governing niche differentiation and community distributions (Gilbert and Parker, 2022). Increasing our understanding of the interrelations between evolutionary processes and niche development bears significant potential for applications in ecologically sound sustainable forest management and forest diversity monitoring (Davies, 2021).

In forest ecosystems, niche differentiation among woody species assemblages arises from variations in species' light, water, and nutrient needs, as well as specific traits of seed dispersal strategies and responses to disturbances and succession (Szymura et al., 2015; Lausch et al., 2019; Akobia et al., 2022). Adaptive limitations result in often overlapping but distinct niche occupancies, causing species turnover along gradients. This turnover leads to gradual change of the phylogenetic signal which, if quantified as dissimilarity, can serve as proxy to explain assembly mechanisms and model resulting distribution patterns. At intermediate scales (i.e. landscape level), classifications based on generalized species assemblages can effectively characterize community distributions to enable modeling and mapping of species assemblages through Earth observation (EO) data (Ferrier and Guisan, 2006; Feilhauer et al., 2011; Feilhauer et al., 2012; Kuenzer et al., 2014). The growing availability of optical remote-sensing data with improved spatial, spectral, and temporal resolution, and advanced data managing tools enable cost-effective analyses of large areas at regular intervals (Foody and Cutler, 2003; Gillespie et al., 2008; Rocchini et al., 2018). When combined with systematic sampling across environmental gradients, it allows for comprehensive assessments of the relative influence of climate, soil, geomorphology and dispersal limitations on species turnover (i.e., β-diversity) of forest communities (Hernández-Stefanoni et al., 2012). Examining the explanatory power of environmental gradients in defining classes of species assemblages provides insights into the processes that constrain patterns of community compositions and improve predictive distribution mapping (Gilbert et al., 2024).

We use data from the first National Forest Inventory of Georgia (Sakartvelo, GNFI), where a relatively high proportion of forest area (44.5 %, MEPA, 2023) and a comparably high degree of conservation regarding natural species distribution and diversity in the respective biome (Novák et al., 2023; Stritih et al., 2024) provide a unique opportunity to link species diversity data with multivariate remote-sensing proxies to test whether diversity patterns can be predicted via environmental gradients within this complex forest ecosystem. We hypothesize that at the regional scale, the composition of woody species assemblages exhibits a distinct phylogenetic signal that reflects habitat conditions and enables the prediction of assemblage membership from site characteristics derived from EO data. To test this hypothesis, we address the following research questions: i) Does hierarchical discrimination of compositional data, considering phylogenetic variability, reflect ecological niche distribution along environmental gradients? ii) Does increasing disturbance intensity influence the prediction accuracy of species assemblages in classification models? iii) Does the relative importance of environmental predictors differ for modeling woody species assemblage membership across sample groups with varying disturbance levels?

## 2. Data and methods

To understand how phylogenetically informed species classifications reflect ecological patterns under varying disturbance regimes, we applied a multistep approach combining hierarchical clustering with classification modeling using environmental gradients derived from EO data. By classifying GNFI subsamples according to quantitative disturbance intensities, we compare how disturbance shapes the resulting cluster hierarchy, coherence and species-cluster relationships. After reducing multicollinearity through principal component analysis (PCA) of EO variable groups, we used the thematic principal components (PC) to model cluster membership via Random Forest (RF) classification. We evaluated prediction accuracy and variable importance across models to examine how phylogenetically informed clustering reflects ecological gradients and responds to disturbance.

### 2.1. Study area

Georgia lies between 41°07' - 43°35'N and 40°04' - 46°44'E and borders the Greater Caucasus to the north and the Lesser Caucasus to the south (Fischer et al., 2018; Nakhutsrishvili et al., 2023; Cortner et al., 2024). The country is characterized by dominant hilly to rugged mountainous terrain with roughly 55 % of the national area located at elevations exceeding 1000 m above sea level (a.s.l.) and around 40 % on slopes with $\geq 20°$ inclination (Mikeladze et al., 2020). The

predominantly mountainous topography of the country represents a geographically diverse region with pronounced environmental gradients that hosts species of independently evolving lineages (Tarkhnishvili et al., 2012; Tarkhnishvili, 2014; Dering et al., 2021). The large number of endemic species found within a relatively limited spatial extent, render Georgia an area of high priority for conservation as part of the Caucasus biodiversity hotspot (Zazanashvili et al., 2001; Myers, 2003; Joppa et al., 2011; Mittermeier et al., 2011). The Greater and Lesser Caucasus moderate Georgia's climate, with mountainous regions experiencing mean annual temperatures of −5 °C – 10 °C and 800 – 1400 mm precipitation (Keggenhoff et al., 2014). Rising to 1000 m a.s.l., the central north-south running Likhi Range acts as a natural climatic boundary that separates the humid, warm climate of Western Georgia (13 – 15°C, <400 – >4000 mm) from the increasingly continental climate of Eastern Georgia (10 – 13 °C, 500 – 600 mm, Denk et al., 2001; Elizbarashvili et al., 2006). Existing forest associations range from Alpine coniferous forests dominated by *Abies nordmanniana* (Steven) Spach. and *Picea orientalis* (L.) Peterm. at higher elevations to open juniper woodland (dominated by *Juniperus polycarpos excelsa* subsp. *polycarpos* (K. Koch) Takht. and *J. foetidissima* Willd.) distributed along the drier areas in the Southeast. At lower elevations, Colchic alder carrs (*Alnus glutinosa* subsp. *barbata* (C. A. Mey.) Yalt) and humid temperate broadleaf forests, including Sweet Chestnut (*Castanea sativa* Mill.), characterize the Western lowlands of the country. Thermophilus to xerophytic mixed oak forests occupy large parts of central Georgia (*Quercus petrea* subsp. *iberica* (Steven ex M. Bieb.) Krassiln., *Carpinus betulus* L., and *Carpinus orientalis* Mill.). Extensive Oriental beech (*Fagus orientalis* Lipsky.) and hornbeam-oriental beech forests complement the main forest associations existent in the country (Dolukhanov, 2010; Patarkalashvili, 2017; Nakhutsrishvili et al., 2021; Novák et al., 2023).

## 2.2. Data

To prepare the dataset for phylogenetically informed clustering and classification, we filtered GNFI plots by forest continuity, species richness, and taxonomic resolution. Sample plot composition was quantified as basal area per species and weighted by interspecies phylogenetic distances.

### 2.2.1. Sampling design of the GNFI

The GNFI is based on a systematic sampling grid of 3.6 km x 3.6 km with a randomly selected origin. Field observations are recorded as cluster samples comprising three sample plots of 0.07 ha each, arranged in an L-shaped configuration with a distance of 100 m along both axes (Fig. 1). Woody species are recorded according to any stems' respective DBH. Stems with DBH ≥ 8 cm are recorded on the inner nested subplot radius of 5 m, whereas stems with DBH ≥ 15 cm and DBH ≥ 30 cm are recorded on subplots with r = 10 m and r = 15 m, respectively (Fig. 1). The GNFI data contains numerous variables describing stand characteristics, i.e. disturbances, and number of stumps (MEPA, 2018). Stumps are measured on the entire sample area (r = 15 m) and classified according to origin ("natural" or "anthropogenic"). Disturbance is recorded per type (e.g. low basal area density, non-systematic wood extraction, etc.) and severity class (1−3). As 18 % of the country's territory is currently not accessible for government officials due to an ongoing political conflict, only approximately 74 % of the national forest area of Georgia (2278,760 ha, Fig. 2) was sampled (MEPA, 2023).

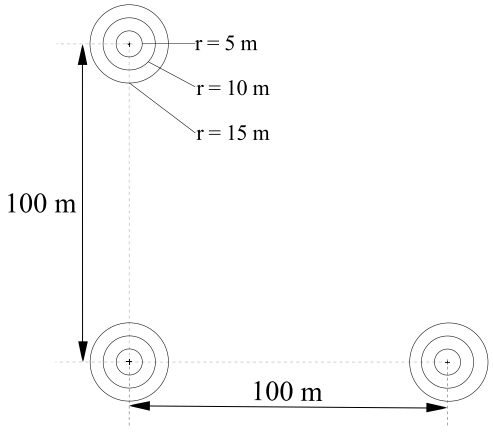


**Fig. 1.** Configuration of cluster plots comprising three sample plots of the National Forest Inventory of Georgia. Woody species are recorded within three nested subplots according to the measured diameter at breast height (DBH, at 1.3 m,). For each stem, tree and stem number, species (if identifiable) and DBH measurement were recorded along with the polar coordinates of the stem axis and other variables (MEPA, 2018).

The complete GNFI dataset contains n = 1773 cluster and m = 4452 sample plot observations with recorded woody species unambiguously identified at species level and are classified as "Tree covered area" and "Forest" according to the local land cover categorization (MEPA, 2023).

### 2.2.2. Data subsampling and diversity

We excluded sample plots intersected by a forest boundary to avoid bias to the compositional data as a result of edge effects (Ries et al., 2004; Willmer et al., 2022) and removed monodominant samples (S = 1), as these cannot be clustered based on species composition. The remaining data set contained 3466 sample plot records (henceforward referred to as "samples", Table 1). We stratified the data according to anthropogenic disturbance (severity > 0), presence of non-natural stumps, signs of cattle grazing or presence of neophytes. Samples with any of the aforementioned attributes > 0 were labeled as "disturbed" ($m_{dis}$ = 2931). Samples with recorded presence of neophytes were labeled as "disturbed with neophytes" ($m_{neo}$ = 197). All other samples were considered "undisturbed" ($m_{undis}$ = 535). To ensure equal sample sizes between undisturbed and disturbed sample groups, we randomly subsampled the disturbed samples ($m_{dis}$ = 535) resulting in a total subsample consisting of 1267 sample plot observations (Table 1), which account for a total sampled area of approximately 90 ha, representing 0.004 % of the total forest area of Georgia (2278,760 ha, MEPA, 2023).

Sample composition was derived as extrapolated sum of basal area (BA, $m^2$ $ha^{-1}$) per species and sample as abundance value (Staudhammer and LeMay, 2001; Cáceres et al., 2019; Yao et al., 2019; Ricotta et al., 2021). Spelling and nomenclature were standardized using the Taxonomic Backbone of World Flora Online (WFO DB, Kindt, 2020) and the Global Biodiversity Information Facility (GBIF Secretariat, 2021). Neophyte species were identified according to the list of invasive species for Georgia available on GBIF (Kolbaia et al., 2020). We constructed a phylogenetic tree by matching the standardized species list with the mega phylogeny of World Plants database (GBOTB. extended.WP.tre, Jin and Qian, 2022; Davies et al., 2023). Pairwise phylogenetic distances were calculated as total branch lengths connecting each pair of species at the terminal nodes of the created tree (Zanne et al., 2014; Kling et al., 2018; Smith and Brown, 2018; Wellenbeck et al., 2024). We calculated dissimilarities between sample plots based on the Discriminating Avalanche (dA), developed by Ganeshaiah and Uma Shaankar (2000) and refined by Hao et al. (2019b) which quantifies interspecies dissimilarity by weighting absolute differences in frequencies of species *i* and *j* in two samples with the corresponding phylogenetic distance (Table 2).

We normalized the resulting dissimilarities according to $dA_{norm} = \frac{dA - dA_{min}}{dA_{max} - dA_{min}}$ (Legendre and Legendre, 2012; Hao et al., 2019a).

### 2.2.3. Environmental Data

We compiled explanatory variables that represent environmental gradients expected to be determinant for species compositions in four thematic groups (Dolukhanov, 2010; Hawkins et al., 2014; Qian et al., 2019; Padullés Cubino et al., 2021). As climate can be seen as the

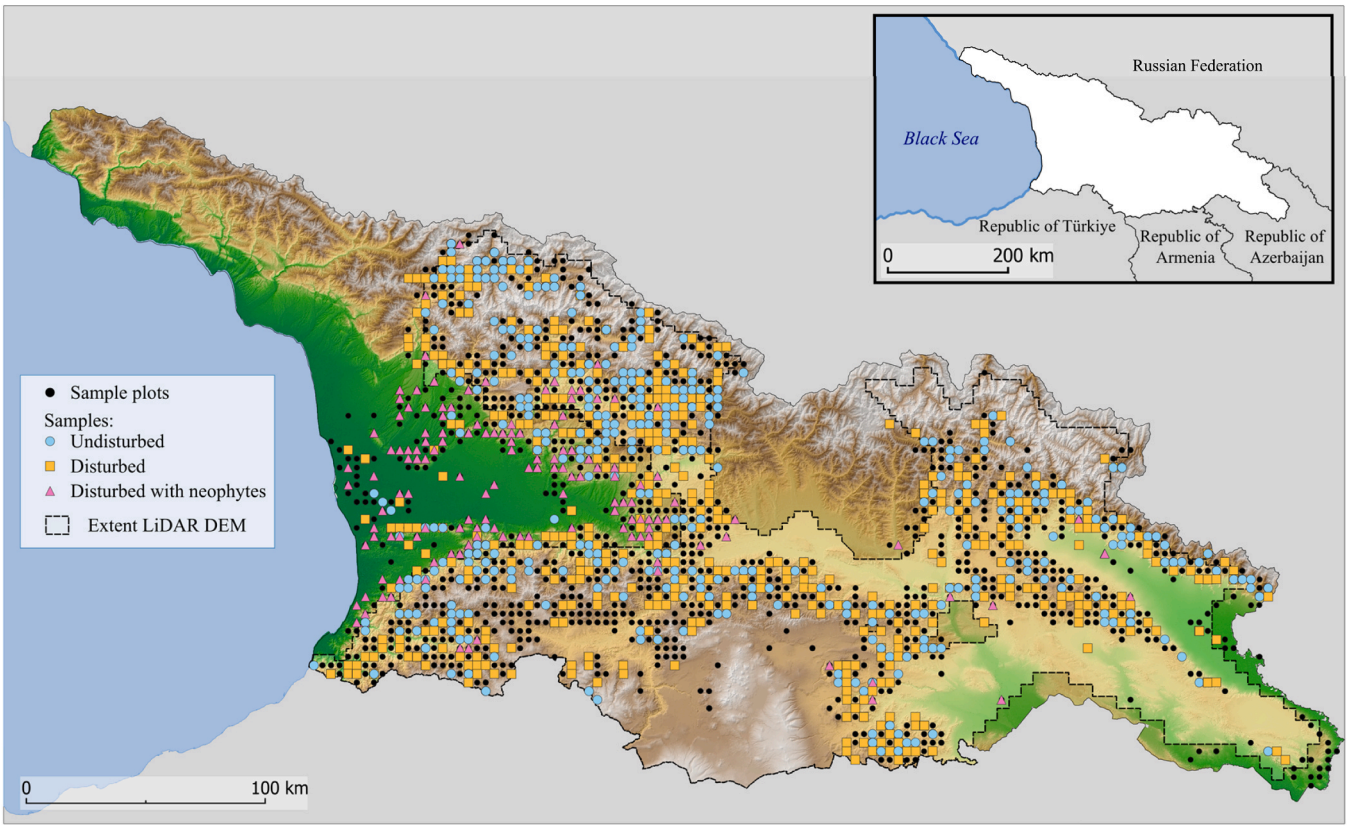


**Fig. 2.** Data overview of sample locations and applied digital elevation models. The SRTM based DEM is available for the entire country area. The dotted line outlines the extent of the LiDAR-based DEM. Sample plots (m = 4452) are shown as point features and represent all accessible sample plots located inside forest, which are not intersected by forest boundaries and contain only records of taxa identified at species or subspecies level.

**Table 1**
Summary statistics of basal area of the subsampled data of the National Forest Inventory of Georgia (2019–2021) per sample group.

| Sample Group | m | S | BA $m^2$ $ha^{-1}$ | | | | |
|---|---|---|---|---|---|---|---|
| | | | Min | Max | Mean[a] | Var.[b] | CV(%)[c] |
| **Undisturbed** | 535 | 58 | 9.14 | 111.43 | 40.04 ( ± 16.292) | 265.431 | 40.69 |
| **Disturbed** | 535 | 61 | 1.74 | 120.79 | 30.85 ( ± 15.740) | 247.746 | 51.02 |
| **Disturbed with neophytes** | 197 | 75 | 1.37 | 82.44 | 20.97 ( ± 13.922) | 193.835 | 66.39 |
| **Summary** | 1267 | 96 | 1.37 | 120.79 | 33.19 ( ± 17.084) | 291.846 | 51.47 |

[a] Mean values are denoted with standard deviation in parenthesis.
[b] Var.=Variance
[c] CV=Coefficient of variation

primary factor influencing tree species distributions (Box, 1996; Zellweger et al., 2015; Araújo et al., 2019; Coelho et al., 2023), we obtained climatic variables from the Copernicus BIOCLIM dataset (Vanuytrecht et al., 2021). To quantify edaphic site conditions, we extracted soil characteristics from the layer-specific SOILGRID database (ISRIC) and calculated mean values across all layers (Hengl et al., 2017; Poggio et al., 2021; Turek et al., 2023; Miller et al., 2024). Topographic variables were derived from two digital elevation models (DEM) of different spatial resolutions and extent (Fig. 2) to account for existing gradients at landscape scale (Sefidi et al., 2016; Gardner et al., 2019; Akobia et al., 2022; Haesen et al., 2023) and micro-topographic gradients, i.e. slope position (Siegert et al., 2016).

**Table 2**
Dissimilarity index used in this study.

| | | | | |
|---|---|---|---|---|
| Discriminating Avalanche (Hao et al., 2019a) | $dA$ | = | $\frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}\Delta_i^{a,b}d_{ij}\Delta_j^{a,b}$ | [1] |

| With | | |
|---|---|---|
| $d_{ij}$ | = | phylogenetic distance between species i and j ($d_{ij} = d_{ji}$ and $d_{ii} = 0$) |
| $\Delta_i^{a,b}$ | = | absolute difference between the frequencies of species *i* in plots a and b ($\lvert p_i^a - p_i^b \rvert$) |
| n | = | number of samples |
| $p_i^a$, $p_i^b$ | = | relative frequencies of species i in plots a and b |

**Table 3**
Formula to estimate heat load (McCune and Keon, 2002).

| | | | |
|---|---|---|---|
| *Heat load* | = | $0.339 + \cos(L) \bullet \cos(S) - 0.196 \bullet \sin(L) \bullet \sin(S) - \sin(S) - 0.482 \bullet \cos(A_f) \bullet \sin(S)$ | [2] |
| Where: | | | |
| L | = | Latitude | |
| $A_f$ | = | Folded azimuth | |
| S | = | Slope [°] | |

We derived estimates of anthropogenic impact from the human footprint raster data provided by the Wildlife Conservation Society (2005); Sanderson et al., (2022) and manually estimated pressure from anthropogenic influence based on proximity to settlements and infrastructure derived from Open Street Map (OSM, OpenStreetMap contributors, 2024). Per sample values derived from raster data were

**Table 4**
List of variables considered as predictors of species assemblages. Resolutions ranged from 5 m to 1224 m. Retained variables are marked by '* '.

| Group | Variable | Brief Description | Range | Unit | Reference | |
|---|---|---|---|---|---|---|
| Climate | BIO01 | Annual mean temp. | −0.28–15.38 | °C | (Vanuytrecht et al., 2021) | |
| | BIO02 | Mean monthly diurnal range | 4.77–10.97 | °C | | * |
| | BIO04 | Temp. seasonality (St. Dev. of the monthly mean temp.) | 17.88 – 33.99 | °C | | * |
| | BIO05 | Maximum daily temp. of the warmest month | 17.69 – 33.98 | °C | | |
| | BIO06 | Minimum daily temp. in coldest month | −4.59–12.07 | °C | | |
| | BIO08 | Mean temp. of wettest quarter | 0.37–17.64 | °C | | * |
| | BIO09 | Mean temp. of driest quarter | −8.86 – 17.91 | °C | | * |
| | BIO12 | Annual mean precipitation | 384.79–2617.71 | mm | | |
| | BIO14 | Minimum mean precipitation of driest month | 5.92 – 81.74 | mm | | |
| | BIO15 | Precipitation seasonality (CV% of monthly precipitation) | 34.14 – 68.65 | % | | * |
| | BIO16 | Mean precipitation in wettest quarter | 156.98 – 930.82 | mm/month | | |
| | BIO18 | Mean precipitation in warmest quarter | 40.15 – 777.79 | mm/month | | * |
| | pot_eva | Potential evaporation annual mean | 55.72 – 90.69.5 | mm/month | | * |
| Topographic | tri | SRTM based Topographic ruggedness index | 0.04 – 26.91 | Non-dimensional | (Riley et al., 1999) | |
| | tpi | LiDAR based Topographic position index | −0.43 – 0.57 | Non-dimensional | (Wilson et al., 2007) | * |
| | rough | SRTM based Roughness index | 0.03 – 26.85 | Non-dimensional | | |
| | g15010 | LiDAR based Geomorphon terrain form (mean): outer radius: 15 m inner radius: 10 m flatness threshold: 0 | 1.0 – 8.54 | Non-dimensional | (Jasiewicz and Stepinski, 2013) | |
| | g40005 | LiDAR based Geomorphon terrain form (mean): outer radius: 40 m inner radius: 5 m flatness threshold: 0 | 1.0 – 9.0 | Non-dimensional | | * |
| | slp | SRTM based slope | 0.1–55.7 | ° | (Fuchs et al., 2017) | * |
| | asp | SRTM based aspect | 0–360 | ° | | |
| | f_asp_hl | Folded aspect heat load | 0.18 – 179.93 | Non-dimensional | (McCune and Keon, 2002) | * |
| | dir | Heat load | 1.35 – 2.64 | Non-dimensional | | * |
| | eastness | Eastward orientation | −1–1 | Non-dimensional | | * |
| | southness | Southward orientation | −1–1 | Non-dimensional | | |
| | std_H | LiDAR based Standardized height | 12.89 – 2122.68 | Non-dimensional | (Conrad et al., 2015) | |
| | slp_H | LiDAR based Slope height | 0.54 – 589.51 | m | | |
| | norm_H | LiDAR based Normalized height | 0–1 | Non-dimensional | | * |
| | v_depth | LiDAR based Valley depth | 0.52 – 599.79 | m | | * |
| | mid_slp | LiDAR based Mid slope position | 0–1 | Non-dimensional | | |
| Soil | bd_mean | Mean bulk density | 0.99–1.52 | kg $dm^{-3}$ | (Poggio et al., 2021) | |
| | cec_mean | Mean cation exchange capacity | 11.9–36.6 | $cmol^{+}$ $kg^{-1}$ | | * |
| | soc_mean | Mean soil organic matter content | 141.2–785.0 | ‰ (g $kg^{-1}$) | | |
| | phh2o_mean | Mean pH-value in $H_2O$ solution | 4.8–7.5 | pH | | * |
| | sand_mean | Mean sand content | 5.9–45.6 | g 100 $g^{-1}$ (%) | | |
| | silt_mean | Mean silt content | 29.8 – 47.9 | g 100 $g^{-1}$ (%) | | * |
| | clay_mean | Mean clay content | 15.9–57.7 | g 100 $g^{-1}$ (%) | | * |
| | nitro_mean | Mean of total nitrogen | 1.59 – 5.99 | g $kg^{-1}$ | | |
| | cfvo_mean | Mean of volumetric fraction of coarse fragments (>2 mm) | 7.18 – 26.4 | $cm^3$ 100 $cm^{-3}$ (vol %) | | * |
| | socs_0_30 | Organic carbon stocks (mean, 0 – 30 cm) | 42.0 – 115.0 | kg $m^{-2}$ | | * |
| | wv0033 | Mean volumetric soil water retention at 330 cm | 267.03 – 379.33 | $10^{-3}$ $cm^3$ $cm^{-3}$ | (Turek et al., 2023) | * |
| | wv0010 | Mean volumetric soil water retention at 100 cm | 346.0 – 439.17 | $10^{-3}$ $cm^3$ $cm^{-3}$ | | * |
| Spatial | ele | SRTM based Elevation | 0.97–2449.4 | m | (Fuchs et al., 2017) | * |
| | X | Easting (UTM 38 N) | 213812–642200 | m | MEPA, (2018) | * |
| | Y | Northing (UTM 38 N) | 4552600–4783000 | m | | * |
| | med.xy | Normalized pairwise distance | 0–1 | Non-dimensional | | * |
| | h_imp_med | Human Impact Index | 124–3999 | Non-dimensional | (Sanderson et al., 2022) | * |
| | prox | Proximity to urban infrastructure | 0–7156 | m | (OpenStreetMap contributors, 2024) | |
| | costs | SRTM based Cumulative costs | 0–6387.7 | Non-dimensional | | * |

Notes: St. Dev. = Standard deviation; temp. = temperature (°C).

aggregated over all cells contained in or crossed by a buffer representing the circular sample plot area (r = 15 m plus the recorded GPS error [m] to account for signal inaccuracies) using the *Zonal Statistics* algorithm of QGIS (QGIS Development Team, 2009).

*2.2.3.1. Climatic variables.* The BIOCLIM dataset (*Global bioclimatic indicators from 1979 to 2018 derived from reanalysis*) contains averages of bioclimatic indicator metrics based on historical reconstructions with a spatial resolution of 1 km (Vanuytrecht et al., 2021). We evaluated 13 bioclimatic variables (Table 4) representing annual or seasonal climatic gradients reflecting limiting factors for species distributions (Gardner et al., 2019).

*2.2.3.2. Soil properties.* We obtained data from SOILGRID, which provides modeled predictions of soil properties as raster files with a resolution of 250 m (Poggio et al., 2021; Turek et al., 2023). We retrieved data for bulk density, cation exchange capacity, pH, soil organic carbon, total nitrogen, organic carbon concentration, soil texture and volumetric water content at 10 kPa and 33 kPa suction in $10^{-3}$ $cm^3$ $cm^{-3}$ (1 mm $m^{-1}$) to evaluate soil water availability (Gardner et al., 2019). For each sample, we obtained means by averaging values over the six reported depths (0–5 cm, 5–15 cm, 15–30 cm, 30–60 cm, 60–100 cm and 100–200 cm, Table 4).

*2.2.3.3. Topography.* Slope, aspect, the Terrain Ruggedness Index (TRI), and Roughness were calculated from the high-resolution Shuttle Radar Topography Mission DEM (SRTM, Farr et al., 2007) that was interpolated to a resolution of 10 m x 10 m and void filled with CGIAR-CSI SRTM ver.4.1 (Riley et al., 1999; Fuchs et al., 2017; Macek et al., 2019; Moudrý et al., 2019). We converted aspect values to "eastness" and "southness" to avoid having circular values and calculated folded aspect ($A_f = 180 - |\text{aspect} - 225|$) to estimate heat load per sample according to McCune and Keon, (2002), (Feilhauer and Schmidtlein, 2009).

Fine-scale topography was derived from a very high resolution (5 m) DEM based on a 2018 light detection and ranging (LiDAR) flight campaign covering 58 % of the total country area (Figure 2). Fine-scale topographic conditions affect the variability of available light, temperature, and soil properties which influence local species distribution patterns (Beatty, 1984; Moeslund et al., 2013; Fazlollahi Mohammadi et al., 2022; Woods and Ortmann, 2024). To account for fine-scale topographic heterogeneity, we calculated the Topographic Position Index (TPI), and classified terrain forms using the r.geomorphon algorithm (GRASS Development Team, 2022) described by Stepinski and Jasiewicz (2011) which classifies topographic structures according to geomorphologic phenotypes (Jasiewicz and Stepinski, 2013; Gioia et al., 2021) from the LiDAR data. We created *geomorphons* with an outer radius around the sample plot center of 25 m and 40 m each with 5 m as inner radius, and 5° and 0° as flatness threshold values, respectively (GRASS Development Team, 2022). Assigned pixel class numbers were averaged per sample to obtain continuous values. To quantify sample position relative to slope, we calculated "standardized height", "slope height", "normalized height", "valley depth" and "mid slope position" using the SAGA algorithm Relative height and slope position (Conrad et al., 2015) with default parameters. The algorithm follows an iterative approach to compute terrain indices based on the vertical distance above the terrain minima and standardized relief-positions, with refinement according to terrain-based watershed effects (Böhner and Selige, 2006). To reduce the influence of outliers, median values were derived for each sample.

*2.2.3.4. Spatial position and anthropogenic influence.* Recorded GPS coordinates and elevation [m a.s.l.] obtained from the SRTM DEM represent spatial positions of samples. We quantified relative isolation as median of all spatial distances obtained from a pairwise spatial distance matrix (Table 4). Distances to sources of anthropogenic influence (roads and settlements) were calculated based on OSM vector files of the road network (primary, secondary and tertiary), buildings and places within the country. Roads were buffered with a distance of 10 m, 8 m, and 4 m, respectively, while buildings were buffered by 100 m. We merged polygons to a binary raster layer (5 m) indicating anthropogenic land use and calculated overall distances as continuous buffer via the SAGA *proximity raster* algorithm (Conrad et al., 2015). Using the 'r.cost' algorithm (GRASS Development Team, 2022), we computed cost values based on a reclassified slope raster representing friction costs as geometric sequence with a common ratio of 2 (0° = 1, 1–5° = 2, 6–10° = 4, 11–20° = 8, 21–30° = 16, 31–40° = 32, 41–45° = 64, ≥45° = 'not accessible'). Using the 'zonal statistics' algorithm we queried median values for all sample areas. To quantify spatial gradients of human influence (Sanderson et al., 2022) on a broader scale, we queried the human footprint raster data (third generation, 300 m resolution) provided by the Wildlife Conservation Society (WCS, https://wcshumanfootprint.org/data-access) per sample.

### 2.3. Cluster analysis

As hierarchical cluster structures of species assemblages resembles the literature's depiction of nested forest community structures (Dolukhanov, 2010; Nakhutsrishvili, 2013; Costanza et al., 2018; Nakhutsrishvili et al., 2023), we discriminated samples in a hierarchical cluster analysis using isometric partitioning (isopam) (Schmidtlein et al., 2010). We applied an extension to the original source code of the isopam algorithm to support dA as dissimilarity index (Wellenbeck et al., 2024). Results of the isopam cluster analyses are reported with metrics quantifying cluster homogeneity (G) and its' standardization across partitions (global.Gs) and the isomap.k parameter, which determines the number of nearest neighbors used in dimensionality reduction. The phi coefficient of species fidelity quantifies the strength of association between species and clusters (Chytrý et al., 2002; Cabido et al., 2018). We calculated the standardized effect size of Mean Pairwise Distance ($SES_{MPD}$) to quantify phylogenetic dispersion within the resulting clusters (Webb et al., 2006; Swenson, 2011a).

### 2.4. Reduction and decorrelation

By applying PCA to each variable group independently, we reduced multicollinearity and created composite predictors (PCs) as thematic explanatory variables. By means of orthogonal transformation, each successive PC captures more inherent variation of the dataset, resulting in a hierarchy of cumulative variance explained (Cruz-Cárdenas et al., 2014). Prior to PCA, we applied forward selection per variable group using pairwise Spearman correlations as these are robust to nonlinear relationships and outliers (Tucker et al., 2017, supplement 1). From each variable pair with a correlation coefficient $|R| \geq 0.8$, one variable was rejected to reduce redundancy (Guisan and Thuiller, 2005; Chytrý et al., 2016; Kavgacı et al., 2023). Table 4 provides an overview of considered and retained variables for modeling.

### 2.5. Random forest modeling

RF models are flexible and perform well with complex datasets (Breiman, 2001; Genuer and Poggi, 2020) to provide robust classification of data based on an ensemble of decision trees built on randomly selected subsets, with final predictions determined by majority voting (Cutler et al., 2007; Waldock et al., 2022; Gilbert et al., 2024; Soley-Guardia et al., 2024). We predicted species assemblage membership within each sample group via RF with variable group PCs as predictors. Stratified classification was applied to ensure proportional cluster representation in bootstrap samples, with default parameters maintained for the number of trees and predictors per split (Breiman, 2001). Out-of-Bag (OOB) error metrics were used to evaluate model

**Table 5**
Formula to derive weighted variable importance according to the principal components used as predictors.

| | | | |
|---|---|---|---|
| Weighted Variable Importance | $W_i$ | $= \sum_{j=1}^{n} (\lvert L_{ij} \rvert \bullet Gini_j)$ | [3] |

With

| | | |
|---|---|---|
| $W_i$ | = | Weighted variable importance for the $j^{th}$ principal component |
| $L_i$ | = | Loadings of the $i^{th}$ variable in each principal component |
| $\lvert L_{ij} \rvert$ | = | Absolute value of the loading for variable *i* on the $j^{th}$ principal component |
| $Gini_j$ | = | Variable importance of the $j^{th}$ principal component |

performance and variable importance was quantified via the Mean Decrease in Gini index (MDG, Breiman, 2001; Afanador et al., 2016; Genuer and Poggi, 2020). Based on MDG, we selected the four most influential PCs from each variable group as explanatory variables for a combined RF classification model. Missing data were imputed using 999 imputations for 99, 96, and 150 values for undisturbed, disturbed, and disturbed with neophytes, respectively (Doove et al., 2014). The higher number of imputations for the latter reflects the concentration of these samples outside the LiDAR-based DEM (Figure 2). 15 RF models were applied in total, first identifying key PCs for each variable group individually and then combining four key PCs from all variable groups to model species assemblage membership for each sample group separately.

### 2.6. Weighted variable importance

To evaluate the contribution of variables to the most predictive PCs in the combined RF model, we quantified variable importance as the cumulative sum of MDG weighted by the corresponding PC loading values (eigenvectors) which accounts for each variable's predictive power in the model by incorporating its absolute contribution to each PC - a common method in Partial Least Squares Regression analysis (Table 5).

### 2.7. Analytical environment and software

The analytical workflow was implemented using established geospatial and statistical platforms throughout all stages of data preparation and modeling. Raster data was processed in the QGIS 3.34.8 environment using GRASS GIS 8.2 and SAGA 9.4.1 (QGIS Development Team, 2009; Conrad et al., 2015; GRASS Development Team, 2022). We used the R Base ver. 4.2.3 implemented in R Studio ver. 2024.09.0 for data analysis (R Core Team, 2024; RStudio Team, 2024). The phylogenetic tree was constructed using the R package V.PHYLOMAKER2 (Jin and Qian, 2019; 2022), PICANTE to derive $SES_{MPD}$ values (Kembel et al., 2010) and cluster analysis was performed using ISOPAM v. 2.0 (Schmidtlein et al., 2024). PCAs were conducted using the R package VEGAN (Oksanen, 2020) and RF modeling was based on the package RANDOMFOREST (Liaw and Wiener, 2002). We imputed missing values using the MICE R package (van Buuren and Groothuis-Oudshoorn, 2011).

## 3. Results

The analyzed GNFI data (m = 1267) consists of 21,303 recorded trees belonging to 96 species of 56 genera and 30 families. Fig. 3 shows the constructed phylogenetic hierarchy for all recorded taxa.

The data subsets according to disturbance contained 8557 tree records of 61 species, spanning 22 genera and 18 families in disturbed samples, whereas disturbed samples with neophytes included 2439 trees from 75 species, covering 50 genera and 26 families. Undisturbed samples consisted of 10,307 trees representing 58 species, 31 genera, and 18 families.

### 3.1. Hierarchical clustering of species assemblages

The cluster analysis of undisturbed samples ($m_{undis}$ = 535) partitioned twelve clusters across four hierarchical levels, whereas clustering of disturbed ($m_{dis}$ = 535) and disturbed samples with neophytes ($m_{neo}$ = 197) identified six and five clusters at two hierarchical levels, respectively. Synoptic tables of the resulting clusters with indicator species frequencies are provided in supplement 3. Whereas all three hierarchical cluster structures depart on three main groups on level I, showing varying levels of coherence, undisturbed samples exhibit the most complex configuration spanning four levels, with broad isomap.k values (100 for partition 3), 55 significant indicator species ($\geq$ threshold G, with $p \leq 0.05$), and a generally higher mean standardized G score (global.Gs: 3.9 ± 2.96 across seven partitions). High species-cluster associations are indicated by a mean phi-value of 0.11 ± 0.28 (n = 37). In contrast, the clusters hierarchies of disturbed samples are organized in a simpler 2-level structure with maximum isomap.k values of 92 (second partition) and 18 (third partition) for disturbed and disturbed samples with neophytes, respectively. Fewer significant indicator species can be identified for disturbed (32) and disturbed with neophytes (25), with mean standardized G scores of 5.1 ± 4.07 and 1.3 ± 1.01, respectively. Mean phi-values of significant indicator species are −0.01 ± 0.20 (n = 29) for disturbed samples and 0.07 ± 0.27 (n = 24) for disturbed samples with neophytes. The mean number of species per cluster (n = 5) was highest for disturbed samples with neophytes (35.0 ± 13.95), intermediate for disturbed samples (24.5 ± 7.79, n = 6), and lowest for undisturbed samples (19.5 ± 8.21, n = 12). $SES_{MPD}$ was highest for undisturbed clusters with −0.16 ± 1.10 (-0.33 ± 0.98 and −0.48 ± 0.73 for disturbed and disturbed samples with neophytes, respectively). As assigned clusters represent homogeneous species assemblages, we evaluated the resulting groups based on BA distributions (>10 %) and labeled them according to BA dominance (Fig. 4).

Whereas discriminated species assemblages on level II showed some similarity for undisturbed and disturbed samples (Fig. 4), clustered species assemblages from disturbed samples with neophytes were not ecologically meaningful as the resulting groups were highly heterogeneous in species BA and lacked a distinct ecological structure (supplement 4a and 4b).

### 3.2. Modeling of species assemblages

As a result of forward selection, we excluded BIO01, BIO05, BIO06, BIO12, BIO14, BIO16 from the climate, bd_mean, nitro_mean, sand_mean, soc_mean from the soil, asp, g15010, mid_slp, slp_H, Std_H, rough, tri from the topography, and prox from the spatial dataset (supplement 5). Consequently, PCAs were computed using seven climate, eight soil, nine topographic, and six spatial variables.

The first two PCs derived from the climate variables (m = 4452) represent 70.4 % of the variance, and the first four PCs together explain 93.6 % (Fig. 5). For soil variables (m = 4446), the first two PCs capture 51.9 % of the cumulative variance and 78.1 % of the variance is explained by the first four PCs. The first two PCs of the topographic variables (m = 3427) explain 54.4 % of the variance, increasing to 79.1 % with the first four PCs. 66.9 % of variance is explained by the

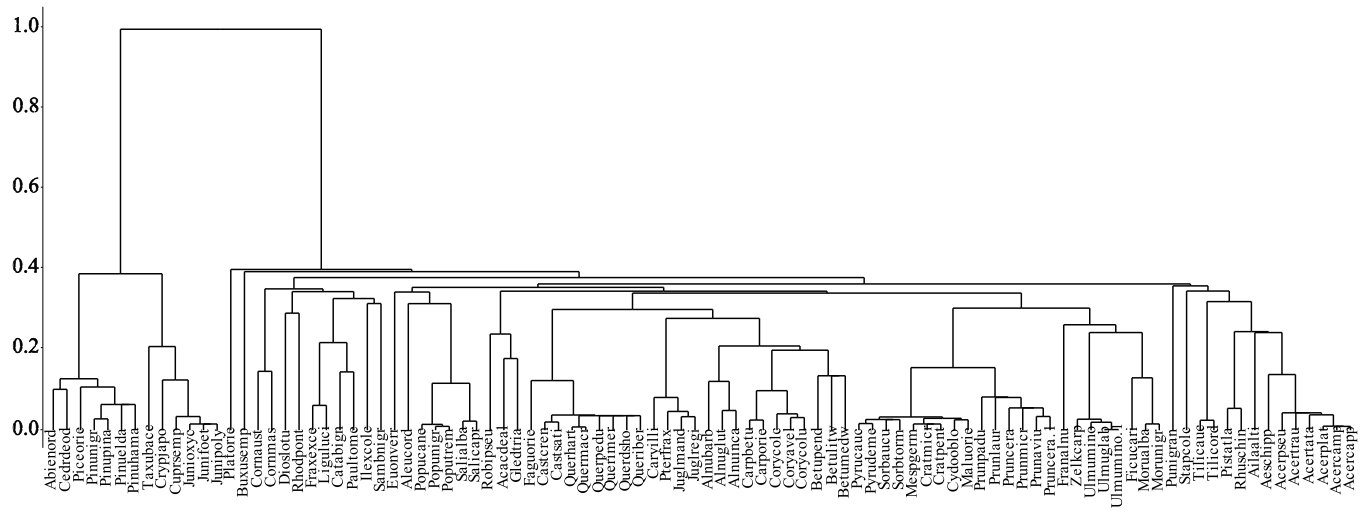


**Fig. 3.** Phylogenic tree of 96 species listed in the subsample (m = 1267) of the National Forest Inventory data of Georgia. Interspecies phylogenetic distances were normalized inside a square matrix of pairwise distance values so that $0 < \text{dist}_{ph} \leq 1$. Species names on the x-axis are abbreviated according to the Cornell Ecology Programs naming convention, using the first four letters of the genus and species name (e.g., Abies nordmanniana = Abienord). For original branch lengths see supplement 2.

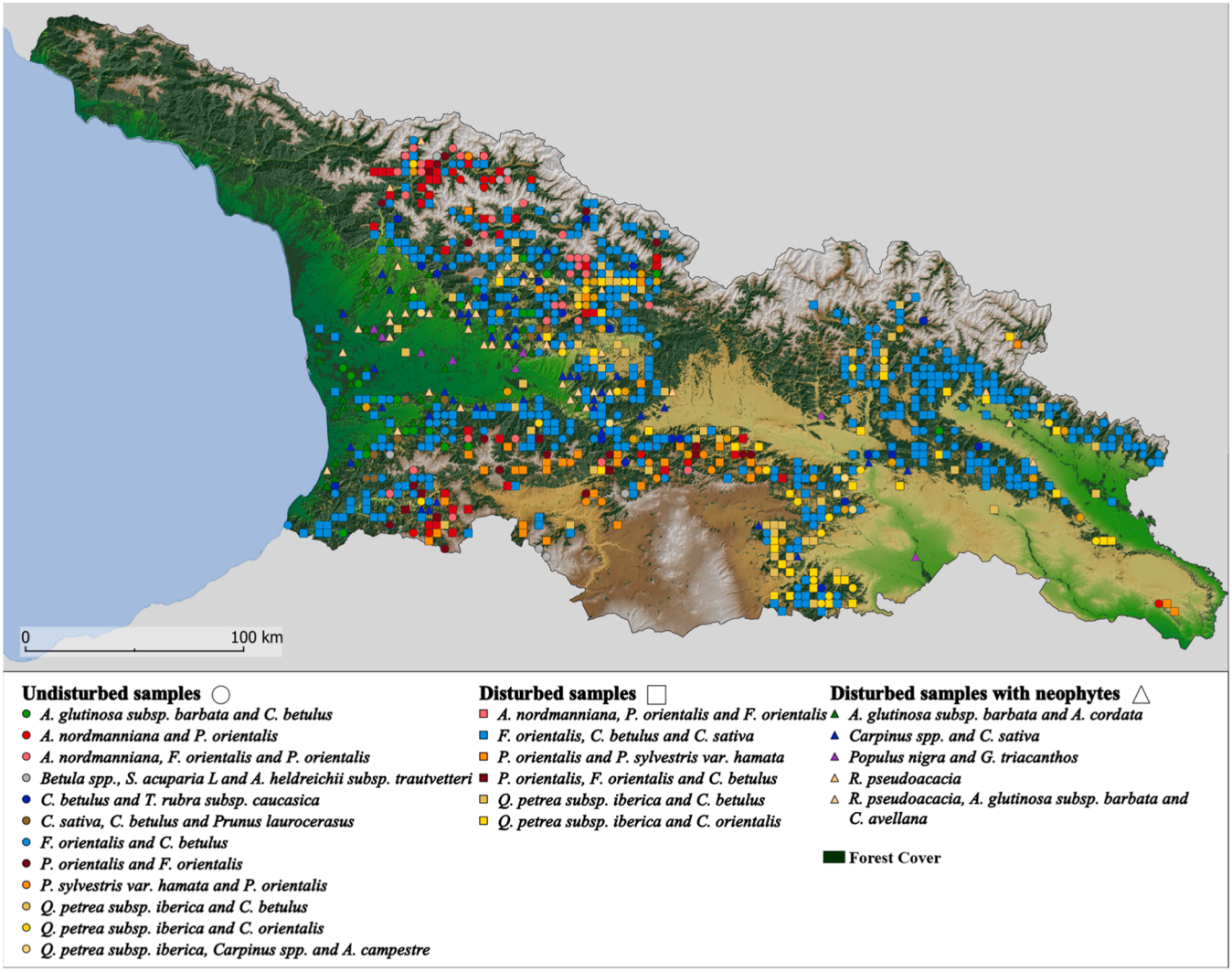


**Fig. 4.** Locations of clustered species assemblages per sample group (m = 1267). Clusters are labeled according to basal area share per species (≥10 %). Green-shaded areas indicate the distribution of forests in Georgia (Griesbach, 2018) for reference.

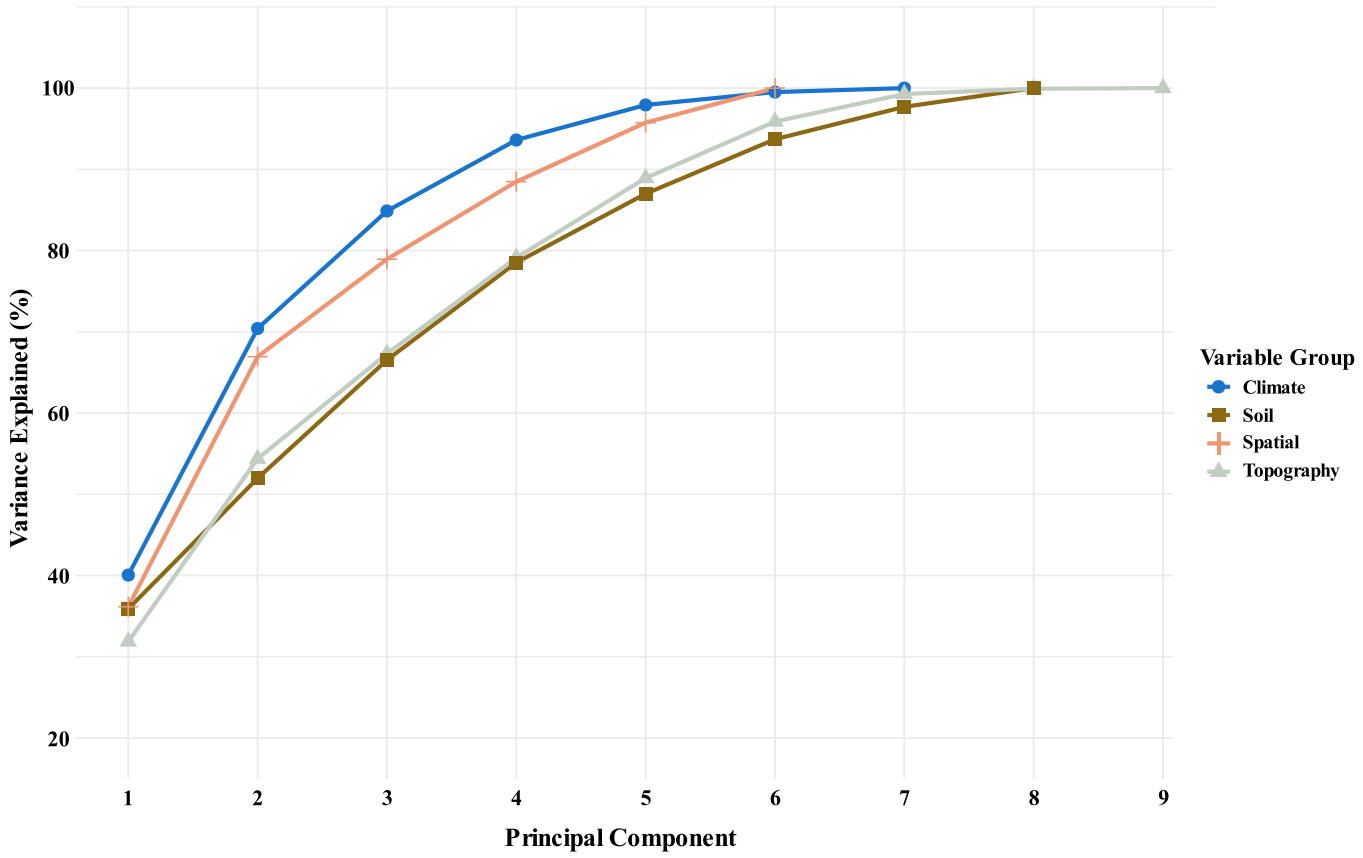


**Fig. 5.** Percentage of explained variance per principal component per variable group.

**Table 6**
Out-of-Bag error estimates for Random Forest classification of species assemblage membership across sample and variable groups. Based on the Mean Decrease Gini index, the most important principal components as predictors have been selected for model optimization.

| | Undisturbed (n = 12) | | Disturbed (n = 6) | | Disturbed with Neophytes (n = 5) | |
|---|---|---|---|---|---|---|
| **Variable group** | **OOB (%)** | **PCs as Predictors*** | **OOB (%)** | **PCs as Predictors*** | **OOB (%)** | **PCs as Predictors*** |
| Climate | 45.8 | 1, 2, 3, 6 | 34.6 | 1, 2, 3, 5 | 57.4 | 1, 2, 3, 7 |
| Soil | 46.7 | 1, 2, 6, 8 | 34.6 | 1, 2, 4, 6 | 55.6 | 1, 2, 4, 7 |
| Topography | 53.5 | 4, 5, 6, 7 | 42.7 | 1, 2, 4, 5 | 48.9 | 2, 5, 7, 9 |
| Spatial | 44.9 | 1, 2, 3, 4 | 34.8 | 1, 3, 5, 6 | 54.8 | 2, 3, 4, 6 |
| Mean OOB (%) | 47.7 ± 3.92 | | 36.7 ± 4.02 | | 54.2 ± 3.68 | |

* Selected according to ranking of variable importance based on Mean Decrease Gini.

first two PCs of the spatial variables and 88.4 % of the first four PCs. Modeling assemblage membership of undisturbed samples using the climate, soil, and spatial variable groups yielded OOB errors below 47 %, whereas the OOB error for the topography variable group was 53.5 % (Table 6).

For predicting assemblage membership of disturbed samples, the lowest OOB error was recorded for the soil and climate variable groups, both at 34.6 %. In contrast, for disturbed samples with neophytes, the topography variable group yielded the lowest OOB error at 48.9 %. Based on the respective MDG per PC we identified the four most important PCs for a combined RF classification model. The combined RF classification model resulted in total OOBs of 46.4 %, 32.5 % and 54.3 % for undisturbed, disturbed and disturbed samples with neophytes (supplement 6). Based on the combined RF model, we calculated individual variable contributions by weighting the cumulative sums of MDG with their corresponding PC loading values (Fig. 6). Variable importance was generally highest in undisturbed samples and lowest for disturbed samples with neophytes.

For predicting cluster membership from undisturbed samples, the highest variable importance values (>45) were observed for the south–north gradient (Y), elevation (ele), and precipitation during the warmest quarter (BIO18). Soil organic carbon stocks in the top 30 cm (socs_0_30) and mean pH (phh20_mean) also showed importance values exceeding 30. For disturbed samples, variables with importance values > 30 included mean temperature of the wettest quarter (BIO08), annual mean potential evaporation (pot_eva), the south–north gradient (Y), precipitation during the warmest quarter (BIO18), median pairwise spatial distance (med.xy), and mean pH (phh20_mean). For disturbed samples with neophytes, variable importance values > 15 were only recorded for the median pairwise spatial distance (med.xy), the west-east gradient (X), mean temperature of the driest quarter (BIO09), elevation (ele), and the standard deviation of monthly mean temperature (BIO04). Additionally, volumetric water content (wv0010_mean, wv0033_mean) received high variable importance values, whereas topographic variables consistently ranked lowest across all sample groups (Fig. 6).

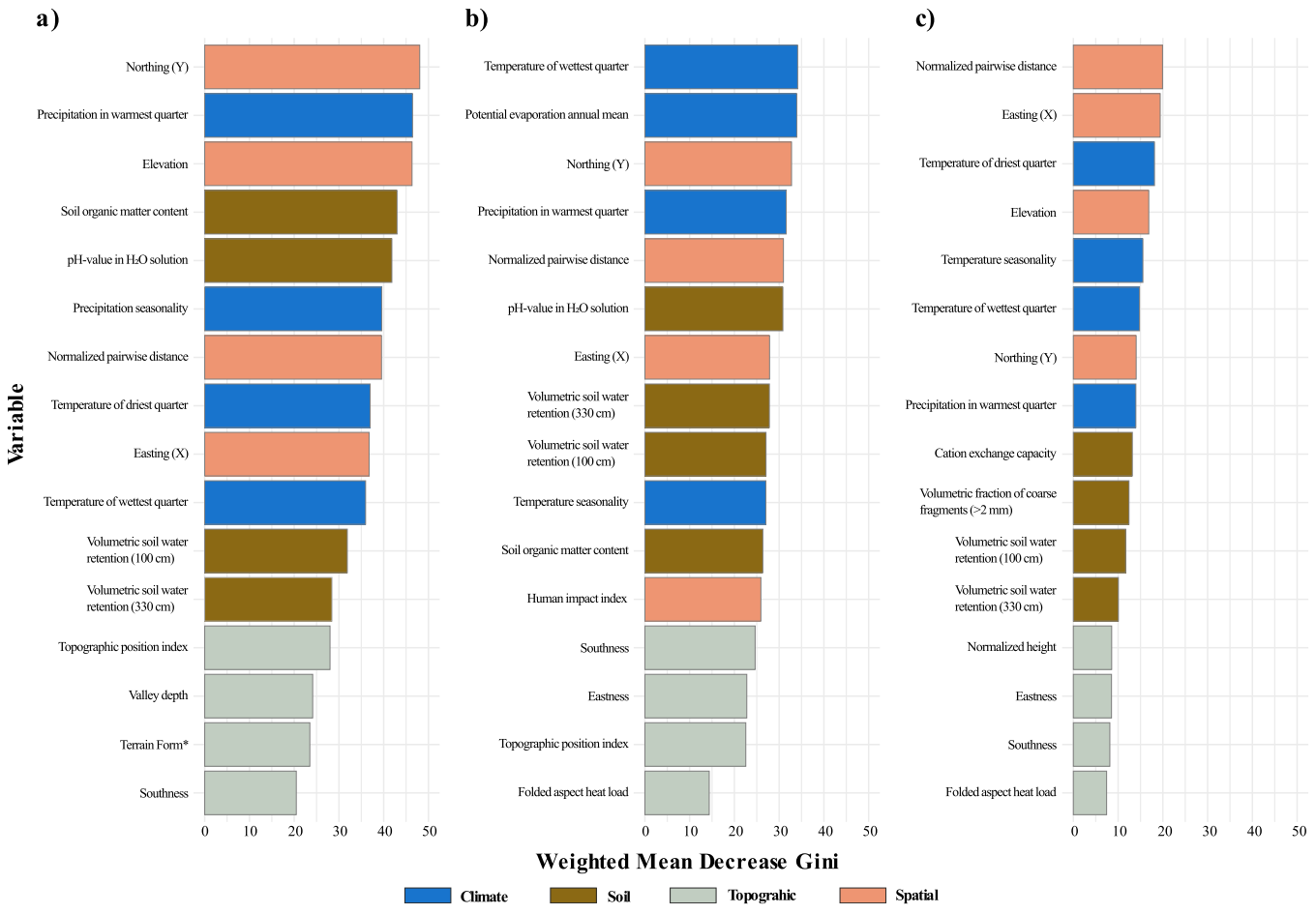


**Fig. 6.** Weighted Mean Decrease Gini indices of the four most important explanatory variables per variable and sample group. Variable importance is weighted according to contribution to the most important principal component to predict cluster membership of twelve species assemblages of undisturbed (a), six of disturbed (b), and five of disturbed samples with neophytes (c), respectively. *geomorphon with 40 m outer radius, 5 m inner radius, and 0 flatness threshold.

## 4. Discussion

We used Random Forest to predict cluster membership in disturbance-stratified sample groups, based on species assemblages derived from phylogenetically informed isometric partitioning. Classification based on site-specific environmental variables performed generally well, with model accuracies varying across variable and sample groups. Combining the four most important PCs from each variable group as predictors moderately improved overall model accuracy, resulting in OOB errors of 46.4 %, 32.5 %, and 54.3 % for undisturbed, disturbed, and disturbed samples with neophytes, respectively. These results confirm our first research question on whether phylogenetic signals in species compositions reflect environmental gradients, in line with other studies (Qian et al., 2013; Qian and Sandel, 2017; Shi et al., 2021; Zhou et al., 2021). Affirming our second research question, model accuracy was highest for undisturbed samples and lowest for samples with neophytes, if class cardinality is considered. As higher disturbance levels reduce the complexity of the resulting cluster hierarchies, predictable assembly patterns are obscured. This is supported by weakened phylogenetic divergence ($SES_{MPD}$), lower species-cluster associations (phi-values) and reduced cluster coherence (higher global G scores). The effects of disturbance in our subsample appear stochastic in type and severity but impede meaningful ecological discrimination beyond six clusters, despite being based on the same number of observations as the undisturbed subsample (m = 535). Disturbances reduce tree density, favor dominance by fewer species, or filter for genetically similar species, and lead to a homogenization of compositional characteristics (Winter et al., 2009; Thompson et al., 2013; Aguirre-Gutiérrez et al., 2020; Gioria et al., 2023; Diniz et al., 2024). Evaluating the explanatory power of environmental variables demonstrates the different distributions of species assemblages per sample groups along environmental gradients (Fig. 6). The importance of the climate, spatial, and soil variable groups was generally higher than topographic variables for all sample groups. The south–north gradient (Y) is an important predictor for disturbed and undisturbed samples, likely reflecting its alignment with the elevational gradient and the evolutionary divergence in the biogeographical histories of the Greater and Lesser Caucasus, characterized by distinct geological and climatic influences (Tarkhnishvili, 2014). While the Greater Caucasus experienced extensive glaciation during the last ice ages, the Lesser Caucasus had a higher snow-line, situated between 2200 and 2300 m a.s.l., resulting in less extensive glaciation (Khazaradze et al., 2018). Among the climate variables, precipitation during the warmest quarter (BIO18), mean temperature of the wettest quarter (BIO08), and annual mean potential evaporation (pot_eva) rank highest for undisturbed and disturbed samples, emphasizing the role of seasonal water balance and climatic extremes during the growing season in shaping woody species assemblages (Gardner et al., 2019). The high ranks of soil pH (phh20_mean) and soil organic carbon stocks (socs_0_30) underline filtering effects of soil edaphic variations (Marage and Gégout, 2009; Walthert and Meier, 2017). These edaphic patterns likely reflect elevational changes in parent material and organic matter turnover. Additionally, the consistently high importance of volumetric water retention (wv0010_mean and wv0033_mean) across all sample groups highlights the critical influence

of water availability gradients in structuring forest species assemblages (Maia et al., 2020). These findings highlight that soil water content variables are critical physiological factors shaping species assemblage distributions, supporting Gardner et al. (2019) who argue that soil moisture is a crucial yet often underrepresented proximal determinant in plant species distribution modeling. The importance of median pairwise spatial distance as a predictor for both disturbed groups suggests a non-stochastic distribution along the respective gradient. In particular, disturbed samples with neophytes cluster in the western part of the country (Figure 2), which may reflect spatial connectivity of disturbance events -whether biotic or abiotic. However, the comparably low prediction accuracy and relatively low variable importance for disturbed samples with neophytes suggest that species assembly is shaped primarily by other factors, most likely related to anthropogenic influences such as land use, introduction pathways or management legacies (Emerson and Gillespie, 2008). Topographic variables consistently show low predictive importance, with relative contributions below 30 % across all sample groups. Since fine-scale topography creates complex mosaics of microhabitats that often exceed the scope of standard observation plot sizes, the selected topographic variables may not have sufficiently captured related gradients. While these environmental predictors are likely relevant in other mountainous forested regions, their influence may vary depending on disturbance regimes, regional species composition, and biogeographic context.

Interspecies dissimilarity can be inflated by deep phylogenetic divergence — notably in gymnosperm–angiosperm mixtures such as *F. orientalis* with *A. nordmanniana* or *P. orientalis* — yet our clustering approach remained effective in distinguishing assemblages (Padullés Cubino et al., 2021; Staab et al., 2021). However, accounting for this systematic bias through separate modeling or statistical adjustment could refine future outcomes.

Community assembly is shaped by dynamic processes across temporal scales as forest communities evolve through succession, environmental changes and responses to natural and anthropogenic disturbances (Gerhold et al., 2015; Davies, 2021). In this context, systematic forest inventories based on permanent sample plots provide a valuable framework for monitoring compositional shifts in community structure over time. In combination with phylogenetic information such inventories offer a pragmatic approach to approximate species assembly patterns in cases where comprehensive floristic inventories are not feasible due to time or resource constraints. However, as forest inventories are primarily optimized for statistically robust estimates of timber volumes, they are constrained by sampling designs, plot configurations, and repetition cycles (Lin et al., 2020; Portier et al., 2022). Minimum DBH-thresholds (e.g. 8 cm, 15 cm and 30 cm based on nested plot radii in Georgia) significantly contribute to omission of species with smaller diameters at plot level (McRoberts et al., 2009; Lin et al., 2020). In addition, bias in identification towards commercial tree species has been reported and accuracy generally depends on the botanical expertise of the field personnel (Lam and Kleinn, 2008). However, standardized protocols for species identification — potentially supported by AI-based biodiversity identification tools or laboratory methods such as DNA barcoding — offer pathways to harmonize assessments with other classification systems, i.e. complete floristic inventories or iBOL projects. Differences in sampling design and abundance estimates between woody species and understory vegetation preclude the integration of lower forest strata, reducing the capacity to detect understory-driven niche signals that contribute to finer-scale differentiation of species assemblage characteristics and environmental positioning. However, integrating understory and herbaceous layers would optimize the characterization of community composition and align with established phytosociological classification systems (Dolukhanov, 2010; Nakhutsrishvili, 2013; Nakhutsrishvili et al., 2023). Including important diagnostic species from these strata is likely to enhance the ecological resolution of species assemblages and improve the detection of fine-scale, gradient-dependent patterns in community composition. While our study relied on phylogenetic dissimilarity, the applied metric is equally compatible with taxonomic or trait-based variation (Hao et al., 2021). Indeed, Cadotte et al. (2013, 2017) suggest that combining phylogenetic and functional diversity improves the detection of assembly patterns along environmental gradients.

The patterns of community distribution along environmental gradients identified in this study are empirically derived but consistent with ecophylogenetic theory. They support the view that community assembly is shaped by processes such as environmental filtering and niche conservatism (Cavender-Bares and Wilczek, 2003). Our findings confirm that phylogenetic dissimilarity in woody species assemblages reflects environmental filtering, making phylogenetic variability a useful proxy for ecological differentiation. Without a theoretical framework linking phylogenetic structure to assembly processes, these patterns would remain largely descriptive and offer limited insight into underlying ecological mechanisms. At the same time, we demonstrate that disturbance-induced shifts are also reflected in the phylogenetic structure of communities (Webb et al., 2002; Davies, 2021). This is evident in the stronger phylogenetic signal and higher classification accuracy found in undisturbed samples, where environmental filtering appears to dominate, compared to more stochastic assemblage patterns in disturbed samples and those containing neophytes. Although our approach cannot fully resolve the specific variables driving community separation, it demonstrates the relevance of scale-dependent assembly rules. Predictive accuracy and phylogenetic distinctiveness vary with disturbance levels, in line with theoretical models of limiting similarity and environmental filtering across spatial scales (Cavender-Bares and Wilczek, 2003; Kraft et al., 2015).

While our results are specific to species composition and disturbance regimes of our study area, the methodological approach is applicable across temperate forest ecosystems. Although species-specific patterns require regional calibration, the underlying framework is transferable across temperate forest ecosystems with relatively intact or naturally structured species compositions. While we relied on readily available and easily derived EO-based variables as predictors, the framework supports flexible integration of additional predictors tailored to regional conditions or management objectives. With appropriate calibration, the proposed framework may serve as a tool for modeling potential natural vegetation, especially in contexts where phylogenetically informed assemblages offer insights into long-term ecological stability under specific environmental conditions. Since current species assemblages reflect not only contemporary environmental conditions but also historical ecological processes and legacy effects, future studies could refine this approach by incorporating dendrochronological analyses, historical climate reconstructions, or time-series EO data to better capture how past environmental variability shapes present-day species assemblages. While beyond the scope of this study, optical EO data capturing phenological or structural traits aligned with phylogenetic structure could potentially improve model-based prediction of species assemblages. Another relevant direction is to assess to which extent the identified species assemblages align with functional groups of similar characteristics relevant to forestry, such as growth dynamics, biomass distribution, or stand structure. Such insights could inform assemblage-based guidelines for sustainable forest management that integrate ecological traits, site conditions, and phylogenetic dynamics in shaping community structure.

The increasing availability of systematic forest inventory data worldwide offers an efficient framework to monitor woody species assemblages as part of forest biodiversity. Our study demonstrates that organizing species assemblages by composition and phylogenetic structure reveals consistent community distribution patterns along environmental gradients. Advancements in DNA sequencing, modeling, and ecological data processing increasingly support phylogenetically informed habitat prediction, improving our understanding of how evolutionary history shapes species assembly across ecological niches (Padullés Cubino et al., 2021). By integrating the phylogenetic

dimension of community assembly mechanisms into classification systems, conservation and forest managers can monitor genetic diversity and its distribution across space and time. This approach not only supports biodiversity conservation but also strengthens strategies for sustainable forest management by linking species composition to ecosystem stability and resilience. Integrating phylogenetic insights into classification approaches allows forest management to move beyond simplistic species-based strategies to more dynamic, ecosystem-based approaches that balance ecological integrity with long-term productivity and sustainable use of forest resources.

## 5. Conclusions

Our study demonstrates that incorporating phylogenetic variability into species assemblage classification provides a robust framework for understanding forest community structure and its relationship to environmental gradients. Phylogenetically informed classification, based on multivariate EO data, aligns well with environmental gradients, especially under low-disturbance scenarios. The decline in prediction accuracy with increasing disturbance reflects the homogenizing effects of disturbance on species composition, which leads to less coherent community structures. These findings emphasize the potential of ecophylogenetics to link woody species assemblages more clearly to environmental conditions and improve forest community classification. This approach can support more targeted conservation strategies, strengthen sustainable forest management, and inform adaptive practices.

## CRediT authorship contribution statement

**Hein Nils:** Supervision, Project administration, Investigation, Conceptualization. **Misof Bernhard:** Writing – review & editing. **Tarkhnishvili David:** Visualization, Conceptualization. **Wellenbeck Alexander:** Writing – review & editing, Writing – original draft, Visualization, Validation, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Janiashvili Zurab:** Writing – review & editing, Data curation. **Schmidtlein Sebastian:** Writing – review & editing, Formal analysis. **Feilhauer Hannes:** Supervision, Methodology, Conceptualization. **Dzadzamia Lasha:** Writing – review & editing, Data curation.

## Declaration of Competing Interest

The authors declare no competing interests.

## Acknowledgments

We thank the Ministry of Environmental Protection and Agriculture, Georgia for granting permission to use inventory and related data from the National Forest Inventory. We also thank Nils Griese for his support during the creation of relevant R functions. This research was conducted within the framework of the Caucasus Barcode of Life project (CaBOL) and funded by the Federal Ministry of Education and Research under grant numbers 01DK20014A and 01DK20014C.

## Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.foreco.2025.122763.

## Data availability

Data will be made available on request.

## References

Afanador, N.L., Smolinska, A., Tran, T.N., Blanchet, L., 2016. Unsupervised random forest: a tutorial with case studies. J. Chemom. 30 (5), 232–241. https://doi.org/10.1002/cem.2790.

Aguirre-Gutiérrez, J., Malhi, Y., Lewis, S.L., Fauset, S., Adu-Bredu, S., Affum-Baffoe, K., et al., 2020. Long-term droughts may drive drier tropical forests towards increased functional, taxonomic and phylogenetic homogeneity. Nat. Commun. 11 (1), 3346. https://doi.org/10.1038/s41467-020-16973-4.

Akobia, I., Janiashvili, Z., Metreveli, V., Zazanashvili, N., Batsatsashvili, K., Ugrekhelidze, K., 2022. Modelling the potential distribution of subalpine birches (Betula spp.) in the Caucasus. Community Ecol. 23 (2), 209–218. https://doi.org/10.1007/s42974-022-00097-4.

Araújo, M.B., Anderson, R.P., Márcia Barbosa, A., Beale, C.M., Dormann, C.F., Early, R., et al., 2019. Standards for distribution models in biodiversity assessments. Sci. Adv. 5 (1), eaat4858. https://doi.org/10.1126/sciadv.aat4858.

Ascanio, A., Bracken, J.T., Stevens, M.H.H., Jezkova, T., 2024. New theoretical and analytical framework for quantifying and classifying ecological niche differentiation. Ecol. Monogr. https://doi.org/10.1002/ecm.1622.

Beatty, S.W., 1984. Influence of Microtopography and Canopy Species on Spatial Patterns of Forest Understory Plants. Ecology 65, 1406–1419. ⟨https://www.jstor.org/stable/1939121⟩.

Böhner, J., Selige, T., 2006. Spatial Prediction of Soil Attributes using Terrain Analysis and Climate Regionalisation. Göttinger Geogr. Abh. 115, 13–27. ⟨https://mediatum.ub.tum.de/doc/1304675/document.pdf⟩.

Bosch, N.E., Wernberg, T., Langlois, T.J., Smale, D.A., Moore, P.J., Franco, J.N., et al., 2021. Niche and neutral assembly mechanisms contribute to latitudinal diversity gradients in reef fishes. J. Biogeogr. 48 (11), 2683–2698. https://doi.org/10.1111/jbi.14237.

Box, O.E., 1996. Predicting physiognomic vegetation types with climate variables. J. Veg. Sci. 7 (3), 309–320. https://doi.org/10.2307/3236274.

Breiman, L., 2001. Random Forests. Mach. Learn. 45, 5–32. https://doi.org/10.1023/A:1010933404324.

Buuren, S., Groothuis-Oudshoorn, K., 2011. mice: Multivariate Imputation by Chained Equations in R. J. Stat. Soft 45 (3), 1–67. https://doi.org/10.18637/jss.v045.i03.

Cabido, M., Zeballos, S.R., Zak, M., Carranza, M.L., Giorgis, M.A., Cantero, J.J., Acosta, A.T.R., 2018. Native woody vegetation in central Argentina: Classification of Chaco and Espinal forests. Appl. Veg. Sci. 21 (2), 298–311. https://doi.org/10.1111/avsc.12369.

Cáceres, M. de, Martín-Alcón, S., González-Olabarria, J.R., Coll, L., 2019. A general method for the classification of forest stands using species composition and vertical and horizontal structure. Ann. For. Sci. 76 (2), 1–19. https://doi.org/10.1007/s13595-019-0824-0.

Cadotte, M.W., 2017. Functional traits explain ecosystem function through opposing mechanisms. Ecol. Lett. 20 (8), 989–996. https://doi.org/10.1111/ele.12796.

Cadotte, M.W., Tucker, C.M., 2017. Should Environmental Filtering be Abandoned? Trends Ecol. Evol. 32 (6), 429–437. https://doi.org/10.1016/j.tree.2017.03.004.

Cavender-Bares, J., Wilczek, A., 2003. Integrating Micro- and Macroevolutionary Processes in Community Ecology. Ecol. Soc. Am. 84, 592–597. ⟨https://www.jstor.org/stable/3107852⟩.

Chirici, G., McRoberts, R.E., Winter, S., Bertini, R., Brändli, U., Asensio, I.A., et al., 2012. National Forest Inventory Contributions to Forest Biodiversity Monitoring. For. Sci. 58 (3), 257–268. https://doi.org/10.5849/forsci.12-003.

Chytrý, M., Tichý, L., Holt, J., Botta-Dukát, Z., 2002. Determination of diagnostic species with statistical fidelity measures. J. Veg. Sci. 13 (1), 79–90. https://doi.org/10.1111/j.1654-1103.2002.tb02025.x.

Chytrý, M., Hennekens, S.M., Jiménez-Alfaro, B., Knollová, I., Dengler, Jansen, F., et al., 2016. European vegetation archive (EVA): an integrated database of European vegetation plots. Appl. Veg. Sci. 19 (1), 173–180. https://doi.org/10.1111/avsc.12191.

Coelho, M.T.P., Barreto, E., Rangel, T.F., Diniz-Filho, J.A.F., Wüest, R.O., Bach, W., et al., 2023. The geography of climate and the global patterns of species diversity. Nature 622 (7983), 537–544. https://doi.org/10.1038/s41586-023-06577-5.

Conrad, O., Bechtel, B., Bock, M., Dietrich, H., Fischer, E., Gerlitz, L., et al., 2015. System for Automated Geoscientific Analyses (SAGA) v. 2.1.4. Geosci. Model Dev. 8 (7), 1991–2007. https://doi.org/10.5194/gmd-8-1991-2015.

Corona, P., Chirici, G., McRoberts, R.E., Winter, S., Barbati, A., 2011. Contribution of large-scale forest inventories to biodiversity assessment and monitoring. For. Ecol. Manag. 262 (11), 2061–2069. https://doi.org/10.1016/j.foreco.2011.08.044.

Cortner, O., Chen, S., Olofsson, P., Gollnow, F., Torchinava, P., Garrett, R.D., 2024. Exploring natural and social drivers of forest degradation in post-Soviet Georgia. Glob. Environ. Change 84, 102775. https://doi.org/10.1016/j.gloenvcha.2023.102775.

Costanza, J.K., Faber-Langendoen, D., Coulston, J.W., Wear, D.N., 2018. Classifying forest inventory data into species-based forest community types at broad extents: exploring tradeoffs among supervised and unsupervised approaches. For. Ecosyst. 5 (1). https://doi.org/10.1186/s40663-017-0123-x.

Cruz-Cárdenas, G., López-Mata, L., Villaseñor, J.L., Ortiz, E., 2014. Potential species distribution modeling and the use of principal component analysis as predictor variables. Rev. Mex. De. Biodivers. 85 (1), 189–199. https://doi.org/10.7550/rmb.36723.

Cutler, D.R., Edwards, T.C., Beard, K.H., Cutler, A., Hess, K.T., Gibson, J., Lawler, J.J., 2007. Random forests for classification in ecology. Ecology 88 (11), 2783–2792. https://doi.org/10.1890/07-0539.1.

Davies, R.W., Ryan, C.M., Harrison, R.D., Dexter, K.G., Ahrends, A., te Beest, M., et al., 2023. Precipitation gradients drive high tree species turnover in the woodlands of

eastern and southern Africa. Ecography 2023 (10). https://doi.org/10.1111/ecog.06720.
Davies, T.J., 2021. Ecophylogenetics redux. Ecol. Lett. 24 (5), 1073–1088. https://doi.org/10.1111/ele.13682.
Denk, T., Frotzler, N., Davitashvili, N., 2001. Vegetational patterns and distribution of relict taxa in humid temperate forests and wetlands of Georgia (Transcaucasia). Biol. J. Linn. Soc. Lond. 72, 287–332. https://doi.org/10.1006/bijl.2000.0502.
Dering, M., Baranowska, M., Beridze, B., Chybicki, I.J., Danelia, I., Iszkuło, G., et al., 2021. The evolutionary heritage and ecological uniqueness of Scots pine in the Caucasus ecoregion is at risk of climate changes. Sci. Rep. 11 (1), 22845. https://doi.org/10.1038/s41598-021-02098-1.
Diniz, É .S., Dias, F.S., Borda-de-Água, L., González, P.M.R., 2024. Anthropogenic disturbance and alien plant invasion drive the phylogenetic impoverishment in riparian vegetation. Biodivers. Conserv 33 (14), 4237–4256. https://doi.org/10.1007/s10531-024-02949-z.
Dolukhanov, A.G. 2010 *Лесная растительность Грузии: (Forest vegetation of Georgia) [in Russian]*. Universal, 484 p. ⟨https://bilib.blogspot.com/2011/12/blog-post_24.html⟩.
Doove, L.L., van Buuren, S., Dusseldorp, E., 2014. Recursive partitioning for missing data imputation in the presence of interaction effects. Comput. Stat. Data Anal. 72, 92–104. https://doi.org/10.1016/j.csda.2013.10.025.
Elizbarashvili, E.S., Chavchanidze, Z.B., Elizbarashvili, M.E., Maglakelidze, R.V., Sulkhanishvili, N.G., Elizbarashvili, S.E., 2006. Soil-climatic zoning of Georgia. Eurasia Soil Sc. 39 (10), 1062–1065. https://doi.org/10.1134/S1064229306100036.
Emerson, B.C., Gillespie, R.G., 2008. Phylogenetic analysis of community assembly and structure over space and time. Trends Ecol. Evol. 23 (11), 619–630. https://doi.org/10.1016/j.tree.2008.07.005.
Ette, J.-S., Sallmannshofer, M., Geburek, T., 2023. Assessing Forest Biodiversity: A Novel Index to Consider Ecosystem, Species, and Genetic Diversity. Forests 14 (4), 709. https://doi.org/10.3390/f14040709.
Farr, T.G., Rosen, P.A., Caro, E., Crippen, R., Duren, R., Hensley, S., et al., 2007. The Shuttle Radar Topography Mission. Rev. Geophys. 45 (2), RG2004. https://doi.org/10.1029/2005RG000183.
Fazlollahi Mohammadi, M., Tobin, B., Jalali, S.G., Kooch, Y., Riemann, R., 2022. Fine-scale topographic influence on the spatial distribution of tree species diameter in old-growth beech (Fagus orientalis Lipsky.) forests, northern Iran. Sci. Rep. 12 (1), 7633. https://doi.org/10.1038/s41598-022-10606-0.
Feilhauer, H., Schmidtlein, S., 2009. Mapping continuous fields of forest alpha and beta diversity. Appl. Veg. Sci. 12 (4), 429–439. https://doi.org/10.1111/j.1654-109X.2009.01037.x.
Feilhauer, H., Faude, U., Schmidtlein, S., 2011. Combining Isomap ordination and imaging spectroscopy to map continuous floristic gradients in a heterogeneous landscape. Remote Sens. Environ. 115 (10), 2513–2524. https://doi.org/10.1016/j.rse.2011.05.011.
Feilhauer, H., He, K.S., Rocchini, D., 2012. Modeling species distribution using niche-based proxies derived from composite bioclimatic variables and MODIS NDVI. Remote Sens. 4 (7), 2057–2075. https://doi.org/10.3390/rs4072057.
Ferrier, S., Guisan, A., 2006. Spatial modelling of biodiversity at the community level. J. Appl. Ecol. 43 (3), 393–404. https://doi.org/10.1111/j.1365-2664.2006.01149.x.
Fischer, E., Gröger, A., Lobin, W., 2018. Illustrated field guide to the flora of Georgia (South Caucasus), 1st edn. University of Koblenz-Landau, p. 830.
Foody, G.M., Cutler, M.E.J., 2003. Tree biodiversity in protected and logged Bornean tropical rain forests and its measurement by satellite remote sensing. J. Biogeogr. 30 (7), 1053–1066. https://doi.org/10.1046/j.1365-2699.2003.00887.x.
Fuchs, H., Kleinn, C., Fehrmann, L., 2017. Establishing the Georgian National Forest Monitoring System: Integrating Remote Sensing. Natl. For. Inventory (NFI) For. Manag. Inventory (FMI) Ga. 22.
Ganeshaiah, K.N., Uma Shaankar, R., 2000. Measuring biological heterogeneity of forest vegetation types: avalanche index as an estimate of biological diversity. Biodivers. Conserv. 9, 953–963. https://doi.org/10.1023/A:1008910918751.
Gardner, A.S., Maclean, I.M., Gaston, K.J., 2019. Climatic predictors of species distributions neglect biophysiologically meaningful variables. Divers. Distrib. 25 (8), 1318–1333. https://doi.org/10.1111/ddi.12939.
GBIF Secretariat, 2021. GBIF Backbone Taxonomy. GBIF Secretariat. ⟨https://www.gbif.org/⟩.
Genuer, R., Poggi, J., 2020. Random forests with R, 97. Springer. ⟨http://www.springer.com/series/6991⟩.
Gerhold, P., Cahill, J.F., Winter, M., Bartish, I.V., Prinzing, A., 2015. Phylogenetic patterns are not proxies of community assembly mechanisms (they are far better). Funct. Ecol. 29 (5), 600–614. https://doi.org/10.1111/1365-2435.12425.
Gilbert, G.S., Parker, I.M., 2022. Phylogenetic Distance Metrics for Studies of Focal Species in Communities: Quantiles and Cumulative Curves. Diversity 14 (7), 521. https://doi.org/10.3390/d14070521.
Gilbert, N.A., Amaral, B.R., Smith, O.M., Williams, P.J., Ceyzyk, S., Ayebare, S., et al., 2024. A century of statistical Ecology. Ecology 1–14. https://doi.org/10.1002/ecy.4283.
Gillerot, L., Grussu, G., Condor-Golec, R., Tavani, R., Dargush, P., Attorre, F., 2021. Progress on incorporating biodiversity monitoring in REDD+ through national forest inventories. Glob. Ecol. Conserv. 32, e01901. https://doi.org/10.1016/j.gecco.2021.e01901.
Gillespie, T.W., Foody, G.M., Rocchini, D., Giorgi, A.P., Saatchi, S., 2008. Measuring and modelling biodiversity from space. Prog. Phys. Geogr.: Earth Environ. 32 (2), 203–221. https://doi.org/10.1177/0309133308093606.
Gioia, D., Danese, M., Corrado, G., Di Leo, P., Minervino Amodio, A., Schiattarella, M., 2021. Assessing the prediction accuracy of geomorphon-based automated landform classification: an example from the ionian coastal belt of Southern Italy. IJGI 10 (11), 725. https://doi.org/10.3390/ijgi10110725.
Gioria, M., Carta, A., Balogianni, V., Fornara, D., Pyšek, P., Osborne, B.A., 2023. Changes in the functional and phylogenetic diversity of above- and below-ground plant communities invaded by two alien herbs. NB 88, 75–101. https://doi.org/10.3897/neobiota.88.109185.
Godoy, O., Rueda, M., 2016. El uso de inventarios forestales para entender la evolución, el mantenimiento, y el funcionamiento de la diversidad de especies. [in Spanish]. ECOS 26 (3), 67–79. https://doi.org/10.7818/ECOS.2016.25-3.09.
GRASS Development Team. 2022 *Geographic Resources Analysis Support System (GRASS GIS) Software, Version 8.2.* Open Source Geospatial Foundation. ⟨https://grass.osgeo.org⟩.
Griesbach, R. 2018 *Support in planning and implementation of national forest inventory of Georgian forests: Mission Report*. unpublished, 26 p.
Guisan, A., Thuiller, W., 2005. Predicting species distribution: offering more than simple habitat models. Ecol. Lett. 8 (9), 993–1009. https://doi.org/10.1111/j.1461-0248.2005.00792.x.
Haesen, S., Lembrechts, J.J., Frenne, P. de, Lenoir, J., Aalto, J., Ashcroft, M.B., et al., 2023. ForestClim-Bioclimatic variables for microclimate temperatures of European forests. Glob. Change Biol. 29 (11), 2886–2892. https://doi.org/10.1111/gcb.16678.
Hao, M., Corral-Rivas, J., González-Elizondo, M.S., Ganeshaiah, K.N., Nava-Miranda, M. G., Zhang, C., et al., 2019a. Assessing biological dissimilarities between five forest communities. For. Ecosyst. 6 (1). https://doi.org/10.1186/s40663-019-0188-9.
Hao, M., Ganeshaiah, K.N., Zhang, C., Zhao, X., Gadow, K. von, 2019b. Discriminating among forest communities based on taxonomic, phylogenetic and trait distances. For. Ecol. Manag. 440, 40–47. https://doi.org/10.1016/j.foreco.2019.03.006.
Hao, M., Gadow, K. von, Alavi, S.J., Álvarez-González, J.G., Baluarte-Vásquez, J.R., Corral-Rivas, J., et al., 2021. A classification of woody communities based on biological dissimilarity. Appl. Veg. Sci. 24 (1). https://doi.org/10.1111/avsc.12565.
Hawkins, B.A., Rueda, M., Rangel, T.F., Field, R., Diniz-Filho, J.A.F., 2014. Community phylogenetics at the biogeographical scale: cold tolerance, niche conservatism and the structure of North American forests. J. Biogeogr. 41 (1), 23–38. https://doi.org/10.1111/jbi.12171.
Hengl, T., Mendes de Jesus, J., Heuvelink, G.B.M., Ruiperez Gonzalez, M., Kilibarda, M., Blagotić, A., et al., 2017. SoilGrids250m: global gridded soil information based on machine learning. PloS One 12 (2), e0169748. https://doi.org/10.1371/journal.pone.0169748.
Henttonen, H.M., Kangas, A., 2015. Optimal plot design in a multipurpose forest inventory. For. Ecosyst. 2 (1). https://doi.org/10.1186/s40663-015-0055-2.
Hernández-Stefanoni, J.L., Gallardo-Cruz, J.A., Meave, J.A., Rocchini, D., Bello-Pineda, J., López-Martínez, J.O., 2012. Modeling α- and β-diversity in a tropical forest from remotely sensed and spatial data. Int. J. Appl. Earth Obs. Geoinf. 19, 359–368. https://doi.org/10.1016/j.jag.2012.04.002.
Heym, M., Uhl, E., Moshammer, R., Dieler, J., Stimm, K., Pretzsch, H., 2021. Utilising forest inventory data for biodiversity assessment. Ecol. Indic. 121, 107196. https://doi.org/10.1016/j.ecolind.2020.107196.
Holt, R., 2009. Bringing the Hutchinsonian niche into the 21st century: ecological and evolutionary perspectives. PNAS 106 (2), 19659–19665. https://doi.org/10.1073/pnas.0905137106.
Jasiewicz, J., Stepinski, T.F., 2013. Geomorphons — a pattern recognition approach to classification and mapping of landforms. Geomorphology 182, 147–156. https://doi.org/10.1016/j.geomorph.2012.11.005.
Jin, Y., Qian, H., 2019. V.PhyloMaker: an R package that can generate very large phylogenies for vascular plants. Ecography 42 (8), 1353–1359. https://doi.org/10.1111/ecog.04434.
Jin, Y., Qian, H., 2022. V.PhyloMaker2: An updated and enlarged R package that can generate very large phylogenies for vascular plants. Plant Divers. 44 (4), 335–339. https://doi.org/10.1016/j.pld.2022.05.005.
Joppa, L.N., Roberts, D.L., Myers, N., Pimm, S.L., 2011. Biodiversity hotspots house most undiscovered plant species. Proc. Natl. Acad. Sci. USA 108 (32), 13171–13176. https://doi.org/10.1073/pnas.1109389108.
Kavgacı, A., Karaköse, M., Keleş, E.S., Balpınar, N., Arslan, M., Yalçın, E., et al., 2023. Classification of forest and shrubland vegetation in central and eastern Euxine Turkey and SW Georgia. Appl. Veg. Sci. 26 (4). https://doi.org/10.1111/avsc.12753.
Keggenhoff, I., Elizbarashvili, M., Amiri-Farahani, A., King, L., 2014. Trends in daily temperature and precipitation extremes over Georgia, 1971–2010. Weather Clim. Extrem. 4, 75–85. https://doi.org/10.1016/j.wace.2014.05.001.
Kembel, S.W., Cowan, P.D., Helmus, M.R., Cornwell, W.K., Morlon, H., Ackerly, D.D., et al., 2010. Picante: R tools for integrating phylogenies and ecology. Bioinforma. (Oxf., Engl. ) 26 (11), 1463–1464. https://doi.org/10.1093/bioinformatics/btq166.
Khazaradze, R., Kharadze, K., Tsikarishvili, K., Chartolani, G., 2018. Ancient Glaciation of the Caucasus. OJG 8 (01), 56–64, 2161-758910.4236/ojg.2018.81004.
Kindt, R., 2020. World.: R. Package Exact. Fuzzy matching Plant names World.Flora Online Taxon. Backbone data 19.
Kling, M.M., Mishler, B.D., Thornhill, A.H., Baldwin, B.G., Ackerly, D.D., 2018. Facets of phylodiversity: evolutionary diversification, divergence and survival as conservation targets. Philos. Trans. R. Soc. Lond. Ser. B, Biol. Sci. 374 (1763). https://doi.org/10.1098/rstb.2017.0397.
Kolbaia, S., Lortkipanidze, B., Kikodze, D., Wong, L.J., Pagad, S., 2020. Global Register of Introduced and Invasive Species - Georgia. Invasive Species Specialist Group ISSG.
Kraft, N.J.B., Adler, P.B., Godoy, O., James, E.C., Fuller, S., Levine, J.M., 2015. Community assembly, coexistence and the environmental filtering metaphor. Funct. Ecol. 29 (5), 592–599. https://doi.org/10.1111/1365-2435.12345.

Kuenzer, C., Ottinger, M., Wegmann, M., Guo, H., Wang, C., Zhang, J., et al., 2014. Earth observation satellite sensors for biodiversity monitoring: potentials and bottlenecks. Int. J. Remote Sens. 35 (18), 6599–6647. https://doi.org/10.1080/01431161.2014.964349.

Kusumoto, B., Kubota, Y., Baselga, A., Gómez-Rodríguez, C., Matthews, T.J., Murphy, D. J., Shiono, T., 2021. Community dissimilarity of angiosperm trees reveals deep-time diversification across tropical and temperate forests. J. Veg. Sci. 32 (2). https://doi.org/10.1111/jvs.13017.

Lam, T.Y., Kleinn, C., 2008. Estimation of tree species richness from large area forest inventory data: Evaluation and comparison of jackknife estimators. For. Ecol. Manag. 255 (3-4), 1002–1010. https://doi.org/10.1016/j.foreco.2007.10.007.

Lausch, A., Baade, J., Bannehr, L., Borg, E., Bumberger, J., Chabrilliat, S., et al., 2019. Linking remote sensing and geodiversity and their traits relevant to biodiversity—part i: soil characteristics. Remote Sens. 11 (20), 2356. https://doi.org/10.3390/rs11202356.

Legendre, P., Legendre, L., 2012. Numerical Ecology, Third English Edition. Elsevier BV, p. 1003.

Liaw, A., Wiener, M., 2002. Classification and Regression by randomForest. R. N. 2 (3), 18–22, 1609-3631.

Lin, H.-T., Lam, T.Y., Gadow, K. von, Kershaw, J.A., 2020. Effects of nested plot designs on assessing stand attributes, species diversity, and spatial forest structures. For. Ecol. Manag. 457, 117658. https://doi.org/10.1016/j.foreco.2019.117658.

Macek, M., Kopecký, M., Wild, J., 2019. Maximum air temperature controlled by landscape topography affects plant species composition in temperate forests. Landsc. Ecol. 34, 2541–2556. https://doi.org/10.1007/s10980-019-00903-x.

Maia, V.A., Souza, C.R. de, Aguiar-Campos, N. de, Fagundes, N.C.A., Santos, A.B.M., Paula, G.G.P. de, et al., 2020. Interactions between climate and soil shape tree community assembly and above-ground woody biomass of tropical dry forests. For. Ecol. Manag. 474, 118348. https://doi.org/10.1016/j.foreco.2020.118348.

Marage, D., Gégout, J.-C., 2009. Importance of soil nutrients in the distribution of forest communities on a large geographical scale. Glob. Ecol. Biogeogr. 18 (1), 88–97. https://doi.org/10.1111/j.1466-8238.2008.00428.x.

McCune, B., Keon, D., 2002. Equations for potential annual direct incident radiation and heat load. J. Veg. Sci. 13 (4), 603–606. https://doi.org/10.1111/j.1654-1103.2002.tb02087.x.

McRoberts, R.E., Tomppo, E., Schadauer, K., Vidal, C., Ståhl, G., Chirici, G., et al., 2009. Harmonizing national forest inventories. J. For. 107 (4), 179–187. https://doi.org/10.1093/jof/107.4.179.

MEPA, 2018. Field manual for the georgian national forest inventory: Part II. Ministry of environmental protection and agriculture. Minist. Environ. Prot. Agric. (MEPA) 74. ⟨https://mepa.gov.ge/En/Files/ViewFile/6652⟩.

MEPA. 2023 *ანგარიში ტყის პირველი ეროვნული აღრიცხვა საქართველოში - First National Forest Inventory in Georgia: Report 2023*. Ministry of Environmental Protection and Agriculture, 123 p. ⟨https://mepa.gov.ge/En/Files/Download/53934⟩.

Mikeladze, G., Gavashelishvili, A., Akobia, I., Metreveli, V., 2020. Estimation of forest cover change using Sentinel-2 multi-spectral imagery in Georgia (the Caucasus). iForest 13 (1), 329–335. https://doi.org/10.3832/ifor3386-013.

Miller, T., Blackwood, C.B., Case, A.L., 2024. Assessing the utility of SoilGrids250 for biogeographic inference of plant populations. Ecol. Evol. 14 (3), e10986. https://doi.org/10.1002/ece3.10986.

Mittermeier, R.A., Turner, W.R., Larsen, F.W., Brooks, T.M., Gascon, C., 2011. Global Biodiversity Conservation: The Critical Role of Hotspots. In: Zachos, F.E., Habel, J.C. (Eds.), Biodiversity Hotspots. Springer Berlin Heidelberg, pp. 3–22.

Moeslund, J.E., Arge, L., Bøcher, P.K., Dalgaard, T., Svenning, J.-C., 2013. Topography as a driver of local terrestrial vascular plant diversity patterns. Nord. J. Bot. 31 (2), 129–144. https://doi.org/10.1111/j.1756-1051.2013.00082.x.

Moudrý, V., Lecours, V., Malavasi, M., Misiuk, B., Gábor, L., Gdulová, K., et al., 2019. Potential pitfalls in rescaling digital terrain model-derived attributes for ecological studies. Ecol. Inform. 54, 100987. https://doi.org/10.1016/j.ecoinf.2019.100987.

Myers, N., 2003. Biodiversity hotspots revisited. BioScience 53 (10), 916–917. https://doi.org/10.1641/0006-3568(2003)053[0916:BHR]2.0.CO;2.

Nakhutsrishvili, G. (Ed.), 2013. *The Vegetation of Georgia (South Caucasus): Geobotany Studies*. 1st edn. Springer Science & Business Media, p. 236.

Nakhutsrishvili, G., Abdaladze, O., Batsatsashvili, K., 2021. Ecological Gradients (West-East) and Vegetation of the Central Great Caucasus. Bocconea 29, 157–168. https://doi.org/10.7320/Bocc29.157.

Nakhutsrishvili, G., Abdaladze, O., Batsatsashvili, K., Dzadzamia, L. (Eds.), 2023. Natural Forests of Georgia: (The South Caucasus). Ilia State University Press, p. 224.

Newton, A.C., Kapos, V., 2002. Biodiversity indicators in national forest inventories. Unasylva 53 (210), 56–64. ⟨https://www.scopus.com/record/display.uri?eid= 2-s2.0-0036958979&origin=inward&txGid=d4e9862d2823739e278c35a31f4f5156⟩.

Norberg, A., Abrego, N., Blanchet, F.G., Adler, F.R., Anderson, B.J., Anttila, J., et al., 2019. A comprehensive evaluation of predictive performance of 33 species distribution models at species and community levels. Ecol. Monogr. 89 (3). https://doi.org/10.1002/ecm.1370.

Novák, P., Kalníková, V., Szokala, D., Aleksanyan, A., Batsatsashvili, K., Fayvush, G., et al., 2023. Transcaucasian VEgetation Database – A Phytosociological database of the Southern Caucasus. VCS 4, 231–240. https://doi.org/10.3897/VCS.105521.

Oksanen, J., 2020. Package 'Vegan". *Community Ecology Package*. Community Ecology Package: Ordination, Diversity and Dissimilarities. CRAN, p. 298. ⟨https://github.com/vegandevs/vegan⟩.

OpenStreetMap contributors, 2024. OpenStreetMap. ⟨https://download.geofabrik.de/⟩.

Padullés Cubino, J., Lososová, Z., Bonari, G., Agrillo, E., Attorre, F., Bergmeier, E., et al., 2021. Phylogenetic structure of European forest vegetation. J. Biogeogr. 48 (4), 903–916. https://doi.org/10.1111/jbi.14046.

Patarkalashvili, T., 2017. Forest biodiversity of Georgia and endangered plant species. Ann. Agrar. Sci. 15 (3), 349–351. https://doi.org/10.1016/j.aasci.2017.06.002.

Pavoine, S., Bonsall, M.B., 2011. Measuring biodiversity to explain community assembly: a unified approach. Biol. Rev. Camb. Philos. Soc. 86 (4), 792–812. https://doi.org/10.1111/j.1469-185X.2010.00171.x.

Poggio, L., Sousa, L.M. de, Batjes, N.H., Heuvelink, G.B.M., Kempen, B., Ribeiro, E., Rossiter, D., 2021. SoilGrids 2.0: producing soil information for the globe with quantified spatial uncertainty. SOIL 7 (1), 217–240. https://doi.org/10.5194/soil-7-217-2021.

Portier, J., Zellweger, F., Zell, J., Alberdi Asensio, I., Bosela, M., Breidenbach, J., et al., 2022. Plot size matters: Toward comparable species richness estimates across plot-based inventories. Ecol. Evol. 12 (6), e8965. https://doi.org/10.1002/ece3.8965.

QGIS Development Team. 2009 *QGIS Geographic Information System*. QGIS Development Team. ⟨http://qgis.osgeo.org⟩.

Qian, H., Sandel, B., 2017. Phylogenetic structure of regional angiosperm assemblages across latitudinal and climatic gradients in North America. Glob. Ecol. Biogeogr. 26 (11), 1258–1269. https://doi.org/10.1111/geb.12634.

Qian, H., Zhang, Y., Zhang, J., Wang, X., 2013. Latitudinal gradients in phylogenetic relatedness of angiosperm trees in North America. Glob. Ecol. Biogeogr. 22 (11), 1183–1191. https://doi.org/10.1111/geb.12069.

Qian, H., Deng, T., Jin, Y., Mao, L., Zhao, D., Ricklefs, R.E., 2019. Phylogenetic dispersion and diversity in regional assemblages of seed plants in China. Proc. Natl. Acad. Sci. USA 116 (46), 23192–23201. https://doi.org/10.1073/pnas.1822153116.

R Core Team. 2024 *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. ⟨https://www.R-project.org/⟩.

Reise, J., Kukulka, F., Flade, M., Winter, S., 2019. Characterising the richness and diversity of forest bird species using National Forest Inventory data in Germany. For. Ecol. Manag. 432, 799–811. https://doi.org/10.1016/j.foreco.2018.10.012.

Ricotta, C., Szeidl, L., Pavoine, S., 2021. Towards a unifying framework for diversity and dissimilarity coefficients. Ecol. Indic. 129, 107971. https://doi.org/10.1016/j.ecolind.2021.107971.

Ries, L., Fletcher, R.J., Battin, J., Sisk, T.D., 2004. Ecological Responses to Habitat Edges: Mechanisms, Models, and Variability Explained. Annu. Rev. Ecol. Evol. Syst. 35 (1), 491–522. https://doi.org/10.1146/annurev.ecolsys.35.112202.130148.

Riley, S.J., DeGloria, S.D., Elliot, R., 1999. A terrain ruggedness index that quantifies topographic heterogeneity. Intermt. J. Sci. 5 (1-4), 23–27. ⟨http://download.osgeo.org/qgis/doc/reference-docs/Terrain_Ruggedness_Index.pdf⟩.

Rocchini, D., Luque, S., Pettorelli, N., Bastin, L., Doktor, D., Faedi, N., et al., 2018. Measuring β-diversity by remote sensing: A challenge for biodiversity monitoring. Methods Ecol. Evol. 9 (8), 1787–1798. https://doi.org/10.1111/2041-210X.12941.

Sanderson, E., Fisher, K., Robinson, N., Sampson, D., Duncan, A., Royte, L., 2022. The march of the human footprint. EcoEvoRxiv 1–54. https://doi.org/10.32942/osf.io/d7rh6.

Schmidtlein et al. (2010) A brute-force approach to vegetation classification. *Journal of Vegetation Science* **21**, 1162–1171. ⟨http://arxiv.org/pdf/1103.2010v1⟩. 10.1088/0004-6256/141/2/44.

Schmidtlein, S., Collison, J., Pfannendoerfer, R., 2024. isopam: Clust. Sites Species Data 10. ⟨https://CRAN.R-project.org/package=isopam⟩.

Sefidi, K., Esfandiary Darabad, F., Azaryan, M., 2016. Effect of topography on tree species composition and volume of coarse woody debris in an Oriental beech (Fagus orientalis Lipsky) old growth forests, northern Iran. iForest 9 (4), 658–665. https://doi.org/10.3832/ifor1080-008.

Shi, W., Wang, Y.-Q., Xiang, W.-S., Li, X.-K., Cao, K.-F., 2021. Environmental filtering and dispersal limitation jointly shaped the taxonomic and phylogenetic beta diversity of natural forests in southern China. Ecol. Evol. 11 (13), 8783–8794. https://doi.org/10.1002/ece3.7711.

Siegert, C.M., Levia, D.F., Hudson, S.A., Dowtin, A.L., Zhang, F., Mitchell, M.J., 2016. Small-scale topographic variability influences tree species distribution and canopy throughfall partitioning in a temperate deciduous forest. For. Ecol. Manag. 359, 109–117. https://doi.org/10.1016/j.foreco.2015.09.028.

Smith, S.A., Brown, J.W., 2018. Constructing a broadly inclusive seed plant phylogeny. Am. J. Bot. 105 (3), 302–314. https://doi.org/10.1002/ajb2.1019.

Soley-Guardia, M., Alvarado-Serrano, D.F., Anderson, R.P., 2024. Top ten hazards to avoid when modeling species distributions: a didactic guide of assumptions, problems, and recommendations. Ecography 2024 (4). https://doi.org/10.1111/ecog.06852.

Staab, M., Liu, X., Assmann, T., Bruelheide, H., Buscot, F., Durka, W., et al., 2021. Tree phylogenetic diversity structures multitrophic communities. Funct. Ecol. 35 (2), 521–534. https://doi.org/10.1111/1365-2435.13722.

Staudhammer, C.L., LeMay, V.M., 2001. Introduction and evaluation of possible indices of stand structural diversity. Can. J. For. Res. 31 (7), 1105–1115. https://doi.org/10.1139/cjfr-31-7-1105.

Stepinski, T.F. and Jasiewicz, J. (2011) Geomorphons - a new approach to classification of landforms. ⟨https://www.researchgate.net/profile/Jaroslaw-Jasiewicz/publication/264850233_Geomorphons_-_A_new_approach_to_classification_of_landforms/links/5639134508aecf1d92a9bd24/Geomorphons-A-new-approach-to-classification-of-landforms.pdf⟩.

Stritih, A., Senf, C., Kuemmerle, T., Munteanu, C., Dzadzamia, L., Stritih, J., et al., 2024. Same, but different: similar states of forest structure in temperate mountain regions of Europe despite different social-ecological forest disturbance regimes. Landsc. Ecol. 39 (6). https://doi.org/10.1007/s10980-024-01908-x.

Swenson, N.G., 2011a. Phylogenetic beta diversity metrics, trait evolution and inferring the functional beta diversity of communities. PloS One 6 (6), e21264. https://doi.org/10.1371/journal.pone.0021264.

Swenson, N.G., 2011b. The role of evolutionary processes in producing biodiversity patterns, and the interrelationships between taxonomic, functional and phylogenetic biodiversity. Am. J. Bot. 98 (3), 472–480. https://doi.org/10.3732/ajb.1000289.
Szymura, T.H., Szymura, M., Macioł, A., 2015. The effect of ecological niche and spatial pattern on the diversity of oak forest vegetation. Plant Ecol. Divers. 8 (4), 505–518. ⟨https://www.tandfonline.com/doi/full/10.1080/17550874.2015.1010186⟩.
Tarkhnishvili, D., 2014. UK ed. edition. Historical Biogeography of the Caucasus. Nova Science Pub Inc, p. 229. ⟨https://novapublishers.com/shop/historical-biogeography-of-the-caucasus/⟩.
Tarkhnishvili, D., Gavashelishvili, A., Mumladze, L., 2012. Palaeoclimatic models help to understand current distribution of Caucasian forest species. Biol. J. Linn. Soc. 105, 231–248. https://doi.org/10.1111/j.1095-8312.2011.01788.x.
R.Studio Team. 2024 *RStudio: Integrated Development for R.*: *RStudio.* Posit Software, P. B. C.
Thompson, I.D., Guariguata, M.R., Okabe, K., Bahamondez, C., Nasi, R., Heymel, V., Sabogal, C., 2013. An Operational Framework for Defining and Monitoring Forest Degradation. Ecol. Soc. 18 (2), 2ß. ⟨http://www.ecologyandsociety.org/vol18/iss2/art20/⟩.
Tobler, W.R., 1970. A computer movie simulating urban growth in the detroit region. Econ. Geogr. 46, 234. https://doi.org/10.2307/143141.
Traub, B., Wüest, R.O., 2020. Analysing the quality of swiss national forest inventory measurements of woody species richness. For. Ecosyst. 7 (1), 37. https://doi.org/10.1186/s40663-020-00252-1.
Tucker, C.M., Cadotte, M.W., Carvalho, S.B., Davies, T.J., Ferrier, S., Fritz, S.A., et al., 2017. A guide to phylogenetic metrics for conservation, community ecology and macroecology. Biol. Rev. Camb. Philos. Soc. 92 (2), 698–715. https://doi.org/10.1111/brv.12252.
Turek, M.E., Poggio, L., Batjes, N.H., Armindo, R.A., van Jong Lier, Q. de, Sousa, L. de, Heuvelink, G.B., 2023. Global mapping of volumetric water retention at 100, 330 and 15 000 cm suction using the WoSIS database. Int. Soil Water Conserv. Res. 11 (2), 225–239. https://doi.org/10.1016/j.iswcr.2022.08.001.
Vanuytrecht, E., Wouters, H., Berckmans, J., Ridder, K., 2021. Downscal. bioclimatic Indic. Sel. Reg. 1979 2018 Deriv. ERA5 reanalysis: Prod. Use Guide 44. ⟨https://cds.climate.copernicus.eu/datasets/sis-biodiversity-era5-regional⟩.
Waldock, C., Stuart-Smith, R.D., Albouy, C., Cheung, W.W.L., Edgar, G.J., Mouillot, D., et al., 2022. A quantitative review of abundance-based species distribution models. Ecography 2022 (1). https://doi.org/10.1111/ecog.05694.
Walthert, L., Meier, E.S., 2017. Tree species distribution in temperate forests is more influenced by soil than by climate. Ecol. Evol. 7 (22), 9473–9484. https://doi.org/10.1002/ece3.3436.
Webb, C.O., 2000. Exploring the phylogenetic structure of ecological communities: an example for rain forest trees. Am. Nat. 156 (2). https://doi.org/10.1086/303378.
Webb, C.O., Ackerly, D.D., McPeek, M.A., Donoghue, M.J., 2002. Phylogenies and community ecology. Annu. Rev. Ecol. Evol. Syst. 33 (1), 475–505. https://doi.org/10.1146/annurev.ecolsys.33.010802.150448.
Webb, C.O., Losos, J.B., Agrawal, A.A., 2006. Integrating phylogenies into community ecology. Ecol. - Spec. Issue 87 (sp7), S1–S2.
Weiher, E., Keddy, P., 2001. Ecological assembly rules: Perspectives, advances, retreats, 1st edn. Cambridge University Press. https://doi.org/10.1017/CBO9780511542237.
Wellenbeck, A., Fehrmann, L., Feilhauer, H., Schmidtlein, S., Misof, B., Hein, N., 2024. Discriminating woody species assemblages from National Forest Inventory data based on phylogeny in Georgia. Ecol. Evol. 14 (7). https://doi.org/10.1002/ece3.11569.
Wildlife Conservation Society. 2005 Global Human Influence Index (HII) Dataset, Last of the Wild Project, Version 2, 2005 (LWP-2):.tif. NASA Socioeconomic Data and Applications Center (SEDAC), Columbia University, Palisades, New York. https://wcshumanfootprint.org/data-access.
Willmer, J.N.G., Püttker, T., Prevedello, J.A., 2022. Global impacts of edge effects on species richness. Biol. Conserv. 272, 109654. https://doi.org/10.1016/j.biocon.2022.109654. ⟨https://www.sciencedirect.com/science/article/pii/S0006320722002075⟩.
Wilson, M.F.J., O'Connell, B., Brown, C., Guinan, J.C., Grehan, A.J., 2007. Multiscale terrain analysis of multibeam bathymetry data for habitat mapping on the continental slope. Mar. Geod. 30 (1-2), 3–35. https://doi.org/10.1080/01490410701295962.
Winter, M., Schweiger, O., Klotz, S., Nentwig, W., Andriopoulos, P., Arianoutsou, M., et al., 2009. Plant extinctions and introductions lead to phylogenetic and taxonomic homogenization of the European flora. PNAS 106 (51), 21721–21725. https://doi.org/10.1073/pnas.0907088106.
Winter, S., Chirici, G., McRoberts, R.E., Hauk, E., Tomppo, E., 2008. Possibilities for harmonizing national forest inventory data for use in forest biodiversity assessments. Forestry 81 (1), 33–44. https://doi.org/10.1093/forestry/cpm042.
Woods, C.L., Ortmann, K., 2024. Microtopographic heterogeneity affects habitat specialization and diversity of understory plants in a northern temperate rainforest. Plant Ecol. https://doi.org/10.1007/s11258-024-01469-8.
Xu, J., Dang, H., Wang, M., Chai, Y., Guo, Y., Chen, Y., et al., 2019. Is phylogeny more useful than functional traits for assessing diversity patterns under community assembly processes? Forests 10 (12), 1159. https://doi.org/10.3390/f10121159.
Yao, J., Zhang, C., Cáceres, M. de, Legendre, P., Zhao, X., 2019. Variation in compositional and structural components of community assemblage and its determinants. J. Veg. Sci. 30 (2), 257–268. https://doi.org/10.1111/jvs.12708.
Zanne, A.E., Tank, D.C., Cornwell, W.K., Eastman, J.M., Smith, S.A., FitzJohn, R.G., et al., 2014. Three keys to the radiation of angiosperms into freezing environments. Nature 506 (7486), 89–92. https://doi.org/10.1038/nature12872.
Zazanashvili, N., Sanadiradze, G., Bukhnikashvili, A., 2001. Caucasus. In: Mittermeier, Russell, A., Myers, N., Goettsch Mittermeier, C. (Eds.), Hotspots. Earth's Biologically Richest and Most Endangered Terrestrial Ecoregions, 1st edn. CEMEX, S.A., Agrupación Sierra Madre. https://doi.org/10.1017/S0376892901270088.
Zellweger, F., Braunisch, V., Morsdorf, F., Baltensweiler, A., Abegg, M., Roth, T., et al., 2015. Disentangling the effects of climate, topography, soil and vegetation on stand-scale species richness in temperate forests. For. Ecol. Manag. 349, 36–44. https://doi.org/10.1016/j.foreco.2015.04.008.
Zhou, W., Zhang, Y., Zhang, S., Yakimov, B.N., Ma, K., 2021. Phylogenetic and functional traits verify the combined effect of deterministic and stochastic processes in the community assembly of temperate forests along an elevational gradient. Forests 12 (5), 591. https://doi.org/10.3390/f12050591.

Alexander Wellenbeck explores synergies between interdisciplinary approaches to integrate forest monitoring with biodiversity conservation. This study is part of his doctoral program at the University of Bonn, conducted within the project Caucasus Barcode of Life (see CaBOL, http://ggbc.eu). The authors are an multidisciplinary team representing Forest Science, Biogeography, Vegetation Ecology, and Evolutionary Biology.