Ecophylogenetic Classification of Forest Communities in Georgia

An Empirical Approach Using National Forest Inventory Data

Dissertation
zur
Erlangung des Doktorgrades (Dr. rer. nat.)
der
Mathematisch-Naturwissenschaftlichen Fakultät
der
Rheinischen Friedrich-Wilhelms-Universität Bonn

vorgelegt von
Alexander Wellenbeck
aus
Köln

Bonn, 2025

Angefertigt mit Genehmigung der Mathematisch-Naturwissenschaftlichen Fakultät der Rheinischen Friedrich-Wilhelms-Universität Bonn

Gutachter/Betreuer: Prof. Dr. Bernhard Misof
Gutachter: Prof. Dr. David Tarkhnishvili

Tag der Promotion: 06.02.2026
Erscheinungsjahr: 2026

Dedicated to my family, Marie, Milan and Samuel

„"...ბუნება ღვთაებრივი ჰარმონიის გამოვლინებაა, ადამიანი კი - ამ ჰარმონიის ნაწილი და როგორიც არ უნდა იყოს მისი საამწუთიერო ინტერესები, იგი საბოლოოდ განწირული იქნება, თუ ბუნების გრძნეულ ენას არ მიუგდებს ყურს, მის უხილავ სულიერებას არ ირწმენს, და მის ტკივილს არ გაიზიარებს!“

— ვაჟა-ფშაველა (1861 – 1915)

“Nature is a manifestation of divine harmony, and the human being — a part of this harmony. No matter what worldly interests a person may have, they will ultimately be doomed if they do not listen to nature’s mystical language, do not believe in its invisible spirituality, and do not share in its pain!”

— Vazha-Pshavela (1861 – 1915)

Summary

The South Caucasus is a biodiversity hotspot characterized by complex biogeographic conditions and a legacy of refugial persistence. Located between the Greater and Lesser Caucasus, the country of Georgia (Sakartvelo) spans a relatively small land area of approximately 69,700 km², yet exhibits an exceptional degree of climatic, geological, and topographic variation. This environmental complexity sustains diverse communities which are characterized by a high degree of species endemism.

With approximately 40% of the country's land area being forested, forest ecosystems extend along multiple biogeographic gradients, exemplifying the region's ecological diversity. Between 2019 and 2021, these ecosystems were systematically assessed by the Government of Georgia within the framework of a National Forest Inventory for the first time to support forest policy and conservation planning. The resulting dataset offers an unprecedented opportunity to investigate patterns of forest diversity in this scientifically underexplored region and its environmental drivers at national scale.

The field of ecophylogenetics explores how evolutionary relatedness shapes species coexistence and biodiversity across environments. This thesis demonstrates how ecophylogenetic concepts can be applied to forest classification and biodiversity monitoring using a real-world, regional-scale case study. Based on systematic sample plot observations from a nation-wide forest assessment and a standardized megaphylogeny, I apply unsupervised cluster analysis based on a phylogenetically informed dissimilarity metric to species compositional data. By evaluating the resulting hierarchical structures, cluster coherence and spatial patterns of grouped species assemblages across environmental gradients, this study provides empirical evidence that quantifying species identity via evolutionary proximity strengthens ecologically meaningful stratification. This integration of ecophylogenetic theory in forest monitoring frameworks demonstrates that moving beyond traditional nominal species classifications offers a more nuanced perspective on community composition.

This cumulative dissertation comprises three research articles. A first comparative analysis (*Discriminating woody species assemblages from National Forest Inventory data based on phylogeny in Georgia* in *Ecology and Evolution*, 14(7), 2024) shows that incorporating phylogenetic information into quantifying dissimilarity for clustering woody species observations yields higher internal cluster coherence than using species-neutral metrics. In addition, by quantifying phylogenetic variability, mapped species assemblages align more clearly with expected biogeographic gradients, underlining the potential for standardized community classification. Building on this, predictive modeling (*Predicting woody species assemblages using ecophylogenetics and Earth observation data* in *Forest Ecology and Management,* Volume 589, 2025) demonstrates that phylogenetically informed cluster membership can be predicted via proxy variables for climate, soil, topography, and spatial configuration derived from Earth observation data. This predictive

potential indicates that species assemblages reflect relative positions along environmental gradients, but these relationships weaken with increasing structural homogenization and the presence of neophytes, reflecting constraints imposed by stand disturbance. Comparing changes in structural stand characteristics along a field-assigned degradation gradient reveals distinct response dynamics and resilience thresholds across phylogenetically informed clusters (Manuscript *Stand structural change along a degradation gradient in the National Forest Inventory of Georgia* under submission in *European Journal of Forest Research*).

In summary, this dissertation investigates how ecophylogenetic concepts can be integrated into the classification of forest communities for the first time. Within the framework of large-scale forest inventories, I operationalize these principles to demonstrate their general applicability and added value for forest classification and biodiversity monitoring. When interspecies phylogenetic variation is quantified for stratification, the spatial distribution of resulting strata aligns with environmental gradients. These findings show that accounting for evolutionary relationships enables the ecological interpretability of community patterns and the environmental factors that shape them. The resulting strata also maintain functional and structural coherence, supporting a broader ecological understanding of community organization. The scalable methodology developed in this thesis can be used to refine forest typologies and to model forest-type-specific dynamics of forest structure and species distribution patterns. This data-driven approach translates ecophylogenetic principles into applied forest monitoring and classification practice, offering a robust framework for regional-scale assessment. In biodiversity hotspots such as the South Caucasus, it provides an evidence-based basis for identifying evolutionarily distinct communities, informing conservation priorities, and stimulating future research on ecosystem responses to environmental change. Fundamentally, it supports a conceptual shift toward more holistic biodiversity evaluation and forest management in line with current trends in ecological research.

Translated Summary (in Georgian)

სამხრეთ კავკასია წარმოადგენს ბიომრავალფეროვნების ერთ-ერთ უმნიშვნელოვანეს ცხელ წერტილს, რომლის ბიოგეოგრაფიული სტრუქტურა აღინიშნება მაღალი სირთულითა და ისტორიული რეფუგიუმების მდგრადობით. საქართველო, მდებარე დიდი და მცირე კავკასიონის სისტემებს შორის, მოიცავს დაახლოებით 69,700 კვადრატულ კილომეტრ ტერიტორიას, თუმცა გამოირჩევა კლიმატური, გეოლოგიური და ტოპოგრაფიული მრავალფეროვნებით. ქვეყნის ტერიტორიის დაახლოებით 40% ტყით არის დაფარული; ტყის ეკოსისტემები განფენილია მრავალი ბიოგეოგრაფიული გრადიენტის გასწვრივ, რაც ასახავს რეგიონის ეკოლოგიურ სირთულეს. 2019–2021 წლებში, აღნიშნული ეკოსისტემები პირველად შეფასდა სისტემურად საქართველოს ეროვნული სატყეო ინვენტარის ფარგლებში, რაც ემსახურებოდა სატყეო პოლიტიკისა და კონსერვაციის დაგეგმვის განმტკიცებას. შედეგად მიღებული მონაცემები ქმნის უპრეცედენტო შესაძლებლობას ტყის ბიომრავალფეროვნების კანონზომიერებებისა და მისი ლანდშაფტური მამოძრავებელი ფაქტორების შესწავლისთვის.

წინამდებარე დისერტაციაში წარმოდგენილია, თუ როგორ შეიძლება ეკოფილოგენეტიკური კონცეფციების გამოყენება ტყის ტიპოლოგიასა და ბიომონიტორინგში რეალური, ლანდშაფტური მასშტაბის მაგალითზე. გამოყენებულია საქართველოს ეროვნული სატყეო ინვენტარის მონაცემები და სტანდარტიზებული მეგაფილოგენია, ხოლო სახეობრივი შემადგენლობა გაანალიზებულია ფილოგენეტიკურად ინფორმირებული დისსიმილარობის მეტრიკისა და უთითებელი კლასტერული ანალიზის მეშვეობით. მიღებული იერარქიული სტრუქტურების, კლასტერების შიდა ერთგვაროვნებისა და სახეობრივ-ეკოლოგიური ერთეულების განაწილების შეფასება გარემო გრადიენტებზე ასახავს იმ ემპირიულ მტკიცებულებებს, რომ სახეობათა იდენტიფიკაცია ევოლუციური სიახლოვის საფუძველზე იძლევა ეკოლოგიურად ფუძემდებლური სტრატიფიკაციის საშუალებას. ეკოფილოგენეტიკური მიდგომის ინტეგრირება სატყეო მონიტორინგის სისტემებში აჩვენებს, რომ სახეობათა ტრადიციული, ნომინალური კლასიფიკაციის გადააზრება შესაძლებელს ხდის საზოგადოებათა სტრუქტურის ღრმად შინაარსობრივ ინტერპრეტაციას.

დისერტაცია შედგება სამი სამეცნიერო ნაშრომისგან. პირველ პუბლიკაციაში წარმოდგენილია შედარებითი ანალიზი, რომელიც აჩვენებს, რომ სახეობათა ფილოგენეტიკური კავშირის გათვალისწინება დისსიმილარობის კვანტიფიკაციისას მნიშვნელოვნად აძლიერებს კლასტერების სტრუქტურულ ერთგვაროვნებას. ფიტოგენეტიკური ცვალებადობის ანალიზით მიღებული საზოგადოებები უკეთესად ასახავენ ბიოგეოგრაფიულ მოლოდინებს, რაც მიუთითებს სტანდარტიზებული საზოგადოებრივი კლასიფიკაციის შესაძლებლობაზე. მეორე პუბლიკაციაში წარმოდგენილია პროგნოზული მოდელირება, რომელიც აჩვენებს, რომ ფიტოგენეტიკურად დადგენილი კლასტერული კუთვნილება შეიძლება შეფასდეს კლიმატური, ნიადაგური,

რელიეფური და სივრცული ცვლადების საფუძველზე, რომლებიც მიღებულია დედამიწის დაკვირვების მონაცემებიდან. ამ ურთიერთობათა ანალიზი მიუთითებს, რომ სახეობრივი ერთეულები ასახავენ მათ გრადიენტულ განლაგებას გარემო პირობების მიმართ, თუმცა ეს კავშირები სუსტდება სტრუქტურული ჰომოგენიზაციისა და ნეოფიტების ზრდის პირობებში, რაც იწვევს სტანდის ეკოლოგიური მოწყვლადობის ზრდას. მესამე პუბლიკაციაში (წარდგენილია განსახილველად) განხილულია სტანდის სტრუქტურული მახასიათებლების ცვლილებები ფიტოგენეტიკურად სტრატიფიცირებულ კლასტერებს შორის, გარემოს მიერ განსაზღვრული დეგრადაციის გრადიენტის გასწვრივ, რაც ავლენს განსხვავებულ რეაგირების დინამიკასა და მდგრადობის ზღვარს.

დისერტაცია ადასტურებს ეკოფილოგენეტიკური კონცეფციების პრაქტიკულ გამოყენებადობას ტყის საზოგადოების კლასიფიკაციისა და მონიტორინგისათვის ფართომასშტაბიანი სატყეო ინვენტარის ფარგლებში. როდესაც სახეობრივი ერთეულების ფილოგენეტიკური მრავალფეროვნება სტრატიფიკაციისთვის ოპერატიულად გამოიყენება, ხისმცენარეთა სივრცითი განაწილება ასახავს გარემო გრადიენტებს. კვლევა მიუთითებს, რომ ევოლუციური კავშირების გათვალისწინება ამაღლებს საზოგადოების ეკოლოგიურ ინტერპრეტაციას და ააშკარავებს მის სტრუქტურულ და ფუნქციურ ერთიანობას. დისერტაციაში გამოყენებული მეთოდოლოგია შეიძლება გამოყენებულ იქნას არსებული სატყეო ტიპოლოგიების დახვეწისა და ტიპსპეციფიკური სტრუქტურული დინამიკისა და სახეობრივი განაწილების მოდელირებისთვის. ეს მონაცემებზე დაფუძნებული მიდგომა აერთიანებს ეკოფილოგენეტიკურ თეორიას სატყეო მართვისა და მონიტორინგის პრაქტიკაში და ქმნის მდგრად ჩარჩოს ფართომასშტაბიანი შეფასებისთვის, რომელიც აერთიანებს სახეობრივ და ფილოგენეტიკურ განზომილებებს. ბიომრავალფეროვნების ცხელი წერტილების, მაგალითად, სამხრეთ კავკასიის შემთხვევაში, ეს მიდგომა იძლევა მტკიცებულებაზე დაფუძნებულ საფუძველს ევოლუციურად განსხვავებული საზოგადოებების იდენტიფიცირებისთვის, კონსერვაციის პრიორიტეტების განსაზღვრისთვის და ეკოსისტემების რეაქციების კვლევის წახალისებისთვის გლობალურ ცვლილებებზე. იგი უწყობს ხელს ბიომრავალფეროვნების ჰოლისტური შეფასებისა და თანამედროვე ეკოლოგიური კვლევის ტენდენციებთან თავსებადი სატყეო მართვის კონცეპტუალურ ტრანსფორმაციას.

Contents

Contents

Figures

Tables

Acronyms and Abbreviations

ASCII	American Standard Code for Information Interchange
a.s.l.	Above sea level
BA	Basal area (m^2 ha^{-1})
DEM	Digital elevation model
DBH	Diameter at breast height (1.3 m)
EO	Earth observation
EU	European Union
FD	Functional diversity
GBIF	Global Biodiversity Information Facility
GIS	Geographic information system
GNFI	National Forest Inventory of Georgia
GoG	Government of Georgia
GPS	Global position system
GRASS	Geographic Resources Analysis Support System
H	Tree height (m)
HTS	High-throughput sequencing
ISOMAP	Isometric mapping
ISOPAM	Isometric ordination space partitioning
LiDAR	Light detection and ranging
MEPA	Ministry of Environmental Protection and Agriculture of Georgia
NFI	National forest inventory
PCA	Principal Component Analysis
PC	Principal Component
PD	Phylogenetic diversity
RF	Random forests
RS	Remote sensing
SAGA	System for Automated Geoscientific Analyses
SRTM	Shuttle Radar Topography Mission
UN-CBD	United Nations Convention on bilogical diversity
UTM	Universal Transverse Mercator

1 Introduction

Global biodiversity is continuously declining as a result of human activity (Vellend et al., 2013; Wiebe & Wilcove, 2025). The magnitude by which human society alters ecosystems globally has led scientists to refer to the current epoch as the Anthropocene (Barnosky et al., 2012; Steffen et al., 2015). Human activities include land-use and land cover change, pollution, overexploitation of resources, the introduction of invasive species, and greenhouse gas emissions, which are reshaping ecosystems around the world. These changes affect the condition, organization, functioning, and evolution of biological systems at multiple scales, resulting in ongoing loss of biodiversity (Fazan et al., 2020; IPBES, 2021; Keck et al., 2025). Biodiversity is distributed unequally across Earths' ecozones (Hagen et al., 2021; Oyebanji et al., 2023; Alzate & Hagen, 2024). Occupying around a third of the planet's land surface, forest ecosystems host the majority of Earth's terrestrial biodiversity (Chirici et al., 2012; FAO, 2020; Gillerot et al., 2021; Heym et al., 2021). The most recent global Forest Resources Assessment (FRA) by the Food and Agriculture Organization (FAO) along with other sources report that the worldwide forest areas continue to diminish (FAO, 2020; Estoque et al., 2022), and biodiversity loss in temperate forest ecosystems has grown more severe over the past decade (Butchart et al., 2010; Paillet, 2017; Da Silva et al., 2019; Forzieri et al., 2022). International agreements and initiatives continue to advocate robust monitoring systems to address and prevent further biodiversity loss in forests and other ecosystems (Newton & Kapos, 2002; Chirici et al., 2012; European Commission, 2021; Hughes & Grumbine, 2023).

Due to the variability of forested areas across climatic zones, national definitions of "forest" differ, depending on ecological, cultural, legal, or land-use criteria (Chazdon et al., 2016). Most definitions emphasize the presence of woody vegetation as a core component, but criteria regarding height, required canopy cover, and minimum land area thresholds may vary (Vidal et al., 2016b). For the FRA of 2020, FAO has defined forest as "land spanning more than 0.5 hectares with trees higher than 5 meters and a canopy cover of more than 10 percent, or trees able to reach these thresholds *in situ*" excluding land that is primarily for urban development or agricultural (FAO, 2018). However, diverting national definitions persist and hinder harmonization of monitoring, reporting, and conservation planning (McRoberts et al., 2009; Trentanovi et al., 2023). The term "biodiversity" has a broad meaning, encompassing multiple levels of observation and being inherently dependent on the temporal and spatial scales at which it is measured (Kaennel, 1998). As such, defining biodiversity is complex, since it relates to various components of ecological complexity and spans a wide array of elements that contribute to its entirety across different dimensions and scales (Legendre & Legendre, 2012; Heydari et al., 2020). Despite lacking a standardized definition, the concept of biodiversity is recognized as one of the most important properties of ecosystems, as it entails numerous services to humans (Ricotta, 2005; Hooper et al., 2005; Legendre & Legendre, 2012; Heydari et al., 2020). Nonetheless, forest biodiversity is subjected to a wide range of threats and pressures, i.e., habitat loss and degradation, shifts in species composition caused by human-induced disturbances, environmental change, invasive species, overexploitation, and pollution (Fazan et al., 2020; Guo et al., 2020; Betts et al., 2022; Wiebe & Wilcove, 2025). As a result, forest biodiversity is declining (Brockerhoff et al., 2017; Helfenstein et al., 2025).

Several international initiatives and agreements are dedicated to preserving biological diversity. For example, biodiversity, its protection, enhancement, management and utilization is among the core topics of international policies as reflected in the United Nations Convention on Biological Diversity (UN-CBD). UN-CBD defines biodiversity as "[...] the variability among living organism from all sources including inter alia, terrestrial, marine and other aquatic ecosystems and the ecological complexes of which they are part; this includes diversity within species, between species and of ecosystems" (Thompson et al., 2009). In the scientific community, a universally accepted definition of biodiversity is lacking, but often referred to as "the variety of all life forms, their ecological roles and the genetic diversity that they contain" (*sensu* Wilcox, 1984; Heydari et al., 2020). In addition to the three components of biodiversity commonly addressed in forest ecosystems, i.e., composition, structure, and function (Alberdi et al., 2010), the genetic variability expressed by species communities has gained increased interest (Gaggiotti et al., 2018). Moving beyond addressing species identity as nominal variable, (as "species-neutral", sensu Chao et al., 2010), incorporating species biological relatedness (taxonomic, phylogenetic, or functional) broadens the means to quantify forest diversity by considering interspecies biological variability (Clarke & Warwick, 1999, 2001; Hao et al., 2019a, Chapter 1.2). Hence, our understanding of biodiversity has started to extend to the taxonomic, phylogenetic, genetic and functional dimensions in which each species contributes uniquely to the overall complexity. This broader view is reflected in the Kunming-Montreal Agreement of the Global Biodiversity Framework, where genetic diversity is explicitly addressed as part of a monitoring framework to be established (UNEP, 2022).

For any type of scientific assessment or survey in forest science, the concept of "forest biodiversity" needs to be operationalized. Accordingly, any monitoring system requires the definition and provision of standardized, applicable protocols observing respective indicators. Within these protocols, specific variables are assessed, and their observed values, aggregates, distributions, or temporal changes serve to inform particular information needs. Some variables directly provide the needed information, while others must be combined into indicators to address information requirements. The inherent complexity of "biological diversity" that manifests across multiple scales requires its different dimensions to be broken down into simple indicators whose attributes can be recorded in the field, derived from several variables, or measured via remote sensing (RS, Chapter 1.4.2). As genetic diversity is recognized as a fundamental level of biodiversity, and the availability of tools for its quantification is continuously increasing, the operationalization of indicators for its assessments is essential.

The field that integrates population genetics, landscape ecology, and spatial statistics to understand how landscape features influence genetic diversity and population structure has been termed landscape genetics (Storfer et al., 2007; Capblancq & Forester, 2021). Driven by advances in high-throughput sequencing (HTS, Zizka et al., 2020), the field has rapidly evolved over the past decade, developing new statistical methods to incorporate genomic data from both plants and animals to study adaptive genetic variation, including gene flow, local adaptation, and responses to environmental change (Holderegger et al., 2010; Schoville et al., 2012; Manel & Holderegger, 2013; Epps & Keyghobadi, 2015). HTS combines several sequencing platforms, thus allowing a rapid and cost-effective processing of large numbers of DNA molecules in parallel, enabling comprehensive studies such as whole genome and transcriptome analyses, mutation detection, and epigenomic profiling (Gobet et al., 2010; Reuter et al., 2015; Eberle et al., 2020; Lee, 2023). Due to its capacity to generate large volumes of high-quality genetic data, this technology has become indispensable in genomics, biodiversity assessments, and studies on community composition, while also facilitating the continuous refinement of phylogenies (Kumar et al., 2022b; Capblancq et al., 2024). Recent applications extend its use to landscape-scale analyses of genetic variation patterns (Bragg et al., 2015; Balkenhol et al., 2019). These approaches offer powerful insights into evolutionary processes, species distributions, and local adaptation, with important applications in conservation and management (Manel & Holderegger, 2013; Kumar et al., 2022a; Lapin et al., 2025).

Whereas intraspecific genetic variability is difficult to measure at large scales, interspecies genetic variability can be directly quantified if taxonomically standardized species records are provided. Integrating genetic variability into standard biodiversity assessments enables the use of phylogenetic dissimilarity between co-occurring species to reveal evolutionary and functional relationships. These patterns remain undetected by traditional diversity measures, which do not incorporate the biological information embedded in species co-occurrence (Ricotta & Pavoine, 2015). The focus here is on phylogenetic diversity (PD), expressed as the ancestral relationships among species, thereby extending directly observed diversity metrics to incorporate their evolutionary lineage structure (Chave et al.).

As advances in HTS and other genomic tools facilitate increasingly rapid genetic data analyses, ever more comprehensive and standardized phylogenies become freely available online. Consequently, species PD can be easily operationalized for community assessments as an indicator for genetic diversity, producing “genetic landscapes” that reflect phylogenetic structure at broad scales. Despite this potential, phylogenetic diversity has rarely been integrated into operational forest inventory systems, which continue to rely predominantly on taxonomic metrics (Chapter

1.4.1). There remains a lack of studies demonstrating how PD can be applied to inform classification and monitoring in real-world forest inventory frameworks.

This study operationalizes PD for forest classification within the framework of large forest inventories (Chapter 1.3). In doing so, I focus on a specific aspect of forest biodiversity, namely interspecies genetic diversity, defined within a phylogenetic framework and assessed across observed species compositions. Hence, this research examines whether the integration of phylogenetically informed diversity metrics can effectively classify forest communities using real-world data from the first National Forest Inventory of Georgia (GNFI, Chapter 1.4.1.1).

To the best of the author's knowledge, no previous study has applied PD metrics to cluster forest assemblages using National Forest Inventory (NFI) data despite the relevance of such inventories for large-scale biodiversity monitoring. This study expands the application of PD by examining spatial patterns of phylogenetic dissimilarity across an entire NFI. Using phylogenetic dissimilarity-based cluster analyses weighted by basal area (BA), species communities are classified, thereby revealing patterns that reflect the phylogenetic structure of woody species assemblages at the landscape level. By combining phylogenetic and structural indicators within a robust forest inventory framework, this research provides scalable biodiversity reference data to guide restoration planning and ecosystem-condition monitoring in temperate forests.

Given the growing emphasis on biodiversity monitoring, this study contributes to the development of indicators suitable for large-scale forest diversity assessments (Chapter 1.4), which if based on permanent observational plots produce invaluable diversity data at low and predictable costs across extensive areas. In line with stipulated commitments of initiatives such as the Kunming–Montreal monitoring architecture (UNEP, 2022) and the objectives of the EU Biodiversity Strategy for 2030 (European Commission, 2021), this study combines forest mensuration with methods from numerical ecology to establish a multidisciplinary and statistically rigorous framework for analyzing forest biodiversity. Drawing on the theoretical framework of ecophylogenetics (Chapter 0), this study explores whether classification approaches that incorporate phylogenetic variability enhance our understanding of community assembly processes when applied to large-scale forest inventory data (Chapter 1.6). This integration is made possible by the rapid expansion of available mega-phylogenies and the analysis tools required to process large datasets.

In conceptual terms, this study contributes to a methodological shift in forest biodiversity research by integrating phylogenetic information into forest classification and analysis. By incorporating evolutionary dissimilarity, the analysis moves beyond conceiving species composition as a mere list of nominal categories, assuming ecological neutrality across taxa. This approach allows for a broadened understanding of community assembly by embracing interspecific variability, niche conservatism, and trait-based filtering processes. It also contributes to the discourse on ecophylogenetics, by investigating if evolutionary relationships can be meaningfully applied to classification at the landscape scale, resulting in interpretable distribution patterns across forested landscapes. The specific research objectives, along with the methodological approach taken to address them, are outlined in detail in Chapter 1.7.

To guide the reader, the following section outlines the structural logic of the dissertation and explains its conceptual rationale. Chapter 0 introduces the concept of ecophylogenetics and briefly summarizes the competing conceptual notions of the field. Forest diversity and its assessment, within the context of forest inventory is presented in Chapter 1.2, followed by a summary on forest classification in Chapter 1.3. Subsequently, I present an overview of forest assessments in Chapter 1.4 ff., a short introduction of the GNFI (Chapter 1.4.1.1), and a brief overview of remote-sensing based methods and Earth observation (EO) data for forest diversity assessment (Chapter 1.4.2). Georgia is presented as the study area of this thesis in Chapter 1.5, including condensed descriptions of current forest management and monitoring issues (Chapter 1.5.3 and 1.5.4., respectively). Chapter 1.6 describes the technical backbone of the research, by outlining the data processing and management environment applied for the analyses of this large dataset. Concluding the introduction, Chapter 1.7 defines the specific research questions and provides an methodological overview of the chosen approaches.

Building on the conceptual considerations presented in this introduction the following Chapters apply the proposed framework to empirical data from GNFI. Chapters 2 to 4 represent the empirical core contribution of this dissertation. Focusing on distinct aspects of PD integration, these Chapters highlight its potential for classification, monitoring, and structural assessment within NFI frameworks on the basis of two peer-reviewed and one submitted research articles. The presented studies demonstrate how PD can be operationalized to improve our understanding of ecological processes and the evolutionary history that shape them at large spatial scales.

The study presented in Chapter 2 demonstrates that accounting for interspecies variability enhances the discrimination of woody species assemblages from GNFI data. By directly comparing the results of unsupervised cluster analyses using both species-neutral and phylogenetically-weighted dissimilarity measures, this research shows that classifications align closely with elevation and forest typology maps, when PD is considered. To support the observation that phylogenetic dissimilarities between woody species assemblages reflect relative position along biogeographic gradients, environmental correlates are used to model membership of phylogenetically grouped clusters in Chapter 3. Providing empirical evidence that cluster membership can be predicted via multivariate EO data and machine learning, the study supports the concepts of ecophylogenetics which link community assembly processes to environmental filtering. However, prediction accuracy is shown to be constrained by external impacts from disturbances or nonnative species occurrences. Chapter 4 explores structural variability between phylogenetically informed clustering along a field-assessed degradation gradient. By quantifying structural changes within clustered forest formation, it demonstrates that responses to increasing degradation are formation specific and non-linear, thereby highlighting a meaningful conceptual distinction between short-term disturbance dynamics and long-term degradation processes in forest assessments from a structural perspective.

An integrative synthesis that consolidates the thesis's core findings is provided in Chapter 5. From a broader perspective, the contributions to an ecophylogenetic understanding of forest biodiversity of this thesis are discussed and overarching conclusions, challenges, and opportunities for future research are highlighted. In Chapter 5.1, the methodological and ecological implications of incorporating phylogenetic structure into forest inventories are compiled, and their relevance for forest classification and management is explored in Chapter 5.2. The final Chapter 5.3 is dedicated to the suitability and potential of large-scale forest inventories as a tool for biodiversity monitoring.

1.1 Ecophylogenetics

The field of phylogenetic ecology (ecophylogenetics) examines how evolutionary history shapes the composition of ecological communities by integrating phylogenetic relationships to investigate the mechanisms underlying local community assembly (Mouquet et al., 2012; Davies, 2021). It is based on the concept that interspecies phylogenetic of co-occurring species are generally not stochastic but determined by environmental factors and niche divergence (Provete, 2013). Emerging in the late 1990s from efforts to incorporate evolutionary history into community ecology, it focuses on the influence of regional and historical processes on local biodiversity patterns (Webb, 2000; Webb et al., 2002).

A central assumption is that phylogenetic clustering occurs because of environmental filtering, meaning that species of similar traits or function also tend to be phylogenetically similar, because of the same environmental factors that shaped them evolutionary, i.e., gymnosperms occupying mostly niches of climatic extremes (Vamosi et al., 2009; Cadotte et al., 2013; Ndiribe et al., 2013, Chapter 2). This phenomenon can be most easily observed along biogeographic gradients on larger scales, e.g., temperature, aridity (e.g., occurrence of Cactaceae) soil nutrient availability, as demonstrated by Fine & Kembel (2011). Whereas environmental filtering is often identified as a key process that separates species with suitable traits for a specific environment, competition mechanisms have also been shown to influence community composition, e.g., when closely related species may mutually exclusive due to overlapping niches (McGill et al., 2006; Godoy et al., 2014; Cadotte, 2023). As a result, phylogenetic overdispersion tends to occur where competitive exclusion or other coexistence mechanisms lead to niche differentiation and larger phylogenetic dissimilarity between co-occurring species at finer spatial scales (Cavender-Bares et al., 2006). However, key debates persist, i.e., the relative influence of environmental filtering compared to competitive interactions (Kraft et al., 2015; Cavender-Bares et al., 2018).

Increasingly, consensus is reached to consider that these processes are not mutually exclusive as both mechanisms may operate concurrently, with relative importances shifting across spatial and environmental gradients (Mayfield & Levine, 2010; McGill & Magurran, 2011; Cadotte, 2017; Davies, 2021).

A second debate revolves around the question to which degree phylogenetic relatedness serves as a reliable proxy variable for trait similarity or functional properties of a species. Based on the assumption that ecological traits are evolutionary conserved, closely related species are thus expected to share more similar trait values than distantly related species. If phylogenetic relatedness reflects ecological similarity, hence, phylogenetic variation can serve as proxy for functional resemblance (Davies, 2021). Consequently, competition for limiting resources would lead to phylogenetic overdispersion by inhibiting the co-existence of closely related species. Many studies report strong phylogenetic signals in key ecological traits, supporting the use of phylogenies in inferring community assembly processes (Cavender-Bares & Wilczek, 2003; Cavender-Bares et al., 2018). The concept of environmental filtering assumes that phylogenetic clustering occurs because close relatives share similar environmental preferences (Davies, 2021). Supporting this notion, other studies have shown that traits are not phylogenetically conserved, but can evolve rapidly or converge across distant lineages, especially in response to environmental pressures or disturbance regimes, thereby weakening the link between phylogeny and species function (Valladares et al., 2015; Mazel et al., 2018).

As a result, the interpretation of scale-dependent patterns resulting from divergent dynamics in phylogenetic community structure remains a subject of ongoing debate (Antão et al., 2019). Observed patterns such as clustering or overdispersion often change depending on the spatial, temporal, or taxonomic resolution of the analysis (Swenson et al., 2006; Qian et al., 2013). Consequently, questions about the most appropriate scale for detecting underlying assembly processes and about the extent to which conclusions can be generalized across different ecological contexts have emerged (Graham & Fine, 2008).

Regardless, the field of ecophylogenetics has solidified the conceptual transition from understanding species identity as independent, nominal category to a perspective of considering relational and evolutionary aspects as part of community structure (Slik et al., 2018; Guillory & Brown, 2021). As phylogenies have become widely available and easily accessible, ecophylogenetics has expanded its relevance beyond community ecology, while continuing to refine its theoretical and methodological foundations. With the numbers of identified species sequenced via gene expression (i.e., HTS) continuing to grow rapidly, phylogenetic concepts have found their way into biodiversity research and conservation (Tucker et al., 2017). Once a surrogate for contexts in which phylogenetic information on recorded species was scarce, the bias inherent in taxonomic relatedness is reduced and with the increasing access to comprehensive ready-for-analysis megaphylogenies, researchers can now retrieve phylogenetic relationships from existing databases, facilitating a broader application of the underlying concepts. E.g., Jin & Qian (2022) report an exponential increase in publications using the tools applied to generate plant phylogenies since 2016, with 180 studies published in 2021 alone.

Ecophylogenetic concepts embed aspects of evolutionary history into ecological analyses focusing on community development and structure. Considering the phylogenetic relatedness of species within ecological communities allows to observe phylogenetic clustering or overdispersion, which reveal evolutionary dynamics (e.g., via shared ancestry) indicative of ecological functions of species, i.e., resource use or dispersal strategies and their respective role within ecosystems (Tarkhnishvili et al., 2012). Across large spatial and temporal scales, these concepts provide means to understand the evolutionary processes that shape the macroecological patterns observed today. By acknowledging that evolutionary relationships among species (whether co-existing or generally exclusive) are not stochastic but manifest in their ecological interactions and consequently shape community structure across spatial, temporal, and environmental scales, the field of ecophylogenetics bridges evolutionary history and ecological understanding and provides a comprehensive framework for studying the dynamics of biodiversity. This conceptual framework forms the basis for the analytical approaches used in Chapters 2 to 4.

1.2 Forest Diversity

To operationalize PD in forest inventories, it is essential to define its position within the broader framework of forest biodiversity dimensions. Understanding forest biodiversity as a multi-dimensional, hierarchical construct requires recognizing its expression across multiple organizational levels (i.e., genetic, taxonomic, phylogenetic, functional, structural, community, ecosystem, biome), each with distinct spatial and temporal scales (Flynn et al., 2011; Naeem et al., 2016; Gaggiotti et al., 2018; Jin et al., 2020). Genetic diversity refers to the intraspecific variability in genetic composition and can be observed on various scales. Beyond intraspecies variation, genetic diversity can also be assessed at interspecific, landscape, biome, or continental scales (Holderegger & Wagner, 2008; Kumar et al., 2022a). Spatially informed representations of genetic variation across a region are used to understand how terrain, habitat, and ecological barriers shape genetic connectivity and population structure (Balkenhol et al., 2019). Taxonomy represents the scientific discipline that differentiates and organizes species in hierarchical systems based on morphology to support communication and understanding of biodiversity and derive biological relationships between species (Bevilacqua et al., 2021). Functional diversity describes the variation in the characteristics of a species (e.g., competitiveness, resource use, and dispersal strategies) that influence ecosystem processes and the co-occurrence of species (Tilman et al., 1997; Petchey & Gaston, 2002, 2006). A widely cited ecological definition describes it as “the value and the range of those species and organismal traits that influence ecosystem functioning” (Laureto et al., 2015).

Diversity is traditionally quantified according to its organization on different spatial levels with alpha (α) diversity representing local diversity within individual sites or samples (within-group), and gamma (γ) diversity describing diversity at the landscape level (Whittaker, 1960; Jost, 2007; Legendre & Legendre, 2012). On an intermediate scale, beta (β) diversity is conceptually distinct, as it quantifies the variation in species (species turnover) between sites or samples (between groups) within a given area and is thus inherently comparative, assessing heterogeneity, rather than diversity in a stricter sense (Tuomisto, 2010; Mammola et al., 2021). It captures how diversity is structured in time and space, or along environmental gradients (Legendre & Legendre, 2012). Beta diversity of communities is expressed as the dissimilarities among all pairs of sampling units. Not being an absolute measure, beta diversity is highly dependent on sampling design and overall scale of assessment (Anderson, 2018; Baselga & Gómez-Rodríguez, 2019).

From a statistical point of view, diversity indicators (i.e. species richness) measure categorical (nominal) variables based on the distribution of distinct species and their relative abundances, independent of any notion of dissimilarity between them (Pielou, 1966; Ricotta & Avena, 2003). In contrast, variance in continuous (metric) variables, simultaneously reflects both the spread and magnitude of differences among observations around the mean (Legendre & Legendre, 2012; Gregorius & Kosman, 2017). Numerous indicators have been developed to quantify diversity across the three aforementioned levels (Jost, 2006; Pellens & Grandcolas, 2016; Ricotta & Podani, 2017). Traditional indicators, which rely on species presence and the evenness of their relative abundances, typically assume that all species are equally distinct by assigning each an attribute value of one (Magurran & McGill, 2011; Chao et al., 2014). Focusing on only two aspects of diversity, these indices yield higher values when there are more species and when individuals are more evenly distributed among them (Magurran, 2005; Chao et al., 2016). However, treating species identity as a purely nominal category inevitably reduces the inherent complexity and falls short to capture the full species diversity by oversimplifying the inherent variation within species assemblages (Leinster & Cobbold, 2012). When species are treated as nominal variables, interspecific differences (and hence diversity) are not explicitly measured, as the following simplified example illustrates: Two sites A and B are assessed, with A containing three species, *Quercus robur* L, *Castanea sativa* Mill. and *Fagus orientalis* L., and on B, *F. orientalis*, *Tilia cordata* Mill. and *Acer platanoides L.* are recorded. Based on species richness alone both sites would be considered equally diverse. If abundances of all five recorded species are similar, even abundance sensitive indices (i.e., Bray-Curtis dissimilarity, Bray & Curtis, 1957) will yield similar diversity estimates for both samples. However, when interspecies dissimilarity is considered, the higher diversity of site B emerges. Site A contains only *Fagaceae*, while site B spans three families (*Fagaceae*, *Malvaceae*, *Sapindaceae*), revealing higher diversity when interspecies dissimilarity is taken into account. Hence, from a conservation perspective, a community composed of phylogenetically or taxonomically distinct species (sample B) is considered more diverse than one dominated by closely related taxa (sample A), assuming equal species richness and respective abundance (Davies & Buckley, 2011; Chao et al., 2016). The idea that traditional species diversity measures based solely on abundance could be expanded to include phylogenetic, functional, or other forms of interspecies dissimilarity was first recognized by Pielou (1975). Phylogenetic dissimilarity can be quantified using evolutionary relationships, whether derived from taxonomic hierarchies or robust phylogenetic trees (Bacaro et al., 2007). Such measures are particularly valuable in conservation, as they reflect the evolutionary history represented within an assemblage (Pellens & Grandcolas, 2016).

Departing from the notion that all species contribute equally to overall diversity, species identity can instead be embedded within a framework that quantifies interspecies relationships. Such frameworks exist for taxonomic, phylogenetic and functional relationships.

Taxonomic Diversity

For taxonomic diversity, species can be assigned attribute values that reflect their taxonomic distinctiveness, typically based on the lengths of branch segments within a taxonomic tree (Clarke & Warwick, 1999; Chao et al., 2014; Hao et al., 2019a). In addition, considering species organization within, e.g., the Linnaean taxonomic hierarchy allows for approximating diversity, i.e., by calculating species-to-genus or species-to-family ratios when organisms are identified only at the genus or family level (Chave et al.; Qian & Jin, 2021). However, the respective ratios can differ substantially across regions and are highly sensitive to sample size and sampling effort, which limits the reliability of this method (Gotelli & Colwell, 2001). One could argue that underlying taxonomical hierarchical relationships are not directly anchored in nature itself, but result from the methods of classification chosen by taxonomists, geneticists or numerical ecologists (Bevilacqua et al., 2021). Accordingly, respective distances between taxa appear arbitrary to a certain extent, as the organization is based on human concepts of distinction or model-based approaches and undergoes frequent adjustments (Geiger et al., 2016). Taxonomy, as categorical hierarchy offers limited resolution of the nuanced dispersion of evolutionary genetic shifts that lead to the establishment of separate species. Clarke & Warwick (1998) introduced average taxonomic distinctness as a measure to quantify within-sample taxonomic breadth, calculated as the average taxonomic distance between all species pairs. This approach was later extended by Clarke et al. (2006) to allow for the inclusion of dissimilarity-based methods by incorporating pairwise taxonomic or phylogenetic relationships between species across samples.

Functional Diversity

Functional diversity (FD) research has expanded at an accelerating pace over the past two decades (Davies et al., 2016; Mammola et al., 2021). FD corresponds to the functional distance between species pairs, expressed either in dissimilarity matrices or in multidimensional trait spaces (Cadotte et al., 2013; Mammola et al., 2021; Lepš & Bello, 2023). When all interspecies distances are treated as equal, each species pair is considered functionally equally distinct, effectively assuming a uniform dissimilarity among functional entities (Mason et al., 2005; Botta-Dukát, 2005; Laliberté & Legendre, 2010; Gaüzère et al., 2022).

Although certain functional traits may reflect underlying evolutionary relationships, taxonomic and functional frameworks typically account for phylogenetic relatedness only indirectly or to a limited extent (Mazel et al., 2018). However, information on shared evolutionary history is fundamental to a more comprehensive understanding of community diversity, as it reflects functional differences among taxa and the potential ecological cohesion or divergence within species assemblages (Srivastava et al., 2012; Zhang et al., 2020; Bevilacqua et al., 2021; Davies, 2021). Quantifying PD allows researchers to estimate evolutionary relatedness, thereby offering a finer resolution of community structure by accounting for deep-time divergence among co-occurring species (Chave et al.).

Phylogenetic Diversity

The concept of PD which considers species not as equal contributors to what we assess as biodiversity, but as representatives of distinct evolutionary lineages was introduced by Faith (1992) around three decades ago. PD has become a foundational concept in ecology for identifying not just how many species are present, but how different they are in evolutionary terms, providing a critical link between biodiversity and ecosystem function with significant implications for conservation management (Cardillo, 2023). PD (*sensu* Faith, 1992) is defined, analogous to taxonomical diversity, as “the total branch-length in the (phylogenetic) tree linking taxa in a sample”. Phylogenetic models, which generate trees linking taxa within a sample, focus conceptually and explicitly on evolutionary pathways through time, whereas cladistic approaches may rely on morphological or other traits to infer relatedness, which may or may not align with evolutionary pathways (Hinchliff et al., 2015). Previous studies proposed that PD could be used as a proxy for FD, based on the assumption that evolutionary similarity of species tends to result in similar traits (Cadotte et al., 2008). Since FD is commonly based on a limited number of measured traits, PD is expected to have a greater explanatory power, as it potentially represents a more inclusive overall measure of functional strategies (Cadotte et al., 2008; Hao et al., 2018; Mammola et al., 2021).

Genetic Diversity

Phylogenetic trees for vascular plants can be constructed using a combination of morphological, fossil, and genetic data (Doyle, 1998; Zanne et al., 2014). While morphological characteristics have historically been essential for resolving early vascular plant relationships (Soltis et al., 1997; Niklas & Crepet, 2020), genetic evidence has become increasingly central to modern phylogenetic analyses (Jin & Qian, 2019, 2022). Modern phylogenies of vascular plants are constructed using a variety of molecular, computational, and statistical approaches, building on foundational studies such as Chase et al. (1993) and the Angiosperm Phylogeny Group (APG I, 1998), which established early plastid and nuclear ribosomal DNA markers for resolving deep evolutionary relationships (Calonje et al., 2009; Smith & Brown, 2018). Advances in sequencing technologies enabled comprehensive phylogenomic analyses, including plastid genomes (Felsenstein, 1981; Jansen et al., 2007) and transcriptomes (Wickett et al., 2014; One Thousand Plant Transcriptomes Initiative, 2019), improving phylogenetic resolution across green plants. Here, supermatrix approaches can combine extensive datasets of multiple gene sequences into a single, comprehensive matrix to reconstruct large-scale phylogenies (Smith & Brown, 2018). In addition, Bayesian methods are be employed for probabilistic modeling to estimate phylogenies from molecular sequence data (Yang & Rannala, 1997). Landmark studies like Zanne et al. (2014) demonstrate the utility of mega-phylogenies for macroevolutionary inferences. Current efforts are based on HTS approaches i.e., DNA barcoding of eukaryotic organisms, which greatly increase practicability, in terms of required time and costs and enhance precision of phylogenetic classification (Hebert et al., 2003; Ratnasingham & Herbert, 2007; Zizka et al., 2020). Using DNA barcoding, research alliances i.e., the “International Barcode of Life” (iBOL), strive to systematically establish comprehensive DNA barcode reference libraries to support biodiversity monitoring and conservation (Adamowicz, 2015). Yet, whereas correlations between abundances and read numbers per sample have been observed, abundances cannot be reliably determined from assessments focused on metabarcoding (Zizka et al., 2020). As specimen collections are rarely systematic and metadata assessment is not standardized, modeling approaches incorporating auxiliary data, e.g., via RS or EO data (Chapter 1.4.2), are constrained (Kissling et al., 2018; König et al., 2019; Kattge et al., 2020). However, the advances in DNA barcoding and phylogenomic techniques, combined with open-science initiatives i.e., iBOL or the “Open Tree of Life” are enabling continuous refinement, and expansion of phylogenetic backbones from which reliable phylogenies can be constructed (Hinchliff et al., 2015; Mirarab & Warnow, 2015; Hao et al., 2021).

Hence, for any list of species representing a recorded species pool, phylogenetic relationships can be derived from localized, frequently updated phylogenies available online (Qian & Jin, 2016; Jin & Qian, 2022). To construct a hierarchical tree reflecting taxonomic relationships from the same list, taxonomic distances between species can be estimated based on e.g., six ranks of the Linnaean classification system: species, genus, family, order, and higher-level groupings such as angiosperms and gymnosperms (Clarke & Warwick, 1998; Hao et al., 2021). When plotting these relationships using dendrograms, branch lengths in the taxonomic tree are typically uniform or arbitrary, reflecting hierarchical ranks rather than evolutionary distance, whereas the phylogenetic tree shows unequal branch lengths that represent estimated divergence times or genetic distances (Figure 1.1, Hao et al., 2019b).

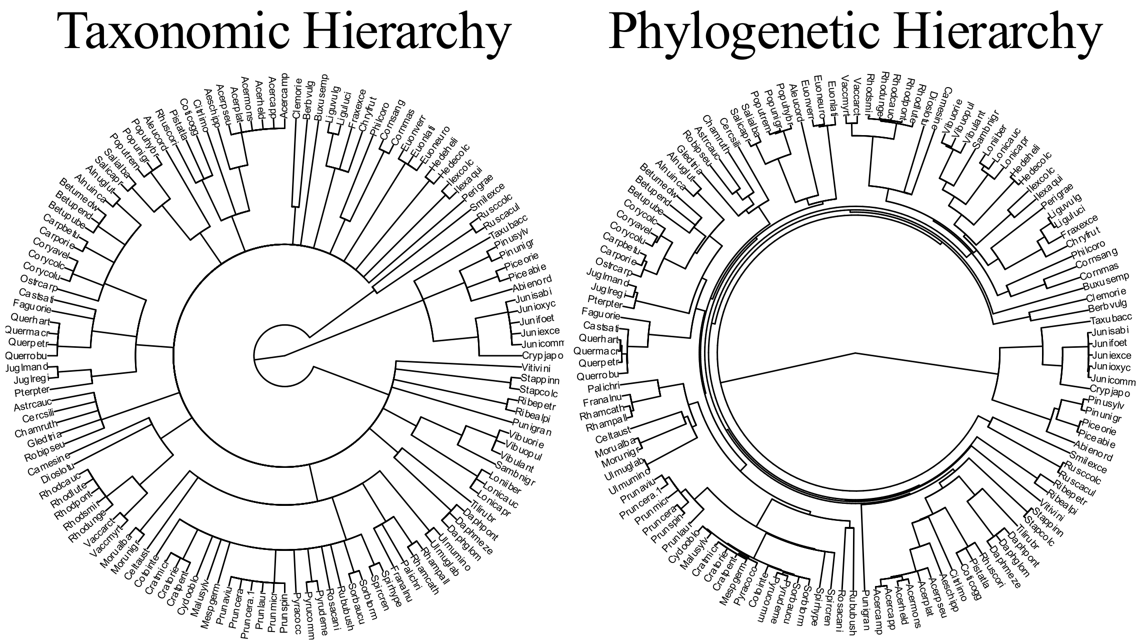


Figure 1.1: Hierarchical relationships of 132 species recorded in a subsample of the Georgian National Forest Inventory. Taxonomic interspecies distances are shown on the left; phylogenetic interspecies distances are shown on the right.

While overall species groupings may appear similar in Figure 1.1, the topology of the phylogenetic tree reflects a more nuanced organization, resulting from the incorporation of genomic data that reveal patterns of evolutionary history not captured by traditional taxonomy based on morphology alone (Soltis & Soltis, 2016; APG IV, 2016). The Linnean taxonomic classification is dynamic and frequently adjusted, especially as complete species inventories and genetic validation through DNA sequencing are increasingly available (Applequist, 2015; Geiger et al., 2016; Zizka et al., 2020). Nonetheless, taxonomy remains a useful proxy where full phylogenies or functional traits of all recorded tree species are not readily available, e.g., for rare species or disputed taxa, such as in many tropical forest ecosystems (Hao et al., 2019b; Gadow et al., 2021).

Based on this conceptual framework interspecies dissimilarity is calculated as the cophenetic distance between each pair of species in a hierarchical cluster structure, i.e., a phylogeny displayed as dendrogram (Potter & Woodall, 2012). The cophenetic distance or similarity between two taxa corresponds to the level in a dendrogram at which both taxa are first grouped within the same cluster (Figure 1.1). Any dendrogram can be represented as a matrix where each entry reflects the cophenetic value for a given pair of taxa. Cophenetic correlation coefficients, which range from –1 to +1, express the degree of correspondence between two such matrices (Legendre & Legendre, 2012). A positive correlation is expected when distances are compared to cophenetic distances (or similarities to cophenetic similarities), while a negative correlation arises when distances are compared to similarities. A higher absolute correlation value indicates stronger agreement between the structures being compared (Kindt & Coe, 2005). The relationship between two hierarchical classifications derived from different datasets, e.g., species composition and environmental gradients, can be statistically tested using matrix correlation or other consensus-based approaches (Legendre & Legendre, 2012). In this context, the Mantel test offers a robust method for assessing the significance of the correlation between two distance matrices, providing a formal statistical framework for evaluating the similarity between underlying structures, i.e., dendrograms (Chapter 2).

Although derived using similar approaches, taxonomic, phylogenetic, and functional diversity fundamentally differ in the aspects of biodiversity they each capture. Taxonomic diversity captures the relationship of species present, whereas FD captures the range of trait differences between them. In contrast, PD captures the evolutionary relationships between lineages, encompassing ancient and recent divergence. While these different facets of interspecies diversity are sometimes congruent, including phylogenetic variability allows for the integration of evolutionary history in biodiversity assessments (Potter & Woodall, 2012; Wilcox et al., 2018; Ali et al., 2025). For example, a forest that contains many species that are taxonomically similar (high species richness) might display low PD if these species are from a narrow evolutionary lineage. In contrast, relatively species poor forests can be phylogenetically diverse if the occurring species are only distantly related. The ability to distinguish between these aspects of diversity has profound implications for conservation management (Zheng et al., 2022).

As such, integrating species distribution, trait, and phylogenetic data offers significant potential to advance community ecology by supporting informed analyses of community assembly processes (Webb et al., 2002; McGill et al., 2006; Simon Véron et al., 2019; Vigués Jorba et al., 2025). Combining multiple, complementary metrics enables more nuanced interpretation of assembly mechanisms, including different modes of competition and the interplay of evolutionary history with ecological filtering (Münkemüller et al., 2020). This approach combines the ecological information embedded in species traits with evolutionary relationships to move beyond species-based descriptions toward a functional characterization of communities (Kraft et al., 2015). Researchers increasingly complement phylogenetic measures with trait information (Pavoine & Bonsall, 2011; Bello et al., 2021) or integrate both into single metrics to estimate community niche structure and overlap (Cadotte et al., 2013).

For the classification of forest communities, the “Discriminating Avalanche” index (dA) was proposed by Hao et al. (2019b; 2019a; 2021) as a beta dissimilarity metric to quantify heterogeneity based on species abundance weighted by proximity of each pair of species according to any biological diversity measure (Tallents et al., 2005). While dA provides a flexible framework for incorporating interspecies biological distances, other phylogenetically and functionally informed dissimilarity metrics exist (Chiu et al., 2014; Chao et al., 2016). Measures like dA enable the quantification of community dissimilarity by combining information on species composition, abundance, and interspecies biological relationships as defined by taxonomic, phylogenetic or functional distance. This allows the index to express lower dissimilarity values when species are closely related, and higher values when they are biologically distinct. Incorporating PD to quantify interspecies dissimilarity, dA accounts for evolutionary differences between species, reflecting functional divergence to the extent that phylogenetic distance approximates functional dissimilarity. The dA index was developed as an extension of earlier “Avalanche indices” to provide a more refined measure of community dissimilarity by incorporating biological heterogeneity among species. It builds on the original Avalanche index proposed by Ganeshaiah et al. (1997) which was further elaborated by Ganeshaiah & Shaankar (2000), and represents a generalization of the PD concept introduced by Faith (1992, Hao et al., 2019b).

To evaluate how extended diversity metrics perform in classifying forest communities using large-scale inventory data, I incorporated dA into unsupervised cluster analyses of GNFI sample-based observations. To account for evolutionary and to an unknown extent functional differences between species, I used PD to weight species abundance, assuming that PD reflects key aspects of community dissimilarity (Webb, 2000; Webb et al., 2002; 2006; Swenson et al., 2012; Cadotte et al., 2013). In Chapter 2, the methodological implications using dA for forest classification are evaluated by juxtaposing quantified dissimilarities with Bray–Curtis dissimilarities which are widely applied in vegetation classification studies.

Structural Diversity

Structural diversity describes the spatial and relational arrangement of components within a more or less well-defined system (Gadow et al., 2012). LaRue et al. (2023b) describe structural diversity as the “volumetric capacity and physical arrangement of biotic components in ecosystems”, that serves as a predictor of productivity and a proxy of niche occupancy. In a more forest specific context, according to Chirici et al. (2011) “structure” refers to vertical and horizontal distribution and physical status (i.e., standing/lying, living/dead/decomposing) of forest biomass.

In forests, a substantial proportion of organisms is represented by woody species, which define forest structure and constitute the bulk share of biomass (Macía, 2008). The structural and compositional data collected at stand level by NFIs provide valuable indicators of biodiversity (McRoberts et al., 2008; Winter et al., 2008; Moreno-Fernández et al., 2025). Key components of stand structure include the spatial distribution of tree stems, stand density, and variation in tree size, amount of deadwood and abundance of tree microhabitats. For example, tree microhabitat abundance and diversity are recognized as indicator for habitat heterogeneity (Asbeck et al., 2021; Larrieu et al., 2022; Mamadashvili et al., 2023; Zemlerová et al., 2023). Following an extensive review of indicators that can be derived from variables commonly assessed in European NFIs that describe forest structure, Storch et al. (2018) present a combined “forest structure index”. Forest structural attributes i.e., tree stem density, canopy stratification, or forest development phases offer direct and proxy variables to model habitat variability and quality (Gadow et al., 2012; Hilmers et al., 2018; Heym et al., 2021). Greater structural complexity is often associated with higher species richness and increased ecological stability (Pretzsch, 2009; Beckschäfer et al., 2013). Accordingly, forest structural diversity is closely linked to habitat provision and species richness across forest-dependent taxa (Steinmann et al., 2011; Chirici et al., 2012; Gao et al., 2014; Seidling et al., 2014). Respective data are frequently used to study how forest structure influences non-tree biodiversity (Gao et al., 2015; Reise et al., 2019; Kozák et al., 2021). Moreno-Fernández et al. (2021) used data from the Spanish NFI to develop a method to quantify forest structure via nearest neighbor indices (e.g., Clark-Evans index of aggregation, Clark & Evans, 1954) and second-order moment functions (i.e., Ripley's K-function, Ripley, 1977) adapted to the nested sample plot design to assess spatial patterns, species mingling, and size differentiation of recorded stems. Based on a subset of the same data, Cáceres et al. (2019) classified forest types based on species, tree diameter, and height and reported that with different structural metrics a similar number of forest types, but diverging class sizes are obtained. They argue that the inclusion of structural diversity captures fine-scale variations in compositional data of stands. Investigating how structural diversity relates to diversity patterns may offer valuable insights, particularly for the identification of fine-scale variation of forest diversity.

Forest communities may exhibit similar species composition in terms of observed richness and respective abundances. However, communities can differ significantly in size structure, as some may contain numerous small-diameter trees, while others may be dominated by fewer individuals that disproportionately account for total niche occupation (LaRue et al., 2023a). In addition, abundance estimates based solely on individual counts ignore existing size differences and may lead to distinct evenness profiles. In forest inventories, a convenient size metric is tree basal area, as it is directly derived from measured stem diameter at breast height (DBH, at 1.3 m), and is highly correlated with tree height and tree volume. To weight niche occupancy of forest communities according to size of constituents, tree basal area can be aggregated per species and extrapolated to total basal area (BA, m^2 ha^{-1}) as structurally weighted abundance value that captures niche occupancy (Staudhammer & LeMay, 2001; Yao et al., 2019; Cáceres et al., 2019).

Together, taxonomic, functional, phylogenetic, and structural diversity offer a comprehensive framework for forest biodiversity assessment. By integrating species distribution, structural attributes, and evolutionary relationships into NFI data, this study aims to operationalize phylogenetic diversity (PD) to improve classification approaches and inform conservation planning at the landscape scale.

1.3 Forest Classification

Vegetation classification aims to summarize the spatial and temporal variation of ecological communities using a limited set of abstract, interpretable entities. Similarly, forest classification attempts to organize forest types based on structural, compositional or functional criteria, according to specific management, conservation or research requirements. Organizing species communities along shared characteristics enables the structuring of the inherent complexity of ecological populations and facilitates the identification of meaningful patterns that allow systematic analyses or different management options. The resulting structures are fundamental for forest science and management, as only a comprehensive ecological understanding enables systematic assessments, long-term monitoring and informed decision making. By reducing the complexity of the object of interest, a systematical approach to infer the dynamics that shape these entities and their distribution is possible across spatial and temporal scales.

However, the criteria used to group entities inherently reflect the objectives of the specific application or discipline, whether for communicating complex patterns, formulating hypotheses, supporting mapping and monitoring efforts, or guiding effective management and conservation strategies (Wohlgemuth, 2000; Cáceres et al., 2015). When natural entities are classified to reduce complexity according to specific properties, other forms of variation are left unaddressed to achieve conceptual coherence (Hao et al., 2019a). As R. H. Whittaker is cited by Dolukhanov (2010) in his influential description of the forest vegetation of Georgia, "the hierarchical relationships are not anchored in nature itself, but are caused by the methods of classification chosen by ecologists and their application"[1]. This observation can be clearly extended to other disciplines, i.e., forest science. Hence, vegetation classification concepts are shaped by specific choices, which are influenced by dynamics outside the ecological realm. These include the need to integrate emerging information needs, new methods, shifting priorities, or changing foundational principles and are thus subject to continuous updates (i.e., change in taxonomic relationship, genetic classification, etc., Cáceres et al., 2015). In addition, the characterization of discriminated entities needs to reflect the information required by different end-users, which can be complemented by context-specific supplementary information related to conservation, vulnerability, management guidance, or ecosystem services. Ultimately, the logic behind the chosen classification approach is primarily shaped by the intended use of the classification system.

Most vegetation classification approaches are grounded in the assumption that ecological complexity is best represented through hierarchical structures, reflecting the nested organization of plant communities across spatial and ecological scales (Jin et al., 2020). These hierarchies often, though not always, correspond to increasing scale, from fine-grained species assemblages to broader vegetation types shaped by climate, topography, and biogeographic history. According to Maurer & McGill , *community* is defined "as all the organisms belonging to a set of species found at a given point in time and space". Meeting this holistic approach is challenging for complete terrestrial community assessments, as many organisms are either too numerous and small (e.g., microorganisms), too mobile, or too infrequent (e.g., birds, large mammals) to be reliably observed

---

[1] This is the author's translation from a German version of Dolukhanov (2010), which cites Whittaker in Russian.

during field-based inventories. Immobile woody species represent a notable exception. Even perennial plants are not always observable at any given time during their life cycle. As a result, two key implications arise: i) referring to the recorded species in forest inventories as “species assemblages” is more appropriate, as they do not capture the entire community and ii) there is a pressing need to infer communities *sensu* Maurer & McGill from a limited set of easily observable species assemblages. While the latter serves more as a conceptual framework than an attainable objective, given the inherent complexity of forest communities, forest classification offers a promising means to reduce the vast number of potential co-occurring taxa that comprise entire communities. Since forest inventories systematically record a wide range of site-specific variables (metadata) to assess correlations between environmental factors and tree growth, this data supports modeling of species assemblages. In combination with structural attributes (i.e., canopy structure, biomass distributions or deadwood) these models may enable inferences about the occurrence of a broader range of taxa.

Over the last decade, vegetation classification has seen renewed global prompting of national and international initiatives to develop standardized classification systems and databases (Peet et al., 2012; Faber-Langendoen et al., 2014; Franklin, 2015; Chytrý et al., 2016; Kattge et al., 2020; Chytrý et al., 2020; Novák et al., 2023a; Chytrý et al., 2024; Song et al., 2024; Knollová et al., 2024). However, available datasets of botanical descriptions are often limited in terms of thematic or temporal coverage, and likely not suitable for integration at multiple scales (Mace et al., 2012), resulting in limited replication potential since reported results are based on small sample sizes which can often not be replicated in similar studies (Ioannidis, 2005). Because ecological surveys are specifically designed for the assessment of biodiversity at a specific location, sample sites are rarely selected according to statistical criteria which limits the possibilities to statistically derive population properties and to compare findings across spatial or temporal scales (Cutko, 2009). As a consequence, being too local or too restricted to specific objectives rather than to provide a foundation for large scale monitoring, repeated assessments rarely attract additional financial resources. In this context, the necessity to build on existing frameworks for data collection for monitoring becomes apparent (Chirici et al., 2012).

Numerous approaches to plot-based vegetation classification have been developed, each reflecting distinct methodological traditions, priorities, and spatial scales (Costanza et al., 2018). Floristic approaches, such as those rooted in the Braun-Blanquet (Braun-Blanquet, 1968) tradition, classify vegetation based on observed species composition and abundance, typically at local to regional scales, and often rely on expert judgment within a hierarchical system (van der Maarel, 1975). The selection of plant groups or strata to include in complex ecosystems like forests is often guided by ecological objectives or technical constraints. For instance, classifications of temperate forest typically focus only on vascular plants. To address such limitations, some approaches classify distinct groups of plants with similar size and habitat use, so called *synusiae*, using separate, tailored protocols. Certain branches of phytosociology have refined the related approaches to classify distinct *synusiae* as ecologically uniform, single-layered assemblages, independently. For example, the “Integrated Synusial Approach*”*, applied in countries such as Switzerland and France, surveys each *synusia* (e.g., tree, shrub, herb, cryptogam layers) separately using distinct protocols. Once these elementary syntaxa are identified, they can be combined to describe and classify complete plant communities based on their synusial composition (Gillet & Julve, 2018). In contrast, physiognomic-structural approaches focus primarily on the physical structure, growth forms (e.g., canopy layers), and floristic composition of vegetation, and are commonly applied at regional to global scales (Küchler, 1949; Faber-Langendoen et al., 2014). Biogeographic or ecozone-based classifications delineate vegetation types based on climatic and physiographic parameters and operate mainly at large continental to global scales. In forestry, dominant species approaches, which group forest vegetation communities according to *dominance types* (based on maximum BA per species in the canopy layer), are most common. In the forestry sector of the former USSR, the classification of Sukachev & Dylis (1964) was widely adopted, combining species dominance in the canopy layer with characteristics of the lower vegetation strata and site properties to create a hierarchical system that captured both compositional and structural aspects of vegetation (Ivanova & Zolotova, 2014). Existing classifications in forestry in Georgia are traditionally based on qualitative assessments, such as groups of vegetation communities or delineated management units according to dominant tree species, stand age and site properties (i.e., site index, *ტყის ბონიტეტი*). The latter is the dominant concept applied in forest management planning, which simplifies the existing complexity of the natural conditions for practical purposes. Dolukhanov (2010) distinguished between “mono-, oligo-, and polydominant forest formations” and observed that the majority of forest communities in Georgia are oligodominant. In the tradition of Soviet and Eastern European ecological science, the concept of *biogeocoenosis* (Petzer, 2023) has long served as a foundational framework for understanding forest ecosystems (Dolukhanov, 2010). Unlike the Western concept of “ecosystem” (the term biogeocoenosis became known from the work of Sukachev & Dylis, 1964), this term describes a tightly integrated unit of living organisms and their physical environment, functioning as a cohesive, self-regulating system defining a specific “forest type” (Ivanova et al., 2022). As Dolukhanov (2010) notes in his typological description of Georgian forests, this framework has been instrumental in characterizing forest formations based on the dominance and interactions of tree species. However, as the biogeocoenotic concept was developed to classify primary forests with limited anthropogenic impact, its application to commercial forestry systems has been criticized to not yield optimal outcomes (Ivanova & Zolotova, 2014; Ivanova et al., 2022). According to this concept, monodominant forests are characterized by a single tree species acting as the primary ecological architect, or *edificator*, shaping the structure and function of the biogeocoenosis (Dolukhanov, 2010). Oligodominant forests, more common across Georgian forests (Figure 1.2), are dominated by two or three tree species. In contrast, polydominant forests involve a more complex interplay among multiple co-dominant species, where no single group of edificators maintains consistent dominance. To illustrate the situation based on available data from Georgia, 4,499 sample plot observations of the GNFI were grouped according to relative BA per species. Samples with relative BA ≥80% were classified as monodominant, samples with a maximum relative BA ≥30%, were classified as oligodominant. All remaining samples (with maximum relative BA <30%) were classified as polydominant. Figure 1.2 shows the distribution of samples by dominance group.

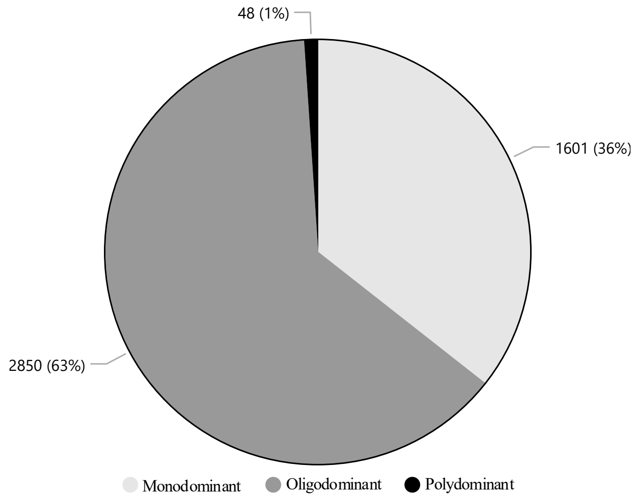


Figure 1.2: Classification of 4,499 sample plots of the dataset of the National Forest Inventory of Georgia per dominance groups according to the terminology used by Dolukhanov (2010).

Based on this classification 36% of all observations are dominated by one single species (Figure 1.2). Hence, according to the GNFI data, the majority of sample plots classify as oligodominant (63%), and an additional 1% resembles polydominant BA distributions, raising doubts about the applicability of a purely dominant species classification concept.

A traditional approach in community ecology and biogeography strives to identify characteristic plant species that are indicative of groups representing trophic sites properties, based on the assumption that specific species reliably indicate specific environmental conditions (Ellenberg & Leuschner, 2010). Hence, an effective indicator species does primarily occur within a specific group of environmentally similar sites (or few closely related groups) and is consistently present in most of the sites within that group (Legendre, 2013). To reflect their ecological preferences, plant species can be linked to abiotic gradients by assigning them values that indicate the point along the gradient where they, on average, reach peak abundance, representing their realized ecological optimum (Di Biase et al., 2023). The obtained values represent the centroid positions of a species' realized optima along one-dimensional environmental gradients and reflect long-term ecological behavior rather than momentary conditions (Schaffers & Sýkora, 2000). As integrative measures, they have been widely used in community ecology to interpret ordination results, track vegetation responses to environmental change, assess the synecology of plant communities, compile regional species pools, estimate occurrence probabilities, and support ecological risk assessment (Schaffers & Sýkora, 2000). For a given plant community, indicator values are typically aggregated into a composite score that characterizes the ecological conditions of the site. A highly influential to forest vegetation community classification in Western European forestry is the related approach by Ellenberg (1974), updated by Ellenberg & Leuschner (2010), whereas in the majority of post-Soviet countries, the indicator values developed by Ramensky et al. (1956) and Tsyganov (1984) are more commonly applied (Zolotova et al., 2023).

Algorithm-based or unsupervised classification (i.e., *cluster analyses*), which rely on resemblance coefficients and are based on (dis-)similarity of compositional attributes offer reproducible, data-driven statistical methods. These approaches are especially useful for data-heavy forest inventories and large-scale monitoring programs (Legendre & Legendre, 2012). Further, integrated approaches combine elements of physiognomic, floristic, and ecological classification, providing flexible systems applicable across multiple spatial scales, depending on data resolution and purpose (Cáceres et al., 2015).

For many years, the most widely used statistical method for identifying indicator species in vegetation classification was the Two-way Indicator Species Analysis TWINSPAN (Hill, 1979; Lötter et al., 2013; Bonari et al., 2021). The algorithm classifies vegetation data by iteratively dividing sites along the main ecological gradients, emphasizing continuous variation rather than discrete groupings. TWINSPAN was one of the first numerical methods to implement an objective, gradient-based classification by combining ordination with species-based diagnostics. It partitions the compositional space along ordination axes derived from reciprocal averaging and then forms clusters that reflect irregular compositional patterns guided by indicator species (Roleček et al., 2009). However, its reliance on a single dominant gradient and the complexity introduced by post hoc adjustments have drawn criticism (Schmidtlein et al., 2010). In contrast, ISOPAM (“Isometric Feature Mapping and Partitioning Around Medoids”) is an unsupervised clustering method that avoids these limitations by focusing on species fidelity and diagnostic species distributions rather than ordination along a single compositional axis. It combines nonlinear dimensionality reduction via ISOMAP ordination (Tenenbaum et al., 2000) with partitioning around medoids (PAM), classifying plots based on indicator species distributions (Schmidtlein et al., 2024). ISOMAP constructs a neighborhood graph of similar plots and computes geodesic distances (as dissimilarities) along shortest paths through this network. This nonlinear dimensionality reduction method preserves global structure, making it useful in contexts with high species turnover. While not a clustering method itself, ISOMAP output can be used to support hierarchical or non-hierarchical classification (Tenenbaum et al., 2000). ISOPAM optimizes cluster structure to maximize indicator species performance, supports both hierarchical and non-hierarchical modes, and allows flexible distance metrics and optional supervision through predefined species or medoids. It performs an exhaustive search across predefined parameter ranges to identify the optimal clustering solution based on specific criteria. It simultaneously optimizes three parameters: the number of clusters for PAM, the number of neighbors in the ISOMAP network, and the number of ISOMAP dimensions used for clustering. By repeatedly searching for groups that maximize species fidelity or indicator species strength until a stopping criterion is met, it produces a low-dimensional ordination less prone to artifacts and better suited for long, nonlinear ecological gradients (Schmidtlein et al., 2010). In hierarchical clustering, the algorithm recursively subdivides groups until the number of indicator species per group falls below a set threshold or group sizes drop below the minimum of three sites. Adjusting these stopping rules allows for more or fewer clustering levels. Upper and lower bounds for these parameters can be user-defined. The algorithm is available as R package (Schmidtlein et al., 2024). To implement clustering via ISOPAM based on dA-dissimilarities, the original source code had to be adjusted manually for this study. ISOPAM has been used in various ecological contexts. Černý et al. (2015) employed the algorithm to classify Korean forest types along environmental gradients. Cabido et al. (2018) classified Chaco and Espinal forests in central Argentina (Cantero et al., 2022, 2020; Zeballos et al., 2023), whereas Hein et al. (2014) applied ISOPAM to classify species composition of spider fauna in Norway. In remote-sensing data analysis, the algorithm was applied to generate soft vegetation classifications by assigning pixels probabilistic memberships to vegetation types, improving the detection of ecological gradients and transitional areas (Feilhauer et al., 2021). Comparing traditional classification with results from ISOPAM, Abe (2021) reported its effectiveness for detecting dominant vegetation patterns in large datasets of Japanese coastal vegetation.

Ultimately, the choice of the classification method remains an informed subjective one and is highly purpose-specific, as the difference in considered variables will yield different groups. From a practical perspective, especially when analyzing large datasets, automated approaches are more effective than manual grouping. Given the multivariate nature and the high variability inherent in forest properties, research into the relationships between essential aspects of diversity is required to help identify essential variables or proxies to be considered for different purposes. This study investigates an approach to classify woody species assemblages based on a combination of structural and evolutionary dissimilarity (Chapter 2). The applied methodology allows the classification of mixed forests using structural *and* phylogenetic properties across large areas. If the resulting groups maintain functional properties expressed as non-stochastic alignment along environmental gradients, a phylogenetically informed approach can be of great value to identify processes of environmental filtering that shape communities, inform conservation management or infer key variables for species distribution modeling (Chapter 3). In addition, defining forest typologies from repeated inventories of permanent sample plots can offer valuable insights into successional dynamics and genetic drift resulting from environmental change (Chapters 3 and 4).

1.4 Forest Assessment

The objective of forest resource assessments on large areas, such as NFIs, is to provide unbiased estimates of key variables and to generate statistics on forest properties. NFIs typically assess forest area, growing stock volumes, and increment to inform forest policy and monitoring. International agreements and initiatives, i.e., the Montréal Process (1998), the Ministerial Conference on the Protection of Forests in Europe (2008), the UN Framework Convention on Climate Change (UN-FCCC) and the UN-CBD, FAO's Global FRA and REDD+ (Reducing Emissions from Deforestation and Forest Degradation) rely on information supplied by NFIs (Vidal et al., 2016a; Breidenbach et al., 2021a). Due to international reporting requirements, biomass, aspects of biodiversity status, land-use, carbon stock and ecosystem services are increasingly considered. However, sampling designs, plot configurations, and assessment methods differ across countries, as do fundamental definitions of forest area and other aspects. Therefore, harmonizing assessment protocols and reporting standards is required to ensure comparability of NFI results (Winter et al., 2008; Tomppo et al., 2010; Zhao et al., 2014).

Forest assessments have a long history and are traditionally based on different approaches to categorization of the forest resource into distinct quantifiable and ultimately manageable entities, with the common objective to support decision-making on different scales. At the stand level, stand-wise evaluation of timber stocks and species composition serves to plan or control forest operations or other silvicultural interventions. Unlike systematic forest inventories, stand-wise assessments are traditionally implemented as qualitative expert-based estimates (forest taxation) with few quantitative assessments that inform about a spatially discrete forest entity (i.e., forest stands). On a larger scale, systematic forest inventories are implemented on forest enterprise or district level, with the aim to provide baseline information for a 10-year management plan (Forest Management Inventory, FMI) and to monitor overall sustainability of interventions. FMIs are often based on a network of permanent sample plots, with closer sampling grids than NFIs and provide results with estimates of statistical precision for different strata (e.g., forest types). NFIs serve as the principal source of forest data for national assessments of the forest resource (Vidal et al., 2016b). The first NFI was implemented in Norway in 1919, motivated by concerns related to unsustainable exploitation (Breidenbach et al., 2021a).

These large-scale assessments have evolved to address increasingly complex information needs, owing to their comprehensive monitoring frameworks and the wide range of ecological, structural, and socio-economic variables they record (Tomppo et al., 2010; Alberdi et al., 2017). Originally designed to provide periodic information of timber stocks on national, regional (NFI) and for forest management on forest district level (FMI), their scope has expanded to include indicators relevant to environmental and ecological policy evaluation (Álvarez-González et al., 2014; Breidenbach et al., 2021a). All NFIs adhere to statistically robust methodologies, with international processes requiring and verifying their scientific soundness to ensure data comparability and reliability across countries (e.g., REDD+). NFIs represent the most comprehensive and unbiased sources of forest information in many countries, due to their large sampling intensity and extensive set of recorded variables.

Given the broad, multidimensional, and multiscale nature of biodiversity, a statistically robust monitoring program must rely on indicators to track changes over time and space (Chapter 1.2). Without such an approach, comprehensive monitoring of large areas would be financially unfeasible. Even so, collecting, managing, and analyzing the data remains costly (Chirici et al., 2012). Hence, from a practical perspective, the synergies between NFIs and other biodiversity assessments offer a valuable opportunity to support biodiversity monitoring across multiple levels. Continuing to expand the scope of NFIs offers a pragmatic approach for facilitating biodiversity assessments in a statistically sound manner, given the repetitive, systematic, and logistically robust nature of NFIs.

1.4.1 Terrestrial Forest Inventories

At an international level, forestry science can draw on data that has been collected in large-scale inventories over the course of a century. The earliest NFIs date back to the early 20th century: 1919 in Norway, 1921 in Finland, 1923 in Sweden, and 1928 in the United States (Fridman et al., 2014; Vidal et al., 2016a; Alekseev et al., 2019). Most European countries developed and implemented NFIs that feature probability sampling designs during the later 20th century. In contrast, many Eastern European countries historically relied on aggregated stand-level data and have only recently adopted sample-based systems (Tomppo et al., 2010; Chirici et al., 2012). Scientists from other disciplines are increasingly recognizing NFI data as a key resource for ecological research, as their growing accessibility provides information across broad environmental gradients. NFIs not only offer a plethora of variables on woody species communities and structure, but being rapidly modernized, they increasingly incorporate edaphic and environmental data that describe the site conditions of sample plot locations. This metadata allows cross-referencing with climate, land-use change, and other data to better understand the drivers and dynamics of forest species diversity decline under global change (Godoy & Rueda, 2016). Therefore, NFIs constitute one of the most comprehensive sources in terms of number of observational units and field-acquired data, for policy formulation, including biodiversity conservation (Alberdi et al., 2010; Chirici et al., 2012). With the continuous shifts in forest policy and management objectives, from simple wood production to sustainable ecosystems, biodiversity and ultimately carbon pool management, demands to the information provided by NFIs have been extended according to contemporary policy requirements and scientific priorities (Tomppo et al., 2010; Corona et al., 2011; Chirici et al., 2011; Kleinn et al., 2020). Over time, timber volume inventories have evolved into comprehensive multipurpose forest monitoring systems that support both national decision-making and international reporting obligations (Kleinn, 2017); [Cannot display reference "Trucios Caciano May, 2020 – Quantifying the uncertainty caused": Template "In-text citation - Thesis - (Default template)" is not defined.]. Permanent sample plots remain the primary source of information for monitoring forest structure, diversity, and dynamics over time, despite rapid advances in predictive modeling approaches. NFIs use systematic sampling approaches that support design-based inference with confidence intervals, providing a statistically robust and transparent foundation for monitoring. As a result, these frameworks are increasingly recognized as a valuable source of reliable data for forest biodiversity monitoring (Winter et al., 2008; Corona et al., 2011; Chirici et al., 2012; Alberdi et al., 2014; Alberdi et al., 2019). However, the underlying sampling design, which comprises aspects of sample selection, configuration, size, shape, and allocation of plots, is defined by the specific monitoring objectives, whether focused on timber resources, biodiversity, carbon stocks, or ecosystem services (Corona et al., 2011). Consequently, challenges persist regarding specific biodiversity objectives, particularly concerning consistent definitions and the need for methodological adaptation (Vidal et al., 2016b).

Whereas NFIs provide precise estimates of tree species proportions at regional to national scales, obtaining spatially explicit information at finer resolutions across large areas is not straightforward (Breidenbach et al., 2021b). Nonetheless, modeling approaches can be used to regionalize and map forest characteristics. With the expanding range of available RS systems and the increasing temporal and spatial resolution of the data acquired by related sensors, there is growing potential to optimize sampling strategies via model-based approaches (Chapter 1.4.2, McRoberts & Tomppo, 2007). For instance, ancillary variables derived from multispectral imagery or LiDAR can help identify areas of high alpha-diversity or strong beta-diversity gradients, thereby enabling stratified or weighted sampling to better capture ecological heterogeneity (Feilhauer & Schmidtlein, 2009).

In this context, the GNFI exemplifies a modern, multipurpose forest monitoring system that builds on a robust sampling and data recording scheme. It provides a high number of permanent sample units described via a wide range of variables (MEPA, 2023). Nonetheless, being the first NFI to be implemented in the country, its full potential to generate information for various disciplines has yet to be realized. For this study, it serves as the primary data source for the presented analyses.

1.4.1.1 The First National Inventory of Georgia (2018-2021)

The GNFI was planned and implemented by the Government of Georgia between 2017 and 2021 to address the lack of reliable data on the country's forest resources. With the analysis results available in 2023, a statistically sound baseline for long-term monitoring and national and international reporting obligations has been established for the first time. The GNFI is mandated to repeat every ten years, according to the current legislation under Georgia's new forest code

(Government of Georgia, 2020). Financed by international donors and implemented by the Ministry of Environmental Protection and Agriculture of Georgia (MEPA), the GNFI is based on a land cover classification and forest area estimation using RS observations from the intersections of a nation-wide 3.6 x 3.6 km UTM sampling grid of randomly selected origin (MEPA, 2023). In line with international standards, field data collection was done by trained field personnel between 2018 and 2021 and employed digital data logging and processing tools from the FAO Open Foris suite to ensure standardized data recording, management and analysis. Forests in Georgia are defined as "land area with a width of not less than 10 meters and an area of not less than 0.5 ha covered with trees higher than 3 meters and a canopy cover of more than 10%, or with trees able to reach these thresholds *in situ*" (Government of Georgia, 2020). This forest definition represents the general forest definition provided by FAO to a large extent. At each grid intersection that does not clearly fall outside a forested area, cluster sample units are established. Within each cluster a comprehensive set of variables across multiple spatial scales and thematic categories is recorded on three sample plots of 0.07 ha each, arranged in an L-shaped configuration with a distance of 100 m along both axes (Figure 2.1). All sample plots are established as permanent observation plots, allowing monitoring of changes in species composition, stand structure, and stem growth comparisons. Sample plots consist of four concentric circular fixed area subplots with radii of 5, 10, 15 and 25 m (MEPA, 2018). As 18% of the country’s territory is currently not accessible for government officials due to an ongoing political conflict (MEPA, 2023), approximately 74% of the national forest area (Figure 4.3) was sampled. The GNFI assesses all woody species as single trees that can consist of multiple stems according to distinct diameter thresholds. Inclusion criteria within the three concentric, nested subplots are defined quantitatively, based on stem distance from the plot center and the corresponding DBH thresholds. Woody species with DBH ≥ 30 cm are recorded in the 15 m plot, whereas those with DBH ≥15 cm and ≥8 cm were recorded within radii of 10 m and 5 m, respectively. In contrast, understory species are defined as “Woody plants that do not and will never create an upper canopy layer under the present conditions” are assessed according to cover percentage within the 10 m radius, but “individuals that meet the definition of a single forest tree, are not described as part of understory” (MEPA, 2018). Regeneration is assessed by counting individuals of “those woody species which will later take part in creation of the main canopy layers” (MEPA, 2018) within two smaller satellite subplots (r = 1.5 m) positioned 5 m from the plot center.

The GNFI data used in this study contained 2006 cluster plot observations (Figure 1.3). As not all subplots are accessible or fall within a forest area, the dataset contained 5870 subplot records.

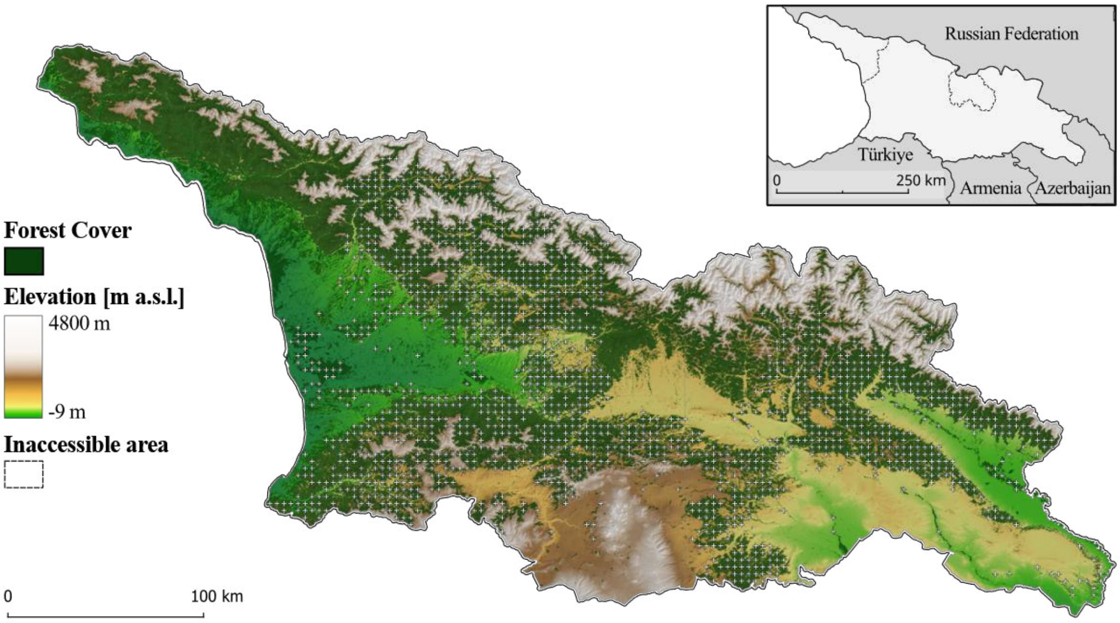


Figure 1.3: Forest cover of Georgia according to Griesbach (2018) with cluster sample plot locations of the first National Forest Inventory of Georgia (n = 2006).

According to the GNFI estimates, forests cover approximately 3,100,500 ha, representing 45% of the entire national territory. 98.5% of the assessed forest area originates from natural regeneration, with only 1.5% being plantations (MEPA, 2023). In total, 24.5% of the forest area is located in or near the territories of Abkhazia and South Ossetia, where field data collection was not possible due to restricted access. An RS-based estimation of overall stocking values in these areas is pending. MEPA reported a mean growing stock of 217.6 m^3 ha^{-1} current mean annual volume increment of 6.0 m^3 ha^{-1} a^{-1} and approximately 22 m^3 ha^{-1} of deadwood volume on average (MEPA, 2023). For reference, European approximate averages are 169.1 m^3 ha^{-1} [74.0 - 409.0 m^3 ha^{-1}] for growing stock density, average annual volume increments range from ~4 – 12 m^3 ha^{-1} a^{-1}, and mean deadwood volume ranges between 7.1 and 33.1 m^3 ha^{-1} (Tomter et al., 2016; Puletti et al., 2019; Forest Europe, 2020). Albeit respective definitions and assessment protocols can differ, the provided values serve indicative purposes.

1.4.2 Remote-Sensing and Earth Observation

Remote sensing (RS) and Earth observation (EO) data offer powerful approaches for assessment and modeling of forest biodiversity across spatial and temporal scales (Innes & Koch, 1998; Yu et al., 2022; Gillespie et al., 2024). The integration of RS into NFIs has evolved over several decades, beginning in the 1990s with efforts to map forest characteristics such as biomass and canopy structure (Tomppo & Katila, 1991; McRoberts & Tomppo, 2007). Advances in RS technology and continuous improvements to sensors in spatial, spectral, and temporal resolution have greatly enhanced the ability to detect forest composition, structure, and phenological change (Fassnacht et al., 2016; Fassnacht et al., 2024; Helfenstein et al., 2025). Yet, despite growing interest, biodiversity-relevant variables remain less explored than standard forest inventory metrics such as biomass or tree height (McRoberts & Tomppo, 2007; Lausch et al., 2016a; Skidmore et al., 2021; Kacic & Kuenzer, 2022). Multitemporal imagery improves the classification of tree species by capturing phenological variability, while wall-to-wall RS data support continuous landscape coverage, allowing for regionalization of forest cover or biodiversity indicators when coupled with ground observations (Gillespie et al., 2008; Feilhauer & Schmidtlein, 2009; Waser et al., 2015; Rocchini et al., 2016; Fassnacht et al., 2022). Here, the easy access to the analysis-ready data catalog of tools like Google Earth Engine facilitates automated time-series analyses, greatly enhancing efficiency and enabling faster classification of forest types, land cover change, and biodiversity patterns (Gorelick et al., 2017). RS supports estimations of ecosystem attributes over vast or inaccessible areas (Khare et al., 2019). While classification of larger spatial entities, i.e., forest and habitat types is feasible, individual tree species classification using remote-sensing remains challenging (Krahwinkler & Rossmann, 2013; Fassnacht et al., 2016).

In biodiversity monitoring, RS is especially useful as a source of ancillary variables that reflect environmental heterogeneity and has been used for habitat type and conservation status prediction (Neumann et al., 2015; Perrone et al., 2023), forest health assessment (Pause et al., 2016), mapping of forest degradation (Thompson et al., 2013) and geodiversity (Lausch et al., 2019), among others. For a deeper review of the synergies between RS and biodiversity research, see Pettorelli et al. (2014). There is ongoing research to explore the relationship between plant and tree species richness and spectral variability captured by RS (Viedma et al., 2012; Rocchini et al., 2021). The spectral variability hypothesis (Rocchini et al., 2018) posits that variability in the spectral signal can serve as a proxy for species richness and habitat diversity, a premise supported by empirical studies (Foody & Cutler, 2003; Féret & Asner, 2014; Torresani et al., 2021; Rocchini et al., 2022). Yet, while Torresani et al. (2019) report generally modest to positive results, they emphasize the need for further testing in different environments to confirm the approach as a generalizable method. A similar approach to monitor forest health was proposed by Lausch et al. (2016b). Key challenges remain in defining suitable indices, scaling, and optimizing the integration of RS with *in situ* observations for biodiversity monitoring (Rocchini et al., 2010; Skidmore et al., 2021; Torresani et al., 2021; Rocchini et al., 2022).

The synergy of forest inventory and mapping via RS can be based on several complementary approaches to enhance accuracy and utility of field data for management and monitoring. If thematic maps of specific areas of interest are available, i.e., areas of high conservation value, or areas with expected high variability of a specific target attribute, stratified sampling can be applied to increase estimate precision (Corona, 2010; Breidenbach et al., 2021b). RS sensor attributes can be correlated with metrics obtained by field-based observation, enabling the spatially explicit modeling of field attributes for larger areas (wall-to-wall mapping). While RS-derived diversity maps can illustrate spatial patterns, they often lack probabilistic validation, which limits their use in statistical inference and national reporting frameworks. To address these limitations, McRoberts (2011) emphasizes the importance of integrating RS data with field observations through model-based estimation frameworks. These approaches enable the derivation of statistically valid estimates accompanied by confidence intervals, thereby enhancing the reliability and transparency of included biodiversity assessments.

Increasingly available EO data derived from RS supports diversity assessments by providing wall-to-wall data on larger scales. A wide range of variables from different disciplines can thus be leveraged for site and habitat modeling, i.e., climatology, soil science, and digital elevation models (DEM, Lausch et al., 2019). These data are modeled from spectral or structural data (e.g., LiDAR, Synthetic Aperture Radar, SAR), increasingly using machine-learning frameworks. Time series aggregates capture dynamic changes e.g., in land-use and provide multitemporal information. Integrated analyses of regional to global time series of physical, index-based, thematic, topographic, and texture-derived variables hold great potential to assess diversity or habitat distribution or long-term shifts and losses (Kuenzer et al., 2014; Lausch et al., 2020). Continuous advances in temporal and spatial resolution, such as daily PlanetScope imagery and time-series analytics, will significantly boost predictive monitoring in scientific applications. These improvements promise to revolutionize predictive environmental modeling enabling near-real-time tracking of diversity shifts, habitat disturbances, and ecosystem responses, offering dynamic tools for conservation and ecosystem science (Ma et al., 2020).

Topography is considered a significant factor influencing the spatial distribution of communities (Amici et al., 2012; Zellweger et al., 2015; Guo et al., 2017; Amatulli et al., 2018; Zarnetske et al., 2019). Topography influences soil fertility and sunlight availability not only in forest ecosystems (Rodrigues et al., 2021; Woods & Ortmann, 2024). This is especially relevant in the predominantly mountainous country of Georgia, where 98% of the forest area is influenced by slope of various degrees according to Patarkalashvili, 2017. To capture indicators of fine-scale terrain features such as slope position, aspect, and terrain form, high-resolution elevation data is invaluable, because these indicators can be derived from DEM (Moudrý et al., 2019). In 2018, a country-wide LiDAR-based flight campaign was implemented in Georgia that produced detailed elevation information in a 5x5 m resolution. To analyze how topography influences species composition, several indices were derived from these raw data to characterize small-scale topographic variability (Chapter 3). For this, 1717 data files (tiles) were available containing point clouds with X, Y, Z coordinates in ASCII format. The data was converted to a DEM using LASTOOLS (Isenburg, 2014) and compared to the freely available SRTM30 DEM (Farr et al., 2007) interpolated to 10 m created by Fuchs et al. (2017) using systematic control points. Preprocessing and control of the derived raster data were carried out in the QGIS 3.20.2 environment (QGIS Development Team, 2009). After file format conversion, the resulting raster was void-filled with GRASS GIS 7.82, (GRASS Development Team, 2022). Figure 1.4 shows the resulting DEM.

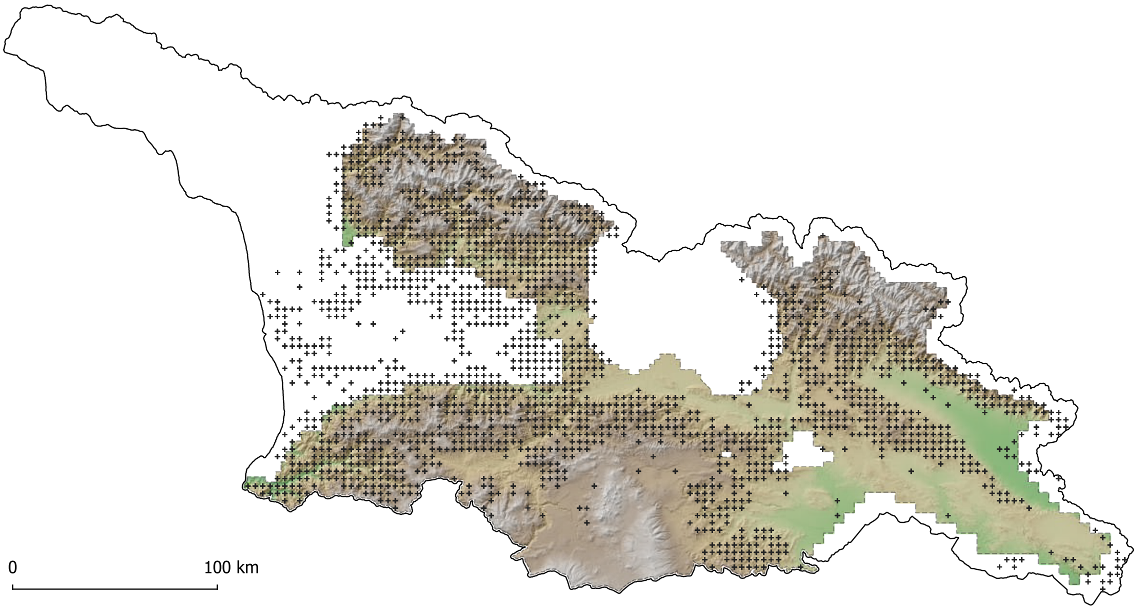
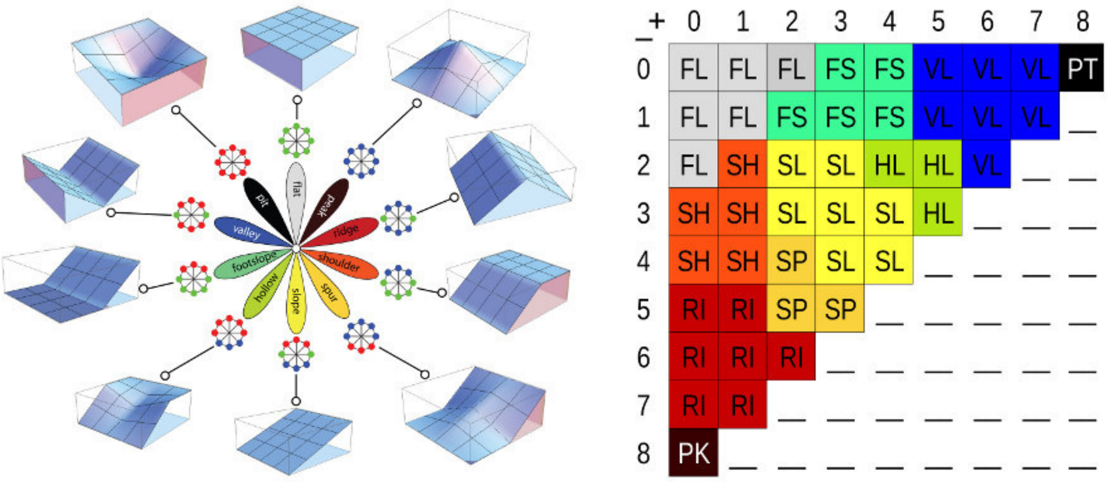


Figure 1.4: Digital elevation model derived from LiDAR raw data (2018) used in topographical analysis of sites locations in this study, incl. sample locations (m = 5870).

The created DEM covered approximately 58% of the country (56% of the national forest area). Subsequently, topographic properties i.e., elevation (m a.s.l.), slope (degrees), and aspect (0° – 360°) were derived for 3578 sample plot locations. This corresponds to 79% of all samples with available forest data. For samples located outside the high-resolution DEM, topographic variables were derived from the SRTM DEM (Figure 1.4). General terrain properties were analyzed using the Topographic Position Index, Terrain Ruggedness Index, and Roughness algorithms available in QGIS (QGIS Development Team, 2009). In addition, the R.GEOMORPHON algorithm available in GRASS (GRASS Development Team, 2022), and the RELATIVE HEIGHTS AND SLOPE POSITION algorithm provided by SAGA, both available in QGIS were used. The RELATIVE HEIGHTS AND SLOPE POSITION algorithm derives five terrain indices that quantify the vertical position of a point in relation to peaks and depressions within the local landscape, identifying features such as slope height, valley depth, and mid-slope position (Conrad et al., 2015). R.GEOMORPHON generates a size-, relief-, and orientation independent classification (Geomorphons, Figure 1.5) of terrain form based on adjacent pixel value differences (Stepinski & Jasiewicz, 2011).



+/−	0	1	2	3	4	5	6	7	8
0	FL	FL	FL	FS	FS	VL	VL	VL	PT
1	FL	FL	FS	FS	FS	VL	VL	VL	_
2	FL	SH	SL	SL	HL	HL	VL	_	_
3	SH	SH	SL	SL	SL	HL	_	_	_
4	SH	SH	SP	SL	SL	_	_	_	_
5	RI	RI	SP	SP	_	_	_	_	_
6	RI	RI	RI	_	_	_	_	_	_
7	RI	RI	_	_	_	_	_	_	_
8	PK	_	_	_	_	_	_	_	_

Figure 1.5: Terrain form classification using the R.GEOMORPHON algorithm available in GRASS of the QGIS environment (adjusted from Jasiewicz & Stepinski, 2013).

Geomorphons were derived from the high-resolution LiDAR-based DEM based on iterative testing of varying radii (15, 25, 30, and 40 m) around each plot center. Based on this test, 15 and 40 m were chosen as outer radius values, each in a combination of 0 and 5 m as inner radius and 0°and 5° as threshold angles (Jasiewicz & Stepinski, 2013).

In this study, RS-based elevation data were prepared and analyzed to characterize fine-scale topographic heterogeneity and support spatial biodiversity modeling (Chapter 2). Additional EO data were retrieved and applied as proxies for predictive modeling (Chapter 3). Given the complex terrain of Georgia’s forested landscapes, these inputs are particularly valuable for quantifying site properties along environmental gradients.

1.5 Study Area

Georgia is located between 41°07’ – 43°35’N and 40°04’ – 46°44’E bordering the Greater Caucasus to the North and the Northern part of the Lesser Caucasus to the South. Constituting part of the Caucasus biodiversity hotspot, the country harbors a rich diversity of habitats and a high degree of endemism (Myers et al., 2000; Mumladze et al., 2020). Since regaining independence in 1991, Georgia has undergone major political and economic transformations that have significantly influenced its natural environment and forestry sector (Gutman & Radeloff, 2017). The transition period was marked by political and institutional instability and unregulated resource use. In recent years, efforts to modernize forest governance, improve biodiversity monitoring, and implement sustainable forest management have gained momentum, supported in part by Georgia’s accession to the UN-CBD in 1994, the adoption of the new Forest Code by the Government of Georgia (2020), and participation in international frameworks such as the European Network of National Forest Inventories (ENFIN). This Chapter provides a brief overview of the country’s environment, its forests, and recent developments in forest governance, scientific research, and monitoring frameworks.

1.5.1 Geography and Climate

Georgia is characterized by dominant hilly to rugged mountainous terrain (with elevations ≥5000 m above sea level, a.s.l.) with roughly 55% of the national area situated above 1000 m a.s.l. and around 40% on slopes with ≥20° inclination (Mikeladze et al., 2020). The Greater Caucasus shields the country from Arctic air masses in winter, while the Southern Caucasus mitigates summer heat from the Southeast. Mountainous regions experience mean temperatures between −5 °C and 10 °C, with precipitation ranging from 800 to 1400 mm (Connor & Kvavadze, 2009). The Likhi Range (up to 1000 m a.s.l.) diverts moist air from the Black Sea, creating a humid, warm climate in Western Georgia and increasingly continental climate in Eastern Georgia. Whereas Western Georgia is characterized by mild winters and hot summers, with mean annual temperatures of 13 – 15 °C and precipitation between <400 and >4000 mm, Eastern Georgian climate is increasingly continental, with mean annual temperatures between 10 °C and 13 °C and lower mean precipitations of 500 – 600 mm (Denk et al., 2001; Elizbarashvili et al., 2006; Keggenhoff et al., 2014; 2017; Nakhutsrishvili et al., 2023). Forest vegetation in Georgia is structured along strong elevational and climatic gradients. The following section provides a generalized descriptive overview of these formations from lowland to subalpine zones.

1.5.2 Main Forest Formations

Forest type formation in Georgia is markedly shaped by elevational and climatic gradients, and variation in soil-forming parent materials (Dolukhanov, 2010). According to Lomsadze et al. (2019) 98% of forests are located on sloped terrain, with 78% situated on slopes ≥35°. A detailed description of the zonation of Georgia's forest vegetation exists by Dolukhanov (2010), which was published 10-years posthumously. A nation-wide map of the natural vegetation has been presented by Bohn et al. (2007). More recent phytosociological descriptions of forest typology published in English language are provided by (Nakhutsrishvili, 2013a; 2023). The number of publications that report regional botanical descriptions of forest vegetation types has increased steadily in the recent past (Abdaladze et al., 2017; Novák et al., 2020; Novák et al., 2021; Nakhutsrishvili et al., 2023; Kavgacı et al., 2023; Novák et al., 2023b). Noteworthy is the recent creation of the phytosociological Transcaucasian Vegetation Database that contains the results of Braun-Blanquet assessments of 2,882 georeferenced plots, spanning observations from 1929 to 2022 (Novák et al., 2023a). To the best of the author's knowledge, no comprehensive and internationally accessible compilation of forest type descriptions exists that focuses on stand structure and growth from a forest science perspective. This highlights the need for classification systems based on systematic data that go beyond phytosociological descriptions and reflect functional and structural forest properties relevant for ecological monitoring and forest management.

The distribution of forest types largely follows an elevational zonation from the humid Colchic lowlands and xeric woodlands to alpine coniferous stands. In the following paragraphs, I present a generalized overview of the major forest formations, focusing on woody species compositions occurring along the vertical and horizontal gradients, based on the available literature reviewed for this study.

Lowland forests (≤500 m a.s.l.)

Lowland forests exemplify the pronounced climatic variation present in the country, because at these elevations the steep climatic gradient is not yet moderated by the influence of elevation. As such, the western zone of the humid Colchic receives high to very high annual precipitation, whereas the eastern lowland valleys are characterized by a more continental, semi-arid climate. The western lowlands are occupied by Colchic forest formations occurring in the catchment basin of the Black Sea, stretching from the extreme northwestern end of the Greater Caucasus in the neighboring Russian Federation along the Black Seas coast to the Northern slopes of the Pontic Mountains in Türkiye (Nakhutsrishvili et al., 2015). According to Dolukhanov (2010), the Colchic lowland forests are humid forests on swampy or waterlogged soils dominated by *Alnus glutinosa* subsp. *barbata* (C.A. Mey.) Yalt. (Denk et al., 2001). Several variants of the humid thermophilic rainforests in the Colchis have developed (Nakhutsrishvili, 2013b; Nakhutsrishvili et al., 2023). On slightly elevated or better-drained sites, *Quercus robur* subsp. *imeretina* (Steven ex Woronow) Menitsky, *F. orientalis* and *Carpinus betulus* L. may appear in association (Goginashvili et al., 2021). The understory layer is typically dense and multilayered, often reaching heights of up to 4 m. It is composed predominantly of evergreen, shade-tolerant woody species, including *Rhododendron ponticum* L., *I. colchica*, and *Prunus laurocerasus* L. Climbing woody species such as *Smilax excelsa* L., *Hedera pastuchovii* Woronow ex Woronow, and *Vitis vinifera* L. are abundant and contribute to the structural complexity. Box et al. (2000) emphasizes the significance of vegetative reproduction among understory species, particularly *Rh. ponticum*, which propagates clonally through layering, forming dense thickets. The authors also highlight the decline of these forests due to drainage, agriculture, and infrastructure development. These forests are considered temperate rainforests due to their high annual precipitation, typically ranging between 1500 and ≥2,500 mm in the coastal lowlands of western Georgia (Nakhutsrishvili, 2013a).

As continentality increases eastward, lowland hardwood or oak-dominated forests receiving increasingly less precipitation occupy the riparian corridors and alluvial plains of central river systems, e.g., along the Alazani and Mtkvari rivers, dominated by *Quercus robur* subsp. *pedunculiflora* (K. Koch) Menitsky. These forests are co-dominated or interspersed with *Fraxinus excelsior* L:, *Acer heldreichii* subsp. *trautvetteri* (Medw.) A.E. Murray, *T. dasystyla* subsp. *caucasica*, and *Pyrus communis* subsp. *caucasica* (Fed.) Browicz, forming a loosely structured canopy. The riparian forests of eastern Georgia host *Pterocarya fraxinifolia* (Poir.) Spach (listed as *Pterocarya pterocarpa* (Michx.) Kunth ex I. Iljinsk in Lachashvili et al., 2022) as a typical species, which can form pure stands (Nakhutsrishvili, 2013a; Nakhutsrishvili et al., 2023; Davitashvili et al., 2024). Dolukhanov (2010) highlights the structural complexity and biogeographic significance of Oak-dominated lowland forests, noting their adaptation to moderately humid conditions, but also the drastic decline of these forests. Other authors stress their conservation value as fragmented relics, with floristic links to the Hyrcanian region (Box et al., 2000; Nakhutsrishvili, 2013a). The understory comprises drought-tolerant shrubs and climbers, including *Crataegus pentagyna* Waldst. & Kit. ex Willd., *Mespilus germanica* L., *Hedera pastuchovii* Woronow ex Woronow, *Clematis vitalba* L., and *Vitis vinifera* L., with *S. excelsa* L occasionally present.

Colline to Montane Forests (500 – 1,200 m a.s.l.)

In the lower foothills of western Georgia, mixed deciduous forests composed of *C. betulus* and *C. sativa* occur between 400 and 1,000 m a.s.l. on well-drained, moderately moist soils (Metreveli et al., 2024). These forests occupy transitional zones between Colchic lowland forests and montane beech stands, dominated by *F. orientalis* and often accompanied by *Quercus petraea* subsp. *polycarpa* (Schur) Soó or especially at higher elevations (Dolukhanov, 2010). *C. sativa* tends to dominate drier slopes, whereas *C. betulus* occupies moister microsites. The understory layer contains mesophilic shrubs such as *Cornus mas* L., *Euonymus latifolius* (L.) Mill., and *Frangula alnus* Mill., though composition varies depending on site moisture, with *Rhododendrum* species constituting common Colchic undergrowth (Nakhutsrishvili et al., 2023). Novák et al. (2021) provide a description of forests dominated or co-dominated by *Carpinus orientalis* Mill. with evergreen understory occurring within this zone on steep limestone slopes at elevations up to 800 m a.s.l.

Gallery forests, typically found along riverbanks in the eastern plains and foothills, are azonal, moisture-dependent communities. *Fraxinus excelsior* L. and *Ulmus minor* Mill. dominate the canopy, often accompanied by *Populus nigra* L., *Acer campestre* L., and *Salix alba* L. in disturbed sections (Kavtaradze et al., 2023). These forests show a tendency to form narrow, linear belts along seasonal or permanent watercourses due to their fragmentation as a result of to river regulation and land conversion (Nakhutsrishvili, 2013a). Floodplain margins and disturbed alluvial zones commonly host riverine strips of *Populus alba* L. and *S. alba*, especially in the eastern lowlands such as the Alazani and Iori valleys. These forests exhibit a pioneer character, forming open stands that rapidly colonize disturbed substrates, with pronounced successional dynamics and sensitivity to changes in the hydrological regime and nonnative species invasion (Kavtaradze et al., 2023). In younger successional stages, these tree species are accompanied by *Tamarix ramosissima* Ledeb. and various shrubs species, but tend to form simple, even-aged stands (Nakhutsrishvili et al., 2023). Mixed broadleaved forests containing *P. communis* subsp. *caucasica*, *Tilia dasystyla* subsp. *caucasica* (V. Engl.) Pigott, and various *Acer* species (e.g., *A. campestre*, *Acer platanoides L.*, *Acer monspessulanum* subsp. *ibericum* (M. Bieb.) Yalt.) form structurally diverse communities in mid to lower elevation (foothill) zones, particularly in the central and western parts of the country. These forests occupy moderately moist to slightly dry locations and represent transitions between drier oak or hornbeam stands and more humid beech or chestnut forests (Nakhutsrishvili, 2013a). The arid easternmost regions of Georgia host xerophytic woodlands dominated by *Juniperus excelsa* subsp. *polycarpos* (K. Koch) Takht. and *Pistacia atlantica* Desf. (Pistachio woodland predominantly occupying the south-eastern part of Iori plateau, Lachashvili et al., 2020) at elevations from approximately 250 to 500 m a.s.l. These open woodlands occur on rocky slopes and shallow soils and are often accompanied by *Rhamnus erythroxyloides* subsp. *erythroxyloides* (listed as *Rhamnus pallasii* Fisch. & C.A. Mey. in Fischer et al., 2018) *Celtis glabrata* Steven ex Planch., and scattered individuals of *Ziziphus jujuba* Mill. These woodlands represent a biogeographic link to Irano-Turanian flora and display exceptional drought tolerance (Lachashvili et al., 2020; Nakhutsrishvili et al., 2023).

Mid-Montane to Subalpine (1,200–1,800 m a.s.l.)

The elevational zone 1,200–1,800 m a.s.l. marks the transition from the mixed montane deciduous forests of lower elevations to the cooler, more humid montane belt, where forest composition becomes increasingly dominated by *F. orientalis* (Nakhutsrishvili, 2013a). It represents the most extensive and ecologically diverse forest zone, especially in the western and central parts of the country (Box et al., 2000). At lower altitudes, *F. orientalis* is associated with broadleaved species like *Q. petraea* subsp. *polycarpa*, *T. dasystyla* subsp. *caucasica* and *Acer* spp. (Urushadze et al.).

In the more humid western parts of Georgia oriental beech forest are characterized by lush Colchic evergreen species that dominate the dense understory i.e., *Rh ponticum*, *Ilex colchica* Pojark., *Viburnum orientale* Pall., *Prunus laurocerasus* L. or *Daphne glomerata* Lam., leading to highly structured forests. In contrast, in eastern Georgia, the understory of *F. orientalis* forests tends to be less dense, with a higher presence of deciduous shrub species, i.e., *Corylus avellana* L., *Euonymus europaeus* L., *Rubus* spp., *C. mas*, *Crataegus monogyna* Jacq., etc. (Nakhutsrishvili, 2013a). As a result of the humidity gradient, the distribution of these forests along the elevational strata is lower and associated species are *Q. petraea* subsp. *polycarpa*, *C. betulus*, *A. campestre* and *T. dasystyla* subsp. *caucasica* in the drier eastern part of the country *Q. petraea* subsp. *polycarpa* may form pure stands below the main distribution range of *F. orientalis*. As humidity increases along the elevational gradient, dominance of *F. orientalis* increases, leading to monospecific stands at higher altitudes (Dolukhanov, 2010; Nakhutsrishvili, 2013a). *Q. petraea* subsp. *polycarpa* and *C. betulus* forests in this elevational strata (1,200 and 1,800 m a.s.l.) are considered transitional communities most developed in the central and eastern parts of the country (Dolukhanov, 2010). These forests are either dominated by *Quercus spp.* or *C. betulus* and *C. orientalis*, depending on local site conditions, with *C. orientalis* typically occupying drier, lower elevations or south-facing slopes. In eastern Georgia, *Q. petraea subsp. iberica* is the dominant species, often forming nearly pure stands or co-dominating with *Carpinus* spp., particularly in lower elevations. Toward more mesic conditions and higher elevations, *C. betulus* increases in importance, sometimes transitioning into *F. orientalis* forests. The understory in these forests differs regionally: in the drier eastern zones, it is typically sparse and composed of drought-tolerant shrubs such as *Berberis vulgaris* L., *Rosa canina* L., and *C. monogyna*. In contrast, in western and more humid locations, the understory layer is denser, frequently dominated by species like *C. mas*, *E. europaeus*, and *C. avellana*.

Above 1,200 m a.s.l. *F. orientalis* occurs mixed with *Abies nordmanniana* (Steven) Spach. and *Picea orientalis* (L.) Peterm. These mixed broadleaved–coniferous forests occupy an intermediate elevational belt and form an important transition zone between the lower montane broadleaved forests (i.e., *F. orientalis* and *Q. petraea* subsp. *polycarpa* dominated types) and subalpine coniferous forests and occur on deep, fertile soils in high precipitation zones, typically between 1,200–1,600 m a.s.l. (Dolukhanov, 2010). *A. nordmanniana* may dominate in sheltered valleys or north-facing slopes, while *F. orientalis* dominates on slightly drier exposures. *P. orientalis* becomes more dominant at the upper margin of this elevational zone, especially on very moist sites, where it occurs mixed with *F. orientalis* and *A. nordmanniana*, and regionally, *Pinus sylvestris* var. *hamata* Steven (Nakhutsrishvili, 2013a; Akhalkatsi et al., 2019). In the Lesser Caucasus, *P. sylvestris* var. *hamata* and *A. nordmanniana* may be more prevalent than *P. orientalis*, reflecting differences in soil and climate. In the western mountains, *P. orientalis* is more dominant and tends to form denser stands, with *F. orientalis* and *A. nordmanniana* as consistent companions.

In eastern Georgia, however, due to increased continentality*,* *F. orientalis* dominates mixed stands, while conifers are often restricted to cooler, moister topographical conditions.

The understory in these forests is marked by high variation in species composition. In western Georgia, the lush, evergreen undergrowth consisting of *Rh. ponticum*, *I. colchica*, and *P. laurocerasus* is equally frequent, whereas in drier areas, especially at lower altitudes, *Lonicera caucasica* Pall and *Rubus* spp. are frequent throughout the zone. In eastern Georgia, the understory is generally less dense and composed of species such as *B. vulgaris*, *L. caucasica*, and *Crataegus* spp. At this elevation zone of the Lesser and eastern Greater Caucasus *Quercus macranthera* Fisch. & C.A. Mey. ex Hohen. may accompany or locally replace *F. orientalis* or *C. betulus* forests on higher, drier slopes (Akobia et al., 2022). In disturbed sites or ecotonal patches, mixed stands occur where *A. nordmanniana, P. orientalis,* and *F. orientalis* intermix with elements such as *Hedlundia armeniaca* (Hedl.) Mezhenskyj and *Acer heldreichii* subsp. *trautvetteri* (Medw.) A.E. Murray*,* contributing to transitional forest types (Box et al., 2000; Dolukhanov, 2010; Akhalkatsi & Tarkhnishvili, 2012; Nakhutsrishvili, 2013a).

Above 1,700 m a.s.l. predominance of conifers increases with *A. nordmanniana* and *P. orientalis* forests forming a distinct vegetation zone at higher montane to subalpine elevations, typically ranging up to 2,100 m a.s.l., with regional variations resulting from exposure, moisture, and slope orientation. Species composition and structure vary between dominant species and between western and eastern Georgia. *P. orientalis* dominated forests are frequently associated with *F. orientalis*. Because both *P. orientalis* and *A. nordmanniana* require relatively high moisture levels, they rarely reach dominance on the drier, more continental slopes in the east of the country (Nakhutsrishvili, 2013a). As a result, their distribution is narrower and increasingly fragmented in the east. Both species sometimes dominate forests on shaded, north-facing slopes or in ravines, where microclimatic conditions favor moisture retention. In western Georgia however, where precipitation is high *P. orientalis* and *A. nordmanniana* frequently co-dominate, forming dense, tall forests with relatively closed canopies. Hence, the origin of the vernacular name "dark coniferous forests" under which these forests are subsumed (Nakhutsrishvili et al., 2021). *P. orientalis* and *A. nordmanniana* dominated forests are found across both the Greater and Lesser Caucasus and *often* develop over acidic soils. On northern or moist slopes, *P. orientalis* and *A. nordmanniana* may dominate almost exclusively, whereas they co-occur with *F. orientalis* or *Betula pubescens* var. *litwinowii* (Doluch.) Ashburner & McAll. on drier or south-facing slopes. The understory of these forests is dense and species-rich in the humid western Greater Caucasus, while it becomes markedly sparser and less diverse under the more continental conditions in the east (Dolukhanov, 2010).

Subalpine (1,800–2,200 m a.s.l.)

Forests of the Upper Montane to Subalpine Zone (1,800–2,200 m a.s.l.) represent a transition between closed montane forests and subalpine meadows along the tree line and are characterized by the increasingly alpine climatic conditions (Nakhutsrishvili, 2013a; Tephnadze et al., 2014). This zone marks the ecological boundary where climatic conditions begin to significantly restrict the distribution of thermophilic species, favoring cold-tolerant conifers and shrubby taxa. Forest types vary across Georgia's west–east precipitation gradient, though the majority are dominated by conifers and cold-adapted broadleaved species, i.e., *B. pubescens* var. *litwinowii*, *Betula medwediewii* Regel, *Betula megrelica* Sosn. and *Betula raddeana* Tratv. and dwarfed growth forms of *F. orientalis* (Akhalkatsi & Tarkhnishvili, 2012; Akobia et al., 2022).

The provided overview existing forest typologies in Georgia is based on qualitative descriptions grounded in phytocoenological methods and nomenclature of vegetation science. Potential distributions of forest types has been mapped by Bohn et al. (2007) on a coarse national scale. All reviewed sources point to the significant human impact on forest communities, typically associated with a high level of degradation.

1.5.3 Forest Degradation and Management

Unsustainable use continues to shape many forests, particularly on erosion-prone slopes and in proximity to settlements (Togonidze, 2015; Patarkalashvili, 2016; Stritih et al., 2024). Regeneration is often constrained by livestock grazing within forest areas, and reforestation efforts have declined markedly since the economic and financial crisis in the country of the 1990s, resulting in persistent degradation risks in specific regions (Olofsson et al., 2010; Patarkalashvili, 2016; Akobia et al., 2022). In addition, forest inventory and data acquisition in the context of forest management planning were significantly interrupted during the same period and uncontrolled timber extraction resulted in structural simplification, low growth rates and reduced timber qualities (Machavariani, 2010). Exacerbating the ongoing degradation, forest and conservation management face structural and institutional challenges, i.e. the absence of modern theoretical frameworks and guidance, empirically grounded classification systems, and shortage of forest science professionals with multidisciplinary backgrounds. For example, comprehensive supraregional approaches to develop systematic forest classifications aligned with modern international typological frameworks and phytogeographical principles are still largely lacking, particularly for forest ecosystems (Goginashvili et al., 2021). Such classifications would enable the integration of structural components and facilitate the quantification of site- and community-specific growth rates derived from reference areas representing intended conditions. Although systematic descriptions that include quantitative structural estimates based on stratified data analysis are still missing, the required forest inventory data is increasingly available (Nakhutsrishvili et al., 2023). These data hold great potential for the systematic investigation of forest structure, successional trajectories, and the ecological impacts of disturbance and degradation. With the successful implementation of the first GNFI and the legal framework solidified by mandate in the New Forest Code (Government of Georgia, 2020) a foundation for permanent, statistically robust data acquisition has been established. The GNFI marks a turning point in forest monitoring in the country by providing the first nationally standardized dataset capable of delivering statistically robust estimates of forest resources, supporting ecologically grounded classification, and enabling productivity modeling, sustainability assessments, and long-term biodiversity monitoring. The first generalized results indicate that Georgia’s forest area is larger than previously assumed (cf. 38.2% Patarkalashvili, 2016), likely reflecting reforestation trends following the discontinuation of crop cultivation and range farming (Hansen et al., 2018; Buchner et al., 2020; Akobia et al., 2022; Stritih et al., 2024).

Furthermore, with the acquisition of standardized FMI data at forest district level, the range of covered forest types is increased. FMIs are usually implemented independently from NFIs and can follow different assessment protocols. This data can be leveraged to create information to reduce uncertainties about forest conditions and address knowledge gaps that currently constrain sustainable forest management (Chalatashvili et al., 2024). Targeted analyses of these data are essential to disentangle forest type-specific structural properties, related growth and regeneration dynamics, and alterations caused by disturbance and degradation. From this, updated and ecologically grounded forest productivity models can be derived which are urgently needed to replace outdated tools, e.g., out-of-date volume tables, essential to forest management in Georgia (Machavariani, pers. comm.). To ensure operational relevance for forest management and conservation, these must be based on up-to-date and regionally adjusted reference values that reflect structural and compositional realities. These values should be stratified by ecological classification systems to provide accurate and sustainable guidelines.

1.5.4 Forest Monitoring and Research

The wealth of ecologically meaningful variables now captured, i.e., vertical structure, shrub and understory composition, and deadwood, offers a solid foundation for advancing management-relevant forest and biodiversity research. The scientific opportunity lies in developing integrative approaches that combine dendrochronology, disturbance ecology, successional modeling, and forest dynamics to strengthen ecological forest science. Already in the 1980s, Dolukhanov (2010) emphasized the need for a standardized forest monitoring system, criticizing the descriptive and inconsistent nature of the 10-year forest management planning cycles. Advocating a multidisciplinary approach, Dolukhanov called for forest type-specific research and to strengthen the integration of ecology and biology into forestry education. With the onset of recurring acquisition of systematically assessed forest data, a wide range of ecological and forestry specific variables have now been assessed on various scales. Georgian forest scientists are increasingly making use of these datasets to address contemporary questions in ecology and forest science (Mikeladze et al., 2020; Akobia et al., 2022; Metreveli et al., 2023; Chalatashvili et al., 2024; Metreveli et al., 2024).

In addition, two marteloscopes of 1 ha, each have been recently established in two distinct forest areas in central and west Georgia (Kruse, 2022). Marteloscopes are permanent sample plots in which all trees are systematically marked and periodically remeasured (Pommerening, 2023); [Cannot display reference "Kruse 2023 – Identifying training needs": Template "In-text citation - Thesis - (Default template)" is not defined.]. They generate high-resolution data on community composition, structure, and habitat features, providing a robust foundation for modeling successional trajectories, and assessing biodiversity and habitat shifts (Kadavý et al., 2024). Designed to train and evaluate tree selection in continuous cover forestry, marteloscopes offer a practical framework to balance ecological reasoning (e.g., habitat value, diversity) and ecosystem services (e.g., provision of timber) to support the development of holistic, sustainability-oriented perspectives in multifunctional forest management and education (Joa et al., 2020; John et al., 2024). As such, marteloscopes hold promise for advancing integrative biodiversity assessment and supporting analyses of forest community assembly and phylogenetic change.

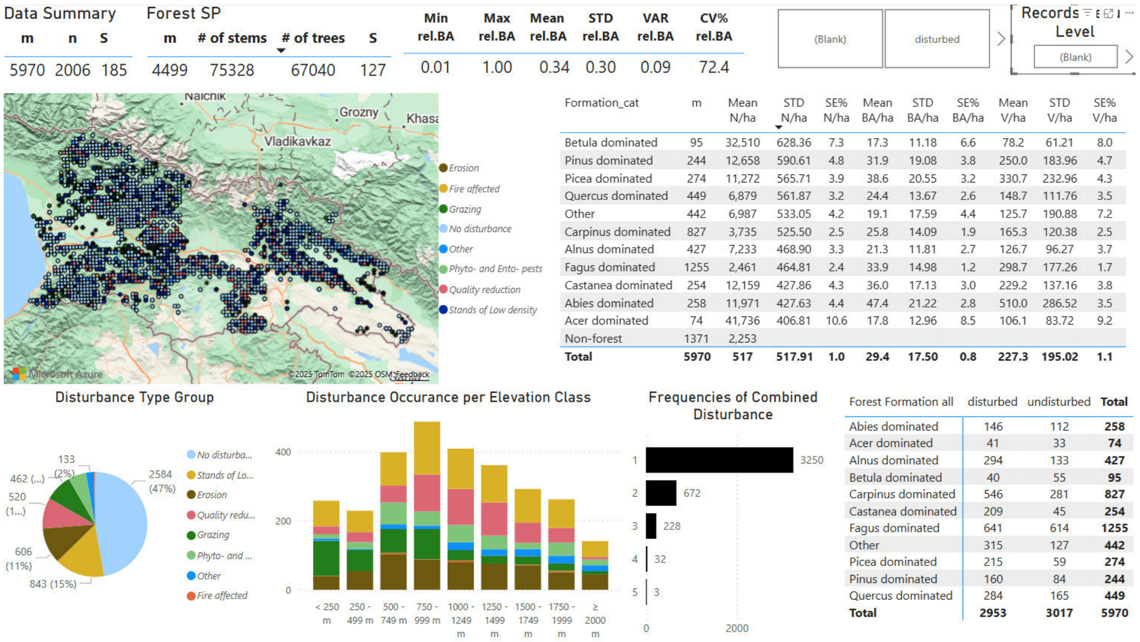
In summary, the wealth of partially permanent observational studies now available in the country provides a solid base of robust data that enables systematic and targeted research to answer pressing ecological questions. Utilizing the available data requires focused processing and analysis with modern data techniques. The rapid development and increasing availability of modeling approaches that rely on large, multivariate datasets make this task particularly challenging. However, multiple opportunities exist to support required advancements in forest science through network collaborations, i.e., ForestGEO (Davies et al., 2021) or ENFIN (Fridman et al., 2014; Vidal et al., 2016a). Similar networks (i.e., European Vegetation Archive, Chytrý et al., 2016) exist for vegetation science, e.g., the TRY vegetation trait database, Kattge et al., 2020.

1.6 Data Analysis

This section outlines the structure, processing environment, and selection criteria used for the analysis of the GNFI data in this study. The dataset comprises 5,970 recorded sample plots organized into 2,006 cluster samples. A total of 67,578 trees were measured, and shrub species were recorded on 4,233 plots. The dataset also includes a wide range of spatially explicit variables relevant for classifying forest habitats. To manage and explore this large and heterogeneous dataset, a solid processing and analysis environment was established. Data files (.xlsx format) from the GNFI were imported into a relational data model constructed in in Microsoft (MS) Power BI. Figure 1.6 illustrates the relational structure used to organize plot-level, species, and environmental attribute tables in MS Power BI.

Figure 1.6: Relational data model used in Power BI, showing table connections and keys within the multivariate dataset (screenshot from the structural analysis described in Chapter 4).

MS Power BI provides a robust platform for integrating, exploring, and visualizing large and complex datasets, supporting efficient early-phase data analysis and interpretation (Becker & Gould, 2019). In this study, the software was used for data processing, aggregation, exploratory analyses, and visualization. The integration with R, combined with straightforward export of aggregated tables as .csv files, allowed seamless transfer of data into RStudio for in-depth analysis. Built-in visualization tools enabled rapid inspection of preliminary patterns prior to conducting detailed statistical analyses in RStudio (RStudio Team, 2024). Efficient exploratory analysis is possible by swift visualization of spatial distributions, as Power BI allows to map geo-referenced data sets (Figure 1.7).



Data Summary

m	n	S
5970	2006	185

Forest SP

m	# of stems	# of trees	S
4499	75328	67040	127

Min rel.BA	Max rel.BA	Mean rel.BA	STD rel.BA	VAR rel.BA	CV% rel.BA
0.01	1.00	0.34	0.30	0.09	72.4

Formation_cat	m	Mean N/ha	STD N/ha	SE% N/ha	Mean BA/ha	STD BA/ha	SE% BA/ha	Mean V/ha	STD V/ha	SE% V/ha
Betula dominated	95	32,510	628.36	7.3	17.3	11.18	6.6	78.2	61.21	8.0
Pinus dominated	244	12,658	590.61	4.8	31.9	19.08	3.8	250.0	183.96	4.7
Picea dominated	274	11,272	565.71	3.9	38.6	20.55	3.2	330.7	232.96	4.3
Quercus dominated	449	6,879	561.87	3.2	24.4	13.67	2.6	148.7	111.76	3.5
Other	442	6,987	533.05	4.2	19.1	17.59	4.4	125.7	190.88	7.2
Carpinus dominated	827	3,735	525.50	2.5	25.8	14.09	1.9	165.3	120.38	2.5
Alnus dominated	427	7,233	468.90	3.3	21.3	11.81	2.7	126.7	96.27	3.7
Fagus dominated	1255	2,461	464.81	2.4	33.9	14.98	1.2	298.7	177.26	1.7
Castanea dominated	254	12,159	427.86	4.3	36.0	17.13	3.0	229.2	137.16	3.8
Abies dominated	258	11,971	427.63	4.4	47.4	21.22	2.8	510.0	286.52	3.5
Acer dominated	74	41,736	406.81	10.6	17.8	12.96	8.5	106.1	83.72	9.2
Non-forest	1371	2,253								
Total	**5970**	**517**	**517.91**	**1.0**	**29.4**	**17.50**	**0.8**	**227.3**	**195.02**	**1.1**

Forest Formation all	disturbed	undisturbed	**Total**
Abies dominated	146	112	**258**
Acer dominated	41	33	**74**
Alnus dominated	294	133	**427**
Betula dominated	40	55	**95**
Carpinus dominated	546	281	**827**
Castanea dominated	209	45	**254**
Fagus dominated	641	614	**1255**
Other	315	127	**442**
Picea dominated	215	59	**274**
Pinus dominated	160	84	**244**
Quercus dominated	284	165	**449**
Total	**2953**	**3017**	**5970**

Figure 1.7: Report view in Microsoft Power BI showing various visualization tools, including summary tables of aggregated values, map visuals, interactive filters, pie, and bar charts (screenshot from the exploratory workspace used for structural analysis in Chapter 4).

Calculations based on interspecific phylogenetic distances require species to be reliably identified at the species level. Before conducting phylogenetic analyses, species identification quality across layers was assessed to determine data suitability. In this dataset, complete species-level identification across all three vegetation layers (regeneration, shrub, and tree) was available for only 39% of sample plots. To ensure consistent data quality and reduce uncertainty, the analysis was restricted to woody species observations in the tree layer, for which 76% of plots had species-level identifications.

Whereas MS Power BI was used for initial data organization, filtering, extrapolation, and compilation, all complex analyses were performed in an R working environment using R Studio (R Core Team, 2024; RStudio Team, 2024). Throughout this thesis, the R working environment served as the central platform for computationally intensive analyses, including the calculation of general diversity metrics during the exploratory analysis, PD metrics, ISOPAM cluster analyses, statistical validations, and predictive modeling using RANDOMFOREST classifications (Breiman, 2001). Given the large multivariate dataset used in this research, dividing the individual work steps among different systems helped maintain clarity, ensured data integrity, and enabled efficient data processing and flexible visualization.

1.7 Research Objectives and Approaches

This Chapter outlines the core research objectives and methodological approach of the thesis. The study aims to develop and evaluate the use of phylogenetic information in combination with systematic forest inventory data to improve biodiversity assessment in large-scale forest monitoring. By classifying forest inventory data from the GNFI as a case study, this research operationalizes PD to group compositional data into ecologically meaningful woody species assemblages. To advance our understanding of the information embedded in phylogenetic variability of forest communities, the study empirically explores the conceptual and methodological potential of this approach in relation to environmental gradients, forest structure, and degradation dynamics.

By incorporating interspecies phylogenetic distances into clustering, the analysis moves beyond traditional diversity measures to capture an evolutionary component of community structure. To my knowledge, this approach has not yet been empirically applied to a large-scale forest inventory dataset. The evaluation focuses on the additional insights gained from phylogenetically informed beta-diversity metrics, compared to species-neutral diversity indices, and how these metrics influence related clustering outcomes. The overarching goal is to assess whether extended diversity measures are suitable for classifying forest communities when applied to NFI data by addressing three research objectives.

The first objective is to evaluate whether phylogenetic information improves the classification of woody species compositional data derived from forest inventory plot observations (Chapter 2). This is addressed by comparing unsupervised clustering results based on a species-neutral dissimilarity index (Bray–Curtis) with those based on a phylogenetically informed metric (dA). The resulting group characteristics and spatial patterns are analyzed using 1059 cluster samples. The following research questions guide this analysis:

- Does the inclusion of interspecies phylogenetic distances in dissimilarity calculations improve the discrimination of species assemblages compared to species-neutral indices?
- To what extent do assemblages identified using phylogenetic dissimilarity better reflect underlying biogeographic gradients than those identified using conventional dissimilarity?
- Can the inclusion of phylogenetic information in dissimilarity measures enhance the ecological interpretability of clustering results?

Building on the phylogenetically informed classification, the second objective evaluates the ecological and spatial coherence of the identified assemblages. Specifically, the second objective aims to assess whether phylogenetically informed assemblages correspond to environmental and biogeographic gradients (Chapter 3). To this end, 3,466 single plot samples were clustered using the same dissimilarity metric (dA). Cluster membership was modeled using a multivariate environmental dataset derived from EO and RS data, which was structured into four thematic groups: climatic, topographic, soil, and spatial variables. To account for compositional shifts under disturbance, NFI samples were grouped into subsets based on disturbance level, and each subset was clustered separately. Random Forest classification models were trained to predict cluster membership of samples, using principal components of each thematic group as independent variables. This modeling approach was used to test whether the observed distribution of species assemblages aligns with biogeographic gradients, leading to the following research questions:

- Can the spatial distribution of woody species assemblages be reliably modeled using site-specific environmental variables?
- Does increasing disturbance affect the compositional characteristics of species assemblages and reduce prediction accuracy in classification models?
- Does the relative importance of environmental predictors differ for modeling woody species assemblage membership across sample groups with varying disturbance levels?

With the first two objectives focusing on classification and modeling, the third objective includes a more practice-oriented perspective by linking forest structural diversity to forest degradation. To investigate how structural variation among woody species assemblages is influenced by degradation, the same cluster analysis was applied. By merging clustered samples with unclustered, monodominant samples into manually defined groups of similar BA distributions, "forest formations" consisting of 4,168 samples were created (Chapter 2). Structural indicators are then compared across these forest formations along a field-assigned degradation severity gradient to evaluate response and to examine technical considerations in assessing forest degradation within the GNFI framework. The following research questions are addressed:

- Does the phylogenetically informed clustering approach reveal structural patterns that reflect the field-assigned degradation severity levels, and can these patterns help distinguish between disturbance and degradation?
- Do stand properties respond differently to degradation severity across forest formations, and do observed changes along the gradient indicate a transition from disturbance to degradation?

This thesis focuses on the methodological integration of biodiversity indicator variables into large-scale forest inventories. It offers an empirical framework for incorporating PD into forest classification and biodiversity monitoring. In addition to contributing insights to empirical biodiversity assessment, it also enhances our understanding of community assembly dynamics by showing how disturbance affects both the structural and phylogenetic composition of forest communities, for example through the presence of nonnative species. In summary, this study introduces an ecophylogenetic perspective to forest inventory analysis to assess and monitor forest-related biodiversity.

1.8 References

Abdaladze, O., Batsatsashvili, K., Körner, C., Nakhutsrishvili, G. and Spehn, E. (eds). 2017 *Plant Diversity in the Central Great Caucasus A Quantitative Assessment.* 1st edn. Springer International Publishing, Cham. 9783319557779, 170 p.

Abe, S. 2021 Comparison of traditional and automated approaches in classification of Japanese coastal beach and dune vegetation. *Society of Vegetation Science* **38**, 67–80. doi.org/10.15031/vegsci.38.67.

Adamowicz, S.J. 2015 International Barcode of Life: Evolution of a global research community. *Genome* **58**, 151–162. 10.1139/gen-2015-0094.

Akhalkatsi, M., Arabuli, G., Asanidze, Z., Goloshvili, T. and Togonidze, N. 2019 Pine Forest on Tree-Line Ecotone in the Mountain Kazbegi in the Georgia (South Caucasus). *Agricultural Research & Technology* **21**. 10.19080/ARTOAJ.2019.21.556149.

Akhalkatsi, M. and Tarkhnishvili, D. 2012 *Habitats of Georgia.* Natura 2000, 119 p. researchgate.net/publication/233151704.

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

Alberdi, I., Cañellas, I. and Condes, S. 2014 A long-scale biodiversity monitoring methodology for Spanish national forest inventory. Application to Álava region. *For. syst.* **23**, 93–110. 10.5424/fs/2014231-04238.

Alberdi, I., Condés, S. and Martínez-Millán, J. 2010 Review of monitoring and assessing ground vegetation biodiversity in national forest inventories. *Environmental monitoring and assessment* **164**, 649–676. 10.1007/s10661-009-0919-4.

Alberdi, I., Nunes, L., Kovac, M., Bonheme, I., Cañellas, I. and Rego, F.C. et al. 2019 The conservation status assessment of Natura 2000 forest habitats in Europe: capabilities, potentials and challenges of national forest inventories data. *Annals of Forest Science* **76**, 1–15. 10.1007/s13595-019-0820-4.

Alberdi, I., Vallejo, R., Álvarez-González, J.G., Condés, S., González-Ferreiro, E. and Guerrero, S. et al. 2017 The multi-objective Spanish National Forest Inventory. *Forest Syst* **26**, e04S. 10.5424/fs/2017262-10577.

Alekseev, A., Tomppo, E., McRoberts, R.E. and Gadow, K. von. 2019 A constructive review of the State Forest Inventory in the Russian Federation. *For. Ecosyst.* **6**. 10.1186/s40663-019-0165-3.

Ali, S., Amin, A., Akhtar, M.S. and Zaman, W. 2025 Phylogenetic Diversity in Forests: Insights into Evolutionary Patterns and Conservation Strategies. *Forests* **16**, 1004. 10.3390/f16061004.

Álvarez-González, J.G., Cañellas, I., Alberdi, I., Gadow, K. von and Ruiz-González, A.D. 2014 National Forest Inventory and forest observational studies in Spain: Applications to forest modeling. *Forest Ecology and Management* **316**, 54–64. 10.1016/j.foreco.2013.09.007.

Alzate, A. and Hagen, O. 2024 Dispersal-diversity feedbacks and their consequences for macroecological patterns. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* **379**, 20230131. 10.1098/rstb.2023.0131.

Amatulli, G., Domisch, S., Tuanmu, M.-N., Parmentier, B., Ranipeta, A., Malczyk, J. and Jetz, W. 2018 A suite of global, cross-scale topographic variables for environmental and biodiversity modeling. *Scientific data* **5**, 180040. 10.1038/sdata.2018.40.

Amici, V., Rocchini, D., Geri, F., Bacaro, G., Marcantonio, M. and Chiarucci, A. 2012 Effects of an afforestation process on plant species richness: A retrogressive analysis. *Ecological Complexity* **9**, 55–62. 10.1016/j.ecocom.2011.11.006.

Anderson, C.B. 2018 Biodiversity monitoring, earth observations and the ecology of scale. *Ecology letters* **21**, 1572–1585. 10.1111/ele.13106.

Antão, L.H., McGill, B., Magurran, A.E., Soares, A.M.V.M. and Dornelas, M. 2019 β-diversity scaling patterns are consistent across metrics and taxa. *Ecography* **42**, 1012–1023. 10.1111/ecog.04117.

APG I. 1998 An Ordinal Classification for the Families of Flowering Plants. *Annals of the Missouri Botanical Garden,* **85**, 531–553. 10.2307/2992015.

APG IV. 2016 An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG IV. *Botanical Journal of the Linnean Society* **181**, 1–20. 10.1111/boj.12385.

Applequist, W.L. 2015 A Brief Review of Recent Controversies in the Taxonomy and Nomenclature of Sambucus nigra sensu lato. *Acta horticulturae* **1061**, 25–33. 10.17660/ActaHortic.2015.1061.1.

Asbeck, T., Großmann, J., Paillet, Y., Winiger, N. and Bauhus, J. 2021 The Use of Tree-Related Microhabitats as Forest Biodiversity Indicators and to Guide Integrated Forest Management. *Curr Forestry Rep* **7**, 59–68. 10.1007/s40725-020-00132-5.

Bacaro, G., Ricotta, C. and Mazzoleni, S. 2007 Measuring Beta-Diversity from Taxonomic Similarity. *Journal of Vegetation Science* **18**, 793–798. academia.edu/download/48919490/Measuring_beta-diversity_from_taxonomic_20160917-3963-1pnlew5.pdf.

Balkenhol, N., Dudaniec, R.Y., Krutovsky, K.V., Johnson, J.S., Cairns, D.M. and Segelbacher, G. et al. 2019 Landscape Genomics: Understanding Relationships Between Environmental Heterogeneity and Genomic Characteristics of Populations. In *Population Genomics: Concepts, Approaches and Applications*. O.P. Rajora (ed). Springer International Publishing. 978-3-030-04589-0, pp. 261–322.

Barnosky, A.D., Hadly, E.A., Bascompte, J., Berlow, E.L., Brown, J.H. and Fortelius, M. et al. 2012 Approaching a state shift in Earth's biosphere. *Nature* **486**, 52–58. 10.1038/nature11018.

Baselga, A. and Gómez-Rodríguez, C. 2019 Diversidad alfa, beta y gamma: ¿cómo medimos diferencias entre comunidades biológicas? [in Spanish]. *Nova Acta Científica Compostelana (Bioloxía)* **26**, 39–45.

Becker, L.T. and Gould, E.M. 2019 Microsoft Power BI: Extending Excel to Manipulate, Analyze, and Visualize Diverse Data. *Serials Review* **45**, 184–188. 10.1080/00987913.2019.1644891.

Beckschäfer, P., Mundhenk, P., Kleinn, C., Ji, Y., Yu, D.W. and Harrison, R.D. 2013 Enhanced Structural Complexity Index: An Improved Index for Describing Forest Structural Complexity. *OJF* **03**, 23–29. 10.4236/ojf.2013.31005.

Bello, F. de, Botta-Dukát, Z., Lepš, J. and Fibich, P. 2021 Towards a more balanced combination of multiple traits when computing functional differences between species. *Methods Ecol Evol* **12**, 443–448. 10.1111/2041-210X.13537.

Betts, M.G., Yang, Z., Hadley, A.S., Smith, A.C., Rousseau, J.S. and Northrup, J.M. et al. 2022 Forest degradation drives widespread avian habitat and population declines. *Nature ecology & evolution* **6**, 709–719. 10.1038/s41559-022-01737-8.

Bevilacqua, S., Anderson, M.J., Ugland, K.I., Somerfield, P.J. and Terlizzi, A. 2021 The use of taxonomic relationships among species in applied ecological research: Baseline, steps forward and future challenges. *Austral Ecology* **46**, 950–964. 10.1111/aec.13061.

Bohn, U., Zazanashhvili, N. and Nakhutsrishvili, G. 2007 The Map of the Natural Vegetation of Europe and its application in the Caucasus Ecoregion. *Bulletin of the Georgian National Academy of Science* **175**, 112–119. science.org.ge/old/moambe/2007-vol1/112-120.pdf.

Bonari, G., Fernández-González, F., Çoban, S., Monteiro-Henriques, T., Bergmeier, E. and Didukh, Y.P. et al. 2021 Classification of the Mediterranean lowland to submontane pine forest vegetation. *Applied Vegetation Science* **24**. 10.1111/avsc.12544.

Botta-Dukát, Z. 2005 Rao's quadratic entropy as a measure of functional diversity based on multiple traits. *Journal of Vegetation Science* **16**, 533–540. 10.1111/j.1654-1103.2005.tb02393.x.

Box, E., Kazue, Fujiwara, Nakhutsrishvili, G., Zazanashvili, N., Liebermann, R. and Miyawaki, A. 2000 Vegetation and Landscapes of Georgia (Caucasus), as a Basis for Landscape Restoration. *Bull. Inst. Environ. Sci. Technol. Natn. Univ.*, 69–102. ynu.repo.nii.ac.jp/record/5442/files/2008kei2000_01.pdf.

Bragg, J.G., Supple, M.A., Andrew, R.L. and Borevitz, J.O. 2015 Genomic variation across landscapes: insights and applications. *The New phytologist* **207**, 953–967. 10.1111/nph.13410.

Braun-Blanquet, J. 1968 L'école phytosociologique zuricho-montpelliéraine et la S.I.G.M.A. *Vegetatio* **16**, 1–78. 10.1007/BF00261357. link.springer.com/article/10.1007/BF00261357.

Bray, J.R. and Curtis, J.T. 1957 An Ordination of the Upland Forest Communities of Southern Wisconsin. *Ecological Monographs* **27**, 325–349. 10.2307/1942268. doi.org/10.2307/1942268.

Breidenbach, J., McRoberts, R.E., Alberdi, I., Antón-Fernández, C. and Tomppo, E. 2021a A century of national forest inventories – informing past, present and future decisions. *Forest ecosystems* **8**, 1–4. 10.1186/s40663-021-00315-x.

Breidenbach, J., Waser, L.T., Debella-Gilo, M., Schumacher, J., Rahlf, J. and Hauglin, M. et al. 2021b National mapping and estimation of forest area by dominant tree species using Sentinel-2 data. *Can. J. For. Res.* **51**, 365–379. 10.1139/cjfr-2020-0170.

Breiman, L. 2001 Random Forests. *Machine Learning* **45**, 5–32. 10.1023/A:1010933404324.

Brockerhoff, E.G., Barbaro, L., Castagneyrol, B., Forrester, D.I., Gardiner, B. and González-Olabarria, J.R. et al. 2017 Forest biodiversity, ecosystem functioning and the provision of ecosystem services. *Biodivers Conserv* **26**, 3005–3035. 10.1007/s10531-017-1453-2.

Buchner, J., Yin, H., Frantz, D., Kuemmerle, T., Askerov, E. and Bakuradze, T. et al. 2020 Land-cover change in the Caucasus Mountains since 1987 based on the topographic correction of multi-temporal Landsat composites. *Remote Sensing of Environment* **248**, 111967. 10.1016/j.rse.2020.111967.

Butchart, S.H.M., Walpole, M., Collen, B., van Strien, A., Scharlemann, J.P.W. and Almond, R.E.A. et al. 2010 Global biodiversity: indicators of recent declines. *Science (New York, N.Y.)* **328**, 1164–1168. 10.1126/science.1187512.

Cabido, M., Zeballos, S.R., Zak, M., Carranza, M.L., Giorgis, M.A., Cantero, J.J. and Acosta, A.T.R. 2018 Native woody vegetation in central Argentina: Classification of Chaco and Espinal forests. *Applied Vegetation Science* **21**, 298–311. 10.1111/avsc.12369.

Cáceres, M. de, Chytrý, M., Agrillo, E., Attorre, F., Botta-Dukát, Z. and Capelo, J. et al. 2015 A comparative framework for broad-scale plot-based vegetation classification. *Applied Vegetation Science* **18**, 543–560. 10.1111/avsc.12179.

Cáceres, M. de, Martín-Alcón, S., González-Olabarria, J.R. and Coll, L. 2019 A general method for the classification of forest stands using species composition and vertical and horizontal structure. *Annals of Forest Science* **76**, 1–19. 10.1007/s13595-019-0824-0.

Cadotte, M.W. 2017 Functional traits explain ecosystem function through opposing mechanisms. *Ecology letters* **20**, 989–996. 10.1111/ele.12796.

Cadotte, M.W. 2023 The interacting influences of competition, composition and diversity determine successional community change. *Journal of Ecology* **111**, 1670–1680. 10.1111/1365-2745.14135.

Cadotte, M.W., Albert, C.H. and Walker, S.C. 2013 The ecology of differences: assessing community assembly with trait and evolutionary distances. *Ecology letters* **16**, 1234–1244. 10.1111/ele.12161.

Cadotte, M.W., Cardinale, B.J. and Oakley, T.H. 2008 Evolutionary history and the effect of biodiversity on plant productivity. *PNAS* **105**, 17012–17017. 10.1073/pnas.0805962105.

Calonje, M., Martín-Bravo, S., Dobeš, C., Gong, W., Jordon-Thaden, I. and Kiefer, C. et al. 2009 Non-coding nuclear DNA markers in phylogenetic reconstruction. *Plant Syst Evol* **282**, 257–280. 10.1007/s00606-008-0031-1.

Cantero, J.J., Zeballos, S.R., Núñez, C.O., Sfragulla, J., Acosta, A.T.R. and Funes, G. et al. 2022 Classification of rocky outcrops plant communities in the mountains of Central Argentina. *Applied Vegetation Science* **25**. 10.1111/avsc.12686.

Capblancq, T. and Forester, B.R. 2021 Redundancy analysis: A Swiss Army Knife for landscape genomics. *Methods Ecol Evol* **12**, 2298–2309. 10.1111/2041-210X.13722.

Capblancq, T., Sękiewicz, K. and Dering, M. 2024 Forest genomics in the Caucasus through the lens of its dominant tree species - Fagus orientalis. *Molecular ecology* **33**, e17475. 10.1111/mec.17475.

Cardillo, M. 2023 Phylogenetic diversity in conservation: A brief history, critical overview, and challenges to progress. *Cambridge prisms. Extinction* **1**, e11. 10.1017/ext.2023.8.

Cavender-Bares, J., Keen, A. and Miles, B. 2006 Phylogenetic structure of Floridian plant communities depends on taxonomic and spatial scale. *Ecology* **87**, 109-122. 10.1890/0012-9658(2006)87[109:PSOFPC]2.0.CO;2.

Cavender-Bares, J., Kothari, S. and Pearse, W. 2018 Evolutionary Ecology of Communities. In *Evolutionary Biology*. J. Cavender-Bares, S. Kothari and W. Pearse (eds). Oxford University Press. 9780199941728, pp. 1–30.

Cavender-Bares, J. and Wilczek, A. 2003 Integrating Micro- and Macroevolutionary Processes in Community Ecology. *Ecological Society of America* **84**, 592–597. jstor.org/stable/3107852.

Černý, T., Kopecký, M., Petřík, P., Song, J.-S., Šrůtek, M. and Valachovič, M. et al. 2015 Classification of Korean forests: patterns along geographic and environmental gradients. *Applied Vegetation Science* **18**, 5–22. 10.1111/avsc.12124.

Chalatashvili, A., Mikeladze, G., Gobronidze, N., Metreveli, V., Marsagishvili, N. and Drossler, L. 2024 Knowledge gaps to manage Caucasian forests: What do Georgian forest scientists and foresters want to know and how to answer such questions? *Annals of Agrarian Science* **22**. journals.org.ge/index.php/aans/article/view/393.

Chao, A., Chiu, C.-H. and Jost, L. 2010 Phylogenetic diversity measures based on Hill numbers. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* **365**, 3599–3609. 10.1098/rstb.2010.0272.

Chao, A., Chiu, C.-H. and Jost, L. 2014 Unifying Species Diversity, Phylogenetic Diversity, Functional Diversity, and Related Similarity and Differentiation Measures Through Hill Numbers. *Annu. Rev. Ecol. Evol. Syst.* **45**, 297–324. 10.1146/annurev-ecolsys-120213-091540.

Chao, A., Chiu, C.-H. and Jost, L. 2016 Phylogenetic Diversity Measures and Their Decomposition: A Framework Based on Hill Numbers. In *Biodiversity Conservation and Phylogenetic Systematics: Preserving our evolutionary heritage in an extinction crisis*. R. Pellens and P. Grandcolas (eds). Springer International Publishing. 978-3-319-22461-9, pp. 141–172.

Chase, M.T.V., Soltis, D.E., Olmstead, R.C., Morgan, D., Les, D.J. and Mishler, B.D. et al. 1993 Phylogenetics of Seed Plants: An Analysis of Nucleotide Sequences from the Plastid Gene rbcL. *Annals of the Missouri Botanical Garden* **80**, 528-548+550-580. scholarworks.boisestate.edu/cgi/viewcontent.cgi?article=1045&context=bio_facpubs.

Chave, J., Chust, G. and Thébaud, C. The importance of phylogenetic structure in biodiversity studies. Chapter Eight. In *Scaling Biodiversity*, pp. 151–167.

Chazdon, R.L., Brancalion, P.H.S., Laestadius, L., Bennett-Curry, A., Buckingham, K. and Kumar, C. et al. 2016 When is a forest a forest? Forest concepts and definitions in the era of forest and landscape restoration. *Ambio* **45**, 538–550. 10.1007/s13280-016-0772-y.

Chirici, G., McRoberts, R.E., Winter, S., Bertini, R., Brändli, U. and Asensio, I.A. et al. 2012 National Forest Inventory Contributions to Forest Biodiversity Monitoring. *Forest Science* **58**, 257–268. 10.5849/forsci.12-003.

Chirici, G., Winter, S. and McRoberts, R.E. 2011 *National Forest Inventories: Contributions to Forest Biodiversity Assessments.* Springer Netherlands, Dordrecht. 978-94-007-0481-7, 220 p.

Chiu, C.-H., Jost, L. and Chao, A. 2014 Phylogenetic beta diversity, similarity, and differentiation measures based on Hill numbers. *Ecological Monographs* **84**, 21–44. doi.org/10.1890/12-0960.1.

Chytrý, M., Hennekens, S.M., Jiménez-Alfaro, B., Knollová, I., Dengler and Jansen, F. et al. 2016 European Vegetation Archive (EVA): an integrated database of European vegetation plots. *Applied Vegetation Science* **19**, 173–180. 10.1111/avsc.12191.

Chytrý, M., Řezníčková, M., Novotný, P., Holubová, D., Preislerová, Z. and Attorre, F. et al. 2024 FloraVeg.EU — An online database of European vegetation, habitats and flora. *Applied Vegetation Science* **27**. 10.1111/avsc.12798.

Chytrý, M., Tichý, L., Hennekens, S.M., Knollová, I., Janssen, J.A.M. and Rodwell, J.S. et al. 2020 EUNIS Habitat Classification: Expert system, characteristic species combinations and distribution maps of European habitats. *Applied Vegetation Science* **23**, 648–675. 10.1111/avsc.12519.

Clark, J. and Evans, C. 1954 Distance to Nearest Neighbor as a Measure of Spatial Relationships in Populations. *Ecology* **35**, 445–453. jstor.org/stable/1931034.

Clarke, K.R., Somerfield, P.J., Airoldi, L. and Warwick, R.M. 2006 Exploring interactions by second-stage community analyses. *Journal of Experimental Marine Biology and Ecology* **338**, 179–192. 10.1016/j.jembe.2006.06.019.

Clarke, K.R. and Warwick, R.M. 1998 A taxonomic distinctness index and its statistical properties. *Journal of Applied Ecology* **35**, 523–531. 10.1046/j.1365-2664.1998.3540523.x.

Clarke, K.R. and Warwick, R.M. 1999 The taxonomic distinctiveness measure of biodiversity: weighting of step length between hierarchical levels. *Mar. Ecol. Prog. Ser.*, 21–29. int-res.com/articles/meps/184/m184p021.pdf.

Clarke, K.R. and Warwick, R.M. 2001 A further biodiversity index applicable to species lists: variation in taxonomic distinctness. *Mar. Ecol. Prog. Ser.* **216**, 265–278. 10.3354/meps216265.

Connor, S.E. and Kvavadze, E.V. 2009 Modelling late Quaternary changes in plant distribution, vegetation and climate using pollen data from Georgia, Caucasus. *Journal of Biogeography* **36**, 529–545. 10.1111/j.1365-2699.2008.02019.x.

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

Corona, P. 2010 Integration of forest mapping and inventory to support forest management. *iForest* **3**, 59–64. 10.3832/ifor0531-003.

Corona, P., Chirici, G., McRoberts, R.E., Winter, S. and Barbati, A. 2011 Contribution of large-scale forest inventories to biodiversity assessment and monitoring. *Forest Ecology and Management* **262**, 2061–2069. 10.1016/j.foreco.2011.08.044.

Costanza, J.K., Faber-Langendoen, D., Coulston, J.W. and Wear, D.N. 2018 Classifying forest inventory data into species-based forest community types at broad extents: exploring tradeoffs among supervised and unsupervised approaches. *Forest ecosystems* **5**. 10.1186/s40663-017-0123-x.

Cutko, A. 2009 *Biodiversity Inventory of Natural Lands: A How-To Manual for Foresters and Biologists: A How-To Manual for Foresters and Biologists.* NatureServe., Arlington, Virginia. www.natureserve.org, 40 p.

Da Silva, L.P., Heleno, R.H., Costa, J.M., Valente, M., Mata, V.A. and Gonçalves, S.C. et al. 2019 Natural woodlands hold more diverse, abundant, and unique biota than novel anthropogenic forests: a multi-group assessment. *Eur J Forest Res* **138**, 461–472. 10.1007/s10342-019-01183-5.

Davies, S.J., Abiem, I., Abu Salim, K., Aguilar, S., Allen, D. and Alonso, A. et al. 2021 ForestGEO: Understanding forest diversity and dynamics through a global observatory network. *Biological Conservation* **253**, 108907. 10.1016/j.biocon.2020.108907.

Davies, T.J. 2021 Ecophylogenetics redux. *Ecology letters* **24**, 1073–1088. 10.1111/ele.13682.

Davies, T.J. and Buckley, L.B. 2011 Phylogenetic diversity as a window into the evolutionary and biogeographic histories of present-day richness gradients for mammals. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* **366**, 2414–2425. 10.1098/rstb.2011.0058.

Davies, T.J., Urban, M.C., Rayfield, B., Cadotte, M.W. and Peres-Neto, P.R. 2016 Deconstructing the relationships between phylogenetic diversity and ecology: a case study on ecosystem functioning. *Ecology* **97**, 2212–2222. 10.1002/ecy.1507.

Davitashvili, N., Bétrisey, S., Janiashvili, Z., Matchutadze, I. and Kozlowski, G. 2024 First description and conservation implications of a unique stand of the Caucasian wingnut in Lapankuri (Georgia). *Feddes Repertorium* **135**, 78–84. 10.1002/fedr.202300031.

Denk, T., Frotzler, N. and Davitashvili, N. 2001 Vegetational patterns and distribution of relict taxa in humid temperate forests and wetlands of Georgia (Transcaucasia). *Biol J Linn Soc Lond*, 287–332. 10.1006/bijl.2000.0502.

Di Biase, L., Tsafack, N., Pace, L. and Fattorini, S. 2023 Ellenberg Indicator Values Disclose Complex Environmental Filtering Processes in Plant Communities along an Elevational Gradient. *Biology* **12**. 10.3390/biology12020161.

Dolukhanov, A.G. 2010 *Лесная растительность Грузии: (Forest vegetation of Georgia) [in Russian].* Universal, Tbilisi, Georgia. 978-9941-17-176-5. bilib.blogspot.com/2011/12/blog-post_24.html, 484 p.

Doyle, J.A. 1998 Phylogeny of vascular plants. *Annual review of ecology and systematics* **29**, 567–599. 10.1146/annurev.ecolsys.29.1.567.

Eberle, J., Ahrens, D., Mayer, C., Niehuis, O. and Misof, B. 2020 A Plea for Standardized Nuclear Markers in Metazoan DNA Taxonomy. *Trends in Ecology & Evolution* **35**, 336–345. 10.1016/j.tree.2019.12.003.

Elizbarashvili, E.S., Chavchanidze, Z.B., Elizbarashvili, M.E., Maglakelidze, R.V., Sulkhanishvili, N.G. and Elizbarashvili, S.E. 2006 Soil-climatic zoning of Georgia. *Eurasian Soil Sc.* **39**, 1062–1065. 10.1134/S1064229306100036.

Elizbarashvili, M., Elizbarashvili, E., Tatishvili, M., Elizbarashvili, S., Meskhia, R. and Kutaladze, N. et al. 2017 Georgian climate change under global warming conditions. *Annals of Agrarian Science* **15**, 17–25. 10.1016/j.aasci.2017.02.001.

Ellenberg, H. 1974 *Zeigerwerte der Gefasspflanzen Mitteleuropas: [in German].* Verlag Erich Goltze KG, Gottingen.

Ellenberg, H. and Leuschner, C. 2010 *Vegetation Mitteleuropas mit den Alpen: in ökologischer, dynamischer und historischer Sicht.* Utb. 978-3-8252-8104-5.

Epps, C.W. and Keyghobadi, N. 2015 Landscape genetics in a changing world: disentangling historical and contemporary influences and inferring change. *Molecular ecology* **24**, 6021–6040. 10.1111/mec.13454.

Estoque, R.C., Dasgupta, R., Winkler, K., Avitabile, V., Johnson, B.A. and Myint, S.W. et al. 2022 Spatiotemporal pattern of global forest change over the past 60 years and the forest transition theory. *Environ. Res. Lett.* **17**, 84022. 10.1088/1748-9326/ac7df5.

European Commission. 2021 *EU Biodiversity Strategy for 2030: Bringing nature back into our lives.* Publications Office of the European Union, Luxembourg. 978-92-76-36472-6, 36 p.

Faber-Langendoen, D., Keeler-Wolf, T., Del Meidinger, Tart, D., Hoagland, B. and Josse, C. et al. 2014 EcoVeg: a new approach to vegetation description and classification. *Ecological Monographs* **84**, 533–561. 10.1890/13-2334.1.

Faith, D.P. 1992 Conservation evaluation and phylogenetic diversity.

FAO. 2018 *Global Forest Resources Assessment 2020: Terms and Definitions*. FRA 2020 Working Paper 186. Food and Agriculture Organization of the United Nations, Rome, 32 p. fao.org.

FAO. 2020 *Global Forest Resources Assessment 2020: Key findings.* FAO, Rome. 978-92-5-132581-0. far.org, 16 p.

Farr, T.G., Rosen, P.A., Caro, E., Crippen, R., Duren, R. and Hensley, S. et al. 2007 The Shuttle Radar Topography Mission. *Reviews of Geophysics* **45**, RG2004. 10.1029/2005RG000183.

Fassnacht, F.E., Latifi, H., Stereńczak, K., Modzelewska, A., Lefsky, M. and Waser, L.T. et al. 2016 Review of studies on tree species classification from remotely sensed data. *Remote Sensing of Environment* **186**, 64–87. 10.1016/j.rse.2016.08.013.

Fassnacht, F.E., Müllerová, J., Conti, L., Malavasi, M. and Schmidtlein, S. 2022 About the link between biodiversity and spectral variation. *Applied Vegetation Science* **25**. 10.1111/avsc.12643.

Fassnacht, F.E., White, J.C., Wulder, M.A. and Næsset, E. 2024 Remote sensing in forestry: current challenges, considerations and directions. *Forestry: An International Journal of Forest Research* **97**, 11–37. 10.1093/forestry/cpad024.

Fazan, L., Song, Y.-G. and Kozlowski, G. 2020 The Woody Planet: From Past Triumph to Manmade Decline. *Plants (Basel, Switzerland)* **9**. 10.3390/plants9111593.

Feilhauer, H. and Schmidtlein, S. 2009 Mapping continuous fields of forest alpha and beta diversity. *Applied Vegetation Science* **12**, 429–439. 10.1111/j.1654-109X.2009.01037.x.

Feilhauer, H., Zlinszky, A., Kania, A., Foody, G.M., Doktor, D., Lausch, A. and Schmidtlein, S. 2021 Let your maps be fuzzy!—Class probabilities and floristic gradients as alternatives to crisp mapping for remote sensing of vegetation. *Remote Sens Ecol Conserv* **7**, 292–305. 10.1002/rse2.188.

Felsenstein, J. 1981 Evolutionary trees from DNA sequences: A maximum likelihood approach. *Journal of molecular evolution* **17**, 368–376. gen-ftp.princeton.edu/sysbio/Felsenstein81.pdf.

Féret, J.-B. and Asner, G.P. 2014 Mapping tropical forest canopy diversity using high-fidelity imaging spectroscopy. *Ecological Applications* **24**, 1289–1296. 10.1890/13-1824.1.

Fine, P.V.A. and Kembel, S.W. 2011 Phylogenetic community structure and phylogenetic turnover across space and edaphic gradients in western Amazonian tree communities. *Ecography* **34**, 552–565. 10.1111/j.1600-0587.2010.06548.x.

Fischer, E., Gröger, A. and Lobin, W. 2018 *Illustrated field guide to the flora of Georgia (South Caucasus).* 1st edn. University of Koblenz-Landau, Koblenz. 978-3-9820257-0-4, 830 p.

Flynn, D.F.B., Mirotchnick, N., Jain, M., Palmer, M.I. and Naeem, S. 2011 Functional and phylogenetic diversity as predictors of biodiversity--ecosystem-function relationships. *Ecology* **92**, 1573–1581. 10.1890/10-1245.1.

Foody, G.M. and Cutler, M.E.J. 2003 Tree biodiversity in protected and logged Bornean tropical rain forests and its measurement by satellite remote sensing. *Journal of Biogeography* **30**, 1053–1066. 10.1046/j.1365-2699.2003.00887.x.

Forest Europe. 2020 *State of Europe's Forests 2020.* Ministerial Conference on the Protection of Forests in Europe, Bratislava, 394 p. www.foresteurope.org.

Forzieri, G., Dakos, V., McDowell, N.G., Ramdane, A. and Cescatti, A. 2022 Emerging signals of declining forest resilience under climate change. *Nature* **608**, 534–539. 10.1038/s41586-022-04959-9.

Franklin, S. 2015 How a national vegetation classification can help ecological research and management. Letter. *Frontiers in Ecology and the Environment* **13**, 185–186. 10.1890/15.WB.006.

Fridman, J., Holm, S., Nilsson, M., Nilsson, P., Ringvall, A. and Ståhl, G. 2014 Adapting National Forest Inventories to changing requirements – the case of the Swedish National Forest Inventory at the turn of the 20th century. *Silva Fenn.* **48**. 10.14214/sf.1095.

Fuchs, H., Kleinn, C. and Fehrmann, L. 2017 *Integrating Remote Sensing, National Forest Inventory (NFI) and Forest Management Inventory (FMI) in Georgia.* ForestEye Research GmbH & Co KG, unpublished, 15 p.

Gadow, K. von, Álvarez González, J.G., Zhang, C., Pukkala, T. and Zhao, X. 2021 *Sustaining forest ecosystems.* Springer, Cham, Switzerland. 9783030587130, 419 p.

Gadow, K. von, Zhang, C.Y., Wehenkel, C., Pommerening, A., Corral-Rivas, J. and Korol, M. et al. 2012 Forest Structure and Diversity. In *Continuous Cover Forestry.* T. Pukkala and K. von Gadow (eds). Springer Netherlands. 978-94-007-2201-9, pp. 29–83.

Gaggiotti, O.E., Chao, A., Peres-Neto, P., Chiu, C.-H., Edwards, C. and Fortin, M.-J. et al. 2018 Diversity from genes to ecosystems: A unifying framework to study variation across biological metrics and scales. *Evolutionary applications* **11**, 1176–1193. 10.1111/eva.12593.

Ganeshaiah, K.N., Chandrashekara, K. and Kumar, A. R. V. 1997 Avalanche index: A new measure of biodiversity based on biological heterogeneity of the communities. *Current Science* **73**, 128–133. jstor.org/stable/24098266.

Ganeshaiah, K.N. and Shaankar, R.U. 2000 Measuring biological heterogeneity of forest vegetation types: avalanche index as an estimate of biological diversity. *Biodiversity and Conservation*, 953–963. doi.org/10.1023/A:1008910918751.

Gao, T., Hedblom, M., Emilsson, T. and Nielsen, A.B. 2014 The role of forest stand structure as biodiversity indicator. *Forest Ecology and Management* **330**, 82–93. 10.1016/j.foreco.2014.07.007.

Gao, T., Nielsen, A.B. and Hedblom, M. 2015 Reviewing the strength of evidence of biodiversity indicators for forest ecosystems in Europe. *Ecological Indicators* **57**, 420–434. 10.1016/j.ecolind.2015.05.028.

Gaüzère, P., O'Connor, L., Botella, C., Poggiato, G., Münkemüller, T. and Pollock, L.J. et al. 2022 The diversity of biotic interactions complements functional and phylogenetic facets of biodiversity. *Current Biology* **32**, 2093-2100.e3. 10.1016/j.cub.2022.03.009.

Geiger, M.F., Astrin, J.J., Borsch, T., Burkhardt, U., Grobe, P. and Hand, R. et al. 2016 How to tackle the molecular species inventory for an industrialized nation-lessons from the first phase of the German Barcode of Life initiative GBOL (2012-2015). *Genome* **59**, 661–670. 10.1139/gen-2015-0185.

Gillerot, L., Grussu, G., Condor-Golec, R., Tavani, R., Dargush, P. and Attorre, F. 2021 Progress on incorporating biodiversity monitoring in REDD+ through national forest inventories. *Global Ecology and Conservation* **32**, e01901. 10.1016/j.gecco.2021.e01901.

Gillespie, T.W., Foody, G.M., Rocchini, D., Giorgi, A.P. and Saatchi, S. 2008 Measuring and modelling biodiversity from space. *Progress in Physical Geography: Earth and Environment* **32**, 203–221. 10.1177/0309133308093606.

Gillespie, T.W., Rogers, M., Robinson, C. and Rocchini, D. 2024 Biodiversity of the World. In *Remote Sensing Handbook, Volume IV.* P.S. Thenkabail (ed). CRC Press. 9781003541172, pp. 191–210.

Gillet, F. and Julve, P. 2018 The integrated synusial approach to vegetation classification and analysis. *phyto* **48**, 141–152. 10.1127/phyto/2017/0164.

Gobet, A., Quince, C. and Ramette, A. 2010 Multivariate Cutoff Level Analysis (MultiCoLA) of large community data sets. *Nucleic acids research* **38**, e155. 10.1093/nar/gkq545.

Godoy, O., Kraft, N.J.B. and Levine, J.M. 2014 Phylogenetic relatedness and the determinants of competitive outcomes. *Ecology letters* **17**, 836–844. 10.1111/ele.12289.

Godoy, O. and 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**, 67–79. 10.7818/ECOS.2016.25-3.09.

Goginashvili, N., Togonidze, N., Tvauri, I., Manvelidze, Z., Memiadze, N., Zerbe, S. and Asanidze, Z. 2021 Diversity and degradation of the vegetation of mountain belt forests of central Adjara (the Lesser Caucasus), Georgia. *J. For. Sci.* **67**, 219–241. 10.17221/80/2020-JFS.

Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D. and Moore, R. 2017 Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote Sensing of Environment* **202**, 18–27. 10.1016/j.rse.2017.06.031.

Gotelli, N.J. and Colwell, R.K. 2001 Quantifying biodiversity: procedures and pitfalls in the measurement and comparison of species richness. *Ecology letters* **4**, 379-391. 10.1046/j.1461-0248.2001.00230.x.

Government of Georgia. 2020 *Forest Code of Georgia*. leap.unep.org/en/countries/ge/national-legislation/forest-code-georgia-no-5949-ss-2020, 60 p.

Graham, C.H. and Fine, P.V.A. 2008 Phylogenetic beta diversity: linking ecological and evolutionary processes across space in time. *Ecology letters* **11**, 1265–1277. 10.1111/j.1461-0248.2008.01256.x.

GRASS Development Team. 2022 *Geographic Resources Analysis Support System (GRASS GIS) Software, Version 8.2.* Open Source Geospatial Foundation. grass.osgeo.org/.

Gregorius, H.-R. and Kosman, E. 2017 On the notion of dispersion: from dispersion to diversity. *Methods Ecol Evol* **8**, 278–287. 10.1111/2041-210X.12665.

Griesbach, R. 2018 *Support in planning and implementation of national forest inventory of Georgian forests: Mission Report,* unpublished, 26 p.

Guillory, W.X. and Brown, J.L. 2021 A New Method for Integrating Ecological Niche Modeling with Phylogenetics to Estimate Ancestral Distributions. *Systematic biology* **70**, 1033–1045. 10.1093/sysbio/syab016.

Guo, W.-Y., Serra-Diaz, J.M., Schrodt, F., Eiserhardt, W.L., Maitner, B.S. and Merow, C. et al. 2020 Half of the world's tree biodiversity is unprotected and is increasingly threatened by human activities. 10.1101/2020.04.21.052464.

Guo, Y., Wang, B., Li, D., Mallik, A.U., Xiang, W. and Ding, T. et al. 2017 Effects of topography and spatial processes on structuring tree species composition in a diverse heterogeneous tropical karst seasonal rainforest. *Flora* **231**, 21–28. 10.1016/j.flora.2017.04.002.

Gutman, G. and Radeloff, V. (eds). 2017 *Land-Cover and Land-Use Changes in Eastern Europe after the Collapse of the Soviet Union in 1991.* 1st edn. Springer International Publishing; Imprint: Springer, Cham. 9783319426389, 1 online resource (VIII, 247 pages 84 illustrations, 65 illustrations in color.).

Hagen, O., Skeels, A., Onstein, R.E., Jetz, W. and Pellissier, L. 2021 Earth history events shaped the evolution of uneven biodiversity across tropical moist forests. *Proceedings of the National Academy of Sciences of the United States of America* **118**. 10.1073/pnas.2026347118.

Hansen, W., Magiera, A., Theissen, T., Waldhardt, R. and Otte, A. 2018 Analysing Betula litwinowii encroachment and reforestation in the Kazbegi region, Greater Caucasus, Georgia. *J Vegetation Science* **29**, 110–123. 10.1111/jvs.12589.

Hao, M., Corral-Rivas, J., González-Elizondo, M.S., Ganeshaiah., K.N., Nava-Miranda, M.G. and Zhang, C. et al. 2019a Assessing biological dissimilarities between five forest communities. *Forest ecosystems* **6**. 10.1186/s40663-019-0188-9.

Hao, M., Gadow, K. von, Alavi, S.J., Álvarez-González, J.G., Baluarte-Vásquez, J.R. and Corral-Rivas, J. et al. 2021 A classification of woody communities based on biological dissimilarity. *Applied Vegetation Science* **24**. 10.1111/avsc.12565.

Hao, M., Ganeshaiah, K.N., Zhang, C., Zhao, X. and Gadow, K. von. 2019b Discriminating among forest communities based on taxonomic, phylogenetic and trait distances. *Forest Ecology and Management* **440**, 40–47. 10.1016/j.foreco.2019.03.006.

Hao, M., Zhang, C., Zhao, X. and Gadow, K. von. 2018 Functional and phylogenetic diversity determine woody productivity in a temperate forest. *Ecology and evolution* **8**, 2395–2406. 10.1002/ece3.3857.

Hebert, P.D.N., Cywinska, A., Ball, S.L. and deWaard, J.R. 2003 Biological identifications through DNA barcodes. *Proceedings. Biological sciences* **270**, 313–321. 10.1098/rspb.2002.2218.

Hein, N., Feilhauer, H., Finch, O.-D., Schmidtlein, S. and Löffler, J. 2014 Snow cover determines the ecology and biogeography of spiders (Araneae) in alpine tundra ecosystems. *Erdkunde* **68**, 157–172. 10.3112/erdkunde.2014.03.01.

Helfenstein, I.S., Sturm, J.T., Schmid, B., Damm, A., Schuman, M.C. and Morsdorf, F. 2025 Satellite Observations Reveal a Positive Relationship Between Trait-Based Diversity and Drought Response in Temperate Forests. *Global change biology* **31**, e70059. 10.1111/gcb.70059.

Heydari, M., Omidipour, R. and Greenlee, J. 2020 Biodiversity, a review of the concept, measurement, opportunities, and challenges. *Journal of Wildlife and Biodiversity* **4**. 10.22120/jwb.2020.123209.1124.

Heym, M., Uhl, E., Moshammer, R., Dieler, J., Stimm, K. and Pretzsch, H. 2021 Utilising forest inventory data for biodiversity assessment. *Ecological Indicators* **121**, 107196. 10.1016/j.ecolind.2020.107196.

Hill, M.O. 1979 TWINSPAN - a FORTRAN program for arranging multivariate data in an ordered two way table by classification of individuals and attributes. *Ecology and Systematics*, 48pp.

Hilmers, T., Friess, N., Bässler, C., Heurich, M., Brandl, R. and Pretzsch, H. et al. 2018 Biodiversity along temperate forest succession. *Journal of Applied Ecology* **55**, 2756–2766. 10.1111/1365-2664.13238.

Hinchliff, C.E., Smith, S.A., Allman, J.F., Burleigh, J.G., Chaudhary, R. and Coghill, L.M. et al. 2015 Synthesis of phylogeny and taxonomy into a comprehensive tree of life. *PNAS* **112**, 12764–12769. 10.1073/pnas.1423041112.

Holderegger, R., Buehler, D., Gugerli, F. and Manel, S. 2010 Landscape genetics of plants. *Trends in Plant Science* **15**, 675–683. 10.1016/j.tplants.2010.09.002.

Holderegger, R. and Wagner, H.H. 2008 Landscape Genetics. *BioScience* **58**, 199-207. 10.1641/B580306.

Hooper, D.U., Chapin, F.S., Ewel, J.J., Hector, A., Inchausti, P. and Lavorel, S. et al. 2005 Effects of Biodiversity on Ecosystem Functionig: Aa Consensus of Current Knowledge. *Ecological Monographs* **75**, 3–35. 10.1890/04-0922.

Hughes, A.C. and Grumbine, R.E. 2023 The Kunming-Montreal Global Biodiversity Framework: what it does and does not do, and how to improve it. *Front. Environ. Sci.* **11**. 10.3389/fenvs.2023.1281536.

Innes, J.L. and Koch, B. 1998 Forest biodiversity and its assessment by remote sensing. *Global Ecology and Biogeography* **7**, 397–419. 10.1046/j.1466-822X.1998.00314.x.

Ioannidis, J.P.A. 2005 Why most published research findings are false. *PLoS medicine* **2**, e124. 10.1371/journal.pmed.0020124.

IPBES. 2021 *Intergovernmental science-policy platform on biodiversity and ecosystem services.: Summary for policy makers of the global assessment report on biodiversity and ecosystem services of the intergovernmental science-policy platform on biodiversity and ecosystem services.* United Nations, IPBES Secretariat, Bonn, Germany, 8 p.

Isenburg, M. 2014 *LAStools: efficient LiDAR processing software.* rapidlasso GmbH, Gilching, Germany. rapidlasso.com/lastools/.

Ivanova, N., Fomin, V. and Kusbach, A. 2022 Experience of Forest Ecological Classification in Assessment of Vegetation Dynamics. *Sustainability* **14**, 3384. 10.3390/su14063384.

Ivanova, N.S. and Zolotova, E.S. 2014 Development of Forest Typology in Russia. *Inter. Jour. of Bio-reso. Stress Manag.* **5**, 298. 10.5958/0976-4038.2014.00572.7.

Jansen, R.K., Cai, Z., Raubeson, L.A., Daniell, H., Pamphilis, C.W. de and Leebens-Mack, J. et al. 2007 Analysis of 81 genes from 64 plastid genomes resolves relationships in angiosperms and identifies genome-scale evolutionary patterns. *PNAS* **104**, 19369–19374. 10.1073/pnas.0709121104.

Jasiewicz, J. and Stepinski, T.F. 2013 Geomorphons — a pattern recognition approach to classification and mapping of landforms. *Geomorphology* **182**, 147–156. 10.1016/j.geomorph.2012.11.005.

Jin, L.S., Yin, D., Fortin, M.-J. and Cadotte, M.W. 2020 The mechanisms generating community phylogenetic patterns change with spatial scale. *Oecologia* **193**, 655–664. 10.1007/s00442-020-04695-9.

Jin, Y. and Qian, H. 2019 V.PhyloMaker: an R package that can generate very large phylogenies for vascular plants. *Ecography* **42**, 1353–1359. 10.1111/ecog.04434.

Jin, Y. and Qian, H. 2022 V.PhyloMaker2: An updated and enlarged R package that can generate very large phylogenies for vascular plants. *Plant diversity* **44**, 335–339. 10.1016/j.pld.2022.05.005.

Joa, B., Paulus, A., Mikoleit, R. and Winkel, G. 2020 Decision Making in Tree Selection – Contemplating Conflicting Goals via Marteloscope Exercises. *Rural Landscapes: Society, Environment, History* **7**. 10.16993/rl.60.

John, M., Wirth, K., Kaufmann, A., Ertelt, H. and Frei, T. 2024 Forest deliberations: Marteloscopes as sites of encounter between climate activists and forest managers. *Forest Policy and Economics* **169**, 103356. 10.1016/j.forpol.2024.103356.

Jost, L. 2006 Entropy and diversity. *Oikos* **113**, 363–375. 10.1111/j.2006.0030-1299.14714.x.

Jost, L. 2007 Partitioning Diversity into Independent Alpha and Beta Components. *Ecology* **88**, 2427–2439.

Kacic, P. and Kuenzer, C. 2022 Forest Biodiversity Monitoring Based on Remotely Sensed Spectral Diversity—A Review. *Remote Sensing* **14**, 5363. 10.3390/rs14215363.

Kadavý, J., Kneiflová, J., Kneifl, M. and Uherková, B. 2024 Using marteloscope in selection forestry - Study case from 'Pokojná hora' (Czech Republic). *J. For. Sci.* **70**, 447–457. 10.17221/29/2024-JFS.

Kaennel, M. 1998 Biodiversity: a Diversity of Definitions. In *Assessment of Biodiversity for Improved Forest Planning*. P. Bachmann, M. Köhl and R. Päivinen (eds). Kluwer Academic Publishers. 978-90-481-4962-9, pp. 71–81.

Kattge, J., Bönisch, G., Díaz, S., Lavorel, S., Prentice, I.C. and Leadley, P. et al. 2020 TRY plant trait database - enhanced coverage and open access. *Global change biology* **26**, 119–188. 10.1111/gcb.14904.

Kavgacı, A., Karaköse, M., Keleş, E.S., Balpınar, N., Arslan, M. and Yalçın, E. et al. 2023 Classification of forest and shrubland vegetation in central and eastern Euxine Turkey and SW Georgia. *Applied Vegetation Science* **26**. 10.1111/avsc.12753.

Kavtaradze, G., Basilidze, L., Pergl, J., Avoiani, E., Vahalik, P. and Aptsiauri, B. et al. 2023 Distribution and environmental impact of alien woody species in lowland riparian forest habitats: Case study in the protected areas of Georgia, South Caucasus. *J. For. Sci.* **69**, 401–414. 10.17221/3/2023-JFS.

Keck, F., Peller, T., Alther, R., Barouillet, C., Blackman, R. and Capo, E. et al. 2025 The global human impact on biodiversity. *Nature* **641**, 395–400. 10.1038/s41586-025-08752-2.

Keggenhoff, I., Elizbarashvili, M., Amiri-Farahani, A. and King, L. 2014 Trends in daily temperature and precipitation extremes over Georgia, 1971–2010. *Weather and Climate Extremes* **4**, 75–85. 10.1016/j.wace.2014.05.001.

Khare, S., Latifi, H. and Rossi, S. 2019 Forest beta-diversity analysis by remote sensing: How scale and sensors affect the Rao's Q index. *Ecological Indicators* **106**, 105520. 10.1016/j.ecolind.2019.105520.

Kindt, R. and Coe, R. 2005 *Tree diversity analysis: A manual and software for common statistical methods for ecological and biodiversity studies.* World Agroforestry Centre (ICRAF), Nairobi, Kenya. www.worldagroforestry.org, 207 p.

Kissling, W.D., Walls, R., Bowser, A., Jones, M.O., Kattge, J. and Agosti, D. et al. 2018 Towards global data products of Essential Biodiversity Variables on species traits. *Nature ecology & evolution* **2**, 1531–1540. 10.1038/s41559-018-0667-3.

Kleinn, C. 2017 The renaissance of National Forest Inventories (NFIs) in the context of the international conventions - a discussion paper on context, background and justification of NFIs. *Pesq. flor. bras.* **37**, 369–379. 10.4336/2017.pfb.37.91.1343.

Kleinn, C., Kändler, G., Polley, H., Riedel, T. and Schmitz, F. 2020 The National Forest Inventory in Germany: Responding to Forest-Related Information Needs. *Allgemeine Forst- und Jagdzeitung* **19**, 97–118. 10.23765/afjz0002062.

Knollová, I., Chytrý, M., Bruelheide, H., Dullinger, S., Jandt, U. and Bernhardt-Römermann, M. et al. 2024 ReSurveyEurope: A database of resurveyed vegetation plots in Europe. *J Vegetation Science* **35**. 10.1111/jvs.13235.

König, C., Weigelt, P., Schrader, J., Taylor, A., Kattge, J. and Kreft, H. 2019 Biodiversity data integration-the significance of data resolution and domain. *PLoS biology* **17**, e3000183. 10.1371/journal.pbio.3000183.

Kozák, D., Svitok, M., Wiezik, M., Mikoláš, M., Thorn, S. and Buechling, A. et al. 2021 Historical Disturbances Determine Current Taxonomic, Functional and Phylogenetic Diversity of Saproxylic Beetle Communities in Temperate Primary Forests. *Ecosystems* **24**, 37–55. 10.1007/s10021-020-00502-x.

Kraft, N.J.B., Adler, P.B., Godoy, O., James, E.C., Fuller, S. and Levine, J.M. 2015 Community assembly, coexistence and the environmental filtering metaphor. *Functional Ecology* **29**, 592–599. 10.1111/1365-2435.12345.

Krahwinkler, P. and Rossmann, J. 2013 Tree Species Classification and Input Data Evaluation. *European Journal of Remote Sensing* **46**, 535–549. 10.5721/EuJRS20134631.

Kruse, L. 2022 *Marteloscopes in Guria and Tianeti: Technical documentation.* Gesellschaft für Internationale Zusammenarberit (GIZ), Tbilisi, Georgia, 36 p.

Kruse, L. 2023 *Identifying training needs for the implementation of Continuous Cover Forestry in Sweden*. M.Sc., Umeå, Sweden. stud.epsilon.slu.se/19023/, 58 p.

Küchler, A.W. 1949 A Physiognomic Classification of Vegetation. *Annals of the Association of American Geographers* **39**, 201–210. jstor.org/stable/2561207.

Kuenzer, C., Ottinger, M., Wegmann, M., Guo, H., Wang, C. and Zhang, J. et al. 2014 Earth observation satellite sensors for biodiversity monitoring: potentials and bottlenecks. *International Journal of Remote Sensing* **35**, 6599–6647. 10.1080/01431161.2014.964349.

Kumar, A., Choudhury, B., Dayanandan, S. and Khan, M.L. (eds). 2022a *Molecular Genetics and Genomics Tools in Biodiversity Conservation.* 1st edn. Springer Singapore; Imprint Springer, Singapore. 9789811660054. link.springer.com, 3281 p.

Kumar, P., Dobriyal, M., Kale, A., Pandey, A.K., Tomar, R.S. and Thounaojam, E. 2022b Calculating forest species diversity with information-theory based indices using sentinel-2A sensor's of Mahavir Swami Wildlife Sanctuary. *PloS one* **17**, e0268018. 10.1371/journal.pone.0268018.

Lachashvili, N., Eradze, N. and Khetsuriani, L. 2020 Floristic composition of pistachio-woodland (Pistacieta atlantici) of Georgia (South Caucasus). *Fl. Medit.* **30**, 39–53. 10.7320/FlMedit30.039.

Lachashvili, N., Kereselidze, K. and Kikvidze, M. 2022 The checklist of trees of Georgia (Caucasus) and their provisional regional assessment according to IUCN categories and criteria. *Fl. Medit.* **32**, 149–188. 10.7320/FlMedit32.149.

Laliberté, E. and Legendre, P. 2010 A distance-based framework for measuring functional diversity from multiple traits. *Ecology* **91**, 299–305. 10.1890/08-2244.1.

Lapin, K., Oettel, J., Braun, M. and Konrad, H. (eds). 2025 *Ecological Connectivity of Forest Ecosystems.* 1st edn. Springer Nature Switzerland; Imprint Springer, Cham. 9783031822063, 211 p.

Larrieu, L., Courbaud, B., Drénou, C., Goulard, M., Bütler, R. and Kozák, D. et al. 2022 Perspectives: Key factors determining the presence of Tree-related Microhabitats: A synthesis of potential factors at site, stand and tree scales, with perspectives for further research. *Forest Ecology and Management* **515**, 120235. 10.1016/j.foreco.2022.120235.

LaRue, E.A., Downing, A.G., Saucedo, S., Rocha, A., Vargas Zesati, S.A., Mata-Silva, V. and Harvey, M.G. 2023a Diversity – volume relationships: adding structural arrangement and volume to species – area relationships across forest macrosystems. *Ecography* **2023**. 10.1111/ecog.06723.

LaRue, E.A., Knott, J.A., Domke, G.M., Chen, H.Y.H., Guo, Q. and Hisano, M. et al. 2023b Structural diversity as a reliable and novel predictor for ecosystem productivity. *Frontiers in Ecol & Environ* **21**, 33–39. 10.1002/fee.2586.

Laureto, L.M.O., Cianciaruso, M.V. and Samia, D.S.M. 2015 Functional diversity: an overview of its history and applicability. *Natureza & Conservação* **13**, 112–116. 10.1016/j.ncon.2015.11.001.

Lausch, A., Baade, J., Bannehr, L., Borg, E., Bumberger, J. and Chabrilliat, S. et al. 2019 Linking Remote Sensing and Geodiversity and Their Traits Relevant to Biodiversity—Part I: Soil Characteristics. *Remote Sensing* **11**, 2356. 10.3390/rs11202356.

Lausch, A., Bannehr, L., Beckmann, M., Boehm, C., Feilhauer, H. and Hacker, J.M. et al. 2016a Linking Earth Observation and taxonomic, structural and functional biodiversity: Local to ecosystem perspectives. *Ecological Indicators* **70**, 317–339. 10.1016/j.ecolind.2016.06.022. sciencedirect.com/science/article/pii/s1470160x16303223.

Lausch, A., Erasmi, S., King, D., Magdon, P. and Heurich, M. 2016b Understanding Forest Health with Remote Sensing -Part I—A Review of Spectral Traits, Processes and Remote-Sensing Characteristics. *Remote Sensing* **8**, 1029. 10.3390/rs8121029. mdpi.com/171038.

Lausch, A., Heurich, M., Magdon, P., Rocchini, D., Schulz, K., Bumberger, J. and King, D.J. 2020 A Range of Earth Observation Techniques for Assessing Plant Diversity. In *Remote Sensing of Plant Biodiversity*. J. Cavender-Bares, J.A. Gamon and P.A. Townsend (eds). Springer International Publishing. 978-3-030-33156-6, pp. 309–348.

Lee, J.-Y. 2023 The Principles and Applications of High-Throughput Sequencing Technologies. *Dev. Reprod.* **27**, 9–24. 10.12717/DR.2023.27.1.9.

Legendre, P. 2013 Indicator Species: Computation. In *Encyclopedia of Biodiversity*. Elsevier. 9780123847201, pp. 264–268.

Legendre, P. and Legendre, L. 2012 *Numerical Ecology*. Third English Edition. Elsevier BV, Great Britain. 978-0-444-53868-0, 1003 p.

Leinster, T. and Cobbold, C.A. 2012 Measuring diversity: the importance of species similarity. *Ecology* **93**, 477–489. 10.1890/10-2402.1.

Lepš, J. and Bello, F. de. 2023 Differences in trait–environment relationships: Implications for community weighted means tests. *Journal of Ecology* **111**, 2328–2341. 10.1111/1365-2745.14172.

Lomsadze, Z., Chitanava, N., Paresishvili, O., Magalashvili, G., Vezirishvili-Nozadze, K. and Makharadze, K. et al. 2019 Georgian Natural Resources (Brief Review). *Annals of Agrarian Science* **17**, 59–74. journals.org.ge/index.php/aans/article/view/299/226.

Lötter, M., Mucina, L. and Witkowski, E. 2013 The classification conundrum: species fidelity as leading criterion in search of a rigorous method to classify a complex forest data set. *Community Ecology* **14**, 121–132. 10.1556/ComEc.14.2013.1.13.

Ma, X., Migliavacca, M., Wirth, C., Bohn, F.J., Huth, A., Richter, R. and Mahecha, M.D. 2020 Monitoring Plant Functional Diversity Using the Reflectance and Echo from Space. *Remote Sensing* **12**, 1248. 10.3390/rs12081248.

Mace, G.M., Norris, K. and Fitter, A.H. 2012 Biodiversity and ecosystem services: a multilayered relationship. *Trends in Ecology & Evolution* **27**, 19–26. 10.1016/j.tree.2011.08.006.

Machavariani, pers. comm.

Machavariani, M. 2010 *Forestry Standards and Practices in Georgia: Technical Report.* ENPI-FLEG Program, USAID, Tbilisi, Georgia, 46 p.

Macía, M.J. 2008 Woody plants diversity, floristic composition and land use history in the Amazonian rain forests of Madidi National Park, Bolivia. *Biodivers Conserv* **17**, 2671–2690. 10.1007/s10531-008-9348-x.

Magurran, A.E. 2005 Biological diversity. *Current Biology* **15**, 116-118. 10.1016/j.cub.2005.02.006.

Magurran, A.E. and McGill, B.J. 2011 *Biological diversity: Frontiers in measurement and assessment.* Oxford University Press, Oxford, New York. 9780199580675, xvii, 345.

Mamadashvili, G., Brin, A., Bässler, C., Chumak, V., Chumak, M. and Deidus, V. et al. 2023 Drivers of tree-related microhabitat profiles in European and Oriental beech forests. *Biological Conservation* **285**, 110245. 10.1016/j.biocon.2023.110245.

Mammola, S., Carmona, C.P., Guillerme, T. and Cardoso, P. 2021 Concepts and applications in functional diversity. *Functional Ecology* **35**, 1869–1885. 10.1111/1365-2435.13882.

Manel, S. and Holderegger, R. 2013 Ten years of landscape genetics. *Trends in Ecology & Evolution* **28**, 614–621. 10.1016/j.tree.2013.05.012.

Mason, N.W.H., Mouillot, D., Lee, W.G. and Wilson, J.B. 2005 Functional richness, functional evenness and functional divergence: the primary components of functional diversity. *Oikos* **111**, 112–118. 10.1111/j.0030-1299.2005.13886.x.

Maurer, B.A. and McGill, B.J. Biological Diversity. Frontiers in measurement and assessment. Measurement of species diversity. In *Biological diversity: frontiers in measurement and assessment*, pp. 55–65.

Mayfield, M.M. and Levine, J.M. 2010 Opposing effects of competitive exclusion on the phylogenetic structure of communities. *Ecology letters* **13**, 1085–1093. 10.1111/j.1461-0248.2010.01509.x.

Mazel, F., Pennell, M.W., Cadotte, M.W., Diaz, S., Dalla Riva, G.V. and Grenyer, R. et al. 2018 Prioritizing phylogenetic diversity captures functional diversity unreliably. *Nature communications* **9**, 2888. 10.1038/s41467-018-05126-3.

McGill, B. and Magurran, A.E. 2011 Measuring the spatial structure of biodiversity. *Biological diversity: frontiers in measurement and assessment*, 152–171. brianmcgill.org/614/reading/mcgill_space.pdf.

McGill, B.J., Enquist, B.J., Weiher, E. and Westoby, M. 2006 Rebuilding community ecology from functional traits. *Trends in Ecology & Evolution* **21**, 178–185. 10.1016/j.tree.2006.02.002.

McRoberts, R.E. 2011 Satellite image-based maps: Scientific inference or pretty pictures? *Remote Sensing of Environment* **115**, 715–724. 10.1016/j.rse.2010.10.013.

McRoberts, R.E., Tomppo, E., Schadauer, K., Vidal, C., Ståhl, G. and Chirici, G. et al. 2009 Harmonizing National Forest Inventories. *Journal of Forestry* **107**, 179–187. doi.org/10.1093/jof/107.4.179.

McRoberts, R.E. and Tomppo, E.O. 2007 Remote sensing support for national forest inventories. *Remote Sensing of Environment* **110**, 412–419. 10.1016/j.rse.2006.09.034.

McRoberts, R.E., Winter, S., Chirici, G., Hauk, E., Pelz, D.R., Moser, W.K. and Hatfield, M.A. 2008 Large-scale spatial patterns of forest structural diversity. *Can. J. For. Res.* **38**, 429–438. 10.1139/X07-154.

MEPA. 2018 *Field Manual for the Georgian National Forest Inventory: Part II (unpublished).* MEPA, Tbilisi, Georgia, 74 p. 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, Tbilisi, Georgia, 123 p. mepa.gov.ge/En/Files/Download/53934.

Metreveli, V., Kreft, H., Akobia, I., Janiashvili, Z., Nonashvili, Z. and Dzadzamia, L. et al. 2023 Potential Distribution and Suitable Habitat for Chestnut (Castanea sativa). *Forests* **14**, 2076. 10.3390/f14102076.

Metreveli, V., Kreft, H. and Gavashelishvili, A. 2024 Environmental covariates of chestnut blight (Cryphonectria parasitica) in Georgia (Caucasus). *Forest Ecology and Management* **569**, 122153. 10.1016/j.foreco.2024.122153.

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

Mirarab, S. and Warnow, T. 2015 ASTRAL-II: coalescent-based species tree estimation with many hundreds of taxa and thousands of genes. *Bioinformatics (Oxford, England)* **31**, i44-52. 10.1093/bioinformatics/btv234.

Moreno-Fernández, D., Breidenbach, J., Cañellas, I., Chirici, G., D'Amico, G. and Ferretti, M. et al. 2025 Enhancing forest biodiversity indicators in inventories through harmonized protocols. *iForest* **18**, 109–120. 10.3832/ifor4778-018.

Moreno-Fernández, D., Cañellas, I., Alberdi, I. and Montes, F. 2021 Improved stand structure characterization from nested plot designs in the Spanish National Forest Inventory. *Forestry: An International Journal of Forest Research* **94**, 244–257. 10.1093/forestry/cpaa031.

Moudrý, V., Lecours, V., Malavasi, M., Misiuk, B., Gábor, L. and Gdulová, K. et al. 2019 Potential pitfalls in rescaling digital terrain model-derived attributes for ecological studies. *Ecological Informatics* **54**, 100987. 10.1016/j.ecoinf.2019.100987.

Mouquet, N., Devictor, V., Meynard, C.N., Munoz, F., Bersier, L.-F. and Chave, J. et al. 2012 Ecophylogenetics: advances and perspectives. *Biological reviews of the Cambridge Philosophical Society* **87**, 769–785. 10.1111/j.1469-185X.2012.00224.x.

Mumladze, L., Japoshvili, B. and Anderson, E.P. 2020 Faunal biodiversity research in the Republic of Georgia: a short review of trends, gaps, and needs in the Caucasus biodiversity hotspot. *Biologia* **75**, 1385–1397. 10.2478/s11756-019-00398-6.

Münkemüller, T., Gallien, L., Pollock, L.J., Barros, C., Carboni, M. and Chalmandrier, L. et al. 2020 Dos and don'ts when inferring assembly rules from diversity patterns. *Global Ecology and Biogeography* **29**, 1212–1229. 10.1111/geb.13098.

Myers, N., Mittermeier, R.A., Mittermeier, C.G., Da Fonseca, G.A.B. and Kent, J. 2000 Biodiversity hotspots for conservation priorities. *Nature* **403**, 853–858. 10.1038/35002501.

Naeem, S., Prager, C., Weeks, B., Varga, A., Flynn, D.F.B. and Griffin, K. et al. 2016 Biodiversity as a multidimensional construct: a review, framework and case study of herbivory's impact on plant biodiversity. *Proceedings. Biological sciences* **283**. 10.1098/rspb.2015.3005.

Nakhutsrishvili, G. 2013a Forest Vegetation of Georgia. In *The Vegetation of Georgia (South Caucasus)*. G. Nakhutsrishvili (ed). Springer Berlin Heidelberg. 978-3-642-29914-8, pp. 35–87.

Nakhutsrishvili, G. (ed). 2013b *The Vegetation of Georgia (South Caucasus): Geobotany Studies.* 1st edn. Springer Science & Business Media, Heidelberg New York Dordrecht London. 978-3-642-29914-8, 236 p.

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

Nakhutsrishvili, G., Abdaladze, O., Batsatsashvili, K. and Dzadzamia, L. (eds). 2023 *Natural Forests of Georgia: (The South Caucasus).* Ilia State University Press, Tbilisi, Georgia. 978-9941-18-445-1, 224 p.

Nakhutsrishvili, G., Zazanashvili, N., Batsatsashvili, K. and Montalvo Mancheno, C., S. 2015 Colchic and Hyrcanian forests of the Caucasus: similarities, differences and conservation status. In *Festschrift for Francesco Maria Raimondo on the occasion of his 70th birthday*, p. 185.

Ndiribe, C., Pellissier, L., Antonelli, S., Dubuis, A., Pottier, J. and Vittoz, P. et al. 2013 Phylogenetic plant community structure along elevation is lineage specific. *Ecology and evolution* **3**, 4925–4939. 10.1002/ece3.868.

Neumann, C., Weiss, G., Schmidtlein, S., Itzerott, S., Lausch, A., Doktor, D. and Brell, M. 2015 Gradient-Based Assessment of Habitat Quality for Spectral Ecosystem Monitoring. *Remote Sensing* **7**, 2871–2898. 10.3390/rs70302871.

Newton, A.C. and Kapos, V. 2002 Biodiversity indicators in national forest inventories. *Unasylva* **53**, 56–64. scopus.com/record/display.uri?eid=2-s2.0-0036958979&origin=inward&txGid=d4e9862d2823739e278c35a31f4f5156.

Niklas, K.J. and Crepet, W.L. 2020 Morphological (and not anatomical or reproductive) features define early vascular plant phylogenetic relationships. *American journal of botany* **107**, 477–488. 10.1002/ajb2.1440.

Novák, P., Kalníková, V., Szokala, D., Aleksanyan, A., Batsatsashvili, K. and Fayvush, G. et al. 2023a Transcaucasian Vegetation Database – a phytosociological database of the Southern Caucasus. *VCS* **4**, 231–240. 10.3897/VCS.105521.

Novák, P., Kalníková, V., Večeřa, M., Štětková, G., Sedláček, V. and Pustková, Š. et al. 2023b Alluvial alder forests of the Greater Caucasus, Georgia: ecology, habitats and variability. *Tuexenia* **43**, 1–28. 10.14471/2023.43.003.

Novák, P., Stupar, V. and Kalníková, V. 2021 Carpinus orientalis forests in Georgian Colchis: First insights. *Tuexenia* **41**, 37–51. 10.14471/2021.41.012.

Novák, P., Zukal, D., Harásek, M., Vlčková, P., Abdaladze, O. and Willner, W. 2020 Ecology and vegetation types of oak-hornbeam and ravine forests of the Eastern Greater Caucasus, Georgia. *Folia Geobot* **55**, 333–349. 10.1007/s12224-020-09386-0.

Olofsson, P., Torchinava, P., Woodcock, C.E., Baccini, A., Houghton, R. and Ozdogan, M. et al. 2010 Implications of land use change on the national terrestrial carbon budget of Georgia. *Carbon balance and management* **5**, 1–13. cbmjournal.com/content/5/1/4.

One Thousand Plant Transcriptomes Initiative. 2019 One thousand plant transcriptomes and the phylogenomics of green plants. *Nature* **574**, 679–685. 10.1038/s41586-019-1693-2.

Oyebanji, O.O., Onditi, K.O., Azevedo, J.A.R., Rahaingoson, F.R., Nneji, L.M. and Adeleye, M.A. et al. 2023 Biogeographic patterns and environmental drivers of species richness in the globally distributed Millettioid/Phaseoloid clade (Fabaceae, subfamily Papilionoideae). *Front. Ecol. Evol.* **11**. 10.3389/fevo.2023.1231553.

Paillet, Y. 2017 Suivis nationaux de biodiversité en forêt en France: une lecture au travers des Variables Essentielles de Biodiversité. [in French]. *Naturae* **6**, 1–11. revue-naturae.fr/.

Patarkalashvili, T. 2016 Some problems of forest management of Georgia. *Annals of Agrarian Science* **14**, 108–113. 10.1016/j.aasci.2016.04.002.

Patarkalashvili, T. 2017 Forest biodiversity of Georgia and endangered plant species. *Annals of Agrarian Science* **15**, 349–351. 10.1016/j.aasci.2017.06.002.

Pause, M., Schweitzer, C., Rosenthal, M., Keuck, V., Bumberger, J. and Dietrich, P. et al. 2016 In Situ/Remote Sensing Integration to Assess Forest Health—A Review. *Remote Sensing* **8**, 471. 10.3390/rs8060471. mdpi.com/2072-4292/8/6/471.

Pavoine, S. and Bonsall, M.B. 2011 Measuring biodiversity to explain community assembly: a unified approach. *Biological reviews of the Cambridge Philosophical Society* **86**, 792–812. 10.1111/j.1469-185X.2010.00171.x.

Peet, R., Lee, M., Jennings, M. and Faber-Langendoen, D. 2012 VegBank – a permanent, open-access archive for vegetation-plot data. *Biodivers. Ecol.* **4**, 233–241. 10.7809/b-e.00080.

Pellens, R. and Grandcolas, P. (eds). 2016 *Biodiversity Conservation and Phylogenetic Systematics: Preserving our evolutionary heritage in an extinction crisis.* Springer, Cham. 9783319224619, 396 p.

Perrone, M., Di Febbraro, M., Conti, L., Divíšek, J., Chytrý, M. and Keil, P. et al. 2023 The relationship between spectral and plant diversity: Disentangling the influence of metrics and habitat types at the landscape scale. *Remote Sensing of Environment* **293**. 10.1016/j.rse.2023.113591.

Petchey, O.L. and Gaston, K.J. 2002 Functional diversity (FD), species richness and community composition. *Ecology letters* **5**, 402–411. 10.1046/j.1461-0248.2002.00339.x.

Petchey, O.L. and Gaston, K.J. 2006 Functional diversity: back to basics and looking forward. *Ecology letters* **9**, 741–758. 10.1111/j.1461-0248.2006.00924.x.

Pettorelli, N., Safi, K. and Turner, W. 2014 Satellite remote sensing, biodiversity research and conservation of the future. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* **369**, 20130190. 10.1098/rstb.2013.0190.

Petzer, T. 2023 *Forum Interdisziplinäre Begriffsgeschichte: Vladimir Sukachev's Concept of Biogeocoenosis.* 12th edn., Berlin, 9 p. publikationen.ub.uni-frankfurt.de/frontdoor/index/index/docId/74810.

Pielou, E.C. 1966 The measurement of diversity in different types of biological collections. *Journal of Theoretical Biology* **15**, 131–144. 10.1016/0022-5193(67)90048-3.

Pielou, E.C. 1975 Ecological models on an environmental gradient. *Applied Statistics (Gupta, RP, ed.)*, 261–269.

Pommerening, A. 2023 Continuous cover forestry: Opportunities for changing forests. *Trees, Forests and People* **14**, 100448. 10.1016/j.tfp.2023.100448.

Potter, K.M. and Woodall, C.W. 2012 Trends over time in tree and seedling phylogenetic Trends over time in tree and seedling phylogenetic diversity indicate regional differences in forest biodiversity changebiodiversity change. *Ecological Applications* **22**, 517–531. srs.fs.usda.gov/pubs/ja/2012/ja_2012_potter_001.pdf.

Pretzsch, H. 2009 *Forest Dynamics, Growth and Yield: From Measurement to Model,.* Springer, Berlin, Heidelberg. 9783540883067, xix, 664.

Provete, D.B. 2013 What Is on the Horizon for Ecophylogenetics? *NatCon* **11**, 7–14. 10.4322/natcon.2013.001.

Puletti, N., Canullo, R., Mattioli, W., Gawryś, R., Corona, P. and Czerepko, J. 2019 A dataset of forest volume deadwood estimates for Europe. *Annals of Forest Science* **76**. 10.1007/s13595-019-0832-0.

QGIS Development Team. 2009 *QGIS Geographic Information System.* QGIS Development Team. qgis.osgeo.org/.

Qian, H. and Jin, Y. 2016 An updated megaphylogeny of plants, a tool for generating plant phylogenies and an analysis of phylogenetic community structure. *JPECOL* **9**, 233–239. 10.1093/jpe/rtv047.

Qian, H. and Jin, Y. 2021 Are phylogenies resolved at the genus level appropriate for studies on phylogenetic structure of species assemblages? *Plant diversity* **43**, 255–263. 10.1016/j.pld.2020.11.005.

Qian, H., Zhang, Y., Zhang, J. and Wang, X. 2013 Latitudinal gradients in phylogenetic relatedness of angiosperm trees in North America. *Global Ecology and Biogeography* **22**, 1183–1191. 10.1111/geb.12069.

R Core Team. 2024 *R: A Language and Environment for Statistical Computing.* R Foundation for Statistical Computing, Vienna, Austria. r-project.org/.

Ramensky, L.G., Tsatsenkin, I.A., Chizhikov, O.N. and Antipin, N.A. 1956 *Экологическая оценка кормовых угодий по растительному покрову: Ecological Assessment of Fodder Lands by Vegetation Cover: [in Russian].* Sel'khozgiz, Moscow, Russia, 472 p.

Ratnasingham, S. and Herbert, P.D.N. 2007 bold: The Barcode of Life Data System (http://www.barcodinglife.org). *Molecular Ecology Notes* **7**, 355–364. 10.1111/j.1471-8286.2006.01678.x.

Reise, J., Kukulka, F., Flade, M. and Winter, S. 2019 Characterising the richness and diversity of forest bird species using National Forest Inventory data in Germany. *Forest Ecology and Management* **432**, 799–811. 10.1016/j.foreco.2018.10.012.

Reuter, J.A., Spacek, D.V. and Snyder, M.P. 2015 High-Throughput Sequencing Technologies. *Molecular Cell* **58**, 586–597. 10.1016/j.molcel.2015.05.004.

Ricotta, C. 2005 Through the jungle of Biological diversity. *Acta Biotheoretica* **53**, 29–38. doi.org/10.1007/s10441-005-7001-6.

Ricotta, C. and Avena, G. 2003 On the relationship between Pielou's evenness and landscape dominance within the context of Hill's diversity profiles. *Ecological Indicators* **2**, 361–365. 10.1016/S1470-160X(03)00005-0.

Ricotta, C. and Pavoine, S. 2015 Measuring similarity among plots including similarity among species: an extension of traditional approaches. *Journal of Vegetation Science* **26**, 1061–1067. 10.1111/jvs.12329.

Ricotta, C. and Podani, J. 2017 On some properties of the Bray-Curtis dissimilarity and their ecological meaning. *Ecological Complexity* **31**, 201–205. 10.1016/j.ecocom.2017.07.003.

Ripley, B.D. 1977 Modelling spatial patterns. *Journal of the Royal Statistical Society: Series B (Methodological)* **39**, 172–192. 10.1111/j.2517-6161.1977.tb01615.x.

Rocchini, D., Bacaro, G., Chirici, G., Da Re, D., Feilhauer, H. and Foody, G.M. et al. 2018 Remotely sensed spatial heterogeneity as an exploratory tool for taxonomic and functional diversity study. *Ecological Indicators* **85**, 983–990. 10.1016/j.ecolind.2017.09.055.

Rocchini, D., Balkenhol, N., Carter, G.A., Foody, G.M., Gillespie, T.W. and He, K.S. et al. 2010 Remotely sensed spectral heterogeneity as a proxy of species diversity: Recent advances and open challenges. *Ecological Informatics* **5**, 318–329. 10.1016/j.ecoinf.2010.06.001.

Rocchini, D., Boyd, D.S., Féret, J.B., Foody, G.M., He, K.S. and Lausch, A. et al. 2016 Satellite remote sensing to monitor species diversity: potential and pitfalls. *Remote Sens Ecol Conserv* **2**, 25–36. 10.1002/rse2.9.

Rocchini, D., Feilhauer, H., Müllerová, J. and Schmidtlein, S. 2022 Remote sensing for vegetation science: A virtual special issue on its power and challenges. *Applied Vegetation Science* **25**. 10.1111/avsc.12677.

Rocchini, D., Salvatori, N., Beierkuhnlein, C., Chiarucci, A., Boissieu, F. de and Förster, M. et al. 2021 From local spectral species to global spectral communities: A benchmark for ecosystem diversity estimate by remote sensing. *Ecological Informatics* **61**, 101195. 10.1016/j.ecoinf.2020.101195.

Rodrigues, A.C., Villa, P.M., Ferreira-Júnior, W.G., Schaefer, C.E.R.G. and Neri, A.V. 2021 Effects of topographic variability and forest attributes on fine-scale soil fertility in late-secondary succession of Atlantic Forest. *Ecol Process* **10**. 10.1186/s13717-021-00333-1.

Roleček, J., Tichý, L., Zelený, D. and Chytrý, M. 2009 Modified TWINSPAN classification in which the hierarchy respects cluster heterogeneity. *J Vegetation Science* **20**, 596–602. 10.1111/j.1654-1103.2009.01062.x.

RStudio Team. 2024 *RStudio: Integrated Development for R.: RStudio.* Posit Software, P. B.C., Boston, MA, US.

Schaffers, A.P. and Sýkora, K.V. 2000 Reliability of Ellenberg indicator values for moisture, nitrogen and soil reaction: a comparison with field measurements. *J Vegetation Science* **11**, 225–244. 10.2307/3236802.

Schmidtlein, S., Collison, J. and Pfannendoerfer, R. 2024 *Package 'isopam': Clustering of Sites with Species Data*, 10 p. cran.r-project.org/package=isopam.

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

Schoville, S.D., Bonin, A., François, O., Lobreaux, S., Melodelima, C. and Manel, S. 2012 Adaptive Genetic Variation on the Landscape: Methods and Cases. *Annu. Rev. Ecol. Evol. Syst.* **43**, 23–43. 10.1146/annurev-ecolsys-110411-160248.

Seidling, W., Travaglini, D., Meyer, P., Waldner, P., Fischer, R. and Granke, O. et al. 2014 Dead wood and stand structure - relationships for forest plots across Europe. *iForest* **7**, 269–281. 10.3832/ifor1057-007.

Simon Véron, Victor Saito, Nélida Padilla-García, Félix Forest and and Yves Bertheau. 2019 The Use of Phylogenetic Diversity in Conservation Biology and Community Ecology: A Common Base but Different Approaches. *The Quarterly Review of Biology* **94**. hal.sorbonne-universite.fr/hal-02290915/file/QRB%20HAL.pdf.

Skidmore, A.K., Coops, N.C., Neinavaz, E., Ali, A., Schaepman, M.E. and Paganini, M. et al. 2021 Priority list of biodiversity metrics to observe from space. *Nature ecology & evolution* **5**, 896–906. 10.1038/s41559-021-01451-x.

Slik, J.W.F., Franklin, J., Arroyo-Rodríguez, V., Field, R., Aguilar, S. and Aguirre, N. et al. 2018 Phylogenetic classification of the world's tropical forests. *Proceedings of the National Academy of Sciences of the United States of America* **115**, 1837–1842. 10.1073/pnas.1714977115.

Smith, S.A. and Brown, J.W. 2018 Constructing a broadly inclusive seed plant phylogeny. *American journal of botany* **105**, 302–314. 10.1002/ajb2.1019.

Soltis, D.E., Soltis, P.S., Nickrent, D.L., Johnson, L.A., Hahn, W.J. and Hoot, S.B. et al. 1997 Angiosperm Phylogeny Inferred from 18S Ribosomal DNA Sequences. *Annals of the Missouri Botanical Garden,* **84**, 1–49. 10.2307/2399952.

Soltis, P.S. and Soltis, D.E. 2016 Ancient WGD events as drivers of key innovations in angiosperms. *Current opinion in plant biology* **30**, 159–165. 10.1016/j.pbi.2016.03.015.

Song, K., Zhang, J., Yan, E., Wang, X. and Song, Y. 2024 EcoFloVCS: An eco-physiognomic-floristic vegetation classification system. *Applied Vegetation Science* **27**. 10.1111/avsc.70005.

Srivastava, D.S., Cadotte, M.W., MacDonald, A.A.M., Marushia, R.G. and Mirotchnick, N. 2012 Phylogenetic diversity and the functioning of ecosystems. *Ecology letters* **15**, 637–648. 10.1111/j.1461-0248.2012.01795.x.

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

Steffen, W., Broadgate, W., Deutsch, L., Gaffney, O. and Ludwig, C. 2015 The trajectory of the Anthropocene: The Great Acceleration. *The Anthropocene Review* **2**, 81–98. 10.1177/2053019614564785.

Steinmann, K., Eggenberg, S., Wohlgemuth, T., Linder, H.P. and Zimmermann, N.E. 2011 Niches and noise—Disentangling habitat diversity and area effect on species diversity. *Ecological Complexity* **8**, 313–319. 10.1016/j.ecocom.2011.06.004.

Stepinski, T.F. and Jasiewicz, J. 2011 Geomorphons - a new approach to classification of landforms. 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.

Storch, F., Dormann, C.F. and Bauhus, J. 2018 Quantifying forest structural diversity based on large-scale inventory data: a new approach to support biodiversity monitoring. *Forest ecosystems* **5**. 10.1186/s40663-018-0151-1.

Storfer, A., Murphy, M.A., Evans, J.S., Goldberg, C.S., Robinson, S. and Spear, S.F. et al. 2007 Putting the 'landscape' in landscape genetics. *Heredity* **98**, 128–142. 10.1038/sj.hdy.6800917.

Stritih, A., Senf, C., Kuemmerle, T., Munteanu, C., Dzadzamia, L. and 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. *Landscape Ecology* **39**. 10.1007/s10980-024-01908-x.

Sukachev, V. and Dylis, N. 1964 Основы лесной биогеоценологии. Fundamentals of forest biogeocoenology [in Russian].

Swenson, N.G., Enquist, B.J., Pither, J., Thompson, J. and Zimmerman, J.K. 2006 The Problem and Promise of Scale Dependency in Community Phylogenetics. *Ecology* **87**, 2418–2424. 10.1890/0012-9658(2006)87[2418:TPAPOS]2.0.CO;2.

Swenson, N.G., Erickson, D.L., Mi, X., Bourg, N.A., Forero-Montaña, J. and Ge, X. et al. 2012 Phylogenetic and functional alpha and beta diversity in temperate and tropical tree communities. *Ecology* **93**, S112-S125. 10.1890/11-0402.1.

Tallents, l.A., Lovett, J.C., Hall, J.B. and Hamilton, A.C. 2005 Phylogenetic diversity of forest trees in the Usambara mountains of Tanzania: correlations with altitude. *Botanical Journal of the Linnean Society* **149**, 217–228. 10.1111/j.1095-8339.2005.00431.x.

Tarkhnishvili, D., Gavashelishvili, A. and Mumladze, L. 2012 Palaeoclimatic models help to understand current distribution of Caucasian forest species. *Biological Journal of the Linnean Society* **105**, 231–248. 10.1111/j.1095-8312.2011.01788.x.

Tenenbaum, J.B., Silva, V. de and Langford, J.C. 2000 A global geometric framework for nonlinear dimensionality reduction. *Science (New York, N.Y.)* **290**, 2319–2323. 10.1126/science.290.5500.2319.

Tephnadze, N., Abdaladze, O., Nakhutsrishvili, G., Simmering, D., Waldhardt, R. and Otte, A. 2014 The impacts of management and site conditions on the phytodiversity of the upper montane and subalpine belts in the Central Greater Caucasus. *phyto* **44**, 255–291. 10.1127/0340-269X/2014/0044-0579.

Thompson, I., Guariguata, M.R., Okabe, K., Bahamondez, C., Nasi, R., Heymel, V. and Sabogal, C. 2013 An Operational Framework for Defining and Monitoring Forest Degradation. *Ecology and Society* **18**, 2ß. ecologyandsociety.org/vol18/iss2/art20/.

Thompson, I., Mackey, B., McNulty, S. and Mosseler, A. 2009 *Forest resilience, biodiversity, and climate change: A synthesis of the Biodiversity / Resilience / Stability Relationship in Forest Ecosystems.* Secretariat of the Convention on Biological Diversity, Montreal. 9292251376, 67 p.

Tilman, D., Knops, J., Wedin, D., Reich, P., Ritchie, M. and Siemann, E. 1997 The Influence of Functional Diversity and Composition on Ecosystem Processes. *Science* **277**, 1300–1302. 10.1126/science.277.5330.1300.

Togonidze, N. 2015 Climate Changes and Anthropogenic Influence on Subalpine Birch Forest **2**, 1394-1375. ijries.org/administrator/components/com_jresearch/files/publications/IJRIES_91_Final.pdf.

Tomppo, E., Gschwantner, T., Lawrence, M. and McRoberts, R.E. 2010 *National Forest Inventories.* Springer Netherlands, Dordrecht. 978-90-481-3232-4, 613 p.

Tomppo, E. and Katila, M. (eds). 1991 *Satellite image-based national forest inventory of finland for publication in the igarss' 91 digest.* Remote Sensing: Global Monitoring for Earth Management.

Tomter, S.M., Kuliešis, A. and Gschwantner, T. 2016 Annual volume increment of the European forests—description and evaluation of the national methods used. *Annals of Forest Science* **73**, 849–856. 10.1007/s13595-016-0557-2.

Torresani, M., Feilhauer, H., Rocchini, D., Féret, J.-B., Zebisch, M. and Tonon, G. 2021 Which optical traits enable an estimation of tree species diversity based on the Spectral Variation Hypothesis? *Applied Vegetation Science* **24**. 10.1111/avsc.12586.

Torresani, M., Rocchini, D., Sonnenschein, R., Zebisch, M., Marcantonio, M., Ricotta, C. and Tonon, G. 2019 Estimating tree species diversity from space in an alpine conifer forest: The Rao's Q diversity index meets the spectral variation hypothesis. *Ecological Informatics* **52**, 26–34. 10.1016/j.ecoinf.2019.04.001.

Trentanovi, G., Campagnaro, T., Sitzia, T., Chianucci, F., Vacchiano, G. and Ammer, C. et al. 2023 Words apart: Standardizing forestry terms and definitions across European biodiversity studies. *Forest ecosystems* **10**, 100128. 10.1016/j.fecs.2023.100128.

Trucios Caciano, R. 2020 *Quantifying the uncertainty caused by sampling, modeling, and field measurements in the estimation of AGB with information of the national forest inventory in Durango, Mexico*. Ph. D. Thesis, Göttingen, 182 p.

Tsyganov, D.N. 1984 *Фитоиндикация экологических режимов в подзоне хвойно-широколиственных лесов: Phytoindication of ecological regimes in the mixed coniferous- broad- leaved forest subzone [in Russian],* Moscow. ilan.ras.ru/wp-content/uploads/2020/03/cueganov-d.n._-fitoindikaciya-yekologicheskih-rez.pdf, 101 p.

Tucker, C.M., Cadotte, M.W., Carvalho, S.B., Davies, T.J., Ferrier, S. and Fritz, S.A. et al. 2017 A guide to phylogenetic metrics for conservation, community ecology and macroecology. *Biological reviews of the Cambridge Philosophical Society* **92**, 698–715. 10.1111/brv.12252.

Tuomisto, H. 2010 A diversity of beta diversities: straightening up a concept gone awry. Part 1. Defining beta diversity as a function of alpha and gamma diversity. *Ecography* **33**, 2–22. 10.1111/j.1600-0587.2009.05880.x.

UNEP. 2022 *Decision adopted by the Conference of the Parties to the Convention on Biological Diversity. Decision adopted by the Conference of the Parties to the Convention on Biological Diversity. 15/4. Kunming-Montreal Global Biodiversity Framework.* Convention on Biological Diversity, 15 p. www.cbd.int/doc/decisions/cop-15/cop-15-dec-04-en.pdf.

Urushadze, T., Manvelidze, Z. and Dolidze, Lasha, Tvauri, Irina *Oriental Beech in Georgia – Present State and Conservation Priorities.* Institude for Forest Genetics, 24 p. vulhm.cz/files/uploads/2019/03/Communicationes_25_2.cast_.pdf.

Valladares, F., Bastias, C.C., Godoy, O., Granda, E. and Escudero, A. 2015 Species coexistence in a changing world. *Frontiers in plant science* **6**, 866. 10.3389/fpls.2015.00866.

Vamosi, S.M., Heard, S.B., Vamosi, J.C. and Webb, C.O. 2009 Emerging patterns in the comparative analysis of phylogenetic community structure. *Molecular ecology* **18**, 572–592. 10.1111/j.1365-294X.2008.04001.x.

van der Maarel, E. 1975 The Braun-Blanquet Approach in Perspective. *Vegetatio* **30**, 213–219. jstor.org/stable/20036872.

Vellend, M., Baeten, L., Myers-Smith, I.H., Elmendorf, S.C., Beauséjour, R. and Brown, C.D. et al. 2013 Global meta-analysis reveals no net change in local-scale plant biodiversity over time. *Proceedings of the National Academy of Sciences of the United States of America* **110**, 19456–19459. 10.1073/pnas.1312779110.

Vidal, C., Alberdi, I., Redmond, J., Vestman, M., Lanz, A. and Schadauer, K. 2016a The role of European National Forest Inventories for international forestry reporting. *Annals of Forest Science* **73**, 793–806. 10.1007/s13595-016-0545-6.

Vidal, C., Alberdi, I.A., Hernández, L. and Redmond, J.J. 2016b *National Forest Inventories: Assessment of Wood Availability and Use.* Springer International Publishing, Cham, Switzerland. 978-3-319-44014-9, 847 p.

Viedma, O., Torres, I., Pérez, B. and Moreno, J.M. 2012 Modeling plant species richness using reflectance and texture data derived from QuickBird in a recently burned area of Central Spain. *Remote Sensing of Environment* **119**, 208–221. 10.1016/j.rse.2011.12.024.

Vigués Jorba, J., Scherrer, D., Duchenne, F., Zellweger, F., Gossner, M.M. and Bollmann, K. 2025 Differential responses of taxonomic, functional and phylogenetic multi-taxa diversity to environmental factors in temperate forest ecosystems. *Ecological Indicators* **178**, 113855. 10.1016/j.ecolind.2025.113855.

Waser, L., Fischer, C., Wang, Z. and Ginzler, C. 2015 Wall-to-Wall Forest Mapping Based on Digital Surface Models from Image-Based Point Clouds and a NFI Forest Definition. *Forests* **6**, 4510–4528. 10.3390/f6124386.

Webb, C.O. 2000 Exploring the Phylogenetic Structure of Ecological Communities: An Example for Rain Forest Trees. *The American naturalist* **156**. 10.1086/303378.

Webb, C.O., Ackerly, D.D., McPeek, M.A. and Donoghue, M.J. 2002 Phylogenies and Community Ecology. *Annu. Rev. Ecol. Evol. Syst.* **33**, 475–505. 10.1146/annurev.ecolsys.33.010802.150448.

Webb, C.O., Losos, J.B. and Agrawal, A.A. 2006 Integrating Phylogenies into Community Ecology. *Ecology - Special Issue* **87**, S1-S2. 10.1890/0012-9658(2006)87[1:IPICE]2.0.CO;2.

Whittaker, R.H. 1960 Vegetation of the Siskiyou Mountains, Oregon and California. *Ecological Monographs* **40**, 279–338. 10.2307/1943563.

Wickett, N.J., Mirarab, S., Nguyen, N., Warnow, T., Carpenter, E. and Matasci, N. et al. 2014 Phylotranscriptomic analysis of the origin and early diversification of land plants. *Proceedings of the National Academy of Sciences of the United States of America* **111**, E4859-68. 10.1073/pnas.1323926111.

Wiebe, R.A. and Wilcove, D.S. 2025 Global biodiversity loss from outsourced deforestation. *Nature* **639**, 389–394. 10.1038/s41586-024-08569-5.

Wilcox, B.A. 1984 In situ conservation of genetic resources: determinants of minimum area requirements. *National parks: conservation and development*, 639–647.

Wilcox, T.M., Schwartz, M.K. and Lowe, W.H. 2018 Evolutionary Community Ecology: Time to Think Outside the (Taxonomic) Box. *Trends in Ecology & Evolution* **33**, 240–250. 10.1016/j.tree.2018.01.014.

Winter, S., Chirici, G., McRoberts, R.E., Hauk, E. and Tomppo, E. 2008 Possibilities for harmonizing national forest inventory data for use in forest biodiversity assessments. *Forestry* **81**, 33–44. 10.1093/forestry/cpm042.

Wohlgemuth, T. 2000 Diskreter und kontinuierlicher Charakter der Vegetation: Waldvegetationsdaten als Referenz. *Bauhinia* **14**, 67–88. zobodat.at/pdf/Bauhinia_14_0077-0088.pdf.

Woods, C.L. and Ortmann, K. 2024 Microtopographic heterogeneity affects habitat specialization and diversity of understory plants in a northern temperate rainforest. *Plant Ecology*. 10.1007/s11258-024-01469-8.

Yang, Z. and Rannala, B. 1997 Bayesian Phylogenetic Inference Using DNA Sequences: A Markov Chain Monte Carlo Method. *Molecular biology and evolution* **14**, 717–724. 10.1093/oxfordjournals.molbev.a025811.

Yao, J., Zhang, C., Cáceres, M. de, Legendre, P. and Zhao, X. 2019 Variation in compositional and structural components of community assemblage and its determinants. *J Vegetation Science* **30**, 257–268. 10.1111/jvs.12708.

Yu, Y., Saatchi, S., Domke, G.M., Walters, B., Woodall, C. and Ganguly, S. et al. 2022 Making the US national forest inventory spatially contiguous and temporally consistent. *Environ. Res. Lett.* **17**, 65002. 10.1088/1748-9326/ac6b47.

Zanne, A.E., Tank, D.C., Cornwell, W.K., Eastman, J.M., Smith, S.A. and FitzJohn, R.G. et al. 2014 Three keys to the radiation of angiosperms into freezing environments. *Nature* **506**, 89–92. 10.1038/nature12872.

Zarnetske, P.L., Read, Q.D., Record, S., Gaddis, K.D., Pau, S. and Hobi, M.L. et al. 2019 Towards connecting biodiversity and geodiversity across scales with satellite remote sensing. *Global Ecology and Biogeography* **28**, 548–556. 10.1111/geb.12887.

Zeballos, S.R., Acosta, A.T.R., Agüero, W.D., Ahumada, R.J., Almirón, M.G. and Argibay, D.S. et al. 2023 Vegetation types of the Arid Chaco in Central-Western Argentina. *VCS* **4**, 167–188. 10.3897/VCS.100532.

Zeballos, S.R., Giorgis, M.A., Cabido, M.R., Acosta, A.T., Del Iglesias, M.R. and Cantero, J.J. 2020 The lowland seasonally dry subtropical forests in central Argentina: vegetation types and a call for conservation. *VCS* **1**, 87–102. 10.3897/VCS/2020/38013.

Zellweger, F., Braunisch, V., Morsdorf, F., Baltensweiler, A., Abegg, M. and Roth, T. et al. 2015 Disentangling the effects of climate, topography, soil and vegetation on stand-scale species richness in temperate forests. *Forest Ecology and Management* **349**, 36–44. 10.1016/j.foreco.2015.04.008.

Zemlerová, V., Kozák, D., Mikoláš, M., Svitok, M., Bače, R. and Smyčková, M. et al. 2023 Natural Disturbances are Essential Determinants of Tree-Related Microhabitat Availability in Temperate Forests. *Ecosystems* **26**, 1260–1274. 10.1007/s10021-023-00830-8.

Zhang, J., Swenson, N.G., Liu, J., Liu, M., Qiao, X. and Jiang, M. 2020 A phylogenetic and trait-based analysis of community assembly in a subtropical forest in central China. *Ecology and evolution* **10**, 8091–8104. 10.1002/ece3.6465.

Zhao, X.H., Corral-Rivas, J., Zhang, C.Y., Temesgen, H. and Gadow, K. von. 2014 Forest observational studies-an essential infrastructure for sustainable use of natural resources. *Forest ecosystems* **1**, 1–10. 10.1186/2197-5620-1-8. forestecosyst.com/content/1/1/8.

Zheng, J., Arif, M., He, X., Ding, D., Zhang, S., Ni, X. and Li, C. 2022 Plant community assembly is jointly shaped by environmental and dispersal filtering along elevation gradients in a semiarid area, China. *Frontiers in plant science* **13**, 1041742. 10.3389/fpls.2022.1041742.

Zizka, V.M., Geiger, M.F. and Leese, F. 2020 DNA metabarcoding of stream invertebrates reveals spatio-temporal variation but consistent status class assessments in a natural and urban river. *Ecological Indicators* **115**, 106383. 10.1016/j.ecolind.2020.106383.

Zolotova, E., Ivanova, N. and Ivanova, S. 2023 Global Overview of Modern Research Based on Ellenberg Indicator Values. *Diversity* **15**, 14. 10.3390/d15010014.

2 Phylogenetically Informed Forest Classification

Classifying forest vegetation types and discriminating species assemblages is central to understanding biodiversity patterns, informing conservation strategies, and supporting ecological monitoring. In this study, I tested whether a phylogenetically informed dissimilarity metric could improve forest community classifications based on the GNFI data. Specifically, I applied the unsupervised clustering algorithm ISOPAM to classify forest cluster-plot (each consisting of three subplots, Chapter 1.4.1.1) sample data, using two dissimilarity metrics: Bray–Curtis index as a species-neutral, and the discriminating Avalanche dissimilarity index (dA), which incorporates interspecific phylogenetic distances (Chapter 1.2). The objective was to evaluate the methodological implications of integrating PD in dissimilarity quantification for hierarchical clustering and to assess how the resulting classifications differ between approaches.

Accordingly, I integrated the GNFI raw data into a relational data model in Microsoft Power BI (Chapter 1.6) to facilitate data cleansing and filtering, perform exploratory data analyses, and to prepare the data by aggregating and extrapolating BA per species and sample. In parallel, I harmonized all recorded species names by applying the package WORLDFLORA (Kindt, 2020) within an R workspace of the processing environment of R Studio (RStudio Team, 2024). Using harmonized species names, I created a community matrix based on BA values per species and sample and migrated it into the R workspace. Here, I generated a phylogeny from the recorded species list, using the V.PHLOMAKER package and computed an interspecies dissimilarity matrix based on the normalized cophenetic distances between each pair of species (Chapter 2.2.3.2, Jin & Qian, 2019). I then calculated two sample-wise dissimilarity matrices: one using Bray–Curtis dissimilarities (via the R package VEGAN, Oksanen, 2020), and one using a custom function based on the equation of dA described by Hao et al. (2019), which was supported by Nils Griese (University of Göttingen). I visualized each dissimilarity matrix as heat map in R (Figure 2.4). Integration of dA into ISOPAM and optimization of the algorithm to my dataset was followed the advice and technical support by the ISOPAM developer Sebastian Schmidtlein (Karlsruhe Institute of Technology, Schmidtlein et al., 2024). Subsequently, I ran ISOPAM based on each dissimilarity matrix separately and exported the resulting cluster characteristics to Power BI for characterization, visualization and initial pattern assessment (Chapter 1.6). Supported by the co-authors, I interpreted the ecological significance of the resulting dissimilarity patterns and stratifications. Using metrics such as Within-Sum-of-Squares, the modified Rand index and average intra-cluster dissimilarity implemented of the R packages MCLUST and FPC, I evaluated cluster coherence of the resulting hierarchies (Scrucca et al., 2016; Hennig, 2023). To assess how accurately each dendrogram captured sample differences, I calculated cophenetic correlation coefficients (Chapter 2.2.3.5). To assess spatial distribution, I mapped clustered sample locations in QGIS (QGIS Development Team, 2009) and compared theses against potential natural vegetation units from Bohn et al.

(2007). Finally, I validated the ecological relevance of the results by visualizing the distribution of elevation values across clusters and applied the non parametric Kruskal–Wallis test for significance.

Subsequently, I prepared all visualizations and optimized them for publication using the vector-based image editing tool INKSCAPE (Yuan et al., 2016). I was responsible for the full preparation, processing, and analysis of the data used in this study. This included harmonizing species names, constructing the community and phylogenetic matrices, calculating dissimilarities, adjusting and running the clustering algorithm, and evaluating and visualizing the results. I also produced all statistical analyses, maps, and visual materials (also see supplemental material of Wellenbeck et al., 2024 available at https://onlinelibrary.wiley.com/doi/full/10.1002/ece3.11569). Finally, I drafted and revised the manuscript, led the peer-review process, and served as corresponding author during submission and revision (see author contributions in the annexed article).

Within the broader context of this thesis, this study establishes a central methodological foundation. It demonstrates that incorporating PD into dissimilarity calculations leads to ecologically improved classifications that align with environmental gradients and biogeographic expectations. These findings provide empirical support for the overarching hypothesis that phylogenetic informed classification enhances the stratification of forest communities, beyond what is possible using species identity alone. The approach developed and applied here directly informs the predictive modeling approach in Publication II (Chapter 3) and highlights the practical potential of phylogenetically informed classification for large-scale forest biodiversity monitoring.

Discriminating Woody Species Assemblages from National Forest Inventory Data based on Phylogeny in Georgia

Abstract

Classifications of forest vegetation types and characterization of related species assemblages are important analytical tools for mapping and diversity monitoring of forest communities. The discrimination of forest communities is often based on β-diversity, which can be quantified via numerous indices to derive compositional dissimilarity between samples. This study aims to test the applicability of an unsupervised classification for National Forest Inventory data by comparing two clustering results based on species-neutral dissimilarity (Bray-Curtis) and the discriminating Avalanche dissimilarity which incorporates interspecies phylogenetic variation. We calculated the mean basal area per hectare and species for 1059 forest plot observations and generated a phylogenetic tree for all observed species. We applied an unsupervised cluster analysis based on both dissimilarities and compared the results regarding performance and statistics and alignment of cluster distributions with biogeographic units. Incorporating genetic variation in the quantification of dissimilarities provides a more nuanced discrimination of woody species assemblages, and results in more coherent clustering result. Favorable clustering statistics include the total number of clusters (23 vs. 20), mean distance within clusters (0.773 vs. 0.343) and Within-Sum-of-Squares (344.13 vs. 112.92). Clusters derived from dissimilarities that account for genetic variation showed a more robust alignment with biogeographic gradients, such as elevation and known habitats. We demonstrate that the applicability of unsupervised classification of species assemblages to large-scale forest inventory data strongly depends on the underlying quantification of dissimilarity. Our results indicate that by incorporating phylogenetic variation, a more precise classification aligned with biogeographic units is attained. This supports the concept that the genetic signal of species assemblages reflects biogeographic patterns and facilitates more precise analyses for mapping, monitoring and management of forest diversity.

Published Article:

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. Ecology and Evolution, 14(7), https://doi.org/10.1002/ece3.11569.

2.1 Introduction

Forests ecosystems host the largest share of terrestrial biodiversity and cover approximately one third of the global land surface (Torresani et al., 2019; FAO, 2020; Gillerot et al., 2021; Heym et al., 2021). In light of increasing pressure on forests due to climate change and the related global loss of biodiversity, also referred to as the “the sixth mass extinction”, with up to 92% of terrestrial endemics being anticipated to be negatively impacted (Manes et al., 2021), reliable approaches to assess and monitor forest diversity are required (Barnosky et al., 2011; Faith, 2013; Palombo, 2021; Cowie et al., 2022). Monitoring should include the quantification of metrics that allow the classification of ecological entities based on their specific level, or degree of biodiversity, and ultimately according to their respective conservation value, which is required by conservationists (Brooks et al., 2015; Zampiglia et al., 2019). Appropriate delineation of forest communities and characterization of related species assemblages across taxonomic groups are important analytical tools for sensible monitoring of species diversity and biodiversity management. On broader scales, such metrics can be provided by means of vegetation classification which aims to group spatial or temporal diversity of species within a finite set of abstract categories (Cáceres et al., 2015). Vegetation classification has proven to provide adequate means for descriptive reporting, communication, and mapping of forest communities and related concepts have responded to changing information needs over time. Consequently, forest type classifications exist for a wide range of targets, i.e., habitat quality, (qualitative assessments for biodiversity management), development over time (i.e., stand classification according to age-classes for forest management) or along biogeographic gradients, and remote sensing - based mapping of ecological communities (Cáceres et al., 2013; Fassnacht et al., 2016; Cáceres et al., 2019b; Hao et al., 2021).

One approach for classifying forest communities focuses on the variation in species compositional characteristics across assemblages of different sites within a geographic area, which is commonly known as β-diversity (Magurran & McGill, 2011; Legendre & Cáceres, 2013). β-diversity can be assessed by the change in species compositional characteristics between sites (i.e., species turnover, Jost, 2010) and a plethora of metrics exist to quantify the degree of dissimilarity between assemblages on various spatial and temporal scales (Ricotta, 2005; Magurran & McGill, 2011; Legendre & Legendre, 2012; Cáceres et al., 2013; Cáceres et al., 2019a). The most common dissimilarity indices are exclusively based on compositional characteristics, i.e., species richness and elements of evenness (Magurran, 2005), while interspecies variability (i.e., phylogenetic, taxonomic, functional, or traits) is not considered (Pavoine et al., 2013; Cáceres et al., 2013; Chiu et al., 2014; Chao et al., 2018; Hao et al., 2019b). In line with the increasing recognition that genetic diversity comprises an integral part of biodiversity, e.g., as stated in the definition of biodiversity by the Intergovernmental Platform on Biodiversity and Ecosystem Services (IPBES, Díaz et al., 2015), literature on how to incorporate phylogenetics as aspect of diversity is growing rapidly

(Chao et al., 2023). Accordingly, several authors have approached forest community classification by accounting for both, compositional data and interspecies phylogenetic variability (i.e., Webb et al., 2002, Capelo, 2020; Hao et al., 2021, Ricotta et al., 2020). As phylogenetically closely related species often share beneficial traits for specific environments, discriminating assemblages based on phylogenetic distances can serve as a proxy for classifying forest communities according to functional roles, environmental diversity, and conservation value (Faith, 2013; Hawkins et al., 2014; Pavoine & Ricotta, 2014; Pavoine, 2016; Tucker et al., 2017; Padullés Cubino et al., 2021; Gilbert & Parker, 2022). Hao *et al.*, (2019b) demonstrated that different patterns emerged if interspecies taxonomic distances are considered for the classification of global forest communities using the discriminating Avalanche index (Ganeshaiah & Shaankar, 2000).

On a smaller scale, National Forest Inventories (NFIs) provide systematic and periodical observations of tree species abundances based on permanent sample units on a country-wide level (Corona et al., 2011). The continuous adaptation of variables assessed during NFIs highlights an increasing emphasis on aspects of biodiversity, enabling ecologists to investigate potentials and limitations of the thus provided data (McRoberts et al., 2009; Corona et al., 2011; Alberdi et al., 2019; Didion, 2020). Incorporating PD of species assemblages extends the perspective on diversity in this context and bears the potential to deepen our understanding of the complex interactions among woody species over large geographical scales.

In the present study, we compare the performance of two dissimilarity indices for the discrimination of forest woody species assemblages when applied to large-scale forest inventory data such as the dataset of the first NFI of Georgia. To this end, we applied unsupervised clustering to the obtained dissimilarity matrices based on a conventional and a dissimilarity index that incorporates interspecies phylogenetic distances, respectively. Apart from statistics for internal evaluation of the resulting classifications, our comparison considered the distribution of discriminated assemblages along biogeographic gradients. Based on the assumption that genetic variability of co-inhabiting species provides a signal that sufficiently reflects site-specific environmental determinants, we investigated whether including this variable to the quantification of dissimilarity results in an improved reflection of biogeographic gradients. To test the general applicability, we incorporated the phylogenetic variability in the classification of a large, real-world dataset and evaluated the results considering cluster cohesiveness and overall interpretability.

2.2 Data and Methods

We compare two dissimilarity indices for the classification of woody species assemblages when applied to NFI data of Georgia. Adhering to the methodological approach underlying the data, we focus on woody species, i.e., all recorded species that meet the specified target diameter at breast height (DBH, at 1.3 m, MEPA, 2018). Consequently, we are referring to woody species even if only species is written hereafter.

2.2.1 Study Area

Georgia is located between the Southern Slopes of the Greater and the Northern part of the Lesser Caucasus, between 41°07' – 43°35'N and 40°04' – 46°44'E (Fischer et al., 2018). The forests of Georgia host large shares of endemic species and form part of the Caucasus biodiversity hotspot (Myers, 2003; Joppa et al., 2011). Existing forest formations range from Alpine coniferous forests dominated by *Abies nordmanniana* (Steven) Spach. and *Picea orientalis* (L.) Peterm. to open juniper woodland (dominated by *Juniperus polycarpos excelsa* subsp. *polycarpos* (K. Koch) Takht. and *J. foetidissima* Willd.), encompassing further thermophilus to xerophytic mixed oak forest (*Quercus petrea* subsp. *iberica* (Steven ex M. Bieb.) Krassiln., *Carpinus betulus* L., *C. orientalis* Mill.), Colchic alder carrs which are dominated by *Alnus glutinosa* subsp. *barbata* (C. A. Mey.) Yalt. and oriental beech (*Fagus orientalis* Lipsky) and hornbeam-oriental beech forests (Bohn et al., 2007; Dolukhanov, 2010; Fischer et al., 2018; Nakhutsrishvili et al., 2021; Novák et al., 2023)

2.2.2 Forest Community Data

Between 2018 and 2021, Georgia implemented its' first NFI based on a systematic sampling grid of 3.6 km x 3.6 km with a randomly selected origin. Sampling units consist of cluster plots (0.21 ha) containing three subplots of 0.0.7 ha each. These subplots are arranged in an L-shaped configuration with a distance of 100 m along both axes (Figure 2.1). As 18% of the country's territory is currently not accessible for government officials due to an ongoing political conflict (MEPA, 2023), clusters were sampled on approximately 74% of the national forest area (Figure 2.2).

2.2.2.1 Assessment of Woody Species

Woody species were recorded per subplot on three concentric nested subplots according to any stems' respective DBH. Stems with DBH ≥30 cm were recorded on the largest plot (r = 15 m). Stems with DBH ≥15 cm and DBH ≥8 cm were recorded on the inner nested plot radii of 10 m radius and 5 m, respectively (Figure 2.1). Numerous variables were recorded per woody species, along with the polar coordinates of the stem axis, species and DBH (MEPA, 2018).

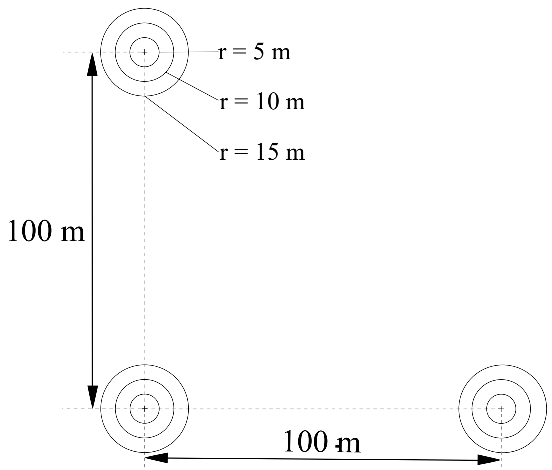


Figure 2.1: Cluster plot configuration of the first National Forest Inventory of Georgia. Each cluster consists of three subplots for tree assessment within three concentric circles according to the measured diameter at breast height (DBH, at 1.3 m, MEPA, 2023).

From the entire NFI dataset (N = 2006), all accessible clusters pertaining to the locally applied land-use class “Forest” and sub-classes “Tree covered area” or “Fire affected forests” were selected for analysis (MEPA, 2018). Subplots with recorded intersection with a forest boundary (“slopover sample plots”) were excluded to avoid including samples with extreme outliers regarding species richness due to edge effects (Willmer et al., 2022). Clusters containing species observations that were not unambiguously identified at species level, e.g., *Deciduous spp.*, *Genera spp.* were omitted, because a precise quantification of the cophenetic distances along the phylogenetic hierarchy is not possible. Clusters containing subplots without woody species observations were excluded. Our sample consists of all cluster plots comprised of observations from three subplots (n = 1059, henceforward referred to as “samples”), which represent 53% of all clusters. Figure 2.2 provides an overview of the spatial distribution of samples used for the analysis.

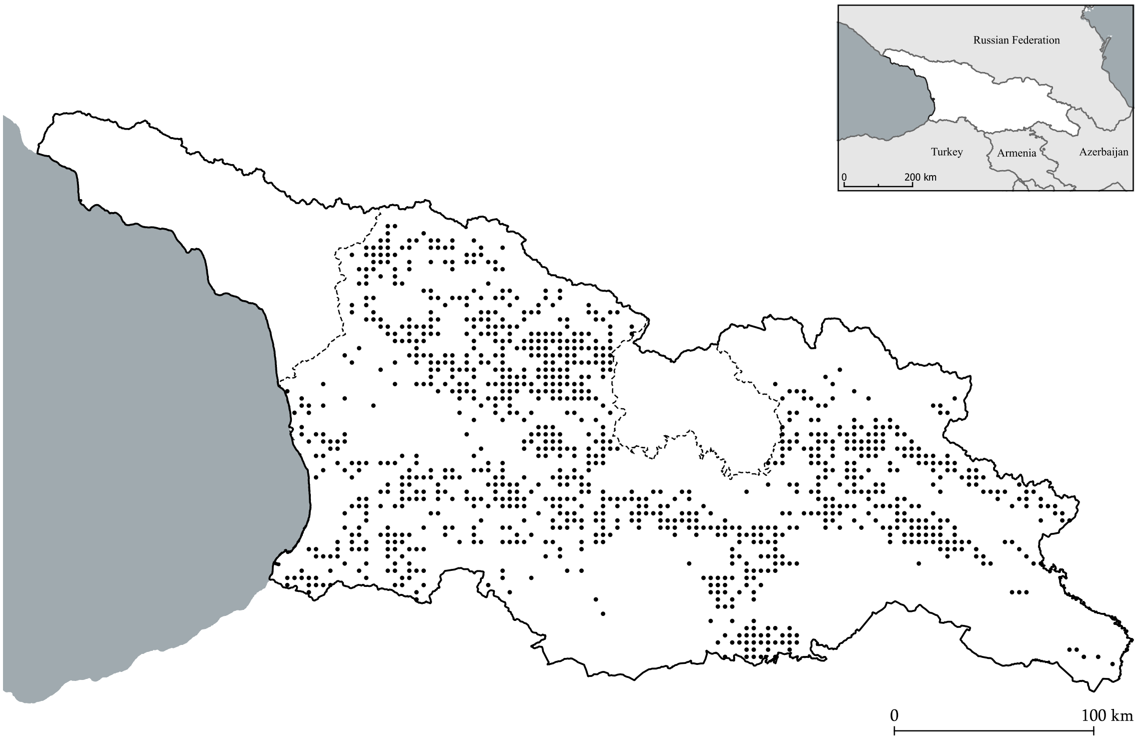


Figure 2.2: Locations of samples used from the National Forest Inventory of the Georgia, n = 1059. Samples (black dots) consist of accessible cluster plots of equal sample size (three subplots) located inside forests, which are not intersected by forest boundaries and contain only records of taxa identified at species level. Dashed lines mark the boundaries of inaccessible areas from where no field data was obtained.

After reprojection of sample locations to UTM38N, WGS84 (EPSG: 32638), sample elevations [m above sea level] were derived from the digital elevation model (DEM) provided by Shuttle Radar Topography Mission (SRTM, Farr et al., 2007). Elevation values were calculated as the median of all raster cell values (30 m x 30 m) contained in or crossed by the circular subplot area (r = 15 m + recorded GPS error [m]) of the Southwestern subplot.

2.2.2.2 Species Diversity Data

Diversity measures are usually based on data representing the compositional variation between species (i.e., occurrence and abundance) in an assemblage (Ricotta et al., 2021). Forest species communities may display similar compositional characteristics, in terms of counts of observed species and respective individuals. However, species can be represented by large numbers of small diameter trees, or stems belonging to the same individual, or fewer individuals but with significant large relative shares of BA. Hence, abundance estimates based on counts of individuals do not take significant differences in the size structure of occurring species into consideration and may result in distinct evenness profiles. We used mean basal area (BA, m^2 ha^{-1}) per species and cluster plot as species abundances to account for the variation in size of the constituents. By weighting compositional data using BA, we incorporate valuable structural information that considers site occupation per species for the quantification of beta diversity (Staudhammer & LeMay, 2001;

McRoberts et al., 2009; Yao et al., 2019; Cáceres et al., 2019a). Consequently, BA of all living stems was aggregated per cluster plot and species and divided by three to obtain mean BA estimates per sample (Cáceres et al., 2015; MEPA, 2018).

2.2.3 Analysis

2.2.3.1 Nomenclature

Spelling and nomenclature of all recorded species were standardized with the Taxonomic Backbone databases of World Flora Online (WFO DB, Kindt, 2020) and the Global Biodiversity Information Facility (GBIF Secretariat, 2021). In cases where species listed in the NFI data did not yield an unequivocal match in the WFO DB, corresponding records were harmonized with Lachashvili et al. (2022) and the nomenclature of the World Plants database (https://www.worldplants.de) to derive names for all species ranked as taxonomically accepted.

2.2.3.2 Phylogenetic Interspecies Distances

A phylogenetic tree encompassing all observed species was constructed by matching the harmonized species list with a mega phylogeny containing 72,570 species of vascular plants according to the World Plants database (GBOTB.extended.WP.tre, Jin & Qian, 2022). The backbone mega phylogeny is based on the species-level phylogeny for vascular plants derived from gene sequencing from seven gene regions and 39 fossil calibrations created by Zanne et al. (2014), that was updated and expanded by Jin & Qian (2022). Following the authors' recommendation to consolidate taxa below species level (e.g., sub-species) with the respective parental species, five infraspecific taxa were combined with their parental species, resulting in the lowest taxonomic unit being species level (Figure 2.3). From the thus created ultrametric phylogenetic tree (Qian & Jin, 2016; Smith & Brown, 2018; Jin & Qian, 2019), cophenetic distances, i.e., the total branch length connecting each pair of species at the terminal nodes of the respective phylogeny were calculated (Kling et al., 2018; Hao et al., 2019a; Bevilacqua et al., 2021).

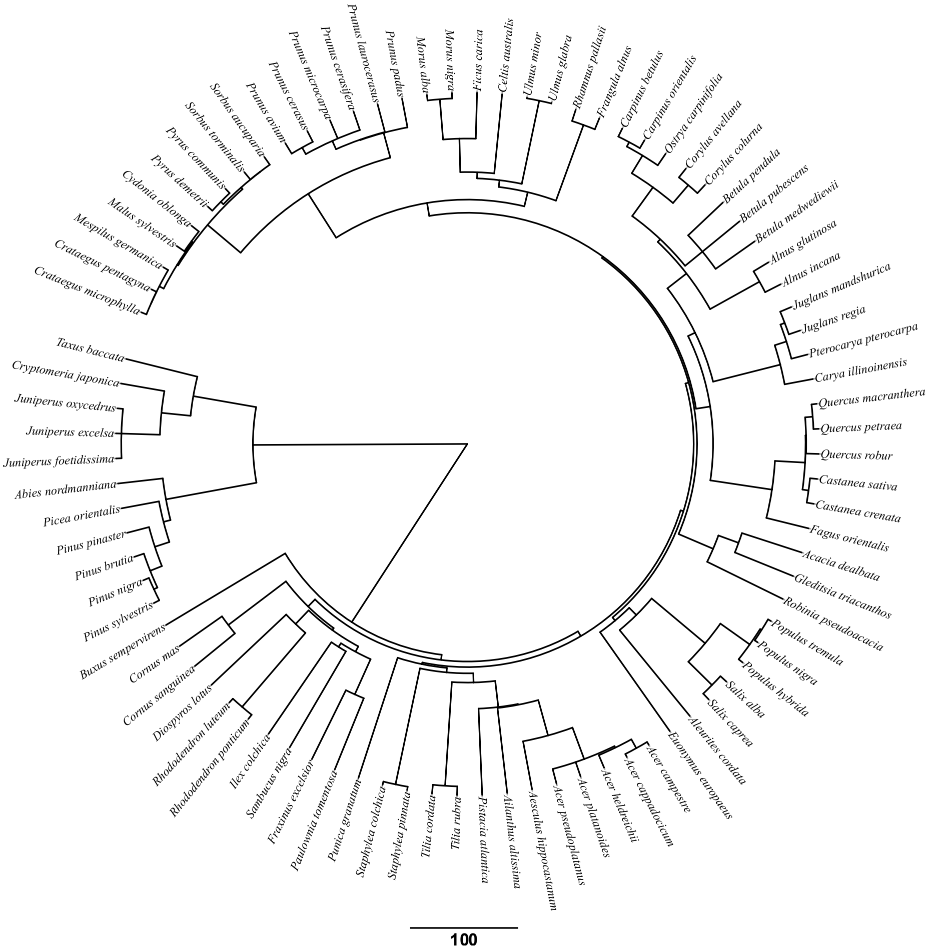


Figure 2.3: Phylogeny for 87 species listed in the data sample of the National Forest Inventory of Georgia. Interspecies phylogenetic distances were calculated as total branch length connecting each pair of species at the terminal nodes of the hierarchy. For respective branch lengths see Figure S1.

Interspecies phylogenetic distances were normalized within a square matrix that contained pairwise distance values [$0 < \text{dist}_{ph} \leq 1$] between each pair of species.

2.2.3.3 Dissimilarity Indices

The discriminating Avalanche index (dA eq. [2] in Table 2.1) developed by Ganeshaiah & Shaankar (2000) and described by Hao et al. (2019a) considers interspecies dissimilarity by multiplying absolute differences in frequencies (numerator in BC) of species *i* and *j* in two samples with a specific distance between species *i* and *j*. We use the phylogenetic distances to weigh the difference in mean basal area between *i* and *j*. Table 2.1 shows the formulas of both indices used in this study.

Table 2.1: Dissimilarity indices used in this study.

Bray-Curtis Dissimilarity Index (1957)	BC	=	$\frac{\sum_{i=1}^{n}\lvert p_i^a - p_i^b\rvert}{\sum_{i=1}^{n}(p_i^a + p_i^b)}$	[1]
Discriminating Avalanche (Hao et al., 2019a)	dA	=	$\frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}\Delta_i^{a,b} d_{ij} \Delta_j^{a,b}$	[2]

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 = total number of sample plots

p_i^a, p_i^b = relative frequencies of species i in plots a and b

As the maximum dissimilarity value obtained by dA equals $\left(1 - \frac{1}{n}\right)$, where n equals the number of species, the resulting dissimilarities were normalized via $x_{norm} = \frac{x - x_{min}}{x_{max} - x_{min}}$, with $x_{\min}$ and $x_{\max}$ representing the minimum and maximum value of dA, respectively (Legendre & Legendre, 2012; Hao et al., 2019a). Consequently, pairwise dissimilarity values of the two resulting dissimilarity matrices (1059 x 1059) range between 0 and 1. Pairwise values of 1 imply that two samples are completely different as they do not share any species, whereas values of 0 indicate two samples are equal in terms of compositional characteristics (Chao et al., 2005; Leyer & Wesche, 2008; Legendre & Legendre, 2012). Prior to clustering, a Mantel test (Mantel, 1967) was performed to check for existing correlations between the two dissimilarity matrices. As the pairwise dissimilarities are not normally distributed, and non-linear relationships between the pairwise dissimilarity values exist, we used the Spearman correlation coefficient with 9999 permutations (Legendre & Legendre, 2012).

2.2.3.4 Isometric Partitioning

The ISOPAM algorithm (Schmidtlein et al., 2010) available in package 'isopam' (v. 2.0, Schmidtlein et al., 2024) combines Isometric feature mapping and Partitioning around medoids (data points that are most centrally located within each cluster with the sum of dissimilarities between medoids and all other data points being minimized) in order to build clusters with a maximum number and fidelity of indicative species. The isomap ordination, which is based on geodesic distances strongly determined by neighborhood definitions, is repeated with different parameter settings. The result is clustered, and the clusters are evaluated according to the criteria mentioned above. These criteria are similar to those used when structuring phytosociological tables (Abe, 2021). In this, and in the use of an ordination, ISOPAM is similar to TWINSPAN (Hill, 1979), but does not involve internal readjustments, uses geodesic distances (taking account of "neighbors of neighbors" in feature space) and works on multidimensional ordination spaces. It has been previously used for large scale classifications of forests (Černý et al., 2015; Cabido et al., 2018; Zeballos et al., 2020) and other systems (Hein et al., 2014; Peterka et al., 2017; Feilhauer et al., 2021; Yu et al., 2022). ISOPAM can be run both unsupervised and supervised (with reference plots). For the current study, the original source code was extended to support dA (Capelo, 2020) and executed on a computer with two Intel Xeon CPU's (E5-2630 v3) and 256 GB RAM using R Statistical Software (v 4.3.2; R Core Team 2023). To ease comparability, we set the maximum number of hierarchy levels to four for both classifications. Subsequently, we extracted lists of indicator species frequencies with levels of significance according to Fisher`s exact test for each cluster using the ISOTAB function, which is part of the ‘isopam’ package. Fidelity ("equalized phi", Tichý & Chitrý 2006) together with Fisher's exact test if the observed frequency is not attained by chance, are the criteria for qualifying as an indicator species in ISOTAB (Schmidtlein et al., 2024).

2.2.3.5 Evaluation of Clustering

To evaluate the correspondence between the original sample dissimilarities and dendrogram distances we calculated cophenetic correlation coefficients for each hierarchical cluster structure (Lapointe & Legendre, 1995; Legendre & Legendre, 2012). The modified Rand index was used to evaluate clustering performance based on the consistency between partitions (Legendre & Legendre, 2012). Cluster homogeneity was evaluated via Within-Sum of Squares (WSS, Hao et al., 2019b) and a comparison of the average distance between and within clusters using the function CLUSTER.STATS of the R package ‘fpc’. To assess relevance of the hierarchies we compared indicator species and the resulting, distributions of relative BA among partitioned groups (Cáceres et al., 2015). Evaluation of correspondence to biogeographic units was based on a comparison of elevational ranges derived from the DEM between groups and the spatial distribution of clusters in relation to forest vegetation type classifications presented by Bohn et al. (2007). We applied the non-parametric Kruskal-Wallis test to check for significance between groups, due to non-normality in the distribution of elevation values within groups (Shapiro test). Henceforward, we are referring to the initial partition at the lowest level of the hierarchy as classes, to the intermediary partitions as branches and to the resulting clusters as assemblages.

2.2.3.6 Data Analysis

The analysis was conducted in R Studio version 2023.09.0-463 (RStudio Team, 2024) using R Base version 4.3.0 (R Core Team, 2024). Harmonization of nomenclature was realized via the R package ‘Worldflora’ version 1.13-2 (Kindt, 2020) and the package ‘V.phylomaker2’ version 0.1.0 was used to match observed species with the phylogenetic backbone (Jin & Qian, 2022). The Mantel test and BC dissimilarities were calculated using the packages ‘vegan’ version 2.6-4 (Oksanen et al., 2022). A custom function was embedded in the adjusted clustering algorithm of the corresponding R package ’isopam’ version 1.2.0 for dA (Schmidtlein et al., 2022). Clustering metrics were obtained using the R packages ‘mclust’, version 6.0.0 (Scrucca et al., 2016) and ‘fpc’, version 2.2-10 (Hennig, 2023).

2.3 Results

Compositional data of n=1,059 samples containing 65,818 living tree observations were analyzed (Table 2.2). In total, 87 species were represented by 52 genera, 29 families, 16 orders and 2 classes.

Table 2.2: Summary statistics of compositional data of cluster plots.

	Species richness				Mean basal area per species [m^2 ha^{-1}]			
n	**Min**	**Max**	**Mean***	**CV%**	**Min**	**Max**	**Mean***	**CV%**
1059	1	12	4.96 (±2.14)	43.08	2.54	79.02	30.4 (±12.79)	42.06

with
Var. = Variance
CV% = Coefficient of variation
*Means are denoted with standard deviation in parenthesis

Angiosperms were represented by 76 species (in 87.4% of all samples) across 26 different families. *Fagaceae* (six species) accounted for the highest number of observations followed by *Betulaceae* with ten species (25.3% and 23.2% of all samples, respectively), *Rosaceae* and *Sapindaceae* (15 and six species, in 11.4% and 10.3% of all samples, respectively). In contrast, gymnosperms (10.3%), were represented by 11 species, belonging to 3 different families, with the largest family being *Pinaceae* (6 species) followed by *Cupressaceae* and *Taxacea*e (four and one species, respectively). Gymnosperms were observed in 33% of all samples.

2.3.1 Pairwise Dissimilarity

Mantel statistics ($r = 0.613$ with $p = 0.0004$) indicated a significant positive correlation between the two dissimilarity matrices. Pairwise dissimilarities according to dA ranged from 0.23 to 0.87, which is almost double the range of that of BC (0.63 to 0.99, Table 2.3).

Table 2.3: Summary statistics of mean pairwise dissimilarities based on Bray-Curtis and discriminating Avalanche.

Index	n	Min	Max	Mean*	Var.	CV%
BC	1059	0.627	0.998	0.789 (±0.099)	0.01	12.55
dA		0.225	0.87	0.383 (±0.147)	0.022	38.38

with
Var. = Variance
CV% = Coefficient of variation
*Means are denoted with standard deviation in parenthesis

Consequently, the mean dissimilarity between samples (0.79 for BC and 0.38 for dA, respectively) and thus overall variation was higher for BC than for dA, as for the latter frequencies of dissimilarity values <1 were more evenly distributed, with very few pairwise dissimilarities of 1. Frequency distributions of sample dissimilarities are provided in Figure S2.

2.3.2 Discrimination of Assemblages

The hierarchical clustering based on BC (HC_{BC}) distinguished 23 assemblages (clusters) across four hierarchical levels (I - IV) and classes. Within HC_{BC}, samples were partitioned into ten and 17 branches at levels II to III, respectively. The hierarchical clustering based on dA (HC_{dA}) contained ten and 15 branches at level II to III and resulted in 20 distinct assemblages over four levels and classes. Two assemblages were not partitioned further below level II in HC_{dA} (Figure 2.4). The number of samples per assemblage ranged from six to 152 for HC_{BC} (Mean: 82±91.5) and three to 163 for HC_{dA} (Mean: 106.1±119.9), respectively. For HC_{BC}, the number of observed species per assemblage ranged from 12 to 54. For HC_{dA}, one assemblage contained only four species, whereas two assemblages encompassed 53 species.

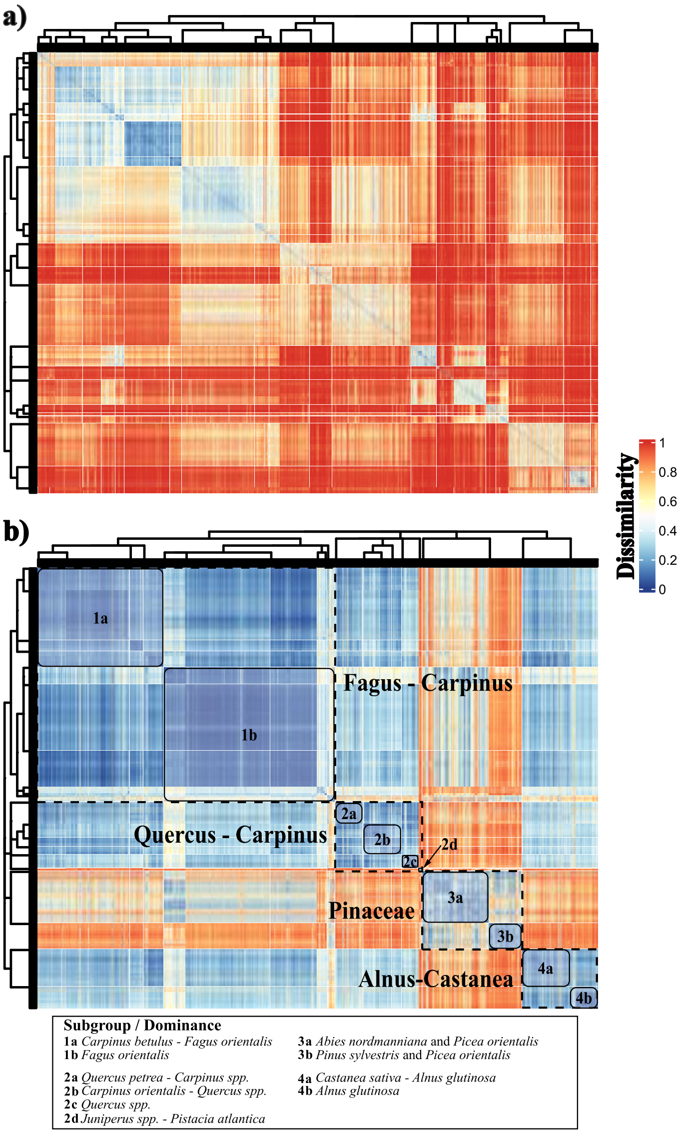


Figure 2.4: Resulting cluster hierarchy from the isopam partitioning (dendrogram) and pairwise dissimilarities of 1059 samples from the National Forest Inventory data of Georgia. Dissimilarities are based on the Bray-Curtis (a) and discriminating Avalanche (b) index. The cell grid is colored according to the dissimilarity values between samples (rows and columns). In total, 23 and 20 assemblages were discriminated for Bray-Curtis and discriminating Avalanche, respectively. At level II of the hierarchical clustering based on the discriminating Avalanche index, ten subgroups distributed over four classes were labeled according to dominance of relative basal area.

The cophenetic correlation coefficients were 0.492 and 0.442 for HC_{BC} and HC_{dA}, respectively. The obtained adjusted Rand index of 0.317 suggests a modest level of similarity between the clustering results. Average distances within clusters ranged from 0.671 to 0.875 and 0.196 to 0.49

for HC_{BC} and HC_{dA}, respectively. Mean distance between clusters was 0.789 for HC_{BC} and 0.384 for HC_{dA}. WSS values of HC_{BC} amounted to 344.13 and 112.92 for HC_{dA}.

2.3.3 Evaluation of Cluster Hierarchies

To evaluate the performance of BC and dA for clustering, we compared the resulting classes, groups and assemblages considering internal metrics of the partitioning process. 77 and 68 indicator species were listed for all partition based on BC and dA, respectively. For a characterization based on indicators and respective frequencies, only highly significant ($p \leq 0.001$) species with total frequencies ≥50% were considered, unless indicated otherwise. Indicator species for both partitions were *A. nordmanniana*, *A. cappadocicum*, *A glutinosa*, *C. betulus*, *Carpinus orientalis* Mill., *C. sativa*, *F. orientalis*, *P. orientalis*, *Q. petraea* subsp. *polycarpa* (Schur) Raus, and *Tilia rubra* subsp. *caucasica* (Rupr.) V. Engl. In addition, *Fraxinus excelsior* L., is an indicator for HC_{dA}. The total number of indicators with frequencies of 100% was 24 and 8 for HC_{BC} and HC_{dA}, respectively. Class one of HC_{BC} is characterized by a high frequency of *F. orientalis* (99%), whereas for HC_{dA}, *F. orientalis* and *C. betulus* are listed with frequencies of 93% and 75% respectively. In class two, the highest frequencies are observed for *Q. petrea* (87%) and *C. betulus* (82%). *P. orientalis* (68%), *F. orientalis* (57%) and *A. nordmanniana* (53%) are the most frequent indicator species in class three, whereas for class four, these are *A. glutinosa* (87%) and *Castanea sativa* Mill. (61%). In HC_{dA}, highly significant indicators in class one are *F. orientalis* (93%) and *C. betulus* (75%), whereas in class two these are *Q. petrea* (84%), *C. orientalis* (61%) and *Fraxinus excelsior* L. (50%). In class three and four of HC_{dA}, *P. orientalis* (79%) and *A. nordmanniana* (64%) and *A. glutinosa* (92%) and *C. sativa* (59%) represent highly significant indicators in class three and four, respectively. Based on these characteristics we labeled the four main classes according to predominant relative BA and are referring to these for ease of readability henceforward as follows: class one is characterized by a dominance of *Fagus*, class two is *Carpinus*-*Quercus* dominated, and class three and four are *Pinaceae* and *Alnus*-*Castanea* dominated, respectively. Synoptic tables of both cluster hierarchies (Figures S3 and S4, respectively) and a detailed description of indicator distributions per partition below level I for HC_{dA} (document S5) are provided as appendices.

2.3.4 Elevation and Spatial Distribution

To evaluate the correspondence of assemblages to existing biogeographic units, we compared the distribution of sample elevations within assemblages. Within the four classes in both hierarchies, sample elevations are distributed similarly. Samples within the *Fagus* dominated group (class one) cover a wide elevational range (≤750 m – >2000 m a.s.l.), however in HC_{BC} 81% of all samples are located between >1000 and 2000 m a.s.l., whereas in HC_{dA} most samples (77%) are located within the lower range of >750 and 1750 m a.s.l. Samples with *Carpinus*-*Quercus* dominance are predominantly located at elevations <1250 m in both hierarchies. The majority of *Pinaceae* dominated samples lie above 1250 m a.s.l., whereas all of *Alnus-Castanea* dominated samples are situated below 1250 m a.s.l., with the majority (70%) positioned below 750 m a.s.l. for HC_{BC} and HC_{dA}, respectively. Over all classes, except for the *Fagus* dominated group, agglomerations of sample elevation values are more pronounced for within assemblages of HC_{dA} than of HC_{BC}. The applied Kruskal-Wallis test revealed highly significant differences ($p \leq 0.001$) between assemblages for both hierarchies (Figure 2.5).

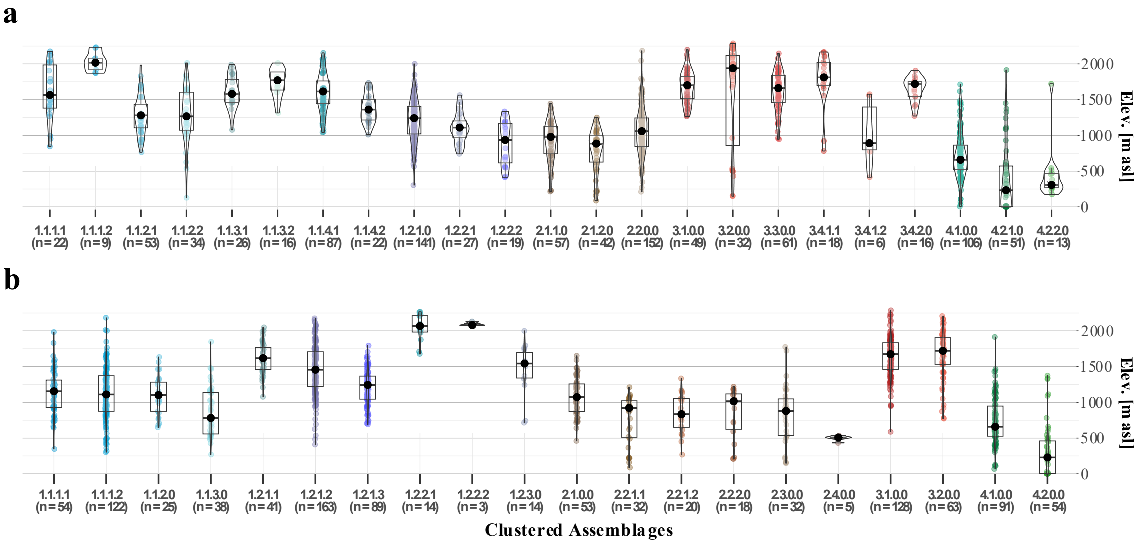


Figure 2.5: Sample elevations [m a.s.l.] from the National Forest Inventory in the Georgia per resulting cluster for Bray-Curtis (a) and discriminating Avalanche (b). The boxes represent interquartile ranges and respective median values (solid line) of sample elevations within each clustered assemblage (colored dots). Resulting p values of the $X^2_{Kruskal\text{-}Wallis}$ *test are* $p = 2.899e\text{-}113$ *and* $p = 1.674e\text{-}115$*, with confidence intervals between 0.544, 1 and 0.538 and 1 for assemblages in the hierarchical clustering based on BC and dA, respectively (n = 1059).*

Spatial distributions of discriminated assemblages show a general alignment along biogeographic units. *Alnus glutinosa* dominated assemblages agglomerate in the humid Alder carrs of Eastern Georgia (Nakhutsrishvili, 2013), whereas the *Pinaceae* dominated assemblages are predominantly located in Montane to Sub-alpine areas of the Lesser and Greater Caucasus. Assemblages dominated by mixed *Quercus spp.*, *Carpinus spp.* and *F. orientalis* are situated at intermediate ranges. Those with high shares of BA of *Quercus spp.* are limited to lower and drier areas, while *F. orientalis* dominated assemblages are located at higher elevations. Interestingly, the five samples assigned to *Juniperus* - *Pistacia* woodland have been clearly discriminated within HC_{dA} that are located in the semi-arid areas of the Southwest of the country (Nakhutsrishvili et al., 2021).

To visually evaluate the spatial distribution of discriminated assemblages, we cut HC_{dA} at level II because 70% of the partitions are not partitioned further below level II, resulting in ten clustered assemblages (Figure 2.4). Based on characteristic indicator species, relative BA distributions, and occupied elevational zones, we labeled each assemblage accordingly and mapped the respective sample location in relation to areas of forest vegetation type classifications according to Bohn et al. (2007, Figure 2.6).

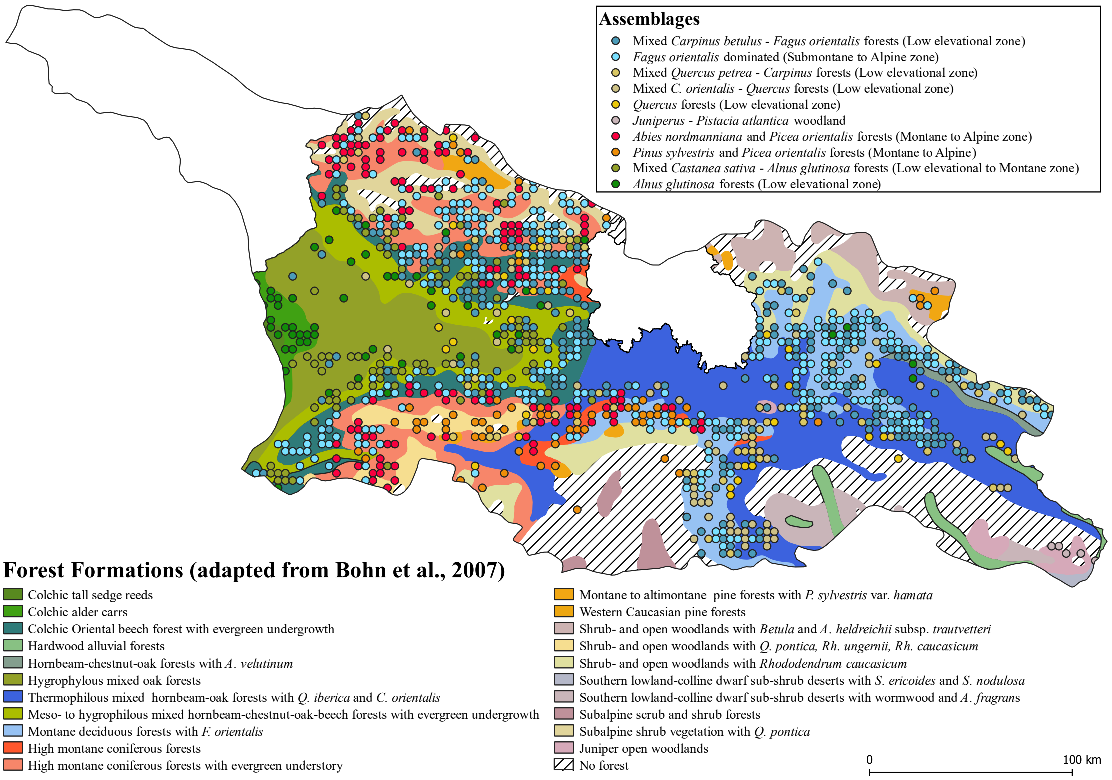


Figure 2.6: Schematic overview of spatial distributions of discriminated assemblages and main forest formations adapted from the vegetation types classification by Bohn et al. (2007). Assemblages are colored according to the respective palette of the four main classes of the cluster hierarchy (blue, beige, red/orange and green for dominance of Fagus, Carpinus-Quercus, Pinaceae and Alnus-Castanea, respectively.

2.4 Discussion

Parting from the assumption that genetic variation of co-inhabiting species provides a signal that reflects site-specific environmental determinants, we contrast the performance of a species-neutral dissimilarity index (BC) with an index that considers interspecies genetic variation (dA) when used in unsupervised classification. Our findings indicate that incorporating interspecies phylogenetic distances in the quantification of dissimilarities results in more coherent and ecologically meaningful classifications of large-scale forest inventory data with high β-diversity.

The Mantel statistics indicate a significant positive correlation between the dissimilarity matrices obtained for each index (r = 0.613, with p = 0.0004), implying that essential patterns of variation among samples are maintained in the respective quantifications. However, frequency distributions and visual inspection of dissimilarities (Figure 2.4) show overall higher dissimilarity values based on BC, reflecting its' sensitivity to species turnover. Whereas the resulting cluster hierarchies maintain a certain level of agreement (cophenetic correlation = 0.511), the fact that a correlation of one signifies complete similarity suggests that the dissimilarity signal resulting from dA is not redundant. Clustering based on dA performs slightly better in preserving the original dissimilarities according to the respective cophenetic correlation coefficients of 0.492 and 0.442 for HC_{BC} and HC_{dA}, respectively. Dendrogram topologies, cluster validation metrics (i.e., WSS of 344.123 and 112.917 for HC_{BC} and HC_{dA}, respectively), and distributions of relative BA among assemblages indicate a higher degree of compactness, separation and yield generally more conceivable clustering results based on dA. Overall, HC_{dA} provided enhanced general interpretability and succeeded in discriminating clearly distinguished assemblages regarding compositional characteristics i.e., the *P. atlantica* and *Juniperus* woodlands of the semi-arid lowlands of Southeastern Georgia. These results support the concept that an extension of variables considered for quantification of dissimilarity leads to a refined conception for diversity classification if genetic variation is considered and are in line with the findings of Hao et al. (2019b) and other authors (Pavoine & Ricotta, 2014; Ricotta & Pavoine, 2015; Capelo, 2020).

As BC represents a “species-neutral” diversity index *sensu* Chao et al. (2010), which assumes that all observed species contribute equally to overall diversity, species turnover constitutes the predominant signal for discrimination, reflected by the significantly higher number of indicator species with frequencies of 100% (24 and 8 for HC_{BC} and HC_{dA}, respectively). Conversely, as dA dissimilarity considers species as phylogenetic units a complete species turnover does not necessarily result in maximum dissimilarity values between two sites, because differences in abundance are weighted by the genetic proximity between species. Assuming that the genetic signal of co-occurring species reflects niche occupation within given ecogeographical areas (Hawkins et al., 2014), the thus refined dissimilarity signal appears to respond to biogeographic gradients in a more interpretable manner.

The validity of the presented approach relies on precise measurements of tree diameters and accurate species identification in the field. While traditional forest science prioritized genus-level information, growing emphasis on diversity-related issues prompted forestry experts to be increasingly trained to provide accurate species identification. The related uncertainties are not design-based issues, but apply to all ecological surveys requiring botanical expert knowledge to ensure taxon detection and validate observations on species level (Lam & Kleinn, 2008; Roswell et al., 2021). Overall, only 2% of all cluster plot observations included individuals that were not identified to species level and had to be excluded. By considering only cluster plot observations of equal sample size ($m = 3$, 64% of all clusters plots) and the exclusion of subplots overlapping with the forest boundary (16%), our analyses are based on a subsample of the NFI data, representing 60% of all cluster plot observations. Hence, conclusions drawn from the presented results should consider e.g., that species exclusively occurring at forest boundaries are excluded. Potential limitations to validity arise from field sampling protocols, as overall subplot size, or sampling effects due to the nested subplot structure (with respective target DBH as inclusion criteria), may introduce bias to the quantification of dissimilarity (Cáceres et al., 2019b). Accordingly, observed numbers of species should be regarded as proxies of true species richness, especially if nested sample plot designs are applied that are based on specific diameter thresholds (Lin et al., 2020). Overall species richness can be assumed to be higher with the contributions of smaller diameter trees being neglected (Corona et al., 2017). Aggregations to cluster plot level could potentially translate to generalization effects, and the loss of information on site-specific environmental factors on smaller scales. The resulting magnitude of impact on the presented results however, is likely to vary according to forest type, topographical condition and management regime (McRoberts et al., 2009).

With continuing advances in whole-genome phylogenetics and functional genomics, information on PD is continuously improving (Kling et al., 2018). Access to comprehensive and standardized phylogenetic mega trees to quantify species genetic relationships is readily available and their application to investigate variation in community compositions is becoming increasingly more common (Gilbert & Parker, 2022; Jin & Qian, 2022).

Our results are of relevance for a wide range of classifications of ecological entities according to conservation value, mapping of ecological communities or other discriminative objectives. The method aligns standard canopy layer data from forest inventories with natural vegetation types according to Bohn et al. (2007), but harmonization with existing forest typologies is limited due to methodological differences, such as the structural vegetation layers considered and the abundance units recorded (e.g., Mucina et al., 2016; Chytrý et al., 2020). Investigating the degree to which the resulting clusters can be aligned with defined syntaxonomical units is an interesting area for future research, especially for the identification of diagnostic species from the shrub and herb layers to refine classifications and the development of practical assessment procedures.

The integration of genetic signals of forest communities during characterization has wide implications for respective approaches to classification. As a proxy indicative for the relationship between species composition and site conditions, interspecies genetic variation extends the scope of forest diversity mapping, management, and monitoring to account for alterations inconceivable by conventional compositional variables. Beyond the potential advantage for streamlining processes by applying unsupervised classification to large datasets, our approach is straightforward and can be readily replicated with comparable data, provided entities are assessed in a systematic manner. Ecological studies are frequently less systematic and constrained by temporal and spatial scales due to the dynamic nature of communities over time and space. This holds true to a lesser degree for assessments of woody species communities within the context of national forest monitoring systems, which are re-sampled in fixed intervals. Hence, from a practical point of view, the resulting network of observational studies provides a valuable framework for a systematic and recurring collection of ecological data, as additional costs and efforts can be embedded into existing structures. Extending the scope of study objectives to systematic assessments of a wider range of botanical and potentially zoological taxa could provide powerful and statistically robust data for analyses of organisms, structural components, and the interrelationships between them.

2.5 Conclusions

We present an approach to discriminate species diversity from NFI data of forest communities with high beta diversity and species turnover. The novelty of the method lies in considering interspecies genetic variability for the quantification of diversity and subsequent classification using an unsupervised clustering algorithm on a country-wide scale. We demonstrate that large-scale forest inventory data can be classified in an ecologically meaningful manner based on mean basal area estimates per species and consideration of phylogenetic dissimilarity between samples. The thus obtained discrimination of species assemblages provides a differentiated picture of existing diversity patterns along expected biogeographic gradients without the need for additional assessments. This approach aligns with a biodiversity concept considering genetic diversity and could potentially be standardized for application to similar datasets, provided systematic data assessment is granted. The presented results should be considered as a step in evaluating to which extent large-area forest inventory data could provide a backbone for extended biodiversity monitoring systems, as discrimination of woody species assemblages allows for systematic delineation of forest ecosystem if genetic variation is considered.

2.6 References

Abe, S. 2021 Comparison of traditional and automated approaches in classification of Japanese coastal beach and dune vegetation. *Society of Vegetation Science* **38**, 67–80. doi.org/10.15031/vegsci.38.67.

Alberdi, I., Nunes, L., Kovac, M., Bonheme, I., Cañellas, I. and Rego, F.C. et al. 2019 The conservation status assessment of Natura 2000 forest habitats in Europe: capabilities, potentials and challenges of national forest inventories data. *Annals of Forest Science* **76**, 1–15. 10.1007/s13595-019-0820-4.

Barnosky, A.D., Matzke, N., Tomiya, S., Wogan, G.O.U., Swartz, B. and Quental, T.B. et al. 2011 Has the Earth's sixth mass extinction already arrived? *Nature* **471**, 51–57. 10.1038/nature09678.

Bevilacqua, S., Anderson, M.J., Ugland, K.I., Somerfield, P.J. and Terlizzi, A. 2021 The use of taxonomic relationships among species in applied ecological research: Baseline, steps forward and future challenges. *Austral Ecology* **46**, 950–964. 10.1111/aec.13061.

Bohn, U., Zazanashhvili, N. and Nakhutsrishvili, G. 2007 The Map of the Natural Vegetation of Europe and its application in the Caucasus Ecoregion. *Bulletin of the Georgian National Academy of Science* **175**, 112–119. science.org.ge/old/moambe/2007-vol1/112-120.pdf.

Bray, J.R. and Curtis, J.T. 1957 An Ordination of the Upland Forest Communities of Southern Wisconsin. *Ecological Monographs* **27**, 325–349. 10.2307/1942268. doi.org/10.2307/1942268.

Brooks, T.M., Cuttelod, A., Faith, D.P., Garcia-Moreno, J., Langhammer, P. and Pérez-Espona, S. 2015 Why and how might genetic and phylogenetic diversity be reflected in the identification of key biodiversity areas? *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* **370**, 20140019. 10.1098/rstb.2014.0019.

Cabido, M., Zeballos, S.R., Zak, M., Carranza, M.L., Giorgis, M.A., Cantero, J.J. and Acosta, A.T.R. 2018 Native woody vegetation in central Argentina: Classification of Chaco and Espinal forests. *Applied Vegetation Science* **21**, 298–311. 10.1111/avsc.12369.

Cáceres, M. de, Chytrý, M., Agrillo, E., Attorre, F., Botta-Dukát, Z. and Capelo, J. et al. 2015 A comparative framework for broad-scale plot-based vegetation classification. *Applied Vegetation Science* **18**, 543–560. 10.1111/avsc.12179.

Cáceres, M. de, Coll, L., Legendre, P., Allen, R.B., Wiser, S.K. and Fortin, M.-J. 2019a Trajectory analysis in community ecology. *Ecological Monographs* **89**. 10.1002/ecm.1350.

Cáceres, M. de, Legendre and He, F. 2013 Dissimilarity measurements and the size structure of ecological communities. *Methods Ecol Evol* **4**, 1167–1177. 10.1111/2041-210X.12116.

Cáceres, M. de, Martín-Alcón, S., González-Olabarria, J.R. and Coll, L. 2019b A general method for the classification of forest stands using species composition and vertical and horizontal structure. *Annals of Forest Science* **76**, 1–19. 10.1007/s13595-019-0824-0.

Capelo, J. 2020 Using species abundance and phylogeny conjointly to approach vegetation classification: A case study on Macaronesia's woody vegetation. *Journal of Vegetation Science* **31**, 616–633. 10.1111/jvs.12886.

Černý, T., Kopecký, M., Petřík, P., Song, J.-S., Šrůtek, M. and Valachovič, M. et al. 2015 Classification of Korean forests: patterns along geographic and environmental gradients. *Applied Vegetation Science* **18**, 5–22. 10.1111/avsc.12124.

Chao, A., Chazdon, R.L., Colwell, R.K. and Shen, T.-J. 2005 A new statistical approach for assessing similarity of species composition with incidence and abundance data. *Ecology letters* **8**, 148–159. 10.1111/j.1461-0248.2004.00707.x.

Chao, A., Chiu, C.-H. and Jost, L. 2010 Phylogenetic diversity measures based on Hill numbers. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* **365**, 3599–3609. 10.1098/rstb.2010.0272.

Chao, A., Chiu, C.-H., Villéger, S., Sun, I.-F., Thorn, S. and Lin, Y.-C. et al. 2018 An attribute-diversity approach to functional diversity, functional beta diversity, and related (dis)similarity measures. *Ecological Monographs* **89**. 10.1002/ecm.1343. doi.org/10.1002/ecm.1343.

Chao, A., Thorn, S., Chiu, C.-H., Moyes, F., Hu, K.-H. and Chazdon, R.L. et al. 2023 Rarefaction and extrapolation with beta diversity under a framework of Hill numbers: The iNEXT.beta3D standardization. *Ecological Monographs*. 10.1002/ecm.1588.

Chiu, C.-H., Jost, L. and Chao, A. 2014 Phylogenetic beta diversity, similarity, and differentiation measures based on Hill numbers. *Ecological Monographs* **84**, 21–44. doi.org/10.1890/12-0960.1.

Chytrý, M., Tichý, L., Hennekens, S.M., Knollová, I., Janssen, J.A.M. and Rodwell, J.S. et al. 2020 EUNIS Habitat Classification: Expert system, characteristic species combinations and distribution maps of European habitats. *Applied Vegetation Science* **23**, 648–675. 10.1111/avsc.12519.

Corona, P., Chirici, G., McRoberts, R.E., Winter, S. and Barbati, A. 2011 Contribution of large-scale forest inventories to biodiversity assessment and monitoring. *Forest Ecology and Management* **262**, 2061–2069. 10.1016/j.foreco.2011.08.044.

Corona, P., Franceschi, S., Pisani, C., Portoghesi, L., Mattioli, W. and Fattorini, L. 2017 Inference on diversity from forest inventories: a review. *Biodivers Conserv* **26**, 3037–3049. 10.1007/s10531-015-1017-2.

Cowie, R.H., Bouchet, P. and Fontaine, B. 2022 The Sixth Mass Extinction: fact, fiction or speculation? *Biological reviews of the Cambridge Philosophical Society* **97**, 640–663. 10.1111/brv.12816.

Díaz, S., Demissew, S., Carabias, J., Joly, C., Lonsdale, M. and Ash, N. et al. 2015 The IPBES Conceptual Framework — connecting nature and people. *Current Opinion in Environmental Sustainability* **14**, 1–16. 10.1016/j.cosust.2014.11.002.

Didion, M. 2020 Extending harmonized national forest inventory herb layer vegetation cover observations to derive comprehensive biomass estimates. *Forest ecosystems* **7**. 10.1186/s40663-020-00230-7.

Dolukhanov, A.G. 2010 *Лесная растительность Грузии: (Forest vegetation of Georgia).* Universal, Tbilisi, Tbilisi, Georgia.

Faith, D.P. 2013 Biodiversity and evolutionary history: useful extensions of the PD phylogenetic diversity assessment framework. *Annals of the New York Academy of Sciences* **1289**, 69–89. 10.1111/nyas.12186.

FAO. 2020 *Global Forest Resources Assessment 2020: Key findings.* FAO, Rome. 978-92-5-132581-0. far.org, 16 p.

Farr, T.G., Rosen, P.A., Caro, E., Crippen, R., Duren, R. and Hensley, S. et al. 2007 The Shuttle Radar Topography Mission. *Reviews of Geophysics* **45**, RG2004. 10.1029/2005RG000183.

Fassnacht, F.E., Latifi, H., Stereńczak, K., Modzelewska, A., Lefsky, M. and Waser, L.T. et al. 2016 Review of studies on tree species classification from remotely sensed data. *Remote Sensing of Environment* **186**, 64–87. 10.1016/j.rse.2016.08.013.

Feilhauer, H., Zlinszky, A., Kania, A., Foody, G.M., Doktor, D., Lausch, A. and Schmidtlein, S. 2021 Let your maps be fuzzy!—Class probabilities and floristic gradients as alternatives to crisp mapping for remote sensing of vegetation. *Remote Sens Ecol Conserv* **7**, 292–305. 10.1002/rse2.188.

Fischer, E., Gröger, A. and Lobin, W. 2018 *Illustrated field guide to the flora of Georgia (South Caucasus).* 1st edn. University of Koblenz-Landau, Koblenz. 978-3-9820257-0-4, 830 p.

Ganeshaiah, K.N. and Shaankar, R.U. 2000 Measuring biological heterogeneity of forest vegetation types: avalanche index as an estimate of biological diversity. *Biodiversity and Conservation*, 953–963. doi.org/10.1023/A:1008910918751.

GBIF Secretariat. 2021 *GBIF Backbone Taxonomy*. gbif.org/.

Gilbert, G.S. and Parker, I.M. 2022 Phylogenetic Distance Metrics for Studies of Focal Species in Communities: Quantiles and Cumulative Curves. *Diversity* **14**, 521. 10.3390/d14070521.

Gillerot, L., Grussu, G., Condor-Golec, R., Tavani, R., Dargush, P. and Attorre, F. 2021 Progress on incorporating biodiversity monitoring in REDD+ through national forest inventories. *Global Ecology and Conservation* **32**, e01901. 10.1016/j.gecco.2021.e01901.

Hao, M., Corral-Rivas, J., González-Elizondo, M.S., Ganeshaiah., K.N., Nava-Miranda, M.G. and Zhang, C. et al. 2019a Assessing biological dissimilarities between five forest communities. *Forest ecosystems* **6**. 10.1186/s40663-019-0188-9.

Hao, M., Gadow, K. von, Alavi, S.J., Álvarez-González, J.G., Baluarte-Vásquez, J.R. and Corral-Rivas, J. et al. 2021 A classification of woody communities based on biological dissimilarity. *Applied Vegetation Science* **24**. 10.1111/avsc.12565.

Hao, M., Ganeshaiah, K.N., Zhang, C., Zhao, X. and Gadow, K. von. 2019b Discriminating among forest communities based on taxonomic, phylogenetic and trait distances. *Forest Ecology and Management* **440**, 40–47. 10.1016/j.foreco.2019.03.006.

Hawkins, B.A., Rueda, M., Rangel, T.F., Field, R. and Diniz-Filho, J.A.F. 2014 Community phylogenetics at the biogeographical scale: cold tolerance, niche conservatism and the structure of North American forests. *Journal of Biogeography* **41**, 23–38. 10.1111/jbi.12171.

Hein, N., Feilhauer, H., Finch, O.-D., Schmidtlein, S. and Löffler, J. 2014 Snow cover determines the ecology and biogeography of spiders (Araneae) in alpine tundra ecosystems. *Erdkunde* **68**, 157–172. 10.3112/erdkunde.2014.03.01.

Hennig, C. 2023 *fpc: Flexible Procedures for Clustering*. cran.r-project.org/package=fpc.

Heym, M., Uhl, E., Moshammer, R., Dieler, J., Stimm, K. and Pretzsch, H. 2021 Utilising forest inventory data for biodiversity assessment. *Ecological Indicators* **121**, 107196. 10.1016/j.ecolind.2020.107196.

Hill, M.O. 1979 TWINSPAN - a FORTRAN program for arranging multivariate data in an ordered two way table by classification of individuals and attributes. *Ecology and Systematics*, 48pp.

Jin, Y. and Qian, H. 2019 V.PhyloMaker: an R package that can generate very large phylogenies for vascular plants. *Ecography* **42**, 1353–1359. 10.1111/ecog.04434.

Jin, Y. and Qian, H. 2022 V.PhyloMaker2: An updated and enlarged R package that can generate very large phylogenies for vascular plants. *Plant diversity* **44**, 335–339. 10.1016/j.pld.2022.05.005.

Joppa, L.N., Roberts, D.L., Myers, N. and Pimm, S.L. 2011 Biodiversity hotspots house most undiscovered plant species. *Proceedings of the National Academy of Sciences of the United States of America* **108**, 13171–13176. 10.1073/pnas.1109389108.

Jost, L. 2010 The Relation between Evenness and Diversity. *Diversity* **2**, 207–232. 10.3390/d2020207.

Kindt, R. 2020 *WorldFlora: An R package for exact and fuzzy matching of plant names against the World Flora Online Taxonomic Backbone data*, 19 p.

Kling, M.M., Mishler, B.D., Thornhill, A.H., Baldwin, B.G. and Ackerly, D.D. 2018 Facets of phylodiversity: evolutionary diversification, divergence and survival as conservation targets. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* **374**. 10.1098/rstb.2017.0397.

Lachashvili, N., Kereselidze, K. and Kikvidze, M. 2022 The checklist of trees of Georgia (Caucasus) and their provisional regional assessment according to IUCN categories and criteria. *Fl. Medit.* **32**, 149–188. 10.7320/FlMedit32.149.

Lam, T.Y. and Kleinn, C. 2008 Estimation of tree species richness from large area forest inventory data: Evaluation and comparison of jackknife estimators. *Forest Ecology and Management* **255**, 1002–1010. 10.1016/j.foreco.2007.10.007.

Lapointe, F.-J. and Legendre, P. 1995 Comparison tests for dendrograms: A comparative evaluation. *journal of Classification* **12**, 265–282. doi.org/10.1007/BF03040858.

Legendre, P. and Cáceres, M. de. 2013 Beta diversity as the variance of community data: dissimilarity coefficients and partitioning. *Ecology letters* **16**, 951–963. 10.1111/ele.12141.

Legendre, P. and Legendre, L. 2012 *Numerical Ecology*. Third English Edition. Elsevier BV, Great Britain. 978-0-444-53868-0, 1003 p.

Leyer, I. and Wesche, K. 2008 *Multivariate Statistik in der Ökologie: Eine Einführung.* Springer, Berlin, Heidelberg. 3540377050, 221 p.

Lin, H.-T., Lam, T.Y., Gadow, K. von and Kershaw, J.A. 2020 Effects of nested plot designs on assessing stand attributes, species diversity, and spatial forest structures. *Forest Ecology and Management* **457**, 117658. 10.1016/j.foreco.2019.117658.

Magurran, A.E. 2005 Biological diversity. *Current Biology* **15**, 116-118. 10.1016/j.cub.2005.02.006.

Magurran, A.E. and McGill, B.J. 2011 *Biological diversity: Frontiers in measurement and assessment.* Oxford University Press, Oxford, New York. 9780199580675, xvii, 345.

Manes, S., Costello, M.J., Beckett, H., Debnath, A., Devenish-Nelson, E. and Grey, K.-A. et al. 2021 Endemism increases species' climate change risk in areas of global biodiversity importance. *Biological Conservation* **257**, 109070. 10.1016/j.biocon.2021.109070.

Mantel, N. 1967 The Detection of Disease Clustering and a Generalized Regression Approach. *Cancer Research* **27**, 209–220.

McRoberts, R.E., Tomppo, E., Schadauer, K., Vidal, C., Ståhl, G. and Chirici, G. et al. 2009 Harmonizing National Forest Inventories. *Journal of Forestry* **107**, 179–187. doi.org/10.1093/jof/107.4.179.

MEPA. 2018 *Field Manual for the Georgian National Forest Inventory: Part II (unpublished).* MEPA, Tbilisi, Georgia, 74 p. 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, Tbilisi, Georgia, 123 p. mepa.gov.ge/En/Files/Download/53934.

Mucina, L., Bültmann, H., Dierßen, K., Theurillat, J.-P., Raus, T. and Čarni, A. et al. 2016 Vegetation of Europe: hierarchical floristic classification system of vascular plant, bryophyte, lichen, and algal communities. *Applied Vegetation Science* **19**, 3–264. 10.1111/avsc.12257.

Myers, N. 2003 Biodiversity Hotspots Revisited. *BioScience* **53**, 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).* Springer Berlin Heidelberg, Berlin, Heidelberg. 978-3-642-29914-8.

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

Novák, P., Kalníková, V., Szokala, D., Aleksanyan, A., Batsatsashvili, K. and Fayvush, G. et al. 2023 Transcaucasian Vegetation Database – a phytosociological database of the Southern Caucasus. *VCS* **4**, 231–240. 10.3897/VCS.105521.

Oksanen, J., Simpson, G., Blanchet, F., Kindt, R., Legendre, P. and Minchin, P. et al. 2022 *vegan: Community Ecology Package*. cran.r-project.org/package=vegan.

Padullés Cubino, J., Lososová, Z., Bonari, G., Agrillo, E., Attorre, F. and Bergmeier, E. et al. 2021 Phylogenetic structure of European forest vegetation. *Journal of Biogeography* **48**, 903–916. 10.1111/jbi.14046.

Palombo, M.R. 2021 Thinking about the Biodiversity Loss in This Changing World. *Geosciences* **11**, 370. 10.3390/geosciences11090370.

Pavoine, S. 2016 A guide through a family of phylogenetic dissimilarity measures among sites. *Oikos* **125**, 1719–1732. 10.1111/oik.03262.

Pavoine, S., Gasc, A., Bonsall, M.B. and Mason, N.W. 2013 Correlations between phylogenetic and functional diversity: mathematical artefacts or true ecological and evolutionary processes? *Journal of Vegetation Science* **24**, 781–793. 10.1111/jvs.12051.

Pavoine, S. and Ricotta, C. 2014 Functional and phylogenetic similarity among communities. *Methods Ecol Evol* **5**, 666–675. 10.1111/2041-210X.12193.

Peterka, T., Hájek, M., Jiroušek, M., Jiménez-Alfaro, B., Aunina, L. and Bergamini, A. et al. 2017 Formalized classification of European fen vegetation at the alliance level. *Applied Vegetation Science* **20**, 124–142. 10.1111/avsc.12271.

Qian, H. and Jin, Y. 2016 An updated megaphylogeny of plants, a tool for generating plant phylogenies and an analysis of phylogenetic community structure. *JPECOL* **9**, 233–239. 10.1093/jpe/rtv047.

R Core Team. 2024 *R: A Language and Environment for Statistical Computing.* R Foundation for Statistical Computing, Vienna, Austria. r-project.org/.

Ricotta, C. 2005 Through the jungle of Biological diversity. *Acta Biotheoretica* **53**, 29–38. doi.org/10.1007/s10441-005-7001-6.

Ricotta, C., Laroche, F., Szeidl, L. and Pavoine, S. 2020 From alpha to beta functional and phylogenetic redundancy. *Methods Ecol Evol* **11**, 487–493. 10.1111/2041-210X.13353.

Ricotta, C. and Pavoine, S. 2015 Measuring similarity among plots including similarity among species: an extension of traditional approaches. *Journal of Vegetation Science* **26**, 1061–1067. 10.1111/jvs.12329.

Ricotta, C., Szeidl, L. and Pavoine, S. 2021 Towards a unifying framework for diversity and dissimilarity coefficients. *Ecological Indicators* **129**, 107971. 10.1016/j.ecolind.2021.107971.

Roswell, M., Dushoff, J. and Winfree, R. 2021 A conceptual guide to measuring species diversity. *Oikos* **130**, 321–338. 10.1111/oik.07202.

RStudio Team. 2024 *RStudio: Integrated Development for R.: RStudio.* Posit Software, P. B.C., Boston, MA, US.

Schmidtlein, S., Collison, J. and Pfannendoerfer, R. 2024 *Package 'isopam': Clustering of Sites with Species Data*, 10 p. cran.r-project.org/package=isopam.

Schmidtlein, S., Collison Jason, Pfannendoerfer, R. and Tichý, L. 2022 *Package 'isopam': Clustering of Sites with Species Data.* Schmidtlein, Sebastian.

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

Scrucca, L., Fop, M., Murphy, T.B. and Raftery, A.E. 2016 mclust 5: clustering, classification and density estimation using Gaussian finite mixture models. *The R Journal* **8**, 289–317. 10.32614/RJ-2016-021.

Smith, S.A. and Brown, J.W. 2018 Constructing a broadly inclusive seed plant phylogeny. *American journal of botany* **105**, 302–314. 10.1002/ajb2.1019.

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

Torresani, M., Rocchini, D., Sonnenschein, R., Zebisch, M., Marcantonio, M., Ricotta, C. and Tonon, G. 2019 Estimating tree species diversity from space in an alpine conifer forest: The Rao's Q diversity index meets the spectral variation hypothesis. *Ecological Informatics* **52**, 26–34. 10.1016/j.ecoinf.2019.04.001.

Tucker, C.M., Cadotte, M.W., Carvalho, S.B., Davies, T.J., Ferrier, S. and Fritz, S.A. et al. 2017 A guide to phylogenetic metrics for conservation, community ecology and macroecology. *Biological reviews of the Cambridge Philosophical Society* **92**, 698–715. 10.1111/brv.12252.

Webb, C.O., Ackerly, D.D., McPeek, M.A. and Donoghue, M.J. 2002 Phylogenies and Community Ecology. *Annu. Rev. Ecol. Evol. Syst.* **33**, 475–505. 10.1146/annurev.ecolsys.33.010802.150448.

Willmer, J.N.G., Püttker, T. and Prevedello, J.A. 2022 Global impacts of edge effects on species richness. *Biological Conservation* **272**, 109654. 10.1016/j.biocon.2022.109654. sciencedirect.com/science/article/pii/S0006320722002075.

Yao, J., Zhang, C., Cáceres, M. de, Legendre, P. and Zhao, X. 2019 Variation in compositional and structural components of community assemblage and its determinants. *J Vegetation Science* **30**, 257–268. 10.1111/jvs.12708.

Yu, H., Fang, G., Rose, K.A., Tang, Y. and Song, X. 2022 Examining Epibenthic Assemblages Associated with Artificial Reefs Using a Species Archetype Approach. *Mar Coast Fish* **14**. 10.1002/mcf2.10206.

Zampiglia, M., Bisconti, R., Maiorano, L., Aloise, G., Siclari, A. and Pellegrino, F. et al. 2019 Drilling Down Hotspots of Intraspecific Diversity to Bring Them Into On-Ground Conservation of Threatened Species. *Front. Ecol. Evol.* **7**. 10.3389/fevo.2019.00205.

Zanne, A.E., Tank, D.C., Cornwell, W.K., Eastman, J.M., Smith, S.A. and FitzJohn, R.G. et al. 2014 Three keys to the radiation of angiosperms into freezing environments. *Nature* **506**, 89–92. 10.1038/nature12872.

Zeballos, S.R., Giorgis, M.A., Cabido, M.R., Acosta, A.T., Del Iglesias, M.R. and Cantero, J.J. 2020 The lowland seasonally dry subtropical forests in central Argentina: vegetation types and a call for conservation. *VCS* **1**, 87–102. 10.3897/VCS/2020/38013.

3 Modeling Phylogenetically Informed Forest Classifications

Publication II builds directly upon the ecophylogenetic stratification framework established in Publication I (Wellenbeck et al., 2024) and extends it into the domain of predictive modeling to forecast cluster membership using environmental proxy variables. To evaluate how accurately the spatial distribution of clustered woody species assemblages can be predicted by environmental correlates, I applied a Random Forest classification modeling approach. By modeling assemblages from environmental variables derived from remote sensing and Earth observation data, this study tests the central hypothesis that community composition defined through evolutionary proximity reflects ecological patterns shaped by environmental filtering processes along biogeographic gradients.

As predictors, I compiled a multivariate dataset representing environmental gradients as spatially explicit raster data within QGIS. Prior to the analysis, I filtered the GNFI dataset in Power BI (Chapter 1.6) by excluding monospecific samples, samples with species not identified at the species level, and samples that were either intersected by a forest boundary or located outside the spatial extent of the area covered by the EO raster data. Contrasting the chosen approach in Chapter 2.2.2.1, I used subplot observations for analyses instead of cluster-plot based samples, to reduce compositional variability within samples. This increased the number of observations to m = 3,466. To account for the influence of disturbance effects (i.e., timber harvest, erosion, etc.) on prediction accuracy, I grouped the samples into subsamples according to disturbance intensity in Power BI. Once each sample was assigned to one of three sample groups (undisturbed, disturbed and disturbed with neophytes, respectively), I compiled separate community matrices for each and transferred them to the R working environment. Because of unequal sample sizes of the sample groups, I randomly selected an equal number of disturbed samples to match the number of undisturbed samples in R. Subsequently, I applied the identical processing chain, including ISOPAM clustering based on dA (Chapters 2.2.3.1 - 2.2.3.4), to all three sample groups. The cluster analyses clearly demonstrated the influence of disturbance, as hierarchical complexity decreased with increasing disturbance intensity (Chapter 3.3.1). Subsequently, I compared the phylogenetic relatedness within woody assemblages across the three cluster hierarchies to test for differences in phylogenetic signal, using SES_{MPD} from the R package PICANTE (Kembel et al., 2010). Homogenization effects led to reduced phylogenetic divergence and weakened species-cluster associations. In parallel, I imported the selected subplots as shapefiles into QGIS, mapped sample group distributions (Figure 3.4) and performed spatial queries against previously prepared Earth observation raster data layers for each sample. From the resulting dataset, which contained more than 30 variables describing site conditions, I visualized correlation matrices in R to filter and identify relevant environmental predictors (see supplemental material in Wellenbeck et al., 2024, available at https://www.sciencedirect.com/science/article/pii/S0378112725002713). Predictor variables for modeling were selected based on ecological relevance and highly intercorrelated variables were excluded based on the discussion with the co-authors. Topographic variables were derived from a high-resolution DEM for classifying terrain forms at small scales and from SRTM data (Chapter 1.4.2). I then loaded a single data table with site property data into the R workspace and organized the predictor variables thematically into climatic conditions, topographic structure, soil properties, and spatial configuration and human influence (Chapter 3.2.2.2.5). I applied Principal Component Analysis (PCA) separately to each of the four thematic groups to reduce the dimensionality of the predictor data sets. Subsequently, I trained three Random Forest classification models, using the package RANDOM FORESTS in R, to predict cluster membership within each sample group based on the principal components of the four thematic predictor groups (Cutler et al., 2007). After discussing the results with the co-authors, I evaluated classification accuracy through overall model error statistics and assessed variable importance to identify which environmental predictor most strongly influenced assemblage discrimination. I performed all statistical evaluations, including model tuning, cross-validation, and analyses of variable importance. In addition, I prepared all figures, drafted and revised the manuscript, incorporated feedback from co-authors, and handled submission as corresponding author (see CRediT authorship contribution statement in the annexed article).

Within the broader conceptual framework of the thesis, this study advances the ecophylogenetic approach by demonstrating that the distribution of phylogenetically informed assemblages can be predicted based on environmental data. Although prediction accuracies were moderate, the study clearly showed that distribution patterns are non-stochastic, with disturbances reducing the phylogenetic signal, thereby limiting detailed clustering and accurate prediction. Consequently, the importance of specific predictors varied among sample groups, though their overall ranking remained largely consistent. These findings underscore the thesis’s overarching theme that the integration of PD into cluster analysis produces coherent groups whose functional properties reflect environmental conditions. This has broad implications for forest biodiversity monitoring and management, enabling targeted conservation of PD or forest science to leverage the inherent relationship between evolutionary proximity and environmental filtering processes.

This study substantiates the role of phylogenetically informed classifications in capturing environmental filtering processes, showing that clustered assemblages align with multivariate environmental gradients across large spatial scales. It further demonstrates that the predictive strength of environmental correlates is constrained by disturbance-driven homogenization, which limits both the resolution of cluster structures and the accuracy of the respective classification models.

Prediction of Woody Species Assemblages using Ecophylogenetics and Earth Observation Data

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.

Published Article:

Wellenbeck, A., Hein, N., Tarkhnishvili, D., Misof, B., Schmidtlein, S., Janiashvili, Z., Dzadzamia, L., Feilhauer, H. (2025), Predicting woody species assemblages using ecophylogenetics and Earth observation data, Forest Ecology and Management, Volume 589, https://doi.org/10.1016/j.foreco.2025.122763

3.1 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 & Kapos, 2002; Chirici et al., 2012; Godoy & 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., 2008; Corona et al., 2011; Traub & 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 & 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 (Swenson, 2011b; Pavoine & Bonsall, 2011). 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 & 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 & Tucker, 2017; Cadotte, 2017) and determine which species co-inhabit sites with similar conditions, forming habitat specific communities (Weiher & 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 (Kusumoto et al., 2021; Bosch 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 biogeographic processes that shape community structure (Webb et al., 2002; Emerson & 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 & 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 & 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 & 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?

3.2 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.

3.2.1 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 ≥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; Mittermeier et al., 2011; Joppa 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).

3.2.2 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.

3.2.2.1 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 (Figure 3.1). 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 (Figure 3.1). 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 (2,278,760 ha, Figure 3.4) was sampled (MEPA, 2023).

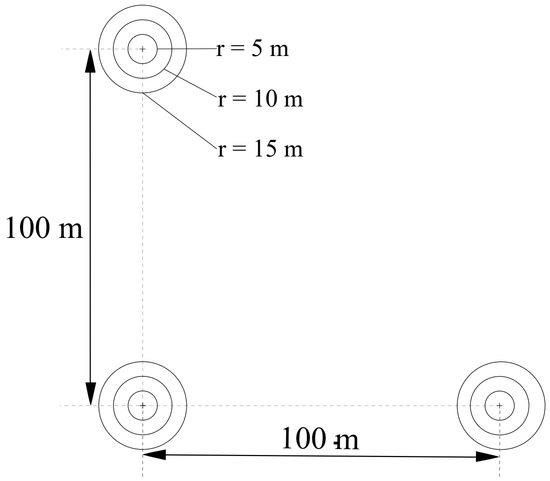


Figure 3.1: Configuration of cluster plots comprising three sample plots of the National Forest Inventory of Georgia. Woody species are recorded within three nested subplots according to the measured diameter at breast height (DBH, at 1.3 m,). For each stem, tree and stem number, species (if identifiable) 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).

3.2.2.2 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 dataset contained 3,466 sample plot records (henceforward referred to as "samples", Table 3.1). 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 3.1), which account for a total sampled area of approximately 90 ha, representing 0.004% of the total forest area of Georgia (2,278,760 ha, MEPA, 2023).

Table 3.1: Summary statistics of analyzed subsampled data.

Sample Group	m	S	Min	Max	BA ha^{-1} Mean*	Var.	CV(%)
Subsamples	1267	96	1.37	120.79	33.19 (±17.084)	291.846	51.47
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

*Mean values are denoted with standard deviation in parenthesis.
Var. = Variance
CV = Coefficient of variation

Sample composition was derived as extrapolated sum of basal area (BA, m^2 ha^{-1}) per species and sample as abundance value (Staudhammer & LeMay, 2001; Yao et al., 2019; Cáceres 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 & 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; Smith & Brown, 2018; Kling et al., 2018; Wellenbeck et al., 2024). We calculated dissimilarities between sample plots based on the Discriminating Avalanche (dA), developed by Ganeshaiah & 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 3.2).

Table 3.2: 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 ($\left(|p_i^a - p_i^b|\right)$)
n = number of samples
p_i^a, p_i^b = relative frequencies of species i in plots a and b

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

3.2.2.2.1 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 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 (Figure 3.2) 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).

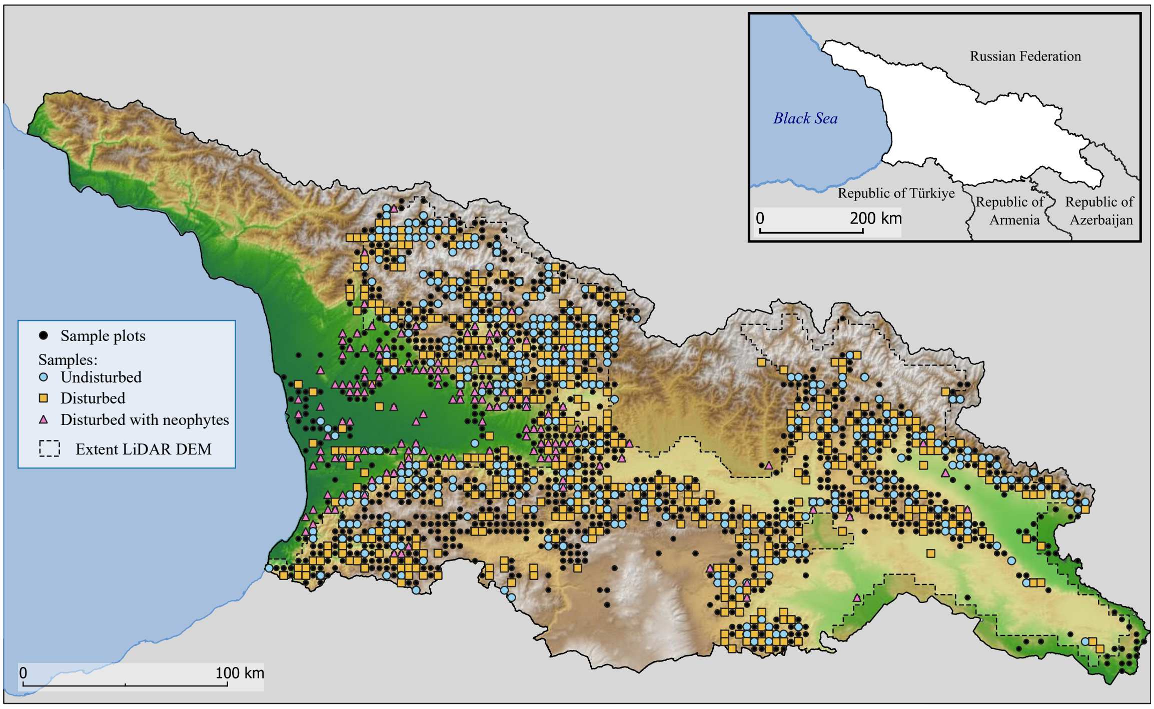


Figure 3.2: Data overview of sample locations and applied digital elevation models. The SRTM based DEM is available for the entire country area. The dotted line outlines the extent of the LiDAR-based DEM. Sample plots (m = 4452) are shown as point features and represent all accessible sample plots located inside forest, which are not intersected by forest boundaries and contain only records of taxa identified at species or subspecies level.

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 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).

3.2.2.2.2 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 3.4) representing annual or seasonal climatic gradients reflecting limiting factors for species distributions (Gardner et al., 2019).

3.2.2.2.3 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 3.4).

3.2.2.2.4 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 & Keon (2002, Feilhauer & Schmidtlein, 2009, Table 3.3).

Table 3.3: Formula to estimate heat load (McCune & Keon, 2002).

$$\textit{Heat load} = 0.339 + \cos(L) \cdot \cos(S) - 0.196 \cdot \sin(L) \cdot \sin(S) - \sin(S) - 0.482 \cdot \cos(A_f) \cdot \sin(S \quad [2]$$

Where:
L = Latitude
A_f = Folded azimuth
S = Slope [°]

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 3.2). 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 & 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 & Jasiewicz (2011) which classifies topographic structures according to geomorphologic phenotypes (Jasiewicz & 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 & Selige, 2006). To reduce the influence of outliers, median values were derived for each sample.

3.2.2.2.5 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 3.4). 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.

Table 3.4: List of variables considered as predictors.

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 & 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 & 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} 100cm^{-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 38N)	213812 - 642200	m	MEPA, 2018	*
	Y	Northing (UTM 38N)	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	OSM	
	costs	SRTM based Cumulative costs	0 - 6387.7	Non-dimensional		*

Notes:
Resolutions ranged from 5 m to 1224 m. Retained variables are marked by '*'.
St. Dev. = Standard deviation; temp. = temperature (°C).

3.2.3 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).

3.2.4 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 & Thuiller, 2005; Chytrý et al., 2016; Kavgacı et al., 2023). Table 3.4 provides an overview of considered and retained variables for modeling.

3.2.5 Random Forest Modeling

RF models are flexible and perform well with complex datasets (Breiman, 2001; Genuer & 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; Soley-Guardia et al., 2024; Gilbert 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 performance and variable importance was quantified via the Mean Decrease in Gini index (MDG, Breiman, 2001; Afanador et al., 2016; Genuer & 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 3.2). 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.

3.2.6 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 3.5).

Table 3.5: Formula to derive weighted variable importance of principal components.

Weighted Variable Importance $$W_i = \sum_{j=1}^{n} \left(|L_{ij}| \cdot Gini_j \right) \quad [3]$$

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

3.2.7 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 (RStudio Team, 2024; R Core Team, 2024). The phylogenetic tree was constructed using the R package V.PHYLOMAKER2 (Jin & 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 & Wiener, 2002). We imputed missing values using the MICE R package (Buuren & Groothuis-Oudshoorn, 2011).

3.3 Results

The analyzed GNFI data (m = 1267) consists of 21,303 recorded trees belonging to 96 species of 56 genera and 30 families. Figure 3.3 shows the constructed phylogenetic hierarchy for all recorded taxa.

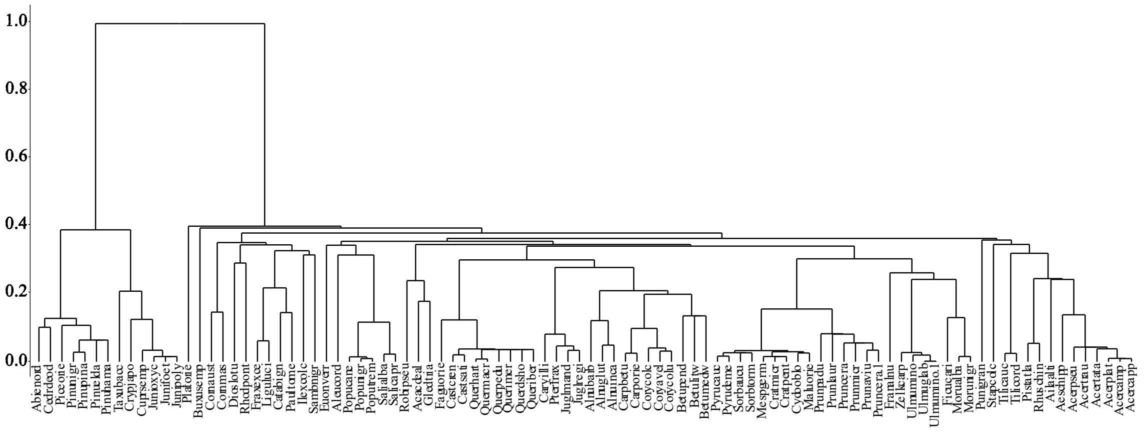


Figure 3.3: Phylogenic tree of 96 species listed in the subsample (m = 1267) of the National Forest Inventory data of Georgia. Interspecies phylogenetic distances were normalized inside a square matrix of pairwise distance values so that $0 < dist_{ph} \leq 1$. Species names on the x-axis are abbreviated 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.

The data subsets according to disturbance contained 8,557 tree records of 61 species, spanning 22 genera and 18 families in disturbed samples, whereas disturbed samples with neophytes included 2,439 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.

3.3.1 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 (≥ 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 (Figure 3.4).

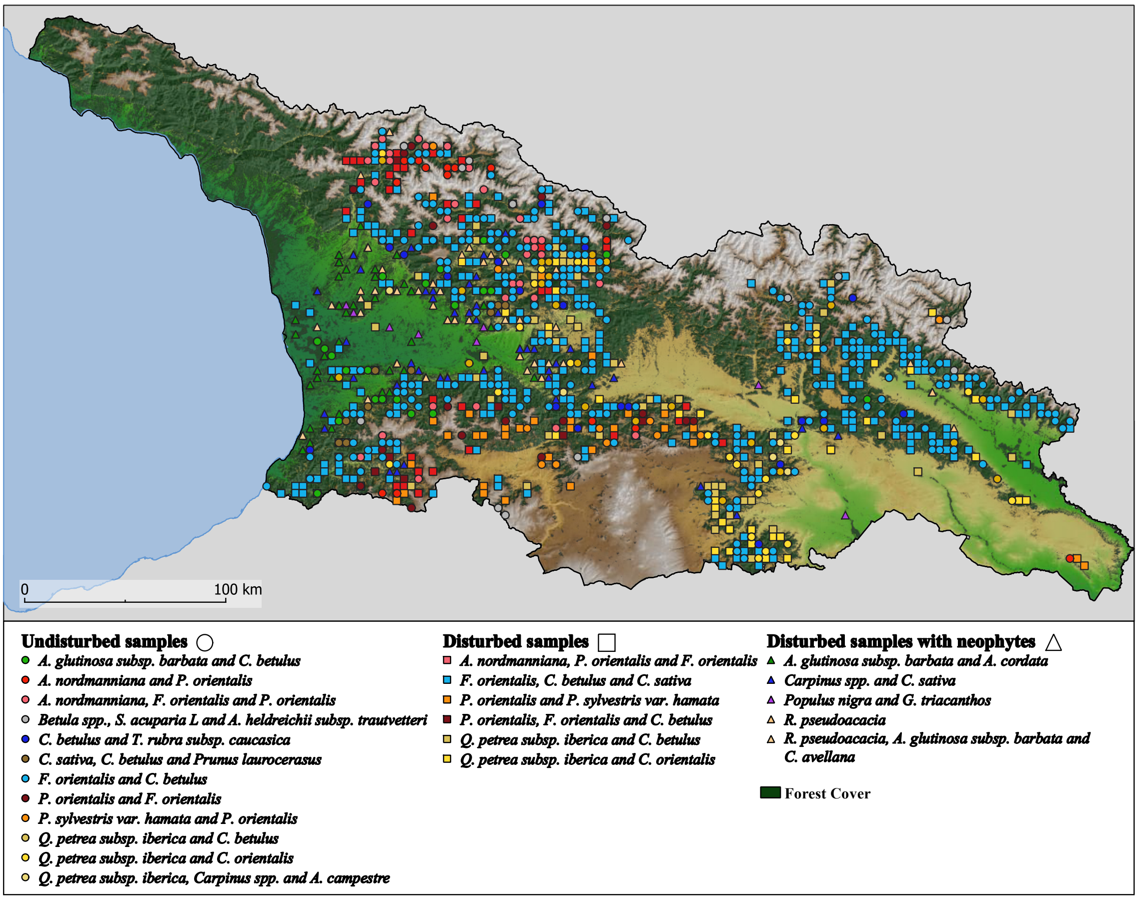


Figure 3.4: Locations of clustered species assemblages per sample group (m = 1267). Clusters are labeled according to basal area share per species (≥10%). Green-shaded areas indicate the distribution of forests in Georgia (Griesbach, 2018) for reference.

Whereas discriminated species assemblages on level II showed some similarity for undisturbed and disturbed samples (Figure 3.4), 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).

3.3.2 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.

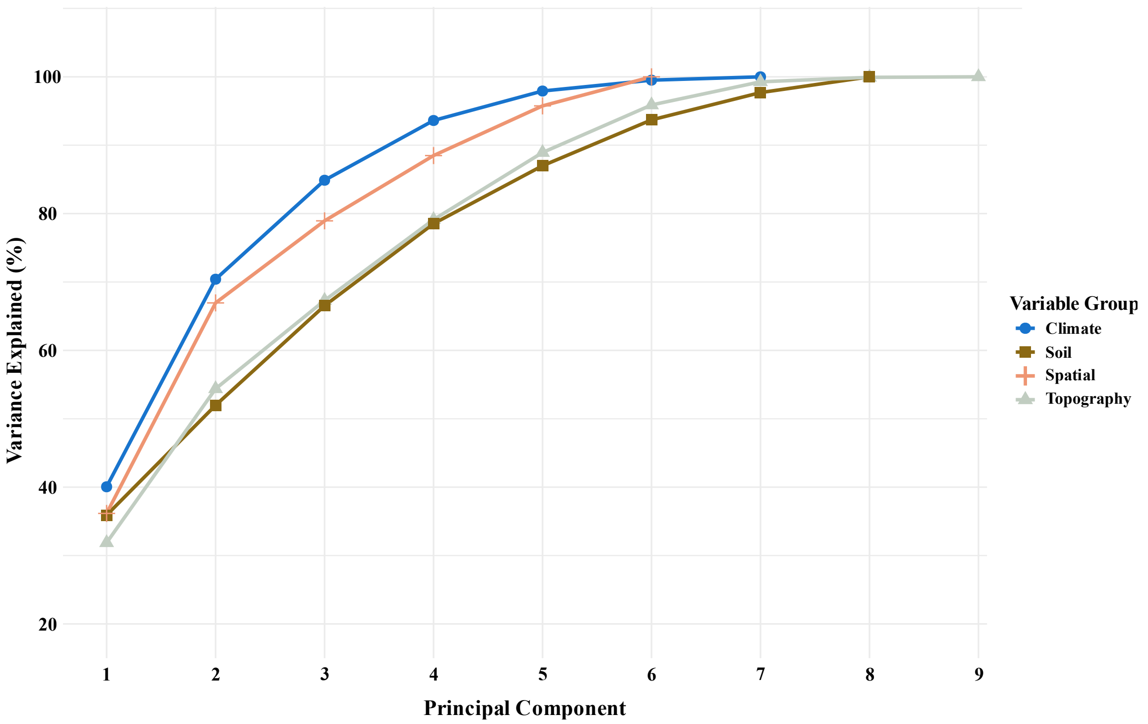


Figure 3.5: Percentage of explained variance per principal component per variable group.

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% (Figure 3.5). 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 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 3.6).

Table 3.6: Out-of-Bag error estimates for Random Forest classification of species assemblage membership.

	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.

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 (Figure 3.6). Variable importance was generally highest in undisturbed samples and lowest for disturbed samples with neophytes.

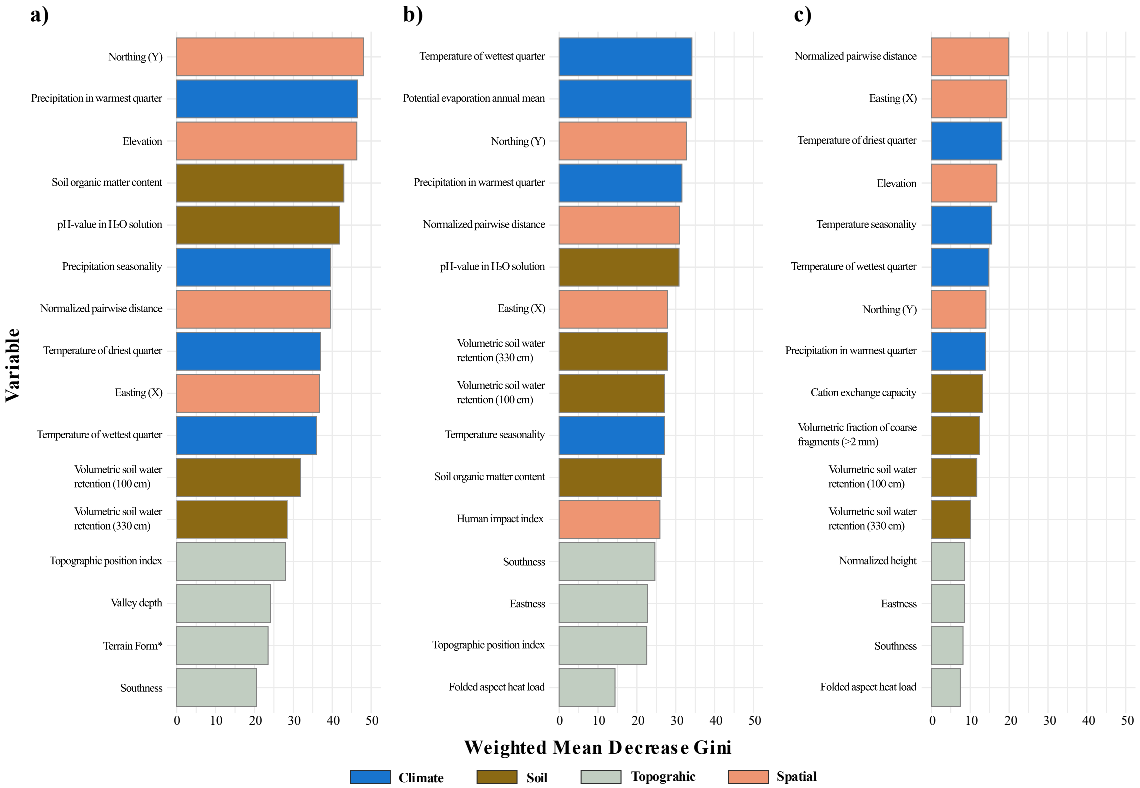


Figure 3.6: Weighted Mean Decrease Gini indices of the four most important explanatory variables per variable and sample group. Variable importance is weighted according to contribution to the most important principal component to predict cluster membership of twelve species assemblages of undisturbed (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.*

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 (Figure 3.6).

3.4 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 & Sandel, 2017; Zhou et al., 2021; Shi 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 (Figure 3.6). 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 biogeographic 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 2,200 and 2,300 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 & Gégout, 2009; Walthert & 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 3.4), 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, i.e., land-use, introduction pathways or management legacies (Emerson & 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 & 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 PD and FD 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 & 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 & 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.

3.5 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.

3.6 References

Afanador, N.L., Smolinska, A., Tran, T.N. and Blanchet, L. 2016 Unsupervised random forest: a tutorial with case studies. *Journal of Chemometrics* **30**, 232–241. 10.1002/cem.2790.

Aguirre-Gutiérrez, J., Malhi, Y., Lewis, S.L., Fauset, S., Adu-Bredu, S. and Affum-Baffoe, K. et al. 2020 Long-term droughts may drive drier tropical forests towards increased functional, taxonomic and phylogenetic homogeneity. *Nature communications* **11**, 3346. 10.1038/s41467-020-16973-4.

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

Araújo, M.B., Anderson, R.P., Márcia Barbosa, A., Beale, C.M., Dormann, C.F. and Early, R. et al. 2019 Standards for distribution models in biodiversity assessments. *Science advances* **5**, eaat4858. 10.1126/sciadv.aat4858.

Ascanio, A., Bracken, J.T., Stevens, M.H.H. and Jezkova, T. 2024 New theoretical and analytical framework for quantifying and classifying ecological niche differentiation. *Ecological Monographs*. 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. jstor.org/stable/1939121.

Böhner, J. and Selige, T. 2006 Spatial Prediction of Soil Attributes using Terrain Analysis and Climate Regionalisation. *Göttinger Geographische Abhandlungen* **115**, 13–27. mediatum.ub.tum.de/doc/1304675/document.pdf.

Bosch, N.E., Wernberg, T., Langlois, T.J., Smale, D.A., Moore, P.J. and Franco, J.N. et al. 2021 Niche and neutral assembly mechanisms contribute to latitudinal diversity gradients in reef fishes. *Journal of Biogeography* **48**, 2683–2698. 10.1111/jbi.14237.

Box, O.E. 1996 Predicting physiognomic vegetation types with climate variables. *Journal of Vegetation Science* **7**, 309–320. 10.2307/3236274.

Breiman, L. 2001 Random Forests. *Machine Learning* **45**, 5–32. 10.1023/A:1010933404324.

Buuren, S. van and Groothuis-Oudshoorn, K. 2011 mice: Multivariate Imputation by Chained Equations in R. *J. Stat. Soft.* **45**, 1–67. 10.18637/jss.v045.i03.

Cabido, M., Zeballos, S.R., Zak, M., Carranza, M.L., Giorgis, M.A., Cantero, J.J. and Acosta, A.T.R. 2018 Native woody vegetation in central Argentina: Classification of Chaco and Espinal forests. *Applied Vegetation Science* **21**, 298–311. 10.1111/avsc.12369.

Cáceres, M. de, Martín-Alcón, S., González-Olabarria, J.R. and Coll, L. 2019 A general method for the classification of forest stands using species composition and vertical and horizontal structure. *Annals of Forest Science* **76**, 1–19. 10.1007/s13595-019-0824-0.

Cadotte, M.W. 2017 Functional traits explain ecosystem function through opposing mechanisms. *Ecology letters* **20**, 989–996. 10.1111/ele.12796.

Cadotte, M.W. and Tucker, C.M. 2017 Should Environmental Filtering be Abandoned? *Trends in Ecology & Evolution* **32**, 429–437. 10.1016/j.tree.2017.03.004.

Cavender-Bares, J. and Wilczek, A. 2003 Integrating Micro- and Macroevolutionary Processes in Community Ecology. *Ecological Society of America* **84**, 592–597. jstor.org/stable/3107852.

Chirici, G., McRoberts, R.E., Winter, S., Bertini, R., Brändli, U. and Asensio, I.A. et al. 2012 National Forest Inventory Contributions to Forest Biodiversity Monitoring. *Forest Science* **58**, 257–268. 10.5849/forsci.12-003.

Chytrý, M., Hennekens, S.M., Jiménez-Alfaro, B., Knollová, I., Dengler and Jansen, F. et al. 2016 European Vegetation Archive (EVA): an integrated database of European vegetation plots. *Applied Vegetation Science* **19**, 173–180. 10.1111/avsc.12191.

Chytrý, M., Tichý, L., Holt, J. and Botta-Dukát, Z. 2002 Determination of diagnostic species with statistical fidelity measures. *J Vegetation Science* **13**, 79–90. 10.1111/j.1654-1103.2002.tb02025.x.

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

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

Corona, P., Chirici, G., McRoberts, R.E., Winter, S. and Barbati, A. 2011 Contribution of large-scale forest inventories to biodiversity assessment and monitoring. *Forest Ecology and Management* **262**, 2061–2069. 10.1016/j.foreco.2011.08.044.

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

Costanza, J.K., Faber-Langendoen, D., Coulston, J.W. and Wear, D.N. 2018 Classifying forest inventory data into species-based forest community types at broad extents: exploring tradeoffs among supervised and unsupervised approaches. *Forest ecosystems* **5**. 10.1186/s40663-017-0123-x.

Cruz-Cárdenas, G., López-Mata, L., Villaseñor, J.L. and Ortiz, E. 2014 Potential species distribution modeling and the use of principal component analysis as predictor variables. *Revista Mexicana de Biodiversidad* **85**, 189–199. 10.7550/rmb.36723.

Cutler, D.R., Edwards, T.C., Beard, K.H., Cutler, A., Hess, K.T., Gibson, J. and Lawler, J.J. 2007 Random forests for classification in ecology. *Ecology* **88**, 2783–2792. 10.1890/07-0539.1.

Davies, R.W., Ryan, C.M., Harrison, R.D., Dexter, K.G., Ahrends, A. and te Beest, M. et al. 2023 Precipitation gradients drive high tree species turnover in the woodlands of eastern and southern Africa. *Ecography* **2023**. 10.1111/ecog.06720.

Davies, T.J. 2021 Ecophylogenetics redux. *Ecology letters* **24**, 1073–1088. 10.1111/ele.13682.

Denk, T., Frotzler, N. and Davitashvili, N. 2001 Vegetational patterns and distribution of relict taxa in humid temperate forests and wetlands of Georgia (Transcaucasia). *Biol J Linn Soc Lond*, 287–332. 10.1006/bijl.2000.0502.

Dering, M., Baranowska, M., Beridze, B., Chybicki, I.J., Danelia, I. and Iszkuło, G. et al. 2021 The evolutionary heritage and ecological uniqueness of Scots pine in the Caucasus ecoregion is at risk of climate changes. *Scientific reports* **11**, 22845. 10.1038/s41598-021-02098-1.

Diniz, É.S., Dias, F.S., Borda-de-Água, L. and González, P.M.R. 2024 Anthropogenic disturbance and alien plant invasion drive the phylogenetic impoverishment in riparian vegetation. *Biodivers Conserv* **33**, 4237–4256. 10.1007/s10531-024-02949-z.

Dolukhanov, A.G. 2010 *Лесная растительность Грузии: (Forest vegetation of Georgia) [in Russian].* Universal, Tbilisi, Georgia. 978-9941-17-176-5. bilib.blogspot.com/2011/12/blog-post_24.html, 484 p.

Doove, L.L., Buuren, S. van and Dusseldorp, E. 2014 Recursive partitioning for missing data imputation in the presence of interaction effects. *Computational Statistics & Data Analysis* **72**, 92–104. 10.1016/j.csda.2013.10.025.

Elizbarashvili, E.S., Chavchanidze, Z.B., Elizbarashvili, M.E., Maglakelidze, R.V., Sulkhanishvili, N.G. and Elizbarashvili, S.E. 2006 Soil-climatic zoning of Georgia. *Eurasian Soil Sc.* **39**, 1062–1065. 10.1134/S1064229306100036.

Emerson, B.C. and Gillespie, R.G. 2008 Phylogenetic analysis of community assembly and structure over space and time. *Trends in Ecology & Evolution* **23**, 619–630. 10.1016/j.tree.2008.07.005.

Ette, J.-S., Sallmannshofer, M. and Geburek, T. 2023 Assessing Forest Biodiversity: A Novel Index to Consider Ecosystem, Species, and Genetic Diversity. *Forests* **14**, 709. 10.3390/f14040709.

Farr, T.G., Rosen, P.A., Caro, E., Crippen, R., Duren, R. and Hensley, S. et al. 2007 The Shuttle Radar Topography Mission. *Reviews of Geophysics* **45**, RG2004. 10.1029/2005RG000183.

Fazlollahi Mohammadi, M., Tobin, B., Jalali, S.G., Kooch, Y. and 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. *Scientific reports* **12**, 7633. 10.1038/s41598-022-10606-0.

Feilhauer, H., Faude, U. and Schmidtlein, S. 2011 Combining Isomap ordination and imaging spectroscopy to map continuous floristic gradients in a heterogeneous landscape. *Remote Sensing of Environment* **115**, 2513–2524. 10.1016/j.rse.2011.05.011.

Feilhauer, H., He, K.S. and Rocchini, D. 2012 Modeling Species Distribution Using Niche-Based Proxies Derived from Composite Bioclimatic Variables and MODIS NDVI. *Remote Sensing* **4**, 2057–2075. 10.3390/rs4072057.

Feilhauer, H. and Schmidtlein, S. 2009 Mapping continuous fields of forest alpha and beta diversity. *Applied Vegetation Science* **12**, 429–439. 10.1111/j.1654-109X.2009.01037.x.

Ferrier, S. and Guisan, A. 2006 Spatial modelling of biodiversity at the community level. *Journal of Applied Ecology* **43**, 393–404. 10.1111/j.1365-2664.2006.01149.x.

Fischer, E., Gröger, A. and Lobin, W. 2018 *Illustrated field guide to the flora of Georgia (South Caucasus).* 1st edn. University of Koblenz-Landau, Koblenz. 978-3-9820257-0-4, 830 p.

Foody, G.M. and Cutler, M.E.J. 2003 Tree biodiversity in protected and logged Bornean tropical rain forests and its measurement by satellite remote sensing. *Journal of Biogeography* **30**, 1053–1066. 10.1046/j.1365-2699.2003.00887.x.

Fuchs, H., Kleinn, C. and Fehrmann, L. 2017 *Establishing the Georgian National Forest Monitoring System: Integrating Remote Sensing, National Forest Inventory (NFI) and Forest Management Inventory (FMI) in Georgia.* Forest Eye Research, GIZ, 22 p.

Ganeshaiah, K.N. and Shaankar, R.U. 2000 Measuring biological heterogeneity of forest vegetation types: avalanche index as an estimate of biological diversity. *Biodiversity and Conservation*, 953–963. doi.org/10.1023/A:1008910918751.

Gardner, A.S., Maclean, I.M. and Gaston, K.J. 2019 Climatic predictors of species distributions neglect biophysiologically meaningful variables. *Diversity and Distributions* **25**, 1318–1333. 10.1111/ddi.12939.

GBIF Secretariat. 2021 *GBIF Backbone Taxonomy*. gbif.org/.

Genuer, R. and Poggi, J. 2020 *Random forests with R.* Springer, Cham. 9783030564841. springer.com/series/6991, 97 p.

Gerhold, P., Cahill, J.F., Winter, M., Bartish, I.V. and Prinzing, A. 2015 Phylogenetic patterns are not proxies of community assembly mechanisms (they are far better). *Functional Ecology* **29**, 600–614. 10.1111/1365-2435.12425.

Gilbert, G.S. and Parker, I.M. 2022 Phylogenetic Distance Metrics for Studies of Focal Species in Communities: Quantiles and Cumulative Curves. *Diversity* **14**, 521. 10.3390/d14070521.

Gilbert, N.A., Amaral, B.R., Smith, O.M., Williams, P.J., Ceyzyk, S. and Ayebare, S. et al. 2024 A century of statistical Ecology. *Ecology*, 1-14. 10.1002/ecy.4283.

Gillespie, T.W., Foody, G.M., Rocchini, D., Giorgi, A.P. and Saatchi, S. 2008 Measuring and modelling biodiversity from space. *Progress in Physical Geography: Earth and Environment* **32**, 203–221. 10.1177/0309133308093606.

Gioia, D., Danese, M., Corrado, G., Di Leo, P., Minervino Amodio, A. and 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**, 725. 10.3390/ijgi10110725.

Gioria, M., Carta, A., Balogianni, V., Fornara, D., Pyšek, P. and 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. 10.3897/neobiota.88.109185.

Godoy, O. and 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**, 67–79. 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. 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. and Thuiller, W. 2005 Predicting species distribution: offering more than simple habitat models. *Ecology letters* **8**, 993–1009. 10.1111/j.1461-0248.2005.00792.x.

Haesen, S., Lembrechts, J.J., Frenne, P. de, Lenoir, J., Aalto, J. and Ashcroft, M.B. et al. 2023 ForestClim-Bioclimatic variables for microclimate temperatures of European forests. *Global change biology* **29**, 2886–2892. 10.1111/gcb.16678.

Hao, M., Corral-Rivas, J., González-Elizondo, M.S., Ganeshaiah., K.N., Nava-Miranda, M.G. and Zhang, C. et al. 2019a Assessing biological dissimilarities between five forest communities. *Forest ecosystems* **6**. 10.1186/s40663-019-0188-9.

Hao, M., Gadow, K. von, Alavi, S.J., Álvarez-González, J.G., Baluarte-Vásquez, J.R. and Corral-Rivas, J. et al. 2021 A classification of woody communities based on biological dissimilarity. *Applied Vegetation Science* **24**. 10.1111/avsc.12565.

Hao, M., Ganeshaiah, K.N., Zhang, C., Zhao, X. and Gadow, K. von. 2019b Discriminating among forest communities based on taxonomic, phylogenetic and trait distances. *Forest Ecology and Management* **440**, 40–47. 10.1016/j.foreco.2019.03.006.

Hawkins, B.A., Rueda, M., Rangel, T.F., Field, R. and Diniz-Filho, J.A.F. 2014 Community phylogenetics at the biogeographical scale: cold tolerance, niche conservatism and the structure of North American forests. *Journal of Biogeography* **41**, 23–38. 10.1111/jbi.12171.

Hengl, T., Mendes de Jesus, J., Heuvelink, G.B.M., Ruiperez Gonzalez, M., Kilibarda, M. and Blagotić, A. et al. 2017 SoilGrids250m: Global gridded soil information based on machine learning. *PloS one* **12**, e0169748. 10.1371/journal.pone.0169748.

Henttonen, H.M. and Kangas, A. 2015 Optimal plot design in a multipurpose forest inventory. *Forest ecosystems* **2**. 10.1186/s40663-015-0055-2.

Hernández-Stefanoni, J.L., Gallardo-Cruz, J.A., Meave, J.A., Rocchini, D., Bello-Pineda, J. and López-Martínez, J.O. 2012 Modeling α- and β-diversity in a tropical forest from remotely sensed and spatial data. *International Journal of Applied Earth Observation and Geoinformation* **19**, 359–368. 10.1016/j.jag.2012.04.002.

Heym, M., Uhl, E., Moshammer, R., Dieler, J., Stimm, K. and Pretzsch, H. 2021 Utilising forest inventory data for biodiversity assessment. *Ecological Indicators* **121**, 107196. 10.1016/j.ecolind.2020.107196.

Holt, R. 2009 Bringing the Hutchinsonian niche into the 21st century: Ecological and evolutionary perspectives. *PNAS* **106**, 19659–19665. 10.1073/pnas.0905137106.

Jasiewicz, J. and Stepinski, T.F. 2013 Geomorphons — a pattern recognition approach to classification and mapping of landforms. *Geomorphology* **182**, 147–156. 10.1016/j.geomorph.2012.11.005.

Jin, Y. and Qian, H. 2019 V.PhyloMaker: an R package that can generate very large phylogenies for vascular plants. *Ecography* **42**, 1353–1359. 10.1111/ecog.04434.

Jin, Y. and Qian, H. 2022 V.PhyloMaker2: An updated and enlarged R package that can generate very large phylogenies for vascular plants. *Plant diversity* **44**, 335–339. 10.1016/j.pld.2022.05.005.

Joppa, L.N., Roberts, D.L., Myers, N. and Pimm, S.L. 2011 Biodiversity hotspots house most undiscovered plant species. *Proceedings of the National Academy of Sciences of the United States of America* **108**, 13171–13176. 10.1073/pnas.1109389108.

Kavgacı, A., Karaköse, M., Keleş, E.S., Balpınar, N., Arslan, M. and Yalçın, E. et al. 2023 Classification of forest and shrubland vegetation in central and eastern Euxine Turkey and SW Georgia. *Applied Vegetation Science* **26**. 10.1111/avsc.12753.

Keggenhoff, I., Elizbarashvili, M., Amiri-Farahani, A. and King, L. 2014 Trends in daily temperature and precipitation extremes over Georgia, 1971–2010. *Weather and Climate Extremes* **4**, 75–85. 10.1016/j.wace.2014.05.001.

Kembel, S.W., Cowan, P.D., Helmus, M.R., Cornwell, W.K., Morlon, H. and Ackerly, D.D. et al. 2010 Picante: R tools for integrating phylogenies and ecology. *Bioinformatics (Oxford, England)* **26**, 1463–1464. 10.1093/bioinformatics/btq166.

Khazaradze, R., Kharadze, K., Tsikarishvili, K. and Chartolani, G. 2018 Ancient Glaciation of the Caucasus. *OJG* **8**, 56–64. 10.4236/ojg.2018.81004.

Kindt, R. 2020 *WorldFlora: An R package for exact and fuzzy matching of plant names against the World Flora Online Taxonomic Backbone data*, 19 p.

Kling, M.M., Mishler, B.D., Thornhill, A.H., Baldwin, B.G. and Ackerly, D.D. 2018 Facets of phylodiversity: evolutionary diversification, divergence and survival as conservation targets. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* **374**. 10.1098/rstb.2017.0397.

Kolbaia, S., Lortkipanidze, B., Kikodze, D., Wong, L.J. and Pagad, S. 2020 *Global Register of Introduced and Invasive Species - Georgia.* IUCN.

Kraft, N.J.B., Adler, P.B., Godoy, O., James, E.C., Fuller, S. and Levine, J.M. 2015 Community assembly, coexistence and the environmental filtering metaphor. *Functional Ecology* **29**, 592–599. 10.1111/1365-2435.12345.

Kuenzer, C., Ottinger, M., Wegmann, M., Guo, H., Wang, C. and Zhang, J. et al. 2014 Earth observation satellite sensors for biodiversity monitoring: potentials and bottlenecks. *International Journal of Remote Sensing* **35**, 6599–6647. 10.1080/01431161.2014.964349.

Kusumoto, B., Kubota, Y., Baselga, A., Gómez-Rodríguez, C., Matthews, T.J., Murphy, D.J. and Shiono, T. 2021 Community dissimilarity of angiosperm trees reveals deep-time diversification across tropical and temperate forests. *Journal of Vegetation Science* **32**. 10.1111/jvs.13017.

Lam, T.Y. and Kleinn, C. 2008 Estimation of tree species richness from large area forest inventory data: Evaluation and comparison of jackknife estimators. *Forest Ecology and Management* **255**, 1002–1010. 10.1016/j.foreco.2007.10.007.

Lausch, A., Baade, J., Bannehr, L., Borg, E., Bumberger, J. and Chabrilliat, S. et al. 2019 Linking Remote Sensing and Geodiversity and Their Traits Relevant to Biodiversity—Part I: Soil Characteristics. *Remote Sensing* **11**, 2356. 10.3390/rs11202356.

Legendre, P. and Legendre, L. 2012 *Numerical Ecology*. Third English Edition. Elsevier BV, Great Britain. 978-0-444-53868-0, 1003 p.

Liaw, A. and Wiener, M. 2002 Classification and Regression by randomForest. *R News* **2**, 18–22.

Lin, H.-T., Lam, T.Y., Gadow, K. von and Kershaw, J.A. 2020 Effects of nested plot designs on assessing stand attributes, species diversity, and spatial forest structures. *Forest Ecology and Management* **457**, 117658. 10.1016/j.foreco.2019.117658.

Macek, M., Kopecký, M. and Wild, J. 2019 Maximum air temperature controlled by landscape topography affects plant species composition in temperate forests. *Landscape Ecology* **34**, 2541–2556. 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. and 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. *Forest Ecology and Management* **474**, 118348. 10.1016/j.foreco.2020.118348.

Marage, D. and Gégout, J.-C. 2009 Importance of soil nutrients in the distribution of forest communities on a large geographical scale. *Global Ecology and Biogeography* **18**, 88–97. 10.1111/j.1466-8238.2008.00428.x.

McCune, B. and Keon, D. 2002 Equations for potential annual direct incident radiation and heat load. *Journal of Vegetation Science* **13**, 603–606. 10.1111/j.1654-1103.2002.tb02087.x.

McRoberts, R.E., Tomppo, E., Schadauer, K., Vidal, C., Ståhl, G. and Chirici, G. et al. 2009 Harmonizing National Forest Inventories. *Journal of Forestry* **107**, 179–187. doi.org/10.1093/jof/107.4.179.

MEPA. 2018 *Field Manual for the Georgian National Forest Inventory: Part II (unpublished).* MEPA, Tbilisi, Georgia, 74 p. 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, Tbilisi, Georgia, 123 p. mepa.gov.ge/En/Files/Download/53934.

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

Miller, T., Blackwood, C.B. and Case, A.L. 2024 Assessing the utility of SoilGrids250 for biogeographic inference of plant populations. *Ecology and evolution* **14**, e10986. 10.1002/ece3.10986.

Mittermeier, R.A., Turner, W.R., Larsen, F.W., Brooks, T.M. and Gascon, C. 2011 Global Biodiversity Conservation: The Critical Role of Hotspots. In *Biodiversity Hotspots*. F.E. Zachos and J.C. Habel (eds). Springer Berlin Heidelberg. 978-3-642-20991-8, pp. 3–22.

Moeslund, J.E., Arge, L., Bøcher, P.K., Dalgaard, T. and Svenning, J.-C. 2013 Topography as a driver of local terrestrial vascular plant diversity patterns. *Nordic Journal of Botany* **31**, 129–144. 10.1111/j.1756-1051.2013.00082.x.

Moudrý, V., Lecours, V., Malavasi, M., Misiuk, B., Gábor, L. and Gdulová, K. et al. 2019 Potential pitfalls in rescaling digital terrain model-derived attributes for ecological studies. *Ecological Informatics* **54**, 100987. 10.1016/j.ecoinf.2019.100987.

Myers, N. 2003 Biodiversity Hotspots Revisited. *BioScience* **53**, 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, Heidelberg New York Dordrecht London. 978-3-642-29914-8, 236 p.

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

Nakhutsrishvili, G., Abdaladze, O., Batsatsashvili, K. and Dzadzamia, L. (eds). 2023 *Natural Forests of Georgia: (The South Caucasus).* Ilia State University Press, Tbilisi, Georgia. 978-9941-18-445-1, 224 p.

Newton, A.C. and Kapos, V. 2002 Biodiversity indicators in national forest inventories. *Unasylva* **53**, 56–64. 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. and Anttila, J. et al. 2019 A comprehensive evaluation of predictive performance of 33 species distribution models at species and community levels. *Ecological Monographs* **89**. 10.1002/ecm.1370.

Novák, P., Kalníková, V., Szokala, D., Aleksanyan, A., Batsatsashvili, K. and Fayvush, G. et al. 2023 Transcaucasian Vegetation Database – a phytosociological database of the Southern Caucasus. *VCS* **4**, 231–240. 10.3897/VCS.105521.

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

OpenStreetMap contributors. 2024 *OpenStreetMap*. download.geofabrik.de/.

Padullés Cubino, J., Lososová, Z., Bonari, G., Agrillo, E., Attorre, F. and Bergmeier, E. et al. 2021 Phylogenetic structure of European forest vegetation. *Journal of Biogeography* **48**, 903–916. 10.1111/jbi.14046.

Patarkalashvili, T. 2017 Forest biodiversity of Georgia and endangered plant species. *Annals of Agrarian Science* **15**, 349–351. 10.1016/j.aasci.2017.06.002.

Pavoine, S. and Bonsall, M.B. 2011 Measuring biodiversity to explain community assembly: a unified approach. *Biological reviews of the Cambridge Philosophical Society* **86**, 792–812. 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. and Rossiter, D. 2021 SoilGrids 2.0: producing soil information for the globe with quantified spatial uncertainty. *SOIL* **7**, 217–240. 10.5194/soil-7-217-2021.

Portier, J., Zellweger, F., Zell, J., Alberdi Asensio, I., Bosela, M. and Breidenbach, J. et al. 2022 Plot size matters: Toward comparable species richness estimates across plot-based inventories. *Ecology and evolution* **12**, e8965. 10.1002/ece3.8965.

QGIS Development Team. 2009 *QGIS Geographic Information System.* QGIS Development Team. qgis.osgeo.org/.

Qian, H., Deng, T., Jin, Y., Mao, L., Zhao, D. and Ricklefs, R.E. 2019 Phylogenetic dispersion and diversity in regional assemblages of seed plants in China. *Proceedings of the National Academy of Sciences of the United States of America* **116**, 23192–23201. 10.1073/pnas.1822153116.

Qian, H. and Sandel, B. 2017 Phylogenetic structure of regional angiosperm assemblages across latitudinal and climatic gradients in North America. *Global Ecology and Biogeography* **26**, 1258–1269. 10.1111/geb.12634.

Qian, H., Zhang, Y., Zhang, J. and Wang, X. 2013 Latitudinal gradients in phylogenetic relatedness of angiosperm trees in North America. *Global Ecology and Biogeography* **22**, 1183–1191. 10.1111/geb.12069.

R Core Team. 2024 *R: A Language and Environment for Statistical Computing.* R Foundation for Statistical Computing, Vienna, Austria. r-project.org/.

Reise, J., Kukulka, F., Flade, M. and Winter, S. 2019 Characterising the richness and diversity of forest bird species using National Forest Inventory data in Germany. *Forest Ecology and Management* **432**, 799–811. 10.1016/j.foreco.2018.10.012.

Ricotta, C., Szeidl, L. and Pavoine, S. 2021 Towards a unifying framework for diversity and dissimilarity coefficients. *Ecological Indicators* **129**, 107971. 10.1016/j.ecolind.2021.107971.

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

Riley, S.J., DeGloria S. D. and Elliot, R. 1999 A Terrain Ruggedness Index that Quantifies Topographic Heterogeneity. *Intermountain Journal of Sciences* **5**, 23–27. download.osgeo.org/qgis/doc/reference-docs/Terrain_Ruggedness_Index.pdf.

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

RStudio Team. 2024 *RStudio: Integrated Development for R.: RStudio.* Posit Software, P. B.C., Boston, MA, US.

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

Schmidtlein, S., Collison, J. and Pfannendoerfer, R. 2024 *Package 'isopam': Clustering of Sites with Species Data*, 10 p. cran.r-project.org/package=isopam.

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

Sefidi, K., Esfandiary Darabad, F. and 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**, 658–665. 10.3832/ifor1080-008.

Shi, W., Wang, Y.-Q., Xiang, W.-S., Li, X.-K. and Cao, K.-F. 2021 Environmental filtering and dispersal limitation jointly shaped the taxonomic and phylogenetic beta diversity of natural forests in southern China. *Ecology and evolution* **11**, 8783–8794. 10.1002/ece3.7711.

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

Smith, S.A. and Brown, J.W. 2018 Constructing a broadly inclusive seed plant phylogeny. *American journal of botany* **105**, 302–314. 10.1002/ajb2.1019.

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

Staab, M., Liu, X., Assmann, T., Bruelheide, H., Buscot, F. and Durka, W. et al. 2021 Tree phylogenetic diversity structures multitrophic communities. *Functional Ecology* **35**, 521–534. 10.1111/1365-2435.13722.

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

Stepinski, T.F. and Jasiewicz, J. 2011 Geomorphons - a new approach to classification of landforms. 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. and 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. *Landscape Ecology* **39**. 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**, e21264. 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. *American journal of botany* **98**, 472–480. 10.3732/ajb.1000289.

Szymura, T.H., Szymura, M. and Macioł, A. 2015 The effect of ecological niche and spatial pattern on the diversity of oak forest vegetation. *Plant Ecology & Diversity* **8**, 505–518. tandfonline.com/doi/full/10.1080/17550874.2015.1010186.

Tarkhnishvili, D. 2014 *Historical Biogeography of the Caucasus.* Nova Science Pub Inc; UK ed. edition. 978-1-63321-910-6. novapublishers.com/shop/historical-biogeography-of-the-caucasus/, 229 p.

Tarkhnishvili, D., Gavashelishvili, A. and Mumladze, L. 2012 Palaeoclimatic models help to understand current distribution of Caucasian forest species. *Biological Journal of the Linnean Society* **105**, 231–248. 10.1111/j.1095-8312.2011.01788.x.

Thompson, I., Guariguata, M.R., Okabe, K., Bahamondez, C., Nasi, R., Heymel, V. and Sabogal, C. 2013 An Operational Framework for Defining and Monitoring Forest Degradation. *Ecology and Society* **18**, 2ß. ecologyandsociety.org/vol18/iss2/art20/.

Tobler, W.R. 1970 A Computer Movie Simulating Urban Growth in the Detroit Region. *Economic Geography* **46**, 234. 10.2307/143141.

Traub, B. and Wüest, R.O. 2020 Analysing the quality of Swiss National Forest Inventory measurements of woody species richness. *Forest ecosystems* **7**, 37. 10.1186/s40663-020-00252-1.

Tucker, C.M., Cadotte, M.W., Carvalho, S.B., Davies, T.J., Ferrier, S. and Fritz, S.A. et al. 2017 A guide to phylogenetic metrics for conservation, community ecology and macroecology. *Biological reviews of the Cambridge Philosophical Society* **92**, 698–715. 10.1111/brv.12252.

Turek, M.E., Poggio, L., Batjes, N.H., Armindo, R.A., van Jong Lier, Q. de, Sousa, L. de and Heuvelink, G.B. 2023 Global mapping of volumetric water retention at 100, 330 and 15 000 cm suction using the WoSIS database. *International Soil and Water Conservation Research* **11**, 225–239. 10.1016/j.iswcr.2022.08.001.

Vanuytrecht, E., Wouters, H., Berckmans, J. and Ridder, K. 2021 *Downscaled bioclimatic indicators for selected regions from 1979 to 2018 derived from ERA5 reanalysis: Product User Guide.* Copernicus Climate Change Service, 44 p. cds.climate.copernicus.eu/datasets/sis-biodiversity-era5-regional.

Waldock, C., Stuart-Smith, R.D., Albouy, C., Cheung, W.W.L., Edgar, G.J. and Mouillot, D. et al. 2022 A quantitative review of abundance-based species distribution models. *Ecography* **2022**. 10.1111/ecog.05694.

Walthert, L. and Meier, E.S. 2017 Tree species distribution in temperate forests is more influenced by soil than by climate. *Ecology and evolution* **7**, 9473–9484. 10.1002/ece3.3436.

Webb, C.O. 2000 Exploring the Phylogenetic Structure of Ecological Communities: An Example for Rain Forest Trees. *The American naturalist* **156**. 10.1086/303378.

Webb, C.O., Ackerly, D.D., McPeek, M.A. and Donoghue, M.J. 2002 Phylogenies and Community Ecology. *Annu. Rev. Ecol. Evol. Syst.* **33**, 475–505. 10.1146/annurev.ecolsys.33.010802.150448.

Webb, C.O., Losos, J.B. and Agrawal, A.A. 2006 Integrating Phylogenies into Community Ecology. *Ecology - Special Issue* **87**, S1-S2. 10.1890/0012-9658(2006)87[1:IPICE]2.0.CO;2.

Weiher, E. and Keddy, P. 2001 *Ecological assembly rules: Perspectives, advances, retreats.* 1st edn. Cambridge University Press, Cambridge, UK. 0-521-65235-9. doi.org/10.1017/CBO9780511542237.

Wellenbeck, A., Fehrmann, L., Feilhauer, H., Schmidtlein, S., Misof, B. and Hein, N. 2024 Discriminating woody species assemblages from National Forest Inventory data based on phylogeny in Georgia. *Ecology and evolution* **14**. 10.1002/ece3.11569.

Wildlife Conservation Society, pers. comm. Global Human Influence Index (HII) Dataset, Last of the Wild Project, Version 2, 2005 (LWP-2):

Willmer, J.N.G., Püttker, T. and Prevedello, J.A. 2022 Global impacts of edge effects on species richness. *Biological Conservation* **272**, 109654. 10.1016/j.biocon.2022.109654. sciencedirect.com/science/article/pii/S0006320722002075.

Winter, M., Schweiger, O., Klotz, S., Nentwig, W., Andriopoulos, P. and Arianoutsou, M. et al. 2009 Plant extinctions and introductions lead to phylogenetic and taxonomic homogenization of the European flora. *PNAS* **106**, 21721–21725. 10.1073/pnas.0907088106.

Winter, S., Chirici, G., McRoberts, R.E., Hauk, E. and Tomppo, E. 2008 Possibilities for harmonizing national forest inventory data for use in forest biodiversity assessments. *Forestry* **81**, 33–44. 10.1093/forestry/cpm042.

Woods, C.L. and Ortmann, K. 2024 Microtopographic heterogeneity affects habitat specialization and diversity of understory plants in a northern temperate rainforest. *Plant Ecology*. 10.1007/s11258-024-01469-8.

Xu, J., Dang, H., Wang, M., Chai, Y., Guo, Y. and Chen, Y. et al. 2019 Is Phylogeny More Useful than Functional Traits for Assessing Diversity Patterns Under Community Assembly Processes? *Forests* **10**, 1159. 10.3390/f10121159.

Yao, J., Zhang, C., Cáceres, M. de, Legendre, P. and Zhao, X. 2019 Variation in compositional and structural components of community assemblage and its determinants. *J Vegetation Science* **30**, 257–268. 10.1111/jvs.12708.

Zanne, A.E., Tank, D.C., Cornwell, W.K., Eastman, J.M., Smith, S.A. and FitzJohn, R.G. et al. 2014 Three keys to the radiation of angiosperms into freezing environments. *Nature* **506**, 89–92. 10.1038/nature12872.

Zazanashvili, N., Sanadiradze, G. and Bukhnikashvili, A. 2001 Caucasus. In *Hotspots. Earth's Biologically Richest and Most Endangered Terrestrial Ecoregions*. 1st edn. R.A. Mittermeier, N. Myers and C. Goettsch Mittermeier (eds). CEMEX, S.A., Agrupación Sierra Madre. 9-686-39758-2. https://doi.org/10.1017/S0376892901270088.

Zellweger, F., Braunisch, V., Morsdorf, F., Baltensweiler, A., Abegg, M. and Roth, T. et al. 2015 Disentangling the effects of climate, topography, soil and vegetation on stand-scale species richness in temperate forests. *Forest Ecology and Management* **349**, 36–44. 10.1016/j.foreco.2015.04.008.

Zhou, W., Zhang, Y., Zhang, S., Yakimov, B.N. and 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**, 591. 10.3390/f12050591.

4 Structural Variation in Phylogenetically Informed Forest Classifications

Article III complements the classification and modeling efforts of the first two publications (Wellenbeck et al., 2024; Wellenbeck et al., 2025) by examining how forest structural diversity varies between phylogenetically informed classifications along a field-assigned degradation severity gradient. It directly addresses the third objective of the thesis by evaluating whether the PD-based classification retains structural coherence under increasing degradation (Chapter 1.7). By comparing changes in structural indicator values across degradation severities and forest formations, the study captures structural responses to degradation severity in a quantitative and ecologically interpretable manner. In addition, this approach allows to evaluate whether all field-assigned "degradation" classes consistently reflect degradation, or whether lower severity rather represent disturbance, because of structural resilience.

I carried out the full data analysis, including filtering and stratifying the GNFI dataset, calculating structural indicators, and performing statistical tests in R. Using the ISOPAM classification derived in Publication II based on m = 3,466 samples, I imported the resulting hierarchy in Power BI (and characterized clusters based on basal area (BA) dominance (Chapter 1.6). To simplify interpretation, I cut the cluster hierarchy and labeled the resulting groups as "forest formations" (Chapter 4.3.1). Subsequently, I assigned monospecific samples manually to forest formations based on BA dominance, increasing the number of grouped samples (m = 4,168). From the GNFI field dataset, I derived structural diversity indicators in Power BI, including tree and regeneration density, basal area per sample, canopy cover, vertical and horizontal distribution of BA, volume increment, and the proportion of deadwood (Chapter 4.2.3.2). I classified each sample according to degradation severity according to the maximum field-assigned degradation severity, ranging from undisturbed to heavily degraded. Subsequently, I migrated the prepared data tables to the R working environment. For each forest formation, I calculated Δ-values as absolute difference of indicator means in relation to undisturbed sample means (including relative change). Based on the distribution of indicator means across degradation severity classes and forest formation, I applied the appropriate statistical test (Welch's t-test or Kruskal–Wallis test) to check for statistical significance (Kruskal & Wallis, 1952). I combined boxplots, including jitter, to visualize the distributions of indicator means and Δ values as bar charts per degradation severity class for each forest formation (Figure 4.4 – Figure 4.7). I created all visualizations using ggplot2 and annotated significances according pairwise comparisons of severity classes (Wickham, 2016). To optimize visualizations, I used INKSCAPE to edit vector-based graphics outside of R (Yuan et al., 2016). Together with the co-authors, I interpreted the results and the ecological implications of the observed trends. I drafted the manuscript, addressed co-author feedback, and led the article's submission process.

The results show distinct structural responses between phylogenetically-informed clusters (forest formations) and extend the ecophylogenetic framework by suggesting that phylogenetic clustering reflects not only compositional but also structural responses to degradation. While some forest formations showed signs of resilience under low degradation severity and maintained structural complexity even under moderate disturbance, others showed marked structural decline with increasing degradation severity. This heterogeneity suggests that PD-based clustering captures meaningful variation in structural response to degradation. In addition, the results provide an empirical foundation to differentiate between disturbance and degradation. However, quantifying threshold values (i.e., Δ-values) that mark negative shifts in successional trajectories to distinguish disturbance from degradation requires further research. This includes defining undisturbed reference values and developing combined indicator approaches.

Whereas Publications I and II focused on classification and predictive modeling, the present study evaluates the ecological relevance of phylogenetically informed clustering in terms of structural diversity It affirms that the developed approach yields not only compositional groupings, but also reflects functional and structural coherence with implications for ecosystem stability. This is highly valuable for developing forest-type-specific, ecosystem-based forest management guidelines.

Stand structural change along a degradation gradient in the National Forest Inventory of Georgia

Abstract

Distinguishing between forest disturbance and degradation is essential for sustainable forest management, and policy formulation as these processes entail different strategies for management based on forest formation-specific resilience thresholds. In this study, we apply a data-driven approach to Georgia's first National Forest Inventory to quantify structural changes along a gradient of field-assigned degradation severity. Using unsupervised, phylogenetically informed clustering we grouped sample plots into similar forest formations and evaluated changes in mean values of key structural indicators—stocking, stand structure, and growth—across severity levels. Our findings show that structural responses are formation-specific and generally non-linear: low severity often reflects resilience or compensation, while moderate and high severity are associated with characteristic patterns of structural decline. Our results highlight the utility of structural metrics in distinguishing disturbance from degradation and underscores the importance of incorporating forest type-specific thresholds into monitoring systems. Based on national forest inventory data, this approach provides a scalable basis for diagnostic monitoring and supports the development of targeted, formation-sensitive forest policy and management planning.

Submitted manuscript:

Wellenbeck, A., Dzadzamia, L., Feldmann, E., Drössler, L., Misof, B., Janiashvili, Z., Hein, N., Stand structural change along a degradation gradient in the National Forest Inventory of Georgia, submitted to European Journal of Forest Research on 06.07.2025 (under review).

4.1 Introduction

Situated between the Greater and Lesser Caucasus, with the Black Sea to the west and the semi-arid Transcaucasian steppes to the east, the forests of Georgia (Sakartvelo) cover roughly 2.3 Mio. ha (MEPA, 2023) and span a wide range of climatic zones despite the country's relatively small size (60,700 km^2). The predominantly mountainous country is recognized for its ecological diversity and forms part of the globally significant Caucasus Biodiversity Hotspot (Myers et al., 2000; Mittermeier et al., 2011). According to the most recent assessments, forests cover approximately 43% of the country's land area (Griesbach, 2018; MEPA, 2023). The combination of topographic complexity and climatic gradients in Georgia has given rise to diverse forest habitats, characterized by distinct ecological conditions and high shares of regional endemism (Zimina, 1978; Akhalkatsi et al., 2014; Mumladze et al., 2019; Nakhutsrishvili et al., 2023). Georgia's forests are distinguished by the presence of unique and uncommon woody species admixtures (Nakhutsrishvili et al., 2023) and were long protected for their ecological functions, with timber imports meeting demand during the Soviet era (UNECE, 2019). Since Georgia's independence in 1991, management has remained small-scale, driven by subsistence needs and continued reliance on forests for grazing and firewood (Patarkalashvili, 2016; Buchner et al., 2020; Stritih et al., 2024; Chalatashvili et al., 2024). Till the early 1990s, forest science was dominated by botanical and biogeographic perspectives, i.e., the most thorough descriptions of forest vegetation classification by Dolukhanov (2010), which was published two decades posthumously (Goginashvili et al., 2021). This foundational classification has been supplemented by more recent regional studies (Bohn et al., 2007; Nakhutsrishvili, 2013), and is progressively being aligned with European vegetation classification standards (Mucina et al., 2016; Chytrý et al., 2020; Preislerová et al., 2024; Knollová et al., 2024). In the last decade, numerous scientific publications focus on specific forest associations on smaller scales (Akhalkatsi et al., 2014; Tephnadze et al., 2014; Akhalkatsi et al., 2019; Novák et al., 2020; Goginashvili et al., 2021; Novák et al., 2021; Akobia et al., 2022; Metreveli et al., 2023). As until recently systematic forest inventory data was unavailable, quantitative analyses of stand structural properties have, to our knowledge, not been carried out (Machavariani, 2010).

While recent advances in forest research, management reforms, and international support have strengthened the implementation of sustainable forestry in Georgia, substantial pressures on forest ecosystems continue (Lomsadze et al., 2019; Chalatashvili et al., 2024). Accordingly, types of degradation were assessed as hierarchical categories with three severity levels during the first National Forest Inventory (GNFI, 2019 – 2021). The results reported indicate that 35.4% (approximately 807,000 ha) of forests show signs of various forms of degradations, with 28.2% and 20% being negatively affected by erosion and cattle grazing, respectively (MEPA, 2023). These impacts are driven by uncontrolled forest use, logging and livestock grazing, exacerbated by a high fuelwood demand and weakened forest management bodies (Machavariani, 2010; Patarkalashvili, 2017; Gutman & Radeloff, 2017; Beridze & Dering, 2021).

Forest disturbance and degradation are closely related yet conceptually distinct processes in forest ecology and management (Thompson et al., 2013; Vásquez-Grandón et al., 2018). Disturbance is broadly defined as a discrete, often stochastic event — either natural or anthropogenic — that alters structure, composition, or function of the forest (Turner, 2010; Fischer et al., 2013; Thom & Seidl, 2016; Patacca et al., 2023). These events, i.e., fire, windthrow and snow damage, avalanches, pest and disease outbreaks or selective logging may temporarily reduce canopy cover or alter competitive dynamics, but are often followed by recovery driven by the forest's intrinsic resilience and self-regulating succession processes (Buma & Wessman, 2011; Mori, 2011; Seidl et al., 2017; Seidl & Turner, 2022). In contrast, degradation represents a more persistent state of ecological decline in forest condition (Thompson et al., 2009; Ghazoul et al., 2015). Numerous definitions of degradation exist (Lund, 2009), but most studies still rely on narrow, carbon-centric frameworks like that of the IPCC (2003), which defines forest degradation as a "direct, human-induced, long-term loss of forest carbon stocks [...] not qualifying as deforestation". Broader definitions consider a wider range of ecological aspects, emphasizing that degraded forests have a diminished capacity to provide key ecosystem services — such as biomass productivity, biodiversity support, or protective functions. These interpretations are more operational for forest management, as they allow targeted responses to be developed and attributed (Sasaki & Putz, 2009; Simula, 2009; Ghazoul et al., 2015; Gunn et al., 2019). Forest degradation can result from chronic pressures such as harvesting, grazing, or pest outbreaks, leading to simplified stand structures (Bahamondez & Thompson, 2016; Kulakowski et al., 2017). While natural or low-impact disturbances may stimulate regeneration and structural adaptation, cumulative or high-impact anthropogenic pressures can push forests beyond resilience thresholds, resulting in arrested succession or simplified structure (Goode et al., 2020; Gough et al., 2022). Distinguishing disturbance from degradation remains conceptually challenging, largely due to variability in disturbance severity, recurrence, and spatial scale (Sturtevant & Fortin, 2021). Such dynamics are more accurately represented along a gradient of ecosystem alteration, with degradation marking a tipping point beyond natural resilience (Tepley et al., 2013). Despite growing recognition of this continuum, locally adapted thresholds to distinguish disturbance from degradation remain scarce, especially in complex, species-rich systems (Senf et al., 2017; Vásquez-Grandón et al., 2018), impeding ecological understanding and the development of targeted restoration strategies.

In this study, we adopt a quantitative approach to differentiate disturbance from degradation across field-assessed severity levels of degradation from the GNFI. To account for site and forest-specific response on various disturbances, we assess changes in basal area, canopy cover, stem density, regeneration, stem size differentiation and vertical distribution, basal area proportions of stumps, and growth metrics (diameter and stand volume increment) across classified forest formations to evaluate whether assigned severity levels indicate reversible dynamics or signs of arrested recovery (Pretzsch et al., 2019). Although remote-sensing based approaches have revealed relatively low deforestation and disturbance rates (Buchner et al., 2020; Chen et al., 2021), field studies suggest more nuanced, localized anthropogenic pressures, especially near settlements (Cortner et al., 2024; Stritih et al., 2024). Hence, quantitative assessments of stand structure are especially valuable in Georgia, where disturbance regimes are typically small-scale and spatially heterogeneous, making them difficult to detect using remote sensing alone (Olofsson et al., 2010; Mikeladze et al., 2020; Stritih et al., 2024). The GNFI provides country-wide systematic sampling data across various forest types including visually assessed "degradation" categories in the field that are divided into three levels of severity. These include biotic and abiotic stressors (e.g., pest outbreaks, timber damage) and offer a practical basis and allow to identify structural alteration. We provide the first structural assessment of selected forest formations in Georgia and examine their relationship to field-assigned degradation levels using recent GNFI data. Our objective was to analyze stand characteristics along a severity gradient of degradation across forest formations and compare observable patterns in stand characteristics with increasing severity. We concentrated our analysis on a general impact gradient (undisturbed, low, moderately, and highly disturbed) as we were particularly interested in formation-specific variation in disturbance patterns, rather than comparing disturbance specific impacts. Hence, the investigated research questions are: i) Do field-assigned degradation severity levels reflect a quantitative gradient in stand properties? ii) Do stand properties respond differently to degradation severity across forest formations? iii) Do observed changes in stand properties along the severity gradient indicate a transition from disturbance to degradation?

4.2 Materials and methods

4.2.1 Study area

Georgia is situated between 41°07'–43°35'N and 40°04'–46°44'E, bordered by the Russian Federation to the north, Türkiye and Armenia to the south, and Azerbaijan to the southeast (Thormann et al., 2019; Nakhutsrishvili et al., 2023). Approximately 55% of the country lies above 1000 m a.s.l., and 40% of its territory is characterized by slopes steeper than 20° (Mikeladze et al., 2020). The Greater and Lesser Caucasus mountain ranges shape Georgia's climate, with mean annual temperatures in mountainous areas ranging from −5 °C to 10 °C and precipitation levels from 800 to 1400 mm (Keggenhoff et al., 2014). The Likhi Range (ca. 1000 m a.s.l.) divides the humid, high-rainfall western regions (13–15 °C, <400–>4000 mm) from the drier, more continental east (10–13 °C, 500–600 mm, Denk et al., 2001; Elizbarashvili et al., 2006). This climatic and topographic variation supports diverse forest types, ranging 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.) of the arid areas of the Southeast. The western lowlands feature humid Colchic forests with (*Alnus glutinosa* subsp. *barbata* (C. A. Mey.) Yalt) and *Castanea sativa* Mill., while central Georgia hosts thermophilic to xerophytic oak forests dominated by *Quercus petraea* subsp. *iberica* (Steven ex M. Bieb.) Krassiln., and hornbeam species (*Carpinus betulus* L., and *Carpinus orientalis* Mill., (Novák et al., 2020; Lachashvili et al., 2020; Novák et al., 2021; Novák et al., 2023b). These transition into mesophytic forests of *Fagus orientalis* Lipsky. and hornbeam–beech associations in more humid regions, forming the core of Georgia's ecological forest (Dolukhanov, 2010; Patarkalashvili, 2017; Nakhutsrishvili et al., 2021; Novák et al., 2023a).

4.2.2 Data preparation

4.2.2.1 Sampling and plot design

Forests in Georgia are defined as "land area with a width of not less than 10 meters and an area of not less than 0.5 ha covered with trees higher than 3 meters and a canopy cover of more than 10%, or with trees able to reach these thresholds in situ" (Government of Georgia, 2020), account for 44.5% (3,100,500 ha) of the country's total area. Field work of the first GNFI (with a 10 year repetition mandate stipulated in the current legislation (Government of Georgia, 2020)) was carried out between 2019 and 2021 (MEPA, 2023). 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 (Figure 4.1). 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 (Figure 1.3) was sampled (MEPA, 2023).

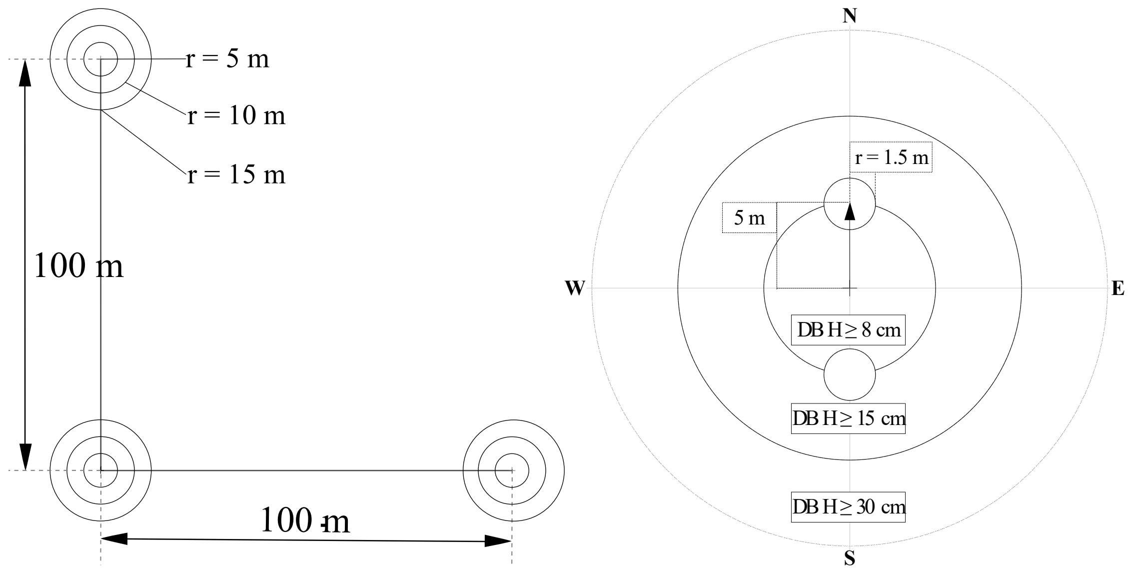


Figure 4.1: Configuration of cluster plots (left) comprising three sample plots of the National Forest Inventory of Georgia. Woody species are assessed within three nested subplots according to the measured diameter at breast height (DBH, at 1.3 m, right). For each stem, tree and stem number, species and DBH measurement are recorded along with the polar coordinates of the stem axis and other variables. Regeneration is assessed on two subplots located 5 m from the sample plot center in a North and South configuration (r = 1.5 m, MEPA, 2018).

Woody species are recorded according to any stems' respective DBH. Stems with DBH ≥8 cm are tallied on the inner nested subplot radius of 5 m, whereas stems with DBH ≥15 cm and DBH ≥30 cm are assessed on subplots with r = 10 m and r = 15 m, respectively. Crown closure is recorded as visual estimates on the same reference area as woody species. Plant counts assess regeneration in two smaller satellite subplots (r =1.5 m) located at 5 m from the plot center (MEPA, 2018, Figure 4.1). On average, increment cores were extracted from 2.8±0.89 stems per sample plot to determine stand age and measure radial growth (MEPA, 2023). In total, 13,356 bore cores were analyzed. Stumps, defined as remaining root collars of trees ≤1.3 m are classified according to origin ("natural" or "anthropogenic") and recorded in the same manner as stems (position, diameter and height), but with a minimum diameter threshold of 10 cm (MEPA, 2018).

4.2.2.2 Forest degradation

At sample plot level, the GNFI collects data on several variables designed to reflect the impacts of external pressures: I) signs of cattle grazing are recorded as "damage of regeneration and ground vegetation by browsing" in a larger 25 m radius observational plot according to three severity classes, II) soil erosion in observed within a 15 m radius and recorded according to origin and severity, III) "forest degradation" is recorded as "changes that has been occurred in the forest and negatively affected on forest structure and functions". Multiple classes can be assigned with the corresponding degree of severity, i.e., "stands of low density", "quality reduction because of non-systemic cuts", "damages caused by phyto- and endopests", "fire affected", "grazing" and "others"

(MEPA, 2018). We use the term *degradation* in accordance with GNFI nomenclature, acknowledging that some recorded impacts may reflect disturbance rather than actual degradation.

4.2.3 Data evaluation

We queried the GNFI dataset ($m = 5970$) to remove all sample plots that were inaccessible ($m = 671$) and not classified as "tree covered area" or "forest" ($m = 1012$) according to the local land cover categorization (MEPA, 2023). To reduce the influence of edge effects on stand structure and species composition, we removed 342 "forest" sample plots that overlapped with the forest boundary (Ries et al., 2004; Willmer et al., 2022). The remaining dataset contained $m = 4168$ sample plot observations ($n = 1717$ cluster), henceforward referred to as "samples" (Figure 4.3). These account for a total sampled area of approximately 294.6 ha, equal to 0.01% of the national forest area (2,278,760 ha, MEPA, 2023). We grouped the data according to recorded degradation severity levels (severity > 0), absence of signs of cattle grazing and anthropogenically induced erosion or stumps that originated from tree cutting operations. Samples with any of the aforementioned attributes >0 were labeled as "disturbed" ($m_{dis} = 2539$). All other samples were considered as "undisturbed" ($m_{undis} = 1629$).

4.2.3.1 Classification of forest formations

For the classification of forest types into broader forest formations, we applied cluster analysis using the *Isometric Partitioning around Medoids* algorithm (ISOPAM, Schmidtlein et al., 2010) of all samples with recorded and complete identification at species levels and with $S_{tr} > 1$. This unsupervised classification method partitions vegetation data iteratively based on ordination and medoid-based clustering. We used an extended version of the ISOPAM algorithm, modified to derive dissimilarities by incorporating phylogenetic distances between woody species (Hao et al., 2019), which ensures that clustering accounts for evolutionary relationships in species composition (Wellenbeck et al., 2024). Monospecific samples ($m = 661$) and samples that contained taxa identification on genus level ($m = 41$) were excluded from the analysis. Consequently, a total of $m = 3,466$ (83%) were clustered based on an abundance matrix containing basal area per hectare (BA [m^2 ha^{-1}]) per recorded woody species and weighted by interspecies phylogenetic distances (supplement 1). Subsequently, the resulting species assemblages were characterized and labeled according to BA proportions (≥10%) and indicator species (Wellenbeck et al., 2024) and classified into "forest formations".

4.2.3.2 Stand Structure

To quantify stand structural characteristics of the classified forest formations, we derived the overall number of stems (N ha^{-1}) and regeneration (N_{reg} ha^{-1}) and BA, along with estimated crown cover percentage (CC%). We calculated proportions of BA per height-class [5 m] (H_p), the quadratic mean diameter (D_q, Curtis & Marshall, 2000) and the share of stump basal area versus the basal area of growing trees (%BA_{stp}). Since %BA_{stp} is derived from DBH and N, this indicator closely mirrors declines in BA and CC%, offering an indirect but robust proxy for removal intensity and reduced timber stocks and i_V. We further compared mean diameter increment (i_d [mm a^{-1}]) and mean stand volume increment (i_V [m^3 ha^{-1} a^{-1}]). All values were aggregated to per hectare values per sample plot. Crown cover plays a key role in regulating light availability throughout the stand profile and influences numerous ecological processes (McElhinny et al., 2005; Modica et al., 2015). In this study, crown cover was recorded as a percentage of area covered by crowns projected to the forest floor at the sample plot center (MEPA, 2018). As crown closure percentage (CC%) is an important proxy for remote-sensing based estimations of BA and stand volume, we compared CC% changes across severity levels. In the analysis of the GNFI, stem volume (V_{st}) is estimated via a three-parametric function derived via regression analysis using species-specific coefficients (*a, b, c*) from national volume tables or statistical fitting (Dees et al., 2020). To quantify existing regeneration, the total of observed seedlings per sample plot were extrapolated to per hectare values.

To quantify vertical structure via distribution analysis, we applied a simple index that calculates the proportion of stems within a specific height class per sample plot. H_p is calculated by assigning stems to distinct height classes (e.g., 0–5 m, 5–10 m, etc.) according to the imputed stem height derived from DBH-H regression analysis (Dees et al., 2020). The proportion of trees in each height class is calculated using the formula:

Table 4.1: Equation applied to quantify vertical stand structure across height-classes.

$$H_p = \frac{\sum n_{H_i}}{\sum n}$$

Where:

n_{H_i} = is the number of stems in height class i.

n = is the total number of stems within the plot.

Low H_p values indicate that stems are highly dispersed among different height classes, suggesting lower densities in vertical strata. A H_p value of 0.5 implies that 50% of the trees are in a specific height class, indicating a balanced distribution across vertical layers. Conversely, H_p values of 1 indicate a uniform vertical structure, as all stems are assigned to one single height class. Forest productivity, measured as mean annual increment can signal degradation when it consistently falls below expected levels for a given forest type and site, particularly if it cannot be restored through management (Bahamondez & Thompson, 2016). Tree ring analysis from extracted bore cores provided annual diameter increment rates (i_d [mm a^{-1}] as ten year average, n = 12,503) and stem age estimates (n = 7,770). Annual stand volume increment (i_V [m^3 ha^{-1} a^{-1}]) was calculated based on measured and estimated diameter increment in the past 10 years and projecting growth rates for the next ten years, adjusting the respective height estimate (Dees et al., 2020).

4.2.4 Analysis

To assess differences in structural characteristics across forest formations and recorded severity levels, we conducted comparative statistical analyses using normality and significance tests. We first compared mean values of each index per forest formation between disturbed samples and undisturbed samples. To test all indices for normality, we applied the Shapiro-Wilk Test (Shapiro & Wilk, 1965) to classify mean value distributions within both groups according to normality (Keren et al., 2020). If $p \leq 0.05$, mean value distribution was classified as “normal”, whereas distributions with $p > 0.05$ were classified as “non-normal”. For mean values that followed normal distributions among groups, we applied Welch’s t-test (Welch, 1947), which accounts for unequal variance between groups. If means were not normally distributed, we applied the Kruskal-Wallis test (Kruskal & Wallis, 1952). Following the selection of significant indices, we applied the same approach across degradation severity levels within each forest formation. For all indices, we calculated absolute differences in means (Δ) along the degradation gradient and visualized sample distributions using boxplots. Statistical significance of Δ values was assessed using the same method as described above (Figure 4.4 - Figure 4.7).

All analyses were conducted using the R base ver. 4.4.0 implemented in R Studio ver. 2024.24.0 (RStudio Team, 2024; R Core Team, 2024). The phylogenetic tree was constructed using the R package V.PHYLOMAKER2 (Jin & Qian, 2019, 2022) and the cluster analysis was performed using ISOPAM ver. 2.0 (Schmidtlein et al., 2024). All spatial data was processed in the QGIS 3.34.8 environment (QGIS Development Team, 2009). Normality was checked using the shapiro.test() function of the package STATS 3.6.2. Welch’s t-test and Kruskal-Wallis tests were implemented using t.test() from base R.

All analyses were conducted using the R base ver. 4.4.0 implemented in R Studio ver. 2024.24.0 (RStudio Team, 2024; R Core Team, 2024). The phylogenetic tree was constructed using the R package V.PHYLOMAKER2 (Jin & Qian, 2019, 2022) and the cluster analysis was performed using ISOPAM ver. 2.0 (Schmidtlein et al., 2024). All spatial data was processed in the QGIS 3.34.8 environment (QGIS Development Team, 2009). For CE computation, we used the FNN ver. 1.1.4.1 (nearest-neighbor calculations) and SPATSTAT ver. 3.1 (spatial pattern analysis) packages in R. Normality was checked using the shapiro.test() function of the package STATS 3.6.2. Welch’s t-test and Kruskal-Wallis tests were implemented using t.test() from base R.

4.3 Results

A total of 3,466 GNFI samples were clustered and labeled based on their basal area (BA) composition. The resulting clusters represent groups of samples with shared species composition and phylogenetic relatedness. These clusters were aggregated into simplified forest formations, and monospecific samples were manually assigned based on dominant species. Mean values and Δ metrics for stocking, structure, and growth variables were then compared across field-assessed degradation severity levels.

4.3.1 Forest formations

Our cluster analysis discriminated 76 species assemblages of samples with similar compositional characteristics and phylogenetic relatedness. The resulting hierarchy is structured in four main groups partitioned on six levels (I-VI, supplement 2). We characterized and labeled each level VI cluster according to dominant genera with BA proportions ≥10% and classified clusters into nine "forest formations", based on BA dominance, namely "*Abies* dominated", "*Picea* dominated", "*Pinus* dominated", "*Fagus* dominated", "*Carpinus* dominated", "*Quercus* dominated", "*Alnus* dominated", "*Castanea* dominated", "Betula dominated", "Acer dominated" and "Other". Subsequently, we allocated monodominant samples (m = 806) manually according to species to the respective formation. Variability of the remaining forest formations contained in "Other" is high as this group encompasses diverse forest types i.e., "*Juniperus* spp. & *Pistacia atlantica*" dominated samples (m = 31), but also *Tilia rubra* subsp. *caucasica* (Rupr.) V. Engl. (m = 69), *Robinia pseudoacacia* L. dominated samples (m = 45), and samples with fast growing pioneer species i.e., *Salix caprea* L., and *Populus tremula* L. (m = 46). In total, 75% of all samples are distributed in formations dominated by broadleaves, whereas 17% are dominated by coniferous genera. Summary statistics per forest formation are provided in supplement 3. Figure 4.2 shows BA proportions across forest formations, along with the distribution of forest formation by elevation class.

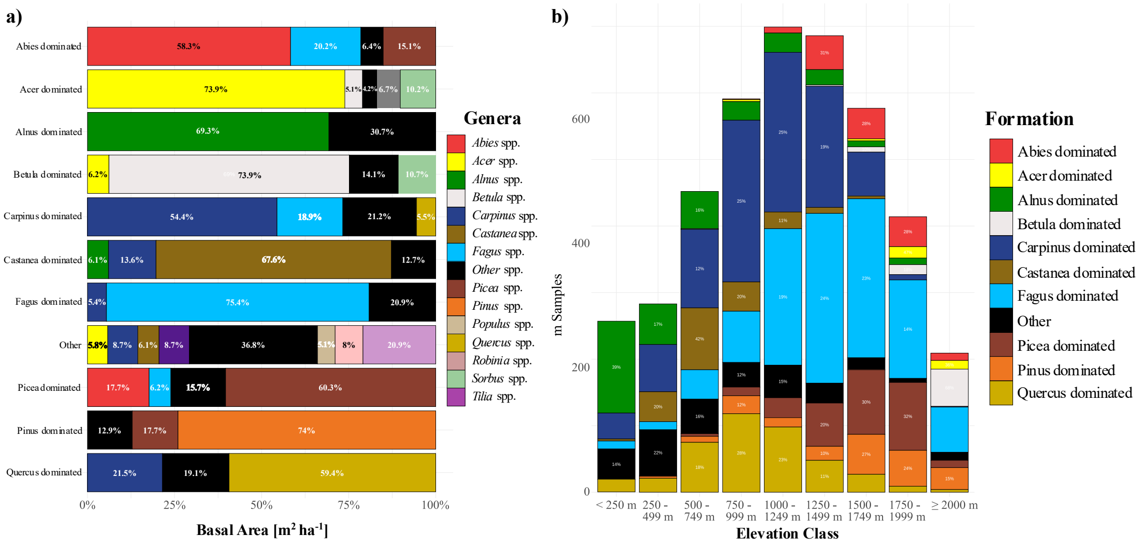


Figure 4.2: Basal area (m^2 ha^{-1}) proportions per genus and forest formation (a). Distributions of samples across elevation classes [m a.s.l] for all forest formation are shown in b). Genera with total shares of ≤ 10% of basal area are grouped as "Other spp.". Percentage values show shares of samples per forest formation (≥10%) that fall within a specific elevation class.

Abies-dominated forest formations occur predominately (>95%) above 1250 m a.s.l.in the western Greater and Lesser Caucasus, typically co-occurring with *P. orientalis*, *F. orientalis* and a minor share of other species, most importantly *Pinus sylvestris subsp. hamata* Steven, *Acer heldreichii* subsp. *trautvetteri* (Medvedev) A. E. Murray, and *C. betulus.* There is limited overlap at these elevations with *Acer*-dominated formations, as these primarily occupy areas above 1500 m a.s.l. (76%) and extend further east across both mountain ranges (Figure 4.2). *A. heldreichii subsp. trautvetteri*, *Acer platanoides* L., *Acer cappadocicum* Gled., *Acer campestre* L. and *Acer pseudoplatanus* L. comprise the dominant species of this forest formation and noteworthy accompanying species include *Sorbus aucuparia* L. (9.5% of total BA) and pioneer species of the genera *Salix*, *Populus* and *Betula*.

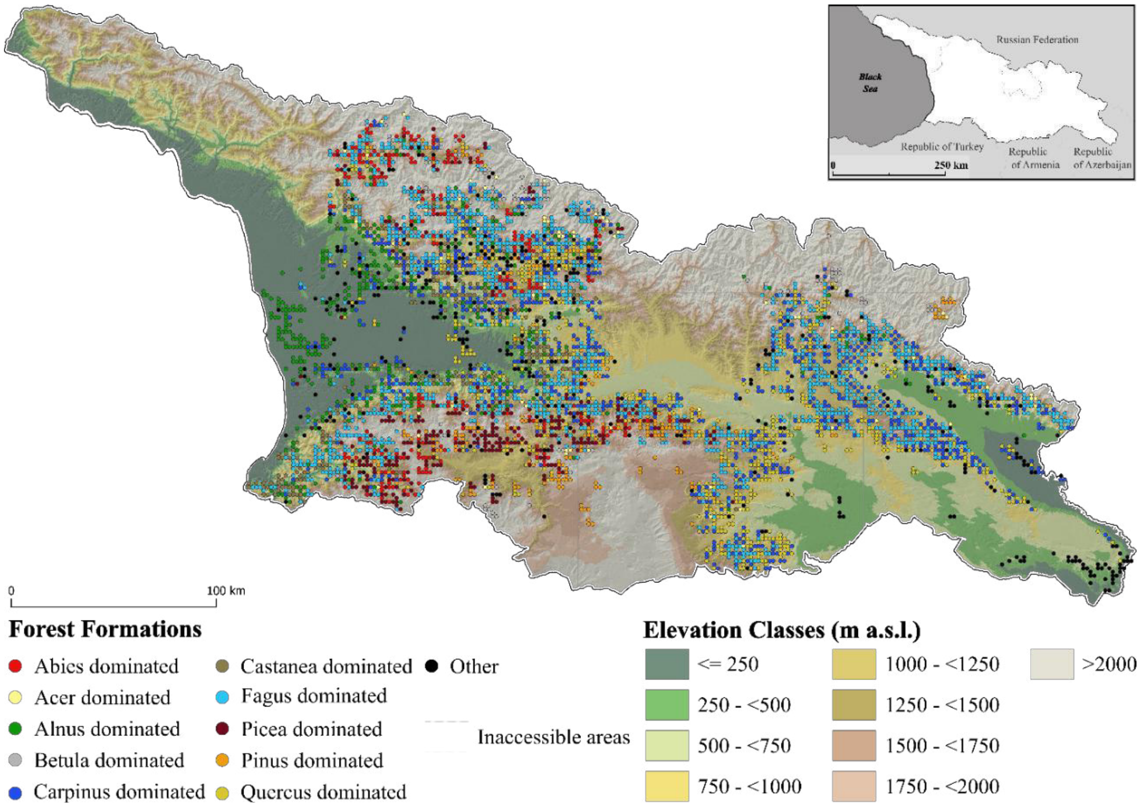


Figure 4.3: Spatial distribution of samples within the country colored according to forest formation. To improve visibility, samples have been shifted manually 1.8 km along the X and Y axis, respectively. For a colored version of this figure please refer to the online version of this article.

The highly diverse *Alnus*-dominated forest formations are primarily concentrated in the western part of the country, particularly within the Colchic lowlands, with 80% of all samples located below 1,000 m a.s.l. These formations account for 34% of all forest samples below 500 m a.s.l. (Figure 4.2) and are dominated by *A. glutinosa* subsp. *barbata* with *Alnus incana* (L.) Moench occurring at higher elevations in the eastern Greater Caucasus, where they form part of riparian forest systems. Within this formation *Alnus* spp. co-occurs with a wide range of other species, likely representing transitions to adjacent formations. In contrast, *Betula*-dominated formations are relatively species poor and are concentrated (89%) above 1750 m a.s.l, accounting for 30% of all samples above 2000 m a.s.l., along with *Fagus* (28%) and *Picea* (15%) formations. Dominated by *Betula* spp., (*Betula pendula* Roth, *Betula medwediewii* Regel, and *Betula pubescens* var. *litwinowii (Doluch.) Ashburner & McAll.*), this formation occurs in association with *S. aucuparia* (11% of total BA). *Carpinus*-dominated formations are primarily (89%) located between 250 and 1500 m a.s.l. and account for 27% of all forest formations between 500 and 1000 m a.s.l. They are associated with *F. orientalis* and *Q. petraea* subsp. *iberica* and many understory and other woody species ($S_{us} = 83$ and $S_{tr} = 64$, respectively), occupying a broad ecological range. *Castanea*-dominated formations are predominantly (94%) located between 250 and 1250 m a.s.l. in the Western part of the country. *Fagus*- dominated formations are distributed between 750 m and 2000 m a.s.l. (89%) and represent the main formation between 1000 and 2000 m a.s.l. (41% of all samples, respectively). Forest formations grouped as “Other” are mostly distributed across elevations below 1000 m a.s.l. (71%), and account for 25% of all formations below 500 m a.s.l. *Picea*-dominated formations occur at elevations between 1000 and 2000 m a.s.l. (89%). *Pinus*-dominated formations are mostly situated above 1500 m a.s.l. (63%) but occur also between 500 and 1000 m a.s.l. (18%). *Quercus*-dominated formations are found across all elevation classes, with 66% of samples concentrated between 500 and 1,500 m a.s.l. Because of limited sample sizes in formations dominated by *Betula* (m = 82), *Acer* (m = 61), and high internal variability in “Other” (m = 346), these formations were excluded from further analyses.

4.3.2 Changes in stand structure

We calculated mean and Δ values (as absolute differences between means) of stand attributes across degradation severity levels. Relative changes in means were also calculated to support interpretation. Variables showing inconsistent trends or high statistical uncertainty were excluded from plotting to improve visual clarity (supplement 4). Additionally, extreme outliers were omitted to allow harmonization of Y-axis scales. Only significant Δ values ($p < 0.05$) were labeled.

Across formations, changes in stand properties are not uniform, as magnitude, pattern, and statistical significance of related Δ values varied among severity levels. Increasing degradation severity resulted in consistent and significant reduction of stocking, expressed as mean BA, and CC% across all formations. In coniferous formations, BA (supplement 4a) and CC% decline in a generally linear pattern, with moderate significance ($p < 0.01$) at low to moderate degradation levels. In contrast, broadleaved formations such as *Fagus*, *Carpinus*, and *Quercus* exhibit more immediate and highly significant reductions ($p < 0.001$), even at low severity (Figure 4.4).

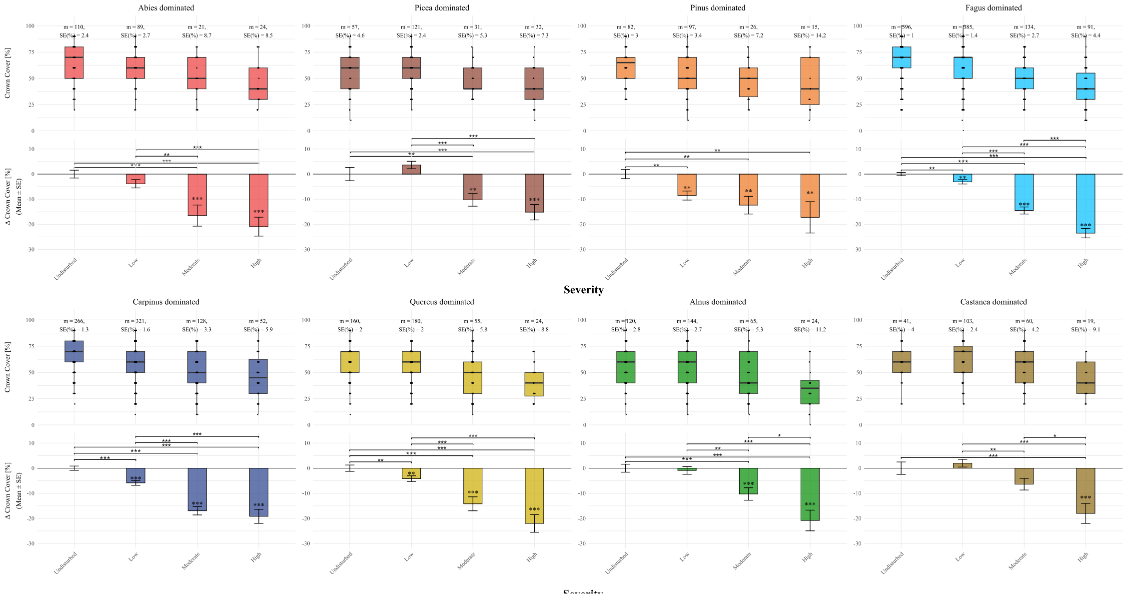


Figure 4.4: Means of crown cover (CC%) and absolute differences (Δ) in mean CC% along the degradation gradient per forest formation. Δ values are derived as absolute differences of CC% across severity levels compared to undisturbed samples ($p \leq 0.05$ = significant (*), $p \leq 0.01$ = moderately significant (**) and $p \leq 0.001$ = highly significant (***)). Error bars represent ± standard error of the mean (SE).

In *Alnus*-dominated formations, mean BA and CC% show highly significant reductions ($p < 0.001$) at moderate and high degradation severity, whereas CC% remains relatively stable between undisturbed and low degradation severity levels. In contrast, in formations dominated by *Castanea* significant changes occur only at high degradation severity, with CC% rising slightly under low degradation severity (62 ± 15.8%) before decreasing to 44± 17.4% at high severity. *Abies*, *Picea*, *Fagus*, *Alnus*, and *Castanea*-dominated formations show a minor, non-significant increase of ~13% in mean N ha^{-1} from undisturbed to low degradation severity, whereas formations dominated by *Pinus*, *Carpinus*, and *Quercus* exhibit reduced mean N ha^{-1} already at low degradation severity (supplement 4b). Horizontal and vertical stand structure tends to decrease with increasing degradation severity, though the magnitude varies by formation and degradation severity level (supplement 4c and 4d). Significant reductions in mean H_p at low severity were found in *Fagus*, *Carpinus*, *Quercus*, and *Alnus*-dominated formations. Mean D_q decreases significantly under high degradation severity, except in *Alnus* and *Pinus* dominated formations. For the latter, no significant changes in mean D_q were observed. $\%BA_{stp}$ increases sharply and consistently with increasing degradation severity, especially at high severity (Figure 4.5). However, along the severity gradient the relative differences in mean $\%BA_{stp}$, do not uniformly decrease across formations. In *Abies* and *Picea-* dominated formations, increases of mean $\%BA_{stp}$ are marked and highly significant. In *Pinus* dominated formations $\Delta\ \%BA_{stp}$ is comparably small, but highly significant ($p \leq 0.001$) compared to undisturbed samples. Among broadleaved formations, *Carpinus* dominated formations show the highest increase, followed by *Quercus* dominated formations. Mean $\%BA_{stp}$

increases in *Alnus* and *Castanea* dominated formations between undisturbed and highly degraded samples are relatively low, but at least moderately significant ($p \leq 0.01$). In *Fagus* dominated formations, all changes in mean $\%BA_{stp}$ showed high statistical significance when compared to means of undisturbed samples.

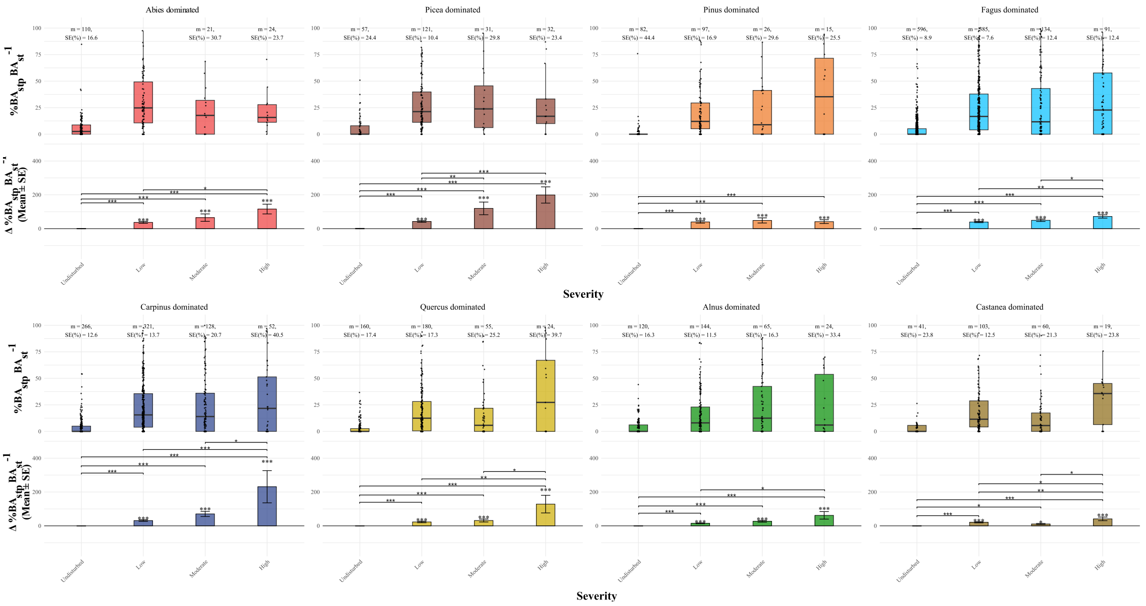


Figure 4.5: Means of stump basal area proportion of in relation to stem basal area ($\%BA_{stp}$) and absolute differences (Δ) in mean $\%BA_{stp}$ along the degradation gradient per forest formation. Δ values are derived as absolute difference between means of degradation severity levels ($p \leq 0.05$ = significant (*), $p \leq 0.01$ = moderately significant (**) and $p \leq 0.001$ = highly significant (***)). Error bars represent ± standard error of the mean (SE).

Distribution of mean N_{reg} ha^{-1} is characterized by high overall variability, with means ranging from 0 (complete absence of regeneration on sample plot) to 19,200 N_{reg} ha^{-1} in *Quercus*-dominated formations of low degradation severity. Most formations show increased regeneration at low to moderate severity, with *Fagus* and *Carpinus*-dominated formations displaying the most consistent and significant ($p < 0.001$) increase. In spite a general increase along the degradation severity gradient, patterns and magnitude of responses is strongly formation dependent (Figure 4.6). Regeneration in *Abies, Picea*, *Quercus* and *Castanea*-dominated formations showed strong and significant increases up to moderate degradation severity, followed by a marked reduction under high degradation severity. In *Carpinus*-dominated formations, the decline under high degradation severity is less pronounced, whereas increased regeneration persists even under high degradation severity in *Alnus* and *Fagus*-dominated formations. Statistically significant differences were primarily detected at low degradation levels ($p < 0.05$). Responses in *Pinus*-dominated formations are minimal and marked by very high variability ($CV\% = 208.7$). *Alnus*-dominated formations showed the lowest mean N_{reg} across the severity gradient, but increasing regeneration persists even under high severity (n.s.).

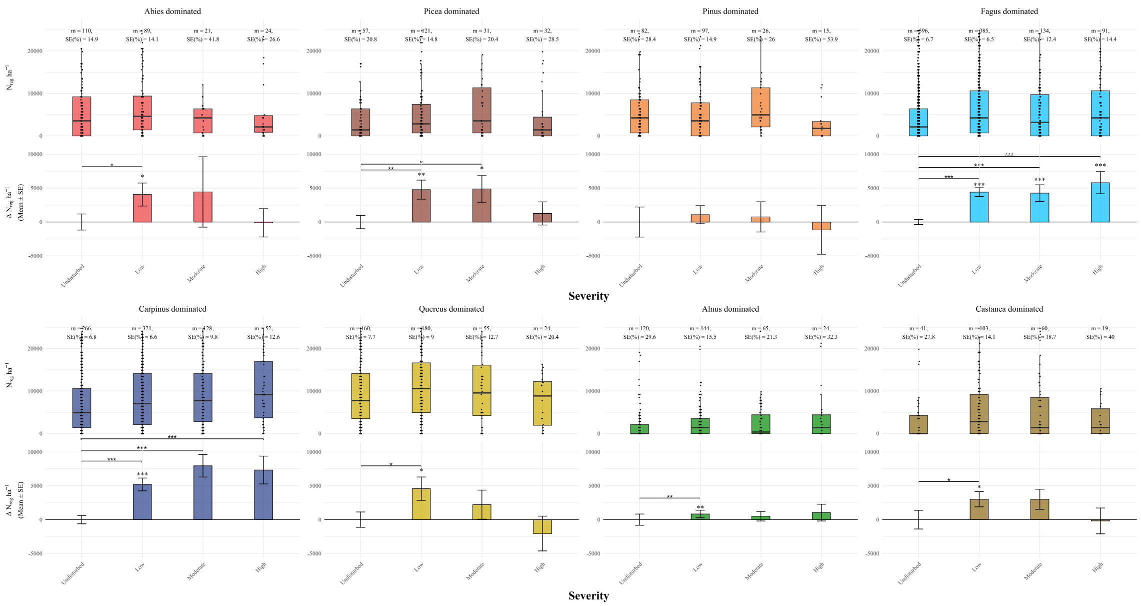


Figure 4.6: Means of seedling density per hectare (N_{reg} ha^{-1}) and absolute differences (Δ) of N_{reg} ha^{-1} across the degradation severity gradient per forest formation. Δ values are derived as absolute difference between means of degradation severity levels ($p \leq 0.05$ = significant (*), $p \leq 0.01$ = moderately significant (**) and $p \leq 0.001$ = highly significant (***)). Error bars represent ± standard error of the mean (SE).

Mean i_d generally increases with degradation severity across all formations, except for *Pinus*-dominated formations that show a minor initial reduction under low degradation severity. Overall, Δ values vary in consistency and magnitude (supplement 4e). Among coniferous formations, *Picea*-dominated formations show a steady, almost linear increase in mean i_d with increasing degradation severity, however changes in mean i_d were not significant for low degradation severity. *Pinus*-dominated formations show a slight decrease in mean i_d at low degradation severity (n.s.), but moderately significant increases thereafter. In contrast, mean i_d is highest at moderate degradation severity in *Abies*-dominated formations. Among broadleaved formations, increases of mean i_d are generally of lower magnitude. *Fagus*-dominated formations show an almost linear and highly significant increase in mean i_d with increasing degradation severity ($p < 0.001$). A similar trajectory along the degradation severity gradient can be observed in *Carpinus*-dominated formations, with a minor decrease at moderate degradation severity. In *Castanea-dominated formations*, mean i_d remains nearly constant at low and moderate degradation severity but increases significantly at high severity levels. *Quercus* and *Alnus-dominated formations* show small significant increases exclusively under low degradation severity. A general pattern of decreasing stand volume increment (i_V) can be observed along the degradation severity gradient (Figure 4.7). However, the trend is not linear across formations. Mean stand volume increment (i_V) declines steadily with increasing degradation severity in *Pinus*, *Fagus* and *Quercus*-dominated formations. In contrast, an initial increase under low degradation severity can be observed in *Abies*, *Picea*, *Alnus* and *Castanea*-dominated formations. Formations dominated by *Carpinus* exhibit the most marked and significant reduction to mean i_V under moderate degradation severity. Statistically significant negative Δ i_V values are observed in *Abies* and *Pinus*-dominated formations under high degradation severity ($p \leq 0.05$), whereas *Fagus*, *Carpinus* and *Quercus* dominated formations show significant reductions already between low and moderate degradation severity ($p < 0.001$). Other formations show less consistent or statistically non-significant changes across severity levels.

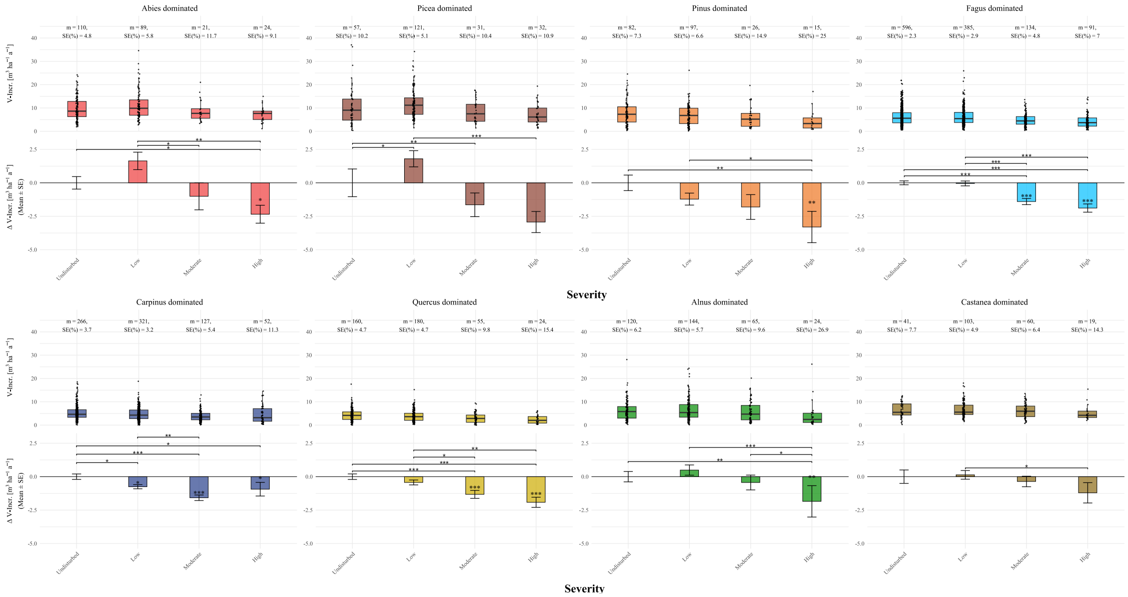


Figure 4.7: Means of stand volume increment (i_V [m^3 ha^{-1} a^{-1}]) and absolute differences (Δ) in mean i_V across the degradation severity gradient per forest formation. Δ values are derived as absolute difference between means of degradation severity levels ($p \leq 0.05$ = significant (*), $p \leq 0.01$ = moderately significant (**) and $p \leq 0.001$ = highly significant (***)). Error bars represent ± standard error of the mean (SE).

4.4 Discussion

For effective forest policy and adaptive sustainable management, it is necessary to distinguish disturbance from degradation, because each demands specific silvicultural responses in relation to the resilience capacity of forest stands. Forest formation-specific patterns in structural alternations can serve as a reference to assess whether observed changes reflect natural variability or point to functional degradation. We compared means and Δ values of variables related to stocking, structure, and growth per forest formation across a field-assigned gradient of degradation severity. Forest formations were derived via aggregation of clustering results to groups of samples with similar compositional data and phylogenetic variability. The spatial distribution of forest formations reflect biogeographic and elevational patterns (Bohn et al., 2007; Dolukhanov, 2010; Nakhutsrishvili, 2013; Wellenbeck et al., 2025).

Although field-assigned severity classes correspond to consistent changes in mean values, the nature and significance of these changes are formation-specific and often non-linear. In line with findings by Meigs et al. (2017), the observed variability of response across variables implies that the nature and intensity of external effect shape structural trajectories of forest formations differently, with similar structural outcomes resulting from different disturbance pathways. However, the trends show a general reduction in overall stocking, increased structural homogenization and loss of stand productivity for all forest formations along the degradation severity gradient (Thompson et al., 2009; Thompson et al., 2013). Whereas compensatory responses at low degradation severity i.e., released radial growth as a reaction to reduced competition and increased regeneration (Čada et al., 2020) suggest resilience dynamics in most formations, significantly reduced stocking and resulting declines in overall volume increment at moderate severity levels point to diminished recovery potential as a result of structural loss (Knapp et al., 2021; Foster et al., 2022). Structural loss is reflected in decreasing mean D_q and increasing mean H_p, indicating reduced size differentiation and vertical homogenization along the degradation gradient (Long, 2009; Goode et al., 2020). Whereas the magnitude of regeneration release is strongly formation-dependent, most formations exhibit increased regeneration under low to moderate degradation severity. As increasing degradation severity results in reduced canopy cover and increased light availability, a corresponding increase in N_{reg} can be observed. This is particularly evident in *Fagus*-dominated formations, where dense canopies typically suppress regeneration (Tinya et al., 2019), but gap dynamics play a key role in facilitating seedling establishment (Feldmann et al., 2018; Zenner et al., 2019; Martin-Benito et al., 2020; Droessler & Wolff, 2023). Coniferous, *Quercus*-, and *Castanea*-dominated formations show regeneration gains at low degradation severity only, with reduction at higher levels and limited significance ($p < 0.05$ for all except *Pinus*). This suggests that while low severity may promote regeneration, higher severity impairs recovery and may exceed resilience thresholds due to additional stand dynamics (e.g., understory dominance, loss of topsoil, etc. Kern et al., 2013).

Building on these findings, we next evaluated Δ value distributions across the degradation severity gradient to assess whether observed dynamics show gradual adjustments or specific thresholds which indicate a transition from disturbance to degradation. Statistical significance in Δ value distribution across forest formations of all variables imply that while low degradation severity generally induces change, these impacts are weaker or non-significant, with 57% of all Δ values reaching significance ($p \leq 0.05$). In contrast, frequency of significance increases to 58% and 75% of responses observed under moderate and high degradation severity, respectively. This severity-dependent pattern suggests that fundamental changes in stocking, structure, and growth do not emerge incrementally but rather manifest once severity exceeds a certain threshold. Hence, a uniform classification of all severities as "degradation" may overlook ecological nuances, as lower severity impacts do not necessarily compromise stand resilience or require management intervention.

Our findings show that responses to degradation differ by forest formation, emphasizing the need to interpret degradation severity in a forest type-specific context that allows for adequate forest policy formulation and targeted management reactions. From a management perspective, it is essential to account for forest formation specific responses, as some forest types potentially exhibit strong resilience to low degradation severity, whereas others show structural decline even under moderate impacts. Sustainable management depends on distinguishing between temporary, self-recovering disturbances and impacts that exceed resilience thresholds, leading to long term forest degradation. However, the nature of degradation types are expected to affect forest structure differently. E.g., in *Castanea*-dominated formations, the high prevalence of pest and pathogen-related disturbances (30% of all samples in this formation) combined with low initial stocking likely contribute to the observed irregularity and reduced specificity of responses (Metreveli et al., 2024). Similarly, *Alnus*-dominated formations, contain both lowland swamp forests and alluvial forests at higher elevation (Figure 4.3), exhibit inconsistent regeneration and structural patterns, which may reflect the inherent variability of this formation and the relatively high grazing pressure (36%) limiting regeneration success despite initial canopy opening. *Abies* and *Picea*-dominated formations, both with above-average shares of erosion-related degradation (~22%), showed steep reductions in stocking, and stand increment under high severity, consistent with expected effects of soil destabilization. *Pinus*-dominated formations, characterized by mixed samples of natural and plantation stands (Lachashvili et al., 2017), display more gradual and heterogeneous structural responses. Moreover, the high incidence of the degradation type "stands of low density" in *Fagus*, *Carpinus*, *Pinus*, and *Quercus*-dominated formations (≥30%) provides context for interpretation of the observed high sensitivity to increasing degradation severity in stocking and spatial structure.

As an observational, sample-based study, this research does not allow causal inference but describes changes in mean values across recorded severity levels. The variation in pattern strength across formations also reflects differences in sample size and within-formation heterogeneity related to site, stand, and species composition. For example, growth indicators are especially variable, reflecting differing site productivity or successional stages. These sources of variation must be considered when interpreting the presented formation-specific responses along the degradation gradient. Hence, our approach inevitably simplifies ecological complexity, as multiple interacting factors —species traits, site conditions, competition, and stand age— influence responses to disturbance and overall resilience. The coarse national sampling grid may obscure fine-scale spatial variation and blend distinct successional trajectories. However, the growing availability of stand structure data from denser permanent plot grids in Georgia enables modeling of forest succession and forest-type specific response dynamics (Pretzsch et al., 2019; Čada et al., 2020; Gough et al., 2022; Patacca et al., 2023). To deepen insight into the explored relationships, effects of combined external factors and influences of stand age, long-term studies examining successional trajectories without and with external effects of varying magnitude are required (Seidl & Rammer, 2017). Further insight could be gained from comparing impacted stands to near-natural reference conditions, as basis for quantifying degradation and the development of forest-type specific stocking guidelines (Bahamondez & Thompson, 2016; Nakhutsrishvili et al., 2023).

With climate-related pressure is assumed to increasingly impact forest ecosystems, conceptual frameworks for sound monitoring of disturbance and degradation that define required management interventions are becoming increasingly important (Martin-Benito et al., 2018; Pretzsch et al., 2020). Integrating remote sensing analyses with quantified structural indicators from field sample assessments to incorporate spatial and temporal scales can provide new insights by combining the synoptic view of satellite data with detailed information from field-based assessments (Senf et al., 2017; Sommerfeld et al., 2018).

Identifying thresholds linked to natural resilience allows for forest-type specific definitions from a structural viewpoint in line with sustainable forest management, focusing on maintaining adaptive capacity in stands that display predictable changes, while addressing degradation in less resilient forest types. Greater recognition of fine-scale natural disturbance processes particular to existing formations enhances targeted management at both the stand level and across broader landscape scales.

4.5 Conclusions

Our findings demonstrate that structural responses to field-assigned degradation severity can be quantified through changes in stand properties, with patterns that are forest formation-specific and non-linear. Low severity often reflects resilience, marked by increased regeneration and stem increment alongside stable or rising productivity in several formations. In contrast, moderate severity is associated with significant declines in stocking and productivity, suggesting a loss of recovery capacity. High severity consistently leads to widespread reductions across all structural indicators. These structural shifts support a conceptual distinction between disturbance and degradation in forest inventory assessments. This supports more targeted, ecologically informed forest policy and silvicultural planning aimed at sustaining multifunctional forest landscapes under increasing anthropogenic and climatic pressures.

4.6 References

Akhalkatsi, M., Arabuli, G., Asanidze, Z., Goloshvili, T. and Togonidze, N. 2019 Pine Forest on Tree-Line Ecotone in the Mountain Kazbegi in the Georgia (South Caucasus). *Agricultural Research & Technology* **21**. 10.19080/ARTOAJ.2019.21.556149.

Akhalkatsi, M., Arabuli, G. and Lorenz, R. 2014 Orchids as indicator species of forest disturbances on limestone quarry in Georgia (South Caucasus). *J. Eur. Orch.* **46**, 123–160.

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

Bahamondez, C. and Thompson, I.D. 2016 Determining forest degradation, ecosystem state and resilience using a standard stand stocking measurement diagram: theory into practice. *Forestry* **89**, 290–300. 10.1093/forestry/cpv052.

Beridze, B. and Dering, M. 2021 Problems and Threats to the Caucasus Forest Ecosystems on the Example of Castanea sativa. *Kosmos* **70**, 19–26. 10.36921/kos.2021_2662.

Bohn, U., Zazanashhvili, N. and Nakhutsrishvili, G. 2007 The Map of the Natural Vegetation of Europe and its application in the Caucasus Ecoregion. *Bulletin of the Georgian National Academy of Science* **175**, 112–119. science.org.ge/old/moambe/2007-vol1/112-120.pdf.

Buchner, J., Yin, H., Frantz, D., Kuemmerle, T., Askerov, E. and Bakuradze, T. et al. 2020 Land-cover change in the Caucasus Mountains since 1987 based on the topographic correction of multi-temporal Landsat composites. *Remote Sensing of Environment* **248**, 111967. 10.1016/j.rse.2020.111967.

Buma, B. and Wessman, C.A. 2011 Disturbance interactions can impact resilience mechanisms of forests. *Ecosphere* **2**, art64. 10.1890/ES11-00038.1.

Čada, V., Trotsiuk, V., Janda, P., Mikoláš, M., Bače, R. and Nagel, T.A. et al. 2020 Quantifying natural disturbances using a large-scale dendrochronological reconstruction to guide forest management. *Ecological Applications* **30**, e02189. 10.1002/eap.2189.

Chalatashvili, A., Mikeladze, G., Gobronidze, N., Metreveli, V., Marsagishvili, N. and Drossler, L. 2024 Knowledge gaps to manage Caucasian forests: What do Georgian forest scientists and foresters want to know and how to answer such questions? *Annals of Agrarian Science* **22**. journals.org.ge/index.php/aans/article/view/393.

Chen, S., Woodcock, C.E., Bullock, E.L., Arévalo, P., Torchinava, P., Peng, S. and Olofsson, P. 2021 Monitoring temperate forest degradation on Google Earth Engine using Landsat time series analysis. *Remote Sensing of Environment* **265**, 112648. 10.1016/j.rse.2021.112648.

Chytrý, M., Tichý, L., Hennekens, S.M., Knollová, I., Janssen, J.A.M. and Rodwell, J.S. et al. 2020 EUNIS Habitat Classification: Expert system, characteristic species combinations and distribution maps of European habitats. *Applied Vegetation Science* **23**, 648–675. 10.1111/avsc.12519.

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

Curtis, R.O. and Marshall, D.D. 2000 Technical Note: Why Quadratic Mean Diameter? *West. J. Appl. For.* **15**, 137–139. 10.1093/wjaf/15.3.137. webpages.uidaho.edu/for274new/pdfs/s6.pdf.

Dees, M., Weinreich, A. and Sperlich, M. 2020 *Data analysis methodology for the national forest inventory of Georgia: Manual (Draft).* UNIQUE land use GmbH, 136 p.

Denk, T., Frotzler, N. and Davitashvili, N. 2001 Vegetational patterns and distribution of relict taxa in humid temperate forests and wetlands of Georgia (Transcaucasia). *Biol J Linn Soc Lond*, 287–332. 10.1006/bijl.2000.0502.

Dolukhanov, A.G. 2010 *Лесная растительность Грузии: (Forest vegetation of Georgia) [in Russian].* Universal, Tbilisi, Georgia. 978-9941-17-176-5. bilib.blogspot.com/2011/12/blog-post_24.html, 484 p.

Droessler, L. and Wolff, T. 2023 Comparison of tree species diversity, deadwood volume and regeneration of managed and old-growth Oriental beech (Fagus orientalis L.) forests in Eastern Georgia. *Caucasiana* **2**, 171–176. 10.3897/caucasiana.2.e106898.

Elizbarashvili, E.S., Chavchanidze, Z.B., Elizbarashvili, M.E., Maglakelidze, R.V., Sulkhanishvili, N.G. and Elizbarashvili, S.E. 2006 Soil-climatic zoning of Georgia. *Eurasian Soil Sc.* **39**, 1062–1065. 10.1134/S1064229306100036.

Feldmann, E., Drößler, L., Hauck, M., Kucbel, S., Pichler, V. and Leuschner, C. 2018 Canopy gap dynamics and tree understory release in a virgin beech forest, Slovakian Carpathians. *Forest Ecology and Management* **415-416**, 38–46. 10.1016/j.foreco.2018.02.022.

Fischer, A., Marshall, P. and Camp, A. 2013 Disturbances in deciduous temperate forest ecosystems of the northern hemisphere: their effects on both recent and future forest development. *Biodivers Conserv* **22**, 1863–1893. 10.1007/s10531-013-0525-1.

Foster, A.C., Wang, J.A., Frost, G.V., Davidson, S.J., Hoy, E. and Turner, K.W. et al. 2022 Disturbances in North American boreal forest and Arctic tundra: impacts, interactions, and responses. *Environ. Res. Lett.* **17**, 113001. 10.1088/1748-9326/ac98d7.

Ghazoul, J., Burivalova, Z., Garcia-Ulloa, J. and King, L.A. 2015 Conceptualizing Forest Degradation. *Trends in Ecology & Evolution* **30**, 622–632. 10.1016/j.tree.2015.08.001.

Goginashvili, N., Togonidze, N., Tvauri, I., Manvelidze, Z., Memiadze, N., Zerbe, S. and Asanidze, Z. 2021 Diversity and degradation of the vegetation of mountain belt forests of central Adjara (the Lesser Caucasus), Georgia. *J. For. Sci.* **67**, 219–241. 10.17221/80/2020-JFS.

Goode, J.D., Barefoot, C.R., Hart, J.L. and Dey, D.C. 2020 Disturbance history, species diversity, and structural complexity of a temperate deciduous forest. *J. For. Res.* **31**, 397–414. 10.1007/s11676-018-0746-y.

Gough, C.M., Atkins, J.W., Fahey, R.T., Curtis, P.S., Bohrer, G. and Hardiman, B.S. et al. 2022 Disturbance has variable effects on the structural complexity of a temperate forest landscape. *Ecological Indicators* **140**, 109004. 10.1016/j.ecolind.2022.109004.

Government of Georgia. 2020 *Forest Code of Georgia*. leap.unep.org/en/countries/ge/national-legislation/forest-code-georgia-no-5949-ss-2020, 60 p.

Griesbach, R. 2018 *Support in planning and implementation of national forest inventory of Georgian forests: Mission Report,* unpublished, 26 p.

Gunn, J.S., Ducey, M.J. and Belair, E. 2019 Evaluating degradation in a North American temperate forest. *Forest Ecology and Management* **432**, 415–426. 10.1016/j.foreco.2018.09.046.

Gutman, G. and Radeloff, V. (eds). 2017 *Land-Cover and Land-Use Changes in Eastern Europe after the Collapse of the Soviet Union in 1991.* 1st edn. Springer International Publishing; Imprint: Springer, Cham. 9783319426389, 1 online resource (VIII, 247 pages 84 illustrations, 65 illustrations in color.).

Hao, M., Ganeshaiah, K.N., Zhang, C., Zhao, X. and Gadow, K. von. 2019 Discriminating among forest communities based on taxonomic, phylogenetic and trait distances. *Forest Ecology and Management* **440**, 40–47. 10.1016/j.foreco.2019.03.006.

IPCC (ed). 2003 *Good practice guidance for land use, land-use change and forestry.* Institute for Global Environmental Strategies (IGES), Hayama, Kanagawa. 4887880030. ipcc-nggip.iges.or.jp, 590 p.

Jin, Y. and Qian, H. 2019 V.PhyloMaker: an R package that can generate very large phylogenies for vascular plants. *Ecography* **42**, 1353–1359. 10.1111/ecog.04434.

Jin, Y. and Qian, H. 2022 V.PhyloMaker2: An updated and enlarged R package that can generate very large phylogenies for vascular plants. *Plant diversity* **44**, 335–339. 10.1016/j.pld.2022.05.005.

Keggenhoff, I., Elizbarashvili, M., Amiri-Farahani, A. and King, L. 2014 Trends in daily temperature and precipitation extremes over Georgia, 1971–2010. *Weather and Climate Extremes* **4**, 75–85. 10.1016/j.wace.2014.05.001.

Keren, S., Svoboda, M., Janda, P. and Nagel, T.A. 2020 Relationships between Structural Indices and Conventional Stand Attributes in an Old-Growth Forest in Southeast Europe. *Forests* **11**, 4. 10.3390/f11010004.

Kern, C.C., D'Amato, A.W. and Strong, T.F. 2013 Diversifying the composition and structure of managed, late-successional forests with harvest gaps: What is the optimal gap size? *Forest Ecology and Management* **304**, 110–120. 10.1016/j.foreco.2013.04.029.

Knapp, E.E., Bernal, A.A., Kane, J.M., Fettig, C.J. and North, M.P. 2021 Variable thinning and prescribed fire influence tree mortality and growth during and after a severe drought. *Forest Ecology and Management* **479**, 118595. 10.1016/j.foreco.2020.118595.

Knollová, I., Chytrý, M., Bruelheide, H., Dullinger, S., Jandt, U. and Bernhardt-Römermann, M. et al. 2024 ReSurveyEurope: A database of resurveyed vegetation plots in Europe. *J Vegetation Science* **35**. 10.1111/jvs.13235.

Kruskal, W.H. and Wallis, W. 1952 Use of Ranks in One-Criterion Variance Analysis. *Journal of the American Statistical Association* **47**, 583–621. 10.1080/01621459.1952.10483441.

Kulakowski, D., Seidl, R., Holeksa, J., Kuuluvainen, T., Nagel, T.A. and Panayotov, M. et al. 2017 A walk on the wild side: Disturbance dynamics and the conservation and management of European mountain forest ecosystems. *Forest Ecology and Management* **388**, 120–131. 10.1016/j.foreco.2016.07.037.

Lachashvili, N., Eradze, N. and Khetsuriani, L. 2020 Floristic composition of pistachio-woodland (Pistacieta atlantici) of Georgia (South Caucasus). *Fl. Medit.* **30**, 39–53. 10.7320/FlMedit30.039.

Lachashvili, N., Eradze, N.V. and Khetsuriani, L.D. 2017 Conspectus of trees and shrubs of Tbilisi environs (East Georgia, South Caucasus). *Annals of Agrarian Science* **15**, 118–129. 10.1016/j.aasci.2016.08.008.

Lomsadze, Z., Chitanava, N., Paresishvili, O., Magalashvili, G., Vezirishvili-Nozadze, K. and Makharadze, K. et al. 2019 Georgian Natural Resources (Brief Review). *Annals of Agrarian Science* **17**, 59–74. journals.org.ge/index.php/aans/article/view/299/226.

Long, J.N. 2009 Emulating natural disturbance regimes as a basis for forest management: A North American view. *Forest Ecology and Management* **257**, 1868–1873. 10.1016/j.foreco.2008.12.019.

Lund, H.G. 2009 *What is a degraded forest?* White Paper on Forest Degradation Definitions Prepared for FAO. Forest Information Services, Gainesville, Virginia, USA., 39 p. http://home.comcast.net/~gyde/2009forest degrade.doc .

Machavariani, M. 2010 *Forestry Standards and Practices in Georgia: Technical Report.* ENPI-FLEG Program, USAID, Tbilisi, Georgia, 46 p.

Martin-Benito, D., Pederson, N., Lanter, C., Köse, N., Doğan, M., Bugmann, H. and Bigler, C. 2020 Disturbances and Climate Drive Structure, Stability, and Growth in Mixed Temperate Old-growth Rainforests in the Caucasus. *Ecosystems* **23**, 1170–1185. 10.1007/s10021-019-00462-x.

Martin-Benito, D., Pederson, N., Köse, N., Doğan, M., Bugmann, H., Mosulishvili, M. and Bigler, C. 2018 Pervasive effects of drought on tree growth across a wide climatic gradient in the temperate forests of the Caucasus. *Global Ecology and Biogeography* **27**, 1314–1325. 10.1111/geb.12799.

McElhinny, C., Gibbons, P., Brack, C. and Bauhus, J. 2005 Forest and woodland stand structural complexity: Its definition and measurement. *Forest Ecology and Management* **218**, 1–24. 10.1016/j.foreco.2005.08.034.

Meigs, G.W., Morrissey, R.C., Bače, R., Chaskovskyy, O., Čada, V. and Després, T. et al. 2017 More ways than one: Mixed-severity disturbance regimes foster structural complexity via multiple developmental pathways. *Forest Ecology and Management* **406**, 410–426. 10.1016/j.foreco.2017.07.051.

MEPA. 2018 *Field Manual for the Georgian National Forest Inventory: Part II (unpublished).* MEPA, Tbilisi, Georgia, 74 p. 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, Tbilisi, Georgia, 123 p. mepa.gov.ge/En/Files/Download/53934.

Metreveli, V., Kreft, H., Akobia, I., Janiashvili, Z., Nonashvili, Z. and Dzadzamia, L. et al. 2023 Potential Distribution and Suitable Habitat for Chestnut (Castanea sativa). *Forests* **14**, 2076. 10.3390/f14102076.

Metreveli, V., Kreft, H. and Gavashelishvili, A. 2024 Environmental covariates of chestnut blight (Cryphonectria parasitica) in Georgia (Caucasus). *Forest Ecology and Management* **569**, 122153. 10.1016/j.foreco.2024.122153.

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

Mittermeier, R.A., Turner, W.R., Larsen, F.W., Brooks, T.M. and Gascon, C. 2011 Global Biodiversity Conservation: The Critical Role of Hotspots. In *Biodiversity Hotspots*. F.E. Zachos and J.C. Habel (eds). Springer Berlin Heidelberg. 978-3-642-20991-8, pp. 3–22.

Modica, G., Merlino, A., Solano, F. and Mercurio, R. 2015 An index for the assessment of degraded Mediterranean forest ecosystems. *For. syst.* **24**, e037. 10.5424/fs/2015243-07855.

Mori, A.S. 2011 Ecosystem management based on natural disturbances: hierarchical context and non-equilibrium paradigm. *Journal of Applied Ecology* **48**, 280–292. 10.1111/j.1365-2664.2010.01956.x.

Mucina, L., Bültmann, H., Dierßen, K., Theurillat, J.-P., Raus, T. and Čarni, A. et al. 2016 Vegetation of Europe: hierarchical floristic classification system of vascular plant, bryophyte, lichen, and algal communities. *Applied Vegetation Science* **19**, 3–264. 10.1111/avsc.12257.

Mumladze, L., Japoshvili, B. and Anderson, E.P. 2019 *Faunal Biodiversity Research in the Republic of Georgia: A Short Review of Trends, Gaps, and Needs in the Caucasus Biodiversity Hotspot*, 18 p.

Myers, N., Mittermeier, R.A., Mittermeier, C.G., Da Fonseca, G.A.B. and Kent, J. 2000 Biodiversity hotspots for conservation priorities. *Nature* **403**, 853–858. 10.1038/35002501.

Nakhutsrishvili, G. (ed). 2013 *The Vegetation of Georgia (South Caucasus): Geobotany Studies.* 1st edn. Springer Science & Business Media, Heidelberg New York Dordrecht London. 978-3-642-29914-8, 236 p.

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

Nakhutsrishvili, G., Abdaladze, O., Batsatsashvili, K. and Dzadzamia, L. (eds). 2023 *Natural Forests of Georgia: (The South Caucasus).* Ilia State University Press, Tbilisi, Georgia. 978-9941-18-445-1, 224 p.

Novák, P., Kalníková, V., Szokala, D., Aleksanyan, A., Batsatsashvili, K. and Fayvush, G. et al. 2023a Transcaucasian Vegetation Database – a phytosociological database of the Southern Caucasus. *VCS* **4**, 231–240. 10.3897/VCS.105521.

Novák, P., Kalníková, V., Večeřa, M., Štětková, G., Sedláček, V. and Pustková, Š. et al. 2023b Alluvial alder forests of the Greater Caucasus, Georgia: ecology, habitats and variability. *Tuexenia* **43**, 1–28. 10.14471/2023.43.003.

Novák, P., Stupar, V. and Kalníková, V. 2021 Carpinus orientalis forests in Georgian Colchis: First insights. *Tuexenia* **41**, 37–51. 10.14471/2021.41.012.

Novák, P., Zukal, D., Harásek, M., Vlčková, P., Abdaladze, O. and Willner, W. 2020 Ecology and vegetation types of oak-hornbeam and ravine forests of the Eastern Greater Caucasus, Georgia. *Folia Geobot* **55**, 333–349. 10.1007/s12224-020-09386-0.

Olofsson, P., Torchinava, P., Woodcock, C.E., Baccini, A., Houghton, R. and Ozdogan, M. et al. 2010 Implications of land use change on the national terrestrial carbon budget of Georgia. *Carbon balance and management* **5**, 1–13. cbmjournal.com/content/5/1/4.

Patacca, M., Lindner, M., Lucas-Borja, M.E., Cordonnier, T., Fidej, G. and Gardiner, B. et al. 2023 Significant increase in natural disturbance impacts on European forests since 1950. *Global change biology* **29**, 1359–1376. 10.1111/gcb.16531.

Patarkalashvili, T. 2016 Some problems of forest management of Georgia. *Annals of Agrarian Science* **14**, 108–113. 10.1016/j.aasci.2016.04.002.

Patarkalashvili, T. 2017 Forest biodiversity of Georgia and endangered plant species. *Annals of Agrarian Science* **15**, 349–351. 10.1016/j.aasci.2017.06.002.

Preislerová, Z., Marcenò, C., Loidi, J., Bonari, G., Borovyk, D. and Gavilán, R.G. et al. 2024 Structural, ecological and biogeographical attributes of European vegetation alliances. *Applied Vegetation Science* **27**. 10.1111/avsc.12766.

Pretzsch, H., Del Río, M., Biber, P., Arcangeli, C., Bielak, K. and Brang, P. et al. 2019 Maintenance of long-term experiments for unique insights into forest growth dynamics and trends: review and perspectives. *Eur J Forest Res* **138**, 165–185. 10.1007/s10342-018-1151-y.

Pretzsch, H., Steckel, M., Heym, M., Biber, P., Ammer, C. and Ehbrecht, M. et al. 2020 Stand growth and structure of mixed-species and monospecific stands of Scots pine (Pinus sylvestris L.) and oak (Q. robur L., Quercus petraea (Matt.) Liebl.) analysed along a productivity gradient through Europe. *Eur J Forest Res* **139**, 349–367. 10.1007/s10342-019-01233-y.

QGIS Development Team. 2009 *QGIS Geographic Information System.* QGIS Development Team. qgis.osgeo.org/.

R Core Team. 2024 *R: A Language and Environment for Statistical Computing.* R Foundation for Statistical Computing, Vienna, Austria. r-project.org/.

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

RStudio Team. 2024 *RStudio: Integrated Development for R.: RStudio.* Posit Software, P. B.C., Boston, MA, US.

Sasaki, N. and Putz, F.E. 2009 Critical need for new definitions of "forest" and "forest degradation" in global climate change agreements. *Conservation Letters* **2**, 226–232. 10.1111/j.1755-263X.2009.00067.x.

Schmidtlein, S., Collison, J. and Pfannendoerfer, R. 2024 *Package 'isopam': Clustering of Sites with Species Data*, 10 p. cran.r-project.org/package=isopam.

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

Seidl, R. and Rammer, W. 2017 Climate change amplifies the interactions between wind and bark beetle disturbances in forest landscapes. *Landscape Ecology* **32**, 1485–1498. 10.1007/s10980-016-0396-4.

Seidl, R., Thom, D., Kautz, M., Martin-Benito, D., Peltoniemi, M. and Vacchiano, G. et al. 2017 Forest disturbances under climate change. *Nature climate change* **7**, 395–402. 10.1038/nclimate3303.

Seidl, R. and Turner, M.G. 2022 Post-disturbance reorganization of forest ecosystems in a changing world. *Proceedings of the National Academy of Sciences of the United States of America* **119**, e2202190119. 10.1073/pnas.2202190119.

Senf, C., Pflugmacher, D., Hostert, P. and Seidl, R. 2017 Using Landsat time series for characterizing forest disturbance dynamics in the coupled human and natural systems of Central Europe. *ISPRS Journal of Photogrammetry and Remote Sensing* **130**, 453–463. 10.1016/j.isprsjprs.2017.07.004.

Shapiro, S.S. and Wilk, M.B. 1965 An Analysis of Variance Test for Normality (Complete Samples). *Biometrika* **52**.

Simula, M. 2009 *Towards Defining Forest Degradation: Comparative Analysis of Existing Definitions.* FAO, 59 p. ardot.fi/Documents_2/Degradationdefintions.pdf.

Sommerfeld, A., Senf, C., Buma, B., D'Amato, A.W., Després, T. and Díaz-Hormazábal, I. et al. 2018 Patterns and drivers of recent disturbances across the temperate forest biome. *Nature communications* **9**, 4355. 10.1038/s41467-018-06788-9.

Stritih, A., Senf, C., Kuemmerle, T., Munteanu, C., Dzadzamia, L. and 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. *Landscape Ecology* **39**. 10.1007/s10980-024-01908-x.

Sturtevant, B.R. and Fortin, M.-J. 2021 Understanding and Modeling Forest Disturbance Interactions at the Landscape Level. *Front. Ecol. Evol.* **9**. 10.3389/fevo.2021.653647.

Tephnadze, N., Abdaladze, O., Nakhutsrishvili, G., Simmering, D., Waldhardt, R. and Otte, A. 2014 The impacts of management and site conditions on the phytodiversity of the upper montane and subalpine belts in the Central Greater Caucasus. *phyto* **44**, 255–291. 10.1127/0340-269X/2014/0044-0579.

Tepley, A.J., Swanson, F.J. and Spies, T.A. 2013 Fire-mediated pathways of stand development in Douglas-fir/western hemlock forests of the Pacific Northwest, USA. *Ecology* **94**, 1729–1743. 10.1890/12-1506.1.

Thom, D. and Seidl, R. 2016 Natural disturbance impacts on ecosystem services and biodiversity in temperate and boreal forests. *Biological reviews of the Cambridge Philosophical Society* **91**, 760–781. 10.1111/brv.12193.

Thompson, I., Guariguata, M.R., Okabe, K., Bahamondez, C., Nasi, R., Heymel, V. and Sabogal, C. 2013 An Operational Framework for Defining and Monitoring Forest Degradation. *Ecology and Society* **18**, 2ß. ecologyandsociety.org/vol18/iss2/art20/.

Thompson, I., Mackey, B., McNulty, S. and Mosseler, A. 2009 *Forest resilience, biodiversity, and climate change: A synthesis of the Biodiversity / Resilience / Stability Relationship in Forest Ecosystems.* Secretariat of the Convention on Biological Diversity, Montreal. 9292251376, 67 p.

Thormann et al. 2019 A prelude to the Caucasus Barcode of Life Platform (CaBOL): Biodiversity Days in Georgia in 2018 and 2019. Report. *Bonn zoological Bulletin* **68**, 275–296.

Tinya, F., Márialigeti, S., Bidló, A. and Ódor, P. 2019 Environmental drivers of the forest regeneration in temperate mixed forests. *Forest Ecology and Management* **433**, 720–728. 10.1016/j.foreco.2018.11.051.

Turner, M.G. 2010 Disturbance and landscape dynamics in a changing world. *Ecology* **91**, 2833–2849. 10.1890/10-0097.1.

UNECE. 2019 *State of Forests in the Caucasus and Central Asia: Geneva Timber and Forest Study Paper*. Overview of forests and sustainable forest management in the Caucasus. United Nations, New York, Geneva. 9789211171600, 75 pages.

Vásquez-Grandón, A., Donoso, P.J. and Gerding, V. 2018 Forest Degradation: When Is a Forest Degraded? *Forests* **9**, 726. 10.3390/f9110726.

Welch, L. 1947 The Generalization of 'Students' Problem When Several Different Population Variances are Involved. *Biometrika* **34**, 28–35.

Wellenbeck, A., Fehrmann, L., Feilhauer, H., Schmidtlein, S., Misof, B. and Hein, N. 2024 Discriminating woody species assemblages from National Forest Inventory data based on phylogeny in Georgia. *Ecology and evolution* **14**. 10.1002/ece3.11569.

Wellenbeck, A., Hein, N., Tarkhnishvili, D., Misof, B., Schmidtlein, S. and Janiashvili, Z. et al. 2025 Predicting woody species assemblages using ecophylogenetics and Earth observation data. *Forest Ecology and Management* **589**. 10.1016/j.foreco.2025.122763.

Willmer, J.N.G., Püttker, T. and Prevedello, J.A. 2022 Global impacts of edge effects on species richness. *Biological Conservation* **272**, 109654. 10.1016/j.biocon.2022.109654. sciencedirect.com/science/article/pii/S0006320722002075.

Zenner, E.K., Peck, J.E. and Sagheb-Talebi, K. 2019 Patchiness in old-growth oriental beech forests across development stages at multiple neighborhood scales. *Eur J Forest Res* **138**, 739–752. 10.1007/s10342-019-01203-4.

Zimina, R.P. 1978 The Main Features of the Caucasian Natural Landscapes and Their Conservation, USSR. *Arctic and Alpine Research* **10**, 479–488. jstor.com/stable/1550782.

5 Synthesis

This thesis investigates the potential of using large-scale forest inventory data to assess woody species diversity by applying phylogenetic information as a biodiversity indicator. Focusing on diversity at the community level, it examines the conceptual and methodological implications of incorporating interspecific phylogenetic relationships into forest classification frameworks. Using data from the GNFI, the study addresses three core objectives: (I) to test the utility of phylogenetically informed assemblages for forest community classification, (II) to evaluate their alignment with environmental and biogeographic gradients, and (III) to investigate how structural variation of grouped species assemblages respond to varying levels of forest degradation (Chapter 1.7).

From a broader perspective, this study bridges standard large-scale forest resource assessments and analytical approaches from community ecology that aim to explain patterns of species assemblage structure and diversity. By integrating phylogenetic information into the classification of woody species communities, it examines the practical relevance of a recent conceptual shift in community ecology using empirical data from a NFI. This multidisciplinary approach highlights the value of accounting for evolutionary relationships in biodiversity research and contributes to a deeper understanding of how forest communities are distributed along environmental and biogeographic gradients. Moreover, it operationalizes PD as a practical component of forest inventory analysis and demonstrates its applicability for classifying compositional data. As the first study to implement this approach in the context of a NFI, it offers a novel perspective on the analysis of tree community variability across large spatial scales. The findings emphasize the potential of shifting from species-centric to phylogenetically structured approaches in ecological classification, opening new avenues for research on how evolutionary relationships reflect functional coherence and ecological fit shaped by environmental gradients. These relationships can serve as indicators of whether species assemblages correspond to environmentally filtered, near-natural conditions or signal altered ecological states due to structural simplification or the introduction of nonnative species.

Although rooted in a multidisciplinary approach, the synthesis presented here frames the interpretation and contextualization of the stated research objectives within the domain of forest science. This perspective highlights emerging research questions and challenges related to general forest assessment, i.e., sample-based representativeness, the integration of biodiversity indicators, and operational feasibility, but also issues specific to the Georgian context. Viewed through the lens of large-scale forest assessment, the following sections discuss the findings with an emphasis on practical insights that can inform the future integration of biodiversity indicators into permanent forest monitoring systems.

5.1 Phylogenetic Diversity in Forest Inventories

Biodiversity is fundamental to both life sciences and ecological conservation, yet it has long been studied in a fragmented manner. Forest science has focused on forest ecosystems, ecology on species, traits, or phylogeny, and genetics emphasizes variation at the gene level. Although biodiversity spans multiple hierarchical levels, from genes to ecosystems, the absence of unified diversity measures has hindered truly integrative research approaches (Gaggiotti et al., 2018; Heydari et al., 2020). This study advances integrative biodiversity research by translating species identities into evolutionary units, enabling differentiation based on phylogenetic relationships within ecological communities. By positioning species identities within evolutionary space, the study applies this integrative concept to forest inventory data from a diverse region characterized by high beta diversity and species turnover. To operationalize PD, concepts from community ecology and forest inventory science are combined to evaluate whether incorporating interspecies evolutionary relationships into classification yields ecologically meaningful results. The methodological novelty lies in incorporating phylogenetic information into the unsupervised classification of woody species assemblages at a national scale. The potential of this approach is demonstrated in Chapter 2 based on comparative cluster analysis of 1059 samples.

In ecology, the difference between pairs of communities is routinely assessed using dissimilarity measures to quantify the degree of community change across time or space (Cáceres et al., 2013). Diversity measures are typically derived from data reflecting the compositional variation within an assemblage, based on species occurrence and abundance (Ricotta et al., 2021). In such approaches, species identities are treated as “species-neutral”, as biological differences between species remain unaddressed, implicitly assuming functional equivalence, meaning that each species contributes equally to the overall observed diversity (Chapter 1.2, Chao et al., 2016). Departing from the traditional assumption that species are functionally or ecologically equivalent, the approach taken here acknowledges that species differ in their evolutionary history and ecological roles. In addition, by using total BA per species as abundance metric, the compositional characteristics of assemblages are quantified by incorporating structural properties that reflect the size of its constituents (Cáceres et al., 2013; Yao et al., 2019). Species composition data based on frequencies alone may be insufficient without a structural component, particularly in woody species communities, where species may be represented by individuals of varying sizes and thus contribute unequally to overall community structure. As a result, ecological communities can exhibit similar species composition in terms of abundance, yet differ substantially in size structure, as reflected by measures such as BA or tree height (H, Cáceres et al., 2019b). In forest inventories, tree size is routinely assessed, typically through direct measurement of DBH, and for a subsample of trees, also H, with the latter being imputed for the full dataset via DBH–H regression models.

Consequently, species assemblage data are complemented by individual-level size information, enabling a more accurate characterization of community structure through structural attributes.

Accordingly, quantifying community dissimilarity using BA weighted by phylogenetic distances integrates both structural variation and evolutionary relatedness into the assessment of diversity. This approach led to the classification of woody species into ecologically cohesive assemblages that reflect not only species co-occurrence, but also relative size distribution and phylogenetic similarity. As these assemblages correspond closely with the known spatial distribution of forest types and major biogeographic gradients, the results suggest a strong phylogenetic signal linked to functional traits and environmental filtering. By incorporating interspecific evolutionary relationships, the approach discriminates forest communities in a way that reflects underlying ecological and biogeographic patterns, using only aggregated basal area per species and phylogenetic information. These results underline the complexity of species communities and emphasize the value of multivariate statistical methods for quantitative analyses of their composition (Legendre & Legendre, 2012; Cáceres et al., 2019a). Operationalizing phylogenetic structure as a measure of diversity in this context demonstrates the potential of forest inventories in monitoring forest biodiversity, but at the same time points to some specific constraints that require further consideration.

Many observed specimens recorded in the GNFI remain taxonomically unresolved. Despite extensive cross-referencing with open-access databases, i.e., World Flora Online (WFO, https://www.worldfloraonline.org/), the Global Biodiversity Information Facility (GBIF, https://www.gbif.org/), and additional references (e.g., Fischer et al., 2018, Lachashvili et al., 2022) taxonomic ambiguities remain. Over the course of study, several taxonomic classifications of recorded species changed in the mentioned online databases. A marked discrepancy was observed between the species identification lists used in GNFI field assessments and the evolving standards of taxonomic harmonization, exemplifying the observed divergence between ecological research and forest expert expertise, already noted by Dolukhanov (2010). This inconsistency highlights the importance of aligning locally recorded species identities with internationally recognized taxonomic backbones, i.e., those used in phylogenetic frameworks (e.g., megaphylogenies) or functional species databases. To address this, comprehensive floristic inventories supported by genetic sequencing such as those undertaken by the CaBOL project within the broader context of the iBOL initiative, are essential for resolving ambiguous taxa and ensuring data compatibility across scientific and applied domains. In this study, the R package WORLDFLORA (Kindt, 2020) was used for semi-automatic species name standardizing of the GNFI list of recorded species, prior to the extraction of species phylogenies (Jin & Qian, 2022). Nonetheless, the taxonomic nomenclature used here is not fully resolved, reflecting the best available information at the time of analysis and remains subject to revision.

Because the primary goal of forest inventories is to estimate timber volume, precise species-level identification has traditionally been a secondary concern. Although the broad spatial coverage may result in high overall species richness, accurate identification of rare or commercially less relevant species is often not prioritized. Consequently, records at the genus level are frequently sufficient for forestry purposes, with non-commercial or infrequent species receiving less attention (Lam & Kleinn, 2008). This bias toward economically relevant species limits biodiversity-focused analyses, particularly in species-rich ecosystems where this can lead to underestimation of diversity and misrepresentation of community composition. As biodiversity gains importance in forest assessments, the demand for taxonomic precision in species identification is increasing. Meeting this demand requires rigorous observer training, standardized species lists, and infrastructures that support laboratory-based identification (Paudel et al., 2021). AI-assisted tools for species identification are becoming increasingly valuable in enhancing accuracy in the field and overall efficiency (Mäder et al., 2021; Schmidt et al., 2022; Jeno et al., 2025).

Understory species composition plays a crucial role in forest classification by capturing fine-scale environmental variation and often serves as an indicator of site conditions (Zilliox & Gosselin, 2014; Kermavnar & Kutnar, 2020). The presented approach, however, is limited to woody species recorded in the canopy layer (DBH ≥ 8 cm). Incorporating understory data into the characterization of PD may reveal more nuanced biogeographic patterns not captured by canopy data alone. Considering understory composition has been shown to enhance the spatial coherence of forest community classifications (Zilliox & Gosselin, 2014). Beyond species composition, understory communities offer additional insights into ecosystem health, responses to forest management, and community assembly processes (Gilbert & Lechowicz, 2004; Tinya et al., 2019; Kermavnar & Kutnar, 2020). This is particularly relevant to the present case study, as the composition of woody understory species is already used in standard phytosociological classifications to distinguish forest types along prevailing biogeographic gradients in Georgia (Dolukhanov, 2010; Nakhutsrishvili, 2013). However, the sampling procedures in place treat canopy and understory species (shrubs) as separate entities and yield compositional data on different scales, which constrains the effective combination of observations. Accordingly, assessment protocols must be adjusted to provide standardized abundance metrics across vegetation layers and enable data harmonization. Using standardized species abundance matrices encompassing compositional properties across layers would increase the resolution of the respective phylogenetic profiles and enable the reconstruction of complex evolutionary relationships among communities and environmental gradients. Future research should refine means to capture phylogenetic variability across vegetation layers. Extending the recorded phylogenetic profiles to additional levels of the existing plant communities would significantly improve the resolution of functional community patterns across environmental gradients. The resulting sharpness would allow the influences of natural and anthropogenic species divergence, as well as the impact of neophytes on the phylogenetic composition of the forest community, to be considered more explicitly.

Sampling optimization remains a central concern for forest inventories. Changing information objectives require adjusted sampling strategies that balance overall costs and accuracy requirements. While NFIs are traditionally optimized to produce robust estimates of timber volumes and growing stock, they can also support species richness estimates, provided that their structural and methodological limitations are acknowledged. To minimize efforts for field sampling while maintaining acceptable estimation accuracy, NFIs typically employ nested plot designs to balance the accurate volume estimates of larger trees with the efficient recording of smaller individuals across varying DBH thresholds (Gadow et al., 2021). Lin et al. (2020) show that standard NFI designs tend to underrepresent rare and small-diameter species, especially in species-rich forests of the tropics. Because forest diversity spans across genetic, taxonomic, and structural dimensions, different accuracy requirements need to be reconciled. As NFIs gain recognition for their role in biodiversity monitoring, efforts to optimize sampling design should be grounded in more holistic concepts. For example, *Zhao et al. (2022)* demonstrate that cluster-based sample plot designs, especially square configurations with ~500 m² subplots, can substantially improve accuracy and precision of species richness estimates compared to single-plot designs in temperate forests. Whereas total species richness can be reliably estimated using statistical approaches that account for undetected species, finding a practical balance between accuracy and efficiency in nested plot designs is essential, to harmonize sampling optimization of various dimensions of diversity (Archaux et al., 2009; Winter et al., 2012). This calls for further research on sampling design optimization for multivariate objectives in forest inventory in the context of biodiversity assessments (Lin et al., 2020).

Structuring forest complexity is essential for effective forest monitoring and management. This study demonstrates how a conceptually novel method of forest classification aligned with ecosystem-based forest management can be applied to forest inventory data. This approach reconceptualizes species identity within a phylogenetic framework, where evolutionary relatedness informs ecological interpretation and classification, in line with biodiversity concepts that integrate genetic diversity. For forest science, this framework enhances the ecological interpretation of forest types by shedding light on the dynamics that shape contemporary woody species communities in an evolutionary context (Wilcox et al., 2018). For management and conservation, it allows the identification of zones with disproportionately high PD worthy of protection or evolutionarily unique stands that can justify adapted management measures. By incorporating phylogenetic information, this approach enables the discrimination of woody species assemblages within a broader ecological context, ultimately supporting systematic delineation of forest ecosystems, with improved alignment according to ecological dynamics.

In this context, NFIs are shown to be well-suited for biodiversity classification using unsupervised phylogenetically informed clustering. The patterns resulting from the cluster analysis prove to be ecologically sound at national level, and no additional data collection is required for the approach. Incorporating PD into classification approaches yields a significantly more nuanced and ecologically grounded representation of forest diversity distribution and allows comparative analyses across regions, provided data collection protocols are harmonized. While standardized integration of PD into forest inventories and respective analyses still presents methodological and practical challenges, it offers a feasible and meaningful step toward a more evolutionarily informed approach to understanding the dynamics of these ecosystems. This has great implications for forest management as this conceptual shift in organizing communities allows deeper insight in the interdependence of biotic and abiotic factors, which is highly relevant for ecosystem-based approaches.

5.2 Ecophylogenetics in Forest Classification and Management

Accurate classification of forest communities, and the species assemblages they comprise, is key to biodiversity assessment, forest management and landscape mapping. Since vegetation classification enables the organization of the complexity of ecosystem dynamics and related functioning, robust and standardized methods for classification are required (Gadow et al., 2021). Organizing forest entities thematically based on specific objectives facilitates the description, comparison and meaningful communication required to understand the dynamics that shape forest diversity on temporal or spatial scales. Recognized vegetation typologies support consistent approaches to mapping, monitoring, management, and conservation of ecosystem diversity. Thematic classifications can be based on functional, structural, compositional, and ecological criteria, as well as on conservation priorities and degradation status, depending on the specific objectives. The field of ecophylogenetics investigates how evolutionary relationships among species shape community distributions along environmental gradients. Hence, when applied to forest classification, it combines ecological criteria with an evolutionary framework, based on genetic divergence that serves a proxy for functional characteristics. Grounded in this conceptual framework, this research introduces a scalable, data-driven approach that integrates forest inventory data with phylogenetic dissimilarity metrics to classify woody species compositions. Applying a straightforward approach of unsupervised classification to NFI data, it demonstrates that genetically informed clustering results in coherent and ecologically meaningful classification of forest species assemblages (Chapter 2). The results further suggest that evolutionary relationships, expressed as phylogenetic dissimilarity, strongly shapes woody species composition. This aligns with Pereira et al. (2021), who argue that incorporating evolutionary history is essential for understanding species composition in transitional zones in the Neotropics. However, increased disturbance (anthropogenic or natural) leads to compositional homogenization and reduced clustering accuracy, a trend further intensified by the presence of neophytes, which can obscure the relationship between phylogenetic composition and environmental factors (Chapter 3). Field-assigned disturbance classes lead to composition-specific structural responses. Whereas low disturbance severity often indicates composition-specific resilience, moderate and high severity levels are linked to declining structure and productivity, reflecting reduced recovery capacity and lasting degradation (Chapter 3). Consequently, while constrained by degree of disturbance and presence of nonnative species, clusters that represent woody species assemblages closely align with major environmental gradients in Georgia. Understanding the underlying processes that shape assemblage distribution can serve as framework to delineate forest species assemblages to support conservation mapping and the development of effective ecosystem monitoring strategies.

The forests of Georgia are distinct within the temperate biome for their high species richness, endemism, and survival of ancient lineages, making them important repositories of evolutionary history (Tarkhnishvili, 2014). Phylogenetic structure in these forests varies along environmental gradients but is also markedly shaped by external effects and increasingly by species introductions (Kavtaradze et al., 2023). Despite these alterations, the general natural distribution of species remains relatively intact, compared to other temperate forest areas, e.g., in central Europe, where large areas have been replanted with nonnative or species that would not occur by natural assembly on specific site (e.g., *Picea abies* (L.) H. Karst. or *Pseudotsuga menziesii* (Mirb.) Franco, Vasadze et al., 2024).

This study demonstrates how interspecific phylogenetic structure reflects the genetic landscape, as evidenced by the spatial clustering of evolutionarily related species assemblages. By parting with traditional forest classification systems that are based on species dominance, this approach produces compositionally coherent units grounded in evolutionary relationships and ecological function. Since closely related species often share similar ecological traits due to evolutionary proximity, a strong phylogenetic signal indicates limited functional divergence (Cadotte et al., 2019). In contrast, distantly related species often display greater niche differentiation, shaped by evolutionary and ecological processes that reduce competition and allow coexistence within communities (Padullés Cubino et al., 2021; Guevara Andino et al., 2021; E-Vojtkó et al., 2023). Hence if traits are evolutionarily conserved, a phylogenetically informed clustering yields species assemblages as functional groups, shaped by niche conservatism and environmental filtering. However, external effects can lead to disruption of the compositional assembly impacting structural or phylogenetic properties, e.g., the introduction of nonnative, evolutionarily distant taxa. As a result, the original position of the community in functional trait space may become distorted, leading to weaker phylogenetic signal. Phylogenetic dissimilarity reflects niche differentiation because it captures the relationship between species traits and environmental conditions. However, this balance between community composition and phylogenetic structure becomes increasingly disrupted under disturbance, which weakens the broader ecological signal (Webb et al., 2006; Cadotte & Tucker, 2017; Diniz et al., 2024). On the other hand, a strong phylogenetic signal may indicate increased resilience as environmental filtering and competition dynamics result in stable communities over time, fluctuating along successional dynamics or in response to changing environmental conditions.

Because assembly mechanisms vary with spatial scale, different phylogenetic patterns are expected to emerge across different spatial extents (Cavender-Bares et al., 2006). At smaller scales, biotic interactions such as competition for limited resources, pathogen pressure, or facilitation may lead to phylogenetic and functional trait overdispersion. In contrast, at broader spatial scales, phylogenetic clustering is often observed due to environmental filtering, where varying abiotic conditions favor the establishment of ecologically similar, closely related species (Tilman et al., 1997; Swenson et al., 2006). The relative contributions of these compositional mechanisms, as well as the levels at which they operate, remain uncertain, especially at the intermediate level where both competition and environmental filtering may operate simultaneously (Arnillas et al., 2021). Understanding the relative processes shaping community assembly across scales remains a critical research challenge (Vellend et al., 2017). Moreover, forest community composition is inherently dynamic. As environmental conditions change and interspecific interactions evolve, communities undergo succession, often accompanied by significant changes in species diversity and composition (Yu et al., 2019). Early-successional stages are typically dominated by fast-growing, often functionally similar species, while late-successional or climax communities tend to support fewer, more specialized species (Cadotte, 2023). The classification approach proposed here is designed to accommodate such dynamics. Through repeated measurements of permanent sample plots, it offers a framework to explore community assembly processes along successional trajectories under varying ecological conditions.

Although the classification algorithm produced results broadly consistent with syntaxonomic classifications, the resulting species assemblages have not been systematically aligned with existing vegetation typologies. However, alignment with broad-scale mapping by Bohn et al. (2007) is demonstrated in Chapter 2.3.4. The proposed approach would need to encompass additional forest strata (i.e., understory and herb layer) to produce clusters that could be harmonized with the existing classifications, particularly the detailed forest vegetation descriptions by Nakhutsrishvili (2013). Extending diversity assessments beyond woody species in the canopy layer is likely to enhance the resolution of the captured phylogenetic-environment relationship. Species occupancy in the herb layer is a classical tool for forest site evaluation, as indicator values reflect stand characteristics i.e., soil humidity, nitrogen availability, light conditions and pH value (Ellenberg, 1974; Wohlgemuth, 2000; Di Biase et al., 2023; Tichý et al., 2023). Incorporating this layer enables more fine-scaled habitat differentiation.

In addition, classification of samples from denser grid configurations is required to capture additional forest types and variations thereof. The presented approach is directly applicable with the existing datasets of denser FMIs without needing further adjustments. At smaller scales however, the integration of additional forest strata might be required to capture functional properties of species assemblages sufficiently (Luo et al., 2019). In addition, different dispersion dynamics of dominant and nondominant species should be considered to account for potential effects on site characteristics caused by dominants (Arnillas et al., 2021).

Unlike traditional classifications, which often lack standardized methods, statistical rigor and formal assignment rules, the methodological framework presented here is transparent and reproducible and supports a seamless integration of new plot records, provided complete species lists are available. The approach is compatible with the increasingly available data from FMIs and the upcoming second NFI in Georgia (2028). In fact, it can be applied in a straightforward manner to most forest inventories. Integrating a phylogenetically informed classification into forest assessments offers a robust foundation for ecologically meaningful stratification. The resulting hierarchy can serve as stratification units for adjusted sampling designs and improve the precision and representativeness of diversity estimates across heterogeneous forest landscapes. Such stratified sampling would increase statistical efficiency by reducing within-stratum variability and enhancing detection of β-diversity among strata (Gimaret-Carpentier et al., 1998). Based on these strata, targeted sampling protocols, e.g., habitat classifications, could be employed. This would allow NFIs not only to quantify PD more accurately, but also to track changes in forest PD and functioning over time.

The findings of this study highlight the potential to strengthen links between systematic, periodically repeated forest assessments and research on species diversity and forest ecology. Refining the resulting classification to incorporate additional vegetation strata allows a more complete capture of the phylogenetic composition of the vegetation community, thereby increasing both the diversity facets monitored and the accuracy of the resulting forest typologies. Controlling for forest typology via permanent monitoring provides critical information on structural and compositional changes across successional trajectories and allows evaluation of the most important factors driving respective changes, including the role of disturbances and types of degradation (Sánchez-Pinillos et al., 2019; Sánchez-Pinillos et al., 2024).

For forest monitoring and growth modeling, the approach allows unsupervised classification of diverse forests over large areas. Plot-based assessments are not suited for detailed characterization and classifications at stand level, but they are well suited to characterize broader populations of interest, such as stratified forest types or formations. Thus, application of a phylogenetic stratification could provide reliable estimates, including confidence intervals to quantify statistical uncertainty, to support forest policy, planning, and decision-making (Bruciamacchie, 2015). The typologies obtained can provide forest managers and policy makers with critical information for adaptive forest and targeted conservation management. Based on these typologies, specific management guidelines can be developed to maintain either a particular species composition or a desired size structure. In addition, dynamic growth models that are adjusted to site-and-community interactions can be derived per forest type, providing means to simulate forest development over time. From these, forest management scenarios can be deduced for specific cost-benefit analyses or estimates of economic framework conditions. Quantifying structural and compositional differences between managed and unmanaged stands can be used to detect deviations from natural succession for degradation monitoring (Cáceres et al., 2019a; Cáceres et al., 2019b).

Hence, by its' dynamic nature the approach can align temporal changes in stands of similar phylogenetic configuration, that result from succession, disturbance, degradation, or recovery. Importantly, the classification integrates functional responses of species assemblages with abiotic site characteristics and allows to derive how environmental filtering and species interactions jointly shape the phylogenetic community structure. Moreover, by demonstrating that abiotic variables predict the occurrence of phylogenetically defined forest units, the study provides a tool for spatial planning and scenario modeling, including anticipating how assemblages may shift under climate change or disturbance regimes. This dual consideration strengthens its relevance for site-specific evaluation for forest science. Crucially, it also enables a more meaningful classification of forest units: rather than relying on arbitrary or traditional categories, i.e., groupings by genus or age, the approach delineates ecologically cohesive forest types that reflect functional relationships and environmental context. This could improve both the ecological validity and practical utility of forest classification methods that embody a holistic, multidisciplinary understanding of ecosystem dynamics and site conditions.

However, a practical integration of PD into forest management requires a thorough understanding of these processes and is often constraint by data deficiencies, and a lack of a standardized methodologies. Nonetheless, increasing the awareness among practitioners holds great potential for developing conservation and management strategies that safeguard not only species richness but also the evolutionary heritage and adaptive capacity of forest ecosystems (Ali et al., 2025). In this regard, the approach aligns closely with the principles of ecosystem-based forestry, which emphasize maintaining ecological function, structural diversity, and resilience. Moreover, it resonates with the concept of biogeocoenosis developed by Sukachev and foundational to forestry in many states of the former USSR (Petzer, 2023). This concept views forests as integrated systems shaped by the interaction of organisms and their physical environment (Sukachev & Dylis, 1964). By classifying forest types based on both structural and phylogenetic composition and environmental context, the method operationalizes this holistic ecological view through a modern, data-driven framework.

Nonetheless, implementing this approach requires a shift in forestry practice and education toward increased ecological integration. This would require to streamline forest science with ecological management principles, moving away from production-focused models toward approaches grounded in ecological understanding. The method connects forestry with forest ecology, allowing for classifications and decisions based on the inherent dynamics of forest ecosystems. Hence, future foresters must be equipped not only with silvicultural expertise, but also with an understanding of the underlying ecological processes, including phylogenetic and successional dynamics, the ability to apply insights from long-term monitoring, and forest management strategies that support biodiversity, ecosystem stability, and evolutionary potential. To achieve this, forest science must adjust its methodological foundations to align with ecological and sustainable forest management and strengthen the ecological aspect of forestry. By applying ecological principles and integrative approaches, new methods can be developed that reflect a holistic, multidisciplinary understanding of ecosystem dynamics and their interactions with the abiotic environment.

5.3 Forest Inventories as Biodiversity Monitoring Systems

Continuous pressures on maintaining biodiversity in the Anthropocene demand effective planning in forest conservation, restoration, and diversity management. Related decisions, however, must be grounded in robust, context-specific knowledge derived from empirical data that aligns with the specific goals and information needs. Ecological models are central to this effort, but their utility depends on a solid empirical understanding of ecosystem dynamics, obtained through data analysis and experimental research. The inherent complexity of the natural processes that shape biodiversity is increasingly explored through multidisciplinary concepts and multivariate approaches that foster a more holistic understanding of related ecosystem dynamics. Policy and management efforts aimed at sustainable forest ecosystems and halting biodiversity decline require robust monitoring systems that generate the empirical data necessary for refining concepts and ecological models capable of capturing biodiversity patterns and related dynamics over time. Developing operational methodologies that structure and quantify biodiversity across spatial and temporal scales is a priority not only for conservation management, especially in the current setting in which species loss outpaces documentation (Bevilacqua et al., 2021). Hence, there is a recognized global need for observation systems that provide consistent, and repeated data on the state and trends of forest biodiversity (Bellingham et al., 2020). However, such systems remain scarce, particularly at national scales, with NFIs representing an exception (Gillerot et al., 2021). Systematic forest inventories, i.e., NFIs, have established themselves as the most experienced and data-rich tools to provide extensive information on woody species diversity in forest ecosystems, with repeated measurements that allow tracking of changes over time and across environmental gradients (Hao et al., 2021). Data generated by these assessments are increasingly available worldwide and provide means to quantify various aspects of forest diversity i.e., structural complexity, and temporal biodiversity changes in forest ecosystems (Winter et al., 2008; Corona et al., 2011; Traub & Wüest, 2020; Heym et al., 2021). Building on NFIs as solid frameworks in place to include the monitoring of additional variables related to forest biodiversity beyond woody species represents an effective strategy for maximizing their utility.

In this context, the concept of Essential Biodiversity Variables (EBVs) was introduced to support consistent, scalable biodiversity monitoring and to bridge the gap between raw ecological observations and global policy indicators (Pereira et al., 2013; Proença et al., 2017; Hardisty et al., 2019). EBVs define a minimum set of measurements needed to track biodiversity change across space and time and are structured into six major classes: genetic composition, species populations, species traits, community composition, ecosystem structure, and ecosystem function (Schmeller et al., 2017; Bellingham et al., 2020). These variables are not only scientifically grounded but also designed to inform national and international biodiversity targets, including those defined by the UN-CBD (Pereira et al., 2013; Proença et al., 2017). NFIs offer a promising foundation for operationalizing EBVs at broad scales. Their standardized sampling frameworks, large spatial extent, and capacity for temporal replication make them particularly well-suited to support EBVs related to species populations, community composition, and ecosystem structure (Bellingham et al., 2020). However, many existing biodiversity monitoring systems underutilize this potential, especially in relation to less frequently addressed dimensions such as genetic, phylogenetic, and functional diversity. For instance, PD, can be derived directly from standardized species lists and the increasing number of phylogenies readily available online. As a result, the classification approach developed in this research offers a direct contribution to EBV-oriented biodiversity monitoring. By clustering forest inventory plots into phylogenetically structured woody species assemblages, the presented methodology captures aspects of both community composition and evolutionary diversity. Especially in ecologically complex and understudied regions such as the South Caucasus, this approach supports the development of scalable, reproducible tools that align national monitoring systems with international biodiversity targets.

While contributing to the understanding of distribution patterns of phylogenetic dissimilarity across environmental gradients, this study also demonstrates that despite being originally designed for timber-oriented assessments, NFI frameworks allow extended biodiversity research. By integrating phylogenetic information, it is possible to obtain a highly differentiated picture of forest diversity patterns without the need for additional field assessments. In addition, because of their systematic nature NFIs allow the integration of remote sensing as auxiliary data to calculate relevant indicators and as carrier data for large area predictions, regionalization or the support of species distribution models (Randin et al., 2020). In combination with the spectral diversity concept, PD can be mapped using spatially explicit sample plot observations as training data (Kacic & Kuenzer, 2022). Mapping PD would delineate areas based on evolutionary diversity enabling the identification of conservation priorities beyond species richness by highlighting areas of high ecological and evolutionary value. Once species identity is quantified in terms of interspecies distances, additional metrics can be derived to quantify PD within communities (e.g., Faith’s PD, MPD, MNTD), from which phylogenetic landscapes can be mapped. In addition, spatially explicit assessments of ecosystem functionality, species-level importance, and extinction risk across taxa is possible (Ma et al., 2019; Mammola et al., 2021; Ette et al., 2023).

For forest policy and management in Georgia, grouping woody species assemblages by genetic similarity offers a standardized, reproducible, and scalable tool to organize ecological complexity and delineate forest types according to ecological and phylogenetic relevance. This is particularly important for monitoring trends that affect biodiversity on various ecological levels because of environmental change or anthropogenic factors, enabling more refined diagnostics and early warning systems. Robust classifications of tree species communities based on a multidimensional diversity approach provide broadened perspectives on forest assessment and the derivation of conservation strategies. Applied at smaller scales, i.e., on forest unit or district levels, the approach could support organization of communities on a finer scale, provided compositional data of additional vegetation strata i.e., understory or herb layer are considered. Whereas empirical evidence supporting this assumption is not yet available, the required data is, given that numerous FMIs with finer spatial resolutions (e.g., 1.8 x 1.8 km and 300 x 300 m, in Akhmeta forest district and Maghraani forest unit, respectively) have been implemented in Georgia. These inventories utilize single sample plot designs with identical configurations, making the subsequent comparative analyses across scales straightforward.

To protect biodiversity within the constraints of limited conservation funding, it is essential to prioritize conservation efforts (Cadotte & Tucker, 2018). The concept of biodiversity hotspots provides a strategy for conserving the greatest biodiversity at the least costs (Myers et al., 2000; Lovett, 2001). Traditionally, this method relies mainly on species distribution patterns to determine which areas are most concentrated with species distribution, which areas have the most obvious endemism, and which areas have the most endangered species (Myers, 2003). By providing a focus on the evolutionary history of different species, this study presents a scalable methodology to identify hotspots of PD across the forest ecosystems. Moreover, the applied approach can support species distribution models and be extended to derive phylogenetically inferred biotic species dependencies which can be linked to hosts, prey, predators or other actors to inform distribution or abundance predictions of nontree taxa (Jetz et al., 2019).

This study demonstrates the added value that NFIs provide for generating information relevant to biodiversity assessment on large scales. NFIs have historically emphasized timber resources, with data collection frameworks primarily structured around the measurement and assessment of trees. This production-oriented legacy characterizes the general sampling designs in many NFIs today, resulting in a disproportionate focus on arboreal components. From a biodiversity perspective however, forests are dynamic networks of interacting species of various taxonomic groups, and the functional contributions of understory plants, fungi, and fauna are no less vital than those of trees. Accordingly, for forest biodiversity monitoring to be ecologically inclusive, it must shift beyond woody species of economic value and address the full spectrum of life forms within forest ecosystems. As a result, biodiversity integration in NFIs is increasingly becoming a matter of concern and options exist to build on indicator variables already collected or easily observed within standard inventory protocols (e.g., forest structure, deadwood, tree microhabitats, large and/or old trees, species diversity, Oettel & Lapin, 2021). Based on these field observations additional indicators for extended biodiversity assessments can be modeled. Still, while these provide a relatively comprehensive set of harmonized indicators for tree-based biodiversity (Winter et al., 2008; McRoberts et al., 2009), significant gaps remain regarding non-tree taxa, but complementary sampling methods (i.e., nest counts, animal sightings, etc.), to capture specific features of diversity beyond woody species are being developed and implemented. Nonetheless, the possibility for reporting harmonized estimates of diversity indicators for nontree life-forms derived from NFI data is very limited in practice. A more integrative monitoring strategy would address these gaps by incorporating standardized assessments across multiple biological groups (e.g., lichens and bryophytes*)*. While understory vegetation is frequently recorded as nontree vegetation, assessment protocols are neither optimized nor standardized. Records of ground vegetation include presence or coverage of herbs, ferns, mosses, lichens, and liverworts, sometimes quantified using Braun-Blanquet (1965) approaches (Chirici et al., 2011). Hence, numerous indicators that can be leveraged for biodiversity assessments beyond woody species are already in place in many forest inventories. These indicators are of extended value for scientists outside the forestry domain. However, conventional monitoring often emphasizes species richness or abundance, metrics that do not reflect evolutionary distinctiveness, trait diversity, or the ecological coherence of species assemblages. Embedding phylogenetic metrics into regular inventory cycles allows to incorporate genetic diversity to be assessed and provides means to delineate communities based on PD, understand processes of lineage diversification and speciation, detect compositional shifts that may affect ecosystem function, and infer habitat distributions and environmental filtering dynamics. By identifying phylogenetically clustered forest units, monitoring programs can track changes in PD over time, compare regions or forest formations based on their evolutionary uniqueness, and prioritize areas where the risk of loss of evolutionary history is most severe. In addition, the integration of genetic data, i.e., through DNA barcoding, into NFIs remains a highly promising step that could expand methodologies for detecting intraspecific genetic patterns and enable more detailed classifications. A synergistic expansion of species detection using systematic DNA barcoding methods within forest inventories holds significant potential to increase the accuracy and completeness of data for estimating forest diversity. A systematic evaluation of existing forest inventories, taking into account available integration options, remains an exciting field of investigation.

Hence, I assume that NFI frameworks will continue to adjust to the information needs required for monitoring trends in forest biodiversity with increasing focus beyond woody species. Relatively straight-forward approaches include the incorporation of systematic assessments of ground vegetation based on standard phytosociological methods (i.e., Braun-Blanquet), species counts on larger observational units or recordings of birds and insects within or among sampling units (i.e., record sightings along transects between sample plots). Other approaches could be based on acoustic monitoring (Alcocer et al., 2022; Beason et al., 2023; Lapin et al., 2025), installations of camera traps and/or adding the acquisition of samples (i.e., topsoil) for subsequent DNA sequencing. However, any approach is limited by logistical constraints and cost implications need to be juxtaposed with the respective precision requirements. Here, additional research focusing on required inventory design extension and optimization related to a defined set of biodiversity indicators is required.

Research on the relationship between the phylogenetic dimension of species identity and the resulting diversity encoded in species assemblages in relation to species assembly processes and environmental filtering, has experienced a remarkable upswing in the last decades. Coined as ecophylogenetics, the discipline that emerged ensuite helps to answer fundamental ecological questions by considering evolutionary relationships between species and presents potential to derive orientations for policy making and management decisions, based on a deeper understanding of ecological dynamics shaping forest ecosystems. By applying ecophylogenetic concepts to GNFI, this thesis provides new findings that answer questions regarding the spatial patterns of species assemblages in forest ecosystems within Georgia and reflects several theories regarding community assembly and environmental and biotic interactions. Due to the systematic nature of the available large dataset along multiple environmental gradients and available online phylogenies it was possible to use PD to classify sample plots into meaningful ecological groups and to prove their spatial distribution aligns with prevailing environmental gradients. Overall, this thesis underlines the importance of large-scale systematic forest inventories for biodiversity monitoring of forest ecosystems.

Experience gained from the three presented studies suggest that community delineation from large-scale forest inventories based on PD is possible and produces clusters that align with biogeographic gradients. The non-stochastic relationship between abiotic site conditions and species assemblage patterns was demonstrated by the predictive analysis to derive cluster membership from environmental data. Comparing structural characteristics of phylogenetically homogeneous forest formations revealed differences among groups in responses along a degradation gradient. Thus, tested on a large multivariate dataset by comparing different aspects of cluster characteristics of distinct class configurations and applying predictive analysis via machine learning, the practicability and usefulness of the proposed approach is demonstrated. The growing availability of high-resolution phylogenies and enhanced capacities to process large multivariate datasets facilitate the easy replication and application of these analyses to additional, expanded, or future datasets (Cadotte et al., 2018). As additional samples from similar inventories with finer grid sizes are available in the country, the approach can be readily applied to gain a better understanding of these relationships on smaller spatial scales.

In conclusion, the ecophylogenetic analysis of forest inventory data offers a scalable, data-driven method for biodiversity assessment, classification, and monitoring. It bridges a conceptual divide between evolutionary ecology and forest science, linking the functional relevance of species identity to environmental gradients. The challenge now lies in deepening this perspective within forest management and respective inventory systems, to develop additional assessment protocols and appropriate methods, and to further operationalize the unveiled relationship between phylogenetic community composition and abiotic conditions. Given Georgia's high ecological complexity and largely natural species composition, the country provides an ideal setting to further develop and apply ecophylogenetic approaches. Located within the Caucasus biodiversity hotspot, Georgia’s forests offer a valuable opportunity to advance integrative biodiversity assessment and safeguard their evolutionary and ecological distinctiveness. Addressing these questions and objectives requires experimental and long term observational datasets, solid information management systems with respective mandates enclosed within a strong institutional framework that combines forest science and ecological expertise.

5.4 References

Alcocer, I., Lima, H., Sugai, L.S.M. and Llusia, D. 2022 Acoustic indices as proxies for biodiversity: a meta-analysis. *Biological reviews of the Cambridge Philosophical Society* **97**, 2209–2236. 10.1111/brv.12890.

Ali, S., Amin, A., Akhtar, M.S. and Zaman, W. 2025 Phylogenetic Diversity in Forests: Insights into Evolutionary Patterns and Conservation Strategies. *Forests* **16**, 1004. 10.3390/f16061004.

Archaux, F., Camaret, S., Dupouey, J.-L., Ulrich, E., Corcket, E. and Bourjot, L. et al. 2009 Can we reliably estimate species richness with large plots? an assessment through calibration training. *Plant Ecol* **203**, 303–315. 10.1007/s11258-008-9551-6.

Arnillas, C.A., Borer, E.T., Seabloom, E.W., Alberti, J., Baez, S. and Bakker, J.D. et al. 2021 Opposing community assembly patterns for dominant and nondominant plant species in herbaceous ecosystems globally. *Ecology and evolution* **11**, 17744–17761. 10.1002/ece3.8266.

Beason, R.D., Riesch, R. and Koricheva, J. 2023 Investigating the effects of tree species diversity and relative density on bird species richness with acoustic indices. *Ecological Indicators* **154**, 110652. 10.1016/j.ecolind.2023.110652.

Bellingham, P.J., Richardson, S.J., Gormley, A.M., Allen, R.B., Cook, A. and Crisp, P.N. et al. 2020 Implementing integrated measurements of Essential Biodiversity Variables at a national scale. *Ecological Solutions and Evidence* **1**. 10.1002/2688-8319.12025.

Bevilacqua, S., Anderson, M.J., Ugland, K.I., Somerfield, P.J. and Terlizzi, A. 2021 The use of taxonomic relationships among species in applied ecological research: Baseline, steps forward and future challenges. *Austral Ecology* **46**, 950–964. 10.1111/aec.13061.

Bohn, U., Zazanashhvili, N. and Nakhutsrishvili, G. 2007 The Map of the Natural Vegetation of Europe and its application in the Caucasus Ecoregion. *Bulletin of the Georgian National Academy of Science* **175**, 112–119. science.org.ge/old/moambe/2007-vol1/112-120.pdf.

Bruciamacchie, M. 2015 Inscrire la production forestière dans le fonctionnement de l'écosystème. [in French]. *Forêt Méditerranéenne* **XXXVI**, 393–396. hal.science/hal-03556651v1.

Cáceres, M. de, Coll, L., Legendre, P., Allen, R.B., Wiser, S.K. and Fortin, M.-J. 2019a Trajectory analysis in community ecology. *Ecological Monographs* **89**. 10.1002/ecm.1350.

Cáceres, M. de, Legendre and He, F. 2013 Dissimilarity measurements and the size structure of ecological communities. *Methods Ecol Evol* **4**, 1167–1177. 10.1111/2041-210X.12116.

Cáceres, M. de, Martín-Alcón, S., González-Olabarria, J.R. and Coll, L. 2019b A general method for the classification of forest stands using species composition and vertical and horizontal structure. *Annals of Forest Science* **76**, 1–19. 10.1007/s13595-019-0824-0.

Cadotte, M.W. 2023 The interacting influences of competition, composition and diversity determine successional community change. *Journal of Ecology* **111**, 1670–1680. 10.1111/1365-2745.14135.

Cadotte, M.W., Campbell, S.E., Li, S., Sodhi, D.S. and Mandrak, N.E. 2018 Preadaptation and Naturalization of Nonnative Species: Darwin’s Two Fundamental Insights into Species Invasion. *Annu. Rev. Plant. Biol. (Annual Review of Plant Biology)*, 661–683. 10.1146/annurev-arplant-042817-040339.

Cadotte, M.W., Carboni, M., Si, X. and Shinichi. T. 2019 Do traits and phylogeny support congruent community diversity patterns and assembly inferences? *Journal of Ecology* **107**, 2065–2077. 10.1111/1365-2745.13247.

Cadotte, M.W. and Tucker, C.M. 2017 Should Environmental Filtering be Abandoned? *Trends in Ecology & Evolution* **32**, 429–437. 10.1016/j.tree.2017.03.004.

Cadotte, M.W. and Tucker, C.M. 2018 Difficult decisions: Strategies for conservation prioritization when taxonomic, phylogenetic and functional diversity are not spatially congruent. *Biological Conservation* **225**, 128–133. 10.1016/j.biocon.2018.06.014.

Cavender-Bares, J., Keen, A. and Miles, B. 2006 Phylogenetic structure of Floridian plant communities depends on taxonomic and spatial scale. *Ecology* **87**, 109-122. 10.1890/0012-9658(2006)87[109:PSOFPC]2.0.CO;2.

Chao, A., Chiu, C.-H. and Jost, L. 2016 Phylogenetic Diversity Measures and Their Decomposition: A Framework Based on Hill Numbers. In *Biodiversity Conservation and Phylogenetic Systematics: Preserving our evolutionary heritage in an extinction crisis*. R. Pellens and P. Grandcolas (eds). Springer International Publishing. 978-3-319-22461-9, pp. 141–172.

Chirici, G., Winter, S. and McRoberts, R.E. 2011 *National Forest Inventories: Contributions to Forest Biodiversity Assessments.* Springer Netherlands, Dordrecht. 978-94-007-0481-7, 220 p.

Corona, P., Chirici, G., McRoberts, R.E., Winter, S. and Barbati, A. 2011 Contribution of large-scale forest inventories to biodiversity assessment and monitoring. *Forest Ecology and Management* **262**, 2061–2069. 10.1016/j.foreco.2011.08.044.

Di Biase, L., Tsafack, N., Pace, L. and Fattorini, S. 2023 Ellenberg Indicator Values Disclose Complex Environmental Filtering Processes in Plant Communities along an Elevational Gradient. *Biology* **12**. 10.3390/biology12020161.

Diniz, É.S., Dias, F.S., Borda-de-Água, L. and González, P.M.R. 2024 Anthropogenic disturbance and alien plant invasion drive the phylogenetic impoverishment in riparian vegetation. *Biodivers Conserv* **33**, 4237–4256. 10.1007/s10531-024-02949-z.

Dolukhanov, A.G. 2010 *Лесная растительность Грузии: (Forest vegetation of Georgia) [in Russian].* Universal, Tbilisi, Georgia. 978-9941-17-176-5. bilib.blogspot.com/2011/12/blog-post_24.html, 484 p.

Ellenberg, H. 1974 *Zeigerwerte der Gefasspflanzen Mitteleuropas: [in German].* Verlag Erich Goltze KG, Gottingen.

Ette, J.-S., Sallmannshofer, M. and Geburek, T. 2023 Assessing Forest Biodiversity: A Novel Index to Consider Ecosystem, Species, and Genetic Diversity. *Forests* **14**, 709. 10.3390/f14040709.

E-Vojtkó, A., Bello, F. de, Lososová, Z. and Götzenberger, L. 2023 Phylogenetic diversity is a weak proxy for functional diversity but they are complementary in explaining community assembly patterns in temperate vegetation. *Journal of Ecology* **111**, 2218–2230. 10.1111/1365-2745.14171.

Fischer, E., Gröger, A. and Lobin, W. 2018 *Illustrated field guide to the flora of Georgia (South Caucasus).* 1st edn. University of Koblenz-Landau, Koblenz. 978-3-9820257-0-4, 830 p.

Gadow, K. von, Álvarez González, J.G., Zhang, C., Pukkala, T. and Zhao, X. 2021 *Sustaining forest ecosystems.* Springer, Cham, Switzerland. 9783030587130, 419 p.

Gaggiotti, O.E., Chao, A., Peres-Neto, P., Chiu, C.-H., Edwards, C. and Fortin, M.-J. et al. 2018 Diversity from genes to ecosystems: A unifying framework to study variation across biological metrics and scales. *Evolutionary applications* **11**, 1176–1193. 10.1111/eva.12593.

Gilbert, B. and Lechowicz, M.J. 2004 Neutrality, niches, and dispersal in a temperate forest understory. *PNAS* **101**, 7651–7656. 10.1073/pnas.0400814101. www.pnas.org/cgi/doi/10.1073pnas.0400814101.

Gillerot, L., Grussu, G., Condor-Golec, R., Tavani, R., Dargush, P. and Attorre, F. 2021 Progress on incorporating biodiversity monitoring in REDD+ through national forest inventories. *Global Ecology and Conservation* **32**, e01901. 10.1016/j.gecco.2021.e01901.

Gimaret-Carpentier, C., Pélissier, R., Pascal, J.-P. and Houllier, F. 1998 Sampling strategies for the assessment of tree species diversity. *Journal of Vegetation Science* **9**, 161–172. 10.2307/3237115.

Guevara Andino, J.E., Pitman, N.C.A., Steege, H. ter, Peralvo, M., Cerón, C. and Fine, P.V.A. 2021 The contribution of environmental and dispersal filters on phylogenetic and taxonomic beta diversity patterns in Amazonian tree communities. *Oecologia* **196**, 1119–1137. 10.1007/s00442-021-04981-0.

Hao, M., Gadow, K. von, Alavi, S.J., Álvarez-González, J.G., Baluarte-Vásquez, J.R. and Corral-Rivas, J. et al. 2021 A classification of woody communities based on biological dissimilarity. *Applied Vegetation Science* **24**. 10.1111/avsc.12565.

Hardisty, A.R., Michener, W.K., Agosti, D., Alonso García, E., Bastin, L. and Belbin, L. et al. 2019 The Bari Manifesto: An interoperability framework for essential biodiversity variables. *Ecological Informatics* **49**, 22–31. 10.1016/j.ecoinf.2018.11.003.

Heydari, M., Omidipour, R. and Greenlee, J. 2020 Biodiversity, a review of the concept, measurement, opportunities, and challenges. *Journal of Wildlife and Biodiversity* **4**. 10.22120/jwb.2020.123209.1124.

Heym, M., Uhl, E., Moshammer, R., Dieler, J., Stimm, K. and Pretzsch, H. 2021 Utilising forest inventory data for biodiversity assessment. *Ecological Indicators* **121**, 107196. 10.1016/j.ecolind.2020.107196.

Jeno, L.M., Grotle Rundereim, K. and Grytnes, J.A. 2025 An experimental comparison between two mobile apps for species identification from the lens of self-determination theory. *Technology, Pedagogy and Education*, 1–19. 10.1080/1475939X.2025.2511039.

Jetz, W., McGeoch, M.A., Guralnick, R., Ferrier, S., Beck, J. and Costello, M.J. et al. 2019 Essential biodiversity variables for mapping and monitoring species populations. *Nature ecology & evolution* **3**, 539–551. 10.1038/s41559-019-0826-1.

Jin, Y. and Qian, H. 2022 V.PhyloMaker2: An updated and enlarged R package that can generate very large phylogenies for vascular plants. *Plant diversity* **44**, 335–339. 10.1016/j.pld.2022.05.005.

Kacic, P. and Kuenzer, C. 2022 Forest Biodiversity Monitoring Based on Remotely Sensed Spectral Diversity—A Review. *Remote Sensing* **14**, 5363. 10.3390/rs14215363.

Kavtaradze, G., Basilidze, L., Pergl, J., Avoiani, E., Vahalik, P. and Aptsiauri, B. et al. 2023 Distribution and environmental impact of alien woody species in lowland riparian forest habitats: Case study in the protected areas of Georgia, South Caucasus. *J. For. Sci.* **69**, 401–414. 10.17221/3/2023-JFS.

Kermavnar, J. and Kutnar, L. 2020 Patterns of Understory Community Assembly and Plant Trait-Environment Relationships in Temperate SE European Forests. *Diversity* **12**, 91. 10.3390/d12030091.

Kindt, R. 2020 *WorldFlora: An R package for exact and fuzzy matching of plant names against the World Flora Online Taxonomic Backbone data*, 19 p.

Lachashvili, N., Kereselidze, K. and Kikvidze, M. 2022 The checklist of trees of Georgia (Caucasus) and their provisional regional assessment according to IUCN categories and criteria. *Fl. Medit.* **32**, 149–188. 10.7320/FlMedit32.149.

Lam, T.Y. and Kleinn, C. 2008 Estimation of tree species richness from large area forest inventory data: Evaluation and comparison of jackknife estimators. *Forest Ecology and Management* **255**, 1002–1010. 10.1016/j.foreco.2007.10.007.

Lapin, K., Oettel, J., Braun, M. and Konrad, H. (eds). 2025 *Ecological Connectivity of Forest Ecosystems.* 1st edn. Springer Nature Switzerland; Imprint Springer, Cham. 9783031822063, 211 p.

Legendre, P. and Legendre, L. 2012 *Numerical Ecology*. Third English Edition. Elsevier BV, Great Britain. 978-0-444-53868-0, 1003 p.

Lin, H.-T., Lam, T.Y., Gadow, K. von and Kershaw, J.A. 2020 Effects of nested plot designs on assessing stand attributes, species diversity, and spatial forest structures. *Forest Ecology and Management* **457**, 117658. 10.1016/j.foreco.2019.117658.

Lovett, J. 2001 Hotspots. Earth's Biologically Richest and Most Endangered Terrestrial Ecoregions EDITED BY RUSSELL A. MITTERMEIER, NORMAN MYERS AND CRISTINA GOETTSCH MITTERMEIER 430 pp., 36 × 29.5 × 3.5 cm, ISBN 9 686 39758 2 hardback, US$ 65.00/GB £45.50, Mexico and Washington DC, USA: CEMEX and Conservation International, distributed by the University of Chicago Press, 2000. *Envir. Conserv.* **28**, 86–94. 10.1017/S0376892901270088.

Luo, Y.-H., Cadotte, M.W., Burgess, K.S., Liu, J., Tan, S.-L. and Xu, K. et al. 2019 Forest community assembly is driven by different strata-dependent mechanisms along an elevational gradient. *Journal of Biogeography* **46**, 2174–2187. 10.1111/jbi.13669.

Ma, X., Mahecha, M.D., Migliavacca, M., van der Plas, F., Benavides, R. and Ratcliffe, S. et al. 2019 Inferring plant functional diversity from space: the potential of Sentinel-2. *Remote Sensing of Environment* **233**, 111368. 10.1016/j.rse.2019.111368.

Mäder, P., Boho, D., Rzanny, M., Seeland, M., Wittich, H.C., Deggelmann, A. and Wäldchen, J. 2021 The Flora Incognita app – Interactive plant species identification. *Methods Ecol Evol* **12**, 1335–1342. 10.1111/2041-210X.13611.

Mammola, S., Carmona, C.P., Guillerme, T. and Cardoso, P. 2021 Concepts and applications in functional diversity. *Functional Ecology* **35**, 1869–1885. 10.1111/1365-2435.13882.

McRoberts, R.E., Tomppo, E., Schadauer, K., Vidal, C., Ståhl, G. and Chirici, G. et al. 2009 Harmonizing National Forest Inventories. *Journal of Forestry* **107**, 179–187. doi.org/10.1093/jof/107.4.179.

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

Myers, N., Mittermeier, R.A., Mittermeier, C.G., Da Fonseca, G.A.B. and Kent, J. 2000 Biodiversity hotspots for conservation priorities. *Nature* **403**, 853–858. 10.1038/35002501.

Nakhutsrishvili, G. 2013 Forest Vegetation of Georgia. In *The Vegetation of Georgia (South Caucasus)*. G. Nakhutsrishvili (ed). Springer Berlin Heidelberg. 978-3-642-29914-8, pp. 35–87.

Oettel, J. and Lapin, K. 2021 Linking forest management and biodiversity indicators to strengthen sustainable forest management in Europe. *Ecological Indicators* **122**, 107275. 10.1016/j.ecolind.2020.107275.

Padullés Cubino, J., Lososová, Z., Bonari, G., Agrillo, E., Attorre, F. and Bergmeier, E. et al. 2021 Phylogenetic structure of European forest vegetation. *Journal of Biogeography* **48**, 903–916. 10.1111/jbi.14046.

Paudel, P., Beckschäfer, P. and Kleinn, C. 2021 Impact of training on different observers in forest inventory. *Banko* **31**, 12–22. 10.3126/banko.v31i1.37338.

Pereira, H.M., Ferrier, S., Walters, M., Geller, G.N., Jongman, R.H.G. and Scholes, R.J. et al. 2013 Ecology. Essential biodiversity variables. *Science (New York, N.Y.)* **339**, 277–278. 10.1126/science.1229931.

Pereira, I.S., Rezende, V.L., Meira-Neto, J.A.A., Clappe, S. and Eisenlohr, P.V. 2021 Phylogenetic structure as a predictive component of beta diversity: Lessons from a comprehensive Neotropical biogeographic transition. *Perspectives in Plant Ecology, Evolution and Systematics* **49**, 125602. 10.1016/j.ppees.2021.125602.

Petzer, T. 2023 *Forum Interdisziplinäre Begriffsgeschichte: Vladimir Sukachev's Concept of Biogeocoenosis.* 12th edn., Berlin, 9 p. publikationen.ub.uni-frankfurt.de/frontdoor/index/index/docId/74810.

Proença, V., Martin, L.J., Pereira, H.M., Fernandez, M., McRae, L. and Belnap, J. et al. 2017 Global biodiversity monitoring: From data sources to Essential Biodiversity Variables. *Biological Conservation* **213**, 256–263. 10.1016/j.biocon.2016.07.014.

Randin, C.F., Ashcroft, M.B., Bolliger, J., Cavender-Bares, J., Coops, N.C. and Dullinger, S. et al. 2020 Monitoring biodiversity in the Anthropocene using remote sensing in species distribution models. *Remote Sensing of Environment* **239**, 111626. 10.1016/j.rse.2019.111626.

Ricotta, C., Szeidl, L. and Pavoine, S. 2021 Towards a unifying framework for diversity and dissimilarity coefficients. *Ecological Indicators* **129**, 107971. 10.1016/j.ecolind.2021.107971.

Sánchez-Pinillos, M., Dakos, V. and Kéfi, S. 2024 Ecological dynamic regimes: A key concept for assessing ecological resilience. *Biological Conservation* **289**, 110409. 10.1016/j.biocon.2023.110409.

Sánchez-Pinillos, M., Leduc, A., Ameztegui, A., Kneeshaw, D., Lloret, F. and Coll, L. 2019 Resistance, Resilience or Change: Post-disturbance Dynamics of Boreal Forests After Insect Outbreaks. *Ecosystems* **22**, 1886–1901. 10.1007/s10021-019-00378-6.

Schmeller, D.S., Mihoub, J.-B., Bowser, A., Arvanitidis, C., Costello, M.J. and Fernandez, M. et al. 2017 An operational definition of essential biodiversity variables. *Biodivers Conserv* **26**, 2967–2972. 10.1007/s10531-017-1386-9.

Schmidt, R., Casario, B., Zipse, P. and Grabosky, J. 2022 An Analysis of the Accuracy of Photo-Based Plant Identification Applications on Fifty-Five Tree Species. *AUF* **48**, 27–43. 10.48044/jauf.2022.003.

Sukachev, V. and Dylis, N. 1964 Основы лесной биогеоценологии. Fundamentals of forest biogeocoenology [in Russian].

Swenson, N.G., Enquist, B.J., Pither, J., Thompson, J. and Zimmerman, J.K. 2006 The Problem and Promise of Scale Dependency in Community Phylogenetics. *Ecology* **87**, 2418–2424. 10.1890/0012-9658(2006)87[2418:TPAPOS]2.0.CO;2.

Tarkhnishvili, D. 2014 *Historical Biogeography of the Caucasus.* Nova Science Pub Inc; UK ed. edition. 978-1-63321-910-6. novapublishers.com/shop/historical-biogeography-of-the-caucasus/, 229 p.

Tichý, L., Axmanová, I., Dengler, J., Guarino, R., Jansen, F. and Midolo, G. et al. 2023 Ellenberg-type indicator values for European vascular plant species. *J Vegetation Science* **34**. 10.1111/jvs.13168.

Tilman, D., Knops, J., Wedin, D., Reich, P., Ritchie, M. and Siemann, E. 1997 The Influence of Functional Diversity and Composition on Ecosystem Processes. *Science* **277**, 1300–1302. 10.1126/science.277.5330.1300.

Tinya, F., Márialigeti, S., Bidló, A. and Ódor, P. 2019 Environmental drivers of the forest regeneration in temperate mixed forests. *Forest Ecology and Management* **433**, 720–728. 10.1016/j.foreco.2018.11.051.

Traub, B. and Wüest, R.O. 2020 Analysing the quality of Swiss National Forest Inventory measurements of woody species richness. *Forest ecosystems* **7**, 37. 10.1186/s40663-020-00252-1.

Vasadze, R., Dumbadze, G. and Gatenadze, S. (eds). 2024 *The Biodiversity of Georgian Forests and the Current Situation*. papers.4science.ge, 9 p.

Vellend, M., Baeten, L., Becker-Scarpitta, A., Boucher-Lalonde, V., McCune, J.L., Messier, J. and Sax, D.F. 2017 Plant Biodiversity Change Across Scales During the Anthropocene. *Annu. Rev. Plant. Biol. (Annual Review of Plant Biology)* **68**, 563–586. doi.org/10.1146/annurev-arplant-042916-040949.

Webb, C.O., Losos, J.B. and Agrawal, A.A. 2006 Integrating Phylogenies into Community Ecology. *Ecology - Special Issue* **87**, S1-S2. 10.1890/0012-9658(2006)87[1:IPICE]2.0.CO;2.

Wilcox, T.M., Schwartz, M.K. and Lowe, W.H. 2018 Evolutionary Community Ecology: Time to Think Outside the (Taxonomic) Box. *Trends in Ecology & Evolution* **33**, 240–250. 10.1016/j.tree.2018.01.014.

Winter, S., Böck, A. and McRoberts, R.E. 2012 Estimating tree species diversity across geographic scales. *Eur J Forest Res* **131**, 441–451. 10.1007/s10342-011-0518-0.

Winter, S., Chirici, G., McRoberts, R.E., Hauk, E. and Tomppo, E. 2008 Possibilities for harmonizing national forest inventory data for use in forest biodiversity assessments. *Forestry* **81**, 33–44. 10.1093/forestry/cpm042.

Wohlgemuth, T. 2000 Diskreter und kontinuierlicher Charakter der Vegetation: Waldvegetationsdaten als Referenz. *Bauhinia* **14**, 67–88. zobodat.at/pdf/Bauhinia_14_0077-0088.pdf.

Yao, J., Zhang, C., Cáceres, M. de, Legendre, P. and Zhao, X. 2019 Variation in compositional and structural components of community assemblage and its determinants. *J Vegetation Science* **30**, 257–268. 10.1111/jvs.12708.

Yu, Q., Rao, X., Ouyang, S., Xu, Y., Hanif, A. and Ni, Z. et al. 2019 Changes in taxonomic and phylogenetic dissimilarity among four subtropical forest communities during 30 years of restoration. *Forest Ecology and Management* **432**, 983–990. 10.1016/j.foreco.2018.10.033.

Zhao, C., Zhu, Y. and Meng, J. 2022 Effects of Plot Design on Estimating Tree Species Richness and Species Diversity. *Forests* **13**, 2003. 10.3390/f13122003.

Zilliox, C. and Gosselin, F. 2014 Tree species diversity and abundance as indicators of understory diversity in French mountain forests: Variations of the relationship in geographical and ecological space. *Forest Ecology and Management* **321**, 105–116. 10.1016/j.foreco.2013.07.049.

Acknowledgements

This thesis constitutes the end of a long journey, marked by a highly nonlinear progression. Before turning to words of thanks, I would like to outline how this journey unfolded. In 2015, I joined the National Forest Agency of Georgia as an integrated expert, following my graduation in Tropical and International Forestry from the University of Göttingen and several years of work in international cooperation. My role focused on institutional support in the development of forest monitoring systems. Coincidentally, the external experts advising the process were instructors from my alma mater. For three years, we explored theoretical frameworks, practical options, and technical feasibilities with our Georgian colleagues. After several pilot inventories marked by demanding fieldwork and intensive discussions, the first National Forest Inventory of Georgia was finally implemented. In 2018, I relocated to Bonn to be nearer to my children, yet the country never truly let me go. I returned to Georgia several times as consultant, until I was invited to pursue a doctoral thesis within the CaBOL (Caucasus Barcode of Life) project. Motivated by my ongoing interest in biodiversity research in the country, I embraced the opportunity. This marked a return to my academic beginnings, as I moved between Bonn and Göttingen and reconnected with the foundations of my earlier work. However, COVID-19 restrictions made fieldwork impossible, and the original research proposal could not be pursued. After a detour exploring systematic metadata using the GeBOL (German Barcode of Life) database, I redirected my focus to the National Forest Inventory. By that time, two years had already gone by and despite the foreseeable ending of project funding, I remained committed to my research, simultaneously working as freelance consultant. Following a change in supervision, I continued my research under the guidance of Prof. Dr. Bernhard Misof at the University of Bonn. The past five years have been characterized by unanticipated events and the continuous effort to balance opposing responsibilities: dedication, adaptation, parenting, earning a livelihood, and periods of deprivation. Yet, they were also years of personal growth, of acquiring new skills, and of gaining a deeper understanding of the complexity of the world around us. This would not have been possible without the unprejudiced support and help of others. Therefore, I would like to express the following:

To my family—thank you for your love, your frustration, and your ability to adapt to a time when I wasn't always present in the ways you needed. I know that my attention, time, and energy were often elsewhere, and that this journey placed limits on what I could offer you— materially and emotionally.

To my supervisors, Prof. Dr. Bernhard Misof and Dr. Nils Hein—I am grateful for their trust and the opportunity to continue my research.

To my parents—your support and understanding made this journey possible. Thank you for giving me the freedom to pursue it.

Special thanks go to all the people that provided technical or just human support: David Tarkhnishvili, Hannes Feilhauer, Sebastian Schmidtlein, Lasha Dzadzamia, Zurab Janiashvili, Lutz Fehrmann, Lars Drössler, Eike Feldmann, Nils Griese, Axel Weinreich, Irakli Sisvadze, Merab Machavariani, Jonas Astrin, Jana Thormann, Björn Rulik, Ximo Mengual, Beka Koberidze, Jan Staubach, Stefano Ricci, Lauri Vesa, the helpful team of AWF Göttingen, Stephanie Steinbach, Jan Petersen, Matthias Dees, Demian Schuck, Christoph Kleinn, Klaus von Gadow, Benjamin Wellenbeck, Rossybel Molina Galeas.

I thank the Ministry of Environmental Protection and Agriculture, Georgia for granting permission to use the data from the National Forest Inventory and all foresters whose dedication in the field ensured the collection of high-quality data. I would also like to thank the German Federal Ministry of Education and Research for funding the first three years of this research, as well as the people behind the open-source communities of QGIS, GRASS, SAGA, LAStools, and R.

Received: 4 March 2024 | Revised: 30 May 2024 | Accepted: 31 May 2024

DOI: 10.1002/ece3.11569

Ecology and Evolution Open Access WILEY

REVIEW ARTICLE

Discriminating woody species assemblages from National Forest Inventory data based on phylogeny in Georgia

Alexander Wellenbeck[1,2] | **Lutz Fehrmann**[2] | **Hannes Feilhauer**[3] | **Sebastian Schmidtlein**[4] | **Bernhard Misof**[1,5] | **Nils Hein**[5]

[1]Systematic Zoology, University of Bonn, Bonn, Germany

[2]Forest Inventory and Remote Sensing, University of Göttingen, Göttingen, Germany

[3]Remote Sensing Centre for Earth System Research (RSC4Earth), Leipzig University, Leipzig, Germany

[4]Institute of Geography and Geoecology, Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany

[5]Leibniz Institute for the Analysis of Biodiversity Change (LIB), Museum Koenig, Bonn, Germany

Correspondence
Alexander Wellenbeck, Systematic Zoology, Meckenheimer Allee 169, Bonn 53115, Germany.
Email: alex.wellenbeck@posteo.de

Funding information
Bundesministerium für Bildung und Forschung, Grant/Award Number: 01DK20014C

Abstract

Classifications of forest vegetation types and characterization of related species assemblages are important analytical tools for mapping and diversity monitoring of forest communities. The discrimination of forest communities is often based on β-diversity, which can be quantified via numerous indices to derive compositional dissimilarity between samples. This study aims to evaluate the applicability of unsupervised classification for National Forest Inventory data from Georgia by comparing two cluster hierarchies. We calculated the mean basal area per hectare for each woody species across 1059 plot observations and quantified interspecies distances for all 87 species. Following an unspuervised cluster analysis, we compared the results derived from the species-neutral dissimilarity (Bray-Curtis) with those based on the Discriminating Avalanche dissimilarity, which incorporates interspecies phylogenetic variation. Incorporating genetic variation in the dissimilarity quantification resulted in a more nuanced discrimination of woody species assemblages and increased cluster coherence. Favorable statistics include the total number of clusters (23 vs. 20), mean distance within clusters (0.773 vs. 0.343), and within sum of squares (344.13 vs. 112.92). Clusters derived from dissimilarities that account for genetic variation showed a more robust alignment with biogeographical units, such as elevation and known habitats. We demonstrate that the applicability of unsupervised classification of species assemblages to large-scale forest inventory data strongly depends on the underlying quantification of dissimilarity. Our results indicate that by incorporating phylogenetic variation, a more precise classification aligned with biogeographic units is attained. This supports the concept that the genetic signal of species assemblages reflects biogeographical patterns and facilitates more precise analyses for mapping, monitoring, and management of forest diversity.

KEYWORDS
beta diversity, community discrimination, dissimilarity, diversity monitoring, National Forest Inventory, phylogeny, Georgia, unsupervised clustering

TAXONOMY CLASSIFICATION
Applied ecology, Biodiversity ecology, Biogeography, Community ecology, Conservation ecology, Population ecology

აბსტრაქტული

ტყის მცენარეულობის ტიპების კლასიფიკაცია და მონათესავე სახეობათა შეკრების დახასიათება მნიშვნელოვანი ანალიტიკური ინსტრუმენტებია ტყის ტიპების აღწერისა და მრავალფეროვნების მონიტორინგისთვის. ტყის ტიპების განსხვავება ხშირად ემყარება β-მრავალფეროვნებას, რომლის რაოდენობრივი დადგენა შესაძლებელია მრავალი ინდექსის მეშვეობით ნიმუშებს შორის კომპოზიციური განსხვავებულობის გამოსათვლელად. ეს კვლევა მიზნად ისახავს შეაფასოს საქართველოს ეროვნული ტყის ინვენტარიზაციის ზედამხედველობის გარეშე კლასიფიკაციის გამოყენებადობა ორი კლასტერული იერარქიის შედარების გზით. ჩვენ გამოვთვალეთ საშუალო ბაზალური ფართობი ჰექტარზე თითოეული მერქნიანი სახეობისთვის 1059 ნაკვეთზე დაკვირვებით და რაოდენობრივად დავადგინეთ სახეობათაშორისი მანძილი 87-ვე სახეობისთვის. ჩვენ შევადარეთ სახეობების ნეიტრალური განსხვავებულობიდან მიღებული შედეგები (ბრეი-კურტისი) ზვავის დისკრიმინაციული განსხვავებულობის საფუძველზე, რომელიც აერთიანებს სახეობათაშორის ფილოგენეტიკურ ვარიაციებს. გენეტიკური ცვალებადობის ჩართვამ განსხვავებულობის რაოდენობრივ განსაზღვრებაში გამოიწვია მერქნიანი სახეობების შეკრების უფრო ნიუანსური განსხვავება და გაზრდილი კლასტერული თანმიმდევრულობა. ხელსაყრელი სტატისტიკა მოიცავს მტევანთა საერთო რაოდენობას (23 v. 20), საშუალო მანძილს მტევნის შიგნით (0.773 vs. 0.343) და კვადრატების ჯამის ფარგლებში (344.13 vs. 112.92). განსხვავებებიდან მიღებული კლასტერებმა, რომლებიც ითვალისწინებენ გენეტიკურ ვარიაციებს, აჩვენეს უფრო მძლავრი გასწორება ბიოგეოგრაფიულ ერთეულებთან, როგორიცაა სიმაღლე და ცნობილი ჰაბიტატები. ჩვენ ვაჩვენებთ, რომ სახეობების შეკრების უკონტროლო კლასიფიკაციის გამოყენებადობა ფართომასშტაბიანი ტყის ინვენტარიზაციის მონაცემებზე მტკიცედ არის დამოკიდებული განსხვავებულობის ფუძემდებლური რაოდენობრივი განსაზღვრაზე. ჩვენი შედეგები მიუთითებს, რომ ფილოგენეტიკური ვარიაციით, უფრო ზუსტი კლასიფიკაციაა შესაძლებელი, რომელიც შეესაბამება ბიოგეოგრაფიულ ერთეულებს. ეს ამტიცებს კონცეფციას, რომ სახეობათა შეკრების გენეტიკური სიგნალი ასახავს ბიოგეოგრაფიულ ნიმუშებს და ხელს უწყობს ტყის მრავალფეროვნების აღწერას მონიტორინგისა და მართვის უფრო ზუსტ ანალიზს.

1 | INTRODUCTION

Forest ecosystems host the largest share of terrestrial biodiversity and cover approximately one-third of the global land surface (FAO, 2020; Gillerot et al., 2021; Heym et al., 2021; Torresani et al., 2019). In light of increasing pressure on forests due to climate change and the related global loss of biodiversity, also referred to as the "the sixth mass extinction," with up to 92% of terrestrial endemics being anticipated to be negatively impacted (Manes et al., 2021), reliable approaches to assess and monitor forest diversity are required (Barnosky et al., 2011; Cowie et al., 2022; Faith, 2013; Palombo, 2021). Monitoring should include the quantification of metrics that allow the classification of ecological entities based on their specific level, or degree of biodiversity, and ultimately according to their respective conservation value, which is required by conservationists (Brooks et al., 2015; Zampiglia et al., 2019). Appropriate delineation of forest communities and characterization of related species assemblages across taxonomic groups are important analytical tools for sensible monitoring of species diversity and biodiversity management. On broader scales, such metrics can be provided by means of vegetation classification which aims to group spatial or temporal diversity of species within a finite set of abstract categories (de Cáceres et al., 2015). Vegetation classification has proven to provide adequate means for descriptive reporting, communication, and mapping of forest communities, and related concepts have responded to changing information needs over time. Consequently, forest-type classifications exist for a wide range of targets, that is, habitat quality (qualitative assessments for biodiversity management), development over time (i.e., stand classification according to age classes for forest management) or along biogeographic gradients, and remote sensing-based mapping of ecological communities (de Cáceres et al., 2013; de Cáceres, Martín-Alcón, et al., 2019; Fassnacht et al., 2016; Hao et al., 2021).

One approach for classifying forest communities focuses on the variation in species compositional characteristics across assemblages of different sites within a geographic area, which is commonly known as β-diversity (Legendre & de Cáceres, 2013; Magurran & McGill, 2011). β-diversity can be assessed by the change in species compositional characteristics between sites (i.e., species turnover, Jost, 2010) and a plethora of metrics exist to quantify the degree of dissimilarity between assemblages on various spatial and temporal scales (de Cáceres et al., 2013; de Cáceres, Coll, et al., 2019; Legendre & Legendre, 2012; Magurran & McGill, 2011; Ricotta, 2005). The most common dissimilarity indices are exclusively based on compositional characteristics, that is, species richness and elements of evenness (Magurran, 2005), while interspecies variability (i.e., phylogenetic, taxonomic, functional, or traits) is not considered (Chao et al., 2018; Chiu et al., 2014; de Cáceres et al., 2013; Hao, Ganeshaiah, et al., 2019; Pavoine et al., 2013). In line with the increasing recognition that genetic diversity comprises an integral part of biodiversity, for example, as stated in the definition of biodiversity by the Intergovernmental Platform on Biodiversity and Ecosystem Services (IPBES, Díaz et al., 2015), literature on how to incorporate phylogenetics as aspect of diversity is growing rapidly (Chao et al., 2023). Accordingly, several authors have approached forest community classification by accounting for both compositional data and interspecies phylogenetic variability (i.e., Capelo, 2020; Hao et al., 2021; Ricotta et al., 2020; Webb et al., 2002). As phylogenetically closely related species often share beneficial traits for specific environments, discriminating assemblages based on phylogenetic distances can serve as a proxy for classifying forest communities according to functional roles, environmental diversity, and conservation value (Faith, 2013; Gilbert & Parker, 2022; Hawkins et al., 2014; Padullés Cubino et al., 2021; Pavoine, 2016; Pavoine & Ricotta, 2014; Tucker et al., 2017). Hao, Ganeshaiah, et al. (2019) demonstrated that different patterns emerged if interspecies taxonomic distances were considered for the classification of global forest communities using the Discriminating Avalanche index (Ganeshaiah & Shaankar, 2000).

On a smaller scale, National Forest Inventories (NFIs) provide systematic and periodical observations of tree species abundances based on permanent sample units on a country-wide level (Corona et al., 2011). The continuous adaptation of variables assessed during NFIs highlights an increasing emphasis on aspects of biodiversity, enabling ecologists to investigate potentials and limitations of the thus provided data (Alberdi et al., 2019; Corona et al., 2011; Didion, 2020; McRoberts et al., 2009). Incorporating phylogenetic diversity of species assemblages extends the perspective on diversity in this context and bears the potential to deepen our understanding of the complex interactions among woody species over large geographical scales.

In the present study, we compare the performance of two dissimilarity indices for the discrimination of forest woody species assemblages when applied to large-scale forest inventory data such as the dataset of the first NFI of Georgia. To this end, we applied unsupervised clustering to the obtained dissimilarity matrices based on a conventional and a dissimilarity index that incorporates interspecies phylogenetic distances, respectively. Apart from statistics for internal evaluation of the resulting classifications, our comparison considered the distribution of discriminated assemblages along biogeographic gradients. Based on the assumption that genetic variability of co-inhabiting species provides a signal that sufficiently reflects site-specific environmental determinants, we investigated whether including this variable in the quantification of dissimilarity results in an improved reflection of biogeographic gradients. To test the general applicability, we incorporated the phylogenetic variability into the classification of a large, real-world dataset and evaluated the results considering cluster cohesiveness and overall interpretability.

2 | DATA AND METHODS

We compare two dissimilarity indices for the classification of woody species assemblages when applied to NFI data of Georgia. Adhering to the methodological approach underlying the data, we focus on woody species, that is, all recorded species that meet the specified target diameter at breast height (DBH, at 1.3 m, MEPA, 2018). Consequently, we are referring to woody species even if only species is written hereafter.

2.1 | Study area

Georgia is located between the Southern Slopes of the Greater and the Northern part of the Lesser Caucasus, between 41°07′ – 43°35′ N and 40°04′ – 46°44′ E (Fischer et al., 2018). The forests of Georgia host large shares of endemic species and form part of the Caucasus biodiversity hotspot (Joppa et al., 2011; Myers, 2003). Existing forest formations range from Alpine coniferous forests dominated by *Abies nordmanniana* (Steven) Spach. and *Picea orientalis* (L.) Peterm. to open juniper woodland (dominated by *Juniperus polycarpos excelsa* subsp. *polycarpos* (K. Koch) Takht. and *J. foetidissima* Willd.), encompassing further thermophilus to xerophytic mixed oak forest (*Quercus petrea* subsp. *iberica* (Steven ex M. Bieb.) Krassiln.,

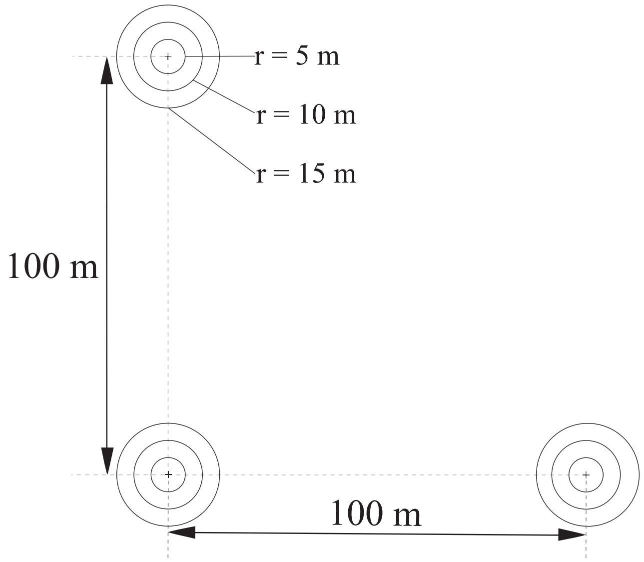


FIGURE 1 Cluster plot configuration of the first National Forest Inventory of Georgia. Each cluster consists of three subplots for tree assessment within three concentric circles according to the measured diameter at breast height (DBH, at 1.3 m, MEPA, 2023).

Carpinus betulus L., and *C. orientalis* Mill.), Colchic alder carrs which are dominated by *Alnus glutinosa* subsp. *barbata* (C. A. Mey.) Yalt. and oriental beech (*Fagus orientalis* Lipsky) and hornbeam-oriental beech forests (Bohn et al., 2007; Dolukhanov, 2010; Fischer et al., 2018; Nakhutsrishvili et al., 2021; Novák et al., 2023).

2.2 | Forest community data

Between 2018 and 2021, Georgia implemented its' first NFI based on a systematic sampling grid of 3.6 × 3.6 km with a randomly selected origin. Sampling units consist of cluster plots (0.21 ha) containing three subplots of 0.0.7 ha each. These subplots are arranged in an L-shaped configuration with a distance of 100 m along both axes (Figure 1). As 18% of the country's territory is currently not accessible for government officials due to an ongoing political conflict (MEPA, 2023), clusters were sampled on approximately 74% of the national forest area (Figure 2).

2.2.1 | Assessment of Woody species

Woody species were recorded per subplot on three concentric nested subplots according to any stems' respective DBH. Stems with DBH ≥ 30 cm were recorded on the largest plot ($r = 15$ m). Stems with DBH ≥ 15 and ≥8 cm were recorded on the inner nested plot radii of 10 m radius and 5 m, respectively (Figure 1). Numerous variables were recorded per woody species, along with the polar coordinates of the stem axis, species, and DBH (MEPA, 2018).

From the entire NFI data set ($N = 2006$), all accessible clusters pertaining to the locally applied land use class "Forest" and sub-classes "Tree covered area" or "Fire affected forests" were selected for analysis (MEPA, 2018). Subplots with recorded intersections with a forest boundary ("slopover sample plots") were excluded to avoid including samples with extreme outliers regarding species richness due to edge effects (Willmer et al., 2022). Clusters containing species observations that were not unambiguously identified at species level, for example, *Deciduous* spp. and *Genera* spp., were omitted because a precise quantification of the cophenetic distances along the phylogenetic hierarchy is not possible. Clusters containing subplots without woody species observations were excluded. Our sample consists of all cluster plots comprised of observations from three subplots ($n = 1059$, henceforward referred to as "samples"), which represent 53% of all clusters. Figure 2 provides an overview of the spatial distribution of samples used for the analysis.

After reprojection of sample locations to UTM38N, WGS84 (EPSG: 32638), sample elevations [m above sea level] were derived from the digital elevation model (DEM) provided by Shuttle Radar Topography Mission (SRTM, Farr et al., 2007). Elevation values were calculated as the median of all raster cell values (30 × 30 m) contained in or crossed by the circular subplot area ($r = 15$ m + recorded GPS error [m]) of the southwestern subplot.

2.2.2 | Species diversity data

Diversity measures are usually based on data representing the compositional variation between species (i.e., occurrence and abundance) in an assemblage (Ricotta et al., 2021). Forest species communities may display similar compositional characteristics, in terms of counts of observed species and respective individuals. However, species can be represented by large numbers of small-diameter trees, or stems belonging to the same individual, or fewer individuals but with significantly large relative shares of total basal area. Hence, abundance estimates based on counts of individuals do not take significant differences in the size structure of occurring species into consideration and may result in distinct evenness profiles. We used mean basal area (BA, m^2/ha) per species and cluster plot as species abundances to account for the variation in size of the constituents. By weighting compositional data using BA, we incorporate valuable structural information that considers site occupation per species for the quantification of β-diversity (de Cáceres, Coll, et al., 2019; McRoberts et al., 2009; Staudhammer & LeMay, 2001; Yao et al., 2019). Consequently, BA of all living stems was aggregated per cluster plot and species and divided by 3 to obtain mean BA estimates per sample (de Cáceres et al., 2015; MEPA, 2018).

2.3 | Analysis

2.3.1 | Nomenclature

Spelling and nomenclature of all recorded species were standardized with the Taxonomic Backbone databases of World Flora Online (WFO DB, Kindt, 2020) and the Global Biodiversity Information Facility (GBIF

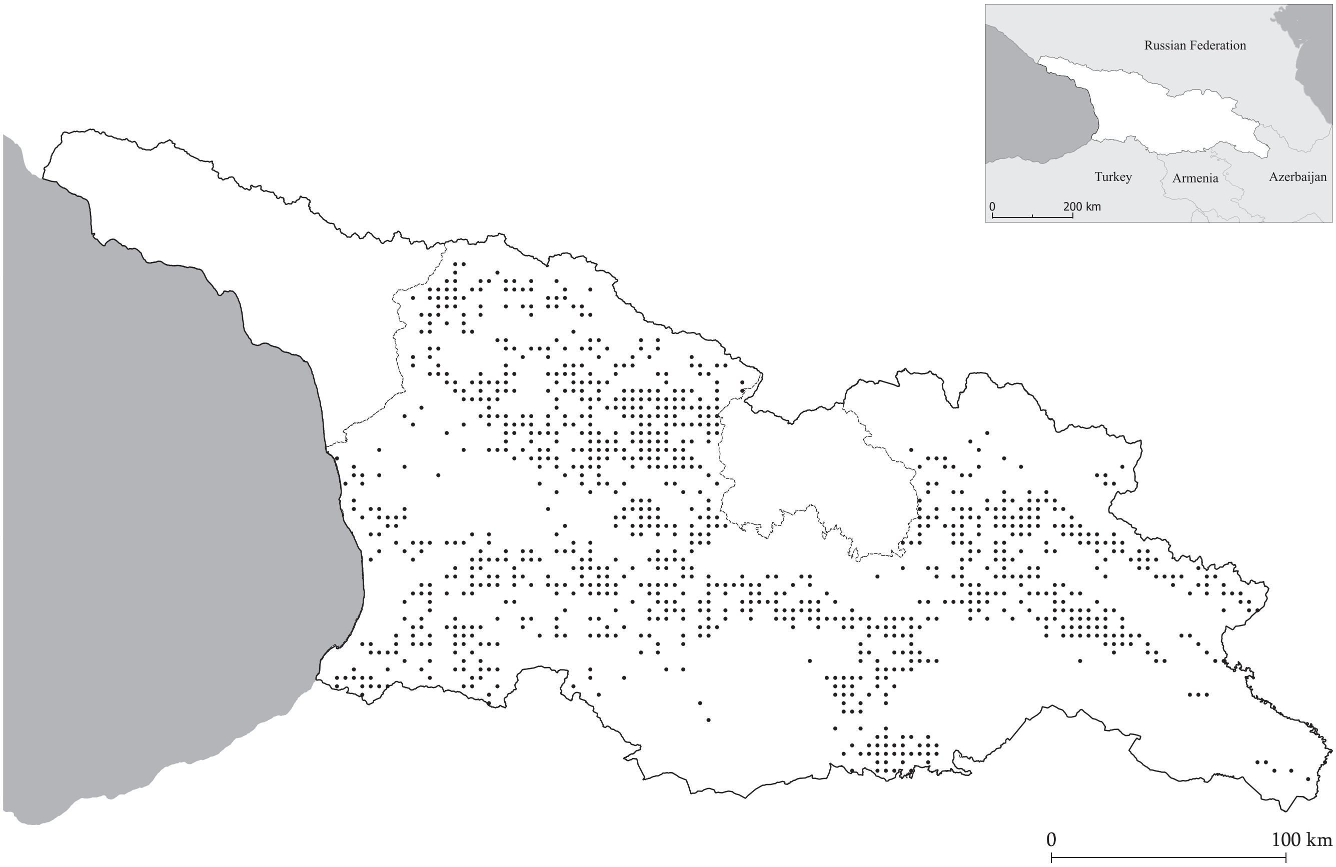


FIGURE 2 Locations of samples used from the National Forest Inventory of Georgia, $n = 1059$. Samples (black dots) consist of accessible cluster plots of equal sample size (three subplots) located inside forests that are not intersected by forest boundaries and contain only records of taxa identified at species level. Dashed lines mark the boundaries of inaccessible areas from where no field data were obtained.

Secretariat, 2021). In cases where species listed in the NFI data did not yield an unequivocal match in the WFO DB, corresponding records were harmonized with Lachashvili et al. (2022) and the nomenclature of the World Plants database (https://www.worldplants.de) to derive names for all species ranked as taxonomically accepted.

2.3.2 | Phylogenetic interspecies distances

A phylogenetic tree encompassing all observed species was constructed by matching the harmonized species list with a mega phylogeny containing 72,570 species of vascular plants according to the World Plants database (GBOTB.extended.WP.tre, Jin & Qian, 2022). The backbone mega phylogeny is based on the species-level phylogeny for vascular plants derived from gene sequencing from 7 gene regions and 39 fossil calibrations created by Zanne et al. (2014), which was updated and expanded by Jin and Qian (2022). Following the authors' recommendation to consolidate taxa below species level (e.g., sub-species) with the respective parental species, five infraspecific taxa were combined with their parental species, resulting in the lowest taxonomic unit being species level (Figure 3). From the thus created ultrametric phylogenetic tree (Jin & Qian, 2019; Qian & Jin, 2016; Smith & Brown, 2018), cophenetic distances, that is, the total branch length connecting each pair of species at the terminal nodes of the respective phylogeny, were calculated (Bevilacqua et al., 2021; Hao, Corral-Rivas, et al., 2019; Kling et al., 2018).

Interspecies phylogenetic distances were normalized within a square matrix that contained pairwise distance values $[0 < \text{dist}_{ph} \leq 1]$ between each pair of species.

2.3.3 | Dissimilarity indices

The Discriminating Avalanche index (dA equation [2] in Table 1) developed by Ganeshaiah and Shaankar (2000) and described by Hao, Corral-Rivas, et al. (2019) considers interspecies dissimilarity by multiplying absolute differences in frequencies (numerator in BC) of species *i* and *j* in two samples with a specific distance between species *i* and *j*. We use the phylogenetic distances to weigh the difference in mean basal area between *i* and *j*. Table 1 shows the formulas of both indices used in this study.

As the maximum dissimilarity value obtained by $\text{dA} = \left(1 - \frac{1}{n}\right)$, where *n* equals the number of species, the resulting dissimilarities were normalized via $x_{\text{norm}} = \frac{x - x_{\min}}{x_{\max} - x_{\min}}$, with $x_{\min}$ and $x_{\max}$ representing the minimum and maximum value of dA, respectively (Hao, Corral-Rivas, et al., 2019; Legendre & Legendre, 2012). Consequently,

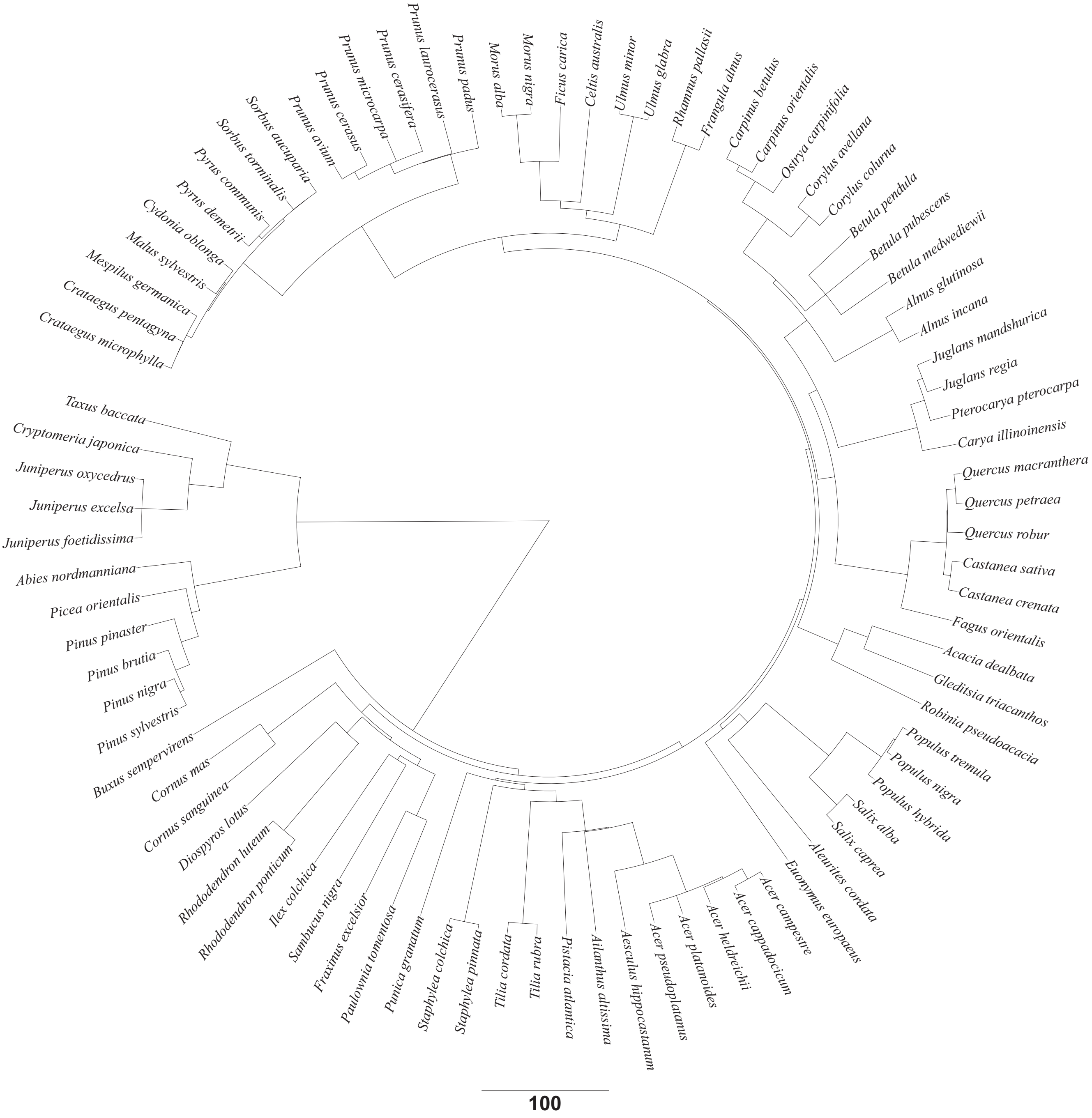


FIGURE 3 Phylogeny for 87 species listed in the data sample of the National Forest Inventory of Georgia. Interspecies phylogenetic distances were calculated as total branch length connecting each pair of species at the terminal nodes of the hierarchy. For respective branch lengths, see Figure S1.

pairwise dissimilarity values of the two resulting dissimilarity matrices (1059×1059) range between 0 and 1. Pairwise values of 1 imply that two samples are completely different as they do not share any species, whereas values of 0 indicate two samples are equal in terms of compositional characteristics (Chao et al., 2005; Legendre & Legendre, 2012; Leyer & Wesche, 2008). Prior to clustering, a Mantel test (Mantel, 1967) was performed to check for existing correlations between the two dissimilarity matrices. As the pairwise dissimilarities are not normally distributed, and non-linear relationships between the pairwise dissimilarity values exist, we used the Spearman correlation coefficient with 9999 permutations (Legendre & Legendre, 2012).

2.3.4 | Isometric partitioning

The ISOPAM algorithm (Schmidtlein et al., 2010) available in package "isopam" (v. 2.0, Schmidtlein et al., 2024) combines isometric feature mapping and partitioning around medoids (data points that are most centrally located within each cluster with the sum of dissimilarities between medoids and all other data points being minimized) in order to build clusters with a maximum number and fidelity of indicative species. The isomap ordination, which is based on geodesic distances strongly determined by neighborhood definitions, is repeated with different parameter settings. The result is clustered, and the clusters are evaluated according to the criteria mentioned above. These criteria are similar to those used when structuring phytosociological tables (Abe, 2021). In this, and the use of an ordination, Isopam is similar to Twinspan (Hill, 1979), but does not involve internal readjustments, uses geodesic distances (taking account of "neighbors of neighbors" in feature space), and works on multidimensional ordination spaces. It has been previously used for large-scale classifications of forests (Cabido et al., 2018; Černý et al., 2015; Zeballos et al., 2020) and other systems (Feilhauer et al., 2021; Hein et al., 2014; Peterka et al., 2017; Yu et al., 2022). Isopam can be run both unsupervised and supervised (with reference plots). For the current study, the original source code was extended to support dA (Capelo, 2020) and executed on a computer with two Intel Xeon CPUs (E5-2630 v3) and 256 GB RAM using R Statistical Software (v 4.3.2; R Core Team, 2023). To ease comparability, we set the maximum number of hierarchy levels to four for both classifications. Subsequently, we extracted lists of indicator species frequencies with levels of significance according to Fisher's exact test for each cluster using the Isotab function, which is part of the "isopam" package. Fidelity ("equalized phi," Tichy & Chitrý, 2006) together with Fisher's exact test if the observed frequency is not attained by chance are the criteria for qualifying as an indicator species in Isotab (Schmidtlein et al., 2024).

2.3.5 | Evaluation of clustering

To evaluate the correspondence between the original sample dissimilarities and dendrogram distances, we calculated cophenetic correlation coefficients for each hierarchical cluster structure (Lapointe & Legendre, 1995; Legendre & Legendre, 2012). The modified Rand index was used to evaluate clustering performance based on the consistency between partitions (Legendre & Legendre, 2012). Cluster homogeneity was evaluated via within sum of squares (WSS, Hao, Ganeshaiah, et al., 2019) and a comparison of the average distance between and within clusters using the function cluster.stats of the R package "fpc." To assess relevance of the hierarchies, we compared indicator species and the resulting distributions of relative BA among partitioned groups (de Cáceres et al., 2015). Evaluation of correspondence to biogeographic units was based on a comparison of elevational ranges derived from the DEM between groups and the spatial distribution of clusters in relation to forest vegetation-type classifications presented by Bohn et al. (2007). We applied the nonparametric Kruskal–Wallis test to check for significance between groups due to nonnormality in the distribution of elevation values within groups (Shapiro test). Henceforward, we are referring to the initial partition at the lowest level of the hierarchy as classes, to the intermediary partitions as branches, and the resulting clusters as assemblages.

TABLE 1 Dissimilarity indices used in this study.

Bray–Curtis dissimilarity index (1957)	$\mathrm{BC}=\frac{\sum_{i=1}^{n}\lvert p_i^a-p_i^b\rvert}{\sum_{i=1}^{n}(p_i^a+p_i^b)}$	[1]
Discriminating Avalanche (Hao, Corral-Rivas, et al., 2019)	$\mathrm{dA}=\frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}\Delta_i^{a,b}d_{ij}\Delta_j^{a,b}$	[2]

Note: 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 $\left(\left|p_i^a-p_i^b\right|\right)$; n = total number of sample plots; p_i^a, p_i^b = relative frequencies of species i in plots a and b.

2.3.6 | Data analysis

The analysis was conducted in R Studio version 2023.09.0-463 (RStudio Team, 2020) using R Base version 4.3.0 (R Core Team, 2023). Harmonization of nomenclature was realized via the R package "Worldflora" version 1.13-2 (Kindt, 2020), and the package "V.phylomaker2" version 0.1.0 was used to match observed species with the phylogenetic backbone (Jin & Qian, 2022). The Mantel test and BC dissimilarities were calculated using the packages "vegan" version 2.6-4 (Oksanen et al., 2022). A custom function was embedded in the adjusted clustering algorithm of the corresponding R package "isopam" version 1.2.0 for dA (Schmidtlein et al., 2022). Clustering metrics were obtained using the R packages "mclust," version 6.0.0 (Scrucca et al., 2016), and "fpc," version 2.2-10 (Hennig, 2023).

3 | RESULTS

Compositional data of $n=1059$ samples containing 65,818 living tree observations were analyzed (Table 2). In total, 87 species were represented by 52 genera, 29 families, 16 orders, and two classes.

TABLE 2 Summary statistics of compositional data of cluster plot observations of the National Forest Inventory of Georgia.

Species richness					Mean basal area per species [m^2/ha]			
n	Min	Max	Mean[a]	CV%	Min	Max	Mean[a]	CV%
1059	1	12	4.96 (±2.14)	43.08	2.54	79.02	30.4 (±12.79)	42.06

Abbreviations: CV%, Coefficient of variation; Var., Variance.

[a] Means are denoted with standard deviation in parenthesis.

Angiosperms were represented by 76 species (in 87.4% of all samples) across 26 different families. *Fagaceae* (six species) accounted for the highest number of observations followed by *Betulaceae* with 10 species (25.3% and 23.2% of all samples, respectively), *Rosaceae*, and *Sapindaceae* (15 and 6 species, in 11.4% and 10.3% of all samples, respectively). In contrast, gymnosperms (10.3%) were represented by 11 species, belonging to three different families, with the largest family being *Pinaceae* (6 species) followed by *Cupressaceae* and *Taxaceae* (4 and 1 species, respectively). Gymnosperms were observed in 33% of all samples.

3.1 | Pairwise dissimilarity

Mantel statistics ($r=.613$ with $p=.0004$) indicated a significant positive correlation between the two dissimilarity matrices. Pairwise dissimilarities according to dA ranged from 0.23 to 0.87, which is almost double the range of that of BC (0.63 to 0.99, Table 3).

Consequently, the mean dissimilarity between samples (0.79 for BC and 0.38 for dA, respectively) and thus overall variation was higher for BC than for dA, as for the latter, frequencies of dissimilarity values <1 were more evenly distributed, with very few pairwise dissimilarities of 1. Frequency distributions of sample dissimilarities are provided in Figure S2.

3.2 | Discrimination of assemblages

The hierarchical clustering based on BC (HC_{BC}) distinguished 23 assemblages (clusters) across four hierarchical levels (I–IV) and classes. Within HC_{BC}, samples were partitioned into 10 and 17 branches at levels II to III, respectively. The hierarchical clustering based on dA (HC_{dA}) contained 10 and 15 branches at levels II to III and resulted in 20 distinct assemblages over four levels and classes. Two assemblages were not partitioned further below level II in HC_{dA} (Figure 4). The number of samples per assemblage ranged from 6 to 152 for HC_{BC} (Mean: 82 ± 91.5) and 3 to 163 for HC_{dA} (Mean: 106.1 ± 119.9), respectively. For HC_{BC}, the number of observed species per assemblage ranged from 12 to 54. For HC_{dA}, one assemblage contained only four species, whereas two assemblages encompassed 53 species.

The cophenetic correlation coefficients were 0.492 and 0.442 for HC_{BC} and HC_{dA}, respectively. The obtained adjusted Rand index of 0.317 suggests a modest level of similarity between the clustering results. Average distances within clusters ranged from 0.671 to 0.875 and 0.196 to 0.49 for HC_{BC} and HC_{dA}, respectively. Mean distance between clusters was 0.789 for HC_{BC} and 0.384 for HC_{dA}. WSS values of HC_{BC} amounted to 344.13 and 112.92 for HC_{dA}.

3.2.1 | Evaluation of cluster hierarchies

To evaluate the performance of BC and dA for clustering, we compared the resulting classes, groups, and assemblages considering internal metrics of the partitioning process. A total of 77 and 68 indicator species were listed for all partitions based on BC and dA, respectively. For a characterization based on indicators and respective frequencies, only highly significant ($p \leq .001$) species with total frequencies ≥50% were considered, unless indicated otherwise. Indicator species for both partitions were *A. nordmanniana*, *A. cappadocicum*, *A. glutinosa*, *C. betulus*, *Carpinus orientalis* Mill., *C. sativa*, *F. orientalis*, *P. orientalis*, *Q. petraea* subsp. *polycarpa* (Schur) Raus, and *Tilia rubra* subsp. *caucasica* (Rupr.) V. Engl. In addition, *Fraxinus excelsior* L. is an indicator for HC_{dA}. The total number of indicators with frequencies of 100% was 24 and 8 for HC_{BC} and HC_{dA}, respectively. Class 1 of HC_{BC} is characterized by a high frequency of *F. orientalis* (99%), whereas for HC_{dA}, *F. orientalis* and *C. betulus* are listed with frequencies of 93% and 75%, respectively. In class 2, the highest frequencies are observed for *Q. petrea* (87%) and *C. betulus* (82%). *P. orientalis* (68%), *F. orientalis* (57%), and *A. nordmanniana* (53%) are the most frequent indicator species in class 3, whereas for class 4, these are *A. glutinosa* (87%) and *Castanea sativa* Mill. (61%). In HC_{dA}, highly significant indicators in class 1 are *F. orientalis* (93%) and *C. betulus* (75%), whereas in class 2, these are *Q. petrea* (84%), *C. orientalis* (61%), and *Fraxinus excelsior* L. (50%). In classes 3 and 4 of HC_{dA}, *P. orientalis* (79%), *A. nordmanniana* (64%), *A. glutinosa* (92%), and *C. sativa* (59%) represent highly significant indicators in classes 3 and 4, respectively. Based on these characteristics, we labeled the four main classes according to predominant relative BA and are referring to these for ease of readability henceforward as follows: class 1 is characterized by a dominance of *Fagus*, class 2 is *Carpinus-Quercus* dominated, and classes 3 and 4 are *Pinaceae* and *Alnus-Castanea* dominated, respectively. Synoptic tables of both cluster hierarchies (Figures S3 and S4, respectively) and a detailed description of indicator distributions per partition below level I for HC_{dA} (Appendix S1) are provided as appendices.

3.2.2 | Elevation and spatial distribution

To evaluate the correspondence of assemblages to existing biogeographic units, we compared the distribution of sample elevations

TABLE 3 Summary statistics of mean pairwise dissimilarities between samples based on Bray–Curtis (BC) and Discriminating Avalanche (dA) of the National Forest Inventory data of Georgia.

Index	*n*	Min	Max	Mean[a]	Var.	CV%
BC	1059	0.627	0.998	0.789 (±0.099)	0.01	12.55
dA		0.225	0.87	0.383 (±0.147)	0.022	38.38

Abbreviations: CV%, Coefficient of variation; Var., Variance.

[a] Means are denoted with standard deviation in parenthesis.

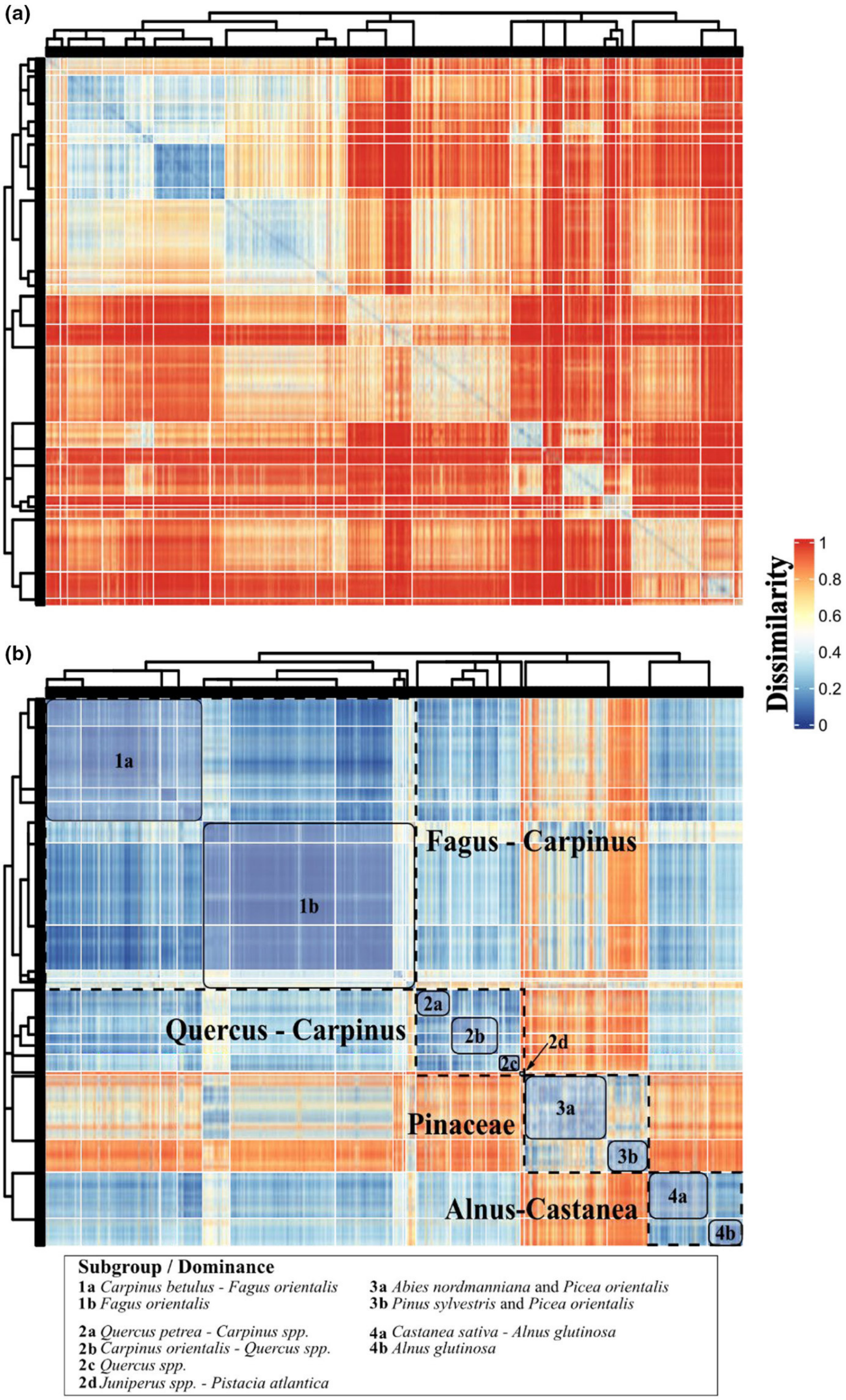


FIGURE 4 Resulting cluster hierarchy from the isopam partitioning (dendrogram) and pairwise dissimilarities of 1059 samples from the National Forest Inventory data of Georgia. Dissimilarities are based on the Bray–Curtis (a) and discriminating avalanche (b) index. The cell grid is colored according to the dissimilarity values between samples (rows and columns). In total, 23 and 20 assemblages were discriminated for Bray–Curtis and discriminating avalanche, respectively. At level II of the hierarchical clustering based on the discriminating avalanche index, 10 subgroups distributed over four classes were labeled according to dominance of relative basal area.

within assemblages. Within the four classes in both hierarchies, sample elevations are distributed similarly. Samples within the *Fagus*-dominated group (class 1) cover a wide elevational range (≤750 m to >2000 m asl), however, in HC_{BC}, 81% of all samples are located between >1000 and 2000 m asl, whereas in HC_{dA}, most samples (77%) are located within the lower range of >750 and 1750 m asl. Samples with *Carpinus-Quercus* dominance are predominantly located at elevations <1250 m in both hierarchies. The majority of *Pinaceae*-dominated samples lie above 1250 m asl, whereas all of *Alnus-Castanea*-dominated samples are situated below 1250 m asl, with the majority (70%) positioned below 750 m asl for HC_{BC} and HC_{dA}, respectively. Overall classes, except for the *Fagus*-dominated group, agglomerations of sample elevation values are more pronounced within assemblages of HC_{dA} than of HC_{BC}. The applied Kruskal–Wallis test revealed highly significant differences ($p \leq .001$) between assemblages for both hierarchies (Figure 5).

Spatial distributions of discriminated assemblages show a general alignment along biogeographical units. *Alnus*

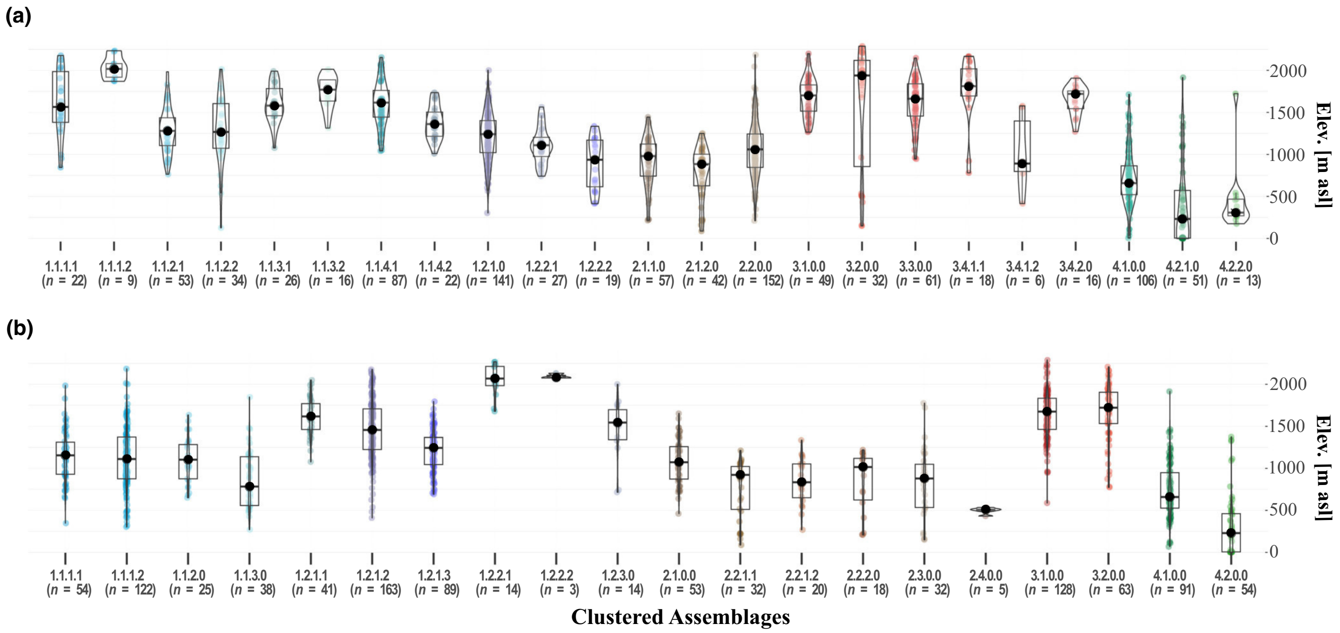


FIGURE 5 Sample elevations [m asl] from the National Forest Inventory in Georgia per resulting cluster for Bray–Curtis (a) and discriminating avalanche (b). The boxes represent interquartile ranges and respective median values (solid line) of sample elevations within each clustered assemblage (colored dots). Resulting p values of the $X^2_{\text{Kruskal–Wallis}}$ test are $p=2.899\text{e-}113$ and $p=1.674\text{e-}115$, with confidence intervals between 0.544, 1 and 0.538, and 1 for assemblages in the hierarchical clustering based on BC and dA, respectively ($n=1059$).

glutinosa-dominated assemblages agglomerate in the humid Alder carrs of Eastern Georgia (Nakhutsrishvili, 2013), whereas the *Pinaceae*-dominated assemblages are predominantly located in montane to subalpine areas of the Lesser and Greater Caucasus. Assemblages dominated by mixed *Quercus* spp., *Carpinus* spp., and *F. orientalis* are situated at intermediate ranges. Those with high shares of BA of *Quercus* spp. are limited to lower and drier areas, while *F. orientalis*-dominated assemblages are located at higher elevations. Interestingly, the five samples assigned to *Juniperus*-*Pistacia* woodland have been clearly discriminated within HC_{dA} that are located in the semi-arid areas of the southwest of the country (Nakhutsrishvili et al., 2021).

To visually evaluate the spatial distribution of discriminated assemblages, we cut HC_{dA} at level II because 70% of the partitions are not partitioned further below level II, resulting in 10 clustered assemblages (Figure 4). Based on characteristic indicator species, relative BA distributions, and occupied elevational zones, we labeled each assemblage accordingly and mapped the respective sample location in relation to areas of forest vegetation-type classifications according to Bohn et al. (2007) (Figure 6).

4 | DISCUSSION

Parting from the assumption that genetic variation of co-inhabiting species provides a signal that reflects site-specific environmental determinants, we contrast the performance of a species-neutral dissimilarity index (BC) with an index that considers interspecies genetic variation (dA) when used in unsupervised classification. Our findings indicate that incorporating interspecies phylogenetic distances in the quantification of dissimilarities results in more coherent and ecologically meaningful classifications of large-scale forest inventory data with high β-diversity.

The Mantel statistics indicate a significant positive correlation between the dissimilarity matrices obtained for each index ($r=.613$, with $p=.0004$), implying that essential patterns of variation among samples are maintained in the respective quantifications. However, frequency distributions and visual inspection of dissimilarities (Figure 4) show overall higher dissimilarity values based on BC, reflecting its' sensitivity to species turnover. Whereas the resulting cluster hierarchies maintain a certain level of agreement (cophenetic correlation = .511), the fact that a correlation of one signifies complete similarity suggests that the dissimilarity signal resulting from dA is not redundant. Clustering based on dA performs slightly better in preserving the original dissimilarities according to the respective cophenetic correlation coefficients of .492 and .442 for HC_{BC} and HC_{dA}, respectively. Dendrogram topologies, cluster validation metrics (i.e., WSS of 344.123 and 112.917 for HC_{BC} and HC_{dA}, respectively), and distributions of relative BA among assemblages indicate a higher degree of compactness, separation, and yield, generally more conceivable clustering results based on dA. Overall, HC_{dA} provided enhanced general interpretability and succeeded in discriminating clearly distinguished assemblages regarding compositional characteristics, that is, the *P. atlantica* and *Juniperus* woodlands of the semi-arid lowlands of Southeastern Georgia. These results support the concept that an extension of variables considered for quantification of dissimilarity leads to a refined conception for diversity classification if genetic variation is considered and are in line with the findings

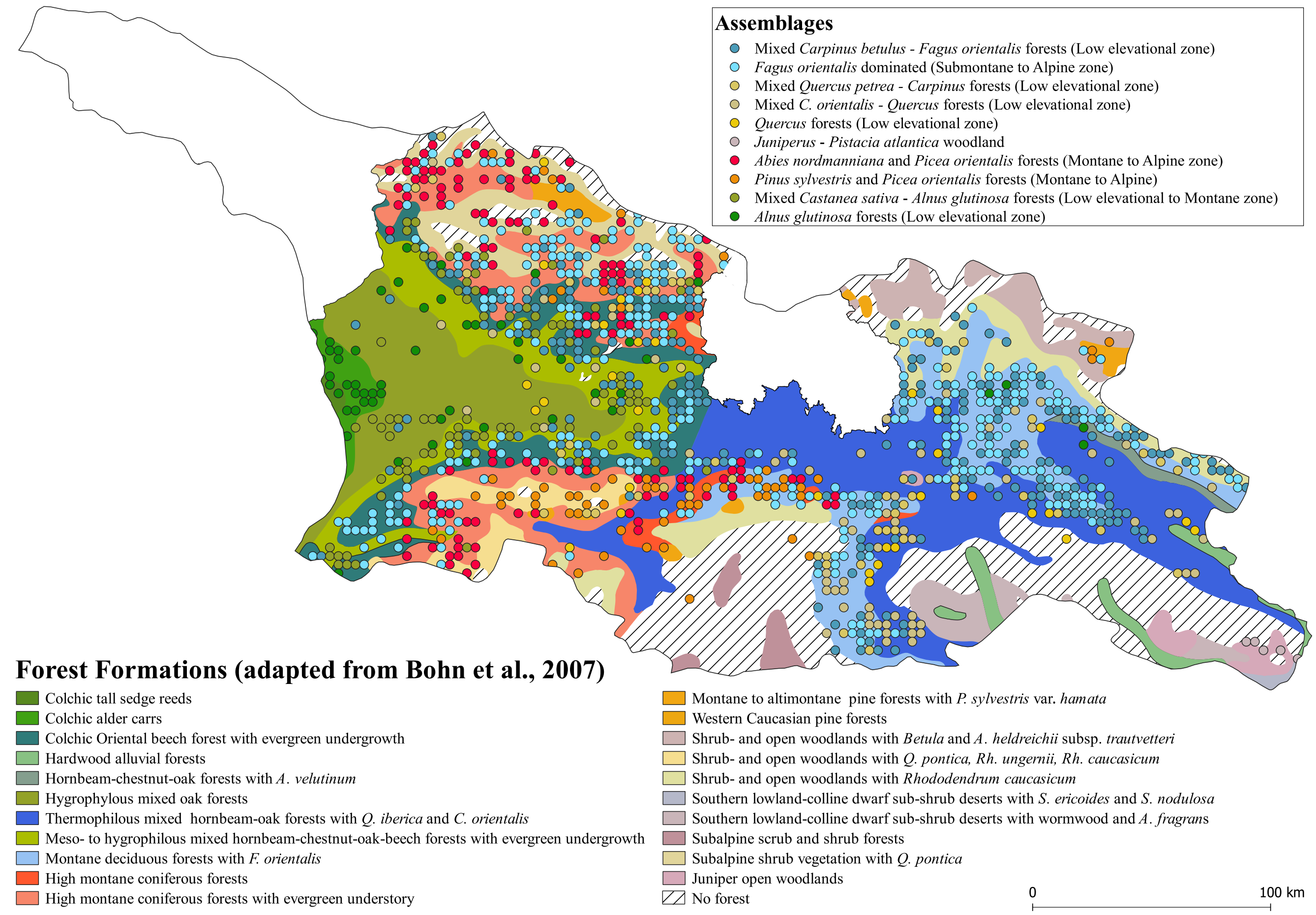


FIGURE 6 Schematic overview of spatial distributions of discriminated assemblages and main forest formations adapted from the vegetation-type classification by Bohn et al. (2007). Assemblages are colored according to the respective palette of the four main classes of the cluster hierarchy (blue, beige, red/orange, and green for dominance of Fagus, Carpinus-Quercus, Pinaceae, and Alnus-Castanea, respectively).

of Hao, Ganeshaiah, et al. (2019) and other authors (Capelo, 2020; Pavoine & Ricotta, 2014; Ricotta & Pavoine, 2015).

As BC represents a “species-neutral” diversity index sensu Chao et al. (2010), which assumes that all observed species contribute equally to overall diversity, and species turnover constitutes the predominant signal for discrimination, reflected by the significantly higher number of indicator species with frequencies of 100% (24 and 8 for HC_{BC} and HC_{dA}, respectively). Conversely, as dA dissimilarity considers species as phylogenetic units, a complete species turnover does not necessarily result in maximum dissimilarity values between two sites because differences in abundance are weighted by the genetic proximity between species. Assuming that the genetic signal of co-occurring species reflects niche occupation within given ecogeographical areas (Hawkins et al., 2014), the thus refined dissimilarity signal appears to respond to biogeographical gradients in a more interpretable manner.

The validity of the presented approach relies on precise measurements of tree diameters and accurate species identification in the field. While traditional forest science prioritized genus-level information, growing emphasis on diversity-related issues prompted forestry experts to be increasingly trained to provide accurate species identification. The related uncertainties are not design-based issues but apply to all ecological surveys requiring botanical expert knowledge to ensure taxon detection and validate observations on species level (Lam & Kleinn, 2008; Roswell et al., 2021). Overall, only 2% of all cluster plot observations included individuals that were not identified to species level and had to be excluded. By considering only cluster plot observations of equal sample size ($m=3$, 64% of all cluster plots) and the exclusion of subplots overlapping with the forest boundary (16%), our analyses are based on a subsample of the NFI data, representing 60% of all cluster plot observations. Hence, conclusions drawn from the presented results should consider, for example, that species exclusively occurring at forest boundaries are excluded. Potential limitations to validity arise from field sampling protocols, as overall subplot size, or sampling effects due to the nested subplot structure (with respective target DBH as inclusion criteria), may introduce bias to the quantification of dissimilarity (de Cáceres, Martín-Alcón, et al., 2019). Accordingly, observed numbers of species should be regarded as proxies of true species richness, especially if nested sample plot designs are applied that are based on specific diameter thresholds (Lin et al., 2020). Overall species richness can be assumed to be higher with the contributions of smaller-diameter trees being neglected (Corona et al., 2017). Aggregations to cluster plot level could potentially translate to generalization effects and the loss of information on site-specific environmental factors on smaller scales. The resulting magnitude of impact on the presented results, however, is likely to vary according to forest type, topographical condition, and management regime (McRoberts et al., 2009).

With continuing advances in whole-genome phylogenetics and functional genomics, information on phylogenetic diversity is continuously improving (Kling et al., 2018). Access to comprehensive and standardized phylogenetic mega trees to quantify species genetic relationships is readily available and their application to investigate variation in community compositions is becoming increasingly more common (Gilbert & Parker, 2022; Jin & Qian, 2022).

Our results are of relevance for a wide range of classifications of ecological entities according to conservation value, mapping of ecological communities, or other discriminative objectives. The method aligns standard canopy layer data from forest inventories with natural vegetation types according to Bohn et al. (2007), but harmonization with existing forest typologies is limited due to methodological differences, such as the structural vegetation layers considered and the abundance units recorded (e.g., Chytrý et al., 2020; Mucina et al., 2016). Investigating the degree to which the resulting clusters can be aligned with defined syntaxonomical units is an interesting area for future research, especially for the identification of diagnostic species from the shrub and herb layers to refine classifications and the development of practical assessment procedures.

The integration of genetic signals of forest communities during characterization has wide implications for respective approaches to classification. As a proxy indicative of the relationship between species composition and site conditions, interspecies genetic variation extends the scope of forest diversity mapping, management, and monitoring to account for alterations inconceivable by conventional compositional variables. Beyond the potential advantage of streamlining processes by applying unsupervised classification to large datasets, our approach is straightforward and can be readily replicated with comparable data, provided entities are assessed in a systematic manner. Ecological studies are frequently less systematic and constrained by temporal and spatial scales due to the dynamic nature of communities over time and space. This holds true to a lesser degree for assessments of woody species communities within the context of national forest monitoring systems, which are resampled in fixed intervals. Hence, from a practical point of view, the resulting network of observational studies provides a valuable framework for a systematic and recurring collection of ecological data, as additional costs and efforts can be embedded into existing structures. Extending the scope of study objectives to systematic assessments of a wider range of botanical and potentially zoological taxa could provide powerful and statistically robust data for analyses of organisms, structural components, and the interrelationships between them.

5 | CONCLUSIONS

We present an approach to discriminate species diversity from NFI data of forest communities with high β-diversity and species turnover. The novelty of the method lies in considering interspecies genetic variability for the quantification of diversity and subsequent classification using an unsupervised clustering algorithm on a country-wide scale. We demonstrate that large-scale forest inventory data can be classified in an ecologically meaningful manner based on mean basal area estimates per species and consideration of phylogenetic dissimilarity between samples. The thus obtained discrimination of species assemblages provides a differentiated picture of existing diversity patterns along expected biogeographical gradients without the need for additional assessments. This approach aligns with a biodiversity concept considering genetic diversity and could potentially be standardized for application to similar datasets, provided systematic data assessment is granted. The presented results should be considered as a step in evaluating to which extent large-area forest inventory data could provide a backbone for extended biodiversity monitoring systems, as discrimination of woody species assemblages allows for systematic delineation of forest ecosystems if genetic variation is considered.

AUTHOR CONTRIBUTIONS

Alexander Wellenbeck: Conceptualization (lead); data curation (lead); formal analysis (lead); investigation (lead); methodology (lead); validation (lead); visualization (lead); writing – original draft (lead); writing – review and editing (lead). **Lutz Fehrmann:** Conceptualization (supporting); formal analysis (supporting); methodology (supporting); supervision (supporting); visualization (supporting); writing – review and editing (supporting). **Hannes Feilhauer:** Formal analysis (supporting); methodology (supporting); writing – review and editing (supporting). **Sebastian Schmidtlein:** Methodology (supporting); software (supporting); writing – review and editing (supporting). **Bernhard Misof:** Methodology (supporting); supervision (supporting); writing – review and editing (supporting). **Nils Hein:** Conceptualization (supporting); investigation (supporting); methodology (supporting); project administration (lead); writing – original draft (supporting); writing – review and editing (supporting).

ACKNOWLEDGMENTS

We thank the Ministry of Environmental Protection and Agriculture, Georgia, for granting permission to use the NFI data. We also thank Nils Griese for his support during the creation of complex R functions. This research was carried out within the framework of the Caucasus Barcode of Life project (CaBOL) and funded by the Federal Ministry of Education and Research under grant number 01DK20014C. Open Access funding enabled and organized by Projekt DEAL.

CONFLICT OF INTEREST STATEMENT

The authors declare no competing interests.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are openly available in Dryad at https://doi.org/10.5061/dryad.wpzgmsbvv.

ORCID

Alexander Wellenbeck https://orcid.org/0009-0008-7007-3922

Nils Hein https://orcid.org/0000-0002-5172-8531

REFERENCES

Abe, S. (2021). Comparison of traditional and automated approaches in classification of Japanese coastal beach and dune vegetation. *Society of Vegetation Science*, *38*, 67–80. https://doi.org/10.15031/vegsci.38.67

Alberdi, I., Nunes, L., Kovac, M., Bonheme, I., Cañellas, I., Rego, F. C., Dias, S., Duarte, I., Notarangelo, M., Rizzo, M., & Gasparini, P. (2019). The conservation status assessment of Natura 2000 forest habitats in Europe: capabilities, potentials and challenges of national forest inventories data. *Annals of Forest Science*, 76(2), 1–15. https://doi.org/10.1007/s13595-019-0820-4

Barnosky, A. D., Matzke, N., Tomiya, S., Wogan, G. O. U., Swartz, B., Quental, T. B., Marshall, C., McGuire, J. L., Lindsey, E. L., Maguire, K. C., Mersey, B., & Ferrer, E. A. (2011). Has the Earth's sixth mass extinction already arrived? *Nature*, *471*(7336), 51–57. https://doi.org/10.1038/nature09678

Bevilacqua, S., Anderson, M. J., Ugland, K. I., Somerfield, P. J., & Terlizzi, A. (2021). The use of taxonomic relationships among species in applied ecological research: Baseline, steps forward and future challenges. *Austral Ecology*, *46*(6), 950–964. https://doi.org/10.1111/aec.13061

Bohn, U., Zazanashhvili, N., & Nakhutsrishvili, G. (2007). The map of the natural vegetation of Europe and its application in the Caucasus ecoregion. *Bulletin of the Georgian National Academy of Science*, *175*, 112–119. http://science.org.ge/old/moambe/2007-vol1/112-120.pdf

Bray, J. R., & Curtis, J. T. (1957). An ordination of the upland Forest communities of southern Wisconsin. *Ecological Monographs*, *27*(4), 325–349. https://doi.org/10.2307/1942268

Brooks, T. M., Cuttelod, A., Faith, D. P., Garcia-Moreno, J., Langhammer, P., & Pérez-Espona, S. (2015). Why and how might genetic and phylogenetic diversity be reflected in the identification of key biodiversity areas? *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *370*(1662), 20140019. https://doi.org/10.1098/rstb.2014.0019

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. *Applied Vegetation Science*, *21*(2), 298–311. https://doi.org/10.1111/avsc.12369

Capelo, J. (2020). Using species abundance and phylogeny conjointly to approach vegetation classification: A case study on Macaronesia's woody vegetation. *Journal of Vegetation Science*, *31*(4), 616–633. https://doi.org/10.1111/jvs.12886

Černý, T., Kopecký, M., Petřík, P., Song, J.-S., Šrůtek, M., Valachovič, M., Altman, J., & Doležal, J. (2015). Classification of Korean forests: Patterns along geographic and environmental gradients. *Applied Vegetation Science*, *18*(1), 5–22. https://doi.org/10.1111/avsc.12124

Chao, A., Chazdon, R. L., Colwell, R. K., & Shen, T.-J. (2005). A new statistical approach for assessing similarity of species composition with incidence and abundance data. *Ecology Letters*, *8*(2), 148–159. https://doi.org/10.1111/j.1461-0248.2004.00707.x

Chao, A., Chiu, C.-H., & Jost, L. (2010). Phylogenetic diversity measures based on Hill numbers. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *365*(1558), 3599–3609. https://doi.org/10.1098/rstb.2010.0272

Chao, A., Chiu, C.-H., Villéger, S., Sun, I.-F., Thorn, S., Lin, Y.-C., Chiang, J.-M., & Sherwin, W. B. (2018). An attribute-diversity approach to functional diversity, functional beta diversity, and related (dis)similarity measures. *Ecological Monographs*, *89*(2), e01343. https://doi.org/10.1002/ecm.1343

Chao, A., Thorn, S., Chiu, C.-H., Moyes, F., Hu, K.-H., Chazdon, R. L., Wu, J., Magnago, L. F. S., Dornelas, M., Zelený, D., Colwell, R. K., & Magurran, A. E. (2023). Rarefaction and extrapolation with beta diversity under a framework of H ill numbers: The iNEXT. beta3D standardization. *Ecological Monographs*, *93*, e1588. https://doi.org/10.1002/ecm.1588

Chiu, C.-H., Jost, L., & Chao, A. (2014). Phylogenetic beta diversity, similarity, and differentiation measures based on Hill numbers. *Ecological Monographs*, *84*(1), 21–44. https://doi.org/10.1890/12-0960.1

Chytrý, M., Tichý, L., Hennekens, S. M., Knollová, I., Janssen, J. A. M., Rodwell, J. S., Peterka, T., Marcenò, C., Landucci, F., Danihelka, J., Hájek, M., Dengler, J., Novák, P., Zukal, D., Jiménez-Alfaro, B., Mucina, L., Abdulhak, S., Aćić, S., Agrillo, E., … Schaminée, J. H. J. (2020). EUNIS habitat classification: Expert system, characteristic species combinations and distribution maps of European habitats. *Applied Vegetation Science*, *23*(4), 648–675. https://doi.org/10.1111/avsc.12519

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

Corona, P., Franceschi, S., Pisani, C., Portoghesi, L., Mattioli, W., & Fattorini, L. (2017). Inference on diversity from forest inventories: A review. *Biodiversity and Conservation*, *26*(13), 3037–3049. https://doi.org/10.1007/s10531-015-1017-2

Cowie, R. H., Bouchet, P., & Fontaine, B. (2022). The sixth mass extinction: Fact, fiction or speculation? *Biological Reviews of the Cambridge Philosophical Society*, *97*(2), 640–663. https://doi.org/10.1111/brv.12816

de Cáceres, M., Chytrý, M., Agrillo, E., Attorre, F., Botta-Dukát, Z., Capelo, J., Czúcz, B., Dengler, J., Ewald, J., Faber-Langendoen, D., Feoli, E., Franklin, S. B., Gavilán, R., Gillet, F., Jansen, F., Jiménez-Alfaro, B., Krestov, P., Landucci, F., Lengyel, A., … Wiser, S. K. (2015). A comparative framework for broad-scale plot-based vegetation classification. *Applied Vegetation Science*, *18*(4), 543–560. https://doi.org/10.1111/avsc.12179

de Cáceres, M., Coll, L., Legendre, P., Allen, R. B., Wiser, S. K., & Fortin, M.-J. (2019). Trajectory analysis in community ecology. *Ecological Monographs*, *89*(2), e01350. https://doi.org/10.1002/ecm.1350

de Cáceres, M., Legendre, P., & He, F. (2013). Dissimilarity measurements and the size structure of ecological communities. *Methods in Ecology and Evolution*, *4*(12), 1167–1177. https://doi.org/10.1111/2041-210X.12116

de Cáceres, M., 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. *Annals of Forest Science*, *76*(2), 1–19. https://doi.org/10.1007/s13595-019-0824-0

Díaz, S., Demissew, S., Carabias, J., Joly, C., Lonsdale, M., Ash, N., Larigauderie, A., Adhikari, J. R., Arico, S., Báldi, A., Bartuska, A., Baste, I. A., Bilgin, A., Brondizio, E., Chan, K. M. A., Figueroa, V.

E., Duraiappah, A., Fischer, M., Hill, R., ... Zlatanova, D. (2015). The IPBES conceptual framework — Connecting nature and people. *Current Opinion in Environmental Sustainability*, *14*, 1–16. https://doi.org/10.1016/j.cosust.2014.11.002

Didion, M. (2020). Extending harmonized national forest inventory herb layer vegetation cover observations to derive comprehensive biomass estimates. *Forest Ecosystems*, *7*(1), 1–14. https://doi.org/10.1186/s40663-020-00230-7

Dolukhanov, A. G. (2010). *Лесная растительность Грузии: (Forest vegetation of Georgia)*. Universal Tbilisi.

Faith, D. P. (2013). Biodiversity and evolutionary history: Useful extensions of the PD phylogenetic diversity assessment framework. *Annals of the New York Academy of Sciences*, *1289*, 69–89. https://doi.org/10.1111/nyas.12186

FAO. (2020). *Global Forest resources assessment 2020: Key findings*. FAO. https://doi.org/10.4060/ca8753en

Farr, T. G., Rosen, P. A., Caro, E., Crippen, R., Duren, R., Hensley, S., & Alsdorf, D. (2007). The shuttle radar topography Mission. *Reviews of Geophysics*, *45*(2), RG2004. https://doi.org/10.1029/2005RG000183

Fassnacht, F. E., Latifi, H., Stereńczak, K., Modzelewska, A., Lefsky, M., Waser, L. T., Straub, C., & Ghosh, A. (2016). Review of studies on tree species classification from remotely sensed data. *Remote Sensing of Environment*, *186*, 64–87. https://doi.org/10.1016/j.rse.2016.08.013

Feilhauer, H., Zlinszky, A., Kania, A., Foody, G. M., Doktor, D., Lausch, A., & Schmidtlein, S. (2021). Let your maps be fuzzy!—Class probabilities and floristic gradients as alternatives to crisp mapping for remote sensing of vegetation. *Remote Sensing in Ecology and Conservation*, *7*(2), 292–305. https://doi.org/10.1002/rse2.188

Fischer, E., Gröger, A., & Lobin, W. (2018). *Illustrated field guide to the flora of Georgia (South Caucasus) (1st edition). Koblenz geographical colloquia series biogeographical monographs: Vol. 3*. University of Koblenz-Landau.

Ganeshaiah, K. N., & Shaankar, R. U. (2000). Measuring biological heterogeneity of forest vegetation types: Avalanche index as an estimate of biological diversity. *Biodiversity and Conservation*, *9*, 953–963. https://doi.org/10.1023/A:1008910918751

GBIF Secretariat. (2021). *Gbif backbone taxonomy*. https://doi.org/10.15468/39OMEI

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

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. *Global Ecology and Conservation*, *32*, e01901. https://doi.org/10.1016/j.gecco.2021.e01901

Hao, M., Corral-Rivas, J., González-Elizondo, M. S., Ganeshaiah, K. N., Nava-Miranda, M. G., Zhang, C., Zhao, X., & von Gadow, K. (2019). Assessing biological dissimilarities between five forest communities. *Forest Ecosystems*, *6*(1), 1–8. https://doi.org/10.1186/s40663-019-0188-9

Hao, M., Gadow, K., Alavi, S. J., Álvarez-González, J. G., Baluarte-Vásquez, J. R., Corral-Rivas, J., Hui, G., Korol, M., Kumar, R., Liang, J., Meyer, P., Remadevi, O. K., Kakkar, R., Liu, W., Zhao, X., & Zhang, C. (2021). A classification of woody communities based on biological dissimilarity. *Applied Vegetation Science*, *24*(1), e12565. https://doi.org/10.1111/avsc.12565

Hao, M., Ganeshaiah, K. N., Zhang, C., Zhao, X., & von Gadow, K. (2019). Discriminating among forest communities based on taxonomic, phylogenetic and trait distances. *Forest Ecology and Management*, *440*, 40–47. https://doi.org/10.1016/j.foreco.2019.03.006

Hawkins, B. A., Rueda, M., Rangel, T. F., Field, R., Diniz-Filho, J. A. F., & Linder, P. (2014). Community phylogenetics at the biogeographical scale: Cold tolerance, niche conservatism and the structure of north American forests. *Journal of Biogeography*, *41*(1), 23–38. https://doi.org/10.1111/jbi.12171

Hein, N., Feilhauer, H., Finch, O.-D., Schmidtlein, S., & Löffler, J. (2014). Snow cover determines the ecology and biogeography of spiders (Araneae) in alpine tundra ecosystems. *Erdkunde*, 68, 157–172. https://doi.org/10.3112/erdkunde.2014.03.01

Hennig, C. (2023). *Fpc: Flexible procedures for clustering* (version 2.2-10,) [computer software]. https://CRAN.R-project.org/package=fpc

Heym, M., Uhl, E., Moshammer, R., Dieler, J., Stimm, K., & Pretzsch, H. (2021). Utilising forest inventory data for biodiversity assessment. *Ecological Indicators*, *121*, 107196. https://doi.org/10.1016/j.ecolind.2020.107196

Hill, M. O. (1979). *TWINSPAN – A FORTRAN program for arranging multivariate data in an ordered two way table by classification of individuals and attributes* (p. 48). Cornell University.

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 Diversity*, *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. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(32), 13171–13176. https://doi.org/10.1073/pnas.1109389108

Jost, L. (2010). The relation between evenness and diversity. *Diversity*, *2*(2), 207–232. https://doi.org/10.3390/d2020207

Kindt, R. (2020). *WorldFlora: An R package for exact and fuzzy matching of plant names against the World Flora Online Taxonomic Backbone data*. https://doi.org/10.1101/2020.02.02.930719

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. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *374*(1763), 20170397. https://doi.org/10.1098/rstb.2017.0397

Lachashvili, N., Kereselidze, K., & Kikvidze, M. (2022). The checklist of trees of Georgia (Caucasus) and their provisional regional assessment according to IUCN categories and criteria. *Flora Mediterranea*, *32*, 149–188. https://doi.org/10.7320/FlMedit32.149

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

Lapointe, F.-J., & Legendre, P. (1995). Comparison tests for dendrograms: A comparative evaluation. *Journal of Classification*, *12*, 265–282. https://doi.org/10.1007/BF03040858

Legendre, P., & de Cáceres, M. (2013). Beta diversity as the variance of community data: Dissimilarity coefficients and partitioning. *Ecology Letters*, *16*(8), 951–963. https://doi.org/10.1111/ele.12141

Legendre, P., & Legendre, L. (2012). *Numerical ecology. Third english edition (third). Developments in environmental modelling: Vol. 24*. Elsevier BV.

Leyer, I., & Wesche, K. (2008). *Multivariate Statistik in der Ökologie*. Springer Berlin Heidelberg.

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

Magurran, A. E. (2005). Biological diversity. *Current Biology*, *15*(4), 116–118. https://doi.org/10.1016/j.cub.2005.02.006

Magurran, A. E., & McGill, B. J. (2011). *Biological diversity: Frontiers in measurement and assessment*. Oxford University Press.

Manes, S., Costello, M. J., Beckett, H., Debnath, A., Devenish-Nelson, E., Grey, K.-A., Jenkins, R., Khan, T. M., Kiessling, W., Krause, C., Maharaj, S. S., Midgley, G. F., Price, J., Talukdar, G., & Vale, M. M. (2021). Endemism increases species' climate change risk in areas of global biodiversity importance. *Biological Conservation*, *257*, 109070. https://doi.org/10.1016/j.biocon.2021.109070

Mantel, N. (1967). The detection of disease clustering and a generalized regression approach. *Cancer Research*, *27*(2), 209–220.

McRoberts, R. E., Tomppo, E., Schadauer, K., Vidal, C., Ståhl, G., Chirici, G., Lanz, A., Cienciala, E., Winter, S., & Smith, B. (2009). Harmonizing National Forest Inventories. *Journal of Forestry*, *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.

MEPA. (2023). ანგარიში ტყის პირველი ეროვნული აღრიცხვა საქართველოში – *First National Forest Inventory in Georgia: Report 2023*. Ministry of Environmental Protection and Agriculture. https://mepa.gov.ge/En/Files/Download/53934

Mucina, L., Bültmann, H., Dierßen, K., Theurillat, J.-P., Raus, T., Čarni, A., Šumberová, K., Willner, W., Dengler, J., García, R. G., Chytrý, M., Hájek, M., Di Pietro, R., Iakushenko, D., Pallas, J., Daniëls, F. J., Bergmeier, E., Santos Guerra, A., Ermakov, N., … Tichý, L. (2016). Vegetation of Europe: hierarchical floristic classification system of vascular plant, bryophyte, lichen, and algal communities. *Applied Vegetation Science*, *19*(S1), 3–264. https://doi.org/10.1111/avsc.12257

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)*. Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-29915-5

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

Novák, P., Kalníková, V., Szokala, D., Aleksanyan, A., Batsatsashvili, K., Fayvush, G., Kolbaia, S., Nakhutsrishvili, G., Sedláček, V., Štěrba, T., & Zukal, D. (2023). Transcaucasian vegetation database – A phytosociological database of the southern Caucasus. *Vegetation Classification and Survey*, *4*, 231–240. https://doi.org/10.3897/VCS.105521

Oksanen, J., Simpson, G., Blanchet, F., Kindt, R., Legendre, P., Minchin, P., O'Hara, R., Solymos, P., & Stevens, M. (2022). *vegan* (Version 2.6-4) [Computer software]. https://CRAN.R-project.org/package=vegan

Padullés Cubino, J., Lososová, Z., Bonari, G., Agrillo, E., Attorre, F., Bergmeier, E., Biurrun, I., Campos, J. A., Čarni, A., Ćuk, M., de Sanctis, M., Indreica, A., Jiménez-Alfaro, B., Khanina, L., Knollová, I., Lenoir, J., Pielech, R., Rašomavičius, V., Škvorc, Ž., … Chytrý, M. (2021). Phylogenetic structure of European forest vegetation. *Journal of Biogeography*, *48*(4), 903–916. https://doi.org/10.1111/jbi.14046

Palombo, M. R. (2021). Thinking about the biodiversity loss in this changing world. *Geosciences*, *11*(9), 370. https://doi.org/10.3390/geosciences11090370

Pavoine, S. (2016). A guide through a family of phylogenetic dissimilarity measures among sites. *Oikos*, *125*(12), 1719–1732. https://doi.org/10.1111/oik.03262

Pavoine, S., Gasc, A., Bonsall, M. B., & Mason, N. W. (2013). Correlations between phylogenetic and functional diversity: Mathematical artefacts or true ecological and evolutionary processes? *Journal of Vegetation Science*, *24*(5), 781–793. https://doi.org/10.1111/jvs.12051

Pavoine, S., & Ricotta, C. (2014). Functional and phylogenetic similarity among communities. *Methods in Ecology and Evolution*, *5*(7), 666–675. https://doi.org/10.1111/2041-210X.12193

Peterka, T., Hájek, M., Jiroušek, M., Jiménez-Alfaro, B., Aunina, L., Bergamini, A., Dítě, D., Felbaba-Klushyna, L., Graf, U., Hájková, P., Hettenbergerová, E., Ivchenko, T. G., Jansen, F., Koroleva, N. E., Lapshina, E. D., Lazarević, P. M., Moen, A., Napreenko, M. G., Pawlikowski, P., … Chytrý, M. (2017). Formalized classification of European fen vegetation at the alliance level. *Applied Vegetation Science*, *20*(1), 124–142. https://doi.org/10.1111/avsc.12271

Qian, H., & Jin, Y. (2016). An updated megaphylogeny of plants, a tool for generating plant phylogenies and an analysis of phylogenetic community structure. *Journal of Plant Ecology*, *9*(2), 233–239. https://doi.org/10.1093/jpe/rtv047

R Core Team. (2023). *R: A language and environment for statistical computing [computer software]*. R Foundation for Statistical Computing. https://www.R-project.org/

Ricotta, C. (2005). Through the jungle of biological diversity. *Acta Biotheoretica*, *53*, 29–38. https://doi.org/10.1007/s10441-005-7001-6

Ricotta, C., Laroche, F., Szeidl, L., & Pavoine, S. (2020). From alpha to beta functional and phylogenetic redundancy. *Methods in Ecology and Evolution*, *11*(4), 487–493. https://doi.org/10.1111/2041-210X.13353

Ricotta, C., & Pavoine, S. (2015). Measuring similarity among plots including similarity among species: An extension of traditional approaches. *Journal of Vegetation Science*, *26*(6), 1061–1067. https://doi.org/10.1111/jvs.12329

Ricotta, C., Szeidl, L., & Pavoine, S. (2021). Towards a unifying framework for diversity and dissimilarity coefficients. *Ecological Indicators*, *129*, 107971. https://doi.org/10.1016/j.ecolind.2021.107971

Roswell, M., Dushoff, J., & Winfree, R. (2021). A conceptual guide to measuring species diversity. *Oikos*, *130*(3), 321–338. https://doi.org/10.1111/oik.07202

RStudio Team. (2020). *RStudio: Integrated development for R (Version 2023.09.0) [Computer software]*. Posit Software, P. B.C.

Schmidtlein, S., Collison, J., & Pfannendoerfer, R. (2024). *Isopam: Clustering of sites with species data* [CRAN]. https://CRAN.R-project.org/package=isopam

Schmidtlein, S., Jason, C., Pfannendoerfer, R., & Tichý, L. (2022). *Package 'isopam' (Version 1.1.9) [Computer software]*. Schmidtlein, Sebastian.

Schmidtlein, S., Tichý, L., Feilhauer, H., & Faude, U. (2010). A brute-force approach to vegetation classification. *Journal of Vegetation Science*, *21*, 1162–1171. https://doi.org/10.1088/0004-6256/141/2/44

Scrucca, L., Fop, M., Murphy, T. B., & Raftery, A. E. (2016). Mclust 5: Clustering, classification and density estimation using gaussian finite mixture models. *The R Journal*, *8*(1), 289–317. https://doi.org/10.32614/RJ-2016-021

Smith, S. A., & Brown, J. W. (2018). Constructing a broadly inclusive seed plant phylogeny. *American Journal of Botany*, *105*(3), 302–314. https://doi.org/10.1002/ajb2.1019

Staudhammer, C. L., & LeMay, V. M. (2001). Introduction and evaluation of possible indices of stand structural diversity. *Canadian Journal of Forest Research*, *31*(7), 1105–1115. https://doi.org/10.1139/cjfr-31-7-1105

Tichy, L., & Chytry, M. (2006). Statistical determination of diagnostic species for site groups of unequal size. *Journal of Vegetation Science*, *17*(6), 809–818. https://doi.org/10.1111/j.1654-1103.2006.tb02504.x

Torresani, M., Rocchini, D., Sonnenschein, R., Zebisch, M., Marcantonio, M., Ricotta, C., & Tonon, G. (2019). Estimating tree species diversity from space in an alpine conifer forest: The Rao's Q diversity index meets the spectral variation hypothesis. *Ecological Informatics*, *52*, 26–34. https://doi.org/10.1016/j.ecoinf.2019.04.001

Tucker, C. M., Cadotte, M. W., Carvalho, S. B., Davies, T. J., Ferrier, S., Fritz, S. A., Grenyer, R., Helmus, M. R., Jin, L. S., Mooers, A. O., Pavoine, S., Purschke, O., Redding, D. W., Rosauer, D. F., Winter, M., & Mazel, F. (2017). A guide to phylogenetic metrics for conservation, community ecology and macroecology. *Biological Reviews of the Cambridge Philosophical Society*, *92*(2), 698–715. https://doi.org/10.1111/brv.12252

Webb, C. O., Ackerly, D. D., McPeek, M. A., & Donoghue, M. J. (2002). Phylogenies and community ecology. *Annual Review of Ecology and Systematics*, *33*(1), 475–505. https://doi.org/10.1146/annurev.ecolsys.33.010802.150448

Willmer, J. N. G., Püttker, T., & Prevedello, J. A. (2022). Global impacts of edge effects on species richness. *Biological Conservation*, *272*, 109654. https://doi.org/10.1016/j.biocon.2022.109654

Yao, J., Zhang, C., de Cáceres, M., Legendre, P., & Zhao, X. (2019). Variation in compositional and structural components of community assemblage and its determinants. *Journal of Vegetation Science*, *30*(2), 257–268. https://doi.org/10.1111/jvs.12708

Yu, H., Fang, G., Rose, K. A., Tang, Y., & Song, X. (2022). Examining epibenthic assemblages associated with artificial reefs using a species archetype approach. *Marine and Coastal Fisheries*, *14*(3), e10206. https://doi.org/10.1002/mcf2.10206

Zampiglia, M., Bisconti, R., Maiorano, L., Aloise, G., Siclari, A., Pellegrino, F., Martino, G., Pezzarossa, A., Chiocchio, A., Martino, C., Nascetti, G., & Canestrelli, D. (2019). Drilling down hotspots of intraspecific diversity to bring them into on-ground conservation of threatened species. *Frontiers in Ecology and Evolution*, *7*, 205. https://doi.org/10.3389/fevo.2019.00205

Zanne, A. E., Tank, D. C., Cornwell, W. K., Eastman, J. M., Smith, S. A., FitzJohn, R. G., McGlinn, D. J., O'Meara, B. C., Moles, A. T., Reich, P. B., Royer, D. L., Soltis, D. E., Stevens, P. F., Westoby, M., Wright, I. J., Aarssen, L., Bertin, R. I., Calaminus, A., Govaerts, R., … Beaulieu, J. M. (2014). Three keys to the radiation of angiosperms into freezing environments. *Nature*, *506*(7486), 89–92. https://doi.org/10.1038/nature12872

Zeballos, S. R., Giorgis, M. A., Cabido, M. R., Acosta, A. T., Del Iglesias, M. R., & Cantero, J. J. (2020). The lowland seasonally dry subtropical forests in central Argentina: Vegetation types and a call for conservation. *Vegetation Classification and Survey*, *1*, 87–102. https://doi.org/10.3897/VCS/2020/38013

SUPPORTING INFORMATION

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

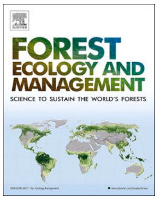
How to cite this article: Wellenbeck, A., Fehrmann, L., Feilhauer, H., Schmidtlein, S., Misof, B., & Hein, N. (2024). Discriminating woody species assemblages from National Forest Inventory data based on phylogeny in Georgia. *Ecology and Evolution*, *14*, e11569. https://doi.org/10.1002/ece3.11569

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

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 ≥ 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).

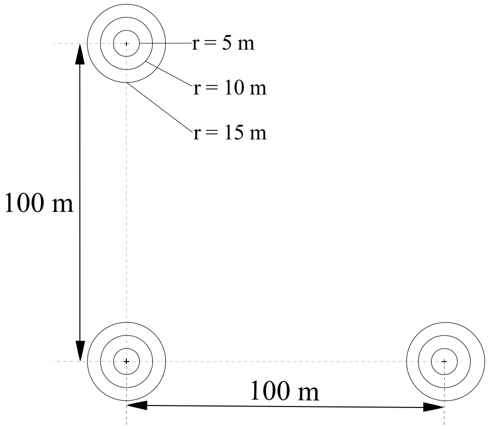


Fig. 1. Configuration of cluster plots comprising three sample plots of the National Forest Inventory of Georgia. Woody species are recorded within three nested subplots according to the measured diameter at breast height (DBH, at 1.3 m,). For each stem, tree and stem number, species (if identifiable) 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

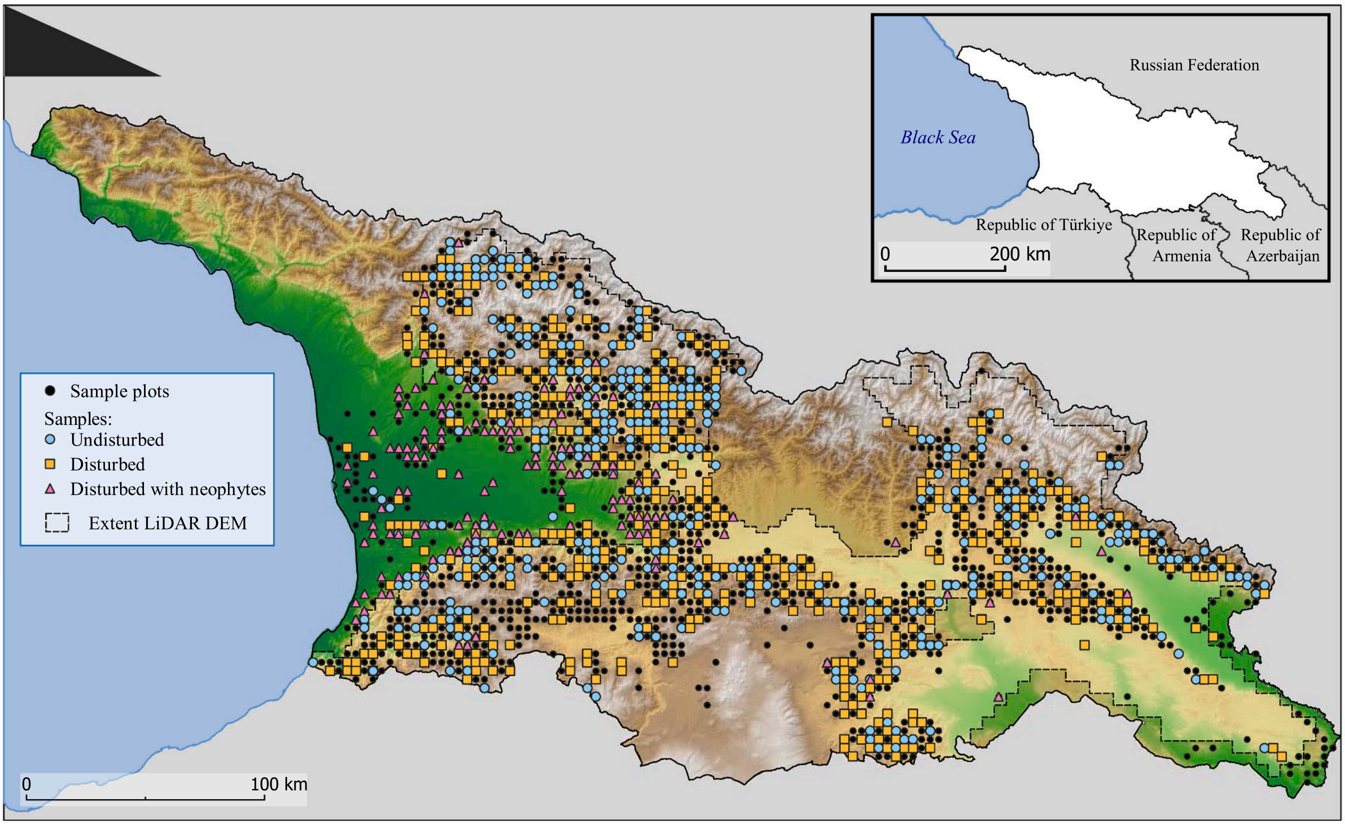


Fig. 2. Data overview of sample locations and applied digital elevation models. The SRTM based DEM is available for the entire country area. The dotted line outlines the extent of the LiDAR-based DEM. Sample plots (m = 4452) are shown as point features and represent all accessible sample plots located 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^{b} **ha**$^{-a}$				
			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

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 ($|p_i^a - p_i^b|$)
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 [°]

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).

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 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

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).

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 ($A_f = 180$ - | 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 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 (≥ 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).

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	

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 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),

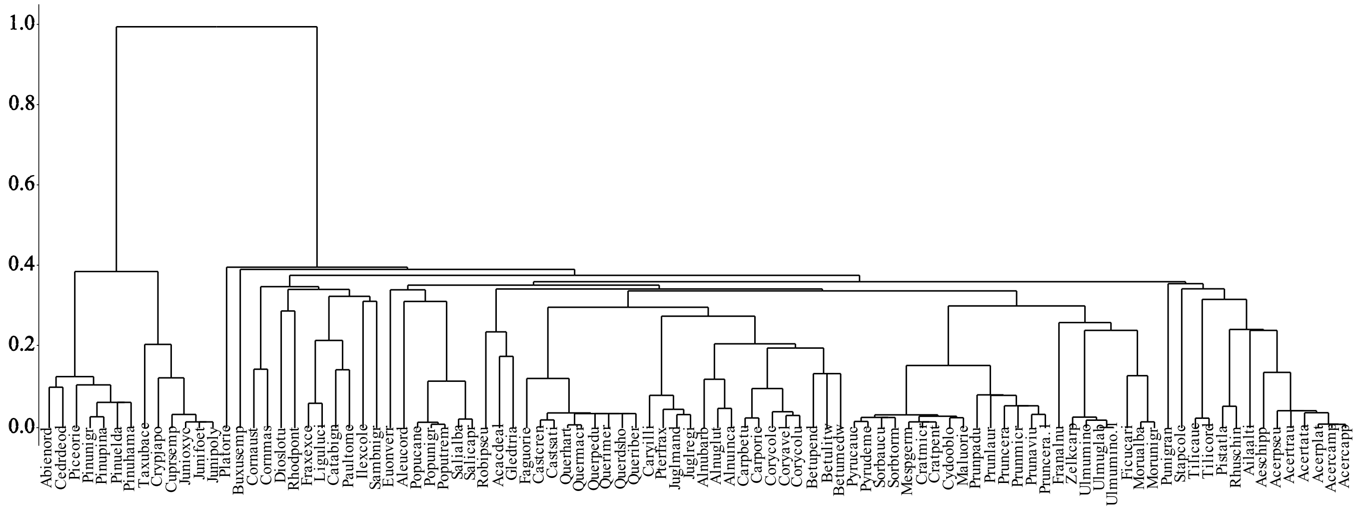


Fig. 3. Phylogenic tree of 96 species listed in the subsample (m = 1267) of the National Forest Inventory data of Georgia. Interspecies phylogenetic distances were normalized inside a square matrix of pairwise distance values so that $0 < \text{dist}_{ph} \leq 1$. Species names on the x-axis are abbreviated 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.

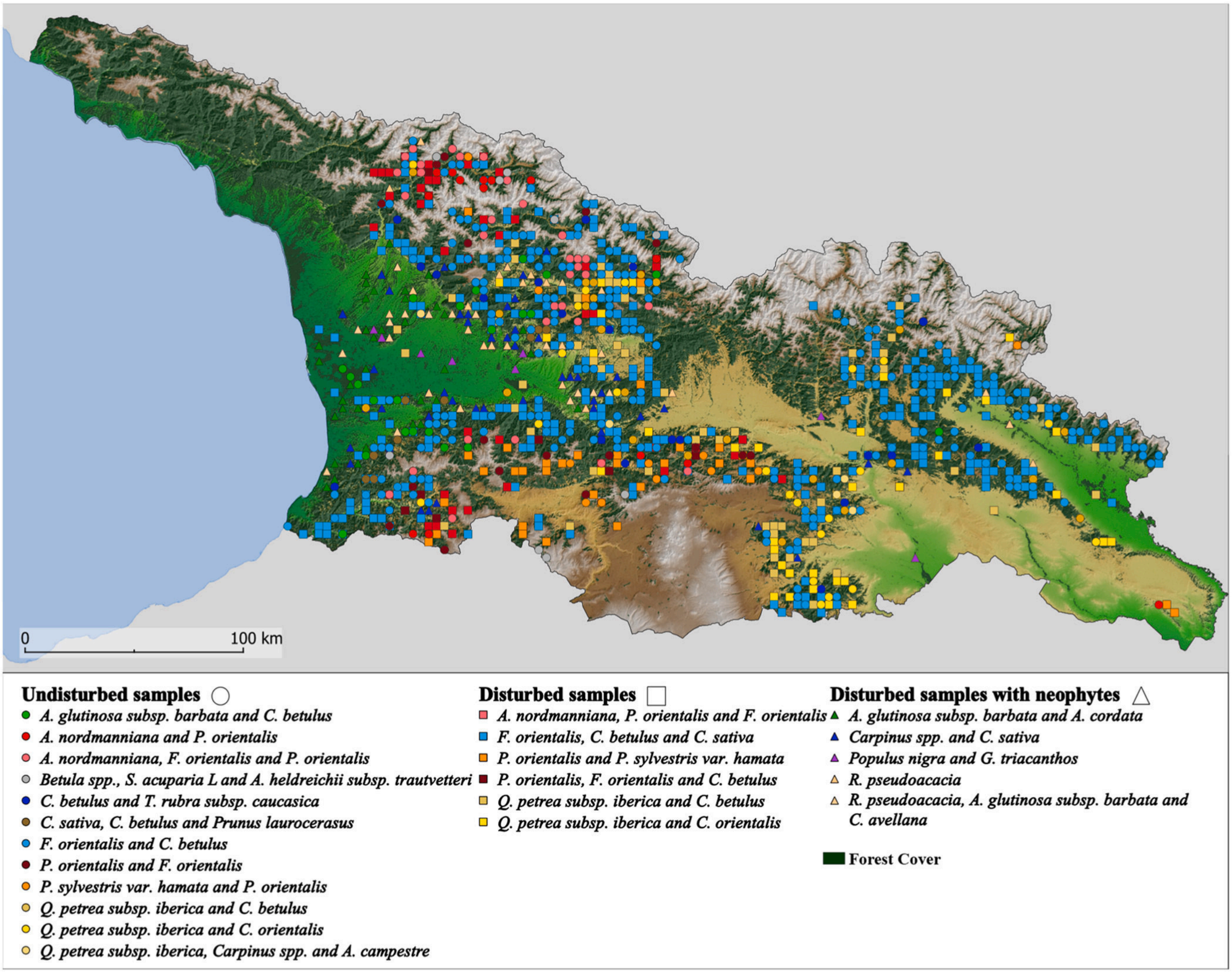


Fig. 4. Locations of clustered species assemblages per sample group (m = 1267). Clusters are labeled according to basal area share per species (≥10 %). Green-shaded areas indicate the distribution of forests in Georgia (Griesbach, 2018) for reference.

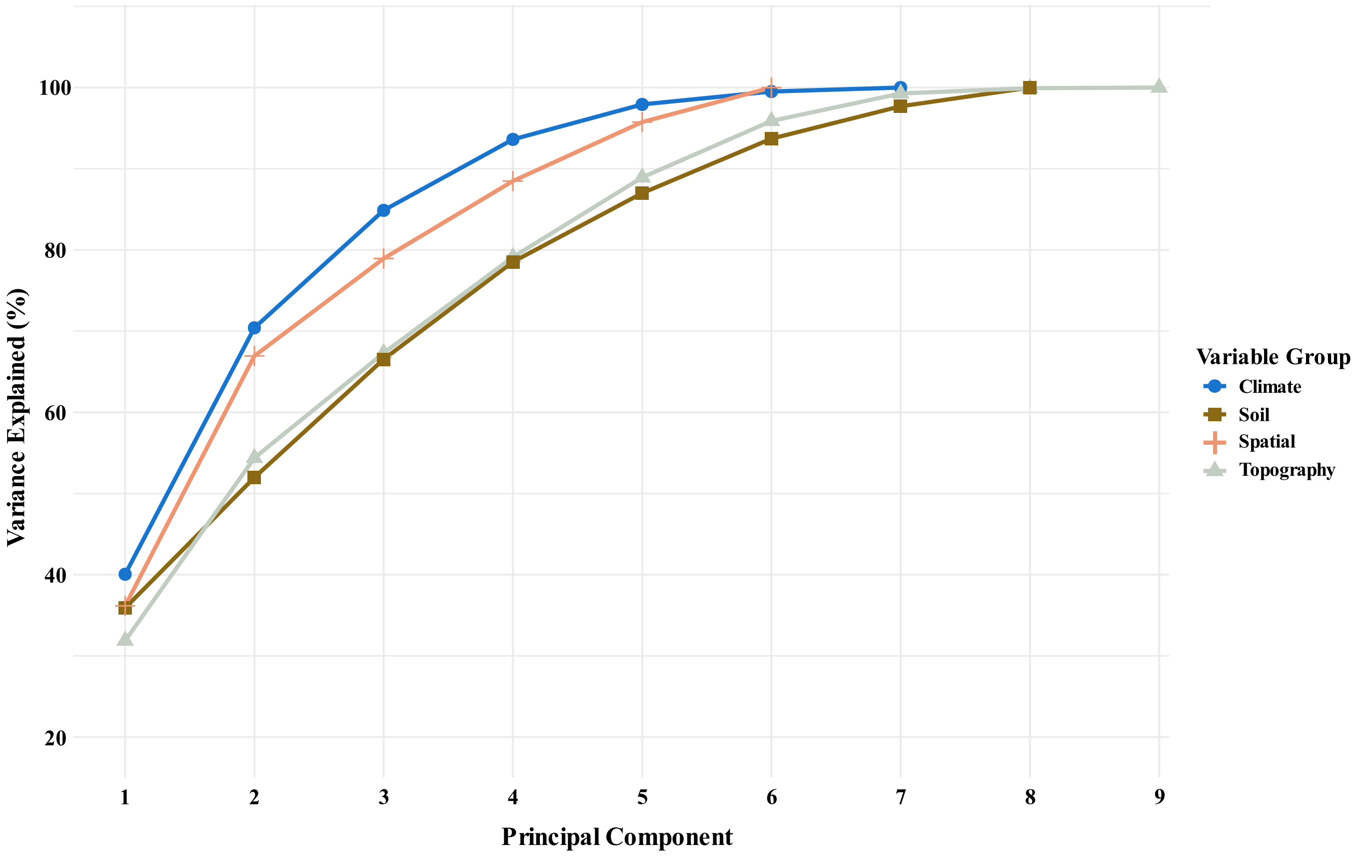


Fig. 5. Percentage of explained variance per principal component per variable group.

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.

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).

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

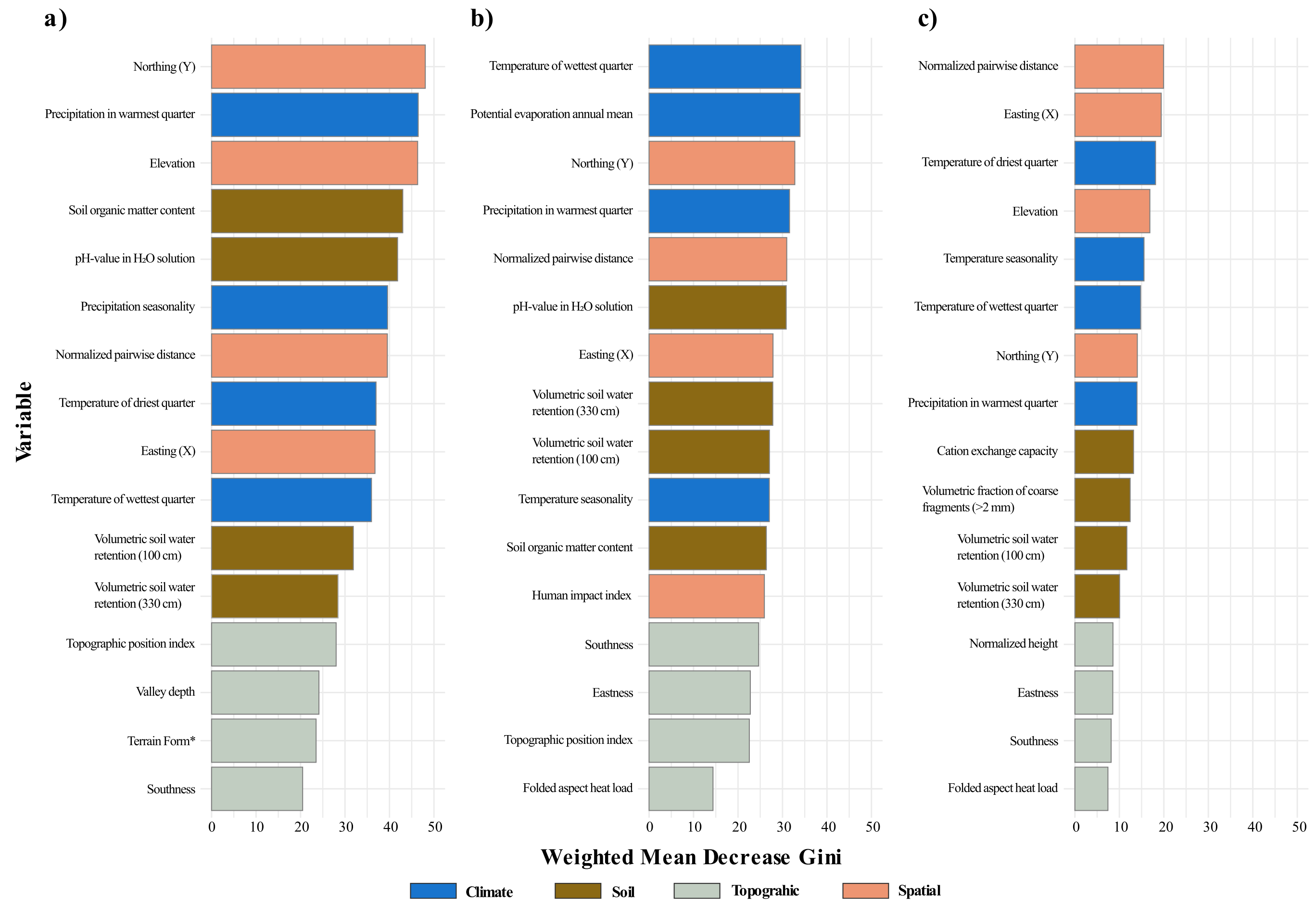


Fig. 6. Weighted Mean Decrease Gini indices of the four most important explanatory variables per variable and sample group. Variable importance is weighted according to contribution to the most important principal component to predict cluster membership of twelve species assemblages of undisturbed (a), 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.

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.