DOI: 10.2478/jlecol-2026-0031 Journal of Landscape Ecology (2026), Vol: 19 / No. 4

COUPLING DIVERSITY INDICES WITH BOOSTED REGRESSION TREES PREDICTS SPECIES ABUNDANCE AND SPATIAL COMMUNITY PATTERNS IN DRY AFROMONTANE FORESTS

GEBREYOHANNES ZENEBE[1]*, AMANUEL ZENEBE[1,2], ATKILT GIRMA[1,2], YEMANE KAHSAI[1], AKLILU NEGUSSIE[4], ROBERT MARCHANT[5], ASTER GEBREKIRSTOS[6] AND EMIRU BIRHANE[1,2,3]

[1]*Institute of Climate and Society (ICS), Mekelle University, P.O. Box 231, Mekelle, Ethiopia.*
[2]*Department of Land Resource Management and Environmental Protection, College of Dryland Agriculture and Natural Resources, Mekelle University, P.O. Box 231, Mekelle, Ethiopia.*
[3]*Faculty of Bioscience and Aquaculture, Nord University, P.O. Box 2501, NO-7729 Steinkjer, Norway, email: emiru.b.hizikias@nord.no, https://orcid.org/0000-0002-8644-5961*
[4]*WeForest, Ogentroostlaan 15, B-3090 Overijse, Belgium.*
[5]*Department of Environment and Geography, University of York, York YO10 5DD, United Kingdom.*
[6]*World Agroforestry Centre (ICRAF), United Nations Avenue, P.O. Box 30677-00100, Nairobi, Kenya.*
**Corresponding author email: zegere777@gmail.com*

Received: 16th August 2025, **Accepted:** 3rd April 2026

ABSTRACT

Understanding how species diversity interacts with spatial environmental gradients to shape landscape-level community patterns is a central question in contemporary landscape ecology. Although predictive modelling of biodiversity is advancing globally, dry Afromontane forests remain poorly represented in analyses that link spatial heterogeneity with community structure. This study addresses this gap by integrating diversity indices with a spatially explicit Boosted Regression Tree modelling framework to predict species abundance and community patterns across the Desa'a Forest, northern Ethiopia. Data from 301 plots were used to quantify woody species richness, abundance, and community composition, and to identify the environmental controls governing the distribution of dominant species. The forest contained ninety-four woody species and 7,207 individuals, with diversity values and plot-level abundances revealing a highly heterogeneous landscape composed of species-poor and species-rich patches. Cluster analysis delineated seven ecological communities, and dominant species such as *Cadia purpurea*, *Tarchonanthus camphoratus*, and *Juniperus procera* contributed disproportionately to spatial variation in forest structure. Boosted Regression Tree models showed that temperature, elevation, precipitation, and slope were the strongest predictors of *Juniperus procera* abundance, explaining sixty-three percent of deviance and revealing a narrow elevational and climatic

niche. These findings clarify how landscape-scale environmental gradients shape species distributions and community mosaics in dry Afromontane forests. By demonstrating a data-driven approach for predicting spatial community structure, this study provides a modelling framework that can enhance conservation planning and landscape-level management in biodiversity-rich mountain ecosystems worldwide.

Keywords: Desa'a forest, Evenness, *Juniperus procera*, northern Ethiopia, Richness.

INTRODUCTION

Understanding how biodiversity responds to environmental change is a central challenge in global ecology. Human activities generate a wide range of environmental pressures, including habitat alteration, pollution, and climate change, and these pressures are causing unprecedented impacts on global biodiversity (Keck *et al*., 2025). Forest ecosystems are particularly vulnerable, as they sustain more than 80 % of the world's terrestrial biodiversity (Karger *et al*., 2021). Within the African context, the Ethiopian Highlands, especially the dry Afromontane forests, stand out as centres of endemism, containing rare and ecologically specialized plant species (Enquist *et al*., 2019). These ecosystems contribute significantly to carbon sequestration, climate regulation, and rural livelihoods (Ahmed *et al*., 2022). However, these forests are increasingly degraded due to a convergence of anthropogenic and climatic pressures (Birhane *et al*., 2023; Tekle & Maryo, 2022). Agricultural expansion, fuelwood harvesting, overgrazing, and settlement have led to fragmentation, edge effects, and biodiversity loss (Dagne & Birhanu, 2023). Additional pressures such as tree cutting, girdling, selective logging, and grazing further disrupt forest structure and regeneration (Aynekulu *et al*., 2016), often shifting community composition toward disturbance-tolerant generalist species at the expense of specialists with narrow ecological niches.

Hutchinson's Niche Theory provides a useful framework for understanding these patterns, where species distributions are shaped by their multidimensional environmental tolerances and competitive interactions. Desa'a Forest, the largest remaining dry Afromontane Forest in northern Ethiopia, exemplifies how these niche dynamics play out in fragmented landscapes. The forest hosts key taxa such as *Juniperus procera, Olea europaea* subsp. cuspidata, and the endangered *Dracaena ombet*, yet woody vegetation cover has declined dramatically from 83 % in 1973 to 39 % in 2015 (Hishe *et al*., 2020). Drivers of this decline include agricultural encroachment, fuelwood extraction, overgrazing, illegal logging, charcoaling, and browsing (Aynekulu *et al*., 2016; Hishe *et al*., 2020). High-elevation specialists like *Juniperus procera* and *Olea europaea L. subsp. cuspidata* (Wall. & G.Don) Cif. are now increasingly restricted to fragmented highland refugia (Kassun *et al*., 2024; Zenebe *et al*., 2024), while lowland shrubs such as *Cadia purpurea* (Forssk.) Aiton and *Tarchonanthus camphoratus* L. encroach into upper elevations, displacing native forest assemblages (Haile *et al*., 2021). These shifts have consequences not only for biodiversity but also for ecosystem services central to local communities, such as hydrological regulation, microclimate moderation, and cultural values (Haile *et al*., 2021). The forest spans multiple agro-ecological zones, making its conservation a regional priority (Aynekulu *et al*., 2012) and an ideal landscape for examining pattern–process interactions.

Although numerous studies in Ethiopian highland forests have quantified plot-level patterns of woody species richness, diversity, and regeneration (Ahmed *et al*., 2022; Birhanu *et al*., 2025; Dagne & Birhanu, 2023; Dessie *et al*., 2025; Gebirehiwot *et al*., 2023; Reshad, 2019; Yinebeb *et al*., 2023), these efforts rely largely on traditional statistical approaches that are limited in their ability to capture nonlinear relationships and interactions among climate,

topography, and disturbance. Three major knowledge gaps, therefore, remain for dry Afromontane systems like Desa'a. First, there is a lack of analytical frameworks that integrate classical diversity metrics with advanced predictive models capable of capturing complex, nonlinear environmental responses. Second, spatially explicit predictions of species abundance and community composition across entire landscapes remain scarce, constraining the identification of priority areas for conservation and restoration. Third, despite their ecological importance, the environmental niche and landscape-scale abundance patterns of keystone species such as *Juniperus procera* remain poorly quantified.

Modeling species abundance effectively is, therefore, a priority in ecological research (Hill *et al*., 2017), as it enables the identification of environmental drivers shaping species distributions through advanced analytical approaches. As noted by Chitayat *et al*. (2024), spatially explicit metrics of tree cover are essential for evaluating habitat characteristics, resource availability, and environmental change. Consequently, analysing the spatial pattern of tree diversity across ecological gradients and land use types can provide critical evidence for targeted biodiversity conservation and management (Samuel *et al*., 2019).

Advances in spatial ecology increasingly rely on machine learning to address such gaps. Global analyses demonstrate that machine learning significantly improves models of plant diversity (Cai *et al*., 2023), and recent work highlights its growing role in sustainable forest management and biodiversity conservation (Espíndola *et al*., 2025). Among these methods, Boosted Regression Trees (BRTs) offer strong advantages for ecological modelling by capturing nonlinear responses, accommodating multicollinearity, and offering high predictive performance even with heterogeneous ecological datasets (Elith *et al*., 2008; Young *et al*., 2011). Their robustness to missing values and ability to identify key environmental predictors make them particularly suitable for modelling species abundance in data-scarce regions (Elith *et al*., 2008; Dedman *et al*., 2017; Irving *et al*., 2019). When combined with community similarity analyses and spatial mapping, BRTs provide a powerful approach for uncovering landscape-level ecological patterns

In this context, the present study seeks to advance landscape ecological understanding by integrating diversity indices with Boosted Regression Tree modelling to examine species abundance and spatial community patterns in the dry Afromontane Forest of Desa'a. The overarching aim is to generate a spatially explicit assessment of how woody species diversity and abundance respond to environmental gradients and disturbance regimes across this heterogeneous landscape. Specifically, the study evaluates patterns of woody species diversity and abundance across ecological and spatial gradients, delineates distinct ecological communities using similarity and cluster analyses, and models the abundance of *Juniperus procera* to quantify the influence of climatic, topographic, and disturbance variables. By producing a spatial prediction map of this key species, the study supports targeted conservation planning and contributes to broader efforts to manage and restore dry Afromontane forests.

Building on established theory and empirical patterns, we hypothesize that woody species diversity and abundance in Desa'a Forest exhibit strong, nonlinear responses to interacting climatic, topographic, and anthropogenic gradients, generating a fine-grained mosaic of species-rich and species-poor communities at the landscape scale. We further hypothesize that these communities can be statistically delineated and differ systematically in their environmental envelopes, with shrub-dominated assemblages associated with warmer, drier, and more heavily disturbed conditions. Finally, we expect *Juniperus procera* to occupy a relatively narrow environmental niche defined by elevation, temperature, precipitation, and slope, and that Boosted Regression Tree modelling will capture these constraints with high predictive accuracy. By coupling traditional diversity metrics with advanced modelling

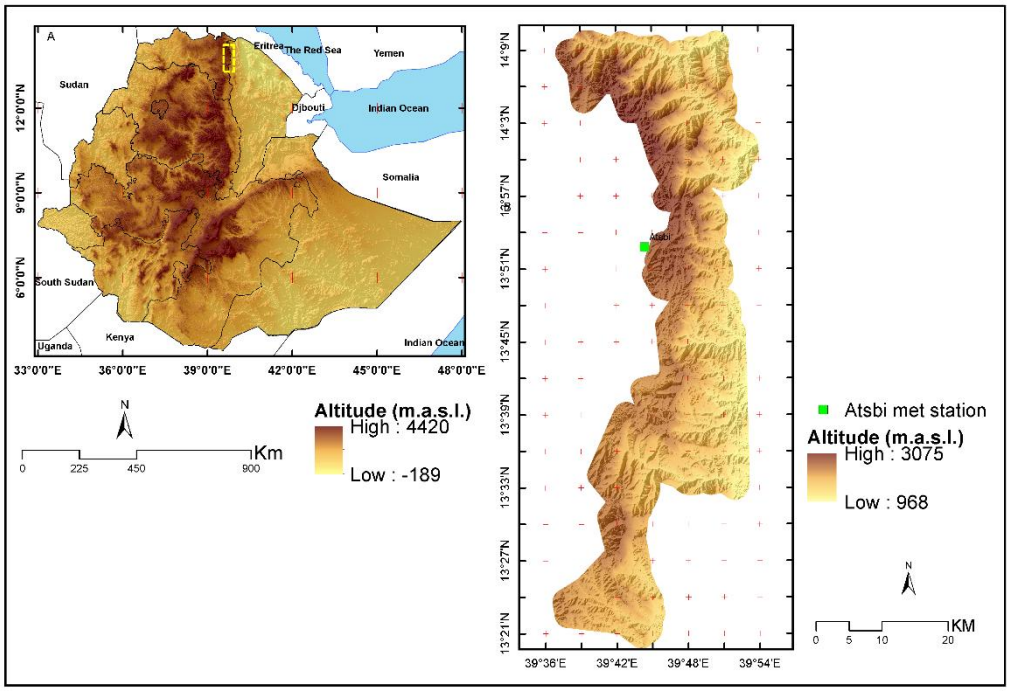
techniques and embedding them within a spatially explicit framework, this study contributes a novel and scalable approach for understanding biodiversity dynamics in fragmented mountain landscapes. The insights generated here have relevance not only for conservation planning in Ethiopia but also for other biodiversity-rich montane regions confronted with rapid environmental change.

METHODOLOGY

Study area

Desa'a Forest lies between 13.30°–14.18° N and 39.40°–39.92° E (Fig. 1), covering approximately 1,438 km² across the boundary of Tigray and Afar regional states (Kassun *et al*., 2024). The landscape is rugged, with elevations ranging from 968 to 3,075 m and slopes reaching up to 75°, forming strong topographic gradients that shape vegetation structure and distribution. The altitudinal and physiographic variation creates a transitional ecological zone linking the cooler highlands of Tigray with the arid Afar lowlands, producing diverse microclimates and moisture regimes.

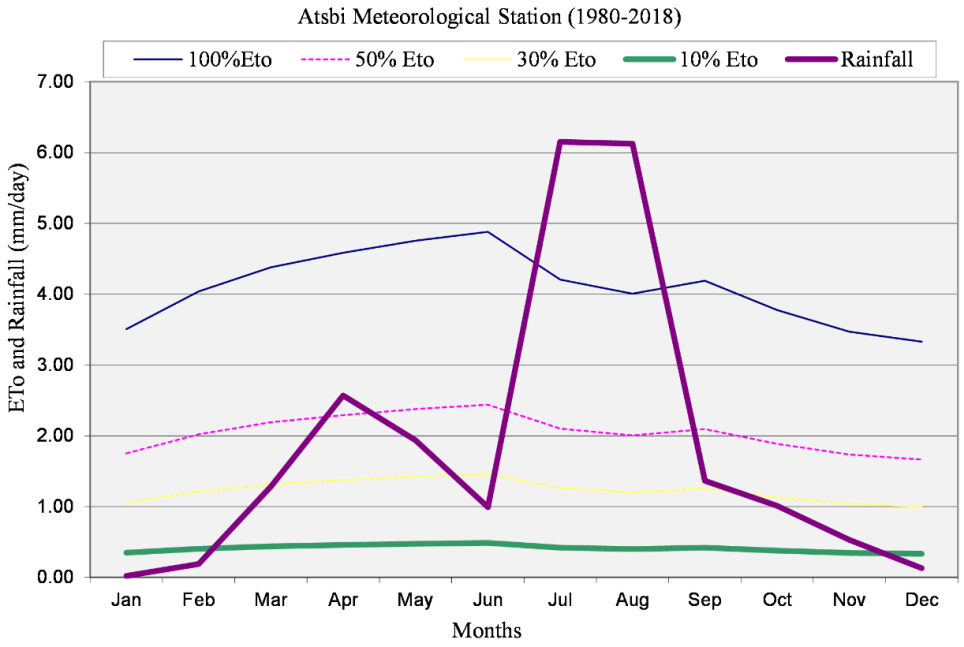
Fig. 1: Location map of the study area



Rainfall averages 500–600 mm annually in the Afar section, while highland areas experience a semi-arid to sub-humid climate (Tilahun *et al*., 2025). Seasonal water balance patterns (Fig. 2) indicate extended dry periods and a short but ecologically critical rainy season that supports peak vegetation growth. Although state-owned, the forest is heavily exploited for fuelwood, charcoal, and grazing (Tadesse *et al*., 2021), resulting in pronounced

disturbance gradients that directly relate to the study's hypotheses on biodiversity patterns and species abundance.

Fig. 2: Monthly rainfall and reference evapotranspiration (ET_0) in mm/day, showing seasonal water balance and periods when rainfall is below or above ET_0 thresholds (10 %, 30 %, 50 %, and 100 %)



Forest vegetation data and sampling method

To test hypothesis 1 (H1), which assumes that nonlinear responses of diversity and abundance to environmental gradients, we used a systematic vegetation inventory dataset collected by the WeForest project (https://www.weforest.org/). Plots were arranged in a 2 km grid, resulting in 303 plots of 20 × 20 m; two were excluded due to spatial inaccuracies, leaving 301 plots for analysis. Within each plot, all woody species were recorded, and trees ≥10 cm DBH and ≥1.5 m height were measured using diameter tapes and callipers. This spatially systematic, landscape-scale sampling design provides broad coverage of environmental gradients and supports the exploration of variation in richness, abundance, and community composition relevant to H1 and H2.

Biodiversity assessment and community analysis

To investigate patterns related to H1 and H2, woody species richness, abundance, Shannon diversity (H'), Simpson diversity (D), and evenness (E) across the 301 plots were quantified using the vegan package in R (Dixon, 2016). These metrics were used to describe how biodiversity varies across environmental gradients. Total abundance was defined as the number of woody individuals per plot, while species richness was calculated as the number of species with non-zero abundance within each plot.

Diversity indices were computed using standard formulations: (see Equations 1–4).
Species Richness (S):

$$S=\sum_{n=1}^{n} I\,(Ni>0)\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ (1)$$

Where: n is the total number of species observed, N_i: abundance (count) of species i, $I(N_i>0)$: indicator function, which is 1 if species i is present (abundance > 0), otherwise 0.
Shannon Diversity Index (H'):

$$H^{'}=\sum_{i=1}^{S} Pi\ln(Pi)\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ (2)$$

Where: S is the total number of species; P_i is the proportion of individuals belonging to species i (individuals of species i divided by total individuals); and $\ln(P_i)$ is the natural logarithm of that proportion.
Simpson Diversity Index (D):

$$D = 1 - \sum_{i=1}^{S}\left(pi^{2}\right)\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ (3)$$

Where: S is the total number of species; and P_i is the proportion of individuals belonging to species i (individuals of species i divided by total individuals). Evenness was calculated as:

Evenness (E): $$E=\frac{H^{'}}{\ln(S)}\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ (4)$$

Where: H′ is the Shannon diversity index; S is the number of species; and ln(S) is the natural logarithm of S, representing the maximum possible entropy when all species are equally abundant.

Spatial visualization of biodiversity patterns was achieved by converting plot coordinates into spatial objects using the sf package in R (Pebesma, 2018). Biodiversity metrics were then linked to these spatial datasets to produce maps of species richness, abundance, Shannon and Simpson indices, and evenness. These maps were generated using ggplot2 (Wickham, 2010), allowing spatial patterns of diversity to be interpreted across the study area - a central component of H1.

To evaluate nonlinear responses of biodiversity metrics along environmental gradients (H1), Generalized Additive Models (GAMs) were applied using the mgcv package in R. GAMs allow smooth, flexible estimation of relationships between biodiversity metrics (species richness, abundance, H′, D, and E) and environmental predictors (Lai *et al*., 2024). Models were structured as:

$$Y_i = \beta_0 + \sum_{k=1}^{p} s_k(X_{ki}) + \varepsilon_i\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ (5)$$

where Y_i is the response variable for observation i; β_0 is the intercept; $s_k(\cdot)$ denotes smooth functions applied to the k-th predictor; X_{ki} represents the value of the k-th environmental predictor for observation i; p is the total number of predictors included in the model; and ε_i is the residual error term. Poisson distributions with log links were used for count data (richness and abundance), and Gaussian distributions for continuous diversity indices (H′, D, E). Flexible responses were inferred when the estimated degrees of freedom (edf) for a smooth

term exceeded 1 and $p < 0.05$. Model adequacy and diagnostics were also assessed. This approach provides a statistical test for nonlinear biodiversity responses along environmental gradients, addressing H1.

To assess sampling completeness and strengthen confidence in observed richness patterns, species accumulation and rarefaction curves were generated from Hellinger-transformed abundance data. This transformation reduces the influence of dominant species and is appropriate for community-level analyses based on Euclidean distances. Sampling adequacy was inferred from the point at which the accumulation curve approached a plateau, defined as fewer than one newly detected species per sampling unit.

To explore hypothesis 2 (H2), which addresses the presence of distinct community groupings, hierarchical cluster analysis was applied to the Hellinger-transformed data using Euclidean distance and Ward's minimum variance method (Borcard *et al*., 2018). The optimal number of clusters was determined through a combination of the Elbow method (based on within-cluster sum of squares) and the Silhouette method, complemented by visual inspection of the dendrogram to ensure a balance between cluster compactness and separation. Differences in environmental space among clusters were evaluated using a Permutational Multivariate Analysis of Variance (PERMANOVA) based on Bray–Curtis dissimilarities. Homogeneity of multivariate dispersion tests were used to assess within-cluster variability, while Kruskal–Wallis tests confirmed that clusters occupy significantly distinct environmental conditions. Cluster structures were visualized using dendextend (Galili, 2015), and rank–abundance curves were used to examine within-cluster dominance and evenness patterns. Community groups were subsequently interpreted relative to environmental variables gradients.

Environmental predictors and spatial modeling of *Juniperus procera* abundance

This component directly addresses hypothesis 3 (H3), which posits that *Juniperus procera* occupies a narrow environmental niche and that Boosted Regression Tree (BRT) models will capture this niche with high predictive accuracy.

Predictor Variables

To evaluate the environmental determinants of *Juniperus procera* abundance and generate spatially explicit predictions, plot-level count data were combined with bioclimatic, topographic, moisture-related, and proximity-based predictors. Nineteen bioclimatic variables were derived from CHELSA daily climate data (1979–2019) using the biovars function in the dismo package in R software. The CHELSA dataset was accessed via the ISIMIP repository at https://data.isimip.org/10.48364/ISIMIP.342217. Topographic predictors (elevation, slope, aspect) were extracted from a DEM, while vegetation moisture stress was represented by median Normalized Difference Moisture Index (Median_NDMI) calculated from Landsat 8 NIR and SWIR bands (see Appendix 1). Proximity metrics, defined as distances to roads, settlements, water sources, and streams, were generated using Euclidean Distance in ArcMap. All predictor layers were standardized to a consistent spatial resolution and projection.

Boosted Regression Tree modeling of Juniperus procera abundance

Boosted Regression Tree (BRT) was selected specifically to evaluate H3, given its ability to model nonlinear relationships and interactions among climate, topography, and disturbance (Elith *et al*., 2008). Models were implemented in R using the gbm (Elith *et al*., 2008), and dismo packages, employing a Poisson distribution, appropriate for modeling

count data (Agis *et al*., 2024). Model settings included a tree complexity of 2, a learning rate of 0.001, a maximum number of 2550 trees, and a bag fraction of 0.75.

Model validation and spatial prediction of Juniperus procera abundance

Model performance was evaluated in the context of H3 by comparing observed counts with predicted values extracted at georeferenced sampling locations. All performance metrics were derived from out-of-sample predictions to ensure unbiased evaluation. Model accuracy for continuous responses was quantified using mean squared error (MSE), root mean squared error (RMSE), mean absolute error (MAE), Pearson correlation (r), and the coefficient of determination (R^2). Two complementary validation strategies were employed. First, ten-fold random cross-validation was conducted, whereby models were trained on 90 % of the data and evaluated on the remaining 10%. Second, spatial block cross-validation was implemented to account for spatial dependence by grouping sampling locations into five clusters (k=5) using k-means clustering; entire clusters were withheld during model training and subsequently used for validation. Model discrimination was further assessed by converting continuous predictions into binary outcomes using an optimal threshold criterion, followed by calculation of the area under the receiver operating characteristic curve (AUC), as well as accuracy, sensitivity, and specificity. Given the spatially structured sampling design, residual spatial autocorrelation was evaluated using Moran's I based on a k-nearest neighbour spatial weights matrix. To facilitate interpretation, variable importance metrics and partial dependence plots were examined to assess the relative influence of key predictors and to evaluate the extent to which *Juniperus procera* is constrained by elevation, temperature, precipitation, and slope, as hypothesized in H3.

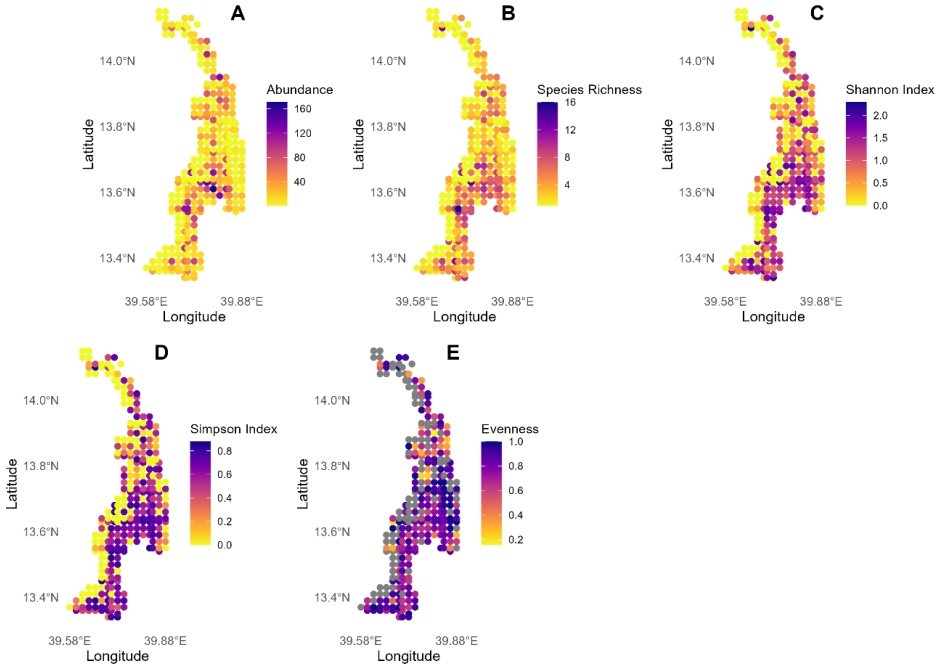
Finally, the BRT model was projected across the study area to generate spatially explicit abundance maps. These maps identify ecological hotspots, potential refugia, and areas of heightened sensitivity, thereby directly supporting conservation planning and forest management objectives.

RESULTS

Spatial variation in woody species abundance and diversity patterns

Woody species abundance and richness showed pronounced spatial heterogeneity across the 301 plots (Fig. 3A–B). Species richness varied from 1 to 16 species per plot (mean ± SD = 3.68 ± 2.9), while total abundance ranged from 1 to 171 individuals per plot (mean ± SD = 23.94 ± 31.1). However, approximately 85 % of the plots contained fewer than 50 individuals, indicating that dense woody stands are spatially restricted and highly localized. Only a small subset of plots exceeded 100 individuals, forming discrete high-density pockets within an otherwise sparsely populated matrix. Across the forest, 94 woody species and 7,207 individuals were recorded, pointing to a relatively diverse but unevenly distributed flora.

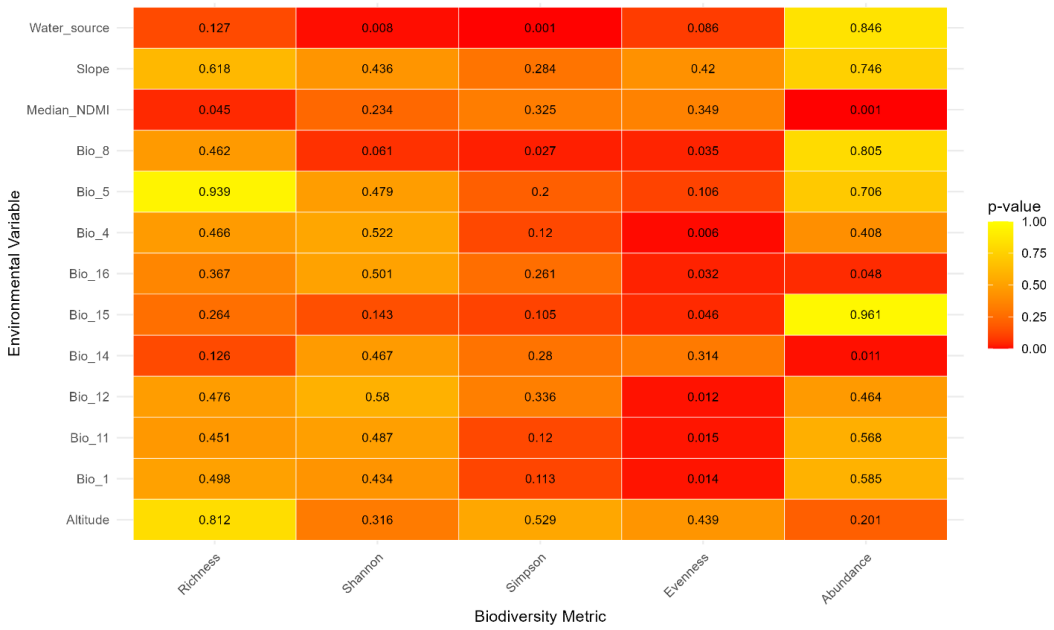
Fig. 3: Spatial variation of total species abundance (A), richness (B), Shannon Index (C), Simpson Index (D), and Evenness (E) across the sampled plots



Spatial patterns in Shannon diversity, Simpson diversity, and evenness revealed a fine-grained mosaic of community structure (Fig. 3C–E). Shannon diversity (H′) ranged from 0 to 2.29 (mean ± SD = 0.75 ± 0.65); while Simpson diversity (D) varied from 0 to 0.88 (mean ± SD = 0.39 ± 0.31), showing broadly similar spatial patterns. Evenness (E) ranged from 0.16 to 1.0 (mean ± SD = 0.76 ± 0.2), but declined in plots dominated by a few highly abundant species.

Generalized Additive Models (GAMs) incorporating a range of environmental predictors showed that most variables had weak or non-significant effects on biodiversity metrics, with significant relationships largely limited to moisture and habitat-related factors (Fig. 4 and Appendix 2). Species richness was influenced by normalized difference moisture index (Median_NDMI; edf = 1.00, p = 0.045), reflecting a primarily linear effect. All other predictors, including temperature and precipitation variables (Bio_5, Bio_8, Bio_12, Bio_16), topography (altitude, slope), and distance to water sources, were non-significant (p > 0.1). Altitude showed a weak, non-significant nonlinear trend (edf = 2.57, p = 0.81). The model explained 27.5% of the deviance. Shannon diversity (H′) was significantly associated with distance to water sources (p = 0.008), suggesting local habitat conditions outweigh broad-scale climatic effects.

Fig. 4: Heatmap of GAM p-values for environmental predictors across biodiversity metrics. Red cells indicate low p-values (high significance) and yellow cells indicate high p-values (non-significant)

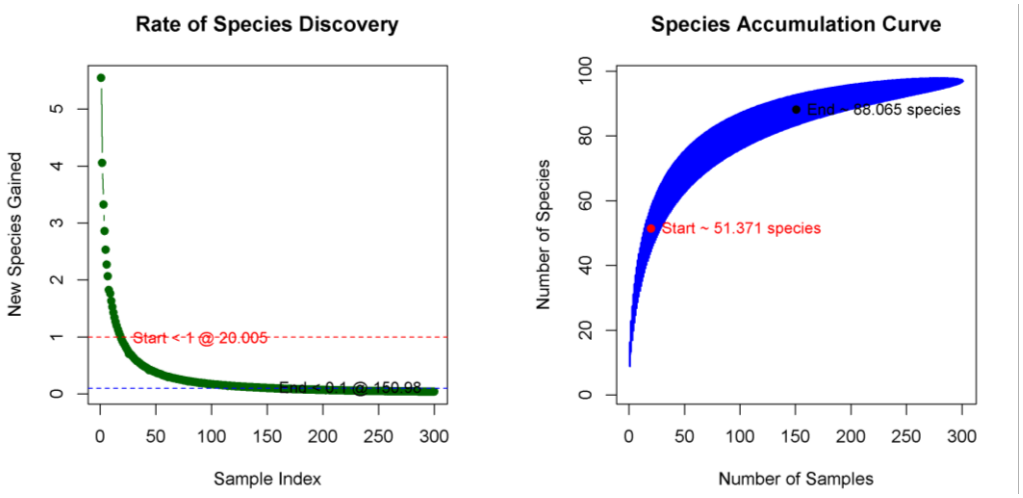


Climatic variables had weak or non-significant influences, although Bio_8 (mean temperature of the wettest quarter) showed a marginal effect ($p = 0.061$). The model explained 31 % of the deviance. Simpson diversity (D) also responded strongly to distance to water sources ($p = 0.001$), with a modest contribution from Bio_8 ($p = 0.027$). The model explained 34.3 % of the deviance. Evenness exhibited broader environmental sensitivity. Temperature (Bio_1, Bio_8, Bio_11), precipitation (Bio_12, Bio_16), and climate variability (Bio_4, Bio_15) had significant or near-significant effects ($p < 0.05$), mostly linear (edf ≈ 1). Distance to water sources showed a marginal effect ($p = 0.086$). The model explained 29.9 % of the deviance. Total abundance was strongly driven by Median_NDMI (edf = 1.00, $p < 0.001$). Significant nonlinear responses were detected for precipitation of the wettest quarter (Bio_16; edf = 4.59, $p = 0.048$) and driest periods (Bio_14; edf = 3.72, $p = 0.011$), highlighting complex responses to precipitation extremes. All other predictors were non-significant. The model explained 37.5 % of the deviance.

Sampling completeness and community structure

Species accumulation analysis confirmed that the sampling effort was sufficient to characterize the woody flora (Fig. 5). The rate of new species detections declined sharply after the first few dozen plots and approached a plateau around 150 plots, where cumulative richness stabilized near the observed total. This indicates that the main patterns of richness and composition are unlikely to be driven by under sampling.

Fig. 5: Species discovery and accumulation curves: (Left) Discovery rate (green) with thresholds at 1 (red) and 0.1 (blue). (Right) Accumulation curve (blue) showing cumulative species richness; shaded area = confidence interval, illustrating sampling completeness and biodiversity trends.



Hierarchical clustering of Hellinger-transformed abundance data identified seven distinct woody plant communities (Fig. 6; Table 1), as supported by both the Elbow and Silhouette methods. Clusters 3 and 4 combined the highest species richness (50 species each) with relatively low total abundance, indicating diverse but sparsely populated assemblages. In contrast, Cluster 7 had comparatively low richness (19 species) but one of the highest total abundances, dominated by a few species. Permutational Multivariate Analysis of Variance (PERMANOVA) confirmed that species composition differed significantly among clusters ($F = 33.85$, $p < 0.001$), while tests for homogeneity of multivariate dispersions indicated significant differences in within-cluster variability ($F = 221.01$, $p < 0.001$), validating that clusters are ecologically meaningful.

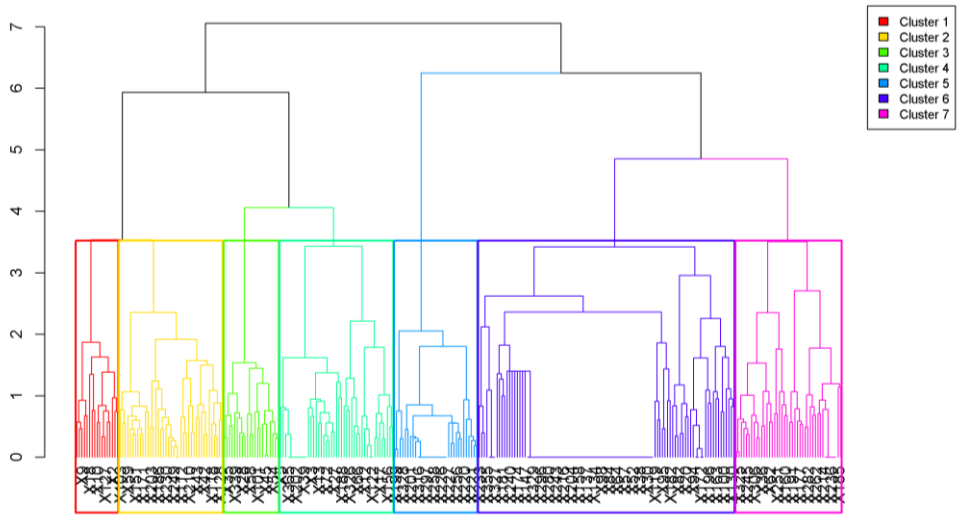
Table 1: Summary of vegetation clusters by species richness, abundance, and dominance.

Cluster	Total Plots	Unique Species	Total Abundance	Dominant Species
1	17	32	197	*Premna resinosa* Hochst. ex Schauer
2	45	41	336	*Vachellia etbaica* Schweinf.
3	101	50	64	*Grewia kakothamnos* K.Schum.
4	41	50	877	*Cadia purpurea* Aiton
5	22	27	466	*Tarchonanthus camphoratus* L.
6	42	41	483	*Dodonaea angustifolia* L.f.
7	33	19	672	*Juniperus procera* Hochst. ex Endl.

Indicator species analysis identified species strongly associated with each cluster, including *Premna resinosa* for Cluster 1, *Vachellia etbaica* for Cluster 2, and *Juniperus procera* for Cluster 7, demonstrating distinct species assemblages. Kruskal-Wallis tests on

environmental variables showed highly significant differences among clusters ($p < 0.001$ for most variables), with marked variation in temperature, precipitation, and altitude ranges, indicating that these communities occupy statistically distinct environmental niches.

Fig. 6: Ward's hierarchical clustering of Hellinger-transformed species abundance in Desa'a forest



Relative abundance was strongly skewed toward a few species, with the 13 most abundant accounting for 83 % of individuals (Table 2). *Cadia purpurea, Tarchonanthus camphoratus*, and *Juniperus procera* alone contributed over 40 % of total abundance, while many species occurred as singletons or in low numbers. This dominance pattern indicates that community mosaics are structured by a limited set of ecologically influential species, supporting the hypothesis that distinct community types reflect contrasting environmental conditions and disturbance regimes.

Table 2: Top 13 abundant species across the 301 plots in Desa'a

Species	Total	Relative Abundance
Cadia purpurea (Forssk.) Aiton	1061	0.1482
Tarchonanthus camphoratus L.	928	0.1297
Juniperus procera Hochst. ex Endl.	894	0.1249
Dodonaea angustifolia L.f.	672	0.0939
Vachellia etbaica (Schweinf.) Kyal. & Boatwr.	665	0.0929
Rhus natalensis Bernh. ex Krauss	414	0.0578
Olea europaea subsp. cuspidata (Wall. ex G.Don) Cif.	332	0.0464
Premna resinosa Hochst. ex Schauer	228	0.0319
Carissa edulis Vahl	213	0.0298
Euclea racemosa subsp. schimperi (A.DC.) F. White	192	0.0268
Calpurnia aurea (Aiton) Benth.	160	0.0224
Grewia kakothamnos K.Schum.	133	0.0186
Maytenus arbutifolia Hochst. ex A. Rich.	113	0.0158

Performance assessment and spatial modeling of *Juniperus procera* abundance

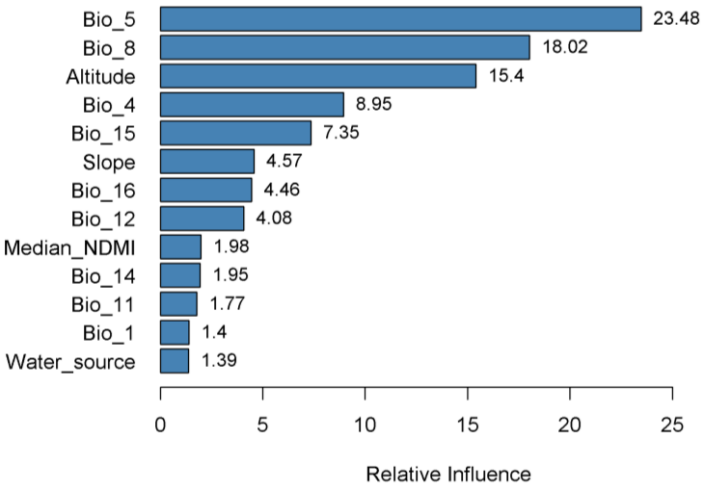
Model performance

The Boosted Regression Tree model for *Juniperus procera* abundance showed strong predictive performance. The optimal model with 2,550 trees, explained 63 % of the deviance (pseudo-R^2 = 0.63), indicating that it captured the major spatial structure in abundance. Predictive accuracy was moderate, with a mean squared error of 30.27, root mean squared error (RMSE) of 5.50, and mean absolute error (MAE) of 4.40. Observed and predicted abundances were closely aligned (Pearsons r = 0.88), confirming a good fit to field data. Residual spatial autocorrelation was negligible (Moran's I = 0.0016, p = 0.44), indicating that spatial dependence did not bias predictions. When sites were classified into high versus low abundance using the median, the model performed well for classification, achieving an area under the curve (AUC) of 0.91, overall accuracy (OA) of 0.82, sensitivity of 0.86, and specificity of 0.80.

Variable importance

Modelled abundance of *Juniperus procera* was driven by a relatively narrow suite of climatic and topographic variables (Fig. 7). Maximum temperature of the warmest month (Bio_5) and mean temperature of the wettest quarter (Bio_8) contributed the most, followed by altitude. Additional influences came from temperature and precipitation seasonality (Bio_4 and Bio_15), slope, precipitation of the wettest quarter (Bio_16), annual precipitation (Bio_12), and moisture availability (Median_NDMI). The top 13 predictors (each >1 % relative influence) accounted for nearly all model explanatory power, indicating that *Juniperus procera* abundance is shaped by a concise yet highly influential environmental niche defined by temperature, precipitation regimes, elevation, and terrain steepness.

Fig.7: Relative influence plot of the top 13 predictor variables



Partial dependence plots revealed clear ecological thresholds in *Juniperus procera* responses to environmental gradients. Predicted abundance peaked at intermediate values of maximum temperature of the warmest month (about 25–27.6 °C) and mean temperature of the wettest quarter (about 14–16.5 °C) and declined outside these ranges. Favourable

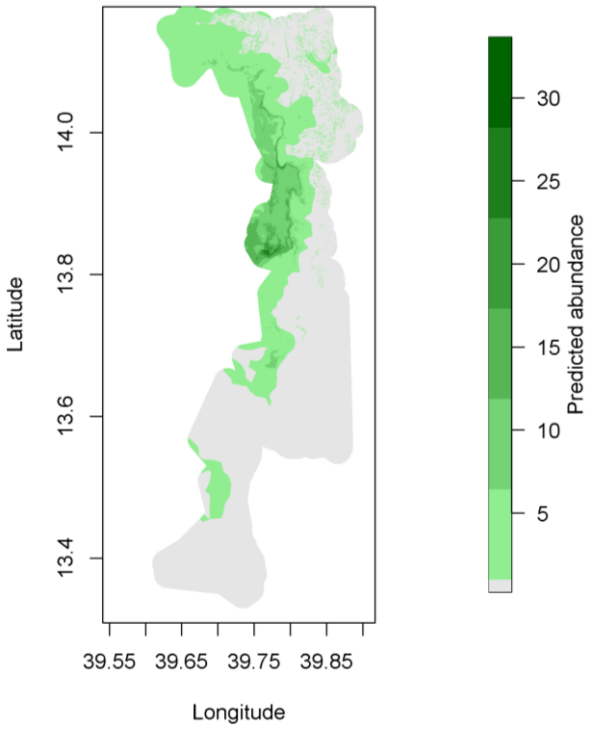
conditions were also associated with moderate temperature seasonality and relatively cool mean temperatures in the coldest quarter.

Precipitation-related variables showed similarly constrained ranges, with the highest predicted abundance at intermediate precipitation seasonality, wettest-quarter precipitation of 375–385 mm, and annual precipitation of 619–680 mm. Topographic and moisture gradients reinforced these patterns: abundance increased sharply between 2,300 and 3,000 m, peaking around 2,500–2,700 m, with suitability rising up to steep slopes (~75%) and stabilizing thereafter (Appendix 3). Optimal Normalized Difference Moisture Index (NDMI) values ranged from slightly negative to near zero, indicating a preference for moderately moist, non-waterlogged conditions. Together, these patterns indicate that *Juniperus procera* occupies a narrow climatic and topographic niche, with distribution tightly constrained by interacting thermal, hydrological, and terrain factors, consistent with the third hypothesis.

Spatial prediction map of Juniperus procera abundance

Field observations recorded *Juniperus procera* in 63 of the 301 plots, with 894 individuals and highly variable plot-level counts. The spatial prediction map derived from the BRT model (Fig. 8) extended these point observations to the full landscape, revealing a strongly fragmented pattern of abundance.

Fig. 8: Spatial Abundance Map of Juniperus procera Predicted by Boosted Regression Trees



Predicted values ranged from near zero to more than 30 individuals per 20 × 20 m pixel, with most of the forest showing low to negligible predicted abundance and a limited number

of high-suitability patches forming potential refugia. These hotspots were concentrated in cooler, high-elevation areas with steep slopes and intermediate moisture conditions, consistent with the identified niche. The spatial prediction, therefore, demonstrates that suitable habitat for *Juniperus procera* is both environmentally constrained and spatially discontinuous, highlighting the vulnerability of this keystone species to further habitat loss and underscoring the value of niche-based modelling for guiding conservation and restoration priorities at the landscape scale.

DISCUSSION

Methodological advances and their scientific value

This study combined classical biodiversity assessment with hierarchical community analysis and machine-learning–based abundance modelling to characterize woody species patterns across one of the largest remaining dry Afromontane forests in northern Ethiopia. This methodological integration directly addresses the need for spatially explicit biodiversity assessments that can capture nonlinear ecological processes in heterogeneous mountain landscapes. Plot-based floristic inventories remain essential for quantifying richness, abundance, and structural attributes, yet their value increases substantially when combined with spatial statistics and predictive modelling. The 301 systematically distributed plots provided a robust empirical foundation for quantifying variability in richness and abundance across steep environmental gradients. The sufficiency of this sampling was supported by the species accumulation curve, which plateaued well before full sampling coverage was reached, consistent with theoretical expectations for well-designed woody vegetation surveys (Dagne & Birhanu, 2023; Wassie *et al.*, 2024).

Cluster analysis using Ward's method effectively delineated ecologically interpretable plant communities, identifying seven distinct clusters. The combination of hierarchical clustering and diversity indices revealed ecological heterogeneity while minimizing the disproportionate influence of rare species through Hellinger transformation, improving the reliability of multivariate analyses. The combined use of PERMANOVA and tests for homogeneity of multivariate dispersions strengthened inference by distinguishing true compositional differences from variability within groups. Furthermore, indicator species analysis added diagnostic power by identifying species-specific associations with clusters. Collectively, this integrative analytical approach represents a methodologically rigorous framework for quantifying community structure and linking species composition to ecological processes, thereby enhancing the scientific robustness and interpretability of biodiversity assessments.

A major methodological contribution of this study lies in applying Boosted Regression Trees (BRTs) to model the abundance of *Juniperus procera*. The BRT captured 63 % of deviance and produced a strong correlation between observed and predicted values (Pearson $r = 0.88$), demonstrating strong predictive power. Machine-learning methods such as BRTs allow for the characterization of complex nonlinear species–environment relationships, outperforming traditional linear models in many ecological contexts (Elith *et al*., 2008; Elith & Leathwick, 2009). Importantly, BRTs are robust to outliers, multicollinearity, and missing data, making them well-suited for regions where environmental data layers may be uneven or incomplete (Dormann *et al*., 2007). The inclusion of climatic, topographic, remote sensing-based indices, and proximity-based predictors offers a holistic environmental envelope for interpreting species abundance patterns. Collectively, this methodological integration of field inventories, hierarchical clustering, and machine-learning modelling

provides strong support for H2 and H3, demonstrating discrete community types and the narrow environmental niche of *Juniperus procera*.

Diversity patterns, community structure, and species niches

Woody species diversity and abundance across gradients

Desa'a forest hosts 94 woody species and 7,207 individuals, highlighting its floristic importance relative to other dry Afromontane forests such as Hurbu Natural Forest (32 species) (Gebirehiwot *et al*., 2023), Hugumburda Forest (79 species) (Aynekulu *et al*., 2016), Endba-Zendy Forest (77 species) (Dessie *et al*., 2025), and Wof Ayzurish Forest (76 species) (Tefera *et al*., 2024). Although it does not reach the exceptionally high species counts of the Guard dry Afromontane forest (137 species) (Dagne & Birhanu, 2023) and forest in the Gozamin district of Gojam (168 species) (Yinebeb *et al*., 2023), Desa'a represents a major floristic refuge in an increasingly fragmented landscape.

Species richness (1–16 species per plot) and abundance (1–171 individuals per plot) exhibited pronounced spatial heterogeneity, consistent with patterns observed in other dry Afromontane systems (Deressa *et al*., 2023; Sherafu *et al*., 2024). Low stand density in most plots (<50 individuals) suggests that large areas of the forest have undergone structural simplification relative to typical Afromontane conditions, where densities often exceed 100 individuals per similar plot sizes (Deressa *et al*., 2023; Woldearegay & Woldu, 2020). Diversity indices revealed gradients from species-poor, dominance-driven stands to more uniformly structured communities. Areas with low Shannon and Simpson values correspond to agricultural and barren regions, reflecting habitat filtering, while higher evenness values point to intact or recovering communities (Magurran, 2004).

Generalized additive models indicated that most broad-scale climatic and topographic variables had weak or non-significant effects on biodiversity, while local-scale moisture availability and proximity to water sources were the strongest predictors. Overall, diversity metrics (richness, Shannon, Simpson, and evenness) showed predominantly linear or weak relationships with environmental gradients, providing limited support for hypothesis 1. Notably, nonlinear effects were mainly observed in total abundance, particularly in response to precipitation extremes, suggesting greater sensitivity to climatic variability. Together, these findings imply that broad-scale climate sets the potential niche space, but local habitat conditions largely determine realized biodiversity patterns, with nonlinear dynamics emerging primarily under extreme environmental conditions.

Community structure along ecological gradients

The results demonstrate that the study area is structured into distinct and ecologically meaningful plant community clusters, indicating non-random patterns of community assembly. Significant differences in species composition among clusters, supported by PERMANOVA, together with clear indicator species associations, confirm that each cluster represents a unique species assemblage. These compositional differences are strongly aligned with environmental variation, as clusters occupy statistically distinct niches defined by temperature, precipitation, and altitude gradients.

The dominance structure further supports this interpretation. Thirteen species accounted for 83 % of total abundance, with *Cadia purpurea, Tarchonanthus camphoratus*, and *Juniperus procera* forming the core of the community. The expansion of *C. purpurea* and *T. camphoratus* is widely associated with forest degradation and canopy opening, consistent with previous findings that disturbance is shifting vegetation from *Juniperus–Olea* dominance toward generalist shrub dominance (Haile *et al*., 2021; Aynekulu, 2012). The

presence of numerous rare species with single occurrences indicates that both dominant and rare taxa contribute to overall ecosystem stability, echoing the argument by Hordijk *et al*. (2024) that conservation actions must safeguard the full spectrum of species to maintain ecosystem resilience. Overall, the findings provide strong support for H2, showing that plant communities are discretely organized along environmental gradients, with some species exhibiting narrow ecological preferences, reinforcing the link between species distribution and specific habitat conditions.

Environmental controls on Juniperus procera abundance

In contrast to community-level diversity, the BRT model provided strong evidence that *Juniperus procera* abundance is controlled by a narrow set of temperature-, precipitation-, and elevation-related thresholds. Optimal abundance occurred under maximum warm-month temperature (Bio_5) of 25–27.6°C, mean wettest-quarter temperature (Bio_8) of 14–16.5°C, and moderate seasonal thermal variability. The dominance of thermal predictors aligns with the species' known sensitivity to high evaporative demand (Siyum *et al*., 2019). The species showed similarly narrow precipitation preferences, with highest suitability at moderate annual totals (619–680 mm) and seasonality values around 1006–1007 mm.

Topography was equally influential, with elevation between 2,500 and 2,700 m displaying peak suitability, reaffirming the established understanding of *Juniperus procera* affinity for cool highland environments (Sertse & Gailing, 2011). Steep slopes (<75 %) and median Normalized Difference Moisture Index (NDMI) ranges of –0.11 to 0.05 indicate a preference for moderately moist but aerated soils. Similar environmental constraints were documented in earlier studies (Abrha *et al*., 2018; Zenebe *et al*., 2024). These results reinforce the hypothesis that *Juniperus. procera* occupies a highly specialized climatic niche and is therefore vulnerable to warming, altered precipitation, and continued landscape disturbance. The spatial prediction map showed fragmented suitable habitat, with most of the forest providing suboptimal conditions, highlighting the need to protect high-suitability refugia.

Residual spatial autocorrelation was negligible, indicating that unmodeled spatial dependence is minimal and unlikely to bias the reported relationships. Because spatial autocorrelation can lead to underestimated standard errors and inflated confidence intervals (Gaspard & Kim, 2019), its absence strengthens confidence in the model results. The negligible residual autocorrelation observed here suggests that the selected predictors and sampling design successfully captured the dominant spatial patterns in species response, supporting the robustness and reliability of the ecological inferences.

International scientific relevance and implications for landscape-scale conservation

The integrated approach used in this study contributes substantially to global landscape ecology by demonstrating how classical floristic data, community analysis, and machine-learning modelling can be combined to reveal spatial biodiversity patterns in data-limited mountain ecosystems. Dry Afromontane forests are major components of the Eastern Afromontane biodiversity hotspot, yet remain underrepresented in global ecological datasets. By providing one of the most comprehensive spatial assessments of woody diversity, community composition, and species-specific abundance modelling in such systems, this study helps fill a major geographic gap.

The identification of strong spatial heterogeneity in diversity, distinct community clusters, and environmentally constrained niches has implications for advancing macroecological and conservation theory. The dominance of disturbance-adaptive shrubs supports global concerns regarding biotic homogenization and the erosion of specialist taxa in human-modified landscapes. The narrow niche of *Juniperus procera* adds to the growing

evidence that montane conifers worldwide are vulnerable to climate warming, moisture decline, and habitat fragmentation.

From a conservation perspective, the spatial prediction map of *Juniperus procera* highlights the fragmented nature of its suitable habitat, identifying high-elevation refugia that should be prioritized for protection and restoration. Integrating these predictions with community cluster information allows planners to target areas of high ecological value and high sensitivity. The methodological framework used here is transferable to other dryland and mountain forests globally, offering a scalable tool for biodiversity modelling in regions where field data are sparse.

In the broader context of global change, these findings underscore the importance of conserving dry Afromontane forests as climate-buffered refugia and reservoirs of endemic biodiversity. As land-use pressures intensify, the insights generated by this study will help inform national and regional forest management strategies, support the design of climate-resilient restoration actions, and contribute to international commitments to biodiversity conservation and ecosystem restoration. Future studies could build on this modelling framework by incorporating functional diversity metrics into Boosted Regression Tree analyses, thereby enhancing the ability to link species traits with environmental gradients.

CONCLUSION

This study demonstrates that the Desa'a Dry Afromontane Forest is structured by strong spatial heterogeneity, with 94 woody species forming sharply contrasting communities across steep environmental gradients. Species richness and abundance peak at mid-elevations and decline at higher altitudes, confirming a clear diversity–elevation signal within a fragmented montane landscape. Cluster analyses revealed seven distinct community types ranging from sparse, disturbance-dominated shrublands to more compositionally balanced mid-elevation assemblages, with indicator species and PERMANOVA confirming discrete and ecologically meaningful clusters, illustrating how environmental filtering and land-use pressures jointly shape landscape-scale community mosaics.

Boosted Regression Tree modelling identified a narrow climatic and topographic niche for *Juniperus procera*, with abundance tightly constrained by temperature, precipitation regimes, elevation, and slope. Residual spatial autocorrelation was negligible, reinforcing the robustness of these predictions. The close match between observed and predicted abundance patterns shows that machine-learning approaches can reliably detect nonlinear constraints on key montane species.

By integrating rigorous plot-based inventories, hierarchical clustering, diversity indices, indicator species analysis, and predictive niche modelling, this study provides a mechanistic understanding of how environmental gradients and disturbance processes generate spatial biodiversity patterns. The findings demonstrate that broad-scale climate sets potential niche space, while local habitat conditions largely determine realized community structure, with nonlinear abundance responses emerging under extreme conditions.

These findings advance landscape ecology by linking species dominance, community turnover, and environmental constraints within a unified spatial framework, enabling predictive assessments of biodiversity in data-limited mountain systems. From a conservation perspective, the study highlights fragmented suitable habitats for *Juniperus procera*, identifies high-elevation refugia, and offers a transferable methodological framework for prioritizing conservation and restoration in dryland mountain forests.

Internationally, the results highlight the vulnerability of montane forests to warming, moisture shifts, and habitat fragmentation, and they offer transferable tools for identifying refugia, mapping priority conservation zones, and designing climate-resilient restoration strategies in mountainous drylands.

ACKNOWLEDGMENTS

The first author sincerely thanks the WeForest project for providing the forest inventory data that formed the basis of this study and also extends gratitude to the UK Department for Environment, Food and Rural Affairs (DEFRA) and the Global Centre on Biodiversity for Climate (GCBC) for their support in data analysis and manuscript draft preparation. We acknowledge the Institute of International Education-Scholars Rescue Fund (IIE-SRF), and Nord University, Faculty of Bioscience and Aquaculture (FBA) for supporting the research stay of Emiru Birhane at Nord University.

CONFLICT OF INTEREST

The authors have declared that they have no conflicts of interest.

REFERENCES

Abrha, H., Birhane, E., Hagos, H., & Manaye, A. (2018). Predicting suitable habitats of endangered Juniperus procera tree under climate change in Northern Ethiopia. *Journal of Sustainable Forestry*, *37*(8), 842–853. https://doi.org/10.1080/10549811.2018.1494000

Agis, D., Tur, J., Moriña, D., Puig, P., & Fernández-Fontelo, A. (2024). good: An R package for modelling count data. *Methods in Ecology and Evolution*, *15*, 2192–2197. https://doi.org/10.1111/2041-210X.14387

Ahmed, S., Lemessa, D., & Seyum, A. (2022). Woody Species Composition , Plant Communities , and Environmental Determinants in Gennemar Dry Afromontane Forest , Southern Ethiopia. *Hindawi Scientific*. https://doi.org/https://doi.org/10.1155/2022/7970435 Research

Aynekulu, E., Aerts, R., Denich, M., Negussie, A., Friis, I., Demissew, S., & Boehmer, H.J. (2016). Plant diversity and regeneration in a disturbed isolated dry Afromontane forest in northern Ethiopia. *Folia Geobotanica*. https://doi.org/10.1007/s12224-016-9247-y

Aynekulu, E., Aerts, R., Moonen, P., Denich, M., Gebrehiwot, K., Vågen, T. G., Mekuria, W., & Boehmer, H. J. (2012). Altitudinal variation and conservation priorities of vegetation along the Great Rift Valley escarpment, northern Ethiopia. *Biodiversity and Conservation*, *21*(10), 2691–2707. https://doi.org/10.1007/s10531-012-0328-9

Birhane, E., Gidey, T., Abrha, H., Brhan, A., Zenebe, A., & Gebresamuel, G Noulèkoun, F. (2023). Impact of land-use and climate change on the population structure and distribution range of the rare and endangered Dracaena ombet and Dobera glabra in northern Ethiopia. *Journal for Nature Conservation*, *76*, 126506.

Birhanu, L., Moges, G., Amsalu, N., & Balzter, H. (2025). Woody species diversity, structure and community distribution along environmental gradients of Seqela Dry Afromontane forest in Northwestern Ethiopia. *PLoS ONE*, *20*(1), 1–21. https://doi.org/10.1371/journal.pone.0313020

Borcard, D., Gillet, F., & Legendre, P. (2018). *Numerical Ecology with R* (R. Gentleman, K.

Hornik, & G. G. Parmigiani (eds.)). Springer New York Dordrecht London Heidelberg. https://doi.org/10.1007/978-1-4419-7976-6

Cai, L., Kreft, H., Taylor, A., Denelle, P., Schrader, J., Essl, F., Kleunen, M. Van, Pergl, J., Py, P., Stein, A., Winter, M., & Barcelona, J. F. (2023). Global models and predictions of plant diversity based on advanced machine learning techniques. *New Phytologist*, *237*, 1432–1445. https://doi.org/10.1111/nph.18533

Chitayat, A.B., Lewis, M., Anyelwisye, M., Laizer, G. S., Stewart, F.A., Wich, S.A., & Piel, A.K. (2024). Development of spatial models and maps for tree species diversity and biomass in a miombo ecosystem, western Tanzania. *Applied Vegetation Science*, *27*(4). https://doi.org/10.1111/avsc.70002

Dagne, Y., & Birhanu, L. (2023). Floristic composition and plant community distribution along environmental gradients in Guard dry Afromontane forest of Northwestern Ethiopia. *BMC Ecology and Evolution*, *23:43*, 1–17. https://doi.org/10.1186/s12862-023-02154-6

Dedman, S., Officer, R., Clarke, M., Reid, D.G., & Brophy, D. (2017). Gbm . auto : A software tool to simplify spatial modelling and Marine Protected Area planning. *PLoS ONE*, *12*(12), 1–16.

Deressa, D., Egigu, M.C., & Sasikumar, J.M. (2023). Population Structure and Regeneration Status of Woody Plant Species in Tulu Korma Dry Afromontane Forest, West Shewa Zone, Oromia, Ethiopia. *Scientifica*. https://doi.org/10.1155/2023/9964663

Dessie, Y., Amsalu, N., Awoke, B., & Gebeyehu, G. (2025). Floral diversity , structural integrity , and regeneration patterns of Endba-Zend dry Afromontane forest in Northwestern Ethiopia. *BMC Ecology and Evolution*, *25:49*, 1 of 21.

Dixon, P. (2016). Computer program review VEGAN , a package of R functions for community ecology. *Journal of Vegetation Science*, *14*(6), 927–930.

Elith, J., & Leathwick, J. R. (2009). Species Distribution Models : Ecological Explanation and Prediction Across Space and Time. *Annu. Rev. Ecol. Evol. Syst*, *40*, 677–697. https://doi.org/10.1146/annurev.ecolsys.110308.120159

Elith, J., Leathwick, J.R., & Hastie, T. (2008). A working guide to boosted regression trees. *Journal of Animal Ecology*, *77*, 802–813. https://doi.org/10.1111/j.1365-2656.2008.01390.x

Enquist, B.J., Feng, X., Boyle, B., Maitner, B., Newman, E.A., Jørgensen, P.M., Roehrdanz, P.R., Thiers, B.M., Burger, J.R., Corlett, R.T., Couvreur, T.L.P., Dauby, G., Donoghue, J.C., Schildhauer, M., Šímová, I., Violle, C., Wieringa, J.J., & Wiser, S.K. (2019). The commonness of rarity : Global and future distribution of rarity across land plants. *SCIENCE ADVANCES*, *5*, 1–14.

Espíndola, R.P., Picanço, M.M., Andrade, L.P. De, Francisco, N., & Ebecken, F. (2025). Applications of Machine Learning Methods in Sustainable Forest Management. *Climate*, *13*(159), 1–16.

F. Dormann, C., M. McPherson, J., B. Araújo, M., Bivand, R., Bolliger, J., Carl, G., G. Davies, R., Hirzel, A., Jetz, W., Daniel Kissling, W., Kühn, I., Ohlemüller, R., R. Peres-Neto, P., Reineking, B., Schröder, B., M. Schurr, F., & Wilson, R. (2007). Methods to account for spatial autocorrelation in the analysis of species distributional data: A review. *Ecography*, *30*(5), 609–628. https://doi.org/10.1111/j.2007.0906-7590.05171.x

Galili, T. (2015). dendextend: an R package for visualizing, adjusting, and comparing trees of hierarchical clustering. *Bioinformatics*. https://doi.org/10.1093/bioinformatics/btv428

Gaspard, G., & Kim, D. (2019). Residual spatial autocorrelation in macroecological and

biogeographical modeling : a review. *Journal of Ecology and Environment*, *43:19*, 1–11.

Gebirehiwot, H. T., Kedanu, A. A., Guangul, A. A., & Adugna, M. T. (2023). Floristic Composition , Structure , And Regeneration Status Of Woody Plant Species In Hurubu Natural Forest , North Shewa , Oromia Region , Ethiopia. *Journal of Landscape Ecology (Czech Republic)*, *16*(1). https://doi.org/10.2478/jlecol-2023-0005

Haile, M., Birhane, E., Rannestad, M.M., & Adaramola, M.S. (2021). Expansive shrubs: Expansion factors and ecological impacts in northern Ethiopia. *Journal for Nature Conservation*, *61*(August 2019), 125996. https://doi.org/10.1016/j.jnc.2021.125996

Hill, L., Hector, A., Hemery, G., Smart, S., Tanadini, M., & Brown, N. (2017). Abundance distributions for tree species in Great Britain : A two- - stage approach to modeling abundance using species distribution modeling and random forest. *Ecology and Evolution*, *7*, 1043–1056. https://doi.org/10.1002/ece3.2661

Hishe, H., Giday, K., Van Orshoven, J., Muys, B., Taheri, F., Azadi, H., Feng, L., Zamani, O., Mirzaei, M., & Witlox, F. (2020). Analysis of Land Use Land Cover Dynamics and Driving Factors in Desa'a Forest in Northern Ethiopia. *Land Use Policy*, *101*, 105039. https://doi.org/10.1016/j.landusepol.2020.105039

Hordijk, I., Bialic-Murphy, L., Lauber, T., Routh, D., Poorter, L., Rivers, M. C., ter Steege, H., Liang, J., Reich, P. B., de-Miguel, S., Nabuurs, G. J., Gamarra, J. G. P., Chen, H. Y. H., Zhou, M., Wiser, S. K., Pretzsch, H., Paquette, A., Picard, N., Hérault, B., … Crowther, T. W. (2024). Dominance and rarity in tree communities across the globe: Patterns, predictors and threats. *Global Ecology and Biogeography*, *33*(10). https://doi.org/10.1111/geb.13889

Irving, K., Jähnig, S.C., Kuemmerlen, M., Irving, K., Jähnig, S.C., Identifying, M.K., & Irving, K. (2019). Identifying and applying an optimum set of environmental variables in species distribution models distribution models. *Inland Waters*, *0*(0), 1–18. https://doi.org/10.1080/20442041.2019.1653111

Karger, D.N., Kessler, M., Lehnert, M., & Jetz, W. (2021). worldwide. *Nature Ecology & Evolution*, *5*(June). https://doi.org/10.1038/s41559-021-01450-y

Kassun, B.W., Kallio, A.M.I., Tr, E., & Rannestad, M.M. (2024). Forest Ecology and Management Vegetation density and altitude determine the supply of dry Afromontane forest ecosystem services : Evidence from Ethiopia. *Forest Ecology and Management*, *552*. https://doi.org/10.1016/j.foreco.2023.121561

Keck, F., Peller, T., Alther, R., Barouillet, C., & Blackman, R. (2025). The global human impact on biodiversity. *Nature*, *641*(February). https://doi.org/10.1038/s41586-025-08752-2

Lai, J., Tang, J., Li, T., Zhang, A., & Mao, L. (2024). Plant Diversity Evaluating the relative importance of predictors in Generalized Additive Models using the gam . hp R package. *Plant Diversity*, *46*(4), 542–546. https://doi.org/10.1016/j.pld.2024.06.002

Magurran, A.E. (2004). Measuring Biological Diversity. *Journal of the Torrey Botanical Society*. https://doi.org/10.2307/4126959

Pebesma, E. (2018). Simple Features for R: Standardized Support for Spatial Vector Data. *The R Journal*, *10*(1), 439–446. https://doi.org/10.32614/RJ-2018-009

Reshad, M. (2019). Woody Species Richness and Diversity at Ades Dry Afromontane Forest of South Eastern Ethiopia. *American Journal of Agriculture and Forestry*, *7*(2), 44. https://doi.org/10.11648/j.ajaf.20190702.12

Samuel, D., Derero, A., & Hadgu, K.M. (2019). Tree species diversity and spatial distribution patterns on agricultural landscapes in sub-humid Oromia, Ethiopia. *Agroforestry Systems*, *93*(3), 1015–1029. https://doi.org/10.1007/s10457-018-0197-7

Sertse, D., & Gailing, O. (2011). Anthropogenic and natural causes influencing population genetic structure of Juniperus procera Hochst . ex Endl . in the Ethiopian highlands. *Genet Resour Crop Evol*, *58*, 849–859. https://doi.org/10.1007/s10722-010-9623-z

Sherafu, Z., Egigu, M.C., & Sasikumar, J.M. (2024). Floristic Composition, Diversity, and Regeneration of Woody Plant Species of Dabal Forest in Southeastern Ethiopia: Implication for Conservation. *Scientifica*, 1–8. https://doi.org/10.1155/2024/7414375

Siyum, Z.G., Ayoade, J.O., Onilude, M.A., & Feyissa, M.T. (2019). Growth trajectories and ages of main tree species in dry Afromontane forest fragments of northern Ethiopia. *Forest Ecosystems*, *6*(15). https://doi.org/https://doi.org/10.1186/s40663-019-0178-y

Tefera, Y., Leulekal, E., & Warkineh, B. (2024). Woody Plant Diversity of Wof Ayzurish Forests, the Dry Afromontane of the North Shewa of Ethiopia: Regeneration and community type with Environmental gradient. *Trees, Forests and People*, *16*(May), 100562. https://doi.org/10.1016/j.tfp.2024.100562

Tekle, T., & Maryo, M. (2022). Ecological Assessment Of Woody Plant Diversity And The Associated Threats In Afromontane Forest Of Ambericho , Southern Ethiopia. *AJournal of Landscape Ecology*, *15*(2).

Tilahun, M., Angassa, A., Bora, Z., Mengistu, S., & Wu, J. (2025). Phenotypic plasticity and genetic variation of Prosopis juliflora (Sw.) DC. across diverse rangelands in northeastern Ethiopia. *Ecological Processes*, *14*(1). https://doi.org/10.1186/s13717-024-00575-9

Wassie, K.B., Walle, G.C., & Alemnew, A.T. (2024). Woody species composition, structure and regeneration status of Alka forest Beyeda District, North Gondar Zone, Amhara Region, Northern Ethiopia. *BMC Plant Biology*, *24*(1). https://doi.org/10.1186/s12870-024-05822-x

Wickham, H. (2010). A Layered Grammar of Graphics. *Journal of Computational and Graphical Statistics*, *19*(1), 3–28. https://doi.org/10.1198/jcgs.2009.07098

Woldearegay, M., & Woldu, Z. (2020). Vegetation structure and regeneration status of the moist, evergreen, afromontane Bore–Anferara–Wadera forest in southern Ethiopia. *Journal of Forestry Research*, *31*(3), 953–967. https://doi.org/10.1007/s11676-018-00872-y

Yinebeb, M., Lulekal, E., & Bekele, T. (2023). Ecological determinants in plant community structure across dry afromontane forest patches of Northwestern Ethiopia. *BMC Ecology and Evolution*, *23:68*, 1–13. https://doi.org/10.1186/s12862-023-02176-0

Young, N.E., Stohlgren, T.J., Evangelista, P.H., Graham, J., & Newman, G. (2011). Regional data refine local predictions: modeling the distribution of plant species abundance on a portion of the central plains. *Environ Monit Assess*, 1–13. https://doi.org/10.1007/s10661-011-2351-9

Zenebe, G., Zenebe, A., Birhane, E., Girma, A., & Shiferaw, H. (2024). Revealing suitable habitats for Juniperus procera and Olea europaea tree species in the remnant dry Afromontane forests of Ethiopia: Insights from ensemble species distribution modeling approach. *Ecology and Evolution*, *14*(10). https://doi.org/10.1002/ece3.70343

APPENDIX

Appendix 1: Input covariate variables used during the BRT based spatial abundance modeling

Short name	Full name	Unit
BIO1	Annual Mean Temperature	°C
BIO2	Mean Diurnal Range (Mean of monthly (max temp - min temp))	°C
BIO3	Isothermality (BIO2/BIO7) (×100)	%
BIO4	Temperature Seasonality (standard deviation ×100)	°C
BIO5	Max Temperature of Warmest Month	°C
BIO6	Min Temperature of Coldest Month	°C
BIO7	Temperature Annual Range (BIO5-BIO6)	°C
BIO8	Mean Temperature of Wettest Quarter	°C
BIO9	Mean Temperature of Driest Quarter	°C
BIO10	Mean Temperature of Warmest Quarter	°C
BIO11	Mean Temperature of Coldest Quarter	°C
BIO12	Annual Precipitation	mm
BIO13	Precipitation of the Wettest Month	mm
BIO14	Precipitation of the Driest Month	mm
BIO15	Precipitation Seasonality (Coefficient of Variation)	mm
BIO16	Precipitation of the Wettest Quarter	mm
BIO17	Precipitation of the Driest Quarter	mm
BIO18	Precipitation of Warmest Quarter	mm
BIO19	Precipitation of the Coldest Quarter	mm
Altitude	SRTM DEM 30 m resolution	Meter
Slope	Slope derived from DEM	% rise
Aspect	Aspect derived from DEM	--
Dist_road	Distance to road access	Meter
Dist_settlements	Distance to settlements	Meter
Dist_streams	Distance to streams	Meter
Median_NDMI	Median Normalized Difference Moisture index derived from Landsat imagery	--
Dist_water_source	Distance to water sources	Meter

Appendix 2: GAM Predictor Significance Across Biodiversity Metrics

Variable	Richness (edf, p)	Shannon (edf, p)	Simpson (edf, p)	Evenness (edf, p)	Abundance (edf, p)
Bio_1	1.00, 0.497	1.00, 0.435	1.00, 0.113	**1.00, 0.014**	1.00, 0.585
Bio_4	1.00, 0.467	1.00, 0.523	1.00, 0.120	**1.29, 0.006**	1.00, 0.408
Bio_5	1.15, 0.939	1.00, 0.479	1.00, 0.200	1.00, 0.106	1.00, 0.706
Bio_8	1.00, 0.462	1.00, 0.061	**1.00, 0.027**	**1.00, 0.035**	1.00, 0.805
Bio_11	1.00, 0.451	1.00, 0.487	1.00, 0.120	**1.00, 0.016**	1.00, 0.567
Bio_12	1.00, 0.476	1.00, 0.580	1.62, 0.336	**2.97, 0.012**	2.39, 0.464
Bio_14	2.75, 0.126	2.54, 0.466	3.77, 0.280	1.06, 0.314	**3.72, 0.011**
Bio_15	1.00, 0.264	5.71, 0.143	5.59, 0.105	**1.52, 0.046**	1.00, 0.962
Bio_16	2.67, 0.367	1.00, 0.501	1.00, 0.262	**1.00, 0.032**	**4.59, 0.048**
Altitude	2.57, 0.812	3.86, 0.316	3.27, 0.529	1.00, 0.439	1.00, 0.201
Slope	1.00, 0.618	1.04, 0.436	1.87, 0.284	1.85, 0.420	1.00, 0.746
Median_NDMI	**1.00, 0.045**	1.00, 0.234	1.00, 0.325	4.27, 0.349	**1.00, 0.001**
Water_source	1.00, 0.127	**1.00, 0.008**	**1.00, 0.001**	5.28, 0.086	1.00, 0.846
Deviance explained	0.275	0.310	0.343	0.299	0.375

Significant predictors ($p < 0.05$) from GAM models for biodiversity metrics are shown with edf and p-values. The bottom row reports model deviance explained for each metric.

Appendix 3: Partial dependence plot showing the effect of predictor variables