Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Ecological niche modeling reveals severe historical range contraction and contemporary refugia for the narrow endemic Arctium pallidivirens in Central Asia

Khikmatullo Ibragimov
Academy of Sciences Republic of Uzbekistan
Khabibullo Shomurodov
Academy of Sciences Republic of Uzbekistan
Bekhruz Khabibullaev
Academy of Sciences Republic of Uzbekistan
Natalya Beshko
Academy of Sciences Republic of Uzbekistan
Khusniddin Abulfayzov
Academy of Sciences Republic of Uzbekistan
Mukhlisa Turopova
Termez State Pedagogical Institute
G`ayrat Sharifov
State Environmental Research Institution “Nuratau State Reserve”
Azamat Yuldoshkhonov
yoldoshxonovazamat9@gmail.com

National University of Uzbekistan

Research Article

Keywords: Arctium pallidivirens, ecological niche modeling, narrow-ranged endemic species, Nuratau Mountain Range, habitat suitability assessment

Posted Date: July 28th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-10490162/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Ecological niche modeling reveals severe historical range contraction and contemporary refugia for the narrow endemic *Arctium pallidivirens* in Central Asia.

Khikmatullo Ibragimov [1], Khabibullo Shomurodov [1], Bekhruz Khabibullaev [1], Natalya Beshko [1], Khusniddin Abulfayzov [1],Mukhlisa Turopova [2], G`ayrat Sharifov [3], Yuldoshkhonov Azamat[4]

1. Laboratory of Conservation and Ecology of Plants, Institute of Botany Academy of Sciences of Republic Uzbekistan, Tashkent 100125, Uzbekistan;
2. Department of Chemistry and Biology Termez State Pedagogical Institute, Termiz, Uzbekistan.
3. State Environmental Research Institution "Nuratau State Reserve".Uzbekistan.
4. National Unversity of Uzbekistan named after Mirzo Ulugbek, Tashkent, Uzbekistan.

Corresponding : yoldoshxonovazamat9@gmail.com and hikmatullai626@gmail.com

Abstract:

Narrow-ranged endemic plant species are highly vulnerable to environmental change, yet their conservation remains poorly understood in Central Asian mountain ecosystems. We assessed the conservation vulnerability of *Arctium pallidivirens,* a rare endemic of the Nuratau Mountains (Uzbekistan), using ensemble species distribution modelling (BIOMOD2). Field surveys and historical herbarium records were integrated with bioclimatic, topographic, and edaphic variables to identify the environmental drivers of habitat suitability. Wet-season precipitation, soil properties, and vegetation productivity were the strongest predictors of species occurrence. The ensemble model identified only 110.16 km² (0.64% of the study area) as suitable habitat. Comparison with the observed distribution polygon (792.95 km²), derived from historical herbarium records and contemporary field observations, revealed a pronounced spatial mismatch, with only 84.81 km² (0.49%) of overlap. This pattern suggests substantial historical range contraction associated with long-term environmental and anthropogenic change. The remaining suitable habitats are concentrated in climatically buffered refugia, indicating that the current distribution reflects historical range contraction rather than an inherently narrow climatic niche. Our findings provide a robust spatial framework for identifying conservation priorities and supporting long-term monitoring of this narrow endemic species.

.**Keywords:** *Arctium pallidivirens*, ecological niche modeling, narrow-ranged endemic species, Nuratau Mountain Range, habitat suitability assessment.

1. Introduction

The conservation of narrow-ranged and ecologically specialized plant species has become one of the major challenges in contemporary biodiversity research. Such species occupy restricted ecological niches and depend strongly on specific climatic, edaphic, and microhabitat conditions, making them highly vulnerable to environmental change. Consequently, identifying the factors governing their spatial distribution is essential for understanding species ecology and developing effective conservation strategies (Kruckeberg & Rabinowitz, 1985; Franklin, 2013). Mountain ecosystems are characterized by pronounced environmental heterogeneity, where temperature, moisture, and soil properties vary considerably over short distances. These steep environmental gradients strongly influence plant distribution at the microhabitat scale, resulting in fragmented populations that often persist in ecologically stable but spatially restricted refugia. Recent studies have shown that climatic variability, particularly in mountain regions, can substantially alter habitat suitability, emphasizing the need to better understand species–environment relationships across multiple spatial scales (Hannah et al., 2002; Pearson & Dawson, 2003; Orsenigo et al., 2022; Fenu et al., 2022, 2023, 2024).

Despite their exceptional floristic diversity and high levels of endemism, the mountain ecosystems of Central Asia remain insufficiently studied with respect to quantitative species–environment relationships. In particular, the combined influence of climatic, edaphic, and vegetation-related factors on the distribution of endemic plant species has received limited attention. Previous studies have identified moisture availability, vegetation productivity, and thermal conditions as important determinants of mountain plant distributions; however, quantitative species-level assessments integrating these environmental drivers remain scarce (Orsenigo et al., 2025; Doniyorova et al., 2025). Consequently, robust spatial assessments are needed to identify the environmental drivers and conservation priorities of narrowly distributed endemic species in this region. Among these poorly studied endemic species, Arctium pallidivirens (Asteraceae) is narrowly restricted to the Nuratau Mountains of Uzbekistan. Owing to its limited distribution and strong ecological specialization, the species is considered particularly vulnerable to environmental change. However, its ecological niche, habitat suitability, and environmental drivers have not yet been quantitatively assessed. This makes A. pallidivirens an appropriate model species for investigating species–environment relationships, habitat suitability, and potential range dynamics in a Central Asian mountain ecosystem. Ecological Niche Modeling (ENM) and Species Distribution Modeling (SDM) have become widely used tools for predicting habitat suitability, identifying environmental drivers of species distributions, and supporting conservation planning (Franklin, 2013; Song et al., 2024; Estopinan et al., 2024). By integrating species occurrence records with climatic, topographic, and edaphic variables, these approaches provide a robust framework for evaluating habitat suitability and ecological constraints. Among them, ensemble modelling implemented in BIOMOD2 combines multiple algorithms to reduce model-specific uncertainty and improve the reliability of predictions (Thuiller et al., 2009; Hao et al., 2019). The objectives of this study were to: (i) model the potential distribution of Arctium pallidivirens using an ensemble species distribution modelling (SDM) approach; (ii) identify the relative importance of climatic, vegetation, and edaphic variables shaping its ecological niche; and (iii) evaluate changes in its distribution by integrating model projections with historical herbarium records (TASH) and contemporary field surveys. We hypothesized that the present distribution of A. pallidivirens represents only a remnant of its historical range and that its persistence is primarily associated with climatically buffered microhabitats and favourable edaphic conditions. By integrating historical occurrence records with contemporary ensemble modelling, this study evaluates evidence of historical range dynamics while identifying environmentally suitable habitats for a narrow endemic species. The findings provide a scientific basis for spatially explicit conservation planning and long-term monitoring in the Nuratau Mountains.

2. Materials and Methods

2.1. Study Area

The Nuratau Mountain Range, which serves as the core focus of our study, is one of the pivotal physical-geographic regions of Central Asia, distinguished by its rich biodiversity, unique landscape architecture, and complex ecological conditions **(Figure 1).** Ranging in elevation from 300 m to 2169 m above sea level, this mountain range extends approximately 250 km along the right bank of the Zarafshan River and encompasses a mountainous topography of nearly 14,000 km 2 . Orographically, the Nuratau Mountains belong to the northwestern spurs of the Turkestan Range—one of the major branches of the Pamir-Alay system—and consist of two parallel, mid-altitude mountain ranges: the Nuratau in the north and the Aktau in the south. These ranges are separated by the Nuratau Valley and a series of low-mountain ridges. Situated within the transition zone between the Pamir-Alay and Tien Shan mountain systems and the Turan Lowland, this range plays a critical role in shaping regional climate and flora dynamics. While the central portion of

the Nuratau Mountains is highly dissected by deep gorges and rocky canyons, the elevation gradually decreases westward, transitioning into the sandy plains of the Kyzylkum Desert. The mountain slopes are incised by numerous mountain streams and dry riverbeds, with the northern slopes exhibiting a significantly steeper and more precipitous topography compared to the southern slopes. Phytogeographically, this area constitutes an essential component of the Irano-Turanian floristic region and comprises three primary subregions: the Nuratau, the Aktau, and the Nuratau relict mountains (Beshko, 2020). Such intricate relief and climatic conditions have facilitated the evolution and persistence of numerous endemic and relict plant species within the Nuratau Mountains. Consequently, the Nuratau Mountain Range stands as a major center of Central Asian endemic flora, serving as a natural refugium for narrow endemic species such as Arctium pallidivirens, and holds profound ecological significance for biodiversity conservation in the region. To fully encompass these ecological gradients and capture the complete historical distribution of the species, the boundaries of the broader Nuratau botanico-geographical district (okrug) were utilized as the spatial domain for modeling. This defined study area incorporates both the core montane ridges and the adjacent environmental matrix, encompassing a total analyzed landscape area of exactly 17,342.20 km2.

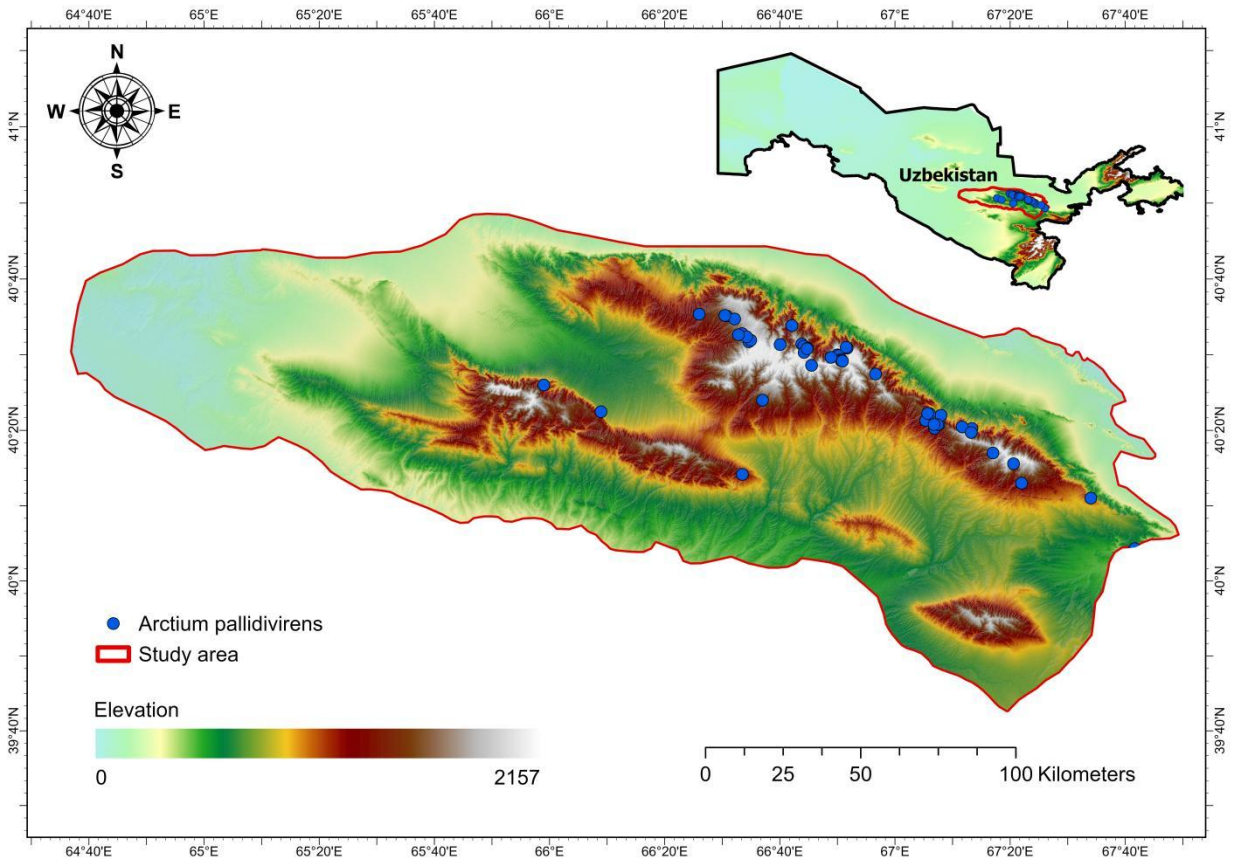


Figure 1. Study area.

2.2. Study Object

Arctium pallidivirens (Kult.) Tscherneva is a perennial herbaceous plant belonging to the Asteraceae family and is a rare species distributed in Central Asia, specifically within the Nuratau Mountain Range. This species primarily grows in the lower and middle mountain zones at elevations ranging from 700 m to 1800 m above sea level, inhabiting rocky and gravelly slopes as well as fine-textured soils. It is adapted to arid and semi-arid climatic conditions and is characterized by its high sensitivity to seasonal moisture availability and edaphic properties

(Beshko, 2020). The species typically occurs as solitary individuals or in small populations, occasionally forming small clusters. In Uzbekistan, approximately 50 coenopopulyations are known, with the total population size estimated to be around 10,000–30,000 individuals. Its Extent of Occurrence (EOO) is calculated at 3,467.2 km 2 , while its Area of Occupancy (AOO) spans 136 km 2 . According to the IUCN Red List categories and criteria, this species is classified as Vulnerable (VU) under the criteria B2ab(ii,iii,iv,v); C2a(i) (Beshko, 2020). The primary threat to the species is associated with livestock grazing pressure, which limits seed reproduction processes and negatively impacts population recruitment. Consequently, the stability of the species' populations is heavily dependent on both ecological and anthropogenic factors (Volis & Beshko, 2023).



Figure 2. Natural appearance of *Arctium pallidivirens* in its habitat. **Photos:** N.Yu. Beshko.

(a) Seedlings and individuals in the early stage of development in natural conditions. **(b)** Individual in the vegetative rosette stage, with leaf formation and color variations. **(c)** General view of the natural population and landscape characteristics of the growth site

(d) Individual in the flowering stage, with a developed generative stem and inflorescence

2.3. Species Occurrence Data.

Species occurrence (presence) data for *Arctium pallidivirens* were compiled from multiple sources. Out of a total of 52 presence records, 17 were obtained through direct field observations, while the remaining 35 were collected from historical herbarium specimens, including the National Herbarium of Uzbekistan (TASH) and the Plantarium database. The herbarium records were carefully verified for taxonomic reliability and spatial accuracy, and records with low coordinate precision or high spatial uncertainty were excluded from the analysis. To minimize spatial autocorrelation and ensure the independence of occurrence records, spatial thinning was applied using a minimum distance threshold of 1 km. Consequently, the initial dataset was reduced from

52 to 34 spatially independent occurrence records, which were retained for all subsequent modelling analyses. Although spatial thinning reduced the number of occurrence records to 34, this sample size is considered sufficient for modelling narrow-ranged endemic species using ensemble SDM approaches. Previous studies have demonstrated that ensemble modelling frameworks incorporating MaxEnt and Generalized Linear Models (GLM) maintain high predictive performance even with relatively small sample sizes (15–30 occurrence records), provided that spatial autocorrelation is minimized through spatial thinning (Wisz et al., 2008; van Proosdij et al., 2016). Pseudo-absence points were generated within the study area using a random sampling method, whereby grid cells containing presence records were explicitly excluded. Detailed information on the pseudo-absence generation strategy and model calibration procedures is provided in the ecological niche modelling section. To evaluate spatial discrepancies and delineate the observed Extent of Occurrence (EOO) shown in Figure 10, a 100% Minimum Convex Polygon (MCP) was constructed in R using the st_convex_hull() function from the sf package. All data preprocessing and statistical analyses were performed in the R statistical environment.

2.4. Remote Sensing and Environmental Variable Generation.

Environmental variables representing topographic, climatic, vegetation, radiation, land cover, and edaphic conditions were compiled to model the potential distribution of the Arctium species. The data were acquired from multiple global remote sensing and gridded databases. Selected variables were partially retrieved and preprocessed via the Google Earth Engine (GEE) platform, while the remaining datasets were downloaded from open-source repositories and processed locally. All predictors were standardized to a common spatial resolution of 1 km (Gorelick et al., 2017). Topographic variables were derived from the Shuttle Radar Topography Mission (SRTM, 30 m) Digital Elevation Model (DEM), from which slope, aspect, and the Topographic Position Index (TPI) were calculated (Farr et al., 2007). Vegetation status was assessed using MODIS MOD13A2 NDVI products (1 km, 16-day composites) to calculate the mean, minimum, and maximum NDVI values (Didan, 2015). Radiation variables were obtained from NASA GLDAS data to compute the mean annual solar radiation (radiation_mean) (Rodell et al., 2004). Land cover was evaluated based on the MODIS land cover product. To indirectly represent anthropogenic pressure, nighttime light intensity (nightlights_mean) was incorporated using VIIRS data, serving as a proxy indicator for human impact. Soil properties, including clay, sand, and silt fractions, pH, nitrogen content, and organic carbon, were retrieved from the SoilGrids database (250 m) and resampled to a 1 km spatial resolution (Hengl et al., 2017). Additionally, to represent long-term climatic conditions, 19 bioclimatic variables (BIO1–BIO19) were obtained from the WorldClim v2.1 database (1 km, 1970–2000) (Fick & Hijmans, 2017). These bioclimatic variables define the fundamental climatic niche of the species and reflect the primary climatic constraints in the modeling process. To mitigate multicollinearity, Pearson correlation analysis ($|r| \geq 0.7$) and Variance Inflation Factor (VIF ≤ 5) filtering methods were applied, resulting in a set of relatively independent variables (Dormann et al., 2013; Feng et al., 2019). From this pool, the final predictors were selected based on their ecological relevance, ease of interpretation, and stability across models. The final predictor set comprised topographic conditions (aspect, slope, and TPI), climatic variables (selected BIOCLIM variables), vegetation productivity (NDVI metrics), soil texture and chemical properties (clay, sand, silt, pH, nitrogen, and organic carbon), radiation (radiation_mean), land cover, and an anthropogenic impact proxy (nightlights_mean).

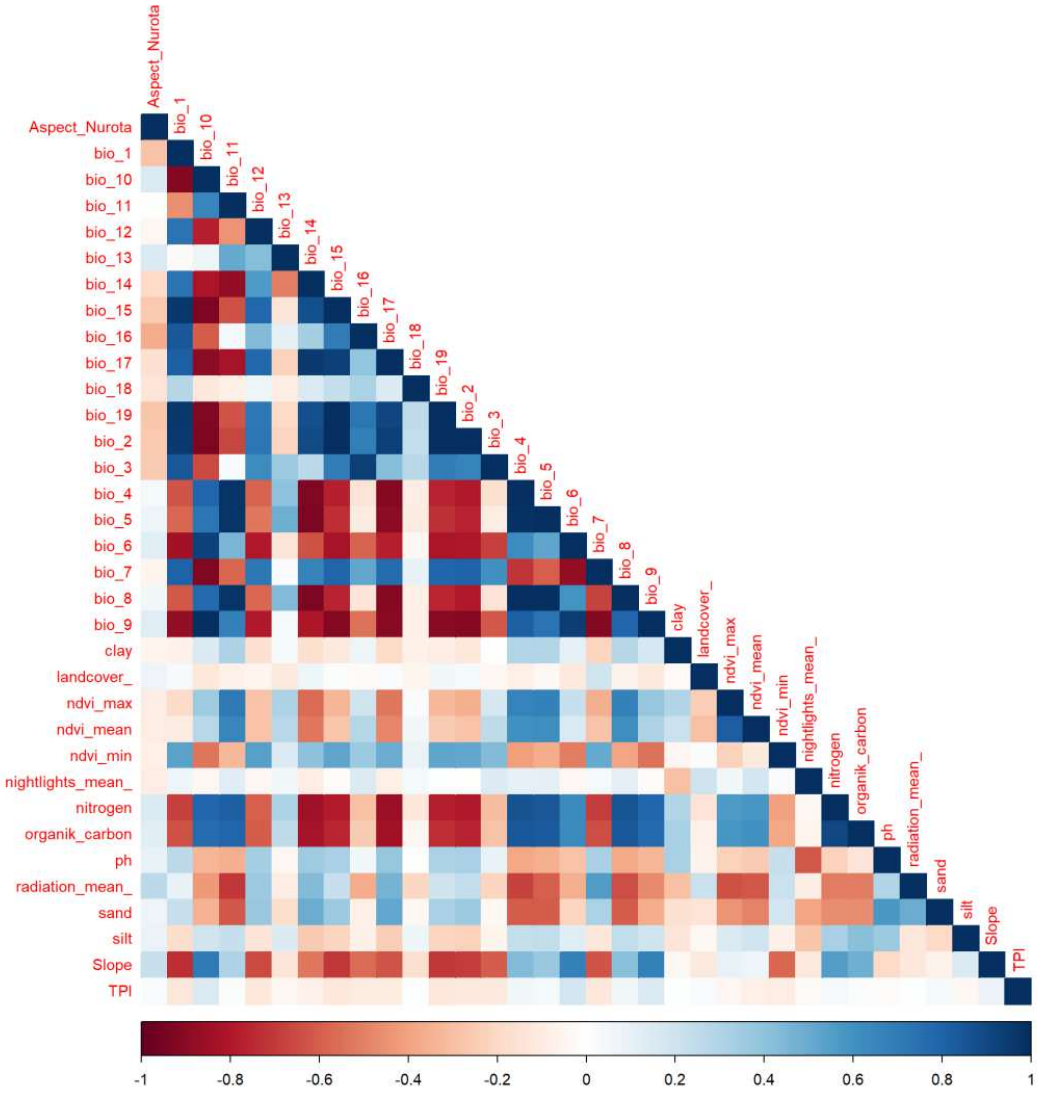


Figure 3. Correlation matrix of environmental predictors.

2.5. Ecological Niche Modeling

2.5.1. Model Construction and Calibration

In this study, the potential distribution of Arctium pallidivirens was assessed using an Ecological Niche Modeling (ENM/SDM) approach. The modeling workflow was implemented within the BIOMOD2 framework in R (Thuiller et al., 2009). Considering model stability and performance efficiency with small sample sizes, four algorithms were selected: Generalized Linear Models (GLM), Generalized Additive Models (GAM), Random Forest (RF), and MaxEnt (via the MAXNET implementation). Since these algorithms encompass diverse statistical and machine learning approaches, they enable a reduction in algorithmic uncertainty and provide a more robust identification of species–environment relationships. To minimize sampling bias and reduce spatial autocorrelation, pseudo-absence points were generated using a randomized background sampling strategy restricted within a spatial buffer surrounding the target study area. To maintain high calibration sensitivity without over-inflating performance metrics, a 1:10 presence-to-pseudo-absence ratio was employed (Barbet-Massin et al., 2012), resulting in 340 pseudo-absence points per iteration. To prevent artifactual bias from a single random selection, pseudo-absences were

generated across 10 independent random sampling iterations. For each iteration, pseudo-absence points were explicitly constrained to grid cells located outside the direct presence buffers. Following multicollinearity and ecological relevance screening, a set of relatively independent and biologically meaningful variables was selected for the final modeling process. Model evaluation was performed using a repeated random-split cross-validation technique, wherein 80% of the data was utilized for model calibration and the remaining 20% was reserved for validation, with this procedure repeated over multiple iterations. Model performance was evaluated strictly based on the validation data using the Area Under the Curve (AUC), True Skill Statistic (TSS), and Cohen's Kappa metrics to assess the generalization capability of the models. Final predictions were generated by combining the outputs across different algorithms and cross-validation runs. The ensemble mean (E mean) approach reduced discrepancies among individual models and enhanced the stability of the final habitat suitability maps. The final ensemble prediction was derived based on continuous suitability scores, and poorly performing model outputs (e.g., TSS <0.4) were explicitly excluded. Furthermore, spatial uncertainty was evaluated based on the variance across models and cross-validation runs, which was visualized as a standalone map. This approach allows for a cautious interpretation of the model projections and highlights geographic zones with high predictive reliability. Overall, the model outputs were interpreted as an index of relative habitat suitability rather than an absolute probability of species occurrence.

2.5.2. Spatial Comparison of Observed and Predicted Distributions

The observed distribution of *Arctium pallidivirens* was delineated based on georeferenced occurrence points collected during field surveys and verified herbarium records. These data represent all currently documented occurrence localities of the species within the study area. To define the spatial boundaries of the observed distribution, a Minimum Convex Polygon (convex hull) encompassing all validated presence points was constructed. This polygon represents the empirically observed range of the species but does not fully capture its realized ecological niche. The predicted potential distribution was evaluated based on the habitat suitability map generated through ensemble modeling. For visualization and spatial analysis, the continuous ensemble projection was converted into a binary (suitable/unsuitable) map using an optimal threshold value (cutoff = 0.445) determined by the maximum True Skill Statistic (TSS / Youden's index). A spatial overlay analysis was then performed between the observed distribution polygon and the binary ensemble suitability map. The analysis revealed a substantial spatial congruence between the predicted suitable habitats and the actual occurrence points. The suitable areas closely aligned with the localities of the observed populations, thereby validating the ecological robustness of the model. Concurrently, the absence of the species in certain predicted suitable areas may be attributed to potential dispersal limitations or unmeasured microhabitat constraints. This approach is descriptive in nature, focusing on quantifying the spatial congruence between the observed and predicted distributions. Rather than establishing causal relationships, this analysis aims to evaluate the discrepancies between the potential and realized geographic ranges of the species.

3.Results

3.1. Model Performance

Model performance was evaluated using a repeated random-split cross-validation technique, wherein 80% of the data was utilized for model calibration and 20% was reserved for validation, with this procedure executed over multiple iterations. Performance metrics were calculated strictly based on the validation datasets to determine the generalization capability of the models. The AUC values across the four modeling algorithms demonstrated high discriminative power. Specifically, the mean AUC score was 0.95±0.04, indicating that all models exhibited a high capacity to differentiate between presence and absence localities of the species. Conversely, threshold-

dependent metrics exhibited relatively lower to moderate evaluation scores. The TSS values averaged around 0.72±0.10, indicating a moderate-to-high capability of the models to distinguish between suitable and unsuitable habitats. Similarly, the Cohen's Kappa values reached 0.59±0.12, demonstrating that the model predictions were significantly better than random chance. The evaluation results indicate that while the discriminative indices (AUC) remained consistently high, a certain degree of variability persisted within the threshold-dependent metrics. This variation can be attributed to uncertainties associated with pseudo-absence data generation and threshold selection. The relatively high performance metrics of the models may be partially driven by the pseudo-absence selection strategy and the structure of the environmental predictors. Therefore, these results should be interpreted with caution. Overall, the model projections were deemed stable and reliable; consequently, subsequent analyses focused on evaluating the spatial congruence derived from the ensemble approach rather than cross-comparing individual models.

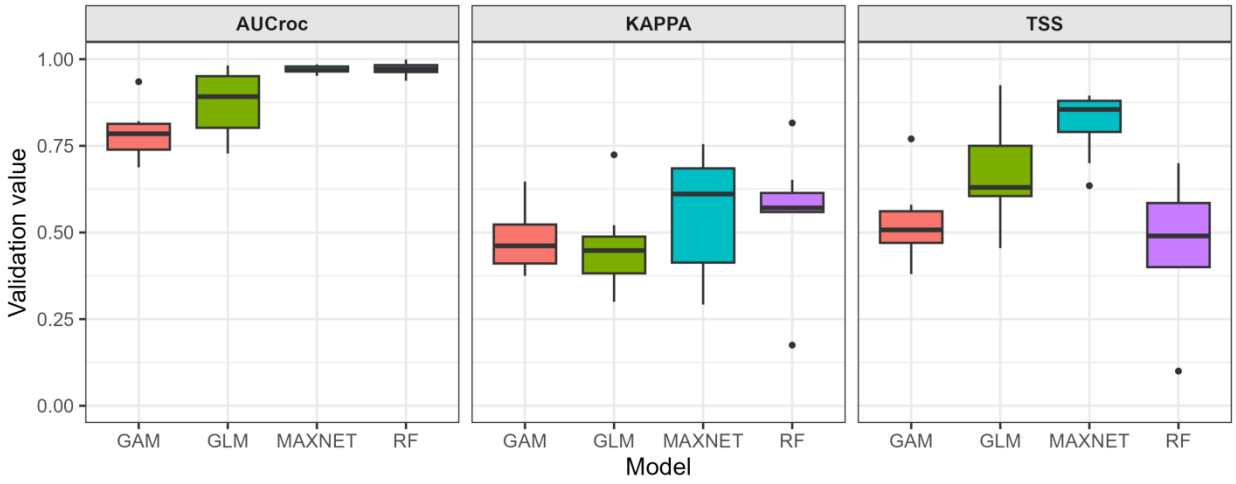


Figure 4. Performance efficiency (AUC, TSS, Kappa) of individual ecological niche models evaluated based on cross-validation.

3.2. Ecological Drivers of *Arctium pallidivirens* Distribution and Their Effects

The relative contribution of environmental variables shaping the spatial distribution of Arctium pallidivirens was evaluated using a permutation-based variable importance analysis. The importance scores of the variables were calculated across the different modeling algorithms and expressed as percentages based on the generalized ensemble outputs (Figure 5). This approach enabled a direct comparison of each environmental factor's contribution to the final model predictions. According to the results, the highest relative importance was exhibited by bio_16 (precipitation of the wettest quarter, 28.8%), indicating that the seasonal moisture regime is a decisive factor governing the species' distribution. This was followed closely by soil-related variables, specifically sand (27.2%), clay (23.3%), and silt (23.1%) fractions, which underscores the species' strong dependency on substrate texture and soil physical properties. Vegetation-related metrics also played a significant role; in particular, the mean NDVI (ndvi_mean, 21.2%) demonstrates that species occurrence is intimately linked to vegetation cover density and productivity. Among other climatic factors, bio_12 (annual precipitation, 19.6%) and bio_13 (precipitation of the wettest month, 11.7%) held prominent importance, confirming the species' high sensitivity to overall moisture availability. Topographic factors exerted a noticeable influence as well: slope steepness (slope, 19.2%), solar radiation (radiation_mean, 16.3%), and aspect (aspect, 8.3%) play essential roles in the species' microhabitat selection. Furthermore, the relatively high importance of nighttime light intensity (nightlights_mean, 17.3%) may indicate an indirect relationship with anthropogenic pressures or localized landscape modifications. Other variables, including the minimum NDVI (11%), land cover (landcover, 10.2%), bio_18 (9.1%),

and the Topographic Position Index (TPI, 7.3%), yielded lower yet still noticeable contributions, indicating that these factors play secondary but context-specific roles in shaping the species' distribution. Overall, these findings demonstrate that the distribution of *Arctium pallidivirens* is primarily driven by moisture regimes, soil texture, and vegetation dynamics, whereas topographic features and anthropogenic factors act as confounding drivers that modulate these primary environmental gradients.

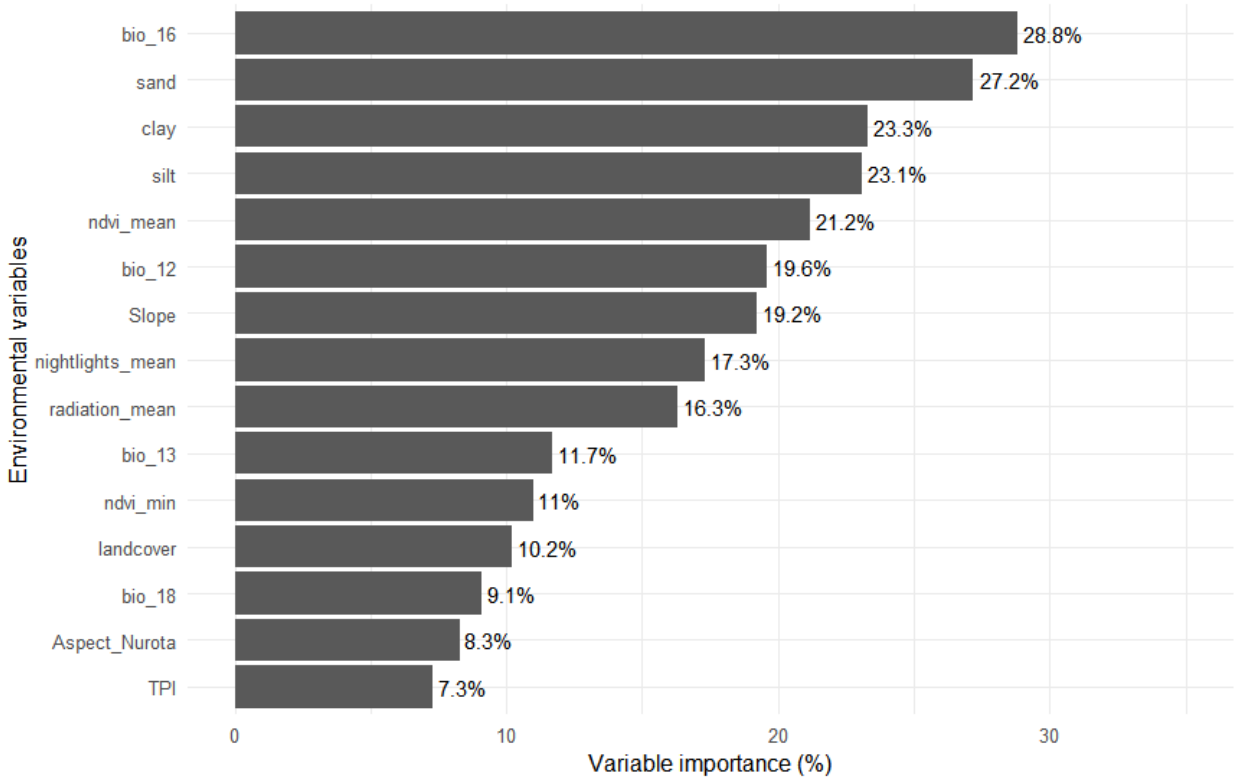


Figure 5. Relative importance (in %) of environmental variables in the *Arctium pallidivirens* distribution model.

The ecological response curves derived from the GLM and MAXNET models for *Arctium pallidivirens* allowed for the determination of functional relationships between primary environmental gradients and habitat suitability. The response curves revealed consistent ecological relationships across both models, with the model outputs exhibiting biologically meaningful response shapes for the key variables **(Figure 6).** For the most influential variable, bio_16 (precipitation of the wettest quarter), a distinct negative relationship was observed in both models: habitat suitability dropped sharply at higher precipitation values, whereas relatively high suitability was recorded at lower values. This finding indicates that the species prefers relatively drier conditions with seasonal moisture availability rather than excessive precipitation regimes. Soil texture variables also displayed clear threshold characteristics. Specifically, for the sand fraction, suitability peaked within a specific intermediate range and decreased abruptly below or above these values. Similarly, the responses to clay and silt fractions demonstrated that the species occurs only within specific soil texture combinations, thereby confirming its substrate selectivity. Regarding the vegetation metric ndvi_mean, a clear positive non-linear relationship was captured. As the NDVI value increased, habitat suitability rose, reaching a maximum within an optimal range before exhibiting a slight decline. This suggests that the species is associated with moderately developed vegetation cover but avoids excessively dense canopy conditions. For another climatic factor, bio_12 (annual precipitation), the model response showed a relatively low

amplitude, signifying that the species is less sensitive to total annual rainfall, while seasonal precipitation (bio_16) plays the predominant role. Overall, the directions of the response curves projected by the GLM and MAXNET models are highly congruent, confirming the reliability of the identified ecological relationships. These results indicate that the ecological niche of *Arctium pallidivirens* is strictly defined by seasonal moisture dynamics, soil texture, and vegetation status, with these factors acting as the primary environmental gradients constraining its spatial distribution.

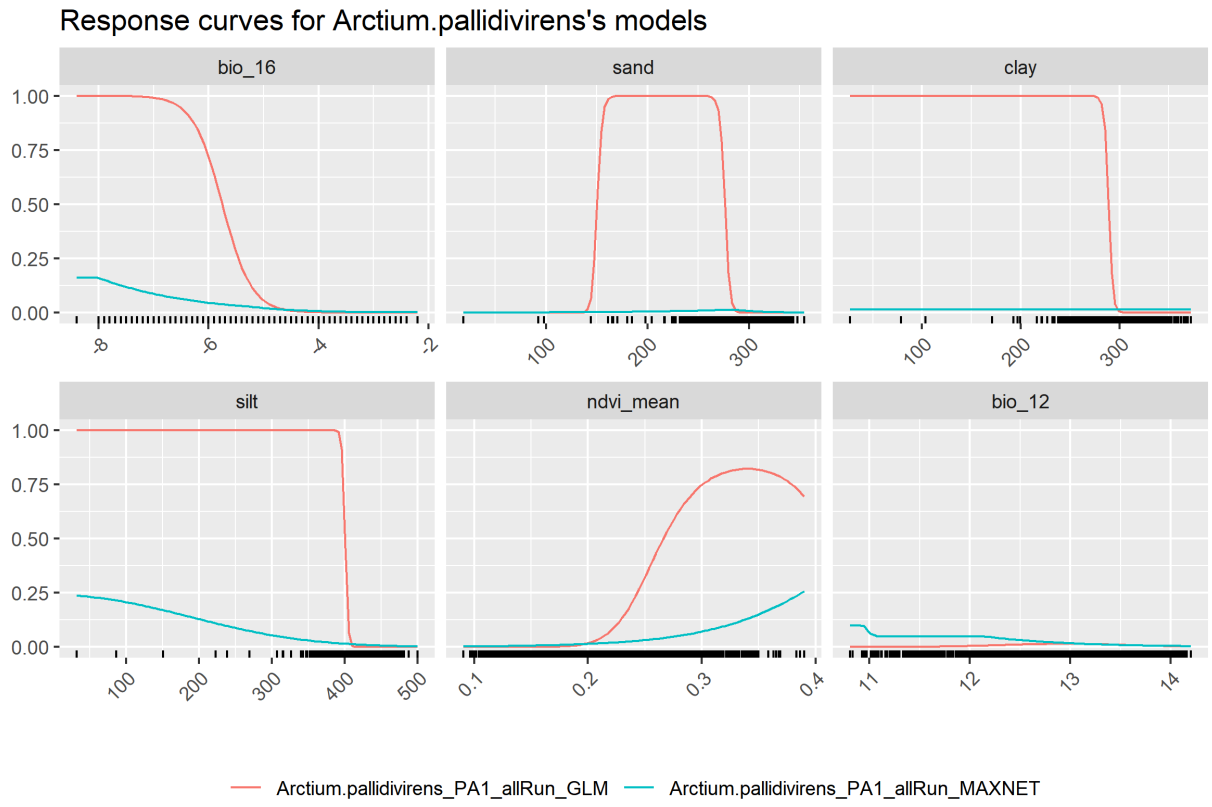


Figure 6. Ecological response curves of *Arctium pallidivirens* along primary environmental gradients.

3.3. Ensemble Projection, Model Congruence, and Spatial Uncertainty of *Arctium pallidivirens*

The potential distribution of *Arctium pallidivirens* was evaluated using an ensemble modeling approach. In this workflow, the outputs from the GLM, GAM, RF, and MAXNET algorithms were integrated to generate a consensus habitat suitability map. This approach successfully minimized the uncertainty associated with individual models, providing a more stable and reliable representation of the species' ecological niche. Spatial congruence analysis across the models indicated that highly suitable habitats are predominantly clustered within the central and eastern portions of the study area, where multiple models concurrently predicted high suitability scores. Specifically, the GLM and MAXNET models delineated high-suitability zones with high precision and distinct spatial clustering; conversely, the GAM model produced relatively smoothed surfaces, while the RF model exhibited a more conservative and fragmented projection. These discrepancies are directly attributable to variations in the underlying mathematical architectures of the algorithms and their respective approaches to defining species–environment relationships.

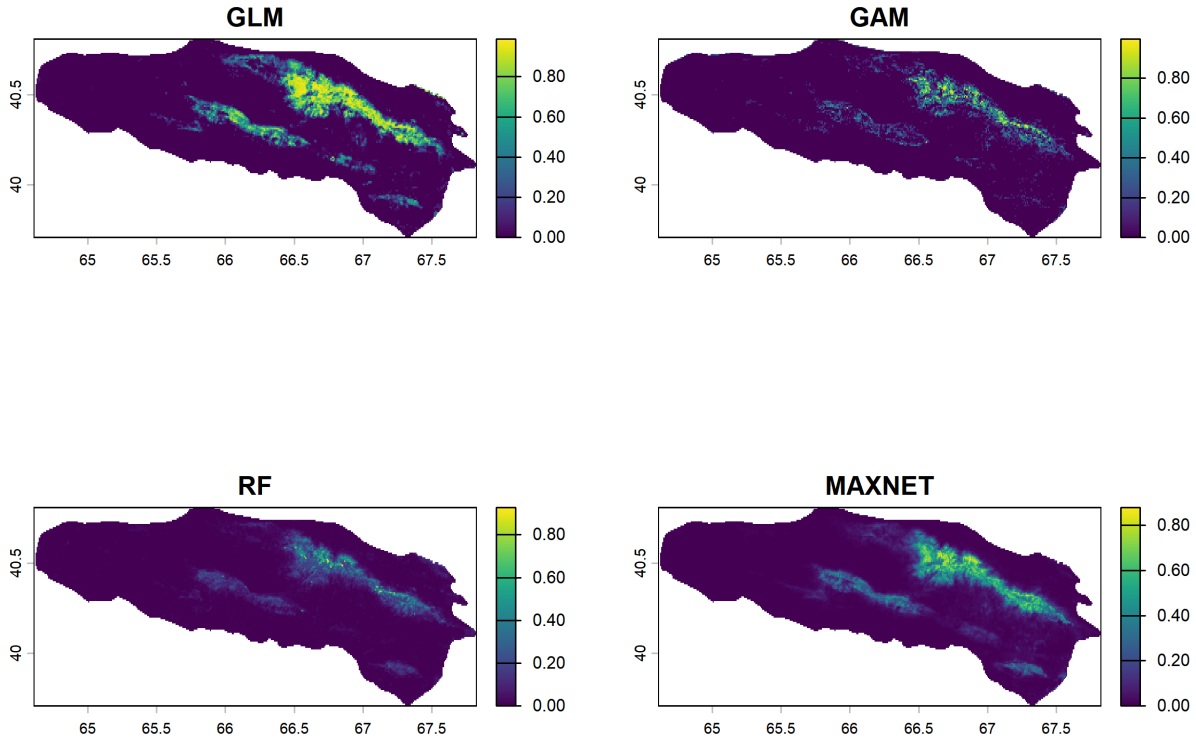


Figure 7. Spatial congruence among GLM, GAM, RF, and MAXNET algorithms predicting habitat suitability for *Arctium pallidivirens* across the study area.

The ensemble suitability map demonstrated that the vast majority of the study area is characterized by unsuitable or low habitat suitability for the species. Specifically, 83.5% of the total area exhibited an extremely low suitability score of <10\%, which encompasses exactly 14,480.74 km2. This overwhelming predominance underscores the species' absence across extensive landscapes and highlights its exceptionally narrow ecological niche (Table 1). Low-suitability areas (10–50%) occupied 13.5% of the region, covering 2,341.20 km2; within these zones, species occurrence is strictly constrained by localized ecological conditions, such as microclimatic variables, moisture availability, and edaphic properties. Moderately suitable habitats (50–80%) were highly limited, accounting for only 2.9% (502.92 km2) of the district. Furthermore, the most optimal habitat conditions (80%) encompassed an extremely minute fraction of the landscape—just 0.1%, covering a mere 17.34 km2. These optimal zones are restricted primarily to highly localized, spatially clustered patches defined by specialized environmental regimes. Overall, the severe restriction of highly suitable zones confirms that *Arctium pallidivirens* operates as a strict ecological specialist under stringent environmental constraints within the Nuratau botanico-geographical district, demonstrating a low potential for natural expansion across broader geographic scales.

Table 1. Area (km 2) and percentage (%) of predicted habitat suitability classes for *Arctium pallidivirens* derived from the ensemble model projection.

Probability class	Area (km²)	% of study area

< 10 %	14,480.74	83.5
10–50 %	2,341.20	13.5
50–80 %	502.92	2.9
> 80 %	17.34	0.1

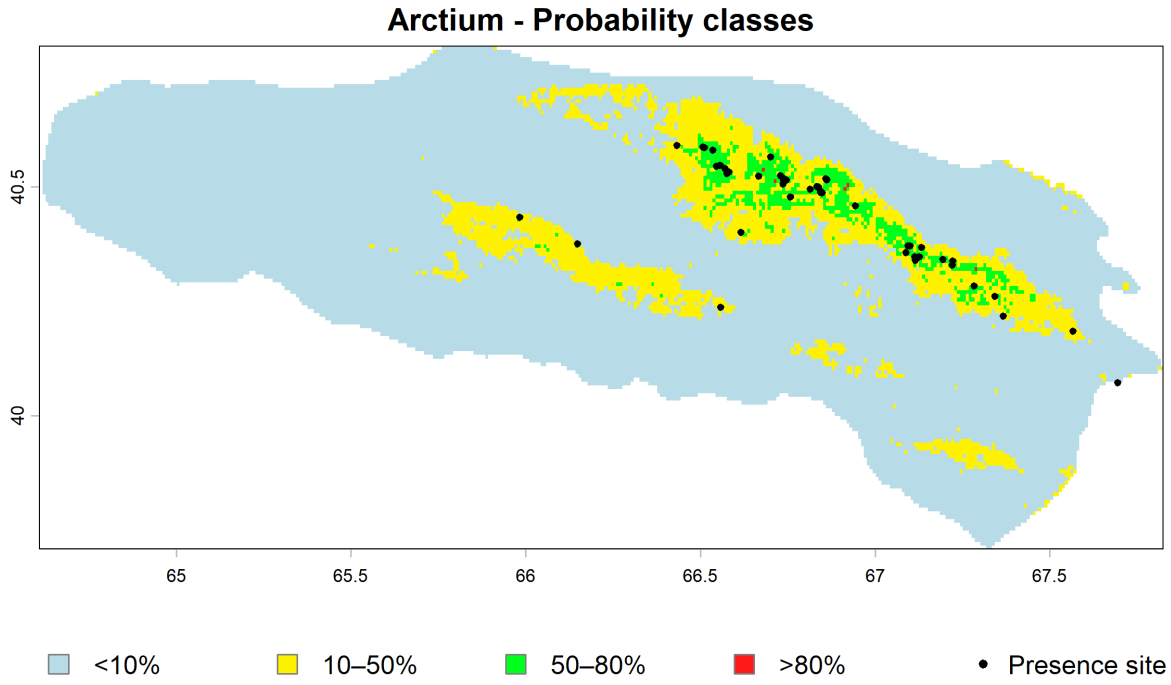


Figure 8. Distribution of predicted presence probability for *Arctium pallidivirens* derived from the ensemble model.

Spatial uncertainty analysis revealed that the reliability of the projections is not uniform across the study area. Areas with low uncertainty were primarily concentrated in the central mountainous regions, indicating highly congruent predictions among the GLM, GAM, RF, and MAXNET models, as well as a stable estimation of habitat suitability. Conversely, zones of high uncertainty were predominantly located within peripheral and transitional areas, reflecting substantial discrepancies among individual model outputs and lower predictive stability. These findings indicate that while the core suitable habitats of Arctium pallidivirens were delineated with high confidence, the peripheral zones remain less reliable due to environmental heterogeneity and inter-model variance. Consequently, regions characterized by both low uncertainty and high habitat suitability should be prioritized in conservation planning, whereas high-uncertainty zones must be interpreted with caution and validated through targeted additional field surveys.

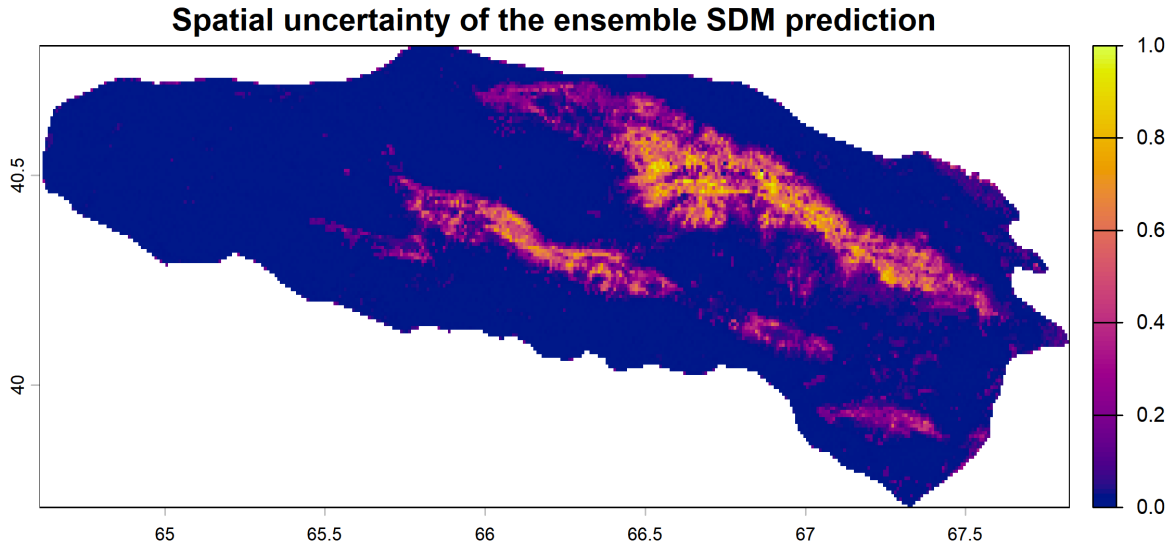


Figure 9. Spatial uncertainty of the ensemble SDM projection for *Arctium pallidivirens*. The uncertainty reflects the variability across modeling algorithms and cross-validation iterations; higher values indicate lower predictive reliability.

3.4. Spatial Congruence of Observed and Predicted Distributions

The total observed distribution range (delineated as a composite polygon encompassing both historical herbarium records from the 1940s–1950s and contemporary field occurrences) covered an area of 792.95 km 2 (4.57% of the study area). The substantial spatial divergence between this footprint and the narrow potential suitability area predicted by the ensemble model (110.16 km 2 , or 0.64%) reflected a high proportion of "observed only" zones, highlighting a pronounced mismatch between the species' historical range and its contemporary ecological refugia. The spatial overlap between the predicted and observed distribution ranges was calculated at 84.81 km 2 , representing 0.49% of the total study region (Table 2). These findings indicate that the suitable habitats identified by the ensemble model accurately capture the contemporary, ecologically optimal environmental regimes for the species.

Table 2. Spatial congruence between observed occurrences and ensemble-predicted suitable habitats for *Arctium pallidivirens* within the study area.

Category	Area (km²)	% of study area	Explanation
Predicted suitable area	110.16	0.64	Area classified as suitable using TSS (Youden) threshold
Observed distribution	792.95	4.57	Area derived from field occurrence records
Overlap (observed + predicted)	84.81	0.49	Area where predicted suitability coincides with observations

Predicted only	25.35	0.15	Suitable area predicted by the model without observations
Observed only	708.14	4.08	Observed occurrences not predicted as suitable

Concurrently, the areas identified by the model as suitable but where the species was not documented ("predicted only") amounted to **25.35 km2 (0.15%)**. Conversely, the regions where the species was historically recorded but not predicted by the modern model ("observed only") encompassed a substantial area of **708.14 km2 (4.08%)**. Rather than indicating a deficiency in model performance, this spatial discrepancy strongly implies that the actual distribution of the species has restricted significantly compared to its historical footprint due to progressive environmental and anthropogenic changes. Ecologically, these results suggest that while macroclimatic and vegetation-related predictors define the species' potential biophysical boundary, the contemporary persistence of *Arctium pallidivirens* is critically governed by fine-scale microhabitat conditions, localized soil heterogeneity, complex microtopography, and potentially, the mitigation of biotic pressures. Consequently, the ensemble model proves highly effective at filtering out human-degraded historical zones and accurately delineating the true extent of the species' current realized ecological niche and contemporary refugia.

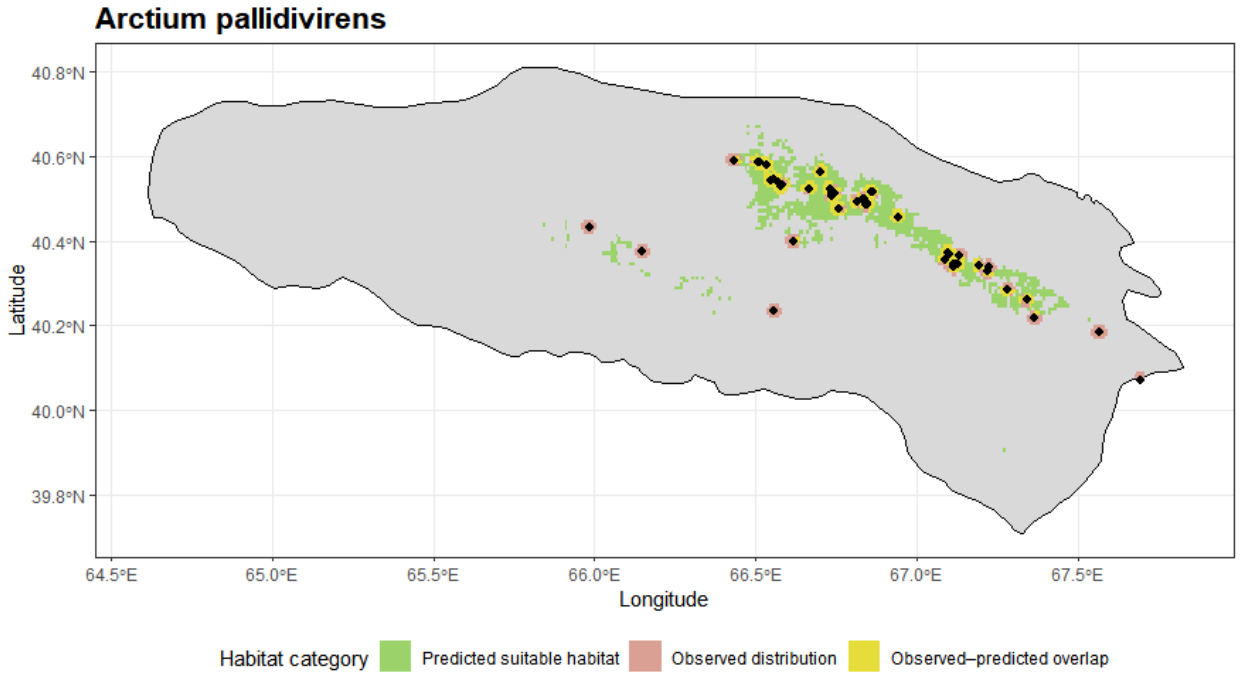


Figure 10. Spatial comparison between the observed distribution range (delineated via a 100% Minimum Convex Polygon using st_convex_hull() in R) and the potential suitable habitat predicted by the BIOMOD2 ensemble species distribution model (SDM) for *Arctium pallidivirens*.

4. Discussion

4.1. Primary Environmental Drivers of *Arctium pallidivirens* Distribution: Comparison with Existing Literature

Our results demonstrate that the spatial distribution of *Arctium pallidivirens* is primarily governed by seasonal moisture regimes, soil properties, and vegetation-associated conditions. These findings are consistent with previous studies showing that the distribution of ecologically specialized and narrow-ranged mountain plant species is controlled by the combined effects of climatic, edaphic, and habitat-related factors (Franklin, 2013; Thuiller et al., 2019). The identification of precipitation of the wettest quarter (Bio16) as the most influential predictor indicates that the distribution of A. pallidivirens is more strongly associated with seasonal moisture availability than with total annual precipitation. This result is consistent with studies demonstrating that plant distributions in semi-arid and continental mountain ecosystems are primarily constrained by water availability during the growing season (Körner, 2021; Miehe et al., 2019). At the same time, the marked decline in habitat suitability under higher Bio16 values suggests that the species is adapted to well-drained gravelly and scree habitats typical of the Nuratau Mountains. Excessive seasonal precipitation in these habitats may promote temporary soil waterlogging, reduce soil aeration, and increase competition from fast-growing mesophytic species, thereby decreasing habitat suitability for *A. pallidivirens*. The high contribution of soil-related variables (sand, clay, and silt fractions) further indicates that the species has a strong dependence on specific edaphic conditions. Our results suggest that A. pallidivirens is restricted to a narrow range of soil textures, confirming its substrate selectivity and the importance of soil physical properties in determining its distribution. This finding agrees with previous studies showing that soil texture regulates mountain plant distributions through its influence on water retention, aeration, and nutrient availability (Hengl et al., 2017). The significant contribution of the Normalized Difference Vegetation Index (NDVI) indicates a close association between *A. pallidivirens* and surrounding vegetation conditions. Rather than representing vegetation density alone, NDVI likely reflects overall microhabitat quality and environmental stability within the habitats occupied by the species. The preference for moderately vegetated sites suggests that *A. pallidivirens* is adapted to open or semi-open habitats rather than densely vegetated communities, consistent with previous ecological interpretations of NDVI in mountain ecosystems (Pettorelli et al., 2011; Lu et al., 2020). Topographic variables, including slope, aspect, and solar radiation, also contributed to explaining the species' distribution by influencing local microclimatic conditions and soil moisture availability (Dobrowski, 2011). In addition, the contribution of the nighttime lights variable should be interpreted cautiously, as it most likely represents an indirect proxy for localized anthropogenic disturbance rather than a direct ecological determinant of the species' distribution. Overall, our findings indicate that the ecological niche of *Arctium pallidivirens* is jointly determined by seasonal moisture availability, soil texture, vegetation structure, and local topographic conditions. These interacting environmental gradients strongly constrain the species' distribution, explaining its highly localized occurrence within the Nuratau Mountains.

4.2. Ecological Interpretation Based on Response Curves.

The response curves provided a functional interpretation of the ecological responses of Arctium pallidivirens along the primary environmental gradients. Both the GLM and MAXNET models revealed a distinct non-linear negative relationship with precipitation of the wettest quarter (Bio16), indicating that habitat suitability declines under elevated seasonal precipitation. This suggests that the species is adapted to habitats with moderate seasonal moisture and possesses a relatively narrow moisture optimum (Randin et al., 2013). The threshold-type response curves for soil variables further confirmed the species' strict edaphic requirements. Habitat suitability reached its maximum within specific combinations of sand, clay, and silt fractions but declined sharply

outside these ranges, highlighting the strong substrate specialization of A. pallidivirens. The response curve for the Normalized Difference Vegetation Index (NDVI) exhibited a positive non-linear relationship, indicating that suitable habitats are associated with moderately developed vegetation cover. Habitat suitability remained low under sparse vegetation conditions and declined slightly at very high NDVI values, suggesting that the species prefers open to semi-open habitats rather than densely vegetated communities (Lu et al., 2020). In contrast, the relatively weak response to annual precipitation (Bio12) further supports that the distribution of A. pallidivirens is influenced more strongly by seasonal moisture availability than by total annual precipitation (Körner, 2021; Miehe et al., 2019). Overall, the response curves confirm that Arctium pallidivirens occupies a narrow ecological niche characterized by well-defined environmental thresholds and a high degree of ecological specialization.

4.3. Study Limitations and Ecological Realities

The findings of this study should be interpreted in light of several methodological considerations and regional ecological realities. Following spatial thinning, the final dataset comprised 34 spatially independent occurrence records. Although limited sample sizes may increase uncertainty in species distribution models, ensemble modelling implemented in the BIOMOD2 framework reduces this uncertainty by integrating predictions from multiple algorithms. A substantial spatial discrepancy was observed between the model-predicted suitable habitat (110.16 km²) and the observed Extent of Occurrence (792.95 km²), primarily due to the large proportion of "observed only" areas. Field surveys combined with the re-evaluation of historical herbarium localities revealed an important ecological pattern. Many sites where *Arctium pallidivirens* was recorded during the 1940s and 1950s, but which were classified as unsuitable by the current model, have experienced pronounced anthropogenic transformation. In these areas, the establishment of settlements, infrastructure development, and intensive livestock grazing have substantially altered natural habitats and likely contributed to the disappearance of previously recorded populations. Therefore, the classification of these historical localities as unsuitable is unlikely to represent a modelling artifact. Instead, it provides evidence consistent with a historical contraction of the species' distribution within the Nuratau Mountains. Consequently, the predominance of "observed only" areas should not be interpreted as model deficiency but rather as an indication that the ensemble SDM effectively identified contemporary climatic refugia where *Arctium pallidivirens* continues to persist despite increasing anthropogenic pressures.

4.4. Conservation Implications

Ensemble modeling outputs facilitated the identification of the most ecologically optimal habitats for *Arctium pallidivirens.* The highly constrained and spatially fragmented nature of these high-suitability zones underscores the species' narrow ecological niche and its critical conservation status. While model projections define the potential climatically and edaphically suitable habitat rather than the absolute realized niche, empirical field observations highlight severe localized threats limiting the species' persistence. Ground-truthing reveals that populations of *Arctium pallidivirens* located outside strictly protected reserves face intense anthropogenic and biotic pressures. Specifically, unmanaged livestock grazing causes physical trampling and microhabitat degradation ***(Figure 11b)***, while severe seed predation by orthopterans (locusts) significantly impairs the species' reproductive output and seed bank replenishment ***(Figure 11a)***.
The areas identified as suitable by the model but currently lacking documented records ("predicted only") should be targeted as potential zones harboring undiscovered populations or isolated micro-refugia, serving as high-priority directions for future field surveys (Raxworthy et al., 2007). Conversely, the high proportion of "observed only" zones provides evidence of substantial historical range contraction, primarily associated with anthropogenic habitat transformation. Consequently, conservation planning should prioritize the protection of stable refugia identified

by the ensemble model while reducing grazing pressure and other localized disturbances to ensure the long-term persistence of this rare endemic species.



Figure 11. Empirical field evidence of direct anthropogenic and biotic pressures on *Arctium pallidivirens* in the Nuratau mountain range: **(a)** seed and inflorescence predation by orthopterans (locusts), and **(b)** habitat degradation and physical trampling caused by unmanaged livestock grazing in non-protected areas.

Conclusion.

This study evaluated the environmental determinants and potential distribution of *Arctium pallidivirens* in the Nuratau Mountains using an ensemble species distribution modelling (SDM) approach. The results demonstrate that the species occupies a narrow ecological niche and depends strongly on seasonal moisture availability, soil properties, and local habitat conditions. The ensemble model identified only 110.16 km² (0.64% of the study area) as currently suitable habitat, whereas the historical Extent of Occurrence reached 792.95 km² (4.57%). This pronounced spatial discrepancy, supported by field verification and historical herbarium records, indicates substantial historical range contraction associated primarily with anthropogenic habitat transformation. The integration of historical occurrence records with ensemble SDM enabled the identification of contemporary refugia where the species continues to persist. These areas represent the highest conservation priority and should be the focus of future monitoring, habitat protection, and field surveys aimed at locating previously undocumented populations. Overall, this study demonstrates the value of integrating historical occurrence data, field observations, and ensemble SDM to improve the conservation assessment of narrowly distributed endemic species. The proposed approach provides a robust spatial framework for identifying priority conservation areas and supporting evidence-based conservation planning in mountain ecosystems.

Author Contributions: K.I. conceived the study, designed the research, carried out field investigations, performed ecological niche modelling and data analyses, and drafted the manuscript. K.S. contributed to botanical surveys, verified herbarium records, and assisted with the ecological interpretation of the results. B.K. processed the environmental datasets, conducted statistical analyses, and contributed to model evaluation. N.B. provided taxonomic expertise, critically revised the manuscript, and contributed to the conservation interpretation. M.T. and Y.A. contributed to the study methodology and participated in writing the Materials and Methods section. G.S. participated in field investigations, facilitated access to the study sites, assisted in locating populations of *Arctium pallidivirens*,and contributed to field data collection. All authors read and approved the final manuscript.

References

Araújo, M. B., & Guisan, A. (2006). Five (or so) challenges for species distribution modelling. *Journal of Biogeography, 33*(10), 1677–1688. https://doi.org/10.1111/j.1365-2699.2006.01584.x

Araújo, M. B., & New, M. (2007). Ensemble forecasting of species distributions. *Trends in Ecology & Evolution, 22*(1), 42–47. https://doi.org/10.1016/j.tree.2006.09.010

Araújo, M. B., et al. (2019). Standards for distribution models in biodiversity assessments. *Science Advances, 5*(1), eaat4858. https://doi.org/10.1126/sciadv.aat4858

Ashcroft, M. B. (2010). Identifying refugia from climate change. *Journal of Biogeography, 37*(8), 1407–1413. https://doi.org/10.1111/j.1365-2699.2010.02300.x

Beshko, N. Yu. (2020). *Flora of Central Asia: The Nuratau Mountains*. Korea National Arboretum.

Barbet-Massin, M., Jiguet, F., Albert, C. H., & Thuiller, W. (2012). Selecting pseudo-absences for species distribution models. *Ecography, 35*(4), 1–12. https://doi.org/10.1111/j.1600-0587.2011.06977.x

Barbet-Massin, M., Jiguet, F., Albert, C. H., & Thuiller, W. (2012). Selecting pseudo-absences for species distribution models: how, where and how many? Methods in Ecology and Evolution, 3(2), 327-338. https://doi.org/10.1111/j.2041-210X.2011.00172.x

Breiman, L. (1996). Random forests. *Machine Learning, 45*(1), 5–32. https://doi.org/10.1023/A:1010933404324

Checklist of the flora of the Nuratau State Nature Reserve. (2025). *GBIF*. https://cloud.gbif.org/uz/resource?r=checklist_nurota&request_locale=ru

Didan, K. (2015). MOD13A2 MODIS/Terra vegetation indices 16-Day L3 global 1 km SIN Grid V006. *NASA EOSDIS Land Processes DAAC*. https://doi.org/10.5067/MODIS/MOD13A2.006

Dullinger, S., et al. (2012). Extinction debt of high-mountain plants under climate change. *Nature Climate Change, 2*, 619–622. https://doi.org/10.1038/nclimate1514

Dormann, C. F., Elith, J., Bacher, S., Buchmann, C., Carl, G., Carré, G., … Lautenbach, S. (2013). Collinearity: A review of methods to deal with it and a simulation study evaluating their performance. *Ecography, 36*(1), 27–46. https://doi.org/10.1111/j.1600-0587.2012.07348.x

Dobrowski, S. Z. (2011). A climatic basis for microrefugia. *Global Change Biology, 17*(2), 1022–1035. https://doi.org/10.1111/j.1365-2486.2010.02245.x

Doniyorova Sh., Kodirov G., Abulfayzov H., Kodirova S. and Muzafarova S. (2025). Investigation of current and future potential distribution of Pamir-Alay endemic species of *Astragalus knorringianus* (Fabaceae). *E3S Web Conf., 670*, 05008. https://doi.org/10.1051/e3sconf/202567005008

Estopinan, J., Bonnet, P., Servajean, M., Munoz, F., & Joly, A. (2024). Modelling species distributions with deep learning to predict plant extinction risk and assess climate change impacts. *arXiv*. https://doi.org/10.48550/arXiv.2401.05470

Farr, T. G., et al. (2007). The Shuttle Radar Topography Mission. *Reviews of Geophysics, 45*(2), RG2004. https://doi.org/10.1029/2005RG000183

Feng, X., Park, D. S., Liang, Y., Pandey, R., & Papeş, M. (2019). Collinearity in ecological niche modeling: Confusions and challenges. *Ecography, 42*, 1–12. https://doi.org/10.1111/ecog.04596

Fenu, G., et al. (2022). Global and Regional IUCN Red List Assessments: 14. *Italian Botanist, 14*, 81–94. https://doi.org/10.3897/italianbotanist.14.97176

Fenu, G., et al. (2023). Global and Regional IUCN Red List Assessments: 16. *Italian Botanist, 16*, 121–133. https://doi.org/10.3897/italianbotanist.16.115947

Fenu, G., et al. (2024). Global and Regional IUCN Red List Assessments: 17. *Italian Botanist, 18*, 141–153. https://doi.org/10.3897/italianbotanist.18.141

Fick, S. E., & Hijmans, R. J. (2017). WorldClim 2. *International Journal of Climatology, 37*(12), 4302–4315. https://doi.org/10.1002/joc.5086

Franklin, J. (2013). *Mapping species distributions: Spatial inference and prediction*. Cambridge University Press.

Funk, C., et al. (2015). The climate hazards infrared precipitation with stations. *Scientific Data, 2*, 150066. https://doi.org/10.1038/sdata.2015.66

Gorelick, N., et al. (2017). Google Earth Engine. *Remote Sensing of Environment, 202*, 18–27. https://doi.org/10.1016/j.rse.2017.06.031

Hannah, L., Midgley, G. F., & Millar, D. (2002). Climate change integrated conservation strategies. *Global Ecology and Biogeography, 11*(6), 485–495. https://doi.org/10.1046/j.1466-822X.2002.00306.x

Hao, T., et al. (2019). Accuracy of SDMs in a changing climate. *Ecography, 42*, 1260–1275. https://doi.org/10.1111/ecog.04222

Hengl, T., et al. (2017). SoilGrids250m. *PLoS ONE, 12*(2), e0169748. https://doi.org/10.1371/journal.pone.0169748

Hersbach, H., et al. (2020). ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society, 146*(730), 1999–2049. https://doi.org/10.1002/qj.3803

Kruckeberg, A. R., & Rabinowitz, D. (1985). Biological aspects of endemism. *Annual Review of Ecology and Systematics, 16*, 447–479. https://doi.org/10.1146/annurev.es.16.110185.002311

Liu, C., White, M., & Newell, G. (2013). Selecting thresholds for prediction of species occurrence. *Journal of Biogeography, 40*(4), 778–789. https://doi.org/10.1111/jbi.12058

Ma, J., Zhang, T., & Li, Y. (2021). Species distribution modeling for conservation planning. *Ecological Modelling, 456*, 109–125. https://doi.org/10.1016/j.ecolmodel.2021.109125

Orsenigo, S., et al. (2022). Global and Regional IUCN Red List Assessments: 13. *Italian Botanist, 13*, 85–94. https://doi.org/10.3897/italianbotanist.13.86714

Orsenigo, S., et al. (2025). Global and Regional IUCN Red List Assessments: 18. *Italian Botanist, 19*, 185–203. https://doi.org/10.3897/italianbotanist.19.156865

Pearson, R. G., & Dawson, T. P. (2003). Predicting the impacts of climate change. *Global Ecology and Biogeography, 12*(5), 361–371. https://doi.org/10.1046/j.1466-822X.2003.00042.x

Rodell, M., et al. (2004). The global land data assimilation system. *Bulletin of the American Meteorological Society, 85*(3), 381–394. https://doi.org/10.1175/BAMS-85-3-381

Song, Y., et al. (2024). Ensemble SDM and phylogeography. *BMC Plant Biology, 24*, 168. https://doi.org/10.1186/s12870-024-04830-1

Thuiller, W., et al. (2009). BIOMOD – ensemble forecasting platform. *Ecography, 32*(3), 369–373. https://doi.org/10.1111/j.1600-0587.2008.05742.x

Thuiller, W., et al. (2005). Climate change threats to plant diversity. *PNAS, 102*(23), 8245–8250. https://doi.org/10.1073/pnas.0409902102

Vermote, E. (2015). MOD09GA MODIS/Terra surface reflectance. *NASA EOSDIS*. https://doi.org/10.5067/MODIS/MOD09GA.006

Volis, S., & Beshko, N. (2023). Preservation of narrow endemics. *Plant Diversity of Central Asia, 2*(2), 82–101. https://doi.org/10.54981/PDCA/vol2_iss2/a3

Wan, Z. (2014). MOD11A2 MODIS/Terra land surface temperature. *NASA EOSDIS*. https://doi.org/10.5067/MODIS/MOD11A2.006

Wisz, M. S., et al. (2008). Sample size effects on SDMs. *Diversity and Distributions, 14*(5), 763–773. https://doi.org/10.1111/j.1472-4642.2008.00482.x

Körner, C. (2007). The use of altitude in ecological research. *Trends in Ecology & Evolution, 22*(11), 569–574. https://doi.org/10.1016/j.tree.2007.09.006

Körner, C. (2021). *Alpine Plant Life* (3rd ed.). Springer.

Lenoir, J., et al. (2017). Local temperatures inferred from plant communities. *Nature, 545*, 223–226. https://doi.org/10.1038/nature22356

Lu, X., et al. (2020). NDVI-based vegetation responses. *Remote Sensing of Environment, 231*, 111272. https://doi.org/10.1016/j.rse.2019.111272

Miehe, G., et al. (2019). Tibetan alpine pasture ecosystem. *Ecological Monographs, 89*(2), e01303. https://doi.org/10.1002/ecm.1303

Pettorelli, N., et al. (2011). The NDVI. *Trends in Ecology & Evolution, 26*(9), 421–429. https://doi.org/10.1016/j.tree.2011.05.005

Randin, C. F., et al. (2009). Climate change and plant distribution. *Global Change Biology, 15*(7), 1689–1703. https://doi.org/10.1111/j.1365-2486.2009.01860.x

Randin, C. F., et al. (2013). Broad-scale niche conservatism. *Journal of Biogeography, 40*(5), 929–941. https://doi.org/10.1111/jbi.12061

Raxworthy, C. J., et al. (2007). Predicting distributions of known and unknown species. *Nature, 426*, 837–841. https://doi.org/10.1038/nature02205

Thuiller, W., et al. (2019). Uncertainty in biodiversity scenario ensembles. *Nature Communications, 10*, 1446. https://doi.org/10.1038/s41467-019-09519-w

Wisz, M. S., et al. (2008). Effects of sample size on SDMs. *Diversity and Distributions, 14*(5), 763–773. https://doi.org/10.1111/j.1472-4642.2008.00482.x