

Article

AI-Driven Conservation of the Endangered Twisted Yew (*Taxus contorta* Griff.) in the Western Himalaya

Salahud Din ¹, Haidar Ali ¹, Thomas Panagopoulos ^{2,*} , Jan Alam ³, Saira Malik ¹ and Hassan Sher ¹

¹ Center for Plant Sciences and Biodiversity, University of Swat, Charbagh 19120, Pakistan; taxuscontorta@gmail.com (S.D.); haidar@uswat.edu.pk (H.A.); sairamalikbotanist11@gmail.com (S.M.)

² Faculty of Science and Technology, University of Algarve, Campus de Gambelas, 8000 Faro, Portugal

³ Department of Botany, Hazara University, Mansehra 21300, Pakistan; janalam@hu.edu.pk

* Correspondence: tpanago@ualg.pt

Abstract

Conserving the threatened West Himalayan endemic *T. contorta* (Taxaceae) is critical due to extinction risks from skewed male- or female-only populations. This study employs ChatGPT-driven artificial intelligence (AI) analysis for textual synthesis and preliminary hypothesis generation to identify favorable propagation sites for *T. contorta* within the Swat district of Pakistan. Over three years (2019–2021), eleven male- or female-only populations of *T. contorta* were surveyed. Environmental data from NASA POWER were analyzed using ChatGPT 3.5 to predict suitable propagation sites, which were then mapped in Google Earth Pro. PCA and hierarchical clustering were applied to identify key environmental variables. Out of 63 generated points, 58 were accurately located in Swat with 92% geographic accuracy, while species-specific general knowledge accuracy was 100%. All points fell within the pre-established *T. contorta* spatial range in Pakistan, with 21 unique sites meeting optimal conditions. Field surveys confirmed 16 new populations. These findings underscore the promising role of AI-driven analysis in conservation planning by identifying and supporting habitat restoration efforts. A bidirectional integration of AI and SDM, combined with remote sensing technologies, represents a novel approach for the effective conservation of endangered plant species.



Academic Editors: Francisco Guerrero and Andreas Reul

Received: 19 August 2025

Revised: 16 September 2025

Accepted: 16 September 2025

Published: 23 September 2025

Citation: Din, S.; Ali, H.; Panagopoulos, T.; Alam, J.; Malik, S.; Sher, H. AI-Driven Conservation of the Endangered Twisted Yew (*Taxus contorta* Griff.) in the Western Himalaya. *Sustainability* **2025**, *17*, 8541. <https://doi.org/10.3390/su17198541>

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Keywords: ChatGPT; propagation site; AI-predicted site; habitat restoration

1. Introduction

T. contorta is a rare species [1], assessed as endangered [2–4], and endemic to the West Himalayas [5]. It is an anti-cancerous medicinal plant [6] and is widely known for producing Taxol (paclitaxel), a secondary metabolite used in the treatment of ovarian and breast cancers [7]. *T. contorta* has been traditionally used for treating headaches, fever, and cough [8]. The plant contains flavonoids, terpenoids, and alkaloids with proven antimicrobial and antioxidant properties [9,10]. *T. contorta* supports wildlife habitat and soil stability on steep terrains, aiding erosion control and forest dynamics [11–13], but is endangered in Pakistan due to overexploitation for medicine, fodder, fuelwood, agriculture, logging, and climate change, with risks from inbreeding and low genetic diversity [14–16]. Conservation is inadequate outside protected areas [17–20], and in Swat, it faces additional threats from illegal Taxol harvesting, timber use, grazing, and weak law enforcement [7,21].

Ecological and environmental research is crucial for conserving species like *T. contorta*. Such studies provide vital data on biodiversity, assess ecological threats, and guide targeted

conservation strategies [21,22]. Climate and soil factors help identify suitable habitats [23], while ecological insights on species richness and distribution patterns inform conservation planning [24]. Effective management also relies on understanding species-specific ecological requirements and ecosystem health [25]. Climate variables are especially important for predicting species distributions and anticipating the effects of environmental change [26]. Restoration efforts depend heavily on such climatic and ecological datasets [22,23].

Recently, artificial intelligence (AI) has emerged as a transformative tool for biodiversity conservation [27,28]. AI applications, such as ChatGPT, are based on a large language model (LLM) using transformer architecture and deep learning, generating outputs through prompt engineering [28]. It enhances conservation efforts by aiding in habitat identification and synthesizing large volumes of ecological and spatial data. These tools integrate satellite imagery, ground surveys, and climate data to model suitable habitats, allowing for informed decision-making in conservation planning [29]. AI models can also forecast future habitat risks due to land use changes, climate effects, and human pressures [30].

Species distribution modeling (SDM) traditionally relies on approaches such as the maximum entropy modeling approach (MaxEnt) and logistic regression, which require detailed occurrence records and environmental layers. These methods also give large spatial scale distribution as presence-absence data and do not give single GPS points, which is not favorable for species like *T. contorta*, whose ecological amplitude ranges from a few meters to a maximum of a few hundred meters. However, in data-deficient regions like Swat, Pakistan, such requirements often limit conservation planning for rare species. In this study, ChatGPT was employed as an AI-driven, knowledge-based tool to rapidly identify potential habitats by integrating ecological traits, climatic preferences, geographic context, and literature-based understanding to urgently propagate *T. contra* to achieve sex equilibrium in single-gender populations. Unlike conventional models, ChatGPT offers immediate site-specific insights without reliance on extensive georeferenced data.

2. Materials and Methods

2.1. Sample Collection Procedures

Eleven populations of *T. contorta* having only male or female individuals were searched and documented in district Swat, Pakistan, in three consecutive years, from 2019 to 2021. Geographic coordinates were recorded and confirmed with Google earth pro.

2.2. Environmental Variables Selection and Data Acquisition

Favorable environmental factors were identified based on the ecological preferences of *T. contorta* within its natural distribution range [5,14,31]. Three-year climatological data were retrieved from the NASA POWER database under the agroclimatology domain, with variables reported using the official names and abbreviations provided by the data source (Table 1). The “POWER Single Point Data Access Widget” supplies near real-time datasets at a spatial resolution of $0.5^\circ \times 0.5^\circ$ for single point coordinates (latitude/longitude). For each parameter, annual means were calculated from the monthly data across three consecutive years (January 2019–December 2021). These annual means were then averaged to obtain a three-year climatological mean for each environmental variable. The same procedure was applied to both observed *T. contorta* occurrence points and AI-predicted locations (Supplementary Materials).

2.3. Derivation of Spatial Data Using ChatGPT

To assess ChatGPT’s species-specific knowledge of *T. contorta*, a series of general questions were initially posed regarding its taxonomy, distribution, geography, conservation status, threats, and coordinates within its natural range in Pakistan. Subsequently, natural

distribution records of *T. contorta*, along with relevant ecological and environmental variables, were provided to ChatGPT to generate suitable latitude and longitude points for potential propagation, with the search restricted to Swat District, Pakistan. As ChatGPT favored decimal degrees (DD) for geographic outputs, field data originally recorded in degrees and decimal minutes (DDM) were converted to DD format prior to input. Environmental parameters were supplied using the standardized abbreviations of NASA POWER, with complete definitions retrieved from the official database. No explicit ecological thresholds (e.g., altitudinal range, climatic limits, or mean temperature ranges) were incorporated into the prompts, apart from limiting the geographic distribution to Swat District. Example prompts included the following: “What do you know about the taxonomy, distribution, conservation status, and threats to *T. contorta* in Pakistan? Also provide some latitude and longitude points for favorable presence of *T. contorta*?” and “Based on these geographic coordinates and ecological data obtained from NASA POWER for actual populations of *T. contorta*, provide geographic coordinates for its propagation, but restrict your search to Swat District, Pakistan.”

Table 1. Environmental variables selected from NASAPOWER for *T. contorta* natural populations and AI-predicted sites.

Category	Parameter Name	Abbreviation	Unit
Solar Radiation	All Sky Surface Photosynthetically Active Radiation Total	ALLSKY_SFC_PAR_TOT	W/m ²
Temperature	Temperature at 2 Meters Maximum	T2M_MAX	°C
	Temperature at 2 Meters Minimum	T2M_MIN	°C
Humidity/Precipitation	Specific Humidity at 2 Meters	QV2M	g/kg
	Relative Humidity at 2 Meters	RH2M	%
	Precipitation Corrected Sum	PRECTOTCORR_SUM	mm
Wind/Pressure	Surface Pressure	PS	kPa
Soil Properties	Surface Soil Wetness	GWETTOP	1
	Profile Soil Moisture	GWETPROF	1
	Root Zone Soil Wetness	GWETROOT	1

2.4. Statistical Analysis Framework

Descriptive statistics (mean, max, and min) were calculated from the actual data. Data was transferred into Z-scores. The Krushal Wallis test was performed to find differences in medians of every parameter. Correlations were obtained using the Pearson correlation coefficient (r). Hierarchical clustering was performed, and a constellation plot along with a cluster summary was obtained. PCA was performed to reduce dimensionality and assess the environmental similarity between AI-predicted and actual *T. contorta* sites based on multivariable ecological data. Statistical analysis was performed with MS Excel 2019, JASP 0.95.0.0, JMP 17, and PAST 5.0.2 software.

2.5. Spatial Analysis Approach

Actual populations and AI-generated points in decimal coordinates were plotted in google earth pro. Images were taken having distance bar and direction.

2.6. Field Validation of AI-Derived Outputs

Field validation of the AI-generated sites was conducted through direct field visits and by engaging local communities through the dissemination of AI-based habitat predictions, followed by their verification through on-site observations.

2.7. Altitudinal Range Determination

Although AI-generated predictions suggested a broader altitudinal range, this study focused on the 1800–3100 m elevation band, as it corresponds with the core ecological range of *T. contorta* based on field records, *T. contorta* occurrence points taken from the Global Biodiversity Information Facility (GBIF; GBIF.org (26 May 2024) GBIF Occurrence Download <https://doi.org/10.15468/dl.fwrsvg>) database in Geospatial Conservation Assessment Tool (GeoCAT 2024), and the literature [3,5,32]. This range also represents the typical elevation of moist temperate forests in the Hindu Kush Himalayas, which provide the most suitable habitat for the species.

2.8. Propagation Methods

Male plant twigs were added to areas that had only female plants, and female plants were added where only males were present. The same approach was used at AI-identified locations. At sites with no existing *Taxus* populations, both male and female plants were planted during December and January of 2024–2025.

3. Results

3.1. Distribution of Natural Populations of *T. contorta*

A total of 11 populations having only male or female individuals were documented and analyzed from 2019 to 2021 in the Hindu Kush-Himalayan region of District Swat, Pakistan. *T. contorta* populations were found on both sides of the river Swat. The lowest altitude was recorded as 2157 m at Miandam, while the highest population was found at Manrai, situated at 2739 m. These populations extended longitudinally from 34.74° E to 35.44° E and latitudinally from 72.15° N to 72.67° N from lower to upper Swat. All the 11 populations came in a polygon of 1958 km². The highest length from one population to another population was 78.2 km, while the width was 38.5 km. The *T. contorta* populations were found to be concentrated primarily in the higher altitude areas along the north side of the River Swat. Out of the 11 populations documented, 6 were located on the north side of the river, while 5 were on the south side of the Swat River (Figure 1).

3.2. Validation of Environmental and Ecological Variables for Natural and AI-Derived *T. contorta* Populations

3.2.1. Descriptive Profiles of Environmental Factors

AI-identified sites were generally found at lower altitudes, as indicated by their higher mean surface pressure (83 kPa) compared to 78 kPa at natural population sites. In terms of atmospheric moisture, AI-predicted areas exhibited slightly higher specific humidity (8 g/kg vs. 7 g/kg), while relative humidity remained similar, though marginally lower at AI sites (63% vs. 65%). This suggests that the AI-selected areas may offer slightly greater overall moisture availability.

Soil moisture profiles across surface and root zones were broadly comparable between both site categories. However, subsurface moisture was slightly higher at AI sites (mean: 0.71) compared to natural populations (mean: 0.68), potentially indicating more favorable water retention conditions. Thermal conditions also differed, with AI sites experiencing warmer maximum temperatures (35 °C vs. 32 °C) and less extreme minimum temperatures (−9 °C vs. −15 °C), thereby offering a more thermally stable environment that may support a broader physiological tolerance range for the species. In terms of precipitation, AI-selected sites received a higher mean annual total (4572 mm) relative to natural population sites (4151 mm), though both groups exhibited a wide range of precipitation values (2655–6020 mm), suggesting overlapping rainfall regimes. Lastly, AI-predicted habitats had slightly lower solar radiation exposure (85 W/m²) compared to natural population

sites (89 W/m^2), possibly indicating more shaded or cloudier microclimatic conditions (Table 2).

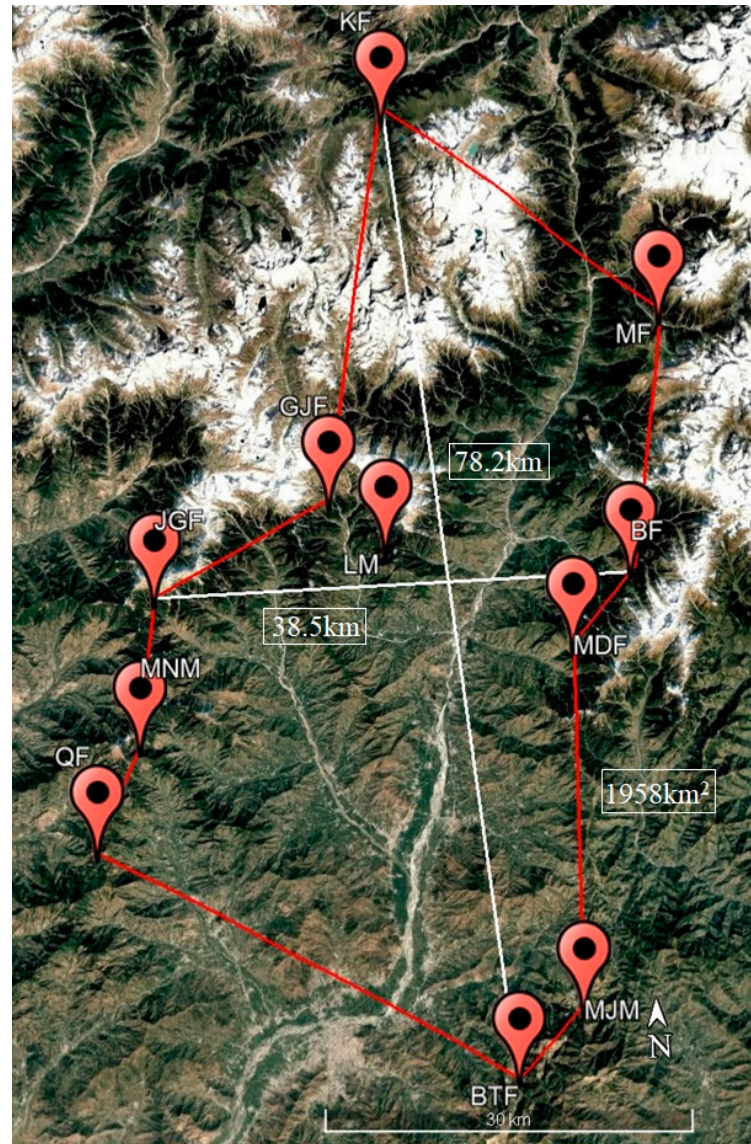


Figure 1. Distribution of natural populations of *T. contorta* in Swat, Pakistan. Locality abbreviations: Kalam (KF), Mankyal (MF), Bashigram (BF), Gabin Jabba (GJF), Jarogo (JGF), Lalko (LM), Miandam (MDF), Malam Jabba (MJM), Manrai (MNM), and Qalagai (QF). F and M, the end of abbreviations, correspond to F (female), M (male).

Table 2. Descriptive summary of ecological and environmental variables for *T. contorta*.

Points	Statistic	V1	V2	V3	V4	V5	V6	V7	V8	V9	V10
AI	Min	66	4	59	0.65	24	−29	0.63	0.65	2655	75
	Max	90	9	69	0.74	39	−4	0.72	0.78	5300	95
	Mean	83	8	63	0.69	35	−9	0.67	0.71	4572	85
CR	Min	66	4	59	0.65	24	−29	0.63	0.65	2655	85
	Max	89	10	69	0.75	40	−4	0.73	0.76	6020	95
	Mean	78	7	65	0.69	32	−15	0.67	0.68	4151	89

Abbreviations: (V1 = PS, V2 = QV2M, V3 = RH2M, V4 = GWETTOP, V5 = T2M_MAX, V6 = T2M_MIN, V7 = GWETPROF, V8 = GWETROOT, V9 = PRECOTCORR_SUM, V10 = ALLSKY_SFC_PAR_TOT). AI (artificial intelligence, ChatGPT points), CR (Current, natural populations).

3.2.2. Correlation Patterns Among Variables

Surface pressure is positively correlated with specific humidity, maximum temperature, minimum temperature, and precipitation. This indicates that areas with higher atmospheric pressure tend to have warmer air, more moisture content, and greater overall rainfall, suggesting favorable growing conditions. Specific humidity also showed a strong positive correlation with both min and max temperatures and precipitation, which supports the idea that warmer areas hold more atmospheric moisture and also experience more precipitation. However, specific humidity is strongly negatively correlated with relative humidity, which implies that although the absolute moisture in the air may be high, the relative humidity can decrease as temperatures rise, since warmer air has a higher capacity to hold water vapor. Relative humidity showed strong negative correlations with almost all other variables, especially surface pressure, specific humidity, both temperatures, and precipitation. This is typical in regions where warmer and more pressurized conditions lead to lower relative humidity, even if absolute moisture levels are high. This inverse relationship helps explain dry air conditions in high-temperature regions (Kalam) despite high total moisture content. Surface soil wetness has weaker correlations with most variables, suggesting it is more immediately influenced by short-term rainfall or evaporation. In contrast, soil moisture and root zone soil wetness showed strong positive correlations with one another and with surface pressure, specific humidity, both temperatures, and precipitation. This implies that long-term stable moisture availability in deeper soil layers is shaped by cumulative environmental conditions, important for root systems. Precipitation is strongly correlated with temperatures, pressure, and deeper soil wetness, showing that areas with more rainfall also have favorable conditions for long-term soil moisture retention. However, photosynthetically active radiation showed a weak correlation with most other variables, suggesting that solar radiation levels vary independently of temperature, humidity, and moisture, likely due to local topography or exposure (Figure 2).

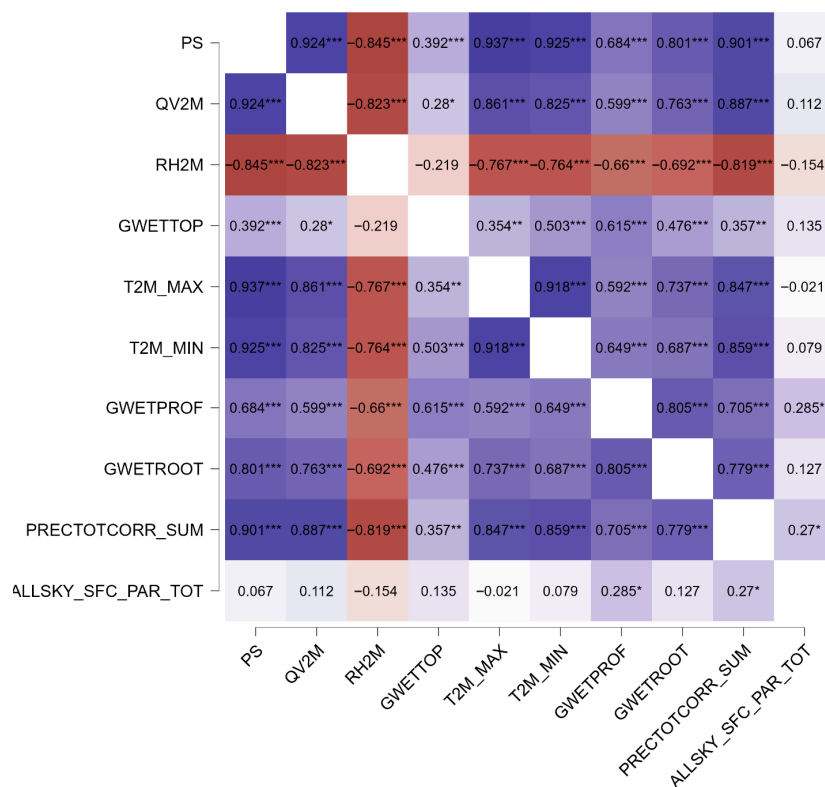


Figure 2. Pearson’s correlation coefficients (r) of environmental variables for AI-derived and field observed populations of *T. contorta*, (Significance level: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$).

The Cronbach's alpha values for the environmental variables indicate good internal consistency, with values ranging from 0.78 to 0.93. Relative humidity and all sky surface photosynthetically active radiation showed the highest alpha values, suggesting they are slightly less consistent with the rest. Overall, the variables demonstrate strong reliability and were found suitable for further multivariate analyses.

3.3. Hierarchical Clustering of Populations Based on Environmental Variables

The two-way cluster analysis constellation plot revealed clear ecological differentiation among both natural populations and ChatGPT-predicted localities, based on key environmental variables. All sites were grouped into two main ecological groups. Group 1 was further divided into seven distinct clusters, while Group 2 formed a single, ecologically unique cluster. In Group 1, Cluster 1 comprised ten AI-predicted localities along with one natural population (Banjoot) and was characterized by higher surface pressure, elevated maximum and minimum temperatures, greater precipitation, and enhanced soil properties. Cluster 2 included twelve AI-predicted sites shaped mainly by high specific humidity and low relative humidity. Cluster 3, consisting of six AI points, was defined by high soil moisture and temperature values, accompanied by reduced photosynthetically active radiation (PAR). Cluster 4 was more diverse, containing nine AI-predicted localities and five natural populations (Miandam, Lalko, Jarogo, Gabbin Jabba, and Bashigram), and reflected moderate ecological conditions with relatively low PAR. Cluster 6, made up of twelve AI points, exhibited moderate environmental values with notably high PAR levels. Cluster 7 included nine AI-predicted sites and three natural populations (Malam Jabba, Manrai, and Qalagai), all showing overall moderate environmental conditions, lower soil moisture levels, and slightly elevated PAR. In contrast, Group 2, represented solely by Cluster 8, comprised three AI-predicted sites and two natural populations (Kalam and Mankyal) and was distinguished by lower surface pressure, lower specific humidity, reduced temperatures and soil values, but higher relative humidity and moderate PAR (Figure 3). These clustering patterns demonstrate that both observed and AI-predicted localities span a broad range of ecological conditions, highlighting the environmental diversity across the study region.

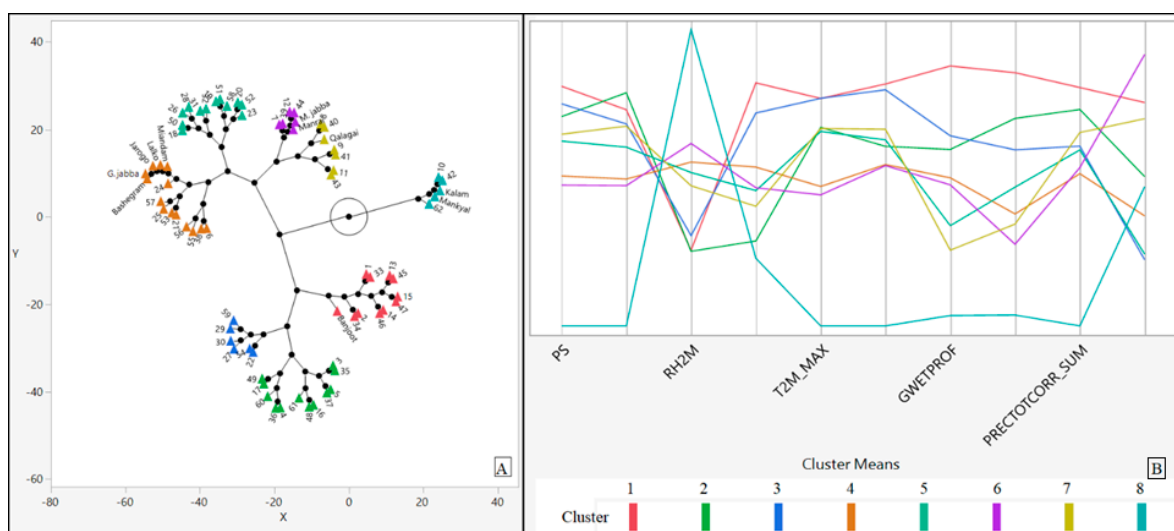


Figure 3. Two-way cluster constellation plot of AI and natural populations of *T. contorta*. (A) Grouping of localities. (B) Cluster means of groups based on environmental variables.

3.4. Principal Component Analysis (PCA) of Population–Environment Relationships

A PCA was conducted to reduce dimensionality in environmental variables of both natural populations and AI-derived locations of *T. contorta*.

3.4.1. Cluster Characteristics of Environmental Variables Derived from PCA

Principal component analysis based on grouping made on hierarchical clustering showed the grouping of natural populations and AI-predicted points into 2 clusters on the basis of environmental conditions. Cluster 1 includes the majority of variables, such as surface pressure, specific humidity, relative humidity, topsoil moisture, maximum and minimum temperatures, root zone soil moisture, soil moisture profile, and precipitation. This cluster explains 81.7% of the variation within its group and contributes 65.4% to the total variation in the data. The most representative variable in this cluster is PS, with a very high R-squared value (0.95), indicating that it is highly representative of the structure within this group. Other variables in Cluster 1 also show high R-squared values with their own cluster (generally above 0.74) and low R-squared values with the next closest cluster, confirming strong internal consistency. Notably, profile soil moisture has the weakest association with its cluster, suggesting it may be more borderline in group similarity compared to others. Cluster 2 includes only two variables: solar radiation and surface soil wetness. This cluster explains 56.7% of its internal variation and contributes 11.3% to the total variation in the dataset. Solar radiation serves as the representative variable for this cluster. Both variables show moderate R-squared values with their own cluster (0.567), and especially for solar radiation, the R-squared value with the next closest cluster is extremely low (0.021), showing it is quite distinct from the variables in Cluster 1. Surface soil wetness, however, shows a higher R-squared with the next closest cluster (0.19), suggesting it is less distinct and may share features with Cluster 1 variables (Table 3).

Table 3. Cluster summary of PCA conducted on ecological and environmental variables of AI and *T. contorta* natural populations.

Cluster	Number of Members	Most Representative Variable	Cluster Proportion of Variation Explained	Total Proportion of Variation Explained	Variable	RSquare with Own Cluster	RSquare with Next Closest	1-RSquare Ratio
1	8	PS	0.817	0.654	PS	0.95	0.093	0.055
					PRECTOTCORR_SUM	0.889	0.173	0.134
					QV2M	0.863	0.068	0.147
					T2M_MAX	0.859	0.049	0.149
					T2M_MIN	0.848	0.149	0.178
					RH2M	0.778	0.061	0.237
					GWETROOT	0.744	0.161	0.305
2		ALLSKY_SFC_PAR_TOT	0.567	0.113	ALLSKY_SFC_PAR_TOT	0.567	0.021	0.442
					GWETTOP	0.567	0.19	0.535

3.4.2. PCA Biplot of Environmental Variables with Population Localities and PCs Variable Loadings

Principal Component Analysis (PCA) revealed distinct environmental gradients explaining variation among the study sites. The first three principal components (PCs) accounted for a cumulative 78.9% of the total variation, with PC1 explaining 67.8%, followed by PC2 (11.8%) and PC3 (6.2%). The biplot illustrated that AI-predicted suitable habitats clustered distinctly from collected sites, primarily along PC1. Variables such as surface pressure, specific humidity, relative humidity, precipitation, maximum and minimum temperatures, and soil moisture at different depths loaded strongly on PC1, suggesting that this axis represents a warm, moist, and pressure-driven environmental gradient. The loading plot further clarified the contribution of variables to each PC. PC2 captured variation due to solar radiation and topsoil moisture, suggesting a secondary gradient linked to sunlight

exposure and surface wetness. PC3 showed moderate influence from soil moisture layers and solar radiation, capturing environmental heterogeneity (Figure 4). Together, these results indicate that combinations of different environmental conditions along different PCs make different microclimates for the growth of *T. contorta* in the area.

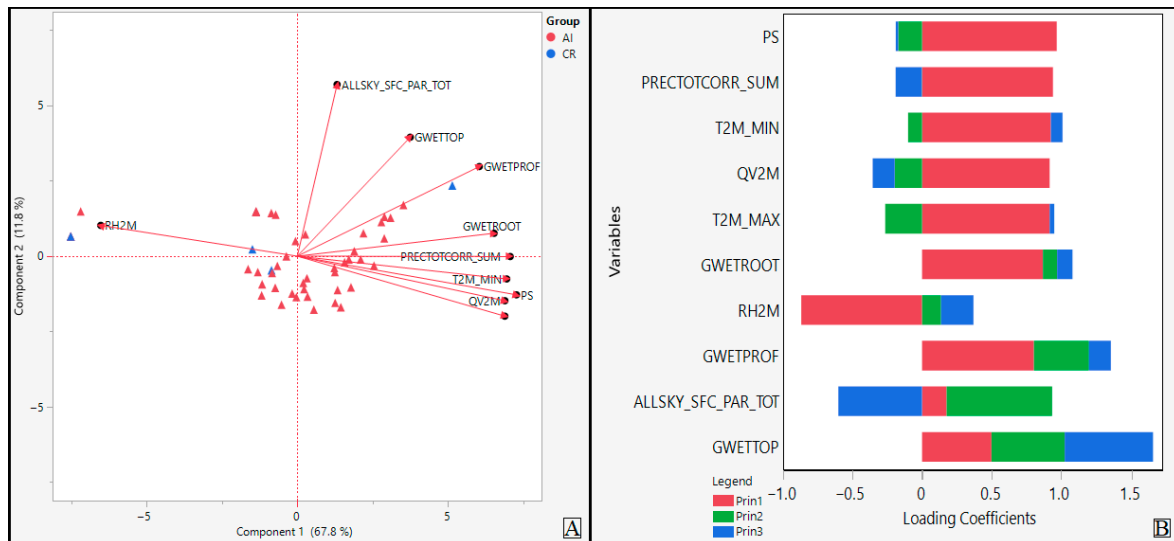


Figure 4. Distribution of localities and environmental conditions on principal components of PCA conducted on ecological and environmental variables of AI and natural *T. contorta* population GPS points. (A) Biplot (B) loading plot.

3.5. AI Model Predictions: Accuracy and Validation

A total of 63 geospatial points were identified in two searches for favorable *T. contorta* propagation sites within three districts: Swat, Chitral, and Upper Dir. Among these, 58 points were located in Swat, four in Chitral, and one in Upper Dir. Of the 63 points, 13 were unique latitude-longitude coordinates, while 25 coordinates were recorded twice. To refine the selection criteria, we established an optimal altitudinal range for *T. contorta* propagation between 1880 m and 3100 m above sea level. In Chitral, two points were identified as unique occurrences, while one was recorded twice. Notably, all Chitral points were located above 3100 m. In Upper Dir, the single identified point was situated below 1800 m, falling outside the optimal range. In Swat, 13 points were located above 3100 m, while 8 were found below 1800 m. The remaining points in Swat fell within the established altitudinal range of 1880–3100 m, suggesting that this district holds the highest potential for *T. contorta* propagation under favorable environmental conditions.

3.6. AI Generated Suitable Locations for *T. contorta* Propagation in Swat

A total of 58 locations were identified in Swat as potential sites for *T. contorta* propagation through AI-based analysis. Among these, 10 locations were identified as single-occurrence points, while 24 were recorded as double-occurrence points. Out of the 58 locations, 13 were situated above 3100 m and 8 were below 1800 m. These 21 points were not considered suitable for *T. contorta* populations despite exhibiting favorable environmental conditions and correlations with identified habitat variables. Specifically, 8 points (comprising two single occurrences and three double occurrences) were located below 1800 m, while 13 points (including five double occurrences and three single occurrences) were positioned above 3100 m.

A total of 37 points, including 5 single-occurrence and 16 double-occurrence points, were found to be within the optimal altitudinal range for *T. contorta* propagation. When considering double-occurrence points as single locations, the total number of favorable sites

was reduced to 21. Field investigations revealed that more than 50% of these AI-predicted sites already supported existing *T. contorta* populations, validating the effectiveness of AI-based site selection for conservation and propagation efforts.

3.6.1. AI-Generated Validation of Natural *T. contorta* Populations

A total of three AI-generated points were found to align precisely with natural *T. contorta* populations in Mankyal and Kalam. Among these, a single occurrence was recorded for Kalam, while Mankyal exhibited a double occurrence, confirming the accuracy of AI-based site selection in identifying existing populations. Additionally, two more points, one single and one double, were identified in Kalam and Mankyal, respectively. However, these locations were situated at higher elevations exceeding 3500 m, beyond the optimal altitudinal range established for *T. contorta* propagation. Despite this, environmental and ecological variables at these sites showed strong correlations with natural *T. contorta* habitat conditions (Figure 5).

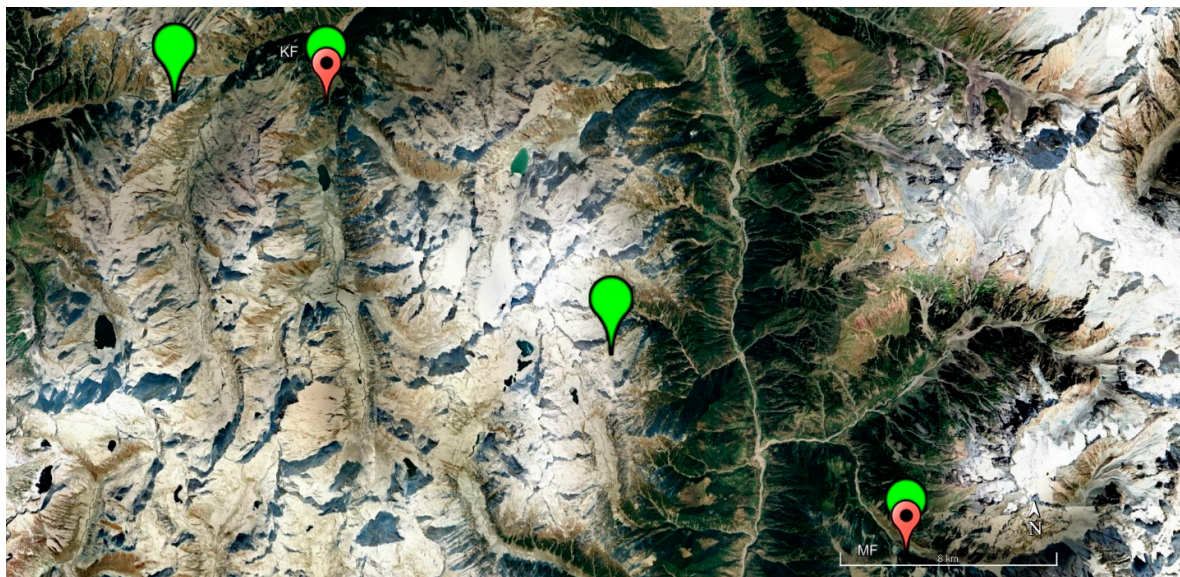


Figure 5. AI-predicted points confirming natural populations of *T. contorta* in Mankyal and Kalam.

3.6.2. AI-Predicted Points at Lalko and Gabbin Jabba

AI-generated predictions identified several potential sites for *T. contorta* populations in the Lalkoo region, with all the predicted points falling within the favorable altitude range of 1800–3100 m. These predictions comprised six double points and two single points, of which two single points and three double points were confirmed to host natural *T. contorta* populations (Figure 6A). Geographically, the predicted points were distributed on both sides of the Lalko stream. On the southern side, there were three points: one single and one double point, with confirmed *T. contorta* populations present at these locations. On the northern side, five double points and one single point were identified, with *T. contorta* populations confirmed at additional double and single points here. A male population was found at a point predicted at Gabbin Jabba (Figure 6B). Environmental conditions at the AI-generated points were analyzed and found to be correlated with the habitat requirements of *T. contorta* natural populations.

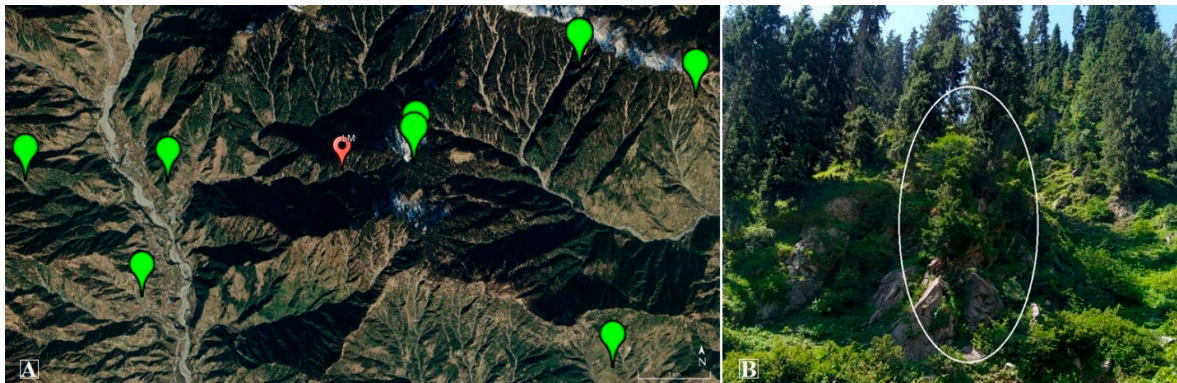


Figure 6. Field validation of AI-predicted points. (A) Distribution of AI-predicted points at Lalko, Swat (green: AI-predicted; red: natural male population). (B) AI-predicted point with a confirmed male *T. contorta* individual at Gabbin Jabba, Swat.

3.6.3. AI-Generated Points Distribution on the North and South Aspects of the Swat River

The distribution of AI-generated points revealed a clear preference for the north-facing aspect of the Swat River. Only three double occurrence points were identified on the south-facing aspect, while all remaining points were located on the north-facing side. Among the points on the southern aspect, one double occurrence point at Bahrain was found to coincide with a natural population of *T. contorta*, while another double-occurrence point was positioned exactly at the known *T. contorta* population in Mankyal (Figure 7). All AI-predicted points on the south-facing slopes were situated within the natural altitudinal range of *T. contorta* (1800–3100 m). On the north-facing aspect, a total of 10 single-occurrence points and 21 double-occurrence points were recorded. Of these, 8 points (comprising two single-occurrence and three double-occurrence points) were located below 1800 m, while 13 points (including five double-occurrence and three single-occurrence points) were situated above 3100 m. These 21 points were excluded from consideration as suitable habitats for *T. contorta* populations. The remaining 31 AI-generated points on the north-facing aspect fell within the natural altitudinal range of *T. contorta*. Field investigations confirmed that more than 50% of these locations already supported natural *T. contorta* populations, further validating the predictive accuracy of AI-generated site selection.



Figure 7. Distribution of AI-generated points and natural populations of *T. contorta* on the north and south aspects of the Swat River (white: AI-predicted; red: existing populations).

4. Discussion

Due to ChatGPT's ability to generate human-like text and answer complex questions, it became one of the fastest-growing and widely accepted applications in the history of the Internet in a short span of time [33]. With the daily use of ChatGPT, developers have been taking advantage of the large amount of data generated in these interactions and have created language models closely adjusted to the language, tone, style, specific needs, and preferences of each user. This has allowed them to generate responses that are more personalized and precise in different disciplines [34]. Currently, there is a wide discussion about the use, scope, limitations, and applications of ChatGPT both in daily life and in education and academia, which is why the number of articles testing its efficiency has increased significantly in recent times [33].

In this study, we assessed the effectiveness of ChatGPT as a resource for locating potential propagation areas and for retrieving both broad and detailed information regarding the geographic range of *T. contorta* and species-specific knowledge. Overall, the findings indicated that ChatGPT responded reliably to inquiries concerning the species' distribution and related ecological knowledge.

4.1. Effectiveness of AI in Identifying Suitable Habitats

Artificial intelligence (AI), including tools like ChatGPT, has demonstrated significant potential in identifying suitable habitats for *T. contorta* (twisted yew), an endangered tree species. Furthermore, AI has proven invaluable in analyzing complex ecological data, leading to significant insights that can aid conservation strategies, such as focusing preservation efforts within areas that exhibit the highest suitability for *T. contorta* [35]. All of the AI-generated points fell within the natural range of *T. contorta*, sharing the same ecological and environmental variables. *T. contorta* is primarily confined to the Hindu Kush-Himalayan ranges in Pakistan, with notable studies by Poudel et al. [14] and Möller et al. [5]. Poudel et al. [14] provided a comprehensive analysis of *T. contorta* across its native range, identifying regions with high probabilities of occurrence. Their study achieved an AUC value of 0.948, indicating excellent model performance in capturing the species' ecological niche. High rainfall supports greater diversity [31], while declining trends threaten *T. wallichiana* habitats [7]. Modeling by Möller et al. [5], Poudel et al. [14], and Rathore et al. [36] reinforces the influence of precipitation and altitude on species adaptation and vulnerability. Crowther et al. [37] further link precipitation to global tree density. Notably, our AI-assisted predictions matched the natural range of *T. contorta* in Swat with 100% accuracy, confirming findings from Poudel and Möller [5,32]. Natural populations and AI-driven environmental variables of precipitation, soil, and variables linked with altitude and topography played a significant role in the geographic distribution and statistical grouping of *T. contorta* throughout the research area. AI tools like ChatGPT show promise for future integration with ecological modeling to enhance conservation efforts.

4.2. Challenges in Species-Specific Knowledge Accuracy and Model Limitations

The findings showed that ChatGPT consistently provided over 60% accuracy in identifying propagation areas when constrained within an altitudinal band of 1800–3100 m. The model demonstrated complete alignment (100%) with known environmental parameters of *T. contorta* habitats. However, mismatches in naming certain localities were noted, likely due to limited availability of regional scientific data, a limitation ChatGPT itself acknowledged. The nomenclature surrounding *Taxus* L. species in Pakistan remains inconsistent in existing literature, and targeted studies remain scarce. Google Scholar searches reveal few direct references to the species, highlighting the restricted scientific documentation available. Despite this, information regarding the species' ecological and historical con-

text is accessible through specialized databases, online platforms, and published works. ChatGPT's ability to aggregate and summarize such information aligns with Fatani [38], who praised its data synthesis capacity, even as Mehnen et al. [34] noted its limitations in retrieving narrowly focused scientific data.

The visibility of a species in scientific literature often reflects its conservation status, with more widely studied taxa generally receiving greater attention [39,40]. It was thus hypothesized that ChatGPT's response quality would correlate with the extent of available literature. However, in this case, no clear relationship was observed between the number of bibliographic entries on *T. contorta* in Google Scholar and the model's performance. Despite sparse references, ChatGPT successfully identified more than 60% of propagation sites within the species' natural distribution range, illustrating its efficiency in filtering and synthesizing relevant data. Species data quality often varies based on taxonomic group, conservation priority, geographic location, and presence in protected zones [3,40,41]. To better understand ChatGPT's ability to capture such data accurately, further studies should be conducted, especially regarding how these variables influence its outputs on species like *T. contorta* in Swat.

One of ChatGPT's strong suits, as evidenced in this study, was its ability to explain geographic concepts, identify species-level distinctions, and provide coding solutions (especially in R) related to spatial distribution analysis. The model correctly handled over 90% of queries in this domain, corroborating Lubiana et al. [42], who noted its effectiveness in bioinformatics, especially in script generation, debugging, and data visualization. Though ChatGPT also provided python and r scripts for species distribution modeling, our application remained limited to single points, and tools like google earth pro were used.

Five of the predicted points for *T. contorta* propagation were located outside the geographic range of the Swat district, which highlights some important limitations of using AI tools for region-specific ecological studies. One possible reason for this issue is prompt ambiguity; although the intention was to restrict predictions within Swat, the language used may not have clearly defined the district boundaries, leading the model to suggest areas beyond the target region. Another factor could be the model's limited understanding of local geographic names. Some place names in Swat may be similar to those in neighboring districts or other regions, and the model, trained primarily on general internet data, may have confused these due to a lack of detailed regional knowledge. Additionally, there may be biases in the training data itself, which tend to overrepresent well-known or frequently mentioned locations while underrepresenting smaller or less documented areas like certain parts of Swat. These limitations suggest that while AI can assist in generating useful initial insights, its outputs should be carefully reviewed and validated, especially when applied to fine-scale geographic tasks requiring local accuracy.

The inconsistency of ChatGPT's responses remains a significant challenge [43]. Variability may stem from training data diversity, question phrasing, and conversational context [44–46]. To minimize these inconsistencies, we tested identical queries across different versions (3.5 and 4.0) on multiple devices and accounts. While the wording and structure of responses varied slightly, the correctness of the information remained stable, mirroring findings from comparative model performance studies by Plevris et al. [47] and Elkhataf [48]. As Wang et al. [49] explain, ChatGPT generates responses by interpreting inputs through learned patterns, without real-time internet access. Question framing significantly affects ChatGPT's responses. Well-structured, detailed inquiries yield more accurate outputs [42], while vague prompts lead to inconsistencies [45,46]. A lack of context, such as using a new account or one-off queries, can reduce the specificity of answers [45]. As such, corroborating AI-generated data with credible sources is strongly advised.

4.3. Integrating AI- and Ground-Based Approaches for Future Conservation of *T. contorta*

Future technologies, especially artificial intelligence (AI), will enhance *T. contorta* conservation by analyzing satellite imagery to detect deforestation and guide land management [29], identifying habitats via remote sensing of altitude, soil moisture, and temperature [41], and monitoring wildlife with camera traps and acoustic sensors [39]. AI's predictive modeling will forecast climate-induced habitat shifts for resource optimization [50], while citizen science platforms will engage communities [27,39,42,51]. Emerging tools include drones for real-time tracking [52,53], climate refugia identification [29,40], illegal trade detection [40], and AI-optimized ecosystem restoration [51,54]. Challenges like data accuracy and bias require human oversight [48,50], with future research refining AI and fostering collaborations to address habitat loss and climate change [28,55–58].

4.4. Comparative Evaluation of Environmental Datasets and Modeling Approaches for Identifying Propagation Sites of *T. contorta*

The combined comparison highlights distinct strengths and limitations across environmental data sources (NASA POWER vs. WorldClim/CHELSA) and modeling approaches (AI-based ChatGPT predictions vs. MaxEnt/logistic regression models).

4.4.1. Environmental Data Platforms

NASA POWER demonstrates a clear advantage in temporal flexibility and climate dynamics. It provides high-resolution, time-sensitive environmental variables such as daily/monthly temperature, humidity, radiation, and soil moisture. This enables researchers to model real-time species–environment interactions, short-term climatic responses, and dynamic ecological patterns [25,59]. In contrast, WorldClim/CHELSA excels in spatial resolution and pre-computed ecological indices (e.g., BIOCLIM variables), making it more suitable for long-term, fine-scale habitat modeling, especially where terrain sensitivity is essential. While NASA POWER includes variables such as radiation, humidity, and soil moisture critical for physiological and ecological modeling, it lacks the pre-processed ecological indices that WorldClim/CHELSA offers. These indices, like isothermality or temperature seasonality, are often essential for niche-based species distribution modeling but are absent in raw NASA outputs and require manual derivation. Thus, integrating both platforms could offer a balanced approach: NASA POWER for temporal variation and real-time ecological dynamics, and WorldClim/CHELSA for static spatial modeling and bioclimatic trends.

4.4.2. Modeling Approaches

In under-sampled, ecologically complex regions like Swat, AI-predicted single GPS points, especially when combined with NASA POWER data, offer a more precise and practical alternative to traditional models like MaxEnt and logistic regression. MaxEnt's effectiveness declines with sparse or biased presence data and is limited by default settings, unreliable pseudo-absence assumptions, and susceptibility to spatial bias and autocorrelation. Additional issues include poor background sampling, multicollinearity, and misleading evaluation metrics [60,61]. Logistic regression also struggles in such contexts due to multicollinearity and high Type I error rates, even with advanced spatial corrections like generalized least squares with spatial correction, generalized additive mixed models, and simultaneous autoregressive models. Simpler spatial models are not well-suited for large-scale ecological applications [62–64]. In contrast, AI (ChatGPT) with NASA POWER enables habitat prediction without prior occurrence records, using ecological reasoning to generate accurate, site-specific GPS points (Table 4). This makes it a robust and flexible tool for conservation planning in data-poor, high-altitude landscapes.

Table 4. Comparative assessment of environmental data sources and modeling approaches for *T. contorta* habitat identification using single GPS points.

S. No	Aspect/Criteria	NASA POWER + AI Points	WorldClim/CHELSA + MaxEnt + Logistic Regression
1	Temporal variation	High, real-time dynamics	Low, based on long-term historical averages (BIOCLIM~1970–2000)
2	Seasonal variation	High, month-wise trends inform propagation suitability	Moderate, seasonal inference based on derived BIO layers
3	Climatic extremes	High, captures real-time extremes for stress tolerant species	Low, smoothed averages, cannot reflect recent shifts
4	Ecological indices	Moderate, AI logic-based inference from raw data	High, uses precomputed 19 BIO variables directly in models
5	Spatial resolution	Moderate, point-based GPS suggestion from AI; NASA~0.5°	High, ~1 km ² resolution in spatial raster output
6	Long term suitability	Moderate, focused on current/future trends	High, good for historical niche modeling
7	Humidity and radiation	High, NASA includes RH, SH, and solar radiation; AI evaluates importance	Low, BIOCLIM does not include radiation or real-time humidity
8	Soil and pressure	High, includes soil moisture, surface pressure, and water stress	Not available, not considered in MaxEnt or BIO layers
9	Data requirement	Low, no species presence/absence required; AI infers zones based on logic	High, needs presence (MaxEnt), or both presence/absence (LogReg) + environmental layers
10	Prediction speed	Very fast, real-time generation of points	Moderate, requires preprocessing and modeling
11	Accuracy	Varies, depends on expert logic and climate realism	High, statistically robust if trained with quality data
12	Field validation	Essential, predictions must be checked on ground	Recommended, especially in extrapolated zones
13	Scalability	High, easily extendable to new regions with little prior data	Limited, needs presence data; scale increases complexity
14	Suitability in data poor regions	Very High, works even with no prior species data	Low, fails without reliable species presence points
15	Expert knowledge integration	High, AI integrates ecological logic, habitat stress, and climate sensitivity	Low to medium, mostly automated, with limited human guided logic
16	Output type	Point based GPS coordinates for propagation	Continuous habitat suitability maps or logistic probability outputs
17	Bias and limitations	May reflect AI/human rule bias; depends on quality of integrated logic	Risk of overfitting, collinearity, or spatial bias if not properly controlled
18	Use for <i>T. contorta</i>	Ideal, identifies viable propagation points in Swat; adapts to ecological niche logic	Good, works where enough presence data exist and environment is modeled accurately

4.5. Considerations, Advantages, and Disadvantages of Using ChatGPT

The advent of the Internet has revolutionized access to specialized scientific knowledge, dramatically increasing the availability of resources. A simple query in an academic database can yield hundreds or even thousands of relevant results. While this wealth of information is beneficial, it can also overwhelm researchers, making it difficult to locate and synthesize relevant data efficiently [50]. This is where tools like ChatGPT become particularly valuable. Recognized for its ability to gather, filter, and summarize information swiftly, ChatGPT offers a practical solution for managing large volumes of content [65].

Despite these benefits, one of the primary criticisms directed at ChatGPT is its lack of proper source citation. Neglecting to cite the original authors risks unethical use of information and may lead to academic misconduct [33]. Additionally, concerns have been raised over the chatbot's tendency to fabricate bibliographic references or misattribute authorship and publication years, which can mislead users relying on its output [66]. Additionally, plagiarism detection tools may struggle to differentiate between AI-generated and human-authored content, further complicating scientific integrity.

In the context of this research on *T. contorta* and the identification of its suitable habitats, ChatGPT demonstrated considerable utility. It provided accurate and well-organized responses, proving adept at distilling complex and scattered data. Particularly when it comes to obtaining quick and precise insights on species characteristics and terminologies relevant to the study, the tool significantly reduced the time and effort required to be compared to conventional search engines. AI-based systems like ChatGPT also show promise in ecological studies, offering assistance in areas such as text simplification, data analysis, decision support, and even modeling ecological phenomena [35,53,55].

One of the major concerns in conservation science applications is the accuracy of the data generated by ChatGPT. The model is trained on vast datasets sourced from the internet, which often contain inaccuracies, biases, and outdated information [67,68]. Since conservation efforts rely heavily on precise ecological data, erroneous outputs can misguide decision-making processes and resource allocation. Moreover, the problem of bias in AI models is significant. As Rodas-Trejo and Ocampo-González [50] discuss, large language models are prone to replicating biases present in their training data. This can result in skewed interpretations of conservation policies, species distributions, and climate change effects, ultimately impacting conservation strategies negatively. The model's inability to fully grasp complex ecological relationships and conservation frameworks can result in oversimplified or irrelevant responses [69]. Liu et al. [70] emphasize that despite advancements in AI capabilities, there remains a substantial gap in contextual intelligence when addressing specialized conservation queries. Conservation efforts often require localized and region specific data, which ChatGPT might not be able to provide accurately due to its generalized training [71]. Conservationists in regions with limited digital representation may find AI-generated outputs lacking in relevance and cultural sensitivity. As Carlini et al. [72] and Nasr et al. [73] highlight, AI models are susceptible to data extraction and privacy risks. Conservation projects often deal with sensitive ecological and geospatial data, and the inadvertent exposure of such information through AI tools poses security challenges. Conservationists should employ ChatGPT as a supplementary tool rather than a primary decision making resource, ensuring that AI-generated insights are validated through peer-reviewed sources and expert consultations [74]. Moreover, improving transparency in AI model training and enhancing domain-specific datasets can help bridge knowledge gaps and reduce inaccuracies.

Despite its potential, the application of ChatGPT in conservation sciences is hindered by challenges related to data accuracy, bias, contextual limitations, and ethical considerations. Addressing these limitations through a combination of AI advancements and human expertise is crucial to leveraging AI tools effectively in the pursuit of sustainable conservation efforts.

5. Conclusions

The results demonstrate the effectiveness of AI, specifically ChatGPT, as a complementary tool in the conservation planning of *T. contorta*, a rare, threatened, and medicinally significant species endemic to the Western Himalayas. By integrating AI-driven analysis with ecological variables, we were able to identify and validate suitable propagation sites

in the Swat district with a high degree of accuracy. Among the 63 AI-identified geospatial points, 58 were located within Swat, and all of them fell within the range of previously established species distribution models. Furthermore, 37 sites were located within the optimal altitudinal band of 1800–3100 m a.s.l., and field validation confirmed 16 new populations, underscoring the practical utility of ChatGPT-assisted site selection. This illustrates that, despite inherent limitations such as lack of potential inaccuracies and static knowledge, ChatGPT can significantly reduce the time and effort required for data collection and analysis in ecological studies. The findings highlight the potential of AI-driven analysis in conservation strategies, offering a scalable approach for identifying critical habitats and providing suggestions for habitat restoration.

However, caution must be exercised when interpreting AI-generated outputs. The dependence on pre-trained datasets and the possibility of incorrect or outdated information necessitate cross-verification with peer-reviewed sources and expert consultation. Ethical considerations, including the proper citation of original data sources and researchers, are also crucial.

The success of this study highlights the value of integrating AI tools into biodiversity conservation, especially in resource-limited and climate-vulnerable regions. Collaboration between conservation scientists and AI developers is essential to enhance model accuracy, ensure up-to-date knowledge integration, and broaden application across other threatened taxa. Future research may benefit from exploring advanced versions of AI models, such as ChatGPT-5, or hybrid systems with real-time internet capabilities for improved performance and reliability. Ultimately, this research establishes a foundational model for AI-supported conservation of endangered plant species and offers a replicable, scalable framework that can inform future efforts in habitat identification, resource prioritization, and ecological monitoring.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/su17198541/s1>, GPS points of both actual populations and AI-predicted points along with environmental and ecological variables are available in excel sheets with the publisher.

Author Contributions: Conceptualization, S.D., H.A. and T.P.; methodology, S.D., H.A. and T.P.; software, S.D.; validation, S.D. and H.A.; investigation: S.D., H.A., J.A. and S.M.; resources, S.D. and H.A.; data curation, S.D., H.A., H.S. and T.P.; writing—original draft preparation, S.D., H.A. and T.P.; writing—review and editing, S.D., H.A. and T.P.; visualization, S.D., H.A. and T.P.; supervision, H.A.; project administration, S.D. and H.A.; funding acquisition, H.A. All authors have read and agreed to the published version of the manuscript.

Funding: Research was performed under HEC-NRPU funded project No. 7303.

Institutional Review Board Statement: The study was conducted in compliance with national and regional regulations for the protection of endangered species. As the plant is endangered, branch sampling for propagation was conducted under strict guidelines, collecting no more than 20 twigs of 20 cm length from any single plant with a diameter greater than 150 cm and a crown cover exceeding 6 m².

Informed Consent Statement: Informed consent was obtained from all individuals involved in the study. Participants were provided with detailed information regarding the objectives, procedures, potential risks, and benefits of the research. They were informed of their right to withdraw at any time without penalty, and their participation was entirely voluntary. All personal data were handled confidentially and used solely for research purposes.

Data Availability Statement: The original contributions presented in this study are included in the article/Supplementary Materials. Further inquiries can be directed to the corresponding authors.

Conflicts of Interest: The authors declare no conflicts of interest.

References

- Rabinowitz, D. Seven forms of rarity. In *Biological Aspects of Rare Plant Conservation*; John Wiley & Sons: Hoboken, NY, USA, 1981.
- Thomas, P. A review of the distribution and conservation status of *Taxus* in the Himalayas, China and Southeast Asia. In *III Jornadas Internacionales Sobre el Tejo (Taxus baccata L.)*; Association Amigos del Tejo y las Tejedas: Madrid, Spain, 2011; pp. 7–274. [[CrossRef](#)]
- Shah, A.; Li, D.Z.; Möller, M.; Gao, L.M.; Hollingsworth, M.L.; Gibby, M. Delimitation of *Taxus fuana* Nan Li & RR Mill (Taxaceae) based on morphological and molecular data. *Taxon* **2008**, *57*, 211–222. [[CrossRef](#)]
- List, I.R. The IUCN Red List of Threatened Species. Available online: <https://www.iucnredlist.org/> (accessed on 18 August 2025).
- Möller, M.; Liu, J.; Li, Y.; Li, J.-H.; Ye, L.-J.; Mill, R.; Thomas, P.; Li, D.-Z.; Gao, L.-M. Repeated intercontinental migrations and recurring hybridizations characterise the evolutionary history of yew (*Taxus* L.). *Mol. Phylogenet. Evol.* **2020**, *153*, 106952. [[CrossRef](#)] [[PubMed](#)]
- Ali, H.A.; Ahmad, H.; Marwat, K.B.; Yousaf, M.; Gul, B.; Khan, I. Trade potential and conservation issues of medicinal plants in District Swat, Pakistan. *Pak. J. Bot.* **2012**, *44*, 1905–1912.
- Iqbal, J.; Meilan, R.; Khan, B. Assessment of risk, extinction, and threats to Himalayan yew in Pakistan. *Saudi J. Biol. Sci.* **2020**, *27*, 762–767. [[CrossRef](#)] [[PubMed](#)]
- Gillani, S.W.; Ahmad, M.; Zafar, M.; Haq, S.M.; Waheed, M.; Manzoor, M.; Shaheen, H.; Sultana, S.; Rehman, F.U.; Makhkamov, T. An insight into indigenous ethnobotanical knowledge of medicinal and aromatic plants from Kashmir Himalayan region. *Ethnobot. Res. Appl.* **2024**, *28*, 1–21. [[CrossRef](#)]
- Kumar, B.; Singh, K.; Sharma, J.; Gairola, S. A comprehensive review of fuelwood resources and their use pattern in rural villages of Western Himalaya, India. *Plant Arch.* **2020**, *20*, 1949–1958.
- Dhyani, S. Are Himalayan ecosystems facing hidden collapse? Assessing the drivers and impacts of change to aid conservation, restoration and conflict resolution challenges. *Biodivers. Conserv.* **2023**, *32*, 3731–3764. [[CrossRef](#)]
- Shaheen, H.; Attique, A.; Riaz, M.T.; Manzoor, M.; Khan, R.W.A.; Riaz, M.T. From biodiversity hotspot to conservation hotspot: Assessing distribution, population structure, associated flora and habitat geography of threatened Himalayan Yew in temperate forest ecosystems of Kashmir. *Biodivers. Conserv.* **2024**, *33*, 553–577. [[CrossRef](#)]
- Bangelesa, F.; Abel, D.; Pollinger, F.; Rai, P.; Ziegler, K.; Ebengo, D.; Tshimanga, R.M.; Ali, M.M.; Knight, J.; Paeth, H. Projected changes in rainfall amount and distribution in the Democratic Republic of Congo—Evidence from an ensemble of high-resolution climate simulations. *Weather Clim. Extrem.* **2023**, *42*, 100620. [[CrossRef](#)]
- Tiwari, O.P.; Sharma, C.M. Anthropogenic disturbance impact on forest composition and dominance-diversity: A case of an ecosensitive region of Garhwal Himalaya, India. *Acta Ecol. Sin.* **2023**, *43*, 662–673. [[CrossRef](#)]
- Poudel, R.C.; Möller, M.; Gao, L.-M.; Ahrends, A.; Baral, S.R.; Liu, J.; Thomas, P.; Li, D.-Z. Using morphological, molecular and climatic data to delimitate yews along the Hindu Kush-Himalaya and adjacent regions. *PLoS ONE* **2012**, *7*, e46873. [[CrossRef](#)]
- Bhardwaj, V. *Taxus wallichiana* Zucc. (Himalayan yew): A medicinal plant exhibiting antibacterial properties. In *Advances in Microbiology, Infectious Diseases and Public Health*; Springer International Publishing: Cham, Switzerland, 2023; Volume 17, pp. 145–153. [[CrossRef](#)]
- Gajurel, J.P.; Werth, S.; Shrestha, K.K.; Scheidegger, C. Species distribution modeling of *Taxus wallichiana* (Himalayan yew) in Nepal Himalaya. *Asian J. Conserv. Biol.* **2014**, *3*, 127–134.
- Pandey, A.; Chandra Sekar, K.; Joshi, B.; Rawal, R.S. Threat assessment of high-value medicinal plants of cold desert areas in Johar valley, Kailash Sacred Landscape, India. *Plant Biosyst. Int. J. Deal. All Asp. Plant Biol.* **2019**, *153*, 39–47. [[CrossRef](#)]
- Paul, A.; Bharali, S.; Khan, M.L.; Tripathi, O.P. Anthropogenic disturbances led to risk of extinction of *Taxus wallichiana* Zuccarini, an endangered medicinal tree in Arunachal Himalaya. *Nat. Areas J.* **2013**, *33*, 447–454. [[CrossRef](#)]
- Bennett, E.M.; Peterson, G.D.; Gordon, L.J. Understanding relationships among multiple ecosystem services. *Ecol. Lett.* **2009**, *12*, 1394–1404. [[CrossRef](#)] [[PubMed](#)]
- Mehta, P.; Sekar, K.C.; Bhatt, D.; Tewari, A.; Bisht, K.; Upadhyay, S.; Negi, V.S.; Soragi, B. Conservation and prioritization of threatened plants in Indian Himalayan Region. *Biodivers. Conserv.* **2020**, *29*, 1723–1745. [[CrossRef](#)]
- Ahmad, S.S.; Abbasi, Q.; Jabeen, R.; Shah, M.T. Decline of conifer forest cover in Pakistan: A GIS approach. *Pak. J. Bot.* **2012**, *44*, 511–514.
- Abdaki, M.; Al-Ozeer, A.Z.; Alobaydy, O.; Al-Tayawi, A.N. Predicting rainfall in Nineveh Governorate in northern Iraq using machine learning time-series forecasting algorithm. *Arab. J. Geosci.* **2023**, *16*, 655. [[CrossRef](#)]
- Abduljaleel, Y.; Chikabvumbwa, S.R.; Haq, F.U. Evaluation and prediction of future droughts with multi-model ensembling of four models under CMIP6 scenarios over Iraq. *Theor. Appl. Climatol.* **2024**, *155*, 131–142. [[CrossRef](#)]
- Abed, S.A.; Halder, B.; Yaseen, Z.M. Investigation of the decadal unplanned urban expansion influenced surface urban heat island study in the Mosul metropolis. *Urban Clim.* **2024**, *54*, 101845. [[CrossRef](#)]

25. Aboelkhair, H.; Morsy, M.; El Afandi, G. Assessment of agroclimatology NASA POWER reanalysis datasets for temperature types and relative humidity at 2 m against ground observations over Egypt. *Adv. Space Res.* **2019**, *64*, 129–142. [[CrossRef](#)]
26. Agyekum, J.; Annor, T.; Quansah, E.; Lamptey, B.; Okafor, G. Extreme precipitation indices over the Volta Basin: CMIP6 model evaluation. *Sci. Afr.* **2022**, *16*, e01181. [[CrossRef](#)]
27. Agathokleous, E.; Saitanis, C.J.; Fang, C.; Yu, Z. Use of ChatGPT: What does it mean for biology and environmental science? *Sci. Total Environ.* **2023**, *888*, 164154. [[CrossRef](#)]
28. Biswas, S.S. Potential use of ChatGPT in global warming. *Ann. Biomed. Eng.* **2023**, *51*, 1126–1127. [[CrossRef](#)] [[PubMed](#)]
29. Powers, R.P.; Jetz, W. Global habitat loss and extinction risk of terrestrial vertebrates under future land-use-change scenarios. *Nat. Clim. Chang.* **2019**, *9*, 323–329. [[CrossRef](#)]
30. Rahman, M.M.; Watanobe, Y. ChatGPT for education and research: Opportunities, threats, and strategies. *Appl. Sci.* **2023**, *13*, 5783. [[CrossRef](#)]
31. Ali, S.; Ali, H.; Baharanchi, O.G.; Sher, H.; Yousefpour, R. Investigating endemic species conservation hotspots based on species distribution models in Swat Region, Hindu Kush Pakistan. *Land* **2024**, *13*, 737. [[CrossRef](#)]
32. Poudel, R.C.; Möller, M.; Li, D.Z.; Shah, A.; Gao, L.M. Genetic diversity, demographical history and conservation aspects of the endangered yew tree *Taxus contorta* (s yn. *Taxus fuana*) in Pakistan. *Tree Genet. Genomes* **2014**, *10*, 653–665. [[CrossRef](#)]
33. Frosolini, A.; Gennaro, P.; Cascino, F.; Gabriele, G. In reference to “Role of Chat GPT in public health”, to highlight the AI’s incorrect reference generation. *Ann. Biomed. Eng.* **2023**, *51*, 2120–2122. [[CrossRef](#)]
34. Mehnen, L.; Gruarin, S.; Vasileva, M.; Knapp, B. ChatGPT as a medical doctor? A diagnostic accuracy study on common and rare diseases. *medRxiv* **2023**. [[CrossRef](#)]
35. Morera, A. Foundation models in shaping the future of ecology. *Ecol. Inform.* **2024**, *80*, 102545. [[CrossRef](#)]
36. Rathore, P.; Roy, A.; Karnatak, H. Modelling the vulnerability of *Taxus wallichiana* to climate change scenarios in South East Asia. *Ecol. Indic.* **2019**, *102*, 199–207. [[CrossRef](#)]
37. Crowther, T.W.; Glick, H.B.; Covey, K.R.; Bettigole, C.; Maynard, D.S.; Thomas, S.M.; Smith, J.R.; Hintler, G.; Duguid, M.C.; Amatulli, G.; et al. Mapping tree density at a global scale. *Nature* **2015**, *525*, 201–205. [[CrossRef](#)]
38. Fatani, B. ChatGPT for future medical and dental research. *Cureus* **2023**, *15*, e37285. [[CrossRef](#)] [[PubMed](#)]
39. Bird, J.P.; Martin, R.; Akçakaya, H.R.; Gilroy, J.; Burfield, I.J.; Garnett, S.T.; Symes, A.; Taylor, J.; Şekercioğlu, Ç.H.; Butchart, S.H.M. Generation lengths of the world’s birds and their implications for extinction risk. *Conserv. Biol.* **2020**, *34*, 1252–1261. [[CrossRef](#)] [[PubMed](#)]
40. Cazalis, V.; Santini, L.; Lucas, P.M.; González-Suárez, M.; Hoffmann, M.; Benítez-López, A.; Pacifici, M.; Schipper, A.M.; Böhm, M.; Zizka, A.; et al. Prioritizing the reassessment of data-deficient species on the IUCN Red List. *Conserv. Biol.* **2023**, *37*, e14139. [[CrossRef](#)] [[PubMed](#)]
41. Boakes, E.H.; McGowan, P.J.; Fuller, R.A.; Chang-qing, D.; Clark, N.E.; O’Connor, K.; Mace, G.M. Distorted views of biodiversity: Spatial and temporal bias in species occurrence data. *PLoS Biol.* **2010**, *8*, e1000385. [[CrossRef](#)]
42. Lubiana, T.; Lopes, R.; Medeiros, P.; Silva, J.C.; Goncalves, A.N.A.; Maracaja-Coutinho, V.; Nakaya, H.I. Ten quick tips for harnessing the power of ChatGPT in computational biology. *PLoS Comput. Biol.* **2023**, *19*, e1011319. [[CrossRef](#)]
43. Liang, J.; Wang, Z.; Ma, Z.; Li, J.; Zhang, Z.; Wu, X.; Wang, B. Online training of large language models: Learn while chatting. *arXiv* **2024**, arXiv:2403.04790. [[CrossRef](#)]
44. Dhanvijay, A.K.D.; Pinjar, M.J.; Dhokane, N.; Sorte, S.R.; Kumari, A.; Mondal, H. Performance of large language models (ChatGPT, Bing Search, and Google Bard) in solving case vignettes in physiology. *Cureus* **2023**, *15*, e42972. [[CrossRef](#)]
45. Nazir, A.; Wang, Z. A comprehensive survey of ChatGPT: Advancements, applications, prospects, and challenges. *Metaradiology* **2023**, *1*, 100022. [[CrossRef](#)]
46. Ray, P.P. ChatGPT: A comprehensive review on background, applications, key challenges, bias, ethics, limitations and future scope. *Internet Things Cyber-Phys. Syst.* **2023**, *3*, 121–154. [[CrossRef](#)]
47. Plevris, V.; Papazafeiropoulos, G.; Jiménez Rios, A. Chatbots put to the test in math and logic problems: A comparison and assessment of ChatGPT-3.5, ChatGPT-4, and Google Bard. *Ai* **2023**, *4*, 949–969. [[CrossRef](#)]
48. Elkhatat, A.M. Evaluating the authenticity of ChatGPT responses: A study on text-matching capabilities. *Int. J. Educ. Integr.* **2023**, *19*, 1–23. [[CrossRef](#)]
49. Wang, H.; Li, J.; Wu, H.; Hovy, E.; Sun, Y. Pre-trained language models and their applications. *Engineering* **2023**, *25*, 51–65. [[CrossRef](#)]
50. Rodas-Trejo, J.; Ocampo-González, P. Assessment of ChatGPT’s potential as an innovative tool in searching for information on wild mammals. *Ecol. Inform.* **2024**, *83*, 102810. [[CrossRef](#)]
51. Zhai, X. ChatGPT for next generation science learning. *XRDS Crossroads ACM Mag. Stud.* **2023**, *29*, 42–46. [[CrossRef](#)]
52. Leorna, S.; Brinkman, T. Human vs. machine: Detecting wildlife in camera trap images. *Ecol. Inform.* **2022**, *72*, 101876. [[CrossRef](#)]
53. Reyhani Haghighi, S.; Pasandideh Saqalaksari, M.; Johnson, S.N. Artificial intelligence in ecology: A commentary on a chatbot’s perspective. *Bull. Ecol. Soc. Am.* **2023**, *104*, e2097. [[CrossRef](#)]

54. Donaldson, M.R.; Burnett, N.J.; Braun, D.C.; Suski, C.D.; Hinch, S.G.; Cooke, S.J.; Kerr, J.T. Taxonomic bias and international biodiversity conservation research. *Facets* **2016**, *1*, 105–113. [[CrossRef](#)]
55. Byeon, J.H.; Kwon, Y.J. The Effect of Outdoor Inquiry Program for Learning Biology Using Digital Twin Technology. *J. Balt. Sci. Educ.* **2023**, *22*, 781–798. [[CrossRef](#)]
56. Jia, X.; Feng, S.; Zhang, H.; Liu, X. Plastome phylogenomics provide insight into the evolution of *Taxus*. *Forests* **2022**, *13*, 1590. [[CrossRef](#)]
57. Li, P.; Zhu, W.; Xie, Z.; Qiao, K. Integration of multiple climate models to predict range shifts and identify management priorities of the endangered *Taxus wallichiana* in the Himalaya-Hengduan Mountain region. *J. For. Res.* **2020**, *31*, 2255–2272. [[CrossRef](#)]
58. Hu, Z.; Liu, S.; Zhong, G.; Lin, H.; Zhou, Z. Modified Mann-Kendall trend test for hydrological time series under the scaling hypothesis and its application. *Hydrol. Sci. J.* **2020**, *65*, 2419–2438. [[CrossRef](#)]
59. Duarte, Y.C.; Sentelhas, P.C. NASA/POWER and DailyGridded weather datasets—How good they are for estimating maize yields in Brazil? *Int. J. Biometeorol.* **2020**, *64*, 319–329. [[CrossRef](#)] [[PubMed](#)]
60. Shcheglovitova, M.; Anderson, R.P. Estimating optimal complexity for ecological niche models: A jackknife approach for species with small sample sizes. *Ecol. Model.* **2013**, *269*, 9–17. [[CrossRef](#)]
61. Syfert, M.M.; Smith, M.J.; Coomes, D.A. The effects of sampling bias and model complexity on the predictive performance of MaxEnt species distribution models. *PLoS ONE* **2013**, *8*, e55158. [[CrossRef](#)]
62. Augustin, N.H.; Muggleston, M.A.; Buckland, S.T. An autologistic model for the spatial distribution of wildlife. *J. Appl. Ecol.* **1996**, *33*, 339–347. [[CrossRef](#)]
63. Brunsdon, C.; Fotheringham, S.; Charlton, M. Geographically weighted regression. *J. R. Stat. Soc. Ser. D (Stat.)* **1998**, *47*, 431–443. [[CrossRef](#)]
64. Betts, M.G.; Ganio, L.M.; Huso, M.M.; Som, N.A.; Huettmann, F.; Bowman, J.; Wintle, B.A. Comment on “Methods to account for spatial autocorrelation in the analysis of species distributional data: A review”. *Ecography* **2009**, *32*, 374–378. [[CrossRef](#)]
65. Magare, D.A.; Patil, M.B. Ethnobotany in the digital age: Opportunities and challenges of traditional knowledge digitization. *World J. Biol. Pharm. Health Sci.* **2025**, *21*, 235–242. [[CrossRef](#)]
66. Domazetoski, V. Enhancing Ecological Knowledge Discovery Using Large Language Models. Master’s Thesis, Georg-August Universität Göttingen, Göttingen, Germany, 2024; pp. 1–71.
67. Lee, N.; Ping, W.; Xu, P.; Patwary, M.; Fung, P.N.; Shoeybi, M.; Catanzaro, B. Factuality enhanced language models for open-ended text generation. *Adv. Neural Inf. Process. Syst.* **2022**, *35*, 34586–34599. [[CrossRef](#)]
68. Zhang, Y.; Warstadt, A.; Li, H.S.; Bowman, S.R. When do you need billions of words of pretraining data? *arXiv* **2020**, arXiv:2011.04946. [[CrossRef](#)]
69. Zuccon, G.; Koopman, B. Dr ChatGPT, tell me what I want to hear: How prompt knowledge impacts health answer correctness. *arXiv* **2023**, arXiv:2302.13793. [[CrossRef](#)]
70. Liu, J.; Liu, C.; Zhou, P.; Lv, R.; Zhou, K.; Zhang, Y. Is chatgpt a good recommender? a preliminary study. *arXiv* **2023**, arXiv:2304.10149. [[CrossRef](#)]
71. Zhang, H.; Wu, C.; Xie, J.; Lyu, Y.; Cai, J.; Carroll, J.M. Redefining qualitative analysis in the AI era: Utilizing ChatGPT for efficient thematic analysis. *arXiv* **2023**, arXiv:2309.10771. [[CrossRef](#)]
72. Carlini, N.; Tramer, F.; Wallace, E.; Jagielski, M.; Herbert-Voss, A.; Lee, K.; Roberts, A.; Song, D. Extracting training data from large language models. In Proceedings of the 30th USENIX Security Symposium (USENIX Security 21), Vancouver, BC, Canada, 11–13 August 2021; pp. 2633–2650. [[CrossRef](#)]
73. Nasr, M.; Shokri, R.; Houmansadr, A. Comprehensive privacy analysis of deep learning: Passive and active white-box inference attacks against centralized and federated learning. In *2019 IEEE Symposium on Security and Privacy (SP)*; IEEE: Piscataway, NJ, USA, 2019; pp. 739–753. [[CrossRef](#)]
74. Terwiesch, C. *Would Chat GPT3 Get a Wharton MBA? A Prediction Based on Its Performance in the Operations Management Course*; Mack Institute for Innovation Management at the Wharton School, University of Pennsylvania: Philadelphia, PA, USA, 2023.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.