

Research

Ecological niche modelling of *Indigofera oblongifolia* (Forssk.): a global machine learning assessment using climatic and non-climatic predictors

Manish Mathur¹ · Preet Mathur²

Received: 1 October 2023 / Accepted: 19 January 2024

Published online: 19 February 2024

© The Author(s) 2024 [OPEN](#)

Abstract

Climate change and other extinction facilitators have caused significant shifts in the distribution patterns of many species during the past few decades. Restoring and protecting lesser-known species may be more challenging without adequate biogeographical information. To address this knowledge gap, the current study set out to determine the global spatial distribution patterns of *Indigofera oblongifolia* (Forssk.) a relatively lesser-known leguminous species. This was accomplished by utilizing three distinct bioclimatic temporal frames (current, 2050, and 2070) and four greenhouse gas scenarios (RCPs 2.6, 4.5, 6.0, and 8.5), in addition to non-climatic predictors such as global livestock population, human modification of terrestrial ecosystems, and global fertilizers application (nitrogen and phosphorus). Furthermore, we evaluate the degree of indigenesness using the geographical area, habitat suitability categories, and number of polygons. This research reveals that climatic predictors outperform non-climatic predictors in terms of improving model quality. Precipitation Seasonality is one of the most important factors influencing this species' optimum habitat suitability up to 150 mm for the current, 2050 RCP 8.5 and 2070-RCPs 2.6, 4.5, and 8.5. Our ellipsoid niche modelling extends the range of precipitation during the wettest quarter and maximum temperature during the warmest month to 637 mm and 26.5–31.80 degrees Celsius, respectively. India has a higher indigenous score in the optimal class than the African region. This findings suggest that the species in question tends to occupy contiguous regions in Africa, while in India, it is dispersed into several smaller meta-populations.

Keywords Ecological niche modelling · Ensemble techniques · *Indigofera oblongifolia* · Legume palatable species · Livestock population · Niche overlap · Random forest

1 Introduction

The phenomenon of climate change plays a pivotal role in shaping the redistribution of species' ranges. It can be posited that alterations in climatic conditions may potentially engender the expansion or contraction of the geographic ranges of said species [1]. Compounding the concerns at hand are the anticipated ramifications of the phenomenon of global warming, which are projected to intensify over the course of the ensuing half-century [2]. Numerous research studies have indicated

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s44274-024-00029-1>.

✉ Preet Mathur, preetm9535@gmail.com; Manish Mathur, eco5320@gmail.com | ¹ICAR-Central Arid Zone Research Institute, 342 003 Jodhpur, India. ²Jodhpur Institute of Engineering and Technology, Jodhpur, India.



that contemporary ecological changes, which are caused by human activities, have led to the displacement of species [3], alterations in phenology [4], and the extinction of certain species [5]. The task of reviving and safeguarding of lesser-known species could prove to be a more arduous undertaking in the absence of sufficient biogeographic data [6, 7]. The integration of biogeography principles, techniques, and evaluations into the realm of conservation biology has significantly bolstered endeavours aimed at preserving the habitats of these particular species.

To comprehend the present and forthcoming path of population dynamics, it is crucial to integrate computational methodologies into predicting the ramifications of climate change on the accessibility of appropriate habitats for a particular species. The utilization of bioclimatic modelling tools has become progressively prevalent in the projection of forthcoming species distributions. The utilization of species distribution models (SDMs) or Ecological Niche Modelling ENM) have become ubiquitous in the scientific community as a means of forecasting the historical, current, and prospective habitat suitability of species through the application of diverse mathematical algorithms [8]. The efficacy of SDMs is predicated upon the interdependence between an organism and its corresponding geographic location, with contemporary models incorporating additional variables such as land-use classifications [9, 10].

Regression and machine learning are two major areas of Species Distribution Modelling (SDM) that rely on presence-absence data [11]. Constraints inherent to any model reduce their ability to anticipate [12]. An over-adjusted model has a lower likelihood of success when used with new data [13]. In contemporary times, novel modelling algorithms have been proposed that advocate the utilization of ensemble modelling. This approach facilitates the identification of the common ground between diverse models, thereby enabling the advantages of employing multiple algorithms while simultaneously revealing uncertainty to the end-users. This has been elucidated in various studies [6, 11, 14–20].

Indigofera oblongifolia is a valuable but underutilized multipurpose leguminous shrub that has traditionally been used for foraging, medicine, herbal tooth-brushing and restoration of mining degraded lands [21–23]. It is a non-thorny erect arid shrub found in open, dry areas with stable sandy soils. It is widely available in Jordan, Yemen, Baherien, Eritrea, Somalia, Egypt, Sudan, Senegal, Angola, Nigeria, Arabia, Baluchistan, Pakistan, Java, Sri Lanka, and India [24, 25], as well as Australia, North and South America [26]. This species can be flourish under the wild conditions (Fig. 1a), under agrarian settings (Fig. 1b) having small flower with distinct petal (Fig. 1c) with small pods (Fig. 1d). Despite the extensive range of ecosystem services and wide niche breadth exhibited by this species in hot arid and semi-arid regions across the globe, a dearth of scientific inquiry has been conducted to model its ecological niche utilizing predictors such as present and future bio-climatic changes, as well as non-climatic predictors like the impact of human activity on terrestrial ecosystems and soil nutrients, specifically nitrogen and phosphorus, from a global vantage point.

Therefore, we used global presence records of *I. oblongifolia* shrub to achieve the following goals: (a) identify the global geographical distribution patterns of this species with climatic change projections across three bio-climatic timeframes (current, 2050, and 2070) and four greenhouse gas scenarios (RCPs 2.6, 4.5, 6.0 and 8.5), (b) to evaluate the effects of the existing global livestock population (cows, goats, and sheep), global fertilizers application (nitrogen and phosphorus), and the impact of anthropogenic variables in terms of human modification of the terrestrial ecosystem on the habitat suitability (c) to evaluate the degree of indigenous (percent indigeneity) by using the geographical area, habitat suitability categories, and number of polygons identified in the preceding first two steps, (d) to quantify the temporal effects of different bio-climatic variables on its fundamental and realized niche. The aforementioned objectives shall serve as a guiding principle in the estimation of the actual area that is conducive to the survival of this species across various habitat classifications. Additionally, the identification of factors and thresholds that influence the current and future distribution patterns of this species shall be undertaken. Consequently, it shall enable us to determine the feasibility of executing comprehensive captive plantations to ensure the perpetual survival of the species and perpetuate the ecological advantages it presently confers.

2 Material and methods

2.1 Data collections

Distributional records for this species were obtained from data repositories such as the Global Biodiversity Information Facility GBIF <https://doi.org/https://doi.org/10.15468/dl.7mf5jt> [27], the Indian Biodiversity Portal (<https://indiabiodiversity.org/species/show/279391>), and published literature [23, 28, 29] and our field work during 2005 to 2014 at various districts of arid and semi-arid areas of Rajasthan, India [30, 31]. The coordinates of these sites were identified on a WGS84 coordinate system with GIS ArcMap [32] software. In order to lessen the effects of spatial autocorrelation and eliminate redundant entries, we employed the spatial thin window of the “Wallace Software”, a Graphical User Interface based on



Fig. 1 *Indigofera oblongifolia*: under wild condition (a); under agrarian condition (b); at flowering stage (c) and at fruiting stage (d)

the R programming language [33], with a thinning distance of 10 km. Utilizing the spatially dispersed presence points of said species, we have ascertained its present IUCN classification, with particular emphasis on its Area of Occupancy (AOO) and Extent of Occurrence (EOO) through employment of the R-based “ConR” programme.

2.2 Bio-climatic (BC) and non-bioclimatic variables

WorldClim version 2.0 [34] provided the observational data used to predict the present and future distributions of species. 19 bioclimatic variables [35] were downloaded at spatial resolution of 30 s ($\sim 1 \text{ km}^2$) and converted to ASCII (or ESRI ASCII) in DIVA-GIS version 7.5 [32] for current as well as two future climatic scenarios (2050-time frame that represents the mean values from 2041 to 2060 and 2070-time frame that represents the mean values from 2061 to 2080 [36]). These datasets pertain to four Representative Concentration Pathways (RCPs), namely RCP 2.6, RCP 4.5, RCP 6.0 and RCP 8.5. RCPs are the spatial and temporal trajectories of future greenhouse gas concentrations and pollutants induced by various human activities. The quartet of Representative Concentration Pathways spans a spectrum from exceedingly elevated (RCP 8.5) to exceedingly diminished (RCP 2.6) forthcoming concentrations. The pathway with the most minimal concentration of greenhouse gases (GHG) is denoted by RCPs 2.6, which is characterized by aggressive mitigation and the lowest emissions. Conversely, RCP 4.5 and RCP 6.0 are classified as intermediate, while RCP 8.5 represents the maximum emission scenario. GHG concentration pathways in which radioactive forcing (global energy imbalance) stabilizes 2.6 W/m^2 , 4.5 W/m^2 , 6.0 W/m^2 and 8.5 W/m^2 , respectively, [37]. Details of each bio-climatic parameter, along-with their units and mathematical expressions see Table S1.

2.3 The global human modification of terrestrial systems (GHMTS)

The Global Human Modification of Terrestrial Systems offers a comprehensive depiction of the worldwide impact of human activity on all landmasses, with the exception of Antarctica, at a resolution of 1 square kilometre. The metric in question is a continuous scale that ranges from 0 to 1, and serves to quantify the degree to which a given landscape has undergone modification. This is achieved through the use of sophisticated modelling techniques that take into account the physical extents of 13 distinct stressors that are attributable to human activity. These stressors are evaluated using

spatially explicit global data sets that are current as of the median year of 2016. Detailed description of the methodology is provided by Kennedy et al. [38, 39]. For present study, the global GHMTS data set in Geo Tiff format was downloaded from Socioeconomic Data and Applications Center (SEDAC) <https://sedac.ciesin.columbia.edu/data/set/lulc-human-modification-terrestrial-systems/data-download> with the intention of evaluating the effects of these human-caused changes on the species' distribution dynamics.

2.4 Global fertilizer (nitrogen and phosphorus and manure production) and livestock population

The global mapping of nitrogen and phosphorus fertilizer application rates (kilograms/hectare), as documented by Potter et al. [40], provides a valuable resource for the determination of nutrient budgets, modelling the impact of nutrient management on crop growth and yield, and tracing the trajectory of nutrients from their extraction and processing from minerals to their eventual application in the agricultural sector. The nutrient application rates for nitrogen (N) and phosphorus (P) in both fertilizer and manure production were obtained from the dataset disseminated by the Socioeconomic Data and Applications Center (SEDAC). <http://sedac.ciesin.columbia.edu/data/collection/fertilizer-and-manure.html>. In these datasets the mean quantity of nitrogen and phosphorus fertilizer administered across all crops situated within the 0.5-degree grid cell. The values of the grid cells are quantified in units of kilograms per hectare (kg/ha), with a range spanning from 0 to 370. In the current study, the aim is to assess the worldwide dispersion tendencies of said species in correlation to the global prevalence of fertilization practices and the generation of organic waste.

2.5 Grided livestock of the World-2015

The dataverse encompasses a spatial dataset ($\sim 1 \text{ km}^2$) pertaining to the distribution of livestock, which holds potential for enhancing development, environmental quality, and health outcomes. The methodology employed in the present study involved the retrieval of sheep, goat, and cattle population densities from the online database <https://data.apps.fao.org/catalog/dataset/glw> (January, 2023) at 5 min arc, in accordance with the approach outlined by Gilbert et al. [41].

2.6 Issue of multicollinearity

The present study utilized the Pearson Correlation Coefficient (r) to scrutinize cross-correlation, and a multicollinearity examination was executed to evaluate the potentiality of over-fitting. In addition, we adhered to the methodology outlined by [42] to eliminate variables that displayed cross correlation coefficient values that were equal to or greater than 0.85. This was accomplished through the utilization of the Niche Tool Box [43, 44]; <https://github.com/luismurao/ntbox>. A singular variable, which exhibits substantial cross-correlation and holds biological relevance to the species, was selected from a set of two alternative variables for the purpose of simplifying model interpretation [20, 45].

2.7 Projection transformation

Due to the disparate origins and resolutions of the Bio-Climatic (BIO) and Non-BIO variables, it is imperative to rectify their projections prior to data extraction and the subsequent prediction of the ensemble model. The achievement of this study was facilitated by employing a sequence of procedures in ArcMap utilizing ArcToolbox. Initially, the projection was established within the "projection and transformation" sub-window of Data Management Tools. The WGS 1984 EASE Geographic Coordinate System (GCS) was employed for this purpose. In order to accurately measure the extent of each habitat suitability class, it is necessary to utilize a Projected Coordinate System (PCS) within the "calculate geometry" function of Arc Map [46]. As a result, we have converted the projections of the habitat class raster file to WGS 1984 web Mercator (auxiliary sphere-3857). This step facilitates the computation of the area encompassed by a particular class, using a unit of measurement specified by the user (in this case, square kilometres were employed).

2.8 Ensemble species distribution modelling

For ensemble species distribution modelling, following algorithms from the Stacked Species Distribution Models (SSDM) R package [47] were used: Generalized Linear Models using Gaussian distribution "GLM", Generalized Additive Model "GAM" [48], Support Vector Machines "SVM", Random Forest "RF" [49, 50], Multivariate Adaptive Spline "MARS", and Maximum

Entropy-Maxent v. 3.4.1 [51, 52]. Artificial Neural Network “ANN” [53], and Classification Tree Analysis “CTA” [54]. We used the default for all studied algorithms [20]. We built the models using 70% of the data (training set) and evaluated the model performance with the rest of the 30% of the data evaluation set [43]. We built the final models with all studied predictors and RCPs using weighted mean separately. For more accurate predictions, models were evaluated using K-fold cross-validation with 10 folds and 10 replications for each algorithm and other settings for different algorithms were used in addition to the default option [55]. In order to reconcile the varying strengths and weaknesses inherent in each algorithm, SSDM undertook the process of averaging multiple model evaluation criteria. This resulted in the declaration of model qualities through an ensemble scoring system [56] that incorporated the following metrics. Area Under the Curve (AUC) of receiver operating characteristics (ROC), true skills statistics (TSS) and Kappa [57]. The criteria for Kappa and TSS values were: excellent, 0.85 to 1; very good, 0.7 to 0.85; good, 0.55 to 0.7; fair, 0.4 to 0.55; and fail, less than 0.4 [2, 58]. When evaluating the AUC statistic, the following grades are used: exceptional (0.90–1.00), very good (0.8–0.9), good (0.7–0.8), fair (0.6–0.7), and poor (0.5–0.6).

2.9 Post distribution analysis

2.9.1 Habitat suitability

The ASCII raster outputs derived from each SDM analysis, utilizing both bio-climatic and non-bioclimatic predictors as previously mentioned, were imported into ArcMap. The cell values, ranging from 0 to 1, were then employed to effectively categorize the habitats for this particular species. Utilizing the constant point break standards for each classification as outlined by Khan et al. [59], we have partitioned the area into four discrete groupings predicated on their degree of suitability: optimal, moderate, marginal, low, and absent/in-appropriate. We employed the raster calculator (Spatial Analyst Tool/Map Algebra/Raster Calculator) methodology to ascertain the total area, measured in square kilometres, that pertains to above suitability classes. We have conducted computations to determine the ratio of every category of habitat in relation to the percentage of the global overall landmass (PTGLA), employing a diverse range of predictors. By utilizing the Variable Importance Percentage (VIP) metric derived from each analysis, we have successfully discerned the preponderant influence of bioclimatic and non-bioclimatic predictors on the distributional pattern of this species [58].

The output of the global distribution of *I. oblongifolia* is subjected to additional filtration through the computation of the area and corresponding PTGLA values under the four aforementioned suitability classes. This is performed at both the continental (Africa) and national (India) levels. The principal aim of this exercise was to ascertain a quantitative measure for the Percent Indigeneity (PI) of the species, utilizing a recently suggested index that assigned weights to the various habitat suitability categories in the following manner:

$$\text{Percent of Indigenous} = \frac{\text{Habitat suitability type} \times \sum_{i=0}^n \text{polygon number}}{\sum_{i=0}^n \text{Area (sq.km)}} \times 100$$

Habitat suitability type scoring factors = optimum = 1; Moderate = 0.75; Marginal = 0.50 and Low = 0.25. $\sum_{i=0}^n \text{polygon number}$ = sum of polygon numbers pertains to a specific suitability class and $\sum_{i=0}^n \text{Area (sq.km)}$, represent the total area (sq. km) occupied by the species under the specific suitability class. High index value represent higher probability to restriction within the confined zone and lesser value represents the species generality.

2.9.2 Ellipsoid niche hypervolume

In order to augment the accuracy of species localization, machine learning models proffer an array of consequential variables. The hypervolumes associated with the niches of this species were quantified through the utilization of the top three predictors across all bioclimatic scenarios and RCPs. For this study, we employed NicheToolBox: Ntbox [60]. Ntbox, a GUI tool written in the R programming language, calls for the raster output of BC variables. By calculating the species' environmental values' centroid and covariance matrix, ellipsoidal models were constructed. It goes from the geographical centre outward to all possible settings in the research region. By employing this methodology, we can ascertain the environmental variables that govern both the fundamental and realized niche of this species.

3 Results

3.1 Data processing and multicollinearity

Through a comprehensive analysis of diverse sources worldwide, we have successfully extracted 334 sites of this species. By employing the spatial thin window feature in Wallace Software's [33] we were able to eliminate all but one instance of a given record within a specific region, utilizing a thinning distance of 10 kms. Ultimately, the ENM development process was completed by incorporating precisely 225 presence points of *I. oblongifolia* that were free of any spatial autocorrelation. These widely scattered coordinates were fed into the ConR software to calculate the present Extent of Occurrence (EOO) and Area of Occupancy (AOO). Figure 2 depicts the exercise's outcomes: an EOO of 50,640,914 km² and an AOO of 900.0 km², with the former indicating a status of least concern and the latter a vulnerable status. The final studied bioclimatic variables and their corresponding VIP values are listed in Table 1. Similarly, the "x" symbol denotes the removed variables in Table 1. Considering their high levels of correlation with other bio-climatic factors, BC-2 and BC-19 were removed from all analyses.

3.2 Model performances

Excellent ensemble model qualities were recorded as AUC approached 0.90 based on current climate and RCPs corresponding to timeframes of 2050 and 2070 (Table 2). Among the individual algorithm, the Random Forest has been identified as the optimal choice among other algorithms, owing to its superior performance in terms of AUC values across all climatic predictors (Tables S2 and S3). It yielded Kappa and TSS values greater than 0.82, indicating high model quality. On the other hand, low Kappa values obtained using the Maxent tool with all climate predictors show that this widely used ENM technique is ineffective for this species. With global human modification of Terrestrial Ecosystem (GHMTS), N and P fertilizer predictors, the ensemble model has good model quality, with AUC values less than 0.80. Random forest and Maxent had the highest and lowest Kappa values for the same predictor, respectively (Table S4). Livestock population densities (cattle, goat, sheep) had higher AUC values (> 0.80) with both the individual algorithm and the ensemble model.

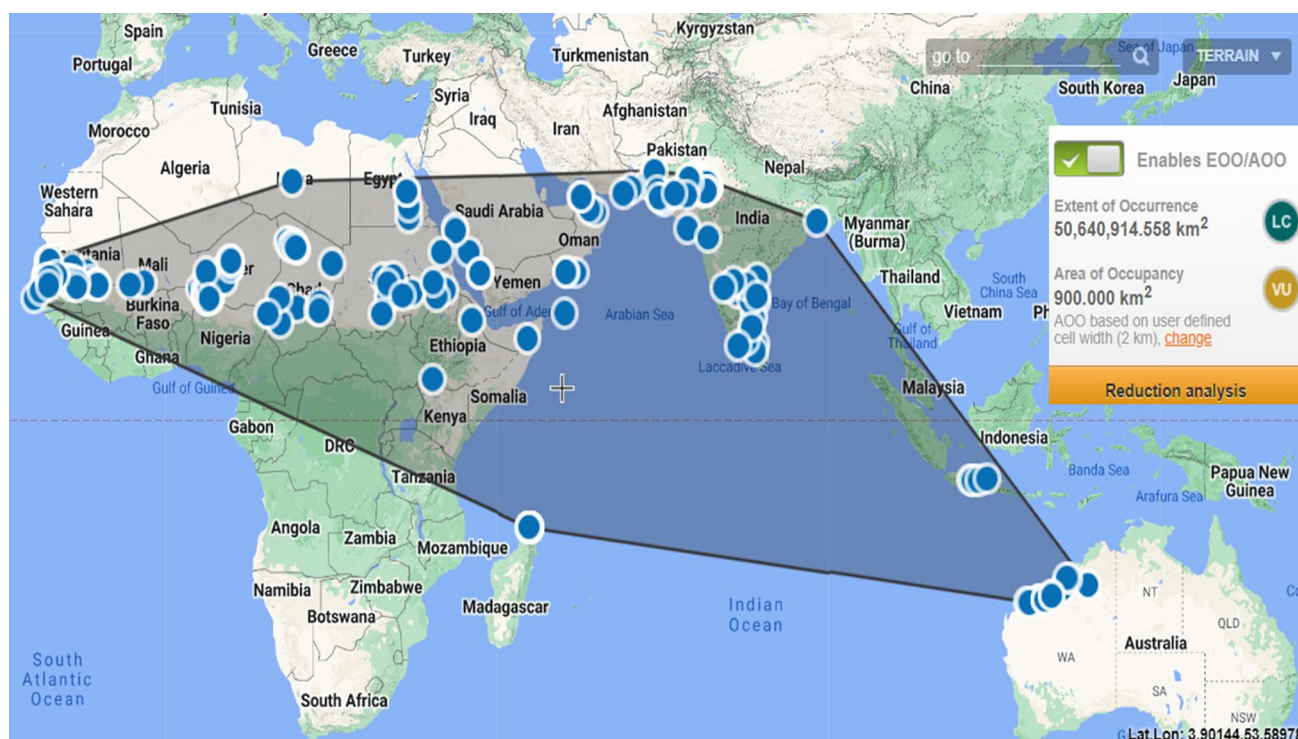


Fig. 2 Graphical display of spatially thinned 225 presence point of *Indigofera oblongifolia* convey Extent of Occurrence (EOO) and Area of Occupancy (AOO)

Table 1 Variable Importance Percentage (VIP) of different bio-climatic variables during current, 2050 and 2070 timeframe with their respective RCPs

Variables	Current	2050 RCPs				2070 RCPs			
		2.6	4.5	6.0	8.5	2.6	4.5	6.0	8.5
BC-1 Annual Mean Temperature (AMT)	x	6.4	x	x	x	x	x	x	x
BC-3 Isothermality (BC2/BC7) ($\times 100$) (Iso)	x	x	4.0	9.19	8.56	6.2	8.4	3.2	15
BC-4 Temperature Seasonality (standard deviation $\times 100$) (TempS)	x	13.57	x	x	x	x	x	x	x
BC-5 Max Temperature of Warmest Month (MaTWaM)	14.27	x	48.3	15.77	7.3	21.2	16.6	x	25
BC-6 Min Temperature of Coldest Month (MiTCM)	6.84	x	1.0	12.55	8.12	x	x	4.8	x
BC-7 Temperature Annual Range (BC 5-BC 6) (TAR)	8.56	3.38	x	x	x	5.1	5.6	x	x
BC-8 Mean Temperature of Wettest Quarter (MeTWaQ)	x	10.38	29.4	8.61	10.2	27.9	26.8	5.4	x
BC-9 Mean Temperature of Driest Quarter (MeTDQ)	x	x	2.4	21.8	3.89	6.4	x	0.6	x
BC-10 Mean Temperature of Warmest Quarter (MeTWaQ)	x	x	x	x	12.8	x	x	41.4	x
BC-11 Mean Temperature of Coldest Quarter (MeTCQ)	11.95	x	x	x	x	x	x	12.0	x
BC-12 Annual Precipitation (AnPr)	18.61	10.42	3.8	x	x	x	x	x	x
BC-13 Precipitation of Wettest Month (PrWeM)	9.05	x	x	x	x	x	3.1	3.0	x
BC-14 Precipitation of Driest Month (PrDM)	x	14.04	0.1	17.41	x	0.7	x	0.9	0.7
BC-15 Precipitation Seasonality (Coefficient of Variation) (PrS)	30.68	x	x	x	28.52	28.2	38.4	28.0	55.8
BC-16 Precipitation of Wettest Quarter (PrWeQ)	x	19.54	0.4	0.8	7.55	3.8	x	x	2.8
BC-17 Precipitation of Driest Quarter (PrDQ)	x	11.92	9.5	6.26	5.49	x	x	x	x
BC-18 Precipitation of Warmest Quarter (PrWaQ)	x	10.24	1.3	7.57	7.08	0.7	1.7	0.7	0.7

X= indicates removal of variable due to significant co-linearity with other variables. BC-2 and BC-19 were excluded from all analysis

3.3 Variable importance percentage

Among the bioclimatic and non-bioclimatic predictors examined, we used the Variable Importance Percentage (VIP) of each analysis to identify which variable affects the distribution pattern of this species the most (Table 1). Our analysis suggested that Precipitation Seasonality (BC-15) was the most influential variables with current, 2050 RCP 8.5, 2070 RCPs 2.6, 4.5 and 8.5. Mean Temperature of Wettest Quarter (BC-8) was identified as second influential variables with 2070 RCP 2.6 and 4.5. Precipitation of Wettest Quarter (BC-16), Maximum Temperature of Warmest Month (BC-5), Mean Temperature of Driest Quarter (BC-9) and Mean Temperature of Warmest Quarter (BC-10) were identified as important variables for distribution of this species during 2050 RCP -2.6, 2050—RCP 4.5, 2050 -RCP 6.0 and 2070-RCP 6.0, respectively (Table 1). Among the human related factors, GHMTS showed the largest impact on the distribution of this species with VIP value of 60.84, followed by nitrogen fertilizer (VIP = 23.64) and Phosphorus

Table 2 Ensemble model qualities parameters with both bioclimatic as well as non-bioclimatic predictors

Predictors	AUC	Sensitivity	Specificity	Kappa	TSS
Current Time Frame	0.933	0.944	0.921	0.770	0.865
2050 RCP 2.6	0.918	0.931	0.904	0.726	0.836
2050 RCP 4.5	0.919	0.928	0.910	0.737	0.838
2050 RCP 6.0	0.914	0.927	0.902	0.723	0.828
2050 RCP 8.5	0.919	0.929	0.909	0.736	0.838
2070 RCP 2.6	0.921	0.923	0.919	0.747	0.842
2070 RCP 4.5	0.911	0.917	0.904	0.720	0.821
2070 RCP 6.0	0.916	0.928	0.905	0.729	0.833
2070 RCP 8.5	0.892	0.900	0.884	0.677	0.784
GHMTS and N & P Fertilizer	0.78	0.78	0.77	0.45	0.55
Livestock Population	0.84	0.85	0.83	0.56	0.67

GHMTS Global Human Modification of Terrestrial Systems

fertilizer with VIP value of 15.51. Similarly, among the livestock population, our analysis revealed greatest impact of goat density (VIP = 64.06) on the distribution of this species followed by sheep and cattle densities with VIP values of 20.17 and 17.75, respectively.

Observable trends in the predicted value of this species with respect to bioclimatic variables are depicted by curves of response in Figs. 3 and 4. We mapped out response curves using the two most significant predictors for each climate epoch and their RCPs. Precipitation Seasonality is defined as the coefficient of variation (the ratio of monthly total precipitation standard deviation to monthly total precipitation means). Based on our research, we know that this species' habitat suitability is generally in sync with this bio-climatic variables up to a factor of 150 for the current (Fig. 3a), 2050-RCP 8.5 (Fig. 3i), and 2070-RCPs 2.6 (Fig. 4a), 4.5 (Fig. 4c), and 8.5 (Fig. 4g). Further, habitat suitability of this species also exhibited synchronization with Precipitation of Wettest Quarter (BC-16) during 2050-RCP 2.6 wherein we found 500 mm is the optimum for this species (Fig. 3c). Maximum Temperature of Warmest Month (30°C) found most suitable with 2050 RCP 4.5 (Fig. 3e). Similarly, Mean Temperature of Driest Quarter between 25 to 30°C supported this species with 2050 RCP 6.0 (Fig. 3g). Mean Temperature of Warmest Quarter (BC-10) particularly 34°C is found most suitable with 2070 RCP 6.0 (Fig. 4e).

3.4 Area (Sq. km × 10³) under different suitability classes and their percent of total geographical land area

3.4.1 Global scenario

Spatial extents of different habitat suitability classes identified for *I. oblongifolia* with help of ensemble modelling are depicted in Fig. 5 (current bio-climatic predictors and non-climatic variables like livestock population and GHMTS as well as nitrogen and phosphorus fertilizers), Fig. 6 (2050 bio-climatic predictors with RCPs 2.6, 4.5, 6.0 and 8.5) and Fig. 7 (2070 bio-climatic predictors (with RCPs 2.6, 4.5, 6.0 and 8.5). Based on cell values in raster out put processed with ArcMap, we can categorize four types of suitability classes with 0.20-point break and these were designated as optimum, moderate, marginal and low [20]. Pictorial representation revealed this species is currently found in Australia, Southern Asia (India and Pakistan), the Arabian Peninsula (including Oman, Muscat, the Area of Qatar, and

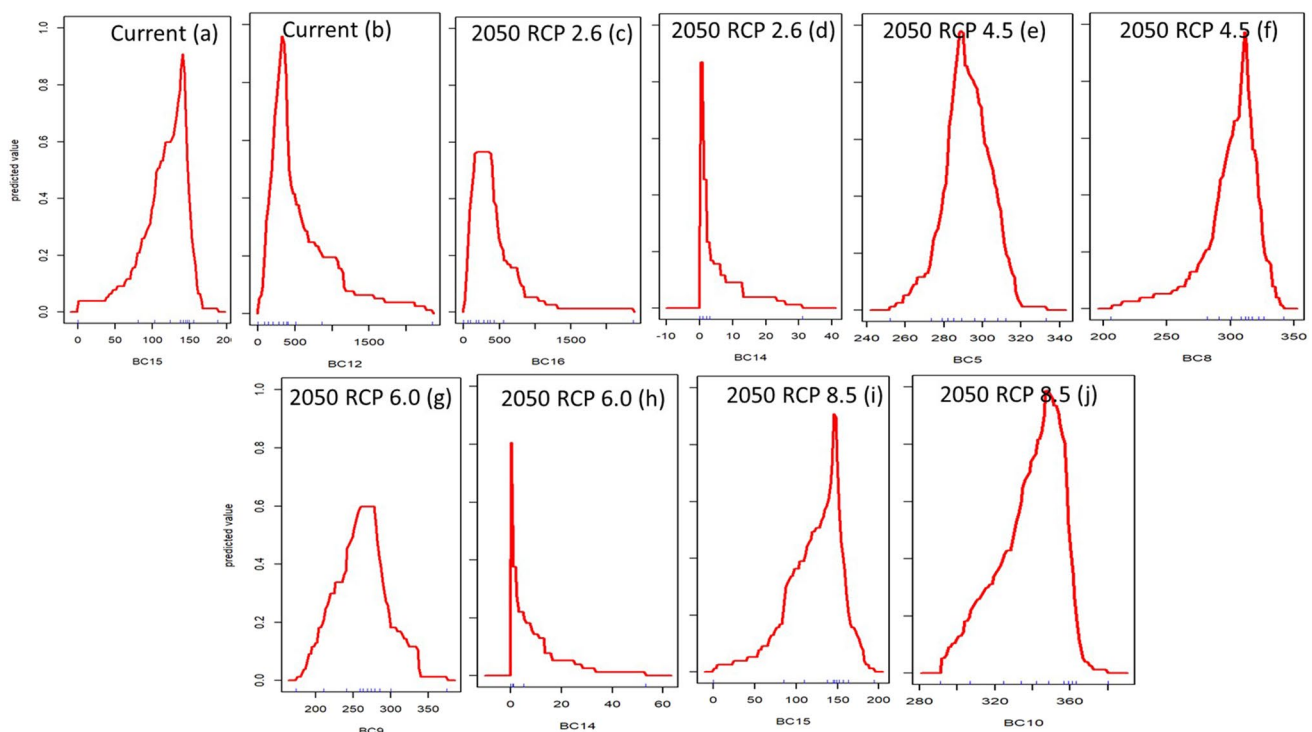


Fig. 3 Response curves showing dependence of habitat suitability on the predictors related to current and 2050 bio-climatic timeframes and their RCPs

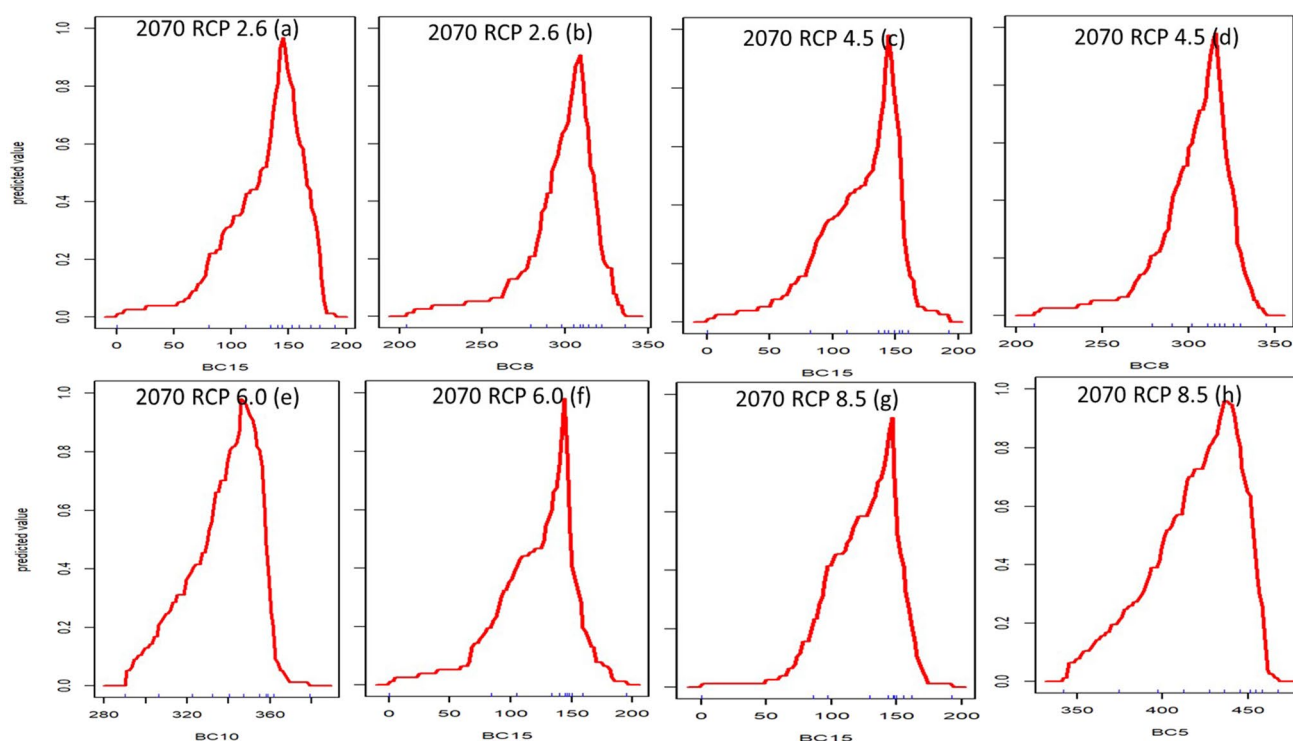


Fig. 4 Response curves showing dependence of habitat suitability on the predictors related to 2070 bio-climatic timeframes and their RCPs

Yemen), and African countries such as Somalia, Sudan, Chad, Niger, Mali, and Senegal (Fig. 5A). This species has been found in Australia at Pardo, Karratha, Port Hedland, Broome, Derby, St. George ranges, and Muller ranges (Fig. 6A, B).

The Indian states of Gujarat, Tamil Nadu, and Andhra Pradesh also saw a reduction in area as a result of this climatic projection. In the Arabian Peninsula, only Oman had a higher concentration. While in Africa, it was missing from Somalia and had shrunk in Sudan, Chad, Niger, Mali, and Senegal (Fig. 7B and C). However, with 2050 RCP 8.5, its optimum suitability areas were drastically reduced from Australia, the Arabian Peninsula, and southern India (Andhra Pradesh and Tamil Nadu). Its primary territory fell under this category in the Indian states of Rajasthan and Gujarat, as well as the Pakistan cities like Hyderabad and Karachi. This species' area has shrunk across Africa, with Sudan, Chad, Niger, Mali, and Senegal reporting shrinkage (Fig. 7D). We observed more shrinkage for this species in the Australian region with 2070 RCP 2.6 (Fig. 7A) compared to 2050. In the climatic timeframe 2070 with RCPs 4.5, 6.0, and 8.5, we observed the complete absence of this species from the Arabian region, as well as a significant shrinkage in width in the African region (primarily Sudan). Reduction from southern India and a small patch of Australia (Fig. 7B, C and D).

Global calculated area ($\text{Sq. km} \times 10^3$) for each suitability class with the current, 2050, 2070 climate and RCPs scenarios as well as studied non-bioclimatic predictors are depicted in Table 3. Under the optimum are we recorded highest area (10,803.64) with GHMTS as well as nitrogen and phosphorus fertilizers), while among the bio-climatic variables, the higher (7638.67) areas under this class were recorded with 2050 RCP 2.6, and the lowest (2910.09) area among all the predictors was recorded with 2070 RCP 6.0.

At a global level and particularly area under optimum suitability class, our data revealed increasing trends during 2050 RCP 2.6 and 4.5 as compared to its current area. However, with other remaining climatic variables and RCPs, we observed declining trends (Table 3). With reference to global total geographical land area (PTGLA) this species has only 2.12, 1.46 and 1.50 percent area under the optimum category quantified GHMTS, N and P, livestock and current climatic conditions, respectively.

Our bioclimatic predictors show that this species occurs in great abundance on both the African continent and in India; therefore, in order to put a numerical value on its indigenous status, we have conducted a thorough examination of its distribution across both landmasses in relation to various climate scenarios.

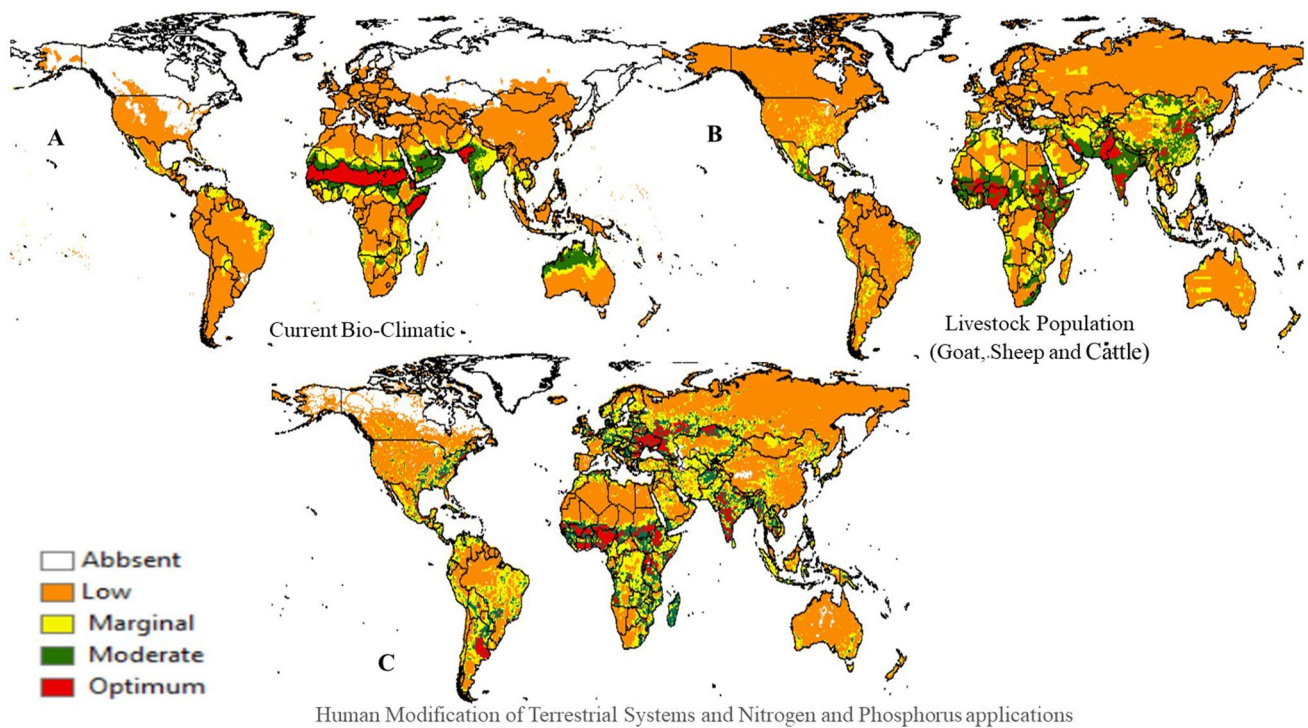


Fig. 5 Habitat suitability of *Indigofera oblongifolia* with current bio-climatic predictors (A) and non-climatic variables like livestock population (B) and human modification of terrestrial ecosystem (GHMTE) as well as nitrogen and phosphorus fertilizers (C)

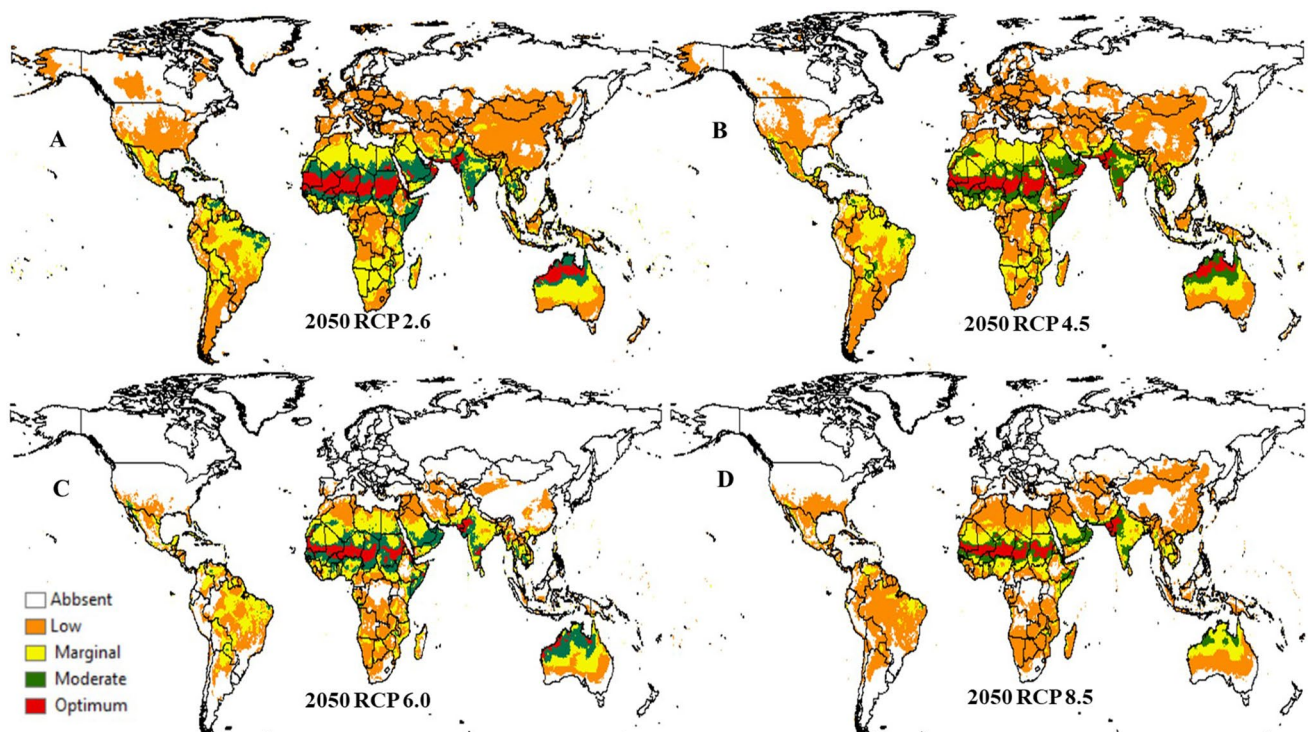


Fig. 6 Habitat suitability of *Indigofera oblongifolia* with 2050 bio-climatic predictors (A) RCP 2.6; (B) RCP 4.5; (C) RCP 6.0 and (D) RCP 8.5

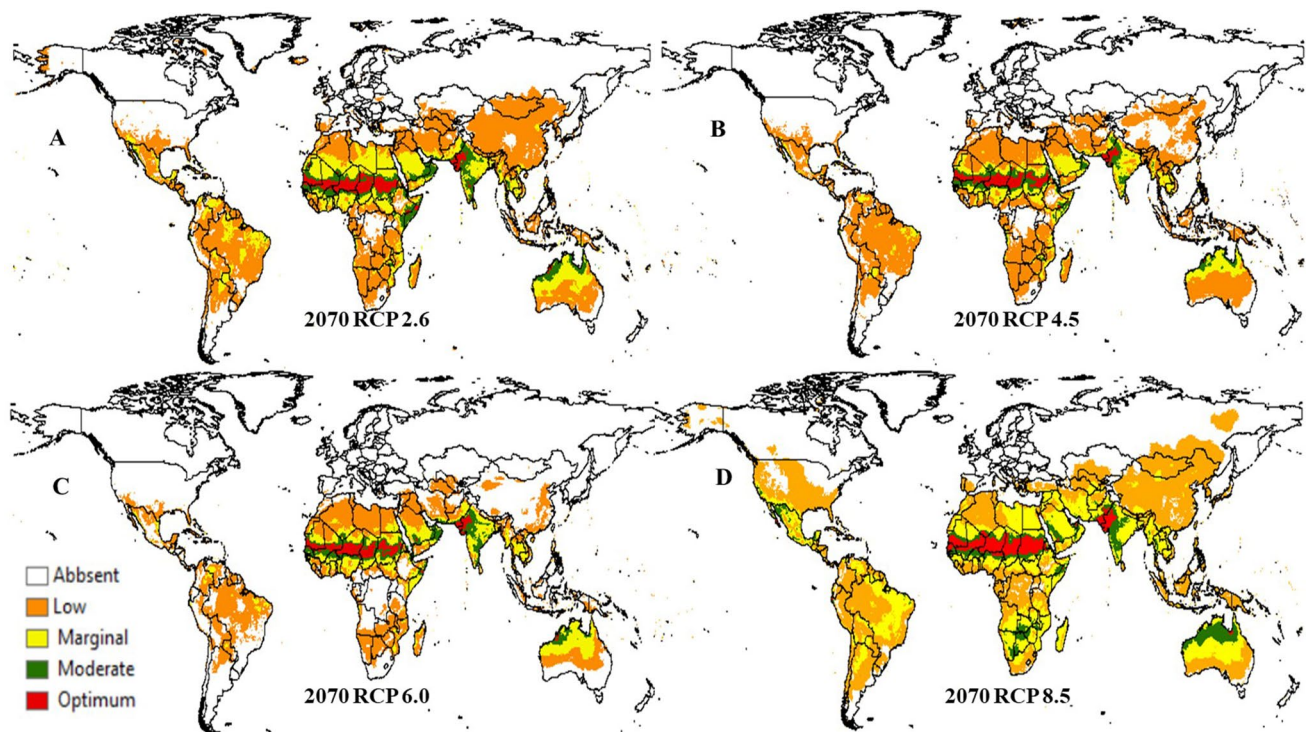


Fig. 7 Habitat suitability of *Indigofera oblongifolia* with 2070 bio-climatic predictors (A) RCP 2.6; (B) RCP 4.5; (C) RCP 6.0 and (D) RCP 8.5

Table 3 Global Area (Sq. km $\times 10^3$) of different suitability class and their percent of total geographical land area (PTGLA) of the world (510,072,000 sq. km.) Area under different habitat suitability classes of *I. oblongifolia* with current bio-climatic conditions, projected 2050 and 2070 bio-climatic prediction with different RCPs and non-bioclimatic variables (GHMTE, N and P and Livestock population)

Climatic and non-climatic variables	Optimum		Moderate		Marginal		Low	
	Area	PTGLA	Area	PTGLA	Area	PTGLA	Area	PTGLA
Current BC	5859.8	1.15	15,644.9	3.07	12,333.1	2.42	116,329.5	22.8
GHMTE, N and P	10,803.6	2.12	22,722.4	4.45	49,347.9	9.67	189,428.7	37.1
Livestock	7471.4	1.46	16,091.3	3.15	3856.6	0.76	129,262.3	25.3
2050 RCP 2.6	7638.6	1.50	14,346.4	2.81	35,943.2	7.05	267,117.8	52.3
2050 RCP 4.5	6443.3	1.26	12,977.4	2.54	32,266.9	6.33	83,572.7	16.3
2050 RCP 6.0	3844.2	0.75	11,288.4	2.21	25,160.9	4.93	34,112.8	6.6
2050 RCP 8.5	3138.9	0.62	6079.6	1.19	16,053.5	3.15	65,266.0	12.8
2070 RCP 2.6	3706.0	0.73	7440.4	1.46	23,250.0	4.56	102,068.2	20.0
2070 RCP 4.5	2938.6	0.58	5005.6	0.98	14,348.3	2.81	60,123.0	11.7
2070 RCP 6.0	2910.0	0.57	5119.0	1.00	14,128.8	2.77	41,188.3	8.0
2070 RCP 8.5	4886.4	0.96	8827.1	1.73	33,753.4	6.62	77,398.4	15.1

3.4.2 Continent scenario (Africa)

Results related to the distribution patterns of this species across the African continent and the area covered by it in various habitat suitability classes under different climatic and RCPs predictors are presented in Fig. 8 and Table 4, respectively. Our research indicated that this species likely exists in the Sudan, Chad, Niger, Mali, and Mauritania, all of which contain suitable habitats ranging from highly suitable to moderate for its survival. However, under RCP 6.0 by the year 2070, we found that countries like Zambia, Angola, Gabon, and the Democratic Republic of the Congo no longer had any indications of this species at all (Fig. 8). In 2070 RCP 8.5, some new areas of moderate suitability appeared in Botswana, Zimbabwe, and Namibia.

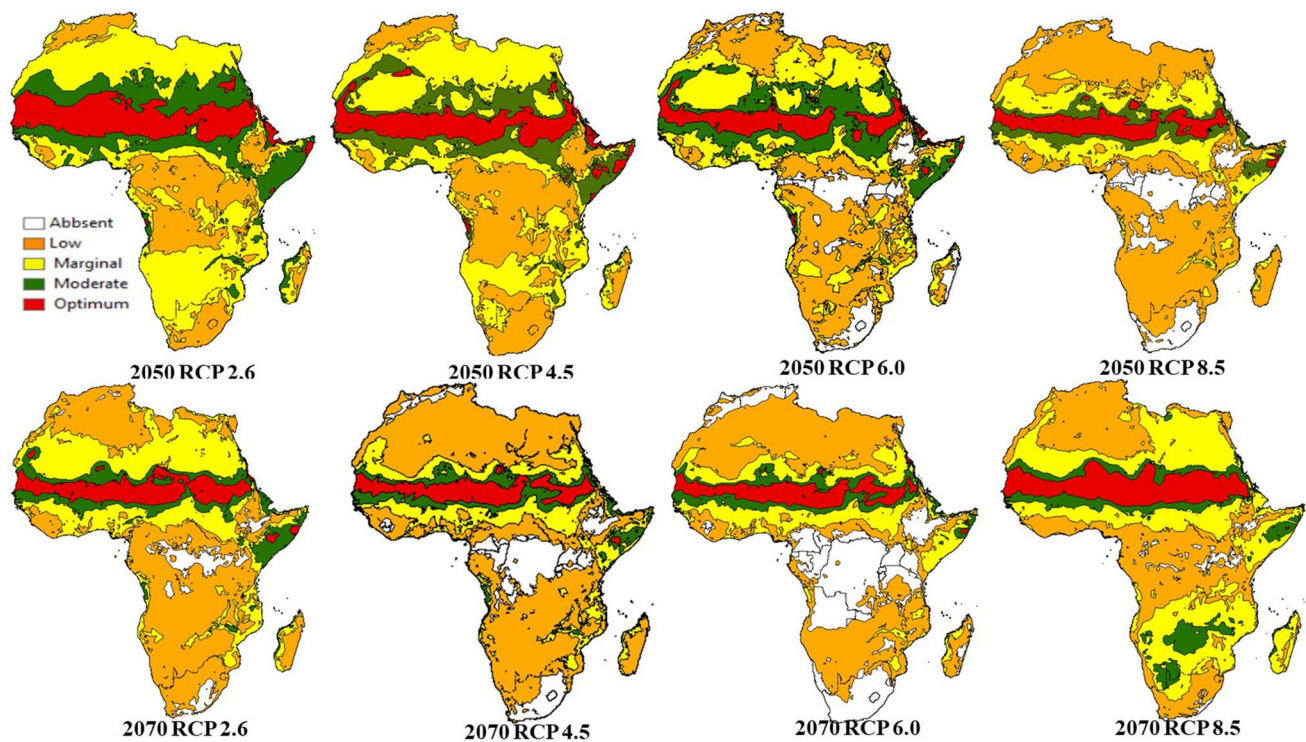


Fig. 8 Habitat suitability of *Indigofera oblongifolia* at Africa continent with 2050 and 2070 bio-climatic predictors and their respective RCPs 2.6, 4.5, 6.0 and 8.5

Table 4 Africa Area (Sq. km $\times 10^3$) of different suitability class and their percent of total geographical land area (PTGLA) of the world (510,072,000 sq. km.). Area under different habitat suitability classes of *I. oblongifolia* with current bio-climatic conditions, projected 2050 and 2070 bio-climatic prediction with different RCPs and non-bioclimatic variables (GHMTE, N and P and Livestock population)

Climatic and non-climatic variables	Optimum		Moderate		Marginal		Low	
	Area	PTGLA	Area	PTGLA	Area	PTGLA	Area	PTGLA
Current BC	4969.2	16.37	7246.6	23.87	5727.5	18.86	15,516.1	51.10
GHMTE, N and P	3485.6	11.48	6604.2	21.75	8680.8	28.59	13,902.9	45.79
Livestock	3546.4	11.68	7660.1	25.23	11,552.7	38.05	10,638.8	35.04
2050 RCP 2.6	5050.6	16.63	7038.8	23.18	13,739.8	45.25	7645.5	25.18
2050 RCP 4.5	4088.9	13.47	5982.7	19.70	13,486.7	44.42	9940.9	32.74
2050 RCP 6.0	2814.5	9.27	5778.4	19.03	10,334.7	34.04	11,056.6	36.41
2050 RCP 8.5	2627.2	8.65	3005.0	9.90	7031.4	23.16	16,777.5	55.25
2070 RCP 2.6	3064.9	10.09	3575.7	11.78	10,298.5	33.92	14,904.1	49.08
2070 RCP 4.5	2469.4	8.13	2869.8	9.45	6146.2	20.24	17,654.4	58.14
2070 RCP 6.0	2379.32	7.84	2433.5	8.01	5966.4	19.65	14,501.1	47.76
2070 RCP 8.5	3986.3	13.13	3740.5	12.32	12,810.4	42.19	12,260.9	40.38

Maximum area (5050.67 Sq. km $\times 10^3$) under the optimum category was recorded with 2050 RCP 2.6 which was 16.63 per cent of total geographical land area of the Africa (Table 4), similarly the lowest area (2379.32 Sq. km $\times 10^3$) under this class was recorded with 2070 RCP 6.0 which is only 7.84% of the total land area of Africa. We also found that, relative to current conditions, areas falling under this category were smaller for all future climate and non-climate scenarios except for the 2050 RCP 2.6. Both non-climatic variables showed more or less similar areas in this class (Table 4), ranging from 11.48 to 11.68 percent of Africa's total land area. However, maximum area (7660.19 Sq. km $\times 10^3$) under the moderate class was observed with livestock component, followed by current (7246.6 Sq. km $\times 10^3$) and 2050 RCP 2.6 (7038.84 Sq. km $\times 10^3$).

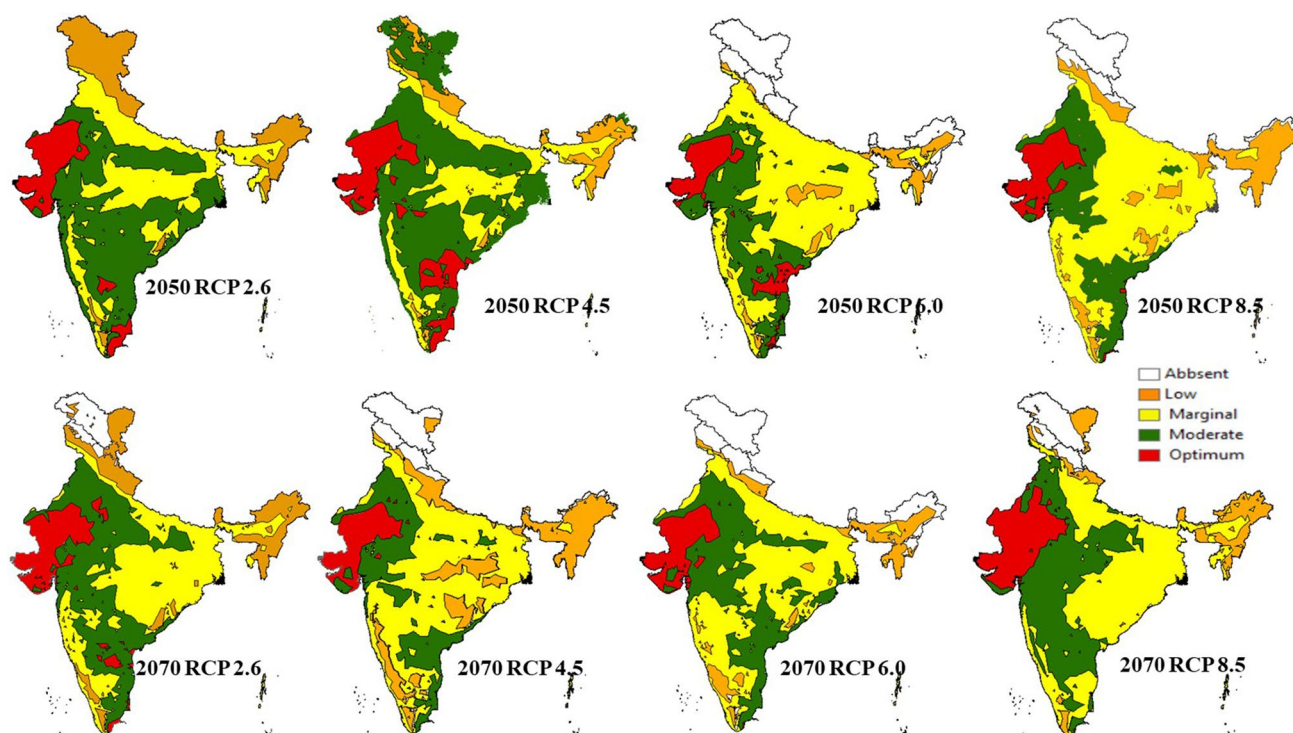


Fig. 9 Habitat suitability of *Indigofera oblongifolia* at India with 2050 and 2070 bio-climatic predictors and their respective RCPs 2.6, 4.5, 6.0 and 8.5

Table 5 India Area (Sq. km $\times 10^3$) of different suitability class and their percent of total geographical land area (PTGLA) of the world (510,072,000 sq. km.). Area under different habitat suitability classes of *I. oblongifolia* with current bio-climatic conditions, projected 2050 and 2070 bio-climatic prediction with different RCPs and non-bioclimatic variables (GHMTE, N and P and Livestock population)

Climatic and non-climatic variables	Optimum		Moderate		Marginal		Low	
	Area	PTGLA	Area	PTGLA	Area	PTGLA	Area	PTGLA
Current BC	398.5	12.12	1141.3	34.72	795.8	24.21	780.1	23.73
GHMTE, N and P	796.1	24.22	1004.4	30.56	911.2	27.72	496.9	15.12
Livestock	949.9	28.90	1033.5	31.44	596.6	18.15	154.3	4.69
2050 RCP 2.6	415.1	12.63	821.1	24.98	587.3	17.87	667.6	20.31
2050 RCP 4.5	578.4	17.60	1911.5	58.15	459.7	13.99	281.2	8.56
2050 RCP 6.0	342.8	10.43	936.5	28.49	1770.2	53.85	300.7	9.15
2050 RCP 8.5	348.9	10.61	908.6	27.64	1209.1	36.78	583.6	17.76
2070 RCP 2.6	438.9	13.35	1025.3	31.19	1106.0	33.65	561.0	17.07
2070 RCP 4.5	324.2	9.86	801.3	24.38	1162.2	35.36	777.1	23.64
2070 RCP 6.0	240.3	7.31	850.3	25.87	934.6	28.43	1164.0	35.41
2070 RCP 8.5	499.9	15.21	1230.9	37.45	964.5	29.34	362.1	11.02

3.4.3 Country level (India)

Results related to the distribution patterns of this species across the India and the area covered by it in various habitat suitability classes under different climatic and RCPs predictors are presented in Fig. 9 and Table 5, respectively. Within India, our analysis revealed clear zonation of different habitat suitability classes; the optimum places for this species were consistently in the west, including the states of Rajasthan and Gujarat; however, under the RCPs of 2.6, 4.5, and 6.0 of 2050, there are also suitable locations in the south, in states like Tamil Nadu and Andhra Pradesh. These southern regions have become moderately suitable with remaining climatic time frame and RCPs. The central, southwestern, and eastern regions of India provide moderate habitat for this species under RCP 2.6 and 4.5 in 2050, but under all other RCPs, these regions will fragment into smaller and smaller patches (Fig. 9). The eastern part of the country has been labelled as a low suitable area, while the northern part of the country is completely devoid of this species. Two very different scenarios for

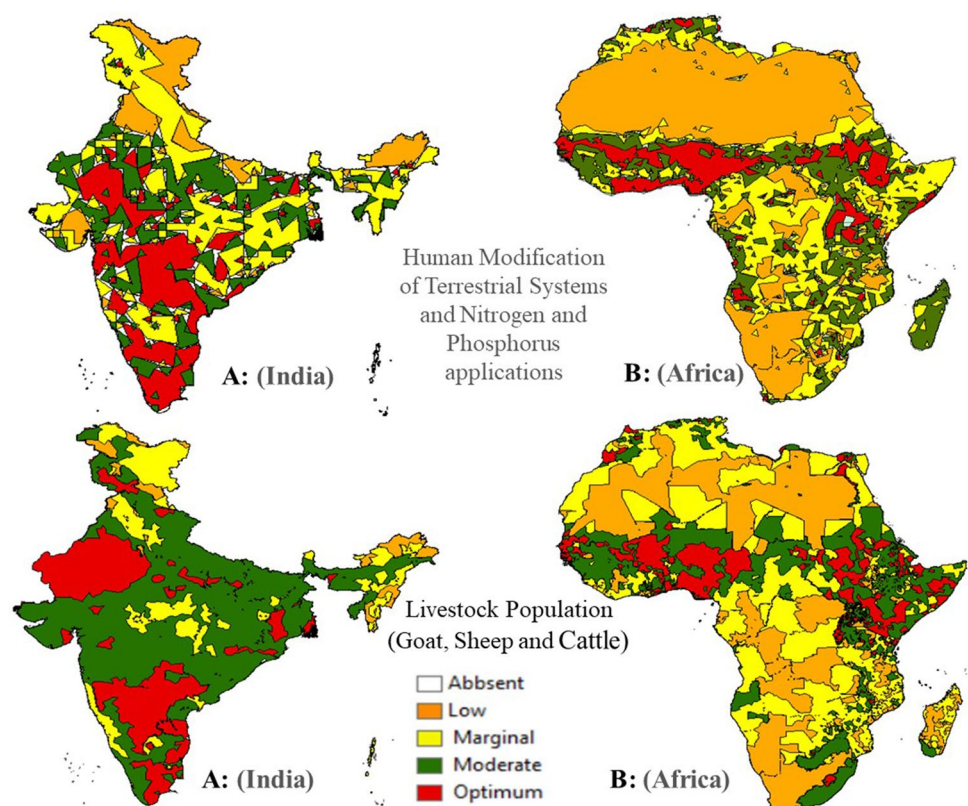
this species were displayed within India in relation to GHMTS, Nitrogen, and Phosphorus fertilizers, and livestock population densities. When we looked at the former non-bioclimatic variables, we found that there were intermingled classes of habitat suitability (Fig. 10). When travelling across India with livestock, we regularly saw large regions that fell into either the optimum (Western and Southern regions) or the moderate (Central, Eastern, and Western regions) categories. As obvious, we recorded the higher areas under the optimum class with non-bioclimatic variables, which were roughly covering 24 to 28% of India total land area. Among the climatic time-frame and RCPs, highest (578.79 Sq. km $\times 10^3$; 17.60%) and lowest (240.39 Sq. km $\times 10^3$; 7.61%) area under this class were recorded with 2050 RCP 4.5 and 2070 RCP 2.6, respectively. Moderate areas for this species covering 24.38% (2070 RCP 4.5) to 58.15 (2050 RCP 4.5).

3.4.4 Percent Indigeneity of *I. oblongifolia*

We calculated the percentage of Indigenous for this species using a weighting pattern of habitat types, the number of polygons (the number of sites occupied by this species), and the total occupied area under the specific suitability class, and the results are shown in Fig. 11 (Africa) and Fig. 12 (Indian region). This is a new index proposed first time and our results indicates that indigenous nature of this species largely affected with number of polygons i.e., higher the occurrence probability at different patches.

Livestock population was the strongest predictor of optimal indigenous production in Africa, followed by GHMTS, nitrogen, and phosphorus. The highest indigenous value (2.06) was recorded with 2050 RCP 6.0 among climatic and RCPs, and the lowest (0.32) was recorded with 2070 RCP 8.5. In comparison to the current bio-climatic predictor, our results show an increase in indigenous nature under this class with 2050 and 2070 and their RCP projections. In comparison to optimum habitat, moderate habitat has more climatic indigenous values. We recorded generality patterns for the low suitability class because we had lower values of this index. The indigenous score in the optimum class was higher in India than in the African region. The highest value (7.90) was recorded with livestock, followed by 2050 RCP 6.0 with a value of 7.0. (Fig. 12).

Fig. 10 Habitat suitability of *Indigofera oblongifolia* at India and Africa with non-bioclimatic variables GHMTES, N and P and Livestock populations



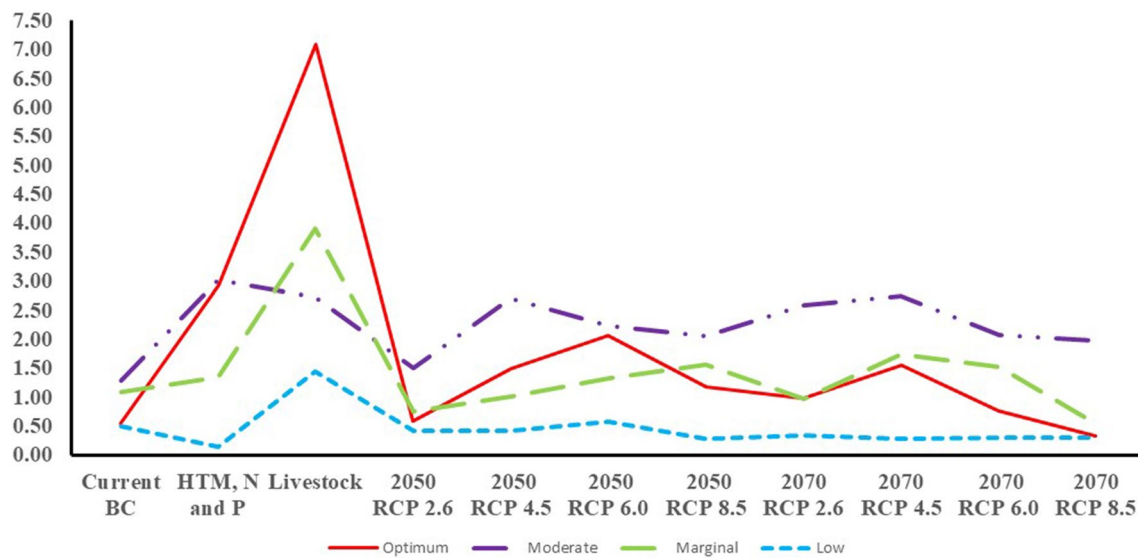


Fig. 11 Per cent Indigeneity of *Indigofera oblongifolia* at Africa continent with different habitat type under different climatic and non-climatic projections

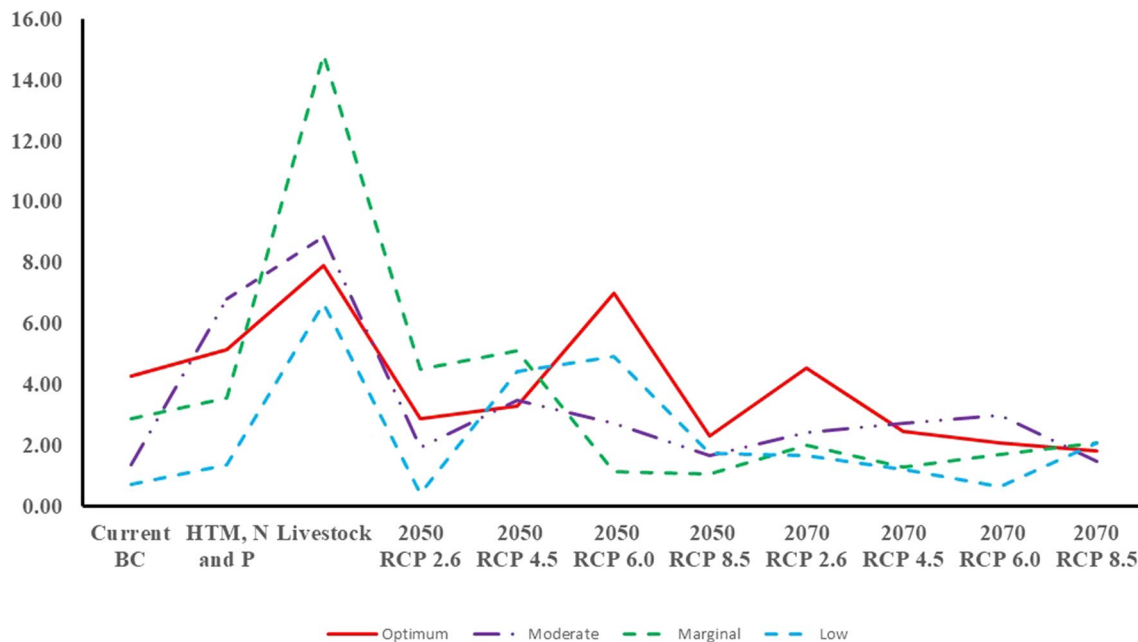


Fig. 12 Per cent Indigeneity of *Indigofera oblongifolia* at India continent with different habitat type under different climatic and non-climatic projections

3.4.5 Ellipsoid niche hypervolume

We developed ellipsoid hypervolume (multidimensional space of resources available for a species) to simulate the species' fundamental niche (defined as a species' ability to persist and reproduce in a broader range of environments when not competing with other species) and its realized niche (when it is in the presence of other interacting species) using occurrence records of *I. oblongifolia* and the most important environmental variables identified from ensemble modeling in form of raster output. This allows us to identify the bio-climatic factors that governs its fundamental and realized niche. The results are display in Fig. 13 (current and 2050 and its four RCPs) and Fig. 14 (2070 and its four RCPs). In both, The blue colour represents niche stability, the green colour represents niche unfilling (the proportion of the native niche

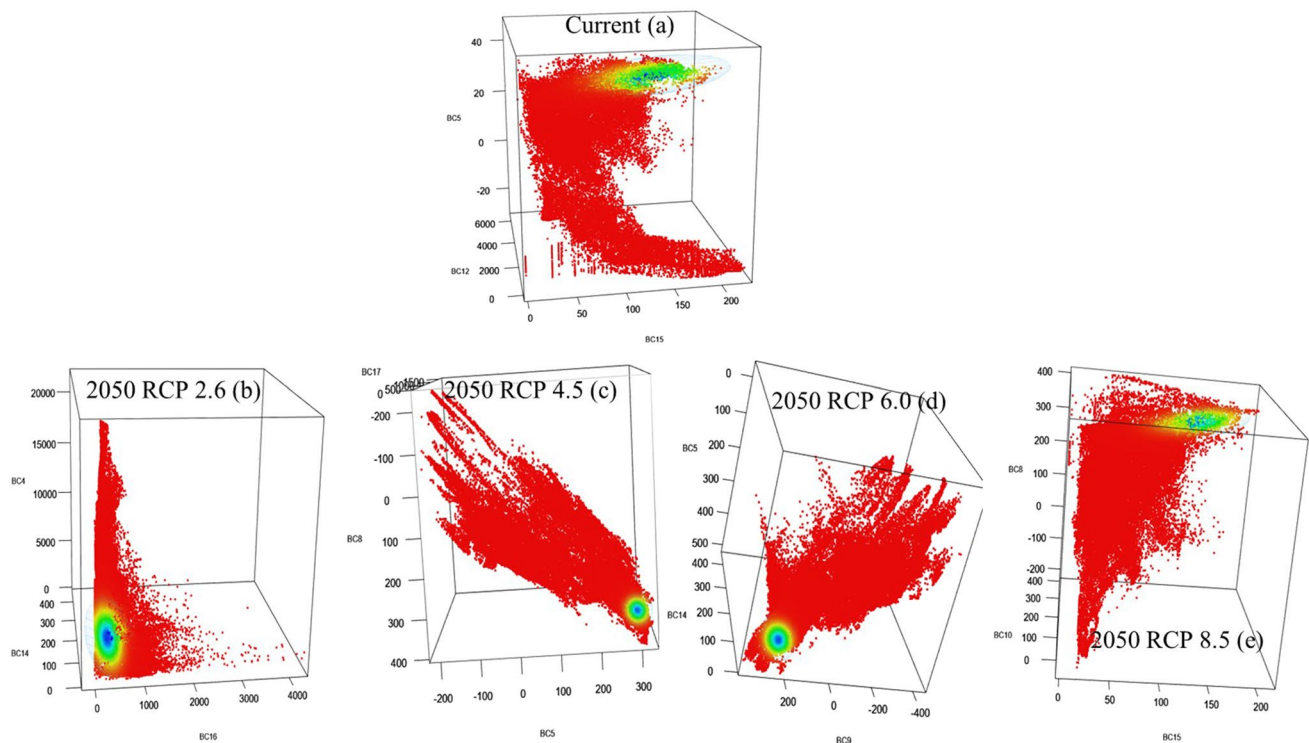


Fig. 13 Graphical representation of *Indigofera oblongifolia* niche hypervolume with three most influential bioclimatic variables pertains to current and 2050 (2.6, 4.5, 6.0 and 8.5 RCP) bioclimatic time-frame. The blue colour represents niche stability, the green colour represents niche unfilling (the proportion of the native niche that does not overlap with the exotic niche), and the red colour represents niche expansion

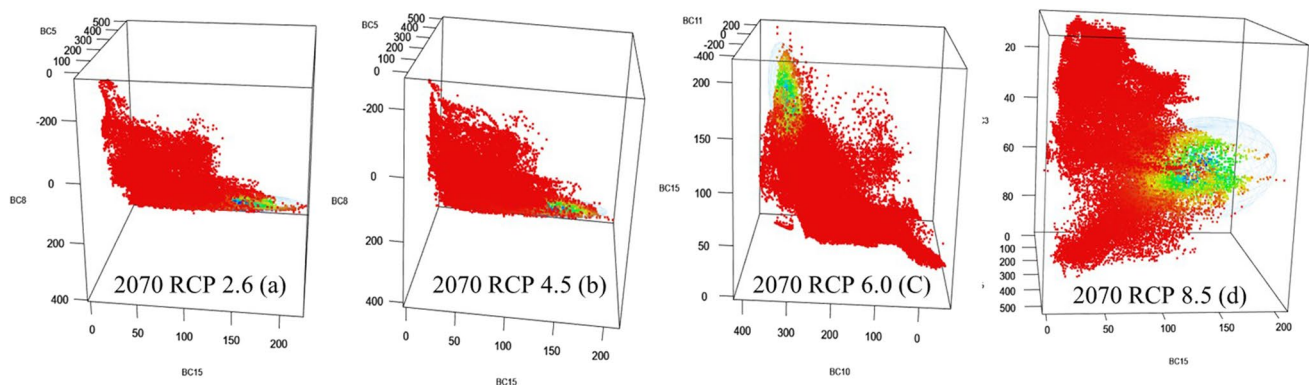


Fig. 14 Graphical representation of *Indigofera oblongifolia* niche hypervolume with three most influential bioclimatic variables pertains to current and 2070 (2.6, 4.5, 6.0 and 8.5 RCP) bioclimatic time-frame. The blue colour represents niche stability, the green colour represents niche unfilling (the proportion of the native niche that does not overlap with the exotic niche), and the red colour represents niche expansion

that does not overlap with the exotic niche), and the red colour represents niche expansion (Ahmed et al. 2019). The size of these zones corresponds to the volume of their niche.

In terms of bioclimatic space, *I. oblongifolia* ellipsoidal niche had a larger hypervolume ($52.81 \times 10^4 \text{ }^\circ\text{C}\cdot\text{mm}^2$) during 2050 RCP 2.6, followed by current ($18.44 \times 10^3 \text{ }^\circ\text{C}\cdot\text{mm}^2$), and was the smallest during the 2050 RCP 6.0 ($23.74 \times 10^2 \text{ }^\circ\text{C}\cdot\text{mm}^2$). However, among the RCPs of 2070 projection, highest hypervolume ($13.32 \times 10^3 \text{ }^\circ\text{C}\cdot\text{mm}^2$) was quantified with RCP 6.0, followed by RCP 2.6 ($11.31 \times 10^3 \text{ }^\circ\text{C}\cdot\text{mm}^2$) and was least ($41.89 \times 10^2 \text{ }^\circ\text{C}\cdot\text{mm}^2$) with RCP 8.5. The influence of environmental variables on niche dynamics is indicated by their centroid values. Their proximity to the centroid indicates that they exert control over species suitability [61].

Values of centroid as well as range of different bio-climatic variables pertain to various time-frames and RCPs are presented in Table 6. With current BC, this species demonstrated the greatest niche expansion from its fundamental niche with annual precipitation (BC-12, i.e., water variable), while precipitation seasonality (BC-15, water availability), and maximum temperature of the warmest month (BC-5) was identified as a facilitator to maintain its fundamental niche areas. Further our data revealed that BC-5 and mean temperature of the wettest quarter was identified for its niche expansion with 2050 RCP 4.5 and 6.0 and 2070 RCPs 2.6, 4.5 and 8.5. While mean temperature of the warmest quarter (BC-10) was identified for niche expansion with RCP 8.5 and RCP 6.0 and 2050 and 2070, respectively. Precipitation seasonality was also identified as a major facilitator to maintain the fundamental niche of this species with all RCPs of 2070 and 8.5 of 2050 (Table 6). Temperature seasonality was the most important factor for niche expansion during 2050 RCP 2.6 and the precipitation of the driest month (BC-14) was identified as the most important governing factor to maintain its fundamental niche area with similar time period.

4 Discussion

The development of current habitat models and forecasting of the forthcoming distribution range of plant species are pivotal components in the creation of recuperation strategies for their indigenous habitats. To devise effective strategies for conservation and management, it is imperative that we acquire a greater understanding of the dispersal patterns of various species and the ramifications of climate change on their potential habitat suitability [62, 63]. The elucidation of the correlation between a species range and its susceptibility to extinction not only enhances our knowledge, but also furnishes us with the elemental parameters that demarcate the geographical expanse of said range [64, 65]. The ecological niches of numerous plant species have been modified due to the escalation of global temperatures in recent decades [66–68].

Regression and machine learning are the two main subfields of species distribution modeling (SDM) that use presence-absence data. Generalized linear models, generalized additive models, and multivariate adaptive regression splines (MARS) are all regression-based techniques. Artificial neural networks (ANNs), classification trees (CARTs), maximum entropy (MaxEnt) algorithms, genetic algorithms (GAs), and random forests (RFs) are all machine learning approaches. Our literature review indicated significant diversity in performance among the aforementioned methods, implying that predicting which sorts of model characteristics increase or do not improve model performance may be problematic. The optimum model will be chosen not just by how the model's assumptions relate to the assembly processes that shape a specific community, but also by other aspects such as the quantity, quality, and spatial organization of the data. Even if two data sets appear to be identical, they may be best modeled using distinct methodologies [11]. In a nutshell, the methodologies differ in terms of species data (absence/presence vs presence-only) and prediction parameters (mechanistic-physiological constraint vs empirical-climatic approach). The shortcomings of any model restrict its predictive power. Over-adjusted models are less likely to be useful when applied to new data [13]. Ensemble approaches, which are more accurate than individual methods, are one option. Thus, among the results of several algorithms, the best-fitting prediction of a species' future range may be used [36]. Ensemble modeling is provided as a method for improving model predictions and determining how comparable several model outcomes are. The essential premise of ensemble (or sometimes consensus) modeling is that different modeling outputs can be viewed as alternative states of the underlying distribution. This method combines projections from multiple models into a single, unified surface for averaging. Several plant species have been studied, including *Abies pindrow*, *Adenocarpus mannii*, *Angylocalyx oligophyllus*, *Baphia maxima*, *Berlinia grandiflora*, *Betula utilis*, *Bradypus variegatus*, *Cassia mannii*, *Cedrus deodara*, *Cynometra hankei*, *Dialium tessmannii*, *Oxybaphus himalaicus*, *Picea smithiana*, *Pinus wallichiana*, *Pseudovigna sulaensis*, *Psophocarpus scandens*, *Quercus ilex*, *R. delavayi*, *Rhododendron arboretum*, *Sclerocarya birrea*, *Thryothorus ludovicianus*, *Vigna comosa*, have demonstrated the predictive accuracy of ensemble techniques for ENM modeling [20].

The current study employed an ensemble machine learning approach to acquire knowledge regarding the distributional tendencies of *I. oblongifolia*, a leguminous species that has received comparatively less scholarly attention. Additionally, we sought to identify the determinants that exert an influence on the aforementioned tendencies across both temporal and spatial dimensions. In the past, there have been endeavours to conduct ecological niche modelling for following leguminous species: *Amorpha canescens* Pursh) *Dalea candida* Michx [69], *Cynometra hankei* Harms, *Pseudovigna sulaensis* R. Clark & Burgt, *Adenocarpus mannii* Hook fil., and *Berlinia grandiflora* (Vahl) Hutch. & Datzel [18]. In contrast to other studies pertaining to leguminous flora, our study distinguishes itself on account of its comprehensive scope, which encompasses a worldwide perspective. Furthermore, we have taken into account a variety of additional

Table 6 Values of centroid (and range) of different bio-climatic variables pertains to various time-frames and RCPs

Bioclimatic variables	Current	2050 RCPs				2070 RCPs			
		2.6	4.5	6.0	8.5	2.6	4.5	6.0	8.5
BC3 Iso	-	-	-	-	-	-	-	-	51.53 (39–57)
BC4 TempS	-	3969.94 (0–5)	-	-	-	-	-	-	-
BC5 (MatWam)	39.38 (30.79–43.34)	-	288.44 (26.5–31.8)	414.95 (36.2–46.6)	-	406.85 (32.8–44.3)	415.36 (33.3–45.6)	-	426.77 (37.0–48.2)
BC8 MeTWeQ	-	-	305.86 (27.8–34.0)	-	112.65 (18.9–34.8)	302.87 (26.8–33.7)	311.32 (28.0–34.3)	-	-
BC9 MeTDQ	-	-	-	260.59 (19.6–32.6)	-	-	-	-	-
BC10 MeTWaQ	-	-	-	-	340.05 (30.0–37.3)	-	-	339.90 (30.1–37.8)	-
BC11 MeTCQ	-	-	-	-	-	-	-	239.2 (13.1–29.9)	-
BC12 AnPr	357.61 (10–840)	-	-	-	-	-	-	-	-
BC14 PrDM	-	1.11 (0–50)	-	1.29 (0–53)	-	-	-	-	-
BC15 PrS	130.33 (79–176.9)	-	-	-	139.48 (89–190)	141.78 (85–196)	137.89 (90–185)	133.83 (87–181)	136.73 (90–182)
BC16 PrWeQ	-	285.04 (0–637)	-	-	-	-	-	-	-
BC17 PrDQ	-	-	3.21 (0–138)	-	-	-	-	-	-

variables, such as anthropogenic activities, as well as the application of nitrogen and phosphorus fertilizers, and the density of livestock populations.

The findings of this study indicate that climatic predictors exhibit greater efficacy than non-climatic predictors in enhancing the quality of models. Moreover, the ecological niches of this particular species are subject to the influence of energy-related parameters such as the maximum temperature attained during the warmest month, the mean temperature during the driest and warmest quarters, and hydrological variables such as the periodicity of precipitation and the amount of precipitation during the wettest quarter. Use of AUC, TSS and Kappa values for ensemble model performance specifically in SDM of plant species have been advocated in many previous studies [12, 59, 70, 71]. The AUC is extensively used to evaluate the accuracy of habitat suitability models, while TSS normalizes total accuracy [12, 72].

Our subsequent efforts involved an attempt to assess its inherent characteristics over temporal dimensions, encompassing a range of bioclimatic factors, and spatial dimensions, utilizing diverse suitability classifications, specifically pertaining to the African continent and the nation of India. In addition to the established criteria for weighting diverse suitability classes, the utilization of area as a denominator and number of polygons as a multiplier in the index provides insight into the higher indigenous score observed in India as compared to the African region within the optimum class. The empirical evidence indicates that this particular species tends to inhabit more contiguous and expansive regions throughout the African continent, whereas in India, it appears to be dispersed into several smaller meta-populations, which are defined as groups of geographically isolated populations of the same species. The aforementioned findings can be comprehended through the lens of the population viability analysis theory, which posits that a densely populated group is comparatively more vulnerable to alterations in their surroundings than a scattered one. Our results offer valuable insights into the criterion for this particular species. We observed that the African optimal area, under varying RCPs and climatic advancements, exhibited a smaller area in comparison to the current optimal area (Table 4), with the exception of 2050 RCP 2.6. This trend was also observed in moderate areas. However, the converse was true for marginal habitats, suggesting a transformation of high-quality habitats into low-quality ones. In the case of India (Table 5), a reversal of the aforementioned trend was observed, wherein the area under the optimum class exhibited an increase at RCP 2.6 and 4.5 in 2050, and RCP 2.6 and 8.5 in 2070.

Contrary to popular belief, the geographic distributions of species are not immutable, but rather fluid, and can undergo changes over time due to environmental fluctuations [73, 74]. The incorporation of the temporal dimension is a crucial aspect in the precise modelling of niches, as evidenced by the findings of Feng et al. [75]. Furthermore, it bears significant implications for the reproducibility of results. In the course of our research, we have expounded upon the ecological niches that have been discerned for *I. oblongifolia*, taking into account its evolutionary trajectory across various climatic time periods and its spatial distribution, with particular emphasis on its global propagation and subsequent concentration at the continental and national scales.

The current study entailed a meticulous examination of the threshold limits of diverse bio-climatic predictors, spanning various time frames and emission scenarios, with respect to the distribution patterns of this species. This was accomplished by means of algorithm-based modelling, specifically ROC curves, as well as by invoking the niche-centroid hypothesis, which posits that the optimal conditions of an ecological fundamental niche are conducive to heightened reproduction and survival rates among individuals. In essence, there exists an inverse correlation between the abundances of various species and their respective distances from the fundamental niche centroid [76, 77]. Thus, in our subsequent methodology, we discerned the efficacious predictors for upholding the fundamental and realized niche of species. An elevated centroid value is indicative of niche expansion, or the realization of a species' potential niche, coupled with a lower abundance. Conversely, a reduced centroid value suggests that a predictor is facilitating the maintenance of its fundamental niche areas, resulting in a higher species abundance [11, 78].

Based on our research, we know that Precipitation Seasonality (BC-15) is the most important factors that affect the optimum habitat suitability of this up to 150 for the current (Fig. 3a), 2050-RCP 8.5 (Fig. 3i), and 2070-RCPs 2.6 (Fig. 4a), 4.5 (Fig. 4c), and 8.5 (Fig. 4g). While, multidimensional space of resources availability (niche hypervolume) for this species indicates that with advancement of climatic-timeframe and RCPs, this parameter, shifted their range from 79–176.9 to 85–196 and hence it may be linked with changes in area under the optimum suitability class with above temporal and emission events. Our ellipsoid niche modelling further elaborates the range of these bioclimatic variables which were up to 637 mm and 26.5–31.8 °C, respectively. Based on such observation, we can link the increase in area under optimum category $7638.67 \text{ Sq. km} \times 10^3$ and $6443.37 \text{ Sq. km} \times 10^3$ under these two RCPs of 2050 as compared to current optimum area $5859.86 \text{ Sq. km} \times 10^3$. Similarly, mean temperature of driest quarter (BC-9) and mean temperature of warmest quarter (BC-10) were found to be most influential during RCP 6.0 of 2050 (with ROC 25 to 30 °C) and 2070 (with ROC 34 °C).

The livestock populations present in the habitats of *I. oblongifolia* exhibit a positive correlation with a greater area under the optimum class in comparison to various bioclimatic and RCPs. This phenomenon can be attributed to the fact that the frequency of livestock trampling leads to an increase in soil bulk densities, thereby promoting soil aeration and ultimately resulting in a higher probability of germination. This finding aligns with the research conducted by Mathur and Sundarmoorthy [31]. Further, *I. oblongifolia*, is propagated from seed. Flowering and fruiting occur from late September to roughly the middle of March [23] Fruits that were mature in April had a better germination rate than those that were immature in June. When heated to 35 °C, 82% of newly harvested seeds germinated, while only 15% of seeds stored for two years did so. Seed viability is affected by a number of variables, such as the success rate of germination and the length of time the seeds are kept in storage.

Further, highest area optimum suitability class was recorded with Global Human Modification of Terrestrial Ecosystem (GHMTE) and with Nitrogen and phosphorus fertilizer application rates (in kilogrammes per hectare). Such results indicate the facilitation behaviour these nutrients on the population dynamics of this species. However, the positive impact of GHMTE on the area under the optimum habitat class can be explain with the Intermediate Disturbance Hypothesis (IDH) and Dynamic Equilibrium Model (DEM). The IDH predicts large species numbers at intermediate levels of disturbance and the DEM predicts that the effect of disturbance depends on the level of productivity [79]. In conclusion, these classical ecological approaches suggested that intermediate disturbance frequency supports higher species population due to a mix of good colonizer and good competitor species. Population fluctuations in *I. oblongifolia* may be less sensitive to human developmental activities because it is typically found in sandy, arid, and semi-arid regions of the world, where biotic activities and other factors associated with habitat destruction are less common than in other, more densely populated and resource-rich regions.

5 Conclusion

The success of conservation and management efforts depends on researchers' ability to track species' movements and predict how climate change and other predictors will affect the suitability of existing habitats. Our analysis realized the importance of species-specific knowledge of eco-physiological processes, such as light energy budgeting and the relationship between moisture and plant performance, for accurate species distribution modelling. Arid and semiarid species can benefit greatly from this physiological essence because of the pulsing nature of the resource release in these environments. Such pulse (rain), inter-pulse (cold period), and non-pulse (summer) events often triggered significant eco-physiological changes within species and could be a deciding factor for the role of species within the plant community. Many questions remain unanswered in the regarding *I oblongifolia*. Despite this, the vast majority of research pertain to the pharmacological potential of this species. That means we can learn a lot about the ecology and physiology of this species (competition and interaction with other community associates). Using a global perspective, we are able to better understand the species' distribution, and we made an effort to draw a link between the species' likely reactions to both current and future predictors.

Acknowledgements Senior author thankful to the Director, ICAR-CAZRI for giving approval to him for attending training on R-Programming that enhance his working capacity using ENM modelling techniques. Miss Preet Mathur (JIET, Jodhpur, India) thankful to their Director for extending their academic help.

Author contributions MM conceived of the study's central hypothesis and its interpretation of results from various machine learning methods. PM performed numerous types of statistical analysis, employing a wide range of software, and transforming the results appropriately for interpretation and visualisation.

Funding Not available.

Availability of data and materials The data sets used and/or analysed during the current study are available from the corresponding author on reasonable request.

Declarations

Ethics approval and consent to participate Not applicable.

Consent for publication Not applicable.

Competing interests The authors declare that they have no competing interests.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Khan S, Verma S. Ensemble modelling to predict the impact of future climate change on the global distribution of *Olea europaea* subsp. *cuspidata*. *Front For Global Change*. 2022;5:977691.
- Ahmad R, Khuroo AA, Hamid M, Charle B, Rashid I. Predicting invasion potential and niche dynamics of *Parthenium hysterophorus* (Congress grass) in India under projected climate changes. *Biodivers Conserv*. 2019. <https://doi.org/10.1007/s10531-019-01775-y>.
- Malhi Y, Franklin J, Seddon N, Solan M, Turner MG, Field CB, Knowlton N. Climate change and ecosystems: threats, opportunities and solutions. *Philos Trans R Soc B*. 2020;375:20190104. <https://doi.org/10.1098/rstb.2019.0104>.
- Negi GCS, Singh P, Singh SP. Atmospheric warming associated phenological earliness does not increase the length of growing season in Himalayan Trees. *Forest Science*. 2021;67:694–700. <https://doi.org/10.1093/forsci/fxab040>.
- Román-Palacios C, Wiens JJ. Recent responses to climate change reveal the drivers of species extinction and survival. *Proc Natl Acad Sci USA*. 2020;117:4211–7. <https://doi.org/10.1073/pnas.1913007117>.
- Zhang K, Zhang Y, Jia D, Tao J. Species distribution modelling of *Sassafras tzumu* and implications for forest management. *Sustainability*. 2020;12:4132.
- Zhang Y, Liu X, Chen G, Hu G. Simulation of urban expansion based on cellular automata and maximum entropy model. *China Earth Sci*. 2020;63:701–12.
- Assefa A, Tibebe A, Bihon A, Dagnachew A, Muktar Y. Ecological niche modelling predicting the potential distribution of African horse sickness virus from 2020 to 2060. *Scientific Rep*. 2022;12:1748. <https://doi.org/10.1038/s41598-022-05826-3>.
- Ab Lah NZ, Yusop Z, Hashim M, Mohd Salim J, Numata S. Predicting the habitat suitability of *Melaleuca cajuputi* based on the maxent species distribution model. *Forests*. 2021;12:1449. <https://doi.org/10.3390/f12111449>.
- Shi X, Yin Q, Sang Z, Zhu Z, Jia Z, Ma L. Prediction of potentially suitable areas for the introduction of *Magnolia wufengensis* under climate change. *Ecol Indi*. 2021;127: 107762. <https://doi.org/10.1016/j.ecolind.2021.107762>.
- Mathur M, Mathur P, Purohit H. Ecological niche modelling of a critically endangered species *Commiphora wightii* (Arn.) Bhandari using bioclimatic and non-bioclimatic variables. *Ecol Proc*. 2023;12:8. <https://doi.org/10.1186/s13717-023-00423-2>.
- Becerra-López J, Ramírez-Bautista A, Romero-Méndez U, Pavón N, Sánchez-Rojas G. Effect of climate change on halophytic grasslands loss and its impact in the viability of *Gopherus flavomarginatus*. *Nat Conserv*. 2017;21:39–55. <https://doi.org/10.3897/natureconservat.21.13614>.
- Merow C, Smith M, Silander JA. Practical guide to maxent for modelling species' distributions: What it does, and why inputs and settings matter. *Ecography*. 2014;36:1058–69.
- Ranjitkar S, Xu J, Kindt SKK. Ensemble forecast of climate suitability for the Trans-Himalayan Nyctaginaceae species. *Ecol Model*. 2014;282:18–24.
- Ranjitkar S, Kindt R, Sujakhu NM, Hart R, Guo W, Yang X, Shrestha KK, Xu J, Luedeling E. Separation of the bioclimatic spaces of Himalayan tree rhododendron species predicted by ensemble suitability models. *Global Ecol Conserv*. 2014;1:2–12.
- Kindt R. Ensemble species distribution modelling with transformed suitability values. *Environ Modelling Soft*. 2018;100:136–45.
- Gilani H, Goheer MA, Ahmad H, Hussain K. Under predicted climate change: distribution and ecological niche modelling of six native tree species in Gilgit-Baltistan, Pakistan. *Ecol Indic*. 2020. <https://doi.org/10.1016/j.ecolind.2019.106049>.
- Oyebanji OO, Salako G, Nneji LM, Oladip SO, Bolarinwa KA, Chukwuma EC, Ayoola AO, Olagunju TE, Ighodalo DJ, Nneji IC. Impact of climate change on the spatial distribution of endemic legume species of the Guineo-Congolian forest, Africa. *Ecol Ind*. 2021. <https://doi.org/10.1016/j.ecolind.2020.107282>.
- Jinga P, Liao Z, Nobis MP (2021) Species distribution modelling that overlooks intraspecific variation is inadequate for proper conservation of marula (*Sclerocarya birrea*, Anacardiaceae). *Global Ecol Conserv* 2021. <https://doi.org/10.1016/j.gecco.2021.e01908>.
- Mathur P, Mathur M. Machine learning ensemble species distribution modelling of an endangered arid land tree *Tecomella undulata*: a global appraisal. *Arab J Geosci*. 2023;16:131. <https://doi.org/10.1007/s12517-023-11229-z>.
- Singh V, Pandey RP. *Ethnobotany of Rajasthan*. India: Scientific publishers, Jodhpur; 1998.
- Quattrocchi U. *CRC world dictionary of medicinal and poisonous plants: common names, scientific names, eponyms, synonyms, and etymology*. CRC Press; 2012.
- Singh JP, Venkatesan K, Patidar A, Kumar M, Swami S, Kumar M Saranya R, Patil NV (2022) *Indigofera oblongifolia* Forssk.- An important underutilized multi-use leguminous shrub of Indian hot arid region. https://assets.researchsquare.com/files/rs-1752860/v1_covered.pdf?c=1657131529.
- Bhandari MM (1990) *Flora of the Indian Desert*. MPS Repros, Jodhpur
- BSI. (2022) *Indigofera oblongifolia*. Botanical survey of India, Kolkata. <https://efloraIndia.gov.in>. Accessed 06 Apr 2022
- Lubbad MY, Al-Quraishy S, Dkhil MA. Antimalarial and antioxidant activities of *Indigofera oblongifolia* on *Plasmodium chabaudi*-induced spleen tissue injury in mice. *Parasitol Res*. 2015;114:3431–8. <https://doi.org/10.1007/s00436-015-4568-y>.
- GBIF.org (14 March 2023) GBIF Occurrence Download <https://doi.org/10.15468/dl.7mf5jt>

28. Sankara RK, Swamy RK, Kumar D, Singh A, Bhat KG (2019) Flora of Peninsular India. <http://peninsula.ces.iisc.ac.in/plants.php?name=Indigofera.oblongifolia>
29. Dkhil MA, Hafiz TA, Thagfan FA, Al-Shaeni EM, Mubarak MA, Khalil M, Abdel-Gaber R, Al-Quraishy S. *Indigofera oblongifolia* protect against trypanosomiasis- induced spleen injury. J Inf Public Health. 2019;12:660–5.
30. Mathur M (2005) *Ecology and Prospecting of some Medicinal Plants of Aphrodisiac Potential*. Ph.D. Thesis, Jai Narain Vyas University, Jodhpur, Rajasthan India. P.412.
31. Mathur M, Sundaramoorthy S. Patterns of herbaceous species richness and productivity along gradients of soil moisture and nutrients in the Indian Thar desert. J of Arid Environ. 2016;125:80–7.
32. Coban HO, Orucu OK, Arslan ES. MaxEnt modelling for predicting the current and future potential geographical distribution of *Quercus libani* Olivier. Sustainability. 2020. <https://doi.org/10.3390/su12072671>.
33. Kass JM, Vilela B, Aiello-Lammens ME, Muscarella R, Merow C, Anderson RP. Wallace: A flexible platform for reproducible modelling of species niches and distributions built for community expansion. Methods Ecol Evol. 2018;9:1151–6. <https://doi.org/10.1111/2041-210X.12945>.
34. Fick SE, Hijmans RJ. WorldClim 2: new 1km spatial resolution climate surfaces for global land areas. Int J Climatol. 2017;37(12):4302–15.
35. Hijmans RJ, Cameron SE, Parra JL, Jones PG, Jarvis A. Very high-resolution in-terpolated climate surfaces for global land area. Int J Climatol. 2005;25(15):1965–8.
36. Zhang Y, Tang J, Ren G, Zhao K, Wang X. Global potential distribution prediction of *Xanthium italicum* based on Maxent Model. Sci Rep. 2021. <https://doi.org/10.1038/s41598-021-96041-z>.
37. Chaturvedi RK, Joshi J, Jayaraman M, Bala G, Ravindranath NH. Multi-model climate change projections for India under representative concentration pathways. Curr Sci. 2012;103(7):791–802.
38. Kennedy CM, Oakleaf JR, Theobald DM, Baruch-Mordo S, Kiesecker J. Managing the middle: a shift in conservation priorities based on the global human modification gradient. Global Change Biol. 2019;25:811–26. <https://doi.org/10.1111/gcb.14549>.
39. Kennedy CM, Oakleaf JRD, Theobald M, Baruch-Mordo S, Kiesecker J (2020) Documentation for the global human modification of terrestrial systems. Palisades, NY: NASA Socioeconomic Data and Applications Centre (SEDAC). <https://doi.org/10.7927/jw1p-am22>.
40. Potter PN, Ramankutty EM, Bennett S, Donner D. Characterizing the spatial patterns of global fertilizer application and manure production. Earth Interac. 2010;14(002):1–22.
41. Gilbert M, Cinardi G, Da Re D, Wint William GR, Wisser D, Robinson, Timothy P (2022) Global sheep, goat and cattle distribution in 2015 (5 minutes of arc <https://doi.org/10.7910/DVN/VZOYHM>; <https://doi.org/10.7910/DVN/YYG6ET>; <https://doi.org/10.7910/DVN/LHBICE>). <https://dataverse.harvard.edu/dataverse/glw>
42. Pradhan P. Strengthening Maxent modelling through screening of redundant explanatory Bioclimatic Variables with Variance inflation factor analysis. Researcher. 2016;8(5):29–34.
43. Obiokara MC, Fourcade Y (2018) Climatic niche and potential distribution of *Tithonia diversifolia* (Hemsl.) A. Gray in Africa. PLoS ONE 13(9): e0202421.
44. Osorio-Olvera L, Yañez-Arenas C, Martínez-Meyer E, Peterson AT. Relationships between population densities and niche-centroid distances in North American birds. Ecol Lett. 2020;23:555–64. <https://doi.org/10.1111/ele.13453>.
45. Kumar S, Stohlgren TJ. MaxEnt modelling for predicting suitable habitat for threatened and endangered tree *Canacomyrica monticola* in New Caledonia. J Ecol Nat Environ. 2009;1:94–8.
46. Padalia H, Srivastava V, Kushwaha SPS. Modelling potential invasion range of alien invasion species, *Hyptis suaveolens* (L) in India: Comparison of MaxEnt and GARP. Ecol Infor. 2014;22:36–43.
47. Jijon JD, Gaudry KH, Constante J, Valencia C. Augmenting the spatial resolution of climate-change temperature projections for city planners and local decision makers. Environ Res Lett. 2021. <https://doi.org/10.1088/1748-9326/abf7f2>.
48. Schmitt S, Pouteau R, Justeau D, de Boissieu F, Brinbaum P. SSDM: An R package to predict distribution of species richness and composition based on stacked species distribution models. Methods Ecol Eval. 2017. <https://doi.org/10.1111/2041-210X.12841>.
49. Wood S. mgcv: Mixed GAM Computation Vehicle with Automatic Smoothness Estimation. 2019. <https://stat.ethz.ch/R-manual/R-devel/library/mgcv/html/OOIndex.html>
50. Breiman L. Random forests. Mach Learn. 2001;45:5–32.
51. Cutler F, Wiener R (2018) randomForest: Breiman and Cutler's random forests for classification and regression. 2018. <https://doi.org/10.1023/A:1010933404324>
52. Layola MRR, Semwal M, Rana TS, Nair NK (2022) Predicting potential suitable habitat for *Ensete glaucum* (Roxb.) Cheesman using Maxent modelling. Flora. <https://doi.org/10.1016/j.flora.2022.152007>.
53. Ripley B, Venables W. nnet: Feed-Forward Neural Networks and Multinomial Log-Linear Models. 2020. <https://cran.r-project.org/web/packages/nnet/nnet.pdf>.
54. Therneau T, Atkinson B, Port BR. rpart: Recursive Partitioning and Regression Trees. 2019. <https://cran.r-project.org/web/packages/rpart/rpart.pdf>.
55. Rajamanickam V, Babel H, Montano-Herrera L, Ehsani A, Stiefel F, Haider S, Presser B, Knapp B. About model validation in bioprocessing. Processes. 2021;9:961. <https://doi.org/10.3390/pr9060961>.
56. Wani IA, Verma S, Mushtaq S, Alsahli AA, Alyemeni MA, Tariq M, Pant S. Ecological analysis and environmental niche modelling of *Dactyloporhiza hatagirea* (D. Don) Soo: a conservation approach for critically endangered medicinal orchid. Saudi J Biol Sci. 2022;28:2109–22.
57. Wani ZA, Ridwan Q, Khan S, Pant S, Siddiqui S, Moustafa M, Ahmad AE, Yassin HM. Changing climatic scenarios anticipate dwindling of suitable habitats for endemic species of Himalaya- prediction of ensemble modelling using *Aconitum heterophyllum* as a model plant. Sustainability. 2022;14:8491. <https://doi.org/10.3390/su14148491>.
58. Xu D, Zhou Z, Wang R, Ye M, Pu B. Modelling the distribution of *Zanthoxylum armatum* in China with MaxEnt modelling. Global Ecol Conserv. 2019. <https://doi.org/10.1016/j.gecco.2019.e00691>.
59. Khan AM, Li Q, Saqib Z, Khan N, Habib T, Khalid N, Majeed M, Tariq A. MaxEnt modelling and impact of climate change on habitat suitability variations of economically important chilgoza pine (*Pinus gerardiana* Wall) in South Asia. Forests. 2022;13:715. <https://doi.org/10.3390/f130507150>.

60. Osorio-Olvera L, Lira-Noriega A, Soberon J, Townsend PA, Falcon M, Contrears-Diaz RG, Martinez-Meyer E, Barve V, Barve N. Ntbox: an R package with graphical user interface for modelling and evaluating multidimensional ecological niches. *Methods Ecol Evol.* 2020;11:1199–206. <https://doi.org/10.1111/2041-210X.13452>.
61. Nunez-Penichet C, Cobos ME, Soberon J. Non-overlapping climatic niches and biogeographic barriers explain disjunct distributions of continental *Urania* moths. *Front Biogeogr.* 2021;13(2): e52142.
62. Mushtaq S, Reshi ZA, Shah M, Charles B. Modelled distribution of an invasive alien plant species differs at different spatio-temporal scales under changing climate: a case study of *Parthenium hysterophorus* L. *Trop Ecol.* 2021;62:10.
63. Zhong Y, Xue Z, Jiang M, Liu B, Wang G. The application of species distribution modelling in wetland restoration: a case study in the Songnen Plain. *Northeast China Ecol Ind.* 2021;121: 107137.
64. Cardillo M, Mace GM, Gittleman JL, Jones KE, Bielby J, Purvis A. The predictability of extinction: Biological and external correlates of decline in mammals. *Proc R Soc Biol Sci.* 2008;275:1441–8.
65. Davies TJ, Purvis A, Gittleman J. Quaternary climate change and the geographic ranges of mammals. *Am Nat.* 2009;174:297–307.
66. Bellard C, Bertelsmeier C, Leadley P, Thuiller W, Courchamp F. Impacts of climate change on the future of biodiversity. *Ecol Lett.* 2012;15:365–77.
67. Telwala Y, Brook BW, Manish K, Pandit MK. Climate-induced elevational range shifts and increase in plant species richness in a himalayan biodiversity epicentre. *PLoS ONE.* 2013;8: e57103.
68. Palacios CR, John CW. Recent responses to climate change reveal the drivers of species extinction and survival. *Proc Natl Acad Sci USA.* 2020;117:4211–7.
69. Lyon NJ, Debinski DM, Rangwala I. Evaluating the utility of species distribution models in informing climate change-resilient grassland restoration strategy. *Front Ecol Evol.* 2019;7:33. <https://doi.org/10.3389/fevo.2019.00033>.
70. Allouche O, Tsoar A, Kadmon R. Assessing the accuracy of species distribution models: prevalence, kappa and the true skill statistic (TSS). *J App Ecol.* 2006;43:1223–32. <https://doi.org/10.1111/j.1365-2664.2006.01214.x>.
71. Jung JB, Park GE, Kim HJ, Huh JH, Um Y. Predicting the habitat suitability for angelica gigas medicinal herb using an ensemble species distribution model. *Forests.* 2023;14:592.
72. Moshobane M, Winter P, Middleton L. Record of naturalized *Ipomoea hederifolia* (Linnaeus 1759) (Convolvulaceae), Scarlet morning-glory in South Africa. *Biol Invasions Rec.* 2022;11:49–56. <https://doi.org/10.3391/bir.2022.11.1.05>.
73. Meier E, Kienast F, Pearman PB, Svenning JC, Thuiller W, Araújo MB, Guisan A, Zimmermann NE. Biotic and abiotic variables show little redundancy in explaining tree species distributions. *Ecography.* 2010;33:1038–48.
74. Page NV, Shanker K. Environment and dispersal influence changes in species composition at different scales in woody plants of the Western Ghats. *India J Veg Sci.* 2018;29:74–83.
75. Feng X, Park DS, Walker C, Peterson AT, Merow C, Papers M. A checklist for maximizing reproducibility of ecological niche models. *Nat Ecol Evol.* 2019;3:1382–94.
76. Martínez-Meyer E, Díaz-Porras D, Peterson AT, Yanez-Arenas C. Ecological niche structure and rangewide abundance patterns of species. *Biol Lett.* 2013;9:20120637. <https://doi.org/10.1098/rsbl.2012.0637>.
77. Machado-Stredel F, Coboas ME, Peterson AT. A simulation-based method for selecting calibration areas for ecological niche models and species distribution models. *Front Biogeogr.* 2021. <https://doi.org/10.21425/F5FBG48814>.
78. Yanez- Arenas C, Martin G, Osorio-Olvera L, Escobar-Lujan J, Castano-Quintero S, Chiappa-Carrara X, Martinez-Meyer E. The abundant niche-centroid hypothesis: key points about unfilled niches and the potential use of supra-specific modelling units. *Biodiv Inf.* 2020. <https://doi.org/10.17161/bi.v15i2.13218>.
79. Svensson JR, Lindegarth M, Jonsson PR, Pavia H. Disturbance-diversity models: what do they predict and how are they tested? *Proc Royal Soc B.* 2012;279:2163–70.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.