Modeling Earth Systems and Environment
https://doi.org/10.1007/s40808-022-01485-5

ORIGINAL ARTICLE

Predicting the effects of climate change on prospective Banj oak (*Quercus leucotrichophora*) dispersal in Kumaun region of Uttarakhand using machine learning algorithms

Zainab Khan[1] · Sk Ajim Ali[1] · Farhana Parvin[1] · Mohd Mohsin[2] · Syed Kausar Shamim[1] · Ateeque Ahmad[1]

Received: 15 March 2022 / Accepted: 6 August 2022

Abstract

All sort of vegetation is highly responsive to climatic factors and therefore distribution and redistribution of vegetation is bound to be affected by the change in the climatic conditions. The present episode of climate change is rapid in nature with it fastest temperature rise in Himalayas after poles on the earth, rendering vegetation of this region vulnerable to redistribution in space and time. Therefore, accurate modeling of the potential distribution of plants native to the Himalayan area is essential. Machine learning has improved the accuracy of species distribution models to a greater extent. The effects of climate change on the spread of Banj oak, a prominent tree species of the mid-Himalayas in Uttarakhand's Kumaun area, were simulated in this study. The generalized linear model (GLM), boosted regression tree (BRT), and maximum entropy (MaxEnt) were used to achieve this. The models' accuracies were calculated and compared. The accuracy was determined using the area under the curve (AUC) and receiver operating characteristics (ROC) curves. The MaxEnt model outperformed the rest two models and therefore it was utilized for modeling and prediction of potential distribution of Banj oak for the present and future. The results with higher accuracy (i.e., AUC > 0.95) model suggested that the areal expansion of potential distribution of Banj oak is going to crunch down by more than 1000 sq. km. as compared to today by the year of 2070, highlighting the gravity of climate change. This areal reduction of broadleaf tree is limited in the lower latitude. Higher altitudes were predicted to enjoy expansion of the aforesaid species. This study is a stand-alone contribution to the species distribution modeling of *Quercus leucotrichophora* in the mid-elevations of the Central Himalayas in India.

Keywords Climate change · Banj oak · Species distribution modeling · Generalized linear model · Boosted regression tree · Maximum entropy

Introduction

The most significant factor influencing the spread of plant communities is climate (Reu et al. 2011). Plants have long been recognized as being particularly sensitive to changes in climatic conditions through time and location (Hansen and Phillips 2015). Therefore, various climatic classifications are deeply embedded with the vegetal occurrence (Belda et al. 2014). Climate change has altered the floral diversity and distribution around the world (Jackson et al. 2009). This distribution and redistribution of vegetation under the impact of climatic conditions can be modeled using various techniques (Foley et al. 2000).

The hotly discussed problem of current climate change is thought to be having a far speedier influence on the spread of vegetation on the earth's surface (Kelly and Goulden 2008). Changing climatic conditions resulted in the redistribution of the occurrence of plant species all over the planet from oceans to terrestrial areas (Cramer et al. 2001). The Himalayan vegetation is especially prone to redistribution as this region is experiencing drastic climate changes and temperature rise is fastest in this region after poles (Panthi et al. 2020).

Because to changes in successional patterns, stand structure, regeneration, and distribution patterns, the forest cover

✉ Sk Ajim Ali
skajimali.saa@gmail.com; skajimali@myamu.ac.in

[1] Department of Geography, Faculty of Science, Aligarh Muslim University, Aligarh 202002, India

[2] Department of Civil Engineering, Faculty of Engineering and Technology, Zakir Husain College of Engineering, Aligarh Muslim University, Aligarh 202002, India

in the Himalayan woodlands is changing (Naudiyal and Schmerbeck 2018). According to a study, Banj oak regrowth is restricted and poor in the Central Himalayas, and the species is likely to be displaced by other species (Dhyani et al. 2020). The Banj oak woodlands have transformed and are fast decaying, according to recent studies (Chakraborty et al. 2018). An increase in forest fires as a result of rising temperatures, climate change, and routine resource exploitation by local communities are being blamed for the decline of Banj oak trees in this region (Verma and Garkoti 2019). Forest fires have been more severe and frequent in the region during the previous two decades, having a significant impact on the slow-growing Banj oak regeneration. Climate change is causing widespread invasive species to cause changes in ecosystem structure and function (Kernan 2015). The decrease and deterioration of the Banj oak forest is projected to disrupt the region's hydrological balances, resulting in water shortages, as seen by the rapid depletion of natural springsheds (Shrestha et al. 2018). Understanding the effects of climate change on the region's dominating Banj oak forests, as well as projecting the future, is critical for developing conservation policies that can help alleviate the region's climate change impacts.

One of the most dominant vascular plants of the Himalayan region is Banj oak. It contains the majority of the region's forest ecology (Singh et al. 2016). Generally, Banj Oak is found at 1500–2500 m altitude of the Himalayan region (Iqbal et al. 2019). It offers excellent habitat to various birds and animals of the sub-montane ecosystems (Shahabuddin 2018). Mature Banj oak trees retain water in their roots and hence are capable in sustaining springs providing water to water-deprived, rain fed subsistent agricultural communities. Due to the enormous ecosystem services provided by this tree, the proper study of its potential spatial distribution in changing climatic conditions is ecologically essential (Singh et al. 2014). Species distribution models (SDMs) help us understand the potential geographic areas for species based on causal effect.

Various studies on the spread of this important tree species have recently been undertaken using different techniques such as maximum entropy (Yang et al. 2013). Boosted regression Tree (BRT) was adopted by Golding and Purse (2016), and generalized linear model (GLM) was utilized by Barbet-Massin et al. (2012) and Lecocq et al. (2019). SDMs can be used for the prediction of potential distribution of this species with higher accuracy (Hernandez et al. 2006). Regression tree is also an accurate method for species distribution modeling (Yu et al. 2020). Mungi et al. (2018) used SDM to model the invasion of *L. camara* in the Himalayas. Sun et al. (2020) used MaxEnt for modeling the distribution of *Quercus* species in China, but they did not provide true value accuracy of the model. While Rathor et al. (2019) modeled combined potential occurrence of *Quercus leucotrichophora*, *Quercus semecarpifolia*, and *Quercus foribunda* in the western Himalayas ensembling GAM, GBM, BRT, CART and ANN. MaxEnt was used to model distribution of Banj oak by Bhandari et al. (2000) in the Garhwal region of Uttarakhand. MaxEnt was also used by Dhyani et al. (2020) to simulate the distribution of *Q. leucotrichophora* in the middle Himalayan area. But, as far as the Kumaun region is concerned, no significant work was been conducted for Kumaun explicitly. Saran et al. (2010), conducted research on *Quercus semecarpifolia* (Brown Oak) using MaxEnt where future niche model assumed + 1, + 2 °C temperature rise and 20 mm precipitation rise instead of more reliable future scenarios of GCMs. The study lacks the validity of adoption of used algorithm.

Through the extensive literature review, it was found that most of the previous studies eliciting potential habitat of *Q. leucotrichophora* are explicitly based on the wider regions and offer limited insight on the species adaptive behavior in the Kumaun. There is limited information on how different models behaved in response to bio climatic variables. However, studies conducted by Bhandari et al. (2000), Dhyani et al. (2018, 2020) on the issue of *Q. leucotrichophora* prodding speculations that this species has great capacity to redistribute itself in space in an attempt to adapt with the changing climatic which led us to contemplate significant spatial changes in the habitat of *Q. leucotrichophora* in Kumaun region in future.

The objective of the present study is to show current potential distribution of *Q. leucotrichophora* and predict future distribution under consistently changing climates in Kumaun region of Central Himalayas with best performing models. GLM, BRT, and MaxEnt algorithms are compared for model accuracy using AUC and ROC curve to determine the best model for predicting the effects of climate change on Banj oak dispersal in Kumaun region.

Study area

Kumaun is an administrative division of Uttarakhand in the central Himalayan region. It is split into six districts and covers a total area of approximately 21,064 km^2. With east longitudes of 80°10′–79°27′ and north latitudes of 30°48′ to 28°52′, the region ranges in altitude from 148 m in the foothills to 7150 m at the peaks (Fig. 1). Microclimatic areas are formed by abrupt altitudinal shifts from south to north. The natural vegetation of the region has a lot of variety across a short distance (Singh and Mal 2014). The region may resemble a temperate zone (over 2000 m height) or a tundra region depending on its height and temperature regimes (Singh 1992). Kumaun's average temperature ranges from 20 to 35 °C, with an average annual precipitation of 120 cm. However, compared to other mountainous locations, Kumaun is

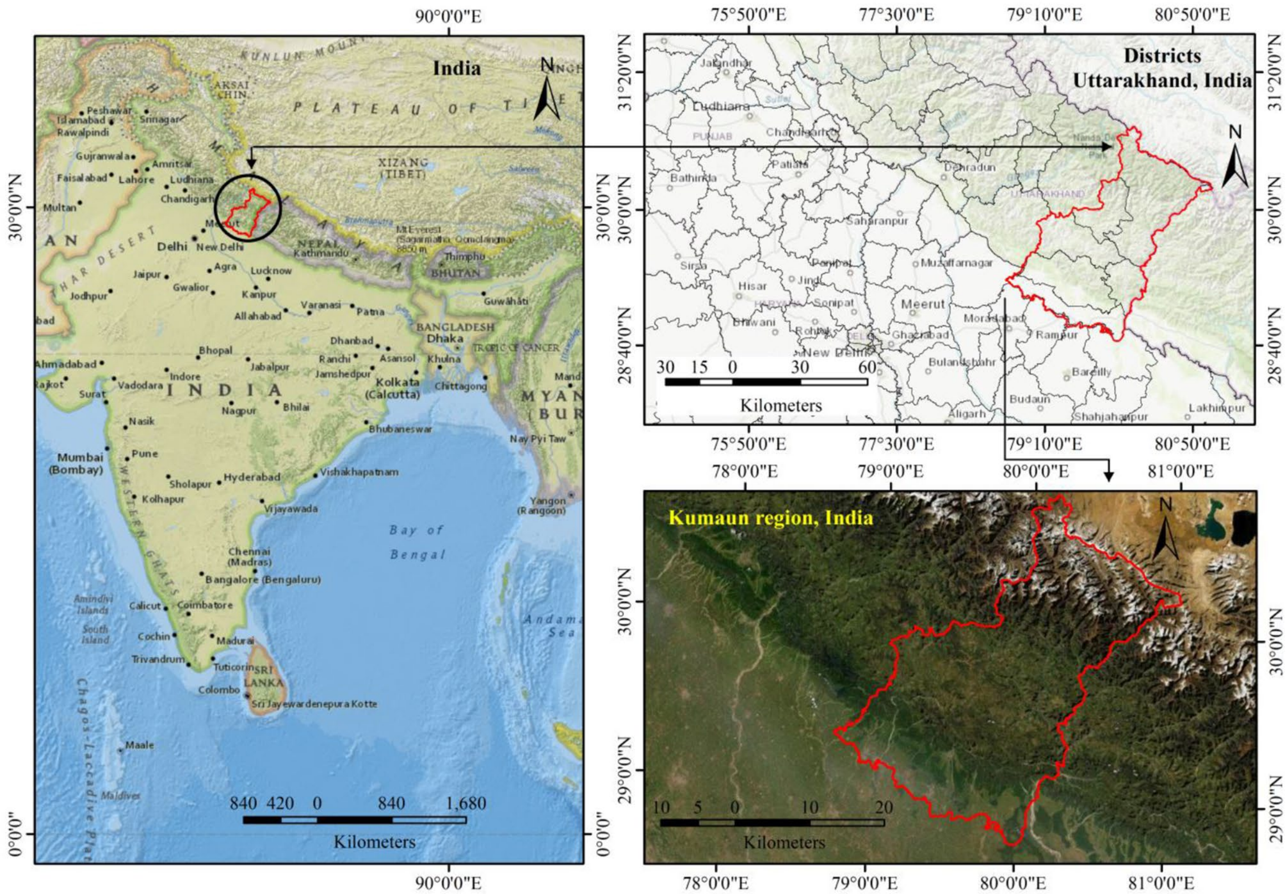

Fig. 1 Location of the study area

a climate change hotspot that has been badly damaged by global warming. It is part of the Himalayas and is the world's second fastest warming area. The rate of temperature rise on Earth has remained steady at 0.74 °C over the previous century, while in the Himalayas, it has grown by 0.6 °C per decade, for a total increase of 6 °C for the century (Jianchu et al. 2009). Precipitation has grown more unpredictable and has been steadily decreasing (Kumar et al. 2021). The region contains a large vegetal cover that is undergoing phenological changes as a result of climate change and plant upward shift (Singh and Mittal 2019).

Local populations in the research region rely on neighboring forests for their subsistence needs, and their economies are built on natural resources. Banj oak is a favored option for fire wood because of its high calorific content, nutritious leaves, and leaf litter for making farm yard manure owing to its fast decomposition (Fig. 2). Agriculture and livestock keeping are two key sources of income for people living at elevations of 1000–3000 m. In Himalayan villages, traditional farming relies heavily on mixed broadleaved trees (Rao and Pant, 2001). Traditional farming practises in the region are giving way to cash farming, resulting in agricultural expansion into forests and the loss of natural forests (Misra et al. 2009). A large number of hydroelectric projects have been planned, built, or are operating in the state, resulting in major biodiversity loss, deforestation, and fragmentation of natural forests, as well as local climate change (Dhyani et al. 2018).

Materials and methods

Data source

Species data

Data for this study were accessed from various online portals. The species occurrence data were collected from GBIF portal (https://www.gbif.org/data). There were 81 occurrence points for the Banj oak. The data were filtered with spatial geographic coordinates.

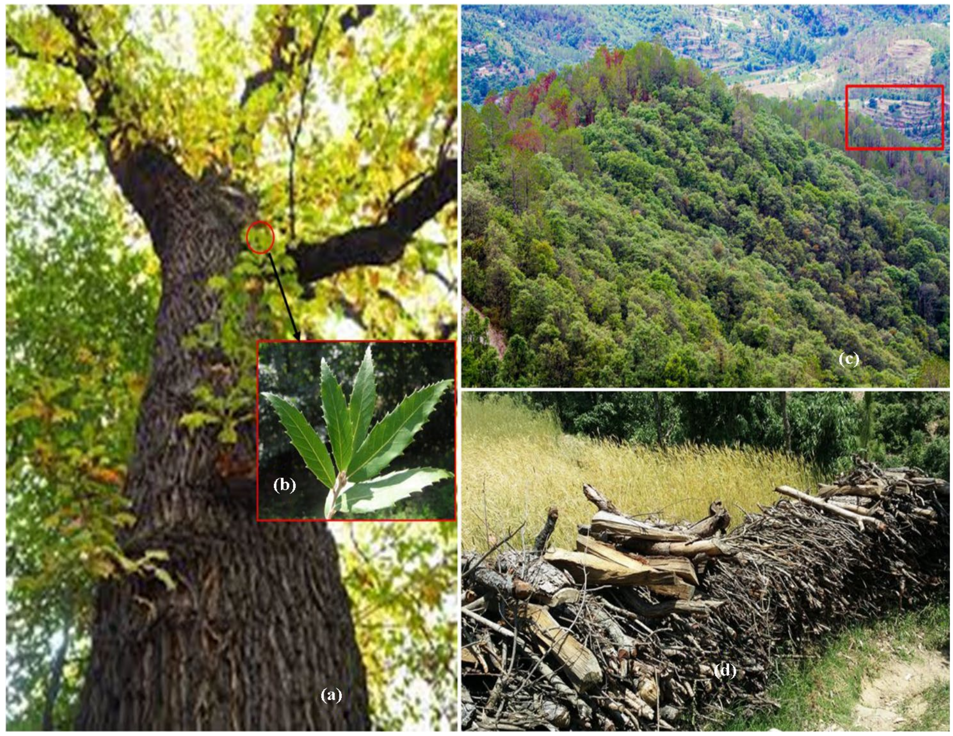

Fig. 2 Banj oak (*Quercus leucotrichophora*) in the study area: **a**, **b** stretch of mixed broadleaved Banj oak forest, **c** aerial view of Banj oak distribution shows the extension of agricultural activities, **d** deforestation of Banj oak by neighboring residents for their subsistence needs

Bioclimatic data

For bioclimatic factors, grid-based-downscaled (calibrated and bias-corrected) general circulation model (GCM) datasets were collected from Worldclim (https://www.worldclim.org/data/bioclim.html). These are simulated dataset widely used in climatic and ecological studies and species distribution modeling (Panagos et al. 2017; Marchi et al. 2019). The current dataset (1979–2013) was created by interpolation of mean monthly climate data from weather stations on different resolution grid while Future dataset is based on nine models i.e., BCC-CSM2-MR, CNRM-CM6-1, CNRM-ESM2-1, CanESM5, GFDL-ESM4, IPSL-CM6A-LR, MIROC-ES2L, MIROC6, MRI-ESM2-0. So, firstly all of the nineteen layers of the bioclimatic datasets at 2.5° spatial resolution were collected that represent current and future climatic condition (Title and Bemmels 2018). These variables represent potential determinants of vegetation distribution. For instance, Rundquist and Harrington (2000) stated that mean annual temperature is a major determinant of vegetation distribution. Temperature seasonality and maximum temperature of the warmest months are other deterring factors (Gwitira et al. 2014). Thus, these factors are important in assessing climatic conditions, identifying yearly trends, seasonality, and temperature and rainfall extremes, and so modeling species distribution (Chakraborty et al. 2016). The potential determinants of vegetation distribution including mean annual temperature are therefore valuable for ecological modeling. All the bioclimatic variables utilized in the study are listed in Table 1.

Methods

In this analysis, open access environment, RStudio was used for accessing and filtering the species occurrence and bioclimatic data sets for ease of usage and accuracy of results. R packages dismo, tidyr, sdm were utilized in this process. Species occurrence presence-only data after downloading was filtered for with geo-coordinates and then transformed into Spatial Points Data Frame. Pseudo absence data were generated by setting n to 1000. Before applying the models, the bioclimatic datasets were also filtered by multicollinearity using vifstep and vifcor functions in R and the threshold was kept 0.9. Setting n to 1000 random pseudo absences and 0.9 as a threshold was adopted as per Naimi et al. (2014). Utilization of vifstep and vifcor helped in reducing the bioclimatic data and only BICLD 2, BICLD 3, BICLD 4, BICLD 8, BICLD 9, BICLD 13, BICLD 14, BICLD 15, BICLD 18 and BICLD 19 were used for present species distribution modeling, whereas BICLD 2, BICLD 3, BICLD 4, BICLD 7, BICLD 8, BICLD 9, BICLD 13, BICLD 14, BICLD 15, BICLD 18, and BICLD 19 were used for future potential species distribution and geospatial analysis.

Table 1 Predictor bioclimatic variables used in this analysis

Predictor variables	Codes	Resolution
Presence-only point	N.A	N.A
Annual mean temperature	BICLD 1	2.5°
Mean diurnal range	BICLD 2	2.5°
Isothermality (BICLD 2/ BICLD 7) (×100)	BICLD 3	2.5°
Temperature seasonality (standard deviation × 100)	BICLD 4	2.5°
Max temperature of warmest month	BICLD 5	2.5°
Min temperature of coldest month	BICLD 6	2.5°
Temperature annual range (BICLD 5- BICLD 6)	BICLD 7	2.5°
Mean temperature of wettest quarter	BICLD 8	2.5°
Mean temperature of driest quarter	BICLD 9	2.5°
Mean temperature of warmest quarter	BICLD 10	2.5°
Mean temperature of coldest quarter	BICLD 11	2.5°
Annual precipitation	BICLD 12	2.5°
Precipitation of wettest month	BICLD 13	2.5°
Precipitation of driest month	BICLD 14	2.5°
Precipitation seasonality (coefficient of variation)	BICLD 15	2.5°
Precipitation of wettest quarter	BICLD 16	2.5°
Precipitation of driest quarter	BICLD 17	2.5°
Precipitation of warmest quarter	BICLD 18	2.5°
Precipitation of coldest quarter	BICLD 19	2.5°

Selection of model prediction

Species distribution models have successfully offer novel possibilities of studying past distribution of organisms and future adaptive behaviors. These models compliment fossil occurrences (Svenning et al. 2011). Several remarkable attempts have been undertaken by various researchers to forecast the spatial distribution of species using various statistical models, including conventional regression approaches, such as logistic regression, resource selection functions, and the Bayesian approach (Convertino et al. 2011), artificial neural networks, ecological niche component analysis, maximum entropy, classification and regression trees, random forests, generalized linear model, and boosted regression trees are examples of machine learning approaches that have lately come into use (Mainali et al. 2015a, b). Each of these approaches has its own set of benefits and drawbacks. For example, one of the most prominent methods, MaxEnt, is less sensitive to small samples (Wisz et al. 2008) but outperforms other methods in term of accuracy (Merow et al. 2013). GLM maintains simplicity in calculations (Graham et al. 2008). BRT has widely been used for predicting future potential species distribution modeling (Lecocq et al. 2019).

However, in the pursuit more robust provenance, scholars found out that different models have different levels of sensitivity to configuration settings and datasets (Hallgren et al. 2019). The justification over the selection of SDM algorithms uncommon and is minimalistic if present. In the current paper, BRT, GLM and MaxEnt, were utilized looking toward their preforming capabilities for predicting present and future *Q. leucotrichophora* in Kumaun seeking to contribute to the available literature on SDM modeling algorithms.

Generalized linear model (GLM)

GLM is a linear model, suitable for handling complicated dataset. Being regression-based model, it is simplistic in nature and oldest in application (Guisan et al. 2002). Graham (2008) first time made attempt to explain the application of this regression-based model extensively for habitat modeling. This model is useful for complicated species distribution modeling issues because it can cope with simultaneous impacts of numerous factors, including mixes of both categorical and continuous variables (Guisan et al. 1999). GLM has been used effectively in modeling of a mean response under non-standard environmental conditions (Khuri et al. 2006). Let, a linear bioclimatic vector or predictor is Xβ, where X denotes a vector of bioclimatic variables and β signifies a vector of estimated parameters plus an intercept (α) which is transformed through a link function to predict the response. The log link function, for example, is frequently employed with species count data and the Poisson distributional family. The model predictions (μ) can be estimated by (Eq. 1)

$$\mu = \exp_{(X\beta),} \text{ which is equivalent to } \log(\mu). \quad (1)$$

Non-linear relationships, on the other hand, can be modeled using quadratic, cubic, or higher-order algebraic expressions.

Boosted regression trees (BRT)

BRT is a relatively new approach that has sprung into the predictive modeling scene in the last decade (Carty 2011). BRT is flexible enough to express typical feature of the ecological and predictive modeling (Hallman and Robinson 2020). The modern regression trees are defined for their ecological applications (De'Ath 2007). This method uses two multiple regression trees to fit data to bring optimal predictive performance (Shabani et al. 2016). In case of predicting extreme climatic event, Pérez Navarro et al. 2019, found BRT performed quite well among GAM, MaxEnt and Mahalnobis. BRT has become one of the reliable SDMs in current research (Eskildsen et al. 2013). This model is well tested for response to colder regions (Hertzog et al. 2014) and has successfully been applied for future species distribution (Mainali et al. 2015a, b). The bioclimatic predictor space is partitioned into rectangles in the BRT model using a set of methods that select locations with the most homogenous predictor responses. The occurrence in BRT model can be expressed as Eq. 2 (Elith et al. 2008):

$$P\left(y = \frac{1}{z}\right) = f(z), \quad (2)$$

where, y is the logit and z is the covariate.

Maximum entropy (MaxEnt)

In statistical terminology, MaxEnt is a correlative model that estimates the ratio (z) $f_1(z)/f(z)$ using baseline sample and covariate data from the occurrence records. MaxEnt is one of highly tested SDM model that has been applied in varied environmental conditions (Kaky et al. 2020) and has been widely applied in futuristic modeling of plants (Phillips and Dubik 2008; Fois et al. 2018; Qin et al. 2020). Let assume that the data are presence-only data, a set of locations in the region L (a random sample). $y = 1$ which denotes presence of species and $y = 0$ denotes absence of species, and z represents environmental covariate or environmental conditions. Thus, define $f_1(z)$ to be the likelihood density of covariates across L (the entire region), and $f_1(z)$ to be the likelihood density of covariates across locations within L (Eq. 3).

$$P = (y = 1|\mathrm{z}) = f_1(z)P(y = 1)/f(z), \quad (3)$$

where , y =1 or indicating presence of species.

Model selection is the process of choosing two or more models with differing bioclimatic predictor criteria. All these three applied models produced different accuracies for the same bioclimatic variable (Aarts et al. 2012) and thus, their comparative accuracy assessment was performed in this study for selecting one more accurate model among them. In that case, AUC and ROC of the applied models i.e., BRT, GLM and MaxEnt were assessed before generating prediction maps. For sampling bootstrap method was adopted after Niami (2020) and an average sample ($n = 10$) were tested for each of the method.

Results

SDM has long been regarded as a valuable method for assessing the influence of climate variability on plant distribution, conservation, and management. In the present analysis, the SDMs predict potential species distribution in space and time using climatic scenarios and other information based on presence, absence, presence-only and pseudo-absence of species occurrence data. These models, as aforementioned, perform differently in different circumstances. Therefore, the accuracy of performance of these models is highly debated.

In the present study, it was found that the best performing model for the current scenarios is MaxEnt with an AUC = 0.99 (Fig. 3a), followed by BRT and GLM with an AUC of 0.98 and 0.97, respectively (Fig. 3b c). For the current scenarios, the MaxEnt model disproves the superiority of BRT over MaxEnt and GLM in this particular study. This is true for the results of future distribution of *Q. leucotrichophora*, where, for future scenarios, MaxEnt also outperformed the other two models with AUC value of 0.98 (Fig. 4a), followed by GLM and BRT with AUC value of 0.95 and 0.93, respectively (Fig. 4b, c). Therefore, MaxEnt can safely be used for modeling the distribution of *Q. leucotrichophora* in Kumaun.

The present and future potential distribution of *Q. leucotrichophora* was also demonstrated in the present analysis with the help of MaxEnt model, the highest performant model as per this study. The thresholds for areas classification were divided into very low, low, moderate, high, and very high classes with value of 0.05, 0.10, 0.20, 0.50, and 0.90 respectively for both present and future period till 2070. Classified maps (Fig. 5c, d) represent varying possibilities of occurrence based on the above-mentioned thresholds.

In the maps of potential species distribution, the probability of present and future potential occurrence of *Q. leucotrichophora* illustrated in Fig. 5a, b. The result confirms a defragmentation of the species in future. The northern limit of *Q. leucotrichophora* seemed to creep northward. The unclassified maps (Fig. 5a, b) on the continuous scale

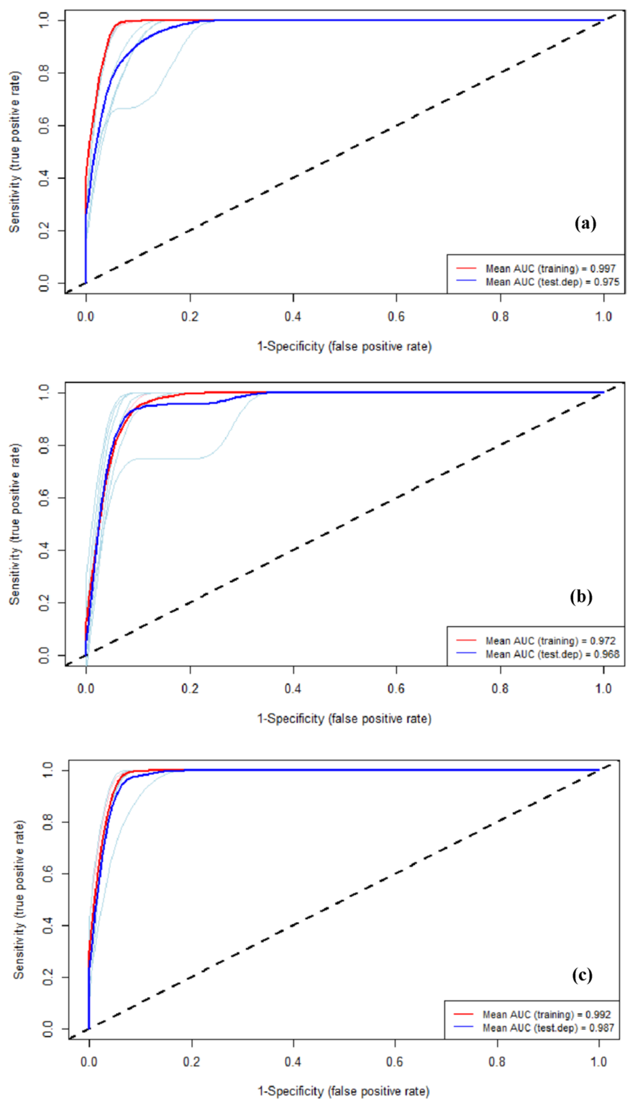

Fig. 3 ROC Plots for the current scenarios **a** BRT-ROC **b** GLM-ROC **c** MaxEnt-ROC

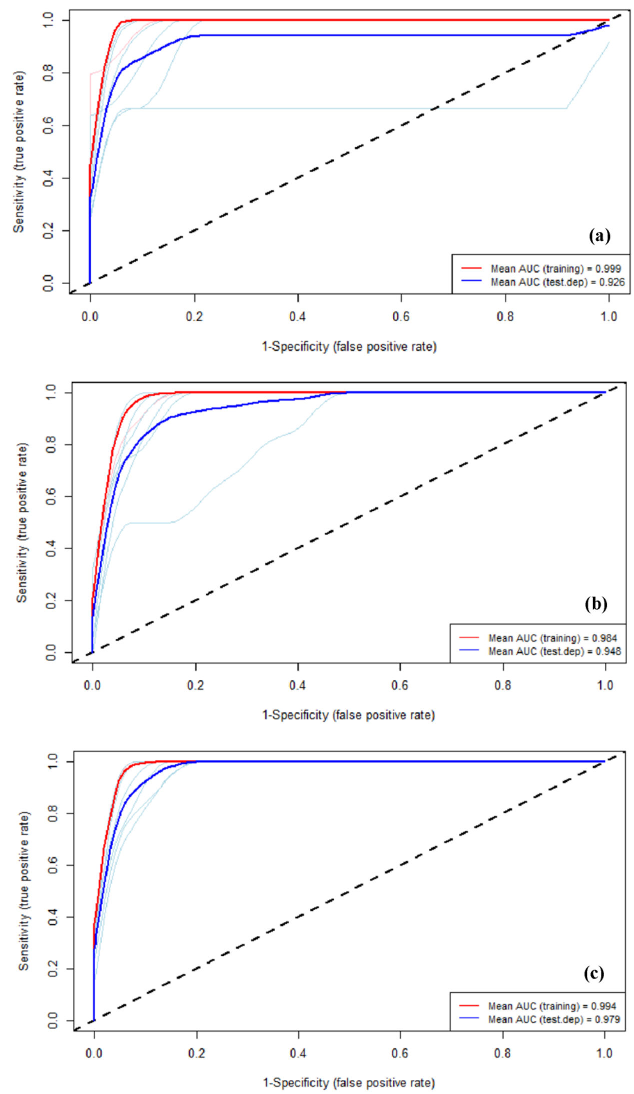

Fig. 4 ROC Plots for Future Scenarios **a** BRT-ROC **b** GLM-ROC **c** MaxEnt-ROC

the more suitable areas were expected to open up to the *Q. leucotrichophora* in the area, though with low probability of occurrence.

Table 2 represents the areal dynamics of occurrence of *Q. leucotrichophora*. It was made clear from the analysis of the table that areas under all the probabilities are going to expand in future, except the area under moderate probability. The areas with very high and very low probabilities of occurrence are going to expand significantly. Currently, approximately 2072 km^2 area with very high probabilities i.e., above 0.90 is there which increased in 2070 by about more than 250 km^2. The results are suggested that area under very low probabilities (0.05) increased by 500 km^2 by 2070 propounding that future climate is going to be more favorable for the *Q. leucotrichophora* at large, but there is a crisis in area in the southern limits of species.

Figure 6 represents the trend in areal dynamics of probabilities. The trend lines of very high (> 0.9) to high (0.5–0.9) probabilities seemed to run parallel to each other. The horizontal line advocates little to no change in area. The very low (< 0.05) and low (0.0.5–0.2) probabilities show an upward trend. While moderate probabilities (0.5 -0.9) represent declining trend in future species distribution.

Discussion

Numerous long-term climate studies in the Himalayan region, notably in Uttarakhand, have discovered considerable differences in air temperatures over time. Piyoosh and Ghosh

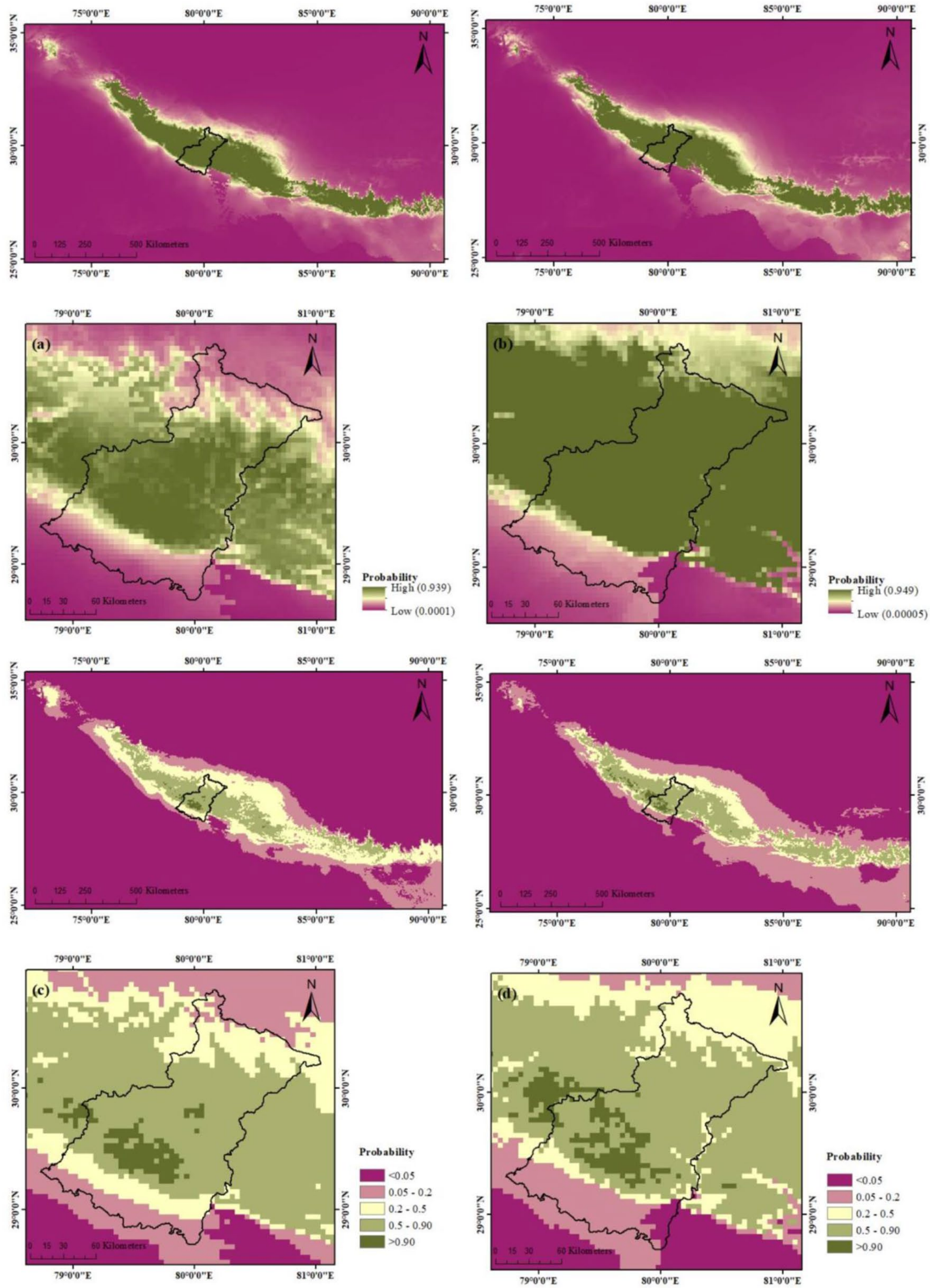

Fig. 5 Probability of occurrence

Table 2 Areal dynamics of probabilities of occurrence

Probabilities	Present area in (km^2)	Future area in (km^2)
Very low < 0.05	19	525
Low 0.05–0.2	2671	3820
Moderate 0.2–0.5	5205	2993
High 0.5–0.9	11,279	11,598
Very high > 0.9	2072	2313

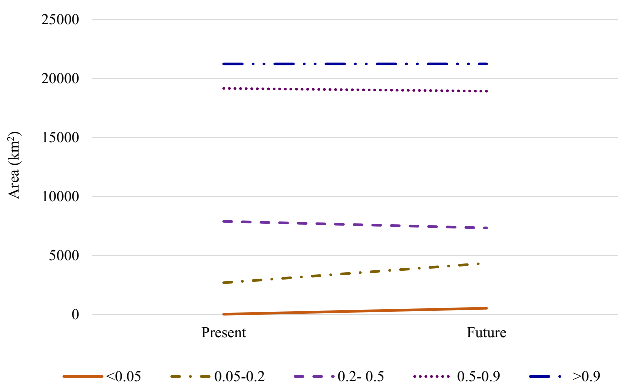

Fig. 6 Areal dynamics of occurrence probability

(2019) looked at fluctuations from 1901 to 2014, whereas Upadhyay et al. (2015) looked at trends from 1985 to 2013. It is well established that climate in the region is changing and temperatures warming (Sabin et al. 2020). However, determining the effect of this climatic inconsistency on a species' ability to migrate is strenuous (Bagaria et al. 2021). SDMs have been used to accurately obtain the result on several faunal (Khadka and James 2017; Singh et al. 2022) as well as floral species' (Manish 2022), risk assessment habitat alteration in Himalayan region and other areas due to climate change (Rather et al. 2022). *Q. leucotrichophora* is bitless focused species in spite of its huge ecological significance in the Himalayan ecosystem. Dhyani et al. (2020, 2021), conducted extensive study on *Q. leucotrichophora* in the central Himalayas. However, the research conducted by Dhyani is not sufficient for the smaller segment of the central Himalayas–Kumaun. In that case, present study offers reliability on the selected algorithms.

By evaluating the accuracy of multiple models to explain the implications of future climate change on the geographical distribution of *Q. leucotrichophora* in the Kumaun region, this study tries to develop a credible model. Previous research by Chakraborty et al. (2016) and Dhyani et al. (2020, 2021) did not compare MaxEnt against other machine learning techniques before predicting future potential distributions. MaxEnt modeling functioned effectively with climatic and other environmental data, producing in high AUC scores (> 0.9) in the targeted region, with the criterion of properly predicting the probable presence and pseudo-absence of the tree.

Climate change has had an influence on species' habitats, including worldwide forest ecosystems (Aitken et al. 2008; Koralewski et al. 2015), and will continue to do so in the Himalayan area (Mainali et al. 2015a, b). Many Himalayan plant species, including the keystone species *Q. leucotrichophora*, are projected to relocate to significantly higher altitudes due to climate change effect in this region (Braunisch et al. 2014; Du et al. 2014). However, there was a restriction in that previous research did not take into account trends in probabilities, resulting in a failure to infer areal information about probabilities.

Because we only used presence-only data from the tree species analyzed, the MaxEnt model, which is a correlative model based on Chakraborty et al. (2016), was used. It has been claimed that combining pseudo-absence with presence improves model predictions greatly. This might entail looking at additional SDMs that could be useful in exhibiting the probability dispersal of Himalayan tree species.

The quality of modeling of potential species distribution is affected by the input data utilized for modeling (occurrence records and environmental factors) (Rocchini et al. 2011). Temperature, for example, is an important indicator that should be incorporated as a predictor variable in potential mountain plant distribution models (Beaumont et al. 2008). It has also been observed that different models perform differently under different conditions (Chakraborty et al. 2016), with some models predicting geographic expansion of potential species distribution and others predicting contraction (Zhu et al. 2012), and therefore accuracy assessment of these models is required before using them for prediction.

It may give appropriate habitat suitability projections in species distribution maps, such as in MaxEnt, with the right choice of any model and meticulous parameter estimations, as well as an adequate sample size of species, geographically uniform occurrence data, and regularization parameters.

Conclusion

In present study, effects of climate change on the potential distribution of *Q. leucotrichophora* have been modeled and future distribution of the species in the Kumaun region of the central Himalayas is predicted. The study result demonstrated efficacy of MaxEnt method while sufficient consideration is given to other methods before adopting the MaxEnt for making predictions. The study propounds a shift of potential occurrence of *Q. leucotrichophora* toward the north of Kumaun and a crunch in the southern part in future compared with present scenario. The study also highlights the possibility of fragmentation in the habitat in the northern extremes. Study's comparative nature for basic SDM algorithms adds to the validity of prior research employing the MaxEnt approach for the Himalayan region. Aside from that, this study is unique in that it includes a trend analysis of probability.

This study, however, has typically utilized three algorithms for distribution modeling. For deeper study of the responsiveness algorithms, further combinations should be examined. Arguably, all SDMs have a major limitation that they do not represent actual occurrence rather potential occurrence of a given species. Therefore, this study does not signify immediate adaption *Q. leucotrichophora* in future neither does it assures current occurrence.

In conclusion, this study validates the MaxEnt method and offers a reliable model for assessing the effect of climate change on *Q. leucotrichophora*, as well as developing ground-level adaptive strategies to mitigate climate change's negative effects for the planning bodies and the ecological studies. It also provides information on the geographical scope of the chance of occurrence of *Q. leucotrichophora* and associated natural capital.

Author contributions ZK, SAA and FP prepared data, developed the methodology, analyzed, and wrote the original manuscript. MM and SKS critically reviewed the manuscript. AA read and revised the manuscript. All authors read and approved the final manuscript.

Funding No funding was received for conducting this study.

Data availability The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Declarations

Animal research (Ethics) Not applicable.

Consent to participate (Ethics) Not applicable.

Consent to publish (Ethics) Not applicable.

Clinical trials registration Not applicable.

Plant reproducibility Not applicable.

References

Aarts G, Fieberg J, Matthiopoulos J (2012) Comparative interpretation of count, presence–absence and point methods for species distribution models. Methods Ecol Evol 3(1):177–187

Aitken SN, Yeaman S, Holliday JA, Wang T, Curtis-McLane S (2008) Adaptation, migration or extirpation: climate change outcomes for tree populations. Evol Appl 1(1):95–111

Bagaria P, Thapa A, Sharma LK, Joshi BD, Singh H, Sharma CM, Chandra K (2021) Distribution modelling and climate change risk assessment strategy for rare Himalayan Galliformes species using archetypal data abundant cohorts for adaptation planning. Clim Risk Manag 31:100264

Barbet-Massin M, Jiguet F, Albert CH, Thuiller W (2012) Selecting pseudo-absences for species distribution models: how, where and how many? Methods Ecol Evol 3(2):327–338

Beaumont LJ, Hughes L, Pitman AJ (2008) Why is the choice of future climate scenarios for species distribution modelling important? Ecol Lett 11(11):1135–1146

Belda M, Holtanová E, Halenka T, Kalvová J (2014) Climate classification revisited: from Köppen to Trewartha. Clim Res 59(1):1–13

Bhandari BS, Mehta JP, Tiwari SC (2000) Dominance and diversity relations of woody vegetation structure along an altitudinal gradient in a montane forest of Garhwal Himalaya. J Trop For Sci 12(1):49–61. https://www.jstor.org/stable/23616403

Braunisch V, Coppes J, Arlettaz R, Suchant R, Zellweger F, Bollmann K (2014) Temperate mountain forest biodiversity under climate change: compensating negative effects by increasing structural complexity. PLoS ONE 9(5):e97718

Carty DM (2011) An analysis of boosted regression trees to predict the strength properties of wood composites. Retrieved from https://trace.tennessee.edu/cgi/viewcontent.cgi?article=2089&context=utk_gradthes&httpsredir=1&referer=. Accessed 12 Jan 2021

Chakraborty A, Joshi PK, Sachdeva K (2016) Predicting distribution of major forest tree species to potential impacts of climate change in the central Himalayan region. Ecol Eng 97:593–609

Chakraborty A, Saha S, Sachdeva K, Joshi PK (2018) Vulnerability of forests in the Himalayan region to climate change impacts and anthropogenic disturbances: a systematic review. Reg Environ Change 18(6):1783–1799. https://doi.org/10.1007/s10113-018-1309-7

Convertino M, Donoghue J, Chu-Agor ML, Kiker G, Munoz-Carpena R, Fischer R, Linkov I (2011) Anthropogenic renourishment feedback on shorebirds: a multispecies Bayesian perspective. Nat Proc 37(8):1184–1194. https://doi.org/10.1016/j.ecoleng.2011.02.019

Cramer W, Bondeau A, Woodward FI, Prentice IC, Betts RA, Brovkin V et al (2001) Global response of terrestrial ecosystem structure and function to CO2 and climate change: results from six dynamic global vegetation models. Glob Change Biol 7(4):357–373

De'Ath G (2007) Boosted trees for ecological modeling and prediction. Ecology 88(1):243–251

Dhyani S, Kadaverugu R, Dhyani D, Verma P, Pujari P (2018) Predicting impacts of climate variability on habitats of *Hippophae salicifolia* (D. Don) (Seabuckthorn) in Central Himalayas: Future challenges. Eco Inform 48:135–146. https://doi.org/10.1016/j.ecoinf.2018.09.003

Dhyani S, Kadaverugu R, Pujari P (2020) Predicting impacts of climate variability on Banj oak (*Quercus leucotrichophora* A. Camus) forests: understanding future implications for Central Himalayas. Reg Environ Change 20(4):1–13

Dhyani A, Kadaverugu R, Nautiyal BP, Nautiyal MC (2021) Predicting the potential distribution of a critically endangered medicinal plant Lilium polyphyllum in Indian Western Himalayan Region. Reg Environ Change 21(2):1–11

Du J, He Z, Yang J, Chen L, Zhu X (2014) Detecting the effects of climate change on canopy phenology in coniferous forests in semi-arid mountain regions of China. Int J Remote Sens 35(17):6490–6507

Elith J, Leathwick JR, Hastie T (2008) A working guide to boosted regression trees. J Anim Ecol 77(4):802–813

Eskildsen A, le Roux PC, Heikkinen RK, Høye TT, Kissling WD, Pöyry J, Luoto M (2013) Testing species distribution models across space and time: high latitude butterflies and recent warming. Glob Ecol Biogeogr 22(12):1293–1303

Fois M, Cuena-Lombraña A, Fenu G, Bacchetta G (2018) Using species distribution models at local scale to guide the search of poorly known species: review, methodological issues and future directions. Ecol Model 385:124–132

Foley JA, Levis S, Costa MH, Cramer W, Pollard D (2000) Incorporating dynamic vegetation cover within global climate models. Ecol Appl 10(6):1620–1632

Golding N, Purse BV (2016) Fast and flexible Bayesian species distribution modelling using Gaussian processes. Methods Ecol Evol 7(5):598–608

Graham JM (2008) The general linear model as structural equation modeling. J Educ Behav Stat 33(4):485–506

Graham CH, Elith J, Hijmans RJ, Guisan A, Townsend Peterson A, Loiselle BA, NCEAS Predicting Species Distributions Working Group (2008) The influence of spatial errors in species occurrence data used in distribution models. J Appl Ecol 45(1):239–247

Guisan A, Weiss SB, Weiss AD (1999) GLM versus CCA spatial modeling of plant species distribution. Plant Ecol 143(1):107–122

Guisan A, Edwards TC Jr, Hastie T (2002) Generalized linear and generalized additive models in studies of species distributions: setting the scene. Ecol Model 157(2–3):89–100

Gwitira I, Murwira A, Shekede MD, Masocha M, Chapano C (2014) Precipitation of the warmest quarter and temperature of the warmest month are key to understanding the effect of climate change on plant species diversity in Southern African savannah. Afr J Ecol 52(2):209–216

Hallgren W, Santana F, Low-Choy S, Zhao Y, Mackey B (2019) Species distribution models can be highly sensitive to algorithm configuration. Ecol Model 408:108719

Hallman TA, Robinson WD (2020) Comparing multi-and single-scale species distribution and abundance models built with the boosted regression tree algorithm. Landsc Ecol 35(5):1161–1174

Hansen AJ, Phillips LB (2015) Which tree species and biome types are most vulnerable to climate change in the US Northern Rocky Mountains? For Ecol Manag 338:68–83

Hernandez PA, Graham CH, Master LL, Albert DL (2006) The effect of sample size and species characteristics on performance of different species distribution modeling methods. Ecography 29(5):773–785

Hertzog LR, Besnard A, Jay-Robert P (2014) Field validation shows bias-corrected pseudo-absence selection is the best method for predictive species-distribution modelling. Divers Distrib 20(12):1403–1413

Iqbal M, Bilal SA, Iqbal A, Almzeb M, Saeed M, Mehreen R et al (2019) Distribution of *Quercus leucotrichophora* at different location and elevation in Jandool Valley, Northern Pakistan. Acta Ecol Sin 39(6):438–442

Jackson ST, Betancourt JL, Booth RK, Gray ST (2009) Ecology and the ratchet of events: climate variability, niche dimensions, and species distributions. Proc Natl Acad Sci 106(Supplement 2):19685–19692

Jianchu X, Shrestha A, Eriksson M (2009) Climate change and its impacts on glaciers and water resource management in the Himalayan Region. In: Assessment of Snow, Glaciers and Water Resources in Asia. International Hydrological Programme of UNESCO and Hydrology and Water Resources Programme of WMO. Koblenz, Germany, 44: 54

Kaky E, Nolan V, Alatawi A, Gilbert F (2020) A comparison between Ensemble and MaxEnt species distribution modelling approaches for conservation: a case study with Egyptian medicinal plants. Eco Inf 60:101150

Kelly AE, Goulden ML (2008) Rapid shifts in plant distribution with recent climate change. Proc Natl Acad Sci 105(33):11823–11826

Kernan M (2015) Climate change and the impact of invasive species on aquatic ecosystems. Aquat Ecosyst Health Manag 18(3):321–333

Khadka KK, James DA (2017) Modeling and mapping the current and future climatic-niche of endangered Himalayan musk deer. Eco Inf 40:1–7

Khuri AI, Mukherjee B, Sinha BK, Ghosh M (2006) Design issues for generalized linear models: a review. Stat Sci 21(3):376–399

Koralewski TE, Wang HH, Grant WE, Byram TD (2015) Plants on the move: assisted migration of forest trees in the face of climate change. For Ecol Manag 344:30–37

Kumar U, Singh S, Bisht JK, Kant L (2021) Use of meteorological data for identification of agricultural drought in Kumaon region of Uttarakhand. J Earth Syst Sci 130(3):1–13

Lecocq T, Harpke A, Rasmont P, Schweiger O (2019) Integrating intraspecific differentiation in species distribution models: consequences on projections of current and future climatically suitable areas of species. Divers Distrib 25(7):1088–1100

Mainali J, All J, Jha PK, Bhuju DR (2015a) Responses of montane forest to climate variability in the central Himalayas of Nepal. Mt Res Dev 35(1):66–77

Mainali KP, Warren DL, Dhileepan K, McConnachie A, Strathie L, Hassan G, Parmesan C (2015b) Projecting future expansion of invasive species: comparing and improving methodologies for species distribution modeling. Glob Change Biol 21(12):4464–4480

Manish K (2022) Medicinal plants in peril due to climate change in the Himalaya. Eco Inform 68:101546

Marchi M, Sinjur I, Bozzano M, Westergren M (2019) Evaluating WorldClim Version 1 (1961–1990) as the baseline for sustainable use of forest and environmental resources in a changing climate. Sustainability 11(11):3043

Merow C, Smith MJ, Silander JA Jr (2013) A practical guide to MaxEnt for modeling species' distributions: what it does, and why inputs and settings matter. Ecography 36(10):1058–1069

Misra S, Maikhuri RK, Dhyani D, Rao KS (2009) Assessment of traditional rights, local interference and natural resource management in Kedarnath Wildlife Sanctuary. Int J Sust Dev World 16(6):404–416. https://doi.org/10.1080/13504500903332008

Mungi NA, Coops NC, Ramesh K, Rawat GS (2018) How global climate change and regional disturbance can expand the invasion risk? Case study of Lantana camara invasion in the Himalaya. Biol Invasions 20(7):1849–1863

Naimi B, Hamm NA, Groen TA, Skidmore AK, Toxopeus AG (2014) Where is positional uncertainty a problem for species distribution modelling? Ecography 37(2):191–203

Naudiyal N, Schmerbeck J (2018) Impacts of anthropogenic disturbances on forest succession in the mid-montane forests of Central Himalaya. Plant Ecol 219(2):169–183

Niami B (2020) sdm R package: species distribution modelling [Video]. YouTube: Biogeoinformatics

Panagos P, Ballabio C, Meusburger K, Spinoni J, Alewell C, Borrelli P (2017) Towards estimates of future rainfall erosivity in Europe based on REDES and WorldClim datasets. J Hydrol 548:251–262

Panthi S, Fan ZX, van der Sleen P, Zuidema PA (2020) Long-term physiological and growth responses of Himalayan fir to environmental change are mediated by mean climate. Glob Change Biol 26(3):1778–1794

Pérez Navarro MÁ, Sapes G, Batllori E, Serra-Diaz JM, Esteve MA, Lloret F (2019) Climatic suitability derived from species distribution models captures community responses to an extreme drought episode. Ecosystems 22(1):77–90

Phillips SJ, Dudík M (2008) Modeling of species distributions with Maxent: new extensions and a comprehensive evaluation. Ecography 31(2):161–175

Piyoosh AK, Ghosh SK (2019) Identification and analysis of recent temporal temperature trends for Dehradun, Uttarakhand, India. Meteorol Atmos Phys 131(4):863–882. https://doi.org/10.1007/s00703-018-0608-3

Qin A, Jin K, Batsaikhan ME, Nyamjav J, Li G, Li J, Xiao W (2020) Predicting the current and future suitable habitats of the main dietary plants of the Gobi Bear using MaxEnt modeling. Glob Ecol Conserv 22:e01032

Rao KS, Pant R (2001) Land use dynamics and landscape change pattern in a typical micro watershed in the mid elevation zone of central Himalaya, India. Agr Ecosyst Environ 86(2):113–124

Rather ZA, Ahmad R, Khuroo AA (2022) Ensemble modelling enables identification of suitable sites for habitat restoration of threatened biodiversity under climate change: a case study of Himalayan Trillium. Ecol Eng 176:106534

Rathore P, Roy A, Karnatak H (2019) Assessing the vulnerability of Oak (Quercus) forest ecosystems under projected climate and land use land cover changes in Western Himalaya. Biodivers Conserv 28(8):2275–2294

Reu B, Proulx R, Bohn K, Dyke JG, Kleidon A, Pavlick R, Schmidtlein S (2011) The role of climate and plant functional trade-offs in shaping global biome and biodiversity patterns. Glob Ecol Biogeogr 20(4):570–581

Rocchini D, Hortal J, Lengyel S, Lobo JM, Jimenez-Valverde A, Ricotta C et al (2011) Accounting for uncertainty when mapping species distributions: the need for maps of ignorance. Prog Phys Geogr 35(2):211–226

Rundquist BC, Harrington JA Jr (2000) The effects of climatic factors on vegetation dynamics of tallgrass and shortgrass cover. Geocarto Int 15(3):33–38

Sabin TP, Krishnan R, Vellore R, Priya P, Borgaonkar HP, Singh BB, Sagar A (2020) Climate change over the Himalayas. In: Krishnan R, Sanjay J, Gnanaseelan C, Mujumdar M, Kulkarni A, Chakraborty S (eds) Assessment of climate change over the Indian region. Springer, Singapore, pp 207–222

Saran S, Joshi R, Sharma S, Padalia H, Dadhwal VK (2010) Geospatial modeling of Brown oak (*Quercus semecarpifolia*) habitats in the Kumaun Himalaya under climate change scenario. J Indian Soc Remote Sens 38(3):535–547

Shabani F, Kumar L, Ahmadi M (2016) A comparison of absolute performance of different correlative and mechanistic species distribution models in an independent area. Ecol Evol 6(16):5973–5986

Shahabuddin G (2018) Birds, Forests and Development in Himalayan Oak forests: a study in progress. Retrieved from https://www.conservationindia.org/articles/birds-forests-and-development-inhimalayan-oak-forests-a-study-in-progress. Accessed 11 Mar 2021

Shrestha RB, Desai J, Mukherji A, Dhakal M, Kulkarni H, Mahamuni K et al (2018) Protocol for reviving springs in the Hindu Kush Himalayas: a practitioner's. manual International Centre for Integrated Mountain Development (ICIMOD)

Singh JS (1992) Forests of Himalaya: structure, functioning and impact of man. CGyanodaya Prakashan

Singh RB, Mal S (2014) Trends and variability of monsoon and other rainfall seasons in Western Himalaya, India. Atmos Sci Lett 15(3):218–226

Singh N, Mittal A (2019) Response of phenological events of *Aesculus indica* Colebr. to climate change along an altitudinal gradient in Kumaun Himalaya, Uttarakhand. Int J Environ 8(1):1–16

Singh V, Thadani R, Tewari A, Ram J (2014) Human influence on banj Oak (*Quercus leucotrichophora*, A. Camus) forests of Central Himalaya. J Sustain for 33(4):373–386

Singh G, Padalia H, Rai ID, Bharti RR, Rawat GS (2016) Spatial extent and conservation status of Banj oak (*Quercus leucotrichophora* A. Camus) forests in Uttarakhand, Western Himalaya. Trop Ecol 57(2):255–262

Singh AP, Chandra A, De K, Uniyal VP, Sathyakumar S (2022) Decreasing potential suitable habitat of bumble bees in the Great Himalayan National Park Conservation area. Orient Insects. https://doi.org/10.1080/00305316.2022.2040631

Sun S, Zhang Y, Huang D, Wang H, Cao Q, Fan P, Wang R (2020) The effect of climate change on the richness distribution pattern of oaks (*Quercus* L.) in China. Sci Total Environ 744:140786

Svenning JC, Fløjgaard C, Marske KA, Nógues-Bravo D, Normand S (2011) Applications of species distribution modeling to paleobiology. Quatern Sci Rev 30(21–22):2930–2947

Title PO, Bemmels JB (2018) ENVIREM: an expanded set of bioclimatic and topographic variables increases flexibility and improves performance of ecological niche modeling. Ecography 41(2):291–307

Upadhyay RG, Ranjan R, Negi PS (2015) Climatic variability and trend at Ranichauri (Uttarakhand). J Agrometeorol 17(2):241

Verma AK, Garkoti SC (2019) Population structure, soil characteristics and carbon stock of the regenerating banj oak forests in Almora, Central Himalaya. For Sci Technol 15(3):117–127. https://doi.org/10.1080/21580103.2019.1620135

Wisz MS, Hijmans RJ, Li J, Peterson AT, Graham CH, Guisan A, NCEAS Predicting Species Distributions Working Group (2008) Effects of sample size on the performance of species distribution models. Divers Distrib 14(5):763–773

Yang XQ, Kushwaha SPS, Saran S, Xu J, Roy PS (2013) Maxent modeling for predicting the potential distribution of medicinal plant, *Justicia adhatoda* L. in Lesser Himalayan foothills. Ecol Eng 51:83–87

Yu H, Cooper AR, Infante DM (2020) Improving species distribution model predictive accuracy using species abundance: application with boosted regression trees. Ecol Model 432:109202

Zhu K, Woodall CW, Clark JS (2012) Failure to migrate: lack of tree range expansion in response to climate change. Glob Change Biol 18(3):1042–1052