Ecological Informatics 40 (2017) 1–7

Contents lists available at ScienceDirect

Ecological Informatics

journal homepage: www.elsevier.com/locate/ecolinf

Modeling and mapping the current and future climatic-niche of endangered Himalayan musk deer

Kapil K. Khadka*, Douglas A. James

Department of Biological Sciences, University of Arkansas, Fayetteville, AR 72701, United States

ARTICLE INFO

Keywords:
Maxent
GARP
Moschidae
Himalayas
Conservation

ABSTRACT

Identification of geographical space enveloped by suitable climatic conditions (i.e., climatic niche) that support species survival over space and time is crucial in conservation biogeography. Numerous algorithms (e.g., Maxent, GARP) with increasing accuracy have been devised and are being employed to overcome the challenges of forecasting climatic niche of species with incomplete information. The current study was conducted to map the distribution of current and future climatic niche of endangered Himalayan musk deer, a species endemic to Asia. Maxent and GARP modeling algorithms were individually employed to forecast current and future climatic niche of the species using randomly collected occurrence records of the species and bioclimatic variables with 30″ resolution from 'WorldClim' datasets. Both the modeling processes performed optimally with regard to AUC and TSS values and forecasted an increase/expansion of climatically-suitable geographical space in the future. A final climatic niche distribution map was produced by combining the binary maps generated from each of the processes to produce a relatively realistic and potentially accurate distribution of climatic niche of the species over space and time. Conservation of forecasted suitable geographical space is recommended and future survey efforts for potentially unexplored populations of the species in the forecasted suitable area are suggested.

1. Introduction

Identification and management of geographical space that supports species' survival is a key to conservation of wildlife in their natural habitats. With biodiversity increasingly being threatened or endangered with extinction from a wide array of anthropogenic disturbances (Millennium Ecosystem Assessment, 2005), the challenges of finding areas that are environmentally conducive to the survival of little known species, have intensified among conservationists. Moreover, the projected climate change and its varying effects on biodiversity over space and time has further amplified the challenge of locating climatically-suitable areas (i.e., climatic niche) in the future for species threatened with extinction (Pounds et al., 1999; Thomas et al., 2004; van Gils et al., 2016). The development of advanced computational algorithms with increasing accuracy have helped in modeling the distribution of species and mapping of potential current and future (e.g., under climate change scenario) environmentally suitable space with wide applications in conservation biogeography (see Guisan and Thuiller, 2005).

Numerous algorithms for modeling techniques of varying applications with data type and availabilities have been developed for species distribution modeling (SDM) (see Elith et al., 2006). All these techniques merely establish relationships between species' known occurrences or absences with environmental characteristics of concern at those geographical space and use those relationships to interpolate between the known occurrences and extrapolate in novel areas or scenarios to forecast a suitable area that support species' survival, under the assumption of niche conservatism and/or stationarity (Holt and Gaines, 1992; Petitgas, 2001). Of numerous modeling algorithms, maximum entropy (Maxent) (Phillips et al., 2006) and genetic algorithm for rule set-based prediction (GARP) (Stockwell and Peters, 1999) have been widely used in conservation biogeography for modeling and mapping the geographical range distribution of species with relatively good success. A wide and increasing applications of these techniques in distribution modeling is probably due to their requirement of species' occurrence records only, which is usually the case with rare, elusive, and declining species with relatively incomplete information about their absences, but demanding an immediate conservation concern.

Himalayan musk deer (*Moschus leucogaster*) and alpine musk deer (*Moschus chrysogaster*), in particular, are confined to high-altitude forests of Bhutan, northern India, Pakistan, Nepal, and China along the Hindukush Himalaya (Green, 1986; Grubb, 2005; Yang et al., 2003). Owing to their small population size and geographic range, the species have been listed in Appendix I of CITES and as endangered in International Union for Conservation of Nature (IUCN) red list. Taking

* Corresponding author.
E-mail addresses: kkkhadka@email.uark.edu (K.K. Khadka), djames@uark.edu (D.A. James).

http://dx.doi.org/10.1016/j.ecoinf.2017.04.009
Received 14 November 2016; Received in revised form 14 April 2017; Accepted 25 April 2017
Available online 29 April 2017
1574-9541/

expert-based range map of IUCN red list as a reference, the species of concern in this study is treated as *Moschus leucogaster*, although both the species are interchangeably treated as Himalayan musk deer and/or alpine musk deer. It is probably due to their small population size, elusive behavior, overlapping habitat, morphological similarities, and lack of genetic studies for species differentiation. Populations of Himalayan musk deer are declining primarily due to habitat loss and overexploitation (Timmins and Duckworth, 2015; Yang et al., 2003), although estimates of the current population size of the species is largely unknown. Moreover, studies on the species so far are scattered and largely locally confined to a small geographical scale, and thus demands for exploration and identification of climatically suitable areas in its whole range that potentially aids in conservation of the species. Also, the effect of predicted climate change on spatial distributions and range-shifts of the species is unknown. Hence, the current study models and maps the current climatically-suitable habitat (i.e., climatic niche) of the species, and attempts to predict the distribution of future geographic range under projected climate change scenario.

2. Methods

2.1. Data preparation

A total of 85 unique geographic coordinates (i.e., latitude and longitude) of species occurrences based on direct observation, fecal pellets, footprints, and resting sites were collected from Nepal, Bhutan, India, and Pakistan. These occurrence data were randomly collected in the potential habitat of the species in each country between 2013 and 2015. When the residuals of the model, after using all the available occurrence points in modeling, were found to be auto-correlated (Moran's I = 0.58, P = 0), they were filtered out to 52 independent occurrence points by spacing them to a minimum of 6 km apart (see Supplementary) as suggested by variogram plot (Dorman et al., 2007). The species occurrence points used for modeling are assumed to represent a full range of climatic conditions in the species' range as those were collected across summer and winter seasons. For predictors, 19 bioclimatic variables with a resolution of 30″ (i.e., ~1 km spatial resolution) from 'WorldClim' datasets (www.worldclim.org, Hijmans et al., 2005) for two time periods, i.e. 'current' and 'future' were used. 'WorldClim' database consists of climate surfaces for global land areas (except Antarctica) interpolated using the thin-plate smoothing spline of observed climates at weather stations. For future scenario, the database consists of projected climate for the years 2050 and 2070, with four different levels of greenhouse gas scenarios, i.e. Representative Concentration Pathways (RCPs). Because of varying level of greenhouse gas concentrations predicted for future and their inherent effect on climate, average climatic surfaces data from three randomly selected "Global Circulation Models" for 3 scenarios (scenario 1 = RCP2.6, scenario 2 = RCP4.5, scenario 3 = RCP8.5) for the year of 2050 were used for projecting the future geographic range of the species. 'scenario 1' (i.e., scenario that targets to limit the increase of global mean temperature to 2 °C, see van Vuuren et al., 2011) and 'scenario 3' (i.e., scenario without climate mitigation target, see Riahi et al., 2011) represent two extreme cases of radiative forcing while 'scenario 2' represent a modest level (i.e., scenario under climate policy) between 'scenario 1' and 'scenario 2'. Spearman's correlation coefficients among the 19 bioclimatic variables in the database were determined, and when the correlation coefficient between the variables was found to be $\geq |0.9|$, only one variable from a set of highly correlated variables was used to reduce the problems due to multi-collinearity (Dormann et al., 2013) (see Supplementary). So, of the 19 bioclimatic variables, 10 bioclimatic variables were used as inputs in modeling processes: annual mean temperature, mean diurnal range, isothermality, temperature seasonality, mean temperature of wettest quarter, annual precipitation, precipitation of driest month, precipitation seasonality, precipitation of warmest quarter, and precipitation of coldest quarter.

Table 1
Geographic area in km^2 forecasted as climatically-suitable by Maxent, GARP and their combinations for different scenarios considered in the study.

	Current	Scenario 1 (RCP2.6)	Scenario 2 (RCP4.5)	Scenario 3 (RCP8.5)
Maxent	2,093,355	2,420,323	2,817,101	1,833,021
GARP	1,184,062	1,477,732	1,853,504	1,594,683
Combined	2,522,094	2,890,376	3,251,760	2,537,413

2.2. Modeling algorithms

Maxent (version 3.3.3 k; http://www.cs.princeton.edu/~schapire/maxent/; Phillips et al., 2006) and GARP (desktop GARP in open Modeller; Muñoz et al., 2009) were employed as the modeling algorithms. Both the processes use known occurrences and pseudo-absences data re-sampled from the area (i.e., set of pixels) within the extent of concern where the species of question is not known to occur. Maxent estimates an unknown probability distribution for occurrence of the species based on the principle of maximum entropy, i.e. it searches for the set of pixels within the study area with environmental characteristics closest to average of the known occurrences (for details about Maxent, see Phillips et al., 2006). Although widely used for modeling species distribution, Maxent appears to suffer the issues of optimal threshold selection and inadequate ability to transfer the model to novel geographic situations (i.e., issue of under-predictions) as noted by Phillips et al. (2006), and also discussed by Peterson et al. (2007), and Royle et al. (2012). GARP on the other hand works in an iterative process of rule selection, evaluation, testing, and incorporation or rejection of the rule based on predictive accuracy from one iteration to the next, i.e., GARP discerns the set of pixels within the study area via evolution of 'if/then' rule to maximize predictivity based on the environmental characteristics of known occurrences (for details about GARP, see Stockwell and Peters, 1999; Peterson and Vieglais, 2001). However, GARP faces issues of poor performance compared to Maxent regarding interpolation of gaps in between the known occurrences (Peterson et al., 2007), and inability to produce unique solution owing to stochastic elements in the algorithm (typical for most machine learning algorithms) (Stockwell and Noble, 1992). With this information in mind, the study generated a *final* climatic niche map of Himalayan musk deer by combining the maps obtained by each of these two techniques for different scenarios. By so doing, it is expected that the predicted climatically-suitable habitat distribution map of the species is more robust and relatively accurate.

For pseudo-absences, Maxent selected only areas with presence locations (i.e., countries with occurrence records) to limit the points to areas that were surveyed for the species, potentially providing the background samples with the same bias as presence locations (Elith et al., 2011). The model was run 50 times, and hence the output of Maxent is an average of 50 replications. Continuous raster map produced with pixel-value ranging 0–1 for habitat suitability was exported to ArcGIS version10.4 (www.esri.com), and a binary map of climatically-suitable and unsuitable areas was created using 'minimum training presence logistic threshold' (value = 0.15).

To handle the variability in models produced by GARP, 3 models were created for each scenario and a composite binary map of climatically-suitable and unsuitable habitat was created where pixels predicted present by at least 2 models were considered "predicted presence" (Anderson et al., 2002; Lim et al., 2002). Lowest training presence threshold was used in GARP for model projection to produce a suitability map. Lastly, a final map was produced by summing up the binary maps generated from each of the modeling processes.