Available online at www.pelagiaresearchlibrary.com

Pelagia Research Library

Pelagia Research Library

European Journal of Experimental Biology, 2015, 5(11):12-19

ISSN: 2248 –9215
CODEN (USA): EJEBAU

Prediction of suitable habitats for *Syzygium caryophyllatum*, an endangered medicinal tree by using species distribution modelling for conservation planning

N. Stalin and Swamy P. S.*

Department of Plant Science, School of Biological Sciences, Madurai Kamaraj University, Madurai, Tamil Nadu, India

---

ABSTRACT

The aim of this study was to use species distribution models (SDMs) to estimate the effects of environmental variables on the habitat suitability of Syzygium caryophyllatum (L.) Alston. SDMs help to identify suitable habitats for the development of threatened plant populations to prevent extinctions, especially in the face of the global environmental change. In the present study three different modelling algorithms were used to predict the habitat suitability of an endangered plant species S. caryophyllatum towards developing conservation strategies. The BIOCLIM, GARP (GARP with the best subsets-new open modeller implementation) and MaxEnt algorithms were run using the Open Modeller Desktop version 1.1.0 software. Jackknife test was used to evaluate the importance of the environmental variables for predictive modelling. Bioclim and GARP models were more accurate with statistically significant AUC (area under the receiver operating characteristic curve) values of 0.99 and 0.97 compared to MaxEnt model which showed the AUC value of 0.91. This approach could be promising in predicting the potential habitat suitability of endangered plant species S. caryophyllatum with minimum number of occurrence points and thus, it can be used as an effective tool for species restoration and conservation planning.

Keywords: *Modelling algorithms, S. caryophyllatum, AUC, Bioclim, MaxEnt, GARP.*

INTRODUCTION

Species distribution models (SDMs) or ecological niche models (ENM) that use environmental factors based on historical collections are increasingly being used to analyze species distributions and also to predict the presence or absence of species in unrecorded areas [1-3]. These models establish relationships between occurrences of species and biophysical and environmental conditions in the study area. SDMs have been used to predict potentially suitable areas for the conservation of endangered and rare species [4-8] for the identification of suitable sites for reintroduction or restoration [9,10] and for assessing potential effects of future climate change on species distributions as well as on local species diversity [11,12]. It is also used to enable the analysis of the impacts of climate change on species, it is essential to quantify the relative importance of climate relative to other descriptors of the environment [13,14].

The ecological niche models have also been used in a wide range of applications such as in locating rare and threatened species habitats [15,16] predicting the spread of crop pests [17] and in estimating the response of species to global climate change [18]. Recent works in this field deals with methodological challenges specific to best ENM-based predictions of suitable areas and identification of conservation priorities[19,20].

The Western Ghats comprises the major portion of the Western Ghats and Sri Lanka Hotspot, one of 34 global biodiversity hotspots for conservation and one of the five on the Indian subcontinent. The area is extraordinarily rich in biodiversity. Although the total area is less than 6 percent of the land area of India, the Western Ghats contains more than 30 percent of fauna and flora found in India. Like other hotspots, the Western Ghats has a high proportion of endemic species. The Western Ghats contains numerous medicinal plants and important genetic resources such as the wild relatives of grains, fruits and spices. In addition to rich biodiversity, the Western Ghats is home to diverse social, religious, and linguistic groups. The high cultural diversity of rituals, customs, and lifestyles has led to the establishment of several religious institutions that strongly influence public opinion and the political decision-making process. Conservation challenges lie in engaging these heterogeneous social groups and involving them in community efforts aimed at biodiversity conservation and consolidation of fragmented habitats in the hotspot.

Syzygium caryophyllatum (L.) Alston., commonly known as Wild black plum, a medium sized tropical evergreen tree belongs to the family Myrtaceae. The vernacular name of this plant is Jangli jamun in Hindi, Kattunjara, Kanipazham or Jnarapazham in Malayalam, Kunta nerale in Kannada. *S.caryophyllatum* is native to India and Sri Lanka; in India the distribution mainly occur in the forests of Western Ghats. Tree grows along margin of evergreen forests or in open formations from low to higher elevations. Fruits are edible, sweet and astringent in taste and they are useful in stomatitis and intestinal disorder. The seeds and bark were dried and its decoction was used for the treatment of Diabetes mellitus [21] . The leaf and bark extracts of this plant are well known for its antibacterial and antioxidant efficacy [22] . Tribal peoples were considering this plant as a boon of nature and its fruits and seeds were consumed by Paniya tribal community of Waynad district, Kerala [23].

Based on the previous reports the extended distribution of this plant species is reported in Kalakad Mundanthurai Tiger Reserve (KMTR) forest located in the Southern Western Ghats in the Tirunelveli district of Tamil Nadu, South India [24] and Annamalai hills that form of the Western Ghats-Sri Lanka biodiversity hotspots [25] . Based on the threat perception due to its habitat loss and human activities natural populations of this species are on the decline mode. Due to these pressures this species has been listed under the endangered category of IUCN Red List. The ecological conditions necessary for the survival of this species can greatly help in conservation scenario. The aim of this study was to predict suitable habitat distribution for threatened tree species *Syzygium caryophyllatum* using known presence observations with three different modelling algorithms (BIOCLIM, GARP and MaxEnt). To identify the environmental factors associated with *S.caryophyllatum* habitat distribution and to predict suitable habitat for reintroduction and future conservation of this species.

MATERIALS AND METHODS

Species occurrence data for Ecological niche modelling

The occurrence points of *S. caryophyllatum* were identified based on the field surveys in the Western Ghats region of Tamilnadu, Kerala, Maharastra and Karnataka states of India and also from the secondary data collected from the literature survey. Thirty two occurrence points of *S. caryophyllatum* were used in the present study. The primary presence only data was used for modelling the distribution of this endangered species.

Environmental data

The environmental variables were nineteen bioclimatic variables used for all the three modelling algorithms (Table 1). These bioclimatic variables were derived from the monthly temperature and rainfall values in order to generate more biologically significant variables. The bioclimatic variables represent, annual trends (e.g. mean annual temperature, annual precipitation), seasonality (e.g. annual range in temperature and precipitation) and extreme or limiting environmental factors (e.g. temperature of the coldest and warmest month, precipitation of wet and dry quarters). These variables were obtained from globally interpolated datasets (source: http://www.worldclim.org) which are presumed to be relevant to plant existence [12, 26, 27]. Analyses were conducted at the 1 x 1 km pixels spatial resolution of the environmental data sets.

Model development

Three different modelling algorithms were used in the present study following the Open Modeller Desktop version 1.1.0 (downloaded from http://openmodeller.sourceforge.net) software. BIOCLIM, GARP (GARP with the best subsets – new open modeller implementation) and MaxEnt algorithms were run using the above software [28-31]. BIOCLIM is one of the earlier modelling techniques, based on climatic envelop theory. For each given environmental variable the algorithm finds the mean and standard deviation (assuming normal distribution) associated with the occurrence points. Each variable has its own envelope represented by the interval [M - CO* StDev, M + C* StDev], where 'M' is the mean; 'CO' is the cut off input parameter; and ’StDev’ is the standard deviation. Besides the envelope, each environmental variable has additional upper and lower limits taken from the maximum and minimum values related to the set of occurrence points [28].

Table 1 Bioclimatic variables used in the model development

Variables	Details
Bio1	Annual mean temperature (º C)
Bio2	Mean diurnal temperature range [mean of monthly (max temp–min temp)]
Bio3	Isothermality (Bio2/Bio7) (×100)
Bio4	Temperature seasonality (standard deviation×100) (º C)
Bio 5	Maximum temperature of warmest month (º C)
Bio 6	Minimum temperature of coldest month(º C)
Bio7	Temperature annual range (P5–P6) (º C)
Bio8	Mean temperature of wettest quarter (º C)
Bio9	Mean temperature of driest quarter (º C)
Bio10	Mean temperature of warmest quarter (º C)
Bio11	Mean temperature of coldest quarter (º C)
Bio12	Annual precipitation (mm)
Bio 13	Precipitation of the wettest month (mm)
Bio 14	Precipitation of the driest month (mm)
Bio15	Precipitation seasonality (coefficient of variation) (mm)
Bio16	Precipitation of wettest quarter (mm)
Bio17	Precipitation of driest quarter (mm)
Bio18	Precipitation of warmest quarter (mm)
Bio19	Precipitation of coldest quarter (mm)

GARP (genetic algorithm for rule set prediction) is an ecological niche modelling method based on a genetic algorithm. This modelling approach predicts the suitable environmental conditions under which the species should be able to maintain populations. For input, GARP uses a set of point localities where the species is known to occur and a set of geographical layers that might limit the specie's capabilities to survive. This model is a random set of mathematical rules which can be read as limiting environmental conditions [31].

The maximum entropy (MaxEnt) approach estimates a target probability distribution of the species by finding the probability distribution of maximum entropy (i.e., that is most spread out or closest to uniform with reference to a set of environmental variables). Default values of different parameters, maximum iterations = 500, convergence threshold = 0.00001 and 50% of data points were used as a random test percentage in the present study [29,30].

Model validation

A receiver operating characteristics (ROC) plot was generated by incorporating the sensitivity values, the true positive fraction against the false positive fraction for all available probability thresholds to measure prediction accuracy of the models output [32-34]. The sensitivity values were calculated using confusion matrix. A curve which maximizes sensitivity against low false positive fraction values is considered as good model which was evaluated by using the area under the curve (AUC). Cross-validated AUC values were summarized to present overall model performance by taking mean AUC values of all model accuracies. The range of AUC is from 0.0 to1.0. A model providing excellent prediction has an AUC higher than 0.9, a fair model has an AUC between 0.7 and 0.9 and a model with AUC below 0.7 is considered poor (Swets 1988). The *Jackknife* procedure was used to assess the importance and percentage of contribution of bioclimatic variables. The final potential species distribution maps had a range of values from 0 to 1 which were regrouped into three classes of potential habitats viz., 'high potential' (>0.6), 'medium potential' (0.2–0.4) and 'low potential' (<0.2).

RESULTS AND DISCUSSION

The prediction of suitable habitats for conservation of the endangered tree species *S. caryophyllatum* was successfully predicted by three species distribution models (BIOCLIM, GARP and MaxEnt). Model outputs varied with the modelling techniques used in the present study (Figure 1 a, b & c). SDMs outputs revealed that, MaxEnt predicted largest area (75.95%) under potential distribution compared to BIOCLIM (1.73%) and GARP (5.26%). Both GARP and MaxEnt showed a wide range of distribution from low to high probability area in India whereas, BIOCLIM output is restricted to Western ghats regions of India. The AUC values for the current potential distribution of *S. caryophyllatum* were high indicating good predictive model performance. BIOCLIM and GARP models showed a good performance with AUC values of 0.99 and 0.97 compared with MaxEnt AUC value of 0.91 (Table 2).

Figure 1 Predicted suitable habitats for *S. caryophyllatum* a) BioClim b) GARP c) MaxEnt

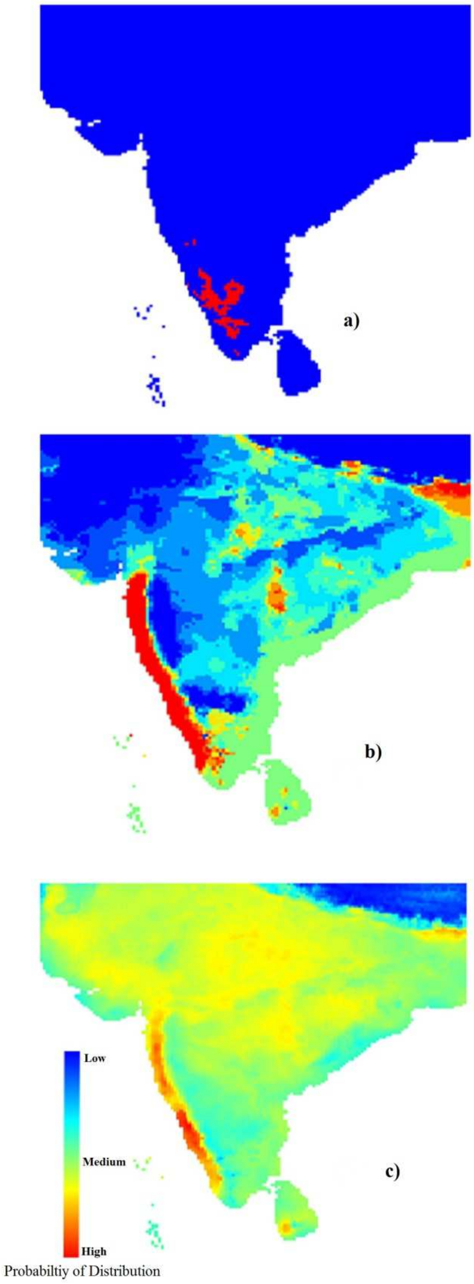


Table 2 Comparative account of various model outputs

Parameters	Bioclim	GARP	MaxEnt
Accuracy (%)	100	96.7742	100
AUC Values	0.99	0.97	0.91
Omission error	0	0.0322581	0
Commision error	0	0	0
Threshold	50 %	20 %	30.598%
% of cells predicted present	1.73229 %	5.26753 %	75.9551%

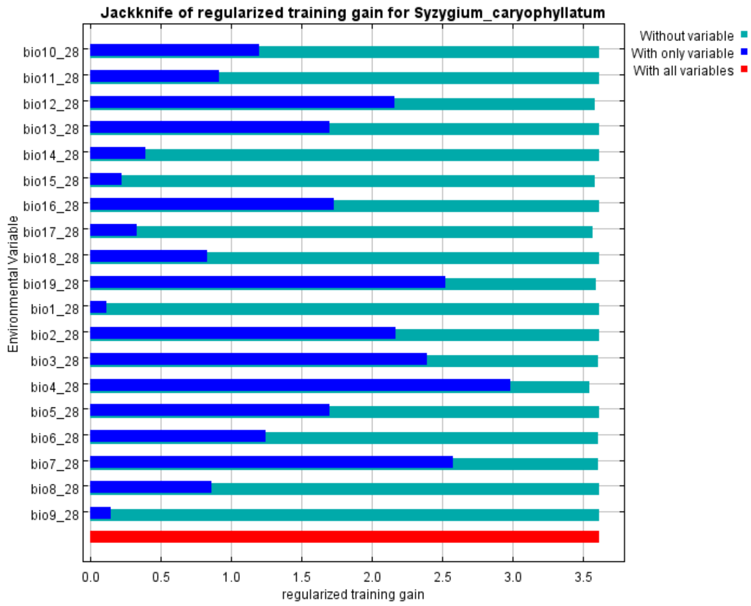
The jacknife test of variable importance in MaxEnt has identified the precipitation of coldest quarter (bioclimatic variable 19) as the most important environmental variable contributed to the model development (Figure 2). Other variables like Isothermality (bioclimatic variable 3), Temperature seasonality (bioclimatic variable 4), Temperature annual range (bioclimatic variable 7), Annual precipitation (bioclimatic variable 12), Precipitation of driest quarter (bioclimatic variable 17), Annual mean temperature (bioclimatic variable 1) and Mean diurnal temperature range (bioclimatic variable 2) also have considerable predictive values with regard to distribution of *S.caryophyllatum*

(Figure. 3). Considering the permutation importance, only 9 variables out of 19 contributed to the model output. Among the 9 variables, temperature seasonality (bio 4) (94 %) had maximum influence on habitat suitability model followed by annual precipitation (bio 12) (3.5%) and precipitation of driest quarter (bio 17) (1.8 %). All three variables together contributed to 99.3% of the variation (Table. 3). Potential distribution maps show various possibilities for conservation and management of this endangered tree species. Both MaxEnt and GARP have shown a wider range of distribution from medium to high probability in South-west and north-east states of India and also showed high probability in some areas of Sri Lanka, it is an another native region of *S. caryophyllatum* whereas, Bioclim distribution is restricted as high probability areas in and around the occurrence points of its existing natural populations.

Table 3 Analysis of percent contribution and permutation importance of Bioclimatic variables to MaxEnt model

Variables	Percentage contribution	Permutation importance
Bio 19	49.5	1.3
Bio 3	17.1	0.1
Bio 4	9.5	94
Bio 7	7.5	0.2
Bio 12	5.3	3.5
Bio 17	3.5	1.8
Bio 1	3.4	0.6
Bio 2	1.4	0
Bio 6	0.8	0.1
Bio 18	0.7	0
Bio 15	0.6	0.3
bio16	0.5	0

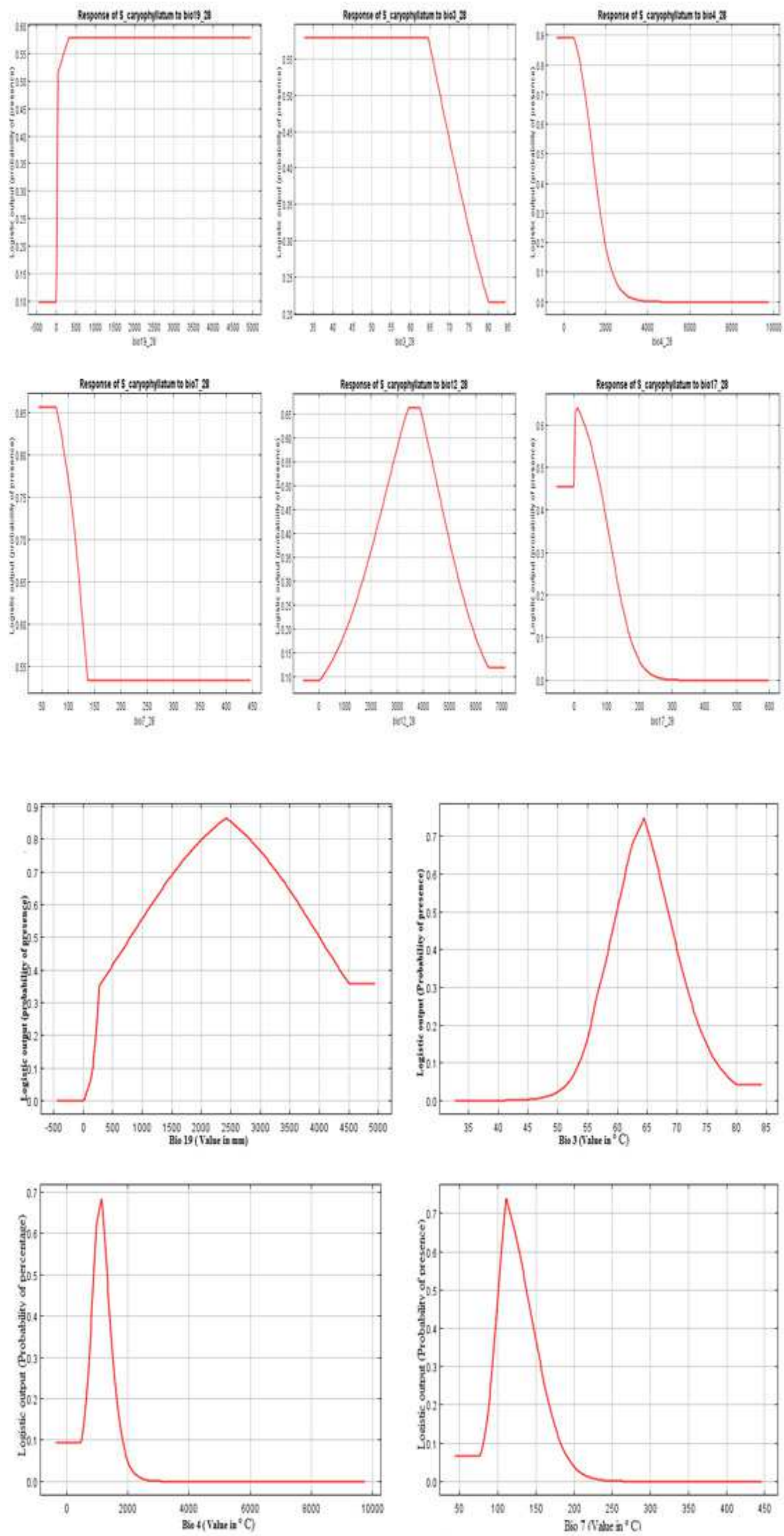
Figure 2 The Jackknife test for evaluation of relative importance of environmental variables for *S. caryophyllatum*



The major role of SDM is to estimate the probability of occurrence of a given species based on observed presence and (or absence locations) as well as environmental and climatic covariates. A common application of this method is to predict species ranges with climate data as predictors. Several studies were done successfully and predict suitable distribution habitats for many threatened plant species using different modelling algorithms like *Artemisia sieberi* and *Artemisia aucheri, Justicia adhatoda ,Coscinium fenestratum, Tapirus pinchaque* and *Monotropa uniflora* [35-39]. Ray et al. (2011) [40] reported the predictive distribution modelling of a rare Himalayan medicinal plant *Berberis aristata* using the three algorithms used in the present study. Also in the previous studies, distribution map of suitable habitat for conservation of *Nepeta septemcrenata* [41] and an endangered tree *Canacomyricca moniticola* [10] using MaxEnt algorithm with low omission error was predicted.

Distribution data on threatened species often have few records and are geographically close together, making it difficult to model their appropriate habitat distribution using commonly used modelling approaches because such data provide limited information for determining the relationships between the species and their environments [10]. Maximum entropy (Maxent) models present good results even for small sample size [42]. Maxent is a multivariate approach to study the geographic distribution of species on a large scale using only presence data of the species [30].

Figure 3 Response curves of the variables that most contributed to explain the potential distribution of *S.caryophyllatum* (Isothermality (Bio 3), Temperature seasonality (Bio 4), Temperature annual range (Bio 7), Annual precipitation (Bio 12), precipitation of the driest quarter (Bio17) and precipitation of the coldest quarter (Bio19)



The presence points of the species were considered as appropriate places for their occurrence. The essential data layers were imported for this model and then statistical analysis was performed by Maxent software to map potential habitat of distribution pattern of threatened and endangered plant species. Maxent has several strong characteristics, it needs only species occurrence data and environmental factors; it can examine factor importance by way of a jackknife procedure and also facilitates model interpretation [30,43].

Geographical data of threatened species provide the degree of species threats, their distributions, and habitat requirements of species. It will also be useful in identifying potentially sensitive or uniquely fragile ecosystems. A threatened species is the one with narrow habitat range, low climate tolerance; specialised adaptation requiring an outside agency for pollination, poor dispersal strategies, few seeds per fruit and poor viability of seeds [44]. The conservation status of a species can best be developed by synthesising information on each of its known populations, viewed together with any information on changes in historical range and evidence of vulnerability of its characteristic habitat. Mapping by such intrinsic features of the land as natural regions and physiographic areas is the best way of presenting plant distributions data. Species and habitat relationship modelling with precise locality data on microclimate, topography and soil in association with site-specific location data of concerned taxa helps in understanding the interrelationships and controls of biotic and abiotic factors on species distribution pattern [45]. Natural populations of *S. caryophyllatum* have always been small, but our modelling distribution prediction results showed potential habitat greater than the area of the actual distribution. These results give an insight into the availability of areas suitable for the species' regeneration, possibly through *ex vitro* conservation planning.

CONCLUSION

Our results of potential habitat distribution maps for *S. caryophyllatum* may help to discover new populations, identify top-priority study sites or set priorities to restore its natural habitat for more effective conservation. Moreover the effective conservation planning is necessary for this tree species for its further existence in the natural forests.

Acknowledgment
First author thank University Grant Commission for the award of meritorious fellowship. Corresponding author thank partial financial support through UGC for the funding of UGC-CAS, DST-IPLS and DST-Purse programme.

REFERENCES

[1] Araujo MB, Pearson RG, Thuiller W, Erhard M, *Glob Change Biol*,**2005**,11, 1504.
[2] Elith J, Leathwick JR, *Annual Review of Ecology, Evolution and Systematics*, **2009,** 40, 677.
[3] Guisan A, Hofer U, *J Biogeogr*, **2003**, 30, 1233.
[4] Gallagher RV, Hughe L, Leishman MR, Wilson PO, *Biol Invasions*, **2010**, 12, 4049.
[5] Papes M, Gaubert P, *Divers Distrib*, **2007**, 13, 890.
[6] Rebelo H, Jones G, *J Appl Ecol*, **2010,** 47**,** 410.
[7] Solano E, Feria TP, *Biodiv and Conser*, **2007,** 16, 188.
[8] Arun KD, Rameshprabu N, Swamy PS, *Paper Proceedings of International Conference on Biodiversity* (ICBD), Colombo, Sri Lanka, **2013,** pp 5.
[9] Klar N, Fernandez N, Kramer-Schadt S, Herrmann M, Trinzen M, Buttner I, Niemitz C, *Biol Conser*, **2008**, 141, 308.
[10] Kumar S, Thomas J, Stohlgren, *J Ecol Natural Environ*, **2009**, 1(4), 94.
[11] Hole DG, Willis SG, Pain DJ, Fishpool LD, Butchart SHM, Collingham YC, Rahbek C, Huntley B, *Ecol Lett*, **2009**, 12, 420.
[12] Pearson RG, Dawson TP, *Global Ecol Biogeogr*, **2003**, 12, 361.
[13] Morueta-Holme N, Flojgaard C, Svenning JC, *PLos One*, **2010**, 5(4), 10.
[14] Newbold T, *Progress in Physical Geography*, **2010**, 34(1), 3.
[15] Jackson CR, Robertson MP, *J Nat Conserv*, **2011**, 19, 87.
[16] Peterson AT, Martínez-Campos C, Nakazawa Y, Martínez-Meyer E, *Transactions of the Royal Society of Tropical Medicine and Hygiene*, **2005**, 99, 647.
[17] Ganeshaiah KN, Barve N, Nath N, Chandrashekara K, Swamy M, *Cur Sci*, **2003**, 85, 1526.
[18] Barve N, Bonilla AJ, Brandes J, Brown JC, Brunsell N, *Rev Mex Biodivers*, **2012**, 83, 817.
[19] Loiselle BA, Howell CA, Graham CH, Goerck JM, Brooks T, Smith KG, Williams PH, *Conserv Biol*, **2003**, 17, 1591.
[20] Peterson AT, Kluza DA, *Anim Conserv*, **2003**, 6, 47.
[21] Ediriweera ERHSS, Ratnasooriya WD, *Ayu*, **2009**, 30(4), 373.
[22] Shilpa KJ, Krishnakumar G, *Int J of Pharm Pharm Sci*, **2012**, 4, 198.
[23] Ratheesh NMK, Anilkumar N, Balakrishnan V, Sivadasan M, Ahmed Alfarhan H, Alatar AA, *J Med Plants Res*, **2011**, 5(15), 3520.
[24] Parthasarathy N, *Biodiv and Conser*, **2009**, 8, 1365.
[25] Muthuramkumar S, Ayyappan N, Parthasarathy N, Mudappa D, Shankar Raman TR, Arthur Selwyn M, Arul Pragasan L, *Biotropica*, **2006**, 38(2), 143.
[26] Hijmans RJ, Cameron S, Parra J, Jones P, Jarvis A, *Int J Climatol*, **2005**, 25, 1965.

[27] Irfan-Ullah M, Giriraj A, Murthy MSR, Peterson AT, *Biodiv Conser*, **2007**, 16(6), 1917.
[28] Busby JR, BIOCLIM - A Bioclimatic Analysis and Prediction System. In: Margules, C.R & M.P. Austin (eds.) Nature Conservation: Cost Effective Biological Surveys and Data Analysis. Canberra: CSIRO, **1991**, pp 64.
[29] Peterson AT, *The Condor,* **2001**, 103(3), 599.
[30] Phillips SJ, Dudik M, Schapire RE, *In proceedings of the 21st International conference on machine learning,* AMC Press, New York, **2004**, pp 655.
[31] Stockwell DRB, Peters D, *Int J Geogr Inf Sci*, **1999**, 13 (2), 143.
[32] Boubli JP, de Lima MG, *Int J Primatol*, **2009**, 30.
[33] Lobo JM, Jime´nez-Valverde A, Real R, *Global Ecol Biogeogr*, **2008**, 17, 145.
[34] VanDerWal J, Shoo LP, Graham C, Williams SE, *Ecol Model*, **2009**, 220, 589.
[35] Hosseini SZ, Kappas M, ZareChahouki MA, *Ecol Inform,* **2013**, 18, 61.
[36] Yang XQ, Kushwaha SPS, Saran S, *Ecol Eng*, **2013**, 51, 83.
[37] Thriveni HN, Srikanth V, Gunaga HN, Babu R, Vasudeva R, *Trop Ecol*,**2015**, 56(1), 101.
[38] Ortega-Andrade HM, Prieto-Torres DA, Gómez-Lora I, Lizcano DJ, *PLoS ONE*, **2015**, 10(3), e0121137.
[39] Pradhan P, Biodiversitas, **2015**, 16(2), 109.
[40] Ray R, Gururaja KV., Ramachandra TV, *J Env Biol*, **2011**, 32(6), 725.
[41] Khafaga O, Hatab EE, Omar K, *Academia Arena*, **2011**, 3(7), 45.
[42] Phillips SJ, Anderson RP, Schapired RE, *Ecol Model*, **2006**, 190, 231.
[43] Scheldeman X, van Zonneveld M, *Biodiversity International*, **2010**, 139.
[44] Nayar MP, Hot spots of endemic plants of India, Nepal and Bhutan, TBGRI, Trivandrum, **1996**, pp 252.
[45] Varghese AO, Joshi AK, Krishna Murthy YVN, *J Indian Soc Remote*, **2010**, 38 (3), 523.