Modeling Spatial Distribution of Lantana Camara – A Comparative Study

Neena Priyanka¹,²,* and P K Joshi¹

¹TERI University, 10 Institutional Area, Vasant Kunj, New Delhi - 110070, India
²Pitney Bowes Software, Logix Cyber Park, First Floor, Tower - A, C 28-29, Sector - 62, NOIDA, UP, 201301, India

Keywords:
Lantana camara; niche modeling; Biomapper; GARP; Maxent

Abstract
Lantana camara L. Was introduced in Indian subcontinent dating back to 1800s as an ornamental plant. Later, it resulted in serious ecological and economic repercussion such as altered native composition, ecological degradation, fire facilitation to areas it invade and particularly to ecologically rich and diverse regions such as protected areas. Such consequences led us to investigate invasion potential distribution of Lantana camara in protected areas namely Jim Corbett and Rajaji National Parks of Western Himalaya, where its invasion has become a menace. Invasion potential distribution was modeled using 3 ecological niche modeling algorithms (i) Biomapper, (ii) GARP, and (iii) Maxent, as regards to their ability to predict the geographic distributions of species. The results mirror the known distribution of the Lantana camara fairly well with various degrees of predictions. Maxent succeeded in anticipating most of the invasion potential distributions whereas GARP and Biomapper produced an odd pattern of over-predictive and under-predictive models respectively. These findings are relevant as model projections provide a useful way to visualize spatial spread and extent of invasion potential distribution. Such predictions are useful for conservation practices including decisions for early detection and formulation of targeted and timely mitigation measures to curb invasion.

1. Introduction
Across the globe, there is a surfeit of Invasive Alien Species (IAS) that create a multitude of harmful effects to native environs that include displacement of native species [1], degradation of habitat [2], alteration in soil properties [3], elimination of wildlife forage [4], alteration of fire regime [5-6] threat to species [7-8] and many more. These host effects caused by IAS suggest that the impacts of invasion will exceed the pace at which understanding about the environment is being enhanced. Thus, worldwide efforts are being carried out to constrain and manage the spread of the invaders. India, with its rich biological diversity, is abode of large number of floral (45,500 plant species) and faunal (91,200 animal species) species [9]. About 40% of the Indian flora constitutes adventive aliens of various origins ranging from American, Asian, Malaysian European and Central Asian

* Corresponding Author (Email: neenapriyanka@gmail.com Tel. (+91) 996 872 3454; Fax: (+91) 120 402 6013).
species and amongst these 25% are invasive [9-11]. A few of these have conspicuously altered vegetation patterns of the country in terrestrial environment [12-13]. Amidst these, *Lantana camara* is one such species that has caused severe repercussions to the native environs.

Originally introduced to Indian subcontinent dating back to 1800s as an ornamental plant, *Lantana camara* (wild sage) has become one of India’s leading invasive plant species [11, 14]. The diversity, broad geographic expansion and wide ecological tolerance are so much that the Invasive Species Specialist Group (ISSG) considers it as among 100 of the "World's Worst" invaders [15]. In spite of such widely known facts that *Lantana camara* is a serious invader to various terrestrial ecosystems, there are few empirical studies in India which accentuate reliable estimates on its distribution and spread [16-17]. Beside these limitations, information if available is fragmented and inadequate, often based on anecdotal evidence. In this context, the major challenge for decision makers is to how to effectively manage the spread of invasive to preserve native biodiversity.

Various studies conducted on invasive aliens prove that once they start expanding the area, little can be done to eradicate them. This call for more research, especially research that will yield quantitative estimates of distribution, is almost universal in the literature of *Lantana camara* invasion studies [14,18-19]. Satellite remote sensing and GIS have successfully been applied to map the distribution of several exposed invasive species, their ecosystems and landscapes [16,20]. However, making detection of cryptic invaders such as *Lantana camara* is complicated as the captured spectral information cannot be directly attributable to these species, thus the prediction of its distribution is difficult [16-17,21]. A combination of remote sensing, GIS and spatial modeling offers potential to detect understory invasion through the development of models and risk maps. However no such attempts have been made in this direction to demonstrate the invasion distribution of *Lantana camara* in India.

To fill this gap, the present study was carried out to predict the distribution of *Lantana camara* along environmental gradients in the north western Himalayan region of India. The region encompasses two important Protected Areas viz., Jim Corbett and Rajaji National Park where the species has been reported to become nuisance. So, to examine its invasion distribution Ecological Niche Modeling (ENM) tool based on Grinnellian concept was applied. The objectives of present study are to (i) test 3 individual modeling algorithms for predicting the invasion distribution of *Lantana camara* and (ii) determine the key variables that are directly contributing to its spread. The obtained information will be useful for creating management strategies to curtail species invasion regime.

2. Material and methods

There are a variety of empirical approaches which are being widely used to map invasion distribution of invaders. A few of these approaches which uses species presence records only are ecological niche modeling [22-24], envelope method (Bioclim) [25-26], Generalized Additive model (GAM) [27-28], logistic regression [29-30], neural networks [31-32] and regression trees [33-34]. Amongst them, ecological niche modeling offers plethora of algorithms such as Genetic Algorithm for Rule set Prediction (GARP), Maximum Entropy (Maxent), Biomapper, Climex, FloraMap, Domain, Envelope which can be employed in deducing invasion distribution of species. In this study, three modeling algorithms namely GARP, Maxent, and Biomapper have been employed to
predict invasion distribution of *Lantana camara*. A brief description of these models have been enlisted in Table 1 and discussed in details below. The special focus of this study, thus, is to use presence-only data for modeling *Lantana camara* distribution in spite being confronted with challenges in model inference relative to presence-absence techniques.

### Table 1. Modeling algorithms families, data and software requirements, and selected references

<table>
<thead>
<tr>
<th>Model</th>
<th>Species data</th>
<th>Environmental data</th>
<th>Software availability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ecological niche factor analysis (Biomapper)</td>
<td>Presence</td>
<td>Continuous</td>
<td><a href="http://www.unil.ch/biomapper">www.unil.ch/biomapper</a></td>
</tr>
<tr>
<td>Genetic algorithm for Rule-set prediction (GARP)</td>
<td>Presence (generates pseudo-absences internally)</td>
<td>Continuous (Possibly categorical)</td>
<td><a href="http://beta.lifemapper.org/desktopgarp">http://beta.lifemapper.org/desktopgarp</a></td>
</tr>
<tr>
<td>Maximum Entropy (Maxent)</td>
<td>Presence</td>
<td>Continuous and/or categorical</td>
<td><a href="http://www.research.att.com/~phillips/density.html">http://www.research.att.com/~phillips/density.html</a></td>
</tr>
</tbody>
</table>

### 2.1. Study Area

The study area, include a portion of Western Himalaya encompassing two most important national Parks viz. Jim Corbett and Rajaji, and lies in the state of Uttarakhand, India. The pristine environs of these two parks are reported to be infested with gregarious *Lantana camara* presence with high density of infestation. The spatial coverage of the study area was deliberately made larger (portion of Western Himalayan region was considered instead of geographical areas of two parks) than the actual range occupied by *Lantana camara* to ensure encompassing of the full extent of the predicted range. Figure 1 depicts the study boundary encompassing Rajaji and Jim Corbett National Park and the adjoining areas used in current research.

The Corbett National Park, (29°50’–30°20’ N and 77°55’-79°80’ E) covering an area of 1318.54 sq. km, lies in two districts– Nainital and Pauri Garhwal. The park is situated at an altitudinal range of 385-1100 m absl with a minimum temperature around 13.1°C and maximum temperature around 38.9°C and average rainfall of 1200 mm. Both parks represent the Shivalik eco-system and the beginning of the vast Indo-Gangetic plains ecosystem, thus representing vegetation of several distinct zones, forest types and diverse flora and fauna. Terrain is approximately 75% hilly and 25% plains.

### 2.2. Input datasets

*Lantana camara* occurrence data comprised of GPS location records from Rajaji and Jim Corbett National Park. Field surveys were conducted twice (March and November, 2009) in the study site to determine the presence of species in the same location during different seasons. For collection of the field records, a base map was prepared using topographic sheets (1:50,000 scale), available satellite data and open source maps. A fishnet
mesh (uniform resolution grid – URG) of 1×1 sq. km was created for the entire study area. The collection location estimates were based on consultation with Park officials, researchers and locals. 302 locations distributed randomly in the Jim Corbett (187) and Rajaji (115) National Parks were selected for data collection and field verification. All locations were physically visited and verified with GPS and base maps. A patch size greater than half a hectare was considered for occurrence collection. All records were geotagged on the base map using the GPS locations. Some of the new locations showing the dominance of the target species were also accounted in the occurrence database.

![Figure 1. Study Area](image)

During the post field data assessment, all point locations collected were overlaid on the URG. Some of the grids having more than one point were identified. This was due to the occurrence of many collection sites in close proximity to each other. Such grids were refined with one location record. The selection of one location was based on nearness to the center of the grid under consideration. As a result, a total of 137 spatially unique points per grid were finally used in the data analysis and modeling. These locations were also validated through detailed review of its establishment in the region by electronic database searches such as Google, Scopus, Science Direct and published works. These 137 occurrence records obtained were plotted using ArcGIS (ESRI).

A suite of 108 environmental variables were considered as potential predictors of Lantana
camara habitat distribution. These variables were selected based on the biological relevance to species distribution and information obtained from other spatial modeling studies carried out on Lantana camara across geographies [35-38]. Monthly temperature (minimum and maximum), precipitation and bioclimatic variables (IPCC 3rd assessment data) defining biophysical tolerances of species were obtained from the WorldClim database [39]. Topographic variable included elevation data obtained from Consortium for Spatial Information (CGIAR-CSI) database. Other topographic variables namely slope, aspect, flow direction, flow accumulation and Compound Topographic Index (CTI) were computed from elevation dataset. Monthly Potential Evapo-transpiration (PET, mm/day) layers were calculated using Thornthwaite equation [20, 40] (cross ref) to model invasion distribution of Lantana camara (equation 1).

\[
\text{PET} = 16 \times dl \times [(10 \times T_a / I)^\alpha]
\]

Where, \(T_a\) = Mean monthly temperature (°C)
\(\alpha = 0.492 + (0.0179)I - (7.71 \times 10^{-5})I^2 + (6.75 \times 10^{-7})I^3\)

\(dl = \text{Day length in hours} / 12z\)
\(= \sum i \text{ (Annual heat index)}\)
\(i = \text{Monthly heat index given by} \quad i = (T_a / 5)^{1.5}\)

Solar radiation was computed using ArcGIS surface analysis tool (Solar Radiation Module - derives incoming solar radiation from raster surface such as DEM). Advanced Very High Resolution Radiometer (AVHRR) vegetation data representing NDVI was acquired from the Global Land Cover Facility (GLCF) website. Soil data, obtained from the International Soil Reference and Information Centre (ISRIC) (http://www.isric.org/data/data-download) was further used to discern invasion distribution of Lantana camara. As the species is associated with land use, and are prominent in the open canopy forest, near water and alongside roads, thus, land use data obtained from Globe cover was used (http://geoserver.isciences.com)

All environmental variables were resampled to 1 km spatial resolution. The environmental variables were subjected to multi-collinearity diagnostics using Pearson product-moment correlation coefficient to reduce the data dimensionality. The Pearson correlation coefficient \(\rho_{XY}\) between two random variables \(X\) and \(Y\) with expected values \(\mu_X\) and \(\mu_Y\) and standard deviations \(\sigma_X\) and \(\sigma_Y\) is defined in equation 2:

\[
\rho_{XY} = \text{Corr} (X, Y) = \frac{\text{cov} (x, y)}{\sigma_x \sigma_y} = \frac{E[(X - \mu_X)(Y - \mu_Y)]}{\sigma_x \sigma_y}
\]

The correlation coefficient values range between +1 and −1 inclusive. The Pearson correlation is +1 in case of a perfect positive (increasing) linear relationship, −1 in case of a perfect decreasing (negative) linear relationship, and some value between −1 and 1 in all other cases, indicating the degree of linear dependence between variables. The closer the coefficient is to either −1 or 1, stronger the correlation between variables and as the value approaches zero degree of correlation decreases and variables are
considered to be unique among the population of variables. For the present study, 400 values based on 135 species occurrence points and 265 randomly generated samples from the study area (representing 10% of the study area cells [201×253]) were examined. Pearson value threshold of ±0.80 was considered as elimination criteria for selecting unique variables [41]. Amongst 108 selected variables, numbers of predictor variables were reduced to 33 and had a Pearson value threshold less than ±0.80. These variables were used in niche modeling.

2.3. Spatial Modeling

The spatial invasion distribution models for *Lantana camara* were built using Biomapper, GARP and Maxent, all of which are niche modeling algorithm for estimating distributions using presence-only data [22-24].

**Biomapper** – Ecological Niche Factor Analysis (ENFA) was performed using Biomapper to construct habitat suitability maps for *Lantana camara*. The experiment was designed to derive factors and develop invasion distribution map that preeminent species niche. For this, all quantitative variables were converted to Idrisi raster format (.rst files). Qualitative variables cannot be used directly as quantitative counterparts. These datasets were converted into semi quantitative scale based on the important feature of the represented categories through coding it with an integer numeric value. Variables were prepared in ArcView 3.3 and then converted to IDRISI raster format (.rst files) for use in Biomapper. Data was first categorized as qualitative or quantitative based on data type which was then normalized using Box-Cox variable transformation. Habitat suitability maps were generated having five factors and ten categories. The output was further reclassified into five categories (i) Very Low (ii) Low (iii) Moderate (iv) High and (v) Very High invasion potential distribution.

**GARP** - For the experimental run of GARP, the convergence condition was set to 0.01, and maximum iterations values to 1000. Using this convergence criterion, 1000 models were generated and amongst them, best model were selected using “best subsets” method. The best subset separates technique by their omission - commission error distinctiveness. The resulting output raster grid was classified into 5 classes (i) Very Low (ii) Low (iii) Moderate (iv) High and (v) Very High invasion potential distribution.

**Maxent** - For Maxent modeling, the parameter settings included convergence criterion set as 0.01, maximum iterations to 1000 with default values of feature types as linear, quadratic, product, categorical, threshold and hinge. The regularization multiplier was set at 1. Output was obtained in the cumulative format which gives a probability of occurrence ranging between 0-100 percent but is presented on a logarithmic scale. The output was classified into 5 classes (i) Very Low (ii) Low (iii) Moderate (iv) High and (v) Very High invasion potential distribution.

The presence records and environmental variables were then subjected to spatial modeling and testing based on defined experimental run parameters. The experimental run derived best models in the form of continuous raster grids to produce a best ensemble estimate of geographic projection of *Lantana camara*. Testing and validation form essence of species spatial distribution models as these enable to assess the predictive performance of models. Hernandez *et al.* [42] suggested that to estimate model accuracy using only presence datasets, multiple evaluation criteria are essential. This is because each measure of testing and validation provides partial elusive “truth” about the models’ prediction abilities.
Under ideal scenario, independent occurrence records should be used for model performance testing [43]. However, in various cases this possibility is ruled out as not sufficient data may be available. Therefore, the best method is to partition the existing sample records randomly into ‘training’ and ‘test’ sets [44], as carried out in this study, thus creating quasi-independent records for testing the model. Hence, for the present study, three evaluation approaches viz. (i) **AUC (Area under Curve)** (ii) **Chi-square test** ($\chi^2$) and (iii) **Kappa statistics** based on quasi-independent test datasets were used in this study.

The AUC of the ROC function is an index of model performance providing overall accuracy through single measure and is threshold independent [45]. The AUC is the probability that a randomly chosen presence site is ranked above a random background site. A random model has AUC of 0.5 while a perfect model should have AUC of 1. AUC classified following Hosmer and Lemeshow, is as follows: AUC = 0.5, no discrimination; 0.7 < AUC < 0.8 = acceptable; 0.8 ≤ AUC < 0.9, excellent and AUC > 0.9, outstanding [46].

Further, models were tested using one-tailed chi-square test. The test is based on observed distribution in field against expected distribution determined by a null hypothesis (equation 3).

$$\chi^2 = \sum \frac{(O - E)^2}{E} \quad (3)$$

Where, $O =$ frequencies observed and $E =$ frequencies expected

The observed frequency of correct and incorrect predictions relative to expected frequency for the test occurrence points form the basis of $\chi^2$ test. At degree of freedom value as 1, the test reveals that how well the test locations predicted the species presence taking into considerations the proportional area predicted present by the model. Since the chi-square statistic summarizes observed and expected random expectations, its magnitude can therefore be used in determining extrapolative capability of model [47].

Kappa Index of Agreement (KIA) measures degree of agreement between two maps both in overall sense and on per category basis. Following Landis and Koch, Kappa values are interpreted as follows: <0: no agreement, 0.0 - 0.2: slight agreement, 0.21- 0.4: fair agreements, 0.41 - 0.60: moderate agreements, 0.61 – 0.80: substantial and 0.81-1: almost perfect agreement. The training derived models obtained from 70% of occurrence points and test derived models obtained from quasi-independent test datasets (30%) were used for computing accuracy assessment and determining kappa range of values. Both these maps were coded as: 0-10 (Very Low), 10-30 (Low), 30-60 (Moderate) and 60-90 (High) and 80-100 (Very High) before accuracy assessment. This approach was used for assessment of the methods’ overall stability in its prediction when incomplete species occurrence datasets is used for predictive modeling.

3. Results

3.1. Invasion distribution of *Lantana camara*

Using 137 occurrence records of *Lantana camara* and 33 environmental variables, the invasion distribution models of *Lantana camara* were generated. Three algorithms viz., Biomapper, Maxent and GARP evaluated in this research were variable in degree of predictions ranging several-fold in
invasion distribution area predicted present for *Lantana camara*. However, models derived using these algorithms were consensus on core areas where species is most likely to occur. Amongst 3 models, Maxent model depicted *Lantana camara* presence gregariously in Himalayan foothills covering parts of Bhabar, Tarai and Shiwalik ranges. The distribution of species was low in middle Himalayan region. It was found to be predominant in regions with altitudinal ranges less than 2000 m above sea level, mean annual precipitation ranging from 800 to 2000 mm, annual temperature regime of 10 to 30°C, soil ranging from sandy to gravelly and clayey loams. *Lantana camara* populations in study region tended to be more towards northern and western limits as compared to southern and eastern regions (Figure 2).

The derived model further illustrated that regions encompassing Haridwar, Pauri Garhwal, Dehradun, parts of Tehri Garhwal and Almora appears to have high infestation of *Lantana camara*. The region in and around Jim Corbett and Rajaji National Parks are highly infested with *Lantana camara* where species is reported to cause nuance already. Most regions of Jim Corbett National Park manifest very high invasion distribution of *Lantana camara*. Invasion is severe in Jhirna, Dhikala and Bijrani zone. Subsequently, species expansion varies from medium to high in south-western zone of park. Other range regions such as Morghatti, Halduparao, Mundaiyapani, Hathikund and Kalagarh appeared to spearhead a more recent movement of species in eastern and northern regions of National Park. In Rajaji National Park alone, very high infestation of weed population is accounted in southern and western regions including Haridwar,
Dholkhand, Motichur and Chilla range. Other regions such as Chillawali, Kansrao, and Mohand, showed a more recent movement of species in north and eastern portions (Figure 2). Thus, Maxent predicted *Lantana camara* potential distribution very well and its modeled distribution is in congruence with reported presence in this region [17] and good uniform coverage compared to models produced with GARP and Biomapper. Hence its prediction can be called as *good prediction* as only small sets of points were left out.

Both algorithms, GARP and Biomapper produced maps that coincided well with the known distribution of *Lantana camara*, although GARP prediction tended to be overly extensive which may be called as expansive predictions i.e. including areas larger than the distribution of test occurrence points (Figure 3). On the contrary, predictions inferred from Biomapper can be called to be limited as only a few presence locations of species were successfully predicted, particularly in the western region of Himalaya encircling small portions of Jim Corbett and Rajaji National Parks (Figure 4).

In both Jim Corbett and Rajaji National Parks trends similar to overall study area was observed across all models (Figure 5). In GARP models, both national parks have been predicted to be almost completely infested with very high density of *Lantana camara* whereas Biomapper models still predicted very low spread in these parks Maxent likely illustrated the moderate scenario of *Lantana camara* infestation prevailing in these parks which is considered close to the real world settings.

![Figure 3. Invasion distribution model of Lantana camara using GARP algorithm](image-url)
Maxent projection estimated 15.76% of the area to be highly invasive. Model derived using GARP predicted 34.9% of the study area as very high invasion potential whereas Biomapper predicted 9.56% of the area.

3.2. Model validation

Testing modeled results with quasi-independent datasets using the AUC scores for these 3 models suggested that GARP models are over predictive and Biomapper models are under-predictive. The AUC values of 0.9456 and 0.7611 respectively for GARP and Biomapper than Maxent models which holds a good prediction with AUC values of 0.985 with training and 0.980 for test datasets. AUC values based on independent testing data coincided well with results of training datasets, thus, showing no significant differences between the two datasets of Maxent models.

The chi-square statistic for Maxent models for *Lantana camara* was 81.45 and hence considered statistically significant. However, predictions involving Biomapper and GARP derived models were significantly not better than a random model with respective values as 51.04 and 63.1, despite broader predictions by GARP models.

Since, Maxent model outperformed in discerning invasion potential distribution of *Lantana camara*, accuracy assessment were carried out only for these models obtained from training and test datasets. The corresponding accuracy assessment shows the frequencies with which classes remained same (along the diagonal) or changed (off-diagonal). A very high agreement between training and test datasets derived models was evident, thereby indicating a high predictive accuracy for generated model. Further, percentage changes within the suitability classes shown in Table 2 reveals that areal range changes in maps remain almost constant despite use of two independent occurrence records. Kappa statistics value of
0.81 (excellent) clearly suggests that derived models are significantly better in correctly predicting most known locations of *Lantana camara* presence in the study region.

**Figure 5. Invasion distribution of Lantana *camara* in Jim Corbett and Rajaji National Park**
Table 2. Accuracy assessment of training and test datasets derived Maxent models

<table>
<thead>
<tr>
<th>Invasion distribution of Lantana camara</th>
<th>Test Data</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Very Low</td>
</tr>
<tr>
<td></td>
<td>(0 - 10)</td>
</tr>
<tr>
<td>Very Low (0 - 10)</td>
<td>87.38</td>
</tr>
<tr>
<td>Low (10 - 30)</td>
<td>6.51</td>
</tr>
<tr>
<td>Moderate (30 - 60)</td>
<td>4.93</td>
</tr>
<tr>
<td>High (60 - 80)</td>
<td>0.59</td>
</tr>
<tr>
<td>Very High (80 - 100)</td>
<td>0.59</td>
</tr>
</tbody>
</table>

3.3. Effect of predictor variables

Results, therefore summarizes that Maxent models predicted significantly better than GARP and Biomapper as shown by better fit to the training data and validation datasets and the accuracy of prediction. Thus, Maxent model was further tested for determining the importance of variables in model development. Jackknifing was performed to determine the environmental variables which have significant contributing effects on Lantana camara invasion distribution. The models were run several times with exclusion of variable which had nil/least significant contribution towards model runs. All variable with training gain above 0.5 were considered and was obtained after 3 runs of Maxent algorithm. The jackknife diagnostics showed that the most important variables (above 5%) explaining presence of Lantana camara were land use (19.3% contribution to Maxent model), NDVI in October (17.4%), Bio 02 - mean diurnal range (13.1%), soil texture (12.9%), NDVI in July (10.0%), and Bio 12 - annual precipitation (8.8%) as these variables presented higher gain i.e. contained the most information compared to other variables. Also, elimination of these resulted in drastic decrease in model gain. Seven other variables which had minor contributions (above 1%) to model build include potential evapotranspiration in March (3.8%), Bio 02 - precipitation seasonality (3.4%), NDVI in January (2.7%), Bio 14 - precipitation of driest period (1.8%), altitude (1.4%), Bio 14 - Precipitation of wettest period (1.3%) and NDVI in April (1.1%).

4. Discussion

Despite the growing exploitation of niche modeling algorithms to predict spatial patterns of ecological invasions, the development of best suitable model using available environmental variables and historical records of species presence have been rarely taken into account for the development of invasive species management plans. In this study, we assessed the invasion distribution models using niche modeling algorithms. The algorithms were used to determine the best models predicting distribution of Lantana camara in Western Himalayan regions.

Various authors have used modeling algorithms such as GARP, Maxent, Biomapper, Climex, FloraMap, Domain, and Envelope to model the ecological niche of Lantana camara. Peng et al., reported on the development of an integrated framework to predict and simulate the dynamics of Lantana camara.
camara in Hokianga, Northland, New Zealand [47]. The result illustrated that Lantana camara is expanding its range by both neighborhood and long distance dispersal. Another attempt was made by Robertson et al. [36] who used PCA-based modeling technique to study invasion of Lantana camara in South Africa, Lesotho and Swaziland. Bioclimatic predictor variables were used in the study and most of the known regions of invasion were determined accurately. Day et al. [35] used Climex model to determine magnitude and direction of spread. The study revealed that 27 countries were identified as not having Lantana camara but have an Environmental Index (EI) greater than 30. This implies that Lantana camara could establish and has the potential to cause harm if introduced. In another study by Lüi, the derived model using Analytic Hierarchy Process (AHP) approaches suggested Lantana camara spread in north China [38].

The present study is in conformance and consistent with various studies cited above for modeling Lantana camara distribution as similar parameters were found to influence distribution of species. Lantana camara distribution was discernible using 3 niche modeling algorithms, however, spectrum of prediction varied. Amongst all, Maxent stood out among the 3 algorithms tested under various sets of evaluation and validation diagnostics. Maxent model characterized most distribution of Lantana camara, by including as many test points thus showing no failures unlike other. The derived model was in congruence with various research and field surveys. Maxent besides generating distribution model, provided an output that illustrates important environmental variable regulating Lantana camara spread in the study region. Thus in context predicting Lantana camara invasion distribution, it is worth noting, that Maxent approach surpassed imposed validation and evaluation diagnostics.

Discrepancies between models and previously recorded range included prediction of suitable habitat in higher ridges of Western Himalayan region by the GARP model and under-prediction of suitable habitat in the lower ridges of Western Himalayan region by Biomapper derived models whereas the extent of the present suitable habitat range of the Lantana camara currently documented and field verified is to be in lower ridges of Western Himalayan region. This incongruence can be attributed either as failure of model to derive ecological niche of the species, with models including regions not ecologically suitable for them, or the derived models are correct, and Lantana camara is not found on the higher ridges of Western Himalayan region because of historical or biological barriers or limited dispersion potential. The latter scenario making range expansions plausible if these limitations are overcome. Moreover, under predicting models may be interpreted as lacking important variables that are essential in discerning the species suitable habitats [48] or relied on inconclusive field data [42]. Thus, well performing models are construed with identified key predictor variables elucidating the habitat suitability of the species [48] which is very much evident with model derived using Maxent algorithm.

The predictor variable analysis reveals that Lantana camara is found to be strongly associated to areas with high forest cover. Furthermore, it was found to be positively related to diurnal temperature, with favorable temperature above 20°C, and associated to areas with annual precipitation above 1000 mm; suitability was limited in lowlands with open forest cover. In addition, Lantana camara is a characteristic weed of human-modified habitats and thus the significant contribution of land use parameter to invasion
distribution model is quite evident. Furthermore, the likely contribution of potential evapo-transpiration is borne out by the fact that *Lantana camara* benefits from soil moisture occurring in semi-arid to normal soil, partially shaded, and irrigated (including roadside runoff) habitats. The observed superiority of elevation gradients is in concordance with the expected behavior of the species since the species flourishes well at lower altitudinal ranges and as it increases, the species occurrence tends to diminish [49-51]. In addition, human disturbance serves as a proxy variable that is dependent on elevation and affect the ecological processes of *Lantana camara* [18]. Thus, the species are mostly encountered in lower altitude as anthropogenic disturbances are high in these regions. The slightly higher annual precipitation and lower mean temperatures in the western Himalayan region imply lower evapo-transpiration that could enable *Lantana camara* to flourish well. Beside these, soil textures are expected to influence its abundance.

5. Conclusion

The present research complements previous studies of comparative analysis of model algorithms w.r.t. plant invasion, however, uniqueness of study lies in the fact that it’s a novel attempt to model *Lantana camara* invasions potential at micro scale which has not been attempted earlier in India. Also, empirical evidence is provided that predictive models calibrated with field data and augmented with habitat describing predictor variables can significantly determine species presence without actual presence in the field and the best models amongst all can be critically assessed and utilized to develop management strategies. One of the major offerings of the current research work is the generation of the first map detailing the modeled potential distributions of the *Lantana camara* based on best model selection algorithms, which can be dynamically updated with evolving information. Our results contribute to the broad ecological understanding and conceptualization of invasive species management as it’s based on efficacy of geospatial predictions. If invasion distribution maps used as early detection tool and management of invasive species in conservation practice, their accuracy and correct interpretation could help in minimizing the ecological implications and economic cost of invasions. Beside these, elucidating the exact region of invasive plants such as *Lantana camara* will be crucial to develop ecological protocols for channelizing the conservation measures and enable to conserve and foresee potential invasions in the future.

The current research had several constraints such as presence-only datasets were for spatial model development and testing; factors such as limited dispersal, speciation, and extinction impacts were not considered in model development though these historic factors limit species geographic distribution than its ecological requirements are seen. In addition, output of ecological niche modeling, such as this one, are dependent on the occurrence records, environmental envelope and resolution at which the study is being conducted. Data for all variables which may be influencing the niche of *Lantana camara* may not be available or the resolution at which study is conducted may not be able capture ecological variations. However, greater knowledge of species and the availability of data may resolve such issues encountered in the present study.

References


[40] Hijmans R. J., Cameron S. E., Parra J. L., Jones P. G., Jarvis A.: Very high resolution interpolated climate surfaces for global land areas. International
DOI: 10.1002/joc.1276


DOI: 10.3390/d2050738

DOI: 10.1016/S0304-3800(01)00396-9

DOI: 10.1016/j.ecolmodel.2006.05.017


DOI: 10.1111/j.0906-7590.2006.04700.x

DOI: 10.1016/S0304-3800(02)00204-1


DOI: 10.1126/science.285.5431.1265


DOI: 10.1111/j.0021-8901.2004.00896.x