Geo-spatial Analysis of Habitat Suitability for Common Leopard (Panthera pardus Linnaeus, 1758) in Shivapuri Nagarjun National Park, Nepal

Bishnu Maharjan1,*, Shahnawaz2, Tej B. Thapa3, Purna Man Shrestha4

1Geographical Information Science and System, University of Salzburg, Austria
2Interfaculty Department of Geoinformatics – Z_GIS, University of Salzburg, Austria
3Interfaculty Central Department of Zoology, Tribhuvan University, Kathmandu, Nepal
4Central Department of Zoology, Tribhuvan University, Kathmandu, Nepal

Copyright©2017 by authors, all rights reserved. Authors agree that this article remains permanently open access under the terms of the Creative Commons Attribution License 4.0 International License

Abstract  Existence of predators like common leopard (Panthera pardus) is associated with high biodiversity, so the protection of their habitats is one of the most effective ways to conserve biodiversity globally. Considering the facts above, the main objective of this research is to predict and map the possible habitat for Common Leopard in Shivapuri Nagarjun National Park by using remote sensing and GIS approach. In order to achieve that, Species Distribution Modeling (MaxEnt) was demonstrated to predict the Common Leopard’s distribution and was applied to figure out possible suitable area in Shivapuri Nagarjun National Park. By using presence – only data of Common Leopard (Panthera pardus) occurrences, 138 observation points alongside several environmental variables which consist Distance from Settlement Area, Forest, Bush, Road, Sparse Forest and Agricultural Area were developed in to MaxEnt Programme. Remotely sensed imagery of ResourceSat-2 imagery used by applying supervised classification in order to determine land use land cover characteristics of the study area. Image processing and feature extraction was done by Erdas Imagine 2011 and maximum likelihood supervised classification was done. The contribution of variable “Distance from Settlement Area (52.4%)” was highest to impact the model. The model performance was accessed through using Receiver Operating Characteristics (ROC) plots and Jackknife tests. The Area under Training data (ROC) curve (RUC) 0.828 and that of Test data ROC curve was 0.678 which is acceptable than the Random Prediction Model (AUC) of 0.5.

Keywords  Species Distribution Modeling, American Standard Code for Information Interchange (ASCII), Global Positioning System (GPS), Maximum Entropy, Receiver Operator Curve (ROC), MaxEnt Modeling, Shivapuri Nagarjun National Park (SNNP)

1. Introduction

Leopards (Panthera pardus) are the most widely distributed wild cats, and occupy a broad variety of habitats, from rainforests to deserts and from the fringes of urban areas to remote mountain ranges (1). The leopard (Panthera pardus) is one of the most widely distributed felids across the forested landscapes of the Indian subcontinent (2). This spotted cat has short powerful limbs, heavy torso, thick neck, and long tail. Large black spots grouped into rosettes on the shoulders, upper arms, back, flanks and haunches, and smaller scattered spots on the lower limbs, head, throat and chest, and the belly has large black blotches. In the world, there are 36 species of wildcat exists. The Leopard have low conservation priority because of their widespread distribution and ecological flexibility, however, global population status is still uncertain (3). The Wild Cat Status Survey (IUCN/SSC Cat Specialist Group) has categorized Leopard as the Near Threatened species (3). Due to habitat conversion or fragmentation, trade of body parts, trend in decreasing number of Leopard, the International Union for Conservation of Nature (IUCN) listed its eight subspecies as endangered or critically among 14 sub species of Leopard (4). With the rise of new powerful statistical techniques and GIS tools, the development of predictive habitat distribution models has rapidly increased in ecology (5). Such models are static and probabilistic in nature, since they statistically relate the geographical distribution of species or communities to their present environment. The tool that can help us study habitat selection at the scale of the species range, and that has been particularly useful in the field of conservation, is habitat suitability modeling (6). Recently, interest in species distribution modeling has increased following the development of new methods for the analysis of presence-only data and the deployment of these methods in user-friendly and powerful computer programs. However,
reliable inference from these powerful tools requires that several assumptions be met, including the assumptions that observed presences are the consequence of random or representative sampling and that detectability during sampling does not vary with the covariates that determine occurrence probability (7). Prediction and mapping of potential suitable habitat for threatened and endangered species is critical for monitoring and restoration of their declining native populations in their natural habitat, artificial introductions, or selecting conservation sites, and conservation and management of their native habitat (8). Habitat-suitability modeling is being increasingly used as a tool for conservation biology. Although studies at large spatial scales are more appropriate for reserve design and management, there is a scarcity of published work on local, high-resolution applications of such models (9).

2. Materials and Methods

In order to develop a model for assessing the distribution of Common Leopard, alongside its habitat contiguity, several other factors need to be taken into consideration. The basic data sets required for predicting the probable habitat for Common Leopard (i.e. study area, physiography, land use land cover etc.) need to be prepared as environmental layers beforehand. Nonetheless, the factors impeding movement of Common Leopard will have strong influence on determining the suitability of habitat.

The research was conducted in Shivapuri National Park which was designated as the 9th national park of Nepal in 2002. It covers an area of 159Km² encompassing two separate forest patches viz. Shivapuri (144km²) and Nagarjun forest (15km²). It is located between 27°45' to 27°52' northern latitude and 85°15' to 85°30' eastern longitudes. It spreads over Kathmandu, Nuwakot, Dhading and Sindhupalchowk districts of Central Nepal region. The extent of the park is demarcated by a 111km long boundary wall and 95 km long ring road. The elevation in the park ranges from 1350m to 2732m and this is the only protected area lying entirely within the Nepal’s mid hills ecosystem. It is located on the northern fringe of Kathmandu valley and lies about 12km away from the capital city. The Shivapuri peak having an elevation of 2732m is the second highest peak around Kathmandu valley and the upper slopes are covered with forest (10). The Shivapuri Nagarjun National Park is situated towards the north of Kathmandu valley and it contains a number of streams and rivulets. This forms one of the primary sources of freshwater supply for Kathmandu valley and it provides over 40 percent of the drinking water to the Kathmandu valley.

2.1. Data Used

I. Common Leopard presence point was collected through village survey, interview, direct and indirect evidences, in Shivapuri Nagarjun National Park.

II. ResourceSat-2 Satellite imagery dated on 25th January 2015 from Indian Space Research Organization (ISRO) having spatial and spectral resolution 5m and Swath Width 70km.

III. Topographic data and land cover of study area from Survey Department, Government of Nepal.
2.2. Maximum Entropy Modeling

As a machine learning method which requires presence only data in modeling, Maximum Entropy (MaxEnt) has high accuracy in predicting geographic distribution of the species. Basically, maximum entropy can be applied to solve the problem in any constraints. The principle of Maximum Entropy in species distribution exposes unknown probability of species occurrence over a set of points or pixels in the study area. The value will be 0 or 1 for plants and range from 0 to 1 to animals which depict the probability of species at every point/pixel. The aim behind this idea lies on the incapability of determining species prevalence only by occurrence data.

2.3. Leopard Presence Point

The presence point of Common Leopard was collected through village survey, interview, direct and indirect evidences, in Shivapuri Nagarjun National Park. The device eTrex 30 channel Global Positioning System (GPS) was used to take the Coordinate of respective Leopard presence point. Then GPS location of total leopard and other points were converted to UTM WGS 84 zone 44N projection for subsequent GIS integration. The relative abundance of leopards was estimated by walking on forest trails, fire line and grassland. The location of leopard presence was recorded through indirect method i.e. scats, pugmark, scraps and opportunistic search method. The combination of these kinds of data was applied in the modeling process by using MaxEnt Program.

2.4. Land Cover

Images used in this study have been captured by LISS-IV sensor aboard Resourcesat-2 of Indian Space Research Organization (ISRO). The images were captured on 25th January 2015 and these were cloud free. LISS-IV provides images of Earth’s surface in high spatial and spectral resolution (spatial resolution 5.8m Swath Width: 70km) which are appropriate for vegetation analysis in mountainous landscape. The data used were a 3 x 12,288 MS mode product which was geometrically (systematically) corrected. The LISS-IV images contain three multispectral bands B2: 0.52-0.59, (green), B3: 0.62-0.68, (red) and B4: 0.77-0.86 (NIR). Based on the visual interpretation and field observations, 6 Land Use Land Cover (LULC) classes were identified on the LISS-IV images and it was classified using Supervised Classification method in ArcGIS 10.2 software. The six LULC classes are following:

i. Dense Forest
ii. Sparse Forest
iii. Road
iv. Agricultural Land
v. Settlement Areas
vi. Bushes/Shrub/Grassland

Supervised classification was performed more than one times using different numbers of training areas for each LULC as well as by varying the size of the training areas. Also supervised classification was performed by merging multiple training areas per class. The results were compared, and it was found that the LULC classification obtained from umerged training areas gave better results. The average overall accuracy for the supervised maximum likelihood classification was 92.523%. The overall Kappa statistics was also high (0.92). Both the producer’s and user’s accuracies were over 90 per cent.
Map 3. Land Use and Land Cover of Shivapuri Nagarjun National Park

Map 4. Euclidean Distance of Study Area
2.5. Distance from Settlement

Despite of Leopards’ characteristics as the most adaptable big cat species, Gunawan et al. (11) described that Leopards tend to keep distance from human settlements approximately more than half a kilometer. Also human-wildlife conflicts are inevitable if the settlements and protected areas are located in close proximity to each-other (12). This creature also tends to remain away from noisy places like roads and paths in this case. In order to provide distance from settlements and roads/paths as an environmental layer to the leopard presence points, Euclidian distance was calculated separately for settlements and roads/paths in ArcGIS 10.2. Similar process was repeated for calculating Euclidian distance from other LULC classes like agricultural land, bushes, dense forest and sparse forest.

2.6. Preparing Environmental Layers for MaxEnt

MaxEnt program requires presence points in Comma Separated Values (CSV) format and the environmental layers in ASCII format for execution of the model. It is important that all of the environmental layer must have the same geospatial parameters. Therefore, the layers of all the above mentioned variables were processed carefully for maintaining exactly the same geospatial properties. As mentioned above, all of the environmental layers were generated from LISS-IV image having cell size of 5.8m. The fractional of the resolution could create cell-shift problem while extracting the layers according to the boundary of the study area leading to some shift in the extent of different environmental layers so it was important to resolve this problem. First of all, the boundary of the study area was converted into rater format at 5m cell size. Then all of the environmental layers were resampled to the same resolution i.e. 5m. The raster format boundary layer was used as the mask for extracting all the environmental layers. Resultantly, all the layers had same spatial extent in the same coordinate system i.e. 5178 column and 2441 rows in WGS_1984_UTM_Zone_44N. Lastly, all the extracted environmental layers were converted from raster to ASCII format in the exactly same bound, cell size and coordinate system and saved in the environmental layers folder.

2.7. Running MaxEnt Model

MaxEnt programme requires two types of inputs for model processing i.e. presence points as ‘samples’ and selected factors influencing the distribution as ‘environmental layers’. For common leopard distribution modeling, a CSV file containing presence points collected during the field observations was given as samples and, on the other side, the folder containing environmental layers in ASCII format was added in MaxEnt’s ‘environmental layers’ menu. MaxEnt requires the user to define environment layers either as ‘categorical’ or as ‘continuous’ depending on the nature of the dataset. In this case, only land use land cover layer contained ‘categorical’ data and all other layers contained continuous data. Next to this, the user needs to select the types of desired outputs e.g. response curve, pictures of predictions, jackknife to measure variable importance, output format and output file type. In order to determine random test percentage, basic setting of MaxEnt has been set to 25 in 1 replicates. On replicated run type menu, bootstrap was chosen in relatively few observations of presence points. Finally, ‘equal training sensitivity and specificity’ was chosen as threshold rule in ‘advanced’ settings while ‘write background predictions’ was ticked in the ‘experimental’ settings to obtain pseudo background data of AUC and TSS calculation process.

2.8. Model Performance

The receiver operating characteristic (ROC) analysis is a threshold independent method and this is used widely for evaluating the accuracy of classification models (13). The ROC curve is generated by plotting sensitivity values (i.e., fraction of true positive) against 1-specificity values (i.e., fraction of false positive) for every possible threshold (14). The AUC is a comparison of the true positive rate and the false positive rate, or how well the model is able to predict presence and absence. MaxEnt uses presence-only data and AUC created for the MaxEnt models shows how well the model is able to distinguish presence from random (15). The value for the AUC ranges from 0 to 1, the closer the value of the AUC is to 1, the better is the fit of the model. An AUC value of 0.5 equals random prediction (16). The area under the ROC curve (AUC) provides a single-value measurement of model performance. Since omission errors reduce sensitivity and commission errors reduce specificity, both types of errors equally reduce the AUC value. While an AUC value of 1 indicates a perfect model, a value of 0.5 indicates a random model. A standard for judging model performance based on AUC values (17).

3. Results

3.1. MaxEnt Output

The result are shown in Figure 1 where ‘A’ shows the Omission rate and predicted area by using cumulative threshold and ‘B’ shows Receiver Operating Characteristics (Sensitivity Vs. 1 - Specificity) on Leopard. In figure 1 (A), X axis shows the cumulative threshold and Y – axis shows fractional values. The red line presents fractions of background predicted, the blue line indicates the omission on training samples, and the skyline shows the omission on test samples whereas the black straight line indicates the predicted omission rate. This graph shows how testing and training omission and predicted area vary with the choice of cumulative threshold. So, omission on test samples is good to match for test data drawn from the MaxEnt distribution itself.
The graphs presented in Figure 1 B display the omission rate and predicted area as a function of cumulative thresholds. The omission rate is calculated both on training presence records as well as on the test records if test data are used. The omission rate should be close to the predicted omission because of the definition of the cumulative thresholds (18). Therefore, analysis of omission rate and predicted areas as function of cumulative threshold showed that omission rate was close to the predicted omission depicting the model to be robust to conduct further analysis. The ROC curves in this model shows high accuracy of the generated model with AUC 0.828 for training data and 0.678 for test data. The red line shows the ‘fit’ of the model to the training data. The value of 0.5 indicates that the performance of this model is no better than random while values greater than 0.5 represent good model performance.
3.2. Analysis of Variable Contributions

<table>
<thead>
<tr>
<th>Variable</th>
<th>Percent Contribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Settlement Area</td>
<td>52.5</td>
</tr>
<tr>
<td>Road</td>
<td>16.5</td>
</tr>
<tr>
<td>Forest Area</td>
<td>15.8</td>
</tr>
<tr>
<td>Agricultural Area</td>
<td>14.1</td>
</tr>
<tr>
<td>Bush</td>
<td>1</td>
</tr>
<tr>
<td>Sparse Forest</td>
<td>0.3</td>
</tr>
</tbody>
</table>

The ‘Variable Contribution’ estimates relative contributions of the selected environmental factors to the MaxEnt model. To determine the first estimate in each iteration of the training algorithm, the increase in regularized gain is added to the contribution of the corresponding variable or subtracted from it if the change to the absolute value of lambda is negative. For the second estimate, for each environmental variable in turn, the values of that variable on training presence and background data are randomly permuted. The model is reevaluated on the permuted data, and the resulting drop in training AUC is shown in the (Table 1) normalized to percentages. As with the variable jackknife, variable contributions should be interpreted with caution when the predictor variables are correlated.

3.3. Jackknife Test

The results of the Jackknife test of variable importance are shown in Figures 2 A, B and C. The environmental variable of highest gain when used in isolation is ‘distance from settlement area’ (Figures 2 A) followed by distances from road, dense forest, agriculture, sparse forest and bush being the least important in predicting the distribution of Common Leopard. The ‘distance from settlement area’ also shows that if it is exclude from the test then the gain drops to the lowest and significantly lower than all other variables. This means ‘distance from settlement area’ is the most important variable in training data that play a significant role in predicting the distribution of Common Leopard in the study area.

Figure 2. Jackknife test of MaxEnt Output: A. Jackknife result of variable in regularized training gain for Common Leopard; B. Jackknife result of variable in the AUC for Common Leopard; C. Jackknife results of variable importance in the test gain for Common Leopard
Geo-spatial Analysis of Habitat Suitability for Common Leopard (*Panthera pardus* Linnaeus, 1758)
in Shivapuri Nagarjun National Park, Nepal
Figure 3. Response curves for the most significant predictors of habitat suitability of Common Leopard according to the MaxEnt model
3.4. Response Curve

The response curves show how the predictions depend on the variables shown in (Figure 3 A-F). The curves show how the logistic prediction changes according to the nature of each environmental variable. In other words, the curves show the marginal effect of changes exactly from one variable to the next whereas the model can obtain advantages in the process of variables changing together.

*P. pardus* as agricultural distance increases the probability of its presence decreases from about 0.55 to almost 0 at distance of 1200m (Figure 3A). The response curve for “distance to agriculture” shows a negative relationship with logistic output for all common Leopards.

As distance to bush increases the chance of finding *P. pardus* also increases. From the response curve of bush – it is apparent that *P. pardus* exists outside the bush as its probability of occurrence starts increasing beyond a distance of 600m (Figure 3B). The probability of finding *P. pardus* doesn’t however depend so much on bush as it doesn’t achieve high probabilities. This might be probably due to the absence of prey species in bush and their requirement for arboreal type of habitat. As leopard hunts cattles normally, it surrounds around settlement zones to hunt domestic animals as well.

In terms of association with forest there is a low probability of finding *P. pardus* as distance increases as shown in the (Figure 3C). *P. pardus* has very high probability of occurrence on the periphery of road up to the distance 800m as shown in the (Figure 3D). Response curve also shows negative relationship with roads.

Generally probability of presence of *P. pardus* also increases up to distance 750m and whereas decreases up to 1200m (Figure 3E). The probabilities of presence have been predicted as low value above 1500m, despite prey abundance in such areas. However *P. pardus* tends to show relatively lower probability of presence in the sparse forest areas above 1400m (Figure 3F).

4. Discussion

4.1. Habitat Suitability Map

MaxEnt generated a habitat suitability map. This map is then classified on the different species occurrence probability threshold class. By using specific probability thresholds which was classified in the map determines habitat suitable parameter in the specific location. However the MaxEnt predicted map was uses colors that indicate predicted probability that conditions are suitable. Warmer colors (red) indicate low probability of suitable conditions for the species and blue indicates high probability. Therefore suitability map was reclassified into four classes; Unsuitable, Moderately Suitable, Suitable and More suitable (Map 5). The Unsuitable category included the areas that have least probability for Common Leopard to occur.

MaxEnt modeling has proven to be very effective at determining habitat use and species distributions for a variety of species and localities. It shows majority of Habitat “Suitable” patches around the SNNP. For enhancing the habitat improvement of Common Leopard in SNNP, certain improvement and interventions has to be carried out. This model also suggest that increasing number of “agricultural land, bushes and Forest” helps in increasing suitable habitat of Common Leopard (*Panthera pardus*). Suitable habitat for Common Leopard (*Panthera pardus*) has been identified and mapped. Agricultural Land, forest, bushes areas are occupied by the Common Leopard which are predicted as suitable habitat for leopard. The leopard prefers to habitat type of agricultural land until bush and forest as these areas have greater affinity towards prey base that serves as food for them. The result of MaxEnt model of Common Leopard has performed the AUC value of 0.99. It is considered as an excellent model because of the value more than 0.9 (21). Therefore, this presence distribution prediction was suitable for Common Leopard. On this model, variables which were represented by environmental layers showed its contribution percentage. Among the 6 variables which have been deliberately chosen, the most important variables of this model consist of settlement area, forest, bush, sparse forest and roads. The potential threat in terms of land use which might be faced by the settlement area becomes the most serious threat to conservation and sustainable development in general. The prediction at the NP’s boundary can be regarded as the potential distribution as well as the prone area for conflict between the leopard and humans (their cattle). Delimit human access to particular zones which are restricted to any disturbances and intensify the survey of wildlife can be the noticeable action to conserve Common Leopard.
Map 5. Habitat Suitability Map

Table 2. Threshold used to predict logistic output into classes

<table>
<thead>
<tr>
<th>Habitat Class</th>
<th>Probability Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unsuitable Habitat</td>
<td>0.00 – 0.20</td>
</tr>
<tr>
<td>Moderately Suitable</td>
<td>0.20 – 0.38</td>
</tr>
<tr>
<td>Suitable</td>
<td>0.38 – 0.60</td>
</tr>
<tr>
<td>More Suitable</td>
<td>0.60 – 0.96</td>
</tr>
</tbody>
</table>

Table 3. Predicted Suitable and Unsuitable areas for Common Leopard

<table>
<thead>
<tr>
<th>Habitat Class</th>
<th>Area (Km²)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unsuitable</td>
<td>41.74</td>
</tr>
<tr>
<td>Least Suitable</td>
<td>47.10</td>
</tr>
<tr>
<td>Suitable</td>
<td>52.27</td>
</tr>
<tr>
<td>More Suitable</td>
<td>13.19</td>
</tr>
</tbody>
</table>

The map was categorized using the threshold 0.00 – 0.20 as unsuitable while from 0.20 – 0.38 as least suitable. The suitability of Common Leopard in Shivapuri Nagarjun National Park is categorized (Table 2). The result of habitat suitability map shows that the total potential suitable habitat for Common Leopard in SNNP is 112.56 Km² while other remains unsuitable habitat for Common Leopard of area 41.74 Km² in SNNP.

As it includes the land use and land cover types there in proximity of Agricultural Land. For achieving this extent of areas as suitable, number of water bodies has to be significantly increased and maintained. When water resource has maintained, automatically the agricultural land becomes increases which significantly increases the occurrence of Common Leopard in Shivapuri Nagarjun National Park. As during supervised classification of Resource Sat II Satellite Imagery, water bodies are explicitly removed because it is not clearly observed in image.

5. Conclusions and Recommendation

There has never been a scientific research conducted in SNNP about Common Leopard. Thus a detailed, scientific study of common leopards is very necessary in SNNP. This could provide an estimation of the number of common leopards in SNNP. The livestock depredation caused by the species should also be addressed during the study. Conservation education must be included in the curriculum of school which provides students the knowledge about the importance of the leopard. Conservation education must also be provided to the villagers about the role the species play in balancing the ecosystem by acting as the supreme predators of hilly region in the food chain. Brochures, posters, leaflets and other publications must be prepared and distributed showing the importance of common leopard and the benefits one can get from it. This could help in making people aware of its importance thus helping in its conservation. Meanwhile local government particularly Kathmandu district can take
this issue of connecting ecology into account on its spatial planning. In order to make sure the prediction of Common Leopard’s presence in SNNP deploying several camera traps within the areas denoted as presence would be beneficial for the next level of wildlife management. By using camera traps, either Common Leopard or other wildlife can be recorded as the main attention in managing the park. Conducting presence – absence survey of Common Leopard in SNNP will give another option of species distribution modelling such as GLM, GAM, and BRT etc. As the prominence of land cover, environmental layer, NDVI and satellite imagery, applying very high resolution of remotely sensed imagery to obtain more detail Landover will produce more precise result in predicting Common Leopards’ distribution and connects human wildlife welfare scenario. As the prominence of Land Use and Land Cover on this model, applying high resolution of remotely sensed imagery to obtain more detail land cover produce more precise result in predicting Common Leopard’s distribution in SNNP. During the study, it was evident that habitat destruction is quite rampant in the 58 study area. Habitat encroachment is one of the main reasons for the leopard to turn its attention towards human settlements that results in livestock depredation. Therefore deforestation and encroachment of the leopard habitat must be discouraged properly. A database must be prepared and maintained by conducting a detailed study about the leopard in the area by the concerned authority which contains everything about the leopard’s situation/condition in the area. For e.g. it’s potential and actual habitat, its natural prey base, prey – predator relationship and so on.

REFERENCES


