Semi-automatic delimitation of Buffer Zones for protected areas: A contribution from Species Distribution Modelling

Cuitlahuac Hernandez-Santiago\textsuperscript{1,2}, Martin Volk\textsuperscript{1}.
\textsuperscript{1}Helmholtz Centre for Environmental Research – UFZ, Leipzig, Germany.
\textsuperscript{2}Universidad del Mar, Oaxaca, Mexico (cuitlahuac.hernandez@ufz.de)

Abstract: Preservation of natural protected areas in Mexico requires human activities regulation in the Buffer Zones (Zona de Influencia). These zones have been defined conceptually by law and delineated spatially using physiographical and anthropogenic criteria. The presented study proposes a demarcation based on species distribution modeling as an ecological criterion to prioritize environmental requirements of protected species in “Laguna de Chacahua” National Park. To achieve this, a model in ArcGIS ModelBuilder was developed and applied for the delimitation of a buffer zone for the National Park. The proposed model involve a) obtaining data of 90 species (mammals, reptiles and amphibians) of which 30 are listed as protected; b) data cleaning and pre-processing of raster layers; c) processing with the species distribution model MaxEnt, d) averaging and generalization of threshold line, and e) post-edition and visualization. The results showed that the delimited area represents 10 times the size of the National Park and imply that species have a wider distribution than the current protected area. As a consequence, the current protected area is probably not large enough to safeguard the studied species. On the base of the newly delimited Buffer Zone, political agreements between park managers and municipal authorities are required to promote sustainable park management that foster protection of registered species.

Keywords: Chacahua; species distribution modelling; buffer zone; national park.

1 INTRODUCTION

Increasing human population growth has led to an extended pressure over continental and marine resources. In terms of biodiversity, this exploitation has induced the loss of several species and created a societal response to find alternatives to avoid or reduce this impact. As many other countries in the World, Mexico has promoted biodiversity preservation through Natural Protected Areas (NPA), which is defined in the D.O.F. [1988] (Spanish acronym of Official Journal of the Federation) as “zones where original environments have not been significantly altered by human activity or require to be preserved or restored”. As stated in D.O.F. [2000], every NPA should demarcate a Buffer Zone (Zona de Influencia) taking into account the social, economic and ecological interactions with the surroundings. Hence, such buffer zones have a major ecological importance for NPA’s due to the interactions in terms of micro-climatic regulation, physiographic and hydrologic processes, genetic diversity, etc.

Buffer Zone was initially included in US legislation in the 1930’s to promote natural parks and reserves protection. During 1976, it was promoted by the UNESCO Programme \textit{Man and the Biosphere} to be included in many countries natural reserve schemes.
According to Martino [2001], their goals have changed from being areas where hunting was allowed, to areas where only practices for conservation of forest and wildlife are permitted to reduce fringe effect in reserves. Shafer [1999b] suggested delimitation criteria where the NPA is defined as core area, surrounded by two external rings (see Figure 1). The ring next to the core area represents the Buffer zone and the external ring would be defined as a Transition zone. According to Ebregt et al. [2000], spatial delimitation for Buffer Zones has traditionally relied on socioeconomic, cultural political or physical criteria defined by managers using an anthropocentric view without a clear methodology for transferring the delimitation model to other cases.

Recently, Domínguez-Cervantes [2009] proposed a conceptual model that includes three main scales and suggested criteria: MACRO, defined by municipal territory where the protected area is located, and the surrounding take reference in the basin and sub-basins; MICRO, defined by biophysical boundaries (hydrology and geomorphology) and anthropic (roads, dams, land tenure and settlements distribution; and LOCAL, defined by sub-zones with specific description of land use dynamics or activities that represent a major thread to the protected area. These scales and criteria were applied using methodologies for evaluation of landscape units in three Biosphere Reserves in Southeast Mexico. Nevertheless, even such methodologies do not include geographic distribution of those species intended to protect.

In order to overcome this deficiency, this study suggests a semi-automatized procedure for delimiting the Buffer Zone and Transition Zone using Species Distribution Modelling (SDM). Methods to create SDM has been developed by several authors such as Carpenter et al. [1993], Hirzel et al. [2002], Lehmann et al. [2003], Pearson et al. [2002] and Phillips et al. [2006] among others. In this work, MaxEnt – developed by Phillips et al. [2006] - was used, since Elith et al. [2006] indicate that this algorithm has proved to be robust to model species distribution using presence-only records for large number of taxa and few records. Therefore the analytical results have provided the information needed to delineate the Buffer Zone.

2 DATA AND METHODS

2.1 Study area

This procedure was tested in the study area “Lagunas de Chacahua” National Park, in the Southwest coastal plains of Oaxaca State, Mexico. It is located in the Municipality of Villa de Tututepec de Melchor Ocampo. The park area is 14,187 hectares and includes a lacustrine system that covers 25% of the park and 28 km of beach. Esparza-Alvarez et al. [2011] reported 32 species of mammals, 45 species of reptiles, 13 species of amphibians and 157 species of birds as a result from recent monitoring that was carried out in the region. According to D.O.F. [2010] 47 from this group were protected species (8 mammals, 20 reptiles, 2 amphibians and 17 birds). The present work considered only mammals, reptiles and amphibians to define the Buffer Zone for the National Park.

2.2 Methodology

The study suggests the use of Species Distribution Models (SDM) as an indicator to define the Buffer and Transition Zone. To achieve this, a workflow using ArcGIS and MaxEnt as main modeling tools has been organized as follows (see Figure 2):

a) Data collection of faunal records from ManisNet and HerpNet repositories, and Universidad del Mar collection.

![Figure 1. Conceptual scheme for NPA’s zones delimitation suggested by Shafer [1999].](image-url)
b) Data cleaning as recommended by Hijmans et al. [2011] and preparing of environmental layers,
c) Modelling with MaxEnt as described by Phillips et al. [2006]. This algorithm fits a maximum entropy probability distribution of species presence-only records using explanatory environmental layers. For the presented work, 75% of records were used for training and 25% for testing. 10,000 random background samples (pseudo-absences) were used from covariate grids and cross-validation with 500 iterations.
d) Model post-processing, which has included averaging of raster models and generalization of Equal training sensitivity and specificity logistic threshold to create the areas with ArcGIS.
e) Final edition of Buffer Zone and visualization in Google Earth.

The running model in ArcGIS ModelBuilder includes the following steps: 1) Clipping and conversion preprocessing of BioClim and land cover data for MaxEnt algorithm; 2) Script to run MaxEnt using a batch file that indicates files and process to be used; 3) Importing and projecting results for processing in ArcGIS; 4) Averaging models according to species class; 5) Creating threshold lines applying Contour tool and MaxEnt output threshold values as interval, this is followed by complexity reduction of resulting lines using Smooth line tool and Polynomial Approximation with Exponential Kernel (PAEK) as method -which according to ESRI [2011] calculates a moving path based on user defined Smoothing tolerance and creates a line that will not pass through the input line vertices-; 6) Clipping resulting to the area of interest; 7) Merging all lines to a single file; 8) Conversion of thresholds lines to polygons using Feature to polygon tool to lines and area of interest files; 9) User intervention is needed to remove extra polygons created by the previous tool and using the thresholds lines as references, define the final delimitation; 10) Applying Identity tool to final delimitation, land cover and municipality layers to identify political adscription and land cover of delimited zones; 11) Exporting of final delimitation file to kml file for visualization in GoogleEarth or Google maps.

3 RESULTS

For Data collection of fauna records, www.manisnet.org and www.herpnet.org provided mammals and reptiles species records from their databases. Both are repositories of specimen records from several international collections.
using a version of Darwin Core 2 conceptual schema to organize the data for every record; they are available in corresponding websites. UMAR holds a collection of records obtained during research projects in the region. ManisNet database has provided 42,451 records and UMAR 19 records of mammals. HerpNet database has provided 56,288 records and UMAR 725 records of reptiles. For amphibian species HerpNet also has provided 11,673 records and UMAR 216 records.

For Data cleaning, we first selected the georeferenced records and removed the duplicates. In the following, we cross-checked the spatial coherence of data and georeferenced the data with the locality description for those species with a small number of records (<20). After the data cleaning process, 856 records of mammals were left, and 2,233 records of reptiles and 1,224 records of amphibians remained. The resulting database contained 4,043 records for Mexico, the average per species were 48; the 10th percentile was 2.89 and, the 90th percentile 109.8. Average training samples was 32 and average sample test 10. Environmental variables such as bioclimatic and elevation data was obtained from Hijmans et al. [2005] in the WorldClim database (www.worldclim.org). The land cover was obtained from GlobCover 2009 (Copyright ESA 2010 & UCLouvain) with description of classes by Bontemps et al. [2011]. For modeling purposes, a subset of these layers was extracted covering Mexico territory.

For modeling with MaxEnt, a script was created to run a batch file that indicates process details. MaxEnt is described by Phillips et al. [2006] as an algorithm that produce probability distributions of species in a graphical logistic model, were predictive values rank from zero (high probability of absence) to one (high probability of presence). The values are presented in a four color palette to indicate prediction strength as follows: dark blue for values close to zero, green where the species distribution is common and yellow to red to indicate sites with best environmental conditions for the species. The algorithm also produces the analytical results of models, in which the explanatory variables with higher values are selected to re-run the process in order to obtain the final models - the selection also should avoid correlated variables to prevent over-parameterization -. Modelling for each species produced a collection of 90 raster layers and analytical results. Model performance was assessed using the test Area Under the Curve (AUC), in which a value equal to 1 is an optimal model and 0.5 a random model. Our models obtained an average Training AUC of 0.973 and average of Test AUC of 0.917. The AUC standard deviation was 0.032.

For model post-processing, species used for this study corresponds to 3 groups (mammals, reptiles and amphibians) with differentiated environmental needs and protection status. To identify how the species distribution differ, all the species were grouped as follows: A) protected species (31 sp.); B) non-protected species (59 sp.); C) protected amphibians (2 sp.); D) non-protected amphibians (9 sp.); E) protected mammals (8 sp.); F) non-protected mammals (24 sp.); G) protected reptiles (21 sp.) and H) non-protected reptiles (24 sp.). To speed up the analysis, the area of interest was extracted from the models to calculate the average for every group using raster calculator. The resulting averaged models (colored images) are presented in Figure 3. Figure 3 also shows the overall pattern of larger areas with high values from East to West along the coast. Protected amphibians obtained the highest values from all models and show two main aggregation areas in the surrounding of Rio Verde to National Park and continue west. To define presence/absence in the resulting models the “Equal training sensitivity and specificity logistic” threshold was used because they were the higher values leading to a more reduced area. Having this threshold, the contour tool was used to create the lines representing the average value for each group and was generalized applying PAEK method with 0.05 decimal degrees of smoothing tolerance. The resulting thresholds are presented as dotted lines also in Figure 3. The results show that trajectories have similar behavior in all the groups from Southeast to Northwest reaching Rio Verde. From here, the delimitation of protected mammals diverges to the south involving the river and the coastal lagoons, and heads then to the North-West away from the coast. The thresholds
lines were converted to polygons using feature to polygon tool and the background layer of land.

![Figure 3. Resulting averaged models for each group that shows species distribution in a color palette, and corresponding thresholds as dotted line.](image)

The Final edition of the Buffer Zone was made taking the central section of the protected mammals’ group threshold polygon, trimming smaller parts along the line and dissecting at East and West where distance to the coastline was shorter and where the model values were lowest. The resulting Buffer Zone covered 132,673 ha, where cultivated areas occupied 42.4% and mosaics of natural and semi-natural vegetation covered 50.9%. In terms of territories, the buffer zone has
included 5 municipalities where Villa de Tututepec de Melchor Ocampo (63.4%) and Santiago Jamiltepec (32.6%) cover the most. The Transition Zone was delimited using the lowest thresholds resulting in an area of 221,402 ha (see Figure 4). The resulting Buffer Zone and Transition Zone can be visualized in GoogleEarth™ downloading and opening one the following buffer_zone.kmz files: http://tiny.cc/buffer_zone (skydrive) or http://tiny.cc/buffer_zones (googledocs).

Figure 4. Delimited zones for conservation of Chacahua National Park. It shows the Buffer Zone that includes the NPA and the surroundings zones that will work as a Transition Zone.

4 DISCUSSION AND CONCLUSIONS

The main objective of this study was to create a model to define the Buffer Zone and the Transition Zone as a tool to help decision makers and managers of natural protected areas to identify relevant regions where conservation strategies can be applied. To achieve this, a workflow was implemented based on MaxEnt and ArcGIS as modeling tools and free available data such as georeferenced species records, bioclimatic and land cover. By using the Equal training sensitivity and specificity logistic threshold, it was possible to delineate the minimum area to promote the conservation of species registered in the park and surroundings.

As remarked by Martin et al. [2009a] there are no clear references about how to delimit buffer zones. Although, previous delimitations have used social, political economical or physical characteristics, our suggested method uses geographical distribution of species as an ecological proxy to define this buffer as recommended by Ebregt et al. [2000].

Similar methods have been applied to identify conservation areas, but this method offers the possibility of using a semi-automatic delimitation of Buffer Zones integrating several tools in a procedure that facilitates repeatability and transferability, using free available data of several species and environmental information. There are several adjustments that can be made in the process, such as the selection of area of analysis, definition of raster calculation, selection of threshold type, the degree of generalization of threshold line and the output format. In this case, the objective was to define a minimum area that facilitate management and effectively promote conservation of species inside the park.
The presented analysis can face difficulties in regions with fast land use changes, like our study area. In such areas, a deep knowledge of species and their preferences to landscape features such as human settlements, infrastructure, field crops, freshwater bodies, and brackish water bodies, etc., need to be integrated to increase predictability validity of resulting models. Another important issue is the lack of georeferenced records, for this work public databases have provide several thousand records, but after the cleaning process, the usable records were as low as 0.3%. In this case, the georeferencing of records - using the locality description in web services as GEO Locate Web Client provided by Rios et al. [2012] - is a helpful tool to improve the number of georeferenced records. Although it needs a careful verification of suggested places, since some descriptions can incur in mlabel places or typos, leading to a wrong georeferencing coordinates.

Considering the results, the analyzed species have a potential distribution area of at least 10 times the size of Chacahuata Lagoon’s National Park. This implies that conservation efforts and management should be oriented also towards the surrounding areas to avoid severe impacts to the park’s fauna. Especially considering the pressure that communities exert over park’s resources, protection of adjacent areas can be even more relevant than the park itself as indicated by Martin et al. [2009b]. This situation is even more important considering that inadequate management of parks can induce dispersion of species as predators that will prey cattle as pointed out by García-Alaniz et al. [2010], or can damage crops as described by Rao et al. [2002] creating unnecessary conflicts.

Finally, alternative methods and datasets can also create a variety of delimitations, even though, the use of indicators such as species distribution can provide an adequate orientation about minimal areas needed to fauna preservation that are coherent with the conservation goals of NPA. In this regard, the present suggestion provides a biological criterion to design a Buffer Zone and supports the process for decision makers.

**ACKNOWLEDGMENTS**

This work was supported by a SEP-PROMEP grant to CHS and the Helmholtz Impulse and Networking Fund through Helmholtz Impulse and Networking Fund through Helmholtz Interdisciplinary Graduate School for Environmental Research (HIGRADE) Bissinger et al. [2008]. The paper has substantially benefitted from comments of Jörg Priess. Thanks to Jesús García Grajales and Alejandra Buenrostro Silva for the UMAR-CFE datasets as well as institutions that provides data through www.herpnet.org and www.manisnet.org.

**REFERENCES**

Bissinger, V. and O. Kolditz, Helmholtz Interdisciplinary Graduate School for Environmental Research (HIGRADE), GAIA, 17(1), 71-73, 2008.


