Title: Development of a Geographic Information System for Niche Model Analysis of Protected Species

Author: Roxanne Broadway

Abstract: The Center for Conservation Biology at the University of California, Riverside, in association with the California Department of Fish & Game and the US Fish and Wildlife Service, is developing a species monitoring protocol and niche model analysis for species listed under the CA and US Endangered Species Acts located in western Riverside County and the Coachella Valley as part of the multi-species habitat conservation plans for these two regions. This report describes the development of (1) the geodatabase that was designed for this project, and (2) a customized application that was used to simplify the multivariate analysis for quantification of habitat suitability.

Introduction: The Center for Conservation biology (CCB) is responsible for the management of the integrated database for the multi-species habitat conservation plans (MSHCP) for western Riverside County and the Coachella Valley, and, thus has gathered data on the spatial distribution and attributes for over 300 species stored in one or more dataset(s) (in Microsoft Excel datasheets). The data originates from three different sources: (1) *historical data* that dates back 150 years and ranges from records of occurrence with no specific location information to detailed collection and location data, (2) current *field data* (for mammals, birds, reptiles, amphibians, insects, and plants) collected by UCR staff and others, and (3) results from *habitat suitability models* (based on environmental factors including soil, elevation, hydrology, vegetation) that indicate the potential locations of species of interest. Thus, the data (and, in particular, the geospatial data) range from unmapable historic data to data that were directly georeferenced or located via modeling. For some of the "unmapable" historic data, the locations are being determined with the use of MaNIS (a georeferencing calculator developed by a consortium of 17 North American institutions to "assign geographic coordinates and maximum error distances for those coordinates to locality descriptions.")

The current problem faced by the users of the CCB datasets is that each dataset was developed to store data for a selected group of species (e.g., reptiles, amphibians, or birds), and each dataset has a unique design with no relationships built into that design that would allow linkage to other datasets. This results in a cumbersome and time consuming process to access, evaluate, analyze and visualize (via mapping) data in combination with other datasets. The CCB requested the development of a relational database for the entire data library used for their MSHCP project.

Given the number of researchers involved in this project and their wide ranging individual research goals, inconsistencies in data formats are inevitable, in

addition to inadvertent inconsistencies within data sets resulting from data entry errors. These inconsistencies had to be reconciled prior to integration into a GIS environment and error checking protocols had to be developed to guard against future data inconsistencies.

Geodatabase Development: Following extensive evaluation of the original datasets and meetings with the users, a personal geodatabase was developed for the entire data library. This geodatabase consists of 16 tables. Five tables focus on attributes of the species (scientific name and community, taxonomy, breeding period, growth habits (of plants), and listing status at the Federal & State and other levels). Three tables contain information about the observations of the species in the field (Observations, Artifacts, Comments). One table describes the site where the observations were made (Site), and six tables provide general information about names, locations, and directions to the observation sites (Point Locale, Point Directions, Sampling Unit ID, Sampling Unit Description, Sampling Unit Locale, Sampling Unit Directions). A final table provides cartographic information (i.e., standard labels) for observation sites (Map Label). Only one of these tables (Sites) contains spatial information. All other tables contain attributes needed to enable selection of appropriate subsets of the data required for analysis, mapping, and field monitoring.

Wherever possible, error checking/data validation was built into the geodatabase through the use of value-coded domains. All of the data within the geodatabase are in the same geographic coordinate system (NAD27) and projection (UTM), and the coordinate domain of the geodatabase is defined by a regional map of the counties in southern California.

Habitat Suitability Analysis: The pre-processing procedure used by the CCB to prepare data for habitat suitability analysis was a multi-stepped, labor-intensive, time-consuming process in ArcGIS that had the potential for the introduction of errors at each step. Following the multi-step process, the data was exported to SAS for analysis using the Mahalanobis distance metric, and then imported into ArcGIS for visualization and interpretation.

To resolve some of these issues, a spatial analysis tool, available as a free download for ArcView 3.x (and requires the spatial analyst extension) can be used to develop the habitat suitability models. The Mahalanobis Distances Extension is can be obtained from Jenness Enterprises via the internet at http://www.jennessent.com/arcview/mahalanobis.htm. Initial tests look promising; the results from the extension are comparable to those produced using the original analytical technique. The benefits of this approach are to (1) facilitate access to data for habitat suitability analysis, (2) eliminate the pre-processing procedure, (3) increase the speed of data access and analysis, and (4) improve the accuracy of analyses.

Author Information:

Roxanne Broadway, MS-GIS Candidate
MS GIS Program
University of Redlands
Redlands, CA
909 335-5128
Roxanne_Broadway@Redlands.edu