Development of a GIS Based Habitat Suitability Model for Southern California Rare Plants

Cristian Sarabia

Masters of Science in Geographic Information Science (MSGISci)

Department of Geography, California State University, Long Beach

Introduction

California's flora is unusually diverse and the much of the state is considered a hotspot of biodiversity with an estimated 6,300 indigenous vascular plants. In recent decades, the amount of flora that has been classified as endangered, threatened, or rare in California has increased, making California the state with the largest number of rare plants in the nation. Geographic information system (GIS) software facilitates the conservation and management of species of concern. By compiling different layers of geospatial information that is pertinent, a database for guidance can be accessed and used by the stakeholders.

The primary goal of this project was to create a model that could identify individual species habitats for several species of concern, including endangered, threatened and rare species in the area managed by the Palos Verdes Peninsula Land Conservancy (PVPLC). This project can then assist stakeholders by identifying potential areas with species of concern, can give guidance on how to appropriately expand the range of these species, and can also identify potential locations where these species may occur that have not yet been identified.

Figure 1. (A) Pictures of *Atriplex pacifica* and (B) *Aphanisma blitoides*

Data and Data Sources

Various forms of data were used in this project and acquired in different ways as seen in Table 1. These inputs were deemed to be the factors with the most influence on habitat suitability for the species included in this study.

Table 1. *List of data and data sources used in the project*

Dataset	Source
Atriplex pacifica point data	Palos Verdes Peninsula Land Conservancy
Aphanisma blitoides point data	Palos Verdes Peninsula Land Conservancy
Soil Survey Geographic Database	National Cooperative Soil Survey
Digital Elevation Model	Los Angeles Region Imagery Acquisition Consortium

Methodology

The set of points for each species was run through the *sample* tool in ArcMap together with the Soil ID classification, LARIAC_DEM, aspect, and slope data as inputs. It is important to note that the slope and aspect were derived from the DEM using the *slope* tool and *aspect* tool. Using the sample tool the soil class, elevation, slope and aspect where existing plants occur were extracted to an excel spreadsheet resulting in co-located values of these variables for the defined locations. Statistics were run for both species and graphs were generated (Figure 3) showing the percentage of field identified individuals of these species found under each variable. This information was used to develop a site suitability model to predict viable locations for these species.

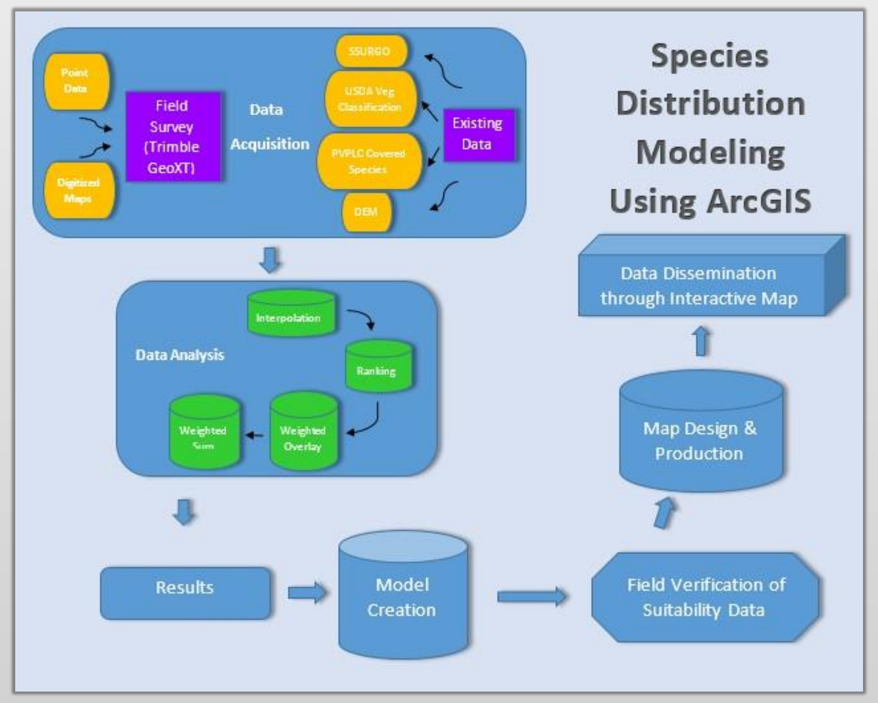

Figure 2. *Species Distribution Model Flowchart*

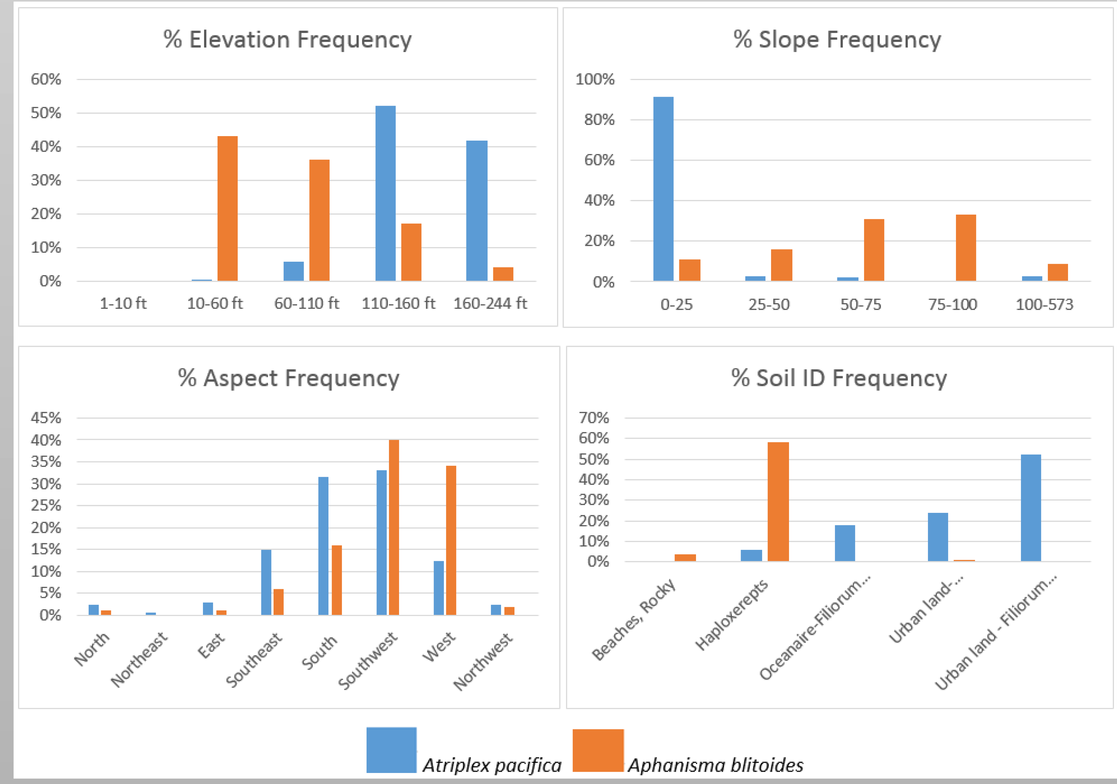

Figure 3. *Atriplex pacifica and Aphanisma blitoides statistics*

Timeline

Task	Proposed Project Timeline	Actual Project Completion
-Completed recommended process trainings	Week of May 8th	Week of May 8th
-Collected species locations and uploaded into ArcMap. -Calculate and analyze surface slope. -Build model and implement tools in ArcGIS.	Week of May 15th	Week of May 29th
-Check for errors in data and check data sites in the field.	Week of May 22nd	Throughout project
-Create final maps and interactive map and check for errors	Week of June 5th	Week of July 12th
-Duplicate Model in QGIS and implement tools	Week of June 5th	Not completed
-Field verify and complete all technical work.	Week of June 12th	Week of July 19th

Results

An ArcGIS ModelBuilder model was built to generate a habitat suitability raster for the species *Aphanisma blitoides* and *Atriplex pacifica*. Each species follows the model in Figure 3 using the same environmental inputs as factors for suitability. The elevation data is converted to slope and aspect, then reclassified and ranked according to statistical information based on each species. The soil data is converted from feature to raster and then all three inputs are put through a weighted overlay and a weighted sum process and finally processed through a raster calculator. The raster calculator generates the final raster with the areas of most likely occurrence for the species. The model uses six different weighting approaches that are integrated into the final result. This improves on the typical approach of using only one weighted overlay as it integrates a selection of equally viable realizations into the final result, generating a probability map of preferred habitat. This model can be reused with other species occurring in coastal locations by changing the rankings within each input reclassification. Other inputs can be added to the model as they become available or are assumed to affect the species.

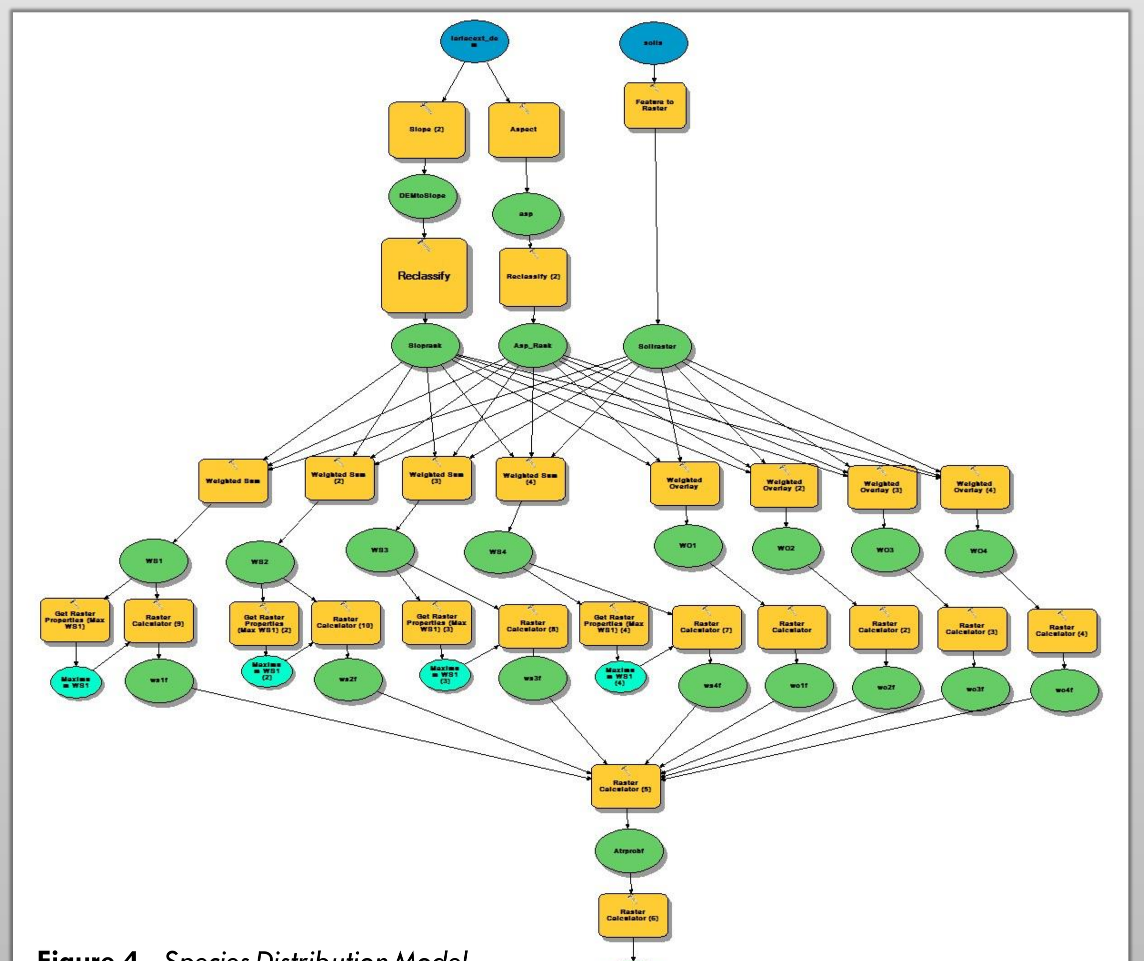

Figure 4. *Species Distribution Model*

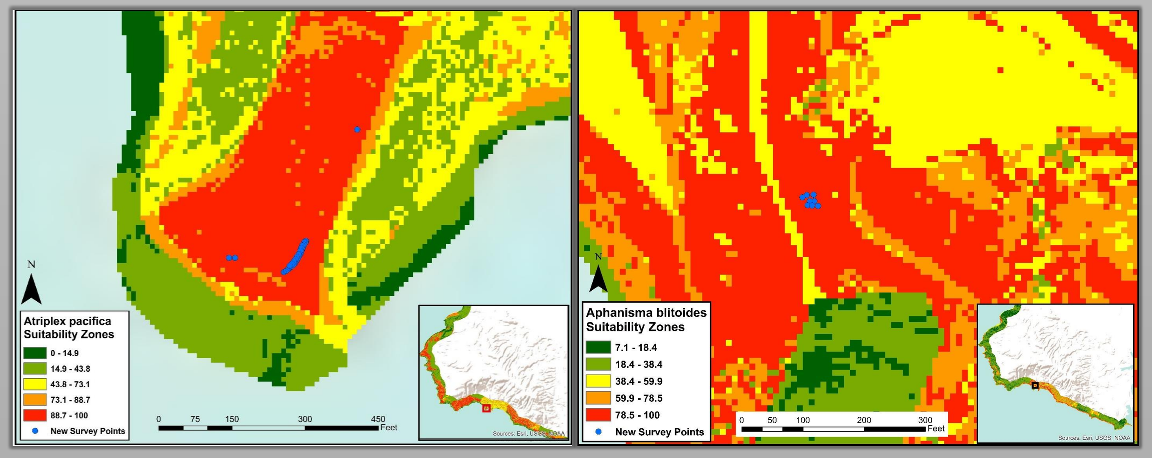

Figure 5. *Model Results*

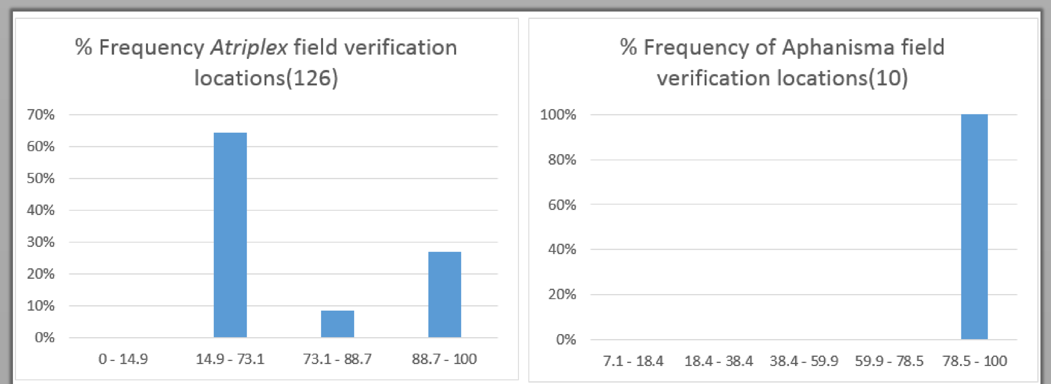

Figure 6. % Frequency of field verification points and model results

Discussion

The results produced by this project are immediately useful to anyone in charge of rare plant management in coastal southern California. The model I built can continue to be refined and newer and more accurate data can be added to it to get better results. The weighted sum/ weighted overlay approach within my model can continue to be changed and analyzed to see what other results can be obtained. There are many potential environmental factors that may be contributing to the distribution of the species analyzed but, I may not be considering them.

Figure 5 in the results section shows a sample of the model results and field verification locations. Figure 6 shows the field verification points collected after the model was run and the frequencies of occurrence within each "probability zone". With only 10 points collected for *Aphanisma blitoides* and 126 for *Atriplex pacifica*, and the minimal distribution throughout the surveyed area, the statistical validity of these results is somewhat limited. I do believe that the location identified by the model as most suitable hold a better chance of the species occurring within them. Due to years of experience with these species and the locations identified, these areas seem to accurately highlight suitable habitat.

Conclusion

Environmental inputs that can be added to the model include accurate vegetation data for my area of interest from the US Forest Service and updated soils data on a detailed scale. I began to work with the vegetation data and realized that some areas were misclassified. If I would have had these inputs, I could have removed developed areas and roads that may be skewing statistics. I continue to wonder if a better soil analysis would be helpful to improve results. A closer survey of the soils of these areas may reveal more diversity and create more micro-type conditions with niche habitats.

The next project to undertake is to try different species within the current model. The best species to analyze would be those that occur within the same extent as my current project. This would minimize the changes that would have to be made to my current model. Another project to undertake is to try the same species I worked on with my model but with different extents in the coastal southern California area. Some areas in mind include San Diego and Catalina Island which hold populations of *Atriplex pacifica* (Calflora). Additionally, creation of a Python script that can facilitate model use would be helpful to the public. Finally, implementing the distribution of the interactive web map to stakeholders would further make use of this model.

Submitted in partial fulfillment of the requirements of the Masters of Science in Geographic Information Science(MSGISci), August 12, 2017.

For additional information please contact: Cristian Sarabia; sporadicgrowth@gmail.com