IDENTIFICATION AND PRIORITIZATION
OF LANDS FOR RESTORATION OF
PIEDMONT PRAIRIE IN NORTH CAROLINA

by

Erica Marie Taecker

Masters project submitted in partial fulfillment of the requirements for the
Master of Forestry and Master of Environmental Management degrees
in the Nicholas School of the Environment and Earth Sciences,
Duke University

2007
ABSTRACT

In the central Piedmont of North Carolina, prairies and savannas were noted by European settlers to have covered a significant portion of the landscape. Piedmont prairie is valued for its extraordinary biodiversity; at least 277 plant species, some endemic, are associated with this unique area. Rich prairie ecosystems in the Piedmont were maintained by both naturally-occurring and human-ignited fires, which created open fields or patches of prairie within oak-pine-hickory or Piedmont longleaf pine forests. Anthropogenic changes to fire regimes and land use have fragmented the Piedmont prairie ecosystem, such that several of its plant species are now federally endangered. Effective conservation of this native ecosystem in our rapidly developing state depends on a solid understanding of its science. Just as importantly, it necessitates the ability for conservation agencies to act efficiently to protect and maintain areas of intact prairie, while quickly identifying and protecting other areas with restoration potential.

This masters project compares the suitability of two multivariate modeling tools, CART (Classification and Regression Tree) and Maxent (Maximum entropy), for predicting the potential geographic distribution of the Piedmont prairie ecosystem in nine Piedmont counties of North Carolina. Natural Heritage “Element Occurrence” point location data of four prairie species were the basis for the models, which considered environmental variables such as elevation, topographic relative moisture index (TRMI), slope, relative aspect, soil clay content, and soil effective cation exchange capacity (ECEC) in the prediction of potential prairie extent. Further, a basic prioritization of the resulting prairie “habitat” patches mapped in GIS highlights areas adjacent to existing protected areas in which to focus conservation and restoration efforts.

The results indicated that the habitat model of prairie created by Maxent reasonably predicts known prairie species occurrences without overgeneralizing the possible distribution of prairie in the study area. Maxent also highlights that ECEC is the most important predictor variable of prairie distribution, followed by clay content. The CART technique resulted in similar accuracy and explanatory variables, but when mapped, “habitat” covered a large proportion of the study area, less useful for targeting regions for further study. The preliminary prioritization suggests that several zones around Charlotte, NC and in Davidson County warrant further investigation for prairie remnants. With sufficient additional information about current land use and cover, the prioritization can be further refined to reduce the effort needed to find suitable sites for the restoration and conservation of Piedmont prairie and its associated forest cover types.
Table of Contents

Abstract ii

Introduction 1

Methods 4

Study Area 4

Figure 1: Study area 5

Data Preparation 5

Figure 2: Prairie element occurrence and random point locations 7
Figure 3: Data sources and variables used 8

Analyses 9

CART Model 9
Maximum Entropy Model 10
Prioritization 12

Results 14

CART Model 14

Figure 4: Complete CART model 14
Figure 5: Cross-validation figures 15
Figure 6: Pruned CART model 15
Figure 7: Mapped prairie distribution as predicted by CART 16

Maxent Model 17

Figure 8: Mapped prairie distribution probability predicted by Maxent 17
Figure 9: Mapped prairie prediction after setting threshold probability value 19

Prioritization 19

Figure 10: Prairie patches adjacent to or intersected by protected areas 20
Figure 11: Prairie patches impacted by roads and urban centers 21

Discussion 21

Conclusions 23
Introduction

In the central Piedmont of North Carolina, prairies and savannas were noted by European settlers to have covered a significant portion of the landscape (USFWS 1994). “We travell'd, this day, about 25 Miles, over pleasant Savanna Ground, high, and dry, having very few trees upon it, and those standing at a great distance. The Land was very good, and free from Grubs or Underwood,” according to John Lawson, one of several North Carolina explorers in the eighteenth century to note the open grassiness of the land (in Lefler, 1967). Remnant prairies and prairie species have been found in many states across the southeastern US, including South Carolina and the “Black Belt” of Mississippi and Alabama (Barone 2005). Piedmont prairie is valued for its extraordinary biodiversity; at least 277 plant species, some endemic, are associated with this unique area (Davis et al. 2002).

Rich prairie ecosystems in the Piedmont were maintained by both naturally-occurring and human-ignited fires, which created open fields or patches of prairie within dry oak or Piedmont longleaf pine forests (NatureServe, 2007). Many landscapes in the Piedmont are characterized by expansive upland “flats,” between firebreaks like streams, which would have made for relatively large natural “fire compartments” (FWS 1994). Prairie species are often found in association with open longleaf pine savannas, the pines like the prairie species requiring disturbance such as fire to reveal bare mineral soil seeds need to germinate. Prairies, like longleaf pine savanna, historically persisted through frequent fire and in some cases grazing animals, which reduced hardwood competition for light and resources during growth (1994). Post-oak blackjack-oak savannas, found in association with Piedmont prairie patches, also have a close relationship with regular fire (1994).
Anthropogenic changes to fire regimes and land use have fragmented the Piedmont prairie ecosystem, such that several of its plant species are now federally endangered (USFWS 1990). Effective conservation of this native ecosystem in our rapidly developing state depends on a solid understanding of its science. Just as importantly, it necessitates the ability for conservation agencies to act efficiently to protect and maintain areas of intact prairie, while quickly identifying and protecting other areas with solid restoration potential.

A database of “Element Occurrences” (EOs) of Piedmont prairie species and associated plants has been developed and maintained in a Geographic Information System (GIS) by the state Natural Heritage Program (NHP). These data are being used extensively in conservation and restoration planning for the Uwharrie mountain region of the central Piedmont, particularly by the inter-agency Greater Uwharrie Conservation Partnership (GUCP).

Considerable progress has been made in identifying, preserving, and improving the existing remnant North Carolina Piedmont prairie patches from the Uwharries to Charlotte, primarily based on the EO inventory. Successful restoration techniques for the Schweinitz’s sunflower, a prairie species, have included direct seeding in rocky microsites (Davis et al. 1999) and the use of prescribed fire at certain times of year (Barden 1994). GIS is also facilitating the prioritization of sites known to contain prairie species for conservation and restoration, as evidenced by the work of the NHP and WRC (Cook, unpublished, 2007). Historic accounts and maps indicating lands covered by prairie and savanna have also been useful in estimating their possible extent.

I am unaware of any project to identify and prioritize potential Piedmont prairie sites suitable for restoration in the NC Piedmont using GIS, specifically those sites for which an EO inventory has not taken place or in which no species of significance have been found at this time. Many prairie remnants are found in power line right-of-ways and along roadsides, locations with
a consistent pattern of disturbance, but this indicates little about the historic distribution of prairie and the fundamental environmental conditions allowing endurance of prairie species. The recovery plan for Schweinitz’s sunflower identifies several soil and terrain features that appear to be common among occurrences of this plant (1994), and others have pointed to a high correlation of prairie with fine soils with a hardpan component, such as the Iredell series (Horan 1995). However, no robust analysis of these variables seems to exist at this time, particularly for the Piedmont prairie ecosystem as a whole.

This project was undertaken to use multivariate and GIS tools to create an accurate and robust spatial habitat model of the geographic distribution of Piedmont prairie “habitat,” using general environmental variables such as soils, slope, and aspect. The habitat model and subsequent maps should provide a reasonable basis for prioritizing sites for future restoration, and at a basic level can help to identify where next to search for prairie EOs. Two main objectives were defined to guide and constrain this analysis:

1. Predict potential Piedmont prairie habitat distribution in the central Piedmont region of NC using environmental variables in a Classification and Regression Tree (“CART”) model and Maximum Entropy (“Maxent”) model, and select the most reasonable results as a basis for prioritization analyses.

2. Map predicted prairie habitat in GIS and use prioritization criteria to identify key potential habitat patches for future inventory, restoration, and conservation.
Methods

Analyses for this project occurred in two main parts: prediction of prairie habitat, and prioritization of potential prairie sites for feasibility of restoration. Prediction of the distribution of Piedmont prairie consisted of modeling habitat from environmental variables associated with EO point data using CART and Maxent procedures. Habitat modeling and prioritization steps utilized ESRI’s ArcGIS 9.1 for data analysis and spatial representation of results.

Study Area

The Piedmont Plateau physiographic region makes up much of central North Carolina, between the western Blue Ridge Mountains and the eastern Coastal Plain. The Piedmont ranges in elevation from around 2,000 feet near the Blue Ridge down to approximately 400 feet above sea level along the Coastal Plain (Langley, 2000). Natural plant communities in the region are diverse and include prairie, heath bluffs, oxbow lakes, and oak-hickory, longleaf pine, and alluvial forests (GUCP et al. 2007). While Piedmont prairie is thought to have historically extended over a large area of the southeastern United States, time, data, and processing limitations confined the study area to a subregion of the North Carolina Piedmont. Nevertheless, these analyses extend over 9 counties or approximately 4,747 square miles to the north and west of the Uwharrie Mountains in Montgomery County to Gaston County west of Charlotte, NC (Fig.1).
Figure 1: Study extent, 9 southern counties of the NC Piedmont

**Data Preparation**

The habitat models are based upon the most current set of Element Occurrence data as maintained in a point shapefile by the NC Natural Heritage Program (2007). EOs as polygons and lines are also available, but the point data is most useful in CART and Maxent analyses. EO data has its drawbacks and biases. Many EOs have been found and recorded as a function of accessibility rather than actual occurrence on a landscape, and are thus in road right-of-ways or on public or protected property. EO data also tends to be highly autocorrelated, since it is more likely for EOs to lie in close proximity to other EOs, again as a function of human access and seed dispersal distances. Finally, using EO data to extrapolate a particular habitat across a
landscape assumes that the EOs are good representatives of that habitat type, and also that the EOs exist in the full range of situations in which they would naturally be found.

Comprehensive vegetation sampling conducted by Davis et al. on remnant Piedmont prairie sites suggests that the following four federally listed species are indicative of current or historic prairie sites: *Echinacea laevigata* (Asteraceae), *Helianthus schweinitzii* (Asteraceae), *Lotus helleri* (Fabaceae), and *Symphiotrichum georgianum* (Asteraceae) (2002). While nearly 300 species are thought to also be representative of the Piedmont prairie ecosystem, the rarity of these four species is consistent with the relative scarcity of remnant prairie patches in the Piedmont. Thus, 232 EO points of these four species were used to develop the CART and Maxent habitat models (Fig. 2). The extent of analysis was determined largely by the geographic spread of these specific EOs. The original study area included the following counties: Montgomery, Randolph, Davidson, Stanly, Union, Anson, Mecklenburg, Rowan, and Gaston. Digitally mapped soil data was unfortunately unavailable for Montgomery County, so it was omitted from the model analyses and subsequent prioritization. While the GUCP is especially interested on protection in the Uwharries region largely contained by Montgomery County, this analysis also includes areas to the west, particularly around Charlotte, NC, where tremendous development pressures make prairie conservation concerns particularly notable.
Environmental data was all obtained in formats compatible with ESRI ArcGIS 9.x, from internet sources maintained by federal agencies (Figure 3). Raster layers of slope, aspect, elevation, and Topographic Relative Moisture Index (Parker 1980) were all created using ArcGIS tools (Appendix A.1) from the 1-arc-second (approximately 30-meter resolution) National Elevation Dataset digital elevation model (DEM) obtained from the US Geologic Survey (USGS) Seamless data viewer (2007). Also downloaded from the site were National Atlas features such as roads, streams, water bodies, cities, towns, and county boundaries, all of which were used to conduct the prioritization portion of the analysis. Soil series polygons and tabular data for counties in the area of interest were found at the Natural Resource Conservation Service (NRCS) Soil survey geographic database (SSURGO) online. According to the NRCS,
these are the “most detailed level of soil geographic data developed by the National Cooperative Soil Survey,” (2007) and are collected by scientists, though they are, like any data, subject to the scale and accuracy at which they were sampled and digitized, and there is some variability from county to county in soil classification.

<table>
<thead>
<tr>
<th>Data Source</th>
<th>Layers / (Variables) Derived</th>
</tr>
</thead>
<tbody>
<tr>
<td>USGS Seamless server</td>
<td>Roads, Cities, Counties, State boundary</td>
</tr>
<tr>
<td></td>
<td>Digital Elevation Model (30-m raster):</td>
</tr>
<tr>
<td></td>
<td>(Elevation)</td>
</tr>
<tr>
<td></td>
<td>(Slope)</td>
</tr>
<tr>
<td></td>
<td>(Relative Aspect)</td>
</tr>
<tr>
<td></td>
<td>(Topographic Relative Moisture Index)</td>
</tr>
<tr>
<td>NRCS SSURGO</td>
<td>Soil survey digital maps for 9 counties</td>
</tr>
<tr>
<td></td>
<td>Soils relational databases for 9 counties</td>
</tr>
<tr>
<td></td>
<td>(Clay content)</td>
</tr>
<tr>
<td></td>
<td>(Effective cation exchange capacity)</td>
</tr>
<tr>
<td>NC Natural Heritage Program</td>
<td>Element Occurrences</td>
</tr>
<tr>
<td></td>
<td>Natural Heritage areas</td>
</tr>
<tr>
<td></td>
<td>Managed (protected) areas</td>
</tr>
</tbody>
</table>

**Figure 3:** Data sources and variables used

After creating the environmental variables from the DEM, all were sampled at the location of the EO points using the Sample tool in the ArcGIS Spatial Analyst extension (Appendix A.1). In addition, a random set of approximately the same number of points (176) was created using the “Create Random Points” tool, and sampled through all variables. The environmental variables of slope, aspect, elevation, TRMI, and soil series were initially used because they are relatively static in space and time and are not frequently altered by humans. Dynamic or anthropogenic variables such as “land cover,” “time since fire/disturbance,” and “distance from roads” tend to correlate with the existence of current prairie EOs, and may have confounded the ability of the model to identify long-term, underlying natural patterns that can help explain the existence of prairie ecosystems historically. Soil data was sampled for effective cation exchange capacity (ECEC) and clay percentage rather than soil type. ECEC, which is correlated with soil pH and
parent material, primarily affects the nutrient availability to plants, while clay content directly influences the water available to plant roots (Brady and Weil 2001). Both variables are described in standard terms on continuous scales in all county soil surveys, whereas soil naming conventions vary drastically between counties and create too many categories for effective modeling in CART or Maxent. All environmental data for EO and Random points were converted into comma-separated tables to facilitate analysis in statistical programs.

**Analyses**

**CART Model**

Classification and Regression Tree (CART) analysis is a relatively new and increasingly common multivariate analysis tool for classifying habitats or vegetation types with respect to known environmental variable values (Breiman et al. 1984). For the purposes of habitat modeling where the response variables are categorical (in this case “Prairie EO” or “Random”) rather than continuous, CART is typically performed as a classification, rather than regression tree. CART assumes no particular distribution of the data, and the algorithm iteratively evaluates each variable for its ability to partition the samples into pure groups (2002). The first split thus partitions the samples into two groups as purely homogenous as possible based on the variable that best accounts for deviance in the entire sample set. The branch to the right contains all samples for which the variable value or range of values is “true,” and the branch to the left is all “false” samples. At each subsequent split, or node, the deviance within the group is reduced according to whichever variable best explains that deviance, and the samples are divided again. The resulting complex tree fully predicts, or over-predicts every sample in the set. The tree is then simplified by pruning it to fewer terminal nodes, ideally to a level in which the
misclassification rate of known “habitat” samples is reasonably low, but the classification efficacy is reasonably high for new samples.

The sampled data (TRMI, elevation, slope, relative aspect, percent clay, ECEC) was read into R statistical software (R Development Core Team 2005), and CART analysis scripts were run to determine what combination of the above variables best predicts potential “Piedmont prairie” (Appendix A.2). The tree was pruned according to the results of cross-validation of the model. In cross validation, 90% of the samples were used to create the tree, and the remaining 10% of the samples were used to test the model. After repeating this cross-validation procedure with different random subsets of data, an estimate of model accuracy was obtained. The tree was pruned of any branches that contributed relatively little to the overall misclassification rate of the model. The remaining branches and their corresponding variable combinations leading to “prairie” were input into ArcGIS as a logical statement (Appendix A.3), and a raster showing cover of predicted Piedmont prairie was mapped.

**Maximum Entropy Model**

Maximum entropy, a well-known concept in machine learning and other fields, has only recently been applied to ecological modeling (Jaynes, 1957, Phillips, 2006). Maxent does not require absence data for samples, and assumes no particular distribution of the data. The principle of maximum entropy places no constraints on the probability distribution of the species or habitat, only assuming that each feature (environmental variable) has the same mean in the approximated distribution as in the empirical observations, in this case element occurrences, of the habitat. In order to do this, Maxent iteratively updates an algorithm that describes the habitat distribution based on computed weights of all the features considered.
All variable layers were prepared for use in Maxent (Phillips 2004), which requires a consistent extent and cell size for proper running of the model. It is important to note that the modification of some layers, particularly clay content and ECEC derived from a finer-scale polygon shapefile, to a lower-resolution raster potentially impacts the accuracy of the model. Set up specifications used for Maxent can be found in Appendix A.4.

Accuracy measures generated by the Maxent program were summarized and compared to those of CART. The resulting layer of predicted probability of habitat from Maxent was imported into ArcGIS, where the continuous cumulative probability map was converted into a discrete “Habitat” vs. “Non-habitat” raster map, and background “No Data” cells were reclassified to “Non-habitat” to create a solid raster image with no empty cells.

In order to determine the threshold probability for discrete “habitat,” an ROC (Receiver Operating Characteristic) curve generated by Maxent was used. The ROC curve indicates the relationship between the specificity, which is the area truly predicted as non-habitat (true negatives), and sensitivity, the area correctly predicted as habitat (true positives) of the model. All map cells with probabilities greater than or equal to that which balanced specificity with sensitivity were mapped as “habitat.” While ROC curves are not typically associated with presence-only modeling, here the analogue for absence is “random” or “background” points (Phillips et al. 2006). The AUC (Area Under Curve) was thus developed using background data instead of absence data, in which case the maximum AUC could only be 0.757 instead of 1.0 for this model (Appendix B.1). An AUC of 1 would mean that the model perfectly discriminates between habitat and non-habitat, while an AUC of 0.5 means the model will predict accurately only 50% of the time, essentially at random (Metz 1978).
Pixelation and was reduced by smoothing the habitat raster using GIS tools. Finally, the entire grid was re-sampled to a larger cell size, which further removed isolated single-pixel patches, and allowed for a more generalized picture of potential habitat patches to facilitate the prioritization analysis.

**Prioritization**

The habitat model with the lowest proportional predicted area which optimized specificity and sensitivity was chosen for further prioritization analyses. Recent layers depicting land cover and use, either from the National Land Cover Dataset (NLCD) or the NC GAP (McKerrow et al. 2006) project, were sought for evaluating existing conditions in areas identified as potential habitat, because patches should be eliminated from consideration if the current land use and cover deviates substantially and irreparably from the desired ecological condition, particularly as a result of development. Even “natural” land covers and uses may be deemed unfeasible for conservation management if they are inaccessible, or if the cost of restoration would be prohibitive. For example, a large patch of potential prairie might be comprised of natural vegetation well suited to restoration to prairie and pine savanna through frequent prescribed fire, but because it is surrounded by urban development, the land cost and difficulty in applying fire to it on a regular basis will make its restoration and purchase unlikely. Neither the most recent 2006 NLCD (2006) nor GAP maps were available for public use in the project study area at the time the analysis was performed. Older land cover maps would have been of little use given the rapid growth in many parts of the study area. The extent of the model also made it unreasonable to download memory-intensive USGS Digital Ortho Quarter Quads (DOQQ), which also would indicate areas inappropriate for restoration or investigation for element occurrences.
Proximity of predicted prairie patches to protected lands was quickly assessed using Natural Heritage shapefiles of managed and protected areas in the region of interest, since many agencies focus conservation efforts on habitat patches adjacent to lands already being managed or protected for such purposes. Patches neighboring protected areas may have similar and relatively intact vegetation and open space, and there may be owners amenable to more protection of natural areas and native ecosystems. Ownership of “protected” lands includes land trusts, state wildlife land, private partners interested in conservation, and the US Forest Service. The raster of Maxent-predicted habitat patches was converted to a polygon shapefile, which could then be evaluated for its intersection and adjacency with protected areas.

A general “burnability” assessment was also considered, in which patches in closer proximity to major highways and urban centers, assumed to have low feasibility of fire restoration, were given a lower priority. This part of the prioritization will be particularly useful after comparing patches to current land cover and after initial field investigation to further evaluate the suitability of patches for fire restoration. It must be noted, however, that prairie restoration methods besides prescribed fire warrant consideration in settings where particularly strong remnant populations of species are found and the impetus for protection outweighs the challenges of restoration.

Collaboration among land management organizations makes large-scale restoration a possibility, particularly when the suitable or potentially suitable habitat to be restored is consolidated into bigger patches. From an ecological perspective, larger areas of “habitat” are better than smaller areas, particularly if they are well-connected, and so the largest patches apparently minimally influenced by urban challenges were also highlighted in the prioritization.

Finally, patches were distinguished according to the presence of known prairie element occurrences, which may be current or historic. Organizations may choose to concentrate efforts
on areas containing prairie elements because of their apparent ease of access or protection status. On the other hand, long-term goals may call for conducting species searches and formulating conservation strategies where no known populations occur.

**Results**

*CART Model*

The complete CART model contained twenty terminal nodes, with an overall misclassification rate of 22.1% (Fig. 4). Using cross-validation, it was determined that misclassification error was mostly minimized after the first five splits, resulting in 6 terminal nodes (Fig. 5).

![Figure 4: Complete CART for prediction of prairie EOs.](image)
The pruned tree revealed the important predictor variables to be ECEC, Relative Aspect, Clay content, Elevation, and Slope (Fig. 6).
Mapped in GIS, the proportional predicted prairie area from the CART model was 0.87 of the total study area, or 4,129 square miles (Fig. 7), which accurately captures most of the known EO points but still misclassifies 26.8% of them as non-habitat (Fig. 8). Predicted habitat covers most of Union county and appears most diffuse in neighboring Anson County. In general, the strongest region of “habitat” according to the CART model lies in a southwest to northeast swath, consistent with the element occurrences used to develop the model.
**Maxent Model**

Maxent created a probabilistic predicted habitat image in which warmer colors indicate more suitable conditions for habitat and cooler colors suggest poorer conditions (Fig. 8).

![Maxent Model](image.png)

**Figure 8**: Habitat distribution predicted from Maxent and projected in GIS; warmer colors indicate higher probability of suitable habitat.

The jackknife assessment excludes one variable at a time and recreates the model to determine the relative contribution of the predictor variable to the habitat distribution (Appendix B.2). It also takes each variable separately to determine whether it alone contributes significantly to the model. As in the CART model, ECEC was found to be the most important predictor of prairie habitat on its own with a 51.2% contribution, followed by Clay with a 34% contribution.
The remaining variables only marginally contributed to refining the Maxent model. Response curves indicate the influence of the variables on the Maxent prediction, notably that prairie vegetation is best supported in soils with an ECEC greater than or equal to seven, and a clay fraction between thirteen and twenty-one percent (Appendix B.3).

The threshold chosen to define the cutoff between “habitat” and “non-habitat” for the Maxent probability model equalized sensitivity and specificity. The cumulative threshold was 35.721, which is the same as the logistic threshold of 0.407. At this level, the omission rate is 0.262, close to the same as the proportional predicted area of 0.272. A total of 1,293 square miles, in 5,907 distinct patches, were predicted as potential “prairie habitat” using this Maxent threshold.

Mapped in GIS at this threshold, predicted prairie habitat is concentrated to several regions similar to the CART map but more defined in Maxent, namely in Mecklenburg, Union, and Gaston counties in the southwestern extent of the study area (Fig. 9). Other areas of note include portions of Rowan County, and eastern Davidson County. Interestingly, the misclassified element occurrences demonstrate no particular spatial patterns, occurring in every county of the study area, often in close proximity to areas predicted as prairie.
Figure 9: Prairie distribution as predicted from Maxent cumulative threshold equalizing specificity and sensitivity; element occurrences misclassified under this scheme appear to demonstrate no particular spatial pattern.

Prioritization

A total of 284 prairie habitat patches derived from Maxent were found to be adjacent to or intersected by currently protected areas (Fig. 10). Most large patches of potential habitat from this analysis fell in Mecklenburg and Union counties, while smaller patches were scattered across the study area.

Of all patches modeled in Maxent, only 39 actually contain known prairie element occurrences, and 20 of those are also adjacent to or inside protected areas (Fig. 10).
Over 300 total patches are intersected by major roads, 41 of which are also on or adjacent to protected lands (Fig. 11). Many patches (1,274) lie within five miles of a major town center, of which 22 patches are on or adjacent to protected lands. Most adjacent patches do not have any of the above human conflicts to fire management, so the top fifteen largest of those remaining patches were highlighted as priorities for investigation (Fig. 11). Despite the high concentration of potential prairie in Mecklenburg County, the largest patches free of city or road challenges are found in Union County. Not emphasized by these methods was the relatively high density of potential prairie patches in northwest Rowan County, which may nevertheless be of significant conservation value.
Discussion

Both CART and Maxent methods created plausible depictions of the possible distribution of Piedmont prairie in the study area, and their results overlapped significantly. However, for the purposes of conservation planning, Maxent results gave a more refined and tractable estimate of prairie habitat. While inconsistencies in soil classification among counties are visible to some degree in the results of both models, soil factors appear to be an undeniable predictor of prairie...
distribution. The importance of ECEC and Clay content in Maxent and CART predictions of prairie is consistent with other research and historic anecdotes.

While every single predicted habitat patch cannot be ground-truthed for its prairie “potential” and “restorability,” the Maxent model can be improved upon as more prairie EOs are found. This type of analysis is perfect for non-absence data sets like this one, since the absence of prairie species in a location does not eliminate the possibility of their existence in the past or future at the same location.

Refining the prioritization can be facilitated with current observations of land use and cover, particularly from the NLCD. Parcel ownership records, accessible for most counties online, can also indicate the current condition of potential habitat patches in some cases. Prioritizing the investigation of possible prairie remnants based on fire probabilities similar to Susan Langley and Cecil Frost’s “fire compartment” concept was not done here, but could assist with narrowing the selection of plausible patches (Langley 2000). From an efficiency standpoint, merging the prairie distribution map with the mapped distribution of other associated plant communities best managed through fire would facilitate prioritization. Complementary projects worth investigating are a longleaf pine distribution as modeled using a Generalized Linear Model (unpublished, USFS) and the NC GAP assessment, which has recently become available and does not use a prairie classification.

The Maxent prediction map of prairie hints that more analysis is needed, particularly in the edge zones showing a higher suitability for Piedmont prairie. Because Montgomery County, where many occurrences of prairie species have been recorded, was left out of this analysis, the distribution of predicted prairie in Uwharrie National Forest and environs remains unknown. That data, in addition to the numerous occurrences of prairie species in South Carolina, could all
be used to improve the model if and when soils data comes online. The Uwharrie area might take priority for the purposes GUCP planning, although model results beyond the Uwharries highlight the need for continued collaboration among regional conservation planning groups.

Conclusions

Conservation organizations increasingly rely upon geospatial tools to define and prioritize sites for protection and restoration activities. Prioritization strategies frequently focus on parcels or landscapes in which conditions are already known from site visits or natural heritage records. Integrating quantitative and testable habitat models into conservation planning and prioritization can truly expand the scope of possibilities for long-term and landscape-scale management of threatened ecosystems and species. This research demonstrates that the Maxent method of habitat modeling gives an estimate of Piedmont prairie distribution in North Carolina that is consistent with historic and scientific species records. Prioritizations based on the Maxent model reinforce the substantial urban threats to Piedmont prairie, supporting the need for continued efforts to conserve and restore this rare ecosystem.
Acknowledgments

The inspiration for and completion of this Masters’ Project could not have been possible without the support and guidance of several people:

I must thank Dr. Dean Urban (Duke University) for sharing his endless store of seemingly innate knowledge of multivariate and geospatial analysis methods. I could not have even begun to think about conducting such research without his courses and guidance.

I am also indebted to my mentor and internship supervisor Will McDow (Environmental Defense) for offering me the Stanback Conservation Internship and for sparking my interest in Uwharrie ecosystems, which directly led to this project. I also thank him for allowing me the freedom and creativity to formulate and re-formulate my objectives and methods as I saw fit, and always trusting in my ability to do what I set out to do.

Without the Stanbacks, my internship would not have been possible; I appreciate their support immensely.

After a hiatus in working on my MP, I owe my renewed enthusiasm in this project to Kiersten (Kacy) Cook (NC WRC), who provided data and optimism. Other members of the Greater Uwharrie Conservation Partnership and the NC Natural Heritage Program are also greatly appreciated for their interest and encouragement.

Doug Sprouse deserves special gratitude for his endless patience in my ongoing academic and professional pursuits, despite the stresses, and for his eager interest in whatever I’m working on and learning. I would not have come this far (with any sanity intact) without his support.

Finally, I thank my family, Doug’s family, and doggies Biscuit and Sissy, for their unconditional love and confidence in my endeavors.
Literature


Greater Uwharries Conservation Partnership et al. (Synthesized by Pohlman, S. and Marcus, J.) 2007. Conservation targets and rankings for the Greater Uwharries Region.


NC Natural Heritage Program. 2006. marea.shp and snha.shp (State Natural Heritage areas and managed areas data files in GIS shapefile format).


Appendices

A.1: Python scripts for data prep in GIS:

# TRMI.py (to create TRMI from DEM)
# Created on: Mon Dec 03 2007 10:21:48 AM
# (generated by ArcGIS/ModelBuilder)

# Import system modules
import sys, string, os, arcgisscripting

# Create the Geoprocessor object
gp = arcgisscripting.create()

# Check out any necessary licenses
gp.CheckOutExtension("spatial")

# Load required toolboxes...
gp.AddToolbox("C:/Program Files/ArcGIS/ArcToolbox/Toolboxes/Spatial Analyst Tools.tbx")

# Set the Geoprocessing environment...
gp.XYResolution = ""
gp.scratchWorkspace = "z:\ED\GIS_DATA"
gp.MTolerance = ""
gp.randomGenerator = "0 ACM599"
gp.outputCoordinateSystem = ""
gp.outputZFlag = "Same As Input"
gp.qualifiedFieldNames = "true"
gp.extent = "+81.178055520587 34.8058333322809 -79.5302674127183 36.0569444434918"
gp.XYTolerance = ""
gp.cellSize = "0.000229"
gp.outputZValue = ""
gp.outputMFlag = "Same As Input"
gp.geographicTransformations = ""
gp.ZResolution = ""
gp.mask = "dem_merge"
gp.workspace = "z:\ED\GIS_DATA"
gp.MResolution = ""
gp.ZTolerance = ""

# Local variables...
dem_merge = "dem_merge"
dem_fill = "z:\ED\GIS_DATA\dem_fill"
aspect = "z:\ED\GIS_DATA\aspect"
flow_dir = "z:\ED\GIS_DATA\flow_dir"
flow_drop = "z:\ED\GIS_DATA\flow_drop"
flow_acc = "z:\ED\GIS_DATA\flow_acc"
flowpath = "z:\ED\GIS_DATA\flowpath"
flowpath_null = "z:\ED\GIS_DATA\flowpath_null"
thin_flow = "z:\ED\GIS_DATA\thin_flow"
flowpath_dir = "z:\\ED\\GIS_DATA\\flowpath_dir"
flow_down = "z:\\ED\\GIS_DATA\\flow_down"
mean_elev = "z:\\ED\\GIS_DATA\\mean_elev"
elev_differ = "z:\\ED\\GIS_DATA\\elev_differ"
hill_tops = "z:\\ED\\GIS_DATA\\hill_tops"
Input_false_raster_or_constant_value = "1"
thin_hill = "z:\\ED\\GIS_DATA\\thin_hill"
ridgetops_dir = "z:\\ED\\GIS_DATA\\ridgetops_dir"
flow_up = "z:\\ED\\GIS_DATA\\flow_up"
slope__2_ = "z:\\ED\\GIS_DATA\\slope"
slope_recl = "z:\\ED\\GIS_DATA\\slope_recl"
asp_recl = "z:\\ED\\GIS_DATA\\asp_recl"
rsp_100 = "z:\\ED\\GIS_DATA\\rsp_100"
rsp_reclass = "z:\\ED\\GIS_DATA\\rsp_reclass"
rsp_reclass1 = "z:\\ED\\GIS_DATA\\rsp_reclass1"
rsp_reclass2 = "z:\\ED\\GIS_DATA\\rsp_reclass2"
trmi = "z:\\ED\\GIS_DATA\\trmi"

# Process: Fill...
gp.Fill_sa(dem_merge, dem_fill, "")

# Process: Flow Direction...
gp.FlowDirection_sa(dem_fill, flow_dir, "NORMAL", flow_drop)

# Process: Single Output Map Algebra (2)...
gp.SingleOutputMapAlgebra_sa("int(aspect(dem_fill))", aspect, "z:\\ED\\GIS_DATA\\dem_fill")

# Process: Reclassify (2)...
gp.Reclassify_sa(aspect, "VALUE", "-1 0;0 17;1 9 18;10 18 19;19 26 20;27 35 19;36 44 18;45 53 17;54 62 16;63 71 15;72 80 14;81 89 13;90 98 12;99 107 11;108" 116 10;117 125 9;126 134 8;135 143 7;144 152 6;153 161 5;162 170 4;171 179 3;180 188 2;189 197 1;198 207 0;208 216 1;217 225 2;226 234 3;235 243 4;244 252 5;253 261 6;262 270 7;271 279 8;280 288 9;289 297 10;298 306 11;307 315 12;316 324 13;325 333 14;334 342 15;343 352 16;352 360 17", asp_recl, "NODATA")

# Process: Slope...
gp.Slope_sa(dem_fill, slope__2_, "DEGREE", "9.00093515197289E-06")

# Process: Reclassify...  

# Process: Flow Accumulation...  
gp.FlowAccumulation_sa(flow_dir, flow_acc, "", "FLOAT")

# Process: Single Output Map Algebra (3)...
gp.SingleOutputMapAlgebra.sa("flow_acc > 25", flowpath, "z:\\ED\\GIS_DATA\\flow_acc")
# Process: Set Null...
gp.SetNull_sa(flowpath, flowpath, flowpath_null, "Value = 0")

# Process: Thin...
gp.Thin_sa(flowpath_null, thin_flow, "ZERO", "NO_FILTER", "ROUND", "2.287641540354E-03")

# Process: Focal Statistics...
gp.FocalStatistics_sa(dem_fill, mean_elev, "Rectangle 10 10 CELL", "MEAN", "DATA")

# Process: Single Output Map Algebra (4)...
gp.SingleOutputMapAlgebra_sa("mean_elev - dem_fill", elev_differ, "z:\ED\GIS_DATA\mean_elev;z:\ED\GIS_DATA\dem_fill")

# Process: Set Null (2)...
gp.SetNull_sa(elev_differ, Input_false_raster_or_constant_value, hill_tops, "Value > -40.0")

# Process: Thin (2)...
gp.Thin_sa(hill_tops, thin_hill, "NODATA", "NO_FILTER", "ROUND", "2.287641540354E-03")

# Process: Con (2)...
gp.Con_sa(thin_hill, flow_dir, ridgetops_dir, ",", "Value < 1")

# Process: Flow Length (2)...
gp.FlowLength_sa(ridgetops_dir, flow_up, "UPSTREAM", ",")

# Process: Con...
gp.Con_sa(thin_flow, flow_dir, flowpath_dir, ",", "Value < 1")

# Process: Flow Length...
gp.FlowLength_sa(flowpath_dir, flow_down, "DOWNSTREAM", ",")

# Process: Single Output Map Algebra...
gp.SingleOutputMapAlgebra_sa("(flow_down / (flow_up + flow_down)) * 100", rsp_100, "z:\\ED\\GIS_DATA\\flow_up;z:\\ED\\GIS_DATA\\flow_down")

# Process: Reclassify (3)...
gp.Reclassify_sa(rsp_100, "Value", "0 20 20 40 40 60 60 10;60 80 5;80 100 0", rsp_reclass, "DATA")

# Process: Single Output Map Algebra (5)...
gp.SingleOutputMapAlgebra_sa("con(thin_hill == 1, 1, rsp_reclass)", rsp_reclass1, "z:\\ED\\GIS_DATA\\rsp_reclass;z:\\ED\\GIS_DATA\\thin_hill")

# Process: Single Output Map Algebra (6)...
gp.SingleOutputMapAlgebra_sa("con (thin_flow==1, 20, rsp_reclass1)", rsp_reclass2, "z:\\ED\\GIS_DATA\\thin_flow;z:\\ED\\GIS_DATA\\rsp_reclass1")

# Process: Single Output Map Algebra (7)...
gp.SingleOutputMapAlgebra_sa("asp_recl + slope_recl + rsp_reclass2", trmi, "z:\\ED\\GIS_DATA\\asp_recl;z:\\ED\\GIS_DATA\\slope_recl;z:\\ED\\GIS_DATA\\rsp_reclass2")
# sampling.py (to sample through EO and random points for env variables)
# Created on: Mon Dec 03 2007 10:21:24 AM
# (generated by ArcGIS/ModelBuilder)
# ---------------------------------------------------------------------------

# Import system modules
import sys, string, os, arcgisscripting

# Create the Geoprocessor object
gp = arcgisscripting.create()

# Check out any necessary licenses
gp.CheckOutExtension("spatial")

# Load required toolboxes...
gp.AddToolbox("C:/Program Files/ArcGIS/ArcToolbox/Toolboxes/Spatial Analyst Tools.tbx")
gp.AddToolbox("C:/Program Files/ArcGIS/ArcToolbox/Toolboxes/Data Management Tools.tbx")
gp.AddToolbox("C:/Program Files/ArcGIS/ArcToolbox/Toolboxes/Analysis Tools.tbx")

# Set the Geoprocessing environment...
gp.XYResolution = ""
gp.scratchWorkspace = "z:\ED\GIS_DATA"
gp.MTolerance = ""
gp.randomGenerator = "0 ACM599"
gp.outputCoordinateSystem = ""
gp.outputZFlag = "Same As Input"
gp.qualifiedFieldNames = "true"
gp.extent = "-81.1780555520587 34.8058333322809 -79.5302674127183 36.0569444434918"
gp.XYTolerance = ""
gp.cellSize = "0.000187"
gp.outputZValue = ""
gp.outputMFlag = "Same As Input"
gp.geographicTransformations = ""
gp.ZResolution = ""
gp.mask = ""
gp.workspace = "z:\ED\GIS_DATA"
gp.MResolution = ""
gp.ZTolerance = ""

# Local variables...
samp_rand = "z:\ED\GIS_DATA\samp_rand"
samp_eos = "z:\ED\GIS_DATA\samp_eos"
trmi = "trmi"
slope = "slope"
aspect = "aspect"
dem_fill = "dem_fill"
eo_proj = "eo_proj"
rel_asp = "z:\ED\GIS_DATA\rel_asp"
rnd_6_cl_proj = "rnd_6_cl_proj"
samp_eos_pt = "samp_eos_pt"
samp_rand_pt = "samp_rand_pt"
random_soil_shp = "z:\ED\GIS_DATA\random_soil.shp"
eos_soil_shp = "z:\ED\GIS_DATA\eos_soil.shp"
eo_pt_dat = "eo_pt_dat"
rand_pt_dat = "rand_pt_dat"
all_soils_proj = "all_soils_proj"

# Process: Single Output Map Algebra...
gp.SingleOutputMapAlgebra_sa("con(aspect > 0, (180 - ABS(aspect - 180)))",
rel_asp, "aspect")

# Process: Sample (2)...
gp.Sample_sa("trmi;slope;dem_fill;z:\ED\GIS_DATA\rel_asp", eo_proj,
samp_eos, "NEAREST")

# Process: Make XY Event Layer...
gp.MakeXYEventLayer_management(samp_eos, "x", "y", samp_eos_pt, "")

# Process: Sample...
gp.Sample_sa("trmi;slope;dem_fill;z:\ED\GIS_DATA\rel_asp", rand_6_cl_proj,
samp_rand, "NEAREST")

# Process: Make XY Event Layer (2)...
gp.MakeXYEventLayer_management(samp_rand, "x", "y", samp_rand_pt, "")

# Process: Intersect...
gp.Intersect_analysis("all_soils_proj #;rand_pt_dat #", random_soil_shp,
"ALL", "", "POINT")

# Process: Intersect (2)...
gp.Intersect_analysis("all_soils_proj #;eo_pt_dat #", eos_soil_shp, "ALL",
"", "POINT")

A.2: CART scripts and results (not figures) from R

```r
#read prairie env data
env.data<-read.csv("Z:/ED/GIS_DATA/data_table.csv",header=T)
env.data
spec<-env.data$SAMPLE
summary(as.factor(spec))
cart.data<-env.data[,,-1]
cart.data

library(tree)
spec.tree<-tree(as.factor(spec)~.,data=cart.data)
plot(spec.tree)
text(spec.tree,cex=0.6)
summary(spec.tree)
Classification tree:
tree(formula = as.factor(spec) ~ ., data = cart.data)
Number of terminal nodes: 20
```
Residual mean deviance: 0.913 = 353 / 387
Misclassification error rate: 0.221 = 90 / 407

> print(spec.tree)

node), split, n, deviance, yval, (yprob)
* denotes terminal node

1) root 407 560.0 eo (0.570 0.430)
   2) ECEC < 4.5 129 160.0 rand (0.318 0.682)
      4) REL_ASP < 97 70 64.0 rand (0.171 0.829)
         8) ELEV < 186.5 20 0.0 rand (0.000 1.000) *
         9) ELEV > 186.5 50 55.0 rand (0.240 0.760)
            18) CLAY < 13.5 11 0.0 rand (0.000 1.000) *
            19) CLAY > 13.5 39 48.0 rand (0.308 0.692) *
      5) REL_ASP > 97 59 82.0 rand (0.492 0.508)
         10) SLOPE < 3.195 26 32.0 rand (0.308 0.692)
            20) CLAY < 13.5 7 0.0 rand (0.000 1.000) *
            21) CLAY > 13.5 19 26.0 rand (0.421 0.579) *
         11) SLOPE > 3.195 33 43.0 eo (0.636 0.364)
            22) REL_ASP < 128.5 11 6.7 eo (0.909 0.091) *
            23) REL_ASP > 128.5 22 30.0 eo (0.500 0.500)
               46) ELEV < 194.5 5 0.0 rand (0.000 1.000) *
               47) ELEV > 194.5 17 22.0 eo (0.647 0.353) *
      3) ECEC > 4.5 278 350.0 eo (0.687 0.313)
         6) CLAY < 17.5 141 150.0 eo (0.766 0.234)
            12) ELEV < 234.5 125 120.0 eo (0.816 0.184)
               24) TRMI < 28.5 97 76.0 eo (0.866 0.134)
                  48) ELEV < 213.5 66 35.0 eo (0.924 0.076)
                     96) SLOPE < 0.725 7 9.6 eo (0.571 0.429) *
                     97) SLOPE > 0.725 59 17.0 eo (0.966 0.034) *
                  49) ELEV > 213.5 31 35.0 eo (0.742 0.258)
                     98) REL_ASP < 137 26 22.0 eo (0.846 0.154)
                        196) ELEV < 226 13 16.0 eo (0.692 0.308)
                               392) SLOPE < 1 8 0.0 eo (1.000 0.000) *
                               393) SLOPE > 1 5 5.0 rand (0.200 0.800) *
                        197) ELEV > 226 13 0.0 eo (1.000 0.000) *
                     99) REL_ASP > 137 5 5.0 rand (0.200 0.800) *
               25) TRMI > 28.5 28 36.0 eo (0.643 0.357)

# cross validation - cv.tree randomly partitions data into 10 subsets and estimates a
# tree using 90% of the data, then predicts group membership for the remaining
# 10%; repeats until all samples have been classified independently.

spec.tree.cv<-cv.tree(spec.tree,FUN=prune.tree,method="misclass")
par(mfrow=c(2,1))
plot(spec.tree.prune.main="Original Prairie Tree")
plot(spec.tree.cv.main="Cross-validated Prairie tree")

spec.pred<-predict.tree(spec.tree,type="class")

spec.pt6<-prune.tree(spec.tree,method="misclass",best=5)

plot(spec.pt6)
text(spec.pt6,cex=0.6)

summary(spec.pt6)
Classification tree:

snip.tree(tree = spec.tree, nodes = c(4, 10, 14, 11, 6))

Variables actually used in tree construction:

[1] "ECEC"  "REL_ASP"  "SLOPE"  "CLAY"  "ELEV"

Number of terminal nodes: 6

Residual mean deviance: 1.16 = 465 / 401

Misclassification error rate: 0.268 = 109 / 407

**A.3: Python scripts for mapping CART model in GIS:**

```python
# Import system modules
import sys, string, os, arcgisscripting

# Create the Geoprocessor object
gp = arcgisscripting.create()

# Check out any necessary licenses
gp.CheckOutExtension("spatial")

# Load required toolboxes...
gp.AddToolbox("C:/Program Files/ArcGIS/ArcToolbox/Toolboxes/Spatial Analyst Tools.tbx")
gp.AddToolbox("C:/Program Files/ArcGIS/ArcToolbox/Toolboxes/Data Management Tools.tbx")
gp.AddToolbox("C:/Program Files/ArcGIS/ArcToolbox/Toolboxes/Analysis Tools.tbx")

# Set the Geoprocessing environment...
gp.XYResolution = ""
gp.scratchWorkspace = "z:\ED\GIS_DATA"
gp.MTolerance = ""
gp.randomGenerator = "0 ACM599"
gp.outputCoordinateSystem = ""
gp.outputZFlag = "Same As Input"
gp.qualifiedFieldNames = "false"
gp.extent = "-81.455197048497 34.8065220263017 -79.5417656848886 36.0267910787041"
gp.XYTolerance = ""
gp.cellSize = "0.000229"
gp.outputZValue = ""
gp.outputMFlag = "Same As Input"
gp.geographicTransformations = ""
gp.ZResolution = ""
gp.mask = ""
gp.workspace = "z:\ED\GIS_DATA"
```
gp.MResolution = ""
gp.ZTolerance = ""

# Local variables...
rel_asp = "rel_asp"
all_soils_proj = "all_soils_proj"
dem_fill = "dem_fill"
slope = "slope"
lessthan_45_shp = "z:\ED\GIS_DATA\lessthan_45.shp"
slope_less = "z:\ED\GIS_DATA\slope_less"
dem_less = "z:\ED\GIS_DATA\dem_less"
asp_less = "z:\ED\GIS_DATA\asp_less"
cart_left = "z:\ED\GIS_DATA\cart_left"
all_soils_proj_2_ = "all_soils_proj"
morethan_45_shp = "z:\ED\GIS_DATA\morethan_45.shp"
clay_shp = "z:\ED\GIS_DATA\clay.shp"
moreclay_shp = "z:\ED\GIS_DATA\moreclay.shp"
elev_cart2 = "z:\ED\GIS_DATA\elev_cart2"
dem_fill_2_ = "dem_fill"
elev_cart3 = "z:\ED\GIS_DATA\elev_cart3"
cart3 = "z:\ED\GIS_DATA\cart3"
cart4 = "z:\ED\GIS_DATA\cart4"
Input_true_raster_or_constant_value_2_ = "1"
Input_false_raster_or_constant_value_2_ = "0"
all_carts = "z:\ED\GIS_DATA\all_carts"
GIS_DATA = "z:\ED\GIS_DATA"
random_soil = "random_soil"
eos_soil = "eos_soil"
acc_eo = "z:\ED\GIS_DATA\acc_eo"
acc_rand = "z:\ED\GIS_DATA\acc_rand"

# Process: Select...
gp.Select_analysis(all_soils_proj, lessthan_45_shp, "\"ECEC\" < 4.5")

# Process: Extract by Mask (2)...
gp.ExtractByMask_sa(dem_fill, lessthan_45_shp, dem_less)

# Process: Select (2)...
gp.Select_analysis(all_soils_proj_2_, morethan_45_shp, "\"ECEC\" >= 4.5")

# Process: Select (3)...
gp.Select_analysis(morethan_45_shp, clay_shp, "CLAY< 17.5")

# Process: Extract by Mask (5)...
gp.ExtractByMask_sa(dem_fill_2_, clay_shp, elev_cart3)

# Process: Con...
gp.Con_sa(elev_cart3, Input_true_raster_or_constant_value_2_, cart3, Input_false_raster_or_constant_value_2_, "Value < 234.5")

# Process: Select (4)...
gp.Select_analysis(morethan_45_shp, moreclay_shp, "CLAY >= 17.5")

# Process: Extract by Mask (4)...
gp.ExtractByMask_sa(dem_fill_2_, moreclay_shp, elev_cart2)
A.4: Maxent model set-up specifications

The follow parameters and settings were used during the run:

202 presence records used for training.

10202 points used to determine the Maxent distribution (background points and presence points).

Environmental layers used: clay(categorical) ecec(categorical) elev relasp slope trmi

Command line:

Feature types used: Linear Quadratic Product Threshold Hinge

Regularization multiplier is 1.0

Regularization values: linear/quadratic/product: 0.050 categorical: 0.250 threshold: 1.000 hinge: 0.500
Species file is z:\ED\MAXENT\eos.csv

Environmental variables from z:\ED\MAXENT\layers

Output directory is z:\ED\MAXENT\outputs2

Output format is Cumulative

Output file type is .asc

Maximum iterations is 500

Convergence threshold is 1.0E-5

Random test percentage is 0

Jackknife selected

Create response curves selected

B.1: Maxent omission/commission analysis
B.2: Maxent variable contributions and jackknife analysis

The following table gives a heuristic estimate of relative contributions of the environmental variables to the Maxent model. To determine the 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. As with the jackknife, variable contributions should be interpreted with caution when the predictor variables are correlated (reprinted from Maxent results).

<table>
<thead>
<tr>
<th>Variable</th>
<th>Percent contribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>ecce</td>
<td>51.2</td>
</tr>
<tr>
<td>clay</td>
<td>34</td>
</tr>
<tr>
<td>elev</td>
<td>5.2</td>
</tr>
<tr>
<td>trmi</td>
<td>4.2</td>
</tr>
<tr>
<td>relasp</td>
<td>2.9</td>
</tr>
<tr>
<td>slope</td>
<td>2.5</td>
</tr>
</tbody>
</table>
The following picture shows the results of the jackknife test of variable importance. The environmental variable with highest gain when used in isolation is ecec, which therefore appears to have the most useful information by itself. The environmental variable that decreases the gain the most when it is omitted is ecec, which therefore appears to have the most information that isn't present in the other variables (reprinted from Maxent results).

**B.3: Response curves for environmental variables**

These curves show how each environmental variable affects the Maxent prediction. The (raw) Maxent model has the form exp(...) / constant, and the curves show how the exponent changes as each environmental variable is varied, keeping all other environmental variables at their average sample value (reprinted from Maxent results).