A MAXIMUM ENTROPY AND GIS APPROACH TO PREDICT POTENTIAL HABITAT FOR NORTHERN BOBWHITES IN THE BLACK BELT PRAIRIE PHYSIOGRAPHIC REGION OF ALABAMA

By

Claude Lee Jenkins

A Thesis submitted to the Graduate Faculty of
Auburn University at Montgomery
in partial fulfillment of the
requirements for the Degree of
Master of Science
in
Geographic Information Systems

Montgomery, Alabama

30 April 2021

A MAXIMUM ENTROPY AND GIS APPROACH TO PREDICT POTENTIAL HABITAT FOR NORTHERN BOBWHITES IN THE BLACK BELT PRAIRIE PHYSIOGRAPHIC REGION OF ALABAMA

By

Claude Lee Jenkins

A thesis submitted to the Graduate Faculty of
Auburn University at Montgomery
in partial fulfillment of the
requirements for the Degree of
Master of Science

In

Geographic Information Systems

Montgomery, Alabama

30 April 2021

Approved by

Hoehun Ha Digitally signed by Hoehun Ha
DN: cn=Hoehun Ha, o=AUM, ou, email=hha@aum.edu, c=US
Date: 2021.04.30 12:25:57 -05'00'

Dr. Hoehun Ha
Thesis Director

Gabe Costa Digitally signed by Gabe Costa
Date: 2021.05.04 08:56:22 -05'00'

Dr. Gabriel Costa
Second Reader

Nicholas Cuba

Dr. Nicholas Cuba
Third Reader

Matthew Ragland Digitally signed by Matthew Ragland
Date: 2021.05.04 11:25:55 -05'00'

Dr. Matthew Ragland
Associate Provost

Acknowledgements

My sincere thanks goes to the late Dr. Terence Winemiller for genuinely accepting me as a non-traditional graduate student. The obvious passion for his work contributed to many enthusiastic conversations about using GIS technology to investigate various spatial-related questions. His expertise and passion will be missed.

I greatly appreciate Dr. Hoehun Ha for accepting me as his graduate student after the passing of Dr. Winemiller. Dr. Ha's clear guidance and skilled advice was precisely what I needed to complete tasks in a successful and timely manner. Without doubt, Dr. Ha is the consummate graduate advisor, and I am sincerely thankful for his help. It was truly a pleasure to work with him.

I will be forever grateful for my wife, Jenise, who quietly endured as I used family time to pursue my graduate degree. My scholarly endeavors would not have been possible without her unwavering support and willingness to forgo family time. Thank you for your understanding, support, and enduring patience. I look forward to family time again!

Table of Contents

List of Figures

List of Tables

Abstract

Northern bobwhite (*Colinus virginianus*) populations have experienced precipitous declines throughout their range. The North American Breeding Bird Survey for Alabama indicates a rate of decline of over six percent per year since 1966. Declining bobwhite populations have been linked to large-scale habitat loss and a reduction in habitat quality due to advances in ecological succession, intensive monoculture agriculture, and intensive timber management practices. Once a very diverse region, modern land-use practices throughout the Black Belt Prairie physiographic region have created a simplified landscape that have left few places where bobwhites can live and thrive. Despite extensive research on bobwhites, there is insufficient knowledge on the spatial distribution of bobwhite habitat on a regional scale, including the Black Belt Prairie physiographic region. Knowing the spatial distribution of suitable bobwhite habitat will allow conservation planners to prioritize efforts and direct limited resources to achieve conservation goals for the region. The goals of this study were to 1) evaluate the maximum entropy modeling approach for predicting northern bobwhite habitat in the Black Belt Prairie physiographic region, and 2) develop a northern bobwhite habitat suitability map for the Black Belt Prairie physiographic region. The results showed that deciduous 1.8 km (1.8 km is the dispersal distance of bobwhites) neighborhood was the most important habitat variable, followed by water 1.8 km neighborhood. Developed 1.8 km neighborhood was the most important anthropogenic variable, ranking third in overall importance. The resultant model denotes potential habitat in the region, and forms the basis for using the maximum entropy and geographic information system approach to predict suitable habitat and prioritize areas for conservation.

Chapter 1: Introduction

1.1 Background

Northern bobwhite (*Colinus virginianus*) populations have experienced precipitous declines throughout their range (Sauer et al. 2000). In the Southeastern United States, the North American Breeding Bird Survey (BBS) indicates that populations have been declining at over five percent per year since 1966 (Figure 1). The BBS is conducted by the United States Geological Survey, Patuxent Environmental Science Center, and provides a consistent measure of bobwhite population trends range-wide (L. Wes Burger. 2001). Surveys for Alabama indicate a rate of decline of over six percent per year during the same time period (Sauer et al. 2017) (Figure 2). However, bobwhite population declines were observed prior to the BBS (Stoddard 1931). Declining bobwhite populations have been attributed to a variety of factors including coyotes, fire ants, avian predators, and pesticides; however, the primary cause has been the cumulative effects of large-scale habitat loss and poor quality habitat associated with advanced ecological succession (Roseberry et al. 1979, Fies et al. 1992), intensive monoculture agriculture (Vance 1976, Exum et al. 1982, Roseberry 1993), and intensive timber management practices (Brennan 1991). Guthery (1997) describes this as a range-wide reduction in useable space. Intensive, modern agricultural practices have created a landscape void of grassy/weedy plant communities essential for nesting and brood rearing, and woody fencerows for protective cover. The reduction in habitat complexity has reduced the capability of the agriculture landscape to support bobwhites (Kabat and Thompson 1963). In forested areas of the Southeast, lack of fire and fire exclusion (Brennan et al. 1998), increase in forest coverage and extent of densely planted pine plantations, and intensive silviculture practices have reduced grassy/weedy areas required for reproduction and foraging (Fies et al. 1992). Once a very diverse region that included tallgrass prairies, upland oak savannas, longleaf grasslands, cedar glades, and other suitable habitat types, modern land use practices throughout the Black Belt Prairie physiographic region have simplified the landscape and left few places where bobwhites can live and thrive. Furthermore, areas with ample space support very few bobwhites due to the lack of proper management.

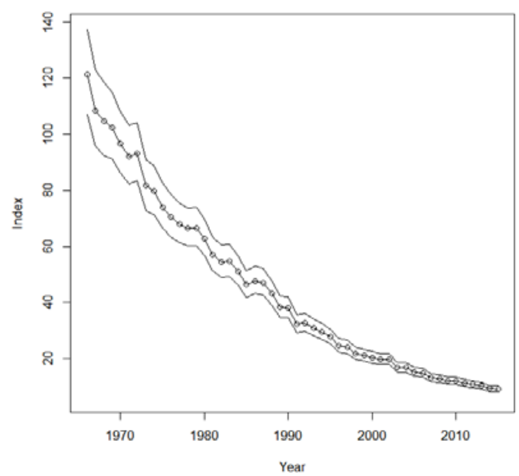


Figure 1. Northern bobwhite population trends in the Southeast, 1966-2015.

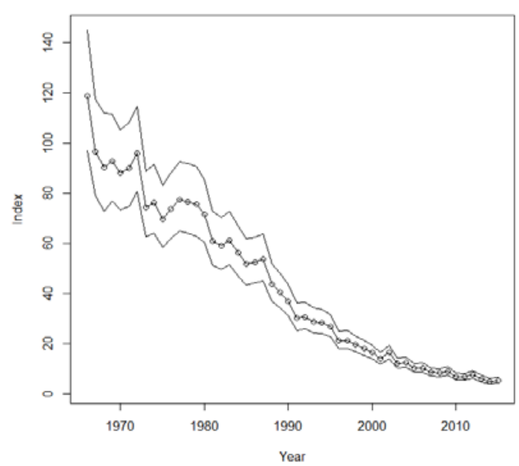


Figure 2. Northern bobwhite population trends in Alabama, 1966-2015.

To address these declines, the Southeast Quail Study Group (SEQSG) published the Northern Bobwhite Conservation Initiative (NBCI) in 2002; the NBCI was a range-wide plan to restore bobwhite populations to the 1980 levels. However, the NBCI was a paper-based plan that had limited utility as a conservation planning tool. For example, wildlife biologists did not have access to a large database that would allow for quick analysis of potential habitat projects and customize conservation planning at a sub-county level. Therefore, in 2008, the National Bobwhite Technical Committee (formally SEQSG) revised the NBCI (renamed the National Bobwhite Conservation Initiative) using a geographic information system (GIS). The revision was accomplished through workshops where participating wildlife biologists developed spatially-explicit estimates of low, medium, and high priority landscapes for bobwhite recovery; this information is referred to as the biologist ranking information (BRI) (Figure 3). The BRI is unique in conservation planning as it relied on expert knowledge of landscape attributes to map priority areas. The GIS-based plan (NBCI ver. 2.0) has tremendous flexibility to aid in conservation planning at regional, state, and local spatial scales (The National Bobwhite Technical Committee 2011).

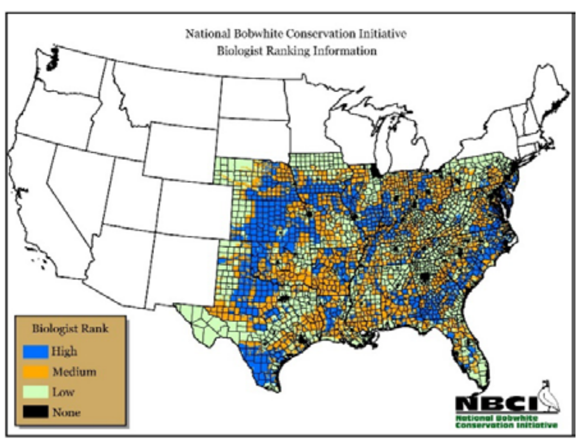


Figure 3. Range-wide biologist ranking information.

Although the NBCI 2.0 was an improvement from the original plan, it did not have the capability to predict the current distribution of the northern bobwhite. In order to make accurate conservation decisions, conservationists need the best information on the bobwhite's distribution status. Habitat suitability models are used to identify suitable areas for bobwhites and other declining species, and can be developed from many types of data, including presence-only data (Anderson and Martinez-Meyer 2004; Phillips et al. 2009; Cianfrani et al. 2010).

Maximum entropy (Maxent) is based on machine learning that is designed to make predictions from presence-only data. It is a common method for modeling the distribution of a species by finding the distribution that is closest to uniform (Phillips et al. 2006). Maxent uses presence-only data and a set of environmental layers (habitat and anthropogenic) to predict the probability that habitat conditions are suitable for species occurrence. Although presence-absence models have been reported to be more accurate in predicting species occurrence, it can be very challenging to collect absence data in the field (i.e. true absence data), particularly at large spatial scales (Cianfrani et al. 2010, Hastie and Fithian 2013). For northern bobwhites, a small, inconspicuous species, it is impossible to collect true absence data at a large scale; therefore, Maxent with presence only data is more feasible to assess habitat suitability in the Black Belt Prairie physiographic region.

The Black Belt Prairie physiographic region of Alabama, particularly the non-forested portion, has received little attention for bobwhites, although they occur throughout the region. This relative lack of attention might be attributed to the size of the Black Belt compared to other landscapes and loss of habitat to other land uses. However, interest in the region from conservationists has increased in recent years; albeit, with no tool to strategically guide bobwhite conservation planning. This study will support the growing interest by identifying the spatial distribution of potential suitable habitat for bobwhites. Furthermore, study results will assist conservation planners with federal and state agencies, and non-government organizations to identify areas to prioritize and direct limited resources, and accomplish bobwhite conservation goals for the region.

1.2 Objectives

Despite the extensive history of northern bobwhite research and management, there is limited knowledge regarding the spatial distribution of suitable habitat at a landscape scale. The National Bobwhite Conservation Initiative Biologist Ranking Information (NBCI BRI) represents the subjective judgement of wildlife experts and provides the fundamental framework for bobwhite conservation at a landscape scale; however, the NBCI BRI does not precisely identify the location and spatial distribution of suitable habitat.

One of the objectives of this study is to determine the spatial distribution of suitable habitat at a landscape scale (Black Belt Prairie physiographic region). Consequently, conservation planners can use this new knowledge to precisely distribute technical expertise, financial assistance, and other resources in a manner that optimizes the return on habitat conservation investment. Conservation investments should be placed in the landscape that have the greatest potential for eliciting a sustained bobwhite population response; therefore, an objective, experimental-based approach to identify suitable bobwhite habitat is needed for spatially explicit allocation of expertise and resources.

The ultimate purpose of this study is to contribute to an enhanced knowledge of suitable bobwhite habitat distribution and a more precise delivery of conservation efforts that equate to an increase in northern bobwhite populations in the Black Belt Prairie physiographic region of Alabama. The specific objectives of this study are to 1) evaluate the maximum entropy modeling approach for predicting northern bobwhite habitat in the Black Belt Prairie physiographic region of Alabama, and 2) develop a northern bobwhite habitat suitability map for the Black Belt Prairie physiographic region.

1.3 Study Outline

The outline for this study followed a path from planning to interpretation. A review of the scientific literature assisted in determining the appropriate habitat and anthropogenic environmental layers used in Maxent. The Alabama Wildlife Federation's landowner database was used to identify locations where northern bobwhite populations are known to occur. Presence data were derived from these locations and were used as samples in the Maxent model. Multiple spatial datasets were gathered from appropriate sources that were required for use in Maxent and ArcGIS. All datasets were processed to prepare them for use in Maxent. A preliminary and final run of the Maxent model produced an output map that was reclassified to produce the likelihood distribution map. The outline of the study is shown in Figure 4.

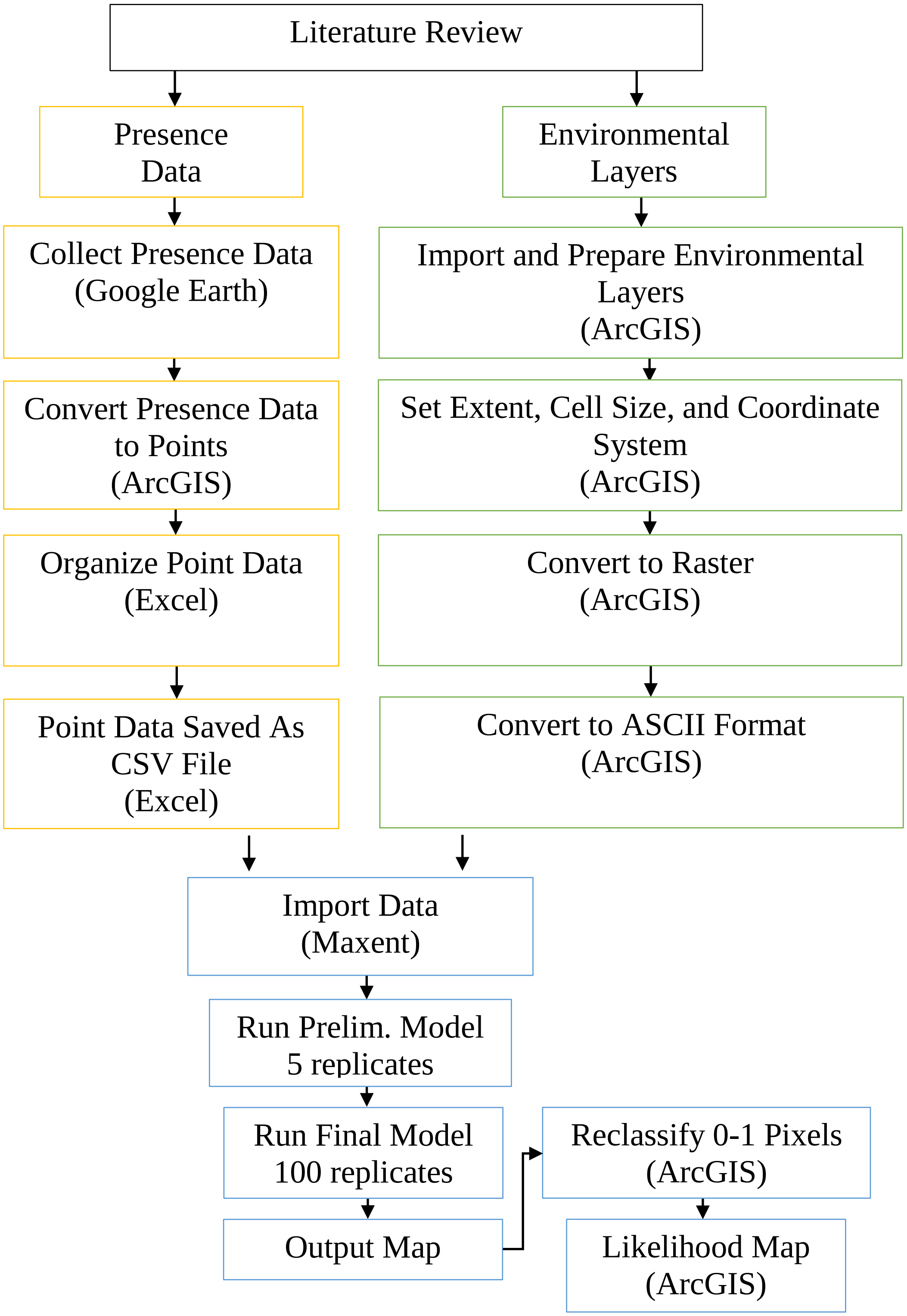


Figure 4. Outline of Study Methodology.

1.4 Literature Review

The spatial distribution of habitats influences the distribution of wildlife populations (Fuller 2012). Therefore, knowing the spatial distribution of suitable northern bobwhite habitats would allow more focused efforts by conservation planners and quail managers (Schairer et al. 1999). Habitat suitability is the likelihood that a species utilizes a particular habitat (Kearney 2006, Wang et al. 2008). Researchers have been using models for decades to determine habitat suitability for bobwhites. Habitat suitability models relate landscape variables to species occurrence (Hirzel and Le Lay 2008) by using presence-only data or presence-absence data (Brotons et al. 2004, Pearce and Boyce 2006). Presence-absence models are more accurate in predicting occurrences; however, collecting true absence data is extremely challenging at a large spatial scale (Cianfrani et al. 2010, Hastie and Fithian 2013). For northern bobwhites, a small, inconspicuous species, true absence data is practically impossible to verify given the potential for bobwhites to be present in an area but not observed; this, of course, results in a biased bobwhite-habitat relationship. Therefore, habitat suitability models with presence-only data are more feasible for large-scale habitat suitability assessments (Brotons et al. 2004, Zimmermann et al. 2010).

Maximum entropy is a relatively new machine learning method that uses presence-only data, thus eliminating the need for true absence data. It has been used successfully to model species distributions (Merow et al. 2013) with a limited number of presence-only data (Phillips et al. 2005; Elith et al. 2011; Merow et al 2013). Maximum entropy has been widely used to create wildlife habitat suitability maps with remarkable predictive accuracy (Phillips et al. 2006, Kampichler et al. 2010). Consequently, the recently developed Maxent modeling program has proven to be a very useful tool for determining species distributions and wildlife-habitat relationships. Although Maxent has become a standard for habitat suitability mapping (Elith et al. 2011), its use to predict habitat suitability for bobwhites as a single species is extremely limited. The northern bobwhite has been studied intensively throughout its range; nonetheless, there are still gaps, particularly in using maximum entropy to predict bobwhite habitat suitability at a landscape scale. Furthermore, no studies have been conducted that used maximum entropy (Maxent) and a Geographic Information System (GIS) to predict bobwhite habitat suitability in the Black Belt Prairie physiographic region of Alabama.

Chapter 2: Methods

2.1 Study Area

The Black Belt Prairie physiographic region is a relatively narrow crescent that is approximately 500 kilometers long and 40 kilometers wide (Barone 2005), extending from southwest Tennessee south through east-central Mississippi and east through central Alabama near the Georgia border (Figure 5). The region is named for the fertile, dark clay soils that were derived from the underlying chalk bedrock. Instead of a continuous, open grassland as in the Great Plains, the historic landscape of the region was heterogeneous, consisting of distinct prairies interspersed through a matrix of different types of upland and bottomland forests (Barone 2005). Forest types include loblolly pine (*Pinus taeda*) plantations, longleaf pine (*Pinus palustris*), upland hardwoods (*Carya* spp. and *Quercus* spp.), mixed pine-hardwoods (*Pinus-Quercus*), and bottomland hardwoods (*Quercus* spp.). The diverse and floristic uniqueness of the Black Belt Prairie physiographic region provided important habitats for numerous species of wildlife, including the northern bobwhite.

Surveys conducted by the General Land Office in the 1830's reported that prairies covered at least 144,000 hectares of the Black Belt region (Barone 2005), with approximately 73,060 hectares in Alabama (Schotz and Barbour 2009). Since that time, more than 99% of the prairies have been converted to agriculture or lost to development (Noss et al. 1995). Due to the region's fertile soil, most of the prairies were converted to agriculture, including row crops and exotic grasses for grazing and haying. The remaining prairies are threatened by erosion, conversion to other land uses, and the encroachment of woody vegetation (e.g. *Juniperus virginiana L.*) due to fire suppression. The fifteen Alabama counties identified by Schotz and Barbour (2009) were used as the study area (Figure 6).

Figure 5. Alabama and Mississippi with the Black Belt region in gray.

Figure 6. The study area includes Alabama counties with Black Belt prairies.

2.2 Maxent

Maxent (version 3.4.1) (Figure 7) was used to develop the maximum entropy model to predict potential habitat for northern bobwhites in the Black Belt Prairie physiographic region of Alabama. Maxent is a model that is based on machine learning that makes predictions about presence-environment relationships from incomplete data (Baldwin 2009); it requires only presence data (i.e. samples) and a set of environmental layers (e.g. water, developed areas, etc.) for the study area. Environmental layers can be both categorical and continuous data. Because Maxent uses presence-only data, there is no need for actual absence data (Baldwin 2009).

By default, Maxent uses a uniform distribution probability that assumes a species has an equal likelihood of being anywhere in the study area. By using a uniform distribution probability, it predicts a distribution that is most spatially spread out (Merow et al. 2013). Maxent was used in this study to estimate the most uniform distribution of bobwhite locations (sample points; presence-only data) compared to the background (environmental layers used as explanatory variables) where presence-absence was not measured. That is, the model's algorithm compares the bobwhite locations to all of the environments in the study area, and defines the environments by sampling a large number of points throughout the study area; the sampled points are called background points (locations). The algorithm will eventually converge to the probability distribution of maximum entropy. The default output, and the output used in this study, is logistic. A logistic output provides a predicted probability of presence, or for this study, a predicted probability of suitable habitat, between 0 and 1 for each pixel in the study area.

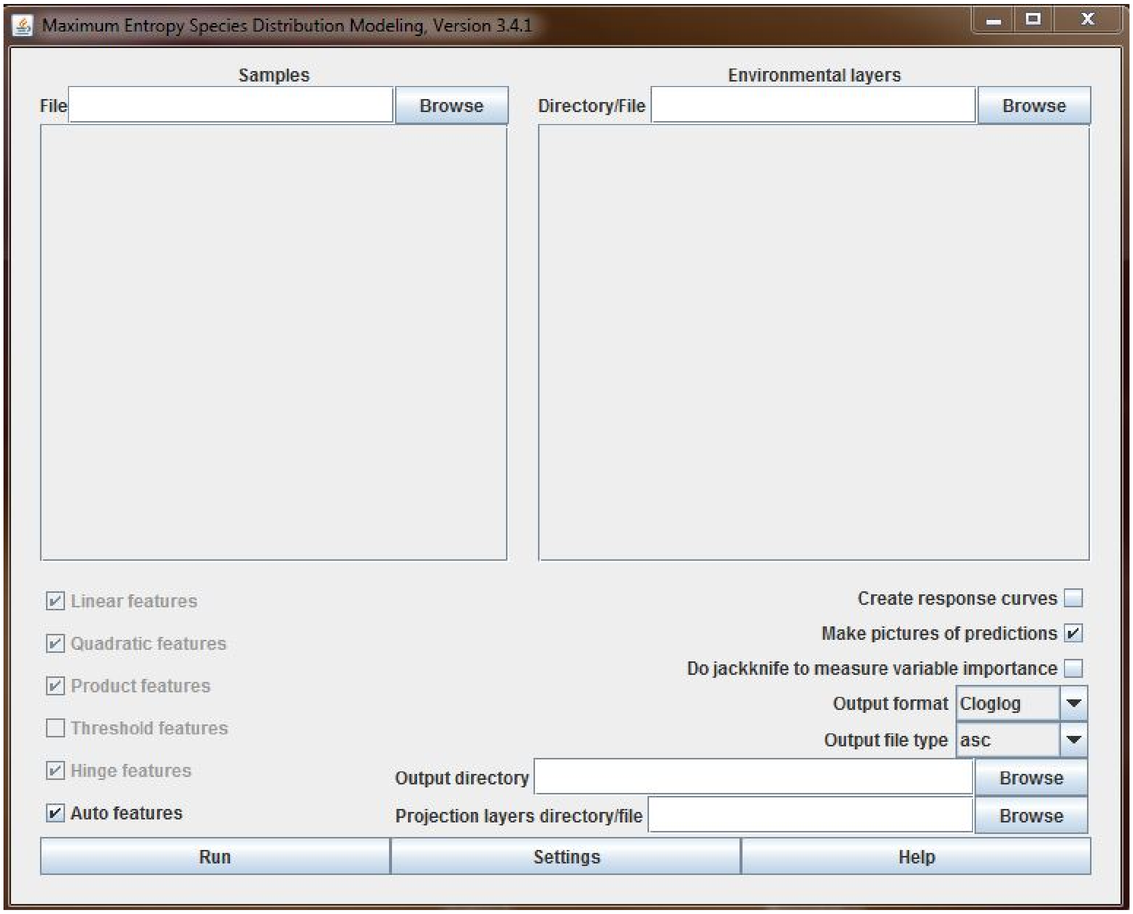


Figure 7. User interface for Maxent (version 3.4.1).

2.3 Presence Data

The Alabama Wildlife Federation's landowner database was used to identify locations where habitat management for northern bobwhites is deliberate and bobwhite populations are considered stable or increasing (Figure 8). The suitable habitat conditions made these locations ideal for collecting presence data. Twenty-six polygons that represent bobwhite locations were digitized in Google Earth and saved as a KML file. To be used in ArcGIS, the KML files were converted to layers using the conversion tool in ArcGIS.

All twenty-six layers were merged into one polygon shapefile using the merge tool in the data management toolbox. For use in Maxent, the polygon was converted to a

30 x 30 meter raster grid. The raster grid was converted into 22,977 points using the raster to point tool in ArcGIS; the tool converts raster grids to points by using the center of each raster grid cell. Table 1 includes the number of points/samples for each location. The x and y coordinates were added to the raster to point shapefile by using the add xy tool in the data management toolbox. The x and y coordinates were copied into an Excel spreadsheet and saved as a comma-separated value (CSV) file (Figure 9) to be used in Maxent; a CSV file is the required format for the samples used in Maxent.

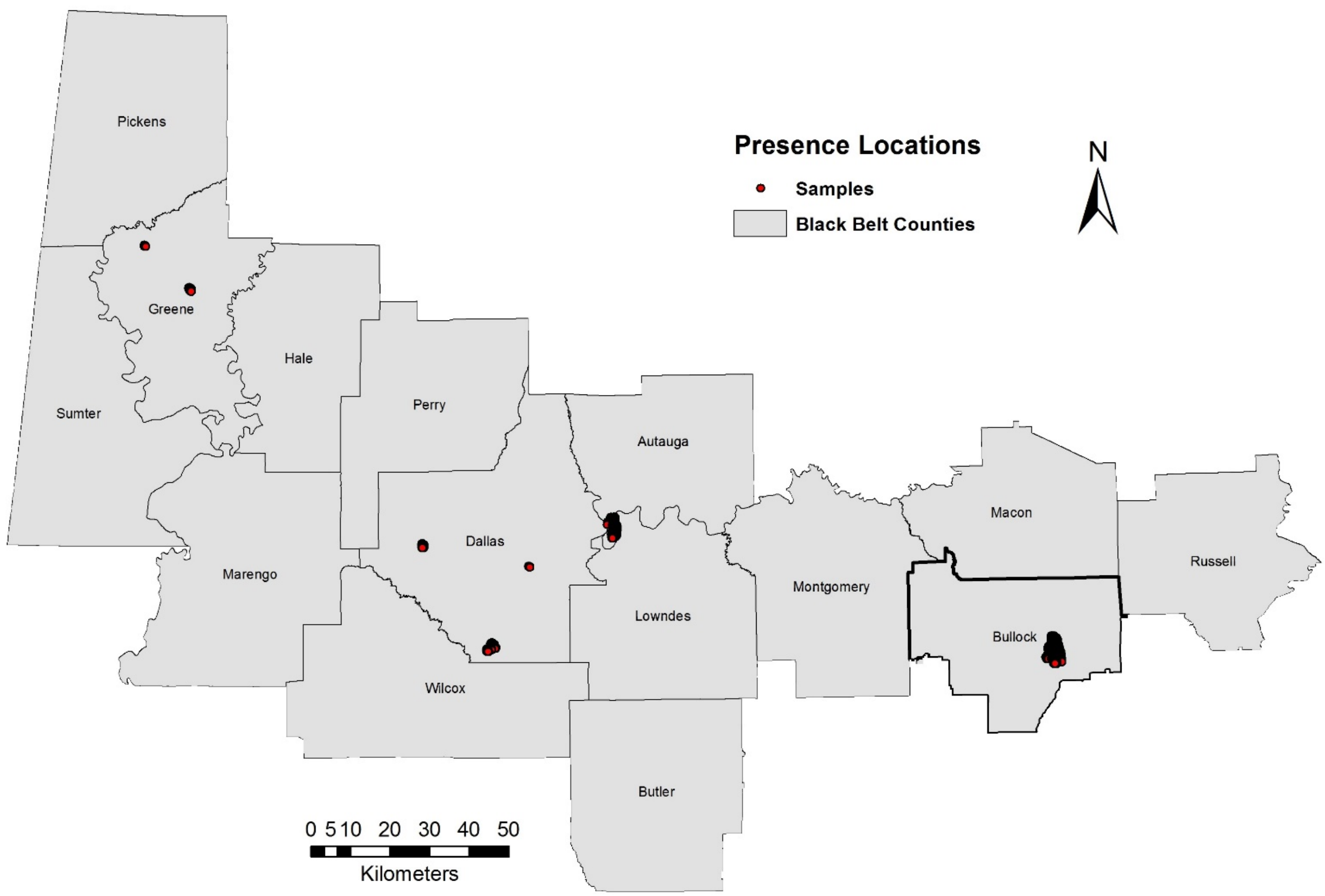
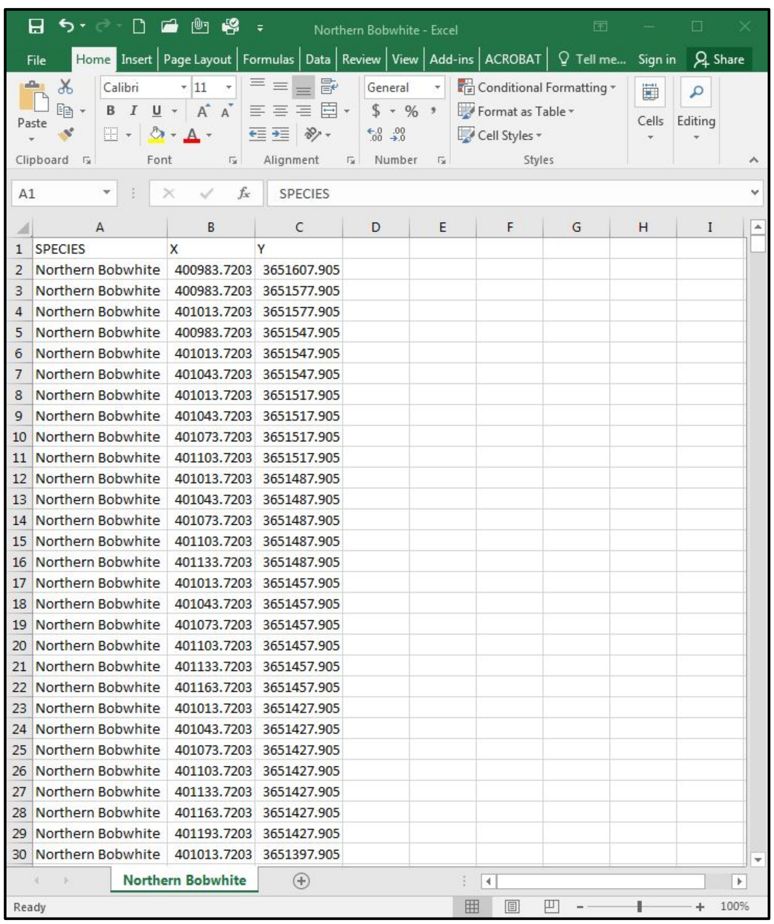


Figure 8. Locations where stable or increasing northern bobwhite populations are known to occur. Presence data were obtained from these locations.

Table 1. Number of samples at each location.

# Locations	County	Samples
1	Autauga	5,687
2	Bullock	14,830
3	Dallas	462
4	Dallas	1,388
5	Dallas	82
6	Greene	108
7	Greene	420
	TOTAL	**22,977**



	A	B	C
1	SPECIES	X	Y
2	Northern Bobwhite	400983.7203	3651607.905
3	Northern Bobwhite	400983.7203	3651577.905
4	Northern Bobwhite	401013.7203	3651577.905
5	Northern Bobwhite	400983.7203	3651547.905
6	Northern Bobwhite	401013.7203	3651547.905
7	Northern Bobwhite	401043.7203	3651547.905
8	Northern Bobwhite	401013.7203	3651517.905
9	Northern Bobwhite	401043.7203	3651517.905
10	Northern Bobwhite	401073.7203	3651517.905
11	Northern Bobwhite	401103.7203	3651517.905
12	Northern Bobwhite	401013.7203	3651487.905
13	Northern Bobwhite	401043.7203	3651487.905
14	Northern Bobwhite	401073.7203	3651487.905
15	Northern Bobwhite	401103.7203	3651487.905
16	Northern Bobwhite	401133.7203	3651487.905
17	Northern Bobwhite	401013.7203	3651457.905
18	Northern Bobwhite	401043.7203	3651457.905
19	Northern Bobwhite	401073.7203	3651457.905
20	Northern Bobwhite	401103.7203	3651457.905
21	Northern Bobwhite	401133.7203	3651457.905
22	Northern Bobwhite	401163.7203	3651457.905
23	Northern Bobwhite	401013.7203	3651427.905
24	Northern Bobwhite	401043.7203	3651427.905
25	Northern Bobwhite	401073.7203	3651427.905
26	Northern Bobwhite	401103.7203	3651427.905
27	Northern Bobwhite	401133.7203	3651427.905
28	Northern Bobwhite	401163.7203	3651427.905
29	Northern Bobwhite	401193.7203	3651427.905
30	Northern Bobwhite	401013.7203	3651397.905

Figure 9. An Excel spreadsheet with presence data that was saved as a CSV file.

2.4 Environmental Layers

Land cover data were obtained from the 2016 National Land Cover Database (NLCD) that is managed by the Multi-Resolution Land Characteristics (MRLC) Consortium (https://www.mrlc.gov). The MRLC is a group of federal agencies who work cooperatively to generate land cover data and information at the national scale for a wide variety of applications. The MRLC provides data on land cover at a 30-meter resolution with 16 land cover classes. All MRLC NLCD data products can be downloaded at no charge (Homer et al. 2015).

The 2016 NLCD 16-class land cover classification scheme was reclassified using the spatial analyst tool in ArcGIS to create a 10-class land cover classification scheme. “Pasture/hay” and “cultivated crops” were combined to create the “agriculture” land cover. “Developed, open space,” “developed, low intensity,” “developed, medium intensity,” and “developed, high intensity” land cover classes were combined to create the “developed” land cover. Lastly, “woody wetlands” and “emergent herbaceous wetlands” were combined to create the “wetland” land cover class (Figure 10). The focal statistics tool in ArcGIS was used to generate Boolean layers for each land cover class. The focal statistics tool is ideal for this operation because it provides options for determining the neighborhood type (circle, rectangle, etc.) and statistic to be calculated (e.g. mean, average, sum, etc.). For each layer, cells of that particular land cover class equal one and all other cells equal zero. Generating Boolean layers from a single, multi-class layer allowed a neighborhood operation to be performed on each land cover class.

Neighborhood data were generated using the focal statistics tool in ArcGIS. Two circle neighborhoods were generated for each land cover class by specifying a radius of

400 meters and 1800 meters. A circle neighborhood of 400 meters (50 hectares) was used to represent the bobwhite's median home range size (Smith and Burger 2003). A circle neighborhood of 1800 meters (1.8 kilometers) was used to represent the bobwhite's dispersal distance (Dimmick 1992). The list of environmental layers is shown in Table 2.

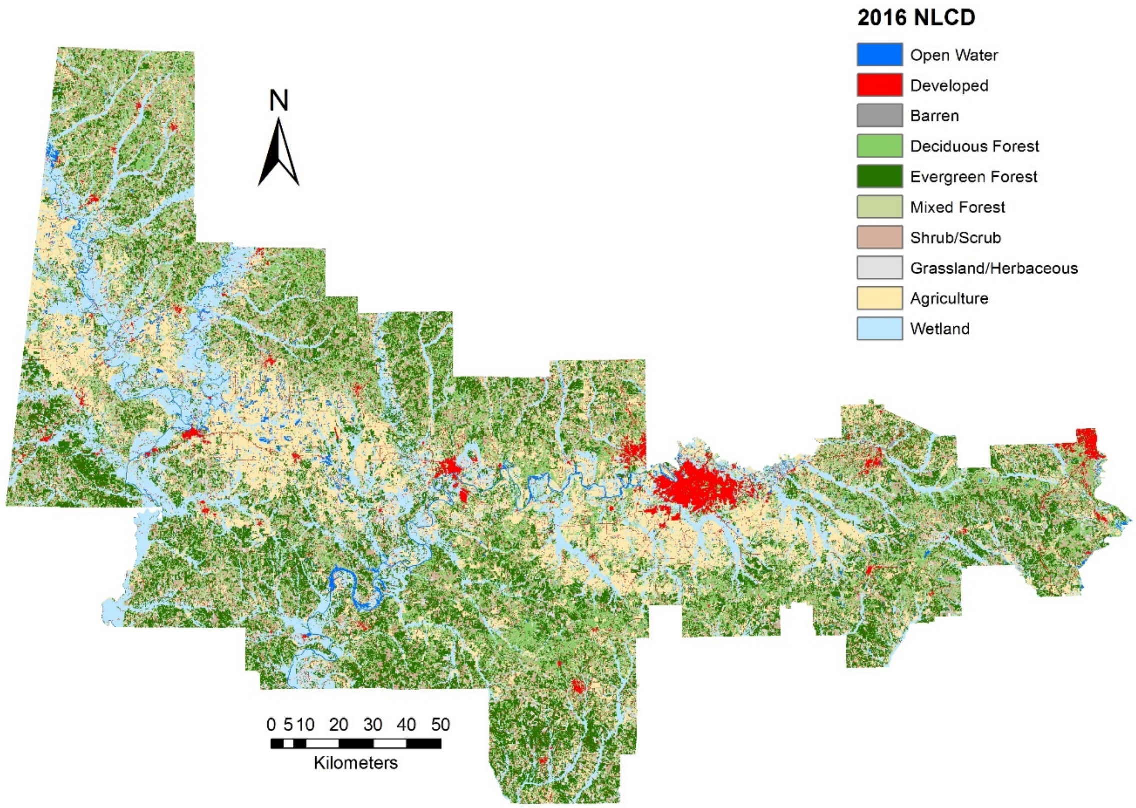


Figure 10. 2016 National Land Cover Database.

Table 2. List of environmental layers/variables used in Maxent.

Layer / Variable	Description	Year
NLCD	National land cover (10 classes)	2016
Elevation	Digital elevation model	2020
Agriculture 1.8km	Total agriculture area within 1.8km	2016
Agriculture 400m	Total agriculture area within 400m	2016
Barren 1.8km	Total barren area within 1.8km	2016
Barren 400m	Total barren area within 400m	2016
Deciduous 1.8km	Total deciduous forest area within 1.8km	2016
Deciduous 400m	Total deciduous forest area within 400m	2016
Developed 1.8km	Total developed area within 1.8km	2016
Developed 400m	Total developed area within 400m	2016
Evergreen 1.8km	Total evergreen forest area within 1.8km	2016
Evergreen 400m	Total evergreen forest area within 400m	2016
Grassland/Herbaceous 1.8km	Total grassland/herbaceous area within 1.8km	2016
Grassland/Herbaceous 400m	Total grassland/herbaceous area within 400m	2016
Mixed 1.8km	Total mixed forest area within 1.8km	2016
Mixed 400m	Total mixed forest area within 400m	2016
Shrub/Scrub 1.8km	Total shrub/scrub area within 1.8km	2016
Shrub/Scrub 400m	Total shrub/scrub area within 400m	2016
Water 1.8km	Total water area within 1.8km	2016
Water 400m	Total water area within 400m	2016
Wetland 1.8km	Total wetland area within 1.8km	2016
Wetland 400m	Total wetland area within 400m	2016

Digital elevation models (DEM) of the study area were obtained from the Geospatial Data Gateway (GDG) (https://www.datagateway.nrcs.usda.gov). The GDG provides access to a repository of hundreds of high resolution environmental and natural resources vector and raster layers. The service is made possible through a partnership between the Natural Resources Conservation Service, Farm Service Agency, and Rural Development. Twelve 30-meter DEMs in raster format were required to cover the study area. All DEMs were combined into one by using the mosaic to new raster tool in ArcGIS (Figure 11).

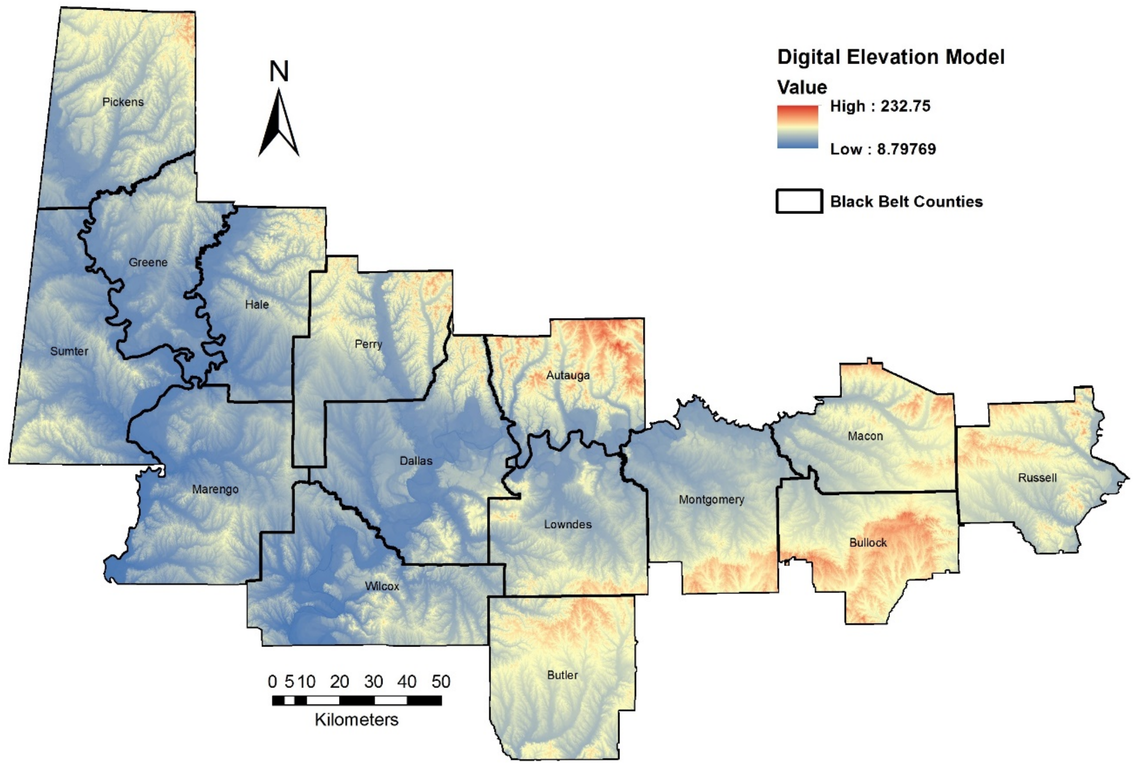


Figure 11. Digital Elevation Model.

2.5 Data Preparation

Maxent is designed to integrate with ArcGIS; however, all data must be in proper formats for smooth integration and to run the model. The extract by mask tool in ArcGIS was used to set the processing extent, cell size, and projection system for all environmental layers. The processing extent was set at 11182 columns x 7495 rows. Cell size was set at 30m x 30m and the projection system was set to WGS 1984_UTM_Zone_16N. The next step in preparing the data required converting all environmental raster layers to ASCII (.asc) format; this is a requirement of Maxent in order to run the model. The raster to ASCII tool in ArcGIS was used to convert the environmental layers to ASCII files.

2.6 Maxent Settings

To begin, Maxent was opened to the main graphical user interface (GUI) where the samples file and environmental layers were uploaded. The samples file contains the presence data in CSV format, and the environmental layers are the elevation and NLCD 2016 layers in ASCII format. The environmental layers must be set as "categorical" or "continuous." Categorical variables are discrete values such as vegetation type and soil type, while continuous variables are measured values such as temperature, precipitation, etc. The elevation and neighborhood ASCII layers were set as continuous while the NLCD 2016 ASCII layer was set as categorical. All settings on the GUI were selected except "Threshold features" as no tolerance or limit exists. "Logistic" was selected as the output format to show a probability that suitable bobwhite habitat will occur within a particular grid cell. The completed main GUI is shown in figure 12.

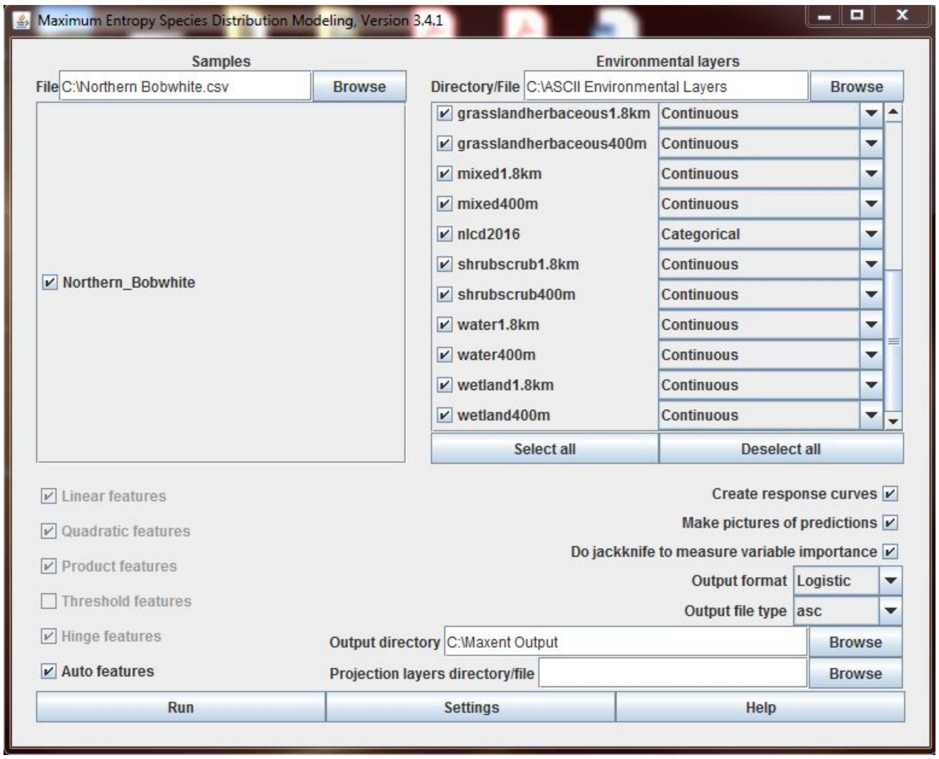


Figure 12. The completed main Maxent graphical user interface.

The Maxent settings tab on the main GUI contains three different tabs: 1) Basic, 2) Advanced, and 3) Experimental, where model parameters are set before running Maxent. Although there are dozens of parameter options among the three tabs, only the parameters selected for this model will be mentioned. The parameters selected for this model are summarized in Table 3.

Table 3. Parameters used to run the Maxent model.

Parameter	**Tab**	**Summary**
Replicates	Basic	Number of times to run the model
Random test percentage	Basic	Percent of presence points set aside as test points to evaluate model performance
Replicate run type	Basic	Testing and training method used when running multiple replicates
Maximum iterations	Advanced	The maximum number of iterations to run
Regularization multiplier	Basic	Reduces model overfitting
Max number of background points	Basic	The number of cells randomly selected for background points

Replicates are the number of times to run a model. Running a model multiple times allows Maxent to average the results from all of the models that were created. Five (5) replicates were chosen for the initial model. To evaluate model performance, an 80-20 split was used for training and testing. That is, 80 percent of the presence data (21,445 points) was used to create the model while 20 percent (1,531 points) was withheld and used to assess the accuracy of the model. Using a certain percentage of presence data to evaluate a model's performance is important. Without these test data, Maxent will use training data (data used to create the model) to evaluate the model, which is a bias method for evaluating model performance.

Maxent offers three different replicate run types (sampling techniques) that can be used as testing methods when running multiple replicates: crossvalidate, bootstrap, and subsample. Crossvalidate was selected for this model. Crossvalidate randomly divides the sample data into different equal-size groups or "folds." Each fold is used separately, and omitted in turn, to test the performance of the training data during each iteration (Friedlaender et al. 2011). The omitted folds are used to evaluate the model. This crossvalidate technique in Maxent is referred to as "K-folds." The number of iterations was set to five (5) to allow Maxent to converge to maximum entropy and determine which environmental layers/variables contribute most to bobwhite habitat. Maxent has a regularization protocol that reduces model overfitting (i.e. predicted distributions that are clustered near presence locations). The regularization multiplier can be adjusted, however, the default setting of one (1) was used for this model as studies have shown that default settings perform as well as adjusted settings (Phillips and Dudik 2008).

Lastly, a bias file (raster grid) was uploaded to determine where Maxent selects the background points (locations). The background points represent the same selection bias as the presence data; this strives to achieve the same environmental bias in both presence and background data sets (Phillips et al. 2009). For this study, 22,977 background points (same number as presence points) were selected from the same counties where presence points (locations) were sampled (Ferrier et al. 2002). A Maxent model that is based on presence and background data with the same bias will not focus on the bias, rather it will focus on the habitats occupied by a species (Dudik et al. 2005).

2.7 Running Maxent

An initial model (five-run replicates) using an 80-20 split was conducted to identify the importance of each environmental layer to the prediction of bobwhite habitat and reduce the number of layers to the most relevant. Irrelevant layers were removed based on the following criteria: 1) the layers decreased training or testing gain, 2) the layers contributed less than one percent to the model, or 3) the layers' response curves were static when all other layers were held at their mean (Ha et al. 2016 and Heumann et al. 2013). Because replicates were set at five, Maxent ran the initial model five times and then averaged the results from all of the models that were generated. Outputs from Maxent for the initial run were a likelihood distribution map, and a variable contribution and permutation table.

A jackknife test was applied to all environmental layers to estimate the contribution of the layers to the model, and determine the training and testing gain for each layer. The test iteratively omits each layer and considers each layer in isolation which allows a comparison of the relative importance of each layer to the habitat distribution. Outputs from the jackknife test include a chart that illustrates the importance of each environmental layer.

The final list of environmental layers was reduced from 22 to 13 (Table 4) using the criteria listed above. This list was used to conduct the 100-run replicates and produce the final model. The parameters that were used in the initial model were used in the final model except for iterations which were set at 100. After the 100-run replicates were completed, all of the outputs created by Maxent were saved in the output folder.

Table 4. Relative contributions of the environmental layers to the Maxent model averaged over 5-run replicates. The variables in bold were used in the final model.

VARIABLE	PERCENT CONTRIBUTION	PERMUTATION IMPORTANCE
Deciduous 1.8km	33.5	43.5
Water 1.8km	22.1	18.5
Developed 1.8km	10.8	1.7
Elevation	6.3	6.2
Mixed 1.8km	5.2	3.9
Wetland 400m	3.4	0.8
NLCD 2016	3	0.1
Wetland 1.8km	2.9	4.6
Barren 1.8km	2.9	4.8
Shrub/scrub 1.8km	2.4	2.9
Grassland/herbaceous 1.8km	2	1.3
Evergreen 1.8km	1.4	5.6
Agriculture 1.8km	1.2	4.3
Shrub/scrub 400m	0.7	0.4
Deciduous 400m	0.7	0.8
Mixed 400m	0.3	0.2
Agriculture 400m	0.3	0.3
Evergreen 400m	0.3	0.3
Barren 400m	0.2	0
Grassland/herbaceous 400m	0.1	0.1
Water 400m	0.1	0
Developed 400m	0	0

Chapter 3: Results

3.1 Model Evaluation

Maxent contains statistical tools to assist in the evaluation of model performance. The most accepted statistical tool for evaluating model performance for Maxent is the area under the curve (AUC) which is determined by the receiver operating curve (ROC). AUC-ROC is a statistic that was adapted to presence-only modeling methodology

(Fielding and Bell 1997). AUC is usually used to determine how a model compares and distinguishes between presence locations and absence locations, but with presence-only data AUC compares presence locations with background locations (Merow et al. 2013). The AUC value ranges from 0 to 1. An AUC value of 0 is under-fitted while an AUC value of 1 is perfectly-fitted. Model fitting is a measure of how well Maxent makes predictions based on the data on which it was trained (i.e. fit the data). A well-fitted model produces more accurate outputs, while an under-fitted model produces less accurate outputs. Model fitting is fundamental to a machine learning model such as Maxent. If a model doesn’t fit the data correctly, the outputs will not be accurate enough for making decisions. The average AUC for the final, 100-replicates model is 0.744 with a standard deviation of 0.011 (Figure 16). The black line in Figure 16 represents a random prediction. That is, the line indicates what would be expected if the model was no better than random (AUC = 0.5). The red line indicates the fit of the model compared to the training data (AUC = 0.744). With an AUC of 0.744, the Maxent model predicts the distribution of habitat for northern bobwhites relatively well. Swets (1988), Graham and Hijmans (2006) and, Pearce and Ferrier (2000) suggests that AUC values between 0.7 and 0.9 are appropriate for many uses.

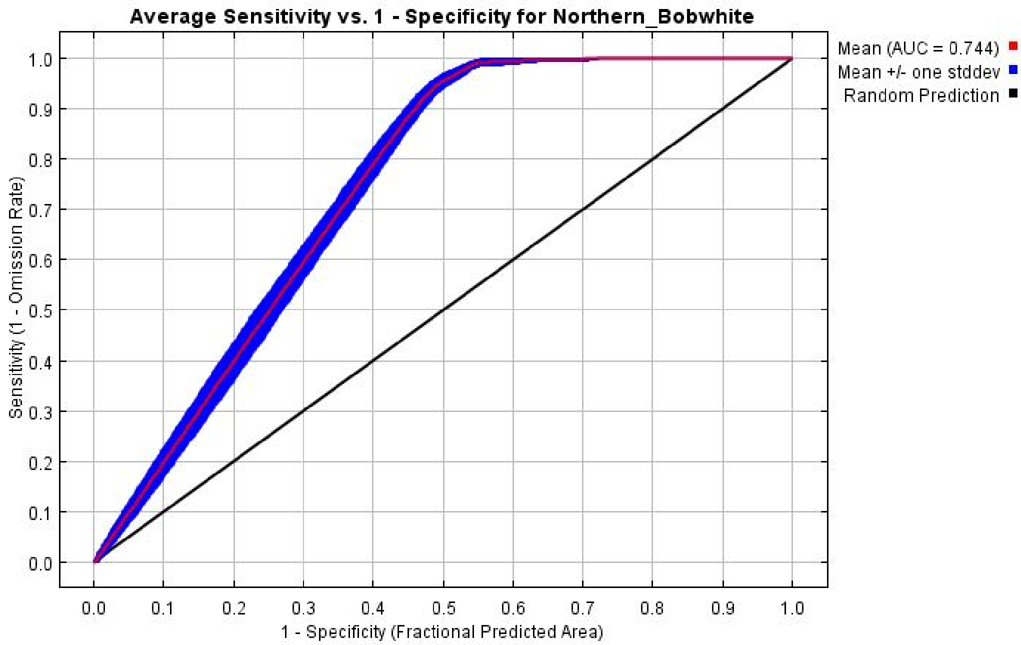


Figure 13. The average AUC for the 100-replicates model.

Another statistical evaluation includes the test omission rate and predicted area as a function of the cumulative threshold that was averaged over the replicate runs. Maxent's diagnostics generate a prediction threshold, predicted values above the threshold are considered to be suitable habitat, while predicted values below the threshold are considered to be unsuitable habitat. Figure 14 shows the predicted area and replicated rates of omission of test samples compared to Maxent's randomly generated prediction omission. The omission rate of test samples should be very close to the predicted omission (Lozar et al. 2018). For this model, the line that represents the omission test data is relatively close to the predicted omission line, indicating that the model performed fairly well.

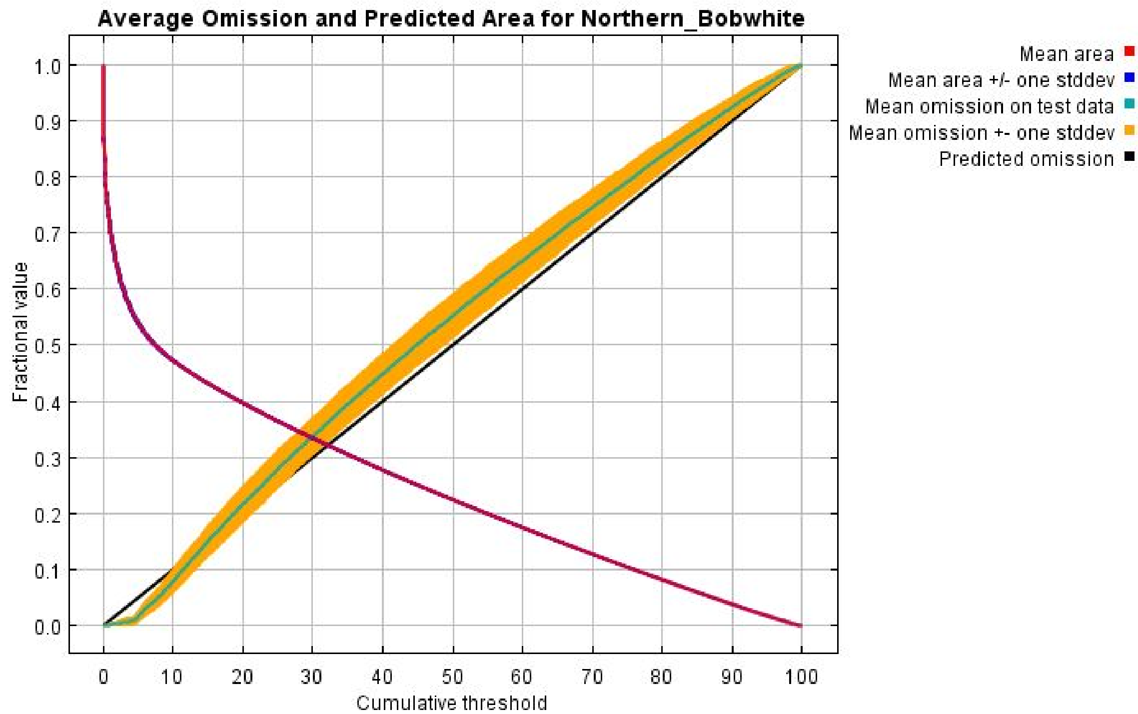


Figure 14. The omission rate and predicted area as a function of the cumulative threshold.

3.2 Variable Contribution

Maxent provides two methods to determine the contribution of the environmental variables to the model: 1) an analysis of variable contributions and 2) a jackknife of regularized training gain (Phillips 2017). While the model is being trained, Maxent keeps track of which environmental variables are contributing most to the model. At the end of the training, Maxent converts the contribution values to percentages and produces an analysis of variable contribution table (Table 7) (Phillips 2017). Variable importance is measured by percent contribution and permutation importance. Percent contribution is determined based on the path variables follow in reaching convergence (Phillips 2017). Simply, percent contribution displays the order in which variables were introduced to the model. Permutation importance is determined by randomly permuting the values for presence and background samples for each variable while holding all other variables constant (Jarnevich et al. 2016). Variables with the highest predictive contribution percentage have the greatest impact on predicting suitable bobwhite habitat. The three most important variables that contributed to the model include: 1) deciduous 1.8km, 2) water 1.8km, and 3) developed 1.8km. Specifically, deciduous 1.8km, water 1.8km, and developed 1.8km contributed 32.6%, 22.8%, and 10.1%, respectively. Noticeably, these results were very similar to the results of the preliminary model. Although the analysis shows that Maxent used the deciduous 1.8km variable more than the others, it doesn’t suggest that deciduous 1.8km is more important to northern bobwhites than the other variables. The contribution of the other variables to the model was relatively minor with predictive contribution percentages between 1.1% and 9.8%. According to the permutation importance, deciduous 1.8km and water 1.8km were high predictors, ranking one and two, respectively.

The jackknife of regularized training gain is useful for determining which variables contribute most to the model (Phillips 2017). That is, it is important for determining which environmental variable has the most information for predicting suitable habitat for northern bobwhites. Multiple models are created when Maxent is running. Initially, each variable is excluded in succession so a model is created with the remaining variables. Afterwards, a model is created using each variable separately. Lastly, a model is created using all of the variables. The results of the jackknife analysis are provided as a bar chart (Figure 15). The red bar at the bottom of the chart represents the model that used all of the variables. The blue bars represent the regularized training gain when each variable was used separately. The green bars represent the regularized training gain when variables were excluded. The variable with the highest gain when used separately ("with only variable") is deciduous 1.8km. This suggests that deciduous 1.8km has the most useful information by itself. The next two variables in order of importance are elevation and developed 1.8km. Although the analysis of variable contributions and jackknife test agree that deciduous 1.8km is the most important variable, they disagree regarding the other variables. The variable that decreases the gain the most when it is excluded ("without variable") is deciduous 1.8km. This suggests that deciduous 1.8km has the most information that is not present in the other variables.

Table 5. Analysis of variable contribution

Variable	Percent contribution	Permutation importance
deciduous1.8km	32.6	29.9
water1.8km	22.8	25.7
developed1.8km	10.1	2
elevation	9.8	11.6
mixed1.8km	5.5	7.9
barren1.8km	3.7	7.2
nlcd2016	3.6	0.4
wetland	2.7	1.6
shrubscrub1.8km	2.6	2
grasslandherbaceous1.8km	2.6	2.2
evergreen1.8km.	1.6	4.9
agriculture1.8km	1.2	2
wetland400m	1.1	2.6

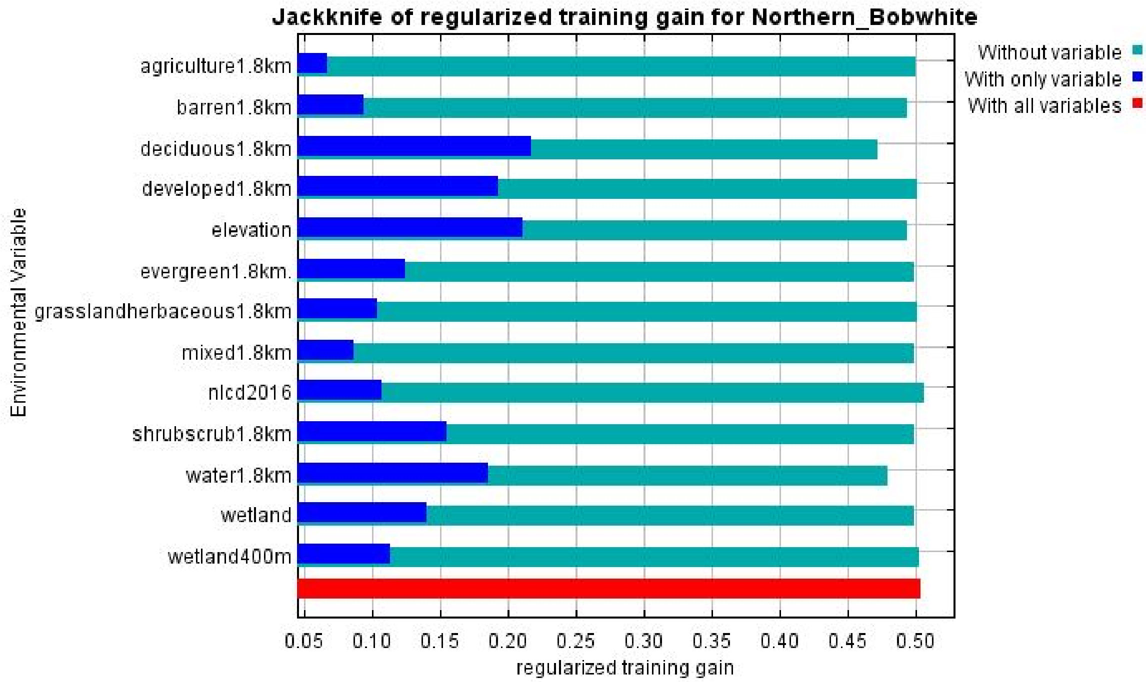


Figure 15. Results of the jackknife analysis for northern bobwhite habitat suitability.

3.3 Response Curves

Maxent produces two sets of response curves for the environmental variables. For each set of curves, Maxent will produce one graph for each environmental variable. Because 13 environmental variables were used for this model, Maxent produced 13 graphs for each set of curves. The first set of curves are referred to as marginal curves which illustrate how the model prediction changes as the value of each environmental variable changes as all other variables remain at their average values. Marginal response curves, however, can be misleading if environmental variables are correlated (Phillips 2017). Therefore, Maxent produces a second set of response curves that are different than the marginal curves. Although both set of response curves may look similar, unlike the marginal curves, the second set of curves were produced using different models. Each curve is a different model that was produced using only one environmental variable while excluding all of the other variables (Phillips 2017). The second set of response curves was used to determine predicted suitability of each environmental variable. The Y-axis for each of the response curves is the predicted probability of presence expressed by the logistic output format. The X-axis includes the number of cells occupied by the environmental variable. The blue line represents the standard deviation.

Figure 16 includes the second set of response curves that were created to represent Maxent's prediction using only that environmental variable. The contribution of the environmental variables varied greatly between models. The environmental variables that contributed most to the final model include, deciduous 1.8km, water 1.8km, and developed 1.8km, respectively. The three variables combined contributed 65.5% to the model. Wetland 400m and agriculture 1.8km contributed least to the final model, respectively. Deciduous 1.8km increased the likelihood of northern bobwhite habitat suitability within 4600 cells (Figure 17a). Cell size is 30m x 30m (0.0009 km2), therefore, Maxent predicted suitable habitat within 4.14 km2 (0.0009 x 4600 = 4.14 km2) or 414 hectares (1023 acres). After 414 hectares, deciduous 1.8km no longer increases the likelihood of bobwhite habitat. The response curve for water 1.8km illustrates that the highest likelihood of suitable habitat is predicted to be at 1600 cells (1.44 km2) or 144 hectares (355 acres), with a sharp decline beyond 1600 cells of water within the 1.8 km neighborhood (Figure 17b). According to the developed 1.8km variable, bobwhite habitat is most likely to occur in the lowest developed areas within the 1.8km neighborhood (Figure 17c). As expected, habitat suitability remains low as developed 1.8km increases.

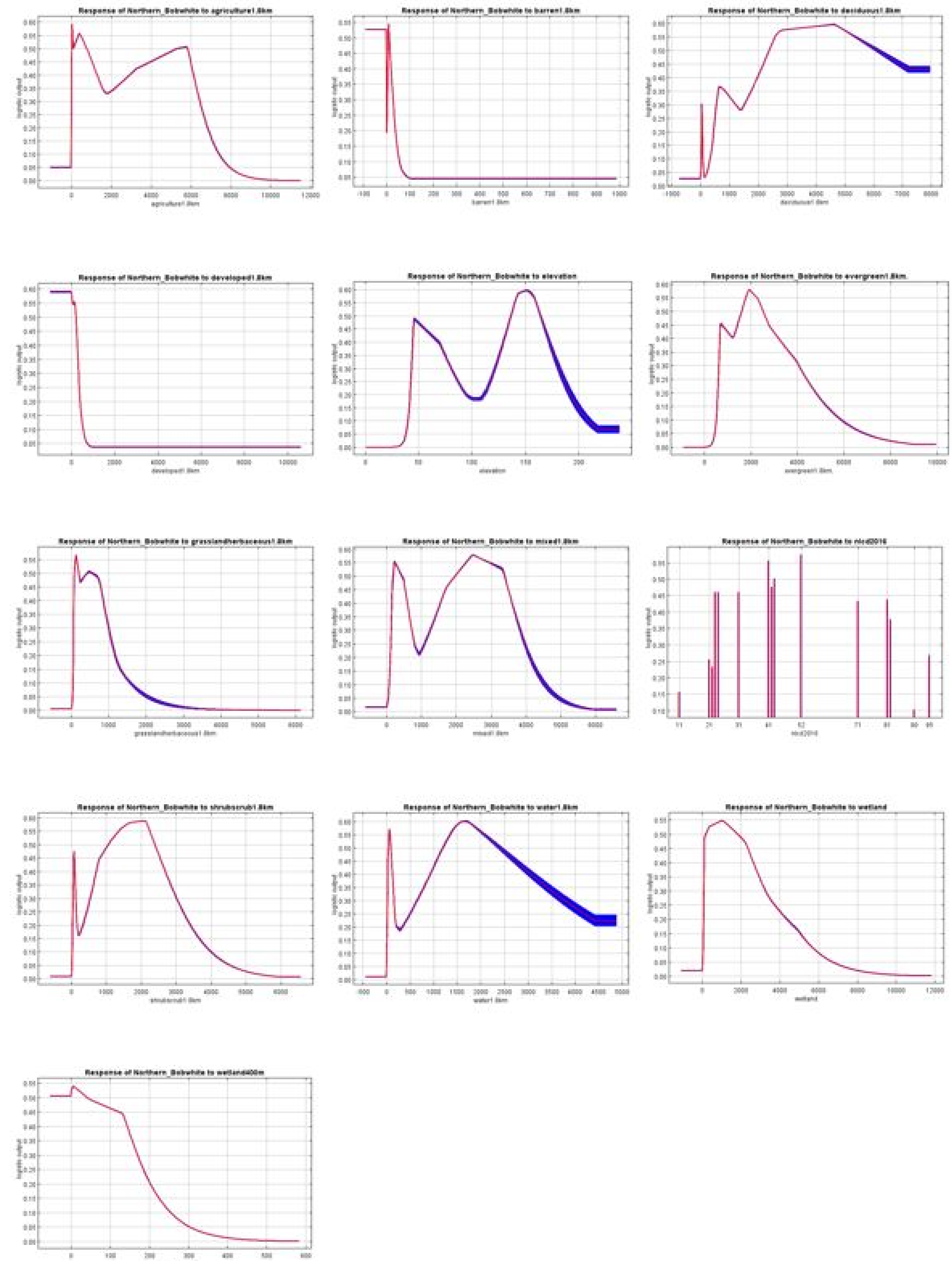


Figure 16. Environmental variable response curves for the final northern bobwhite model.

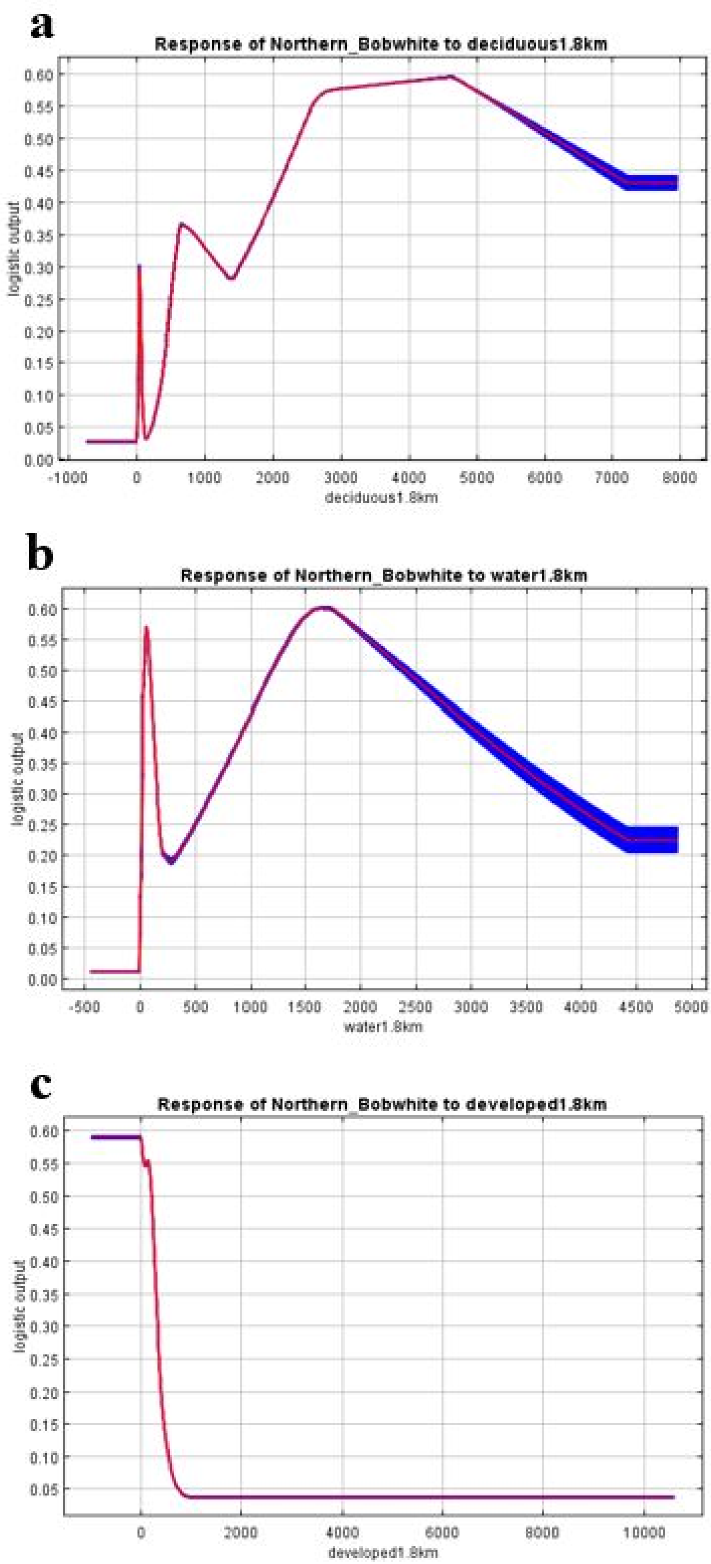


Figure 17. Environmental variable response curves that contributed most to the model are a) deciduous 1.8km, b) water 1.8km, and c) developed 1.8km.

3.4 Likelihood Distribution

The initial output of Maxent is an html file that includes two png images of the predicted habitat. The two images, a mean and standard deviation of the output grids, are rated on a continuous scale between 0.0 and 1.0. Values that are closer to 0 indicate less potential suitable habitat while values that are closer to 1 indicate higher potential suitability. Although Maxent produces results in a png format, the images were converted to raster images using the conversion tool in ArcGIS for improved visualization and evaluation. Finally, the symbology of the raster was changed to create a map of the likelihood of bobwhite habitat in the Black Belt Prairie physiographic region (Figure 18). In addition to the numerical values, the dark blue color indicates a low probability of suitable habitat while red indicates a high probability of suitable habitat.

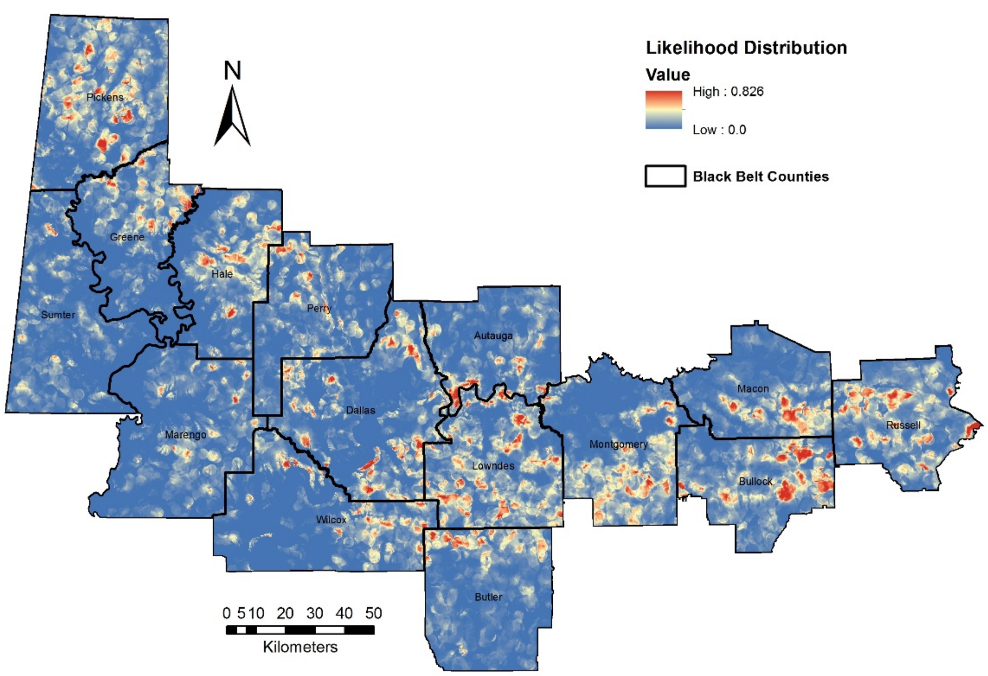


Figure 18. The likelihood distribution of suitable habitat.

Maxent automatically calculates predicted areas and training omission rates for cumulative and logistic thresholds. Each threshold was used to obtain a binary (suitable vs unsuitable) prediction for bobwhite habitat suitability. Table 6 includes the predicted areas and testing omission rates using different threshold levels. Table labels include:

- Thresholds - (cumulative and logistic) used for the model criteria
- Description - briefly describes the model criteria
- Predicted area - the proportion of the total area predicted to be suitable habitat
- Training omission rate - the rate of failure to predict suitable habitat where it is known to occur

The training omission rates for all model criteria were relatively low except for 0.3287. The relatively low training omission rates indicate acceptable results (Lozar et al. 2018).

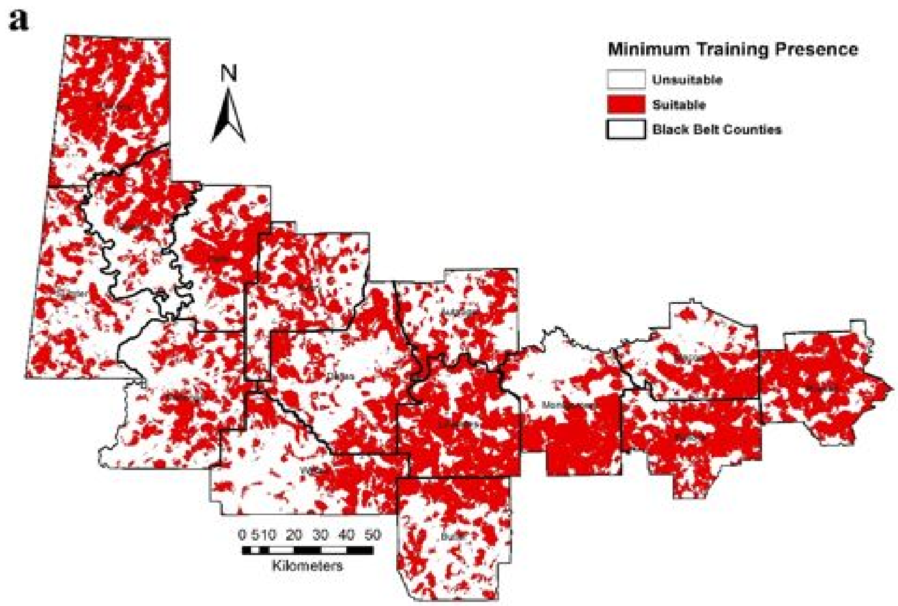
Table 6. Predicted areas and testing omission rates using different threshold methods.

Cumulative Threshold	Logistic Threshold	Description	Predicted Area	Training Omission Rate
1	0.0604	Fixed cumulative value 1	0.71	0.0018
5	0.2292	Fixed cumulative value 5	0.5397	0.016
10	0.3942	Fixed cumulative value 10	0.4726	0.08
0.7119	0.0483	Minimum training presence	0.7409	0
11.4047	0.4176	10 percentile training presence	0.4601	0.1
29.8104	0.5057	Equal training sensitivity and specificity	0.3356	0.3356
7.6946	0.3331	Maximum training sensitivity plus specificity	0.4969	0.0484
3.6407	0.17	Balance training omission, predicted area, and threshold value	0.5741	0.0076
2.6959	0.1281	Equate entropy of thresholded and original distributions	0.6061	0.0062

It is often beneficial in conservation planning to develop binary models of habitat distributions (Abade et al. 2014). A binary output can allow for a quick assessment of habitat suitability. Consequently, the continuous results (i.e. likelihood distribution) were converted into a binary distribution by selecting the appropriate threshold. For this study, a threshold is the minimum level at which habitat is predicted to be suitable. Predicted values above the threshold are considered to be suitable habitat, whereas predicted values below the threshold are considered to be unsuitable habitat. Thresholds are selected to provide a preferred balance between omission and commission errors (Hernandez et al. 2006; Fielding and Bell 1997). However, this is a rather arbitrary means of selecting a threshold. Unfortunately, selecting an appropriate threshold is not clear when using presence-only data (Pearson et al. 2007).

Threshold selection is a possible bias in modeling (Phillips and Dudik 2008; Bean et al. 2012; Syfert et al. 2013; Nenzen and Araugo 2011). Liu et al. (2013) evaluated the suitability of 13 threshold selection methods using presence only data for simulated species. The researchers found that the maximum training sensitivity plus specificity method was promising for presence only data. However, the study results contrasted with Norris (2014) that maximum training sensitivity plus specificity resulted in the greatest omission error and increased loss of suitable areas among all the predicted distributions of lowland tapir. Norris found that the minimum training presence and fixed cumulative value 1 methods had the lowest threshold values and near zero omission errors; therefore, these methods were selected as the most appropriate to identify suitable and unsuitable areas for lowland tapir. According to Glover-Kapfer (2015), the maximum test sensitivity plus specificity logistic threshold should be used rather than the commonly recommended maximum training sensitivity plus specificity logistic threshold.

The area of suitable habitat was compared using all nine of the threshold selection methods produced by Maxent (Table 6). Four of the methods considerably under predicted habitat suitability: maximum training sensitivity plus specificity, equal training sensitivity and specificity, 10 percentile training presence, and fixed cumulative value 10. Although arbitrary, the decision to consider these selection methods inappropriate was based on knowledge of the study area. Three of the methods considerably under predicted habitat suitability along the perimeter of multiple counties in the study area: equate entropy of thresholded and original distributions, balance training omission, predicted area, and threshold value, and fixed cumulative value 5. The decision to consider these selection methods inappropriate for use was based on knowledge of the study area. Although arbitrary, these methods clearly missed predicting a considerable amount of suitable habitat. Minimum training presence (MTP) and fixed cumulative value 1 (FCV1) were likely the most appropriate methods for selecting a threshold (Figure 19). MTP and FCV1 had a logistic threshold of 0.0483 and 0.0604, respectively. The training omission rate for MTP and FCV1 was 0 and 0.0018, respectively. For MTP, 15 of the presence locations were below the logistic threshold of 0.0483, whereas 79 of the 22,977 presence locations were below the logistic threshold of 0.0604 for FCV1. The binary models indicate that MTP comprised 53% of the study area while FCV1 comprised 47% (Table 7). Determining an appropriate threshold selection method for this study was more arbitrary than science. It involved a subjective examination of model predictions to determine if they make sense relative to northern bobwhite habitat ecology and knowledge of the study area.



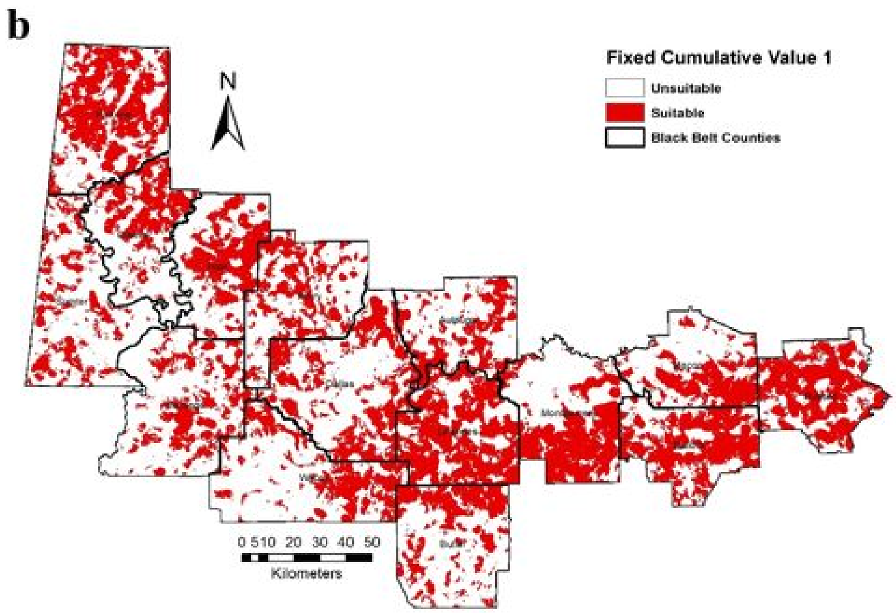


Figure 19. Binary distributions produced by a) MTP and b) FCV1 threshold methods.

Table 7. Area and percentage of study area predicted as suitable habitat using the MTP and FCV1 threshold selection methods.

Threshold Method	Unsuitable	Suitable	% study area (suitable)
MTP	1,581,600 ha	1,399,800 ha	53%
	3,908,218 ac	3,458,981 ac	
FCV1	1,758,200 ha	1,223,100 ha	47%
	4,344,606 ac	3,022,345 ac	

Chapter 4: Discussion

4.1 Implications for Variable Contribution

Maxent determines which variables make the greatest contribution to the model by conducting an analysis of variable contributions and a jackknife analysis. The analysis of variable contributions determined deciduous 1.8km, water 1.8km, and developed 1.8km, respectively, to be the variables that had the highest predictive contribution percentage, thus having the greatest impact on predicting suitable northern bobwhite habitat. Although Maxent used the deciduous 1.8km variable more than the others, it doesn't suggest that it is more important to bobwhites than the other variables; certainly, there are other variables that are far more important for bobwhite survival and reproductive success. As expected, predicted suitable bobwhite habitat remained consistently low as developed 1.8km increased. The contribution of other variables to the model was minor with relatively low contribution percentages. Because the majority of presence locations were taken in evergreen land cover, it is a surprise that evergreen 1.8km contributed only 1.6 percent to the model. Furthermore, the northern bobwhite is adapted to early successional habitat that includes an abundance of herbaceous ground cover and shrubs; consequently, the contribution percentages of shrubscrub 1.8km, evergreen 1.8km, and grasslandherbaceous 1.8km were expected to be significantly higher.

According to the jackknife analysis, deciduous 1.8km is a strong predictor of suitable bobwhite habitat as it had the highest gain when used separately ("with only variable"). This suggests that deciduous 1.8km has the most useful information by itself. Furthermore, deciduous 1.8km decreases the gain the most when it is excluded ("without variable"), which suggests that it has the most information that is not present in the other variables. Although the analysis of variable contributions and jackknife test agree that deciduous 1.8km is the most important variable, they disagree regarding the other variables. For example, elevation was the second most important variable in the jackknife analysis, while it contributed only 9.8 percent according to the variable contribution analysis. Although the results of this study indicate that Maxent can be used to predict suitable bobwhite habitat at a landscape scale, model performance could potentially be improved by reducing the number of variables to include only those variables that satisfy the strictest ecological niche of northern bobwhites.

4.2 Implications for Response Curves

The response curves agreed with the variable contribution analysis that deciduous 1.8km, water 1.8km, and developed 1.8km, respectively, contributed most to the model. An initial view of the deciduous 1.8km response curve may seem questionable as it predicts an increase in likelihood of suitable bobwhite habitat within 4,600 cells. However, as expected, the likelihood of suitable bobwhite habitat declined beyond 4,600 cells. For example, bobwhites consume and benefit from mast produced by deciduous trees; however, deciduous forests are not essential for bobwhite reproduction and survival. Consequently, bobwhites will avoid relatively large deciduous habitats, instead, they utilize relatively small patches of deciduous habitats that are interspersed into a matrix of evergreen forests, shrubs, and grassland habitats. Similarly, an initial view of the water 1.8km response curve may seem questionable as it predicts an increase in likelihood of suitable bobwhite habitat between 250 and 1,600 cells; however, as expected, habitat suitability sharply declines beyond 1,600 cells. For example, bobwhites in Alabama achieve their daily water requirements from their diet (insects, vegetation, and fruits); however, free-standing water is important during periods of extreme drought. Because of Alabama's humid subtropical climate (Beck et al. 2018), water variables could potentially be excluded from the model – unlike west Texas where the climate is dry and free-standing water is critical for bobwhite survival. As expected, the likelihood of suitable bobwhite habitat is relatively high when developed 1.8km is low, and a consistently low likelihood in suitability as developed 1.8km increases. Similarly, the likelihood of suitable habitat is relatively high when barren 1.8km is low, and a consistently low likelihood in suitability as barren 1.8km increases. Surprisingly, the evergreen 1.8km response curve shows a sharp decline in habitat suitability as evergreen 1.8km increases. This is a surprise because evergreen forests (upland forests) provide the best opportunity for bobwhite management.

4.3 Implications for Likelihood Distribution

An initial inspection of the continuous likelihood distribution map indicates that known areas of stable or increasing bobwhite populations were accurately predicted to be high. Although Maxent was not used in this study to predict occupancy, one would expect Maxent to predict a high likelihood distribution for areas that are known to have stable or increasing populations because bobwhite populations are closely associated with habitat suitability. However, upon close inspection of the map, some areas with a relatively high likelihood distribution are obviously poorly suited for bobwhites; this discrepancy was determined by using aerial imagery and site inspections. For example, a relatively large area in east-central Russell County has a high likelihood distribution; however, the area is a flooded wetland. Maxent may have under predicted the likelihood distribution along the outer edges of the study area. If so, a possible explanation could be due to model overfitting. That is, the predicted distributions are clustered near presence locations leaving the outer edges under predicted. The default setting of one was used for the regularization multiplier; however, adjusting the setting could potentially reduce model overfitting.

Minimum training presence (MTP) and fixed cumulative value 1 (FCV1) are likely the most appropriate methods for selecting a threshold to obtain a binary (suitable vs unsuitable) prediction for bobwhite habitat suitability. For this study, a threshold is the minimum level at which habitat is predicted to be suitable. It is often beneficial in conservation planning to develop binary models (Abade et al. 2014); however, determining an appropriate threshold method for this study was more arbitrary than science. As with the continuous likelihood distribution, the MTP and FCV1 binary maps indicate potentially under predicted areas along the outer edges of the study area. Furthermore, with the MTP and FCV1 thresholds showing 53 percent and 47 percent, respectively, of the study area to be suitable for bobwhites, undue confidence may be placed in the predictions. Choosing biologically meaningful thresholds may depend on prevalence or population density (Merow et al. 2013), which is unknown in this study.

Therefore, thresholds based on arbitrary selection may have little utility in predicting bobwhite habitat suitability on a landscape scale.

4.4 Implications for Management

Understanding spatial distributions of habitats is critical to management and conservation planning of wildlife populations (De Knegt et al. 2011). This research suggests that Maxent can be used to understand the spatial distribution of potential suitable habitat for northern bobwhites. Using the predictive capability of Maxent, this research will allow conservation planners to focus their work in areas of the Black Belt Prairie physiographic region that have a high likelihood of supporting bobwhite populations. By focusing on areas with a high likelihood of a bobwhite population response, managers can reduce the possibility of investing management efforts in lower-quality areas. Furthermore, the identification of high-quality areas based on the ecological requirements of northern bobwhites provide areas for potential reintroduction efforts, and public outreach and education. The likelihood distribution model can be used to correlate high-quality areas with land ownership (public vs. private) for a more precise delivery of outreach and management efforts. Lastly, this model can be used as a monitoring tool to determine changes in the spatial distribution of habitat over time caused by changes in land use patterns and habitat management.

Chapter 5: Conclusion

This research concluded that Maxent can be used to model suitable habitat for northern bobwhites by using presence-only data and environmental variables. This model's AUC, the most commonly used metric to determine model quality, indicates a "good" and useful model (Swets 1988). Therefore, this model is useful for predicting habitat suitability for northern bobwhites in the Black Belt Prairie physiographic region. There is widespread disparity regarding the use of default settings versus model tuning. Because of the disparity and lack of information on using Maxent to predict bobwhite habitat suitability, the default settings were largely used for this study. Consequently, this approach produced a useful model for predicting suitable bobwhite habitat. However, model performance could potentially be improved by tuning certain parameters.

Literature Cited

Abade, L., D. W. Macdonald, and A. J. Dickman. 2014. Using landscape and bioclimatic features to predict the distribution of lions, leopards and spotted hyaenas in Tanzania's Ruaha landscape. PLoS One 9:e96261. doi: 10.1371/journal.pone.0096261.

Anderson, R. P., and E. Martinez-Meyer. 2004. Modeling species' geographic distributions for preliminary conservation assessments: an implementation with the spiny pocket mice (*Heteromys*) of Ecuador. Biological Conservation. 116:167-179.

Baldwin, R.A. 2009. Use of Maximum Entropy Modeling in Wildlife Research. *Entropy* 11:854-866.

Barone, J.A. 2005. Historical presence and distribution of prairies in the Black Belt of Mississippi and Alabama. Castanea 70:170-183.

Bean, W. T., Stafford, R. and Brashares, J. S. 2012. The effects of small sample size and sample bias on threshold selection and accuracy assessment of species distribution models. Ecography 35:250-258.

Beck, H. E., N. E. Zimmermann, T. R. McVicar, N. Vergopolan, A. Berg, E. F. Wood. 2018. Present and future Koppen-Geiger climate classification maps at 1-km resolution. Scientific Data 5:180214, doi:10.1038/sdata.2018.214.

Brennan, L. A. 1991. How can we reverse the northern bobwhite decline? Wildlife Society Bulletin 19:544–555.

Brennan, L.A., R. T. Engstrom, W. E. Palmer, S. M. Hermann, G. A. Hurst, L. W. Burger, and C. L. Hardy. 1998. Whither wildlife without fire? Transactions North American Wildlife and Natural Resources Conference 63:1–11.

Brotons, L., W. Thuiller, M. Arauja, and A. H. Hirzel. 2004. Presence-absence versus presence only modeling methods for predicting bird habitat suitability. Ecography 27:437-448.

Cianfrani, C., G. Le Lay, A. H. Hirzel, and A. Loy. 2010. Do habitat suitability models reliably predict the recovery areas of threatened species? Journal of Applied Ecology 47:421-430.

De Knegt, H. J., F. Van Langevelde, A. K. Skidmore, A. Delsink, R. Slotow, S. Henley, G. Bucini, W. F. De Boer, M. B. Coughenour, C. C. Grant, I. M. A. Heitkönig, M. Henley, N.M. Knox, E. M. Kohi, E. Mwakiwa, B. R. Page, M. Peel, Y. Pretorius, S. E. Van Wieren, and H. H. T. Prins. 2011. The spatial scaling of habitat selection by African elephants. Journal of Animal Ecology 80:270-281.

Dimmick, R. W. 1992. Northern Bobwhite (*Colinus virginianus*): Section 4.1.3, US Army Corps of Engineers Wildlife Resources Management Manual, Technical Report EL-92-18, US Army Engineer Waterways Experiment Station, Vicksburg, Miss.

Dudik, M., S. J. Phillips, and R. E. Schapire. 2005. Correcting sample selection bias in maximum entropy density estimation. Pages 323-330 in Advances in neural information processing systems 18. MIT Press, Cambridge, Massachusetts, USA.

Elith, J. 2002. Quantitative methods for modeling species habitat: Comparative performance and an application to Australian plants. In: S. Ferson and M. Burgman (eds.), *Quantitative Methods for Conservation Biology*. Springer, pp 39-58.

Elith, J., C. H. Graham, R. P. Anderson, M. Dudik, S. Ferrier, A. Guisan, R. J. Hijmans, F. Huettmann, J. R. Leathwick, A. Lehmann, J. Li, L. G. Lohmann, B. A. Loiselle, G. Manion, C. Moritz, M. Nakamura, Y. Nakazawa, J. M. Overton, A. T. Peterson, S. J. Phillips, K. S. Richardson, R. Scachetti-Pereira, R. E. Schapire, J. Soberon, S. Williams, M. S. Wisz, N. E. Zimmermann. 2006. Novel methods improve prediction of species' distributions from occurrence data. Ecography 29:129-151.

Elith, J., S. J. Phillips, T. Hastie, M. Dudik, Y. E. Chee, and C. J. Yates. 2011. A statistical explanation of MaxEnt for ecologists. Diversity and Distributions 17:43-57.

Exum, J. H., R. W. Dimmick, and B. L. Deardon. 1982. Land use and bobwhite populations in an agricultural system in west Tennessee. Proceedings of the National Quail Symposium 2:6–12.

Ferrier, S., G. Watson, J. Pearce, and M. Drielsma. 2002. Extended statistical approaches to modelling spatial pattern in biodiversity in northeast New South Wales. 1. Species-level modelling. Biodiversity and Conservation 11:2275-2307.

Fielding, A.H., Bell, J.F., 1997. A review of methods for the assessment of prediction errors in conservation presence/absence models. Environ. Conserv. 24, 38-49.

Fuller, R. J. 2012. Birds and Habitat: Relationships in changing landscapes. University Press, Cambridge, UK.

Fies, M. L., I. L. Kenyon, and J. V. Gwynn. 1992. Effects of changing land use patterns on bobwhite quail habitat in Virginia. Virginia Journal of Science 23:143–155.

Friedlaender, A.R., Johnston, D.W., Fraser, W.R., Burns, J., Halpin, P.N., Costa, D.P. 2011. Ecological niche modeling of sympatric krill predators around Marguerite Bay, Western Antarctic Peninsula. Deep-Sea Res II 58:1729-1740.
Glover-Kapfer, P. 2015. A training manual for habitat suitability and connectivity modeling using tigers (*Panthera tigris*) in Bhutan as example. 10.13140/RG.2.2.34804.86409.

Grahma, C. H. and R. J. Hijmans. 2006. A comparison of methods for mapping species ranges and species richness. Global Ecology and Biogeography 15:578-587.

Guthery, F. S. 1997. A philosophy of habitat management for northern bobwhite. Journal of Wildlife Management 61: 291–301.

Ha, H., Heumann, B. W., Liesch, M., Wang, X., 2016. Modelling potential conservation easement locations using physical and socio-economic factors: a case-study from south-east Michigan. Appl. Geogr. 75, 104-115.

Hastie, T., and W. Fithian. 2013. Inference from presence-only data; the ongoing controversy. Ecography 36:864-867.

Hernandez, P. A., Graham, C. H., Master, L. L., Albert, D. L. 2006. The effect of sample size and species characteristics on performance of different species distribution modeling methods. Ecography. 29:773-785.

Heumann, B. W., Walsh, S. J., Verdery, A. M., McDaniel, P. M., Rindfuss, R. R. 2013. Land suitability modelling using a geographic socio-environmental niche-based approach: a case study from Northeastern Thailand. Ann. Assoc. Am. Geogr. 103, 764-784.

Hirzel, A. H., and G. Le Lay. 2008. Habitat suitability modelling and niche theory. Journal of Applied Ecology 45:1372-1381.

Homer, C.G., Dewitz, J.A., Yang, L., Jin, S., Danielson, P., Xian, G., Coulston, J., Herold, N.D., Wickham, J.D., and Megown, K., 2015, Completion of the 2011 National Land Cover Database for the conterminous United States-Representing a decade of land cover change information. *Photogrammetric Engineering and Remote Sensing*, v. 81, no. 5, p. 345-354

Jarnevich, C. S., T. R. Holcombe, B. A. Grisham, J. Timmer, C. W. Boal, M. Butler, J. Pitman, S. Kyle, D. Klute, G. Beauprez, A. Janus, and B. Van Pelt. 2016. Assessing range-wide habitat suitability for the Lesser Prairie-Chicken. Avian Conservation and Ecology 11(1):2.

Kabat, C., and D. R. Thompson. 1963. Wisconsin quail, 1834– 1962: population dynamics and habitat management. Wisconsin Conservation Department Technical Bulletin 30. Madison.

Kampichler, C., R. Wieland, S. Calme, H. Weissenberger, and S. Arriaga-Weiss. 2010. Classification in conservation biology: a comparison of five machine-learning methods. Ecological Informatics 5:441-450.

Kearney, M. 2006. Habitat, environment and niche: what are we modelling? Oikos 115:186-191.

Liu, C., White, M. and Newell, G. 2013. Selecting thresholds for the prediction of species occurrence with presence-only data. Journal of Biogeography 40:778-789.

Lozar, Robert C., Scott A. Tweddale, Charles R. Ehlschlaeger, Carey L. Baxter, and Jeffrey A. Burkhalter. 2018. Testing Maximum Entropy Analysis to Define Population Distributions. ERDC/CERL TR-18-22.

L. Wes Burger. 2001. Quail management: issues, concerns, and solutions for public and private lands—a southeastern perspective.

Mark D. Smith and Loren W. Burger Jr. “Multiresolution approach to wildlife habitat modeling using remotely sensed imagery”, Proc. SPIE 5153, Ecosystems’ Dynamics, Agricultural Remote Sensing and Modeling, and Site-Specific Agriculture, (22 December 2003); https://doi.org/10.1117/12.506409.

Merow, C., M. J. Smith, and J. A. Silander, Jr. 2013. A practical guide to MaxEnt for modeling species’ distributions: what it does, and why inputs and settings matter. Ecography 36:1-12.

Nenzen, H. K. and Araujo, M. 2011. Choice of threshold alters projections of species range shifts under climate change. Ecological Modelling 222:3346-3354.

Norris, D. 2014. Model thresholds are more important than presence location type: Understanding the distribution of lowland tapir (*Tapirus terrestris*) in a continuous Atlantic forest of southeast Brazil. Tropical Conservation Science Vol.7(3):529-547.

Noss, R. F., E.T. LaRoe III, and J.M. Scott. 1995. Endangered ecosystems of the United States: A preliminary assessment of loss and degradation. United States Department of Interior, National Biological Service, Biological Report 28, Washington, DC.

Pearce, J. L., and M. S. Boyce. 2006. Modelling distribution and abundance with presence-only data. Journal of Applied Ecology 43:405-412.

Pearce, J., and S. Ferrier. 2000. Evaluating the predictive performance of habitat models developed using logistic regression. *Ecological Modelling.* 133:225-245.

Pearson, R. G., Raxworthy, C. J., Nakamura, M., Peterson, A. T. 2007. Predicting species distributions from small numbers of occurrence records: a test case using cryptic geckos in Madagascar. J. Biogeography 34:102-117.

Phillips, S. J. 2017. A brief tutorial on Maxent. Available from url: http://biodiversityinformatics.amnh.org/open_source/maxent/. Accessed on January 11, 2021.

Phillips, S. J., Dudik, M. 2008. Modeling of species distributions with Maxent: New extensions and a comprehensive evaluation. Ecography 31, 161-175.
Phillips, S. J., R. P. Anderson, and R. E. Schapire. 2006. Maximum entropy modeling of species geographic distributions. Ecological Modeling. 190:231-259.

Phillips, S. J., M. Dudik, J. Elith, C. H. Graham, A. Lehmann, J. Leathwick, and S. Ferrier. 2009. Sample selection bias and presence-only distribution models: implications for background and pseudo-absence data. Ecological Applications. 19: 181-197.

Roseberry, J. L. 1993. Bobwhite and the ‘‘new’’ biology. Proceedings of the National Quail Symposium 3:16–20.
Roseberry, B. G. Peterjohn, and W. D. Klimstra. 1979. Dynamics of an unexploited bobwhite population in deteriorating habitat. Journal of Wildlife Management 43:306–315.

Sauer, J.R., D.K. Niven, J.E. Hines, D.J. Ziolkowski, Jr, K.L. Pardieck, J.E. Fallon, and W.A. Link. 2017. The North American Breeding Bird Survey, Results and Analysis 1966-2015. Version 2.07.2017 USGS Patuxent Wildlife Research Center, Laurel, MD.

Sauer, J. R., J. E. Hines, I. Thomas, J. Fallon, and G. Gough. 2000. The North American Breeding Bird Survey, Results and Analysis 1966–1999. Version 98.1, United States Geological Survey, Patuxent Wildlife Research Center, Laurel, Maryland.

Schairer, G. L., R. H. Wynne, M. L. Fies, S. D. Klopfer. 1999. Predicting Landscape Quality for Northern Bobwhite from Classified Landsat Imagery. Proceedings of the Southeastern Association of Fish and Wildlife Agencies 53:243-256.

Schotz, A.R. and M.S. Barbour. 2009. Ecological assessment and terrestrial vertebrate surveys for Black Belt Prairies in Alabama. Unpublished report submitted to the Alabama Department of Conservation and Natural Resources, Division of Wildlife & Freshwater Fisheries, Montgomery, Alabama. Alabama Natural Heritage Program, Auburn, Alabama. 139 pages. (report available at: http://www.alnhp.org/reports/Prairie_SWG_Final_Report.PDF)

Stoddard, H. L. 1931. The Bobwhite Quail: Its Habits, Preservation, and Increase. Scribner's, New York.

Swets, J. A. 1988. Measuring the accuracy of diagnostic systems. *Science* 240(4857):1285-1293.

Syfert, M. M., Smith, M. J. and Coomes, D. A. 2013. The effects of sampling bias and model complexity on the predictive performance on maxent species distribution models. Plos One 8:e55158.

The National Bobwhite Technical Committee. 2011. Palmer, W.E., T.M. Terhune, and D.F. McKenzie (eds). The National Bobwhite Conservation Initiative: A range-wide plan for re-covering bobwhites. National Bobwhite Technical Committee Technical Publication, ver. 2.0, Knoxville, TN.

Vance, D. R. 1976. Changes in land-use and wildlife populations in southeastern Illinois. Wildlife Society Bulletin 4:11–15.

Wang, Y. H., K. C. Yang, C. L. Bridgman, and L. K. Lin. 2008. Habitat suitability modelling to correlate gene flow with landscape connectivity. Landscape Ecology.

Zimmermann, N. E., T. C. Edwards, C. H. Graham, P. B. Pearman, and J. C. Svenning. 2010. New trends in species distribution modelling. Ecography 33:985-989.