Multi-scale Habitat Relationships of Grassland Vertebrates in the East Gulf Coastal Plain and Understanding the Consequences of Error in Broad-scale Land Cover Data

by

Kevin W. Ryer

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

Auburn, Alabama
August 6, 2016

Keywords: grassland, small mammals, occupancy, error, track-tubes, land cover

Approved by

Conor P. McGowan, Chair, U.S. Geological Survey, Alabama Cooperative Fish and Wildlife
Research Unit Assistant Leader and Associate Research Professor of Wildlife Sciences
Robert A. Gitzen, Assistant Professor of Wildlife Sciences
Craig Guyer, Professor of Biological Sciences

Abstract

Grasslands in the southeastern United States are an ecologically important habitat that has experienced habitat loss and degradation. Conservation efforts have increased the utilization of remotely sensed land cover data to inform management actions of these grasslands. Vertebrate conservation is a main objective of the conservation of grasslands in the region. We used occupancy estimation to determine multi-scale habitat relationships for two small mammal species, the hispid cotton rat (*Sigmodon hispidus*) and oldfield mouse (*Peromyscus polionotus*). We also evaluated the effects of error in remotely sensed land cover data on estimated habitat relationships. Our results indicate broad-scale landscape composition is an important habitat factor to consider for small mammal conservation; however, not accounting for error in land cover data could result in biased parameter estimates.

Acknowledgments

I would like to thank my major professor Dr. Conor McGowan providing this opportunity and his guidance and support during this process. I would also like to thank committee member Dr. Robert Gitzen for his analysis expertise and Dr. Craig Guyer for his guidance and ecological expertise. I need to thank the Gulf Coastal Plain and Ozark Landscape Conservation Cooperative for funding this project, and Chelsea Warner and Nick Sharp for their technical assistance, I also need to thank Orin Robinson for his discussion and comments that improved this thesis, Patrick Farrell for his helpful suggestions at the start of this project, Eric Lonsdorf for assisting with spatial analysis, and everyone who assisted with accessing land. I also need to thank my family and friends for supporting my pursuits.

Table of Contents

List of Tables

List of Figures

Chapter 1:
Habitat Relationships of Grassland-associated Small Mammals in the East Gulf Coastal Plain

ABSTRACT

Conservation and restoration of priority habitats and vertebrate species are major foci of many conservation organizations and agencies. The ability to access broad-scale spatial data from satellites and databases has expanded the scope of many conservation plans to include entire ecoregions. We utilized broad and fine-scale habitat data to create occupancy models to evaluate the relationships between multi-scale habitat features and occupancy of grassland-associated vertebrates in the East Gulf Coastal Plain. We surveyed 57 sites during April-July 2015 for 8 focal species associated with grasslands in the East Gulf Coastal Plain. We used visual encounter surveys to survey for squamates and track-tubes to survey for small mammals. We examined relationships between occupancy and both fine- and broad-scale habitat features. Fine-scale attributes, which were measured at each site, included grass height, ground cover, and tree density. Broad-scale features included proportion of grassland and 4 different classes of forest (e.g. deciduous, evergreen, mixed, woody wetlands) within 200 m, 500 m, 1 km 3km, and 5 km radii around each site. Features measured at these radii were proportion of grassland and forest relative to all other habitat types within the given radii. Network connectivity measurements, such as number of grassland patches and number of links, were also included in the occupancy models. The most frequently detected focal species were the hispid cotton rat (*Sigmodon hispidus*) and the oldfield mouse (*Peromyscus polionotus*). We did not detect any of

the 4 squamate focal species. The hispid cotton rat was positively associated with the proportion of woody wetlands within 5 and 3 km. The oldfield mouse had no strong association with any of the habitat variables we examined. In this study, broad-scale habitat features seemed to possess more weight towards explaining occupancy of the two small mammals than the fine-scale habitat features.

INTRODUCTION

Grasslands of the Gulf Coastal Plains have experienced > 90% area loss since the 1800's due to land conversion, fire suppression, urban expansion, and other human activities (Wilson 1981, Noss et al. 1995, Noss 2013). More specifically the grasslands found in the Black Belt prairie region of Alabama and Mississippi were converted for agricultural use in the late 1800s and early 1900s and now <1% of native grassland remains (Barone 2005). The remaining grassland patches in Alabama and Mississippi are highly fragmented and connectivity between the patches has been greatly affected by the anthropogenic habitat alterations in the surrounding landscape such as agriculture, urbanization and silviculture (Noss 2013). These remnant grassland patches are also currently threatened with degradation due to fire suppression. Without fire as an ecological influence, these grasslands become invaded by the fire intolerant red cedar (*Juniperus virginiana*) and rapidly deteriorate by losing herbaceous cover (Stipe and Bragg 1989). The grasslands found in Alabama and Mississippi are important habitat for some vertebrates including small mammals. Habitat loss and fragmentation are implicated as the cause of drastic declines of some small mammal populations in the southeastern United States (Humphrey and Barbour 1981, Hafner et al. 1998). Exploring the relationships between small mammal populations and broad-scale habitat attributes may provide better insight into how broad-scale conservation efforts will affect small mammal populations.

However, limited information exists on small mammal habitat selection at multiple

spatial scales, especially spatial scales larger than home ranges (Means 2006, Panzacchi et al.

2010). Many studies have focused specifically on fine-scale vegetation conditions to analyze

small mammal habitat selection (Jorgensen 2004). However there have been some studies that

have expanded the spatial scale to investigate other factors such as broad-scale habitat attributes,

and they have found the occupancy of small mammals may be influenced by broad-scale habitat

attributes (Fitzgibbon 1997, Moore and Swihart 2005, Swihart et al. 2007, Urban and Swihart

2009). One broad-scale habitat attribute that could affect occupancy, is habitat fragmentation

which directly affects dispersal by increasing the distance needed to reach suitable habitat and

reducing the proportion of individuals that can disperse, thus potentially decreasing the ability to

occupancy suitable habitat in highly fragmented landscapes (Diffendorfer 1995). Research also

indicates that habitat proportions and patterns, are two important factors that may affect small

mammal population size. In a simulation study conducted by Flather and Bevers (2002) when

large amounts of suitable habitat are available, the amount of habitat can explain 96% of the

variation in population size; however, when the amount of suitable habitat falls below 30%-50%,

the pattern of suitable habitat increases in explaining variation in population size. There seems to

be a need to produce more empirical evidence of the relationships between broad-scale habitat

attributes and small mammals to both increase our understanding of how occupancy is

influenced and provide more specific data for the creation of conservation actions.

One approach to modeling the relationships between habitat and species, which also

incorporates imperfect detection of the species, is occupancy estimation (Mackenzie et al. 2002).

Occupancy probability is the probability that a species is present at a site. However, in most

circumstances the ability to detect a species is imperfect due to most species being

inconspicuous. Detection probability is the probability of a species being detected during a survey given that the species is present. Incorporating detection probabilities into occupancy models has improved occupancy estimates by removing the assumption that a non-detection represents an absence (MacKenzie et al. 2002). We used occupancy estimation to model multi-scale habitat relationships of grassland-associated species.

For this study, we focused on grassland-associated vertebrate population occupancy and habitat associations in the portion of the Gulf Coastal Plain that is east of the Mississippi River known as the East Gulf Coastal Plain (EGCP). We focused on two small mammals, the oldfield mouse (*Peromyscus polionotus*) which is a habitat specialist, restricted to grasslands, and is of moderate conservation concern in Alabama and the hispid cotton rat (*Sigmodon hispidus*, here after cotton rat) which is a habitat generalist, found in a variety of habitats, including grasslands. Although this species is of low direct conservation concern, it is an important prey item for many avian raptors, game species, and large-bodied snakes (Beasom and Moore 1977, Tyler and Jensen 1981, Preston 1990, Baker 1991, Baxley and Qualls 2009). We viewed occupancy of cotton rats to have a positive relationship with occupancy of cotton rat predators. Our objectives were to evaluate four hypotheses related to the conservation and management of small mammals that utilize grassland habitat in the Gulf Coastal Plain. First, we hypothesized that there is a positive relationship between small mammal occupancy and the amount of grassland in the surrounding landscape. Habitat loss is a cause of small mammal population declines and, therefore, an increase in grassland habitat in the surrounding landscape will cause an increase in these populations at any given site (Andrén 1994). Second, there will be a negative relationship between small mammal occupancy and the amount of forest in the surrounding landscape. An increase in the amount of forest around a grassland site will reduce the amount of other less

suitable herbaceous habitats such as shrub and hay/pasture (Clark et al. 1998). Third, small mammal occupancy will be influenced by site-level vegetation conditions. Tall, herbaceous vegetation can reduce avian predation risk and cotton rat occurrence is associated with abundant grass cover (Sheffield et al. 2001, Mengak and Guynn 2003). Finally, connectivity of grassland patches will influence small mammal occupancy. Evaluating these four hypotheses will give conservation planners a better capacity to design broad-scale grassland conservation plans that will benefit grassland-associated small mammals.

STUDY AREA

The study was conducted in portions of the EGCP occurring within the states of Alabama and Mississippi (Figure 1.1). More specifically a majority of the sites were located in the Black Belt Prairie region. The land cover of the EGCP is composed of 56% forest, 12% shrub, 10% cultivated crops, 10% hay/pasture, 7% developed, 4% grassland, and 1% water (Homer et al. 2015). Alabama's and Mississippi's climate is classified a subtropical with an average annual temperature of 65°F and average annual rainfall of 53.05 inches (PRISM Climate Group). The soil in the coastal plain of Alabama and Mississippi is composed of loamy or clayey subsoil and a sandy loam or loam surface layer. The soils of Black Belt Prairie region are composed of mostly acid and alkaline clayey soils. The soils in the Black Belt Prairie region inhibit the survival of many species of trees, allowing prairies to form.

METHODS

Patch and Plot Selection

Grassland patches from the 2011 NLCD that were ≥1 hectare in size were used as a target population of potential survey patches. These patches were separated into two categories based on surrounding 2011 NLCD land cover, <50% and ≥50% forested land cover in a 1-km buffered area. For logistical reasons five polygons with a radius of 50 km were created to subsample from all grassland patches in the study area. We created 50 x 50 m grids that were overlaid on the patches that we had sampled using ArcMap 10.2. We numbered each cell that contained grassland land cover and used a random number generator to select 50 x 50 m plots. If selected plots were < 100 m away from each other, then one plot would be removed and resampled, so that no plots were closer than 100 m from each other. Each patch had two or more plots randomly selected; however, due to logistical constraints and patch size, only four patches had more than three plots surveyed.

Focal Species Selection

The selection of focal species was guided by GCPOLCC "Integrated Science Agenda", literature review and expert elicitation. We evaluated two functional groups for potential focal species (e.g. squamates and small mammals). To select the focal species used in this study we used the natural history of species found in the study area and then identified the species with habitat requirements that could potentially be met with the proposed grassland restoration. Habitat usage and home range size were used to determine habitat requirements and population density was used to assess the probability of detection (Table 1.1). All focal species that were selected for this study are detectable with the proposed survey methods described below.

Vertebrate Surveys

We conducted repeated presence/absence surveys for vertebrates between March and July 2015. We used track-tubes and time-constrained visual encounter surveys to sample for small mammals at each site, visiting each site three times. Use of two different types of survey protocols was necessary since we were surveying for multiple species that have different activity patterns.

Track-tubes can be equally as effective at detecting small mammals as live-trapping methods, but are less expensive, easier to deploy, cause no animal mortality, and eliminate zoonosis exposure (Drennan et al. 1998, Glennon et al. 2002). The track-tubes used in this research were, with minimal modifications, based on the design used by Glennon et al. (2002). The track-tubes were made of 2 pieces of plastic rain gutter (4.5" x 2.5" x 12") duct taped together on one edge to create a hinged tube; 3/8" binder clips were used to attach the other edge and create the tube shape (Figure 1.2). Aluminum flashing (2.9" x 14" x 0.01") was attached to the floor of the tube to allow a flat surface for an animal to walk on. Contact paper (3" x 11") was taped to the aluminum at each end with painters tape. The side of the contact paper with the adhesive faced away from the aluminum so that the animals entering the tube would walk on the side with the adhesive. Felt squares (2" x 3") were placed at each entrance of the tube and served as pads for the ink. The ink consisted of 2 parts carbon powder and 1 part mineral oil; animals entered the tube by walking across the ink-filled felt pad and left foot prints on the adhesive contact paper. Sunflower seeds were used to attract the animals into the tube. At each plot 9 track-tubes were placed on two 50 m transects that were perpendicular to each other and crossed in the center (25 m). The track-tubes were placed in-line with each transect at 0 m, 12.5 m, 37.5 m, and 50 m (Figure 1.3). One track-tube was placed at the center of the two transects. Track-

tubes were checked every 48 hours during an 8-day sampling period. A total of 3 sampling occasions occurred at each plot.

Visual encounter surveys (VES) are a common survey method for squamate species. This survey method is inexpensive and requires less time than common passive survey methods such as drift fences and pitfall traps. We employed a 30 minute time limit for each of the 3 visual encounter surveys at each 50 x 50 m site. We walked in a straight transect parallel to 2 sides of the site and when we would came in contact with the perpendicular sides of the site we would shift over 5 meters turnaround and walk in a straight transect parallel to the last transect (Figure 1.4).

Vegetation Surveys

We measured the vegetative conditions in each plot to assess how these fine-scale parameters are associated with animal occupancy. The vegetation measurements that were taken include grass height, ground coverage, and tree density. These site attributes are defined by the GCPO LCC as criteria for "the desired ecological state" (GCPO LCC. 2014). At each 50 x 50 m plot, two 50 m measuring tapes were placed in the four cardinal directions, with the 25 m mark on each measuring tape crossing the center of the plot. Current-year growth and residual grass heights were measured at 12 points in each plot. The measurements were taken at 5 m, 15 m, and 20 m from center on each leg of the line intercept (Figure 1.5). At those marks we measured the height of the tallest piece of current-year and residual grass in a 1 m radius. We measured ground cover using Daubenmire quadrats (20 cm x 50 cm; $1\text{-}m^2$) to determine percent grass, forb, shrub, litter and bare ground cover (Daubenmire 1959). The quadrats were centered on the left side of the legs looking out from the center and were placed at the 7.5 m, 12.5 m, 20 m, 32.5 m, 37.5 m, and 45 m marks on each of the two 50 m measuring tapes. We evaluated shrub coverage using

the line intercept method. Each time a shrub intersected the 50 m measuring tape, measurements were taken indicating the start and end of the intersection of the shrub with the transect line. We used 6 trees per 50 m^2 as the tree density threshold for determining if the plot met the GCPO LCC's vegetative condition for amount of trees allowed in a plot.

Land Cover

We used ArcMap 10.2 (ArcGIS Desktop: Release 10.2. Environmental Systems Research Institute, Inc., Redlands, CA), Geospatial Modeling Environment Ver. 0.7.3.0 (Beyer 2012) and 2011 NLCD to determine land cover composition in specified buffered radii (200 m, 500 m, 1 km, 3 km, and 5 km) around the surveyed sites. At each plot we created circular buffer polygons in ArcMap for each radius. We then used the "isectpolyrst" command in Geospatial Modeling Environment to determine land cover composition in each buffer polygon. Percent of habitat was measured for grassland, pine forest, deciduous forest, mixed forest, and woody wetlands at the specified radii listed above.

Network connectivity was a measure of grassland network size around each site. We used the methods from Farrell (2015) to calculate the number of grassland patches and linkages in a network with distances of 500 m, 1000 m, and 1500 m. Distances chosen were to reflect probable/possible dispersal movements of the oldfield mouse and cotton rat (Bowne et al. 1999, Oddy et al. 1999, Swilling and Wooten 2002).

Statistical Analysis

Forty candidate models relating species occupancy and detectability were created for each species. These models were designed to identify habitat attributes that are correlated to occupancy probability of the focal species. The local-scale vegetation measurements, as well as

the broad-scale landscape metrics were used as covariates in single-season occupancy models for both species (Table 1.2). Model creation and analysis was conducted with the package "unmarked", a statistical package produced for use in R (Fiske and Chandler 2011, R Core Team 2014). We used a sequential modeling building process wherein we first evaluated detection probability covariates and then assessed occupancy probability covariates. We limited the number of covariates incorporated into each model because we had very few encounters of any species. We incorporated grass height or bare ground density into a single season model with occupancy as a constant parameter to determine covariate structure of detection models of the oldfield mouse and cotton rat using a model selection analysis (Burnham and Anderson 2002). We hypothesize increased grass height would affect observer vision during visual encounter surveys and increased bare ground would increase the ability of an observer to detect vertebrates. We observed in the field that increased bare ground helped facilitate track-tube placement to be flush with the ground, which could affect detectability. The best detection model was selected and used as the base model to assess occupancy models. Akaike information criterion adjusted for small sample size (AICc) measures the quality of statistical models (Burnham and Anderson 2002) and offers an analytical method for comparing models of occupancy and detection. We used $\leq 2.00 \; \Delta$ AICc as model selection criteria to determine which models best fits the data and model weight (w) to interpret relative likelihood of a model given the data and candidate models (Burnham and Anderson 2002, Arnold 2010). This model selection technique allowed us to determine the covariates that best explain variation in occupancy and detection.

RESULTS

We surveyed 57 grassland plots (28 plots had <50% forested land cover in their 1km buffered area, 29 plots had ≥50% forested land cover in their 1km buffered area). We observed 8

mammal species, 3 amphibian species, and 5 reptile species for a total of 16 species (Table 1.3).

Nine species were only observed with visual encounter surveys, 4 species were observed with

only track-tubes, and 3 species were observed with both methods. We recorded 86 small

mammal detections during 342 track-tube nights or 8,208 track-tube hours. The cotton rat and

oldfield mouse were the most frequently detected focal species in this survey, with detections

recorded at 6 and 3 sites, respectively. The probability of detecting the cotton rat was 0.15 (0.02

– 0.57; 95% CI) based on the intercept only model. The probability of detecting the oldfield

mouse was 0.27 (0.06 – 0.67; 95% CI) when bare ground cover was 11.8%; the mean bare

ground cover measured during this study.

Broad-scale Habitat Composition

The occupancy of the oldfield mouse was not associated with the amount of grassland

and pine forest surrounding the site, but was negatively associated with the amount of woody

wetland and positively associated with mixed forest surrounding the site (Table 1.4). In the top

model for the oldfield mouse ($w = 0.17$) occupancy was negatively associated with proportion of

woody wetlands within 5 km, however, the confidence interval on the beta parameter correlation

coefficient included zero. Oldfield mice were 0.66 (0.39 – 1.13: 95% CI) times as likely to use

sites for every 1% increase in woody wetlands within 5 km of the site (Figure 1.6), thus

exhibiting a decrease in occupancy as woody wetlands increases. The second best model ($w =$

0.16) had the proportion of mixed forest within 200 m as an occupancy covariate for the oldfield

mouse. Oldfield mice were 1.21(1.01 – 1.45: 95% CI) times as likely to use sites for every 1%

increase in mixed forest within 200 m of the site (Figure 1.7). The constant model for the

oldfield mouse ($w = 0.01$) was not one of the top models. Occupancy of the cotton rat was

associated with the proportion of woody wetlands around the site (Table 1.5). In the top 2 models ($w = 0.44$ and $w = 0.27$), cotton rats were 1.18 (1.03 – 1.34: 95% CI) times as likely to use sites for every 1% increase in woody wetlands within 5 km of the site and 1.14 (1.01 – 1.29: 95% CI) times as likely to use sites for every 1% increase in woody wetlands within 3 km of the site (Figure 1.8 and Figure 1.9). The amount of grasslands, mixed forest, and pine forest were not associated with cotton rat occupancy. The constant model for the cotton rat ($w = 0.01$) was not one of the top models.

Mean grassland proportion was greatest at 200 m (0.10 ± 0.16 SD) and smallest at 5 km (0.03 ± 0.02 SD). The mean for pine forest proportion was greatest at 5 km (0.12 ± 0.07 SD) and smallest at 200 m (0.05 ± 0.08 SD). Similarly, the mean for deciduous forest proportion was greatest at 5 km (0.16 ± 0.11 SD) and smallest at 200 m (0.09 ± 0.16 SD). The mean for mixed forest proportion was greatest at 3 km (0.06 ± 0.05 SD) and smallest at 200 m (0.04 ± 0.06 SD). The mean for woody wetland proportion was greatest at 5 km (0.14 ± 0.10 SD) and smallest at 200 m (0.06 ± 0.12 SD).

Plot Vegetation Condition and Patch Connectivity

According to our model selection analysis, occupancy of the oldfield mouse and cotton rat were not associated with grass height, grass density, forb density, bare ground density, and tree density. Furthermore, occupancy of both species was not associated with the number of linkages between grassland patches or the number of grassland patches in a network. Oldfield mouse detection was associated with bare ground density. We observed that the oldfield mouse was 1.12 (0.93 – 1.36: 95% CI) times as likely to be detected for every 1% increase in bare ground.

The plot vegetation conditions had high variability and only one site met all 5 of the vegetation criteria for the GCPO LCC's "desired ecological state". Plots met 4 criteria (n=13), 3 criteria (n=18), 2 criteria (n=13), 1 criteria (n=9), or zero criteria (n=3) of the criteria listed for the GCPO LCC's "desired ecological state". The shrub density criteria (<20% cover) was met at 50 of the 57 sites. The tree density criteria (<6 trees/50 m^2) was met at 36 of the 57 sites. Herbaceous vegetation density criteria (>75% cover) was met at 33 of the 57 plots. The bare ground density criteria (>5% and <20% cover) was met at 25 of the 75 plots. The grass height criteria (4 – 6 feet) was met at 2 of the 57 plots surveyed. The mean grass height for all plots was 69.5 cm (± 28.5 SD), mean grass cover was 41.0% (± 22.8 SD), mean forb cover was 28.2% (± 23.4 SD), and mean bare ground was 11.8% (± 15.9 SD).

DISCUSSION

Occupancy of the cotton rat was not associated with the amount of grassland in the surrounding landscape; however, cotton rat occupancy was positively associated with woody wetlands in the surrounding landscape. These findings suggest that woody wetlands in the landscape may facilitate cotton rat occupancy in the EGCP. The cotton rat can persist in many different types of habitats with herbaceous vegetation other than grasslands (Flehartly and Mares 1973, Best and Dusi 2014). Our results agree with previous studies that have found cotton rats to be the most abundant small mammal in regenerating bottomland forests (Chamberlain and Leopold 2003). The use of only the cotton rat as surrogate species for its grassland associated cryptic predators may not be appropriate because of the apparent non-association cotton rats have with grasslands.

The directionality of the relationship of woody wetlands and oldfield mouse occupancy could not be determined with certainty because the 95 % confidence interval around the estimate included zero, however the mean estimate and the model selection analysis indicate a negative relationship. The results from the models with mixed forest are difficult to interpret but they could be specious due to the small number of encounters in our study. If these estimates are accurate, it would support the observations from the central part of the species range where it inhabits mixed forest (Best and Dusi 2014). The oldfield mouse is a habitat specialist that requires sandy soils to build burrows and herbaceous vegetation as a food resource (Best and Dusi 2014). A majority of our sites were located in Black Belt Prairie region where soil is composed mostly of calcium carbonate and clay which is distinct from the sandier soils in other regions in the EGCP. The 3 plots were the oldfield mouse was detected were in the regions in the study area with sandier soils south of the Black Belt Prairie region. Soil type may have explanatory abilities for oldfield mouse occupancy that could not be identified because of its lack of inclusion in models during the study. A lack of understanding of how soil type affects occupancy could serve as a limiting factor for locations of grassland restoration projects.

Many of our plots did not meet all of the desired vegetation conditions of grasslands described by the GCPO LCC. We hypothesize one of the reasons for this is because we used remotely sensed land cover data to randomly select plots in the grassland land cover class. Accuracy of land cover data can be a low for rare habitat types such as grasslands in the EGCP (Wickham 2013). Some of our plots were located on land used for even-aged timber management. Land cover data most likely classified the land as grassland because the classification process must have taken place shortly after timber harvest which removes almost all trees and allows herbaceous vegetation to grow. The desired vegetative condition of grass

height (1.22 – 1.83 m) was only met at 2 plots. We hypothesize the reasons for this is due to recent prescribed fires, haying and mowing operations, overgrowth, and surveys being conducted early in the growing season. The desired vegetative condition of bare ground density (5 -20%) was the met at 25 of the 57 plots. We hypothesize the reason is because some of our plots were located on land used for haying which may have had nonnative cool season grasses. Cool season grasses do not bunch like the native warm season grasses which increase bare ground density. These results suggest using remotely sensed land cover data to describe grasslands in the landscape may not be appropriate.

Sample size can influence the results of occupancy estimation models by biasing estimates when sample size is small (Mackenzie 2002). Our ability to build more complex models by adding more covariates was limited by the small sample size. Freedman (1983) demonstrated how the number of predictor variables in a regression model should be of the same order of magnitude as the number of data points use in the model. Miller (2002) expanded on these findings and found that coefficient estimates are biased away from zero when more predictor variables are included in a regression model than should be included based on the number of data points.

This study was designed to try and detect small mammals and squamates that utilize grassland patches in the EGCP. We selected VES, an active survey technique, to survey for squmates due to constraints on time and personnel. We intended to increase the number of sites surveyed and survey a large geographic distribution. Drift-fence arrays and pitfall traps, both passive survey techniques, have higher detection probabilities for squamates (Case and Fisher 2001), however these techniques require large amounts of time and personnel to setup and monitor. Drift-fence arrays and pitfall traps would apply significant logistical restrictions on the

number of sites we could survey, so our study design intentionally sacrificed detection probability at a single site in favor of surveying a greater number of sites over a larger spatial extent. The squamate species we encountered during the study were black racer (*Coluber constrictor*), green anole (*Anolis carolinensis*), skink (*Eumeces spp.*) and eastern fence lizard (*Sceloporus undulatus*). Focal species squamates were never detected during surveys, and we conclude that the VES methods may be inappropriate for surveying these kinds of species with only 3 surveys per plot. Mackenzie and Royle (2005) suggest increasing the number of surveys as detection probability decreases. However detection probability is difficult to know at the start of a study without an extensive pilot study first being conducted.

Our results indicate that broad-scale habitat features may be important for small mammal occupancy in the EGCP. The amount of woody wetlands and mixed forest were explanatory variables for the occupancy of the cotton rat and the oldfield mouse. However, broad-scale habitat features are seldom examined as explanatory variables for small mammal population patterns (Michel et al. 2007). Fine-scale vegetation conditions are known to be able to explain small mammal abundance (Bowman et al. 2001). As seen with our results, broad and fine-scale habitat variables may not be mutually exclusive in their explanatory ability of small mammal population patterns, thus multiple scales should be considered when trying to identify influences of population pattern. Additional research should focus on how features of broad-scale habitat could influence the occupancy of small mammals and other terrestrial vertebrates.

Management Implications

Our results indicate small mammal occupancy is influenced by broad-scale habitat features. Previous studies of the old field mouse and the cotton rat have largely focused on habitat

selection at very small spatial scales, and while our results do not contradict those previous results, our study does indicate that larger spatial scales should be considered in habitat research and management decisions. Land managers who want to increase small mammal occupancy may want to consider broad-scale habitat features when selecting sites for restoration and conservation. Our results indicate cotton rat occupancy may have a positive association with woody wetlands in the surrounding landscape and oldfield mouse occupancy may have a positive association with mixed forest in the surrounding landscape. Certain habitat types that are not usually considered important for a species may have some facilitative or inhibitory properties that can influence occupancy of nearby grassland sites.

Literature Cited

Andrén, H. 1994. Effects of habitat fragmentation on birds and mammals in landscapes with different proportions of suitable habitat: a review. Oikos 71:355–366.

Arnold, T. W. 2010. Uninformative parameters and model selection using Akaike's information criterion. Journal of Wildlife Management 74:1175–1178.

Baker, R. H. 1991. Naturalists mammalian prey of the common barn-owl (*Tyto alba*) along the Texas coast. The Southwestern Naturalist 36:343–347.

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

Baxley, D. L., and C. P. Qualls. 2009. Black pine snake (*Pituophis melanoleucus lodingi*): spatial ecology and associations between habitat use and prey dynamics. Journal of Herpetology 43:284–293.

Beasley, J. C., T. L. Devault, M. I. Retamosa, and O. E. Rhodes Jr. 2007. A hierarchical analysis of habitat selection by raccoons in Northern Indiana. Journal of Wildlife Management 71:1125–1133.

Beasom, S. L., and R. A. Moore. 1977. Bobcat food habit response to a change in prey abundance. The Southwestern Naturalist 21:451–457.

Beyer, H.L. 2012. Geospatial Modelling Environment (Version 0.7.3.0). (software). URL: http://www.spatialecology.com/gme.

Bowman, J., G. Forbes, and T. Dilworth. 2001. Landscape context and small-mammal abundance in a managed forest. Forest Ecology and Management 140:249–255.

Bowne, D., J. Peles, and G. Barrett. 1999. Effects of landscape spatial structure on movement patterns of the hispid cotton rat (*Sigmodon hispidus*). Landscape Ecology 14:53–65.

Burnham, K. P., and D. R. Anderson. 2002. Model selection and multimodel inference: a practical information-theoretic approach. Springer Science & Business Media, New York, NY. 488 pp.

Chamberlain, M. J., and B. D. Leopold. 2003. Effects of a flood on relative abundance and diversity of small mammals in a regenerating bottomland hardwood forest. The Southwestern Naturalist 48:306–309.

Clark, B. K., B. S. Clark, T. R. Homerding, and W. E. Munsterman. 1998. Communities of small mammals in six grass-dominated habitats of southeastern Oklahoma. The American Midland Naturalist 139:262–268.

Daubenmire, R., 1959. A canopy-coverage method of vegetational analysis. Northwest Science 33:43-64.

Diffendorfer, J. E., M. S. Gaines, and R. D. Holt. 1995. Habitat fragmentation and movements of three small mammals (*Sigmodon*, *Microtus*, and *Peromyscus*). Ecology 76:827–839.

Drennan, J. E., P. Beier, and N. L. Dodd. 1998. Use of track stations to index abundance of sciurids. Journal of Mammalogy 79:352–359.

Farrell, P. D. 2015. The use of a small scale study and regional data sources to understand grassland bird relationships. M.S. Thesis. Auburn University, Auburn, AL.115 pp.

Flather, C. H., and M. Bevers. 2002. Patchy reaction diffusion and population abundance: the relative importance of habitat amount and arrangement. The American Naturalist 159:40–56.

Fleharty, E. D., and M. A. Mares. 1973. Habitat preference and spatial relations of *Sigmodon hispidus* on a remnant prairie in West-Central Kansas. The Southwestern Naturalist 18:21–29.

Fiske, I. J., and R. B. Chandler. 2011. unmarked : An R package for fitting hierarchical models of wildlife occurrence and abundance. Journal of Statistical Software 43:1–23.

Fitzgibbon, C. D. 1997. Small mammals in farm woodlands: the effects of habitat, isolation and surrounding land-use patterns. Journal of Applied Ecology 34:530–539.

Freedman, D. A. 1983. A note on screening regression equations. The American Statistician 37:152–155.

GCPO LCC. 2009. Gulf Coastal Plains and Ozarks Landscape Conservation Cooperative development and operations plan. www.fws.gov/southeast/LCC/GulfPlains/pdf/ GCPOLCCDevOpPlanFinal12112009.pdf. Accessed 16 June 2016.

GCPO LCC. 2014. Gulf Coastal Plains and Ozarks Landscape Conservation Cooperative integrated science agenda v. 4. http://api.ning.com/files/2zaTVXiHsQQp52wJb3mdPb8GOamb-9txWAxWdNdeB-cvArmNWFifu-kHE7h3Br61ZnH4zGSYwt9P*vLWiRg*TKZRAajjmE3F/ DraftIntegratedScienceAgendaV4.pdf. Accessed 16 June 2016.

Glennon, M. J., W. F. Porter, and C. L. Demers. 2002. An alternative field technique for estimating diversity of small-mammal populations. Journal of Mammalogy 83:734–742.

Hafner, D. J., E. Yensen, and G. L. Kirkland Jr.1998. North American rodents. Status survey and conservation action plan. IUCN/SSC Rodent Specialist Group. IUCN, Gland, Switzerland 171pp.

Homer, C.G., J.A. Dewitz, L. Yang, S. Jin, P. Danielson, G. Xian, J. Coulston, N.D. Herold, J.D. Wickham, and K. Megown. 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 81:345-354.

Humphrey, S. R. and D. B. Barbour. 1981. Status and habitat of 3 subspecies of *Perornyscus polionotus* in Florida. Journal of Mammalogy. 62:840-844.

Jorgensen, E. E. 2004. Small mammal use of microhabitat reviewed. Journal of Mammalogy 85:531–539.

Mackenzie, D. I., J. D. Nichols, G. B. Lachman, S. Droege, J. A. Royle, and C. A. Langtimm. 2002. Estimating site occupancy rates when detection probabilities are less than one. 83:2248–2255.

Mackenzie, D. I., and J. A. Royle. 2005. Designing occupancy studies: General advice and allocating survey effort. Journal of Applied Ecology 42:1105–1114.

Means, D. B. 2006. Vertebrate faunal diversity of longleaf pine ecosystems. Pp. 157–213, *In* S. Jose, E. J. Jokela, and D. L. Miller (Eds.). The Longleaf Pine Ecosystem: Ecology, Silviculture, and Restoration. Springer, New York. 438 pp.

Mengak, M. T., and D. C. Guynn. 2003. Small mammal microhabitat use on young loblolly pine regeneration areas. Forest Ecology and Management 173:309–317.

Michel, N., F. Burel, P. Legendre, and A. Butet. 2007. Role of habitat and landscape in structuring small mammal assemblages in hedgerow networks of contrasted farming landscapes in Brittany, France. Landscape Ecology 22:1241–1253.

Miller, A. 2002. Subset selection in regression. CRC Press, Boca Raton, FL. 256 pp.

Moore, J. E., and R. K. Swihart. 2005. Modeling patch occupancy by forest rodents: incorporating detectability and spatial autocorrelation with hierarchically structured data. Journal of Wildlife Management 69:933–949.

Noss, R. F. 2013. Forgotten grasslands of the South: natural history and conservation. Island Press. Washington, D.C., USA. 320 pp.

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. Biological report 28. National Biological Service, Washington, D.C., USA. 83 pp.

Oddy, D. M., M. A. Hensley, J. A. Provancha, and R. B. Smith. 1999. Long-distance dispersal of a southeastern beach mouse (*Peromyscus polionotus niveiventris*) at Cape Canaveral, Florida. Florida Field Naturalist 8:124–125.

Oehler, J. D., and J. a. Litvaitis. 1996. The role of spatial scale in understanding responses of medium-sized carnivores to forest fragmentation. Canadian Journal of Zoology 74:2070–2079.

Panzacchi, M., J. D. C. Linnell, C. Melis, M. Odden, J. Odden, L. Gorini, and R. Andersen. 2010. Effect of land-use on small mammal abundance and diversity in a forest–farmland mosaic landscape in south-eastern Norway. Forest Ecology and Management 259:1536–1545.

Preston, C. R. 1990. Distribution of raptor foraging in relation to prey biomass and habitat structure. The Condor 92:107–112.

PRISM Climate Group, Oregon State University, http://prism.oregonstate.edu, created 12 Jul 2016.

R Core Team. 2014. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/

Sheffield, L. M., J. R. Crait, W. D. Edge, and G. Wang. 2001. Response of American kestrels and gray-tailed voles to vegetation height and supplemental perches. Canadian Journal of Zoology 79:380–385.

Swihart, R. K., J. R. Goheen, S. A. Schnelker, and C. E. Rizkalla. 2007. Testing the generality of patch and landscape-level predictors of tree squirrel occurrence at a regional scale. Journal of Mammalogy 88:564–572.

Swilling, W. R., and M. C. Wooten. 2002. Subadult dispersal in a monogamous species : the alabama beach mouse (*Peromyscus polionotus ammobates*). Journal of Mammalogy 83:252–259.

Tyler, J. D., and J. F. Jensen. 1981. Notes on foods of great horned owls (*Bubo virginianus*) in Jackson County, Oklahoma. Pg. 28-30, *In* G. W. Todd (Ed.). Proceedings of the Oklahoma Academy of Science Vol. 61. 110 pp.

Urban, N. a., and R. K. Swihart. 2009. Multiscale perspectives on occupancy of meadow jumping mice in landscapes dominated by agriculture. Journal of Mammalogy 90:1431–1439.

Wilson, T. H. 1981. Natural history of the Black Belt prairie. Journal of the Alabama Academy of Science 52:10-19

Wickham, J. D., S. V Stehman, L. Gass, J. Dewitz, J. A. Fry, and T. G. Wade. 2013. Remote sensing of environment accuracy assessment of NLCD 2006 land cover and impervious surface. Remote Sensing of Environment 130:294–304.

TABLES AND FIGURES

Table 1.1. Final focal species list for the project after consideration of relevant selection criteria. Home range size (small: ≤5 ha, medium: >5 ha but ≤10 ha, large: >10 ha). Density (low ≤4 individuals per ha, medium: >5 but ≤ 10 individuals per ha, high: >10 individuals per ha). Conservation Concern is based on the species status provided by the Alabama Department of Conservation and Natural Resources. GCPO LCC signifies if the species is listed in the GCPO LCC Integrated Science Agenda as a representative of grassland habitat.

Functional Group	Common Name	Scientific Name	Habitat Use	Home Range	Density	Conservation Concern	GCPO LCC
Squamates	Six-lined racerunner	*Aspidoscelis sexlineata*	open habitats	Small	Low	Moderate	YES
	Corn snake	*Pantherophis guttata guttata*	dry uplands	Large	Unknown	Moderate	YES
	Slender glass lizard	*Ophisaurus ventralis*	open habitats	Unknown	Unknown	Moderate	YES
	Eastern kingsnake	*Lampropeltis getula*	open pine savannas, macrohabitat generalist	Large	Unknown	High	YES
Small Mammals	Eastern harvest mouse	*Reithrodontomys humulis*	Fallow fields	Small	Medium	Moderate	YES
	Oldfield mouse	*Peromyscus polionotus*	Grassland, sandy soil	Small	Unknown	Moderate	YES
	Southeastern shrew	*Sorex longirostris*	old field, dry upland, shrubby areas	Small	Unknown	Moderate	YES
	Hispid cotton rat	*Sigmodon hispidus*	Grass-dominated habitats	Medium	High	Lowest	NO

Table 1.2. Occupancy covariates used in the single season analysis of oldfield mouse (*Peromyscus polionotus*) and cotton rat (*Sigmodon hispidus*) occupancy in the Mississippi and Alabama portions of the East Gulf Coastal Plain from March 2015 – July 2015. Model 6 is the constant occupancy model.

	(ψ) Occupancy Covariates		(ψ) Occupancy Covariates
1	Woody Wetlands 1 km	21	Grassland 3 km
2	Woody Wetlands 200 m	22	Grassland 500 m
3	Woody Wetlands 3 km	23	Grassland 5 km
4	Woody Wetlands 5 km	24	Herb Cover
5	Woody Wetlands 500 m	25	Mixed Forest 1 km
6	.	26	Mixed Forest 200 m
7	Bare Ground	27	Mixed Forest 3 km
8	Deciduous Forest 1 km	28	Mixed Forest 5 km
9	Deciduous Forest 200 m	29	Mixed Forest 500 m
10	Deciduous Forest 3 km	30	Evergreen Forest 1 km
11	Deciduous Forest 5 km	31	Evergreen Forest 200 m
12	Deciduous Forest 500 m	32	Evergreen Forest 3 km
13	All Forest 1 km	33	Evergreen Forest 5 km
14	All Forest 200 m	34	Evergreen Forest 500 m
15	All Forest 3 km	35	Net Size 1 km
16	All Forest 500 m	36	Net Size 1.5 km
17	All Forest 5 km	37	Net Size 500 m
18	Grass Height	38	Num Links 1 km
19	Grassland 1 km	39	Num Links 1.5 km
20	Grassland 200 m	40	Num Links 500 m

Table 1.3. Species detected during surveys in the Mississippi and Alabama portions of the East Gulf Coastal Plain from March 2015 – July 2015. This list includes the number of sites that the species was detected at and the detection method.

Species	Number of Sites Detected	Detection Method Track-tube	VES
Mammals			
Sigmodon hispidus	6	X	X
Peromyscus polionotus	3	X	X
Peromyscus spp.	11	X	
Didelphis virginiana	3	X	
Procyon lotor	12	X	
Microtus spp.	2	X	
Sorex spp./ Blarina spp./ Cryptotis spp.	4	X	X
Sylvilagus spp.	2		X
Reptiles			
Anolis carolinensis	6		X
Coluber constrictor	1		X
Sceloporus undulatus	1		X
Eumeces spp.	1		X
Terrapene carolina triunguis	1		X
Amphibians			
Hyla squirella	1		X
Anaxyrus terrestris	1		X
Pseudacris spp.	1		X

24

Table 1.4. Summary of single season occupancy (ψ) models for oldfield mouse (*Peromyscus polionotus*) at 57 plots in the Mississippi and Alabama portions of the East Gulf Coastal Plain from March 2015 – July 2015. Models are structured with occupancy (ψ) and detection (p) components and are reported with AICc, difference of model AICc to best model (AICc), model likelihood, model weight (*w*), and cumulative model weights (Cum. *w*).

Model	AICc	Δ AICc	Likelihood	*w*	Cum. *w*
Ψ(Woody Wetlands 5km) p(Bare Ground)	31.97	0.00	1.00	0.17	0.17
Ψ(Mixed Forest 200m) p(Bare Ground)	32.08	0.12	0.94	0.16	0.32
Ψ(Mixed Forest 5km) p(Bare Ground)	33.15	1.18	0.55	0.09	0.41
Ψ(Woody Wetlands 3km) p(Bare Ground)	34.36	2.39	0.30	0.05	0.46
Ψ(Mixed Forest 500m) p(Bare Ground)	34.38	2.42	0.30	0.05	0.51
Ψ(Grass Height) p(Bare Ground)	34.68	2.71	0.26	0.04	0.56
Ψ(Evergreen Forest 5km) p(Bare Ground)	35.32	3.35	0.19	0.03	0.59
Ψ(Mixed Forest 3km) p(Bare Ground)	35.38	3.42	0.18	0.03	0.62
Ψ(.) p(Bare Ground)	35.63	3.66	0.16	0.03	0.64
Ψ(Grassland 1km) p(Bare Ground)	35.93	3.96	0.14	0.02	0.67
Ψ(Mixed Forest 1km) p(Bare Ground)	36.19	4.22	0.12	0.02	0.69
Ψ(Evergreen Forest 3km) p(Bare Ground)	36.20	4.24	0.12	0.02	0.71
Ψ(Woody Wetlands 200m) p(Bare Ground)	36.21	4.24	0.12	0.02	0.73
Ψ(Grassland 3km) p(Bare Ground)	36.48	4.51	0.10	0.02	0.74
Ψ(Num Links 500m) p(Bare Ground)	36.50	4.53	0.10	0.02	0.76
Ψ(All Forest 5km) p(Bare Ground)	36.52	4.55	0.10	0.02	0.78
Ψ(Deciduous Forest 5km) p(Bare Ground)	36.93	4.97	0.08	0.01	0.79
Ψ(Grassland 5km) p(Bare Ground)	37.10	5.13	0.08	0.01	0.80
Ψ(Grassland 200m) p(Bare Ground)	37.14	5.17	0.08	0.01	0.82
Ψ(Woody Wetlands 1km) p(Bare Ground)	37.47	5.50	0.06	0.01	0.83
Ψ(All Forest 3km) p(Bare Ground)	37.47	5.50	0.06	0.01	0.84
Ψ(Woody Wetlands 500m) p(Bare Ground)	37.52	5.56	0.06	0.01	0.85

Ψ(Deciduous Forest 3km) p(Bare Ground)	37.56	5.59	0.06	0.01	0.86
Ψ(Grassland 500m) p(Bare Ground)	37.58	5.61	0.06	0.01	0.87
Ψ(Deciduous Forest 200m) p(Bare Ground)	37.67	5.70	0.06	0.01	0.88
Ψ(Evergreen Forest 500m) p(Bare Ground)	37.69	5.72	0.06	0.01	0.89
Ψ(All Forest 500m) p(Bare Ground)	37.70	5.73	0.06	0.01	0.90
Ψ(Evergreen Forest 1km) p(Bare Ground)	37.72	5.75	0.06	0.01	0.91
Ψ(Num Links 1km) p(Bare Ground)	37.78	5.81	0.05	0.01	0.92
Ψ(All Forest 200m) p(Bare Ground)	37.85	5.89	0.05	0.01	0.92
Ψ(All Forest 1km) p(Bare Ground)	37.90	5.93	0.05	0.01	0.93
Ψ(Net Size 1km) p(Bare Ground)	37.92	5.95	0.05	0.01	0.94
Ψ(Deciduous Forest 500m) p(Bare Ground)	37.93	5.96	0.05	0.01	0.95
Ψ(Evergreen Forest 200m) p(Bare Ground)	37.94	5.97	0.05	0.01	0.96
Ψ(Net Size 500m) p(Bare Ground)	37.94	5.97	0.05	0.01	0.97
Ψ(Herb Cover) p(Bare Ground)	37.94	5.98	0.05	0.01	0.97
Ψ(Num Links 1.5km) p(Bare Ground)	37.95	5.98	0.05	0.01	0.98
Ψ(Deciduous Forest 1km) p(Bare Ground)	37.95	5.98	0.05	0.01	0.99
Ψ(Bare Ground) p(Bare Ground)	37.95	5.98	0.05	0.01	1.00
Ψ(Net Size 1.5km) p(Bare Ground)	44.43	12.46	0.00	0.00	1.00

Table 1.5. Summary of single season occupancy (ψ) models for hispid cotton rat (*Sigmodon hispidus*) at 57 plots in the Mississippi and Alabama portions of the East Gulf Coastal Plain from March 2015 – July 2015. Models are structured with occupancy (ψ) and detection (p) components and are reported with AICc, difference of model AICc to best model (AICc), model likelihood, model weight (*w*), and cumulative model weights (Cum. *w*).

Model	AICc	Δ AICc	Likelihood	*w*	Cum. *w*
Ψ(Woody Wetlands 5km) p(.)	53.19	0	1.00	0.44	0.44
Ψ(Woody Wetlands 3km) p(.)	54.21	1.02	0.60	0.27	0.71
Ψ(Woody Wetlands 1km) p(.)	56.11	2.92	0.23	0.10	0.81
Ψ(Woody Wetlands 500m) p(.)	56.62	3.43	0.18	0.08	0.89
Ψ(Woody Wetlands 200m) p(.)	59.1	5.91	0.05	0.02	0.91
Ψ(Mixed Forest km) p(.)	60.44	7.25	0.03	0.01	0.92
Ψ(Grassland 1km) p(.)	61.05	7.86	0.02	0.01	0.93
Ψ(.) p(.)	61.29	8.1	0.02	0.01	0.94
Ψ(Evergreen Forest500m) p(.)	61.92	8.73	0.01	0.01	0.95
Ψ(Mixed Forest 200m) p(.)	62.34	9.15	0.01	0.00	0.95
Ψ(Mixed Forest 5km) p(.)	62.54	9.35	0.01	0.00	0.96
Ψ(Deciduous Forest 5km) p(.)	62.68	9.49	0.01	0.00	0.96
Ψ(All Forest 1km) p(.)	62.96	9.77	0.01	0.00	0.96
Ψ(All Forest 3km) p(.)	62.96	9.77	0.01	0.00	0.97
Ψ(Deciduous Forest 3km) p(.)	62.99	9.8	0.01	0.00	0.97
Ψ(Evergreen Forest 200m) p(.)	63.05	9.86	0.01	0.00	0.97
Ψ(Grassland 500m) p(.)	63.07	9.88	0.01	0.00	0.98
Ψ(All Forest 5km) p(.)	63.21	10.02	0.01	0.00	0.98
Ψ(Grassland 200m) p(.)	63.27	10.08	0.01	0.00	0.98
Ψ(Herb Cover) p(.)	63.32	10.13	0.01	0.00	0.98
Ψ(Deciduous Forest 500m) p(.)	63.36	10.17	0.01	0.00	0.99
Ψ(All Forest 500m) p(.)	63.4	10.21	0.01	0.00	0.99

Ψ(All Forest 200m) p(.)	63.43	10.24	0.01	0.00	0.99
Ψ(Deciduous Forest 1km) p(.)	63.45	10.26	0.01	0.00	0.99
Ψ(Deciduous Forest 200m) p(.)	63.51	10.32	0.01	0.00	1
Ψ(Grassland 5km) p(.)	63.51	10.32	0.01	0.00	1
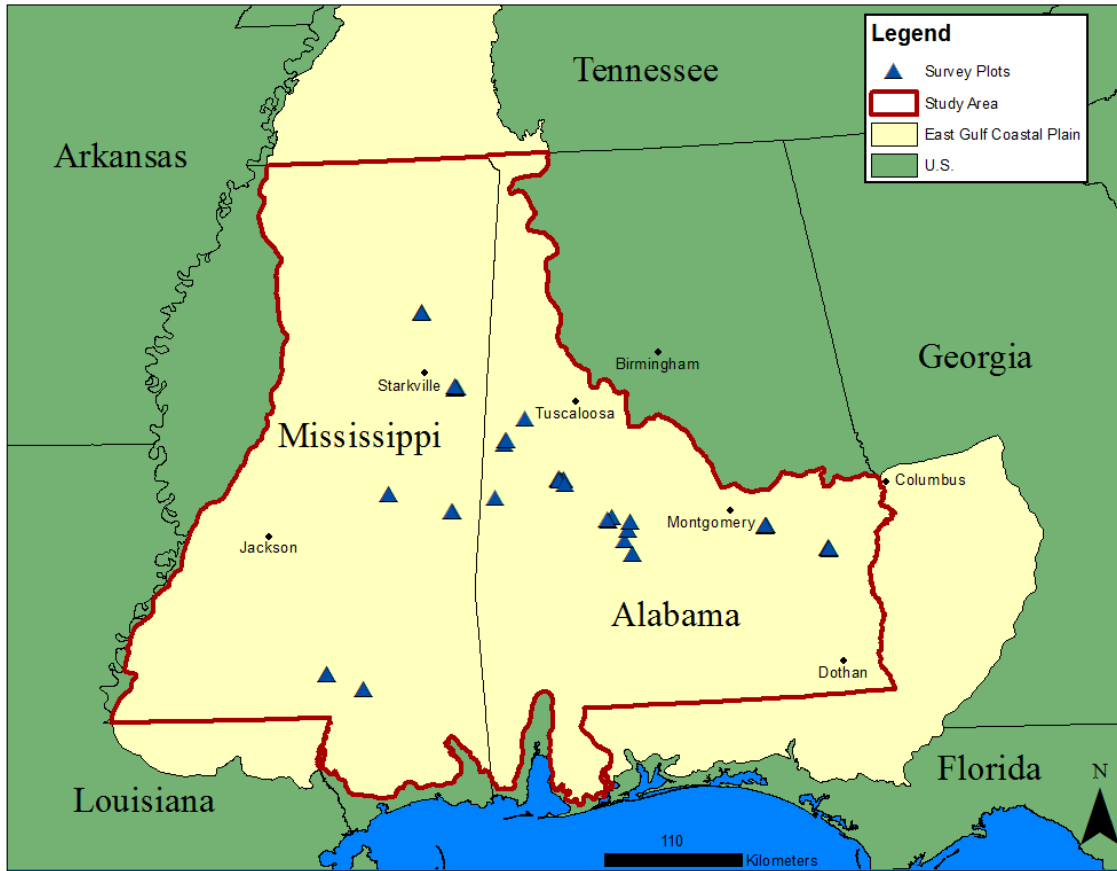
Figure 1.1. The study area boundary with the location of 57 plots surveyed for focal species and vegetation conditions. The blue triangles are the plots surveyed from March 2015 – July 2015 and used in our occupancy analysis.

Figure 1.2. An example of the track-tubes used in this study.
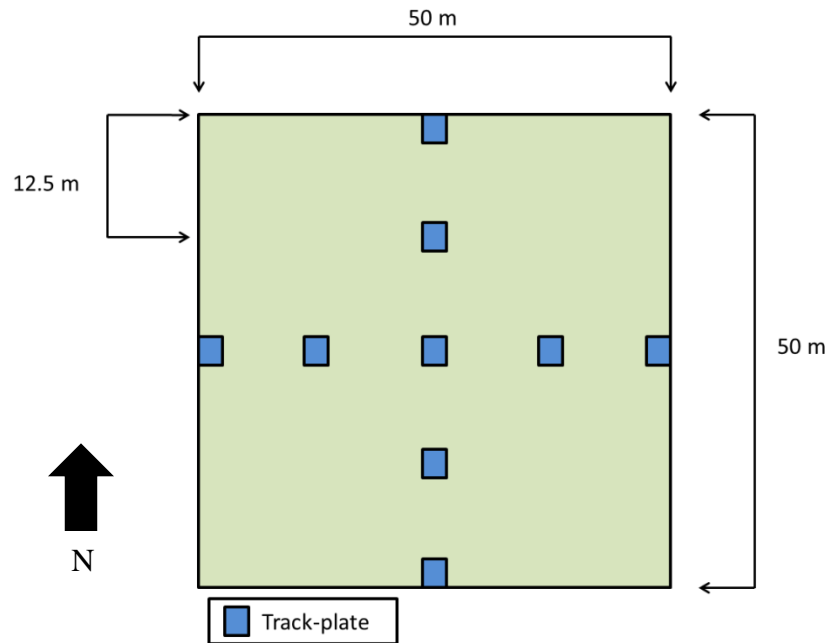
Figure 1.3. Schematic of track-tube placement used at each of 57 plots in the Mississippi and Alabama portions of the East Gulf Coastal Plain in March 2015 – July 2015.
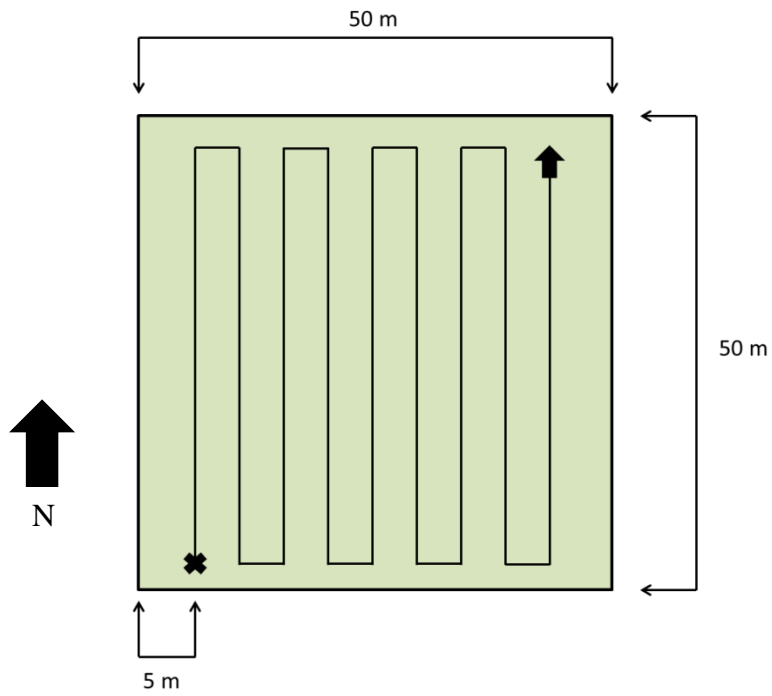
Figure 1.4. Survey pattern for time-constrained visual encounter surveys at 57 plots in the Mississippi and Alabama portions of the East Gulf Coastal Plain in March 2015 – July 2015.
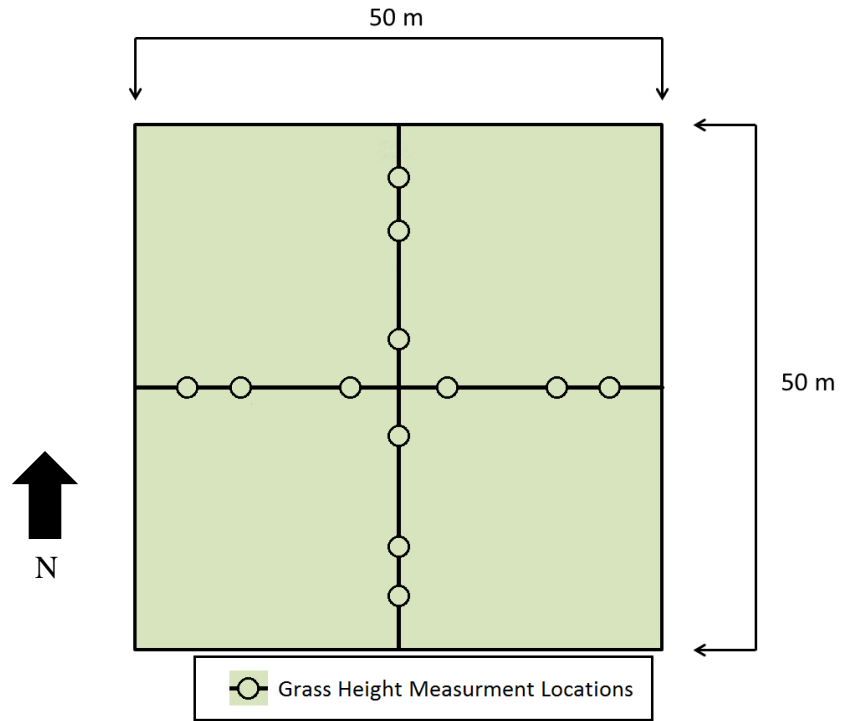
Figure 1.5. Design for measuring current-growth and residual grass heights at each of 57 plots in

the Mississippi and Alabama portions of the East Gulf Coastal Plain in March 2015 – July 2015.

Circles represent 1 meter radius inside which the tallest grass stem was measured.
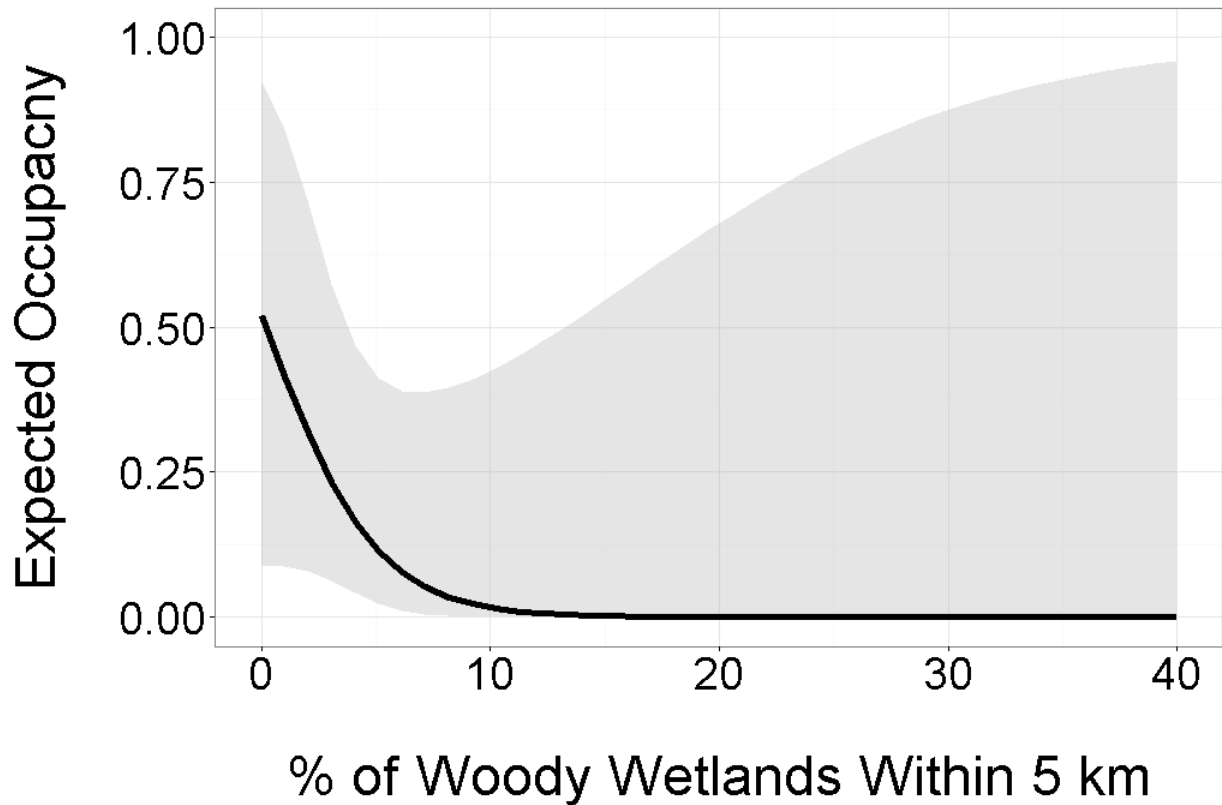
Figure 1.6. Expected occupancy probability of oldfield mouse (*Peromyscus polionotus*) vs. percentage of woody wetlands within 5 km in the Mississippi and Alabama portions in the East Gulf Coastal Plain in 2015. Shading represents the 95% prediction interval of occupancy probability.
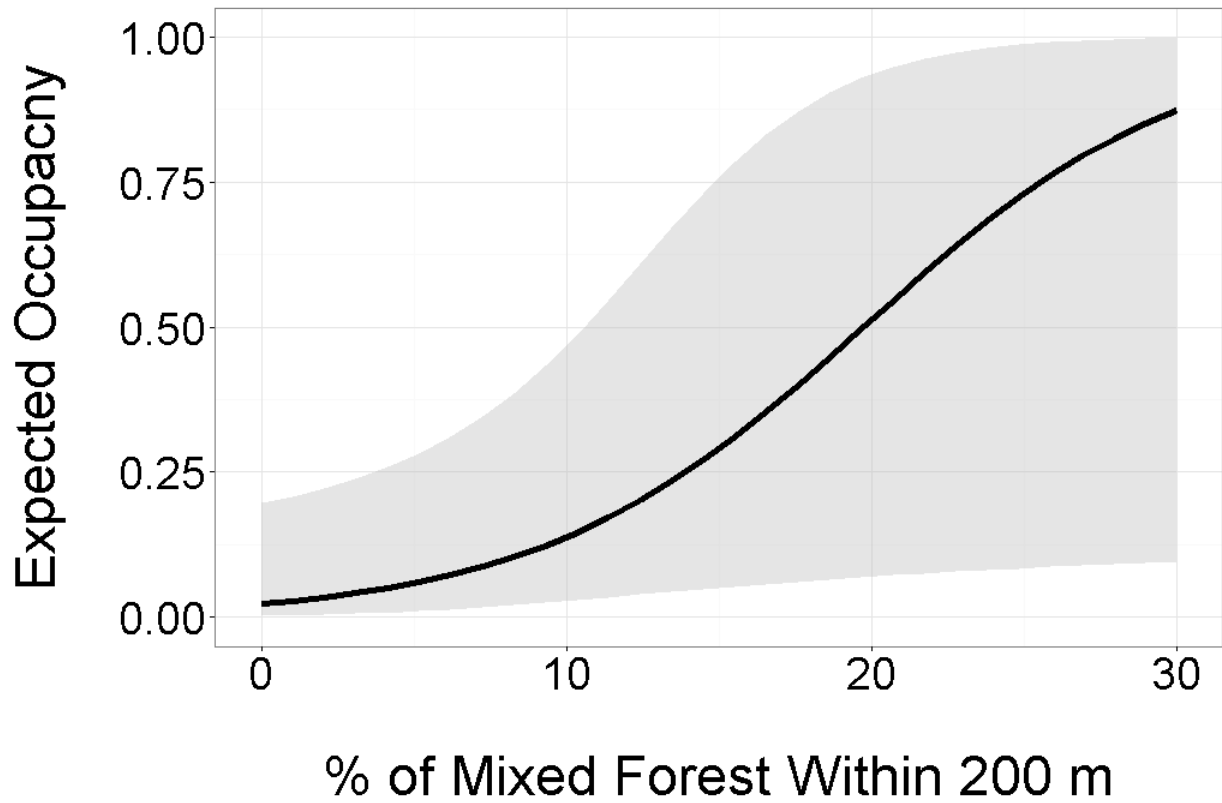
Figure 1.7. Expected occupancy probability of oldfield mouse (*Peromyscus polionotus*) vs. percentage of mixed forest within 200 m in the Mississippi and Alabama portions in the East Gulf Coastal Plain in 2015. Shading represents the 95% prediction interval of occupancy probability.
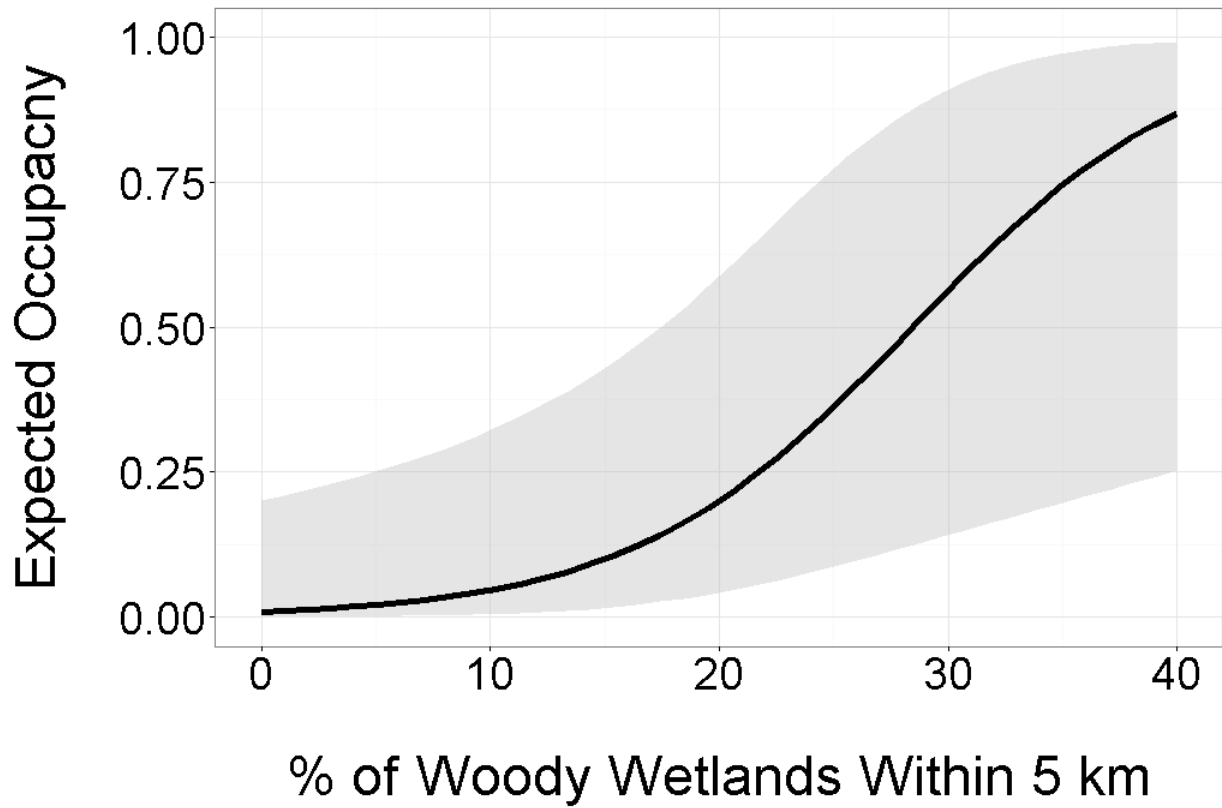
Figure 1.8. Expected occupancy probability of hispid cotton rat (*Sigmodon hispidus*) vs. percentage of woody wetlands within 5 km in the Mississippi and Alabama portions in the East Gulf Coastal Plain in 2015. Shading represents the 95% prediction interval of occupancy probability.
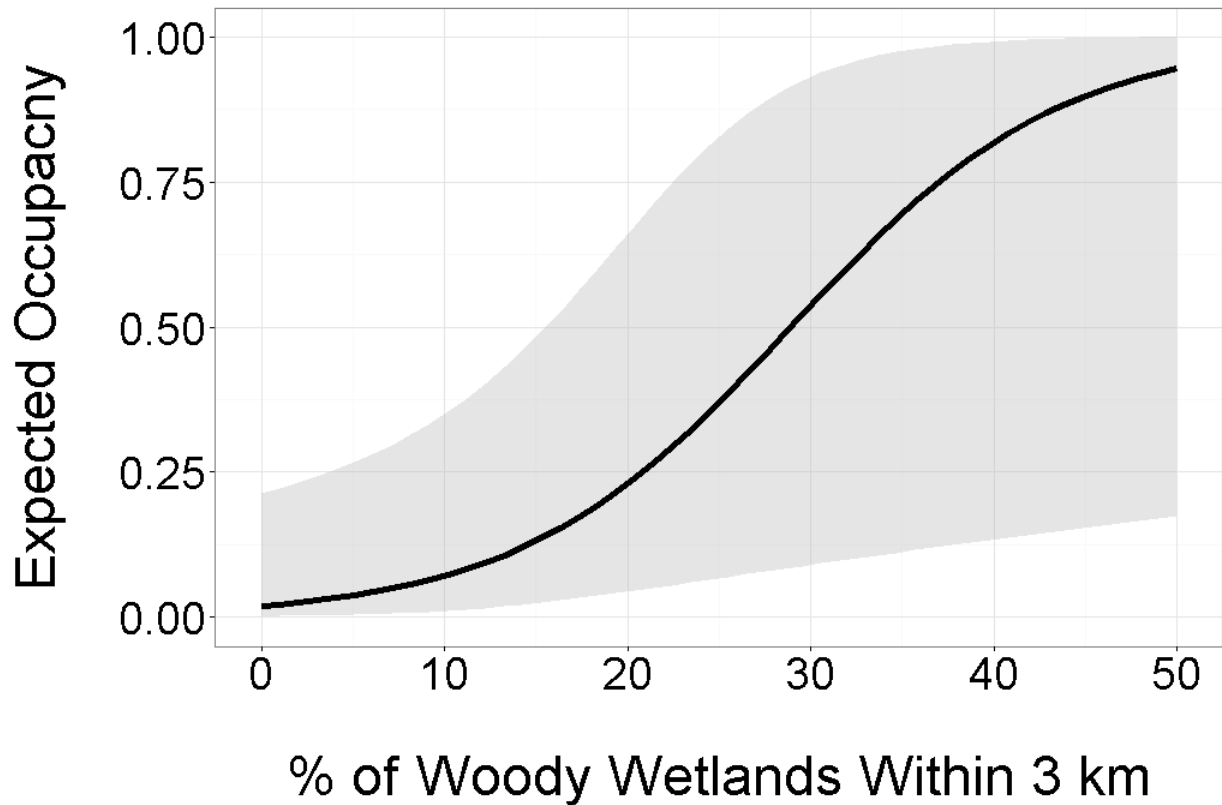
Figure 1.9. Expected occupancy probability of hispid cotton rat (*Sigmodon hispidus*) vs. percentage of woody wetlands within 3 km in the Mississippi and Alabama portions in the East Gulf Coastal Plain in 2015. Shading represents the 95% prediction interval of occupancy probability.

Chapter 2:
The Effects of Error in Remotely Sensed Land Cover Maps on Habitat Relationship Models

ABSTRACT

The use of remotely sensed land cover data continues to grow in ecological research and conservation decision making. However, misclassification of land cover types, a type of error, is an inherent part of the process of creating land cover maps from remotely sensed data. The overall misclassification of land cover maps may misrepresent misclassification in some sub-regions of the land cover map. In the southeastern United States, land cover changes occur quickly over time, due to clear-cutting of timberlands, rapid, natural succession rates and other factors. These temporal changes may not be fully represented in a remotely sensed land cover data products due to a time lag between when classification occurs and the time of interest. In this study, I used simulations to evaluate the effects of error in land cover data on occupancy estimation results. Occupancy estimation related to land cover data can provide estimates of focal species habitat associations for conservation decision makers to determine where and what habitat to manage on the landscape. However error in land cover classification may bias occupancy analyses and lead to erroneous inference with respect to landscape scale habitat management. My simulations used 12 error scenarios, 2 habitat covariates, and 2 sets of predefined parameter estimates. The 12 error scenarios were representative of the accuracy assessment of the 2006 National Land Cover Database, the land cover change index from 2011 National Land Cover Database or simulated misclassification rates. The 2 habitat covariates were measures of proportions of a common habitat class (i.e., forest) and a rare habitat class (i.e.,

grassland). One set of the predefined parameter estimates gave a hypothetical species a weak

negative occupancy association to grassland land cover and a strong positive occupancy

association forest land cover; the other set gave a strong positive occupancy association to

grassland land cover and a weak negative occupancy association to forest land cover. I found

regression coefficient biases are increased when the land cover class of interest is of lower

proportions in the landscape. As the strength of the relationship between occupancy and

covariates increases the precision of the regression coefficients decrease. Ecological researchers

and conservation biologists who use remotely sensed land cover data to estimate habitat

relationships of species should utilize measured land cover error to help estimate covariates that

are closer to truth.

INTRODUCTION

In ecology and conservation biology, remote sensing is increasingly being used to

support habitat and landscape conservation decisions and to benefit species and ecosystems

(Horning et al. 2010; Rose et al. 2015; Turner et al. 2015). Remote sensing is the acquisition of

data from scanning the Earth with satellites. Remotely sensed data products include land cover

maps, elevational maps, bathymetric maps, and climatic maps (Horning et al. 2010). Land cover

maps are commonly used in ecology and conservation biology to examine the broad-scale habitat

relationships of vertebrate species (Greenwald et al. 2009; Steen et al. 2012; Urbanek and

Nielsen 2013). Many land cover maps are produced using satellite spectral data combined with

ancillary data, such as digital elevation models, to refine classifications (Vogelmann 1998). The

spectral data received from satellites is used in a decision tree analysis to delineate classes of

land cover. The decision tree uses the spectral signature of a grid cell, i.e., the variation in

reflectance of wavelengths of light, to determine the most likely class of land cover for that cell.

A widely used land cover database in North America is the National Land Cover Database (NLCD), in which data is collected, maintained and updated by United States Geologic Survey every 5 years (Homer et al. 2015).

Researchers are able to use remotely sensed data to find ecological patterns and processes at large scales, which can be measured and used in ecological models. Measurements of these patterns and processes include errors from the remote sensing process that are seldom discussed in an ecological context (Glenn and Ripple 2004, Shao and Wu 2008). Accounting for measurement error in remotely sensed data may allow conservation decision-makers to make better informed decisions that could improve conservation efforts while reducing the risk of misappropriating conservation resources. For instance if a habitat metric is considered to have a stronger relationship with a species than it actually does, the value of areas with less than ideal conditions will be overestimated. The misclassification of land cover data from remotely sensed data is an inherent part of the process that is used to produce these products. It has been suggested that user accuracy and producer accuracy be measures of misclassification used for thematic land cover maps (Liu et al. 2007). The user accuracy is the probability of a classified cell on the map represents that class on the ground and the producers accuracy is the probability of the ground being correctly classified in the map (Story and Congalton 1986). Misclassification can occur when the spectral signature of multiple habitats is similar, and the sensors on the satellites lack the resolution to differentiate between the habitats (Horning et al. 2010). Grass-dominated classes such as grassland/herbaceous, cultivated crops, emergent wetland, hay/pasture, and developed open space accounted for a high percentage of user accuracy error (26%) in the 2006 NLCD (Wickham et al. 2013). These classes all have herbaceous vegetation with similar spectral signatures which causes the large amount of misclassification.

Additionally, classification error can occur when map products are not frequently updated to reflect land cover change during the time of interest (Gallant 2009). Natural processes (e.g. wildland fire, forest disease, etc.) and anthropogenic influences (e.g. forest harvesting, urbanization, agricultural conversion, etc.) can change the land cover over time. According to the 2011 NLCD land cover of the conterminous United States is mostly composed of 21.61% shrub/scrub, 15.51% cultivated crops, 14.56% grassland/herbaceous, 11.56% evergreen forest, and 10.84% deciduous forest (Homer et al. 2015). This composition is not considerably different than the composition of the 2006 NLCD; 1.24% total change in all 16 land cover classes (Fry et al. 2011) over the 5 year period between product releases. Some areas of the United States can see a greater rate of land cover change than others (Sleeter et al. 2013; Homer et al. 2015). In the southern United States 40% of the forested land cover is used for harvesting of timber (Oswalt and Smith 2014). Land cover in this area can change frequently due to timber harvesting operations and the rapid regrowth of loblolly pine (*Pinus taeda*).

Understanding habitat relationships for vertebrate species of conservation concern can be an important part of conservation and management for those species. By determining these relationships, researchers can identify how habitat patterns influence a species occurrence patterns and use that information to improve habitat features and composition for species of conservation concern. Occupancy estimation and modeling is one method used to evaluate habitat features and their influence on the probability of a species occurring at a location. Metrics calculated from remotely sensed land cover data can be used as predictor variables in habitat relationship studies that use occupancy models to predict the probability of species occurrence (Duren et al. 2011, Sadoti et al. 2013, and Chapter 1). The output of these models include regression coefficients that can be used to predict occupancy at sites that are not surveyed based

on the measured covariates at those unsurveyed sites. Regression coefficients can be used in predictive models to evaluate current and future distributions of species (Kéry et al. 2013).

Measurement error in covariates used in linear and non-linear models can have two major effects on our interpretation of the models (Carroll et al. 2006). First, covariate error can cause estimators of regression coefficients from the models to be biased (Cochran 1968). Second, the statistical model can lose precision and we can lose ability to discriminate among alternative models, therefore reducing our ability to correctly identify the relationships of the variables (Tosteson et al. 2003). Biases could lead to incorrect conclusions and negatively affect conservation efforts based on such models. For example, with inaccurate covariate estimates, a predictive model of habitat use and landscape occupancy may provide inaccurate predictions leading to mismanaged habitat. However, there have been few studies specifically quantifying how serious this problem is for realistic levels of misclassification.

Using simulations, I evaluated the effects of land cover misclassification on regression coefficients in binomial generalized linear models (GLMs). I had three primary objectives. First I evaluated known misclassification estimates and known temporal changes in land cover data, and determined if biases occur in occupancy model regression coefficients. Second, I attempted to determine what level of land cover misclassification resulted in the true regression coefficient of interest being omitted from the 95% confidence interval. Third, I wanted to evaluate predictive models that incorporate regression coefficients from occupancy models that use error laden land cover data.

METHODS

The East Gulf Coastal Plain (EGCP) is located in the southeastern United States and was used to produce realistic error and habitat proportions for this study. The EGCP is home to many different types of habitats that are rare in their abundance but are considered ecologically important for many species (GCPO LCC. 2009). Land cover of the EGCP is composed of 56% forest, 12% shrub, 10% cultivated crops, 10% hay/pasture, 7% developed, 4% grassland, and 1% water (Homer et al. 2015). I clipped the 2011 NLCD to the EGCP land area using ArcMap 10.2 to measure region specific land cover composition change and land cover class change. I also measured land cover proportions in the EGCP so that the land cover covariates used in this analysis closely reflected reality. I used ArcMap 10.2 to randomly sample 100 points in the EGCP and measured the proportion of each land cover class in a 1 km buffered area around each point. Using the 100 sample points, I calculated the means for 3 land cover proportions: grassland, the sum of forest (deciduous, evergreen, and mixed forests), and the sum of all other classes. Grassland land cover was chosen because it is found in low abundance but is considered a high conservation priority. Forest land cover was chosen because it if found to be very abundant in the EGCP.

I created a simulation in program R 3.0.1 (R Core Team 2014) that simulated occupancy data using 2 land cover proportions as covariates that were manipulated to simulate the effects of land cover classification error. The occupancy model had 3 parameters and 2 covariates. The simulation ran for 1,000 iterations with 500 sites and 1 observation period for each of the 12 scenarios used in this study.

Habitat Covariate Generation

I generated habitat proportions (*d*) at each of the 500 sites (*i*) using three single trial random draws from beta distributions (d_{i1} = (α = 1, β = 8), d_{i2}= (α = 4, β = 6), d_{i3}= (α = 4, β = 6)). The shape parameters for the beta distributions were based on the actual land cover data described above. The d_{i1} distribution represents a land cover class with small proportions in the landscape (i.e. grasslands in the EGCP), d_{i2} represents land cover with large proportions in the landscape (i.e. forests in the EGCP), and d_{i3} represents the proportion all other land cover types. The values of these 500 habitat proportion sets were constrained to sum to 1 by normalizing the proportion of habitat in each type at each site so that

$$\frac{d_{ij}}{\sum_{j=1}^{3} d_{ij}} = k_{ij} \tag{1}$$

where $k_{i,j}$ is the normalized proportion of habitat type j at site i. I subsequently created a vector (*s*)

$$s_i = \begin{matrix} k_{i1} \\ k_{i2} \\ k_{i3} \end{matrix} \tag{2}$$

so that *s* is the vector of 3 land cover proportions measured from a land cover map. I created an error matrix (*T*) for each scenario (*c*) so that

$$\begin{bmatrix} g_g & g_f & g_o \\ f_g & f_f & f_o \\ o_g & o_f & o_o \end{bmatrix} = T_c \tag{3}$$

where g_g is the proportion of k_{i1} that is correctly classified as k_{i1}, f_g is the proportion of k_{i1} that is incorrectly classified as k_{i2}, o_g is the proportion of k_{i1} that is incorrectly classified as k_{i3}, g_f is the proportion of k_{i2} that is incorrectly classified as k_{i1}, f_f is the proportion of k_{i2} that is correctly classified as k_{i2}, o_f is the proportion of k_{i2} that is incorrectly classified as k_{i3}, g_o is the proportion

of k_{i3} that is incorrectly classified as k_{i1}, f_o is the proportion of k_{i3} that is incorrectly classified as k_{i2}, o_o is the proportion of k_{i3} that is correctly classified as k_{i3}. This matrix represents the misclassification found in the land cover data in each simulation scenario. All values in T were drawn from a beta distributions that used land cover accuracy values as a mean and a standard deviation of 0.05 to create alpha and beta shape parameters for each distribution.

I examined 12 scenarios where error was removed form land cover covariates (Table 2.1). Scenario A had no misclassification error, scenario B used the error matrix from a user's accuracy assessment of the conterminous United States from the 2006 NLCD (Wickham 2013) and scenarios C – G varied the accuracy of correctly identified land cover in increments of 10% from 0.9 to 0.5 with incorrectly identified land cover being evenly split between the other two land cover covariates. Scenarios H – L used the proportions of land cover change from 2006 to 2011 NLCD in 1 year intervals, for example scenario H is 1/5 and scenario I is 2/5 of the total proportions of land cover change. I then multiplied A and s to create a vector (e)

$$T_c \times s_i = e_{ic} = \begin{matrix} x_{ic1} \\ x_{ic2} \\ x_{ic3} \end{matrix} \tag{4}$$

so that e is the vector of 3 estimated land cover proportions (x) at each site at the time of survey, given the accuracy of each scenario. For example, scenario A had no misclassification, i.e., $g_1=f_2=o_3=1$ and $g_2=g_3=f_1=f_3=o_1=o_2=0$, then

$$e_{iA} = s_{i.} \tag{5}$$

Habitat relationship parameter values

I defined a vector (b)

$$b_v = \begin{matrix} h_{v1} \\ h_{v2} \\ h_{v3} \end{matrix} \tag{6}$$

44

of parameter values, (h_1, h_2, and h_3) were the true regression coefficients for habitat associations used to simulate species occupancy in relation to habitat proportions at each of the 500 points. Specifically h_1 was the true intercept, h_2 was the true regression coefficient for proportion of grassland, and h_3 was the true regression coefficient for proportion of forest. One vector (b_1) contained true regression coefficients that gave grasslands relatively weak negative relationships ($h_{12} = -2$), forests relatively strong positive relationships ($h_{13} = 6$), and the intercept was equal to zero ($h_{11} = 0$). The other vector (b_2) contained regression coefficients that gave grasslands relatively strong positive relationships ($h_{22} = 6$), forests relatively weak negative relationships ($h_{23} = -2$), and the intercept was equal to zero ($h_{11} = 0$).

Occupancy Simulation

The parameter values b were used in a logit link function that includes e_{ic} to generate true probability of occupancy (ψ_i) for each scenario.

$$\psi_i = \frac{\exp(h_1 + h_2 x_{ic1} + h_3 x_{ic2})}{1 + \exp(h_1 + h_2 x_{ic1} + h_3 x_{ic2})} \tag{7}$$

The third habitat proportion, other, was not included in this function to prevent over parameterization. I performed a Bernoulli trial for presence (p_i) that used ψ_i as the probability of presence for site i.

$$p_i \sim Bernoulli(\psi_i) \tag{8}$$

All absences were treated as true absences (i.e., the simulation did not have a detection process). The simulated presence/absence data (p) and land cover covariates (s) for each iteration of the simulation were incorporated into a binomial GLM to obtain regression coefficients that explain habitat relationships. Thus, I had 1.000 estimates of occupancy, regression coefficients

and standard error. The average of the 1,000 estimated regression coefficients for each

simulation was calculated along with the 95% confidence interval. I created a vector (u)

$$u_c = \begin{matrix} r_{c1} \\ r_{c2} \\ r_{c3} \end{matrix} \tag{9}$$

so that u is the vector of the 3 average estimated regression coefficients from the 1000 iterations

with scenario c. For example, r_{F3} describes the average estimated regression coefficient from

scenario F that corresponds to forest land cover proportions. The coefficient estimates were

compared to the values from the true underlying occupancy model. For example, if b_1 was used

then r_{F3} is compared to 6 and if b_2 was used then r_{F3} is compared to -2. The average of the 1,000

standard errors for each simulation was calculated to evaluate the precision of the estimated

regression coefficients. I also calculated the percentage of 1,000 iterations where the true

regression coefficient was omitted from the 95% confidence interval for each individual

parameter estimate in the simulation.

To put the changes in regression parameters estimates due land cover classification error

in an applied context I created comparative maps to examine the differences in predicted

occupancy under different land cover error scenarios. I used Program R to create predicted

occupancy maps for the scenarios using u in a logit link function to calculate ψ. The covariates

used to calculate ψ where the proportion of grasslands and forest within 1 km of each cell in the

2011 NLCD. These covariates were calculated similarly to the methods used to create the

covariates used in the simulations. The map was a small extent of the 2011 NLCD in the EGCP

to visualize the effects of the scenarios on predicted occupancy. I used a non-quantitative visual

comparison of predicted occupancy maps to examine the spatial arrangement of occupancy hot

spots (e.g., ψ > 0.8) among scenarios.

RESULTS

Land cover composition in the EGCP has changed 3.29% from 2006 to 2011. A total of 91.54% of the EGCP land cover in 2006 was the same class in 2011. Three habitat classes (shrub/scrub, evergreen forest, and grassland/herbaceous) account for 65.72% of the total land cover change in the EGCP. From 2006 to 2011 50.58% of the grassland/herbaceous land cover changed to a different class in the EGCP; 30.14% to shrub/scrub, 16.03% to evergreen.

I ran a total of 24 simulations, i.e., scenarios A-L with b_1 and scenarios A-L with b_2. All scenarios except for scenario A produced biased average regression coefficients. The most extreme bias of an estimated regression coefficient occurred in scenario L when the true parameter value for grassland was equal to -2 and the 95% confidence interval for the average estimated regression coefficient spanned zero (Figure 2.1). Biases increased for all estimated regression coefficients as the time since land cover classification increased (e.g., scenarios H – L). However, biases increased more greatly for the estimated regression coefficients of grassland than they did for forest (Figure 2.2, 2.3, 2.4). The strength and direction of the bias experienced by the intercepts were similar for b_1 and b_2 in scenarios C – G (Figure 2.5 and 2.6). However, the strength and direction of the bias experienced by the intercepts for b_1 in scenarios I – L was more extreme and in opposite directions than the bias experienced by the intercepts for b_2.

For parameter set b_1, where the intercept is equal to 0, the grassland parameter is equal to -2 and the forest parameter is equal to 6, the true regression coefficient was outside the 95% confidence intervals for the forest estimate in scenarios E, F, and G (Figure 2.7). Also for b_1, the 95% confidence intervals for the grassland estimates for all scenarios did include the true regression coefficient; however, they all included zero (Figure 2.8). For parameter set b_2, where the intercept is equal to 0, the grassland parameter is equal to 6 and the forest parameter is equal

to -2, the true regression coefficient was outside the 95% confidence intervals for the grassland estimate in scenarios E, F, G, K, and L (Figure 2.9). Furthermore, for b_2, the 95% confidence intervals for the forest estimates for all scenarios did include the true regression coefficient; however, the 95% confidence intervals of scenarios D, E, F, and G did include zero (Figure 2.10). All scenarios with parameter set b_1 provided 95% confidence intervals that bracketed the true intercept (Figure 2.11). The 95% confidence interval for the intercept of scenario G with parameter set b_2 did not include zero which also means it did not include the true intercept (Figure 2.12).

In summary, for parameter set b_1, all scenarios gave 95% confidence intervals that included zero for the grassland parameter estimate, and the true regression coefficient was found in all 95% confidence intervals created from the average standard error of scenarios A, B, C, D, H, I, J, K, and L. For parameter set b_2 scenarios A, B, C, H, I, and J produced 95% confidence intervals created from the average standard error, which included the true regression coefficients and did not include zero for the grassland estimate and forest estimate.

The true parameter value for forest (6) was omitted from >50% of the 95% confidence intervals estimated in the 1,000 iterations for simulations of scenarios E – G (Table 2.2). The true parameter value for grassland (6) was omitted from >50% of the 95% confidence intervals estimated in the 1,000 iterations for simulations of scenarios E – G, K, and L (Table 2.3).

The predicted occupancy maps visualized the change in regression coefficients from each scenario in a specific location. Scenario A produced a map that predicted occupancy when the land cover map correctly classifies the land cover at the time of animal sampling (Figure 2.13) and serves as the most accurate prediction of animal occupancy. The map for scenario J predicted occupancy when the land cover map was accurately classified 3 years before the time

48

of animal sampling (Figure 2.14). The map for scenario E predicted occupancy when mapping accuracy is 70% (i.e., 70% of pixels are correctly classified) for all habitat classes in the map extent (Figure 2.15). The three scenarios show differing occupancy maps of the same geographic area. I did not conduct quantitative comparisons of the maps, but instead used the maps for visual comparisons of the spatial arrangement of occupancy hot spots. Compared to the map from scenario A, the map from scenario J loses the areas with very high ($\psi > 0.8$) occupancy probability. That is, when the land cover was three years out of data, probable occupancy was under estimated for areas with high occupancy probability. While the map from scenario E retains very similar areas of high occupancy probability ($\psi > 0.8$), the low and intermediate areas of occupancy probability ($0.2 < \psi < 0.6$) appear to vary between the maps from scenario A and E (Figure 2.13, 2.15). That is, when 30% of the pixels are misclassified, occupancy probability was under estimated areas with low to moderate occupancy.

DISCUSSION

   This simulation study provided evidence that low to moderate levels of misclassification in land cover data can affect occupancy estimation and assessment of habitat relationships. Simulations run with scenario B produced average regression coefficients that were bias (e.g., did not include the true parameter value). Simulations run with scenarios C – G, which specified accuracy of the land cover classes to be 90 – 50%, also produced average regression coefficients that were bias. Simulations run with scenarios H – L (i.e., land cover change in one year intervals) produced average regression coefficients that were bias; however, the grassland regression coefficients had more extreme bias than the forest coefficients.

49

I found the accuracy of the specific land cover class of interest was not the only determinate of regression coefficient bias when using land cover proportions. The accuracy of the other land cover classes will affect the regression coefficient by influencing the proportion of land cover of interest. For example, if true land cover in a specified area is composed of 50% grassland and 50% forest and 15% of grassland is misclassified as forest and 15% of forest is misclassified as grassland then grassland and forest proportions stay the same. This is most likely the reason why with parameter set b_2 (e.g. Intercept = 0, Grassland = 6, Forest = -2) the mean regression coefficients for scenario E and J differ, even though the class accuracy for grassland is similar (i.e., 70% and 69.7% respectively). Scenario E had 15% of the high proportion of forest and other be reclassified as grassland while scenario J had < 3% of forest and other be reclassified as grassland. Grassland in scenario E had a large net gain while grassland in scenario J had a very small net loss. Many studies have found there is a positive relationship of land cover class abundance and accuracy (Wickham et al. 2004, Wickham 2010). In this study I used the accuracy of grassland land cover (i.e. a rarity in the EGCP) and forest land cover (i.e. common in the EGCP) to guide the creation of scenarios that described misclassification. I did not create scenarios that showed a negative relationship of land class abundance and accuracy, because of this, there is difficulty in describing the effects of land cover proportion separately from accuracy. That is, it is difficult to distinguish whether the parameter estimate biases in my simulation study were due to the fact that there was error added to the land cover proportions or if because when the proportions of a specific habitat type are quite small to start with, small changes might have larger effects on parameter estimates. To separate theses effects I performed additional analyses using 50% and 90% accuracy for both land cover proportions to gain preliminary results that would indicate the effect of starting land cover proportion and land cover

50

accuracy. With this post hoc partial assessment, results indicated that land cover accuracy affects regression coefficient bias and habitat proportion affects the variance of the regression coefficient. This is of concern since only 7 of the 16 land cover classes in the 2006 NLCD exceed 80% user accuracy (Wickham 2013). These additional analyses suggest that when a land cover type of interest is limited or rare in the study area, the effects of misclassification error in remotely sensed data might be a greater concern.

In the EGCP land cover changes frequently for 3 land cover classes (shrub/scrub, evergreen forest, and grassland/herbaceous). This is most likely due to timberland clearcutting and the long growing season that allows for quicker successional changes. Evergreen forests are cut down for industry and an open patch is left where grasses, vines, and woody shrubs grow. The open patch is eventually replanted with trees or natural regeneration will occur and trees will grow. The patch then becomes an evergreen forest again. This study used 5 scenarios that had estimated land cover change data from a freely available and widely used land cover product. All of the scenarios produced average regression coefficients that were bias. The most extreme bias occurred in the grassland parameter estimates. This is most likely because grasslands in the EGCP experience relatively high rates of land cover change. The rates of these changes explored here may be much more predictable with silviculture models that are able to estimate young tree height for each year (Zhang et al. 1996). Researchers could incorporate these known rates into land cover covariate estimates used in occupancy or other habitat relationship research. However, ecological time lag could counteract the effects of the highly dynamic landscape in the southeast, by allowing species to persist in an area for several years after land cover change (Metzger et al. 2009).

I viewed this simulation study as the first step in developing an understanding of how misclassification might affect occupancy covariate analyses and future simulation analyses might expand their assessment to look at more than two habitat classes. Combining 16 habitat classes into 3 simplified the simulation and results, allowing for easier interpretation. Furthermore, I used only land cover proportions as covariates in the simulation to keep the simulation relatively simple for interpretation. However, land cover proportions are not the only remotely sensed land cover measurements used in occupancy modeling (Prugh et al. 2008). Habitat pattern metrics, which consider spatial patterns of land cover, are also affected by misclassification and the other errors in land cover data described in this study. The error found in land cover data has spatial properties and is not randomly distributed across the land cover map. Knowing the spatial distribution of error may allow insight into the causes of error and improve our understanding of how error affects land cover metrics. There are known correlations between habitat abundance and several habitat pattern metrics (Wang et al. 2014). Burnicki (2012) found error in land cover change maps impacted the number of patches and mean nearest neighbor metrics, while having little impact on patch shape and size metrics. It may be sensible to continue this research with covariates that measure spatially referenced land cover data such as total core area, mean nearest-neighbor distance, and total edge.

Management Implications

Habitat use and occupancy models are important tools used by researchers to understand the relationships of species and the habitats in the landscape (Morrison et al. 2012). However, when the habitat data used in models has known errors, it is up to the researchers to determine if the error in the habitat data will affect the interpretation and results of the model. Here, I

demonstrated with simulated data how errors in remotely sensed land cover data can affect the results from a binomial GLM that evaluates habitat relationships. The regression coefficients that are used to predict occupancy can be easily biased with error in land cover data. The biases in the scenarios studied here weakened the habitat relationships adding ambiguity to the occupancy predictions. The predicted occupancy map for scenario A, no error, indicated several patches with predicted occupancy >80%. However when using the current presence/absence data with a 3 year old map (i.e., scenario J) under the assumption that the species is occupying the landscape based on the current land cover and not the land cover in the 3 year old map, the patches of predicted occupancy >80% disappeared. If highly mobile species can change their occupancy in the landscape at the same rate as land cover change then estimations of habitat relationships could be incorrect when using older land cover data. A decision maker that was given the map of scenario A would most likely prioritize the northeast portion for conservation efforts. If the decision maker was give the map of scenario J they may be more reluctant to prioritize conservation efforts in the northeast even though the true predicted occupancy is much greater than shown. Without a rigorous quantitative analysis of a broad spatial areas I cannot conclude that this pattern will hold true beyond the small area that I mapped, but I can infer that predicted occupancy has the potential to change on the landscape which could affect management decisions with respect to spatial arrangement of habitat acquisition, restoration or other habitat management decisions. For effective conservation actions, decision makers may need consider the accuracy of the remote sensing data used in models that create the regression coefficients that will be incorporated into management decisions.

Literature Citied

Burnicki, A. C. 2012. Impact of error on landscape pattern analyses performed on land-cover change maps. Landscape Ecology 27:713–729.

Carroll, R.J., D. Ruppert, L.A. Stefanski, and C.M. Crainiceanu. 2006. Measurement error in nonlinear models: a modern perspective. CRC press.

Cochran, W.G., 1968. Errors of measurement in statistics. Technometrics, 10(4), pp.637-666.

Duren, K. R., J. J. Buler, W. Jones, and C. K. Williams. 2011. An improved multi-scale approach to modeling habitat occupancy of northern bobwhite. Journal of Wildlife Management 75:1700–1709.

Gallant, A. L. 2009. What you should know about land-cover data. Journal of Wildlife Management 73:796–805.

GCPO LCC. 2009. Gulf Coastal Plains and Ozarks Landscape Conservation Cooperative development and operations plan. www.fws.gov/southeast/LCC/GulfPlains/pdf/ GCPOLCCDevOpPlanFinal12112009.pdf. Accessed 16 June 2016.

Glenn, E. M., and W. J. Ripple. 2004. On using digital maps to assess wildlife habitat. Wildlife Society Bulletin 32:852–860.

Greenwald, K. R., J. L. Purrenhage, and W. K. Savage. 2009. Landcover predicts isolation in *Ambystoma* salamanders across region and species. Biological Conservation 142:2493–2500.

Homer, C.G., J.A. Dewitz, L. Yang, S. Jin, P. Danielson, G. Xian, J. Coulston, N.D. Herold, J.D. Wickham, and K. Megown. 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 81:345-354.

Horning, N., J.A. Robinson, E.J. Sterling, W. Turner, S. Spector. 2010. Remote sensing for ecology and conservation: A Handbook of Techniques. Oxford University Press, New York, NY.

Kéry, M., G. Guillera-Arroita, and J. J. Lahoz-Monfort. 2013. Analysing and mapping species range dynamics using occupancy models. Journal of Biogeography 40:1463–1474.

Liu, C., P. Frazier, and L. Kumar. 2007. Comparative assessment of the measures of thematic classification accuracy. Remote Sensing of Environment 107:606–616.

Metzger, J. P., A. C. Martensen, M. Dixo, L. C. Bernacci, M. C. Ribeiro, A. M. G. Teixeira, and R. Pardini. 2009. Time-lag in biological responses to landscape changes in a highly dynamic Atlantic forest region. Biological Conservation 142:1166–1177.

Morrison, M.L., Marcot, B. and Mannan, W., 2012. Wildlife-habitat relationships: concepts and applications. Island Press.

Oswalt, S.N., W.B. Smith. 2014. U.S. forest resource facts and historical trends. U.S. Forest Service, Washington, D.C., USA.

Prugh, L. R., K. E. Hodges, A. R. E. Sinclair, and J. S. Brashares. 2008. Effect of habitat area and isolation on fragmented animal populations. Proceedings of the National Academy of Sciences of the United States of America 105:20770–5.

R Core Team. 2014. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/

Rose, R. A., D. Byler, J. R. Eastman, E. Fleishman, G. Geller, S. Goetz, L. Guild, H. Hamilton, M. Hansen, R. Headley, J. Hewson, N. Horning, B. A. Kaplin, N. Laporte, A. Leidner, P. Leimgruber, J. Morisette, J. Musinsky, L. Pintea, A. Prados, V. C. Radeloff, M. Rowen, S. Saatchi, S. Schill, K. Tabor, W. Turner, A. Vodacek, J. Vogelmann, M. Wegmann, D. Wilkie, and C. Wilson. 2015. Ten ways remote sensing can contribute to conservation. Conservation Biology 29:350–359.

Sadoti, G., B. Zuckerberg, M. A. Jarzyna, and W. F. Porter. 2013. Applying occupancy estimation and modelling to the analysis of atlas data. Diversity and Distributions 19:804–814.

Shao, G., and J. Wu. 2008. On the accuracy of landscape pattern analysis using remote sensing data. Landscape Ecology 23:505–511.

Sleeter, B. M., T. L. Sohl, T. R. Loveland, R. F. Auch, W. Acevedo, M. a. Drummond, K. L. Sayler, and S. V. Stehman. 2013. Land-cover change in the conterminous United States from 1973 to 2000. Global Environmental Change 23:733–748.

Steen, D. a, C. J. W. McClure, J. C. Brock, D. C. Rudolph, J. B. Pierce, J. R. Lee, W. J. Humphries, B. B. Gregory, W. B. Sutton, L. L. Smith, D. L. Baxley, D. J. Stevenson, and C. Guyer. 2012. Landscape-level influences of terrestrial snake occupancy within the southeastern United States. Ecological applications 22:1084–97.

Story, M., and R. G. Congalton. 1986. Accuracy assessment: A user's perspective. Photogrammetric Engineering and Remote Sensing, 52:397−399.

Tosteson, T. D., J. S. Buzas, E. Demidenko, and M. Karagas. 2003. Power and sample size calculations for generalized regression models with covariate measurement error. Statistics in Medicine 22:1069–1082.

Urbanek, R. E., and C. K. Nielsen. 2013. Influence of landscape factors on density of suburban white-tailed deer. Landscape and Urban Planning 114:28–36.

Vogelmann, J. E., T. L. Sohl, P. V Campbell, and S. D. M. 1998. Regional land cover characterization using Landsat Thematic Mapper data and ancillary data sources. Environmental Monitoring and Assessment 51:415–428.

Wang, X., F. G. Blanchet, and N. Koper. 2014. Measuring habitat fragmentation: An evaluation of landscape pattern metrics. Methods in Ecology and Evolution 5:634–646.

Wickham, J. D., S. V. Stehman, J. H. Smith, and L. Yang. 2004. Thematic accuracy of the 1992 national land-cover data for the western United States. Remote Sensing of Environment 91: 452-468.

Wickham, J. D., S. V. Stehman, J. A. Fry, J. H. Smith, and C. G. Homer. 2010. Thematic accuracy of the NLCD 2001 land cover for the conterminous United States. Remote Sensing of Environment 114:1286–1296.

Wickham, J. D., S. V Stehman, L. Gass, J. Dewitz, J. A. Fry, and T. G. Wade. 2013. Accuracy assessment of NLCD 2006 land cover and impervious surface. Remote Sensing of Environment 130:294–304.

Zhang, S., H. E. Burkhart, and R. L. Amateis. 1996. Modeling individual tree growth for juvenile loblolly pine plantations. Forest Ecology and Management 89:157–172.

TABLES AND FIGURES

Table 2.1. The scenarios (A-L) that were used in a simulation to evaluate the effects misclassification in land cover data has on regression coefficients from occupancy models. The values are means that were used, plus a standard deviation of 0.05, to create a beta distribution to draw an error value for the simulation.

Scenario	grassland to grassland	grassland to forest	grassland to other	forest to grassland	forest to forest	forest to other	other to grassland	other to forest	other to other
A	1.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00	1.00
B	0.75	0.03	0.22	0.01	0.91	0.08	0.03	0.08	0.89
C	0.90	0.05	0.05	0.05	0.90	0.05	0.05	0.05	0.90
D	0.80	0.10	0.10	0.10	0.80	0.10	0.10	0.10	0.80
E	0.70	0.15	0.15	0.15	0.70	0.15	0.15	0.15	0.70
F	0.60	0.20	0.20	0.20	0.60	0.20	0.20	0.20	0.60
G	0.50	0.25	0.25	0.25	0.50	0.25	0.25	0.25	0.50
K	0.90	0.04	0.06	0.01	0.98	0.01	0.00	0.01	0.99
L	0.80	0.08	0.13	0.01	0.97	0.02	0.00	0.02	0.98
M	0.70	0.11	0.19	0.02	0.95	0.03	0.00	0.02	0.97
N	0.60	0.15	0.25	0.03	0.93	0.04	0.00	0.03	0.96
O	0.49	0.19	0.32	0.04	0.91	0.05	0.00	0.04	0.96

Table 2.2. Percent of the 1,000 iterations for each simulation where the true parameter value was omitted from the 95% confidence interval. The true parameter values used in these simulations were $h_{21} = 0$, $h_{22} = -2$. $h_{23} = 6$.

Scenario	Intercept	Grassland	Forest
A	4%	6%	4%
B	9%	5%	10%
C	9%	7%	15%
D	12%	8%	34%
E	19%	10%	60%
F	32%	17%	87%
G	48%	22%	99%
H	5%	7%	7%
I	8%	10%	8%
J	7%	11%	8%
K	8%	16%	9%
L	7%	22%	12%

Table 2.3. Percent of the 1,000 iterations for each simulation where the true parameter value was omitted from the 95% confidence interval. The true parameter values used in these simulations were $h_{21} = 0$, $h_{22} = 6$. $h_{23} = -2$.

Scenario	Intercept	Grassland	Forest
A	6%	5%	6%
B	5%	33%	6%
C	6%	14%	7%
D	18%	38%	13%
E	32%	61%	26%
F	48%	85%	33%
G	62%	95%	46%
H	6%	9%	5%
I	6%	26%	8%
J	4%	46%	6%
K	6%	66%	8%
L	6%	87%	11%
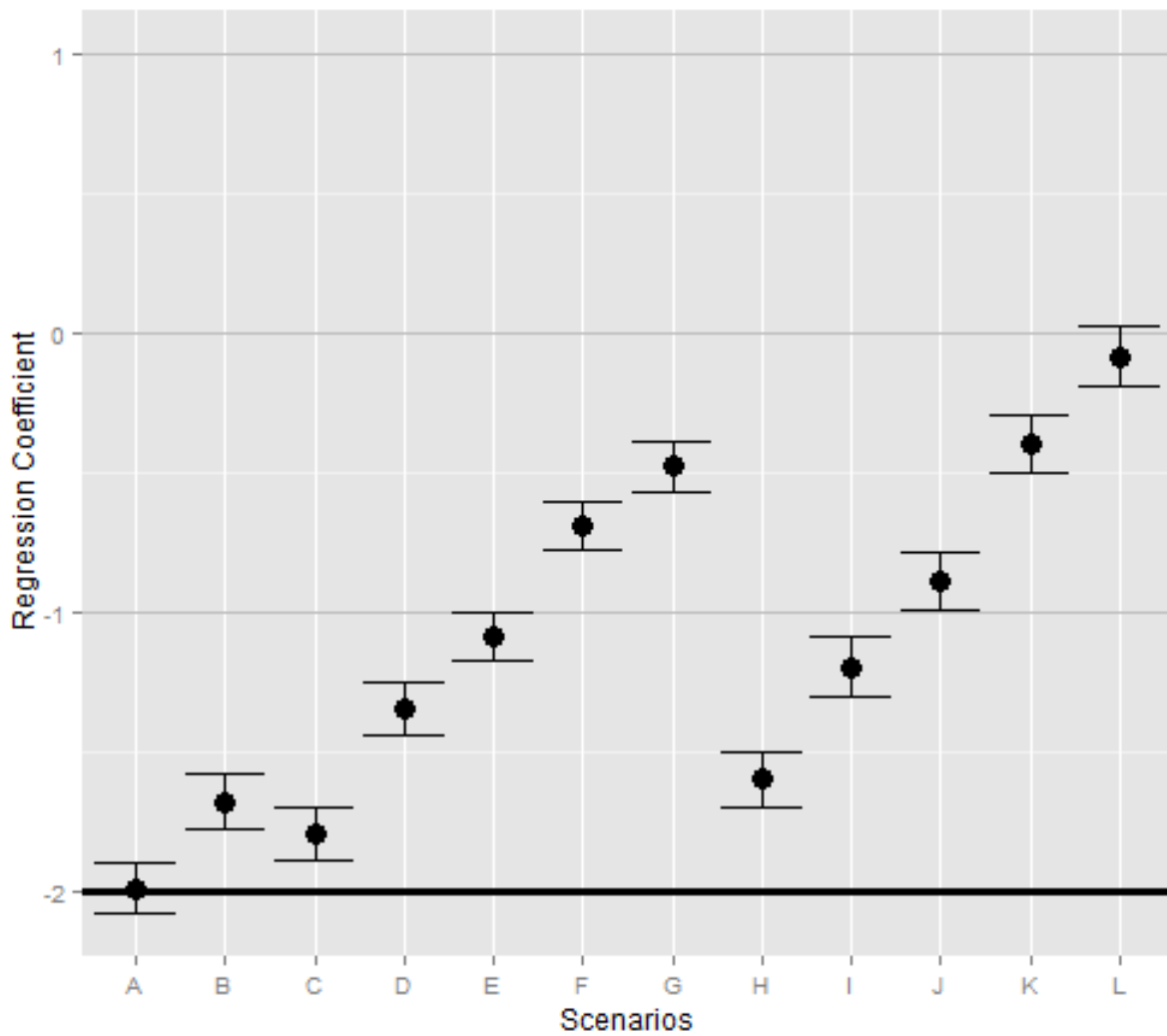
Figure 2.1. The average regression coefficients and 95% confidence intervals from 1000 replicates for the grassland parameter from binomial generalized linear regression using land cover proportions that have been changed to reflect different classification errors. The true parameter values used in these simulations were $h_{21} = 0$, $h_{22} = -2$. $h_{23} = 6$.
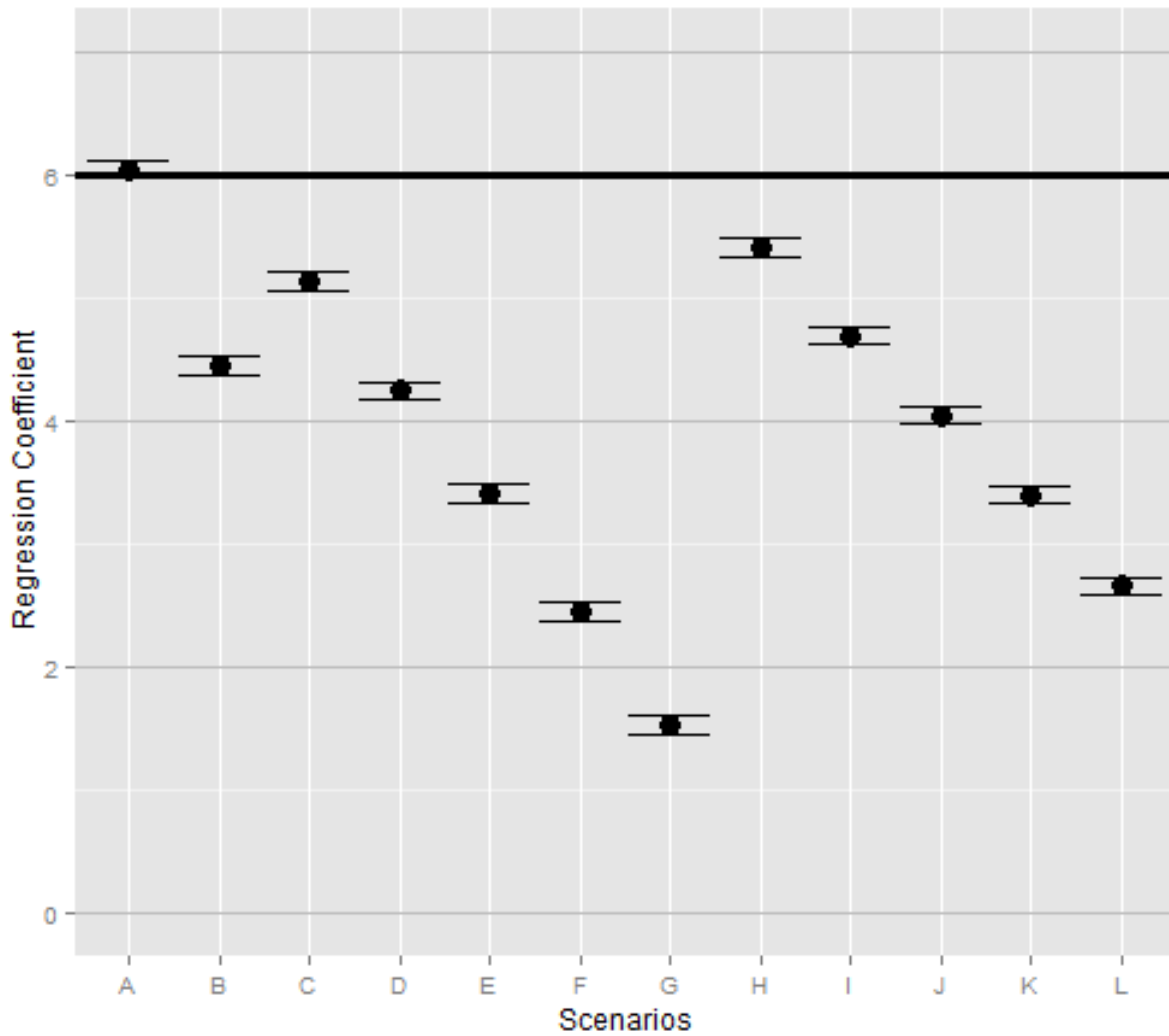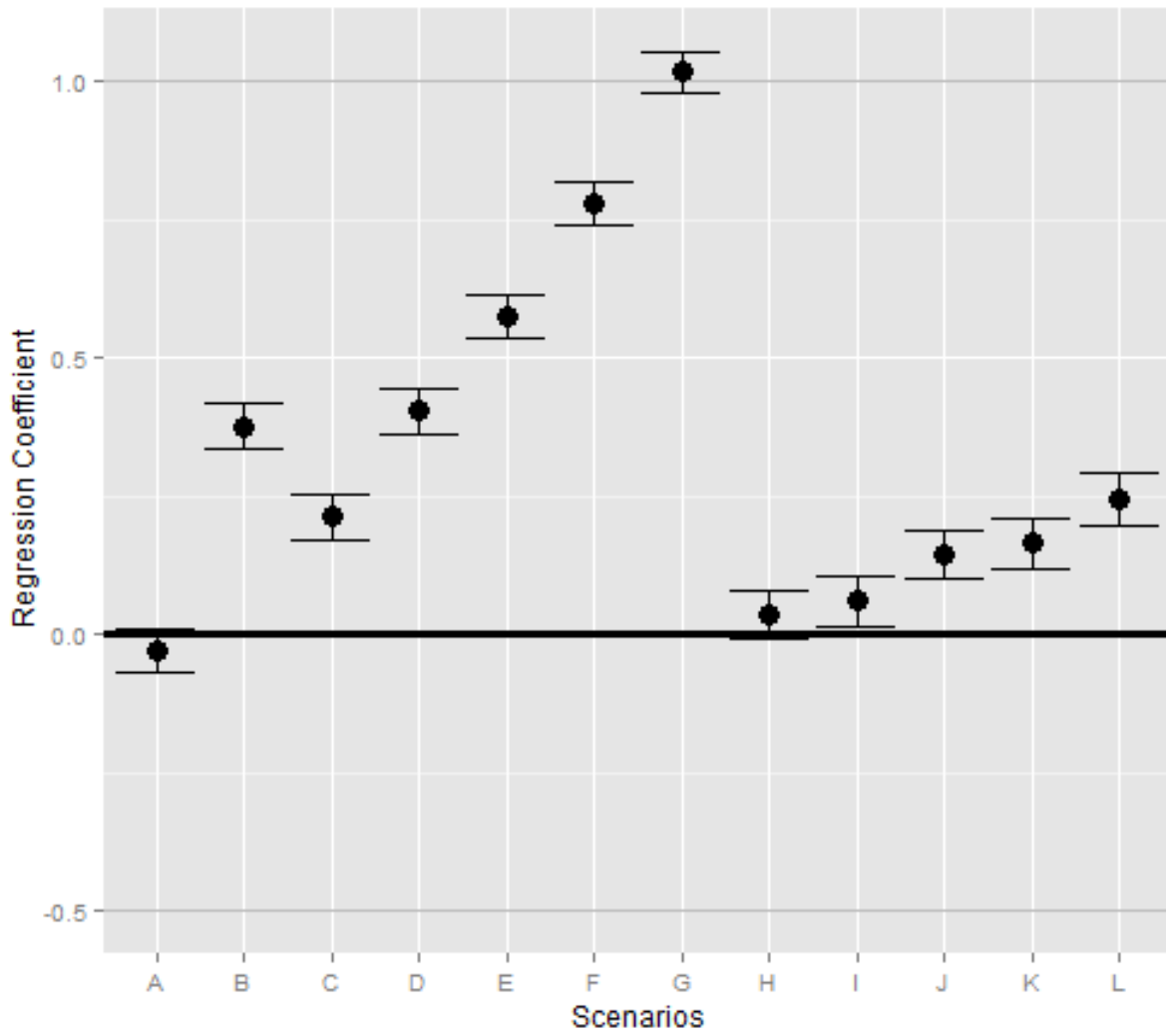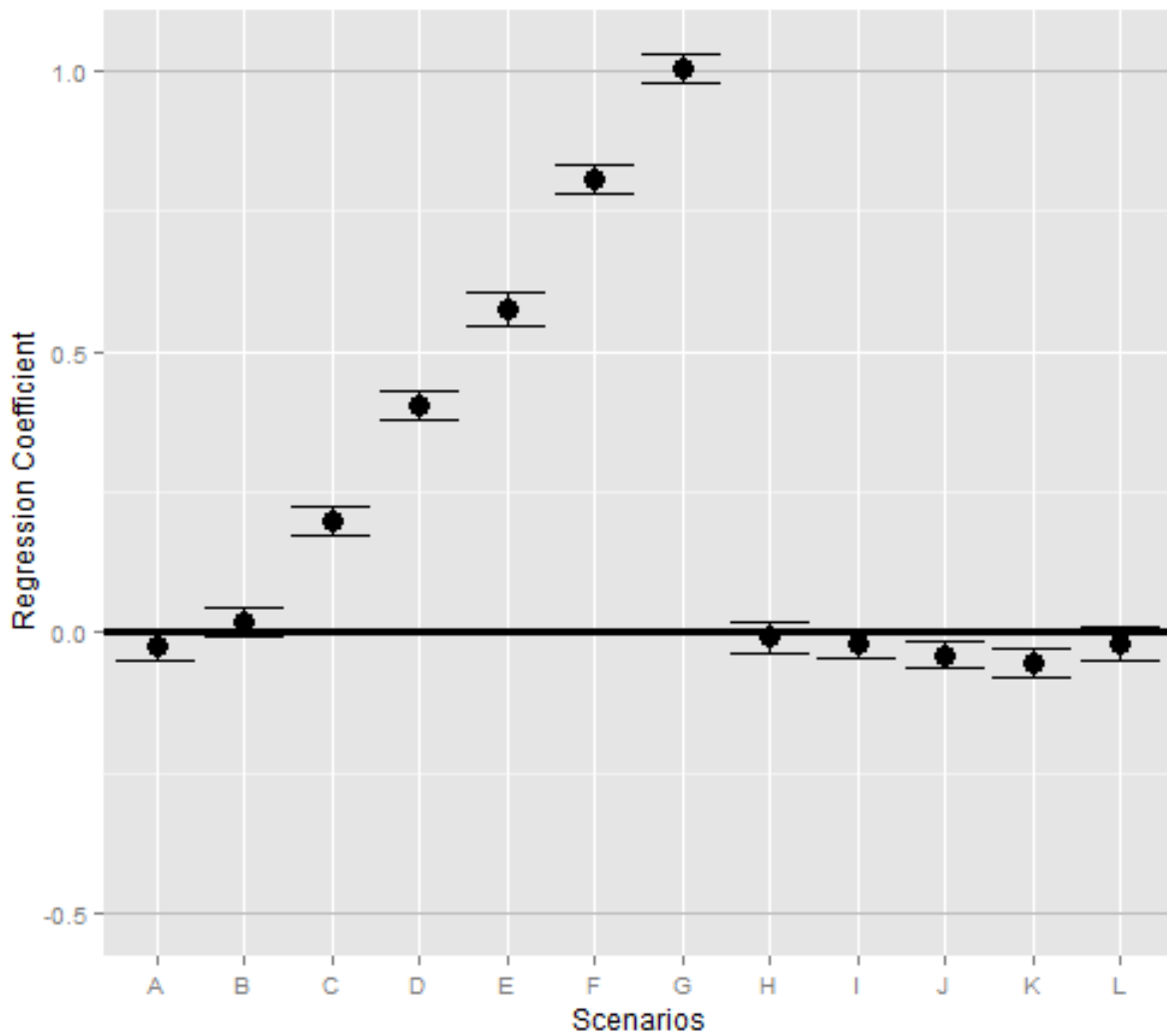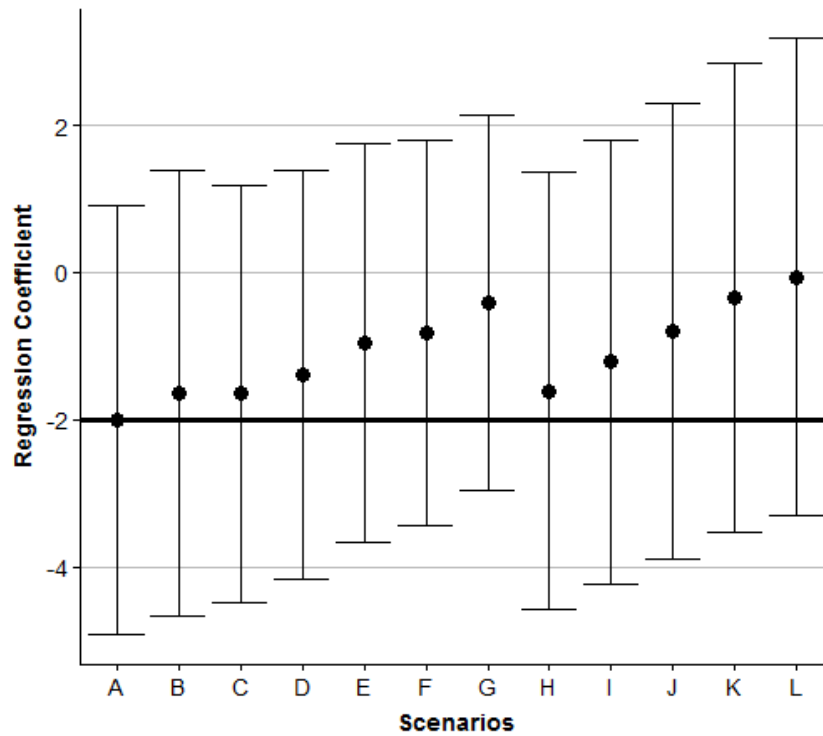
Figure 2.2. The average regression coefficients and 95% confidence intervals from 1000 replicates for the grassland parameter from binomial generalized linear regression using land cover proportions that have been changed to reflect different classification errors. The true parameter values used in these simulations were $h_{21} = 0$, $h_{22} = 6$. $h_{23} = -2$.
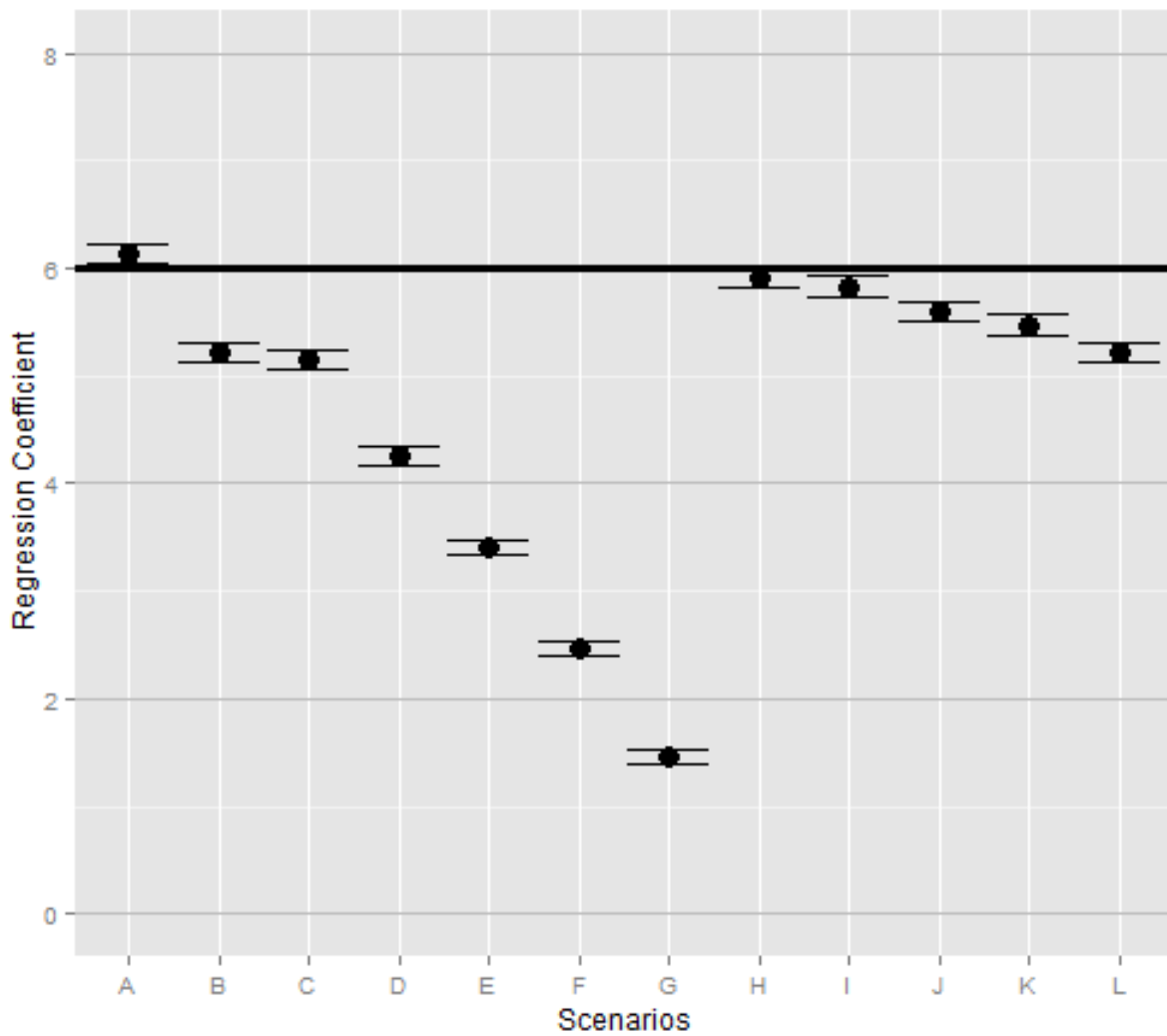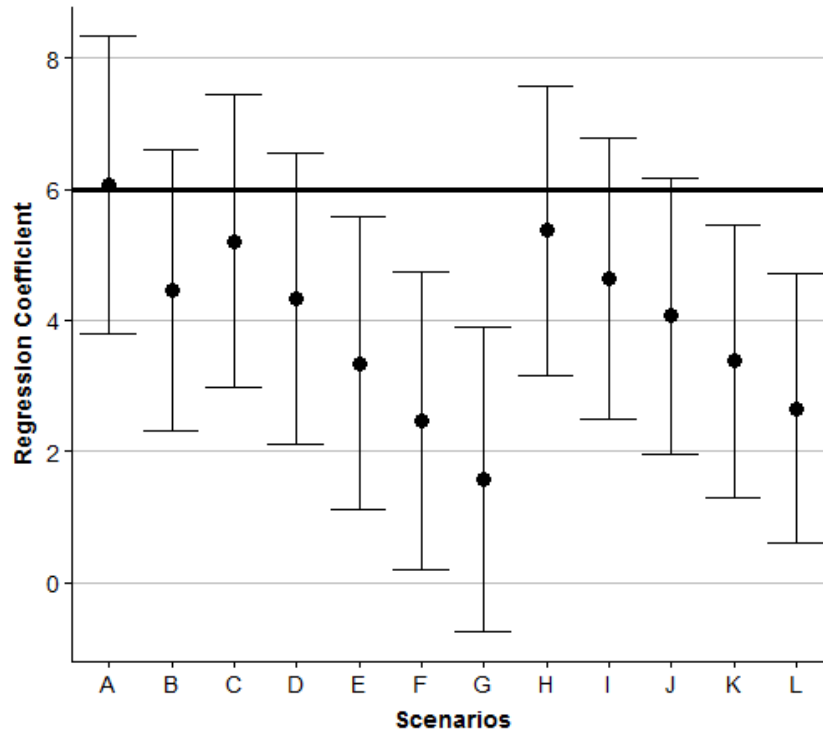
Figure 2.3. The average regression coefficients and 95% confidence intervals from 1000 replicates for the forest parameter from binomial generalized linear regression using land cover proportions that have been changed to reflect different classification errors. The true parameter values used in these simulations were $h_{21} = 0$, $h_{22} = -2$. $h_{23} = 6$.
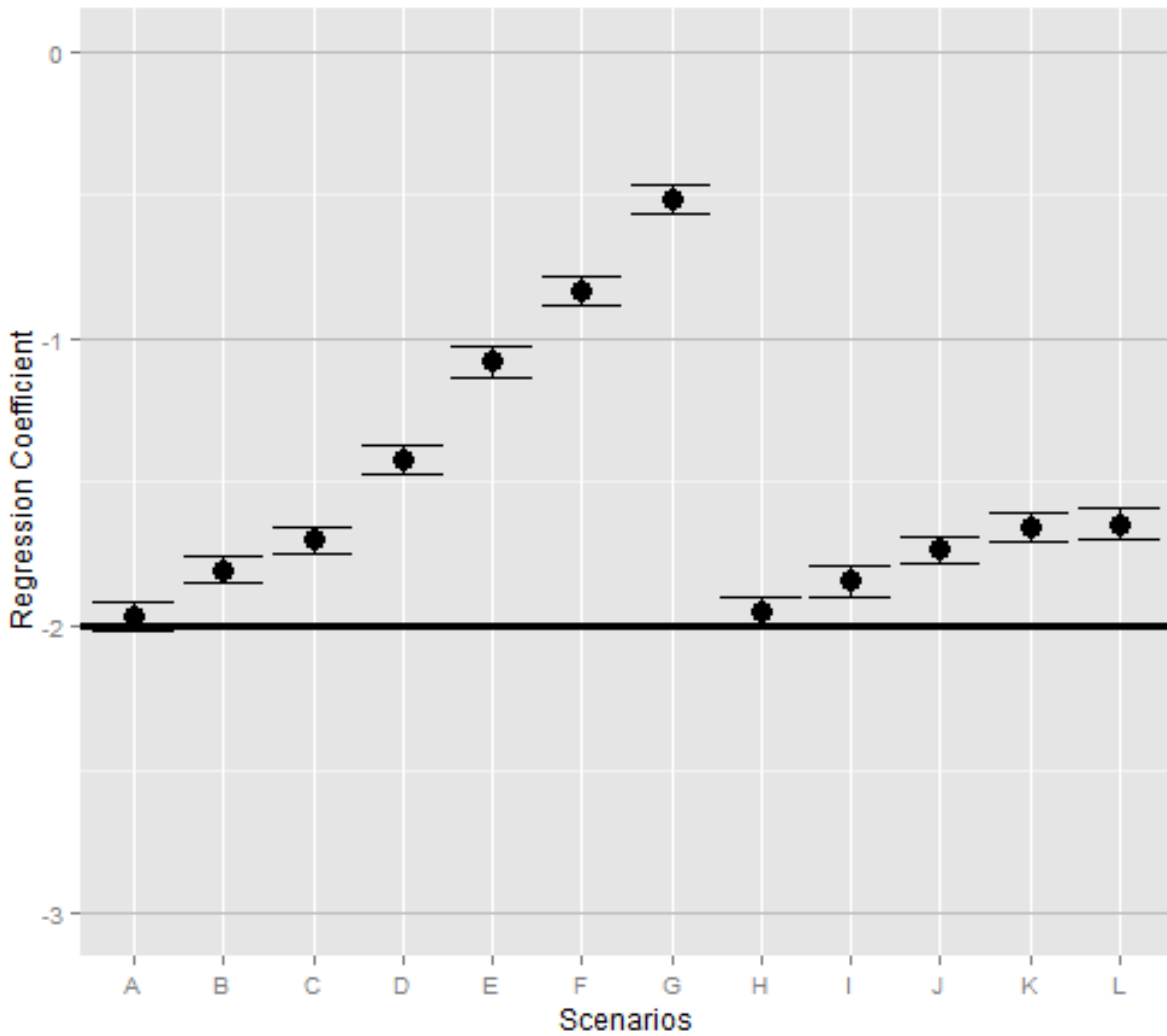
Figure 2.4. The average regression coefficients and 95% confidence intervals from 1000 replicates for the forest parameter from binomial generalized linear regression using land cover proportions that have been changed to reflect different classification errors. The true parameter values used in these simulations were $h_{21} = 0$, $h_{22} = 6$. $h_{23} = -2$.
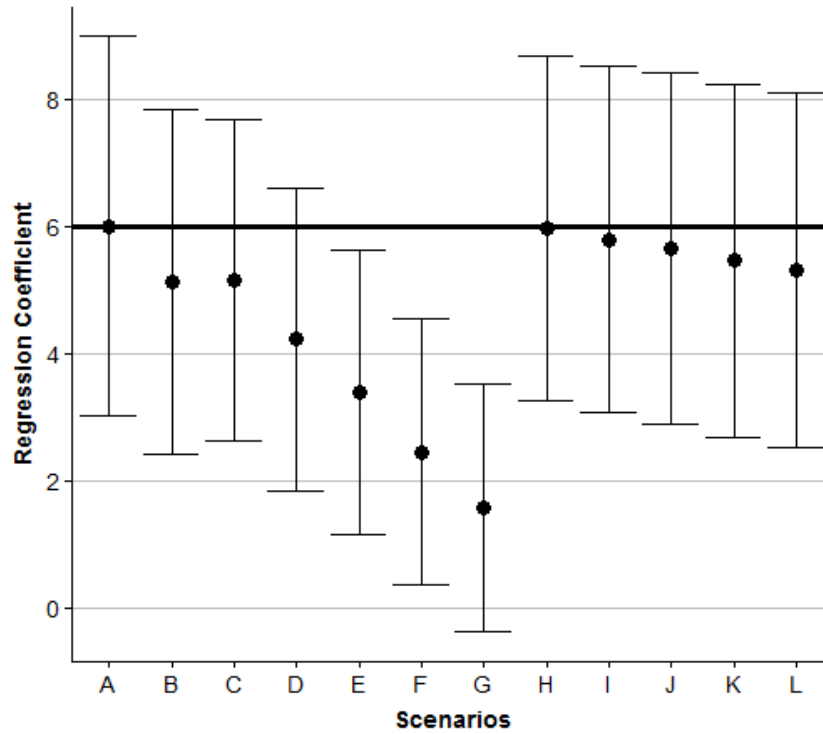
Figure 2.5. The average regression coefficients and 95% confidence intervals from 1000 replicates for the intercept from binomial generalized linear regression using land cover proportions that have been changed to reflect different classification errors. The true parameter values used in these simulations were $h_{21} = 0$, $h_{22} = -2$. $h_{23} = 6$.
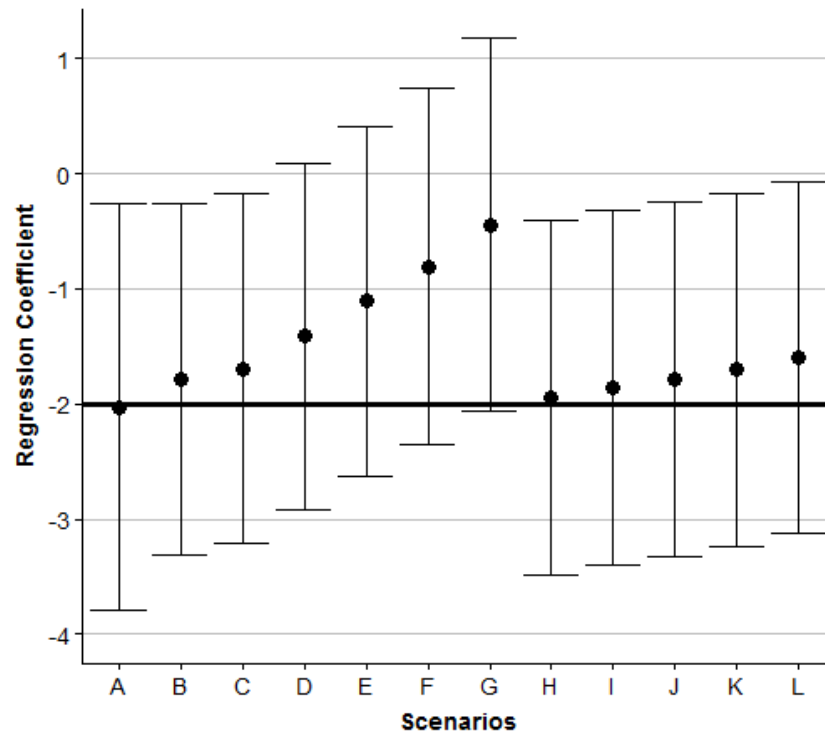
Figure 2.6. The average regression coefficients and 95% confidence intervals from 1000 replicates for the intercept from binomial generalized linear regression using land cover proportions that have been changed to reflect different classification errors. The true parameter values used in these simulations were $h_{21} = 0$, $h_{22} = 6$. $h_{23} = -2$.
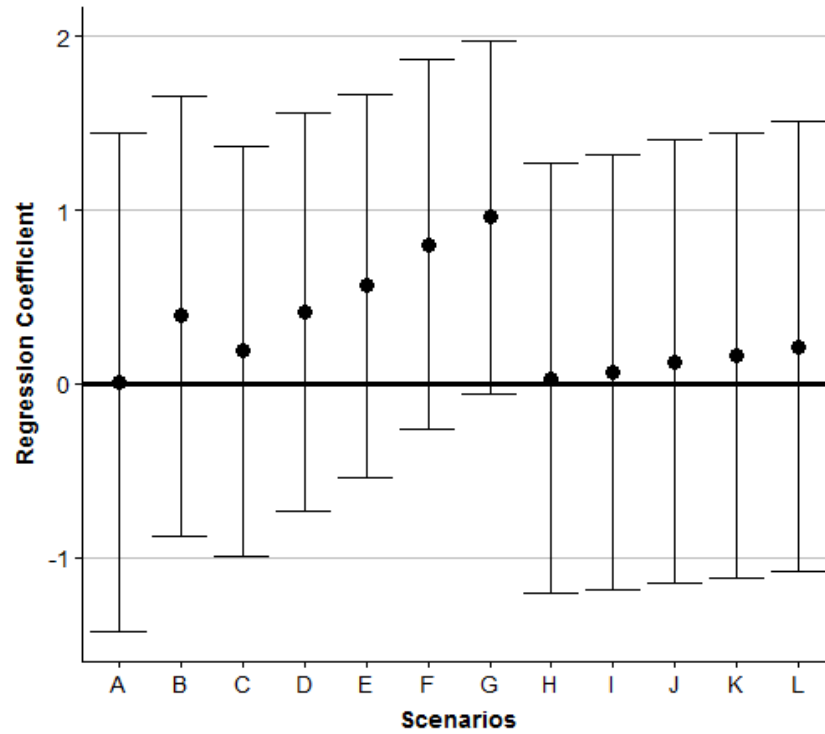
Figure 2.7. The 95% confidence intervals calculated from the average standard error from 1,000 iterations for the grassland parameter from binomial generalized linear regression using land cover proportions that have been changed to reflect different classification errors. The true parameter values used in these simulations were $h_{21} = 0$, $h_{22} = -2$. $h_{23} = 6$.

Figure 2.8. The 95% confidence intervals calculated from the average standard error from 1,000 iterations for the grassland parameter from binomial generalized linear regression using land cover proportions that have been changed to reflect different classification errors. The true parameter values used in these simulations were $h_{21} = 0$, $h_{22} = 6$. $h_{23} = -2$.
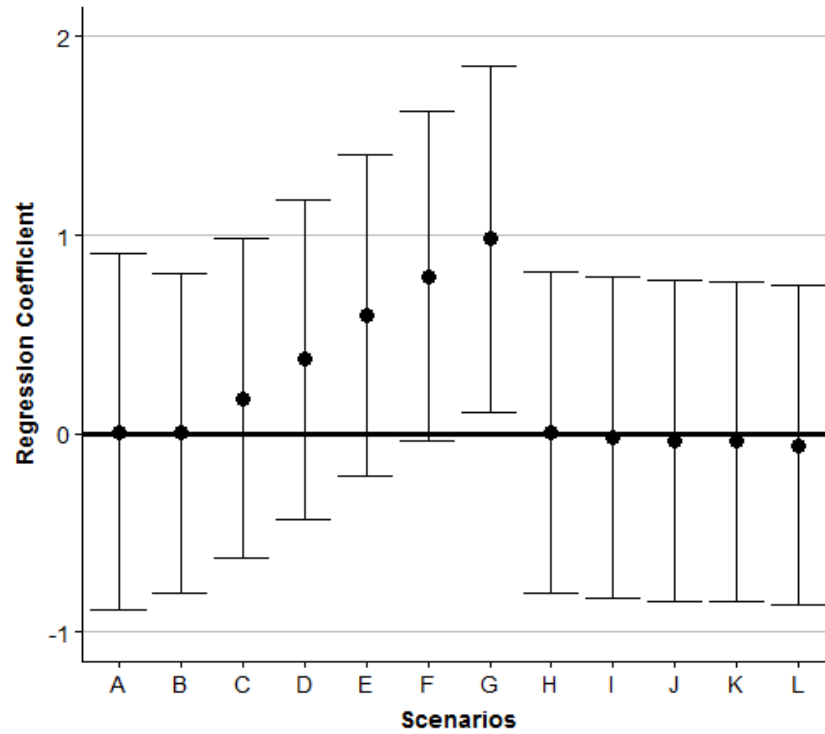
Figure 2.9. The 95% confidence intervals calculated from the average standard error from 1,000 iterations for the forest parameter from binomial generalized linear regression using land cover proportions that have been changed to reflect different classification errors. The true parameter values used in these simulations were $h_{21} = 0$, $h_{22} = -2$. $h_{23} = 6$.

Figure 2.10. The 95% confidence intervals calculated from the average standard error from 1,000 iterations for the forest parameter from binomial generalized linear regression using land cover proportions that have been changed to reflect different classification errors. The true parameter values used in these simulations were $h_{21} = 0$, $h_{22} = 6$. $h_{23} = -2$.

Figure 2.11. The 95% confidence intervals calculated from the average standard error from 1,000 iterations for the intercept from binomial generalized linear regression using land cover proportions that have been changed to reflect different classification errors. The true parameter values used in these simulations were $h_{21} = 0$, $h_{22} = -2$. $h_{23} = 6$.

Figure 2.12. The 95% confidence intervals calculated from the average standard error from 1,000 iterations for the intercept from binomial generalized linear regression using land cover proportions that have been changed to reflect different classification errors. The true parameter values used in these simulations were $h_{21} = 0$, $h_{22} = 6$. $h_{23} = -2$.
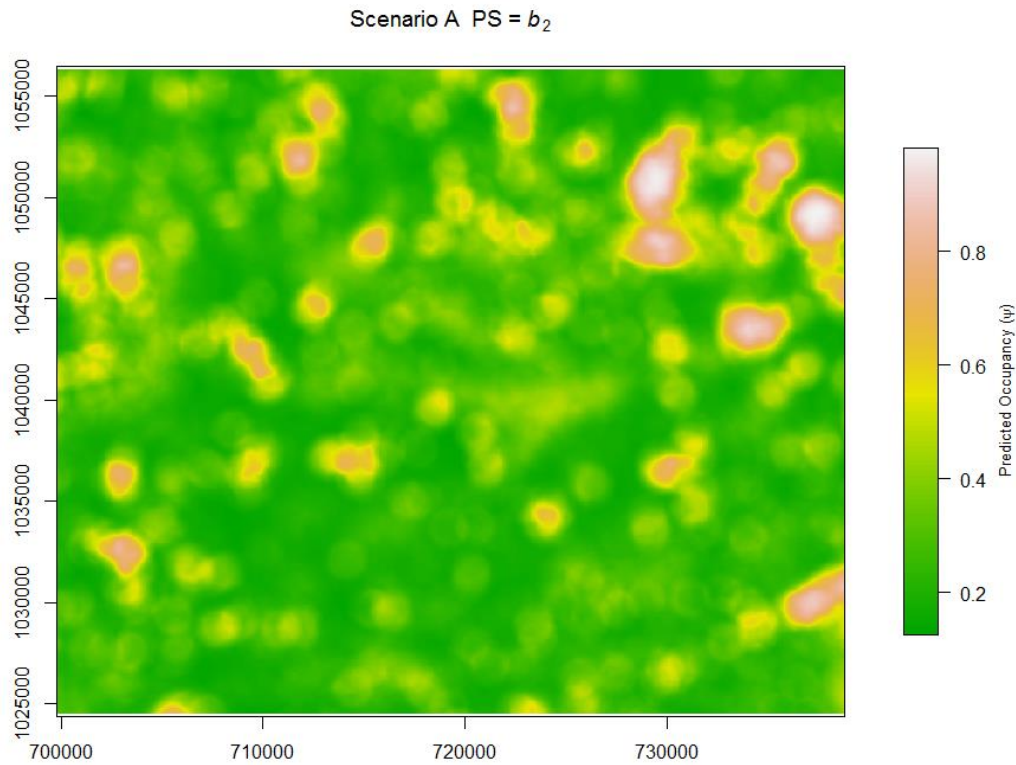
Figure 2.13. A map that visualizes the predicted probability of occupancy in a geographic space clipped from the 2011 NLCD in the EGCP using parameter set b2, under scenario A. Scenario A occurs when the land cover map has no error.
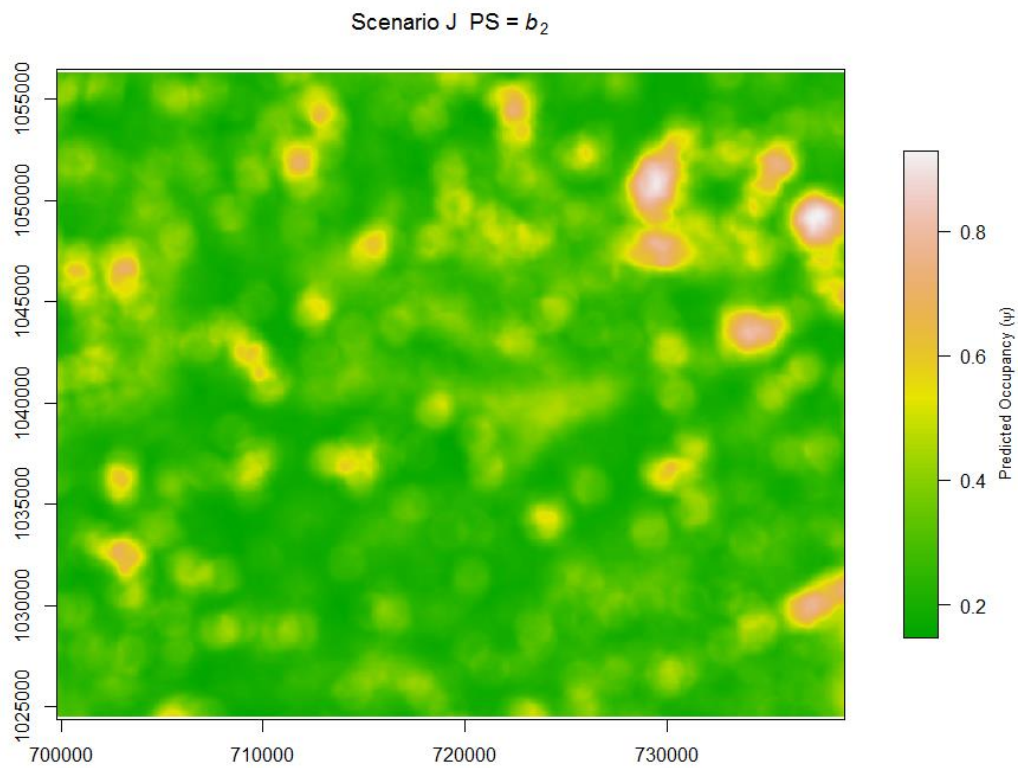
Figure 2.14. A map that visualizes the predicted probability of occupancy in a geographic space clipped from the 2011 NLCD in the EGCP using parameter set b2, under scenario J. Scenario J occurs when current presence/absence data is modeled with a map created 3 years prior that does not take into account land cover change over time.
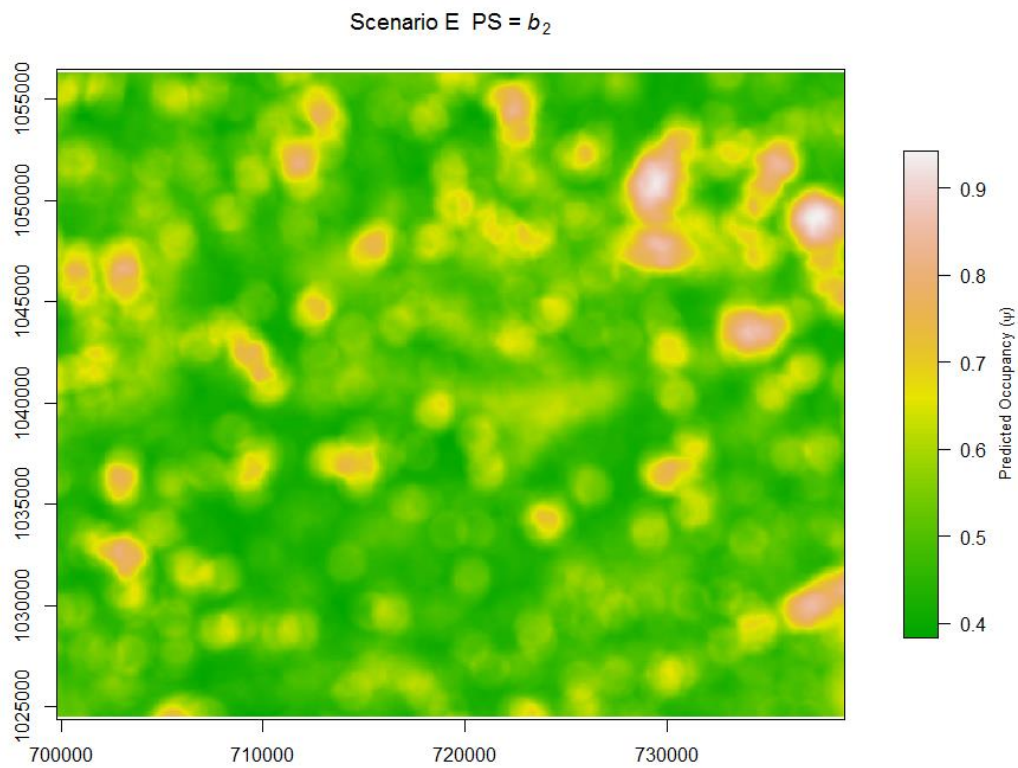
Figure 2.15. A map that visualizes the predicted probability of occupancy in a geographic space clipped from the 2011 NLCD in the EGCP using parameter set b2, under scenario E. Scenario E is modeled with a user accuracy of 70% for each habitat covariate.