Habitat Models and Predictions for Gopher Tortoises in Georgia

Final Report to USGS Cooperative Research Units Program
Project title: A Spatially-Explicit Population Modeling Framework to Support Conservation Decision Making for Gopher Tortoises in Georgia
Grant: G12AC20363
Research Work Order: GA RWO #113
Performance Period: July 1, 2012 - June 30, 2017

Bryan L. Nuse[1,2], Rachel L. Bormann[1], Jeffrey Hepinstall-Cymerman[1], and Clinton T. Moore[2,3]

[1]Warnell School of Forestry and Natural Resources, University of Georgia, Athens GA 30602-2152
[2]Georgia Cooperative Fish and Wildlife Research Unit, Warnell School of Forestry and Natural Resources, University of Georgia
[3]US Geological Survey

December 1, 2017

Contents

1 — Introduction

1.1. Project Summary

This report details the activities and results from the USGS Cooperative Research Units Program funded project entitled, *A Spatially-Explicit Population Modeling Framework to Support Conservation Decision Making for Gopher Tortoises in Georgia*, which ran from July 2012 to June 2017. This project was part of a larger set of projects whereby our team at the University of Georgia has attempted to map gopher tortoise (*Gopherus polyphemus*) habitat across its range in Georgia, develop population models using field survey data, and develop a decision support tool to assist Georgia Department of Natural Resources (GADNR) in prioritizing conservation planning efforts for the gopher tortoise.

The goal of this project was to develop models that predict response of gopher tortoise populations to conservation actions; such models could be used as predictive components of a formal decision tool to target conservation actions on the ground. Other projects conducted by our lab focused on mechanistic population models and decision analysis methods. Under this project, we focused on modeling gopher tortoise habitat suitability in the state of Georgia. While gopher tortoise habitat suitability models have been reported elsewhere, there existed no model built from a broad spatial base of monitoring data. Our work derives from surveys of gopher tortoises conducted by GADNR and its partners at sites across the state since 2008, and we use model predictors based on remotely-collected soil and habitat characteristics.

Our work is presented in three main sections. First, we develop habitat suitability models based only on soil and topographic predictors (i.e., ignoring vegetative components), with the purpose of comparing predictive ability of models that employ expert-derived soil suitability indices as predictors against those that use physical soil attributes as predictors. Second, we develop habitat suitability models that incorporate soil, topographic, and habitat predictors within four major ecoregions of Georgia. From the best of these models, we develop maps of suitable habitat across the tortoise range in Georgia. Third, we describe an open-source relational database that we built that integrates gopher tortoise occurrence records originally distributed across separate files.

1.2. Background

The gopher tortoise is a keystone species of southern pine forests, particularly of longleaf pine (*Pinus palustris*) savannas. It is also a candidate for Threatened status under the Endangered Species Act in the eastern portion of its range. Prevention of species listing is a shared goal of diverse public and private entities throughout the region, but this outcome may be achieved only by demonstrating that conservation measures are reducing population threats and are leading the species on a path to temporal and spatial persistence.

The U.S. Fish and Wildlife Service and other federal landholders, as well as state natural resource agencies in Florida, Georgia, Alabama, and South Carolina, are attempting to ascertain the status of gopher tortoise populations in each state. In Georgia, 79 sites on a combination of federal, state, and private lands already have been surveyed.

Survey efforts are ongoing; however, the focus is on estimating sizes and densities of known gopher tortoise populations, not on gopher tortoise habitat relationships, and so possess several limitations. First, the surveys have been conducted only on areas that are known to contain gopher tortoise burrows (i.e., are known suitable habitat), so we have limited data to test predictions of poor or unsuitable habitat. Second, limited vegetation data are collected during these surveys, making it difficult to draw inferences as to what specific habitat characteristics beyond soil type are driving gopher tortoise habitat suitability and population density. Many factors influence gopher tortoise populations, some of which are directly related to the production of quality habitat. Sites that are currently occupied clearly indicate that the site is currently suitable. The converse, i.e., sites are not currently occupied because the site is deemed unsuitable, is not necessarily true. Sites can be unoccupied for a variety of historical reasons such as land use, collection by humans, disease, or surrounding land uses leading to small populations susceptible to demographic or environmental stochasticity. Our ability to draw conclusions about habitat suitability and population status at a static point in time is limited by these facts.

We are attempting to gather sufficient data to produce maps of current habitat suitability as well as a better understanding of what management actions could move currently unsuitable habitat into a state able to support gopher tortoises. Population models, similarly, need to account for the movement (both natural and aided by people) of individuals.

Our broader gopher tortoise modeling effort is three-pronged: map suitable habitat; predict the metapopulation size for any given network of habitat patches; and develop a decision support tool to assist GADNR in prioritizing current and future gopher tortoise habitat to maintain future populations. Because our habitat model is the basis for our population model, it must adequately represent potential tortoise habitat. Gopher tortoises require a relatively open understory to facilitate movement, a somewhat open canopy to allow light penetration to the ground layer, and herbaceous ground cover for food. These conditions persist more readily on low quality sandy sites, but must be maintained through prescribed fire or mechanical methods in higher quality sites. As such, management actions can change poor quality tortoise habitat into higher quality habitat, provided the underlying soil type is appropriate. For example, a focus of related work done by our lab is a test of the hypothesis that financial incentives provided to landowners lead to restoration or maintenance of high quality habitat on private lands. We are attempting to build both habitat and population models in spatially explicit and temporally dynamic form, to account for the potential of land to transition in and out of suitable habitat for gopher tortoises with changes in overstory, midstory, and understory structure and composition. However, we are limited in our knowledge of exactly which structural and floristic qualities of vegetation drive habitat suitability.

As mentioned, Georgia DNR is currently surveying known gopher tortoise populations in the state across a range of land owner types and management regimes. These data are being shared with our team, and they form the empirical basis of both our habitat and population modeling. The DNR team uses a crew of three to survey sites using standard line distance transect sampling methodology (Nomani et al. (2008), Smith et al. (2009), Stober and Smith (2010)). Transects are only run on suitable soils within a study area (i.e., transects are not run through wetlands). All found burrows are examined with an endoscopic inspection camera to determine the presence of an animal, but descriptive notes regarding vegetation are only made for the entire study area.

Previous models predicting habitat suitability relied on NLCD land cover data to determine suitable overstory and ground conditions. However, the NLCD has limited land cover classes that do not adequately depict the specific vegetation conditions required by the tortoise. In earlier work (Hepinstall-Cymerman et al., 2017), our team used vegetation data collected in the field at a sample of gopher tortoise population sites to classify a time series of Landsat 8 imagery into six variables depicting overstory, midstory, understory, and grass-forb structure of a 30-m pixel. Models that incorporated the imagery-based vegetation metrics as predictors outperformed models that incorporated NLCD predictors, confirming that gopher tortoises use

habitat components not easily portrayed in coarse land cover types.

The U.S. Fish and Wildlife Service and USDA Natural Resources Conservation Service (2012) collaborated on the creation of a soil classification system for gopher tortoises. The system rated each of several soil properties for their importance to gopher tortoise life history requirements, so that a given soil type receives a score in the interval [0,1] and classified into discrete levels: Highly Suitable, Moderately Suitable, Less Suitable, Marginal, and Unsuitable. Since the publication of the first version in 2012, the classification system and the levels have been revised several times (NRCS, 2017*a*).

The soil classification system is acknowledged to be an imperfect predictor of site suitability for gopher tortoises, as the system does not account for other key components such as vegetation cover and soil temperature (NRCS, 2017*a*). However, because soils are considered a limiting factor for where tortoises occur on the landscape, the classification system is described as a kind of filter for conservation efforts, where efforts might be steered away from sites with a poor rating class regardless of the ability to bring about desired vegetation conditions on the site.

The soil classification system is in widespread use by conservation agencies in the region. However, the system is built on the composite knowledge of experts and has not been extensively field tested. The system has undergone repeated revision, suggesting response to counterfactual evidence to predictions generated by the system.

1.3. Objectives

The first objective of our work was to model gopher tortoise presence/pseudo-absence based on soil and topographic information only, for the purpose of comparing models using soil rating as a predictor to models using physical soil attributes as predictors. Under the premise that a model based only on soil characteristics serves as an adequate first filter for conservation efforts, we questioned whether the expert-derived soil rating performs that task better than direct use of some of the soil attributes that comprise the rating. The soil rating system is essentially the specification of a certain set of soil attributes and complex interactions among them that serve to indicate different degrees of suitability. Our interest was in comparing models using this system to models using soil attributes themselves in potential interactive and quadratic combinations.

The second objective of our work was development of habitat suitability models using soil, topographic, and imagery-based vegetation predictors. In this work, we considered models containing predictors based either on soil ratings or soil attributes, multiple two-way interactions among effects, quadratic effects, and predictors summarized at any of three levels of spatial scaling. We conducted model fitting within each of four Major Land Resource Areas (MLRA) that roughly correspond to ecoregion divisions of Georgia (NRCS, 2005). Averaging across well-performing models, we produced a habitat suitability map for the state of Georgia.

The third objective of our work was creation of a relational database of gopher tortoise burrow survey data by integrating data maintained in separate GIS shape files. We built the database on an open-source platform, requiring no proprietary software to run. The database is fully searchable and presents data summaries and export files for a diversity of needs. As tortoise monitoring now occurs throughout the range, the database architecture can be replicated elsewhere, ultimately facilitating the creation of a common data set across the species range.

2 — Comparing the utility of soil attributes against a soil suitability rating system, in modeling tortoise habitat

2.1. Introduction

In this section, we model gopher tortoise habitat using soil and topographic variables only, for the purpose of comparing models using soil ratings as a predictor against those using physical soil attributes as predictors. Soils are considered the limiting factor for gopher tortoises due to their need to dig deep burrows (USFWS, 1990). These soils include well-drained, sandy soils with a low clay content (Auffenberg and Franz, 1982; Cox et al., 1987). Topography is also considered a key component to modeling tortoise habitat (Jones and Dorr, 2004; Baskaran et al., 2006; USFWS and NRCS, 2012). Tortoises are often found in more xeric uplands/sandhills and ridges as opposed to mesic bottomlands or wetlands (McRae, 1981; Auffenberg and Franz, 1982; Diemer, 1986); some populations are known to defy this expectation, however (Castellón et al., 2012).

In 2012, the USFWS and NRCS developed a list of eight physical soil attributes (ponding frequency, flooding frequency, composition by medium and coarse gravel fragments, depth to annual high water table, depth to restrictive layer, % sand, % clay, and % slope) important to habitat suitability in the tortoise's federally threatened range (USFWS and NRCS, 2012). A recent revision of the index (version 6.6) incorporates potential evapotranspiration, in order to accommodate areas where tortoises use soils with higher water tables (NRCS, 2017*a*). This index is presented as a continuous value between 0 and 1, but breakpoints may optionally be applied to obtain discrete suitability classes as in previous versions.

Spatial scale (resolution) is an important component of any habitat analysis. The scale at which tortoises respond to soils and topography is unknown. To explore this relationship, we considered three scales for both soils suitability and topography within the size of a potential tortoise home range: 3 pixels by 3 pixels (0.81 ha; all raster maps used 30-m cells), 9 by 9 (7.29 ha), and 15 by 15 (20.25 ha).

We used logistic regression to explore differences between the values of soils and topographic variables at known gopher tortoise burrow locations, and at randomly placed pseudo-absences. Our objectives were to:

1. Compare models using the expert-derived USFWS/NRCS soil type classifications (soil suitability) with models using physical soil attributes,

2. Seek the most efficacious spatial scale, for each important soil and topographic variable, in predicting burrow occurrence.

2.2. Methods

2.2.1. Tortoise presence and pseudo-absence points

Burrow locations ($n = 20{,}231$) were obtained during line transect distance surveys (Nomani et al., 2008; Smith et al., 2009) performed from 2008 to 2017 by the Georgia Department of Natural Resources (DNR) and the J.W. Jones Ecological Research Center at Ichauway (Jones Center). All detected burrows, including those that were unoccupied or abandoned, were considered to be presences since they indicated appropriateness of the soil for digging.

Logistic regression requires absences, and in lieu of actual absence locations, we randomly generated pseudo-absence points equal in the number to that of the burrows. Pseudo-absences were placed within 5-km buffer zones surrounding each survey property boundary, but outside of a 300-m buffer surrounding each burrow. In this way we attempted to capture soil conditions in areas potentially accessible to but not used by tortoises, while avoiding choosing pseudo-absences in unknown tortoise populations (e.g., on unsurveyed private lands).

To account for potential spatial autocorrelation in environmental predictor variables, we generated correlograms of the predictors at burrows and used these to determine a minimum inter-point distance of 150 m between presence points. We applied this minimum distance to the dataset of burrows by randomly dropping one burrow from each pair with an inter-point distance less than the minimum (see section 5.3). After applying this spatial filter and eliminating any points whose spatial coordinates fell outside the study area, 4,712 presences found at 79 survey properties remained. The pseudo-absences were also subject to the minimum pairwise distance of 150 m.

2.2.2. Major Land Resource Areas

Because of differences between ecoregions with respect to elevation, topography, soil attributes, water and land cover, we performed separate analyses for each of the four NRCS-designated Major Land Resource Areas (MLRAs) within the tortoise's range in Georgia (Figure 2.1; NRCS 2005). The focal MLRAs are the Georgia Sand Hills (hereafter, *Sandhills*; 813 burrows), the Southern Coastal Plain (*Coastal Plain*; 2,565), the Atlantic Coast Flatwoods (*Flatwoods*; 1,256), and the Tidewater (78).

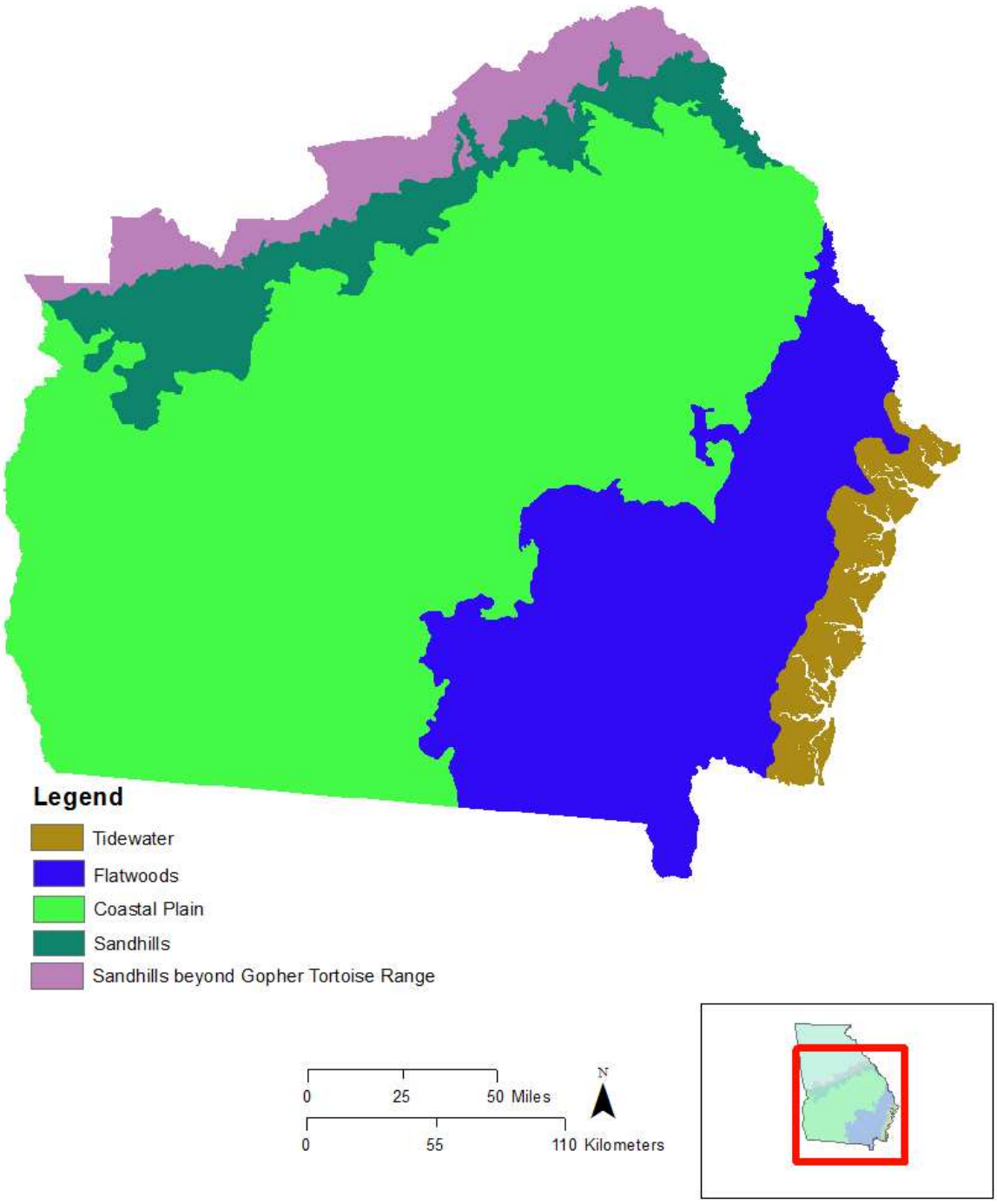

Figure 2.1.: Major Land Resource Areas (MLRAs) that intersect the gopher tortoise's range in Georgia. The abbreviated names used in this report are shown (see section 2.2.2 for full names used by the NRCS).

2.2.3. Environmental Variables

We obtained soil attribute data from the Soil Survey Geographic Database (SSURGO), mapped via the rasterized version, gSSURGO (NRCS, 2013). We resampled the 10-m resolution data to 30 m in order to match the lowest resolution data used in our analysis.

For the soils suitability models, we used the 2017 revision (version 6.6; NRCS 2017*a*) of a gopher tortoise soil suitability index developed by USFWS and NRCS (2012). We used the

index value directly, which is presented on the interval $[0, 1]$. We created three scales of soils suitability habitat by averaging the index value within a moving window in ArcGIS 10.1: 3 pixels by 3 (0.81 ha), 9 by 9 (7.29 ha), and 15 by 15 (20.25 ha) (symbols, when used as model variables: $\texttt{soil}_3$, $\texttt{soil}_9$ and $\texttt{soil}_{15}$, respectively).

From those available in the SSURGO dataset, we selected soil attributes that were considered to be important to gopher tortoise life history requirements (USFWS and NRCS, 2012), and were sufficiently complete (lacked missingness). Of the NRCS attributes listed, we used percent sand (symbol: `sand`) and annual minimum depth to water table (`wtd`). Percent sand is a continuous variable on [0, 100]. Depth to water table was divided into three categorical variables according to the breakpoints in the USFWS/NRCS report: 0-20 cm (unsuitable; $\texttt{wtd}_{0-20}$), 21-80 cm (marginal to moderately suitable; $\texttt{wtd}_{21-80}$), and > 80 cm (highly suitable; $\texttt{wtd}_{>80}$) (USFWS and NRCS, 2012). After examining the cross-tabulations of categories and presence/pseudo-absence, we reduced the levels of `wtd` to two ($\texttt{wtd}_{\leq 80}$ and $\texttt{wtd}_{>80}$) in the Sandhills and Tidewater areas due to few presences in the 0-20 cm and 21-80 cm categories. We interpreted all missing values as indicating a depth greater than 200 cm (the maximum depth recorded).

We did not use % clay or % silt due to high Pearson's correlations ($|r| > 0.7$) with % sand. We did not use the categorical variables summarizing flood frequency or ponding frequency, because of a lack of variation, and missingness. We did not make use of the continuous measures of gravel fragment size or depth to restrictive layer due to a large number of missing values in the tortoise's range in Georgia.

In order to obtain a more complete picture of the effect of topography on tortoise habitat, we used a topographic position index (TPI), calculated using the ArcGIS tools provided by Jenness (2006) rather than the percent slope value indicated by SSURGO. The TPI is a continuous variable calculated from a 30-m digital elevation model (DEM) for the state of Georgia (resampled from 10-m DEM obtained from the USGS National Map: https://nationalmap.gov/), and takes into account the difference between a pixel's elevation and the average elevation of the neighborhood around that pixel, in addition to its slope. Positive TPI values mean the cell is higher than its surroundings (tends towards ridgetops), while negative values mean it is lower (tends towards valleys). Values near zero could mean either a flat area or a mid-slope area, depending on the scale used. Topographic position index is scale-dependent, and is determined to a degree by the neighborhood used in the analysis. For example, what may be a ridge at a small scale could be part of a valley at a larger scale. We smoothed the TPI map similarly to `soil`, at the same three scales: 3x3, 9x9, and 15x15 pixels ($\texttt{tpi}_3$, $\texttt{tpi}_9$ and $\texttt{tpi}_{15}$, respectively).

Correlations (Pearson's $|r|$) between the continuous variables we used in logistic regression models were all less than 0.70. Before model-fitting, each continuous variable was first centered around a central value: its mean, or the midpoint of its admissible range (50% for `sand`, 0.5 for `soil`). The variable was then divided by a scaling factor: its standard deviation, in the case of `tpi`; 100 for `sand`; 1 for `soil`.

2.2.4. Models & Model Evaluation

We related the presence and pseudo-absence of gopher tortoise burrows to our environmental variables (Table 2.1) using multiple logistic regression. We developed a candidate set of models that when ranked via Akaike's Information Criterion (AIC), would allow us to assess variable importance, prevalence of non-linear effects, and the degree of model complexity required to model burrow occurrence. The same set of models was fit in each MLRA. *Soil attribute* models contained only terms in: TPI at each of the three scales ($\texttt{tpi}_3$, $\texttt{tpi}_9$, $\texttt{tpi}_{15}$), `sand`, and `wtd`. *Soil suitability* models contained only terms in: soils suitability and TPI, each at three scales ($\texttt{soil}_3$, $\texttt{soil}_9$, $\texttt{soil}_{15}$; $\texttt{tpi}_3$, $\texttt{tpi}_9$, $\texttt{tpi}_{15}$). The *soil attribute* model set contained all possible univariate models, models with all possible quadratic terms for the continuous variables (with the constituent main effect), and models with all possible two-way interactions among main effects. The *soil suitability* model set contained all possible univariate models, as well as models

with TPI in quadratic form ($\texttt{tpi}_3^2$, $\texttt{tpi}_9^2$, $\texttt{tpi}_{15}^2$). In this model set, we did not consider models in which soil suitability and TPI were both included but at different scales. Models were fit using the `glm()` function in program R (R Core Team, 2017).

We evaluated model predictive value with 5-fold cross validation (a random selection of 80% of the data for a training set and 20% for a test set), with 100 repetitions. We report the median value of the area under receiver operating characteristic curve (AUC), calculated with the R package ROCR (Sing et al., 2005).

2.3. Results and Discussion

In general, more complex models were favored according to AIC in all MLRAs except the Tidewater area, larger scales were favored over smaller scales, and soils attribute models were favored over soils suitability models (Tables 2.2, 2.4, 2.6, 2.8). The higher rankings of the 15x15 scale models indicate that burrow locations are modeled better by larger topographic features rather than minor undulations in the landscape across the tortoise's range in Georgia.

In three (Sandhills, Coastal Plain, Flatwoods) of the four MLRAs, the identical model was indicated as the top model, garnering an overwhelming degree of AIC weight ($\omega \geq 0.969$). This model was the most complex fitted, containing main effects in `wtd`, `sand`, and $\texttt{tpi}_{15}$; quadratic effects in `sand` and $\texttt{tpi}_{15}$; and all two-way interactions involving `sand`. The estimated coefficients varied in their state of statistical significance (at $\alpha = 0.05$). A second model in the same form but involving $\texttt{tpi}_9$ was either the second or third-ranked model in these sets.

In the Sandhills MLRA, no other model scored within 10 AIC points of the top-ranked one (Table 2.2). This model included effects of $\texttt{sand:tpi}_{15}$ and $\texttt{sand}^2$ but neither was significant at the $\alpha = 0.05$ level (Table 2.3). The large positive coefficient for `sand` indicates that the probability of a burrow occurrence increases sharply linearly as % sand increases. However, the large negative coefficient for the `sand:wtd` interaction indicates that the rate of increase in probability with sand content is higher when the water table depth is > 80 cm than when it is ≤ 80 cm (i.e., regardless of sand content, tortoises are not burrowing in an area that has a minimum annual water table depth ≤ 80 cm nearly as often as when that depth is > 80 cm; Figure 2.2). The significant negative (concave down) coefficient for $\texttt{tpi}_{15}^2$ indicates burrow probabilities are higher outside of broad valleys (low TPI values) but the rate of increase levels off as the topography changes to mid-slope and then to ridge areas (Figure 2.2; right panel).

In the Coastal Plain MLRA, some support ($\Delta\text{AIC} \leq 10$) existed for the same complex model at the 9x9 scale (Table 2.4). The model included effects of the $\texttt{sand:tpi}_{15}$ and `sand:wtd` interactions—though neither was significant at $\alpha = 0.05$ level—and the quadratic effect $\texttt{tpi}_{15}^2$ (Table 2.5). The significant negative value of the $\texttt{tpi}_{15}^2$ coefficient helps describe a peak of occurrence probability at moderately positive values of `tpi` (mid-slopes and broad ridges; Figure 2.3). The large positive coefficient for the quadratic sand variable $\texttt{sand}^2$ indicates that the probability of a burrow occurrence increases sharply in a quadratic fashion (concave up) as `sand` increases. The apparent negative relationship between sand composition and burrow presence at low sand composition amounts (Figure 2.3) is not real: it is an artifact of few samples at low sand percentage and the poor ability of the quadratic model to represent a strictly increasing pattern of curvature. The probability of burrow occurrence decreases as `wtd` decreases, with a large negative coefficient for the lowest `wtd` category compared to the deeper categories (Figure 2.3).

In the Flatwoods MLRA, none but the most complex model was supported ($\Delta\text{AIC} \leq 10$) (Table 2.6). Here, all the interactions were significant, but $\texttt{tpi}_{15}^2$ was not (Table 2.7). Similar to the Coastal Plain, the large positive coefficient for the $\texttt{sand}^2$ variable indicates that the probability of a burrow occurrence increases sharply in a quadratic fashion (concave up) as % sand increases. The small positive coefficient for $\texttt{tpi}_{15}$ indicates burrow probability increases in a positive direction from low TPI values (valleys) to mid-slopes, to ridges (highest TPI values).

For the `sand:wtd` interaction, the coefficients indicate that the relationship between sand content and burrow occurrence is positive in deeper sands ($\texttt{wtd}_{>80}$) and negative in shallower sand classes ($\texttt{wtd}_{\leq 80}$). In this MLRA, the $\texttt{sand:tpi}_{15}$ interaction is significant and negative, indicating that at the highest TPI values (ridgetops), the positive effect of an increase in % sand is reduced by the benefit of having a high, dry landscape (Figure 2.4).

For the Tidewater MLRA, two models of simple form (main effects only of `wtd`, `sand`, and `tpi`) collectively accounted for 0.839 of model weight, with ambiguity in the choice of scale for `tpi` (Table 2.6). The top-ranked model favored `tpi` at the 9x9 scale ($\omega = 0.520$), and the second-ranked model favored `tpi` at the 15x15 scale ($\omega = 0.319$). The top four models collectively accounted for 0.965 of model weight. The third and fourth-ranked models also included the main effects `wtd`, `sand`, and `tpi` but included both two-way interactions involving `sand`. Again, these two models were indeterminate in choice of scale, either 9x9 or 15x15, for `tpi`. Although all three effects appear in the top model, only `wtd` and `tpi` were significant ($p \leq 0.05$; Table 2.9). This suggests that tortoises do not have a wide range of appropriate sandy soils and are burrowing in some suboptimal conditions wherever the topography is supportive and soils are not wet, regardless of sand composition (Figure 2.5). The prominence of `tpi` at the 9x9 scale, relative to other MLRAs, suggests that tortoises are choosing habitat at smaller scales in this MLRA, which contains large expanses of non-habitat and many connectivity barriers (rivers and other large bodies of water and US Interstate 95).

According to the cross-validation results, all top models have high predictive value (median AUC > 70%; Tables 2.2 - 2.8). The Tidewater area has unexpectedly high AUC values (median AUC > 95%), perhaps attributable to our method of choosing pseudo-absences. The 5-km buffer may have been too large for this MLRA and the pseudo-absences may have been chosen from areas outside the environmental envelope where tortoises would be found.

Overall, the models using soil attributes available in the SSURGO database confirm the general soil habitat characteristics considered to be important by the literature for gopher tortoise habitat. The soils suitability classifications undoubtedly have high utility for mapping general tortoise habitat and choosing survey areas in Georgia, but our results suggest that incorporating physical soil attributes where the detail is available may be beneficial in other modeling efforts. Finally, we observed that larger scales for topographic attributes and soil suitability classifications were most useful for prediction, in areas outside of the Tidewater region.

2.3.1. Predictive maps

To visualize our model predictions on the study landscape, we produced a map for each MLRA that represents the average predicted burrow occurrence probability, averaged across the top models (those with ΔAIC < 2). In ArcGIS Raster Calculator, we first made a predictive map for each of the top models for each MLRA, by making a linear combination of the appropriate predictor raster layers with the estimated coefficients from that model. Topographic position index layers were resampled to 30 m to match all other layers.

We then took a weighted average of the (logit-scale) values of the linear predictors from the set of top models, with the weights being the models' corresponding AIC weights (Tables 2.2–2.8). The average linear predictor values were transformed into probabilities, and the composite prediction layers for each MLRA were mosaicked into a single map (Figures 2.6, 3.8–3.9). To facilitate additional analysis, this continuous layer ($\in [0, 1]$) was classified into *suitable* (≥ 0.50) and *unsuitable* (< 0.50) (Figure 2.7).

Table 2.1.: Characteristics of variables in the gopher tortoise burrow dataset, including pseudo-absences used in this chapter, in each of the four MLRAs.
Notes:
1) For `burrow`, the binary response variable, the variable name column gives the categories: 1 for burrows, 0 for pseudo-absences.
2) For `wtd`, the categories are specified separately for each MLRA, since they used differing classifications.
3) The `sites` row gives the number of tortoise survey sites in that MLRA.
4) Comparing with Table 3.2 reveals two differences in the dataset used in Chapter 3, versus the one described here. In the Chapter 3 analysis, 1) twice as many pseudo-absences were used in the Tidewater MLRA; and 2) two categories of water table depth (`wtd` $\in \{\leq 80, > 80\}$) were used for all MLRAs.
5) Sum across the `sites` row is 90, which is greater than the number of sites (79); this is because some sites (and/or associated pseudo-absence annuli) extend across an MLRA boundary.

	Coastal Plain		Flatwoods		Sandhills		Tidewater	
Variable	Mean	Range/count	Mean	Range/count	Mean	Range/count	Mean	Range/count
$\texttt{burrow}_0$		2565		1256		813		78
$\texttt{burrow}_1$		2565		1256		813		78
`sites`		51		23		13		3
`wtd`		*0-20:* 645		*0-20:* 893		$\leq$*80:* 92		$\leq$*80:* 95
		21-80: 380		*21-80:* 737				
		$>$*80:* 4105		$>$*80:* 882		$>$*80:* 1534		$>$*80:* 61
`sand`	77.0	[5.3, 97.9]	83.7	[0, 98.4]	83.7	[5.3, 95.0]	75.2	[5.5, 98.4]
$\texttt{tpi}_3$	0.1	[-4, 5]	0.1	[-3, 2]	0.1	[-3, 4]	0.0	[-2, 1]
$\texttt{tpi}_9$	0.3	[-14, 13]	0.2	[-4, 6]	0.4	[-10, 12]	0.1	[-2, 2]
$\texttt{tpi}_{15}$	0.5	[-18, 18]	0.4	[-4, 7]	0.7	[-17, 16]	0.2	[-2, 4]
$\texttt{soils}_3$	0.67	[0, 1]	0.49	[0, 1]	0.75	[0, 1]	0.40	[0, 1]
$\texttt{soils}_9$	0.63	[0, 1]	0.44	[0, 1]	0.73	[0, 1]	0.37	[0, 1]
$\texttt{soils}_{15}$	0.61	[0, 1]	0.40	[0, 1]	0.71	[0, 1]	0.35	[0, 1]

Table 2.2.: Complete list of models of burrow occurrence, in the **Sandhills** MLRA, ranked by AIC.
Column definitions:
1) Column ω shows the AIC weights, calculated as follows: if, for the ith model of a set numbering M, $q_i = e^{(-0.5*\Delta AIC_i)}$, then $\omega_i = q_i/(\sum_{i=1}^{M} q_i)$, with $\Delta AIC_i = AIC_i - \min(AIC)$. Column K gives the number of model parameters (including the intercept).
2) Column AUC shows the median value of the area under the receiver operator characteristic curve, over 100 repetitions of a 5-fold cross validation.
3) Columns representing variables are filled when that variable appears in the model, with the following codes: the italicized number shows the scale of the variable (3,9,15); a "1" indicates that only the main effect of the variable was included; a "2" indicates that a quadratic term for the variable was included; and capital letters show that the variable participated in a two-way interaction with a second variable matched to the letter shown (letters with which variables are matched appear above their names in column labels). For example the value "*3:* 2,A,D" shows that that column's variable appeared in the model smoothed at the 3x3 scale, that the model included the quadratic term of the variable, and that it was included in two-way interactions with both `wtd` and `soil`.

			A	B	C	D	
AIC	ω	K	`wtd`	`sand`	`tpi`	`soil`	AUC
1832.6	0.995	8	B	2,A,C	*15:* 2,B		0.75
1843.8	0.004	4			*15:* 2	*15*	0.75
1847.6	0.001	8	B	2,A,C	*9:* 2,B		0.76
1848.5	0.000	4			*9:* 2	*9*	0.74
1850.9	0.000	5		2	*15:* 2		0.75
1862.8	0.000	4	1	1	*15*		0.75
1863.5	0.000	6	B	A,C	*15:* B		0.75
1872.2	0.000	6	B	A,C	*9:* B		0.75
1872.2	0.000	4	1	1	*9*		0.74
1872.3	0.000	5		2	*9:* 2		0.74
1880.4	0.000	8	B	2,A,C	*3:* 2,B		0.75
1883.9	0.000	4			*3:* 2	*3*	0.72
1887.2	0.000	3		1	*15*		0.75
1889.1	0.000	4		C	*15:* B		0.74
1890.0	0.000	6	B	A,C	*3:* B		0.74
1890.0	0.000	4	1	1	*3*		0.74
1894.2	0.000	2				*3*	0.72
1896.2	0.000	4	B	A			0.73
1897.0	0.000	2				*9*	0.73
1897.4	0.000	3	1	1			0.74
1901.0	0.000	3		1	*9*		0.74
1902.6	0.000	4		C	*9:* B		0.74
1915.4	0.000	5		2	*3:* 2		0.73

Continued next page...

Table 2.2.: **Sandhills** continued.

			A	B	C	D	
AIC	ω	K	`wtd`	`sand`	`tpi`	`soil`	AUC
1927.1	0.000	3		1	*3*		0.73
1928.7	0.000	4		C	*3:* B		0.73
1929.5	0.000	2				*15*	0.73
1938.3	0.000	2		1			0.73
1940.3	0.000	3		2			0.73
2099.3	0.000	3	1		*15*		0.63
2099.7	0.000	3			*15:* 2		0.61
2117.4	0.000	3	1		*9*		0.62
2127.2	0.000	3			*9:* 2		0.62
2139.7	0.000	3	1		*3*		0.59
2144.8	0.000	2	1				0.56
2169.3	0.000	2			*15*		0.62
2191.9	0.000	3			*3:* 2		0.59
2199.7	0.000	2			*9*		0.61
2242.3	0.000	2			*3*		0.55
2256.1	0.000	1					0.50

Table 2.3.: Parameter estimates for the top-ranked model in the **Sandhills** MLRA (see Table 2.2). A colon ':' indicates an interaction.
Column definitions:
1) For each term, the point estimate and standard error are given on the logit scale.
2) For the intercept, corresponding probability estimate is provided, with its associated 95% confidence interval. Note that the intercept represents the reference level of `wtd`, i.e. $\texttt{wtd}_{>80}$; the value shown thus corresponds to the case where `wtd` > 80 cm and all other variables are at their means. For all other terms, the estimate is exponentiated to give the odds ratio, with confidence interval. For these terms, statistical significance at the $\alpha = 0.05$ level is inferred by exclusion of 1.0 from the confidence interval.
3) Continuous variables were centered and/or scaled before model fitting: `tpi` was centered by subtracting the mean (Table 2.1), then divided by the standard deviation; `sand` was centered by subtracting 50%, then divided by 100. These scaling factors represent the units of change associated with the given odds ratio. For example, with `sand` at its mean and `wtd` at its reference level ($\texttt{wtd}_{>80}$), an increase of 3.37 in the $\texttt{tpi}_{15}$ variable would increase the expected odds of finding a burrow 1.50 * 0.76 = 1.14 times (i.e., product of odds ratios for $\texttt{tpi}_{15}$ and $\texttt{tpi}_{15}^2$: the interactions may be ignored when other variables are at their means/reference levels).

Term	Estimate	SE	Odds ratio	Probability	95% CI	Unit change
`Intercept`	-2.88	0.41		0.05	[0.02, 0.11]	
`sand`	10.41	2.75	3.3×10^4		[184, 8.9×10^6]	100%
$\texttt{wtd}_{\leq 80}$	-0.24	1.09	0.79		[0.05, 5.27]	
$\texttt{tpi}_{15}$	0.41	0.31	1.50		[0.82, 2.75]	3.37
$\texttt{sand}^2$	-4.44	4.58	0.01		[0.00, 78.0]	
$\texttt{tpi}_{15}^2$	-0.27	0.05	0.76		[0.69, 0.84]	
$\texttt{sand:wtd}_{\leq 80}$	-8.03	3.97	0.00		[0.00, 1.92]	
$\texttt{sand:tpi}_{15}$	0.42	0.84	1.52		[0.30, 8.04]	

Table 2.4.: Complete list of models of burrow occurrence, in the **Coastal Plain** MLRA, ranked by AIC. Columns defined as in Table 2.2.

			A	B	C	D	
AIC	ω	K	`wtd`	`sand`	`tpi`	`soil`	AUC
5761.7	0.969	10	B	2,A,C	*15:* 2,B		0.78
5768.6	0.031	10	B	2,A,C	*9:* 2,B		0.79
5816.6	0.000	10	B	2,A,C	*3:* 2,B		0.78
5877.8	0.000	8	B	A,C	*15:* B		0.78
5887.0	0.000	8	B	A,C	*9:* B		0.78
5895.7	0.000	5	1	1	*15*		0.78
5903.9	0.000	5	1	1	*9*		0.78
5933.5	0.000	8	B	A,C	*3:* B		0.78
5951.3	0.000	5	1	1	*3*		0.78
5982.5	0.000	6	B	A			0.78
6002.0	0.000	4	1	1			0.78
6046.8	0.000	4			*9:* 2	*9*	0.73
6047.2	0.000	5		2	*15:* 2		0.76
6073.1	0.000	4			*3:* 2	*3*	0.73
6077.3	0.000	5		2	*9:* 2		0.75
6128.2	0.000	4			*15:* 2	*15*	0.73
6169.8	0.000	2				*3*	0.71
6183.0	0.000	3		1	*15*		0.75
6184.6	0.000	4		C	*15:* B		0.74
6198.4	0.000	5		2	*3:* 2		0.74
6215.4	0.000	3		1	*9*		0.75
6217.0	0.000	4		C	*9:* B		0.74
6282.8	0.000	2				*9*	0.71
6314.3	0.000	3		2			0.72
6323.9	0.000	4	1		*15*		0.69
6331.3	0.000	3		1	*3*		0.74
6332.1	0.000	4	1		*9*		0.68
6332.1	0.000	4		C	*3:* B		0.74
6388.8	0.000	4	1		*3*		0.67
6430.4	0.000	2		1			0.73
6434.2	0.000	2				*15*	0.69
6448.5	0.000	3	1				0.64

Continued next page...

Table 2.4.: **Coastal Plain** continued.

			A	B	C	D	
AIC	ω	K	`wtd`	`sand`	`tpi`	`soil`	AUC
6729.5	0.000	3			*15:* 2		0.65
6777.1	0.000	3			*9:* 2		0.64
6779.8	0.000	2			*15*		0.65
6824.7	0.000	2			*9*		0.64
6958.3	0.000	3			*3:* 2		0.59
6984.9	0.000	2			*3*		0.60
7113.7	0.000	1					0.50

Table 2.5.: Parameter estimates for the top-ranked model in the **Coastal Plain** MLRA (see Table 2.4). See Table 2.3 for column definitions, and Table 2.1 for variable means and ranges.

Term	Estimate	SE	Odds ratio	Probability	95% CI	Unit change
`Intercept`	-0.78	0.08		0.31	[0.28, 0.35]	
`sand`	-0.07	0.49	0.93		[0.36, 2.42]	100%
$\texttt{wtd}_{0-20}$	-2.02	0.21	0.13		[0.09, 0.20]	
$\texttt{wtd}_{21-80}$	-0.66	0.26	0.52		[0.31, 0.84]	
$\texttt{tpi}_{15}$	0.36	0.08	1.44		[1.24, 1.67]	2.15
$\texttt{sand}^2$	10.40	1.03	3.2×10^4		$[4.3 \times 10^3, 2.4 \times 10^5]$	
$\texttt{tpi}_{15}^2$	-0.07	0.01	0.94		[0.91, 0.96]	
$\texttt{sand:wtd}_{0-20}$	0.37	0.71	1.45		[0.36, 5.90]	
$\texttt{sand:wtd}_{21-80}$	-1.50	0.82	0.22		[0.05, 1.15]	
$\texttt{sand:tpi}_{15}$	0.24	0.24	1.27		[0.80, 2.03]	

Table 2.6.: Complete list of models of burrow occurrence, in the **Flatwoods** MLRA, ranked by AIC. Columns defined as in Table 2.2.

			A	B	C	D	
AIC	ω	K	`wtd`	`sand`	`tpi`	`soil`	AUC
1809.8	1.000	10	B	2,A,C	*15:* 2,B		0.92
1831.2	0.000	8	B	A,C	*15:* B		0.92
1860.4	0.000	10	B	2,A,C	*9:* 2,B		0.92
1864.8	0.000	5	1	1	*15*		0.92
1885.7	0.000	8	B	A,C	*9:* B		0.92
1913.2	0.000	5	1	1	*9*		0.91
1921.2	0.000	10	B	2,A,C	*3:* 2,B		0.91
1943.9	0.000	8	B	A,C	*3:* B		0.90
1959.5	0.000	5	1	1	*3*		0.90
1974.9	0.000	4			*15:* 2	*15*	0.91
2008.9	0.000	6	B	A			0.90
2011.5	0.000	4			*9:* 2	*9*	0.90
2015.4	0.000	4	1	1			0.90
2066.7	0.000	4	1		*15*		0.90
2119.6	0.000	4	1		*9*		0.89
2121.2	0.000	4			*3:* 2	*3*	0.88
2168.1	0.000	4	1		*3*		0.88
2185.5	0.000	2				*3*	0.86
2193.5	0.000	2				*9*	0.88
2224.2	0.000	3	1				0.86
2266.4	0.000	2				*15*	0.87
2329.0	0.000	5		2	*15:* 2		0.87
2336.5	0.000	4		C	*15:* B		0.87
2355.0	0.000	3		1	*15*		0.87
2439.4	0.000	5		2	*9:* 2		0.85
2489.5	0.000	4		C	*9:* B		0.85
2505.7	0.000	3		1	*9*		0.85
2601.4	0.000	5		2	*3:* 2		0.82
2654.8	0.000	4		C	*3:* B		0.81
2667.2	0.000	3		1	*3*		0.81
2750.4	0.000	3		2			0.77
2788.1	0.000	2		1			0.77

Continued next page...

Table 2.6.: **Flatwoods** continued.

			A	B	C	D	
AIC	ω	K	`wtd`	`sand`	`tpi`	`soil`	AUC
2877.7	0.000	2			*15*		0.80
2879.6	0.000	3			*15:* 2		0.80
3050.2	0.000	3			*9:* 2		0.75
3091.9	0.000	2			*9*		0.75
3281.7	0.000	3			*3:* 2		0.65
3321.4	0.000	2			*3*		0.65
3484.4	0.000	1					0.50

Table 2.7.: Parameter estimates for the top-ranked model in the **Flatwoods** MLRA (see Table 2.6). See Table 2.3 for column definitions, and Table 2.1 for variable means and ranges.

Term	Estimate	SE	Odds ratio	Probability	95% CI	Unit change
`Intercept`	-3.58	0.78		0.03	[0.01, 0.10]	
`sand`	5.69	2.43	295.6		[3.15, 4.8×10^4]	100%
$\texttt{wtd}_{0-20}$	-2.93	1.01	0.05		[0.01, 0.41]	
$\texttt{wtd}_{21-80}$	2.14	0.82	8.53		[1.85, 47.55]	
$\texttt{tpi}_{15}$	2.13	0.29	8.43		[4.91, 15.71]	0.95
$\texttt{sand}^2$	16.3	2.91	1.2×10^7		[3.3×10^4 , 3.5×10^9]	
$\texttt{tpi}_{15}^2$	-0.05	0.03	0.95		[0.90, 1.02]	
$\texttt{sand:wtd}_{0-20}$	-1.00	2.44	0.37		[0.00, 40.7]	
$\texttt{sand:wtd}_{21-80}$	-8.49	2.01	0.00		[0.00, 0.01]	
$\texttt{sand:tpi}_{15}$	-3.30	0.69	0.04		[0.01, 0.13]	

Table 2.8.: Complete list of models of burrow occurrence, in the **Tidewater** MLRA, ranked by AIC. Columns defined as in Table 2.2.

			A	B	C	D	
AIC	ω	K	`wtd`	`sand`	`tpi`	`soil`	AUC
75.6	0.520	4	1	1	*9*		0.97
76.6	0.319	4	1	1	*15*		0.97
79.6	0.073	6	B	A,C	*9:* B		0.97
80.2	0.053	6	B	A,C	*15:* B		0.97
83.0	0.013	8	B	2,A,C	*9:* 2,B		0.96
83.5	0.010	8	B	2,A,C	*15:* 2,B		0.95
84.4	0.006	4	1	1	*3*		0.95
86.4	0.002	3	1	1			0.96
88.4	0.001	6	B	A,C	*3:* B		0.95
88.4	0.001	4	B	A			0.95
92.0	0.000	8	B	2,A,C	*3:* 2,B		0.93
94.3	0.000	3	1		*15*		0.96
97.0	0.000	5		2	*15:* 2		0.94
98.3	0.000	5		2	*9:* 2		0.94
101.6	0.000	3	1		*9*		0.94
115.8	0.000	3	1		*3*		0.88
116.7	0.000	2	1				0.87
118.2	0.000	5		2	*3:* 2		0.88
120.0	0.000	3		2			0.87
130.9	0.000	3		1	*15*		0.87
132.8	0.000	4		C	*15:* B		0.88
133.3	0.000	3		1	*9*		0.86
134.4	0.000	4		C	*9:* B		0.86
150.5	0.000	3		1	*3*		0.75
151.8	0.000	4		C	*3:* B		0.76
155.3	0.000	2		1			0.67
172.8	0.000	2			*15*		0.80
174.8	0.000	3			*15:* 2		0.80
175.7	0.000	4			*15:* 2	*15*	0.79
191.2	0.000	4			*9:* 2	*9*	0.73
191.5	0.000	2			*9*		0.72
193.1	0.000	3			*9:* 2		0.74

Continued next page...

Table 2.8.: **Tidewater** continued.

			A	B	C	D	
AIC	ω	K	`wtd`	`sand`	`tpi`	`soil`	AUC
209.4	0.000	4			*3:* 2	*3*	0.64
211.0	0.000	2				*3*	0.62
212.4	0.000	2				*9*	0.66
213.3	0.000	2				*15*	0.64
216.1	0.000	2			*3*		0.60
217.4	0.000	3			*3:* 2		0.59
218.3	0.000	1					0.50

Table 2.9.: Parameter estimates for the top-ranked model in the **Tidewater** MLRA (see Table 2.8). See Table 2.3 for column definitions, and Table 2.1 for variable means and ranges.

Term	Estimate	SE	Odds ratio	Probability	95% CI	Unit change
`Intercept`	-2.28	4.46		0.09	[0.00, 0.89]	
`sand`	13.64	10.59	8.4×10^5		[0.001, 8.7×10^{14}]	100%
$\texttt{wtd}_{\leq 80}$	-4.69	1.02	0.01		[0.00, 0.04]	
$\texttt{tpi}_9$	2.55	1.11	12.8		[2.29, 163]	0.38

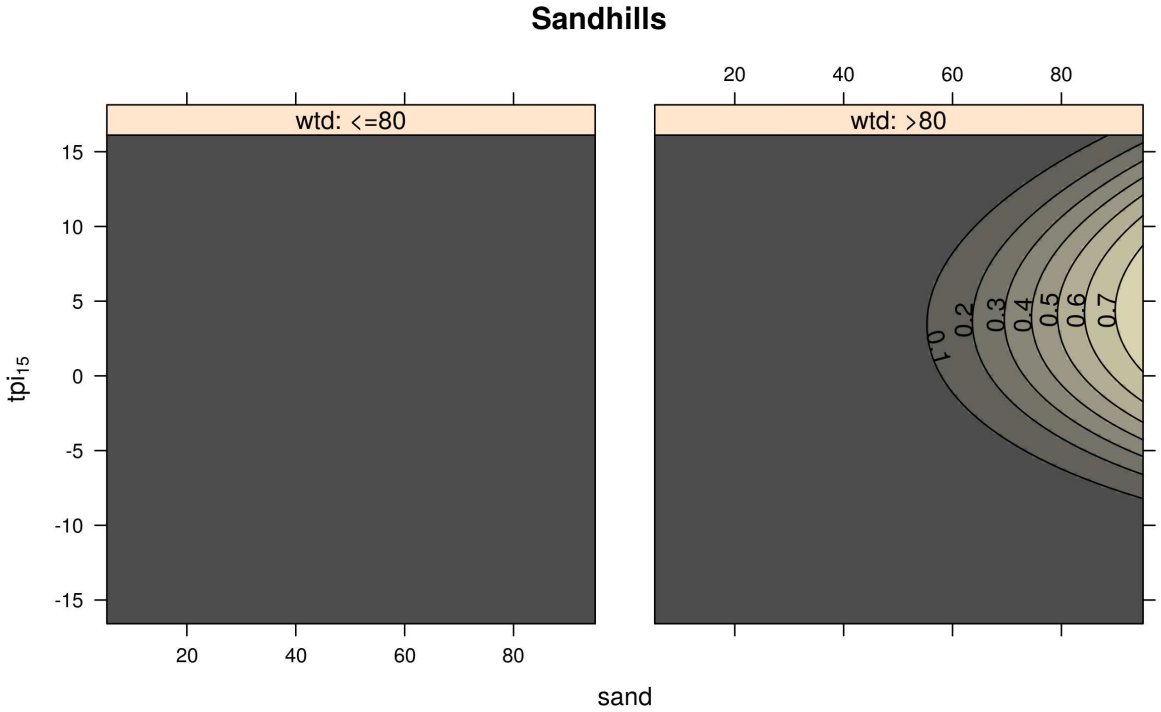

Figure 2.2.: Contour plots of burrow occurrence probability, for the **Sandhills** top model in Table 2.2. Each plot shows the expected occurrence probability, over the ranges of `sand` and $\texttt{tpi}_{15}$, at one level of the `wtd` variable, annual minimum depth to water table (cm). Probability values throughout the left panel ("`wtd`: $\leq$80") are all < 0.1.

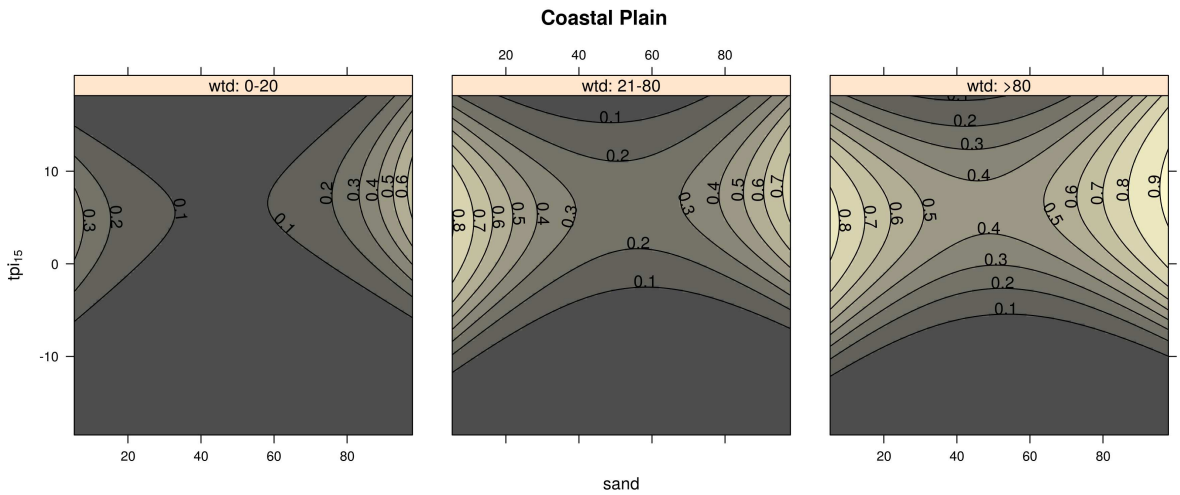

Figure 2.3.: Contour plots of burrow occurrence probability, for the **Coastal Plain** top model in Table 2.4. Each plot shows the expected occurrence probability, over the ranges of `sand` and $\texttt{tpi}_{15}$, at one level of the `wtd` variable, annual minimum depth to water table (cm).

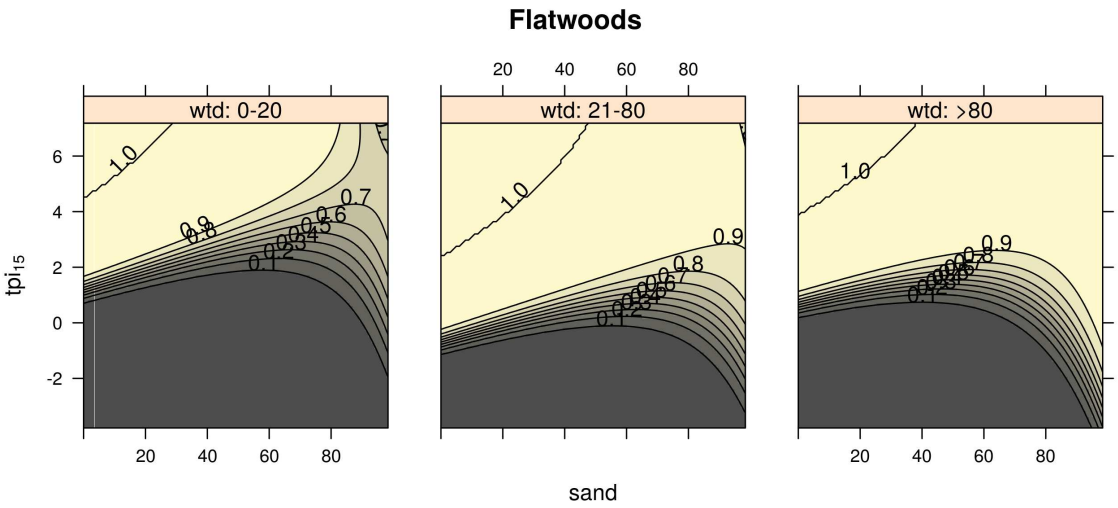

Figure 2.4.: Contour plots of burrow occurrence probability, for the **Flatwoods** top model in Table 2.6. Each plot shows the expected occurrence probability, over the ranges of `sand` and $\texttt{tpi}_{15}$, at one level of the `wtd` variable, annual minimum depth to water table (cm).

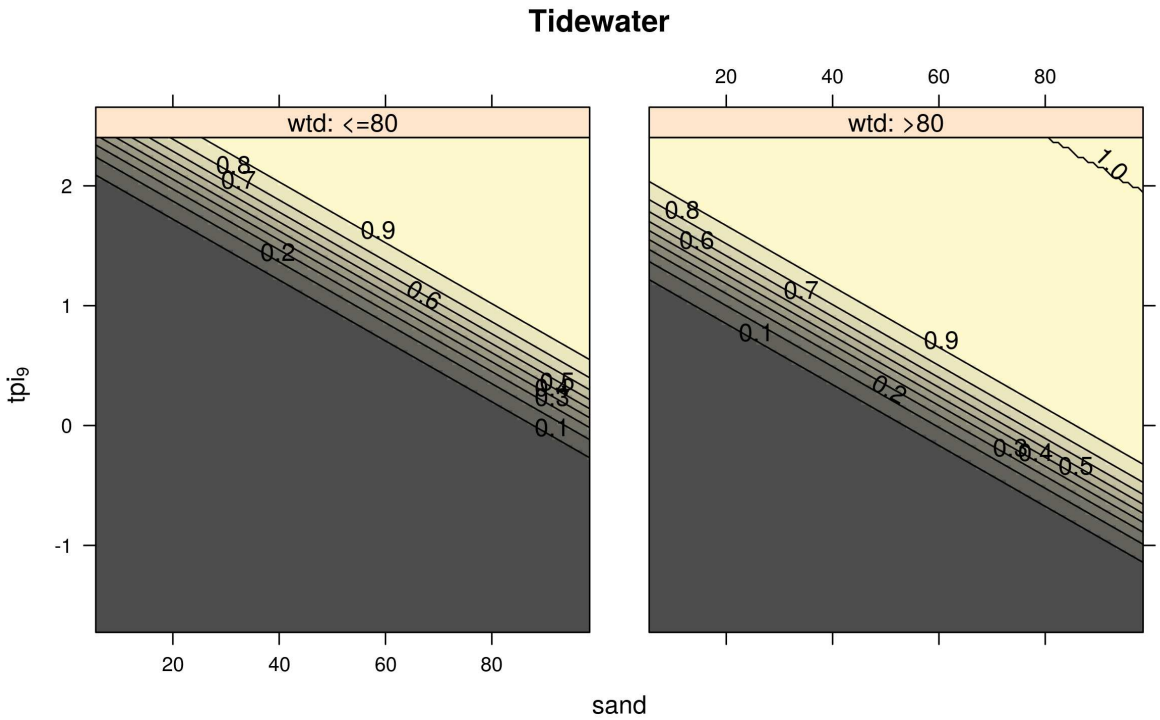

Figure 2.5.: Contour plots of burrow occurrence probability, for the **Tidewater** top model in Table 2.8. Each plot shows the expected occurrence probability, over the ranges of `sand` and $\texttt{tpi}_{9}$, at one level of the `wtd` variable, annual minimum depth to water table (cm).

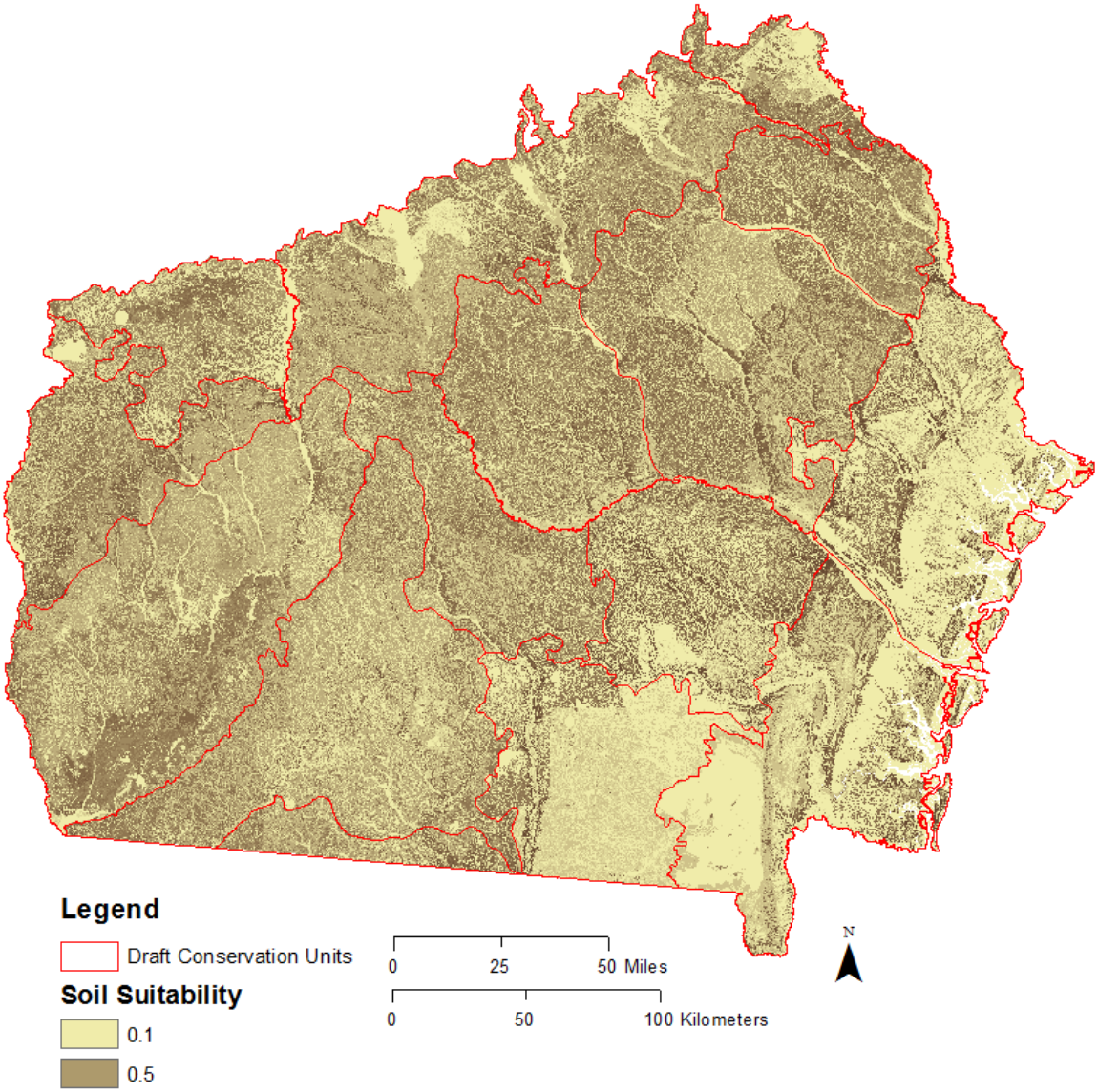

Figure 2.6.: Continuous (0.0–1.0) prediction of habitat suitability from the weighted composite model for each MLRA across the gopher tortoise range in Georgia. See section 2.3.1 for a description of this and the following maps.

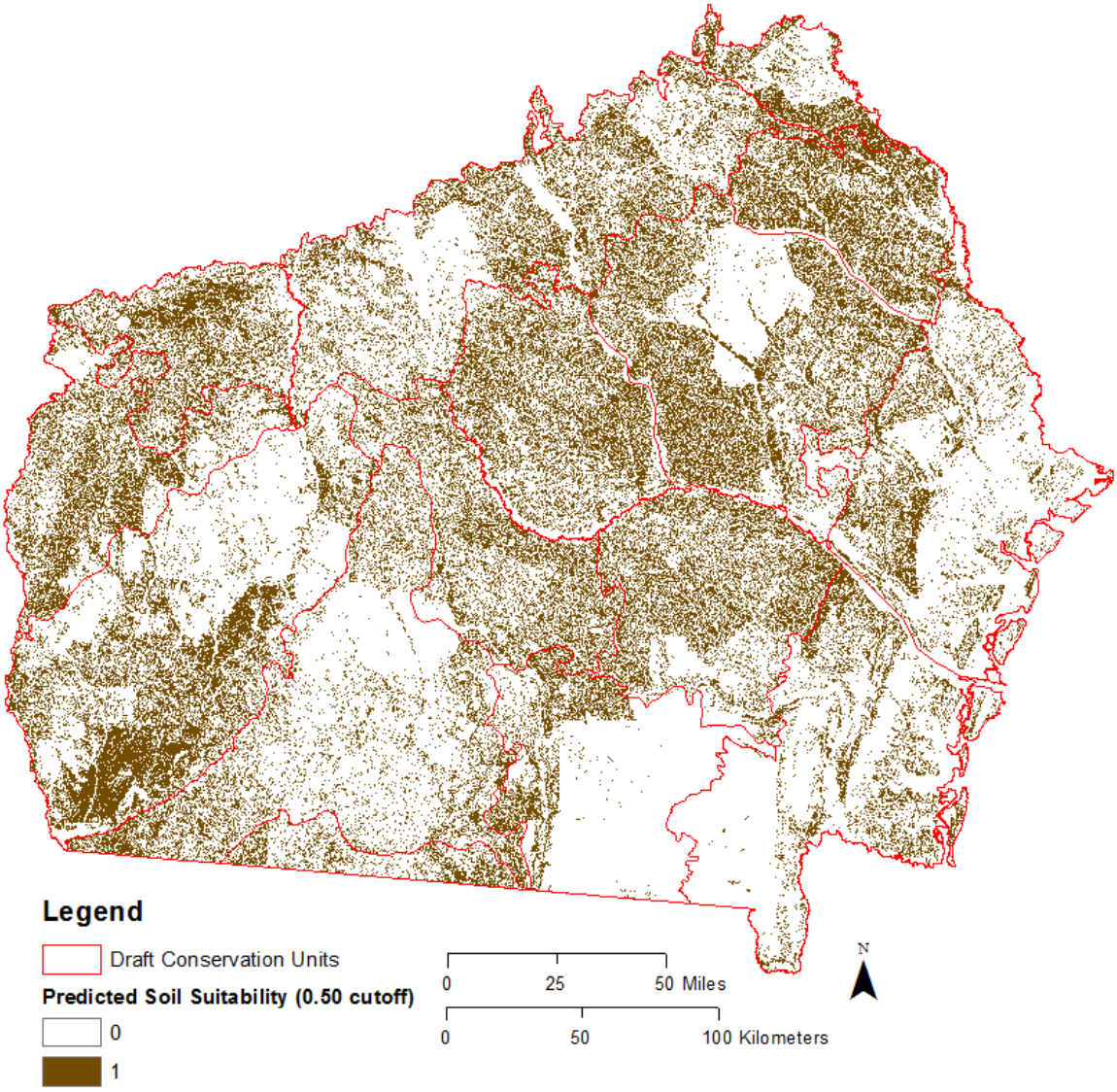

Figure 2.7.: Binary (*Habitat* = 1, *Not Habitat* = 0) classification of gopher tortoise habitat across its range in Georgia using a 0.50 cutoff value from the continuous habitat suitability prediction (Figure 2.6).

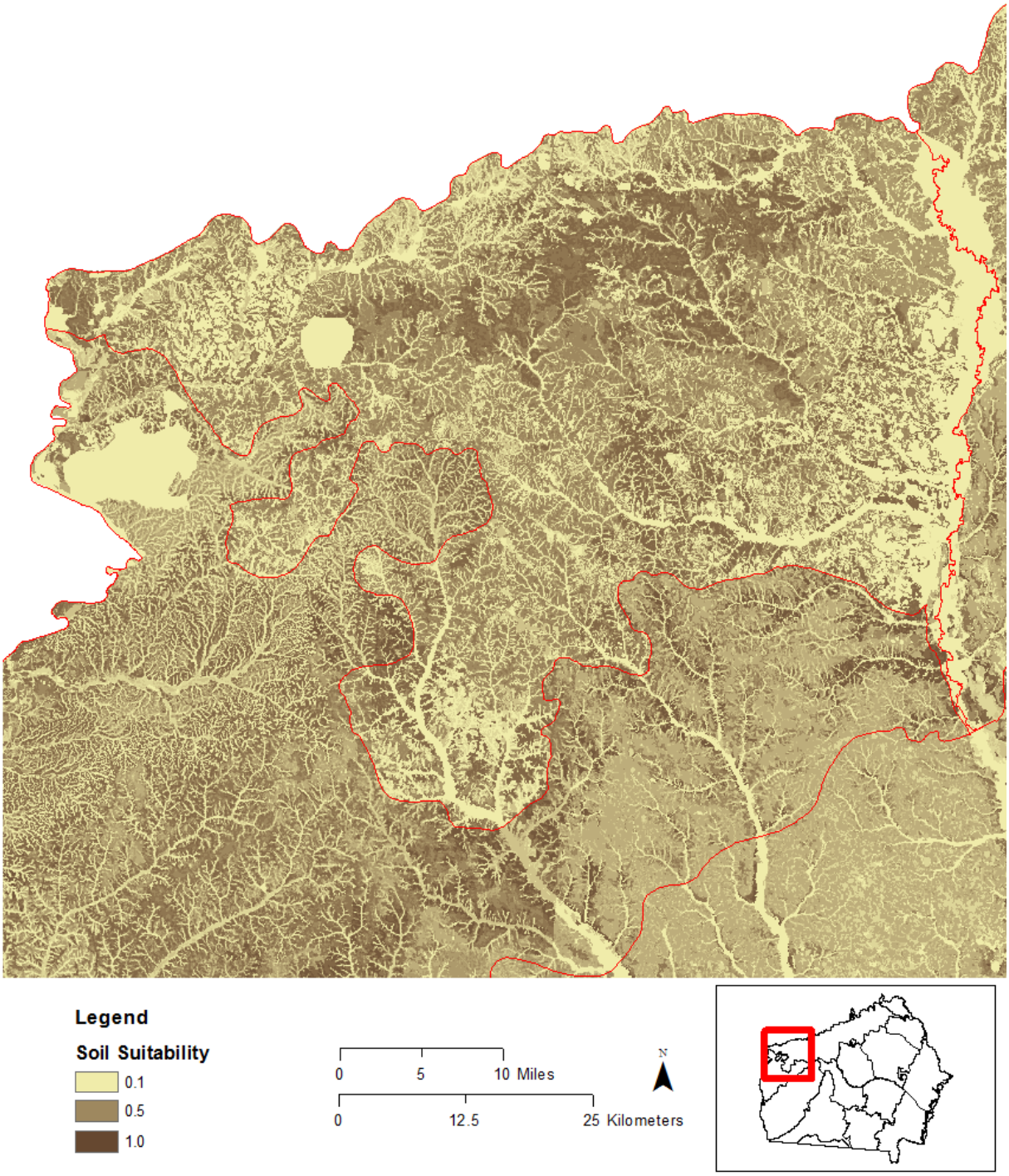

Figure 2.8.: Zoomed-in detail of the map in Figure 2.6, showing one of Georgia's draft tortoise conservation units outlined in red.

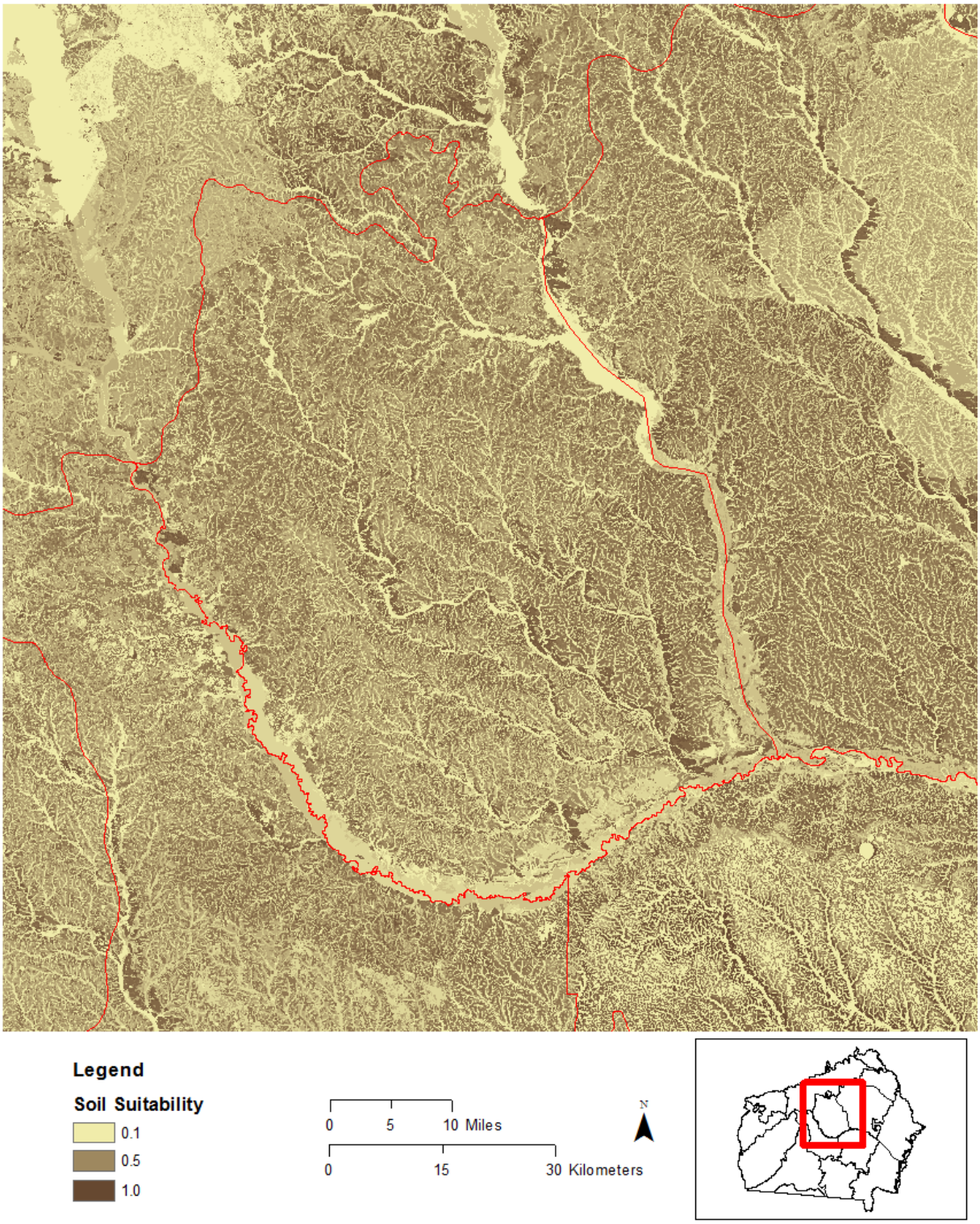

Figure 2.9.: Zoomed-in detail of the map in Figure 2.6, showing another of Georgia's draft tortoise conservation units outlined in red.

3 — Combining soil and remotely-sensed vegetation predictors to improve a presence-only tortoise habitat model

3.1. Introduction

Having identified a number of soil-related variables with high utility in discriminating burrows from pseudo-absences (Chapter 2), we combined these edaphic and landform predictors with a set that were vegetation-based, to construct models that could take advantage of both relatively permanent (soil, landform) and more transient (vegetation) landscape features. As in Chapter 2, we fit and ranked models separately within each Major Land Resource Area intersecting the gopher tortoise's range in Georgia (MLRA; see Section 2.2.2). Our objectives for this model selection exercise were to:

1. seek a confidence set of models of burrow occurrence (relative to pseudo-absences) with good predictive ability, for each MLRA, and
2. identify important vegetation and soil predictors in each MLRA, when these appeared together in candidate models
3. understand the shapes of the response, across the ranges of important variables

We expected the addition of vegetation-based predictors to improve models' predictive ability, since tortoises respond to both soil and vegetation conditions (Nussear and Tuberville, 2014). We also expected that the soil-based variables appearing in the confidence set of models for each MLRA would match those determined to be important in Section 2.3.

3.2. Variables

We collected a number of soil- and vegetation-based variables with established value in predicting the occurrence of gopher tortoise burrows in Georgia. These were used in candidate logistic regression models (Section 3.3).

3.2.1. Response

The response variable was identical to that used in Chapter 2, a binary indication of burrows versus pseudo-absences. Locations of burrows—both active and inactive—were collected during line transect distance sampling performed at 79 sites, by the Georgia DNR. Pseudo-absence placement and spatial filtering of burrow and pseudo-absence locations, is described in Section 2.2.1. The only difference, compared to the analysis described in Chapter 2, was that whereas there the number of pseudo-absences used was equal to the number of burrows in each MLRA, here we used twice as many pseudo-absences as burrows in the Tidewater MLRA. This change was made to more thoroughly sample the area within the annuli from which pseudo-absences were chosen, there being so few burrows in the Tidewater.

3.2.2. Soil-based predictors

We used the same soil variables identified and described in Section 2.2.3: the NRCS-produced gopher tortoise soil suitability index (symbolized by `soil`; USFWS and NRCS 2012; NRCS 2017*a*); percent sand (`sand`) and annual minimum depth to water table (`wtd`) from the National Soil Information System database (NRCS, 2017*b*); and a topographic position index (`tpi`; Jenness 2006). In contrast to the treatment of `wtd` in Chapter 2, where the number of levels of `wtd` varied across MLRA, we here used a two-level categorical version of depth to water table in each of the four MLRAs: ≤ 80 cm ($\texttt{wtd}_{\leq 80}$) and > 80 cm ($\texttt{wtd}_{>80}$).

3.2.3. Phenology-derived vegetation predictors

Gopher tortoises require both suitable soils and appropriate vegetation conditions. In particular, tortoises need sufficient forage in the herb layer, which in turn requires sufficient insolation of the forest floor during the growing season (MacDonald and Mushinsky, 1988; Mushinsky, 2014; Nussear and Tuberville, 2014). In pine forests inhabited by tortoises in Georgia, several general forest attributes can influence the availability of tortoise food resources, including canopy closure, and vegetation density in the understory and shrub strata. These attributes are not necessarily observable in extensive landcover map products such as the National Land Cover Dataset.

Previously, we investigated the feasibility of deriving (Hepinstall-Cymerman et al., 2017) vegetation attributes important to tortoise habitat quality, from temporal series of Landsat images. This exercise resulted in a set of predicted maps for the entire study area, of forest characteristics beneath the canopy. We used field observations as the basis for our vegetation attribute maps, and descriptions of the field and analytical methods are given in Appendices A–C.

We made use of the same set of vegetation attributes as were used in Hepinstall-Cymerman et al. (2017) (Table 3.1; Appendix C). In that work, we found an improvement in the ability of these attributes to predict burrow occurrence, when the raster layers representing them were first masked by a map of landcover types broadly inhabitable by gopher tortoises. This masking layer was generated by reclassifying the National Landcover Dataset (NLCD): pixels of classes 42, 43, 52 and 7 (evergreen forest, mixed forest, shrub/scrub, grassland/herbaceous) were assigned a value of 1, and all others 0. Masking improved the vegetation predictors' value, because spectral and phenological signatures were at times similar between habitats that are known to support tortoise populations, and those that do not. For example `fbgr` values in the unmasked map were similar in open pine stands with relatively dense ground cover and in agricultural fields.

3.2.4. Predictor variable smoothing

Animal populations may be expected to respond to habitat conditions at different spatial scales. To accommodate uncertainty as to the appropriate scale to use for each of our candidate variables, we smoothed all continuous variables except `sand` at three scales. Mean values were taken over moving windows of three sizes: 3 pixels by 3 (0.81 ha), 9 by 9 (7.29 ha), and 15 by 15 (20.25 ha). This mirrors the smoothing done in Chapter 2, described in section 2.2.3. The `sand` predictor was left unsmoothed because ease of burrowing is expected to depend only on local soil conditions. The final list of candidate predictors is shown in Table 3.2.

3.3. Model construction, ranking, validation

Our predictors were chosen based on *a priori* knowledge of habitat components that promote gopher tortoise occurrence. However, we had no *a priori* basis for building a small set of candidate models representing specific hypotheses. We therefore built all possible logistic regression models from our predictors, under the following constraints:

1. pairs of correlated predictors, with Pearson's $|r| > 0.6$, were not allowed in the same model, and

2. all two-way interactions and quadratic effects (for continuous variables) were allowed, provided the constituent main effects were also present.

Because all smoothed versions of the same variable (e.g., $\texttt{soil}_3$, $\texttt{soil}_9$, $\texttt{soil}_{15}$) were correlated with one another (i.e., $|r| > 0.6$), they never appeared in the same model with one another.

The same set of models ($n = 1{,}598{,}787$) was fit to burrow and pseudo-absence data in each of the four MLRAs. Models were fit using the `glm()` function in program R (R Core Team, 2017), and ranked via Akaike's Information Criterion (AIC) (Burnham and Anderson, 2002).

As in section 2.2.4, each model in the confidence set (ΔAIC < 4) underwent 100 iterations of five-fold cross-validation, and the area under the receiver operator characteristic curve (AUC) was calculated.

3.4. Results and Discussion

3.4.1. Model ranking, parameter estimates

We examined models within a confidence set defined as ΔAIC < 4 (Burnham and Anderson, 2002) (Tables 3.3–3.9). Confidence sets contained many models; this was unsurprising considering the enormous number of fitted models. Model complexity was high in all MLRAs: all favored models included multiple interactions and quadratic effects. Within each MLRA's confidence set, models did not differ in terms of the variables they contained; they did differ, however, in terms of smoothing scale of those variables, and particular two-way interactions and quadratic terms.

Whereas the analysis in Chapter 2 ranked models higher that contained soil attributes (`wtd`, `sand`) than those containing the NRCS soil index (`soil`), the presence of vegetation attributes seems to have engendered a context in which models based on soil index out-perform those based on soil attributes. The exception was the Sandhills MLRA, where `sand` and broad-scale $\texttt{tpi}_{15}$ were present in all favored models.

The one main effect that appeared in all models in all MLRAs' confidence sets was `pine`, the measure of pine dominance across all forest strata. The smoothing scale was not consistent across MLRA, but the coefficient for this main effect was always positive (Tables 3.4–3.10). Where a $\texttt{pine}^2$ effect was included, it was negative, suggesting a rapid increase in the benefit of increased `pine` to habitat suitability up to some threshold, beyond which the growth in the benefit is reduced (Sandhills and Coastal Plain; Figures 3.1 and 3.2). This relationship is complicated, in the Flatwoods and Tidewater MLRAs, by the presence of interactions involving `pine` (Figures 3.3 and 3.4).

Sandhills

Most confidence set models in the Sandhills MLRA contained main and quadratic effects of percent sand (`sand`), topographic position index smoothed at the 15x15 (20.25 ha) scale ($\texttt{tpi}_{15}$), prevalence of forbs and grasses in the herb layer smoothed at the 9x9 (7.29 ha) scale ($\texttt{fbgr}_9$), and our index of pine dominance (`pine`) at either the 15x15 or 9x9 scale (Table 3.3). All variables were involved in at least one interaction, in every favored model.

Parameter estimates for the top model are given in Table 3.4. As shown in Figure 3.1, burrow occurrence probability increases with increased sand content of the soil, up to some threshold where the benefit of increased sand plateaus. This response is a consequence of the negative value of $\texttt{sand}^2$; the exact shape of the curve is dependent on the values of the other variables in the model.

Intermediate values of `fbgr` are favored, which may indicate that areas occupied by tortoises tend to have considerable forage in the herb layer, as well as woody species (if more productive, mesic) or bare soil (less productive, xeric). This pattern is consistent across all the MLRAs where `fbgr` appears (i.e., all except Flatwoods; Figures 3.1, 3.2, 3.4).

Topographic position index induced a similar pattern in occurrence probability, with intermediate values favored. This is in contrast to the Flatwoods result, where higher values of $\texttt{tpi}_{15}$ are unequivocally favored. This could be because high relief areas are common in the Sandhills, whereas there is generally less relief in the Flatwoods: perhaps tortoises may be more reliably found on the highest ground in that flatter region.

Coastal Plain

All confidence set models in the Coastal Plain contained these main effects and associated quadratic terms: $\texttt{soil}_3$, $\texttt{fbgr}_3$, $\texttt{shrb}_{15}$, $\texttt{pine}_{15}$ (Table 3.5). Every model contained several two-way interactions. In the top-ranked model (Table 3.6) occurrence probability increased with $\texttt{soil}_3$ and $\texttt{pine}_{15}$, but was maximal at intermediate values of $\texttt{fbgr}_3$.

Burrow occurrence responded positively to the broadly-smoothed measure of vegetation density in the shrub layer ($\texttt{shrb}_{15}$), when the value of $\texttt{pine}_{15}$ was low; where pine dominance was moderate or high, however, occurrence probability was maximal at intermediate values of $\texttt{shrb}_{15}$ (Figure 3.2). This complexity in the effect of shrub-stratum density may result from the fact that several classes of plants occupy the shrub stratum in a pine forest: species with a true shrub habit; hardwood tree saplings; regenerating pine saplings. If suitable habitats with lower pine dominance are those with lower productivity (more xeric), stunted trees could also contribute to shrub stratum density; and in this case, they may not be inordinately shading the herb layer, since the whole stature of the forest is reduced.

The top model for the Tidewater also included a $\texttt{shrb}_{15}\texttt{:pine}_{15}$ interaction (Figure 3.4), where the general pattern across $\texttt{shrb}_{15}$ was much the same as in the Coastal Plain; however, the slope of the 'ridge' was there reversed, so that the highest burrow occurrence probability corresponded to the lowest values of $\texttt{pine}_{15}$ and the highest values of $\texttt{shrb}_{15}$: again, this may be due to stunted oaks on more xeric sites.

Flatwoods

Most confidence set models in the Flatwoods contained these main effects and associated quadratic terms: $\texttt{tpi}_{15}$, $\texttt{soil}_9$ (sometimes $\texttt{soil}_{15}$), $\texttt{covr}_3$, $\texttt{pine}_{15}$ (Table 3.7). Every model contained several two-way interactions. Occurrence probability generally increased with `soil` across the confidence set; this relationship is typified in the top model (Table 3.8).

Topographic position index, `tpi`, contains information about both slope and relative elevation, and ought to be more informative than slope alone, in describing tortoise habitat (see Section 2.3). Slope is one element of the NRCS soil index (USFWS and NRCS, 2012), and it may be argued that `soil` and `tpi` should not be allowed in the same model together (though they were not excessively correlated in the datasets). They did co-occur only in the confidence set of the Flatwoods, where a negative quadratic effect of both, as well as a negative-valued interaction between them, describes a plateau-shaped response across their ranges (Figure 3.3) indicating a compensatory relationship: higher quality soils promote higher occurrence probability at lower values of `tpi`; higher `tpi` promotes occurrence even at lower values of `soil`.

Interactions with $\texttt{pine}_{15}$ were interesting here. Occurrence probability rose rapidly with $\texttt{pine}_{15}$ or $\texttt{soil}_9$ when the other was at a low value, but more gradually when that was otherwise. The response across the ranges of $\texttt{pine}_{15}$ and maximum vegetation cover (across forest strata), $\texttt{covr}_3$, displayed a saddle shape, with occurrence probability highest when $\texttt{pine}_{15}$ was high and $\texttt{covr}_3$ low, or *vice versa*. At high values of $\texttt{pine}_{15}$, occurrence probability increases swiftly as

$\texttt{tpi}_{15}$ passes 0; at lower $\texttt{pine}_{15}$ values, the increase with $\texttt{tpi}_{15}$ is much slower; at the lowest $\texttt{pine}_{15}$ values even high $\texttt{tpi}_{15}$ values cannot raise the response probability above about 0.5.

Finally, very high values of $\texttt{covr}_3$ retard the increase in occurrence probability with increased $\texttt{tpi}_{15}$. This seems sensible, as dense vegetation in any stratum would shade the herb layer and degrade habitat.

Tidewater

All confidence set models in the Tidewater contained these main effects and associated quadratic terms: $\texttt{fbgr}_{15}$, $\texttt{shrb}_{15}$, $\texttt{pine}_9$ (Table 3.9). The soil index $\texttt{soil}_3$ appeared in every model, but often without an accompanying quadratic term. Every model contained several two-way interactions. Occurrence probability increased with $\texttt{soil}_3$.

As previously mentioned, occurrence probability here was maximal at intermediate values of $\texttt{shrb}_{15}$ and $\texttt{fbgr}_{15}$ (Figure 3.4). It was highest too when $\texttt{fbgr}_{15}$ was intermediate and $\texttt{pine}_9$ high. And as mentioned above, probability was maximal when $\texttt{pine}_9$ was lowest and $\texttt{shrb}_{15}$ highest.

Parameter estimates for several effects in the top Tidewater model were very large (Table 3.10). However, standard errors were not obviously outsized. The large estimates seem to be driven by the presence of interactions with $\texttt{fbgr}_{15}$, and it may be that this is a case of quasi-separation: one region of the $\texttt{fbgr}_{15}$-$\texttt{shrb}_{15}$ parameter plane, for instance, contains all the presences (Figure 3.5). The model attempts to describe a single 'mesa' of occurrence probability in this region. Small sample size, and a small number of surveys from which observations came, may also contribute to this issue.

3.4.2. Validation

Confidence set models had good to very high AUC values, averaged over 100 iterations of five-fold cross-validation. Median AUC values were in the interval [0.83, 0.84] for Sandhills, 0.89 for Coastal Plain; 0.92 for Flatwoods; and in [0.97, 0.98] for Tidewater. Interestingly, well-supported models in the MLRAs with less topographical relief (Flatwoods, Tidewater) showed higher predictive value, than those in MLRAs with more relief (Sandhills, Coastal Plain). The exceptionally high AUC in the Tidewater may have resulted from a scarcity of similar habitats, to those supporting surveyed gopher tortoise populations, within the 5-km annuli used for selecting pseudo-absences. Adding new tortoise populations to the Tidewater dataset would likely decrease the apparent predictive values of the confidence set models, by our measure.

3.4.3. Predictive maps

To visualize our model predictions on the study landscape, we produced a map for each MLRA that represents the average predicted burrow occurrence probability, averaged across the top models (those with $\Delta\text{AIC} < 2$). In ArcGIS Raster Calculator, we first made a predictive map for each of the top models for each MLRA, by making a linear combination of the appropriate predictor raster layers with the estimated coefficients from that model. Topographic position index layers were resampled to 30 m to match all other layers.

We then took a weighted average of the (logit-scale) values of the linear predictors from the set of top models, with the weights being the models' corresponding AIC weights (Tables 3.3, 3.5, 3.7, 3.9) normalized to sum to unity across the top models ($\Delta\text{AIC} < 2$) used to produce the map. The average linear predictor values were transformed into probabilities, and the composite prediction layers for each MLRA were mosaicked into a single map (Figure 3.6). Note that there is uncertainty in this and all subsequent predictive maps due both to parametric uncertainty within models, and to model uncertainty: there was no clear winning model in Tables 3.3–3.9. Therefore, assessed values at every point on maps shown in Figures 3.6–3.11 carry with them an associated degree of confidence; we have not shown these uncertainties.

To facilitate additional analysis, the continuous layer ($\widehat{\Pr}(\text{occurrence}) = \hat{o} \in [0, 1]$) was classified into *suitable* ($\hat{o} \geq 0.50$) and *unsuitable* ($\hat{o} < 0.50$) (Figure 3.7). Two additional analyses were done using this binary layer. First, because 100 hectares (≈250 acres) is a stated minimum goal for maintaining a viable gopher tortoise population (GTC, 2013), all patches of habitat smaller than this size (using an 8-way neighborhood rule) were eliminated (Figures 3.8–3.9).

A second use of the binary suitability map attributed parcel data, initially for two draft conservation planning units where complete parcel boundaries were available, with the total area of predicted gopher tortoise habitat (of any patch size) found within the parcel boundary (based on a parcel id attribute) (Figure 3.10–3.11). This preliminary analysis shows which parcels have large amounts of habitat within their boundaries. However, care must be taken with interpreting these results as some parcels span multiple physical patches and the predicted habitat within these parcels may be in fragments smaller than the total area reported.

Many additional analyses are possible with the predictions we have generated. For example, a "habitat capacity" analysis could be developed by combining the predicted habitat map (Figure 3.6) with soil suitability maps, similar to Figures 3.8–3.9, to indicate areas where soil conditions are suitable but vegetation conditions would require management actions to return the area to suitable habitat. Finally, EPA Level IV ecoregions (Omernik and Griffith, 2014) have entered common use as gopher tortoise conservation planning units. Table 4.1 gives estimates of total habitat area within each ecoregion in Georgia (or the intersection of the ecoregion with the state, where the former extends beyond Georgia's boundary).

Table 3.1.: Plot-level summaries of vegetation metrics originally measured in quadrats or subplots. In all cases, the summary is obtained by taking the mean across all sample units. (This table is very similar to Table A.3; but the variable `other` has here been omitted, since it was not included in the present analysis.)

Symbol	**Name**	**Sample unit**	**Measurement**
`fbgr`	Forb + grass cover	Quadrat, 2–100 cm layer	Sum of "forb" and "grass" percent cover
`dens`	Densiometer	Subplot	Mean of readings in four cardinal directions
`covr`	Maximum cover	Subplot	Maximum cover value across all strata
`shrb`	Shrub cover	Subplot	Percent cover value in the shrub stratum
`pine`	Pine dominance	Subplot	Pine-dominated strata assigned a value of 1, mixed 0 and hardwood -1; then a weighted average taken across strata, with the percent cover values* as the weights

* In a given stratum, cover value is for all vegetation; vegetation in that stratum is then labeled as being predominantly "pine", "mixed" or "hardwood".

Table 3.2.: Characteristics of variables in the gopher tortoise burrow dataset, including pseudo-absences used in this chapter, in each of the four MLRAs. The variable name column gives the categories, when a variable was categorical; in this case, the range/count column displays the number of cases with that value. Variable `burrow` is the response, and is 1 for burrows, 0 for pseudo-absences. The `sites` row gives the number of tortoise survey sites in that MLRA. Compare with Table 2.1. Note that the sum across the `sites` row is 90, which is greater than the number of sites (79); this is because some sites (and/or associated pseudo-absence annuli) extend across an MLRA boundary.

	Coastal Plain		Flatwoods		Sandhills		Tidewater	
Variable	Mean	Range/count	Mean	Range/count	Mean	Range/count	Mean	Range/count
$\texttt{burrow}_0$		2536		1242		796		156
$\texttt{burrow}_1$		2536		1242		796		78
`sites`		51		23		13		3
$\texttt{wt\_cat}_{\leq 80}$		989		1589		95		164
$\texttt{wt\_cat}_{>80}$		4083		895		1497		70
`sand`	77.1	[5.3, 97.9]	83.1	[0, 98.4]	84.0	[27.1, 95.0]	69.8	[7.7, 98.4]
$\texttt{tpi}_3$	0.07	[-2.8, 3.5]	0.07	[-2.5, 2.3]	0.06	[-3.4, 4.4]	0.01	[-0.8, 0.7]
$\texttt{tpi}_9$	0.33	[-11.6, 13.2]	0.24	[-4.7, 6.2]	0.34	[-10.2, 12.3]	0.08	[-1.0, 2.4]
$\texttt{tpi}_{15}$	0.54	[-18.5, 18.2]	0.40	[-5.2, 7.2]	0.65	[-16.6, 16.1]	0.15	[-1.7, 3.8]

Continued next page...

Table 3.2.: Variable characteristics continued.

	Coastal Plain		Flatwoods		Sandhills		Tidewater	
Variable	Mean	Range/count	Mean	Range/count	Mean	Range/count	Mean	Range/count
$\texttt{soil}_3$	0.63	[0, 1]	0.39	[0, 1]	0.82	[0, 1]	0.29	[0, 1]
$\texttt{soil}_9$	0.59	[0, 1]	0.35	[0, 1]	0.80	[0, 1]	0.28	[0, 1]
$\texttt{soil}_{15}$	0.57	[0, 1]	0.32	[0, 1]	0.79	[0, 1]	0.27	[0, 1]
$\texttt{covr}_3$	14.7	[0, 38.6]	14.2	[0, 40.8]	11.4	[0, 37.6]	15.2	[0, 38.9]
$\texttt{covr}_9$	14.6	[0, 37.3]	14.0	[0, 36.7]	11.5	[0, 35.1]	15.1	[0, 38.3]
$\texttt{covr}_{15}$	14.4	[0, 36.4]	13.7	[0, 35.1]	11.6	[0, 32.8]	14.9	[0, 36.3]
$\texttt{dens}_3$	35.2	[0, 85.5]	34.4	[0, 84.7]	26.7	[0, 83.8]	35.1	[0, 84.5]
$\texttt{dens}_9$	34.9	[0, 82.1]	33.9	[0, 84.0]	27.1	[0, 81.4]	34.9	[0, 82.4]
$\texttt{dens}_{15}$	34.3	[0, 79.2]	33.0	[0, 79.1]	27.2	[0, 76.0]	34.3	[0, 79.1]
$\texttt{fbgr}_3$	16.7	[0, 64.0]	14.1	[0, 59.5]	18.1	[0, 63.9]	4.0	[0, 33.5]
$\texttt{fbgr}_9$	15.6	[0, 61.3]	13.0	[0, 52.4]	17.5	[0, 61.9]	3.7	[0, 27.6]
$\texttt{fbgr}_{15}$	14.9	[0, 60.5]	12.2	[0, 43.4]	17.1	[0, 58.0]	3.7	[0, 25.1]
$\texttt{shrb}_3$	18.6	[0, 58.3]	18.8	[0, 63.2]	14.3	[0, 45.6]	16.8	[0, 53.5]
$\texttt{shrb}_9$	18.4	[0, 51.3]	18.2	[0, 53.1]	14.5	[0, 42.1]	16.5	[0, 51.2]
$\texttt{shrb}_{15}$	18.1	[0, 48.8]	17.7	[0, 52.9]	14.5	[0, 36.5]	16.2	[0, 48.0]
$\texttt{pine}_3$	0.06	[-0.59, 0.57]	0.09	[-0.66, 0.55]	0.10	[-0.58, 0.52]	-0.10	[-0.69, 0.36]
$\texttt{pine}_9$	0.05	[-0.49, 0.53]	0.08	[-0.63, 0.52]	0.09	[-0.53, 0.47]	-0.10	[-0.60, 0.28]
$\texttt{pine}_{15}$	0.05	[-0.39, 0.49]	0.08	[-0.61, 0.51]	0.08	[-0.47, 0.41]	-0.09	[-0.49, 0.26]

Table 3.3.: Top models of burrow occurrence, in the **Sandhills** MLRA, ranked by AIC. Models within 4 AIC points of the top model are shown.
Column definitions:
1) Column ω shows the AIC weights, calculated as follows: if, for the ith model of a set numbering M, $q_i = e^{(-0.5*\Delta AIC_i)}$, then $w_i = q_i/(\sum_{i=1}^{M} q_i)$, with $\Delta AIC_i = AIC_i - \min(AIC)$.
2) Column K gives the number of model parameters (including the intercept).
3) Columns representing variables are filled when that variable appears in the model, with the following codes: the italicized number shows the scale of the variable (3,9,15); a "2" indicates that a quadratic term for the variable was included; and capital letters show that the variable participated in a two-way interaction with a second variable matched to the letter shown (letters with which variables are matched appear above their names in column labels). For example the value "*9:* 2,B,C" shows that that column's variable appeared in the model smoothed at the 9x9 scale, that the model included the quadratic term of the variable, and that it was included in two-way interactions with both `sand` (labeled B in the table) and `tpi` (labeled C).

			A	B	C	D	E	F	G	H	I
AIC	ω	K	`wtd`	`sand`	`tpi`	`soil`	`covr`	`dens`	`fbgr`	`shrb`	`pine`
1534.1	0.041	12		2,G,I	*15:* 2,G				*9:* 2,B,C		*15:* 2,B
1534.7	0.030	11		2,I	*15:* 2,G				*9:* 2,C		*15:* 2,B
1534.8	0.028	12		2,G,I	*15:* 2,G				*9:* 2,B,C		*9:* 2,B
1535.0	0.026	11		2,I	*15:* 2,G				*9:* 2,C		*9:* 2,B
1535.5	0.020	13		2,C,G,I	*15:* 2,B,G				*9:* 2,B,C		*15:* 2,B
1535.7	0.019	13		2,G,I	*15:* 2,G,I				*9:* 2,B,C		*15:* 2,B,C
1535.9	0.017	13		2,C,G,I	*15:* 2,B,G				*9:* 2,B,C		*9:* 2,B
1536.0	0.016	12		2,I	*15:* 2,G				*9:* 2,C,I		*9:* 2,B,G
1536.1	0.015	13		2,G,I	*15:* 2,G				*9:* 2,B,C,I		*15:* 2,B,G
1536.1	0.015	13		2,G,I	*15:* 2,G,I				*9:* 2,B,C		*9:* 2,B,C
1536.2	0.014	12		2,I	*15:* 2,G,I				*9:* 2,C		*9:* 2,B,C
1536.3	0.014	13		2,G,I	*15:* 2,G				*9:* 2,B,C,I		*9:* 2,B,G
1536.3	0.014	12		2,I	*15:* 2,G,I				*9:* 2,C		*15:* 2,B,C
1536.4	0.013	11		G,I	*15:* 2,G				*9:* 2,B,C		*15:* 2,B
1536.4	0.013	12		2,C,I	*15:* 2,B,G				*9:* 2,C		*15:* 2,B
1536.4	0.013	12		2,C,I	*15:* 2,B,G				*9:* 2,C		*9:* 2,B
1536.5	0.012	12		2,G,I	*15:* 2,G				*15:* 2,B,C		*9:* 2,B
1536.6	0.012	11		G,I	*15:* 2,G				*9:* 2,B,C		*9:* 2,B
1536.6	0.011	10		I	*15:* 2,G				*9:* 2,C		*9:* 2,B
1536.7	0.011	12		2,I	*15:* 2,G				*9:* 2,C,I		*15:* 2,B,G
1536.8	0.011	10		I	*15:* 2,G				*9:* 2,C		*15:* 2,B
1536.8	0.010	11		2,I	*15:* 2,G				*15:* 2,C		*9:* 2,B
1537.0	0.010	14		2,C,G,I	*15:* 2,B,G,I				*9:* 2,B,C		*9:* 2,B,C
1537.0	0.009	14		2,C,G,I	*15:* 2,B,G,I				*9:* 2,B,C		*15:* 2,B,C

Continued next page...

Table 3.3.: **Sandhills** top models continued.

			A	B	C	D	E	F	G	H	I
AIC	ω	K	`wtd`	`sand`	`tpi`	`soil`	`covr`	`dens`	`fbgr`	`shrb`	`pine`
1537.2	0.009	13		2,C,G,I	*15:* 2,B,G				*15:* 2,B,C		*9:* 2,B
1537.4	0.008	12		2,I	*15:* 2,G				*9:* 2,C,I		*3:* 2,B,G
1537.4	0.008	14		2,C,G,I	*15:* 2,B,G				*9:* 2,B,C,I		*9:* 2,B,G
1537.4	0.008	13		2,G,I	*15:* 2,G,I				*15:* 2,B,C		*9:* 2,B,C
1537.4	0.008	13		2,I	*15:* 2,G,I				*9:* 2,C,I		*9:* 2,B,C,G
1537.4	0.008	13		2,C,I	*15:* 2,B,G				*9:* 2,C,I		*9:* 2,B,G
1537.5	0.007	14		2,C,G,I	*15:* 2,B,G				*9:* 2,B,C,I		*15:* 2,B,G
1537.5	0.007	13		2,C,I	*15:* 2,B,G,I				*9:* 2,C		*9:* 2,B,C
1537.6	0.007	14		2,G,I	*15:* 2,G,I				*9:* 2,B,C,I		*15:* 2,B,C,G
1537.7	0.007	12		2,I	*15:* 2,G,I				*15:* 2,C		*9:* 2,B,C
1537.7	0.007	14		2,G,I	*15:* 2,G,I				*9:* 2,B,C,I		*9:* 2,B,C,G
1537.7	0.007	11		I	*15:* 2,G,I				*9:* 2,C		*9:* 2,B,C
1537.7	0.007	12		G,I	*15:* 2,G,I				*9:* 2,B,C		*9:* 2,B,C
1537.8	0.006	12		G,I	*15:* 2,G,I				*9:* 2,B,C		*15:* 2,B,C
1537.8	0.006	11		I	*15:* 2,G				*9:* 2,C,I		*9:* 2,B,G
1537.8	0.006	11		2,I	*15:* 2,G				*9:* 2,C		*3:* 2,B
1537.8	0.006	14		2,C,G,I	*15:* 2,B,G,I				*15:* 2,B,C		*9:* 2,B,C
1537.9	0.006	12		C,G,I	*15:* 2,B,G				*9:* 2,B,C		*9:* 2,B
1537.9	0.006	13		2,C,I	*15:* 2,B,G,I				*9:* 2,C		*15:* 2,B,C
1538.0	0.006	12		C,G,I	*15:* 2,B,G				*9:* 2,B,C		*15:* 2,B

Table 3.4.: Parameter estimates for the top **Sandhills** model of burrow occurrence (see Table 3.3).
Column definitions:
1) For each term, the point estimate and standard error are given on the logit scale.
2) For the intercept, the corresponding probability estimate is provided, with its associated 95% confidence interval; the value shown reflects the case where all predictor variables are at their means. For all other terms, the estimate is exponentiated to give the odds ratio, with confidence interval. For these terms, statistical significance at the $\alpha = 0.05$ level is inferred by exclusion of 1.0 from the confidence interval.
3) Each continuous variable was centered around its mean and divided by its standard deviation before model fitting: this standard deviation represents the unit change associated with the given odds ratio. For example, with all other variables held at their means, an increase of 20.00% in the `sand` variable increase the odds of finding a burrow 5.75 * 0.69 = 3.97 times (i.e., product of odds ratios for `sand` and $\texttt{sand}^2$: the interactions may be ignored when other variables are at their means). See Table 3.2 for variable means and ranges.

Term	Estimate	SE	Odds ratio	Probability	95% CI	Unit change
Intercept	-1.35	0.14		0.21	[0.16, 0.25]	
`sand`	1.75	0.20	5.75		[4.01, 8.68]	20.00
$\texttt{tpi}_{15}$	0.40	0.06	1.48		[1.33, 1.67]	2.03
$\texttt{fbgr}_{9}$	0.83	0.15	2.30		[1.73, 3.07]	10.70
$\texttt{pine}_{15}$	0.99	0.15	2.68		[2.01, 3.65]	0.12
$\texttt{sand}^2$	-0.37	0.20	0.69		[0.47, 1.01]	
$\texttt{tpi}^2_{15}$	-0.07	0.02	0.93		[0.9, 0.96]	
$\texttt{fbgr}^2_{9}$	-0.19	0.04	0.83		[0.77, 0.89]	
$\texttt{pine}^2_{15}$	-0.23	0.06	0.80		[0.71, 0.89]	
$\texttt{sand:fbgr}_{9}$	0.21	0.17	1.23		[0.89, 1.72]	
$\texttt{sand:pine}_{15}$	-0.49	0.18	0.61		[0.42, 0.87]	
$\texttt{tpi}_{15}\texttt{:fbgr}_{9}$	-0.15	0.04	0.86		[0.8, 0.93]	

Table 3.5.: Top models of burrow occurrence, in the **Coastal Plain** MLRA, ranked by AIC. Models within 4 AIC points of the top model are shown. Column definitions as in Table 3.3.

			A	B	C	D	E	F	G	H	I
AIC	ω	K	`wtd`	`sand`	`tpi`	`soil`	`covr`	`dens`	`fbgr`	`shrb`	`pine`
4191.4	0.139	13				*3:* 2,G,I			*3:* 2,D,I	*15:* 2,I	*15:* 2,D,G,H
4191.9	0.109	14				*3:* 2,G,I			*3:* 2,D,H,I	*15:* 2,G,I	*15:* 2,D,G,H
4192.5	0.082	14				*3:* 2,G,H,I			*3:* 2,D,I	*15:* 2,D,I	*15:* 2,D,G,H
4192.5	0.081	12				*3:* 2,G			*3:* 2,D,I	*15:* 2,I	*15:* 2,G,H
4192.9	0.067	15				*3:* 2,G,H,I			*3:* 2,D,H,I	*15:* 2,D,G,I	*15:* 2,D,G,H
4193.1	0.061	13				*3:* 2,G			*3:* 2,D,H,I	*15:* 2,G,I	*15:* 2,G,H
4193.7	0.045	12				*3:* 2,G,I			*3:* 2,D	*15:* 2,I	*15:* 2,D,H
4193.9	0.040	13				*3:* 2,G,H			*3:* 2,D,I	*15:* 2,D,I	*15:* 2,G,H
4194.4	0.031	14				*3:* 2,G,H			*3:* 2,D,H,I	*15:* 2,D,G,I	*15:* 2,G,H
4194.5	0.030	13				*3:* 2,G,I			*3:* 2,D,H	*15:* 2,G,I	*15:* 2,D,H
4194.6	0.028	13				*3:* 2,G,H,I			*3:* 2,D	*15:* 2,D,I	*15:* 2,D,H
4194.6	0.028	11				*3:* 2,G			*3:* 2,D	*15:* 2,I	*15:* 2,H
4195.3	0.020	14				*3:* 2,G,H,I			*3:* 2,D,H	*15:* 2,D,G,I	*15:* 2,D,H

Table 3.6.: Parameter estimates for the top **Coastal Plain** model of burrow occurrence (see Table 3.5). Column definitions as in Table 3.4.

Term	Estimate	SE	Odds ratio	Probability	95% CI	Unit change
Intercept	-1.82	0.07		0.14	[0.12, 0.16]	
$\texttt{soil}_3$	1.02	0.05	2.78		[2.53, 3.05]	0.41
$\texttt{fbgr}_3$	0.83	0.06	2.29		[2.05, 2.58]	13.10
$\texttt{shrb}_{15}$	1.02	0.06	2.78		[2.48, 3.11]	9.75
$\texttt{pine}_{15}$	0.64	0.06	1.90		[1.68, 2.14]	0.12
$\texttt{soil}_3^2$	0.30	0.05	1.35		[1.22, 1.50]	
$\texttt{fbgr}_3^2$	-0.23	0.03	0.79		[0.75, 0.84]	
$\texttt{shrb}_{15}^2$	-0.24	0.04	0.78		[0.73, 0.84]	
$\texttt{pine}_{15}^2$	-0.07	0.02	0.93		[0.89, 0.98]	
$\texttt{soil}_3\texttt{:fbgr}_3$	-0.12	0.04	0.88		[0.82, 0.96]	
$\texttt{soil}_3\texttt{:pine}_{15}$	-0.07	0.04	0.93		[0.86, 1.00]	
$\texttt{fbgr}_3\texttt{:pine}_{15}$	0.08	0.04	1.08		[1.00, 1.16]	
$\texttt{shrb}_{15}\texttt{:pine}_{15}$	-0.18	0.04	0.84		[0.77, 0.91]	

Table 3.7.: Top models of burrow occurrence, in the **Flatwoods** MLRA, ranked by AIC. Models within 4 AIC points of the top model are shown. Column definitions as in Table 3.3.

			A	B	C	D	E	F	G	H	I
AIC	ω	K	`wtd`	`sand`	`tpi`	`soil`	`covr`	`dens`	`fbgr`	`shrb`	`pine`
1736.8	0.090	13			*15:* 2,D,E,I	*9:* 2,C,I	*3:* 2,C,I				*15:* C,D,E
1737.9	0.050	14			*15:* 2,D,E,I	*9:* 2,C,I	*3:* 2,C,I				*15:* 2,C,D,E
1738.2	0.044	14			*15:* 2,D,E,I	*9:* 2,C,E,I	*3:* 2,C,D,I				*15:* C,D,E
1738.5	0.038	13			*15:* 2,D,E,I	*15:* 2,C,I	*3:* 2,C,I				*15:* C,D,E
1738.6	0.036	12			*15:* 2,D,I	*9:* 2,C,I	*3:* 2,I				*15:* C,D,E
1738.6	0.035	12			*15:* 2,D,E	*9:* 2,C,I	*3:* 2,C,I				*15:* D,E
1738.9	0.031	12			*15:* D,E,I	*9:* 2,C,I	*3:* 2,C,I				*15:* C,D,E
1738.9	0.030	13			*15:* 2,D,E	*9:* 2,C,I	*3:* 2,C,I				*15:* 2,D,E
1739.3	0.025	14			*15:* 2,D,E,I	*15:* 2,C,I	*3:* 2,C,I				*15:* 2,C,D,E
1739.4	0.024	11			*15:* D,I	*9:* 2,C,I	*3:* 2,I				*15:* C,D,E
1739.4	0.023	13			*15:* 2,D,I	*9:* 2,C,E,I	*3:* 2,D,I				*15:* C,D,E
1739.5	0.023	15			*15:* 2,D,E,I	*9:* 2,C,E,I	*3:* 2,C,D,I				*15:* 2,C,D,E
1739.8	0.020	13			*15:* 2,D,E	*15:* 2,C,I	*3:* 2,C,I				*15:* 2,D,E
1739.8	0.019	13			*15:* 2,D,I	*9:* 2,C,I	*3:* 2,I				*15:* 2,C,D,E
1740.0	0.018	14			*15:* 2,D,E,I	*15:* 2,C,E,I	*3:* 2,C,D,I				*15:* C,D,E
1740.0	0.018	12			*15:* 2,D,E	*15:* 2,C,I	*3:* 2,C,I				*15:* D,E
1740.0	0.018	13			*15:* 2,D,E	*9:* 2,C,E,I	*3:* 2,C,D,I				*15:* D,E
1740.1	0.017	12			*15:* D,I	*9:* 2,C,E,I	*3:* 2,D,I				*15:* C,D,E
1740.1	0.017	13			*15:* D,E,I	*9:* 2,C,I	*3:* 2,C,I				*15:* 2,C,D,E
1740.1	0.017	13			*15:* D,E,I	*9:* 2,C,E,I	*3:* 2,C,D,I				*15:* C,D,E
1740.5	0.014	14			*15:* 2,D,E	*9:* 2,C,E,I	*3:* 2,C,D,I				*15:* 2,D,E
1740.7	0.013	12			*15:* D,I	*9:* 2,C,I	*3:* 2,I				*15:* 2,C,D,E

Table 3.8.: Parameter estimates for the top **Flatwoods** model of burrow occurrence (see Table 3.7). Column definitions as in Table 3.4.

Term	Estimate	SE	Odds ratio	Probability	95% CI	Unit change
Intercept	-0.38	0.09		0.41	[0.37, 0.45]	
$\texttt{tpi}_{15}$	2.58	0.16	13.20		[9.72, 18.1]	2.03
$\texttt{soil}_{9}$	1.54	0.06	4.69		[4.19, 5.25]	0.36
$\texttt{covr}_{3}$	0.80	0.07	2.22		[1.94, 2.55]	11.30
$\texttt{pine}_{15}$	0.48	0.05	1.62		[1.47, 1.79]	0.12
$\texttt{tpi}_{15}^2$	-0.37	0.11	0.69		[0.56, 0.86]	
$\texttt{soil}_{9}^2$	-0.28	0.06	0.76		[0.67, 0.86]	
$\texttt{covr}_{3}^2$	-0.77	0.06	0.46		[0.41, 0.52]	
$\texttt{tpi}_{15}\texttt{:soil}_{9}$	-1.37	0.13	0.25		[0.2, 0.33]	
$\texttt{tpi}_{15}\texttt{:cov}_{3}$	-0.19	0.13	0.83		[0.64, 1.08]	
$\texttt{tpi}_{15}\texttt{:pine}_{15}$	0.16	0.12	1.18		[0.93, 1.49]	
$\texttt{soil}_{9}\texttt{:pine}_{15}$	-0.27	0.05	0.76		[0.7, 0.83]	
$\texttt{covr}_{3}\texttt{:pine}_{15}$	-0.55	0.05	0.58		[0.52, 0.64]	

Table 3.9.: Top models of burrow occurrence, in the **Tidewater** MLRA, ranked by AIC. Models within 4 AIC points of the top model are shown. Column definitions as in Table 3.3.

			A	B	C	D	E	F	G	H	I
AIC	ω	K	`wtd`	`sand`	`tpi`	`soil`	`covr`	`dens`	`fbgr`	`shrb`	`pine`
79.7	0.069	11				*3*			*15:* 2,H,I	*15:* 2,G,I	*9:* 2,G,H
80.3	0.051	12				*3:* I			*15:* 2,H,I	*15:* 2,G,I	*9:* 2,D,G,H
81.6	0.027	12				*3:* H			*15:* 2,H,I	*15:* 2,D,G,I	*9:* 2,G,H
81.6	0.027	12				*3:* H,I			*15:* 2,H,I	*15:* D,G,I	*9:* 2,D,G,H
81.6	0.027	11				*3:* H			*15:* 2,H,I	*15:* D,G,I	*9:* 2,G,H
81.6	0.027	12				*3:* G			*15:* 2,D,H,I	*15:* 2,G,I	*9:* 2,G,H
81.7	0.025	12				*3:* 2,			*15:* 2,H,I	*15:* 2,G,I	*9:* 2,G,H
82.0	0.022	13				*3:* 2,I			*15:* 2,H,I	*15:* 2,G,I	*9:* 2,D,G,H
82.1	0.021	12				*3:* 2,I			*15:* 2,H,I	*15:* G,I	*9:* 2,D,G,H
82.2	0.020	13				*3:* H,I			*15:* 2,H,I	*15:* 2,D,G,I	*9:* 2,D,G,H
82.3	0.020	13				*3:* G,I			*15:* 2,D,H,I	*15:* 2,G,I	*9:* 2,D,G,H
82.3	0.019	11				*3:* I			*15:* 2,H,I	*15:* G,I	*9:* 2,D,G,H
82.4	0.018	11				*3:* 2,I			*15:* 2,I	*15:* I	*9:* 2,D,G,H
82.5	0.017	13				*3:* 2,G,I			*15:* 2,D,H,I	*15:* G,I	*9:* 2,D,G,H
83.2	0.012	13				*3:* 2,H,I			*15:* 2,H,I	*15:* D,G,I	*9:* 2,D,G,H
83.3	0.012	12				*3:* G,H			*15:* 2,D,H,I	*15:* D,G,I	*9:* 2,G,H
83.3	0.012	12				*3:* G,I			*15:* 2,D,H,I	*15:* G,I	*9:* 2,D,G,H
83.3	0.012	10				*3:* I			*15:* 2,I	*15:* I	*9:* 2,D,G,H
83.3	0.011	13				*3:* G,H,I			*15:* 2,D,H,I	*15:* D,G,I	*9:* 2,D,G,H
83.5	0.011	13				*3:* G,H			*15:* 2,D,H,I	*15:* 2,D,G,I	*9:* 2,G,H
83.6	0.010	13				*3:* 2,G			*15:* 2,D,H,I	*15:* 2,G,I	*9:* 2,G,H
83.6	0.010	13				*3:* 2,H			*15:* 2,H,I	*15:* 2,D,G,I	*9:* 2,G,H
83.6	0.010	12				*3:* 2,H			*15:* 2,H,I	*15:* D,G,I	*9:* 2,G,H
83.6	0.010	10				*3*			*15:* 2,H,I	*15:* G,I	*9:* 2,G,H
83.7	0.010	14				*3:* 2,G,I			*15:* 2,D,H,I	*15:* 2,G,I	*9:* 2,D,G,H

Table 3.10.: Parameter estimates for the top **Tidewater** model of burrow occurrence (see Table 3.9). Column definitions as in Table 3.4.

Term	Estimate	SE	Odds ratio	Probability	95% CI	Unit change
Intercept	-19.60	4.87		0.00	[≈0, ≈0]	
$\texttt{soil}_3$	4.00	0.79	54.60		[14.4, 336]	0.41
$\texttt{fbgr}_{15}$	-48.20	12.00	0.00		[≈0, ≈0]	9.54
$\texttt{shrb}_{15}$	14.70	4.18	2.30×10^6		[1540, 2.59×10^{10}]	9.75
$\texttt{pine}_9$	9.04	2.09	8.4×10^3		[226, 9.45×10^5]	0.14
$\texttt{fbgr}_{15}^2$	-29.20	7.60	0.00		[≈0, ≈0]	
$\texttt{shrb}_{15}^2$	-2.35	0.93	0.10		[0.01, 0.50]	
$\texttt{pine}_9^2$	-0.78	0.26	0.46		[0.25, 0.73]	
$\texttt{fbgr}_{15}\texttt{:shrb}_{15}$	12.20	3.94	2.05×10^5		[183, 1.26×10^9]	
$\texttt{fbgr}_{15}\texttt{:pine}_9$	9.58	2.19	1.45×10^4		[324, 2.05×10^6]	
$\texttt{shrb}_{15}\texttt{:pine}_9$	-2.30	0.80	0.10		[0.02, 0.42]	

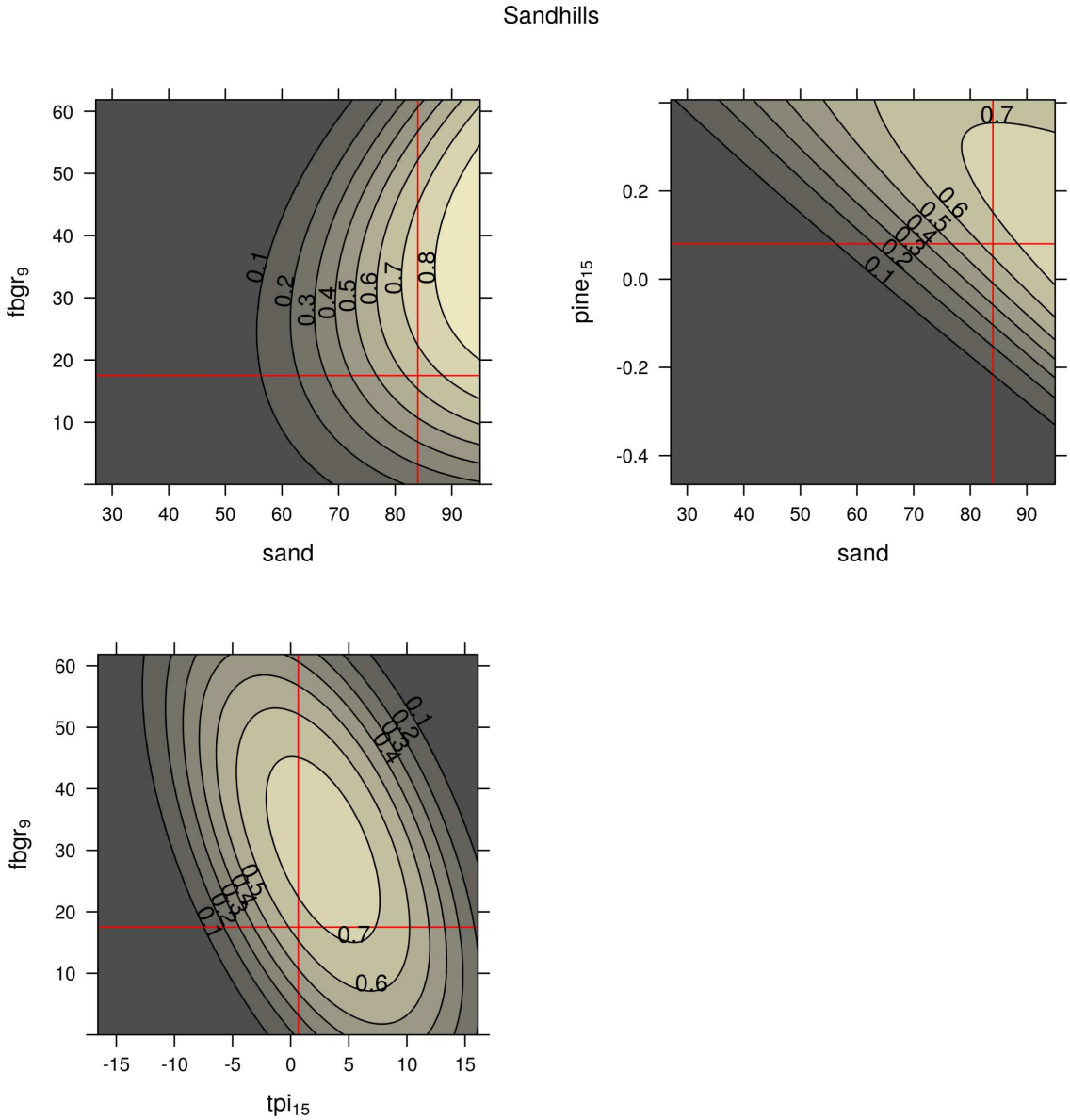

Figure 3.1.: Contour plots of expected burrow occurrence probability, for the **Sandhills** top model in Table 3.3 (parameter estimates in Table 3.4). Each plot shows the expected occurrence probability across the ranges of each variable involved in a two-way interaction. Red lines indicate the means of each variable.

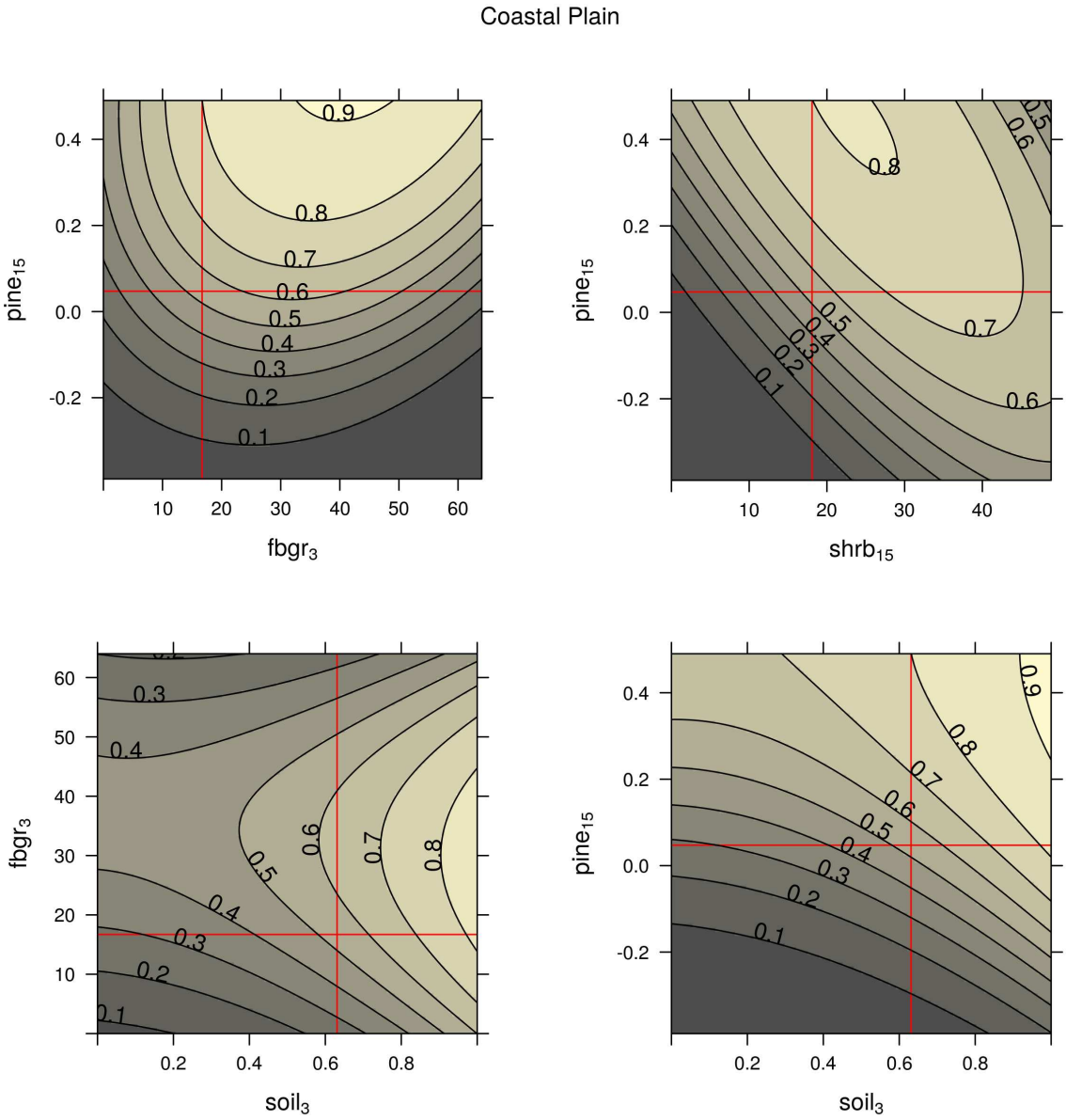

Figure 3.2.: Contour plots of expected burrow occurrence probability, for the **Coastal Plain** top model in Table 3.5 (parameter estimates in Table 3.6). Each plot shows the expected occurrence probability across the ranges of each variable involved in a two-way interaction. Red lines indicate the means of each variable.

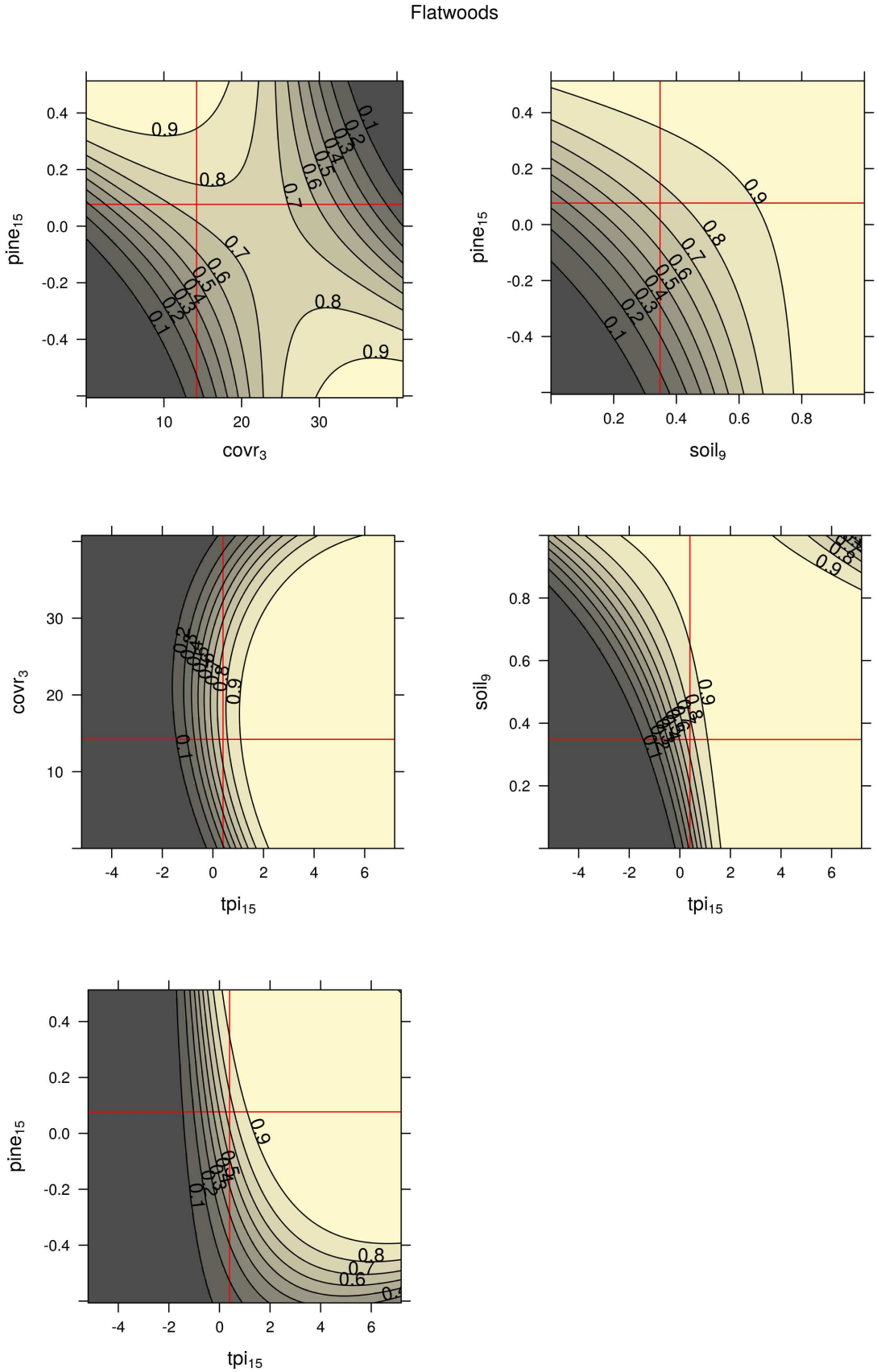

Figure 3.3.: Contour plots of expected burrow occurrence probability, for the **Flatwoods** top model in Table 3.7 (parameter estimates in Table 3.8). Each plot shows the expected occurrence probability across the ranges of each variable involved in a two-way interaction. Red lines indicate the means of each variable.

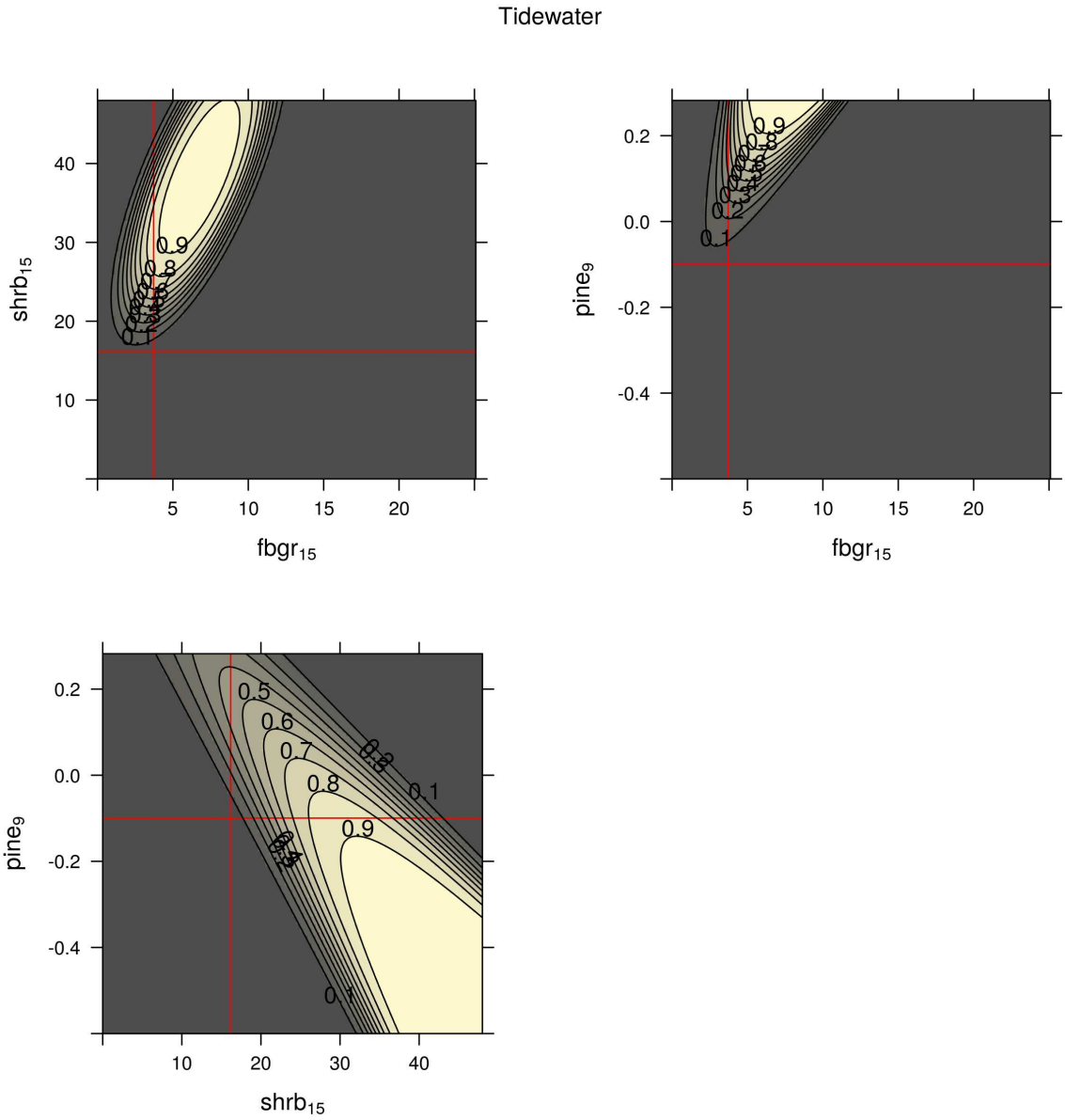

Figure 3.4.: Contour plots of expected burrow occurrence probability, for the **Tidewater** top model in Table 3.9 (parameter estimates in Table 3.10). Each plot shows the expected occurrence probability across the ranges of each variable involved in a two-way interaction. Red lines indicate the means of each variable.

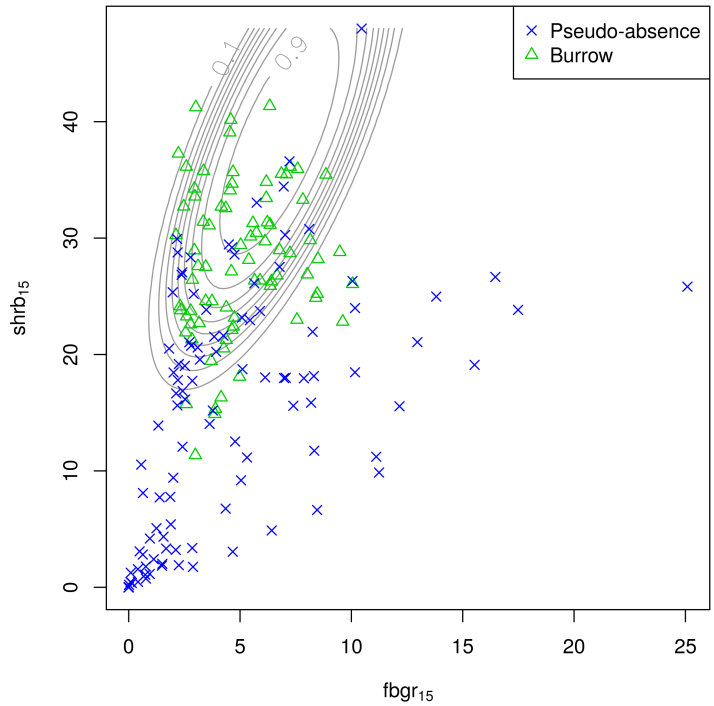

Figure 3.5.: Burrows and pseudo-absences in the Tidewater MLRA, plotted in the $\texttt{shrb}_{15}$-$\texttt{fbgr}_{15}$ plane. Burrows are restricted to a small portion of the plane, and this may contribute to the inflated parameter estimates in Table 3.10. Gray contours show the expected burrow occurrence probability according to the top model in Table 3.9; they correspond to the contours in Figure 3.4.

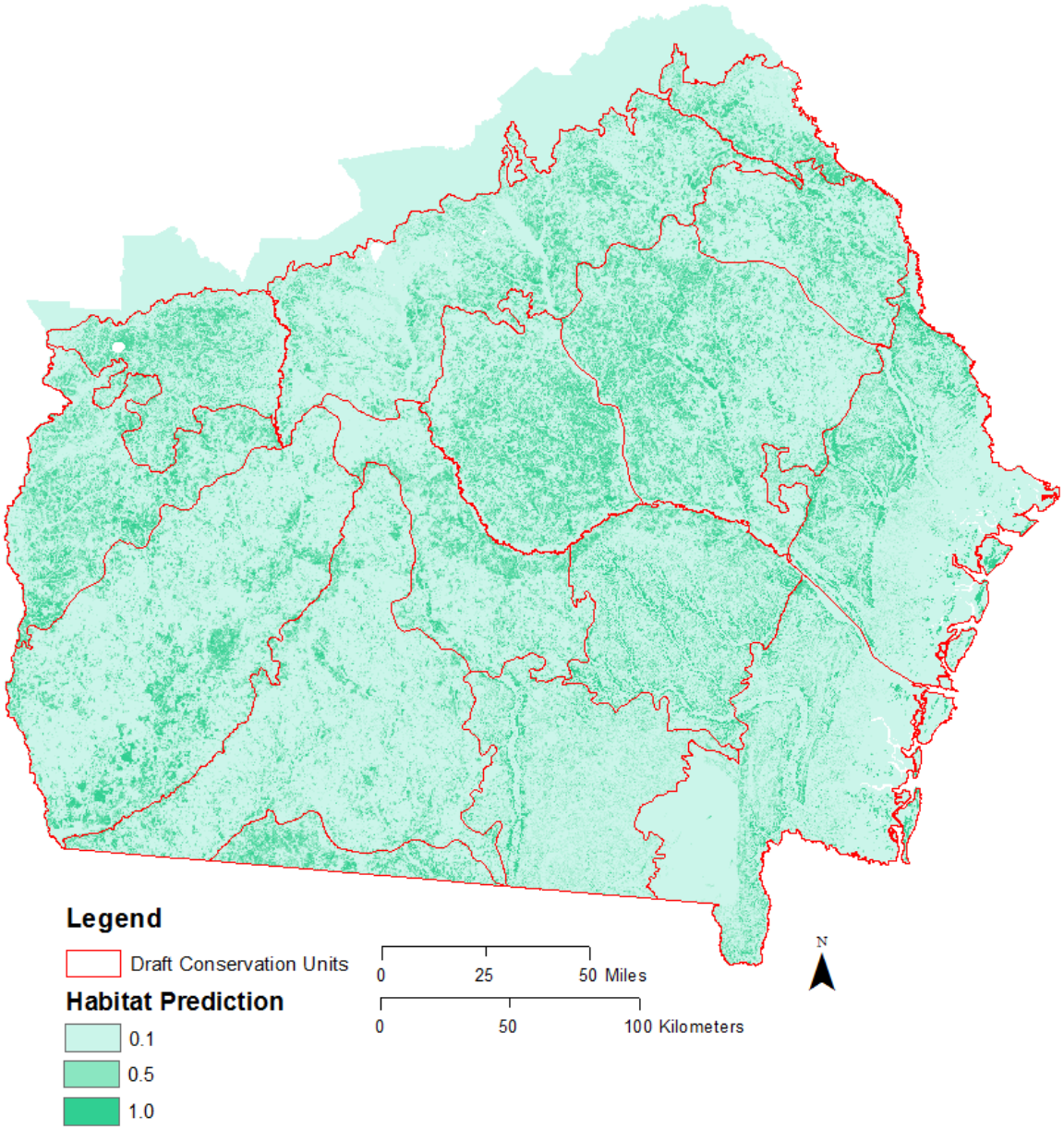

Figure 3.6.: Continuous (0.0–1.0) prediction of habitat suitability from the weighted composite model for each MLRA across the gopher tortoise range in Georgia. See section 3.4.3 for a description of this and the following maps.

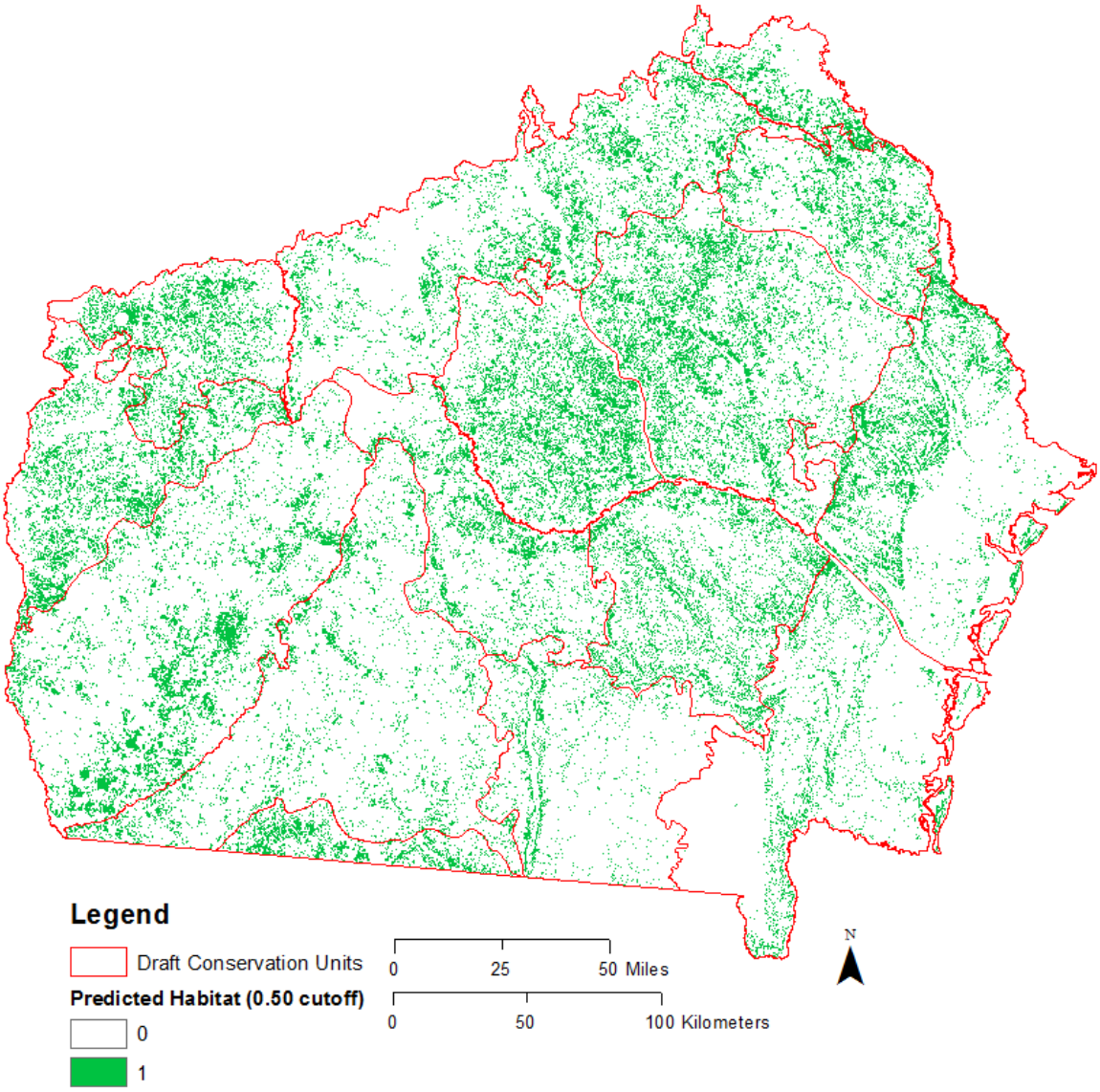

Figure 3.7.: Binary (*Habitat* = 1, *Not Habitat* = 0) classification of gopher tortoise habitat across its range in Georgia using a 0.50 cutoff value from the continuous habitat suitability prediction (Figure 3.6).

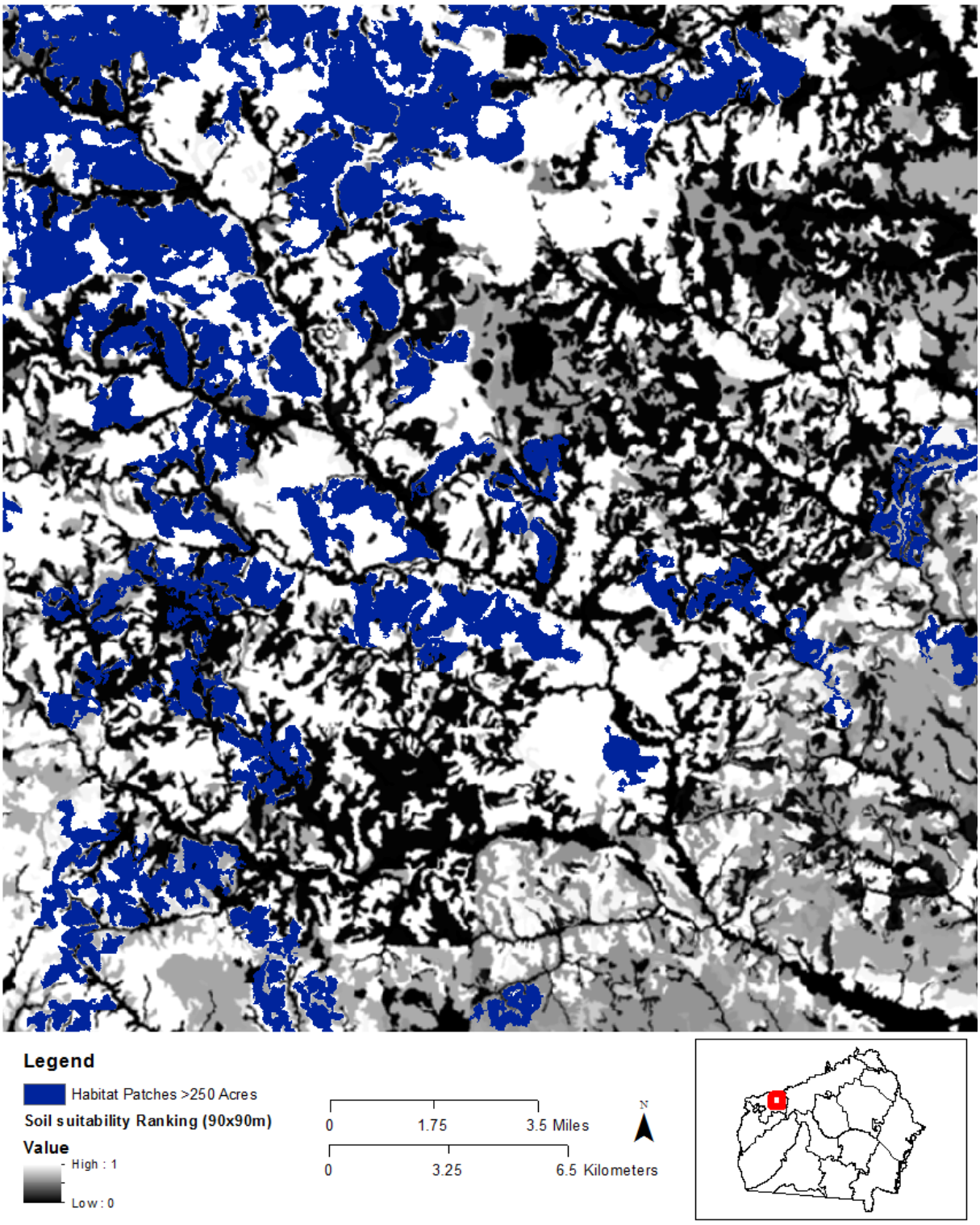

Figure 3.8.: Zoomed-in view of large (>100 ha) patches of suitable habitat in the foreground and average soil suitability ranking for a 90x90-m window, in one of a set of gopher tortoise conservation planning units (inset map). Light-shaded areas in the underlying suitability layer (grayscale), are predicted to have high capacity to support gopher tortoises, given appropriate vegetation management (possibly including restoration).

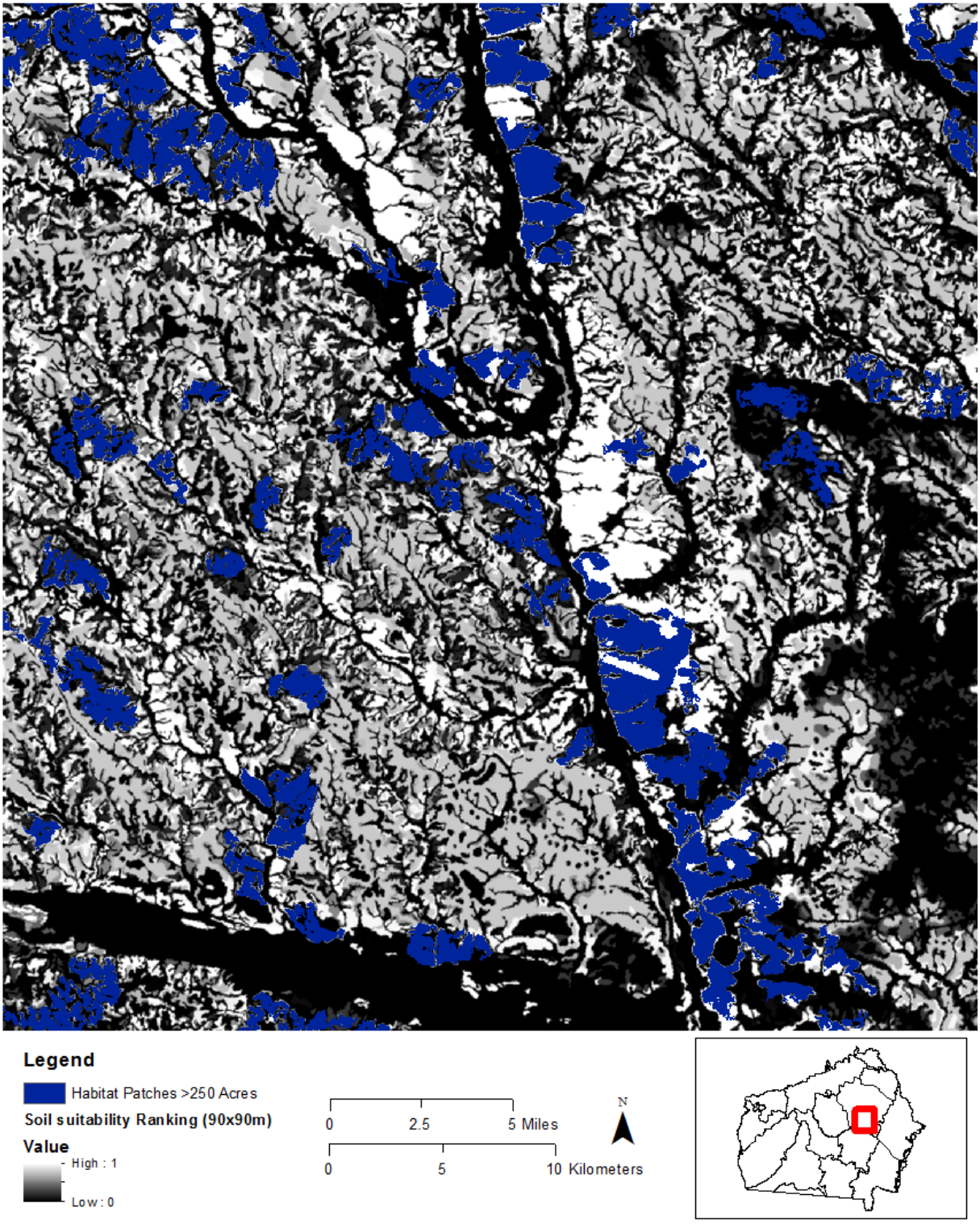

Figure 3.9.: Zoomed-in view of large (>100 ha) patches of suitable habitat in the foreground and average soil suitability ranking for a 90x90-m window, in one of a set of gopher tortoise conservation planning units (inset map). Light-shaded areas in the underlying suitability layer (grayscale), are predicted to have high capacity to support gopher tortoises, given appropriate vegetation management (possibly including restoration).

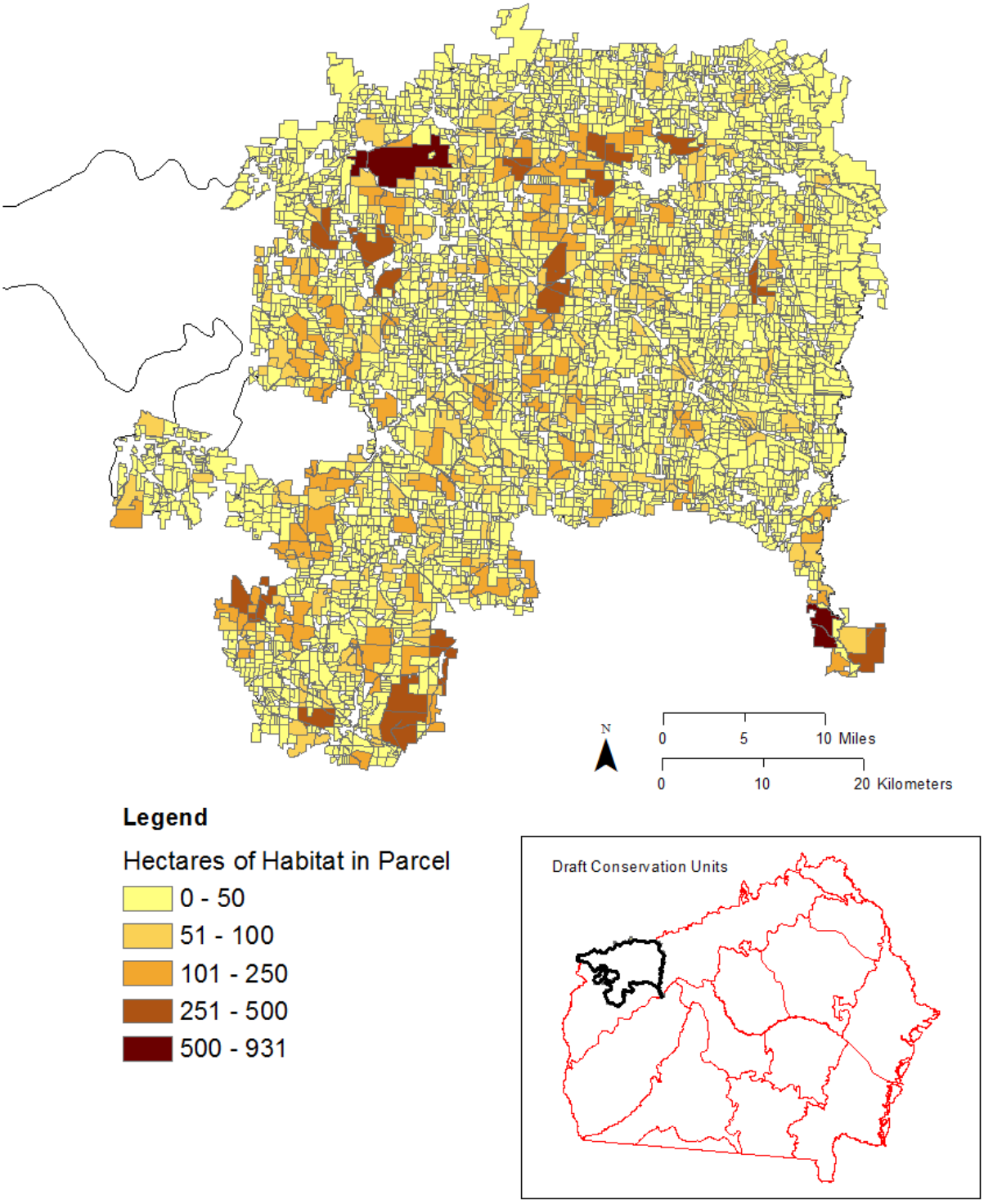

Figure 3.10.: Map depicting amount of predicted habitat in parcels, in one of a set of gopher tortoise conservation planning units (inset map). Note: since this was done by parcel ID, some parcels that exist as separate units will have the total area of the combined parcel polygons considered.

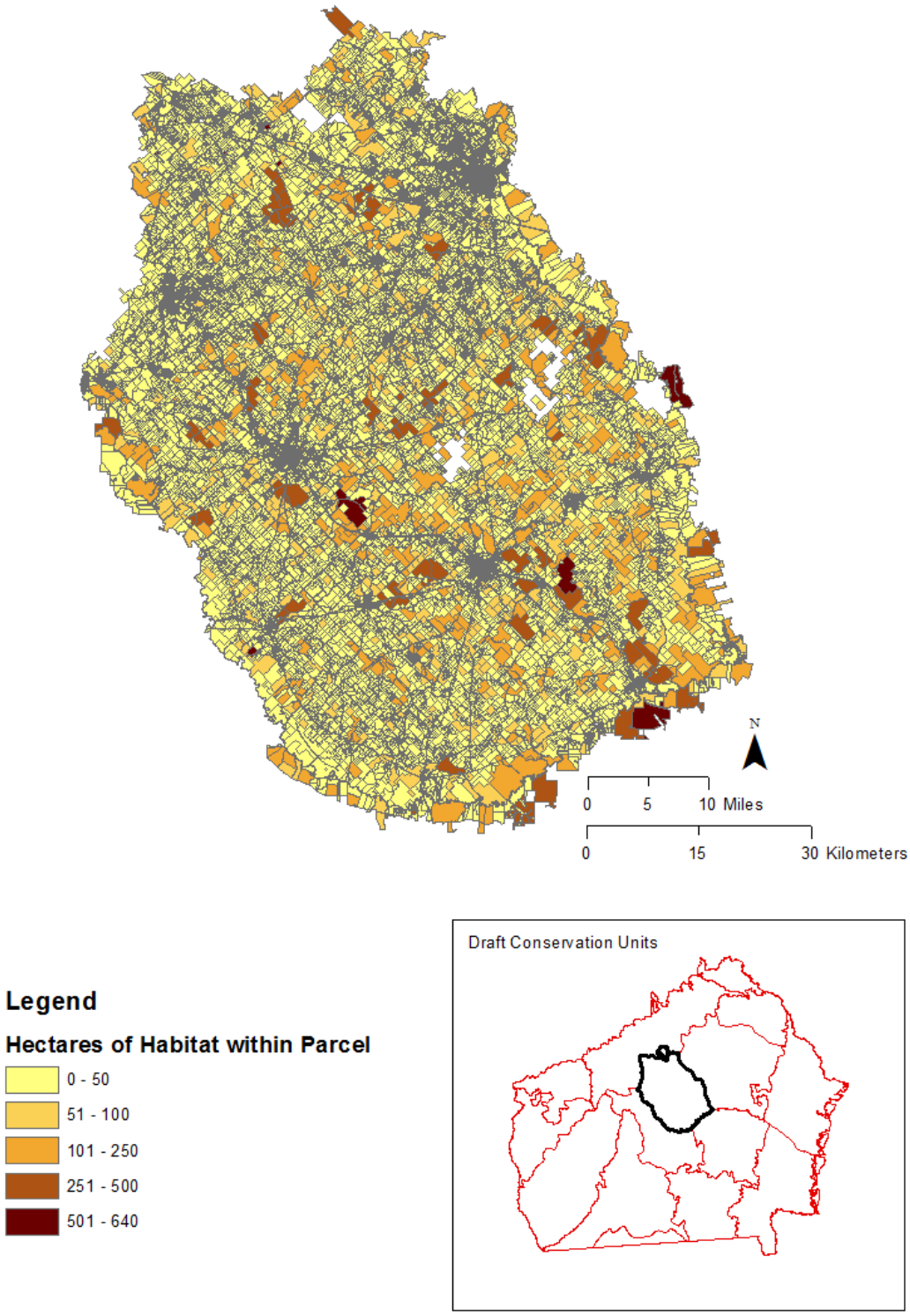

Figure 3.11.: Map depicting amount of predicted habitat in parcels, in one of a set of gopher tortoise conservation planning units (inset map). Note: since this was done by parcel ID, some parcels that exist as separate units will have the total area of the combined parcel polygons considered.

4 — Habitat modeling synthesis and conclusions

4.1. Predictor variables

In Chapter 2 we established that specific attributes of soils may be more useful in predicting gopher tortoise burrow locations in Georgia, than a composite index of soil suitability in common use (USFWS and NRCS, 2012; NRCS, 2017*a*). However, as shown in Chapter 3, this general conclusion must be amended when predictors representing vegetation conditions are available: in that case, model selection in all MLRAs except the Sandhills favored the NRCS soil index over specific soil attributes (`wtd`, `sand`). A possible explanation for this result, is that the information in the soil and vegetation attributes is somewhat redundant: both are finer-grained than the soil index, being better able to distinguish between areas that share a similar soil index value. However, vegetation is transient and closely tied to changes in land use and management, whereas soil properties are comparatively permanent (though even these may be affected by major disruptions such as tilling, grading or erosion).

The relationships we observed between soil and vegetation predictors, have several implications. First, as mentioned in Section 3.4.3, maps of suitable abiotic conditions (soil, slope) might be used to understand "habitat capacity," or potential to support suitable habitat, with appropriate management or restoration. For this purpose, specific soil attributes showed superior predictive value compared to the soil index in Georgia, and the use of the soil index as a "coarse filter" for habitat potential appears to be no better than use of the specific soil attributes.

Second, our results here and in Hepinstall-Cymerman et al. (2017) support making as many of the following substitutions as possible, when modeling gopher tortoise habitat: 1) topographic position index, for slope; 2) phenology-based vegetation attributes, for landcover maps such as reclassified NLCD; and 3) specific soil attributes, for the NRCS soil suitability index. But adding phenology-based vegetation attributes appears to ameliorate any shortcomings of the soil index as a predictor. In other words, the value of that index is dependent upon the availability of other high-utility predictors of habitat condition.

4.2. Comparison of predictive maps

Favored models that drew exclusively on soil-derived predictor variables (Section 2.3.1, Figure 2.6) consistently predicted more suitable area (Table 4.1) in larger patches (e.g., Figure 4.1), than the models that used both soil- and vegetation-derived predictors (Section 3.4.3, Figure 3.6). This result is generally expected, since large expanses of appropriate soils are not currently in an appropriate land use class and management regime to support tortoises. Of greater interest perhaps is the discrepancy between the predictions. As seen in Table 4.1, when considering binary predictions (averaged over top models and using a threshold of 0.5) the total area of suitable habitat is approximately half that of suitable soils in the Sandhills and Coastal Plain MLRAs; closer two thirds in the Flatwoods; and about one third in Tidewater. Although the area of suitable soils might be interpreted as representing a ceiling on the area in Georgia inhabitable

by tortoises, in practice many such areas would be impossible or enormously costly to restore. A large part of the predicted suitable habitat might still require considerable resources to restore to optimal condition.

A second difference is clear, in comparing the statewide maps shown in Figures 2.6–2.7 and 3.6–3.7: in the former pair of maps (Chapter 2), some county boundaries are distinctly visible (Bryan/Bulloch, Emanuel, Richmond, Twiggs/Wilkinson, Ware), and seemingly artificial differences in habitat suitability are apparent at the county level; in the latter pair (Chapter 3), such differences are not obvious. The soil surveys that inform the SSURGO database are primarily conducted and quality-assured at the county level. Discrepancies in soil properties, spatial arrangement, and even soil type ascribed to the same soil formation, where it crosses county borders, is a known and challenging issue in soil survey mapping (Thompson et al., 2012). Perhaps unsurprisingly, habitat predictions produced from SSURGO-derived soil properties alone appear to be more sensitive to such inter-county data artifacts, than when the predictors comprised the NRCS soil suitability index and vegetation attributes. Until further spatial 'harmonization' of SSURGO data is achieved within the gopher tortoise's range (Thompson et al., 2012), it may be necessary to rely on the soil index—along with vegetation characteristics—to predict tortoise habitat, merely to account for county-level differences in attribution of soil formations.

4.3. Considerations for field work

Our results in Chapters 2 and 3 indicate two fruitful areas for future field work. First, model selection and cross-validation results (Tables 2.8 and 3.9) and parameter estimates (Tables 2.9 and 3.10) point to a need for more gopher tortoise burrow surveys in the Tidewater MLRA: some parameter estimates were unrealistic, and AUC values used to assess cross-validation results were much higher than expected. Such surveys would not necessarily need to follow a line transect distance modeling methodology.

A field assessment of model predictions, based on visitation to a statistically drawn collection of sites, would help to validate these models. Vegetation observations during such visits could be used to refine models underpinning the phenology-based vegetation variables we used as predictors in Chapter 3 (Appendix C).

4.4. Caveats

Finally, it is worth emphasizing two major caveats to our conclusions. First, gopher tortoise survey sites in Georgia are not chosen randomly. Rather, sites are surveyed after they are known to contain populations of tortoises (investigating tortoise habitat characteristics is not a primary objective of these surveys). This non-random sampling scheme may have led to an artificial increase in the precision of our parameter estimates, and possibly bias in those estimates as well.

Secondly, and because surveys were conducted only where tortoises had been previously observed, we were constrained to use pseudo-absences rather than true absences, a so-called 'presence-only' modeling approach and one that, though it is widely used (Elith et al., 2011), has been criticized (Yackulic et al., 2013). We were forced to make decisions about how and where in the landscape to distribute those pseudo-absences, and our choices may have influenced our model selection results and models' parameter estimates. It is possible, for instance, that a number of our pseudo-absences fell within real but as yet unsurveyed tortoise populations. Some procedure for formally identifying areas known not to support gopher tortoises, would be of great value to future gopher tortoise habitat modeling. However, conclusions based upon our presence-only models have their own inherent value, because they reflect habitats that are both *suitable* and *accessible*. The latter kind of area may be more easily found, surveyed, put

under some form of protection, and ultimately managed in a manner that encourages tortoise population stability.

Table 4.1.: Predicted gopher tortoise habitat within Level IV Ecoregions (US Environmental Protection Agency). The ecoregions are perfect subdivisions of the NRCS MLRAs, with the exception of the 'Southeastern Floodplains and Low Terraces', which crosses the boundary of the Sandhills and Coastal Plain MLRAs. Where MLRAs or ecoregions cross state boundaries, absolute areas in hectares are given only for that part of the region contained in Georgia. Percent habitat columns give the percent area of the ecoregion or MLRA that intersects Georgia, that is predicted to be suitable habitat, when the prediction values are subjected to a cutoff threshold of 0.5 (see Figures 2.7 and 3.7). The *Soil only* column derives from the model-averaged prediction from Chapter 2; column *Soil and vegetation* derives from the corresponding prediction from Chapter 3.

				Percent habitat	
MLRA	EPA code	Ecoregion name	Area in GA	*Soil only*	*Soil and vegetation*
Sandhills (excluding floodplains)	65c	Sand Hills	723,505	36.1	17.9
Coastal Plain (excluding floodplains)	65d	Southern Hilly Gulf Coastal Plain	265,715	39.9	21.2
	65g	Dougherty Plain	1,072,496	31.3	14.2
	65h	Tifton Upland	1,008,228	20.1	9.3
	65k	Coastal Plain Red Uplands	838,074	32.0	20.0
	65l	Atlantic Southern Loam Plains	2,233,777	40.8	21.9
	65o	Tallahassee Hills/Valdosta Limesink	140,949	29.8	26.0
		Total	**5,559,241**	**33.9**	**17.9**
Flatwoods	75e	Okefenokee Plains	590,538	13.8	7.2
	75f	Sea Island Flatwoods	1,016,021	22.1	16.3
	75g	Okefenokee Swamp	179,214	0.5	0.2
	75h	Bacon Terraces	476,914	36.4	19.5
	75i	Floodplains and Low Terraces	71,975	13.7	10.5
		Total	**2,334,662**	**21.0**	**13.2**
Tidewater	75j	Sea Islands/Coastal Marsh	312,682	17.0	6.5
Sandhills & Coastal Plain	65p	Southeastern Floodplains and Low Terraces	276,422	13.8	7.6

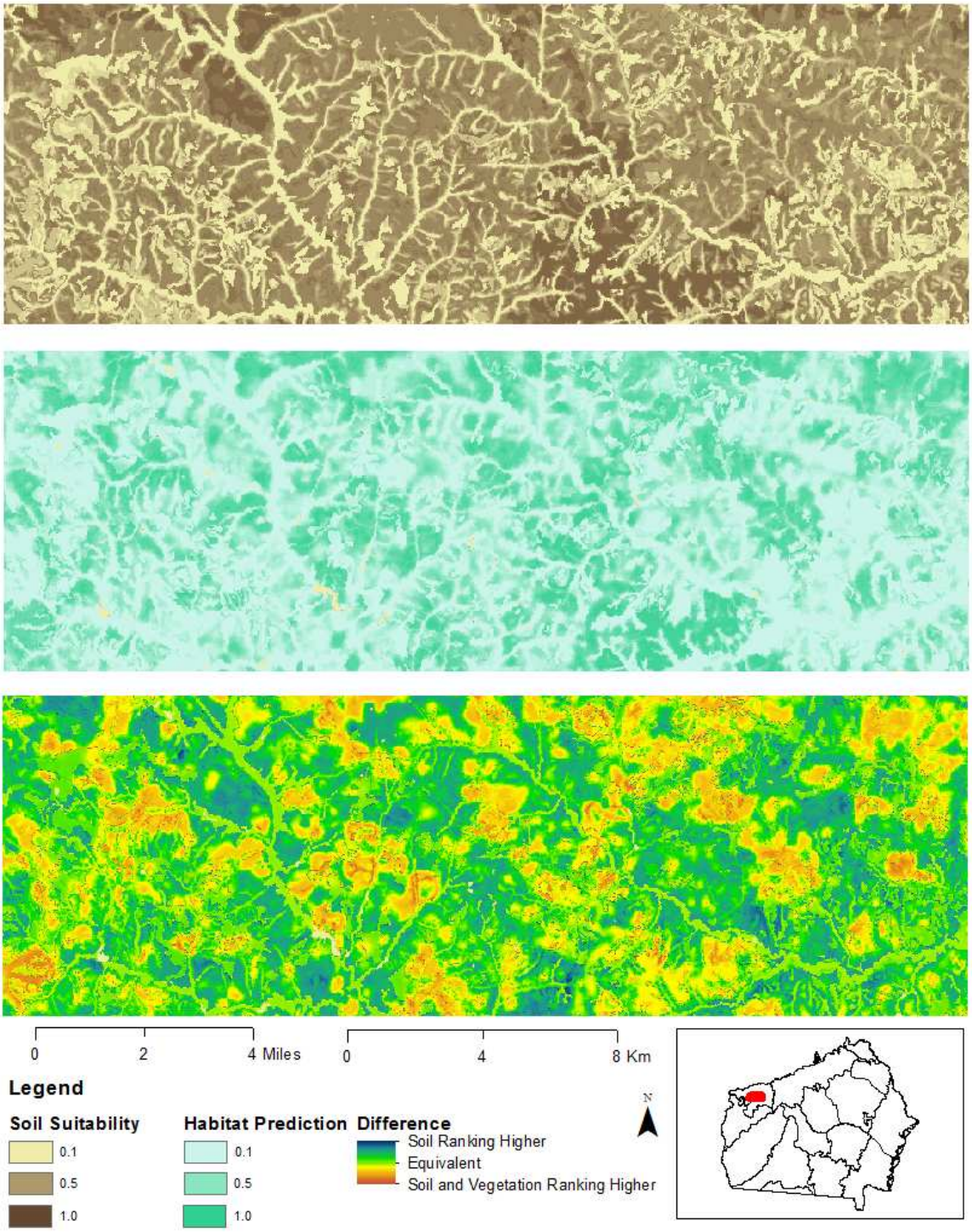

Figure 4.1.: Details of the same area (outlined in red, in inset map), showing: **top panel**, the model-averaged predictive map using only soil-derived predictor variables (Chapter 2; Figure 3.6); **middle panel**, the model-averaged prediction using predictors derived from both soil and vegetation characteristics (Chapter 3; Figure 2.6); and **bottom panel**, the difference between them (bottom panel).

5 — Database description

5.1. Introduction

Spatial data obtained from line transect distance surveys (LTDS) of gopher tortoise populations in Georgia consist primarily of:

1. burrow locations and occupancy status;
2. transect line locations;
3. the "sampling frame", or the set of polygons within which transects were laid out; and
4. the property boundary.

Georgia DNR stores these data mainly as shapefiles, so that each piece of information belonging to each survey resides in a separate file with an independent attribute table.

To facilitate our group's modeling work we consolidated these separate, independent files into a single, spatially-enabled relational database. We consider this database (the architecture and the data it holds) a useful product of this project, and so we describe the design here.

Holding disparate pieces of data in a relational database rather than separate files presents three main advantages:

1. centralization – all the data are held in the same unit, so the subunits ("tables" in the parlance of relational databases) cannot be misplaced or duplicated;
2. efficiency – writing and executing queries within and across data types –via table joins– is straightforward;
3. integrity –
   - data are difficult to accidentally alter or delete
   - values within columns may be strictly standardized and controlled (e.g., to guarantee uniqueness), via column constraints
   - relationships across columns may be controlled, via table constraints
   - relationships across tables may be constrained, via foreign keys

Foreign keys are a mechanism to ensure *referential integrity* across tables. A *child* table's foreign key consists of one or more columns, that are constrained to contain only values that appear in a corresponding column (or columns) of a *parent* table. In turn, the column(s) in the parent table must be subject to a unique constraint. Foreign keys are a fundamental component of a relational database, and they are wholly lacking from any system of data storage that relies on distinct files.

The database was built using PostgreSQL (aka "Postgres"; PostgreSQL 2017), an open source database management system distributed free under an open source license (similar to BSD or MIT licenses). To accomplish the spatial processing necessary to produce data usable for

statistical modeling, we made use of the PostGIS (PostGIS, 2017) spatial extension. PostGIS provides all the functionality of a GIS, and is reasonably fast for vector operations.

Postgres employs the notion of a *schema*, as a container for database objects. Schemas are the only level of hierarchy available, for use in organizing objects such as tables, views and functions. Table names may be prepended with the name of the schema that contains the object; and in this section, all database object names will be printed in tele-type: `schema.table`. Column names will be enclosed in square brackets: `[column]`.

5.2. Database Design

Our initial assessment of the components of the LTDS surveys suggested the following hierarchy of the spatial elements, with each level containing all the elements listed below it (GIS data type in parentheses):

1. property boundary (polygon),
2. sample frame (polygon),
3. transects (line),
4. burrows (point).

Initial versions of the database reflected this hierarchy in the foreign key structure. However, we soon found this design to be inadequate, because some properties had been re-surveyed using *different* property boundaries, and/or sample frames, and/or transects. A particular set of burrows, then, could no longer be uniquely described by stating that it "belongs to Tract A," for example. If the original survey of Tract A was done in 2014, and the new one in 2016 with a different sample frame and transects, say, then the new burrow set could only be uniquely identified as "belonging to the 2016 survey, which was done on the 2016 version of the transects in the 2016 version of the sample frame, in Tract A." We revised our foreign key structure to reflect a scheme in which burrows belong to a survey, which is performed using a set of transects, which reside in a sample frame within a property boundary, within an identified property. Surveys performed at the same property could use the same set of transects, or different sets, so they appear below transects in the hierarchy. Surveys generate burrow locations, however, so burrows are in the lowest position (Table 5.1).

Table 5.1.: The tables used to store data gathered during LTDS surveys. The primary key is shown for each; the key is composite in all cases but the `survey.sites` table at the top of the hierarchy. The foreign key for each table points to the primary key in the table above it.

Schema	Table	Primary key
survey	sites	[site]
bndry	bndries	[site],[bndry]
samp_fr	samp_frms	[site],[bndry],[frame]
trans	trans_diss	[site],[bndry],[frame],[trans]
survey	surveys	[site],[bndry],[frame],[trans],[visit]
burrows	all_burrows	[site],[bndry],[frame],[trans],[visit],[entry_order]

The database relies upon *natural keys*, meaning that values appearing in the key column(s) act as the key values. The one exception is the burrow observation table `burrows.all_burrows`, which uses the serial `[entry_order]` column, representing an arbitrary or *synthetic* key.

All tables but `survey.sites` use *composite* primary keys, meaning the key is composed of multiple columns. Each table contains a column `[key_sum]` which acts as a readable summary of the primary key for a given row. This `[key_sum]` column is created automatically by a trigger function, when a new row is added to the table (or the table is updated); it is purely a convenience.

PostgreSQL uses the term *view* to refer to a saved query. Views are named, and treated for most purposes exactly like tables. A number of views appear in the schema "burrows", to filter burrow observations in preparation for their use in distance analysis (Table 5.2).

Table 5.2.: Views in schema `burrows`, to progressively filter the set of burrow records, to generate a set appropriate for distance analysis.

View	Purpose
`dist_calc`	Calculate the distance from each burrow, to the nearest transect (uses PostGIS spatial operations).
`dist_ok`	Filter out burrows that had `[invalid] = 'Y'` in the `burrows.all_burrows` table; OR were off the transect and NOT within 15 map units (meters) of the end of the transect. In other words, all valid burrows that are not too far off the end of a transect are retained.
`dist_ok_occu`	Retain only those burrows in `burrows.dist_ok` that were occupied. This table serves as the basis for distance analysis, to estimate tortoise density and abundance.
`dist_ok_occu_soil`	Add soil suitability index and ranking to records in `burrows.dist_ok_occu`.

Information on Georgia soils, from NRCS's Soil Survey Geographic (SSURGO) database, is held in the `soil` schema, including a polygonized version of the gridded SSURGO (gSSURGO) raster for the Georgia coastal plain. This allows direct spatial queries to extract soil properties in relation to LTDS objects such as sample frames, transects and burrows.

5.3. Function to support habitat modeling

Placement of pseudo-absences for use in the habitat modeling described in Chapters 2 and 3 was carried out by a function within the database. Spatial filtering of burrows and pseudo-absences was also carried out by this function, named `habmod.pseudo_abs`.

The following query was used to generate the pseudo-absences used in all habitat modeling in this report. It calls `habmod.pseudo_abs`, placing pseudo-absences within polygons held in the table `habmod.buffers`. Areas beyond the boundaries of the MLRA polygons held in `soil.ga_mlra_diss` are excluded. The ratio of accepted pseudo-absences to burrows is 3:1. The resulting points are saved in the table `habmod.pts` (this target table is hard-wired into the function `habmod.pseudo_abs`). The minimum distance between any point in the output is set at 150 meters.

```
WITH un_one AS
    (SELECT ST_Union(geom) AS geom FROM habmod.buffers),
  un_two AS
    (SELECT ST_Union(geom) AS geom FROM soil.ga_mlra_diss),
```

```
  intrsct AS
    (SELECT ST_Intersection(A.geom, B.geom) AS geom FROM un_one A, un_two B)
SELECT habmod.pseudo_abs(geom, 3, 150)
  FROM intrsct;
```

6 — Products and Synergistic Activities

6.1. Products directly and indirectly derived from this work

6.1.1. Presentations directly related to this project

Nuse, B. L., N. Jafari, R. L. Bormann, A. D. Wright, C. T. Moore, J. Hepinstall-Cymerman, and M. Elliott. 2016. Accommodating space and demographic uncertainty, in pursuit of gopher tortoise conservation in Georgia. Invited seminar. Warnell School of Forestry and Natural Resources, University of Georgia. 27 October 2016, Athens, GA.

Bormann, R., J. Hepinstall-Cymerman, B. Nuse, C. Moore, and M. Elliott. 2015. (Poster) Modeling gopher tortoise (*Gopherus polyphemus*) distribution in Georgia. 37th Annual Gopher Tortoise Council Meeting, 16-18 October 2015, Covington, LA.

Bormann, R., J. Hepinstall-Cymerman, Clinton T. Moore, and Matt Elliott. 2015. Modeling gopher tortoise habitat in Georgia and analyzing the role of private landowners in tortoise conservation. 9th World Congress of the International Association for Landscape Ecology. 5-10 July 2015, Portland, OR.

Bormann, R., J. Hepinstall-Cymerman, L. German, J. Rice, C. Moore, and M. Elliott. 2014. (Poster) Modeling gopher tortoise habitat and habitat connectivity in Georgia, and analyzing the role of private landowners in their conservation. 36th Annual Gopher Tortoise Council Meeting, 17-19 October 2014, Albany, GA.

6.1.2. Products related to gopher tortoise projects funded by other entities

Manuscripts

Jafari, N., B. L. Nuse, C. T. Moore, B. Dilkina, and J. Hepinstall-Cymerman. 2017. Achieving full connectivity of sites in the Multiperiod Reserve Network Design Problem. Computers and Operations Research 81:119-127. http://dx.doi.org/10.1016/j.cor.2016.12.017

Reports

Hepinstall-Cymerman, J., R. L. Bormann, T. J. Prebyl, and B. L. Nuse. 2017. NCASI Final Report: Validating Gopher Tortoise Habitat and Population Predictions for Georgia.

Presentations and Posters

Gaya, H., L. L. Smith, and C. T. Moore. 2017. Improving line-transect distance sampling (LTDS) for gopher tortoise (*Gopherus polyphemus*) populations. 39th Annual Gopher Tortoise Council Meeting, 13-15 October 2017, Edgefield, SC.

Prebyl, T.J., L. Smith, J. Martin, J. Hepinstall-Cymerman, and C. Moore. 2017. Working gophers: Evaluating conservation efforts for gopher tortoise on private lands. 39th Annual Gopher Tortoise Council Meeting, 13-15 October 2017, Edgefield, SC.

Hepinstall-Cymerman, J., T. Prebyl, B. Nuse, and C. Moore. 2017. Predicting gopher tortoise habitat connectivity in Georgia from forest structure, soils, and land cover. The Wildlife Society 24th Annual Conference, 23-27 September 2017, Albuquerque, NM.

Gaya, H., L. L. Smith, and C. Moore. 2017. (Poster) Improving line-transect distance sampling (LTDS) for gopher tortoise (*Gopherus polyphemus*) populations. The Wildlife Society 24th Annual Conference, 23-27 September 2017, Albuquerque, NM.

Prebyl, T. J., L. Smith, J. Martin, J. Hepinstall-Cymerman, and C. Moore. 2017. (Poster) Evaluating private lands conservation practices for gopher tortoise and savanna-like ecosystems in the southeastern coastal plain. The Wildlife Society 24th Annual Conference, 23-27 September 2017, Albuquerque, NM.

Hepinstall-Cymerman, J., T. Prebyl, B. Nuse, C. Moore, and R. Bormann. 2017. Modeling gopher tortoise habitat and habitat connectivity in Georgia using multi-season Landsat imagery and Circuitscape. US-IALE 2017 Annual Meeting, 9-13 April 2017, Baltimore, MD.

Nuse, B. L., C. T. Moore, N. Jafari, J. Hepinstall-Cymerman, and M. Elliot. 2017. Optimal reserves for spatially structured gopher tortoise populations, under forecast landscape conditions in Georgia, USA. US-IALE 2017 Annual Meeting, 9-13 April 2017, Baltimore, MD.

Hepinstall-Cymerman, J., R. Bormann, B. Nuse, C. Moore, and M. Elliott. 2016. Mapping gopher tortoise habitat in Georgia using multi-season Landsat imagery and multiple statistical modeling techniques. The Wildlife Society 23rd Annual Conference, 15-19 October 2016, Raleigh, NC.

Nuse, B. L., C. T. Moore, N. Jafari, J. Hepinstall-Cymerman, and M. Elliott. 2016. Informing reserve design with demography for gopher tortoise conservation planning in Georgia, USA. The Wildlife Society 23rd Annual Conference, 15-19 October 2016, Raleigh, NC.

Wright, A. D., J. Hepinstall-Cymerman, L. L. Smith, C. T. Moore, and R. B. Chandler. 2016. Long-term population ecology and large-scale movement patterns of gopher tortoises in southwestern Georgia: A spatial capture-recapture approach. The Wildlife Society 23rd Annual Conference, 15-19 October 2016, Raleigh, NC.

Hepinstall-Cymerman, J. 2016. Geospatial data and techniques for understanding the resource selection, landscape ecology, and movement ecology of terrestrial vertebrates. NCSU Center for Geospatial Analytics Seminar Series, Invited Seminar, Raleigh, NC.

Nuse, B. L., C. T. Moore, and J. Hepinstall-Cymerman. 2016. Predicting spatial structure in gopher tortoise (*Gopherus polyphemus*) populations in Georgia, USA. US-IALE 2016 Annual Meeting, 3-7 April 2016, Asheville, NC.

Wright, A. D., J. Hepinstall-Cymerman, L. L. Smith, and C. T. Moore. 2016. Long-term population ecology and large-scale movement patterns of gopher tortoises (*Gopherus polyphemus*) in southwestern Georgia. US-IALE 2016 Annual Meeting, 3-7 April 2016, Asheville, NC.

Moore, C. T., J. Hepinstall-Cymerman, B. L. Nuse, R. Bormann, A. D. Wright, and N. Jafari. 2016. UGA Moore-Hepinstall gopher tortoise research: projects update. Presentation to U.S. Fish and Wildlife Service, Southeast Climate Science Center, and state agency partners, 12 February 2016, Athens, GA.

Wright, A. D., J. Hepinstall-Cymerman, L. L. Smith, and C. T. Moore. 2016. (Poster) Long-term population ecology and large-scale movement patterns of gopher tortoises (*Gopherus polyphemus*) in southwestern Georgia. Annual Scientific Advisory Committee Meeting of the Joseph W. Jones Ecological Research Center, 4 February 2016, Newton, GA.

Nuse, B. L., J. Hepinstall-Cymerman, C. T. Moore, and M. Elliott. 2015. Linking gopher tortoise processes and landscape resistance to assess functional connectivity of gopher tortoise populations in Georgia. Annual Meeting of the Southeastern Association of Fish and Wildlife Agencies, 1-4 Nov 2015, Asheville, NC.

Wright, A. D., J. Hepinstall-Cymerman, L. L. Smith, and C. T. Moore. 2015. Long-term population ecology and large-scale movement patterns of gopher tortoises (*Gopherus polyphemus*) in southwestern Georgia. Annual Meeting of the Southeastern Association of Fish and Wildlife Agencies, 1-4 Nov 2015, Asheville, NC.

Jafari, N., and C. Moore. 2015. Land development uncertainties in the dynamic reserve network design problem. INFORMS Annual Meeting, 1-4 Nov 2015, Philadelphia, PA.

Wright, A. D., J. Hepinstall-Cymerman, L. L. Smith, and C. T. Moore. 2015. Long-term population ecology and large-scale movement patterns of gopher tortoises (*Gopherus polyphemus*) in southwestern Georgia. 37th Annual Gopher Tortoise Council Meeting, 16-18 October 2015, Covington, LA.

Wright, A. D., J. Hepinstall-Cymerman, L. Smith, and C. T. Moore. 2015. (Poster) Data are scarce but action is necessary: Using agent-based models for conservation. Annual Student Conference on Conservation Science, 7-9 October 2015, New York, NY.

Nuse, B. L., C. T. Moore, J. Hepinstall-Cymerman, and M. Elliott. 2015. (Poster) Comparing hypotheses about missed burrows in gopher tortoise line transect surveys: Bayesian distance analysis with data augmentation. 2015 Ecological Society of America Annual Meeting, 9-14 August 2015, Baltimore, MD.

Jafari, N., and C. T. Moore. 2015. Achieving the spatial connectivity of parcels in the dynamic reserve network design problem. International Symposium on Mathematical Programming, 12-17 July 2015, Pittsburgh, PA.

Nuse, B. L., J. Hepinstall-Cymerman, C. T. Moore, and M. Elliott. 2015. Assessing functional connectivity in the face of uncertainty about population processes: a Bayesian modeling approach applied to conservation of the gopher tortoise in Georgia, USA. International Association of Landscape Ecology World Congress, 5-10 July 2015, Portland, OR.

Wright, A. D., J. Hepinstall-Cymerman, L. L. Smith, and C. T. Moore. 2015. Estimating dispersal rates and landscape resistance to understand functional connectivity for a long-lived endangered species. 9th World Congress of the International Association for Landscape Ecology. 5-10 July 2015, Portland, OR.

Moore, C. 2015. Adaptive landscape planning and decision framework for gopher tortoise conservation: Project update. Meeting of the Gopher Tortoise Candidate Conservation Agreement Team, 10 June 2015, Spanish Fort, AL.

Wright, A. D., J. Hepinstall-Cymerman, L. L. Smith, and C. T. Moore. 2015. Estimating dispersal rates and landscape resistance to understand functional connectivity for a long-lived endangered species. Project meeting with Southeast Climate Science Center, 19-21 May 2015, Raleigh, NC.

Moore, C. T. 2015. Reserve planning for conservation of the gopher tortoise in Georgia. Presentation to the Southeast Regional Partnership for Planning and Sustainability, 4 March 2015, Atlanta, GA.

Wright, A. D., J. Hepinstall-Cymerman, L. L. Smith, and C. T. Moore. 2015. Long-term population ecology and large-scale movement patterns of gopher tortoises (*Gopherus polyphemus*) in southwestern Georgia. Warnell Graduate Student Symposium, 18-20 February 2015, Athens, GA.

Jafari, N., C. T. Moore, and J. Hepinstall-Cymerman. 2015. Solution alternatives to achieve parcel connectivity in the dynamic reserve design problem. Twenty-Ninth Conference of the Association for the Advancement of Artificial Intelligence (AAAI-15), Workshop on Computational Sustainability. 25 January 2015, Austin TX.

Moore, C. T. 2015. Sustainability in the eyes of a decision analyst: three case studies in conservation. 2nd Annual Sustainability Science Symposium, Center for Integrative Conservation Research, University of Georgia, 23 January 2015, Athens, GA.

Wright, A. D., J. Hepinstall-Cymerman, L. L. Smith, and C. T. Moore. 2015. (Poster) What to do when data is scarce and conservation action is necessary: Using spatially-explicit individual-based models to understand the ecology of threatened/endangered species. 2nd Annual Sustainability Science Symposium, Center for Integrative Conservation Research, University of Georgia, 23 January 2015, Athens, GA.

Jafari, N., and C. T. Moore. 2014. Designing a dynamic reserve problem under stochastic uncertainties. INFORMS Annual Meeting, 9-12 November 2014, San Francisco, CA.

Nuse, B. L., and C. T. Moore. 2014. (Poster) A Bayesian population model to inform gopher tortoise conservation planning in Georgia. 36th Annual Gopher Tortoise Council Meeting, 17-19 October 2014, Albany, GA.

Wright, A. D., J. Hepinstall-Cymerman, L. L. Smith, and C. T. Moore. 2014. (Poster) Long-term population ecology and large-scale movement patterns of gopher tortoises (*Gopherus polyphemus*) in southwest Georgia. 36th Annual Gopher Tortoise Council Meeting, 17-19 October 2014, Albany, GA.

Moore, C. 2014. Adaptive landscape planning and decision framework for gopher tortoise conservation. Meeting of the Gopher Tortoise Candidate Conservation Agreement Team, 9 July 2014, Savannah, GA.

Moore, C. 2014. An adaptive landscape planning and decision framework for gopher tortoise conservation: Overview, progress, and reserve design needs. Meeting to consider research needs

in dynamic reserve design, 24 April 2014, Gainesville, FL.

Moore, C. T. 2014. Conservation through a decision analytic lens. Invited seminar, Clemson University, 6 March 2016, Clemson, SC.

Moore, C., and J. Hepinstall-Cymerman. 2013. (Poster) A landscape planning and decision framework for gopher tortoise conservation in Georgia. 35th Annual Gopher Tortoise Council Meeting, 11-13 October 2013, Ponte Vedra Beach, FL.

Technical Assistance and Professional Service

2013–present	Members of our lab are serving as invited experts on an interagency technical working group to define population benchmarks and range-wide targets for gopher tortoise conservation and recovery efforts in the non-listed portion of the range.
2015	Performed an analysis to support the Department of Defense Gopher Tortoise Conservation and Crediting Strategy, of tortoise densities in surveyed areas in Georgia, producing density estimates for the soil suitability classes used in the NRCS rating system (USFWS and NRCS, 2012).
2015	Participated in a NRCS-sponsored workshop to offer input to the soils suitability ranking system.
March 2017	Conducted a meeting with Georgia Department of Natural Resources to offer input on the development of gopher tortoise conservation planning regions within the state.
2017	Working with researchers in the Department of Geography at the University of Georgia to inform their efforts to map areas of conflict between gopher tortoise habitat and siting suitability for solar arrays.

7 — Related Ongoing Work

We have other ongoing projects that will benefit from the products developed through this work. We are developing population predictions using gopher tortoise line transect distance survey data from GADNR, and the habitat maps described in this report. The habitat maps and population predictions will be combined with parcel boundaries to assist GADNR with prioritizing areas to target conservation activities for gopher tortoise in the future. The results of this work will also be informative in an effort being conducted by the Georgia Cooperative Fish and Wildlife Research Unit to construct distribution and abundance models for several herpetofauna of concern, including the gopher tortoise, in the southeastern US. We also expect these results to be informative to a project now getting underway that evaluates effects of private landowner incentive programs for gopher tortoise conservation. These projects are funded through several different grants from multiple sources.

8 — Acknowledgments

We used imagery-based vegetation predictors based on modeling work conducted by Thomas Prebyl (Appendix C) of the Warnell School of Forestry and Natural Resources, University of Georgia. We thank Matt Elliott of Georgia Department of Natural Resources for access to data and his encyclopedic knowledge of tortoise populations in the state. We acknowledge Lora Smith of the Joseph W. Jones Ecological Research Center for her advice and insight into gopher tortoise biology and habitat ecology.

Funding received through this work provided material support to Ms. Rachel Bormann, a PhD student in the Warnell School of Forestry and Natural Resources, and furthered her education and professional training.

9 — References

Aresco, M. J., and C. Guyer. 1999*a*. Burrow abandonment by Gopher Tortoises in slash pine plantations of the Conecuh National Forest. Journal of Wildlife Management **63**:26–35.

Aresco, M. J., and C. Guyer. 1999*b*. Growth of the Tortoise *Gopherus polyphemus* in slash pine plantations of Southcentral Alabama. Herpetologica **55**:499–506.

Auffenberg, W., and R. Franz, 1982. The Status and Distribution of the Gopher Tortoise. Pages 95–124 *in* R. Bury, editor. North American Tortoises: Conservation and Ecology, volume 12 of *Wildlife Research Report*. US Fish and Wildlife Service, Washington DC.

Baskaran, L., V. Dale, R. Efroymson, and W. Birkhead. 2006. Habitat Modeling Within a Regional Context: An Example Using Gopher Tortoise. American Midland Naturalist **155**:335–351.

Birkhead, R., C. Guyer, S. M. Hermann, and W. K. Michener. 2005. Patterns of folivory and seed ingestion by Gopher Tortoises (*Gopherus polyphemus*) in a Southeastern pine savanna. American Midland Naturalist **154**:143–151.

Breiman, L. 2001. Random forests. Machine Learning **45**:5–32.

Breininger, D. R., P. A. Schmalzer, and C. R. Hinkle. 1994. Gopher tortoise (*Gopherus polyphemus*) densities in coastal scrub and slash pine flatwoods in Florida. Journal of Herpetology **28**:60–65.

Burnham, K. P., and D. R. Anderson. 2002. Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach. Springer.

Castellón, T. D., B. B. Rothermel, and S. Z. Nomani. 2012. Gopher tortoise (*Gopherus polyphemus*) burrow densities in scrub and flatwoods habitats of peninsular Florida. Chelonian Conservation and Biology **11**:153–161.

Cox, J., D. Inkley, and R. Kautz, 1987. Ecology and habitat protection needs of gopher tortoise (*Gopherus polyphemus*) populations found on lands slated for large-scale development in Florida. Nongame Wildlife Program Technical Report 4, Florida Game and Fresh Water Fish Commission, Tallahassee, Florida.

Diemer, J. E. 1986. The ecology and management of the gopher tortoise in the southeastern United States. Herpetologica **42**:125–133.

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

GTC, 2013. Gopher Tortoise Minimum Viable Population and Minimum Reserve Size Working Group Report. Technical report, Gopher Tortoise Council.

Hepinstall-Cymerman, J., R. L. Bormann, T. J. Prebyl, and B. L. Nuse, 2017. NCASI Final Report: Validating Gopher Tortoise Habitat and Population Predictions for Georgia.

Jenness, J., 2006. Topographic Position Index extension for ArcView 3.x, v. 1.2. Jenness Enterprises. URL http://www.jennessent.com/arcview/tpi.htm.

Jones, J. C., and B. Dorr. 2004. Habitat associations of Gopher Tortoise burrows on industrial timberlands. Wildlife Society Bulletin **32**:456–464.

Kaczor, S. A., and D. C. Hartnett. 1990. Gopher tortoise (*Gopherus polyphemus*) effects on soils and vegetation in a Florida sandhill community. American Midland Naturalist **123**:100–111.

Kent, M. 2011. Vegetation Description and Data Analysis: A Practical Approach. Second edition. Wiley.

Kowal, V. A., A. Schmolke, R. Kanagaraj, and D. Bruggeman. 2014. Resource selection probability functions for gopher tortoise: providing a management tool applicable across the species' range. Environmental Management **53**:594–605.

LANDFIRE, 2013. LANDFIRE Existing Vegetation Type layer. Technical report, U.S. Geological Survey. URL http://landfire.cr.usgs.gov/viewer/.

Liaw, A., and M. Wiener. 2002. Classification and regression by randomForest. R News **2**:18–22.

Lohoefener, R., and L. Lohmeier. 1981. Comparison of gopher tortoise (*Gopherus polyphemus*) habitats in young slash pine and old longleaf pine areas of southern Mississippi. Journal of Herpetology **15**:239–242.

MacDonald, L. A., and H. R. Mushinsky. 1988. Foraging ecology of the gopher tortoise, *Gopherus polyphemus*, in a sandhill habitat. Herpetologica **44**:345–353.

Masek, J. G., E. F. Vermote, N. E. Saleous, R. Wolfe, F. G. Hall, K. F. Huemmrich, F. Gao, J. Kutler, and L. Teng-Kui. 2006. A Landsat surface reflectance dataset for North America, 1990-2000. IEEE Geoscience and Remote Sensing Letters **3**:68–72.

McCoy, E. D., H. R. Mushinsky, and J. Lindzey. 2006. Declines of the gopher tortoise on protected lands. Biological Conservation **128**:120–127.

McRae, W. 1981. Movement patterns and home range of the gopher tortoise. American Midland Naturalist **106**:165–179.

Mushinsky, H. R., 2014. Growth patterns of North American tortoises, Chapter 7, pages 53–59 . Johns Hopkins University Press.

Nomani, S. Z., R. R. Carthy, and M. K. Oli. 2008. Comparison of methods for estimating abundance of gopher tortoises. Applied Herpetology **5**:13–31.

NRCS, 2005. Major Land Resource Area (MLRA). URL https://www.nrcs.usda.gov/wps/portal/nrcs/detail/soils/survey/geo/?cid=nrcs142p2_053625.

NRCS, 2013. The Gridded Soil Survey Geographic (gSSURGO) Database for Georgia. U.S. Department of Agriculture, Natural Resources Conservation Service.

NRCS, 2017*a*. Gopher Tortoise Habitat Suitability Rating (v6.6). Unpublished report, National Resources Conservation Service.

NRCS, 2017*b*. National Cooperative Soil Characterization Database. Online. URL https://www.nrcs.usda.gov/wps/portal/nrcs/detail/soils/survey/tools/?cid=nrcs142p2_053552.

Nussear, K. E., and T. D. Tuberville, 2014. Habitat characteristics of North American tortoises, Chapter 9, pages 77–84 . Johns Hopkins University Press.

Omernik, J., and G. Griffith. 2014. Ecoregions of the conterminous United States: evolution of a hierarchical spatial framework. Environmental Management **54**:1249–1266.

PostGIS, 2017. PostGIS 2.4.2dev Manual. PostGIS Development Group. URL http://postgis.net/docs/manual-2.4/.

PostgreSQL, 2017. PostgreSQL 9.6.5 Documentation. PostgreSQL Global Development Group. URL https://www.postgresql.org/files/documentation/pdf/9.6/postgresql-9.6-US.pdf.

PRISM, 2012. United States Average Monthly or Annual Precipitation, 1981-2010. Technical report, PRISM Climate Group. URL http://prism.oregonstate.edu.

R Core Team, 2017. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.

Sing, T., O. Sander, N. Beerenwinkel, and T. Lengauer. 2005. ROCR: visualizing classifier performance in R. Bioinformatics **21**:78–81. URL http://rocr.bioinf.mpi-sb.mpg.de.

Smith, L. L., J. M. Linehan, J. M. Stober, M. J. Elliott, and J. B. Jensen. 2009. An evaluation of distance sampling for large-scale gopher tortoise surveys in Georgia, USA. Applied Herpetology **6**:355–368.

Stober, J. M., and L. L. Smith. 2010. Total counts versus line transects for estimating abundance of small gopher tortoise populations. Journal of Wildlife Management **74**:1595–1600.

Thompson, J., T. Nauman, N. Odgers, Z. Libohova, and J. Hempel, 2012. Harmonization of legacy soil maps in North America: Status, trends and implications for digital soil mapping efforts. Pages 92–102 *in* B. Minasny, B. P. Malone, and A. B. McBratney, editors. Digital Soil Assessments and Beyond: Proceedings of the 5th Global Workshop on Digital Soil Mapping, Sydney, Australia. CRC Press.

USFWS, 1990. Gopher Tortoise Recovery Plan. Technical Report 28, US Fish and Wildlife Service, Jackson, Mississippi.

USFWS, and NRCS, 2012. Gopher tortoise (*Gopherus polyphemus*) soil classifications for the federally listed range using the National Soil Information System Database. Unpublished report, U.S. Fish and Wildlife Service and Natural Resources Conservation Service.

USGS, 2016. Landsat Spectral Indices Product Guide Version 3.3. US Geological Survey. URL https://landsat.usgs.gov/sites/default/files/documents/si_product_guide.pdf.

Wigley, T. B., C. W. Hedman, C. Loehle, M. Register, J. R. Poirier, and P. E. Durfield. 2012. Density of gopher tortoise burrows on commercial forestland in Alabama and Mississippi. Southern Journal of Applied Forestry **36**:38–43.

Yackulic, C. B., R. Chandler, E. F. Zipkin, J. A. Royle, J. D. Nichols, E. H. C. Grant, and S. Veran. 2013. Presence-only modelling using MAXENT: when can we trust the inferences? Methods in Ecology and Evolution **4**:236–243.

Yager, L., M. Hinderliter, C. Heise, and D. Epperson. 2006. Gopher Tortoise response to habitat management by prescribed burning. Journal of Wildlife Management **71**:428–443.

Appendix A — Vegetation surveys to support development of phenology-based vegetation attributes

This appendix describes vegetation surveys conducted to support the development of a predictive model of vegetation attributes derived from remotely-sensed data (described in Appendix B). Collection of these data were supported under a grant from the National Council for Air and Stream Improvement. Full details on sample site selection, stratification, and vegetation monitoring are provided in Hepinstall-Cymerman et al. (2017).

To develop our vegetation field methods, we drew on published sources (Appendix B) as well as discussions with tortoise ecology experts Matt Elliott (GADNR) and Lora Smith (Joseph Jones Ecological Research Center; JJERC), and members of each of their gopher tortoise survey crews. A literature review of gopher tortoise habitat vegetation and diet was used to determine the appropriate level of detail to gather in assessing ground vegetation < 1 m in height.

A.1. Site selection for sampling protocol development

In late summer-early fall 2015, field sites in Georgia were visited to conduct vegetation sampling: an industrial timber property, GADNR-owned wildlife management areas, and property owned by The Nature Conservancy (Table A.2). These sites were chosen because gopher tortoise surveys were actively underway; or because access could be otherwise arranged. In all 42 vegetation plots were surveyed with 5 subplots at each location. Plots were sited based on access and proximity to the tortoise survey crew.

A.2. Within-site stratification

Based on the results of the fall 2015 survey work, we developed a stratification approach derived from remotely sensed imagery to ensure field plots sampled in spring 2016 captured the range of vegetation and spectral characteristics within each potential sampling property. We used the Landsat 8 Enhanced Vegetation Index (EVI) product to identify seasonal spectral characteristics of vegetation at study properties (Masek et al., 2006; USGS, 2016). We acquired all available Landsat 8 EVI products from Feb 2013 to Mar 2016 with less than 40% cloud cover and used the supplied CFmask band to remove all pixels that were not classified as clear (free of clouds and cloud shadows). We next created seasonal maximum greeness rasters by calculating the maximum cloud-free pixel EVI value across all available rasters for the winter season (Julian days ≥ 340 or ≤ 85) and summer season (Julian days ≥ 130 and ≤ 250). Additionally we created a deciduous index raster by subtracting the winter maximum greeness from the summer maximum greeness and dividing by the maximum annual greeness:

$$\text{deciduous index} = \frac{\text{summer maximum greeness} - \text{winter maximum greeness}}{\max(\text{summer maximum greeness, winter maximum greeness})}$$

Within each property boundary to be sampled in spring 2016, we created a 60 x 60 m grid of potential sampling points in our GIS. Points were excluded from land cover classes which comprised less than 0.5% of all known active or occupied burrow locations using the LANDFIRE 2013 Existing Vegetation Type layer (LANDFIRE, 2013). Roads were also excluded from potential plot locations to avoid including spectral signatures of roads in later modeling efforts despite the road land cover class containing 4.4% of active or occupied burrows.

At the remaining grid points within each property we extracted raster values from the two seasonal greenness layers and deciduous index described above. We then performed k-means clustering in program R (R Core Team, 2017) to classify each point to one of eight clusters based on extracted spectral data. We chose eight for the number of clusters as we felt it adequately captured the variability in landscape conditions, while still allowing a practical amount of time to sample from sites in each cluster.

For each property to be visited for vegetation sampling in spring 2016, points were opportunistically selected from each of the eight clusters, though an effort was made to sample from across the range of the property. We also attempted to select plots in locations where points of a given cluster type were clumped spatially as we believed these locations to be more representative of the k-means cluster than single isolated points. Specifically, in the field we attempted to select a location around or near known tortoise populations, if that was known, or where GADNR was planning on surveying in the future (expected habitat). Within one location (walking distance or very short driving distance for a day or half-day), we attempted to opportunistically select from each of the 8 clusters. We moved to a new part of the property each day and selected sites again using the same criteria.

A.3. Vegetation sampling methods

We selected sampling plots at the sites with a goal of obtaining a range of stand overstory and understory conditions as well as a 4:1 ratio of points inside likely habitat to outside habitat. In the areas where tortoise surveys were to be conducted, likely habitat was overlaid with a 25- to 100-m^2 fishnet and a 100- to 300-m buffer and 10% of the polygons were chosen randomly as potential survey plots.

At each polygon, we randomly selected a center point, taking five pictures at the center point (one in each cardinal direction, and one vertically up), and measured the following items:

1. GPS location using a Flint S Series GPS unit.

2. Basal area (from the center point) using a 10-factor wedge prism counting "in" trees and 1/2 for every "borderline" tree $\geq$ 10 cm diameter.

3. Canopy cover (from the center point), by taking four densiometer readings, one in each cardinal direction, using a Model-A Forest Spherical Densiometer.

4. Vertical structure, by dividing the vegetation within the plot into the following vertical strata (if distinctly present): herbaceous (0–0.5 m), shrub, lower understory, upper understory, and overstory. We measured the maximum height of the shrub, understory, and overstory layers for a 15-m radius around the center point using either a Nikon Forestry 550 Laser Rangefinder/Hypsometer or a Suunto PM-5/360 PC Clinometer.

5. Dominant shrub, understory, and overstory type: pine (P) ($\geq$ 75% pine), hardwood (H) ($\geq$ 75% hardwood), mixed (M) ($<$ 75% pine and $<$ 75% hardwood), other (O) (not pine or hardwood), or none (N). Other type was recorded in notes.

6. If the dominant overstory was pine, we recorded the pine species.

We also gathered the following data at five 5-m radius subplots, with their centers located: at the center point of the plot, and ∼7.5 m in each cardinal direction from the plot center. Pictures were taken of the ground in each subplot, from a height of approximately 1.5 m.

1. In a 1x1-m quadrat at the center of the subplot, we estimated percent cover in each of two height layers: (0–2 cm and 2–100 cm), estimated by eye in 5% intervals for the following categories:
   a) Percent bare ground
   b) Litter
   c) Forbs
   d) Grasses, sedges
   e) Other greens (items possibly edible by tortoises) including woody vine leaves, shrub leaves, cacti (prickly pear), ferns, palms, tree seedlings
   f) Other including stems, lichen, woody debris
2. Percent cover estimated by eye in 10% intervals for each of the vertical strata (ground, 0–0.5 m; shrub; lower understory; upper understory; overstory) for a 5m radius around each point.

Kent (2011) (Ch.3 "The description of vegetation in the field") provided general guidance for our methods.

A.4. Results

In all, between fall 2015 and spring 2016, 188 plots (135 in known suitable habitat; 53 in presumed unsuitable habitat) were sampled for vegetation using the sampling protocol described in Appendix B. The sites selected for sampling in spring 2016 were identified through an analysis of multi-season Landsat images (described above) and stratified across levels of understory and overstory percent coverage (Table A.1). Specific site attributes are detailed in Table A.2.

Table A.1.: Number of sites sampled by understory and overstory percent cover. "Low": $< 60\%$; "Moderate": $\geq 60, < 80\%$; "High" $\geq 80\%$.

	Overstory		
Understory	Low	Moderate	High
Low	41	40	27
Moderate	11	19	22
High	2	10	10

A.5. Aggregated vegetation metrics

To bring the field-measured vegetation metrics (Appendix B) up to the same spatial grain of the 30-m pixel Landsat imagery, we aggregated measurements made within subplots up to the plot level. These summaries are described in Table A.3.

Table A.2.: Sites,with ownership/management status as of 2016. "WMA" = Wildlife Management Area, "NA" = Natural Area.

Property	Georgia County	Public access	Ownership/management
Altama WMA	Glynn	Yes	Former privately-owned hunting property, recently purchased by TNC (GADNR to purchase), GADNR managed
Black Creek NA	Taylor	Yes	GADNR owned and managed
Flat Tub WMA	Coffee	Yes	GADNR owned and managed
Flournoy	Marion/Talbot	No	Privately owned and managed
Fort Perry	Marion	No	Former privately-owned hunting property, now ownership/management shared by TNC, GADNR
Ichauway	Baker	No	Privately owned and managed
Little Satilla WMA	Wayne	Yes	Part Weyerhaeuser owned/managed, part Rayonier
Moody Forest NA	Appling	Yes	TNC/GADNR jointly owned and managed
Parkers Mill	Marion	No	TNC owned and managed
Plum Creek, Hayner Pasture	Glynn	No	Weyerhaeuser owned and managed
Plum Creek, Randolph	Randolph	No	Weyerhaeuser owned and managed
Rovig	Marion	No	TNC owned and managed
Sansavilla WMA	Wayne	Yes	Conservation Fund owned (GADNR to purchase), GADNR managed
WR Bean	Marion	No	TNC owned and managed

Table A.3.: Plot-level summaries of vegetation metrics originally measured in quadrats or subplots. In all cases, the summary is obtained by taking the mean across all sample units.

Symbol	**Name**	**Sample unit**	**Measurement**
`fbgr`	Forb + grass cover	Quadrat, 2–100 cm layer	Sum of "forb" and "grass" percent cover
`dens`	Densiometer	Subplot	Mean of readings in four cardinal directions
`other`	Other cover	Quadrat, 2–100 cm layer	Sum of the "other", "other evergreen" and "other deciduous" percent cover
`covr`	Maximum cover	Subplot	Maximum cover value across all strata
`shrb`	Shrub cover	Subplot	Percent cover value in the shrub stratum
`pine`	Pine dominance	Subplot	Pine-dominated strata assigned a value of 1, mixed 0 and hardwood -1; then a weighted average taken across strata, with the percent cover values* as the weights

*In a given stratum, cover value is for all vegetation; vegetation in that stratum is then labeled as being predominantly "pine", "mixed" or "hardwood".

Appendix B — Studies characterizing the vegetation of gopher tortoise habitat

Lohoefener and Lohmeier (1981)
Objective: determine if tortoises are affected by a sudden transition from longleaf to slash pine (∼54 ha, Desoto National Forest).
Vegetation sampling methods: Line transect, 5 transects 38 m in longleaf, 6 transects 44 m in slash. Area per species of "forage plants", apparently defined as grasses and forbs.

MacDonald and Mushinsky (1988)
Objective: determine the diet of a population of gopher tortoises, University of South Florida (USF) Ecological Research Area.
Vegetation sampling methods: 36 variable-length transects, total sampled area = 100 m^2. All live rooted plants identified and counted.

Kaczor and Hartnett (1990)
Objective: assess effect of gopher tortoise activity on micro-environment and vegetation (2.1 ha, USF Ecological Research Area). *Vegetation sampling methods:* track mortality and canopy cover per species on 27 recently abandoned mounds and 2 1x1-m plots within 5 m.

Breininger et al. (1994)
Objective: compare gopher tortoise density in scrub and pine flatwoods, examine influence of time since the last fire on gopher tortoise density (Kennedy Space Center).
Vegetation sampling methods: 112 30x50-m plots by stratified random design, 20 points 5 m apart on 4 parallel lines 5 m apart, point-intercept method, measured species of trees, shrubs, and herbs, shrub height at each point.

Aresco and Guyer (1999*a*)
Objective: determine frequency of burrow abandonment in pine plantations, if abandonments are associated with vegetation changes (6 sites Conecuh NF).
Vegetation sampling methods: 123 burrows, 60 random plots. Circular plots (5-m radius): # trees/ha, BA, dbh, tree species ID. Circular plot (2.5-m radius): # shrub stems 1–2m in height. Quadrats (5 per burrow/plot, 1x1m): species and % canopy cover of plants ≤1m. Plants were grouped into categories: tree seedlings, woody shrubs, legume forbs, non-legume forbs, grasses, woody vines.

Aresco and Guyer (1999*b*)
Objective: describe gopher tortoise growth rates and estimate age to maturity in habitats heavily impacted by forestry operations, compare growth patterns of gopher tortoises to those published in other studies, compare forage quality data to previous publications.
Vegetation sampling methods: same as Aresco and Guyer (1999*a*) above.

Jones and Dorr (2004)
Objective: report burrow densities, determine primary habitat conditions for burrows on corporate timberland (11 counties in AL and MS).
Vegetation sampling methods: 10x20-m quadrat at transect midpoint or at each burrow on transect, 2759 20x250-m transects. Verified dominant pine type. Measured % cover of plants in 3 categories (>6m, 1–6m, ≤1m) by densiometer (>6m) and by eye at 10 1x1-m plots. Derived total canopy cover from upper and mid story cover.

Birkhead et al. (2005)
Objective: document folivory and frugivory by gopher tortoises (Jones Center).
Vegetation sampling methods: four 1x1-m quadrats surrounding 16 burrows. All plants within the frame identified.

Yager et al. (2006)
Objective: determine if fire would restore habitat suitable for gopher tortoises, the response of gopher tortoises to habitat restoration (8 sites, Camp Shelby, MS).
Vegetation sampling methods: transects 75 m apart, sample points every 25 m along transects, at each sample point, 1x1-m quadrat. Measured % cover understory veg <3.5 m by growth by growth form categories (total herbaceous, woody vines and shrubs, graminoids-grass and grass likes, forbs, legumes) and species using Peet's modified Daubenmire method. Counted woody stems and classified by diameter width. Measured overstory cover (densiometer). Determined overall forest type.

McCoy et al. (2006)
Objective: compare size and structure of gopher tortoise populations from late 1980's to early 2000's (10 sites, ∼2000 ha, central/northern FL).
Vegetation sampling methods: 7-m wide transects covering ∼10% of area. 1x1-m quadrats every 50 m along every other transect. Measured % cover by grasses, live herbaceous, woody/dead, bare and presence/absence of 1–3 m and >3 m canopy cover.

Wigley et al. (2012)
Objective: locate burrows, relate burrow density to forest structural characteristics.
Vegetation sampling methods: % cover and plant species or functional group for overstory (>10.2 cm dbh), midstory (<10.2 dbh, >1.4m tall), and understory (<1.4m tall). Previously collected data available included BA (pine and hardwood), dbh, tree height, stems per ha.

Kowal et al. (2014)
Note: vegetation data not used directly in this publication.
Objective: used vegetation data from previous vegetation inventories to validate the use of NDVI to describe vegetation.
Vegetation sampling methods: random plot per stand across Fort Benning (∼1 plot per acre). Measured pine and hardwood BA. Measured percent composition of herbaceous, bare ground, pine straw, and woody vegetation within a 3.4-m radius around sampling points. The height of the majority of hardwood midstory vegetation within an 11.3-m radius around sampling points was categorized as low (<2.1 m), medium (2.1–4.6 m), or tall (>4.6 m), and hardwood midstory density was categorized as sparse, moderate, or dense.

Smith, Lora & Joseph Jones Ecological Research Center tortoise survey crew
Note: ongoing vegetation sampling in conjunction with burrow surveys.
Vegetation sampling methods: One random point per transect (∼50–400m spacing). Collect basal area, canopy cover (densiometer), dominant overstory type (P/H/M/Other/None). Mid-

story (1–3 m, 5-m radius) percent cover and type (P/H/Shrubs/M/Other/None). Ground cover ($\leq$1m, 1m radius) percent cover and type (bare, litter, grasses, vines, woody (shrubs), other (e.g., cactus).

Appendix C — Relating ground-based vegetation data to multi-season Landsat spectral indices

This appendix describes the development of a predictive model of vegetation attributes, parameterized from field-collected data (Appendix A). This work was supported under a grant from the National Council for Air and Stream Improvement. Full details on model development and assessment are provided in Hepinstall-Cymerman et al. (2017).

C.1. Remotely sensed data acquisition and processing

We utilized the applications-ready Landsat 8 Surface Reflectance-Derived Spectral Indices produced by the U.S. Geologic Survey (Masek et al., 2006; USGS, 2016). These products are pre-processed to facilitate analysis of vegetation change and include the following spectral indices: Normalized Difference Vegetation Index (NDVI), Enhanced Vegetation Index (EVI), Soil Adjusted Vegetation Index (SAVI), Modified Soil Adjusted Vegetation Index (MSAVI), Normalized Difference Moisture Index (NDMI), Normalized Burn Ratio (NBR), and Normalized Burn Ratio 2 (NBR2). We acquired these indices for all available Landsat 8 scenes overlapping our study area from Jan 2014 to Sep 2016, excepting those which had more than 40 percent cloud cover, 491 unique scenes in all. For each scene we used the CFmask band to mask all pixels that were not classified as "clear" (free of clouds or cloud shadows); data from masked pixels were not used in subsequent analysis.

In an initial step in the modeling process we needed to determine how to best extract seasonal changes in spectral indices from the remotely sensed data that would reflect vegetation patterns of interest. Given the available cloud-free imagery in 2015-2016 we determined that a 90-day window would be the shortest time frame from which we could extract seasonal measures of each spectral index and still ensure complete temporal coverage of the annual cycle and geographic coverage across our study area. As such we then proceeded to extract the values of all spectral indices at our field locations and computed the mean values across all available scenes within 90-day moving windows so that we could visualize the annual trend in each spectral index. We plotted these spectral trends with respect to field sites containing differing levels of canopy closure and ground-level vegetation as a means to identify the seasons and spectral indices that provided the greatest discernibility for these vegetation attributes (Figure C.1).

After this preliminary analysis we determined the vegetation indices (NDVI, EVI, SAVI, and MSAVI) were largely redundant and that the best separation was apparent using EVI. In addition the NBR and NBR2 indices appeared to provide additional spectral information that was not highly correlated to EVI. For each of these three indices (EVI, NBR, and NBR2) we next selected three 90 day time periods that captured the seasonal trend in the index and used our GIS to create a raster representation for each time period/index combination. For each index we also created difference rasters between the three selected time periods. This resulted in a total of 15 seasonal spectral indices (Table C.1) derived from the Landsat 8 data that we used as predictor variables in our subsequent models of vegetation attributes. Finally, we also

acquired 30-year mean annual temperature data (PRISM, 2012) and re-sampled the raster to match the resolution of our seasonal spectral indices.

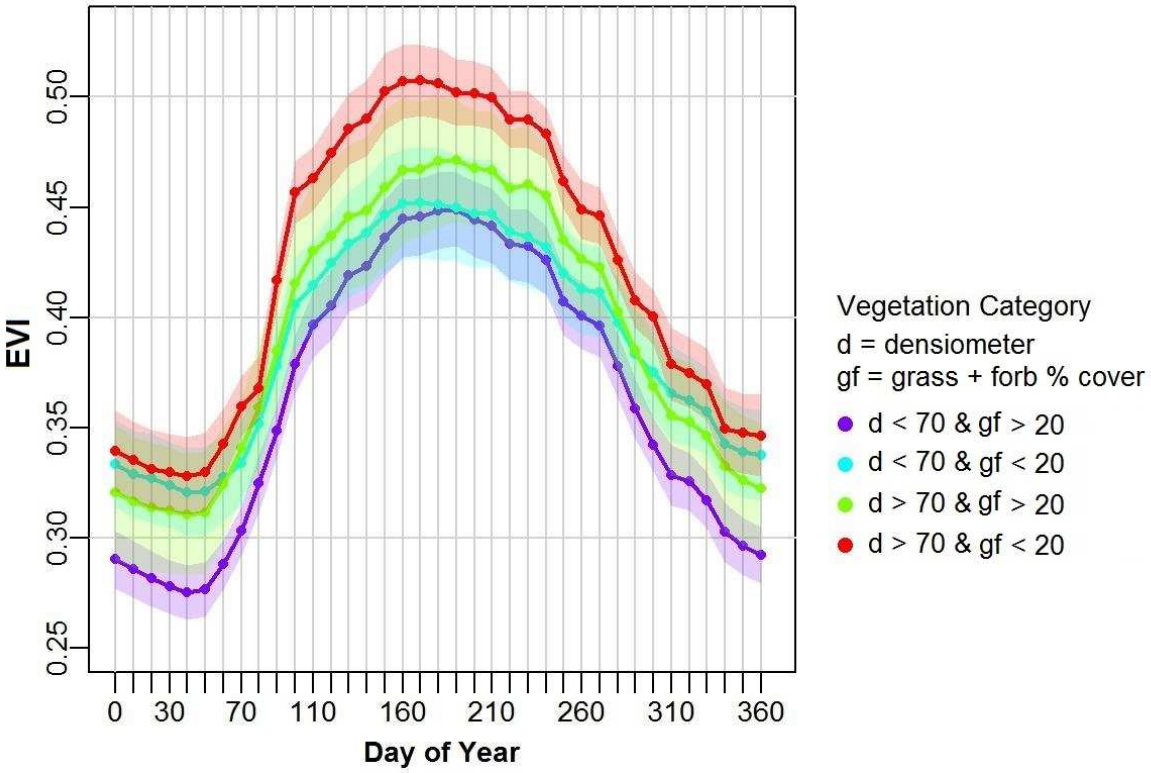

Figure C.1.: Example of an annual trend in EVI averaged across all available scenes using a 90-day moving window. Four vegetation categories are shown, points indicate mean values, and shading indicates 95% confidence intervals.

C.2. Predicting vegetation characteristics

Prior to modeling, all field sites were inspected using 2015 aerial imagery and the most recent Landsat image to identify where forest harvest appeared to have occurred after the field measurements were collected ($n = 14$), sites that were less than 60 meters to a hard edge (e.g. row-crop field, roads, water) ($n = 9$), or both ($n = 8$). Such sites were removed from further analysis, leaving a total of 154 field sites to use as training data. We used random forests (RF) regression (Breiman, 2001) to model field vegetation summaries (see Table A.3) using our 15 seasonal spectral indices and mean annual temperature as predictor variables. Each of the six field vegetation summaries (`covr`, `dens`, `fbgr`, `other`, `shrb`, `pine`) were modeled separately using the package randomForest version 4.6-12 (Liaw and Wiener, 2002) in R version 3.3.0 (R Core Team, 2017). All RF models were fit using the following parameters: ntree = 10000 (number of trees), mtry = 4 (number of random predictors sampled for splitting at each node), nodesize = 5 (maximum size of terminal nodes). We evaluated the predictive accuracy of each RF model by computing error statistics on the out-of-bag predictions (data points which were withheld during fitting of each RF tree). Error statistics include root mean squared error (RMSE), bias (mean error), and pseudo-R^2 which was equal to (1 - MSE / variance(y)) where y is the response variable. Inspection of our initial results indicated the observed and predicted data for the `other` vegetation summary followed geographic patterns in mean annual temperature, which we were

Table C.1.: Days-of-year selected for seasonal spectral indices. Single values indicate the seasonal spectral index was based on 90-day window centered on the specified day of year. Two values indicate a difference was calculated between the two dates indicated using the above 90-day windows.

	Spectral Index		
Season	EVI	NBR	NBR2
Early	30	50	50
Mid	160	230	180
Late	230	300	300
Early Difference	30 - 160	50 - 230	50 - 180
Late Difference	230 - 160	300 - 230	300 - 180

Table C.2.: Model results and error statistics. Name = indicates the model response variable (described in Table A.3), RMSE = root mean squared error, Bias = mean error, pseudo-R^2 = (1 - MSE / variance), Observed Range = observed range in response variable. RMSE, ME, and pseudo-R^2 were calculated from out-of-bag samples.

Name	RMSE	Bias	Pseudo-R^2	Observed Range
`covr`	6.843	0.028	0.553	[0, 50]
`dens`	13.648	-0.097	0.601	[0, 95]
`fbgr`	15.890	0.505	0.504	[0, 81]
`other`	14.208	0.164	0.444	[0, 78]
`shrb`	18.465	0.311	0.258	[0, 89]
`pine`	0.369	0.007	0.443	[-1, 1]

concerned may have been a sampling artifact. Thus we compared the fit of models of `other` with and without temperature as a predictor variable and found negligible differences; as such we did not include temperature in our final model of `other`.

C.3. Results and discussion of predictive vegetation modeling

RF models of field vegetation summaries varied in predictive accuracy but in most cases explained a substantial proportion of the observed variability (Table C.2). The two best-fitting RF models were those that modeled canopy conditions: `dens` and `covr`, which had pseudo-R^2 values of 0.60 and 0.55 respectively. RSME values for `dens` and `covr` were also low (13.65 and 6.84 respectively). RF models of the ground-level vegetation—represented by `fbgr` and `other`—had lower predictive ability than the canopy models but were still able to explain considerable variability with pseudo-R^2 values of 0.50 and 0.44 respectively. The lowest performing model was that of `shrb` (percent shrub cover) with a pseudo-R^2 value of 0.26 and RMSE = 18.5. Our measure of model bias was biologically negligible in all cases (Table C.2).

We considered the possibility that the model for `fbgr` may be overly sensitive to canopy conditions and investigated this possibility by comparing the correlation between `fbgr` and `dens` in our observed and predicted data. While we found the predicted correlation to be slightly higher than in the observed data ($r = -0.54$ *versus* $r = -0.50$), we do not feel this

indicative of strong bias, as the `fbgr` predictions varied widely with `dens` (Figure C.2).

Our results indicate that a substantial amount of the variability in canopy and understory vegetation conditions in upland coastal plain forests can be explained using remotely sensed Landsat 8 imagery. In five out of the six field vegetation summaries we modeled (i.e., all but `shrb`), the RF models explained between approximately 45–60% of the variability in site conditions. These models had very low bias though model precision was moderate with RMSE values in models estimating percent cover ranging between 6.8 and 15.9. The applicability of these models will depend on the scientific question and necessary precision; but in general the models we developed appear capable of distinguishing closed canopy forest stands from stands characterized by open canopies with grasses and forbs in the understory. In contrast to predicting canopy and ground-level conditions our model of mean percent shrub cover was much less accurate. This could be driven by either a lack of a unique spectral signal from the shrubs present in this study area or could be because the field measurements failed to adequately capture shrub coverage. It is possible that the percent cover of patchily distributed shrubs could be less accurately captured by the field observations than the other vegetation measurements. It is also worth acknowledging that while the number of field sites used to train these models ($n = 154$), was sufficient to achieve acceptable predictive accuracy in most cases we expect further accuracy gains could be achieved with a larger training dataset.

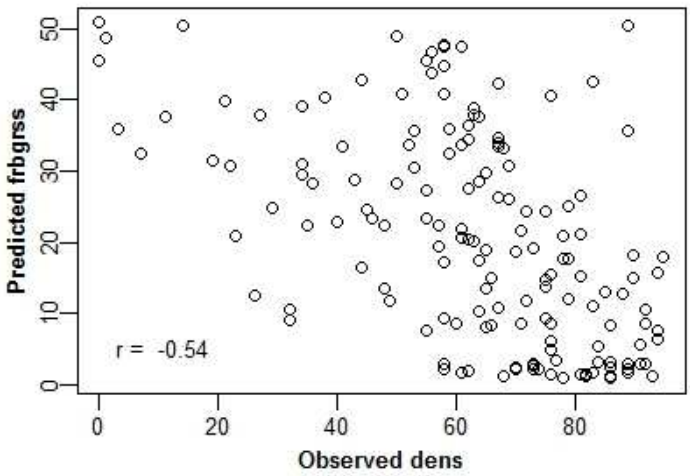

Figure C.2.: Relationship between predicted forb and grass cover `fbgr`, and observed densiometer `dens`.