7-2012

Species Distribution Model for the Invasive Plant Japanese Stiltgrass

Christopher Walter Tyson
College of William and Mary

Follow this and additional works at: https://scholarworks.wm.edu/honorstheses

Recommended Citation

https://scholarworks.wm.edu/honorstheses/476

This Honors Thesis is brought to you for free and open access by the Theses, Dissertations, & Master Projects at W&M ScholarWorks. It has been accepted for inclusion in Undergraduate Honors Theses by an authorized administrator of W&M ScholarWorks. For more information, please contact scholarworks@wm.edu.
Spatial Application of a Species Distribution Model for the Invasive Plant

Microstegium vimineum in the Chesapeake Bay Lowlands Ecoregion

An honors thesis submitted in partial fulfillment of the requirement for the degree of Bachelors of Science in Biology from The College of William and Mary

By

Christopher Walter Tyson

Dr. Matthias Leu, Committee Chair

Dr. Martha Case,

Dr. Randolph Chambers

Dr. Stuart Hamilton

Williamsburg, VA

April 11, 2012
ABSTRACT

_Microstegium vimineum_ (Trin.) A. Camus (Japanese stiltgrass) is a widespread shade-tolerant exotic plant species throughout much of the eastern United States. Where it spreads, _Microstegium_ profoundly affects ecological functions, altering soil chemistry and hydrology, displacing native flora, and thus reducing flora and fauna diversity. Successful control of this noxious species is highly contingent on early detection before large seed banks are established. As such, identifying areas at risk for invasion would allow conservation managers to better apply resources for maintenance and control. In this study, I developed a predictive species distribution model for _Microstegium_. Based on field surveys of 160 points throughout the Chesapeake Bay lowlands and Geographic Information System (GIS) analysis of landscape features, I developed a spatial model predicting patches likely to be currently or in the future invaded by _Microstegium_. I identified a suite of 11 landscape features and metrics that were important explaining the distribution of _Microstegium_. Habitat type variables (viz. proportion of clear-cut lands within a 2-km radius, Atlantic mesic forests within a 270-m radius, and dry-mesic oak forests within a 2-km radius) primarily had the strongest predictive value on _Microstegium_ distribution. Additionally, several anthropogenic features, such as distance to various road types, and the distance to water bodies were identified as predictive of sites likely to be invaded by _Microstegium_. Ultimately, my model had high predictive success for sites unoccupied by _Microstegium_, but only low predictive success for invaded sites. While species distribution models are inherently limited by their inability to distinguish between current and future occupancy, this model can likely be improved through a more stratified sampling scheme and additional field surveys.
INTRODUCTION

In addition to direct habitat loss, invasions by exotic species constitute one of the most severe threats to biodiversity across the globe. Nearly half of threatened or endangered plant and animal species are so due to an exotic, invasive species of some sort (Wilcove et al., 1998; Pimentel, 2005); an estimated $120 billion per year is spent on exotic maintenance or is lost due to damage (Pimentel, 2005). Exotic plant invasions, especially, are recognized as profound agents in altering ecological processes by displacing native flora and thus habitat for animals, which ultimately leads to biodiversity loss, as either extirpations or even extinctions. Once established, exotic plants are difficult to eradicate and require substantial resources. As such, identifying exotics with the potential to become invasive as well as the factors associated with invasion are essential to both prevention and eradication programs.

*Microstegium virens* (Japanese stiltgrass) is one exotic invader that has gained notoriety for increasing habitat homogeneity and biodiversity declines throughout deciduous forests in the eastern United States (Adams and Engelhardt, 2009; Flory and Clay, 2010). *Microstegium* was first reported in Tennessee in 1918, at which time *Microstegium* was used as a packing material for imported Chinese porcelain (Barworth et al., 2003). Since its introduction, *Microstegium* has rapidly expanded and now can be found in nearly all of states east of the Mississippi River. As an annual, shade-tolerant, C₄ species, *Microstegium* is an aggressive competitor that can colonize a range of soil and light conditions (Claridge and Franklin, 2010; Droste et al., 2010). Seed production for *Microstegium* is variable between populations, but is generally high, with upwards of 1000 seeds per tiller (Claridge and Franklin, 2002).
Chasmogamous (cross-pollinated) and cleistogamous (self-pollinated) seeds are both produced, with chasmogamous being dominant (Gibson et al., 2002; Huebner, 2003). The literature suggests a variety of potential dispersal mechanisms for seeds (and sometimes whole plants) including water, wind, animals (notably, white-tailed deer [Odocoileus virginianus] and humans); water, however, is generally considered to be the dominant vector (Tu, 2000; Swearingen, 2004; Romanello, 2009). Seeds are reported to remain viable for an average of three to five years and as such may persist in the forest understory for considerable time before conditions induce germination (Swearingen, 2004; Romanello, 2009).

Within ecosystems, *Microstegium* has been implicated in a variety of negative effects across trophic levels with the foremost cited being a reduction of native plant diversity and biomass, suppression of forest succession, alteration of soil nutrient cycles, drawdown of the water table, and the reduction of arthropod community diversity (Adams and Engelhardt, 2009; Flory and Clay, 2010; Civitello et al., 2008; Freyer, 2011). Given the ties of these processes to higher trophic levels, such as small mammal communities or bird populations, *Microstegium* has the potential to engender massive ecosystem changes, leading some to argue that *Microstegium* is an ecosystem engineer in invaded regions (Baiser et al, 2008).

Until recently, the prevailing focus on exotic plant invasion has been biased towards early-successional species. Recent research, however, suggests that shade-tolerant species (e.g. plants more capable of expanding into forests) may be equally invasive and would therefore present a threat to even protected natural areas distant from anthropogenic features (Martin et al, 2009). While roads and other mechanisms of disturbance are implicated in early-
successional invasions, the general understanding of how anthropogenic activities influence late-successional exotic plant invasion is quite limited (Gebard and Belnap, 2003; von der Lippe and Kowarik, 2008). I take *Microstegium* to be a potential model organism for shade-tolerant species and as such hope to learn general patterns applicable to shade-tolerant invasive plants.

An understanding of life-history traits and the current literature on *Microstegium* distribution suggest several factors that are likely associated with invasion patterns. *Microstegium’s* essential niche requirements primarily consist of medium to low light, medium to high soil moisture and reduced leaf litter (Nord et al., 2010; Warren et al., 2011). Another factor potentially influencing *Microstegium* involves over-browsing by white-tailed deer, which may result in reduced competition from native flora (Tilghman, 1989; Baiser et al., 2008). Moreover, like most invasive species, anthropogenic features (e.g. roads, development) are often associated with exotic plant invasion (Gelbard and Belnap, 2003; Christen and Matlack, 2006). Consequently, factors such as canopy cover, local hydrology, deer population density, and disturbance are a priori possible influences in *Microstegium* invasion (Kuhman, 2010). Identifying landscape level proxies of these local processes will allows us to form predictive models of *Microstegium* establishment and invasion.

While some studies have identified factors influencing *Microstegium* distribution at the local scale, the question of how this particular invasive responds to factors at the landscape scale remains largely unanswered. Many studies increasingly recognize that identifying the relevant spatial and temporal scale is crucial for understanding an organism’s potential to become an invasive (Johnson 1980, Morris 1987, Levin 1992). As these scales are likely to be
highly individuated, they must be explored on a case by case basis and within a particular context. Consequently, understanding how any specific local variable influences *Microstegium* must be critically examined in order to capture how *Microstegium* responds to landscape features. Characterizing the effect of landscape features (e.g. anthropogenic, abiotic, habitat) on the distribution of *Microstegium* within a spatial framework ultimately allows one to model and thus map areas potentially susceptible to invasion (Pauchard, 2006).

In this paper, I applied results from a two-year study to develop a predictive model of *Microstegium* occurrence. Specifically, my objectives were to first identify general landscape features associated with *Microstegium* occurrence. Given limited knowledge of the salient variables and scales influencing *Microstegium* dispersal, I used a hierarchical modeling approach to determine the most parsimonious model (Aldridge et al. 2012). Second, I spatially applied the final model to my study area in a Geographical Information System (GIS). And third, I validated the final model with independent data. Ultimately, the map developed in this paper will inform local and regional management of *Microstegium* by delineating areas where it currently occurs and identifying areas with a high potential of being invaded.

**METHODOLOGY**

*Site selection*

I selected sampling sites to reflect a gradient of human disturbance based on road density, level of development, and a wildland-urban interface (WUI) classification. I used the WUI categorization developed by Radeloff et al. (2005), which classifies areas according to
housing density, percent forest cover, and distance to forested areas. The intent of allocating sampling sites across such a gradient was to produce a meaningful variance of invaded sites, based on a priori assumptions about the nature of Microstegium invasion. As such, field sites ranged from protected conservation lands to state parks to residential areas. The study area extent was based off the Chesapeake Bay Lowlands ecoregion (The Nature Conservancy, 2007), which is characterized by temperate broadleaf and mixed forests with substantial floodplain and mesic forest types. Within the ecoregion, the study area was further delimited by the James River to the south, the Rappahannock River to the north, the Piedmont ecoregion to the west, and the Chesapeake Bay to the east (Figure 1 / Figure 2). I restricted my analysis to this area to ensure that results would have equal validity across the entire extent. I randomly generated survey points within study sites with a minimum distance of 600 m between points to reduce landscape autocorrelation effects (Legendre, 1993; Nielson et al., 2011). The 600-m inter-point distance was based on an evaluation of forest patch and flood plain distributions.

Sampling protocol

Sampling occurred between June and late July during the summers of 2010 and 2011. Given that Microstegium germinates as early as March in Virginia, this survey was sufficient for assessing the presence or absence at each site (Fryer, 2011). Each point center was located using a handheld GPSmap 90Cx Garmin, which is accurate up to +/- 3m. Field plots consisted of two 30-m transects arranged in the cardinal directions.

I assessed Microstegium cover using the line-intercept method (Herrick et al., 2009). I dropped a wooden dowel of 1-m height every 1-m on the transect and noted each plant species
touching the dowel (59 measurements per plot). I estimated *Microstegium* biomass (percent cover and height) via 0.5-m² quadrats placed every 5 m along the transect (11 measurements per plot). I also conducted visual searches within a 15-m radius of the plot center to ensure that exotic species not detected on transects were still reported. To further describe the local vegetation, I also collected various measurements, which are not reported here as my study focused on landscape rather than local factors influencing *Microstegium* invasion.

In addition to vegetation structure, I also assessed white-tailed deer abundance via fecal pellet counts along transects at each plot. Fecal pellet counts were taken during the summers of 2010 and 2011 along two 120-m long transects. As preliminary sampling during 2010 revealed that 30-m long transects resulted in insufficient observations, each vegetation transect was extended by an additional 30-meters to the north and west. For each pellet observation I measured the distance from the transect to the pellet mass and classified pellet observations as either single, groups or clusters (i.e. combined groups of an indistinct number of pellets). Using program Distance 6.0, I estimated pellet density per hectare adjusted for detection probability for each site (Thomas et al., 2009). I included time of day, wind speed, and temperature for each observation as these variables could influence detection probability. I also measured pellet decay during the time of sampling to adjust density estimates for potential pellet loss. Pellet decay was measured on the campus of the College of William and Mary in the forest patch next to Barrett Hall. I selected 40 groups of pellets in 2010 and 2011 and counted the number of remaining pellets every three days during June and July.
Landscape Analysis

Aside from micro-plot characteristics, various spatial features were assessed based on an extensive review of the invasive literature, specific life-history traits of Microstegium, and anecdotal evidence. Landscape feature data were compiled from various online resources (Table 1) and were stored in ArcGIS 10 (ESRI, 2011). I chose a 30-m spatial resolution for my study based on the coarsest land cover dataset.

I considered three predictor variables classes: abiotic, anthropogenic, and habitat. In the abiotic category, variables consisted of elevation, distance to and density of hydrological features, focal flow, soil pH, water content at one meter, and a terrain roughness index (Cole and Weltzin, 2004; Fryer, 2011; Kuhman, 2011). Anthropogenic variables included distance to and density of roads and traffic volume (Chirsten and Matlack, 2006; von der Lippe and Kowarik, 2008). I defined habitat variables as biotic habitat variables or land cover variables. As such, habitat variables included various metrics of landscape composition and configuration as well as land cover type proportion and hunting frequency (Martin 2009; Baiser 2008).

I used land cover classifications based on the Southeastern Gap Analysis Project (http://www.basic.ncsu.edu/segap/), but reclassified these types within the context of Microstegium invasion (NatureServe, 2007). My new reclassification focused on developed and mesic land cover types given the facultative relationship between such areas and Microstegium invasion (Fryer, 2011). Ultimately, I derived nine land cover types (Table 2 / Figure 3). Landscape metrics and habitat proportions were based on these reclassifications.
In order to characterize landscape composition and configuration, I employed various metrics at the landscape and class level. Class level metrics involve reference to specific land cover types, whereas landscape level metrics took the landscape as a whole, only focusing on patches in general. Metric selection was aimed at assessing likely features influencing *Microstegium* invasion. As such, general metrics of aggregation (indirect fragmentation) and weighted-edge density were considered relevant. For class level metrics (see Table 3), I measured aggregation index (AI) and contrast weighted edge-density (CWED), based on a contrast-weighted matrix sensitive to developed and mesic habitat types. Cushman et al. (2008) found that these metrics form a parsimonious suite of highly universal land and class scale features. At the landscape level, I assessed the Shannon diversity index (SHDI), contagion (CONTAG) and landscape shape index (LSI), each at the six scales (Table 4). I calculated landscape and class metrics as well as land cover proportion using Fragstats version 3.3 in combination with Fragstats Batch (McGarigal, 2002).

To spatially estimate the effects of deer browsing, I developed a map of hunted areas as a proxy variable with the expectation that hunting frequency and browsing intensity would be negatively correlated. Sites were scored 0 through 3 (0: never hunted, 1: open for periodic, but not seasonal hunting, 2: private lands that are potentially seasonally hunted, 3: public access lands that are seasonally hunted). Where possible I reclassified sites to this scheme based on their exact hunting policy, and where not on the wildland-urban index (Radeloff, 2005).

Given the limited information on the response of *Microstegium* to spatial patterns, appropriate landscape variables were measured at a range of scales. Canopy openness, road
density, hydrological feature density, land cover type, traffic volume and landscape metrics were all assessed within six extents, namely, 90-m, 270-m, 540-m, 1-km, 2-km and 3-km, based on prior research on exotic spread as well as the home range for white-tailed deer, which has been estimated at approximately a 3-km radius (see Hirth, 1977; Nielsen et al., 2011). Furthermore, both Euclidean distance and exponential decay functions were calculated for distance-based variables, such as nearest road or nearest stream. Decay functions were of the form \((1-e^{\alpha/d})\) where, \(d\) was the distance in meters to a feature of interest and \(\alpha\) was set to either: 90-m, 270-m or 540-m, corresponding to my selected scales. Decay distances were included as ecological processes may not follow a linear relationship (Nielsen, 2011). The Moran’s I Index revealed no issues of spatial auto-correlation between landscape features.

**Model Development**

I used a hierarchical modeling approach to determine the most parsimonious model. Given the novelty of this modeling approach and a lack of knowledge concerning *Microstegium* landscape distribution, traditional *a priori* modeling and hypothesis testing were infeasible. As such, model development and selection followed an Akaike Information Criterion (AIC) approach (Burnham and Anderson, 2002), which penalizes models based on the number of predictor variables. I identified the best variables individually within each of abiotic, biotic, and human stressor sub-models and then carried forward the top variables of each sub-model to determine the top-model (Leu et al., 2011, Aldridge et al. 2012). As such, an overabundance of variables was initially identified as potentially salient. To simplify the variable pool and avoid multicolinearity issues, highly correlating variables based on Spearman rank coefficients
(≥|0.7|) were removed from the analyses (Aldridge and Boyce, 2007). Additionally, variables with less than 20 non-zero (or insufficiently variable) data values were discarded to avoid perfect fit. I also checked for nonlinear relationships between Microstegium occurrence and predictor variables using scatter plots; no apparent quadratic relationships were found.

Together the three variable classes formed a total 88 predictor variables. Generalized linear regression models with a logistic link function were run for each variable. Model structure was of the form:

\[ w(x) = \exp(\beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_k x_k) \]

where \( w(x) \) is the probability of selection for model parameters, \( \beta \) and \( x \) are the coefficients and model parameters, respectively (Aldridge, 2012). Based on the AIC score, the optimal scale and distance measure for each variable was then identified and carried forward to establish the best model for each category of interest (Burnham and Anderson, 2002). Following the univariate analysis, variables were separated into their respective categories: biotic, abiotic, and human stressors. Multivariate models were then formed from within these variable categories. I restricted the model combinations to seven variables based on the number of presence sites; approximately one variable per ten observations (Hosmer and Lemeshow, 2000). Running a further set of generalized linear regression models yielded a top model for each category. Variables from these three top models were then placed in the top model pool. From this pool, all combinations of variables were combined to identify the top models. Furthermore, to produce a more robust predictive model, I model averaged variables of all
models with combined cumulative weights of ≥ 0.9 (Burnham and Anderson, 2002). Model weights were readjusted based on the models included in the final set of models.

**Spatial Application**

Applying the overall model formula using the Raster Calculator tool within ArcMap, the model-averaged beta values were spatially applied for each variable in the final variable pool, yielding a final predictive occurrence surface for *Microstegium* across the study area. This probabilistic map was reclassified into a binary output representing presence / absence based on the specificity-sensitivity threshold (Liu et al., 2005). Given that water bodies and high urban areas were not sampled, such areas were removed from the model output map.

**Model Evaluation**

I assessed the fit of the final model against 41 independent survey sites visited during the summer of 2011. These validation sites were selected to mirror the disturbance gradient of the original 160 sites (Figure 1). Various approaches have been proposed for evaluating species distribution models (Johnson et al., 2006; Leu et al., 2011). Using receiver operating characteristic (ROC) (Metz 1978) analysis to derive presence and absence probability bins, I transformed the continuous probability of occurrence predicted by my model into a binary presence / absence model (Hebel and McCarter, 2006, Leu et al. 2005). The expected invasion status predicted by the final model at each validation point was then compared to observed values.
RESULTS

*Microstegium* was present at 67 of the 160 sites surveyed between 2010 and 2011. I found neither presence nor biomass of *Microstegium* to correlate with the date (Spearman rho: 0.025 and -0.017, respectively), which suggests that the survey scheme was appropriate for assessing whether *Microstegium* was indeed present. Of the 67 invaded plots determined via area searches, I only detected *Microstegium* on 48 (71.6%) transects. The average height was 19 cm (n = 271) for all *Microstegium* found along transects. For the validation plots, I found a total of 14 invaded sites (35%, n = 41). The average height was 22 cm (n = 48).

Fecal pellets were used to estimate the density of white-tailed deer. The average pellet count at each point was 54 (n=160, with 101 pellet encounters). Detection probability was 0.34 based on a half-normal cosine function. Plot deer pellet count negatively correlated with the average index of hunting pressure within a 270-m radius (Spearman rho: -0.22, p= 0.01). This significant correlation substantiates my spatial application of hunting intensity.

*Landscape Models*

After assessing variables for scale, type (linear or decay) and correlations, I retained 28 variables (see Table 5). Habitat variables were primarily the top predictor variables (both in terms of AIC and magnitude of effect). From the univariate analysis, the proportion of clear-cut lands within a 2 km-radius (CNT_5_2000m) was the top predictor variable (Table 5, Figure 4). After all category variable combinations were run, the total variable pool was further reduced to 11. From the abiotic set, the single variable model of distance to water bodies was identified as the top model (Table 6). The combination of distance to primary roads, rail and the decay
distances to local roads and secondary roads, with asymptotes of 90-m and 270-m, respectively, was identified as the top anthropogenic model (Table 7). Lastly, five combined habitat variables were identified as the top habitat model. The proportion of clear-cut, Atlantic mesic forests (Figure 5), dry-mesic (Figure 6), within a 2-km, 2-km, and 270-m radius, respectively, along with the aggregation index of developed herbaceous and Atlantic coastal plain land cover was carried forward as the top habitat model (Table 8). I opted to include deer hunting, a proxy for deer browsing, as a separate variable category given its status as both a habitat and disturbance predictor. From these 11 variables, I developed final candidate models (K ≤ 7). All models generated add substantially lower AIC value from the null model. From the large number of models generated, weighted averages were low (Table 9). All variables ultimately identified by the hierarchical modeling approach had strong justification from the literature. The model averaged beta-values suggest that *Microstegium* is likely to invade areas with: low proportion of clear-cut land within a 2-km radius, high proportion of Atlantic mesic forest habitat within a 270 m radius, close to water bodies and secondary roads, maximally aggregated Atlantic coastal plain forests, and dispersed developed herbaceous habitat (Table 10).

Validation

The spatially applied model produced a continuous probability of invasion for *Microstegium* across the landscape, which ranged from virtually one to near zero (Figure 7). The final model had excellent predictive capabilities (ROC = 0.81). At the sensitivity-specificity threshold 75.6% of points were classified correctly. The cut-off between absence and presence
was (0.99833). The transformed binary presence / absence model (Figure 8) for Microstegium had moderately high success in predicting unoccupied sites (0.75), but low success in predicting occupied sites (0.53). As such, expected versus observed results indicate marginal overall success for the presence / absence model (Table 11).
DISCUSSION

The final model comprised a suite of variables that are influential in identifying key habitat for *Microstegium*. My model ultimately had high predictive success for unoccupied sites, but relatively low predictive success for occupied sites. Landscape variables selected during the hierarchical modeling approach had strong support from the literature, which validates my approach, as well as the plausibility of the model as a whole.

*Habitat Variables*

Among the most influential in predicting invasion probability of *Microstegium* were land cover variables, namely habitat proportion and landscape metrics. Certainly the most dominant of this set was the proportion of clear-cut lands within a 2-km radius (β-Value: -284.5168), which showed a profoundly negative relationship to *Microstegium* distribution. Clear-cut stands are often characterized by disturbed soils, low moisture, high direct light, higher elevation and thick accumulations of pine needles (NatureServe, 2007). Given that *Microstegium*’s niche requirements primarily consist of mesic soils and moderate shade, clear-cut stands are likely to be resistant to *Microstegium* invasion (Freyer, 2011). Furthermore, low elevation sites are often a target for *Microstegium* seed dispersal, thus potentially excluding *Microstegium* from higher elevation clear-cut sites (Nord, 2010). And while logged areas are highly disturbed sites, thick pine needle mats likely serve as a barrier to *Microstegium* implantation (Schramm and Joan, 2010). Taken together, each of these factors makes clear-cuts highly resilient to *Microstegium* invasion. That the proportion of clear-cut stands was found to be the most relevant at a 2-km radius is most likely explained by a threshold effect of
clear-cut lands and Microstegium occupancy. Specifically, small-scale clear-cut stands are not likely to be indicative of large clear-cut operations, which are most fully inhospitable to Microstegium. As such, when there is a high proportion of a clear-cut land within a 2-km radius, it can be expected that Microstegium is largely absent. Still, our clear-cut classification did not distinguish between the age of the clear cut nor did it consider the type of forest that was logged. Incorporating both of these factors might refine the predicted relationship between clear cuts and Microstegium invasion.

Two other habitat proportion variables were identified as dominant predictors of Microstegium habitat: Atlantic coastal mesic forest cover within a 270-m radius (β-Value: 26.6952) and dry-mesic oak forest cover within a 2-km radius (β-Value: 7.2768). Both variables are well supported by previous accounts of Microstegium invasion. Atlantic coastal plain mesic hardwood and mixed forests are characterized by non-wetland, low elevation, mesic flats that lie between drier pine uplands and floodplains (NatureServe, 2007). Such areas are likely to accumulate seeds via runoff from higher elevations making Atlantic coastal mesic forests likely areas for Microstegium invasion given their vulnerability to seeding and also in terms of niche preferences (Nord et al., 2010; Kuhman, 2010). That a relatively local scale was most influential is intuitive considering that patches of Atlantic coastal plain forest typically run alongside tributaries and are bordered by ridges. As such, more aggregated tracts of Atlantic coastal plain forests are likely to be found within a smaller scale. Dry-mesic oak forest is roughly similar habitat to Atlantic coastal mesic forest, yet differs in that this system may occur at higher elevations and is generally considered unfavorable for vegetation due to thin soil (NatureServe, 2007). Not surprisingly, this system was positively correlated with Microstegium, but to a lesser
extent. The 2-km scale identified as relevant for dry-mesic oak forest can perhaps be explained by the distribution of these areas as generally surrounding more hospitable lowlands. As such, higher proportions of dry-mesic oak forest at a 2-km scale are likely to be indicative of prime Microstegium habitat along tributaries in the area.

In addition to land cover proportions, two landscape metrics were identified as predictors of Microstegium invasion from within the habitat variable category. Both metrics were class aggregation indices, which relate the degree to which a focal class is aggregated within the landscape. Edges, either naturally occurring or due to anthropogenic fragmentation, are reportedly associated with Microstegium invasion (Christen and Matlack, 2006; Rauschert et al., 2010). The aggregation index of Atlantic coastal plain forests within a 2-km radius showed a marginally positive correlation with Microstegium invasion (β-Value: 0.0338). This relationship is likely explained by the fact that Atlantic coastal plain forests tend to be naturally aggregated alongside rivers and streams, and consequently the aggregation index at a large scale simply reflects greater hydrological activity, which is conducive to Microstegium transport and establishment (Figure 5). The aggregation index of developed herbaceous cover at a 3-km radius (β-Value: -0.1662) was also identified as a predictor Microstegium habitat. Developed herbaceous cover was a composite classification that primarily consisted of developed open space, which includes golf courses and recreational parks. This combination of anthropogenic disturbance and retained habitat make developed herbaceous cover especially vulnerable to invasive species (Huebner, 2003; Kuhman, 2010). My model identified a slightly negative relationship between aggregated patches of developed herbaceous space. Dispersed developed herbaceous cover inherently includes more edges and is likely fragmented by
anthropogenic features such as roads, which are known to facilitate invasion (Christen and Matlack, 2006; Freyer, 2011). Conversely, aggregated developed herbaceous cover will present fewer edges and is likely to include larger patches of open areas, subject to direct sunlight (e.g. golf courses), which justifies the negative relationship found by the model.

Abiotic Variables

The final model identified Euclidean distance from water bodies (β-Value: -0.0000052) as the only relevant, abiotic predictor of Microstegium distribution. The water bodies feature class included all swamps, marshes, lakes and ponds as classified by Cowardin et al. (1979). This variable had substantial a priori support based on facts of Microstegium’s niche specifications and its dispersal mechanisms. Water is largely recognized as the primary dispersal vector for Microstegium seeds, which can tolerate prolonged periods of submersion (Hunt and Zaremba, 1992; Virginia DCR, 2002; Romanello, 2009). The water bodies feature class was limited as it included both fresh and salt water. This can potentially explain the low beta-value given that the latter case would diminish Microstegium habitat (Hopfensperger and Baldwin, 2009). Thus a more restricted focus on fresh water areas might improve my model for Microstegium.

Additionally, although elevation was not ultimately included in the model, this parameter could supplement the shortcomings of water bodies as a predictor variable. Given that Microstegium is not likely to be found immediately alongside water bodies, but at slightly higher elevations, elevation and distance to water may interact and should be included as an interactive term in future models predicting Microstegium occurrence.
White-tailed Deer Hunting

Compared to other habitat-type variables hunting ($\beta$-Value: -0.3363) had a relatively minimal predictive value on Microstegium occupancy. Still, there is substantial evidence in both the literature and anecdotal observations from study sites to suggest that higher densities of white-tailed deer greatly increase an ecosystem’s susceptibility to invasion (Tilghman, 1989; Baiser et al., 2008; Schramm, 2008). Given this interesting and relatively unexplored relationship, I developed a novel spatial map for hunting intensity, which could be applied to map the effect of white-tailed deer hunting on Microstegium distribution. My index of hunting intensity was an informal characterization based on the surrounding land use or on the hunting policy of a site. Originally, I hoped to directly assess the relationship between Microstegium and white-tailed deer at each site. Over-browsing and invasive establishment, however, are temporally distinct, which makes it difficult to capture this relationship. While fecal pellet data was collected at each site, additional years of surveying are necessary in order to obtain reliable estimates of white-tailed deer use. While my spatial layer of hunting intensity significantly correlated with my field data of fecal pellet counts, further refinement of this map of hunted lands will improve spatial application of the relationship between hunting and Microstegium.

Anthropogenic Variables

Though less meaningful in their predictive power than habitat variables, several anthropogenic features were retained in the final model as significant predictor variables for Microstegium occupancy. Of these, the decay distance from secondary roads with a 270-m asymptote was the most influential ($\beta$-Value: -4.3070). Secondary roads are main arteries,
usually in the national, state or county highway system. Intersections with other roads or driveways are primarily at-grade. Such roads have a long disturbance history, which may explain their strong association with Microstegium. This relationship, however, is likely to diminish as the distance from this feature increases, as has been shown for other exotic plant invasions (Gelbard and Belnap, 2003; Christen and Matlack, 2006). Similarly, the decay distance from local roads with a 90-m asymptote (β-Value: -2.0909) was identified as predictive of Microstegium invasion, yet only within a relatively short distance. Considering that local roads are generally paved and non-arterial (e.g. neighborhood roads), with a generally shorter disturbance history, this finding is unsurprising (Mortensen and Rauchert, 2009; Nielson, 2011). Aside from sources of disturbance, road verges are known to create favorable conditions for Microstegium by collecting water runoff (Hunt and Zaremba, 1992). As such, invasions are likely to be centralized around roadways, with the possibility to expand over time. Both the Euclidean distance from railways and from primary roads (β-Values: 0.0001 and -0.0001, respectively), were identified as minimally predictive in the final model. The positive relationship between Euclidean distance from railways and Microstegium presence is most likely an artifice of the distribution of sampling sites. Plots close to railroads tended to be highly developed and as such offered limited habitat, while points further from tracks were more likely to be a mix of disturbed and yet viable habitat, which were highly suited to Microstegium invasion.

As roads potentially account for both seed dispersal and establishment of Microstegium, it is unsurprising that local, primary and secondary roads were all identified as predictive of Microstegium invasion. Though traffic estimates were included in the univariate analysis, they were ultimately not found to be significantly predictive, despite prior findings that might
suggest otherwise (Von der Lippe and Kowarik, 2008). A likely explanation is that the road classifications indirectly accounted for traffic volume, making this variable class redundant.

CONCLUSIONS

Though it is reassuring that all variables selected via a hierarchical modeling approach had strong support from the literature (and are congruent with the original expectations), my final model could be improved by refining the stratification of sampled sites and increasing sample size, which would allow for a more discerning model. An inherent limitation of species distribution models in general is the inability to distinguish between currently occupied sites and sites that may become occupied in the future (Franklin, 2009). Consequently, while my validation found that some sites were unoccupied where the model predicted invasion, it is impossible to know whether this is the result of error in the model, or simply that these sites are likely to be invaded in the future. While I did not take samples of the seed bank, such measurements would be a step towards refining estimates of occupancy and improving a distribution model for Microstegium.

Additionally, it is worth remembering that this model is limited to being predictive and not explanatory. Determining that Microstegium is less likely to be in clear-cut stands rather than in Atlantic mesic hardwood forests does not imply that Atlantic mesic hardwood habitat causes invasion to occur. Similarly, while distance to ORV trails may negatively correlate with Microstegium presence, it does not follow that building ORV trails will exhibit a negative effect on Microstegium. My model of presence and absence merely reflects associations between
certain landscape features and the likelihood of finding Microstegium. Determining the actual causes of Microstegium establishment is another matter. In some cases, it may be that the features correlated with Microstegium invasion are simultaneously causes for invasion. This might be the case in road development, where seeds are transported to a site by construction equipment and are also established through disturbance. Conversely, land cover types, which may be highly correlated with invasion by Microstegium are not causes in themselves, but require a separate factor to explain why some patches are invaded and some patches are not.

Future projects could address Microstegium at a regional level, which would have more broad scale applications. Given that this project was limited to a specific ecoregion, it is difficult to extrapolate outside of the study area. Conducting multi-state surveys would illuminate more large-scale processes, which could then be applied at a regional level. Still, while such a large-scale understanding would no doubt be valuable, changes in climate and land cover across the landscape might problematize drawing conclusions at too broad of an extent and could also limit the feasibility of such a project.

A critical next step in refining a species distribution model for Microstegium is to explore the relationship between the presence of seeds and germination. Identifying sites that are likely to manifest Microstegium invasions versus sites that are merely likely to be seeded would greatly improve any model of Microstegium distribution.

Furthermore, while I tested landscape effects at multiple scales, additional scales might better capture landscape effects on Microstegium distribution. My choice of extents was partially based on prior landscape-scale exotic research and on the maximum home range size.
of the white-tailed deer. The most descriptive scale, however, is likely to vary with the landscape feature, and as such would need to be investigated on a very individual basis. Refining the scale for each landscape feature would better capture the way in which *Microstegium* responds to the landscape.

Though my modeling effort focused on spatially applicable features, local variables warrant further attention. In a preliminary analysis of landscape and local variables, I identified landscape variables as far more predictive of *Microstegium* presence than the site variables I collected. A potential explanation for this finding was that some of the relevant site descriptions (e.g. available water content within the soil or soil pH) were not measured directly, but were represented by landscape variables such as distance to hydrological feature or elevation or habitat type. Still, identifying influential local variables might better inform the selection of correlating landscape features.

This model constitutes a primary effort in determining factors associated with *Microstegium* invasion. However, there remain other variables to consider and to include in my hierarchical modeling approach. Ultimately, it will be necessary to establish variables that are informative to conservation managers for determining areas with high probability of invasive presence and ideally can also be used to inform design plans and other active strategies for exotic control. My model does confirm that shade-tolerant exotic plant species have the potential to be invasive and may become established independent of anthropogenic effects. As such, even areas without seeming causes of invasion may still be susceptible to invasion. This knowledge
reinforces the need to improve our ability to identify areas of likely invasion in order to optimize exotic control efforts.

From this study, I have identified key predictive variables in determining likely locations for *Microstegium* invasion. Given the aggressive nature of this invasive species, early detection and treatment is essential for effective management (Flory and Lewis, 2009). Applying the variables identified through my analysis will allow conservation and park managers to more effectively allocate resources for exotic control by focusing efforts on areas likely to currently invaded or become invaded in the future.
ACKNOWLEDGMENTS

First and foremost, I would like to thank Matthias Leu, who has been the best mentor and advocate that a budding researcher could hope for. Before I could distinguish between crab grass and Japanese stiltgrass, Matthias supported my interest in ecological research. His dedication is contagious and is certainly responsible for my taking on this project. I will always be thankful for his example.

The fieldwork portion of this project was tough and without the help and support of Matt Feresten, Corbett Wicks, and Stephanie Wilson, it would have been far more so. We endured many hot summer days together, and I will always be thankful for the company they provided. I should also like to thank the other members of the Applied Conservation and Ecological Research lab who have been helpful in the research process, offering various suggestions and advice along the way.

Additionally, I should like to thank my committee members, Dr. M. Case, Dr. R. Chambers and Dr. S. Hamilton for taking the time to assist in this project. I know their various backgrounds will be greatly helpful for improving the final product.

Funding for this project was provided by the Howard Hughes Medical Institute, the Charles Center’s Dintersmith Honors Grant and the Virginia Crouch Memorial award. Thank you for making this experience possible.

Thank you to my close friends and family who have been supportive during the long evenings in the lab. I certainly could not have managed alone.
LITERATURE CITED


McGarigal, K., Cushman, S.A., Neel, M.C., Ene, E., 2002. FRAGSTATS: Spatial Pattern Analysis Program for Categorical Maps. Computer software program produced by the authors at the University of Massachusetts, Amherst, available at the following website: http://www.umass.edu/landeco/research/fragstats/fragstats.html


Pauchard, A., Shea, K. 2006. Integrating the study of non-native plant invasions across spatial scales. Biological Invasions 8:399–413


### Table 1: List of spatial data sources used in this study.

<table>
<thead>
<tr>
<th>Spatial data</th>
<th>Category</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Land cover</td>
<td>Habitat</td>
<td><a href="http://www.basic.ncsu.edu/segap/">http://www.basic.ncsu.edu/segap/</a></td>
</tr>
<tr>
<td>Wildland Urban Interface</td>
<td>Habitat</td>
<td><a href="http://silvis.forest.wisc.edu//maps/wui/state">http://silvis.forest.wisc.edu//maps/wui/state</a></td>
</tr>
</tbody>
</table>
Table 2: Land cover reclassifications from the Southeast GAP project ([http://www.basic.ncsu.edu/segap/](http://www.basic.ncsu.edu/segap/))

<table>
<thead>
<tr>
<th>Reclassified Cover Name</th>
<th>Total Area (km²)</th>
<th>Composite Cover Types</th>
<th>Area (km²)</th>
</tr>
</thead>
</table>
| Developed Herbaceous Space | 1536.6609       | Developed Open Space  
Quarry/Strip Mine/Gravel Pit  
Pasture/Hay  
Row Crop                              | 238.8771   |
| Salt water wetland      | 156.8421        | Atlantic Coastal Plain Northern Tidal Salt Marsh  
Atlantic Coastal Plain Northern Maritime Forest | 127.6308  |
| Developed               | 439.5330        | Low Intensity Developed  
Medium Intensity Developed  
High Intensity Developed  
Bare Soil                 | 261.6615   |
| Dry Pine Forest         | 1004.5233       | Southern Piedmont Dry Oak-(Pine) Forest - Hardwood Modifier  
Evergreen Plantations or Managed Pine (can include dense successional regrowth)  
Southern Piedmont Dry Oak-(Pine) Forest - Loblolly Pine Modifier  
Southern Piedmont Dry Oak-(Pine) Forest - Mixed Modifier | 30.9951   |
| Clear Cut               | 664.4745        | Successional Shrub/Scrub (Clear Cut)  
Successional Shrub/Scrub (Other)  
Clearcut - Grassland/Herbaceous  
Other - Herbaceous          | 592.1118   |
| Open Water              | 524.2257        | Open Water (Fresh)  
Open Water (Brackish/Salt)           | 106.6545   |
| Dry-Mesic Oak Forest    | 1319.5323       | Atlantic Coastal Plain Dry and Dry-Mesic Oak Forest                                     | 1319.5323 |
| Floodplain forests      | 1165.6881       | Unconsolidated Shore (Lake/River/Pond)  
Atlantic Coastal Plain Blackwater Stream Floodplain Forest - Forest Modifier  
Atlantic Coastal Plain Small Blackwater River Floodplain Forest  
Southern Piedmont Large Floodplain Forest - Forest Modifier  
Southern Piedmont Small Floodplain and Riparian Forest  
Atlantic Coastal Plain Northern Basin Swamp and Wet Hardwood Forest  
Atlantic Coastal Plain Northern Tidal Wooded Swamp  
Atlantic Coastal Plain Northern Fresh and Olgohaline Tidal Marsh | 6.92469   |
| Atlantic Coastal Plain Mesic Hardwood and Mixed Forest | 1356.6762 | Atlantic Coastal Plain Mesic Hardwood and Mixed Forest                                | 1356.6762 |
Table 3: Definitions of class level metrics.

<table>
<thead>
<tr>
<th>Metric</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Contrast-Weighted Edge Density (CWED)</td>
<td>Contrast-weighted edge density standardizes edge to a per unit area basis that facilitates comparison among landscapes of varying size.</td>
</tr>
<tr>
<td>Aggregation Index (AI)</td>
<td>Aggregation index is calculated from an adjacency matrix, which shows the frequency with which different pairs of patch types (including like adjacencies between the same patch type) appear side-by-side on the map. Aggregation index takes into account only the like adjacencies involving the focal class, not adjacencies with other patch types. In addition, in contrast to all of the other metrics based on adjacencies, the aggregation index is based on like adjacencies tallied using the single-count method, in which each cell side is counted only once.</td>
</tr>
</tbody>
</table>

Table 4: Definition of landscape metrics

<table>
<thead>
<tr>
<th>Metric</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shannon’s Diversity Index (SHDI)</td>
<td>Shannon’s diversity index is a popular measure of diversity in community ecology, applied here to landscapes. Shannon’s index is somewhat more sensitive to rare patch types than Simpson’s diversity index.</td>
</tr>
<tr>
<td>Contagion Index (CONTAG)</td>
<td>Contagion is inversely related to edge density. When edge density is very low, for example, when a single class occupies a very large percentage of the landscape, contagion is high, and vice versa. In addition, note that contagion is affected by both the dispersion and interspersion of patch types. Low levels of patch type dispersion (i.e., high proportion of like adjacencies) and low levels of patch type interspersion (i.e., inequitable distribution of pairwise adjacencies) result in high contagion, and vice versa.</td>
</tr>
<tr>
<td>Landscape Shape Index</td>
<td>Landscape shape index provides a simple measure of class aggregation or clumpiness and, as such, is very similar to the Aggregation index. The differences lie in whether aggregation is measured via class edge (or perimeter) surfaces or via internal like adjacencies (as in AI). Since these surface counts are inversely related to each other (i.e., holding area constant, as the perimeter count increases, the internal adjacency count must decrease, and vice versa), these metrics largely measure the same thing.</td>
</tr>
</tbody>
</table>
Table 5: Retained predictor variables from the univariate analyses

<table>
<thead>
<tr>
<th>Variable</th>
<th>Variable Category</th>
<th>Description</th>
<th>AIC</th>
<th>Δ AIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>CNT_2000m_5</td>
<td>Habitat</td>
<td>Proportion of clear-cut cover within 2 km radius</td>
<td>186.67</td>
<td>0</td>
</tr>
<tr>
<td>NEAR_DIST_PRIMARY</td>
<td>Anthropogenic</td>
<td>Euclidean distance to primary roads</td>
<td>204.77</td>
<td>18.1</td>
</tr>
<tr>
<td>Metric_3000_Al_Devolved_herbaceous</td>
<td>Habitat</td>
<td>Aggregation index of developed herbaceous cover at 3 km</td>
<td>208.72</td>
<td>22.05</td>
</tr>
<tr>
<td>Hunting</td>
<td>Hunting</td>
<td>Index of white-tailed deer hunting intensity</td>
<td>209.09</td>
<td>22.42</td>
</tr>
<tr>
<td>ORV_Dist_Decay_540</td>
<td>Anthropogenic</td>
<td>Decay distance (1-e540/d) to off-road vehicle roads</td>
<td>209.43</td>
<td>22.76</td>
</tr>
<tr>
<td>CNT_270m_9</td>
<td>Habitat</td>
<td>Proportion of Atlantic coastal plain forest cover within 270 m radius</td>
<td>210.4</td>
<td>23.73</td>
</tr>
<tr>
<td>Water bodies_NEAR_DIST</td>
<td>Abiotic</td>
<td>Euclidean distance to nearest water body</td>
<td>211.37</td>
<td>24.7</td>
</tr>
<tr>
<td>Roads_540m_density</td>
<td>Anthropogenic</td>
<td>Density of all roads within 540 m radius</td>
<td>211.87</td>
<td>25.2</td>
</tr>
<tr>
<td>Metric_3000_Al_Dry_Pine_Forest</td>
<td>Habitat</td>
<td>Aggregation index of dry-pine forest cover at 3 km</td>
<td>212.36</td>
<td>25.69</td>
</tr>
<tr>
<td>CNT_2000m_7</td>
<td>Habitat</td>
<td>Proportion of dry-mesic oak cover within 2 km radius</td>
<td>212.88</td>
<td>26.21</td>
</tr>
<tr>
<td>CNT_540m_4</td>
<td>Habitat</td>
<td>Proportion of dry-pine forest cover within 540 m radius</td>
<td>213.21</td>
<td>26.54</td>
</tr>
<tr>
<td>Local_Dist_Decay_90</td>
<td>Anthropogenic</td>
<td>Decay distance (1-e90/d) to local roads</td>
<td>213.39</td>
<td>26.72</td>
</tr>
<tr>
<td>CNT_3000m_6</td>
<td>Abiotic</td>
<td>Proportion of open water cover within 3 km radius</td>
<td>213.94</td>
<td>27.27</td>
</tr>
<tr>
<td>NEAR_DIST_TRAIN</td>
<td>Anthropogenic</td>
<td>Euclidean distance to train tracks</td>
<td>214.55</td>
<td>27.88</td>
</tr>
<tr>
<td>CNT_270m_8</td>
<td>Habitat</td>
<td>Proportion of floodplain forests within 270 m radius</td>
<td>216.26</td>
<td>29.59</td>
</tr>
<tr>
<td>DEM_10m</td>
<td>Abiotic</td>
<td>Elevation at a 10m resolution</td>
<td>216.51</td>
<td>29.84</td>
</tr>
<tr>
<td>Secondary_Dist_Decay_270</td>
<td>Anthropogenic</td>
<td>Decay distance (1-e270/d) to secondary roads</td>
<td>216.6</td>
<td>29.93</td>
</tr>
<tr>
<td>Hydro_density_3000</td>
<td>Abiotic</td>
<td>Density of hydrological features within 3 km</td>
<td>216.75</td>
<td>30.08</td>
</tr>
<tr>
<td>CNT_3000m_2</td>
<td>Habitat</td>
<td>Proportion of salt water wetland cover within 3 km radius</td>
<td>217.43</td>
<td>30.76</td>
</tr>
<tr>
<td>Metric_2000_Al_Atlantic_coastal_plain</td>
<td>Habitat</td>
<td>Aggregation index of Atlantic coastal plain forest cover at 2 km</td>
<td>218.19</td>
<td>31.52</td>
</tr>
<tr>
<td>MEAN_TRAFFIC_540</td>
<td>Anthropogenic</td>
<td>Average traffic within 540 m radius</td>
<td>218.48</td>
<td>31.81</td>
</tr>
<tr>
<td>CNT_1000m_3</td>
<td>Habitat</td>
<td>Proportion of developed land cover within 3 km radius</td>
<td>220.35</td>
<td>33.68</td>
</tr>
<tr>
<td>TRI_270</td>
<td>Abiotic</td>
<td>Terrain roughness index at 270 m radius</td>
<td>220.97</td>
<td>34.3</td>
</tr>
<tr>
<td>Canopy_3000</td>
<td>Habitat</td>
<td>Canopy cover within 3 km radius</td>
<td>221.01</td>
<td>34.34</td>
</tr>
<tr>
<td>Focal_Flow</td>
<td>Abiotic</td>
<td>Index of hydrological flow</td>
<td>221.28</td>
<td>34.61</td>
</tr>
<tr>
<td>Streams_Decay_90</td>
<td>Abiotic</td>
<td>Decay distance (1-e90/d) to streams</td>
<td>221.38</td>
<td>34.71</td>
</tr>
<tr>
<td>AWC_100cm</td>
<td>Abiotic</td>
<td>Available water content at 100 cm</td>
<td>221.53</td>
<td>34.86</td>
</tr>
<tr>
<td>soil_pH</td>
<td>Abiotic</td>
<td>Soil pH</td>
<td>221.55</td>
<td>34.88</td>
</tr>
</tbody>
</table>
Table 6: Top ten combined abiotic sub-models.

<table>
<thead>
<tr>
<th>Rank</th>
<th>Combined Abiotic Variables</th>
<th>AIC</th>
<th>Δ AIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Water bodies_NEAR_DIST</td>
<td>211.37</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>Water bodies_NEAR_DIST+ Focal_Flow</td>
<td>212.35</td>
<td>0.98</td>
</tr>
<tr>
<td>3</td>
<td>DEM_10m+ Water bodies_NEAR_DIST</td>
<td>212.62</td>
<td>1.25</td>
</tr>
<tr>
<td>4</td>
<td>Water bodies_NEAR_DIST+ Streams_Decay_90</td>
<td>212.72</td>
<td>1.35</td>
</tr>
<tr>
<td>5</td>
<td>Water bodies_NEAR_DIST+ CNT_3000m_6</td>
<td>212.84</td>
<td>1.47</td>
</tr>
<tr>
<td>6</td>
<td>Water bodies_NEAR_DIST+ Hydro_density_3000</td>
<td>212.92</td>
<td>1.55</td>
</tr>
<tr>
<td>7</td>
<td>Water bodies_NEAR_DIST+ TRI_270</td>
<td>213.24</td>
<td>1.87</td>
</tr>
<tr>
<td>8</td>
<td>Water bodies_NEAR_DIST+ soil_pH</td>
<td>213.28</td>
<td>1.91</td>
</tr>
<tr>
<td>9</td>
<td>Water bodies_NEAR_DIST+ AWC_100cm</td>
<td>213.36</td>
<td>1.99</td>
</tr>
<tr>
<td>10</td>
<td>DEM_10m+ Water bodies_NEAR_DIST+ Focal_Flow</td>
<td>213.68</td>
<td>2.31</td>
</tr>
</tbody>
</table>

Table 7: Top ten combined anthropogenic sub-models.

<table>
<thead>
<tr>
<th>Rank</th>
<th>Combined Anthropogenic Variables</th>
<th>AIC</th>
<th>Δ AIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>NEAR_DIST_PRIMARY+ Local_Dist_Decay_90+ NEAR_DIST_TRAIN+ Secondary_Dist_Decay_270</td>
<td>194.2</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>NEAR_DIST_PRIMARY+ Local_Dist_Decay_90+ ORV_Dist_Decay_540+ NEAR_DIST_TRAIN+ Secondary_Dist_Decay_270</td>
<td>194.73</td>
<td>0.53</td>
</tr>
<tr>
<td>3</td>
<td>NEAR_DIST_PRIMARY+ Roads_540m_density+ Local_Dist_Decay_90+ NEAR_DIST_TRAIN+ Secondary_Dist_Decay_270</td>
<td>195.66</td>
<td>1.46</td>
</tr>
<tr>
<td>4</td>
<td>NEAR_DIST_PRIMARY+ Local_Dist_Decay_90+ ORV_Dist_Decay_540+ NEAR_DIST_TRAIN</td>
<td>196.02</td>
<td>1.82</td>
</tr>
<tr>
<td>5</td>
<td>NEAR_DIST_PRIMARY+ Local_Dist_Decay_90+ NEAR_DIST_TRAIN+ Secondary_Dist_Decay_270+ MEAN_TRAFFIC_540</td>
<td>196.07</td>
<td>1.87</td>
</tr>
<tr>
<td>6</td>
<td>NEAR_DIST_PRIMARY+ Roads_540m_density+ Local_Dist_Decay_90+ ORV_Dist_Decay_540+ NEAR_DIST_TRAIN</td>
<td>196.22</td>
<td>2.02</td>
</tr>
<tr>
<td>7</td>
<td>NEAR_DIST_PRIMARY+ Roads_540m_density+ Local_Dist_Decay_90+ ORV_Dist_Decay_540+ NEAR_DIST_TRAIN+ Secondary_Dist_Decay_270</td>
<td>196.23</td>
<td>2.03</td>
</tr>
<tr>
<td>8</td>
<td>NEAR_DIST_PRIMARY+ Roads_540m_density+ ORV_Dist_Decay_540+ NEAR_DIST_TRAIN</td>
<td>196.38</td>
<td>2.18</td>
</tr>
<tr>
<td>9</td>
<td>NEAR_DIST_PRIMARY+ Roads_540m_density+ Local_Dist_Decay_90+ NEAR_DIST_TRAIN</td>
<td>196.43</td>
<td>2.23</td>
</tr>
<tr>
<td>10</td>
<td>NEAR_DIST_PRIMARY+ Local_Dist_Decay_90+ ORV_Dist_Decay_540+ NEAR_DIST_TRAIN+ Secondary_Dist_Decay_270+ MEAN_TRAFFIC_540</td>
<td>196.49</td>
<td>2.29</td>
</tr>
</tbody>
</table>
Table 8: Top ten combined habitat sub-models.

<table>
<thead>
<tr>
<th>Rank</th>
<th>Combined Habitat Variables</th>
<th>AIC</th>
<th>Δ AIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>CNT_2000m_5+ Metric_3000_AI_Developed_herbaceous+ CNT_270m_9+ CNT_2000m_7+ Metric_2000_AI_ATLANTIC_coastal_plain</td>
<td>175.82</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>CNT_2000m_5+ Metric_3000_AI_Developed_herbaceous+ CNT_270m_9+ Metric_2000_AI_ATLANTIC_coastal_plain</td>
<td>175.94</td>
<td>0.12</td>
</tr>
<tr>
<td>3</td>
<td>CNT_2000m_5+ Metric_3000_AI_Developed_herbaceous+ CNT_270m_9+ CNT_3000m_2+ Metric_2000_AI_ATLANTIC_coastal_plain</td>
<td>176</td>
<td>0.18</td>
</tr>
<tr>
<td>4</td>
<td>CNT_2000m_5+ Metric_3000_AI_Developed_herbaceous+ CNT_270m_9+</td>
<td>176.17</td>
<td>0.35</td>
</tr>
<tr>
<td>5</td>
<td>CNT_2000m_5+ Metric_3000_AI_Developed_herbaceous+ CNT_270m_9+ Canopy_3000+ CNT_540m_4+ CNT_270m_8+ Metric_2000_AI_ATLANTIC_coastal_plain</td>
<td>176.18</td>
<td>0.36</td>
</tr>
<tr>
<td>6</td>
<td>CNT_2000m_5+ Metric_3000_AI_Developed_herbaceous+ CNT_270m_9+ CNT_540m_4+ CNT_270m_8+ Metric_2000_AI_ATLANTIC_coastal_plain</td>
<td>176.33</td>
<td>0.51</td>
</tr>
<tr>
<td>7</td>
<td>CNT_2000m_5+ Metric_3000_AI_Developed_herbaceous+ CNT_270m_9+ Canopy_3000 Metric_2000_AI_ATLANTIC_coastal_plain</td>
<td>176.6</td>
<td>0.78</td>
</tr>
<tr>
<td>8</td>
<td>CNT_2000m_5+ Metric_3000_AI_Developed_herbaceous+ CNT_270m_9+ Canopy_3000+ CNT_3000m_2 Metric_2000_AI_ATLANTIC_coastal_plain</td>
<td>176.62</td>
<td>0.8</td>
</tr>
<tr>
<td>9</td>
<td>CNT_2000m_5+ Metric_3000_AI_Developed_herbaceous+ CNT_270m_9+ Canopy_3000+ CNT_3000m_2+ Metric_2000_AI_ATLANTIC_coastal_plain</td>
<td>176.65</td>
<td>0.83</td>
</tr>
<tr>
<td>10</td>
<td>CNT_2000m_5+ Metric_3000_AI_Developed_herbaceous+ CNT_270m_9+ CNT_270m_8+ Metric_2000_AI_ATLANTIC_coastal_plain</td>
<td>176.77</td>
<td>0.95</td>
</tr>
</tbody>
</table>
Table 9: Final candidate models up to 2 ΔAIC. Models are ordered according to Δ AIC. K indicates the number of parameters in the model. The Akaike weights (W) represents the likelihood of the model being the best out of all models evaluated (n=311). This set of models within 2 Δ AIC reaches a cumulative Akaike weight of 0.2118. The null model is provided as reference in the last row.

<table>
<thead>
<tr>
<th>Variables</th>
<th>AIC</th>
<th>Δ AIC</th>
<th>K</th>
<th>W/</th>
<th>Cum W/</th>
</tr>
</thead>
<tbody>
<tr>
<td>CNT_2000m_5+ Water bodies_NEAR_DIST+ Metric_3000_AI_Developed_herbaceous+</td>
<td>174.32</td>
<td>0</td>
<td>7</td>
<td>0.02159</td>
<td>0.021595</td>
</tr>
<tr>
<td>CNT_2000m_5+ Water bodies_NEAR_DIST+ Metric_3000_AI_Developed_herbaceous+</td>
<td>174.65</td>
<td>0.33</td>
<td>5</td>
<td>0.01831</td>
<td>0.039905</td>
</tr>
<tr>
<td>CNT_2000m_5+ Water bodies_NEAR_DIST+ Metric_3000_AI_Developed_herbaceous+</td>
<td>175.00</td>
<td>0.68</td>
<td>6</td>
<td>0.01537</td>
<td>0.055276</td>
</tr>
<tr>
<td>CNT_2000m_5+ Water bodies_NEAR_DIST+ Metric_3000_AI_Developed_herbaceous+</td>
<td>175.33</td>
<td>1.01</td>
<td>7</td>
<td>0.01303</td>
<td>0.068308</td>
</tr>
<tr>
<td>CNT_2000m_5+ Water bodies_NEAR_DIST+ Metric_3000_AI_Developed_herbaceous+</td>
<td>175.57</td>
<td>1.25</td>
<td>7</td>
<td>0.01156</td>
<td>0.079867</td>
</tr>
<tr>
<td>CNT_2000m_5+ Water bodies_NEAR_DIST+ Metric_3000_AI_Developed_herbaceous+</td>
<td>175.82</td>
<td>1.5</td>
<td>6</td>
<td>0.01020</td>
<td>0.090068</td>
</tr>
<tr>
<td>CNT_2000m_5+ Water bodies_NEAR_DIST+ Metric_3000_AI_Developed_herbaceous+</td>
<td>175.85</td>
<td>1.53</td>
<td>6</td>
<td>0.01005</td>
<td>0.100157</td>
</tr>
<tr>
<td>CNT_2000m_5+ Water bodies_NEAR_DIST+ Metric_3000_AI_Developed_herbaceous+</td>
<td>175.94</td>
<td>1.62</td>
<td>4</td>
<td>0.00960</td>
<td>0.109723</td>
</tr>
<tr>
<td>CNT_2000m_5+ Water bodies_NEAR_DIST+ Metric_3000_AI_Developed_herbaceous+</td>
<td>176.01</td>
<td>1.69</td>
<td>4</td>
<td>0.00928</td>
<td>0.119000</td>
</tr>
<tr>
<td>CNT_2000m_5+ Water bodies_NEAR_DIST+ Metric_3000_AI_Developed_herbaceous+</td>
<td>176.04</td>
<td>1.72</td>
<td>6</td>
<td>0.00914</td>
<td>0.128138</td>
</tr>
<tr>
<td>CNT_2000m_5+ Water bodies_NEAR_DIST+ CNT_270m_9+ Secondary_Dist_Decay_270</td>
<td>176.14</td>
<td>1.82</td>
<td>4</td>
<td>0.00869</td>
<td>0.136830</td>
</tr>
<tr>
<td>CNT_2000m_5+ Water bodies_NEAR_DIST+ CNT_270m_9+ CNT_2000m_7+ Metric_2000_AI</td>
<td>176.16</td>
<td>1.84</td>
<td>6</td>
<td>0.00861</td>
<td>0.145436</td>
</tr>
<tr>
<td>CNT_2000m_5+ Water bodies_NEAR_DIST+ CNT_270m_9+ CNT_2000m_5+ Metric_2000_AI</td>
<td>176.17</td>
<td>1.85</td>
<td>3</td>
<td>0.00856</td>
<td>0.153999</td>
</tr>
<tr>
<td>CNT_2000m_5+ Water bodies_NEAR_DIST+ CNT_270m_9+ CNT_2000m_5+ Metric_2000_AI</td>
<td>176.18</td>
<td>1.86</td>
<td>5</td>
<td>0.00852</td>
<td>0.162519</td>
</tr>
<tr>
<td>CNT_2000m_5+ Water bodies_NEAR_DIST+ CNT_270m_9+ Metric_2000_AI_Developed_herbaceous+</td>
<td>176.19</td>
<td>1.87</td>
<td>5</td>
<td>0.00848</td>
<td>0.170997</td>
</tr>
<tr>
<td>CNT_2000m_5+ Water bodies_NEAR_DIST+ Metric_3000_AI_Developed_herbaceous+</td>
<td>176.19</td>
<td>1.87</td>
<td>6</td>
<td>0.00848</td>
<td>0.179475</td>
</tr>
<tr>
<td>CNT_2000m_5+ NEAR_DIST+ Water bodies_NEAR_DIST+ Metric_3000_AI_Developed_herbaceous+</td>
<td>176.24</td>
<td>1.92</td>
<td>6</td>
<td>0.00827</td>
<td>0.187744</td>
</tr>
<tr>
<td>CNT_2000m_5+ NEAR_DIST+ Water bodies_NEAR_DIST+ Local_Dist_Decay_90+ Metric_2000_AI</td>
<td>176.26</td>
<td>1.94</td>
<td>7</td>
<td>0.00819</td>
<td>0.195930</td>
</tr>
<tr>
<td>CNT_2000m_5+ NEAR_DIST+ Water bodies_NEAR_DIST+ Local_Dist_Decay_90+ Metric_2000_AI</td>
<td>176.29</td>
<td>1.97</td>
<td>5</td>
<td>0.00807</td>
<td>0.203994</td>
</tr>
<tr>
<td>CNT_2000m_5+ NEAR_DIST+ Water bodies_NEAR_DIST+ Metric_3000_AI_Developed_herbaceous+</td>
<td>176.36</td>
<td>2.04</td>
<td>6</td>
<td>0.00779</td>
<td>0.211781</td>
</tr>
<tr>
<td>Null (Intercept)</td>
<td>219.56</td>
<td>45.24</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
</tbody>
</table>
Table 10: Final model variables with β-value and model averaged β-values. The model averaged β-values are derived from models with AIC weights summing to 0.9 (n=311).

<table>
<thead>
<tr>
<th>Variable</th>
<th>β-Value</th>
<th>Averaged β-Value</th>
<th>Variable Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Intercept)</td>
<td>0.4477</td>
<td>3.3519</td>
<td>Intercept</td>
</tr>
<tr>
<td>CNT_2000m_5</td>
<td>-38.0052</td>
<td>-284.5168</td>
<td>Proportion of clear-cut cover within 2 km radius</td>
</tr>
<tr>
<td>CNT_2000m_7</td>
<td>0.9720</td>
<td>7.2768</td>
<td>Proportion of dry-mesic oak cover within 2 km radius</td>
</tr>
<tr>
<td>CNT_270m_9</td>
<td>3.5659</td>
<td>26.6952</td>
<td>Proportion of Atlantic coastal plain forest cover within 270 m radius</td>
</tr>
<tr>
<td>Hunting</td>
<td>-0.0449</td>
<td>-0.3363</td>
<td>Hunting intensity</td>
</tr>
<tr>
<td>Local_Dist_Decay_90</td>
<td>-0.2793</td>
<td>-2.0909</td>
<td>Decay distance $(1-e^{\frac{d}{270}})$ to local roads</td>
</tr>
<tr>
<td>Metric_2000_AI_Atlantic_coastal_plain</td>
<td>0.0338</td>
<td>0.2532</td>
<td>Aggregation index of Atlantic coastal plain forest cover at 2 km</td>
</tr>
<tr>
<td>Metric_3000_AI_Developed_herbaceous</td>
<td>-0.0222</td>
<td>-0.1662</td>
<td>Aggregation index of developed herbaceous cover at 3 km</td>
</tr>
<tr>
<td>NEAR_DIST_PRIMARY</td>
<td>0.0000</td>
<td>-0.0001</td>
<td>Euclidean distance to primary roads</td>
</tr>
<tr>
<td>NEAR_DIST_TRAIN</td>
<td>0.0000</td>
<td>0.0001</td>
<td>Euclidean distance to train tracks</td>
</tr>
<tr>
<td>Secondary_Dist_Decay_270</td>
<td>-0.5753</td>
<td>-4.3070</td>
<td>Decay distance $(1-e^{\frac{d}{270}})$ to secondary roads</td>
</tr>
<tr>
<td>Water bodies_NEAR_DIST</td>
<td>-0.00022</td>
<td>-0.0000052</td>
<td>Euclidean distance to water bodies</td>
</tr>
</tbody>
</table>

Table 11: Contingency table for the expected occupancy based on the model prediction versus the observed occupancy from the validation survey points.

<table>
<thead>
<tr>
<th>Contingency Table</th>
<th>Observed Unoccupied</th>
<th>Observed Occupied</th>
</tr>
</thead>
<tbody>
<tr>
<td>Expected Unoccupied</td>
<td>18</td>
<td>6</td>
</tr>
<tr>
<td>Expected Occupied</td>
<td>9</td>
<td>8</td>
</tr>
</tbody>
</table>
Figure 1: Distributions of study plots and model validation plots in relation to roads within the study area.
Figure 2: Regional map with the study area in context.
Figure 3: Distribution of land cover types within the study area as classified by the SE Gap Analysis project.
Figure 4: Distribution of clear-cut lands within a 2-km radius in relation to study plots. The upper insert depicts the distribution of clear-cut lands as classified by the SE Gap Analysis project.
Figure 5: Distribution of Atlantic coastal plain mesic forest within a 270-m radius in relation to study plots. The upper insert depicts the distribution of Atlantic coastal plain mesic forest as classified by the SE Gap Analysis project.
Figure 6: The distribution of dry-mesic oak forest within a 2-km radius in relation to study plots. The upper insert depicts the distribution of dry-mesic oak forest as classified by the SE Gap Analysis project.
Figure 7: The continuous probability model output based on applying the beta-values for all landscape variables predicted by the model output. The insert depicts a close extent where the patches have a range of values for predicted invasion.
Figure 8: The binary model output reclassified from the continuous probability model output with a cut-off of 0.999833. The insert depicts the same extent as in Figure 7, but as a binary classification.