

HABITAT MODELING OF INVASIVE PLANTS IN  
GREAT SMOKY MOUNTAINS NATIONAL PARK

by  
Michelle Ann Cawley

A thesis submitted to the faculty of the University of North Carolina  
at Chapel Hill in partial fulfillment of the requirements for the degree  
of Masters of Arts in the Curriculum in Ecology.

Chapel Hill  
2003

Approved by:

---

Advisor: Peter S. White

---

Reader: Robert K. Peet





©2003  
Michelle Ann Cawley  
ALL RIGHTS RESERVED

## ABSTRACT

Michelle Ann Cawley: Habitat modeling of invasive plants in  
Great Smoky Mountains National Park.  
(Under the direction of Peter S. White.)

Using Mahalanobis distance, I modeled the spread of non-native plants in Great Smoky Mountains National Park (GRSM). I asked whether the distribution of 13 exotics was correlated with landscape-level variables including elevation, distance from corridors and habitat variables. I found that invasion sites can be distinguished and that spatial signals were stronger than environmental signals at this stage of the invasion process. The habitat model may be useful in focusing monitoring on sites that resemble current invasion sites.

I also surveyed *Paulownia tomentosa* along roads near GRSM to determine whether this species has a patchy or continuous distribution and whether all roadside habitat is suitable for invasion by this species. I found evidence of a patchy mode of spread and that roadside habitat with *Paulownia* can be distinguished from all roadside habitat. The patchy distribution of suitable habitat suggests that existing patches will not coalesce over time.

## ACKNOWLEDGEMENTS

I would like to thank my advisor Dr. Peter White, who once said: “I live for my graduate students” and never showed me anything to the contrary. He is a patient mentor, who infused all of our discussions with enthusiasm, imagination, and humor. Many thanks to Dr. Robert Peet whose insights into making my research more generally applicable were invaluable and to Dr. Dean Urban who originally inspired me to pursue a spatial approach to ecological research, which is the best compliment I can think to offer a teacher. My committee helped me get from point A to point B, but without Joan Tilghman, Dr. Richard Bilsborrow, and Dr. Seth Reice I would never have found point A, for that I am most grateful.

This research would not have been possible without the help of Dr. Jack Weiss. As a scientist, a statistician, and a patient teacher he embodies everything a graduate student needs. I credit him with the fact that meetings that spanned whole afternoons discussing various statistical approaches were truly enjoyable.

I am grateful to Kristine Johnson, Scott Kichman and Richard Schulz from GRSM who provided the database for this research and were always willing to offer interpretive support. I am also indebted to Evan Hammer, who answered countless GIS-based questions and wrote a script for use in ArcView. Funding support from the Ecology curriculum, Biology Department, Southern Research Circle, and GRSM made it possible for me to take time to develop fully my research interests.

Cottie Pasternak and Karen Henry were always able to come up with what I needed, whether it was funding, a laptop, a GPS unit, or guidance through the arcane bureaucracy of graduate school. I feel lucky to have had them on my side.

My friends and family never seemed to waver in their confidence of me. I am especially grateful to Eileen Vandenburg whose friendship and insights as a scientist I treasure. Elizabeth Gregory made my first research trip memorable; I am fortunate to have a friend outside of graduate school that has such a genuine interest in my research. Many thanks to Jamila Batchelder, Wade Hoiland, Brad Lamphere, David Vandermast, and Reba Vidra; they made the whole journey more fun. I also want to thank my fellow plant ecologists and Curriculum students who shared ideas, road trips, and hikes with me; I feel privileged to consider myself one of your peers.

To those who knew me before my first camping trip (Mom, John, Michael, Kara, and Jardena), thank you for accepting how I have changed.

Finally, to my husband Tom who cheerfully adjusted to what began as a temporary interruption of our lives, but turned into a new path. I cannot imagine a better friend to meander through life with.

## TABLE OF CONTENTS

	Page
LIST OF TABLES.....	ix
LIST OF FIGURES.....	xi
LIST OF ABBREVIATIONS .....	xiii
 <b>Chapter</b>	
I. Introduction.....	1
II. Habitat modeling of invasive plants in Great Smoky Mountains National Park using Mahalanobis distance.....	4
2.1 Introduction.....	4
2.2 Methods.....	6
2.21 Study Site.....	6
2.22 History of Invasive Plants in GRSM.....	6
2.23 Dataset.....	7
2.24 Univariate Analyses.....	7
2.25 Multivariate Model Formation.....	9
2.26 Role of Individual Variables.....	10
2.27 Fit of the Habitat Model.....	11
2.28 Logistic Regression.....	12
2.29 Variations in the Model by Species type.....	13

2.2.10 Accessibility of Known Invasion Sites.....	13
2.3 Results.....	14
2.3.1 Univariate Randomization Tests.....	14
2.3.2 Distribution of Invasion-Prone Habitat.....	15
2.3.3 Fit of Habitat Model.....	15
2.3.4 Role of Individual Variables.....	16
2.3.5 Role of Elevation.....	17
2.3.6 Role of Distance from Roads and Streams.....	17
2.3.7 Role of Disturbance History.....	18
2.3.8 Variation of Environmental Variables across Quartiles of $D^2$ .....	19
2.3.9 Patterns among Subgroups.....	20
2.3.10 Validity of Invasive Plant Dataset.....	21
2.4 Discussion.....	21
2.4.1 Modeling Habitat using Landscape-Level Variables.....	21
2.4.2 Fit of the Habitat Model.....	22
2.4.3 Variables that Affect Distribution of Suitable Habitat.....	23
2.4.4 Role of Elevation.....	24
2.4.5 Role of Distance to Roads.....	24
2.4.6 Role of Distance to Streams.....	25
2.4.7 Role of Disturbance History.....	25
2.4.8 Validity of Invasive Plant Dataset.....	27
2.4.9 Conclusion.....	28
2.5 Future Research.....	29

III. Invasion pattern of <i>Paulownia tomentosa</i> along roads surrounding Great Smoky Mountains National Park.....	81
3.1 Introduction.....	81
3.2 Methods.....	82
3.2.1 Study Area.....	82
3.2.2 Study Species.....	83
3.2.3 Data Set.....	84
3.2.4 Univariate Randomization Tests.....	85
3.2.5 Evaluation of Field Variables.....	86
3.2.6 Habitat Suitability Model.....	87
3.2.7 Detecting Pattern of Spread.....	88
3.3 Results.....	88
3.3.1 Distribution of <i>Paulownia</i> .....	88
3.3.2 Habitat Model for <i>Paulownia</i> .....	89
3.3.3 Invasion Pattern of <i>Paulownia</i> .....	89
3.4 Discussion.....	90
3.4.1 Drought tolerance in <i>Paulownia</i> .....	90
3.4.2 Soil Preferences of <i>Paulownia</i> .....	91
3.4.3 Pattern of Spread.....	91
3.4.4 Relevance of Habitat Suitability Model.....	92
3.4.5 Microsite Conditions along Roads.....	93
3.5 Future research.....	93
REFERENCES .....	115

## LIST OF TABLES

		Page
Table 2.1	Summary information for target exotic species in Great Smoky Mountains National Park.....	31
Table 2.2	Variables used to characterize habitat of target exotic species in Great Smoky Mountains National Park (GRSM).....	32
Table 2.3	Correlation matrix of independent variables used to model habitat of target exotic species .....	33
Table 2.4	Results of randomization tests on full dataset ( <i>i.e.</i> , all 13 target exotic species) allowing random points to fall anywhere in the Park.....	34
Table 2.5	Results of restricted randomization tests on full dataset ( <i>i.e.</i> , all 13 target exotic species) only including random points less than 1,000 meters from roads.....	35
Table 2.6	Range of $D^2$ values for all pixels in GRSM and for sites with target exotic species.....	36
Table 2.7	Logistic regression model for predicting the occurrence of target exotic species in Deep Creek Catchment.....	37
Table 2.8	Logistic regression model for predicting the occurrence of <i>Microstegium vimineum</i> in Deep Creek Catchment.....	37
Table 2.9	Logistic regression model for predicting the occurrence of target exotic species in Great Smoky Mountains National Park.....	38
Table 2.10	Table 2.10. Results of 10 iterations of stepwise logistic regression.....	39
Table 2.11	Table 2.11. Effect of removing variables from final logistic regression model.....	40
Table 2.12	Summary of analyses where each variable showed an effect.....	41
Table 2.13	Range of $D^2$ values for all pixels ( <i>i.e.</i> , 2,285,376) in Great Smoky Mountains National Park when omitting variables from the Mahalanobis distance habitat model.....	42

Table 2.14	Range of $D^2$ values for invasion sites (i.e., 345) when omitting variables from the Mahalanobis distance habitat model.....	42
Table 2.15	Distribution of $D^2$ values by disturbance history.....	43
Table 2.16	Results of randomization tests for evergreen and semi-evergreen target exotic species.....	44
Table 2.17	Results of randomization tests for deciduous target exotic species.....	45
Table 2.18	Results of randomization tests for target exotics that spread primarily by seed.....	46
Table 2.19	Results of randomization tests for target exotic species that have primarily vegetative spread.....	47
Table 2.20	Variation in $D^2$ values by distance from road.....	48
Table 3.1	Variables used to characterize habitat of <i>Paulownia tomentosa</i> along roads near Great Smoky Mountains National Park (GRSM).....	95
Table 3.2	Results of randomization tests for <i>Paulownia tomentosa</i> along roads.....	96
Table 3.3	Two Sample <i>t</i> -test comparing the means for soil depth (cm) between random points and sites with <i>Paulownia tomentosa</i> .....	97
Table 3.4	Two Sample <i>t</i> -test comparing the means for slope angle (°) between random points and sites with <i>Paulownia tomentosa</i> .....	97
Table 3.5	Range of $D^2$ values for all sites within 500 meters of roads and sites with <i>Paulownia tomentosa</i> .....	98

## LIST OF FIGURES

		Page
Figure 2.1	Location of Great Smoky Mountains National Park, North Carolina/Tennessee.....	49
Figure 2.2	Pattern of human history of disturbance in Great Smoky Mountains National Park.....	51
Figure 2.3	Distribution of roads near Great Smoky Mountains National Park.....	53
Figure 2.4	Distribution of target exotic species in Great Smoky Mountains National Park.....	55
Figure 2.5	Distribution of suitable habitat for target exotic species in Great Smoky Mountains National Park.....	57
Figure 2.6	Results of testing model against presence/absence of target exotic species in Deep Creek Catchment.....	59
Figure 2.7	Results of testing model against presence/absence of <i>Microstegium vimineum</i> in Deep Creek Catchment.....	61
Figure 2.8	Dice-Leraas diagrams showing change in independent variables across quartiles of $D^2$ .....	63
Figure 2.9	Change in suitable habitat when omitting elevation from the Mahalanobis distance habitat model.....	74
Figure 2.10	Change in suitable habitat when omitting distance from roads from the Mahalanobis distance habitat model.....	76
Figure 2.11	Distribution of percentage of target exotic species by percentage of landscape in each category of disturbance history.....	78
Figure 2.12	Distribution of disturbance history by mean elevation.....	80
Figure 3.1	Distribution of roads surrounding Great Smoky Mountains National Park.....	99
Figure 3.2	Conceptual models of invasion patterns along corridors.....	101
Figure 3.3	Distribution of sites colonized by <i>Paulownia tomentosa</i> along roads near Great Smoky Mountains National Park.....	103

Figure 3.4	Box-plot comparing the mean of soil depth at random sites and sites with <i>Paulownia tomentosa</i> .....	105
Figure 3.5	Box-plot comparing the mean of slope angle at random sites and sites with <i>Paulownia tomentosa</i> .....	107
Figure 3.6	Distribution of roadside habitat similar to sites invaded by <i>Paulownia tomentosa</i> ( $D^2 < 125$ ).....	109
Figure 3.7	Distribution of roadside habitat similar to sites invaded by <i>Paulownia tomentosa</i> ( $D^2 < 60$ ).....	111
Figure 3.8	Cumulative frequency curve of nearest neighbor distance (NND) between patches with <i>Paulownia tomentosa</i> and random sites and nearest <i>Paulownia</i> patch.....	113

## LIST OF ABBREVIATIONS

BTA	Beers (1966) transformed aspect
CDF	Cumulative density function
$D^2$	Mahalanobis Distance statistic
DEM	Digital Elevation Model
DH	Disturbance History
DR	Distance to Roads
DS	Distance to Streams
GIS	Geographic Information Systems
GPS	Global Positioning System
GRSM	Great Smoky Mountains National Park
LFI	Landform Index
NND	Nearest Neighbor Distance
RSP	Relative Slope Position
SI	Solar Incidence
SWI	Shannon-Weiner Index of Topographic Complexity
TCI	Topographic Convergence Index
TRMI	Total Relative Moisture Index
TSI	Terrain Shape Index

## Chapter I: Introduction

Given the conservation implications, improving the ability to predict which species will become invasive is one of the more active areas of invasion ecology. Early approaches like Baker's (1974) list of 12 characters associated with an ideal weed are the basis for more recent efforts to predict which species will become invasive (Burke and Grime 1996, Rejmanek and Richardson 1996, Reichard and Hamilton 1997). However, in most landscapes invasions are already in process and methods to improve control efforts of existing invaders are needed.

Burke and Grime (1996) noted that the colonizer and the colonized are partners in the process. This emphasizes the need to consider whether certain types of communities are more readily invaded; this debate has been centered on whether species-rich or species-poor communities are more invulnerable (Tilman 1997, Stohlgren *et al.* 1999a, Stohlgren *et al.* 1999b). However, irrespective of species diversity, certain environmental and landscape features (*e.g.*, elevation, corridors, moisture, and sunlight) may also be associated with invulnerability at the landscape scale.

Few studies have investigated invasion patterns at the landscape scale (Wiser *et al.* 1998, Parendes 1997). Evaluating patterns of invasion at this scale may give more insight into areas that may be vulnerable to future invasions. Further, studies of invasion at the landscape scale allow researchers to pose spatially explicit questions. Cadenasso and Pickett (2001) noted that few studies have investigated spatially explicit dynamics of the invasion

process. They argued that understanding the spatial processes underlying invasion patterns would aid managers in maintaining the integrity of biological communities.

Non-native plants represent 23 percent of the plant species in Great Smoky Mountains National Park (GRSM) (K. Johnson, GRSM, *pers. comm.*); of these two percent or approximately 50 species are target species (*i.e.*, invasive) and are actively controlled by NPS personnel. Much research has concluded that invasive species threaten native biodiversity (Zalba *et al.* 2000, Lodge *et al.* 1998, Rundel and Sturmer 1998). Exotic invasions occur globally. However, species richness increases from the poles to the tropics (Brown and Gibson 1983); therefore, it is plausible that exotic species richness may also be higher with decreasing latitude. In North America, some of the most notorious examples of invasion have occurred in the South, including *Pueraria lobata* (Kudzu), *Eichhornia crassipes* (Water hyacinth) and *Melaleuca leucadendron* (Melaleuca). At a minimum, Southern ecosystems have more risk associated with exotic invasions, given their higher native species richness. GRSM staff has responded to the growing threat of non-native plants by developing a system to map, monitor, and control invasive plants beginning in the mid-1980s.

In Chapter 2 of this thesis I derive a habitat model for invasive species in GRSM. I use the locations of 13 of the most common non-native species based on Park records to develop a Mahalanobis habitat model. Further, I model subsets of this group of 13 species to determine whether variations driven by each species' unique niche are detectable. I compare the results of the Mahalanobis model to a model derived using logistic regression.

*Paulownia tomentosa* is a conspicuous invader that surrounds GRSM along roads. In Chapter 3 of this thesis, I model the habitat of roadside occurrences of *Paulownia* using eight

landscape-level environmental variables and two field variables. Further, although *Paulownia* is very common along roads its distribution appears to be interrupted. I also addressed whether this species has a satellite or continuous pattern of spread along roads surrounding GRSM.

## **Chapter II: Habitat modeling of invasive plants in Great Smoky Mountains National Park using Mahalanobis distance**

### **2.1 Introduction**

Controlling non-native plants is a necessary component of any effort to preserve natural habitats. Modeling potentially suitable habitat for exotic plants across a landscape may allow land managers to focus their control efforts and shape land-use planning efforts. Further, examining invasion at the landscape scale offers a chance to understand the spatial processes driving invasions (*e.g.*, dispersal, the role of connectivity) and how suitable habitat is distributed.

Invasibility has been associated with landscape and environmental features including streams (Williams 1993a, DeFerrari and Naiman 1994, Planty-Tabacchi *et al.* 1996, Hood and Naiman 2000), forest edges (Brothers and Spingarn 1992, Hester and Hobbs 1992, Hobbs and Huenneke 1992, Tyser and Worley 1992, Zink *et al.* 1995, Parendes 1997, Morgan 1998, Parendes and Jones 2000, Meekins and McCarthy 2001), lower elevations (Brown and Carter 1998, Wiser *et al.* 1998, Lesica and Miles 1999, Hood and Naiman 2000), higher moisture availability (Woods 1993, Pysek and Prach 1993, D'Antonio 1993, DeFerrari and Naiman 1994, Wiser *et al.* 1998, Brown and Carter 1998, Lesica and Miles 1999, Hood and Naiman 2000, Meekins and McCarthy 2001), soil fertility (Huenneke *et al.* 1990, Woods 1993, Pysek and Prach 1993, Wiser *et al.* 1998, Morgan 1998, Stohlgren *et al.* 1999a, Stohlgren *et al.* 1999b, Thompson *et al.* 2001) and disturbance (Denslow 1980, Hobbs and Huenneke 1992, Bergelson *et al.* 1993, Morgan 1998, Nuzzo 1999, Knapp and Canham

2000). In this study, I examined the relationship between historic invasions in Great Smoky Mountains National Park and landscape characteristics.

As a consequence of the availability of spatial data at the landscape scale and technological advances for analysis (*e.g.*, Geographic Information Systems (GIS)), much effort has gone into modeling species habitat, which in turn has resulted in a wealth of literature on the subject (Johnson 1990, Clark *et al.* 1993, Wilds 1996, Knick and Dyer 1997, Corsi *et al.* 1999, Boetsch 2000, Rouget *et al.* 2001, van Manen *et al.* 2002, Scott, P.J. *et al.* 2002). A tradeoff exists between modeling habitat using a species-specific approach versus a group of species that behave similarly. A species-specific approach to modeling habitat allows the nuances of a given species' niche to be examined; however, using several species to model habitat may result in a model that is more generally applicable.

In this project, I used a database on historic sites for 13 of the most common invasive alien plants in the National Park to ask the following questions:

1. What are the roles of spatial variables (*e.g.*, disturbance history and distance from corridors) and environmental variables (particularly elevation, moisture availability and solar incidence) in the distribution of invasive plants?
2. What percent of the Park's landscape is particularly vulnerable to invasion?
3. Do different groups of invasive species (deciduous, evergreen, vegetative spreaders, seed spreaders) vary in their pattern of invasion across the landscape?
4. Is the model of invasion prone habitat general enough to predict sites invaded by *Microstegium vimineum*, a common invader not used to build the model?

## **2.2 Methods**

### 2.2.1 Study Site

Great Smoky Mountains National Park (GRSM) has an area of 206,119 ha and straddles the border between Tennessee and North Carolina in the southern Appalachians (Figure 2.1). Elevations range from 266 to 2,027 m, with the higher elevations being largely confined to the eastern half of the park. Precipitation increases with elevation in the park (Shanks 1954). The park's climate is classified as temperate rainforest at elevations with the highest precipitation (Thornewaite 1948); lower elevations that have less precipitation are classified as humid mesothermal (Shanks 1954). July mean temperatures range from 22° to 14° C and January mean temperatures range from 4° to -2° C (Shanks 1954).

The park is characterized by mountainous terrain with steep slopes, narrow valleys and coves, and fast-flowing streams (King and Stupka 1950). This topographically complex terrain gives rise to a range of habitats (Whittaker 1956), and is therefore, rich in native species diversity. In general native plant species richness decreases with elevation and increases with site moisture (Whittaker 1956).

The main sources of human disturbance in the park have been logging and settlement with approximately two-thirds of the park landscape affected (Figure 2.2) (Pyle 1985, Pyle 1988). Roads also constitute a form of disturbance. Most roads in GRSM are close the park's border; however, Newfoundgap Road dissects the park and several smaller roads traverse the western part of GRSM (Figure 2.3).

### 2.2.2 History of Invasive Plants in GRSM

The Park has an active program to control the spread of the more than 50 non-native plants. This program developed a database that includes spatial data on past and current

invasion sites. Park employees identified invasion sites in the field. Areas near known invasion sites are also searched for the presence of non-native plants (Kristine Johnson, GRSM, *pers. comm.*). The park control program was instituted in the late 1980s at which time invasives were already established at most, if not all, the sites recorded in the database. My analysis thus examines the characteristics of precontrol invasions in the park.

### 2.2.3 Dataset

The data used for this research were based on location information for 13 of the most widespread non-native plants (hereafter “target exotic species”) documented by the park database used to track and monitor invasion sites. These species were found at 355 sites (Table 2.1, Figure 2.4). All of the species are woody with the exception of *Verbascum thapsus*. I included species that were present at greater than 10 sites in the Park.

I used a database of 13 GIS habitat variables to characterize the habitat conditions of the 355 sample locations (Table 2.2). I chose these variables based on availability and previous use in modeling species distributions in GRSM (Wilds 1996, Boetsch 2000, van Manen *et al.* 2002). All GIS coverages were continuous variables based on 30 m pixels, with the exception of disturbance history. This categorical variable is based on historical records of logging and settlement in the area that is now GRSM (Pyle 1885, Pyle 1988). Logging is categorized by intensity (*i.e.*, sites with a history of light, selective, or heavy cutting). Sites with a history of settlement or no disturbance are also identified in the dataset. Finally, I computed a correlation matrix for all continuous variables (Table 2.3).

### 2.2.4 Univariate Analyses

I used randomization tests written for ArcView Version 3.2 (ESRI, Redlands, CA) to test whether invasion sites had a particular signature with respect to landscape-level

variables. I asked whether the target exotic species are found at sites with significantly lower elevation and closer to corridors (*i.e.*, roads and streams) compared to random sites. Further, I asked whether target exotic species are found at sites with significantly higher solar insolation (*i.e.*, BTA, TSI, and LFI) and moisture availability (*i.e.*, lower RSP, slope and TCI, higher TRMI, and closer to streams) than would be expected at random. Finally, I tested whether the sites with target exotic species had lower topographic complexity (*i.e.*, SWI) compared to random sites.

To run a randomization test I identified a chosen number of random points in the park. I set the number of random points equal to the number of sites with target exotic species. The number of random points varied depending on whether I was looking at values for all target exotic species or a subset (Table 2.1). Next, I calculated the average value of a chosen variable for all of the random points. These two steps were repeated 1,500 times giving 1,500 sets of random points each having an associated mean value of the independent variables. I ran the randomization test for all 13 independent variables (Table 2.2) for the full dataset (*i.e.*, 13 species at 355 sites).

I compared the average value of a variable for the invasion sites against the random iterations. Because my hypotheses for exotic habitat versus random points were one-tailed I calculated the  $p$ -value by dividing the number of observations that were greater than or equal to the observation for invasion sites (or when applicable less than or equal to the observation for invasion sites) by the total number of observations (*i.e.*, 1,501). Because I was using variations of the same dataset repeatedly, I used a Bonferroni correction to interpret the  $p$ -value. The Bonferroni correction is conservative and recalculates alpha by dividing the number of times a dataset is used. For this research I used an alpha of 0.05 and used

variations of the same data (Table 2.2) 62 times. This means that  $p$ -values less than 0.000806 (*i.e.*, 0.05/62) were significant at the 0.05 level.

To address the fact that most invasion sites are found near roads, I used a variation of the randomization tests described above by restricting the placement of random points to within 1,000 meters of roads.

To test whether disturbance history is correlated invasibility, I used a Chi-Square test to compare the distribution of all sites by disturbance history and sites with non-native plants.

### 2.2.5 Multivariate Model Formation

I built multivariate models of suitable habitat using the Mahalanobis Distance statistic ( $D^2$ ) (Rao 1952, Clark *et al.* 1993, Knick and Dyer 1997, Corsi *et al.* 1999, Boetsch 2000, van Manen *et al.* 2002). Mahalanobis Distance is a multivariate technique that only requires presence data. Presence only approaches are ecologically valid for studying invasion patterns, given that sites without invasives are not necessarily unsuitable habitat, rather they may be sites that have yet to be colonized (McCay 2001).  $D^2$  is a dimensionless statistic calculated as follows:

$$D^2 = (\underline{x} - \underline{\hat{u}})' \Sigma^{-1} (\underline{x} - \underline{\hat{u}})$$

where  $\underline{x}$  is a vector of habitat characteristics for each cell in the GIS grid,  $\underline{\hat{u}}$  is the mean vector of habitat characteristics of the invasion sites, and  $\Sigma^{-1}$  is the inverse of the variance-covariance matrix calculated from the invasion sites. The statistic represents the standard squared distance between a set of sample variates,  $x$ , and “ideal” habitat, represented by  $\hat{u}$ . Because  $D^2$  is calculated using the variance-covariance matrix, collinearity among independent variables is not an issue.

$D^2$  gives a measure of dissimilarity such that low values of  $D^2$  are most similar to the sample sites (*i.e.*, sites with invasives) and high values of  $D^2$  have the highest dissimilarity to the sample sites. I calculated a value of  $D^2$  for each 30 m pixel in GRSM using the Grid function in ArcInfo Version 8.1 (ESRI, Redlands, CA). I used S-plus version 6.0 (Insightful Corp., Seattle, WA) to calculate the covariance matrix for all independent variables in the model and to invert the covariance matrix.

Because the distribution of  $D^2$  consistently had a long tail, I defined the lowest 95 percent of  $D^2$  values as “suitable” habitat. For each habitat model I built, I included only those variables that were significant in the randomization tests. However, I included BTA in the model of the full dataset because it was marginally significant.

#### 2.2.6 Role of Individual Variables

To assess how individual variables affected the full model I looked at the change in each independent variable included in the model across quartiles of  $D^2$  using Dice-Leraas diagrams (Dice and Leraas 1936, Sokal 1965, Jones 1988). The 1<sup>st</sup> quartile of  $D^2$  values represents the most suitable habitat, while the 4<sup>th</sup> quartile of  $D^2$  values represents the least suitable habitat. I wanted to know whether there was a change in the mean values of variables as habitat became less similar to invasion sites. As well as mean, Dice-Leraas diagrams include other descriptive statistics (*i.e.*, minimum, maximum, median and standard deviation), which may change as similarity to invasive habitat decreases. Viewing changes in variables by quartile of  $D^2$  is arbitrary, but gives a more detailed picture of how variables change as habitat becomes less similar to invasion sites than only comparing sites above and below one value of  $D^2$  (*i.e.*, “suitable” and “unsuitable” habitat).

To examine the spatial effects of individual variables in the pattern of invasion prone habitat, I looked at the change in the distribution of “suitable” habitat when variables were omitted from the model. To determine which variables to use in these map analyses, I referred to the Dice-Leraas diagrams and chose variables where the mean shifted obviously as habitat became less similar to invasion sites.

### 2.2.7 Fit of the Habitat Model

I tested the habitat model by sampling throughout a range of  $D^2$  values in one catchment. I chose the Deep Creek catchment in North Carolina for sampling because of ease of access to a range of  $D^2$  values. High elevation areas, where unsuitable invasion habitat was concentrated according to the model, are accessible off high elevation roads in the northern part of this catchment. Sampling took place in July 2002 at 195 sites along trails. I sampled every 50 meters along approximately 30 kilometers of trails to maximize the number of sites I could visit. At each location, I recorded the location using a handheld GPS, whether any target invasive plants were present, and whether *M. vimineum* was present. I used the distribution of *M. vimineum* to test the model’s general applicability because this species is common, but data on its distribution are independent from the data used to build the model.

I was able to identify a  $D^2$  value for each site using the location data. I used logistic regression of exotics (present or absent) against  $D^2$  to assess the fit of the model. Next, I calculated the empirical probabilities for the presence of exotics using deciles of the  $D^2$  distribution. In each decile the probability is calculated by dividing the number of exotic plots by the total number of plots in that decile. Finally, I built a loess curve (locally estimated scatterplot smoothing) which is fit to the empirical probabilities calculated for the

$D^2$  deciles. The loess procedure fits a line using only data in a local neighborhood of each point. If the loess curve deviates markedly from the logistic regression curve it is likely that an inappropriate model was used for the logistic regression.

Using SAS version 8e (SAS Institute Inc., Cary, NC) I did a Hosmer-Lemeshow goodness-of-fit test based on the deciles of the predicted probability distribution to determine whether there is evidence of a significant lack of fit from the linear logit model.

### 2.2.8 Logistic Regression

I used stepwise Logistic regression (Hosmer and Lemeshow 1989, ter Braak and Looman 1995) to model the habitat of the target exotic species using SAS version 8e (SAS Institute Inc., Cary, NC). I used this method as an additional means to determine the most important variables controlling the distribution of the target exotic species and to compare results with the model derived using Mahalanobis distance. I derived my response variable using the distribution of target exotic species (presence) compared to the same number of random points (absence). To address the criticism that logistic regression “over fits” the data, I repeated this procedure with 10 sets of random points. For the final model, I included variables that were significant in the majority of 10 iterations of random points.

Using this model I looked at misclassifications of the data. Because invasion is continuous, sites without target exotics that the model classified as “good” exotic habitat are not a concern. I was more interested in sites with a target exotic species that do not appear to be “good” invasion habitat. I found the  $D^2$  value for sites with target exotics that had a low probability (*i.e.*,  $p < 25\%$ ) of having an exotic according to the logit model

### 2.2.9 Variations in the Model by Species Type

To determine whether it is legitimate to group all thirteen species to model invasion habitat, I used the same methods as described for the full dataset (*i.e.*, randomization tests and habitat model derived using Mahalanobis Distance) on subsets of species. Because most species are only present at a small percentage of sites, I grouped species with similar dispersal or leaf type as follows:

- ◆ Primarily seed dispersal (9 species at 209 sites)
- ◆ Primarily vegetative spread (3 species at 228 sites)
- ◆ Deciduous species (9 species at 232 sites)
- ◆ Evergreen and semi-evergreen species (4 species at 216 sites)

### 2.2.10 Accessibility of Known Invasion Sites

Many recorded invasion sites are near roads. Roads and other access points make non-native plants easier to observe, which suggests that identified invasion sites may be biased by accessibility. To determine whether the GRSM database of invasion sites was biased by access I sampled for species presence/absence at 40 sites that were at least 100 meters from roads and on average 125 meters from a trail and at 40 random sites that were adjacent to roads. I sampled along NC 129 and NC 28 for the roadside sites. I chose these highways because they are adjacent to the park border and run through “suitable” habitat. Further, development along these highways is low, which may interfere with invasibility.

I chose random distances along the road to determine sample sites. I used the vehicle odometer to determine when to stop. For off-road sites, I sampled off trails in the Deep Creek catchment. I used a pedometer to choose random points to exit the trail in a random direction and walked approximately 50 meters into the forest.

At each site I recorded the location using a handheld GPS and whether any of the non-native plants tracked by the park were present. All sites were in habitat classified as “suitable” by the multivariate model derived using Mahalanobis distance. Data collection took place in July 2002.

I used a Pearson Chi-Square test to determine whether the likelihood of encountering a non-native plant is significantly different between roadside and off-road sites.

## **2.3 Results**

### 2.3.1 Univariate Randomization Tests

The full randomization tests on the complete dataset (*i.e.*, all 13 target exotic species) showed that all variables except for LFI, Solar Incidence and BTA were significantly different for invasion sites compared to random points (Table 2.4). In summary, invasion sites based on all 13 target exotic species were at sites with lower elevations, higher moisture (TCI, TRMI), and higher micro-scale topographic exposure (TSI). Further, invasion sites were found more often than expected at random on less steep slopes, lower relative slope positions, closer to roads and streams and at sites with less topographic complexity. However, invasion sites did not appear any more often than would be expected at random on sites with south and west facing slopes (BTA) or sites with higher levels of meso-scale exposure (LFI).

The restricted randomization tests (*i.e.*, random points were all within 1,000 meters of roads) showed similar results as the full randomization tests (Table 2.5)). However, BTA and LFI were only significant using the restricted randomization tests and SWI was not significant according to the restricted randomization test.

### 2.3.2 Distribution of Invasion-Prone Habitat

$D^2$  values ranged from 3 to 5,651 with a mean of 222.9 (SD = 154.5).  $D^2$  values for invasion sites with at least one target exotic species ranged from 17 to 513 with a mean of 99.9 (SD = 60.8) (Table 2.6). Ninety-five percent of invasion sites had a  $D^2$  value below 215 (*i.e.*, “suitable” habitat) and approximately 58 percent of GRSM has a  $D^2$  below 215 (Figure 2.5). The distribution of suitable habitat is concentrated along the Park’s borders and in the northwestern part of the park (Figure 2.5).

Defining habitat as suitable below  $D^2$  of 513 (*i.e.*, the largest  $D^2$  value for an invasions site) shows that 95 percent of GRSM has a  $D^2$  below this value. Although this map gives a different picture than using a  $D^2$  of 215 as a cutoff and shows that most of GRSM is potentially suitable for invasion, it is still worthwhile to focus on invasion prone habitat that has a higher likelihood of invasibility (*i.e.*, sites with a higher similarity to existing invasion sites).

### 2.3.3 Fit of Habitat Model

$D^2$  values for the 195 sample locations to test the model ranged from 26 to 837. Target exotic species were found at 14 sites, which ranged in  $D^2$  from 26 to 154. The loess curve calculated from the  $D^2$  deciles does not deviate markedly from the logistic regression curve built using the presence/absence of target exotic species, suggesting that the model is appropriate (Figure 2.6). Further, the Hosmer-Lemeshow goodness of fit statistic indicated an appropriate fit of the model (Table 2.7). There are few data points with target exotics from the test plots, so evaluating the model is difficult.

*Microstegium vimineum* was found at 31 sites, which ranged in  $D^2$  from 29 to 193. Evaluating the model’s generality using these independent data showed similar results

(Figure 2.7) as using presence/absence of target exotics. The model was generally applicable enough to correctly identify sites with *M. vimineum* as “suitable” for invasives (Table 2.8).

The results of the analysis using logistic regression also provide support for the suitability of the Mahalanobis habitat model (Table 2.9). Nineteen sites had a target exotic but did not appear to be “good” invasion habitat (*i.e.*, probability < 25 percent) according to the logit model. These sites ranged in  $D^2$  from 93 to 513 (*i.e.*, the highest  $D^2$  for an exotic site), and had an average  $D^2$  of 201.6 which was twice as high as the average  $D^2$  for all target exotic species.

#### 2.3.4 Role of Individual Variables

The Dice-Leraas diagrams (Figure 2.8) show that elevation, distance to roads, and distance to streams vary by  $D^2$  quartile. The mean for these variables is lower at sites most similar to invasion sites (*i.e.*, 1<sup>st</sup> quartile of  $D^2$  values) compared to sites least similar to invasion sites (*i.e.*, 4<sup>th</sup> quartile of  $D^2$  values). None of these variables is highly correlated with any of the remaining variables (Table 2.3) which means that removing them from the model should remove their effect on the distribution of habitat.

Using stepwise logistic regression consistently resulted in a model that included elevation, distance to roads, distance to streams, and disturbance history (Table 2.10). Further, TCI, Slope and LFI were consistently included in the logit model (Table 2.10). The logit model showed that LFI was significant, although the full randomization test did not show that this variable was significantly different between random points and sites with target exotic species. Further, the change in mean LFI across quartile of  $D^2$  was not exceptional (Figure 2.8c).

Distance to roads and distance to streams did not cause a significant change when removed from the logistic model (Table 2.11). However, elevation, disturbance history, LFI, and slope did cause a significant change when removed from the logistic model (Table 2.11).

Overall, elevation, distance to roads, distance to streams, and disturbance history appear to be important according to multiple analyses including, the randomization tests, Dice-Leraas diagrams, and multiple iterations of logistic regression (Table 2.12)

### 2.3.5 Role of Elevation

Elevation stands out as the variable that shifts the most as  $D^2$  values change (Figure 2.8b). Sites most similar to invasion sites (*i.e.*, 1<sup>st</sup> quartile of  $D^2$  values) are found at lower elevations than less suitable habitat and are never found at the highest elevations. However, low elevation sites are not necessarily classified as suitable habitat (Figure 2.8b) meaning that sites least similar to invasion sites (*i.e.*, 4<sup>th</sup> quartile of  $D^2$  values) were also found at low elevations.

The habitat model with elevation omitted showed a range of  $D^2$  values from 3 to 5,591 with a mean  $D^2$  of 141.9 (SD = 101.2) (Table 2.13). The mean value for invasion sites was not very different when elevation was omitted from the model (Table 2.14). Comparing the distribution of “suitable” habitat between the maps generated using the full model and when excluding elevation from the model shows that 76 percent of the Park was classified similarly according to both models and approximately 23 percent of the Park went from unsuitable to suitable when omitting elevation (Figure 2.9).

### 2.3.6 Role of Distance from Roads and Streams

Distance from roads and distance from streams also appear to control the invasibility of a site given that low  $D^2$  values had a lower mean distance from roads and streams than did

high  $D^2$  values (Figures 8d and 8f). All sites near corridors are not invasion prone as sites least similar to invasion sites (*i.e.*, 4<sup>th</sup> quartile of  $D^2$  values) were also found near corridors. However, sites most similar to invasion sites (*i.e.*, 1<sup>st</sup> quartile of  $D^2$  values) were never found at sites very far from corridors.

When omitting distance from roads from the model,  $D^2$  values ranged from 15 to 1,509 with a mean value of 251.3 (SD = 235.8) (Table 2.13). The distribution of “suitable” habitat using this model showed that 94 percent of park is classified as suitable. Comparing how habitat is classified by the full model compared to omitting roads shows that for approximately 60 percent of the Park both models classified habitat similarly. However, a large portion of the interior of the Park (about 38 percent of the Park’s area) was not suitable under the full model, but became suitable when roads were omitted from the model (Figure 2.10). These sites are on average twice as far from roads as sites that were classified the same by both models.

When omitting distance from streams from the model  $D^2$  values ranged from 3 to 5,597 with a mean value of 194.8 (SD = 130.6). The mean value for invasion sites was similar when omitting distance from streams compared to the full model (Table 2.14). The distribution of “suitable” habitat when omitting distance from streams showed a similar pattern to the full model (Figure 2.5). Further, I found that suitable and unsuitable habitat was classified similarly by these two models 92 percent of the time.

### 2.3.7 Role of Disturbance History

The actual distribution of invasion sites by disturbance history is significantly different from the expected distribution based on percentage of land within each category (Chi-Square = 185.078, 4 *df*,  $p < 0.0001$ ) (Figure 2.11). Most invasion sites occur on land

that was formerly settled or subject to light logging. Invasion sites are rare on land that was subject to heavy cutting or historically undisturbed by humans. The highest elevations with respect to disturbance history are on undisturbed sites and sites that were subject to heavy cutting (Figure 2.12). The other categories of disturbance (*i.e.*, selective cut, light cut and settlement) had comparable values for mean elevation (Figure 2.12). The distribution of sites by disturbance history in the park shows that sites with a history of light cutting are concentrated near the park border and sites historically undisturbed or subject to heavy logging are concentrated in the interior of the park (Figure 2.2).

Although I did not include disturbance history in the multivariate habitat models, sites with a history of settlement had the lowest mean  $D^2$  and sites with a history of no disturbance and heavy logging had the highest mean  $D^2$  values (Table 2.15).

#### 2.3.8 Variation of Environmental Variables across Quartiles of $D^2$

The mean for several variables, including BTA, RSP, TCI, SWI, Slope TRMI, TSI do not change across quartiles of  $D^2$  (*i.e.*, as habitat becomes less similar to invasion sites) (Figure 2.8). Further, several of these variables, including BTA (Figure 2.8a) and RSP (Figure 2.8e) showed no change in their range across quartiles of  $D^2$ ; for these variables suitable and unsuitable habitat cannot be distinguished. Although suitable habitat often cannot be differentiated from unsuitable habitat for several variables because the range overlaps as  $D^2$  changes, it is worth noting that sites most similar to invasion sites (*i.e.*, first quartile of  $D^2$  values) are never found at elevations above 1,200 m, at sites greater than 1,200 m from streams, at sites greater than 6,000 m from roads, or at sites with extreme values of TSI (Figure 2.8).

### 2.3.9 Patterns among Subgroups

The model was built using the mean values of variables from the invasion sites from 13 target exotic species. These species have a range of life histories, but comparing the mean values of variables of subgroups of the thirteen target exotics shows little variation by dispersal type or leaf type (Tables 2.16-2.19).

The average values for evergreen (Table 2.16) and deciduous species (Table 2.17) for the 13 variables were very similar.  $D^2$  values for evergreen species ranged from 19 to 566 and 95 percent of sites had a  $D^2$  below 183.  $D^2$  values for deciduous species ranged from 27 to 323 and 95 percent of sites had a  $D^2$  below 213. Based on the randomization tests, aspect was the only variable that differed between the two groups (Tables 2.16 and 2.17).

The models of the distribution of suitable habitat (*i.e.*, lowest 95 percent of  $D^2$  values) for deciduous and evergreen species show a similar pattern, although the model of suitable habitat for deciduous species is slightly more extensive. Overall, 82 percent of the park is classified similarly by these two models and 17 percent of the park is suitable for deciduous species, but not evergreen species.

There were differences between the mean values for seed spreaders and vegetative spreaders, and a significant difference in their patterns of suitable habitat. Exposure (*i.e.*, BTA) was only significant for vegetative spreaders and SWI was only significant for seed spreaders (Tables 2.18 and 2.19).

$D^2$  values for species that primarily spread vegetatively ranged from 21 to 1,379 and 95 percent of sites had a  $D^2$  less than 678.  $D^2$  values for species that disperse primarily by seed ranged from 19 to 433 and 95 percent of sites had a  $D^2$  of less than 215.

The distribution of suitable habitat (*i.e.*, lowest 95 percent of  $D^2$  values) for seed spreaders shows a similar pattern to the overall model. However, suitable habitat for vegetative spreaders is 93 percent of the Park.

#### 2.3.10 Validity of Invasive Plant Dataset

Sites along roads had a significantly higher chance of supporting a target exotic species compared to sites visited that were far from roads (Table 2.4). Nineteen of the forty roadside sites had at least one target exotic, whereas none of the off-road sites had exotics. All road and off-road sites had  $D^2$  values that were less than the highest  $D^2$  value for an invasion site (*i.e.*, 513) suggesting that all sites were within the parameters of suitable habitat as defined by the model. The Pearson Chi-Square test was significant for the road versus off-road sites (Table 2.20).

Roads in the park do not sample the landscape randomly; they are concentrated towards lower elevations (Figure 2.3). However, more than half of all roads found in the park have a mean or maximum elevation above 616 meters, which is the average elevation for invasion sites.

## **2.4 Discussion**

### 2.4.1 Modeling Habitat using Landscape-Level Variables

Invasion prone sites can be identified using landscape level GIS data. The habitat of invasion sites was outside the range of what would be expected at random for the 13 target exotic species for all variables except LFI, BTA, and Solar Incidence.

Unmeasured environmental variables are unavoidable with a landscape level approach to modeling habitat. There is incomplete data on parent geology and soil type for the park; therefore, I did not include these variables in the model. Further, several studies

have shown that soil nutrients are related to invasibility (Huenneke *et al.* 1990, Pysek and Prach 1993, Woods 1993, Morgan 1998, Wiser *et al.* 1998, Stohlgren *et al.* 1999a, Stohlgren *et al.* 1999b, Thompson *et al.* 2001). If soil variables covary with other measured variables, their effect may be included in the model, but it is likely that including soil nutrient data would improve the model's specificity. To build a predictive surface of invasion prone habitat it was necessary to have data on all nine independent variables for each 30 m<sup>2</sup> location in the park (*i.e.*, more than two million pixels); requiring two million plus data values prohibits using field variables. Modeling habitat using the Mahalanobis distance statistic is a predictive tool that is powerful largely because it can be extrapolated to a landscape scale; this approach represents a tradeoff between specificity and extent.

#### 2.4.2 Fit of the Habitat Model

The habitat model derived from the 13 target exotic species held up when tested using the distribution of *M. vimineum*. Although this species occupies its own unique niche, it has a similar response to environmental and spatial variables as the 13 target exotics used to build the model.

The model's generality is also supported by the fact that the differences among subgroups of species are not striking with the exception of species that spread vegetatively. One possible reason for this outcome is because all of the species are successful invaders, and therefore, do not sample the full continuum of life histories found in all plant species. Indeed, several studies have found that successful invaders share certain life history characteristics (*e.g.*, age at first reproduction, growth rates, number of seeds) (Baker 1974, Reichard and Hamilton 1997). This may be true because the pool of plant species that has a chance at becoming an invader may also be non-random, meaning that species that were

intentionally introduced by horticulture only sample part of the range of potential life histories. Further, sites that are prone to invasion may not include the full range of habitats found across a landscape.

The model for species that spread primarily vegetatively is not very effective. Given their dispersal limitation, these species are tightly linked with their original distribution. This subgroup, which includes *Hedera helix*, *Pueraria montana*, and *Vinca minor*, has a history of human association. Therefore, the current distribution of these species likely reflects where they were originally introduced by humans. This in effect muddles their ecological address and suggests that modeling these species in this landscape is futile.

The effect of including species that spread vegetatively in the full habitat model is negligible. The model for species that spread primarily by seed and the full model are comparable, which suggests that the distribution of species that spread primarily by seed is driving the pattern of the full model. Further, I excluded the highest five percent of  $D^2$  values for target exotics when defining suitable habitat, which were largely sites with vegetative spreaders.

#### 2.4.3 Variables that Affect Distribution of Suitable Habitat

The variables that consistently appear to be important in defining invasion prone habitat are elevation, distance from roads, and distance from streams. These variables are not highly correlated with any of the other variables included in the model. Further, the actual distribution of invasion sites by disturbance history is much different than would be expected at random. Each of these four variables has a significant spatial component unlike the other variables, which are largely environmental controls.

#### 2.4.4 Role of Elevation

Clearly elevation is important in part due to environmental parameters. The environment changes significantly with elevation in the Park (Whittaker 1956); many of the species used to model invasive habitat are not characteristic of high elevation habitats. At some upper end the ability to colonize a site will be determined by the environment found at high elevations versus the spatial processes that drive recruitment.

Spatially, low elevations are aggregated at the edge of the park where recruitment limitation is less of a control. Further, there may be an aspect of connectivity in the explanation of why elevation matters. The theory of island biogeography (MacArthur and Wilson 1967) has been used to equate mountaintops with islands because these high elevation sites are surrounded by (potentially) unfavorable sites much like islands. In this sense, isolation may explain why there are few high elevation invasion sites or even species that are successful invaders at high elevations.

The connectivity of suitable habitat in the western part of GRSM means that invasion has the potential to proceed rapidly in this area. Further, low elevation sites in this part of the park will likely be invulnerable, irrespective of why elevation matters.

#### 2.4.5 Role of Distance to Roads

Like elevation the mechanisms as to why distance to roads matters in the model may be spatial and environmental. Roadsides may be invasion prone because they are found at lower elevations and elevation has spatial and environmental components. Although roads are skewed towards lower elevations, higher elevation roads also appear to be invasion prone. Roadsides, like other disturbance corridors, are sites where the availability of resources including sunlight, moisture, and nutrients may deviate favorably from the

surrounding landscape (Schonewald-Cox and Buechner 1992, Tyser and Worley 1992, Zink *et al.* 1995, Forman and Alexander 1998, Parendes and Jones 2000). Spatially, roads in and near the park may matter because of their proximity to the park border, and therefore, a seed source. Finally, roadsides may create suitable invasion habitat that is connected, much like low elevation sites in the western part of GRSM.

Most roads in this landscape create a disturbance through suitable habitat that is near a seed source. In this sense, roads provide all of the necessary ingredients for successful invasion. Although the data suggest that roads affect invasibility of a site, questions remain as to the mechanisms behind why roads are important.

#### 2.4.6 Role of Distance to Streams

The discussion relative to roads also largely applies to streams. Riparian areas experience regular disturbances in the form of flooding, which affects the distribution of resources including nutrients, sunlight, and moisture. Streams are skewed toward lower elevations although not as heavily as roads. Finally, riparian areas form a network of connected habitat that is suitable for invasion.

#### 2.4.7 Role of Disturbance History

The disturbance variable was based on the historical land uses of logging or settlement. The mechanism driving the effect from settlement sites may be two-fold, unlike logging. First, disturbance associated with sites that were settled by humans could provide non-native species a suitable invasion window (Johnstone 1986). This mechanism could theoretically operate in sites that were settled or logged. However, settlement sites may also have higher rates of invasion because humans may have intentionally introduced non-native species at these sites.

At first glance it appears that sites with a history of settlement and logging are more prone to invasion, and that sites that were historically undisturbed (*i.e.*, not settled or logged) are not as invasible. However, the pattern of invasion by disturbance history may be artifactual because sites that were settled are aggregated near streams, less common in the interior of the Park, and are found at lower elevations. This suggests that higher invasion rates may be the result of high recruitment (*i.e.*, close to the park border), favorable environment (*i.e.* low elevations), or higher connectivity, not disturbance legacy. Further, undisturbed sites are located in remote high elevation areas, suggesting that environmental factors and isolation may control the low invasion rates at these sites. The idea that disturbance history is artifactual is supported by the fact that sites historically subjected to heavy logging, which are remote and found at high elevations, have low rates of invasion. The mechanism for invasion on formerly logged sites would likely be that logging provided a suitable invasion window. Low invasion rates at these sites do not support this mechanism, rather low invasion rates may be the result of unsuitable environment and isolation as was proposed for low invasion rates on undisturbed sites.

It is unclear how the effect of disturbance history is influenced by environment, connectivity, and recruitment. This dataset does not allow the relative effects of spatial and environmental factors to be teased apart. Following the progression of an invasion over time on sites where disturbance history is known may be a way to get at this issue. Further, experimentation may also allow one to control for environment, connectivity and recruitment. Finally, using a statistical approach that allows one to control for the effect of space (*e.g.*, Mantel's  $r$ ) would also be useful in addressing this question.

#### 2.4.8 Validity of Invasive Plant Dataset

Although roads are concentrated towards lower elevations a significant portion of roads are still higher than the “ideal” elevation used to define invasion prone habitat. There are two possible explanations for why invasion sites are closer to roads than would be expected at random.

The ecological explanation of the pattern of invasion sites along roads is twofold. First, non-native plants are common in the landscape surrounding GRSM (*pers. obs.*) but non-native plants are rare inside park borders. This suggests that the source of invasive plants in the park is the area surrounding the park, meaning that the park border is comparable to the leading edge of the invasion. Also, disturbance associated with the park edge and roads is an ideal place for non-native plants to become established (Brothers and Spingarn 1992, Parendes 1997) whereas areas with high light availability and soil disturbance (both characteristic of roadsides) are less common in the interior of the park.

Another possibility to explain the pattern of invasion sites is that sites are biased by accessibility. This would mean that park efforts to identify invasion sites are concentrated along access corridors. Conversations with park employees suggest that backcountry sites are not overlooked as possible invasion sites (Kristine Johnson, GRSM, *pers. comm.*). Further, the Pearson chi-square test of the presence of exotics along roads and remote sites was significant. Finally, I did not find any exotics at sites that were remote from roads nor did I see evidence of target exotics at any places off trails while traveling to and from random sample sites.

#### 2.4.9 Conclusion

A combination of spatial and environmental factors is correlated with the distribution of non-native species in GRSM. Although I found environmental variables to be significant, spatial signals from the habitat models coupled with the fact that GRSM staff is employing a strategy of recruitment limitation to minimize invasion suggest that spatial controls are more important at this stage of invasion. This is consistent with the findings of Wiser *et al.* (1998) who found that early in the invasion process there is a weaker relationship with site variables. Further, Grice *et al.* (2000) found that a strong influence of historical factors in explaining the pattern of shrub invasion in tropical savannas.

Assuming spatial factors are limiting non-native species distributions, it may only be a matter of time until sites categorized as unsuitable are invaded. Meaning the current definition of a 'suitable' site may be biased because non-native species may not have reached the full range of suitable habitat. However, it is likely that environment does have some effect on a species ability to invade a site. Certainly spatial controls alone cannot explain a species distribution across a landscape with such a wide range of environments as found in GRSM. At some abiotic threshold the ability of a species to get to a site will not matter because at some extreme sites will fall outside the competitive tolerance for these invaders.

It appears that it is valid to model this group of non-native species concurrently. In general, I found that the unique niche of individual species was undetectable when modeling at the landscape scale. This result may provide support for the hypotheses that species that become successful invaders share certain characteristics and that not all habitats are equally invisable.

## 2.5 Future Research

Using a GIS to model habitat sacrifices specificity, while maximizing extent. Future attempts to model invasion prone habitat will gain insight by including field variables, particularly data on soil nutrients which other research has shown to affect habitat invasibility (Huenneke *et al.* 1990, Woods 1993, Pysek and Prach 1993, Wiser *et al.* 1998, Morgan 1998, Stohlgren *et al.* 1999a, Stohlgren *et al.* 1999b, Thompson *et al.* 2001) and would be a worthwhile complement to this study.

Further, this research has raised the question of how the relative effects of spatial and environmental controls are weighted. Designing a study that controls for the effect of space (*e.g.*, Mantel's  $r$ ) is one potential way to untangle the effects of spatial and environmental controls. Such a study must include field variables so that the effect of space is not overestimated (*i.e.*, unmeasured environmental variables will inflate the effect of space on the distribution of invasive species).

Much research has found that roadsides are prone to invasion. The follow-up question often posed is: do roadsides serve as a jumping off point for non-native species to colonize adjacent natural habitat (Parendes 1997, Tyser and Worley 1992)? In contrast, this research has raised the question of what mechanisms are driving invasion along roads. Future research could address the role of connectivity of roadside habitat as well as changes in resource levels along these disturbance corridors. The idea of connectivity can also be explored with respect to riparian habitat and low versus high elevation sites. If connectivity of suitable habitat matters, then invasion should proceed faster in flatter terrain versus mountainous terrain.

The current study modeled habitat for a group of 13 species. Looking at subgroups of the target exotic species, suggested this approach was valid; however, the general result that each species' niche did not affect the full habitat model could be explored further by modeling each species individually.

Table 2.1. Summary information for target exotic species in Great Smoky Mountains National Park.

<b>Species</b>	<b>Mode of Spread</b>	<b>Leaf Type</b>	<b>Number of invasion sites</b>
<i>Ailanthus altissima</i>	Both	Deciduous	11
<i>Albizia julibrissin</i>	Mostly seed	Deciduous	16
<i>Celastrus orbiculatus</i>	Mostly seed	Deciduous	19
<i>Hedera helix</i>	Mostly vegetative	Evergreen	25
<i>Ligustrum vulgare</i>	Mostly seed	Semi-evergreen	49
<i>Lonicera japonica</i>	Mostly seed	Semi-evergreen	31
<i>Paulownia tomentosa</i>	Mostly seed	Deciduous	12
<i>Populus alba</i>	Mostly seed	Deciduous	20
<i>Pueraria montana</i>	Mostly vegetative	Deciduous	92
<i>Rosa multiflora</i>	Mostly seed	Deciduous	36
<i>Verbascum thapsus</i>	All seed	Deciduous	26
<i>Vinca minor</i>	Mostly vegetative	Evergreen	111
<i>Wisteria floribunda</i>	Mostly seed	Deciduous	11

Table 2.2. Variables used to characterize habitat of target exotic species in Great Smoky Mountains National Park (GRSM).

<b>Variable</b>	<b>Description</b> <sup>1</sup>	<b>Range in GRSM</b>	<b>Source</b>
Aspect (BTA)	Aspect transformed using: $1 + \cos(45 - \text{aspect})$	0.0-2.0	Beers <i>et al.</i> (1966)
Distance to Roads	Distance from roads (m)	0-11,653	Calculated using Find Distance Command
Distance to Streams	Distance from mainstreams (m)	0-2,809	Calculated using Find Distance Command
Disturbance History (DH) <sup>2</sup>	Based on settlement and logging history	Categorical	Pyle (1985), Pyle (1988)
Elevation	Digital Elevation Model (DEM)	266-2,027	U.S.G.S. DEM Data
Landform Index (LFI)	Meso-scale topographic exposure	-1-67	McNab (1993)
Relative Slope Position (RSP)	Relative Slope Position (%)	0-100	Wilds (1996)
Shannon-Weiner Index (SWI)	Index of topographic complexity	13-35	Miller (1986)
Slope	Slope steepness (degrees)	0.0-61.0	Calculated from DEM
Solar Incidence	Index of exposure to sunlight; approximated for the solar equinox; considers both nearby and distant shadowing factors	0-254	Calculated from DEM using Hillshade Command
Terrain Shape Index (TSI)	Index of micro-scale topographic exposure	-169-91	McNab (1989)
Topographic Convergence Index (TCI)	Simulates the flow accumulation of water; $TCI = \ln(A / \tan B)$ , where A is drained surface area and B is drained surface slope	15-200	Beven and Kirkby (1979), Wolock and McCabe (1995), Halpin (1995),
Topographic Relative Moisture Index (TRMI)	Index of moisture considering the effects of slope position, aspect, and elevation	0-85	Parker 1882

<sup>1</sup> All variables were continuous with spatial resolutions of 30 m, except for disturbance history which was categorical

<sup>2</sup> Disturbance History included the following classes: no history of settlement or logging; settlement, and light, selective, or heavy cutting.

Table 2.3. Correlation matrix of independent variables used to model habitat of target exotic species. Correlations greater than 0.5 are in bold.

	TSI	TRMI	TCI	SWI	Slope	RSP	Elevation	BTA	DS	DR	LFI	SI
TSI	1	0.3799	<b>0.7353</b>	0.0656	-0.1695	-0.4904	-0.1300	0.0277	-0.0653	-0.0612	0.1406	-0.0077
TRMI	--	1	<b>0.5437</b>	-0.1698	-0.4374	-0.7139	-0.0728	-0.4790	-0.2873	-0.0994	0.0001	0.3251
TCI	--	--	1	-0.1042	-0.4308	<b>-0.5912</b>	-0.0744	-0.0015	-0.1249	-0.0806	0.0241	0.1453
SWI	--	--	--	1	0.4339	-0.1830	0.0092	0.1704	0.1537	0.2372	0.4249	-0.2029
Slope	--	--	--	--	1	0.2404	0.1426	-0.0830	0.2305	0.1746	0.3932	-0.4051
RSP	--	--	--	--	--	1	0.0681	-0.0673	0.1625	-0.0510	-0.2700	-0.1423
Elevation	--	--	--	--	--	--	1	-0.0224	0.0708	0.1822	0.0704	-0.0831
BTA	--	--	--	--	--	--	--	1	0.1805	0.0753	0.0176	0.0560
DS	--	--	--	--	--	--	--	--	1	0.1910	-0.1893	-0.0681
DR	--	--	--	--	--	--	--	--	--	1	0.0858	-0.1418
LFI	--	--	--	--	--	--	--	--	--	--	1	-0.1645
SI	--	--	--	--	--	--	--	--	--	--	--	1

Table 2.4. Results of randomization tests on full dataset (*i.e.*, all 13 target exotic species) allowing random points to fall anywhere in the Park.

Variable	<i>p</i> -value <sup>1</sup>	Interpretation <sup>2</sup>	Average Value for Invasion Sites
BTA	0.0060	No relationship	1.125
Distance to Roads	<b>0.0007</b>	Closer than expected	1033.094
Distance to Streams	<b>0.0007</b>	Closer than expected	198.797
Elevation	<b>0.0007</b>	Lower than expected	616.085
LFI	0.1018	No relationship	23.042
RSP	<b>0.0007</b>	Lower than expected	46.625
Slope	<b>0.0007</b>	Lower than expected	168.961
Solar Incidence	0.1277	No relationship	12.530
SWI	<b>0.0007</b>	Lower than expected	27.183
TCI	<b>0.0007</b>	Higher than expected	81.192
TRMI	<b>0.0007</b>	Higher than expected	30.949
TSI	<b>0.0007</b>	Higher than expected	-17.051

<sup>1</sup> Note that *p*-values are the same because most variables had only one observation (*i.e.*, the average for invasion sites) outside the range of expected values. *P*-values less than 0.0008 are significant with 95 percent confidence and are in bold.

<sup>2</sup> Refers to invasion sites.

Table 2.5. Results of restricted randomization tests on full dataset (*i.e.*, all 13 target exotic species) only including random points less than 1,000 meters from roads.

Variable	<i>p</i> -value <sup>1</sup>	Interpretation <sup>2</sup>
BTA	<b>0.0007</b>	Higher than expected
Elevation	<b>0.0007</b>	Lower than expected
LFI	<b>0.0007</b>	Higher than expected
RSP	<b>0.0007</b>	Lower than expected
Slope	<b>0.0007</b>	Lower than expected
Solar Incidence	0.3331	No relationship
SWI	0.2032	No relationship
TCI	<b>0.0007</b>	Higher than expected
TRMI	<b>0.0007</b>	Higher than expected
TSI	<b>0.0007</b>	Higher than expected

<sup>1</sup> Note that *p*-values are the same because most variables had only one observation (*i.e.*, the average for invasion sites) outside the range of expected values. *P*-values less than 0.0008 are significant with 95 percent confidence and are in bold.

<sup>2</sup> Refers to invasion sites.

Table 2.6. Range of  $D^2$  values for all pixels in GRSM and for sites with target exotic species.

	<b>Number of Sites</b>	<b>Min</b>	<b>Max</b>	<b>Range</b>	<b>Mean</b>	<b>Std</b>	<b>Median</b>
All Sites	2,285,376	3	5,651	5,648	222.9	154.5	183
Sites with target exotic species <sup>1</sup>	345	17	513	496	99.9	60.8	86

<sup>1</sup> Some sites do not have a  $D^2$  value because of missing data for one or more variables. This resulted in 10 of 355 sites with target exotic species having no value for  $D^2$ .

Table 2.7. Logistic regression model for predicting the occurrence of target exotic species in Deep Creek Catchment

<b>Variable</b>	<b>Parameter Estimate</b>	<b>Standard Error</b>	<b>Wald Chi-Square</b>	<b>Pr &gt; Chi-Square</b>
Intercept	0.2774	0.7080	0.1535	0.6953
$D^2$	-0.0305	0.0103	8.6654	0.0032

Hosmer and Lemeshow Goodness-of-Fit Test = 5.9496, 7 df,  $p = 0.5457$

Table 2.8. Logistic regression model for predicting the occurrence of *Microstegium vimineum* in Deep Creek Catchment

<b>Variable</b>	<b>Parameter Estimate</b>	<b>Standard Error</b>	<b>Wald Chi-Square</b>	<b>Pr &gt; Chi-Square</b>
Intercept	1.1143	0.4941	5.0851	0.0241
$D^2$	-0.0245	0.0057	18.7167	<0.0001

Hosmer and Lemeshow Goodness-of-Fit Test = 8.1468, 7 df,  $p = 0.3198$

Table 2.9. Logistic regression model for predicting the occurrence of target exotic species in Great Smoky Mountains National Park. Variables were included in final model if they were significant in more than five iterations of random points (Table 2.10). *P*-value for Disturbance History (DH) variables is based on entire DH construct. Parameter estimates for DH are interpreted against no disturbance.

<b>Variable</b>	<b>Parameter Estimate</b>	<b>Standard Error</b>	<b>Wald Chi-Square</b>	<b>Pr &gt; Chi-Square</b>
Intercept	1.5405	0.726	4.5025	0.0338
DH (Heavy Cut)	1.163	0.3248	12.8226	<0.0001
DH (Light Cut)	0.6618	0.3221	4.2206	<0.0001
DH (Selective Cut)	-0.7195	0.4215	2.9133	<0.0001
DH (Settlement)	-0.8619	0.4121	4.3745	<0.0001
Distance to Roads	-0.00016	0.00006	7.1215	0.0076
Distance to Streams	-0.00091	0.000341	7.1959	0.0073
Elevation	-0.00355	0.000602	34.7501	<0.0001
LFI	0.0936	0.0216	18.808	<0.0001
Slope	-0.0808	0.0164	24.2983	<0.0001
TCI	0.0086	0.00521	2.7206	0.0991

-2 Log L (intercept only): 984.269; -2 Log L (Intercept plus covariates): 501.894

Table 2.10. Results of 10 iterations of stepwise Logistic regression. Each iteration was based on a new selection of 355 random points compared to values for the 355 sites with target exotic species. Elevation, Distance from roads (DR), Distance from streams (DS), slope, Disturbance History (DH), LFI, and TCI were used to build the final model because they were significant in the majority of iterations.

Iteration	-2 log L	% concordant	Elevation	DR	DS	Slope	DH <sup>1</sup>	LFI	TCI	SI	SWI	BTA	TRMI	TSI	RSP
1	531.170	91.3	√	√	√	√	√	√	√	√					
2	545.806	90.8	√	√	√	√	√	√	√	√		√			
3	536.718	91.0	√	√	√	√	√	√							
4	560.973	90.4	√	√	√	√	√	√							
5	541.850	90.7	√	√	√	√	√	√	√		√			√	
6	509.838	92.3	√	√	√	√		√		√					
7	532.676	91.2	√	√	√	√	√		√				√		
8	582.180	89.4	√	√	√	√	√	√		√					
9	549.553	90.6	√	√	√	√	√		√		√				
10	544.160	90.9	√	√	√	√	√	√	√						
<b>Total iterations significant</b>			10	10	10	10	9	8	6	4	2	1	1	1	0

<sup>1</sup> Disturbance History (DH) is a categorical variable which makes it invalid for use in a stepwise approach. Significance of the entire DH construct was determined by subtracting the -2 Log L with DH in the model (501.894) from each iteration's value of -2 Log L and comparing to a Chi-Square distribution, *df* 5, *p* < 0.05.

Table 2.11. Effect of removing variables from final logistic regression model.

<b>Variable Omitted</b>	<b>% Concordant</b>	<b>- 2 Log L</b>	<b>Chi-Square<sup>1</sup></b>	<b>Pr &gt; Chi-Square</b>	<b>df</b>
None ( <i>i.e.</i> , Full Model)	92.3	501.894	--	--	11
Disturbance History	90.9	544.160	42.266	< <b>0.05</b>	5
Distance to Roads	92.0	509.223	7.329	Ns	1
Distance to Streams	92.0	509.405	7.511	Ns	1
Elevation	91.0	542.497	40.603	< <b>0.05</b>	1
LFI	91.5	521.429	19.535	< <b>0.05</b>	1
Slope	91.5	527.155	25.261	< <b>0.05</b>	1
TCI	92.2	504.672	2.778	Ns	1

<sup>1</sup> Chi-square calculated by subtracting -2 Log L for full model (*i.e.*, 501.894) from -2 Log L for reduced model.

Table 2.12. Summary of analyses where each variable showed an effect. Boxes with a check indicate that the variable was significant or showed a change (*i.e.*, mean or range shift with D2). Boxes with ‘- -’ indicate that the variable was not tested or included in this analysis.

Test	Variable												
	BTA	DH <sup>1</sup>	DR	DS	Elevation	LFI	RSP	Slope	SI	SWI	TCI	TRMI	TSI
Full randomization (Table 2.4)		√	√	√	√		√	√		√	√	√	√
Restricted randomization (Table 2.5)	√	--	--	--	√	√	√	√			√	√	√
Shift in mean across quartiles of $D^2$ (Figure 2.8)		--	√	√	√				--				
Shift in range across quartile of $D^2$ (Figure 2.8)		--	√	√	√	√		√	--	√	√	√	√
Map Calculations	--	--	√		√	--	--	--	--	--	--	--	--
Significant in > 5 iterations of stepwise Logistic model (Table 2.10)		√	√	√	√	√		√			√		
Significant when omitted from final Logistic model (Table 2.11)	--	√			√	√	--	√	--	--		--	--
<b>% of analyses where variable an effect</b>	<b>20</b>	<b>100</b>	<b>83</b>	<b>67</b>	<b>100</b>	<b>67</b>	<b>40</b>	<b>83</b>	<b>0</b>	<b>40</b>	<b>67</b>	<b>60</b>	<b>60</b>

<sup>1</sup> Disturbance History (DH) was not included in the randomization tests. The check in this box is based on results of Chi-Square test (Figure 2.11).

Table 2.13. Range of  $D^2$  values for all pixels (*i.e.*, 2,285,376) in Great Smoky Mountains National Park when omitting variables from the Mahalanobis distance habitat model.

<b>Omitted Variable</b>	<b>Min</b>	<b>Max</b>	<b>Range</b>	<b>Mean</b>	<b>Std</b>	<b>Median</b>
None ( <i>i.e.</i> , full model)	3	5,651	5,648	222.9	154.5	183
Distance from Streams	3	5,597	5,594	194.8	130.6	161
Distance from Roads	1	17,534	5,758	314.9	254.3	252
Elevation	3	5,591	5,588	141.9	101.2	118

Table 2.14. Range of  $D^2$  values for invasion sites (*i.e.*, 345) when omitting variables from the Mahalanobis distance habitat model.

<b>Omitted Variable</b>	<b>Min</b>	<b>Max</b>	<b>Range</b>	<b>Mean</b>	<b>Std</b>	<b>Median</b>
None ( <i>i.e.</i> , full model)	17	513	496	99.9	60.8	86
Distance from Streams	10	372	362	89.6	48.6	80
Distance from Roads	15	1,509	1,494	251.3	235.8	175
Elevation	13	505	492	82.7	56.1	75

Table 2.15. Distribution of  $D^2$  values by disturbance history.

<b>Disturbance History</b>	<b>Min</b>	<b>Max</b>	<b>Range</b>	<b>Mean</b>	<b>Std</b>	<b>Median</b>
Undisturbed	17	5,651	5,634	324.9	163.0	301
Settlement	4	4,723	4,719	144.9	99.1	118
Light cut	4	4,148	4,144	204.1	152.5	162
Selective Cut	3	1,226	1,223	164.1	113.8	131
Heavy Cut	7	1,267	1,260	237.6	137.5	213

Table 2.16. Results of randomization tests for evergreen and semi-evergreen target exotic species.

Variable	<i>p</i> -value <sup>1</sup>	Interpretation <sup>2</sup>	Mean value for invasion sites
BTA	0.1218	No relationship	1.084
Distance to Roads	<b>0.0007</b>	Closer than expected	1,149.411
Distance to Streams	<b>0.0007</b>	Closer than expected	186.621
Elevation	<b>0.0007</b>	Lower than expected	618.523
LFI	0.1756	No relationship	23.068
RSP	<b>0.0007</b>	Lower than expected	43.244
Slope	<b>0.0007</b>	Lower than expected	12.394
Solar Incidence	0.1697	No relationship	169.222
SWI	0.0400	No Relationship	27.091
TCI	<b>0.0007</b>	Higher than expected	83.855
TRMI	<b>0.0007</b>	Higher than expected	32.534
TSI	<b>0.0007</b>	Higher than expected	-15.773

<sup>1</sup> *P*-values less than 0.0008 are significant with 95 percent confidence and are in bold. Note that *p*-values are the same because most variables had only one observation (*i.e.*, the average for invasion sites with evergreen or semi-evergreen species) outside the range of expected values.

<sup>2</sup> Refers to invasion sites with evergreen or semi-evergreen species.

Table 2.17. Results of randomization tests for deciduous target exotic species.

<b>Variable</b>	<b><i>p</i>-value<sup>1</sup></b>	<b>Interpretation<sup>2</sup></b>	<b>Mean value for Invasion sites</b>
BTA	<b>0.0007</b>	Higher than expected	1.178
Distance to Roads	<b>0.0007</b>	Closer than expected	986.181
Distance to Streams	<b>0.0007</b>	Closer than expected	206.527
Elevation	<b>0.0007</b>	Lower than expected	612.462
LFI	0.2055	No relationship	23.100
RSP	<b>0.0007</b>	Lower than expected	47.657
Slope	<b>0.0007</b>	Lower than expected	12.534
Solar Incidence	0.1816	No relationship	168.667
SWI	0.0579	No relationship	27.383
TCI	<b>0.0007</b>	Higher than expected	78.638
TRMI	<b>0.0007</b>	Higher than expected	29.990
TSI	<b>0.0007</b>	Higher than expected	-17.990

<sup>1</sup> *P*-values less than 0.0008 are significant with 95 percent confidence and are in bold. Note that *p*-values are the same because most variables had only one observation (*i.e.*, the average for invasion sites with deciduous species) outside the range of expected values.

<sup>2</sup> Refers to invasion sites with deciduous species.

Table 2.18. Results of randomization tests for target exotics that spread primarily by seed.

Variable	<i>p</i> -value <sup>1</sup>	Interpretation <sup>2</sup>	Mean value for invasion sites
BTA	0.1277	No relationship	1.085
Distance to Roads	<b>0.0007</b>	Closer than expected	774.702
Distance to Streams	<b>0.0007</b>	Closer than expected	161.556
Elevation	<b>0.0007</b>	Lower than expected	589.646
LFI	0.5369	No relationship	22.474
RSP	<b>0.0007</b>	Lower than expected	45.057
Slope	<b>0.0007</b>	Lower than expected	11.373
Solar Incidence	0.0699	No relationship	171.354
SWI	<b>0.0007</b>	Lower than expected	26.931
TCI	<b>0.0007</b>	Higher than expected	82.544
TRMI	<b>0.0007</b>	Higher than expected	32.257
TSI	<b>0.0007</b>	Higher than expected	-16.794

<sup>1</sup> *P*-values less than 0.004 are significant with 95 percent confidence and are in bold. Note that *p*-values are the same because most variables had only one observation (*i.e.*, the average for invasion sites with species that are primarily seed spreaders) outside the range of expected values.

<sup>2</sup> Refers to invasion sites with species that spread primarily by seed.

Table 2.19. Results of randomization tests for target exotic species that have primarily vegetative spread.

Variable	<i>p</i> -value <sup>1</sup>	Interpretation <sup>2</sup>	Mean value for invasion sites
BTA	<b>0.0007</b>	Higher than expected	1.171
Distance to Roads	<b>0.0007</b>	Closer than expected	1281.764
Distance to Streams	<b>0.0007</b>	Closer than expected	227.376
Elevation	<b>0.0007</b>	Lower than expected	632.217
LFI	0.0140	No relationship	23.725
RSP	<b>0.0007</b>	Lower than expected	46.507
Slope	<b>0.0007</b>	Lower than expected	13.578
Solar Incidence	0.4511	No relationship	165.856
SWI	0.1856	No relationship	27.512
TCI	<b>0.0007</b>	Higher than expected	79.691
TRMI	<b>0.0007</b>	Higher than expected	29.971
TSI	<b>0.0007</b>	Higher than expected	-17.1353

<sup>1</sup> *P*-values less than 0.004 are significant with 95 percent confidence and are in bold. Note that *p*-values are the same because most variables had only one observation (*i.e.*, the average for invasion sites with target exotic species that have primarily vegetative spread) outside the range of expected values.

<sup>2</sup> Refers to invasion sites with target exotic species that have primarily vegetative spread.

Table 2.20. Variation in  $D^2$  values by distance from road.  $D^2$  values were calculated using multivariate model of habitat similarity.

<b>Site Type</b>	<b>Mean <math>D^2</math></b>	<b>Maximum <math>D^2</math></b>	<b>Number of sites with target exotic species</b>
Near Roads	143	519	19
Far from Roads	88	422	0

Pearson Chi-Square = 22.3641, 1 *df*, *p*-value < 0.001.

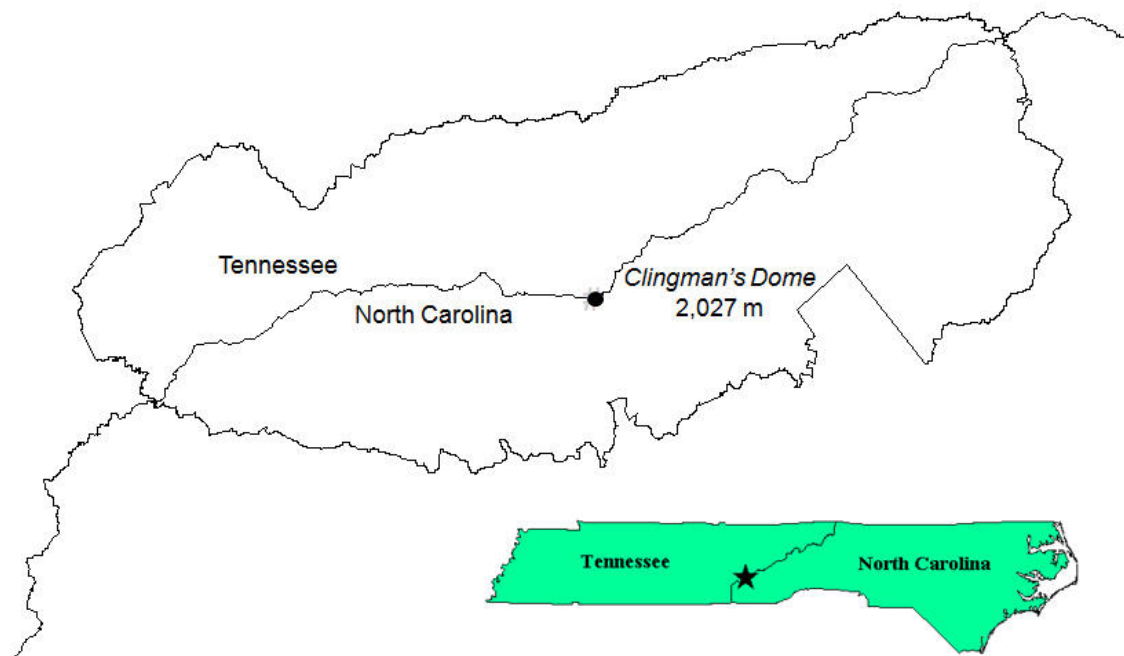


Figure 2.1. Location of Great Smoky Mountains National Park, North Carolina/Tennessee.  
Study area for development of habitat model.



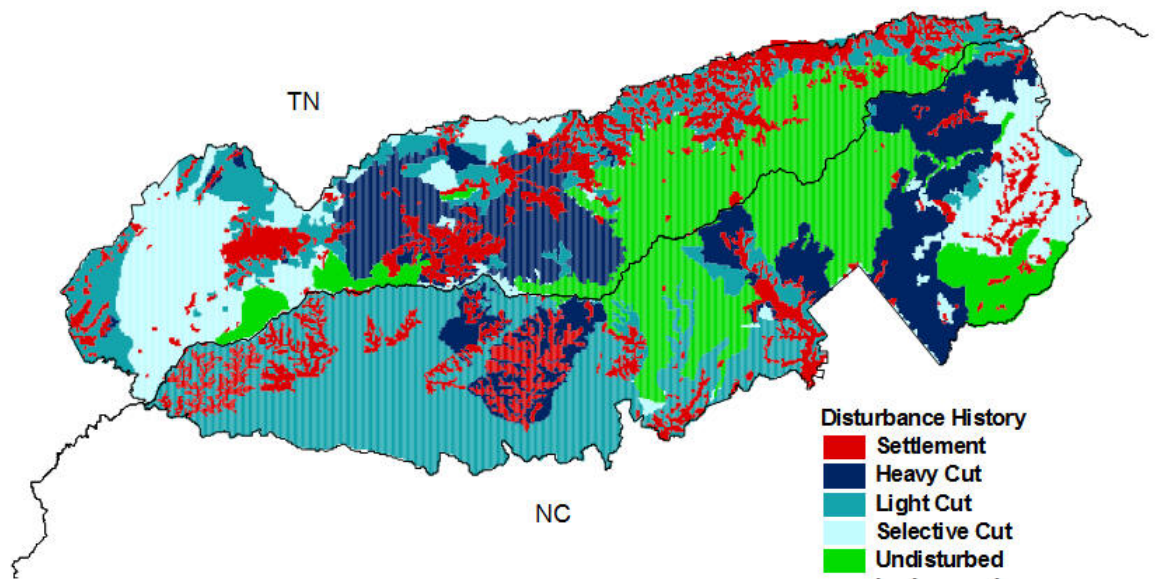


Figure 2.2. Pattern of human history of disturbance in Great Smoky Mountains National Park.



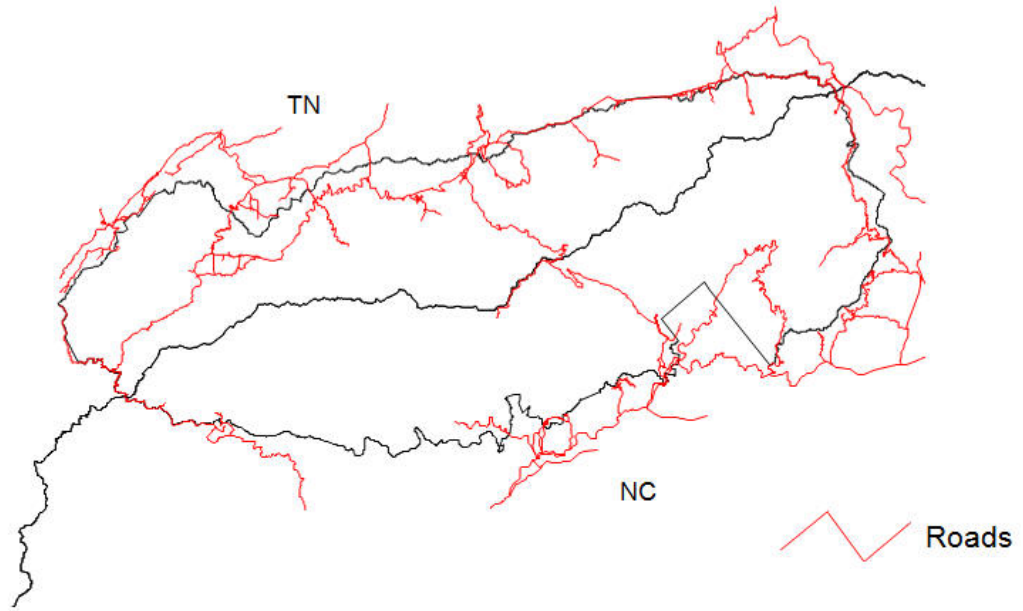


Figure 2.3. Distribution of roads near Great Smoky Mountains National Park.



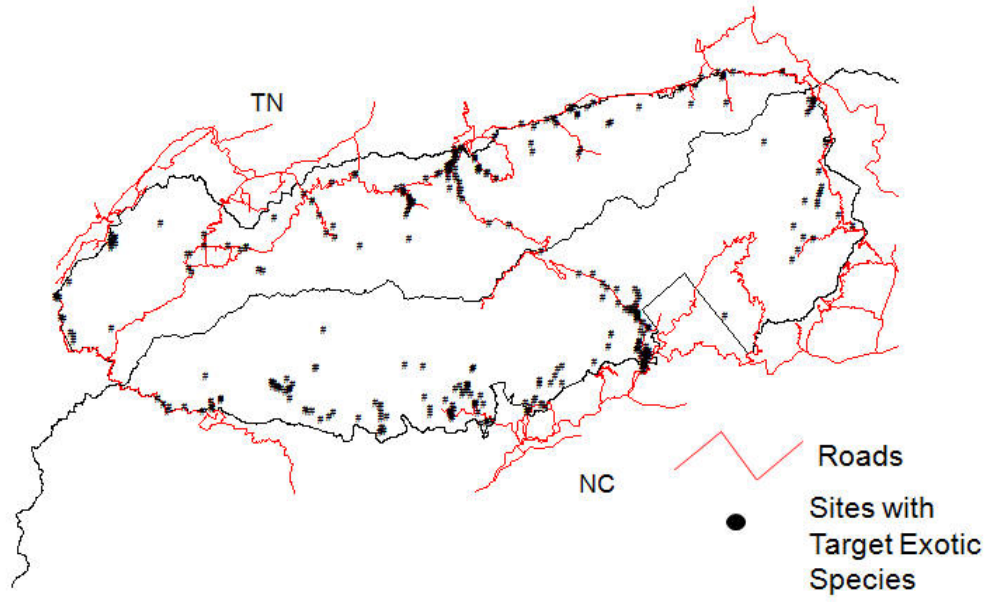


Figure 2.4. Distribution of target exotic species in Great Smoky Mountains National Park.



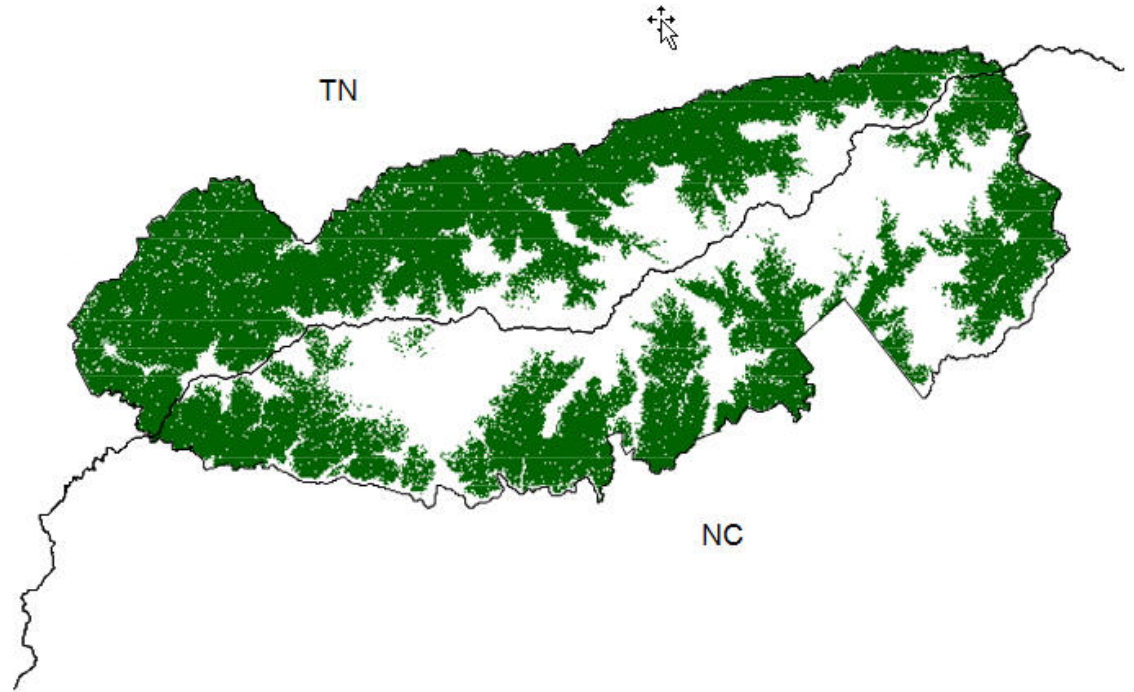


Figure 2.5. Distribution of suitable habitat for target exotic species in Great Smoky Mountains National Park. “Suitable” habitat was defined as sites with a  $D^2 < 215$  according to the Mahalanobis distance model and is shaded gray. Ninety-five of sites with a target exotic species had a  $D^2$  below 215.



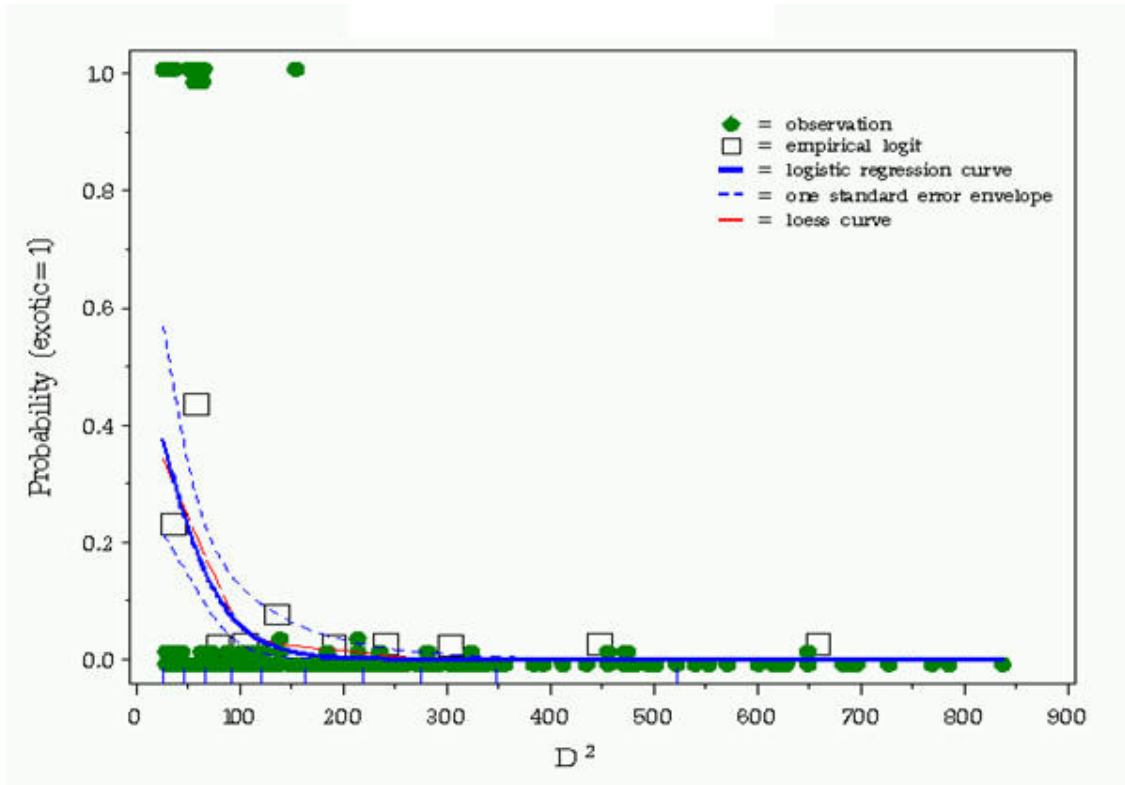


Figure 2.6. Results of testing model against presence/absence of target exotic species in Deep Creek Catchment. The loess curve (locally estimated scatterplot smoothing) is fit to the empirical probabilities calculated for the  $D^2$  deciles. The loess procedure fits a line using only data in a local neighborhood of each point. The loess curve does not deviate markedly from the logistic regression curve, which suggests an appropriate model was used for the logistic regression.



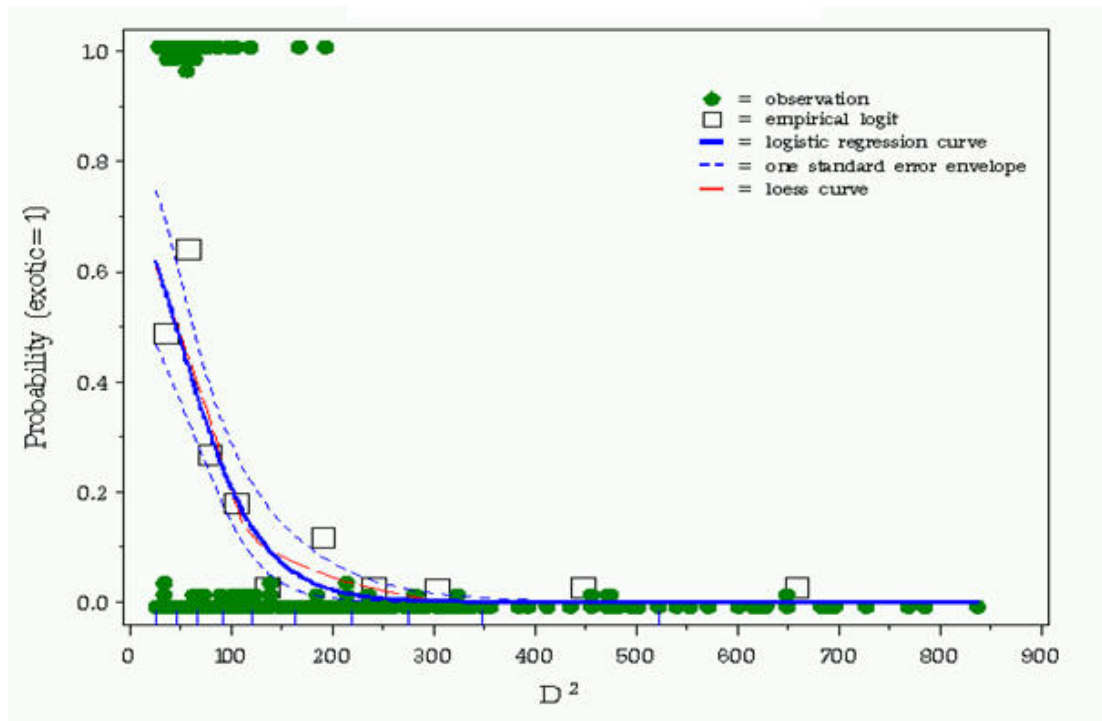


Figure 2.7. Results of testing model against presence/absence of *Microstegium vimineum* in Deep Creek Catchment. The loess curve (locally estimated scatterplot smoothing) is fit to the empirical probabilities calculated for the  $D^2$  deciles. The loess procedure fits a line using only data in a local neighborhood of each point. The loess curve does not deviate markedly from the logistic regression curve, which suggests an appropriate model was used for the logistic regression.



Figure 2.8. Dice-Leraas diagrams showing change in independent variables across quartiles of  $D^2$ . The 1<sup>st</sup> quartile of values includes sites that are most similar to invasion sites; the 4<sup>th</sup> quartile of values includes sites least similar to invasion sites. The mean for Elevation (b), Distance from roads (d), and Distance from streams (f) increases as habitat becomes less similar to invasion sites. These variables as well as Slope (e), SWI (g), TCI (h), and TSI (i) also show a shift in the range of habitat from similar to dissimilar sites. In each case the range for the 1<sup>st</sup> quartile of values (*i.e.*, sites most similar to invasion sites) is smaller than the range for the 4<sup>th</sup> quartile of values (*i.e.*, sites least similar to invasion sites)

Figure 2.8a BTA

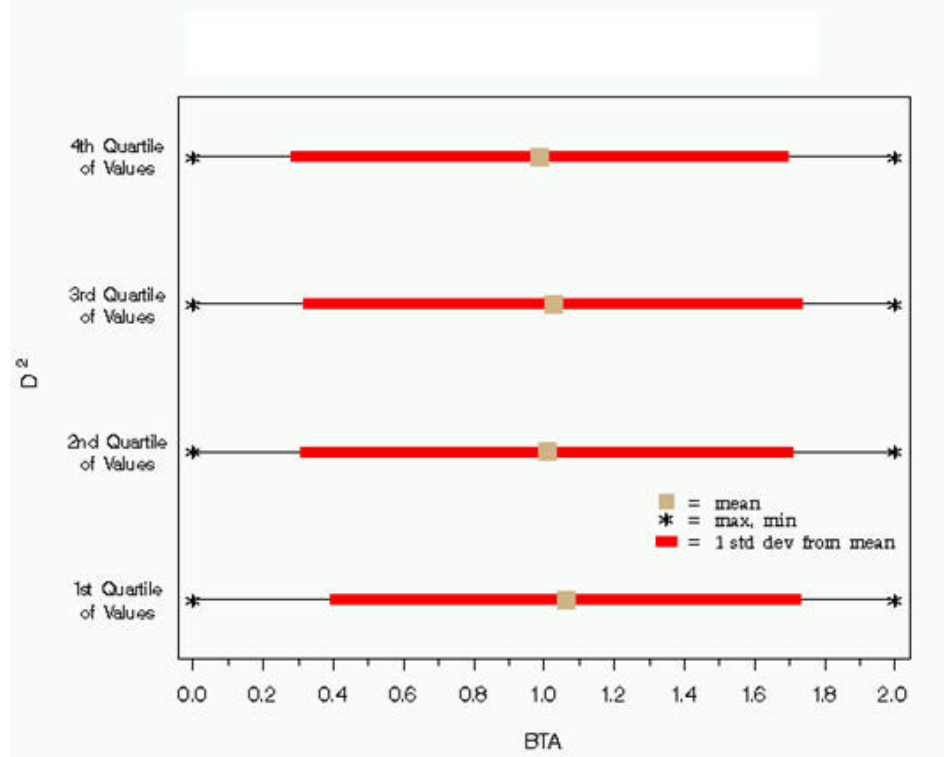


Figure 2.8b Elevation

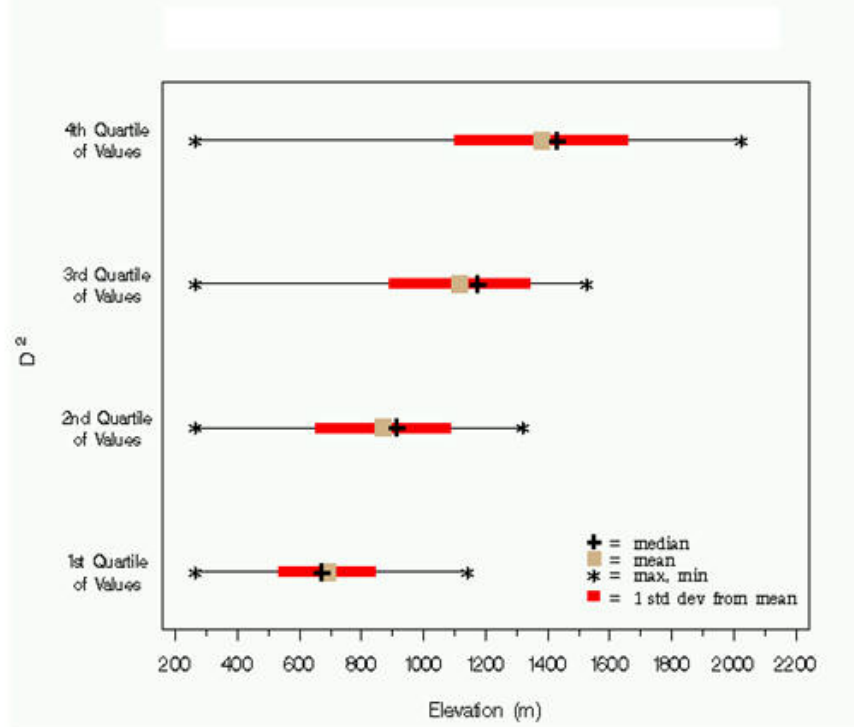


Figure 2.8c LFI

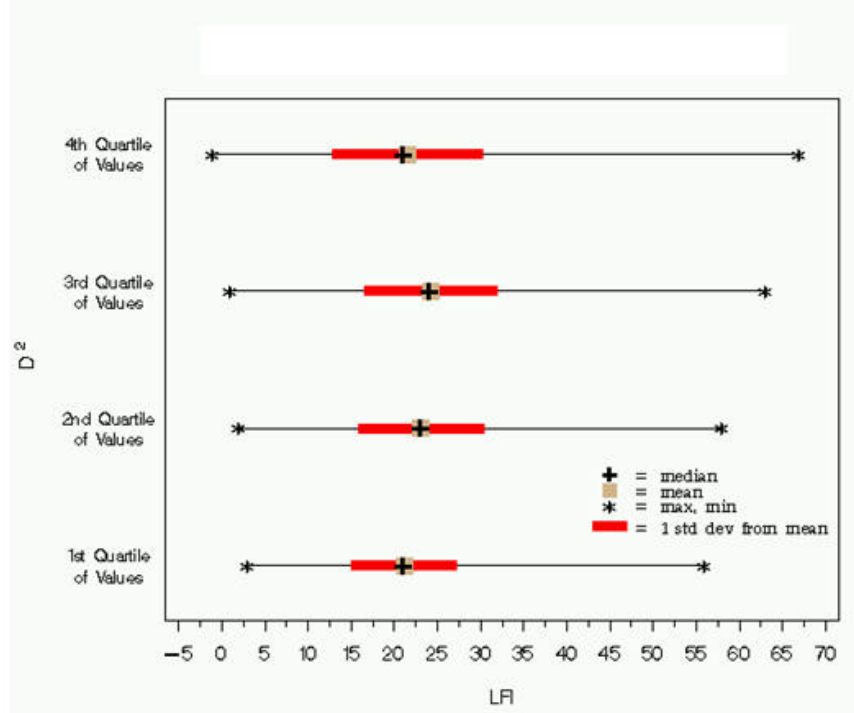


Figure 2.8d Distance from Roads

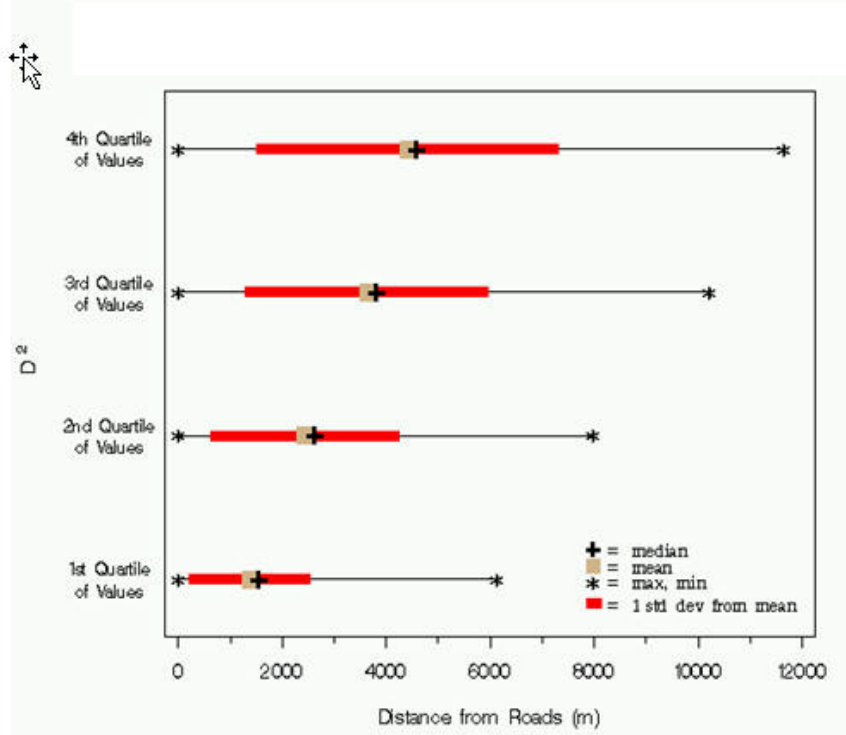


Figure 2.8e RSP

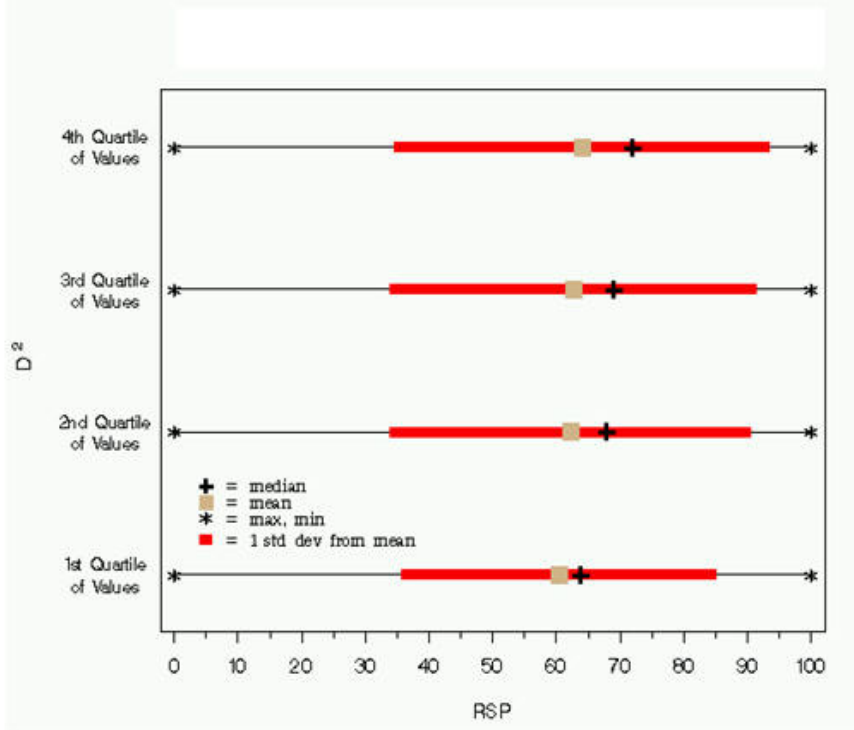


Figure 2.8f Slope

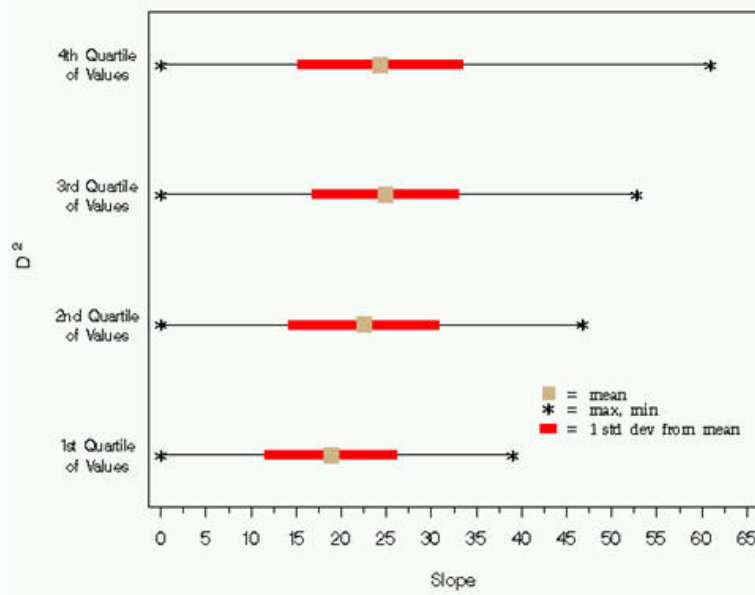


Figure 2.8g Distance from Streams

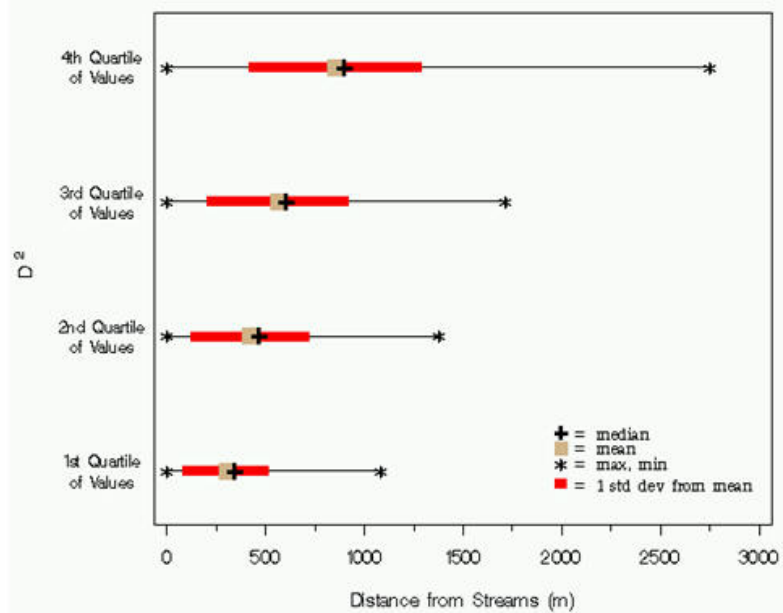


Figure 2.8h SWI

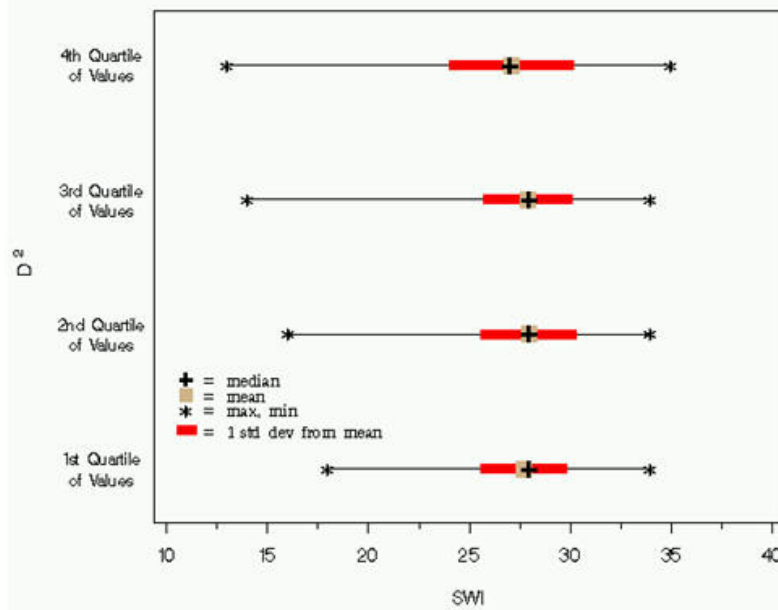


Figure 2.8i TCI

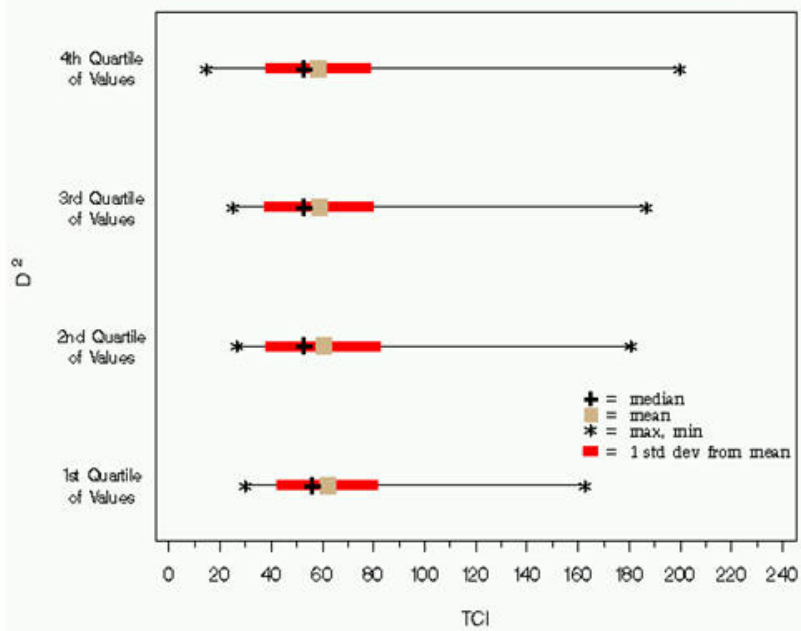
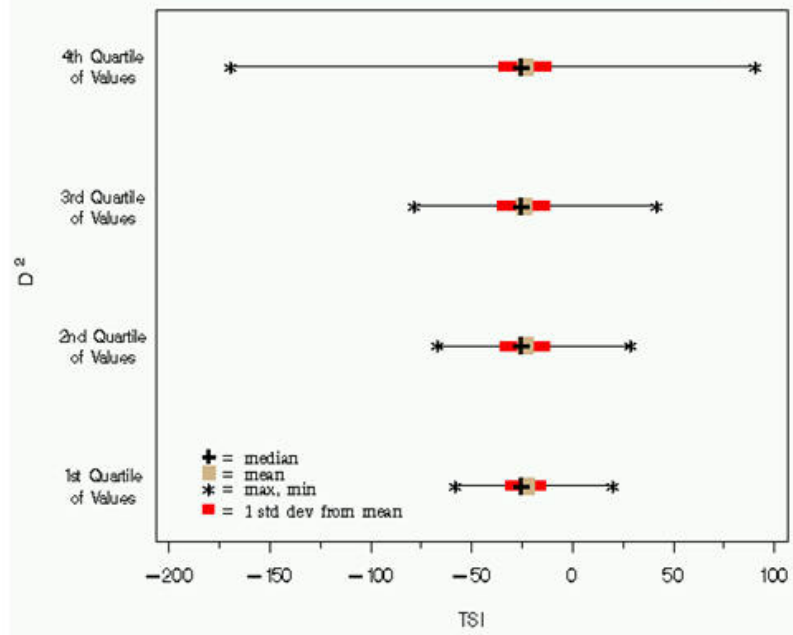


Figure 2.8j TSI













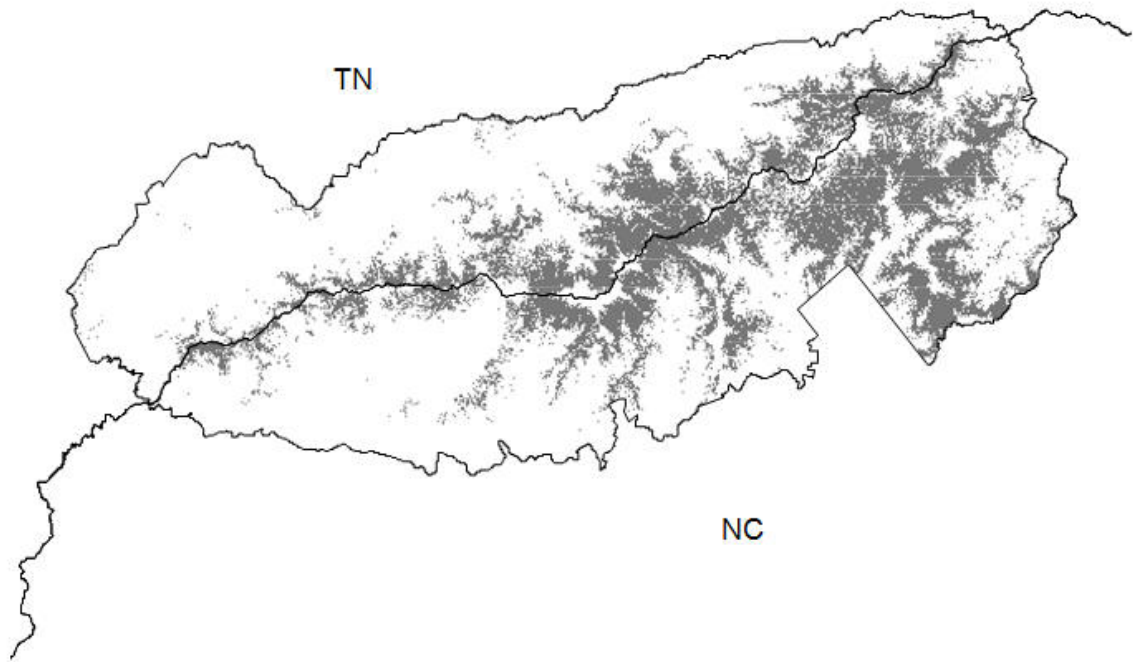


Figure 2.9. Change in suitable habitat when omitting elevation from the Mahalanobis distance habitat model. Shaded sites became “suitable” habitat when elevation was omitted from the model. “Suitable” habitat is defined as  $D^2$  values below the 95<sup>th</sup> percentile for sites with a target exotic species.



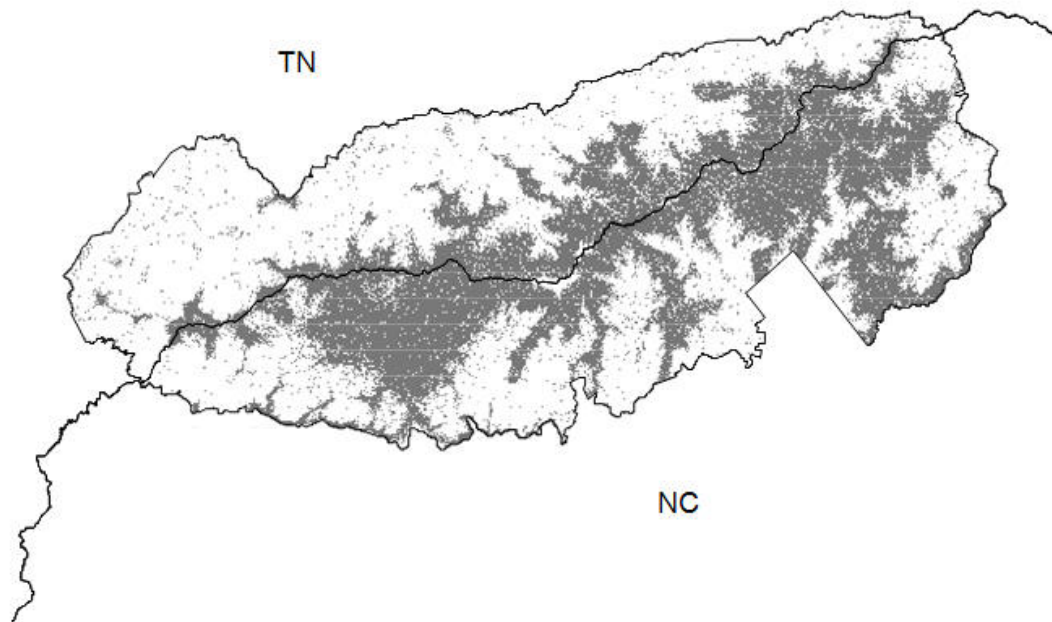


Figure 2.10. Change in suitable habitat when omitting distance from roads from the Mahalanobis distance habitat model. Shaded sites became “suitable” habitat when distance from roads was omitted from the model. “Suitable” habitat is defined as  $D^2$  values below the 95<sup>th</sup> percentile for sites with a target exotic species.



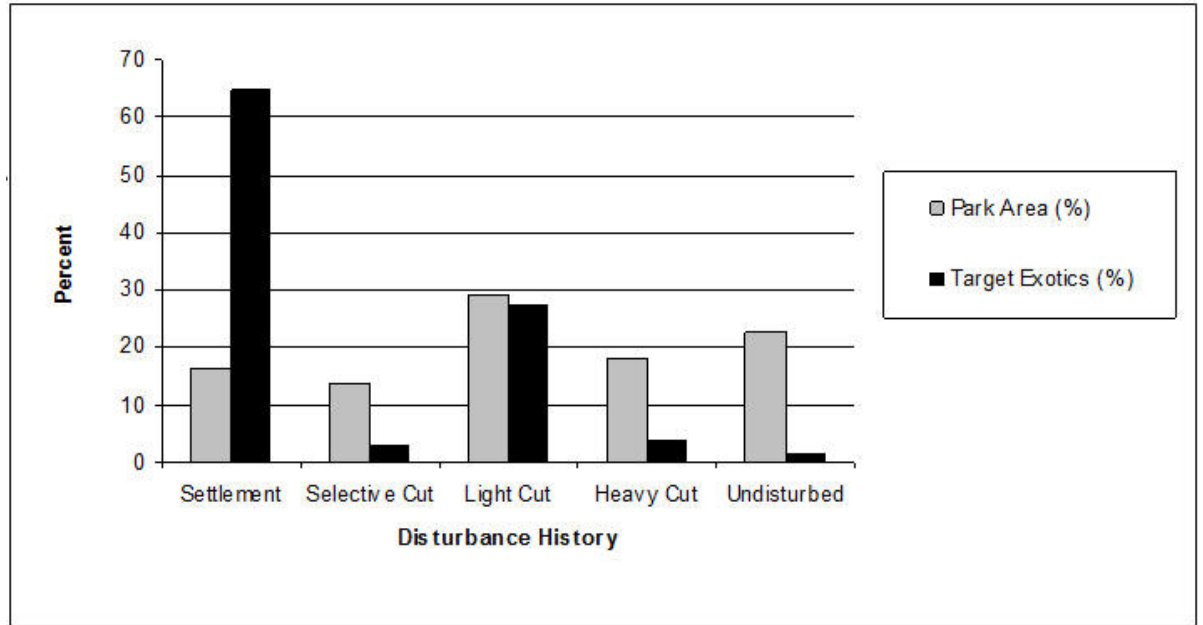


Figure 2.11. Distribution of percentage of target exotic species by percentage of landscape in each category of disturbance history. Sites with a history of no disturbance, selective cutting, and heavy cutting have less target exotic species than would be expected based on percent of Park area in these categories. Invasion sites are most likely to fall on sites with a history of settlement. Chi-square = 185.078, 4 *df*,  $p < 0.0001$ . (Significance test based on entire disturbance history construct, not a particular category.)



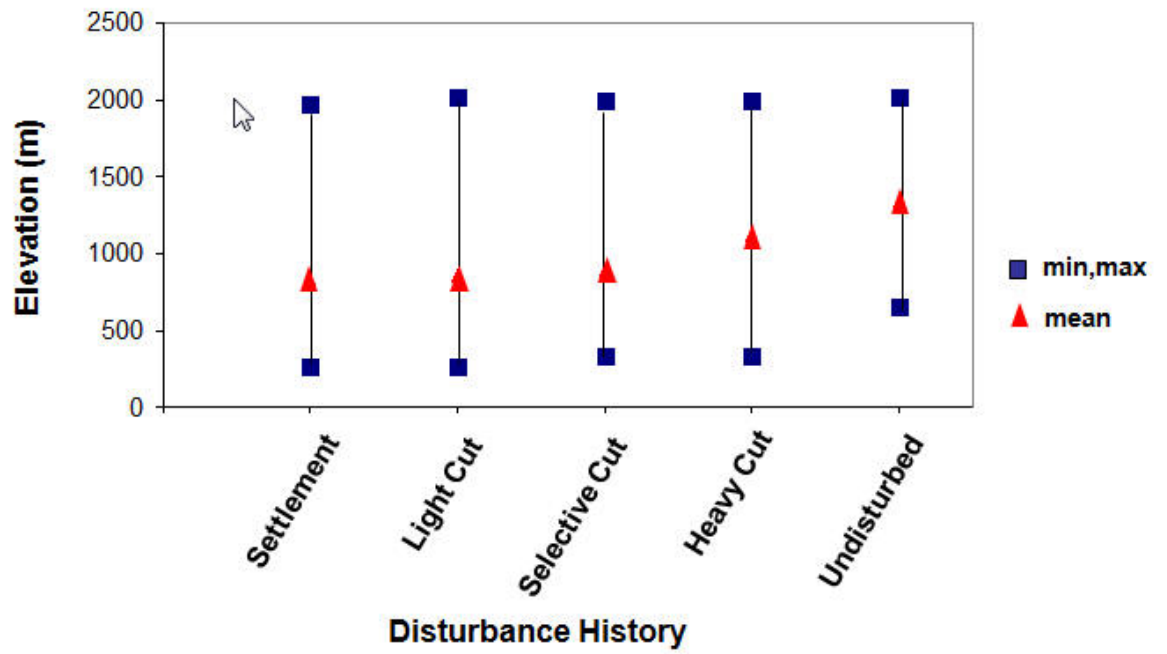


Figure 2.12. Distribution of disturbance history by mean elevation. Sites with a history of no disturbance and heavy cutting are found at higher mean elevations in the Park.

## **Chapter III: Invasion pattern of *Paulownia tomentosa* along roads surrounding Great Smoky Mountains National Park**

### **3.1 Introduction**

Several studies have shown that edges, including roadsides are prone to invasions by non-native plants (Brothers and Spingarn 1992, Hester and Hobbs 1992, Hobbs and Huenneke 1992, Tyser and Worley 1992, Zink *et al.* 1995, Parendes 1997, Morgan 1998, Parendes and Jones 2000, Meekins and McCarthy 2001). The mechanisms driving invasion along roads may include changes in resource levels including light (Parendes 1997, Parendes and Jones 2000), moisture (Forman and Alexander 1998), and nutrients (Schonewald-Cox and Buechner 1992, Forman and Alexander 1998). Further, road building causes a disturbance and maintaining roads causes a continual disturbance and disturbance has been linked with invasibility in many contexts (Denslow 1980, Hobbs and Huenneke 1992, Bergelson *et al.* 1993, Morgan 1998, Nuzzo 1999, Knapp and Canham 2000). Finally, in theory roadsides are connected habitat that passes through sites that may not be invasion prone (Schonewald-Cox and Buechner 1992, Forman and Alexander 1998). Bergelson *et al.* (1993) found that invasion proceeded faster along disturbance corridors compared to a patchy distribution of disturbance.

In addition to understanding the mechanisms behind invasion, several studies have addressed the pattern of spread (Andow *et al.* 1990, Parendes 1997, Nuzzo 1999, Grice *et al.* 2000, Rouget *et al.* 2001). Understanding the pattern of invasion may give insight into how to slow down the rate of invasion (Moody and Mack 1988, Clark *et al.* 1998)

*Paulownia tomentosa* (or princess tree, hereafter *Paulownia*) was introduced as an ornamental in the mid 1800s (Carpenter *et al.* 1983). As an escape, this species is most often found on roadsides, waste places, and other disturbed areas (Radford *et al.* 1968, Carpenter *et al.* 1983, Williams 1993a, Langdon and Johnson 1994). In 1975, over a century after it was introduced, a survey along 430 km of road near Great Smoky Mountains National Park found eight *Paulownia* trees (Baron *et al.* 1975). Currently the same area is estimated to have thousands of individuals (Langdon and Johnson 1994), although invasion of *Paulownia* along roads surrounding GRSM does not appear to be uninterrupted. Invasion of non-native plants along roads is particularly insidious because the high edge to area ratio of roadside habitat means that most seeds with fall into uninvaded habitat (Moody and Mack 1988). This is particularly true for roads surrounding GRSM.

For this research, I used the current distribution of *Paulownia* along roads surrounding GRSM to ask the following questions:

1. Does *Paulownia* colonize only certain roadside habitat?
2. What is the pattern of spread of *Paulownia* along roadsides?

## **3.2 Methods**

### **3.2.1 Study Area**

The study area for this research was the roadside habitat along approximately 650 kilometers of roads surrounding Great Smoky Mountains National Park (GRSM) (Figure 3.1). GRSM has an area of 206,119 ha and straddles the border between Tennessee and North Carolina in the southern Appalachians. Precipitation increases with elevation in this landscape (Shanks 1954). The park's climate is classified as temperate rainforest at elevations with the highest precipitation (Thornewaite 1948); lower elevations that have less

precipitation are classified as humid mesothermal (Shanks 1954). July mean temperatures range from 22° to 14° C and January mean temperatures range from 4° to -2° C (Shanks 1954).

The region is characterized by mountainous terrain with steep slopes, narrow valleys and coves, and fast-flowing streams (King and Stupka 1950). This topographically complex terrain gives rise to a range of habitats (Whittaker 1956), and is therefore, rich in native species diversity.

Roads constitute a form of disturbance in and around the park. Most roads near GRSM are close to the park's border; however, Newfoundgap Road dissects the park and several smaller roads also traverse the western part of GRSM. Most roads are two-lane, although the section of Interstate 40 at the North Carolina-Tennessee border is less than two kilometers from the Park's border. Further, several road-widening projects creating four lane highways have recently been completed or are currently underway in the study area.

### 3.2.2 Study Species

*Paulownia*, a species native to China, has several characteristics associated with successful invaders (Baker 1974), including copious seed production – as many as 20 million seeds per tree in a year (Tang *et al.* 1980), early age at first reproduction – between eight and 10 years (Carpenter *et al.* 1983), and small wind-dispersed seeds known to disperse long distances (Langdon and Johnson 1994).

*Paulownia* is shade-intolerant and sometimes thought to be drought resistant because it is common on steep rock faces with shallow soil (Carpenter *et al.* 1983, Williams 1993b, Langdon and Johnson 1994). However, other research suggests that *Paulownia* prefers deep soil (Tang *et al.* 1980). Overall, *Paulownia*'s regeneration niche (Grubb 1977) is a high light

environment with little leaf litter (Tang *et al.* 1980). The microsite conditions associated with its regeneration are not found beneath a closed canopy and most populations appear to have a unimodal age distribution (Williams 1993a, *pers. obs.*).

### 3.2.3 Data Set

I mapped all *Paulownia* trees on the roads surrounding GRSM. I surveyed all roads at least two times and most roads three times. *Paulownia* is easy to identify along roads; the species has large, heart-shaped leaves and the fruit is persistent. The species may be confused with *Tilia americana* but at 30 m or less (the width of a road margin) the differences are obvious.

I recorded location information using a handheld GPS unit. For trees that were more than approximately five meters apart, I recorded unique location information. When multiple individuals were found together, I recorded one location value and the number of individuals present at that location.

For each location I also recorded the soil depth around the base of one tree at three places and the slope angle. Soil depth was not necessarily distance to bedrock, but rather the depth of soil until I hit rock. The road survey took place in July and August 2001.

I used a database of eight GIS habitat variables to characterize the habitat conditions of sites with *Paulownia* (Table 3.1). I chose these variables based on availability and previous use in modeling species distributions in GRSM (Wilds 1996, Boetsch 2000, van Manen *et al.* 2002). All GIS coverages were continuous variables based on 30 meter pixels. Further, I collected soil depth and slope angle from random locations along roads; I chose random distances to sample along roads and used the vehicle odometer to determine when to stop. For soil depth, I used the average of three measurements from a random roadside site.

### 3.2.4 Univariate Randomization Tests

I used randomization tests written for ArcView Version 3.2 (ESRI, Redlands, CA) to test whether sites with *Paulownia* had a particular signature with respect to landscape level variables. I asked whether *Paulownia* is found at sites with significantly lower elevation and moisture availability (*i.e.*, lower TRMI, lower TCI, higher RSP) compared to random roadside sites. Further, I asked whether *Paulownia* is found at sites with significantly higher solar insolation (*i.e.*, BTA, TSI, LFI). Further, given *Paulownia*'s success on steep rock faces, I expected that *Paulownia* would be found at roadside sites with greater topographic complexity (*i.e.*, SWI).

To run a randomization test I identified a chosen number of random points adjacent to roads. I set the number of random points equal to the number of unique sites with *Paulownia* (*i.e.*, 261). Next, I calculated the average value of a chosen variable for all of the random points. These two steps were repeated 1500 times giving 1500 sets of random points each having an associated mean value of the independent variables. I ran the randomization test for all eight independent variables (Table 3.1).

I compared the average value of a variable for sites with *Paulownia* against the random iterations. Because my hypotheses for *Paulownia* sites versus random points were one-tailed I calculated the *p*-value by dividing the number of observations that were greater than or equal to the observation for invasion sites (or when applicable less than or equal to the observation for invasion sites) by the total number of observations (*i.e.*, 1501). Because I was using variations of the same dataset repeatedly, I used a Bonferroni correction to interpret the *p*-value. The Bonferroni correction is conservative and recalculates alpha by dividing the number of times a dataset is used. For this research I used an alpha of 0.05 and

used the dataset 62 times. This means that  $p$ -values less than 0.0008 (*i.e.*, 0.05/62) were significant at the 0.05 level.

### 3.2.5 Evaluation of Field Variables

I used two sample t-tests in S-plus (Insightful Corp., Seattle, WA) to test whether *Paulownia* is found at sites with significantly lower soil depth and significantly higher slope angle compared to random sites. Further, I generated box plots to compare the means between random sites and sites known to have *Paulownia*.

### 3.2.6 Habitat Suitability Model

To assess the favorability of roadside habitat for *Paulownia* I generated a habitat model using the Mahalanobis Distance statistic ( $D^2$ ) (Rao 1952, Clark *et al.* 1993, Knick and Dyer 1997, Corsi *et al.* 1999, Boetsch 2000, van Manen *et al.* 2002). Mahalanobis Distance is a multivariate technique that only requires presence data. Presence only approaches are ecologically valid for studying invasion patterns, given that sites without invasives are not necessarily unsuitable habitat, rather they may be sites that have yet to be colonized.  $D^2$  is a dimensionless statistic calculated as follows:

$$D^2 = (\underline{x} - \hat{\underline{u}})' \Sigma^{-1} (\underline{x} - \hat{\underline{u}})$$

where  $\underline{x}$  is a vector of habitat characteristics for each cell in the GIS grid,  $\hat{\underline{u}}$  is the mean vector of habitat characteristics of the invasion sites, and  $\Sigma^{-1}$  is the inverse of the variance-covariance matrix calculated from the invasion sites. The statistic represents the standard squared distance between a set of sample variates,  $\underline{x}$ , and “ideal” habitat, represented by  $\hat{\underline{u}}$ . Because  $D^2$  is calculated using the variance-covariance matrix, collinearity among independent variables is not an issue.

$D^2$  gives a measure of dissimilarity such that low values of  $D^2$  are most similar to the sample sites (*i.e.*, sites with *Paulownia*) and high values of  $D^2$  have the highest dissimilarity to the sample sites. I calculated a value of  $D^2$  for each 30 m pixel within 500 meters of roads near GRSM using the Grid function in ArcInfo version 8.1 (ESRI, Redlands, CA). I used S-plus version 6.0 (Insightful Corp., Seattle) to calculate the covariance matrix for all independent variables in the model and to invert the covariance matrix. I only included those variables that were significant in the randomization tests in the habitat model.

### 3.2.7 Detecting Pattern of Spread

I used nearest neighbor distance (NND) to assess *Paulownia*'s pattern of spread along roads. Invasion along roads may be conceptualized in two ways (Baker 1974, Baker 1986, Parendes 1997); plants may spread in a stream-like pattern (*i.e.*, as an advancing wave front that is channeled by the roadside corridor) (Carey 1996) or in a satellite pattern such that invasion is characterized by the establishment of populations distant from the initial invasion point (Bergelson *et al.* 1993, Perrins *et al.* 1993, Buell *et al.* 1995, Clark *et al.* 1998, Nuzzo 1999) (Figure 3.2).

Using NND to detect the pattern of spread one would expect that a stream-like pattern of spread would result in consistently small nearest neighbor distances, whereas a satellite pattern of spread would result in at least some large values for nearest neighbor distances.

I used the distribution of *Paulownia* along routes 19, 28 and 129, because there is little development and these roads are close to the Park border. Further, *Paulownia* is common from Bryson City, NC to the western edge of GRSM near the Foothills Parkway.

Because satellite populations are likely more than one individual, I calculated NND for patches of *Paulownia* such that locations with multiple *Paulownia* were treated as one

site or patch. I calculated the distance to the nearest patch using Network Analyst in ArcView Version 3.2 (ESRI, Redlands, CA). I used distance along the highway rather than Euclidean distance because it was more appropriate ecologically. Finally, I built a cumulative density function (CDF) using the NND between *Paulownia* and random points and between two *Paulownia* patches.

### **3.3 Results**

#### 3.3.1 Distribution of *Paulownia*

The distribution of *Paulownia* along roads surrounding GRSM was not continuous (Figure 3.3). Invasion was heavy at the North Carolina-Tennessee border near I-40 and the eastern part of GRSM. I recorded 38 percent of all identified trees within five kilometers of the NC-TN border at I-40. Another band of invasion was found on the western part of the park from Bryson City, NC to the Foothills Parkway where I recorded 58 percent of all *Paulownia* trees.

*Paulownia* was found at lower elevations and at higher levels of solar insolation (*i.e.*, LFI, TSI) than random points (Table 3.2). *Paulownia* was also found at sites with higher moisture availability (*i.e.*, TCI, RSP) and at sites with higher topographic complexity (*i.e.*, SWI). Aspect and TRMI were not significantly different between random sites and sites with *Paulownia* (Table 3.2).

Comparing soil depth between random sites and sites with P shows that soil depth was significantly lower near *Paulownia* (Figure 3.4, Table 3.3). Slope angle was not significantly different between random roadside sites and those with *Paulownia* (Figure 3.5, Table 3.4).

### 3.3.2 Habitat Model for *Paulownia*

The habitat model built using Mahalanobis Distance showed that sites with *Paulownia* ranged in  $D^2$  from 12 to 345. Ninety-five percent of sites had a  $D^2$  below 125 and on average *Paulownia* was found at a  $D^2$  of 60. In contrast, all sites near roads ranged in  $D^2$  from 0 to 2,923 with an average  $D^2$  of 118 for all roadside habitat (Table 3.5)

The distribution of *Paulownia* habitat along roads shows that all roadsides are not the same from the perspective of *Paulownia*. Viewing roadsides above and below  $D^2$  of 125 (*i.e.*, 95<sup>th</sup> percentile of  $D^2$  for *Paulownia* sites) some patchiness in habitat suitability is evident (Figure 3.6). Further, viewing roadsides as greater than or less than the mean  $D^2$  value for *Paulownia* (*i.e.*, 60) shows an even patchier distribution of habitat (Figure 3.7).

### 3.3.3 Invasion Pattern of *Paulownia*

Most *Paulownia* trees are found very close to conspecifics (*i.e.*, on average 28 meters apart) and patches usually consist of multiple individuals. The distance between patches of *Paulownia* ranged between 4 and 5,625 meters. On average patches were 437 meters apart. In contrast, the distance between *Paulownia* and random points along roads ranged between 4 and 8,172 meters and on average were 1,698 meters apart.

The cumulative frequency for *Paulownia* patches and random sites shows evidence of clumping among *Paulownia* in that the curve for *Paulownia* is skewed towards lower distances between patches (Figure 3.8). However, I also observed large distances between patches, which provide support for a satellite pattern of spread in this species. I would expect that a stream-like pattern of spread would show only small distances between patches and would not show evidence of distinct patches, which I observed.

### 3.4 Discussion

#### 3.4.1 Drought tolerance in *Paulownia*

Not all roadside habitat is equivalent with respect to *Paulownia*. The research is unclear as to whether *Paulownia* is drought-tolerant (Carpenter *et al.* 1983, Williams 1993a, Williams 1993b, Langdon and Johnson 1994). Evidence for this hypothesis comes from the fact this species has been observed growing on steep, exposed rock faces (Williams 1993b, *pers. obs.*) and has colonized xeric forests burned as part of restoration efforts for *Pinus pungens* habitat (K. Johnson, GRSM, *pers. comm.*, Langdon and Johnson 1994). However, I did not find support for this hypothesis. In fact, I found that *Paulownia* was found at sites with significantly higher moisture availability (*i.e.*, higher TCI, and lower RSP) and showed no relationship with TRMI.

Grading associated with road building may increase moisture availability along roads. In some instances, this may have been the case for sites with *Paulownia*, although approximately 44 percent of the *Paulownia* I observed while completing the road survey were growing on exposed rock. The relationship between rocks and moisture is unclear. *Paulownia* unquestionably does well on exposed rock, a niche with little competition from other species, but equating exposed rock with drought-tolerance is speculative. One argument to the contrary is that *Paulownia* does not have hearty, small leaves characteristic of drought-tolerant species. Further, rocks may act to contain water and minimize evaporation, which may increase moisture availability for plants that are able to grow at these sites. Data on *Paulownia*'s rooting depth and experimentation related to drought-tolerance are needed to shed more light on this species drought tolerance.

### 3.4.2 Soil Preferences of *Paulownia*

The research is ambiguous as to *Paulownia*'s soil preferences (Tang *et al.* 1980, Carpenter *et al.* 1983, Williams 1993a, Williams 1993b, Langdon and Johnson 1994). *Paulownia* clearly can tolerate shallow soils and was correlated with rocks according to this study. The question of how soil nutrients affect the distribution of this species remains unanswered. It was prohibitive to include field variables in the habitat model derived using Mahalanobis distance. Modeling habitat using the Mahalanobis distance statistic is a predictive tool that is powerful largely because it can be extrapolated to a landscape scale; this approach represents a tradeoff between specificity and extent.

### 3.4.3 Pattern of Spread

I found evidence of a patchy or satellite pattern of spread in *Paulownia*, although it is possible that over time the patches will coalesce and appear as more of a stream-like pattern. However, I also found evidence that roadside habitat is not continuously suitable, which is likely due to the topographic complexity of the landscape. If the mechanism driving the satellite pattern is the patchiness of suitable habitat it means that areas between patches may not be invaded over time.

Irrespective of whether patches coalesce over time, focusing control efforts on large, obvious populations of *Paulownia* at the expense of smaller populations may make this species more difficult to control as a faster rate of spread is associated with a long distance dispersal (Clark *et al.* 1998). Using a model of invasive spread, Moody and Mack (1988) found that treating even 30 percent of satellite populations greatly improved overall effectiveness of control measures. This was true in part because dispersal from a large

invasion site is more likely to land in an invaded area, whereas dispersal from small foci is more likely to land in neighboring uninfested sites (Moody and Mack 1988).

Interpretation of the pattern of spread is subject to a bias given that non-roadside populations were not included in this analysis. Although this may change the picture somewhat, other research (Williams 1993a, Williams 1993b, Langdon and Johnson 1994) and information from a GRSM database used to track locations of known infestations does not show that *Paulownia* is common in forested areas. Further, I never observed *Paulownia* in undisturbed areas bordering existing populations or even in gaps that were easily within dispersal range of hundreds of flowering individuals. Although *Paulownia* has colonized burn scars in the Park (Kristine Johnson, GRSM, *pers. comm.*), its presence at these remote sites will likely provide more support for the satellite pattern of spread.

#### 3.4.4 Relevance of Habitat Suitability Model

Not all roadside habitat is ideal for invasion by *Paulownia*. This suggests that roadside corridors are not necessarily continuous habitat, which is particularly true for roads that pass through mountainous terrain. Unlike, conservation of rare species where habitat connectivity is sought, a patchy distribution of suitable habitat is advantageous when controlling spread of invasive species as it should slow the rate of percolation across the landscape (Turner 1989, Turner *et al.* 1989, Bergelson *et al.* 1993). However, species that can disperse easily through a patchy distribution of suitable sites will not be slowed in their rate of spread. It is possible that the patchiness of suitable habitat has little effect on *Paulownia*, with its many small windblown seeds.

### 3.4.5 Microsite Conditions along Roads

The GIS data used to derive the habitat model and in the randomization tests are 30 m resolution and are largely based on a digital elevation model that does not account for how roads may change microsite conditions. Roads clearly affect available sunlight. Further roads may also cause changes in moisture and nutrient availability. Field data on microsite conditions along roads would yield a more unequivocal picture of how roadside conditions vary.

### **3.5 Future research**

This research has raised additional questions with regard to the distribution of *Paulownia* along roads. Although it appears that *Paulownia* is spreading in a satellite pattern time-series data would help clarify this question. Further, it would be interesting to address whether a patchy distribution of suitable habitat matters for the rate of *Paulownia* invasion given its dispersal capabilities. Comparing the rate of spread along roads in mountainous (*i.e.*, patchy habitat) and flat (*i.e.*, higher connectivity) terrain could give some indication of how the distribution of suitable habitat affects the rate of spread. Finally, follow-up field research to Moody and Mack's (1988) model on the effect of controlling satellite populations on the rate of spread would be relevant to conservation efforts.

The current study shows evidence that *Paulownia* is associated with sites that have higher moisture availability. This is in contrast to hypotheses put forth by other research (Carpenter *et al.* 1983, Williams 1993b, Langdon and Johnson 1994). Experimental data on *Paulownia*'s drought-tolerance would be helpful in settling this issue. Further, an investigation of how *Paulownia* meets its moisture requirements on steep, exposed rock faces could shed light on whether it is drought tolerant.

Finally, the hypothesis that roadsides are not necessarily continuous habitat could be explored further using field data that measures microsite conditions along roads.

Table 3.1. Variables used to characterize habitat of *Paulownia tomentosa* along roads near Great Smoky Mountains National Park (GRSM).

<b>Variable</b>	<b>Description <sup>1</sup></b>	<b>Range in GRSM</b>	<b>Source</b>
Aspect (BTA)	Aspect transformed using: $1 + \cos(45 - \text{aspect})$	0.0-2.0	Beers <i>et al.</i> 1966
Elevation	Digital Elevation Model (DEM)	266-2,027	U.S.G.S. DEM Data
Landform Index (LFI)	Meso-scale topographic exposure	-4-75	McNab (1993)
Relative Slope Position (RSP)	Relative Slope Position (%)	0-100	Wilds (1996)
Shannon-Weiner Index (SWI)	Index of topographic complexity	13-35	Miller (1986)
Terrain Shape Index (TSI)	Index of micro-scale topographic exposure	-171-130	McNab (1989)
Topographic Convergence Index (TCI)	Simulates the flow accumulation of water; $TCI = \ln(A / \tan B)$ , where A is drained surface area and B is drained surface slope	8-200	Beven and Kirkby (1979), Wolock and McCabe (1995), Halpin (1995),
Topographic Relative Moisture Index (TRMI)	Index of moisture considering the effects of slope position, aspect, and elevation	0-91	Parker 1882

<sup>1</sup> All variables were continuous with spatial resolutions of 30 m.

Table 3.2. Results of randomization tests for *Paulownia tomentosa* along roads.

<b>Variable</b>	<b><i>p</i>-value<sup>1</sup></b>	<b>Interpretation<sup>2</sup></b>	<b>Average Value for <i>Paulownia</i> Sites</b>
BTA	0.0500	No relationship	1.085
Elevation	<b>0.0007</b>	Lower than expected	506.992
LFI	<b>0.0007</b>	Higher than expected	24.794
RSP	<b>0.0007</b>	Lower than expected	54.635
SWI	<b>0.0007</b>	Higher than expected	28.211
TCI	<b>0.0007</b>	Higher than expected	74.096
TRMI	0.7548	No relationship	26.494
TSI	<b>0.0007</b>	Higher than expected	-18.020

<sup>1</sup> *P*-values less than 0.0008 are significant with 95 percent confidence and are in bold. Note that *p*-values are the same because most variables had only one observation (*i.e.*, the average for sites with *Paulownia*) outside the range of expected values.

<sup>2</sup> Refers to sites with *Paulownia*.

Table 3.3. Two Sample *t*-test comparing the means for soil depth (cm) between random points and sites with *Paulownia tomentosa*.

<b>Group</b>	<b>N</b>	<b>Mean</b>	<b>Std. Dev.</b>	<b>Std. Error</b>
Random Points	234	11.93	7.26	0.47
<i>Paulownia tomentosa</i>	180	3.09	5.86	0.44

*t*-statistic 13.0698, 411 *df*, *p* = 0.0001

Table 3.4. Two Sample *t*-test comparing the means for slope angle (°) between random points and sites with *Paulownia tomentosa*.

<b>Group</b>	<b>N</b>	<b>Mean</b>	<b>Std. Dev.</b>	<b>Std. Error</b>
Random Points	234	31.34	15.63	1.02
<i>Paulownia tomentosa</i>	192	32.26	21.86	1.58

*t*-statistic -0.488, 366 *df*, *p* < 0.3129

Table 3.5. Range of  $D^2$  values for all sites within 500 meters of roads and sites with *Paulownia tomentosa*.

	<b>Number of Sites</b>	<b>Min</b>	<b>Max</b>	<b>Range</b>	<b>Mean</b>	<b>Std</b>	<b>Median</b>
All Sites	888,359	0	2,923	2,923	117.8	115.2	77
Sites with <i>Paulownia</i> <sup>1</sup>	246	12	345	333	59.5	36.4	53

<sup>1</sup> Some sites do not have a  $D^2$  value because of missing data for one or more variables. This resulted in 15 of 261 sites with *Paulownia tomentosa* having no value for  $D^2$ .

Figure 3.1. Distribution of roads surrounding Great Smoky Mountains National Park.

Roadside habitat near GRSM was the study area for investigating invasion of *Paulownia tomentosa* along roads.

Insert Figure 3.1

Figure 3.2. Conceptual models of invasion patterns along corridors. Model A shows a continuous pattern of spread such that invasion proceeds as an advancing wave front. Model B shows a satellite or patchy pattern of spread such that new invasion sites occur at a distance from existing invasion sites.

Insert Figure 3.2

Figure 3.3. Distribution of sites colonized by *Paulownia tomentosa* along roads near Great Smoky Mountains National Park. Section A is the portion of GRSM closest to Interstate 40 and included 38 percent of all recorded *Paulownia* along roads. Section B from Bryson City, NC to Foothills Parkway in Tennessee included 58 percent of all recorded *Paulownia* along roads.

Insert Figure 3.3

Figure 3.4. Box-plot comparing the mean of soil depth at random sites and sites with *Paulownia tomentosa*. Soil depth (cm) was averaged from three measurements at base of tree or a random site and was based on depth of soil to rock (not necessarily bedrock). Sites with *Paulownia* had significantly lower soil depth compared to random sites ( $t$ -statistic 13.0698, 411  $df$ ,  $p = 0.0001$ )

Insert Figure 3.4

Figure 3.5. Box-plot comparing the mean of slope angle at random sites and sites with *Paulownia tomentosa*. Sites with *Paulownia* did vary significantly in slope angle (°) from random sites ( $t$ -statistic -0.488, 366  $df$ ,  $p < 0.3129$ ).

Insert figure 3.5

Figure 3.6. Distribution of roadside habitat similar to sites invaded by *Paulownia tomentosa* ( $D^2 < 125$ ). According to the Mahalanobis distance roadside habitat model 95 % of sites with *Paulownia* had  $D^2$  value below 125; these sites are shaded in black.

Insert figure 3.6

Figure 3.7. Distribution of roadside habitat similar to sites invaded by *Paulownia tomentosa* ( $D^2 < 60$ ). According to the Mahalanobis distance roadside habitat model sites with *Paulownia* had a mean  $D^2$  value of 60; these sites are shaded in black.

Insert figure 3.7

Figure 3.8. Cumulative frequency curve of nearest neighbor distance (NND) between patches with *Paulownia tomentosa* and random sites and nearest *Paulownia* patch. NND was defined as distance along roads instead of Euclidean distance. Curve for NND between *Paulownia* patches is to the left of curve for random sites indicating some degree of clustering of *Paulownia* patches. However, some patches are distance from other patches suggesting that this species exhibits a satellite or patchy pattern of spread.

Insert figure 3.8

## REFERENCES

- Andow, D.A., P.M. Kareiva, S.A. Levin and A. Okubo. 1990. Spread of invading organisms. *Landscape Ecology* 4(2/3): 177-188.
- Baker, H.G. 1974. The evolution of weeds. *Annual Review of Ecology and Systematics* 5: 1-24.
- Baker, H.G. 1986. Patterns of plant invasion in North America. Pages 44-57 in H.A. Mooney and J.A. Drake editors. Ecology of biological invasions of North America and Hawaii. Springer-Verlag, New York, New York, USA.
- Baron, J., C. Dombrowski, and S.P. Bratton. 1975. The status of five woody plants in the Tennessee District, Great Smoky Mountains National Park. Management Report No. 2, U.S. Department of the Interior, National Park Service, Great Smoky Mountains National Park, Gatlinburg, Tennessee. 26 pages.
- Beers, T.W., P.E. Dress, and L.C. Wensel. 1966. Aspect transformation in site productivity research. *Journal of Forestry* 64: 691-692.
- Bergelson, J., J.A. Newman, and E.M. Floresroux. 1993. Rates of weed spread in spatially heterogeneous environments. *Ecology* 74(4): 999-1011.
- Beven, K.J. and M.J. Kirkby. 1979. A physically based, variable contributing area model of basin hydrology. *Hydrological Science Bulletin* 24: 43-69.
- Boetsch, J.R. 2000. Modeling the occurrence of rare plant populations at the landscape scale. MS Thesis. University of North Carolina, Chapel Hill, USA.
- Brothers, T.S. and A. Spingarn. 1992. Forest fragmentation and alien plant invasion of central Indiana old-growth forests. *Conservation Biology* 6(1): 91-100.
- Brown, J.R. and J. Carter. 1998. Spatial and temporal patterns of exotic shrub invasion in an Australian tropical grassland. *Landscape Ecology* 13: 93-102.
- Brown, J.S. and A.C. Gibson. 1983. Biogeography. C.V. Mosby, St. Louis, MO, USA.
- Buell, A.C., A.J. Pickart, and J.D. Stuart. 1995. Introduction history and invasion patterns of *Ammophila arenaria* on the north coast of California. *Conservation Biology* 9(6): 1587-1593.
- Burke, M.J.W. and J.P. Grime. 1996. An experimental study of plant community invasibility. *Ecology* 77(3): 776-790.
- Cadenasso, M.L. and S.T.A. Pickett. 2001. Effect of edge structure on the flux of species into forest interiors. *Conservation Biology* 15(1): 91-97.

- Carey, J.R. 1996. The incipient Mediterranean fruit fly population in California: implications for invasion biology. *Ecology* 77(6): 1690-1697.
- Carpenter, S.B., M.J. Immel, and N.D. Smith. 1983. Effect of photoperiod on the growth and photosynthetic capacity of *Paulownia* seedlings. *Castanea* 48(1): 13-18.
- Clark, J.D., J.E. Dunn, and K.G. Smith. 1993. A multivariate model of female black bear habitat for use in a geographic information system. *Journal of Wildlife Management* 57(3): 519-526.
- Clark, J.S., C. Fastie, G. Hurtt, S.T. Jackson, C. Johnson, G.A. King, M. Lewis, J. Lynch, S. Pacala, C. Prentice, E.W. Schupp, T. Webb III, and P. Wycoff. 1998. Reid's paradox of rapid plant migration. *BioScience* 48(1): 13-24.
- Corsi, F, E. Duprè, and L. Boitani. 1999. A large-scale model of wolf distribution in Italy for conservation planning. *Conservation Biology* 13(1): 150-159.
- D'Antonio, C.M. 1993. Mechanisms controlling invasion of coastal plant communities by the alien succulent *Carpobrotus edulis*. *Ecology* 74(1): 83-95.
- DeFerrari, C.M. and R.J. Naiman. 1994. A multi-scale assessment of the occurrence of exotic plants on the Olympic Peninsula, Washington. *Journal of Vegetation Science* 5: 247-258.
- Denslow, J.S. 1980. Patterns of plant diversity during succession under different disturbance regimes. *Oecologia* (Berl.) 46: 18-21.
- Dice, L. R. and Leraas, H. J. 1936. A graphic method for comparing several sets of measurements. Publications from the Laboratory for Vertebrate Zoology 3, University of Michigan Museum of Zoology. 3 pages.
- Forman, R.T.T. and L.E. Alexander. 1998. Roads and their major ecological effects. *Annual Review of Ecology and Systematics* 29: 207-231.
- Grubb, P.J. 1977. The maintenance of species-richness in plant communities: the importance of the regeneration niche. *Biological Reviews* 52: 107-145.
- Halpin, P.N. 1995. A cross-scale analysis of environmental gradients and forest patterns in the Giant Sequoia – Mixed Conifer forest of the Sierra Nevada. Ph.D. Dissertation, University of Virginia, Charlottesville, USA.
- Hester, A.J. and R.J. Hobbs. 1992. Influence of fire and soil nutrients on native and non-native annuals at remnant vegetation edges in the Western Australian wheatbelt. *Journal of Vegetation Science* 3: 101-108.

- Hobbs, R.J. and L.F. Huenneke. 1992. Disturbance, diversity, and invasion: implications for conservation. *Conservation Biology* 6(3): 324-337.
- Hood, W.G. and R.J. Naiman. 2000. Vulnerability of riparian zones to invasion by exotic vascular plants. *Plant Ecology* 148: 105-114.
- Hosmer, D.W. Jr. and S. Lemeshow. 1989. Applied Logistic Regression. John Wiley and Sons, New York, New York, USA.
- Huenneke, L.F., S.P. Hamburg, R. Koide, H.A. Mooney, P.M. Vitousek. 1990. Effects of soil resources on plant invasion and community structure in Californian serpentine grassland. *Ecology* 71(2): 478-491.
- Johnson, L.B., 1900. Analyzing spatial and temporal phenomena using geographical information systems. *Landscape Ecology* 4(1): 31-43.
- Johnstone, I.M. 1986. Plant invasion windows: a time-based classification of invasion potential. *Biological Reviews* 61: 369-394.
- Jones, Brian. 1988. Biostatistics in paleontology. *Geoscience Canada* 15(1): 3–22.
- King, P.B. and A. Stupka. 1950. The Great Smoky Mountains – their geology and natural history. *Scientific Monthly* 71: 31-43.
- Knapp, L.B. and C.D. Canham. 2000. Invasion of old-growth forest in New York by *Ailanthus Altissima*: sapling growth and recruitment in canopy gaps. *Journal of the Torrey Botanical Society* 127(4): 307-315.
- Knick, S.T. and D.L. Dyer. 1997. Distribution of black-tailed jackrabbit habitat determined by GIS in southwester Idaho. *Journal of Wildlife Management* 61(1): 75-85.
- Langdon, K.R. and K.D. Johnson. 1994. Additional notes on invasiveness of *Paulownia tomentosa* in natural areas. *Natural Areas Journal* 14(2): 139-140.
- Lesica, P. and S. Miles. 1999. Russian olive invasion into cottonwood forests along a regulated river in north-central Montana. *Canadian Journal of Botany* 77: 1077-1083.
- Lodge, D. M., R.A. Stein, K.M. Brown, A.P. Covich, C. Bronmark, J.E. Garvey, and S.P. Klosiewski. 1998. Predicting impact of freshwater exotic species on native biodiversity: Challenges in spatial scaling. *Australian Journal of Ecology* 23(1): 53-67.
- McCay, D.H. 2001. Spatial patterns of sand pine invasion into longleaf pine forests in the Florida Panhandle. *Landscape Ecology* 16: 89-98.
- MacArthur, R. and E.O. Wilson. 1967. The Theory of Island Biogeography. Princeton University Press, Princeton, NJ, USA.

- McNab, W.H. 1989. Terrain Shape Index: quantifying effect of minor landforms on tree height. *Forest Science* 35: 91-104.
- McNab, W.H. 1993. A topographic index to quantify the effect of mesoscale landform on site productivity. *Canadian Journal of Forest Research* 23: 1100-1107.
- Meekins, J.F. and B.C. McCarthy. 2001. Effect of environmental variation on the invasion success of a non-indigenous forest herb. *Ecological Applications* 11(5): 1336-1348.
- Miller, R. 1986. Predicting rare plant distributions patterns in the southern Appalachians of the southeastern USA. *Journal of Biogeography* 13: 293-311.
- Moody, M.E. and R.N. Mack. 1988. Controlling the spread of plant invasions: the importance of nascent foci. *Journal of Applied Ecology* 25(3): 1009-1021.
- Morgan, J.W. 1998. Patterns of invasion of an urban remnant of a species-rich grassland in southeastern Australia by non-native plant species. *Journal of Vegetation Science* 9: 181-190.
- Nuzzo, V. 1999. Invasion pattern of the herb garlic mustard (*Alliaria petiolata*) in high quality forests. *Biological Invasions* 1: 169-179.
- Parendes, L.A. 1997. Spatial patterns of invasion by exotic plants in a forested landscape. Ph.D. Dissertation. Oregon State University, Corvallis, USA.
- Parendes, L.A. and J.A. Jones. 2000. Role of light availability and dispersal in exotic plant invasion along roads and streams in the H.J. Andrews Experimental Forest, Oregon. *Conservation Biology* 14(1): 64-75.
- Parker, A.J. 1982. The topographic relative moisture index: an approach to soil-moisture assessment in mountainous terrain. *Physical Geography* 3: 160-168.
- Perrins, J., A. Fitter, and M. Williamson. 1993. Population biology and rates of invasion of three introduced *Impatiens* species in the British Isles. *Journal of Biogeography* 20: 33-44.
- Planty-Tabacchi, A.M., E. Tabacchi, R.J. Naiman, C. DeFerrari, and H. Decamps. 1996. Invasibility of species-rich communities in riparian zones. *Conservation Biology* 10(2): 598-607.
- Pyle, C. 1985. Vegetation disturbance history of Great Smoky Mountains National Park: An analysis of archival maps and records. U.S. Department of the Interior, National Park Service, Research/Resources Management Report SER-77, Gatlinburg, Tennessee, USA.

- Pyle, C. 1988. The type and extent of anthropogenic vegetation disturbance in the Great Smoky Mountains before National Park Service acquisition. *Castanea* 53: 225-235.
- Pysek, P. and K. Prach. 1993. Plant invasions and the role of riparian habitats: a comparison of four species alien to central Europe. *Journal of Biogeography* 20: 413-420.
- Radford, A.E., H.E. Ahles, and C.R. Bell. 1968. Manual of the vascular flora of the Carolinas. The University of North Carolina Press, Chapel Hill, North Carolina, USA.
- Rao, C.R. 1952. Statistical methods in biometric research. J. Wiley and Sons, New York, New York.
- Reichard, S.H. and C.W. Hamilton. 1997. Predicting invasions of woody plants introduced into North America. *Conservation Biology* 11(1): 193-203.
- Rejmanek, M. and D.M. Richardson. 1996. What attributes make some plant species more invasive? *Ecology* 77(6): 1655-1661.
- Rouget, M., D.M. Richardson, S.J. Milton, and D. Polakow. 2001. Predicting invasion dynamics of four alien *Pinus* species in a highly fragmented semi-arid shrubland in South Africa. *Plant Ecology* 152: 79-92.
- Rundel, P.W. and S.B. Sturmer. 1998. Native plant diversity in riparian communities of the Santa Monica Mountains, California. *Madrono* 45(2): 93-100.
- Schonewald-Cox, C. and M. Buechner. 1992. Park protection and public roads. Pages 375-395 in Fiedler P.L. and S. K. Jain, editors. Conservation biology: The theory and practice of native conservation, preservation and management. Chapman and Hall, New York, New York, USA.
- Shanks, R.E., 1954. Climates of the Great Smoky Mountains. *Ecology* 35(3): 354-361.
- Sokal, Robert R. 1965. Statistical methods in systematics. *Biological Reviews* 40: 337-391.
- Stohlgren, T.J., D. Binkley, G.W. Chong, M.A. Kalkan, L.D. Schell, K.A. Bull, U. Otsuki, G. Newman, M. Bashkin, and Y. Son. 1999a. Exotic plant species invade hot spots of native plant diversity. *Ecological Monographs* 69(1): 25-46.
- Stohlgren, T.J., L.D. Schell and B.V. Heuvel. 1999a. How grazing and soil quality affect native and exotic plant diversity in Rocky Mountain grasslands. *Ecological Applications* 9(1): 45-64.
- Tang, R.C., S.B. Carpenter, R.F. Wittwer and D.H. Graves. 1980. *Paulownia* – A crop tree for wood products and reclamation of surface-mined land. *Southern Journal of Applied Forestry* 4: 19-24.

- ter Braak, C.J.F. and C.W.N. Looman. 1995. Regression. *In: Data Analysis in Community and Landscape Ecology*. R.H.G. Jongman, C.F.J. ter Braak, and O.F.R. van Tongeren, editors. Cambridge University Press.
- Thompson, K., J.G. Hodgson, J.P. Grime and M.J. Burke. 2001. Plant traits and temporal scale: evidence from a 5-year invasion experiment using native species. *Journal of Ecology* 89: 1054-1060.
- Thornewaite, C.W. 1948. An approach toward a rational classification of climate. *Geog. Review* 38: 55-94.
- Tilman, D. 1997. Community invasibility, recruitment limitation, and grassland biodiversity. *Ecology* 78(1): 81-92.
- Turner, M.G. 1989. Landscape Ecology: The effect of Pattern on Process. *Annual Review of Ecology and Systematics* 20: 171-197.
- Turner, M.G., R.H. Gardner, V.H. Dale, and R.V. O'Neill. 1989. Predicting the spread of disturbance across heterogeneous landscapes. *OIKOS* 55: 121-129.
- Tyser, R.W. and C.A. Worley. 1992. Alien flora in grasslands adjacent to road and trail corridors in Glacier National Park, Montana (U.S.A.). *Conservation Biology* 6(2): 253-262.
- van Manen, F.T., J.D. Clark, S.E. Schlarbaum, K. Johnson, and G. Taylor. 2002. A model to predict the occurrence of surviving butternut trees in the southern Blue Ridge mountains. Pages 491-497 in J.M. Scott, P.J. Heglund, M.L. Morrison, J.B. Haufler, M.G. Raphael, W.A. Wall, and F.B. Samson, editors. Predicting species occurrences: Issues of accuracy and scale. Island Press, Washington D.C., USA.
- Whittaker, R.H. 1956. Vegetation of the Great Smoky Mountains. *Ecological Monographs* 26(1): 1-80.
- Wilds, S. P. 1996. Gradient analysis of the distribution of flowering dogwood (*Cornus florida* L.) and dogwood anthracnose (*Discula destructiva* Redlin.) in western Great Smoky Mountains National Park. M.A. Thesis. University of North Carolina, Chapel Hill, USA.
- Williams, C.E. 1993a. Age structure and importance of naturalized *Paulownia tomentosa* in a central Virginia streamside forest. *Castanea* 58(4): 243-249.
- Williams, C.E. 1993b. The exotic empress tree, *Paulownia tomentosa*: an invasive pest of forests? *Natural Areas Journal* 13(3): 221-222.
- Wiser, S.K., R.B. Allen, P.W. Clinton, and K.H. Platt. 1998. Community structure and forest invasion by an exotic herb over 23 years. *Ecology* 79(6): 2071-2081.

- Wolock, D.M. and G.J. McCabe, Jr. 1995. Comparison of single and multiple flow direction algorithms for computing topographic parameters in TOPMODEL. *Water Resources Research* 31: 1315-1324.
- Woods, K.D. 1993. Effects of invasion by *Lonicera tatarica* L. on herbs and tree seedlings in four New England forests. *American Midland Naturalist*. 130(1): 62-74.
- Zalba, S.M., M.I. Sonaglioni, C.A. Compagnoni, and C.J. Belenguer. 2000. Using a habitat model to assess the risk of invasion by an exotic plant. *Biological Conservation* 93(2): 203-208.
- Zink, T.A., M.F. Allen, B. Heindl-Tenhunen, and E.B. Allen. 1995. The effect of a disturbance corridor on an ecological reserve. *Restoration Ecology* 3(4): 304-310.