IMPACTS OF URBANIZATION AND LANDSCAPE CHANGE ON NATIVE PLANT DIVERSITY AND THE DISTRIBUTION OF EXOTIC FOREST INVADERS

by

Amy Davis

A dissertation submitted to the faculty of
The University of North Carolina at Charlotte
in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in
Geography and Urban Regional Analysis

Charlotte

2014

Approved by:

____
Dr. John Kupfer

____
Dr. Ming- Chun Lee

____
Dr. Ross Meentemeyer

____
Dr. Jean-Claude Thill

ABSTRACT

AMY DAVIS. Impacts of urbanization and landscape change on native plant diversity and the distribution of exotic forest invaders. (Under the direction of JEAN-CLAUDE THILL)

This dissertation addresses three questions focused on enhancing our understanding of the impacts of urbanization and land cover change on patterns of native biodiversity, the distribution of exotic plant species in forests, and the resistance of forest communities to invasion. First, I examine whether human-mediated invasion pressure, quantified using a novel variable, the residential force of invasion (rFOI) can be used to improve species distribution models (SDMs) of exotic forest understory invaders. My results show that human mediated invasion pressure influences the distribution of forest invaders and that including rFOI significantly improves model performance. This research also demonstrates that high invasion pressure substantially increases the risk of invasion in habitats that prior to considering invasion pressure, were identified as unsuitable. Second, using a multi-scalar approach, I investigate the influence of multi-temporal trajectories of forest cover change on patterns of native plant diversity in forests. Temporal trajectories of forest cover change are derived from a longitudinal assessment of forest cover across four time intervals (1938-1956, 1956-1975, 1975-1997, 1997-2009) and classified into distinct types of forest cover change reflecting the timing, extent and nature of the change (deforestation, reforestation or no change). My results show that while measures of forest cover change derived using a single time interval (1938-2009) did not influence native diversity, the typology of multi-temporal forest cover change is a significant explanatory factor of patterns of native diversity, after controlling for other environmental and landscape covariates. Furthermore, from these results, I have identified a specific type of trajectory, “rapid reforestation between 1957 and 1975”, that is associated with the highest native diversity observed within in the study extent. Finally, I investigate the direct and indirect effects of urbanization, environmental factors, and landscape proximity on the ability of forest communities to resist invasions using a structural equation modeling framework. My results support the hypothesis that urbanization has negative direct and indirect effects on invasion resistance via its interaction with propagule pressure and habitat eutrophication, thereby providing a mechanism to explain the increased susceptibility of urban forests to invasions. Taken together, the results of this work provide much needed empirical evidence linking anthropogenic factors with patterns of native biodiversity, the distribution of exotic species and decreased resistance to invasions and show that careful consideration of how anthropogenic factors may influence specific ecological processes can both advance our knowledge of both invasion and urban ecology, as well as result in better predictive models.

DEDICATION

For Kent Davis, without you, this could not have been done.

ACKNOWLEDGMENTS

This dissertation has benefited greatly from the guidance, methodological insights, and contributions of Dr. Jean-Claude Thill. I am also grateful to Dr. Meentemeyer for his intellectual contributions to the topics that structure this dissertation and for providing the resources necessary over the years in order to accomplish this work. The comments and suggestions of John Kupfer also have strengthened this dissertation. The field work completed for this dissertation could not have been completed without the assistance of numerous individuals in the field including Jon Watkins, Emily Henke, Tara Weller, Kent Davis, and Monica Dorning. Jon Watkins has also helped me with the space and logistics of processing hundreds of soil samples. I would also like to thank Stephen Bowman who cheerfully assisted me on numerous occasions with obtaining signatures, navigating paperwork and other requirements. I'd like to thank Dr. Wenwu Tang for providing me a dedicated workstation and the computing resources at the Center for Applied GIS after the departure of Dr. Ross Meentemeyer. This research was funded in part by the UNC Charlotte Graduate Assistant Support Plan, UNC Charlotte Department of Geography BRIDGES Fellowship in Geography, the North Carolina Association of Environmental Professionals (NCAEP) Scholarship and the Renaissance Computing Institute (RENCI).

TABLE OF CONTENTS

CHAPTER 1: INTRODUCTION

Landscape change and biological invasions are two major consequences of urbanization, which in turn have serious ramifications to biodiversity. Habitat loss and fragmentation induced by landscape change, coupled with biological invasions, represent the biggest drivers of global and local biodiversity loss (Hooper et al., 2005). Human quality of life and well-being are intrinsically linked to the ecosystem services provided by biodiversity, such as food, clean air, water quality protection, and recreational opportunities (Diaz et al., 2006; Mace et al., 2012; Cardinale et al., 2012). Since biodiversity facilitates ecosystem resilience and the provisioning of ecosystem services, the maintenance of biodiversity is of paramount importance within the context of landscape sustainability. Currently, about half of the world's populations resides in cities and this number is projected to increase by another 2.5 billion by 2050 (United Nations, 2014). Given our current and projected state of urbanization and the negative impacts urbanization has on biodiversity, how can we manage and protect the biodiversity of forests in urban and urbanizing landscapes in order to sustain the provisioning of ecosystem services they provide? This challenge is especially apparent in the southeastern United States, where biodiversity is threatened by both of these drivers via the transformation of forest and other natural habitats to accommodate the demand for housing and related amenities, and widespread exotic species invasions in forests that limit the recruitment and establishment of native species. Small forest remnants embedded in urban landscapes, however altered they may be, are increasingly being recognized for their existing and potential roles as providing refuge to several bird and wildlife species, and the provisioning of ecosystem services necessary if cities are to be sustainable, including flood mitigation, air pollution removal and recreation (Nowak & Dwyer, 2007; Escobedo et al., 2011; Pataki et al., 2011). However, urban forests may be highly susceptible to biodiversity loss and invasions by exotic species as a result of the impacts of urbanization that manifest themselves primarily through landscape change, eutrophication and higher invasion pressure. Although there have been several studies demonstrating that exotic species richness is correlated with metrics of urbanization or other anthropogenic factors (e.g. housing density, population density, income) it is difficult to draw robust inferences from these data/studies due to being encumbered by one more of the following qualities making them ill-suited to the task. These are 1) very coarse spatial grain which precludes distinguishing habitat types (thus one cannot deduce the impacts of urbanization to invasibility of a specific habitat type) 2) lack of sufficient landscape scale replication in the sampling design; 3) reliance of secondary data such as herbarium records; 4) including heavily altered human dominated habitats such as residential yards, gardens and other areas of primarily ornamental landscaping which obfuscates the impacts of urbanization/anthropogenic factors to forests or other primarily “natural” habitats and the potential effects of spatial autocorrelation are ignored. This dissertation uses a robust, stratified random sampling design in which numerous forests (40) are sampled for plant diversity and invasive species along an urban to rural gradient across the study extent in order to incorporate landscape scale replication and account for the effects of environmental and landscape heterogeneity, as well as spatial dependence to investigate the impacts of urbanization and of landscape change that accompanies human settlements and shifts in production systems on native diversity and on the distribution of exotic species in forests, and how these anthropogenic drivers may influence forest vulnerability to invasions. The rapidly growing Charlotte, North Carolina metropolitan area serves as a case study.

The second chapter investigates whether dispersal and anthropogenic factors can be used to improve the accuracy of spatially explicit models for predicting the distribution of invasive species in forests across the Charlotte metropolitan area. Accurate models are important since the detection of understory invaders using remote sensing based approaches has not been very effective to date. Since the effects of anthropogenic factors on exotic species distributions have yet to be explicitly investigated, the objective of this chapter is to investigate the extent to which human-mediated invasion pressure measured using a novel metric of anthropogenic influence, the residential force of invasion (rFOI), explains invasion patterns in metropolitan forests and whether it notably improves the accuracy of species distribution models (SDM). Based on the premise that single family residences serve as external sources of invasive propagules, the potential rFOI at a given location, *i* is measured by the cumulative sum of the inverse weighted distances from *i,* to every single family house, weighted by age, within the study extent. Although the focus of the research is on the role of human-mediated invasion pressure, we also account for endogenous invasion pressure (local-scale invasion pressure arising from dispersing invaders from nearby sites) as we expect that an SDM applied in an urbanizing landscape is likely to have its predictive accuracy degraded by variations in both sources of invasion pressure, if they are not included. This has important implications for conservation management, as models that do not account for invasion pressure may fail to identify habitats of special concern as being at risk of invasion.

The third chapter presents an approach to classify multi-temporal trajectories of forest cover change derived from a longitudinal assessment of forest cover into a typology of forest cover change; we then propose a multivariate statistical model to test the impacts of this typology on forest native plant diversity. Over time, human-environmental interactions have largely impacted patterns of forest cover resulting in both deforestation and reforestation, in the southeastern United States. This history of human-environment interactions and the intermediate disturbances they generate are unlikely to be represented when measuring forest change using only a single time interval as is commonly done (e.g., Albuquerque & Rueda, 2010). Disturbance is any event that results in the eradication of biomass and the clearing of habitat space thereby permitting new residents to utilize resources (Roxburgh et al., 2004). According to the intermediate disturbance hypothesis, intermediate levels of disturbance can promote the coexistence of species in a community, resulting in higher species richness (Connell, 1978; Roxburg et al., 2004), by providing resources (e.g. light as the result of a canopy gap) and maximizing opportunities for colonization through the provision of space and nutrients (Davis et al., 2000). Too much disturbance or too little disturbance is detrimental to species richness (Connell, 1978; Gardner et al., 2008). However, despite potential application of the intermediate disturbance hypothesis to explain patterns of plant diversity observed in heterogeneous landscapes, no studies to date have investigated how patterns of forest cover change through time can influence native diversity. To address, this, we investigate the effects of the typology of multi-temporal forest cover change trajectories, obtained from a longitudinal assessment of forest cover using k-means clustering, on native plant diversity. In accordance with the intermediate disturbance hypothesis, we expect that trajectories that suggest slight to moderate disturbances to have higher diversity; and trajectories that are indicative of severe disturbances to be associated with lower diversity as compared to a trajectory of little to no change. K-means clustering is an unsupervised, data-driven method that can be used to find the natural sub-classes of a complex or lengthy data set (Kulik et al., 2011) and can also be applied to longitudinal data sets (Genolini & Falissard, 2011).

To obtain diversity data, we survey 177 random plots located within 40 forest fragments located throughout the Charlotte, NC region using a nested sampling design in which plots were nested within fragments. We obtain forest cover trajectories by assessing the amount of forest cover present at each plot for 5 time periods: 1938, 1956, 1975, 1997 and 2009, for nested spatial extents (100, 200, 400, 750 and 1000m diameter) surrounding each plot, in order to investigate how the effects of forest cover changes on native diversity varies with scale. We also derive a measure of overall forest change using just a single time interval, 1938-2009, using the same nested spatial extents. We then investigate the effects of forest cover change trajectories assessed at each plot location, on native plant diversity, while accounting for other landscape and environmental factors that are known to influence diversity patterns metrics such as soil fertility, topographical characteristics, and landscape structure, using linear mixed models. This research identifies a typology of forest change trajectories spanning five decades and the relevant geographic scale at which they influence native diversity, allowing us to understand the impacts of past spatio-temporal patterns of anthropogenic disturbance on current levels of diversity.

The fourth chapter provides an analysis of the direct and indirect effects of urbanization (using road density and housing age as metrics) on the ability of forest communities to resist invasions, while accounting for environmental factors (slope, topographic moisture, canopy closure and soil fertility), and landscape proximity, using a structural equation modeling framework. Several broad-scale studies focused on understanding the distribution of exotic and native species along rural-urban gradients have documented the increase of exotic species with increasing urbanization (e.g., Burton et al., 2005; Bartuszevige et al., 2006; Duguay et al., 2007), but few studies of invasibility have so far utilized anthropogenic factors as predictors. Thus, while the existence of links between urbanization and exotic species richness is well established, the underlying mechanisms driving these relationships have not been investigated. Anthropogenic activities associated with urbanization have been linked to eutrophication (nutrient enrichment of habitats) and propagule pressure (a composite of the number of releases of the invader into the novel environment and the number of propagules (seeds) released during each event. In this context, we expect that urbanization directly influences community resistance to invasion via the positive relationship between urbanization and propagule pressure. Human activities such as transportation and gardening introduce exotic species to novel sites both accidentally and intentionally (Lockwood, 2005) and may explain why metrics of urbanization such as housing and road density and exotic species richness are positively correlated (Burton et al., 2005; Bartuszevige et al., 2006; Duguay et al., 2007; Ranta & Vilijanen, 2011). We also expect urbanization has indirect negative effects via soil fertility on invasion resistance. The links between urbanization and the eutrophication to forest ecosystems are well documented and result from increased atmospheric deposition of nitrogen and nutrient loading resulting from anthropogenic activities such as fertilizers applied to lawns carried by stormwater runoff (Stevens et al., 2004; Sutton et al., 2011). Eutrophication has been linked both to increased habitat invasibility as well as biodiversity loss (Davis et al., 2000; Parepa et al., 2013; Burke & Grime, 1996). We integrate these hypotheses into a framework that also considers the roles of landscape habitat proximity, abiotic factors related to light and moisture availability, soil fertility, and local habitat disturbance, on invasion resistance. To test our system of integrated hypotheses, we use structural equation modeling (SEM), which is a multivariate statistical approach that unlike other multivariate methods such as ordination can be used for hypothesis testing (Grace, 2010). The advantages of SEM include the ability to distinguish between direct and indirect effects, and to simultaneously evaluate a network of hypothesized relationships among variables. This is the first study to investigate the potential direct and indirect effects of urbanization on invasion resistance.

Chapter 5 presents the conclusions of this dissertation research and potential avenues for future research. Chapters 2-4 have been prepared and organized in such a way to resemble as closely as possible a manuscript for submission to a peer-reviewed journal. As such, these manuscripts often refer to “we” instead of “I” to reflect the contributions of the co-authors of the manuscripts.

CHAPTER 2: BEYOND HABITAT SUITABILITY: THE ROLE OF INVASION PRESSURE ON THE DISTRIBUTION OF FOREST INVADERS

Introduction

Maintaining and enhancing the ecosystem services provided by urban forests are a critical element of management strategies aimed at achieving landscape sustainability. Although the ecological value of remnant urban forests is often overlooked, these forests are recognized to provide several key ecosystem services, including 1) habitat for a variety of birds and wildlife; 2) clean air via carbon sequestration and the removal of volatile organic compounds, 3) protection of drinking water by filtering runoff and reducing erosion; and 4) health and mental well being by offering a place for humans to recreate and experience nature (Nowak & Dwyer, 2007; Escobedo et al., 2011). What sets apart urban forests from other forest ecosystems is that the continued provisioning of these ecosystem services is unceasingly threatened by invasions of shade-tolerant exotic shrubs. When dense thickets of invasive cover form from uncontrolled invasion foci, native plant species are displaced (Ens & French, 2008; Merriam & Feil, 2003), which in turn results in cascading changes in ecosystem structure and function that ultimately can negatively impact ecosystem services (Haddad et al., 2009; Hanula et al., 2009; Hart & Holmes, 2013).

In order to develop effective management strategies to protect the conservation values and ecosystem services provided by urban forests, landscape-scale information on the potential distribution of invaders is needed and this can be accomplished via species distribution modeling (Peterson, 2003; Galien et al., 2010). Species distribution models (SDMs) utilize ecological niche theory to statistically link species occurrences as identified from field surveys with environmental and climatic variables theorized to influence the likelihood of invasion. This is commonly done in conjunction with a geographic information system (GIS) in order to map habitat suitability (Peterson, 2003; Guisan & Thuiller, 2005).

Although habitat suitability alone is insufficient for an invasion to occur, the focus of species distribution modeling has largely been on predicting the risk of invasion based on habitat suitability, without explicit consideration for the role of dispersal or invasion pressure in influencing invasion success (van Teeffelen & Ovaskaine, 2007). However, the organism must actually reach the site in order for an invasion to occur and the probability of arrival at any given location is influenced by invasion pressure (Tanentzap & Bazely, 2009; Shea & Chesson, 2002). Invasion pressure refers to the number of propagules arriving at a given location from invaded locations over a given duration; it has been shown to explain invasion success in experimental invasion studies. Thus, investigation in the role of invasion pressure in SDMs is warranted.

Two sources of invasion pressure can be distinguished: exogenous, where the source of the invasive propagules is external to the forest, and endogenous, where the source of invasive propagules is mainly from within the forest. The main form of exogenous invasion pressure we are concerned with is human-mediated, and arises from the intentional import and use of exotic plants for gardening and landscaping (Kowarik, 2003; Krivanek et al., 2006). For example some of the worst understory forest shrub invaders on the eastern seaboard of the United States have been widely planted as ornamentals or hedges by homeowners, suggesting an abundant source of exotic propagules that readily escape from yards and gardens into forests via dispersal by birds and deer (Stromayer et al., 1998; Bartuszevige et al., 2006; Gosper et al., 2005, Williams et al., 2008).

In contrast, endogenous invasion pressure is a measure of a (mostly) fine-scale contagious process or local spread that reflects the likelihood of invasive propagules arriving at a given location from a nearby invaded location, which presumably belongs to the same community or habitat (Vaclavik et al., 2012). High endogenous invasion pressure can result in invasion success, even when either biotic or abiotic conditions are unfavorable to the invader (Von Holle, 2005; Brown & Peet, 2003). In addition, the predictive performance of SDMs of invasive species has been shown to be substantially enhanced by the explicit inclusion of endogenous invasion pressure (Havel et al., 2002; Meentemeyer et al., 2008; Vaclavik and Meentemeyer, 2009). The effects of human-mediated invasion pressure on the distribution of exotic plants in forests and on the performance of SDMs remain to be explicitly considered.

The objective of this paper is to investigate the extent to which human-mediated invasion pressure explains invasion patterns in metropolitan forests and whether it notably improves SDM accuracy. Although the focus of the research is on the role of human-mediated invasion pressure, we also account for endogenous invasion pressure as we expect that an SDM applied in an urbanizing landscape that relies solely on ecological niche modeling is likely to have its predictive accuracy degraded by variations in both sources of invasion pressure, if they are not included. We measure human-mediated invasion pressure as the residential force of invasion (rFOI). Based on the premise that single family residences serve as external sources of invasive propagules, the potential rFOI at a given location, *i* is measured by the cumulative sum of the inverse weighted distances from *i*, to every single family house, weighted by age, within the study extent. We expect that explicitly weighting single family residences by distance and age should result in a more realistic and more effective proxy of human-mediated invasion pressure in urban landscapes than a simple measure of housing density. We measure endogenous pressure using the "force of invasion" method described by Meentemeyer et al. (2008). Systematic model errors are anticipated when either source of invasion pressure is omitted. Two situations are possible in this respect: 1) an omission error (incorrectly predicted absences) occurs when a high invasion pressure results in more actual invaders in areas of very low suitability; and 2) commission errors (incorrectly predicted presences) are generated when low invasion pressure prevails in areas of high suitability (Table 1). Presences are correctly predicted when areas of high suitability are accompanied by high invasion pressure and absences are correctly predicted in areas of low suitability overlap with low invasion pressure (Table 1).

Table 1: Expected prediction outcome by habitat suitability and unaccounted invasion pressure

Predicted habitat suitability	**degree of unaccounted invasion pressure**	**expected prediction outcome**
high suitability	high invasion pressure	presence
high suitability	low invasion pressure	commission
low suitability	high invasion pressure	omission
low suitability	low invasion pressure	absence

We expect that accounting for both sources of invasion pressure would enhance model performance beyond what can be achieved with accounting for a single source only. To test this hypothesis, we surveyed forests located along an urban to rural gradient in Charlotte, NC for the presence/absence for two of the region's worst forest invaders that are also prevalent in residential yards, namely privet and autumn olive. We then evaluate the performance of four models for each species, which in turn include only environmental predictors (niche model), niche + human-mediated invasion pressure, niche + endogenous invasion pressure, and finally niche + human-mediated invasion pressure and endogenous invasion pressure. We also assess the level of spatial dependence present in the residuals of each model as a means to detect model misspecification as far as invasion pressure is concerned. Indeed, the influence of both exogenous and endogenous invasion pressure on the space-time process of diffusion of an invasive species and on its resulting spatial distribution is theoretically justified. Accordingly, since invasion pressure is by construct a spatial dependence driver, we expect to see a reduction in spatial dependence with the addition of human-mediated invasion pressure and endogenous invasion pressure as predictors, as compared to the niche model.

The rest of this chapter is organized as follows. First, we describe the study area, provide background on the case study species autumn olive and privet; then we describe how data were collected in the field for autumn olive and privet, we describe how the environmental predictors, residential force of invasion, and forest force of invasion are derived, and how the models to compare the effects of residential force of invasion and forest force of invasion are developed, assessed, and validated. The next section reports on the results. Finally, we discuss our findings and the implications for researchers and land managers, and present our conclusions.

Methods

1.1. Study Area

Charlotte, North Carolina, is one of the ten fastest growing cities in the United States, with a 2010 population of over 730, 000 (2010 U.S. Population Census). It is located in the Piedmont physiographic province which is characterized by gently rolling terrain, erosion prone soils, and forests dominated by mixed hardwood and pine. The city boundaries encompass much of Mecklenburg County. Rapid population growth and an expanding human footprint that is characteristic of sprawl have consumed much of the forests and agricultural land in the area (Meentemeyer et al., 2013). The forests are largely mixed deciduous, dominated by oaks and hickories, and are highly fragmented. The majority of the larger forest tracts remaining in the area are owned by Mecklenburg County.

1.2. Study Species

Ligustrum sinense (Chinese privet) was first imported for ornamental use in 1852 (Dirr, 1990). It was reported as being naturalized in forests throughout the North Carolina Piedmont as early as the 1930s (Radford et al., 1965). This semi-evergreen to evergreen shrub is widely used as hedge today as it tolerates shade, heat, drought and the clay soils that are characteristic of the area. The shrub produces small bluish-black drupes in the late Fall that are consumed by birds (Wilcox & Beck, 2007) and deer (Stromayer et al., 1998). Invasion by Chinese privet is a threat to biodiversity because it is capable of forming dense thickets which crowd out native vegetation and prevents forest regeneration (Merriam & Feil, 2002). This species has been reported as invasive throughout the eastern United States.

Elaeagnus umbellata (Autumn olive) is native to Korea, China and Japan (Latham, 1963). It was first imported from Asia in 1830 and was widely available in commercial nurseries by 1917 (Latham, 1963). It was extensively planted by the United States Soil Conservation Service for erosion control. Like Chinese privet, autumn olive has a wide distribution, ranging from Maine to Georgia, and has been reported to be an aggressive invader in the southeastern forests of the piedmont and mountains of the states of Virginia, North Carolina, South Carolina and Georgia.

1.3. Field Data Collection

To examine the effects of residential development and habitat factors on the probability of either *Ligustrum* or *Elaeangus* presence, we sampled 345 $100m^2$ field plots in patches of primarily deciduous forests stratified across three classes of building density: urban, suburban and rural in Mecklenburg County during 2009-2012 (Figure 1). To estimate building density, we extracted the centroids from 2011 countywide parcel data containing buildings (Mecklenburg County Geospatial Information Services) and used a moving window to map the density for every grid cell based on the number of buildings contained within a 1-km neighborhood of the grid cell. We assigned building density into 3 classes of development intensity: urban (<0.1 ha per building), suburban (> 0.1 ha and < 0.68 per building) and rural (> 0.68 ha per building) as described by Theobald (2005).

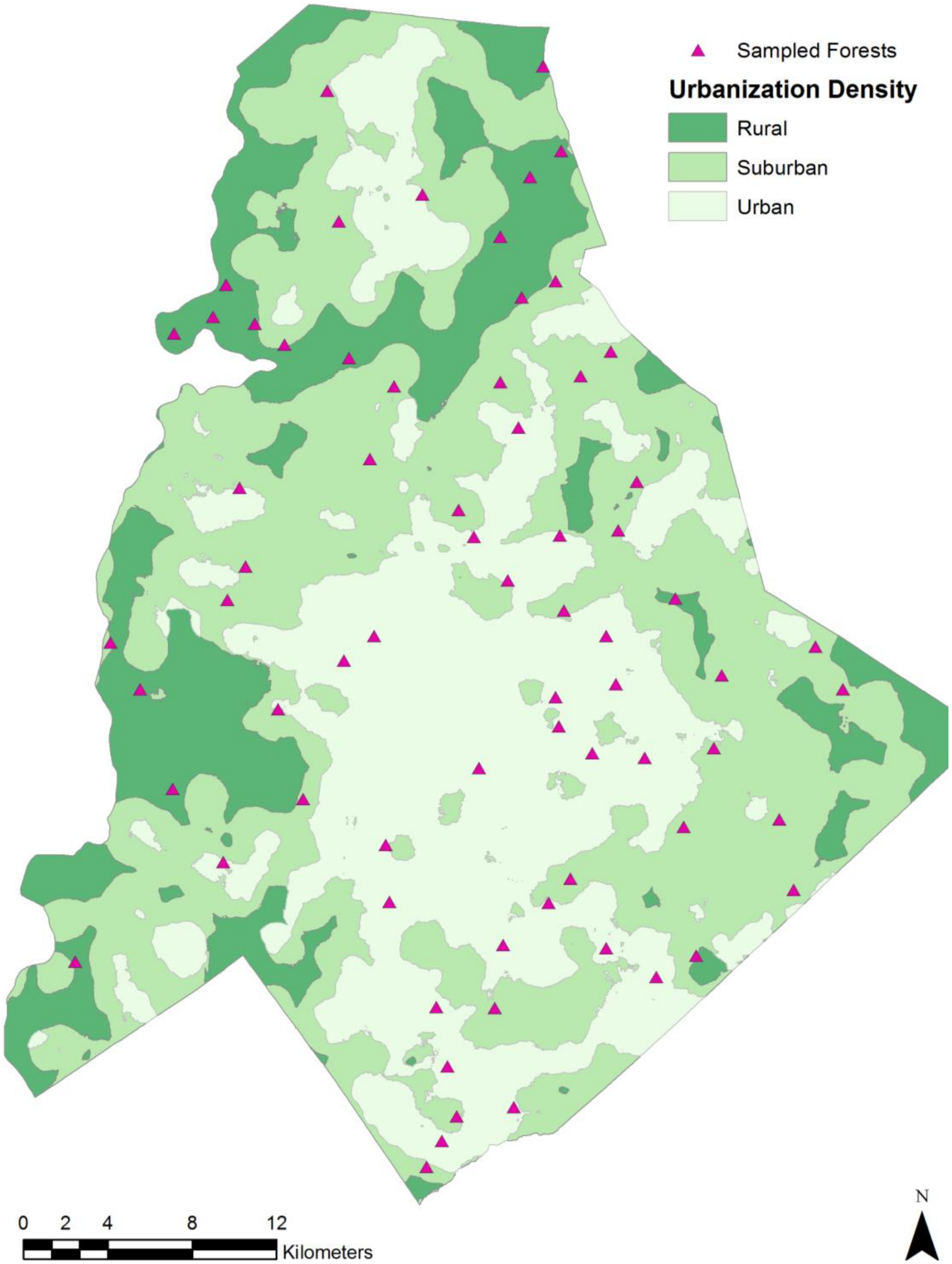


Figure 1: Study extent showing the distribution of forest patches sampled for Chinese privet and autumn olive.

We used a stratified random sampling method to identify forest patches in each class to ensure that the number of forest patches sampled reflects their spatial distribution by class within the study extent. Forest patches selected were a minimum of 2 hectares, and were delineated by roads, cleared utility right-of-ways, streams and any other non-forest cover type that separated forests by a minimum of 10 meters. Each forest sampled contained a minimum of 3 plots and a maximum of 10 plots, as determined by forest patch size to ensure that the spatial heterogeneity within each patch was well represented. Most forests contained 5-6 plots. The plots were randomly located in a geospatial database prior to conducting field work. Plots were located in situ using a handheld global positioning system (GPS). We recorded the presence of each of the two study species as "1" and their absence as "0", and if present, assigned scores of one to five to each plot according to the percentage of coverage by each of the two species: 1-20, 21-40, 41-60, 61-80, or 81-100%.

Thirty patches embedded in urban areas were sampled (136 plots), 31 suburban forests (139) and 11 rural forests (50). To augment our study, an additional 84 plots from 14 rural forests and 4 suburban forests were obtained as a random subset from the Mecklenburg County Department of Natural Resources invasive survey conducted from 2003 to 2009. A random subset was obtained due to the systematic sampling method employed by this survey. As a result, the complete dataset used in this study encompasses a total of 429 observations.

1.4. Independent Variables

1.4.1. Environmental Predictors

Eight environmental variables are selected as potential predictors of autumn olive or privet presence as indicated by theory, previous research, or based on the authors' field observations. These environmental predictors are quantified to reflect three broad classes of resource availability (light, moisture, and edaphic factors) that influence habitat suitability and thus can govern the distribution of invaders. Weather variables were not collected as both study species have a geographical range that at minimum spans the eastern U.S., and thus it can be assumed that weather would have little effect on more localized distributions.

Light availability is estimated using two different metrics of solar insolation: the annual mean potential solar radiation (SI) which measures the average solar insolation received at a given location for a whole year based on monthly intervals; and the mean potential index solar radiation (ISI), which is an estimate of the solar insolation received at a given location, based only on seasonal intervals, namely the winter and summer solstices and spring and fall equinoxes. Both are derived from a 10m digital elevation model (DEM) using the area solar radiation tool in the ArcGIS 10 software application. The percent tree canopy coverage is created by the Multi-Resolution Land Characteristics (MRLC) Consortium and is based on the 2001 National Land Cover Database (NLCD).

Moisture availability is estimated using topographic moisture index, topographic curvature, relative slope position and aspect. All of these metrics are derived from a 10m resolution DEM. The topographic moisture index is calculated as the natural log of the ratio of the upslope contributing area to the slope (Moore et al., 1991). Curvature is derived using the spatial analyst toolbox in ArcGIS 10 and is a measure of the "slope of the slope" and delineates convex, concave and flat areas (Moore et al., 1991). Relative slope position provides a measure of the relative position of each pixel/geographic location as compared to its neighbors using a moving window. A moving window of 100 m is used in this process. Aspect is transformed using the method described in Beers (1966).

Soil wetness capacity and rock type are used to measure variability in edaphic factors. Soil wetness capacity is derived from soil data obtained from the Natural Resources Conservation Service Soil Survey Geographic Database (SSURGO). To capture the potential effects of varying geology, the bedrock classification of the lithology geospatial dataset created by the North Carolina Geological Survey is used.

Distance to forest edge is mapped using a supervised classification of forest cover of 2009 satellite imagery of the study extent. This variable captures both disturbance and the possibility of higher resource variability and is frequently a factor in invasive distributions.

The dataset includes 429 observations for privet, and due to 3 missing data points, the dataset is reduced to 426 records for autumn olive. We recorded 153 presences of privet and 132 presences of autumn olive. Descriptive statistics for each species and the predictors described above can be found in Table 2.

Pearson's correlations among the set of response and predictor variables are examined to check for multicollinearity. As expected, single-family housing density and residential force of invasion are strongly correlated ($r > 0.8$); thus these variables are not

Table 2: Descriptive statistics of predictors used to model the distribution of autumn olive and Chinese privet.

Variable	Description	Min	Max	Mean	Std. dev
Autumn olive					
AO (0/1)	Presence/absence of autum olive	0	1	0.33	0.47
SI	Annual Mean Solar Radiation (WH/m^2)	989577	1679760	1489587	131635.8
CC	Percent Canopy Closure	0	97	80.94	21.99
SFHD 1000	Single family housing density within 1000 m	0.00	427.81	169.35	131.04
rFOI 1.5	Residential force of invasion with α= 1.5	0.0134	0.0783	0.042	0.016
aoFOI 1.0	Autumn olive force of invasion with α= 1.0	0.0017	0.0609	0.0065	0.0065
SWC	Soil wetness capacity (low vs high)	na	na	na	na
Chinese privet					
Priv(0/1)	Presence/absence of privet	0	1	0.36	0.48
TMI	Topographic moisture index	7.39	23.73	10.73	2.56
RSP	Relative slope position	0.05	100.05	43.33	24.11
ISI	Index Mean Solar Radiation (WH/m^2)	3156.71	3922.57	3654.38	109.93
rFOI1.0	Residential force of invasion with α= 1.0	3.06	8.50	5.84	1.40
SFHD 1500 m	Single family housing density within 1500 m	0.00	416.28	174.62	119.36
pFOI	Privet force of invasion with α= 1.5	0.00	0.00106	0.00005	0.00012
pFOI*RSP	Privet FOI interaction with Rel. slope position	0.00	0.11155	0.00688	0.01501

included together in the same model. A table listing the Pearson's correlation among the final set of environmental factors for each species can be found in Appendix 1.

1.4.2. Invasion Pressure

To examine the effects of human mediated invasion pressure on species dispersion, a new variable capturing the residential force of invasion is constructed. No complete and exhaustive dataset on landscaping use of autumn olive and Chinese privet on residential grounds is available. Short of having access to such a fine-grained dataset, every single-family residence *k* is regarded as a potential source of autumn olive and Chinese privet seeds. We exclude other building types from consideration, such as commercial buildings and multi-family dwellings, as these structures are often not accompanied by landscaped yards, whereas single-family residences almost always have these. Thus, residential force of invasion is estimated based on the accumulated presence of single-family residences (k) weighted by their age (w_k) given in the georeferenced property records (Mecklenburg County Geospatial Information Services), using an inverse distance dispersal kernel ($\frac{1}{d_{ik}^{\alpha}}$) with the following form:

$$\mathrm{rFOI}_i = \sum_{k=1}^{N} \frac{1}{d_{ik}^{\alpha}} \mathrm{w}_k \quad . \tag{1}$$

The euclidean distance between each single family residence k and a forest cell i is d_{ik}. The α parameter modifies the degree to which the likelihood of arrival from residential sites to forest sites decreases as the distance between them increases. The amount of invasive propagules received by each forest cell is the sum of the output of invasive propagules estimated for each residence. Given that these species have been present in the area for several decades (Radford et al., 1965) and given that previous work has indicated that the number of exotic species present in urban habitats is correlated with housing age (Hope et al., 2003; Gavier-Pizarro, 2010), it is postulated that the dispersal of Chinese privet and autumn olive is also influenced by housing age. Thus, the residential force of invasion (weighted rFOI) incorporates a weight by the housing age and we compare it to residential force of invasion that does not account for age (unweighted rFOI) using the likelihood profiling technique. The unweighted rFOI is formulated as follows:

$$\mathrm{rFOI}_i = \sum_{k=1}^{N} \frac{1}{d_{ik}^{\alpha}} \quad . \tag{2}$$

In order to parameterize α, we have tested values ranging from 0.5 to 2.5 in increments of 0.5 using likelihood profiling for each of the models shown in equations 1 and 2 (Havel et al., 2002; Meentemeyer et al., 2008) on the training data (see below for a description of the training data). The α parameter resulting in the least negative model log-likelihood is selected as the optimal parameter value (Havel et al., 2002). The results of the likelihood profiling analysis show that the optimal α value for rFOI is 1.0 for privet and 1.5 for autumn olive (Table 3). Weighting rFOI by housing age is preferred as it results in the minimal log-likelihood value.

Table 3: Results from likelihood profiling analysis to calibrate the α parameter for residential weighted and unweighted force of invasion (rFOI) for Chinese privet and autumn olive.

	Chinese privet		autumn olive	
α	**weighted**	**unweighted**	**weighted**	**unweighted**
0.5	-203.49	-203.77	-179.90	-178.88
1	-200.63	-203.67	-176.89	-178.28
1.5	-202.02	-203.02	-173.93	-178.10
2	-206.13	-206.27	-176.12	-179.28
2.5	-208.30	-208.29	-184.92	-184.19

Furthermore, we have compared the residential force of invasion with a simpler metric of human invasion pressure that is computationally faster to derive, namely single family housing density. In order to determine the scale at which housing density is most relevant ecologically, we have mapped housing density using circular neighborhoods of 200, 500, 750, 1000 and 1500 meter extents. The optimal spatial extents for the effects of single family housing density on the invaders is 1500 meters for privet and 1000 meters for autumn olive, as revealed by likelihood profiling analysis (Table 4).

Table 4: Results from likelihood profiling analysis to examine how the influence of single-family housing density (sfhd) on the probability of invasion by Chinese privet or autumn olive varies with spatial extent.

spatial extent	Chinese privet	autumn olive
200	-207.31	-183.79
500	-207.66	-177.01
750	-201.54	-173.51
1000	-200.12	-172.34
1500	-198.94	-173.08

In order to examine the influence of fine-scale dispersal primarily arising from nearby invaded sites within forests, on the likelihood of invader presence, we have derived the forest force of invasion (Havel et al., 2002; Meentemeyer et al., 2008) for both target species. Similarly to the derivation of the rFOI, the forest force of invasion (fFOI) is estimated as the distance (d_{ik}) from every known invaded cell present in the training data, k, to every other cell (i) excluding itself in the study extent using an inverse distance dispersal kernel modified by the β parameter, weighted by the invasive cover (W_{ick}) with the following form:

$$fFOI_i = \sum_{k-1}^{n} \frac{1}{d_{ik}^{\beta}} W_{ick}. \qquad (3)$$

The β parameter modifies the degree to which the likelihood of arrival from invaded sites to uninvaded sites decreases as the distance between the sites increases. The best-fit β value is obtained again using likelihood profiling on the training data as described above. The inclusion of the forest force of invasion parameter should account for the spatial dependence that would be expected to occur as a result of dispersal of the invader. The invasive cover is transformed from percentage categories to an ordinal variable, with values ranging from 1-5, with 1 representing the lowest cover category of

1-20% and 5 representing the highest cover of 81-100%. The model for the non-weighted fFOI is:

$$fFOI_i = \sum_{k-1}^{n} \frac{1}{d_{ik}^{\beta}} . \tag{4}$$

We use a neighborhood of 1500 meters to account for the scope of spatial dispersion around each site, as 99% of fleshy fruited seeds released from shrubs have been shown to be dispersed within this radius (Vittoz & Engler, 2007). Forest force of invasion weighted by cover is superior to its unweighted counterpart for both species (Table 5). The optimal β value for the force of invasion is 1.5 for privet and 1.0 for autumn olive.

Table 5: Results from likelihood profiling analysis to calibrate the β parameter for forest force of invasion (fFOI) for Chinese privet and autumn olive.

	Chinese privet		autumn olive	
β	**weighted**	**unweighted**	**weighted**	**unweighted**
0.5	-184.51	-199.71	-154.32	-147.22
1	-178.39	-195.65	-131.95	-136.47
1.5	-176.43	-194.53	-136.91	-146.35
2	-178.52	-194.5	-147.81	-155.31
2.5	-180.39	-194.8	-155.91	-160.29

2.5 Model Development

The training models are constructed using a random subset of 70% of the complete dataset. The remaining 30% of the data, the “testing set”, is used to validate the models developed in the training phase. Generalized linear modeling using the probit-link function is used to develop our models. Parametric estimation is completed using the “stats” package of R, version 2.15.3 (R Core Team, 2012). The probit specification is widely used in spatial econometrics and other applied fields in the estimation of spatial models with binary dependent variables. In probit models, the probability that a location will be invaded is calculated as:

$$\text{Prob}\ (Y = 1|X) = \Phi\ X\beta + \varepsilon \tag{5}$$

The link function is indicated by Φ, the cumulative standard normal distribution, X represents a vector of predictors and β are the parameters to be estimated. Probit models differ from logit in that the true value of the dependent variable is a latent variable with a Gaussian distribution between 0 and 1. The probit specification is theoretically more appropriate for models predicting the distribution of invasive species since the risk of invasion is not directly observed, and is instead a latent variable with continuous values between 0 and 1 (LeSage & Pace, 2009).

The observed values are the binary outcomes of the latent variable, in which the outcome is 1 when the value meets or exceeds a given threshold and is 0 when it is less than the threshold:

$$y_i = \begin{cases} 1 \text{ if } z_i \geq t, \\ 0 \text{ if } z_i < t \end{cases} \tag{6}$$

where y_i represents a binary outcome, in our case, presence or absence of the invader, z_i is the hidden Gaussian variable, and t is the threshold at which if z_i meets or exceeds, then $y_i = 1$. Traditionally, values above 0.5 are scored as “1” or present, and values below 0.5 are scored as “0” or absent. Using a threshold of 0.5 has been shown to be insufficient for predicting the distribution of species whose prevalence is substantially lower than 0.5 as is privet and autumn olive. The optimal threshold to score plots as being invaded (1) or not invaded (0) is derived for each model from receiver operating characteristic (ROC) plots using the ROCR package in R (Sing et al., 2005). Probit and logit specifications usually generate nearly identical results in terms of AIC and log-likelihood in the absence of predictors with extreme-values (Hahn & Soyer, 2005). In order, to check this assumption, we compare the predictive accuracy of probit models to their logit model counterparts and find that the results are nearly identical (not shown).

The base model is the niche model, which consists only of environmental predictors as is the standard practice with ecological niche modeling (Jimenez-Valverde et al., 2011). A parsimonious model is constructed using the Akaike information criterion (AIC) to select the best set of environmental predictors, with significance at $p \leq 0.05$. To avoid overfitting, we use likelihood ratio tests, so as to compare the fit of the full model (k, number of parameters) with simpler nested models (with fewer than k parameters).

To isolate the relative contribution of human-mediated invasion pressure (rFOI or single family housing density) and forest force of invasion to improving model performance, two additional models that expand on the niche model are constructed: 1) niche + residential force of invasion or single family housing density; and 2) niche + forest force of invasion. The final model adds forest force of invasion and either residential force of invasion or single family housing density to the niche model to account for both fine and broad scale sources of propagule pressure (Tables 6a and 6b).

Table 6a: Names and specifications of models developed to isolate the relative contributions of environmental factors, residential force of invasion and forest force of invasion for autumn olive distribution modeling.

Model Name	**Model Specification**
autumn olive niche	SI * CC + SWC
niche + residential FOI	SI * CC + SWC + rFOI 1.5+ rFOI 1.5^2
niche + aoFOI	SI * CC + SWC + aoFOI 1.0
niche + aoFOI + residential FOI	SI * CC + SWC + aoFOI 1.0 + rFOI 1.5 + rFOI 1.5^2
niche + SFHD 1000	SI * CC + SWC + SFHD 1000 + SFHD 1000^2
niche + SFHD 1000 + aoFOI	SI * CC + SWC + SFHD 1000 + SFHD 1000^2 + aoFOI

Table 6b: Names and specifications of models developed to isolate the relative contributions of environmental factors, residential force of invasion, and forest force of invasion for Chinese privet distribution modeling.

Model Name	**Model Specification**
privet niche	RSP + TMI + ISI + ISI^2
niche + residential FOI	RSP + TMI + ISI + ISI^2+ rFOI 1.0+ rFOI 1.0^2
niche + privetFOI	RSP + TMI + ISI + ISI^2 + pFOI 1.5 + (RSP * pFOI 1.5)
niche + resIdentialFOI + privetFOI	RSP + TMI + ISI + ISI^2+ rFOI 1.0+ rFOI 1.0^2 + pFOI 1.5 + (RSP * pFOI 1.5)
niche + SFHD 1500	RSP + TMI + ISI + ISI^2 + SFHD 1500 + SFHD 1500^2
niche + SFHD 1500 + privetFOI	RSP + TMI + ISI + ISI^2 + SFHD 1500 + SFHD 1500^2 + pFOI 1.5 + (RSP * pFOI 1.5)

2.6. Model Validation

Models are evaluated on the basis of omission error and commission error, overall predictive accuracy, and area under the curve (AUC) (Manel et al., 2001). In order to verify that the models are working as hypothesized, following Sargent (2013), we examine the predictive outcomes of niche, niche + residential force of invasion and niche + forest force of invasion. If the model performs according to the hypothesis, niche model omission errors that convert to presence after accounting for either residential or forest

FOI should have higher force of invasion than the ones that remain as omission errors. Niche model commission errors that convert to absence after accounting for either residential or privet force of invasion should have lower force of invasion than the ones that remain as commission errors. To test whether the difference in means is statistically significant, we use Welch's t-test, which is a robust test allowing for unequal variances (Ruxton, 2006).

2.7. Assessment of Spatial Dependence

It is common for the residuals of SDM models to exhibit spatial autocorrelation (Dormann et al., 2007; Vaclavik & Meentemeyer, 2012). Spatial lag models regard spatial autocorrelation as the result of a meaningful process that results in some spatial clustering of the data values. The spatial dependence present in model residuals is estimated and included in the model as additional predictor which operates according to proximity like a spillover process, the spatial lag effect, in order to obtain well-behaved parameter estimates (Anselin, 2002). The forest FOI is similar to a spatial lag effect in that we expect our data to exhibit spatial autocorrelation as the result of unaccounted endogenous invasion pressure in which the presence of privet or autumn olive influences the probability of presence in neighboring locations. The difference is that we are using ecological information a priori to guide the specification of the forest FOI and testing whether the forest FOI as a measure of invasion pressure can explain spatial dependence rather than deriving it from the data a posteriori as is the case with the spatial lag parameter. If our hypothesis is correct, we expect SDMs developed for autumn olive and privet that do not include measures of endogenous invasion pressure to have significant spatial autocorrelation, as a result of unaccounted invasion pressure. We also expect that accounting for human-mediated invasion pressure should result in a decrease of spatial dependence since these variables are hypothesized to also influence dispersal patterns. In order to assess the degree of spatial autocorrelation (SAC) that may be present, we calculate the Moran’s I statistic for the generalized residuals of each of the SDM models (Gourieroux, 1987):

$$\frac{\varphi(x_i,\theta)}{\Phi(x_i,\theta)[1-\Phi(x_i,\theta)]}[yi-\Phi(x_i,\theta)] \tag{7}$$

where φ is the probability distribution function and Φ is the cumulative distribution function for the standard normal distribution. The distribution under the null hypothesis of the Moran’s I test statistic based on the generalized residuals of the probit model is asymptotically standard normal and can be used for hypothesis testing (Amaral et al., 2013; Bille, 2013; Kelejian & Prucha, 2001; Gourieroux, 1987). We use a spatial weight matrix based on inverse distance weighted neighbors that are no further than 5,000 meters apart, which is the minimum distance required to ensure all observations have at least one neighbor.

Modeling Results

3.1. Niche Models

The results of the statistical analysis with the niche models (Tables 7a and 7b) indicate that the best set of environmental predictors for privet presence are relative slope position (RSP), topographic moisture index (TMI) and potential index mean solar radiation (ISI). Privet prefers lower topographic positions that receive less solar radiation (Table 7b). Although preliminary tests show that the topographic moisture index is not significant, it is kept in the model because privet is known to prefer moist locations

(Radford, 1965; Merriam, 2002) and this is supported by bivariate analysis with privet and TMI. The niche model for privet performs poorly overall, exhibiting the highest omission rates and lowest commission error rates for both training and testing sets as compared to the other models (Table 7b).

Table 7a: Probit regression coefficients for all the models estimated for autumn olive.

	Model					
Predictors	**Niche**	**rFOI**	**aoFOI**	**sfhd**	**rFOI + aoFOI**	**sfhd +aoFOI**
(Intercept)	-0.669***	-0.826***	-0.533*	-0.936***	-0.459	-0.536*
SI	-0.342***	-0.242**	-0.250*	-0.231*	-0.227*	-0.217*
CC	0.450***	0.385**	0.306*	0.360**	0.289*	0.273*
SI * CC	0.218*	0.214*	0.185	0.213*	0.186	0.19
SWC(low)	-0.184	-0.415	-0.39	-0.460	-0.416	-0.483
SWC(med)	0.345	0.239	0.084	0.256	0.001	0.005
rFOI 1.5	na	-0.404***	na	na	-0.060	na
rFOI 1.5^2	na	0.254***	na	na	0.075	na
SFHD 1000	na	na	na	-0.495***	na	-0.117
SFHD 1000^2	na	na	na	0.361**	na	0.174
log aoFOI 1.0	na	na	1.343***	na	1.537***	1.487***
rFOI 1.5 * aoFOI 1.0	na	na	na	na	0.428*	na
SFHD 1000 * aoFOI 1.0	na	na	na	na	na	0.462*

Significance levels: ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05

Table 7b: Probit regression coefficients for all the models estimated for Chinese privet.

	Model					
Predictors	**Niche**	**rFOI**	**pFOI**	**sfhd**	**rFOI + pFOI**	**sfhd + pFOI**
(Intercept)	-0.210*	-0.385**	-0.138	-0.418**	-0.332*	-0.323*
RSP	-0.420***	-0.385***	-0.282**	-0.406***	-0.253*	-0.276**
TMI	-0.051	-0.014	-0.011	-0.054	0.029	-0.011
ISI	0.075	0.045	0.043	0.061	0.008	0.043
ISI^2	-0.251**	-0.226**	-0.254**	-0.233**	-0.234**	-0.244**
rFOI 1.0	na	0.155*	na	na	0.104	na
rFOI 1.0^2	na	0.148	na	na	0.171	na
pFOI 1.5	na	na	0.868***	na	0.861***	0.819***
pFOI 1.5 * RSP	na	na	0.418**	na	0.426**	0.429**
sfhd 1500	na	na	na	0.244**	na	0.128
sfhd 1500^2	na	na	na	0.178	na	0.167

Significance levels: ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05

The best set of environmental predictors for autumn olive presence are canopy closure (CC), annual potential mean solar radiation (SI), and soil wetness capacity (SWC). Autumn olive prefers cool, shady sites that have less solar radiation and higher canopy closure and is discouraged from invading areas with low soil wetness capacity. Like for privet, the niche model for autumn olive has the lowest accuracy and the highest omission and commission error rates as compared to the models that include either/or residential FOI or forest FOI.

3.2. Models with Human-Mediated Invasion Pressure

An examination of the individual contribution of residential force of invasion to model performance reveals a substantial reduction in omission error (Tables 8a and 8b) for both privet (testing omission error from 0.44 to 0.26) and autumn olive (testing omission error from 0.50 to 0.34), suggesting that residential force of invasion influences the distribution of these invaders significantly. An even greater reduction in omission error for privet invasion is achieved with single family housing density as a predictor instead of residential FOI (omission = 0.20).

The addition of either residential FOI or single family housing density generates an increase in commission errors for privet, which is often the trade-off for lower omission rates (Pearce & Ferrier, 2000). The inclusion of residential FOI to autumn olive niche models results in a decrease of omission and increase of commission errors whereas single family housing density results in no change of omission and decrease of commission errors (Table 8a). This suggests that residential FOI and single family housing density are not analogous. These two variables measure distinct drivers of dispersion, at least as far as autumn olive is concerned.

3.3. Models with Forest FOI

Accounting for forest force of invasion more strongly reduces omission error than with either residential FOI or single family housing density for both autumn olive and privet (Tables 8a, 8b). As with residential FOI, the addition of privet FOI to the niche model for privet presence also results in higher commission error than in the niche model alone (Table 8b). Conversely, the addition of autumn olive FOI greatly reduces commission error as compared to the niche model for autumn olive presence (Table 8a).

Table 8a: Evaluation of probit regression models predicting the distribution of autumn olive.

model	AIC	threshold	omission	commission	accuracy	AUC	Morans I
niche (SI + SWC * CC)	354.82	0.38	0.33 (0.5)	0.34 (0.41)	0.67 (0.57)	0.71 (0.6)	0.155***
niche + residentialFOI	327.55	0.29	0.22 (0.34)	0.35 (0.44)	0.69 (0.58)	0.78 (0.67)	0.112***
niche + aoFOI	254.29	0.29	0.19 (0.22)	0.19 (0.29)	0.81 (0.72)	0.88 (0.8)	-0.015
niche + aoFOI + residentialFOI	255.00	0.34	0.2 (0.22)	0.17 (0.26)	0.82 (0.75)	0.89 (0.83)	-0.021
niche + SFHD 1000	321.70	0.37	0.38 (0.5)	0.2 (0.2)	0.74 (0.72)	0.78 (0.66)	0.091***
niche + SFHD 1000 + aoFOI	253.73	0.34	0.21 (0.22)	0.17 (0.24)	0.82 (0.76)	0.89 (0.82)	-0.020

* numbers in parentheses refer to testing data withheld from the original training set.
Significance levels: '***' 0.001 '**' 0.01 '*' 0.05

Table 8b: Evaluation of probit regression models predicting the distribution of Chinese privet.

model	AIC	threshold	omission	commission	accuracy	AUC	Morans I
niche (RSP + TMI + SI +SI^2)	378.47	0.39	0.35 (0.44)	0.25 (0.19)	0.72 (0.72)	0.73 (0.77)	0.158***
niche + residentialFOI	374.34	0.34	0.24 (0.26)	0.30 (0.27)	0.73 (0.73)	0.75 (0.84)	0.146***
niche + privetFOI	329.34	0.31	0.17 (0.13)	0.28 (0.24)	0.76 (0.80)	0.83 (0.89)	-0.028
niche + residentialFOI + privetFOI	327.18	0.32	0.17 (0.10)	0.26 (0.26)	0.77 (0.80)	0.84 (0.91)	-0.036
niche + SFHD 1500	369.05	0.35	0.25 (0.20)	0.29 (0.26)	0.72 (0.76)	0.76 (0.86)	0.115***
niche + SFHD 1500 + privetFOI	328.73	0.3	0.16 (0.13)	0.30 (0.29)	0.70 (0.77)	0.83 (0.88)	-0.034

* numbers in parentheses refer to testing data withheld from the original training set.
Significance levels: '***' 0.001 '**' 0.01 '*' 0.05

3.4. Models with Human-Mediated Invasion Pressure and Forest FOI

As evidenced in Table 8b, the best predictive model for privet presence accounts for both residential and forest force of invasion, possessing the highest AUC (training = 0.84, testing = 0.91), accuracy (training = 0.77, testing = 0.80), and the lowest omission rate (training = 0.17, testing = 0.10). Including both residential and forest force of invasion also results in the best predictive model for autumn olive, with the highest AUC (training = 0.89, testing = 0.83). This model includes an interaction term between forest force of invasion and residential force of invasion which has a positive effect on the presence of autumn olive (Table 8a).

3.5. Spatial Dependence

Significant spatial autocorrelation is observed in the niche model residuals for autumn olive (Table 8a) and for privet (Table 8b). The addition of privet FOI substantially reduces spatial autocorrelation present in model residuals to a level where it is not statistically significant. Autumn olive FOI is also associated with a substantial reduction in spatial dependence, but a small (0.09) value of statistically significant Moran's I remains, indicating a possible misspecification of the model. This is alleviated by log-transforming autumn olive FOI, which improves model fit, and like privet FOI, generates residuals free of statistically significant spatial autocorrelation. This suggests that unaccounted endogenous invasion pressure is the source of the spatial dependence in the niche model residuals for both species.

The addition of residential FOI also serves to reduce the spatial autocorrelation present in privet and autumn olive niche model residuals, with a 27.7% reduction for autumn olive and an 8.0% reduction for privet, although spatial autocorrelation remains statistically significant at the 0.001 level for both models. The addition of residential FOI to models that include forest FOI does not further reduce residual spatial dependence.

Model Validation

3.6. Residential Force of Invasion

We investigate how the omission and commission generated by the niche models for privet and autumn olive are resolved after accounting for residential FOI. In line with our expectations, the niche omission errors that converted to presences after accounting for residential FOI have significantly higher levels of residential FOI as compared to omission errors that did not convert in the privet rFOI models (Welch's t-test for training (32.91) = 8.01, $p \leq 0.001$, testing(10.04) = 3.27, $p = 0.0084$). Relatively few privet niche commission errors are explained by residential FOI (4 out of 20 (8%), in training; 1 out of 13 (7.6% in testing). This is not surprising as the mean residential force of invasion is not appreciably different between absences (mean = 0.038, sd = 0.013) and commission (0.049, sd = 0.018) in the test set (Table 9a), suggesting that unaccounted residential FOI is not the cause of commission errors in this case. Unaccounted residential FOI also explains some of the omission errors generated by the autumn olive niche model. However, unlike our results for privet, the autumn olive niche model omission errors that were converted to presence after accounting for residential FOI, experience significantly lower residential FOI as compared to omission errors that did not convert (Welch's t-test for training (18.45) = 8.46, $p \leq 0.001$, testing(5.24) = 3.94, p = 0.01; Table 9a). This phenomenon is observed with commission errors, as those that are converted to absences after accounting for residential FOI unexpectedly have higher mean residential FOI as compared to those remaining as commission errors (Welch's t-test for training (61.98) = 2.84, $p \leq 0.006$, testing (24.04) = 3.29, p = 0.003. This suggests that contrary to our expectations, autumn olive invasion is more likely in areas with low residential FOI.

Table 9a: Results of Welch's t-tests comparing the mean residential FOI of omission/commission errors from the niche models that are converted to presences/absences after including rFOI in the models, with those that remain as omission or commission for both training and testing data sets.

		N	Mean rFOI	SD	DF	Dfdens	T	p
Privet training	omission	27	5.61736	1.28137	1	32.90887	8.0147	<.0001
	presence	13	7.73993	0.34818				
	absence	4	4.56982	0.29568	1	19.89923	5.9182	<.0001
	commission	46	6.0484	1.36598				
Privet testing	omission	10	6.17871	1.50976	1	10.04093	3.2686	0.0084
	presence	7	7.7846	0.30682				
	absence	1	NA	NA		NA	NA	NA
	commission	12	5.5726025	1.098562		NA	NA	NA
AO training	omission	17	0.049495	0.013477	1	18.45378	8.4573	<.0001
	presence	11	0.020774	0.003054				
	absence	16	0.049552	0.00691	1	61.98099	2.8416	0.0061
	commission	55	0.041327	0.017223				
AO testing	omission	6	0.047971	0.017851	1	5.237131	3.935	0.01
	presence	6	0.018956	0.00275				
	absence	7	0.0557	0.00571	1	24.03786	3.2907	0.0031
	commission	36	0.045033	0.014512				

3.7. Forest Force of Invasion

Similarly, we investigate how the omission and commission generated by the niche models for each of the two invasive species are resolved after accounting for privet FOI and autumn olive FOI, respectively. If the pFOI and aoFOI models perform as intended, than the privet/autumn olive omission errors associated with high levels of their respective forest force of invasion should be identified as presence in the pFOI and aoFOI models. Similar to what we have seen with unaccounted residential FOI, several of the locations predicted by the privet niche model as omission errors are associated with high unaccounted privet FOI. These locations become correctly identified as presences in the pFOI model. On the contrary, those with low privet FOI remain counted as omission errors (Table 9b). The difference in means between the newly identified presences and the remaining omission errors is statistically significant for both training and testing sets (Welch's t-test for training (21.58) = 6.46, $p \leq 0.001$, testing (12.0) = 3.46, $p = 0.004$) (Table 9b). Accounting for privet FOI also results in a reduction of niche commission errors. As expected, the errors that are converted to absences have a much lower privet FOI as compared to those that remain commission cases (Welch's t-test (43.72) = 4.09, $p \leq 0.001$) (Table 9b). No commission errors are converted to absence in the test set, however the mean privet FOI is not significantly different between absences and commission ($p = 0.23$) (Table 9b). This suggests that the commission errors in the testing set cannot be explained by privet FOI.

As with privet, the autumn olive niche omission errors that converted to presences after accounting for autumn olive FOI had higher mean autumn olive FOI than those errors that did not convert, although this difference was only significant for the training set (Welch's t-test for training (19.27) = 5.7, $p \leq 0.001$, testing (7.86) = 1.85, $p = 0.1$) (Table 9b). In contrast to privet, the inclusion of autumn olive FOI results in a substantial reduction of commission error. The commission errors that are converted to absences are linked to significantly lower autumn olive FOI than those that did not convert (Welch's t-test for training (34.05) = 6.18, $p \leq 0.001$, testing (21.54) = 3.93, $p \leq 0.001$) (Table 9b). This indicates that some niche commission errors (falsely predicted absences) can be explained by very low forest FOI, as we hypothesized in Table 1.

Table 9b: Results comparing the mean forest FOI for Chinese privet and autumn olive using the same testing approach as described for Table 9a.

		N	Mean fFOI	SD	DF	Dfdens	T	p
Privet training	omission	18	0.000055	0.000044	1	21.59	6.46	<.001
	presence	22	0.000627	0.000412				
	absence	7	0.000012	0.000014	1	43.72	4.09	<.001
	commission	43	0.000161	0.000238				
Privet testing	omission	5	0.000163	0.000253	1	12.85	3.46	0.004
	presence	12	0.001501	0.00128				
	absence	0	NA	NA		NA	NA	NA
	commission	13	0.000097	0.000116		NA	NA	NA
AO training	omission	10	0.003643	0.001299	1	19.27	5.70	<.001
	presence	18	0.012829	0.006606				
	absence	39	0.002904	0.000646	1	34.05	6.18	<.001
	commission	32	0.005856	0.002639				
AO testing	omission	5	0.005507	0.002403	1	7.86	1.85	0.1
	presence	7	0.010729	0.006926				
	absence	24	0.002894	0.000782	1	21.54	3.93	<.001
	commission	19	0.004992	0.002222				

Discussion and Conclusions

We investigated whether the addition of human-mediated invasion pressure and endogenous invasion pressure to SDMs developed for autumn olive and Chinese privet better explains invasion patterns in metropolitan forests and improves model accuracy. Our results show that 1) human-mediated invasion pressure does influence the likelihood of invasion by either autumn olive or Chinese privet and its inclusion generated a significant decrease in omission errors as compared to niche only models; and 2) the highest predictive accuracy and lowest omission error rates are obtained for both species when both human-mediated and endogenous sources of invasion pressure are included in the models. We discuss each of these results in more detail below.

A comparison of models with and without residential FOI revealed that the addition of residential FOI substantially boosted the performance of niche models for Chinese privet (Table 8b) and to a lesser degree, autumn olive (Table 8a) by lowering omission errors. However, the role of residential FOI on the distribution of these invaders differed by species. Our results show that although residential FOI increases the risk of invasion by Chinese privet, it decreases the risk of invasion by autumn olive (Table 9a). This trend is likely explained by the predominance of autumn olive in rural (N = 211) as opposed to urban forests (N = 28). The niche models developed in this study significantly underpredict the distribution of the invaders, and residential FOI corrects for this at the expense of commission error. Including residential FOI in SDMs applied to metropolitan landscapes is useful for increasing the detection of an invader when niche models are performing poorly, as they likely will with species with wide environmental tolerances (Evangelista et al., 2008). To the best of our knowledge, this is the first study to explicitly investigate the role of anthropogenic variables on the performance of spatially explicit predictive models such as SDMs.

The explanatory power of residential FOI was surpassed by estimates of endogenous invasion pressure for both species. Our results link high and low estimated endogenous invasion pressure as measured by forest FOI to omission and commission errors, respectively, for autumn olive and privet. Thus, not surprisingly, the addition of forest FOI to models that also accounted for residential FOI resulted in the best predictive models overall, thus supporting the hypothesis that accounting for both sources of invasion pressure will provide the best predictive accuracy. This was significantly better than previous results reported for autumn olive (AUC = 0.75, omission = 0.20) and

Chinese privet (AUC = 0.72, omission = 0.38) (Lemke et al., 2012). Accounting for forest FOI has previously been shown to improve the performance of SDMs of invasions by *Daphnia lumholtzi* (Havel et al., 2002) and *Phytopthora ramorum* (Meentemeyer et al., 2008). Although forest FOI is a stronger predictor of invasion success than residential FOI, a practical advantage to residential FOI is that all the values of residential FOI can readily be determined for the entire study extent, i.e. all of the cells belonging to residences are known, whereas the forest FOI has some uncertainty as our knowledge of invader presence is limited to the locations that were sampled. Non-sampled locations may harbor presences that were not included in the forest FOI. The predictive accuracy of non-sampled locations that are far from known presences (i.e. greater than 1500m) is less certain. The 1500 m distance reflects the maximum distance that 99% of fleshy fruited seeds like those of privet of autumn olive will disperse (Vittoz & Engler, 2007).

The niche models developed in this study significantly underpredict the distribution of the invaders and the residuals of these models exhibit significant spatial autocorrelation. Spatial autocorrelation can be the result of an omitted abiotic or biotic variable, or poor model specification (Austin, 2007). Although dispersal and other biotic processes have been acknowledged as a potential source of autocorrelation (Legendre, 1993; Bahn et al., 2008), the residual spatial autocorrelation is often handled with an autocovariate or filtered out, without explicit inquiry/investigation as to the origins of the spatial dependence (Miller and Franklin, 2010; De Knegt et al., 2010; van Teeffelen & Ovaskaine, 2007). The use of autocovariates has garnered recent controversy, as although they may improve prediction (Augustin et al., 1996; Betts et al., 2009) their impacts on inference of the other parameters in the models are unclear (Dormann et al., 2007).

However, ignoring spatial dependence can also lead to false conclusions (Legendre, 1993). Although our approach for accounting for endogenous invasion pressure via the forest FOI is essentially an autologistic method (Augustin et al., 1996), it differs from the traditional use of autocovariates in the sense that we have incorporated invasive cover as an additional attribute of the response variable, as we expect that it positively influences invasive pressure. Spatial autocorrelation is often attributed to the tendency of the species to cluster in similar environmental conditions, or in other words because the environmental conditions are clustered, so are the distribution patterns of the organism. If this was true in our case, we would expect to see high commission errors in our niche models rather than the high rate of omission errors observed. The addition of forest FOI to niche models not only resulted in a vast improvement of predictive accuracy, but greatly reduced the level of spatial autocorrelation present in the residuals, suggesting that the failure to account for endogenous invasion pressure is the cause of both the high omission error and spatial dependence present in the niche models. Our finding that endogenous invasion pressure is a major predictor of invasion success is in line with previous findings and theory (Eschtruth & Battles, 2011; Simberloff, 2009) and suggests that the theoretically motivated use of autologistic methods is appropriate for predictive models. Our results indicate that forest FOI explains invader presence in sites identified by the simpler niche model as unsuitable. This implies that high numbers of immigrants from invaded habitats are arriving and successfully establishing in marginal habitats. This has been observed with plant invasions in grasslands (Thomson, 2007), and forests (Brown & Peet, 2003; Von Holle & Simberloff, 2005). Marginal habitats that receive high amounts of seed rain from known invaded locations or from residences likely to harbor the invader, are more susceptible to invasions than what otherwise would be predicted by models that do not estimate propagule pressure. However, forest FOI does not explain commission errors as well suggesting that despite the elimination of spatial dependence via the inclusion of forest FOI, the possibility of an omitted variable that make these locations more resistant to invasions exists, or that these locations simply haven't been successfully invaded yet (Lobo et al., 2010).

This is the first study that explicitly investigates the potential links between invasion pressure and the performance of SDMs in urban landscapes. We have shown that high invasion pressure arising from either endogenous or anthropogenic sources increases the risk of invasion in habitats that would otherwise be identified as unsuitable for invasion. We have also demonstrated that omission and commission errors of the niche model are associated with unaccounted human-mediated invasion pressure and endogenous invasion pressure. Thus, it is imperative that predictive models, especially when applied to metropolitan landscapes, incorporate a spatially explicit understanding of the contribution of both sources of invasion pressure to the risk of invasion to increase the chances that these habitats at risk of invasion are not left unidentified. This conclusion is strengthened by the fact that these results were achieved with two different invasive shrub species across a heterogeneous metropolitan area. This approach could be used to assist land managers in prioritizing forest remnants for detection and control of invaders in order to protect biodiversity and has the potential to improve the accuracy of SDMs developed for other widespread sub-canopy generalist invaders of metropolitan forests (e.g. *Lonicera maacki, Rhamnus cathartica*). However, our results point to the daunting task that land managers may face in the sense that even if they successfully eradicated invasives from their properties, forests in urbanizing landscapes are continuously subjected to external invasion pressure from residences, which is outside their realm of control. This suggests that forests that are experiencing high human-mediated invasion pressure will require frequent monitoring in order to keep invasions in check.

CHAPTER 3: INFLUENCE OF MULTI-TEMPORAL FOREST COVER CHANGE TRAJECTORIES ON FOREST PLANT DIVERSITY

Introduction

Human activities such as the construction of roads, the conversion of forest to intense agricultural uses or impervious surface, are major drivers of forest loss, which have negative consequences on biodiversity worldwide (Foley et al., 2005). These human induced changes result in decreased habitat area, the creation of new edges that encourage an influx of weedy and invasive species, increased dispersal distances and isolated populations (Fahrig, 2003; Gibson et al., 2011; Didham et al., 2007; With, 2002). Understanding and quantifying the effects of land cover change on native biodiversity is one of the main goals of landscape ecology (Turner, 2005).

Small forest remnants embedded in urban landscapes, however altered they may be, offer refuge to various bird and wildlife species and still have an important role in the provisioning of ecosystem services, including wildlife habitat, flood mitigation, air pollution removal and recreation (Nowak & Dwyer, 2007; Escobedo et al., 2011; Pataki et al., 2011). Since biodiversity facilitates ecosystem resilience and the provisioning of ecosystem services, the maintenance of biodiversity is of paramount importance within the context of landscape sustainability (Folke et al., 2004; Hooper et al., 2005). As a first step towards developing forest management strategies to enhance and protect diversity, it is vital to understand how forest biodiversity has responded to forest cover change dynamics that have occurred over time. Although previous studies have investigated the effects of historical patterns of landscape forest structure on plant biodiversity (e.g., Bellemare, 2002; Von Holle & Motzkin, 2007; D'Orangeville et al., 2008), and studies have linked cross-sectional measures of forest loss and forest fragmentation to declines in biodiversity (Arroyo-Rodriguez et al., 2013; Aguiliar-Santelises & del Castillo, 2013; Lung et al., 2012), no studies to date have investigated how patterns of forest cover change through time can influence native diversity. In this context, the goal of this paper is to analyze the relationship between temporal trajectories of forest cover change and patterns of forest plant diversity. The Charlotte, North Carolina, region provides the geographic setting for this research.

Over time, human-environmental interactions have deeply impacted patterns of forest cover and change in the southeastern region of the United States. This pattern can be summarized in three phases: 1) widespread deforestation occurred during the 18th and 19th centuries to accommodate agriculture, 2) subsequent agricultural abandonment beginning in the late 19th and early 20th centuries as a result of poor agricultural practices in which forest regenerated on fallow fields, and 3) deforestation as a result of expanding urbanization after World War II (Peet & Christensen, 1984; Wear & Greis, 2002). Disturbance is generated when an event (such as forest clearing, windthrow, etc.) results in the eradication of biomass and the clearing of habitat space, thereby permitting new residents to utilize resources (Roxburgh et al., 2004). Intermediate fluxes in forest cover are a type of disturbance, generated as result of human-environment interactions. According to the intermediate disturbance hypothesis, the timing and spatial extent of disturbances can influence levels of biodiversity (Mallik et al., 2013). Intermediate levels of disturbance can promote the coexistence of species in a community, resulting in higher species richness (Connell, 1978; Roxburg et al., 2004), by providing resources (e.g. light as the result of a canopy gap) and maximizing opportunities for colonization through the provision of space and nutrients (Davis et al., 2000). Too much disturbance or too little disturbance is detrimental to species richness (Connell, 1978; Gardner & Engelhardt, 2008). These intermediate fluctuations in forest cover are unlikely to be apprehended when land cover change is measured using a single time interval as is commonly done (e.g. de Albuquerque & Rueda, 2010), especially when this interval is large. Thus, the main drawback to assessing the impacts of land cover change on biodiversity using a single interval spanning a long time period, is the inability to evaluate the impacts of intermediate anthropogenic disturbances that characterize human-dominated landscapes on biodiversity.

Several statistical methods have been used successfully to predict and explain land cover change, including linear regression (e.g., Meyfroidt et al., 2010; de Espindola et al., 2012), logistic regression (e.g. Verburg et al., 2002; Huang et al., 2009; Schneider & Pontius, 2001), spatial Markov chain models (e.g. Lambin, 1997; Guan et al., 2011) and others. The latter two methods use transition matrices and can handle sequences of land use/land cover data at multiple discrete time steps to predict the probability of land use change (Muller & Middleton, 1994). The availability of remotely sensed data has enabled the investigation of spatial phenomena via the acquisition of a time series of aerial and/or satellite imagery covering the same geographic extent at fine temporal and spatial resolutions. The causal agents of land use/land cover change have been studied extensively (e.g. Agrawal et al., 2002; Lambin et al., 2001), but the study of impacts of land cover change trajectories on ecological processes or natural phenomena is just beginning to gain traction.

For the purpose of this research, we define a trajectory as a sequence of changes, or lack thereof, assessed for more than one time interval. Progress in this area of research may have been hindered by methodological challenges to the representation of spatially explicit change over time. Peuquet (2005) pointed out that the only data model available within a geographic information system (GIS) for spatiotemporal data is a "temporal series of spatially registered snapshots", in which spatial phenomena are assessed over a series of discrete time periods, thus corresponding to a longitudinal survey or repeated measures approach. These time series are processed using an overlay analysis to obtain a single layer that reflects the accumulation of the sequences of change that have taken place for a given location. This method utilizes an object-oriented approach where the spatial object is defined by the change or accumulation of changes that have occurred, and has been referred to as space-time composites (Langran & Chrisman, 1988). Each spatial unit whether represented using grid cells or a polygon, represents a succession of the changes that occurred. For example if land use change is assessed for 4 time periods, and the possible land use states for each time period are "forest", "agriculture" or "urban", then a potential change accumulation object is "agriculture-forest-forest-urban". In this way, the trajectory defines the object.

Although predictive models that employ Markov chains and transition matrices can easily handle a time series of land cover data, it is much more difficult to distill time series of land cover data into meaningful typologies or categories of spatio-temporal change that can then be utilized in other modeling approaches. For example, Guns &

Vanacker (2013) summarized land cover changes examined across 4 time intervals on the basis of a conversion from a natural cover type to a degraded or non-natural cover type. Although they have recorded the interval at which the change occurred, the trajectories are grouped according to the type of change that occurred, consequently ignoring the time it occurred. This in effect, obscures the temporal component of this dataset. The advantage of this approach is that they had distilled their spatially explicit time series dataset into a relative small number of possible trajectories (5), which could then be used as predictors in logistic regression models to understand the role of land use change trajectories on landslide risk. On the other hand, the work of Feng et al. (2014) demonstrates the potential enormity of the number of possible trajectories that can be obtained from a time series of land cover assessments. They tracked the land use types present in satellite imagery using 5 time periods, and assigned each pixel a numeric value representing the trajectory it experienced, following the change accumulation approach described above. They derived a total of 2,218 trajectories and were only interested in those trajectories in which urbanization occurred (only 530 relevant trajectories met this condition). Only 21 of these trajectories occupied an area greater than 0.5% of the study extent. Instead of incorporating these trajectories in a single model, they calculated the individual contribution of each one of these trajectories to the urban heat island effect in the Xiamen City, China metropolitan area. This approach demonstrates that the number of unique trajectories describing the succession of land use types or states of the system that occurred can be quite numerous if the trajectories are not collapsed into broader groups. However, the ad-hoc grouping of trajectory data into categories, a process that is frequently based on the nature of the change, regardless of the magnitude or timing of the change, may obscure the very reason a longitudinal land cover assessment was conducted in the first place. It may also result in the data being organized in ways that do not adequately represent meaningful similarity within groups. In addition, the classification of such large numbers of trajectories may prove cognitively challenging.

As an alternative to ad-hoc or supervised classification approaches, k-means clustering is an unsupervised, data-driven method that can be used to find the natural sub-classes of a complex or lengthy data set (Kulik et al., 2011) and can also be applied to longitudinal data sets (Genolini & Falissard, 2011), although this has not yet been not applied to the clustering of longitudinal land cover data. The advantages of using an unsupervised method such as k-means clustering include the ability to group data not just based on the type and rate of the change, but also according to the timing and sequencing.

In this research, we investigate the effects of a typology of forest cover change trajectories on native plant diversity. A typology is obtained by k-means clustering of a longitudinal assessment of forest cover. To obtain data on native plant diversity, we surveyed 177 random plots located within 40 forest fragments located throughout the Charlotte, NC region using a nested sampling design in which plots were nested within fragments. We obtained forest cover trajectories by assessing the amount of forest cover present at each plot for each of the following 5 time periods: 1938, 1956, 1975, 1997 and 2009. This is also done for nested spatial extents (100, 200, 400, 750 and 1000m diameter) surrounding each plot in order to investigate how the effects of forest cover changes on native diversity varies with scale. In addition, we derive a more commonly used metric of forest cover change: forest cover change measured using 2 time steps (1938-2009 Forest Change), with the same nested spatial extents. We then investigate the effects of forest cover change trajectories or 1938-2009 Forest Change, assessed at each plot location, on native plant diversity, while accounting for other landscape and environmental factors that are known to influence diversity pattern metrics such as soil fertility, topographical characteristics, and landscape structure, using linear mixed models. In accordance with the intermediate disturbance hypothesis, we expect trajectories that suggest slight to moderate disturbances to have higher diversity; and trajectories that are indicative of severe disturbances to be associated with lower diversity as compared to a trajectory marked by little to no change. This research has the potential to identify forest change trajectories and the relevant geographical scale that are associated with the lowest or highest diversity. Its significance stems from the ability to understand the impacts of past spatio-temporal patterns of anthropogenic disturbance on current levels of diversity.

The remainder of this paper is organized as follows. First, we present an overview of the modeling framework to obtain a meaningful typology of forest cover change trajectories. We then provide background information on the kml algorithm, and describe the study system, the sampling design used to collect data on native diversity and soil fertility, and the derivation of soil moisture and solar irradiation in a GIS. We also describe our land cover mapping procedure and the derivation of metrics of landscape structure and forest change trajectories using these data. We conclude our methods section with a description of how linear mixed models were implemented, along with likelihood ratio tests in order to assess the contribution of forest cover change trajectories to explaining native plant diversity. Finally, we provide our results and a discussion of their implications for land cover change research.

Methods

2.1 Modeling Framework

In order to understand how the diversity of native forest plants responds to distinct forest cover change trajectories, we calculate from field data our response variable, the Shannon Index of native diversity (StdH). This index takes into account both the number of unique species present and their abundance (Shannon, 1948).

$$H' = -1 * \sum_{i=1}^{S} (p_i \ln p_i) \qquad (1)$$

Where H' is the diversity index, S is species richness (the total number of unique species present) and p_i is the proportion of the total sample belonging to the *i*th species. The minimum H' that can be obtained is 0, in which case there is a single species in the whole sample. When the sample is dominated by the abundance of one species, in the presence of additional species, the index approaches 0. The maximum H' value that can be obtained is bounded by the species richness of the sample and is obtained when the species present have identical abundance.

In this study, we use a nested sampling design (described below) to collect data on native plant diversity, in which several plots were sampled within forest patches. Since our sample plots are clustered within forest patches, the data are not truly independent, as sample plots within patches are likely to be correlated due to being exposed to a similar environmental and landscape context.

Grouped data are common in experimental or observational ecological studies and result from collecting multiple data samples within sites, experimental blocks or individuals. These types of data are referred to as “multilevel”, as the lower level observations or measurements are nested within higher level groups. Data collected from blocks or groups violate assumptions of independence and can lead to Type I errors, when the grouping structure is ignored. Linear mixed models (also known as multilevel models or hierarchical linear models) provide a statistically robust way to analyze multilevel/grouped data (Pinheiro & Bates, 2000; Zuur et al., 2009).

The general model structure for a two-level model, where Y_{ij} is the response variable vector of responses for the ith level-1 unit nested within the jth level-2 unit is (Pinheiro & Bates, 2000):

$$\text{Level-1: } Y_{ij} = \beta_{0j} + B_{1j}X_{1ij} + \varepsilon_{ij;}$$

$$\text{Level-2: } \beta_{0j} = \beta_0 + u_{ij} \tag{2}$$

Where β_{0j} is the mean for the jth group, B is a vector of linear coefficients corresponding to a vector of predictors X and ε_{ij} represents the within-group errors which are assumed to be normally distributed with a mean of 0 and constant variance; β_0 is the intercept for the jth group and u_{ij} represents the between-group errors, also assumed normally distributed with a mean of 0 and constant variance. Fixed effects refer to the experimental treatment or explanatory variables of interest (Zuur et al., 2009). Random effects refer to the variation among groups. Here, we use linear mixed models to quantify the effects of distinct forest cover change trajectories on the Shannon index of native plant diversity, while also accounting for environmental and landscape covariates. Using this approach, the independent variable (typologies of forest cover change trajectories) and environmental and landscape covariates are fixed effects, and the grouping factor, which is the forest patch in which plots were samples is a random effect. In this way the correlation of native diversity measured within forest patches is accounted for when assessing the effects of explanatory variables and mitigates the risk of Type I errors expected from using a statistical procedure that ignores the grouped/clustered structure of the data (Pinheiro & Bates, 2000). We repeat this approach using 1938-2009 Forest Change instead of the typologies of forest cover change trajectories as the independent variable, to explain native diversity. All the models are estimated using a maximum likelihood procedure. Maximum likelihood is preferred over restricted maximum likelihood when comparing fixed effects terms in mixed models (Bates, 2010).

In order to identify the spatial extent under which the trajectories or 1938-2009 Forest Change have the most influence on native diversity, we first construct the base linear mixed model containing only environmental and landscape covariates. Then we compare the base model to a model that also contains either the forest cover trajectories variable derived at one of the five nested spatial extents or 1938-2009 Forest Change derived at one of the five nested spatial extents, using likelihood ratio tests.

In addition to collecting data on native plant diversity, we also collected data on one of the environmental covariates, soil fertility in the field. The remaining environmental covariates (soil moisture and solar irradiation) as well as the landscape covariates (1938 mean landscape proximity, 1938 total forest area, 1938 number of forest patches, 2009 mean landscape proximity, 2009 total forest area, 2009 number of forest patches, and matrix change) were derived through GIS analysis using relatively straightforward methods described in more detail below. However, the derivation of the main independent variable of interest, typologies of forest cover change trajectories, involves numerous steps. These are outlined as follows 1) obtain and georectify historical aerial photos corresponding to plot locations, 2) create land cover maps by digitizing the land cover types present in each aerial photo 3) delineate nested spatial buffers in a GIS around each plot, 4) summarize the area of the smallest buffer covered by forest for each time period, and repeat for each increasing buffer size in order to examine the effects of the spatial scale at which forest cover trajectories are assessed on native diversity, 5) classify the longitudinal forest cover data into groups representing the major types of forest cover change trajectories that were present for each spatial extent using the kml algorithm. This provides the distinct typologies of forest cover change trajectories that are used as a predictor in linear mixed models.

The kml algorithm is a non-parametric k-means clustering algorithm that accommodates longitudinal data; it is available in the kml software package for the R language (Genolini & Falissard, 2011). The principle of a k-means clustering algorithm is to repeatedly cluster data until the within-cluster variance is minimized and the between-cluster variance is maximized. One disadvantage of k-means is that the number of clusters must be specified a priori. However, the kml software package offers multiple cluster validation methods that can be used to choose the optimal number of clusters for each spatial extent. These methods each utilize a criterion that assesses how well a given number of clusters has partitioned the data broadly operating on their ability to minimize within-group variation and maximize between-group variation (e.g. Calinski-Harabatz (1974), Davies-Bouldin (1979), Ray-Turi (1999)), or the lowest deviance as measured using information theoretic methods: Akaike information criterion (AIC), Akaike corrected information criterion (AICc) and the Bayesian information criterion (BIC). We describe our study system and these steps in more detail below.

2.2. Study System

The greater Charlotte, NC metropolitan area is rapidly growing, having experienced a 35% increase in population over the last decade (2010 U.S. Population Census). It is located in the piedmont physiographic province which is characterized by gently rolling terrain, erosion prone soils, and forests dominated by mixed hardwood and pine (Christensen & Peet, 1984). The region has experienced a trajectory of deforestation, reforestation and deforestation due to intense human-environmental interactions over the past three centuries. Broad scale clearing of forests for agricultural use began in the 18th century, and by the mid-19th century, nearly all of the arable land in the Piedmont was being tilled. Huge amounts of topsoil were lost due to poor agricultural practices, with estimates of erosion for cotton fields approximating 9 tons per acre per year (Trimble, 1974). Due to widespread erosion, declining fertility, and the collapse of the cotton market in the 1930s, cultivation was largely abandoned, resulting in widespread reforestation either by natural processes or the planting of loblolly pine (*Pinus taeda*) (Trimble, 1974; Hart, 1980). Virtually no old-growth forest remains, as almost the entire region has at one time been under cultivation. The second wave of deforestation is the result of urban growth, which is consuming the remaining open space in the area (Meentemeyer et al., 2013). The forests that remain in the region potentially harbor significant biological diversity and serve as refuges for bird and wildlife.

2.3. Sampling Design & Field Data Collection

In order to ensure that the full spectrum of urbanization and associated land use changes within the study extent are represented, we sampled 177 plots across 40 forest patches located across a rural-urban gradient to obtain data on native plant diversity. The rural-urban gradient spans the county seat of Charlotte, NC and the adjacent county of

Cabarrus (Figure 1). To map building density, we extracted the centroids from 2011 countywide parcel data containing commercial or residential buildings (Mecklenburg County Geospatial Information Services, Cabarrus County Land Records Information System) and used a moving window approach to map the density for every $100m^2$ grid cell based on the number of buildings contained within a 1-km neighborhood of each grid cell. We assigned building density into 3 classes of land-use intensity: urban (<0.1 ha per building), suburban (> 0.1 ha and < 0.68 per building) and rural (> 0.68 ha per building) as described by Theobald (2005). Next, 40 forest fragments were selected across the land-use strata using a stratified random sampling strategy (Figure 1). Prior to field work, 3-5 10m x 10m plot locations were randomly selected within each forest, a minimum of 100 m apart, and were located in situ using a Trimble GeoXT GPS receiver. The diameter at breast height (DBH) and species were recorded for all woody stems that were a minimum of 1cm DBH (Burton and Samuelson, 2005). We used these data to calculate the response variable, the Shannon diversity index based on native species for each plot (Shannon, 1948). Three soil cores were taken at a depth of 0-10 cm and were pooled. A total of 177 $100m^2$ plots were sampled.

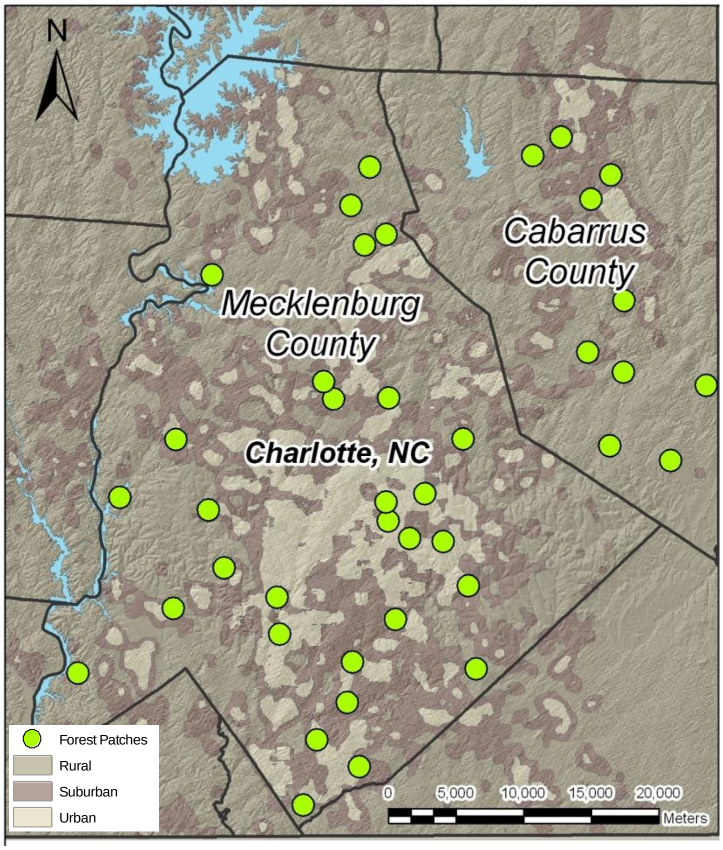


Figure 1: Study extent showing the location of forests that were sampled within the greater Charlotte metropolitan area.

2.4 Predictor Variables

2.4.1. Land cover mapping

Digital aerial photographs from 5 different time periods (1938, 1957, 1975, 1997 and 2009) corresponding to sampled plot locations were obtained from Mecklenburg and Cabarrus county GIS departments. We georectified aerial photos for the plots located in Mecklenburg County to 2009 satellite imagery. The Cabarrus county aerial photos had been georectified previously by Cabarrus County GIS personnel. We mapped the land cover types present within a 1,000m diameter buffer of each plot for each of the 5 time steps using “heads up” digitizing in a GIS. Land cover within each buffer was coded into 4 broad classes, namely agriculture, development, forest, or water, so as to keep classification errors to a minimum (Table 2). Various measures (the landscape covariates) designed to capture the change of land-cover and landscape over the study period were derived from these basic statistics; these measures are introduced in the next section. Also, in order to examine how the effects of forest cover change on native diversity varies with spatial extent, four smaller nested buffers (100, 200, 400 and 750 meters) were derived as a subset from the 1,000m buffer, for a total of 5 spatial extents.

To create a spatiotemporal dataset of forest cover for each sampled plot, the area of forest in hectares contained within each 1,000m plot buffer at each of the 5 time steps was extracted through GIS processing. This was repeated for the remaining 4 spatial extents embedded within the 1,000m buffers (100m, 200m, 400m, and 750m). The coverage for the 1957 aerial photos for Mecklenburg County was incomplete for 10 forest fragments. Rather than omit them from the dataset, the missing forest cover values were derived using a robust multiple imputation method described by Honaker et al. (2011). In order to summarize the amount of deforestation, reforestation and persistent forest for each time interval (1938-1956, 1956-1975, 1975-1997, 1997-2009, and 1938-2009) we used overlay analysis of our 1,000 m buffer land cover maps to detect changes in forest cover between 2 successive time periods. For example areas that were mapped as forest in 1938 were classified as “deforestation” if they were no longer forest in 1956.

2.4.2 Landscape Covariates

We derived three metrics summarizing historical forest landscape structure using Fragstats version 4.1 on the basis of the 1938 forest cover data obtained in the previous step: total area forested in 1938 within 1 km of each plot, number of forest patches within 1 km of each plot in 1938, and landscape proximity for 1938 (McGarigal et al., 2012). Using 2009 forest cover data, we obtained the same metrics to also summarize contemporary landscape structure. The proximity index is a dimensionless number that provides a relative measure of forest connectivity within a landscape, where lower numbers designate more isolated patches, and higher numbers designate less isolated patches.

We also measured the area of agricultural land that converted to development from 1938 to 2009 as a metric of matrix change. The matrix is defined as the largest most interconnected patch in the landscape in which smaller patches appear to be embedded (Forman, 1995). In this study, the matrix is agriculture in 1938, as that was the dominant land use type. We know that the matrix transitions to development, as developed/urban land is the dominant land use type by 2009. With this variable, we will be able to test whether conversion of agriculture to development measured at a landscape scale has a negative effect on native diversity, independently from forest loss (Table 1).

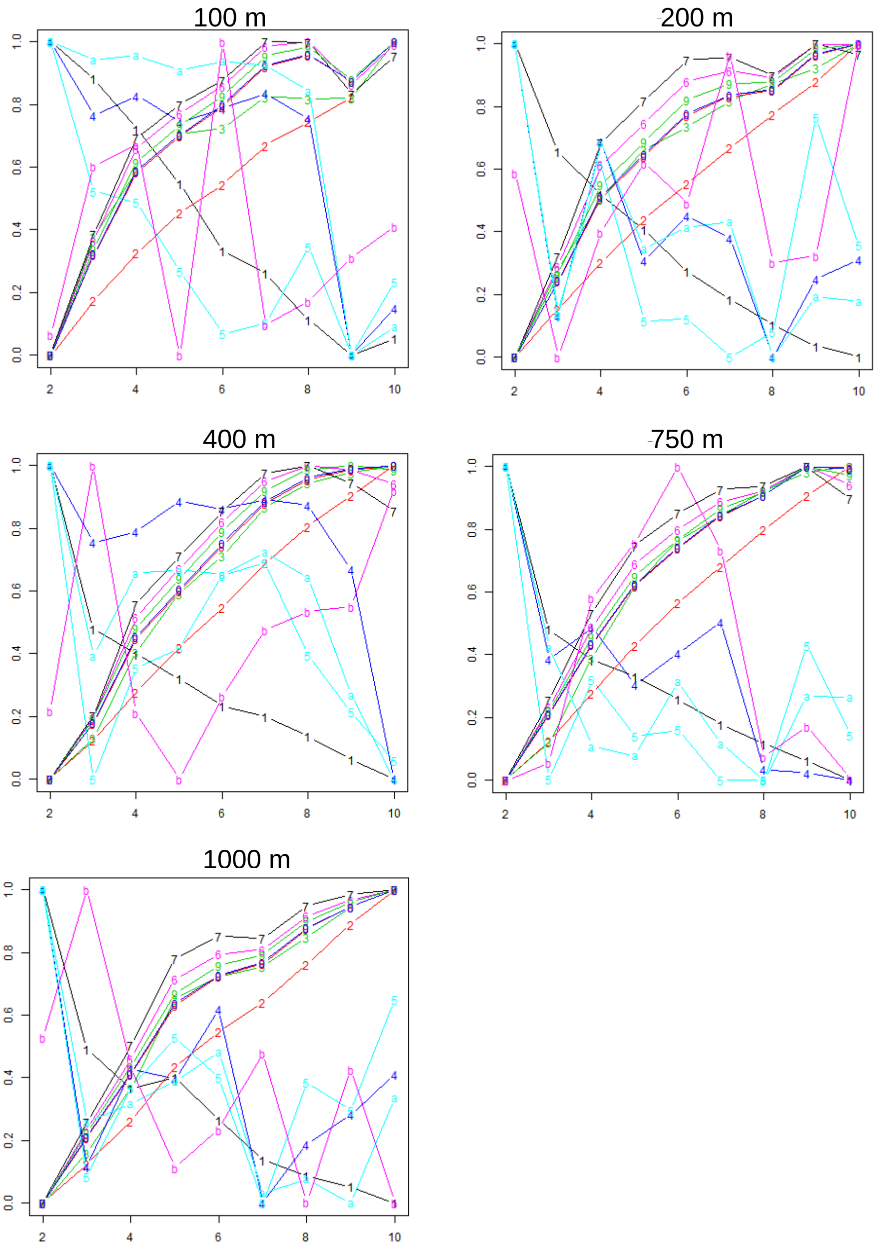
Table 1: Descriptive statistics for all the variables examined.

Variable	Description	Min	Max	Mean	Std. dev
StdH	Shannon diversity index calculated for native species at the stand level for each plot	0.010	0.067	0.038	0.009
TrH	Shannon diversity index calculated for native species at the transgressive level for each plot	0.000	1.463	0.547	0.240
pH	Soil pH	3.800	6.300	5.029	0.517
N	Soil nitrogen (ppm)	36.000	128.000	90.026	15.174
P	Soil potassium (ppm)	5.000	59.000	17.058	9.340
Ca	Soil calcium (ppm)	127.000	2913.000	767.410	544.506
K	Soil phosphorus (ppm)	18.000	337.000	77.173	41.246
TMI	Topographic moisture index	7.050	23.720	10.525	2.502
SI	Potential mean solar radiation	3252.000	3840.190	3641.918	119.269
1938ForArea	Total area (ha) forested in 1938 within 1 km of each plot	0.000	47.316	17.731	11.719
1938NP	Number of forest patches in 1938 within 1 km of each plot	0.000	10.000	3.853	1.893
1938Prox	Mean landscape proximity in 1938	0.000	448.700	28.343	51.167
2009ForArea	Total area (ha) forested in 2009 within 1 km of each plot	5.427	74.972	31.392	15.382
2009NP	Number of forest patches in 2009 within 1 km of each plot	1.000	11.000	4.404	1.914
2009Prox	Mean landscape proximity in 2009	0.000	742.972	106.388	152.017
MatChange	Area within 1 km of each plot that converted from agriculture to development from 1938 to 2009	0.000	58.020	18.701	14.401
traj100	Major forest change trajectories derived within 100 m of each plot (categorical variable)	na	na	na	na
traj200	Major forest change trajectories derived within 200 m of each plot (categorical variable)	na	na	na	na
traj400	Major forest change trajectories derived within 400 m of each plot (categorical variable)	na	na	na	na
traj750	Major forest change trajectories derived within 750 m of each plot (categorical variable)	na	na	na	na
100_FC	Overall change in forest area from 1938 to 2009 within 100 m of each plot	-0.250	0.780	0.226	0.293
200_FC	Overall change in forest area from 1938 to 2009 within 200 m of each plot	-1.080	3.140	0.785	1.048
400_FC	Overall change in forest area from 1938 to 2009 within 400 m of each plot	-5.090	9.950	2.205	3.422
750_FC	Overall change in forest area from 1938 to 2009 within 750 m of each plot	-14.170	25.660	4.910	9.293
1000_FC	Overall change in forest area from 1938 to 2009 within 1 km of each plot	-21.480	38.580	6.616	14.466

2.4.3 Forest Change Trajectories

The input of the kml clustering algorithm is a matrix of n rows, where n is the number of sample plots (177), by m columns, where m is the number of time steps (5). Each column corresponds to a vector of the area of forest cover measured at each of the 177 plots for a certain time step, and each row contains five cross-sectional measures of forest cover between 1938 and 2009. When these trajectories are processed, we repeatedly apply the kml algorithm to obtain unique partitions of the forest cover trajectories, with each partition corresponding to a specific number of clusters (i.e. from 2 to 10) derived by kml. A total of 9 different partitions are obtained for each spatial extent. We assess the quality of each partition according to 11 individual criteria or cluster validation methods, using a graphical interface provided by the kml package, for each spatial extent. We report the results of the 11 cluster quality indexes, as a single criterion may be unreliable, and agreement among multiple criteria can be helpful in choosing the optimal number of clusters (Genolini & Falissard, 2011). All the criteria are standardized on a scale of 0 and 1, so that the optimal clustering is identified by the largest criterion. The optimal number of clusters was seven for both the 100 and 200 m spatial extents, eight for 400 m, nine for 750 m, and six for 1000 m (Figure 2).

Once the optimal number of clusters for each extent is identified, each trajectory class can be described qualitatively in terms of the nature of the change (reforestation, deforestation, or no change) and the timing of the change, e.g. reforestation between 1938 and 1957 (Table 2). The trajectory class is a predictor in the multivariate model of native plant biodiversity. A total of 13 trajectory classes are derived across the 100, 200, 400, 750, and 1000 m spatial extents and are listed in Table 2. Of these, 6 trajectory classes appear in 3 or more spatial extents; these are the most common trajectory classes derived. The mean trajectory (forest cover plotted over time) representative of each of the 6 most common trajectory classes are shown in Figure 3. Four mean trajectories are identified for a single spatial extent only, with two of these associated with the 750 m extent. Trajectories that explain less than 5% of the data are assumed to be miscellaneous as they describe less than 8 plots in our dataset. Comparing the same trajectories across scales is revealing: the percentage of trajectories classified as little to no change in forestation decreases as spatial extent increases, indicating that the probability of forest change increases with scale. Deforestation is more evident at larger spatial extents given that 3 mean deforestation trajectories are only identified at the 750 m extent. The percentage of plots classified as rapid reforestation between 1957 and 1975, or rapid reforestation between 1975 and 1997 trajectories are relatively stable across the 100-400 meter spatial extents, suggesting that these trajectories are scale invariant to at least 400 meters.


100 m

200 m

400 m

750 m

1000 m

Figure 2: Graphical output from the Kml package showing the cluster quality values of 11 criteria, standardized on scale of 0 to 1 (y axis) for each number of clusters derived (x axis) for choosing the optimal number of clusters for each spatial extent. For each criterion the optimal number of clusters is the one that has the highest cluster quality value as indicated by the first maxima. We select the optimal number of clusters based on those that have the highest number of criteria indicating the highest cluster quality values. The criteria are: “1” Calinski-Harabatz 1, “2” Calinski-Harabatz 2, “3” Calinski-Harabatz 3, “4” Ray uri, “5” Davies Bouldin, “6” Bayesian Information Criterion (BIC) 1, “7” BIC 2, “8” Akaike Information Criterion (AIC) 1, “9” Corrected Akaike Information Criterion (AICc) 1, “a” AICc2, “b” Posterior probability.

Table 2: Mean forest change trajectories derived for each spatial extent and relative distribution (%) across the study area.

Trajectory Class	100	200	400	750	1000
little to no change	53.70	34.50	26.00	7.91	na
rapid reforestation between 1938 and 1975	16.40	14.70	11.90	13.60	15.30
reforestation until 1957, deforestation after	8.47	20.30	13.60	13.60	22.00
rapid reforestation between 1975 and 1997	7.91	7.91	6.21	na	na
rapid reforestation between 1957 and 1975	7.34	7.91	6.78	11.90	19.20
reforestation after 1938	4.52	7.34	6.21	19.20	22.00
rapid deforestation until 1975, reforestation after	1.69	na	na	na	na
slight deforestation between 1938 and 1975, reforestation after	na	7.34	11.90	na	na
reforestation-deforestation after19 57 (but less than above)	na	na	17.50	14.70	na
reforestation to 1957, rapid deforestation after	na	na	na	3.39	na
rapid reforestation from 1938 to 1975, rapid deforestation after	na	na	na	2.82	na
reforestation to 1975,deforestation after	na	na	na	13.00	12.4
reforestation from 1938-1975, deforestation 1975-1997, stable after	na	na	na	na	9.04

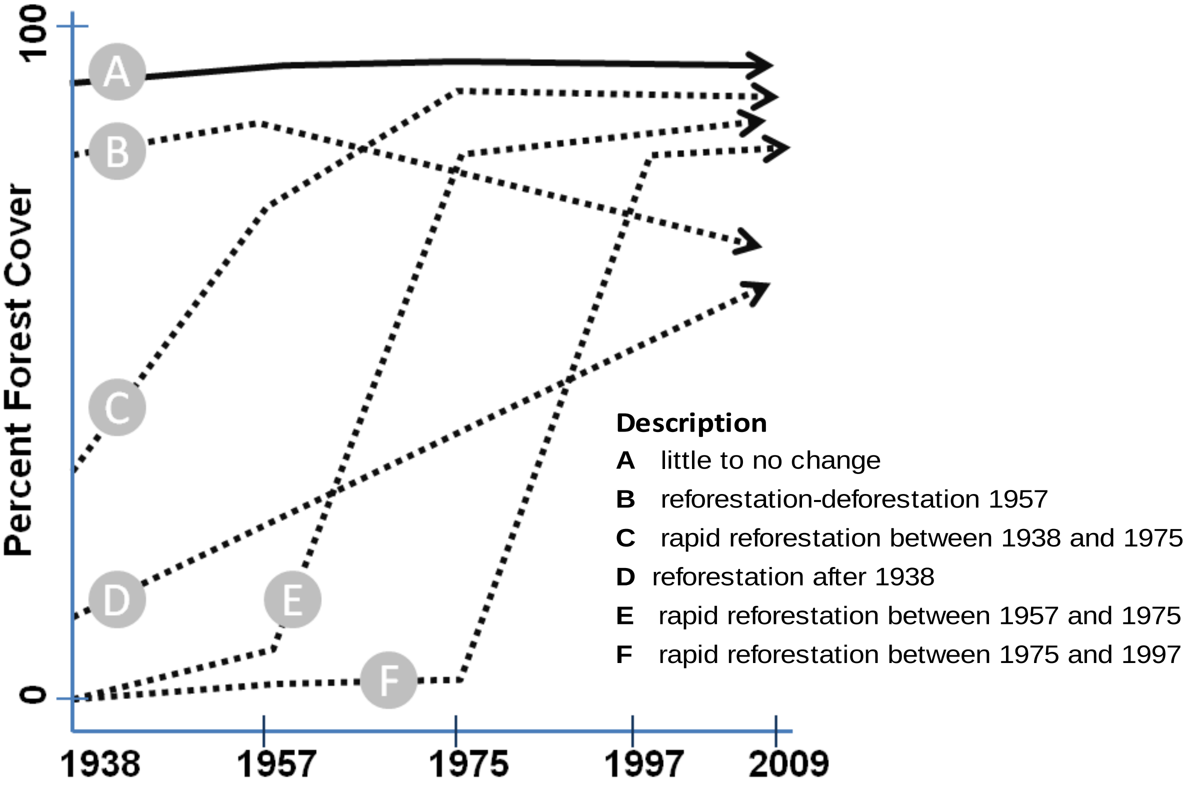


Figure 3: Schematic diagram showing how each of the six most common trajectory classes of longitudinal forest cover change derived using k-means clustering differs in the amount of forest cover measured for each time step, and the rate of change (slope) between time steps.

2.4.4. Environmental Covariates

To account for the effects of environmental heterogeneity on patterns on native diversity, soil fertility, soil moisture and solar irradiation are used as environmental covariates in our linear mixed models and their derivation is described below.

2.4.4.1. Soil Fertility

The soil samples collected in the field were sent to Brookside Laboratory for analysis to quantify the pH and the levels of available nitrogen, potassium, phosphorus, and calcium (Table 1). These five soil factors have been shown to influence species richness and vegetation composition in Piedmont forests (Peet and Christensen, 1980; Taverna et al., 2005). In order to reduce the collinearity among these variables, we reduced these 5 measurements into 2 predictors reflecting overall soil fertility using principal components analysis (Ye & Wright, 2010) in the JMP Pro 10 statistical analysis package (SAS, 2012) . The 2 retained principal components explain 42.1% and 22.7% of the data variance, respectively. Variables with significant loadings ($p < 0.05$) on the 1st principal component are pH ($R^2 = 0.41$), N ($R^2 = 0.36$), Ca ($R^2 = 0.67$), K ($R^2 = 0.62$); and on the 2nd principal component, they are pH ($R^2 = 0.25$), N ($R^2 = 0.28$), and P ($R^2 = 0.59$). These principal components are labeled SFv1 and SFv2, respectively.

2.4.4.2. Soil Moisture and Solar Irradiation

Soil moisture has also been shown to influence species richness in this region (Peet and Christensen, 1980). We derived the topographic moisture index from a 10m resolution digital elevation model (DEM) as the natural log of the ratio of the upslope contributing area and slope (Moore et al., 1991). In addition, the annual potential mean solar radiation, which can influence tree species diversity (e.g., Leathwick et al., 1998; Currie, 1991), was calculated using the method described by Dubayah (1994).

2.5. Linear Mixed Modeling

All the models of native plant diversity are estimated with the lme4 package for R (Bates, 2013). We construct a null model containing only environmental predictors (topographic moisture index, potential mean solar radiation and soil factor vectors SFv1 and SFv2) and random effects. Since our landscape structure metrics are moderately correlated (Table 3), we use likelihood ratio tests to assess the ability of each of our landscape covariates to further improve the fit of our null model for StdH. We include only significant predictors as identified from the likelihood ratio tests to construct a base model that accounts for environmental factors as well as landscape structure. We then assess the ability of either the forest cover change trajectory clusters or 1938-2009 forest change, at each spatial extent, to explain native diversity as compared to the base model, using likelihood ratio tests. We use the method described by Nakagawa & Schielzeth (2012) to estimate the R^2 of the significant models. This method reports both the marginal R^2 which is the percent variance explained by the fixed effects and the conditional R^2, which is the percent variance explained by both fixed and random effects, to determine at which spatial extent forest change trajectories explain the most variation in Shannon native diversity.

In order to identify the effect of individual trajectories on StdH, we obtain coefficient estimates and their probability values for each parameter using Markov-Chain Monte Carlo simulations (n = 100,000) using the language R package (Baayen, 2011) with the “little to no change” trajectory type as the reference level. We plot the estimated partial effects of trajectory

Table 3: Pearson's correlation of all the variables used in this study.

Predictor	StdH	SFv1	SFv2	TMI	SI	1938ForArea	1938NP	1938Prox	2009ForArea	2009NP	2009Prox	MatChange	traj100	traj200	traj400	traj750	traj1000	100_FC	200_FC	400_FC	750_FC	1000_FC
StdH	1	-0.215	-0.113	-0.052	0.006	-0.005	0.166	-0.028	0.100	0.038	0.134	-0.204	0.128	0.008	0.132	0.091	0.048	0.123	0.124	0.094	0.115	0.146
SFv1	-0.215	1	0.004	0.030	#####	-0.122	-0.073	-0.044	-0.168	0.006	-0.042	0.241	-0.077	-0.032	-0.091	-0.112	-0.002	-0.132	-0.170	-0.156	-0.208	-0.272
SFv2	-0.113	0.004	1	-0.040	0.073	-0.114	-0.073	-0.094	-0.202	0.080	0.011	0.262	0.010	0.008	-0.055	-0.142	-0.273	-0.043	-0.072	-0.058	-0.031	-0.067
TMI	-0.052	0.030	-0.040	1	0.191	-0.198	-0.139	-0.151	-0.141	-0.153	-0.140	0.135	0.128	0.078	0.177	-0.038	0.018	0.118	0.119	0.089	0.002	-0.071
SI	0.006	-0.077	0.073	0.191	1	0.065	-0.021	0.107	-0.055	-0.138	-0.147	0.117	0.159	0.109	0.116	-0.013	-0.173	0.189	0.136	0.072	0.033	0.010
1938ForArea	-0.005	-0.122	-0.114	-0.198	0.065	1	0.103	0.640	0.458	-0.087	0.251	-0.384	-0.338	-0.327	-0.353	0.254	0.381	-0.297	-0.269	-0.229	-0.128	-0.023
1938NP	0.166	-0.073	-0.073	-0.139	#####	0.103	1	0.095	-0.046	0.328	-0.069	-0.026	0.010	0.009	0.080	-0.236	-0.177	0.106	0.110	0.116	0.183	0.181
1938Prox	-0.028	-0.044	-0.094	-0.151	0.107	0.640	0.095	1	0.455	-0.181	0.094	-0.322	-0.174	-0.160	-0.238	0.198	0.178	-0.182	-0.159	-0.071	0.063	0.151
2009ForArea	0.100	-0.168	-0.202	-0.141	#####	0.458	-0.046	0.455	1	-0.309	0.646	-0.786	-0.180	-0.120	-0.200	0.213	0.400	-0.048	0.126	0.347	0.525	0.613
2009NP	0.038	0.006	0.080	-0.153	#####	-0.087	0.328	-0.181	-0.309	1	-0.109	0.025	-0.034	-0.040	0.075	-0.079	-0.031	0.005	-0.050	-0.100	-0.115	-0.110
2009Prox	0.134	-0.042	0.011	-0.140	#####	0.251	-0.069	0.094	0.646	-0.109	1	-0.498	-0.100	0.032	-0.085	0.222	0.238	-0.036	0.072	0.208	0.282	0.321
MatChange	-0.204	0.241	0.262	0.135	0.117	-0.384	-0.026	-0.322	-0.786	0.025	-0.498	1	0.247	0.223	0.146	-0.366	-0.552	0.059	-0.081	-0.231	-0.336	-0.420
traj100	0.128	-0.077	0.010	0.128	0.159	-0.338	0.010	-0.174	-0.180	-0.034	-0.100	0.247	1	0.576	0.541	-0.216	-0.376	0.642	0.570	0.443	0.298	0.216
traj200	0.008	-0.032	0.008	0.078	0.109	-0.327	0.009	-0.160	-0.120	-0.040	0.032	0.223	0.576	1	0.574	-0.234	-0.387	0.521	0.550	0.459	0.311	0.242
traj400	0.132	-0.091	-0.055	0.177	0.116	-0.353	0.080	-0.238	-0.200	0.075	-0.085	0.146	0.541	0.574	1	-0.272	-0.392	0.585	0.583	0.522	0.397	0.319
traj750	0.091	-0.112	-0.142	-0.038	#####	0.254	-0.236	0.198	0.213	-0.079	0.222	-0.366	-0.216	-0.234	-0.272	1	0.513	-0.238	-0.230	-0.251	-0.326	-0.272
traj1000	0.048	-0.002	-0.273	0.018	#####	0.381	-0.177	0.178	0.400	-0.031	0.238	-0.552	-0.376	-0.387	-0.392	0.513	1	-0.267	-0.201	-0.180	-0.212	-0.174
1938_2009_FC100	0.123	-0.132	-0.043	0.118	0.189	-0.297	0.106	-0.182	-0.048	0.005	-0.036	0.059	0.642	0.521	0.585	-0.238	-0.267	1	0.913	0.673	0.460	0.351
1938_2009_FC200	0.124	-0.170	-0.072	0.119	0.136	-0.269	0.110	-0.159	0.126	-0.050	0.072	-0.081	0.570	0.550	0.583	-0.230	-0.201	0.913	1	0.857	0.633	0.510
1938_2009_FC400	0.094	-0.156	-0.058	0.089	0.072	-0.229	0.116	-0.071	0.347	-0.100	0.208	-0.231	0.443	0.459	0.522	-0.251	-0.180	0.673	0.857	1	0.862	0.742
1938_2009_FC750	0.115	-0.208	-0.031	0.002	0.033	-0.128	0.183	0.063	0.525	-0.115	0.282	-0.336	0.298	0.311	0.397	-0.326	-0.212	0.460	0.633	0.862	1	0.948
1938_2009_FC1000	0.146	-0.272	-0.067	-0.071	0.010	-0.023	0.181	0.151	0.613	-0.110	0.321	-0.420	0.216	0.242	0.319	-0.272	-0.174	0.351	0.510	0.742	0.948	1

type on StdH with their estimated highest probability density (HPD) interval to assess the significance of trajectory type on StdH. We estimate the HPD interval of each parameter using Markov-chain Monte Carlo (MCMC) sampling of their respective posterior distribution. The HPD is formed from 95% of the parameter values with the highest posterior probability. This Bayesian approach to a confidence interval has a more intuitive interpretation than a confidence interval estimated using a classical statistical approach, in that the interval represents the boundaries of which we believe the true parameter estimate is contained with 95% confidence (Hosmer et al., 2013).

Results

3.1 Longitudinal Land Cover Changes

We summarize the area in hectares covered by agriculture, development, forest and water within a 1km buffer of each plot, for each of the 5 time steps (Table 4). For this descriptive assessment, we do not impute the missing land cover values for 1956; thus the approximately 3300 hectares corresponding to missing aerial photos are not classified. However, regardless of how these missing 3300 hectares would be allocated to specific land cover types, general land cover trends can be inferred from these data. Namely, the area devoted to agriculture steadily declines from 1938 to 2009, while during the same time developed area increases, and forest cover appears to increase from 1938 and until 1975, after which it decreases. Fluctuations in water are probably due to periodic variations in water levels and the draining/refilling of ponds.

Table 4: Summary of cumulative land cover area for a 1 km spatial extent of each plot, in hectares by year.

Landcover	1938	1956	1975	1997	2009
Agriculture	11454.26	6890.86	4287.7	2018.4	1186.63
Development	57.81	1168.79	4438.92	7893.29	9729.16
Forest	5200.41	5326.79	7945.45	6780.74	5753.07
Water	226.25	235.48	266.68	245.62	268.65
missing	na	3316.83	na	na	na

3.2 Changes in Forest Cover

We further investigate the nature of the forest cover changes observed in Table4, by examining the amount of reforestation, deforestation and persistent forest present within 1 km of each plot for each time interval (1938-1956, 1956-1975, 1975-1997, 1997-2009). The greatest amount of reforestation occurred between 1956-1975, but sharply decreases to less than half that amount during 1975 to 1997, and declines even further between 1997-2009 (Figure 4). 1975 to 1997 has the greatest amount of net deforestation observed as well as the highest amount of persistent forest. The area of persistent forest declined slightly during 1997-2009, but remains higher than the amount observed for the period covering 1938-1975. Reforestation is higher than deforestation until 1956-1975, when deforestation starts to exceed reforestation. This trend in deforestation continues to 2009. However, these trends are not evident when using the 1938-2009 time interval to assess patterns of forest change, in which reforestation is shown to exceed deforestation (Figure 4).

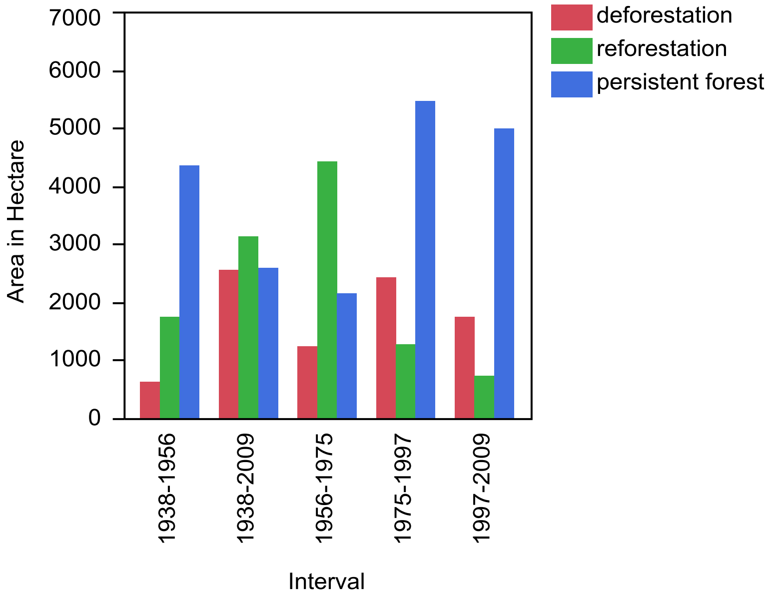


Figure 4: Total area of deforestation, reforestation and persistent forest by time interval within a 1 km spatial extent of each plot location.

3.3 Linear Mixed Modeling Analysis

Seven variables reflecting landscape structure are examined for their potential as significant predictors of StdH while accounting for environmental heterogeneity, namely 1938 mean landscape proximity, 1938 total forest area, 1938 number of forest patches, 2009 mean landscape proximity, 2009 total forest area, 2009 number of forest patches, and matrix change (Table 5). Only matrix change, 1938 number of forest patches and 2009 mean landscape proximity are identified as significant using likelihood ratio tests. However, matrix change and 2009 mean landscape proximity exhibit moderate correlation (-0.50) (Table 3) and including both parameters does not yield a better fit as compared to including one of these only, as indicated by likelihood ratio tests (data not shown). Given that matrix change is substantially more correlated with the environmental predictors in the model, as well as the forest change trajectories as compared to 2009 mean landscape proximity (Table 3), we use only 1938 number of forest patches and

2009 mean landscape proximity to represent landscape structure in our modeling work. This specification serves as base model against which other specifications are compared.

Table 5: Results of likelihood ratio tests to examine individual significance of landscape covariates as compared to the null model of only environmental predictors.

predictors	Df	AIC	BIC	log lik	Chisq	Chi Df	Pr(>Chisq)
SFv1 + SFv2 + TMI + SI + (1|site)	7	-1034.1	-1013	524.03			
SFv1 + SFv2 + TMI + SI + 1938ForAr + (1|site)	8	-1032.2	-1008	524.09	0.1126	1	0.737
SFv1 + SFv2 + TMI + SI + 1938NP + (1|site)	8	-1034.9	-1010	525.43	2.803	1	0.094
SFv1 + SFv2 + TMI + SI + 01938Prox + (1|site)	8	-1032.6	-1008	524.32	0.576	1	0.448
SFv1 + SFv2 + TMI + SI + 2009ForAr + (1|site)	8	-1033.1	-1009	524.55	1.0426	1	0.307
SFv1 + SFv2 + TMI + SI + 2009NP + (1|site)	8	-1032.1	-1008	524.04	0.0127	1	0.910
SFv1 + SFv2 + TMI + SI + 2009Prox + (1|site)	8	-1036.6	-1012	526.3	4.5303	1	0.033
SFv1 + SFv2 + TMI + SI + MatChange + (1|site)	8	-1036.5	-1012	526.26	4.4629	1	0.035

The results of the likelihood ratio tests comparing the effects of adding a categorical fixed effects term describing forest change trajectories to the base model described above (environmental predictors + landscape structure) illustrate that the addition of forest change trajectories derived at either the 100 or 200 m spatial extent results in a better predictive model of Shannon native diversity than the base model alone (Table 6). Forest change trajectories derived at the 200m spatial extent are slightly better than trajectories at the 100 m spatial extent at explaining native diversity, having the overall lowest AIC (-1042.5) and the highest marginal R^2 (0.18). The 400 m spatial extent is found to only be marginally statistically significant ($p = 0.08$). Trajectories derived at the 750 m and 1000 m spatial extents did not result in a statistically significant improvement of the base model. Forest change measured using only 2 time steps (1938-2009 Forest Change) does not significantly influence StdH at any of the spatial extents as indicated by likelihood ratio tests at $p < 0.05$, but the 100 m spatial extent is marginally significant at $p < 0.1$ ($p = 0.098$) (Table 7).

Table 6: Results of likelihood ratio tests to examine individual significance of forest cover change clusters derived at one of five nested spatial extents as compared to the base model explaining StdH.

model	# of clusters	predictors	Df	AIC	BIC	log lik	Chisq	Chi Df	Pr(>Chisq)	Marg. R2	Cond.R2
null		SFv1 + SFv2 + TMI + SI + (1|site)	7	-1034.1	-1013	524.03				0.06	0.1
base		SFv1 + SFv2 + TMI + SI + 1938 NP + 2009Prox + (1|site)	9	-1038	-1010	527.98	7.8853	2	0.019	0.1	0.13
100	7	SFv1 + SFv2 + TMI + SI + 1938 NP + 2009Prox + traj100 + (1|site)	15	-1040	-994.1	535	14.055	6	0.029	0.17	0.2
200	7	SFv1 + SFv2 + TMI + SI + 1938 NP + 2009Prox + traj200 + (1|site)	15	-1042.5	-996.6	536.25	16.542	6	0.011	0.18	0.21
400	8	SFv1 + SFv2 + TMI + SI + 1938 NP + 2009Prox + traj400 + (1|site)	16	-1036.6	-987.6	534.32	12.682	7	0.080	0.16	0.19
750	9	SFv1 + SFv2 + TMI + SI + 1938 NP + 2009Prox + traj750 + (1|site)	17	-1031	-979	532.52	9.0851	8	0.335	0.14	0.17
1000	6	SFv1 + SFv2 + TMI + SI + 1938 NP + 2009Prox + traj1000 + (1|site)	14	-1030.2	-987.3	529.1	2.2491	5	0.814	0.11	0.14

Table 7: Results of likelihood ratio tests to examine individual significance of the 1938-2009 Forest Change variable derived at one of five nested spatial extents as compared to the base model explaining StdH.

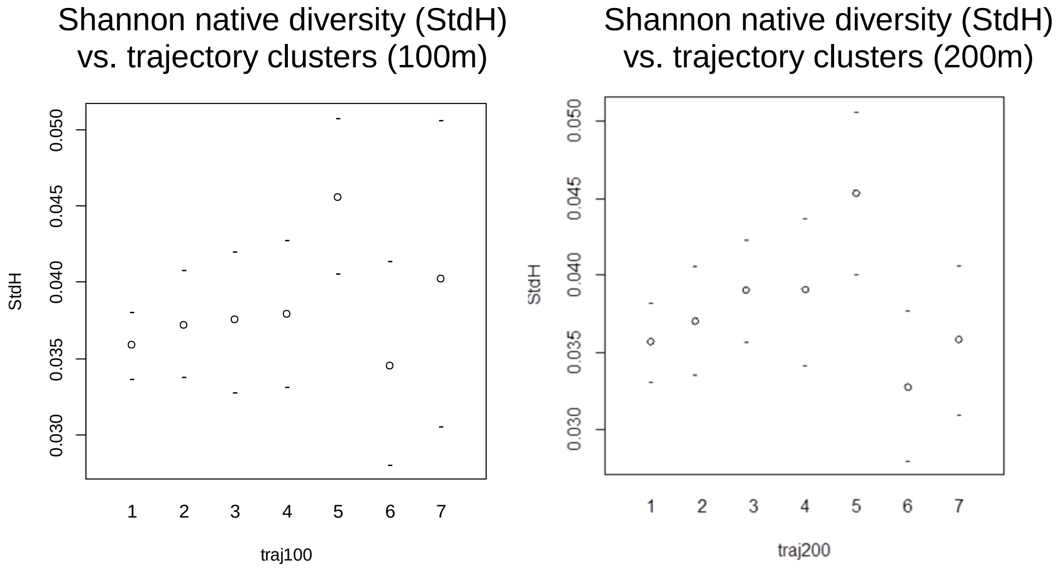
model	predictors	Df	AIC	BIC	log lik	Chisq	Chi Df	Pr(>Chisq)
base	SFv1 + SFv2 + TMI + SI + 1938 NP + 2009Prox + (1|site)	9	-1038	-1010	527.98	7.8853	2	0.019
100	SFv1 + SFv2 + TMI + SI + 1938 NP + 2009Prox + 100_FC + (1|site)	10	-1038.7	-1008	529.34	2.732	1	0.098
200	SFv1 + SFv2 + TMI + SI + 1938 NP + 2009Prox + 200_FC + (1|site)	10	-1037.9	-1007	528.95	1.9416	1	0.164
400	SFv1 + SFv2 + TMI + SI + 1938 NP + 2009Prox + 400_FC + (1|site)	10	-1036.5	-1006	528.26	0.5626	1	0.453
750	SFv1 + SFv2 + TMI + SI + 1938 NP + 2009Prox + 750_FC + (1|site)	10	-1036	-1005	528.02	0.0863	1	0.769
1000	SFv1 + SFv2 + TMI + SI + 1938 NP + 2009Prox + 1000_FC + (1|site)	10	-1036.2	-1006	528.09	0.219	1	0.640

We obtain parameter values for the base model of StdH and for the models augmented by trajectory information derived at the 100 m and 200 m spatial extents, respectively, that are identified in the previous step as being significantly better than the base model (Table 8). SFv1 negatively affects StdH and is the only statistically significant environmental predictor in the base model for stand level native diversity. No environmental predictors are statistically significant in models of StdH that include forest change trajectories. 2009 Prox is positively associated with StdH. We assess the significance of trajectory class on StdH using the “little to no change” trajectory as the reference level (Figure 5), and only the “rapid reforestation between 1957 and 1975” trajectory is associated with a statistically significant increase in StdH. No trajectory classes are associated with a statistically significant decrease in diversity. These findings are visualized in Figure 5, which provides a plot the estimated partial effects of trajectory type derived at the 100 m and 200 m spatial extents on StdH ,with their estimated highest probability density (HPD) interval.

Table 8: Coefficients for linear mixed models with forest cover change trajectories clusters derived at 100 m and 200 m spatial extents, the only significant spatial extents for StdH.

Parameter	**base model**	**100**	**200**
(Intercept)	0.019119	0.022003	0.011693
SFv1	-0.001201*	-0.001002*	-0.000968*
SFv2	-0.000974	-0.000926	-0.00138*
TMI	0.000035	0.000029	0.000135
SI	0.000392	0.00029	0.000538
1938 NP	0.000663	0.000582	0.000608
2009 Prox	0.000010*	0.000011*	0.000011*
rapid reforestation between 1938 and 1975	**na**	0.000984	0.000741
rapid reforestation between 1957 and 1975	**na**	0.009443***	0.009659***
rapid reforestation between 1975 and 1997	**na**	0.001511	0.002412
reforestation after 1938	**na**	-0.001151	-0.003324
reforestation-deforestation 1957 (1)	**na**	0.001044	0.002626
slight deforestation /stable between 38 and 75	**na**	**na**	0.000119
rapid deforestation until 1975, than reforestation	**na**	0.005309	**na**

Significance levels: ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05



Legend

1 little to no change
2 rapid reforestation between 1938 and 1975
3 reforestation-deforestation after 1957
4 rapid reforestation between 1975 and 1997
5 rapid reforestation between 1957 and 1975
6 reforestation after 1938
7 (100) rapid deforestation until 1975, reforestation after
7 (200) slight deforestation between 38 and 75, reforestation after

Figure 5: Partial effects of individual forest change trajectory type derived at the 100 m and 200 m spatial extents on mean Shannon native diversity. MCMC-based 95% highest posterior density intervals are indicated by tick marks trajectory type. *Trajectories with MCMC-based p values <0.05.

Discussion & Conclusions

We assessed the relationship of forest cover change trajectories derived using 5 time steps on forest native plant diversity across nested spatial extents. Our results show that only forest cover change trajectories derived at 100 m and 200 m spatial extents are statistically significant at $p < 0.05$, suggesting that native diversity is more sensitive to forest change at local rather than landscape scales. Until now, the effects of forest cover change trajectories derived using multiple time intervals on biodiversity has not been investigated. Thus we compared forest cover change trajectories to a more commonly used metric of forest cover change, namely forest change derived using only one time interval (1938-2009), across multiple spatial extents in their ability to explain stand level diversity. It is noteworthy that this metric of forest change measured over a single time interval, was not identified as a significant predictor of stand level diversity at any spatial extent, but trajectories of forest change derived at the 100 m and 200 m spatial extents had statistically significant impacts on stand level diversity. This highlights the importance of quantifying and representing intermediate patterns of forest cover change when evaluating the potential impacts of forest change on ecological phenomena. Our results show that the timing and rate of the change may have important implications for diversity patterns, but these elements may be lost when collapsing the change dynamics across time, as is commonly done.

We also investigated whether specific trajectory classes of forest cover change are associated with the lowest or highest diversity. Although we did identify one trajectory, the “rapid reforestation between 1975 and 1975”, to be associated with the highest observed mean of native diversity, no trajectories had a statistically significant relationship with the lowest mean of native diversity. This is probably due to the fact that relatively few of the trajectory classes were characterized by deforestation. In addition, the amount of deforestation, and hence disturbance that occurred may have not been severe enough to negatively impact diversity. However, it is possible that the negative effects associated with trajectories characterized by deforestation such as the “reforestation until 1957, deforestation after” trajectory, may have still yet to be realized

(Hahs et al., 2009; Hylander & Ehrlén, 2013; Vellend et al., 2006), due to extinction debt. Extinction debt refers to the phenomenon of a delayed negative response of a species or community of species to habitat change. Thus, the results our research are inconclusive as to the effects of deforestation on native biodiversity, and further study that includes more sites that have experienced moderate to severe deforestation is warranted.

We summarized the amount of deforestation, reforestation and persistent forest for each time interval used in this study (1938-1956, 1956-1975, 1975-1997, 1997-2009) and compared these trends to the amount of deforestation, reforestation and persistent forest observed using a time interval capturing only the first and last time period used in this study (1938-2009), in order to provide a description of the overall changes that occurred across the study extent and to compare how these changes differ when assessed using four smaller time intervals as compared to one large time interval. When assessing patterns of forest change using the 1938-2009 interval, reforestation is shown as the dominant trend, exceeding deforestation. This finding is supported by the notion of a "forest transition" in which the trend of deforestation in a given region is replaced by net reforestation, often due to agricultural abandonment and a shift from an agrarian to an industrial based economy (Mather & Needle, 1998; Drummond & Loveland, 2010). However, the assessment of forest cover using this time interval overlooks the spatiotemporal fluxes in forest cover. In our analysis, these fluxes were observed using 4 smaller time intervals covering the same time period overall (1938-1956, 1956-1975, 1975-1997, 1997-2009). As evidenced in Figure 4, our results have shown that reforestation has peaked during 1956-1975, and since then deforestation has outpaced reforestation. Thus the trend of net reforestation observed for the 1938-2009 interval does not actually reflect recent trends, and is a legacy of the abundance of reforestation that occurred in 1956-1975. This suggests that important land use trends are more likely to be revealed if they are assessed using a sequence of finer temporal intervals as compared to a single and undifferentiated interval. Deforestation is likely to continue in the study system, given that agricultural and forest lands are the two types of open space available to accommodate urban growth and development, and that the vast majority of agricultural holdings have already been converted to development, which will likely preclude reforestation. However, the few working farms that remain are likely on highly productive soils and are capable of harboring significant diversity if forests were allowed to regenerate, as agricultural lands with soils of low productivity were likely to have been the first to be abandoned (Kleipeis et al., 2013; Pimentel et al., 1992).

In this study, we were able to identify the forest cover change trajectory associated with the highest levels of plant diversity and the spatial extents at which forest cover change best explained variation in plant diversity. In addition, we described a method to identify and describe the major types of forest cover trajectories using a k-means clustering method for longitudinal data. Our method also has the potential to identify land cover changes which are associated with the most severe impacts to diversity, though none were identified in this study, which we attribute to lack of dramatic deforestation evident in the trajectories obtained for our study system. Assessments of deforestation and forest fragmentation often cite the potential for biodiversity loss as a motivation for the study (e.g., Nagendra et al., 2009; Sanchez-Chero, 2013; Trisurat et al., 2013), but relatively few studies have so far directly investigated the links between forest change (usually assessed as deforestation or fragmentation) and field assessments of biodiversity. Our research has addressed this gap. We have presented a method that can be used to analyze the effects of land cover change assessed over multiple temporal datasets on biodiversity or on other metrics of ecosystem function and structure in order to understand the cumulative impacts of past anthropogenic induced landscape change and inform decision making with regard to biodiversity management.

CHAPTER 4: A STRUCTURAL EQUATION MODELING APPROACH TO UNDERSTANDING RESISTANCE TO INVASIONS IN URBAN FORESTS.

Introduction

Exotic plant invasions appear to be everywhere. Some notable examples are cheatgrass invasions of the grasslands of the western United States. (Bagchi et al., 2013), purple loosestrife in the wetlands across North America (Welk, 2004), and fields and trees draped with kudzu in the southeastern United States (Hickman & Lerdau, 2013). However, they are not present quite everywhere, with some locations remaining non-invaded. Hence the question that has long troubled researchers is to know what makes these communities resistant to invasions while others are afflicted by invasive species.

Three main theories have been proposed to explain the observed variation in habitat resistance to invasions: the diversity invasibility hypothesis (Elton, 1958), and more recently resource availability (Davis et al, 2000) and propagule pressure (Lockwood et al., 2005). Much of the research addressing resistance to invasions has focused on characterizing native-exotic richness relationships with the expectation that communities with high levels of resident native diversity have fewer available niches and thus should be more resistant to invasions. This theory is commonly referred to as biotic resistance or the diversity-invasibility hypothesis (Elton, 1958). It has been validated empirically by a number of small scale studies (e.g., Kennedy et al., 2002; Brown & Peet, 2003; van Ruijven, et al., 2003), but has been rejected in larger scale studies, where the trend has been to find a positive correlation between diversity and invasibility instead (e.g., Tanentzap et al., 2010; Brown & Peet, 2003; Zeiter & Stampfli, 2012; Stohlgren et al., 1999).

To address the lack of consistency of findings with regard to the external validity of the diversity-invasibility hypothesis, two “unifying” theories have been advanced, fluctuating resource availability (Davis et al., 2000) and propagule pressure (Lockwood et al., 2005; Simberloff, 2009), which are now discussed in some detail. The fluctuating resource availability hypothesis proposes that an increase in the availability of necessary resources in a community will make that community less resistant to invasions since there are resources available for a newly arriving invader to utilize (Davis et al., 2000). The availability of excess resources can be attributed either of two causes. A first cause may be the decreased uptake/utilization of resources by the resident community stemming from a disturbance that reduced the resident vegetation (e.g. herbivory, disease). Also, it may follow from a disturbance that increases the availability of one or more nutrients or resources. For instance, overabundant rainfall results in increased soil moisture, along the same line, newly created canopy gaps resulting from a windthrow or timber harvest result in increased light, and high atmospheric deposition as a consequence of anthropogenic activities results in increased soil nitrogen in urban environments. According to this theory, a community’s resistance to invasion is not only dependent on resident species richness, but also on the resource supply which is likely not consistent across time or space, particularly in communities that experience frequent disturbance. Soil macronutrients and pH, to which we broadly refer to as “soil fertility”, have long been regarded as limiting factors of plant growth, influencing both plant diversity as well as invasibility (Tilman, 1984; Burke & Grime, 1996; Dukes & Mooney, 1999). Other environmental resources, such as light availability and moisture, can also influence invasion resistance, having either positive or negative effects on natives and exotics, depending on the traits of the invader and the recipient community. For example light limitation can promote invasions by shade tolerant species while negatively affecting shade intolerant natives (Cole & Weltzin, 2005; Reinhart et al., 2006; Martin et al., 2009). The dynamics of resource availability are not feasible to measure in large scale observation studies, and thus resource availability is measured cross-sectionally, instead of longitudinally. However, disturbance is closely linked to the idea of fluctuating resources, as the sudden increase in resource availability, as discussed above, is often tied to a disturbance of some kind. Based on this assumption, it is possible to ascertain whether a forest community has experienced deforestation, a specific type of disturbance, using a time series of aerial photos of the community. Local habitat disturbance, such as deforestation can promote invasions of shade tolerant shrubs that take advantage of the newly available resources such as light and space, becoming established while light availability is plentiful, and then persist under the canopy after subsequent reforestation. Historical land use disturbances that result in a canopy gap have been shown to foster invasions by woody plant species; on the other hand, forest habitat that has remained intact seems to discourage invasions (DeGasperis & Motzkin, 2007; Mosher et al., 2009).

The diversity-invasibility hypothesis and the fluctuating resources hypothesis are both focused on the intrinsic qualities of communities, both biotic and abiotic, that make them more susceptible or more resistant to invasions. However, habitat suitability alone is insufficient for an invasion to occur: the invader must actually arrive at the site in question (Tanentzap & Bazely, 2009; Shea & Chesson, 2002). This likelihood of arrival is dependent on propagule pressure. Propagule pressure is a composite of the number of releases of the invader into the novel environment and of the number of propagules (seeds) released during each event (Lockwood et al., 2005). Strong propagule pressure can overcome biotic resistance and marginal environmental conditions, as high seed volumes are linked to establishment success. Propagule pressure has been demonstrated to be a consistent predictor of invasion success (Simberloff, 2009).

Other factors that influence the distribution of exotic species such as landscape proximity and past landscape disturbance may also affect invasion resistance. The proximity of forest patches in the landscape can foster landscape-scale exchange of both native and exotic species (With, 2002; Kumar et al., 2006; Hanski, 1998; Verheyen et al., 2003). Landscape scale patterns of habitat connectivity are linked to the dispersal of native species while fragmented patchy landscapes are thought to facilitate dispersal by invasive species. Reduced connectivity is a consequence of habitat loss that typically accompanies urbanization and may be a path through which urbanization can indirectly exert a negative influence on the resistance to invasions by simultaneously promoting invasions and discouraging the persistence of natives that are poor dispersers.

However, as it is currently applied, the diversity-invasibility framework, even when expanded to include consideration of resource availability and propagule pressure, presents at least two shortcomings. First, native-exotic richness relationships are flawed measures of invasibility because, while the number of exotic species does in part reflect a community’s predisposition to be invaded, it only reflects the presence of a species and not its abundance. This can be misleading as a community might have a single exotic species that happens to be dominant, or conversely, exhibit very low abundance of any of a number of exotic species. The second issue is that, although several broad-scale studies focused on understanding the distribution of exotic and native species along rural-urban gradients have documented the increase of exotic species with increasing urbanization (e.g., Burton et al., 2005; Bartuszevige et al., 2006; Duguay et al., 2007), few studies of invasibility have so far utilized anthropogenic factors as predictors. While the existence of links between urbanization and exotic species richness is now well established, the underlying mechanisms driving this relationship remain to be investigated however. In this regard, the processes of propagule pressure and eutrophication have particularly been linked to anthropogenic activities that accompany urbanization. These two processes have strong theoretical and empirical links to invasibility of urban communities and have been found to lower the resistance to invasions by providing resources to invaders and by increasing the number of arriving invaders, respectively (Tanentzap & Bazely, 2009; Shea & Chesson, 2002; Pysek et al., 2010). We expect that urbanization directly influences community resistance to invasion via the positive relationship between urbanization and propagule pressure. Human activities such as transportation and gardening introduce exotic species to novel sites, both accidentally and intentionally (Pysek et al., 2010), and may explain why anthropogenic factors and exotic species richness are positively correlated (Burton et al., 2005; Bartuszevige et al., 2006; Duguay et al., 2007; Ranta & Vilijanen, 2011). For example**,** metrics reflecting the intensity of human activity such as housing density (Gavier-Pizarro et al., 2010) and road density (Lilley & Villend, 2009), have all been shown to be positively correlated with exotic plant richness. This in turn suggests that the likelihood of invasive dispersal to forests and other habitats increases with the concentration of human activity. Although it is not feasible to directly measure propagule pressure, it can be estimated via proxy variables that reflect the spatial concentration of anthropogenic activity (Pysek et al., 2010). We also expect that urbanization has indirect negative effects via resource availability (measured as soil fertility) on invasion resistance. The links between urbanization and the eutrophication to forest ecosystems are well documented and result from increased atmospheric deposition of nitrogen and nutrient loading resulting from anthropogenic activities such as fertilizers applied to lawns carried by stormwater runoff (Stevens et al., 2004; Sutton et al., 2011). Atmospheric deposition is higher near busy roads and highways as well as in urban forests. Eutrophication of habitats is linked to increased susceptibility to invasions by altering the spatial and temporal patterns of resource availability (Davis et al., 2000; Grimm et al., 2008). In addition, while a surplus of nutrients can facilitate invasions by exotic species, they are also associated with declines in native species richness (Parepa et al., 2013; Burke & Grime, 1996).

Furthermore, a truly integrated approach to understanding and modeling invasion resistance must address the issue of spatial dependence, which is pervasive in environmental and ecological data (Legendre, 1993; Dormann, 2007). In the case of unaccounted spatial dependence, the correlation between two variables can be inflated if their values are also correlated with space, thus potentially resulting in a Type I error. We expect spatial dependence to be present in our system as the result of the contagious process of invasive spread, coupled with the likelihood of habitat conditions being correlated with space. In other words, communities located next to highly invaded communities are more likely to be invaded either by dispersal of the invader from invaded sites to adjacent sites and/or increased habitat susceptibility to invasions;

likewise, highly resistant communities are likely situated near communities with a low degree of invasion (Dark, 2004). We account for spatial dependence in our system by using the spatial lag model to estimate the spatial dependence among neighboring observations and explicitly incorporate spatially lagged invasion resistance as predictor in order to distinguish the effects of spatial dependence from the other predictors.

In this research, we investigate the hypothesis that urbanization has both direct effects on community resistance to invasions via propagule pressure, and indirect on effects via resource availability. The analysis uses a structural equation modeling framework that integrates other factors known to influence invasibility, namely abiotic factors related to light and moisture availability, topographic slope, history of habitat disturbance, landscape proximity and spatial dependence of neighboring observations. This series of structural hypotheses, forming a network of causal pathways, is shown in Figure 1. Each path is labeled with the theorized mechanism underlying the relationship between the associated pair of predictors. To test our hypothesized system of invasion resistance, we collected data on native and exotic species diversity from 177 plots located within 40 forest fragments located throughout the Charlotte, NC region. We measured resistance to invasions using a metric that represents both the richness and abundance of exotic species as compared to the total richness and abundance of all species present in a plot. Thus instead of using native species richness as a predictor in the model as is commonly done in invasibility studies, we incorporate native species richness into the response variable, where we assume that the ratio of exotic to all species present is the outcome of biotic resistance (Guo and Symtad, 2008). First, we investigate whether invasion resistance, as we define it, is significantly different among urban, suburban and rural forests. Then, we evaluate the fit of our hypothesized model with real data using structural equation modeling (SEM), a multivariate method that can be used to test hypotheses about systems, not just one response variable (Grace et al., 2010). To our knowledge, no studies have evaluated the role of urbanization, and its relationship with other factors influencing invasibility in resistance to invasions.

The remainder of this paper is organized as follows. First we present our modeling framework which includes background on structural equation modeling, our model building process, and the process used to investigate and account for spatial dependence in our model of invasion resistance. We then describe our study system and sampling design, and the methods used to derive invasion resistance, and the environmental and landscape covariates. The next section presents the results, including the estimated indirect and direct coefficients for all the predictors in the model within the context of spatial dependence, and model fit statistics. Finally, we discuss our findings within the larger context of urban ecology and resistance to invasions and present our conclusions.

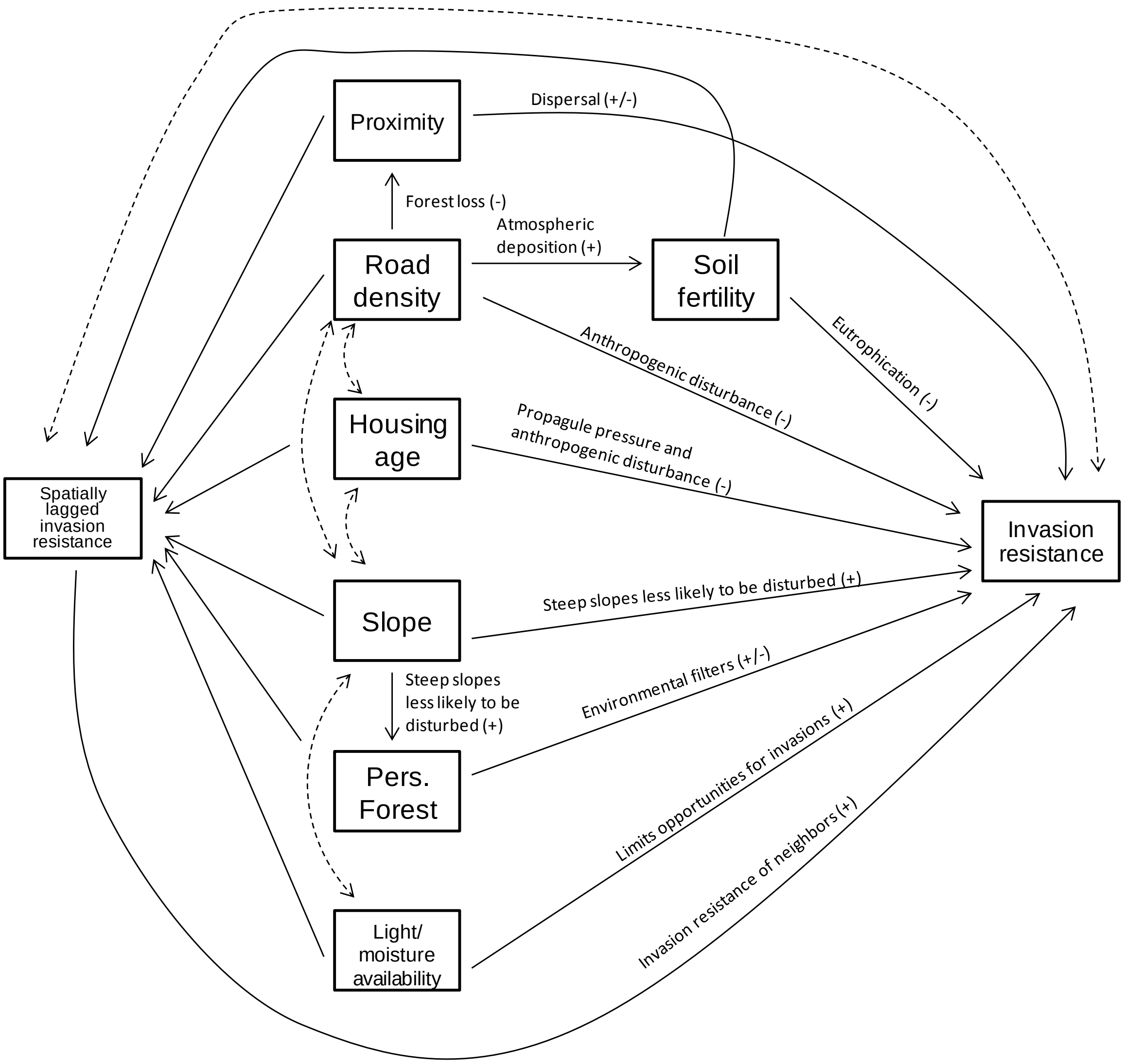


Figure 1: Model showing the hypothesized relationships linking different predictors to influence resistance to invasions. Dashed, double-headed arrows indicate expected covariances (unanalyzed correlations).

Methods

2.1. Modeling Framework

2.1.1. Structural Equation Modeling

To test our system of integrated hypotheses, we use SEM which unlike other multivariate methods such as ordination, can be used for hypothesis testing (Grace, 2010). The causal processes under investigation are represented by a series of linear regression equations. The predicted structural relationships among exogenous (predictors whose values are independent of the values of other predictors in the model) and endogenous variables (variables whose values are dependent on the values of other predictors in the model), can be represented graphically, promoting a clear conceptualization and communication of theorized relationships (Shipley, 2002; Grace et al., 2010). SEM can be used to evaluate how well the hypothesized path model fits with observed data. Path analysis is a type of structural equation model that does not use latent variables and only uses observed variables. This approach is appropriate in cases such as ours, where there is relatively low correlation among predictors (Table 1) which is necessary for the construct of latent variables (Bollen, 1989). The coefficients estimated using multiple linear regression correspond to the coefficients estimated for the direct effects by SEM. The difference between SEM and linear multiple regression is that only the direct effects are estimated using linear multiple regression and the indirect effects "which are simply the unresolved causal relationships between the predictors, are ignored" (Shipley, 2002). The advantage of SEM over multiple linear regression is that with SEM we can decompose the variation among the predictors to simultaneously evaluate a network of hypothesized relationships among variables, which are represented by a series of linear regression equations, into direct and indirect effects. For a model in which x affects y, and y affects z, the indirect effects of x on z via y is the product of the coefficients (these would be considered direct effects) estimated for x->y and y->z. The technique of SEM allows us to represent and estimate "unanalyzed associations" between predictors, which are typically represented with curved, double-headed arrows. These covariances are expected to occur, but are not surmised to be the result of a causal relationship, or it is unknown which predictor is actually the endogenous variable. Specifically, the technique of SEM allows us to investigate the direct and indirect effects of urbanization on invasion resistance and to test theories about *how* urbanization influences invasion resistance, not just test the statistical significance of urbanization as a predictor of invasion resistance.

2.1.2. Model Building

Each endogenous variable (soil fertility, invasion resistance, spatially lagged invasion resistance, persistent forest area, and proximity) in the model shown in Figure 1, along with their respective predictor variables corresponds to a specific linear regression sub-model. The endogenous and associated exogenous variables associated with each sub-model can be found in Table 2.

In order to build the hypothesized model shown in Figure 1, we collected data on invasion resistance and soil fertility in the field, and derived the remaining predictors (metrics of urbanization, landscape proximity, light and moisture availability, and area of persistent forest) in a geographic information system (GIS). We use Pearson's correlation as an initial test of statistical significance to determine whether a predictor should be included in the final model. We use the Moran's I statistic to assess whether the residuals from each linear regression sub-model, demonstrate statistically significant spatial dependence. This test evaluates whether the residuals exhibit significant spatial clustering as compared to the null hypothesis of no clustering (Cliff & Ord, 1972). The spatial lag model is a theory-driven method for directly incorporating the effects of spatial dependence into regression models in order to obtain less biased coefficients, and is appropriate when the values of the response variable are influenced by the values taken by this variable in neighboring locations (Anselin, 2002). The spatial lag model has the form:

$$\mathbf{y_i = \rho W y_i + X_i \beta + \varepsilon_i, \ \varepsilon \sim N(0,\sigma^2 I)} \qquad (1)$$

Where y is the $n \times 1$ vector of the dependent variable (invasion resistance) and X represents an $n \times \boldsymbol{k}$ matrix of explanatory variables, i is a geographic location, Wy_i is the spatially lagged predictor of y_i, in which W is the n x n contiguity matrix defining the neighbors of location i, ρ is the spatial dependence parameter and is a measure of how correlated the value of y_i is with its neighbors. The errors are assumed to follow a normal distribution with constant variance. Clustered data such as ours are often handled using a multilevel modeling approach to avoid Type I errors resulting from ignoring the lack of independence of the errors induced by the clustering. Multilevel models adjust for the clustered errors allowing for correct statistical inference (Ryu, 2014). The disadvantage of multilevel models to account for clustered data is that in order to obtain valid standard fit statistics, the statistics must be obtained for each individual level of the models (Hox & Maas, 2004; Ryu, 2014). In addition, multilevel models treat spatial autocorrelation as a “nuisance”, arriving solely from the sampling design, not from, or in part due to an ecological process. The levels are predefined, thus the effects of spatial autocorrelation are filtered out, rather than being revealed endogenously through the autocorrelation structure of the data. The advantage of using spatial lag models is that we can explicitly account for spatial dependence in our model and examine its role in invasion resistance. We account for spatial autocorrelation by obtaining parameters for each sub-model using spatial lag models. In order to explicitly understand the direct and indirect effects of the neighboring values of invasion resistance on invasion resistance, we add a spatial lag (the average of the neighboring values) of invasion resistance, as a predictor in the structural equation model (Figure 3). Finally we estimate the structural equation model shown in Figure 3 and evaluate model fit using the Stata 13 statistical software (Stata Corp. 2013). We describe each of these in more detail below.

2.2. Data Collection

2.2.1 Study Extent, Sampling Design

We carry out our investigation in the greater Charlotte, North Carolina area. The Charlotte metropolitan area is the 10th fastest growing city in the United States, with a population over 1 million. This rapid population growth has been accompanied by low and mid-density development patterns that have consumed much of the forests and agricultural land in the vicinity. The area is characterized by gently rolling terrain, erosion prone soils, and forests dominated by mixed hardwood and pine. In order to ensure that the full spectrum of urbanization within the study extent is represented, we sampled woody plant species from 40 forest patches located across a rural-urban gradient spanning the county seat of Charlotte, NC and the adjacent county of Cabarrus (Figure 2). The rural-urban gradient was developed based on building densities obtained from parcel data (Mecklenburg County Geospatial Information Services; Cabarrus County Land

Records Information System) using a 1 km moving window in a GIS. From this, three classes of land use intensity strata were derived: rural, suburban and urban using the classification developed by Theobald (2005). The distribution of the three land use intensity classes is urban: 28.3%, suburban: 43.2% and rural: 28.5%.

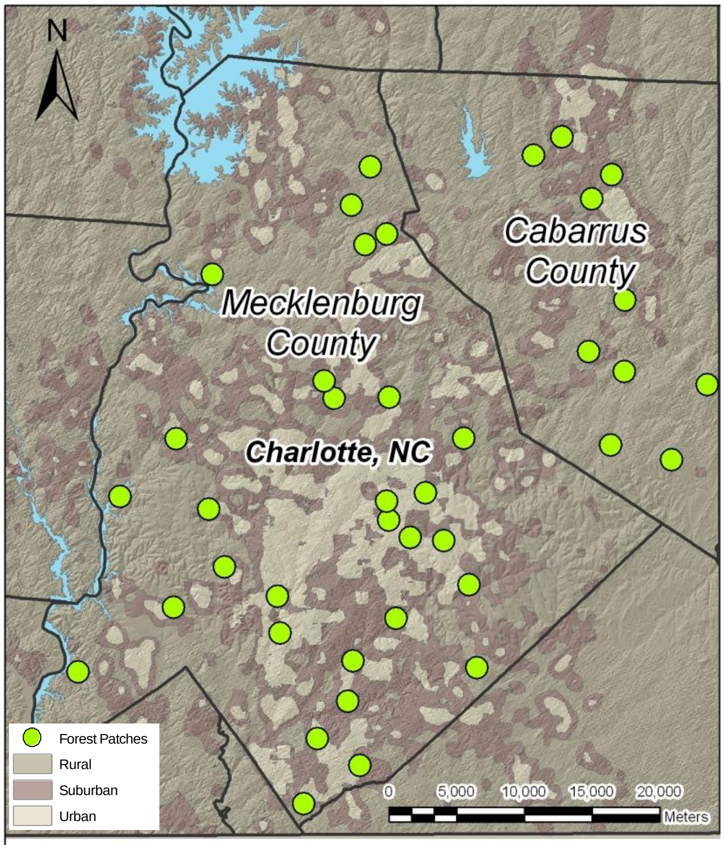


Figure 2: Study extent showing the location of forests that were sampled within the greater Charlotte metropolitan area.

2.2.2. Deriving Invasion Resistance

To assess invasion resistance, we collected data on both native and exotic species richness and abundance from 3-5 randomly located plots nested within each forest patch, for a total of 177 plots. Species composition and count were recorded for all woody species within the plot that were at least 1.3 meters in height. This constitutes the “forest stand” (Burton & Samuelson, 2005). The transgressive layer (stems less than 1.3 m in height) was sampled using five 1 m^2 randomly located sub-plots in each of the 177 plots. Guo and Symtad (2008) argued that instead of focusing solely on the number of unique exotic species present in a community, both the number and abundance of exotic species should be considered when assessing invasibility. By including measures of exotic species abundance, one can assess the degree of invasion that has occurred. The degree of invasion observed in a community reflects both the intrinsic habitat vulnerability to invasions as determined by biotic factors such as the richness and competitive ability of the resident species, as well as extrinsic factors such as propagule pressure and disturbance that influence the likelihood of arrival of an exotic species. We measure the degree of invasion p as the proportion of unique exotic species present in each plot (exotic richness), weighted by their abundance (the number present of each exotic species), to the product of the total richness and abundance of each species present in each plot:

$$p = (\text{exotic richness} * \text{exotic abundance})/(\text{total species richness} * \text{total abundance}) \quad (2)$$

As invasion resistance is the degree to which a community is not invaded, we obtain invasion resistance as $1 - p$, and apply a logit transformation to map the proportion data (Crawley, 2005), which are bounded by 0 and 1 to the real number line, giving us:

$$\mathbf{Logit(1- p)} \qquad (3)$$

Given that that there is a lack of evidence to support the existence of habitats that are 100% resistant to invasion, especially when exposed to high propagule pressure (Williamson, 1996; Von Holle & Simberloff, 2005), we regard uninvaded plots as being

99.99% resistant to invasions as opposed to 100%. Therefore, our final metric for invasion resistance is

logit (0.9999 – p) (4)

In addition, the advantage of using count as a measure of abundance is that we can incorporate species richness and abundance, measured at both the stand level and the transgressive level, in a single metric.

2.2.3. Soil Fertility

At each field plot, three 10m depth soil cores were taken from different locations and were pooled to form a representative sample for each plot. The pooled soil samples were sent to Brookside Laboratories (New Bremen, OH) to quantify the pH and the levels of available nitrogen, potassium, phosphorus, and calcium.. These five soil factors have been shown to influence species richness and vegetation composition in Piedmont forests (Peet and Christensen, 1980; Taverna et al., 2005). In order to reduce the dimensionality of our soil data, we use principal components analysis to reduce these 5 variables into 2 predictors reflecting overall soil quality (Ye & Wright, 2010). The 2 principal components explain 42.1% and 22.7% of the data variance, respectively. Variables with significant loadings ($p < 0.05$) on the first principal component (SFv1) are pH (R^2 = 0.41), nitrogen (R^2 = 0.36), calcium (R^2 = 0.67), potassium (R^2= 0.62); and on the second principal component are pH (R^2 = 0.25), nitrogen (R^2= 0.28), and phosphorous (R^2= 0.59). These principal components are labeled SFv1 and SFv2, respectively.

2.2.4. Light and Moisture Availability

Light availability is estimated using the potential mean index of solar radiation (PMSI), which is an estimate of the solar insolation received at a given location, based only on seasonal intervals, namely the winter and summer solstices and spring and fall equinoxes. The PMSI is derived from a 10m digital elevation model (DEM) using the area solar radiation tool in the ArcGIS 10 software application. Canopy closure is measured using the percent tree canopy coverage obtained from the Multi-Resolution Land Characteristics (MRLC) Consortium and is based on the 2001 National Land Cover Database (NLCD). Moisture availability is estimated using the topographic moisture index, and is derived from a 10m resolution DEM. The topographic moisture index (TMI) is calculated as the natural log of the ratio of the upslope contributing area to the slope (Moore et al., 1991). Slope as measured in degrees is also derived from a 10m resolution DEM.

2.2.5. Landscape Proximity

We derive the landscape proximity index for a 1 km^2 area within each plot using a supervised classification of forest cover of 2009 satellite imagery of the study extent using Fragstats v4.1 (McGarigal et al., 2012). The proximity index is a dimensionless number that provides a relative measure of forest connectivity within a landscape, where lower numbers are associated with forest patches that are more isolated, and higher numbers indicate less isolated patches.

2.2.6. Area of Persistent Forest

We use the area of persistent forest as a metric to reflect the degree of past landscape disturbance. Digital aerial photographs from 5 different time periods (1938, 1957, 1975, 1997 and 2009) corresponding to plot locations were obtained from Mecklenburg and Cabarrus county GIS departments. We map the forest cover present within a 100 m diameter buffer of each plot for each of the 5 time steps using “heads up” digitizing in ArcGIS, thus creating a separate layer of forest cover for each year. A spatial overlay is produced, in which each of these layers is stacked on the previous time layer in chronological order. The area of persistent forest cover is delineated on the basis that forest cover remains present in the same spatial location as evident in the spatial overlay and remains undisturbed (i.e. did not transition from deciduous to coniferous forest) throughout all five time periods.

2.2.7. Metrics of Urbanization

To map road density, we use a moving window to map the density for every grid cell based on road centerlines obtained for the study extent within a 1-km neighborhood of the grid cell. The average single family housing age is derived from georeferenced property records (Mecklenburg County Geospatial Information Services; Cabarrus County Land Records Information System) also using a 1-km moving window.

2.3 Pearson's Correlation

We evaluate the relationship among our hypothesized predictors and the response variable using Pearson's correlation (Table 1). Canopy closure, PMSI, sFv2 and Edge density have very low, non-significant correlation with invasion resistance and are omitted from our models on that basis.

Table 1. Pearson's correlation among the variables examined in this study.

Type	**Predictor**	Invasion resistance	Road density	Housing Age	Canopy closure	PMSI	TMI	Slope	sFv1	s Fv2	2009 Proximity	Pers. For Area
Response	Invasion resistance	1.000	-0.171	-0.243	0.068	-0.073	-0.210	0.156	-0.235	0.011	0.201	0.343
Urban	Road density	-0.171	1.000	0.162	-0.119	0.159	0.117	-0.224	0.176	0.234	-0.498	-0.110
Urban	Housing Age	-0.243	0.162	1.000	-0.053	-0.040	0.093	0.154	0.067	0.068	-0.145	0.020
Resource availability	Canopy closure	0.068	-0.119	-0.053	1.000	0.058	0.080	0.012	-0.022	-0.034	0.173	0.237
Resource availability	PMSI	-0.073	0.159	-0.040	0.058	1.000	0.191	-0.477	-0.079	0.089	-0.147	-0.156
Resource availability	TMI	-0.210	0.117	0.093	0.080	0.191	1.000	-0.492	0.033	-0.049	-0.140	-0.095
Slope	Slope	0.156	-0.224	0.154	0.012	-0.477	-0.492	1.000	0.016	0.003	0.279	0.284
Soil nutrient availability	sFv1	-0.235	0.176	0.067	-0.022	-0.079	0.033	0.016	1.000	0.015	-0.062	0.046
Soil nutrient availability	s Fv2	0.011	0.234	0.068	-0.034	0.089	-0.049	0.003	0.015	1.000	-0.037	0.045
Landscape connectivity	Proximity	0.201	-0.498	-0.145	0.173	-0.147	-0.140	0.279	-0.062	-0.037	1.000	-0.012
Persistent forest	Pers. For Area	0.343	-0.110	0.020	0.237	-0.156	-0.095	0.284	0.046	0.045	-0.012	1.000

2.4 Spatial Dependence

We confirm the presence of significant spatial autocorrelation in regression residuals for each sub-model using Moran's I statistic in the spdep package in R (Bivand et al., 2013) (Table 2).

Table 2: Results of Moran's I tests for each of the sub-models.

Endogenous	**Exogenous**		
		Moran's I	**p**
soil fertility			
	Road density	0.13	0.02
Invasion resistance			
	Soil fertility	0.26	<0.001
	Proximity		
	Pers. Forest Area		
	TMI		
	Slope		
	Housing age		
	Road density		
Proximity			
	Slope	0.51	<0.001
	Road density		
Pers. Forest Area			
	Slope	0.3	<0.001

Significance levels: '***' 0.001 '**' 0.01 '*' 0.05

In order to account for spatial dependence in our SEM, we use spatial lag models to obtain parameter estimates for each of our sub-models using the maximum likelihood estimation via the lagsarlm command in the spdep package in R. The lagsarlm command estimates the spatial dependence parameter, ρ, using a two-stage procedure in order to avoid obtaining biased parameter estimates and standard errors that would result from not adequately accounting for the endogeneity introduced by including a spatial lag of the response variable. Since the forest patches have 3-5 samples each, we use a contiguity matrix based on the 4 nearest neighbors of $y_{i,}$ weighted by their inverse distances to y_i when estimating the spatial lag model. In order to understand the direct and indirect effects of the neighboring values of invasion resistance on invasion resistance, we use the spatially lagged predictor of invasive resistance (Wy_i) as an endogenous variable in the structural equation model (Figure 1).

Specifically, we are testing whether the predictors that affect invasion resistance are also significantly correlated with the spatial lag of invasion resistance. If so, this suggests that the neighbors of i are similarly influenced by the predictors, and/or that the environmental conditions of the neighbors of i are similar to those at i. We estimate the following goodness of fit measures for our overall model using Stata 13: $\chi 2$ statistic, root mean squared error (RMSEA), the Tucker - Lewis Index (TLI) and the standardized root mean square residual (SRMR).

Results

Prior to conducting SEM analysis, we use a one-way analysis of variance (ANOVA) to confirm that invasion resistance significantly differs by level of urbanization as measured by urban, suburban and rural land use intensity classes. The analysis is significant, *F*(2, 174) = 3.06, *MSE* = 14.81, $p < 0.05$. Post-hoc Tukey’s HSD tests show that invasion resistance for urban forests is significantly lower as compared to rural forests. None of the other comparisons are significant.

Our conceptual model presented in Figure 1 is supported by our data as indicated by the χ^2 statistic ($\chi^2 = 23.07$, df = 22, $p = 0.398$). The χ^2 statistic assesses the fit of the model, in relation to a model that fits the covariances perfectly (Grace & Bollen, 2005). Thus, our model is not statistically different from a model with a “perfect” fit. All of the additional fit statistics also support that this is a good fitting model. The RMSEA is 0.017 (values below 0.05 indicate a good fit); the TLI is 0.992, (values greater than 0.95 indicate a good fit), and finally the SRMR is less than 0.08, also indicating a good fit (Hooper et al., 2008). However, our tested model differs slightly from the conceptual model (Table 1) because metrics of light availability were found not to be correlated with invasion to resistance, and thus were not included in the model specification tested and shown in Figure 3.

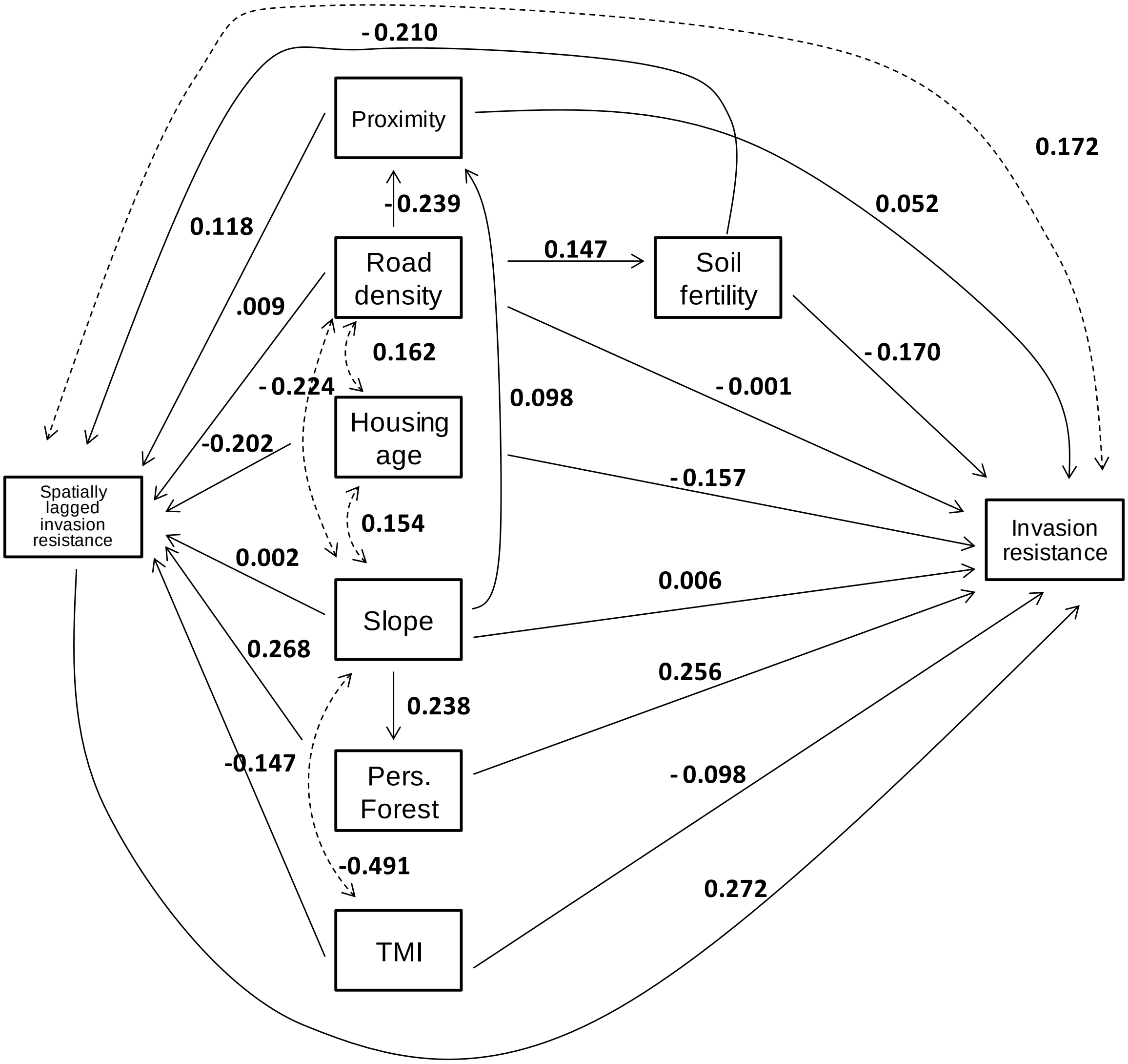


Figure 3: Final structural equation model showing the relationships among the predictors of invasion resistance. Dashed, double-headed arrows indicate covariances. The path coefficients are standardized and correspond to direct effects.

Housing age has statistically significant direct effects (DE) on invasion resistance (DE = 0.157, p=0.015), but road density is not found to have significant direct effects (Table 4).

Table 4: Parameters for the final structural equation model for direct, indirect and total effects.

Endogenous	**Exogenous**	**Std. Coef**		
		Direct	**Indirect**	**Total**
soil fertility				
	Road density	0.147*	na	0.147*
Invasion resistance				
	Spatially lagged invasion resistance	0.272***	na	0.272***
	Soil fertility	-0.17**	-0.057**	-0.227***
	Proximity	0.052	0.032	0.084***
	Pers. Forest Area	0.256***	0.073***	0.329***
	TMI	-0.098	-0.04	-0.138***
	Slope	0.006	0.087***	0.093***
	Housing age	-0.157*	-0.055**	-0.212***
	Road density	-0.001	-0.051**	-0.052**
Spatially lagged invasion resistance				
	Soil fertility	-0.21**	na	-0.21**
	Proximity	0.118	na	0.118
	Pers. Forest Area	0.268***	na	0.268***
	TMI	-0.147	na	-0.147
	Slope	0.002	0.075***	0.077
	Housing age	-0.202**	na	-0.202***
	Road density	0.009	-0.059**	-0.05
Proximity				
	Slope	0.098	na	0.098
	Road density	-0.239***	na	-0.239***
Pers. Forest Area				
	Slope	0.238***	na	0.238***

Significance levels: '***' 0.001 '**' 0.01 '*' 0.05

However, road density has statistically significant indirect effects on invasion resistance. There are four pathways that potentially contribute to the coefficient calculated for the indirect effects of road density on invasion resistance. They are 1) road density via soil fertility, 2) road density via soil fertility via spatially lagged invasion resistance, 3) road density via landscape proximity and 4) road density via landscape proximity via spatially lagged invasion resistance. Since there are a number of different sources of indirect effects of road density on invasion resistance, we have decomposed the sources of variation (Table 5). The indirect effect of road density via soil fertility is clearly the main contributor of the negative indirect effects of road density on invasion resistance, supporting our hypothesis that urbanization has indirect negative effects on invasion resistance via habitat eutrophication.

Table 5: Decomposition of the indirect effect of road density on invasion resistance by source.

Path of Indirect Effects	**Std. Coef**
Road density via soil fertility	-0.025
Road density via soil fertility via spatially lagged invasion resistance	-0.008
Road density via proximity	-0.012
Road density via proximity via spatially lagged invasion resistance	-0.007
Total indirect effects	-0.052

The only measure of resource availability in the model that has statistically significant effects on invasion resistance is soil fertility; slope and TMI are not significant. However, slope has a positive indirect effect on invasion resistance (IE = 0.087, p = <0.001), via the effects of slope on landscape proximity and persistent forest area. As was found with invasion resistance, housing age and soil fertility also have significant negative direct effect on spatially lagged invasion resistance, while similarly, persistent forest area has a significant positive direct effect, and road density, slope and TMI and landscape proximity are not associated with a significant direct effect on invasion resistance. This indicates that the predictors of invasion resistance are similarly correlated with the values of invasion resistance at neighboring sites as represented by spatially lagged invasion resistance. Overall, spatially lagged invasion resistance is the strongest predictor of invasion resistance as indicated by a comparison of the standardized coefficients derived for each predictor (Table 4). The area of persistent forest is the second strongest predictor of invasion resistance (Table 4), thus supporting our hypothesis that the area of locally intact habitat discourages invasions. As we expected, the area of persistent forest increases positively with slope, suggesting that steeper areas are less likely to be disturbed. Slope, which does not have a statistically significant direct effect on invasion resistance, has indirect positive effects on invasion resistance via its effects on the area of persistent forest.

Despite being correlated with invasion resistance as shown in Table 1, TMI and landscape proximity are not associated with statistically significant direct effects in our model, suggesting that these predictors do not directly influence invasion resistance. Both of these predictors have indirect effects on invasion resistance, albeit, statistically insignificant effects, that contribute to the total effects measured for these predictors, which are statistically significant. Whereas the same suite of predictors have the same direction of effect and similar level of statistical significance for both invasion resistance and spatially lagged invasion resistance, the total effects of landscape proximity, TMI and slope, and road density are statistically significant for invasion resistance, but not for spatially lagged invasion resistance. This can be explained, in part, due to the absence of indirect effects to be considered for landscape proximity and TMI, and that the direct and indirect coefficients for the predictors of spatially lagged invasion resistance were lower than those observed for invasion resistance.

Discussion and Conclusions

In this research, we confirmed that urban forests have significantly lower invasion resistance as compared to rural forests and we then investigated a causal model to explain this phenomenon using SEM. We tested the hypothesis that urbanization has both direct and indirect negative effects on invasion resistance, using an SEM framework that integrates other factors that are known to influence invasibility, namely measures of resource availability (soil fertility and TMI), topographic slope, area of persistent forest, landscape proximity and spatial dependence of neighboring observations.

Housing age and road density were the two metrics of urbanization tested in the model, and only housing age had statistically significant direct effects on invasion resistance. The significant negative path coefficient between housing age and resistance to invasions provides insight into the mechanisms by which urbanization negatively influences resistance to invasions. This supports our hypothesis that locations surrounded by older housing have been receiving an influx of exotic propagules arriving from residential yards and gardens longer than those that are surrounded by younger housing. We also tested the hypothesis that urbanization has indirect negative effects on invasion resistance and this was also supported by our findings. Specifically, we investigated whether urbanization has negative consequences for resistance to invasions via habitat eutrophication by assessing whether road density is associated with increased soil fertility. Our results show that road density has a significant positive direct effect on soil fertility and soil fertility has a direct negative effect on invasion resistance. The path coefficient for the indirect effect of road density on invasion resistance via soil fertility accounts for 50% of the indirect effects of road density on invasion resistance. This is in line with previous studies that have observed an urbanization-eutrophication link in soils obtained from forests located an urban-rural gradient (Pouyat et al., 1995; Lovett et al., 2000) and provides a plausible mechanism of how urbanization can be detrimental to native diversity and resistance to invasions.

Although the spatially lagged invasion resistance is a slightly stronger predictor of invasion resistance than persistent forest area, with regard to direct effects, the total effects coefficient for persistent forest area exceeds that for spatially lagged invasion resistance. This indicates that the lack of disturbance in the immediate vicinity is a strong protective mechanism against invasions and that this mechanism explains community invasion resistance better than the spatial dependence of neighboring observations. Mosher et al. (2009) also found that communities that had persistent tracts of forest cover were more resistant to invasions. Similarly, DeGasperis & Motzkin (2007) linked past disturbance in forest cover during the early 20th century to the distribution and abundance of an invasive shrub. These findings along with ours, suggest that forest cover disturbance represents a major opportunity for both the arrival of invasive species via available space and newly available light as a result of a canopy gap.

Soil fertility was the strongest negative predictor of invasion resistance, supporting our expectation that eutrophication is detrimental to invasion resistance (Table 4). The role of eutrophication on lowering resistance to invasions may be two-fold, namely 1) facilitating establishment by invasive species, as eutrophication implies the availability of excess nutrients available for utilization by invaders (Burke & Grime, 1996), and, 2) several studies have linked eutrophication, especially in the form of excess nitrogen to declines in native plant diversity (Smith et al., 1999; Bobbink et al., 2010; De Schrijver et al., 2011).

Slope is often considered a proxy or indirect measure of site conditions and thus is frequently included in invasion studies. However slope is often correlated with other predictors characterizing landscape structure or anthropogenic modification. In this study, slope was positively correlated with area of persistent forest and had a negative covariance with road density, thus revealing the indirect role of slope in influencing resistance to invasions.

In our model, we examined the effects of the variables used to predict invasion resistance on spatially lagged invasion resistance. Our results revealed that the predictors of invasion resistance had a similar but overall lower effect on spatially lagged invasion resistance, suggesting that neighboring communities are exposed to similar habitat conditions. In many cases, accounting for spatial dependence in ecological models often results in the loss of statistical significance of environmental and landscape predictors (Dormann, 2007). However, our case, soil fertility, persistent forest area, and housing age have strong direct effects on invasion resistance, even when accounting for spatial dependence. In addition, all the predictors of invasion resistance in this study, have statistically significant total effects (which is the sum of the direct and indirect effects). This is due in part to the indirect effects of the predictors via spatially lagged invasion resistance, which would not have been apparent if spatially lagged invasion resistance was not explicitly a predictor in our SEM model. Thus an advantage of explicitly incorporating a spatially lagged predictor in a SEM model is the ability to decompose the effects of spatial dependence from that of the other predictors.

To the best of our knowledge, this study is the first to investigate the potential direct and indirect effects of urbanization on invasion resistance. These results emphasized the importance of integrating multiple hypotheses when investigating invasion resistance and have furthered our understanding of what makes communities more or less resistant to invasions. Our work has shown that the area of forest that has remained undisturbed for at least the past 75 years, is a strong predictor of invasion resistance and suggests that intact forest in the immediate vicinity of a plot confers a protective benefit. Conversely, we have shown how urbanization can negatively affect invasion resistance, as our results show that eutrophication increases with road density and has negative ramifications for forest ecosystems, in which excess soil nutrients can be “too much of a good thing” (Sutton et al., 2011) resulting in decreased resistance to invasions. Also, housing age was shown to be linked to invasion resistance, in which forests surrounded by older housing have lowered resistance to invasions. Other than the indirect effects of slope, no abiotic factors outside of soil fertility were shown to influence invasion resistance, once other factors were accounted for. However, we do not suggest these metrics be excluded from future studies. Rather, more study is warranted in which the roles of these environmental predictors are considered using an integrated approach that also accounts for the impacts of urbanization.

CHAPTER 5: CONCLUSIONS

This dissertation examines and quantifies the impacts of urbanization and landscape change to patterns of native biodiversity and the distribution of exotic species by focusing on three questions that heretofore have not been addressed, yet have the potential to further our understanding of the consequences of anthropogenic activities to ecosystem processes. The following questions were addressed : 1) Does human-mediated invasion pressure influence the distribution of forest invaders in urban landscapes?; 2) Do multi-temporal forest change trajectories, classified into a typology, influence forest native diversity, if so, what spatial extent?; and 3) What are the direct and indirect effects of metrics of urbanization on forest community invasion resistance? The unifying thread in this work is the desire to understand and quantify the impacts of urbanization to native plant biodiversity and vulnerability of forests to exotic species invasions, with the intent of obtaining much- needed empirical data in order to guide decision making about the management of forest biodiversity in order to sustain the provisioning of ecosystem services given that the urban footprint is expected to dramatically expand by 2030 (Seto et al., 2012). Within this context, each question addresses a missing piece of how urbanization and landscape change influences native biodiversity and exotic species invasions. This work is among the first to quantify the relationships between urbanization and landscape change on exotic and native forest plant diversity, using a statistically robust methodology that employs a stratified random vegetation data collection survey along an urban to rural gradient to ensure both adequate replication and that the landscape context surrounding the forest samples reflects the full spectrum of landscape heterogeneity and land-use intensity present in the study extent, the Charlotte, NC metropolitan area.

One of the fundamental differences of this work as compared to other urban to rural gradient studies is the method used to derive the gradient. Urban to rural gradient studies have been criticized for relying on over simplified metrics of urbanization such as distance from the central business district or percent impervious surface (Alberti et al., 2007). I have addressed this by delineating the urban to rural gradient using a 1 km moving window to map both building and road density. This produces an urban to rural gradient that reflects the intensity of development within 1 km of every 10m grid cell in the study extent, capturing both the intensity of human activity and development that is not always reflected by percent impervious surface, as well as the “leap frog” patterns of development that are missed when assuming urbanization decreases steadily from the central business district (Ewing, 2008).

Although biological invasions are often cited as major cause of biodiversity loss, it is difficult to prove this using data from cross-sectional observational studies, as it cannot be distinguished whether invasions are the cause of biodiversity loss, or that areas of low biodiversity were simply more invasible (as is predicted by the diversity-invasibility hypothesis) and thus low biodiversity was present prior to invader arrival. Furthermore, habitat loss and fragmentation induced by urbanization, as well as the chronic anthropogenic disturbance generated by urban development to ecosystem processes are also thought to result in biodiversity loss (Fahrig, 2003; Fischer and

Lindenmayer, 2006) as well as facilitate invasions by exotic species (McKinney, 2008; Kowarik, 2011). In this study, I found no statistically significant correlation between any metrics of native and exotic diversity. This is not surprising, as other studies have also noted an absence of a native-exotic species relationship, and the presence of a significant negative relationship is not robust if landscape context and anthropogenic factors that also affect the distribution and persistence of plant species were not accounted for (Fridley et al., 2007). In addition, the majority of studies fail to take propagule pressure into account, which has been a consistent predictor of invasibility success (Pysek et al., 2010; Simberloff, 2009) and has been shown to overwhelm abiotic and biotic resistance to invasions (Von Holle & Simberloff, 2005). Thus, conclusions drawn from invasibility studies that do not consider propagule pressure are often inconsistent. In this work, I employed the theoretical link between propagule pressure and anthropogenic activity to develop a spatially explicit variable of human-mediated invasion pressure, the rFOI. This metric not only represents the intensity of surrounding development, but also reflects the likelihood of dispersal of an invader from single family residences as weighted by distance and age of housing (based on knowledge of how long the species has been used an ornamental by homeowners) unlike a more commonly used measure of housing density, which only reflects the intensity of surrounding development. The rFOI is a better proxy, theoretically, of the impacts of humans on the invasion process as compared to housing density because it explicitly links dispersal potential of the invader with the number, distance, and age of residences. This effort was justified, as my results show that the rFOI significantly better detect the presence of Autumn olive and Chinese privet as compared to housing density. This is the first study to explicitly consider the role of human-mediated invasion pressure on the distribution of exotic species and on the performance of SDMs. As a result, I have shown that there is empirical evidence justifying the theoretical link between human activity and exotic species invasions and how inclusion of variables that link human activity to invasive dispersal can be used to improve the accuracy of spatially explicit predictive models of invasions such as SDMs.

The results of my fourth dissertation chapter in which I investigated hypotheses rooted in urban and invasion ecology to assess the impacts of urbanization to forest invasion resistance further support the role of anthropogenic factors in influencing ecological processes. Until now, the underlying mechanisms explaining the link between susceptibility of forests to invasions and urbanization have not been investigated. Here, I synthesized the findings of different disciplines that showed: 1) the nitrification of forest soils along an urban to rural gradient (Pouyat et al., 1995; Lovett et al., 2000) that road density and increased atmospheric deposition of nitrogen and sulfur are linked, primarily via automobile emissions (Fenn et al, 2003; Redling et al., 2013) that habitat eutrophication resulting from increased atmospheric deposition negatively impacts native species richness (Maskell et al., 2010; Stevens et al., 2010; Clark & Tilman, 2008); and 4) that excess nitrogen makes habitats more invasible according to Davis'(2000) resource availability theory to generate the hypothesis that urbanization (measured via road density) has negative impacts on invasion resistance as a result of increased atmospheric deposition resulting in habitat eutrophication (measured via soil fertility). While the results of the structural equation model did not support a direct effect of road density on invasion resistance, road density is shown to have statistically significant indirect negative effects on invasion resistance via its positive effects on soil fertility. These results support my hypothesis that urbanization negatively effects invasion resistance via habitat eutrophication, which has been shown in other studies to result in biodiversity loss and increased susceptibility to invasions. Aspects of roads have been previously shown to be positively associated with exotic species presence, but this relationship has been attributed to the role of roads as dispersal corridors for exotic species (Christen & Matlack, 2005; Gelbard & Belnap, 2003) While this mechanism is likely for weedy species that either grow alongside roads or possess tiny seeds that are easily transported via vehicular traffic (Von der Lippe & Kowarik, 2007), it is less likely for species such as slower growing/maturing shrubs that are largely not found along road sides and are mowed before they can set seed. Flory & Clay (2006) found a significant positive correlation within 10 m of forest interiors and invasive shrub species but this relationship can easily be explained by the increased availability of light, as these sites would also be close to forest edges. However this study does not address the role of light availability or other edge effects in influencing invasive shrub distributions. Therefore, caution is warranted when investigating the relationship of metrics of urbanization to ecological processes such as invasion, to ensure that the question at hand is testing a specific hypothesized ecological relationship underlying observed spatial phenomena, otherwise the informative power of the study is limited to just describing spatial relationships (McDonnell & Hahs, 2008).

In a similar vein, my dissertation further investigated the link between housing and exotic species invasions (Gavier-Pizzaro et al., 2010). My results indicate a statistically significant negative effect of housing age on invasion resistance, suggesting that locations surrounded by older housing have been receiving an influx of exotic propagules arriving from residential yards and gardens longer than those that are surrounded by younger housing. The temporal component represented by housing age also may reflect the duration of anthropogenic disturbances other than escaping exotic species originating from residential development. However, it cannot be distinguished whether housing age is a proxy of propagule pressure or anthropogenic disturbance, or both. Above, I discussed how human-mediated invasion pressure measured as the residential force of invasion is a better predictor of invasive shrub presence than single family housing density. Up until now, the underlying mechanisms explaining the observed correlation between urbanization and exotic species richness have not been investigated, and my results address this gap. Specifically, this study is the first to investigate the potential direct and indirect effects of urbanization on invasion resistance.

In this dissertation, I investigated a suite of metrics reflecting both landscape structure and landscape configuration. Several of these metrics such as distance to forest edge, patch size, and proximity have been linked to the distribution of exotic species (Deckers et al., 2005; Bartuszevige et al., 2006; Hutchinson & Vankat, 1998), but none of these metrics were significant predictors in my research with regard to either exotic species distributions or forest invasion resistance once spatial dependence was accounted for. This suggests that the results of studies that investigate the effects of landscape scale factors on invasibility without controlling for spatial autocorrelation potentially not reliable and may suffer from a Type I error. This is not to say that landscape factors have not been found to be important in this research. To the contrary, one of the strongest predictors of invasion resistance is the area of forest persisting since at least 1938 surrounding the sampled communities.

I have also provided a methodological framework to derive and assess the impacts of typologies of multi- temporal forest cover change trajectories on patterns of native biodiversity. My results demonstrate that the impacts of a typology of forest cover change trajectories on native plant diversity are significant at only the 100 and 200 meter spatial extents, suggesting that diversity is more sensitive to local scale changes in forest cover. In addition, my results demonstrate that studies employing measures of forest cover change using one large time interval, as is commonly done, may fail to identify the underlying relationship between native biodiversity and forest change that would otherwise be exposed if multi-temporal trajectories of forest cover change based on smaller time intervals were used instead. The methodology described in this work can be used to analyze the effects of land cover change assessed over multiple cross-sectional datasets on biodiversity or on other metrics of ecosystem function and structure in order to understand the cumulative impacts of past anthropogenic induced landscape change and inform decision making with regard to biodiversity management.

In conclusion, my doctoral training in geographic information science and urban-regional analysis has allowed me to apply a unique perspective to problem of biological invasions and biodiversity loss in urban landscapes. I have coupled ecological theory with geographical methods to further our understanding of how ecological processes such as the dispersal of invaders are linked to spatial dependence and can be used to dramatically improve the accuracy of species distribution models for generalist invaders. In addition, I have shown how anthropogenic factors such as human-mediated invasion pressure, explains the distribution of invasive species in areas of low environmental suitability in urban forests and thus should not continue to be overlooked in invasion ecology. Using this approach, I also used structural equation modeling to test hypothesized mechanisms of how metrics of urbanization interact with other ecosystem processes (eutrophication and atmospheric deposition) to make urban forest communities less resistant to invasions. As a result, I have produced much needed empirical data regarding the impacts of anthropogenic factors on invasive species spread in forests, the scale at which landscape change impacts patterns of native plant diversity in forests, and how urbanization lowers invasion resistance in forest communities. I show that careful consideration of how anthropogenic factors may influence specific ecological processes can both advance our knowledge of invasion and urban ecology, as well as result in better predictive models that can be used to manage species invasions and protect native biodiversity in urban forests.

BIBLIOGRAPHY

Agarwal, C., Green, G. M., Grove, J.M., Evans, T. P., & Schweik, C. M. (2002). A review and assessment of land-use change models: dynamics of space, time, and human choice. Gen. Tech. Rep. NE-297. Newton Square, PA: U.S. Department of Agriculture, Forest Service, Northeastern Research Station. 61 p.

Aguilar-Santelises, R., & del Castillo, R. F. (2013). Factors affecting woody plant species diversity of fragmented seasonally dry oak forests in the Mixteca Alta, Oaxaca, Mexico. Revista Mexicana de Biodiversidad, 84(2), 575-590.

Alberti, M., Booth, D., Hill, K., Coburn, B., Avolio, C., Coe, S., & Spirandelli, D. (2007). The impact of urban patterns on aquatic ecosystems: an empirical analysis in Puget lowland sub-basins. Landscape and Urban Planning, 80(4), 345-361.

Amaral, P. V., Anselin, L., & Arribas-Bel, D. (2013). Testing for spatial error dependence in probit models. Letters in Spatial and Resource Sciences, 6, 91-101.

Anselin, L. (2002). Under the hood issues in the specification and interpretation of spatial regression models. Agricultural Economics, 27, 247-267.

Anselin, L., & Rey, S. (1991). Properties of tests for spatial dependence in linear regression models. Geographical Analysis, 23(2), 112-131.

Arroyo-Rodríguez, V., Rös, M., Escobar, F., Melo, F. P., Santos, B. A., Tabarelli, M., & Chazdon, R. (2013). Plant β-diversity in fragmented rain forests: testing floristic homogenization and differentiation hypotheses. Journal of Ecology, 101(6), 1449-1458.

Augustin, N. H., Mugglestone, M. A., & Buckland, S. T. (1996). An autologistic model for the spatial distribution of wildlife. Journal of Applied Ecology, 33, 339-347.

Austin M. (2007). Species distribution models and ecological theory: A critical assessment and some possible new approaches. Ecological Modelling, 200, 1-19.

Baayen, R. H. (2011). languageR: Data sets and functions with" Analyzing Linguistic Data: A practical introduction to statistics". 2011. R package version, 1.

Bagchi, S., Briske, D. D., Bestelmeyer, B. T., Ben Wu, X. (2013), Assessing resilience and state-transition models with historical records of cheatgrass *Bromus tectorum* invasion in North American sagebrush-steppe. Journal of Applied Ecology, 50, 1131–1141.

Bahn, V., Krohn, W. B., & O’Connor, R. J. (2008). Dispersal leads to spatial autocorrelation in species distributions: a simulation model. Ecological Modelling, 213, 285-292.

Bartuszevige, A. M., & Gorchov, D. L. (2006). Avian seed dispersal of an invasive shrub. Biological Invasions, 8, 1013-1022.

Bates, D. M. (2010). lme4: Mixed-effects modeling with R. URL: //lme4.r-forge.r-project.org/book/

Bates, D., Maechler, M., & Bolker, B. (2013). lme4: Linear mixed-effects models using S4 classes, R package, version 0.999999-0, 2012. URL://cran.r-project.org/web/packages/lme4/index.html

Beers, T., Dress, P. & Wensel, L. (1966). Aspect transformation in site productivity research. Journal of Forestry, 64, 691-692.

Bellemare, J., Motzkin, G., & Foster, D.R. (2002). Legacies of the agricultural past in the forested present: an assessment of historical land-use effects on rich mesic forests. Journal of Biogeography, 29,1401-1420.

Beron, K. J., & Vijverberg, W. P. (2004). Probit in a spatial context: a Monte Carlo analysis. In Advances in spatial econometrics (pp. 169-195). Springer Berlin Heidelberg.

Betts, M. G., Ganio, L. M., Huso, M. M., Som, N. A., Huettmann, F., Bowman, J., & Wintle, B. A. (2009). Comment on "Methods to account for spatial autocorrelation in the analysis of species distributional data: a review". Ecography, 32, 374-37.

Billé, A. G. (2013). Computational issues in the estimation of the spatial probit model: A comparison of various estimators. Review of Regional Studies, 43, 131-154.

Bivand, R. (2013). spdep: Spatial dependence: weighting schemes, statistics and models. R package version 0.5-68.URL://CRAN.R-project.org/package=spdep.

Bobbink, R., Hicks, K., Galloway, J., Spranger, T., Alkemade, R., Ashmore, M., Bustamante, M., Cinderby, S., Davidson, E., Dentener, F., Emmett, B., Erisman, J-W., Fenn, M., Gilliam, F., Nordin, A., Pardo, L., & De Vries, W. (2010). Global assessment of nitrogen deposition effects on terrestrial plant diversity: a synthesis. Ecological applications, 20(1), 30-59.

Bollen, K. A. (1989). Structural equations with latent variables. John Wiley & Sons. Oxford, England.

Brown, R. L., & Peet, R. K. (2003). Diversity and invasibility of southern Appalachian plant communities. Ecology, 84, 32-39.

Burke, M. J., & Grime, J. P. (1996). An experimental study of plant community invasibility. Ecology, 77(3), 776-790.

Burton, M. L., Samuelson, L. J., & Pan, S. (2005). Riparian woody plant diversity and forest structure along an urban-rural gradient. Urban Ecosystems, 8(1), 93-106.

Caliński, T., & Harabasz, J. (1974). A dendrite method for cluster analysis. Communications in Statistics-theory and Methods, 3(1), 1-27.

Christen, D., & Matlack, G. (2006). The role of roadsides in plant invasions: a demographic approach. Conservation Biology, 20(2), 385-391.

Christensen, N. L., & Peet, R. K. 1984. Convergence during secondary forest succession. Journal of Ecology 72 (1), 25-36.

Clark, C. M., & Tilman, D. (2008). Loss of plant species after chronic low-level nitrogen deposition to prairie grasslands. Nature, 451(7179), 712-715.

Cliff, A., & Ord, K. (1972). Testing for spatial autocorrelation among regression residuals. Geographical Analysis, 4(3), 267-284.

Cole, P. G., & Weltzin, J. F. (2005). Light limitation creates patchy distribution of an invasive grass in eastern deciduous forests. Biological Invasions, 7(3), 477-488.

Connell, J. H. (1978). Diversity in tropical rainforests and coral reefs. Science, 199, 1302-1310.

Consiglio, T., Schatz, G. E., Mcpherson, G., Lowry, P. P., Rabenantoandro, J., Rogers, Z. S., Rabevohitra, R., and Rabehevitra, D. (2006). Deforestation and plant diversity of Madagascar's littoral forests. Conservation Biology, 20(6), 1799-1803.

Crawley, Michael, J. (2005). Statistics: An introduction using R. John Wiley and Sons.

Currie, D. J. (1991). Energy and large-scale patterns of animal-and plant-species richness. American Naturalist, 137(1), 27-49.

Dark, S. J. (2004). The biogeography of invasive alien plants in California: an application of GIS and spatial regression analysis. Diversity and Distributions, 10(1), 1-9.

Davies, D. L., & Bouldin, D. W. (1979). A Cluster Separation Measure. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1(2), 224-227.

Davis, M. A., Grime, J. P. & Thompson, K. (2000). Fluctuating resources in plant communities: a general theory of invasibility. Journal of Ecology, 88, 528–534.

de Albuquerque, F. S., & Rueda, M. (2010). Forest loss and fragmentation effects on woody plant species richness in Great Britain. Forest Ecology and Management, 260(4), 472-479.

Deckers, B., Verheyen, K., Hermy, M., & Muys, B. (2005). Effects of landscape structure on the invasive spread of black cherry Prunus serotina in an agricultural landscape in Flanders, Belgium. Ecography, 28(1), 99-109.

De Espindola, G. M., De Aguiar, A. P. D., Pebesma, E., Câmara, G., & Fonseca, L. (2012). Agricultural land use dynamics in the Brazilian Amazon based on remote sensing and census data. Applied Geography, 32(2), 240-252.

DeGasperis, B. G., & Motzkin, G. (2007). Windows of opportunity: historical and ecological controls on Berberis thunbergii invasions. Ecology, 88(12), 3115-3125.

De Knegt, H. J., Van Langevelde, F., Coughenour, M. B., Skidmore, A. K., De Boer, W. F., Heitkönig, I. M. A., Knox, N. M., Slotow, R., Van der Waal, C., &, Prins, H. H. T. (2010). Spatial autocorrelation and the scaling of species-environment relationships. Ecology, 91, 2455-2465.

de Lima, R. F., Dallimer, M., Atkinson, P. W., & Barlow, J. (2013). Biodiversity and land-use change: understanding the complex responses of an endemic-rich bird assemblage. Diversity and Distributions, 19(4), 411-422.

De Schrijver, A., De Frenne, P., Ampoorter, E., Van Nevel, L., Demey, A., Wuyts, K., & Verheyen, K. (2011). Cumulative nitrogen input drives species loss in terrestrial ecosystems. Global Ecology and Biogeography, 20(6), 803-816.

Díaz, S., Fargione, J., Chapin III, F. S., & Tilman, D. (2006). Biodiversity loss threatens human well-being. PLoS biology, 4(8), e277.

Didham, R. K., Tylianakis, J. M., Gemmell, N. J., Rand, T. A., & Ewers, R. M. (2007). Interactive effects of habitat modification and species invasion on native species decline. Trends in Ecology & Evolution, 22(9), 489-496.

Dirr, M. A. (1998). Manual of Woody Landscape Plants: Their Identification, Ornamental Characteristics, Culture, Propagation and Uses, Fifth Edition. Stipes Publishing Co., Champaign, IL. pp. 338–341.

D'Orangeville, L, Bouchard, A., & Cogliastro, A. (2008). Post-agricultural forests: Landscape patterns add to stand-scale factors in causing insufficient hardwood regeneration. Forest Ecology and Management, 255, 1637-1646.

Dormann, C. F. (2007). Effects of incorporating spatial autocorrelation into the analysis of species distribution data. Global Ecology and Biogeography, 16, 129-138.

Dormann, C., M McPherson, J., B Araújo, M., Bivand, R., Bolliger, J., Carl, G., & Wilson, R. (2007). Methods to account for spatial autocorrelation in the analysis of species distributional data: a review. Ecography, 30(5), 609-628.

Drummond, M. A., & Loveland, T. R. (2010). Land-use pressure and a transition to forest-cover loss in the eastern United States. Bioscience, 60(4), 286-298.

Dubayah, R. C. (1994). Modeling a solar radiation topoclimatology for the Rio Grande River Basin. Journal of Vegetation Science, 5(5), 627-640.

Duguay, S., Eigenbrod, F., & Fahrig, L. (2007). Effects of surrounding urbanization on non-native flora in small forest patches. Landscape Ecology, 22(4), 589-599.

Dukes, J. S., & Mooney, H. A. (1999). Does global change increase the success of biological invaders? Trends in Ecology & Evolution, 14(4), 135-139.
Elton, C.S. (1958). The ecology of invasions by animals and plants. Methuen, London.

Ens, E-J., & French, K. (2008). Exotic woody invader limits the recruitment of three indigenous plant species. Biological Conservation, 141, 590-595.

Eschtruth, A. K., & Battles, J. J. (2011). The importance of quantifying propagule pressure to understand invasion: an examination of riparian forest invasibility. Ecology, 92, 1314-1322.

Escobedo, F. J., Kroeger, T., & Wagner, J. E. (2011). Urban forests and pollution mitigation: analyzing ecosystem services and disservices. Environmental Pollution, 159, 2078-2087.

Evangelista, P. H., Kumar, S., Stohlgren, T. J., Jarnevich, C. S., Crall, A. W., Norman III, J. B. & Barnett, D. T. (2008). Modelling invasion for a habitat generalist and a specialist plant species. Diversity and Distributions, 14, 808–817.

Ewing, R. H. (2008). Characteristics, causes, and effects of sprawl: A literature review. In Urban Ecology: An International Perspective on the Interaction Between Humans and Nature (pp. 519-535). Springer New York .

Fahrig, L. (2003). Effects of habitat fragmentation on biodiversity. Annual Review of Ecology, Evolution, and Systematics, 34, 487-515.

Feng, H., Zhao, X., Chen, F., & Wu, L. (2014). Using land use change trajectories to quantify the effects of urbanization on urban heat island. Advances in Space Research, 53(3), 463-473.

Fenn, M. E., Haeuber, R., Tonnesen, G. S., Baron, J. S., Grossman-Clarke, S., Hope, D., Jaffe, D. A., Copeland, S., Geiser, L., Rueth, H. M., & Sickman, J. O. (2003). Nitrogen emissions, deposition, and monitoring in the western United States. BioScience, 53(4), 391-403.

Fischer, J., & Lindenmayer, D. B. (2007). Landscape modification and habitat fragmentation: a synthesis. Global Ecology and Biogeography, 16(3), 265-280.

Foley, J. A., DeFries, R., Asner, G. P., Barford, C., Bonan, G., Carpenter, S. R., Stuart Chapin,F., Coe, M.T., Daily G. C., Gibbs H. K., Helkowski, J. H., Holloway,T., Howard, E. A., Kucharik, C. J., Monfreda, C., Patz, J. A., Colin Prentice, I., Ramankutty, N., & Snyder, P. K. (2005). Global consequences of land use. Science, 309(5734), 570-574.

Folke, C., Carpenter, S., Walker, B., Scheffer, M., Elmqvist, T., Gunderson, L., & Holling, C. S. (2004). Regime shifts, resilience, and biodiversity in ecosystem management. Annual Review of Ecology, Evolution, and Systematics, 557-581.

Forman, R. T. (1995). Some general principles of landscape and regional ecology. Landscape Ecology, 10(3), 133-142.

Gallien, L., Münkemüller, T., Albert, C. H., Boulangeat, I., & Thuiller, W. (2010). Predicting potential distributions of invasive species: where to go from here? Diversity and Distributions, 16, 331–342.

Gardner, R. H., & Engelhardt, K. A. (2008). Spatial processes that maintain biodiversity in plant communities. Perspectives in Plant Ecology, Evolution and Systematics, 9(3), 211-228.

Gavier Pizarro, G. I., Stewart,S. I., Huebner ,C., Keuler, N. S., & Radeloff, V. C. (2010). Housing is positively associated with invasive exotic plant species richness in New England, USA. Ecological Applications, 20, 1913–1925.

Gelbard, J. L., & Belnap, J. (2003). Roads as conduits for exotic plant invasions in a semiarid landscape. Conservation Biology, 17(2), 420-432

Genolini, C., & Falissard, B.(2010). Kml: k-means for longitudinal data. Computational Statistics, 25(2), 317–328.

Gibson, Luke, Lynam,A .J., Bradshaw, C. J. A., He, F., Bickford, D. P., Woodruff, D. S., Bumrungsri, S., & Laurance,W. F. (2013). Near-complete extinction of native small mammal fauna 25 years after forest fragmentation. Science, 341(6153), 1508-1510.

Gosper, C. R., Stansbury, C. D., & Vivian-Smith, G. (2005). Seed dispersal of fleshy-fruited invasive plants by birds: contributing factors and management options. Diversity and Distributions, 11, 549–558.

Gourieroux, C., Monfort, A., Renault, E., & Trognon, A. (1987). Generalised residuals. Journal of Econometrics, 34, 5-32.

Grace, J. B., & Bollen, K. A. (2005). Interpreting the results from multiple regression and structural equation models. Bulletin of the Ecological Society of America, 86(4), 283-295.

Grace, J. B., Anderson, T. M., Olff, H., & Scheiner, S. M. (2010). On the specification of structural equation models for ecological systems. Ecological Monographs, 80(1), 67-87.

Grimm, N. B., Faeth, S. H., Golubiewski, N. E., Redman, C. L., Wu, J., Bai, X., & Briggs, J. M. (2008). Global change and the ecology of cities. Science, 319(5864), 756-760.

Guan, D., Li, H., Inohae, T., Su, W., Nagaie, T., & Hokao, K. (2011). Modeling urban land use change by the integration of cellular automaton and Markov model. Ecological Modelling, 222(20), 3761-3772.

Guisan, A., & Thuiller, W. 2005. Predicting species distribution: offering more than simple habitat models. Ecology Letters, 8, 993–1009.

Guns, M., & Vanacker, V. (2013). Forest cover change trajectories and their impact on landslide occurrence in the tropical Andes. Environmental Earth Sciences, 70(7), 2941-2952.

Guo, Q., & Symstad, A. (2008). A two-part measure of degree of invasion for cross-community comparisons. Conservation Biology, 22(3), 666-672.

Haddad, N. M., Crutsinger, G. M., Gross, K., Haarstad, J., Knops, M. H., & Tilman, D. (2009). Plant species loss decreases arthropod diversity and shifts trophic structure. Ecology Letters 12, 1029-1039.

Hahn, E. D., & Soyer, R. (2005). Probit and logit models: Differences in the multivariate realm. Working Paper, George Washington University.

Hahs, A. K., McDonnell, M. J., McCarthy, M. A., Vesk, P. A., Corlett, R. T., Norton, B. A., Clemants, S. E., Duncan, R. P., Thompson, K., Schwartz, M. W., & Williams, N. S.G. (2009). A global synthesis of plant extinction rates in urban areas. Ecology Letters, 12(11), 1165-1173.

Hanski, I. (1998). Metapopulation dynamics. Nature, 396(6706), 41-49.

Hanula, J. L., S. Horn, and Taylor, J.W. (2009). Chinese privet (*Ligustrum sinense*) removal and its effects on native plant communities of riparian forests. Invasive Plant Science and Management, 2, 292–300.

Hart, J. L., & Holmes, B. N. (2013). Relationships between *Ligustrum sinense* invasion, biodiversity, and development in a mixed bottomland forest. Invasive Plant Science and Management, 6, 175-186

Hart, John Fraser. (1980). Land Use Change in a Piedmont County. Annals of the Association of American Geographers, 70(4), 492-527.

Havel, J. E., Shurin, J. B., & Jones, J. R. (2002). Estimating dispersal from patterns of spread: spatial and local control of lake invasions. Ecology, 83, 3306-3318.

Hickman, J. E., & Lerdau, M. T. (2013). Biogeochemical impacts of the northward expansion of kudzu under climate change: the importance of ecological context. Ecosphere, 4 (10), article 121. Retrieved from WWW: esajournals.org/doi/full/10.1890/ES13-00142.1.

Hooper, D., Coughlan, J., & Mullen, M. R. (2008). Structural Equation Modelling: Guidelines for Determining Model Fit. Electronic Journal of Business Research Methods, 6(1), 53-60.

Hooper, D.U., Chapin, F.S., Ewel, J.J., Hector, A., Inchausti, P., Lavorel, S., Lawton, J.H., Lodge, D.M., Loreau, M., Naeem, S., Schmid, B., Seta, H., Symstad, J., Vandermeer, J., & Wardle, D.A. (2005). Effects of biodiversity on ecosystem functioning: A consensus of current knowledge. Ecological Monographs, 75(1), 3–35.

Hope, D., Gries, C., Zhu, W., Fagan, W. F., Redman, C. L., Grimm, N. B., Nelson, A. L, Martin, C., & Kinzig, A. (2003). Socioeconomics drive urban plant diversity. Proceedings of the National Academy of Sciences, 100, 8788-8792.

Hosmer Jr, D. W., Lemeshow, S., & Sturdivant, R. X. (2013). Applied logistic regression. John Wiley & Sons. Hoboken.

Hox, J., & Maas, C. (2004). Multilevel structural equation models: The limited information approach and the multivariate multilevel approach. In K.van Montfort, J. Oud & A. Satorra (Eds.), Recent Developments on Structural Equation Models (pp. 135-149). Springer, Dordrecht, Netherlands.

Huang, B., Zhang, L., & Wu, B. (2009). Spatiotemporal analysis of rural–urban land conversion. International Journal of Geographical Information Science, 23(3), 379-398.

Hutchinson, T. F., & Vankat, J. L. (1998). Landscape structure and spread of the exotic shrub *Lonicera maackii* (Amur honeysuckle) in southwestern Ohio forests. The American Midland Naturalist, 139(2), 383-390.

Hylander, K., & Ehrlén, J. (2013). The mechanisms causing extinction debts. Trends in Ecology & Evolution, 28(6), 341-346.

Jiménez-Valverde, A., Peterson, A. T., Soberón, J., Overton, J. M., Aragón, P., & Lobo, J. M. (2011). Use of niche models in invasive species risk assessments. Biological Invasions, 13, 2785-2797.

Kelejian, H. H., & Prucha, I. R. (2001). On the asymptotic distribution of the Moran I test statistic with applications. Journal of Econometrics, 104, 219-257.

Kennedy, T. A., Naeem, S., Howe, K. M., Knops, J. M., Tilman, D., & Reich, P. (2002). Biodiversity as a barrier to ecological invasion. Nature, 417(6889), 636-638.

Kidane, Y., Stahlmann, R., & Beierkuhnlein, C. (2012). Vegetation dynamics, and land use and land cover change in the Bale Mountains, Ethiopia. Environmental Monitoring and Assessment, 184(12), 7473-7489.

Klepeis, P., Scull, P., LaLonde, T., Svajlenka, N., & Gill, N. (2013). Changing forest recovery dynamics in the northeastern United States. Area, 45(2), 239-248.

Kowarik, I. (2003). Human agency in biological invasions: secondary releases foster naturalisation and population expansion of alien plant species. Biological Invasions, 5, 293-312.

Kowarik, I. (2011). Novel urban ecosystems, biodiversity, and conservation. Environmental Pollution, 159(8), 1974-1983.

Krauss, J., Bommarco, R., Guardiola, M., Heikkinen, R. K., Helm, A., Kuussaari, M., Lindborg, R., Öckinger, E., Pärtel, M., Pino, J., Pöyry, J., Raatikainen, K. M., Sang, A., Stefanescu, C., Teder, T., Zobel, M., & Steffan-Dewenter, I. (2010), Habitat fragmentation causes immediate and time-delayed biodiversity loss at different trophic levels. Ecology Letters, 13(5), 597–605.

Krivanke, M., Pysek, P., & Jarosik, V. (2006). Planting history and propagule pressure as predictors of invasion by woody species in a temperate region. Conservation Biology, 20, 1487-1498.

Kulik, I., Hornsby Stewart, K., & Bishop, I.D. (2011). Modeling geospatial trend changes in vegetation monitoring data. Computers, Environment and Urban Systems 35: 45-56.

Kumar, S., Stohlgren, T. J., & Chong, G. W. (2006). Spatial heterogeneity influences native and nonnative plant species richness. Ecology, 87(12), 3186-3199.

Lambin, E. F. (1997). Modelling and monitoring land-cover change processes in tropical regions. Progress in Physical Geography, 21(3), 375-393.

Lambin, E. F., Turner, B. L., Geist, H. J., Agbola, S. B., Angelsen, A., Bruce, J. W., Coomes, O.T., Dirzo R., Fischer G., Folke C., George, P.S., Homewood K., Imbernon J., Leemans R., Li X., Moran E.F., Mortimore, M., Ramakrishnan P.S., Richards, J.F., Skånes, H., Steffen, W., Stone, G.D., Svedin, U., Veldkamp, T.A., Vogel, C., and Xu, J. (2001). The causes of land-use and land-cover change: moving beyond the myths. Global Environmental Change, 11(4), 261-269.

Langran, G., & Chrisman, N. R. (1988). A framework for temporal geographic information. Cartographica: The International Journal for Geographic Information and Geovisualization, 25(3), 1-14.

Legendre, P. (1993). Spatial autocorrelation: trouble or new paradigm? Ecology, 74, 1659-1673.

Lemke, D., Coulston, J. W., Hulme, P., Paterson, C., & Brown, J. A. (2012). Invasive potential of invasive plants in the forest of the southern region, United States. In: Morin, Randall S.; Liknes, Greg C., comps. Moving from status to trends: Forest Inventory and Analysis (FIA) symposium 2012; 2012 December 4-6; Baltimore, MD. Gen. Tech. Rep. NRS.

LeSage, J., & Pace, R. K. (2009). Introduction to spatial econometrics. CRC Press. Boca Raton.

Lobo, J.M., Jime´nez-Valverde,A., & Hortal, J. (2010). The uncertain nature of absences and their importance in species distribution modeling. Ecography, 33, 103-114

Lockwood, J. L., Cassey, P., & Blackburn, T. M. (2009). The more you introduce the more you get: the role of colonization pressure and propagule pressure in invasion ecology. Diversity and Distributions, 15, 904-910.

Lockwood, J. L., Cassey, P., & Blackburn, T. (2005). The role of propagule pressure in explaining species invasions. Trends in Ecology & Evolution, 20(5), 223-228.

Lovett, G. M., Traynor, M. M., Pouyat, R. V., Carreiro, M. M., Zhu, W. X., & Baxter, J. W. (2000). Atmospheric deposition to oak forests along an urban-rural gradient. Environmental Science & Technology, 34(20), 4294-4300.

Leathwick, J. R., Burns, B. R., & Clarkson, B. D. (1998). Environmental correlates of tree alpha-diversity in New Zealand primary forests. Ecography, 21(3), 235-246.

Lele, N., & Joshi, P. K. (2009). Analyzing deforestation rates, spatial forest cover changes and identifying critical areas of forest cover changes in North-East India during 1972–1999. Environmental Monitoring and Assessment, 156(1-4), 159-170.

Lung, T., Peters, M. K., Farwig, N., Böhning-Gaese, K., & Schaab, G. (2012). Combining long-term land cover time series and field observations for spatially explicit predictions on changes in tropical forest biodiversity. International Journal of Remote Sensing, 33(1), 13-40.

Mace, G. M., Norris, K., & Fitter, A. H. (2012). Biodiversity and ecosystem services: a multilayered relationship. Trends in ecology & evolution, 27(1), 19-26.

Mallik, Azim U. (2013). Predicting Plant Diversity Response to Disturbance: Applicability of the Intermediate Disturbance Hypothesis and Mass Ratio Hypothesis. Critical Reviews in Plant Sciences 32(6), 383-395.

Manel, S., Williams, H. C., & Ormerod, S. J. (2001). Evaluating presence–absence models in ecology: the need to account for prevalence. Journal of Applied Ecology, 38, 921-931.

Martin, P. H., Canham, C. D., & Marks, P. L. (2008). Why forests appear resistant to exotic plant invasions: intentional introductions, stand dynamics, and the role of shade tolerance. Frontiers in Ecology and the Environment, 7(3), 142-149.

Maskell, L. C., Smart, S. M., Bullock, J. M., Thompson, K., & Stevens, C. J. (2010). Nitrogen deposition causes widespread loss of species richness in British habitats. Global Change Biology, 16(2), 671-679.

Mather, A. S., & Needle, C. L. (1998). The forest transition: a theoretical basis. Area, 30(2), 117-124.

McDonnell, M. J., & Hahs, A. K. (2008). The use of gradient analysis studies in advancing our understanding of the ecology of urbanizing landscapes: current status and future directions. Landscape Ecology, 23(10), 1143-1155.

McGarigal, K., S. A. Cushman, & Ene, E. (2012). FRAGSTATS v4: Spatial Pattern Analysis Program for Categorical and Continuous Maps. Computer software program produced by the authors at the University of Massachusetts, Amherst. WWW:umass.edu/landeco/research/fragstats/fragstats.html.

McKinney, M. L. (2008). Effects of urbanization on species richness: a review of plants and animals. Urban Ecosystems, 11(2), 161-176.

Meentemeyer, R. K., Anacker, B. L., Mark, W., & Rizzo, D. M. (2008). Early detection of emerging forest disease using dispersal estimation and ecological niche modeling. Ecological Applications, 18, 377-390.

Meentemeyer, R.K., Tang, W.W., Dorning, M.A., Vogler, J.B., Cunniffe, N.J., & Shoemaker, D.A. (2013). FUTURES: Multilevel Simulations of Emerging Urban-Rural Landscape Structure Using a Stochastic Patch-Growing Algorithm. Annals of the Association of American Geographers, 103, 785-807

Merriam, R. W., & Feil, E. (2002). The potential impact of an introduced shrub on native plant diversity and forest regeneration. Biological Invasions, 4, 369-373.

Meyfroidt, P., Rudel, T. K., & Lambin, E. F. (2010). Forest transitions, trade, and the global displacement of land use. Proceedings of the National Academy of Sciences, 107(49), 20917-20922.

Miller, J. A., & Franklin, J. (2010). Incorporating Spatial Autocorrelation in Species Distribution Models. In Handbook of Applied Spatial Analysis (pp. 685-702). Springer Berlin Heidelberg.

Moore, I. D., R. B. Grayson, & A. R. Landson. (1991). Digital Terrain Modelling: A Review of Hydrological, Geomorphological, and Biological Applications. Hydrological Processes, 5, 3–30.

Mosher, E. S., Silander Jr, J. A., & Latimer, A. M. (2009). The role of land-use history in major invasions by woody plant species in the northeastern North American landscape. Biological Invasions, 11(10), 2317-2328.

Mueller, R. O. (1996). Linear Regression and Classical Path Analysis. In Basic Principles of Structural Equation Modeling (pp. 1-61). Springer New York.

Muller, M. R., & Middleton, J. (1994). A Markov model of land-use change dynamics in the Niagara Region, Ontario, Canada. Landscape Ecology, 9(2), 151-157.

Nagendra, H., Paul, S., Pareeth, S., & Dutt, S. (2009). Landscapes of protection: forest change and fragmentation in Northern West Bengal, India. Environmental Management, 44(5), 853-864.

Nakagawa, S., & Schielzeth, H. (2013). A general and simple method for obtaining R2 from generalized linear mixed-effects models. Methods in Ecology and Evolution, 4(2), 133-142.

Nowak, D. J., & Dwyer, J. F. (2007). Understanding the benefits and costs of urban forest ecosystems. In Urban and Community Forestry in the Northeast (pp. 25-46). Springer, Dordrecht.

Pataki, D. E., Carreiro, M. M., Cherrier, J., Grulke, N. E., Jennings, V., Pincetl, S., Pouyat, R.V, Whitlow T.H., & Zipperer, W. C. (2011). Coupling biogeochemical cycles in urban environments: ecosystem services, green solutions, and misconceptions. Frontiers in Ecology and the Environment, 9(1), 27-36.

Parepa, M., Fischer, M., & Bossdorf, O. (2013). Environmental variability promotes plant invasion. Nature Communications, 4, 1604.

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

Peet, R. K., & Christensen, N. L. (1980). Hardwood forest vegetation of the North Carolina Piedmont. Veroffentlichungen des Geobotanischen Institutes der ETH, Stiftung Rubel, (69), 14-39.

Peterson, A. T. (2003). Predicting the geography of species' invasions via ecological niche modeling. The Quarterly Review of Biology, 78, 419-433.

Pimentel,D., Stachow, U., Takacs, D. A., Brubaker, H. W., Dumas, A. R., Meaney, J. J., O'Neil, J. A. S., Onsi, D. E., & Corzilius, D. B. (1992). Conserving biological diversity in agricultural/forestry systems. BioScience, 42(5), 354-362.

Pinheiro, J. C., & Bates, D. M. (2000). Mixed-effects Models in S and S-PLUS. Springer-Verlag, New York.

Pouyat, R. V., McDonnell, M. J., & Pickett, S. T. A. (1995). Soil characteristics of oak stands along an urban-rural land-use gradient. Journal of Environmental Quality, 24(3), 516-526.

Pyšek, P., Jarošík, V., Hulme, P. E., Kühn, I., Wild, J., Arianoutsou, M., Bacherf, S., Chirong, F., Didžiulish, V., Essli, F., Genovesij, P., Gherardik, F., Hejdaa,M., t Karkl, S., Lambdonm, P.W., Desprez-Loustaun,M-L., Nentwigo,W., Pergla, J., Poboljšajp, K., Rabitschi, W., Roquesq, A., Royr, D.B., Shirleys, S., Solarzt, W., Vilàu, M., & Winter, M. (2010). Disentangling the role of environmental and human pressures on biological invasions across Europe. Proceedings of the National Academy of Sciences, 107(27), 12157-12162.

R Core Team. (2012). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL://www.R-project.org/.

Radford, A. E., Ahles, H. E., & Bell, C. R. (1968). Manual of the vascular flora of the Carolinas. University of North Carolina Press, Chapel Hill.

Ray, S., & Turi, R. H. (1999). Determination of number of clusters in k-means clustering and application in colour image segmentation. In Proceedings of the 4th international conference on advances in pattern recognition and digital techniques (pp. 137-143).

Redling, K., Elliott, E., Bain, D., & Sherwell, J. (2013). Highway contributions to reactive nitrogen deposition: tracing the fate of vehicular NOx using stable isotopes and plant biomonitors. Biogeochemistry, 116(1-3), 261-274.

Reinhart, K. O., Gurnee, J., Tirado, R., & Callaway, R. M. (2006). Invasion through quantitative effects: intense shade drives native decline and invasive success. Ecological Applications, 16(5), 1821-1831.

Roxburgh, S. H., Shea, K., & Wilson, J. B. (2004). The intermediate disturbance hypothesis: patch dynamics and mechanisms of species coexistence. Ecology, 85(2), 359-371.

Ruxton, G. D. (2006). The unequal variance t-test is an underused alternative to Student's t-test and the Mann–Whitney U test. Behavioral Ecology, 17, 688-690.

Ryu, E. (2014). Model fit evaluation in multilevel structural equation models. Frontiers in psychology, 5.

Sanchez-Cuervo, A. M., & Aide, T. M. (2013). Identifying hotspots of deforestation and reforestation in Colombia (2001-2010): implications for protected areas. Ecosphere, 4(11), article 143, //dx.doi.org/10.1890/ES13-00207.1.

Sargent, R. G. (2013). Verification and validation of simulation models. Journal of Simulation, 7, 12-24.

Schneider, L. C., & Pontius Jr, G. (2001). Modeling land-use change in the Ipswich watershed, Massachusetts, USA. Agriculture, Ecosystems & Environment, 85(1), 83-94.

Seto, K. C., Güneralp, B., & Hutyra, L. R. (2012). Global forecasts of urban expansion to 2030 and direct impacts on biodiversity and carbon pools. Proceedings of the National Academy of Sciences, 109(40), 16083-16088.

Shannon, C. E. (1948). A mathematical theory of communication. Bell Syst. Tech. J, 27, 379-423.

Shea, K., & Chesson, P. (2002). Community ecology theory as a framework for biological invasions. Trends in Ecology & Evolution, 17(4), 170-176.

Shipley, B. (2002). Cause and correlation in biology: a user's guide to path analysis, structural equations and causal inference. Cambridge University Press, Cambridge.

Simberloff, D. (2009). The role of propagule pressure in biological invasions. Annual Review of Ecology, Evolution, and Systematics, 40, 81-102.

Sing, T., Sander, O., Beerenwinkel, N., & Lengauer, T. (2005). ROCR: visualizing classifier performance in R. Bioinformatics, 21, 3940-3941.

Stevens, C. J., Dise, N. B., Mountford, J. O., & Gowing, D. J. (2004). Impact of nitrogen deposition on the species richness of grasslands. Science, 303(5665), 1876-1879.

Stevens, C. J., Duprè, C., Dorland, E., Gaudnik, C., Gowing, D. J., Bleeker, A., Diekmann, M., artin, Alard, D., Bobbink, R., Fowler, D., Corcket, E., Mountford, J.O., Vandvik, V., Aarrestad, P.A., Muller, S., & Dise, N. B. (2010). Nitrogen deposition threatens species richness of grasslands across Europe. Environmental Pollution, 158(9), 2940-2945.

Stohlgren, T. J., Binkley, D., Chong, G. W., Kalkhan, M. A., Schell, L. D., Bull, K. A., Otsuki, Y., Newman, G., Bashkin, M., and Son, Y. (1999). Exotic plant species invade hot spots of native plant diversity. Ecological Monographs, 69(1), 25-46.

Stromayer, K. A. K., Warren, R. J., Johnson, A. S., Hale, P.E., Rogers, C. L., & Tucker, C. L., (1998). Chinese privet and the feeding ecology of white-tailed deer: The role of an exotic plant. Journal of Wildlife Management 62, 1321-1329.

Sutton, M. A., Oenema, O., Erisman, J. W., Leip, A., van Grinsven, H., & Winiwarter, W. (2011). Too much of a good thing. Nature, 472(7342), 159-161

Tanentzap, A. J., & Bazely, D. R. (2009). Propagule pressure and resource availability determine plant community invasibility in a temperate forest understorey. Oikos, 118, 300-308.

Tanentzap, A. J., Bazely, D. R., & Lafortezza, R. (2010). Diversity–invasibility relationships across multiple scales in disturbed forest understoreys. Biological Invasions, 12(7), 2105-2116.

Taverna, K., Urban, D. L., & McDonald, R. I. (2005). Modeling landscape vegetation pattern in response to historic land-use: a hypothesis-driven approach for the North Carolina Piedmont, USA. Landscape Ecology, 20(6), 689-702.

Theobald, D. M. (2005). Landscape Patterns of Exurban Growth in the USA from 1980 to 2020. Ecology and Society, 10(1), 32.WWW: ecologyandsociety.org/vol10/iss1/art32/

Thomson, D. M. (2007). Do source-sink dynamics promote the spread of an invasive grass into a novel habitat? Ecology, 88, 3126-3134.

Tilman, G. D. (1984). Plant dominance along an experimental nutrient gradient. Ecology, 65(5), 1445-1453.

Trimble, S.W. (1974). Man-induced soil erosion on the Southern Piedmont: 1700–1970. Soil Conservation Society of America, Ankeny, IA.

Trisurat, Y., Alkemade, R., & Verburg, P. H. (2010). Projecting land-use change and its consequences for biodiversity in Northern Thailand. Environmental Management, 45(3), 626-639.

Turner, M. G. (2005). Landscape ecology: what is the state of the science? Annual Review of Ecology, Evolution, and Systematics, 36, 319-344.
United Nations, Department of Economic and Social Affairs, Population Division (2014). World Urbanization Prospects: The 2014 Revision.

Václavík, T., & Meentemeyer, R. K. (2009). Invasive species distribution modeling (iSDM): Are absence data and dispersal constraints needed to predict actual distributions? Ecological Modelling, 220, 3248-3258.

Václavík, T., Kupfer, J. A., & Meentemeyer, R. K. (2012). Accounting for multi-scale spatial autocorrelation improves performance of invasive species distribution modelling (iSDM). Journal of Biogeography, 39, 42-55.

Van Teeffelen, A. J., & Ovaskainen, O. (2007). Can the cause of aggregation be inferred from species distributions?. Oikos, 116, 4-16.

Vellend, M., Verheyen, K., Jacquemyn, H., Kolb, A., Van Calster, H., Peterken, G., & Hermy, M. (2006). Extinction debt of forest plants persists for more than a century following habitat fragmentation. Ecology, 87(3), 542-548.

Verburg, P. H., Soepboer, W., Veldkamp, A., Limpiada, R., Espaldon, V., & Mastura, S. S. (2002). Modeling the spatial dynamics of regional land use: the CLUE-S model. Environmental Management, 30(3), 391-405.

Verheyen, K., Guntenspergen, G. R., Biesbrouck, B., & Hermy, M. (2003). An integrated analysis of the effects of past land use on forest herb colonization at the landscape scale. Journal of Ecology, 91(5), 731-742.

Vittoz, P., & Engler, R. (2007). Seed dispersal distances: a typology based on dispersal modes and plant traits. Botanica Helvetica, 117, 109-124.

Von der Lippe, M., & Kowarik, I. (2007). Long distance dispersal of plants by vehicles as a driver of plant invasions. Conservation Biology, 21(4), 986-996.

Von Holle, B., & Motzkin, G. (2007). Historical land use and environmental determinants of nonnative plant distribution in coastal southern New England. Biological Conservation, 136, 33-43.

Von Holle, B., & Simberloff, D. (2005). Ecological resistance to biological invasion overwhelmed by propagule pressure. Ecology, 86(12), 3212-3218.

Wear, D. N., & Greis, J. G. (2002). Southern forest resource assessment-technical report. Asheville, NC: U.S. Department of Agriculture, Forest Service, Southern Research Station.

Welk, E. (2004). Constraints in range predictions of invasive plant species due to non-equilibrium distribution patterns: Purple loosestrife (*Lythrum salicaria*) in North America. Ecological Modelling, 179(4), 551-567.

Williamson, M. (1996). Biological Invasions Chapman & Hall, London.

With, K.A. (2002). The landscape ecology of invasive spread. Conservation Biology, 16 (5), 1192-1203.

Ye, R., & Wright, A. L. (2010). Multivariate analysis of chemical and microbial properties in histosols as influenced by land-use types. Soil and Tillage Research, 110(1), 94-100.

Zeiter, M., & Stampfli, A. (2012). Positive diversity–invasibility relationship in species-rich semi-natural grassland at the neighbourhood scale. Annals of botany, 110(7), 1385-1393.

Zuur, A., Ieno, E. N., Walker, N., Saveliev, A. A., & Smith, G. M. (2009). Mixed effects models and extensions in ecology with R. Springer, New York.

APPENDIX A: PEARSON'S CORRELATION OF PREDICTORS USED IN SPECIES DISTRIBUTION MODELS OF CHINESE PRIVET AND AUTUMN OLIVE

Table 1a. Pearson correlation for the final set of predictors used to predict the distribution of autumn olive.

		AO (0/1)	SI	CC	SFHD 1000	SFHD 1000^2	rFOI 1.5	rFOI 1.5^2	aoFOI 1.0	SWC
AO (0/1)	Pearson Correlation	1	-0.19	0.19	-0.34	0.13	-0.33	0.17	0.53	
	Sig.		0.0011	<0.001	0.0292	0	0.0025	<0.001	<0.001	
SI	Pearson Correlation	-0.19	1	-0.01	0.25	-0.07	0.27	-0.06	-0.12	0.370
	Sig.	0.0012		0	0.2154	<0.001	0.3417	0.04	0.04	<0.001
CC	Pearson Correlation	0.19	-0.01	1	-0.14	0	-0.12	0.01	0.17	0.095
	Sig.	0.0011	0.8455		0.9894	0.033	0.8307	0.0033	0.0033	0.237
SFHD 1000	Pearson Correlation	-0.34	0.25	-0.14	1	0.24	0.89	0.03	-0.42	0.173
	Sig.	<0.001	<0.001	0.0133		<0.001	0.5518	<0.001	<0.001	0.010
SFHD 1000^2	Pearson Correlation	0.13	-0.07	0	0.24	1	0.1	0.62	0.14	0.14
	Sig.	0.0292	0.2154	0.9894	<0.001		<0.001	0.0142	0.0142	0.048
rFOI 1.5	Pearson Correlation	-0.33	0.27	-0.12	0.89	0.1	1	0.05	-0.43	0.167
	Sig.	<0.001	<0.001	0.0331	0	0.0792		<0.001	<0.001	0.013
rFOI 1.5^2	Pearson Correlation	0.17	-0.06	0.01	0.03	0.62	0.05	1	0.2	0.077
	Sig.	0.0025	0.3417	0.8307	0.5518	<0.001	0.3486		0.0004	0.4
aoFOI 1.0	Pearson Correlation	0.53	-0.12	0.17	-0.42	0.14	-0.43	0.2	1	0.070
	Sig.	<0.001	0.04	0.0033	<0.001	0.0142	<0.001	0.0004		0.418
SWC	Pearson Correlation		0.370	0.095	0.173	0.14	0.167	0.077	0.070	1
	Sig.		<0.001	0.237	0.010	0.048	0.013	0.4	0.418	

Table 1b. Pearson correlation for the final set of predictors used to predict the distribution of Chinese privet.

		Priv(0/1)	TMI	RSP	SI	SI^2	rFOI 1.0	rFOI 1.0^2	SFHD 1500m	SFHD 1500m2	pFOI	pFOI*RSP
Priv(0/1)	Pearson Correlation	1	0.21	-0.30	0.12	-0.2	0.19	0.12	0.24	0.10	0.38	0.35
	Sig.		0.00	0.04	0.00	0.00	0.03	0.00	0.08	0.00	0.00	0.00
TMI	Pearson Correlation	0.21	1	-0.55	0.16	-0.27	0.10	-0.13	0.14	-0.03	0.10	-0.10
	Sig.	0.00		0.00	0.00	0.07	0.02	0.01	0.54	0.07	0.07	0.07
RSP	Pearson Correlation	-0.30	-0.55	1.00	0.06	0.02	-0.12	-0.01	-0.11	-0.03	-0.19	0.08
	Sig.	0.00	0.00		0.76	0.03	0.80	0.05	0.61	0.00	0.16	0.16
SI	Pearson Correlation	0.12	0.16	0.06	1	-1	0.12	0.10	0.14	-0.01	0.10	0.08
	Sig.	0.04	0.00	0.30		0.04	0.08	0.02	0.85	0.07	0.13	0.13
SI^2	Pearson Correlation	-0.20	-0.27	0.02	-0.55	1.00	-0.18	0.00	-0.20	0.04	-0.10	-0.06
	Sig.	0.00	0.00	0.76	0.00		0.97	0.00	0.47	0.09	0.28	0.28
rFOI 1.0	Pearson Correlation	0.19	0.10	-0.12	0.12	-0.18	1	0	0.81	-0.12	0.17	0.12
	Sig.	0.00	0.07	0.03	0.04	0.00		0.00	0.04	0.00	0.04	0.04
rFOI 1.0^2	Pearson Correlation	0.12	-0.13	-0.01	0.10	0.00	0.04	1.00	-0.03	0.52	0.02	0.03
	Sig.	0.03	0.02	0.80	0.08	0.97	0.47		0.00	0.78	0.63	0.63
SFHD 1500m	Pearson Correlation	0.24	0.14	-0.11	0.14	-0.20	0.81	-0.03	1	0	0.21	0.21
	Sig.	0.00	0.01	0.05	0.02	0.00	0.00	0.62		0.00	0.00	0.00
SFHD 1500m^2	Pearson Correlation	0.10	-0.03	-0.03	-0.01	0.04	-0.12	0.52	-0.04	1	0.03	0.06
	Sig.	0.08	0.54	0.61	0.85	0.47	0.04	0.00	0.44		0.28	0.28
pFOI 1.5	Pearson Correlation	0.38	0.10	-0.19	0.10	-0.10	0.17	0.02	0.21	0.03	1	0.83
	Sig.	0.00	0.07	0.00	0.07	0.09	0.00	0.78	0.00	0.64		0.00
pFOI*RSP	Pearson Correlation	0.35	-0.10	0.08	0.08	-0.06	0.12	0.03	0.21	0.06	0.83	1
	Sig.	0.00	0.07	0.16	0.13	0.28	0.04	0.63	0.00	0.28	0.00	