2013

Effects of Climate Change on the Habitats of the Invasive Species *Ailanthus altissima* Along the Appalachian Trail

John Clark
University of Rhode Island, jclark817@gmail.com

Follow this and additional works at: <https://digitalcommons.uri.edu/theses>

Recommended Citation

Clark, John, "Effects of Climate Change on the Habitats of the Invasive Species *Ailanthus altissima* Along the Appalachian Trail" (2013). *Open Access Master's Theses*. Paper 79.
<https://digitalcommons.uri.edu/theses/79>

This Thesis is brought to you for free and open access by DigitalCommons@URI. It has been accepted for inclusion in Open Access Master's Theses by an authorized administrator of DigitalCommons@URI. For more information, please contact digitalcommons-group@uri.edu.

**EFFECTS OF CLIMATE CHANGE ON THE
HABITATS OF THE INVASIVE SPECIES *AILANTHUS ALTISSIMA*
ALONG THE APPALACHIAN TRAIL**

**BY
JOHN CLARK**

**A THESIS SUBMITTED IN PARTIAL FULFILMENT OF THE
REQUIREMENTS FOR THE DEGREE OF
MASTER OF SCIENCE
IN
ENVIRONMENTAL SCIENCE**

UNIVERSITY OF RHODE ISLAND

2013

MASTER OF ENVIRONMENTAL SCIENCE THESIS
OF
JOHN CLARK

APPROVED:

Thesis Committee:

Major Professor Yeqiao Wang

Peter V. August

Keith T. Killingbeck

Nasser H. Zawia
DEAN OF THE GRADUATE SCHOOL

UNIVERSITY OF RHODE ISLAND
2013

ABSTRACT

The Appalachian National Scenic Trail (A.T.) is a footpath stretching from Springer Mountain in Georgia to Mount Katahdin in Maine and spanning over 3,500 km of peaks, valleys, and ridges. The A.T.'s gradients in elevation, latitude, and moisture and north-south alignment represent a continental scale cross-section, or "MEGA-Transect," of eastern U.S. forest and alpine areas and offer a setting for collecting scientific data on the health of ecosystems and species that inhabit them.

The Appalachian Trail Decision Support System, or A.T.-DSS, is an Internet-based implementation and dissemination toolset directed at enhancing the decision-making process for managing natural resources. The A.T.-DSS provides a coherent framework for monitoring, reporting, and forecasting ecological conditions by integrating NASA multi-platform sensor data, NASA Terrestrial Observation and Prediction System (TOPS) models, and *in situ* measurements from A.T. MEGA-Transect partners.

The purpose of this research is to develop a prototype habitat suitability model for the invasive species tree-of-heaven (*Ailanthus altissima* (Mill.) Swingle), an exotic tree species pervasive throughout the United States due to its rapid growth, high fecundity, hardy tolerance, and strong competitive ability. This prototype model demonstrates the capabilities of the A.T.-DSS by leveraging seamless geospatial data and climate models from TOPS along with ground based Forest Inventory and Analysis data from the USDA Forest Service to model the current and potential future distributions of suitable *Ailanthus* habitats within the A.T. landscape.

Analysis of the FIA records revealed that *Ailanthus* was most abundant in the Mid-Atlantic States and tended to occur at lower elevations, closer to roadways, and in younger forest stands. Maximum entropy modeling (Maxent) was used to relate the observed distribution of *Ailanthus* to an array of geospatial data layers representing environmental conditions, termed environmental variables. Significant relationships were detected for land cover (developed areas, canopy cover) and topographic (elevation, slope) variables. However, climatic variables were consistently the highest performing predictors, and revealed a preference for warmer and drier regions.

Projected precipitation and temperature data based on scenarios from the Intergovernmental Panel on Climate Change for the period 2095-2099 were substituted for current climate variables to examine potential trends in the distribution of suitable *Ailanthus* habitats. The resulting models indicate that total suitable area will increase from 56% to 82% of the study area. Additionally, the mean elevation of suitable habitats will increase by 59 m and the mean latitude will shift north by 49 km. The predicted changes were most dramatic along the New England section of the A.T.

ACKNOWLEDGEMENTS

First and foremost I'd like to thank my advisor, Dr. Yeqiao Wang. Without his vision, guidance, and wisdom this project would not exist. I'd also like to thank my outstanding committee members: Drs. Pete August, Keith Killingbeck, and Charlie Roman; whose indispensable input shepherded this thesis into its final form.

This work is just one small component of an enormous undertaking to establish a decision support system for the Appalachian Trail, or A.T.-DSS. The project was funded by NASA's Science Mission Directorate through the Applied Science Program (ROSES-2008) under "Decision Support through Earth Science Research Results." Our project brought together an incredibly diverse team of skilled individuals, including Rama Nemani, Forrest Melton, Hiro Hashimoto, and Sam Hiatt of NASA's Terrestrial Observation and Prediction System; Fred Dieffenbach, Brian Mitchell, and Casey Reese of the National Park Service; Glenn Holcombe and Marcia McNiff of the U.S. Geological Survey; Ken Stolte of the U.S. Forest Service; and Matt Robinson of the Appalachian Trail Conservancy. I'm thrilled to have worked with an outstanding group of professionals throughout the course of this ambitious project.

Despite our geographic absence from the trail, the University of Rhode Island was also strongly represented on the A.T.-DSS team. Beyond Dr. Wang and myself, Drs. August and Peter Paton patiently imparted their acumen as we worked to untangle ecological knots. From the Environmental Data Center, Roland Duhaime developed our viewshed monitoring tool, Chris Damon developed our mapping viewers, and Chuck LaBash provided a fertile infrastructure for system development. From outside the department, Fu Luo from the Electric Engineering Department

crafted a website to embed the project within, while Jianjun Zhao joined us from China's Northeast Normal University to investigate phenology. I'd also like to thank my field crew. While the data we collected was not used in this thesis, Dan Evans and Sam Kellog were kingly throughout my crash course in A.T. culture, and even managed to nudge me around most mossy holes.

I cannot overstate how much I've enjoyed my time with the folks of the Environmental Data Center, Laboratory for Terrestrial Remote Sensing, and Natural Resources Science Department. I've never had to look far for a fresh perspective, sympathetic shoulder, or unexpected rubberband. Listing you all here would break the printer, and then I'd have to fix it. I'd especially like to thank Pete, Y.Q., and Chuck, who startled me with the revelation that ecology and computers aren't immiscible, and proceeded to enable my fledging addiction. And, of course, I'm infinitely grateful for the tireless aid of Debra Bourassa, without whom I'd manage to botch gravity.

Finally, I'd like to thank my family and friends for their limitless support throughout this endeavor. Especially my mother, Liz Argo, whose unique blend of stoicism and optimism has never ceased to amaze me. It probably helped that she went back and got her bachelor's right after mine, then nearly lapped me as she chased after a grad degree of her own. Talk about motivation. Finally, I dedicate this work to the memory of my father, Jimi Clark, whose legacy galvanized me to pursue a career investigating the world around me.

PREFACE

This document follows the standard thesis formatting requirements defined by the Graduate School of the University of Rhode Island. Readers may prefer to access a digital version of this document for ease of navigation and to view its numerous cartographic representations at higher resolutions. A Portable Document Format (.pdf) copy is available online from the Appalachian Trail Decision Support System website (accessible at [<www.edc.uri.edu/ATMT-DSS/>](http://www.edc.uri.edu/ATMT-DSS/) at the time of submission).

TABLE OF CONTENTS

Abstract.....	ii
Acknowledgements.....	iv
Preface.....	vi
Table of Contents	vii
List of Tables	ix
List of Figures.....	x
Introduction	1
Overview.....	1
<i>Ailanthus altissima</i>	4
Climate Change	8
Species Distribution Modeling	9
Methodology	12
Study Area.....	12
Data Sources	13
Ground Based Observations	13
Elevation Data.....	14
Landcover Data.....	15
Soils Data.....	16
Climate Data	16
Data Preparation	17
Habitat Modeling.....	18
Maxent Parameters	18

Variable Selection.....	19
Model Evaluation.....	21
Habitat Projection	23
Results	24
Analysis of Environmental Variables	24
Model Selection and Projection.....	25
Distribution and Trends	27
Discussion.....	28
Environmental Variables.....	29
Bioclimatic.....	29
Topographic	30
Land Cover	31
Scale	32
Distribution and Trends	28
Limitations	35
Conclusion	37
Tables	39
Figures.....	51
Appendices.....	81
Appendix I: Maps of All Environmental Variables	81
Appendix II: Maxent Log File.....	118
Bibliography	131

LIST OF TABLES

Table 1: Geospatial Data Sources	39
Table 2: Environmental Variables	40
Table 3: Correlation of Topographic and Landcover Variables	42
Table 4: Correlation of Bioclimatic Variables	43
Table 5: Environmental Variables included in Candidate Maxent Models	44
Table 6: Parameters and Evaluation of Candidate Maxent Models	45
Table 7: Maxent Model Variable Contributions	46
Table 8: Change in Bioclimatic Variables from Baseline to Projected Conditions	47
Table 9: Maxent Model Thresholds	48
Table 10: Current and Projected Area of Suitable Habitats	49
Table 11: Current and Projected Elevation and Latitude of Suitable Habitats	50

LIST OF FIGURES

Figure 1: A.T. Hydrologic Unit Code (HUC)-10 Shell	51
Figure 2: Ecological Provinces of the A.T.....	52
Figure 3: Perturbed Forest Inventory & Analysis (FIA) Plot Distribution	53
Figure 4: Elevation.....	54
Figure 5: Distance to Developed Landcover.....	55
Figure 6: Wetland and Open Water	56
Figure 7: Canopy Cover	57
Figure 8: CMIP5 Representative Concentration Pathways.....	58
Figure 9: Annual Mean Temperature.....	59
Figure 10: Annual Precipitation	60
Figure 11: Distribution of Projected Changes in Temperature	61
Figure 12: Distribution of Projected Changes in Precipitation	62
Figure 13: Forest Inventory & Analysis (FIA) Histograms, Boxplots, and T-tests.....	63
Figure 14: Environmental Variable Histograms, Boxplots, and T-tests	64
Figure 15: Marginal Variable Response Curves	69
Figure 16: Isolated Variable Response Curves	70
Figure 17: Variable Jackknifes.....	71
Figure 18: Model Performance (AUC)	72
Figure 19: Maxent Distribution for Current Conditions, Continuous.....	73
Figure 20: Maxent Distribution for Projected Conditions, Continuous.....	74
Figure 21: Multivariate Similarity Surface (MESS)	75
Figure 22: Most Dissimilar Variable (MoD)	76

Figure 23: Effects of Variable Clamping on Projection.....	77
Figure 24: Maxent Distribution for Current Conditions, Binary	78
Figure 25: Maxent Distribution for Projected Conditions, Binary	79
Figure 26: Change in Suitable Habitats, Binary	80

INTRODUCTION

Overview

The Appalachian National Scenic Trail (A.T.) is a footpath stretching from Springer Mountain in Georgia to Mount Katahdin in Maine and spanning over 3,500 km of peaks, valleys, and ridges along the Appalachian Mountains. It intersects 14 states; 8 National Forests; 6 units of the National Park System; more than 70 State Park, Forest, and Game Management units; and 287 local jurisdictions. The A.T. passes through some of the largest and least fragmented forest blocks remaining in the eastern United States (Dufour and Crisfield 2008); forests containing rich biological diversity and the headwaters of important water resources.

The A.T.'s north-south alignment and gradients of elevation, latitude, and moisture represent a continental scale cross-section, or "MEGA-Transect," of eastern U.S. forest and alpine areas, offering a setting for collecting scientific data on the structure, function, species composition, and condition of ecosystems. The high elevation setting of the A.T. provides an ideal landscape for the early detection of undesirable changes in the natural resources of the eastern United States; for example, development encroachment, acid precipitation, invasions of exotic species, and climate change impacts.

The Appalachian Trail (A.T.) Decision Support System (DSS) or A.T.-DSS is an Internet-based implementation and dissemination toolset accessible at <http://www.edc.uri.edu/ATMT-DSS>. Funded by NASA, the purpose of the A.T. DSS is to facilitate decision-making for the National Park Service (NPS) Appalachian National Scenic Trail (APPA), the Appalachian Trail Conservancy (ATC), and the

U.S. Forest Service (USFS), as well as provide a means to convey meaningful information to the public. The NPS Inventory and Monitoring (I&M) program is designed to develop and implement long-term natural resource monitoring and create a targeted decision support system aimed at selecting a suite of reliable and representative metrics, or *Vital Signs*, to provide long-term data on ecosystem trends. Among the Vital Signs defined by the I&M program and A.T. MEGA-Transect partners, the A.T.-DSS targets *phenology and climate change, forest health, and landscape dynamics* for system development, data preparation, and modeling.

The A.T.-DSS integrates NASA multi-platform sensor data, NASA Terrestrial Observation and Prediction System (TOPS) models, and *in situ* measurements from A.T. MEGA-Transect partners to support the resource management decision-making process. TOPS is a modeling framework that combines operational satellite data, microclimate mapping, and ecosystem simulation models to characterize ecosystem status and trends (Nemani 2009). By integrating NASA's Earth Observation System data and modeling products that link climate models (e.g., through TOPS) and ecological models (e.g., habitat suitability) with *in situ* observations (e.g., USFS Forest Inventory and Analysis data), the A.T.-DSS creates a coherent framework for data integration, monitoring, reporting and forecasting to improve the effectiveness of decision-making for managing the A.T. to conserve natural resources.

The objectives of the A.T. DSS include:

1. Develop a comprehensive set of seamless indicator data layers consistent with selected A.T. Vital Signs.

2. Establish a ground monitoring system to complement TOPS and integrate NASA data with field observations.
3. Assess historical, current, and forecasted ecosystem conditions and trends by coupling TOPS with habitat modeling.
4. Develop an Internet-based implementation and dissemination system for data visualization, sharing, and management to facilitate collaboration and promote public understanding of the A.T. environment.

Habitat modeling, Objective 3, is highly pertinent to the selected NPS I&M Vital Signs and exhibits the utility of the system by leveraging seamless geospatial data and climate models from TOPS along with ground based Forest Inventory and Analysis (FIA) data from the USDA Forest Service to model the current and potential distributions of important species within the A.T. To demonstrate these capabilities, a prototype habitat suitability model was developed for the invasive tree species *Ailanthus altissima* (Mill.) Swingle. Commonly referred to as Tree-of-Heaven, *Ailanthus* is a deciduous member of the Simaroubaceae family native to the temperate regions of central China. It is an exotic tree species pervasive throughout the United States due to its rapid growth, high fecundity, hardy tolerance, and strong competitive ability. Efforts to remove *Ailanthus* populations are most successful with early intervention, making regional monitoring a vital component of effective management.

The distribution of potential suitable *Ailanthus* habitats can be estimated across a broad landscape scale by combining field based observations, remote sensing data products, and statistical modeling algorithms. Once a model is established, alternative scenarios, such as climate change predictions, can be incorporated to examine the

impacts of potential shifts in habitat distributions. Modeling *Ailanthus* will facilitate the examination of an important process driving ecological change within native communities, while demonstrating the effectiveness of leveraging A.T.-DSS resources to support regional conservation and management goals. Therefore, the objectives of this research are threefold:

1. Relate field-based observations of the distribution of *Ailanthus* to a set of environmental variables.
2. Map the current distribution of suitable habitats and identify high-risk regions along the A.T.
3. Integrate projected precipitation and temperature data from TOPS based on IPCC climate change scenarios to simulate potential shifts in the distribution of *Ailanthus* habitats.

Ailanthus altissima

In its native range of central China, *Ailanthus* is valued for its rich cultural history and many traditional medicinal uses. However, many aliases used for the species in the Western world reflect a more dubious association, including “ghetto palm,” “stink tree,” and even “tree from hell.” *Ailanthus* was introduced to the eastern U.S. in 1784, when a gardener planted seeds in Philadelphia imported from England. It quickly became a popular ornamental due to its rapid growth, exotic appearance, and hardiness. In particular, it became a common shade tree in urban areas due to its high resilience and tolerance for pollution (Hu 1979, Hoshovsky 1988). *Ailanthus* is shade intolerant and often exploits gaps in the canopy to become established. Due to these traits, *Ailanthus* is often observed in urban settings and along transportation

corridors, as well as other disturbances. The species' tenacity in the midst of harsh urban conditions serves as a central metaphor for the novel A Tree Grows in Brooklyn (Smith 2009).

Ailanthus features reddish-brown branches laden with alternate pinnately compound leaves comprised of four to thirty-five leaflets (Miller 1990). The leaflets are lanceolate with two to four glandular teeth near a rounded base. Removing leaves reveals a distinctive heart-shaped scar. *Ailanthus* has a smooth, light-gray trunk and can reach heights of 17 to 27 meters. The species is dioecious, typically having male and female flowers on separate individuals, and bears small white to greenish-yellow flowers from mid-April to July. Male plants produce three to four times as many flowers as females and emit a strong odor to attract insect pollinators (Hu 1979, Miller 1990).

In late summer, *Ailanthus* develops many clusters of winged seeds, or samaras, which disperse from fall through spring. Mature females may release as many as 300,000 samaras (Hoshovsky 1988). Airborne samaras often disperse over distances greater than 100 m (Landenberger et al. 2006) and secondary wind dispersal facilitated by urban road corridors as far as 456 m (Kowarik and von der Lippe 2011). Samaras have germinated successfully after being submerged for extended periods of time, suggesting river corridors may provide a secondary pathway for dispersal (Kowarik and Säumel 2008).

Ailanthus has successfully colonized every continent with the exception of Antarctica. Suitable climates range from temperate to subtropical and humid to arid (Miller 1990). It is particularly common in temperate regions with typical conditions

consisting of long, warm growing seasons; regular winter frost; and annual precipitation >500mm (Kowarik and Saumel 2007). Populations have been recorded within 42 states across the U.S., from Florida to Oregon and from New Mexico to Maine (EDDMaps 2011, USDA 2012). *Ailanthus* is susceptible to frost damage, particularly juveniles, thus restricting *Ailanthus* from higher latitudes and elevations (Fryer 2010). *Ailanthus* is relatively drought hardy, though extended dry periods exclude the species from extremely arid regions. The species is sometimes reported to have a root system vulnerable to flooding (Miller 1990), well others consider it relatively tolerant (Fryer 2010).

Within its introduced range, *Ailanthus* is strongly associated with urban areas and transportation corridors (Hu 1979, Landenberger et al. 2009), due in part to its historic role in urban forestry. The species is also very hardy, able to tolerate harsh environmental pollutants and thrive across a wide range of poor soils (Miller 1990), and has even been considered for strip mine reclamation. These traits, along with shade intolerance, make *Ailanthus* ideally suited to colonize the ruderal conditions found in human-impacted or otherwise disturbed areas. Seedlings develop rapidly once established and are capable of growing over a meter the first year (Hoshovsky 1988). *Ailanthus* can sprout from roots as well as stumps and often forms dense thickets. In addition to shading competitors, *Ailanthus* releases allelopathic compounds into adjacent soils that suppress the development of competing seedlings (Lawrence et al. 1991, Gómez-Aparicio and Canham 2008).

This strong competitive ability enables *Ailanthus* to severely impact native communities within its introduced range. Disturbance, both natural and

anthropogenic, is a cyclical process within the landscape. Forest canopy gaps resulting from disturbance allow sunlight to penetrate to the understory, facilitating the establishment of early successional species. As succession proceeds, these pioneers are gradually replaced by more shade tolerant species. This ongoing cycle of disturbance and succession maintains heterogeneity within the landscape (Connell and Slatyer 1977). By interrupting this process and suppressing native species (Fryer 2010), *Ailanthus* modifies the vegetative community (Hejda et al. 2009), and by extension alters resources and ecosystem services other species depend upon .

The economic and environmental costs of a biological invasion can be significant (Pimentel et al. 2005, Cardinale et al. 2012). Invasive species disrupt the balance of ecosystems by outcompeting and displacing native species (Mack et al. 2000). Biodiversity is lost and habitats are damaged, habitats that rare and endangered species may depend on (Benning et al. 2002). As the community changes, essential ecological functions alter and ecosystem services, such as drinking water filtration (Brauman et al. 2007) and timber production, are degraded (Charles and Dukes 2007, Vilà et al. 2011). Finally, eliminating native vegetation diminishes the aesthetic quality of a region. This is particularly significant for the A.T., where natural vistas are highly valued by the public (Shriver et al. 2005). Proactive management efforts are needed to mitigate the damaging spread of *Ailanthus*.

Removing *Ailanthus* allows native vegetation to recover (Burch and Zedaker 2003) and prevents further dispersal. However, management efforts are confounded by the tree's ability to resprout from roots and stumps. Physical control methods, such as pulling or cutting, must be sure to remove the entire root system or the remnants

will rapidly regenerate. While some seedlings may be removed entirely by hand, the stumps of more mature plants must be treated with herbicide to prevent resprouting (Hoshovsky 1988, Fryer 2010). Early detection is crucial for minimizing the costs of control programs and the risks of further dispersal and establishment. A greater understanding of the processes and patterns of *Ailanthus* invasion within the landscape is needed to inform effective management programs (Peterson and Vieglais 2001, Byers et al. 2002, Thuiller et al. 2005).

Climate Change

Climate change is a widely recognized phenomenon (IPCC 2007) with significant implications for the spread, impact, and management of invasive species (Walther et al. 2009, Dukes 2011). Climate change alters temperature and precipitation patterns, resource availability (CO₂, N), and affects management decisions and practices in land-cover and land-use (Bradley et al. 2010). Hellman et al. (2008) identify five groups of potential interactions between climate change and biological invasion: altered pathways of introduction, likelihood of new invasions, distribution of existing invasions, impacts of invasion, and effectiveness of management strategies. It is challenging to incorporate the full extent of complex factors driving *Ailanthus* invasion, especially potential inter-specific interactions. However, the broad geographic range, high dispersal, and rapid growth of *Ailanthus* suggest that it will adapt to changing conditions more readily than most native species, giving this invasive species a decisive competitive advantage as the frequency of disturbances within the landscape increases (Dale et al. 2001).

One of the focuses of this study is the assessment of the direct effects of climate change, i.e. temperature and precipitation trends, on *Ailanthus* habitat suitability. Warming trends have been predicted to correspond with horizontal migrations of vegetation averaging 0.43 km yr^{-1} across a wide variety of ecosystems (Loarie et al. 2009). In particular, for the A.T. region, temperatures are predicted to increase by 2°C to 6°C by the end of the 21st century (Hashimoto et al. 2011). While studying all of the significant factors that influence *Ailanthus* is beyond the scope of this project, modeling the climatic envelope of *Ailanthus* will provide insight on the future distribution of suitable habitat and potential ecological impacts (Pearson and Dawson 2003, Jeschke and Strayer 2008, Dukes 2011).

Species Distribution Modeling

Innovative statistical methods and advances in Geographic Information System (GIS) technology have led to the emergence of species distribution modeling as an important ecological tool (Guisan and Zimmermann 2000). Ecological niche theory examines the relationship between species fitness and environmental conditions (Hutchinson 1957); species distribution modeling extends this paradigm into geographic space by linking species distribution to spatial variability (Austin 2002, Brotons et al. 2004, Hirzel and Le Lay 2008). Species distribution modeling is a rapidly developing and highly diverse field, and may be alternatively referred to as bioclimatic models, climate envelopes, ecological niche models, habitat models, resource selection models, or range maps, often with varied emphases and interpretations (Elith and Leathwick 2009, Sillero 2011). These models generate spatially explicit predictions of species occurrence, typically by comparing

environmental variables between species presence and absence locations. Common techniques range from simple environmental envelopes, e.g., BIOCLIM (Busby 1991); to machine learning based algorithms, e.g. Genetic Algorithm for Rule-set Production (GARP) (Stockwell 1999). Species distribution modeling has been used to examine a wide variety of populations and scenarios; including risk assessments for invasive species (Peterson 2003, Dullinger et al. 2009, Jones et al. 2010, Robinson et al. 2010) and the potential impacts of climate change (Kriticos et al. 2003, Thomas et al. 2004, Jarnevich and Stohlgren 2008, Bradley et al. 2009, Elith et al. 2010). Albright et al. (2009) used herbarium records and generalized linear regression to predict the distribution of *Ailanthus* in both the U.S. and China. This study will build on these results using more detailed presence records, higher spatial resolution, alternative statistical techniques, and climate projections.

Modeling the suitable habitats of an invasive species presents a unique challenge. A major assumption underlying most models is that the absence of a species from a particular area indicates that conditions found there are unsuitable for the species (Lobo et al. 2010). However, by definition the distribution of an invasive population may not have reached equilibrium within the landscape (Sakai et al. 2001, Robinson et al. 2010). The absence of an invasive species from a particular location may not indicate unsuitable conditions, but rather that the species simply hasn't been introduced or dispersed into that area. These characteristics of invasive populations necessitate the use of 'presence-only' species distribution modeling techniques (Elith and Leathwick 2009).

Maximum entropy (Maxent) modeling is a machine learning based method for predicting species geographic distributions from presence-only data (Phillips et al. 2006). Several comparative studies of species distribution modeling methods have ranked Maxent among the top modeling approaches (Hernandez et al. 2006, Hijmans and Graham 2006, Elith et al. 2006, Elith and Graham 2009). With Maxent, the true distribution of a species is estimated as a probability distribution across all sites within the study area. The probability distribution adheres to a set of constraints derived from the presence data while maximizing entropy. The maximum entropy distribution is that which draws the least inferences beyond the available information, i.e. the most spread out or closest to uniform (Phillips et al. 2006). A simple example is the normal distribution, which maximizes entropy within the constraints dictated by the given information: the mean and standard deviation. With species distribution modeling, the set of constraints are functions of environmental variables. That is, the mean environmental conditions predicted by the model should be close to the conditions observed at presence locations (Phillips and Dudik 2008, Elith et al. 2011).

The environmental variables used to represent ecological conditions within the study area may also be termed independent variables, covariates, predictors, or inputs. Due to the complex relationship between the species and environment, functions fitted by Maxent are typically non-linear composites of many transformations of the covariates, termed features. Feature classes that may be fit to the distribution include linear, quadratic, product, threshold, categorical, and hinge transformations. Maxent begins with a uniform probability distribution and repeatedly adjusts the weights of features to maximize the probability, or gain, of the observed species presence points

(Phillips et al. 2006). To limit over-fitting, gain is reduced by a regularization parameter that penalizes complex features. The model iterates until the increase in gain falls below the convergence threshold or the maximum number of iterations are reached (Phillips 2005, Phillips et al. 2006, Elith et al. 2011). Once a suitability model has been established, projected climate data can be substituted for current conditions to examine potential shifts in the distribution of suitable *Ailanthus* habitats (Elith et al. 2010).

METHODOLOGY

Study Area

The A.T. is an open and complex system. The spatial extent of the A.T.-DSS is adopted from a boundary defined by the NPS and USGS. It was established by selecting all 10-digit Hydrological Unit Code (HUC-10) watersheds within 5 statute miles of the A.T. land base, termed the A.T. HUC-10 shell (Dieffenbach 2003). The shell provides an ecologically relevant boundary around the A.T. for habitat suitability modeling (Figure 1).

Bailey's ecoregion provinces were used to further delineate the study A.T.-shell into sub-units (Figure 2), facilitating the closer examination of trends in habitat distributions across the study area. Bailey's ecoregions are a hierarchical classification system which groups areas with similar climates and dominant potential vegetation (Bailey 1998).

Data Sources

Ground Based Observations

Ground-based observations of *Ailanthus* were provided by the Forest Inventory and Analysis (FIA) program of the USDA Forest Service (Woudenberg et al. 2010). The FIA program was established to provide a comprehensive inventory and analysis of the present and prospective conditions of and requirements for the renewable resources of the forest and rangelands of the U.S. (USFS 2005a). Forest monitoring is a central component of FIA, which provides a nationwide systematic sample of a wide array of measurements on forested ecosystems (USFS 2005b). Phase 1 of the systematic sample uses remote sensing to stratify land cover in the United States into forested and non-forested lands. Phase 2 establishes one field sample per 2,000 ha of forest, with 15% of plots measured each year in eastern states. FIA plots consist of a cluster of four circular 24-foot radius subplots spaced out in a fixed pattern (USFS 2005c). Measurements include the species, size, and condition of trees within the plot, as well as physiographic site attributes (Woudenberg et al. 2010).

To protect the privacy of private forest landowners, a portion of the FIA plot locations are altered before the records are made publically available. The majority of plots are perturbed, or ‘fuzzed,’ which adjusts the plot coordinates to within a 1.6 km radius of the true location. A smaller portion of plot records are swapped with a plot with similar ownership and ecological condition (Lister et al. 2005). However, Coulston et al. (2004) found that 95% of perturbed plot locations were within 0.8 km of the true locations. Furthermore, the uncertainty introduced into spatial models by

perturbation decreases as ancillary data spatial resolution decreases and spatial autocorrelation increases (Coulston et al. 2006).

Within the A.T.-shell, 3,926 FIA plots were visited and measured between 2002 and 2010, and observations of *Ailanthus* were recorded at 136 locations (Figure 3). In addition to the plot coordinates, several attributes were retained from the FIA records to examine the characteristics of sites colonized by *Ailanthus* and compare them to the overall study area. Plot attributes included elevation, aspect, slope, distance to improved road, land ownership, water on plot, physiographic class, stand age, stand size, and basal area of live trees (Woudenberg et al. 2010).

Elevation Data

Topographic information within the A.T.-shell was supplied by the National Elevation Dataset (NED), a 30-meter resolution digital elevation model (DEM) produced by the USGS with seamless coverage across the conterminous United States. The NED is compiled from the best publically available elevation data and undergoes rigorous accuracy assessments (Gesch et al. 2002). Individual tiles spanning the study area were acquired from seamless.usgs.gov, mosaicked, and clipped to the boundary of the A.T.-shell (Figure 4).

Several additional variables were derived from the NED to further characterize topography using ArcGIS 10 Spatial Analyst Tools and the Geomorphometry and Gradient Metrics toolbox (Evans and Oakleaf 2011). Slope calculates the maximum rate of change in elevation from the focal raster cell to its neighbors within an 800-m radius. Slope position subtracts a focal mean of elevation from the original elevation raster. The compound topographic index (CTI) is a steady state wetness index and is a

function of slope and upstream contributing area. The topographic radiation aspect index (TRASP) transforms circular aspect into a continuous variable better suited for modeling. Cooler and wetter north- northeast orientations are assigned values close to zero, while hotter and dryer south- southwest orientations are closer to one. The heat load index (HLI) is similar to TRASP, but also accounts for slope steepness.

Landcover Data

Landcover was used to represent the distribution of important cultural and biological features within the landscape. The National Land Cover Database (NLCD) is a product of the Multi-Resolution Land Characteristics (MRLC) Consortium, a partnership of Federal agencies led by the USGS. Land cover classifications are based on Landsat imagery and an array of ancillary data layers, and are available with seamless coverage across the conterminous U.S. from <http://www.mrlc.gov> for the years 1992, 2001, and 2006 at 30-m spatial resolution (Fry et al. 2011). Potentially significant classes were extracted from NLCD2006; including developed (Figure 5), agricultural, wetland, and open water areas (Figure 6). Layers were also generated measuring the distance from each pixel to the nearest agricultural and developed feature, respectively, to reflect their strong association with *Ailanthus* dispersal. In addition, a layer from NLCD2001 for percent tree canopy (Figure 7) was used to examine the shade intolerance of *Ailanthus*.

Soils Data

The USGS created the USSOILS dataset by combining the many individual mapping units comprising the State Soil Geographic Database (STATSGO), a collection of detailed soil surveys managed by the NRCS, into a seamless coverage of polygons spanning the conterminous U.S. (Schwarz and Alexander 1995). The polygons of soil mapping units were transformed to a 300-m resolution raster within ArcMap. Individual raster layers were extracted for soil drainage class, flood frequency, and hydric soils.

Climate Data

NASA's Terrestrial Observation and Prediction System (TOPS) provided baseline and projected climate data from the Intergovernmental Panel on Climate Change (IPCC) Fifth Assessment Report (AR5) Coupled Model Intercomparison Project Phase 5 (CMIP5) (Taylor et al. 2012). The CMIP5 is an ensemble of 16 individual General Circulation Models (GCMs) that predict future conditions under a set of alternative scenarios defined by Representative Concentration Pathways (RCPs) (Figure 8). RCPs represent the atmospheric concentration of greenhouse gasses, or radiative forcing values, in the year 2100 resulting from future scenarios with varying levels of global emissions and mitigation. RCP6.0 was selected for *Ailanthus* modeling, as it represents a moderate increase in radiative forcing that stabilizes by 2100 due to technologies and strategies for reducing greenhouse gas emissions (Vuuren et al. 2011). An ensemble of CMIP5 data were downscaled to 250 m and subset to the A.T.-shell for two time periods, a 1950-2005 baseline and projections for 2090-2095. Multidimensional ERDAS Imagine rasters for average monthly maximum

temperature, minimum temperature, and precipitation were created, with an individual band for each month.

In addition to annual mean temperature (Figure 9) and precipitation (Figure 10), a set of 19 bioclimatic variables were derived reflecting annual trends, seasonality, and extreme or limiting environmental factors. The bioclimatic variables were calculated for both the current and projected sets of climate data using the “biovars” function of the R package “dismo” (Hijmans et al. 2012). The variables are intended to provide more biologically meaningful measures of conditions that are likely to restrict the range of *Ailanthus*. Finally, the distribution of climatic changes within the A.T.-shell was visualized by subtracting present variable rasters from projected variable raster within ArcGIS to generate difference surfaces (Figure 11, Figure 12).

Data Preparation

Data processing, analysis, and modeling were conducted using a Dell PowerEdge T310 system running Microsoft Windows Server 2003 x64 with an Intel Xeon X3460 processor and 16 GB of RAM. All environmental layers were preprocessed to conform to a uniform spatial extent, resolution, and geographic projection, and converted to ESRI ASCII grid format prior to Maxent modeling. Data were prepared using ArcGIS 10 and custom tools were created using the ModelBuilder module to streamline workflows. Layers with global and continental extents were first clipped (subset) to the eastern U.S. to expedite subsequent processing. All layers were reprojected to the Albers equal-area conic projection,

using the “NAD_1983_To_WGS_1984_5” transformation as needed, and stored in ERDAS .img format.

The layers were clipped using the shapefile of the A.T.-shell boundary with a one-mile buffer. Retaining data within a buffer around the shell circumvents distortions caused by edge effects near the boundary of the study area. Operations were performed at this stage to derive additional layers (e.g., heat load index from elevation) and transform layers (e.g. natural log of distance to development). The layers were then resampled to 300 m using a snap raster template to ensure cell (pixel) alignment agreed perfectly between layers. Bilinear interpolation was used for downscaling continuous variables with resolution coarser than 300 m.

The layers were then adjusted to reflect the uncertainty introduced by FIA plot location perturbation by calculating focal statistics with a moving window radius of 800 m. For continuous variables, the mean value of cells within 800 m of the focal cell was calculated, and for categorical variables the total counts for each category of interest were tallied. Finally, the layers were clipped to the A.T.-shell and exported to ESRI ASCII grid format. See Table 1 for the geospatial data sources, Table 2 for a list of derived environmental variables, and Appendix I for cartographic representations of all environmental variables.

Habitat Modeling

Maxent Parameters

The predictive performance of Maxent is influenced by the choice of feature types fitted to the environmental variables and the regularization constants used to control overfitting. While complex models may accurately predict populations at

equilibrium, simpler models are likely better suited for range shifting populations (Elith et al. 2010, Jiménez-Valverde et al. 2011). The default Maxent settings for feature classes and regularization are adapted from a study that tuned parameters based on datasets for 226 species across 6 regions, and have been shown perform well across a wide range of applications (Elith et al. 2006, Phillips and Dudik 2008). These default settings were retained with the exception that feature classes were limited to hinge features. Hinge features form piece-wise linear functions and have been shown to improve model performance (Phillips and Dudik 2008), while providing simple, smooth models appropriate for predicting the projected distribution of range shifting species (Elith et al. 2010).

Variable Selection

For an initial assessment of *Ailanthus* habitat characteristics, FIA plot locations within the A.T.-shell were imported into ArcGIS to append values from the collection of environmental variables to the point data attributes. Histograms, boxplots, and t-tests were constructed to compare the distributions of *Ailanthus* presence and absence points across the sets of FIA attributes (Figure 13) and environmental variables (Figure 14).

Incorporating a large number of variables into a Maxent model may lead to complex solutions that obscure important ecological relationships, resulting in unexpected or erroneous behavior when extrapolating the model to future conditions (Mac Nally 2000, Warren and Seifert 2011, Elith et al. 2011). Pearson correlation coefficients were calculated for every pairwise combination of topographic and landcover environmental variable rasters (Table 3), as well as bioclimatic variables

(Table 4) using ENMtools (Warren et al. 2010). The number of variables was reduced iteratively by evaluating an initial model incorporating many variables, eliminating variables with weak or inapt contributions, and running the revised model. An extensive array of tools was used to evaluate variable performance; including variable response curves, percent contribution and permutation, jackknifing, and Pearson correlation coefficients.

Marginal variable response curves plot the change in logistic prediction from varying the value of one environmental variable while holding all other variables constant at their average sample value (Figure 15). Strongly correlated variables may confound the interpretation of marginal response curves, as the actual Maxent model can incorporate features where variables change together. Isolated variable response curves (Figure 16) represent a model incorporating only the focal variable and may be more informative when dealing with highly correlated variables (Phillips et al. 2006). The shape of the response curves is highly informative. Variables with complex surfaces may indicate overfitting, while sharp increases or decreases near the limits of the environmental range sampled increase uncertainty when extrapolating the model to projected conditions. Finally, the appropriateness of the response curve shape should be considered in the context of the ecological understanding of *Ailanthus* (Austin 2002, Austin 2007, Elith et al. 2010).

The percent contribution and permutation importance of each variable are also provided with the Maxent model output. Percent contribution is calculated as the training algorithm iterates by adding the increase in regularized gain, or subtracting if the absolute value of lambda is negative. Permutation importance is determined by

randomly changing the values of the focal variable for the training and background data, reevaluating the model with each permuted variable in turn, and recording the corresponding drop in training AUC normalized to a percentage. As with marginal response curves, the interpretation of variable contributions is confounded by high correlation (Phillips 2005, Elith et al. 2011).

Maxent also evaluates the set of input environmental variables by performing jackknife tests (Figure 17). The jackknife tests compare the regularized training gain, test gain, and test AUC for a set of models created while withholding each variable in turn and with each variable in isolation. Variables may perform well in isolation but make little difference on the overall model prediction, indicating that they are good predictors but contain little information not present in the other variables. In other cases, jackknife tests may indicate that withholding the variable actually increases model performance. Finally, a variable which performs well for the training data but badly with the test data is poor at generalizing, and therefore less transferable (Phillips 2005, Elith et al. 2011).

Model Evaluation

For each candidate set of variables, a 10-fold (replicate) cross-validated model was generated with 122 of the *Ailanthus* presence points used for model training and the remaining 14 for testing. As with variable selection, a variety of methods were used to assess model performance. Models were evaluated based on their performance on test data, parameter complexity, ecological consistency, and degree of extrapolation required when projecting to future conditions.

To compare model performance, the Maxent package determines the area under the curve (AUC) of the receiver operating characteristics (ROC) for each model. The ROC curve is constructed by plotting model sensitivity and specificity (Figure 18). Sensitivity is a function of the omission rate, i.e. the rate training or test presence points incorrectly classified as unsuitable. Specificity is typically the commission rate, or rate of absence points incorrectly classified as suitable. However, given the lack of absence data for presence only modeling, specificity is instead derived from the fraction of the study area predicted as suitable (Phillips 2005). While the application of AUC to presence only modeling is not without limitations (Lobo et al. 2008), it nevertheless provides a threshold independent measure of model predictive performance on withheld test data. In addition, comparing predictive performance on training versus test data, as well as the standard deviation of scores across replicates, provides an indication of model transferability. A model with high training but low test AUC is likely overfit to the training data, and may perform poorly when extrapolated to new environmental space (e.g., climate projections) (Warren and Seifert 2011).

Model complexity was assessed using sample-size corrected Akaike information criteria (AIC_c), as proposed by Warren and Seifert (2011) and implemented within ENMtools (Warren et al. 2010). AIC_c is determined from model log likelihood (the product of suitability scores across all presence points) penalized by the number of parameters (the complexity of features applied to the environmental variables). Models with lower AIC_c scores balance high predictive performance with low complexity, and are likely more appropriate for extrapolating to future conditions.

Hinge features are penalized more heavily than other feature classes as they incorporate more parameters. Therefore, AIC_c scores were only compared between models including the same selection of feature classes (Warren and Seifert 2011).

Habitat Projection

Once a model was selected, TOPS AR5 GCM data were substituted for current climate variables and suitability recalculated. The projections used variable clamping to ensure that the values of projected variables were restricted to the range of values encountered while training the model under current conditions. In addition to the projected distribution, Maxent provides outputs for evaluating divergence of current and projected variables, as well as the influence of variable clamping.

A threshold must be applied to the continuous probability distribution to provide a binary map of suitable and unsuitable locations. The Maxent output includes several common thresholds and their omission rates. Of these, “Balance training omission, predicted area and threshold value,” henceforth the Balance threshold (Bt), provided the best compromise between overfitting and over-commissioning:

$$\text{Bt} = \text{Minimize } 6 * \text{training omission rate} + .04 * \text{cumulative threshold} + 1.6 * \text{fractional predicted area}$$

The threshold was applied to both current and projected distributions by reclassifying the two raster files within ArcGIS, and their total suitable area, mean latitude, and mean elevation were calculated. The previous metrics were also calculated within each ecological province intersecting the overall A.T.-shell (Figure 2), facilitating the comparison of regions with similar conditions across time.

RESULTS

Analysis of Environmental Variables

Comparing the FIA attributes throughout the study area with the subset of *Ailanthus* presence points revealed patterns that reflected the habitat preferences of *Ailanthus* as reported in the literature (Figure 13). Topographically, *Ailanthus* was generally observed at sites with lower elevations (FIA_ELEV). FIA field crews also assigned a xeric, mesic, or hydric physiographic class to each site (FIA_PHYSCLCD). *Ailanthus* sites were most frequently mesic, sometimes xeric, and rarely hydric. Similarly, *Ailanthus* was less frequent at sites where water bodies (e.g., permanent or temporary streams) were recorded (FIA_WATERCD). Culturally, *Ailanthus* was found at sites closer to roadways (FIA_RDISTCD) and located on private lands more often than publically held (FIA_OWNGRCD). Biologically, *Ailanthus* was more prevalent in younger forest stands (FIA_STANDAGE).

A similar analysis of the environmental variables (Figure 14) further illustrated the characteristics of *Ailanthus* habitats. *Ailanthus* was more likely to occur in warmer (bio1) and dryer (bio12) climates. Elevation (dem) mirrored the distribution of the FIA plots, with a clear preference for lower sites. However, variable contributions were low, likely due to correlation with other variables. Slope position (slopepos) was highly correlated with elevation and shared a similar distribution. Aspect (trasp) was slightly skewed towards southern exposures, with trasp outperforming hli. *Ailanthus* was found closer to agriculture (agdist) and development (devdist), but these variables only performed moderately within Maxent.

No trends were clear for the compound topographic index (CTI) or soil flood frequency (soil_fldfreq). Hydric soil (soil_hydric), drainage class (soil_drain), and open water/wetlands (nlcd_wet) revealed a preference for dryer areas.

Nine highly correlated ($r > 0.90$) bioclimatic variables were removed (Table 4), while no topographic or landcover variables were eliminated (Table 3). Strongly performing or ecologically limiting variables were selected over generalized or erratically performing variables. Notable exceptions were the mean temperature of coldest quarter (bio11) and temperature seasonality (bio4). Both variables were retained despite high correlation ($r = 0.91$) due to the ecological significance of bio11 and exceptional performance of bio4 in preliminary jackknife tests (Figure 17).

Maxent consistently selected climate variables as the most important predictors of suitability. Mean temperature of coldest quarter (bio11) ranked highest for percent contribution and permutation importance across model runs (Table 7). Jackknife tests for temperature seasonality (bio4, Figure 17) had the highest gain when used in isolation and the largest decrease in gain when omitted, indicating that seasonality contains information both most useful by itself and not present in other variables. Conversely, some variables reduced test gain, such as the compound topographic index (cti) and distance to development (devdist), and were removed from subsequent models.

Model Selection and Projection

Model “4bio_4topo” (Figure 19) was selected as the highest performer from an array of 15 models incorporating alternative sets of environmental variables (Table 5). This model had the lowest AIC_c score (3707.5), a high mean test AUC (0.85), and low

test AUC standard deviation (0.034) between replicates. The model's AIC_c score outranked models incorporating both more and less variables, suggesting its level of complexity was nearest to optimal, while the low standard deviation indicates a robust model with high transferability (Table 6).

The model incorporated a limited set of environmental variables (Table 7) with clear ecological interpretations. Mean temperature of coldest quarter (bio11) made the largest contribution to the model (40.1%), with an isolated variable response curve indicating high suitability for sites with warmer winters (Figure 16) but decreasing rapidly within the conditions of the trails southern extremes. Temperature seasonality (bio4) contributed 27.1% of regularized gain, with response curves reflecting a preference for the moderate seasonal variation over the seasonal extremes of the northern A.T. or the steady warmth of to the south (Figure 16). Mean temperature of wettest quarter (bio8) made the third largest contribution (22.7%) with a marginal response curve similar to bio11, suitability increasing with temperature with a sharp decrease at the upper extreme of temperatures sampled (Figure 16). Slope and aspect (trasp) contributed 4.6% and 3.1% of regularized gain, respectively, with a preference for moderate slopes and drier, sunnier aspects. While the remaining variables appear to have contributed very little (nlcd_wet = 0.9%, bio19 = 0.8%, dem = 0.6%), their importance may have been obscured due to correlation with other variables.

Bioclimatic variables calculated from TOPS AR5 GCM RCP 6.0 data were substituted for current climate variables and suitability was recalculated (Figure 20). From the 1950-2005 baseline to the 2095-2099 projection, mean temperature seasonality (bio4) increased by 3.84×10^{-3} °C, mean temperature of coldest quarter

(bio8) increased by 2.18 °C, mean temperature of wettest quarter (bio11) increased by 4.29 °C, and mean precipitation of coldest quarter (bio19) increased by 532×10^{-4} m (Table 8). The multivariate similarity surface (MESS) revealed that the projected variables were outside the range encountered during training in three regions (Figure 21). Mean temperature of coldest quarter (bio8) was the most dissimilar variable across the New Jersey and Virginia trail sections, while mean temperature of wettest quarter (bio11) was most dissimilar in the southern section (Figure 22). If clamping had not restricted projected variables to the range of values encountered during training, the model would predict a dramatic decrease in suitability throughout the mid-Atlantic (Figure 23).

Distribution and Trends

The “Balance” threshold value provided by Maxent (Table 9) was used to reclassify the continuous raster surfaces into two discrete suitability classes. All cells with a logistic probability greater than the threshold value of 0.047 were predicted as suitable, and all cells below unsuitable (Figure 24). For the current distribution, this produced a fractional predicted area of .560 and a training omission rate of 0. The threshold was also applied to the projected distribution (Figure 25), and a change map was created to visualize shifts in suitability (Figure 26).

The Maxent model of current conditions (Figure 24) predicts that 60,044 km², or 56%, of the A.T.-shell is potentially suitable for *Ailanthus* colonization. By 2095, the suitable areas are projected to expand to 89,066 km² (82%), an increase of 48% (Figure 25). Ecoregion province M211 – Adirondack-New England Mixed Forest exhibited the most dramatic increase, from 2% to 50% total area. Suitable area also

increased in provinces 221 – Eastern Broadleaf Forest and M221 – Central Appalachian Broadleaf Forest by +15% and +25%, respectively. The significance of any trends observed in provinces 211 and 231 are limited due to the small portion of the A.T.-shell they encompass (7% of A.T.-shell, combined) (Table 10).

The mean elevation of suitable areas is projected to increase by 59 m (from 391 m to 449 m) and the mean latitude to shift north by 49 km. Elevation increased in province M211 by 28% (96 m). Ranges shifted north in provinces M211 (108 km) and 221 (36 km), and south in M221 (61 km) (Table 11).

DISCUSSION

Distribution and Trends

The distribution of suitable habitats estimated by the Maxent model (Figure 19) largely coincides with the existing knowledge of *Ailanthus* distribution within the eastern United States (EDDMaps 2011, USDA 2012). The majority of locations with high suitability fall within the Virginian and mid-Atlantic sections of the A.T, which is expected given the distribution of FIA presence points. Conditions in these regions are ideal, with moderate to low rainfall, low elevations, mild winters, and abundant development. Suitability decreases as the trail moves south into the Smoky Mountains and elevation and precipitation increase, and development thins. To the north, suitability again decreases as elevation increases and temperature drops. The Northeast is predicted to contain the least suitable areas along the A.T. While *Ailanthus* invasions are reported throughout the Northeast (Hu 1979, EDDMaps 2011, USDA 2012), and historically abundant in New England, these records predominately

occur within the low elevations and dense population centers along the Atlantic seaboard, rather than the remote, mountainous regions the A.T. passes through. The increase in suitability predicted as the A.T. leaves the Kittatinny Mountains in New Jersey and approaches the New York City metro area seems to support this conclusion.

Estimating the future distribution of suitable *Ailanthus* habitats by integrating climate projections reveals several interesting trends. Overall, there is a 48% percent increase in suitable area, representing a dramatic increase in the potential extent of *Ailanthus* invasion. Subdividing the A.T.-shell by ecological province delineates the area into units with similar environmental conditions and ecological communities, providing insight into the processes driving this expansion. The most dramatic increase occurs in M211: Adirondack-New England Mixed Forest-Coniferous Forest-Alpine Meadow, where warmer temperatures allow *Ailanthus* to expand north with a 49-km increase in mean latitude, as well as to higher elevations. Conversely, the average latitude in M221: Central Appalachian Broadleaf Forest-Coniferous Forest-Meadow actually shifts south, and average elevation increases, as *Ailanthus* migrates into the Smoky Mountains.

Environmental Variables

Bioclimatic

The performance of various environmental variables assessed throughout Maxent modeling indicates that the distribution of suitable *Ailanthus* habitats is primarily constrained by climate conditions at a regional scale. The mean temperature of the coldest quarter (bio11) and temperature seasonality (bio4) were particularly

significant (Table 7), with a preference for warmer and milder conditions (Figure 16). *Ailanthus* saplings are highly vulnerable to frost mortality (Miller 1990, Kowarik and Saumel 2007) and annual die-backs may restrict occurrence to lower elevations and warmer regions (Fryer 2010), a limiting factor that the marginal response curve for bio11 appears to reflect (Figure 15). Kowarik and Saumel (2007) also note that while *Ailanthus* tolerates a wide range of climatic conditions, temperature seasonality strongly affects growth, dispersal, and survival.

Variables relating temperature to precipitation also performed well (Table 7). While annual mean temperature (bio1) and annual precipitation (bio12) performed well during preliminary modeling, the more nuanced climate variables were more capable of capturing the extreme factors that limit the success of *Ailanthus*. *Ailanthus*'s preference for high mean temperatures during the wettest quarter (bio8, Figure 16) may indicate increased mortality due to frost stress and mechanical damage associated with winter storms (Lemon 1961), or it may simply reflect a broader preference for warmer climates. While there are conflicting accounts of *Ailanthus* flood vulnerability (Miller 1990, Kowarik and Säumel 2008), its exceptional drought tolerance is well established (Trifilò et al. 2004, Kowarik and Saumel 2007). This trait is evident within the response curve for the rainfall of the coldest quarter (bio19, Figure 16), and appears to support the frequently reported preference for drier soils.

Topographic

Topographic variables were also included in the final distribution model. Slope made the largest contribution (Table 7), with suitability being highest at moderate gradients (Figure 16). An extensive root system allows *Ailanthus* to

colonize rough terrain and steep slopes (Fryer 2010). These extreme areas may coincide with decreases in canopy density and increase access to direct sunlight, a primary *Ailanthus* habitat requirement (Kowarik and Saumel 2007). However, it should be noted that the maximum slope was only 26 degrees due to the variable's 300-m pixel resolution and the adjustments made to reflect FIA plot location perturbation.

The heat load index (HLI) was discarded in favor of the topographic radiation aspect index (TRASP), as HLI's incorporation of slope was inappropriate given the inclusion of slope as a separate variable. The variable response curves for TRASP (Figure 16) indicates that *Ailanthus* prefers the sun exposure, and therefore increased temperature and decreased humidity, of the south-southwest facing slopes. Suitability decreased in the presence of wetlands (nlcd_wet, Figure 16), further indicating a preference for dryer sites. While compound topographic index (CTI) was theoretically well suited to identify wet and dry positions in the landscape, it too was inhibited by the coarse resolution and the locational fuzzing treatment.

Finally, the response curve for elevation (dem, Figure 16) clearly reflects *Ailanthus*'s characteristic association with low laying, mild, heavily developed areas. While elevation made the smallest lowest contribution (Table 7), it had the fourth highest permutation importance. Its influence was likely diminished due to correlation with other variables better suited to discriminate the underlying mechanisms disrupting *Ailanthus* establishment, such as low temperatures and frost mortality.

Land Cover

Land cover variables, while potentially significant, proved difficult to incorporate into the model. The association between *Ailanthus* and urban areas and canopy cover is very prevalent throughout the literature (Hu 1979, Miller 1990, Landenberger et al. 2009) and apparent from the analysis of FIA plot data (Figure 13). Plotting an isolated variable response curve illustrates the relationship: suitability decreases exponentially as the distance to development increases. However, when distance to development (devdist) and canopy cover standard deviation (lfcc_std) were added to the final model (4bioalt_4topo_2lc, Table 5), they each only made an infinitesimal contribution of 0.6%. Furthermore, the overall test AUC of the model decreased when the two land cover variables were included (Table 6). One factor suppressing the importance of land cover variables may be correlation with bioclimatic and topographic variables. The regions within the A.T.-shell furthest removed from urban areas also contain some of its most extreme conditions. Maine to the north and the Smoky Mountains to the south contain remote areas, but are also at very high elevations with low temperatures and high rainfall, respectively, as well as large tracts of forest. In other words, the broad spatial extent of the A.T.-shell contains regional-scale sociogeographical patterns that obscure potential finer-scale relationships between land cover and *Ailanthus* habitat suitability.

Scale Considerations

One of spatial ecology's fundamental quandaries is reconciling information obtained at disparate spatial scales into a common resolution for analysis (Openshaw 1983, Turner et al. 1989, Wiens 1989, Levin 1992, Wu 2004). There is often a large degree of variation in the spatial resolutions of the datasets readily available for

analysis. Model resolution is limited by the coarsest dataset, the locational uncertainty of FIA species occurrences in this instance, and predictor variables with finer grains (i.e., smaller pixel size) must be downscaled to coarser scales. A 300-m pixel size was selected for modeling as an optimal compromise between fine-grain land cover and topography data, coarse bioclimatic data, and the spatial uncertainty of FIA plot locations. In the case of the variable for distance to development (devdist), the NLCD source contains four separate classes of development intensity at a 30-m native resolution. To generate a variable suitable for modeling, the thematic classes were extracted and dissolved into a unified mask of urban areas, the distance from each cell in the study area to the nearest cell of the urban mask was calculated, the resulting 30-m distance raster was aggregated to 300 m, and lastly an 800-m radius moving average was taken to reflect the perturbation of FIA plot locations.

Information is inevitably lost or altered with each manipulation, whether it be the intensity of development or the complexity of the wildland-urban interface (Turner et al. 1989), resulting in a distorted representation of the underlying regional patterns the variable strives to reflect. Features within the landscape, such as transportation corridors or ridges, may be exaggerated or suppressed depending on the algorithms used (Wu 2004). The order in which operations are performed can have a dramatic effect; aggregating the urban areas to 300 m before calculating distance results in a substantially different product. Transforming predictor variables to new scales requires a clear understanding of the operations applied to the data, as well as the ecological process the output is attempting to characterize.

However, scale issues may arise even when transformations are carefully managed to minimize distortions, as the underlying ecological processes may themselves be dependent on the scale of observation (Wiens 1989, Levin 1992, Guisan and Thuiller 2005). For instance, the observed patterns of light shift radically when moving from the perspective of a mite amid topsoil to a raptor circling high above the landscape. While the pattern of light can be accurately measured throughout the intervening scales, only a limited domain is relevant to the canopy cover processes that influence *Ailanthus* establishment (Austin and Van Niel 2011). A related consideration is the extent of the study area, as mentioned previously. Continental-scale distributions are typically driven by broad climatic patterns, but have little predictive power on localized models, where variations in topography and land cover variables exert far more influence (Guisan and Zimmermann 2000, Elith and Leathwick 2009).

The poor performance of the finer-scaled predictors is likely due to a combination of these factors. The model's coarse scale may lie beyond the domain where observed patterns reflect the ecological processes relevant to *Ailanthus*. The patterns of vegetation cover, soil saturation, and human disturbance within one 300-m pixel can vary widely, and do so at scales that are likely to influence *Ailanthus*. Coupled with distortions from downscaling and FIA spatial error, it's unsurprising that incorporating land cover variables into the final model decreased its overall performance. The Maxent distribution is fit to training data misrepresenting the actual conditions, referred to as forced-matching (Guisan et al. 2007). The drop in

performance would likely be more pronounced if the extent was reduced while retaining the coarse grain.

Ultimately, the intended purpose of a model should determine its scale within the limitations imposed by the spatial accuracy of the species observation data (Guisan and Thuiller 2005, Austin 2007). While increasing model resolution may reveal fine scale ecological processes, a priority of this study was to examine the potential influence of climate change. To that end, the broad extent and coarse-grained predictor variables used for this model were well-suited to investigate the influence of extensive environmental gradients on *Ailanthus* habitats within the A.T.-shell.

Limitations

This distribution model makes significant assumptions by relying primarily on climate data. While the importance of biological, cultural, and topographic features is evident from the literature (Kota et al. 2006, Landenberger et al. 2009, Fryer 2010) and analysis of *Ailanthus* FIA records; these variables proved difficult to implement within the Maxent model for several reasons. Circumventing these issues by increasing the resolution of variables and acquiring locationally-accurate plot records would allow the model to discriminate suitable patches within the broad regions predicted by this model. Unfortunately, model resolution is constrained by current hardware and software performance limitations (particularly across a study area as expansive as the A.T.) and federal privacy regulations precluded the use of the true FIA plot locations. However, hierarchical modeling approaches may hold the key to integrating both small and large scale processes across broad spatial extents, and are

currently an active topic of discussion (Pearson et al. 2004, Guisan and Thuiller 2005, Jones et al. 2010).

Projections of the distribution model are similarly impaired by their lack of biological (Araújo and Luoto 2007) and anthropogenic interactions. Land-cover change will affect *Ailanthus* dispersal pathways and establishment opportunities (Dale 1997), but specific patterns are difficult to predict with certainty. Biological interactions, such as interspecific competition, may also limit invasion. However, the hardy traits of *Ailanthus* (e.g., rapid growth, high fecundity, and robust environmental tolerances) suggest its competitive advantage over native species will only increase as climate change alters the frequency and magnitude of disturbances (Dukes and Mooney 1999, Dale et al. 2001). While incorporating more interactions would augment this model, a climatic projection is suitable for an initial investigation of potential shifts in the distribution of suitable *Ailanthus* habitats.

Bioclimatic variables with projected values extending beyond the range of current values encountered while training the Maxent model further complicate predictions (Figure 21, Figure 22). In these cases, the model must either extrapolate features beyond the range they were parameterized within or ‘clamp’ the bioclimatic values and hold them constant at the upper limit of current conditions (Figure 23). Expanding the study area to incorporate a broader range of conditions would partially mitigate this issue, but only at the cost of decreasing the model’s ability to discriminate within the A.T.-shell (Elith et al. 2010).

Additional methods and independent test data are needed to further validate the model. The reliability of using the area under curve (AUC) of the receiving operator

curve (ROC) (Figure 18) to assess the accuracy of presence only species distribution models has been questioned (Lobo et al. 2008). However, no clear alternative metric has emerged, and the issue remains an active area of discussion within the modeling community (Warren and Seifert 2011). Independent test data, *Ailanthus* presence records other than FIA data, would provide a more robust evaluation of the model's predictive performance. While several alternative databases do exist, they lack the regional coverage or spatial accuracy of FIA. Similarly, comparing Maxent distribution models with additional modeling techniques or constructing model ensembles is increasingly prevalent (Araujo and New 2007, Stohlgren et al. 2010), and may provide valuable insight on the behavior and accuracy of predicting the suitable habitats of *Ailanthus*.

Conclusion

Ailanthus is a problematic invasive with important implications for forest health and landscape dynamics within the A.T.-shell. The model projection indicates the potential extent of *Ailanthus* invasion will increase significantly as the climate changes. Mapping the distribution of suitable habitats facilitates a quantitative assessment of the potential impacts of *Ailanthus* on biodiversity and ecosystem services within the Appalachian Trail corridor. In particular, further investigation is needed to determine how the biological communities of sensitive high elevation areas and northern forests will be affected by the introduction of novel competitors.

This habitat suitability model successfully integrates the resources of the A.T.-DSS to select a set of environmental variables that define *Ailanthus* habitat suitability, map the estimated current distribution of suitable habitats, and examine the potential

effects of climate change on biological invasion. The FIA database provided accurate, abundant, and detailed ground observations of *Ailanthus* populations. Although the relatively coarse grain of the model may have obscured some relationships, geospatial data proved to be a valuable tool for determining the environmental factors restricting the range of *Ailanthus*. In particular, the seamless climate data products provided by TOPS were a powerful and accessible resource. As a prototype application of the A.T.-DSS, this research demonstrates the utility of coupling *in situ* and geospatial data with innovative statistical techniques to investigate important ecological processes within the landscape. This modeling approach establishes a framework that can be effectively adopted to examine the distribution of additional important species in the region and inform efforts to conserve natural resources within the Appalachian Trail Corridor.

TABLES

Table 1: Geospatial Data Sources

Source agency, description, format, resolution, and retrieval location of geospatial datasets used in this study.

Dataset	Source	Description	Format	Resolution	Location
A.T. Centerline	ATC	Representation of the Appalachian Trail compiled from GPS ground survey.	Line	n/a	http://www.appalachiantrail.org/
Watershed Boundary Dataset	USGS	Hydrologic unit delineations at the HUC-10 level.	Polygon	n/a	http://nhd.usgs.gov/data.html
A.T.-shell	NPS	Study area comprised of subset HUC-10 watersheds intersecting the A.T. centerline.	Polygon	n/a	Derived
Forest Inventory & Analysis	USDA FS	Point shapefile adapted from the FIA Access Database, subset encompassing only plots where <i>Ailanthus</i> has been observed.	Point	n/a	http://apps.fs.fed.us/fiadb-downloads/datamart.html
Digital Elevation Model	USGS	Surface elevation from the National Elevation Dataset.	Raster	30m	http://edna.usgs.gov/
National Land Cover Database	MRLC	2006 land cover classifications.	Raster	30m	http://www.mrlc.gov/
Current and Projected Climate Conditions	TOPS	Current conditions reflect 1950-2005 averages. Projections are downscaled ensemble averages for 2095-2099 from the IPCC AR5 CIMP5.	Raster	250m	http://ecocast.arc.nasa.gov/
USSoils	USGS	Soils data compiled from STATSGO	Polygon	n/a	http://water.usgs.gov/GIS/metadata/usgswr/d/XML/ussoils.xml

Table 2: Environmental Variables

List of environmental variables evaluated for habitat modeling including abbreviation, title, units, and source dataset.

Variable	Title	Units	Source
agdist	Distance to Agricultural Landcover	m	NLCD06
agsum	Sum of Agricultural Landcover	%	NLCD06
bio1	Annual Mean Temperature	°C	TOPS
bio2	Mean Diurnal Range (Mean of monthly (max temp - min temp))	°C	TOPS
bio3	Isothermality (bio2/bio7) (* 100)	%	TOPS
bio4	Temperature Seasonality (standard deviation *100)	$1 \times 10^{-3} \text{ }^{\circ}\text{C}$	TOPS
bio5	Max Temperature of Warmest Month	°C	TOPS
bio6	Min Temperature of Coldest Month	°C	TOPS
bio7	Temperature Annual Range (bio5-bio6)	°C	TOPS
bio8	Mean Temperature of Wettest Quarter	°C	TOPS
bio9	Mean Temperature of Driest Quarter	°C	TOPS
bio10	Mean Temperature of Warmest Quarter	°C	TOPS
bio11	Mean Temperature of Coldest Quarter	°C	TOPS
bio12	Annual Precipitation	$1 \times 10^{-4} \text{ m}$	TOPS
bio13	Precipitation of Wettest Month	$1 \times 10^{-4} \text{ m}$	TOPS
bio14	Precipitation of Driest Month	$1 \times 10^{-4} \text{ m}$	TOPS
bio15	Precipitation Seasonality (Coefficient of Variation)	%	TOPS
bio16	Precipitation of Wettest Quarter	$1 \times 10^{-4} \text{ m}$	TOPS
bio17	Precipitation of Driest Quarter	$1 \times 10^{-4} \text{ m}$	TOPS
bio18	Precipitation of Warmest Quarter	$1 \times 10^{-4} \text{ m}$	TOPS
bio19	Precipitation of Coldest Quarter	$1 \times 10^{-4} \text{ m}$	TOPS

Variable	Title	Units	Source
cti	Compound Topographic Index	n/a	DEM
dem	Elevation	m	DEM
devdist	Distance to Developed Landcover	m	NLCD06
devsum	Sum of Developed Landcover	%	NLCD06
hli	Heat Load Index	n/a	DEM
lfcc	Canopy Cover Mean	%	LANDFIRE
lfcc_min	Canopy Cover Minimum	%	LANDFIRE
lfcc_std	Canopy Cover Standard Deviation	%	LANDFIRE
nlcd_wet	Wetland and Open Water Landcover	0-1	NLCD06
slope	Slope	°	DEM
slopepos	Slope Position (800-m radius focal mean elevation – elevation)	m	DEM
soil_drain	Soil Drainage Class	0-7	USSOILS
soil_fldfreq	Flood Frequency	0-4	USSOILS
soil_hydric	Hydric Soils	0-1	USSOILS
trasp	Topographic Radiation Aspect Index	0-1	DEM

(Table 2 Continued)

Table 3: Correlation of Topographic and Landcover Variables

Matrix of Pearson's correlation coefficients calculated for the full array of environmental variables. Red and blue shading indicate positive and negative correlation, respectively, with intensity increasing with the degree of correlation.

	trasp	agdist_mean	agsum	cti	dem	devdist	devsum	hli	lfcc_mean	lfcc_min	lfcc_std	nlcd_wet	slope	slopepos	soil_drain	soil_fldfreq	soil_hydric
trasp		0.00	0.02	0.01	-0.01	-0.01	0.01	0.01	-0.02	-0.02	0.01	0.00	-0.03	0.00	0.01	0.01	0.00
agdist			-0.21	0.09	0.04	0.54	-0.14	-0.02	0.09	0.08	-0.27	0.16	-0.04	0.11	0.22	-0.13	0.43
agsum				0.23	-0.28	-0.23	0.03	-0.29	-0.78	-0.43	0.31	-0.09	-0.40	-0.27	-0.01	0.00	-0.09
cti					-0.38	-0.03	0.17	-0.50	-0.44	-0.39	0.16	0.39	-0.49	-0.37	-0.01	-0.31	0.11
dem						0.16	-0.23	0.50	0.43	0.43	-0.40	-0.20	0.68	0.77	-0.19	0.13	-0.26
devdist							-0.23	0.12	0.20	0.25	-0.38	0.08	0.15	0.22	0.14	-0.06	0.28
devsum								-0.20	-0.44	-0.34	0.18	-0.04	-0.25	-0.21	0.01	0.00	-0.06
hli									0.43	0.46	-0.29	-0.18	0.74	0.45	-0.15	0.12	-0.18
lfcc_mean										0.66	-0.37	-0.17	0.57	0.40	0.06	0.25	0.02
lfcc_min											-0.43	-0.18	0.55	0.43	-0.05	0.14	-0.07
lfcc_std												0.07	-0.42	-0.46	-0.10	-0.13	-0.12
nlcd_wet													-0.21	-0.15	-0.06	-0.37	0.13
slope														0.62	-0.19	0.17	-0.23
slopepos															-0.03	0.15	-0.13
soil_drain																0.51	0.74
soil_fldfreq																	0.12
soil_hydric																	

Table 4: Correlation of Bioclimatic Variables

Matrix of Pearson's correlation coefficients calculated for the full array of environmental variables. Red and blue shading indicate positive and negative correlation, respectively, with intensity increasing with the degree of correlation.

	bio1	bio2	bio3	bio4	bio5	bio6	bio7	bio8	bio9	bio10	bio11	bio12	bio13	bio14	bio15	bio16	bio17	bio18	bio19
bio1		0.42	0.78	-0.83	0.89	0.98	-0.80	0.35	0.78	0.95	0.98	0.10	0.13	-0.07	0.21	0.15	0.05	0.06	0.22
bio2			0.74	-0.43	0.44	0.34	-0.18	-0.03	0.53	0.35	0.44	0.07	0.14	-0.17	0.36	0.15	-0.02	0.05	0.23
bio3				-0.92	0.54	0.80	-0.79	0.05	0.88	0.58	0.85	0.36	0.43	0.07	0.40	0.44	0.26	0.35	0.51
bio4					-0.50	-0.90	0.96	-0.10	-0.88	-0.60	-0.91	-0.44	-0.48	-0.20	-0.30	-0.50	-0.37	-0.43	-0.54
bio5						0.79	-0.45	0.45	0.53	0.98	0.80	-0.20	-0.18	-0.28	0.09	-0.15	-0.21	-0.25	-0.08
bio6							-0.90	0.32	0.81	0.88	0.99	0.20	0.22	0.03	0.20	0.24	0.15	0.17	0.30
bio7								-0.14	-0.81	-0.59	-0.88	-0.43	-0.46	-0.24	-0.22	-0.47	-0.38	-0.43	-0.50
bio8									-0.06	0.45	0.28	-0.55	-0.55	-0.56	0.09	-0.52	-0.56	-0.35	-0.58
bio9										0.61	0.84	0.44	0.51	0.15	0.41	0.52	0.34	0.38	0.61
bio10											0.87	-0.11	-0.09	-0.21	0.12	-0.07	-0.13	-0.16	0.00
bio11												0.21	0.25	0.02	0.24	0.27	0.16	0.18	0.33
bio12													0.97	0.85	0.05	0.98	0.96	0.93	0.96
bio13														0.76	0.24	0.99	0.89	0.91	0.95
bio14															-0.40	0.75	0.95	0.72	0.75
bio15																0.24	-0.21	0.19	0.16
bio16																	0.89	0.92	0.97
bio17																		0.85	0.89
bio18																			0.85
bio19																			

Table 5: Environmental Variables included in Candidate Maxent Models

Matrix of the environmental variables incorporated into each candidate habitat suitability model. Model names are derived from the number of bioclimatic, topographic, and land-cover variables selected. For descriptions of the variables, see Table 2.

[illegible]

Table 6: Parameters and Evaluation of Candidate Maxent Models

Models were evaluated using a variety of methods. Sample size corrected Akaike information criteria (AIC_c) is determined from model log likelihood (the product of suitability scores across all presence points) penalized by the number of parameters (the complexity of features applied to the environmental variables). Models with lower AIC_c scores balance high predictive performance with low complexity, and are likely more appropriate for extrapolating to future conditions. The area under the curve (AUC) of the receiver operating characteristics (ROC) provides a threshold independent measure of model predictive performance on withheld test data. The values listed below are averaged across the 10 replications performed for each model.

Model	Log Likelihood	Parameters	AICc Score	Mean Test AUC	AUC SD
4bio_4topo	-1751.24	56.8	3707.533	0.85	0.034
2bio_1topo	-1796.72	40.9	3713.173	0.812	0.035
4bioalt_4topo_2lc	-1749.24	61.8	3733.685	0.847	0.034
5bioalt_5topo	-1747.56	68.2	3777.135	0.848	0.045
10bioalt_6topo	-1735.15	72.2	3788.426	0.855	0.047
5bioalt_4topo	-1747.74	69.8	3794.235	0.849	0.044
4bioalt_3topo	-1752.98	67.8	3796.161	0.85	0.041
5bioalt2_3topo	-1750.84	69.4	3804.34	0.852	0.04
6bioalt_6topo	-1743.5	73.1	3817.243	0.848	0.046
allbio_alltopo	-1732.14	76.2	3821.316	0.847	0.039
5bio_5topo	-1751.37	74.4	3862.39	0.848	0.045
10bio_5topo	-1739.94	79.4	3886.28	0.851	0.048
4bio_5topo	-1752.8	76.9	3886.495	0.842	0.046
10bio_5topo_4lc	-1734.39	88.3	3997.353	0.844	0.046
allbio_alltopo_alllc	-1722.52	92.5	4079.972	0.844	0.049

Table 7: Maxent Model Variable Contributions

The relative contributions of the environmental variables for the selected model. Percent contribution is calculated as the training algorithm iterates by adding the increase in regularized gain or subtracting if the absolute value of lambda is negative. Permutation importance is determined by randomly changing the values of the focal variable for the training and background data, reevaluating the model with each permuted variable in turn, and recording the corresponding drop in training AUC normalized to a percentage. Interpretation of variable contributions may be confounded by high correlation.

Variable	Percent contribution	Permutation importance
bio11	40.1	34.7
bio4	27.1	43.9
bio8	22.7	6
slope	4.6	5.1
trasp	3.1	2
nlcd_wet	0.9	0.7
bio19	0.8	2.4
dem	0.6	5.3

Table 8: Change in Bioclimatic Variables from Baseline to Projected Conditions

Changes from the 1950-2005 baseline to the CIMP5 RCP6.0 2090-2095 climate projection for bioclimatic variables included in the final model.

Variable	Min			Max			SD			Mean		
	Current	RCP60	Change	Current	RCP60	Change	Current	RCP60	Change	Current	RCP60	Change
bio4 (1 x 10 ⁻³ °C)	665.325	707.27	41.945	1142.73	1099.58	-43.15	110.85	82.21	-28.64	896.67	900.51	3.84
bio8 (°C)	-8.8	-5.57	3.23	23.84	27.8	3.96	6.6	8.31	1.71	14.95	17.13	2.18
bio11 (°C)	-14.32	-8.92	5.4	6.21	9.22	3.01	4.59	3.82	-0.77	-2.13	2.16	4.29
bio19 (1 x 10 ⁻⁴ m)	2038	2479	441	9310	9679	369	894	850	-44	3286	3818	532

Table 9: Maxent Model Thresholds

Common thresholds applied to the continuous probability distribution output by Maxent to derive binary classes of predicted suitability. Fractional predicted area is the fraction of the total study area classified suitable, while training omission rate is the fraction of training points (presences) predicted as unsuitable. The “Balance” threshold was applied to the selected model.

Cumulative threshold	Logistic threshold	Description	Fractional predicted area	Training omission rate
1	0.04	Fixed cumulative value 1	0.577	0
5	0.16	Fixed cumulative value 5	0.427	0.051
10	0.234	Fixed cumulative value 10	0.35	0.11
1.208	0.047	Minimum training presence	0.56	0
8.217	0.212	10 percentile training presence	0.374	0.096
24.111	0.353	Equal training sensitivity and specificity	0.221	0.221
19.547	0.317	Maximum training sensitivity plus specificity	0.255	0.176
1.208	0.047	Balance training omission, predicted area and threshold value	0.56	0
7.715	0.204	Equate entropy of thresholded and original distributions	0.381	0.088

Table 10: Current and Projected Area of Suitable Habitats

Change in predicted suitable area from the 1950-2005 baseline to the CIMP5 RCP6.0 2090-2095 climate projection. Ecoregion provinces in grey italics incorporated only a small portion of the trail and were not interpreted. See Figure 2 for the locations of Bailey's ecoregion provinces intersecting the study area.

Province	Suitable Area (km ²)						
	Total Area	Current	%	Projected	%	Change	%
<i>211</i>	<i>4478</i>	<i>1283</i>	<i>28.7%</i>	<i>2372</i>	<i>53.0%</i>	<i>1089</i>	<i>84.8%</i>
221	20013	17211	86.0%	19802	98.9%	2591	15.1%
<i>231</i>	<i>2831</i>	<i>1577</i>	<i>55.7%</i>	<i>2824</i>	<i>99.7%</i>	<i>1247</i>	<i>79.1%</i>
M211	29746	624	2.1%	14969	50.3%	14345	2298.3%
M221	51004	39348	77.1%	49098	96.3%	9750	24.8%
A.T.-shell	108072	60044	55.6%	89066	82.4%	29022	48.3%

Table 11: Current and Projected Elevation and Latitude of Suitable Habitats

Change in the elevation and latitude of predicted suitable area from the 1950-2005 baseline to the CIMP5 RCP6.0 2090-2095 climate projection. Ecoregion provinces in grey italics incorporated only a small portion of the trail and were not interpreted. See Figure 2 for the locations of Bailey's ecoregion provinces intersecting the study area.

Province	Mean Elevation (meters)				Mean Latitude			
	Current	Projected	Change	%	Current	Projected	Change (deg)	Change (km)
<i>211</i>	<i>289</i>	<i>414</i>	<i>125</i>	<i>43.3%</i>	<i>41.26</i>	<i>41.41</i>	<i>0.14</i>	<i>17</i>
221	195	201	6	3.0%	40.31	40.63	0.31	36
<i>231</i>	<i>267</i>	<i>348</i>	<i>81</i>	<i>30.1%</i>	<i>37.17</i>	<i>36.02</i>	<i>-1.15</i>	<i>-128</i>
M211	340	436	96	28.2%	42.49	43.46	0.97	108
M221	491	561	70	14.2%	38.23	37.68	-0.55	-61
A.T.-shell	391	449	59	15.1%	38.91	39.35	0.44	49

FIGURES

Figure 1: A.T. Hydrologic Unit Code (HUC)-10 Shell

Ten-digit Hydrological Unit Code (HUC-10) watersheds within 5 statute miles of the A.T. land base, termed the A.T. HUC-10 shell.

Figure 2: Ecological Provinces of the A.T.

Bailey's ecoregions are a hierarchical classification system which groups areas with similar climates and dominant potential vegetation (Bailey 1998).

Figure 3: Perturbed Forest Inventory & Analysis (FIA) Plot Distribution

The Forest Inventory and Analysis (FIA) program of the USDA Forest Service provides a detailed, systematic record of forest vegetation. Between 2002 and 2010, 3,926 plots were surveyed and *Ailanthus* was observed at 136 locations.

Figure 4: Elevation

Digital elevation model derived from the 30-meter National Elevation Dataset (NED) produced by the USGS.

Figure 5: Distance to Developed Landcover

Distance from the focal pixel to the nearest developed area derived from the National Land Cover Database (NLCD).

Figure 6: Wetland and Open Water

The distribution of wetland and open water areas derived from the National Land Cover Database (NLCD).

Figure 7: Canopy Cover

The distribution of canopy cover derived from the National Land Cover Database (NLCD).

Figure 8: CMIP5 Representative Concentration Pathways

The Coupled Model Intercomparison Project Phase 5 (CMIP5) is an ensemble of 16 individual General Circulation Models (GCMs) that predict future conditions under a set of alternative scenarios defined by Representative Concentration Pathways (RCPs). RCPs represent the atmospheric concentration of greenhouse gasses, or radiative forcing values, in the year 2100 resulting from future scenarios with varying levels of global emissions and mitigation. RCP6.0 was selected for *Ailanthus* modeling, as it represents a moderate increase in radiative forcing that stabilizes by 2100 due to technologies and strategies for reducing greenhouse gas emissions.

<https://upload.wikimedia.org/wikipedia/en/3/3e/All_forcing_agents_CO2_equivalent_concentration.png>

Figure 9: Annual Mean Temperature

Annual mean temperature from the 1950-2005 baseline climate data.

Figure 10: Annual Precipitation

Annual precipitation from the 1950-2005 baseline climate data.

Figure 11: Distribution of Projected Changes in Temperature

Change in annual mean temperature from the 1950-2005 baseline to the CIMP5 RCP6.0 2090-2095 climate projection.

Figure 12: Distribution of Projected Changes in Precipitation

Change in annual precipitation from the 1950-2005 baseline to the CIMP5 RCP6.0 2090-2095 climate projection.

Figure 13: Forest Inventory & Analysis (FIA) Histograms, Boxplots, and T-tests
Statistical comparison of presence (red) and absence (green) sites using attributes extracted from the FIA database.

Figure 14: Environmental Variable Histograms, Boxplots, and T-tests
Statistical comparison of presence (red) and absence (green) sites using values extracted from the environmental variable rasters. See Table 2 for units.

Welch Two Sample t-test

data: FIAvar\$Presence and FIAvar\$Absence
 $t = -7.7421$, $df = 166.922$, $p\text{-value} = 8.936e-13$
 alternative hypothesis: true difference in means is not equal to 0
 95 percent confidence interval:
 -1169.969 -694.515
 sample estimates:
 mean of x mean of y
 13640.91 14573.15

Welch Two Sample t-test

data: FIAvar\$Presence and FIAvar\$Absence
 $t = -0.1032$, $df = 143.776$, $p\text{-value} = 0.918$
 alternative hypothesis: true difference in means is not equal to 0
 95 percent confidence interval:
 -0.2973748 0.2678568
 sample estimates:
 mean of x mean of y
 7.033681 7.048440

Welch Two Sample t-test

data: FIAvar\$Presence and FIAvar\$Absence
 $t = -7.0128$, $df = 148.891$, $p\text{-value} = 7.61e-11$
 alternative hypothesis: true difference in means is not equal to 0
 95 percent confidence interval:
 -173.1255 -97.0086
 sample estimates:
 mean of x mean of y
 352.5009 487.5680

Welch Two Sample t-test

data: FI Δ var $\{$ Presence and FI Δ var $\{$ Absence
 $t = -13.4377$, $df = 340.974$, $p\text{-value} < 2.2e-16$
 alternative hypothesis: true difference in means is not equal to 0
 95 percent confidence interval:
 -511.0399 -380.5352
 sample estimates:
 mean of x mean of y
 323.7922 769.5798

Welch Two Sample t-test

data: FI Δ var $\{$ Presence and FI Δ var $\{$ Absence
 $t = 0.2479$, $df = 143.923$, $p\text{-value} = 0.8046$
 alternative hypothesis: true difference in means is not equal to 0
 95 percent confidence interval:
 -46.28159 59.55341
 sample estimates:
 mean of x mean of y
 5933.110 5926.474

Welch Two Sample t-test

data: FI Δ var $\{$ Presence and FI Δ var $\{$ Absence
 $t = -3.9815$, $df = 160.142$, $p\text{-value} = 0.0001037$
 alternative hypothesis: true difference in means is not equal to 0
 95 percent confidence interval:
 -0.07858288 -0.02647302
 sample estimates:
 mean of x mean of y
 0.02205882 0.07458678

Welch Two Sample t-test

data: FI\Avar\Presence and FI\Avar\Absence
 $t = -0.846$, $df = 143.654$, $p\text{-value} = 0.399$
 alternative hypothesis: true difference in means is not equal to 0
 95 percent confidence interval:
 -0.9100366 0.3645214
 sample estimates:
 mean of x mean of y
 5.267024 5.539782

Welch Two Sample t-test

data: FI\Avar\Presence and FI\Avar\Absence
 $t = -4.1869$, $df = 146.293$, $p\text{-value} = 4.863e-05$
 alternative hypothesis: true difference in means is not equal to 0
 95 percent confidence interval:
 -95.30333 -34.18273
 sample estimates:
 mean of x mean of y
 -63.9782093 0.7648207

Welch Two Sample t-test

data: FI\Avar\Presence and FI\Avar\Absence
 $t = -3.2828$, $df = 149.52$, $p\text{-value} = 0.001280$
 alternative hypothesis: true difference in means is not equal to 0
 95 percent confidence interval:
 -0.21894019 -0.05440891
 sample estimates:
 mean of x mean of y
 3.028676 3.165351

Welch Two Sample t-test

data: FIavar6Presence and FIavar6Absence
 $t = -3.8235$, $df = 143.906$, $p\text{-value} = 0.0001953$
 alternative hypothesis: true difference in means is not equal to 0
 95 percent confidence interval:
 -0.04911289 -0.01563917
 sample estimates:
 mean of x mean of y
 0.02500000 0.05737603

Welch Two Sample t-test

data: FIavar6Presence and FIavar6Absence
 $t = 2.0917$, $df = 143.835$, $p\text{-value} = 0.03822$
 alternative hypothesis: true difference in means is not equal to 0
 95 percent confidence interval:
 0.001736324 0.061371207
 sample estimates:
 mean of x mean of y
 0.5397168 0.5081630

Figure 15: Marginal Variable Response Curves

Marginal variable response curves plot the change in logistic prediction from varying the value of one variable while holding all other variables constant at their average sample values. Strongly correlated variables may confound the interpretation of marginal response curves. See Table 2 for variable units.

Figure 16: Isolated Variable Response Curves

Isolated variable response curves represent a model incorporating only the focal variable and may be more informative when dealing with highly correlated variables. See Table 2 for variable units.

Figure 17: Variable Jackknifes

The jackknife tests compare the regularized training gain, test gain, and test AUC for a set of models created while withholding each variable in turn and with each variable in isolation.

Figure 18: Model Performance (AUC)

The area under the curve (AUC) of the receiver operating characteristics (ROC) provides a threshold independent measure of model predictive performance on withheld test data. The ROC curve is constructed by plotting model sensitivity and specificity. Sensitivity is a function of the omission rate, i.e. the rate training or test presence points incorrectly classified as unsuitable. Specificity is typically the commission rate, or rate of absence points incorrectly classified as suitable. However, given the lack of absence data for presence only modeling, specificity is instead derived from the fraction of the study area predicted as suitable. The mean and standard deviation reflect the variation in the 10-fold Maxent model replicates

Figure 19: Maxent Distribution for Current Conditions, Continuous

The continuous Maxent probability distribution of *Ailanthus* habitat suitability for model 4bio_4topo under current conditions. Green areas are predicted to have low suitability, while red areas indicate high suitability.

Figure 20: Maxent Distribution for Projected Conditions, Continuous

The continuous Maxent probability distribution of *Ailanthus* habitat suitability for model 4bio_4topo under projected climate conditions. Green areas are predicted to have low suitability, while red areas indicate high suitability.

Figure 21: Multivariate Similarity Surface (MESS)

The multivariate similarity surface calculates how similar the projected bioclimatic variables are to the baseline variables. Negative values indicate at least one projected variable is outside the range conditions, and are therefore novel conditions.

Figure 22: Most Dissimilar Variable (MoD)

The most dissimilar variable indicates which projected bioclimatic variable is furthest outside the range of current conditions.

Figure 23: Effects of Variable Clamping on Projection

Variable clamping restricts the projected variables to the range of values encountered in the current climate values. This surface reflects the absolute difference in predictions when using vs. not using clamping. Higher values indicate clamping has had a larger effect on predicted suitable.

Figure 24: Maxent Distribution for Current Conditions, Binary

Binary suitability produced by applying the “balance” threshold to the Maxent probability distribution of *Ailanthus* habitat suitability for model 4bio_4topo under current conditions. Green areas are predicted to have low suitability, while red areas indicate high suitability.

Figure 25: Maxent Distribution for Projected Conditions, Binary

Binary suitability produced by applying the “balance” threshold to the Maxent probability distribution of *Ailanthus* habitat suitability for model 4bio_4topo under projected climate conditions. Green areas are predicted to have low suitability, while red areas indicate high suitability.

Figure 26: Change in Suitable Habitats, Binary

Change in binary suitability from the 1950-2005 baseline to the CIMP5 RCP6.0 2090-2095 climate projection for Maxent model 4bio_4topo. Yellow areas are predicted suitable under current conditions, red areas are predicted to become suitable by the 2090s, while green areas will remain poorly suitable.

APPENDICES

Appendix I: Maps of All Environmental Variables

1.	Distance to Agricultural Landcover (agdist)	82
2.	Sum of Agricultural Landcover (agsum)	83
3.	Compound Topographic Index (cti)	84
4.	Annual Mean Temperature (bio1)	85
5.	Mean Diurnal Range (bio2)	86
6.	Isothermality (bio3)	87
7.	Temperature Seasonality (bio4).....	88
8.	Max Temperature of Warmest Month (bio5)	89
9.	Min Temperature of Coldest Month (bio6)	90
10.	Temperature Annual Range (bio7).....	91
11.	Mean Temperature of Wettest Quarter (bio8).....	92
12.	Mean Temperature of Driest Quarter (bio9)	93
13.	Mean Temperature of Warmest Quarter (bio10)	94
14.	Mean Temperature of Coldest Quarter (bio11).....	95
15.	Annual Precipitation (bio12).....	96
16.	Precipitation of Wettest Month (bio13)	97
17.	Precipitation of Driest Month (bio14).....	98
18.	Precipitation Seasonality (bio15).....	99
19.	Precipitation of Wettest Quarter (bio16).....	100
20.	Precipitation of Driest Quarter (bio17)	101
21.	Precipitation of Warmest Quarter (bio18).....	102
22.	Precipitation of Coldest Quarter (bio19).....	103
23.	Elevation (dem).....	104
24.	Distance to Developed Landcover (devdist)	105
25.	Sum of Developed Landcover (devsum).....	106
26.	Heat Load Index (hli)	107
27.	Canopy Cover Mean (lfcc).....	108
28.	Canopy Cover Minimum (lfcc_min).....	109
29.	Canopy Cover Standard Deviation (lfcc_std)	110
30.	Wetland and Open Water Landcover (nlcd_wet).....	111
31.	Slope (slope)	112
32.	Slope Position (slopepos).....	113
33.	Soil Drainage Class (soil_drain)	114
34.	Flood Frequency (soil_fldfreq)	115
35.	Hydric Soils (soil_hydric)	116
36.	Topographic Radiation Aspect Index (trasp)	117

1. Distance to Agricultural Landcover (agdist)

2. Sum of Agricultural Landcover (agsum)

3. Compound Topographic Index (cti)

4. Annual Mean Temperature (bio1)

5. Mean Diurnal Range (bio2)

6. Isothermality (bio3)

7. Temperature Seasonality (bio4)

8. Max Temperature of Warmest Month (bio5)

9. Min Temperature of Coldest Month (bio6)

10. Temperature Annual Range (bio7)

11. Mean Temperature of Wettest Quarter (bio8)

12. Mean Temperature of Driest Quarter (bio9)

13. Mean Temperature of Warmest Quarter (bio10)

14. Mean Temperature of Coldest Quarter (bio11)

15. Annual Precipitation (bio12)

16. Precipitation of Wettest Month (bio13)

17. Precipitation of Driest Month (bio14)

18. Precipitation Seasonality (bio15)

19. Precipitation of Wettest Quarter (bio16)

20. Precipitation of Driest Quarter (bio17)

21. Precipitation of Warmest Quarter (bio18)

22. Precipitation of Coldest Quarter (bio19)

23. Elevation (dem)

24. Distance to Developed Landcover (devdist)

25. Sum of Developed Landcover (devsum)

26. Heat Load Index (hli)

27. Canopy Cover Mean (lfcc)

28. Canopy Cover Minimum (Ifcc_min)

29. Canopy Cover Standard Deviation (Ifcc_std)

30. Wetland and Open Water Landcover (nlcd_wet)

31. Slope (slope)

32. Slope Position (slopepos)

33. Soil Drainage Class (soil_drain)

34. Flood Frequency (soil_fldfreq)

35. Hydric Soils (soil_hydric)

36. Topographic Radiation Aspect Index (trasp)

Appendix II: Maxent Log File

Fri Mar 22 19:37:52 EDT 2013

MaxEnt version 3.3.3k

Checking header of E:\maxent\vars\projection\bio11_300m.asc

Checking header of E:\maxent\vars\projection\bio19_300m.asc

Checking header of E:\maxent\vars\projection\bio4_300m.asc

Checking header of E:\maxent\vars\projection\bio8_300m.asc

Checking header of E:\maxent\vars\projection\dem_300m_atshell.asc

Checking header of E:\maxent\vars\projection\nlcd_wet_300m_atshell.asc

Checking header of E:\maxent\vars\projection\slope_800mn_300m_atshell.asc

Checking header of E:\maxent\vars\projection\trasp_800mn_300m_atshell.asc

Reading samples from ToH20120229_met_trimduplicates_exact.csv

Read samples: max memory 3817799680, total allocated 1071906816, free 552569392, used 519337424, increment 291726680

Extractor: max memory 3817799680, total allocated 1071906816, free 552560896, used 519345920, increment 8496

Extracting random background and sample data

Time since start: 11.828

1202652 points with values for all grids

Adding samples to background in feature space

Command line used: nowarnings noprefixes -E -E Ailanthus responsecurves
jackknife outputformat=raw

outputdirectory=E:\maxent\result\spring13\4bio_4topo_final

projectionlayers=E:\maxent\vars\projection\rcp60

samplesfile=E:\maxent\ToH20120229_met_trimduplicates_exact.csv

environmentallayers=E:\maxent\vars\projection nowarnings noaskoverwrite

replicates=10 nolinear noquadratic noproduct nothreshold noautofeature noprefixes -N

agdist_mean_300m_atshell -N agsum_300m_atshell -N bio10_300m -N bio12_300m

-N bio13_300m -N bio14_300m -N bio15_300m -N bio16_300m -N bio17_300m -N

bio18_300m -N bio1_300m -N bio2_300m -N bio3_300m -N bio5_300m -N

bio6_300m -N bio7_300m -N bio9_300m -N cti_300m_atshell -N

devdist_mean_300m_atshell -N devdist_p1ln_300m_atshell -N devsum_300m_atshell

-N hli_300m_atshell -N lfcc_mean_300m_atshell -N lfcc_min_300m_atshell -N

lfcc_std_300m_atshell -N slopepos_800rad_300m_atshell -N soil_drain_300m_atshell

-N soil fldfreq_300m_atshell -N soil_hydric_300m_atshell -t nlcd_wet_300m_atshell

Command line to repeat: java density.MaxEnt nowarnings noprefixes responsecurves

jackknife outputdirectory=E:\maxent\result\spring13\4bio_4topo_final

projectionlayers=E:\maxent\vars\projection\rcp60

samplesfile=E:\maxent\ToH20120229_met_trimduplicates_exact.csv

environmentallayers=E:\maxent\vars\projection nowarnings noaskoverwrite nolinear

noquadratic noproduct nothreshold noautofeature noprefixes -N

agdist_mean_300m_atshell -N agsum_300m_atshell -N bio10_300m -N bio12_300m

-N bio13_300m -N bio14_300m -N bio15_300m -N bio16_300m -N bio17_300m -N

bio18_300m -N bio1_300m -N bio2_300m -N bio3_300m -N bio5_300m -N

bio6_300m -N bio7_300m -N bio9_300m -N cti_300m_atshell -N

devdist_mean_300m_atshell -N devdist_p1ln_300m_atshell -N devsum_300m_atshell

-N hli_300m_atshell -N lfcc_mean_300m_atshell -N lfcc_min_300m_atshell -N
 lfcc_std_300m_atshell -N slopepos_800rad_300m_atshell -N soil_drain_300m_atshell
 -N soil fldfreq_300m_atshell -N soil_hydric_300m_atshell -t nlcd_wet_300m_atshell
 Species: Ailanthus
 Layers: bio11_300m bio19_300m bio4_300m bio8_300m dem_300m_atshell
 nlcd_wet_300m_atshell slope_800mn_300m_atshell trasp_800mn_300m_atshell
 Layertypes: Continuous Continuous Continuous Continuous Continuous Categorical
 Continuous Continuous
 responsecurves: true
 jackknife: true
 outputdirectory: E:\maxent\result\spring13\4bio_4topo_final
 projectionlayers: E:\maxent\vars\projection\rcp60
 samplesfile: E:\maxent\ToH20120229_met_trimduplicates_exact.csv
 environmentalayers: E:\maxent\vars\projection
 warnings: false
 askoverwrite: false
 linear: false
 quadratic: false
 product: false
 threshold: false
 autofeature: false
 prefixes: false

getSamples: max memory 3817799680, total allocated 1071906816, free 545010584,
 used 526896232, increment 7550312

Making features

makeFeatures: max memory 3817799680, total allocated 1071906816, free
 544964160, used 526942656, increment 46424

Ailanthus:

Regularization values: linear/quadratic/product: 0.050, categorical: 0.250, threshold:
 1.000, hinge: 0.500

136 samples

Density: max memory 3817799680, total allocated 1071906816, free 530741864, used
 541164952, increment 14222296

linearPredictor: max memory 3817799680, total allocated 1071906816, free
 530741864, used 541164952, increment 0

FeaturedSpace: max memory 3817799680, total allocated 1071906816, free
 530741864, used 541164952, increment 0

Sequential: max memory 3817799680, total allocated 1071906816, free 530741864,
 used 541164952, increment 0

Initial loss: 9.2237500588861

Time since start: 45.625

480: time = 33.156000 loss = 8.258853

Resulting gain: 0.9648973010358137

Projecting...

Writing file E:\maxent\result\spring13\4bio_4topo_final\Ailanthus.asc

Time since start: 79.266
 Writing E:\maxent\result\spring13\4bio_4topo_final\plots\Ailanthus.png
 Time since start: 87.688
 Projecting...
 Writing file E:\maxent\result\spring13\4bio_4topo_final\Ailanthus_rcp60.asc
 Writing file
 E:\maxent\result\spring13\4bio_4topo_final\Ailanthus_rcp60_clamping.asc
 Time since start: 130.391
 Writing E:\maxent\result\spring13\4bio_4topo_final\plots\Ailanthus_rcp60.png
 Time since start: 138.86
 Writing
 E:\maxent\result\spring13\4bio_4topo_final\plots\Ailanthus_rcp60_clamping.png
 Time since start: 147.235
 Writing file E:\maxent\result\spring13\4bio_4topo_final\Ailanthus_rcp60_novel.asc
 Writing file
 E:\maxent\result\spring13\4bio_4topo_final\Ailanthus_rcp60_novel_limiting.asc
 Time since start: 203.797
 Ailanthus response curves
 Time since start: 205.391
 Response curve: only bio11_300m
 Making features
 makeFeatures: max memory 3817799680, total allocated 1308295168, free 1018046384, used 290248784, increment -250916168
 Ailanthus bio11_300m:
 Regularization values: linear/quadratic/product: 0.050, categorical: 0.250, threshold: 1.000, hinge: 0.500
 136 samples
 Density: max memory 3817799680, total allocated 1308295168, free 1018029080, used 290266088, increment 17304
 linearPredictor: max memory 3817799680, total allocated 1308295168, free 1018029080, used 290266088, increment 0
 FeaturedSpace: max memory 3817799680, total allocated 1308295168, free 1018029080, used 290266088, increment 0
 Sequential: max memory 3817799680, total allocated 1308295168, free 1018029080, used 290266088, increment 0
 Initial loss: 9.2237500588861
 Time since start: 209.75
 320: time = 4.328000 loss = 8.741746
 Resulting gain: 0.4820042634993431
 Ailanthus response curves
 Response curve: only bio19_300m
 Making features
 makeFeatures: max memory 3817799680, total allocated 1287847936, free 864493128, used 423354808, increment 133088720
 Ailanthus bio19_300m:

Regularization values: linear/quadratic/product: 0.050, categorical: 0.250, threshold:
 1.000, hinge: 0.500
 136 samples
 Density: max memory 3817799680, total allocated 1287847936, free 864466248, used
 423381688, increment 26880
 linearPredictor: max memory 3817799680, total allocated 1287847936, free
 864466248, used 423381688, increment 0
 FeaturedSpace: max memory 3817799680, total allocated 1287847936, free
 864466248, used 423381688, increment 0
 Sequential: max memory 3817799680, total allocated 1287847936, free 864466248,
 used 423381688, increment 0
 Initial loss: 9.2237500588861
 Time since start: 212.547
 180: time = 2.531000 loss = 9.117629
 Resulting gain: 0.10612091163523019
 Ailanthus response curves
 Response curve: only bio4_300m
 Making features
 makeFeatures: max memory 3817799680, total allocated 1267990528, free
 1018765240, used 249225288, increment -174156400
 Ailanthus bio4_300m:
 Regularization values: linear/quadratic/product: 0.050, categorical: 0.250, threshold:
 1.000, hinge: 0.500
 136 samples
 Density: max memory 3817799680, total allocated 1267990528, free 1018740768,
 used 249249760, increment 24472
 linearPredictor: max memory 3817799680, total allocated 1267990528, free
 1018740768, used 249249760, increment 0
 FeaturedSpace: max memory 3817799680, total allocated 1267990528, free
 1018740768, used 249249760, increment 0
 Sequential: max memory 3817799680, total allocated 1267990528, free 1018740768,
 used 249249760, increment 0
 Initial loss: 9.2237500588861
 Time since start: 217.25
 320: time = 4.468000 loss = 8.407032
 Resulting gain: 0.8167185188093811
 Ailanthus response curves
 Response curve: only bio8_300m
 Making features
 makeFeatures: max memory 3817799680, total allocated 1250295808, free
 854682720, used 395613088, increment 146363328
 Ailanthus bio8_300m:
 Regularization values: linear/quadratic/product: 0.050, categorical: 0.250, threshold:
 1.000, hinge: 0.500
 136 samples

Density: max memory 3817799680, total allocated 1250295808, free 854682720, used 395613088, increment 0
 linearPredictor: max memory 3817799680, total allocated 1250295808, free 854682720, used 395613088, increment 0
 FeaturedSpace: max memory 3817799680, total allocated 1250295808, free 854682720, used 395613088, increment 0
 Sequential: max memory 3817799680, total allocated 1250295808, free 854682720, used 395613088, increment 0
 Initial loss: 9.2237500588861
 Time since start: 222.485
 300: time = 4.985000 loss = 8.718646
 Resulting gain: 0.5051037672986656
 Ailanthus response curves
 Response curve: only dem_300m_atshell
 Making features
 makeFeatures: max memory 3817799680, total allocated 1232142336, free 700039568, used 532102768, increment 136489680
 Ailanthus dem_300m_atshell:
 Regularization values: linear/quadratic/product: 0.050, categorical: 0.250, threshold: 1.000, hinge: 0.500
 136 samples
 Density: max memory 3817799680, total allocated 1232142336, free 700039568, used 532102768, increment 0
 linearPredictor: max memory 3817799680, total allocated 1232142336, free 700039568, used 532102768, increment 0
 FeaturedSpace: max memory 3817799680, total allocated 1232142336, free 700039568, used 532102768, increment 0
 Sequential: max memory 3817799680, total allocated 1232142336, free 700039568, used 532102768, increment 0
 Initial loss: 9.2237500588861
 Time since start: 225.391
 180: time = 2.609000 loss = 9.153857
 Resulting gain: 0.06989312910660495
 Ailanthus response curves
 Response curve: only nlcd_wet_300m_atshell
 Making features
 makeFeatures: max memory 3817799680, total allocated 1215299584, free 785845768, used 429453816, increment -102648952
 Ailanthus nlcd_wet_300m_atshell:
 Regularization values: linear/quadratic/product: 0.050, categorical: 0.250, threshold: 1.000, hinge: 0.500
 136 samples
 Density: max memory 3817799680, total allocated 1215299584, free 785845768, used 429453816, increment 0
 linearPredictor: max memory 3817799680, total allocated 1215299584, free 785845768, used 429453816, increment 0

FeaturedSpace: max memory 3817799680, total allocated 1215299584, free 785845768, used 429453816, increment 0
 Sequential: max memory 3817799680, total allocated 1215299584, free 785845768, used 429453816, increment 0
 Initial loss: 9.2237500588861
 60: time = 0.078000 loss = 9.192508
 Resulting gain: 0.031241803299451476
 Ailanthus response curves
 Response curve: only slope_800mn_300m_atshell
 Making features
 makeFeatures: max memory 3817799680, total allocated 1215299584, free 785486832, used 429812752, increment 358936
 Ailanthus slope_800mn_300m_atshell:
 Regularization values: linear/quadratic/product: 0.050, categorical: 0.250, threshold: 1.000, hinge: 0.500
 136 samples
 Density: max memory 3817799680, total allocated 1215299584, free 775481968, used 439817616, increment 10004864
 linearPredictor: max memory 3817799680, total allocated 1215299584, free 775481968, used 439817616, increment 0
 FeaturedSpace: max memory 3817799680, total allocated 1215299584, free 775481968, used 439817616, increment 0
 Sequential: max memory 3817799680, total allocated 1215299584, free 775481968, used 439817616, increment 0
 Initial loss: 9.2237500588861
 Time since start: 227.078
 80: time = 1.218000 loss = 9.214601
 Resulting gain: 0.009149497901907111
 Ailanthus response curves
 Response curve: only trasp_800mn_300m_atshell
 Making features
 makeFeatures: max memory 3817799680, total allocated 1215299584, free 595486632, used 619812952, increment 179995336
 Ailanthus trasp_800mn_300m_atshell:
 Regularization values: linear/quadratic/product: 0.050, categorical: 0.250, threshold: 1.000, hinge: 0.500
 136 samples
 Density: max memory 3817799680, total allocated 1215299584, free 595450784, used 619848800, increment 35848
 linearPredictor: max memory 3817799680, total allocated 1215299584, free 595450784, used 619848800, increment 0
 FeaturedSpace: max memory 3817799680, total allocated 1215299584, free 595450784, used 619848800, increment 0
 Sequential: max memory 3817799680, total allocated 1215299584, free 595450784, used 619848800, increment 0
 Initial loss: 9.2237500588861

Time since start: 229.328
 140: time = 2.000000 loss = 9.209168
 Resulting gain: 0.014582479705529394
 Ailanthus response curves
 Time since start: 231.422
 Jackknife: leave bio11_300m out
 Making features
 makeFeatures: max memory 3817799680, total allocated 1198456832, free 711244704, used 487212128, increment -132636672
 Ailanthus bio11_300m:
 Regularization values: linear/quadratic/product: 0.050, categorical: 0.250, threshold: 1.000, hinge: 0.500
 136 samples
 Density: max memory 3817799680, total allocated 1198456832, free 701567680, used 496889152, increment 9677024
 linearPredictor: max memory 3817799680, total allocated 1198456832, free 701567680, used 496889152, increment 0
 FeaturedSpace: max memory 3817799680, total allocated 1198456832, free 701567680, used 496889152, increment 0
 Sequential: max memory 3817799680, total allocated 1198456832, free 701567680, used 496889152, increment 0
 Initial loss: 9.2237500588861
 Time since start: 257.828
 440: time = 26.171000 loss = 8.280164
 Jackknife: leave bio19_300m out
 Making features
 makeFeatures: max memory 3817799680, total allocated 1333723136, free 892368696, used 441354440, increment -55534712
 Ailanthus bio19_300m:
 Regularization values: linear/quadratic/product: 0.050, categorical: 0.250, threshold: 1.000, hinge: 0.500
 136 samples
 Density: max memory 3817799680, total allocated 1333723136, free 892359272, used 441363864, increment 9424
 linearPredictor: max memory 3817799680, total allocated 1333723136, free 892359272, used 441363864, increment 0
 FeaturedSpace: max memory 3817799680, total allocated 1333723136, free 892359272, used 441363864, increment 0
 Sequential: max memory 3817799680, total allocated 1333723136, free 892359272, used 441363864, increment 0
 Initial loss: 9.2237500588861
 Time since start: 289.25
 500: time = 31.203000 loss = 8.272370
 Jackknife: leave bio4_300m out
 Making features

makeFeatures: max memory 3817799680, total allocated 1113718784, free 618224512, used 495494272, increment 54130408
 Ailanthus bio4_300m:
 Regularization values: linear/quadratic/product: 0.050, categorical: 0.250, threshold: 1.000, hinge: 0.500
 136 samples
 Density: max memory 3817799680, total allocated 1113718784, free 610168192, used 503550592, increment 8056320
 linearPredictor: max memory 3817799680, total allocated 1113718784, free 610168192, used 503550592, increment 0
 FeaturedSpace: max memory 3817799680, total allocated 1113718784, free 610168192, used 503550592, increment 0
 Sequential: max memory 3817799680, total allocated 1113718784, free 610168192, used 503550592, increment 0
 Initial loss: 9.2237500588861
 Time since start: 320.219
 500: time = 30.734000 loss = 8.377422
 Jackknife: leave bio8_300m out
 Making features
 makeFeatures: max memory 3817799680, total allocated 1174601728, free 876521960, used 298079768, increment -205470824
 Ailanthus bio8_300m:
 Regularization values: linear/quadratic/product: 0.050, categorical: 0.250, threshold: 1.000, hinge: 0.500
 136 samples
 Density: max memory 3817799680, total allocated 1174601728, free 867253160, used 307348568, increment 9268800
 linearPredictor: max memory 3817799680, total allocated 1174601728, free 867253160, used 307348568, increment 0
 FeaturedSpace: max memory 3817799680, total allocated 1174601728, free 867253160, used 307348568, increment 0
 Sequential: max memory 3817799680, total allocated 1174601728, free 867253160, used 307348568, increment 0
 Initial loss: 9.2237500588861
 Time since start: 351.11
 500: time = 30.641000 loss = 8.274713
 Jackknife: leave dem_300m_atshell out
 Making features
 makeFeatures: max memory 3817799680, total allocated 1224802304, free 833299992, used 391502312, increment 84153744
 Ailanthus dem_300m_atshell:
 Regularization values: linear/quadratic/product: 0.050, categorical: 0.250, threshold: 1.000, hinge: 0.500
 136 samples
 Density: max memory 3817799680, total allocated 1224802304, free 823104072, used 401698232, increment 10195920

linearPredictor: max memory 3817799680, total allocated 1224802304, free 823104072, used 401698232, increment 0
 FeaturedSpace: max memory 3817799680, total allocated 1224802304, free 823104072, used 401698232, increment 0
 Sequential: max memory 3817799680, total allocated 1224802304, free 823104072, used 401698232, increment 0
 Initial loss: 9.2237500588861
 Time since start: 376.61
 440: time = 25.266000 loss = 8.269345
 Jackknife: leave nlcd_wet_300m_atshell out
 Making features
 makeFeatures: max memory 3817799680, total allocated 1283981312, free 826047984, used 457933328, increment 56235096
 Ailanthus nlcd_wet_300m_atshell:
 Regularization values: linear/quadratic/product: 0.050, categorical: 0.250, threshold: 1.000, hinge: 0.500
 136 samples
 Density: max memory 3817799680, total allocated 1283981312, free 814671136, used 469310176, increment 11376848
 linearPredictor: max memory 3817799680, total allocated 1283981312, free 814671136, used 469310176, increment 0
 FeaturedSpace: max memory 3817799680, total allocated 1283981312, free 814671136, used 469310176, increment 0
 Sequential: max memory 3817799680, total allocated 1283981312, free 814671136, used 469310176, increment 0
 Initial loss: 9.2237500588861
 Time since start: 406.125
 500: time = 29.406000 loss = 8.263119
 Jackknife: leave slope_800mn_300m_atshell out
 Making features
 makeFeatures: max memory 3817799680, total allocated 1239547904, free 983075536, used 256472368, increment -212837808
 Ailanthus slope_800mn_300m_atshell:
 Regularization values: linear/quadratic/product: 0.050, categorical: 0.250, threshold: 1.000, hinge: 0.500
 136 samples
 Density: max memory 3817799680, total allocated 1239547904, free 983063448, used 256484456, increment 12088
 linearPredictor: max memory 3817799680, total allocated 1239547904, free 983063448, used 256484456, increment 0
 FeaturedSpace: max memory 3817799680, total allocated 1239547904, free 983063448, used 256484456, increment 0
 Sequential: max memory 3817799680, total allocated 1239547904, free 983063448, used 256484456, increment 0
 Initial loss: 9.2237500588861
 Time since start: 430.328

440: time = 24.015000 loss = 8.304572
 Jackknife: leave trasp_800mn_300m_atshell out
 Making features
 makeFeatures: max memory 3817799680, total allocated 1259667456, free 852771016, used 406896440, increment 150411984
 Ailanthus trasp_800mn_300m_atshell:
 Regularization values: linear/quadratic/product: 0.050, categorical: 0.250, threshold: 1.000, hinge: 0.500
 136 samples
 Density: max memory 3817799680, total allocated 1259667456, free 841842520, used 417824936, increment 10928496
 linearPredictor: max memory 3817799680, total allocated 1259667456, free 841842520, used 417824936, increment 0
 FeaturedSpace: max memory 3817799680, total allocated 1259667456, free 841842520, used 417824936, increment 0
 Sequential: max memory 3817799680, total allocated 1259667456, free 841842520, used 417824936, increment 0
 Initial loss: 9.2237500588861
 Time since start: 457.703
 440: time = 27.140000 loss = 8.287244
 Jackknife: only bio11_300m
 Making features
 makeFeatures: max memory 3817799680, total allocated 1072627712, free 711222176, used 361405536, increment -56419400
 Ailanthus bio11_300m:
 Regularization values: linear/quadratic/product: 0.050, categorical: 0.250, threshold: 1.000, hinge: 0.500
 136 samples
 Density: max memory 3817799680, total allocated 1072627712, free 711204328, used 361423384, increment 17848
 linearPredictor: max memory 3817799680, total allocated 1072627712, free 711204328, used 361423384, increment 0
 FeaturedSpace: max memory 3817799680, total allocated 1072627712, free 711204328, used 361423384, increment 0
 Sequential: max memory 3817799680, total allocated 1072627712, free 711204328, used 361423384, increment 0
 Initial loss: 9.2237500588861
 Time since start: 461.953
 320: time = 4.203000 loss = 8.741746
 Res.gain: 0.4820042634993431
 Jackknife: only bio19_300m
 Making features
 makeFeatures: max memory 3817799680, total allocated 1218183168, free 1010874008, used 207309160, increment -154114224
 Ailanthus bio19_300m:

Regularization values: linear/quadratic/product: 0.050, categorical: 0.250, threshold:
 1.000, hinge: 0.500
 136 samples
 Density: max memory 3817799680, total allocated 1218183168, free 1010851232,
 used 207331936, increment 22776
 linearPredictor: max memory 3817799680, total allocated 1218183168, free
 1010851232, used 207331936, increment 0
 FeaturedSpace: max memory 3817799680, total allocated 1218183168, free
 1010851232, used 207331936, increment 0
 Sequential: max memory 3817799680, total allocated 1218183168, free 1010851232,
 used 207331936, increment 0
 Initial loss: 9.2237500588861
 Time since start: 464.469
 180: time = 2.469000 loss = 9.117629
 Res.gain: 0.10612091163523019
 Jackknife: only bio4_300m
 Making features
 makeFeatures: max memory 3817799680, total allocated 1218183168, free
 627382888, used 590800280, increment 383468344
 Ailanthus bio4_300m:
 Regularization values: linear/quadratic/product: 0.050, categorical: 0.250, threshold:
 1.000, hinge: 0.500
 136 samples
 Density: max memory 3817799680, total allocated 1218183168, free 627360288, used
 590822880, increment 22600
 linearPredictor: max memory 3817799680, total allocated 1218183168, free
 627360288, used 590822880, increment 0
 FeaturedSpace: max memory 3817799680, total allocated 1218183168, free
 627360288, used 590822880, increment 0
 Sequential: max memory 3817799680, total allocated 1218183168, free 627360288,
 used 590822880, increment 0
 Initial loss: 9.2237500588861
 Time since start: 468.844
 320: time = 4.344000 loss = 8.407032
 Res.gain: 0.8167185188093811
 Jackknife: only bio8_300m
 Making features
 makeFeatures: max memory 3817799680, total allocated 1185742848, free
 874070200, used 311672648, increment -279150232
 Ailanthus bio8_300m:
 Regularization values: linear/quadratic/product: 0.050, categorical: 0.250, threshold:
 1.000, hinge: 0.500
 136 samples
 Density: max memory 3817799680, total allocated 1185742848, free 874038096, used
 311704752, increment 32104

linearPredictor: max memory 3817799680, total allocated 1185742848, free 874038096, used 311704752, increment 0
 FeaturedSpace: max memory 3817799680, total allocated 1185742848, free 874038096, used 311704752, increment 0
 Sequential: max memory 3817799680, total allocated 1185742848, free 874038096, used 311704752, increment 0
 Initial loss: 9.2237500588861
 Time since start: 473.735
 300: time = 4.844000 loss = 8.718646
 Res.gain: 0.5051037672986656
 Jackknife: only dem_300m_atshell
 Making features
 makeFeatures: max memory 3817799680, total allocated 1171914752, free 661495072, used 510419680, increment 198714928
 Ailanthus dem_300m_atshell:
 Regularization values: linear/quadratic/product: 0.050, categorical: 0.250, threshold: 1.000, hinge: 0.500
 136 samples
 Density: max memory 3817799680, total allocated 1171914752, free 661495072, used 510419680, increment 0
 linearPredictor: max memory 3817799680, total allocated 1171914752, free 661495072, used 510419680, increment 0
 FeaturedSpace: max memory 3817799680, total allocated 1171914752, free 661495072, used 510419680, increment 0
 Sequential: max memory 3817799680, total allocated 1171914752, free 661495072, used 510419680, increment 0
 Initial loss: 9.2237500588861
 Time since start: 476.25
 180: time = 2.484000 loss = 9.153857
 Res.gain: 0.06989312910660495
 Jackknife: only nlcd_wet_300m_atshell
 Making features
 makeFeatures: max memory 3817799680, total allocated 1157431296, free 709982096, used 447449200, increment -62970480
 Ailanthus nlcd_wet_300m_atshell:
 Regularization values: linear/quadratic/product: 0.050, categorical: 0.250, threshold: 1.000, hinge: 0.500
 136 samples
 Density: max memory 3817799680, total allocated 1157431296, free 709948360, used 447482936, increment 33736
 linearPredictor: max memory 3817799680, total allocated 1157431296, free 709948360, used 447482936, increment 0
 FeaturedSpace: max memory 3817799680, total allocated 1157431296, free 709948360, used 447482936, increment 0
 Sequential: max memory 3817799680, total allocated 1157431296, free 709948360, used 447482936, increment 0

Initial loss: 9.2237500588861
 60: time = 0.047000 loss = 9.192508
 Res.gain: 0.031241803299451476
 Jackknife: only slope_800mn_300m_atshell
 Making features
 makeFeatures: max memory 3817799680, total allocated 1157431296, free 709610408, used 447820888, increment 337952
 Ailanthus slope_800mn_300m_atshell:
 Regularization values: linear/quadratic/product: 0.050, categorical: 0.250, threshold: 1.000, hinge: 0.500
 136 samples
 Density: max memory 3817799680, total allocated 1157431296, free 700762872, used 456668424, increment 8847536
 linearPredictor: max memory 3817799680, total allocated 1157431296, free 700762872, used 456668424, increment 0
 FeaturedSpace: max memory 3817799680, total allocated 1157431296, free 700762872, used 456668424, increment 0
 Sequential: max memory 3817799680, total allocated 1157431296, free 700762872, used 456668424, increment 0
 Initial loss: 9.2237500588861
 80: time = 0.485000 loss = 9.214601
 Res.gain: 0.009149497901907111
 Jackknife: only trasp_800mn_300m_atshell
 Making features
 makeFeatures: max memory 3817799680, total allocated 1157431296, free 523979664, used 633451632, increment 176783208
 Ailanthus trasp_800mn_300m_atshell:
 Regularization values: linear/quadratic/product: 0.050, categorical: 0.250, threshold: 1.000, hinge: 0.500
 136 samples
 Density: max memory 3817799680, total allocated 1157431296, free 523945928, used 633485368, increment 33736
 linearPredictor: max memory 3817799680, total allocated 1157431296, free 523945928, used 633485368, increment 0
 FeaturedSpace: max memory 3817799680, total allocated 1157431296, free 523945928, used 633485368, increment 0
 Sequential: max memory 3817799680, total allocated 1157431296, free 523945928, used 633485368, increment 0
 Initial loss: 9.2237500588861
 Time since start: 478.125
 140: time = 1.328000 loss = 9.209168
 Res.gain: 0.014582479705529394
 Ending

BIBLIOGRAPHY

- Albright, T. P., H. Chen, L. Chen, and Q. Guo. 2009. The ecological niche and reciprocal prediction of the disjunct distribution of an invasive species: the example of *Ailanthus altissima*. *Biological Invasions* 12:2413–2427.
- Araújo, M. B., and M. Luoto. 2007. The importance of biotic interactions for modelling species distributions under climate change. *Global Ecology and Biogeography* 16:743–753.
- Araujo, M., and M. New. 2007. Ensemble forecasting of species distributions. *Trends in Ecology & Evolution* 22:42–47.
- Austin, M. 2007. Species distribution models and ecological theory: A critical assessment and some possible new approaches. *Ecological Modelling* 200:1–19.
- Austin, M. P. 2002. Spatial prediction of species distribution: an interface between ecological theory and statistical modelling. *Ecological modelling* 157:101–118.
- Austin, M. P., and K. P. Van Niel. 2011. Improving species distribution models for climate change studies: variable selection and scale. *Journal of Biogeography* 38:1–8.
- Bailey, R. G. 1998. *Ecoregions of North America*. U.S. Department of Agriculture, Forest Service.
- Benning, T. L., D. LaPointe, C. T. Atkinson, and P. M. Vitousek. 2002. Interactions of climate change with biological invasions and land use in the Hawaiian Islands: modeling the fate of endemic birds using a geographic information system.

Proceedings of the National Academy of Sciences of the United States of America 99:14246–14249.

Bradley, B. A., D. M. Blumenthal, D. S. Wilcove, and L. H. Ziska. 2010. Predicting plant invasions in an era of global change. *Trends in Ecology & Evolution* 25:310–318.

Bradley, B. A., D. S. Wilcove, and M. Oppenheimer. 2009. Climate change increases risk of plant invasion in the Eastern United States. *Biological Invasions* 12:1855–1872.

Brauman, K. A., G. C. Daily, T. K. Duarte, and H. A. Mooney. 2007. The nature and value of ecosystem services: an overview highlighting hydrologic services. *Annual Review of Environment and Resources* 32:67–98.

Brotons, L., W. Thuiller, M. B. Araújo, and A. H. Hirzel. 2004. Presence-absence versus presence-only modelling methods for predicting bird habitat suitability. *Ecography* 27:437–448.

Burch, P. L., and S. M. Zedaker. 2003. Removing the invasive tree *Ailanthus altissima* and restoring natural cover. *Journal of Arboriculture* 29:18–24.

Busby, J. 1991. BIOCLIM: a bioclimate analysis and prediction system. *Plant Protection Quarterly* 6:8.

Byers, J. E., S. Reichard, J. M. Randall, I. M. Parker, C. S. Smith, W. M. Lonsdale, I. A. E. Atkinson, T. R. Seastedt, M. Williamson, E. Chornesky, and D. Hayes. 2002. Directing research to reduce the impacts of nonindigenous species. *Conservation Biology* 16:630–640.

- Cardinale, B. J., J. E. Duffy, A. Gonzalez, D. U. Hooper, C. Perrings, P. Venail, A. Narwani, G. M. Mace, D. Tilman, D. A. Wardle, A. P. Kinzig, G. C. Daily, M. Loreau, J. B. Grace, A. Larigauderie, D. S. Srivastava, and S. Naeem. 2012. Biodiversity loss and its impact on humanity. *Nature* 486:59–67.
- Charles, H., and J. S. Dukes. 2007. Impacts of invasive species on ecosystem services. *Biological Invasions*:217–237.
- Connell, J. H., and R. O. Slatyer. 1977. Mechanisms of succession in natural communities and their role in community stability and organization. *American naturalist*:1119–1144.
- Coulston, J. W., G. A. Reams, R. E. McRoberts, and W. D. Smith. 2004. Practical considerations when using perturbed forest inventory plot locations to develop spatial models: a case study. *Proceedings of the Sixth Annual Forest Inventory and Analysis Symposium*.
- Coulston, J. W., K. H. Riitters, R. E. McRoberts, G. A. Reams, and W. D. Smith. 2006. True versus perturbed forest inventory plot locations for modeling: a simulation study. *Canadian Journal of Forest Research* 36:801–807.
- Dale, V. H. 1997. The relationship between land-use change and climate change. *Ecological applications* 7:753–769.
- Dale, V. H., L. A. Joyce, S. McNulty, R. P. Neilson, M. P. Ayres, M. D. Flannigan, P. J. Hanson, L. C. Irland, A. E. Lugo, C. J. Peterson, D. Simberloff, F. J. Swanson, B. J. Stocks, and B. M. Wotton. 2001. Climate change and forest disturbances. *BioScience* 51:723–734.

- Dieffenbach, F. 2003. APPA_HUC10_Shell Metadata. National Park Service
Northeast Temperate Network.
- Dufour, C., and E. Crisfield (Eds.). 2008. The Appalachian Trail MEGA-Transect.
Appalachian Trail Conservancy.
- Dukes, J. S. 2011. Responses of invasive species to a changing climate and
atmosphere. *Fifty Years of Invasion Ecology*:345–357.
- Dukes, J. S., and H. A. Mooney. 1999. Does global change increase the success of
biological invaders? *Trends in Ecology & Evolution* 14:135–139.
- Dullinger, S., I. Kleinbauer, J. Peterseil, M. Smolik, and F. Essl. 2009. Niche based
distribution modelling of an invasive alien plant: effects of population status,
propagule pressure and invasion history. *Biological Invasions* 11:2401–2414.
- EDDMaps. 2011. Early Detection & Distribution Mapping System. The University of
Georgia - Center for Invasive Species and Ecosystem Health.
- Elith, J., and C. H. Graham. 2009. Do they? How do they? WHY do they differ? On
finding reasons for differing performances of species distribution models.
Ecography 32:66–77.
- Elith, J., C. H. Graham, R. P. Anderson, M. Dudík, S. Ferrier, A. Guisan, R. J.
Hijmans, F. Huettmann, J. R. Leathwick, A. Lehmann, J. Li, L. G. Lohmann,
B. A. Loiselle, G. Manion, C. Moritz, M. Nakamura, Y. Nakazawa, J. McC. M.
Overton, A. Townsend Peterson, S. J. Phillips, K. Richardson, R. Scachetti-
Pereira, R. E. Schapire, J. Soberón, S. Williams, M. S. Wisz, and N. E.
Zimmermann. 2006. Novel methods improve prediction of species'
distributions from occurrence data. *Ecography* 29:129–151.

- Elith, J., M. Kearney, and S. Phillips. 2010. The art of modelling range-shifting species. *Methods in Ecology and Evolution* 1:330–342.
- Elith, J., and J. R. Leathwick. 2009. Species distribution models: ecological explanation and prediction across space and time. *Annual Review of Ecology, Evolution, and Systematics* 40:677–697.
- Elith, J., S. J. Phillips, T. Hastie, M. Dudík, Y. E. Chee, and C. J. Yates. 2011. A statistical explanation of MaxEnt for ecologists. *Diversity and Distributions* 17:43–57.
- Evans, J. S., and J. Oakleaf. 2011. Geomorphometry and Gradient Metrics Toolbox.
- Fry, J. A., G. Xian, S. Jin, J. A. Dewitz, C. G. Homer, L. Yang, C. A. Barnes, N. D. Herold, and J. D. Wickham. 2011. Completion of the 2006 National Land Cover Database for the Conterminous United States. *Photogrammetric Engineering and Remote Sensing* 77:858–864.
- Fryer, J. L. 2010. *Ailanthus altissima*. Fire Effects Information System. U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station, Fire Sciences Laboratory.
- Gesch, D., M. Oimoen, S. Greenlee, C. Nelson, M. Steuck, and D. Tyler. 2002. The National Elevation Dataset. *PE & RS- Photogrammetric Engineering & Remote Sensing* 68:5–11.
- Gómez-Aparicio, L., and C. D. Canham. 2008. Neighbourhood analyses of the allelopathic effects of the invasive tree *Ailanthus altissima* in temperate forests. *Journal of Ecology* 96:447–458.

- Guisan, A., C. H. Graham, J. Elith, F. Huettmann, and the NCEAS Species Distribution Modelling Group. 2007. Sensitivity of predictive species distribution models to change in grain size. *Diversity and Distributions* 13:332–340.
- Guisan, A., and W. Thuiller. 2005. Predicting species distribution: offering more than simple habitat models. *Ecology letters* 8:993–1009.
- Guisan, A., and N. E. Zimmermann. 2000. Predictive habitat distribution models in ecology. *Ecological Modelling* 135:147–186.
- Hashimoto, H., S. H. Hiatt, C. Milesi, F. S. Melton, A. R. Michaelis, P. Votava, W. Wang, and R. R. Nemani. 2011. Monitoring and forecasting ecosystem climate impacts on ecosystem dynamics using the Terrestrial Observation and Prediction System. *in* Y. Q. Wang, editor. *Remote Sensing of Protected Lands*. CRC Press.
- Hejda, M., P. Pyšek, and V. Jarošík. 2009. Impact of invasive plants on the species richness, diversity and composition of invaded communities. *Journal of Ecology* 97:393–403.
- Hellmann, J. J., J. E. Byers, B. G. Bierwagen, and J. S. Dukes. 2008. Five potential consequences of climate change for invasive species. *Conservation Biology* 22:534–543.
- Hernandez, P. A., C. H. Graham, L. L. Master, and D. L. Albert. 2006. The effect of sample size and species characteristics on performance of different species distribution modeling methods. *Ecography* 29:773–785.

- Hijmans, R. J., and C. H. Graham. 2006. The ability of climate envelope models to predict the effect of climate change on species distributions. *Global Change Biology* 12:2272–2281.
- Hijmans, R. J., S. Phillips, J. Leathwick, and J. Elith. 2012. Package “dismo.”
- Hirzel, A. H., and G. Le Lay. 2008. Habitat suitability modelling and niche theory. *Journal of Applied Ecology* 45:1372–1381.
- Hoshovsky, M. C. 1988. Element stewardship abstract for *Ailanthus altissima*. The Nature Conservancy, USA.
- Hu, S. 1979. *Ailanthus*. *Arnoldia* 39:29–50.
- Hutchinson, G. E. 1957. Concluding remarks. Pages 415–427 Cold Spring Harbor Symposia on Quantitative Biology.
- IPCC. 2007. Climate change 2007: the physical science basis: contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change. (S. Solomon, Ed.). Cambridge University Press, Cambridge; New York.
- Jarnevich, C. S., and T. J. Stohlgren. 2008. Near term climate projections for invasive species distributions. *Biological Invasions* 11:1373–1379.
- Jeschke, J. M., and D. L. Strayer. 2008. Usefulness of Bioclimatic Models for Studying Climate Change and Invasive Species. *Annals of the New York Academy of Sciences* 1134:1–24.
- Jiménez-Valverde, A., A. T. Peterson, J. Soberón, J. M. Overton, P. Aragón, and J. M. Lobo. 2011. Use of niche models in invasive species risk assessments. *Biological Invasions* 13:2785–2797.

- Jones, C. C., S. A. Acker, and C. B. Halpern. 2010. Combining local-and large-scale models to predict the distributions of invasive plant species. *Ecological Applications* 20:311–326.
- Kota, N. L., R. E. Landenberger, and J. B. McGraw. 2006. Germination and early growth of *Ailanthus* and tulip poplar in three levels of forest disturbance. *Biological Invasions* 9:197–211.
- Kowarik, I., and M. von der Lippe. 2011. Secondary wind dispersal enhances long-distance dispersal of an invasive species in urban road corridors. *NeoBiota* 9:49.
- Kowarik, I., and I. Saumel. 2007. Biological flora of Central Europe: *Ailanthus altissima* (Mill.) Swingle. *Perspectives in Plant Ecology, Evolution and Systematics* 8:207–237.
- Kowarik, I., and I. Säumel. 2008. Water dispersal as an additional pathway to invasions by the primarily wind-dispersed tree *Ailanthus altissima*. *Plant Ecology* 198:241–252.
- Kriticos, D., R. Sutherst, J. Brown, S. Adkins, and G. Maywald. 2003. Climate change and the potential distribution of an invasive alien plant: *Acacia nilotica* ssp. *indica* in Australia. *Journal of Applied Ecology* 40:111–124.
- Landenberger, R. E., N. L. Kota, and J. B. McGraw. 2006. Seed dispersal of the non-native invasive tree *Ailanthus altissima* into contrasting environments. *Plant Ecology* 192:55–70.

- Landenberger, R. E., T. A. Warner, and J. B. McGraw. 2009. Spatial patterns of female *Ailanthus altissima* across an urban-to-rural land use gradient. *Urban Ecosystems* 12:437–448.
- Lawrence, J. G., A. Colwell, and O. J. Sexton. 1991. The Ecological Impact of Allelopathy in *Ailanthus altissima* (Simaroubaceae). *American Journal of Botany* 78:pp. 948–958.
- Lemon, P. C. 1961. Forest Ecology of Ice Storms. *Bulletin of the Torrey Botanical Club* 88:pp. 21–29.
- Levin, S. A. 1992. The Problem of Pattern and Scale in Ecology: The Robert H. MacArthur Award Lecture. *Ecology* 73:pp. 1943–1967.
- Lister, A., C. Scott, S. King, M. Hoppus, B. Butler, and D. Griffith. 2005. Strategies for preserving owner privacy in the national information management system of the USDA Forest Service's Forest Inventory and Analysis unit. U.S. Department of Agriculture, Forest Service 352:163.
- Loarie, S. R., P. B. Duffy, H. Hamilton, G. P. Asner, C. B. Field, and D. D. Ackerly. 2009. The velocity of climate change. *Nature* 462:1052–1055.
- Lobo, J. M., A. Jiménez-Valverde, and J. Hortal. 2010. The uncertain nature of absences and their importance in species distribution modelling. *Ecography* 33:103–114.
- Lobo, J. M., A. Jiménez-Valverde, and R. Real. 2008. AUC: a misleading measure of the performance of predictive distribution models. *Global Ecology and Biogeography* 17:145–151.

- Mac Nally, R. 2000. Regression and model-building in conservation biology, biogeography and ecology: the distinction between—and reconciliation of—“predictive” and “explanatory” models. *Biodiversity & Conservation* 9:655–671.
- Mack, R. N., D. Simberloff, W. Mark Lonsdale, H. Evans, M. Clout, and F. A. Bazzaz. 2000. Biotic invasions: Causes, epidemiology, global consequences, and control. *Ecological Applications* 10:689–710.
- Miller, J. H. 1990. *Ailanthus altissima* (Mill.) Swingle *ailanthus*. Pages 101–104 in R. M. Burns and B. H. Honkala, editors. *Silvics of North America Vol. 2. Hardwoods*. USDA Forest Service, Washington, DC.
- Openshaw, S. 1983. *The modifiable areal unit problem*. Geo books Norwich.
- Pearson, R. G., and T. P. Dawson. 2003. Predicting the impacts of climate change on the distribution of species: are bioclimate envelope models useful? *Global Ecology and Biogeography* 12:361–371.
- Pearson, R. G., T. P. Dawson, and C. Liu. 2004. Modelling species distributions in Britain: a hierarchical integration of climate and land-cover data. *Ecography* 27:285–298.
- Peterson, A. T. 2003. Predicting the geography of species’ invasions via ecological niche modeling. *Quarterly Review of Biology* 78:419–433.
- Peterson, A. T., and D. A. Vieglais. 2001. Predicting species invasions using ecological niche modeling: new approaches from bioinformatics attack a pressing problem. *BioScience* 51:363–371.
- Phillips, S. 2005. A brief tutorial on Maxent. AT&T Research.

- Phillips, S. J., R. P. Anderson, and R. E. Schapire. 2006. Maximum entropy modeling of species geographic distributions. *Ecological Modelling* 190:231–259.
- Phillips, S. J., and M. Dudik. 2008. Modeling of species distributions with Maxent: new extensions and a comprehensive evaluation. *Ecography* 31:161–175.
- Pimentel, D., R. Zuniga, and D. Morrison. 2005. Update on the environmental and economic costs associated with alien-invasive species in the United States. *Ecological Economics* 52:273–288.
- Robinson, T. P., R. D. van Klinken, and G. Metternicht. 2010. Comparison of alternative strategies for invasive species distribution modeling. *Ecological Modelling* 221:2261–2269.
- Sakai, A. K., F. W. Allendorf, J. S. Holt, D. M. Lodge, J. Molofsky, K. A. With, S. Baughman, R. J. Cabin, J. E. Cohen, N. C. Ellstrand, D. E. McCauley, P. O’Neil, I. M. Parker, J. N. Thompson, and S. G. Weller. 2001. The population biology of invasive species. *Annual Review of Ecology and Systematics* 32:305–332.
- Schwarz, G. E., and R. B. Alexander. 1995. Soils data for the Conterminous United States Derived from the NRCS State Soil Geographic (STATSGO) Data Base. <http://water.usgs.gov/GIS/metadata/usgswrd/XML/ussoils.xml>.
- Shriver, G., T. Maniero, K. Schwarzkopf, D. Lambert, F. Dieffenback, D. Owen, Y. Q. Wang, J. Nugrand-Marzilli, G. Tierney, C. Reese, and T. Moore. 2005, November. Appalachian Trail Vital Signs. National Park Service.

- Sillero, N. 2011. What does ecological modelling model? A proposed classification of ecological niche models based on their underlying methods. *Ecological Modelling* 222:1343–1346.
- Smith, B. 2009. *A Tree Grows in Brooklyn*. HarperCollins.
- Stockwell, D. 1999. The GARP modelling system: problems and solutions to automated spatial prediction. *International Journal of Geographical Information Science* 13:143–158.
- Stohlgren, T. J., P. Ma, S. Kumar, M. Rocca, J. T. Morisette, C. S. Jarnevich, and N. Benson. 2010. Ensemble Habitat Mapping of Invasive Plant Species. *Risk Analysis* 30:224–235.
- Taylor, K. E., R. J. Stouffer, and G. A. Meehl. 2012. An Overview of CMIP5 and the Experiment Design. *Bulletin of the American Meteorological Society* 93:485–498.
- Thomas, C. D., A. Cameron, R. E. Green, M. Bakkenes, L. J. Beaumont, Y. C. Collingham, B. F. Erasmus, M. F. De Siqueira, A. Grainger, and L. Hannah. 2004. Extinction risk from climate change. *Nature* 427:145–148.
- Thuiller, W., D. M. Richardson, P. PYŠEK, G. F. Midgley, G. O. Hughes, and M. Rouget. 2005. Niche-based modelling as a tool for predicting the risk of alien plant invasions at a global scale. *Global Change Biology* 11:2234–2250.
- Trifilò, P., F. Raimondo, A. Nardini, M. L. Gullo, and S. Salleo. 2004. Drought resistance of *Ailanthus altissima*: root hydraulics and water relations. *Tree physiology* 24:107–114.

- Turner, M. G., R. V. O'Neill, R. H. Gardner, and B. T. Milne. 1989. Effects of changing spatial scale on the analysis of landscape pattern. *Landscape ecology* 3:153–162.
- USDA, N. 2012. The PLANTS Database. National Plant Data Team, Greensboro, NC 27401-4901 USA.
- USFS. 2005a. What is Forest Inventory and Analysis?
http://www.fia.fs.fed.us/library/fact-sheets/overview/What_is_FIA_FS.pdf.
- USFS. 2005b. FIA Data Collection. http://www.fia.fs.fed.us/library/fact-sheets/data-collections/FIA_Data_Collection.pdf.
- USFS. 2005c. Sampling and Plot Design.
- Vilà, M., J. L. Espinar, M. Hejda, P. E. Hulme, V. Jarošík, J. L. Maron, J. Pergl, U. Schaffner, Y. Sun, and P. Pyšek. 2011. Ecological impacts of invasive alien plants: a meta-analysis of their effects on species, communities and ecosystems. *Ecology Letters* 14:702–708.
- Vuuren, D. P., J. Edmonds, M. Kainuma, K. Riahi, A. Thomson, K. Hibbard, G. C. Hurtt, T. Kram, V. Krey, J.-F. Lamarque, T. Masui, M. Meinshausen, N. Nakicenovic, S. J. Smith, and S. K. Rose. 2011. The representative concentration pathways: an overview. *Climatic Change* 109:5–31.
- Walther, G.-R., A. Roques, P. E. Hulme, M. T. Sykes, P. Pyšek, I. Kühn, M. Zobel, S. Bacher, Z. Botta-Dukát, and H. Bugmann. 2009. Alien species in a warmer world: risks and opportunities. *Trends in Ecology & Evolution* 24:686–693.
- Warren, D. L., R. E. Glor, and M. Turelli. 2010. ENMTools: a toolbox for comparative studies of environmental niche models. *Ecography*.

- Warren, D. L., and S. N. Seifert. 2011. Ecological niche modeling in Maxent: the importance of model complexity and the performance of model selection criteria. *Ecological Applications* 21:335–342.
- Wiens, J. A. 1989. Spatial scaling in ecology. *Functional ecology* 3:385–397.
- Woudenberg, S. W., B. L. Conkling, B. M. O’Connell, E. B. LaPoint, J. A. Turner, and K. L. Waddell. 2010. The Forest Inventory and Analysis Database: Database description and users manual version 4.0 for Phase 2. US Department of Agriculture, Forest Service, Rocky Mountain Research Station.
- Wu, J. 2004. Effects of changing scale on landscape pattern analysis: scaling relations. *Landscape Ecology* 19:125–138.